The EMBO Journal Vol.18 No.16 pp.4498–4504, 1999

The crystal structure of the complex of replication protein A subunits RPA32 and RPA14 reveals a mechanism for single-stranded DNA binding

Alexey Bochkarev[1,2], Elena Bochkareva[2,3], Lori Frappier[1,4] and Aled M.Edwards[3,4,5]

[1]Department of Medical Genetics and Microbiology, University of Toronto, 1 King's College Circle, Toronto, Ontario M5S 1A8, [3]Department of Medical Biophysics, University of Toronto, Ontario Cancer Institute, 610 University Avenue, Toronto, Ontario M5G 2M9 and [5]Banting and Best Department of Medical Research, C.H.Best Institute, 112 College Street, University of Toronto, Toronto, Ontario M5G 1L6, Canada

[2]Present address: Department of Biochemistry and Molecular Biology, University of Oklahoma Health Sciences Center, PO Box 26901, BMSB 853, Oklahoma City, OK 73190, USA

[4]Corresponding authors
e-mail: lori.frappier@utoronto.ca

Replication protein A (RPA), the eukaryote single-stranded DNA-binding protein (SSB), is a heterotrimer. The largest subunit, RPA70, which harbours the major DNA-binding activity, has two DNA-binding domains that each adopt an OB-fold. The complex of the two smaller subunits, RPA32 and RPA14, has weak DNA-binding activity but the mechanism of DNA binding is unknown. We have determined the crystal structure of the proteolytic core of RPA32 and RPA14, which consists of the central two-thirds of RPA32 and the entire RPA14 subunit. The structure revealed that RPA14 and the central part of RPA32 are structural homologues. Each subunit contains a central OB-fold domain, which also resembles the DNA-binding domains in RPA70; an N-terminal extension that interacts with the central OB-fold domain; and a C-terminal helix that mediate heterodimerization via a helix–helix interaction. The OB-fold of RPA32, but not RPA14, possesses additional similarity to the RPA70 DNA-binding domains, supporting a DNA-binding role for RPA32. The discovery of a third and fourth OB-fold in RPA suggests that the quaternary structure of SSBs, which in Bacteria and Archaea are also tetramers of OB-folds, is conserved in evolution. The structure also suggests a mechanism for RPA trimer formation.

Keywords: crystal structure/OB-fold/replication protein A/single-stranded DNA

Introduction

Replication protein A (RPA), the nuclear single-stranded DNA (ssDNA)-binding protein (SSB), is a core component of the DNA replication, repair and recombination machineries (Wold, 1997). Within these complexes, RPA is thought to play two roles: to bind the ssDNA thereby stabilizing the unwound DNA, and to facilitate the assembly of the complex through direct protein–protein interactions. RPA is composed of three subunits (RPA70, RPA32 and RPA14) each of which is conserved in all eukaryotic cells (Brill and Stillman, 1991). The DNA-binding activity of RPA is essential for its function. The high-affinity DNA-binding domain is located within the central portion of the RPA70 subunit (Gomes and Wold, 1995, 1996; Pfuetzner *et al*., 1997), which is composed of two structurally homologous DNA-binding domains (Bochkarev *et al*., 1997). Each of these domains interacts with four nucleotides of DNA in a defined polarity with respect to the 5′ and 3′ ends of the DNA. The RPA70 ssDNA-binding domains share a common fold, known as the OB (oligonucleotide–oligosaccharide-binding) fold, with many other ssDNA- and RNA-binding proteins (Murzin, 1993; Bycroft *et al*., 1997). These structural comparisons revealed that the mechanism of single-stranded nucleic acid binding appears to be conserved from bacteriophage to man.

The complete ssDNA-binding mechanism of RPA, which originally was thought to be effected by the two domains in the RPA70 subunit, now appears to be more complex. Five observations suggest that that the RPA heterotrimer contains additional sites of DNA binding. First, the RPA heterotrimer has 10- to 50-fold higher affinity for ssDNA than do the two RPA70 domains alone (Kim *et al*., 1994; Pfuetzner *et al*., 1997). Secondly, under certain conditions, the heterotrimer can bind to ssDNA with an occluded binding site of 30 nucleotides (Kim *et al*., 1994; Kim and Wold, 1995), far larger than the eight nucleotide binding site that would have been predicted solely from the crystal structure of the RPA70 domains (Bochkarev *et al*., 1997). Thirdly, the purified complex of RPA14 and RPA32 (RPA14/32) as well as the C-terminal region of RPA70, contains weak DNA-binding activity (Bochkareva *et al*., 1998; Brill and Bastin-Shanower, 1998; Sibenaller *et al*., 1998). Fourthly, in the context of the heterotrimer, RPA32 can be cross-linked to DNA (Philipova *et al*., 1996; Lavrik *et al*., 1998; Mass *et al*., 1998). Finally, sequence and secondary structure comparisons revealed that the central part of RPA32, and possibly RPA14, are similar to the ssDNA-binding domains of RPA70 (Philipova *et al*., 1996).

