Biochemistry

pubs.acs.org/biochemistry Article

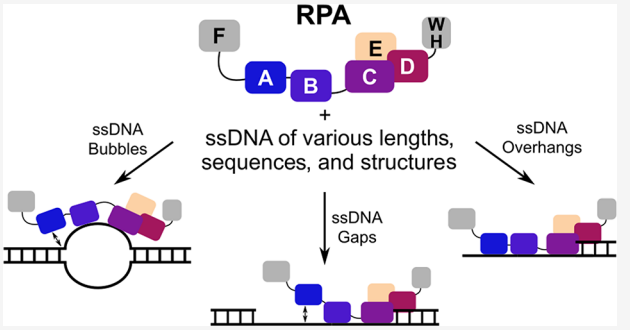
Replication Protein A Utilizes Differential Engagement of Its DNA-Binding Domains to Bind Biologically Relevant ssDNAs in Diverse Binding Modes

Thomas A. Wieser and Deborah S. Wuttke*

Cite This: https://doi.org/10.1021/acs.biochem.2c00504 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Replication protein A (RPA) is a ubiquitous ssDNA-binding protein that functions in many DNA processing pathways to maintain genome integrity. Recent studies suggest that RPA forms a highly dynamic complex with ssDNA that can engage with DNA in many modes that are orchestrated by the differential engagement of the four DNA-binding domains (DBDs) in RPA. To understand how these modes influence RPA interaction with biologically relevant ligands, we performed a comprehensive and systematic evaluation of RPA's binding to a diverse set of ssDNA ligands that varied in sequence, length, and structure. These equilibrium binding data show that WT RPA binds structured ssDNA ligands differently from its engagement with minimal ssDNAs. Next, we investigated each DBD's contributions to RPA's binding modes through mutation of conserved, functionally important aromatic residues. Mutations in DBD-A and -B have a much larger effect on binding when ssDNA is embedded into DNA secondary structures compared to their association with unstructured minimal ssDNA. As our data support a complex interplay of binding modes, it is critical to define the trimer core DBDs' role in binding these biologically relevant ligands. We found that DBD-C is important for engaging DNA with diverse binding modes, including, unexpectedly, at short ssDNAs. Thus, RPA uses its constituent DBDs to bind biologically diverse ligands in unanticipated ways. These findings lead to a better understanding of how RPA carries out its functions at diverse locations of the genome and suggest a mechanism through which dynamic recognition can impact differential downstream outcomes.



Human replication protein A (RPA) is an essential single-stranded DNA (ssDNA) binding protein complex that is highly conserved across eukaryotes.[1−6] RPA is required for proper execution of major DNA metabolic pathways, including replication, repair, and homologous recombination (HR). RPA most notably acts as a first responder to DNA damage and DNA replication as it coats ssDNA to prevent uncontrolled nuclease activity, stabilizes the ssDNA, and supports the recruitment and assembly of various DNA processing proteins.[1,7−10] Additionally, RPA has been found to function at the single-stranded region of telomeres.[11−16] The role RPA plays in discriminating which DNA processing pathway is executed remains poorly understood. Given RPA's wide range of functions, it is important to understand how RPA binds ssDNA in the context of DNA structure as RPA may use this as a mechanism to discriminate which downstream pathway is activated.

RPA is composed of a large subunit (RPA70), medium subunit (RPA32), and small subunit (RPA14) (Figure 1).[17−22] All three subunits contain oligosaccharide/oligonucleotide (OB) folds, designated as OB A−F, a subset of which are denoted as the core DNA-binding domains (DBDs) (Figure 1).[2,21,23−25] OB-F of RPA70 and the phosphorylation domain (P) and winged helix−turn−helix (WH) domains of RPA32 mediate protein−protein interactions.[1,9,17,19−22,26−29] DBDs C, D, and OB-E interact via a triple α-helical bundle to comprise the trimer core of the three-subunit complex. As revealed in the *U. maydis* RPA crystal structure, RPA binds long ssDNAs through an extensive binding interface composed of four core DBDs spread across RPA70 and RPA32 (Figure 1).[17]

DBDs A and B have been characterized as the high-affinity DBDs, while DBDs C and D are classified as the low-affinity domains.[30−33] In addition to making electrostatic contacts with ssDNA, each DBD contains two highly conserved aromatic residues that mediate base-stacking interactions with ssDNA (Figures 1A−D, S1).[17] Although considered a nonspecific ssDNA binding protein, early studies found RPA to preferentially bind pyrimidine-rich sequences over purine-rich ssDNA. Hence, RPA's ssDNA binding has been extensively

Received: August 30, 2022
Revised: October 7, 2022

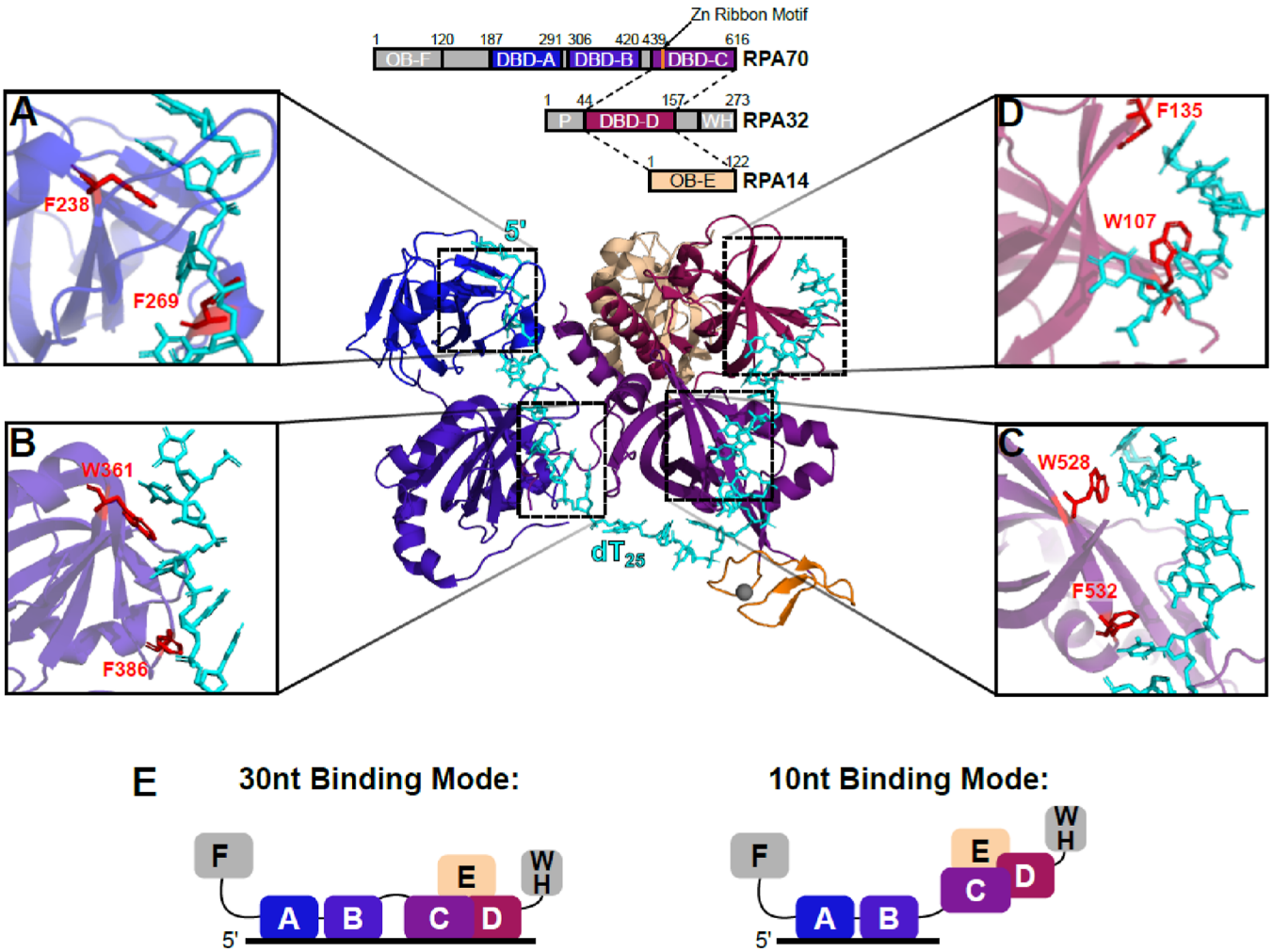


Figure 1. RPA has an extensive ssDNA binding interface that utilizes aromatic base-stacking interactions. Shown is the crystal structure of *U. maydis* RPA bound to ssDNA dT25 (PDB: 4GNX). The structure is colored in correspondence to the above domain map. Domains that make up the RPA trimer core are connected with dashed lines in the domain map, and domains not present in the crystal structure are colored gray. Insets (A–D) correspond to DBD-A, DBD-B, DBD-C, and DBD-D, respectively, and show the conserved aromatic residues that can participate in base-stacking interactions with ssDNA for each DBD. The domain map and amino acid residues in insets (A–D) are numbered in accordance with the human RPA protein. For inset (B), residues 328–340 were removed for clarity. (E) RPA has two canonical binding modes: a high-affinity 30 nt mode in which all four core DBDs are engaged with ssDNA and a lower-affinity 10 nt mode in which only the high-affinity DBDs, DBD-A and -B, are bound to ssDNA.

studied against minimal polythymine (dT_x) or polycytosine (dC_x) ssDNA.[34] Fully bound RPA has an occlusion site of ~30 nucleotides (nts) and was characterized to associate with ssDNA in two distinct, canonical modes: (i) a high-affinity (K_d = 0.2–5 nM) 30 nt mode with all four DBDs bound to DNA and (ii) a lower-affinity (K_d = 5–50 nM) 10 nt binding mode with only DBDs A and B engaged with ssDNA (Figure 1E).[17,21,25,35−37] Originally, it was unclear how the 10 nt mode was employed in the biological context and consequently has been studied less than RPA's 30 nt mode. Additionally, RPA binds ssDNA with a specific polarity where DBD-A is positioned toward the 5′ end of the ssDNA.[38,39] RPA's binding polarity has important functional implications as RPA aids in correct positioning of DNA processing factors, such as XPG in nucleotide excision repair or BLM/DNA2 during HR.[40−45] Although RPA binding has been extensively characterized against dT_x ssDNA, less is known about how it interacts with ligands of varying structures and/or sequences.

RPA's trimer core has been shown to have diverse binding features to ssDNA. Previous studies showed that the trimer core, in the absence of the high-affinity DBDs, was unable to bind ssDNA, but it could bind primer/template junctions with 5′ protruding ssDNA with near full-length protein affinity.[31,46] In the context of full-length RPA, it has been proposed that the trimer core can compete with the high-affinity DBDs OB-A and OB-B for binding on short stretches (<15 nts) of ssDNA.[47] However, to our knowledge, there is no direct evidence supporting trimer core interactions with short stretches of ssDNA in the context of full-length protein.

Recent single-molecule studies have provided complementary insights into dynamics and mechanisms of RPA–ssDNA binding that led to an elaboration on the canonical, static models of binding (see Cladwell and Spies for a comprehensive review[48]). A study examining RPA dynamics at DNA forks shows that RPA only needs to establish a "toehold" of 2–3 nts on ssDNA to facilitate binding of the rest of the RPA molecule.[49] Connected by a short flexible linker, high-affinity DBDs A and B display independent yet coordinated binding to ssDNA.[50] This characteristic is thought to be an important aspect of RPA's modularity acting as a flexible scaffold allows

https://doi.org/10.1021/acs.biochem.2c00504
Biochemistry XXXX, XXX, XXX–XXX

assembly of DNA processing factors. Once bound to ssDNA, DBDs A and B have been shown to be highly dynamic.[36,47,51] The DBDs can undergo rearrangement once bound to ssDNA, and it is postulated that individual DBDs can undergo "microscopic" dissociation and re-binding events.[47,50] Importantly, these DBD dynamics differ in the context of full length compared to truncated constructs containing only DBDs A and B, indicating that individual DBD binding behavior is strongly influenced by the other subunits in RPA.[30,32,47] Together, available studies support the notion that RPA's diverse ssDNA-binding characteristics are influenced by all the DBDs in the complex, along with the length and structural context of its ssDNA ligand. Moreover, these findings highlight the need to study RPA in the context of full-length protein as truncated constructs do not fully recapitulate RPA's ssDNA-binding activity.

