Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Using barium ions for heavy-atom derivatization and phasing of xylanase II from *Trichoderma longibrachiatum*

Natalia Moiseeva and Marc Allaire*

National Synchrotron Light Source, Brookhaven National Laboratory, Building 725D, Upton, NY 11973, USA

Correspondence e-mail: allaire@bnl.gov

Received 30 May 2007
Accepted 13 July 2007

This paper describes the use of barium chloride to produce a heavy-atom derivative of xylanase II crystals from *Trichoderma longibrachiatum*, which was obtained either by cocrystallization or soaking. SAD phasing led to interpretable electron-density maps that allowed unambiguous chain tracing. In the best case, with a data set collected at 9.5 keV, 88% of the residues were built, with 83% of the side chains assigned. The barium ions are found to mainly interact with main-chain carbonyl groups and water molecules. It is suggested that barium ions could also be used as a potential anomalous scatterer in the quick cryosoaking procedure for phasing.

1. Introduction

The determination of phase estimates is an important step in macromolecular structure solution using X-ray crystallography. Traditional soaking methods include a broad range of heavy-metal compounds and require optimization of the soaking time. The current emphasis on high-throughput biological structure determination has led to the development of methods to prepare heavy-atom derivatives that are more reliable and less disruptive than traditional heavy-atom soaking.

The multiple-wavelength anomalous diffraction (MAD) method at the Se *K* absorption edge combined with the incorporation of Se in proteins by replacing methionine with selenomethionine has become a routine method in macromolecular crystallography (Hendrickson & Ogata, 1997). Alternative methods to obtain derivatives for anomalous diffraction experiments have been proposed, such as the 'quick cryosoaking' procedure for derivatization by the incorporation of monovalent halide anions (Dauter *et al.*, 2000) and monovalent and polyvalent cations (Nagem *et al.*, 2001) as well as the use of noble gases such as Xe and Kr (Vitali *et al.*, 1991; Schiltz *et al.*, 1994; Cohen *et al.*, 2001). The anomalous signal from the S atoms in sulfur-rich proteins has increasingly been used to solve structures using the SAD method (Hendrickson & Teeter, 1981; Wang, 1985; Dauter *et al.*, 1999). Recently, traditional heavy-atom derivatization has been revisited by introducing short soaking (Sun & Radaev, 2002; Sun *et al.*, 2002). Techniques such as mass spectrometry (Cohen *et al.*, 2000) or native-gel electrophoresis (Boggon & Shapiro, 2000) have been used to screen for derivatives.

Barium is a small stable divalent ion (ionic radius 1.35 Å; Sharpe, 1992). The barium absorption edges (*K* edge 37.4 keV, *L* edges 5.2–6.0 keV) lie beyond the easily accessible wavelength range on commonly available synchrotron beamlines. However, barium has a significant anomalous signal component in the energy range 7–13 keV (wavelength range ∼1.7–1.0 Å) which can be exploited for the phasing of protein structures. The present paper is an attempt to answer the following: (i) whether these ions can be introduced under gentle conditions into protein crystals and (ii) whether there will be a

sufficient number of ordered barium sites that can be used for phasing and for automated model building. The protein used in this study is xylanase II from *Trichoderma longibrachiatum* (PDB code 1xyp; Törrönen & Rouvinen, 1995).

2. Experimental

Xylanase (xylanase II from *T. longibrachiatum*) was crystallized as described previously (Moiseeva & Allaire, 2004). The barium chloride used for heavy-atom derivatization was of analytical grade (Sigma–Aldrich, USA). The Ba-derivative xylanase crystals were obtained either by cocrystallization using 0.2 *M* $BaCl_2$ instead of 0.2 *M* $CaCl_2$ in crystallization wells or by soaking native-grown crystals in crystallization buffer containing 0.5 *M* $BaCl_2$ instead of $CaCl_2$ for 2–3 h. Prior to data collection, the crystals were cryoprotected by the addition of 30%(*v*/*v*) glycerol to the barium-containing crystallization buffer followed by flash-freezing the crystal in a stream of nitrogen maintained at 100 K.

