# Lawrence Berkeley National Laboratory

## Lawrence Berkeley National Laboratory

### Title

A new structural framework for integrating replication protein A into DNA processing machinery

### Permalink

https://escholarship.org/uc/item/838515bz

### Author

Brosey, Chris

### Publication Date

2013-01-17

eScholarship.org Powered by the California Digital Library
University of California

*Published online 8 January 2013* *Nucleic Acids Research, 2013, Vol. 41, No. 4* ***2313–2327***
*doi:10.1093/nar/gks1332*

# A new structural framework for integrating replication protein A into DNA processing machinery

**Chris A. Brosey[1], Chunli Yan[2], Susan E. Tsutakawa[3], William T. Heller[4,5], Robert P. Rambo[3], John A. Tainer[3,6,7], Ivaylo Ivanov[2,*] and Walter J. Chazin[1,8,*]**

[1]Department of Biochemistry, Center for Structural Biology, Vanderbilt University, Nashville, TN 37232, USA, [2]Department of Chemistry, Georgia State University, Atlanta, GA 30302-4098, USA, [3]Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA, [4]Center for Structural Molecular Biology, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA, [5]Biology & Soft Matter Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA, [6]Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037, USA, [7]The Skaggs Institute for Chemical Biology, La Jolla, CA 92037, USA and [8]Department of Chemistry, Vanderbilt University, Nashville, TN 37232, USA

Received September 5, 2012; Revised November 26, 2012; Accepted November 27, 2012

## ABSTRACT

**By coupling the protection and organization of single-stranded DNA (ssDNA) with recruitment and alignment of DNA processing factors, replication protein A (RPA) lies at the heart of dynamic multi-protein DNA processing machinery. Nevertheless, how RPA coordinates biochemical functions of its eight domains remains unknown. We examined the structural biochemistry of RPA's DNA-binding activity, combining small-angle X-ray and neutron scattering with all-atom molecular dynamics simulations to investigate the architecture of RPA's DNA-binding core. The scattering data reveal compaction promoted by DNA binding; DNA-free RPA exists in an ensemble of states with inter-domain mobility and becomes progressively more condensed and less dynamic on binding ssDNA. Our results contrast with previous models proposing RPA initially binds ssDNA in a condensed state and becomes more extended as it fully engages the substrate. Moreover, the consensus view that RPA engages ssDNA in initial, intermediate and final stages conflicts with our data revealing that RPA undergoes two (not three) transitions as it binds ssDNA with no evidence for a discrete intermediate state. These results form a framework for understanding how RPA integrates the ssDNA substrate into DNA processing machinery, provides substrate access to its binding partners and promotes the progression and selection of DNA processing pathways.**

## INTRODUCTION

Replication protein A (RPA) is a modular multi-domain protein that functions in a wide range of DNA processing pathways required to maintain and propagate the genome of all living organisms. RPA functions by interfacing with dynamic multi-protein machinery and acts as a central hub that links many DNA transactions. RPA provides the primary single-stranded DNA (ssDNA) binding activity in eukaryotes and also serves as a scaffold and coordinator of DNA processing machinery (1,2). Binding of ssDNA is critical for shielding DNA strands from endonuclease activity and preventing the formation of disruptive secondary structures. RPA couples this activity to the recruitment of DNA processing factors, thereby providing a platform for organization of DNA processing machinery and managing access to the DNA substrate. Conjectures have been made about how protein interactions couple to the DNA binding activity of RPA (2), but uncertainty remains about how this might occur, and there is no structural framework to work from for intact RPA.

Despite its central importance in DNA processing machinery, little information is available on the physical basis for the coordination of RPA functions. This is in large part because its modular nature poses a significant

*To whom correspondence should be addressed. Tel: 1+ 615 936 2210; Fax: 1+ 615 936 2211; Email: walter.chazin@vanderbilt.edu
Correspondence may also be addressed to Ivaylo Ivanov. Tel: 1+ 404 413 5529; Fax: 1+ 404 413 5505; Email: iivanov@gsu.edu
Present address:
Chris A. Brosey, Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, MO 63110, USA.

© The Author(s) 2013. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/3.0/), which permits non-commercial reuse, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com.

challenge for current structural methods. RPA is a heterotrimer of RPA70, RPA32 and RPA14 subunits organized into five structural modules connected by flexible linkers (70N, 70AB, 70C/32D/14, 32N, 32C; see Figure 1). The trimer core has one domain from each subunit, from which extend the remaining modules. Nuclear magnetic resonance (NMR) experiments on intact RPA have shown that the other modules are structurally independent of the core (3). The dynamic independence of the structural modules makes techniques, such as X-ray crystallography challenging to apply to full-length RPA and unlikely to fully capture the functionally relevant ensemble in solution.

The binding of ssDNA by RPA has been studied for >20 years, and it is generally held that RPA has three discrete DNA-binding modes (2,4,5). Four domains (70A, 70B, 70C and 32D) are known to engage ssDNA with progressively higher affinity and 5′–3′ polarity, respectively. An initial 8–10 nucleotide (nt)-binding mode involves the tandem domains 70A and 70B, which are connected by a short 10 residue flexible linker. A poorly characterized intermediate binding mode has been suggested, encompassing an excluded site size of 12–23 nt. In addition to 70AB, this mode is presumed to engage 70C. The final binding mode engages a second domain from the trimer core, 32D, and occludes up to 30 nt of ssDNA (6,7).

As DNA processing proceeds, RPA must navigate between its different DNA-binding states. The differences in the number of domains directly contacting the DNA in the three ssDNA binding modes are expected to result in significant differences in the spatial organization of the DNA-binding apparatus of RPA. X-ray diffraction, NMR and scattering studies of isolated 70AB have provided insight into the initial ssDNA-binding mode (8–11). Binding of 8–10 nt of ssDNA aligns and compacts the domains, although the complex remains dynamic, presumably as a consequence of torsional motion between the two domains (11). The extent to which isolated 70AB typifies the action of intact RPA during DNA binding is not known. Here, we combine small angle X-ray and neutron scattering (SAXS and SANS) experiments with all-atom molecular dynamics (MD) simulations to determine the effects of ssDNA binding on the architecture of RPA. These results provide new perspectives on how coupling of protein and DNA-binding activities drive changes in the architecture and progression of DNA processing machinery.

## MATERIALS AND METHODS

### Materials

The pET15b vector containing human RPA DNA-binding core (RPA-DBC, $RPA70_{181–616}/32_{43–171}$/14, RPA70ABC/32D/14) was a kind gift of A. Bochkarev. Thrombin cleavable, 6×-His fusion tags precede the N-termini of the 70ABC and 14 subunits. Active thrombin was purchased from CALBIOCHEM. All ssDNA substrates-$dCCAC_7$, $dCCAC_{17}$, $dCCAC_{21}$, $dCCAC_{24}$ and $dCCAC_{27}$ (d10, d20, d24, d27 and d30, respectively), as well as fluorescently modified $dC_{10}$, $dC_{20}$ and $dC_{30}$ oligonucleotides were purchased from Integrated DNA Technologies (IDT) with standard desalting purification and were resuspended in sterile distilled water before use (sequences were validated by electrospray ionization mass spectrometry (ESI-MS) analysis performed by IDT). Full-length RPA used in these studies was prepared as described (3) and provided as a kind gift by Dr S. Michael Shell.

### Expression and purification of recombinant RPA-DBC

Expression and purification of the RPA-DBC have been described previously by the Bochkarev laboratory during its initial biochemical description of the construct (12,13). Modifications to this protocol, along with characterization of monodispersity by size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALS) are provided in Supplementary Materials.

### Fluorescence anisotropy ssDNA-binding assays

The ssDNA-binding activity of RPA-DBC was assessed by a rise in fluorescence anisotropy as increasing amounts of protein were added to polycytidine substrates labelled at their 5′-ends with 6-carboxyfluorescein (5′-FAM-$dC_{10}$, -$dC_{20}$, -$dC_{30}$). Triplicate serial dilutions of protein (0–0.5 μM) were prepared in 384-well plates with 20 mM HEPES–KOH (pH 7.5), 200 mM NaCl and 10 mM β-mercaptoethanol, then mixed with fluorescently labelled ssDNA (final concentration 25 nM). Polarized fluorescent intensities were measured with a Spectramax M5 plate reader (Molecular Machines) at excitation and emission wavelengths of 492 and 520 nm, respectively, for 100 s (1 reading/s) and averaged. Dissociation constants ($K_d$) were calculated by fitting the data to a simple two-state binding model in KaleidaGraph (v. 3.51).

### Preparation of RPA-DBC/ssDNA complexes for SAXS

Purified RPA-DBC was concentrated to 7–12 mg/ml, combined with 1.5–2-fold molar excess ssDNA substrate (d10, d20, d24, d27 or d30) and incubated on ice for 6–18 h (d10, d20) or at room temperature for 30 min (d24, d27, d30). To remove excess ssDNA and ensure homogeneous complexes, samples (300 μl) were injected onto a Superdex 200 HR 10/30 gel filtration column (GE Healthcare) equilibrated overnight in 20 mM HEPES–KOH (pH 7.5), 200 mM NaCl, 5 mM dithiothreotol (DTT) and 2% glycerol. Samples for RPA-DBC, RPA-DBC/d10 and RPA-DBC/d20 eluted as single peaks. Samples for RPA-DBC/d24, RPA-DBC/d27 and RPA-DBC/d30 eluted with high-molecular weight shoulders, but conservative fractionation (290 μl fractions) of each elution profile allowed each 2–3 ml peak width to be completely isolated from multiply DNA-bound species or free ssDNA. The presence of both protein and DNA in each fraction was confirmed by ultraviolet absorbance readings. Calculations based on the $K_d$ values determined for ssDNA binding to RPA-DBC and the concentrations of protein and DNA used for sample preparation indicated that the amount of free protein was <3% for the 10mer complex and ≤1% for the longer ssDNA substrates. Experimental extinction coefficients calculated

from each gel filtration fraction used for SAXS confirmed 1:1 stoichiometry for the purified complexes.

### Small-angle X-ray scattering data collection and analysis

SAXS data were collected at the SIBYLS beamline 12.3.1 at the Advanced Light Source, Lawrence Berkeley National Laboratory. Scattering measurements were performed on 20 μl samples at 15°C using a Hamilton robot for loading samples from a 96-well plate into a helium-purged sample chamber (14,15). Data were collected on the original gel filtration fractions from each SEC run, as well as concentration series from fractions sampled from the later eluting half of each SEC elution peak (~2–8-fold concentration). Fractions before the SEC void volume were used for buffer subtraction of the original 1× gel filtration fractions, and concentrator eluates were used for buffer subtraction of each concentration series (2–8×).

SAXS experiments were performed using an X-ray beam from a multilayer monochromator of 12 keV (λ 1.3 Å) covering the following momentum transfer range: $0.011\,\text{Å}^{-1} < q < 0.322\,\text{Å}^{-1}$, where $q$ is defined as $q = 4\pi \sin(\theta/2)/\lambda$ with scattering angle θ and wavelength λ. The multilayer monochromator provides increased X-ray flux, allowing stronger signals for lower protein concentrations. Sequential exposures (0.5, 0.5, 2, 5 or 6 and 0.5 s) were taken, and data were monitored for radiation-dependent aggregation. All SAXS data were collected using the MarCCD 165 detector in fast frame transfer mode and reduced via normalization to the incident beam intensity. Standard procedures were used for processing the data and are described in detail in the Supplementary Materials. SAXS data from this publication have been submitted to the BIOSIS database (http://bioisis.net).

