View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by epublications@Marquette

Marquette University
e-Publications@Marquette

Biological Sciences Faculty Research and Publications

Biological Sciences, Department of

2-1-2015

Intrinsically Disordered C-Terminal Tails of *E. coli* Single-Stranded DNA Binding Protein Regulate Cooperative Binding to Single-Stranded DNA

Alexander G. Kozlov
Washington University School of Medicine in St. Louis

Elizabeth A. Weiland
Washington University School of Medicine in St. Louis

Anuradha Mittal
Washington University in St Louis

Vince Waldman
Washington University School of Medicine in St. Louis

Edwin Antony
Marquette University, edwin.antony@marquette.edu

See next page for additional authors

Accepted version. *Journal of Molecular Biology*, Vol. 427, No. 4 (February 2015): 763-774. DOI. © 2015 Elsevier. Used with permission.
Edwin Antony was affiliated with Utah State University at the time of publication.
NOTICE: this is the author's version of a work that was accepted for publication in *Journal of Molecular Biology.* Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in *Journal of Molecular Biology.* VOL 427, ISSUE 4, February 2015, DOI.

Authors
Alexander G. Kozlov, Elizabeth A. Weiland, Anuradha Mittal, Vince Waldman, Edwin Antony, Nicole Fazio, Rohit V. Pappu, and Timothy M. Lohman

This article is available at e-Publications@Marquette: https://epublications.marquette.edu/bio_fac/423

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

Intrinsically Disordered C-terminal Tails of *E. coli* Single-Stranded DNA Binding Protein Regulate Cooperative Binding to Single-stranded DNA

Alexander G. Kozlov
Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine
St. Louis, MO
Elizabeth Weiland
Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine
St. Louis, MO
Anuradha Mittal
2Department of Biomedical Engineering & Center for Biological Systems Engineering, Washington University in St. Louis
St. Louis, MO
Vince Waldman
Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine
St. Louis, MO

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

Edwin Antony
Department of Chemistry and Biochemistry, Utah State University,
Logan, UT
Nicole Fazio
Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine
St. Louis, MO
Rohit V. Pappu
Department of Biomedical Engineering & Center for Biological Systems Engineering, Washington University in St. Louis,
St. Louis, MO
Timothy M. Lohman
Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine
St. Louis, MO

Abstract: The homotetrameric *E. coli* single stranded DNA binding (SSB) protein plays a central role in DNA replication, repair and recombination. *E. coli* SSB can bind to long single-stranded (ss) DNA in multiple binding modes using all four subunits ($(SSB)_{65}$ mode) or only two subunits ($(SSB)_{35}$ binding mode), with the binding mode preference regulated by salt concentration and SSB binding density. These binding modes display very different ssDNA binding properties with the $(SSB)_{35}$ mode displaying highly cooperative binding to ssDNA. SSB tetramers also bind an array of partner proteins, recruiting them to their sites of action. This is achieved through interactions with the last 9 amino acids (acidic tip) of the intrinsically disordered linkers (IDLs) within the four C-terminal tails connected to the ssDNA binding domains. Here we show that the amino acid composition and length of the IDL affects the ssDNA binding mode preferences of SSB protein. Surprisingly the number of IDLs and the lengths of individual IDLs together with the acidic tip contribute to highly cooperative binding in the $(SSB)_{35}$ binding mode. Atomistic simulations suggest that cooperative binding correlates with preference of IDLs for globular conformations.

Introduction

Single stranded DNA binding proteins (SSBs) play essential roles in all aspects of DNA replication, recombination and repair. They bind selectively to single-stranded (ss)DNA with high affinity, protecting

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

ssDNA intermediates from degradation and inhibiting formation of DNA secondary structures {Chase, 1986; Meyer, 1990; Lohman and Ferrari, 1994; Wold, 1997}. *E. coli* SSB (EcSSB) protein also interacts with at least fourteen metabolic proteins {Shereda, 2008} (SSB interacting proteins (SIPs)) and serves as central hubs in regulating genome maintenance.

*Ec*SSB protein functions as a homotetramer, with each subunit (177 amino acids) comprised of two domains: an N-terminal DNA binding domain (112 amino acids), containing an oligonucleotide/oligosaccharide binding fold (OB fold), referred to here as the core, and a C-terminal domain composed of a flexible, intrinsically disordered (ID) linker (56 amino acids (Fig. 1A)) and a nine-residue linear motif (the "tip") that is the primary site of interaction of SSB with the SIPs {Shereda, 2008}.

Due to its four potential DNA binding sites, the tetrameric core of *Ec*SSB binds ssDNA in multiple modes that differ in the number of subunits used to contact the DNA. Two of the major modes are denoted as (SSB)35 and $(SSB)_{65}$, where the subscripts indicate the average number of ssDNA nucleotides occluded per bound SSB tetramer {Lohman, 1985; Bujalowski and Lohman, 1986}. The relative stabilities of these binding modes depend on salt concentration and type {Lohman, 1985; Bujalowski and Lohman, 1986} as well as protein to DNA ratio {Chrysogelos, 1982 ; Griffith, 1984; Bujalowski, 1988; Roy, 2007}. In the (SSB)65 mode, favored at [NaCl]>0.2M or $[Mg^{2+}]$>10mM, ~65 nucleotides of ssDNA wrap around all four subunits of the tetramer (see Fig. 1B), while displaying only "limited" cooperativity between adjacent tetramers {Lohman et al, 1986; Overman, Bujalowski, Lohman, 1988; Ferrari et al, 1994; Raghunathan, 2000}. In this binding mode SSB protein can diffuse along ssDNA {Roy et al, 2009; Zhou et al, 2011}, enabling transient destabilization of DNA hairpins and facilitating RecA filament formation {Roy et al, 2009}. These activities are functionally important for DNA repair and/or recombination.

In the $(SSB)_{35}$ mode, favored at [NaCl]<0.02M or $[Mg^{2+}]$<1mM and high SSB to DNA ratios, SSB binds ssDNA using an average of only two subunits, but with high cooperativity {Griffith, 1984 ; Lohman, 1985; Lohman et al, 1986; Lohman, 1990; Ferrari et al, 1994}, a property believed to be important during DNA replication

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

{Lohman et al TIBS, 1988}. In its $(SSB)_{35}$ mode, SSB can undergo direct transfer between separate DNA molecules {Kozlov and Lohman, 2002 b} or intersegment transfer between distant sites within the same DNA molecule {Lee and Ha, 2014}, an activity that may be important for SSB recycling during replication {Kozlov and Lohman, 2002}.

The four unstructured C-termini of EcSSB can also influence the relative stability of the different SSB-ssDNA binding modes. SSB variants with only one or two C-terminal tails shift the binding mode preference toward the $(SSB)_{35}$ mode {Antony et al, 2013}. The highly conserved, negatively charged "tip" of the SSB C-terminus can interact with the DNA binding site within the SSB core at moderate salt concentrations {Kozlov and Lohman, 2010; Su and Mason, 2014} and thus reduce ssDNA binding affinity. Removal of the C-terminal tip also favors the $(SSB)_{35}$ binding mode {Kozlov and Lohman, 2010; Roy, 2007}. The $(SSB)_{35}$ mode seems also to be favored when the C-terminal tip interacts with some SIPs (e.g. PriC {Wessel, 2013} and PriA {Bhattacharyya, 2014}). Interestingly, the *Plasmodium falciparum (Pf)* SSB tetramer, which has a similar organization in terms of sequence and architecture, but possesses a C-terminal tail with a higher fraction of charged residues {Antony et al, 2012 a}, does not form a stable $(SSB)_{35}$ binding mode {Antony et al, 2012 b}.

The regions responsible for highly cooperative binding of SSB tetramers to ssDNA have not yet been identified. Here, we report the surprising result that highly cooperative binding of *Ec*SSB to ssDNA in its $(SSB)_{35}$ mode involves the C-terminal tails and is affected by the length, composition and number of IDLs and by the acidic tip. Complete removal of the linker or dramatic changes to its amino acid composition eliminates highly cooperative binding to ssDNA.

Results

Design of SSB tail variants. Biochemical studies suggest that the C terminal tails (residues 113-177, see Fig. 1A) of *E. coli* SSB lack any substantial structure in its apo form and when it is bound to ssDNA based on {Williams, 1983}. The absence of electron density for residues beyond 112 in any of the x-ray crystal structures {Raghunathan, 1997; Raghunathan, 2000; Savvides, 2004} further confirms the designation of intrinsic disorder for the C-terminal tails.

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

In fact, the C-terminal tails are disordered in all crystal structures of SSB analyzed so far {Arif and Vijayan, 2012; Antony et al, 2012}. While the 9 amino acid tip of the EcSSB C-terminus is conserved among many bacterial SSBs {Shereda, 2008}, the IDL (56 amino acids for *E. coli* SSB (see Fig 1A)) varies considerably in length from 25 to 135 aa (Fig. S5). ? In order to examine the effect of linker length and composition on the DNA binding properties of SSB we expressed and purified the SSB linker variants summarized in Fig. 1C to examine.

