4172–4181 *Nucleic Acids Research, 2005, Vol. 33, No. 13*
doi:10.1093/nar/gki736

Solution structure of the DNA-binding domain of RPA from *Saccharomyces cerevisiae* and its interaction with single-stranded DNA and SV40 T antigen

Chin-Ju Park, Joon-Hwa Lee and Byong-Seok Choi*

Department of Chemistry, National Creative Research Initiative Center, Korea Advanced Institute of Science and Technology, 373-1, Guseong-dong, Yuseong-gu, Daejon 305-701, Korea

Received April 13, 2005; Revised June 23, 2005; Accepted July 9, 2005 PDB no. 1YNX

ABSTRACT

Replication protein A (RPA) is a three-subunit complex with multiple roles in DNA metabolism. DNA-binding domain A in the large subunit of human RPA (hRPA70A) binds to single-stranded DNA (ssDNA) and is responsible for the species-specific RPA–T antigen (T-ag) interaction required for Simian virus 40 replication. Although *Saccharomyces cerevisiae* RPA70A (scRPA70A) shares high sequence homology with hRPA70A, the two are not functionally equivalent. To elucidate the similarities and differences between these two homologous proteins, we determined the solution structure of scRPA70A, which closely resembled the structure of hRPA70A. The structure of ssDNA-bound scRPA70A, as simulated by residual dipolar coupling-based homology modeling, suggested that the positioning of the ssDNA is the same for scRPA70A and hRPA70A, although the conformational changes that occur in the two proteins upon ssDNA binding are not identical. NMR titrations of hRPA70A with T-ag showed that the T-ag binding surface is separate from the ssDNA-binding region and is more neutral than the corresponding part of scRPA70A. These differences might account for the species-specific nature of the hRPA70A–T-ag interaction. Our results provide insight into how these two homologous RPA proteins can exhibit functional differences, but still both retain their ability to bind ssDNA.

INTRODUCTION

Replication protein A (RPA) is a single-stranded (ss) DNA-binding (SSB) protein that is likely to be the eukaryotic equivalent of the *Escherichia coli* SSB protein (1). RPA is involved in many aspects of eukaryotic DNA metabolism, such as DNA replication, nucleotide excision repair, mismatch repair and genetic recombination [reviewed in (2)].

In humans, as in other eukaryotes, RPA consists of three distinct, evolutionarily well-conserved subunits, RPA70 (70 kDa), RPA32 (32 kDa) and RPA14 (14 kDa) (3). RPA70 interacts with ssDNA as well as with other proteins, such as XPA and Rad51 (4–6). The SSB activity of human RPA70 (hRPA70) resides mainly in the central region of the protein (amino acid residues 175–420, called hRPA70AB), which also interacts with other proteins (2). The N-terminal domain (residues ~1–170) participates in specific protein–protein interactions, and the C-terminal domain (residues ~450–616) interacts with hRPA32 and hRPA14 (7). hRPA32, which is regulated by cell-cycle-dependent N-terminal phosphorylation, can also bind to ssDNA and interacts with several protein partners, such as XPA and Rad52 (8–11).

Saccharomyces cerevisiae RPA (scRPA) has also been studied not only with biochemical methods, but also with a genetic approach. The scRPA70 subunit also bears the major SSB activity in its central part (scRPA70AB; residues 180–416), and the composition of its domains is matched precisely with that of hRPA70 (12,13). scRPA has similar ssDNA-binding properties to human RPA. It can interact with ssDNA non-sequence specifically, keeping preference to polypyrimidine, such as hRPA (14). Genetic studies in *S.cerevisiae* have shown that the genes encoding all three RPA subunits are essential for

*To whom correspondence should be addressed at Department of Chemistry, Center for Repair System of Damaged DNA, Korea Advanced Institute of Science and Technology, 373-1, Guseong-dong, Yuseong-gu, Daejon 305-701 Korea. Tel: +82 42 869 2828; Fax: +82 42 869 2810; Email: byongseok.choi@kaist.ac.kr
Present address:
Joon-Hwa Lee, Department of Chemistry and Biochemistry, University of Colorado at Boulder, Boulder, CO 80309, USA

The authors wish it to be known that, in their opinion, the first two authors should be regarded as First Authors

viability (15,16). Mutational analyses with scRPA revealed that some mutants are defective in intrachromosomal recombination (17), the recombinational repair of induced double-stranded breaks (18) and meiotic recombination (19), whereas others exhibit either mutator phenotypes (20) or a stimulation of recombination among direct repeats, ultraviolet (UV) irradiation and methyl methanesulfonate sensitivities, growth thermosensitivity and replication defects (18). From the results of these genetic studies, several hypotheses regarding the *in vivo* functions of scRPA have been put forth (19). However, the absence of 3D structures for each of the proteins that make up the scRPA complex has made these hypotheses difficult to prove.

Although the two RPA homologs from human and *S.cerevisiae* are highly similar (Figure 1), one should be careful not to apply what is known about the structure of hRPA to scRPA, because the two proteins are not functionally equivalent. First, antibodies to hRPA do not cross-react with scRPA (21). This indicates that the surface antigens of the homologs vary significantly. Second, none of the genes that encode the three subunits of hRPA can complement the corresponding null mutations in yeast (15). Third, hRPA can support Simian virus 40 (SV40) DNA replication *in vitro*, while scRPA cannot (3,22).

In an *in vitro* SV40 replication system, RPA is required for SV40 large tumor antigen (T-ag)-dependent, origin-dependent unwinding of the SV40 origin of replication. During initiation, RPA interacts with the SV40 T-ag and the DNA polymerase δ–primase complex (Pol/Prim) (23), resulting in the formation of an initiation complex at the replication origin. Physical interactions between RPA, T-ag and Pol/Prim are necessary for the initiation of SV40 DNA replication (24) and for the function of RPA in the elongation stage (25). Moreover, hRPA is required for the initiation of DNA replication (which is detected by the synthesis of small RNA primers by Pol/Prim), and it cannot be replaced by scRPA (26), suggesting a specific role for hRPA in SV40 DNA replication. T-ag was shown to bind to RPA through the major ssDNA-binding domain in hRPA70 and possibly through the C-terminus of RPA32 (27,28). The mapped RPA-binding site in SV40 T-ag is embedded within the origin DNA-binding domain (OBD) of T antigen, whose solution structure is available (29,30). Finally, it has been reported that scRPA has a reduced binding affinity for SV40 T-ag compared with hRPA (31). Given this collection of findings, we deemed that any new structural information for scRPA70A (residues 181–294; DNA-binding domain A in RPA70 from *S.cerevisiae*) can be used, not only for explaining the molecular bases of the scRPA70A genetic studies, but also for providing clues to elucidate how these closely related proteins function so differently (32,33).

In this study, we determined the solution structure of the free form of scRPA70A by multidimensional heteronuclear NMR spectroscopy and deduced a model of its ssDNA-bound forms employing residual dipolar coupling (RDC)-based homology modeling. Comparison with the hRPA70A structure indicated that the ssDNA-binding regions are conserved through evolution, even though the conformational changes that occur upon ssDNA binding are not identical. Chemical shift perturbation experiments with the SV40 T-ag OBD indicated that the SV40 T-ag OBD interaction surface on RPA70 is separated from the RPA70 ssDNA-binding site. These results provide insight into the mechanisms by which these two homologous domains differ in their ability to interact with T-ag, while displaying similarities in their modes of ssDNA binding.