SANS experiments were performed using the Bio-SANS instrument at the High Flux Isotope Reactor of Oak Ridge National Laboratory (16), as described in the Supplementary Materials.

### Interpreting RPA-DBC P(r) distributions

A model of free RPA-DBC was constructed from crystal structures of 70AB (PDB ID: 1FGU) and the trimer core (PDB ID:). Residues for the B–C linker were added in PyMOL (17) and joined using the Modeller (9v4) interface in Chimera (18). Multiple inter-domain arrangements were generated from these starting models using BILBO-MD (19), from which models were selected that matched the experimental $D_{max}$ value of free RPA-DBC. Inter-domain distances were assessed in PyMOL and compared with features in the experimental P(r) distribution. Assignment of P(r) features to specific inter-domain distances were confirmed by simulating P(r) distances from the models (using the FoXS server and GNOM) and examining the impact of removing a given domain from the model (Supplementary Figure S3). A similar process was used to investigate P(r) distributions of the 10-nt and 20-nt complexes. The 10-nt model was generated in a manner similar to the free protein, with substitution of DNA-bound coordinates for 70AB (1JMC), while the 20-nt model was taken from the trajectory of the compact 20-nt MD simulation.

### All-atom molecular dynamics simulations

Crystal structures for 70AB, 70AB/dC8 and RPA70C/32D/14 were obtained from the RCSB Protein Data Bank [PDB ID: 1FGU (9), 1JMC (8) and 1L1O (13)]. Five models were constructed: (i) DNA-free RPA-DBC; (ii) RPA-DBC/dC10; (iii) RPA-DBC/dC20 compact; (iv) RPA-DBC/dC20 extended; and (v) RPA-DBC/dC30 (two independent simulations performed). The length of the ssDNA was adjusted to correspond to the initial (dC10) and final (dC30) modes of binding ssDNA, and one of the putative intermediate states (dC20). The DNA-free RPA-DBC model was built by connecting 70AB (9) and RPA70C/32D/14 (13). The RPA-DBC/dC10 complex was created directly from the 70AB/$dC_8$ (8) and RPA70C/32D/14 crystal structure (13). For the 20mer models, two limiting cases were considered: the first engaged only the RPA70A, B and C domains with nine bases positioned between the 70AB and RPA70C/32D/14 ('extended'). In the compact RPA-DBC/dC20 (and also in the RPA-DBC/dC30 complex) all four domains, 70A, 70B, 70C and 32D, together directly contact ssDNA. For the RPA-DBC/dC20 and RPA-DBC/dC30 complexes, the RPA70C and 32D domains were assumed to bind ssDNA in a binding mode identical to the crystal structure of 70AB (8). The RPA70B domains were aligned to the RPA70C and 32D domains, respectively, by using homologous structural elements, and the position of the ssDNA fragment was identified based on this alignment. The residues Phe532, Tyr581 from RPA70C and residues Trp107, Phe135 from 32D were found in position to stack with the ssDNA bases. Missing loops and the linker between 70B and 70C domains of RPA were built with the program Modeller 9v4 (20).

Hydrogen atoms were introduced using the tLeap module of AMBER 11 (21). To accelerate sampling of the conformational ensemble for each of the systems, we carried out the simulation without inclusion of explicit water using a modified Generalized Born implicit solvent (GBIS) model (22,23), which dramatically reduces the number of degrees of freedom in the system and speeds up the sampling of domain motions. All systems were minimized for 5000 steps with backbone atoms fixed, followed by 5000 steps of minimization with harmonic restraints to remove unfavourable contacts. The systems were then gradually brought up to 300 K and run for 50 ps while keeping the protein backbone restrained. The ssDNA substrates were 'fixed' into each DNA-binding domain during the equilibration period only by defining distance restraints that maintained canonical base-stacking interactions between aromatic residues and DNA bases based on the initial X-ray crystal structures (8). The equilibration was continued for another 50 ns, and the harmonic restraints were gradually released. Production runs were continued for 140 ns for the RPA-DBC system and 200 ns for all ssDNA-containing systems. Two independent simulations for RPA-DBC/

dC30 system were carried out. All simulations used an integration time step of 2 fs with the SHAKE algorithm being applied to fix the bonds between hydrogen atoms and heavy atoms in the systems. The r-RESPA multiple time step method (24) was adopted with a 2 fs time step for bonded, 2 fs for short-range non-bonded and 4 fs for long-range electrostatic interactions. The cut-off for non-bonded interactions and computation of the effective Born radii was set to 18 Å. Dielectric constants of 1 (interior) and 78.5 (exterior) were used in all GBIS-MD simulations. Ionic strength was set to 0.15 M to mimic physiological conditions. All simulations were performed using the NAMD 2.8 code (25,26) with the AMBER Parm99SB parameter set (27) containing the basic force field for nucleic acids and proteins, as well as the refined parameters for backbone dihedrals for protein (SB) and nucleic acids dihedrals (BSC0) on Hopper II, a Cray XE6 system at the National Energy Research Scientific Computing Center.

### Analysis of MD simulations

Computation of scattering profiles from the simulation and comparison to the experimental data used the FoXS code (28). For each of the five models, we computed theoretical scattering profiles for all conformations in the trajectory (last 100 000 frames or 200 ns, except for the DNA-free model with 140 ns and 70 000 frames). The first 50 ns were excluded as equilibration. The theoretical scattering curves were then averaged over the entire ensemble, and these average profiles were compared with the experimental scattering data. This approach of calculating the X-ray scattering from the entire trajectory is distinct from minimal ensemble searches, which typically select only three to five models based on ensemble fits to the data. Averaging over the entire conformational ensemble is expected to better represent solution-phase scattering compared to using a single conformer. To further characterize the structural ensemble, root mean square deviation (rmsd)-based clustering was performed using PTRAJ utility in AMBER 11 and based on a previously reported clustering algorithm (29). The 100 000 frames taken from the trajectory were clustered according to the RMSD, using a cut-off of 6 Å. Data were further analysed using the PTRAJ utility and custom VMD TCL scripts. The 2D histogram figures were generated using Origin 8.0.

## RESULTS

The DNA-binding apparatus of RPA consists of the tandem domains 70AB and the trimer core (Figure 1). As the other three RPA domains (70N, 32N, 32C) are structurally independent (3,11,30,31), we used a construct containing only the DNA-binding core (70ABC/32D/14, RPA-DBC) (12,13) (Figure 1C and D). This construct facilitates interpretation of the scattering data because the extra non-DNA–binding domains would add complexity to the analysis without improving the information content in regard to the binding of ssDNA. The long linkers to the 32C and 70N domains are flexible, and 32N is dynamically

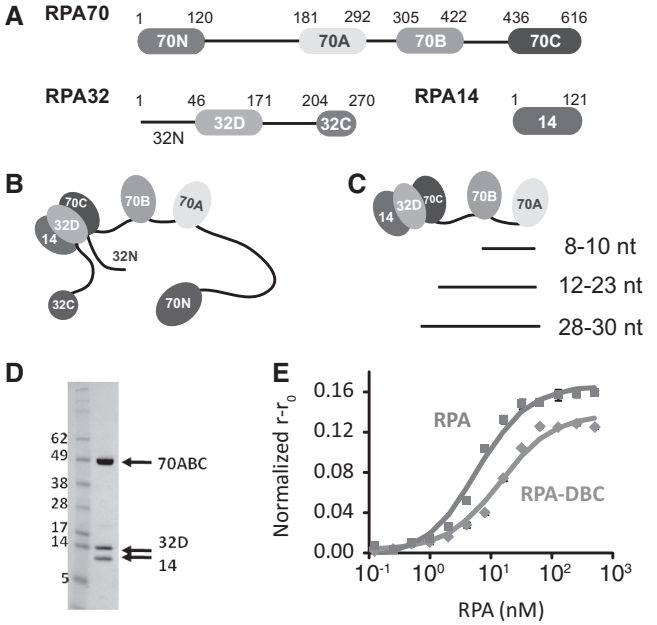


**Figure 1.** Modular, multi-domain RPA has an independent DNA-binding core. (**A**) Domain organization of the RPA70, RPA32 and RPA14 subunits. (**B**) The RPA trimer interface involves domains 70C, 32D and 14. Flexible linkers connect the remaining domains. (**C**) Three proposed modes of RPA binding of ssDNA: the initial 8–10 nt, intermediate 12–23 nt and final 28–30 nt modes engage domains 70A–70B, 70A–70B–70C and 70A–70B–70C–32D, respectively. (**D**) Sodium dodecyl sulphate–polyacrylamide gel electrophoresis of purified RPA-DBC, with molecular weight standards (left lane) in kDa. (**E**) Binding curves from fluorescence anisotropy assays of RPA (squares) ($K_d$ 5.8 ± 0.4 nM) and RPA-DBC (diamonds) ($K_d$ 14.2 ± 0.2 nM) binding to $dC_{30}$. Solid lines are fits to a two-state binding model.

disordered; thus, the architecture of the isolated DBC in the absence and presence of ssDNA is expected to be the same as that in full-length RPA.

To test that this model system accurately reproduces the behaviour of full-length RPA, ssDNA-binding affinities were measured using a fluorescence anisotropy assay. To ensure binding of 70A at the 5′-end of each substrate, poly-dC substrates were used with a single adenine at position 3, as reported previously (11). As anticipated, the affinities of RPA and RPA-DBC were similar (Figure 1E) and consistent with those reported previously (32). Subsequent SEC-MALS analysis confirmed the monodispersity of RPA-DBC alone and in complex with its ssDNA substrates (Supplementary Figure S1) and established optimal solution conditions for examining this system by small-angle scattering (33).