Diffraction data were collected from one crystal of cocrystallized and one crystal of soaked Ba-derivatized xylanase at five different wavelengths, with the longest wavelength collected first, on beamline X6A (Allaire *et al.*, 2003) at the National Synchrotron Light Source at Brookhaven National Laboratory (Upton, NY, USA) using an ADSC Q210 CCD detector. The wavelengths were chosen to cover a range of 7.3–12.7 keV. The diffraction data were collected over a total angular rotation range of 360° with an oscillation range of 1.0° and an exposure time of 10 s. No attempts were made to measure Bijvoet-related reflections close in time or on the same diffraction image. The diffraction data were processed and scaled using *HKL*-2000 (Otwinowski & Minor, 1997) and further data analysis used the *CCP*4 suite of programs (Collaborative Computational Project, Number 4, 1994).

In order to evaluate the ability of the Ba derivatives to solve the xylanase structure, *de novo* SAD phasing was undertaken on each data set. Data were used up to the maximum resolution available at the particular wavelength, *i.e.* no common resolution cutoff was applied. Barium anomalous scattering factors were estimated using the program *CROSSEC* (Cromer, 1983). The expected amount of anomalous signal (Hendrickson & Ogata, 1997) was calculated with $N = 2$ barium sites. The observed amount of anomalous signal was calculated using *phenix.xtriage* (Zwart *et al.*, 2005). The program *SOLVE* (Terwilliger & Berendzen, 1999) was used to locate anomalous Ba sites, to refine the positions, occupancies and *B* factors of the barium ions and to derive estimates of the protein phases. No attempts were made to fix the occupancies and *B* factors to the same values for each data set.

The phases calculated by *SOLVE* were further improved using *RESOLVE* (Terwilliger, 2000) and automated *RESOLVE* model building was attempted in each case. We applied the SAD standard protocol provided as an example in the *RESOLVE* manual. A complete model was built automatically with *ARP*/*wARP* (Perrakis *et al.*, 1999) using the last data set for the Ba-cocrystallized crystal, *i.e.* collected at 0.98 Å wavelength with the highest (∼1.2 Å) resolution limits, and subjected to several rounds of refinement with *SHELXL* (Sheldrick & Schneider, 1997) and cycles of manual rebuilding with *Coot* (Emsley & Cowtan, 2004).

Table 1
Data-collection statistics.

Values in parentheses are for the last resolution shell.

(*a*) Crystal grown in 0.2 *M* $BaCl_2$.

Data set	1	2	3	4	5
X-ray energy (keV)	7.29	8.00	9.50	11.00	12.66
Wavelength (Å)	1.70	1.55	1.31	1.13	0.98
Space group	*I*222	*I*222	*I*222	*I*222	*I*222
Unit-cell parameters					
a (Å)	66.60	66.62	66.65	66.69	66.72
b (Å)	67.81	67.83	67.86	67.89	67.92
c (Å)	79.56	79.58	79.61	79.65	79.69
Mosaicity range (°)	0.19–0.28	0.18–0.28	0.19–0.29	0.21–0.31	0.22–0.32
Resolution limits (Å)	20–2.10 (2.17–2.10)	20–1.90 (1.97–1.90)	20–1.60 (1.66–1.60)	20–1.40 (1.45–1.40)	20–1.20 (1.24–1.20)
Observations (total/unique)	138940/10820	186962/14517	320589/24124	500642/35874	793685/56817
Average redundancy	12.8	12.9	13.3	14.0	14.0
Average $I/\sigma(I)$	88.4 (60.4)	94.5 (59.4)	94.8 (49.3)	90.1 (34.5)	74.8 (14.1)
Completeness ($I > 0$) (%)	100 (100)	100 (99.8)	99.9 (99.9)	99.8 (99.4)	99.3 (97.8)
R_{merge} (%)	4.9 (7.1)	4.7 (6.9)	4.4 (7.6)	4.1 (9.4)	4.2 (15.8)

(*b*) Crystal soaked in 0.5 *M* $BaCl_2$.