The domain structure of RPA14/32 has been determined by genetic and biochemical analysis. RPA32 comprises three domains: a phosphorylated N-terminal domain (Brush *et al*., 1994, 1996), a central domain predicted to contain the ssDNA-binding activity (Philipova *et al*., 1996; Bochkareva *et al*., 1998; Sibenaller *et al*., 1998) and a C-terminal protein interaction domain (Lee *et al*., 1995; Nagelhus *et al*., 1997). The central domain of RPA32 is sufficient to interact with RPA14 and is the only part of RPA32 required for viability in yeast (Philipova *et al*., 1996; Bochkareva *et al*., 1998). The N-terminal domain, which comprises ~40 residues, is phosphorylated in a cell cycle-dependent fashion and is probably a substrate

for the phosphatidyl inositol 3′ kinase family, which includes DNA-PK and the ATM kinase (Brush *et al*., 1994, 1996). The C-terminal domain, which comprises ~100 residues, mediates the interaction of RPA with the xeroderma pigmentosum-A protein (XPA) (Lee *et al*., 1995) (A.M.Edwards, unpublished). The domain structure of RPA14 is less clear; genetic studies revealed that only the C-terminal half of the subunit is required for viability in yeast (Philipova *et al*., 1996), yet biochemical studies have shown that the subunit is highly resistant to proteolysis (Bochkareva *et al*., 1998), suggesting that it is composed of a single structural domain.

The protease-resistant core of the RPA14/32, which comprises the central domain of RPA32 and the entire RPA14 subunit, displays micromolar affinity for ssDNA (Bochkareva *et al*., 1998; Sibenaller *et al*., 1998). Interestingly, in the human system, the presence of either the N- or the C-terminal domain of RPA32 in the RPA14/32 core dramatically reduces ssDNA binding (Bochkareva *et al*., 1998; E.Bochkareva, unpublished), suggesting that the DNA binding is subject to negative regulation. As a step toward understanding the function and regulation of the complex of the smaller two subunits, we initiated a structural analysis of the RPA14/32 core complex.

Results and discussion

Structure determination

The core of the complex of RPA32 and RPA14 (RPA14/$32_{43–171}$) was defined by limited proteolysis using trypsin (Bochkareva *et al*., 1998). The resulting fragments were co-expressed in *Escherichia coli* and the complex purified to homogeneity. The RPA14/$32_{43–171}$ complex crystallized in the $P2_12_12_1$ space group with two dimers in the asymmetric unit. The structure was solved from native and selenomethionine-containing crystals using a combination of multiwavelength anomalous dispersion phasing and non-crystallographic averaging. All residues in the interval of 3–116 (of total 1–121) from RPA14 and of 43–171 from $RPA32_{43–171}$ were traced in the final structure except for residues 110–117 in RPA32 (part of the loop between β-strands 3 and 4). The final 2.5 Å structure was refined to an *R*-factor of 21.2% and a free *R*-factor of 30.3% (Table I).

Table I. Refinement statistics

Experimental phase calculation (3.0 Å)	
No. of Se atoms	17
FOM	0.72
Structure refinement	
Resolution (Å)	8.0–2.5
No. of reflections with $F > 2\sigma$	20084
R/R_{free} (%)	21.2/30.3
No. of atoms (3668)	
Water molecules	120
Model r.m.s. deviation from ideality	
Bonds (Å)	0.012
Angles (°)	1.75

General description of the RPA14/$32_{43–171}$ dimer

The structure of the RPA14/$32_{43–171}$ dimer is shown in Figure 1. RPA14 and the core of RPA32 ($RPA32_{43–171}$) are structural homologues; each comprises an OB-fold flanked by a conserved extension at the N-terminus and an α-helix at the C-terminus (Figures 1 and 2). The dimerization interface is composed primarily of hydrophobic contacts between the two C-terminal helices. The helices are apposed in parallel and at an angle of ~25°. The extensive network of hydrophobic dimerization interactions includes residues F94, L98, Y99, A102 I105, I106, F109, F112 and Y113 from RPA14 and residues M155, I159, L160, V162, I163 and M167 from RPA32. Other contacts between the two subunits include an interaction between the N-terminal extension of RPA14 (amino acids 8–10) and a segment of RPA32 comprising amino acids 97–100, which connects the second and third β-strands (the L23 loop). The importance of this latter interaction in complex formation is not clear as the N-terminal half of yeast RPA14 can be deleted without loss of viability in yeast (Philipova *et al*., 1996). There is little interaction between the OB-fold cores of each subunit, suggesting that the two domains might adopt alternative relative orientations.

Conservation of structure among RPA14, RPA32 and RPA70

RPA14 and RPA32 each contain an OB-fold similar to those observed in other single-stranded nucleic acid-binding proteins, including RPA70 (Figure 2). The N-terminal region is also conserved in the two DNA-binding domains of RPA70 (domains A and B). The C-terminal helix is conserved in only one of the two RPA70 domains, domain B [RPA70(b), Figure 2].

