Cell, Vol. 103, 449–456, October 27, 2000, Copyright ©2000 by Cell Press

Structural Basis for the Recognition of DNA Repair Proteins UNG2, XPA, and RAD52 by Replication Factor RPA

Georges Mer,[1] Alexey Bochkarev,[4,5] Rajesh Gupta,[3] Elena Bochkareva,[5] Lori Frappier,[4] C. James Ingles,[3] Aled M. Edwards,[2,3,7] and Walter J. Chazin[1,6,7]

[1]Department of Molecular Biology
The Scripps Research Institute
La Jolla, California 92037
[2]Department of Medical Biophysics
Ontario Cancer Institute
Toronto, Ontario M5G 2M9
Canada
[3]Banting and Best Department of Medical Research
C. H. Best Institute
Toronto, Ontario M5G 1L6
Canada
[4]Department of Medical Genetics and Microbiology
University of Toronto
Toronto, Ontario M5S 1A8
Canada
[5]Department of Biochemistry and Molecular Biology
University of Oklahoma Health Science Center
Oklahoma City, Oklahoma 73190
[6]Department of Biochemistry
Vanderbilt University Medical Center
Nashville, Tennesse 37232

Summary

Replication protein A (RPA), the nuclear ssDNA-binding protein in eukaryotes, is essential to DNA replication, recombination, and repair. We have shown that a globular domain at the C terminus of subunit RPA32 contains a specific surface that interacts in a similar manner with the DNA repair enzyme UNG2 and repair factors XPA and RAD52, each of which functions in a different repair pathway. NMR structures of the RPA32 domain, free and in complex with the minimal interaction domain of UNG2, were determined, defining a common structural basis for linking RPA to the nucleotide excision, base excision, and recombinational pathways of repairing damaged DNA. Our findings support a hand-off model for the assembly and coordination of different components of the DNA repair machinery.

Introduction

Replication protein A (RPA), the eukaryotic, single-stranded, DNA binding protein, is essential to nucleotide excision repair (NER) (Coverley et al., 1991; Sancar, 1996) and the repair of double-strand breaks by homologous recombination (Park et al., 1996; Sugiyama et al., 1998). RPA very likely also plays a significant role in postreplicative base excision repair (BER) (DeMott et al., 1998; Otterlei et al., 1999). In these pathways, RPA is thought to mediate the coordinated assembly of the DNA repair apparatus at sites of DNA damage through specific interactions with key repair proteins (Wold, 1997; Iftode et al., 1999). This is concomitant to RPA binding tightly to single-stranded (ss) DNA (Bochkarev et al., 1997), which presumably decreases ssDNA entropy (Wuite et al., 2000), and thereby stimulates DNA repair and replication. All of these functions are associated with the two larger subunits of RPA (RPA32 and RPA70). The smaller subunit (RPA14) is believed to have a structural role in assembly of the RPA heterotrimer.

Several critical observations implicate the participation of RPA32 in the three repair pathways. First, in NER, RPA32 binds specifically to the xeroderma pigmentosum damage-recognition protein XPA, facilitating XPA interaction with the damaged site (He et al., 1995). Deletion of the C-terminal region of RPA32 results in the inhibition of NER activity in vitro (Stigger et al., 1998; Lee et al., 1999). Second, RPA32 was shown to interact with the recombination protein RAD52 (Park et al., 1996; Hays et al., 1998); with its C terminus being crucial to this interaction and to homologous recombination (Park et al., 1996). Third, RPA32 interacts with the major human uracil-DNA glycosylase (UNG2), a key enzyme in BER (Nagelhus et al., 1997; Otterlei et al., 1999). Combined with the observation that UNG2 and RPA colocalize in replication foci (Otterlei et al., 1999), this suggests a role for RPA in coupling BER to DNA replication.

To investigate the functions of RPA32 in multiple DNA repair pathways, we determined the three-dimensional (3D) solution structure of the C-terminal region of human RPA32, $RPA32_{172–270}$, both free and in complex with a sixteen amino acid peptide fragment ($UNG2_{73–88}$) encompassing the known RPA binding region of UNG2 (Otterlei et al., 1999). The 3D structure of the complex served as a basis to identify the RPA32 binding sequences in XPA and RAD52. High resolution NMR was used to demonstrate that the interaction interfaces of UNG2, XPA, and RAD52 with RPA32 are similar. Based on these results, we propose a competition-based protein switch mechanism to assemble requisite proteins at sites of DNA damage.

Results and Discussion

$RPA32_{172–270}$ Contains a Winged Helix-Loop-Helix Domain Involved in Protein–Protein Interactions

Following upon the studies showing that the C-terminal region of RPA32 is required for DNA repair, a $RPA32_{172–270}$ construct was subcloned, expressed, and purified for structural characterization. The first 32 residues of this construct are unstructured, as evidenced by the random coil values of NMR chemical shifts and the absence of long-range nuclear Overhauser enhancements (NOEs). This observation is consistent with the proposal that this region may be a flexible linker to the ssDNA binding

[7]To whom correspondence should be addressed (email: aled.edwards@utoronto.ca [A. M. E.]; walter.chazin@vanderbilt.edu [W. J. C.]).