The effect of binding ssDNA on the architecture of RPA was investigated using both SAXS and SANS, powerful low-resolution techniques for characterizing the structure of proteins and protein complexes in solution (34–36). SAXS data were acquired on isolated RPA-DBC and complexes with 10-, 20-, 24-, 27- and 30-nt substrates, whereas complementary SANS contrast variation experiments were performed only on the 30-nt complex (Supplementary Figure S2). Guinier analysis verified an absence of sample aggregation in all cases (Supplementary Figures S2 and S6). Joint analysis of the

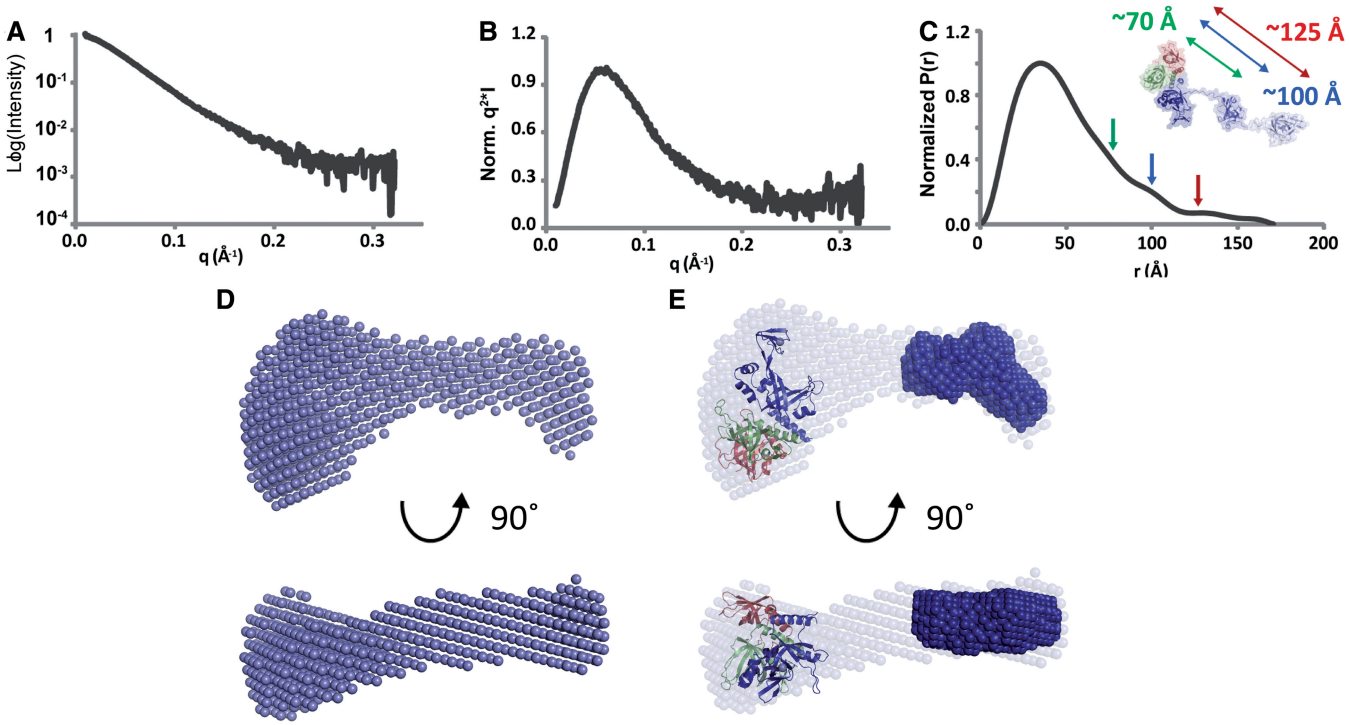


**Figure 2.** DNA-free RPA-DBC has a dynamic but condensed architecture. (**A**) Experimental scattering profile at ~2.4 mg/ml, (**B**) Kratky analysis and (**C**) P(r) distribution of RPA-DBC. (Inset) Arrows identify features in the P(r) function corresponding to distances between domain centres-of-mass: 70A (red), 70B (green) and 70AB (blue) to-trimer core. (**D**) Averaged *ab initio* molecular envelope of RPA-DBC from 15 runs of GASBOR. (**E**) Docking of the 70AB scattering envelope (dark blue) and trimer core crystal structure (ribbon) into the RPA-DBC envelope. The bottom panel in D and E is rotated 90° towards the viewer.

SANS and SAXS data for the 30-nt complex was found to be in full agreement with the analysis performed on the SAXS data alone (Supplementary Figure S2). The lower signal-to-noise of the SANS data resulting from the inherently lower flux available on SANS instruments did not provide additional constraints on the models. As only SAXS data were collected for isolated RPA-DBC and the remaining ssDNA complexes, the results described later in the text are based on the SAXS data alone.

### The RPA DNA-binding core exhibits a dynamic, but not fully extended, architecture

The scattering profile, I(q), of free RPA-DBC reveals a dynamic architecture with flexibility between RPA70A, RPA70B and the trimer core, as plots of I(q) versus q lack the fine structure expected for a fixed spatial arrangement (Figure 2A) (37,38). Convergence in the Kratky transformation at high q-values indicates a significant percentage of ordered globular structure (Figure 2B) (34). Interestingly, Guinier analysis yields an $R_g$ value of 38.8 Å, substantially smaller than that back-calculated from models in which the linkers between domains are fully extended (~53 Å) (e.g. Supplementary Figure S3). This suggests that although RPA-DBC occupies a range of architectures, it favours those that are more condensed than fully extended.

The distance distribution function, P(r), of free RPA-DBC exhibits a maximum at ~34 Å originating from paired electron distances within the OB-fold domains (Figure 2C). Shallow secondary shoulders are observed at ~70, ~100 and ~130 Å, which we attribute to paired electron distances between trimer core and 70B, 70AB and 70 A, respectively (Figure 2C inset, Supplementary Figure S3). The considerable smoothing of these secondary peaks is consistent with dynamic averaging of paired electron distances between RPA-DBC modules (38). The $D_{max}$ of 171 Å indicates a finite population of architectures in the ensemble with full extension of the A–B and B–C linkers (Figure 2C and Supplementary Figure S3), also consistent with dynamic averaging. The presence of inter-domain fluctuations agrees with previous SAXS and NMR studies on RPA70AB and the full-length protein (3,10,11).

To gain further insights into the ensemble of RPA-DBC architectures, *ab initio* low-resolution structural models (molecular envelopes) were generated from the data and analysed using the existing high-resolution structures of 70AB and the trimer core (8,9,13). Examination of the RPA-DBC scattering envelope reveals a 'bilobal' organization (Figure 2D). The globular trimer core can be accommodated by the larger globular end of the envelope, whereas the extended protrusion is suggestive of the 70A and 70B domains, which are expected to be dynamic in the absence of ssDNA (3,11). Automated docking of the high-resolution structures using the *collage* module of SITUS supports these domain assignments within the molecular envelope (Figure 2E). Nevertheless, 70AB crystal structures fail to fit into the

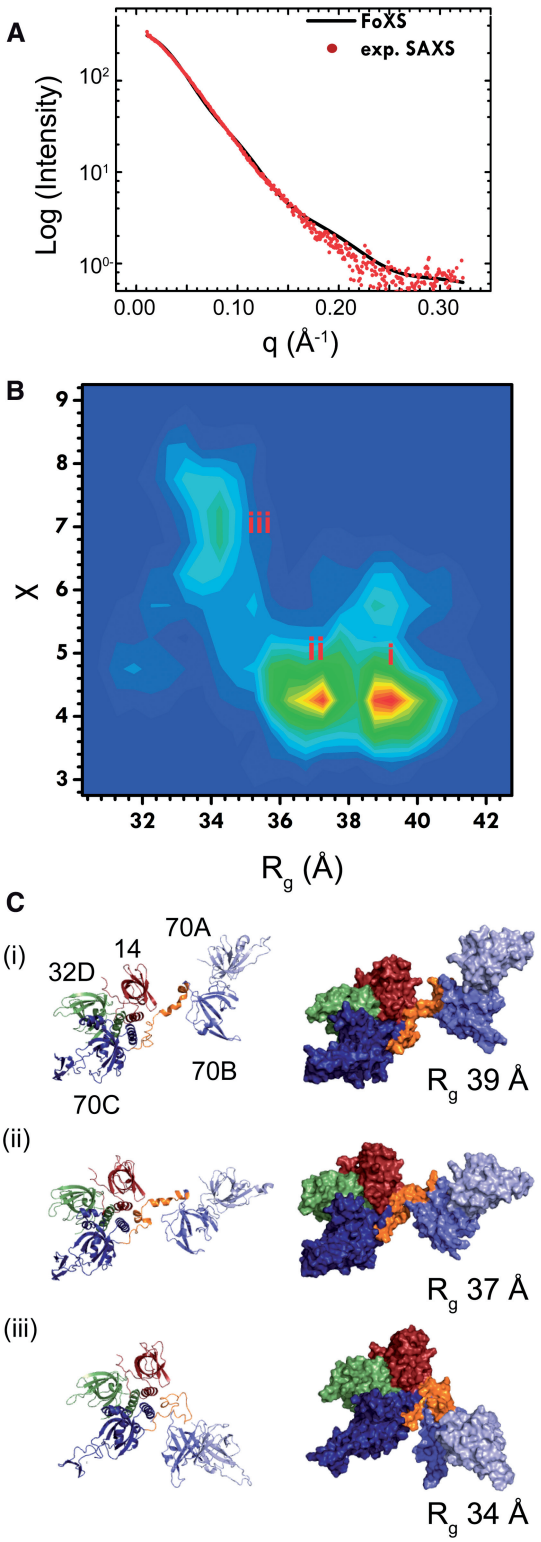


**Figure 3.** MD simulations provide atomic insights into the architecture of DNA-free RPA-DBC. (**A**) Overlay of the experimental scattering curve (red) with the averaged scattering curve (black) calculated from the 70 000 MD models ($\chi = 3.2$). (**B**) 2D histogram comparing goodness-of-fit to the experimental data ($\chi$) versus radius-of-gyration ($R_g$) for each model. Roman numerals indicate the primary populations observed in the plot. (**C**) Ribbon (left) and surface (right) diagrams of representative models of RPA-DBC from each $R_g$ population in 3B. Colour-coding: 70A (light blue), 70B (blue), 70C (dark blue), 32D (green), 14 (red) and B–C linker (orange).

corresponding region of the RPA-DBC envelope, which we attribute to heterogeneity arising from dynamic inter-domain fluctuations (Supplementary Figure S4). In contrast, the low-resolution SAXS envelope of 70AB (11), which incorporates the inherent inter-domain dynamics, provides an excellent fit to the relevant portion of the RPA-DBC envelope (Figure 2E). The remainder of the extended density between the 70AB and trimer core is presumed to arise from fluctuating residues in the B–C linker, although we note that molecular envelopes generally tend to over-estimate extended architectures.

### Inter-domain linkers limit extension between DNA-binding domains

The presence of inter-domain dynamics within the DNA-binding core suggests that RPA-DBC can assume a range of inter-domain arrangements in solution. At the same time, its notably low $R_g$ value indicates that RPA-DBC has a significant preference for more condensed architectures over those that are fully extended. To investigate the conformational space accessible to RPA-DBC, all-atom MD simulations were performed on the DNA-free protein.

Discerning molecular details about the structural dynamics of RPA-DBC poses a fundamental challenge for simulation: the computational cost of all-atom simulations during the time needed to fully characterize the movements of RPA domains is prohibitive. To resolve this problem, 140 ns of all-atom simulations were performed with a GBIS model. The use of GBIS has two principal benefits. First, the number of degrees of freedom in the system is dramatically reduced resulting in simulation speed up. Second, the absence of explicit solvent molecules eliminates their kinetic hindrance of the motions of the RPA domains resulting in accelerated sampling of conformational space. To evaluate the simulation globally, scattering profiles were back-calculated from the full trajectory and averaged. The averaged profile was then compared with the experimental scattering data. Variations in RPA-DBC architecture throughout the trajectory were also examined (Figure 3 and Supplementary Figures S5 and S9).

