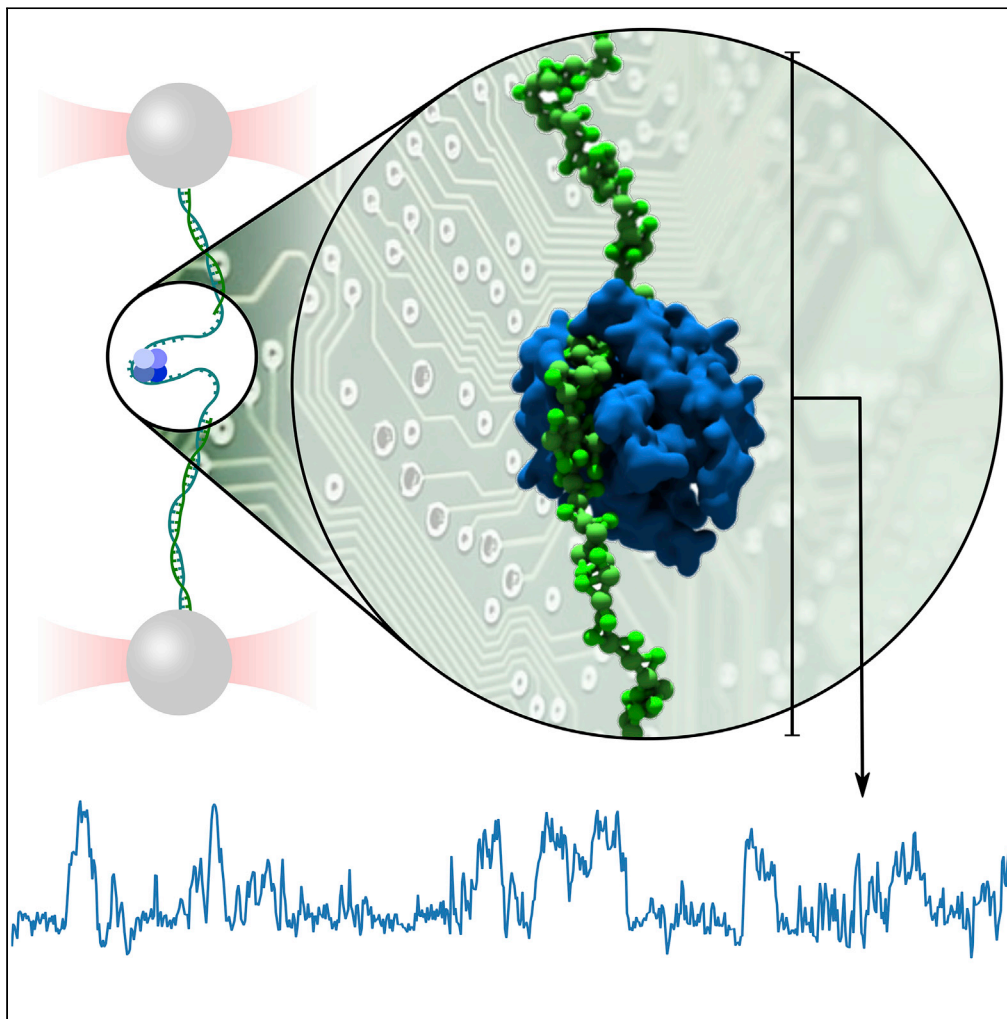


Article

Single-molecule biophysics experiments *in silico*:
Toward a physical model of a replisome

Christopher
Maffeo, Han-Yi
Chou, Aleksei
Aksimentiev

aksiment@illinois.edu

Highlights

Coarse-grained model
derived from all-atom
simulation recapitulates
experiments

Model reproduces the
elastic response to force
and exchange dynamics

Model reveals structure of
intermediate states
usually inaccessible to
experiment

Model applied to viral
replisome with trombone
loop containing tens of
SSB proteins

Maffeo et al., iScience 25,
104264
May 20, 2022 © 2022 The
Author(s).
[https://doi.org/10.1016/
j.isci.2022.104264](https://doi.org/10.1016/j.isci.2022.104264)

Article

Single-molecule biophysics experiments
in silico: Toward a physical model of a replisomeChristopher Maffeo,^{1,2} Han-Yi Chou,¹ and Aleksei Aksimentiev^{1,2,3,*}

SUMMARY

The interpretation of single-molecule experiments is frequently aided by computational modeling of biomolecular dynamics. The growth of computing power and ongoing validation of computational models suggest that it soon may be possible to replace some experiments outright with computational mimics. Here, we offer a blueprint for performing single-molecule studies *in silico* using a DNA-binding protein as a test bed. We demonstrate how atomistic simulations, typically limited to sub-millisecond durations and zeptoliter volumes, can guide development of a coarse-grained model for use in simulations that mimic single-molecule experiments. We apply the model to recapitulate, *in silico*, force-extension characterization of protein binding to single-stranded DNA and protein and DNA replacement assays, providing a detailed portrait of the underlying mechanics. Finally, we use the model to simulate the trombone loop of a replication fork, a large complex of proteins and DNA.

INTRODUCTION

Single-molecule experiments can characterize physical and biochemical properties of individual biomolecules and biomolecular assemblies to provide key insights into the mechanisms that underlie biological function (Tinoco and Bustamante, 2002; Dudko et al., 2008; Neuman and Nagy, 2008; Zhou et al., 2010; Robertson et al., 2007; Kim and Ha, 2013). However, most experimental techniques that operate on individual biomolecules, including force spectroscopy (Dudko et al., 2008; Neuman and Nagy, 2008) and fluorescent microscopy (Roy et al., 2008; Stein et al., 2011), can simultaneously probe a very limited number of degrees of freedom. Such limitations can be problematic when it comes to interpreting the results of single-molecule measurements because the functional behavior of a biomolecule may depend on the state of one or more “hidden” degrees of freedom, such as the conformation of a protein (Meinhold and Wright, 2011; Stein et al., 2011). Physics-based computational modeling of single-molecule experiments can greatly aid functional interpretation of the measurements by resolving degrees of freedom that are neither controlled nor observed in the experiment.

Among the most popular computational approaches, explicit solvent all-atom MD simulations have been applied widely to study the molecular mechanisms underlying biological and biotechnological systems (Karplus and McCammon, 2002; Lindahl and Sansom, 2008; Klepeis et al., 2009; Hollingsworth and Dror, 2018; Maffeo et al., 2019). However, the explicit representation of each atom limits routine application of the method to small systems and/or relatively short timescales. Modeling larger systems for longer timescales requires dramatic reduction of the degrees of freedom, or coarse-graining, of a system. In contrast to all-atom simulations, where there are relatively few models available that are all general in nature and have each been developed using similar approaches over decades, coarse-grained (CG) models are extremely diverse in their parametrization and domain of application.

The parametrization of a CG model can, in general, be classified as either top-down or bottom-up. Many CG models have been parameterized using a top-down approach, where the interaction sites and the energy potentials associated with them are chosen using a mixture of intuition, physical law, and trial-and-error until the model produces desired, usually experimentally measured, properties for some test systems. Examples of such models include: the general-purpose Martini model (De Jong et al., 2013); the oxDNA (Ouldrige et al., 2011; Šulc et al., 2012) and 3SPN2 (Hinckley et al., 2013) models for DNA; the AWSEM structured protein model (Davtyan et al., 2012); the various models for disordered peptides from the Hummer (Kim and Hummer, 2008), Onck (Ghavami et al., 2016), Best (Zheng et al., 2016), Papoian (Wu et al.,

¹Department of Physics, University of Illinois at Urbana-Champaign, 1110 W Green St, Urbana, 61801 IL, USA

²Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, 405 N Matthews Avenue, Urbana, 61801 IL, USA

³Lead contact

*Correspondence: aksiment@illinois.edu

<https://doi.org/10.1016/j.isci.2022.104264>



2018), and Mittal (Shea et al., 2021) labs; rigid body and elastic network models of proteins used in virus capsid assembly studies (Nguyen and Brooks, 2008; Laro et al., 2018; Grime et al., 2016); and rigid body protein models of the bacterial cytoplasm from the Elcock lab (McGuffee and Elcock, 2010) and of chromatin from the Schlick lab (Arya and Schlick, 2006), among others. The Levy lab developed a top-down model with emphasis on electrostatic and aromatic ring interactions to predict the binding geometry of a DNA strand to single-stranded DNA-binding protein (SSB) proteins of several species, showing good agreement with available crystallographic structures (Mishra and Levy, 2015; Pal and Levy, 2019).

At the opposite end of the spectrum for model development, bottom-up coarse-graining starts with simulating a detailed model of the system, typically all-atom, to gather statistics. After selection of CG interaction sites, one of several approaches is used to optimize the CG interaction potentials until agreement of forces, configuration, or general thermodynamic observables is obtained between the CG model and the detailed model. CG potentials are commonly split into bonded and non-bonded parts with non-bonded potentials usually chosen to be pairwise. Collective variable-based potentials, for example those dependent on density, are sometimes used (Louis, 2002; Wagner et al., 2017). Pairwise force-matching optimization strategies include the simplex algorithm (Meyer et al., 2000), iterative Boltzmann inversion (Reith et al., 2003), inverse Monte Carlo (Lyubartsev and Laaksonen, 1995), and machine learning (Gkeka et al., 2020), among others (Sun et al., 2021).

Compared to top-down, a bottom-up model will precisely capture the effective interactions between CG sites, and arguably it may be less biased by the expectations of the model developer. On the other hand, a bottom-up model may be criticized for inheriting bias from the underlying detailed model and for lacking predictive power because the model primarily reproduces what is already known from simulation of the underlying detailed model. Nevertheless, there are cases where bottom-up coarse-graining has been successfully applied to investigate systems at length scales and timescales much larger and longer than those that could be studied with an underlying all-atom model (McGuffee and Elcock, 2010; Grime et al., 2016; Torre and Takada, 2020; Takada et al., 2015). For example, our own CG model of ssDNA (Maffeo et al., 2014) was parametrized using 10 μ s of all-atom MD simulation of a 60-nucleotide (nt) DNA strand performed on a supercomputer, yet it enabled workstation-powered studies of a 1000-nt DNA strand transport through a dual nanopore systems (Choudhary et al., 2020) for durations of hundreds of microseconds, simulations presently impossible to carry out using an explicit-solvent atomistic description. However, it should be noted that bottom-up models will inherit any deficiencies of the underlying higher resolution model. Thus, in the case of our ssDNA model, introducing weak non-bonded repulsion was needed to swell the DNA strand to a size consistent with small-angle x-ray scattering experiments (Sim et al., 2012; Maffeo et al., 2014).

Here, we demonstrate the extension of our CG ssDNA model to include SSB, a globular protein that plays a key role during DNA replication and repair processes by preventing secondary structure in the nascently unwound ssDNA and by recruiting repair proteins to sites of DNA damage. We model the protein as a rigid body, reducing the degrees of freedom for each SSB molecule to six. The interactions between CG ssDNA beads and the SSB were obtained using the iterative Boltzmann inversion (IBI) method to match structural distributions from all-atom simulations and assuming the protein–DNA interactions to be short-ranged. The resulting model is shown to be in quantitative agreement with the experimentally measured elasticity of an ssDNA–SSB complex. The model was then used to study competition between ssDNA fragments and SSB molecules in systems that mimic experimental assays. Finally, we introduced SSB–SSB interactions using a top-down approach and used the model to investigate the possible structure of a trombone loop within the DNA replication fork.

RESULTS

Seeking access to timescales that would enable *in silico* modeling of constructs used in experiments to probe SSB–ssDNA and SSB–SSB interactions, we set out to develop a CG model of SSB using all-atom MD simulations to guide the process. The construction of the model is described below.