Data set	1	2	3	4	5
X-ray energy (keV)	7.29	8.00	9.50	11.00	12.66
Wavelength (Å)	1.70	1.55	1.31	1.13	0.98
Space group	*I*222	*I*222	*I*222	*I*222	*I*222
Unit-cell parameters					
a (Å)	66.56	66.56	66.57	66.59	66.62
b (Å)	67.82	67.83	67.85	67.88	67.91
c (Å)	79.46	79.47	79.49	79.52	79.56
Mosaicity range (°)	0.35–0.60	0.35–0.59	0.35–0.57	0.37–0.59	0.37–0.61
Resolution limits (Å)	20–2.10 (2.17–2.10)	20–1.90 (1.97–1.90)	20–1.60 (1.66–1.60)	20–1.40 (1.45–1.40)	20–1.25 (1.29–1.25)
Observations (total/unique)	143384/10825	191642/14504	326849/24100	502323/35780	722560/50152
Average redundancy	13.2	13.2	13.6	14.0	14.4
Average $I/\sigma(I)$	75.2 (33.3)	81.1 (31.1)	84.4 (21.9)	78.2 (14.0)	61.9 (7.9)
Completeness ($I > 0$) (%)	99.8 (99.5)	99.8 (98.9)	99.8 (98.8)	99.3 (97.6)	98.1 (95.7)
R_{merge} (%)	6.5 (12.0)	6.2 (12.1)	5.6 (14.3)	5.1 (19.7)	5.2 (32.8)

3. Results and discussion

Data-collection statistics for all data sets collected from cocrystallized and soaked Ba-derivative crystals at five different wavelengths/energies (1.7 Å/7.3 keV, 1.55 Å/8.0 keV, 1.31 Å/9.5 keV, 1.13 Å/11.0 keV, 0.98 Å/12.7 keV) are summarized in Table 1. High redundancies (12.8–14.4) were obtained with overall completenesses higher than 98%, $I/\sigma(I)$ values ranging from 61.9 to 95.8 and good-quality merging residuals of below 6.5%. Comparison of the cocrystallized with the soaked derivative data sets reveal slight changes in the unit-cell axes, a twofold increase in mosaicity and an ∼1% increase in R_{merge}. Isomorphous differences electron-density map calculated between the first and last data sets did not reveal significant signs of radiation damage during data collection.

Automated location of barium positions using *SOLVE* identified the same two sites from the data sets collected from the barium-soaked crystal and the barium-cocrystallized form. The occupancies refined to different values for each data set and ranged from 0.55 to 0.89 for one site and from 0.23 to 0.39 for the second site. Additionally, up to two further barium sites with lower occupancies (0.13–

0.19) were also located in some data sets. The temperature factors for the barium sites refined to values between 7.6 and 23.7 Å^2. The overall mean figure of merit and Z score ranged from 0.38 to 0.53 and from 5.7 to 22.6, respectively. In one instance (the barium-grown crystal collected at 1.13 Å/11 keV) no sites could be automatically located; in this case the two common sites were used to proceed to automated map interpretation.

The initial phases obtained from each independent SAD experiment were sufficient for automatic tracing of the protein structure using the program *RESOLVE*. Table 2 shows the statistics of the *RESOLVE* density-modification procedure and automated model building. For all data sets a significant fraction (41–88%) of chain tracing was correctly identified. For both cases of barium-derivative preparation, the maximum number of residues built with side chains occurred using the data set collected at 1.31 Å (9.5 keV). In the best case the peptide backbone of 88% of the residues were built, with 83% of the side chain assigned.

The positions of the Ba sites were validated by inspection of the anomalous difference Fourier map (Fig. 1*a*). Five additional peaks (5σ level) were visible in the anomalous map. One of these peaks corresponds to the S atom of methionine residue 169. The other peaks were assigned to different atom types on the basis of their binding environment (Fig. 1*b*). Two of these additional peaks were interpreted as additional weak Ba sites. The remaining two peaks were assigned as Cl sites. The Cl sites were distinguished from the Ba sites based on their differing binding patterns (Fig. 2), as described by Dauter *et al.* (1999).

Each barium binds at the protein surface as shown in Fig. 1(*b*). Barium-ion interactions are mostly made with main-chain carbonyl groups and water molecules as shown in Fig. 2. For the most occupied barium additional interactions involve the side chains of histidine and asparagine. Specifically, this Ba (Ba400, Fig. 2*a*) is coordinated by six nearest neighbors: the carboxyl group of Phe93 (O) at a distance of 2.64 Å, the side chains of His144 (ND1) and Asn92 (OD1) at distances of 3.01 and 2.87 Å and three waters at distances of 2.74, 2.84 and 2.88 Å from the Ba site. It is found that the number of interactions formed by the barium sites correlates with their occupancies, with the two most occupied sites having six-coordinated interactions (Figs. 2*a* and 2*b*) and the lower occupancy sites having four-coordinated interactions (Figs. 2*c*–2*f*). In general, the coordination of Ba sites in this study is similar to the coordination of Cs and Gd sites obtained using the quick cryosoaking derivatization procedure (Nagem *et al.*, 2001).

