Biochemical and Biophysical Research Communications **254,** 109–113 (1999)
Article ID bbrc.1998.9890, available online at http://www.idealibrary.com on IDEAL®

Extended X-Ray Absorption Fine Structure Evidence for a Single Metal Binding Domain in *Xenopus laevis* Nucleotide Excision Repair Protein XPA

Garry W. Buchko, Lilia M. Iakoucheva, Michael A. Kennedy, Eric J. Ackerman,[1] and Nancy J. Hess
Environmental Molecular Sciences Laboratory, Biosciences Department, and Biogeochemistry Resources, Pacific Northwest National Laboratories, Richland, Washington 99352

Received November 9, 1998

Nucleotide excision repair (NER) is an important cellular mechanism, conserved from bacteria to humans, responsible for eliminating multiple types of structurally distinct DNA lesions from the genome. The protein XPA appears to play a central role in NER, recognizing and/or verifyingdamaged DNA and recruiting other proteins, including RPA, ERCC1, and TFIIH, to repair the damage. Sequence analysis and genetic evidence suggest that zinc, which is essential for DNA binding, is associated with a C4-type motif, C-X_2-C-X_{17}-C-X_2-C. Sequence analysis suggests that a second, H2C2-type zinc-binding motif may be present near the C-terminal. Seventy percent of the amino acid sequence of *Xenopus laevis* XPA (xXPA) is identical to human XPA and both putative zinc-binding motifs are conserved in all known XPA proteins. Electrospray ionization–mass spectroscopy data show that xXPA contains only one zinc atom per molecule. EXAFS spectra collected on full-length xXPA in frozen (77 K) 15% glycerol aqueous solution unequivocally show that the zinc atom is coordinated to four sulfur atoms with an average Zn—S bond length of 2.33 ± 0.02 Å. Together, the EXAFS and mass spectroscopy data indicate that xXPA contains just one C4-type zinc-binding motif.

Many chemically and structurally distinct DNA lesions produced by a broad variety of physical and chemical DNA damaging agents, including natural toxins, man-made carcinogens, and UV and ionizing radiation (1, 2), are repaired by the ubiquitous, multienzyme, nucleotide excision repair (NER) pathway (3, 4). One of the major proteins involved in NER is XPA. Mutations to the human XPA gene (5) results in the most severe forms of the syndrome xeroderma pigmentosum (XP) (6), an autosomal recessive disease often characterized by neurological abnormalities, hypersensitivity to sunlight, and elevated levels of skin cancer. In humans, the XPA gene product (hXPA) is a 31-kDa protein of 273 amino acids (7) with no inherent catalytic activity. However, hXPA preferentially binds DNA damaged by UV, cisplatin and osmium tetraoxide with respect to undamaged DNA (8–10) and specifically interacts with other NER proteins including ERCC1 (11), TFIIH (12), and RPA (13), suggesting that XPA plays a central, multifunctional role in NER: recognizing and/or verifying (14) damaged-DNA and recruiting other proteins to repair the damage.

The mechanism of the XPA-DNA interactions are of great interest. The location of the DNA binding domain of hXPA has been reduced to a 122-amino-acid region between M98 and F219 (hXPA-MBD) (15). Sequence analysis shows that the hXPA-MBD contains a single class IV, C4-type, zinc-binding motif, C105-X_2-C108-X_{17}-C126-X_2-C129 (7). Zinc-binding domains play a major role in eukaryotic protein-nucleic acid interactions (16) and genetic evidence suggests that it is an essential functional motif in XPA because the replacement of C105, C108, C126, or C129 with a Ser severely reduces NER (10, 17).

Extended X-ray absorption fine structure (EXAFS) spectroscopy has been successfully used to characterize the local structural environment around metals in a variety of metalloproteins (18–20) by unambiguously identifying the element at the metal center and by providing information on interatomic distances, the number and type of ligand atoms, and the statistical and/or thermal disorder of the shells of atoms around the metal center (21, 22). Recently, we have obtained EXAFS spectra on a lyophilized sample of hXPA-MBD which unambiguously shows that hXPA-MBD contains a metal-binding motif and that the metal is coordinated to the sulfur atom of 4 cysteine residues (23). While a zinc-metal site has been unambiguously iden-

[1] To whom correspondence should be addressed at Pacific Northwest National Laboratories, Biosciences Department, Richland, WA 99352. Fax: 509-376-2149. E-mail: eric.ackerman@pnl.gov.