In light of recent advancements in understanding how the interplay of individual DBDs modulate RPA's ssDNA-binding properties, we sought to systematically determine RPA's ssDNA-binding characteristics in the context of ssDNA sequence, length, and structure. Long stretches of purely unstructured ssDNA ligands (minimal ssDNA) rarely exist on their own in the cell as transiently exposed ssDNA is rapidly bound and processed by various proteins. Instead, RPA typically encounters ssDNA embedded in various secondary structures, including replication bubbles, ssDNA overhangs (OHs) that occur during double-stranded DNA break processing, and ssDNA gaps arising from nucleotide repair pathways.[52−56] RPA has been shown to behave differently when presented with the same ssDNA sequence in the context of different structures and lengths.[51] It is important to understand how RPA binds ssDNA in varied structural contexts as the ssDNA secondary structure imposes conformational constraints that are not present in linear minimal ssDNA. Moreover, most studies examining RPA ssDNA-binding did so using homopolymers, dT_x or dC_x. Given the role of aromatic residues in base-stacking and likely conferring the pyrimidine preference, it is important to examine binding to more biologically relevant DNA sequences and structures RPA is likely to encounter in a cell.

Prior equilibrium-based binding experiments used to understand each DBD's contributions to binding used a divide-and-conquer strategy using truncated RPA constructs, which have missed the dynamic features evident in single-molecule experiments.[31,57,58] We instead use a systematic mutation of each DBD in RPA to understand the binding contribution of each in the context of full-length protein. This strategy allowed us to probe how these contributions vary in the context of the suite of ligands functionally relevant to RPA. In this study, we used fluorescence anisotropy (FA) to benchmark RPA ssDNA binding to a comprehensive set of exclusively ssDNA ligands that vary in sequence and length. To understand how structure influences binding, we embedded these ssDNA sequences into three different ssDNA-containing structures, ssDNA bubbles, gaps, and OHs, and modulated the length of the ssDNA regions. In this study, we systematically mutated aromatic residues in the DNA-binding interface to block the participation of a given DBD in binding and characterized each mutant's binding against a library of 28 unique ssDNA-containing ligands. This comprehensive evaluation of RPA's ssDNA binding has revealed that RPA accommodates structured ssDNA differently than minimal ssDNA and aromatic residues play differential roles in accommodating structured ligands. Furthermore, we find that DBD-C aromatic residues are important modulators of RPA ssDNA binding and are needed for association with structured ligands with short stretches of ssDNA available to bind. This study highlights RPA's many different binding modes in the context of biologically relevant ligands and may lead to a better understanding of how RPA structural ensembles discriminate between different DNA processing pathways.

MATERIALS AND METHODS

Cloning and Purification of Human RPA. Recombinant RPA was expressed from a pET15b plasmid containing all three subunit genes, RPA70, 32, and 14 (plasmid courtesy of the Chazin lab).[59] RPA70 and RPA14 contain an N-terminal 6×-His tag, while RPA32 was untagged. Mutant RPA plasmids were generated by site-directed mutagenesis using mutagenic primers, and mutations were confirmed by Sanger sequencing. Recombinant RPA was expressed in and purified from *E. coli* BL21(DE3)-pLysS expression cells (Novagen). Briefly, *E. coli* were transformed with the RPA plasmid via heat shock. 4× 1 L cultures were grown in 2×YT media with 0.1 mg/mL ampicillin and 0.34 mg/mL chloramphenicol at 37 °C for 3.5 h to an OD_{600} of 1.0. Cultures were then incubated in an ice bath for 30 min before inducing protein expression with 1 mM isopropyl-β-D-thiogalactopyranoside. After 16−18 h expression at 18 °C, cells were harvested by centrifugation (6000×*g*, 15 min, 4 °C), frozen, and stored at −20 °C until purification.

Recombinant RPA was purified as reported previously.[59] Briefly, 2 L of cell pellets was resuspended in 60 mL of lysis buffer [20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.5, 0.5 M NaCl, 10 μM $ZnCl_2$, 5 mM β-mercaptoethanol, 10 mM imidazole] with one complete ethylenediaminetetraacetic acid (EDTA)-free EASYpack protease inhibitor tablet (Roche) using a tissue Dounce homogenizer. Cells were lysed on ice using a Misonix Sonicator 3000 for 5 min in 5 s pulses with 5 s off between pulses. Lysate was cleared by centrifugation (27,000*g*, 25 min, 4 °C). Before affinity purification via a nickel-NTA column, the lysate was subjected to a 0.1% polyethylenimine precipitation for 1 h at 4 °C to precipitate out RPA bound to DNA. The precipitate was cleared by centrifugation (27,000*g*, 25 min, 4 °C) and incubated with 12 mL of precleared Ni-NTA beads for 1 h at 4 °C, after which unbound proteins were cleared by gravity flow. Beads were washed three times, first with 50 mL of lysis buffer, second with 50 mL of modified lysis buffer containing 30 mM imidazole, and last with 50 mL of modified lysis buffer containing 50 mM imidazole. Protein was eluted with 50 mL of lysis buffer containing 300 mM imidazole. Elution was concentrated at 4 °C using spin concentrators (Vivaspin Turbo) to a final volume of 0.5−1 mL, loaded onto a HiLoad 16/600 Superdex 200 size exclusion column (GE Healthcare), and eluted with fast protein liquid chromatography buffer (20 mM HEPES pH 7.5, 100 mM NaCl, 10 μM $ZnCl_2$, 5 mM β-mercaptoethanol, 0.2 M arginine). Fractions collected containing RPA heterotrimer were pooled, concentrated to 100−200 μM, aliquoted, flash-frozen, and stored at −70 °C until they were used for binding. RPA was quantified using UV−vis spectroscopy, obtaining the A_{280} and using the extinction coefficient, ε_{red} = 87210. Yields for each RPA purification typically ranged between 1.5 and 2 mg/L of growth.

https://doi.org/10.1021/acs.biochem.2c00504
Biochemistry XXXX, XXX, XXX−XXX

Mutant RPAs were generated via site direct mutagenesis using mutagenic primers with 1−2 mismatch nucleotides to mutate the desired amino acid to alanine. Briefly, the plasmid containing WT RPA sequence was amplified via PCR using desired mutagenic primers. All mutagenic primers were purchased from Integrated DNA Technologies (IDT), and sequences can be found in Table S1. After PCR, the sample is treated with DPN1 at 37 °C for 1 h to digest parent plasmid. Following DPN1 treatment, the linear mutant plasmid is ligated together using Gibson assembly (100 mM Tris pH 7.5, 10 mM $MgCl_2$, 0.2 mM of each dNTP, 10 mM dithiothreitol, 0.3 g/mL PEG-8000, 1 mM NAD, 0.005 U/μL T5 exonuclease, 0.03 U/μL Phusion polymerase, 5 U/μL Taq ligase) at 55 °C for 1 h.[60] Ligated plasmids were transformed into DH5α competent cells and selected for using ampicillin. After mutant plasmid amplification, the plasmid was sequenced using Sanger sequencing to confirm mutation. Double mutant plasmids were generated as above, except in two rounds, where round 1 contained one set of mutagenic primers and round 2 contained the other set of mutagenic primers. Mutant RPAs were purified as described above. For mutant RPA quantification, the ε_{red} was adjusted appropriately to compensate for mutation of the aromatic residue to alanine.

DNA Oligonucleotide Preparation. All unlabeled and fluorescently labeled DNA oligonucleotides were purchased from IDT and were subsequently gel-purified before use. Oligonucleotides are fluorescently labeled at the 5′ end with 6-carboxyfluorescein (FAM). ssDNA oligonucleotides used for binding were 10 or 30 nt in length and varied in sequence to reflect representative sequences encountered in vivo (Table S2). Ligands of "random" DNA sequence were generated online using a random sequence generator.[61] Sequences were then computationally screened to assess for propensity to form secondary structures to ensure that the oligonucleotides would be linear under our working conditions. Gap, bubble, and OH structured DNA oligonucleotides were prepared by annealing up to three partially complementary oligonucleotides, with one being fluorescently labeled. Structures had varying 10, 20, or 30 nt ssDNA regions adjacent to partially duplex regions. The sequence of the ssDNA regions is random with a 50:50 mixture of purine and pyrimidine nucleotides (Table S2). The length of the duplex region was modulated to ensure the same total charge regardless of the length of the ssDNA region. For each of the DNA structures, 2 μM of the indicated oligos was mixed in an annealing buffer (30 mM HEPES pH 7.5, 150 mM NaCl, and 1 mM EDTA) and was annealed using a thermocycler. The annealing reaction consisted of initial 95 °C denaturing for 5 min and then was slowly cooled at a rate of 4 °C/min to a final temperature of 25 °C. Structures were verified by native PAGE gel (Figure S2) and stored at 4 °C for future usage.

FA Binding Assay. To perform the binding assays, ssDNA and structured DNA ligand concentrations were kept constant and as low as possible while still maintaining good signal to noise (100−800 pM, see Figure S3A and Table S3). To ensure that the addition of FAM to ssDNA did not affect binding, binding to free fluorophores and competition experiments with unlabeled ssDNA were conducted (Figure S3). Nucleotide ligands were diluted to 2× final concentration in a binding buffer (20 mM HEPES pH 7.5, 200 mM NaCl, 1 mM β-mercaptoethanol). Protein was diluted separately in the binding buffer to 2× final concentration. DNA and protein were mixed in a 1:1 volume ratio in a 20 μL reaction in a flat-bottom low-flange 384-well black NBS polystyrene plate (Corning) and were allowed to reach equilibrium at room temperature in the dark for 30 min. Perpendicular ($I_\perp$) and parallel ($I_\parallel$) fluorescence intensities were measured using a ClarioStar Plus FP plate reader (BMG Labtech), and anisotropy values were calculated for each protein titration point where anisotropy = $(I_\parallel - I_\perp)/(I_\parallel + 2 \times I_\perp)$ and correlates directly with fraction bound for a 1:1 system. Unless otherwise noted, 1:1 binding stoichiometry is assumed. Associated anisotropy was plotted as a function of the log of protein concentration. To determine the apparent dissociation constant, K_D, the data were fit to the quadratic binding isotherm (eq 1) since both ligand and protein concentrations are known

$$\text{anistropy} = S\left(\frac{([L]_T + [P]_T + K_D) - \sqrt{([L]_T + [P]_T + K_D)^2 - 4([L]_T[P]_T)}}{2[L]_T}\right) + O \quad (1)$$

where S is the difference between the maximum and minimum anisotropies measured, O is an offset factor equivalent to the minimum anisotropy measurement, and $[L]_T$ and $[P]_T$ are the known total concentrations of ssDNA and RPA, respectively. The produced anisotropy signal is linearly related to the fraction bound for a 1:1 system (Figure S4). Briefly, RPA was titrated (0.5−1000 nM) and allowed to reach equilibrium with a constant amount of 30 nt ssDNA ligand (25 nM), after which anisotropy measurements were taken. The unsaturated binding region of the curve fits a linear regression well ($R^2 = 0.96$) and a binding stoichiometry of 0.99 ± 0.13 ($n = 3$). For single binding events, anisotropy values can be normalized to fraction bound

$$\text{fraction bound} = \frac{A - O}{S} \quad (2)$$

where A is the measured anisotropy value.

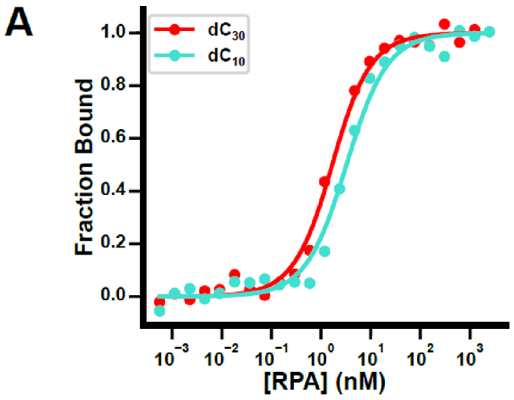
In binding experiments where two distinct binding events are observed, anisotropy values were fit using an independent two-site binding model to extract the apparent dissociation constants for the first and second binding events, $K_{D,1}$ and $K_{D,2}$, respectively (eq 3)

$$\text{anisotropy} = S_1\left(\frac{[P]}{[P] + K_{D,1}}\right) + S_2\left(\frac{[P]}{[P] + K_{D,2}}\right) + O \quad (3)$$

where S_1 is the difference between the absolute minimum anisotropy measurement and the maximum anisotropy of the first binding event (first plateau) and S_2 is the difference between the maximum anisotropy of the first binding event and the absolute maximum anisotropy measured. All data were fit using an in-house binding script. All binding reactions were performed in triplicate or more using different protein dilutions on separate days. Standard errors of the mean (SEMs) were calculated and reported. Statistical significance is determined for differences between averages of apparent dissociation constants using Student's two-tailed paired t-test. The average K_D ± SEM for all proteins binding ssDNA ligands and structured ligands are reported in Tables S3 and S4, respectively.

Competition Experiments. Competition experiments were set up as in FA binding assays. Briefly, a constant amount of WT RPA (6 nM) was incubated with a constant

https://doi.org/10.1021/acs.biochem.2c00504
Biochemistry XXXX, XXX, XXX−XXX



B

WT RPA vs ssDNA: Average K_D (nM) ± S.E.M.

