

Rapid Long-distance Migration of RPA on Single Stranded DNA Occurs Through Intersegmental Transfer Utilizing Multivalent Interactions

Sushil Pangeni^{1,2}, Gargi Biswas^{3,†}, Vikas Kaushik^{4,†}, Sahiti Kuppa⁴, Olivia Yang⁵, Chang-Ting Lin⁵, Garima Mishra⁶, Yaakov Levy³, Edwin Antony^{4,*} and Taekjip Ha^{1,2,7,8,*}

1 - TC Jenkins Department of Biophysics, Johns Hopkins University, Baltimore, MD, USA

2 - Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA, USA

3 - Department of Chemical and Structural Biology, Weizmann Institute of Science, Rehovot, Israel

4 - Department of Biochemistry and Molecular Biology, St. Louis University School of Medicine, St. Louis, MO, USA

5 - Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, MD, USA

6 - Department of Physics, Ashoka University, Sonapat, Haryana, India

7 - Department of Pediatrics, Harvard Medical School, Boston, MA, USA

8 - Howard Hughes Medical Institute, Boston, MA, USA

Correspondence to Edwin Antony and Taekjip Ha: edwin.antony@health.slu.edu (E. Antony), taekjip.ha@childrens.harvard.edu (T. Ha)

<https://doi.org/10.1016/j.jmb.2024.168491>

Edited by Gerhard Hummer

Abstract

Replication Protein A (RPA) is a single stranded DNA (ssDNA) binding protein that coordinates diverse DNA metabolic processes including DNA replication, repair, and recombination. RPA is a heterotrimeric protein with six functional oligosaccharide/oligonucleotide (OB) domains and flexible linkers. Flexibility enables RPA to adopt multiple configurations and is thought to modulate its function. Here, using single molecule confocal fluorescence microscopy combined with optical tweezers and coarse-grained molecular dynamics simulations, we investigated the diffusional migration of single RPA molecules on ssDNA under tension. The diffusion coefficient D is the highest (20,000 nucleotides²/s) at 3 pN tension and in 100 mM KCl and markedly decreases when tension or salt concentration increases. We attribute the tension effect to intersegmental transfer which is hindered by DNA stretching and the salt effect to an increase in binding site size and interaction energy of RPA-ssDNA. Our integrative study allowed us to estimate the size and frequency of intersegmental transfer events that occur through transient bridging of distant sites on DNA by multiple binding sites on RPA. Interestingly, deletion of RPA trimeric core still allowed significant ssDNA binding although the reduced contact area made RPA 15-fold more mobile. Finally, we characterized the effect of RPA crowding on RPA migration. These findings reveal how the high affinity RPA-ssDNA interactions are remodeled to yield access, a key step in several DNA metabolic processes.

© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Introduction

Replication protein A was first reported as an important for *simian vacuolating virus 40* (SV40)

replication in human cell extracts,^{1,2} and is highly conserved in eukaryotes.^{3–6} RPA serves as a housekeeping single stranded DNA (ssDNA) binding protein that maintains ssDNA stability during

numerous biological processes where ssDNA is transiently exposed.^{5,6} RPA performs or influences many processes including protection of ssDNA from nucleases,⁷ resolving secondary structures like hairpins,^{8,9} G- quadruplexes,^{10,11} and R-Loops.¹² If unprotected by RPA, spontaneous mutations can occur to the exposed ssDNA.^{13,14} RPA also marks the site of DNA damage by binding to ssDNA. RPA is required for the recruitment of ATR (ataxia-telangiectasia mutated- and Rad3-related) to damaged sites and ATR-mediated damage checkpoint activation.^{15,16} RPA can interact with more than three dozen RPA interacting proteins that are essential for DNA metabolism and genome maintenance.^{16,17}

RPA binds to ssDNA with very high affinity (dissociation constant $K_D < 10^{-10}$ M) but must be removed or redistributed from ssDNA to make room for other DNA processing proteins. For example, RPA can help recruit helicases like HELB to ssDNA which in turn can facilitate the removal of RPA during DNA replication and recombination.¹⁸ Other mediator proteins such as Rad52 remodel select domains of RPA and gain access to the ssDNA.¹⁹ RPA is a heterotrimeric protein composed of RPA70, RPA32 and RPA14 subunits where the numbers denote the apparent molecular weight of each subunit. The subunits are composed of six oligosaccharide/oligonucleotide binding (OB) fold domains that are connected by flexible linkers of varying lengths. RPA70 houses OB-F, OB-A, OB-B and OB-C. OB-D resides in RPA32 along with a winged-helix (wh) domain. RPA14 harbors the final OB domain (OB-E) (Figure 1A). Four of these domains (OB-A, B, C, and D) primarily coordinate DNA binding and are termed DNA binding domains (DBDs: A, B, C and D). OB-F and the wh-domain coordinate protein-protein interactions in RPA70 and RPA32, respectively and are also termed protein-interaction domain (PIDs: PID^{70N} and PID^{32C}).^{20–26} RPA is held together as a constitutive heterotrimer through interactions between DBD-C (RPA70), DBD-D (RPA32), and OB-E (RPA14). While OB-E shows structural changes in the presence of ssDNA, no direct evidence for DNA binding has been observed.²⁴ The DBDs and PIDs are connected by flexible linkers (Figure 1B), contributing to the structural flexibility of RPA.²⁵

The prevailing models for RPA-ssDNA interactions posit that the DBDs function as dynamic units with a complex array of binding, dissociation, and remodeling properties for each. Thus, RPA interacting proteins can selectively remodel one or more DBDs and PIDs and gain access to the DNA. In addition, the intrinsic flexibility of the linkers enables the DBDs and PIDs to be arranged in multiple configurations depending on the DNA substrate it encounters and the associated biological function. Recent work suggests that the overall configurational

properties of RPA can be envisioned from the perspective of two halves of RPA: a dynamic and less-dynamic half.²⁴ OB-F, DBD-A and DBD-B along with the long F-A linker (all in RPA70) are more dynamic and are positioned off the ssDNA when multiple RPA molecules bind to DNA at high density. In contrast, DBD-C (RPA70), DBD-D (RPA32), and OB-E (RPA14), which form the trimerization core (Tri-C), constitute a less-dynamic half that binds more stably to ssDNA.^{7,13,17,24–30,24}

An established approach to describe the DNA binding properties of RPA invokes binding mode transitions. Binding modes are defined by the occluded site size for RPA, defined as the number of nucleotides required to saturate the DNA binding site under a given condition. As each DBD can dynamically associate or dissociate from DNA, accessibility of the DNA to other interacting proteins will be dictated by the binding mode of RPA. Affinity based measurements made with isolated DBDs suggested that RPA initially bind ssDNA with a low binding mode of 8–12 nt involving mainly DBD-A and DBD-B. Binding of additional DBDs (DBD-C and DBD-D) switches RPA to a higher binding mode of 17–35 nt that involves all DBDs.^{8,25,28,31–34} A switch or transition between the binding modes can be modulated in solution by changing the ionic strength.^{8,33} The low binding mode is observed at ~100 mM NaCl and the higher binding mode is achieved at ~600 mM NaCl.⁸ Assignment of which DBDs drive the two binding modes came from affinity measurements made with isolated DBDs. However, recent work shows that these affinity measurements are more complex when considered from the context of full-length RPA where DBDs-A and B are more dynamic.³⁵

Diffusion of RPA on ssDNA is another feature that has been experimentally observed.⁸ Yet, how RPA diffuses on long ssDNA and the underlying mechanistic basis are poorly understood. Single molecule measurements are ideally suited for highly dynamic proteins such as RPA.^{25,34,36–41} Single molecule FRET (smFRET)⁴² has been used to study RPA-ssDNA interactions and revealed the dynamic nature of RPA on ssDNA where each DBD can undergo microscopic association and dissociation while RPA is bound to the ssDNA^{34,35,37,43} and RNA.³⁷ Magnetic tweezers have been used to measure the force regulated dynamics of RPA at a DNA replication fork-like structure.⁴⁴ Single molecule DNA curtains have also been used to study RPA dynamics,^{19,36,41,45–47} and uncovered that RPA can undergo facilitated exchange in the presence of free excess RPA, Rad51, or SSB in solution.^{41,47}

Diffusional migration of RPA on ssDNA was discovered also using single molecule methods.^{8,29} Nguyen *et al* have shown diffusion of RPA along short (≤ 120 nt) strands of DNA using smFRET⁸ and found that as the length of DNA

