DNA damage alters binding conformations of *E. coli* single-stranded DNA binding protein

Michael Morse[1†], Francesco Navarro Roby[2†], Mansi Kinare[2], James McIsaac[2], Mark C. Williams[1*], Penny J. Beuning[2*]

[1] Department of Physics, Northeastern University, Boston, MA, 02115, USA

[2] Department of Chemistry and Chemical Biology, Northeastern University, Boston, MA, 02115, USA

† These authors contributed equally to this work

* Corresponding Authors: p.beuning@northeastern.edu and ma.williams@northeastern.edu

Keywords: atomic force microscopy, single molecule imaging, cooperative binding, abasic site

Single-stranded DNA binding proteins (SSBs) are essential cellular components, binding to transiently exposed regions of single-stranded DNA (ssDNA) with high affinity and sequence non-specificity to coordinate DNA repair and replication. *E. coli* SSB (*Ec*SSB) is a homotetramer that wraps variable lengths of ssDNA in multiple conformations (typically occupying either 65 or 35 nts), which is well-studied across experimental conditions of substrate length, salt, pH, temperature, etc. In this work, we use atomic force microscopy (AFM) to investigate the binding of SSB to individual ssDNA molecules. We introduce non-canonical DNA bases that mimic naturally occurring DNA damage, synthetic abasic sites, as well as a non-DNA linker into our experimental constructs at sites predicted to interact with *Ec*SSB. By measuring the fraction of DNA molecules with *Ec*SSB bound as well as the volume of protein bound per DNA molecule, we determine the protein binding affinity, cooperativity, and conformation. We find that with only one damaged nucleotide, the binding of *Ec*SSB is unchanged relative to its binding to undamaged DNA. In the presence of either two tandem abasic sites or a non-DNA spacer, however, the binding affinity associated with a single *Ec*SSB tetramer occupying the full substrate in the 65 nt mode is greatly reduced. In contrast, the binding of two *Ec*SSB tetramers, each in the 35 nt mode, is preserved. Changes in the binding and cooperative behaviors of *Ec*SSB across these constructs can inform how genomic repair and replication processes may change as environmental damage accumulates in DNA.

Statement of Significance:

Single stranded binding proteins (SSBs) bind transiently exposed single-stranded DNA (ssDNA) during DNA replication, recombination, and repair. SSBs both protect ssDNA from degradation and recruit additional proteins to aid in essential cellular processes. *E. coli* SSB (*Ec*SSB), a well-studied model system, binds ssDNA in multiple conformations, occluding variable lengths of substrate. We examine *Ec*SSB binding to ssDNA substrates at a single molecule level and find that modifying DNA to imitate naturally occurring DNA damage alters the preferred binding conformation of *Ec*SSB without reducing its high binding affinity. Our results suggest that *Ec*SSB can bind damaged ssDNA in a site directed manner that could help facilitate specific remediation of individual bases.

Introduction

Single stranded binding proteins (SSBs) are a class of proteins that bind preferentially to single stranded DNA (ssDNA) with high affinity. This binding specificity allows SSBs to quickly and stably bind regions of ssDNA that are transiently exposed during essential cellular processes such as DNA replication, recombination, and repair (1-4). The presence of SSB prevents the formation of secondary structure that can inhibit polymerization and degradation by nucleases. SSBs can also recruit other proteins to perform genome maintenance functions (5-9).

The SSB of *E. coli* (*Ec*SSB), perhaps the most well studied SSB, is a stable homotetramer, with each 177 amino acid, 18.9 kDa subunit containing an oligonucleotide/oligosaccharide binding (OB) fold and a disordered C-terminal tail (10,11). The OB folds are structured, both individually and when forming a tetramer (12,13), and each can independently bind ssDNA substrates. Thus, depending on substrate length and solution conditions, a single ssDNA can wrap around the OB fold tetramer in different conformations. In its largest binding size conformation, an *Ec*SSB tetramer can accommodate a single 65 nucleotide (nt) ssDNA that binds to each OB fold as it wraps around the tetramer surface. This conformation is most stable *in vitro* at high salt concentrations and low ratios of protein to DNA (14). At lower salt concentrations and in the presence of excess protein, however, *Ec*SSB can bind to a 35 nt length of ssDNA (15), such that more tetramers can be accommodated on a substrate of defined length. In the 35 nt state, not all OB folds directly interact with the ssDNA, and a structure has been resolved in which two 35 nt ssDNAs bind to one *Ec*SSB tetramer (12). Besides the main 65 and 35 states, other less stable binding states have been proposed or observed under different experimental conditions (16,17). In contrast, the C-terminal tail, which consists of an acidic tip attached to the OB fold by a long, disordered linker, does not function primarily through direct interaction with ssDNA substrates. Instead, the C-terminal tail primarily interacts with other proteins, including other *Ec*SSB tetramers (18-21).

While *Ec*SSB must be able to bind ssDNA in a sequence independent manner, such as when it cycles over the full genome during replication, there is possibility for substrate/sequence specific effects. In addition, *Ec*SSB plays a vital role in DNA repair (22) and localizes in response to DNA damage (23). DNA damage can result from chemical reactions, radiation exposure, and enzymatic activity (24). In particular, depurination of DNA bases results in apurinic/apyrimidinic (AP) sites, also known as abasic sites. Unrepaired AP sites can stall DNA polymerization at the replication fork and result in mutation (25-28). *Ec*SSB has been shown to play a role in the SOS response to DNA damage, interacting with repair proteins such as RecA (29), RadD (30), and Exonuclease I (31), and cells deficient in SSB display increased mutagenesis (32). It is less understood, however, if *Ec*SSB itself binds to sites of DNA damage differently than to undamaged, canonical DNA bases.

To test the hypothesis that sites of DNA damage modulate *Ec*SSB binding, we directly observe *Ec*SSB binding *in vitro* at a single molecule level to ssDNA substrates with modified bases that mimic DNA damage. Due to the prevalence of abasic sites in DNA, occurring spontaneously

approximately once per generation in *E. coli* and more frequently under stress conditions (33,34), we chose to utilize a stable abasic site mimic in these studies. We utilized a 67 nt long sequence which can accommodate one tetramer in the 65 state or two in the 35 state (without excess unbound ssDNA) and chose bases for modification expected to interact closely with bound proteins based on structural models (12). While *Ec*SSB is able to bind these damaged substrates with nM affinity, similar to its binding to undamaged DNA, we find the exact binding conformation is modulated, favoring the simultaneous binding of two proteins flanking the damage site, even under conditions for which the binding of a single protein is favored for the undamaged ssDNA substrate. These results suggest that the exact spatial binding pattern of *Ec*SSB along a longer length of ssDNA could be influenced by specific locations of DNA damage, which in turn could provide a mechanism to direct DNA repair machinery.

Materials and Methods

Protein purification

WT *Ec*SSB was purified as previously described (16,35). The plasmid encoding WT *Ec*SSB pEAW134 was a gift from Dr. Mark Sutton of the University at Buffalo. Briefly, *Ec*SSB was expressed in *E. coli* BL21 Tuner cells, precipitated with Polymin P followed by ammonium sulfate, and then further purified on an ssDNA-cellulose column. Protein concentration was determined spectroscopically using an extinction coefficient of $\varepsilon_{280} = 1.13\times10^5$ $M^{-1}cm^{-1}$ for the *Ec*SSB tetramer (35).

