Journal of Biomolecular NMR, 10 (1997) 313–314
KLUWER/ESCOM

© 1997 *Kluwer Academic Publishers. Printed in The Netherlands.*

J-Bio NMR 494

Human nucleotide excision repair protein XPA: Expression and NMR backbone assignments of the 14.7 kDa minimal damaged DNA binding domain (Met^{98}–Phe^{219})

Garry W. Buchko, Shuisong Ni, Brian D. Thrall and Michael A. Kennedy*

Pacific Northwest National Laboratory, Environmental Molecular Sciences Laboratory and Molecular Biosciences Department, 902 Battelle Boulevard, Richland, WA 99352, U.S.A.

Received 22 July 1997
Accepted 15 August 1997

Keywords: Zinc-binding domain; DNA repair; XPA; Backbone assignments

Biological context

A wide variety of chemically and structurally distinct DNA lesions are repaired in the genome of prokaryotes and eukaryotes by the ubiquitous nucleotide excision repair (NER) pathway (Gunz et al., 1995). At least three human cancer-prone genetic instability syndromes arise from mutations in NER genes: trichothiodystrophy, Cockayne's syndrome, and xeroderma pigmentosum (XP). XP is an autosomal recessive human disease in which sunlight induces skin disorders, including cancer (Cleaver, 1968). One of the most severe XP repair deficiency syndromes arises from mutations in the XPA gene which codes for a 31 kDa protein (273 a.a.) that is involved in the initial damage recognition steps of NER (Tanaka et al., 1990; Robins et al., 1991). Crucial to XPA's ability to recognize damaged DNA is a C4-type zinc-binding motif, C105-X_2-C108-X_{17}-C126-X_2-C129 (Asahina et al., 1994). Proton NMR solution studies of a synthetic 41-residue peptide containing the zinc-binding core of XPA (Asp^{101}–Lys^{141} = zXPA-41) (Buchko and Kennedy, 1997) suggest that zXPA-41 adopts a structure similar to the DNA binding domain of the chicken erythroid transcription factor GATA-1 when bound to its cognate DNA target sequence (Omichinski et al., 1993). Although the structures of the zinc-binding core of XPA and GATA-1 appear similar, the motif in GATA-1 requires only 28 additional C-terminal residues to bind its consensus DNA target while XPA requires an additional 82 C-terminal residues to bind damaged DNA (Kuraoka et al., 1996). Collectively, the 122-residue XPA fragment from Met^{98} to Phe^{219} is referred to as the XPA minimal damaged DNA binding domain (XPA-MBD) (Kuraoka et al., 1996). To better understand how the extra 82 residues in XPA associate with the zinc-binding core to recognize a broad spectrum of DNA lesions, the structure of XPA-MBD is being studied by NMR spectroscopy.

Methods and Results

The cDNA for XPA-MBD was cloned into the vector pET-11d, transfected into the host *E. coli* bacterial strain BL21(DE3)pLysS, and over-expressed in minimal medium (Miller) containing $^{15}NH_4Cl$ and [$^{13}C_6$]-D-glucose. The majority of XPA-MBD was expressed in the soluble fraction with a yield of approximately 60 mg/l. Pure XPA-MBD (>90%) was obtained using a one-step protocol on a POROS HQ/M strong anion exchange column. Analysis on a TSK-GEL G3000SW$_{XL}$ (7.8 mm × 30 cm, 5 µm) size exclusion column (50 mM K_2HPO_4, 200 mM NaCl, pH 7.0) showed one peak with a molecular weight expected for a monomeric 15 kDa protein. Western blot analysis of the purified protein using monoclonal XPA antibodies (Santa Cruz Biotechnology, CA) tested positive. Mass spectrometry showed the stoichiometric presence of one zinc atom per molecule of XPA-MBD. DNA filter binding assays using UV-damaged and undamaged linearized plasmid DNA (pET-11d) indicated that XPA-MBD was active, having a fourfold greater affinity for UV-damaged DNA (data not shown).

A 1.5 mM NMR sample of XPA-MBD was prepared in 600 µl of 90% H_2O/10% D_2O in the following NMR buffer: 20 mM K_2HPO_4, 100 mM KCl, 25 mM DTT-d_{10}, 50 µM ZnAc, 50 µM NaN_3, pH 7.3. 2D HSQC and 3D

*To whom correspondence should be addressed.
Supplementary material available from the authors: tables containing the ^{15}N, $^{13}C^{\alpha}$, ^{13}CO, ^{1}HN, and $^{1}H^{\alpha}$ chemical shifts.