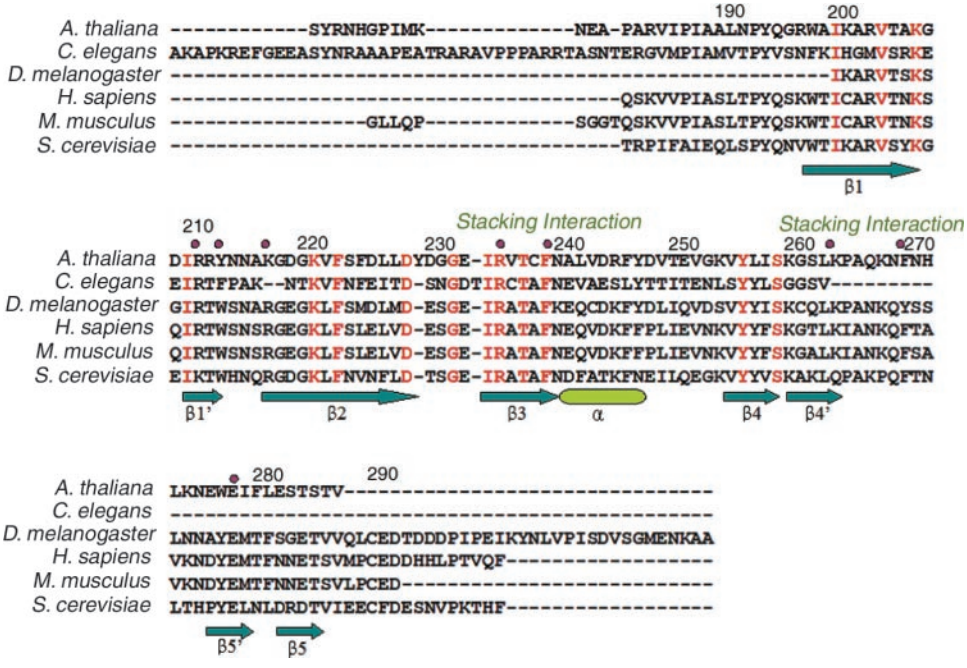

Figure 1. Sequence alignment of RPA70A from several organisms with the corresponding secondary structural elements noted. The alignment was performed with the T-coffee server (61). Residues are numbered according to RPA70A from *S.cerevisiae*. The residues in red are highly conserved across the organisms. The dots indicate residues that interact with ssDNA.

MATERIALS AND METHODS

Sample preparations

The gene encoding scRPA70A(*RPA70*$_{181–294}$) was cloned into the pET14b vector as an N-terminal histidine-tagged fusion (Novagen), and the construct was used to transform *E.coli* strain BL21(DE3)pLysS. Uniformly ^{15}N- and ^{15}N/^{13}C-labeled proteins were obtained by growing the transformed *E.coli* cells in M9-minimal media containing $^{15}NH_4Cl$ (Cambridge Isotopes Inc.) and unlabeled/$^{13}C_6$-labeled-D-glucose (Cambridge Isotope Inc.) as the sole nitrogen and carbon sources, respectively. The labeled proteins were initially purified with a Ni-NTA affinity column (Pharmacia Inc.). After the thrombin digestion reaction, samples were loaded onto a Superdex-75 (Pharmacia Inc.) gel filtration FPLC column. The purity and homogeneity of all samples were confirmed by SDS–PAGE. hRPA70A was expressed and purified as described previously (6).

The gene encoding the SV40 T-ag OBD (*T-ag-OBD*$_{131–260}$) was cloned into the pGEX-1λT vector (Novagen) as a glutathione *S*-transferase-tagged fusion, and the construct was used to transform *E.coli* strain BL21. The SV40 T-ag OBD was purified on a Glutathione Sepharose Uniflow column (Clontech, Inc.). After the thrombin digestion reaction, the protein was purified with on a Superdex-75 (Pharmacia Inc.) gel filtration FPLC column. The purity and homogeneity of the SV40 T-ag OBD sample were confirmed by SDS–PAGE.

The DNA oligonucleotide TT-10 (5′ GCGATTGCG 3′) was purchased from Bioneer, Inc. (Daejon, Korea). Oligonucleotides were purified by reversed-phase high-performance liquid chromatography. The purity and homogeneity of the products were confirmed by matrix-assisted laser desorption ionization spectroscopy.

NMR spectroscopy and solution structure calculation

All NMR experiments were performed on a Varian Inova 600 MHz spectrometer (KAIST, Daejon) equipped with a triple-resonance ^{1}H/^{13}C/^{15}N probe.

The 2D ^{15}N/^{1}H HSQC and 3D ^{15}N-edited NOESY-HSQC were acquired on a uniformly ^{15}N-labeled sample in a 90% H_2O/10% D_2O solution containing 20 mM sodium phosphate, 100 mM NaCl and 2 mM DTT (pH 7.0) at 27°C. The 3D CBCA(CO)NH, HNCACB, HNCO, HCCH-TOCSY, HCCH-TOCSY-NNH and ^{13}C-edited NOESY-HSQC data were collected for an ^{15}N/^{13}C-labeled sample in the same buffer and under the same conditions as described above. The data were processed with NMRPipe (34) and analyzed with the program SPARKY (http://www.cgl.ucsf.edu/home/sparky).

The NOE distance restraints were derived from cross-peaks in the ^{15}N-edited NOESY-HSQC and the ^{13}C-edited NOESY-HSQC. The dihedral angle restraints were extracted from TALOS chemical shift analysis (35) and intensity-modulated ^{15}N/^{1}H HSQC experiments (36). Structure calculation was initially performed using CYANA, which combines automated assignment of NOE cross-peaks and structure calculation (37). On the basis of distance restraints derived from direct CYANA output, structure calculations were also carried out using the internal variable module (38) of XPLOR-NIH (39). The database-derived Rama potential function in XPLOR was used in the calculation (40). The overall results were very similar using both types of software. The 22 structures with the lowest energies calculated with XPLOR-NIH were retained and validated by the program PROCHECK-NMR (41). Experimental data and structural statistics are summarized in Table 1. Structures were visualized using the program MOLMOL (42) and the Swiss-PDB viewer application (43). POV-ray was used for the rendering.

Table 1. Structural statistics for scRPA70A structure determination

NOE upper distance limits	1138
Intra-residual	283
Medium-range ($\|i - j\| < 4$)	401
Long-range ($\|i - j\| > 4$)	454
Dihedral angle constraints	161
Hydrogen bonds	70
NOE violations	
Number > 0.2 Å	0
Mean deviation from ideal covalent geometry	
Bond lengths	0.0027 Å
Bond angles	0.49°
Improper	0.36°
Ramachandran analysis	
Most favored (%)	82.4
Additionally (%)	15.7
Generously (%)	2.0
Disallowed (%)	0.0
Coordinate precision (residues 197–210, 219–264 and 276–288)	
Backbone (Å)	0.90 ± 0.26
Heavy atom (Å)	1.44 ± 0.26