Beads-and-grids models of ssDNA–SSB assembly

To model an ssDNA–SSB complex, we begin by selecting an existing CG model of ssDNA that reproduces the experimentally measured elastic properties of DNA, namely a model developed several years ago by our lab (Maffeo et al., 2014). In that model, each nucleotide is represented by two beads. The interaction

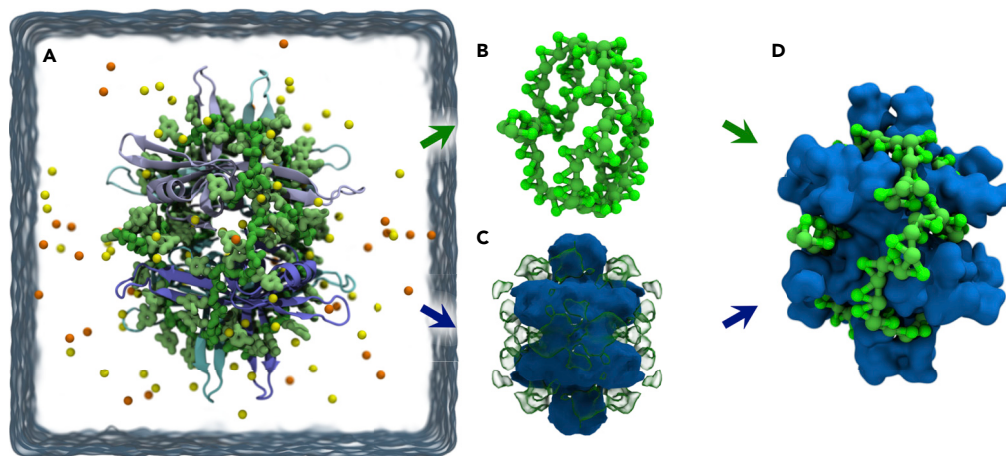


Figure 1. Beads-and-grids model of ssDNA–SSB assembly

(A) All-atom model of SSB (blue and cyan) bound to ssDNA (green) in a 100 mM electrolyte solution (semi-transparent surface) containing cations (Na^+ ; yellow) and anions (Cl^- ; orange).

(B) Beads-on-a-string representation of the ssDNA molecule depicted in panel A. Two beads, P (phosphate; green) and B (base; light green), represent each nucleotide. The beads interact through a set of tabulated potentials that were optimized to reproduce distributions extracted from atomistic simulations and the experimentally inferred radius of gyration of unstructured ssDNA. Solvent is represented implicitly through the non-bonded bead-bead interaction potentials. See (Maffeo et al., 2014) for additional details.

(C) Grids-based model of an SSB particle. The physical SSB particle features a mass and a moment of inertia and can translate and rotate in response to external forces and torques. For each type of the CG beads used in the ssDNA model, a three-dimensional potential grid is tethered to the SSB particle to prescribe the interaction between the protein and that bead type. A cross-section of the potential for P beads is depicted. At a given timestep, for each force applied to an ssDNA bead, the opposite force and corresponding torque are applied to the SSB particle.

(D) Beads-and-grids representation of the SSB–ssDNA complex.

potentials were parametrized iteratively to match target structural distributions obtained from all-atom simulations, including bond and angle distributions and pair distance distributions. A small correction to the resulting non-bonded potentials was applied to swell the CG ssDNA so that its radius of gyration matched that found in experiment, presumably correcting a systematic bias in the underlying all-atom force field. The final version of the model quantitatively reproduced the elastic response of a DNA molecule to applied tension (Smith et al., 1996; Saleh et al., 2009), although it was not specifically parameterized to do so.

The successful application of the iterative coarse-graining procedure for constructing our ssDNA model led us to apply the same approach to SSB. First, we chose to represent the protein as a rigid body without internal degrees of freedom because conformational fluctuations of the core SSB protein were seen to be small in all-atom simulations (Maffeo and Aksimentiev, 2017). Further, all high-resolution SSB tetramer structures published to date (Raghunathan et al., 1997, 2000; Matsumoto et al., 2000; Savvides et al., 2004) adopt similar conformations, suggesting that the conformation of SSB is relatively static. Nevertheless, we note that the rigid body approximation eliminates any possible conformational transitions within the protein. To support the iterative Boltzmann inversion procedure, we decided to use tabulated three-dimensional grid-based potentials to describe the interactions between each CG DNA bead type and the protein. Figure 1 provides a pictorial representation of the all-atom and CG model of an ssDNA–SSB complex. Note that, in building our model of SSB, we have neglected its intrinsically disordered C-terminal tails which play a role in recruitment of auxiliary replication and repair proteins and also slightly modulate SSB binding to ssDNA and its cooperativity (Antony et al., 2013; Su et al., 2014; Kozlov et al., 2010, 2015, 2017). While adding explicit representation of the flexible tails to our model is straightforward from a technical standpoint, the parameterization of the model, in particular, of the tails' interactions with auxiliary proteins and with other tails, is challenging because of the lack of the required experimental data. Because the acidic fragments of C-terminal tail tips were unable to compete effectively with same-length ssDNA fragments for contact with the SSB surface in our prior all-atom MD simulation study (Maffeo and Aksimentiev, 2017), we believe their inclusion would negligibly influence the DNA-binding behavior of the model.

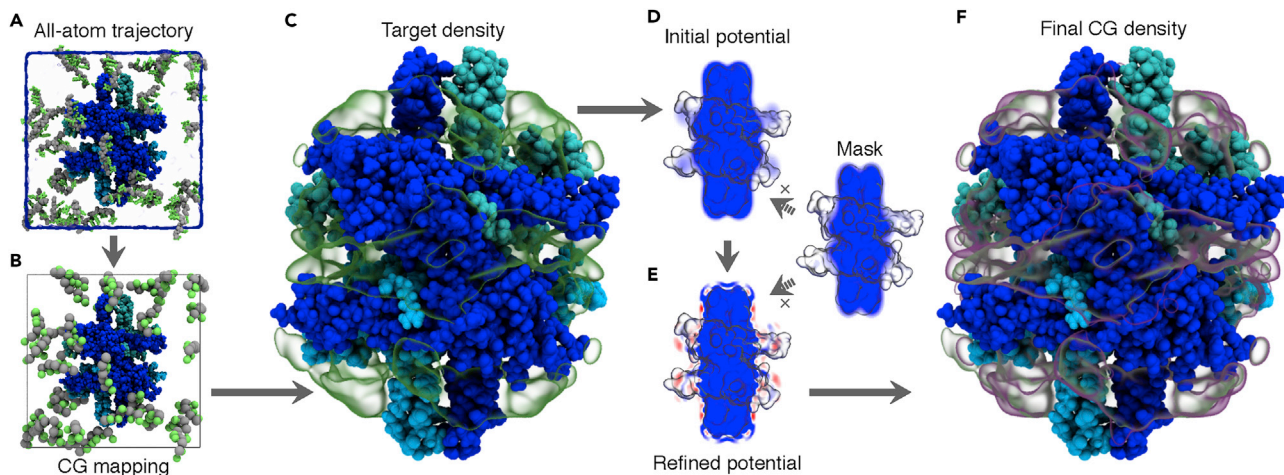


Figure 2. Parametrization of the CG model of SSB

(A) Generation of initial target data for parametrization through all-atom MD simulations of SSB surrounded by 31 dT₅ ssDNA fragments. The SSB protein was restrained to its initial coordinates. See also [Video S1](#).

(B) Atomic coordinates of the ssDNA fragments from the MD trajectories were mapped into a CG representation.

(C) Target density of each CG ssDNA bead type was extracted from the mapped trajectory.

(D–F) Iterative Boltzmann inversion of the target density. Boltzmann inversion provided an initial estimate for the CG SSB–ssDNA interaction potential (D). Each CG potential was refined (E) until they produced a CG density (F) that matched the target density, see also [Video S2](#), [Figures S1](#) and [S2](#). Two-dimensional slices of the 3D potentials and densities shown are for the P ssDNA beads, and the SSB protein surface is depicted as a semi-transparent blue molecular surface. A mask, depicted between panels D and E, was generated from the blurred density of SSB atoms and was used to restrict the IBI potential to vicinity of the protein surface. See also [Figure S3](#).

Refinement of the model against all-atom simulation data

To generate target distributions required for the parametrization of our CG model, we probed the interactions between SSB and short fragments of ssDNA using the all-atom MD method. A typical simulation system contained 31 fragments of dT₅ distributed around an SSB molecule in a cubic volume of 100 mM NaCl electrolyte, [Figure 2A](#). After equilibration, the average concentration of ions in regions 4 nm away from SSB atoms was 106 and 285 mM for Cl[−] and Na⁺, respectively, with the difference being caused by the negatively charged DNA nucleotides not bound to the surface of the SSB. Five replicate systems were constructed differing by the initial placement of the fragments and run in parallel, generating trajectories that amounted to 14 μs of aggregate simulation time. During each simulation, the protein was restrained to its initial coordinates while the DNA fragments were observed to spontaneously bind to and dissociate from the SSB surface, [Video S1](#). The resulting simulation trajectories were then mapped into CG DNA trajectories as described previously ([Maffeo et al., 2014](#)), [Figure 2B](#). From the CG representation of the all-atom trajectories, the three-dimensional, position-dependent target density was computed for each type of the DNA beads and averaged over the symmetry axes of the SSB homotetramer, [Figure 2C](#).

We began parameterization of our CG model of the ssDNA–SSB assembly by building six CG models of the above all-atom systems matching their nucleic acid composition and dimensions. The SSB protein was represented using a set of trial grid-based potentials that acted on the ssDNA beads. Naive application of the IBI procedure produced a model that resulted in excessive binding of DNA under tension, [Figure S1](#). Hence, we altered the IBI method to utilize a mask based on the SSB protein atom density that smoothly extended 0.5 nm from the protein, allowing only shorter-ranged interactions to emerge from the procedure. The initial guess for the potentials was obtained through Boltzmann inversion of the target densities, multiplied by the mask, [Figure 2D](#). The trial potentials were applied to the CG DNA beads during subsequent simulations performed using the ARBD software package ([Comer and Aksimentiev, 2012](#)) developed in-house. The resulting CG trajectories, lasting a total of 200 ns, were analyzed to produce density maps of ssDNA CG beads. For a given bead type, the ratio of the CG density to the target density provided a space-dependent correction to the SSB potential

$$\Delta u = -k_B T \ln \frac{\rho_{\text{target}} + \rho_0}{\rho_{\text{CG}} + \rho_0} \times M,$$

where the mask M ranged from zero away from the protein to one near the protein and $\rho_0 = 5 \times 10^{-7}$ beads/ \AA^3 protected against numerical instability. Note that one nanosecond of a CG simulation corresponds to considerably longer real-world time interval because of the smoothing of the interaction potentials (Zwanzig, 1988). By repeating the process for a total of 189 iterations, CG potentials (Figure 2E) were obtained that produced CG densities (Figures 2F, S2 and Video S2) in excellent agreement with the target densities. As a test, we simulated a dT₁₅₀ strand interacting with an SSB protein for 4 μs . During the simulation, the DNA strand wrapped the SSB protein, with an average of ~ 62 nt binding the protein (Figure S3), in slight excess but close to the number of nucleotides required to quench the natural fluorescence of SSB under similar conditions (Bell et al., 2015; Dubiel et al., 2019).

Experimental validation of the model

Force spectroscopy, usually employing one or more optical or magnetic traps, can be used to probe the mechanical properties of a biomolecular complex. The mechanical work required to disassemble a binary complex of biomolecules, which is equivalent to the binding free energy of the complex, can sometimes be inferred from experimentally measured force-extension curves. The Chemla lab characterized the free energy landscape of SSB binding to ssDNA using an experimental assay depicted schematically in Figures 3A and 3B (Suksombat et al., 2015). The extension of a DNA construct under tension is measured before (X_{ssDNA}) and after (X_{SSB}) SSB binding. The difference between these two extensions, $X_{\text{ssDNA}} - X_{\text{SSB}}$, indicates how much the DNA construct shrinks due to SSB binding. At low tensions, both X_{ssDNA} and X_{SSB} approach zero, and the difference between them is small. Under a moderate tension (~ 5 pN), the DNA is stretched against an entropic force. The DNA molecule exhibits less extension when SSB is bound to it because the ssDNA bound to the SSB is not unavailable for stretching. At slightly higher forces, SSB partially unbinds from ssDNA. At even higher forces, SSB dissociates from the DNA and $X_{\text{ssDNA}} - X_{\text{SSB}}$ approaches zero.