0006-291X/99 $30.00

```
hXPA MAAADGALPE AAALEQPAEL PASVRASIER KRQRALMLRQ ARLAARPYSA  50
xXPA *EPEPEPEQ* *NK-*EEKI* S*A***K*** N********* ****C***PT  49

     TAAAATGGMA NVKAAPKIID TGGGFILEEE EEEEQKIGKV VHQPGPVMEF 100
     GE-----*IS T***P**V** S****FI*** *A***HVEN* *R*****L*C  94

     DYVICEECGK EFMDSYLMNH FDLPTCDNCR DADDKHKLIT KTEAKQEYLL 150
     **L******* D******S** ***AV**S** **EE****** R********* 144

     KDCDLEKREP PLKFIVKKNP HHSQWGDMKL YLKLQIVKRS LEVWGSQEAL 200
     ****ID**** V****L**** *NTH****** ***A*VI*** ******E*** 194

     EEAKEVRQEN REKMKQKKFD KKVKELRRAV RSSVWKRETI VHQHEYGPEE 250
     *******KD* *D******** ********T* ***L**K*AS G********* 244

     NLEDDMYRKT CTMCGHELTY EKM                               273
     HV*E*S*K** *IT**Y*MN* ***                               267
```

Putative Zinc-Binding Domains:
C4-type: $C105X_2CX_{17}CX_2C129$
H2C2-type: $H242X_1HX_{16}CX_2C264$

FIG. 1. Amino acid alignment of hXPA and xXPA using the MegAlign program. Relative to hXPA, only the non-identical residues in xXPA are shown and the conserved residues are underlined. The general amino acid sequences of the 2 putative zinc-binding regions are provided below the sequence. Potential zinc-binding residues are highlighted in bold.

tified in the minimal DNA binding domain of hXPA, and its solution structure determined by NMR/distance geometry methods (24, 25), there is still some uncertainty about the number of zinc atoms bound to full-length XPA (10) because there is a possible second, H2C2-type, zinc-binding domain near the C-terminal (H242-C264) (7). To resolve the issue, we have obtained EXAFS data on full-length *Xenopus laevis* XPA (xXPA), a protein whose amino acid sequence shares 70% identity and 85% similarity (conserved replacement of amino acids) to hXPA (26) (Fig. 1).

MATERIALS AND METHODS

Preparation of xXPA. Plasmid pXPACXE1 containing the full-length *Xenopus* XPA cDNA was a kind gift from Professor Kenji Kohno (Nara Institute of Science and Technology, Japan). The complete coding sequence was cloned into the vector pET-11a and transfected into the host *Escherichia coli* bacterial strain BL21(DE3)pLysS (Novagen). After expression, three chromatographic steps were used to purify xXPA: anion exchange, cation exchange, and hydrophobic interaction chromatography (complete expression and purification protocols will be described fully elsewhere (Iakoucheva and Ackerman, in preparation)). Approximately 90 mg of ~99.5% pure xXPA, as determined by SDS polyacrylamide gel electrophoresis and Coomassie Blue staining, was obtained from 120 g of frozen cells (wet weight). The protein was dialyzed against storage buffer (50 mM Tris–HCl, 200 mM KCl, 10 mM DTT, 15% glycerol, pH 7.5), concentrated to 25 mg/ml using centrifugal ultrafiltration (Millipore, Bedford, MA), and then quickly frozen in liquid nitrogen for storage at −80°C. Mass determination of xXPA by electrospray ionization mass spectral (Finnigan TSQ 7000 triple-quadrupole mass spectrometer, San Jose, CA) confirmed the predicted molecular weight for a protein containing a 1:1 stoichiometric ratio of zinc.