This similarity in structure of RPA14 and RPA32 to the RPA70 ssDNA-binding domains suggests that the RPA32 and RPA14 proteins may possess a similar structural mechanism for DNA binding. In support of this hypothesis, RPA32 has additional hallmarks of a ssDNA-binding domain, including two conserved aromatic residues, which in the RPA70 domains were shown to stack with the DNA bases (Figures 2 and 3).

The conservation of the C-terminal helix in RPA14 and RPA32 and its structural similarity to domain B of RPA70 suggests that the helices are functionally important components of RPA (Bochkarev *et al*., 1997). The importance of this helix in the structure of domain B from RPA70 had not been appreciated in our previous studies for several reasons. First, the helix is not conserved in domain A of RPA70, which is almost identical in structure and function to the core of domain B. Secondly, similarly placed helices were not apparent in any other SSB whose structure had been solved. Thirdly, the helix did not make extensive contacts with the OB-fold in domain B; it had been thought possible that the apposition of the helix and the domain B OB-fold could simply have been a crystal packing effect. Finally, the helix does not make contact with the DNA. However, in the light of the new observation that the helix is conserved in three of the four OB-folds in RPA, and that structural conservation underlies function, these helices must play a conserved role in RPA function. Nevertheless, it is important to note that although the position of the C-terminal helices is conserved, the relative orientation of the C-terminal helix in domain B of RPA70 is opposite to that of the helices in both RPA32 and RPA14.

The N-terminal extensions of RPA14, RPA32 and each RPA70 domain, which also are conserved in structure,

Fig. 1. Structure of the RPA32/RPA14 dimer. A stereo ribbon representation of the structures of RPA14 and $RPA32_{45–170}$ in green and yellow, respectively. The N- and C-termini of each fragment are designated.

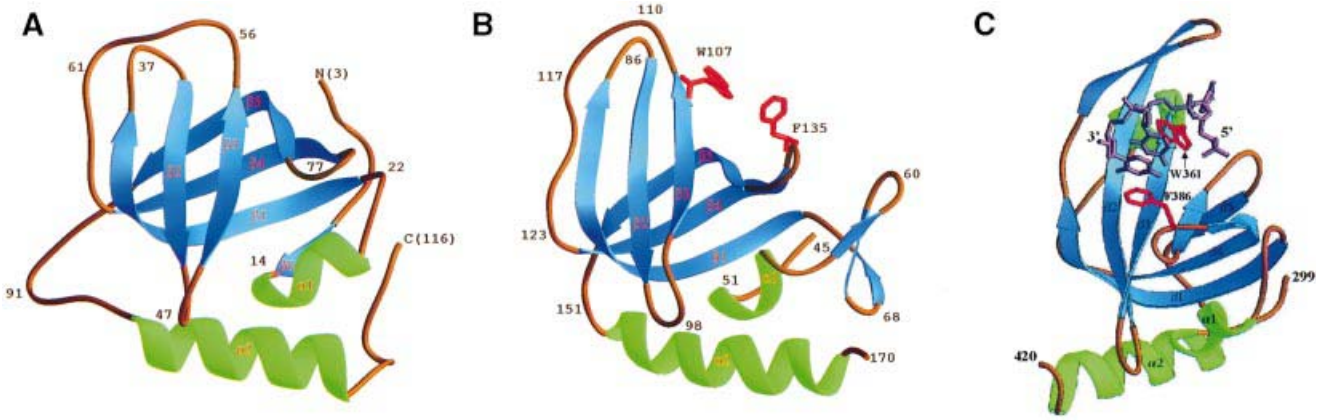

Fig. 2. Comparison of the structures of $RPA14_{3–116}$, $RPA32_{45–170}$ and $RPA70(b)_{299–420}$. An alignment of three OB-folds in $RPA14_{3–116}$ (**A**), $RPA32_{45–170}$ (**B**) and $RPA70(b)_{299–420}$ (**C**), highlighting the conserved C-terminal α-helices and the N-terminal extensions. The β-strands are coloured in blue, α-helices in green and loops in gold. The secondary structure elements are named according to the convention previously described for OB-folds (Murzin, 1993; Bochkarev *et al*., 1997; Bycroft *et al*., 1997). Amino acid positions are indicated by the numbers, and the N- and C-termini are designated. For $RPA32_{45–170}$ and $RPA70(b)_{299–420}$, residues that are either predicted or have been shown to stack with the DNA bases are indicated and coloured in red. For $RPA70_{299–420}$, whose structure was determined in the presence of oligodeoxycytosine, the crystallographically determined structure of three bases is shown in purple. The 5′ and 3′ ends of the ssDNA are indicated.

have two characteristics in common. First, they contain either a short β-strand or an extended chain that interacts with the first β-strand of the OB-fold. Secondly, they contain short helical segments that interact with the loops connecting the second and third (L23) and the fourth and fifth (L45) β-strands of the OB-folds. For example, in RPA32, there are hydrogen bond contacts between residues D96 (L23) and K139 (L45) and the peptide backbone of I51 and A57 in the N-terminus (Figure 4). In RPA14, residue D45 in the L23 loop forms a hydrogen bond with A14 in the N-terminal extension, and there are extensive hydrophobic interactions between the L45 loop and the N-terminus. The L45 loop plays an important role in ssDNA binding; it clamps down on the ssDNA, forming one of the walls of the DNA-binding groove, and supplies an aromatic residue that stacks with the DNA bases (Figure 2). The strong interaction of the N-terminal extensions with this loop suggests that the N-terminal region may regulate DNA binding.

