Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The absorption edge of protein-bound mercury and a double-edge strategy for HgMAD data acquisition

Xinhua Ji,* Jaroslaw Blaszczyk and Xin Chen†

Macromolecular Crystallography Laboratory, National Cancer Institute, Frederick, MD 21702, USA

† Present address: Department of Molecular Biology, Princeton University, Princeton, NJ 08544, USA.

Correspondence e-mail: jix@ncifcrf.gov

Received 8 December 2000
Accepted 16 April 2001

The L_{III} absorption edge of protein-bound mercury (Hg) has been experimentally determined using X-ray data collection from a crystal. This absorption edge is 12 291 eV, 4 eV higher than the theoretical value of elemental Hg. Considering the possible shift of the Hg absorption edge with the chemical environment in different protein crystals, a double-edge strategy for multiwavelength anomalous diffraction (MAD) data collection has been developed. The approach provides a convenient way to optimize the dispersive signal between a remote wavelength and two edge wavelengths separated from each other by 4 eV. The dispersive signals derived from both edges are used, along with anomalous signals, in MAD phasing and phase refinement. This approach has been used in the crystal structure determination of three proteins containing one Hg atom per 186–196 amino-acid residues at 2.0, 2.6 and 2.7 Å resolution. A set of four wavelengths is recommended for HgMAD data acquisition: 1.0087 Å (12 291 eV, edge1), 1.0084 Å (12 295 eV, edge2), 1.0064 Å (12 320 eV, peak) and 0.9918 Å (12 500 eV, remote). Although it is no longer necessary to determine the L_{III} absorption edge of protein-bound Hg experimentally, an initial fluorescence scan on the crystal for data collection is still necessary to verify the existence of Hg in the crystal.

1. Introduction

Phase determination from multiwavelength anomalous diffraction (MAD) measurements has now become a routine approach in solving macromolecular crystal structures (Hendrickson & Ogata, 1997). In principle, MAD measurements can be performed using any anomalously scattering element. However, technical limitations reduce the choices to those with an X-ray absorption edge in the wavelength range feasible at a synchrotron beamline. The most commonly used absorption edges are the *K* edges of Se, Br, Fe and Cu and the L_{III} edges of Hg, Au, Pt and lanthanides. Se is usually introduced into macromolecules by genetic engineering (Hendrickson *et al.*, 1990; Doublie, 1997), Br can be introduced by a quick soak of protein crystals (Dauter & Dauter, 1999; Dauter *et al.*, 2000), Fe and Cu are inherent in native proteins and heavy metals are usually soaked into protein crystals, as in the multiple isomorphous replacement (MIR) method. A complete list of the accessible absorption edges for MAD experiments with a conventional beamline has been published by Hendrickson & Ogata (1997).

In principle, two wavelengths, with the measurements of $F(+)_{\lambda1}$, $F(-)_{\lambda1}$, $F(+)_{\lambda2}$ and $F(-)_{\lambda2}$, allow unique phase determination (Peterson *et al.*, 1996; Evans & Wilson, 1999;

