

Partial wrapping of single-stranded DNA by replication protein A and modulation through phosphorylation

Rahul Chadda^{1,†}, Vikas Kaushik^{1,†}, Iram Munir Ahmad², Jaigeeth Deveryshetty¹, Alex S. Holehouse³, Snorri Th. Sigurdsson², Gargi Biswas⁴, Yaakov Levy⁴, Brian Bothner⁵, Richard B. Cooley⁶, Ryan A. Mehl⁶, Reza Dastvan¹, Sofia Origanti⁷ and Edwin Antony^{1,*}

¹Department of Biochemistry and Molecular Biology, Saint Louis University School of Medicine, St. Louis, MO 63104, USA

²Department of Chemistry, Science Institute, University of Iceland, 107 Reykjavik, Iceland

³Department of Biochemistry and Molecular Biophysics, Washington University in Saint Louis School of Medicine, St. Louis, MO 63110, USA

⁴Department of Chemical and Structural Biology, Weizmann Institute of Science, Rehovot, Israel

⁵Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT 59717, USA

⁶Department of Biochemistry and Biophysics, Oregon State University, Corvallis, OR 97331, USA

⁷Department of Biology, Saint Louis University, St. Louis, MO 63103, USA

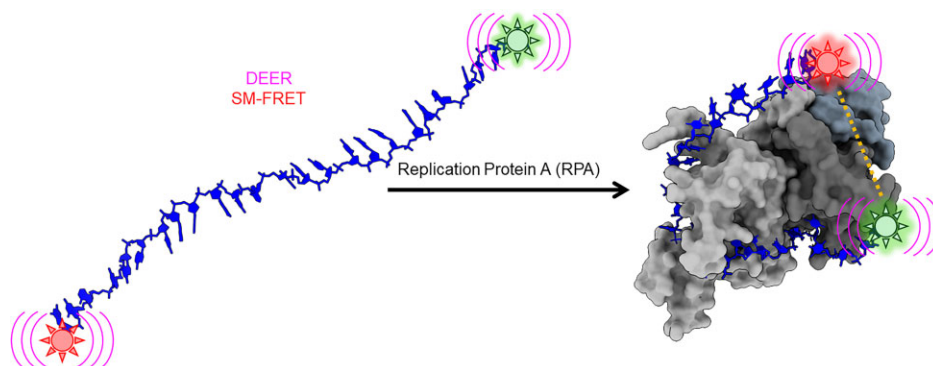
*To whom correspondence should be addressed. Tel: +1 314 977 9257; Fax: +1 314 977 9206; Email: edwin.antony@health.slu.edu

[†]The first two authors should be regarded as Joint First Authors.

Abstract

Single-stranded DNA (ssDNA) intermediates which emerge during DNA metabolic processes are shielded by replication protein A (RPA). RPA binds to ssDNA and acts as a gatekeeper to direct the ssDNA towards downstream DNA metabolic pathways with exceptional specificity. Understanding the mechanistic basis for such RPA-dependent functional specificity requires knowledge of the structural conformation of ssDNA when RPA-bound. Previous studies suggested a stretching of ssDNA by RPA. However, structural investigations uncovered a partial wrapping of ssDNA around RPA. Therefore, to reconcile the models, in this study, we measured the end-to-end distances of free ssDNA and RPA–ssDNA complexes using single-molecule FRET and double electron–electron resonance (DEER) spectroscopy and found only a small systematic increase in the end-to-end distance of ssDNA upon RPA binding. This change does not align with a linear stretching model but rather supports partial wrapping of ssDNA around the contour of DNA binding domains of RPA. Furthermore, we reveal how phosphorylation at the key Ser-384 site in the RPA70 subunit provides access to the wrapped ssDNA by remodeling the DNA-binding domains. These findings establish a precise structural model for RPA-bound ssDNA, providing valuable insights into how RPA facilitates the remodeling of ssDNA for subsequent downstream processes.

Graphical abstract



Introduction

Replication protein A (RPA) is an essential eukaryotic single-stranded DNA (ssDNA) binding protein that sequesters transiently exposed ssDNA during various DNA metabolic processes (1–4). RPA binds to ssDNA with high affinity (K_d

$\sim 10^{-10}$ M) (5,6), resolves secondary structures, and shields it from nucleolytic degradation (7–10). Notably, the assembly of RPA filaments on ssDNA acts as a trigger for the DNA damage checkpoint (11–14). RPA–ssDNA filaments also function as an interaction hub by recruiting over thirty proteins/enzymes

Received: April 1, 2024. Revised: May 30, 2024. Editorial Decision: June 18, 2024. Accepted: June 25, 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of Nucleic Acids Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

and promoting their assembly on DNA with correct binding polarity (15–17). This multifaceted functionality is enacted through six oligonucleotide/oligosaccharide binding (OB) domains housed within a heterotrimeric complex (Figure 1A) (18–20).

Human RPA is a constitutive heterotrimer composed of three subunits RPA70, RPA32 and RPA14. Six OB domains (A–F) are spread across the three subunits. OB-F, A, B and C reside in RPA70 and are connected by disordered linkers. RPA32 harbors OB-D and a winged helix (wh) motif. OB-E is a structural domain in RPA14 and holds the three subunits together as part of a trimerization core (Tri-C) along with OB-C & OB-D (21). Domains A, B, C & D primarily coordinate ssDNA interactions and are termed DNA-binding domains (DBDs). OB-F and wh coordinate protein-protein interactions and are called protein-interaction domains (PIDs). Dynamic rearrangements of the DBDs and PIDs are observed upon binding to DNA and in response to post-translational modifications (22–25). Each DBD exhibits a moderate affinity for ssDNA binding. Consequently, their collective action engenders a remarkably high affinity and stoichiometric binding of RPA to ssDNA (1). The intrinsic dynamic binding/dissociation characteristics of each DBD transiently unveil pockets within the buried ssDNA, enabling access for incoming proteins while the RPA complex remains attached to the ssDNA (19,26,27).

One fascinating aspect of RPA–ssDNA interactions and the proposed roles in DNA metabolism is the sequestration of the DNA. If the ssDNA is buried under RPA due to the high-affinity interactions, how do incoming proteins gain access to the DNA? This problem is enhanced when specific structures associated with the DNA need to be recognized. For example, during homologous recombination (HR), a double strand break is corrected by resection of one strand and thus producing a DNA substrate with a long 3′ overhang and a ssDNA junction (28,29). The ssDNA is coated by RPA and a Rad51 nucleoprotein must be formed on the ssDNA to promote subsequent steps in HR. So, the position of RPA on the DNA with respect to the ss–dsDNA junction likely plays a key role in initiating Rad51 binding. These ssDNA overhangs can be up to 2000 nt long (30,31), thus positioning of the 3′ end of the ssDNA and the ss–dsDNA junction are critical. Two crucial parameters play a pivotal role in unravelling the intricate RPA–ssDNA interactions: a) the length of the ssDNA, and b) the shaping or contour of the RPA–ssDNA complex. The length of the ssDNA determines the number of assembled RPA molecules, subsequently influencing the variety and number of interactors recruited during a specific cellular DNA metabolic process (6). Concurrently, the shape of the complex determines the positioning of the ss–ds junction, or fork, respective to the 3′ or 5′ termini. If the DNA is stretched by RPA, then the spacing between two positions on the DNA will be distant. Conversely, if DNA is wrapped by RPA, then this distance would be much closer. On average, the occluded site-size for a single human RPA heterotrimer is ~18–25 nucleotides (nt) (32). Therefore, the density of RPA molecules can be reasonably inferred to be proportional to the length of the ssDNA. Additionally, employing poly-pyrimidine (dT) ssDNA as experimental substrates further mitigates interference arising from secondary structures.

Surprisingly, there is a lack of experimental consensus regarding the configuration of ssDNA within the RPA–ssDNA

complex. In both X-ray and CryoEM structures of diverse RPA–ssDNA complexes, the ssDNA displays a ‘C’ shape, exhibiting end-to-end distances typically ranging between ~5 and 8 nm (Figure 1B) (18–20). Remarkably, in all these structures, the contour of the ssDNA along the DBDs of RPA is similar including comparable end-to-end distances. However, the structural findings could be influenced by the utilization of truncated RPA (18), or a limited subset of its DBDs (21), alongside potential artifacts originating from crystal packing (crystallography), or dynamics-induced factors (CryoEM) (19). In contrast, biochemical and bulk FRET measurements suggest a complete linearization of even short 30 nucleotide (nt) long ssDNA (19). However, bulk FRET assays are often a readout of multiple underlying populations, especially at relatively high protein to ssDNA ratios.

To mitigate experimental biases or limitations, and accurately determine the contour of ssDNA bound to RPA, we employed solution-based single-molecule confocal FRET and double electron–electron resonance (DEER) spectroscopy (33–36) to directly measure the end-to-end distance of ssDNA, both in the absence and presence of RPA. On average, we observe only a small ~3 nm increase in the ssDNA end-to-end distance in the RPA-bound complex. Our results suggest a partial wrapping of ssDNA around RPA, exhibiting a contour closer to that observed in the crystal structure (18). Therefore, models depicting RPA–ssDNA interactions during DNA metabolism should consider the significant curvature induced in the DNA lattice upon RPA binding. Furthermore, we show that the post-translational modification of RPA through phosphorylation at Ser-384 in the RPA70 subunit introduces substantial changes in the arrangement of the DBDs and PIDs while causing only minimal changes in the pattern of ssDNA wrapping. Thus, access to internal ssDNA segments can be made available by remodeling the RPA domains to serve specific DNA metabolic roles without substantially altering the path, or shape, of DNA.

Materials and methods

Preparation of oligonucleotides

Unlabeled, Cyanine-5, and Cyanine-3 end-labeled poly-(dT) ssDNA oligonucleotides of various lengths (dT)_n (*n* = 15, 30, 45, 60, 80 or 97 thymidine bases; [Supplementary Table S1](#)) were purchased from Integrated DNA Technologies (Iowa). For the synthesis of the doubly spin-labeled (dT)₂₂ and (dT)₅₀ oligonucleotides, 2′-aminouridine nucleotides were incorporated at specific sites ([Supplementary Table S1](#)) and post-synthetically spin labeled with an isothiocyanate derivative of an isoindoline nitroxide as described (37).