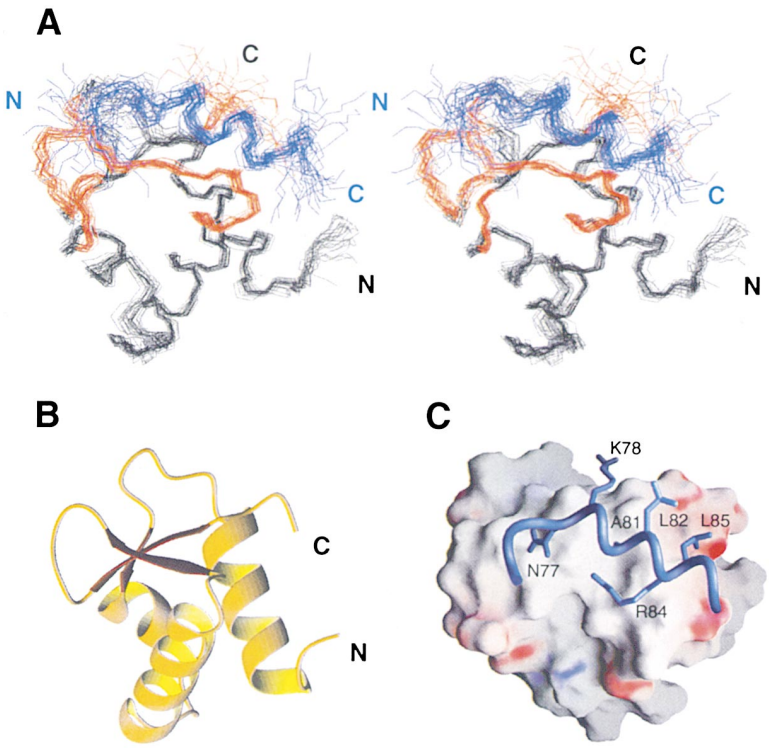

Figure 1. Structure of RPA32C Free and in Complex with $UNG2_{73–88}$

(A) Stereoview of the backbone (N, C^{α}, and C′) of 30 superimposed NMR structures with the lowest restraint violation energies of RPA32C complexed with $UNG2_{73–88}$. $UNG2_{73–88}$ is blue and RPA32C is red and black. The red portion shows perturbed ^{15}N-HSQC chemical shifts upon binding $UNG2_{73–88}$. Each model was superimposed onto the average structure using residues 207–219 and 223–266 of RPA32C and 77–85 of $UNG2_{73–88}$. The rmsd is 0.49 Å for the backbone atoms N, C^{α}, and C′, and 1.18 Å for all heavy atoms (not shown).

(B) Ribbon diagram (Koradi et al., 1996) of the structure of free RPA32C shown in the same orientation as in A. The three α helices, H-I (residues 207–217), H-II (227–233), and H-III (239–252) are depicted in yellow. The antiparallel β sheet involving residues 225 and 226 (S-I), 255 through 258 (S-II), and 263 through 266 (S-III) is shown in brown.

(C) Molecular surface representation of the electrostatic potential (positive in blue and negative in red) of RPA32C calculated in GRASP (Nicholls et al., 1993). The $UNG2_{73–88}$ peptide, depicting the side chains involved in intermolecular NOEs, is shown as a blue tube.

domain (residues 40–171) of RPA32 (Bochkareva et al., 1998). In contrast, the C-terminal residues 204–270, which we term RPA32C, adopt a compact globular α/β domain structure consisting of a right-handed three-helix bundle and a short three-stranded antiparallel β sheet (Figure 1B). Unexpectedly, the entire α/β fold is indistinguishable from the so-called winged helix-loop-helix (wHTH) DNA-recognition module present in several proteins in the structural database (DALI) (Holm and Sander, 1993). The closest similarity is with the globular domain of histone H5 (Ramakrishnan et al., 1993), where best-fit superposition of 61 C^{α} atoms gives an atomic root-mean-square deviation (rmsd) of 1.9 Å. One noticeable feature of RPA32C is that it binds other proteins through its β sheet region (vide infra) whereas most DNA binding wHTH proteins utilize the third α helix for DNA recognition. There may be wHTH proteins that bind both DNA and other proteins through a single domain.

UNG2 Associates with RPA32 via an Induced α Helix

In the presence of $UNG2_{73–88}$, the structure of $RPA32_{172–270}$ is largely unaffected. The N-terminal portion of $RPA32_{172–270}$ remains unfolded, indicating that this segment is not required for formation of the complex. In the folded region (RPA32C), the rmsd between free and complexed protein is 0.98 Å over all heavy atoms (Figure 1A).

The $UNG2_{73–88}$ peptide alone has no stable structure, but adopts a helical conformation upon binding $RPA32_{172–270}$. This induced helix extends from Asn77 to Ala87, as evidenced by several sequential and medium-range NOEs characteristic of α-helical structure (data not shown). The $UNG2_{73–88}$ peptide contacts the second and third strands of the β sheet of RPA32C and also the C-terminal end of helix H-III (Figure 1A). Fifty-five intermolecular NOEs were assigned unambiguously in a 3D ^{13}C-filtered, ^{12}C-edited NOE spectrum of this complex.

Ala81, Leu82, Leu85, and the aliphatic portions of Lys78 and Arg84 from the peptide are splayed across a complementary surface of RPA32C comprising residues Gly253, His254, Tyr256, Ser257, Thr258, Asp260, Thr267, and Asp268, all of which exhibit intermolecular NOEs with the peptide (Figure 2A). For instance, the aromatic ring of Tyr256 occupies a central position at the surface of the β sheet and participates in van der Waals contacts with the aliphatic protons of Lys78. There is a significant electrostatic component to the interaction (Figure 1C): five basic residues of the peptide, clustered at both ends of the central hydrophobic region, align well with two acidic patches on RPA32C's binding site. The spatial proximity of Arg73, Arg76, and Arg88 of UNG2 to the acidic patches created by Asp260, Asp261, and Asp262; and Asp247, Glu252, Asp268, and Glu270 is illustrated in Figure 2B. Also shown is the favorable orientation of Arg84 toward helix H-III macrodipole in RPA32C. In addition, the guanidino group of Arg84 forms hydrogen bonds with the backbone carbonyl of Ser250 or Asn251 in 80% of the structures. The side chain carbonyl of Asn77 forms a hydrogen bond with the hydroxyl of Ser257 in 25% of the structures.

The relative orientation of the DNA binding and protein interaction parts of UNG2 was modeled by linking the 3D structure of RPA32C-$UNG2_{73–88}$ to the crystal structure of UNG2's catalytic domain ($UNG2_{91–313}$) bound to DNA (Slupphaug et al., 1996) (Figure 3). It shows that $UNG2_{73–88}$ is located opposite the DNA binding site. This

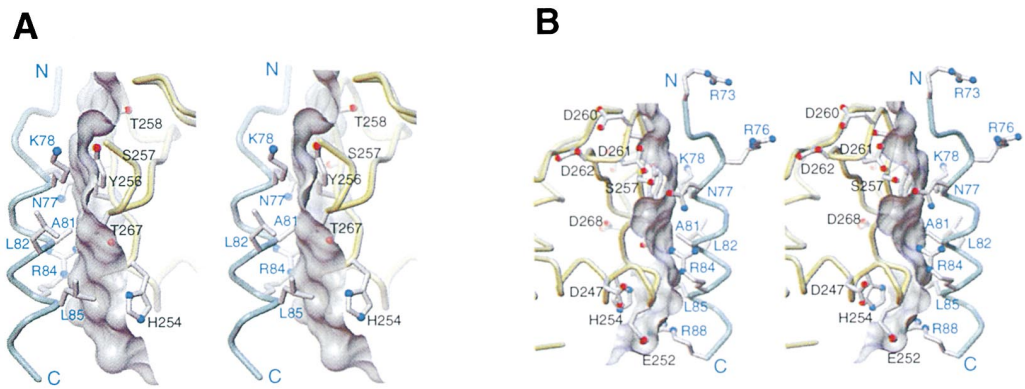

Figure 2. Details of RPA32C-UNG2$_{73–88}$ Interaction

(A) The UNG2$_{73–88}$ peptide is depicted in blue and the protein in yellow. In gray is the molecular surface of the protein generated with MSMS (Sanner et al., 1996). The side chains involved in intermolecular NOEs are displayed.