DNA binding substrates

Hybrid ds/ssDNA substrates were prepared as previously described (16). A 268 bp dsDNA was produced by PCR amplification using pUC19 plasmid template and Taq DNA polymerase (New England Biolabs, NEB). The primers are listed in Table 1, and 30X PCR cycles of denaturing (95 °C, 30 s), annealing (56 °C, 30 s), and extension (68 °C, 60 s) were performed. The product was digested by BamHI (37 °C, 4 h), resulting in a 248 bp dsDNA with a 4 nt 5′ overhang. The cut dsDNA was then incubated with a 10X molar excess of ssDNA substrates and linker oligos (Table 1). The DNA mixture was heated to 50 °C for 5 min then gradually cooled to 16 °C, allowing the linker oligo to anneal to both the dsDNA overhang and the ssDNA substrate. The constructs were ligated overnight (4 °C, 16 h) with T4 DNA ligase. The sample was gel purified to remove excess ssDNA, ensuring all ssDNA in the final product is ligated to a dsDNA marker. The final ligated product is 260 bp of dsDNA with a 67 nt ssDNA overhang. In addition to the ds/ss DNA hybrid, ligation of the dsDNA to itself produced a 500 bp dsDNA and this additional gel band was also excised for use as a molecular calibration ruler. All DNA oligos and enzymes were purchased from Integrated DNA Technologies and New England Biolabs, respectively.

AFM imaging and analysis

Varying concentrations of *Ec*SSB were added to DNA binding substrates diluted to a concentration of 1 nM in a buffer containing 145 mM NaCl, 5 mM NaOH, 100 µM spermidine, and 10 mM HEPES (pH 7.5). Samples were incubated for 5 min at 37 °C, then 5 µL of solution was deposited on a freshly cleaved mica surface. After 1 min of deposition time, the mica was washed with an excess of DI water and then dried with Argon gas. The sample was imaged with a MultiMode 8 AFM and Nanoscope V controller (Bruker) using peak force tapping mode and analyzed using Gwyddion software (version 2.55). For each condition measured (specific ssDNA substrate and *Ec*SSB concentration), three or more biological replicates of *Ec*SSB-ssDNA incubation were performed. For each incubation, the surface was imaged in multiple locations to observe a large number of ssDNA substrates. Though the exact number of substrates in each image frame naturally varies, an average of 408 substrates (ranging from 219 to 558) were imaged per condition. Standard error of the mean (error bars in plots) was calculated based on deviations in average binding fraction and protein volume per distinct incubation.

Results

Measuring *Ec*SSB-ssDNA binding

The binding of unlabeled protein to unlabeled, short ssDNA substrates is difficult to detect at a single molecule level by many popular methods. For AFM imaging specifically, the large size of the *Ec*SSB tetramer (75.5 kDa) obscures the presence of any tightly bound ssDNA substrate small enough to be fully bound by the protein (~65 nt or ~20 kDa). Thus, free proteins and proteins bound by unlabeled oligos are nearly indistinguishable, preventing accurate measurement of binding affinities. In this study, we utilize a ds/ssDNA hybrid (36), where the ssDNA substrate of interest is ligated to the end of a dsDNA marker (Fig 1A). The rigidity of dsDNA both prevents its binding to *Ec*SSB directly and increases its visibility in AFM imaging, where the 260 bp marker appears as an 85 nm long line. Additionally, the dsDNA region controls the hybrid construct's migration through a gel, ensuring unligated ssDNA, which migrates further through the gel, is completely removed from the sample during purification. Thus, *Ec*SSB tetramers bound to the ssDNA substrate will only colocalize with one end of the dsDNA marker (Fig 1B), enabling accurate numeration of ssDNA substrates with and without protein bound. This method allows for the ligation of any ssDNA substrate to the dsDNA marker and for this project we utilize a 67 nt sequence from M13 bacteriophage which our lab has previously used in polymerase assays (37). The sequence has 54% GC content and contains limited secondary structure (no large stable hairpins) and can be considered representative of the mixed base composition, naturally occurring sequences with which *Ec*SSB interacts. The length was chosen such that either one *Ec*SSB tetramer can bind ~67 nt or two tetramers can bind ~33.5 nt each, as the exact binding site size of the two conformations have been determined to be 35±2 nt and 65±3 nt (12,38,39). We intentionally chose a substrate on the

smaller end of this range to limit protein shifting on the substrate (40,41), which could impact which nucleotides interact with specific protein residues.

To measure the binding affinity of *Ec*SSB to our ssDNA substrates, we incubated varying concentrations of *Ec*SSB with a fixed 1 nM concentration of DNA in a 150 mM Na buffer, as moderate salt conditions allow *Ec*SSB to bind ssDNA in both the 35 and 65 nt mode. Samples were incubated at 37 °C for 5 min to ensure equilibrium binding, in accordance with previous kinetic measurements showing equilibration of ssDNA with 100 pM *Ec*SSB occurring on a 100 s timescale (16). The number of ssDNA substrates bound or unbound by *Ec*SSB are counted at each protein concentration (Fig 1D). As expected, the fraction of ssDNA that are bound by protein increases with *Ec*SSB concentration (*c*), with a trend that can be well fit as:

(1) $$f(c) = \frac{c}{c + (K_D)}$$

Here, K_D is an effective dissociation constant, the protein concentration at which half the substrates are bound. Note, the concentration used in this analysis is the concentration of free protein after the system equilibrates, not the concentration initially added to the incubation. The free protein concentration is calculated by subtracting the product of total ssDNA concentration (1 nM), the fraction of ssDNA bound (*f*), and the average bound state from the total protein concentration:

(2) $$c_{free} = c_{total} - \left([ssDNA] \cdot f \cdot \langle Bound\ State \rangle\right)$$

The bound state is either one or two proteins per bound ssDNA and the average value of which is measured for each condition (discussed later). This simple binding isotherm fits the experimental data within experimental error (Reduced $\chi^2 \approx 1$) with the ssDNA substrates transitioning from mostly unbound to mostly bound around a free protein concentration of 1 nM.

*Ec*SSB binding to damaged ssDNA

The binding affinity experiment was repeated using different ssDNA substrates (Fig 2A), each the same length, but with the 27th and/or the 28th base (from the 5′ end) modified. When a single base was replaced with an AP site, we observed nearly identical *Ec*SSB binding for these substrates (Fig 2B). The fraction of ssDNA bound is still well fit by a simple binding isotherm with an effective K_D similar to that for binding to the undamaged ssDNA.

When both the 27th and 28th bases were replaced with AP sites, however, a different binding behavior was observed (Fig 2C). To further test specificity, we also utilized an ssDNA with both the 27th and 28th bases replaced by a triethylene glycol spacer (SP), which gave the same results. While comparable levels of ssDNA binding are observed for protein concentrations above K_D (approaching saturation), the fractional binding at lower protein concentrations is reduced. As a result, we observe a sharp transition in binding, where the ssDNA goes from

mostly free to mostly bound over a small *Ec*SSB concentration increase. Thus, fitting with the simple binding isotherm again returns the same approximate K_D (Fig 2D), but is a poor fit to the data (Reduced $\chi^2 >> 1$, Fig 2E). One possibility is that the removal of the 27th and 28th C bases in the damaged constructs reduces the ability of the ssDNA to form secondary structure, and stable secondary structures would be expected to inhibit *Ec*SSB binding. However, the opposite effect is observed, with less binding observed for the damaged constructs, suggesting the removal of ssDNA secondary structure is not the primary cause of the altered binding behavior. The more likely issue is that this model does not take into account the multiple binding modes of *Ec*SSB, and especially the cooperativity associated with the simultaneous binding of two proteins. Thus, we must examine the binding stoichiometry of the ssDNA-*Ec*SSB complexes to better understand this behavior.