Purification of RPA and site-specific phosphoserine incorporation

Human RPA was produced using plasmid pET-Duet1-hRPA^{Syn}-70C-His coding for a poly-His affinity tag at the C-terminus of RPA70. The open reading frames for RPA70, RPA32 and RPA14 were codon-optimized for overexpression in *E. coli* (GenScript Inc). RPA70 was engineered into multiple cloning site (MCS) 1 while RPA32 and RPA14 were cloned into MCS2. RPA was purified as described (38) with the following modifications. Briefly, the plasmid was transformed into BL21 (DE3) cells and transformants were selected

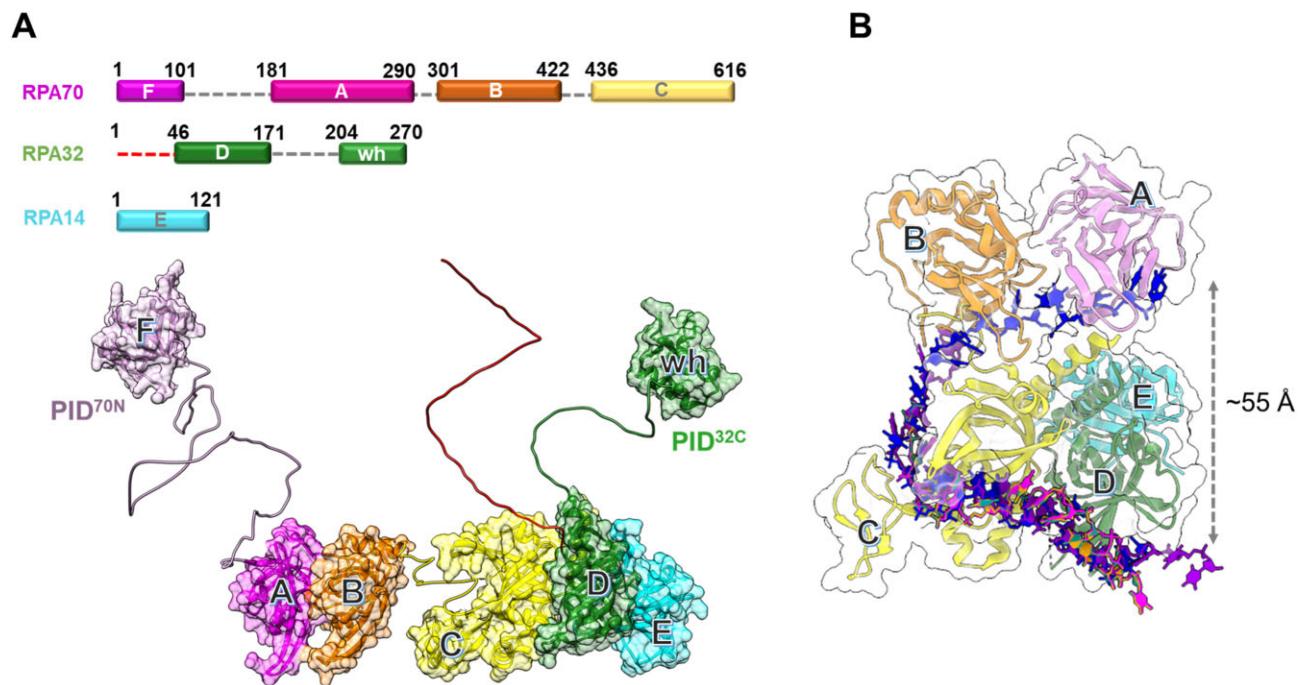


Figure 1. Architecture of the replication protein A (RPA) – ssDNA complex. **(A)** The three subunits of RPA—RPA70, RPA32 and RPA14 house several oligonucleotide/oligosaccharide binding (OB) domains that are classified as either DNA binding domains (DBDs: DBD-A, DBD-B, DBD-C and DBD-D) or protein-interaction domains (OB-F/PID^{70N} and winged helix (wh)/PID^{32C}). A structural model of RPA is shown and was generated using information from known structures of the OB-domains and AlphaFold2 models of the linkers. **(B)** Crystal structure of *Ustilago maydis* RPA bound to ssDNA is shown (PDB 4GNX). This structure lacks OB-F, the F-A linker, wh and the D-wh linker. ssDNA from CryoEM structures of *Saccharomyces cerevisiae* RPA (PDB 6I52) and *Pyrococcus abyssi* RPA (8OEL) are superimposed and reveal an end-to-end DNA distance of ~55 Å. Structural data supports partial wrapping of ssDNA around the OB domains.

using ampicillin (100 µg/ml). A single colony was inoculated in 1 L of Luria Broth and incubated at 37°C without shaking for 20-24 hrs. Cells were then grown at 37°C with shaking at 250 rpm, until the OD₆₀₀ reached 0.6 and then induced with 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Induction was carried out at 37°C for 3 h. Harvested cells were resuspended in 30 ml/l cell resuspension buffer (30 mM HEPES, pH 7.8, 300 mM KCl, 0.02% v/v Tween-20, 1.5× protease inhibitor cocktail, 1 mM PMSF and 10% (v/v) glycerol). Cells were lysed with 0.4 mg/ml lysozyme for 30 min at 4°C followed by sonication in the cold room. The samples were maintained at 4°C through the entirety of the remaining purification. The clarified lysate was fractionated over a Ni²⁺-NTA agarose column (Gold Biotechnology Inc.). RPA was eluted using cell resuspension buffer containing 400 mM imidazole following a wash with lysis buffer containing 2M NaCl to remove any trace non-specifically bound nucleic acids. Fractions containing RPA were pooled and diluted with H₀ buffer (30 mM HEPES, pH 7.8, 0.02% v/v Tween-20, 1.5× protease inhibitor cocktail, 10% (v/v) glycerol and 0.25 mM EDTA pH 8.0) to match the conductivity of buffer H₁₀₀ (H₀ + 100 mM KCl), and further fractionated over a fast-flow Heparin column (Cytiva Inc.). RPA was eluted using a linear gradient of H₁₀₀–H₁₅₀₀ buffers (subscripts denote the mM concentrations of KCl), and fractions containing RPA were pooled and concentrated using an Amicon spin concentrator (30 kDa molecular weight cut-off). The concentrated RPA was fractionated over a HiLoad 26/600 Superdex-200 column (Cytiva Inc.) using RPA storage buffer (30 mM HEPES, pH 7.8, 200 mM KCl, 0.25 mM EDTA, 0.01% v/v Tween-20 and 10% (v/v) glycerol). Purified RPA was flash frozen using liquid nitrogen and

stored at –80°C. RPA concentration was measured spectroscopically using $\epsilon_{280} = 87\,410\text{ M}^{-1}\text{cm}^{-1}$.

Human RPA-pSer³⁸⁴, carrying a phospho-serine at position 384 in RPA70, was expressed and purified using genetic code expansion (39,40). First, a fragment containing the RPA70, RPA32 and RPA14 open reading frames was subcloned into an RSF-Duet1 plasmid and an amber suppression codon (TAG) was engineered at the position corresponding to Ser-384 in the RPA70 subunit using Q5 site-directed mutagenesis (NEB). pRSF-Duet1-hRPA-70C-His-S384TAG and pKW2-EFsep (41) plasmids were co-transformed in to BL21 (DE3) Δ serB *E. coli* cells and transformants were selected using chloramphenicol (25 µg/ml) and kanamycin (50 µg/ml). A starter culture was prepared by inoculating colonies into ZY-non inducing media (ZY-NIM; Supplementary Table S2A) followed by overnight growth with shaking at 250 rpm at 37°C. 1% of the overnight starter culture was added to ZY-auto induction media (ZY-AIM; Supplementary Table S2B) and grown until the OD₆₀₀ reached 1.5. The temperature was reduced to 20°C and the cultures were grown for an additional 20 h. Cells were then harvested by centrifugation at 2057 xg for 20 min and the cell pellet was resuspended with cell-resuspension buffer (30 mM HEPES, pH 7.8, 300 mM KCl, 0.02% v/v Tween-20, 1.5× protease inhibitor cocktail, 1 mM PMSF, 10% (v/v) glycerol, 50 mM sodium fluoride, 10 mM sodium pyrophosphate, and 1 mM sodium orthovanadate). RPA-pSer³⁸⁴ was purified as described above for wild type RPA. Typical yields for RPA-pSer³⁸⁴ are ~6.8 mg of pure protein/L of culture compared to ~20–25 mg/l for wild-type unmodified RPA. Phosphoserine incorporation was confirmed using mass spectrometry and analysis on Phos-tag SDS PAGE.

Analysis of pSer incorporation using Phos-Tag SDS-PAGE

Incorporation of pSer into RPA70 was assessed using Phos-tag SDS PAGE (Fujifilm Inc.). A gradient SDS-PAGE gel (10–14%) was prepared by adding 0.05 mM of the Phos-tag reagent and 0.1 mM MnCl₂ to the resolving gel (before pouring the gel). Other steps for preparing the resolving and stacking gels, and running the gel, are as defined for standard SDS-PAGE analysis. Gels are stained using Coomassie stain. The band carrying pSer in RPA70 runs with slower electrophoretic mobility in Phos-tag SDS PAGE analysis.

Single-molecule FRET measurements

smFRET data were collected on an EI-FLEX bench-top microscope from Exciting Instruments Ltd (Sheffield, UK). For measurement, 100 μl of a fluorescent sample droplet was placed onto a no.1 thickness coverslip and excited with alternating 520 and 638 nm lasers at 0.22 and 0.15 mW power, respectively. Experiments were carried out in buffer containing 50 mM Tris-acetate, pH 7.5, 50 mM KCl, 5 mM MgCl₂, 10% (v/v) glycerol and 0.1 mg/ml BSA. Lasers were sequentially turned ON for 45 μs for each measurement and separated by a dark period of 5 μs for a total of 40 mins of acquisition. Fluorescence emission photons from freely diffusing molecules were collected using an Olympus 60× (1.2 N.A.) water-immersion objective, focused onto a 20 μm pinhole. 20 pM of fluorescent sample was suitable for a burst rate of 1 Hz. After passing through the pinhole, the photons were split using a 640 nm long-pass filter, cleaned up using 572 and 680 nm band-pass filters, and focused onto respective avalanche photodiodes. The photon arrival times, and respective detector were saved in HDF5 data format for offline analysis.

After photoirradiation, the background counts from the buffer were comparable to that of DI water at 2–3 counts per second (cps), whereas typical bursts comprised of 50–100 photons. The photoirradiation of buffer was performed using a 100W LED flood light. The entire set-up (i.e. flood light and buffers taken in covered glass beakers) was placed in a cold-room to minimize heating of samples. Analysis of smFRET data was performed in an Anaconda environment, with Jupyter notebooks, using FRETbursts Python package (42). Single molecule photon emission bursts were identified using a dual channel burst search (DCBS) algorithm as previously described (L = 10, and F = 45 for both channels). The background estimated from an exponential fit to inter-photon delays greater than 1.5 ms was subtracted. The compensation for spectral crosstalk (a), compensation factor for different detection efficiencies between donor and acceptor channel (g), and compensation for direct excitation (d) for EI-FLEX were estimated as 0.0938, 1.591 and 0.05824, respectively. The same factors for measurements on the Picoquant Micro Time (MT200; Picoquant Inc., Germany) instrument were 0.05, 0.85 and 0.1, respectively.

Continuous wave (CW)-EPR and DEER spectroscopy

Continuous wave (CW)-EPR spectra of spin-labeled ssDNA samples ± RPA were collected at room temperature on a Bruker EMX spectrometer operating at X-band frequency (9.5 GHz) using 2 mW incident power and a modulation amplitude of 1 G. DEER spectroscopy was performed on an Elexsys E580 EPR spectrometer operating at Q-band fre-

quency (33.9 GHz) with the dead-time free four-pulse sequence at 83 K. Pulse lengths were 20 ns ($\pi/2$) and 40 ns (π) for the probe pulses and 40 ns for the pump pulse. The frequency separation was 63 MHz. Samples for DEER analysis were cryoprotected with 24% (vol/vol) glycerol and flash-frozen in liquid nitrogen. Primary DEER decays were analyzed using a home-written software (DeerA, Dr Richard Stein, Vanderbilt University) operating in the Matlab (MathWorks) environment as previously described (43). Briefly, the software carries out analysis of the DEER decays obtained under different conditions for the same spin-labeled pair. The distance distribution is assumed to consist of a sum of Gaussians, the number and population of which are determined based on a statistical criterion.

