Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

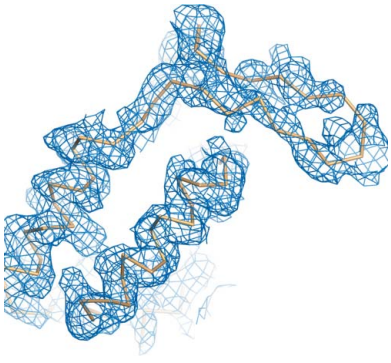
Structure determination by multiple-wavelength anomalous dispersion (MAD) at the Pr L_{III} edge

Sandra Puehringer,[a,b] Michael Hellmig,[a] Sunbin Liu,[b] Manfred S. Weiss,[a] Markus C. Wahl[b] and Uwe Mueller[a]*

[a]Institute F-I2, Macromolecular Crystallography, Helmholtz-Zentrum Berlin für Materialien und Energie, Albert-Einstein-Strasse 15, 12489 Berlin, Germany, and [b]Fachbereich Biologie, Chemie, Pharmazie, Institut für Chemie und Biochemie, AG Strukturbiochemie, Freie Universität Berlin, Takustrasse 6, 14195 Berlin, Germany

Correspondence e-mail: uwe.mueller@helmholtz-berlin.de

Received 5 April 2012
Accepted 5 June 2012



The use of longer X-ray wavelengths in macromolecular crystallography has grown significantly over the past few years. The main reason for this increased use of longer wavelengths has been to utilize the anomalous signal from sulfur, providing a means for the experimental phasing of native proteins. Here, another possible application of longer X-ray wavelengths is presented: MAD at the L_{III} edges of various lanthanide compounds. A first experiment at the L_{III} edge of Pr was conducted on HZB MX beamline BL14.2 and resulted in the successful structure determination of the C-terminal domain of a spliceosomal protein. This experiment demonstrates that L_{III} edges of lanthanides constitute potentially attractive targets for long-wavelength MAD experiments.

1. Introduction

Although the majority of new macromolecular structures deposited in the Protein Data Bank (Berman *et al.*, 2000) nowadays are solved by molecular replacement (Long *et al.*, 2008), experimental phasing by single-wavelength anomalous dispersion (SAD; Hendrickson & Teeter, 1981; Wang, 1985) and multiple-wavelength anomalous dispersion (MAD; Hendrickson, 1991) are still today's methods of choice for determining the structure of a macromolecule for which no homologous model exists. After its inception in the early 1990s, the method of substituting methionine residues by selenomethionine (SeMet; Hendrickson *et al.*, 1990) has now fully developed into the standard method of producing heavy-atom derivatives for MAD phasing.

There are cases, however, when this procedure is not applicable, either because the expression of an SeMet derivative is unsuccessful or because the material cannot be produced recombinantly. In such instances, other derivatization procedures have to be employed: either cocrystallization or soaking of the protein crystal with heavy-atom-containing compounds (Rupp, 2009).

While the classical isomorphous replacement approach is centred on the heavy atoms Hg, Au, Pt *etc.*, lanthanide salts are also occasionally used in the MAD (Weis *et al.*, 1991; Shapiro *et al.*, 1995; Hendrickson, 1999), SAD or SIRAS/MIRAS methods (Nagem *et al.*, 2001; Vennila & Velmurugan, 2011). Originally, lanthanide ions were proposed to substitute calcium in calcium-binding sites (Hendrickson, 1999), but their use is certainly not limited to this. More recently, newly developed lanthanide complexes such as $[Eu(DPA)_3]^{3-}$ (Talon *et al.*, 2011) or $[Ln(DPA)_3]^{3-}$ (Pompidor *et al.*, 2010) have attracted renewed attention, which may be attributed to the very large anomalous signal that can be obtained at the white line of many lanthanide L_{III} absorption edges.

As can be seen from the values presented in Table 1, the L_{III} edges of the lanthanides are mostly located at energies which are accessible at many synchrotron beamlines around the world (Djinović Carugo *et al.*, 2005) but which have not been greatly used owing to the experimental difficulties that such rather long-wavelength experiments entail. Nevertheless, mostly owing to the increasing effort directed into the S-SAD method, such wavelengths are now more frequently used.