*Ec*SSB binding stoichiometry and cooperativity

AFM imaging provides additional information through volumetric measurement of objects. By integrating over all pixels associated with an object (multiplying pixel height by area and summing over all pixels), the volume sterically occupied by an object can be measured. While the exact value is influenced by some conditions external to the measured object (particularly the size of the AFM tip itself), it has been shown that measured volumes of multiple proteins scale linearly with their molecular weights when calibrated with a common fixed marker, such as DNA (42). Correspondingly, we independently establish this linear relationship using proteins studied in our lab (16,43,44) and a 500 bp DNA as a calibration marker (Fig 3A). We verified this method can distinguish between one and two bound *Ec*SSB tetramers by incubating *Ec*SSB with the undamaged ssDNA substrate in 20 and 300 mM Na buffer, where the 35 and 65 nt binding modes, respectively, are known to predominate (45). AFM images reveal the bound protein complexes to be noticeably larger in 20 mM Na (Fig 3B) than in 300 mM Na buffer (Fig 3C). By measuring the average volume of these protein clusters under both conditions and converting the volume to an estimated molecular mass using the same 500 bp DNA marker, we find that the protein-ssDNA complexes are in fact consistent with either one or two tetramers bound to the substrate (Fig 3D).

We measured the average protein volume for all observed *Ec*SSB concentrations and used this conversion process to determine the average binding stoichiometry. For the undamaged ssDNA, only one tetramer is present for bound substrates when the concentration of ssDNA is equal to or exceeds the concentration of *Ec*SSB tetramers (Fig 4A). As more *Ec*SSB is added to the system, however, more ssDNA substrates are bound by two tetramers. At the highest *Ec*SSB concentration measured (10 nM), the average protein cluster size is consistent with all substrates being bound by two tetramers. A similar transition has been observed at moderate salt concentrations using labeled ssDNA in a FRET assay (45). In contrast, for the damaged ssDNA substrates harboring two tandem abasic sites or the spacer, we observe mostly volumes consistent with two *Ec*SSB tetramers bound to ssDNA even at low *Ec*SSB concentration (Fig 4B). Even when the ssDNA and *Ec*SSB tetramers are at equimolar concentration (1 nM), we observe more substrates bound by two tetramers than by one; most substrates are protein free as shown in Fig 2C, which explains how there is enough *Ec*SSB in the system to achieve this

stoichiometry. Thus, in both the measurements of ssDNA binding fraction (Fig 2) and *Ec*SSB stoichiometry (Fig 4), we observe similar binding behavior at high *Ec*SSB concentrations, where binding of two tetramers in the 35 mode predominates, regardless of DNA damage, but at low *Ec*SSB concentrations the otherwise preferred binding of one tetramer in the 65 mode is inhibited by the tandem DNA damage sites.

Discussion

Multimode model of *Ec*SSB binding conformation

Due to *Ec*SSB's multiple binding modes and the ability of the ssDNA to accommodate up to two tetramers, each substrate can be in one of four conformations: unbound, bound by one tetramer in the 35 or 65 wrapping mode, or bound by two tetramers both in the 35 mode (Fig 5A). Transitions between these states occur when a tetramer binds or dissociates from a substrate or when an already bound tetramer swaps its binding conformation. A model that describes the occupancy of each state, accounting for cooperative binding between neighboring proteins in the 35 mode has been previously developed (15). The average number of tetramers bound per substrate (ν) can solved for in terms of effective association constants (K_{35} and K_{65}) and cooperativity parameter (ω_{35})

$$\nu = \frac{(S_1 K_{35} + K_{65}) P_f + 2\omega_{35}(K_{35} P_f)^2}{1 + (S_1 K_{35} + K_{65}) P_f + \omega_{35}(K_{35} P_f)^2} \quad (3)$$

Note S_1 is a unitless statistical factor enumerating the number of exact sites where the tetramer can bind based on the excess length of substrate relative to the binding site size (S_1 = 67 - 35 + 1 = 33 for this system).

We apply this quantitative model to our data (Fig 5B). The average number of proteins bound per ssDNA substrate is calculated by multiplying the fraction of DNA substrates that are bound by protein (Fig 1 and 2) by the average number of *Ec*SSB tetramers for each bound substrate (Fig 4). Note this is also equivalent to calculating the average number of proteins per substrate using the average protein volume for all substrates where the protein volume for unbound substrates is zero. For undamaged DNA, this value increases gradually, first approaching one tetramer per ssDNA as *Ec*SSB concentration surpasses the effective K_D and then approaching two tetramers per ssDNA at higher concentrations.

Eq. 3 fits the data, though the exact fitting parameters are not well defined (a large range of values fit within error and reduced $\chi^2 < 1$). In particular, since the values of K_{35} and ω_{35} are multiplied in the squared term, their values are directly dependent on that of the other. We set the minimum value for the cooperativity factor ($\omega_{35} = 10^5$) from the original work (15) which used comparable conditions (69 nt poly dA substrate, 125 mM NaCl, pH 8.1, 25 °C versus our conditions of 67 nt mixed base substrate, 145 mM NaCl + 5 mM NaOH, pH 7.5, 37 °C), and found best fit parameters of $K_{65} = (1.22 \pm 0.67) \times 10^8$ M^{-1} and $K_{35} = (9.96 \pm 1.01) \times 10^5$ M^{-1} for ω_{35} =

2×10^5 with a reduced χ^2 value of 0.762. A similar trend was observed in the original work with the association constant for the 65 state multiple orders of magnitude larger than the 35 state ($K_{65} = 1.6\times10^8$ M^{-1} and $K_{35} = 1.6\times10^5$ M^{-1}) (15). Thus, while our assay does not directly measure *Ec*SSB wrapping conformation, this model predicts that singly bound tetramers predominately occupy the 65 state, and the 35 state is primarily observed when two tetramers bind the same substrate. This is consistent with published results from a FRET assay with the termini of the substrate labeled that directly detects ssDNA conformation (45).

For the ssDNA substrates with two damage sites, a different *Ec*SSB binding response is observed (Fig 5C). Compared to the undamaged substrate, significantly less binding is observed in the ~1:1 ssDNA to protein regime, where the substrate should be predominately bound by tetramers in the 65 nt state. In contrast, full binding is recovered at high *Ec*SSB concentrations where more substrates should be bound by two tetramers, each in the 35 nt state. As a result, a sharper transition from ~0 tetramers per substrate at low *Ec*SSB concentrations to ~2 tetramers per substrate at high *Ec*SSB concentrations is observed, with a narrower concentration range of *Ec*SSB showing binding of ~1 tetramer per substrate. The simplest explanation that explains these features is that the binding affinity associated with the non-cooperative 65 state is reduced, while the ability of *Ec*SSB to bind in the cooperative 35 state is unchanged or even slightly enhanced. Altering the parameters of Eq. 3 accordingly (reducing K_{65}) can improve the fit, though least-squares fitting returns a non-physical value of $K_{65} < 0$. Instead, if we constrain K_{65} to non-negative values, the best fit values for K_{35} are not-significantly different than the undamaged substrate. However, these best fits still produce $\chi^2 > 1$, suggesting there may be more complex effects occurring that cannot be described by Eq. 3.

Biological implications of damage site directed binding

The binding conformations of *Ec*SSB are denoted the 65 and 35 nt state in reference to their total binding site size (*i.e.* the occluded length of substrate to which other *Ec*SSB tetramers cannot bind). Not all nts along this length interact with the protein to the same degree, however, and both structures and models of ssDNA bound to *Ec*SSB show some nts are in direct contact with the *Ec*SSB OB folds (12), while others span the distances between sites as the ssDNA wraps around the tetramer. As such, the degree to which *Ec*SSB binding is altered by modified bases likely depends on their exact position. On a substrate significantly longer than its binding site size, an *Ec*SSB tetramer could slide via reptation (40,41), changing the specific nts interacting with specific amino acids. For this reason, we utilized a 67 nt substrate in this study, which restricts the binding of a protein in the single 65 nt state or two proteins in the 35 nt state with minimal sliding. A structure based on the model of the 65 nt state (Fig 5D) shows that the 27th and 28th base (from the 5' end) of a bound ssDNA (12) are stacked between a phenylalanine (residue 60) of one *Ec*SSB subunit and a tryptophan (residue 40) of another subunit. When both these nts are replaced with abasic sites or a non-DNA linker this interaction is lost, which is likely responsible for the reduced binding observed under conditions in which

non-cooperative binding of the 65 mode should predominate due to the near equimolar ratio of ssDNA and *Ec*SSB tetramers.

