JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Communication
pubs.acs.org/JACS

Redefining the DNA-Binding Domain of Human XPA

Norie Sugitani,‡ Steven M. Shell,‡ Sarah E. Soss, and Walter J. Chazin*

Departments of Biochemistry and Chemistry, and Center for Structural Biology, Vanderbilt University, Nashville, Tennessee 37232-8725, United States

Supporting Information

ABSTRACT: Xeroderma pigmentosum complementation group A (XPA) protein plays a critical role in the repair of DNA damage via the nucleotide excision repair (NER) pathway. XPA serves as a scaffold for NER, interacting with several other NER proteins as well as the DNA substrate. The critical importance of XPA is underscored by its association with the most severe clinical phenotypes of the genetic disorder Xeroderma pigmentosum. Many of these disease-associated mutations map to the XPA_{98-219} DNA-binding domain (DBD) first reported ~20 years ago. Although multiple solution NMR structures of XPA_{98-219} have been determined, the molecular basis for the interaction of this domain with DNA is only poorly characterized. In this report, we demonstrate using a fluorescence anisotropy DNA-binding assay that the previously reported XPA DBD binds DNA with substantially weaker affinity than the full-length protein. In-depth analysis of the XPA sequence suggested that the original DBD construct lacks critical basic charge and helical elements at its C-terminus. Generation and analysis of a series of C-terminal extensions beyond residue 219 yielded a stable, soluble human XPA_{98-239} construct that binds to a Y-shaped ssDNA−dsDNA junction and other substrates with the same affinity as the full-length protein. Two-dimensional $^{15}N-^{1}H$ NMR suggested XPA_{98-239} contains the same globular core as XPA_{98-219} and likely undergoes a conformational change upon binding DNA. Together, our results demonstrate that the XPA DBD should be redefined and that XPA_{98-239} is a suitable model to examine the DNA binding activity of human XPA.

Nucleotide excision repair (NER) is a highly versatile DNA damage repair pathway that is able to remove bulky DNA lesions arising from exposure to sunlight, endogenous metabolites, and various environmental toxins.[1,2] Defects in NER result in the genetic disease xeroderma pigmentosum (XP), a spectrum of disorders characterized by hypersensitivity to sunlight, dramatically increased incidents of skin cancer, and neurological disorders.[3−5] NER in humans involves the coordinated action of ~30 proteins, including seven that were identified on the basis of their direct association with specific XP disorders (XPA−XPG).[6−8] Among these, the essential XP complementation group A protein (XPA) is associated with the most severe clinical XP phenotypes, leading to neurodegenerative disorders, accelerated aging, and cancer.[3,5]

Despite its key importance to NER, XPA has no known enzymatic function.[9] However, XPA is known to bind to DNA and a number of other NER proteins, suggesting that it serves as a scaffold for the complex multiprotein NER machinery.[8,10−13] Genetic and biochemical studies suggest that DNA binding by XPA is crucial for the proper function of NER. Moreover, a number of XPA mutations associated with severe XP symptoms map to residues in the DNA-binding domain (DBD).[5,9,14,15] Nevertheless, there has yet to be any systematic biophysical and structural characterization of the interactions between XPA and DNA.

The discovery of the human XPA DBD was reported nearly 20 years ago. Biochemical studies revealed a protease-resistant domain within residues 98−219 that was associated with binding of DNA.[16] Additional studies suggested that, relative to ssDNA or dsDNA, XPA binds preferentially to DNA containing ssDNA−dsDNA junctions.[19,20] This observation was of particular interest because NER requires unwinding of the DNA duplex, which creates ssDNA−dsDNA junctions. Two solution NMR structures of XPA_{98-219} were subsequently determined, revealing a globular core spanning residues 98−198.[17,21] NMR chemical shift analysis was also used to investigate binding of a 9 nucleotide (nt) ssDNA substrate, which enabled mapping of the interaction to a shallow basic cleft in XPA_{98-219}.[18] However, the affinity for this substrate is extremely weak (K_d estimated to be several mM), which leads to considerable doubt about whether this model accurately represents how XPA interacts with DNA. Nevertheless, this study has remained the prevailing model to explain how XPA binds to the NER bubble.[22] We therefore set out to structurally characterize the interactions between XPA_{98-219} and a high-affinity DNA substrate.