Lanthanide derivatization can be achieved using either the quick-soaking protocol (Dauter *et al.*, 2000) at high concentration, by conventional soaking or by cocrystallization. As standard

components in several commercially available additive and heavy-atom screens, the elements of the lanthanide series are coming to the fore for the improvement of crystal growth or diffraction quality. Such lanthanide–protein complexes can be used directly for phasing without further manipulation of the crystal. It has been shown that the development of better X-ray facilities in combination with more powerful software often makes it possible to phase the structure even after the collection of data at only one wavelength using the SAD approach when anomalous scatterers are present in the protein (Dauter *et al.*, 2002). As lanthanide cations show very large anomalous scattering at their L_{III} edges, in most cases the collection of a single data set might be sufficient for use of the SAD approach for phasing. However, in cases where only few anomalous scatterers are bound and where only a low-resolution data set can be obtained, one may have to resort to a full MAD analysis.

This paper describes a lanthanide anomalous phasing experiment which was performed on BL14.2 of the Joint Berlin MX Laboratory at BESSY II of the Helmholtz-Zentrum Berlin für Materialien und Energie (Mueller *et al.*, 2012). We have used the MAD approach to determine the structure of a 130-residue protein which is present as three copies in the asymmetric unit using bound praseodymium ions.

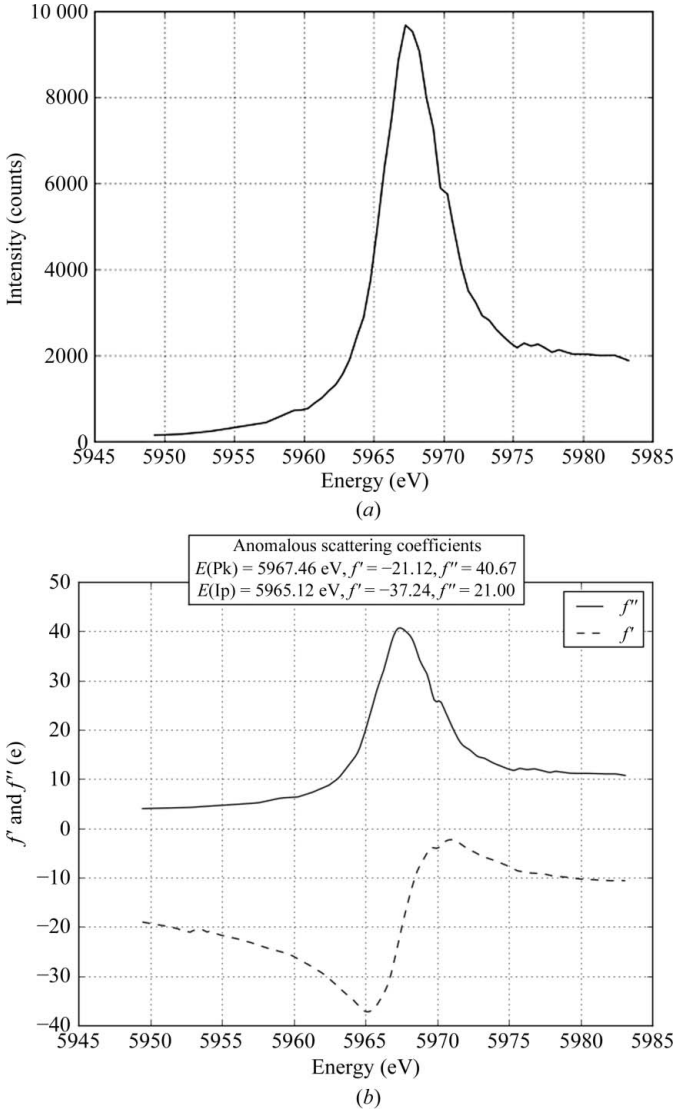


Figure 1
(*a*) Experimental X-ray fluorescence scan across the Pr L_{III} edge. (*b*) f' and f'' scattering factors calculated from the experimental edge-scan data using *CHOOCH* (Evans & Pettifer, 2001).