Despite the global similarities between the structures of RPA14, RPA32 and the RPA70 domains, there are also significant differences. Some of these differences are likely to be due to the fact that the structure of the RPA70 domains, but not the RPA14/32 structure, was determined as a complex with ssDNA. Other differences are probably functionally important.

One major difference between the RPA14/32 and RPA70 structures is in the position of the L45 loop. This difference probably reflects differences between the bound and free OB-folds. Structures of several OB-fold nucleic acid-binding proteins have been determined; many were solved in the absence of nucleic acid, and two, RPA70 and aspartyl tRNA-synthetase, contained a well-defined nucleic acid (Ruff *et al*., 1991; Bochkarev *et al*., 1997). To date, the

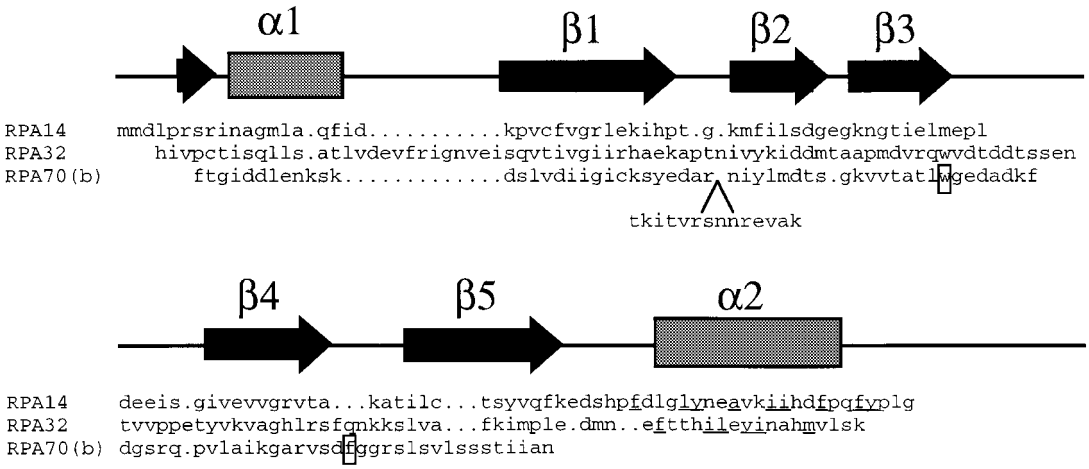

Fig. 3. Structure-based sequence alignment of RPA14, RPA32 and RPA70(b). The sequences of RPA14 (residues 4–116), RPA32 (residues 45–171) and RPA70 (residues 302–402) were aligned based on the three-dimensional structure. The positions of the secondary structure elements are shown above, with arrows representing β-strands and rectangles representing α-helices. Residues that mediate the dimerization of RPA32 and RPA14 are underlined. Residues in RPA70 known to interact with the DNA are boxed.

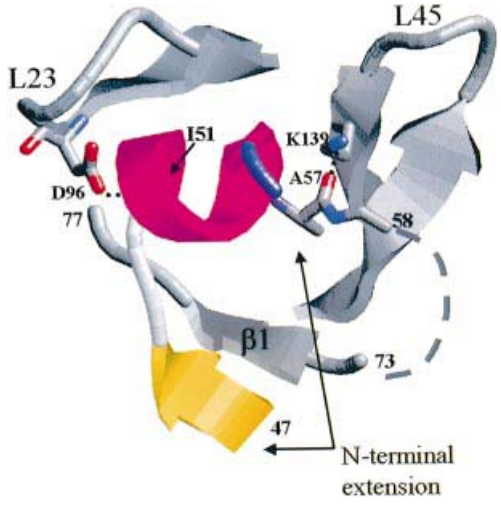

Fig. 4. Interaction of the N-terminal part of the RPA32 fragment with the RPA32 OB-fold. The structure of the N-terminal part of the RPA32 fragment whose structure was determined, and that corresponds to residues 47–77, is shown. Residues 58–73 were removed for the sake of clarity. The β-strand and the α-helix in the N-terminal part of RPA32 are shown in yellow and pink, respectively. Amino acids highlighted include residues I51 and A57, which form hydrogen bonds with D96 and K139 in the L23 and L45 loops, respectively.

bound and free OB-fold structures have differed mainly in the relative positions of the L45 loop. In all of the structures determined in the absence of DNA, the loop extends away from the core of the OB-fold (open position). In the RPA70 and aspartyl-tRNA synthetase structures, the loop folds back to stabilize the bound nucleic acid (closed position). Accordingly, in both RPA14 and RPA32 subunits, the L45 loops do not fold back onto the β-sheet but rather are found in a slightly extended position. The correlation between the position of the L45 loop and the DNA-binding state strengthens the hypothesis that the 'clamping' of the L45 loop is a critical step in the process of ssDNA binding, and suggests also that the position of the loops in RPA14 and RPA32 will change upon DNA binding.

