Journal of Biomolecular NMR, **14:** 321–331, 1999.
KLUWER/ESCOM

Human replication protein A: Global fold of the N-terminal RPA-70 domain reveals a basic cleft and flexible C-terminal linker†

Doris M. Jacobs[a], Andrew S. Lipton[a], Nancy G. Isern[a], Gary W. Daughdrill[a], David F. Lowry[a,*], Xavier Gomes[b] & Marc S. Wold[b]
[a]*Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, 902 Battelle Boulevard, Richland, WA 99352, U.S.A.;* [b]*Department of Biochemistry, University of Iowa College of Medicine, Iowa City, IA 52240, U.S.A.*

Received 13 May 1999; Accepted 21 June 1999

Key words: backbone assignments, DNA recombination, DNA repair, DNA replication, RPA, RPA70, secondary structure

Abstract

Human Replication Protein A (hsRPA) is required for multiple cellular processes in DNA metabolism including DNA repair, replication and recombination. It binds single-stranded DNA with high affinity and interacts specifically with multiple proteins. hsRPA forms a heterotrimeric complex composed of 70-, 32- and 14-kDa subunits (henceforth RPA70, RPA32, and RPA14). The N-terminal 168 residues of RPA70 form a structurally distinct domain that stimulates DNA polymerase α activity, interacts with several transcriptional activators including tumor suppressor p53, and during the cell cycle it signals escape from the DNA damage induced G2/M checkpoint. We have solved the global fold of the fragment corresponding to this domain (RPA70Δ169) and we find residues 8–108 of the N-terminal domain are structured. The remaining C-terminal residues are unstructured and may form a flexible linker to the DNA-binding domain of RPA70. The globular region forms a five-stranded anti-parallel β-barrel. The ends of the barrel are capped by short helices. Two loops on one side of the barrel form a large basic cleft which is a likely site for binding the acidic motifs of transcriptional activators. Many lethal or conditional lethal yeast point mutants map to this cleft, whereas no mutations with severe phenotype have been found in the linker region.

Abbreviations: C′, carbonyl resonance; CBCA(CO)NH, $C^{\beta}_{(i-1)}$ to $C^{\alpha}_{(i-1)}$ to C′ to amide N/HN correlation experiment; DTT, dithiothreitol; EDTA, sodium salt of ethylenediaminetetraacetic acid; HBCBCACOCAHA, $H^{\alpha\beta}$ to $C^{\alpha\beta}$ to C^{α} to C′ to C^{α} to H^{α} intraresidue correlation experiment; HNCACB, amide HN/N to $C^{\alpha}_{i/(i-1)}$, $C^{\beta}_{i/(i-1)}$ correlation experiment; HNCO, amide HN/N to $C'_{(i-1)}$ correlation experiment; HMQC, heteronuclear multiple quantum correlation; HSQC, heteronuclear single quantum correlation; IPTG, isopropyl β-D-thiogalactopyranoside; KPi, equimolar KH_2PO_4 (potassium-dihydrogen-phosphate)/K_2HPO_4 (dipotassium-hydrogen-phosphate); NaN_3, sodium azide; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; PMSF, phenylmethanesulfonyl fluoride; RPA, Replication Protein A, SSB, single-stranded binding protein; TOCSY, total correlation spectroscopy; TPPI, time proportional phase incrementation; Tris, tris(hydroxymethyl)aminomethane; 2D, 3D, 4D, two-, three-, four-dimensional.

Introduction

Homo sapiens Replication Protein A (henceforth hsRPA) also known as Replication Factor A or human SSB is the most abundant single-stranded DNA-binding protein in human cells (Wold and Kelly, 1988; Fairman and Stillman, 1988; Wold, 1997; Wobbe et al., 1987; for reviews) It is required for multi-

*To whom correspondence should be addressed. E-mail: david.lowry@pnl.gov.

ple processes in eukaryotic DNA metabolism such as DNA replication, repair and recombination. hsRPA is required for both initiation and elongation phases of DNA replication. hsRPA binds non-specifically to single-stranded DNA (ssDNA) and has an ATP-independent DNA helix destabilizing activity (Georgaki et al., 1992).

Unlike other SSBs, RPA contains protein interaction domains essential for function. For instance, interactions with DNA polymerase α and SV40 are essential for initiation of DNA replication (Braun et al., 1997; Weisshart et al., 1998), and during nucleotide excision repair hsRPA binds to the repair factors XPA, XPG and ERCC1/XPF. Furthermore, the interaction of hsRPA with Rad52 is essential for homologous recombination. Recently, it was reported that in the yeast *Saccharomyces cerevisiae*, Rad52 stimulates DNA strand exchange by recruiting Rad51 to a complex of RPA with single-stranded DNA. Rad52 mediates competition between Rad51 and RPA for DNA binding (Kanaar and Hoeijmakers, 1998). hsRPA also interacts with transcriptional activators such as GAL4, VP16 and tumor suppressor p53. Finally, RPA was also recently identified as a critical link between DNA double strand break repair and the escape from cell cycle arrest (Lee et al., 1998).

The 18 kDa N-terminal domain of RPA70 (residues 1–168) is a protein-protein interaction domain. This domain specifically interacts with transcription factors and DNA polymerase α and the importance of these interactions for RPA functions is confirmed by point mutants in the corresponding domain of the yeast scRPA70 gene (RFA1). We can reasonably map the yeast mutations onto the human gene product since 31% of the amino acids are identical and 44% of the amino acids are similar to the hsRPA1 gene encoding the large subunit of RPA (Erdile et al., 1991; Brill and Stillman, 1991; Umbricht et al., 1993). The point mutants are either lethal or have a deleterious effect on RPA function, thus causing slow or conditional growth as well as defects in DNA repair and recombination (Longhese et al., 1994; Firmenich et al., 1995; Umezu et al., 1998).

In an effort to understand how protein interactions of hsRPA contribute to DNA metabolism, we determined the global fold of RPA70Δ169. Yeast replication and repair mutants mapped onto this fold cluster around a large basic cleft. The basic cleft is formed by five arginines and one lysine in two long loops that extend from one end of a β-barrel. We propose that the basic cleft is a binding surface for transcriptional activators and repair factors. The other yeast mutants are near the helix 3–flexible linker transition, suggesting that the linker plays a direct role in RPA function.