Ligand Length	Ligand dC_X	Random 1	Random 2	Random 3	Random 4	hTel	yTel
30 nts	0.9 ± 0.2	1.2 ± 0.2	2.5 ± 0.2	2.3 ± 0.1	3.4 ± 0.5	1.0 ± 0.1	0.5 ± 0.1
10 nts	3.0 ± 0.3	15.9 ± 0.6	29.8 ± 2.4	17.3 ± 0.6	21.6 ± 2.7	16.2 ± 0.9	4.4 ± 0.4

Figure 2. WT RPA binding ssDNA. (A) Representative FA binding isotherms of WT RPA binding ssDNA dC_{30} (red) and dC_{10} (cyan). Data were fit to the quadratic binding isotherm to extract the dissociation constant, K_D. (B) Tabulated dissociation constants ± SEM of WT RPA binding ssDNA ligands of different sequences and lengths. Random ligands 1−4 correspond to ligands RanS1−4 and RanL1−4 in Table S2 for 10 and 30 nt ssDNA, respectively.

amount of labeled ssDNA (30 nt random 2; 0.8 nM). A serial dilution over of unlabeled competitor 30 nt ssDNA was performed over 6 $\log_{10}$ scales and then added to the labeled, bound complex. Reactions were allowed to reach equilibrium at room temperature overnight before data collection on a plate reader. Anisotropies were calculated for each titration point, and data were fit with a single-site competitor model to extract inhibitor constant, K_i

$$\text{anisotropy} = S - O \times \left(\frac{[\text{uL}]}{[\text{uL}] + K_i} \right) \quad (4)$$

where uL is the unlabeled competitor ligand.

Electrophoretic Mobility Shift Assays. Electrophoretic mobility shift assays (EMSAs) were performed using the same fluorescently labeled ligands as in FA assays. Binding reactions were prepared in a modified binding buffer (20 mM Tris pH 7.5, 200 mM NaCl, 1 mM β-mercaptoethanol, 5% glycerol), and the final ligand concentration was 2.5 nM. Reactions were allowed to reach equilibrium at 4 °C in the dark for 30 min. Binding reactions were separated on a native 1× TBE 3.5−10% or 6.5% polyacrylamide gel at 200 V for 40 min at 4 °C. Gels were imaged using a Typhoon FLA 9500 imager (GE Healthcare) using 473 nM excitation.

Mass Photometry. Mass photometry was used as an orthogonal method to examine stoichiometry between RPA and 30 nt 3′ OH ssDNA. Data were collected on the Refeyen Two^{MP} (Refeyen) mass photometer. Protein and ligand dilutions were made in the binding buffer, as used in FA experiments, and were allowed to reach equilibrium at room temperature. In order to obtain a sufficient signal for bound species, ligand concentrations were kept constant at 25 nM and the RPA concentration varied from 25 to 500 nM. Complex sizes are determined by fitting to a standard curve constructed from protein standards.

RESULTS

RPA is More Sensitive to DNA Sequence when Binding Short, Minimal ssDNAs. Recent single-molecule studies suggest that there are at least two RPA−ssDNA binding modes with differing kinetic stabilities, with only the most kinetically stable complex being detectable by EMSA.[51] Thus, we used an FA-based assay for our binding studies since it is a true equilibrium method ideal for detecting any kinetically labile states with the capacity for high-throughput measurements. Using FA, we generated highly reproducible binding isotherms (Figure 2A) allowing extraction of the apparent dissociation constants, K_D, for RPA−ssDNA binding. In almost all cases, the data were well fit with one binding event, suggesting that the multiple binding modes detected by smTIRF are not distinguishable events in solution under our conditions.[51] Our measured apparent dissociation constants for dC_{30} and dC_{10} were 0.9 ± 0.2 nM and 3.0 ± 0.3 nM, respectively (Figure 2B), consistent with previously reported apparent K_D's for RPA binding 30−35 and 8−15 nt poly-dT or poly-dC homopolymers which range from 0.2 to 5.0 and 1 to 50 nM, respectively.[23,30,33,34,50,51,62,63] The close agreement of our values with published data using other methods suggests that FA is a robust method to determine the K_D's of RPA−ssDNA complexes.

Binding was performed with 10 and 30 nt minimal ssDNA ligands that represent a variety of sequences: canonical poly-dC (dC_x), four different biologically representative "random" DNA sequences representing the genome complexity, and human (hTel) and yeast (yTel) telomeric sequences (Table S2). We found that RPA binds minimal 30 nt ssDNA ligands with similar affinities ranging from K_D = 0.5 to 3.4 nM (Figure 2B) with a slight preference (2- to 3-fold) toward a pyrimidine-rich sequence compared to a random DNA sequence (Figure 2B). The pyrimidine preference is amplified to 5- to 10-fold in the context of 10 nt ssDNA ligands (Figure 2B) as noted previously.[33] Moreover, binding to 10 nt random sequence ligands was more variable than their 30 nt counterparts, suggesting that RPA is generally more sensitive to DNA sequence when only short stretches of DNA are available to bind. Unexpectedly, RPA bound the yeast telomeric sequence, which contains a mixture of purine and pyrimidine bases, with similar affinity to dC_{10} (K_D = 4.4 nM vs K_D = 3.3 nM,

https://doi.org/10.1021/acs.biochem.2c00504
Biochemistry XXXX, XXX, XXX−XXX

respectively). Examination of the crystal structure of RPA bound to dC_8 ssDNA (PDB: 1JMC) and the 10 nt yTel sequence reveals that three of the four aromatic residues in DBDs A and B could similarly participate in base stacking interactions with three thymine nucleotides (yTel sequence: TGTGGGTGTG or TGTGGGTGTG)[21] which may help RPA form high-affinity interactions with short spans of ssDNA. We use dissociation constants to random ssDNA ligands as RPA's baseline affinities to minimal 30 (K_D = 1−2 nM) and 10 (K_D = 15−30 nM) nt ssDNA for further comparisons.

RPA Binding of Minimal Random ssDNA is Altered in the Context of Biologically Relevant Secondary Structures. Functionally, RPA encounters a wide variety of ssDNA-containing DNA secondary structures, including ssDNA bubbles, ssDNA gaps, and ssDNA OHs, rather than isolated ssDNA.[64] Previous studies have examined RPA binding to dT_x ssDNA in the context of structure; however, the length and sequence of the ssDNA region were rarely modulated.[44,46,51] To identify the impact of these structures on RPA binding activity, we embedded minimal, random ssDNA regions that reflect the types of sequence encountered by RPA in vivo into the tested secondary structures and vary the length of the ssDNA region from 10 to 30 nts (Figure 3).

When presented with ssDNA gaps of 10, 20, or 30 nts (Gap10, Gap20, and Gap30, respectively), we found that RPA bound all three structures indiscriminately with a $K_D \approx 10$ nM. This was a surprising observation given that RPA binds minimal 30 nt random ssDNA significantly tighter than their 10 nt counterparts. This would suggest that WT RPA accommodates all the gap structures similarly in an ssDNA length-independent manner for the size of DNA gaps we studied. The observed K_D's for gapped DNA most closely reflect the affinity of RPA toward random, minimal 10 nt ssDNA substrates, raising the possibility that RPA engages with gap structures in a conformation that reflects RPA's canonical 10 nt binding mode (Figure 1E). Structural constraints of the ss/dsDNA junctions may limit RPA's ability to adopt a fully engaged mode of binding. Our observations with a randomized sequence differ from a previous study which examined RPA binding to a gap structure with a dT_{20} ssDNA region; however, the differences in observed K_D can be explained by the sequence of the ssDNA region as our data show that RPA favors 10 nt polypyrimidine sequence 5- to 10-fold over random DNA.[51]

Unlike ssDNA gaps, RPA displayed a differentiated binding profile when binding ssDNA bubbles of various lengths. Binding affinities were measured with ssDNA bubbles of 10, 20, and 30 nts in length (Bub10, Bub20, and Bub30, respectively). As expected from RPA's minimal ssDNA-binding characteristics, we observed an attenuation in binding affinity as the size of the ssDNA bubble was decreased, with measured dissociation constants of 5.2 ± 0.3, 12.7 ± 1.3, and 63.2 ± 5.2 nM for Bub30, Bub20, and Bub10, respectively. While RPA binds 30 and 20 nt bubbles with similar affinity to 30 and 10 nt minimal ssDNA, respectively, the drop in affinity becomes more drastic as the size of the ssDNA bubble shrinks to 10 nt. Topological constraints on the bubble may only present a smaller track of ssDNA available for binding, resulting in lower affinities for 20 and 10 nt bubbles. Previously, Chen et al. examined RPA binding to a dT_{20} ssDNA bubble. Using smTIRF, they were unable to detect binding between the immobilized RPA and their bubble ligand; however, they were able to detect binding via EMSA.[51] It is possible that immobilization of RPA affects the flexibility of its DBDs and this flexibility is needed to form a bound complex with bubble ssDNA.

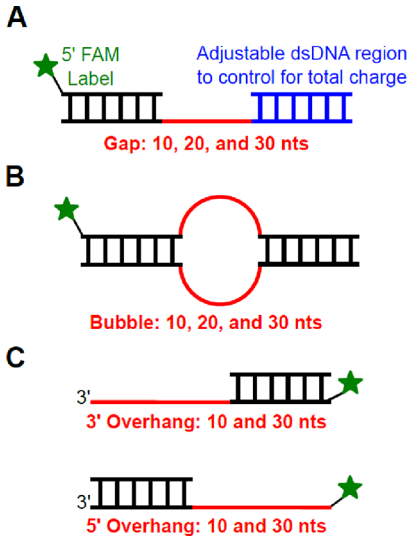


Figure 3. (A) Diagram of the ssDNA gap structure. The ssDNA region (red) varied in lengths of 10, 20, and 30 nts. The adjustable dsDNA region (blue) was used to control for total charge of the different gap structures, so observed changes in binding affinity are not an artifact rising from the total charge of the structure. All ssDNA regions are of random sequence composition. The FAM label is on the 5′ end of a 15 nt ssDNA oligo termed gap label, which is annealed to all the gap structures. (B) Diagram of the ssDNA bubble (Bub) structure. The ssDNA region (red) varied in lengths of 10, 20, and 30 nts. All ssDNA regions are of random sequence composition and are not complementary to the antiparallel strand to ensure bubble formation. All bubble structures have the same total charge. The FAM label is on the 5′ end of a 60 nt ssDNA oligo termed Bub label, which is annealed to all the bubble structures. (C) Diagram of the ssDNA OH structures. OH structures were constructed to have ssDNA OHs of 10 and 30 nts in length. Since RPA binds ssDNA with polarity, 5′ and 3′ ssDNA OHs for each length were generated. All OH structures have the same total charge. For 10 nt OH structures, the FAM label is on the 5′ end of a 40 nt ssDNA oligo termed 10 nt label, which is annealed to all 10 nt OH structures. For 30 nt OH structures, the FAM label is on the 5′ end of a 50 nt ssDNA oligo termed 30 nt label, which is annealed to all 30 nt OH structures.

Next, OH structures, duplex DNAs with either 5′ and 3′ ssDNA ends, were constructed to assess the role of polarity in this context. Examining binding to 10 nt OHs, RPA bound 5′ and 3′ OHs similarly, with K_D = 15.8 ± 2.2 nM and K_D = 23.2 ± 3.9 nM, respectively. The dissociation constants to both structures are similar to those measured for the minimal 10 nt ssDNA (K_D = 15−30 nM), suggesting that RPA engages with 10 nt OHs in a manner similar to 10 nt ssDNAs. RPA has a slight preference for binding the 10 nt 5′ protruding ssDNA over the 3′ ligand, although the difference between affinities is less than 2-fold. There are discrepancies in the literature on RPA preference for binding 5′ or 3′ protruding ssDNA, potentially arising from the ligand preparation method selected.[38,44] Due to RPA's binding polarity, steric constraints of the ss/dsDNA junction may have a larger effect on 3′ protruding ssDNA, which could decrease RPA's affinity for the ligand. While our data with WT RPA show that it can bind both 10 nt OH polarities with similar affinities and a slight preference for the 5′ ssDNA, our subsequent mutational analysis indicates that RPA is engaging with each ligand

https://doi.org/10.1021/acs.biochem.2c00504
Biochemistry XXXX, XXX, XXX−XXX

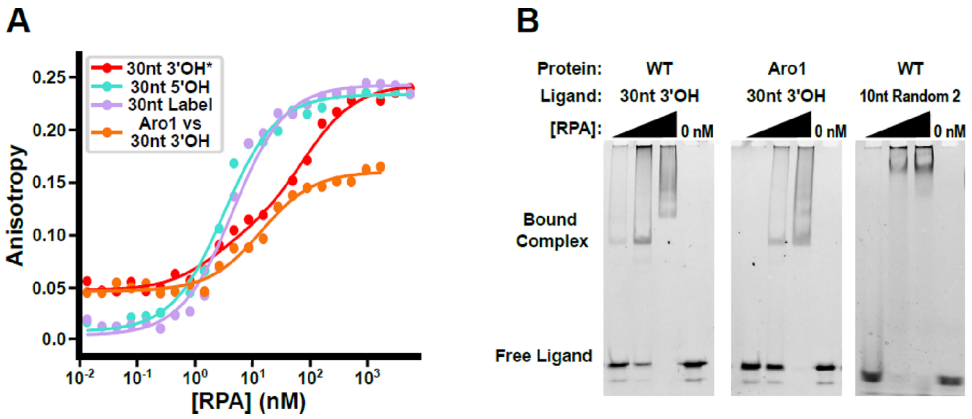


Figure 4. Two binding events are observed for binding 30 nt 3′ OHs. (A) Representative binding isotherms of WT RPA binding the 30 nt OH structures and 30 nt label, and Aro1 binding 30 nt 3′ OH. WT RPA shows two binding events when binding the 30 nt 3′ OH substrate (red), while select aromatic mutants only show one binding event (example: Aro1, orange). All bindings that showed one binding event were fit to the quadratic binding isotherm to extract the dissociation constant, K_D, while all two-binding event isotherms were fit with an independent two-site binding model (see methods). (B) EMSA of WT and Aro1 RPA binding 30 nt 3′ OH structure (panels I and II, respectively) and WT RPA binding 10 nt ssDNA random 3 (panel III). Panels I and II were separated on a native 3.5−10% polyacrylamide gel, and panel III was separated on a native 7% polyacrylamide gel. The final concentrations of RPA from low to high are 4.2, 25, and 150 nM. The ligand concentration was held constant at a final concentration of 2.5 nM.