The size and character of the loop that connects the first and second β-strand are also different between RPA14, RPA32 and the RPA70 domains. In the RPA70 DNA-binding domains, this loop, which is quite long, wraps over the ssDNA and makes extensive contacts with the phosphate backbone. In both the RPA14 and RPA32 structures, the L12 loop is significantly shorter than those found in RPA70. Thus the RPA14 and RPA32 loops would be predicted to support fewer hydrogen-bonding interactions. The inability of RPA14 and RPA32 to make this type of interaction might explain the relatively weak affinity of RPA14/32 for ssDNA, as compared with the RPA70 domains.

Possible mechanism of DNA binding

The RPA heterotrimer interacts with ssDNA in at least two modes that are thought to occur sequentially. In the first mode, there is an 8–10 nucleotide interaction site that has high affinity but is relatively unstable compared with the second mode of interaction, which is stable and occludes 30 nucleotides (Kim *et al*., 1994; Kim and Wold, 1995; Bochkarev *et al*., 1997). The transition from the first to the second mode is accompanied by a conformational change that can be detected biochemically and by electron microscopy (Blackwell *et al*., 1996).

The new structural data for RPA suggest a structural mechanism for these multiple binding modes. For two reasons, it is likely that the initial interaction of RPA with DNA is through the RPA70 subunit. First, RPA70 contains the highest affinity for ssDNA (Gomes and Wold, 1995, 1996). Secondly, the central fragment of RPA70, which is composed of two OB-folds, forms an eight nucleotide binding site (Bochkarev *et al*., 1997). We now suggest that the transition from the eight to the 30 nucleotide binding mode is associated in part with the filling of the RPA32 or RPA14 OB-folds. It is also possible that the RPA70 C-terminal domain of RPA70, which has been postulated to contain an OB-fold (Brill and Bastin-Shanower, 1998), may play a role in this structural transition. Three lines of experimental data suggest that RPA32 interacts with DNA, and probably in the same manner as does RPA70. First, the complex of RPA32 and

RPA14 has ssDNA-binding activity (Bochkareva *et al*., 1998; Sibenaller *et al*., 1998). Secondly, RPA32 can be cross-linked to DNA (Philipova *et al*., 1996). Thirdly, we have found that RPA32 is structurally very similar to the RPA70 DNA-binding domain and possesses the key conserved aromatic residues which, in RPA70, stack with the DNA bases.

A role for RPA14 in ssDNA binding is less clear. Although our structural data suggest that RPA14 is conserved in structure with other SSBs, unlike for RPA32, there is no experimental evidence supporting an RPA14–DNA interaction. RPA14 has not been shown to cross-link to DNA (Philipova *et al*., 1996). Furthermore, RPA14 contains neither the conserved aromatic residues that are important for stacking (RPA14 contains a methionine and a lysine residue in their place), nor a basic loop that connects the first and second β-strands, which would be predicted to fold over and interact with the phosphate backbone. Finally, the N-terminal half of RPA14, which includes a large part of the OB-fold, is not essential for yeast cell viability (Philipova *et al*., 1996). In the context of these structural and genetic observations, it is difficult to reconcile an essential DNA-binding role for RPA14.

Another possible role for the RPA14—a protein-interacting domain—was suggested by the recently reported structure of the heterodimeric SSB protein that binds to telomeric DNA (Horvath *et al*., 1998). Like RPA, the telomere-binding SSB is also composed of four OB-folds. As revealed by the crystal structure, three are intimately involved in binding DNA; the fourth interacts with an extended peptide from another subunit of the heterodimer. The relevance of the interaction of a peptide with an OB-fold remains to be established. This interaction may regulate DNA binding by competition, may be an artefact of crystallization or may represent a completely different mode of OB-fold function.

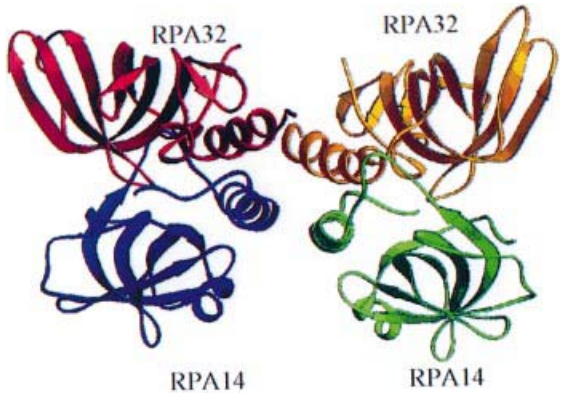

Fig. 5. Structure of crystallographic dimer of RPA14/$32_{45–170}$. A ribbon representation of the crystallographic dimer of RPA14/$32_{45–170}$ that highlights the four-helix bundle that maintains the crystallographic dimer. The RPA14 subunits are shown in blue and green and the RPA32 subunits in red and yellow. The crystallographic dimer is viewed along an axis perpendicular to the non-crystallographic axis of symmetry.