Residual dipolar coupling-based homology modeling of the scRPA70A–TT-10 complex

Three types of dipolar couplings were measured for the ^{15}N/^{13}C-labeled scRPA70A-unlabeled TT-10 oligonucleotide complex crystalline medium, which consisted of 10 mg/ml of the filamentous phage Pf1 (Asla Labs, Riga, Latvia) in a 90% H_2O/10% D_2O solution containing 20 mM sodium phosphate, 100 mM NaCl and 2 mM DTT (pH 7.0) at 27°C. $^1D_{NH}$ were measured in 2D ^{15}N/^{1}H IPAP-HSQC (44). $^1D_{NCO}$ and $^2D_{HNCO}$ were measured in 3D TROSY-based HNCO pulse sequences (45). 75 $^1D_{NH}$, 83 $^1D_{NCO}$ and 70 $^2D_{HNCO}$ were collected and used as the restraints. The normalization factors employed for $^1D_{NCO}$ and $^2D_{HNCO}$ relative to $^1D_{NH}$ were 9.04 and 3.04, respectively (46). The refinement was performed using XPLOR-NIH based on the protocols for calmodulin (47). The structure ensemble of scRPA70A was used as the starting model. A weak Rama potential (40) and a pseudopotential for the radius of gyration (48) were applied in the simulated annealing. The force constant for the RDC restraints started at 0.005 kcal/mol and finally reached to 5 kcal/mol in the first simulated annealing protocol. The value 3 kcal/mol was used as the force constant for the RDC restraints in the second simulated annealing protocol. All other force constants and procedures were the same as those used in the protocols for calmodulin (47).

NMR titration of scRPA70A and hRPA70A with ssDNA and the SV40 T-ag OBD

The 2D ^{15}N/^{1}H HSQC spectra were acquired on uniformly ^{15}N-labeled samples of scRPA70A and hRPA70A in a 90%

H_2O/10% D_2O solution containing 20 mM sodium phosphate, 100 mM NaCl and 2 mM DTT (pH 7.0) at 27°C. Titrations of TT-10 were carried out with both scRPA70A and hRPA70A. The detailed procedure is same as the one described in our previous report (6).

For the SV40 T-ag OBD titration, the concentrations of both hRPA70A and scRPA70A were 0.15 mM. Aliquots of 1.0 mM SV40 T-ag OBD solution, dissolved in the same buffer used for the hRPA70A and scRPA70A samples, were added directly to the NMR cell, and the samples were allowed to equilibrate for several minutes. Changes in average amide chemical shifts ($\Delta\delta_{avg}$) were calculated by the formula $|\Delta H| + |\Delta N|/a$, where ΔH and ΔN are the linear change along the ^{1}H and ^{15}N axes, respectively, and $a = 8$ for glycine and $a = 6$ for all other amino acids (49). The studied complexes were at saturation at the endpoint of titration. Chemical shift perturbations were determined to be significant if they were >0.02 p.p.m. for RPA70A.

RESULTS

Structure determination and overall structures

Standard triple-resonance experiments were used to assign the spectrum of scRPA70A as described in Materials and Methods. Backbone assignment was completed, and side-chain assignment was almost completed. The ^{15}N/^{1}H HSQC spectrum of uniformly ^{15}N-labeled scRPA70A and assignments of the backbone amides are shown and subset of the ^{1}H/^{15}N HSQC cross-peaks observed for free scRPA70A (black) shifted upon the addition of TT-10 (red) (Figure 2).

The structure calculations for scRPA70A were performed using 1138 interproton distance restraints and 161 dihedral angle restraints (Table 1 and Figure 3). A final set of 22 lowest energy structures was selected from 100 calculations. There were no violations larger than 0.2 Å and 5° for the NOEs and dihedral angles, respectively. The superimposed backbone traces of 22 structures were well aligned, as shown in Figure 3A, where residues in the loops (residues 213–217 and residues 265–274) and residues at the N-terminus (residues 181–196) and C-terminus (residues 290–294) were somewhat scattered, probably because they are flexible in nature (see Supplementary Figure 1). The root-mean-square deviation (r.m.s.d.) for backbone heavy atoms in residues in regular secondary elements was 0.90 Å and for all heavy atoms was 1.02 Å. The solution structure of scRPA70A was deposited in the Protein Data Bank (PDB ID 1YNX).

Figure 3B shows a ribbon diagram of the lowest energy structure of scRPA70A, revealing that the domain has a typical oligonucleotide/oligosaccharide binding (OB)-fold, which consists of five stranded β-barrel and one α-helix. As expected from the high sequence similarity, the architecture of scRPA70A closely resembled the crystal structure of hRPA70A (32,33). The Cα atoms in the secondary structural part of the scRPA70A structure were superimposed with their equivalents in the hRPA70A. The r.m.s.d. for the Cα atoms between two proteins was 1.58 Å. Figure 3C shows a superposition of scRPA70A and hRPA70A. While the regular structural regions were well aligned each other, the L12 and L45 loops showed bigger differences. The L12 loop of scRPA70A was bent into the core of the molecule, while the corresponding region of hRPA70A had a rather straight orientation. In contrast, the L45 loop of scRPA70A showed a more opened conformation compared with that of hRPA70A. However, the spacing of the two loops in scRPA70A did not differ significantly from that of the human protein, because the two loops of scRPA70A were shifted forward.

Although the flexible loops of scRPA70A and hRPA70A showed some differences, the structural and sequence comparisons of the two proteins suggested similar modes of ssDNA binding. In hRPA70A crystal structure, intimate contacts with ssDNA were made by R210, R216, R234, F238,

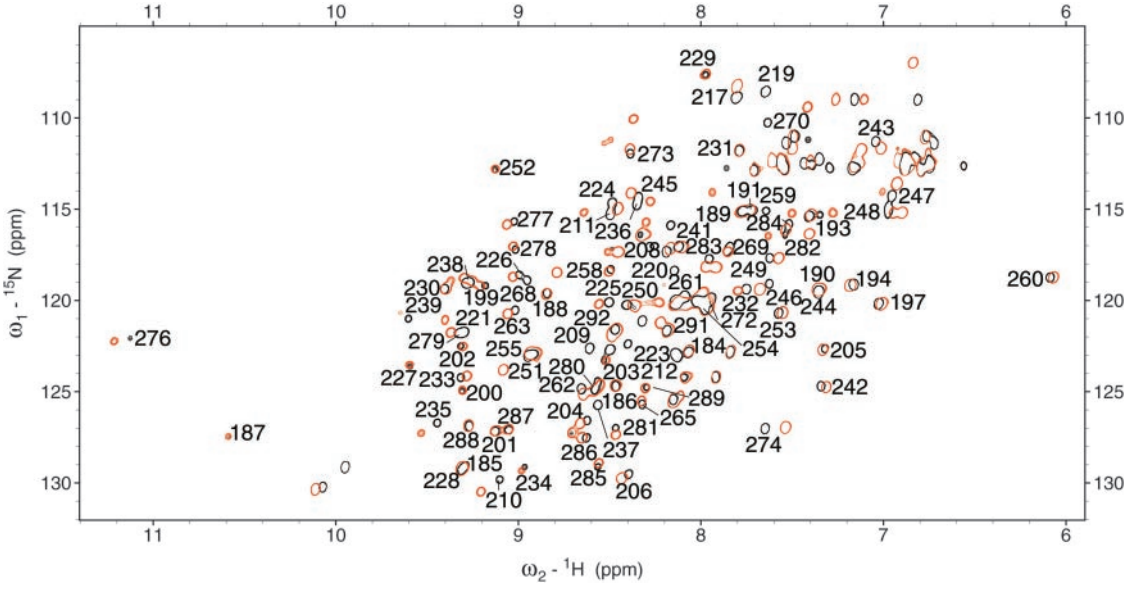

Figure 2. The 2D ^{15}N/^{1}H HSQC spectrum of scRPA70A and assignments of the backbone amides. ^{1}H/^{15}N cross-peaks of scRPA70A in the absence (black) and in the presence of the TT-10 (red).