The averaged scattering profile of the 70 000 models sampled shows reasonable correspondence to the experimental scattering data (Figure 3A), considering the complexity of sampling the architecture of large flexible biomolecules and the potential biases in the modelling from the force field and the implicit solvent model. As expected for a protein with flexible linkers, there is no single conformation that perfectly fits the data, and many conformations agree with the data equally well (Figure 3B). Notably, although the simulation samples architectures across a relatively broad range of $R_g$ values (~30–43 Å), the majority of architectures are close to the experimental $R_g$ of 38.8 Å (Figure 3B). To better characterize the inter-domain arrangements, distances between 70A and 70B and between 70B and the trimer core were calculated for each model, and the magnitude of oscillation in these distances was plotted (Supplementary Figures S5 and S9). In addition, the inter-domain dihedral angle

describing the spatial relationship of 70A, 70B and the trimer core was calculated and plotted versus the distance between 70B and the trimer core (Supplementary Figure S5). The A–B and B–C inter-domain distance ranges (~28–43 and ~40–75 Å, respectively) are consistently less than the distances when the A–B and B–C linkers are fully extended (~54 and ~70 Å, respectively). Similarly, the dihedral angle range (~90–150°) is less than the 180° value for an extended linear arrangement, revealing that the domains occupy an arc. The oscillations in these parameters are notable, for example, the 70A–70B distance varies broadly over the course of the simulation with multiple states in between (Supplementary Figures S5 and S9). Overall, the simulations predict that although RPA-DBC is found to substantially favour architectures that are more compact than fully extended, 70A, 70B and the trimer core are uncoupled and moving independently in the DNA-free protein.

To identify specific structural features that might drive the preference for less extended architectures in RPA-DBC, we inspected representative models, which, as expected, were only partially extended and exhibited a curved arrangement of domains (Figure 3C). Two representative features observed consistently across the simulation were seen in both structures: (i) a transient helix in the B–C linker, which results in shorter distances between 70B and 70C and (ii) the C-terminal helix of 70B uncoils and merges into the B–C linker. Examination of the trajectory in its entirety highlighted several transient non-specific interactions involving the B–C linker with the adjacent 70B domain and trimer core, which also stabilize less extended architectures. Although interpreting the relevance of such contacts is limited by the use of implicit solvent in the simulation, these interactions constrain the inter-domain sampling and, therefore, contribute to the higher representation of models with less extended architectures. Hence, these interactions may explain the preference for condensed over fully extended architectures indicated by the SAXS data.

Overall, the simulation provides reasonable agreement with the experimental data, for example, reproducing the preference of DNA-free RPA-DBC for more compact than fully extended architectures. However, there are some differences. The most significant is that the simulations do not sample a population of fully extended architectures, as reflected in the gradual slope of the *P(r)* function, as it extends out ~170 Å (Figure 2C). Although there appear to be some limitations in exploring accessible conformational space, the simulations do provide excellent models for identifying specific structural features at the origin of the key characteristics of RPA-DBC architecture.

### Transition to the ~10-nt binding mode results in compaction of RPA-DBC

SAXS analysis of RPA-DBC bound to the 10-nt substrate reveals a significant reduction in the $R_g$ value and $D_{max}$ in the P(r) distribution relative to the DNA-free protein, showing a DNA-induced change in architecture leading to a compaction of RPA-DBC (Table 1, $\Delta R_g$ 2.5 Å, $\Delta D_{max}$ 47 Å, Figure 4B). Compaction of RPA-DBC is consistent with the known involvement of the high affinity RPA70AB domains in the first mode of DNA-binding. Although we find that binding of RPA-DBC to 10-nt of ssDNA is associated with a significant overall compaction, detailed analysis of the SAXS data reveals that RPA-DBC retains inter-domain flexibility but not to the full extent of the DNA-free protein (*infra vide*).

**Table 1.** Summary of measurements from SAXS Studies

| | Guinier $R_g$ (Å)[a] | Real-space $R_g$ (Å)[b] | $D_{max}$ (Å)[b] | Porod volume (Å$^3$)[c] |
|---|---|---|---|---|
| RPA-DBC | 38.8 (0.2) | 42.1 | 171 | 167 905 |
| d10 | 36.3 (0.1) | 37.5 | 124 | 149 693 |
| d20 | 34.9 (0.07) | 35.1 | 114 | 136 743 |
| d24 | 34.9 (0.04) | 35.5 | 127 | 128 189 |
| d27 | 34.9 (0.05) | 35.2 | 119 | 134 010 |
| d30 | 34.9 (0.06) | 35.4 | 116 | 139 054 |

[a]Gunier radius-of-gyration ($R_g$) determined using the Primus software package (62). Uncertainties are given in parentheses.
[b]Real-space $R_g$ and $D_{max}$ were determined from calculated pairwise distance distribution functions using GNOM.
[c]Porod volumes were calculated directly from the experimental data according to equations in (39).

The P(r) distribution of the 10-nt complex exhibits the single maximum associated with the distances within the OB-fold domains noted earlier in the text and in addition, a shallow shoulder at ~70 Å that is attributed to distances between the trimer core and the paired 70AB domains (Supplementary Figure S3). The explicit shoulders at ~100 and ~130 Å and the shallow slope out to ~170 Å observed for the DNA-free protein are no longer evident. The a*b initio* molecular envelope calculated from the scattering data (Figure 4C) for the 10mer complex is more compact than for the DNA-free protein, which correlates with the changes observed in $R_g$ and the P(r) function. Docking of the crystal structures of the trimer core and 70AB-$dC_8$ into the molecular envelope provides a better overall fit than that observed for the DNA-free protein. Although significant condensation of the envelope is observed in the extended region fit to RPA70AB (Figure 4C, right), residual density remains between 70AB and trimer core. This residual density could be due to flexibility between these modules, although as noted earlier in the text, we can not be certain because of the tendency towards over-estimation of extended architectures in molecular envelope calculations. Because the compaction of the P(r) function relative to the DNA-free protein indicates fully extended architectures are not sampled, the residual flexibility in the 10-nt complex must involve more limited extensions and torsional motions between the RPA70AB and trimer core modules.

To gain atomic insight into the binding to 10 nt of ssDNA, MD simulations were performed and analysed in the same manner as DNA-free RPA-DBC. The averaged scattering profile calculated from the ensemble of 100 000 models shows general agreement with the

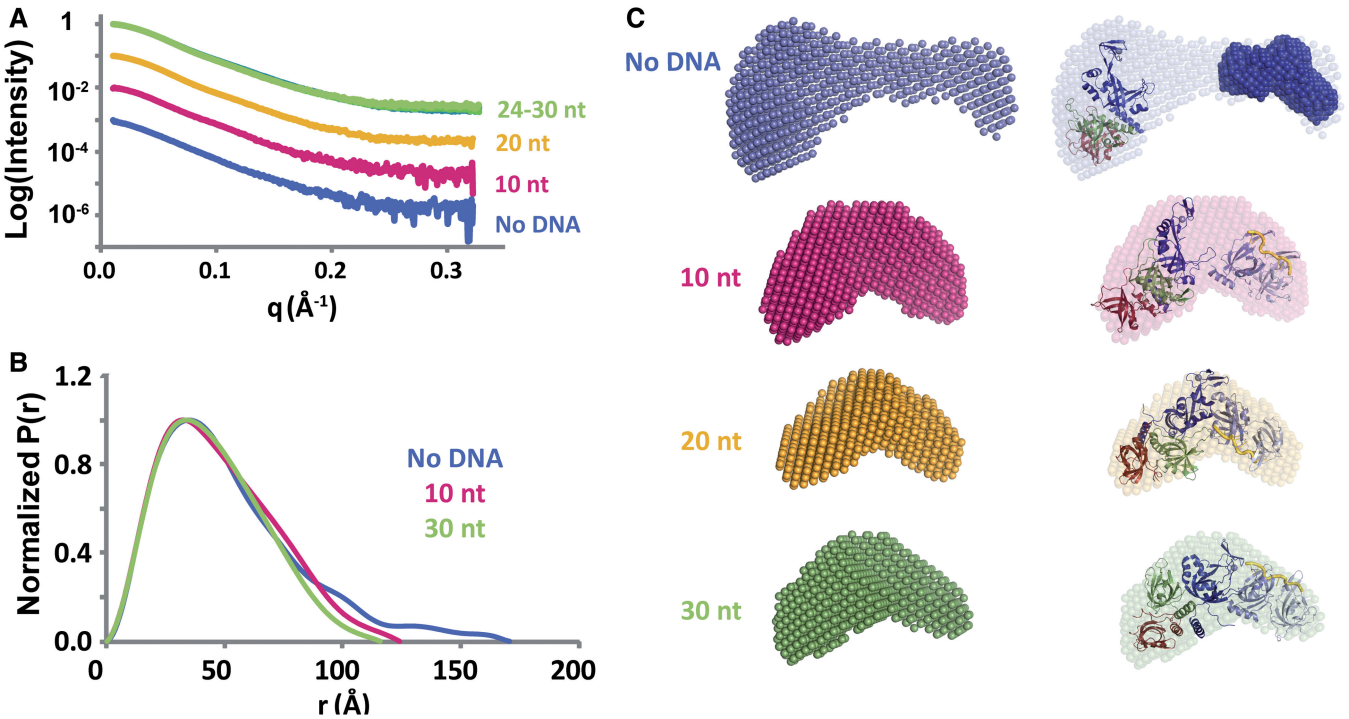


**Figure 4.** SAXS shows progressive compaction of RPA-DBC architecture induced by binding to progressively longer ssDNA substrates. (**A**) Experimental scattering profiles of RPA-DBC alone (blue) and bound to 10-nt (magenta), 20-nt (yellow) and 24–30-nt (green) substrates. The 24–30-nt complexes are indistinguishable. (**B**) P(r) distributions for RPA-DBC alone (blue) and bound to 10-nt (magenta) and 30-nt (green) substrates. (**C**) Averaged *ab initio* molecular envelopes of RPA-DBC and DNA-bound complexes from 15 independent runs of GASBOR. Panels at left display molecular envelopes with solid spheres. Panels at right display transparent molecular envelopes with 70AB/$dC_8$ (1JMC) and trimer core (1L1O) crystal structures docked (ribbon), which are coloured as 70A (light blue), 70B (blue), 70C (dark blue), 32D (green), 14 (red) and ssDNA (gold). The scattering envelope of DNA-free RPA-DBC shown at the top uses the 70AB scattering envelope (dark blue).

experimental data (Figure 5A). A broad overall range in $R_g$ was observed (33–46 Å), highly reminiscent of the simulations for the DNA-free protein (Figure 5B). The narrower 70A–70B distance distribution (23–28 Å) and the reduced oscillations in the 70A–70B distance range (~23–27 Å) reflect the spatial constraint between these domains imposed by the binding of DNA (Supplementary Figures S7 and S9). In contrast, inspection of the 70B-trimer core distance reveals substantial excursions (~30 Å) and low-frequency oscillations, which suggest that 70AB and the trimer core are flexibly tethered and move independently (Supplementary Figures S7 and S9). A shift in the distribution of 70B-trimer core distances relative to the DNA-free protein is observed. This reflects differences in non-specific interactions between the B–C linker and neighbouring domains observed in the simulation, rather than a direct effect of binding ssDNA (Supplementary Figure S7). Notably, binding of ssDNA to 70A and 70B did not impact the transient helical character of the B–C linker and uncoiling of the 70B C-terminal helix seen in the DNA-free protein (Figure 5C).

