7892 *Biochemistry* **2009,** *48,* 7892–7905
DOI: 10.1021/bi801896s

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Prediction of and Experimental Support for the Three-Dimensional Structure of Replication Protein A[†]

Jonathan Eric Nuss,[‡] Deacon John Sweeney, and Gerald Michael Alter*

Department of Biochemistry and Molecular Biology and Biomedical Sciences Ph.D. Program, Wright State University, Dayton, Ohio 45435 [‡] Present address: Department of Bacteriology, Army Medical Research Institute of Infectious Diseases, Frederick, MD 21702-5011.

Received October 8, 2008; Revised Manuscript Received July 18, 2009

ABSTRACT: Replication protein A (RPA) is a heterotrimeric, multidomain, single-stranded DNA binding protein that is essential for DNA replication, repair, and recombination. Crystallographic and NMR studies on RPA protein fragments have provided structures for all domains; however, intact heterotrimeric RPA has resisted crystallization, and a complete protein structure has not yet been described. In this study, computational methods and experimental reactivity information (MRAN) were used to model the complete structure of RPA. To accomplish this, models of RPA's globular domains and its domain-linking regions were docked in various orders. We also determined rates of proteolytic cleavage and amino acid side chain chemical modifications in native, solution state RPA. These experimental data were used to select alternate modeling intermediates and final structural models, leading to a single model most consistent with our results. Using molecular dynamics simulations and multiple rounds of simulated annealing, we then relaxed this structural model and examined its flexibility. The family of resultant models is consistent with other, previously published, critical lines of evidence and with experimental reactivity data presented herein.

Replication protein A (RPA)[1] is a single-stranded DNA (ssDNA) binding protein that plays a critical role in DNA replication, repair, and recombination as well as in cell cycle control (*1*−*3*). In each of these processes, RPA interacts with specific proteins, forming different multimolecular complexes with DNA. For example, in addition to binding directly with replication and repair proteins, RPA interacts with a host of regulatory proteins, including cell cycle control proteins (*4*), cyclins (*5*), and p53 (*6*). RPA is also phosphorylated in a cell cycle-dependent fashion and hyperphosphorylated in response to DNA damage (*7*). Recent studies show that hyperphosphorylated RPA displays altered DNA binding affinity and a weakened ability to associate with several protein binding partners (*7*−*9*). Thus, not only is RPA a central player in many aspects of cellular oligonucleotide metabolism, but apparently it is also a major regulatory protein.

RPA is also a heterotrimeric protein, composed of subunits with apparent molecular masses of 70, 32, and 14 kDa (*1*), and has been the target of many structural studies. Proteolytic digestion (*10*), sequence alignment data (*11*), and three-dimensional (3D) structural analysis indicate that its three subunits all contain oligonucleotide binding (OB) fold domains. These characteristic, clamplike, OB folds include five-stranded β-barrel structures often involved in binding anionic polymers such as ssDNA (*12*).

Though native RPA has resisted X-ray crystallographic analysis, fragments containing each domain have been examined by either NMR or X-ray crystallography. Early crystallographic studies on the central portion of RPA 70 (*13*) revealed two OB folds connected by a short linker region. These were later shown by X-ray crystallography to change their juxtaposition upon binding ssDNA (*13*, *14*). The structures of the N-terminal domain of RPA 70 (*15*, *16*) and of the C-terminal region in the 32 kDa subunit (*17*) were determined by NMR and shown to contain an OB fold and a winged helix−loop−helix domain, respectively. The structures of RPA 14, bound to the central domain of RPA 32 (*18*), and of these two smaller subunits complexed with the C-terminal domain of RPA 70 (*19*) have recently been reported. This trimeric structure contains an interface involving three OB folds, one from each subunit, providing the minimum trimerization core of RPA. Herein, we employed MRAN (modification reactivity analysis) methodologies (*20*), to extend this structural information and to propose a 3D model for the complete heterotrimeric RPA protein.

[†]This study was supported in part by the Biomedical Sciences Ph.D. Program of Wright State University (fellowship for J.E.N. and D.J.S.) and the Department of Defense, Contract DAMD17-00-C-0020.

*To whom correspondence should be addressed: Department of Biochemistry and Molecular Biology, 162 Diggs Laboratory, 3640 Colonel Glenn Hwy., Dayton, OH 45435-0001. E-mail: gerald.alter@wright.edu. Phone: (937) 775-2504. Fax: (937) 775-3485.

[1]Abbreviations: RPA, replication protein A; RPA 14, RPA 32, and RPA 70, 14, 32, and 70 kDa subunits of RPA, respectively; domain A, 70 kDa subunit, residues 181−300; domain B, 70 kDa subunit, residues 301−420; domain C, 70 kDa subunit, residues 439−616, PDB entry 1L1Oc; domain D, 32 kDa subunit, residues 44−171, PDB entry 1L1Ob; domain E, 14 kDa subunit, PDB entry 1L1Oa; domain F, 70 kDa subunit, residues 1−114, PDB entry 1EWI; DPU, 32 kDa subunit, residues 204−270 (C-terminal domain), PDB entry 1DPU; NTM, alternate model for domain F structure (*1*); TMC, trimerization core (domains C−E), PDB entry 1L1O; 32L, 32 kDa subunit, residues 172−203; 32N, 32 kDa subunit, residues 1−43; 70L1, 70 kDa subunit, residues 115−182; 70L2, 70 kDa subunit, residues 421−438; FF, empirical force field; MD, molecular dynamics; MRAN, modification reactivity analysis; MS, mass spectroscopy; rmsd, root-mean-square deviation; SA, simulated annealing.

 Published on Web 07/21/2009

MATERIALS AND METHODS

Chemicals. Immobilized trypsin and V8 protease and sulfo-*N*-hydroxysuccinimide acetate (sulfo-NHS-acetate) were purchased from Pierce (Rockford, IL). Soluble sequencing-grade proteases (trypsin, chymotrypsin, and V8 protease) were purchased from Roche (Indianapolis, IN). α-Cyano-4-hydroxycinnaminic acid (CHCA), sinnapinic acid (SPA), and all other chemicals were purchased from Sigma (St. Louis, MO). Water was deionized using a Milipore (Bedford, MA) Mili Q Plus water purification system.

RPA Purification. Plasmid p11d-tRPA, which contains DNA corresponding to all RPA subunits, was obtained as a generous gift from M. Wold (Department of Biochemistry, University of Iowa Carver College of Medicine, Iowa City, IA). It was used to transform the BL21(DE3) strain of *Escherichia coli*. The transformed cells were then induced to express human recombinant RPA, and the protein was purified with a slight modification of a previously described procedure (*21*). To ensure bound zinc was retained in the protein, EDTA was omitted from all buffers, and a zinc-containing buffer (termed Zn buffer) [20 mM Tris (pH 7.8), 10 mM DTT, 0.01% IGEPAL CA 630 (octylphenoxylpolyethoxyethanol), and 10 μM $ZnCl_2$] was used throughout the purification (*22*).

Proteolytic and Chemical Modification Time Courses. Chemical modification time courses, limited proteolysis, and exhaustive protein digestions were performed as previously described (*20*). They are briefly described below.

(*i*) *Limited proteolysis* was performed on aliquots of purified RPA ([RPA] ≈ 0.5 mg/mL) by using agarose-gel-immobilized trypsin or *Staphylococcus aureus* V8 protease, in Zn buffer, at pH 7.8. Routinely, ~100 mg of the protease gel slurry was added to 250 μL of purified RPA. When V8 protease was used, conditions were chosen so that cleavage occurred adjacent to glutamic acid residues only. Reactions were stopped at different times (2, 5, 10, 20, 40, and 80 min) by centrifuging reaction mixtures at low speed and decanting the supernatant. Hydrolysate (1 μL) corresponding to different time points was saved for mass spectrometric analysis.

(*ii*) *Sulfo-NHS-Acetate Modification*. RPA (10 μM) was dialyzed versus 20 mM HEPES buffer (pH 8.0), and then a sulfo-NHS-acetate stock solution (100 mM) was added to achieve a final concentration of 1 mM. Aliquots were withdrawn from the reaction mixture at various times (2, 5, 10, 20, 40, and 80 min), and the reaction was quenched by addition of an excess of glycine (10 mM). Samples were then exhaustively digested with an appropriate protease and the digests analyzed by MS. For these experiments, trypsin, V8 protease, and chymotrypsin were used to achieve maximum sequence coverage.

(*iii*) *Modification by Hydroxyl Radical*. RPA was dialyzed versus 20 mM HEPES (pH 7.8) containing 10 μM $ZnCl_2$ in preparation for oxidative labeling. Hydroxyl radicals were generated in situ by γ radiolysis of water. Dialyzed samples (~0.37 mg/mL), in plastic microcentrifuge tubes, were placed in front of a cesium-137 source (0.5 Gy/s). At various times, aliquots were removed and the extent of RPA hydroxylation was determined by MS analysis, after exhaustive digestion with trypsin.

(*iv*) *Exhaustive Proteolysis*. After modification, and prior to exhaustive digestion, RPA was denatured by addition of acetonitrile to a final concentration of 20% and incubation at 37 °C for 30 min. The samples were then diluted 1:1 with 25 mM ammonium bicarbonate (pH 8.0) to decrease the acetonitrile concentration to 10%. Sequencing-grade trypsin, V8 protease, or chymotrypsin was added at a protease/RPA ratio of 1/40 (w/w), followed by overnight incubation at 37 °C.

Mass Spectrometry. Peptides generated by limited proteolysis, or from exhaustive digestions, were analyzed by MALDI-TOF using a Ciphergen (Fremont, CA) Protein Chip System, model PBS II. Mass spectra were recorded after 1 μL of a protein digest (typically, 0.25−0.50 mg/mL before digestion) had been placed on a protein chip spot and 0.5 μL of matrix solution had been added. Laser desorption protocols were optimized to analyze peptides with masses of < 10 kDa and for peptides with masses of > 10 kDa (laser intensities equal to 160 and 273, respectively). At least 100 individual spectra were averaged in all instances, and peak masses were determined with the centroid tool included in the Protein Chip analysis package (Ciphergen User's Guide, Vermillion Inc., Fremont, CA). It should be noted that peptides with masses of < 1 kDa were difficult to observe because of interference effects caused by matrix ions.