We began by setting up crystallization trials for human XPA_{98-219} in complex with a Y-shaped ssDNA−dsDNA junction substrate that contains an 8 basepair duplex extended by two noncomplementary 12 nt ssDNA arms on one end of the duplex (Figure S1). After standard screening of conditions, crystals were obtained that diffracted to 2.2 Å. A concern arose during the course of refining the data when it was realized that, based on the Matthews coefficient, the volume of the unit cell was not sufficient to contain the mass of the protein and the DNA substrate. The molecule in the crystal was assumed to be the protein because the volume of the asymmetric unit at 42% solvent content could accommodate only one molecule of 15 kDa XA_{98-219}. Moreover, an absorption peak at 9.67 keV indicated the presence of zinc in the crystal, presumably from the XPA_{98-219} zinc motif. Before progressing with further

Received: March 25, 2014
Published: July 21, 2014

   dx.doi.org/10.1021/ja503020f | J. Am. Chem. Soc. 2014, 136, 10830−10833

refinement, we decided to determine the affinity of $XA_{98–219}$ for the DNA substrate.

To this end, a fluorescence anisotropy (FA) assay was employed to directly compare the DNA-binding activity of full-length human XPA and $XPA_{98–219}$ (Figure 1). Fluorescein

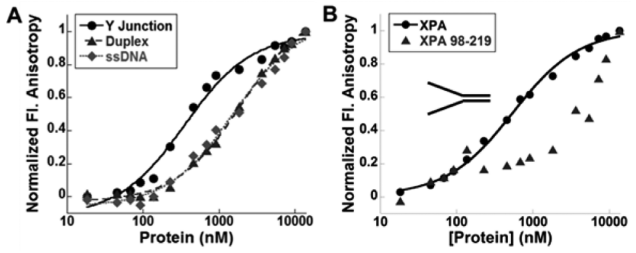

Figure 1. (A) Fluorescence anisotropy assay of the binding of XPA to Y-shaped ssDNA–dsDNA junction (black circles, solid line), duplex (gray triangles, dashed line), and ssDNA (light gray diamonds, dotted line). (B) Comparison of the binding of the Y-shaped ssDNA–dsDNA junction by full-length XPA (black circles, solid line) and $XPA_{98–219}$ (gray triangles). The concentration of the FITC-tagged DNA substrate was 20 nM, and measurements were performed at room temperature in a buffer containing 20 mM HEPES at pH 7.5, 75 mM KCl, 5 mM $MgCl_2$, 5% glycerol, and 1 mM dithiothreitol.

isothiocyanate (FITC)-modified ssDNA, dsDNA, and Y-shaped ssDNA–dsDNA junction substrates (Figure S1) were used for these analyses. Substrates of 20 nucleotides were selected for this analysis, as this corresponds to the approximate length of DNA predicted to be occluded by one molecule of XPA.[20] The results we obtained for full-length XPA were consistent with previous reports; Figure 1A shows that XPA binds a Y-shaped ssDNA–dsDNA junction (0.29 ± 0.09 μM) with higher affinity than dsDNA (1.7 ± 0.6 μM) or ssDNA (1.5 ± 0.2 μM). In stark contrast, $XPA_{98–219}$ had substantially weaker DNA binding affinity for all three substrates, so weak that it was not possible to extract a K_d value even for the highest affinity Y-shaped ssDNA–dsDNA junction (Figure 1B). In order to verify that $XPA_{98–219}$ was properly folded, a ^{15}N-enriched sample was prepared, and a ^{15}N–^{1}H HSQC spectrum was acquired. Comparison to the previously reported spectra for this construct[17,18,21] confirmed that our sample of $XPA_{98–219}$ was properly folded and free of aggregation (Figure S2, red spectrum). Taken together, these results demonstrate that $XPA_{98–219}$ lacks critical elements necessary to reproduce the full DNA-binding activity of XPA. Thus, the widely accepted view that $XPA_{98–219}$ is the DBD must be revised.