If RPA14 is not important for ssDNA binding, the additional ssDNA-binding domain may reside in the C-terminal domain (CTD) of RPA70 (extending from residues 420 to 617). Three lines of evidence support a ssDNA-binding role for this part of RPA. First, the co-expression of this part of RPA70 with the RPA14/32 complex (to create a heterotrimer of RPA14/32 and RPA70-CTD) increased the DNA-binding affinity of RPA14/32 5-fold (Bochkareva *et al*., 1998). Secondly, mutation of the RPA70-CTD decreases the ssDNA-binding affinity of the corresponding mutant RPA heterotrimeric complex (Walther *et al*., 1999). Thirdly, the purified RPA70 C-terminus has weak, but significant, ssDNA-binding activity (Brill and Bastin-Shanower, 1998).

Regulation of DNA binding

The DNA-binding activity of the human RPA14/32 complex is inhibited by either the N-terminal 40 residues of RPA32 or the C-terminal 100 residues of RPA32 (Bochkareva *et al*., 1998). The mechanism of inhibition is unknown, and the inhibitory fragments were not present in the RPA14/$32_{43–171}$ complex whose structure was determined. However, the structure revealed a possible mechanism by which these parts of the molecule could modulate DNA binding. The N- and C-termini of the core of RPA32, from which the inhibitory domains would initiate, are closely apposed in the structure and are also near the open L45 loop. Since the L45 loop is predicted to be a significant contributor to DNA binding, it is possible that the RPA32 inhibitory domains affect the mobility of this loop and, therefore, the DNA-binding activity of this subunit.

Model for heterotrimer formation

The RPA14/$32_{43–171}$ interacts with the C-terminal domain of the RPA70 subunit; co-expressing these three fragments is sufficient for trimerization (Bochkareva *et al*., 1998).

Table II. Data collection

Experiment	MAD1				MAD2				Native
Unit cell									
a (Å)	62.4				65.4				65.7
b (Å)	77.5				75.7				76.6
c (Å)	115.0				119.0				119.4
Data set	1	2	3	4	1	2	3	4	1
Wavelength (Å)	0.9792	0.9794	0.964	0.985	0.9792	0.9794	0.964	0.985	1.54
Anomalous or isomorphous	ano	ano	ano	iso	ano	ano	ano	iso	iso
Resolution	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.4
No. of reflections measured	88440	81990	110558	97485	92387	96029	104670	103996	123441
No. of independent reflections	34463	32711	36360	18741	36212	36194	37344	20353	23882
R-factor (overall/outer shell)	6.3/21.7	6.9/15.2	6.8/19.8	7.6/22.1	5.9/30.2	5.4/23.7	4.9/21.8	5.8/26.2	4.4/16.2
I/σ	17.1/6.4	12.3/5.3	15.7/5.8	17.0/7.9	15.5/2.6	16.0/3.0	19.2/3.7	24.7/4.3	34.4/7.0
Completeness	93.7/84.7	88.2/83.9	98.3/95.1	94.4/96.0	92.8/83.5	92.8/83.7	95.7/89.6	97.6/95.0	98.6/96.9

The structure of the C-terminal region of RPA70 is unknown, but the structure of the crystallographic dimer of RPA14/$32_{43–171}$ (Figure 5) suggests a mechanism for heterotrimer formation. The RPA14/$32_{43–171}$ dimer is stabilized largely through interactions between the two parallel C-terminal helices. In the crystal, this helix duo interacts in antiparallel fashion with two helices from another heterodimer, forming a four-helix packing arrangement that stabilizes the dimer of RPA14/$32_{43–171}$. This four-helix bundle formed by RPA14/$32_{43–171}$ may mimic a similar structure in the heterotrimer, and suggests that RPA70 might contribute one or two helices to the trimerization interface. Indeed, secondary structure predictions of RPA70 suggest that the topology of the C-terminal region is consistent with an OB-fold and, moreover, that the extreme C-terminus has helical character (Brill and Bastin-Shanower, 1998).

Materials and methods

Plasmid construction and protein purification

The complex of hexahistidine-tagged human RPA14 and $RPA32_{43–171}$ was expressed and purified as previously described (Bochkareva *et al*., 1998). To incorporate selenomethionine into the complex, BL21(λDE3) cells were grown in a methionine-depleted medium as previously described (Van Duyne *et al*., 1993). Digestion of the complex with thrombin resulted in the inclusion of four residues, GSHM, at the N-terminus of hRPA14. The protein was concentrated using a Centricon ultrafiltration membrane and stored frozen at 15 mg/ml in a solution containing 1 mM HEPES (pH 7.5), 50 mM NaCl, 10 mM dithiothreitol (DTT).