González *et al.*, 1999). However, to optimize both dispersive (f') and Bijvoet (f'') differences, at least three wavelengths are needed (Hendrickson & Ogata, 1997). To choose the wavelengths, the absorption spectrum of the crystal for data collection is essential, because the absorption-edge value shifts with the chemical environment. Although sharp edges for most commonly used anomalously scattering elements can be observed using protein crystals, the absorption edge for Hg is very difficult to determine in a similar fashion because the Hg L_{III} absorption spectrum obtained from an Hg-containing protein crystal, which is directly proportional to the imaginary component of the anomalous scattering factor, is usually flat with many turning points. Therefore, the real component of the scattering factor, which is derived from the absorption spectrum by using the Kramers–Kronig transformation, usually has multiple minima. The first experimental measurement of X-ray fluorescence from the L_{III} edge of Hg was performed with a frozen solution of Hg-derivatized GMP synthetase (Tesmer *et al.*, 1994), revealing that the L_{III} absorption edge for cysteinyl Hg in GMP synthetase is very similar to the theoretical edge for elemental Hg, corresponding to a wavelength of 1.009 Å. Thereafter, the edge wavelength used in various HgMAD experiments ranged from 1.0088 to 1.0095 Å (see, for example, Hubbard *et al.*, 1994; Krishna *et al.*, 1994; Tesmer *et al.*, 1994; Benson *et al.*, 1995; Friedman *et al.*, 1995; Georgiadis *et al.*, 1995; Goldberg *et al.*, 1995; Stebbins *et al.*, 1995). In 50% of cases, the HgMAD data were not sufficient for structure determination and therefore other sources of phase information were necessary (see, for example, Benson *et al.*, 1995; Georgiadis *et al.*, 1995; Goldberg *et al.*, 1995; Stebbins *et al.*, 1995). Recently, we have observed the L_{III} absorption edge of protein-bound Hg with a single crystal and developed a double-edge approach for HgMAD data acquisition in three experiments. This approach has provided sufficient phase information for the crystal structure determination of all three macromolecules: ERA (GTPase-dependent cell-cycle regulator; Chen *et al.*, 1999), GK42, (dimeric unligated guanylate kinase; Blaszczyk *et al.*, in preparation) and GK51 (monomeric unligated guanylate kinase; Blaszczyk *et al.*, 2001). With the development and accessibility of third-generation synchrotron beamlines, a typical four-wavelength data collection (a scan of 120° per wavelength at 2.25 Å resolution) takes only 23 min to perform (Walsh, Dementieva *et al.*, 1999; Walsh, Evans *et al.*, 1999). Therefore, the number of wavelengths to use in a MAD experiment is not necessarily a limiting factor. Of course, longer data-collection time is expected for a second-generation synchrotron light source.

2. Materials and methods

2.1. Hg derivatization

Crystals of ERA, GK42 and GK51 were Hg derivatized by soaking with the crystallization reservoir solution containing 10–20 m*M* ethylmercury phosphate (EMP). Soaking was carried out at controlled room temperature and the soaking time ranged from 1 to 24 h, optimized according to pilot experiments conducted with in-house X-ray diffraction facilities.

2.2. Fluorescence scan, wavelength selection and data collection

Hg-containing crystals were used for the fluorescence scan. The scan near the L_{III} absorption edge of Hg was performed with a fluorescence detector (a photomultiplier tube from Thorn EMI Electron Tubes Inc. with a phosphor on the face) mounted on beamline X9B at the National Synchrotron Light Source (NSLS), Brookhaven National Laboratory. Of the three Hg-derivatized crystals, the ERA crystal gave the best absorption curve (Fig. 1*a*), which was much smoother than those for GK42 and GK51 (Fig. 1*b*). The Kramers–Kronig transformation of the ERA absorption curve resulted in a single minimum that defined the accurate position of the L_{III}