The observation that $XPA_{98–219}$ does not recapitulate the full activity of XPA led us to perform analyses of the primary sequence to search for indications that other residues might contribute to DNA binding. Having previously showed that the N-terminal domain of human XPA is disordered,[23] we focused on the sequence extending toward the C-terminus. Interestingly, secondary structure predictions indicated that the C-terminus of $XPA_{98–219}$ is located in the midst of a long helical element, with a high probability for helical secondary structure extending well beyond F219 (Figure S3). Moreover, there are several lysine and arginine residues in the region C-terminal to F219 that presumably enhance DNA binding affinity through electrostatic interaction with negatively charged DNA. Based on these insights, a series of C-terminally extended constructs were prepared (Figure S4). In all, six different human XPA constructs were cloned into bacterial expression vectors. After screening for soluble expression in *E. coli*, we assessed the solubility and stability of each construct; on this basis, $XPA_{98–239}$ was selected for further analysis.

To determine if the extra C-terminal residues were important for binding DNA, the affinity of $XPA_{98–239}$ for the 20 nt Y-shaped ssDNA–dsDNA junction, dsDNA, and ssDNA substrates (Figure S1) was measured using the FA assay (Figures 2A and S5). Notably, these data provided K_d values of 0.29 ± 0.08, 1.3 ± 0.2, and 1.5 ± 0.8 μM, respectively, very similar to those for the full-length XPA, including the