Materials and methods

Production and purification of ^{15}N- and $^{13}C/^{15}N$-labeled RPA70Δ169

Forty μL aliquots of the *E. coli* bacterial strain BL21 or DL39/DE3 were transformed with the plasmid vector p11d-RPA-70Δ169-616 (Gomes and Wold, 1995). For ^{15}N-labeling the protein was over-expressed in CELTONE® -N (Martek Biosciences Corp.) and for ^{13}C/^{15}N-labeling the ISOGRO^{TM} Powder-Growth Medium (ISOTEC Inc.) was used. In both cases only 5 g/1 L of the powder and 1 g/1 L of D-glucose were added. For the doubled labeled samples D-glucose was replaced by [$^{13}C_6$]-D-glucose. In order to assist with assignment of NMR spectra, three protein samples incorporating one specific ^{15}N labeled amino acid each (glycine, alanine, and leucine) were produced in defined media as outlined by Muchmore et al. (1989) except the amounts of the amino acids (other than the labeled amino acid), the purines and pyrimidines were halved, but the amounts of the other components were as described. Cells were harvested 3 h after induction by centrifugation at 3000 rpm for 20 min and frozen at −80 °C.

The cells from 1 L of media were thawed and suspended in 20 mL TED-buffer (100 mM Tris, 100 mM Tris-HCl, 0.25 mM EDTA, 1 mM DTT, pH 7.6) containing 9 mg PMSF dissolved in 500 μL ethanol. The suspension was lysed by three passes through a French press. Cell debris was removed by centrifugation at 17 000 rpm for 30–60 min. The supernatant containing the soluble labeled protein was loaded onto S-Sepharose, the diluent buffer was TED, and the eluting buffer was TED with 1000 mM NaCl. The fraction containing the protein product eluted at 50 mM NaCl and was concentrated to 900 μL using a Centriprep 10 concentrator (Amicon Inc.). Since the expression of specifically ^{15}N-labeled protein was low and the protein unstable, an additional size exclusion column was run to increase purity. The buffer used for size exclusion was 50 mM KPi, 100 mM KCl, 10 mM DTT, pH 6.5. The purity of the protein was determined by using 8–14% SDS-polyacrylamide gels and analyzing them by Coomassie staining. The purification protocol results in a RPA70Δ169 fraction with a purity of

>90%. The yield of the uniformly labeled samples was approximately 17 mg of protein from 1 L of cell culture.

Uniformly ^{15}N-labeled- and ^{13}C/^{15}N-labeled NMR samples were prepared at pH 7.6 and 6.5 each. The following buffers were used: 100 mM Tris, 100 mM Tris-HCl, 0.25 mM EDTA, 50 mM NaCl, 0.2% NaN_3 and 1 mM DTT at pH 7.6; 50 mM KPi, 100 mM KCl, 0.2% NaN_3 and 1 mM DTT-d_{10} at pH 6.5. The final NMR samples contained 90% H_2O/10% D_2O with a protein concentration of 2 mM, each. Due to aggregation during aging, NMR spectra were recorded with freshly prepared samples.

NMR data collection

(1) *General*

All spectra were performed on Varian 750-, 600- and 500-Unityplus spectrometers equipped with a triple-resonance ^{1}H/^{13}C/^{15}N probe and a gradient amplifier. All experiments were conducted at 25 °C and at pH 6.5 and/or pH 7.6. Water signals were suppressed by using enhanced-sensitivity pulsed field gradient NMR techniques. Most spectra were collected with reduced spectral widths in the indirectly detected dimensions with States-TPPI hypercomplex (Marion et al., 1989a) phase incremention and initial delays of half-dwell time. Unless otherwise stated, in all heteronuclear 2D experiments, WALTZ16 modulation (Shaka et al., 1983) was used for ^{15}N decoupling and GARP (Shaka et al., 1985) was used for ^{13}C decoupling. The ^{1}H, ^{13}C and ^{15}N chemical shifts were referenced according to the method described by Wishart et al. (1995) using external DSS (7,3-(trimethylsilyl)-1-propanesulfonic acid, sodium salt = 0 ppm). The ^{1}H, ^{15}N, ^{13}C and ^{13}C′ carriers were centered at 4.77 ppm, 120.5 ppm, 46.5 ppm and 177.3 ppm, respectively. All NMR data were processed using the FELIX'95 program distributed by Molecular Simulations Inc. (San Diego, CA). After time domain convolution difference water reduction (Marion et al., 1989b) the data sets were apodized with either a 90° skewed sine-bell window function or a 90° skewed sine-bell squared window function in all dimensions. The baselines in the directly detected dimension were corrected with a polynomial function of fourth order.

(2) *^{1}H-^{15}N 2D NMR*

Two-dimensional, gradient enhanced, water flip-back ^{1}H-^{15}N HSQC spectra (Kay et al., 1992; Zhang et al., 1994) were acquired on uniformly ^{15}N- or ^{13}C/^{15}N-labeled samples in 90% H_2O/10% D_2O. Typically the data were collected with spectral widths of 13.5 ppm in t_1 (64 complex points) and 13.1 ppm in t_2 (512 complex points). The final digital resolution in the 512 × 128 matrix was 8.0 Hz/pt for ^{15}N and 9.6 Hz/pt for $^1H^N$.

Amide exchange rates were measured by dissolving the uniformly ^{15}N-labeled protein in D_2O at pH 6.5 and acquiring a series of 2D ^{1}H-^{15}N HSQC spectra at 9, 22, 41, 63, 93, 168, 288, 408 and 1213 min after H_2O exchange. Very slowly exchanging amide signals were observed in spectra obtained from a sample dissolved in D_2O for at least several days. $^3J(H^N\text{-}H^\alpha)$ coupling constants were measured semiquantitatively (>7 Hz) from a 2D ^{1}H-^{15}N HMQC-J experiment (Kay and Bax, 1990; Mueller, 1979). The spectrum was reduced to the same spectral widths used for all ^{1}H-^{15}N correlation measurements. The collection of 200 complex points results in an acquisition time (t_1) of 195 ms for ^{15}N. After zerofilling, the final digital resolution was 9.6 Hz/pt for ^{1}H and 2.0 Hz/pt for ^{15}N.

2D heteronuclear ^{15}N{^{1}H} NOE spectra (Farrow et al., 1994; Barbato et al., 1992) were recorded in the presence and absence of ^{1}H saturation (t_{sat} = 3 s). The spectra were collected with spectral widths of 18.2 ppm in t_1 (^{15}N; 160 complex points) and 12.8 ppm in t_2 (^{1}H; 512 complex points). Eighty transients per increment were acquired for suitable sensitivity and the recycle time was set to 2 s to ensure equilibrium ^{15}N magnetization. The final size of the matrix (512 × 256) results in a digital resolution for F_1 (^{15}N) and F_2 ($^1H^N$) of 6.3 Hz/pt and 3.6 Hz/pt, respectively.