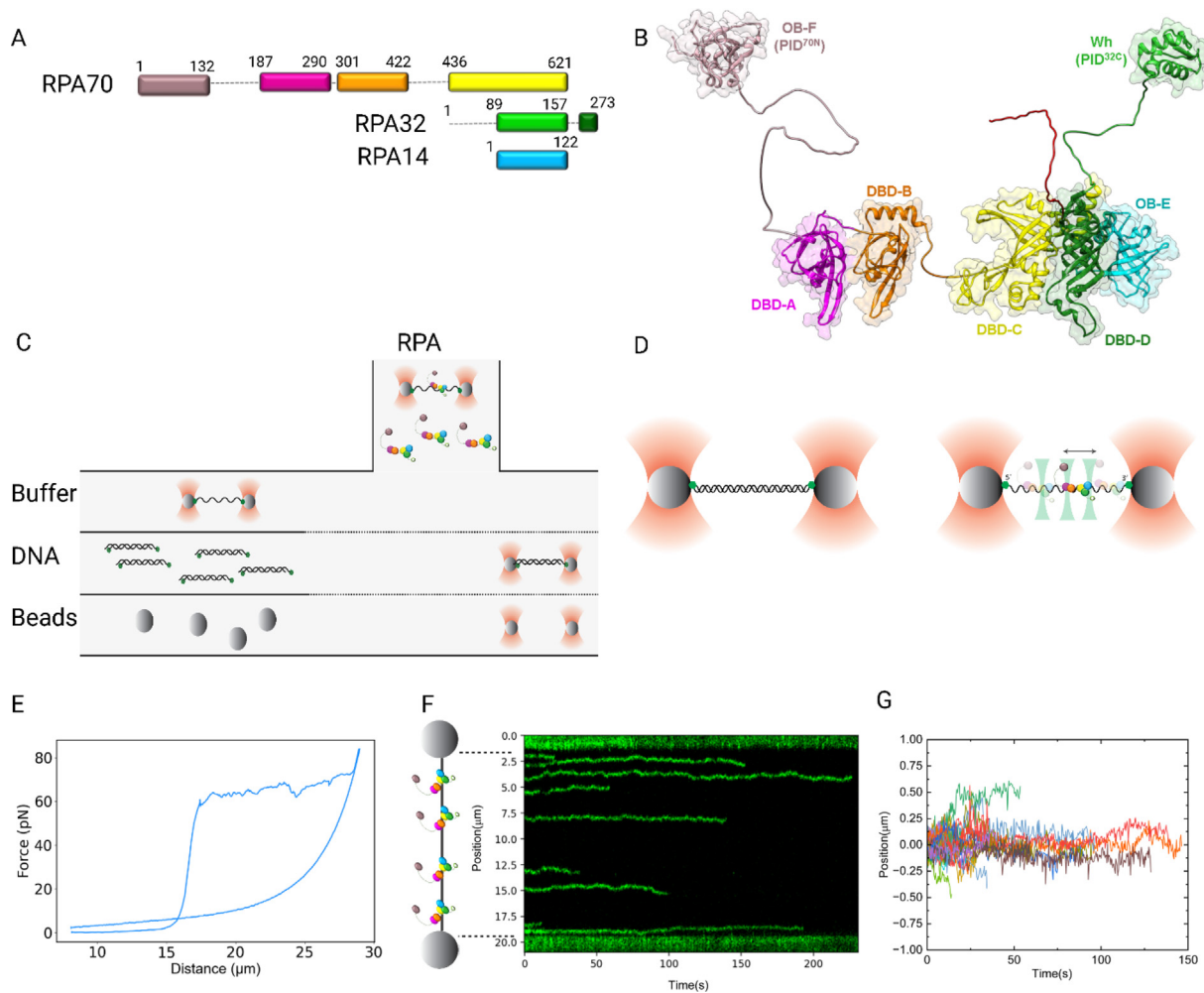


Figure 1. Experimental Scheme: A) Schematic of RPA subunits and their respective domains. B) A model for RPA generated from the structures of the individual OB-domains. Intervening linkers were generated using AlphaFold. C) Laminar flow system of commercial optical trap(C-Trap). D)Optical trap illustration. E) Force extension curve of 48.5 kbp lambda DNA to generate the ssDNA. F) Kymograph showing RPA diffusion along ssDNA. RPA DBD D diffusion along ss-DNA under 8 pN tension on ssDNA (10 mM HEPES pH7.5,2mM MgCl₂,100 mM KCl). G) Diffusion traces of many RPA molecules as in F. Diffusion start position was adjusted to position zero. Colors for each subunit of RPA match for A and B and the following illustrations.

was increased, the diffusion coefficient, D , can be determined more precisely.⁸ Although previous studies have demonstrated changes in binding mode with varying ionic strength, the effect of binding mode on the movement of RPA on DNA is unknown.⁸ *E. coli* single stranded DNA binding protein (SSB), which plays similar roles in bacteria, has been shown to have significantly higher diffusion coefficients on long stretches of DNA compared to short DNA, potentially due to intersegmental transfer between distant sites on DNA.^{48,49} Studying RPA movement on long stretches of ssDNA will help us to examine the effect of the presence of other proteins including other RPA interacting proteins, and the role of intersegmental transfer. Although the effect of DNA tension on SSB diffusion

on DNA has been studied previously,⁴⁸ diffusive property of yeast RPA on ssDNA longer than 120 oligonucleotides length, or under mechanical tension, has not been reported yet. A recent study by Mersch *et al* have investigated diffusion of human RPA (hRPA) on 20 kbp long ssDNA and reported that the hRPA diffusion does not depend significantly on tension or salt concentration.⁵⁰

Diffusion of RPA and SSB has been extensively studied using computation tools.^{48,5,51,52} We have previously shown force dependent SSB diffusion on ssDNA using coarse grained molecular simulation that aligned with experimental data.⁴⁸ While previous computational work on RPA diffusion at the small length scale of the ssDNA less than 150 nucleotides provided the molecular details of

protein interactions with DNA during diffusion,^{51–54} they were unable to examine long-range jumps between DNA segments.

In this study, we aimed to understand the diffusion of *Saccharomyces cerevisiae* RPA under various salt conditions and under tension applied across ssDNA. RPA diffuses along ssDNA with one dimensional random walk and its diffusion is reduced with higher ionic strength or with increasing tension. Our findings help us understand how single stranded binding proteins may migrate on single stranded DNA.

Results

RPA can diffuse over a long distance on DNA before dissociation

Saccharomyces cerevisiae RPA (RPA) was labeled in one of the DBDs with an MB453 fluorophore.⁵⁵ The fluorophore was positioned either on DBD-A (RPA-DBD-A^{MB453}) or DBD-D (RPA-DBD-D^{MB453}) using non-canonical amino acids as described.³⁵ To investigate the mechanism of RPA diffusion, we performed confocal fluorescence imaging of fluorescently-labeled RPA on ssDNA under mechanical tension using the LUMICKS C-Trap, an instrument that combines dual optical traps with confocal scanning fluorescence microscopy with single fluorophore sensitivity.⁵⁶ To create long ssDNA, lambda phage DNA was tethered to two streptavidin beads using biotin on the same strand and was mechanically denatured. Formation of ~50 kb ssDNA was confirmed by fitting to a Freely Jointed Chain Model (FJC).⁵⁷

Movement of RPA along the ssDNA is recorded as a function of time and visualized as a kymograph (Figure 1F). At time zero, each spot corresponds to one RPA molecule bound to the ssDNA ([RPA] = 10 pM). Over time, RPA moves back and forth in an apparent diffusive movement, consistent with one dimensional random walk. The displacement (X) relative to the initial position is converted into nucleotides using the extension of 48.6 kb ssDNA at a given force (Figure 1E). RPA can diffuse over a long distance on DNA before dissociation or photobleaching (Figure 1F). Mean square displacement (MSD) plotted vs. time t showed an initial linear increase which was fitted using $MSD = 2 \cdot D \cdot t$ (where D is diffusion coefficient and t is time) to estimate D (Figure S1).⁵⁸

RPA diffusion is affected by DNA tension and salt concentration

We examined the effect of DNA tension on RPA diffusion using two fluorescent versions of RPA labeled at either DBD-A (RPA-DBD-A^{MB453}) or DBD-D (RPA-DBD-D^{MB453}). We first calculated the diffusion coefficient expressed in $\mu\text{m}^2/\text{s}$, as measured in the laboratory frame along the direction of force. As shown in Figure 2A and 2B

we observed little dependence on force. Next, we accounted for a change in the base per rise (BPR) between nucleotides with force. For example, at 16.2 μm extension at 5 pN, BPR would be 3.4 Å but at higher force with extension of 20 μm , BPR would be 4.1 Å. When we converted the observed displacement into nucleotides then calculated the diffusion coefficient in units of nt^2/s , we observed a large force dependence (Figure 2C, 2D). D values were similar for both RPA-DBD-A^{MB453} and RPA-DBD-D^{MB453} and decreased with increasing tension on ssDNA (Figure 2D, 2E). For example, for RPA-DBD-D^{MB453}, D decreased ~10 fold from 22,000 nt^2/s at 3 pN to 2,300 nt^2/s at 28 pN (100 mM KCl) (Figure 2D), reminiscent of slower diffusive movements with increasing tension observed for *E. coli* SSB⁴⁸. These data also show that conjugation of the MB543 on either DBD-A or DBD-D does not differentially influence their diffusion properties (compare Figure 2D and Figure 2E), consistent with their similarities observed in the kinetics and thermodynamics of their ssDNA binding properties.^{24,35} Thus, we used RPA-DBD-D^{MB453} for further single molecule analysis and will refer to it as RPA^f.

At a low force (3 pN) and high salt (600 mM KCl), RPA^f had a diffusion coefficient of 17,000 nt^2/s (Figure 2D) which is approximately 5 times larger than what was previously reported for human RPA at 600 mM KCl (2,800 nt^2/s) in the absence of tension on 120 nt ssDNA.⁸ At this diffusion coefficient, an RPA can migrate over about 200 nt in 1 second. At forces higher than 18 pN, the diffusion coefficient for yeast RPA^f decreases to ~3000 nt^2/s at 600 mM KCl which is comparable to that of human RPA diffusion on short DNA at zero force (Figure 2D, 2E). One possible explanation is that for long ssDNA, RPA may use intersegmental transfer between distant sites, a process that would be hindered when the DNA is stretched mechanically.

Next, we looked at the effect of ionic strength. RPA became less mobile with increasing salt concentration for all force values tested (Figure 2D, 2E). For example, D for RPA^f decreased by a factor of three when [KCl] increased from 100 mM to 600 mM (Figure 2D, 2E). Salt concentration influences the way various combinations of DBDs associate with ssDNA, with higher salt promoting binding of more DBDs.^{8,29,33} If more DBDs bind to the DNA, RPA may diffuse slower due to an enhanced number of contacts.

RPA switches from the 22 nt binding mode to the 30 nt binding mode when salt concentration increases from 100 mM to 600 mM, but the binding mode stays at 22 nt below 100 mM down to 1 mM.^{8,33} Therefore, we performed the experiment under low salt concentrations to further test if the effect of salt concentration on RPA diffusion is indeed due to changes in binding mode. We did not see a significant change in D between 1 mM