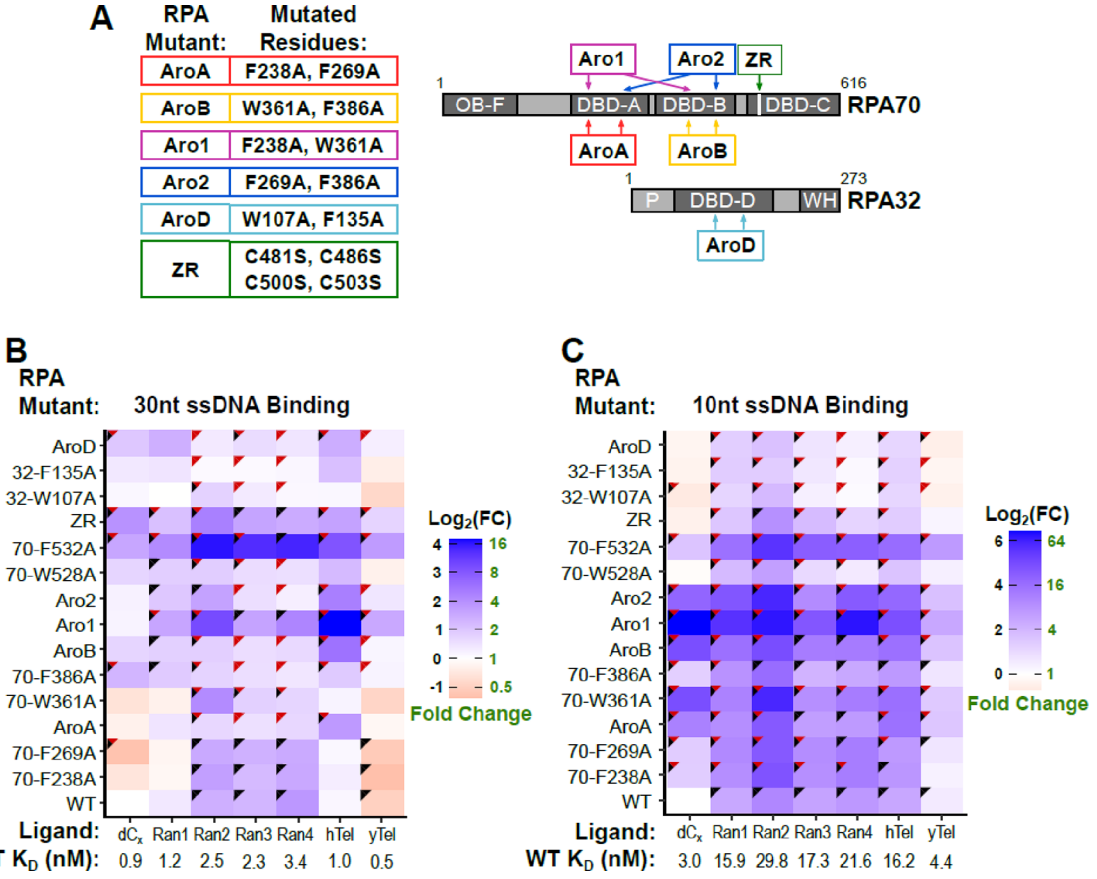


Figure 5. Heatmap of RPA aromatic mutants binding with ssDNA. (A) Single and double aromatic mutants of RPA were generated for use in binding experiments. Shown is a diagram of double and quadruple RPA mutants with their mutant name and mutated residues (left). Mutations are mapped to the corresponding DBDs on RPA domain map (right). ZR = zinc ribbon; the ZR mutant has all four zinc-coordinating cystines mutated to serines. The ssDNA binding profile for all aromatic mutants against (B) 30 nt ssDNA ligands and (C) 10 nt ssDNA ligands. Dissociation constants are represented as a heatmap of the $\log_2$ of the fold change (FC) relative to WT RPA affinity for dC_{30} and dC_{10} for 30 and 10 nt ligands, respectively. Colored triangle in the top right corner represents significance: black = significantly different from WT affinity for dC_x; red = significantly different from WT affinity for ligands of the same type; black and red = significantly different from WT affinity for both dC_x and ligands of the same type. Random ligands 1−4 are abbreviated Ran1−4. The average WT RPA K_D is reported below each ligand. Tabulated FC values for 30 and 10 nt ssDNA can be found in Tables S6 and S7, respectively.

https://doi.org/10.1021/acs.biochem.2c00504
Biochemistry XXXX, XXX, XXX−XXX

differently and can more easily accommodate short 5′ ssDNA OHs (see below).

As the protruding ssDNA increases in size to 30 nt, we observe different binding behaviors for 5′ and 3′ OHs. Specifically, one binding event is observed for the 5′ OH, but two distinct binding events are observed for the 3′ OH (Figure 4A). This was unexpected as we have not observed multiple distinct binding events for any other of the ligands used in this analysis. RPA binds 30 nt 5′ OH with a tight affinity (K_D = 2.7 ± 0.2 nM), similar to that of random 30 nt ssDNAs. The 30 nt 3′ OH binding isotherm was fit with an independent two-site binding model to extract the dissociation constants for the first higher-affinity and second lower-affinity binding events, $K_{D,1}$ and $K_{D,2}$, respectively. The first binding event observed when associating with 30 nt 3′ OH is also of high affinity, $K_{D,1}$ = 2.2 ± 0.3 nM. Thus, unlike the 10 nt OH ligands, there is no obvious preference for either polarity for the high-affinity binding event, which is in agreement with previous studies.[44] The second binding event is significantly weaker ($K_{D,2}$ = 50.3 ± 10.3 nM) than the first and could represent the binding of another RPA monomer on the ssDNA. To address the stoichiometry of binding, we performed EMSAs to better define the nature of the second binding event (Figure 4B panel I). At low concentrations of RPA, a single faster migrating bound species is observed, consistent with the first detected binding event (Figure 4B panel I, lanes 1 and 2). At higher concentrations of RPA, we see the appearance of a second, slower migrating species, indicating the binding of another RPA monomer on the DNA (Figure 4B panel I, lane 3). An orthogonal strategy using mass photometry was done to confirm the oligomeric state of the secondary binding event and shows 2:1 ratio of RPA to ssDNA at higher concentrations of RPA (Figure S5). The oligomerization of RPA on ssDNA is prevented when only short (≤10 nts) stretches of ssDNA are available to bind (Figure 4B panel III). Overall, these data suggest that RPA binds minimal and structured ssDNA in distinct modes.

Aromatic Residues in DBD-A and DBD-B are Critical for High-Affinity Interactions with Embedded Minimal ssDNA. RPA may achieve high-affinity recognition of diverse ssDNAs through a plastic binding interface composed of its multiple OB-folds. In this model, individual OB-folds are needed to promote specific modes of recognition or play a compensatory role in ssDNA binding. As conserved aromatic residues are critical for ssDNA recognition, disruption of these interactions is a tool to define the role of individual OB-folds in recognition.[65] Functional characterization of mutants lacking these aromatic residues in the high-affinity DBD-A and -B has revealed separation of function RPA mutants, highlighting the importance of aromatic residues in RPA functions.[58,66,67] Previously developed double aromatic RPA mutants in DBD-A and -B, AroA (F238A/F269A), AroB (W361A/F386A), Aro1 (F238A/W361A), and Aro2 (F269A/F386A), are ideally suited for testing disruptions to ssDNA binding (see Figure 5A for RPA mutant list). We used these mutants, along with single aromatic mutations, in the context of full-length RPA to address which of RPA's DBDs are essential for its association with our library of ssDNA ligands to reveal how disruption of individual OB-fold ssDNA interactions affect overall recognition of various ssDNAs.

When binding minimal 30 nt ssDNA ligands, both single and double aromatic RPA mutants in DBDs A and B displayed similar binding characteristics to WT (Figure 5B), consistent with prior studies on these mutants using >dT_{25}.[51,58,66,67] One exception is the interaction between the double aromatic mutants and the 30 nt human telomere oligonucleotide. RPA mutants AroA, AroB, and Aro2 bound hTel 3- to 5-fold weaker than WT ($K_{D,mutants}$ = 3.3−5.4 nM), while Aro1 RPA binds even more weakly (K_D = 16.6 ± 3.5 nM). This specific behavior at telomere sequences could have important implications for RPA interactions with human telomeres in vivo, specifically RPA's ability to melt guanine-quadruplexes (G-quads) and protect the G-rich 3′ ssDNA OH at telomere ends.[12−14] As reported for binding polypyrimidine sequences, our data suggest that mutation of the aromatic residues generally has a negligible effect on binding random minimal 30 nt ssDNA, although there are certain sequence contexts where this appears to be not the case.[32]

The aromatic residues in DBD-A and -B also play a prominent role in driving high-affinity complexation with ssDNA when binding short stretches of ssDNA (≤10 nts) (Figure 5C). We found that, in general, double aromatic mutants bound short, random ssDNA ligands 2- to 10-fold weaker than WT RPA. Interestingly, the most exaggerated differences in binding were observed when binding RPA's preferred dC_{10} substrate, with fold-changes (FCs)in binding ranging from 11- to 85-fold weaker than WT. This suggests that the polypyrimidine ligands previously used for binding may not accurately capture the range of RPA's true ssDNA-binding activity. This likely explains why past studies were unable to detect binding between aromatic mutants and dT_{15} ligands.[51,67] Mixed double aromatic mutants Aro1 and Aro2 bound minimal, random 10 nt ssDNAs weaker than either AroA or AroB, suggesting coupling between mutations at these sites such that mutating one aromatic in both DBDs is more detrimental than mutating both aromatics in a single DBD. Aro1 RPA bound 10 nt ligands the weakest among the mutants with dissociation constants typically in the 100−200 nM range (Figure 5C). This suggests that the region of the OB-fold that makes the most 5′ interaction with ssDNA in each DBD is important for forming high-affinity complexes with short, minimal ssDNA.

The impact of aromatic mutations in DBD-A and -B is amplified when binding ssDNA embedded in DNA secondary structures. All single and double aromatic mutants in DBD-A and -B, apart from F386A, bound all the ssDNA gap structures significantly weaker than WT RPA, typically 5- to 15-fold weaker (Figure 6). These aromatic mutations also had a dramatic effect on binding to small ssDNA bubbles. Except for Aro1 and F386A RPA, all other mutants had a modest impact on binding Bub30 and Bub20, binding approximately 2- to 3-fold and 3- to 8-fold weaker relative WT RPA, respectively (Figure 7). Surprisingly no DBD-A or DBD-B mutants (except for F386A) formed detectable bound complexes with Bub10 (Figure 7). Aro1 RPA bound bubble structures particularly weak, binding Bub30 10-fold weaker than WT, and was unable to form detectable bound complexes with Bub20 or Bub10.