Table 2
Phasing statistics.

(*a*) Crystal grown in 0.2 *M* $BaCl_2$.

Data set	1	2	3	4	5
Wavelength (Å)	1.70	1.55	1.31	1.13	0.98
Anomalous f''	9.90	8.54	6.45	5.05	3.98
Expected amount of anomalous signal (Ba = 2)	0.077	0.066	0.050	0.039	0.031
Observed amount of anomalous signal	0.074	0.066	0.055	0.048	0.041
SOLVE: FOM $\langle m \rangle$	0.50	0.51	0.53	—	0.42
SOLVE: Z score	5.8	7.0	22.6	—	6.5
RESOLVE: overall FOM	0.80	0.81	0.83	0.80†	0.74
Residues built with side chains (%)	11	31	83	36†	24
Total residues built (%)	75	79	88	70†	54

(*b*) Crystal soaked in 0.5 *M* $BaCl_2$,

Data set	1	2	3	4	5
Wavelength (Å)	1.70	1.55	1.31	1.13	0.98
Anomalous f''	9.90	8.54	6.45	5.05	3.98
Expected amount of anomalous signal (Ba = 2)	0.077	0.066	0.050	0.039	0.031
Observed amount of anomalous signal	0.073	0.065	0.053	0.046	0.044
SOLVE: FOM $\langle m \rangle$	0.50	0.50	0.50	0.46	0.38
SOLVE: Z score	6.5	5.9	6.8	6.8	6.8
RESOLVE: overall FOM	0.79	0.79	0.79	0.72	0.67
Residues built with side chains (%)	47	47	60	49	13
Total residues built (%)	88	71	79	70	41

† Results from *SOLVE* run in 'addsolve' mode using the top two common sites followed by *RESOLVE* build.

The xylanase protein crystal used in this study represents a straightforward case with well diffracting crystals. Two Ba sites provide sufficient anomalous signal for automatic model building. We found six Ba-binding sites per 190 amino acids. Interactions are formed with main-chain carbonyl O atoms and water molecules. These common interactions suggest that other proteins may have appropriate sites for Ba binding. It may be of interest to collect the anomalous signal at different wavelengths for cross-comparison. In this study, the best results were obtained using data sets collected at an energy of 9.5 keV. In difficult cases, these multiple-wavelength data sets could potentially be analysed together in order to improve redundancies and statistics and provide a better estimate of the phases.