We examined SSB variants with 16, 37, 47 and 54 amino acids deleted from the linker region (SSBΔ151-166, SSBΔ130-166, SSBΔ120-168 and SSBΔ115-168, respectively, see Fig. 1C), while retaining the last 9 amino acids of the tip. The maximum deletion construct (lacking 54 amino acids) contains only two glycine residues between the DNA binding core and the tip (see Fig. 1C-ii), and we refer to this as SSB-GG. We also examined SSBΔC8 (see Fig. 1C-iii), which has a full-length linker but lacks the last 8 amino acids of the tip {Kozlov et al, 2010}. An additional variant, SSB-EPE was made (Fig. 1C-iv), that has the *Ec*-core (1-112 a.a.) and the *Ec*-tip, connected by the linker from *Plasmodium falciparum (Pf)* SSB that is longer (80 amino acids) and has a higher fraction of charged residues than the *E. coli* linker {Antony et al, 2012 a}.

Hydrodynamic properties of the SSB linker variants.

We examined the assembly state of all SSB variants using sedimentation velocity at low and moderate salt concentrations (10 mM and 0.30 M NaCl)), conditions that favor the (SSB)35 or (SSB)65 binding modes, respectively, for wt *E. coli* SSB {Lohman and Overman, 1985; Bujalowski and Lohman, 1986}. All variants displayed a single, symmetric peak in a continuous sedimentation (c(s)) analysis consistent with a single species. In such cases accurate estimates of molecular mass and frictional coefficient ratios can be obtained {Schuck, 1998; Dam, Schuck, 2004} (see Table 1). The estimated molecular masses agree well with those expected for tetramers calculated from the amino acid composition (±2.5kDa). At 0.30 M NaCl the estimated frictional coefficient ratios (f/f_o) increase with increasing linker length suggesting that the hydrodynamic radii increase with increasing linker length.

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

At lower salt concentrations (10 mM NaCl) $s_{20,w}$ for wt SSB (4.56S), shows a ~10% increase over its value (4.16 S) in 0.30 M NaCl. This corresponds to a decrease in f/f_o from 1.60 in 0.30 M NaCl to 1.46 in 10 mM NaCl, presumably due to a greater compaction of the tails resulting from the interaction of the tips with the DNA binding sites within the tetrameric core, an interaction that is favored at lower salt concentrations {Kozlov and Lohman, 2010; Su, Mason, 2014}. The fact that the SSB-GG variant that is completely missing the linker has nearly the same f/f_o (1.41) as does wtSSB in 10 mM NaCl also suggests that the linkers in the wtSSB are not as extended in 10 mM NaCl as they are in 0.30 M NaCl. However, the opposite effect of [NaCl] is observed for the SSB-EPE construct, which shows an increase in f/f_o from 1.76 at 0.30 M to 1.89 at 10 mM NaCl, suggesting that the *Pf*-linker is more extended than the *Ec* linker at both [NaCl], but much more so at 10 mM NaCl.

Atomistic simulations of the Intrinsically Disordered Linkers.

We used atomistic simulations based on the ABSINTH implicit solvation model {Vitalis, Pappu, 2009} to assess the relationships between linker length and sequence and conformational properties. This approach has yielded accurate assessments of sequence-ensemble relationships for a range of intrinsically disordered proteins{Mao, Pappu, 2010; Das, Pappu, 2012; Mao, Pappu, 2013; Das, Pappu, 2013; Pappu, Crick, 2008}. Internal scaling profiles provide an assessment of the most likely value for the mean spatial separation (R_{ij}) for pairs of residues i and j that are |j–i| apart in the linear sequence. If the sequence predominantly adopts globules, then the profile of (R_{ij}) will show plateauing behavior, and the plateau value provides an assessment of the density of the globule. Sequences whose conformational distributions are congruent with those of statistical random coils will exhibit power-law behavior. Specifically, $(R_{ij}) \sim |j-i|^{v}$, where the value of v will approach either 0.5, if chain-solvent and intra-chain interactions are perfectly counterbalanced as in a Flory random coil, or 0.59 if the chain statistics are congruent with those of a self-avoiding random walk. In Fig. 2, EcEc and PfPf refer to sequences that include the C-terminal linker and tip from *Ec*SSB and *Pf*SSB, respectively. The SSB core was not included in the simulations.

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

The conformational properties of EcEc are clearly distinct from those of PfPf (Fig. 2A). The EcEc sequence adopts compact globular conformations as evidenced by the plateauing behavior of (R_{ij}) and the similarity of this profile to that of a maximally compact reference globule. The internal scaling profiles for the internal deletion variants, SSBΔ151-166, Δ130-166, and Δ120-166, are congruent with the profile for the wild type EcEc construct. The implication is that internal deletions within EcEc do not compromise the intrinsic preferences for compact globular conformations. In contrast, the internal scaling profile for PfPf conforms to that of a Flory random coil implying a preference for expanded and aspherical conformations.

Fig. 2B shows a set of six cumulative distribution functions for the end-to-end distance, R_{ee}. These functions quantify the probabilities associated with realizing a Ree value that is less than or equal to some threshold. The probability of realizing a specific value for R_{ee} shows minimal variation between EcEc and its internal deletion constructs. In contrast, the simulations predict that (R_{ee}) will be approximately four times larger for PfPf than for EcEc. These simulation results clearly suggest that the *E. coli* linkers yield compact, dense globules whereas the linker from *P. falciparum* forms expanded conformations whose properties are congruent with canonical Flory random coils. These results are compatible with the inferences drawn from the sedimentation velocity results presented above.

DNA binding properties of the SSB linker variants.

What do the differences in conformational properties between *E. coli* and *P. falciparum* imply for ssDNA binding? Similarly, what do the similarities between the conformational properties of the wild type *E. coli* linker and those of internal deletions imply for ssDNA binding? We examined the ssDNA binding mode preferences of the SSB linker variants by determining their occluded site sizes at low and moderate salt concentrations by monitoring the quenching of the SSB Trp fluorescence upon binding poly(dT) {Lohman and Overman, 1985; Bujalowski and Lohman, 1986}. At 0.30 M NaCl (Fig. 3A) the extents of Trp quenching (Q_{max}~90-95%) and occluded site sizes (n_{GG}=60±2 nt and n_{EPE}=69±2 nt) suggest that the SSB-GG and SSB-EPE variants form fully wrapped (SSB)65 complexes similar to wtSSB (n_{wt}=65±2 nt). Experiments performed with the linker deletion variants and SSBΔC8 show the same behavior (see Supplementary Fig. S1-A).

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

At 10 mM NaCl wtSSB binds in its $(SSB)_{35}$ binding mode with a 35 nt site size {Lohman and Overman, 1985}. The SSB-GG variant with no linker also shows a 35 nt site size (nGG=35±2 nt; Qmax=0.59±0.01) (Fig. 3B). Interestingly, the SSB-GG variant displays a very stable $(SSB)_{35}$ binding mode at 10 mM NaCl, in contrast to wtSSB that shows a slow re-equilibration to a higher site size mode upon further addition of poly(dT) due to the metastability of its $(SSB)_{35}$ mode at low SSB to DNA binding densities {Lohman and Overman, 1985; Lohman, Overman and Datta, 1986; Lohman and Ferrari, 1994}. The variants lacking only part of the linker (Δ120-166, Δ130-166 and Δ151-166,) demonstrate binding behavior that are similar to wtSSB (see supplementary Fig. S1-B). At 10 mM NaCl the SSB-EPE variant shows a larger occluded site size (n_{EPE}=55±2 nt; Fig. 3B). This is the same site size determined for the wt *Pf*SSB under similar conditions {Antony, 2012 b}. This unexpected result suggests that the *Pf* linker, and not the *Pf* DNA binding core, prevents formation of the $(SSB)_{35}$ binding mode. Therefore, total removal of the *Ec* linker favors the $(SSB)_{35}$ mode, whereas replacement of the *Ec* linker with the *Pf*-linker favors formation of the higher site size complexes.

wtSSB can form a high affinity (stoichiometric) fully wrapped 1:1 complex with $(dT)_{70}$ at both low and high salt concentrations {Bujalowski, Lohman, 1989 b; Ferrari, Bujalowski and Lohman, 1994; Roy et al, 2007}. In this complex the $(dT)_{70}$ interacts with all four subunits, quenching ~90% of the SSB Trp fluorescence. The wtSSB tetramer can also bind two molecules of $(dT)_{35}$, but with a negative cooperativity that increases as the salt concentration decreases and this is partly responsible for formation of the $(SSB)_{35}$ binding mode on poly(dT) at low [NaCl] {Lohman, Bujalowski, 1988; Bujalowski, Lohman, 1989 b; Lohman, Bujalowski, 1994; Lohman, Ferrari, 1994}. At 0.30 M NaCl (supplementary Fig. S2-A) both SSB-GG and SSB-EPE form fully wrapped stoichiometric complexes with $(dT)_{70}$ as observed for wtSSB. All variants show high affinity binding to $(dT)_{70}$ under conditions that lower the affinity such that an equilibrium constant can be measured (2 M NaBr) (Fig. S2-C). However, at 10mM NaCl, SSB-GG does not form a 1:1 complex with $(dT)_{70}$, but rather 2 SSB-GG tetramers bind to $(dT)_{70}$, even at high DNA to protein ratios (Fig. S2-B), consistent with a very stable $(SSB)_{35}$ binding mode as seen for binding to poly(dT) (Fig. 3B). Consistent with this observation, SSB-GG shows a much larger negative cooperativity for binding a second

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

molecule of $(dT)_{35}$ at both 0.3 M and 10 mM NaCl than does wtSSB (see supplementary Fig. S2-D and E). The SSB-EPE variant shows moderate negative cooperativity for binding of a second molecule of $(dT)_{35}$ at 0.30 M NaCl (Fig.S2-D), although its inability to bind second molecule of dT_{35} at low salt (Fig. S2-E) is due to the fact that this variant is unable to form low site size mode in these conditions but forms 55 mode instead (Fig. 3B).