Table 1
Theoretical values of the L_{III} absorption edges of the lanthanide ions according to Cromer & Liberman (1981).

Element	L_{III} edge (keV)	L_{III} edge (Å)
Lanthanum	5.484	2.2638
Cerium	5.723	2.1689
Praseodymium	5.964	2.0812
Neodymium	6.208	1.9994
Promethium	6.459	1.9217
Samarium	6.716	1.8482
Europium	6.977	1.7791
Gadolinium	7.243	1.7137
Terbium	7.514	1.6519
Dysprosium	7.790	1.5934
Holmium	8.071	1.5379
Erbium	8.358	1.4851
Thulium	8.648	1.4353
Ytterbium	8.944	1.3878
Lutetium	9.244	1.3428

Table 2
Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

	Peak	Inflection point	High-energy remote
X-ray energy (keV)	5.9675	5.9651	13.5
Wavelength (Å)	2.077	2.078	0.918
Oscillation range (°)	1	1	1
No. of images	165	135	120
Space group	$C222_1$	$C222_1$	$C222_1$
Resolution range (Å)	50–3.95 (4.05–3.95)	50–3.95 (4.05–3.95)	50–3.42 (3.51–3.42)
Unit-cell parameters (Å)	a = 87.65, b = 161.12, c = 105.24	a = 87.66, b = 161.42, c = 105.41	a = 87.71, b = 161.29, c = 105.38
Mosaicity (°)	0.18	0.18	0.21
Observed reflections	44133 (3253)	36726 (2731)	50600 (3757)
Unique reflections	12480 (922)	12470 (924)	19306 (1437)
Completeness (%)	98.6 (98.2)	98.2 (97.6)	98.6 (99.2)
Multiplicity	3.5 (3.5)	2.9 (3.0)	2.6 (2.6)
$\langle I/\sigma(I)\rangle$	9.2 (2.6)	8.8 (2.1)	12.4 (2.0)
R_{merge} (%)	11.8 (56.8)	10.8 (62.5)	7.0 (56.7)
$R_{r.i.m.}/R_{meas}$ (%)	14.0 (67.0)	13.3 (76.7)	8.8 (70.9)
$R_{p.i.m.}$ (%)	7.4 (25.8)	6.9 (32.8)	4.1 (30.6)
CC_{anom}, ∞–8 Å (%)	95.0	89.5	55.9
CC_{anom} at 4.4 Å (%)	31.9	20.5	4.8

2. Materials and methods

2.1. Crystallization

Crystallization of the spliceosomal protein was performed using the sitting-drop vapour-diffusion method at 291 K. Initial crystal hits were obtained using a 1:1 mixture of 15 mg ml^{-1} protein solution and reservoir consisting of 0.1 *M* sodium acetate, 0.1 *M* HEPES pH 7.5, 12%(*w*/*v*) PEG 4000 (from The PEGs II Suite, Qiagen). After optimization by varying the pH and the concentration of PEG 4000, crystals were produced in a droplet from conditions consisting of 0.1 *M* sodium acetate, 0.1 *M* HEPES pH 7.4–7.8, 14–16%(*w*/*v*) PEG 4000 within 2–5 d. The Hampton Research Additive Screen was used for further improvement of diffraction power. Well diffracting crystals were obtained by cocrystallization of the protein with 10 m*M* praseodymium(III) acetate hydrate. Prior to data collection, the crystals were cryoprotected in reservoir solution containing 26%(*v*/*v*) ethylene glycol and subsequently flash-cooled in liquid nitrogen. No heavy atoms were present in the cryosolution.