Crystallization and data collection

Crystals of RPA14/$32_{43–171}$ were grown by vapour diffusion in hanging drops. The drops were formed by mixing an equal volume of the protein solution with a reservoir solution that contained 0.1 M MES (pH 6.5), 0.75 M ammonium sulfate, 20% PEG 8K and 10 mM DTT. Crystals, with space group $P2_12_12_1$, formed within hours and typically grew to 0.2×0.2×0.5 mm in a few days. The crystals diffracted to 2.3 Å and contained two heterodimers per asymmetric unit. For low temperature data collection, the crystals were flash-transferred to a cryoprotectant (reservoir solution with 35% PEG 8K) for a few seconds, mounted in a 20 mm cryo loop (Hampton Research) and frozen in a stream of N_2 gas. Native data were collected from a crystal maintained at –175°C on an R-axis IV area detector with a rotating anode X-ray source and Cu Kα radiation. The native diffraction data were integrated with DENZO and the intensities scaled with SCALEPACK (Otwinowski and Minor, 1997). Two multiwavelength diffraction data sets (MAD1 and MAD2) from two different selenomethionine-containing crystals were collected at CHESS (beam line F2, ADSC Quantum-4K CCD detector). The images were processed with Mosfilm and scaled with Scala. Later, MAD experiments were reprocessed with a version of DENZO/SCALEPACK that supports ADSC Quantum-4 data format. The two crystals, which possessed slightly different unit cell parameters (see Table II), were ordered to 2.5 Å. The native crystal parameters were close to that observed in the MAD2 experiment.

Phasing and refinement

The diffraction data were phased using a combination of MAD and non-crystallographic (NC) averaging. Twenty potential selenium sites were identified by analysing 4 Å resolution MAD1 anomalous differences collected at the peak (maximum of f'') of the Se anomalous spectrum. The selenium sites were found by direct methods using the SnB ('shake and bake') program (Miller *et al*., 1994). The positions of 10 of these sites were confirmed by uncovering five pairwise relationships with axes of NC symmetry; five Se were superimposed with another five with an r.m.s. deviation of 0.5 Å. The 10 pairwise-related selenium sites were used to calculate initial MAD phases (Furey and Swaminathan, 1997), from which four more pairwise and three more unpaired selenium sites were identified by difference Fourier map analysis. The 17 selenomethionine sites (FOM 0.72) were used to calculate 3.0 Å phases, which were improved by solvent flattening (Furey and Swaminathan, 1997). At this point, the electron density map was readily interpretable. Seventy percent of the model was built into this map using the program O (Jones *et al*., 1991). The position of the NC operator was calculated based on seven pairwise Se relationships and then was refined using the experimental map. The models of the two crystallographically independent dimers were then used to build two molecular masks for NC averaging.

The low resolution MAD1 phases were then used to locate the selenium atoms in the MAD2 experimental data. The MAD2 Se coordinates were refined and the 3 Å map calculated. The quality of this map was improved over the MAD1 experimental map, but several features of the molecule were absent. Cross-map averaging was applied to the MAD1 and MAD2 data sets. In the final map, there was continuous density for the entire RPA14 subunit and for most of the RPA32 subunit. The L34 loops of both RPA32 subunits (amino acids 110–117) in both MAD1 and MAD2 experiments were disordered and not included in the final model.

The model was refined against the 2.5 Å native data using XPLOR (Brünger, 1992). Strongly bound water molecules were then incorporated. The refined model, which contains 3668 non-hydrogen atoms and 120 water molecules, has an r.m.s. deviation from ideality of 0.012 Å and 1.75° for bond distances and angles, respectively. The final *R*-factor is 21.2% for a 20 084 reflections with $F > 2\sigma$ between 8 and 2.5 Å. A Ramachandran plot shows no violations for either monomer, with >83% of all residues in the most favoured regions. A model quality assessment made with the PROCHECK program indicated that all the parameters are within or better than estimated limits for 2.5 Å resolution structures in the Protein Data Bank.

Acknowledgements

We would like to thank the staff at CHESS for their assistance during data collection. A.B. is grateful to organizers of the 1997 'EMBO practical course of MAD data collection and analysis'. This research was supported by grants to L.F. and A.M.E from the National Cancer Institute of Canada (NCIC), which receives funds from the Canadian Cancer Society, and the Medical Research Council of Canada (MRC), and to A.B. from the Presbyterian Health Foundation. A.M.E. and L.F. hold MRC Scientist Awards. A.B. was supported by an NCIC post-doctoral fellowship.

References

Blackwell,L.J., Borowiec,J.A. and Mastrangelo,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.

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.

Bochkareva,E., Frappier,L., Edwards,A.M. and Bochkarev,A. (1998) The RPA32 subunit of human replication protein A contains a single-stranded DNA-binding domain. *J. Biol. Chem.*, **273**, 3932–3936.

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.