Confocal smFRET with alternating-laser excitation (ALEX)

The photons emitted during a transit through confocal volume are called a burst and can be used to estimate FRET efficiency (E).

$$E = \frac{f_{Dex}^{Aem}}{(f_{Dex}^{Dem} + f_{Dex}^{Aem})} \quad (1)$$

This uncorrected *E* is a ratio of number of sensitized acceptor emission photons i.e. via energy transfer (f_{Dex}^{Aem}) and the sum of number of photons in the donor channel after donor excitation (f_{Dex}^{Dem}) and the number of photons in the acceptor channel after donor excitation (f_{Dex}^{Aem}). However, to convert the *E* into distances, spectral crosstalk must be taken into account. This especially poses a problem in the low *E* regimen, i.e. distinguishing low *E* species from donor alone species is a huge challenge. And this is where alternating laser excitation (ALEX) solves the problem. During an ALEX scheme, both donor and acceptor fluorophores are rapidly excited in alternating fashion. The diffusion coefficient of biomolecules in dilute solutions is of the order of hundreds of $\mu m^2/s$. And, so typically an alternation frequency of 20KHz is more than sufficient to excite both donor and acceptor fluorophores multiple times during the transit of a single molecule through the confocal volume (~1 ms). The f_{Aex}^{Aem} thus estimated can be used to calculate another informative quantity about the population of diffusing single molecules named raw stoichiometry.

$$S = \frac{(f_{Dex}^{Aem} + f_{Dex}^{Dem})}{(f_{Dex}^{Dem} + f_{Dex}^{Aem} + f_{Aex}^{Aem})} \quad (2)$$

Stoichiometry can be best understood in terms of ratio of total fluorescence photons recorded after donor wavelength excitation to total fluorescence photons recorded after direct donor and acceptor excitation. For donor-only and acceptor-only species the *S* is close to 1 and 0, respectively. Similarly, the donor and acceptor laser intensity are tuned such that double-labeled molecules scale to an *S* value of 0.5. Furthermore, binding of a protein like RPA to Cy5/Cy3 labeled ssDNA could lead to a reduction in *E* (i.e. DNA ends are brought further apart), or an increase in *S* (increase in donor quantum yield due to PIFE) or both. Finally, in order to convert *E* into distances, three additional steps were taken: (1) Background intensity in all three photon streams f_{Dex}^{Aem} , f_{Dex}^{Dem} , and f_{Aex}^{Aem} is estimated via mean count rate and subtracted taking the length of burst into account. (2) Two channel cross-talk factors are estimated (a) bleed-through of donor-emission into

the acceptor channel and (b) the direct excitation of acceptor molecules by the donor-laser. These are estimated from the donor-, and acceptor- only molecules in the background-corrected ES histogram, and finally (3) gamma correction, which takes into account the differences in the detection efficiencies of donor and acceptor, their quantum, and the transmission efficiencies of the optical elements etc. The corrected FRET values were converted into distance following the relationship between energy transfer efficiency (E) and donor-to-acceptor separation (r):

$$E = \frac{1}{1 + \left(\frac{r}{R_0}\right)^6} \quad (3)$$

The R_0 or the donor-acceptor distance at which the energy transfer efficiency is 50% was taken to be 5.4 nm. The end-to-end distance for a fully linearized ssDNA molecule was estimated simply by multiplying l_d or rise per base, taken to be 0.67 nm, with the number of the bases in the dT polymer.

Crosslinking mass spectrometry (XL-MS) analysis

RPA or RPApSer³⁸⁴ (10 μ M) in the absence or presence of ssDNA [(dT)₂₅] (10 μ M) were incubated with 5 mM BS3 crosslinker at room temperature for 15 min in 20 μ l reaction buffer (50 mM HEPES pH 7.8, 100 mM KCl, 10% glycerol). The crosslinking reaction was quenched with 2 μ l of 1M ammonium acetate for 15 min and the samples were separated on SDS-PAGE. Gel bands were excised and destained with a 50 mM ammonium bicarbonate and 50% acetonitrile mixture and reduced with a mixture of 100 mM dithiothreitol and 25 mM bicarbonate for 30 min at 56°C. The reaction was subsequently exchanged for the alkylation step with 55 mM iodoacetamide and 25 mM ammonium bicarbonate and incubated in the dark at room temperature for 25 min. The solution was then washed with the 50 mM ammonium bicarbonate and 50% acetonitrile mixture. The gel pieces were then first dehydrated with 100% acetonitrile and then rehydrated with sequence grade trypsin solution (0.6 μ g, Promega) and incubated overnight at 37°C. The reaction was quenched with 10 μ l of 50% acetonitrile and 0.1% formic acid (FA, Sigma) and transferred to new microfuge tubes, vortexed for 5 min, and centrifuged at 15 000 rpm for 30 min. The extracted and dried peptide was reconstituted with 0.1% FA in water and injected onto a Neo trap cartridge coupled with an analytical column (75 μ m ID \times 50 cm PepMap Neo C18, 2 μ m). Samples were separated using a linear gradient of solvent A (0.1% formic acid in water) and solvent B (0.1% formic acid in ACN) over 120 mins using a Vanquish Neo UHPLC System coupled to an Orbitrap Eclipse Tribrid Mass Spectrometer with FAIMS Pro Duo interface (Thermo Fisher Scientific). The acquired MS/MS data was queried for cross-link identification against the sequence of five target proteins using Proteome Discoverer v3.0 with the XlinkX node, applying a 1% FDR for cross-link validation. Data were visualized using xiVIEW (44) and plotted in Inkscape.

Mass photometry (MP) measurements

All measurements were carried out on a TwoMP instrument (Refeyn Ltd.) as described before (45). Briefly, glass coverslips (No. 1.5H thickness, 24 \times 50 mm, VWR) were cleaned by sonication in isopropanol followed by deionized water and dried using a nitrogen gas stream. For each round of measurement, a clean coverslip was placed on the oil-immersion objective lens

(Olympus PlanApo N, 1.42 NA, 60 \times), with a holey-silicone gasket (Refeyn Ltd) adhered on the top surface of the coverslip. All dilutions and measurements were performed at room temperature (23 \pm 2°C) in 1 \times Mg²⁺/Ca²⁺ buffer (20 mM HEPES, pH 7.5, 150 mM KCl, 5 mM MgCl₂, 5 mM CaCl₂) supplemented with 1 mM DTT. Samples of DNA alone, RPA alone, or DNA-RPA mixtures were allowed to equilibrate at 23 \pm 2°C for 5 min after which 1 μ l of the respective sample was quickly diluted in 15 μ l of buffer. The newly adhered spots were video recorded for 1 min. High contrast (light-scattering) events corresponding to single particle landings on the coverslip were analyzed further. A known mass standard (β -amylase, Sigma A8781-1VL) was used to convert image contrast-signal into mass units. Histograms were plotted from all the data gathered during the 1 min video interval and non-linear least squares fit to single gaussian function to extract mean mass and error.

Coarse-grained molecular dynamics (MD) simulations

To explore the interaction between RPA and ssDNA at the molecular level, a coarse-grained model has been utilized. In this model, each protein residue was represented by two beads positioned at the C $_{\alpha}$ and C $_{\beta}$ locations. For charged amino acids (K, R, H, D and E), charges were placed at the C $_{\beta}$ position. The coarse-grained model for ssDNA comprised three beads representing each nucleotide: one at the geometric center of phosphate (P), one at sugar (S) and one at base (B). While the S and B beads were neutrally charged, the P bead carried a negative charge. Simulations were conducted using a native topology-based model, where the internal energy of the protein, ssDNA and protein-ssDNA interaction was denoted as E_{prot} , E_{ssDNA} and $E_{ssDNA-Protein}$, respectively. These energies can be expressed by the following equations:

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

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

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

The potential energy of E^{Bond} , E^{Angle} , $E^{Dihedral}$, $E^{Electrostatics}$, $E^{Repulsions}$ can be expressed by $\sum_{bonds} K_{bonds} (b_{ij} - b_{ij}^0)^2$, $\sum_{angles} K_{angles} (\theta_{ijk} - \theta_{ijk}^0)^2$, $\sum_{dihedrals} K_{dihedrals} [1 - \cos(\phi_{ijkl} - \phi_{ijkl}^0) - \cos(3\Phi_{ijkl} - \Phi_{ijkl}^0)]$, $\sum_{i,j} K_{electrostatics} B(\kappa) \frac{q_i q_j \exp^{-\kappa r}}{\epsilon_r r_{ij}}$ and $\sum_{i \neq j} K_{contacts} \left(\frac{C_{ij}}{r_{ij}}\right)^{12}$ respectively. The $E_{ssDNA-Protein}^{Aromatic}$, $E_{prot}^{Native\ Contacts}$, $E_{ssDNA}^{Stacking}$, $E_{ssDNA-Protein}^{Aromatic}$ are expressed by a Lennard-Jones interaction term designated by, $\sum_{i \neq j} K_{contacts} [5 \left(\frac{A_{ij}}{r_{ij}}\right)^{12} - 6 \left(\frac{A_{ij}}{r_{ij}}\right)^{10}]$. All

parameters described in these equations are adopted from previous studies (46,47). The repulsion radius of P, S and B beads is set to 1.9 Å. The stacking interaction potential between protein and ssDNA beads in contact in the crystal structure (PDB: 4GNX) (18) is denoted as $K_{contacts} = 3.0$,

while for those not in contact in the crystal structure, it is set to 0.5. This configuration ensures that RPA does not diffuse through ssDNA. The salt concentration in the Debye–Hückel model of electrostatics is set to 0.01 M. Additionally, the flexibility of ssDNA in response to salt concentration is represented by the dihedral potential (K_{dihedral}) between four consecutive phosphate beads in the ssDNA, set to either 0.0 or 0.4 to simulate high and low salt conditions, respectively.

Results

Solution based single molecule FRET with alternating-laser excitation (ALEX) provides an excellent read out of end-to-end ssDNA distances

Single molecule FRET (smFRET) has been used to obtain accurate end-to-end distances for double-stranded nucleic acids (48). Similar measurements of conformational flexibility for ssDNA have been measured using single molecule total internal reflection (smTIRF) microscopy where overhang DNA with varying length of ssDNA were tethered to glass slides (49). Here, we used a benchtop microscope that uses ALEX (50) for in-solution smFRET measurements to capture the conformations of ssDNA and RPA–ssDNA complexes (51). First, to test whether this instrumental setup accurately captures the conformational sampling of ssDNA, we measured the changes in FRET as a function of ssDNA length. For these experiments, all ssDNA substrates carried Cy3 (donor) and Cy5 (acceptor) fluorophores at the 3' versus 5' ends, respectively (3'-Cy3-(dT)_{xx}-Cy5-5', where xx = number of nucleotides). A plot of the FRET efficiency versus length of end-labeled poly (dT) substrates shows an excellent monotonic relationship with high and low FRET captured for (dT)₁₅ and (dT)₉₅, respectively (Figure 2). These measurements obtained using a benchtop EI-FLEX microscope are in excellent agreement with similar data collected using a Picoquant MicroTime 200 confocal microscope (Supplementary Figure S1 and Supplementary Table S3).

RPA binding does not produce a complete linearization of ssDNA

To assess the change in ssDNA shape upon RPA binding, we quantified the FRET changes in the RPA–ssDNA complex. First, in these assays, DNA alone produces a FRET signal proportional to the length (Figure 2A). Upon binding to RPA, the FRET signal decreases because of an increase in distance between the donor and acceptor (Figure 3). In addition, since the concentration of RPA and ssDNA used in these experiments are low (pM range), we collected data on each ssDNA substrate as a function of increasing RPA concentrations (Figures 3A–H, S2–S4 and Supplementary Tables S4 and S5). Furthermore, to ensure that we only assessed 1:1 RPA:DNA complexes (molar ratios of RPA trimer bound to one molecule of DNA), we performed parallel mass photometry analysis of the complexes (Figures 3I–L). Under all RPA concentrations tested, we captured predominantly stoichiometric RPA–ssDNA complexes. Thus, our interpretation of end-to-end ssDNA distances from the FRET measurements is not influenced by the binding of multiple RPA molecules. The RPA:DNA binding stoichiometries measured using MP (nM concentrations) are recapitulated in analytical ultracentrifugation analysis at higher protein concentrations (μM; data not shown).

smFRET measurements were performed on four ssDNA substrates of varying lengths [(dT)_{xx}, xx = 15, 25, 30 or 45 nt]. When these measurements are plotted as a function of ssDNA length, the results mirror the data for DNA alone (Figure 4A). However, for every ssDNA substrate, an overall increase in end-to-end distance of $\sim 3.1 \pm 0.2$ nm is observed (Supplementary Table S5 and Figure 4B). This surprising result does not agree with the canonical models described for RPA where the DBDs are arranged in a sequential fashion with the ssDNA stretched in a linear manner. Instead, the shorter end-to-end distance measurements better agree with observations in the structural studies where DNA is wrapped around the DBDs, as shown in Figure 1B (18–20). Thus, ssDNA is not linearized, but rather follows the intrinsic curvature of RPA (Figure 1B). The second observation is that the ~ 3.1 nm shift in distance is observed across the increasing length of ssDNA (Figure 4B). This suggests that DNA occupancy along the curvature of the DBDs is uniform and there is no looping or extrusion of the ssDNA. Finally, the minimal change in distances between the shortest (dT)₁₅ and the longest (dT)₄₅ ssDNA substrates supports our model that the trimerization core (Tri-C) of RPA contributes most to ssDNA binding stability (and in this case contour or path of wrapping) whereas the F–A–B domains are intrinsically more dynamic.