Matrix solutions were prepared by dissolving the energy-absorbing material in a solvent consisting of 50% acetonitrile, 0.1% trifluoroacetic acid, and water. A saturated solution of sinnapinic acid (SPA) was used when peptides with masses of > 10 kDa were analyzed, while 33 mM α-cyano-4-hydroxycinnaminic acid (CHCA) was used as a matrix solution for peptides with masses of < 10 kDa. Scaling was performed using RPA 14 as a standard for spectra in the $m/z \geq 10000$ range, while angiotensin II and ubiquitin standards were added to samples when spectra were collected in the $m/z < 5000$ range. Peaks in MS spectra were routinely associated with proteolytic fragments on the basis of their mass. However, when the assignment was ambiguous, the peptides in question were isolated and submitted to secondary exhaustive digestion (*23*).

(*i*) *Determination of Proteolytic Reaction Rate Constants*. Rate constants for limited proteolysis reactions were determined as previously described (*20*). Generally, MS peak amplitudes, corresponding to peptides of interest, were normalized to appropriate internal standards and these intensities plotted versus reaction time. Reaction time courses were analyzed using kinetic schemes involving either a single first-order process or multiple linked first-order reactions. Cross determinations of rate constants frequently occurred since more than a single peptide could result from any particular proteolytic reaction site. In these instances, an average rate constant was reported together with an estimate of error.

(*ii*) *Determination of Chemical Reaction Rate Constants*. Rate constants for covalent modification, by both hydroxyl radicals and sulfo-NHS-acetate, were determined as described in the Supporting Information and in ref *20*. Briefly, samples for each time point were exhaustively digested and the products analyzed by MALDI-TOF mass spectrometry. Novel MS peaks were used to identify reactive peptides that had characteristic changes in mass, increments of +16 for hydroxylation and +42 for acetylation when V8 protease was used for postmodification exhaustive digestion. Rate constants for each modification were determined by fitting the time course of the fractional decrease in the unreacted peptide to a first-order kinetic model (*20*).

Molecular Dynamics Simulations. The AMBER software suite (Regents of the University of California) was used for molecular dynamics (MD) simulations (*24*, *25*), as described previously (*26*). Protein structures were typically energy-minimized by using a combination of steepest descent and conjugate

gradient steps. The resulting structures were encased in a water box which was 10 Å larger on each edge than the centered protein molecule. The protein and water were "heated" in successive steps until 300 K was reached. Each step included an increase in temperature of 50 °C and a molecular dynamic simulation to distribute energy about the system (normally 10000 1.5 fs steps). We then performed a 100 ps production MD simulation. Individual structures representing 1 ps steps were superimposed using a root-mean-square deviation (rmsd) fitting routine. To identify potentially mobile portions in the structural model, we calculated rmsds for each α carbon and then decorated PDB models according to rmsds. Amino acid α carbons with standard deviations of > 3 Å were colored magenta, 2.5−3.0 Å red, 2.0−2.5 Å orange, 1.5−2.0 Å yellow, 1.0−1.5 Å green, 0.5−1.0 Å cyan, and < 0.5 Å blue.

Simulated Annealing. Using the AMBER software suite, our most appropriate full-length RPA model (from the MD analysis) was subjected to multiple rounds of simulated annealing (SA). In this process, unhydrated structures were warmed in 50 K steps [500 calculations per step (500 × 1.5 fs)] to 600 K. The structures were then cooled in 100 K steps to 300 K. At each cooling step, 2000 MD steps (1.5 fs × 2000) were performed to allow the system energy to distribute. The system coupling (tautp) was decreased in successive steps to ensure a gradual distribution of energy throughout the structure. Finally, the degree of coupling was reduced over 20000 MD steps (1.5 fs × 20000), resulting in structures, near the end of this process, that were constant. Coordinates of the corresponding structure(s) were then captured. For several structures, the resultant models were subjected to additional energy minimization using 2000 steps of steepest descent followed by 3000 steps using the conjugate gradient method. Explicit waters were not included in these calculations. For each round of SA, a different specific structure of model 4 was used as a starting structure. These structures included the model at points in the MD experiment just mentioned, and at points in the warming of the model in preparation for MD.

Scoring Models. Model scoring has been described previously (*20*) and is based on the expectation that exposed potential targets for modification are reactive, those that are buried are not, and partially exposed targets will have limited reactivity (*10*, *27*). Briefly, probed residues were grouped into four reactivity classes according to the magnitude of their modification rate constants. For each reagent, these were designated 3, 2, and 1 for reactive residues with decreasing rate constant values and −3 for unreactive residues. Within each reactivity class, rate constants usually varied ~4-fold. The reactivity number for each residue that was probed is given in Table 1.

To determine solvent accessibility, the relative surface area of each reactivity-classified amino acid in each model was calculated using the GET AREA server (http://www.scsb.utmb.edu/cgi-bin/get_a_form.tcl). Amino acid residues were then classified as accessible if the ratio was at least 30% and buried if the ratio was less than 20%. If a residue's side chain ratio was between 20 and 30%, it was considered to be partially accessible (*28*). Probed residues were then assigned an exposure score of 1 or −1 if they were accessible or buried, respectively. For sulfo-NHS modifications, only the accessibility of ε-amino groups was considered.

Scoring involved assigning a structure score to each probed residue. This score was the product of the reactivity class number (3, 2, 1, or −3) and an exposure score for each amino acid classified either as fully exposed or buried (1 or −1, respectively). Partially exposed amino acids were assigned structural scores of −1 if their reactivity scores were 3 or −3, 0 if their reactivity scores were 2, and 1 if their reactivity scores were 1. Thus, reactive amino acids had more positive structure scores when their exposure in a particular model matched anticipation, based on the residue's reactivity classification. Structure scores for RPA models were then calculated by summing the individual structure scores of all probed residues in the model.

Table 1: Relative Rate Constants for Modification of RPA

	subunit	residue		k^a	VF[b]	R^c	no. det.[d]	std. dev.[e]	rx. class[f]
trypsin	70	111	K	3.0	X	0.95	3	3.76	3
		157	K	0.57		0.98	1		2
		167	K	0.55		0.98	1		2
		431	K	4.0	X	0.95	2	4.87	3
	32	37	K	0.47		0.99	1		2
		38	K	0.48		0.99	3	0.35	2
		40	R	0.07		0.97	2	0.03	2
		42	R	0.44		0.99	2	0.08	2
		180	R	0.91	X	0.93	5	0.61	3
		217	K	0.68		0.96	3	0.32	3
		231	K	0.31		0.86	1		2
		243	K	0.15		0.85	1		2
		265	K	0.15		0.85	1		2
V8	70	7	E	0.16		0.91	1		2
		120	E	0.15		0.94	2	0.01	2
		418	E	0.25		0.98	1		3
		584	E	0.02		0.91	1		1
	32	7	E	0.49		0.99	1		3
		36	E	0.38		0.99	1		3
		190	E	0.92	X	0.96	3	1.18	3
		223	E	0.50	X	0.89	3	0.14	3
		252	E	0.08		0.99	1		2
NHS	70	103	K	0.22		0.78			2
		111	K	0.31		0.78			3
		157	K	0.34		0.83			3
		431	K	0.07		0.94			2
	32	37	K	0.45		0.87			3
		171	K	0.33		0.88			3
		217	K	0.13		0.94			2
		231	K	0.12		0.97			2
		243	K	0.07		0.94			2
		265	K	0.24		0.72			2
OH radical	70	118	Y	0.81		0.76			3
		151	M	0.81		0.76			3
	32	188	M	1		0.88			3
		194	F	1		0.88			3
		199	F	1		0.88			3
		200	M	1		0.88			3
		219	C	0					−3
		227	F	0					−3
		248	F	0.2		0.86			1
		256	Y	0.2		0.86			1
		264	F	0.2		0.86			1

[a] For each reagent, rate constants were normalized. Maximum values range from 1 to 4. [b] VF indicates the value of the faster of the pair of rate constants determined using a linked kinetic scheme that is larger than what can be reliably determined by our analysis (see Materials and Methods). A value 5 times that of the slower rate constant was assigned. This may underestimate the true value of the faster rate constant. [c] *R* parameter describing the quality of the respective fit. [d] Number of nonreplicate determinations of the rate constant values used to estimate the value of the relative rate constant in the *k* column of this table. [e] Standard deviation associated with the average value of the relative rate constant. [f] Value of reactivity score given to a model site based on the observed rate constant, *k*.

RESULTS

Rates of Covalent Modification. Rate constants summarizing the kinetics of acetylation of lysine side chains with

sulfo-NHS-acetate and the hydroxylation of sulfur-containing, as well as aromatic amino acids, for the nondomain regions of RPA are listed in Table 1. Similarly, rate constants describing proteolytic cleavage of RPA on the carboxyl side of lysine and arginine (trypsin) as well as glutamic acid (V8 protease) in nondomain regions are also summarized in Table 1. For the sake of clarity, these results are also combined with previously reported reactivity results for amino acids in domain regions of RPA (*20*) and are presented as a complete data set in Table 2 of the Supporting Information. Controls assessing the intactness of the protein's three-dimensional structure during proteolytic time courses are described and shown in the Supporting Information. They indicate that the protein remains substantially intact during limited proteolytic time courses.

Construction of RPA Structural Models. Complete structure models of RPA were assembled (predicted) using a strategy of successive pairwise dockings involving molecular models of both domain structures and nondomain regions. The relevance of reported domain structural models (*12*−*18*) to the native protein's conformation has been confirmed by the criterion of MRAN (*20*). Structural models for the four large nondomain regions of RPA (>20 amino acids) were predicted using the nontemplate method implemented on the HMMSTR-Rosetta server (*29*). High-scoring models were screened using MRAN, i.e., by comparing predicted reactivity based on a specific model with observed reactivity. A brief discussion of the process is presented as Supporting Information together with a summary of results (Figure 11 and Table 3).

(*i*) *Docking RPA Domain and Nondomain Sequences*. We assembled complete RPA models using two different core structures. The use of core structures followed from the observation that proteolysis of RPA resulted in a stable structure containing domains from all subunits (*19*). We assembled our first core (Figure 1, PC) by starting with a model for the complete 32 kDa subunit and docking models of the 14 kDa subunit as well as the C domain to it. On the basis of the core model, we constructed three different structural models for RPA by using alternate 3D models for the A, B, and F domains to complete the RPA structure (Figure 1, 1,2,3). Our second core (TMC) was used to construct five additional models of the complete RPA structure. This core contained only domains C, D, and E and had the reported structure determined by X-ray crystallographic analysis of a RPA fragment (PDB entry 1L10) (Figure 1, TMC). The five different structural models resulted from alternate docking orientations of structural segments at various steps in the process being approximately equally favorable with respect to energy and to exposure of the residues we found most reactive. Specifically, additions of DPU and its associated linker region to the core's 32 kDa subunit produced three different principle intermediate structures (Figure 1, I, II, and III). Two of these led to complete models (Figure 1, 4 and 5). Alternate ways of adding the A and B domains with their associated linker to the third intermediate structure produced two different principle intermediate structures (Figure 1, row 3). One of these led directly to a completed RPA structure (Figure 1, 6). Finally, addition of the F domain to the second intermediate resulted in two different final structures (Figure 1, 7 and 8).