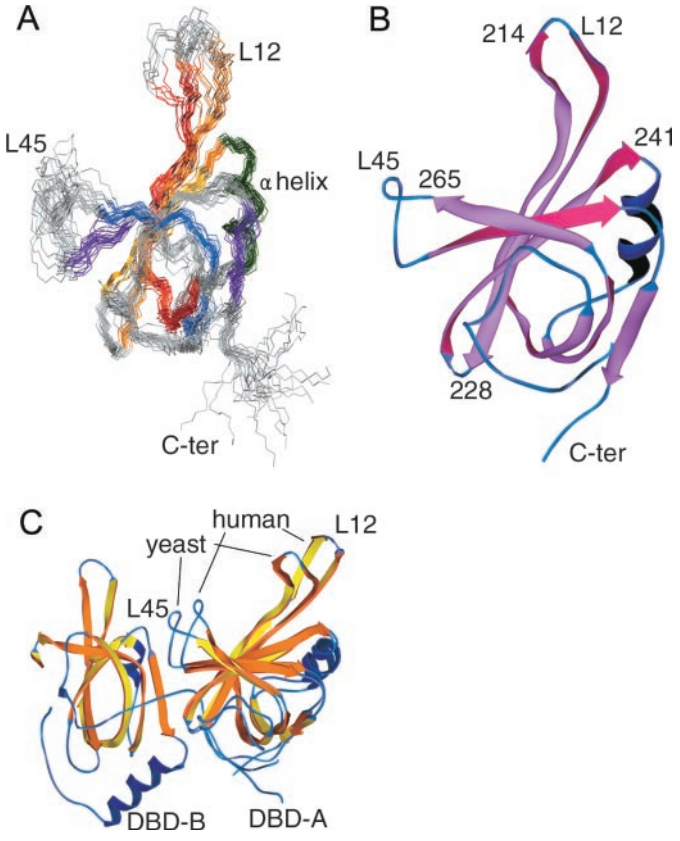

Figure 3. (**A**) Superimposed backbone traces of 22 scRPA70A structures. β1, β1′ are shown in red, β2 in orange, β3 in yellow, α in green, β4, β4′ in blue and β5, β5′ in violet. (**B**) Ribbon diagram of the lowest energy structure of scRPA70A, revealing the existence of a typical OB fold, which consists of five stranded β-barrel and one α-helix. (**C**) Superposition of scRPA70A and hRPA70A 3D structures. The structure of hRPA70A was derived from 1fgu.pdb (33).

Q263, F269 and E277. These amino acids are conserved or conservatively substituted in scRPA70A (described in detail below). A previous report that hRPA70NAB (the N-terminal region and RPA70AB) has nearly identical DNA-binding properties (the occluded binding site size and binding affinity) to scRPA70NAB, where the largest distinction observed between the two fragments was a 2-fold difference in affinity for $(dT)_{30}$, also suggest that structural differences in the flexible loops of the homologous proteins do not significantly influence the ssDNA-binding properties (8).

Interaction with ssDNA

In order to address the conformational rearrangement of scRPA70A upon ssDNA binding, we performed RDC-based homology modeling for the scRPA70A–ssDNA complex instead of doing the traditional NOE-based structure calculations. We considered that it would be difficult to assign resonances of ssDNA and to obtain intermolecular NOEs in a ssDNA–protein complex. Chou *et al*. (47) first demonstrated that the dipolar refinement protocol can be used to define large conformational rearrangements that may take place when sample conditions are altered. The protocol was applied to the protein calmodulin, which experiences a conformational change upon Ca^{2+} binding. We hypothesized that ssDNA binding to scRPA70A induces observable conformational changes in the protein. So we applied the same strategy used for calmodulin to the scRPA70A–ssDNA complex. The calculated free scRPA70A structures were used as the starting model, and the measured dipolar coupling restraints of the scRPA70A–ssDNA complex were used in two sequential low-temperature, simulated annealing procedures. In order to exclude the pitfalls of potential effects of conformational averaging on the magnitude of the observed RDCs, in the calculation, we did not include the RDC values of the residues in the tip of L12 and L45 loops, which are known to be flexible. Also, we performed an ^{15}N relaxation study for the validation of the model (see Supplementary Figure 1).

The magnitudes of the normalized dipolar couplings in the complex at a pf1 concentration of 10 mg/ml ranged from −16.0 to 19.0 Hz. Using the histogram approach, the alignment tensor magnitude, D_a, was estimated at 9.5 with a rhombicity, R, of 0.37 (50). After the second simulated annealing, we selected the lowest energy structure and analyzed the backbone φ, ψ angles. Figure 4A shows the angle differences between the lowest energy structure of free scRPA70A and of the ssDNA-bound form. The residues in the L12 and L45 loops showed large differences in both the φ and ψ angles. The data were consistent with the chemical shift perturbation of scRPA70A observed in the presence of TT-10 ssDNA (Figure 4B). These results imply that ssDNA binding to scRPA70A effects changes to amino acid residues in the L12 and L45 loops. Superimposition of backbone traces (excluding the turns and loops) of the RDC-based model with the lowest energy structure of free scRPA70A yielded an r.m.s.d. of 1.56 Å.

From studies of the crystal structure of hRPA70AB, the internal flexibility of the various domains has been addressed (33). In the case of hRPA70B (residues ∼305–422), the L12 and L45 loops showed large shifts upon ssDNA binding. In contrast, for hRPA70A, conformational changes upon ssDNA binding were observed only in L12; the L45 loop did not move significantly, possibly because of a crystal packing effect. However, a dynamics study of hRPA70A using NMR relaxations showed that both the L12 and L45 loops are flexible in the absence of ssDNA and that their motions are influenced by ssDNA binding (51). From the structure calculations of the free form of scRPA70A and ^{15}N relaxation study, we found that the L12 and L45 loops of scRPA70A also have internal flexibilities (see Supplementary Figure 1). Similar to the phenomenon observed for hRPA70A, the main conformational changes in scRPA70A upon ssDNA binding occurred in the two loops, L12 and L45. While the L12 loop showed only moderate reorientation, larger changes were observed in the L45 region. Amino acid residues in the β4′ region (A260, K261, L262 and Q263) showed large displacements upon ssDNA binding, which moved the L45 loop closer to the L12 loop, compared with the free scRPA70A structure. The loop movement observed for scRPA70A was similar to that which occurs in hRPA70A. Therefore, these results show that the RDC-based homology modeling method is suitable for the detection of conformational changes in the protein partner upon ssDNA binding.