(B) The complex was rotated by 180° around the peptide's principal axis. All acidic side chains of RPA32C and basic side chains of UNG2$_{73–88}$ are displayed.

supports the modular construction of UNG2 in which the N-terminal region seems to be dedicated to multiple interactions with other proteins (Otterlei et al., 1999). In vivo, repair enzymes and accessory proteins are thought to function in concert as part of a network of successive short-lived protein–protein interactions (Parikh et al., 1999). The structure of RPA32C-UNG2$_{73–88}$ illustrates this concept.

XPA and RAD52 Bind RPA32 in the Same Manner as UNG2

Employing the 3D structure of RPA32$_{172–270}$-UNG2$_{73–88}$ as a basis, we characterized the interaction of RPA32 with the NER protein XPA. As a first step, affinity chromatography and mass spectrometry were used to identify the segments of XPA that bound RPA32. Fragments of human XPA generated by limited proteolysis were passed over a column of an immobilized complex of RPA14 and either full-length RPA32 or C-terminally truncated (lacking residues 172–270) RPA32Δ. XPA was retained only when the column contained full-length RPA32, confirming the necessity of the C-terminal region of RPA32 for the RPA32-XPA interaction (Figure 4A). All bound XPA fragments contained residues 1–98, suggesting that this N-terminal segment contains the essential RPA32 binding domain. As expected, an XPA$_{1–98}$ construct bound specifically to RPA32/14 but not to RPA32Δ/14 (Figure 4B).

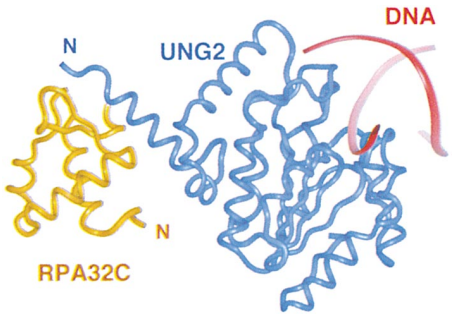

Figure 3. Model for the RPA32C-UNG2-DNA Complex

The NMR structure of RPA32C-UNG2$_{73–88}$ was linked to the crystal structure of UNG2$_{91–313}$ (catalytic domain) bound to DNA (Slupphaug et al., 1996). The interacting α helix of UNG2 to RPA32C extends away from the catalytic domain. This is consistent with limited proteolysis data showing that the N-terminal segment of UNG2 corresponding to the interacting α helix is cleaved away, even in the presence of RPA (Nagelhus et al., 1997).

The interaction of XPA$_{1–98}$ with RPA32$_{172–270}$ was then examined more extensively using NMR spectroscopy (Figure 5A). The surface of RPA32$_{172–270}$ that contacts XPA$_{1–98}$ was determined by monitoring the changes in the ^{15}N- and ^{1}H-amide chemical shifts in the ^{15}N-heteronuclear single-quantum coherence (HSQC) spectra of RPA32$_{172–270}$ upon titration with unlabeled XPA$_{1–98}$. The addition of XPA$_{1–98}$ caused fifteen ^{15}N/^{1}H correlations to broaden and four to shift (Figure 5A). These affected resonances of RPA32$_{172–270}$ correspond to residues located within or near the β sheet region; akin to those perturbed upon addition of the UNG2$_{73–88}$ peptide. As we have shown for RPA32$_{172–270}$-UNG2$_{73–88}$ (Figure 1A), perturbations of ^{15}N and ^{1}H chemical shifts are a very sensitive indicator of the interaction surface of RPA32C. The similarity between the titration spectra of RPA32$_{172–270}$ with UNG2$_{73–88}$ and XPA$_{1–98}$ indicates that the two molecules interact with RPA32$_{172–270}$ in a virtually identical manner.

A reverse titration with ^{15}N-labeled XPA$_{1–98}$ and unlabeled RPA$_{172–270}$ was also carried out to identify the region of XPA$_{1–98}$ that binds RPA32$_{172–270}$. Inspection of ^{15}N-HSQC spectra of XPA$_{1–98}$ indicated that two sets of side chain amide resonances were affected (data not shown). This observation, taken together with the results of the RPA32$_{172–270}$-UNG2$_{73–88}$ structure, suggested that only a short amino acid segment could be accommodated on RPA32C's binding surface. Consequently, the binding site was tentatively localized to a positively charged region of XPA that contains two glutamines (Gln33 and Gln40). The corresponding peptide, XPA$_{29–46}$, was synthesized and used in binding studies with ^{15}N-labeled RPA32$_{172–270}$ (Figure 5A). XPA$_{29–46}$ altered the chemical shifts of those residues of RPA32$_{172–270}$ that were perturbed in the presence of XPA$_{1–98}$ (and UNG2$_{73–88}$), demonstrating that XPA$_{29–46}$ is sufficient to generate the

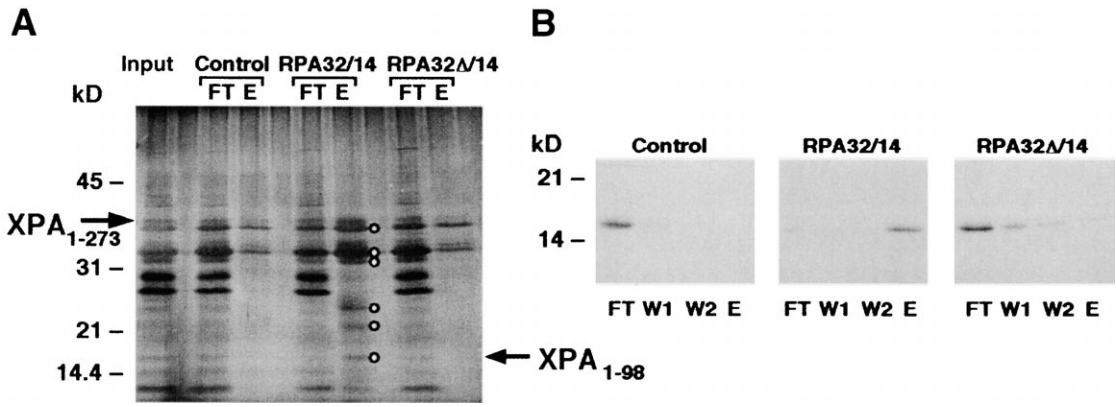

Figure 4. Identification of the RPA32 Binding Domain of XPA

(A) Affinity chromatography of partially proteolyzed human XPA on RPA columns. Recombinant XPA was purified from bacterial cells (He et al., 1995) and a mixture of XPA fragments generated by partial proteolysis with trypsin (input) was passed through columns which either contained no bound ligand (Control), the RPA32/14 heterodimer, or the same complex with RPA32 missing residues 172–270 (RPA32Δ/14). The flow-through (FT) and eluted fragments with 1M NaCl (E) were resolved by denaturing gel electrophoresis and visualized by silver staining. The XPA fragments that bound specifically to the RPA32/14 column and whose masses were determined using matrix-assisted laser desorption/ionization-time of flight mass spectrometry are indicated.