A more complete description of the assembly of RPA structures is provided as Supporting Information.

(*ii*) *Model Characteristics*. Models resulting from the assembly process described above are shown in Figures 12 and 14 of the Supporting Information. All are characterized by having a compact core, including the C, D, and E domains and a looser association of the A, B, and perhaps F domains (Figure 14 of the Supporting Information). The A and B domains, as represented by the 1FGU (PDB entry) model, seem most tenuously associated with other parts of the protein. In fact, in some cases the A and B domains seem to be unattached to any of the remaining domains of the protein.

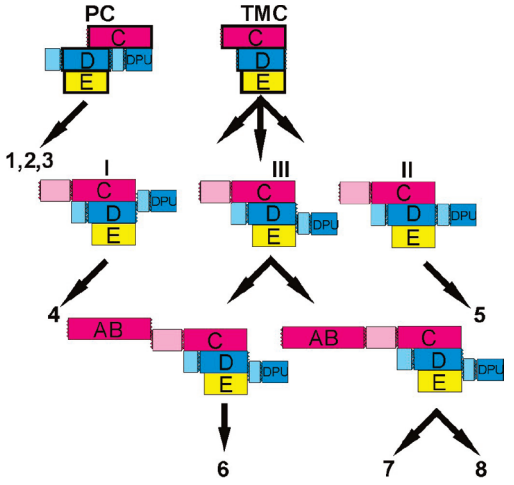

FIGURE 1: Schematic view of the assembly of RPA structural models. Segments of RPA structure which are portions of the protein's 14, 32, and 72 kDa subunits are colored in hues of yellow, blue, and red, respectively. Domain regions are more intensely colored than nondomain regions and contain a letter identifying the domain, where possible. Cores (first row) and principle intermediates (second and third rows) are represented by cartoon figures. Arrows indicate paths resulting in complete candidate RPA structure models (Arabic numbers). PC corresponds to the predicted core structure and TMC to the minimum trimerization core (experimentally determined in an RPA fragment). Roman numerals indicate intermediates that arose upon addition of C-terminal portions of the 32 kDa subunit (32L and DBU). This is indicated by different positions of these regions in the figures in row 2. Alternate intermediates in the bottom row resulted from addition of components of the 70 kDa subunit and are indicated by different positions of the red segments. The bold boxed portion of the PC core represents the PTMC model used in Table 2. (Please note that arrangements of structural segments in the cartoons presented in this figure are not designed to reproduce the actual structural features of the intermediates.)

(*iii*) *Preliminary Analysis of Candidate Protein Structures*. We evaluated docking results at intermediate stages of assembly based on the most reactive residues (reactivity class 3). This was accomplished by requiring that these residues be exposed to solvent. Mismatches between observed and anticipated reactivities of residues in other reactivity classes were apparent. These were useful in selecting the best, from alternate intermediate and final models. For example, residues 70−560E and 70−568R are part of both "core" models but with different solvent exposures (red surface in the left panel of Figure 2, Table 2). These residues are exposed in the TMC core as well as in all structure models built upon it (Table 2), consistent with their substantial observed reactivity (Table 2 of the Supporting Information). In contrast, even in a truncated model of our initial core, which contains only the same domains as the TMC core (bold-boxed portion of PC, Figure 1; PTMC in Table 2), 70−560E and 70−568R still have very low degrees of solvent exposure. They also have low degrees of solvent exposure in models built upon this core. We conclude that these models (tabulated in the left three columns under models in Table 2) are flawed. Differences in these core models are shown graphically in

Table 2: Solvent Accessibility of Representative Residues in Precursor and Full RPA Models

Model Type	Residue	Rx Class[a]	Model[b] or Relative Solvent Accessible Surface of Residue							
Precursor			PTMC[c]			TMC				
	70-560	E	14.3	14.3	14.3	71.8	71.8	71.8	71.8	69.2
	70-568	R	19.3	19.3	19.3	42.8	42.8	42.8	45.8	41.1
Complete			**1**	**2**	**3**	**6**	**7**	**8**	**5**	**4**
	70-560	E 2	14	14	14	74	72	72	72	69
	70-568	R 2	19	19	19	36	36	36	43	42
Precursor							**III**		**II**	**I**
	32-180	R					7.6		53.1	48
Complete			**1**	**2**	**3**	**6**	**7**	**8**	**5**	**4**
	32-180	R 2	5	8	10	4	8	7	53	78
	14-104	K -3	80	65	86	33	86	87	80	0

[a] Reactivity class/score. [b] Models named as summarized in Figure 1 and the Supporting Information. [c] Predicted core corresponding to removal of 32L, 32N, and DPU from 32+Cr in the Supporting Information (Figure 11, line 1). This makes a core of C, D, and E domains analogous to the TMC core.

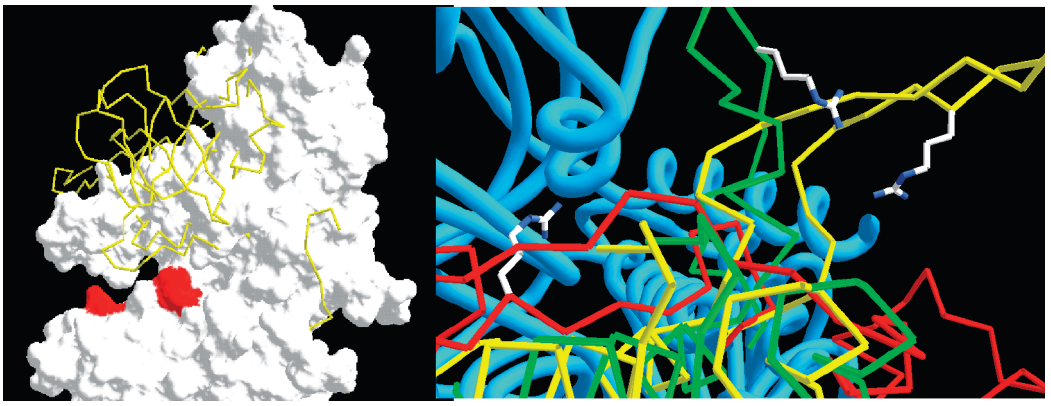

FIGURE 2: Fatal flaw analysis of alternate models of RPA assembly intermediates. In the left panel, the model of the observed "trimerization core" structure is represented by the white surface while the predicted trimerization core model is shown as a yellow stick model. Predicted and observed models are aligned so that they overlap very well in the region of residues 70−560E and 70−568R. In the right panel, alternate models with differing positions or orientations of domain DPU and 32L are shown. These structures were aligned using the observed trimerization core portion of the structure common to all of them (blue tube). Conformations of DPU and 32L in the analogous intermediates leading to model 4 (yellow), model 5 (green), and model 7 (red) are shown in their various positions. The 32−180R side chain position in respective models is also shown (CPK-colored side chains).

the left panel of Figure 2. Residues 70−560E and 70−568R from the respective models are superimposed and represented as a red surface. The backbone structure of the predicted core is represented by the yellow tubular structure, and the experimentally derived TMC model is represented by the white surface. The yellow structure more nearly occludes residues 70−560E and 70−568R than the white surface.

During the next step in the assembly of models based on the TMC core, three nondomain sequences and the C-terminal domain of RPA, DPU (PDB entry 1DPU) were added. While two of the nondomain sequences were near the periphery of the model, the third, which covalently connected DPU to the core, as well as DPU itself made intimate contact with the central core. Three different orientations of the latter region and DPU passed initial screening based on exposure of reactivity class 3 residues [Figure 1, row 2 (I, II, III)]. These models are superimposed in the right panel of Figure 2. The substantial reactivity of residues 32−180R (CPK-colored side chain in all models) is consistent with this residue's exposure in two of the intermediate models, I (Figure 1; Figure 2, yellow Cα trace) and II (Figure 1; Figure 2, green Cα trace) (Table 2, two right-most columns under the models heading). The third intermediate model, III (Figure 1; Figure 2, red Cα trace), is inconsistent with these reactivity results since the region of residues 32−180R is virtually buried in that model and all structural models derived from it (Table 2; Figure 2, right panel). Of the structures that are relevant after application of the criteria just described (Table 2, two right-most columns under the models heading), the absence of reactivity at residues 14−104K suggests that model 4 (right-most column in Table 2) is most relevant to native RPA. In this model, the region of residues 14−104K is buried while the same residue is solvent accessible in the alternative model (Table 2, model 5).

Throughout the RPA structure, the "fatal flaw" analysis just described is repeated with different residues modified as part of this study. It is an effective screen of candidate structures; however, the approach is time-intensive and tends to focus on

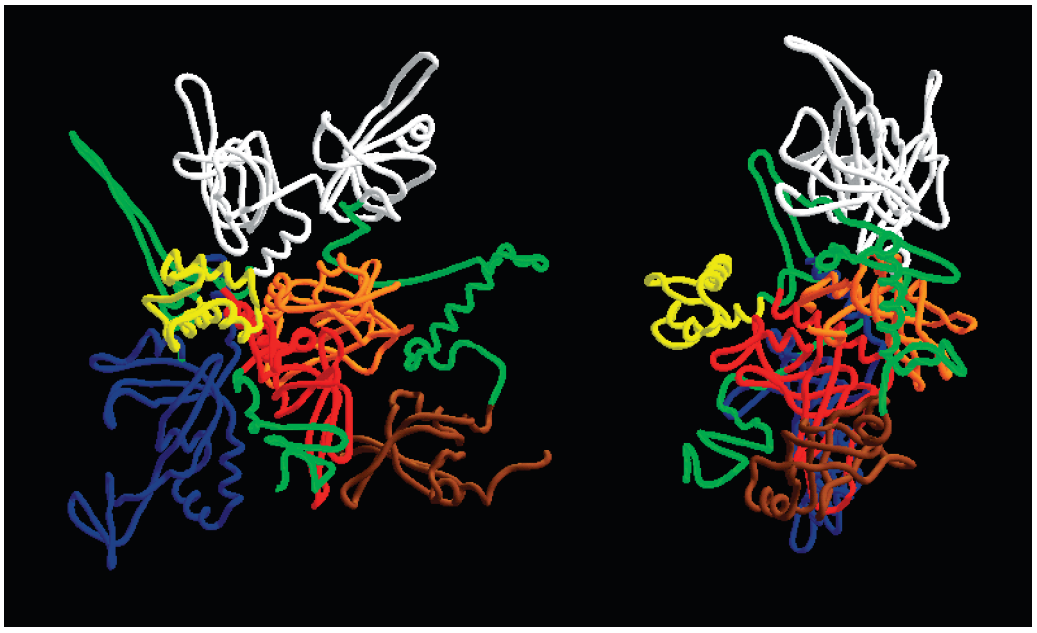

FIGURE 3: Tube respresentation of the backbone of the "best" assembled RPA structure, efll. Domains A and B are colored white. Domains C−F are colored blue, red, orange, and brown, respectively. Large (> 20 amino acids) nondomain regions are colored green. Relative to the view in the left panel, the model in the right panel is rotated 90° clockwise about a vertical axis through the model.