To test how well our CG model can reproduce the results of single-molecule experiments, we performed force-extension experiments on the SSB–ssDNA complex *in silico*. In each simulation, a dT₇₀ molecule was stretched by applying opposite constant forces to the P beads at each end of the molecule. First, the simulations were performed in the absence of SSB, Video S3. The extension during these simulations reached a steady-state value within ~ 50 ns. Low, average, and high extension configurations under a 5 pN force are shown in Figure 3A and a typical extension trace is plotted in Figure 3C.

We began our simulations of the force-extension dependence of an ssDNA–SSB complex from a configuration where ssDNA was wrapped around the SSB, Video S4. Complete dissociation of SSB from ssDNA was not observed, even at the highest applied tension studied (20 pN). Within 50 ns of the beginning of each simulation, the ssDNA extension reached a steady state. The protein–DNA complex was seen to be highly dynamic in the simulation, in particular when the tension force was near 5 pN. Low, average, and high extension configurations under a 5 pN tension are shown in Figure 3B along with an extension trace, Figure 3C.

For each simulation, the extension of the ssDNA molecule was obtained directly from the trajectory coordinates and then time-averaged to yield a single value. The difference between extensions without and with SSB as a function of applied force yielded a curve that could be directly compared to the experimentally obtained curve, Figure 3D. Although our model was not parameterized to reproduce the behavior of an SSB–ssDNA complex under applied force, we observed very close quantitative agreement between our simulation results and experiment.

Kinetics of SSB–ssDNA interactions

The process of coarse-graining effectively smoothens the interaction potentials, resulting in faster kinetics (Zwanzig, 1988). For the lengths of ssDNA considered in this study, the rate of end-to-end collisions of an ssDNA molecule is 5–50 times higher than in the corresponding experiment (Maffeo et al., 2014), whereas the timescale of free diffusion of an SSB particle in solution is exactly same as in experiment, prescribed by the diffusion constant. The difference in the rates of relative speedup observed for different components of the CG system can impact the quality of the kinetic data extracted from the simulations. For example, single-molecule experiments suggest that diffusion of SSB along DNA occurs through reptation (Zhou et al., 2011), whereby a bulge of DNA dissociates from the protein surface and propagates by breaking contacts on one side while forming new contacts on the other side of the bulge. While such bulges can be

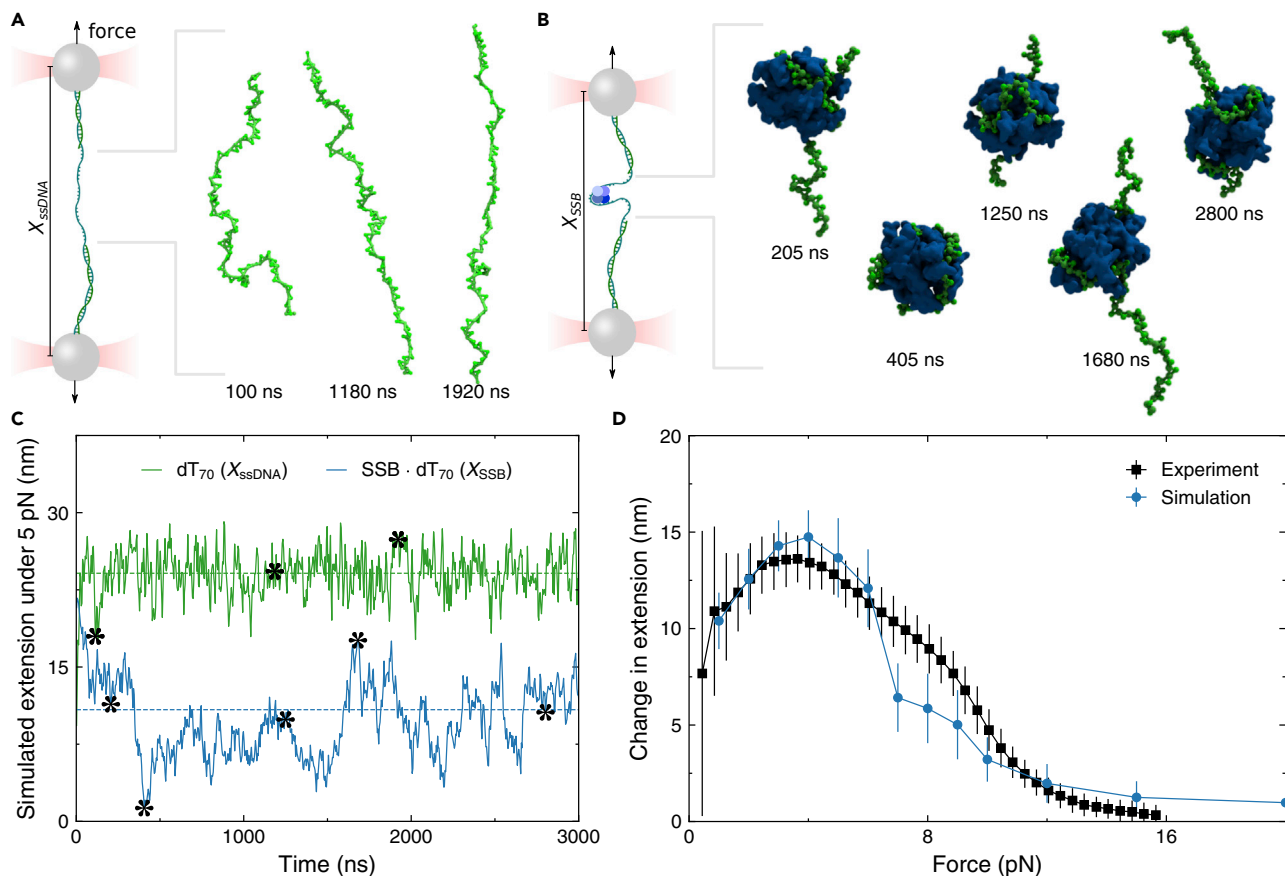


Figure 3. Force-extension dependence of an ssDNA-SSB complex measured *in vitro* and *in silico*

(A) Schematic of a single-molecule experiment including a pair of optical traps that has been used to study SSB (Suksombat et al., 2015). A 5 pN force was applied in opposite directions to the ends of a 70-nt ssDNA molecule in CG simulations representing the experiment. Exemplary configurations of the DNA are depicted on the right with the DNA represented by small green spheres connected by rods. See also Video S3.

(B) Simulation trajectory mimicking the experiment. A 5 pN force was applied in opposite directions to the ends of a 70-nt ssDNA molecule while an SSB protein was bound to the ssDNA. The extension of the ssDNA increases and decreases as the SSB molecule wraps and releases the ssDNA. Exemplary configurations of the ssDNA-SSB complex are depicted on the right, with the protein represented as a blue molecular surface. See also Video S4.

(C) Simulated extension of ssDNA in complex with SSB (blue) and in isolation (green). The asterisks denote the moments of the trajectory depicted in panels A and B. The dotted lines indicate the corresponding trajectory-average values.

(D) Average change in dT_{70} extension upon SSB dissociation experimentally measured *in vitro* (black) and *in silico* (blue) as a function of applied tension. Each simulated data point was obtained by time-averaging DNA extension and taking the difference between the values obtained with and without SSB. Experimental data were provided by the Chemla lab (Suksombat et al., 2015).

observed directly in all-atom MD simulations (Maffeo and Aksimentiev, 2017), diffusion in our model mainly occurs by sliding without reptation because the binding surface roughness was coarse-grained away. The thermodynamic quantities, however, are not expected to be affected by the timescale differences.

Besides resulting in a large boost in the rate of SSB diffusion along the ssDNA molecule, the artificially accelerated dynamics at the SSB-ssDNA interface causes rapid transitions between wrapping geometries of ssDNA on the SSB surface, Figures 4A and 4B. The extension of SSB-bound ssDNA under 5 pN tension was seen to change in discrete steps in single-molecule experiments (Suksombat et al., 2015), while, in our simulations, the transitions occurred so that steps could only barely be resolved, Figure 4B. Nevertheless, the likelihood of the protein occupying a given wrapping geometry is a thermodynamic property that is expected to be accurately captured by the model. Under tension, the most prevalent wrapping geometry at low force includes ssDNA making a “C” shape on the SSB surface, in contrast with the baseball-seam wrapping geometry suggested by the X-ray crystallographic structure (Raghuathan et al., 2000), which is the second most prevalent wrapping geometry. As the force increases, it becomes increasingly likely that the SSB loses contact with one or more of the four high-affinity binding sites, also called oligonucleotide binding (OB) folds, Figure 4C.

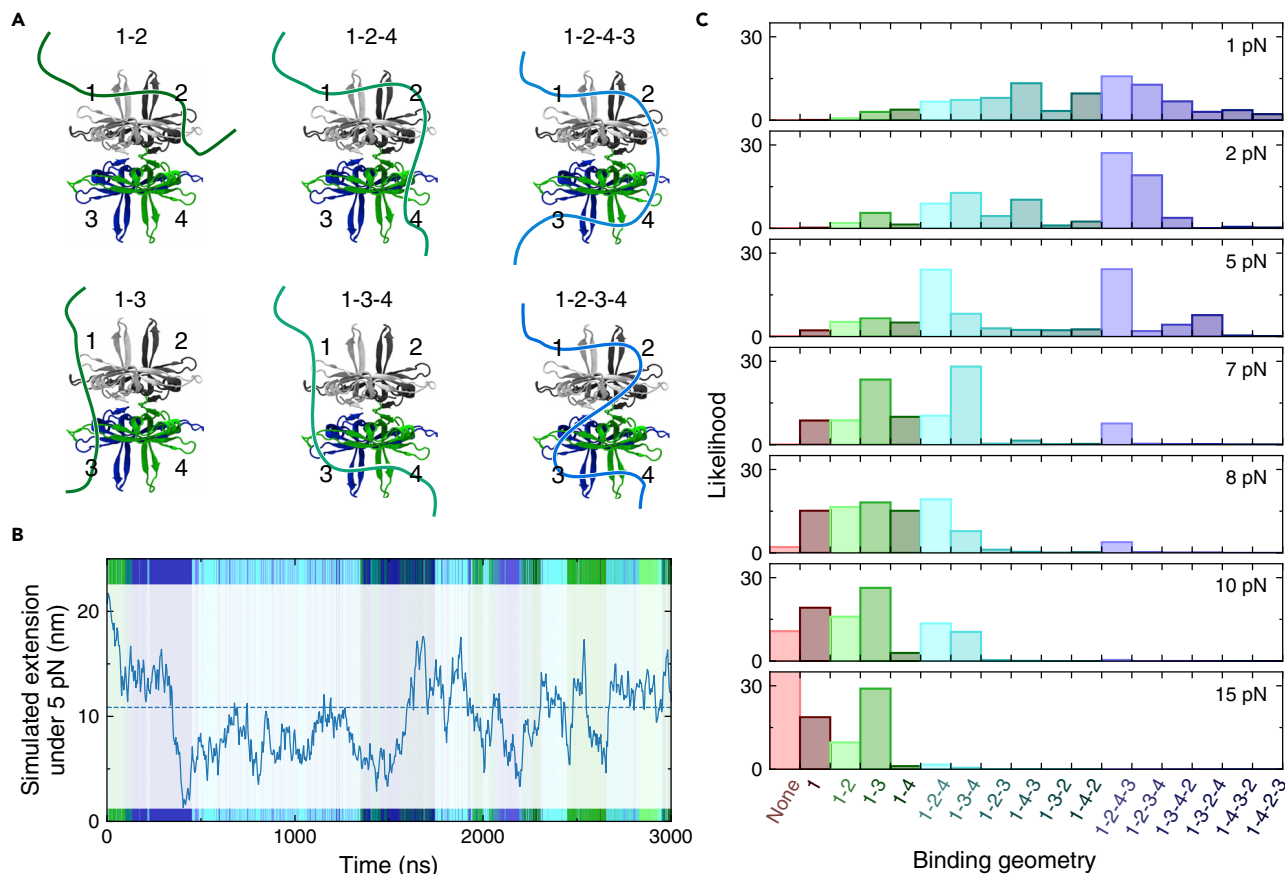


Figure 4. Wrapping geometry of ssDNA on SSB observed in CG simulations

(A) Labeling scheme used to classify the wrapping geometry. Each of the four high-affinity OB folds is assigned an integer. The wrapping geometry can be specified by the sequence in which the DNA strand contacts the OB folds. The symmetry of the SSB tetramer allows the first OB fold contacted by the 5' end of the strand to be labeled "1".

