Biochemistry **1999,** *38,* 15116−15128

Interactions of Human Nucleotide Excision Repair Protein XPA with DNA and RPA70ΔC327: Chemical Shift Mapping and ^{15}N NMR Relaxation Studies[†]

Garry W. Buchko,[‡,§] Gary W. Daughdrill,[‡] Robert de Lorimier,[§] Sudha Rao B. K.,[‡] Nancy G. Isern,[‡]
Jody M. Lingbeck,[||] John-Stephen Taylor,[||] Marc S. Wold,[⊥] Miriam Gochin,[#] Leonard D. Spicer,[§]
David F. Lowry,[‡] and Michael A. Kennedy*,[‡]

Pacific Northwest National Laboratories, Environmental Molecular Sciences Laboratory, Richland, Washington 99352, Departments of Biochemistry and Radiology, Duke University Medical Center, Durham, North Carolina 27710, Department of Biochemistry, University of Iowa College of Medicine, Iowa City, Iowa 52240, Department of Chemistry, Washington University, St. Louis, Missouri 63130, and Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143

Received July 28, 1999; Revised Manuscript Received September 9, 1999

ABSTRACT: Human XPA is an essential component in the multienzyme nucleotide excision repair (NER) pathway. The solution structure of the minimal DNA binding domain of XPA (XPA-MBD: M98-F219) was recently determined [Buchko et al. (1998) *Nucleic Acids Res. 26*, 2779−2788, Ikegami et al. (1998) *Nat. Struct. Biol. 5*, 701−706] and shown to consist of a compact zinc-binding core and a loop-rich C-terminal subdomain connected by a linker sequence. Here, the solution structure of XPA-MBD was further refined using an entirely new class of restraints based on pseudocontact shifts measured in cobalt-substituted XPA-MBD. Using this structure, the surface of XPA-MBD which interacts with DNA and a fragment of the largest subunit of replication protein A (RPA70ΔC327: M1-Y326) was determined using chemical shift mapping. DNA binding in XPA-MBD was highly localized in the loop-rich subdomain for DNA *with or without a lesion* [dihydrothymidine (dhT) or 6−4-thymidine-cytidine (64TC)], or with DNA in *single- or double-stranded form*, indicating that the character of the lesion itself is not the driving force for XPA binding DNA. RPA70ΔC327 was found to contact regions in *both* the zinc-binding and loop-rich subdomains. Some overlap of the DNA and RPA70ΔC327 binding regions was observed in the loop-rich subdomain, indicating a possible cooperative DNA-binding mode between XPA and RPA70ΔC327. To complement the chemical shift mapping data, the backbone dynamics of free XPA-MBD and XPA-MBD bound to DNA oligomers containing dhT or 64TC lesions were investigated using ^{15}N NMR relaxation data. The dynamic analyses for the XPA-MBD complexes with DNA revealed localized increases and decreases in S^2 and an increase in the global correlation time. Regions of XPA-MBD with the largest increases in S^2 overlapped regions having the largest chemical shifts changes upon binding DNA, indicating that the loop-rich subdomain becomes more rigid upon binding DNA. Interestingly, S^2 decreased for some residues in the zinc-binding core upon DNA association, indicating a possible concerted structural rearrangement on binding DNA.

Genomic DNA is constantly being damaged by exogenous and endogenous agents which subsequently lead to carcinogenesis, aging, and genetic diseases (*1*, *2*). Nucleotide excision repair (NER)[1] is a ubiquitous cellular response to certain types of DNA damage and is responsible for removing many structurally unrelated DNA lesions from the genome (*3*, *4*). At least 30 polypeptides are involved in NER (*5*) with mutations to NER genes associated with three identified cancer-prone genetic instability syndromes: trichothiodystrophy, Cockayne's syndrome, and xeroderma pigmentosum (XP) (*1*). One class of the latter syndrome, an autosomal recessive disease characterized by sun (UV) hypersensitivity, is a result of mutations to the XPA gene, a gene that encodes for a 31 kDa protein of 273 amino acids (*6*, *7*). The protein XPA is essential for NER activity, playing a central, multifunctional role in the process (*3*, *4*, *8*). XPA has been reported to preferentially bind damaged over

---

[†] This work was performed under the auspices of the U.S. Department of Energy (Contract DE-AC06-76RLO1830) and was supported by the Department of Energy Office of Biological and Environmental Research Program under Grant 249311 KP11-01-01.

* To whom correspondence should be addressed. Phone: (509) 372-2168. Fax: (509) 376-2303. E-mail: ma_kennedy@pnl.gov.

[‡] Pacific Northwest National Laboratories.

[§] Duke University Medical Center.

[||] Washington University.

[⊥] University of Iowa College of Medicine.

[#] University of California.

[1] Abbreviations: NER, nucleotide excision repair; XPA-MBD, minimal DNA-binding domain of human NER protein XPA (M98-F219); CoXPA-MBD, cobalt-substituted XPA-MBD; RPA, human replication protein A; RPA70, 70 kDa subunit of RPA; RPA70ΔC169 (M1−A168); RPA70ΔC327 (M1-Y326); T_1, longitudinal relaxation time; T_2, transverse relaxation time; S^2, generalized order parameter; τ_m, overall correlation time; τ_e, effective correlation time; R_{ex}, ^{15}N exchange contribution; RMSD, root mean standard deviation; ni, number of increments; sw, sweep width; nt, number of transients; NOE, nuclear Overhauser effect; HSQC, heteronuclear single quantum coherence; d9, dCCAATAACC; dhT, dihydrothymidine; 64TC, 6-4-thymidine-cytidine photoproduct; DTT, dithiothreitol.

10.1021/bi991755p CCC: $18.00

undamaged DNA (*9−11*) and is known to interact with other NER proteins including ERCC1 (*12*), TFIIH (*13*), RPA (*14*), and XPC-HR23B (*15*, *16*). While implicated in recognizing and/or verifying damaged DNA (*16*) and recruiting other proteins to repair the damage, the mechanisms by which XPA performs its multifunctional roles are not understood.

Because XPA has been shown to have an affinity for damaged DNA (*9−11*) and an extensive variety of chemically and structurally distinct DNA lesions are repaired by NER (*2*, *17*, *18*), there is considerable interest in understanding the mechanism of the XPA−DNA interactions. The minimal region of XPA that can still bind DNA has been pinpointed to a 122 amino acid region between M98 and F219 (XPA-MBD) (*19*). Sequence analysis (*7*) and EXAFS spectroscopy (*20*) show that XPA-MBD contains a single class IV, C4-type, zinc-binding motif, C105-X_2-C108-X_{17}-C126-X_2-C129. The three-dimensional structure of the 122-residue minimal DNA-binding domain of XPA has recently been determined using multidimensional NMR protocols in two laboratories (*21*, *22*). In general, the major features of the structures generated by both groups are preserved.

Insights into the structural basis for the functions of XPA may be obtained by chemical shift mapping (*23−25*). Protein−ligand interactions often produce changes in the chemical environment of the nuclei at the interface of ligand binding which are accompanied by perturbations in the chemical shifts of the backbone $^1H^N$ and ^{15}N resonances (*24−26*). By identifying the resonances that undergo a binding-dependent chemical shift or intensity perturbation, it is possible to map the location of the binding site onto the three-dimensional structure of the protein.

A better understanding of the role of XPA in NER may also be obtained from a detailed description of the internal protein dynamics. Many biological functions, such as protein−ligand interactions, protein folding, and enzyme catalysis, are dependent on the protein's intramolecular motions (*27*). Using isotopically ^{15}N-enriched proteins and advances in heteronuclear NMR relaxation techniques, it is possible to probe the motions of individual protein residues on the picosecond to second (or longer) time scale (*28−31*). Experimentally measured $^1H^N$-^{15}N steady-state NOEs and ^{15}N longitudinal (T_1) and transverse (T_2) relaxation times can be used to describe the motion of individual NH bond vectors, using the model-free formalism of Lipari and Szabo (*32*, *33*), in terms of generalized order parameters (S^2), effective correlation times for internal motion (τ_e), and ^{15}N exchange contributions (R_{ex}). An optimized value for the overall molecular rotational correlation time, τ_m, is also obtained from the model-free analysis of the relaxation data.

As part of our ongoing efforts to better understand the functions of XPA: (i) the solution structure of free XPA-MBD was refined using additional distance restraints obtained from 4D NMR experiments and pseudocontact shift restraints measured in cobalt substituted XPA-MBD (*34*), (ii) the DNA and RPA70 binding surfaces on uniformly ^{15}N-labeled XPA-MBD were investigated by chemical shift mapping, and (iii) the backbone dynamics of uniformly ^{15}N-labeled XPA-MBD were investigated using two-dimensional inverse detected heteronuclear $^1H^N$-^{15}N NMR spectroscopy. Chemical shift maps for XPA-MBD bound to (i) single-stranded DNA containing a single dihydrothymidine adduct (dhT), (ii) double-stranded DNA containing a single dihydrothymidine adduct (dhT2x), (iii) undamaged single-stranded dhT (the dihydrothymidine replaced with a thymine) (d9), and (iv) double-stranded DNA containing a 6-4 photoproduct of 5′-thymidine-cytosine-3′ (64TC) are presented. The chemical structure of the DNA lesion in these DNA oligomers represent grossly (64TC) (*35−37*) and mildly (dhT) (*38, 39*) DNA distorting lesions. XPA is known to bind DNA more tightly in the presence of RPA (*40*, *41*), a heterotrimer composed of a 70, 32, and 14 kDa subunit. Two XPA-binding regions have been identified in RPA, a strong site at the C-terminus of RPA32 and a weaker site between residues 236−382 in RPA70 (*42*). Mapping experiments for XPA binding to a fragment of RPA70 that includes the N-terminal 61% of the weak XPA-binding region, RPA70ΔC327, are also included here. In addition to the chemical shift mapping, relaxation data for free XPA-MBD and XPA-MBD associated with dhT, dhT2x, and 64TC are presented and compared to the chemical shift maps. Together, the chemical shift mapping and relaxation data are discussed in relation to XPA's role in NER and in relation to models of how XPA interacts with DNA and RPA.

MATERIALS AND METHODS

Expression and Purification of XPA-MBD and Fragments of RPA70. Uniformly ^{15}N-labeled XPA-MBD was prepared as described previously (*43*) and exchanged into 90% H_2O/10% D_2O in the following NMR buffer: 20 mM K_2HPO_4, 100 mM KCl, 25 mM perdeuterated DTT, 50 μM NaN_3, pH 7.3. XPA-MBD containing cobalt (CoXPA-MBD) instead of zinc was prepared similarly by substituting cobalt nitrate for zinc acetate when inducing protein expression in minimal medium (*20*). The purified CoXPA-MBD sample was exchanged into the same NMR buffer with 2.0 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) substituting for 25 mM DTT. Fragments containing the N-terminal 168 and 326 amino acid residues of the RPA70 subunit, RPA70ΔC169 and RPA70ΔC327, respectively, were prepared following established protocols (*44*) and exchanged into the XPA-MBD NMR buffer.