C-terminal linkers affect cooperative binding of SSB to ssDNA.

E. coli SSB is able to bind with high cooperativity to ssDNA, resulting in formation of protein clusters on long ssDNA {Sigal, 1972; Ruyechan, Wetmur, 1975; Schneider, Wetmur, 1982; Lohman et al, 1986}, but only at low salt concentrations {Lohman et al, 1986}. Highly cooperative binding to ssDNA requires SSB binding in its $(SSB)_{35}$ binding mode; the $(SSB)_{65}$ binding mode displays only a limited cooperativity {Lohman, Overman and Datta, 1986; Overman, Bujalowski, Lohman, 1988; Ferrari et al, 1994}. However, the regions of SSB protein responsible for highly cooperative binding have not previously been identified.

We used two qualitative methods to examine whether the SSB linker modifications affect the ability of SSB to bind with high cooperativity to M13 mp18 phage ssDNA (~7.25 knt). The first is an agarose gel electrophoretic mobility shift assay (EMSA) {Lohman, Overman and Datta, 1986} and the second is a sedimentation velocity assay. For a non-cooperative or low cooperative binding protein, all of the DNA in the population will have nearly the same amount of protein bound and thus all of the DNA migrates at the same rate in the gel and only one DNA band is observed for a given protein to DNA ratio. However, highly cooperative binding of wtSSB to M13 phage ssDNA is evident as a bimodal distribution of ssDNA, resulting from some ssDNA molecules having near saturating amounts of SSB bound while other ssDNA has little bound protein {Lohman, Overman and Datta, 1986}.

The behaviors of the SSB variants lacking the linker (SSB-GG) or with the *E. coli* linker replaced by the *Pf*-linker (SS-EPE) are compared with wtSSB at low and moderate salt (10mM NaCl and 0.30 M NaCl). SSB-DNA complexes were formed at varying SSB to DNA ratios at the indicated NaCl concentration, equilibrated for 1 hour, and

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

loaded onto the gel, followed by electrophoresis (see Materials and Methods for details). The hallmark of highly cooperative binding of wtSSB to ssM13 DNA at low salt (10 mM NaCl) is a bi-modal distribution of ssDNA as shown in Fig. 4A-i, where at intermediate SSB to DNA ratios (e.g., R35 =0.3), free ssDNA exists in the same population as ssDNA that is nearly saturated with SSB {Lohman, Overman and Datta, 1986}. As observed previously {Lohman, Overman and Datta, 1986}, this highly cooperative binding is eliminated at higher [NaCl] (0.30 M), where the ssDNA migrates as a single diffuse band at all SSB to DNA ratios (Figure S3A-i). The same experiments performed at 10 mM NaCl with the SSB variant that is either missing the linker (SSB-GG) (Fig. 4A-ii) or has the linker replaced with the *Pf*-linker (SSB-EPE) (Fig. 4A-iii) do not show bi-modal distributions indicating the absence of highly cooperative binding. At 0.30 M NaCl the binding patterns are the same for all of these variants since highly cooperative binding of wtSSB is eliminated (Fig. S3-A). We also examined the cooperative binding behavior of three SSB variants, SSBΔ151-166, SSBΔ130-166 and SSBΔ120-166, in which parts of the SSB linker have been deleted. Each of these SSB variants still show bimodal DNA distributions at 10 mM NaCl indicative of highly cooperative behavior (Fig. S3-B).

In the EMSA experiments, the SSB-ssDNA complexes are allowed to equilibrate for 1 hr before loading onto the agarose gel for electrophoresis. However, since the electrophoresis running buffer (20 mM-Tris-Acetate, pH 8.3, 0.5mM EDTA) differs from the buffer in which the SSB-M13ssDNA complexes were formed it is possible that the change in salt conditions can affect distribution of the complexes during electrophoresis. For this reason we also examined cooperative binding using sedimentation velocity analytical ultracentrifugation, since those experiments can be performed in the identical buffer conditions used to form the SSB-ssDNA complexes. These sedimentation velocity experiments were analyzed using a c(s) distribution analysis {Schuck, 1998; Dam, Schuck, 2004}.

At low salt concentration (10 mM NaCl) M13 ssDNA sediments as a slow moving (14.0 S), but sharp symmetrical peak in a c(s) distribution analysis (Fig. 4B, green). At 0.30 M NaCl, M13 ssDNA sediments as a single peak, but with a much larger sedimentation coefficient (~31 S) (data not shown) indicative of a lower frictional

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

coefficient that results from compaction of the ssDNA that is mediated by the increased electrostatic screening.

As a control, we first examined M13ssDNA complexes with wtSSB formed in 10 mM NaCl at less than saturating protein to DNA ratios (Fig. 4B-i). At a low protein to DNA ratio (R_{35}=0.3), where R_{35} is the fraction of ssDNA saturated by SSB if it were all to bind in its $(SSB)_{35}$ binding mode, the c(s) profile shows a bimodal distribution of ssDNA, with ~16% of the ssDNA having a sedimentation coefficient ~54 S, indicating a high SSB binding density, whereas ~53% of the DNA sediments with s~20-22 S indicating little bound SSB (see Supplementary Table S1 ?). The remainder of the ssDNA sediments between these two peaks reflecting a wide variation in the amount of SSB bound. At saturating concentrations of wtSSB, the ssDNA moves with a sedimentation coefficient of the high molecular weight complex, ~55 S (90% at R_{35}=0.8). These results are qualitatively consistent with the EMSA results (Fig 4A).

We next examined the SSB linker variants at a protein to ssDNA ratio of R_{35}=0.3 at 10 mM NaCl. Fig. 4B-ii shows that the ssDNA bound with either SSB-GG (no linker) or SSB-EPE (*Pf*-linker), sediments as single peaks with s ~24 S and ~30 S, respectively indicating the absence of high cooperativity. The SSBΔ130-166 and SSBΔ120-166 variants missing 37 and 47 amino acids from the linker, respectively (Fig. 4B-iii) display a bi-modal c(s) profiles similar to wtSSB indicating high cooperativity. The shift in peak position for the highly cooperative complexes of SSBΔ130-166 (44 S) and SSBΔ120-166 (40 S) relative to wtSSB (~54 S) reflects the lower molecular mass of the SSB due to the partial deletion of the linkers (Table 1). We also examined the SSBΔC8 variant that is missing the last 8 amino acids of the C-terminal tip. The c(s) profile indicates that binding is cooperative at 10 mM NaCl (Fig. 4B-iv), however, the c(s) distribution is clearly less cooperative than wtSSB (the fraction of the highly cooperative species decreases from 16% to ~7%, see Table S1). Hence, the acidic tip also appears to play some role in cooperative binding.

We have previously described SSB variants in which either two subunits (SSB linked dimer-Drl) or all four subunits (SSB linked tetramer-Drl) are covalently linked {Antony et al, 2013}. The amino acid linkers connecting the OB-folds in these constructs is from the *D. radiodurans* SSB, hence these variants are designated SSB-LD-Drl and

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

SSB-LT-Drl, respectively {Antony et al, 2013}. These variants possess either only two (SSB-LD-Drl) or one (SSB-LT-Drl) C-terminal tails, respectively {Antony et al, 2013} and we examined their cooperative binding by sedimentation velocity at 10 mM NaCl (R35 = 0.3). Figure 4B-iii shows that the LD-Drl variant with two tails displays the bi-modal ssDNA profile indicative of high cooperativity similar to wtSSB (16% highly cooperative species, Supplementary Table S1). However, the LT-Drl variant with only one C-terminal tail does not display the bi-modal behavior, although it still binds with more cooperativity than SSB-GG (Fig. 4B-iv, Table S1)). Hence, reducing the number of C-terminal tails to one significantly reduces its ability to bind ssDNA with high cooperativity.