(B) Specific interaction of XPA_{1-98} with RPA32/14. Purified XPA_{1-98} was passed over similar affinity columns used in Panel A. The flow-through (FT), successive washes (W1 and W2), and the fractions eluted (E) were resolved by denaturing gel electrophoresis and visualized by staining with Coomassie blue.

full set of intermolecular interactions with $RPA32_{172-270}$. Similar to $UNG2_{73-88}$, a folding transition of XPA_{29-46} to an α-helical conformation is evident from the changes in ^1H chemical shifts of the peptide upon binding $RPA32_{172-270}$ (data not shown). Interestingly, XPA_{29-46} also encompasses the nuclear localization signal (NLS) of XPA (Miyamoto et al., 1992).

The $UNG2_{73-88}$ and XPA_{29-46} peptide sequences can be aligned using the amino acids of $UNG2_{73-88}$ that participate in intermolecular contacts with RPA32C (Figure 5B). The alignment illustrates how XPA_{29-46} interacts with RPA32. It is interesting that although both XPA_{29-46} and $UNG2_{73-88}$ terminate with Arg-Leu-Ala-Ala-Arg, these regions are in fact not aligned in a structural context (Figure 5B). The structural analysis reveals that these regions would not occupy corresponding positions in the respective RPA32C complexes. Interestingly this sequence forms part of an α helix in all seven occurrences found in other protein structures from the Protein Data Bank; it may bias peptides toward helical conformations and thereby facilitate their interaction with RPA32C.

The exchange between the bound and free states of XPA_{29-46} and $UNG2_{73-88}$ peptides is slow on the NMR chemical shift time scale, indicative of a dissociation constant (K_d) $\leq 10^{-6}$ M. The binding affinity of $RPA32_{172-270}$ for XPA_{1-98} was weaker with a K_d of 5×10^{-6} M. This dissociation constant was determined by fitting the progressive change in chemical shift of the amino nitrogen of Gln33 as ^{15}N-labeled XPA_{1-98} was titrated with unlabeled $RPA_{172-270}$. Theses values demonstrate that RPA32 is a significant contributor to the overall affinity of the RPA-XPA interaction with a reported K_d of 1.9×10^{-8} M (Saijo et al., 1996). The additional contribution to binding may involve a direct interaction between the DNA binding domains of XPA and RPA70 (Ikegami et al., 1998; Buchko et al., 1999).

We also considered whether RAD52 shares a common mode of interaction with UNG2 and XPA. In human RAD52, the interacting domain has been localized to a 60 amino acid stretch between residues 221 and 280 (Park et al., 1996). Examination of this region of RAD52 sequence showed a pattern of similarity with XPA_{29-46} and $UNG2_{73-88}$ (Figure 5B): the segment 257–274 is positively charged with a high propensity to form an α helix. It also contains the sequence Arg-Gln-Lys, which is homologous to three of the contact residues of $UNG2_{73-88}$ and XPA_{29-46} in their respective interactions with RPA32C. Titration of the corresponding peptide ($RAD52_{257-274}$) into ^{15}N-labeled $RPA32_{172-270}$ showed that it interacted specifically with the same surface of the protein as did $UNG2_{73-88}$ and XPA_{29-46} (Figure 5A). The exchange between free and bound $RAD52_{257-274}$ is fast on the time scale of chemical shifts, showing that it binds more weakly than $UNG2_{73-88}$ and XPA_{29-46}. RAD52 and RPA function in DNA recombination as heterooligomeric complexes in which RAD52 forms multimeric ring-like structures interacting with RPA-coated ssDNA (Shinohara et al., 1998). The lower affinity of RPA for RAD52 might thus be compensated by oligomerization, which would lead to a higher effective affinity.

RPA32C Is an Independent Module in the Context of RPA32-RPA14 Complex

The presence of a 32 amino acid flexible segment at the N terminus of $RPA32_{172-270}$ (vide supra) suggests that RPA32C is an independent module that does not contact the adjacent regions of RPA. To test this hypothesis, we performed NMR studies on a complex of full length RPA32 and RPA14. If the RPA32/14 complex were a single globular unit, we would not expect to observe any NMR signals in a conventional HSQC experiment because the size of the complex would lead to fast relaxation and correspondingly broad resonance lines. The ^{15}N-HSQC spectrum of RPA32/14 (Figure 6) does