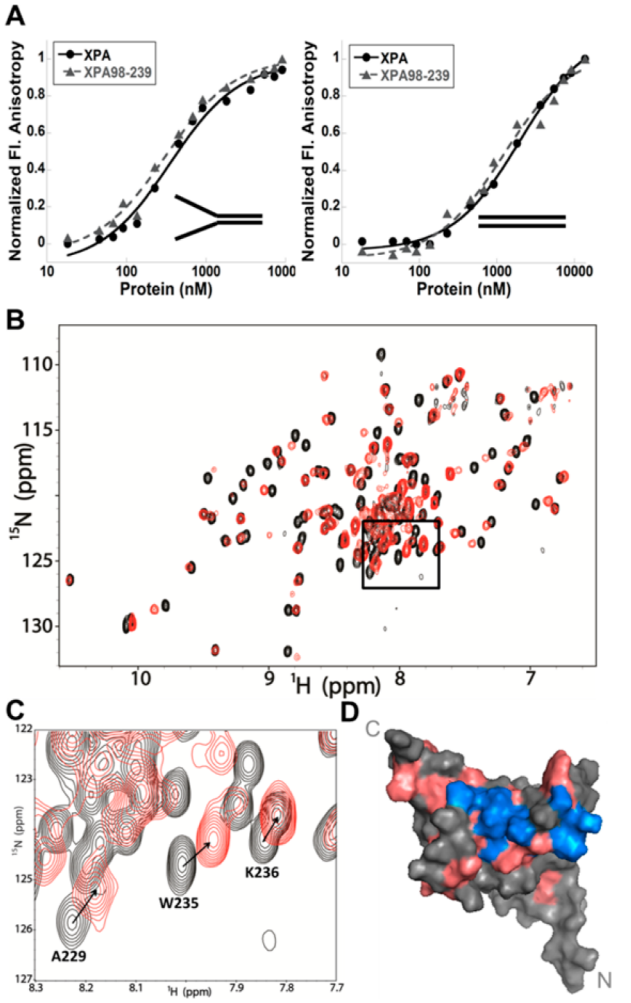

Figure 2. (A) Fluorescence anisotropy assay of the binding of the Y-shaped ssDNA–dsDNA junction (left) and duplex (right) substrates by full-length XPA (black circles, solid line) and $XPA_{98–239}$ (gray triangles, dashed line). The conditions were the same as in Figure 1. (B) 900 MHz ^{15}N–^{1}H TROSY HSQC spectra of $XPA_{98–239}$ obtained in the absence (black) and presence (red) of an equimolar amount of Y-shaped ssDNA–dsDNA junction substrate. The data were acquired at 35 °C in a buffer containing 20 mM Tris at pH 7.0, 150 mM KCl, 1 mM DTT. (C) Zoomed-in view of the boxed region of (B) showing perturbations of cross peaks from A229, W235, and K236 in the C-terminal extension. (D) Map of NMR chemical shift perturbations on a surface representation of $XPA_{98–219}$ (PDB ID: 1d4u). Residues identified in the study of $XPA_{98–219}$ binding a 9 nt ssDNA substrate[18] are colored blue. Additional residues with significant perturbations in the study of $XPA_{98–239}$ binding the Y-shaped ssDNA–dsDNA junction substrate are colored salmon. See Supporting Information for a detailed description of how residues with significant perturbations were identified.

dx.doi.org/10.1021/ja503020f | J. Am. Chem. Soc. 2014, 136, 10830–10833

preference for the Y-shaped ssDNA−dsDNA junction over dsDNA or ssDNA.[19,20] These results indicate XPA_{98-239} is a more suitable model for XPA DBD than XPA_{98-219}.

To verify that XPA_{98-239} occupies a stable conformation and is not aggregated, ^{15}N-enriched XPA_{98-239} was prepared, and a 2D $^{15}N-^{1}H$ HSQC spectrum was acquired. XPA_{98-239} is seen to have the characteristics of a stably folded 17 kDa protein, with narrow line widths and spectral dispersion evident in the ^{1}H dimension (Figure 2B, black spectrum). An overlay of the spectra of XPA_{98-219} and XPA_{98-239} (Figure S2) strongly suggests that they adopt a similar topology and that the new construct contains the globular core spanning residues 98−198. Since many of the peaks overlap, a significant number of the previously reported resonance assignments for XPA_{98-219} could be transferred to XPA_{98-239}. There are 16 extra cross peaks in the HSQC spectrum of XPA_{98-239}, and a limited number of them could be assigned (N. Sugitani, S. E. Soss, and W. J. Chazin, unpublished results). These cross peaks from the C-terminal extension have the same line shape as the other peaks in the spectrum and are narrowly dispersed. Although consistent with the prediction of helical character, the available data are not sufficient to formally assign the structure of the C-terminal extension.

To further characterize the interaction of XPA_{98-239} with DNA, we monitored a titration of the Y-shaped ssDNA−dsDNA junction substrate using 2D $^{15}N-^{1}H$ HSQC NMR. This analysis showed perturbation of a select number of cross peaks in the spectrum that saturate at a ratio of ∼1:1 (Figure S6), consistent with specific binding of this DNA substrate with low μM affinity. Comparison to the corresponding titration of XPA_{98-219} confirms that XPA_{98-239} binds the substrate much more strongly; the shorter construct is far from saturation at the 1:1 ratio and in fact does not saturate even at a substrate ratio of 5:1 (Figure S6). These observations support the proposal that the XPA DBD had been incorrectly assigned.