Table 3: Reactivity Scores for Candidate Structural Models

model[a]	score[b]	model[a]	score[b]
4	208	3	178
5	202	7	170
6	190	1	158
2	186	8	158

[a] Models named as in Figure 1. [b] Scoring as described in Materials and Methods.

information from a limited portion of the protein structure. An alternative is to score models on the basis of all residues whose reactivities were monitored.

(*iv*) *Reactivity-Based Scoring of Candidate Structures*. The eight completed structural models presented here (Figures 12 and 14 of the Supporting Information) were scored according to how well the observed reactivity of each probed amino acid residue matched its predicted reactivity. The prediction was based on the solvent accessibility of each amino acid in each model (Materials and Methods and ref *20*). For these analyses, results from all probed side chains in a particular model were combined to judge model relevance. Results in Table 3 show that the highest scoring model is model 4. An orthographic tube rendering of model 4 is shown in Figure 3. This is the same model favored on the basis of the fatal flaw observations mentioned above (Table 2).

(*v*) *Domain Interactions*. Domains of the best scoring RPA structural models interact directly with other domains except in the case of domain A (Table 4). This domain interacts with domain B virtually entirely through the short unstructured amino acid sequence, ABL (Figure 3 of the Supporting Information). Similarly, a portion of the large nondomain region of the 70 kDa subunit (70L1 of Figure 5) mediates the domain's remaining interactions, principally with domain E. In this model, no other large nondomain sequences are involved in mediating domain−domain interactions.

(*vi*) *Molecular Dynamics Simulations*. To relax the structural model (*30*) and to evaluate conformational flexibility that might characterize portions of RPA's structure, molecular dynamics (MD) simulations were performed on structural model 4. Results are summarized in Figure 4, where blue and cyan colors denote the least movement (root-mean-square deviation) and yellow and red the most movement (see Materials and Methods). Generally, nondomain regions are located on the periphery and are among the most dynamic portions of the model (yellow and red at the middle left, middle right, and middle bottom edges). In contrast, the trimerization core domains of the RPA model (Figure 4, left view, central and bottom left, predominantly cyan and blue) are not very flexible in the time frame of the simulations used. However, some movement in the A and B domains (labeled A and B, respectively, in the left panel of Figure 4) is apparent (predominantly green, some cyan, and some yellow). In fact, from these MD simulations, the cores of both the A and B domain folds are more mobile than other domain folds in the model. In addition, domain A appears to undergo larger motions than domain B.

Table 4: RPA Domain−Fragment Interactions[a]

	E	32N	D	32L	DPU	F	70L1	A	B	70L2	C
E			X			X	X		X	X(sm)	X
32N			X								
D					X	X				X	X
32L					±				X(sm)		X(sm)
DPU											
F							X				
70L1								X			
A									±[b]		
B										X	X
70L2											X
C											

[a] Portions of RPA structure named as described in the abbreviations and Supporting Information (Figure 1). Interactions involving four or more, fewer than 4, or one or two amino acid side chains are indicated by X, X(sm), and ±, respectively. [b] Interaction between these domains is mediated by a short unstructured sequence (ABL in Figures 9 and 11 of the Supporting Information).

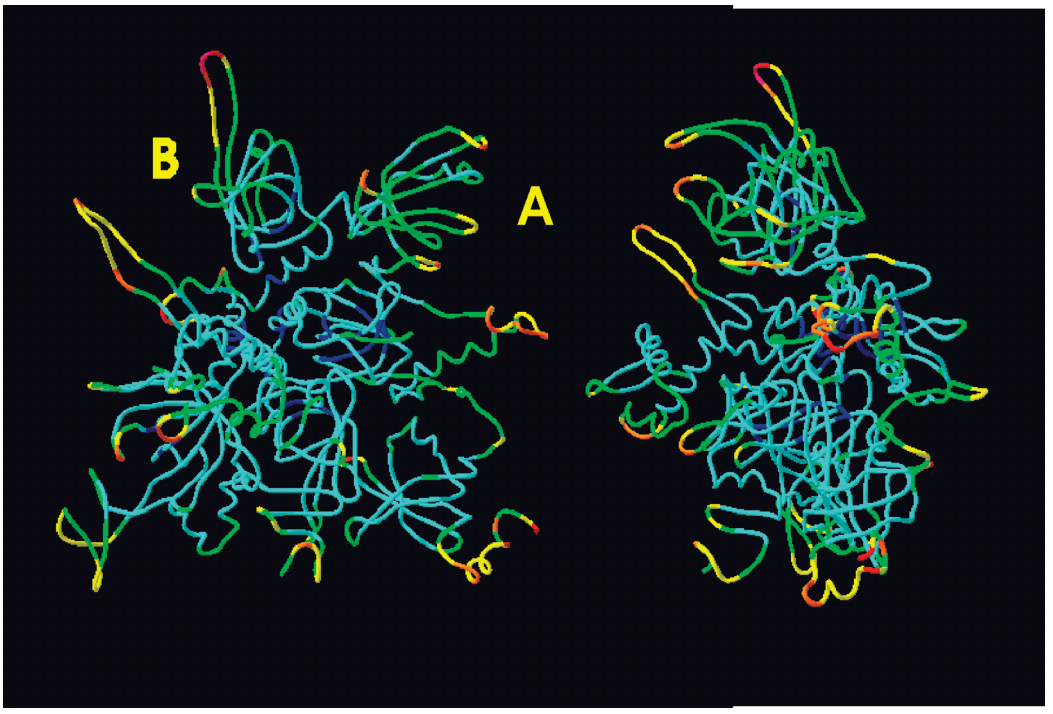

FIGURE 4: Orthogonal views of the backbone structure of structural model 4 colored according to MD simulation results. Red denotes the greatest variation of amino acid α carbons during MD simulation and blue the smallest (see Materials and Methods). Domains A and B are indicated by the respective letters. In the right panel (model rotated 90° clockwise about a vertical axis relative to left panel), domain A substantially eclipses domain B.

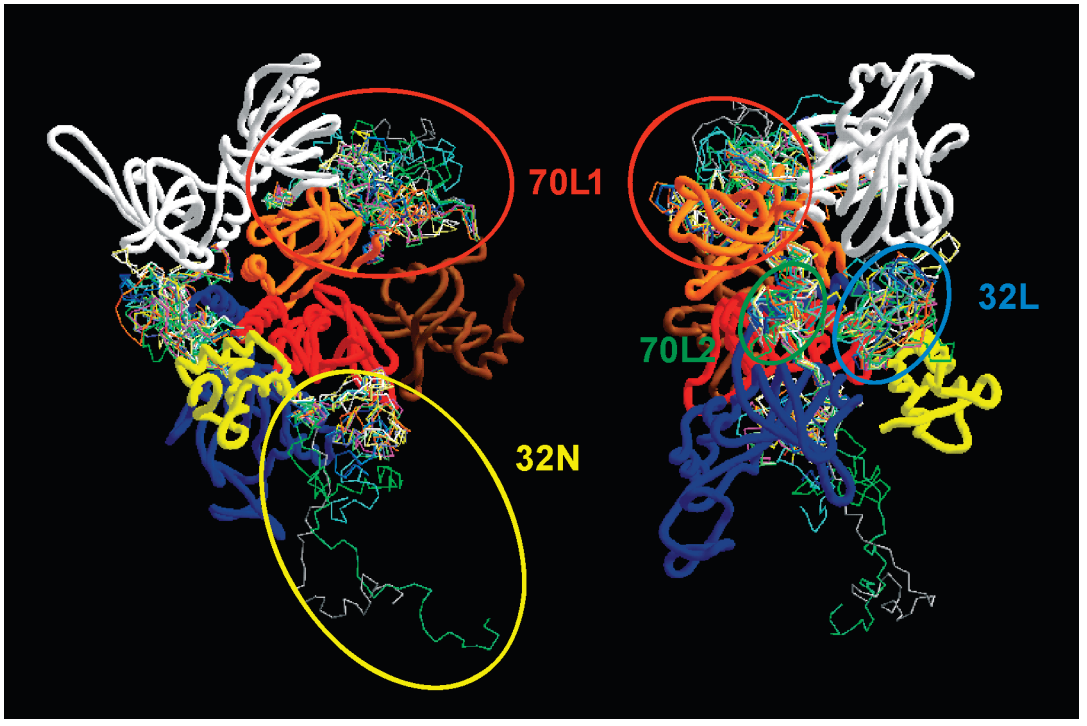

FIGURE 5: Superposition of 17 RPA structural models derived from structural model 4 upon energy minimization, molecular dynamics simulation, and simulated annealing manipulations. Domain regions (which closely overlapped in all models) are shown as a single tube and are colored as in Figure 3. α carbon traces of individual models of the major nondomain regions are shown as stick figures. The PRA structures in the right panel are rotated by 90° counterclockwise about the vertical axis relative to the structure in the left panel.

(*vii*) *Simulated Annealing*. To further explore the conformational properties of our RPA model's nondomain regions, we performed simulated annealing (SA). Initial results recapitulated MD results, in that most domains moved much less than interdomain sequences. Therefore, domain locations were fixed during the subsequent analyses. A total of 11 SA cycles were performed, one each on 11 different, but very similar, structure 4 models selected from model warming and MD simulations (Materials and Methods). Results from these experiments demonstrate that the peptide backbones of the four large

nondomain sequences generally did not return to their initial conformations. Furthermore, elements of secondary structure in these regions neither formed nor re-formed during simulations. These observations are consistent with the nondomain portions of RPA being intrinsically random coil, or disordered, as predicted on the basis of primary structure analysis (*31*−*33*). Conformations of the nondomain sequences are shown in Figure 5 (Cα traces) together with the protein's domain regions (represented as a single tube). Figure 5 shows that the 32L region is very close to the 70L2 region in these models. (See the abbreviations and/or Figure 9 of the Supporting Information for the nomenclature used to describe RPA structure segments.) However, in each model, the nondomain sequences were sufficiently far apart that they did not exhibit direct contact during the simulations. The 70L1 and 32N regions are both far from other nondomain sequence regions. Therefore, the movements of these four large nondomain regions are judged to be essentially independent of each other.