Aromatic mutations also tended to have larger effects on binding 10 nt OHs relative to minimal 10 nt binding. Mutants typically bound short OHs 4- to 34-fold weaker relative to WT and tended to prefer binding the 5′ OHs over its 3′ counterpart, except for AroA and Aro1 RPA (Figure 8A). As the OH length increases to 30 nts, all aromatic mutants, except for Aro1 RPA, have virtually no effect on binding 30 nt 5′ OH or the first, high-affinity binding event with the 3′ structure (Figure 8B), suggesting that these mutants are engaging with

https://doi.org/10.1021/acs.biochem.2c00504
Biochemistry XXXX, XXX, XXX−XXX

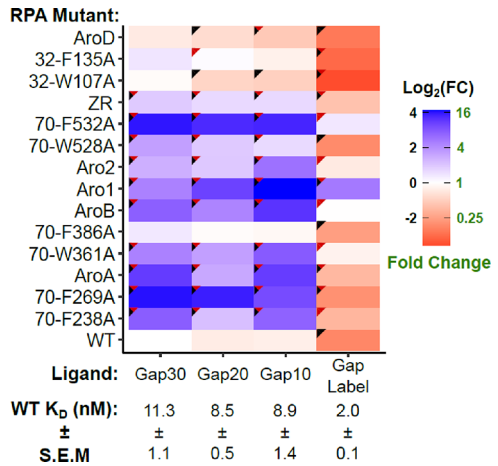


Figure 6. Heatmap of WT and aromatic mutants binding ssDNA gap structures. Dissociation constants are represented as a heatmap of the $\log_2$ of the FC relative to WT RPA affinity for Gap30 ligand. The colored triangle in the top right corner represents significance: black = significantly different from WT affinity for Gap30; red = significantly different from WT affinity for ligands of the same type; black and red = significantly different from WT affinity for both Gap30 and ligands of the same type. The length of the ssDNA region is denoted by the number following Gap. Gap30, Gap20, and Gap10 have ssDNA regions of lengths 30, 20, and 10 nts, respectively. The average WT RPA K_D ± SEM is reported below each ligand. Tabulated FC values for ssDNA gaps can be found in Table S8.

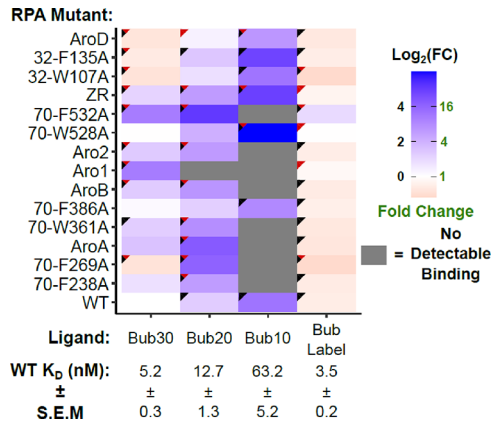


Figure 7. Heatmap of WT and aromatic mutants binding ssDNA bubble structures. Dissociation constants are represented as a heatmap of the $\log_2$ of the FC relative to WT RPA affinity for Bub30 ligand. The colored triangle in the top right corner represents significance: black = significantly different from WT affinity for Bub30; red = significantly different from WT affinity for ligands of the same type; black and red = significantly different from WT affinity for both Bub30 and ligands of the same type. The length of the ssDNA region is denoted by the number following Bub. Bub30, Bub20, and Bub10 have ssDNA regions of lengths 30, 20, and 10 nts, respectively. The average WT RPA K_D ± SEM is reported below each ligand. Tabulated FC values for ssDNA bubbles can be found in Table S9.

these ligands in a manner similar to 30 nt minimal ssDNAs. The second binding event observed with the 30 nt 3′ OH is, however, more greatly affected by aromatic mutations. The large effect on $K_{D,2}$ suggests that the second binding event is consistent with RPA engaging with available ssDNA in a 10 nt-like binding mode (Figure 1E). Mutants F269A, AroA, and Aro1 RPA had no detectable second binding event (Figure 4A,B panel II). An interesting observation is that F269A RPA does not have a detectable second binding event, while one of its double mutants, Aro2 RPA, does. This suggests that F269 is important for facilitating the second binding event, but there is some unknown synergistic effect of mutating both F269 and F386 that allows for RPA regain this binding property. Overall, mutational analysis of the aromatic residues in the high-affinity DBD-A and -B has revealed that these aromatic residues are important for forming high-affinity interactions with 10 nt ssDNAs, especially in the context of secondary structures, but are dispensable for binding long ssDNAs, consistent with previous observations.[37,51,58]

DBD-C is Crucial for Recognition of Short ssDNA Oligonucleotides and Structured ssDNAs. Recent single-molecule studies suggest that the trimer core (DBDs C, D, and OB-E, Figure 1) is important for modulating individual DBD behavior when bound to ssDNA.[47] DBD-C uniquely contains a zinc ribbon (ZR) domain, which has been implicated in DNA replication, protein stability, and, to a lesser extent, DNA binding.[58,68] Bulk biochemical studies have focused on the isolated trimer core, which assumes that the functions of DBDs A and B are separable from the trimer-forming domains.[37,46,58] This assumption may not hold true in all contexts as our data and other available data suggest a complex interplay of binding modes at structured ligands.[47]

To better understand the role the trimer core plays, we mutated the conserved aromatic ssDNA-contacting residues in DBD-C as well as the ZR (Figure 5A) and tested binding to our library of ssDNA ligands. We found that mutation of the conserved aromatic residue W528 in DBD-C had little effect on minimal 30 or 10 nt ssDNA binding (Figure 5B,C). In contrast, mutation of F532 appears to make RPA more sensitive to ssDNA sequence when binding in the 30 nt mode, as evident by the weaker (15-fold) binding to random DNA sequences. This is consistent with the engagement of this residue in the context of longer ssDNAs, as seen in the crystal structure (Figure 1).[17] Unexpectedly, mutation of F532 has a greater effect when binding to 10 nt ssDNAs, with affinities 3- to 5-fold weaker than WT RPA (Figure 5C). This is surprising given that the canonical model of the 10 nt mode has only DBD-A and -B contacting ssDNA (see Figure 1E); thus, no impact was expected upon mutation of DBD-C. Our data suggest that DBD-C, specifically F532, is important for binding short stretches of ssDNA in the context of full-length protein.

The aromatic residues in DBD-C, especially F532 and to a lesser extent W528, gain greater importance when considering association with ssDNA structures. Mutating W528 to alanine had a modest effect on binding ssDNA gaps, binding all gaps 2- to 3-fold weaker than WT. When presented with ssDNA bubbles, W528A RPA bound Bub30 and Bub20 with similar affinity to WT but had significantly reduced affinity to Bub10. As observed with minimal ssDNAs, F532 participates to a greater extent in binding these structured ligands, with mutant F532A RPA binding to all gap structures approximately 17-fold weaker relative to WT (Figure 6). Additionally, F532A RPA bound all bubble structures significantly weaker than WT and surprisingly did not bind Bub10 to any detectable degree up to an RPA concentration of 5 μM (Figure 7). Another unexpected observation was that F532A RPA was unable to bind 10 nt 3′ OH structure but could bind the 10 nt 5′ ligand (Figure 8A). This was the only RPA mutant studied here that was unable to bind the 10 nt 3′ OH ligand. These mutations in DBD-C had marginal impacts on the high-affinity binding events for 30 nt OH ligands but had a much greater effect on

https://doi.org/10.1021/acs.biochem.2c00504
Biochemistry XXXX, XXX, XXX−XXX

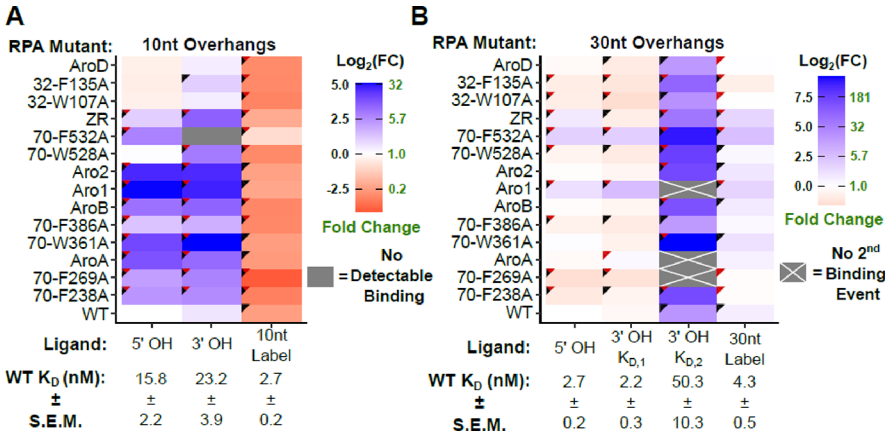


Figure 8. Heatmap of WT and aromatic mutants binding (A) 10 nt ssDNA OH and (B) 30 nt ssDNA OH structures. Dissociation constants are represented as a heatmap of the $\log_2$ of the FC relative to WT RPA affinity for 10 nt 5′ OH and 30 nt 5′ OH ligand for (A,B), respectively. For (A), the colored triangle in the top right corner represents significance: black = significantly different from WT affinity for 10 nt 5′ OH; red = significantly different from WT affinity for ligands of the same type; black and red = significantly different from WT affinity for both 10 nt 5′ OH and ligands of the same type. For (B), the colored triangle in the top right corner represents significance: black = significantly different from WT affinity for 30 nt 5′ OH; red = significantly different from WT affinity for ligands of the same type; black and red = significantly different from WT affinity for both 30 nt 5′ OH and ligands of the same type. The polarity of the ssDNA region is denoted by either a 5′ or 3′ for 5′ ssDNA OHs and 3′ ssDNA OHs, respectively. The average WT RPA K_D ± SEM is reported below each ligand. Tabulated FC values for 10 and 30 nt ssDNA OHs can be found in Tables S10 and S11, respectively.

the weaker, second binding event (Figure 8B). Overall, these data suggest that DBD-C is important and necessary for forming high-affinity complexes when only short stretches of ssDNA are available to bind. It may contribute to binding by providing redundant yet necessary binding modes, so RPA can stay bound to or effectively rebind ssDNA.[51] Our data suggest that residue F532 in DBD-C is important for facilitating the switch to this redundant binding mode. Together, these data challenge the static view of the canonical 10 nt binding mode as mutations in DBD-C impact binding to short stretches of ssDNA and suggest that the trimer core participates in a 10 nt binding mode.

The ZR has been implicated in protein−protein interactions and overall stability of the RPA complex but has been shown to have minimal impact on binding dT_{30} ssDNA.[58] However, the ZR partially wraps around the ssDNA (Figure 1) and makes electrostatic contacts with the ssDNA backbone in some structures of RPA.[17,36] Mutation of the ZR (C481S/C486S/C500S/C503S) has little effect on binding either 30 or 10 nt ssDNA, consistent with previous studies (Figure 5).[58] We sought to investigate if the ZR has a role in accommodation of structured ssDNAs. The lack of an effect on minimal ssDNA binding extended to our panel of structured ligands. ZR-RPA bound all ssDNA structures approximately 2-fold weaker than WT, with the exception of 10 nt 3′ OH (∼8-fold weaker relative to WT). Interestingly, this suggests that the ZR may be needed for binding short ssDNA ligands less than 10 nts in length but otherwise does not play a significant role in binding longer minimal or structured ssDNAs.

Aromatic Residues in DBD-D Modulate RPA's Sequence Sensitivity. RPA32 contains the last of the four ssDNA-binding DBDs, DBD-D, which is flanked by an N-terminal unstructured phosphorylation domain and a C-terminal WH domain, both of which have be implicated in modulating protein−protein interactions (Figure 1).[10,31,42,69] Like DBD-C, DBD-D's overall contributions to RPA−ssDNA binding in the context of full-length protein are not well understood, even though DBD-D shows extensive contacts to ssDNA in the crystal structure.[17] Prior studies examining DBD-D's ssDNA properties were typically done in the context of either the isolated trimer core or RPA32/14 dimer.[31,46,58] To our knowledge, only one study has examined the effects of mutating the conserved aromatic residues in DBD-D in context of full-length protein using the canonical ligands.[37] Using *S. cerevisiae* RPA, similar binding affinities between WT and OB-D mutant (W107A/F135A) against dT_{12} and dT_{40} ligands were observed, suggesting that these residues do not contribute to binding.[31,37,58] However, recent dynamic and structural studies suggest that it may play a larger role in binding than originally appreciated.[36,47] Given that its participation in the dynamic modes of binding might not be accurately recapitulated by isolated domains, we examined mutant DBD-D ssDNA binding in the context of the full-length protein.[47]