DEER spectroscopy confirms wrapping of DNA along the curvature dictated by the DNA binding domains of RPA

To obtain another independent experimental validation of our findings from the smFRET analysis, we performed double electron-electron resonance (DEER) spectroscopy (52,53). This experiment requires the incorporation of two spin labels, one on each end of the oligonucleotide. 2'-aminouridine was introduced during chemical synthesis of the oligonucleotides and post-synthetically spin-labeled with an isothiocyanate derivative of an isoindoline nitroxide (Figure 5A and Supplementary Table S1) (37). The distance between the two spin centers was measured using DEER spectroscopy in the absence and presence of RPA. For a (dT)₂₂ substrate, DNA alone produces a broad distribution of distances with two populations (Figure 5B and Supplementary Table S6) and these measurements are in excellent agreement with our smFRET data. When RPA is added in equimolar amounts to form a stoichiometric RPA:DNA (1:1) complex, the end-to-end distance increases to $\sim 6.83 \pm 1.5$ nm (Figure 5C). This measurement is again in very good agreement with our smFRET observations (Figures 4 and 5). We also checked if this phenomenon was specific to human RPA by performing DEER and smFRET measurements for the *Saccharomyces cerevisiae* RPA–ssDNA complex. On a (dT)₂₂ oligonucleotide, the end-to-end distance in DEER measurements was 6.9 ± 1.8 nm (Figure 4). In smFRET measurements on a (dT)₂₅ oligonucleotide, we observed an end-to-end distance of 8.5 ± 0.82 nm (Figure 4). These distances agree well with the smFRET measurements. These distances are similar between the human and yeast RPA, suggesting that the mechanism of ssDNA wrapping is likely conserved across eukaryotic RPA.

Computational prediction of the shape of the wrapped ssDNA

In principle, ssDNA should behave as a flexible polymer, in particular when the sequence cannot form secondary structures through intra-strand base-pairing; thus the choice of

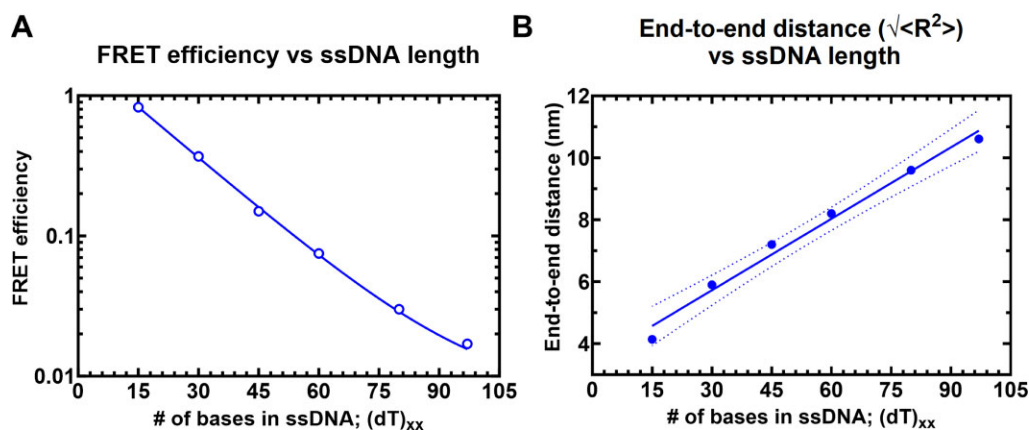


Figure 2. Solution confocal-based FRET measurements accurately report on ssDNA end-to-end distances. **(A)** FRET efficiencies were calculated based on single molecule measurements of Cy3 (donor) and Cy5 (acceptor) fluorescence. Data were collected on a series of ssDNA substrates (30 pM) of increasing lengths. Energy transfer events were detected as bursts of photons as the molecules transited the confocal volume. The mean of the distribution is plotted in panel A as a function of length of poly-thymidine (dT)_{xx}. **(B)** The end-to-end distances were calculated based on a R_0 value of 5.4 nm for the Cy3/Cy5 pair and Eq. (3) described in the Methods. The dotted lines represent a 95% confidence interval for the linear fit.

poly-dT (49). The dependence of the end-to-end distance for a flexible polymer on the number of monomer units (here nucleotides) generally can be described in terms of polymer scaling laws. These laws define the dimensions of unfolded and intrinsically disordered macromolecules and are modeled based on the composition of the nucleotides and salt conditions (54). We therefore sought to determine if our unbound dT constructs showed characteristic polymer scaling behavior. Fitting the root-mean-squared end-to-end distance against the number of bases reveals extremely good agreement with a polymer scaling law of $R_c = R_0 N^\nu$, where R_0 is 1.07 nm and ν is 0.50 (Figure 6A). This would suggest that under the solution conditions examined, ssDNA (dT) behaves as a flexible chain that conforms to the statistical properties of a Gaussian chain. To further assess the validity of our experimental measurements and conclusions, we compared end-to-end distance obtained previously using a confocal-based smFRET set up by Chen *et al.* (55). Gratifyingly, for (dT)₄₀ under matching solution conditions (50 mM NaCl), the data from Chen *et al.* lies directly over the polymer fit generated from our smFRET data (Figure 6A).

If the bound-state conformation were to form a stretched linear extension, we would expect major deviations from the polymer scaling and an increase in distance between the DNA ends. Instead, bound-state derived distances reveal an identical dependency on the number of nucleotides, shifted by ~3 nm more expanded than the unbound state (Figure 6B). Both bound and unbound distances are far from the trend expected for fully linearized DNA. Taken together, our results suggest the configurational/conformational properties of the RPA-bound ssDNA are relatively similar to unbound, with the exception of a systematic shift in the end-to-end distance, consistent with dT wrapping around RPA in a dynamic state that preserves the DNA's intrinsic flexibility (Figure 6C).

Post-translational modification of human RPA by Aurora kinase B rearranges the domains without affecting the wrapping of ssDNA

Since ssDNA wraps around the contour of the DNA-binding domains (DBDs) of RPA, we wondered how RPA-interacting

proteins might gain access to the buried ssDNA. There are two important features that need to be considered: The first is the intrinsic dynamic interactions between the DBDs and ssDNA, and the second is the configurational rearrangements of the domains with respect to each other and the ssDNA. We, and others, have shown that the domains possess different dynamic properties, with DBD-A and DBD-B being more dynamic on ssDNA compared to DBD-C and DBD-D (19,22,23,26). Using C-trap experiments, we recently showed that the rates of diffusion for RPA on long stretches of ssDNA are regulated by the trimerization core (46). Thus, incoming proteins can access the buried ssDNA through transient dissociation of one or more DBDs. The second mode of ssDNA access might be provided through post-translational modifications of the DBDs or the disordered linkers (22,26). Using hydrogen-deuterium exchange mass spectrometry, we recently showed that the domains of RPA are not splayed apart but are tightly organized along with the protein-interaction domains and this configuration is altered upon phosphorylation by Aurora kinase B (22). This modification is specific to RPA functions during mitosis. To gain a better understanding of these configurational changes and how the phosphorylated RPA alters the wrapping of ssDNA, we performed crosslinking mass spectrometry (XL-MS) of RPA and phosphorylated-RPA in the absence/presence of ssDNA. Aurora kinase B phosphorylates RPA at a single Ser-384 position in the large RPA70 subunit (22). Using genetic code expansion (39,40) we generated site-specific phospho-serine (pSer) modified RPA at Ser-384 (RPA-pSer³⁸⁴). This approach produced 100% phosphorylated RPA as seen by complete shift of the RPA-pSer³⁸⁴ band in Phos-Tag SDS-PAGE analysis (Figure 7A). Site-specific incorporation was also confirmed through mass spectrometry (Supplementary Figure S5) and by western blotting using a pSer-384 specific antibody (Figure 7B) (22). XL-MS of RPA or RPA-pSer³⁸⁴ were performed with a bis (sulfosuccinimidyl)suberate (BS3) crosslinker in the absence or presence of a (dT)₂₅ ssDNA substrate. BS3 crosslinks primary amines (Lys residues) that are within 15 Å (56).

We made several key observations: (i) Extensive crosslinks (XLs) are captured between the three subunits (RPA70, RPA32 and RPA14) and between the DNA binding do-

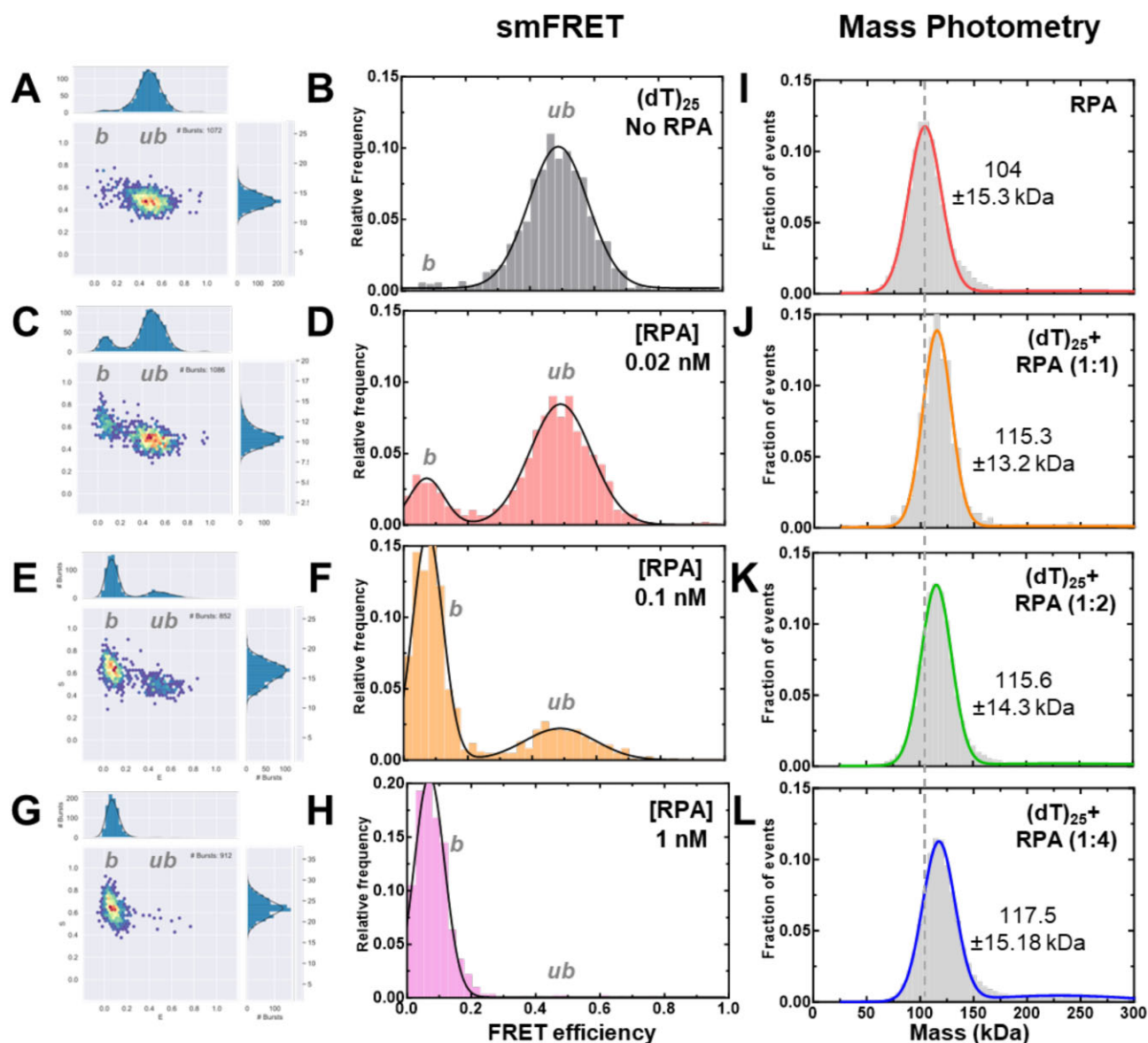


Figure 3. Concentration dependence of RPA-ssDNA complexes. **A-H**) FRET analysis of $(dT)_{25}$ ssDNA bound to increasing concentrations of RPA show a shift from the unbound to bound complex. As RPA concentrations are increased, a complete shift to the bound population is observed. Ratios are defined as one molecule of ssDNA: number of RPA trimers (molar ratio). **I-L**) Mass photometry analysis of RPA and RPA- $(dT)_{25}$ complexes show formation of predominantly single RPA bound $(dT)_{25}$ complexes. The dotted line serves as a reference point for the mass of free RPA in solution as seen in panel I. The measured mass for RPA and the RPA-DNA complexes (1:1 stoichiometry) are noted. In all conditions tested here, one RPA molecule binds to one molecule of ssDNA.