2.2. Data collection and processing

All diffraction experiments were carried out on BL14.2 at BESSY II. In order to determine the peak and inflection-point wavelengths, an X-ray absorption-edge scan was recorded using an XFLASH X-ray fluorescence detector (Bruker AXS, Berlin). The beamline-control software of BL14.2 (Mueller *et al.*, 2012) was modified in order to allow the collection of an edge scan at the theoretical Pr L_{III} edge (Fig. 1*a*) using the standard GUI. The program *CHOOCH* (Evans & Pettifer, 2001) was used to calculate the dispersive and anomalous scattering factors f' and f'' from the obtained scan. Subsequently, complete diffraction data sets were collected at three wavelengths for MAD phasing (Table 2). Peak and inflection-point data sets were collected at a crystal-to-detector distance of 140 mm and wavelengths corresponding to the Pr L_{III}-edge f'' maximum (5.967 keV, 2.0777 Å) and f' minimum (5.965 keV, 2.0785 Å), respectively. A third data set was recorded at a high-energy remote wavelength of 13.5 keV (0.9184 Å). The data were collected using an oscillation increment of 1°. An additional data set was subsequently collected to 2.7 Å resolution. This data set was used later during refinement and model building (unpublished data). All data were processed using *XDS* (Kabsch, 2010). $R_{p.i.m.}$ values were calculated using the program *RMERGE* (Weiss, 2001), which is available from MSW upon request.

2.3. Structure solution

Substructure solution, phasing and density modification were performed with the programs *SHELXC*/*D*/*E* (Schneider & Sheldrick, 2002; Sheldrick, 2008). *SHELXC* used 44 117 reflections to 3.95 Å resolution, 36 726 reflections to 3.95 Å resolution and 62 175 reflections to 3.2 Å resolution from the peak, inflection-point and high-energy remote data sets, respectively, to calculate F_A values and to create the required input files for heavy-atom search and phasing. 300 trials were preformed using *SHELXD* (Usón & Sheldrick, 1999; Schneider & Sheldrick, 2002) to search for ten Pr sites to a resolution of 4.4 Å. Phasing and density modification were performed with *SHELXE* using a solvent content of 50% and 200 cycles of density modification. Evaluation of the map and manual model building were performed using the program *Coot* (Emsley & Cowtan, 2004; Emsley *et al.*, 2010). All figures were produced with the program *PyMOL* (DeLano, 2004).

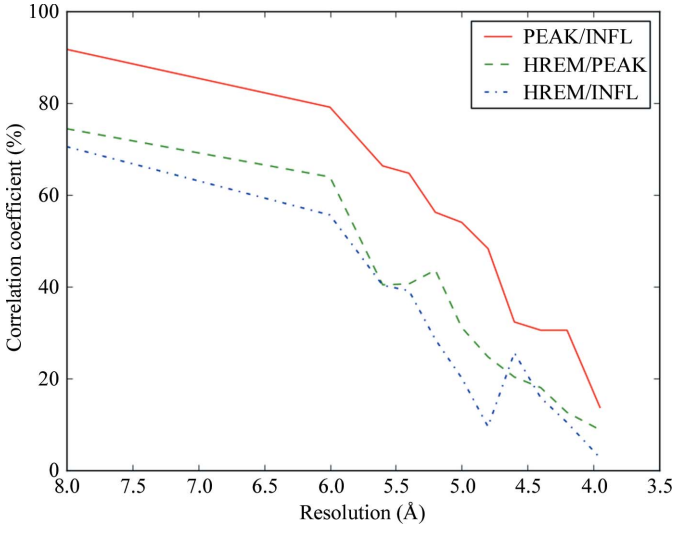
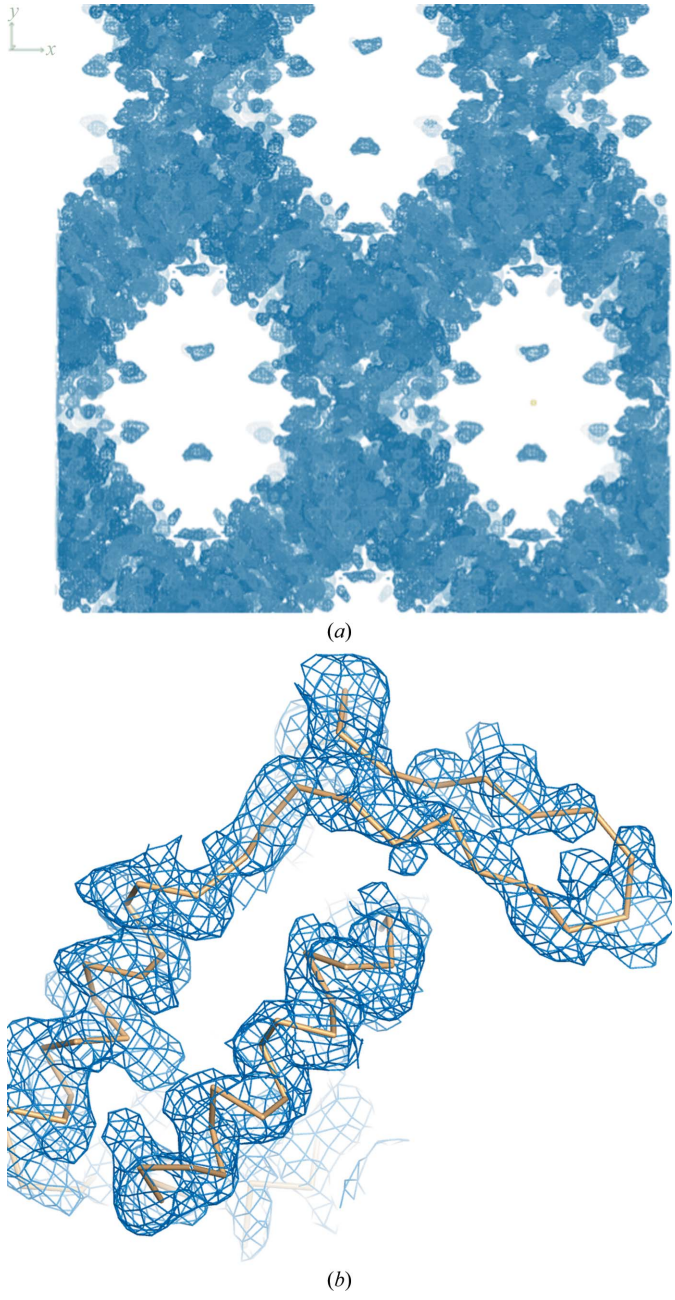