Because we used only RDC constraints for the backbone bonds, it was hard to address the precise orientation of the side

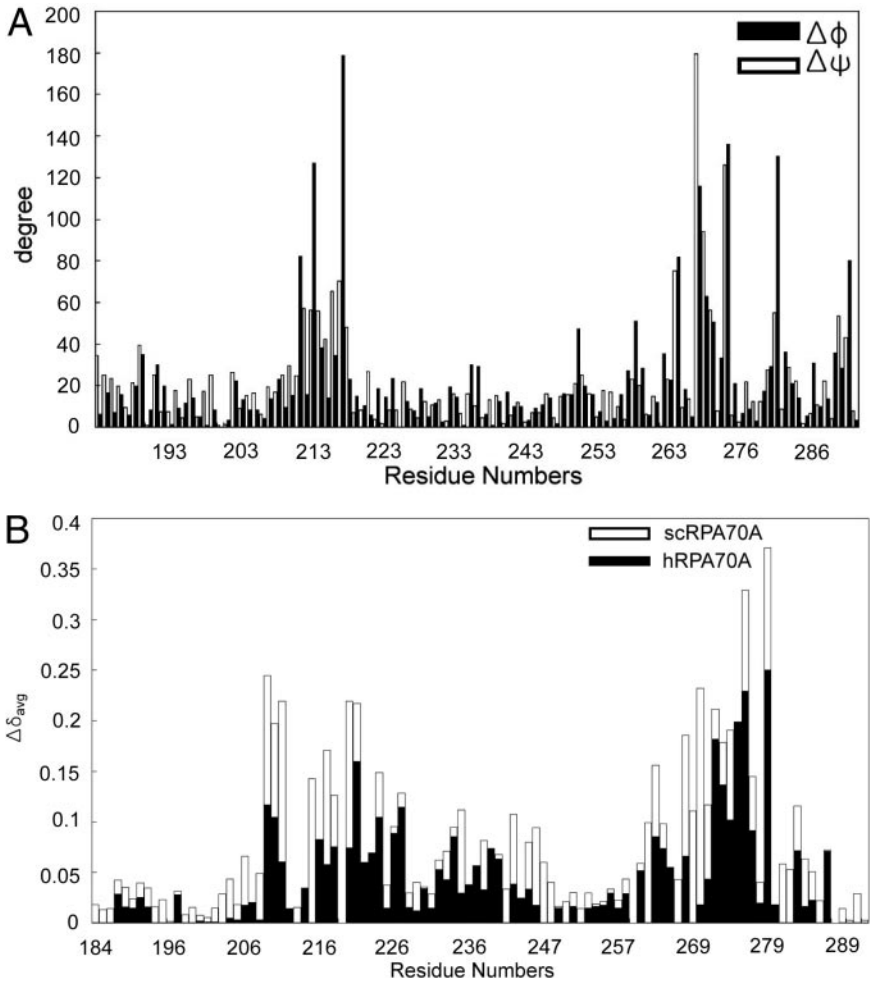

Figure 4. (**A**) Backbone angle differences between the lowest energy structures of scRPA70A in the free and ssDNA-bound forms. (**B**) The average chemical shift changes ($\Delta\delta_{avg}$) in the ^{1}H and ^{15}N resonances of hRPA70A (closed bar) and scRPA70A (open bar) upon addition of TT-10 ssDNA.

chains in the protein. However, it was possible to compare the spatial positions of the backbone atoms. In doing so, we found that the positions of residues critical for ssDNA interactions in hRPA70A were well matched in scRPA70A, e.g. two phenylalanines (F238 and F269) that make stacking interactions with DNA bases in hRPA70A lay in similar positions in scRPA70A. Also, the positions of R210 (K210 in scRPA70A) and E277 were similar in the two proteins. However, in scRPA70A, Q263 was positioned more closely to the spaces in the structure where ssDNA resided than is K263 in hRPA70A. In addition, for scRPA70A, R234 sat further from the ssDNA than does the same residue of hRPA70A. Finally, in accordance with the conformational change in the L12 loop upon ssDNA binding, the relative position of R216 in scRPA70A differed from that of R216 in hRPA70A. Consequently, these changes resulted in a slightly narrower ssDNA-binding arc in scRPA70A relative to hRPA70A (Figure 5A).

The other changes that occurred upon ssDNA binding to scRPA70A were in the helix region (residues 239–248). The α helices in each of the two structures had slightly different orientations (Figure 5B). This observation could imply that ssDNA binding to scRPA70A induces conformational changes, not only in residues in the binding arc, but also in other residues that do not make direct contact with ssDNA.

The residues of hRPA70A that are affected by TT-10 binding have been mapped previously in chemical shift perturbation experiments (6); they are located on the β2, β3, β4′ and β5′ strands and in the L45 loop, and these results are consistent with the crystal structure of hRPA70A (32). When we performed chemical shift perturbation with scRPA70A and TT-10, we observed both similarities and differences as compared with the data for hRPA70A. The perturbation of similar residues in both scRPA70A and hRPA70A upon ssDNA binding implies that the positioning of ssDNA bound to scRPA70A is similar to the positioning of ssDNA bound to hRPA70A, i.e. the binding sites are similar for the two proteins. In addition to the equivalent positions of amino acid residues perturbed by ssDNA binding that were mapped in both scRPA70A and hRPA70A, residues in the helix between β3 and β4 were also substantially perturbed in scRPA70A, but not in hRPA70A (Figures 4B and 5C). From our data herein and those of a previous study (52), we hypothesize that the residues in the helix region of scRPA70A do not contact ssDNA directly. Therefore, the perturbations we observed in the helix region of scRPA70A

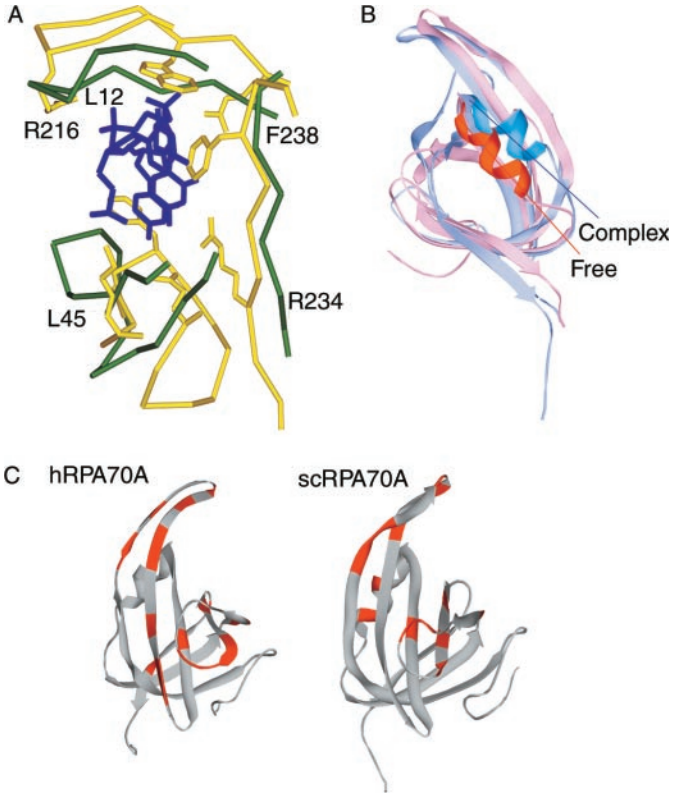

Figure 5. (**A**) Superimposed backbones of the DNA-binding arcs of the ssDNA-bound forms of scRPA70A (green) and of the hRPA70A (yellow)–dC_8 ssDNA (blue) crystal structure. The hRPA70A and dC_8 ssDNA coordinates were derived from 1jmc.pdb (32). (**B**) A ribbon diagram of scRPA70A in which structural changes between the free (pink) and ssDNA-bound (blue) forms are shown. (**C**) A ribbon diagram of RPA70A affected by ssDNA binding. Residues perturbed by ssDNA binding are shown in red. hRPA70A (left), scRPA70A (right).