(3) *Sequence-specific residue assignments*

The backbone nuclei ($^1H^N$, ^{15}N, $^{13}C^\alpha$, ^{13}C′, $^1H^N$) and the $^{13}C^\beta$-nuclei were assigned sequentially in two different manners using data from two sets of experiments, namely 3D HNCACB (Muhandiram and Kay, 1994; Kay et al., 1992; Wittekind and Mueller, 1993) and 3D CBCA(CO)NH (Muhandiram and Kay, 1994; Kay et al., 1992; Grzesiek and Bax, 1992a) as well as 3D HNCO (Muhandiram and Kay, 1994; Kay et al., 1990, 1992; Grzesiek and Bax, 1992b) and 3D HBCBCACOCAHA (Kay, 1993). The 3D HNCACB and 3D CBCA(CO)NH experiments were acquired at 750 MHz with spectral widths of 59.8, 13.5, and 13.1 ppm for F_1 ($^{13}C^{\alpha\beta}$; 44 complex points), F_2 (^{15}N; 25 complex points) and F_3 (^{1}H; 512 complex points), respectively, and with 128 scans per

increment. The $^{13}C'$ decoupling in the 3D HNCACB experiment and the $^{13}C^{\alpha}$ decoupling in the 3D CBCACONNH experiment was accomplished using a SEDUCE-1 decoupling scheme (McCoy and Mueller, 1992a,b) centered at 175 ppm and 54 ppm, respectively. Mirror image linear prediction was applied in the ^{15}N dimension. The digital resolutions of the $^{13}C^{\alpha\beta}$ dimension were 44.1 Hz/pt in the 3D HNCACB spectrum and 88.1 Hz/pt in the 3D CBCA(CO)NH spectrum. In both spectra the final digital resolutions in the ^{15}N and ^{1}H dimensions were 8.0 Hz/pt and 9.6 Hz/pt, respectively. The acquired data matrix of the 3D HNCO spectrum, which was recorded with 64 scans per increment, consisted of 52 (t_1, $^{13}C'$) $\times$ 25 (t_2, ^{15}N) $\times$ 512 (t_3, ^{1}H) complex points with spectral widths of 12.0 ppm in the $^{13}C'$ dimension, 13.5 ppm in the ^{15}N dimension and 13.1 ppm in the ^{1}H dimension. The digital resolution in the final 512 $\times$ 256 $\times$ 128 matrix was 8.9 Hz/pt for $^{13}C'$ (F_1), 8.0 Hz/pt for ^{15}N (F_2) and 9.6 Hz/pt for ^{1}H (F_3). The 3D HBCBCACOCAHA was recorded with 64 scans on the 750 MHz spectrometer. The acquisition times in t_1 ($^{13}C^{\alpha\beta}$; 44 complex points), t_2 ($^{13}C'$; 52 complex points) and t_3 (^{1}H; 512 complex points) were 3.9, 23.0 and 56.9 ms, respectively, with a final digital resolution of F_1 (44.1 Hz/pt), F_2 (17.7 Hz/pt) and F_3 (8.8 Hz/pt). Linear prediction was used to extend the $^{13}C^{\alpha\beta}$ dimension from 44 to 64 points followed by zerofilling to 256 complex points.

(4) *NOE assignments*

NOEs were assigned from the 3D ^{15}N-edited NOESY-HMQC spectrum (Marion et al., 1989c; Zuiderweg and Fesik, 1989) and 3D $^{15}N/^{13}C$-edited NOESY-HSQC spectrum (Pascal et al., 1994) with mixing times of 150 and 100 ms, respectively. The former spectrum was recorded with 24, the latter with 32 scans per increment. The spectral widths of the ^{15}N-edited NOESY-HMQC spectrum were 16.7 ppm in both ^{1}H dimensions and 15.1 ppm in the ^{15}N dimension. The acquired data matrix consisted of 128 (t_1; ^{1}H) $\times$ 32 (t_2; ^{15}N) $\times$ 512 (t_3; ^{1}H) complex points with a final digital resolution of F_1 (39 Hz/pt), F_2 (7.1 Hz/pt) and F_3 (9.8 Hz/pt). The 3D $^{15}N/^{13}C$-edited NOESY-HSQC spectrum was recorded on the 600 MHz spectrometer with spectral widths of 9000.9 Hz and 7500.5 Hz in both ^{1}H dimensions and 4527.0 Hz in the $^{15}N/^{13}C$-dimension. The final digital resolution was 29.3 Hz/pt (F_1; ^{1}H), 70.5 Hz/pt ($F_{2/3}$; $^{15}N/^{13}C$) and 8.8 Hz/pt (F_4; ^{1}H). Intraresidue proton chemical shifts were assigned using data from the 3D ^{15}N-edited TOCSY-HSQC experiment (Zhang et al., 1994) with the same spectral widths used in the ^{15}N-edited NOESY-HMQC spectrum. The mixing time was 57 ms and 24 scans per increment were employed. The digital resolution of the final matrix corresponds to that of the ^{15}N-edited NOESY-HMQC spectrum except that the digital resolution of the ^{15}N dimension was doubled. Experiments for sidechain assignments based on carbon TOCSY failed to give correlations beyond the β-position for most residues. Therefore, sidechain assignments were completed by comparing the ^{15}N-edited TOCSY and the $^{15}N/^{13}C$ NOESY experiments.

(5) *Structure calculations*

Structures were calculated with the program XPLOR 3.1 (Brunger, 1992) on a Silicon Graphics Octane with 175 MHz R-10 000 processors. We started with 40 extended structures (residues 1–112). The coordinates were sub-embedded and then fully embedded using 1052 NOE restraints, 58 dihedral restraints, and 46 hydrogen bond restraints. The list of NOE restraints comprised 406 intraresidue, 302 sequential, 152 medium range ($1 < |i - j| < 5$), and 146 long range ($|i - j| > 4$) restraints. The embedding was followed by hybrid distance geometry/simulated annealing and parameters for this step were the same as in the dgsa.inp example in the XPLOR 3.1 manual except we used 30 000 high temperature steps, 200 000 cooling steps and a 0.002 time step in the Verlet dynamics. Eight low-energy structures were chosen for additional simulated annealing refinement at 1000 K with cooling steps varied from 30 000 to 300 000.

Results and discussion

Spectra and assignments

Under the applied experimental conditions the N-terminal fragment of RPA70 remains stable in solution at pH 7.6 for several months, whereas at pH 6.5 the protein precipitates within a few weeks. Due to intermediate conformation exchange at the higher pH value, the spectra at pH 6.5 were needed to obtain additional assignment and NOE data. However, the dynamics of several loops are pH-dependent and the motion at pH 7.6 is probably more physiologically relevant than the motion at pH 6.5. Therefore most of the NMR experiments were carried out at pH 7.6 as well as at pH 6.5.