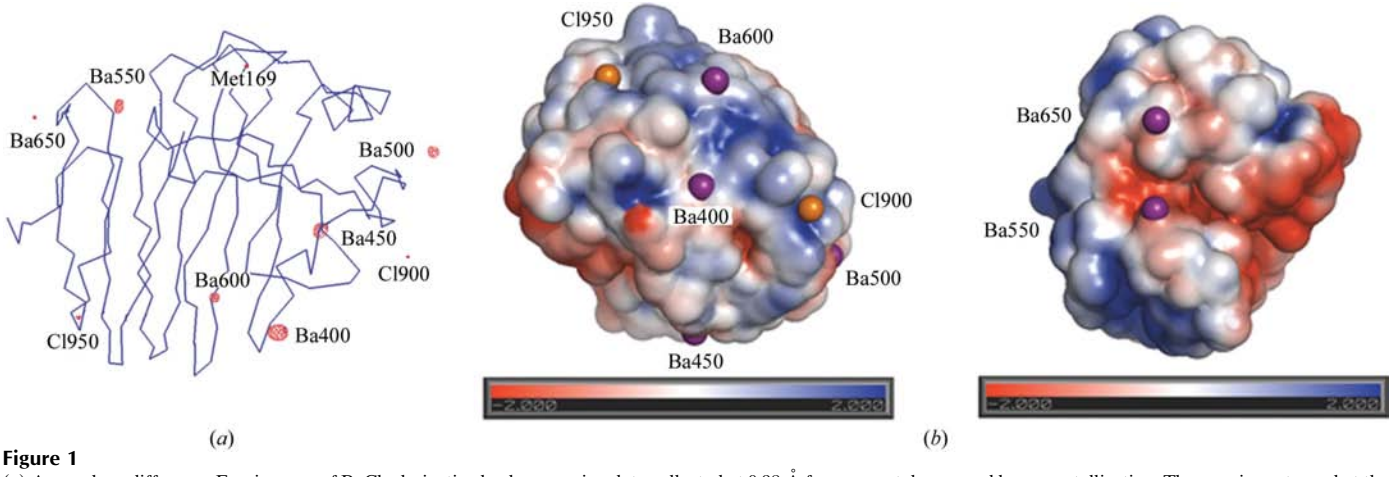

Figure 1
(*a*) Anomalous difference Fourier map of $BaCl_2$-derivatized xylanase using data collected at 0.98 Å from a crystal prepared by cocrystallization. The map is contoured at the 5σ level and was visualized in *PyMOL* (DeLano, 2002). (*b*) Distribution of Ba^{2+} and Cl^- ions around the xylanase electrostatic surface shown for two opposite faces of protein at the scale ± 2 kT/e. The color map goes from red (negative electrostatic potential) to blue (positive electrostatic potential). The electrostatic surface potential was calculated using *APBS* (Baker *et al.*, 2001) and was visualized in *PyMOL* (DeLano, 2002).