(B) Wrapping geometry during simulation of SSB-dT₇₀ complex under 5 pN tension. The blue trace depicts the instantaneous extension of the complex. At each moment in time, the color in the background of the plot indicates the wrapping geometry using the same color scheme specified in panel C.

(C) Force dependence of the wrapping geometry. The likelihood of the system occupying a given wrapping geometry is shown for the tension force ranging from 1 to 15 pN. At low force, wrapping geometries where ssDNA contacts all four OB folds are most likely. At high forces, wrapping geometries with only two OB fold contacts are more likely.

Transfer of SSB from one DNA strand to another

Single-molecule experiments can be designed to probe biologically important behavior of biomolecules. For example, DNA replication enzymes generally operate only in the 5'-to-3' direction so that replication on one of the strands is performed in short (~ 500 bp) sections between priming sites called Okazaki fragments. The ssDNA exposed in these fragments is coated by SSB proteins that bind tightly, yet almost paradoxically must easily be repositioned as the polymerase causes the ssDNA loop to shrink. Analysis of the diffusion of a fluorescently labeled SSB molecule demonstrates that the motion of SSB bound to a long ssDNA construct can be so fast that SSB must be hopping along different, non-continuous fragments of the same ssDNA construct (Lee et al., 2014), and not sliding or rolling (Roy et al., 2007; Zhou et al., 2011). Subsequent single-molecule experiments characterized the kinetics of exchanging one DNA strand bound to an SSB molecule for another (Yang and Ha, 2018). These experiments demonstrate that SSB can jump from one region of ssDNA to another within the replication fork, allowing it to move out of the way of an active polymerase. The experiments primarily tell us that the exchange process happens, but it remains difficult to determine the steps that typically occur during the exchange, including such details as the structure of the encounter complex where two strands simultaneously bind one SSB protein.

Our CG model of SSB and ssDNA enables simulations of such spontaneous transfer processes, which are much too slow to simulate using the all-atom MD approach. As a demonstration, we constructed a system containing a

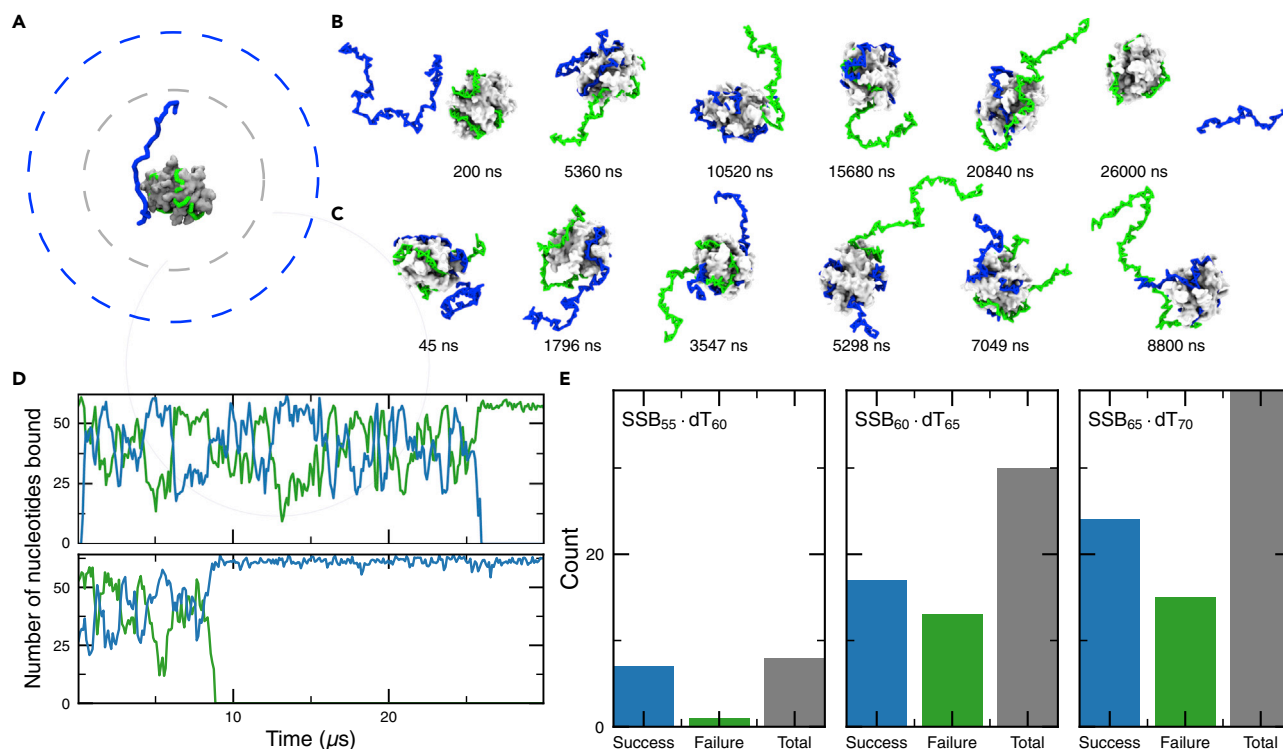


Figure 5. Transfer of SSB between two different-length DNA strands

(A) System containing SSB and two DNA strands, one initially in complex with SSB (green; 65 nt) and one invading strand (blue; 70 nt). The SSB is shown as a gray molecular surface, and the dashed lines represent spherical potentials used to confine the SSB (gray) and DNA strands (blue).

(B) Encounter that results in the dissociation of the invading strand. See also [Video S5](#).

(C) Encounter that results in the dissociation of the incumbent strand. See also [Video S6](#).

(D) Number of nucleotides bound to SSB from each strand during the simulations depicted in panels C and D. The number for initially bound and invading strands is respectively shown using green and blue lines. A nucleotide was considered bound to the SSB if a P bead of ssDNA was found within 1 nm of the non-hydrogen atoms of SSB. The CG representation of SSB was replaced by its all-atom representation for this analysis.

(E) Statistical outcome of encounter. By repeating the encounter simulation multiple times, the likelihood of an encounter resulting in a successful transfer of the SSB to the invading strand could be determined. The outcome of an encounter was also determined for dT₆₅ invading an SSB–dT₆₀ complex and for dT₆₀ invading an SSB–dT₅₅ complex.

complex of SSB and dT₆₅ and a free dT₇₀ molecule, confining the free DNA molecule to reside within a sphere of a 25 nm radius using a half harmonic potential (spring constant $k = 1 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ per bead), whereas the center of the SSB was confined to a smaller, 10 nm radius sphere ($k = 8 \text{ kcal}^{-1} \text{ mol}^{-1} \text{ \AA}^{-2}$), [Figure 5A](#). The incumbent DNA strand was not confined and was free to leave the immediate volume around the SSB, effectively preventing rebinding. The confinement potential creates a high effective concentration of $\sim 25 \mu\text{M}$ for the invading dT₇₀ molecule, which enhances the rate at which an encounter complexes form. During the simulation, we observed the 70-nt strand to bind to the SSB and compete for high-affinity area at the SSB surface until either the invading strand dissociates, [Figure 5B](#), or the incumbent strand dissociates [Figure 5C](#). The strand exchange process is well characterized by the time series of the number of nucleotides of each strand bound to the SSB, [Figure 5D](#).

The process of SSB transfer took less than a day to simulate on a single workstation. Hence, it was feasible to replicate the simulation many times in parallel to generate results that could be expressed in statistical terms. Forty-eight trajectories were produced, and the time series of the number of nucleotides bound to SSB was extracted for each strand in each trajectory. The outcome of attempted transfer events was identified algorithmically as successful or failed, depending on whether the incumbent or the invading strand dissociated from the complex, [Figure 5E](#). Representative simulation trajectories of failed and successful transfer events are highlighted in [Videos S5](#) and [S6](#). Similar simulations were carried out to examine the competition between shorter dT₆₀ and dT₆₅ strands and dT₅₅ and dT₆₀ strands.

Once an encounter complex is formed, it appears that the likelihood of successfully transferring the SSB is sensitive to the overall length of the two strands, with dT₇₀ displacing dT₆₅ $\sim 60\%$ of the time and dT₆₀

displacing dT₅₅ nearly 90% of the time. The likelihood of the invading strand displacing the incumbent strand appears to have a nonlinear relationship with chain length because long DNA strands can more easily form contacts with three or four high-affinity OB folds. For example, when a dT₅₅ strand was bound to the SSB, the DNA made contact with three or more OB folds just 40% of the time, compared to 56% for a dT₆₀ strand and 66% for a dT₆₅ strand.

Exchange of SSB proteins bound to ssDNA

Single-molecule experiments from the Ha lab demonstrated the ease with which SSB can be repositioned or replaced by another protein (Roy et al., 2009; Yang and Ha, 2018). In a typical experimental assay, a fluorescently labeled SSB molecule is initially bound to a surface-tethered DNA construct that has an exposed dT₇₀ strand. The construct has a 5' end of the ssDNA embedded within a dsDNA duplex which in turn is tethered to a glass slide, leaving the 3' end of dT₇₀ exposed to solution. When a second SSB molecule encounters the SSB–ssDNA complex, it can replace the incumbent SSB (Yang and Ha, 2018). Three states could be detected via the measurement of the FRET efficiency, corresponding to a single labeled SSB molecule bound to the DNA, two SSBs bound to the DNA, and, finally, an unlabeled SSB from the solution bound to the DNA. However, the experiments are unable to detect intermediates that likely occur during the exchange process where two SSBs are simultaneously bound to a single DNA construct.

We recreated this experimental assay *in silico*. In typical simulations, a DNA construct was tethered to a half harmonic repulsive surface. One SSB molecule was partially wrapped with the ssDNA part of the construct and a second SSB particle was placed 8 Å away from the nearest DNA bead, Figure 6A. To describe the SSB–SSB interactions, we used a previously developed protocol (Singharoy et al., 2019) that clusters atoms according to the CHARMM36 (Huang and MacKerell Jr, 2013) Lennard-Jones (LJ) parameters, constructs approximate 3D LJ potential maps, an electrostatic potential map from APBS (Holst and Saied, 1996; Jurrus et al., 2018), and maps with the distribution of corresponding LJ and electrostatic charges. Note that the sum over an LJ charge map equals the number of atoms of the corresponding LJ type in the SSB protein; negative LJ charges are never employed. If a voxel of the charge map (generalized, not necessarily electrostatic) of one SSB overlaps with the potential map of the other, a force is computed as the charge of that voxel multiplied by a negative gradient of the potential at that voxel, with the latter being determined using linear interpolation of the potential map. The corresponding torque is computed as a cross product between the vector connecting the center of the SSB (with the charge map) with the center of the charge voxel and the force on the charge voxel. For efficiency, grid-grid interactions were only computed every 20 steps. We reiterate that our model of SSB neglects the disordered C-terminal tails, which have been shown to promote cooperative binding of SSBs to a long DNA fragment (Kozlov et al., 2015).

Immediately after the simulation started, the second, upper SSB molecule became bound to the ssDNA, Figure 6B. After 4.4 μs, the free end of the ssDNA begins to make contact with the lower SSB molecule (the one initially bound near the dsDNA) as the upper SSB spontaneously loses contacts with the DNA. Within 200 ns, the fluctuations cause the upper SSB to move toward the surface to which the DNA is tethered so that it can compete with the lower SSB for contacts with the ssDNA near the duplex, first three panels of Figure 6B. This competition continues for another hundred nanoseconds, whereupon the DNA near the ssDNA–dsDNA junction spontaneously unbinds from the lower SSB, such that exchange is complete, last two panels of Figure 6B. The entire exchange process is highlighted in Video S7. Examination of the contacts formed between the ssDNA and the SSBs allows the exchange to be detected algorithmically, Figure 6C. Forty-seven additional replicas of the system, starting from the same initial conformation were then independently simulated for a similar duration to determine whether the exchange event could be observed repeatedly and to obtain an estimate of the exchange rate.