main (DBDs A, B, C and D), the protein interaction domains (OB-F and wh), and several connecting linkers of non-phosphorylated RPA (Figure 7C and Supplementary Figures S6–S9). Of particular interest are the XLs between OB-F and F-A linker with DBD-A, DBD-B, DBD-C, DBD-D and the wh domains. This finding suggests that all the domains are situated close together in a compacted configuration. (ii) Comparison of RPA in the absence and presence of ssDNA shows crosslinks unique to each condition (denoted by the blue lines in Figure 7), suggesting changes in the configurations/conformations or ssDNA occluded Lys residues. However, the overall contacts between the domains are still observed (Figures 7C and D). The data suggest that ssDNA wraps around the compacted structural architecture of RPA without the need to unravel the domains, as would be ex-

pected if the ssDNA were to be linearly stretched. (iii) Surprisingly, introduction of phosphoserine at position 384 (DBD-B; RPA70) produces significantly new contacts within RPA (Figure 7E and Supplementary Figure S6). (iv) Finally, the largest changes in XL patterns are observed upon ssDNA binding to RPA-pSer³⁸⁴ suggesting that a single post-translational modification can bring about large-scale configurational changes leading to an altered ssDNA-RPA-pSer³⁸⁴ complex (Figure 7F and Supplementary Figure S7).

To understand how ssDNA is shaped when bound to RPA-pSer³⁸⁴, we measured the end-to-end distance using DEER (Figure 5). We observe two distributions for RPA-pSer³⁸⁴. The first overlaps with wild-type RPA and ~70% of the phosphorylated RPA is in this population (~7 nm). In addition, there is a 30% increase in the second population where the end-to-

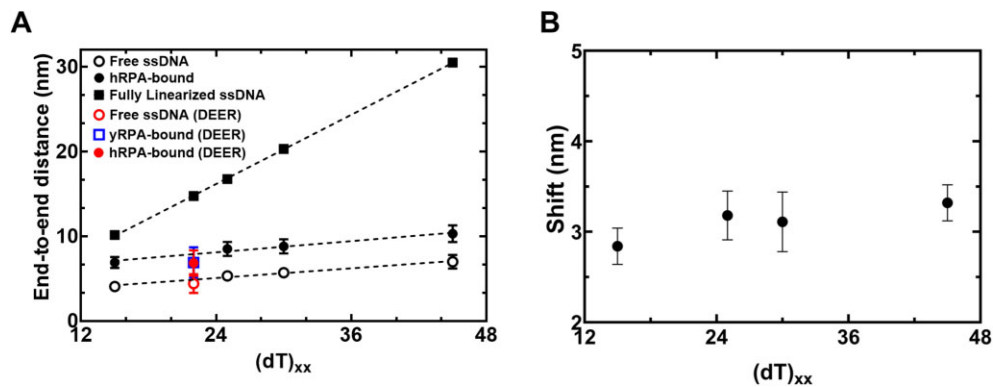


Figure 4. RPA binding produces a modest 3 nm increase in end-to-end distance. **(A)** The estimated end-to-end distance between the 3' and 5' ends of ssDNA is plotted as a function of nucleotides which make up the chain. Closed squares represent theoretically calculated end-to-end distances assuming the ssDNA was completely linearized. Open and closed circles represent experimental end-to-end distance measurements from smFRET analysis ssDNA -/+ RPA, respectively. The figure also shows that end-to-end distances of yRPA-bound, hRPA-bound, and free $(dT)_{22}$ ssDNA, measured using DEER spectroscopy. The DEER measurements fall on the respective trend lines suggesting good agreement between experimental measurements performed using two independent biophysical approaches. **(B)** A $\sim 3.1 \pm 0.2$ nm shift is observed in the end-to-end distances between the free ssDNA and RPA-bound ssDNA. This ssDNA-independent uniform expansion suggests that the average dimensions of the DNA ensemble are increasing due to interactions with RPA.

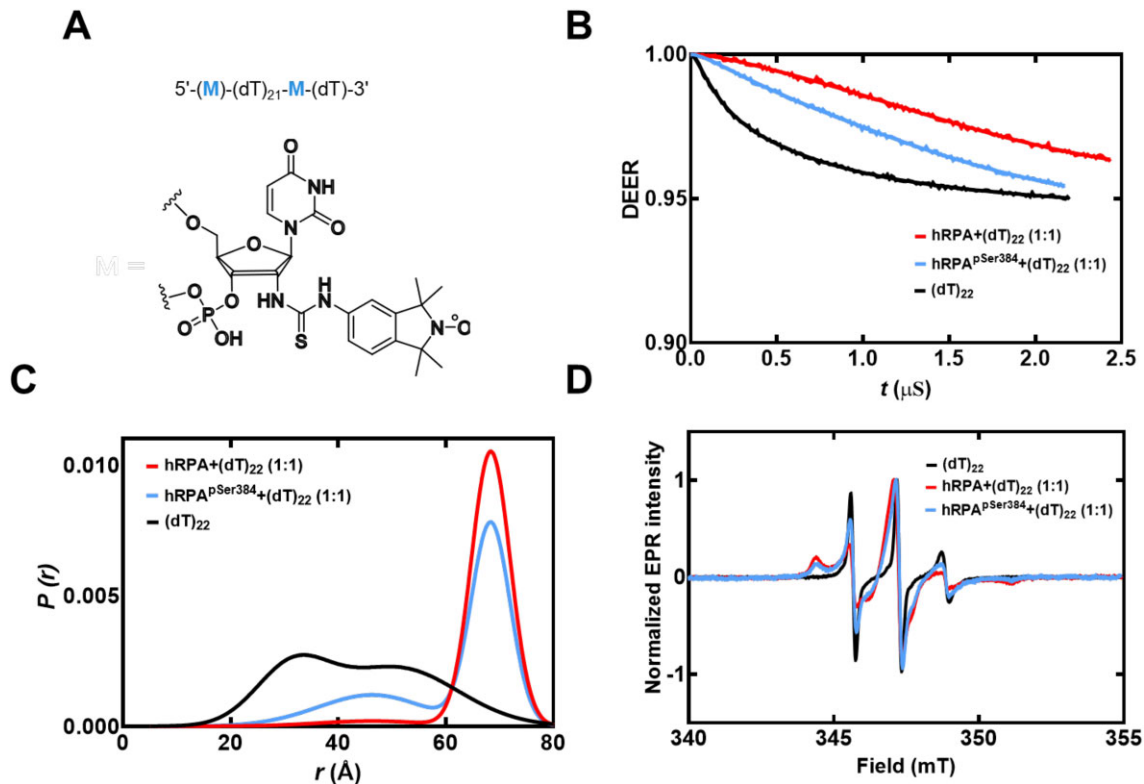


Figure 5. DEER spectroscopy of RPA and RPA-pSer³⁸⁴ bound to ssDNA. **(A)** Structure of the isoindoline nitroxide spin label and the position on the (dT) oligonucleotide. **(B)** Raw DEER decays and fits are presented for the experimentally determined distance distributions $P(r)$ **(C)**. **(D)** CW EPR spectra of labeled ssDNA in the absence and presence of RPA and RPA-pSer³⁸⁴.

end distance is much shorter (~ 5 nm). Along with the XL-MS data, we interpret these changes as remodeling of the DBDs without large-scale changes to the wrapping pattern of the ssDNA. Thus, from a functional standpoint, phosphorylation remodels RPA such that segments of ssDNA are made accessible to RPA-interacting proteins with minimal alterations to the wrapping of DNA. For RPA-pSer³⁸⁴, these interactions are specific to mitosis.

Coarse-grained MD simulations showcase the binding properties of one versus two RPA-bound ssDNA complexes

The end-to-end distances we have thus far experimentally measured were for free ssDNA or ssDNA bound to a single RPA. Due to limitations in the maximal end-to-end distances that can be measured using FRET or DEER, we were not able to extend these investigations to scenarios where more than

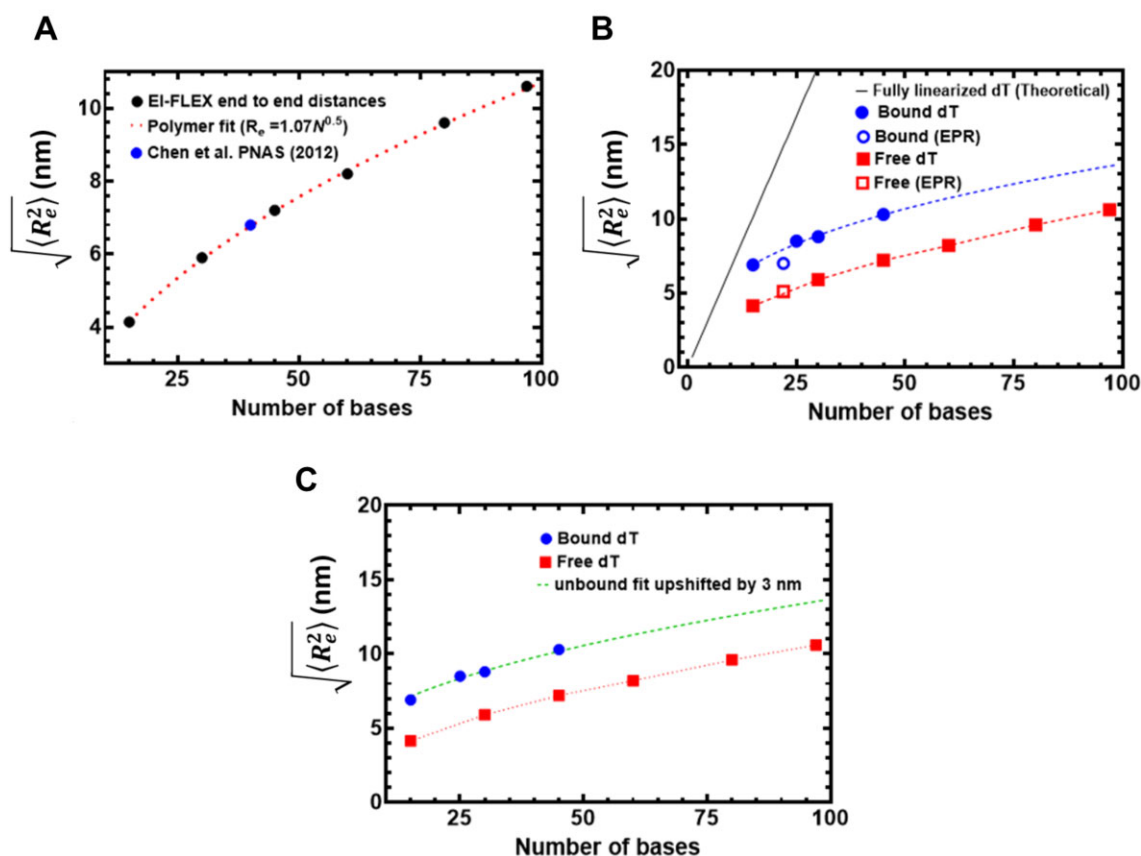


Figure 6. Computational predictions of ssDNA wrapping agree well with experimental measurements. **(A)** Black circles are data, and the red dashed line is a fit to those data using a polymer scaling model. The blue circle is data reporting on the end-to-end distance for (dT)₄₀ made by Chen et al. 2011. **(B)** Blue circles are from RPA-bound (dT) oligonucleotides, red squares are from RPA-free (dT) oligonucleotides. Data include distances measured by both EPR and smFRET experiments. Blue and red dashed lines are polymer model fits to the data, respectively. The black line is the theoretical expected end-to-end distance if the DNA were fully linearized and stretched to its contour length. **(C)** The root mean squared end-to-end distance dependence on the number of bases is essentially identical for the bound and unbound states, as highlighted by the fact that the exact polymer fit upshifted by 3.1 ± 0.2 nm fully describes the bound dT dependency.