The 32N region is extended in some SA models and was much more compact in others (Figure 5). We gauged extension of this region in each model by calculating its surface area after it was excised from the rest of the model. Surface area values for the 32N region varied by 40%. Similar surface area values for the 70L1 region in each model varied by less than 20%, while the surface areas for the 32L and 70L2 regions varied considerably less.

(*viii*) *Energies of Models Generated during MD and SA Analyses*. Empirical force field (FF) energies of models produced during MD and SA refinement of structural model 4 (best scoring) were calculated using the SPDBV implementation of the GROMOS force field (*34*). While it is unlikely that the calculated energies are sufficiently accurate to make their absolute values meaningful, their relative values and their trends are useful. These FF energies decreased when model 4 was hydrated, minimized, and warmed to 300 °C in preparation for MD simulations. However, the models generated during MD simulations did not have appreciably lower FF energies. Similarly, models generated at the end of SA "runs" all had approximately the same energies, and those were virtually the same as for models resulting from MD. When the SA models were subjected to another round of hydrated energy minimization, their energies decreased again and by virtually the same amount. As judged by superimposition of models before and after this final minimization, this energy decrease caused virtually no change in the respective model's α carbon structures. Representative models overlapped within a 0.4 Å rmsd. Modest, but larger, changes in structural models resulted from the initial hydration and warming of the model 4 structure (2.19 Å rmsd) and subsequent MD simulation (100 ps) (3.1 Å rmsd total).

FF energies for the large nondomain portions of models were positive, except in the HM70, HM80, and HM90 models (Figure 6, top panel). These values are generally consistent with the classification of these regions as being intrinsically disordered (*35*). A linkage between compactness, as judged by the isolated region's surface area and FF energy, was observed for the 32N region (Figure 6, bottom panel). In fact, we found a positive correlation between FF energy and the region's extension (R^2 = 0.78), suggesting that a compact conformation of this region is favored (Figure 6, bottom panel).

Amino Acids with Unusual Reactivity. The best-scoring RPA models were readily identified using only solvent accessibility as a criterion for judging how well observed and expected

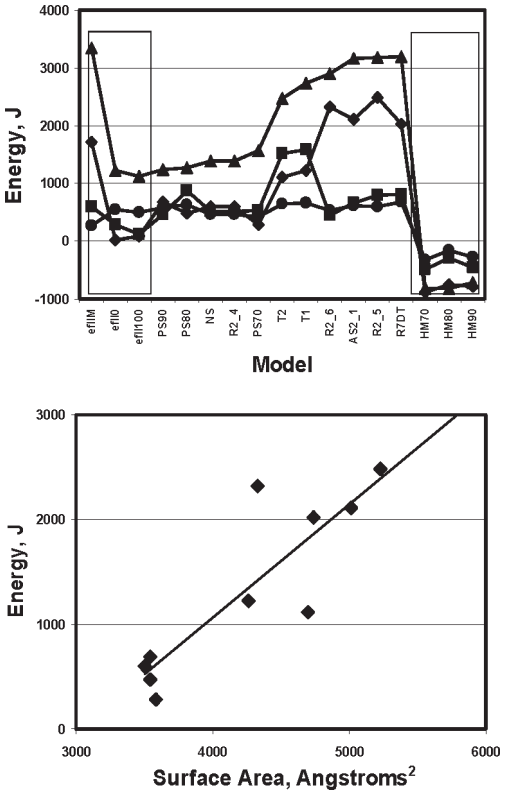

FIGURE 6: In the top panel are shown FF energy estimates of the 32N (◆), 32L (■), 70L1 (▲), and 70L2 (●) nondomain regions in the of RPA structural models shown in Figure 5. The horizontal axis labels identify specific structural models. Energy estimates were made using the Swiss PDB viewer implementation of GROMOS96. The bottom panel shows a correlation between the total force field energy of the 32N nondomain region of RPA structural models and their respective surface areas. The straight line is a correlation line and has a correlation coefficient of 0.78.

reactivity matched (Tables 2 and 3). However, not all amino acids had positive structure scores in all of the assembled models, and SA did not result in a model that made them all positive. The HM80 model was among a group of lowest-energy models derived from structural model 4 by SA and subsequent energy minimization, yet 14 of the 156 measured reactions had negative structure scores. Of these, five had structure scores of −1, which indicates only slight disagreement with our expectation, six had scores of −2, indicating apparently significant disagreement, and three seemed to be in substantial disagreement, with structure scores of −3. Features of these 14 reaction sites are summarized in Table 5 and discussed below.

(*i*) *Multiple Reactions on a Single Peptide*. Negative scores were assigned to several amino acids that are potentially reactive to hydroxyl radicals (see OH in the reagent column of Table 5). In these cases, more than one potential hydroxylation target was present in the peptides that were analyzed (peptide indicated in the relevant footnotes of Table 5). Therefore, we calculated percent solvent accessibility, by taking the combined exposure of all potential "highly reactive" hydroxylation targets in each tryptic peptide (*31*, *36*, *37*), divided by the maximum potential exposure of the same residues in each model. This was used for calculating a "peptide" structure score, which was then assigned to each reactive site in the peptide. However, the results can also be explained in terms of one residue in these peptides being sufficiently exposed to be reactive (70−415, 70−498, and 70−567 in

Table 5: Amino Acids with Unexpected Reactivities in the HM80 RPA Structural Model

Residue	R[a]	Reagent[b]	Exposure			ES	Alt	Score[g]	
			HM80[c]	All Models[d]	TMC/ (Frag)[e]	Envir.[f]	Dock	Rx	Struct.
70-414	W[h]	H	OH	N	0-5	(36)		1	-1
70-415	F[h]	H	OH	Y	8-34	(35)		1	1
70-498	Y[i]	H	OH	Y	17-31	21		1	1
70-500	C[i]	H	OH	N	0-2	0		1	-1
70-558	F[j]	H	OH	N	0-3	30		1	-1
70-562	F[j]	H	OH	N	0-3	4		1	-1
70-567	F[j]	H	OH	Y	1-30	16		1	1
32-65	R	H	T	N	1-6	40		2	-2
14-39	K	H	NHS	N	0-95	89(40)		2	-2
70-234	R	H	T	N	22-30	(21.7)		3	-1
70-612	R	H	T	N	2-43	65		2	-2
70-344	R	H	T	N	8-12	10		2	-2
32-145	K	H	NHS	N	0-5	3	568R (4.2) 564N (3.0)	2	-2
70-367	K	L	NHS	Y	36-91	101	363E (2.9) 364D(3.1)	-3	-3
70-449	K	L	NHS	Y	25-49	49(24)	525E (2.8) 527Q(4.0)	-3	-3
32-81	R	H	T	N	6-12	13	Exp[k]	3	-3
32-84	E	H	V8	N	8-18	69	Exp[k]	2	-2

[a] Reactivity higher (H) or lower (L) than anticipated. [b] Reagent used to modify RPA: OH, hydroxyl free radical; T, trypsin; NHS, sulfo-*N*-hydroxysuccinimide acetate; V8, V8 protease. [c] The exposure of the indicated residue in HM80 is consistent (Y) or not consistent (N) with residue reactivity observations. [d] Range of exposure of the indicated residue in all models resulting form SA refinement of best assembled RPA structure. [e] Exposure of the indicated residue: in a X-ray crystallographic structure of either the appropriate RPA fragment or (in parentheses) in the reported RPA trimerization core (TMC). [f] Charged and polar residues in the proximity of the ε-amine of the indicated residue together with the distance between the amino nitrogen and the closest charged atom of the indicated residue's side chain. [g] Reactivity (RX) and structure (Struct.) score of the indicated residue, determined and assigned as described in Materials and Methods. [h] One of multiple potentially reactive residues on polypeptide fragment 70, residues 413−431. [i] One of multiple potentially reactive residues on polypeptide fragment 70, residues 498−502. [j] One of multiple potentially reactive residues on polypeptide fragment 70, residues 554−568. [k] Residues having >20% relative solvent accessibility values in structure models with domain F in alternate docking locations.

Table 5), while other residue side chains are buried. This would account for the unanticipated reactivity of four of the amino acids listed in Table 5 (top two gray boxes) and is consistent with the proposed structure.

(*ii*) *Multiple or Flexible Conformations*. In the HM80 model of RPA, several residues in different peptides were either buried or had little solvent exposure but were nonetheless reactive. However, viewing SA results as a cadre of RPA conformations, reactive residues may be buried in several, or even most, conformational models, as long as they are solvent-exposed in other models. For example, at locations 14−39K (sulfo-NHS-acetate-reactive) and 70−234R and 70−612R (both trypsin-reactive) (Table 5, bottom gray box), these residues are sufficiently exposed, in at least some of the SA models, to account for their enhanced reactivity ("Exposure: All Models" column). Another residue, 32−65R (top white box under "Exposure" columns), is located at the interface between domain D and the C-terminal portion of DPU, which normally covers 32−65R (Figure 7, top left panel). The interface between these domains is relatively polar, and they are attached to RPA through a peptide that is not part of any secondary structure. This suggests that there are no strong interactions immobilizing these portions of the RPA structure and that the domain juxtapositions are likely to vary with respect to each other. Accordingly, this would make substantial solvent exposure very likely, even if only on a transitory basis, and could account for the unanticipated reactivity of 32−65R in our structural model.

Amino acid 70−344R is another residue whose reactivity is anomalous (Table 5, bottom white box spanning the "Exposure" columns). This residue is "buried" in all models derived from X-ray crystallography of RPA fragments containing domains A and B (PDB entries 1FGU and 1JMC, respectively). It is not involved in interactions between domains or between domains and interdomain sequences used to construct our RPA model [Table 5, TMC/(Frag) column] or in the assembly of our best SA-derived models. It seems unlikely, therefore, that assembly of our model is responsible for the unexpectedly high reactivity of 70−344R. Rather, we assert that RPA's native solution state structure varies from the crystal structure,

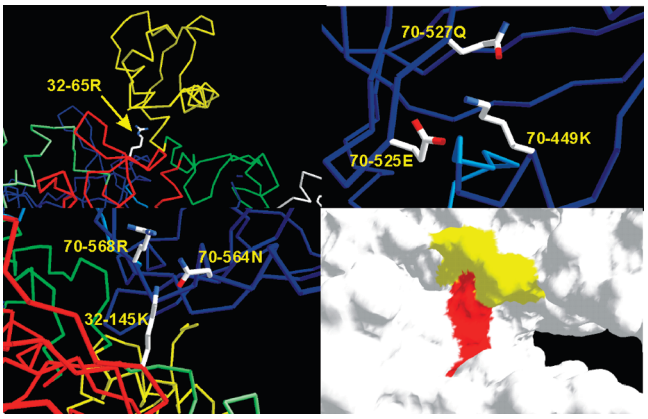

FIGURE 7: α carbon (top panels and bottom left panel) and Connolly surface (bottom right panel) presentation of the HM80 RPA structural model in the region of 32−65R (top left panel), 70−449 (top right panel), and 32−145K (bottom panels). In the bottom right panel, the surface contributed by 32−145K is colored red, the surface contributed by 70−568R and 70−564N is colored yellow, and the remaining surface is colored white.

perhaps by loss of some of the two-stranded β-sheet in the region of 70−344R. This could help explain the increased level of exposure of this residue with only modest structural changes in RPA.