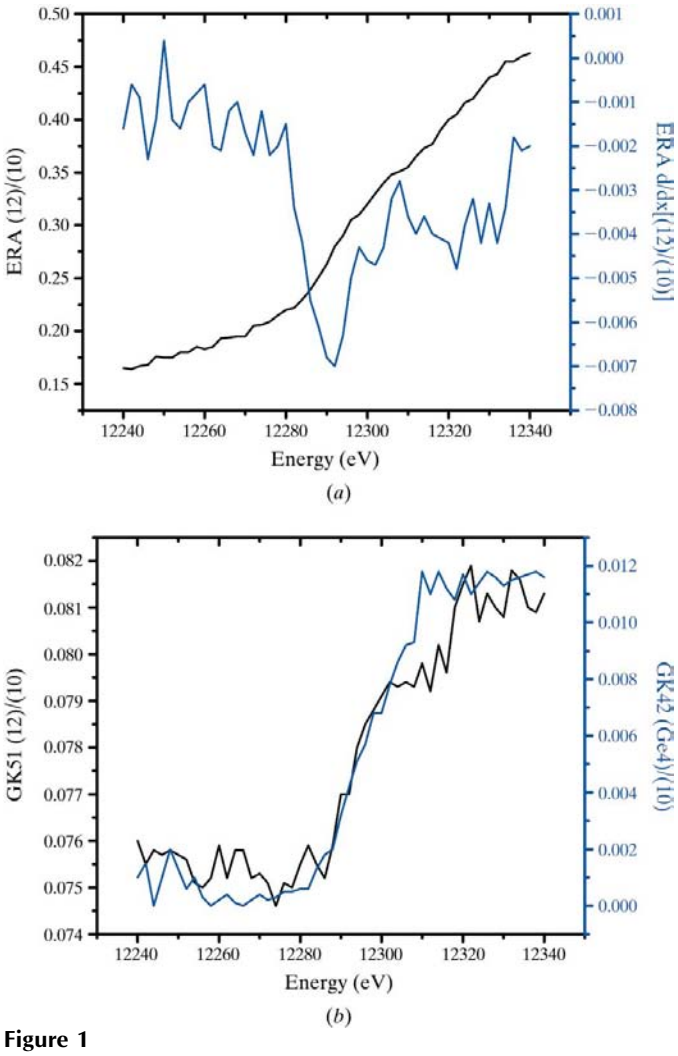

Figure 1
X-ray fluorescence spectra. (*a*) Absorption curve (black) and its Kramers–Kronig transformation (blue) for ERA (Chen *et al.*, 1999). (*b*) Absorption curves for GK42 (blue; Blaszczyk *et al.*, in preparation) and GK51 (black; Blaszczyk *et al.*, 2001).

absorption edge of ERA-bound Hg, which was 12 291 eV (Fig. 1*a*), 4 eV higher than the theoretical value and corresponding to a wavelength of 1.0087 Å. In contrast, multiple minima were observed for the Kramers–Kronig transformation of the absorption curves of GK42 and GK51; no accurate absorption edge could be experimentally derived for GK42 and GK51. Considering the possible shift of the absorption edge with the chemical environment and the difficulty in obtaining the Hg absorption edge experimentally (Fig. 1), we tested a second edge value that was 4 eV higher or lower than 12 291 eV (Table 1). All three HgMAD data sets were collected on NSLS beamline X9B on a Quantum-4 CCD detector at cryogenic temperature (100 K) maintained with an Oxford Cryosystem. A randomly oriented single crystal was used for the entire data collection with four wavelengths. Complete data were collected at each wavelength with a single pass. Neither the mirror-symmetry protocol (Hendrickson & Ogata, 1997) nor the inverse-beam method (Hendrickson & Ogata, 1997) was used. X-ray data were processed using *HKL*2000 (Otwinowski & Minor, 1997).