Another recent study has also investigated whether modulating a fixed length ssDNA substrate can impact *Ec*SSB binding modes by reversing the polarity of the ssDNA backbone (46). When a single reverse-polarity phosphodiester linkage was inserted into the middle of a 70 nt poly dT substrate, *Ec*SSB continued to bind with high affinity. However, reversing the polarity between every nt prohibited binding in the 65 mode and the cooperativity of the 35 mode was greatly reduced. It was proposed that to accommodate this modified substrate and bind stably, the ssDNA follows a unique path around the *Ec*SSB tetramer. It is possible a similar process occurs due to the presence of abasic sites and non-DNA spacers in our assays. That is, in addition or instead of these damage sites modulating the affinities of the canonical wrapping states, a modified ssDNA-*Ec*SSB complex may be formed that is responsible for the unique binding response we observe.

*Ec*SSB functions *in vivo* by binding variable length segments of ssDNA, such as Okazaki fragments that grow and shrink in length during DNA replication. *Ec*SSB's high local concentration (47) and binding affinity result in complete saturation of exposed ssDNA, with the total number of proteins equal to the substrate length divided by the average binding site size. Thus, while a single protein could bind in many positions along a long substrate, the full protein lattice ensures that all nts along the length are in close proximity to at least one *Ec*SSB. However, as our results show, the exact binding pattern could be altered by the presence of DNA damage. The ssDNA could remain saturated with protein (*i.e.* there is no contiguous length of protein-free ssDNA long enough to accommodate an additional tetramer), but the damaged nt(s) could reside either between neighboring tetramers or between OB folds within a single tetramer, rather than being tightly bound. Further studies are needed to determine if the exact positioning of *Ec*SSB can affect the ability to recruit repair proteins, but it is plausible that certain nts may be more accessible when not tightly held by the *Ec*SSB OB fold.

Conclusions

We have shown that the presence of modified bases that mimic naturally occurring DNA damage can alter the binding conformation of *Ec*SSB without preventing protein saturation. Our results are consistent with *Ec*SSB maximizing direct contact between undamaged DNA and its OB fold domains, leaving sites of DNA damage less tightly bound. While previous studies have shown that *Ec*SSB binding conformation can be modulated *in vitro* by changing conditions such as salt concentration and temperature that do not change dramatically *in vivo*, the accumulation of DNA damage is a plausible mechanism to alter *Ec*SSB binding conformation during normal cellular conditions.

Author Contributions

MM, MCW, and PJB designed the research. MM, FNR, and MK performed experiments. MM and FNR analyzed data. MM and JM produced experimental materials. MM, FNR, and PJB wrote the original manuscript. All authors reviewed and revised the manuscript.

Declaration of Interests

The authors declare no competing interests.

Acknowledgements

This work was supported by National Science Foundation grants MCB-1817712 (MCW) and MCB-1615946 (PJB). FNR and MK were supported by the Research Opportunities for Undergraduates: Training in Environmental Health Sciences (ROUTES) program under grant number R25ES025496 from the National Institute of Environmental Health Sciences.

Figure 1: AFM imaging of *Ec*SSB binding to ssDNA. (A) A ds/ss DNA hybrid is constructed by ligating a restriction enzyme digested dsDNA to a target ssDNA sequence using a linker oligo complementary to both. The end product contains 260 bp of dsDNA and a 67 nt 5' ssDNA overhang. The constructs are incubated with varying concentrations of *Ec*SSB, enabling binding specifically to the ssDNA end. (B) DNA/protein solutions are imaged using AFM. Colocalization of the *Ec*SSB tetramer (white spots) with the end of the dsDNA marker (red lines) indicates bound ssDNA substrates. (B) increased *Ec*SSB concentration results in a greater fraction of substrates bound. (D) The fraction of ssDNA substrates bound as a function of *Ec*SSB concentration is well fit (reduced $\chi^2 \approx 1$) by a simple binding isotherm (Eq. 1, dotted line). Error bars are SEM for N≥3 biological replicates.

Figure 2: *Ec*SSB binding to damaged ssDNA. (A) Binding experiments are repeated for 67 nt ssDNA substrates with the 27^{th} and/or 28^{th} (magenta) bases modified. These C bases (undamaged) are replaced with either an abasic site (AP) or non-DNA spacer (SP). (B) When a single base is replaced by a stable abasic site (blue 27, green 28) both substrates are well fit by Eq. 1 (dotted lines) and exhibit the same binding behavior as the undamaged DNA (black dotted line fit from Fig 1D). (C) Replacing both the 27^{th} and 28^{th} bases with two abasic sites (red) or one triethylene glycol spacer (yellow) results in a sharper transition between mostly unbound and mostly bound substrates. Comparing all five DNA substrates, the damage sites do not alter the apparent binding affinity K_D to a significant degree (D), but the substrates with two sites modified are no longer fit by Eq. 1, as measured by reduced χ^2 (E).

Figure 3: Salt-dependent binding modes of *Ec*SSB. (A) The integrated volume of a DNA (red) or proteins (blue) as measured by AFM is directly proportional to the known molecular weight of the constructs, if the imaging tip and solution conditions are conserved. When high concentrations of *Ec*SSB (10 nM) are incubated with the undamaged ssDNA at low (B) or high (C) salt concentrations, the size of the bound protein cluster on the substrate varies, with larger volumes at low salt. (D) The average sizes of the protein clusters under both conditions are converted to a measured molecular weight (blue bars), using a conversion factor determined by the apparent size of the 500 bp DNA construct measured under the same conditions (black dashed line, panel A). Comparing these values to the molecular weight of the ssDNA substrate (20.7 kDa) with either one (red) or two (green) *Ec*SSB tetramers bound (75.5 kDa each) confirms that one *Ec*SSB (65 nt mode) binds the ssDNA at high salt and two *Ec*SSB (35 mode) bind the ssDNA at low salt.

Figure 4: Bound *Ec*SSB stoichiometry. (A) For each protein-bound ssDNA substrate imaged, the number of tetramers present is calculated using volumetric methods. The average binding stoichiometry is plotted as a function free *Ec*SSB concentration, showing that ssDNA is typically bound by one tetramer at low concentrations and two tetramers at high concentrations. A sigmoidal function (dashed line) is plotted as a guide to the eye. (B) Compared to undamaged DNA (same dashed line as in A), the tandem damage site constructs show an increase in ssDNA bound by two tetramers even at low protein concentration.

Figure 5: Multistate binding measurements and model. (A) Cartoon showing the four possible states of the ssDNA substrate and potential transitions. (B) The average number of *Ec*SSB tetramers bound per ssDNA substrate is calculated by multiplying the fraction of substrates bound (Fig 1D) by the average binding stoichiometry (Fig 4A). The data are fit using Eq. 3 (15), with the best fit parameters displayed. (C) The average number of *Ec*SSB tetramers per damaged ssDNA substrate shows a sharper transition, which is not fully captured by Eq. 3 (reduced $\chi^2 > 1$) even with K_{65} reduced to 0, particularly in the region where equimolar concentrations of ssDNA and *Ec*SSB tetramers are present (~1 nM). (D) Structure of *Ec*SSB tetramer (12), with individual subunits in yellow and green, bound to two 35-mer oligos (cyan), which was used to model *Ec*SSB wrapping modes (12). The 27^{th} and 28^{th} nt (red) stack between residues Trp40 and Phe60 of two adjacent subunits when ssDNA binds in the 65 nt mode. (E) Cartoon showing the *Ec*SSB 65 nt binding mode, in which the ssDNA substrate fully occupies all four OB folds on one tetramer (left), and the 35 nt binding mode, where the substrate fully occupies one OB fold and partially occupies two OB folds on two tetramers (right). The 27 and 28 nt of the substrate strongly intact with protein in the 65 nt conformation but are located between strong interaction sites in the 35 nt conformation.