Synthesis of the DNA Dodecamer Containing the 64TC Photoproduct. Both d(GAGTATCATGAG) and its complimentary strand d(CTCATGATACTC) were purchased from Integrated DNA Technologies, Inc. (Coralville, IA). The purity of the sequences was verified by one-dimensional ^1H NMR spectroscopy. The 64TC photoproduct was prepared by the irradiation of 30 mL aliquots of a 6.2 μM solution of d(GAGTATCATGAG) in double-distilled H_2O. Each aliquot was first degassed by bubbling N_2 through the solution for 5 min and then transferring the solution to a polystyrene Petri dish and sealing it in a transparent plastic bag (Gladlock Zipper, Oakland, CA) purged with N_2. The DNA was exposed, on ice, for 3 h to 254 nm light at a distance of 6 cm from a Spectroline model XX-15F lamp (1500 $\mu W/cm^2$) fitted with a Spectroline Longlife 250−375 nm filter (Westbury, NY). The irradiated DNA was lyophilized, redissolved in H_2O, and then injected onto a semipreparative, Dynamax C-18 column (300 Å, 5 μm particle size, 10 mm ID × 25 cm L) (Woburn, MA). The 64TC photoproduct was isolated over 50 min using a 12−32% methanol gradient in 75 mM phosphate buffer (pH 6.6) at a flow rate of 1.7 mL/min. The 64TC product eluted at 34 min with an overall yield of 13%. The DNA was then desalted by reinjection on the same column preequilibrated with water (eluting with 50% acetonitrile). The pure 64TC dodecamer (1 mg) was annealed to an equimolar amount of the complimentary strand in 300 μL of 10 mM Na_2HPO_4, (pD 7.0), 0.01% NaN_3, and 250 mM NaCl in either 100% D_2O (Aldrich, Milwaukee, WI) or 10% D_2O in H_2O.

Synthesis of the DNA Nonanomer Containing Dihydrothymidine. The free base, 5,6-dihydrothymidine, was purchased from the Sigma Chemical Co. (St. Louis, MI) and converted into the 5′-*O*-4,4′-dimethoxytrityl-3′-β-cyanoethyl phosphoramidite following established protocols (*45*, *46*). The dhT-phosphoramidite was then incorporated into the sequence dCCAA(dhT)AACC (10 μmol scale) using solid-state phosphoramidite chemistry with an Applied Biosystems 380B DNA synthesizer (Foster City, CA). The product was purified and desalted on a Waters semipreparative C-18 reversed-phase column (Milford, MA). The sample purity was verified by mass spectroscopy and one-dimensional ^{1}H NMR spectroscopy. Half of the purified dhT nonanomer was then annealed to an equimolar amount of the complimentary strand to generate the double-stranded product, dhT2x. The complimentary strand to dhT and a nonanomer containing a normal dT in the position of dhT, dCCAATAACC (d9), were synthesized following a similar protocol.

NMR Spectroscopy. All the NMR data were collected at 25 °C on Varian 750-, 600-, and 500-Unityplus spectrometers equipped with a triple-resonance ^{1}H/^{13}C/^{15}N probe and a gradient amplifier. Aliquots of RPA70ΔC169 and RPA70ΔC327 were titrated into a 0.2 mM XPA-MBD sample in molar ratios of 0:1, 0.25:1, 0.5:1, 1:1, and 2:1 (RPA70:XPA-MBD). Aliquots of desalted d9, dhT, and dhT2x were titrated into a 1.0 mM XPA-MBD sample stepwise (4 or 5 additions) until the DNA and protein were approximately equimolar. Because of the limited amount of 64TC and the difficulty in its preparation, titration of 64TC into the protein was not practical. Instead, an ∼1:1 molar ratio of XPA-MBD and 64TC was prepared directly in a volume of 220 μL in a Shigemi tube (Tokyo, Japan). A 1.0 mM sample of CoXPA-MBD was prepared to obtain cobalt pseudocontact shifts. Two-dimensional ^{1}H-^{15}N HSQC (*47*, *48*) spectra were collected for CoXPA-MBD and after each addition of RPA70 or DNA to assist mapping of the changes in the $^1H^N$ and ^{15}N chemical shifts. The $^1H^N$ and ^{15}N cross-peaks in the ^{1}H-^{15}N HSQC spectrum of free XPA-MBD were assigned previously (*43*). To verify the assignment of the XPA-MBD ^{1}H-^{15}N HSQC cross-peaks after the final addition of DNA and for CoXPA-MBD, three-dimensional ^{15}N-edited NOESY-HSQC (*48*) spectra were recorded. The HSQC data was compared to the published sequence-specific backbone and side-chain resonance assignments (^{15}N, $^1H^N$, $^1H^\alpha$, $^1H^\beta$, $^1H^\gamma$, and $^1H^{other}$) for free XPA-MBD (*21*).

Additional NOE-based distance restraints were obtained from 4D ^{13}C-^{13}C (*49*)- and ^{15}N-^{15}N (*50*, *51*)-edited NOESY spectra collected at a proton resonance frequency of 600 and 750 MHz, respectively. For the 4D ^{13}C-^{13}C-edited experiment, 58 complex t_1, 16 complex t_2, and 16 complex t_3 increments were recorded using four scans, with spectral widths of 6000 Hz (F1,^{1}H), 4200 Hz (F2, ^{1}H), 3000 Hz (F3,^{13}C), and 3000 Hz (F4,^{13}C). For the 4D ^{15}N-^{15}N-edited experiment 112 complex t_1, 12 complex t_2, and 21 complex t_3 increments were recorded using four scans, with spectral widths of 13 750 Hz (F1,^{1}H), 9870 Hz (F2, ^{1}H), 1500 Hz (F3,^{1}H), and 1750 Hz (F4,^{1}H). Both data sets were zero filling to 512 × 128 × 64 × 64K. With the exception of the 4D ^{15}N-^{15}N-edited NOESY experiment, all the pulse sequences were obtained by permission from the Lewis Kay pulse sequence library (University of Toronto).

Solution Structure Refinement of XPA-MBD Using Co^{2+} Pseudocontact Shifts. An unpaired electron at a paramagnetic metal center can interact with the surrounding nuclear environment in two major ways. One is a scalar interaction, known as the contact shift, which propagates no further than four or five bonds from the metal center. The second is a dipolar (through space) interaction which is manifested in a pseudocontact shift of neighboring nucleic that has a $1/r^3$ distance dependence (*34*). Consequently, pseudocontact shifts contain distance information that extends up to 20 Å from the metal site, unlike proton−proton NOEs that have a $1/r^6$ distance dependence and contain distance information that extends out to 5−7 Å. If the magnitude and orientation of the **g**-tensor and the initial coordinates of the effected nuclei are known, the long-distance information in the pseudocontact shifts can be used to refine NMR solution structures (*34*). The substitution of Zn^{2+} with Co^{2+} in the metal center of a protein is expected to have little effect on the protein's solution structure (*34*). Indeed, EXAFS spectroscopy showed that the metal in XPA-MBD (zinc) and CoXPA-MBD was tetrahedrally coordinated to the sulfur atom of four cysteine residues with metal-sulfur bond lengths that were identical at 2.34 Å (*20*).

The magnitude and orientation of the **g**-tensor was determined using PSEUDOC, a least-squares fitting method that uses paramagnetic chemical shifts and PDB coordinates to calculate the axial ($\Delta\chi_{ax}$) and equatorial ($\Delta\chi_{eq}$) values of the susceptibility tensor (*34*). This calculation places the metal center at the origin of a Cartesian coordination system with the *z*-axis going through the C8 sulfur atom, the *y*-axis through the C11 sulfur atom, and the *x*-axis perpendicular to the *z*- and *y*-axes. Using 115 pseudocontact shifts from $^1H^N$, ^{15}N, and $^1H^\alpha$ nuclei, the following values were obtained for the five $\Delta\chi$-tensor parameters after 11 iterations: $\Delta\chi_{ax} = -133.1$ ppm, $\Delta\chi_{eq} = -1470$ ppm, $\alpha = 28$, $\beta = 110$, and $\gamma = -109$. The RMSD for the calculated versus the observed chemical shifts was 0.05 ± 0.04 ppm. These $\Delta\chi$-tensor values were used for the structure refinement.

Initial coordinates for XPA-MBD were obtained from the average, previously published structures for the two subdomains (*21*). These initial structures were refined using an X-PLOR program (*52*) modified to include distance restraints based on Co^{2+} pseudocontact chemical shifts of $^1H^N$, $^1H^\alpha$, and ^{15}N nuclei. The distance and dihedral angle restraints used to generate the initial structures (*21*), along with an additional 144 long distance restraints (i to $i+4 > 5$) obtained from subsequent 4D ^{15}N-^{15}N- and ^{13}C-^{13}C-edited NOESY experiments and 115 distance restraints based on Co^{2+} pseudocontact shifts, were used in the refinement. This process involved iterative calculations using 1×10^4 simulated steps and annealing temperatures of 500 K to reach energy minimum. Such a schedule monotonically reduced the total energy from 2150 to 800 kcal and the paramagnetic shift energy from 955 to 250 kcal. Using this approach, an ensemble of 10 structures was calculated. The average pairwise RMSD of these 10 structures to the mean structure was 1.13 ± 0.30 Å for the backbone atoms and 1.81 ± 0.24

Å for all atoms (F100−V206). The introduction of the Co^{2+}-based distance restraints into the refinement calculations did not result in any additional violations in the NOE-based distance restraints. PROCHECK Ramachandran analysis of the 10 structures showed only 2.2% of the residues were in disallowed regions. Simultaneous calculations made without Co^{2+} pseudocontact shift restraints resulted in an ensemble of 10 structures whose final energies were 250 kcal lower, but, converged poorly (>4 Å RMSD for the backbone atoms). The refined structures were deposited in the RCSB Protein Data Bank, PDB code 1D4U.

Backbone ^{15}N Relaxation Measurements. All the two-dimensional ^{1}H-^{15}N HSQC spectra collected for the ^{15}N relaxation measurements were obtained at a proton resonance frequency of 500.13 MHz, 25 °C, using identical spectrometer parameters (2K data points, ni = 64). The data were processed with Felix97 (MSI, San Diego, CA) software.