In each simulation, both SSBs were constantly in contact with the ssDNA, except for one replica where complete dissociation of an SSB molecule was observed. Taken together, the simulations provided a statistical view of the exchange process within the intermediate complex, allowing the rate of exchange to be determined ($\sim 0.28 \mu\text{s}^{-1}$) from a fit to the dwell time distribution, Figure 6D. We stress, however, that the absolute exchange rate determined from our CG simulations should be interpreted cautiously because of the smoothing of the underlying free energy landscape, which results in a process-dependent acceleration of the system's dynamics (Zwanzig, 1988). Rather than examining the absolute rates, such simulations could be used to compare the exchange rates, for example, probing how the exchange rate depends on the length of the ssDNA or perhaps on modifications of the SSB structure. Our simulation approach could

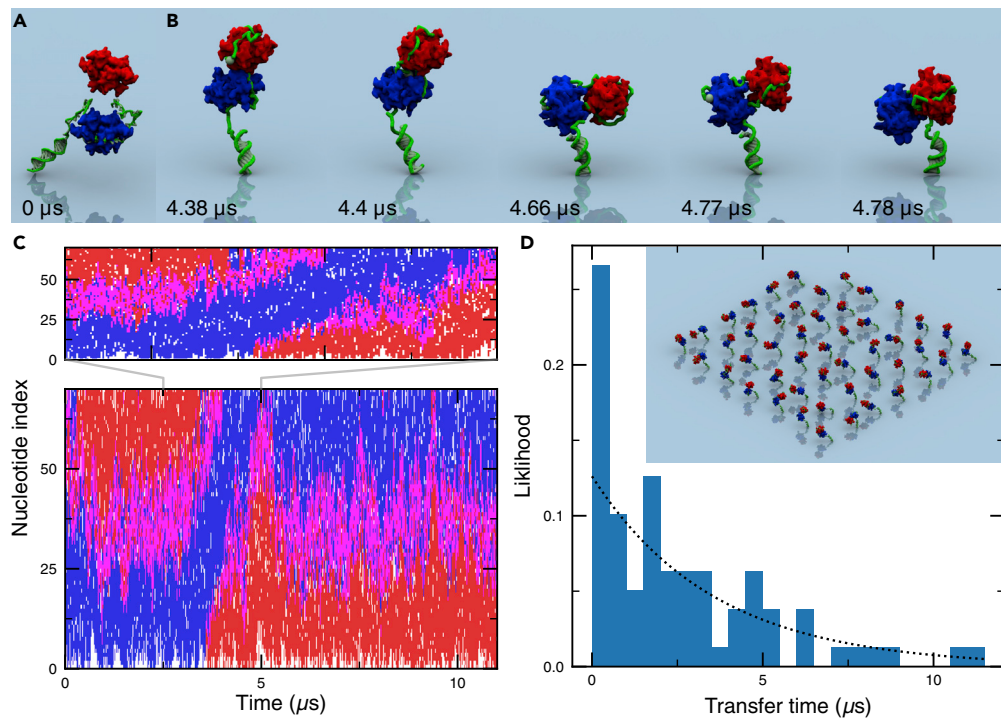


Figure 6. Exchange of two SSB proteins bound to the same ssDNA fragment

(A) System containing a DNA construct tethered to a plane by a harmonic potential with two SSB molecules in close proximity.

(B) Snapshots illustrating the simulated process of SSB exchange. See also [Video S7](#).

(C) Exemplary trace depicting the instantaneous contact status of each ssDNA nucleotide of the DNA construct. For a given configuration, each nucleotide is either bound to the SSB initially placed near the ssDNA–dsDNA junction (blue), to the SSB initially bound to the free 3' end of the ssDNA (red), to both SSB proteins (magenta), or neither SSB (white). A nucleotide was considered to be in contact with the SSB if its P bead was within 1 nm of any non-hydrogen SSB atoms. The CG representation of SSB was replaced by its all-atom equivalent for this analysis. The top graph shows a zoomed-in view of the bottom graph.

(D) Probability of observing an exchange event versus simulation time. The exchange events were detected algorithmically by determining the median index of nucleotides bound to each SSB protein. Identical initial configurations were used to initiate 48 independent simulations, differing by the random generator seed. Multiple transfer events within a continuous simulation trajectory were considered as independent events. The inset illustrates instantaneous configurations of the 48 simulation systems; each system was simulated in isolation from the others. Dotted line shows an exponential fit to the distribution.

be generalized to model competitive binding of SSB and other DNA-binding proteins, such as RecA, to the same DNA fragment, providing insights into the microscopic mechanisms of DNA repair ([Maffeo et al., 2019](#)).

Model of the T7 bacteriophage trombone loop saturated with bacterial SSB

According to the current model of DNA replication in prokaryotes and some viruses, the parental dsDNA is unwound by a helicase that is in complex with a pair of DNA polymerase molecules. One polymerase replicates the leading strand, traversing the DNA in the same direction that the helicase moves, while the other polymerase replicates the lagging strand in the opposite direction, [Figure 7A](#). The lagging strand polymerase therefore must “back-stitch”, replicating the DNA in a piecewise manner while the “trombone” loop grows, on one end, from ssDNA released by the helicase and, on the other, from daughter dsDNA at the polymerase. The main biological function of SSB proteins is to protect the lagging strand exposed during DNA replication from degradation by endonucleases and secondary structure formation.

We constructed a physical model of SSB proteins in the context of the replication fork, using the minimalist replication machinery of the T7 bacteriophage as our target system. We manually docked structured parts

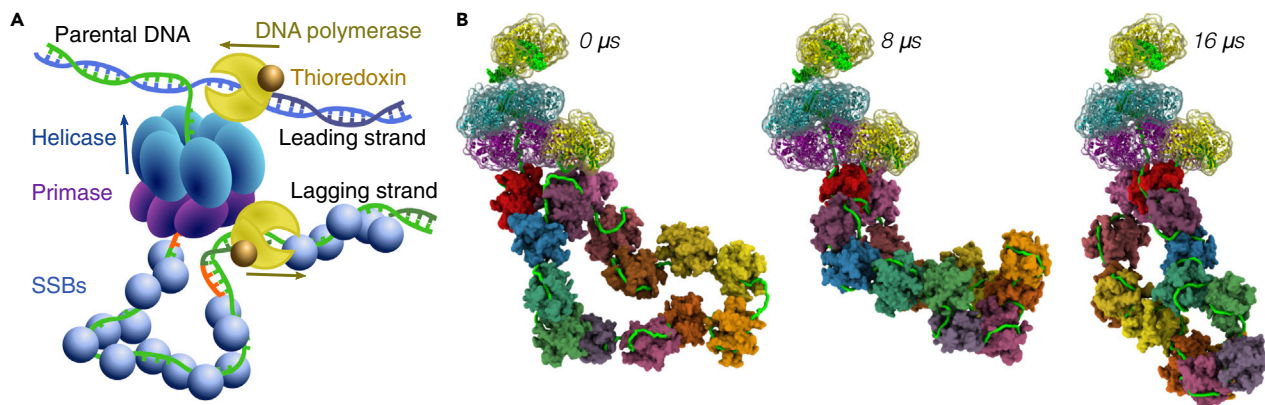


Figure 7. Physical model of the trombone loop of the DNA replication fork

(A) Schematic depicting the T7 DNA replication fork.

(B) Snapshots of the simulation of the trombone loop. Semi-transparent molecular surfaces represent the helicase (teal) in complex with the primase (purple) and polymerase (yellow) proteins, which could be resolved in cryo-EM and are represented in the simulation by grid-based potentials that act on the ssDNA and SSB proteins. Double-stranded DNA and ssDNA beads bound to the above proteins are depicted but were not represented explicitly in the simulations. The remaining ssDNA (green beads and rods) is coated with SSB proteins, each drawn as rainbow-colored molecular surface. See also [Videos S8](#) and [S9](#).

of the helicase, primase, thioredoxin, and polymerase proteins into the cryo-EM map of the replication fork and used the procedure described in the previous section for SSB-SSB interactions to convert the atomic models into potential maps that would act on the SSB rigid body proteins. In addition, we used the cryo-EM map as a potential to sterically repel ssDNA from the structured parts of the replication fork. We then constructed a model using these potential maps with a 500-nt loop of ssDNA saturated with 14 SSB molecules to represent the trombone loop of the replication fork. The model of the loop was constructed by applying transformations to a previously equilibrated model of the ssDNA-SSB construct. The DNA loop was anchored at locations resolved by cryo-EM near the primase on the one end and the lagging-strand polymerase at the other. Our model represents the state of the replication fork immediately after the lagging polymerase binds a primed template strand and as such does not include an explicit bead-based representation of the dsDNA. To the best of our knowledge, this is the first structurally correct physical model of a trombone loop.

During the 16 μ s simulation, all SSB proteins remained bound to the ssDNA of the trombone loop, whereas the loop as a whole was seen to undergo significant rearrangements, [Figure 7B](#) and [Video S8](#). Most prominently, the loop became more compact with SSB proteins sometimes making contact with more than one region of the ssDNA strand, effectively sharing the strand. After 7 μ s, the second SSB from the polymerase dissociated spontaneously from the DNA for 300 ns before rebinding at the same location, [Video S9](#). Interestingly, a single SSB protein became stably bound to both ends of the loop where it enters the structured region of the replication fork, a prediction that could be tested experimentally.

DISCUSSION

Using a bottom-up approach, we have developed a CG model of a structured DNA-binding protein, SSB, compatible with a previously developed 2-bead-per-nucleotide model of ssDNA. The key methodological innovation of our approach is that the protein is modeled as a rigid body decorated with grid-based potential maps that have been derived to reproduce results of all-atom MD simulations. Furthermore, we have extended our grid-based formulation of interaction potentials to describe interactions between multiple SSB proteins interacting with each other, with ssDNA, and with proteins represented by a stationary grid potential derived from 3D cryo-electron microscopy maps. Thus, we believe our model provides a preliminary framework for accurate physical modeling of DNA replication and repair processes.

Development of interaction potentials between ssDNA and SSB required substantial computational resources as it relied on brute-force sampling of ssDNA configurations around the SSB. Using enhanced sampling methods, such as the adaptive biasing force ([Comer et al., 2015](#)), may considerably reduce the computational cost of the 3D potential parameterization. Our parameterization of SSB-SSB potentials,

as well as the 3D potentials governing the interactions between SSB and the structurally resolved parts of the replication fork, was done using a computationally inexpensive approach that combined an approximate treatment of van der Waals interactions and Poisson–Boltzmann electrostatics (Singharoy et al., 2019). In doing so, we have neglected solvent-mediated interactions, which may be important for accurate description of the specific and nonspecific binding of biomolecules. Note that our brute-force approach accounts for the solvent-mediated effects. Beyond reproducing results that could be determined experimentally, the model revealed the statistics of the force-dependent wrapping geometry, which cannot be easily obtained using experimental techniques. Similarly, the processes of DNA transfer and SSB exchange pathways could be directly observed, which would be extremely challenging using single-molecule experimental techniques.

While reproducing thermodynamic quantities, accurate modeling of kinetic processes remains an outstanding challenge for CG models of heterogeneous biomolecular systems, with our work being no exception. In simple terms, once a heterogeneous system has been coarse-grained, each process occurring in the system will be sped up by the smoothing of the free energy landscape. The path traveled along the free energy landscape by the system for one process might be smoothed more than for another process, which will give a larger speedup for the former. Hence, each process effectively runs on its own clock, which makes prediction or extraction of the reaction rates extremely challenging. As an example, the sliding motion of ssDNA on the SSB surface appears to be accelerated significantly more than other processes, such as the association/dissociation of ssDNA from the SSB surface. A solution to this type of problem may require extra knowledge of the kinetics of the processes of interest. In the case of SSB, one could use experimentally known rates to slow ssDNA sliding by introducing additional friction terms that oppose DNA bead motion along the SSB surface or by adopting a stateful description of nucleotide binding the SSB surface, in either case tuning parameters to match experimental data.