Fig. 1. Two-dimensional HSQC spectrum of XPA-MBD (1.5 mM, NMR buffer, pH 7.3, 30 °C) collected at a ^{1}H resonance frequency of 750 MHz. To optimize resolution in the nitrogen dimension, a ^{15}N spectral width of 1350 Hz was used which resulted in the cross peaks for the underlined assignments folding back into the spectrum (sw1 = 1350 Hz, sw = 10 000 Hz, nt = 8, ni = 32, np = 1024). Cross-peak assignments labelled with an asterisk are side-chain amide resonances of Asn or Gln. The three cross peaks labelled only with ? are backbone amide resonances which cannot be unambiguously assigned.

HMQC-NOESY, HNCA, CBCA(CO)NH, HNCACB, HNCO, HCCH-TOCSY, and HCCH-TOCSY-NNH data* were collected at 30 °C using Varian 750- and 500-Unityplus spectrometers. The data were processed using Felix95 (MSI, San Diego, CA, U.S.A.) software. The ^{1}H, ^{13}C, and ^{15}N chemical shifts were referenced according to the method of Wishart et al. (1995) using external DSS (DSS = 0.0 ppm).

Extent of assignments and data deposition

Figure 1 is a 2D HSQC spectrum of XPA-MBD collected at a ^{1}H resonance frequency of 750 MHz. The wide chemical shift dispersion of the amide protons (~4 ppm) is characteristic of a protein that has adopted a unique secondary structure. The ^{1}HN resonances in Fig. 1 were assigned primarily using the CBCA(CO)NH and HNCACB data in conjunction with the amide-to-amide region of the HMQC-NOESY spectrum. The HNCA spectrum was used to verify the assignments. Using these assignments, the ^{13}CO and H$^{\alpha}$ resonances were determined from the HNCO, HCCH-TOCSY, and HCCH-TOCSY-NNH experiments. Only nine of the cross peaks in Fig. 1 could not be unambiguously assigned: four side-chain amide resonances of Asn^{210} and Asn^{169}, two side-chain resonances of Gln^{208}, Gln^{216}, or Gln^{185}, and three backbone ^{1}HN resonances.

*Pulse sequences for all experiments except the HMQC-NOESY (Varian User Library) were obtained with permission from the public accessible ftp site of Lewis Kay at the University of Toronto, Canada.

In total, 104 out of the 117 possible ^{1}HN resonances (122 residues minus four prolines and the terminal amide) were observed (89%) and 101 of these were assigned (86%). Of the missing 16 resonances, six are in a gap between Gln^{208}–Lys^{213} at the C-terminal end. The remaining gaps include three out of four residues between His^{171}–Gln^{174} and three two-residue gaps at Asp^{152}–Cys^{153}, Arg^{158}–Glu^{159}, and Lys^{217}–Lys^{218}. Because the majority of the missing ^{1}HN resonances are at the C-terminal end, it is unlikely that the absence of these resonances will seriously affect the structure calculations. Note that while the ^{1}HN resonances between Tyr^{116}–Leu^{123} were not observable in zXPA-41 at pH 6.3 (Buchko and Kennedy, 1997), these amide resonances are observable in the larger XPA fragment studied here at an elevated pH, indicating that the 83 additional C-terminal residues affect the stability of the zinc-binding core region. On the basis of the backbone assignments of XPA-MBD, 3D structure calculations are in progress using ^{1}H-^{1}H NOE-based distance restraints in distance geometry/simulated annealing protocols.

Acknowledgements

We gratefully acknowledge the helpful discussions with Dr. David F. Lowry. 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. G.W.B. and S.N. were supported by Associated Western Universities Inc., Northwest Division (AWU NW), under Grant DE-FG06-92RL-12451 with the U.S. Department of Energy.

References

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.

Buchko, G.W. and Kennedy, M.A. (1997) *J. Biomol. Struct. Dyn.*, **14**, 677–690.

Cleaver, J.E. (1968) *Nature*, **218**, 652–656.

Gunz, D., Hess, M.T. and Naegeli, H. (1995) *J. Biol. Chem.*, **271**, 25089–25098.

Kuraoka, I., Morita, E.H., Saijo, M., Matsuda, T., Morikawa, K., Shirakawa, M. and Tanaka, K. (1996) *Mutat. Res.*, **362**, 87–95.

Omichinski, J.G., Clore, G.M., Schaad, O., Felsenfeld, G., Trainor, C., Appella, E., Stahl, S.J. and Gronenborn, A.M. (1993) *Science*, **261**, 438–446.

Robins, P., Jones, C.J., Biggerstaff, M., Lindahl, T. and Wood, R.D. (1991) *EMBO J.*, **10**, 3913–3921.

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.

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.