In vivo Complementation and DNA Damage Sensitivity of SSB Linker Variants.

We examined the ability of SSB variants to function in *E. coli* by testing their ability to complement the loss of wtSSB *in vivo* by using a complementation assay developed by Porter {Porter, Black, 1991; Antony et al, 2013} (see Materials and Methods for details). Our results indicate that the genes expressing all of the SSB variants functionally complement the wtSSB gene *in vivo*, with the exception of SSBΔC8 as shown previously {Curth, Greipel, 1996}. Therefore, it appears that the presence of a long C-terminal linker length is not essential for *E. coli* growth. These results also indicate that highly cooperative binding of SSB to ssDNA is not essential for *E. coli* survival.

We further tested *E. coli* expressing the SSB variants for their ability to recover from DNA damage. *E. coli* cells expressing SSB linker variants were exposed to varying degrees of UV irradiation (see Materials and Methods). Exposure to UV irradiation leads to formation of DNA breaks and base damage including crosslinks {Bonura, Smith, 1975}. Interestingly, we found that the ability of the cells to recover after UV-induced damage depends on the length of the linker (Figure S4). While the cells expressing SSB variants with partial linker deletions (SSBΔ130-166, SSBΔ151-166) or have Plasmodium linker instead (SSB-EPE), show a similar ability to recover as wtSSB, cells expressing SSB variants that are completely missing the linker (SSB-GG) or most of it (SSBΔ120-166) are much more sensitive to UV

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

irradiation. This indicates that a linker of sufficient length is important for some DNA repair pathways. Whether this is due to a need for cooperative binding or if the phenotype reflects a need for a minimum length tether remains to be determined.

Discussion

Cooperative binding to ssDNA of SSB proteins involved in DNA replication was first demonstrated for the phage T4 gene 32 protein {Alberts, Frey, 1972; Kowalczykowski, von Hippel, 1981; Lohman, 1984; Giedroc, 1986} and subsequently for the *E. coli* SSB protein {Sigal, 1972; Ruyechan, Wetmur, 1975; Schneider, Wetmur, 1982; Lohman, Datta, 1986}. However, despite years of study, the regions of SSB protein responsible for highly cooperative binding had not previously been identified, although it was hypothesized based on crystal structures that interactions between the L_{45} loops might be involved {Raghunathan et al, 2000}. It also was not known whether highly cooperative binding is essential for *E. coli* survival. In fact, the eukaryotic SSB analogue, Replication Protein A (RPA) does not display cooperative binding to ssDNA {Kim and Wold, 1995; Kumaran et al, 2006}.

Although the DNA binding domain (OB-fold) and the 9 amino acid tip are essential for *E. coli* survival there was little known about the function or properties of the 56 amino acid linker that connects the OB-fold and the tip. This linker lacks structure, as indicated by the absence of electron density in crystal structures of SSB, even when bound to ssDNA {Savvides, 2004}. Our studies reveal several surprising results. First, complete elimination of the linker (SSB-GG) eliminates highly cooperative binding of SSB to ssDNA *in vitro*. Second, removal of the acidic C-terminal tip (last 9 amino acids) reduces, but does not eliminate cooperative binding. Third, an SSB variant with two C-terminal tails retains cooperative binding, whereas a one tail variant shows a reduction in cooperative binding. Fourth, replacement of the 56 amino acid *Ec*SSB linker with the 80 amino acid *Pf*-linker (SSB-EPE) eliminates the $(SSB)_{35}$ binding mode and cooperativity. Fifth, the SSB linker and high cooperativity are not essential for *E. coli* survival, although cooperative binding may function in other roles not central to survival.

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

Three components seem to influence highly cooperative binding of SSB to ssDNA. 1. SSB tetramers need to be bound in the $(SSB)_{35}$ mode, although this is not sufficient; 2. IDLs of appropriate length, number and amino acid composition are needed; 3. the acidic tips of SSB tetramers also facilitate cooperative binding. The length of the linker region seems less important since linkers with lengths of 9, 19 or 40 amino acids still display highly cooperative binding. In this regard, a two-linker variant, SSB linked dimer, retains high cooperativity. Although the one tail variant loses its ability to bind with high cooperativity, it still shows some residual cooperativity. Interestingly, unlike the two-tailed variant, expression of the single-tailed SSB variant shows a dominant lethal phenotype. Single tailed SSB also shows defects in coupled leading and lagging strand replication and does not support replication restart *in vivo* {Antony et al, 2013}.

Previous studies have shown that at low to moderate salt concentrations, the 9 amino acid tips of the C-termini of SSB can interact with the ssDNA binding sites within the DNA binding core (OB-fold) and compete weakly for ssDNA binding {Kozlov et al, 2010; Shishmarev, Dixon, 2014}. In the $(SSB)_{35}$ binding mode an average of two ssDNA binding sites are unoccupied, hence, one possible mechanism for cooperative binding, depicted in Fig. 5, is that the 9 amino acid tip(s) from one $(SSB)_{35}$-bound tetramer, interacts with a ssDNA binding site in an adjacent $(SSB)_{35}$-bound tetramer, thus promoting clustering of SSB on the ssDNA. However, we also show that an SSB in which only the last 8 amino acids of the tip are deleted (SSBΔC8) still binds cooperatively, although it is reduced. Thus, both the acidic tip as well as an ID linker of appropriate length and amino acid composition contribute to highly cooperative ssDNA binding. The fact that an SSB with no linker, but retaining the acidic tip (SSB-GG) does not show cooperative binding may indicate the need for a minimal length IDL to allow the tip to reach an unoccupied ssDNA site on an adjacent SSB tetramer on the ssDNA. The fact that a one-tailed variant loses high cooperativity supports this proposal. Although *E. coli* can survive with an SSB lacking the linker regions, as long as the 9 amino acid tip is retained, our studies indicate that sensitivity of *E. coli* to UV irradiation and the DNA damaging agent MMS increases as the SSB linker is shortened. Hence, the length of the linkers are important for some DNA repair processes.

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

Many proteins involved in signaling and regulation use intrinsically disordered regions (IDR) as interaction sites{Dunker, 2008}. However, our observations that the SSB C-terminal tails play roles in cooperative binding of SSB to ssDNA represents a novel function of an IDR. Recent computational studies have classified IDPs based on their amino acid compositions{Das, Pappu, 2013}. Our computational studies suggest that the sequence of the disordered linker plus tip from *E. coli* should form heterogeneous distributions of globules and this is supported by hydrodynamic measurements. For strong polyampholytes such as the C-terminal linker plus tip from *P. falciparum*, the patterning of oppositely charged residues gives rise to conformations whose statistical properties are congruent with those of Flory random coils.

Our survey of 134 bacterial SSB proteins indicates that the lengths of C-terminal linkers plus tips vary from 25 to 135 residues, with the large majority being 55-65 residues long. Based on their amino acid compositions, a majority of these sequences are classified as globule formers according to the diagram-of-states (Fig. S5). We therefore expect that theses variations in linker lengths will not lead to significant changes in the overall dimensions of globule forming sequences. Further, since globule formation is a proxy for strong self-interactions {Pappu, Crick, 2008}, it is reasonable to propose that globule, as opposed to coil-forming linkers promote intra- and inter-tetramer associations that in turn promote positive cooperativity in ssDNA binding. In essence, we propose a model where strongly self-associating linkers with acidic tips can mediate two types of interactions between adjacent SSBs. These include interactions between globular linkers, which we propose is encouraged by the high local concentration of SSBs, and previously documented interactions between acidic tips and OB cores. Upon ssDNA binding, the linker and tip-mediated interactions between SSBs might serve as "links in a chain" that facilitate the formation of high molecular weight clusters. The links, we propose, are strengthened by increasing the numbers of tails and this is consistent with data we report for the wtSSB versus the two- and one-tailed variants.

Materials and Methods.

Reagents and Buffers. All buffers were prepared with reagent grade chemicals and distilled water that was subsequently treated with

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

a Milli Q (Millipore, Bedford, MA) water purification system. Buffer T is 10 mM Tris , pH 8.1 (25□C), 0.1 mM Na_3EDTA. The gel electrophoresis buffer is 20 mM Tris-Acetate (pH 8.3), 0.5 mM Na_3EDTA (Sigma-Aldrich, T4038-1L).

Oligodeoxynucleotides were synthesized and purified as described Ferrari, Bujalowski and Lohman, 1994} and were ≤ 98% pure as judged by denaturing gel electrophoresis. The poly(dT) (Midland certified reagent company, Midland, TX((Catalog #P-2004, Lot number 071308)), had an average length of ~ 1000 (Midland Certified Reagent Company). DNA samples were dialyzed extensively before use. Single stranded circular M13 mp18 DNA was purchased from New England Biolabs (Catalog #N4040S) and was supplied in 10 mM Tris-HCl (pH 7.5), 1 mM EDTA. Concentrations of the nucleic acids were determined spectrophotometrically using the extinction coefficients: poly(dT), dT_{35} and dT_{70}: ε_{260} = 8.1×10^3 M^{-1} (nucleotide) cm^{-1} {Kowalczykowski, Lonberg, 1981} single stranded M13 DNA: ε_{259} = 7370 M^{-1} cm^{-1} {Berkowitz, Day, 1974}.