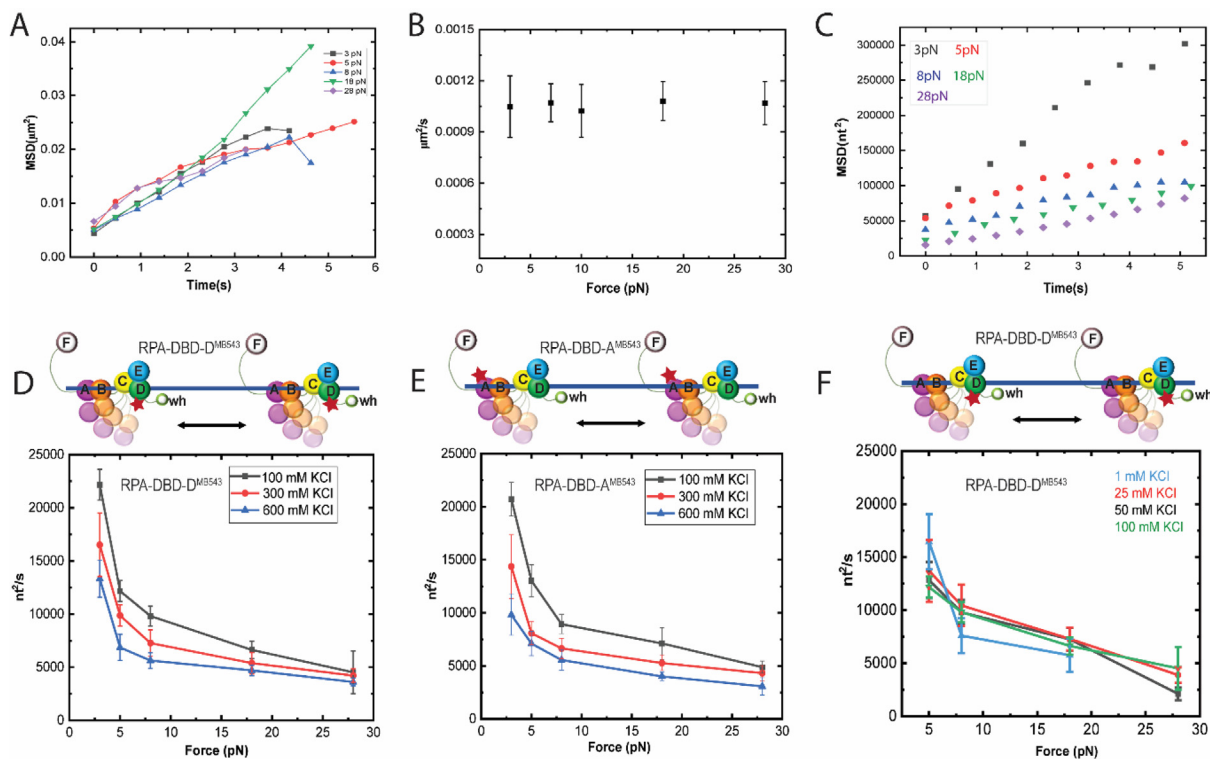


Figure 2. Apparent Diffusion Coefficient of *Saccharomyces cerevisiae* RPA (RPA) studied as a function of tension on DNA filament and change in salt conditions. A) Mean Squared Displacement (MSD) in micrometer vs time for various forces 3–28 pN. Black 3 pN, Red 5 pN, Blue 8 pN, Green 18 pN, and Purple 28 pN. B) Diffusion coefficient measured as micrometer. C) Mean Squared Displacement (MSD) in nucleotides vs time for various forces 3–28pN. Black 3 pN, red 5pN, blue 8 pN, green 18 pN, and purple 28 pN. D) Diffusion coefficient of RPA labelled with MB453 dye at DNA binding domain D (RPA-DBD-D^{MB543}) as indicated by red star. E) Diffusion coefficient of full-length RPA protein labelled with MB453 dye at DNA binding domain A (RPA-DBD-A^{MB543}) as indicated by red star. F) Diffusion coefficient of RPA-DBD-D^{MB543} at the lower salt concentrations. Experiments were performed at three salt concentrations ranging from 100 – 600 mM KCl while tension was varied from 3 pN to 28 pN. All the error bars are Standard Error of Mean (SEM). Protein concentration is 10 pM with experimental buffer being 100 mM KCl (varied), 10 mM HEPES pH7.5, 2 mM MgCl₂.

and 100 mM KCl (Figure 2F). Therefore, our data are consistent with the model where RPA is more mobile in the 22 nt binding mode compared to the 30 nt binding mode.

Coarse grained simulations of RPA diffusion

To understand the effect of force and salt on the diffusion of RPA on ssDNA in molecular details, we used coarse-grained molecular simulations. We started by using different starting conformations of a fixed length ssDNA, where the end-to-end distances were different. Each of the systems were simulated keeping the ends of the DNA fixed in space, while the other parts of the DNA and the protein were allowed to move freely. This represents the experimental scheme for optical trap where traps were fixed at a specific distance with ssDNA held between two traps. Tuning of the end-to-end distance effectively control the force on the DNA where higher end-to-

end distance represents higher force on ssDNA (Figure S2).

A representative simulation of RPA on 860 nt ssDNA (0.5 pN force, 30 mM salt) is shown as five snapshots of RPA-bound ssDNA's conformation. In snapshot 2, RPA bridges two distant sites that are separated by 168 nt (Figure 3A). Around this time point, the index of the nucleotide closest to the protein center of mass (COM) changes rapidly back and forth by $\Delta nt = 168$, eventually resulting in a large jump in position by that amount (Figure S3). We proposed that the looped state, with the loop size of 168 nt, is an intermediate for intersegmental transfer. Snapshot 4 shows another looped state of a 51 nt loop, resulting in a smaller scale intersegmental transfer. The size of intersegmental transfer is widely distributed and became progressively smaller as the force increased (Figure 3B). At high force, the DNA is more stretched, which hinders the formation of a

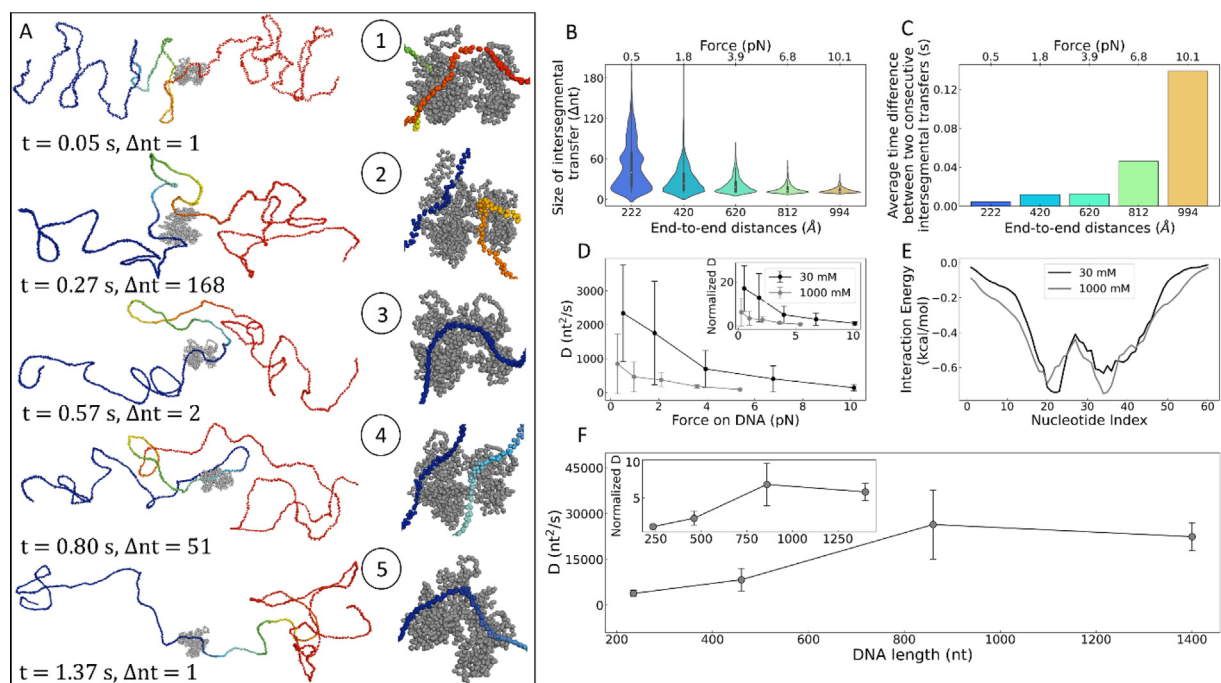


Figure 3. Diffusion of RPA on ssDNA through intersegmental transfer. A) Five representative snapshots of RPA-ssDNA binding sampled variation of the nucleotide index (nt) closest to the centre-of-mass (COM) of the RPA protein is shown throughout a representative simulation trajectory for ssDNA of length of 860 nucleotides ssDNA with a force of 0.5 pN on the ssDNA (which corresponds to an end-to-end distance of 817 Å. B-C) Characteristics of the intersegmental transfer events of RPA when diffusion along ssDNA (235 nt long) under force applied on its ends. These simulations were performed at 30 mM. B) Violin plots of the length of ssDNA, Δnt , skipped in intersegmental transfer of RPA when diffusing along ssDNA at five different forces applied on its two ends [the applied force (shown in the upper X axis) results with end-to-end distance varying between 222 and 994 Å, (shown in the lower X axis)]. The simulations at different force applied on the two ssDNA ends are shown with different color. C) The kinetics of the intersegmental transfer is indicated by the average time difference between any two consecutive intersegmental transfer events (having length $\Delta nt > 10$) for the five systems at 30 mM. D-E) Variation of diffusion coefficient at different force and salt concentrations. D) Variations of diffusion coefficient for diffusion of RPA along ssDNA for different strength of force applied on its two ends, at low (black) and high (grey) salt concentrations. The inset shows the diffusion coefficients normalized to the value at high applied force. The error bars represent the standard deviation in the diffusion coefficient obtained from 30 identical simulations. E) The potential energy of the interaction between RPA and ssDNA of 60 nucleotides as a function of the ssDNA nucleotide index for low and high salt concentrations. A ssDNA of 60 nucleotides have been utilized to focus on the protein-DNA interaction at the binding site as per the RPA-ssDNA crystal structure. F) Variation of RPA diffusion coefficient as a function of ssDNA length. The inset shows the normalized diffusion coefficient relative to the diffusion coefficients for the shortest simulated ssDNA. The error bars represent the standard deviation of the diffusion coefficients obtained from 30 independent trajectories.

large ssDNA loop thereby reducing the probability of large intersegmental transfer events. Indeed, the wait time between any two consecutive intersegmental transfers with $\Delta nt > 10$ dramatically increases as the force increase (Figure 3C). A similar force dependence was observed also for high salt simulations (1000 mM) (Figure S4D, S4E and S4F) with the main difference being that in higher salt, the intersegmental transfer is often shorter in jump size and their frequency is lower, consistent with lower mobility in high salt we observed experimentally (Figure 2). The diffusion coefficient D calculated from the simulated trajectories also showed salt and force dependence (Figure 3D). In

high salt, a 15-fold increase of the force resulted in a ~ 6 folds decrease in the D value, analogous to a 9-fold increase in force causing a ~ 4 -fold decrease in D observed experimentally (Figure 2, 3D). At the low salt simulation, D was 3–4 times greater than in the high salt simulation, similarly to the fractional change in D determined experimentally (Figures 2, 3D).

To gain insights into the molecular origin of the salt effect, we calculated the binding energy vs nucleotide position at the binding interface RPA-ssDNA for 235 nt long ssDNA (Figure 3E). For most of the regions of the protein-ssDNA interface, the interaction is more stable (i.e., more negative interaction energy) at higher salt

concentration. The tighter interaction between RPA and ssDNA at high salt is consistent with the higher binding mode. We do not completely understand the nature of contacts between the domains of RPA as salt could also influence such interactions and add another layer of complexity in deciphering these transitions.