EXAFS spectroscopy. EXAFS measurements were conducted at the Stanford Synchrotron Radiation Laboratory under dedicated operating conditions (3.0 GeV and 40 to 90 mA current). A 2 mM sample of xXPA containing 15% glycerol was loaded into a specially designed 125 μL cell holder and frozen at ~77 K with liquid nitrogen. Spectra for xXPA were collected at the Zn K-edge in the fluorescence mode, up to a photoelectron wavevector of 13 $Å^{-1}$, using an energy resolving 13-element Ge detector. Energy calibration for zinc was made by assigning the first inflection point in the absorption edge of Zn foil to 9659 eV. The absorption spectrum was normalized by fitting polynomials through the pre- and post-edge regions. At E_0, the value of the extrapolated pre-edge was set to zero and the difference between the extrapolations of the pre- and post-edge polynomials was set to unity.

The EXAFS oscillations were extracted by fitting a polynomial spline function through the postedge region and normalizing the difference between this approximation of the solitary-atom EXAFS and the actual data with the absorption decrease calculated using the McMaster tables (27). Fourier transforms were taken over photoelectron wavevector ranges that varied on the basis of the signal-to-noise ratio for each element. EXAFS nodes were selected as endpoints to the Fourier transform range and a two-sigma-wide Gaussian window was used to dampen the EXAFS oscillations at the end points.

The phase and amplitude for the cation-sulfur path were calculated using the *ab initio* code FEFF7.02 (28). The normalized phase and amplitudes of the zinc-sulfur scattering paths were used to fit the experimentally measured EXAFS. The ZnS F-4 M 3 structure was used to approximate the metal cation tetrahedral environment in both proteins. The number of sulfur atoms determined by using the normalized phase and amplitude to fit the ZnS EXAFS data was multiplied by a scale factor of 0.9 to give a sulfur atom value of 3.9 atoms for ZnS, a value in good agreement with the expected value of 4.0 atoms. Typically, scale factors between 0.8 and 1.0 are used to match the values calculated by FEFF7.02 (28).

RESULTS AND DISCUSSION

The recombinant xXPA was shown to be active by its ability to restore NER in *Xenopus* oocyte nuclear extracts in which cellular xXPA had been inactivated by the addition of an affinity-purified rabbit polyclonal antibody specific for xXPA (29).

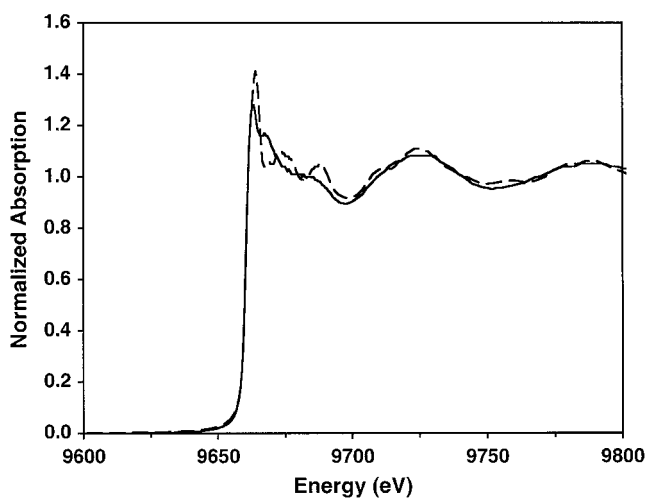

FIG. 2. Overlay of the Zn K-edge XANES spectra for xXPA in frozen (77 K) 15% glycerol aqueous solution (solid line) and the ZnS standard (300 K) (dashed line).

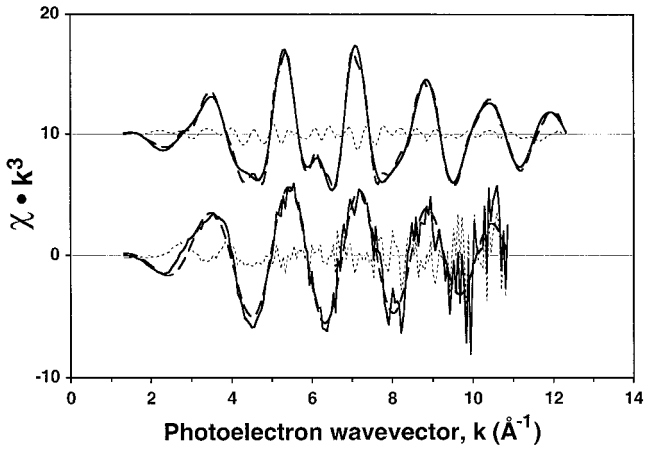

FIG. 3. EXAFS spectra for xXPA (bottom) and the ZnS standard (top) shown in Fig. 2. Solid line, experimental data; dashed line, fitted data; dotted line, residual data (experimental minus fitted).