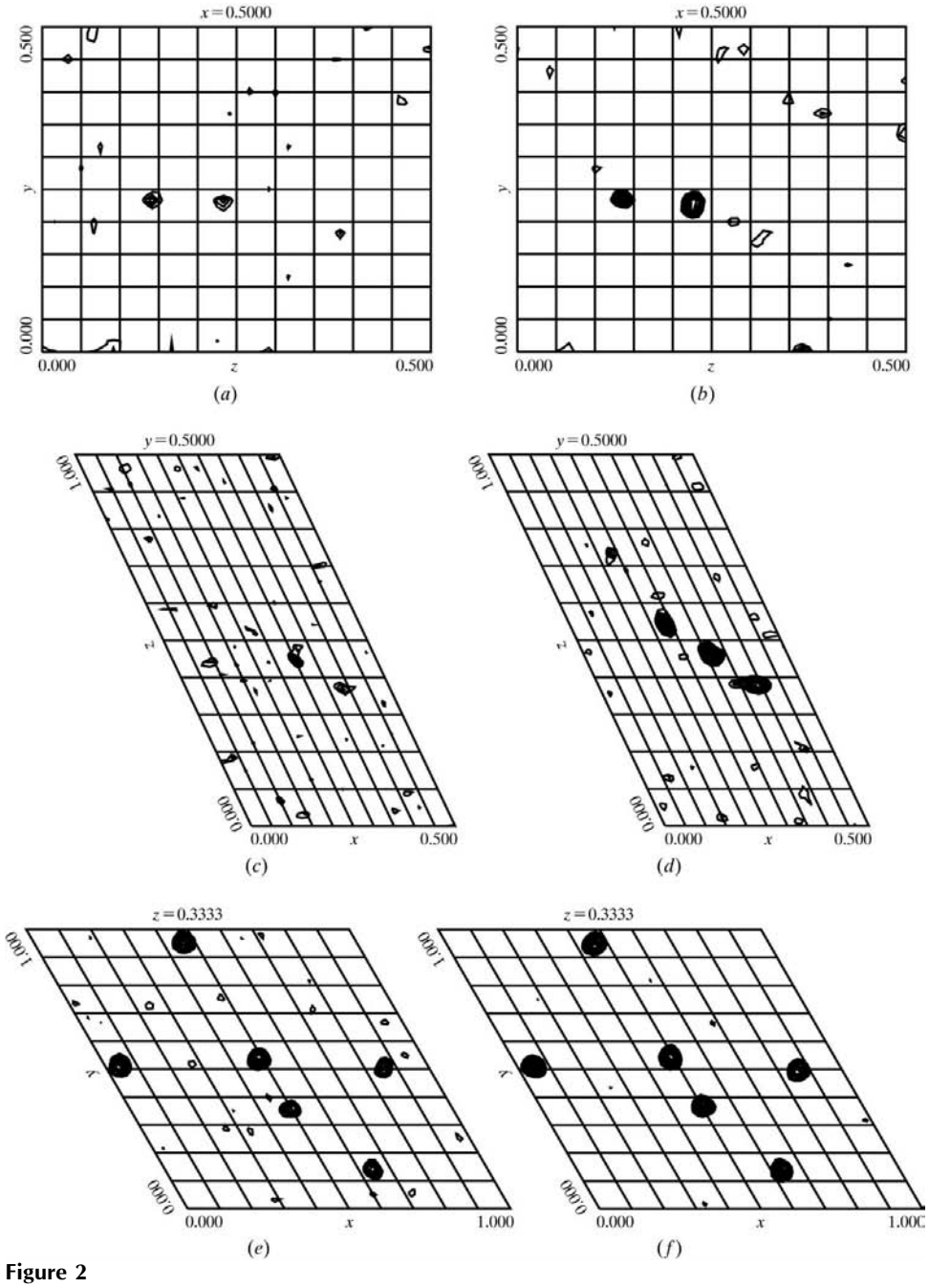

Figure 2
Difference Patterson maps. Harker sections for (*a*, *b*) GK42 (Blaszczyk *et al.*, in preparation), (*c*, *d*) ERA (Chen *et al.*, 1999) and (*e*, *f*) GK51 (Blaszczyk *et al.*, 2001). (*a*), (*c*) and (*e*) are dispersive Pattersons and (*b*), (*d*) and (*f*) are anomalous Pattersons, respectively. The maps were generated using *PHASES* (Furey & Swaminathan, 1997).

2.3. Location of Hg atoms and phase refinement

The Hg atoms in the three crystals were located either *via* the interpretation of difference Patterson maps (Fig. 2) or with the program package *SOLVE* (Terwilliger & Berendzen, 1999). A pseudo-MIR approach (Ramakrishnan & Biou, 1997) embedded in *PHASES* (Furey & Swaminathan, 1997) was then used to refine the positional parameters of Hg sites and to calculate the phase angles. The remote wavelength was used as the 'native' and the remaining wavelengths were used as 'derivatives'. The resulting phases were improved with the density-modification procedure *DM* (Cowtan, 1994). Phase improvement was also carried out with the program package *SHARP* (de La Fortelle & Bricogne, 1997), which resulted in experimental electron-density maps of very high quality.

3. Results

3.1. Hg derivatization of crystals with EMP soaking

Both GK and ERA contain a single free cysteine residue per polypeptide chain, which readily binds an Hg atom. All three crystals have two molecules in the asymmetric unit and therefore two Hg atoms were located in each of the three cases. In addition, a half-occupied Hg atom was located near a His residue in ERA; thus, a total of three Hg atoms was found in the asymmetric unit of ERA (Table 1).

Table 1
HgMAD phasing statistics of GK42 (Blaszczyk *et al.*, in preparation), ERA (Chen *et al.*, 1999) and GK51 (Blaszczyk *et al.*, 2001).

NA, not applicable.