Oligo	Sequence (5’-3’)
PCR Primer 1	CAGGTCGACTCTAGAGGATCCC
PCR Primer 2	ACTGAGAGTGCACCATATGCG
Linker	GATCGGGAAGGG
ssDNA Substrate	CGTTACTCAGATCAGGCCTGCGAAGA*XY*TGGGCGTCCGGCTGCAGCTGTACTATC ATATGCCTATATCCCTTCCC

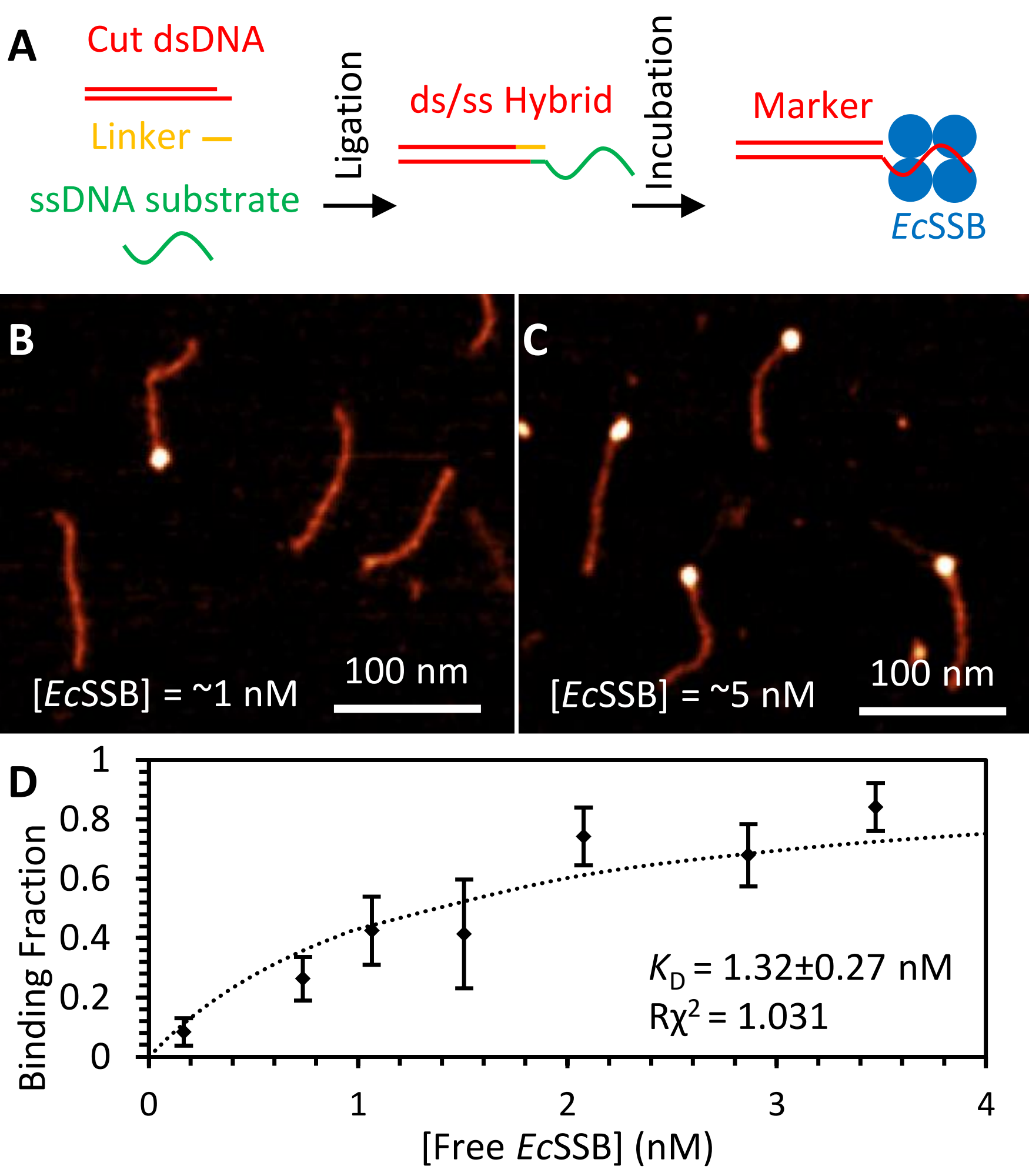
Table 1: DNA oligos used for construction of DNA substrates. PCR primers generate a 268 bp product from pUC19. The linker oligo is complementary to both the 5’ overhang generated by BamHI digestion and the 3’ end of the ssDNA substrates (underlined). For ssDNA substrates, the 27th (X) and 28th (Y) bases (italic) are both C (undamaged), one is replaced with an apurinic/apyrimidinic (AP) site (AP27 and AP28), both are replaced by an AP site (AP27AP28), or both are replaced with a triethylene glycol spacer (SP2728).