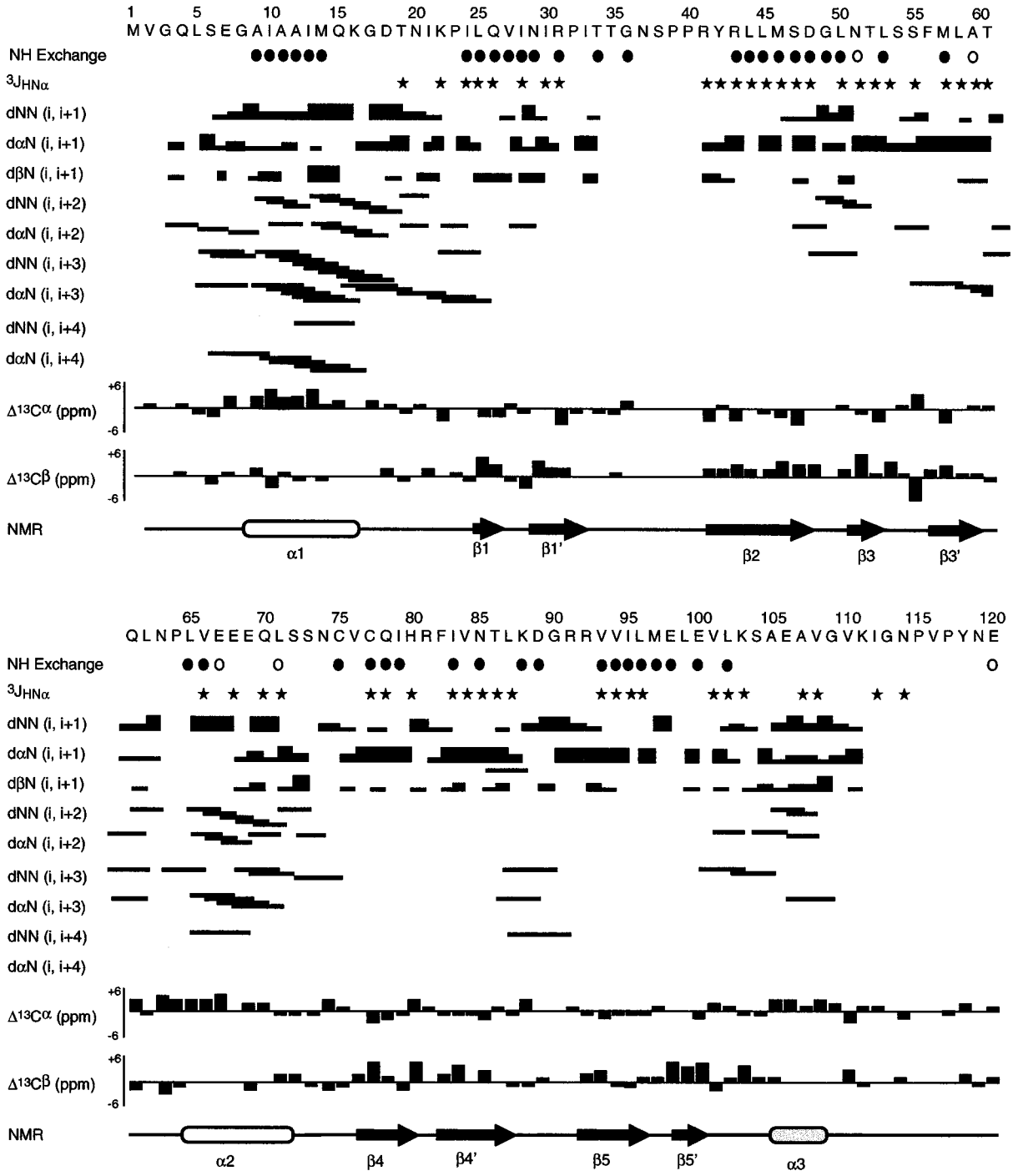

Figure 1. Diagram of the sequential and medium-range NOE connectivities, labile hydrogen exchange rates, three-bond ^3J (H^N-H^α) coupling constants and variation of $^{13}C^\alpha$ and $^{13}C^\beta$ chemical shifts from random coil values for residues 1 to 120 of RPA70Δ169. The NOE correlations were determined from the 3D ^{15}N-edited NOESY-HMQC and 3D $^{15}N/^{13}$C-edited NOESY-HSQC spectra at 25 °C and pH 6.5. The height of the bars reflects the strength of the NOE correlation as strong, medium or weak. Filled circles represent H^N nuclei that are still present after 20 h at pH 6.5, 25 °C, while open and gray circles are nuclei detectable after 93 and 408 min, respectively. ^3J (H^N-H^α) coupling constants ($\geq$7 Hz) are indicated by filled stars. Deviations of $^{13}C^\beta$ chemical shifts from random coil values are self-explanatory. Secondary structure deduced from the NOESY data is indicated below the NOE data: helix by oval boxes, transient helix by hatched oval boxes, β-strands by arrows, turns and loops by lines.