Electrospray ionization mass spectra for xXPA confirmed a molecular weight predicted for a protein containing a 1:1 stoichiometric molar ratio of zinc. Such a ratio is in agreement with the value reported for the yeast *Saccharomyces cerevisiae* homologue of XPA, RAD14, using atomic absorption spectroscopy (30). Given the amino acid sequence similarities between xXPA, hXPA, and RAD14 (Fig. 1), we speculate that the conflicting 1:2 molar protein:zinc ratio observed by Asahina *et al.* (1994) is likely a result of sample preparation or instrumentation shortcomings.

The X-ray absorption near-edge structure (XANES) spectra for xXPA and ZnS is shown in Fig. 2. XANES spectra provide information on the metal valence and its site symmetry as well as serving as a fingerprint for the local structure around the absorber. For example, XANES can be used to determine the oxidation state of the absorbing element by measuring the energy shift of the absorption edge. With higher oxidation states the absorption edge shifts to higher energy by a few eV. Figure 2 illustrates that the absorption edge energies for xXPA and the ZnS standard are identical, with an absorption edge energy of 9660.91 eV, corresponding to the Zn^{2+} oxidation state. Note that directly above the absorption edge the XANES profile of the ZnS standard displays more spectral features than xXPA due to backscattering of Zn next nearest-neighbors which are absent in xXPA. However, further above the edge (>9700 eV), the spectrum is dominated by low frequency oscillations that result from the nearest-neighbor sulfur atoms.

The experimental (solid line) and fitted (dashed line) EXAFS spectra for xXPA and ZnS are shown in Fig. 3 between the photoelectron wavevector region of 1.3 to 11.0 Å for xXPA and 1.3 to 12.6 Å for ZnS. Analysis of the EXAFS oscillations provides quantitative information on the number and chemical identity of neighboring atoms as well as their distance from the absorber. The EXAFS for xXPA and ZnS are very similar even though the ZnS EXAFS contain an additional low amplitude, high frequency contribution from the second and third nearest-neighbor Zn atoms. The xXPA EXAFS are well fit with a single sulfur scattering wave as shown in Fig. 3 by the lack of oscillations in the small residual spectra (dotted line) remaining upon subtraction of the experimental data from the fit. The ZnS EXAFS are equally well fitted with a linear combination of individual scattering waves from the nearest neighbor sulfur and next-nearest neighbor zinc atoms. Table 1 contains the quantitative fitting results and shows that in both xXPA and the nearest-neighbors of the ZnS standard, the zinc is coordinated by four sulfur atoms with identical Zn—S bond lengths of 2.33 ± 0.2 Å.

Fourier transforms of the EXAFS oscillations for xXPA and ZnS are compared in Fig. 4. Because the peaks in the Fourier transforms are uncorrected for phase shifts they appear 0.2 to 0.5 Å shorter than the actual distance from the absorber to the neighboring atoms. The major peak observed in the Fourier transforms of the EXAFS for xXPA and ZnS is due to the backscattering amplitude of sulfur atoms. The position of the major peak is nearly identical for xXPA and ZnS indicating that the Zn-S bond length is very similar in the two compounds. The second peak observed in the Fourier transform of the EXAFS of ZnS is due to second nearest-neighbor Zn atoms.

The Zn-S bond length of 2.33 Å determined for xXPA is identical to the bond length obtained from EXAFS

TABLE 1

Fitting Results to the EXAFS Spectra for xXPA (15% Frozen Aqueous Glycerol), hXPA-MBD (Lyophilized), and ZnS[a]

Sample	Zn-S distance (Å)	Number of sulfur atoms	Debye waller	ΔE_0	r^2
xXPA	2.33 ± 0.02	4.3 ± 1.0	0.0025 ± 0.0004	−1.0 ± 3.3	1.07865
hXPA-MBD	2.34 ± 0.01	4.4 ± 0.9	0.0009 ± 0.0016	−1.4 ± 3.1	0.8114
ZnS	2.33 ± 0.02	3.9 ± 0.7	0.0016 ± 0.0004	−2.7 ± 2.4	0.8893

[a] The uncertainties were determined as the quantity that a parameter could be varied from the value giving the best fit as to increase the r^2 value by 10% from its minimum value. Data for hXPA-MBD are from Hess *et al.*, (23).