Nitrogen-15 T_1 and T_2 values were measured in triplicate using previously described experiments (*30*). T_1 values were obtained by using eight time delays of 5.55, 72.2, 161, 272, 405, 583, 838, and 1280 ms, while T_2 values were obtained by using seven time delays of 15.8, 32.6, 47.3, 78.9, 110, 142, and 189 ms. Both T_1 and T_2 were obtained by fitting the measured peak volumes to a two-parameter function

$$I(T) = I_o \exp(-T/T_i)$$

where $I(T)$ is the intensity after a delay of time T, I_o is the intensity corresponding to the ^{15}N magnetization at the beginning of the relaxation period T, and T_i represents the spin−spin ($i = 1$) and spin−lattice ($i = 2$) relaxation times.

The steady-state $^{1}H^{N}$-^{15}N NOE values (*30*) were measured in duplicate from the ratios of the cross-peak volumes in spectra recorded with (I_{sat}) and without (I_{unsat}) 3 s of proton saturation.

$$\mathrm{NOE} = I_{sat}/I_{unsat}$$

Initial Estimation of the Overall Rotational Correlation Time, τ_m. An estimation of the overall correlation time (τ_m) was determined from the T_1/T_2 ratios for each residue (*28, 29*) using the precaution of excluding values with an T_1/T_2 ratio ±1 standard deviation from the mean. In free XPA-MBD, T_1/T_2 ratios for 80 residues resulted in an average τ_m of 9.27 ± 2.28 ns. Nineteen residues greater than or less than ±2.28 ns from the mean were removed and the average of the smaller set was 9.27 ± 1.34 ns. Similarly, τ_m estimates of 19.76 ± 4.21 (58), 18.31 ± 4.01 (53), and 13.97 ± 1.67 (51) ns were obtained for XPA-MBD in the presence of dhT, dhT2x, and 64TC, respectively (number in parentheses is the number of residues available to calculate the average).

MODELFREE Analysis of Internal Motion. In proteins, it is generally assumed that the relaxation properties of the backbone amide ^{15}N nuclear spins are governed by dipolar interactions with directly bonded amide protons and by chemical shift anisotropy effects. Under such assumptions, the spectral density function, $J(\omega)$, can be modeled on different time scales with the measured relaxation parameters T_1, T_2, and the $^{1}H^{N}$-^{15}N steady-state NOE (*29, 32, 33*). The program MODELFREE (version 4.0) (*31*) uses these Lipari-Szabo expressions to calculate the motional parameters S^2 [the generalized order parameter, where $S^2 = S_f^2 S_s^2$, the square of the order parameters for internal motions on the fast (f) and slow (s) time scales], τ_e (the effective correlation time for internal motion), and R_{ex} (the ^{15}N chemical exchange term) using the relaxation data as input. The motional parameters were optimized for five models (Table 1) starting with the estimated value for the global τ_m. For each residue, the simplest model meeting the goodness-of-fit criteria outlined by Mandel et al. (*31*) was chosen using a fixed value for τ_m. After selecting the model for each ^{15}N nuclear spin, both τ_m and the internal motional parameters were optimized simultaneously. If $\Delta\tau_e > \tau_e$, then a simpler model was substituted and the simultaneous optimization step repeated. Values for τ_m were optimized iteratively by fixing the optimized τ_m value calculated by MODELFREE and reselecting the model for each ^{15}N nuclear spin using the selection criteria. The process was repeated until the optimized value for τ_m varied by < 0.2 ns. MODELFREE analysis was not performed on XPA-MBD in the presence of dhT2x because the sample deteriorated before the $^{1}H^{N}$-^{15}N heteronuclear NOE data could be collected.

Table 1: Summary of the Model-Free Spectral Density Models Used to Fit T_1, T_2, and ^{1}H-^{15}N Heteronuclear NOE Data for Free XPA-MBD and XPA-MBD Associated with Single-Stranded DNA Containing Dihydrothymidine (dhT) and Double-Stranded DNA Containing the 64TC Photoproduct (64TC)

model parameters	XPA-MBD	dhT	64TC
S^2	54	40	53
S^2 and τ_e	5	7	3
S^2 and R_{ex}	4	0	2
S^2, τ_e, and R_{ex}	0	0	0
S_s^2, S_f^2, and τ_e	7	14	3
not fit	7	4	6
total	77	65	67

RESULTS

Solution Structure of XPA-MBD. Figure 1A is a superposition of the 10 calculated structures with the lowest energies which illustrates the convergence of the final set of refined structures. Figure 1B is a ribbon diagram of the average structure of the ensemble with the major secondary features labeled. Details of the structural features of XPA-MBD have been discussed previously (*21, 22*). Figure 2A plots the pairwise RMSD to the mean structure for each residue and illustrates the improvements to the calculated structures afforded by the additional NOE- and Co^{2+}-based restraints. The solid line is the pairwise RMSD for the refined XPA-MBD structures and the dashed line is the pairwise RMSD for the two subdomains calculated prior to the refinements (*21*). Even with the additional restraints, the uncertainty in the structures is still localized to regions in the C-terminal subdomain. Such uncertainty is reflected in a smaller RMSD of 0.61 ± 0.12 Å when the disordered loops in the “loop-rich” subdomain are removed from the backbone RMSD calculation. Figure 2B shows the $^{1}H^{N}$ and ^{15}N residues shifted by the substitution of cobalt for zinc at the metal center. As expected, the strongest shifts are localized in the zinc-binding subdomain. Note that while restraints obtained from the latter data played a substantial role in generating structures that converged better, the additional NOE-based restraints from the new 4D data sets were also necessary.

Chemical Shift Mapping of the DNA-Binding and RPA-Binding Interfaces of XPA-MBD. Figure 3 is a superposition

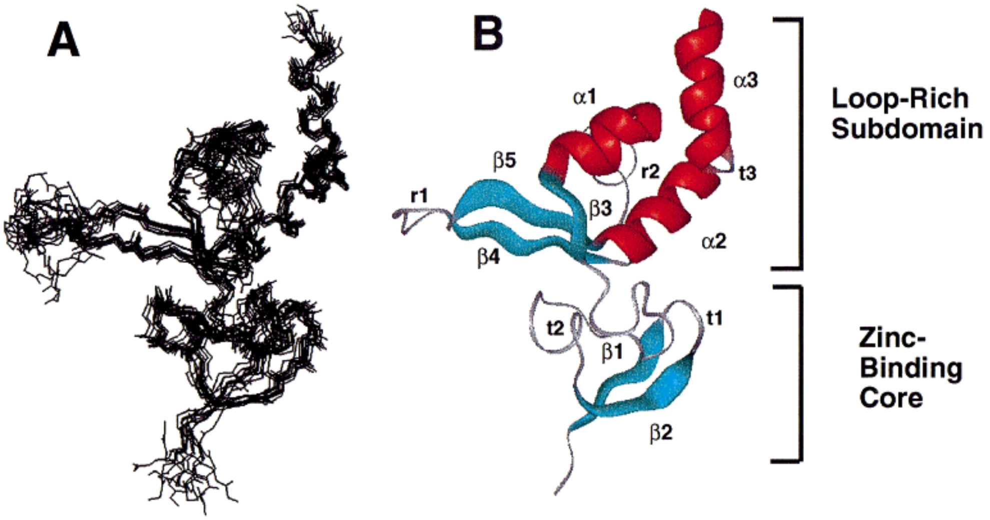

FIGURE 1: (A) Conformational ensemble of the backbone atoms ($N-C^{\alpha}-C=O$) of 10 refined structures for XPA-MBD superimposed on the average structure. (B) Ribbon representation of the average structure of the ensemble on the left with the secondary structural features labeled. Residues D101−K137 = zinc-binding core; residues L138−F219 = loop-rich subdomain.

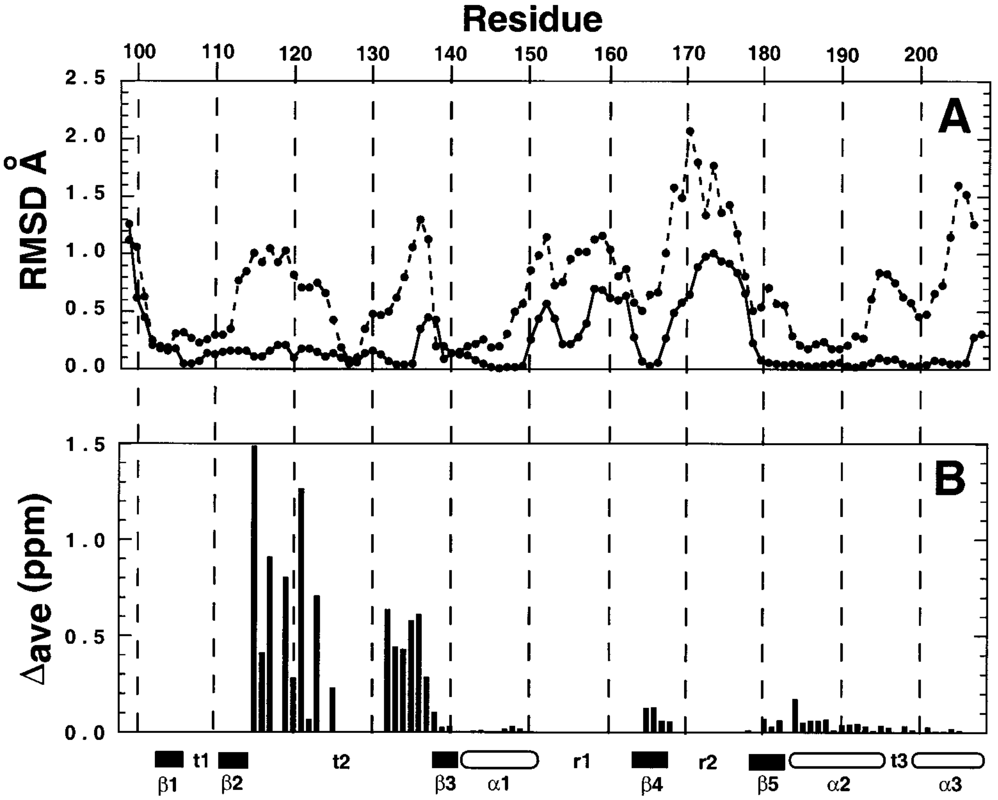

FIGURE 2: (A) Plots of the mean pairwise RMSDs to the average structure for each residue of XPA-MBD. The solid line is for the ensemble in Figure 1A, the dashed line is for the previously published structures for the individual subdomains (*21*). The data were generated by moving a window of three residues along the sequence and plotting the mean pairwise RMSD (Å) over the central residue. (B) Combined average chemical shift changes in the $^{1}H^{N}$ and ^{15}N resonances of XPA-MBD upon substituting Zn^{2+} with Co^{2+} at the metal center. The average chemical shift change = $\Delta_{ave} = [((\Delta^{1}H^{N})^{2} + (\Delta^{15}N/5)^{2})/2)]^{1/2}$.