Overall, the comparison of the experimental data for the 10-nt complex to that predicted from the computational ensemble shows them to be in qualitative agreement to an extent similar to that for the DNA-free protein. However, there are differences. We focus on comparison of the $P(r)$ distribution, in particular at ~100 Å and beyond, as this provides the most striking change reflecting the compaction on binding of ssDNA. The independent movement of the RPA70AB module with respect to the trimer core gives rise to the observed shoulder at 70 Å in the P(r) distributions. The computed P(r) has a much more pronounced feature at 70 Å compared with the more compact experimental P(r) (Supplementary Figure S11). This difference suggests that the MD simulation did not fully capture the distribution of compact versus extended structures characteristic of the ensemble in solution. As RPA70AB remains bound to the ssDNA throughout the simulation, this 'extra' degree of compaction would seem to arise from some form of allosteric effect of binding DNA. However, substantial dynamic averaging as we observe for RPA-DBC can lead to an underestimation of the population of more extended architectures in a P(r) distribution (37,38). Hence, additional information is required to elucidate the detailed molecular basis for the changes induced in RPA-DBC in the initial mode of binding to ssDNA.

### A second transition in RPA-DBC on binding ≥20 nt

Scattering analysis of the complex with the 20-nt ssDNA substrate reveals further compaction of RPA-DBC (Table 1). A reduction of 1.3 Å in $R_g$ and 10 Å in $D_{max}$ relative to the 10-nt complex is observed, which we assign to the trimer core engaging the ssDNA. The P(r) function

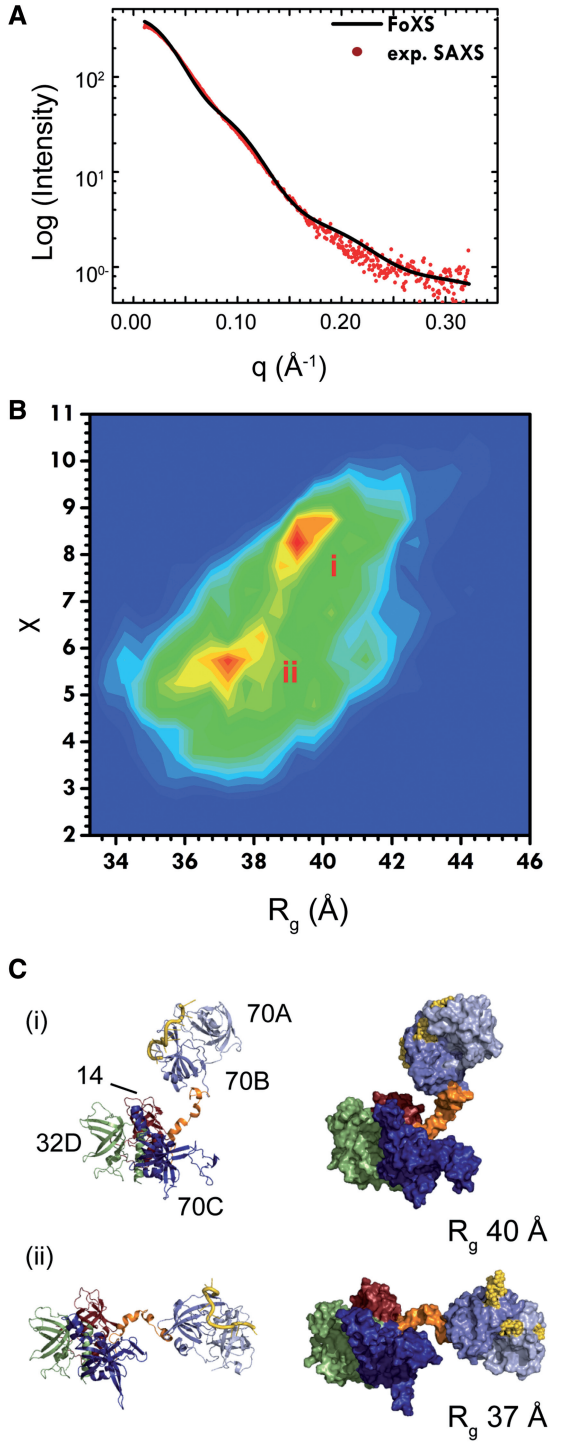


**Figure 5.** Atomic insights into restriction of motion and compaction within the 70AB module on binding to a 10-nt ssDNA substrate. (**A**) Overlay of the experimental scattering curve (red) with the averaged scattering curve (black) calculated from the 100 000 MD models ($\chi$ = 5.2). (**B**) 2D histogram comparing goodness-of-fit to the experimental data ($\chi$) versus radius-of-gyration ($R_g$) for each model. Roman numerals indicate the primary populations observed in the plot. (**C**) Representative models of RPA-DBC from each $R_g$ population in panel 5B displayed as described in Figure 3C. Colour-coding: 70A (light blue), 70B (blue), 70C (dark blue), 32D (green), 14 (red), B–C linker (orange) and ssDNA (gold).

for the 20-nt complex is similar to the 10-nt complex but is narrower and more symmetric, particularly at r values >50 Å (Figure 4B and Supplementary Figure S6). Moreover, the molecular envelope derived from the scattering data is well fit by 70AB-$dC_8$ and trimer core crystal structures (Figure 4C, right panel). These observations indicate RPA-DBC favours more compact architectures in the 20-nt complex than in the 10-nt complex (Supplementary Figure S3).

Surprisingly, no further changes were found in any of the scattering parameters for RPA-DBC in complex with 24-, 27- and 30-nt substrates relative to the 20-nt complex (Figure 4, Supplementary Figure S6 and Table 1). Examination of the *ab initio* molecular envelopes confirms that the compaction of RPA-DBC is the same for the 20-, 24-, 27- and 30-nt complexes. Overall, the primary architectural changes induced by binding ≥20-nt correlate to two major structural transitions: repositioning and alignment of 70A and 70B domains and closer positioning of the trimer core and 70AB.

The progressive compaction of RPA-DBC seems to be accompanied by a reduction in inter-domain flexibility. For example, the spatial volumes of the scattering envelopes become increasingly condensed with each structural transition and are well fit by the volumes of the high-resolution domain structures as soon as the trimer core is engaged (Figure 4C). This decrease in envelope volume is consistent with a reduction in inter-domain fluctuations (37). Direct extraction of the volume from experimental data supports our hypotheses of DNA-dependent compaction of the protein. SAXS data limited to the linear portion of the Porod-Debye region was used to estimate particle volumes for DBC alone (167 905 $Å^3$), DBC + 10 nt (149 693 $Å^3$), DBC + 20 nt (136 743 $Å^3$), DBC + 24 nt (128 189 $Å^3$) DBC + 27 nt (134 010 $Å^3$) and DBC + 30 nt (139 054 $Å^3$) (Table 1) (39). Although the addition of ssDNA should increase the volume, we saw significant DNA-induced decreases in volume as ssDNA is added to the 10- and 20-nt substrates, even when subtracting the increasing volumes for the DNA (Supplementary Figure S10). These correspond well to the two primary transitions in RPA-DBC architecture, consistent with a reduction in inter-domain fluctuations and compaction of RPA-DBC as the DNA-binding core engages ssDNA. The 24-, 27- and 30- complexes are similar to each other and to the 20-nt complex. The downwards and then upwards trends in these volumes are attributed to full engagement of RPA-DBC and subsequent redistribution of domains on the substrate.

## Models of the architecture and persistence of motion after the trimer core is engaged

To obtain insights into the molecular architectures associated with the longer DNA substrates, we performed MD simulations of RPA-DBC in complex with 20- and 30-nt. The 20-nt intermediate binding mode has been thought to involve ssDNA binding to just 70A, 70B and 70C, but manual modelling revealed it is physically possible to also engage 32D. Consequently, separate starting models representing the extremes of binding

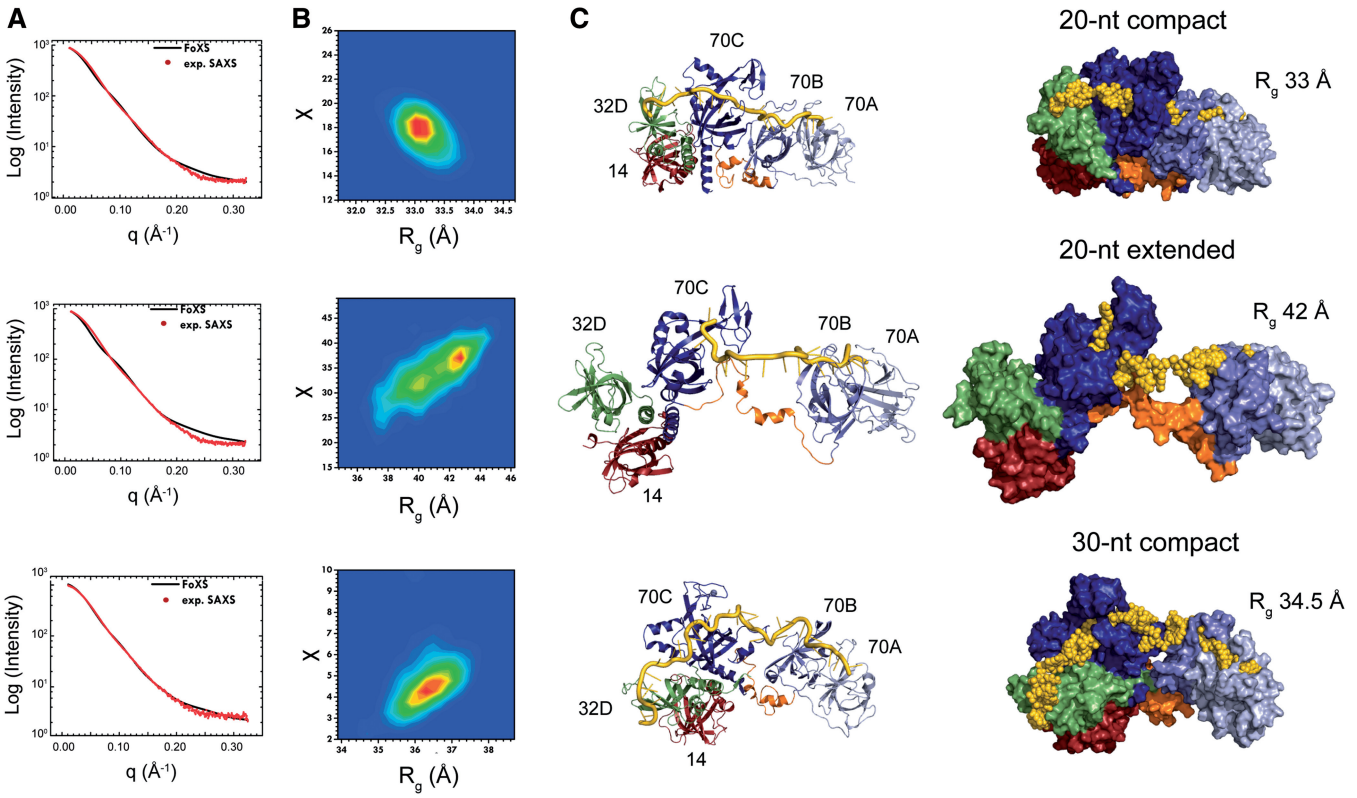


**Figure 6.** Atomic insights into the compaction of RPA-DBC on binding to 20-nt and 30-nt ssDNA substrates. (**A**) Overlay of the experimental scattering curve (red) with the averaged scattering curve (black) calculated from 100 000 MD models ($\chi$ = 20.7 for 20 nt compact; $\chi$ = 32.8 for 20 nt extended; $\chi$ = 4.6 for 30 nt). (**B**) 2D histogram comparing goodness-of-fit to the experimental data ($\chi$) versus radius-of-gyration ($R_g$) for each model. (**C**) Representative models from clustering analysis of the trajectories displayed as described in Figure 5C. Colour-coding: 70A (light blue), 70B (blue), 70C (dark blue), 32D (green), 14 (red), B–C linker (orange) and ssDNA (gold).

were created: the first with all four domains bound to the substrate ('compact') and a second engaging only A, B and C with 9 nt between 70B and 70C ('extended') (Supplementary Figure S8).