Figure 2
Anomalous correlation coefficients CC_{anom} (Einspahr & Weiss, 2012) between the three data sets.

3. Results and discussion

Initial crystallization experiments with the 130-residue spliceosomal protein produced small and weakly diffracting crystals. After several fine-screening approaches, crystallization trials using an additive screen (Hampton Research) improved the crystal quality substantially. To verify the presence of praseodymium in the crystal, an absorption-edge scan at the L_{III} edge was performed and the anomalous scattering coefficients were calculated (Fig. 1*b*). The peak, inflection-point and high-energy remote data sets exhibited



(*a*)

(*b*)

Figure 3
(*a*) Experimental 3.2 Å resolution MAD electron-density map contoured at 1.5σ. (*b*) Helical region contoured at 1.5σ with manually placed main-chain atoms.

significant anomalous signal up to a maximum resolution of 4.4 Å (Fig. 2). Substructure determination was carried out assuming ten praseodymium sites in the asymmetric unit. *SHELXD* located 14 sites, only three of which showed significant occupancies of 100, 4 and 36%. For phase determination, all 14 sites were used. Phasing and density modification with *SHELXE* led to refined occupancies of the three major sites of 100, 32 and 27%. This corresponds to phasing contributions of about 85, 9 and 6%, respectively, as derived from the estimated length of the heavy-atom vector for individual sites and for all three sites together. The final correlation coefficient of the map was around 85%, with a pseudo-free correlation coefficient of around 69%. The map showed clear electron-density difference between protein and solvent regions. Helical regions could be identified easily and the protein main chain was placed manually (Fig. 3). For comparison, a SAD phasing experiment using the peak data set alone was carried out to verify the necessity of the second- and third-wavelength data sets for successful structure solution. Although the substructure could be determined and *SHELXE* could phase the structure using the SAD strategy, the resulting experimental electron density exhibited a lower map correlation coefficient and was hardly interpretable.

The structure model consisted of three molecules per asymmetric unit, with each molecule binding one Pr cation. The final 2.7 Å resolution model and an anomalous difference electron-density map calculated from the peak Pr data set showed that the Pr ions are bound at the surface of the protein in all three molecules. They are coordinated by two adjacent aspartate side chains and are not involved in any crystal contacts.