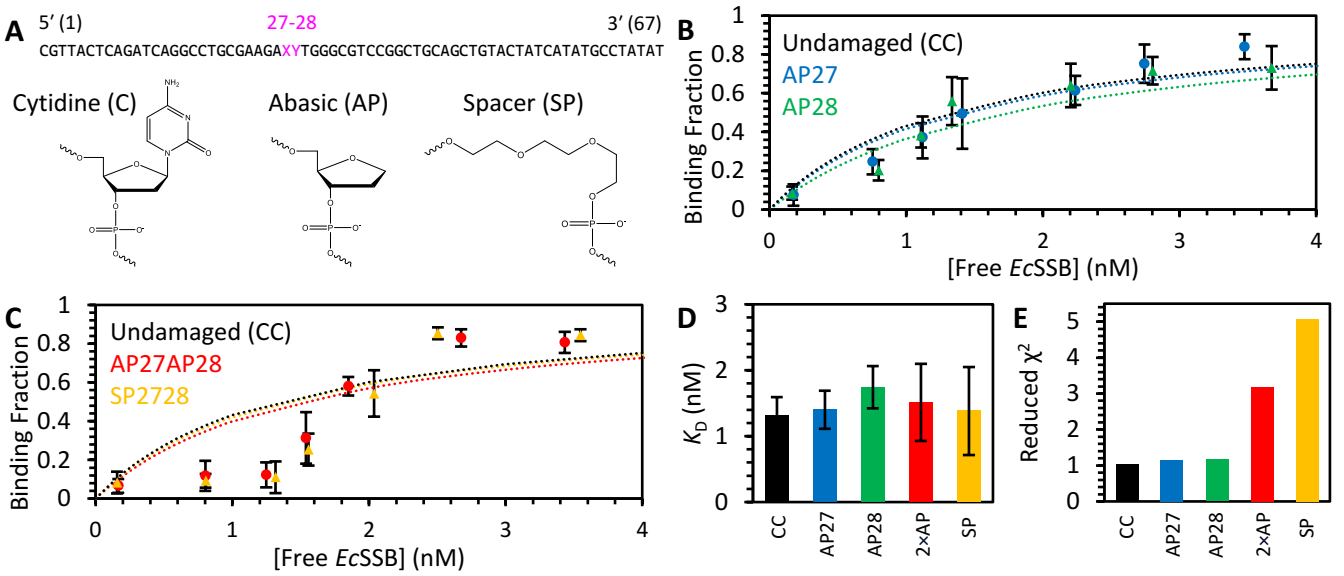
References

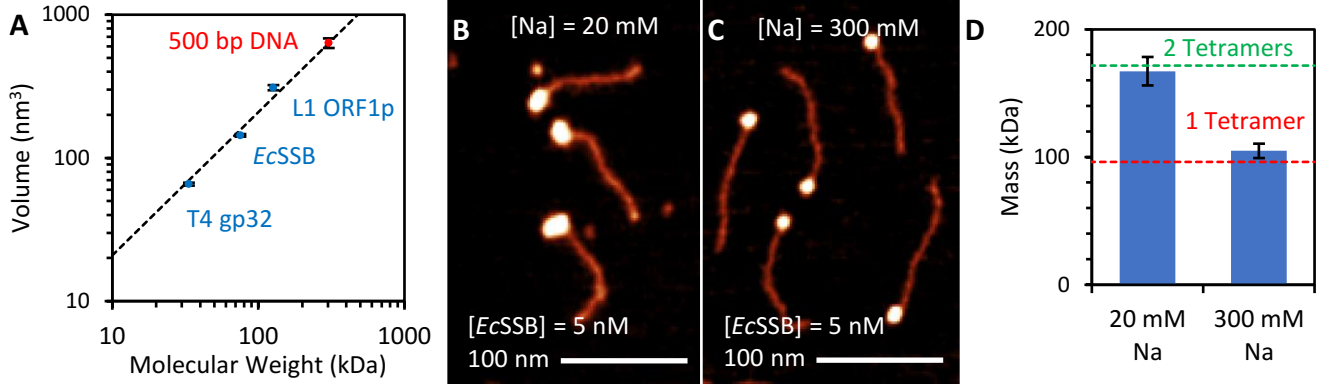
1. Sigal, N., H. Delius, T. Kornberg, M. L. Gefter, and B. Alberts. 1972. A DNA-unwinding protein isolated from *Escherichia coli*: its interaction with DNA and with DNA polymerases. *Proc. Natl. Acad. Sci. USA*. 69(12):3537-3541.
2. Molineux, I. J., and M. L. Gefter. 1975. Properties of the *Escherichia coli* DNA-binding (unwinding) protein interaction with nucleolytic enzymes and DNA. *J. Mol. Biol.* 98(4):811-825.
3. Weiner, J. H., L. L. Bertsch, and A. Kornberg. 1975. The deoxyribonucleic acid unwinding protein of *Escherichia coli*. Properties and functions in replication. *J. Biol. Chem.* 250(6):1972-1980.
4. Chrysogelos, S., and J. Griffith. 1982. *Escherichia coli* single-strand binding protein organizes single-stranded DNA in nucleosome-like units. *Proc. Natl. Acad. Sci. USA*. 79(19):5803-5807.
5. Kuzminov, A. 1999. Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage λ. *Microbiol. Mol. Biol. Rev.* 63(4):751-813.
6. Shereda, R. D., A. G. Kozlov, T. M. Lohman, M. M. Cox, and J. L. Keck. 2008. SSB as an organizer/mobilizer of genome maintenance complexes. *Crit. Rev. Biochem. Mol. Biol.* 43(5):289-318.
7. Shereda, R. D., D. A. Bernstein, and J. L. Keck. 2007. A central role for SSB in *Escherichia coli* RecQ DNA helicase function. *J. Biol. Chem.* 282(26):19247-19258.
8. Han, E. S., D. L. Cooper, N. S. Persky, V. A. Sutera Jr, R. D. Whitaker, M. L. Montello, and S. T. Lovett. 2006. RecJ exonuclease: substrates, products and interaction with SSB. *Nucleic Acids Res.* 34(4):1084-1091.
9. Furukohri, A., Y. Nishikawa, M. Tatsumi Akiyama, and H. Maki. 2012. Interaction between *Escherichia coli* DNA polymerase IV and single-stranded DNA-binding protein is required for DNA synthesis on SSB-coated DNA. *Nucleic Acids Res.* 40(13):6039-6048.
10. Greipel, J., G. Maass, and F. Mayer. 1987. Complexes of the single-stranded DNA-binding protein from *Escherichia coli* (Eco SSB) with poly (dT): an investigation of their structure and internal dynamics by means of electron microscopy and NMR. *Biophys. Chem.* 26(2-3):149-161.
11. Savvides, S. N., S. Raghunathan, K. Fütterer, A. G. Kozlov, T. M. Lohman, and G. Waksman. 2004. The C-terminal domain of full-length *E. coli* SSB is disordered even when bound to DNA. *Protein Sci.* 13(7):1942-1947.
12. Raghunathan, S., A. G. Kozlov, T. M. Lohman, and G. Waksman. 2000. Structure of the DNA binding domain of *E. coli* SSB bound to ssDNA. *Nat. Struct. Biol.* 7(8):648-652, doi: 10.1038/77943.
13. Raghunathan, S., C. S. Ricard, T. M. Lohman, and G. Waksman. 1997. 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*. 94(13):6652-6657.
14. Lohman, T. M., L. B. Overman, and S. Datta. 1986. Salt-dependent changes in the DNA binding co-operativity of *Escherichia coli* single strand binding protein. *J. Mol. Biol.* 187(4):603-615.
15. Ferrari, M. E., W. Bujalowski, and T. M. Lohman. 1994. Co-operative binding of *Escherichia coli* SSB tetramers to single-stranded DNA in the (SSB) 35 binding mode. *J. Mol. Biol.* 236(1):106-123.
16. Naufer, M. N., M. Morse, G. B. Möller, J. McIsaac, I. Rouzina, P. J. Beuning, and M. C. Williams. 2021. Multiprotein *E. coli* SSB–ssDNA complex shows both stable binding and rapid dissociation due to interprotein interactions. *Nucleic Acids Res.* 49(3):1532-1549.
17. Suksombat, S., R. Khafizov, A. G. Kozlov, T. M. Lohman, and Y. R. Chemla. 2015. Structural dynamics of *E. coli* single-stranded DNA binding protein reveal DNA wrapping and unwrapping pathways. *elife*. 4:e08193.