SSB tail variants. SSB Δ120-166, Δ130-166, Δ151-166 and SSB GG were generated with the Agilent Technologies Lightning Site-Directed Mutagenesis Kit using *E. coli* wtSSB with the native operon (cloned into pET21a) as the template. The EPE construct was an evolution involving 5 intermediates generated using the *E. coli* or *P. falciparum* SSB as a template.

Mutants were expressed using an auto-induction protocol {Studier, 2005}. An overnight culture was diluted 1:100 into flasks containing 500 ml of ZYM 5052 media {Studier, 2005} with ampicillin and grown at 37°C with shaking at 300 rpm. The cells were harvested in the morning and washed before storage at -80°C until purification.

All variants were purified (>98% pure) as described {Lohman, 1986; Bujalowski, 1991}. Protein concentrations were determined spectrophotometrically {Lohman and Overman, 1985} (buffer T, 0.2M NaCl): ε_{280}=1.13×10^5 M^{-1} cm^{-1} for wtSSB, SSBΔC8, SSBΔ151-166 and ε_{280}=8.98×10^4 M^{-1} cm^{-1} for SSBΔ130-166, SSBΔ120-166, GG and EPE.

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

Agarose gel electrophoresis. Agarose gel electrophoresis to examine the cooperativity of SSB-ssM13 DNA complexes was performed as described {Lohman, Overman, Datta, 1986} using 0.5% agarose gels (14 cm horizontal gels). The total reaction volume of each sample was 30 µl. In a series of experiments, the DNA concentration was held constant (0.8-1.5 nmole nucleotides), while varying the protein concentration. The pre-equilibration solution conditions used are given in the text for each experiment. Electrophoresis was carried out at room temperature (22° C) at constant voltage (~1 V/cm) for 3 to 3.5 h. After completion of electrophoresis, the gel was soaked in a 1 M NaCl buffer T to dissociate bound SSB and the DNA was visualized by staining with ethidium bromide (2 µg/ml solution) and destained for 2-3 hrs at 4° C in buffer T + 2 M NaCl.

Fluorescence measurements. Titrations of SSB constructs with ssDNA were performed by monitoring the intrinsic tryptophan fluorescence {Lohman, 1992; Kozlov, Galletto and Lohman, 2012}. The titration data for dT35 were fit to a two-site binding model using "Scientist" software (Micromath, St.Louis) {Kozlov, Lohman, 2012}

Analytical Sedimentation. Sedimentation velocity experiments were performed using an Optima XL-A analytical ultracentrifuge and an An50Ti rotor (Beckman Instruments, Inc., Fullerton, CA) at 25°C. All protein samples were dialyzed extensively versus buffer T containing the indicated [NaCl]. Experiments were performed at 1.5 - 3 µM (SSB tetramers), except for SSBΔC8, which has much lower solubility in 10 mM NaCl (≤0.3 µM)). Protein solution (380 µl) and buffer (392µl) were loaded in the appropriate sectors of an Epon charcoal-filled two-sector centerpiece and centrifuged at 42000 RPM while monitoring absorbance at 280 nm (or 230-235 nm for the samples with low solubility). The traces were analyzed using the program SEDFIT {Schuck, 1998; Dam, Schuck, 2004} to obtain the c(s) distributions and frictional coefficient ratios. The molecular weight estimates are based on the values $\bar{\upsilon}_i$ calculated from the amino acid compositions using the program SEDINTERP. Values of $s_{20,w}$, presented in Table 1, were converted from fitted experimental values of $s_{25,salt}$ using corresponding values of viscosity and density and assuming that υ is constant.

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

The sedimentation velocity experiments with M13 phage ssDNA were performed at constant DNA concentration (25 or 50 μM nucleotides) while varying protein concentration to achieve different protein/DNA ratios calculated assuming a 35 nucleotide site size (R_{35}=[P]/[M13]$_{nts}$×35) at 15000 RPM (25°C) monitoring absorbance at 260 nm. The contribution of SSB absorbance at R_{35}=0.3) is <~5%. The data were analyzed using SEDFIT {Schuck, 1998; Dam, Schuck, 2004}.

In Vivo Complementation Experiments: *ssb* complementation experiments were performed as described previously {Porter, Black, 1991; Antony et al, 2013}. All experiments were repeated at least twice. Plasmids were also sequenced to confirm that no mutations occurred.

UV Sensitivity experiments. RDP317 containing SSB mutants after the above passage series were tested for UV sensitivity. A colony was picked from a freshly streaked plate and grown overnight. Serial 10-fold dilutions were made and 5μl of the 10^{-2} through 10^{-7} were spotted onto an LB agar plate containing kanamycin and ampicillin. Spotted cultures were allowed to dry into the plate before being exposed to UV. The plates were exposed to shortwave light from a Mineral Light Lamp Model UVGL-25 set to approximately 4 inches above for 6, 12, 18, 24, or 30 seconds or non-exposed.

Atomistic simulations: All simulations were performed using the CAMPARI modeling package (http://campari.sourceforge.net). The forcefield parameters were obtained from the abs_3.2_opls.prm file. In the simulations, all polypeptide atoms and solution ions were modeled in atomic detail. The sequence constructs for the simulations were:

i. *Ac*GGRQGGGAPAGGNIGGGQPQGGWGQPQQPQGGNQFSGGAQSR PQQSAPAAPSNEPPMDFDDDIPF*Nme* designated as **EcEc**;
ii. *Ac*GGRQGGGAPAGGNIGGGQPQGGWGQPQQPQGGNQFSGGPPMDF DDDIPF*Nme* designated as **EcEc**(Δ**151-166**);
iii. *Ac*GGRQGGGAPAGGNIGGGPPMDFDDDIPFNme designated as **EcEc**(Δ**130-166**);
iv. *Ac*GGRQGGGPPMDFDDDIPF*Nme* designated as **EcEc**(Δ**120-166**); and

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

v. *Ac*DDKRNFNQRNNSNNINSENQQHINNEHINNNNINNGNDFMPLNSN DKIIEDKEFTDRLDDNNEENNFQSNSETFDKQEGIYDKMNVQEFEE*Nme* designated as **PfPf**.

Here, *Ac* and *Nme* refer to N-acetyl and N'-methylamide, which we use as capping groups for the N- and C-termini, respectively. Parameters for the mobile Na^+ and Cl^- ions are those of Mao and Pappu (REF8 this is actually Mitsutake, 2003). The screening model was set to mode 2 and the long-range electrostatics correction model designated by the keyword LRER_MC was set to mode 1 for the PP sequence given its high FCR and this parameter for was set to mode 3 for the *E. coli* linker plus tip sequences. The degrees-of-freedom for the Metropolis Monte Carlo simulation were the backbone and sidechain torsion angles, and rigid body translations and rotations of the proteins and solution ions. The effects of solvent-mediated interactions were modeled using the ABSINTH implicit solvation model and forcefield paradigm {Vitalis, Pappu, 2009}. Spherical cutoffs of 10 Å and 14 Å were respectively used for Lennard-Jones interactions and electrostatic interactions between atoms bearing partial charges that belong to neutral charge groups. No cutoffs were deployed for calculating the electrostatic interaction between atoms belonging to charge groups bearing a net charge.

In each simulation, the polypeptide plus solution ions were part of a spherical droplet and the diameter of the droplet was set to be equal to the contour length of the chain, which ensures the avoidance of confinement or finite size artifacts. Neutralizing counterions and excess ion pairs were included to mimic a NaCl concentration of 15 mM. To enhance conformational sampling, we used thermal replica exchange {Mitsutake, 2003} with a schedule that is similar to that used in previous work {Mao, Pappu, 2013}. For each peptide construct, we performed three independent replica exchange simulations. Each simulation comprises of a total 6.1×10^7 independent Monte Carlo moves, with the results of the initial 10^6 moves being set aside as equilibration. Swaps between neighboring thermal replicas were attempted once every 5×10^4 steps. Snapshots were saved once every 5×10^4 steps whereas the polymeric properties were computed once every 500 steps and saved to disk once every 2×10^4 steps.

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

Acknowledgments.

We thank T. Ho for synthesis and purification of the oligodeoxynucleotides. This research was supported in part by the NIH (GM030498 to TML) and NSF (MCB 1121867 to RVP).

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jmb.2014.12.020.