of the ^{1}H-^{15}N HSQC spectra of XPA-MBD in the absence and presence of dhT2x. A subset of the HSQC cross-peaks observed for free XPA-MBD (red) shift upon the addition of an equimolar concentration of dhT2x (blue). Figure 4A summarizes the average chemical shift changes effected in XPA-MBD upon the addition of d9, dhT, dhT2x, and 64TC. In general, all four DNAs perturb the chemical shifts of the same $^{1}H^{N}$ and ^{15}N resonances and the largest chemical shift perturbations are clustered in the loop-rich subdomain: the triple-strand antiparallel β-sheet, loop r2, and helices α1 and α3. *Most interestingly, these perturbations occur regardless of whether the DNA contains a lesion (d9 versus dhT, dhT2x, and 64TC) and regardless of whether the DNA is single or double stranded.* Figure 4A does not show the direction of the $^{1}H^{N}$ and ^{15}N chemical shift changes upon the addition of DNA, which, in general, are the same for each residue and for each DNA, suggesting a common mode of DNA binding. Note that while the magnitude of the average changes in the $^{1}H^{N}$ and ^{15}N chemical shifts do differ between the damaged DNAs, these differences should not be overinterpreted. Even

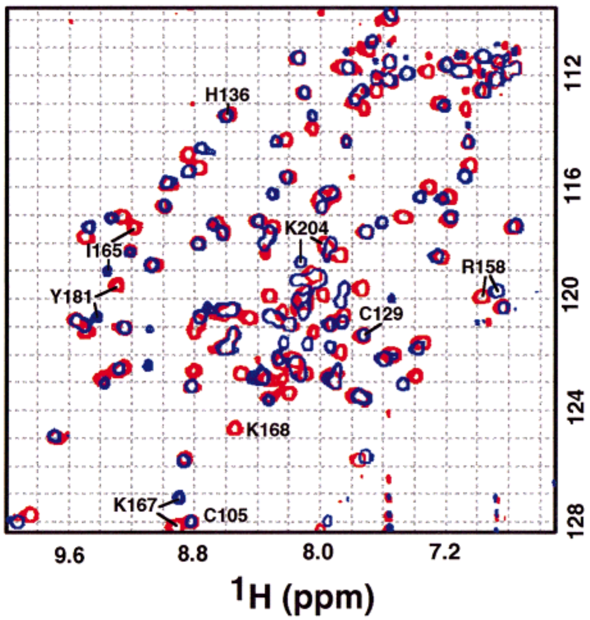

FIGURE 3: Comparison of the ^{1}H-^{15}N HSQC spectra of free XPA-MBD in the presence (blue) and absence (red) of dhT2x, 25 °C. Representative resonances for residues in different regions of the protein are labeled. C105, C129, and H136 are residues in the zinc-binding core and their $^1H^N$ and ^{15}N resonances do not change in the presence of DNA. I165 (β4), K167 (β4), Y181 (β5), and K204 (α3) are residues in the loop-rich subdomain whose cross-peaks move significantly in the presence of DNA. The cross-peak for K168 (β4) disappears. The cross-peak for R158, located in loop r1, changes moderately.

though the d9, dhT, and dhT2x concentrations were approximately equimolar with XPA-MBD at the final titration point, plots of the change in selected $^1H^N$ and ^{15}N resonances as a function of the amount of added DNA suggest that saturation had not yet been reached. Furthermore, it was not possible to exclusively determine if dhT2x and 64TC were binding XPA-MBD as a single or double strand because the melting temperature of dhT2x is ~45 °C and 25.1 °C for 64TC.

Interactions between XPA-MBD and RPA70 were studied using the N-terminal 168 residues (RPA70ΔC169) and N-terminal 326 residues (RPA70ΔC327) of RPA70. It is important to note that RPA70ΔC327 lacks the C-terminal 55 residues (out of 146) identified as the XPA binding site in RPA70 (*42*). The addition of RPA70ΔC167 to XPA-MBD did not affect the chemical shifts or cross-peak intensities in the $^1H^N$-^{15}N HSQC spectra of XPA-MBD indicating that the two proteins did not interact. However, the addition of RPA70ΔC327 to ^{15}N-labeled XPA-MBD resulted in a broadening of the XPA-MBD $^1H^N$-^{15}N HSQC cross-peaks. Such signal losses are observed when a stable large molecular weight complex forms, the exchange rate between the free and ligand-bound states is similar to the chemical shift difference between the two states, or, when ligand binding increases the $^1H^N$ exchange rate with water (*47*). The results with the two fragments of RPA70 concur with earlier studies which indicate that RPA70 residues 236−382 are necessary for binding XPA (*42*). Figure 4B plots the normalized change in intensity of the $^1H^N$-^{15}N HSQC cross-peaks of XPA-MBD upon the addition of RPA70ΔC327. Residue intensity changes near 1 indicate no, or little, interaction while intensity changes greater than 1.5 are indicative of a binding interface. While DNA perturbed the $^1H^N$ and ^{15}N XPA-MBD backbone resonances *only* in the loop-rich subdomain (Figure 4A), RPA70ΔC327 perturbed XPA-MBD resonances in *both* subdomains (Figure 4B).

Backbone Dynamics of Free XPA-MBD. A total of 104 of the 117 expected $^1H^N$-^{15}N cross-peaks were observed in the ^{1}H-^{15}N HSQC spectrum of XPA-MBD and 100 of them were assigned (*21*, *43*). Due to resonance overlap or weak cross-peak intensities, 24 residues were excluded from the relaxation analysis. Thus, complete quantitative relaxation measurements were made for 77 residues of free XPA-MBD. The top panel in Supporting Information Figures 1, 2, and 3 show the T_1, T_2, and heteronuclear NOE data, respectively, for free XPA-MBD. Most of T_1 values lie between 0.6 and 0.8 s with an overall average value of 0.77 ± 0.11 s. The majority of the T_2 values lie between 60 and 100 ms with an average time of 85 ± 19 ms between residues I104 and R207. Those T_2 values that lie above the 60−100 ms range are clustered primarily at the N-terminus (F100−Y102) and in loop r2 (N169−D177), suggesting these regions are more dynamic than the rest of the molecule. Interestingly, the few T_2 values determined for residues in loop r1 fall within the average T_2 range observed for the entire protein, suggesting that this loop is not more dynamic than the overall protein. The majority of the $^1H^N$-^{15}N heteronuclear NOEs fall in the 0.6−0.8 range, indicating that internal motions on the fast (picosecond) time scale are restricted. The exceptions are at the N-terminus (F100−Y103) and loop r2, regions with greater flexibility as suggested by the T_2 data.

Before the relaxation data could be analyzed with MODELFREE, it was necessary to determine if an axially symmetric or an isotropic model should be used in the calculations. The principal components of the inertia tensor for XPA-MBD were determined using the program pdbinertia_1.0 (*31*). The ratios of the principle moments were 1.00:0.81:0.50, values that vary significantly from a sphere and indicate that the protein is a prolate ellipsoid. The asymmetry in XPA-MBD was further verified using the program R2R1_1.1 (*31*) and the data selection criteria outlined to estimate the overall τ_m. The program indicated a statistically better fit for the relaxation data using an axially symmetric model over an isotropic model ($F = 10.0$), with values of 10.9 ns, 1.41, 0.2°, and 0.6° for τ_m, $D_\parallel/D_\perp$, θ, and ϕ, respectively. The latter values were used as starting parameters in the MODELFREE analysis of the relaxation data for XPA-MBD.

In the final analysis, 70 amide vectors for XPA-MBD were satisfactorily modeled, on the basis of statistical criteria outlined by Mandel et al. (*31*), with an axially symmetric diffusion tensor model. A summary of the MODELFREE spectral density models used to fit the relaxation data is reported in Table 1. The majority of the amide vectors (77%) for which MODELFREE modeling was attempted were fit best by the simplest, one parameter model, S^2. Overall, using all the models, the average value for S^2 was 0.78 ± 0.14. As shown in the top panel of Figure 5, most of the S^2 values fall between 0.65 and 0.9. Except for R130, all residues in regions of the protein with a well-defined secondary structure fell within, or above, this S^2 range. Of the 12 residues in loop r1, two residues are prolines, and the amides of three residues are exchange broadened and are not observed (*21*, *43*). However, the S^2 values for four out of the five tractable

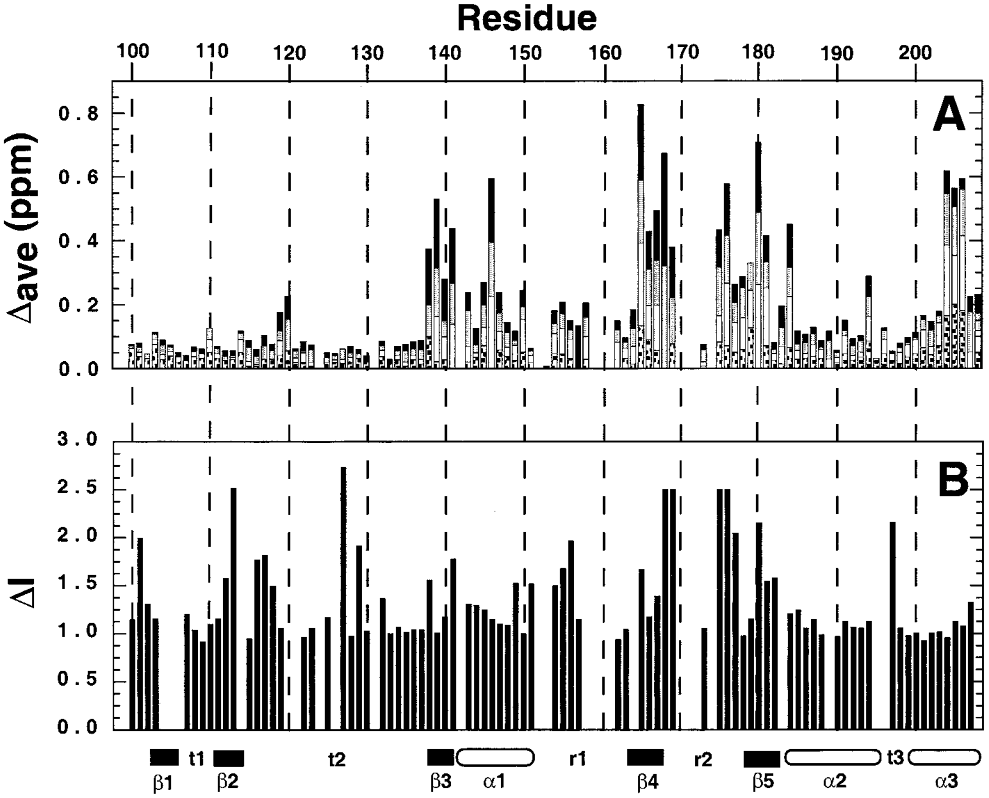

FIGURE 4: (A) Combined average chemical shift changes in the $^{1}H^{N}$ and ^{15}N resonances of XPA-MBD upon binding d9 (black), dhT (grey), dhT2x (white), and 64TC (black spots) plotted against residue number. The average chemical shift change = $\Delta_{ave} = [((\Delta^{1}H^{N})^2 + (\Delta^{15}N/5)^2)/2)]^{1/2}$. (B) Decrease in intensity of the XPA-MBD $^{1}H^{N}$-^{15}N HSQC cross-peaks in the presence of excess RPA70ΔC327 plotted against residue number. $\Delta I = [((\log I_o - \log I_{RPA}))/\log I_o]\alpha$ where I_o = free XPA-MBD, I_{RPA} = RPA70ΔC327, and α = scaling factor = 5.