The ssDNA used in the experimental study is much longer than the 235 nt ssDNA used in the simulations thus far. To explore the effect of the ssDNA length on the characteristics of intersegmental transfer, we performed simulations with different ssDNA lengths (235, 460, 860 and 1400 nucleotides). As the DNA length increases from 235 nucleotides to 860 nucleotides, there is an increase in the diffusion coefficient D (Figure 3F), likely due to higher probability of large size intersegmental transfer events (Figure 3B, S4D). Beyond 860 nucleotides, the change in the D value was negligible (Figure 3F). The saturation of the intersegmental transfer events is also reflected in the size distribution and occurrence probability of intersegmental transfer (Figure S5B, S5C). This result suggests that for a better comparison to experimental results we should use simulations obtained using the longer ssDNA (860 nt and 1400 nt). They showed a 10-fold increase in the D values (Figure 3G, inset plot) compared to 235 nt long DNA, suggesting that D values with the short ssDNA should be multiplied by factor of by 10 to compare to the experiments. Indeed, we obtained a diffusion coefficient ~ 2500 nt^2/s with 235 nucleotides DNA (Figure 3F), while in experiments for a 48.5 kbp ssDNA the measured D value obtained was ~ 10 times higher (Figure 2D, 2E).

The 'dynamic half' of RPA consisting of OB-F, DBD-A, and DBD-B has higher mobility

Diffusion of RPA is proposed to occur through dynamic binding and rearrangements of the individual DBDs.^{24,35} These domains were recently classified to function as two halves: "FAB", composed of OB-F, DBD-A, and DBD-B, is considered 'dynamic' based on their higher propensity to detach from ssDNA. The "trimerization core" (Tri-C), composed of DBD-C, DBD-D and DBD-E, is considered 'less-dynamic', contributing more to the stability of RPA-ssDNA interactions.^{24,30,35} FAB, lacking the Tri-C, does not show significant DNA dependent protection from proteolysis, likely because its interaction with DNA is not as stable as the full-length protein, frequently exposing the protein surface.²⁴ The occlusion size of FAB on ssDNA is proposed to be ~ 8 nt.^{8,25} To test the contributions of these two halves to RPA diffusion, we measured the diffusion properties of FAB. FAB is highly diffusive with diffusion coefficient of 210,000 nt^2/s at 5 pN which is about 15 times larger compared to the full-length RPA constructs under same tension (Figure 4). FAB binding to DNA is more

transient, with its average residence time on DNA of 18 seconds before dissociation compared to the less diffusive WT RPA with an average residence time of one minute (Figure 4C, 4D, 4E). The larger diffusion coefficient for FAB is likely due to fewer nucleotide contacts with ssDNA, also causing fast dissociation. FAB also showed similar salt and tension dependence of diffusion coefficient as in the full-length proteins (Figure 4A, Figure 2). Unfortunately, such experiments with TriC (DBD-C, RPA32, and RPA 14) alone were not possible due to the instability of the complex when isolated.

RPA diffusion under RPA crowding

Next, we looked at the effect of excess unlabeled RPA WT (full length RPA with no label in any DBD) in the diffusional properties of RPA^f (Figure 5A) in order to mimic the crowding effect provided by additional RPA proteins bound to the same DNA^{35,59}. At unlabeled RPA WT to RPA^f ratio of 15 or higher, D decreased measurably. D dropped from 13,000 nt^2/s in the absence of excess RPA to 900 nt^2/s when 100 times excess unlabeled RPA was available, probably because the available free space is constrained (Figure 5). As the concentration of RPA was increased, the effective distance traversed by a single RPA decreased; for example, RPA diffused up to 200 nts for WT to RPA^f ratio of 30 compared to up to ~ 2000 nts for WT to RPA^f ratio of 1 likely due to collisions with other RPA molecules consistent with the decrease in diffusion coefficient (Figure 5B, 5C).

Discussion

Using optical trap combined with single molecule confocal fluorescence microscopy, we directly observed the diffusional movements of RPA along mechanically stretched ssDNA in real time. The diffusion coefficient is 21,000 nt^2/s at 3 pN at physiological salt concentration of 100 mM KCl. Coarse grained simulations provide evidence for the importance of intersegmental transfer.

We observed a remarkable decrease in diffusion coefficient with increasing salt concentration. RPA can adopt different ssDNA binding modes on ssDNA as per the ionic condition. Nguyen *et al*⁶ showed that as salt concentration increases, the occlusion size, the number of nucleotide contacts made by human RPA on ssDNA, increases. Kumaran *et al* proposed a similar binding mode switch for yeast RPA, a lower DNA binding mode of 18 nt in low salt and a higher DNA binding mode of 28 nt in high salt.³³ The decrease in D we observed when KCl concentration increased from 100 mM to 600 mM could be a result of such switch in binding mode.^{8,33} The resulting depletion of DNA binding sites on the RPA surface would hinder the formation of a transient bridge to a distant site on DNA, reducing RPA mobility. Consistent with this

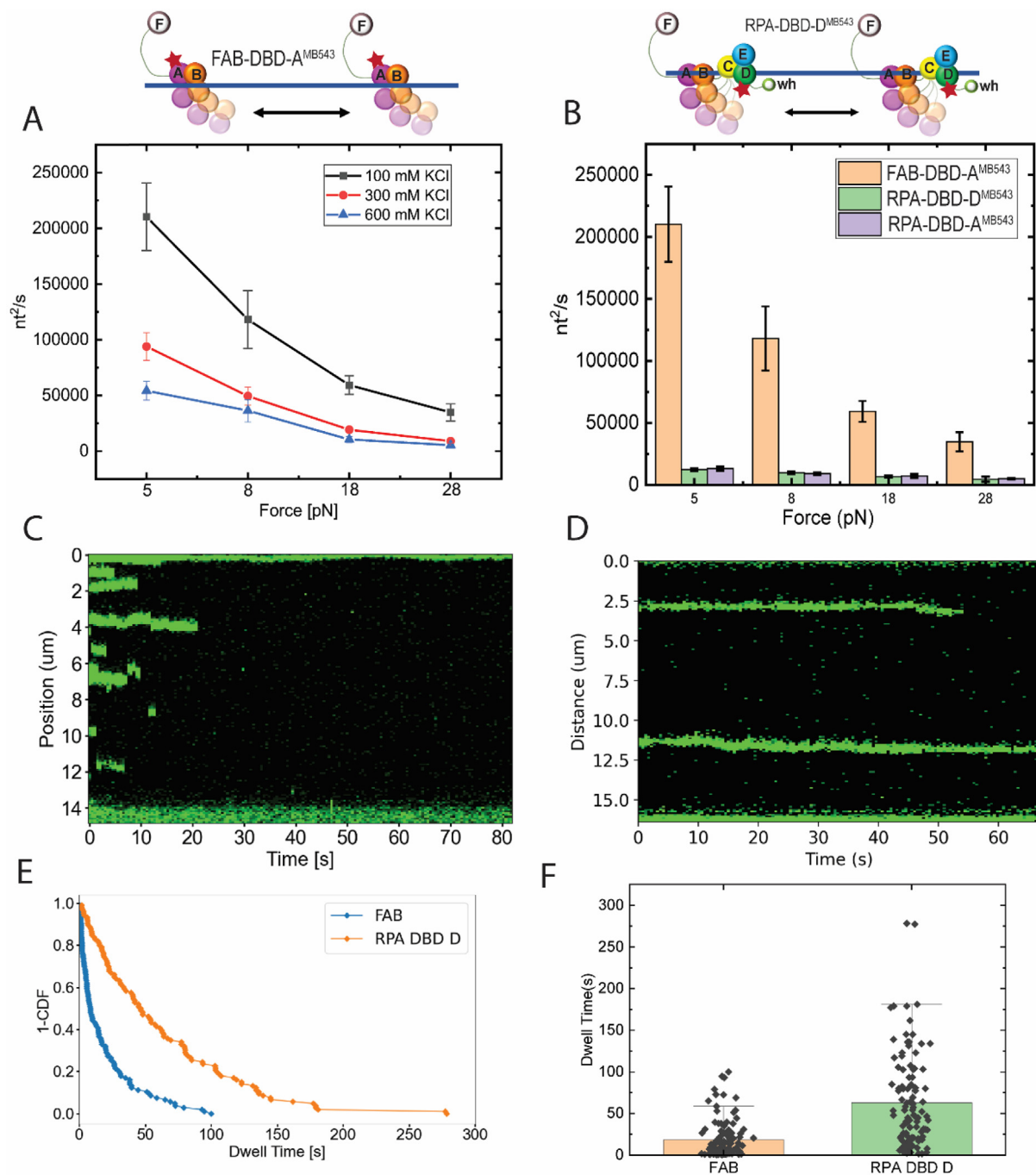


Figure 4. Apparent Diffusion Coefficient of *saccharomyces cerevisiae* RPA (RPA) studied as a function of tension on DNA filament and change in salt conditions. A) Diffusion of truncated protein with only FAB domains from RPA70. B) Comparison of diffusion coefficient of four different variants of RPA studied. RPA DBD A and RPA DBD D data are taken as in Figure 2 while FAB data is taken from Figure 4 A. The Schematic of RPA shown on top is a representative RPA diffusion where RPA is labeled at DNA binding domain D (RPA-DBD-D^{MB543}) as in Figure 2D.C) A typical kymograph for FAB. FAB is very diffusive in short range and dissociates from DNA. D) A typical kymograph of full-length RPA. E) A 1-CDF plot depicting how fast FAB dissociates from DNA compared to full length RPA. F) Dwell time of RPA full length and FAB on ssDNA. Salt conditions ranged from 100 mM KCL to 600 mM KCl while tension ranged from 3pN to 28pN with error bars being SEM. FAB concentration was 1 nM. Other buffer components were 10 mM HEPES pH7.5, 2mM MgCl₂.

explanation, we see no difference in D in salt concentration ranging from 1 mM to 100 mM

(Figure 2F), the range for which Nguyen *et al* did not see any change in binding mode.