(*iii*) *Electrostatic and Hydrogen Bonding Influences*. The unexpected reactivity of residues 32−145K, 70−367K, and 70−449K seems to arise from electrostatic and hydrogen bonding interactions (white box in the "ES Environment" column of Table 5). Amino acids 70−367K and 70−449K are exposed in our SA models and yet are unreactive (Table 5, "Reactivity" column). Both residues are within direct interaction distance of at least one carboxylic acid side chain (Table 5; for 70−449K, see Figure 7, top right panel) (70−367K not shown). Because of electrostatic interactions with respective negatively charged carboxylic acid groups in these locations (*36*), the positively charged protonated states of 70−367K and 70−449K's ε-amino groups would be favored. Since a protonated primary amine is dramatically less reactive with respect to nucleophilic acetylation than a depronated one, this explains the lack of reactivity at 70−367K and −449K (*27*). In contrast, 32−145K readily reacts with sulfo-NHS-acetate but is not significantly exposed in our best models (Figure 7, bottom panels, red portion of the molecular surface in the bottom right panel). Solvent access to this residue seems to be primarily blocked by residues 70−564N and 70−568R (Figure 7, bottom panels, yellow portion of the surface in the bottom right panel) in our models. The proximity (4−5 Å) of a positively charged arginine side chain (70−568R), coupled with the absence of negatively charged side chains in this area, favors an uncharged, deprotonated, state for 32−145K. Further, since the amide oxygen of 70−564N is within hydrogen bonding distance of 32−145K's ε-amino nitrogen, the asparagine can be an effective H-bond acceptor. This would tend to promote the stability of 32−145K's deprotonated state and more chemically reactive state. Additionally, only modest changes in the location of 70−564N could further increase 32−145K's level of side chain exposure and its reactivity. Finally, the electrostatic attraction arising from the presence of positively charged 70−568R in the vicinity should increase the local concentration of the negatively charged sulfo-NHS-acetate reagent. In combination then, these circumstances can account for the high chemical reactivity of 32−145K.

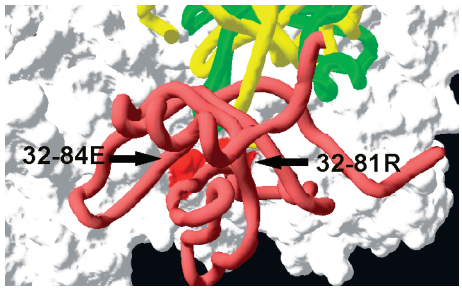

FIGURE 8: Occlusion of 32−81R and 32−84E (red surface) by domain F when it is bound to the remainder of the protein as in the HM80 model (brown tube), but not when it is bound at two alternative low-energy domain F binding sites (yellow and green tubes). Truncated RPA lacking domain F is represented by the white molecular surface.

(*iv*) *Alternate or Multiple Domain Binding Sites*. Finally, the reactivity observed at 32−81R and 32−84E was not expected since these residues are buried by the association of domain F with the core of RPA (Table 5, white box in the "Alt dock" column). Given that we found a number of similar energy docking results for the binding of domain F in our RPA models, it seems possible that it might transiently occupy alternate locations. In fact, removing domain F from the HM80 model, followed by redocking, identified alternate docking sites on RPA's surface. At least two of these sites are located where covalent bonds could reasonably connect domain F to the rest of the protein. The green and yellow tubes in Figure 8 indicate domain F bound at alternate sites in the vicinity of 32−81R and 32−84E (arrows; both colored cyan). The brown tube indicates domain F positioned as in the HM80 model. In both alternate models, 32−81R and −84E are substantially more exposed to solvent than in the HM80 model, results consistent with their reactivity.

In sum, our proteolytic and chemical modification results are in very good agreement with the RPA structural model we present here.

DISCUSSION

Here we report a predicted structural model for RPA that has been tested by 156 chemical and proteolytic reactions at 136 sites throughout the protein's structure (Table 1, Figures 1−8, and Table 1 of the Supporting Information). The model was assembled hierarchically using experimentally determined models of constituent domains and non-template-based predicted models for the protein's large (> 20 amino acids) nondomain regions (Figure 1 and Figures 9−11 of the Supporting Information). The methods and strategies used, along with MRAN, have been described previously (*20*). Our proteolytic and chemical modification results are in very good agreement with the predicted RPA structural model.

Previously Reported Chemical and Enzymatic Modifications. Researchers have probed RPA for many years using proteolytic and chemical modifications. Most studies, however, did not analyze reaction and cleavage locations at the amino acid level, and virtually none examined the kinetics of the reactions. Yet, other researchers reported results involving 60 sites on RPA that are relevant to this study. They used the same reagent we did (trypsin) or reagents with chemical specificity similar to those used here (*7*, *19*, *22*, *33*, *38*). Although reaction conditions frequently varied and reaction rates were not examined, the

Table 6: Comparisons between Our and Previously Reported Modification Results

site			reagent used[a]					
subunit	residue		here[e]	previously[f]	identification method[b]	res. solvent exposure	agreement[c]	ref[d]
32	38	K	T	T[g]	S		A	*20*, *37*
32	180	R	T	T	SDS		A	*20*, *37*
70	167	K	T	T[g]	S		A[h]	*37*
70	344	R	T	T	SDS		A[h]	*37*
70	354	K	T	T	SDS		A[h]	*37*
70	367	K	T	T	SDS		A[h]	*37*
70	389	R	T	T	SDS		A	*37*
70	410	K	T	T	S		A	*20*
70	431	K	T	T	SDS		A	*22*, *37*
70	443	K	T	T[g]	Inferred		A[h,i]	*22*, *37*
70	489	K	T	T[g]	SDS		A	*37*
70	612	R	T	T	Inferred		A[j]	*22*
70	412	R	T	T	S		NA[k]	*22*
14	33	K	NHS_S	NHS_B	MS		A	*38*
14	49	K	NHS_S	NHS_B	MS		A	*38*
32	38	K	T	NHS_B[g]	MS		A	*38*
32	40	R	T	HPG	MS		A	*8*
32	93	K	NHS_S	NHS_B	MS		A	*38*
32	138	K	T	NHS_B	MS		A	*38*
32	139	K	T	NHS_B	MS		A	*38*
70	88	K	NHS_S	NHS_B	MS		A	*38*
70	92	R	T	HPG	MS		A	*8*
70	206	K	NHS_S	NHS_B	MS		A	*38*
70	216	R	T	HPG	MS		A	*8*
70	220	K	NHS_S	NHS_B	MS		A	*38*
70	331	K	T	NHS_B	MS		A	*38*
70	335	R	T	HPG	MS		A	*8*
70	379	K	T	NHS_B	MS		A	*38*
70	431	K	NHS_S	T	SDS		A	*20*
70	443	K	NHS_S	T[g]	Inferred		A	*22*
70	472	R	T	HPG	MS		A	*8*
70	489	K	NHS_S,T	NHS_B[g]	MS		A	*38*
70	502	K	NHS_S	NHS_B	MS		A	*38*
70	573	R	T	HPG	MS		A	*8*
70	575	R	T	HPG	MS		A	*8*
70	577	K	NHS_S	NHS_B	MS		A	*38*
70	586	R	T	HPG	MS		A	*8*
70	588	K	NHS_S	NHS_B	MS		A	*38*
70	585	K	NHS_S	NHS_B	MS		A	*38*
70	600	R	T	HPG	MS		A	*8*
70	605	R	T	HPG	MS		A	*8*
70	611	R	T	HPG	MS		A	*8*
70	461	Y	OH	Xtn	SDS		A[l]	*8*
14	30	R		HPG	MS	29	C	*8*
32	133	R		HPG	MS	34	C	*8*
70	91	R		HPG	MS	24	C	*8*
70	159	Y		Xtn	MS	49	C	*8*
70	163	K		NHS_B[g]	MS	65	C	*38*
70	165	F		Xtn	MS	40	C	*8*
70	167	K		NHS_B[g]	MS	28	C	*38*
70	210	R		HPG	MS	67	C	*8*
70	212	W		Xtn	MS	29	C	*8*
70	244	K		NHS_B	MS	29	C	*38*
70	263	K		NHS_B	MS	37	C	*38*
70	324	K		NHS_B	MS	23	C	*38*
70	326	Y		Xtn	SDS	39	C	*8*
70	343	K		NHS_B	MS	26	C	*38*
70	347	Y		Xtn	SDS	43	C	*8*
70	382	R		HPG	MS	32	C	*8*
70	478	Y		Xtn	SDS	23	C	*8*
70	604	R		HPG	MS	19	C	*8*
70	183	K		NHS_B	MS	2	I	*38*
70	202	R		HPG	MS	6	I	*8*
70	259	K		NHS_B	MS	6	I	*38*

[a] Reagents used to modify RPA include the following: T, trypsin; V8, *S. aureus* strain V8 protease; NHS_S, sulfo-*N*-hydroxysuccinimide acetate; OH, hydroxyl free radical; HPG, hydroxyphenyl glyoxal; NHS_B, acetyl-*N*-hydroxysuccinimide biotin; Xtn, chymotrypsin. [b] Method of peptide identification leading to the assignment of a particular reaction site: S, chemical peptide sequencing; MS, mass spectrometric determination of peptide molecular weight; SDS, SDS−PAGE determination of peptide molecular weight; Inferred, no or very slow reaction at indicated residue. [c] A, reactive/nonreactive residue reported here and that reported by other researchers agree, whether identical reagents were used or ones with similar reactivity were employed in this and previous studies; C, previously reported results are consistent with residue exposure in our structural model(s), though reactivity at the indicated residue was not observed by us; I, reaction at the indicated residue reported in a previous study is apparently inconsistent with its exposure in our RPA model(s). [d] Literature references for previously reported results. [e] Reagent used in this study. [f] Reagent used in the previous study. [g] The reactivity toward multiple reagents relevant to this study has been reported at this site. [h] Agreement with at least one of several possible reactive residues proposed in the previous report. [i] Very low reactivity observed in this study. [j] Limited reactivity observed in this study; no reactivity reported in the previous, single-time point study. [k] No reactivity observed in this study; others reported reactivity only in a RPA fragment. [l] Hydroxylation rate of the entire peptide only was determined, though the peptide (70:450−472) contains multiple potentially reactive sites (461Y, 462F, and 470Y).