residues in loop r1 are greater than 0.7, confirming, as indicated by the T_1 and T_2 data, that loop r1 tumbles with a rate similar to the overall molecule. Only two of the nine residues in loop r2 were fit by MODELFREE, and their S^2 values are less than the overall S^2 average. While the MODELFREE data for loop r2 is small, the trend suggests that r2 may be more mobile relative to the overall protein backbone, a conclusion reached with the T_2 and heteronuclear NOE data. XPA-MBD also contains a long "turn", t_2, between the first β-sheet (β1 and β2) and the start of the triple-strand antiparallel β-sheet. Many of the t_2 amide resonances are weak relative to the majority of the XPA-MBD $^{1}H^{N}$ resonances, suggesting that this region experiences motion that produces averaging on the chemical shift time scale (*21*). However, all the S^2 values determined for residues in t_2 are above 0.65, indicating that this turn is well-defined as indirectly suggested by the low pairwise RMSD values from the structure calculations (Figure 2A).

Exchange terms, R_{ex}, were necessary to produce a better fit in only four of the amide vectors (Table 1). These terms were not concentrated in any area of the protein, suggesting that no significant conformational or chemical exchange was occurring. Ikegami et al. (*53*) recently reported using 34 R_{ex} terms to fit the relaxation data for XPA-MBD. However, $T_{1\rho}$ rates measured as a function of the strength of the spin-lock field (*54*) did not detect any residues undergoing conformational exchange. A two time-scale model requiring fast internal motion terms, τ_e and S_f^2, was required to better fit some of the amide vectors at both the N- and C-termini of XPA-MBD, indicating that both ends of the protein were more dynamic than the overall backbone.

The final values for τ_m, $D_\parallel/D_\perp$, θ, and ϕ obtained after two rounds of MODELFREE iterations were 10.6 ± 0.1 ns, 1.41 ± 0.11, −50 ± 10°, and −84 ± 15°, respectively. The overall correlation time determined for free XPA-MBD, 10.6 ns, is within the range expected for a 14.7 kDa protein (*24*, *55*). The calculated rotational diffusion anisotropy of 1.41 is similar to a recently reported value of 1.39 for free XPA-MBD (*53*) and is consistent with an elliptical topology.

Backbone Dynamics of XPA-MBD in the Presence of DNA. Complete relaxation measurements were made for 65 and 67 residues of XPA-MBD in the presence of dhT and 64TC, respectively. Relative to free XPA-MBD, fewer residues were tractable due to a decrease in intensity of most of the $^{1}H^{N}$-^{15}N HSQC cross-peaks. Such an observation itself suggests that both dhT and 64TC formed a complex with the protein. Because the XPA-MBD sample containing dhT2x deteriorated before the heteronuclear NOE data could be collected, only T_1 and T_2 data are presented for this complex.

Supporting Information Figures 1, 2, and 3 contain the relaxation data for XPA-MBD in the presence of dhT, dhT2x, and 64TC. Most of the T_1 values lie between 0.7 and 0.9 s in the presence of 64TC and between 0.9 and 1.1 s in the presence of dhT and dhT2x. Table 2 shows that the overall average T_1 values for XPA-MBD increases in the presence of DNA relative to free XPA-MBD (0.77 s). The average T_1 values for dhT and dhT2x are similar (~1.02 s), while the average value for 64TC is smaller (0.88 s). Most of the T_2 values lie between 50 and 70 ms in the presence of dhT, dhT2x, and 64TC. Relative to free XPA-MBD, the average T_2 value decreases for all the complexes to a value of approximately 60 ms. The T_2 values for residues D101 and

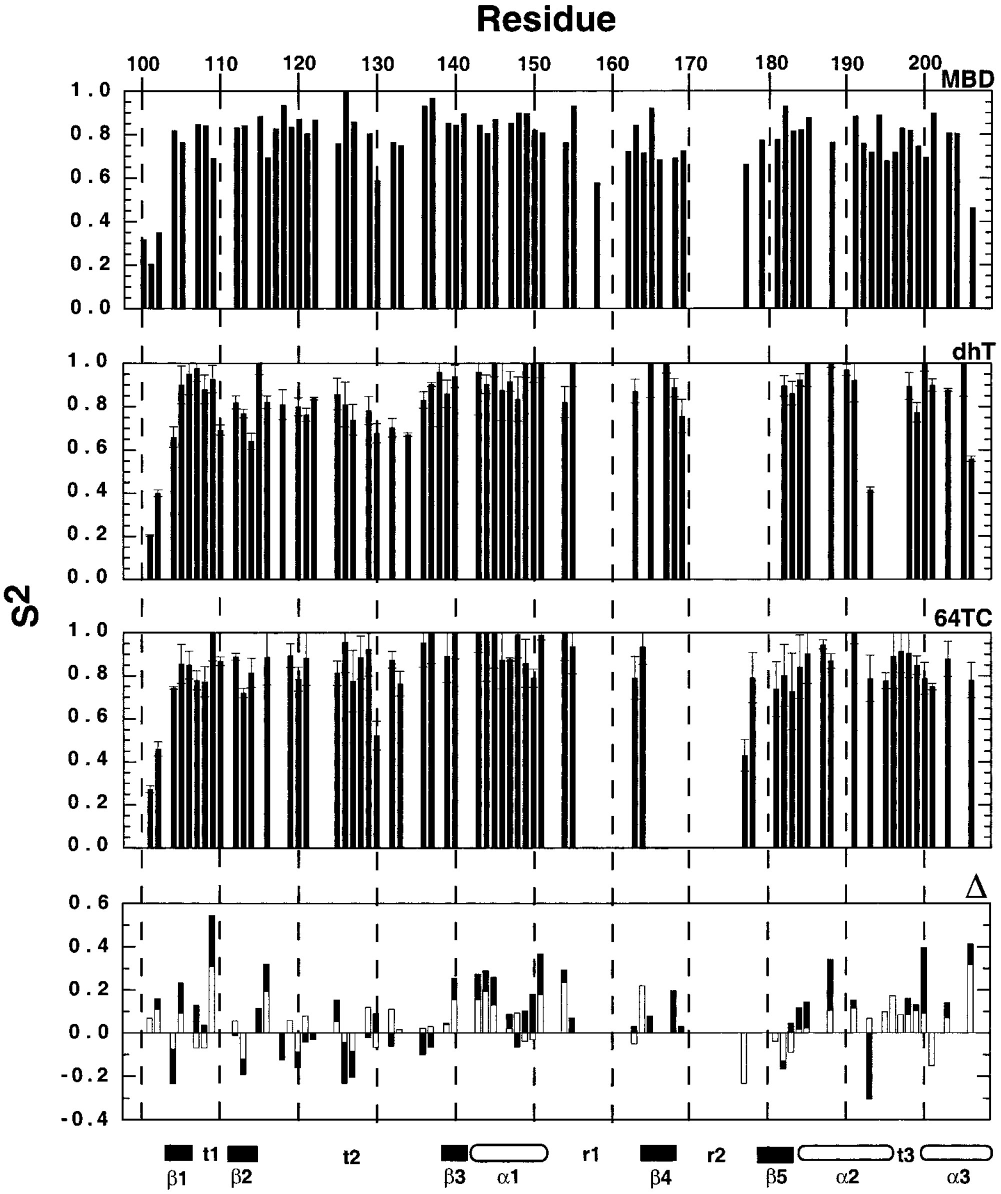

FIGURE 5: Optimized model-free order parameter values, S^2, determined for the amide vectors of free XPA-MBD (top) and XPA-MBD in the presence of dhT and 64TC, using the relaxation data from Supporting Information Figures 1, 2, and 3. (Δ) Difference between the generalized order parameters, S^2, for free XPA-MBD and in the presence of dhT (solid) and 64TC (open) (DNA associated XPA-MBD minus free XPA-MBD). Only residues whose relaxation parameters were determined for free XPA-MBD and DNA bound XPA-MBD are included in the plot.

Table 2: Rotational Correlation Times (τ_m) and Average T_1, T_2, and S^2 Values[a]

parameter	XPA-MBD	dhT	dhT2x	64TC
estimated τ_m (ns)	9.27 ± 1.34 (61)	19.76 ± 4.21 (58)	18.31 ± 4.01 (53)	13.97 ± 1.67 (51)
modelfree τ_m (ns)	10.6 ± 0.1	17.5 ± 0.2		13.5 ± 0.2
T_1 (all residues) (s)	0.77 ± 0.11 (83)	1.01 ± 0.15 (81)	1.04 ± 0.15 (77)	0.88 ± 0.14 (73)
T_2 (I104−R207) (ms)	85 ± 19 (85)	59 ± 38 (76)	58 ± 15 (71)	62 ± 12 (71)
S^2 (all)	0.78 ± 0.14	0.84 ± 0.15		0.84 ± 0.14
S^2 (>0.60)	0.81 ± 0.08	0.87 ± 0.10		0.87 ± 0.09

[a] The estimated τ_m was calculated as described in the text. The value in parentheses is the number of residues used in the calculations.