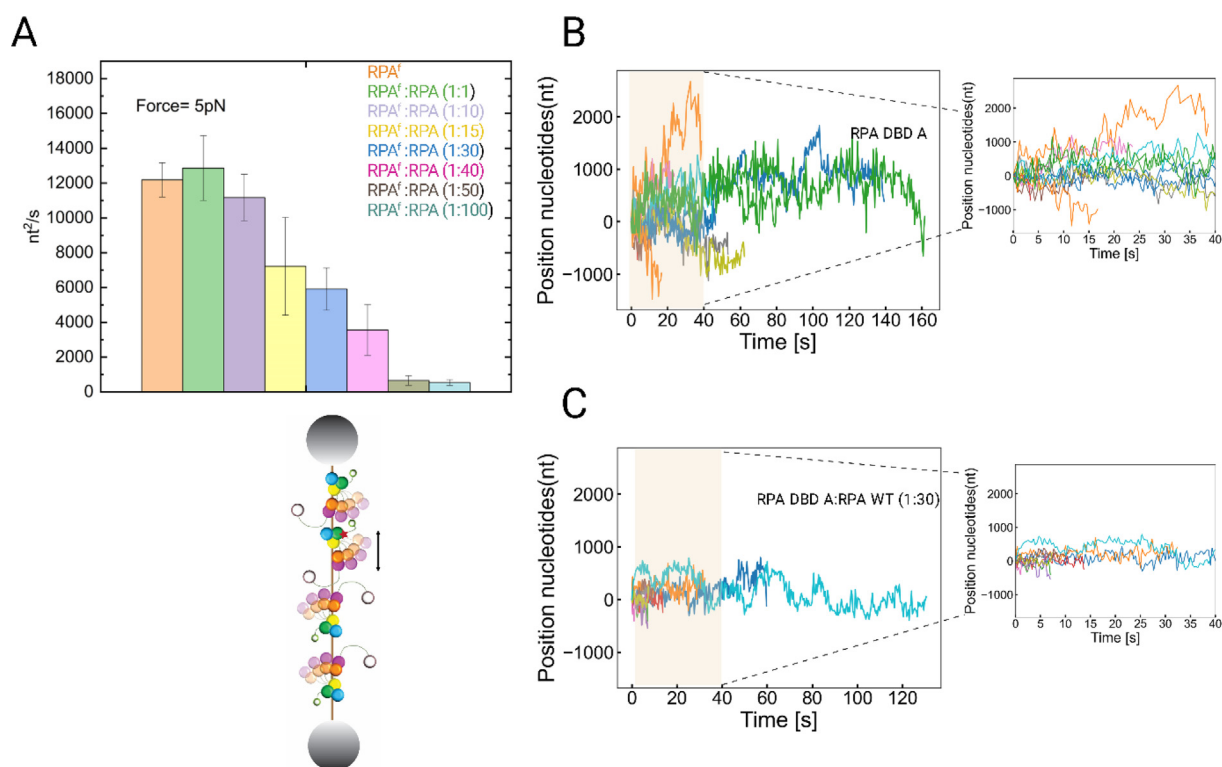


Figure 5. RPA diffusion under the influence of free unlabeled RPA and chaperone Rtt105. A) Apparent diffusion coefficient of RPA DBD D (RPA labeled at DBD D as indicated by red star) under the presence of various concentrations of unlabeled RPA WT. Desired ratio of proteins were applied to the same microfluidic channel by incubating together for 1–2 min on ice. Error bars are SEM. B) Distance traversed by RPA under the condition where there are no collisions with other RPA molecules. C) Distance traversed by RPA^f where WT to RPA^f ratio of 30 was used.

Persistence length of ssDNA decreases with increasing salt,⁶⁰ making ssDNA more flexible through screening electrostatic repulsion between negatively charged backbone phosphates.^{60,61} Higher flexibility may make it possible to obtain tighter contact with DNA binding sites spread over the RPA surface. Indeed, our molecular dynamics simulations based on the crystal structure of RPA-ssDNA complex²⁶ suggest that the higher salt concentration induces stronger interaction (Figure 3E). Overall, the salt effect on diffusion of RPA could be governed by the binding mode of RPA, ssDNA flexibility and RPA-ssDNA interaction strength. Additionally, recent cross-linking and hydrogen–deuterium mass spectrometry experiments reveal extensive interactions between the DBDs and PIDs of RPA.^{24,62,63} Thus, additional salt-induced changes to such intra-RPA contacts could also influence the diffusion properties.

An earlier report of force dependent diffusion of ssDNA binding protein SSB⁴⁸ showed that the diffusion coefficient (D) dropped by a factor of 6 as the force was increased from 3 pN to 5 pN.⁴⁸ Because SSB dissociated from ssDNA at ~ 8 pN, the range of tension examined was small.^{48,64–66} The DNA construct used for the SSB study could not form

secondary structure as ssDNA were generated using rolling circle amplifications of a short sequence devoid of secondary structure. In the current study of yeast RPA, while diffusion was slower than SSB, it was higher than the values reported for human RPA by Nguyen *et al.*⁸: i.e. 17,000 nt²/s in our study vs 2800 nt²/s in theirs, both in 600 mM KCl. They also used 120 nt oligos with no secondary structure possibilities.⁸ Our experiments performed using lambda DNA would allow some secondary structures to form at lower forces, potentially making intersegmental transfer more effective.

The size of intersegmental transfer also decreases with increasing force as we show through molecular simulations (Figure 3B). In the absence of force, DBDs of RPA can bind distant sites at the same time and effectively stabilize the condensed DNA.⁶¹ Recent study on human RPA did not observe a significant influence of force on its diffusion.⁵⁰ Our study on yeast RPA showed a clear force dependence on the diffusion of the protein both in experiment and simulations. Although human RPA and yeast RPA are structurally similar, experimental observations have suggested that yeast RPA is more dynamic than human RPA.^{35,37,67,68} This could also explain the

differential effect of force on diffusion of human RPA and yeast RPA.⁵² We note that the diffusion coefficient expressed in $\mu\text{m}^2/\text{s}$, that is in the distance measured in the laboratory frame along the direction of force, showed little dependence on force (Figure 2B). It is when we convert that distance to the number of nucleotides that we see a strong force dependence (Figure 2D).

Another possible explanation for force dependent decrease in diffusion can be the reptation mechanism. Reptation involves DNA bulge formation on the RPA surface, storing 1–7 nt of ssDNA involving aromatic interactions between protein and ssDNA. Bulge dissolution *i.e.* release of 1–7 nt in the bulge leads to RPA diffusion on ssDNA.^{5,51,52} There has not been direct experimental evidence for this mechanism in the case of RPA, but bulge formation has been reported as a possible diffusion mechanism for SSB in fluorescence-force spectroscopy analysis⁶⁹ and also in molecular simulation studies of RPA.⁵² As we increase the tension on ssDNA, the possibility of RPA forming a bulge will be diminished as ssDNA is stretched. Conversely, under low tension where relaxed DNA is present, RPA can diffuse faster with bulge formation increasing the diffusion.

We found that when high density of RPA is present on ssDNA, 1D random motion of individual RPA molecules becomes restricted, leading to an overall decrease in diffusion coefficient (Figure 5). Whether the crowing effect mainly acts through inhibiting local migration in small, a few nt steps, as is expected for reptation, or by hindering larger scale movements such as intersegmental transfer is presently unknown.

Our study suggests that full length RPA is less dynamic than FAB, a truncated RPA that lacks the trimeric core. FAB was 15 times more mobile than full length RPA (Figure 2D, 4A). Further studies are needed to determine if the higher mobility is due to an increase in size or frequency of intersegmental transfer.

Methods

Protein Purification: Replication Protein A was purified and labeled as detailed in.^{36,55}

Optical trap combined with confocal microscopy

48.5 kbp Lambda DNA construct was prepared with three biotins on either end of DNA.⁵⁷ Lambda DNA was purchased from Roche Inc. Short oligos to anneal to the sticky ends were purchased from Integrated DNA Technologies Inc. DNA were stored in TE buffer (10 mM Tris-HCl, pH 8.0 and 0.1 mM EDTA). Optical trap experiments were performed using a commercial dual optical trap combined with confocal microscopy and microfluidics [C-trap] from Lumicks BV Inc. The microfluidic chamber was pas-

sivated at the start of every experiment day using the protocol: 0.1 % BSA (Sigma) was flowed at 1 bar for 5 minutes then at 0.4 bar pressure for 25 minutes, followed by a 10-minute rinse with Milli-Q water at 1 bar pressure for 5 min and 0.4 bar pressure for 5 minutes, followed by 0.5 % Pluronic F-127 flowed at 1 bar pressure for 5 min then 0.4 bar pressure for 25-minutes, followed 1 bar pressure for 5 min and 0.4 bar pressure for 5 minutes with 1X PBS.

Streptavidin coated polystyrene particle beads of average size 4.8 μM [0.5 % w/v] (Spherotech Inc.) were diluted 1:250 in 1X Phosphate Saline Buffer (PBS) that contains 137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4 , and 2 mM KH_2PO_4 (10X buffer purchased from Thermofisher Scientific USA) and 1–2 nM of DNA were made in 1X PBS. DNA was captured between two streptavidin beads and mechanically denatured by moving one bead to the overstretched region to create ssDNA. ssDNA was confirmed by fitting force-distance (FD) curve to Freely Jointed Chain model [FJC] (contour length 48.5 kbp / 27.160 μm ; persistence length 0.9 nm; stretch modulus 1000 pN) in real time.^{70,71} DNA was held for 5 s in the fully ssDNA state then returned to a required tension on ssDNA position for the fluorescence experiments.

RPA-DBD-D-MB543, RPA-DBD-A-MB543, and RPA WT were stored at storage buffer (30 mM HEPES pH 7.8, 200 mM KCl, 0.02 % Tween-20, 10 % glycerol, and 0.2 mM EDTA pH 8.0) and diluted to 1 nM with experimental buffer (30 mM HEPES pH 7.8, 100 mM KCl, 6 % Glycerol, 5 mM MgCl_2) just before the experiment then 10 pM final concentration of RPA was used during the experiments. Imaging buffer 0.8 % (w/v) dextrose, 165 U/mL glucose oxidase, 2170 U/mL catalase, and 2–3 mM Trolox was used to increase the fluorescence lifetime of the fluorophores. Imaging settings were 2–3 ms exposure time (per pixel), red excitation 638 nm, and green excitation 561 nm. Pixel size was kept constant at 100 nm. Excitation wavelengths were Green (575–625 nm), Red (670–730 nm). Laser power was selected to maintain the 5 micro-watt power at the objective. Note that our experiments required pixel time of 3 ms to capture a single RPA molecule. Reducing the pixel time would often require higher RPA concentration of protein where less bright RPA could not be captured but will still contribute to the diffusion of RPA being observed.

DNA oligos

Lambda DNA (Purchased from Roche).

Oligo 1: GGG CGG CGA CCT GGA CAA.

oligo 2: AGG TCG CCG CCC TTT TTT T/iBiodT/ T/iBiodT/T/iBiodT/.

oligo 3: iBiodT/T/iBiodT/ T/iBiodT/T TTT TTT AGA GTA CTG TAC GAT CTA GCA TCA ATC TTG TCC.