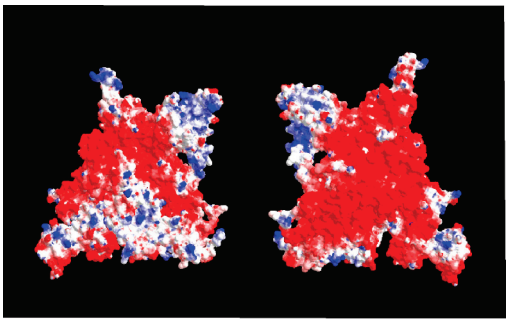

FIGURE 9: Molecular graphic of the Connolly surface for the RPA model HM80 colored according to electrostatic potential. In the left panel, RPA is oriented in the same way as in the left panel of Figure 2. In the right panel, the model is rotated 180° about a vertical axis through the model. Electrostatic calculations were performed using the Swiss PDB Viewer implementation of GROMOS96.

agreement with our proposed model at 57 of the 60 sites is very good (see Table 6). The three problematic sites are found at interfaces involving domain A and are consistent with a flexible, multiconformational attachment of that domain to the RPA (see below). Then, previous studies confirm our results in general. In fact, they add 18 novel reactive sites to the RPA model, all of which are consistent with the model we propose.

RPA Structural Model. The RPA structural model presented here contains stable domains and highly flexible nondomain regions. Its overall structure is discoidal, and its surface is predominantly negatively charged with neutral and positive patches coinciding with ssDNA or protein binding sites (Figure 9). This leaves one face of the structure largely negative, and a seemingly attractive target for interactions with positively charged molecules.

(*i*) *Domains*. Structural models of RPA [both before (Figure 3) and after (Figure 5) refinement by MD and SA methods] have DNA binding sites [domains A and B (white), C (blue), and D (red)] exposed to solvent. With the exception of domain D, they are on the protein's periphery. This placement is consistent with sufficiently long ssDNA binding simultaneously to domains A−D (*1*, *39*). Three of four apparent protein binding sites on RPA, domain F (brown), domain A (right white domain), and domain DPU (yellow) (*3*, *6*, *18*, *40*−*46*), are also exposed and accessible to protein ligands. We will discuss the fourth binding site below. (See the abbreviations and Figure 9 of the Supporting Information for the nomenclature describing RPA structural segments.)

(*ii*) *Long Nondomain Regions*. Four relatively long (>20 amino acids) regions of the RPA primary structure are coil or intrinsically disorganized as judged by primary structure analysis (*35*, *47*, *48*) (Figure 3, green tubes; Figure 5, Cα trace regions of graphic). These are also on the periphery of the RPA model and with the exception of 70L1 do not mediate domain−domain interactions in our model (Table 4 and Figures 3 and 4). Simulated annealing analyses indicate that each of these nondomain regions has variable conformations (Figure 5) and that individual conformations have marginal stability as judged by calculated force field energies, yet the overall RPA conformation is quite stable. These nondomain regions occur spatially in the areas between domains B and C, between domains C and D, and between domains E/F and A. Interestingly, 32N, the intrinsically disordered N-terminal region of RPA 32, assumes conformations that range form being rather compact to quite extended (Figure 5). Our SA results suggest that this nondomain region has a lower energy in more compact conformations (Figure 6).

(*iii*) *Flexibility and Multiple Juxtapositions of Domains*. Some portions of our RPA model seem likely to be flexible or have multiple distinct conformations. Specifically, interactions between either domain A (right white domain in Figure 3 and left panels of Figure 5, respectively) or domain F (brown, Figures 3 and 4) and the remainder of RPA involve some amino acid residues with different reactivity in our study and those of others (Table 6) (*7*, *38*) or have reactivities implying multiple binding locations (domain F, Table 5). This notion is consistent with our RPA model in which domain A's interaction with the remainder of RPA is mediated by nondomain regions, including an intrinsically disorganized region (70L1) (Table 4) and a short linker between domains A and B. MD simulations (Figure 4) indicate that domain A is more flexible than the rest of the protein and, further, that its neighbor, domain B, is itself somewhat flexible, though it is rather rigidly attached to the protein core through its C-terminal helix [in the "triple helix region" (*19*)].

Functional Inferences. (*i*) *ssDNA Binding*. To accommodate reported ssDNA binding behavior, the ssDNA binding site arrangement in our model must evolve as progressively longer pieces of ssDNA bind (*13*, *14*, *39*). Assuming that three nucleotides are sufficient to fill each binding site, domain A must reorient relative to domain B for an eight-nuclotide ssDNA to completely occupy domain A and B ssDNA-binding sites as reported by Zou (*39*). We anticipate that the new orientation is very similar to the structure reported for a RPA fragment containing domains A and B bound to an eight-nucleotide ssDNA oligomer (*13*). Subunits A and B must again reorient relative to domain C to bring these sites close enough so that a 16-nucleotide length of ssDNA occupies all sites simultaneously (*39*). This structural change may correspond to the large ssDNA-binding-linked conformational change in RPA originally observed in electron microscopic studies (*49*). Finally, to bind to the last ssDNA binding site, D, the unstructured N-terminal region of subunit 32 must be extended to vacate the pathway between the ssDNA binding site of domains C and D (Figure 5). We anticipate the energy required to drive these changes in domain juxtaposition and 32N conformation must derive from binding energy developed by ssDNA−RPA interactions.

Interestingly, the model presented here does not preclude simultaneous binding of multiple ssDNA molecules to RPA or a single ssDNA molecule binding to different sites. This feature is consistent with reports of 5′ overhang binding to RPA using domain C and D ssDNA binding sites (*50*, *51*) and suggests that novel ssDNA binding modes may exist.

(*ii*) *Phosphorylation*. Through the action of multiple DNA damage-dependent and cell cycle-dependent phosphorylation pathways (*3*, *4*, *7*, *8*, *49*), as many as five negatively charged phosphates are placed in a short (~30) amino acid sequence of RPA (the 23N regions). This region is devoid of positively charged amino acid side chains. An intriguing aspect of this portion of the model we present here is that it likely acts as a gate. When it is in its most stable unphosphorylated conformation (Figure 6), which is compact, it blocks the ssDNA binding path linking domains C and D. When the region is in an extended conformation, it opens this path. We speculate that phosphorylation shifts 32N's conformational equilibria toward extended conformations like those in Figure 6. Such

a conformation would minimize unfavorable charge−charge interactions among phosphates in this region. Interestingly, the extended C−D gate sequence is also sufficiently long to readily interact with the A, B, or C domain ssDNA binding sites as argued by Zou (*7*), Wold, and others (*4*, *6*, *52*) as well as with the RPA 32 N-terminal (DPU) and F domain protein binding areas. Regardless of the domain to which phosphorylated 32N binds, it will interfere with ssDNA or protein binding by direct competition or, in the case of domain F, by blocking the path between domains C and D.

(*iii*) *Protein Binding*. Several protein binding areas on RPA have been reported. They are located (A) in a region of domain A colocalized with its ssDNA binding site (Figure 5, right white domain, left panel) (*46*), (B) in a region consisting of domain A, some of domain B, and an unstructured region C-terminal to domain A (*42*) [together this area, hereafter termed A2p, is on the opposite side of the domains' ssDNA binding sites (Figure 5, at the junction of the white domains with the rest of RPA)], (C) on the C-terminal winged helical domain of RPA32 (DPU) (Figure 5, yellow domain) (*18*, *43*), and (D) in the basic groove of domain F (Figure 5, brown domain) (*40*, *41*, *53*). The binding sites on the A, F, and DPU domains are readily accessible to the solvent and are substantially separated. We anticipate they are available for multiple, perhaps simultaneous binding interactions, with the caveat that the bulk of these ligands may cause mutual interference. However, the region forming A2p is substantially occluded. Liganding at that site requires a change in the conformation of the RPA model shown here, in which the A and B domains dissociate from a portion of the RPA core to expose the A2p binding site. Such a dissociation seems reasonable in view of the number of docking structures we found with domains A and B loosely associated with the rest of RPA, and the reorientation of domains A and B associated with ssDNA binding of a RPA fragment (*16*).

(*iv*) *Interactions among Binding Sites*. We anticipate binding competition between most ssDNAs and proteins binding at domain A since most ssDNA associates at that site in combination with other sites. We also anticipate that the association of the phosphorylated 32N region of RPA with any of the binding areas mentioned above will compete with ssDNA or protein association at the respective sites (*52*).

Interactions between a protein (SV 40 T antigen helicase) binding at the A2p site and ssDNA binding at domains A and B have been reported (*42*). Without occupying the same sites, this protein as well as ssDNA must change the orientation of domains A and B (see above). We suggest this reorientation organizes domains A and B to optimize ssDNA association while exposing the A2p protein association site (see above). This synergistic interaction between ssDNA binding and proteins that associate at the A2p site provides a mechanism by which protein and ssDNA binding may be allosterically modulated. Further, it provides a means by which proteins associating with RPA may change binding locations as more avid binding sites become unmasked.

In total, the RPA structural model we report here is a structural prediction supported by experimental evidence. The modeling process we used builds on and supplements elegant X-ray and NMR structural studies of RPA fragments. Importantly, our model is consistent with our results as well as other authors' chemical modification and proteolysis results. Taken together, these observations constitute a substantial collection of experimental probes that test the structural relevance of our model at many locations throughout its structure. We suggest this improves the confidence in this model as a structural context for RPA's biological functions. The organization of ssDNA and protein binding sites on this structural model rationalizes or predicts a number of interactions among ligands associating with RPA. Characterizaiton of some of these complexes is in progress in our laboratory.

NOTE ADDED IN PROOF

Since the initial submission of this work, a communication suggesting the architecture of RPA has been published (*54*). That paper emphasizes the RPA's potential for dynamic positioning of domains that are not part of the protein's core.