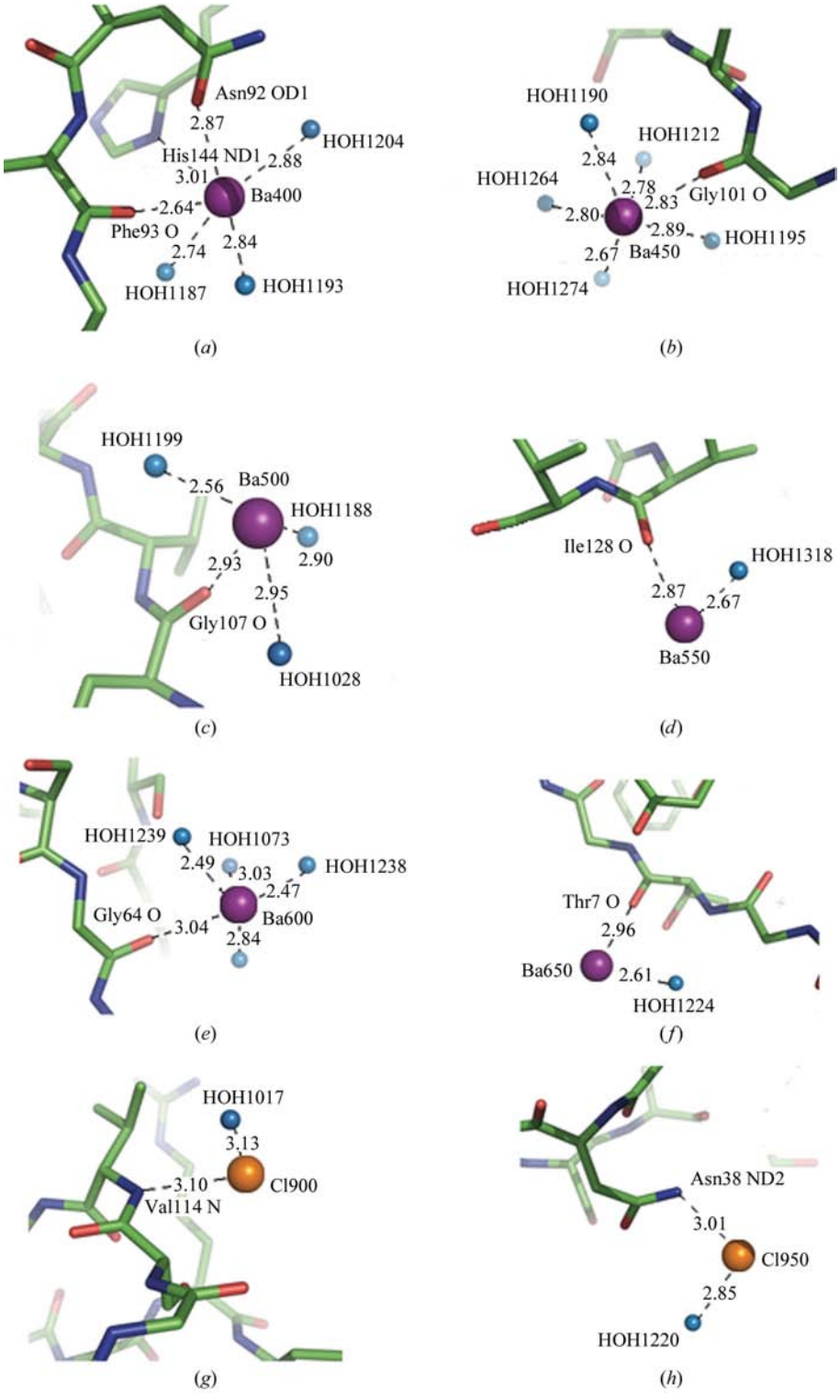

Figure 2
Binding of Ba^{2+} and Cl^{-} ions to the xylanase. Ba^{2+} ions are shown as purple spheres, Cl^{-} ions as orange spheres and water molecules as sky-blue spheres. The strong anomalous Ba^{2+} sites (*a*, *b*) are six-coordinated by neighboring residues and water molecules, whereas weaker sites (*c*–*f*) are four- and two-coordinated. Cl^{-} sites (*g*, *h*) are two-coordinated. This figure was prepared using *PyMOL* (DeLano, 2002).

This work was supported by the Department of Energy under contract No. DE-AC02-98CH10886, the National Institutes of Health/National Institute of General Medical Sciences under agreement GM-0080 and the Brookhaven National Laboratory Laboratory-Directed Research and Development program.

References

Allaire, M. *et al.* (2003). *Synchrotron Radiat. News*, **16**, 20–25.
Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.
Boggon, T. J. & Shapiro, L. (2000). *Structure Fold. Des.* **8**, R143–R149.
Cohen, A., Ellis, P., Kresge, N. & Soltis, S. M. (2001). *Acta Cryst.* D**57**, 233–238.
Cohen, S. L., Padovan, J. C. & Chait, B. T. (2000). *Anal. Chem.* **72**, 574–579.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cromer, D. T. (1983). *J. Appl. Cryst.* **16**, 437.
Dauter, Z., Dauter, M., de La Fortelle, E., Bricogne, G. & Sheldrick, G. (1999). *J. Mol. Biol.* **289**, 83–92.
Dauter, Z., Dauter, M. & Rajashankar, K. R. (2000). *Acta Cryst.* D**56**, 232–237.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* DeLano Scientific, San Carlos, CA, USA. http://www.pymol.org.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Hendrickson, W. A. & Ogata, C. M. (1997). *Methods Enzymol.* **276**, 494–523.
Hendrickson, W. A. & Teeter, M. M. (1981). *Nature (London)*, **290**, 107–113.
Moiseeva, N. & Allaire, M. (2004). *Acta Cryst.* D**60**, 1275–1277.
Nagem, R. A. P., Dauter, Z. & Polikarpov, I. (2001). *Acta Cryst.* D**57**, 996–1002.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Sharpe, A. G. (1992). *Inorganic Chemistry.* New York: John Wiley.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
Schiltz, M., Prangé, T. & Fourme, R. (1994). *J. Appl. Cryst.* **27**, 950–960.
Sun, P. D. & Radaev, S. (2002). *Acta Cryst.* D**58**, 1099–1103.
Sun, P. D., Radaev, S. & Kattah, M. (2002). *Acta Cryst.* D**58**, 1092–1098.
Terwilliger, T. C. (2000). *Acta Cryst.* D**56**, 965–972.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Törrönen, A. & Rouvinen, J. (1995). *Biochemistry*, **34**, 847–856.
Vitali, J., Robbins, A. H., Almo, S. C. & Tilton, R. F. (1991). *J. Appl. Cryst.* **24**, 931–935.
Wang, B.-C. (1985). *Methods Enzymol.* **115**, 90–112.
Zwart, P. H., Grosse-Kunstleve, R. W. & Adams, P. D. (2005). *CCP4 Newsl.* **42**, contribution 10.