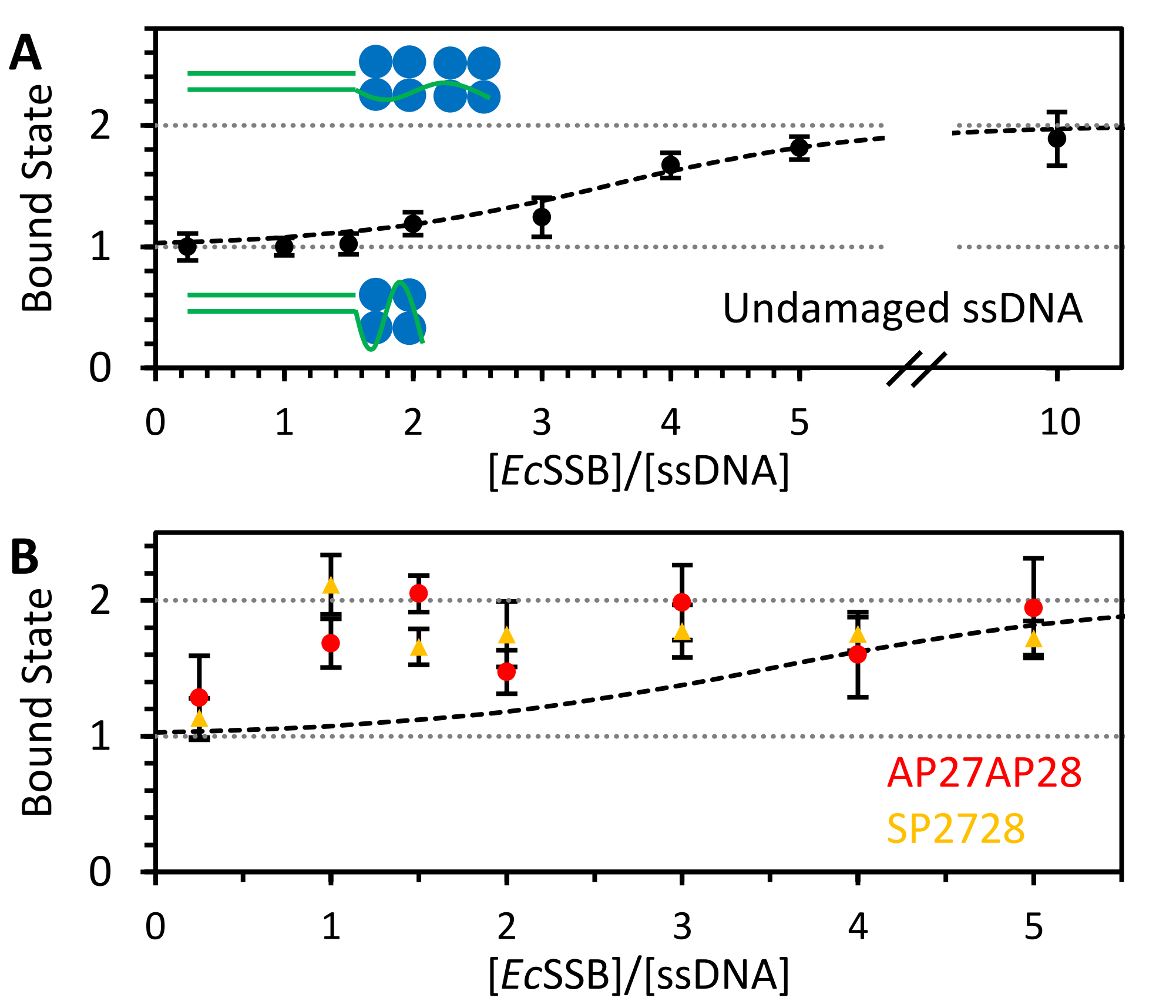
18. Antony, E., E. Weiland, Q. Yuan, C. M. Manhart, B. Nguyen, A. G. Kozlov, C. S. McHenry, and T. M. Lohman. 2013. Multiple C-terminal tails within a single *E. coli* SSB homotetramer coordinate DNA replication and repair. *J. Mol. Biol.* 425(23):4802-4819.
19. Shinn, M. K., A. G. Kozlov, and T. M. Lohman. 2021. Allosteric effects of SSB C-terminal tail on assembly of *E. coli* RecOR proteins. *Nucleic Acids Res.* 49(4):1987-2004.
20. Kozlov, A. G., M. M. Cox, and T. M. Lohman. 2010. Regulation of single-stranded DNA binding by the C termini of *Escherichia coli* single-stranded DNA-binding (SSB) protein. *J. Biol. Chem.* 285(22):17246-17252.
21. Kozlov, A. G., E. Weiland, A. Mittal, V. Waldman, E. Antony, N. Fazio, R. V. Pappu, and T. M. Lohman. 2015. Intrinsically disordered C-terminal tails of *E. coli* single-stranded DNA binding protein regulate cooperative binding to single-stranded DNA. *J. Mol. Biol.* 427(4):763-774.
22. Meyer, R. R., and P. S. Laine. 1990. The single-stranded DNA-binding protein of *Escherichia coli*. *Microbiol. Rev.* 54(4):342-380.
23. Zhao, T., Y. Liu, Z. Wang, R. He, J. Xiang Zhang, F. Xu, M. Lei, M. B. Deci, J. Nguyen, and P. R. Bianco. 2019. Super-resolution imaging reveals changes in *Escherichia coli* SSB localization in response to DNA damage. *Genes Cells*. 24(12):814-826.
24. Friedberg, E. C., G. C. Walker, W. Siede, R. D. Wood, R. A. Schultz, and T. Ellenberger. 2005. DNA repair and mutagenesis. American Society for Microbiology Press.
25. Janion, C., A. Sikora, A. Nowosielska, and E. Grzesiuk. 2003. *E. coli* BW535, a triple mutant for the DNA repair genes *xth*, *nth*, and *nfo*, chronically induces the SOS response. *Environ. Mol. Mutag.* 41(4):237-242.
26. Kunkel, T. A. 1984. Mutational specificity of depurination. *Proc. Natl. Acad. Sci. USA*. 81(5):1494-1498.
27. Loeb, L. A., and B. D. Preston. 1986. Mutagenesis by apurinic/apyrimidinic sites. *Annu. Rev. Genet.* 20(1):201-230.
28. Schaaper, R. M., and L. A. Loeb. 1981. Depurination causes mutations in SOS-induced cells. *Proc. Natl. Acad. Sci. USA*. 78(3):1773-1777.
29. Umezu, K., N.-W. Chi, and R. D. Kolodner. 1993. Biochemical interaction of the *Escherichia coli* RecF, RecO, and RecR proteins with RecA protein and single-stranded DNA binding protein. *Proc. Natl. Acad. Sci. USA*. 90(9):3875-3879.
30. Chen, S. H., R. T. Byrne-Nash, and M. M. Cox. 2016. *Escherichia coli* RadD protein functionally interacts with the single-stranded DNA-binding protein. *J. Biol. Chem.* 291(39):20779-20786.
31. Sandigursky, M., F. Mendez, R. E. Bases, T. Matsumoto, and W. A. Franklin. 1996. Protein-protein interactions between the *Escherichia coli* single-stranded DNA-binding protein and exonuclease I. *Radiat. Res.* 145(5):619-623.
32. Whittier, R. F., and J. W. Chase. 1981. DNA repair in *E. coli* strains deficient in single-strand DNA binding protein. *Mol. Gen. Genet.* 183(2):341-347.
33. Lindahl, T., and B. Nyberg. 1972. Rate of depurination of native deoxyribonucleic acid. *Biochemistry*. 11(19):3610-3618.
34. Lindahl, T. 1993. Instability and decay of the primary structure of DNA. *Nature*. 362(6422):709-715.
35. Lohman, T. M., J. M. Green, and R. S. Beyer. 1986. Large-scale overproduction and rapid purification of the *Escherichia coli ssb* gene product. Expression of the ssb gene under. lambda. PL control. *Biochemistry*. 25(1):21-25.
36. Shlyakhtenko, L. S., A. Y. Lushnikov, M. Li, L. Lackey, R. S. Harris, and Y. L. Lyubchenko. 2011. Atomic force microscopy studies provide direct evidence for dimerization of the HIV restriction factor APOBEC3G. *J. Biol. Chem.* 286(5):3387-3395.