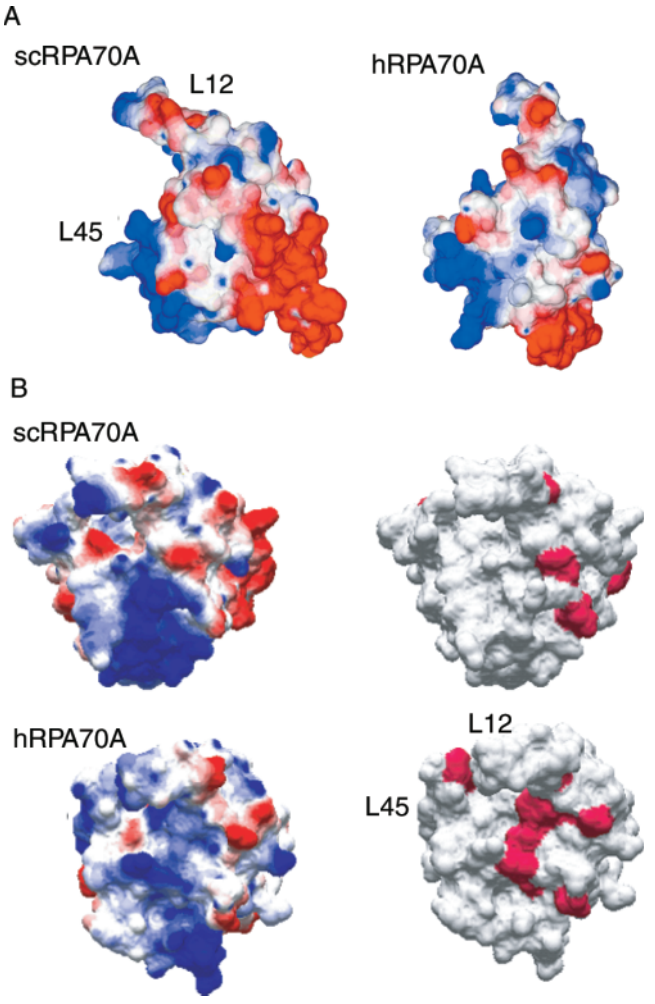

Figure 6. (**A**) Charge distributions of scRPA70A (left) and hRPA70A (right). The positive charges are shown in blue, and the negative charges are shown in red. (**B**) The charge distributions and molecular surfaces of RPA70A affected by T-ag OBD interaction. Residues perturbed by T-ag OBD binding are shown in red. hRPA70A (bottom), scRPA70A (top).

may result from indirect conformational changes that occur upon ssDNA binding. This hypothesis is consistent with the unexpected conformational changes in this part of the protein with ssDNA binding that were revealed in the RDC-based model structure. Taken together, our results suggest that the orientations of the residues involved in ssDNA binding are conserved in both hRPA70A and scRPA70A; however, the conformational changes that occur in the two proteins upon ssDNA binding are not identical.

Interaction with the SV40 T-ag OBD

Our structural data showed that the ssDNA-binding arcs of human and *S.cerevisiae* RPA70A are conserved. Moreover, the charge distributions in the binding arc were also similar. Both surfaces of the ssDNA-binding regions were basic, and these properties could permit the phosphates backbones of the DNA strands to interact with both domains in similar manners. However, Figure 6A shows that, except for the ssDNA-binding arc, the surfaces of the two proteins differ. The surface that consists of the α-helix and the C-terminal region (from β5′) in scRPA70A was highly acidic, while the equivalent part of hRPA70A is less acidic and relatively neutral. Also, the amino acid sequences in these regions are less homologous compared with other parts of the two proteins (53) (Figure 1). On the basis of this information, we hypothesized that (i) the SV40 T-ag binding surface of RPA70A does not overlap with the ssDNA-binding region and (ii) the different charge distributions displayed by the surfaces of the two proteins, which include the α helices and the C-terminal region, contribute to the ability of the proteins to discriminate among interaction partners.

To test this hypothesis, we performed chemical shift perturbation experiments on hRPA70A and scRPA70A in the presence of the SV40 T-ag OBD. We titrated unlabeled SV40 T-ag OBD into solutions of each of the ^{15}N-labeled RPA70A protein separately, starting with a molar ratio of 1:0 (labeled/unlabeled) and progressing to a ratio of 1:10, and collected spectra at each addition. The studied complexes were at saturation at the titration endpoint. The estimated K_d for hRPA70A was ∼200 μM and scRPA70A had a substantially lower binding affinity for the OBD than does hRPA70A (data not shown). In hRPA70A, residues G219, E240, N252, G260, T261, F269, V272, K273, T279 and E283 exhibited significant chemical shift perturbations, while only residues

G219, F222, N271, D281, T284 and D290 showed notable chemical shift changes in scRPA70A. The residues that showed substantial changes only in hRPA70A (E240, G260, T261 and T279) were located near each other and were more neutral than the same parts of scRPA70A (D240, A260, K261 and N279). We then mapped the SV40 T-ag OBD binding surface in each protein and found that it was separated from the ssDNA-binding region (Figure 6B). Even though a few residues in the L12 and L45 loops (G219 and F269 of hRPA70A) also showed substantial chemical shift changes, these alternations did not appear to result from direct protein–protein contacts. One reason for this conclusion is that it is hard to access to G219 directly from the outside of the protein. Because of the natural flexibilities of the loop regions described above, the observed chemical shift changes of the residues in these loops might result from indirect conformational changes in the protein due to T-ag binding.

From our results, we suggest that the higher T-ag binding affinity of hRPA70A, relative to scRPA70A, results from the different surface characters detected in the two proteins. Furthermore, we speculate that these properties account for the species-specificity observed in SV40 replication.

DISCUSSIONS

scRPA70A is one of the proteins most homologous to hRPA70A, and, thus, it has been assumed that the structure of scRPA70A will be very similar to that of hRPA70A. However, biochemical experiments have highlighted functional differences between the two proteins. These findings prompted us to determine the solution structure of scRPA70A, which could provide clues as to the origins of the functional dissimilarities. As expected, the overall structure of scRPA70A was quite similar to that of hRPA70A; the intricate details, however, revealed some important differences.

Because the hRPA70A structure was determined by X-ray crystallography, we must consider the possibility that the subtle structural distinctions between the two homologous proteins might result from the methodological differences. However, when we used chemical shift perturbations and RDC-based homology modeling to assess the conformational changes in the structures of both RPA70A proteins, differences between the *S.cerevisiae* and human proteins were observed. scRPA70A showed indirect conformational changes that did not occur in hRPA70A and exhibited a narrower ssDNA-binding arc than did hRPA70A. These findings imply that scRPA70A experiences more changes upon ssDNA binding than does hRPA70A. Considering that a cofolding event that involves the loop structures and single-stranded nucleic acids occurs in at least one of the binding partners concurrent with complex formation (53), it is intriguing to note that the binding affinity of scRPA70NAB for ssDNA is slightly higher than that of hRPA70NAB (~2-fold), which might be related to the additional conformational changes observed for scRPA70A upon binding to ssDNA (8).

OB folds, common feature of ssDNA-binding modules, have been studied in many proteins. For non-base-specific ssDNA interactions, OB folds retain flexibilities in its binding sites. hRPA70A, Pf3, T4gp32 and scRPA70A all show this common characteristic (51,54,55). The hydrogen bonding interaction between protein and ssDNA is also used in non-base-specific binding as well as in base-specific interaction. However, the flexible binding sites on the protein make it possible to adjust base-specific contacts along the different DNA sequences. The sequence-specific ssDNA-binding proteins, such as POT1, have much more extensive contact with ssDNA by using hydrogen bonds (56,57).