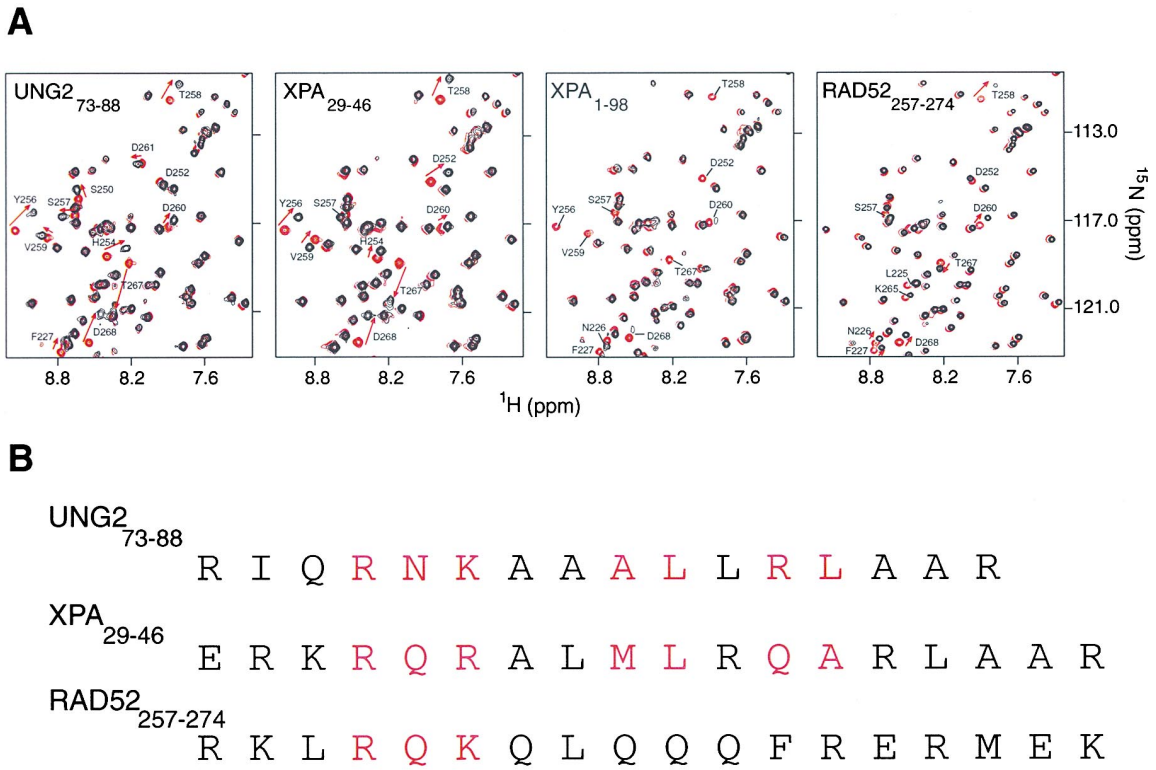

Figure 5. Interaction of $RPA32_{172–270}$ with Multiple Repair Proteins

(A) Superimposed region of the ^{15}N-HSQC spectra of RPA32 (0.25 mM) in its free form (red) and complexed to, starting from the left, $UNG2_{73–88}$ (molar ratio 1:1), $XPA_{29–46}$ (molar ratio 1:1), $XPA_{1–98}$ (molar ratio molar ratio 1:2), and $RAD52_{257–274}$ (molar ratio molar ratio 1:8). All experiments were recorded at 500 MHz, pH 7.0, and 25°C except for $XPA_{29–46}$ (pH 7.5) and $RAD52_{257–274}$ (600 MHz).

(B) Amino acid sequences of $UNG2_{73–88}$, $XPA_{29–46}$, and $RAD52_{257–274}$. In red are the residues of $UNG2_{73–88}$ involved in intermolecular NOEs and the corresponding alignments with $XPA_{29–46}$ and $RAD52_{257–274}$.

however show signals that correspond mostly to the previously assigned spectrum of $RPA_{172–270}$. The remaining few additional resonances are attributed to the N-terminal ~40 amino acids containing the essential RPA32 phosphorylation site, based on the large number of glycine resonances (9) that could be identified in the spectrum (Figure 6). The absence of other signals from RPA32/14 indicates that the core of RPA32/14 complex behaves as a single large entity, in agreement with the crystal structure (Bochkarev et al., 1999). The appearance of peaks corresponding to $RPA32_{1–40}$ and $RPA32_{172–270}$ suggests that these two regions tumble independently of the core of the complex. The unrestricted motion of RPA32C could be a source of plasticity for RPA to bind different protein targets. We cannot, however, rule out that the high mobility of the domain is inhibited in the full heterotrimeric complex.

RPA32C: Mediating the Assembly of DNA Repair Complexes via a Hand-Off Mechanism

RPA is involved in both early and later steps of DNA repair within the nucleotide excision, double-strand break point and possibly base-excision DNA repair pathways. In early steps of both NER and recombinational repair, RPA binds to the ssDNA opposite or adjacent to the site of DNA damage and interacts with one of the damage-recognition proteins (XPA or RAD52). Subsequently, in NER and perhaps in other pathways, RPA interacts with the excision nucleases. In the later stage, RPA is known to be required for DNA polymerization, where it makes specific contacts with the repair DNA polymerases (Longhese et al., 1994).

This involvement of RPA throughout the repair reactions is orchestrated by a series of RPA–ssDNA and RPA–protein interactions. These interactions assist in both the assembly of the repair complexes and in the structural transitions that occur along the reaction pathway. In the early stages of the assembly of the repair complexes, RPA plays an architectural role. Since RPA binds ssDNA with high affinity in a defined orientation with respect to the 5′ and 3′ polarity (Bochkarev et al., 1997), and RPA binding to DNA is an early event in the assembly of repair complexes, binding of RPA to DNA is likely to establish the initial polarity of assembled protein–DNA complexes. The initial architectural role not only includes RPA interaction with ssDNA but also with one of the lesion-recognizing proteins, UNG2, XPA, or RAD52. Subsequent repair processes likewise rely on a series of RPA-mediated protein interactions. For example, in NER, RPA interacts specifically with the excision endonucleases XPF-ERCC1 and XPG, as well as with the repair DNA polymerases (Aboussekhra et al., 1995; Matsunaga et al., 1996). The directionality of these subsequent protein interactions, some involving RPA32C,

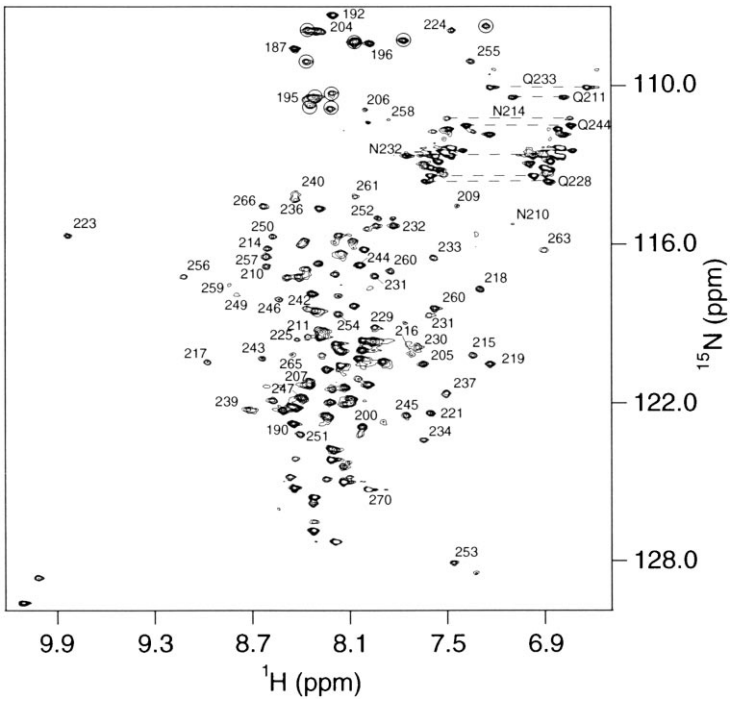

Figure 6. ^{15}N-^{1}H Correlation Spectrum of the RPA32–RPA14 Complex

^{15}N-HSQC spectrum of RPA32/14 recorded at 800 MHz, 25°C, and pH 7.5. The resonances assigned to $RPA32_{172–270}$ are labeled. The signals attributed to glycine residues of $RPA32_{1–40}$ are circled.

is also established by the initial binding of RPA to ssDNA. The RPA heterotrimer comprises four DNA binding modules, three in RPA70 and one in RPA32 (Iftode et al., 1999). Recent studies suggest that the RPA32 ssDNA binding domain is positioned at the 3′-end of the RPA/DNA complex (de Laat et al., 1998; Iftode et al., 1999), placing the RPA32C protein interaction module also at the 3′-end of the RPA complex. This positioning of RPA32C is probably key to the integrity of the repair assembly.