References

[1] Chase JW, Williams KR. Single-stranded DNA binding proteins required for DNA replication. *Annu Rev Biochem* 1986;55:103–36.
[2] Meyer RR, Laine PS. The single-stranded DNA-binding protein of *Escherichia coli. Microbiol Rev* 1990;54:342–80.
[3] Lohman TM, Ferrari ME. *Escherichia coli* single-stranded DNA-binding protein: multiple DNA-binding modes and cooperativities. *Annu Rev Biochem* 1994;63:527–70.
[4] Wold MS. Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu Rev Biochem* 1997;66:61–92.
[5] Shereda RD, Kozlov AG, Lohman TM, Cox MM, Keck JL. SSB as an organizer/mobilizer of genome maintenance complexes. *Crit Rev Biochem Mol Biol* 2008;43:289–318.
[6] 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;260:3594–603.
[7] Bujalowski W, Lohman TM. *Escherichia coli* single-strand binding protein forms multiple, distinct complexes with singlestranded DNA. Biochemistry 1986;25:7799–802.
[8] Chrysogelos S, Griffith J. *Escherichia coli* single-strand binding protein organizes single-stranded DNA in nucleosome-like units. *Proc Natl Acad Sci USA* 1982;79:5803–7.
[9] Griffith JD, Harris LD, Register J. Visualization of SSB-ssDNA complexes active in the assembly of stable RecA-DNA filaments. *Cold Spring Harb Symp Quant Biol* 1984;49:553–9.
[10] Bujalowski W, Overman LB, Lohman TM. Binding mode transitions of *Escherichia coli* single strand binding protein-single-stranded DNA complexes. Cation, anion, pH, and binding density effects. *J Biol Chem* 1988;263:4629–40.
[11] Roy R, Kozlov AG, Lohman TM, Ha T. Dynamic structural rearrangements between DNA binding modes of *E. coli* SSB protein. *J Mol Biol* 2007;369:1244–57.
[12] Lohman TM, Overman LB, Datta S. Salt-dependent changes in the DNA binding co-operativity of *Escherichia coli* single strand binding protein. *J Mol Biol* 1986;187:603–15.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

[13] Overman LB, Bujalowski W, Lohman TM. Equilibrium binding of *Escherichia coli* single-strand binding protein to single-stranded nucleic acids in the (SSB)65 binding mode. Cation and anion effects and polynucleotide specificity. *Biochemistry* 1988;27:456–71.
[14] Ferrari ME, Bujalowski W, Lohman TM. Co-operative binding of *Escherichia coli* SSB tetramers to single-stranded DNA in the (SSB)35 binding mode. *J Mol Biol* 1994;236:106–23.
[15] 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;7:648–52.
[16] Roy R, Kozlov AG, Lohman TM, Ha T. SSB protein diffusion on single-stranded DNA stimulates RecA filament formation. *Nature* 2009;461:1092–7.
[17] Zhou R, Kozlov AG, Roy R, Zhang J, Korolev S, Lohman TM, et al. SSB functions as a sliding platform that migrates on DNA via reptation. *Cell* 2011;146:222–32.
[18] Alberts B, Frey L, DeliusH. Isolation and characterization of gene 5 protein of filamentous bacterial viruses. *J Mol Biol* 1972;68:139–52.
[19] Kowalczykowski SC, Lonberg N, Newport JW, von Hippel PH. Interactions of bacteriophage T4-coded gene 32 protein with nucleic acids. I. Characterization of the binding interactions. *J Mol Biol* 1981;145:75–104.
[20] Giedroc DP, Keating KM, Williams KR, Konigsberg WH, Coleman JE. Gene 32 protein, the single-stranded DNA binding protein frombacteriophage T4, is a zincmetalloprotein. *Proc Natl Acad Sci USA* 1986;83:8452–6.
[21] Lohman TM, Bujalowski W, Overman LB. *E. coli* single strand binding protein: a new look at helix-destabilizing proteins. *Trends Biochem Sci* 1988;13:250–5.
[22] Kozlov AG, Lohman TM. Kinetic mechanism of direct transfer of *Escherichia coli* SSB tetramers between single-stranded DNA molecules. *Biochemistry* 2002;41:11611–27.
[23] Lee KS, Marciel AB, Kozlov AG, Schroeder CM, Lohman TM, Ha T. Ultrafast redistribution of *E. coli* SSB along long single-stranded DNA via intersegment transfer. J Mol Biol 2014;426:2413–21.
[24] Antony E, Weiland E, Yuan Q, Manhart CM, Nguyen B, Kozlov AG, et al. Multiple C-terminal tails within a single *E. coli* SSB homotetramer coordinate DNA replication and repair. *J Mol Biol* 2013;425:4802–19.
[25] Kozlov AG, Cox MM, Lohman TM. Regulation of single-stranded DNA binding by the C termini of *Escherichia coli* single-strandedDNA-binding (SSB) protein. *J Biol Chem* 2010;285:17246–52.
[26] Su XC, Wang Y, Yagi H, Shishmarev D, Mason CE, Smith PJ, et al. Bound or free: interaction of the C-terminal domain of *Escherichia coli* single-strandedDNA-binding protein (SSB) with the tetrameric core of SSB. *Biochemistry* 2014;53:1925–34.
[27] Wessel SR, Marceau AH, Massoni SC, Zhou R, Ha T, Sandler SJ, et al. PriC-mediated DNA replication restart requires PriC complex formation with the single-stranded DNA-binding protein. *J Biol Chem* 2013;288:17569–78.

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

[28] Bhattacharyya B, George NP, Thurmes TM, Zhou R, Jani N, Wessel SR, et al. Structural mechanisms of PriA-mediated DNA replication restart. *Proc Natl Acad Sci USA* 2014;111:1373–8.
[29] Antony E, Weiland EA, Korolev S, Lohman TM. *Plasmodium falciparum* SSB tetramer wraps single-stranded DNA with similar topology but opposite polarity to *E. coli* SSB. *J Mol Biol* 2012;420:269–83.
[30] Antony E, Kozlov AG, Nguyen B, Lohman TM. *Plasmodium falciparum* SSB tetramer binds single-stranded DNA only in a fully wrapped mode. *J Mol Biol* 2012;420:284–95.
[31] Williams KR, Spicer EK, LoPresti MB, Guggenheimer RA, Chase JW. Limited proteolysis studies on the *Escherichia coli* single-stranded DNA binding protein. Evidence for a functionally homologous domain in both the *Escherichia coli* and T4 DNA binding proteins. *J Biol Chem* 1983;258:3346–55.
[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-Å resolution. *Proc Natl Acad Sci USA* 1997;94:6652–7.
[33] Savvides SN, Raghunathan S, Futterer K, Kozlov AG, Lohman TM, Waksman G. The C-terminal domain of full-length *E. coli* SSB is disordered even when bound to DNA. *Protein Sci* 2004;13:1942–7.
[34] Arif SM, Vijayan M. Structural diversity based on variability in quaternary association. A case study involving eubacterial and related SSBs. *Methods Mol Biol* 2012;922:23–35.
[35] Dam J, Schuck P. Calculating sedimentation coefficient distributions by direct modeling of sedimentation velocity concentration profiles. *Methods Enzymol* 2004;384:185–212.
[36] Vitalis A, Pappu RV. ABSINTH: a new continuum solvation model for simulations of polypeptides in aqueous solutions. *J Comput Chem* 2009;30:673–99.
[37] Mao AH, Crick SL, Vitalis A, Chicoine CL, Pappu RV. Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc Natl Acad Sci* USA 2010;107:8183–8.
[38] Mao AH, Lyle N, Pappu RV. Describing sequence-ensemble relationships for intrinsically disordered proteins. *Biochem J* 2013;449:307–18.
[39] Das RK, Pappu RV. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc Natl Acad Sci USA* 2013;110:13392–7.
[40] Bujalowski W, Lohman TM. Negative co-operativity in *Escherichia coli* single strand binding protein-oligonucleotide interactions. II. Salt, temperature and oligonucleotide length effects. *J Mol Biol* 1989;207:269–88.
[41] Lohman TM, Bujalowski W. Effects of base composition on the negative cooperativity and binding mode transitions of *Escherichia coli* SSB-single-stranded DNA complexes. *Biochemistry* 1994;33:6167–76.
[42] Sigal N, Delius H, Kornberg T, Gefter ML, Alberts B. A DNA unwinding protein isolated from *Escherichia coli*: its interaction with DNA and with DNA polymerases. *Proc Natl Acad Sci USA* 1972;69:3537–41.