	GK42	ERA	GK51	Optimized
Space group	$P2_12_12_1$	$P2_1$	$P3_1$	
Unit-cell parameters				
a, b, c (Å)	44, 77, 92	87, 66, 88	53, 53, 117	
α, β, γ (°)	90, 90, 90	90, 115, 90	90, 90, 120	
No. Hg/No. amino acids	2/372	3/590	2/372	
Resolution (Å)	2.0	2.6	2.7	
Wavelength selection (Å)				
Edge1	1.0087	1.0087	1.0087	1.0087
Edge2	1.0090	1.0090	1.0084	1.0084
Peak	1.0072	1.0076	1.0076	1.0064
Remote	0.9958	0.9918	0.9918	0.9918
Phasing power/figure of merit				
$\Delta f'_{\rm edge1-remote}$	1.38/0.26	0.57/0.17	2.14/0.39	
$\Delta f'_{\rm edge2-remote}$	1.51/0.25	0.93/0.24	2.11/0.38	
$f''_{\rm peak}$	3.14/0.47	2.27/0.38	2.52/0.42	
$f''_{\rm edge1}$	2.81/0.45	1.67/0.34	1.29/0.27	
$f''_{\rm edge2}$	1.96/0.33	2.07/0.37	2.31/0.41	
$f''_{\rm remote}$	3.09/0.47	1.35/0.32	2.95/0.45	
Overall	NA/0.65	NA/0.59	NA/0.77	

3.2. Data collection

X-ray data statistics are presented in Table 2. The HgMAD data, with a redundancy of 3–5 and overall completeness >98%, contain sufficient phase information for the crystal structure determination of all three proteins. At cryogenic temperatures, MAD data collection could be performed using only one crystal. Special techniques for MAD data collection (mirror symmetry and inverse beam) were not used, which reduced the data-collection time significantly.

3.3. Double-edge approach

We have experimentally determined the $L_{\rm III}$ absorption edge of protein-bound Hg, 1.0087 Å (edge1), using the Hg-derivatized ERA crystal (Fig. 1*a*), which has been used in three HgMAD experiments. The second edge (edge2) is 1.0090 Å for GK42 and ERA, which is 4 eV lower than edge1, whereas it is 1.0084 Å for GK51, which is 4 eV higher than edge1. The data statistics show that both edge2 wavelengths contain dispersive information comparable to that of edge1, whereas the edge2 with higher energy provides more anomalous information than the edge2 with lower energy (Table 1). Therefore, the use of two edges not only tolerates the shift of the absorption edge of protein-bound Hg with different chemical environments and at any synchrotron beamline, but also provides more phase information (Table 1).

3.4. Peak wavelength

The absorption peak wavelength is 1.0072 Å for GK42 and 1.0076 Å for both ERA and GK51; the statistics show that the shorter wavelength provides the stronger anomalous signal (Table 1). Therefore, the peak wavelength should be chosen with even higher energy; the absorption curves for GK42 and GK51 suggest a value of ∼12 320 eV, corresponding to a wavelength of ∼1.0064 Å.

Table 2
Data-collection statistics (averaged over four wavelengths) for GK42 (Blaszczyk *et al.*, in preparation), ERA (Chen *et al.*, 1999) and GK51 (Blaszczyk *et al.*, 2001).

	GK42	ERA	GK51
Resolution (Å)	2.0	2.6	2.7
Redundancy	3.4	5.5	2.8
Overall completeness (%)	98.6	99.5	99.4
Last-shell† completeness (%)	94.3	99.1	93.6
Overall $R_{\rm sym}$	0.071	0.068	0.080
Last-shell† $R_{\rm sym}$	0.570	0.488	0.240
Resolution where $R_{\rm sym} > 0.20$ (Å)	2.00–2.31	2.60–2.85	2.70–2.85
Resolution where $I/\sigma(I) < 3.0$ (Å)	2.00–2.15	2.60–2.80	2.70–2.75

† The outermost shell with highest resolution: 2.00–2.07 Å for GK42, 2.60–2.69 Å for ERA and 2.70–2.75 Å for GK51.

3.5. Remote wavelength

The remote wavelength is 0.9958 Å for GK42 and 0.9918 Å for both ERA and GK51 and the statistics favor the shorter wavelength (Table 1). Considering the increased systematic error and decreased anomalous information with the remoteness of the wavelength, 0.9918 Å (12 500 eV) is appropriate, which is 180 eV higher than the peak wavelength (12 320 eV), as suggested in the previous section.