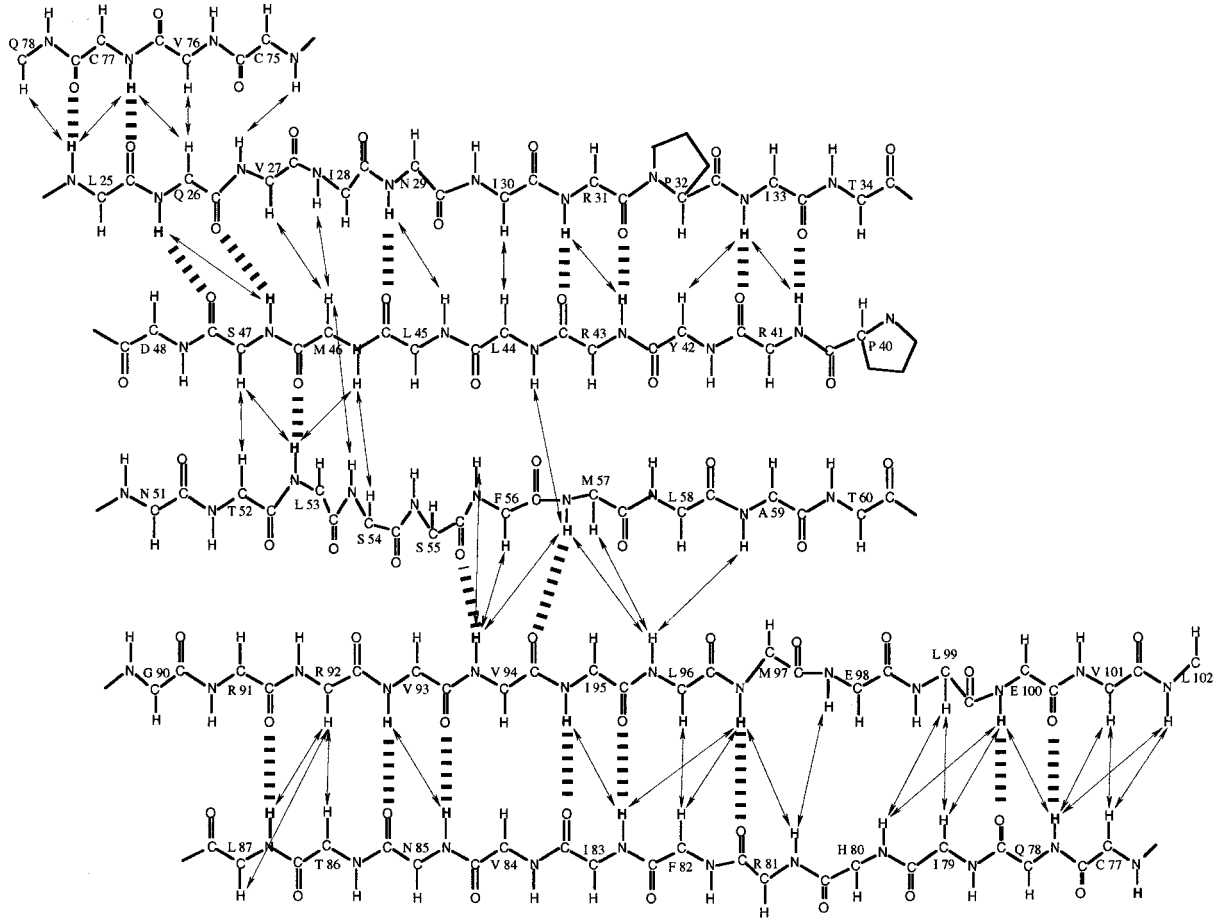

Figure 2. NOE contacts observed in the five-stranded β-sheet of the core of RPA70Δ169 and the hydrogen-bonding donor-acceptor pairs in the same region. Arrows denote dNN, dαN and dαα NOEs. Dashed lines represent hydrogen bonds derived from the hydrogen exchange and NOE data. Only amide residues with resonances remaining >20 h were considered for assignment as a hydrogen bond donor.

(1) *^{1}H, ^{13}C and ^{15}N resonance assignments*

The $^{1}H^{N}$ resonances were assigned primarily using the HNCACB and CBCA(CO)NH data. The amino acid types were identified by their $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts. Not all ^{1}H resonances of the amino acid side chains appeared in this spectrum. Using the data at pH 6.5, 10 additional residues were assigned, namely Gly-3, Asn-20, Thr-34, Thr-35, Gly-36, Tyr-42, Ser-54, Gln-61, Lys-88 and Asn-114. Except for breaks at prolines, we were able to correlate $^{13}C^{\alpha}$, $^{13}C^{\beta}$, ^{15}N and $^{1}H^{N}$ signals for the first 138 residues. Only the amide resonances of the residues Asn-37 and Ser-38 could not be assigned because of conformation exchange. The remaining unassigned resonances lack proton chemical shift dispersion and are assumed to be from residues 139–168. Proline $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts were obtained from the CBCA(CO)NH spectrum. Most of the assignments were verified by combining the 3D HNCO, 3D HBCBCACOCAHA and ^{15}N edited TOCSY-HSQC spectra. Since the 3D HNCO and 3D HBCBCACOCAHA spectra were only recorded at pH 7.6, not all connectivities could be confirmed. In addition, the C′ resonances of some residues could not be assigned to their intraresidue $^{13}C^{\alpha\beta}$ and $^{1}H^{\alpha}$ resonances due to crowded regions and degeneracy with the water resonance in the $^{13}C^{\alpha\beta}$-$^{1}H^{\alpha}$ planes of the 3D HBCBCACOCAHA spectrum.

(2) *Identification of secondary structure of RPA70Δ169*

Figure 1 diagrams the short- and medium-range NOEs, the $^{13}C^{\alpha\beta}$ chemical shift deviations from random coil values, ^{3}J (H^{N}-H^{α}) coupling constants ≥7 Hz as well as the slow and intermediate D_2O exchange rates of the N-terminal 120 residues of

Table 1. Restraint violation and energy statistics for 8 RPA-Δ169 structures

	Ensemble	Mean[a]
Energy (kcal/mol)		
Total	336.4 ± 23.5	314.0
Bond lengths	19.3 ± 2.2	18.2
Bond angles	185 ± 6.1	181.3
Improper	22.8 ± 3.5	19.0
van der Waals	34.3 ± 7.8	26.2
Distance restraint violation	74.4 ± 5.1	69.0
Dihedral restraint violation	0.25 ± 0.0	0.20
NOE violations		
No. > 0.1 Å	35 ± 4	33
Maximum violation (Å)	0.37 ± 0.04	0.41
Dihedral angle violations		
No > 0.1°	5 ± 1	6
Maximum (degrees)	0.52 ± 0.12	0.40
Rms deviations from ideality		
Bond lengths	0.0033 ± 0.0002	0.0032
Bond angles (degrees)	0.609 ± 0.010	0.603
Impropers (degrees)	0.414 ± 0.03	0.379
Rms deviations from experimental data		
Distance restraints (1052 NOE + 46 H-bond)	0.037 ± 0.001	0.036
Dihedral angle restraints (58)	0.079 ± 0.014	0.078

Procheck results on representative structure	
Most favored region	48.5%
Additionally allowed region	45.5%
Generously allowed region	6.1%
Disallowed region	0.0%

[a]Values are reported for the structure that best matches the experimental data.

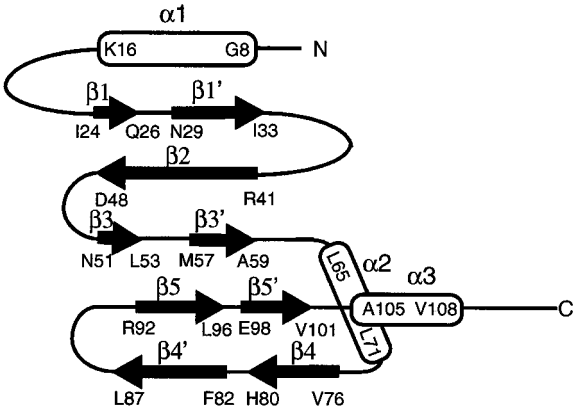

Figure 3. Schematic diagram of the secondary structure of RPA70Δ169. Residues are denoted by 1-letter amino acid code and residue number. Elements of secondary structure are labeled with Greek letters. Helices are indicated by oval boxes, β-strands by arrows, turn, loop and bulges by lines.