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

[43] Ruyechan WT, Wetmur JG. Studies on the cooperative binding of the *Escherichia coli* DNA unwinding protein to single-stranded DNA. *Biochemistry* 1975;14:5529–34.
[44] Schneider RJ, Wetmur JG. Kinetics of transfer of *Escherichia coli* single strand deoxyribonucleic acid binding protein between single-stranded deoxyribonucleic acid molecules. *Biochemistry* 1982;21:608–15.
[45] Porter RD, Black S. The single-stranded-DNA-binding protein encoded by the *Escherichia coli* F factor can complement a deletion of the chromosomal ssb gene. *J Bacteriol* 1991;173:2720–3.
[46] 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;24:2706–11.
[47] Bonura T, Smith KC. Enzymatic production of deoxyribonucleic acid double-strand breaks after ultraviolet irradiation of *Escherichia coli* K-12. *J Bacteriol* 1975;121:511–7.
[48] Lohman TM. Kinetics and mechanism of dissociation of cooperatively bound T4 gene 32 protein single-stranded nucleic acid complexes. 1. Irreversible dissociation induced by sodium-chloride concentration jumps. *Biochemistry* 1984;23:4656–65.
[49] Pappu RV, Wang X, Vitalis A, Crick SL. A polymer physics perspective on driving forces and mechanisms for protein aggregation. *Arch Biochem Biophys* 2008;469:132–41.
[50] Berkowitz SA, Day LA. Molecular weight of single-stranded fd bacteriophage DNA. High speed equilibrium sedimentation and light scattering measurements. *Biochemistry* 1974;13:4825–31.
[51] Studier FW. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* 2005;41:207–34.
[52] Lohman TM, Green JM, Beyer RS. Large-scale overproduction and rapid purification of the *Escherichia coli* ssb gene product. Expression of the ssb gene under lambda PL control. *Biochemistry* 1986;25:21–5.
[53] Bujalowski W, Lohman TM. Monomer-tetramer equilibrium of the *Escherichia coli* ssb-1 mutant single strand binding protein. *J Biol Chem* 1991;266:1616–26.
[54] Kozlov AG, Galletto R, Lohman TM. SSB-DNA binding monitored by fluorescence intensity and anisotropy. *Methods Mol Biol* 2012;922:55–83.
[55] Mao AH, Pappu RV. Crystal lattice properties fully determine short-range interaction parameters for alkali and halide ions. *J Chem Phys* 2012;137:064104.
[56] Mitsutake A, Sugita Y, Okamoto Y. Replica-exchange multi-canonical and multi-canonical replica-exchange Monte Carlo simulations of peptides. II. Application to a more complex system. *J Chem Phys* 2003;118:6676–88.

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

Figure Legends

Figure 1. Structural organization of *Ec*SSB.

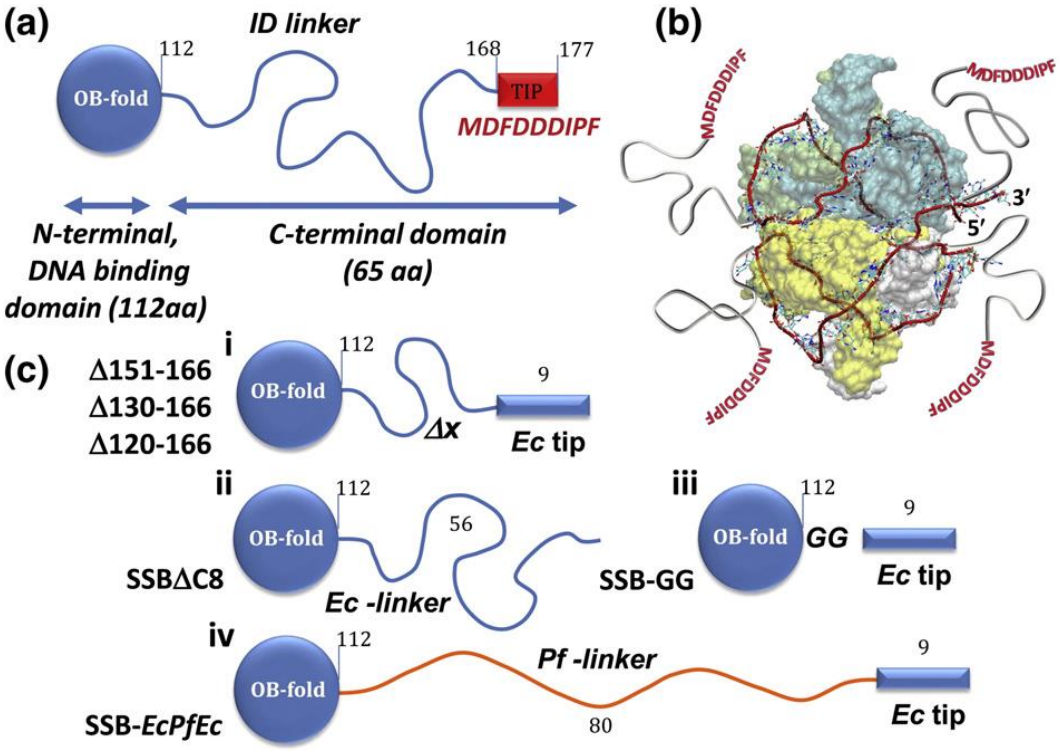

Fig. 1. Structural organization of *E. coli* SSB linker variants. (a) wt*Ec*SSB subunits (177 aa) are composed of an N-terminal DNA binding domain (OB fold) (1–112 residues) and a C-terminal IDLs (56 aa) and a 9-residue acidic "tip". (b) Structural model of 65 nt of ssDNA (red ribbon) bound to the *Ec*SSB tetramer [15] with the addition of the C-terminal tails (shown in gray) that are not observed in the crystal structure. (c) Design of SSB linker variants containing the *Ec core* (1–112 aa), but varying C-terminal tails: missing part of the linker (i), the whole linker [GG (ii)], acidic "tip" [SSBΔC8 (iii)] or replacing the *Ec* linker with the *Pf* linker [*EcPfEc*, (iv)].

Figure 2.

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

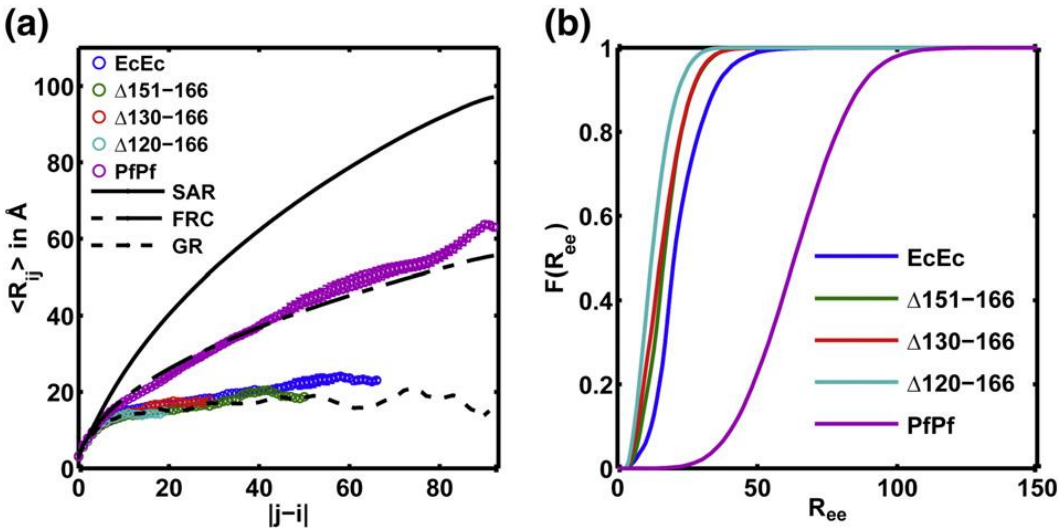

Fig. 2. Results from simulations summarizing the conformational properties of IDLs. (a) Plot of the internal scaling profiles for the four EcEc linker tail variants and the PfPf linker + tail. The profiles for globular (GR), Flory random coils (FRC) and self-avoiding reference (SAR) ensembles are also shown for calibration. (b) Plot of the cumulative distribution functions of Ree values for each of the linker variants. The mean Ree values are 22.6 ± 0.7 Å for EcEc, 17.8 ± 0.8 Å for EcEc(Δ151-166), 16.8 ± 1.9 Å for EcEc(Δ130-166), 13.7 ± 0.1 Å for EcEc(Δ120-166) and 63.6 ± 0.9 Å for PfPf.

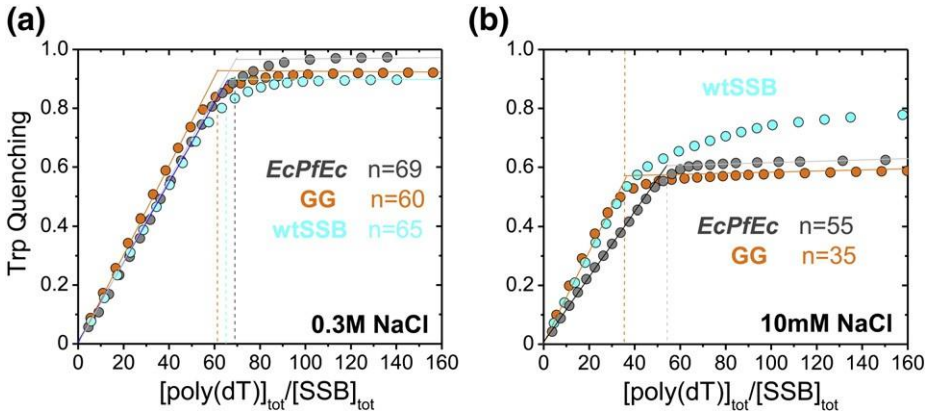

Fig. 3. Occluded site sizes of SSB linker variants on poly(dT). Results of titrations of 0.30 μM wtSSB (cyan), SSB-GG (orange) and SSB-*EcPfEc* (dark gray) with poly(dT) monitoring SSB Trp fluorescence quenching in buffer T, 25 °C with (a) 0.30 M NaCl or (b) 10 mM NaCl, to estimate the occluded site sizes (n) for the SSB variants bound to poly(dT) [54].