The progression through each of the different DNA repair reactions requires the assembly and disassembly of specific protein complexes. For example, in the NER pathway, RPA is at the center of a set of dynamic interactions with XPA, XPG, and XPF-ERCC1. After cooperative recognition of the DNA damage by RPA and XPA, the XPG and XPF-ERCC1 nucleases are recruited to the 3′ and 5′ sides of the site of DNA damage, respectively, where each clips the damaged DNA strand (de Laat et al., 1998). The positioning of both of these nucleases on the DNA is facilitated by direct protein–protein interactions with RPA (de Laat et al., 1998). Progression through the NER pathway may involve RPA, possibly through its RPA32C domain "handing-off" XPA for XPF-ERCC1. In the later stages of the repair reaction, RPA might substitute its interaction with this nuclease for one with either DNA polymerase δ or ϵ to perform the repair synthesis reactions.

This hypothesis suggests that the dynamic assembly of the DNA repair machinery might be orchestrated by multiple interactions with RPA and its RPA32C domain, perhaps using a competition-based protein switch mechanism where proteins vie for binding RPA in the context of RPA-coated ssDNA. In other words, RPA32C might serve as the initial anchor and exchange point for a number of DNA repair proteins. A similar protein competition model was proposed for the function of RPA in redefining in time the successive protein complexes involved in the DNA replication process (Yuzhakov et al., 1999). As we learn more about the mechanisms for the various DNA repair pathways, and the molecular and kinetic details of all of the protein–protein interactions in the repair complexes, this "handing-off" mechanism may well prove to be a general means for the successive transfers to RPA of one protein after another as the repair process proceeds.

Finally, that RPA32C interacts similarly with XPA, UNG2, and RAD52 implies that the RPA32C interactions with the NER, BER, and recombinational machineries are exclusionary. Since certain DNA lesions can be repaired by more than one pathway, a critical point is reached when one is selected over the others. The exclusive nature of the RPA32C interaction might provide the structural basis for the choice of a specific DNA repair pathway.

Conclusion

This work shows that the participation of RPA in different aspects of DNA metabolism is attained through a certain degree of structural similarity in the RPA binding sites of unrelated proteins. The 3D structure of $RPA32_{172–270}$ in complex with $UNG2_{73–88}$ provides a framework for understanding the molecular interactions of XPA and RAD52 with RPA32. Our study demonstrates that UNG2, XPA, and RAD52 share the same binding surface on RPA32, and strongly suggests that all three proteins also share the same mode of binding, typified by an induced structural transition to an α helix. This structural convergence among UNG2, XPA, and RAD52, representative of three different pathways of DNA repair, provides new evidence in support of the requirement of RPA in

BER, and supports a hand-off model for the assembly and coordination of DNA repair complexes.

Experimental Procedures

Overexpression and Purification of $RPA_{172–270}$ and $XPA_{1–98}$

Human $RPA_{172–270}$ and $XPA_{1–98}$ were cloned into the T7 polymerase expression vector, pET15b (Novagen), and overexpressed in the *E. coli* strain BL21(DE3) pLysS as a fusion to an N-terminal six-histidine tag. Both proteins were purified by a combination of nickel-chelate affinity chromatography and, after removal of the histidine tag by digestion with thrombin, anion exchange high-performance liquid chromatography (HPLC). An additional reversed-phase HPLC purification step was done on $RPA_{172–270}$ before preparing the $RPA_{172–270}$–$UNG2_{73–88}$ complex for structural studies. Four amino acids, GSHM, persist from the fusion proteins and extend from the N-termini of $RPA_{172–270}$ and $XPA_{1–98}$. Unlabeled, uniformly ^{15}N- and $^{15}N/^{13}C$- and/or 10% ^{13}C-isotope labeled $RPA_{172–270}$ was obtained by growing cells in LB broth or M9-minimal media containing $^{15}NH_4Cl$ and $^{13}C_6$-glucose. $XPA_{1–98}$ was prepared unlabeled and uniformly ^{15}N-labeled only.

Overexpression and Purification of RPA32/14

RPA14 and RPA32 were coexpressed in M9 minimal medium containing $^{15}NH_4Cl$ and purified as previously described (Bochkareva et al., 1998). The sample for NMR studies contained ~0.4 mM RPA32/14, 10 mM sodium phosphate buffer (pH 7.5), 150 mM NaCl, 7.5 mM DTT, and 10 μM $ZnSO_4$.

Synthesis of $UNG2_{73–88}$, $XPA_{29–46}$, and $RAD52_{257–274}$

All peptides were synthesized on a Gilson AMS 422 synthesizer using Fmoc chemistry and purified by reversed-phase HPLC.

NMR Spectroscopy

All NMR spectra of $RPA_{172–270}$ and its complexes were recorded on samples containing 1 to 2 mM $RPA_{172–270}$, 20 mM sodium phosphate buffer (pH 7.0 or pH 7.5), 50 mM NaCl, and 5 or 10 mM DTT dissolved in either 7% or 99.99% D_2O. NMR experiments were recorded at 25°C on Bruker AMX 500, AMX 600, DRX 600, DMX 750, and DRX 800 MHz spectrometers. The ^1H, ^{13}C, and ^{15}N resonances of free $RPA32_{172–270}$ (Mer et al., 2000) were assigned from a series of 3D experiments (Bax and Grzesiek, 1993): ^{15}N-edited TOCSY, HNCACB, CBCA(CO)NH, C(CO)NH-TOCSY, HBHA(CBCACO)NH, and H(CCO)NH-TOCSY. The assignments of free $RPA32_{172–270}$ were used to assign the corresponding chemical shifts in the complex. $UNG2_{73–88}$ resonances were assigned using [F1-^{13}C, F2-^{13}C]-filtered TOCSY with 25, 50, and 66 ms of isotropic mixing, [F1-$^{13}C/^{15}N$, F2-$^{13}C/^{15}N$]-filtered, [F1-^{13}C, F2-^{13}C]-filtered, and [F1-^{15}N]-filtered NOESY spectra with 150 ms mixing time (Otting and Wüthrich, 1990; Ikura and Bax, 1992). Sixty methylene protons were stereospecifically assigned from the analysis of intraresidue NOEs in a short mixing time (τ_m = 30 ms) 2D ^1H NOESY combined with information on $^3J_{HN-H\alpha}$ and $^3J_{H\alpha-H\beta}$ coupling constants obtained from 3D HNHB and HACAHB-COSY experiments, respectively. Stereospecific assignments of the methyl groups of Leu and Val residues were obtained by nonrandom ^{13}C labeling (Senn et al., 1989). Interproton distances (361 intraresidue, 314 sequential, 291 medium-range, and 288 long-range for free $RPA32_{172–270}$; 361 intraresidue, 328 sequential, 304 medium-range, and 343 long-range, including 55 intermolecular for the complex) were derived from 2D ^1H NOESY (τ_m = 80 ms), 3D ^{15}N-edited NOESY (τ_m = 80 ms), 3D ^{13}C-edited NOESY (τ_m = 100 ms), and [F1-$^{13}C/^{15}N$, F2-$^{13}C/^{15}N$]-filtered NOESY (τ_m = 150 ms) spectra. Intermolecular NOEs were assigned using 2D and 3D ^{13}C-edited, ^{13}C-filtered NOESY experiments (τ_m = 200 ms) (Zwalen et al., 1997). Fifty-five ϕ dihedral angle restraints were obtained based on $^3J_{HN-H\alpha}$ coupling constants and 40 ψ restraints were assigned to residues for which C^α and H^α chemical shifts indicated helical conformations. Thirty-three χ_1 and 6 χ_2 torsion angle restraints were derived from analysis of stereospecific assignments and intraresidue NOEs. Twelve pairs of hydrogen bond restraints were included based on the identification of slowly exchanging amide protons and characteristic NOEs. All NMR spectra were processed with Felix97 (Molecular Simulations, Inc.).