one RPA is bound. So, we turned to coarse-grain molecular dynamics (MD) simulations to gain a better understanding of such complexes. First, free ssDNA of varying lengths was simulated and the end-to-end distances across varying lengths of ssDNA ranging from 15 to 80 nucleotides are shown (Figure 8A). The data shown in pink (Figure 8A) attempts to replicate the scenario observed in the original crystal structure PDB: 4GNX (18), where the central 23 nucleotides were constrained to a distance of 55 Å for the ssDNA, without the presence of RPA in the simulation. In contrast, the data in grey (Figure 8A) illustrates the extreme scenario where the 23 nucleotides interacting with RPA are arranged linearly, while the remaining ssDNA is free to move. This simulation was conducted solely with ssDNA in the absence of RPA. Both the grey and pink plots were simulated for DNA lengths ranging from 25 to 100 nucleotides (Figure 8A). Next, we performed simulations with two RPA molecules at different initial configurations, along with varying lengths of ssDNA ranging from 60 to 100 nucleotides. The distribution of end-to-end distances for free ssDNA, ssDNA bound to a single RPA, and ssDNA bound to two RPAs on a (dT)₈₀ ssDNA shows a significant overlap of end-to-end distances in the three mentioned cases (Figure 8B). Additionally, snapshots from the simulations illustrating some instances of end-to-end distances for free ssDNA, ssDNA bound to a single RPA, and ssDNA bound to two

RPAs are shown (Figure 8C and Supplementary Movies S1–S3). These movies provide an excellent view of how ssDNA is wrapped along the DNA binding domains of RPA while retaining the intrinsic flexibility and curvature that are experimentally captured in smFRET and DEER analysis.

Discussion

RPA binds ssDNA with high affinity and coats the substrate to form a regulatory protein-interaction hub to orchestrate a variety of DNA metabolic processes (1,3). RPA possesses multiple DNA binding and protein interaction domains that are connected by disordered linkers and spread across its heterotrimeric architecture. An incoming RPA-interacting protein could access the ssDNA buried under RPA by remodeling on one or more domains without the need to displace RPA (57,58). To better understand how such remodeling occurs, knowledge of how ssDNA is bound by RPA is required. Canonically, models for RPA postulate that the DNA binding domains (DBDs) are assembled in a linear array to stretch the DNA leading to multiple binding modes (59). In contrast, structural studies show that the ssDNA is bent around the DBDs (18–20). In solution, an ensemble of states ranging from linear to the bent form can also be envisioned to exist in equilibrium (57,60).

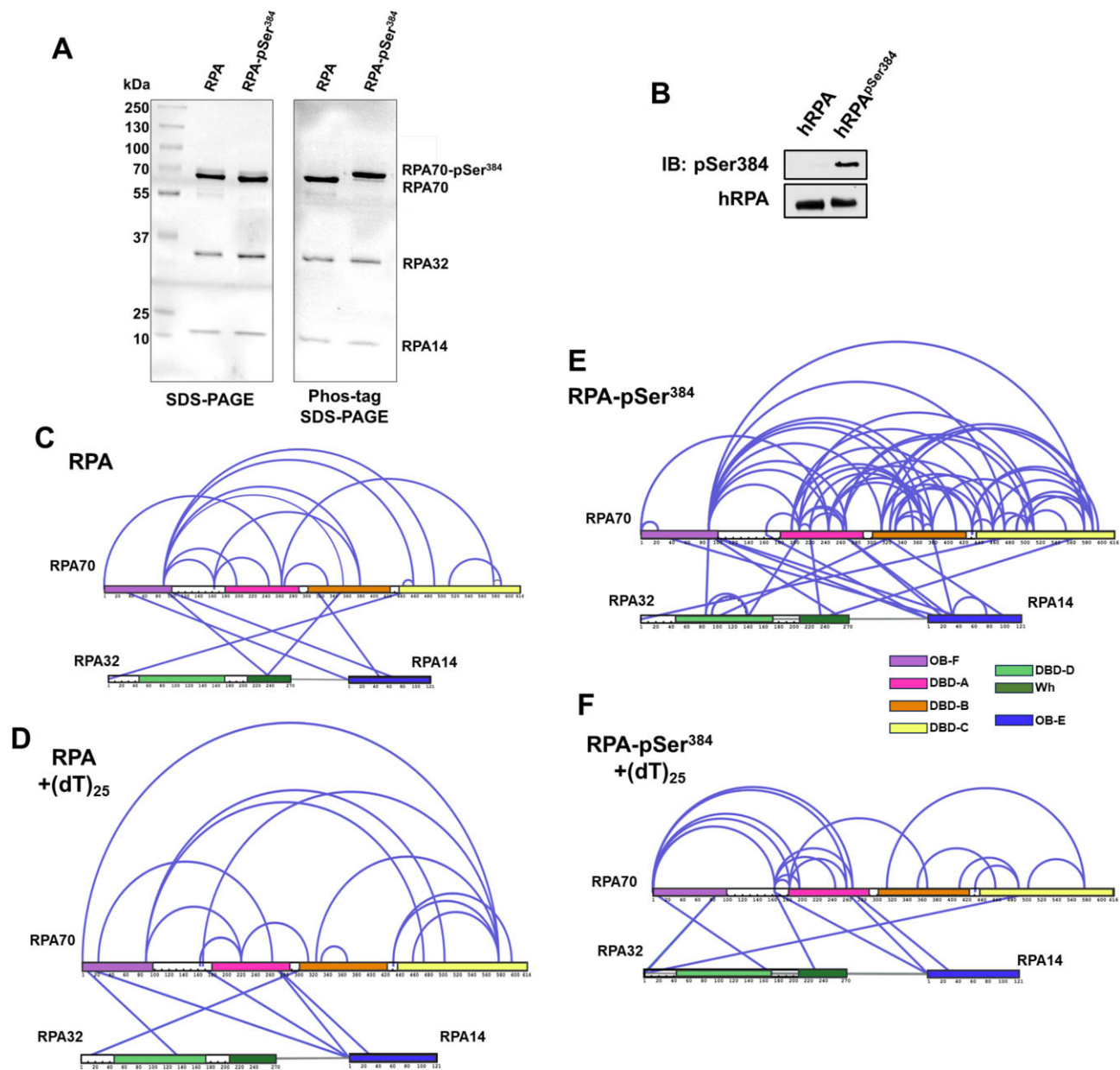


Figure 7. Phosphorylation at a single position in RPA70 alters access to the ssDNA with minimal alternations of end-to-end distance. **(A)** SDS-PAGE and Phos-tag SDS-PAGE analysis of RPA and RPA-pSer³⁸⁴ proteins show selective shift of the RPA70 band only in the RPA-pSer³⁸⁴ sample confirming 100% incorporation of pSer. **(B)** Western blotting of RPA and RPA-pSer³⁸⁴ with an antibody specific to pSer³⁸⁴ confirms the site-specific phosphorylation. **(C, D)** Cross-linking mass spectrometry (XLMS) analysis of RPA in the absence or presence of ssDNA (dT)₂₅. A direct comparison of the XLS between RPA and the RPA-(dT)₂₅ experiments is presented with XLS unique to each condition denoted in blue. **(E)** A similar comparative XLMS analysis RPA-pSer³⁸⁴ and the RPA-pSer³⁸⁴-(dT)₂₅ complex is shown and the XLS unique to each condition is shown in blue. The XL patterns show differences in ssDNA driven changes upon phosphorylation.

From the perspective of the binding properties of the individual domains, recent data from our group and others support the idea of RPA behaving as two functional halves with a dynamic F-A-B half (OB-F, DBD-A and DBD-B) and a less-dynamic Tri-C half (DBD-C, RPA32 and RPA14) (1,19,22,26,27,61). Here Tri-C is modeled to provide ssDNA binding stability to RPA. Support for this model also arises from single-molecule C-trap experiments where the diffusion rate of RPA is dictated by the Tri-C half (46). Here, we show that ssDNA wraps around RPA and end-to-end distance measurements suggest uniform contact through-

out the ssDNA-RPA complex. The data argues against linearization of ssDNA and an array-like assembly of the DBDs.

Crosslinking MS data supports a structural model for RPA where both the DBDs and protein-interaction domains (PIDs) are in close proximity. Crosslinks were also detected between the disordered OB-F:DBD-A linker and almost all DBDs and PIDs, suggesting a compacted structure for human RPA (Figure 8 and Supplementary Figure S6). The end-to-end distance measurements, performed as a function of ssDNA length, mirror that of free ssDNA in solution with a ~3 nm shift. This