4. Conclusions

We have determined the $L_{\rm III}$ absorption edge of protein-bound Hg using an ERA crystal and optimized a set of four wavelengths for HgMAD data acquisition: 1.0087 Å (12 291 eV, edge1), 1.0084 Å (12 295 eV, edge2), 1.0064 Å (12 320 eV, peak) and 0.9918 Å (12 500 eV, remote), with which at least six items of optimized phase information can be derived, including two dispersive differences and four anomalous signals (Table 1). This set of four wavelengths can generally be used in every HgMAD experiment, although an initial fluorescence scan is still necessary to verify the presence of Hg in the crystal.

We thank Dr Zbigniew Dauter for assistance during data collection at the synchrotron beamline X9B at Brookhaven National Laboratory and extensive discussion during the development of the double-edge strategy in HgMAD experiments, and Dr Alexander Wlodawer and Ms Anne Arthur for critical reading of the manuscript.

References

Benson, T. E., Filman, D. J., Walsh, C. T. & Hogle, J. M. (1995). *Nature Struct. Biol.* **2**, 644–653.
Blaszczyk, J., Li, Y., Shi, G., Yan, H. & Ji, X. (2001). *J. Mol. Biol.* **307**, 247–257.
Chen, X., Court, D. L. & Ji, X. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 8396–8410.

Cowtan, K. (1994). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **31**, 34–38.
Dauter, Z. & Dauter, M. (1999). *J. Mol. Biol.* **289**, 93–101.
Dauter, Z., Dauter, M. & Rajashankar, K. R. (2000). *Acta Cryst.* D**56**, 232–237.
Doublie, S. (1997). *Methods Enzymol.* **276**, 523–530.
Evans, G. & Wilson, K. S. (1999). *Acta Cryst.* D**55**, 67–76.
Friedman, A. M., Fischmann, T. O. & Steitz, T. A. (1995). *Science*, **268**, 1721–1727.
Furey, W. & Swaminathan, S. (1997). *Methods Enzymol.* **277**, 590–620.
Georgiadis, M. M., Jessen, S. M., Ogata, C. M., Telesnitsky, A., Goff, S. P. & Hendrickson, W. E. (1995). *Structure*, **3**, 879–892.
Goldberg, J., Huang, H.-B., Kwon, Y.-G., Greengard, P., Nairn, A. C. & Kuriyan, J. (1995). *Nature (London)*, **376**, 745–753.
González, A., Pédelacq, J., Solà, M., Gomis-Rüth, F. X., Coll, M., Samama, J. & Benini, S. (1999). *Acta Cryst.* D**55**, 1449–1458.
Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.
Hendrickson, W. A. & Ogata, C. M. (1997). *Methods Enzymol.* **276**, 494–523.
Hubbard, S. R., Wei, L., Ellis, L. & Hendrickson, W. A. (1994). *Nature (London)*, **372**, 746–754.
Krishna, T. S., Kong, X.-P., Gary, S., Burgers, P. M. & Kuriyan, J. (1994). *Cell*, **79**, 1233–1243.
La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Peterson, M. R., Harrop, S. J., McSweeney, S. M., Leonard, G. A., Thompson, A. W., Hunter, W. N. & Helliwell, J. R. (1996). *J. Synchrotron Rad.* **3**, 24–34.
Ramakrishnan, V. & Biou, V. (1997). *Methods Enzymol.* **276**, 538–557.
Stebbins, C. E., Borukhov, S., Orlova, M., Polyakov, A., Goldfarb, A. & Darst, S. A. (1995). *Nature (London)*, **373**, 636–640.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Tesmer, J. J., Stemmler, T. L., Penner-Hahn, J. E., Davisson, V. J. & Smith, J. L. (1994). *Proteins Struct. Funct. Genet.* **18**, 394–403.
Walsh, M. A., Dementieva, I., Evans, G., Sanishvili, R. & Joachimiak, A. (1999). *Acta Cryst.* D**55**, 1168–1173.
Walsh, M. A., Evans, G., Sanishvili, R., Dementieva, I. & Joachimiak, A. (1999). *Acta Cryst.* D**55**, 1726–1732.