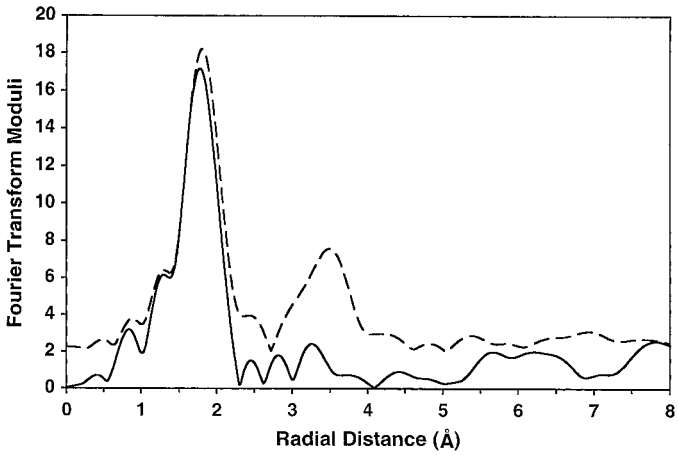

FIG. 4. Fourier transforms of the xXPA and ZnS EXAFS profiles over the photoelectron wavevector region from ~2 to ~11 Å. The major peak in each of the Fourier transforms is due to sulfur atoms. The second peak in the ZnS standard consists of second nearest-neighbor Zn atoms. The peaks in the Fourier transforms are uncorrected for phase shifts and appear 0.2 to 0.5 Å shorter than the actual distance from the absorber to the neighboring atoms.

data for another C4-type, zinc-binding domain in the UvrA protein of *Escherichia coli*, a damage recognition subunit of ABC excision nuclease (18). Like UvrA, there is no evidence that the metal in xXPA is coordinated to the imidazole nitrogen of a histidine residue. Proteins containing zinc-binding motifs tetrahedrally coordinated to one (C3H1) or two (C2H2) histidines are common (16, 19). Such EXAFS profiles display oscillations composed of more than a single component due to the presence of a shorter Zn-N bond (1.94 Å) and the corresponding Fourier transforms have additional peaks (20). Consequently, the second, H2C2-type zinc-binding motif predicted in XPA (7) does not exist. The EXAFS data, in combination with the electrospray mass spectrometry data, indicate that xXPA contains only one C4-type zinc binding domain (Table 1).

The Zn-S bond length of 2.33 Å determined for full-length xXPA is, within the uncertainty limits, identical to a value of 2.34 Å determined for a lyophilized (23) and a frozen 15% glycerol aqueous (Hess and Buchko, unpublished results) sample of the minimal DNA binding domain of human XPA. Like xXPA, the EXAFS and electrospray mass spectrometry data show that hXPA-MBD contains one C4-type zinc-binding domain (Table 1) (23) and a substantial body of indirect evidence indicates that C105, C108, C126 and C129 are the 4 cysteine residues involved in metal chelation (10,17). Because the amino acid sequence of xXPA shares 70% identity and 85% similarity to hXPA, with 97 of the 122 residues identical in the minimal DNA binding region, it is likely that C99, C102, C120, and C123 are the residues involved in metal chelation in xXPA.