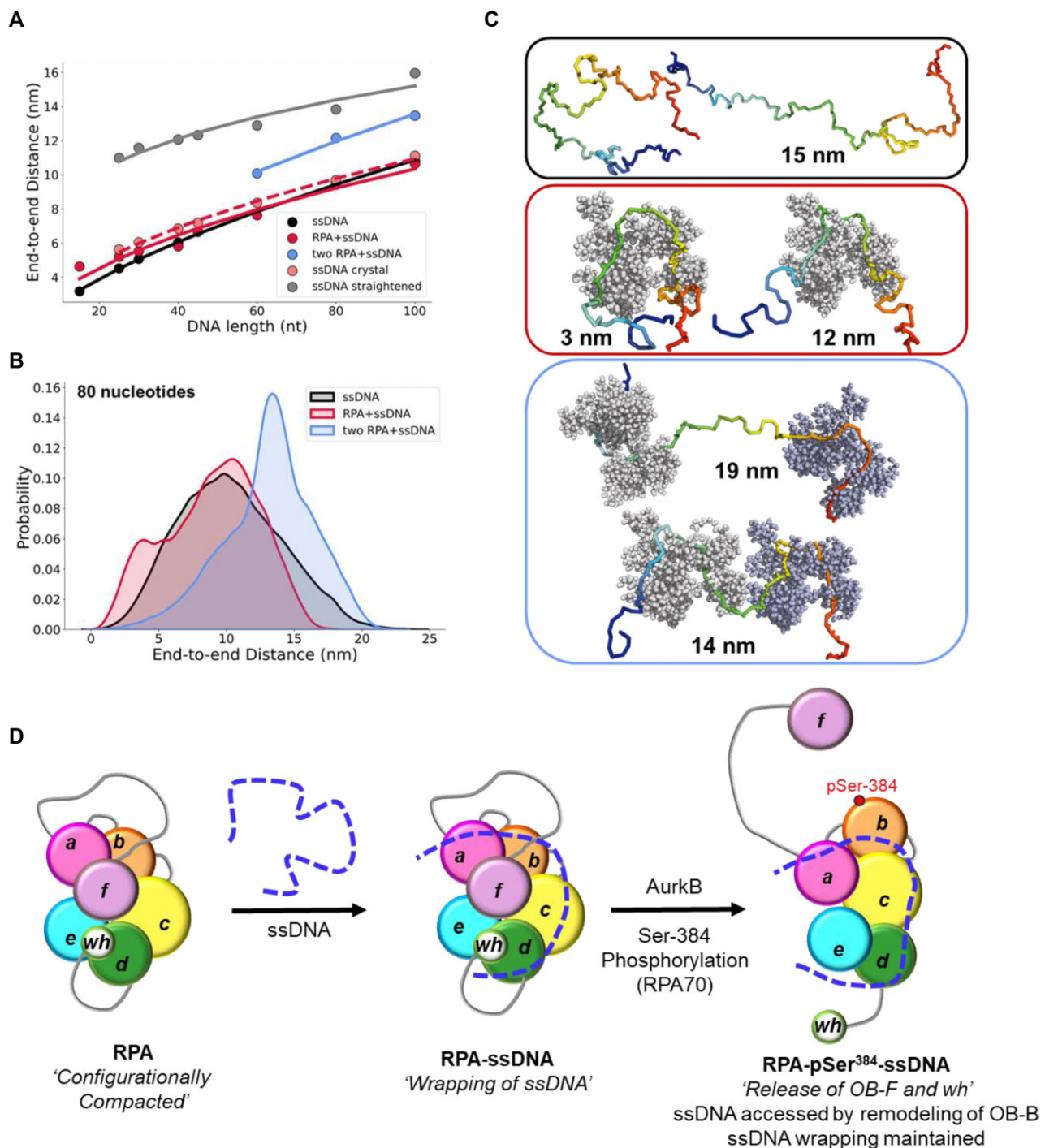


Figure 8. Conformational analysis of ssDNA-RPA complexes from coarse-grained molecular dynamics simulation and a model for ssDNA wrapping by RPA. **(A)** Mean end-to-end distances for ssDNA of varying lengths in different conditions: free ssDNA (black), ssDNA bound to a single RPA (red), and ssDNA bound to two RPAs (blue). Additionally, ssDNA was studied in isolation (i.e. in the absence of explicit RPA) but with a restraint applied on the end-to-end distance of the middle 23 nucleotides, which was set to be 55 Å (mimicking curved binding to RPA) or 138 Å (mimicking stretched conformations of the ssDNA). The coarse-grained simulations were based on *Ustilago maydis* RPA (PDB: 4GNX), but the five OB-domains were considered as non-dynamic for simplicity. **(B)** Distribution of end-to-end distances of ssDNA (dT)₈₀ nucleotides in length, modeled as free ssDNA (black), ssDNA bound to a single RPA (red), and ssDNA bound to two RPAs (blue). The distributions are based on multiple long simulations for each system. The distribution of the end-to-end distances for ssDNA interacting with two RPA molecules combines simulations in which the excess ssDNA is placed either as a linker between the two RPAs or as flanking ssDNA at the ends. **(C)** Snapshots from coarse-grained simulations showing varying end-to-end distances for 80-nucleotide ssDNA in different states: free ssDNA, bound to a single RPA, and bound to two RPAs. Please also refer to videos provided in the Supplemental Information. **(D)** The OB-domains of RPA are depicted along with the connecting disordered linkers. The XL-MS data suggest that the domains are compacted together, and ssDNA is wrapped around this architecture. The domains are remodeled upon phosphorylation by Aurora kinase B at position Ser-384 in the RPA70 subunit. This modification releases the OB-F and wh domains and promotes RPA interactions with other proteins. While the wrapping of ssDNA is not altered, remodeling allows access to the ssDNA.

data suggest that ssDNA wraps around the compacted RPA core without much of a change in the intrinsic flexibility of ssDNA, likely owing to the dynamic interactions with RPA. The fact our data fit well to a flexible polymer model in both the bound and unbound states (i.e. end-to-end distance scales with a number of nucleotides proportional to a scaling exponent) suggests that the DNA is not in a fixed structure when bound to RPA (or when unbound, as expected), but is highly flexible, and exists in a collection of different conformations that rapidly interconvert between one another. As such, the uniform expansion suggests that the average dimension of this ensemble increases. In other words, the DNA is not ‘stretching’ because this implies it is tightly extended, instead, the average dimensions of the DNA ensemble are increasing due to interactions with RPA. This explanation fits perfectly with the wrapped conformation observed in the structural studies where the ssDNA follows along the curvature imposed by the DNA binding domains with the Tri-C providing the central stability to maintaining this architecture. From the context of the individual domains, the more-dynamic versus less-dynamic ssDNA binding attributes assigned to the domains of RPA (23) also explain how the OB-A & OB-B domains can be dynamic with rapid on-off rates while the curvature of the ssDNA is maintained by the less-dynamic Tri-C core (composed of OB-C, OB-D and OB-E).

In such a model, access to internal regions in the ssDNA or the ends can be provided to incoming RPA-interacting proteins by rearranging one or more DBDs (62,63). Such rearrangements could be promoted by post-translational modification (s) of one or more DBDs (Figure 8). As a proof of concept, we here show how phosphorylation of RPA at a single position in DBD-B (Ser-384 in RPA70) drives changes in ssDNA access. Aurora kinase B phosphorylates RPA at Ser-384 during mitosis to suppress homologous recombination and to facilitate chromosome segregation by maintaining Aurora B activity (22). Surprisingly, the end-to-end ssDNA distance does not change in RPA-pSer³⁸⁴, but the patterns of crosslinking between the domains in XL-MS are strikingly altered. Thus, the DBDs and PIDs have been repositioned or remodeled through phosphorylation without altering the overall wrapping path of ssDNA (Figure 8). Such changes would grant access to internal regions of the ssDNA wrapped by RPA upon phosphorylation.

In summary, measurement of end-to-end distances using smFRET and DEER spectroscopy supports a model where ssDNA is wrapped around a compacted structure of RPA with remodeling of the domains enacted through post-translational modifications. Thus, RPA can be differentially modulated to serve varying DNA metabolic needs without major changes to the ssDNA organization within the complex. How these changes transpire within the context of multiple RPA molecules bound to longer ssDNA remains to be established. The MD simulations presented here show how multiple RPA molecules can be engaged on longer ssDNA while retaining the wrapped architecture within the DNA binding domains. At the moment, experimentally capturing how ssDNA wrapping or stretching changes upon binding of multiple RPA molecules is not feasible as the changes in end-to-end distances described here are already at the limits of detection for the smFRET and DEER methodologies. Thus, structural approaches such as CryoEM will be needed to further address such questions.

Data availability

Raw data are available in the associated Supplemental Data File. Plasmids for protein overproduction as available upon request. Code for computational modeling can be accessed through these repositories: https://github.com/holehouse-lab/supportingdata/tree/master/2024/chadda_kaushik_2024 and <https://doi.org/10.5281/zenodo.12205956>.

Supplementary data

Supplementary Data are available at NAR Online.

Acknowledgements

Authors thank the lab members and Dr Timothy Craggs (Exciting Instruments) for technical advice and critical reading of the manuscript. We thank Dr Greg Sabat, University of Wisconsin-Madison, for phospho-proteomic MS analysis. XL-MS data were obtained at the Mass Spectrometry Technology Access Center (MTAC) at Washington University in St. Louis. We thank Dr Byoung-Kyu Cho at MTAC for collecting the XL-MS spectra.

Author contributions Rahul Chadda: Formal analysis, Methodology, Validation, Writing—original draft. Vikas Kaushik: Protein Purification, pSer incorporation, XL-MS experiments and data analysis, Writing—review & editing. Jaigeeth Deveryshetty, Edwin Antony, & Brian Bothner: Cross-linking mass spectrometry and structural analysis, Writing—review & editing. Alex Holehouse: Formal analysis, Methodology, Validation, Writing—original draft. Iram Munir Ahmad and Snorri Th.d Sigurdsson: Design and synthesis of spin-labeled oligonucleotides, Writing—review & editing. Gargi Biswas and Yaakov Levy: MD simulation of RPA-ssDNA interactions, Writing—review & editing. Richard Cooley and Ryan Mehl: Design of cells and strategies for pSer-incorporation. Writing—review & editing. Reza Dastvan: EPR - Formal analysis, Methodology, Validation, Writing—review & editing. Sofia Origanti: Conceptualization, Formal analysis, Methodology, Validation, Writing—review & editing. Edwin Antony: Conceptualization, Formal analysis, Methodology, Validation, Writing—original draft.

Funding

National Institutes of Health [R35-GM149320 and S10-OD030343 to E.A., R01-GM145783 to R.D., R01-GM143179 to S.O., DP2-CA290639-01 to A.S.H.]; S.Th.S. acknowledges financial support from the Icelandic Research Fund [206708]; Acquisition of the EI-FLEX microscope was supported through a grant from the Department of Energy, Office of Basic Energy Sciences [DE-SC0020965 to E.A.]; Y.L. was supported by grants from the Israeli Science Foundation [2072/22]; Estate of Gerald Alexander; pSer incorporation was supported in part by the GCE4All Biomedical Technology Development and Dissemination Center supported by National Institute of General Medical Science grant [RM1-GM144227 to R.M]; XL-MS data collection supported by the Washington University Institute of Clinical and Translational Sciences which is, in part, supported by the NIH/National Center for Advancing Translational Sciences (NCATS), CTSA [UL1TR002345]. Funding for open access charge: National Institutes of Health [R35-GM149320].

Conflict of interest statement

None declared.