Data analysis and sharing

Data was analyzed using custom python script detailed in Pylake API from Lumicks.⁷² Python version 3.8.6 is used. Origin pro 2021b was also used. Some illustrations were made using Biorender.com. Similarly, all the codes are available in Github (<https://github.com/spangeni/RPA-Diffusion.git>). All the data has been included in the manuscript. Raw data will be made available upon reasonable request.

Simulation methods

1. Generating protein coarse-grained model

To investigate the interaction of RPA with ssDNA at the molecular level, a coarse-grained model was used. In this model, each of the protein residues are represented by two beads placed at the C_α and C_β positions. For the charged amino acids (K, R, H, D and E), respective charges were placed at the C_β position. Simulations have been performed using a native topology-based model which included nonspecific electrostatics (non-native) interactions and used Lennard-Jones potential to represent native contact interactions. For protein modelling, an approach similar to that described in Refs.^{51,53} have been utilized, where the internal energy of the protein is designated by E_{prot} , where,

$$E_{prot}(\Gamma, \Gamma_0) = E_{prot}^{Bond} + E_{prot}^{Angle} + E_{prot}^{Dihedral} + E_{prot}^{NativeContacts} + E_{prot}^{Electrostatics} + E_{prot}^{Repulsions}$$

where Γ denotes a particular conformation during the simulation trajectory and Γ_0 denotes the native conformation. The potential energy of any conformation during the simulation trajectory consists of the following terms,

$$E_{prot}(\Gamma, \Gamma_0) = \sum_{bonds} K_{bonds} (b_{ij} - b_{ij}^0)^2 + \sum_{angles} K_{angles} (\theta_{ijk} - \theta_{ijk}^0)^2 + \sum_{dihedrals} K_{dihedrals} [1 - \cos(\phi_{ijkl} - \phi_{ijkl}^0) - \cos(3\Phi_{ijkl} - \Phi_{ijkl}^0)] + \sum_{i \neq j} K_{contacts} [5 \left(\frac{A_{ij}}{r_{ij}}\right)^{12} - 6 \left(\frac{A_{ij}}{r_{ij}}\right)^{10}] + \sum_{i,j} K_{electrostatics} B(\kappa) \frac{q_i q_j \exp^{-\kappa r}}{\epsilon_r r_{ij}} + \sum_{i \neq j} K_{contacts} \left(\frac{C_{ij}}{r_{ij}}\right)^{12}$$

where K_{bonds} , K_{angles} , $K_{dihedrals}$, $K_{contacts}$, and $K_{repulsion} = 100 \text{ Kcal mol}^{-1} \text{ \AA}^{-2}$, 20 Kcal mol^{-1} , 1 Kcal mol^{-1} , 1 Kcal mol^{-1} , and 1 Kcal mol^{-1} , respectively. Additionally, b_{ij} is the distance between bonded beads i and j in \AA , and b_{ij}^0 is the optimal distance between bonded beads i and j in angstroms; θ_{ijk} is the angle between subsequently bonded beads i - k in radians, θ_{ijk}^0 is the optimal angle between subsequently bonded beads i - k in radians, ϕ_{ijkl} is the dihedral angle between subsequently bonded backbone beads i - l in

radians, and ϕ_{ijkl}^0 is the optimal dihedral angle between subsequently bonded backbone beads i - l in radians. A_{ij} is the optimal distance between beads i and j in contact in angstroms, and r_{ij} is the distance between beads i and j in angstroms in a given conformation along the trajectory. Optimal values were calculated from the atomic coordinates of the relevant Protein Data Bank (PDB) structure file. C_{ij} is the sum of radii for any two beads not forming a native contact; the repulsion radius of the backbone bead (C_α) was 1.9 \AA , and the radius of the side-chain bead (C_β) was set to 1.5 \AA . The electrostatic interactions were modeled by the Debye-Hückel potential, and parameters from previous studies^{51,53} have been used. The simulations were run at a temperature where protein remains folded and fluctuate around the native state.

2. Generating ssDNA coarse-grained model

In the coarse-grained model of the single-stranded DNA (ssDNA)⁷³ each of the nucleotide was represented by three beads, positioned at the geometric centre of phosphate (P), sugar (S) and base (B); while the S and B bead were neutrally charged, the P bead beard a negative charge. The internal potential energy of the ssDNA was designated by E_{ssDNA} , which is given by the following expression,

$$E_{ssDNA}(\Gamma, \Gamma_0) = E_{ssDNA}^{Bond} + E_{ssDNA}^{Angle} + E_{ssDNA}^{Dihedral} + E_{ssDNA}^{BasePairing} + E_{ssDNA}^{Stacking} + E_{ssDNA}^{Repulsions}$$

In this expression, the term E_{ssDNA}^{Bond} represents the contribution from the backbone consisting of covalently linked phosphate (P) and sugar (S), also, covalently linked sugar (S) and base (B). E_{ssDNA}^{Angle} ssDNA is the potential for bond angles that is applied between all the following three neighboring beads: $(P_i-S_i-B_i)$, $(B_i-S_i-P_{i+1})$, $(P_i-S_i-P_{i+1})$, and $(S_i-P_{i+1}-S_{i+1})$ with $K_{angles} = 20$. $E_{ssDNA}^{Dihedral}$ consists of two kinds of dihedral potentials. One is the potential of the dihedral angle formed by the beads B_i , S_i , S_{i+1} , and B_{i+1} with $K_{Dihedral} = 0.5$. The other is the potential from the dihedral angle formed by four consecutive phosphate beads $(P_i, P_{i+1}, P_{i+2}$ and $P_{i+3})$; and this $K_{Dihedral}$ is either 0.7 or 0 which makes the persistence length of the DNA 32 and 17 \AA respectively.

$E_{ssDNA}^{BasePairing}$ arises from the contribution of the interaction between complementary base-pairs, which is absent in case of ssDNA, so $E_{ssDNA}^{BasePairing} = 0$. $E_{ssDNA}^{Stacking}$ arises from the interaction between two consecutive bases, B_i and B_{i+1} and can be expressed as $\epsilon_{B-B} [5 \left(\frac{r_{0ij}}{r_{ij}}\right)^{12} - 6 \left(\frac{r_{0ij}}{r_{ij}}\right)^{10}]$, where r_{0ij} is the typical distance⁷⁴ between two consecutive bases (i.e., 3.6 \AA) and strength of interaction ϵ_{B-B} depends upon the identity of bases (i.e., whether A, T, C or G). The expression for the repulsion term, $E_{ssDNA}^{Repulsions}$ is similar to the

repulsion term in case of proteins, only the radii of the base, phosphate, and sugar are 1.5, 3.7, and 3.7 Å respectively and the repulsion is applied between two beads which satisfies the condition $l - |j| > 6$.

3. Modelling protein-ssDNA interactions

The interaction energy between protein and single stranded DNA is given by,

$$E_{ssDNA-Protein}(\Gamma, \Gamma_0) = E_{ssDNA-Protein}^{Electrostatics} + E_{ssDNA-Protein}^{Aromatic} + E_{ssDNA-ssDNA}^{Repulsions}$$

The repulsion is applied between all beads of the protein and all beads of the ssDNA. The electrostatic interactions acting between all the charged beads in the system are modeled by the Debye-Hückel potential. These interactions are nonspecific, and the phosphate groups of the ssDNA can interact with any charged residue in the protein. The aromatic interaction, like base stacking, is modeled by the L-J potential, with a base-aromatic amino acid interaction strength of ϵ_{B-AA} . The parameter $\epsilon_{B-AA} = 0$ for all the nonaromatic amino acids, and $\epsilon_{B-AA} > 0$, only for aromatic amino acids (W, F, Y, and H). The strength of interaction ϵ_{B-AA} for aromatic amino acids depends on their identities.

4. Modelling force

Next, to understand the effect of force on ssDNA in molecular details, we started with the RPA-ssDNA coarse-grained model obtained from the original crystal structure (PDB ID; 4GNX). Since, in the crystal structure only 29 nucleotide long ssDNA was present, we modelled the ssDNA as longer one (initially 235 nucleotides) by keeping the RPA-bound fragment of 29 nucleotides in the middle and identical to the original crystal structure. Now, after generating the coarse-grained model, pulling force was applied to both ends of the DNA to obtain different initial structures having different end-to-end distances representing different forces. To that end, five different end-to-end distances were utilized, namely, 222, 420, 620, 812 and 994 Å. This end-to-end distances were converted to respective force values using the FJC model,⁷¹ which resulted in forces 0.5, 1.8, 3.9, 6.8 and 10.1 pN respectively on the ssDNA.

Now, with these five initial structures, molecular dynamics simulation was performed keeping both ends of the ssDNA fixed in space while the protein and the rest of the DNA was allowed to move freely. The temperature was 0.4 (arbitrary units) and the salt concentration was kept at 0.065 mM during the simulation, and the simulations were run 5×10^7 simulation steps, saving the coordinates at each 1000 steps. The simulation temperature is chosen in such a way that the protein remains folded during the simulation and fluctuates around its native state. Importantly, the salt concentration mentioned before, regulates the electrostatics interaction between the protein and

DNA, but since in our model, no electrostatic interaction is present between the DNA beads, this salt concentration cannot confer flexibility or rigidity to the DNA molecule which is sensitive to the salt concentration.

5. Modelling effect of salt

To model the effect of salt in the RPA-ssDNA interactions, we modified the persistence length of the DNA that is salt dependent. From previous studies,^{53,60} it is well known that, as the salt concentration is increased, the ssDNA becomes more flexible and its persistence length decreases. Also, in a similar coarse-grained model of ssDNA, the persistence length was modified by changing the dihedral constant K_ϕ .⁵³ In our default model, the dihedral constant was 0.7, which resulted in a persistence length of 32 Å.⁵³ As per the experimental studies by Murphy et al, this persistence length of the ssDNA is equivalent to that in the 30 mM salt.^{53,60} Thus, to increase the DNA flexibility, which would correspond to higher salt concentration, we modified the strength of the dihedral constant K_ϕ to 0, which resulted in a persistence length of 17 Å, the flexibility of which would correspond to that in the salt concentration 1000 mM. It is noteworthy that, since the force-extension relationship according to the FJC model explicitly depends upon the persistence length of the DNA, the conversion of the end-to-end distances to force between the two ends depends on the salt concentration. Therefore, the end-to-end distances of 222, 420, 620, 812 and 994 Å correspond to forces of 0.5, 1.8, 3.9, 6.8 and 10.1 pN for salt concentration of 30 mM ($K_\phi = 0.7$) and to values of 0.3, 1.0, 2.1, 3.6 and 5.4 pN for salt concentration of 1000 mM (i.e., $K_\phi = 0$).