Structure Calculations

Fifty structures of $RPA32_{172–270}$ and $UNG2_{73–88}$ were calculated with the program DIANA using the REDAC (Güntert and Wüthrich, 1991) strategy and refined by simulated annealing using AMBER 4.1 (Pearlman et al., 1995). The calculation protocol consisted of 3000 steps of energy minimization prior to 20 ps of restrained simulated annealing. The starting orientations of the complex were generated by positioning the peptide and the protein 30 Å apart using NAB (Macke and Case, 1998). The two molecules were then docked by restrained molecular dynamics for 20 ps, during which the intermolecular restraints were ramped on from a force constant of 0 to 32 kcal mol^{-1} $Å^{-2}$. All restraints within the peptide and protein were maintained with force constants of 32 kcal mol^{-1} $Å^{-2}$. A final representative ensemble of 30 structures for both free and bound $RPA32_{172–270}$ were selected for analysis based on restraint violation energies. The mean restraint violation and total Amber energies of the complex were 4.6 ± 0.7 kcal mol^{-1} and −1751 ± 32 kcal mol^{-1}, respectively. No distance and angle restraints were violated by more than 0.2 Å and 5°, respectively. The quality of the final structures was analyzed with Procheck-NMR, and revealed that 88.1% of the nonglycine and nonproline residues were in the most favorable region of the Ramachandran plot, and 9.8% in additionally allowed regions. The final ensemble of structures of free $RPA32_{172–270}$ was of similar quality and precision as the complex.

Ligand Titrations

Ligand titrations presented in Figure 5A were performed by recording a series of 2D ^{15}N- and ^{13}C-HSQC spectra on 0.25 mM $^{15}N/10\%$ ^{13}C-isotope labeled $RPA32_{172–270}$ with increasing concentration of ligand ranging from 0 to 2 mM. All samples were prepared in 20 mM sodium phosphate buffer, 50 mM NaCl, and 5 mM DTT. The titration spectra were recorded at 25°C and pH 7.0 for $UNG2_{73–88}$, $XPA_{1–98}$, and $RAD52_{257–274}$. The pH was 7.5 for the titration with $XPA_{29–46}$.

Acknowledgments

We thank C. Arrowsmith, D. Case, J. Tainer, M. V. Botuyan, J. Christodoulou, P. Fagan, S. Parikh, and S. Veeraraghavan for fruitful discussions; M. Pique for assistance with graphics; L. Mäler, J. Smith, and C. Weber for assistance with Amber; J. Chung for help in setup of NMR experiments; and M. V. Botuyan for protein purification. L. F. and A. M. E. are Medical Research Council of Canada Scientists. This research was supported by grants from the National Cancer Institute of Canada to C. J. I., L. F., and A. M. E., the National Institutes of Health to A. B., and the National Science Foundation (to W. J. C.). G. M. acknowledges the support of postdoctoral fellowships from the French Association for Cancer Research, the Philippe Foundation, Inc., and the Human Frontier Science Program Organization.

Received June 1, 2000; revised September 5, 2000.

References

Aboussekhra, A., Biggerstaff, M., Shivji, M.K., Vilpo, J.A., Moncollin, V., Podust, V.N., Protic, M., Hubscher, U., Egly, J.-M., and Wood, R.D. (1995). Mammalian DNA nucleotide excision repair reconstituted with purified protein components. Cell *80*, 859–868.

Bax, A., and Grzesiek, S. (1993). Methodological advances in protein NMR. Acc. Chem. Res. *26*, 131–138.

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.

Bochkarev, A., Pfuetzner, R.A., 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.

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.

Buchko, G.W., Daughdrill, G.W., de Lorimier, R., Rao, B.K., Isern, N.G., Lingbeck, J.M., Taylor, J.S., Wold, M.S., Gochin, M., Spicer, L.D., et al. (1999). Interactions of human nucleotide excision repair protein XPA with DNA and RPA70 Delta C327: chemical shift mapping and ^{15}N NMR relaxation studies. Biochemistry *38*, 15116–15128.

Coverley, D., Kenny, M.K., Munn, M., Rupp, W.D., Lane, D.P., and Wood, R.D. (1991). Requirement for the replication protein SSB in human DNA excision repair. Nature *349*, 538–541.

De Laat, W.L., Appeldoorn, E., Sugasawa, K., Weterings, E., Jaspers, N.G.J., and Hoeijmakers, J.H.J. (1998). DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. Genes Dev. *12*, 2598–2609.

DeMott, M.S., Zigman, S., and Bambara, R.A. (1998). Replication protein A stimulates long patch DBA base excision repair. J. Biol. Chem. *273*, 27492–27498.

Güntert, P., and Wüthrich, K. (1991). Improved efficiency of protein structure calculations from NMR data using the program DIANA with redundant dihedral angle constraints. J. Biomol. NMR *1*, 447–456.

Hays, S.L., Firmenich, A.A., Massey, P., Banerjee, R., and Berg, P. (1998). Studies of the interaction between RAD52 protein and the yeast single-stranded DNA binding protein RPA. Mol. Cell. Biol. *18*, 4400–4406.

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.

Holm, L., and Sander, C. (1993). Protein structure comparison by alignment of distance matrices. J. Mol. Biol. *233*, 123–138.