SUPPORTING INFORMATION AVAILABLE

A complete discussion of analytical models used to quantify poteolytic and chemical modification time courses, examples of this analysis applied to our data presented here, as well as methods for correlating resulting rate constants with specific reaction sites in RPA structures is presented. Also included are proteolytic reaction controls as well as a comprehensive table of chemical and enzymatic modification results used to select the best from among alternate intermediate and complete RPA structures. The selection of models for interdomain regions is described, and molecular graphics of these models are presented. Finally, the scheme used to "assemble" candidate RPA models is described in detail, and molecular graphics are presented. This material is available free of charge via the Internet at http://pubs.acs.org.

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. Iftode, C., Daniely, Y., and Borowiec, J. A. (1999) Replication protein A (RPA): The eukaryotic SSB. *Crit. Rev. Biochem. Mol. Biol. 34*, 141–180.
3. Zou, Y., Liu, Y., Wu, X., and Shell, S. M. (2006) Functions of human replication protein A (RPA): From DNA replication to DNA damage and stress responses. *J. Cell. Physiol. 208*, 267–273.
4. Binz, S. K., Sheehan, A. M., and Wold, M. S. (2004) Replication protein A phosphorylation and the cellular response to DNA damage. *DNA Repair 3*, 1015–1024.
5. Wu, X., Shell, S. M., and Zou, Y. (2005) Interaction and colocalization of Rad9/Rad1/Hus1 checkpoint complex with replication protein A in human cells. *Oncogene 24*, 4728–4735.
6. Bochkareva, E., Kaustov, L., Ayed, A., Yi, G. S., Lu, Y., Pineda-Lucena, A., Liao, J. C., Okorokov, A. L., Milner, J., Arrowsmith, C. H., and Bochkarev, A. (2005) Single-stranded DNA mimicry in the p53 transactivation domain interaction with replication protein A. *Proc. Natl. Acad. Sci. U.S.A. 102*, 15412–15417.
7. Liu, Y., Kvaratskhelia, M., Hess, S., Qu, Y., and Zou, Y. (2005) Modulation of replication protein A function by its hyperphosphorylation-induced conformational change involving DNA binding domain B. *J. Biol. Chem. 280*, 32775–32783.
8. Oakley, G. G., Patrick, S. M., Yao, J., Carty, M. P., Turchi, J. J., and Dixon, K. (2003) RPA phosphorylation in mitosis alters DNA binding and protein-protein interactions. *Biochemistry 42*, 3255–3264.
9. Vassin, V. M., Wold, M. S., and Borowiec, J. A. (2004) Replication protein A (RPA) phosphorylation prevents RPA association with replication centers. *Mol. Cell. Biol. 24*, 1930–1943.
10. Gomes, X. V., Henricksen, L. A., and Wold, M. S. (1996) Proteolytic mapping of human replication protein A: Evidence for multiple structural domains and a conformational change upon interaction with single-stranded DNA. *Biochemistry 35*, 5586–5595.
11. Brill, S. J., and Bastin-Shanower, S. (1998) Identification and characterization of the fourth single-stranded-DNA binding domain of replication protein A. *Mol. Cell. Biol. 18*, 7225–7234.

12. Murzin, A. G. (1993) OB(oligonucleotide/oligosaccharide binding)-fold: Common structural and functional solution for non-homologous sequences. *EMBO J. 12*, 861–867.
13. 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.
14. 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.
15. Jacobs, D. M., Lipton, A. S., Isern, N. G., Daughdrill, G. W., Lowry, D. F., Gomes, X., and Wold, M. S. (1999) Human replication protein A: Global fold of the N-terminal RPA-70 domain reveals a basic cleft and flexible C-terminal linker. *J. Biomol. NMR 14*, 321–331.
16. Olson, K. E., Narayanaswami, P., Vise, P. D., Lowry, D. F., Wold, M. S., and Daughdrill, G. W. (2005) Secondary structure and dynamics of an intrinsically unstructured linker domain. *J. Biomol. Struct. Dyn. 23*, 113–124.
17. Mer, G., Bochkarev, A., Gupta, R., Bochkareva, E., Frappier, L., Ingles, C. J., Edwards, A. M., and Chazin, W. J. (2000) Structural basis for the recognition of DNA repair proteins UNG2, XPA, and RAD52 by replication factor RPA. *Cell 103*, 449–456.
18. Bochkarev, A., Bochkareva, E., Frappier, L., and Edwards, A. M. (1999) The crystal structure of the complex of replication protein A subunits RPA32 and RPA14 reveals a mechanism for single-stranded DNA binding. *EMBO J. 18*, 4498–4504.
19. 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.
20. Nuss, J. E., Sweeney, D. J., and Alter, G. M. (2006) Reactivity-based analysis of domain structures in native replication protein A. *Biochemistry 45*, 9804–9818.
21. Henricksen, L. A., Umbricht, C. B., and Wold, M. S. (1994) Recombinant replication protein A: Expression, complex formation, and functional characterization. *J. Biol. Chem. 269*, 11121–11132 [erratum, (1994) *J. Biol. Chem. 269*, 16519].
22. Bochkareva, E., Korolev, S., and Bochkarev, A. (2000) The role for zinc in replication protein A. *J. Biol. Chem. 275*, 27332–27338.
23. Nuss, J. E., and Alter, G. M. (2004) Denaturation of replication protein A reveals an alternative conformation with intact domain structure and oligonucleotide binding activity. *Protein Sci. 13*, 1365–1378.
24. Case, D. A., Cheatham, T. E.III, Darden, T., Gohlke, H., Luo, R., Merz, K. M.; Jr., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The Amber biomolecular simulation programs. *J. Comput. Chem. 26*, 1668–1688.
25. Cheatham, T. E.III, and Young, M. A. (2000) Molecular dynamics simulation of nucleic acids: Successes, limitations, and promise. *Biopolymers 56*, 232–256.
26. Letner, C., and Alter, G. M. (1994) The Evolution of Lactate Dehydrogenase Conformation During Molecular Dynamics Simulation , ARInternet.
27. Means, G., and Feeney, R. (1971) Chemical Modification of Proteins , Holden-Day Inc., San Francisco.
28. Fraczkiewicz, R., and Braun, W. (1998) Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J. Comput. Chem. 19*, 319–333.
29. Bystroff, C., and Shao, Y. (2002) Fully automated ab initio protein structure prediction using I-SITES, HMMSTR and ROSETTA. *Bioinformatics 18* (Suppl. 1), S54–S61.
30. Lee, M. R., Tsai, J., Baker, D., and Kollman, P. A. (2001) Molecular dynamics in the endgame of protein structure prediction. *J. Mol. Biol. 313*, 417–430.
31. Nukuna, B. N., Sun, G., and Anderson, V. E. (2004) Hydroxyl radical oxidation of cytochrome c by aerobic radiolysis. *Free Radical Biol. Med. 37*, 1203–1213.
32. Cantor, C., and Schimmel, P. (1980) Biphysical Chemistry Part 1: The Conformation of Biologcial Macromolecules , W. H. Freeman and Co., San Francisco.
33. Pestryakov, R., Weisshart, K., Schlott, B., Khodyreva, S., Kremmer, E., Grossie, F., Larvrik, O., and Nasheuer, H. (2003) The C-Terminal RPA70 and the Central RPA32 Domains are Involved in the Interactions with the 3′-End of a Primer-Template DNA. *J. Biol. Chem. 278*, 17515–17524.
34. Guex, N., and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis 18*, 2714–2723.
35. Li, X., Romero, P., Rani, M., Dunker, A. K., and Obradovic, Z. (1999) Predicting Protein Disorder for N-, C-, and Internal Regions. *Genome Inf. Ser. 10*, 30–40.
36. Takamoto, K., and Chance, M. R. (2006) Radiolytic protein footprinting with mass spectrometry to probe the structure of macromolecular complexes. *Annu. Rev. Biophys. Biomol. Struct. 35*, 251–276.
37. Maleknia, S. D., Ralston, C. Y., Brenowitz, M. D., Downard, K. M., and Chance, M. R. (2001) Determination of macromolecular folding and structure by synchrotron X-ray radiolysis techniques. *Anal. Biochem. 289*, 103–115.
38. Shell, S. M., Hess, S., Kvaratskhelia, M., and Zou, Y. (2005) Mass spectrometric identification of lysines involved in the interaction of human replication protein a with single-stranded DNA. *Biochemistry 44*, 971–978.
39. 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.
40. 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.
41. 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.
42. 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.
43. 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. *Nat. Struct. Mol. Biol. 12*, 332–339.
44. Doherty, K. M., Sommers, J. A., Gray, M. D., Lee, J. W., von Kobbe, C., Thoma, N. H., Kureekattil, R. P., Kenny, M. K., and Brosh, R. M. Jr. (2005) Physical and functional mapping of the replication protein a interaction domain of the werner and bloom syndrome helicases. *J. Biol. Chem. 280*, 29494–29505.
45. 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.
46. Daughdrill, G. W., Buchko, G. W., Botuyan, M. V., Arrowsmith, C., Wold, M. S., Kennedy, M. A., 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.
47. Romero, P., Obradovic, Z., Kissinger, C. R., Villafranca, J. E., Garner, E., Guilliot, S., and Dunker, A. K. (1998) Thousands of proteins likely to have long disordered regions. *Pac. Symp. Biocomput. '98,* 437–448.
48. Linding, R., Jensen, L. J., Diella, F., Bork, P., Gibson, T. J., and Russell, R. B. (2003) Protein disorder prediction: Implications for structural proteomics. *Structure 11*, 1453–1459.
49. Blackwell, L. J., Borowiec, J. A., and Masrangelo, I. A. (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.
50. Dickson, A. M., Krasikova, Y., Pestryakov, P., Lavrik, O., and Wold, M. S. (2009) Essential functions of the 32 kDa subunit of yeast replication protein A. *Nucleic Acids Res. 37*, 2313–2326.
51. Pestryakov, P. E., Khlimankov, D. Y., Bochkareva, E., Bochkarev, A., and Lavrik, O. I. (2004) Human replication protein A (RPA) binds a primer-template junction in the absence of its major ssDNA-binding domains. *Nucleic Acids Res. 32*, 1894–1903.
52. Binz, S. K., and Wold, M. S. (2008) Regulatory functions of the N-terminal domain of the 70-kDa subunit of replication protein A (RPA). *J. Biol. Chem. 283*, 21559–21570.
53. Xu, X., Vaithiyalingam, S., Glick, G. G., Mordes, D. A., Chazin, W. J., and Cortez, D. (2008) The basic cleft of RPA70N binds multiple checkpoint proteins, including RAD9, to regulate ATR signaling. *Mol. Cell. Biol. 28*, 7345–7353.
54. 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.