RPA70Δ169. These data were correlated and used to identify secondary structure. The segment α1 including residues Gly-8 to Lys-16 shows all the evidence for an α-helical conformation: strong dNN($i, i + 1$), weak dαN($i, i + 1$), strong dβN($i, i + 1$), medium-range NOEs, consistent $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ secondary chemical shifts and the absence of ^{3}J (H^{N}-H^{α}) coupling constants. The segments α2 and α3 including residues Leu-65 to Leu-71 and Ala-105 to Val-108, respectively, can also be identified as α-helices mainly because of strong dNN($i, i + 1$) NOEs, medium-range NOEs and the upfield $^{13}C^{\alpha}$ chemical shifts. However, in comparison to the α1-helix, fewer medium-range NOEs are observed and ^{3}J (H^{N}-H^{α}) coupling constants ≥7 Hz are found for some residues in the α2 and α3 segments indicating relatively unstable helices. The exchange rates for H^{N} within the helical sections demonstrate that amino acid residues in the α1-helix exchange their amide protons slowly whereas

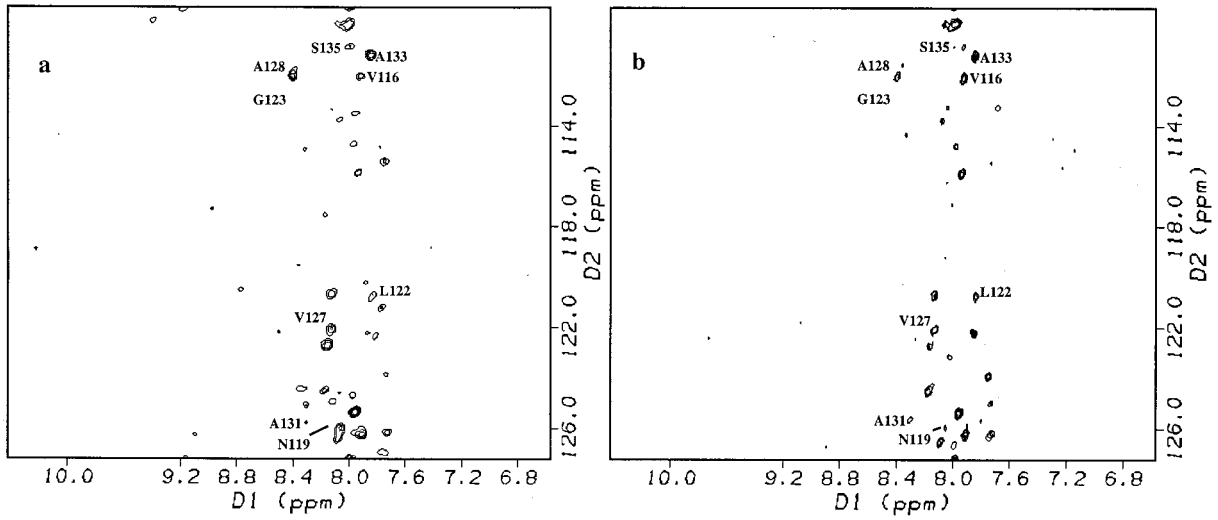

Figure 4. (a) A heteronuclear ^{15}N-^{1}H NOE spectrum labeled with assigned resonances from the C-terminal linker region. All labeled resonances are negative relative to a control spectrum, indicating fast internal backbone motion in the linker. Almost all other unlabeled resonances are also negative and not from the globular region, we assume they are also from the linker. (b) The longest (1665 ms) timepoint of the ^{15}N-T$_1$ experiment shows that the resonances in (a) also have long ^{15}N-T$_1$'s; another strong indication that the C-terminus is flexible. Both experiments were acquired at 500 MHz, 25 °C, recycle delay of 2.0 s using pulse sequences kindly provided by the Kay laboratory (Farrow et al., 1994).

no slow exchanging amides are observed in the α3-helix. Residues in the α2-helix have slow, intermediate and fast exchange rates.

In addition to the three helices, five segments characterized by strong dαN($i, i + 1$), weak or absent dNN($i, i + 1$), and few medium-range NOEs were observed indicating that the residues adopt extended conformations. Furthermore, the $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ secondary chemical shifts indicate extended strands. For most residues in these five segments the ^{3}J (H^{N}-H^{α}) coupling constants are $\geq$7 Hz. The pattern of NOE connectivities verifies that the five segments form a β-sheet. Figure 2 shows the NOE patterns for the sheet and Figure 3 shows the topology of RPA70Δ169. The β1-, β2- and β3-strands as well as the β4- and β5-strands are antiparallel, whereas the β5- and β3-strands are parallel. All β-strands except the β2-strand are broken by turns characterized by strong dNN($i, i+1$) NOEs and intraresidue dαN NOEs at residues Gln-26, Val-27, Ile-28, Leu-53, Ser-54, Phe-56, Met-57, His-80, Phe-82 and Glu-98 (data not shown). Furthermore, the $\Delta^{13}C^{\alpha}$ and $\Delta^{13}C^{\beta}$ values change signs due to transition from extended conformations to turn-like conformations. The intraresidue dαN NOEs of Leu-45, Ser-47, Asn-45, Cys-77, Asn-85 and Thr-86 demonstrate that residues within the strands are twisted. Finally, unambiguous connections between two corners of the sheet (β4-β1) indicate that the sheet forms a barrel. Strips of slowly exchanging amides are observed in the β1-, β2-, β4- and β5-strands, thus representing the central core of the five-stranded β-sheet. As can be seen in Figure 1, alternating fast and slow exchange rates are observed at Asn-29 (slow), Ile-30 (fast) and Arg-31 (slow) as well as Phe-82 (fast), Ile-83 (slow), Val-84 (fast), Asn-85 (slow) and Thr-86 (fast) as well as at Glu-98 (slow), Leu-99 (fast) and Glu-100 (slow). Therefore, the ends of the β1′, β4′ and β5′ strands are solvent exposed. Since hydrogen bonds across the β-strands also cause slow amide exchange rates, hydrogen bonds are drawn in Figure 2 whenever possible. Amides from residues Thr-52 to Ala-58 are mainly subject to fast exchange rates, except the amino acid Met-57 shows slow exchange and Leu-53 intermediate exchange. The β3 and β3′ strands in the middle of the β-sheet are destabilized possibly because of a half-twist. Together, the data indicate that this β-sheet forms a five-stranded antiparallel β-barrel with strand 3,3′ as the parallel strand necessary for barrels with an odd number of strands.

Both assigned and unassigned residues from Val-116 to Ala-168 have no observable ^{1}H-^{1}H NOE cross-peaks that would indicate secondary structure. The secondary shifts of the $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ resonances for these residues are also relatively small in this region of the protein (Table, Supplementary Material). Furthermore, these amides exchange rapidly with D_2O sol-