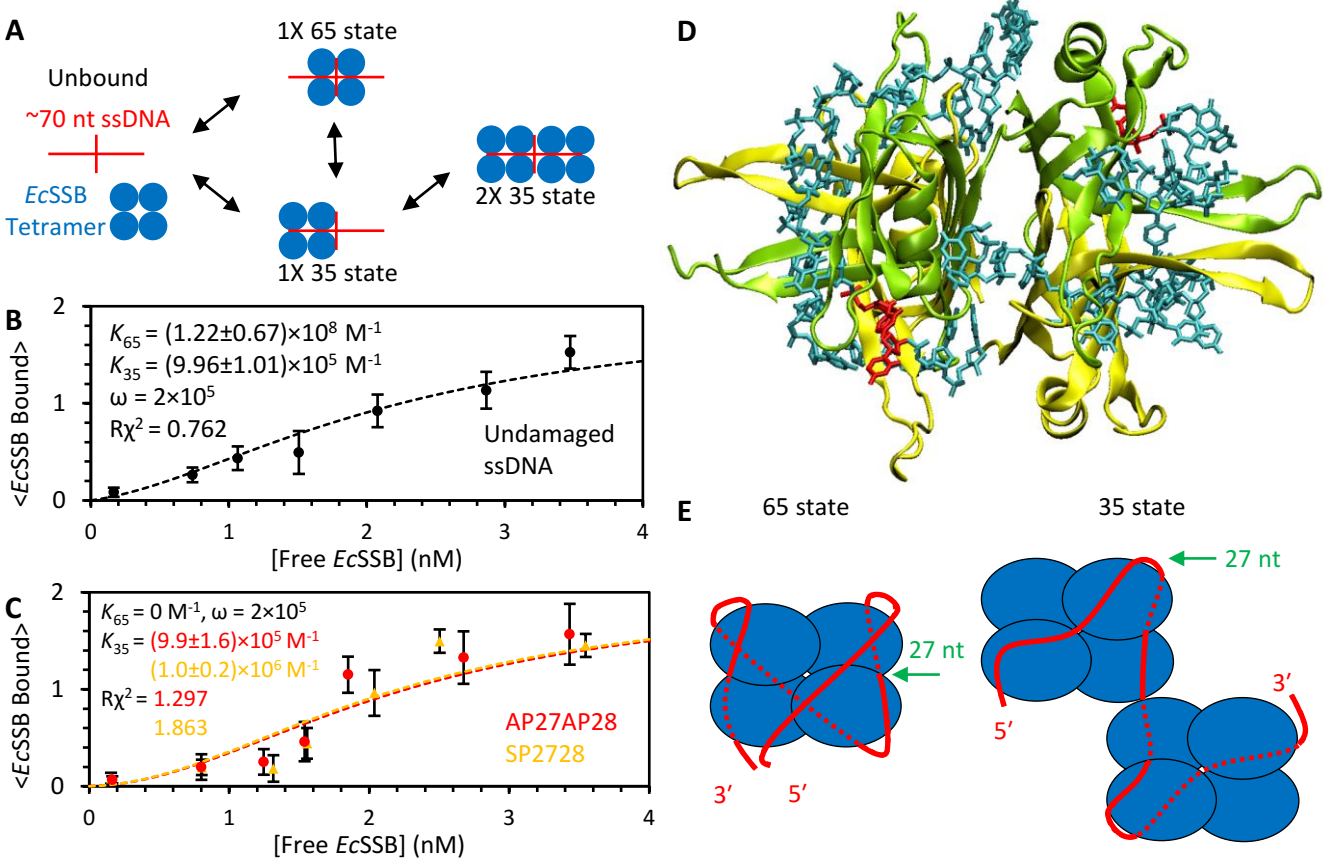
37. Antczak, N. M., A. R. Walker, H. R. Stern, E. M. Leddin, C. Palad, T. A. Coulther, R. J. Swett, G. A. Cisneros, and P. J. Beuning. 2018. Characterization of nine cancer-associated variants in human DNA polymerase κ. *Chem. Res. Toxicol.* 31(8):697-711.
38. Bujalowski, W., and T. M. Lohman. 1986. *Escherichia coli* single-strand binding protein forms multiple, distinct complexes with single-stranded DNA. *Biochemistry*. 25(24):7799-7802.
39. Lohman, T. M., and L. B. Overman. 1985. Two binding modes in *Escherichia coli* single strand binding protein-single stranded DNA complexes. Modulation by NaCl concentration. *J. Biol. Chem.* 260(6):3594-3603.
40. Zhou, R., A. G. Kozlov, R. Roy, J. Zhang, S. Korolev, T. M. Lohman, and T. Ha. 2011. SSB functions as a sliding platform that migrates on DNA via reptation. *Cell*. 146(2):222-232.
41. Lee, K. S., A. B. Marciel, A. G. Kozlov, C. M. Schroeder, T. M. Lohman, and T. Ha. 2014. Ultrafast redistribution of *E. coli* SSB along long single-stranded DNA via intersegment transfer. *J. Mol. Biol.* 426(13):2413-2421.
42. Fuentes-Perez, M. E., M. S. Dillingham, and F. Moreno-Herrero. 2013. AFM volumetric methods for the characterization of proteins and nucleic acids. *Methods*. 60(2):113-121.
43. Cashen, B. A., M. N. Naufer, M. Morse, C. E. Jones, M. C. Williams, and A. V. Furano. 2022. The L1-ORF1p coiled coil enables formation of a tightly compacted nucleic acid-bound complex that is associated with retrotransposition. *Nucleic Acids Res.* 50(15):8690-8699.
44. Cashen, B. A., M. Morse, I. Rouzina, Richard L. Karpel, and Mark C. Williams. 2023. Dynamic structure of T4 gene 32 protein filaments facilitates rapid noncooperative protein dissociation. *Nucleic Acids Res.*
45. Roy, R., A. G. Kozlov, T. M. Lohman, and T. Ha. 2007. Dynamic structural rearrangements between DNA binding modes of *E. coli* SSB protein. *J. Mol. Biol.* 369(5):1244-1257.
46. Kozlov, A. G., and T. M. Lohman. 2021. Probing *E. coli* SSB protein-DNA topology by reversing DNA backbone polarity. *Biophys. J.* 120(8):1522-1533.
47. Schmidt, A., K. Kochanowski, S. Vedelaar, E. Ahrné, B. Volkmer, L. Callipo, K. Knoops, M. Bauer, R. Aebersold, and M. Heinemann. 2016. The quantitative and condition-dependent *Escherichia coli* proteome. *Nat. Biotechnol.* 34(1):104-110.


A
Cut dsDNA
Linker
ssDNA substrate
Ligation
ds/ss Hybrid
Incubation
Marker
EcSSB
B
[EcSSB] = ~1 nM
100 nm
C
[EcSSB] = ~5 nM
100 nm
D
Binding Fraction
0
0.2
0.4
0.6
0.8
1
0
1
2
3
4
[Free EcSSB] (nM)
K_D = 1.32±0.27 nM
Rχ^2 = 1.031


A
5′ (1)
27-28
3′ (67)
CGTTACTCAGATCAGGCCTGCGAAGAXYTGGGCGTCCGGCTGCAGCTGTACTATCATATGCCTATAT
Cytidine (C)
Abasic (AP)
Spacer (SP)
B
Undamaged (CC)
AP27
AP28
Binding Fraction
[Free EcSSB] (nM)
C
Undamaged (CC)
AP27AP28
SP2728
Binding Fraction
[Free EcSSB] (nM)
D
K_D (nM)
CC
AP27
AP28
2×AP
SP
E
Reduced χ²
CC
AP27
AP28
2×AP
SP


A
1000
100
10
Volume (nm3)
500 bp DNA
L1 ORF1p
EcSSB
T4 gp32
10
100
1000
Molecular Weight (kDa)
B
[Na] = 20 mM
[EcSSB] = 5 nM
100 nm
C
[Na] = 300 mM
[EcSSB] = 5 nM
100 nm
D
200
100
0
Mass (kDa)
2 Tetramers
1 Tetramer
20 mM Na
300 mM Na


A
Bound State
0
1
2
Undamaged ssDNA
0
1
2
3
4
5
10
[EcSSB]/[ssDNA]
B
Bound State
0
1
2
AP27AP28
SP2728
0
1
2
3
4
5
[EcSSB]/[ssDNA]


A
1X 65 state
Unbound
~70 nt ssDNA
EcSSB
Tetramer
2X 35 state
1X 35 state
B
K65 = (1.22±0.67)×10^8 M^-1
K35 = (9.96±1.01)×10^5 M^-1
ω = 2×10^5
Rχ^2 = 0.762
Undamaged
ssDNA
<EcSSB Bound>
[Free EcSSB] (nM)
C
K65 = 0 M^-1, ω = 2×10^5
K35 = (9.9±1.6)×10^5 M^-1
(1.0±0.2)×10^6 M^-1
Rχ^2 = 1.297
1.863
AP27AP28
SP2728
<EcSSB Bound>
[Free EcSSB] (nM)
D
E
65 state
35 state
27 nt
27 nt
3′
5′
3′
5′