Figure N3: Summary of atomistic simulation results. Panel A shows the internal scaling profiles for the four *E. coli* constructs and one Pf construct. The error bars denote estimates in the errors of the mean Rij values. The reference profiles shown in black were obtained by performing simulations for the PP sequence with only Lennard-Jones interactions turned on (the globule reference), as a statistical coil (the Flory random coil reference), and with only steric repulsions turned on (the self avoiding reference).

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

Panel B shows cumulative distribution functions (CDFs) for each of the sequence constructs. Each CDF quantifies the probability associated with realizing a particular value of Ree such F(X) is the probability that the sequence in question will have an Ree value that is less than or equal to X. For example, according to the CDFs, the probability that Ree ≤ 50Å is approximately 0.2 for the PP sequence and is essentially unity for all the EE constructs. Panel C shows CDFs that enable comparative assessments of different sequence constructs adopting particular values for Rg/N1/2, which is a chain-length (N) normalized value of the radius of gyration. This enables a comparison of the density distributions for different sequences.

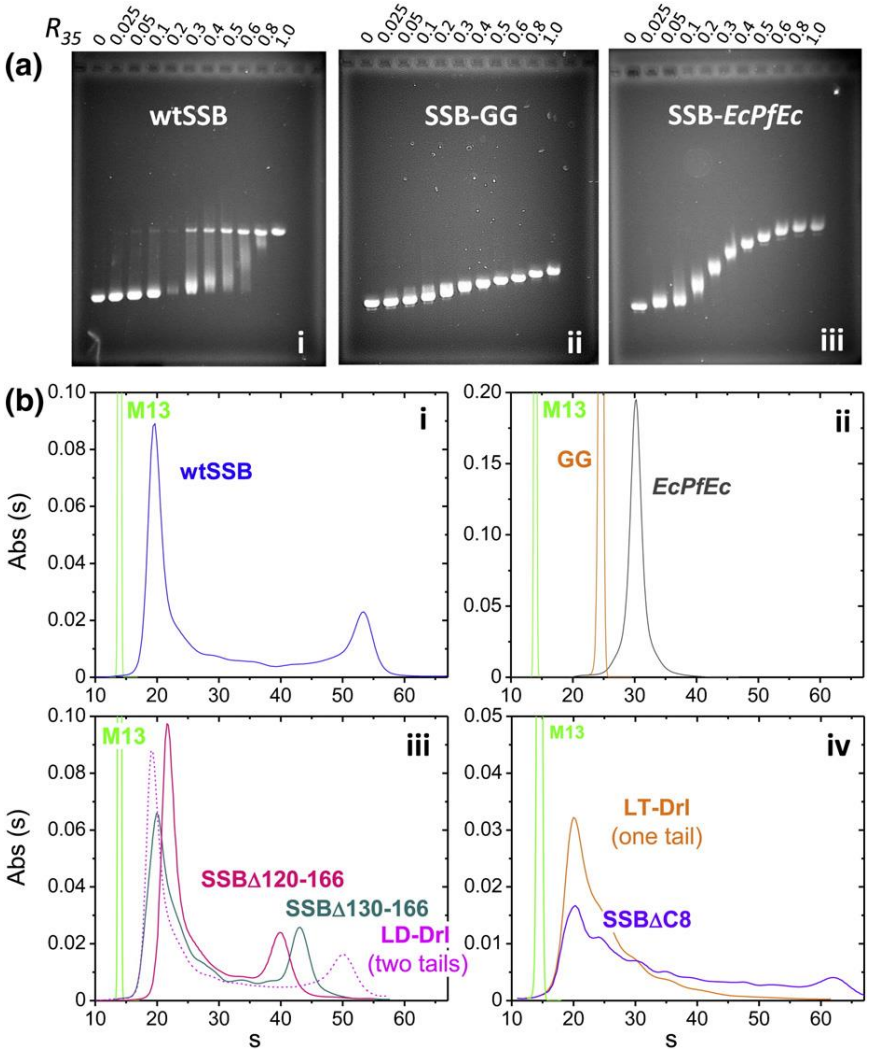

Fig. 4. Cooperativity of SSB linker variants bound to M13 ssDNA. (a) EMSAs of SSB–M13 mp18 ssDNA complexes in buffer T, 10 mM NaCl, 22°, formed at different protein-to-DNA ratios: R_{35} = 35[SSB]$_{tot}$/[M13nts]$_{tot.}$ wtSSB (a-i) shows a bimodal distribution of bound DNA at intermediate ratios (R_{35} = 0.05–0.7) indicative of highly cooperative binding, whereas SSB-GG (a-ii) and SSB-*EcPfEc* (a-iii) show single band at all protein-to-DNA ratios indicative of low or no cooperativity. (b) Sedimentation velocity profiles of SSB variants bound to M13 ssDNA at R_{35} = 0.3 (buffer T, 10 mM NaCl, 25 °C). Bimodal distribution of wtSSB–M13 complexes (b-i) indicates highly cooperative binding. SSB-GG and SSB-*EcPfEc* (b-ii) bind with low cooperativity.

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

Intermediate linker deletion constructs SSBΔ130-166 and SSBΔ120-166, as well as two-tailed variant LD-DrI, retain high cooperativity (b-iii). SSB missing conserved acidic "tip" (SSBΔC8) and single tail SSB construct, LT-DrI (b-iv), show decreased cooperativity (see Table S1 for quantification).

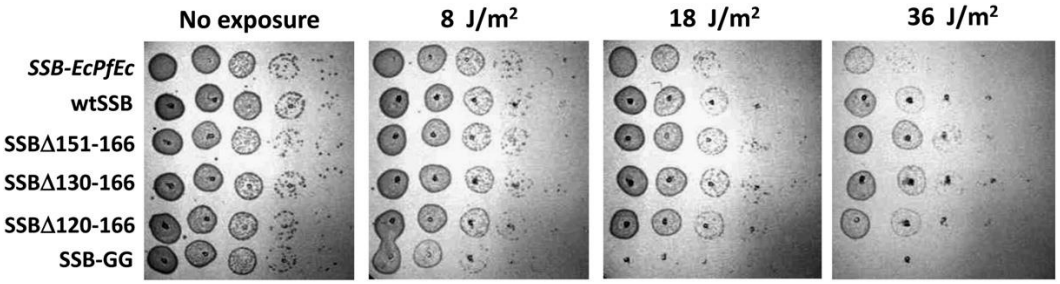

Fig. 5. Sensitivity of *E. coli* strains carrying the SSB linker variants to UV irradiation. The sensitivity to uv irradiation of *E. coli* cells carrying the SSB linker deletion variants SSBΔ151-166, SSBΔ130-166 and SSBΔ120-166 is very similar to wtSSB, although sensitivity increases for the SSB-*EcPfEc* variant and even more for the SSB-GG variant.

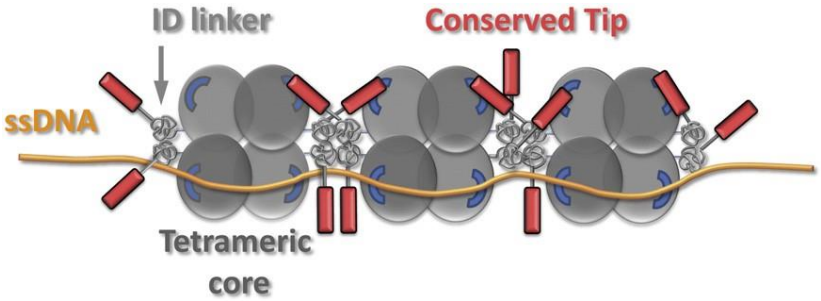

Fig. 6. Model for highly cooperative ssDNA binding in the (SSB)35 mode. Binding of SSB to ssDNA in its $(SSB)_{35}$ binding mode leaves DNA binding sites unoccupied in two subunits. Negatively charged tips of two C-terminal tails bind to the unoccupied subunits of adjacent tetramers. Positive cooperativity is enhanced due to inter-chain interactions between globular IDLs of adjacent tetramers.

Address correspondence to:
Department of Biochemistry and Molecular Biophysics, Box 8231
Washington University School of Medicine
660 South Euclid Ave.
St. Louis, M0 63110
E-mail: lohman@biochem.wustl.edu
Tel: (314)-362-4393
FAX: (314)-362-7183

Journal of Molecular Biology, Vol. 427, No. 4 (February 27, 2015): pg. 763-774. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.