A few other limitations of our model should be appreciated. First, the employed ssDNA model was parameterized against all-atom simulations of dT_{60} , and hence lacks sequence-dependent energetics or the possibility of forming static secondary structure. Additionally, the procedure used to develop the model coarse-grains away the explicit representation of ions, making the resulting model representative of a particular thermodynamic state, in this case a fixed chemical potential with 100 mM NaCl and a temperature of 291 K. Similarly, the parameters emerging from the IBI procedure may be influenced by the concentration of DNA nucleotides in the all-atom system, which were chosen to represent DNA-saturating conditions.

For the future, we envision an extension of the model presented here, where the proteins of the replication fork are mobile and some are able to stochastically alter the state of nearby DNA, for example unwinding the DNA or synthesizing the complementary strand, in a tension-dependent manner. While such simulations may at first appear to be destined for the distant future, we believe they are within the grasp of current computational technology. Indeed, fifty years of biochemistry has provided us with extensive data regarding the individual biochemical steps, while single-molecule experiments have shown how individual biomolecules behave. Reproducing single-molecule biochemistry *in silico* will be the next enabling step toward accurate physical modeling of multi-component biological assemblies.

Limitations of the study

The CG model of SSB–DNA interaction is based primarily upon all-atom MD simulations and therefore inherits flaws from the atomistic force field. The model is parameterized to reproduce a particular thermodynamic state of the system, e.g. its ion and/or nucleotide concentrations. The model assumes ssDNA to lack secondary structure. The model neglects C-terminal SSB tails and incorporates atomistic SSB–SSB interactions approximately. The coarse-graining procedure accelerates the dynamics in the system in a process-dependent manner, for example eliminating energetic barriers to sliding diffusion but not dissociation.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact

- Materials availability
- Data and code availability
- **METHOD DETAILS**
 - General simulation protocols
 - Generation of target data for ssDNA–SSB interaction parametrization
 - Refinement of CG SSB potentials
 - Motion of the SSB particle

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2022.104264>.

ACKNOWLEDGMENTS

This work was supported by the grants from the National Science Foundation (PHY-1430124), and the National Institutes of Health (GM137015 and P41-GM104601). The supercomputer time was provided through XSEDE Allocation Grant MCA05S028 and the Leadership Resource Allocation MCB20012 on Frontera of the Texas Advanced Computing Center. A.A. and C.M. thank Taekjip Ha and Yann Chemla for insightful discussions and sharing single molecule data.

AUTHOR CONTRIBUTIONS

A.A. and C.M. designed the research. H-Y.C. and C.M. performed the research. C.M., H-Y.C., and A.A. wrote the manuscript.

DECLARATIONS OF INTERESTS

None to declare.

Received: November 16, 2021

Revised: March 23, 2022

Accepted: April 12, 2022

Published: May 20, 2022

REFERENCES

- Andersen, H.C. (1983). RATTLE: a “velocity” version of the SHAKE algorithm for molecular dynamics calculations. *J. Comput. Phys.* 52, 24–34. [https://doi.org/10.1016/0021-9991\(83\)90014-1](https://doi.org/10.1016/0021-9991(83)90014-1).
- Antony, E., Weiland, E., Yuan, Q., Manhart, C.M., Nguyen, B., Kozlov, A.G., McHenry, C.S., and Lohman, T.M. (2013). Multiple C-terminal tails within a single *E. coli* SSB homotetramer coordinate DNA replication and repair. *J. Mol. Biol.* 425, 4802–4819. <https://doi.org/10.1016/j.jmb.2013.08.021>.
- Arya, G., and Schlick, T. (2006). Role of histone tails in chromatin folding revealed by a mesoscopic oligonucleosome model. *Proc. Natl. Acad. Sci. U S A.* 103, 16236–16241. <https://doi.org/10.1073/pnas.0604817103>.
- Batcho, P.F., Case, D.A., and Schlick, T. (2001). Optimized particle-mesh Ewald/multiple-time step integration for molecular dynamics simulations. *J. Chem. Phys.* 115, 4003–4018. <https://doi.org/10.1063/1.1389854>.
- Beglov, D., and Roux, B. (1994). Finite representation of an infinite bulk system: solvent boundary potential for computer simulations. *J. Chem. Phys.* 100, 9050–9063. <https://doi.org/10.1063/1.466711>.
- Bell, J.C., Liu, B., and Kowalczykowski, S.C. (2015). Imaging and energetics of single SSB-ssDNA molecules reveal intramolecular condensation and insight into RecOR function. *Elife* 4, e08646. <https://doi.org/10.7554/elife.08646>.
- Brünger, A.T. (1992). X-PLOR, Version 3.1: A System for X-Ray Crystallography and NMR (The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University).
- Choudhary, A., Joshi, H., Chou, H.-Y., Sarthak, K., Wilson, J., Maffeo, C., and Aksimentiev, A. (2020). High-fidelity capture, threading, and infinite-depth sequencing of single DNA molecules with a double-nanopore system. *ACS Nano* 14, 15566–15576. <https://doi.org/10.1021/acsnano.0c06191>.
- Comer, J., and Aksimentiev, A. (2012). Predicting the DNA sequence dependence of nanopore ion current using atomic-resolution Brownian dynamics. *J. Phys. Chem. C* 116, 3376–3393. <https://doi.org/10.1021/jp210641j>.
- Comer, J., Gumbart, J.C., Hénin, J., Lelièvre, T., Pohorille, A., and Chipot, C. (2015). The adaptive biasing force method: everything you always wanted to know but were afraid to ask. *J. Phys. Chem. B* 119, 1129–1151. <https://doi.org/10.1021/jp506633n>.
- Davtyan, A., Schafer, N.P., Zheng, W., Clementi, C., Wolynes, P.G., and Papoian, G.A. (2012). AWSEM-MD: protein structure prediction using coarse-grained physical potentials and bioinformatically based local structure biasing. *J. Phys. Chem. B* 116, 8494–8503. <https://doi.org/10.1021/jp212541y>.
- De Jong, D.H., Singh, G., Bennett, W.F.D., Arnarez, C., Wassenaar, T.A., Schäfer, L.V., Periole, X., Tieleman, D.P., and Marrink, S.J. (2013). Improved parameters for the Martini coarse-grained protein force field. *J. Chem. Theor. Comput.* 9, 687–697. <https://doi.org/10.1021/ct300646g>.
- Denning, E.J., Priyakumar, U.D., Nilsson, L., and MacKerell, A.D., Jr. (2011). Impact of 2-hydroxyl sampling on the conformational properties of RNA: update of the CHARMM all-atom additive force field for RNA. *J. Comput. Chem.* 32, 1929–1943. <https://doi.org/10.1002/jcc.21777>.
- Dubiel, K., Myers, A.R., Kozlov, A.G., Yang, O., Zhang, J., Ha, T., Lohman, T.M., and Keck, J.L. (2019). Structural mechanisms of cooperative DNA binding by bacterial single-stranded DNA-binding proteins. *J. Mol. Biol.* 431, 178–195. <https://doi.org/10.1016/j.jmb.2018.11.019>.
- Dudko, O.K., Hummer, G., and Szabo, A. (2008). Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proc.*