ACKNOWLEDGMENTS

We thank Drs. Naxing Xu, Jim Bruce, and Richard D. Smith (PNNL) for the mass spectral analyses and Dr. Britt Hedman (SSRL) for useful discussions. The Stanford Synchrotron Radiation Laboratory is funded by Basic Energy Sciences, Department of Energy. 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 L.M.I. 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

1. Coverley, D., Kenny, M. K., Munn, M., Rupp, W. D., Lane, D. P., and Wood, R. D. (1991) *Nature* **349,** 538–541.
2. Satoh, M. S., Jones, C. J., Wood, R. D., and Lindahl, T. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 6335–6339.
3. Sancar, A. (1995) *J. Biol. Chem.* **270,** 15915–15918.
4. Wood, R. D. (1997) *J. Biol. Chem.* **272,** 23454–23468.
5. Tanaka, K. (1993) *Jpn. J. Hum. Genet.* **38,** 1–14.
6. 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.), pp. 4493–4419, McGraw-Hill, New York.
7. 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.
8. Robins, P., Jones, C. J., Biggerstaff, M., Lindahl, T., and Wood, R. D. (1991) *EMBO J.* **10,** 3913–12104.
9. Jones, C. J., and Wood, R. D. (1993) *Biochemistry* **32,** 12096–12104.
10. 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.
11. Li, L., Elledge, S. J., Peterson, C. A., Bales, E. S., and Legerski, R. J. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 5012–5016.
12. Park, C-H., Mu, D., Reardon, J. T., and Sancar, A. (1995) *J. Biol. Chem.* **270,** 4896–4901.
13. Matsuda, T., Saijo, M., Kuraoka, I., Kobayashi, T., Nakatsu, Y., Nagai, A., Enjoji, T., Masutani, C., Sugasawa, K., Hanoaka, F., Yasui, A., and Tanaka, K. (1995) *J. Biol. Chem.* **270,** 4152–4157.
14. 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.
15. Kuraoka, I, Morita, E. H., Saijo, M., Matsuda, T., Morikawa, K. Shirakawa, M., and Tanaka, K. (1996) *Mutat. Res.* **362,** 87–95.
16. Berg, J. M., and Shi, Y. (1996) *Science* **271,** 1081–1086.
17. Miyamoto, I., Miura, N., Niwa, H., Miyazaki, J., and Tanaka, K. (1992) *J. Biol. Chem.* **67,** 12182–12187.
18. Navaratnam, S., Myles, G. M., Strange R. W., and Sancar, A. (1989) *J. Biol. Chem.* **264,** 16067–16071.
19. Summers, M. F., Henderson, L. E., Chance, M. R., Bess, J. W., Jr., South, T. L., Blake, P. R., Sagi, I., Perez-Alvarado, G., Sowder, R.C., III. Hare, D. R., and Arthur, L. O. (1992) *Protein Sci.* **1,** 563-574.
20. Strange, R. W., Murphy, L. M., Karlsson, B. G., Reinhammar, B., and Hasnain, S. S. (1996) *Biochemistry* **35,** 16391–16398.
21. Teo, B. K. (1981) *in* EXAFS Spectroscopy: Techniques and Applications (Teo, B. K., and Joy, D. C., Eds.), pp 13–58, Plenum, New York.
22. Brown, G. E., Jr., Calas, G., Waychunas, G. A., and Petiau, J. (1988) *Rev. Mineral.* **18,** 431–512.

23. 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.
24. Buchko, G. W., Ni, S., Thrall, B. D., and Kennedy, M. A. (1998) *Nucleic Acids Res.* **26,** 2779–2788.
25. Ikegami, T., Kuraoka, I., Saijo, M., Kodo, N., Kyogoku, Y., Morikawa, K., Tanaka, K., and Shirakawa, M. (1998) *Nature Struct. Biol.* **5,** 701–706.
26. Shimamoto, T., Kohno, K., Tanaka, K., and Okada, Y. (1991) *Biochem. Biophys. Res. Commun.* **181,** 1231–1237.
27. McMaster, W. H., Kerr del Grande, N., Mallett, J. H., and Hubbell, J. H. (1969) *in* Compilation of X-Ray Cross Sections, pp. 350, Univ. California, Livermore, CA.
28. Rehr, J. J., Mustre de Leon, J., Zabinsky, S. I., and Albers, R. C. (1991) *J. Am. Chem. Soc.* **113,** 5135–5140.
29. Oda, N., Saxena, J. K., Jenkins, T. M., Prasad, R., Wilson, S. H., and Ackerman, E. J. (1996) *J. Biol. Chem.* **271,** 13816–13820.
30. Guzder, S. N., Sung, P., Prakash, L., and Prakash, S. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 5433–5437.