As far as OB fold as a protein interaction module is concerned, we found that surfaces formed by amino acid residues from the α-helix to the C-terminus of hRPA70A interacted with T-ag OBD. This binding surface is separated from the ssDNA-binding surface and from the Rad51 and XPA binding regions of hRPA70A, which overlap with the ssDNA-binding surface (4,5). These findings imply that the T-ag OBD will not compete with ssDNA binding to RPA70A, while Rad51 and XPA will (4,5). A previous biochemical study demonstrated that RPA bound to ssDNA showed a reduced affinity for T-ag binding, which suggests that ssDNA binding can modulate hRPA70A's ability to form protein–protein interactions (58). From this previous result and those described here, we suggest that the observed modulation of the hRPA70A–T-ag OBD interaction by ssDNA does not result from direct obstruction of the T-ag binding surface on hRPA70A.

In a similar vein, substantial chemical shift changes in amino acid residues in the L12 and L45 loops upon T-ag binding to hRPA70A raise the possibility that hRPA70A–T-ag OBD interaction modulate the ssDNA–RPA70A interaction by inducing indirect conformational changes in the DNA-binding region. The reciprocal regulation of ssDNA- and protein partner-binding implies that these two key biochemical functions of RPA70A should not be considered independently. An earlier study showed that the D228Y mutation in scRPA70A conferred a slow-growth phenotype and UV sensitivity on *S.cerevisiae*, and the mutant version of scRPA displayed reduced ssDNA-binding affinity (59). Because the mutated residue is not in the ssDNA-binding site, it is unlikely that the mutation inhibits ssDNA binding directly. One explanation is that the mutation affects the protein structure in such a way that it influences the ability of scRPA to interact with ssDNA.

Regarding the species-specific interaction between the T-ag OBD and hRPA70A, the main interacting forces between these two proteins are not electrostatic, as the T-ag binding surface of hRPA70A is relatively neutral. The equivalent region of scRPA70A has more charged residues, and we speculate that the differences between these two surfaces determine whether or not they can bind T-ag. A recent study has addressed the mechanisms at work behind the interaction between the C-terminal domain of the hRPA32 subunit (hRPA32C) and the T-ag OBD (60). In this interaction, the binding surface of hRPA32C has an acidic character, and the T-ag OBD provides a complementary basic surface. scRPA32C lacks the acidic residues necessary for binding to the T-ag OBD. From this recent study and our current results, we hypothesize that a weak, two-point interaction occurs between hRPA and T-ag at surfaces on hRPA that are not found in homologous proteins from other species.

From the comparison between the structures of hRPA70A and scRPA70A in this study, we suggest that evolutionary pressures may have specifically exploited, for specific protein–protein interactions, a region of the RPA70 protein distal to the ssDNA-binding region. Taken together, the results

of our study of scRPA70A presented here and previously published biochemical data provide insight into how similar protein structures can display distinct protein–protein interactions while retaining similar ssDNA-binding affinities.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Prof P. Bullock for the SV40 T-ag OBD expression plasmids and Dr R. Legerski for providing the scRPA expression vector. This work was supported by the National Creative Research Initiative Program to B.-S.C. from the Ministry of Science and Technology, Korea. C.-J.P. was supported partially by the BK21 project. Funding to pay the Open Access publication charges for this article was provided by National Creative Research Initiative Program from the Ministry of Science and Technology, Korea.

Conflict of interest statement. None declared.

REFERENCES

1. Wold,M.S. and Kelly,T. (1988) Purification and characterization of replication protein A, a cellular protein required for *in vitro* replication of simian virus 40 DNA. *Proc. Natl Acad. Sci. USA*, **85**, 2523–2527.
2. 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.
3. Fairman,M.P. and Stillman,B. (1988) Cellular factors required for multiple stages of SV40 DNA replication *in vitro*. *EMBO J.*, **7**, 1211–1218.
4. 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.
5. 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.
6. Lee,J.H., Park,C.J., Arunkumar,A.I., Chazin,W.J. and Choi,B.S. (2003) NMR study on the interaction between RPA and DNA decamer containing *cis*-syn cyclobutane pyrimidine dimer in the presence of XPA: implication for damage verification and strand-specific dual incision in nucleotide excision repair. *Nucleic Acids Res.*, **31**, 4747–4754.
7. Lin,Y.L., Chen,C., Keshav,K.F., Winchester,E. and Dutta,A. (1996) Dissection of functional domains of the human DNA replication protein complex replication protein A. *J. Biol. Chem.*, **271**, 17190–17198.
8. 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.
9. 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.
10. He,Z., Henricksen,L.A., Wold,M.S. and Ingles,C.J. (1995) RPA involvement in the damage-recognition and incision steps of nucleotide excision repair. *Nature*, **374**, 566–569.
11. Park,M.S., Ludwig,D.L., Stigger,E. and Lee,S.H. (1996) Physical interaction between human RAD52 and RPA is required for homologous recombination in mammalian cells. *J. Biol. Chem.*, **271**, 18996–19000.
12. Heyer,W.D., Rao,M.R., Erdile,L.F., Kelly,T.J. and Kolodner,R.D. (1990) An essential *Saccharomyces cerevisiae* single-stranded DNA binding protein is homologous to the large subunit of human RP-A. *EMBO J.*, **9**, 2321–2329.
13. 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.
14. 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.
15. Heyer,W.D., Rao,M.R., Erdile,L.F., Kelly,T.J. and Kolodner,R.D. (1990) An essential *Saccharomyces cerevisiae* single-stranded DNA binding protein is homologous to the large subunit of human RP-A. *EMBO J.*, **9**, 2321–2329.
16. Longhese,M.P., Neecke,H., Paciotti,V., Lucchini,G. and Plevani,P. (1996) The 70 kDa subunit of replication protein A is required for the G1/S and intra-S DNA damage checkpoints in budding yeast. *Nucleic Acids Res.*, **24**, 3533–3537.
17. Longhese,M.P., Plevani,P. and Lucchini,G. (1994) Replication factor A is required *in vivo* for DNA replication, repair, and recombination. *Mol. Cell. Biol.*, **14**, 7884–7890.
18. Umezu,K., Sugawara,N., Chen,C., Haber,J.E. and Kolodner,R.D. (1998) Genetic analysis of yeast RPA1 reveals its multiple functions in DNA metabolism. *Genetics*, **148**, 989–1005.
19. Soustelle,C., Vedel,M., Kolodner,R. and Nicolas,A. (2002) Replication protein A is required for meiotic recombination in *Saccharomyces cerevisiae*. *Genetics*, **161**, 535–547.
20. Chen,C., Umezu,K. and Kolodner,R.D. (1998) Chromosomal rearrangements occur in *S.cerevisiae* rfa1 mutator mutants due to mutagenic lesions processed by double-strand-break repair. *Mol. Cell*, **2**, 9–22.
21. Din,S., Brill,S.J., Fairman,M.P. and Stillman,B. (1990) Cell-cycle-regulated phosphorylation of DNA replication factor A from human and yeast cells. *Genes Dev.*, **4**, 968–977.
22. Melendy,T. and Stillman,B. (1993) An interaction between replication protein A and SV40 T antigen appears essential for primosome assembly during SV40 DNA replication. *J. Biol. Chem.*, **268**, 3389–3395.
23. 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.
24. Lee,S.H. and Kim,D.K. (1995) The role of the 34-kDa subunit of human replication protein A in simian virus 40 DNA replication *in vitro*. *J. Biol. Chem.*, **270**, 12801–12807.
25. Melendy,T. and Stillman,B. (1993) An interaction between replication protein A and SV40 T antigen appears essential for primosome assembly during SV40 DNA replication. *J. Biol. Chem.*, **268**, 3389–3395.
26. Matsumoto,T., Eki,T. and Hurwitz,J. (1990) Studies on the initiation and elongation reactions in the simian virus 40 DNA replication system. *Proc. Natl Acad. Sci. USA*, **87**, 9712–9716.
27. Iftode,C., Daniely,Y. and Borowiec,J.A. (1999) Replication protein A (RPA): the eukaryotic SSB. *Crit. Rev. Biochem. Mol. Biol.*, **34**, 141–180.
28. Wang,M., Park,J.S., Ishiai,M., Hurwitz,J. and Lee,S.H. (2000) Species specificity of human RPA in simian virus 40 DNA replication lies in T-antigen-dependent RNA primer synthesis. *Nucleic Acids Res.*, **28**, 4742–4749.
29. Weisshart,K., Taneja,P. and Fanning,E. (1998) The replication protein A binding site in simian virus 40 (SV40) T antigen and its role in the initial steps of SV40 DNA replication. *J. Virol.*, **72**, 9771–9781.
30. Luo,X., Sanford,D.G., Bullock,P.A. and Bachovchin,W.W. (1996) Solution structure of the origin DNA-binding domain of SV40 T-antigen. *Nature Struct. Biol.*, **3**, 1034–1039.
31. Braun,K.A., Lao,Y., He,Z., Ingles,C.J. and Wold,M.S. (1997) Role of protein–protein interactions in the function of replication protein A (RPA): RPA modulates the activity of DNA polymerase alpha by multiple mechanisms. *Biochemistry*, **36**, 8443–8454.
32. 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.
33. 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.
34. Delaglio,F., Grzesiek,S., Vuister,G.W., Zhu,G., Pfeifer,J. and Bax,A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
35. Cornilescu,G., Delaglio,F. and Bax,A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR*, **13**, 289–302.