For the compact 20-nt simulation, the fit of the averaged scattering profile from 100 000 models to the experimental scattering data was not as good as observed for the DNA-free and 10-nt complex (Figure 6A, top). Inspection of the distribution of $R_g$ values indicates a highly uniform architecture that is more compact than that detected experimentally (compare ~33 Å from simulation versus ~35 Å from experiment) (Figure 6B, top). The narrow ranges of 70AB-trimer core distances, inter-domain angles and oscillations observed in the simulation indicate that the system is highly constrained (Supplementary Figures S8 and S9), which is inconsistent with the experimental data. These results strongly suggest that architectures in which 32D is engaged on 20-nt of DNA are not significantly populated.

The averaged scattering profile of the 100 000 models from the extended 20-nt simulation also did not fit as well to the experimental scattering data as was observed for the DNA-free and 10-nt binding simulations (Figure 6A, middle). Although the distribution of $R_g$ values, inter-domain distances and angles is reminiscent of that for the DNA-free protein and 10-nt complex (Figure 6B middle and Supplementary Figures S8 and S9), the most common architectures have $R_g$ values >40 Å, well in excess of the 35 Å measured experimentally. In fact, there is a greater population of extended architectures in this simulation than for the DNA-free and 10-nt complex (Figure 6C, middle). Examination of the MD trajectory reveals that this arises because fewer contacts are made between the B–C linker and the adjacent domains, allowing the unbound portions of the ssDNA and the B–C linker to sample a much wider range of conformations.

Overall, the simulations on the 20-nt complex support a model in which RPA engages both 70AB and the trimer core, although neither of the two limiting cases fit the data as well as was observed for the DNA-free or the 10-nt binding mode. As the assumptions inherent to all simulations are the same, this suggests that the two starting models and subsequent trajectories for the 20-nt complex were not representative of the complex in solution. Moreover, the binding of RPA-DBC to 20-nt of ssDNA seems not to correspond to the simple model in which domains A, B and C are engaged on the substrate.

The averaged scattering profile from 100 000 models of the 30-nt complex agrees with the scattering data to an extent that is comparable with the DNA-free protein and

10-nt complex. Remarkably, although the DNA is longer and the starting structure for the extended 20- and 30-nt complexes both incorporated nine bases between 70AB and the trimer core (Supplementary Figure S8), the agreement with the experimental data is significantly better for the 30-nt complex and follows the pattern of the 10-nt complex. An important role for transient non-specific interactions involving the B–C linker is observed (Figure 6C, bottom), although in this case, interactions with the ssDNA substrate are dominant. Overall, the 30-nt simulation reproduces the bilobal architecture of the *ab initio* envelope with a convex path for the DNA. A smaller magnitude and narrower distribution of $R_g$ values relative to the DNA-free protein and 10-nt complex is consistent with the trend observed in the experimental data (Figure 6A, bottom). The agreement between simulation and experiment is also reflected in parameters such as the shorter distances between 70B and the trimer core (Supplementary Figure S8). The relatively broad range of distances, angles and oscillations between 70AB and the trimer core indicates that some degree of flexibility along the B–C linker is retained (Supplementary Figures S8 and S9). Analysis of the simulations and experimental data suggests this likely corresponds to torsional motions (e.g. twisting of the domains) and not extensive excursions from the average architecture.

## DISCUSSION

Knowledge of how intact human RPA engages ssDNA substrates and transitions between its various functional states has remained inaccessible using conventional structural techniques. Such dynamic modular proteins are challenging to analyse but can be characterized by small angle scattering and computational methods (40–42). The combination of these methods enables interpretation of the low resolution SAXS data in terms of underlying structural models. The application of this approach to the binding of RPA to ssDNA provided information about (i) the transitions in the architectural ensembles of RPA as it binds ssDNA; (ii) hypotheses about the atomic interactions involved; and (iii) a first estimate of the 'path' followed by the ssDNA. Remodelling the architecture of modular multi-domain proteins is central to their assembly and function in the dynamic multi-protein machinery (41,43). The approach used here to characterize RPA architecture in different functional states should be broadly applicable to these systems.

Our studies show that the binding of RPA on ssDNA results in a progressive compaction of the protein. Because RPA is a modular protein with flexible linkers, its architecture is best described in terms of an ensemble of population-weighted architectural states, which might include low populations of architectures that diverge significantly from the mean. The DNA-free protein samples a large range of architectures because of the flexibility between the 70A and 70B domains and between the 70AB module and the trimer core (Figure 7A). The initial 8–10-nt interaction mode causes the 70A and 70B domains to compact significantly, as observed in previous

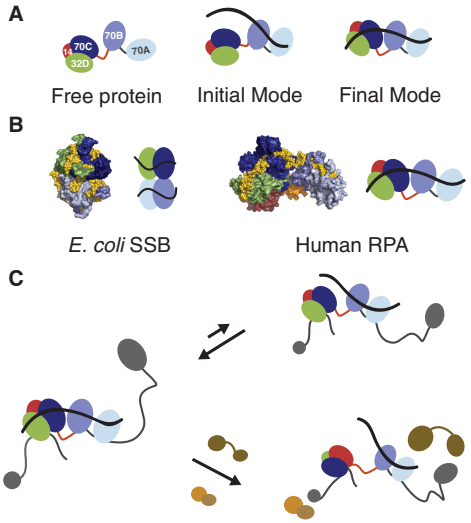


**Figure 7.** Models for modulation of RPA-ssDNA interactions. (**A**) Two major transitions in RPA architecture as it progressively binds ssDNA: the initial binding mode that compacts and aligns the 70A and 70B domains, the final binding mode that compacts the DNA-binding apparatus as it engages the trimer core. (**B**) Comparison of *E. coli* (left) and eukaryotic (right) SSB architectures. Surface representations (left) were generated from *E. coli* SSB coordinates 1EYG and the most populated MD cluster model for the 30-nt complex of human RPA-DBC. (**C**) Proposed mechanisms for displacement of RPA from ssDNA. The weaker affinity of the 32D domain could enable spontaneous dissociation of the trimer core from the DNA; however, the close proximity to 70AB is retained in the 10-nt binding mode, which promotes reassociation (top). The coupling of protein interactions at domains 70N, 32C and 70AB could shift the equilibrium from final to initial binding mode by sterically limiting reassociation of the trimer core and by increasing the distance between the trimer core and 70AB (bottom).

X-ray crystal structures and SAXS studies of RPA70AB. A second major compaction occurs when the trimer core is engaged in binding ssDNA. Even binding to as few as 20 nt of ssDNA is sufficient to cause this transition. Notably, no further significant changes occur as RPA-DBC binds up to 30 nt of ssDNA. Overall, our results show that RPA-DBC occupies an ensemble of architectures in solution that can be best viewed as an equilibrium between a range of extended, intermediate and compact states with binding to ssDNA binding driving the equilibrium towards more compact conformations.

The binding of RPA to ssDNA is more complex than that predicted by a simple model of consecutive engagement of the four known DNA-binding domains. SAXS data showed that RPA-DBC favours more compact as opposed to fully extended architectures. The simulations suggested that the linker between the 70B and 70C domains may be critical to the preference for more condensed states, as the B–C linker was found to form transient helical structure and participate in a range of transient interactions with domains 70B and 70C as well as the DNA. We observed that the 10-nt complex sampled a range of architectures that was more restricted relative to the DNA-free protein. Detailed analysis of the SAXS and simulation results suggested that the extent of compaction may not be solely explained by effects on RPA70AB. However, the challenges of interpreting

SAXS data for multi-domain proteins makes it difficult to discern if this observation is a by-product of significant inter-domain dynamics in RPA-DBC or if there is an allosteric effect of ssDNA binding on the B–C linker that inhibits sampling of fully extended RPA-DBC architectures. The SAXS data showed that binding of RPA-DBC to 20 nt of ssDNA engages the trimer core, but the simulations suggested that fixed binding to domains 70C and/or 70D is insufficient to explain the experimental data. And finally, we note the similarity observed in the SAXS data for the binding of RPA-DBC to 20, 24, 27 and 30 nt of ssDNA was not anticipated.

One key finding from our results is that the DNA-binding apparatus of RPA undergoes two architectural transitions as it binds to progressively longer stretches of ssDNA. This result is inconsistent with the prevailing view of three modes of RPA binding to ssDNA with distinct architectures (2). Moreover, it has been long held that the 30-nt binding mode is fully extended, and the 10-nt binding mode is compact (2,44). In fact, we find no evidence that the 10-nt binding mode is more compact than the 30-nt binding mode—our data show just the opposite. The previously reported compact architecture for the 10-nt binding mode was dependent on gluteraldehyde cross-linking of RPA in nucleoprotein complexes, whose structure was monitored indirectly by scanning transmission electron microscopy (44). Treatment with cross-linker may have caused RPA to be trapped into compact architectures that could only bind 8–10 nt.

In addition to correcting the model for the initial 10-nt and final 30-nt binding modes, our results show there is no distinct architecture associated with the intermediate 12–23-nt binding mode. The evidence in support of this intermediate binding mode included photo-cross-linking of RPA subunits to DNA substrates and testing the DNA-binding capability of domain-specific point mutants (4,5). Although these approaches establish the participation of a particular domain in ssDNA binding, they are indirect and do not reflect the structural organization of the molecule. Our direct physical analysis by SAXS is consistent with RPA having only two not three DNA-bound architectural states.

The architecture and binding of ssDNA by RPA are fundamentally different from that of the simpler and well-studied homotetrameric ssDNA-binding proteins (SSBs) found in prokaryotes (45). Crystal structures of *Escherichia coli* SSB reveal a compact back-to-back and base-to-base arrangement of the four OB-fold domains (46). As a result, SSB DNA-binding clefts face opposite directions relative to each other, compelling ssDNA substrates to encircle the SSB tetramer to occupy two (35-nt state) or all four (65-nt state) binding domains (Figure 7B). In contrast, flexible linking of RPA's DNA-binding domains allows each binding surface to orient in a similar direction and enables the core to organize in a convex manner around the ssDNA. This opens up the possibility for RPA to function in two ways, either like SSBs by inducing ssDNA to organize around the protein or conversely, by adapting its architecture in response to the DNA (Figure 7B). Our results support the latter. The greater versatility afforded by the flexible arrangement of OB-folds permits RPA to accommodate variations in the available length and binding context of DNA substrates across multiple DNA processing pathways. Such differences between prokaryotes and eukaryotes highlight the importance of modular structure for the more complex eukaryotic DNA processing machinery.