Mutation of individual or both conserved surface aromatic residues in DBD-D, W107, and F135 reduces RPA's sequence sensitivity, especially in the 10 nt mode (Figure 5C). DBD-D mutants bound all 10 nt minimal ligands with affinities similar to WT binding dC_{10} (Figure 5C), resulting in the loss of sequence discrimination. This result was highly unanticipated in light of the previous characterization of DBD-D contributions to binding.[31,37,58] Additionally, our DBD-D mutant binding data to minimal ssDNAs lend further support to the idea that the trimer core participates in binding short (≤10 nts) ssDNAs. Next, we determined if DBD-D is needed for recognition of the structured DNA ligands. In contrast to observations with 10 nt ligands, nearly all RPA32 mutants bound structured ssDNAs with near-WT affinity (Figures 5678). The one exception is AroD RPA (W107A/F135A) association with bubble ligands, which bound all bubble sizes 2-fold tighter than their respective WT affinity (Figure 7). Mutation of the aromatic residues to alanine may allow for conformational flexibility of the ssDNA as the base-stacking interactions are eliminated. This plasticity allows for the DNA to adopt a more thermodynamically favored bound state to be formed with DBD-D and/or RPA's other DBDs. Together

https://doi.org/10.1021/acs.biochem.2c00504
Biochemistry XXXX, XXX, XXX−XXX

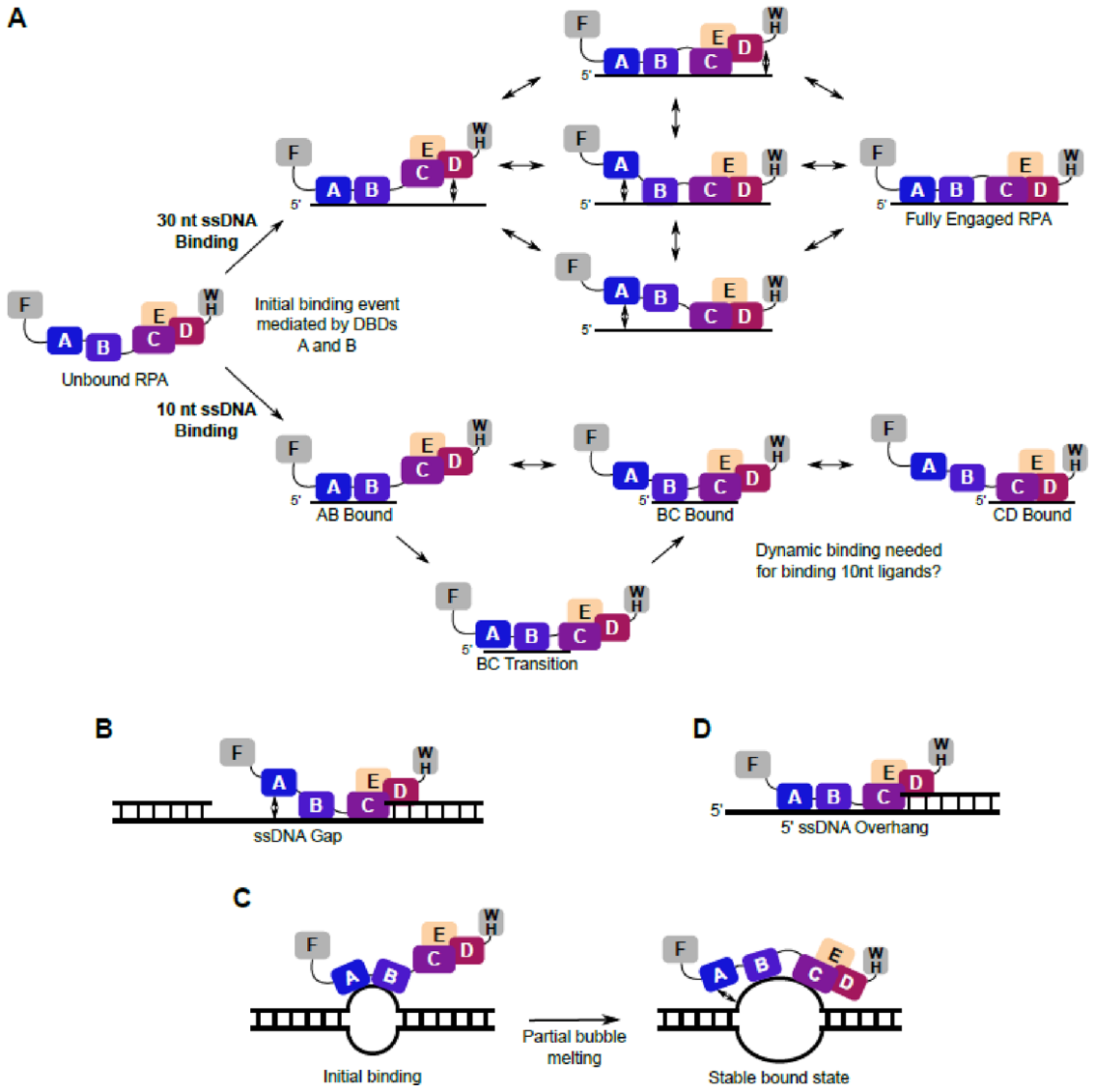


Figure 9. RPA engages with ssDNA in diverse binding modes. (A) Schematic of dynamic RPA ssDNA binding to 30 nt (top) and 10 nt (bottom) minimal ssDNA ligands. The dynamic binding to 30 nt ligands is based on a model from (48). Potential binding models for RPA association with ssDNA gaps (B), bubbles (C), and 5′ OHs (D).

with the data from DBD-C, our analysis supports the model of differential DBD engagement where the trimer core is a key participant in binding ssDNA, both long and short.

DISCUSSION

RPA Trimer Core Aids in Binding of Short Stretches of ssDNA. Prior analysis has suggested a sequential model of RPA binding to long ssDNAs where the high-affinity domains bind ssDNA first, followed by trimer core DBD binding.[17,23] In contrast, the observation of additional bound states in single-molecule experiments suggests a more dynamic binding mechanism, with individual DBDs dissociating and associating with long ssDNA (see Figure 9A top).[35,47] Using in-solution binding experiments, our mutational analysis of aromatic residues in trimer core DBDs C and D suggests that the trimer core does in fact participate in binding short stretches of ssDNA as mutations in this region unexpectedly led to significant differences in binding short minimal and structured ssDNAs.

The length constraint inherent in short ssDNA (≤10 nts) would mean that the bound DBDs need to hand off the ssDNA to the unbound DBDs to stay macroscopically associated with ssDNA. DBD-C positions itself as the intermediate to facilitate the handing off the ssDNA from DBDs A/B to the trimer core (Figure 9A BC bound). Our data pinpoint F532 as essential for this handoff mechanism by feeding ssDNA into the binding interface of DBD-C (Figure 9A BC transition), while W528 assists in guiding the ssDNA to DBD-D (Figure 9A CD bound). As with long ssDNAs,[47,48] these different bound states are most likely in dynamic equilibrium with one another. Given this mechanism of binding to short ssDNAs, we looked to see if we can explain F532A RPA's discrepancy in binding the 10 nt OH structures, binding the 5′ but not the 3′ OH. Given RPA's binding polarity and steric hindrance caused by the ss/dsDNA junction, the effective binding site for the 3′ protruding ssDNA is likely smaller than 10 nts. With sufficiently short ssDNA, F532A RPA's DBD-C cannot not get a toehold on the DNA, preventing the handoff of ssDNA. Without this handoff, RPA cannot form a detectable bound

https://doi.org/10.1021/acs.biochem.2c00504
Biochemistry XXXX, XXX, XXX−XXX

complex. As for binding the 10 nt 5′ OH, due to positioning of the DBDs along the ligand, RPA may be able to use alternative binding modes that allow it bind with modest affinity (discussed further below, see Figure 9D).

Our data predict that the ability of the trimer core to bind ssDNA is important for forming a stable RPA−ssDNA bound complex. This is supported by stopped-flow analysis of full-length *S. cerevisiae* RPA, which suggests that the trimer core may out-compete DBDs A and B for binding when only short stretches of ssDNA are available to bind.[47] Using cryogenic electron microscopy to analyze full-length *S. cerevisiae* RPA, Yates and colleagues were only able to resolve the trimer core bound to ssDNA.[36] Together, these results suggest that the trimer core bound to ssDNA is a critical bound state in many RPA functions.

RPA Uses Diverse Binding Modes to Engage with ssDNA Embedded in Various Structural Contexts. Recent dynamic studies suggest that RPA can engage with ssDNA in a multitude of binding modes with various amounts DBD engagement, challenging the canonical binding modes.[17,21,23,48] Structural studies in related systems have shown that proteins containing tandem OB-folds, such as sequence-specific telomere end protection proteins, can utilize plastic modes of ssDNA recognition to obtain high-affinity interactions with ligands of varying sequences.[70−73] Crystallographic structural data show that *S. pombe* Pot1's C-terminal OB-fold undergoes local or global rearrangement of both the protein and ssDNA that allow it accommodate a degenerate telomeric sequence.[71] *O. nova* TEBP employs a nucleotide shuffling mechanism in which it can contort noncognate ssDNA in a way that allows TEBP to bind it in a pattern similar to binding its cognate ligand.[72] Likewise, RPA's multiple DBDs may be able to adopt various conformations not fully recapitulated by the available crystallographic data. The hallmark of this plastic binding is local, and in some instances global, rearrangement of the ssDNA and/or protein.[71,72] RPA may be to employ similar mechanisms of recognition to bind the wide variety of ssDNA ligands it encounters in the cell. Available in-solution data suggest that RPA can adopt different binding modes depending on ssDNA length and structure. Small-angle X-ray scattering data propose different compaction states of RPA when bound to dT_x ssDNA of various lengths.[50] Additionally, binding to structured ssDNA ligands suggests that some DBDs have enhanced binding depending on the presence of DNA structure (discussed further below).[46,51]

While RPA binds minimal 30 nt ssDNA with higher affinity in a sequence-independent mode, it binds minimal 10 nt ssDNA with lower affinity and higher sequence sensitivity. This would suggest that RPA engages with long and short minimal ssDNA ligands in a distinct manner. With longer stretches of ssDNA available to bind, RPA may utilize redundant binding modes across all four DBDs that allow it to achieve high affinity and low sequence specificity.[51] Mutational analysis of RPA's binding interface has highlighted that when presented with long minimal ssDNAs, other DBDs can help compensate for loss of affinity incurred by mutation of one OB-fold; however, this compensatory effect is lost when the ligand length is ≤10 nts. Importantly, our data show that the use of the canonical polypyrimidine ssDNA does not fully capture RPA's binding activity, especially when assessing binding to short ssDNA ligands.

This plasticity appears to be deployed when RPA engages with biologically representative minimal ssDNA embedded into DNA gaps, bubbles, and OHs. RPA bound these structured ssDNAs in a diverse manner that is largely distinct from its minimal ssDNA-binding activity. RPA bound ssDNA gaps in a length-independent manner, which was not the case for binding bubbles or OHs. Available kinetic data show that RPA binds a dT_{20} ssDNA gap with similar association and dissociation rates as minimal ssDNA, suggesting that RPA binds these ligands in a similar manner.[51] We note, however, that this kinetic analysis was done with homopolymers of dT_x, and as our data suggest, RPA sensitivity to the DNA sequence increases as the length of the ssDNA decreases. Revaluation of binding kinetics in the presence of ssDNA of mixed sequences may help elucidate how/if RPA binds these ligands with different rates. It is possible that constraints of the junctions on both sides of the ssDNA gap prevent RPA from engaging in the fully bound mode. Alternatively, previous data suggest that the ss/dsDNA junction is the trimer core's preferred ligand and this could explain why WT RPA binds all tested ssDNA gap lengths with similar affinities (Figure 9B).[46] Further structural analysis of RPA bound to gapped ssDNA ligands is needed in order to elucidate how it is truly engaged with these structures.

RPA was very sensitive to the length of ssDNA when embedded into DNA bubbles. The 5′-most aromatic residues in each DBD-A and -B, F238 and W361, respectively, are more critical for binding ssDNA in a structured context with high affinity. Previous analysis of OB-A and OB-B aromatic mutants shows that these constructs were unable to fully melt 20 nt ssDNA bubbles.[51] We propose that these base-stacking interactions help facilitate the coordinated ssDNA binding of DBDs A and B which provide the energy to relieve topological strain on bubble ssDNA, allowing for the formation of a stably bound complex (Figure 9C).

Of the DNA structures examined, RPA engages with ssDNA OHs in a manner that most closely resembles its minimal ssDNA-binding activity. One key difference was the observation of a second binding event when binding the 3′ 30 nt OH but not for a ligand with the OH in the opposite polarity. RPA has been shown to undergo concentration-dependent exchange on ssDNA.[74] We hypothesize that at higher RPA concentrations, multiple RPA molecules can bind to the same piece of ssDNA, although it is unclear what mode of ssDNA engagement each RPA molecule is in and how dimerization impacts RPA engagement. Given RPA's binding polarity, it is likely that the 5′ OH positions the trimer core near the ss/dsDNA junction, which may allow for a more stable binding conformation that prevents binding of another RPA molecule (Figure 9D). We note, however, that our models do not fully represent the structural conformation of RPA bound to these ligands and this is just one of many possible mechanisms of action. Further structural and kinetic analysis will need to be done to assess the full scope of conformations RPA adopts on these ligands.