Y102 at the N-terminus of XPA-MBD in the presence of all three DNA are above the average range, likely due to the same dynamic flexibility observed in the unstructured terminus of free XPA-MBD. The only other residues with T_2 values significantly greater than the average T_2 range are E192 and V193 (α2) in the presence of dhT, which are more similar to those for free XPA-MBD than the values observed in the presence of dhT2x and 64TC. Perhaps this represents a localized difference in XPA-MBD association with the three DNAs. In the presence of dhT and dhT2x, the smallest T_2 values are observed at, or near, the residues of the triple-strand β-sheet, regions associated with large $^1H^N$ and ^{15}N chemical shift perturbations upon DNA association. Heteronuclear NOE data were only obtained for XPA-MBD in the presence of dhT and 64TC. As with free XPA-MBD, most of the heteronuclear NOEs fall in the 0.6−0.8 range indicating that in the presence of DNA internal motions in the fast (picosecond) time scale are restricted. Relative to free XPA-MBD, the only significant difference is that K179 (β5) displays a negative NOE in the presence of both dhT and 64TC.

Upon the addition of DNA to XPA-MBD, the $^1H^N$ and ^{15}N line widths were observed to increase which reflects the formation of a larger biomolecular complex (*42*). An increase in line width is associated with a decrease in T_2 and an increase in T_1, trends that were observed upon the addition of all the DNAs to XPA-MBD (Table 2 and Supporting Information Figures 1 and 2). DNA−protein complex formation was further confirmed by collecting ^1H-^{15}N HSQC spectra at different temperatures. Lowering the temperature stepwise from 25 to 7 °C resulted in a significant increase in the $^1H^N$ and ^{15}N line widths to the extent that many ^1H-^{15}N HSQC cross-peaks either were not observed or were not assignable at 7 °C in the presence of all 3 DNAs. On the other hand, the $^1H^N$-^{15}N HSQC spectra of free XPA-MBD was still well resolved at 5 °C (data not shown).

The MODELFREE analyses of the data for the complexes with dhT and 64TC were performed using an isotropic model because a solution structure for XPA-MBD bound to DNA does not yet exist. While XPA-MBD alone clearly adopts the shape of a prolate ellipsoid, it is reasonable to assume a more spherical geometry for the protein−DNA complex. Furthermore, it is unlikely that major structural changes occur to XPA-MBD upon DNA binding since no changes in the proteolytic digestion pattern of full-length XPA is observed upon binding to single-stranded DNA (*19*), and there were no significant differences between the ^{15}N-edited NOESY spectra of free XPA-MBD and DNA bound XPA-MBD. In the final analysis, 61 amide vectors were satisfactorily modeled for XPA-MBD in the presence of dhT and 64TC using an isotropic model. A summary of the MODELFREE spectral density models used to fit the relaxation data is reported in Table 1. As with free XPA-MBD, the majority of the amide vectors was fit best by a one parameter model, S^2. A higher percentage of the amide vectors for XPA-MBD in the presence of 64TC (79%) were fit by S^2 alone than in the presence of dhT (62%). The addition of dhT and 64TC to XPA-MBD resulted in an increase in the average value for S^2 to 0.84 relative to an average value of 0.78 observed for free XPA-MBD. As illustrated in the middle two panels of Figure 5, most of the S^2 values for XPA-MBD fall between 0.7 and 1.0 in the presence of DNA. The bottom panel of Figure 5 plots the change in S^2 as a function of residue number for the XPA-MBD NH bond vectors effected by the addition of dhT and 64TC. Except for turn t_2, backbone motion is generally more restricted throughout XPA-MBD even though the chemical shift mapping experiments indicate that the DNA is making contacts exclusively with the loop-rich subdomain (Figure 4). The increase in S^2 upon the addition of DNA is consistent with DNA binding restricting motion of the backbone NH bond vectors on the picosecond time scale (*47, 56*).

Table 1 indicates that no exchange terms, R_{ex}, were required to fit the relaxation data for XPA-MBD in the presence of dhT, and only two terms were used for 64TC. Therefore, no conformational or chemical exchange was occurring upon the addition of DNA. Table 1 also indicates that XPA-MBD bound to dhT required more fast internal motion terms to fit the relaxation data than 64TC. A cluster of fast internal motion terms, τ_e and S_f^2, were only observed at the N-terminus of XPA-MBD in both DNA complexes, indicating that this region, as observed in free XPA-MBD, is more mobile that the overall backbone. Aside from the latter residues, the remaining residues requiring τ_e or S_f^2 terms were spread throughout the protein.

Following three to four iterations, the final values of τ_m calculated for XPA-MBD in the presence of dhT and 64TC are 17.5 ± 0.2 and 13.5 ± 0.2 ns, respectively. Binding of a single strand of dhT to XPA-MBD would increase the molecular mass by ~2.9 kDa, and the binding of a double-strand of 64TC would increase the molecular mass by ~7.8 kDa. The observed increase in τ_m for XPA-MBD with the addition of both DNAs is expected due to the formation of a larger molecular weight complex. However, the τ_m value for the XPA-MBD-64TC complex should be larger than for the XPA-MBD-dhT complex. While the τ_m value determined for the protein−64TC complex is in the range expected for a 16.7 (1×) or 21.6 (2×) kDa protein (*24*), the value observed in the presence of dhT is larger than expected. Note that the τ_m value optimized for dhT (17.5 ns) using MODELFREE is near the value estimated for XPA-MBD in the presence of dhT2x (18.3 ns) (Table 2). The observation of larger τ_m values for dhT and dhT2x appears to correlate with larger T_1 values for both DNAs (Table 2). While the average T_2 values for XPA-MBD decreased by similar amounts in the presence of all three DNAs, the T_1 values for XPA-MBD increased more in the presence of dhT and dhT2x than with 64TC. The reason for such an observation is not obvious. However, the finding of larger τ_m values for dhT and dhT2x relative to 64TC suggests one of two things: either the amount of 64TC added to the XPA-MBD sample was not sufficient to saturate the protein to the same level as for dhT and dhT2x, or the global structure of the DNA−protein complex is different upon binding the 64TC and dhT adduct.

Mapping of the DNA- and RPA70-Binding Sites onto the Structure of XPA-MBD. Figure 6A is the average backbone structure for XPA-MBD (Figure 1A), highlighting the residues whose $^1H^N$ and ^{15}N chemical shifts are most perturbed by the addition of DNA (Figure 4A). The triple-strand antiparallel β-sheet and helices α1 and α3 are most affected and these regions cluster together to define a contiguous surface on XPA-MBD. Figure 6B is a structural map of XPA-MBD highlighting the residues whose order parameter, S^2, increased (blue) or decreased (red) by >0.1 in the presence of dhT. In general, the largest increases in S^2 are clustered in the same region of the loop-rich subdomain where the largest chemical shift perturbations of the $^1H^N$ and ^{15}N chemical shifts are observed, strongly suggesting that this is the surface responsible for binding DNA. Note that while the backbone of XPA-MBD is generally less mobile in the presence of DNA, some regions in the zinc-binding subdomain actually become more mobile. Perhaps such changes in the zinc-binding core play a role in recruiting RPA or other NER proteins to the XPA−DNA complex.

Previously, we postulated that the disorder observed in the loop-rich subdomain, r1, r2, and r3 (Q208-F219), might confer flexibility to the protein to allow it to recognize such a wide range of DNA lesions and other NER proteins (*21*).

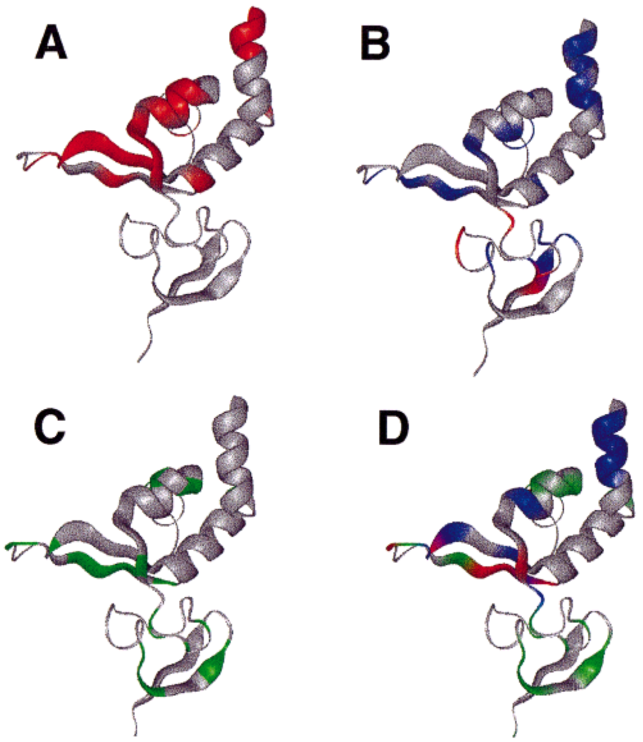

FIGURE 6: Summary of chemical shift mapping and relaxation experiments using ^{15}N-labeled XPA-MBD. (A) Residues with the largest ^{1}H^{N} and ^{15}N chemical shifts changes ($\Sigma\Delta_{ave} > 0.1$ ppm) following titration with various DNA oligomers (see text) are indicated in red. (B) Residues with order parameters that increase by more than 0.1 unit are colored in blue, those that decrease by more than 0.1 units are colored in red. (C) Residues whose ^{1}H^{N}-^{15}N HSQC cross-peak intensity decrease most ($\Delta I > 0.1$, Figure 4B) following titration with RPA70ΔC327 are colored in green. (D) Summary of chemical shift mapping results: residues that shift in the presence of DNA and RPA70ΔC327 are colored in red, those that shift with RPA only are colored in green, and those that shift with DNA only are colored in blue.

Analysis of the relaxation data suggests that the backbone motion in r1 is as restricted as the motion in the majority of the backbone of XPA-MBD, while r2 may be more flexible. Loop r3 is so flexible that the residues in this region were not used in the structure calculations (*21*, *22*). Addition of DNA to XPA-MBD caused changes in the ^{1}H^{N} and ^{15}N chemical shifts of residues extending from β-strands β4 and β5 into loop r2 while those in r1 were effected only modestly. Additional, unassigned, ^{1}H^{N}-^{15}N cross-peaks were also observed in the HSQC spectra, which may indicate that r3 also orders about the DNA. The substantial chemical shift perturbations observed toward the C-terminus of α3 and the new ^{1}H^{N}−^{15}N HSQC cross-peaks may indicate that r3 also orders about the DNA. Interestingly, it was previously shown that α3 is essential for DNA binding (*19*). Truncated XPA missing α3, S49−W195, no longer possessed any DNA binding activity unlike the extended sequence, S49−F219. Hence, the two most flexible loops in free XPA-MBD are affiliated with binding DNA and become less flexible upon DNA association. This was observed regardless of the type of DNA lesion and regardless of whether the DNA was single or double stranded.