Investigations of the diffusion of bacterial SSBs along ssDNA suggest a mechanism for initiating displacement of SSBs from their ssDNA substrates (47,48). In this model, transient thermal dissociation of single DNA-binding domains combines with limited re-association to the ssDNA. For RPA, 32D possesses the weakest affinity of the four DNA-binding domains (49) and is the most plausible mediator of initial dissociation. Once 32D and 70C are disengaged, there can be an increase in the separation of the trimer core from 70AB (Figure 7C). It is conceivable that this promotes diffusion along the ssDNA and thereby enables access to the substrate. However, our observation that RPA-DBC does not significantly populate highly extended architectures suggests the relatively close proximity of 70AB and the trimer core likely promotes subsequent re-engagement of the trimer core onto the DNA (Figure 7C).

If the 32C and 70D domains are readily re-engaged, how is it possible for RPA to be displaced from ssDNA? We have previously reported that the interaction of RPA with other DNA processing proteins within the DNA processing machinery can promote or inhibit its binding to ssDNA [e.g. (50)]. In our model, these interactions shift the equilibrium between multiple relatively compact RPA architectural states, facilitating binding or release of ssDNA. Bacterial SSBs have been proposed to use interaction with their disordered C-terminal tails to promote release of ssDNA (51). The primary protein recruitment modules in RPA are 70N and 32C, which are independent structural modules. However, in most, if not all, cases, RPA binding partners interact via multiple contacts, also engaging the 70AB structural module. The interaction with 70AB alone could promote dissociation from the ssDNA either via direct competition for the DNA binding sites or allosteric effects from binding to the A–B linker (50,52,53). However, these interactions are invariably much weaker than 70AB interaction with ssDNA, and it is difficult to imagine they are sufficient to cause RPA to dissociate.

Our data suggest an alternative mechanism: the involvement of RPA70AB and one or both of the 70N and 32C protein interaction modules places steric or other constraints on RPA architecture that lead to *increased* separation of 70AB and the trimer core (Figure 7C). This increased separation would amplify intrinsic dissociation of the trimer core from the ssDNA, lower the overall affinity, and increase the probability of diffusion along the ssDNA and/or provide access to a competing DNA-binding domain from the binding partner (Figure 7C). The increasing recognition that RPA and other DNA processing proteins interact through multiple contact points [e.g. SV40 Tag helicase (50,53), polymerase α- primase (54,55), XPA (31,56), Rad52 (31,57)] suggests that this mechanism for protein-mediated displacement of

RPA from DNA could be conserved across DNA processing pathways. It is anticipated that this mechanism for obtaining access to DNA will be generally applicable, in light of the increasing recognition of the importance of architectural remodelling in DNA processing machinery (58–60).

A crystal structure of the RPA-DBC homologue from the fungus *Ustilago maydis* in complex with $dT_{32}$ and $dT_{62}$ in two space groups was published during the final revisions of this manuscript (61). A key feature in this structure is the collapsed quaternary structure driven by contacts between RPA70B, RPA70C, 10 of the B–C linker residues and the ssDNA intervening between the nucleotides bound in the B and C domains. Comparison with the SAXS curves of human RPA-DBC bound to 10–30 nt of ssDNA indicates the conformational ensemble in solution differs from the highly condensed crystal structure (Supplementary Figure S12). However, crystal-like compact architectures may be populated in solution. Notably, interfaces in the crystal packing are larger than the interfaces within the DBC in the structure. This observation suggests that this architecture is not likely to be highly populated in solution, which is consistent with our SAXS data.

Overall, the SAXS/MD and crystallographic studies provide highly complementary information. For example, both support condensation of the architecture of RPA on binding ssDNA and place the four binding sites for ssDNA along a curved trajectory on one face of the DBC. The crystal structure provides a DBC structure at near-atomic resolution. It also highlights interactions involving domains B and C, the B–C linker and the nucleotides between those bound in domains B and C, which are also found in the MD simulations. On the other hand, the SAXS/MD and crystal structure data provide two different models for RPA function: (i) in the crystal structure, keystone interactions centred on the B–C linker create a unique four-way interface that stabilizes a closed conformation for 30-nt bound RPA, whereas (ii) in the SAXS/MD data, a dynamic two-state condensation on ssDNA binding is defined with considerable flexibility. Importantly, the flexibility in our model naturally incorporates interactions with multiple RPA partners whose modular interfaces may stabilize a given functional architecture in concert with DNA binding. Our SAXS and computational results in combination with the crystal structures marks a new era in understanding architectural remodelling of RPA. Testable hypotheses can now be generated for the roles of the B–C linker, allosteric compaction of the DNA binding core promoted by the binding of DNA and architectural changes induced by partner protein binding, as well as a framework for understanding how RPA functions in so many DNA processing pathways.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table 1, Supplementary Figures 1–12, Supplementary Methods and Supplementary References [62–68].

## ACKNOWLEDGEMENTS

The pET15b plasmid for RPA-DBC was a kind gift of Alexey Bochkarev. The authors gratefully acknowledge Dr Miaw-Sheue Tsai of the EMB Core of the Structural Biology of DNA Repair (SBDR) program, who supervised construction of the two alternative versions of RPA-DBC not reported here, and Dr S. Michael Shell and Dr. Nicholas P. George for helpful discussion. The authors also thank Baokhanh N. Ho and Syiedah Korre for helping set up the systems and perform the MD simulations, and Dr Nikola Pavletich for providing the coordinates of the *Ustilago maydis* RPA-DBC crystal structures. Norie Sugitani collected ESI-MS data, and Dr W. Hayes McDonald performed the proteomics analysis.

## FUNDING

National Institutes of Health operating [R01 GM65484 to W.J.C.; R01 GM46312 to J.A.T.; P01 CA092584 to J.A.T. and W.J.C.]; National Science Foundation [NSF-CAREER MCB-1149521 to I.I.]; Georgia State University (to I.I.); NIH training T32 GM80320 (to C.A.B.); NIH centre grants [P30 ES00267 to the Vanderbilt Center in Molecular Toxicology and P30 CA068485 to the Vanderbilt Ingram Cancer Center]. Computational resources: NSF XSEDE program [CHE110042, in part]; National Energy Research Scientific Computing Center supported by the DOE Office of Science [DE-AC02-05CH11231, in part]. The X-ray scattering technology and applications to the determination of macromolecular shapes and conformations at the SIBYLS beamline (12.3.1) at the Advanced Light Source, Lawrence Berkeley National Laboratory: U.S. Department of Energy (DOE) program Integrated Diffraction Analysis Technologies (IDAT) (in part); NIH grant Macromolecular Insights on Nucleic acids Optimized by Scattering [R01GM105404, in part]. Support from the U.S. Department of Energy for the research at Oak Ridge National Laboratory was provided to the Center for Structural Molecular Biology [FWP ERKP291, Office of Biological and Environmental Research] and the High Flux Isotope Reactor [contract DE-AC05-00OR22725, Scientific User Facilities Division, Office of Basic Energy Sciences]. Funding for open access charge: National Institutes of Health [R01 GM65484].

*Conflict of interest statement*. None declared.

## REFERENCES

1. 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.
2. 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.