- Natl. Acad. Sci. U S A. 105, 15755–15760. <https://doi.org/10.1073/pnas.0806085105>.
- Dullweber, A., Leimkuhler, B., and McLachlan, R. (1997). Symplectic splitting methods for rigid body molecular dynamics. *J. Chem. Phys.* 107, 5840–5851. <https://doi.org/10.1063/1.474310>.
- Foloppe, N., and MacKerell Jr, A.D. (2000). All-atom empirical force field for nucleic acids: I. parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comput. Chem.* 21, 86–104. [https://doi.org/10.1002/\(sici\)1096-987x\(20000130\)21:2<86::aid-jcc2>3.0.co;2-g](https://doi.org/10.1002/(sici)1096-987x(20000130)21:2<86::aid-jcc2>3.0.co;2-g).
- Gao, Y., Cui, Y., Fox, T., Lin, S., Wang, H., de Val, N., Zhou, Z.H., and Yang, W. (2019). Structures and operating principles of the replisome. *Science* 363, eaav7003. <https://doi.org/10.1126/science.aav7003>.
- Ghavamian, A., van der Giessen, E., and Onck, P.R. (2016). Energetics of transport through the nuclear pore complex. *PLoS One* 11, e0148876. <https://doi.org/10.1371/journal.pone.0148876>.
- Gkeka, P., Stoltz, G., Barati Farimani, A., Belkacemi, Z., Ceriotti, M., Chodera, J.D., Dinner, A.R., Ferguson, A.L., Maillet, J.-B., Minoux, H., et al. (2020). Machine learning force fields and coarse-grained variables in molecular dynamics: application to materials and biological systems. *J. Chem. Theor. Comput.* 16, 4757–4775. <https://doi.org/10.1021/acs.jctc.0c00355>.
- Gordon, D., Hoyles, M., and Chung, S.-H. (2009). Algorithm for rigid-body Brownian dynamics. *Phys. Rev. E* 80, 066703. <https://doi.org/10.1103/PhysRevE.80.066703>.
- Grime, J.M.A., Dama, J.F., Ganser-Pornillos, B.K., Woodward, C.L., Jensen, G.J., Yeager, M., and Voth, G.A. (2016). Coarse-grained simulation reveals key features of HIV-1 capsid self-assembly. *Nat. Commun.* 7, 11568. <https://doi.org/10.1038/ncomms11568>.
- Hart, K., Foloppe, N., Baker, C.M., Denning, E.J., Nilsson, L., and MacKerell, A.D., Jr. (2012). Optimization of the CHARMM additive force field for DNA: improved treatment of the BI/BII conformational equilibrium. *J. Chem. Theor. Comput.* 8, 348–362. <https://doi.org/10.1021/ct200723y>.
- Hinckley, D.M., Freeman, G.S., Whitmer, J.K., and de Pablo, J.J. (2013). An experimentally-informed coarse-grained 3-site-per-nucleotide model of DNA: structure, thermodynamics, and dynamics of hybridization. *J. Chem. Phys.* 139, 144903. <https://doi.org/10.1063/1.4822042>.
- Hollingsworth, S.A., and Dror, R.O. (2018). Molecular dynamics simulation for all. *Neuron* 99, 1129–1143. <https://doi.org/10.1016/j.neuron.2018.08.011>.
- Holst, M.J., and Saied, F. (1996). Numerical solution of the nonlinear Poisson–Boltzmann equation: developing more robust and efficient methods. *J. Comput. Chem.* 16, 337–364. <https://doi.org/10.1002/jcc.540160308>.
- Huang, J., and MacKerell, A.D., Jr. (2013). Charmm36 all-atom additive protein force field: validation based on comparison to nmr data. *J. Comput. Chem.* 34, 2135–2145. <https://doi.org/10.1002/jcc.23354>.
- Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: Visual molecular dynamics. *J. Mol. Graphics* 14, 33–38. [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5).
- Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926–935. <https://doi.org/10.1063/1.445869>.
- Jurrus, E., Engel, D., Star, K., Monson, K., Brandi, J., Felberg, L.E., Brookes, D.H., Wilson, L., Chen, J., Liles, K., et al. (2018). Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* 27, 112–128. <https://doi.org/10.1002/pro.3280>.
- Karplus, M., and McCammon, J.A. (2002). Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* 9, 646–652. <https://doi.org/10.1038/nsb0902-646>.
- Kim, H., and Ha, T. (2013). Single-molecule nanometry for biological physics. *Rep. Progr. Phys.* 76, 016601. <https://doi.org/10.1088/0034-4885/76/1/016601>.
- Kim, Y.C., and Hummer, G. (2008). Coarse-grained models for simulations of multiprotein complexes: application to ubiquitin binding. *J. Mol. Biol.* 375, 1416–1433. <https://doi.org/10.1016/j.jmb.2007.11.063>.
- Klepeis, J.L., Lindorff-Larsen, K., Dror, R.O., and Shaw, D.E. (2009). Long-timescale molecular dynamics simulations of protein structure and function. *Curr. Opin. Struct. Biol.* 19, 120–127. <https://doi.org/10.1016/j.sbi.2009.03.004>.
- Kozlov, A.G., Cox, M.M., and Lohman, T.M. (2010). Regulation of single-stranded DNA binding by the C termini of Escherichia coli single-stranded DNA-binding (SSB) protein. *J. Biol. Chem.* 285, 17246–17252. <https://doi.org/10.1074/jbc.M110.118273>.
- Kozlov, A.G., Shinn, M.K., Weiland, E.A., and Lohman, T.M. (2017). Glutamate promotes SSB protein–protein interactions via intrinsically disordered regions. *J. Mol. Biol.* 429, 2790–2801. <https://doi.org/10.1016/j.jmb.2017.07.021>.
- Kozlov, A.G., Weiland, E., Mittal, A., Waldman, V., Antony, E., Fazio, N., Pappu, R.V., and Lohman, T.M. (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, 763–774. <https://doi.org/10.1016/j.jmb.2014.12.020>.
- Lázaro, G.R., Mukhopadhyay, S., and Hagan, M.F. (2018). Why enveloped viruses need cores—the contribution of a nucleocapsid core to viral budding. *Biophys. J.* 114, 619–630. <https://doi.org/10.1016/j.bpj.2017.11.3782>.
- Lee, K.S., Marciel, A.B., Kozlov, A.G., Schroeder, C.M., Lohman, T.M., and Ha, T. (2014). Ultrafast redistribution of *E. coli* SSB along long single-stranded DNA via intersegment transfer. *J. Mol. Biol.* 426, 2413–2421. <https://doi.org/10.1016/j.jmb.2014.04.023>.
- Leimkuhler, B., and Matthews, C. (2016). Efficient molecular dynamics using geodesic integration and solvent–solute splitting. *Proc. R. Soc. A* 472, 20160138. <https://doi.org/10.1098/rspa.2016.0138>.
- Lindahl, E., and Sansom, M.S.P. (2008). Membrane proteins: molecular dynamics simulations. *Curr. Opin. Struct. Biol.* 18, 425–431. <https://doi.org/10.1016/j.sbi.2008.02.003>.
- Louis, A.A. (2002). Beware of density dependent pair potentials. *J. Phys. Condens. Matter* 14, 9187–9206. <https://doi.org/10.1088/0953-8984/14/40/311>.
- Lyubartsev, A.P., and Laaksonen, A. (1995). Calculation of effective interaction potentials from radial distribution functions: a reverse Monte Carlo approach. *Phys. Rev. E* 52, 3730–3737. <https://doi.org/10.1103/PhysRevE.52.3730>.
- MacKerell, A.D., Jr., and Banavali, N.K. (2000). All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution. *J. Comput. Chem.* 21, 105–120. [https://doi.org/10.1002/\(sici\)1096-987x\(20000130\)21:2<105::aid-jcc3>3.0.co;2-p](https://doi.org/10.1002/(sici)1096-987x(20000130)21:2<105::aid-jcc3>3.0.co;2-p).
- Maffeo, C., and Aksimentiev, A. (2017). Molecular mechanism of DNA association with single-stranded DNA binding protein. *Nucleic Acids Res.* 45, 12125–12139. <https://doi.org/10.1093/nar/gkx917>.
- Maffeo, C., Chou, H.-Y., and Aksimentiev, A. (2019). Molecular mechanisms of DNA replication and repair machinery: insights from microscopic simulations. *Adv. Theor. Simul.* 2, 1800191. <https://doi.org/10.1002/adts.201800191>.
- Maffeo, C., Ngo, T.T.M., Ha, T., and Aksimentiev, A. (2014). A coarse-grained model of unstructured single-stranded DNA derived from atomistic simulation and single-molecule experiment. *J. Chem. Theor. Comput.* 10, 2891–2896. <https://doi.org/10.1021/ct500193u>.
- Martyna, G.J., Tobias, D.J., and Klein, M.L. (1994). Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* 101, 4177–4189. <https://doi.org/10.1063/1.467468>.
- Matsumoto, T., Morimoto, Y., Shibata, N., Kinebuchi, T., Shimamoto, N., Tsukihara, T., and Yasuoka, N. (2000). Roles of functional loops and the CD-terminal segment of a single-stranded DNA binding protein elucidated by X-ray structure analysis. *J. Biochem.* 127, 329–335. <https://doi.org/10.1093/oxfordjournals.jbchem.a022611>.
- McGuffee, S.R., and Elcock, A.H. (2010). Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm. *PLoS Comput. Biol.* 6, e1000694. <https://doi.org/10.1371/journal.pcbi.1000694>.
- Meinhold, D.W., and Wright, P.E. (2011). Measurement of protein unfolding/refolding kinetics and structural characterization of hidden intermediates by NMR relaxation dispersion. *Proc. Natl. Acad. Sci. U. S. A.* 108, 9078–9083. <https://doi.org/10.1073/pnas.1105682108>.
- Meyer, H., Biermann, O., Faller, R., Reith, D., and Müller-Plathe, F. (2000). Coarse graining of nonbonded inter-particle potentials using automatic simplex optimization to fit structural properties. *J. Chem. Phys.* 113, 6264–6275. <https://doi.org/10.1063/1.1308542>.
- Mishra, G., and Levy, Y. (2015). Molecular determinants of the interactions between proteins and ssDNA. *Proc. Natl. Acad. Sci. U. S. A.*

- 112, 5033–5038. <https://doi.org/10.1073/pnas.1416355112>.
- Miyamoto, S., and Kollman, P.A. (1992). Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* 13, 952–962. <https://doi.org/10.1002/jcc.540130805>.
- Neuman, K.C., and Nagy, A. (2008). Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat. Methods* 5, 491–505. <https://doi.org/10.1038/nmeth.1218>.
- Nguyen, H.D., and Brooks, C.L. (2008). Generalized structural polymorphism in self-assembled viral particles. *Nano Lett.* 8, 4574–4581. <https://doi.org/10.1021/nl802828v>.
- Ortega, A., Amorós, D., and de La Torre, J.G. (2011). Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys. J.* 101, 892–898. <https://doi.org/10.1016/j.bpj.2011.06.046>.
- Ouldrige, T.E., Louis, A.A., and Doye, J.P.K. (2011). Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model. *J. Chem. Phys.* 134, 085101. <https://doi.org/10.1063/1.3552946>.
- Pal, A., and Levy, Y. (2019). Structure, stability and specificity of the binding of ssDNA and ssRNA with proteins. *Plos Comput. Biol.* 15, e1006768. <https://doi.org/10.1371/journal.pcbi.1006768>.
- Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26, 1781–1802. <https://doi.org/10.1002/jcc.20289>.
- Raghunathan, S., Kozlov, A.G., Lohman, T.M., and Waksman, G. (2000). Structure of the DNA binding domain of *E. coli* SSB bound to ssDNA. *Nat. Struct. Mol. Biol.* 7, 648–652. <https://doi.org/10.1038/77943>.
- Raghunathan, S., Ricard, C.S., Lohman, T.M., and Waksman, G. (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. U. S. A.* 94, 6652–6657. <https://doi.org/10.1073/pnas.94.13.6652>.
- Reith, D., Pütz, M., and Müller-Plathe, F. (2003). Deriving effective mesoscale potentials from atomistic simulations. *J. Comput. Chem.* 24, 1624–1636. <https://doi.org/10.1002/jcc.10307>.
- Robertson, J.W.F., Rodrigues, C.G., Stanford, V.M., Rubinson, K.A., Kraslinikov, O.V., and Kasianowicz, J.J. (2007). Single-molecule mass spectrometry in solution using a solitary nanopore. *Proc. Natl. Acad. Sci. U. S. A.* 104, 8207–8211. <https://doi.org/10.1073/pnas.0611085104>.
- Roy, R., Hohng, S., and Ha, T. (2008). A practical guide to single-molecule FRET. *Nat. Methods* 5, 507–516. <https://doi.org/10.1038/nmeth.1208>.
- Roy, R., Kozlov, A.G., Lohman, T.M., and Ha, T. (2007). Dynamic structural rearrangements between DNA binding modes of *E. coli* SSB protein. *J. Mol. Biol.* 369, 1244–1257. <https://doi.org/10.1016/j.jmb.2007.03.079>.
- Roy, R., Kozlov, A.G., Lohman, T.M., and Ha, T. (2009). SSB protein diffusion on single-stranded DNA stimulates RecA filament formation. *Nature* 461, 1092–1097. <https://doi.org/10.1038/nature08442>.
- Saleh, O.A., McIntosh, D.B., Pincus, P., and Ribbeck, N. (2009). Nonlinear low-force elasticity of single-stranded DNA molecules. *Phys. Rev. Lett.* 102, 068301. <https://doi.org/10.1103/PhysRevLett.102.068301>.
- Savvides, S.N., Raghunathan, S., Fütterer, K., Kozlov, A.G., Lohman, T.M., and Waksman, G. (2004). The C-terminal domain of full-length *E. coli* SSB is disordered even when bound to DNA. *Protein Sci.* 13, 1942–1947. <https://doi.org/10.1110/ps.04661904>.
- Shea, J.-E., Best, R.B., and Mittal, J. (2021). Physics-based computational and theoretical approaches to intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* 67, 219–225. <https://doi.org/10.1016/j.sbi.2020.12.012>.
- Sim, A.Y.L., Lipfert, J., Herschlag, D., and Doniach, S. (2012). Salt dependence of the radius of gyration and flexibility of single-stranded DNA in solution probed by small-angle x-ray scattering. *Phys. Rev. E* 86, 021901–021905. <https://doi.org/10.1103/PhysRevE.86.021901>.
- Singharoy, A., Maffeo, C., Delgado-Magnero, K.H., Swainsbury, D.J., Sener, M., Kleinekathöfer, U., Vant, J.W., Nguyen, J., Hitchcock, A., Isralewitz, B., et al. (2019). Atoms to phenotypes: molecular design principles of cellular energy metabolism. *Cell* 179, 1098–1111.e23. <https://doi.org/10.1016/j.cell.2019.10.021>.
- Smith, S.B., Cui, Y., and Bustamante, C. (1996). Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science* 271, 795–799. <https://doi.org/10.1126/science.271.5250.795>.
- Stein, I.H., Steinhauer, C., and Tinnefeld, P. (2011). Single-molecule four-color FRET visualizes energy-transfer paths on DNA origami. *J. Am. Chem. Soc.* 133, 4193–4195. <https://doi.org/10.1021/ja1105464>.
- Su, X.-C., Wang, Y., Yagi, H., Shishmarev, D., Mason, C.E., Smith, P.J., Vandevenne, M., Dixon, N.E., and Otting, G. (2014). Bound or free: interaction of the c-terminal domain of *escherichia coli* single-stranded DNA-binding protein (SSB) with the tetrameric core of SSB. *Biochemistry* 53, 1925–1934. <https://doi.org/10.1021/bi5001867>.
- Suksombat, S., Khafizov, R., Kozlov, A.G., Lohman, T.M., and Chempla, Y.R. (2015). Structural dynamics of *e. coli* single-stranded DNA binding protein reveal DNA wrapping and unwrapping pathways. *eLife* 4, 1–23. <https://doi.org/10.7554/eLife.08193>.
- Šulc, P., Romano, F., Ouldrige, T.E., Rovigatti, L., Doye, J.P.K., and Louis, A.A. (2012). Sequence-dependent thermodynamics of a coarse-grained DNA model. *J. Chem. Phys.* 137, 135101. <https://doi.org/10.1063/1.4754132>.
- Sun, T., Minhas, V., Korolev, N., Mirzoev, A., Lyubartsev, A.P., and Nordenskiöld, L. (2021). Bottom-up coarse-grained modeling of DNA. *Front. Mol. Biosci.* 8, 645527. <https://doi.org/10.3389/fmolb.2021.645527>.
- Takada, S., Kanada, R., Tan, C., Terakawa, T., Li, W., and Kenzaki, H. (2015). Modeling structural dynamics of biomolecular complexes by coarse-grained molecular simulations. *Acc. Chem. Res.* 48, 3026–3035. <https://doi.org/10.1021/acs.accounts.5b00338>.
- Tinoco Jr, I., and Bustamante, C. (2002). The effect of force on thermodynamics and kinetics of single molecule reactions. *Biophys. Chem.* 101–102, 513–533. [https://doi.org/10.1016/s0301-4622\(02\)00177-1](https://doi.org/10.1016/s0301-4622(02)00177-1).
- Ugarte La Torre, D., and Takada, S. (2020). Coarse-grained implicit solvent lipid force field with a compatible resolution to the C α protein representation. *J. Chem. Phys.* 153, 205101. <https://doi.org/10.1063/5.0026342>.
- Wagner, J.W., Dannenhoffer-Lafage, T., Jin, J., and Voth, G.A. (2017). Extending the range and physical accuracy of coarse-grained models: Order parameter dependent interactions. *J. Chem. Phys.* 147, 044113. <https://doi.org/10.1063/1.4995946>.
- Wu, H., Wolynes, P.G., and Papoian, G.A. (2018). AWSEM-IDP: a coarse-grained force field for intrinsically disordered proteins. *J. Phys. Chem. B* 122, 11115–11125. <https://doi.org/10.1021/acs.jpcc.8b05791>.
- Yang, O., and Ha, T. (2018). Single-molecule studies of ssDNA-binding proteins exchange. *Methods Enzymol.* 600, 463–477. <https://doi.org/10.1016/bs.mie.2017.11.017>.
- Yoo, J., and Aksimentiev, A. (2012). Improved parametrization of Li, Na, K, and Mg ions for all-atom molecular dynamics simulations of nucleic acid systems. *J. Phys. Chem. Lett.* 3, 45–50. <https://doi.org/10.1021/jz201501a>.
- Zheng, W., Borgia, A., Buholzer, K., Grishaev, A., Schuler, B., and Best, R.B. (2016). Probing the action of chemical denaturant on an intrinsically disordered protein by simulation and experiment. *J. Am. Chem. Soc.* 138, 11702–11713. <https://doi.org/10.1021/jacs.6b05443>.
- Zhou, R., Kozlov, A.G., Roy, R., Zhang, J., Korolev, S., Lohman, T.M., and Ha, T. (2011). SSB functions as a sliding platform that migrates on DNA via reptation. *Cell* 146, 222–232. <https://doi.org/10.1016/j.cell.2011.06.036>.
- Zhou, R., Schlierf, M., and Ha, T. (2010). Force-fluorescence spectroscopy at the single-molecule level. *Methods Enzymol.* 475, 405–426. First edition chapter 6. [https://doi.org/10.1016/S0076-6879\(10\)75016-3](https://doi.org/10.1016/S0076-6879(10)75016-3).
- Zwanzig, R. (1988). Diffusion in a rough potential. *Proc. Natl. Acad. Sci. U S A.* 85, 2029–2030. <https://doi.org/10.1073/pnas.85.7.2029>.