Using the data from Figure 4B, Figure 6C maps onto the three-dimensional backbone structure of XPA-MBD (green) the resonances for which the ^{1}H^{N}-^{15}N HSQC cross-peak intensities were most affected by the addition of RPA70ΔC327. The β-sheets in both the zinc-binding and loop-rich subdomains, together with loop r2, occupy a face of the protein which may represent the RPA70ΔC327 binding surface on XPA-MBD. Figure 6D summarizes the chemical shift mapping results for both RPA70ΔC327 and the DNAs. Note that RPA70ΔC327 and DNA share a binding interface in the triple-strand β-sheet and loop r2 as indicated by the residues colored red. The nonshared sites are primarily in the zinc-binding core for RPA70ΔC327 (green) and exclusively the loop-rich subdomain for the DNAs (blue).

DISCUSSION

The data presented here show that XPA-MBD binds DNA with or without a lesion and in single- or double-stranded form. Hence, XPA contains a general DNA binding capacity that is localized to a very specific region of the protein. The location of the DNA-binding site is corroborated by the relaxation data that shows that the largest increases in S^2 overlap regions with the largest ^{1}H^{N} and ^{15}N chemical shift perturbations. The binding of DNA oligomers containing a mildly (dhT) (*38*, *39*) and grossly (64TC) distorting adduct (*35*−*37*), as well as undamaged DNA, suggests that the protein does not recognize specific features of a DNA lesion. Such an observation is not surprising given XPA's limited selectivity for damaged over undamaged DNA (*57*). If XPA-MBD, by itself, is not recognizing specific features of a DNA lesion, how is XPA-MBD recognizing damaged DNA? Gunz et al. (*58*) provided evidence for a "thermodynamic probing of the duplex" with XPA recognizing DNA lesions that destabilize the DNA helix. In this model, recognition of DNA damage requires only an enhanced binding affinity for exposed DNA caused by the lesion. This model is consistent with the general DNA-binding capacity of XPA-MBD observed in the mapping experiments presented here. Ikegami et al. (*22*) noted that the loop-rich subdomain of XPA-MBD contains a highly "basic cleft". Such basic surfaces are especially prominent in proteins that bind to DNA with little sequence specificity (*59*) and are typically observed at protein−nucleic acid interfaces (*60*). The major driving force for this type of protein−DNA association is intermolecular electrostatic interactions together with the associated release of counterions and bound water. In XPA, specific, positively charged amino acid side chains probably make direct contact with the sequential, negatively charged phosphate groups in the DNA backbone. As highlighted in Figure 7, the lysine and arginine residues in the loop-rich domain of XPA-MBD form a nearly continuous line that is more of a basic protrusion than a "basic cleft". Except for the gap between the C-terminal α3 and the bulk of the basic residues, the interresidue distance between adjacent basic headgroups is 3−5 Å. This distance is compatible with the 4−6 Å distance between phosphate groups on the backbone of single-stranded DNA. Such binding of single-stranded DNA to XPA agrees with models proposing that damage recognition depends on the degree of helix distortion and associated single-strandedness caused by the damage to DNA (*8*, *58*). While we do not know if the double-stranded DNA oligonucleotides added to XPA-MBD bound to the protein in double- or single-stranded form, the flexibility in loops r2 and r3 may allow XPA to accommodate the phosphodiester backbone of both

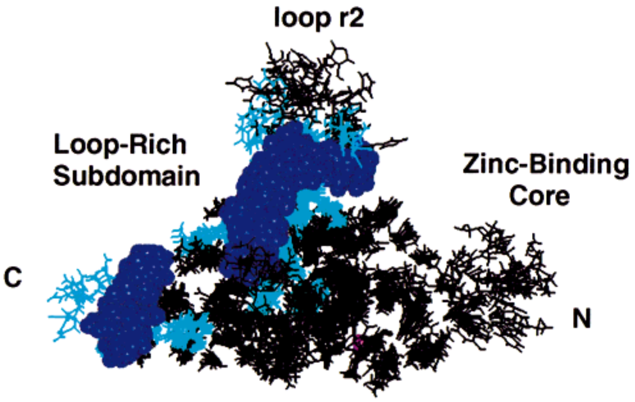

FIGURE 7: Superposition of the side chains of the ensemble of XPA-MBD structures shown in Figure 1A. The side chains of the residues whose $^{1}H^{N}$ and ^{15}N chemical shifts are perturbed most by the addition of DNA (Figure 6A) are colored blue with the positively charged Lys and Arg side chains highlighted and colored dark blue.

double- and single-stranded DNA on the basic protrusion of XPA. Interestingly, the tertiary structure of the loop-rich subdomain responsible for this DNA interaction has not been observed before since a DALI search (*61*) for similar structures produced no results with a *Z* number greater than 2. This observation suggests that the DNA binding mode of XPA is novel.

XPA is one of three NER proteins that have been shown to have a moderate preference for binding damaged DNA over undamaged DNA. The other two proteins are RPA (*62*, *63*) and XPC (*16*, *64*). However, none of these proteins, individually, bind damage DNA with enough specificity to account for the high specificity of lesion removal observed by NER in vivo (*57*). Instead, it has been proposed that the high specificity observed in vivo may be due to these proteins acting cooperatively to recognize damaged DNA. Weak, nonspecific binding of the basic surface of XPA to the phosphodiester backbone of DNA, exposed by helix distortion and associated single-strandedness caused by damage to the DNA, may serve to reduce the cooperative search for specific DNA damage to essentially one dimension (*65*). Even with proteins that bind to a specific DNA sequence, like the *lac* repressor, such electrostatic interactions serve to bind the DNA loosely until the protein can lock itself into the correct DNA-binding site (*66*).

It is known that an RPA−XPA complex binds damaged DNA about 4−10 times more tightly and selectively than either protein does individually with $K_d \approx 3 \times 10^{-6}$ for both XPA and RPA (*40*, *41*). This observation supports the sequential recognition model for the specificity of damage recognition by NER (*57*). According to Sancar's model (*57*), the specificity of binding is due to a cascade effect whereby the selectivities of RPA and XPA are multiplicative instead of additive (as they would be in a concerted recognition model). While the RPA−XPA complex binds DNA with some small degree of positive cooperatively, i.e., the RPA−XPA complex has a higher affinity for DNA than either protein does individually, the degree of positive cooperativity is much smaller than one might expect. For example, if the DNA binding energy of the RPA−XPA heterodimer was equal to the sum of the individual binding energies, then the DNA-binding affinity would be expected to increase by the product of the individual binding constants which would produce picomolar binding. Rather, the observed binding affinity is increased only by ∼10-fold. Therefore, there must be a structural explanation for why potentially tighter DNA binding in the complex is not observed. *One possibility is that there may be a feature of the RPA−XPA association that increases DNA-binding specificity at the expense of DNA-binding energy.* A clue to the structural interactions that determine DNA-binding affinity and specificity may be contained in the mutual chemical shift mapping site for RPA70ΔC327 and DNA on XPA (Figures 4 and 6D) in the loop-rich subdomain. There are a few possible scenarios as to how this shared binding site might account for the increased selectivity of the RPA−XPA complex. It may be that the RPA−XPA interface is exclusively in the zinc-binding core, but RPA effects conformational changes in the loop-rich subdomain of XPA that enhance XPA binding to damaged DNA, or it may be that there is an RPA−XPA interface in the loop-rich subdomain of XPA which together bind damaged DNA more tightly than either protein does individually. The loss in total DNA-binding energies when RPA and XPA are complexed may be the price for the greater specificity the complex has for damaged DNA compared to either protein individually. The picture is likely more complicated than this because RPA contains two XPA-binding sites, a strong site at the C-terminus of RPA32 and a weaker site between residues 236−382 in RPA70 (*42*). However, the weaker XPA-binding site in RPA70 overlaps the stronger of two single-strand DNA-binding sites in RPA.

It has recently been reported that an XPC-HR23B complex bound damaged DNA tightly enough to permit DNase I footprinting (*16*). Furthermore, it was reported that preincubation of damaged DNA with XPC-HR23B followed by the addition of other repair factors resulted in a higher rate of repair relative to damaged DNA preincubated with XPA or RPA plus XPA. These authors concluded that an XPC-HR23B−DNA complex was the initiating step in NER with XPA primarily serving a proofreading function (*16*). A more recent attempt to repeat these experiments did confirm that XPC-HR23B bound damaged DNA more tightly than XPA, RPA, or RPA plus XPA (*57*). However, preincubation with XPC-HR23B slowed the rate of repair relative to preincubation with XPA, RPA, or XPA plus RPA (*57*) leading to the conclusion that an XPA-RPA−DNA complex initiates NER. Regardless of which model proves to be correct, it is known that RPA and XPA are both required for lesion recognition and lesion removal by NER, regardless of the type of lesion (*67*). The chemical shift mapping and relaxation results presented here lay a foundation for the molecular organization of the XPA-RPA−DNA complex.

While XPA's role in recognizing damaged DNA is important, it is not XPA's only known NER function (*57*). After the initial RPA-XPA−DNA complex is formed, XPA orchestrates other NER proteins into the proper position to effect the dual incision of the 24−32 nucleotide repair patch. Such orchestration involves recruiting TFIIH, with or without XPC-HR23B, to the preincision site. TFIIH and RPA then recruit XPG (3′ endonuclease) while RPA and XPA jointly recruit XPF-ERCCI (5′ endonuclease) (*57*). Thus, XPA and RPA are intimately involved in positioning both the 5′ and 3′ endonucleases that carry out the dual incision of the damaged DNA. Efforts are underway in our laboratory to build on the structural models presented here to better understand the structural basis for XPA's important functional roles in NER.

ACKNOWLEDGMENT

We thank Dr. Andrew S. Lipton for programming assistance, Drs. John R. Cort and Kathleen McAteer for useful discussions, Drs. Brian D. Thrall, Shuisong Ni, Vicki Burnett, and David L. Springer for constructing XPA-MBD clones, and finally, Drs. Paul D. Ellis and Jeffrey D. Saffer for making many resources available to conduct the research.