Iftode, C., Daniely, Y., and Borowiec, J.A. (1999). Replication protein A (RPA): the eukaryotic SSB. Crit. Rev. Biochem. Mol. Biol. *34*, 141–180.

Ikegami, T., Kuraoka, I., Saijo, M., Kodo, N., Kyogoku, Y., Morikawa, K., Tanaka, K., and Shirakawa, M. (1998). Solution structure of the DNA- and RPA-binding domain of the human repair factor XPA. Nat. Struct. Biol. *5*, 701–706.

Ikura, M., and Bax, A. (1992). Isotope-filtered 2D NMR of a protein-peptide complex: study of a skeletal muscle myosin light chain kinase fragment bound to calmodulin. J. Am. Chem. Soc. *114*, 2433–2440.

Koradi, R., Billeter, M., and Wüthrich, K. (1996). MOLMOL: A program for display and analysis of macromolecular structures. J. Mol. Graphics *14*, 51–55.

Lee, B.E., Sung, J.W., Kim, D.K., Lee, J.R., Kim, N.D., Kang, S.W., and Kim, D.K. (1999). Functional studies on the interaction between human replication protein A and Xeroderma pigmentosum group A complementing protein (XPA). Mol. Cell *9*, 185–190.

Longhese, M.P., Plevani, P., and Luchini, G. (1994). Replication factor A is required in vivo for DNA replication, repair, and recombination. Mol. Cell. Biol. *14*, 7884–7890.

Macke, T., and Case, D.A. (1998). Modeling Unusual Nucleic Acid Structures, N.B. Leontes and J. Santalucia, Jr., eds. (Washington, DC: American Chemical Society), pp. 379–393.

Matsunaga, T., Park, C.-H., Bessho, T., Mu, D., and Sancar, A. (1996). Replication protein A confers structure-specific endonuclease activities to the XPF-ERCC1 and XPG subunits of human DNA repair excision nuclease. J. Biol. Chem. *271*, 11047–11050.

Mer, G., Edwards, A.M., and Chazin, W.J. (2000). ^{1}H, ^{15}N and ^{13}C resonance assignments for the C-terminal protein interaction region of the 32 kDa subunit of human replication protein A. J. Biomol. NMR *17*, 179–180.

Miyamoto, I., Miura, N., Niva, H., Miyazki, J., and Tanaka, K. (1992). Mutational analysis of the structure and function of the xeroderma pigmentosum group A complementing protein. J. Biol. Chem. *267*, 12182–12187.

Nagelhus, T.A., Haug, T., Singh, K.K., Keshav, K.F., Skorpen, F., Otterlei, M., Bharati, S., Lindmo, T., Benichou, S., Benarous, R., and Krokan, H.E. (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.

Nicholls, A., Bharadwj, R., and Honig, B. (1993). GRASP: graphical representation and analysis of surface properties. Biophys. J. *64*, 166–170.

Otterlei, M., Warbrick, E., Nagelhus, T.A., Haug, T., Slupphaug, G., Akbari, M., Aas, P.A., Steinsbekk, K., Bakke, O., and Krokan, H.E. (1999). Post-replicative base excision repair in replication foci. EMBO J. *18*, 3834–3844.

Otting, G., and Wüthrich, K. (1990). Heteronuclear filters in two-dimensional [^{1}H,^{1}H]-NMR spectroscopy: combined use with isotope labelling for studies of macromolecular conformation and intermolecular interactions. Quat. Rev. Biophys. *23*, 39–96.

Parikh, S.S., Mol, C.D., Hosfield, D.J., and Tainer, J.A. (1999). Envisioning the molecular choreography of DNA base excision repair. Curr. Opin. Struct. Biol. *9*, 37–47.

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.

Pearlman, D.A., Case, D.A., Caldwell, J.W., Ross, W.S., Cheatham, T.E., III, Ferguson, D.M., Seibel, G.L., Singh, U.C., Weiner, P.K., and Kollman, P.A. (1995). AMBER 4.1. (San Francisco: University of California).

Ramakrishnan, V., Finch, J.T., Graziano, V., Lee, P.L., and Sweet, R.M. (1993). Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. Nature *362*, 219–223.

Saijo, M., Kuraoka, I., Masutani, C., Hanaoka, F., and Tanaka, K. (1996). Sequential binding of DNA repair proteins RPA and ERCC1 to XPA *in vitro*. Nucleic Acids Res. *24*, 4719–4724.

Sancar, A. (1996). DNA excision repair. Annu. Rev. Biochem. *65*, 43–81.

Sanner, M.F., Olson, A.J., and Sphener, J.-C. (1996). Reduced surface: an efficient way to compute molecular surfaces. Biopolymers *38*, 305–320.

Senn, H., Werner, B., Messerle, B.A., Weber, C., Traber, R., and Wüthrich, K. (1989). Stereospecific assignment of the methyl ^{1}H NMR lines of valine and leucine in polypeptides by nonrandom ^{13}C labelling. FEBS Lett. *249*, 113–118.

Shinohara, A., Shinohara, M., Ohta, T., Matsuda, S., and Ogawa, T. (1998). Rad52 forms ring structures and co-operates with RPA in single-strand DNA annealing. Genes Cells *3*, 145–156.

Slupphaug, G., Mol, C.D., Kavli, B., Arvai, A.S., Krokan, H.E., and Tainer, J. (1996). A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. Nature *384*, 87–92.

Stigger, E., Drissi, R., and Lee, S.-H. (1998). Functional analysis of human replication protein A in nucleotide excision repair. J. Biol. Chem. *273*, 9337–9343.

Sugiyama, T., New, J.H., and Kowalczykowski, S.C. (1998). DNA annealing by Rad52 protein is stimulated by specific interaction with the complex of replication protein A and single-stranded DNA. Proc. Natl. Acad. Sci. USA *95*, 6049–6054.

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.

Wuite, G.J.L., Smith, S.B., Young, M., Keller, D., and Bustamante, C. (2000). Single-molecule studies of the effect of template tension on T7 DNA polymerase activity. Nature *404*, 103–106.

Yuzhakov, A., Kelman, Z., Hurwitz, J., and O'Donnell, M. (1999). Multiple competition rections for RPA order the assembly of the DNA polymerase δ holoenzyme. EMBO J. *18*, 6189–6199.

Zwalen, C., Legault, P., Vincent, S.J.F., Greenblatt, J., Konrat, R., and Kay, L.E. (1997). Methods for measurement of intermolecular NOEs by multinuclear NMR spectroscopy: application to a bacteriophage λ N-peptide/*boxB* RNA complex. J. Am. Chem. Soc. *119*, 6711–6721.

Protein Data Bank ID Code

The coordinates of RPA32C in complex with $UNG2_{73-88}$ were deposited in the Protein Data Bank with the ID code 1DPU.