Pathway Discrimination by RPA. RPA participates in a wide range of DNA metabolic pathways, all of which have unique ssDNA contexts that RPA must interact with. RPA's highly modular and dynamic ssDNA binding is thought to be essential for proper functions as it allows for engagement with biologically diverse ligands and assembly of auxiliary protein complexes onto DNA. The aromatic residues that mediate base-stacking interactions are important for modulating this dynamic behavior of RPA.[32,37,51] Molecular dynamic simulations suggest that the aromatic residues act as a DNA

https://doi.org/10.1021/acs.biochem.2c00504
Biochemistry XXXX, XXX, XXX−XXX

anchoring point, which would allow for microscopic dissociation events without full unbinding of the RPA complex.[75] Mutation of two conserved aromatic residues in DBDs A and/or B results in only a modest defect in 30 nt ssDNA binding (2- to 7-fold weaker) but dramatically alters the dynamic state of RPA bound to dT_{35} ssDNA ligands, suggesting that aromatic residues are important for the dynamic, differential DBD binding.[51,66,67] Cells expressing RPA with these mutations reveal a separation of function phenotypes where RPA could support DNA replication but was unable to support DNA repair.[66,67] This is not a simple affinity effect as mutation of polar contact residues in DBD-A results in a nearly 100-fold decrease in 30 nt ssDNA-binding affinity but has no impact on RPA functions in vivo.[66] These findings indicate that the aromatic residues act to help modulate RPA binding modes which are needed to guide RPA functions.

Different DNA processing pathways present various structured ssDNA intermediates that RPA must engage with. Correct positioning of RPA on these intermediates is important for post-translational modification of RPA and for downstream DNA processing as other proteins, such as nucleases, need to be loaded in a way that allows for proper end resection of DNA.[43,76] One interesting aspect of RPA function is its ssDNA protective capacity while balancing its removal from ssDNA for completion of the downstream DNA metabolic pathway. Examples of rapid RPA displacement include displacement from replication start sites by DNA polymerase-α and removal of RPA by Rad52/51 proteins during HR.[25] From a functional standpoint, the lower-affinity mode of RPA engagement with short ($\leq$10 nts) stretches of ssDNA may be important for the rapid removal of RPA from ssDNA by other DNA processing factors in order to carry out their function.

Considering RPA's function at telomeres, RPA has been shown to aid in telomerase elongation of telomeres in vitro, but at too high concentrations, it can be inhibitory to telomerase processivity.[14] Telomeres present unique ssDNA secondary structures, G-quads, and aromatic residues in RPA may play an important function in resolving these structures, similar to RPA's ability to bind ssDNA bubbles. Resolution of G-quads allows for telomerase to bind ssDNA telomeres and perform elongation.[12] Displacement of RPA from telomeric ssDNA is also crucial for proper telomere maintenance. It has been shown that RPA occupancy at telomeres is dependent on a dynamic equilibrium with hnRNPA1, POT1, and TERRA RNA, where hnRNPA1 mediates that handoff of ssDNA from RPA to POT1.[77] RPA's ability to employ differential engagement with ssDNA could be key to its displacement from telomeric DNA. Modulation of RPA's affinity and/or specificity toward telomeric ssDNA may result in increased RPA occupancy at telomeres. Thus, RPA could compete and occlude binding sites from other telomere binding proteins, resulting in aberrant telomeres.

CONCLUSIONS

In summary, this study highlights RPA's diverse modes of engagement with biologically relevant ssDNA structures and the role of aromatic residues within each DBD in accommodating these various ligands. Overall, our binding data to minimal and structured ligands show that RPA can discriminate ssDNA on a variety of axes and, in some cases, engages with structured ssDNA in modes distinct from its interactions with minimal ssDNA. Consistent with past results, we found aromatic residues in DBDs A and B to be important for binding ssDNA, especially in the context of structure. Additionally, mutational analysis of trimer core DBDs has revealed that they play a more significant role in binding short stretches of ssDNA than originally appreciated. Data from this and recent single-molecule studies reveal RPA's dynamic binding characteristics which challenge the static canonical modes of RPA ssDNA binding. Modulation of these binding modes, especially transitions from high-to-low affinity states (and vice versa), could further influence or be influenced by post-translational modifications, such as phosphorylation or SUMOylation, as well as subsequent protein−protein interactions. These insights revealed in this study lay the groundwork for determining how each ssDNA-binding mode drives the in vivo function of RPA.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.biochem.2c00504.

Structural alignment of RPA crystal structures, structure oligo gels, supporting FA data, mass photometry data, tabulated data for dissociation constants and FCs for all binding data, ligand sequences used for binding, mutagenic primer sequences, and anisotropy ranges for binding (PDF)

Accession Codes

RPA70: Uniprot P27694-1, RPA32: Uniprot P15927-1, RPA14: Uniprot P35244-1.

AUTHOR INFORMATION

Corresponding Author

Deborah S. Wuttke − *Department of Biochemistry, University of Colorado Boulder, Boulder, Colorado 80309, United States;* orcid.org/0000-0002-8158-8795; Phone: 303-492-4576; Email: Deborah.wuttke@colorado.edu

Author

Thomas A. Wieser − *Department of Biochemistry, University of Colorado Boulder, Boulder, Colorado 80309, United States;* orcid.org/0000-0002-2390-8631

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.biochem.2c00504

Funding

The work presented here was supported by the National Institute of Health Grants (R01GM139274, D.S.W.), the T32 Molecular Biophysics Predoctoral Training Program (T32GM065103), and the National Science Foundation (MCB1716425, D.S.W.).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We'd like to thank the Walter Chazin lab for gifting us the RPA plasmid. We also thank the Robert Batey Lab at CU-Boulder for access to their plate reader on which the FP data were collected. Additionally, we would like to thank Nickolaus Lammer for writing a program for the facile analysis of binding data (available at https://github.com/nicklammer/AnisotropyBindingFit) and Alexandra Barbour for contributing

the data for Figure S4. We would also like to thank the NIH Biophysics Training program for providing a platform to present and provide feedback on this work. We thank the Shared Instruments Pool (RRID: SCR_018986) of the Department of Biochemistry at the University of Colorado Boulder for the use of the Typhoon FLA 9500 imager. The Typhoon imager is funded by NIH Shared Instrumentation grant S10OD21603.

REFERENCES

(1) Wold, M. S. REPLICATION PROTEIN A: A Heterotrimeric, Single-Stranded DNA-Binding Protein Required for Eukaryotic DNA Metabolism. *Annu. Rev. Biochem.* **1997**, *66*, 61−92.
(2) Wold, M. S.; Kelly, T. 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.* **1988**, *85*, 2523−2527.
(3) Wobbe, C. R.; Weissbach, L.; Borowiec, J. A.; Dean, F. B.; Murakami, Y.; Bullock, P.; Hurwitz, J. Replication of Simian Virus 40 Origin-Containing DNA in Vitro with Purified Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 1834−1838.
(4) Fairman, M. P.; Stillman, B. Cellular Factors Required for Multiple Stages of SV40 DNA Replication in Vitro. *EMBO J.* **1988**, *7*, 1211−1218.
(5) Ashton, N. W.; Bolderson, E.; Cubeddu, L.; O'Byrne, K. J.; Richard, D. J. Human Single-Stranded DNA Binding Proteins Are Essential for Maintaining Genomic Stability. *BMC Mol. Biol.* **2013**, *14*, 9.
(6) Brill, S. J.; Stillman, B. Yeast Replication Factor-A Functions in the Unwinding of the SV40 Origin of DNA Replication. *Nature* **1989**, *342*, 92−95.
(7) Chen, R.; Wold, M. S. Replication Protein A: Single-Stranded DNA's First Responder: Dynamic DNA-Interactions Allow Replication Protein A to Direct Single-Strand DNA Intermediates into Different Pathways for Synthesis or Repair. *BioEssays* **2014**, *36*, 1156−1161.
(8) Krasikova, Y. S.; Rechkunova, N. I.; Lavrik, O. I. Replication Protein A as a Major Eukaryotic Single-Stranded DNA-Binding Protein and Its Role in DNA Repair. *Mol. Biol.* **2016**, *50*, 649−662.
(9) Maréchal, A.; Zou, L. RPA-Coated Single-Stranded DNA as a Platform for Post-Translational Modifications in the DNA Damage Response. *Cell Res.* **2015**, *25*, 9−23.
(10) Oakley, G.; Patrick, S. Replication Protein A: Directing Traffic at the Intersection of Replication and Repair. *Front. Biosci.* **2010**, *15*, 883−900.
(11) Grudic, A.; Jul-Larsen, Å.; Haring, S. J.; Wold, M. S.; Lønning, P. E.; Bjerkvig, R.; Bøe, S. O. Replication Protein A Prevents Accumulation of Single-Stranded Telomeric DNA in Cells That Use Alternative Lengthening of Telomeres. *Nucleic Acids Res.* **2007**, *35*, 7267−7278.
(12) Salas, T. R.; Petruseva, I.; Lavrik, O.; Bourdoncle, A.; Mergny, J.-L.; Favre, A.; Saintomé, C. Human Replication Protein A Unfolds Telomeric G-Quadruplexes. *Nucleic Acids Res.* **2006**, *34*, 4857−4865.
(13) Schramke, V.; Luciano, P.; Brevet, V.; Guillot, S.; Corda, Y.; Longhese, M. P.; Gilson, E.; Géli, V. RPA Regulates Telomerase Action by Providing Est1p Access to Chromosome Ends. *Nat. Genet.* **2004**, *36*, 46−54.
(14) Rubtsova, M. P.; Skvortsov, D. A.; Petruseva, I. O.; Lavrik, O. I.; Spirin, P. V.; Prasolov, V. S.; Kisseljov, F. L.; Dontsova, O. A. Replication Protein A Modulates the Activity of Human Telomerase in Vitro. *Biochemistry* **2009**, *74*, 92−96.
(15) Prakash, A.; Natarajan, A.; Marky, L. A.; Ouellette, M. M.; Borgstahl, G. E. O. Identification of the DNA-Binding Domains of Human Replication Protein A That Recognize G-Quadruplex DNA. *J. Nucleic Acids* **2011**, *2011*, 896947.
(16) Lin, C.-Y. G.; Näger, A. C.; Lunardi, T.; Vančevska, A.; Lossaint, G.; Lingner, J. The Human Telomeric Proteome during Telomere Replication. *Nucleic Acids Res.* **2021**, *49*, 12119.
(17) Fan, J.; Pavletich, N. P. Structure and Conformational Change of a Replication Protein A Heterotrimer Bound to SsDNA. *Genes Dev.* **2012**, *26*, 2337−2347.
(18) Bochkareva, E.; Korolev, S.; Bochkarev, A. The Role for Zinc in Replication Protein A. *J. Biol. Chem.* **2000**, *275*, 27332−27338.
(19) Bochkareva, E.; Kaustov, L.; Ayed, A.; Yi, G.-S.; Lu, Y.; Pineda-Lucena, A.; Liao, J. C. C.; Okorokov, A. L.; Milner, J.; Arrowsmith, C. H.; Bochkarev, A. Single-Stranded DNA Mimicry in the P53 Transactivation Domain Interaction with Replication Protein A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 15412−15417.
(20) Bochkareva, E.; Elena, E.; Korolev, S.; Lees-Miller, S. P.; Bochkarev, A. Structure of the RPA Trimerization Core and Its Role in the Multistep DNA-Binding Mechanism of RPA. *EMBO J.* **2002**, *21*, 1855−1863.
(21) Bochkarev, A.; Pfuetzner, R. A.; Edwards, A. M.; Frappier, L. Structure of the Single-Stranded-DNA-Binding Domain of Replication Protein A Bound to DNA. *Nature* **1997**, *385*, 176−181.
(22) Pfuetzner, R. A.; Bochkarev, A.; Frappier, L.; Edwards, A. M. Replication Protein A: CHARACTERIZATION AND CRYSTALLIZATION OF THE DNA BINDING DOMAIN. *J. Biol. Chem.* **1997**, *272*, 430−434.
(23) Blackwell, L. J.; Borowiec, J. A. Human Replication Protein A Binds Single-Stranded DNA in Two Distinct Complexes. *Mol. Cell. Biol.* **1994**, *14*, 3993−4001.
(24) Bochkareva, E.; Korolev, S.; Lees-Miller, S. P.; Bochkarev, A. Structure of the RPA Trimerization Core and Its Role in the Multistep DNA-Binding Mechanism of RPA. *EMBO J.* **2002**, *21*, 1855−1863.
(25) Fanning, E.; Klimovich, V.; Nager, A. A Dynamic Model for Replication Protein A (RPA) Function in DNA Processing Pathways. *Nucleic Acids Res.* **2006**, *34*, 4126−4137.
(26) Bastin-Shanower, S. A.; Brill, S. J. Functional Analysis of the Four DNA Binding Domains of Replication Protein A: THE ROLE OF RPA2 IN SsDNA BINDING. *J. Biol. Chem.* **2001**, *276*, 36446−36453.
(27) Gomes, X. V.; Wold, M. S. Stuctural Analysis of Human Replication Protein A: Mapping Functional Domains of the 70-KDa Subunit. *J. Biol. Chem.* **1995**, *270*, 4534−4543.
(28) Braun, K. A.; Lao, Y.; He, Z.; Ingles, C. J.; Wold, M. S. Role of Protein−Protein Interactions in the Function of Replication Protein A (RPA): RPA Modulates the Activity of DNA Polymerase α by Multiple Mechanisms. *Biochemistry* **1997**, *36*, 8443−8454.
(29) Chen, R.; Subramanyam, S.; Elcock, A. H.; Spies, M.; Wold, M. S. Dynamic Binding of Replication Protein a Is Required for DNA Repair. *Nucleic Acids Res.* **2016**, *44*, 5758−5772.
(30) Arunkumar, A. I.; Stauffer, M. E.; Bochkareva, E.; Bochkarev, A.; Chazin, W. J. Independent and Coordinated Functions of Replication Protein A Tandem High Affinity Single-Stranded DNA Binding Domains. *J. Biol. Chem.* **2003**, *278*, 41077−41082.
(31) Bochkareva, E.; Frappier, L.; Edwards, A. M.; Bochkarev, A. The RPA32 Subunit of Human Replication Protein A Contains a Single-Stranded DNA-Binding Domain. *J. Biol. Chem.* **1998**, *273*, 3932−3936.
(32) Wyka, I. M.; Dhar, K.; Binz, S. K.; Wold, M. S. Replication Protein A Interactions with DNA: Differential Binding of the Core Domains and Analysis of the DNA Interaction Surface. *Biochemistry* **2003**, *42*, 12909−12918.
(33) Kim, C.; Snyder, R. O.; Wold, M. S. Binding Properties of Replication Protein A from Human and Yeast Cells. *Mol. Cell. Biol.* **1992**, *12*, 3050−3059.
(34) Kim, C.; Paulus, B. F.; Wold, M. S. Interactions of Human Replication Protein A with Oligonucleotides. *Biochemistry* **1994**, *33*, 14197−14206.
(35) Dueva, R.; Iliakis, G. Replication Protein A: A Multifunctional Protein with Roles in DNA Replication, Repair and Beyond. *NAR Cancer* **2020**, *2*, zcaa022.
(36) Yates, L. A.; Aramayo, R. J.; Pokhrel, N.; Caldwell, C. C.; Kaplan, J. A.; Perera, R. L.; Spies, M.; Antony, E.; Zhang, X. A