6. Effect of DNA length

As mentioned previously, the ssDNA in our simulations is of length of 235 nucleotides. Since the ssDNA in the experiments is much longer, which may affect the probability of intersegmental transfer, we simulated the diffusion of RPA in several systems for varying length of ssDNA. We used ssDNA of length of 235, 460, 860 and 1400 nucleotides. For these four systems, the end-to-end distances were kept as 222, 435, 817 and 1320 Å, respectively, which corresponds to similar force values of ~ 0.5 pN, for each case.

7. Calculation of diffusion coefficient and intersegmental transfer

The diffusion of RPA along ssDNA is characterized by the diffusion coefficient. In order to do that, first, the distance of all the ssDNA nucleotides were calculated from the protein centre of mass (COM) for every saved trajectory throughout the simulation. The nucleotide index closest to the protein COM in each step was identified as M. Now, the three-dimensional diffusion coefficient (D) was calculated from the mean square displacement (MSD) of the centre-of-mass of RPA moving along the ssDNA:

$$MSD(n, N) = \sum_{i=1}^{N-n} (M_{i+n} - M_i)^2 = 2Dn\Delta t$$

where N is the trajectory length in time steps, n is the measurement window ranging from 1 to N , Δt is the time step interval, and M is the nucleotide index closest to the protein COM at the corresponding time step. The diffusion coefficient was calculated between time frames 200 to 800 for diffusion on DNA.

Further, to convert the simulation time steps in real time unit, we took a small fragment of the ssDNA consisting of 20 nucleotides to monitor its dynamics. The coarse-grained model of this 20-nucleotide ssDNA was simulated for 1×10^6 simulation time steps allowing the whole ssDNA to move freely in space. We ran 50 copies of the same simulation to obtain the lifetime for forming a closed circular ssDNA (i.e., formation of a contacts between the two ends) in the units of simulation time steps. This lifetime was matched with the experimental lifetime in real units in seconds obtained from the experiments by Kim et al.,⁷⁴ and we obtained the equivalent time units in seconds for a simulation time step. We found that each simulation time step corresponds to 3.267×10^{-8} s. Thus, the diffusion coefficient obtained in the units of nt^2 /timesteps were then converted to nt^2 /s.

To calculate the size of the intersegmental transfers (Δnt , which corresponds to the length of the ssDNA bridged during the jump), the absolute value of the difference of nucleotide index (M) closest to protein COM between two consecutive saved trajectories (specifically 1000 consecutive simulation steps) were calculated which can be expressed as follows,

$$(\Delta nt)_i = |M_i - M_{i+1}|$$

where, M_i represents the nucleotide index closest to the protein COM at i^{th} simulation time step, while M_{i+1} stands for the nucleotide index closest to the protein COM at $i + 1$ th simulation time step. We define this difference of nucleotide index as intersegmental transfer if $\Delta nt > 10$.

Fundings

Research reported in this publication was supported by grants from the National Institutes of Health: NIGMS R35 GM149320 to E.A. and NIGMS R35 GM122569 to T.H. The Office of the Director of the National Institutes of Health under award number S10OD025221 to Scot Kuo allowed access to LUMICKS C-Trap. S.K. was supported by a fellowship from the National Cancer Institute F99 CA274696. Y.L. was supported by the Israeli Science Foundation (2072/22) and by a research grant from the Estate of Gerald Alexander. T.H. is an investigator of the Howard Hughes Medical Institute.

CRedit authorship contribution statement

Sushil Pangeni: Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis, Data curation, Conceptualization. **Gargi Biswas:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Vikas Kaushik:** Writing – review & editing, Resources, Methodology, Data curation. **Sahiti Kuppa:** Writing – review & editing, Data curation, Conceptualization. **Olivia Yang:** Writing – review & editing, Supervision, Data curation, Conceptualization. **Chang-Ting Lin:** Writing – review & editing, Visualization, Data curation. **Garima Mishra:** Software, Data curation. **Yaakov Levy:** Writing – review & editing, Validation, Supervision, Software, Methodology, Data curation. **Edwin Antony:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Project administration, Data curation, Conceptualization. **Taekjip Ha:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Funding acquisition, Data curation.

DATA AVAILABILITY

Data will be made available on request.

DECLARATION OF COMPETING INTEREST

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Chang-Ting Lin was previously employed by LUMICKS B.V and Olivia Yang is currently employed by LUMICKS B.V. Y. L. holds The Morton and Gladys Pickman professional chair in Structural Biology.

Acknowledgements:

We would like to thank the team at LUMICKS for their technical help especially Suresh Ramakoti, and Momcilo Gavrilo for initial C-trap training to the author. This article is subject to HHMI's Open Access to Publications policy. HHMI lab heads have previously granted a nonexclusive CC BY 4.0 license to the public and a sublicensable license to HHMI in their research articles. Pursuant to those licenses, the author-accepted manuscript of this article can be made freely available under a CC BY 4.0 license immediately upon publication.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jmb.2024.168491>.

Received 28 December 2023;
Accepted 8 February 2024;
Available online 14 February 2024

Keywords:

Replication protein A (RPA);
Single Molecule Biophysics;
Diffusion;
MD simulation

† These two authors made equal contributions.

References

- Weissart, K., Taneja, P., Fanning, E., (1998). The replication protein A binding site in simian virus 40 (SV40) T antigen and its role in the initial steps of SV40 DNA replication. *J. Virol.* **72**, 9771.
- Wold, M.S., Kelly, T., (1988). Purification and characterization of replication protein A, a cellular protein required for in vitro replication of simian virus 40 DNA. *PNAS* **85**, 2523–2527.
- Brill, S.J., Stillman, B., (1991). Replication factor-A from *Saccharomyces cerevisiae* is encoded by three essential genes coordinately expressed at S phase. *Genes Dev.* **5**, 1589–1600.
- Brill, S.J., Stillman, B., (1989). Yeast replication factor-A functions in the unwinding of the SV40 origin of DNA replication. *Nature* **342**, 92–95.
- Fanning, E., Klimovich, V., Nager, A.R., (2006). A dynamic model for replication protein A (RPA) function in DNA processing pathways. *Nucleic Acids Res.* **34**, 4126–4137.
- Dueva, R., Iliakis, G., (2020). Replication protein A: a multifunctional protein with roles in DNA replication, repair and beyond. *NAR Cancer* **2**.
- Bochkareva, E., Belegu, V., Korolev, S., Bochkarev, A., (2001). Structure of the major single-stranded DNA-binding domain of replication protein A suggests a dynamic mechanism for DNA binding. *EMBO J.* **20**, 612–618.
- Nguyen, B., Sokoloski, J., Galletto, R., Elson, E.L., Wold, M.S., Lohman, T.M., (2014). Diffusion of human replication protein A along single-stranded DNA. *J. Mol. Biol.* **426**, 3246–3261.
- Prakash, A., Borgstahl, G.E.O., Prakash, A., Borgstahl, G. E.O., (2012). The structure and function of replication protein A in DNA replication. *Subcell. Biochem.* **62**, 171–196.
- Deng, S.K., Chen, H., Symington, L.S., (2015). Replication protein A prevents promiscuous annealing between short sequence homologies: Implications for genome integrity. *Bioessays* **37**, 305.
- Wang, Y.R., Guo, T.T., Zheng, Y.T., Lai, C.W., Sun, B., Xi, X.G., Hou, X.M., (2021). Replication protein A plays multifaceted roles complementary to specialized helicases in processing G-quadruplex DNA. *iScience* **24**, 102493
- Nguyen, H.D., Yadav, T., Giri, S., Saez, B., Graubert, T.A., Zou, L., (2017). Functions of replication protein A as a sensor of R loops and a regulator of RNaseH1. *Mol. Cell* **65**, 832–847.e4.
- Sabatinos, S.A., Forsburg, S.L., (2015). Managing single-stranded DNA during replication stress in fission yeast. *Biomol.* **5**, 2123–2139.
- Ashton, N.W., Bolderson, E., Cubeddu, L., O'Byrne, K.J., Richard, D.J., (2013). Human single-stranded DNA binding proteins are essential for maintaining genomic stability. *BMC Mol. Biol.* **14**, 1–20.
- Namiki, Y., Zou, L., (2006). ATRIP associates with replication protein A-coated ssDNA through multiple interactions. *PNAS* **103**, 580–585.
- Maréchal, A., Zou, L., (2015). RPA-coated single-stranded DNA as a platform for post-translational modifications in the DNA damage response. *Cell Res.* **25**, 9–23.
- Chen, R., Wold, M.S., (2014). Replication protein A: single-stranded DNA's first responder. *Bioessays* **36**, 1156–1161.
- Hormeno, S., Wilkinson, O.J., Aicart-Ramos, C., Kuppa, S., Antony, E., Dillingham, M.S., Moreno-Herrero, F., (2022). Human HELB is a processive motor protein that catalyzes RPA clearance from single-stranded DNA. *Proc. Natl. Acad. Sci.*, 119.
- Gibb, B., Ye, L.F., Kwon, Y., Niu, H., Sung, P., Greene, E. C., (2014). Protein dynamics during presynaptic-complex assembly on individual single-stranded DNA molecules. *Nature Struct. Mol. Biol.* **21**, 893–900.
- Jacobs, D.M., Lipton, A.S., Isern, N.G., Daughdrill, G.W., Lowry, D.F., Gomes, X., Wold, M.S., (1999). Human replication protein A: global fold of the N-terminal RPA-70 domain reveals a basic cleft and flexible C-terminal linker†. *J. Biomol. NMR* **14**, 321–331.
- Li, S., Dong, Z., Yang, S., Feng, J., Li, Q., (2019). Chaperoning RPA during DNA metabolism. *Curr. Genet.* **65**, 857–864.
- Yang, Q. et al, (2017). G9a coordinates with the RPA complex to promote DNA damage repair and cell survival. *PNAS* **114**, E6054–E6063.
- Nguyen, D.D., Kim, E.Y., Sang, P.B., Chai, W., (2020). Roles of OB-Fold Proteins in Replication Stress. *Front. Cell Dev. Biol.* **8**, 954.
- Ahmad, F., Patterson, A., Deveryshetty, J., Mattice, J.R., Pokhrel, N., Bothner, B., Antony, E., (2021). Hydrogen-deuterium exchange reveals a dynamic DNA-binding map of replication protein A. *Nucleic Acids Res.* **49**, 1455–1469.
- Yates, L.A., Aramayo, R.J., Pokhrel, N., Caldwell, C.C., Kaplan, J.A., Perera, R.L., Spies, M., Antony, E., et al., (2018). A structural and dynamic model for the assembly of Replication Protein A on single-stranded DNA. *Nature Commun.* **9**, 1–14.
- Bochkareva, E., Korolev, S., Lees-Miller, S.P., Bochkarev, A., (2002). Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA. *EMBO J.* **21**, 1855–1863.
- Bochkareva, E., Korolev, S., Bochkarev, A., (2000). The role for zinc in replication protein A *. *J. Biol. Chem.* **275**, 27332–27338.
- Fan, J., Pavletich, N.P., (2012). Structure and conformational change of a replication protein A heterotrimer bound to ssDNA. *Genes Dev.* **26**, 2337–2347.
- Chen, R., Subramanyam, S., Elcock, A.H., Spies, M., Wold, M.S., (2016). Dynamic binding of replication protein a is required for DNA repair. *Nucleic Acids Res.* **44**, 5758–5772.