Brill,S.J. and Stillman,B. (1991) Replication factor-A from *Saccharomyces cerevisiae* is encoded by three essential genes coordinately expressed at S phase. *Genes Dev.*, **5**, 1589–1600.

Brünger,A.T. (1992) *X-PLOR, Version 3.1: A System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT.

Brush,G.S., Anderson,C.W. and Kelly,T.J. (1994) The DNA-activated protein kinase is required for the phosphorylation of replication protein A during simian virus 40 DNA replication. *Proc. Natl Acad. Sci. USA*, **91**, 12520–12524.

Brush,G.S., Morrow,D.M., Hieter,P. and Kelly,T.J. (1996) The ATM homologue MEC1 is required for phosphorylation of replication protein A in yeast. *Proc. Natl Acad. Sci. USA*, **93**, 15075–15080.

Bycroft,M., Hubbard,T.J.P., Proctor,M., Freund,S.M.V. and Murzin,A.G. (1997) The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. *Cell*, **88**, 235–242.

Furey,W. and Swaminathan,S. (1997) PHASES-95: A program package for processing and analyzing diffraction data from macromolecules. In Sweet,C.C.R. (ed.), *Methods in Enzymology: Macromolecular Crystallography*. Academic Press, Orlando, FL, Vol. 277, pp. 590–620.

Gomes,X.V. and Wold,M.S. (1995) Structural analysis of human replication protein A. Mapping functional domains of the 70-kDa subunit. *J. Biol. Chem.*, **270**, 4534–4543.

Gomes,X.V. and Wold,M.S. (1996) Functional domains of the 70-kilodalton subunit of human replication protein A. *Biochemistry*, **35**, 10558–10568.

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.

Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.*, **A47**, 110–119.

Kim,C. and Wold,M.S. (1995) Recombinant human replication protein A binds to polynucleotides with low cooperativity. *Biochemistry*, **34**, 2058–2064.

Kim,C., Paulus,B.F. and Wold,M.S. (1994) Interactions of human replication protein A with oligonucleotides. *Biochemistry*, **33**, 14197–14206.

Lavrik,O.I., Kolpashchikov,D.M., Nasheuer,H.P., Weisshart,K. and Favre,A. (1998) Middle subunit of replication protein A contacts growing RNA–DNA primers in replicating simian virus 40 chromosomes. *Mol. Cell. Biol.*, **18**, 6399–6407.

Lee,S.H., Kim,D.K. and Drissi,R. (1995) Human xeroderma pigmentosum group A protein interacts with human replication protein A and inhibits DNA replication. *J. Biol. Chem.*, **270**, 21800–21805.

Mass,G., Nethanel,T. and Kaufmann,G. (1998) The middle subunit of replication protein A contacts growing RNA-DNA primers in replicating simian virus 40 chromosomes. *Mol. Cell. Biol.*, **18**, 6399–407.

Miller,R., Gallo,S.M., Khalak,H.G. and Weeks,C.M. (1994) SnB: crystal structure determination via shake-and-bake. *J. Appl. Crystallogr.*, **27**, 613–621.

Murzin,A.G. (1993) OB (oligonucleotide/oligosaccharide)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.*, **12**, 861–867.

Nagelhus,T.A. *et al*. (1997) A sequence in the N-terminal region of human uracil-DNA glycosylase with homology to XPA interacts with the C-terminal part of the 34-kDa subunit of replication protein A. *J. Biol. Chem.*, **272**, 6561–6566.

Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. In Carter,J.C.W. and Sweet,R.M. (eds), *Methods in Enzymology: Macromolecular Crystallography, Part A*. Academic Press, Orlando, FL, Vol. 276, pp. 307–326.

Pfuetzner,R.A., Bochkarev,A., Frappier,L. and Edwards,A.M. (1997) Replication protein A. Characterization and crystallization of the DNA binding domain. *J. Biol. Chem.*, **272**, 430–434.

Philipova,D., Mullen,J.R., Maniar,H.S., Lu,J., Gu,C. and Brill,S.J. (1996) A hierarchy of SSB protomers in replication protein A. *Genes Dev.*, **10**, 2222–2233.

Ruff,M., Krishnaswamy,S., Boeglin,M., Poterszman,A., Mitschler,A., Podjarny,A., Rees,B., Thierry,J.C. and Moras,D. (1991) Class II aminoacyl transfer RNA synthetases: crystal structure of yeast aspartyl-tRNA synthetase complexed with tRNA (Asp). *Science*, **252**, 1682–1689.

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.

Van Duyne,G.D., Standaert,R.F., Karplus,P.A., Schreiber,S.L. and Clardy,J. (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol.*, **229**, 105–124.

Walther,A.P., Gomes,X.V., Lao,Y., Lee,C.G. and Wold,M.S. (1999) Replication protein A interactions with DNA. 1. Functions of the DNA-binding and zinc-finger domains of the 70-kDa subunit. *Biochemistry*, **38**, 3963–3973.

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.

Received April 20, 1999; revised and accepted June 22, 1999