STAR★METHODS

KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|--------------------------------|---|------------|
| Deposited data | | |
| SSB PDB | RCSB.org | 1EYG |
| Software and algorithms | | |
| NAMD | ks.uiuc.edu/Research/namd | 2.12 |
| ARBD | bionano.physics.illinois.edu/arbdcg | 1.0a |
| Simulation configuration files | gitlab.engr.illinois.edu/tbgl/pubdata/cg-ssb | 1d5e4db |

RESOURCE AVAILABILITY

Lead contact

Further information and requests for software and and trajectories should be directed to and will be fulfilled by the lead contact, Aleksei Aksimentiev (aksiment@illinois.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- All data reported in this paper will be shared by the [lead contact](#) upon request.
- All original code has been deposited at <https://gitlab.engr.illinois.edu/tbgl/pubdata/cg-ssb> and is publicly available as of the date of publication. DOIs are listed in the [key resources table](#).
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

METHOD DETAILS

General simulation protocols

All atomistic MD simulations were performed using the program NAMD (Phillips et al., 2005), the TIP3P model of water (Jorgensen et al., 1983), standard parameters for ions (Beglov and Roux, 1994), periodic boundary conditions, particle-mesh Ewald (PME) full electrostatics with a PME grid density of about 1 Å per grid point. In the simulations of SSB surrounded by a solution of dT₅ fragments, the CHARMM36 force field (Hart et al., 2012; Denning et al., 2011; MacKerell and Banavali, 2000; Foloppe and MacKerell Jr., 2000) was employed with NBFIX corrections applied to ion–nucleic acid interactions and to the arginine and lysine amine–nucleic acid interactions (Yoo and Aksimentiev, 2012). Van der Waals and short-range electrostatic energies were calculated using a smooth (7–8 Å) cutoff. Integration was performed using 2–2–6 fs multiple timestepping (Batcho et al., 2001). To enable 2-fs timestepping for bonded interactions, water bonds (and angles) and non-water covalent bonds with hydrogens were held rigid using the SETTLE (Miyamoto and Kollman, 1992) and RATTLE (Andersen, 1983) algorithms, respectively.

Steric clashes that were introduced during the assembly of the system were removed through minimization using a conjugate gradient method. Subsequent simulations were performed in the NPT ensemble; a temperature of 291 K was maintained by applying Langevin forces (Brünger, 1992) to all non-hydrogen atoms (1 ps⁻¹ damping coefficient) and a pressure of 1 bar was maintained by Nose'-Hoover Langevin piston pressure control (Martyna et al., 1994). The temperature was held constant using a Langevin thermostat (Phillips et al., 2005) applied to all non-hydrogen atoms; the Langevin damping constant was set to 0.1 ps⁻¹.

All CG simulations were performed using the ARBD simulation package (Comer and Aksimentiev, 2012). The two-beads-per-nucleotide model of ssDNA previously developed (Maffeo et al., 2014) using the IBI protocol was used in all CG simulations, including the tabulated potentials, masses, damping coefficients and all-atom mapping scheme. The temperature in ARBD was set to 291 K. The integration of the equations

of motion was performed using a geodesic BAOAB (Leimkuhler and Matthews, 2016) integrator that includes a Langevin thermostat. Linear interpolation of all bonded and nonbonded potentials was employed, including three-dimensional grid-based potentials used to represent SSB proteins. The protocols of the rigid body dynamics are described in the last section of STAR Methods.

Generation of target data for ssDNA–SSB interaction parametrization

As described in the main text, we performed atomic simulations to generate target distributions for the CG parametrization procedure. To enhance throughput, ten atomic models were simulated, each containing 31 dT₅ fragments distributed around an SSB molecule in a small cubic volume of 100 mM NaCl electrolyte, 10.6 nm on each side. The composition of the models was identical but the initial configurations were unique in each system because of the random initial arrangement of the DNA fragments. During the simulations, which lasted over 14 μ s in total, the backbone atoms of residues forming α -helices and β -sheets were harmonically restrained about their initial coordinates (1 kcal mol⁻¹ Å⁻² spring constant). From the CG representation of the all-atom trajectories, the three-dimensional, position-dependent target density was computed for each DNA bead type and averaged over the symmetry axes of the SSB homotetramer. The target density calculations were carried out over a 2 Å grid using the volmap plugin of VMD (Humphrey et al., 1996) and a 4 Å-width Gaussian for each DNA bead.

Refinement of CG SSB potentials

First, we employed a naive application of the IBI (Reith et al., 2003) procedure. An initial CG potential for each bead type was obtained by applying Boltzmann inversion to the corresponding target density, divided by a factor of ten to crudely account for the potential being applied to each bead of the same type within a CG DNA fragment. Following that, a procedure was applied to each potential to ensure that, on average, no force is applied at the boundary and to eliminate large forces due to extremely low bead density where the SSB core prevents access. Specifically, the potential was adjusted by uniformly subtracting the average value at the edge of the grid and then by setting the potential to 20 $k_B T$ at voxels with a bead density < 10⁻⁶ beads/Å³. Initial trial potentials were obtained by applying Boltzmann inversion to one-tenth of the target density of that bead type.

The IBI (Reith et al., 2003) procedure was then applied with each iteration consisting of a CG simulation, the calculation of a CG density map, followed by an update of the CG potential. For each iteration, CG simulations of six systems lasted a total of \sim 80 ns and the trajectories allowed CG density maps to be computed using the same protocol that yielded the target densities. The SSB potential for each bead type was updated as described in the main text according to the expression,

$$\Delta u = - a k_B T \ln \frac{\rho_{\text{target}} + \rho_0}{\rho_{\text{CG}} + \rho_0},$$

where the scaling factor $a = 0.1$ ensured gradual convergence, and $\rho_0 = 10^{-6}$ beads/Å³ protected against numerical instability. The correction to the potential was spatially smoothed using a Gaussian filter with a width of 2 Å. The potential after adding the update was adjusted using the same procedure described for the initial trial potential, ensuring no average force at the boundary and no large forces at or near the SSB core. All software for manipulating the density and potential grids was developed in-house using C++ and Python.

At the end of the naive IBI procedure, the resulting model was found to bind too much DNA under tension compared with experiment, Figure S1. Speculating that the excess binding could be caused in part by attractive regions of the potentials that extended far (> 1 nm) from the protein surface, which counters physical intuition, we repeated the IBI refinement as described above, except for a few modifications described below. First, a higher resolution (0.5 Å) was used to recompute the target density grids with a slightly smaller 3 Å Gaussian width for each bead. The same resolution was used to compute the average density of non-hydrogen SSB atoms observed during the all-atom MD simulations. To produce a mask that restricts the IBI adjustments to region near the protein surface, the SSB atom density was smoothed using the Voltool plugin of VMD with a 6 Å Gaussian width, which was then averaged over the SSB symmetry axes. The resulting grid values were then linearly transformed and clamped so that values ≤ 0.001 became zero, and values ≥ 0.025 became one, providing our mask, which extended about 6 Å from the protein surface. The mask was used to scale the initial IBI potentials as well as all IBI updates, which were otherwise performed as described above, except the prefactor of 0.1 was changed to one. Finally, a Debye–Hückel correction was

applied to the DNA P beads as the difference of two Yukawa potentials to better match the 182 mM electrolyte condition of the all-atom systems in comparison to the 100 mM condition that the DNA model was developed under. After 189 IBI iterations, the resulting potentials were used in production simulations. The Debye–Hückel correction was not adopted in the production simulations.

Motion of the SSB particle

Upon completion of the refinement procedure, the potentials reproduced the all-atom distribution of dT₅ fragments bound to SSB. In subsequent simulations, we modeled the SSB protein in ARBD as a rigid body having a mass and a moment of inertia determined from the all-atom model. At each timestep, the force vector applied to each ssDNA bead by the grid was multiplied by -1 to obtain the force \vec{F} applied by the bead on the SSB; the corresponding torque was obtained as a cross product of the vector connecting the center of the SSB with the DNA bead and the force vector. A Langevin force and torque (Gordon et al., 2009) was added to the total force and torque to mimic the effect of the solvent with damping coefficients corresponding to the diagonal elements of the diffusion tensor obtained from Hydropro (Ortega et al., 2011). Finally, these forces and torques were used to perform integration of the rigid body's equation of motion to update position and orientation of the SSB molecule using a symplectic algorithm (Dullweber et al., 1997).

In the simulations that featured multiple SSB molecules, forces were evaluated between all pairs of SSB proteins using a protocol (Singharoy et al., 2019) that clusters atoms types into three groups according to their CHARMM36 (Huang and MacKerell Jr, 2013) Lennard-Jones (LJ) parameters, constructs approximate 3D LJ potential maps (1 Å resolution), an electrostatic potential map from APBS (Holst and Saied, 1996; Jurrus et al., 2018), and maps with the distribution of the corresponding LJ and electrostatic charges. If a voxel of the charge map of one SSB overlaps with the potential map of the other, a force is computed for that voxel from the product of the charge value with the negative gradient of the potential determined using a linear interpolation of the potential map; the torque is determined as a cross product between the vector leading from the origin of the SSB protein to the charge voxel and the force on that voxel. The force and the torque is scaled by -1 and, for the torque, transformed to apply forces and torques consistent with the Newton third law to the second SSB. Due to the relatively slow motion of the SSB proteins, the forces and torques were recalculated every 20 simulation steps.