30. Brosey, C.A., Soss, S.E., Brooks, S., Yan, C., Ivanov, I., Dorai, K., Chazin, W.J., (2015). Functional dynamics in replication protein A DNA binding and protein recruitment domains. *Structure* **23**, 1028–1038.
31. Brosey, C.A., Yan, C., Tsutakawa, S.E., Heller, W.T., Rambo, R.P., Tainer, J.A., Ivanov, I., Chazin, W.J., (2013). A new structural framework for integrating replication protein A into DNA processing machinery. *Nucleic Acids Res.* **41**, 2313–2327.
32. Blackwell, L.J., Borowiec, J.A., (1994). Human replication protein A binds single-stranded DNA in two distinct complexes. *Mol. Cell Biol.* **14**, 3993–4001.
33. Kumaran, S., Kozlov, A.G., Lohman, T.M., (2006). *Saccharomyces cerevisiae* replication protein A binds to single-stranded DNA in multiple salt-dependent modes. *Biochemistry* **45**, 11958–11973.
34. Caldwell, C.C., Spies, M., (2020). Dynamic elements of replication protein A at the crossroads of DNA replication, recombination, and repair. *Crit. Rev. Biochem. Mol. Biol.* **55**, 482–507.
35. Pokhrel, N. et al, (2019). Dynamics and selective remodeling of the DNA-binding domains of RPA. *Nature Struct. Mol. Biol.* **26**, 129–136.
36. Pokhrel, N., Origanti, S., Davenport, E.P., Gandhi, D., Kaniecki, K., Mehl, R.A., Greene, E.C., Dockendorff, C., et al., (2017). Monitoring replication protein A (RPA) dynamics in homologous recombination through site-specific incorporation of non-canonical amino acids. *Nucleic Acids Res.* **45**, 9413–9426.
37. Wang, Q.M., Yang, Y.T., Wang, Y.R., Gao, B., Xi, X., Hou, X.M., (2019). Human replication protein A induces dynamic changes in single-stranded DNA and RNA structures. *J. Biol. Chem.* **294**, 13915–13927.
38. Kaur, G., Spenkelink, L.M., (2021). Biology on track: single-molecule visualisation of protein dynamics on linear DNA substrates. *Essays Biochem.* **65**, 5–16.
39. Takahashi, S., Kawasaki, S., Miyata, H., Kurita, H., Mizuno, T., Matsuura, S.I., Mizuno, A., Oshige, M., et al., (2014). A new direct single-molecule observation method for DNA synthesis reaction using fluorescent replication protein A. *Sensors* **14**, 5174–5182.
40. Bain, F.E., Fischer, L.A., Chen, R., Wold, M.S., (2018). Single-molecule analysis of replication protein A-DNA interactions. *Methods Enzymol.* **600**, 439–461.
41. Ma, C.J., Gibb, B., Kwon, Y., Sung, P., Greene, E.C., (2017). Protein dynamics of human RPA and RAD51 on ssDNA during assembly and disassembly of the RAD51 filament. *Nucleic Acids Res.* **45**, 749–761.
42. Ha, T., Enderle, T., Ogle, D.F., Chemla, D.S., Selvin, P. R., Weiss, S., (1996). Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc. Natl. Acad. Sci.* **93**, 6264–6268.
43. Grimme, J.M., Spies, M., (2011). FRET-based assays to monitor DNA binding and annealing by Rad52 recombination mediator protein. *Methods Mol. Biol.* **745**, 463–483.
44. Kemmerich, F.E., Daldrop, P., Pinto, C., Levikova, M., Cejka, P., Seidel, R., (2016). Force regulated dynamics of RPA on a DNA fork. *Nucleic Acids Res.* **44**, 5837–5848.
45. Collins, B.E., Ye, L.F., Duzdevich, D., Greene, E.C., (2014). DNA curtains: Novel tools for imaging protein-nucleic acid interactions at the single-molecule level. *Methods Cell Biol.* **123**, 217–234.
46. Finkelstein, I.J., Greene, E.C., (2011). Supported lipid bilayers and DNA curtains for high-throughput single-molecule studies. *Methods Mol. Biol.* **745**, 447–461.
47. Gibb, B., Ye, L.F., Gergoudis, S.C., Kwon, Y., Niu, H., Sung, P., Greene, E.C., (2014). Concentration-dependent exchange of replication protein A on single-stranded DNA revealed by single-molecule imaging. *PLoS One* **9**, e87922.
48. Lee, K.S., Marciel, A.B., Kozlov, A.G., Schroeder, C.M., Lohman, T.M., Ha, T., (2014). Ultrafast redistribution of *E. coli* SSB along long single-stranded DNA via intersegment transfer. *J. Mol. Biol.* **426**, 2413.
49. Roy, R., Kozlov, A.G., Lohman, T.M., Ha, T., (2009). SSB protein diffusion on single-stranded DNA stimulates RecA filament formation. *Nature* **461**, 1092–1097.
50. Mersch, K.N., Sokoloski, J.E., Nguyen, B., Galletto, R., Lohman, T.M., (2023). “Helicase” activity promoted through dynamic interactions between a ssDNA translocase and a diffusing SSB protein. *Proc. Natl. Acad. Sci.* **120**, e2216777120
51. Mishra, G., Bigman, L.S., Levy, Y., (2020). ssDNA diffuses along replication protein A via a reptation mechanism. *Nucleic Acids Res.* **48**, 1701–1714.
52. Mondal, A., Bhattacharjee, A., (2020). Mechanism of dynamic binding of replication protein A to ssDNA. *J. Chem. Inf. Model.* **60**, 5057–5069.
53. Mishra, G., Levy, Y., (2015). Molecular determinants of the interactions between proteins and ssDNA Molecular determinants of the interactions between proteins and ssDNA. *PNAS* **112**, 5033–5038.
54. Bigman, L.S., Levy, Y., (2023). Protein diffusion along protein and DNA Lattices: role of electrostatics and disordered regions. *Annu. Rev. Biophys.* **52**, 463–486.
55. Kuppa, S., Pokhrel, N., Corless, E., Origanti, S., Antony, E., (2021). Generation of Fluorescent Versions of *Saccharomyces cerevisiae* RPA to Study the Conformational Dynamics of Its ssDNA-Binding Domains. *In: Methods in Molecular Biology*. Humana, New York, NY, pp. 151–168.
56. Lee, K.S., Balci, H., Jia, H., Lohman, T.M., Ha, T., (2013). Direct imaging of single UvrD helicase dynamics on long single-stranded DNA. *Nature Commun.* **4**, 1878.
57. Candelli, A., Hoekstra, T.P., Farge, G., Gross, P., Peterman, E.J.G., Wuite, G.J.L., (2013). A toolbox for generating single-stranded DNA in optical tweezers experiments. *Biopolymers* **99**, 611–620.
58. Michalet, X., Berglund, A.J., (2012). Optimal diffusion coefficient estimation in single-particle tracking. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.* **85**, 061916
59. Kim, C., Paulus, B.F., Wold, M.S., (2002). Interactions of human replication protein A with oligonucleotides. *Biochemistry* **33**, 14197–14206.
60. Murphy, M.C., Rasnik, I., Cheng, W., Lohman, T.M., Ha, T., (2004). Probing single-stranded DNA Conformational flexibility using fluorescence spectroscopy. *Biophys. J.* **86**, 2530–2537.
61. Chen, J., Le, S., Basu, A., Chazin, W.J., Yan, J., (2015). Mechanochemical regulations of RPA’s binding to ssDNA. *Sci. Rep.* **5**, 9296.
62. Roshan, P. et al, (2023). An Aurora B-RPA signaling axis secures chromosome segregation fidelity. *Nature Commun.*, 14.

63. Kuppa, S. et al, (2022). Rtt105 regulates RPA function by configurationally stapling the flexible domains. *Nature Commun.*, 13.
64. Zhao, W., Ali, M.M., Brook, M.A., Li, Y., (2008). Rolling circle amplification: applications in nanotechnology and biodetection with functional nucleic acids. *Angew. Chem. Int. Ed.* **47**, 6330–6337.
65. Rei, E., Hölzel, R., Bier, F.F., (2009). Synthesis and stretching of rolling circle amplification products in a flow-through system. *Small* **5**, 2316–2322.
66. Brockman, C., Kim, S.J., Schroeder, C.M., (2011). Direct observation of single flexible polymers using single stranded DNA. *Soft Matter* **7**, 8005–8012.
67. Jiang, X. et al, (2006). Structural mechanism of RPA loading on DNA during activation of a simple pre-replication complex. *EMBO J.* **25**, 5516–5526.
68. Cai, L., Roginskaya, M., Qu, Y., Yang, Z., Xu, Y., Zou, Y., (2007). Structural characterization of human RPA sequential binding to single-stranded DNA using ssDNA as a molecular ruler. *Biochemistry* **46**, 8226–8233.
69. Zhou, R., Kozlov, A.G., Roy, R., Zhang, J., Korolev, S., Lohman, T.M., Ha, T., (2011). SSB functions as a sliding platform that migrates on DNA via reptation. *Cell* **146**, 222–232.
70. Smith, S.B., Cui, Y., Bustamante, C., (1996). Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science (80-)* **271**, 795–799.
71. Wang, M.D., Yin, H., Landick, R., Gelles, J., Block, S.M., (1997). Stretching DNA with optical tweezers. *Biophys. J.* **72**, 1335.
72. Vanlier, J., Moldovan, D., Pauszek, R., Broekmans, O., Mirone, A., Aafke, Lamerton, S., (2022). *lumicks/pylake v0.11.1*
73. Pal, A., Levy, Y., (2019). Structure, stability and specificity of the binding of ssDNA and ssRNA with proteins. *PLoS Comput. Biol.* **15**, e1006768
74. Kim, S.H., Lee, T.-H., (2021). Conformational dynamics of poly(T) single-stranded DNA at the single-molecule level. *J. Phys. Chem. Letter* **12**, 4576–4584.