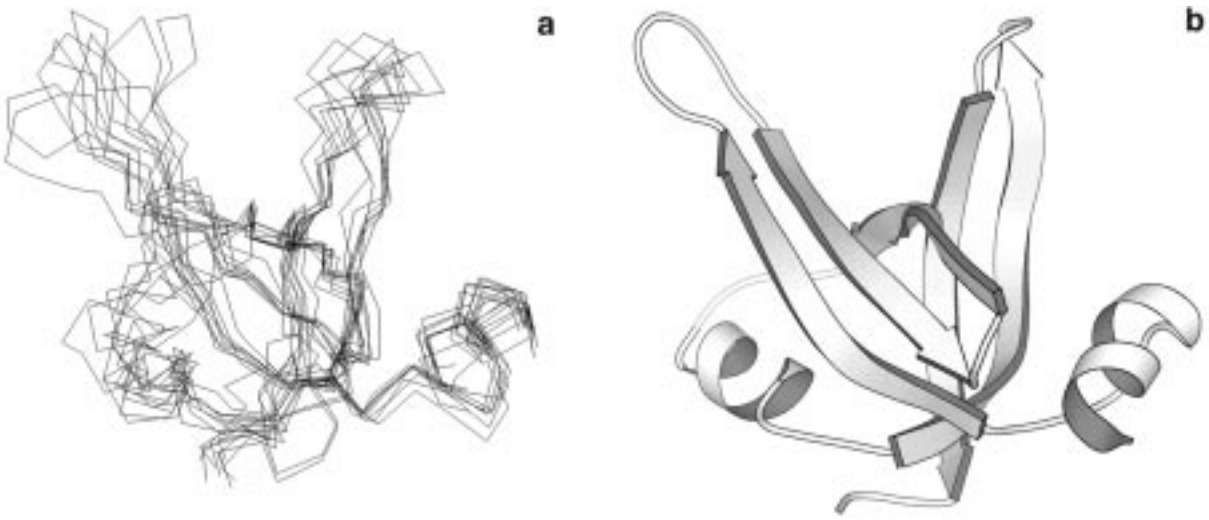

Figure 5. (a) C_α traces (residues 8–105) of the ensemble of 8 structures aligned with the mean structure. Only residues in the helices and sheet were used for the alignment. (b) The Molscript (Kraulis, 1991) ribbon diagram is shown for the representative structure in the same orientation as the ensemble.

vent and have intense exchange crosspeaks with water in the HSQC-TOCSY spectrum. The ^{15}N-^{1}H NOEs are negative and the ^{15}N-T_1's are long, therefore these amino acids are highly mobile (Figure 4a,b). Consequently, these data suggest that there is a significant degree of mobility in the protein for the C-terminal 53 residues of RPA70Δ169. Due to the high flexibility of these backbone amides, this segment presumably represents a linker to the tandem SSB domains in RPA70.

Calculation of the global fold

We initially calculated 40 structures for residues 1–114. Two folds resulted corresponding to the two possible ways that the β-barrel could form. Sixteen of the forty structures corresponded to a β-barrel without a hydrophobic core and only barely satisfying the core NOE restraint between Met-46 and Phe-56 sidechains. The remaining 24 structures corresponded to a β-barrel with a pronounced hydrophobic core and easily satisfied core restraint. We chose 8 of these 24 structures for further refinement. These 8 structures converged with no distance restraint violations >0.5 Å and no dihedral angle restraint violations $>1.0°$. The rmsd to the average structure was 3.32 Å for all non-hydrogen atoms and 2.83 Å for all backbone atoms. The β-core (25–32, 42–57, 76–86, 92–102) converged to 1.19 ± 0.35 Å while the helices and core (8–30, 42–57, 76-86) converged to 1.49 ± 0.38 Å (Figure 5, Table 1).

Description of the fold

The five-stranded antiparallel β-barrel, similar to *S. nuclease*, is capped on both ends by the two stable helices, as in papain domain 2 (Richardson, 1981). There are two long loops (residues 31–42 and 87–92) that extend from one side of the barrel and define a basic cleft of five arginines and one lysine (Figure 6). The loop comprising residues 31–42 is fluctuating on an intermediate timescale as indicated by pH sensitive peak intensities, missing peaks, and $T_1\rho$ measurements (Daughdrill and Lowry, unpublished observations).

Conclusions

In summary, we present the NMR assignment, secondary structure, and global fold of the N-terminal domain of RPA70, a fragment which plays an essential role in RPA-protein interactions. The N-terminal 8–108 amino acid residues form a five-stranded β-barrel capped on both ends by short helices. A 60 residue long flexible segment links the structured domain with the tandem DNA-binding domains of RPA70. This linker sequence is loosely associated with the structured domain and appears to mediate coupling or signalling between the N-terminal domain and the tandem SSB domains (Daughdrill et al., in preparation). Several yeast lethal, conditional lethal, and cell cycle control point mutants have mutations that map to

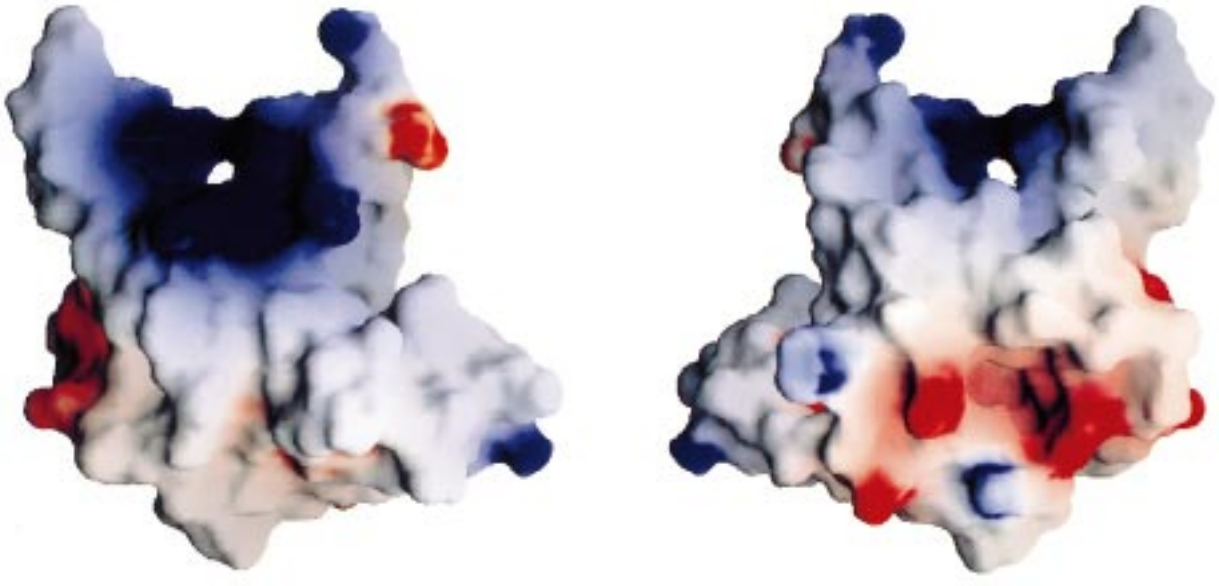

Figure 6. A GRASP (Nicholls et al., 1991) surface potential rendering of residues 8–105. Positive and negative charge potential surfaces are blue and red respectively. Molecules in the left and right views are rotated 180° about the vertical axis. The molecule on the left is in the same orientation as the structures in Figure 5. Note the large positively charged (blue) cleft.

the structured region and the basic cleft that we have identified. Point mutations have only recently been identified in the tail region and they cause relatively mild phenotypes, as might be expected for a flexible region.