36. Permi,P., Kilpelainen,I., Annila,A. and Heikkinen,S. (2000) Intensity modulated HSQC and HMQC: two simple methods to measure 3J(HNH)alpha in proteins. *J. Biomol. NMR*, **16**, 29–37.
37. Herrmann,T., Guntert,P. and Wuthrich,K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.*, **319**, 209–227.
38. Schwieters,C.D. and Clore,G.M. (2001) Internal coordinates for molecular dynamics and minimization in structure determination and refinement. *J. Magn. Reson.*, **152**, 288–302.
39. Schwieters,C.D., Kuszewski,J.J., Tjandra,N. and Marius Clore,G. (2003) The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.*, **160**, 65–73.
40. Kuszewski,J., Gronenborn,A.M. and Clore,G.M. (1997) Improvements and extensions in the conformational database potential for the refinement of NMR and X-ray structures of proteins and nucleic acids. *J. Magn. Reson.*, **125**, 171–177.
41. Laskowski,R.A., Rullmannn,J.A., MacArthur,M.W., Kaptein,R. and Thornton,J.M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR*, **8**, 477–486.
42. Koradi,R., Billeter,M. and Wuthrich,K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.*, **14**, 51–55, 29–32.
43. Guex,N. and Peitsch,M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714–2723.
44. Ottiger,M., Delaglio,F. and Bax,A. (1998) Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson.*, **131**, 373–378.
45. Yang,D., Venters,R.A., Mueller,G.A., Choy,W.Y. and Kay,L.E. (1999) TROSY-based HNCO pulse sequences for the measurement of 1HN-15N, 15N-13CO, 1HN-13CO, 13CO-13Ca and 1HN-13Ca dipolar couplings in 15N, 13C, 2H-labeled proteins. *J. Biomol. NMR*, **14**, 333–343.
46. Bewley,C.A., Gustafson,K.R., Boyd,M.R., Covell,D.G., Bax,A., Clore,G.M. and Gronenborn,A.M. (1998) Solution structure of cyanovirin-N, a potent HIV-inactivating protein. *Nature Struct. Biol.*, **5**, 571–578.
47. Chou,J.J., Li,S. and Bax,A. (2000) Study of conformational rearrangement and refinement of structural homology models by the use of heteronuclear dipolar couplings. *J. Biomol. NMR*, **18**, 217–227.
48. Kuszewski,J., Gronenborn,A.M. and Clore,G.M. (1999) Improving the packing and accuracy of NMR structures with a pseudopotential for the radius of gyration. *J. Am. Chem. Soc.*, **121**, 2337–2338.
49. Rooney,L.M. and Sachchidanand, Werner,J.M. (2004) Characterizing domain interfaces by NMR. *Methods Mol. Biol.*, **278**, 123–138.
50. Clore,G.M., Gronenborn,A.M. and Bax,A. (1998) A robust method for determining the magnitude of the fully asymmetric alignment tensor of oriented macromolecules in the absence of structural information. *J. Magn. Reson.*, **133**, 216–221.
51. Bhattacharya,S., Botuyan,M.V., Hsu,F., Shan,X., Arunkumar,A.I., Arrowsmith,C.H., Edwards,A.M. and Chazin,W.J. (2002) Characterization of binding-induced changes in dynamics suggests a model for sequence-nonspecific binding of ssDNA by replication protein A. *Protein Sci.*, **11**, 2316–2325.
52. 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.
53. Theobald,D.L., Mitton-Fry,R.M. and Wuttke,D.S. (2003) Nucleic acid recognition by OB-fold proteins. *Annu. Rev. Biophys. Biomol. Struct.*, **32**, 115–133.
54. Horstink,L.M., Abseher,R., Nilges,M. and Hilbers,C.W. (1999) Functionally important correlated motions in the single-stranded DNA-binding protein encoded by filamentous phage Pf3. *J. Mol. Biol.*, **287**, 569–577.
55. Shamoo,Y., Friedman,A.M., Parsons,M.R., Konigsberg,W.H. and Steitz,T.A. (1995) Crystal structure of a replication fork single-stranded DNA binding protein (T4 gp32) complexed to DNA. *Nature*, **376**, 362–366.
56. Lei,M., Podell,E.R. and Cech,T.R. (2004) Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. *Nature Struct. Mol. Biol.*, **11**, 1223–1229.
57. Horvath,M.P., Schweiker,V.L., Bevilacqua,J.M., Ruggles,J.A. and Schultz,S.C. (1998) Crystal structure of the Oxytricha nova telomere end binding protein complexed with single strand DNA. *Cell*, **95**, 963–974.
58. Loo,Y.M. and Melendy,T. (2004) Recruitment of replication protein A by the papillomavirus E1 protein and modulation by single-stranded DNA. *J. Virol.*, **78**, 1605–1615.
59. Smith,J. and Rothstein,R. (1995) A mutation in the gene encoding the *Saccharomyces cerevisiae* single-stranded DNA-binding protein Rfa1 stimulates a RAD52-independent pathway for direct-repeat recombination. *Mol. Cell. Biol.*, **15**, 1632–1641.
60. 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.
61. Notredame,C., Higgins,D.G. and Heringa,J. (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.*, **302**, 205–217.