The NMR titration data also enabled us to test the validity of the previous model for the DNA binding site of XPA. In the previous study of a 9 nt ssDNA substrate binding to XPA_{98-219}, chemical shift perturbations in fast exchange between the free and bound states were observed for 13 residues, and 3 others were exchange broadened. Cross peaks from a larger number of residues are perturbed in the titration with the much higher affinity 20 nt Y-shaped ssDNA−dsDNA junction substrate. Consistent with the higher affinity for the Y-shaped substrate, both chemical shift perturbations in fast exchange and line broadening of resonances in intermediate exchange were observed. Figure 2D maps the residues exhibiting significant perturbations on the previously determined NMR structure of the globular core (residues 98−198); beyond the residues previously assigned to the DNA binding site in the study of XPA_{98-219} with 9 nt ssDNA (blue), the titration with Y-shaped substrate identified many additional perturbed residues (salmon). The latter include several additional residues in and around the basic cleft (residues L191, K204, and R207). One additional critical observation was the perturbation of cross peaks from three residues in the C-terminal extension (A229, W235, and K236, Figure 2C), which strongly supports our proposal of the need for the C-terminal extension for full DNA binding activity.

Our results show that XPA_{98-239} contains the full DNA-binding apparatus of human XPA, thereby redefining the XPA DBD. Mutations of residues between F219 and T239 are associated with severe XP disorders, which implies this region of the protein is critical to the function of XPA.[5,9] The incorrect assignment of XPA_{98-219} as the DBD may help explain the lack of substantial progress in elucidating the molecular mechanisms of XPA action in NER over the past 20 years.[22] Moreover, our studies of the more physiologically relevant ssDNA−dsDNA junction substrate clearly demonstrate the previous model for the XPA DNA binding site was incomplete. The new XPA_{98-239} DBD provides an excellent target for high-resolution structural and biophysical investigations of the XPA−DNA complex that can better define its role in NER. Additionally, as increased NER activity is often associated with loss of effectiveness of multiple classes of current anticancer treatments, such as radiation therapy and cisplatin,[24,25] XPA has been identified as a possible target for therapeutic intervention due to its critical role in NER.[26] The availability of structural information greatly enhances the pace of drug discovery. Hence, given the high-quality NMR data presented here, XPA_{98-239} has potential as a valuable reagent for structural analyses directed to the design and validation of novel small-molecule inhibitors and probes.

■ ASSOCIATED CONTENT

Supporting Information

Experimental methods, Figures S1−S6, and Table S1. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

walter.chazin@vanderbilt.edu

Author Contributions

‡N.S. and S.M.S. contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Nicholas P. George, Matthew K. Thompson, and Brandt F. Eichman for assistance in analyzing X-ray diffraction data, Markus Voehler for valuable insights for NMR experiments, Joshua Bauer and Sonja Brooks for their help with FA assay, and M. Wade Calcutt and Anindita Basu for assistance with mass spectroscopy. This work was supported by funding from NIH grants R01 ES1065561 and PO1 CA092584. Access to facilities was supported by P30 ES00267 to the Vanderbilt Center in Molecular Toxicology and P30 CA068485 to the Vanderbilt Ingram Cancer Center. S.M.S. is supported by postdoctoral fellowship 119569-PF-11-271-01-DMC from the American Cancer Society. Support for acquisition and upgrade of the NMR instrumentation was obtained from shared instrumentation grants from the NIH (S10 RR025677) and NSF (DBI-0922862).

■ REFERENCES

(1) Gillet, L. C.; Scharer, O. D. *Chem. Rev.* **2006**, *106*, 253−76.
(2) Truglio, J. J.; Croteau, D. L.; Van Houten, B.; Kisker, C. *Chem. Rev.* **2006**, *106*, 233−52.
(3) Cleaver, J. E. *Nat. Rev. Cancer* **2005**, *5*, 564−73.
(4) Gratchev, A. *Adv. Exp. Med. Biol.* **2008**, *637*, 113−9.
(5) Hengge, U. R.; Emmert, S. *Adv. Exp. Med. Biol.* **2008**, *637*, 10−8.
(6) Park, C. J.; Choi, B. S. *FEBS J.* **2006**, *273*, 1600−8.
(7) Riedl, T.; Hanaoka, F.; Egly, J. M. *EMBO J.* **2003**, *22*, 5293−303.
(8) Shell, S. M.; Zou, Y. *Adv. Exp. Med. Biol.* **2008**, *637*, 103−12.
(9) Cleaver, J. E.; States, J. C. *Biochem. J.* **1997**, *328* (Pt.1), 1−12.