https://doi.org/10.1021/acs.biochem.2c00504
Biochemistry XXXX, XXX, XXX−XXX

Structural and Dynamic Model for the Assembly of Replication Protein A on Single-Stranded DNA. *Nat. Commun.* **2018**, *9*, 5447.
(37) Bastin-Shanower, S. A.; Brill, S. J. Functional Analysis of the Four DNA Binding Domains of Replication Protein A. *J. Biol. Chem.* **2001**, *276*, 36446−36453.
(38) Iftode, C.; Borowiec, J. A. 5′ → 3′ Molecular Polarity of Human Replication Protein A (hRPA) Binding to Pseudo-Origin DNA Substrates. *Biochemistry* **2000**, *39*, 11970−11981.
(39) Kolpashchikov, D. M. Polarity of Human Replication Protein A Binding to DNA. *Nucleic Acids Res.* **2001**, *29*, 373−379.
(40) Krasikova, Y. S.; Rechkunova, N. I.; Lavrik, O. I. Replication Protein A as a Major Eukaryotic Single-Stranded DNA-Binding Protein and Its Role in DNA Repair. *Mol. Biol.* **2016**, *50*, 649−662.
(41) Li, Y.; Shen, J.; Niu, H. DNA Duplex Recognition Activates Exo1 Nuclease Activity. *J. Biol. Chem.* **2019**, *294*, 11559−11567.
(42) Soniat, M. M.; Myler, L. R.; Kuo, H.-C.; Paull, T. T.; Finkelstein, I. J. RPA Phosphorylation Inhibits DNA Resection. *Mol. Cell* **2019**, *75*, 145−153.
(43) Zhou, C.; Pourmal, S.; Pavletich, N. P. Dna2 Nuclease-Helicase Structure, Mechanism and Regulation by Rpa. *eLife* **2015**, *4*, No. e09832.
(44) de Laat, W. L.; Appeldoorn, E.; Sugasawa, K.; Weterings, E.; Jaspers, N. G. J.; Hoeijmakers, J. H. J. DNA-Binding Polarity of Human Replication Protein A Positions Nucleases in Nucleotide Excision Repair. *Genes Dev.* **1998**, *12*, 2598−2609.
(45) Topolska-Woś, 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.; Chazin, W. J. A Key Interaction with RPA Orients XPA in NER Complexes. *Nucleic Acids Res.* **2020**, *48*, 2173−2188.
(46) Pestryakov, P. E.; Khlimankov, D. Y.; Bochkareva, E.; Bochkarev, A.; Lavrik, O. I. Human Replication Protein A (RPA) Binds a Primer-Template Junction in the Absence of Its Major SsDNA-Binding Domains. *Nucleic Acids Res.* **2004**, *32*, 1894−1903.
(47) Pokhrel, N.; Caldwell, C. C.; Corless, E. I.; Tillison, E. A.; Tibbs, J.; Jocic, N.; Tabei, S. M. A.; Wold, M. S.; Spies, M.; Antony, E. Dynamics and Selective Remodeling of the DNA-Binding Domains of RPA. *Nat. Struct. Mol. Biol.* **2019**, *26*, 129−136.
(48) Caldwell, C. C.; Spies, M. Dynamic Elements of Replication Protein A at the Crossroads of DNA Replication, Recombination, and Repair. *Crit. Rev. Biochem. Mol. Biol.* **2020**, *55*, 482−507.
(49) Kemmerich, F. E.; Daldrop, P.; Pinto, C.; Levikova, M.; Cejka, P.; Seidel, R. Force Regulated Dynamics of RPA on a DNA Fork. *Nucleic Acids Res.* **2016**, *44*, 5837−5848.
(50) Brosey, C. A.; Yan, C.; Tsutakawa, S. E.; Heller, W. T.; Rambo, R. P.; Tainer, J. A.; Ivanov, I.; Chazin, W. J. A New Structural Framework for Integrating Replication Protein A into DNA Processing Machinery. *Nucleic Acids Res.* **2013**, *41*, 2313−2327.
(51) Chen, R.; Subramanyam, S.; Elcock, A. H.; Spies, M.; Wold, M. S. Dynamic Binding of Replication Protein a Is Required for DNA Repair. *Nucleic Acids Res.* **2016**, *44*, 5758−5772.
(52) Branum, M. E.; Reardon, J. T.; Sancar, A. DNA Repair Excision Nuclease Attacks Undamaged DNA. *J. Biol. Chem.* **2001**, *276*, 25421−25426.
(53) Hanawalt, P. C. Subpathways of Nucleotide Excision Repair and Their Regulation. *Oncogene* **2002**, *21*, 8949−8956.
(54) Hegde, M. L.; Hazra, T. K.; Mitra, S. Early Steps in the DNA Base Excision/Single-Strand Interruption Repair Pathway in Mammalian Cells. *Cell Res.* **2008**, *18*, 27−47.
(55) Kubota, Y.; Nash, R. A.; Klungland, A.; Schär, P.; Barnes, D. E.; Lindahl, T. Reconstitution of DNA Base Excision-Repair with Purified Human Proteins: Interaction between DNA Polymerase Beta and the XRCC1 Protein. *EMBO J.* **1996**, *15*, 6662−6670.
(56) Mocquet, V.; Lainé, J. P.; Riedl, T.; Yajin, Z.; Lee, M. Y.; Egly, J. M. Sequential Recruitment of the Repair Factors during NER: The Role of XPG in Initiating the Resynthesis Step. *EMBO J.* **2008**, *27*, 155−167.
(57) Gomes, X. V.; Wold, M. S. Structural Analysis of Human Replication Protein A. *J. Biol. Chem.* **1995**, *270*, 4534−4543.
(58) Walther, A. P.; Gomes, X. V.; Lao, Y.; Lee, C. G.; Wold, M. S. Replication Protein A Interactions with DNA. 1. Functions of the DNA-Binding and Zinc-Finger Domains of the 70-KDa Subunit. *Biochemistry* **1999**, *38*, 3963−3973.
(59) Brosey, C. A.; Chagot, M.-E.; Chazin, W. J.Preparation of the Modular Multi-Domain Protein RPA for Study by NMR Spectroscopy. In *Protein NMR Techniques*; Shekhtman, A., Burz, D. S., Eds.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2012; Vol. *831*, pp 181−195.
(60) Gibson, D. G.; Young, L.; Chuang, R.-Y.; Venter, J. C.; Hutchison, C. A.; Smith, H. O. Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nat. Methods* **2009**, *6*, 343−345.
(61) Stothard, P. The Sequence Manipulation Suite: JavaScript Programs for Analyzing and Formatting Protein and DNA Sequences. *BioTechniques* **2000**, *28*, 1102−1104.
(62) Kim, C.; Wold, M. S. Recombinant Human Replication Protein A Binds to Polynucleotides with Low Cooperativity. *Biochemistry* **1995**, *34*, 2058−2064.
(63) Wang, Q.-M.; Yang, Y.-T.; Wang, Y.-R.; Gao, B.; Xi, X.; Hou, X.-M. Human Replication Protein A Induces Dynamic Changes in Single-Stranded DNA and RNA Structures. *J. Biol. Chem.* **2019**, *294*, 13915−13927.
(64) Ashton, N. W.; Bolderson, E.; Cubeddu, L.; O'Byrne, K. J.; Richard, D. J. Human Single-Stranded DNA Binding Proteins Are Essential for Maintaining Genomic Stability. *BMC Mol. Biol.* **2013**, *14*, 9.
(65) Mishra, G.; Levy, Y. Molecular Determinants of the Interactions between Proteins and SsDNA. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 5033−5038.
(66) Haring, S. J.; Mason, A. C.; Binz, S. K.; Wold, M. S. Cellular Functions of Human RPA1: MULTIPLE ROLES OF DOMAINS IN REPLICATION, REPAIR, AND CHECKPOINTS. *J. Biol. Chem.* **2008**, *283*, 19095−19111.
(67) Hass, C. S.; Lam, K.; Wold, M. S. Repair-Specific Functions of Replication Protein A. *J. Biol. Chem.* **2012**, *287*, 3908−3918.
(68) Bochkareva, E.; Korolev, S.; Bochkarev, A. The Role for Zinc in Replication Protein A. *J. Biol. Chem.* **2000**, *275*, 27332.
(69) Oakley, G. G.; Tillison, K.; Opiyo, S. A.; Glanzer, J. G.; Horn, J. M.; Patrick, S. M. Physical Interaction between Replication Protein A (RPA) and MRN: Involvement of RPA2 Phosphorylation and the N-Terminus of RPA1. *Biochemistry* **2009**, *48*, 7473−7481.
(70) Hom, R. A.; Wuttke, D. S. Human CST Prefers G-Rich but Not Necessarily Telomeric Sequences. *Biochemistry* **2017**, *56*, 4210−4218.
(71) Dickey, T. H.; McKercher, M. A.; Wuttke, D. S. Nonspecific Recognition Is Achieved in Pot1pC through the Use of Multiple Binding Modes. *Structure* **2013**, *21*, 121−132.
(72) Theobald, D. L.; Schultz, S. C. Nucleotide Shuffling and SsDNA Recognition in Oxytricha Nova Telomere End-Binding Protein Complexes. *EMBO J.* **2003**, *22*, 4314−4324.
(73) Lim, C. J.; Barbour, A. T.; Zaug, A. J.; Goodrich, K. J.; McKay, A. E.; Wuttke, D. S.; Cech, T. R. The Structure of Human CST Reveals a Decameric Assembly Bound to Telomeric DNA. *Science* **2020**, *368*, 1081−1085.
(74) Gibb, B.; Ye, L. F.; Gergoudis, S. C.; Kwon, Y.; Niu, H.; Sung, P.; Greene, E. C. Concentration-Dependent Exchange of Replication Protein A on Single-Stranded DNA Revealed by Single-Molecule Imaging. *PLoS One* **2014**, *9*, No. e87922.
(75) Mishra, G.; Bigman, L. S.; Levy, Y. SsDNA Diffuses along Replication Protein A via a Reptation Mechanism. *Nucleic Acids Res.* **2020**, *48*, 1701−1714.
(76) Cappadocia, L.; Kochańczyk, T.; Lima, C. D. DNA asymmetry promotes SUMO modification of the single-stranded DNA-binding protein RPA. *EMBO J.* **2021**, *40*, No. e103787.
(77) Flynn, R. L.; Centore, R. C.; O'Sullivan, R. J.; Rai, R.; Tse, A.; Songyang, Z.; Chang, S.; Karlseder, J.; Zou, L. TERRA and HnRNPA1 Orchestrate an RPA-to-POT1 Switch on Telomeric Single-Stranded DNA. *Nature* **2011**, *471*, 532−536.

https://doi.org/10.1021/acs.biochem.2c00504
Biochemistry XXXX, XXX, XXX−XXX