References

- Caldwell,C.C. and 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.
- Wold,M.S. and Kelly,T. (1988) Purification and characterization of replication protein A, a cellular protein required for in vitro replication of simian virus 40 DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **85**, 2523–2527.
- Wold,M.S. (1997) Replication protein A: A heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu. Rev. Biochem.*, **66**, 61–92.
- Iftode,C., Daniely,Y. and Borowiec,J.A. (1999) Replication protein A (RPA): the eukaryotic SSB. *Crit. Rev. Biochem. Mol. Biol.*, **34**, 141–180.
- Kim,C., Paulus,B.F. and Wold,M.S. (1994) Interactions of human replication protein A with oligonucleotides. *Biochemistry*, **33**, 14197–14206.
- Kumaran,S., Kozlov,A.G. and Lohman,T.M. (2006) Saccharomyces cerevisiae replication protein A binds to single-stranded DNA in multiple salt-dependent modes. *Biochemistry*, **45**, 11958–11973.
- Treuner,K., Ramsperger,U. and Knippers,R. (1996) Replication protein A induces the unwinding of long double-stranded DNA regions. *J. Mol. Biol.*, **259**, 104–112.
- Lao,Y., Lee,C.G. and Wold,M.S. (1999) Replication protein A interactions with DNA. 2. Characterization of double-stranded DNA-binding/helix-destabilization activities and the role of the zinc-finger domain in DNA interactions. *Biochemistry*, **38**, 3974–3984.
- Bartos,J.D., Willmott,L.J., Binz,S.K., Wold,M.S. and Bambara,R.A. (2008) Catalysis of strand annealing by replication protein A derives from its strand melting properties. *J. Biol. Chem.*, **283**, 21758–21768.
- Salas,T.R., Petruseva,I., Lavrik,O., Bourdoncle,A., Mergny,J.L., Favre,A. and Saintome,C. (2006) Human replication protein A unfolds telomeric G-quadruplexes. *Nucleic Acids Res.*, **34**, 4857–4865.
- Marechal,A. and 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.
- Namiki,Y. and Zou,L. (2006) ATRIP associates with replication protein A-coated ssDNA through multiple interactions. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 580–585.
- Zou,L. and Elledge,S.J. (2003) Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science*, **300**, 1542–1548.
- Zou,L., Liu,D. and Elledge,S.J. (2003) Replication protein A-mediated recruitment and activation of Rad17 complexes. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 13827–13832.
- de Laat,W.L., Appeldoorn,E., Sugawara,K., Weterings,E., Jaspers,N.G. and Hoeijmakers,J.H. (1998) DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Genes Dev.*, **12**, 2598–2609.
- Acharya,A., Kasaciunaite,K., Gose,M., Kissling,V., Guerois,R., Seidel,R. and Cejka,P. (2021) Distinct RPA domains promote recruitment and the helicase-nuclease activities of Dna2. *Nat. Commun.*, **12**, 6521.
- Zhou,C., Pourmal,S. and Pavletich,N.P. (2015) Dna2 nuclease-helicase structure, mechanism and regulation by Rpa. *eLife*, **4**, e09832.
- Fan,J. and Pavletich,N.P. (2012) Structure and conformational change of a replication protein A heterotrimer bound to ssDNA. *Genes Dev.*, **26**, 2337–2347.
- Yates,L.A., Aramayo,R.J., Pokhrel,N., Caldwell,C.C., Kaplan,J.A., Perera,R.L., Spies,M., Antony,E. and Zhang,X. (2018) A structural and dynamic model for the assembly of replication protein A on single-stranded DNA. *Nat. Commun.*, **9**, 5447.
- Madru,C., Martinez-Carranza,M., Laurent,S., Alberti,A.C., Chevreuil,M., Raynal,B., Haouz,A., Le Meur,R.A., Delarue,M., Henneke,G., et al. (2023) DNA-binding mechanism and evolution of replication protein A. *Nat. Commun.*, **14**, 2326.
- Bochkareva,E., Korolev,S., Lees-Miller,S.P. and Bochkareva,A. (2002) Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA. *EMBO J.*, **21**, 1855–1863.
- Roshan,P., Kuppa,S., Mattice,J.R., Kaushik,V., Chadda,R., Pokhrel,N., Tumala,B.R., Biswas,A., Bothner,B., Antony,E., et al. (2023) An Aurora B-RPA signaling axis secures chromosome segregation fidelity. *Nat. Commun.*, **14**, 3008.
- Ahmad,F., Patterson,A., Deveryshetty,J., Mattice,J.R., Pokhrel,N., Bothner,B. and Antony,E. (2021) Hydrogen-deuterium exchange reveals a dynamic DNA-binding map of replication protein A. *Nucleic Acids Res.*, **49**, 1455–1469.
- Chen,J., Le,S., Basu,A., Chazin,W.J. and Yan,J. (2015) Mechanochemical regulations of RPA's binding to ssDNA. *Sci. Rep.*, **5**, 9296.
- Sugitani,N. and Chazin,W.J. (2015) Characteristics and concepts of dynamic hub proteins in DNA processing machinery from studies of RPA. *Prog. Biophys. Mol. Biol.*, **117**, 206–211.
- Pokhrel,N., Caldwell,C.C., Corless,E.I., Tillison,E.A., Tibbs,J., Jovic,N., Tabei,S.M.A., Wold,M.S., Spies,M. and Antony,E. (2019) Dynamics and selective remodeling of the DNA-binding domains of RPA. *Nat. Struct. Mol. Biol.*, **26**, 129–136.
- Pokhrel,N., Origanti,S., Davenport,E.P., Gandhi,D., Kaniecki,K., Mehl,R.A., Greene,E.C., Dockendorff,C. and Antony,E. (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.
- San Filippo,J., Sung,P. and Klein,H. (2008) Mechanism of eukaryotic homologous recombination. *Annu. Rev. Biochem.*, **77**, 229–257.
- Heyer,W.D., Ehmsen,K.T. and Liu,J. (2010) Regulation of homologous recombination in eukaryotes. *Annu. Rev. Genet.*, **44**, 113–139.
- Symington,L.S. (2014) End resection at double-strand breaks: mechanism and regulation. *Cold Spring Harb. Perspect. Biol.*, **6**, a016436.
- Sugawara,N. and Haber,J.E. (1992) Characterization of double-strand break-induced recombination: homology requirements and single-stranded DNA formation. *Mol. Cell. Biol.*, **12**, 563–575.
- Nguyen,B., Sokoloski,J., Galletto,R., Elson,E.L., Wold,M.S. and Lohman,T.M. (2014) Diffusion of human replication protein A along single-stranded DNA. *J. Mol. Biol.*, **426**, 3246–3261.
- Schiemann,O., Heubach,C.A., Abdullin,D., Ackermann,K., Azarkh,M., Bagryanskaya,E.G., Drescher,M., Endeward,B., Freed,J.H., Galazzo,L., et al. (2021) Benchmark test and guidelines for DEER/PELDOR experiments on nitroxide-labeled biomolecules. *J. Am. Chem. Soc.*, **143**, 17875–17890.
- Stelzl,L.S., Erlenbach,N., Heinz,M., Prisner,T.F. and Hummer,G. (2017) Resolving the conformational dynamics of DNA with angstrom resolution by pulsed electron-electron double resonance and molecular dynamics. *J. Am. Chem. Soc.*, **139**, 11674–11677.
- Marko,A., Denysenkov,V., Margraf,D., Cekan,P., Schiemann,O., Sigurdsson,S.T. and Prisner,T.F. (2011) Conformational flexibility of DNA. *J. Am. Chem. Soc.*, **133**, 13375–13379.
- Prisner,T.F., Marko,A. and Sigurdsson,S.T. (2015) Conformational dynamics of nucleic acid molecules studied by PELDOR spectroscopy with rigid spin labels. *J. Magn. Reson.*, **252**, 187–198.
- Saha,S., Jagtap,A.P. and Sigurdsson,S.T. (2015) Site-directed spin labeling of 2'-amino groups in RNA with isoindoline nitroxides

- that are resistant to reduction. *Chem. Commun.*, **51**, 13142–13145.
38. Binz, S.K., Dickson, A.M., Haring, S.J. and Wold, M.S. (2006) Functional assays for replication protein A (RPA). *Methods Enzymol.*, **409**, 11–38.
 39. Zhu, P., Mehl, R.A. and Cooley, R.B. (2022) Site-specific incorporation of phosphoserine into recombinant proteins in *Escherichia coli*. *Bio. Protoc.*, **12**, e4541.
 40. Rogerson, D.T., Sachdeva, A., Wang, K., Haq, T., Kazlauskaitė, A., Hancock, S.M., Huguenin-Dezot, N., Muqit, M.M., Fry, A.M., Bayliss, R., *et al.* (2015) Efficient genetic encoding of phosphoserine and its nonhydrolyzable analog. *Nat. Chem. Biol.*, **11**, 496–503.
 41. Zhu, P., Gafken, P.R., Mehl, R.A. and Cooley, R.B. (2019) A highly versatile expression system for the production of multiply phosphorylated proteins. *ACS Chem. Biol.*, **14**, 1564–1572.
 42. Ingargiola, A., Lerner, E., Chung, S., Weiss, S. and Michalet, X. (2016) FRETbursts: an open source toolkit for analysis of freely-diffusing single-molecule FRET. *PLoS One*, **11**, e0160716.
 43. Hustedt, E.J., Stein, R.A. and McHaourab, H.S. (2021) Protein functional dynamics from the rigorous global analysis of DEER data: conditions, components, and conformations. *J. Gen. Physiol.*, **153**, e201711954.
 44. Graham, M., Combe, C., Kolbowski, L. and Rappsilber, J. (2019) xiView: a common platform for the downstream analysis of Crosslinking Mass Spectrometry data. bioRxiv doi: <https://doi.org/10.1101/561829>, 26 February 2019, preprint: not peer reviewed..
 45. Hoitsma, N.M., Norris, J., Khoang, T.H., Kaushik, V., Chadda, R., Antony, E., Hedglin, M. and Freudenthal, B.D. (2023) Mechanistic insight into AP-endonuclease 1 cleavage of abasic sites at stalled replication fork mimics. *Nucleic Acids Res.*, **51**, 6738–6753.
 46. Pangeni, S., Biswas, G., Kaushik, V., Kuppa, S., Yang, O., Lin, C.T., Mishra, G., Levy, Y., Antony, E. and Ha, T. (2024) Rapid long-distance migration of RPA on single stranded DNA occurs through intersegmental transfer utilizing multivalent interactions. *J. Mol. Biol.*, **436**, 168491.
 47. Mishra, G., Bigman, L.S. and Levy, Y. (2020) ssDNA diffuses along replication protein A via a reptation mechanism. *Nucleic Acids Res.*, **48**, 1701–1714.
 48. Hellenkamp, B., Schmid, S., Doroshenko, O., Opanasyuk, O., Kuhnemuth, R., Rezaei Adariani, S., Ambrose, B., Aznauryan, M., Barth, A., Birkedal, V., *et al.* (2018) Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study. *Nat. Methods*, **15**, 669–676.
 49. Murphy, M.C., Rasnik, I., Cheng, W., Lohman, T.M. and Ha, T. (2004) Probing single-stranded DNA conformational flexibility using fluorescence spectroscopy. *Biophys. J.*, **86**, 2530–2537.
 50. Hohlbein, J., Craggs, T.D. and Cordes, T. (2014) Alternating-laser excitation: single-molecule FRET and beyond. *Chem. Soc. Rev.*, **43**, 1156–1171.
 51. Ambrose, B., Baxter, J.M., Cully, J., Willmott, M., Steele, E.M., Bateman, B.C., Martin-Fernandez, M.L., Cadby, A., Shewring, J., Aaldering, M., *et al.* (2020) The smfBox is an open-source platform for single-molecule FRET. *Nat. Commun.*, **11**, 5641.
 52. Heinz, M., Erlenbach, N., Stelzl, L.S., Thierolf, G., Kamble, N.R., Sigurdsson, S.T., Prisner, T.F. and Hummer, G. (2020) High-resolution EPR distance measurements on RNA and DNA with the non-covalent G spin label. *Nucleic Acids Res.*, **48**, 924–933.
 53. Igbaria-Jaber, Y., Hofmann, L., Gevorkyan-Airapetov, L., Shenberger, Y. and Ruthstein, S. (2023) Revealing the DNA Binding Modes of CsoR by EPR Spectroscopy. *ACS Omega*, **8**, 39886–39895.
 54. Hofmann, H., Soranno, A., Borgia, A., Gast, K., Nettels, D. and Schuler, B. (2012) Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 16155–16160.
 55. Chen, H., Meisburger, S.P., Pabit, S.A., Sutton, J.L., Webb, W.W. and Pollack, L. (2012) Ionic strength-dependent persistence lengths of single-stranded RNA and DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 799–804.
 56. Gong, Z., Ye, S.X. and Tang, C. (2020) Tightening the Crosslinking Distance Restraints for Better Resolution of Protein Structure and Dynamics. *Structure*, **28**, 1160–1167.
 57. Wieser, T.A. and Wuttke, D.S. (2022) Replication protein A utilizes differential engagement of its DNA-binding domains to bind biologically relevant ssDNAs in diverse binding modes. *Biochemistry*, **61**, 2592–2606.
 58. Fanning, E., Klimovich, V. and Nager, A.R. (2006) A dynamic model for replication protein A (RPA) function in DNA processing pathways. *Nucleic Acids Res.*, **34**, 4126–4137.
 59. Chen, R., Subramanyam, S., Elcock, A.H., Spies, M. and Wold, M.S. (2016) Dynamic binding of replication protein a is required for DNA repair. *Nucleic Acids Res.*, **44**, 5758–5772.
 60. Wang, Q.M., Yang, Y.T., Wang, Y.R., Gao, B., Xi, X.G. and Hou, X.M. (2019) Human replication protein A induces dynamic changes in single-stranded DNA and RNA structures. *J. Biol. Chem.*, **294**, 13915–13927.
 61. Kuppa, S., Deveryshetty, J., Chadda, R., Mattice, J.R., Pokhrel, N., Kaushik, V., Patterson, A., Dhingra, N., Pangeni, S., Sadauskas, M.K., *et al.* (2022) Rtt105 regulates RPA function by configurationally stapling the flexible domains. *Nat. Commun.*, **13**, 5152.
 62. Topolska-Wos, A.M., Sugitani, N., Cordoba, J.J., Le Meur, K.V., Le Meur, R.A., Kim, H.S., Yeo, J.E., Rosenberg, D., Hammel, M., Schärer, O.D., *et al.* (2020) A key interaction with RPA orients XPA in NER complexes. *Nucleic Acids Res.*, **48**, 2173–2188.
 63. Mer, G., Bochkarev, A., Gupta, R., Bochkareva, E., Frappier, L., Ingles, C.J., Edwards, A.M. and Chazin, W.J. (2000) Structural basis for the recognition of DNA repair proteins UNG2, XPA, and RAD52 by replication factor RPA. *Cell*, **103**, 449–456.