3. Brosey,C.A., Chagot,M.-E., Ehrhardt,M., Pretto,D.I., Weiner,B.E. and Chazin,W.J. (2009) NMR analysis of the architecture and functional remodeling of a modular multidomain protein, RPA. *J. Am. Chem. Soc.*, **131**, 6346–6347.
4. Lavrik,O.I., Kolpashchikov,D.M., Weisshart,K., Nasheuer,H.-P., Khodyreva,S.N. and Favre,A. (1999) RPA subunit arrangement near the 3′-end of the primer is modulated by the length of the template strand and cooperative protein interactions. *Nucleic Acids Res.*, **27**, 4235–4240.
5. Bastin-Shanower,S.A. and Brill,S.J. (2001) Functional analysis of the four DNA binding domains of replication protein A. The role of RPA2 in ssDNA binding. *J. Biol. Chem.*, **276**, 36446–36453.
6. Kim,C., Snyder,R.O. and Wold,M.S. (1992) Binding properties of replication protein A from human and yeast cells. *Mol. Cell. Biol.*, **12**, 3050–3059.
7. Kim,C., Paulus,B.F. and Wold,M.S. (1994) Interactions of human replication protein A with oligonucleotides. *Biochemistry*, **33**, 14197–14206.
8. Bochkarev,A., Pfuetzner,R.A., Edwards,A.M. and Frappier,L. (1997) Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA. *Nature*, **385**, 176–181.
9. Bochkareva,E., Belegu,V., Korolev,S. and Bochkarev,A. (2001) Structure of the major single-stranded DNA-binding domain of replication protein A suggests a dynamic mechanism for DNA binding. *EMBO J.*, **20**, 612–618.
10. Arunkumar,A.I., Stauffer,M.E., Bochkareva,E., Bochkarev,A. and Chazin,W.J. (2003) Independent and coordinated functions of replication protein A tandem high affinity single-stranded DNA binding domains. *J. Biol. Chem.*, **278**, 41077–41082.
11. Pretto,D.I., Tsutakawa,S., Brosey,C.A., Castillo,A., Chagot,M.-E., Smith,J.A., Tainer,J.A. and Chazin,W.J. (2010) Structural dynamics and single-stranded DNA binding activity of the three N-terminal domains of the large subunit of replication protein A from small angle X-ray scattering. *Biochemistry*, **49**, 2880–2889.
12. Bochkareva,E., Korolev,S. and Bochkarev,A. (2000) The role for zinc in replication protein A. *J. Biol. Chem.*, **275**, 27332–27338.
13. Bochkareva,E., Korolev,S., Lees-Miller,S.P. and Bochkarev,A. (2002) Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA. *EMBO J.*, **21**, 1855–1863.
14. Hura,G.L., Menon,A.L., Hammel,M., Rambo,R.P., Poole,F.L., Tsutakawa,S.E., Jenney,F.E., Classen,S., Frankel,K.A., Hopkins,R.C. *et al.* (2009) Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). *Nat. Methods*, **6**, 606–612.
15. Classen,S., Rodic,I., Holton,J., Hura,G.L., Hammel,M. and Tainer,J.A. (2010) Software for the high-throughput collection of SAXS data using an enhanced Blu-Ice/DCS control system. *J. Synchrotron Radiat.*, **17**, 774–781.
16. Lynn,G.W., Heller,W., Urban,V., Wignall,G.D., Weiss,K. and Myles,D.A.A. (2006) Bio-SANS—a dedicated facility for neutron structural biology at Oak Ridge National Laboratory. *Phys. B*, **385–386**, 880–882.
17. The PyMOL Molecular Graphics System, Version 1.4, Schrödinger, LLC.
18. Fiser,A., Do,R.K. and Sali,A. (2000) Modeling of loops in protein structures. *Prot. Sci.*, **9**, 1753–1773.
19. Pelikan,M., Hura,G.L. and Hammel,M. (2009) Structure and flexibility within proteins as identified through small angle X-ray scattering. *Gen. Physiol. Biophys.*, **28**, 174–189.
20. Martí-Renom,M.A., Stuart,A.C., Fiser,A., Sánchez,R., Melo,F. and Sali,A. (2000) Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.*, **29**, 291–325.
21. Case,D.A., Cheatham,T.E., Darden,T., Gohlke,H., Luo,R., Merz,K.M., Onufriev,A., Simmerling,C., Wang,B. and Woods,R.J. (2005) The Amber biomolecular simulation programs. *J. Comput. Chem.*, **26**, 1668–1688.
22. Onufriev,A., Bashford,D. and Case,D.A. (2000) Modification of the generalized Born model suitable for macromolecules. *J. Phys. Chem. B*, **104**, 3712–3720.
23. Onufriev,A., Bashford,D. and Case,D.A. (2004) Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins*, **55**, 383–394.
24. Tuckerman,M., Berne,B.J. and Martyna,G.J. (1992) Reversible multiple time scale molecular dynamics. *J. Chem. Phys.*, **97**, 1990.
25. Kalé,L., Skeel,R., Bhandarkar,M., Brunner,R., Gursoy,A., Krawetz,N., Phillips,J., Shinozaki,A., Varadarajan,K. and Schulten,K. (1999) NAMD2: greater scalability for parallel molecular dynamics. *J. Comput. Phys.*, **151**, 283–312.
26. Phillips,J.C., Braun,R., Wang,W., Gumbart,J., Tajkhorshid,E., Villa,E., Chipot,C., Skeel,R.D., Kalé,L. and Schulten,K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem.*, **26**, 1781–802.
27. Hornak,V., Abel,R., Okur,A., Strockbine,B., Roitberg,A. and Simmerling,C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins*, **65**, 712–725.
28. Schneidman-Duhovny,D., Hammel,M. and Sali,A. (2010) FoXS: a web server for rapid computation and fitting of SAXS profiles. *Nucleic Acids Res.*, **38**, W540–W544.
29. Shao,J., Tanner,S.W., Thompson,N. and Cheatham,T.E. (2007) Clustering molecular dynamics trajectories: 1. characterizing the performance of different clustering algorithms. *J. Chem. Theory Comput.*, **3**, 2312–2334.
30. Daughdrill,G.W., Ackerman,J., Isern,N.G., Botuyan,M.V., Arrowsmith,C.H., Wold,M.S. and Lowery,D.F. (2001) The weak interdomain coupling observed in the 70 kDa subunit of human replication protein A is unaffected by ssDNA binding. *Nucleic Acids Res.*, **29**, 3270–3276.
31. 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 A. *Cell*, **103**, 449–456.
32. Cai,L., Roginskaya,M., Qu,Y., Yang,Z., Xu,Y. and Zou,Y. (2007) Structural characterization of human RPA sequential binding to single-stranded DNA using ssDNA as a molecular ruler. *Biochemistry*, **46**, 8226–8233.
33. Brosey,C.A., Tsutakawa,S.E. and Chazin,W.J. (2012) Sample preparation methods to analyze DNA-induced structural changes in replication protein A. *Methods Mol. Biol.*, **922**, 101–122.
34. Putnam,C.D., Hammel,M., Hura,G.L. and Tainer,J.A. (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations, and assemblies in solution. *Q. Rev. Biophys.*, **40**, 191–285.
35. Heller,W.T. (2010) Small-angle neutron scattering and contrast variation: a powerful combination for studying biological structures. *J. Appl. Crystall.*, **D66**, 1213–1217.
36. Whitten,A.E. and Trewhella,J. (2009) Small-angle scattering and neutron contrast variation for studying bio-molecular complexes. *Methods Mol. Biol.*, **544**, 307–323.
37. Heller,W.T. (2005) Influence of multiple well defined conformations on small-angle scattering of proteins in solution. *Acta crystallogr. D*, **61**, 33–44.
38. Bernadó,P. (2010) Effect of interdomain dynamics on the structure determination of modular proteins by small-angle scattering. *Eur. Biophys. J.*, **39**, 769–780.
39. Rambo,R.P. and Tainer,J.A. (2011) Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers*, **95**, 559–571.
40. Rambo,R.P. and Tainer,J.A. (2010) Bridging the solution divide: comprehensive structural analyses of dynamic RNA, DNA, and protein assemblies by small-angle X-ray scattering. *Curr. Opin. Struct.Biol.*, **20**, 128–137.
41. Perry,J.J.P., Cotner-Gohara,E., Ellenberger,T. and Tainer,J.A. (2010) Structural dynamics in DNA damage signaling and repair. *Curr. Opin. Struct. Biol.*, **20**, 11.
42. Tsutakawa,S.E., Van Wynsberghe,A.W., Freudenthal,B.D., Weinacht,C.P., Gakhar,L., Washington,M.T., Zhuang,Z., Tainer,J.A. and Ivanov,I. (2011) Solution X-ray scattering combined with computational modeling reveals multiple conformations of covalently bound ubiquitin on PCNA. *Proc. Natl Acad. Sci. USA.*, **108**, 17672–17677.
43. Stauffer,M.E. and Chazin,W.J. (2004) Structural mechanisms of DNA replication, repair, and recombination. *J. Biol. Chem.*, **279**, 30915–30918.

44. Blackwell,L., Borowiec,J. and Masrangelo,I. (1996) Single-stranded-DNA binding alters human replication protein A structure and facilitates interaction with DNA-dependent protein kinase. *Mol. Cell. Biol.*, **16**, 4798–4807.
45. Richard,D.J., Bolderson,E. and Khanna,K.K. (2009) Multiple human single-stranded DNA binding proteins function in genome maintenance: structural, biochemical, and functional analysis. *Crit. Rev. Biochem. Mol. Biol.*, **44**, 98–116.
46. Raghunathan,S., Kozlov,A.G., Lohman,T.M. and Waksman,G. (2000) Structure of the DNA binding domain of E. coli SSB bound to ssDNA. *Nature Struct. Biol.*, **7**, 648–652.
47. Roy,R., Kozlov,A.G., Lohman,T.M. and Ha,T. (2009) SSB protein diffusion on single-stranded DNA stimulates RecA filament formation. *Nature*, **461**, 1092–1097.
48. Zhou,R., Kozlov,A.G., Roy,R., Zhang,J., Korolev,S., Lohman,T.M. and Ha,T. (2011) SSB functions as a sliding platform that migrates on DNA via reptation. *Cell*, **146**, 222–232.
49. Sibenaller,Z.A., Sorensen,B.R. and Wold,M.S. (1998) The 32- and 14-kilodalton subunits of replication protein A are responsible for species-specific interactions with single-stranded DNA. *Biochemistry*, **37**, 12496–12506.
50. Jiang,X., Klimovich,V., Arunkumar,A.I., Hysinger,E.B., Wang,Y., Ott,R.D., Guler,G.D., Weiner,B., Chazin,W.J. and Fanning,E. (2006) Structural mechanism of RPA loading on DNA during activation of a simple pre-replication complex. *EMBO J.*, **25**, 5516–5526.
51. Shereda,R.D., Kozlov,A.G., Lohman,T.M., Cox,M.M. and Keck,J.L. (2008) SSB as an organizer/mobilizer of genome maintenance complexes. *Crit. Rev. Biochem. Mol. Bio.*, **43**, 289–318.
52. Stauffer,M.E. and Chazin,W.J. (2004) Physical interaction between replication protein A and Rad51 promotes exchange on single-stranded DNA. *J. Biol. Chem.*, **279**, 25638–25645.
53. Arunkumar,A.I., Klimovich,V., Jiang,X., Ott,R.D., Mizoue,L., Fanning,E. and Chazin,W.J. (2005) Insights into hRPA32 C-terminal domain–mediated assembly of the simian virus 40 replisome. *Nature Struct. Mol. Biol.*, **12**, 332–339.
54. Vaithiyalingam,S., Warren,E.M., Eichman,B.F. and Chazin,W.J. (2010) Insights into eukaryotic DNA priming from the structure and functional interactions of the 4Fe-4S cluster domain of human DNA primase. *Proc. Nat Acad. Sci. USA.*, **107**, 13684–9.
55. Dornreiter,I., Erdile,L.F., Gilbert,I.U., von Winkler,D., Kelly,T.J. and Fanning,E. (1992) Interaction of DNA polymerase alpha-primase with cellular replication protein A and SV40 T antigen. *EMBO J.*, **11**, 769–776.
56. Daughdrill,G.W., Buchko,G.W., Botuyan,M.V., Arrowsmith,C.H., Wold,M.S. and Lowry,D.F. (2003) Chemical shift changes provide evidence for overlapping single-stranded DNA- and XPA-binding sites on the 70 kDa subunit of human replication protein A. *Nucleic Acids Res.*, **31**, 4176–4183.
57. Jackson,D., Dhar,K., Wahl,J.K., Wold,M.S. and Borgstahl,G.E. (2002) Analysis of the human replication protein A: Rad52 Complex: Evidence for Crosstalk Between RPA32, RPA70, Rad52 and DNA. *J. Mol. Biol.*, **321**, 133–148.
58. Sarker,A.H., Tsutakawa,S.E., Kostek,S., Ng,C., Shin,D.S., Peris,M., Campeau,E., Tainer,J.A., Nogales,E. and Cooper,P.K. (2005) Recognition of RNA polymerase II and transcription bubbles by XPG, CSB, and TFIIH: insights for transcription-coupled repair and Cockayne Syndrome. *Mol. Cell*, **20**, 187–198.
59. Williams,R.S., Dodson,G.E., Limbo,O., Yamada,Y., Williams,J.S., Guenther,G., Classen,S., Glover,J.N.M., Iwasaki,H., Russell,P. *et al.* (2009) Nbs1 flexibly tethers Ctp1 and Mre11-Rad50 to coordinate DNA double-strand break processing and repair. *Cell*, **139**, 87–99.
60. Tubbs,J.L., Latypov,V., Kanugula,S., Butt,A., Melikishvili,M., Kraehenbuehl,R., Fleck,O., Marriott,A., Watson,A.J., Verbeek,B. *et al.* (2009) Flipping of alkylated DNA damage bridges base and nucleotide excision repair. *Nature*, **459**, 808–813.
61. 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.
62. Konarev,P.V., Volkov,V.V., Sokolova,A.V., Koch,M.H.J. and Svergun,D.I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystall.*, **36**, 1277–1282.
63. Petoukhov,M.V., Konarev,P.V., Kikhney,A.G. and Svergun,D.I. (2007) ATSAS 2.1–towards automated and web-supported small-angle scattering data analysis. *J. Appl. Crystall.*, **40**, s223–s228.
64. Dmitri,I.S., Petoukhov,M.V. and Koch,M.H.J. (2001) Determination of domain structure of proteins from x-ray solution scattering. *Biophys J.*, **80**, 2946–2953.
65. Franke,D. and Svergun,D.I. (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Crystall.*, **42**, 342–346.
66. Birmanns,S., Rusu,M. and Wriggers,W. (2011) Using Sculptor and Situs for simultaneous assembly of atomic components into low-resolution shapes. *J. Struct. Biol.*, **173**, 428–435.
67. Wignall,G.D. and Bates,F.S. (1987) Absolute calibration of small-angle neutron scattering data. *J. Appl. Crystall.*, **20**, 28–40.
68. Tjioe,E. and Heller,W.T. (2007) ORNL_SAS: software for calculation of small-angle scattering intensities of proteins and protein complexes. *J. Appl. Crystall.*, **40**, 782–785.

## DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.