The pronounced basic cleft of the β-barrel may be a binding surface that recognizes complementary acidic surfaces. In addition, the intermediate timescale motion in one of these loops and the clustering of deleterious point mutations at or near the basic cleft imply a functional role for this surface. The structural data presented here taken together with the biochemical and genetic data of other groups suggests the N-terminal domain of RPA70 may participate in coordinating replication and repair with the cell cycle by interacting with the acidic activation motifs of transcriptional regulators and repair proteins such as p53, VP16, Gal4 and XPG (Wold, 1997; Lin et al., 1996).

Acknowledgements

We thank G. Buchko and M. Kennedy for help with spectrometers and XPLOR. This work was performed under the auspices of the U.S. Department of Energy (Contract DE-AC06-76RLO1830). D.M.J. was supported by Associated Western Universities, Inc. under Grant DE-FG06-92RL-12451, DE-FG07-93ER-75912 or DE-FG07-94ID-13228 with the U.S. Department of Energy.

References

Barbato, G., Ikura, M., Torchia, D. and Bax, A. (1992) *Biochemistry*, **31**, 5269–5278.

Braun, K.A., Lao, Y., He, Z., Ingles, J. and Wold, M.S. (1997) *Biochemistry*, **36**, 8443–8454.

Brill, S.J. and Stillman, B. (1991) *Genes Dev.*, **5**, 1589–1600.

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

Erdile, L.F., Heyer, W.-D., Kolodner, R. and Kelly, T.J. (1991) *J. Biol. Chem.*, **266**, 12090–12098.

Fairman, M.P. and Stillman, B. (1988) *EMBO J.*, **7**, 1211–1218.

Farrow, N.A., Muhandiram, R., Singer, A.U., Pascal, S.M. and Kay, C.M. (1994) *Biochemistry*, **33**, 5984–6003.

Firmenich, A.A., Elias-Arnanz, M. and Berg, P. (1995) *Mol. Cell. Biol.*, **15**, 1620–1631.

Georgaki, A., Strack, B., Podust, V. and Huebscher, U. (1992) *FEBS Lett.*, **308**, 240–244.

Gomes, X.V. and Wold, M.S. (1995) *J. Biol. Chem.*, **270**, 4534–4543.

Grzesiek, S. and Bax, A. (1992a) *J. Am. Chem. Soc.*, **114**, 6291–6293.

Grzesiek, S. and Bax, A. (1992b) *J. Magn. Reson.*, **96**, 432–440.

Kanaar, R. and Hoeijmakers, J.H.J. (1998) *Nature*, **391**, 335–337.

Kay, L.E. and Bax, A. (1990) *J. Magn. Reson.*, **86**, 110–126.

Kay L.E., Ikura, M., Tschudin, R. and Bax, A. (1990) *J. Magn. Reson.*, **89**, 496–514.

Kay, L.E., Keifer, P. and Saarinen, T. (1992) *J. Am. Chem. Soc.*, **114**, 10663–10665.

Kay, L.E. (1993) *J. Am. Chem. Soc.*, **115**, 2055–2057.

Kraulis, P.J. (1991) *J. Appl. Crystallogr*. **24**, 946–950.
Lee, S.E., Moore, J.K., Holmes, A., Umezu, K., Kolodner, R.D. and Haber, J.E. (1998) *Cell*, **94**, 399–409.
Lin, Y., Chen, C., Keshav, K.F., Winchester, E. and Dutta, A. (1996) *J. Biol. Chem*., **271**, 17190–17198.
Longhese, M.P., Plevani, P. and Lucchini, G. (1994) *Mol. Cell. Biol.*, **14**, 7884–7890.
Marion, D., Ikura, M., Tschudin, R. and Bax, A. (1989a) *J. Magn. Reson.*, **85**, 393–399.
Marion, D., Ikura, M. and Bax, A. (1989b) *J. Magn. Reson.*, **84**, 425–430.
Marion, D., Kay, L.E., Sparks, S.W., Torchia, D.A. and Bax, A. (1989c) *J. Am. Chem. Soc*., **111**, 1515–1517.
McCoy, M.A. and Mueller, L. (1992a) *J. Am. Chem. Soc.*, **114**, 2108–2111.
McCoy, M.A. and Mueller, L. (1992b) *J. Magn. Reson.*, **98**, 674–679.
Muchmore, D.C., McIntosh, L.P., Russell, C.B., Anderson, D.E. and Dahlquist, F.W. (1989) *Methods Enzymol*., **177**, 44–73.
Mueller, L. (1979) *J. Am Chem. Soc.*, **101**, 4481–4484.
Muhandiram, D.R. and Kay, L.E. (1994) *J. Magn. Reson.*, **B103**, 203–216.
Nicholls, A., Sharp, K.A. and Honig, B. (1991) *Proteins Struct. Funct. Genet.*, **11**, 281–296.
Pascal, S.M., Muhandiram, D.R., Yamazari, T., Forman-Kay, J.D. and Kay, L.E. (1994) *J. Magn. Reson.*, **B103**, 197–201.
Richardson, J.S. (1981) *Adv. Protein Chem*., **34**, 167–339.
Shaka, A.J., Keeler, J. and Freeman, R. (1983) *J. Magn. Reson.*, **53**, 313–340.
Shaka, A.J., Barker, P. and Freeman, R. (1985) *J. Magn. Reson.*, **64**, 547–552.
Umbricht, C.B., Erdile, L.F., Jabs, E.W. and Kelly, T.J. (1993) *J. Biomol. Chem.*, **268**, 6131–6138.
Umezu, K., Sugawara, N., Chen, C., Haber, J.E. and Kolodner, R.D. (1998) *Genetics*, **148**, 989–1005.
Weisshart, K., Taneja, P. and Fanning, E. (1998) *J. Virology*, **72**, 9771–9781.
Wishart, D.S., Bigam, C.G., Yao, J., Abildgaard, F., Dyson, H.J., Oldfield, E., Markley, J.L. and Sykes, B.D. (1995) *J. Biomol. NMR*, **6**, 135–140.
Wittekind, M. and Mueller, L. (1993) *J. Magn. Reson.*, **B101**, 201–205.
Wobbe, C.R., Weissbach, L., Borowiec, J. A., Dean, F. B., Murakami, Y., Bullock, P. and Hurwitz, J. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 1834–1838.
Wold, M.S. and Kelly, T. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 2523–2527.
Wold, M.S. (1997) *Annu. Rev. Biochem.*, **66**, 61–91.
Zuiderweg, E.R.P. and Fesik, S.W. (1989) *Biochemistry*, **28**, 2387–2391.
Zhang, O., Kay, L.E., Olivier, P. and Forman-Kay, J.D. (1994) *J. Biomol. NMR*, **4**, 845–858.