SUPPORTING INFORMATION AVAILABLE

Three figures showing the plots of the (1) ^{15}N T_1 values, (2) ^{15}N T_2 values, and (3) ^{1}H^{N}-^{15}N heteronuclear NOE values for XPA-MBD alone and in the presence of dhT, dhT2x, and 64TC. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Cleaver, J. E., and Kraemer, K. H. (1989) in *The Metabolic Basis of Inherited Disease* (Scriver, S. C., Beaudet, A. L., Sly, W. S., and Valle, D., Eds.) McGraw-Hill Book Co., New York.
2. Friedberg, E. C., Walker, G. C., and Siede, W. (1995) *DNA Repair and Mutagenesis*, pp 283−365, American Society for Microbiology, Washington, DC.
3. Sancar, A. (1995) *J. Biol. Chem. 270*, 15915−15918.
4. Wood, R. D. (1997) *J. Biol. Chem. 272*, 23465−23468.
5. Mu, D., Park, C. H., Matsunaga, T., Hsu, D. S., Reardon, J. T., and Sancar, A. (1995) *J. Biol. Chem. 270*, 2415−2418.
6. Tanaka, K., Miura, N., Satokata, I., Miyamoto, I., Yoshida, M. C., Satoh, Y., Kondo, S., Yasui, A., Okayama, H., and Okada, Y. (1990) *Nature 348*, 73−76.
7. Tanaka, K. (1993) *Jpn. J. Hum. Genet. 38*, 1−14.
8. Cleaver, J. E., and States, J. C. (1997) *Biochem. J. 328*, 1−12.
9. Robins, P., Jones, C. J., Biggerstaff, M., Lindahl, T., and Wood, R. D. (1991) *EMBO J. 10*, 3913−3921.
10. Jones, C. J., and Wood, R. D. (1993) *Biochemistry 32*, 12096−12104.
11. Asahina, H., Kuraoka, I., Shirakawa, M., Morita, E. H., Miura, N., Miyamoto, I., Ohtsuka, E., Okada, Y., and Tanaka, K. (1994) *Mutat. Res. 315*, 229−238.
12. Li, L., Elledge, S. J., Peterson, C. A., Bales, E. S., and Legerski, R. J. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 5012−5016.
13. Park, C.-H., Mu, D., Reardon, J. T., and Sancar, A. (1995) *J. Biol. Chem. 270*, 4896−4901.
14. Matsuda, T., Saijo, M., Kuraoka, I., Kobayashi, T., Nakatsu, Y., Nagai, A., Enjoji, T., Masutani, C., Sugasawa, K., Hanoaka, F., Yasui, A., and Tanaka, T. (1995) *J. Biol. Chem. 270*, 4152−4157.
15. Li, R.-Y., Calsou, P., Jones, C. J., and Salles, B. (1998) *J. Mol. Biol. 281*, 211−218.
16. Sagasawa, K., Ng, J. M. Y., Masutani, C., Iwai, S., van der Spek, P. J., Eker, A. P. M., Hanaoka, F., Bootsma, D., and Hoeijmakers, J. H. J. (1998) *Mol. Cell 2*, 223−232.
17. Coverley, D., Kenny, M., Munn, M. Rupp, W. D., Lane, D. P., and Wood, R. D. (1991) *Nature 349*, 538−541.
18. Satoh, M. S., Jones, C. J., Wood, R. D., and Lindahl, T. (1993) *Proc. Natl. Acad. Sci U.S.A. 90*, 6335−6339.
19. Kuraoka, I., Monta, E. H., Saijo, M., Matsuda, T., Monkawa, K., Shirakawa, M., and Tanaka, K. (1996) *Mutat. Res. 362*, 87−95.
20. Hess, N. J., Buchko, G. W., Conradson, S. D., Espinosa, F. J., Ni, S., Thrall, B. D., and Kennedy, M. A. (1998) *Protein Sci. 7*, 1970−1975.
21. Buchko, G. W., Ni, S., Thrall, B. D., and Kennedy, M. A. (1998) *Nucleic Acids Res. 26*, 2779−2788.
22. Ikegami, T., Kuraoka, I., Saijo, M., Kodo, N., Kyogoku, Y., Morikawa, K., Tanaka, K., and Shirakawa, M. (1998) *Nat. Struct. Biol. 5*, 701−706.
23. Chen, T., Reizer, J., Saier, M. H., Jr., Fairbrother, W. J., and Wright, P. E. (1993) *Biochemistry 32*, 32−37.
24. Lee, A. L., Volkman, B. F., Robertson, S. A., Rudner, D. Z., Barbash, D. A., Cline, T. W., Kanaar, R., Rio, D. C., and Wemmer, D. E. (1997) *Biochemistry 36*, 14306−14317.
25. Foster, M. P., Wuttke, D. S., Clemens, K. R., Jahnke, W., Radhakrishnan, I., Tennant, L., Reymond, M., Chung, J., and Wright, P. E. (1998) *J. Biomol. NMR 12*, 51−71.
26. Shukar, S. B., Hajduk, P. J., Meadows, R. P., and Fesik, S. W. (1996) *Science 274*, 1531−1534.
27. Williams, R. J. P. (1989) *Eur. J. Biochem. 183*, 479−497.
28. Kay, L. E., Torchia, D. A., and Bax, A. (1989) *Biochemistry 28*, 8972−8979.
29. Clore, G. M., Driscoll, P. C., Wingfield, P. T., and Gronenborn, A. M. (1990) *Biochemistry 29*, 7387−7401.
30. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) *Biochemistry 33*, 5984−6003.
31. Mandel, A. M., Akke, M., and Palmer, A. G., III (1995) *J. Mol. Biol. 246*, 144−163.
32. Lipari, G., and Szabo, A. (1982) *J. Am. Chem. Soc. 104*, 4546−4559.
33. Lipari, G., and Szabo, A. (1982) *J. Am. Chem. Soc. 104*, 4559−4570.
34. Gochin, M., and Roder, H. (1995) *Protein Sci. 4*, 296−305.
35. Taylor, J.-S., Garrett, D. S., and Cohrs, M. P. (1988) *Biochemistry 27*, 7206−7215.
36. Kim, J.-K., Patel, D., and Choi, B.-S. (1995) *Photochem. Photobiol. 62*, 44−50.
37. Spector, T. I., Cheatham, T. E., III, and Kollman, P. A. (1997) *J. Am. Chem. Soc. 119*, 7095−7104.
38. Miller, J. H., Cooney, M., Maiskiewicz, K., and Osman, R. (1998) in *Molecular Modeling of Nucleic Acids*, (Leontis, N. B., and Santa Lucia, J., Eds.) pp 312−328, ACS Symp. Ser. 682, ACS, Washington DC.
39. Baleja, J. D., Buchko, G. W., Weinfeld, M., and Sykes, B. D. (1993) *J. Biomol. Struct. Dyn. 8*, 747−762.
40. He, Z., Henrickson, L. A., Wold, M. S., and Ingles, C. J. (1995) *Nature 374*, 566−569.
41. Li, L., Lu, X., Peterson, C. A., and Legerski, R. J. (1995) *Mol. Cell. Biol. 15*, 5396−5402.
42. Walther, A. P., Gomes, X. V., Lao, Y., Lee, C. G., and Wold, M. S. (1999) *Biochemistry 38*, 3963−3973.
43. Buchko, G. W., Ni, S., Thrall, B. D., and Kennedy, M. A. (1997) *J. Biol. NMR 10*, 313−314.
44. Gomes, X. V., and Wood, M. S. (1995) *J. Biol. Chem. 270*, 4534−4543.
45. Hare, D. R., Wemmer, D. E., Chou, S.-H., Drobny, G. P., and Reid, B. R. (1983) *J. Mol. Biol. 171*, 319−336.
46. Caruthers, M. H. (1983) *Science 230*, 281−285.
47. Kay, L. E., Muhandiram, D. R., Farrow, N. A., Aubin, Y., and Forman-Kay, J. D. (1996) *Biochemistry 35*, 361−368.
48. Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) *J. Biol. NMR 4*, 845−858.
49. Vuister, G. W., Clore, G. M., Gronenborn, A. M., Powers, R., Garrett, D. S., Tschudin, R., and Bax, A. (1993) *J. Magn. Reson. B 101*, 210−213.
50. Constantine, K. L., Mueller, L., Goldfarb, V., Wittekind, M., Metzler, W. J., Yanchunas, J., Robertson, J. G., Malley, M. F., Friedrichs, M. S., and Farmer, B. T., II (1997) *J. Mol. Biol. 267*, 1223−1246.
51. Farmer, B. T., II, and Mueller, L. (1994) *J. Biomol. NMR 4*, 673−687.
52. Brünger, A. T. (1993) *X-PLOR Manual, Version 3.1*, Yale University Press, New Haven, CT.
53. Ikegami, T., Kuroaka, I., Saijo, M., Kodo, N., Kyogoku, Y., Morikawa, K., Tanaka, K., and Shirakawa, M. (1999) *J. Biochem. 125*, 495−506.
54. Peng, J. W., and Wagner, G. (1994) *Methods Enzymol. 239*, 563−654.

55. Cheng, J.-W., Lepre, C. A., Chambers, S. P., Fulghum, J. R., Thomson, J. A., and Moore, J. M. (1993) *Biochemistry 32*, 9000−9010.
56. Zhang, W., Smithgall, T. E., and Gmeiner, W. H. (1998) *Biochemistry 37*, 7119−7126.
57. Wakasugi, M., and Sancar, A. (1999) *J. Biol. Chem. 274*, 18759−18768.
58. Gunz, D., Hess, M. T., and Naegeli, H. (1996) *J. Biol. Chem. 271*, 25089−25098.
59. Nadassy, K., Wodak, S. J., and Janin, J. (1999) *Biochemistry 38*, 1999−2017.
60. Creighton, T. E. (1993) *Proteins: Structure and Molecular Properties*, pp 355−360, W. H. Freeman & Company, New York.
61. Holm, L., and Sander, C. (1998) *Nucleic Acids Res. 26*, 316−319.
62. Clugston, C. K., McLaughlin, K., Kenny, M. K., and Brown, R. (1992) *Cancer Res. 52*, 6375−6379.
63. Burns, J. L., Guzder, S. N., Sung, P., Prakash, S., and Prakash, L. (1996) *J. Biol. Chem. 271*, 11607−11610.
64. Reardon, J. T., Mu, D., and Sancar, A., (1996) *J. Biol. Chem. 271*, 19451−19456.
65. McGhee, J. D., and von Hoppel, P. H. (1974) *J. Mol. Biol. 86*, 469−489.
66. Oehler, S., Eismann, E. R., Kramer, H., and Muller-Hill, B. (1990) *EMBO J. 9*, 973−979.
67. Mu, D., Hsu, D. S., and Sancar, A. (1996) *J. Biol. Chem. 271*, 8285−8294.

BI991755P