(10) Camenisch, U.; Nageli, H. *Adv. Exp. Med. Biol.* **2008**, *637*, 28−38.
(11) Tsodikov, O. V.; Ivanov, D.; Orelli, B.; Staresincic, L.; Shoshani, I.; Oberman, R.; Scharer, O. D.; Wagner, G.; Ellenberger, T. *EMBO J.* **2007**, *26*, 4768−76.
(12) Krasikova, Y. S.; Rechkunova, N. I.; Maltseva, E. A.; Petruseva, I. O.; Lavrik, O. I. *Nucleic Acids Res.* **2010**, *38*, 8083−94.
(13) Missura, M.; Buterin, T.; Hindges, R.; Hubscher, U.; Kasparkova, J.; Brabec, V.; Naegeli, H. *EMBO J.* **2001**, *20*, 3554−64.
(14) Bradford, P. T.; Goldstein, A. M.; Tamura, D.; Khan, S. G.; Ueda, T.; Boyle, J.; Oh, K. S.; Imoto, K.; Inui, H.; Moriwaki, S.; Emmert, S.; Pike, K. M.; Raziuddin, A.; Plona, T. M.; DiGiovanna, J. J.; Tucker, M. A.; Kraemer, K. H. *J. Med. Genet.* **2011**, *48*, 168−76.
(15) DiGiovanna, J. J.; Kraemer, K. H. *J. Invest. Dermatol.* **2012**, *132*, 785−96.
(16) Kuraoka, I.; Morita, E. H.; Saijo, M.; Matsuda, T.; Morikawa, K.; Shirakawa, M.; Tanaka, K. *Mutat. Res.* **1996**, *362*, 87−95.
(17) Ikegami, T.; Kuraoka, I.; Saijo, M.; Kodo, N.; Kyogoku, Y.; Morikawa, K.; Tanaka, K.; Shirakawa, M. *Nat. Struct. Biol.* **1998**, *5*, 701−6.
(18) Buchko, G. W.; Tung, C. S.; McAteer, K.; Isern, N. G.; Spicer, L. D.; Kennedy, M. A. *Nucleic Acids Res.* **2001**, *29*, 2635−43.
(19) Hey, T.; Lipps, G.; Krauss, G. *Biochemistry* **2001**, *40*, 2901−10.
(20) Yang, Z.; Roginskaya, M.; Colis, L. C.; Basu, A. K.; Shell, S. M.; Liu, Y.; Musich, P. R.; Harris, C. M.; Harris, T. M.; Zou, Y. *Biochemistry* **2006**, *45*, 15921−30.
(21) Buchko, G. W.; Daughdrill, G. W.; de Lorimier, R.; Rao, B. K.; Isern, N. G.; Lingbeck, J. M.; Taylor, J. S.; Wold, M. S.; Gochin, M.; Spicer, L. D.; Lowry, D. F.; Kennedy, M. A. *Biochemistry* **1999**, *38*, 15116−28.
(22) Petruseva, I. O.; Evdokimov, A. N.; Lavrik, O. I. *Acta Naturae* **2014**, *6*, 23−34.
(23) Mer, G.; Bochkarev, A.; Gupta, R.; Bochkareva, E.; Frappier, L.; Ingles, C. J.; Edwards, A. M.; Chazin, W. J. *Cell* **2000**, *103*, 449−56.
(24) Muggia, F. M.; Fojo, T. *J. Chemother.* **2004**, *16* (Suppl.4), 77−82.
(25) Wozniak, K.; Blasiak, J. *Acta Biochim. Pol.* **2002**, *49*, 583−96.
(26) Neher, T. M.; Shuck, S. C.; Liu, J. Y.; Zhang, J. T.; Turchi, J. J. *ACS Chem. Biol.* **2010**, *5*, 953−65.