Only one of the Pr sites was found to be fully occupied. This confirmed the result from the substructure search in *SHELXD*. In essence, this shows that one fully occupied Pr site with a phasing contribution of 85% was sufficient to phase an ∼400-amino-acid residue peptide structure.

4. Conclusion

Despite the fact that diffraction experiments using long X-ray wavelengths are more difficult than standard macromolecular crystallography experiments (Djinović Carugo *et al.*, 2005), it could be demonstrated that such experiments may be extremely powerful for macromolecular phase determination. In the present experiment, we showed that MAD experiments at the praseodymium $L_{\rm III}$ edge can yield useful phase information even at mediocre resolution and with only a few bound anomalous scatterers. In cases when resolution is limiting, a full MAD analysis may be required for successful structure determination. To our knowledge, the MAD experiment performed at the Joint Berlin MX Laboratory beamline BL14.2 of HZB is the first MAD experiment to be reported at the praseodymium $L_{\rm III}$ edge. Phasing was successful with only one fully occupied site per 390 residues, which illustrates the phasing potential of the lanthanides.

5. Outlook

We are currently attempting to expand the limits of the approach by conducting MAD experiments at $L_{\rm III}$ edges of different lanthanides, such as lanthanum and cerium, at even longer wavelengths. Initial experiments have revealed that with the present beamline stability and goniometer limitations, data collections at wavelengths up to 2.4 Å are feasible. This makes it possible to collect data at the $L_{\rm III}$ edges of all lanthanides, starting at the La edge at around 2.3 Å (Table 1).

We acknowledge access to beamline BL14.2 of the BESSY II storage ring (Berlin, Germany) *via* the Joint Berlin MX-Laboratory, sponsored by the Helmholtz Zentrum Berlin für Materialien und Energie, the Freie Universität Berlin, the Humboldt-Universität zu Berlin, the Max-Delbrück Centrum, and the Leibniz-Institut für Molekulare Pharmakologie. SP was supported by the FWF-fellowship J3173-N17.

References

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
Cromer, D. T. & Liberman, D. A. (1981). *Acta Cryst.* A**37**, 267–268.
Dauter, Z., Dauter, M. & Dodson, E. J. (2002). *Acta Cryst.* D**58**, 494–506.
Dauter, Z., Dauter, M. & Rajashankar, K. R. (2000). *Acta Cryst.* D**56**, 232–237.
DeLano, W. L. (2004). *Abstr. Pap. Am. Chem. Soc.* **228**, 030-CHED.
Djinović Carugo, K., Helliwell, J. R., Stuhrmann, H. & Weiss, M. S. (2005). *J. Synchrotron Rad.* **12**, 410–419.
Einspahr, H. M. & Weiss, M. S. (2012). *International Tables for Crystallography*, Vol. F, 2nd ed., edited by E. Arnold, D. M. Himmel & M. G. Rossmann, pp. 64–74. Chichester: Wiley.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Evans, G. & Pettifer, R. (2001). *J. Appl. Cryst.* **34**, 82–86.
Hendrickson, W. A. (1991). *Science*, **254**, 51–58.
Hendrickson, W. A. (1999). *J. Synchrotron Rad.* **6**, 845–851.
Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.
Hendrickson, W. A. & Teeter, M. M. (1981). *Nature (London)*, **290**, 107–113.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Long, F., Vagin, A. A., Young, P. & Murshudov, G. N. (2008). *Acta Cryst.* D**64**, 125–132.
Mueller, U., Darowski, N., Fuchs, M. R., Förster, R., Hellmig, M., Paithankar, K. S., Pühringer, S., Steffien, M., Zocher, G. & Weiss, M. S. (2012). *J. Synchrotron Rad.* **19**, 442–449.
Nagem, R. A. P., Dauter, Z. & Polikarpov, I. (2001). *Acta Cryst.* D**57**, 996–1002.
Pompidor, G., Maury, O., Vicat, J. & Kahn, R. (2010). *Acta Cryst.* D**66**, 762–769.
Rupp, B. (2009). *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology.* New York: Garland Science.
Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* D**58**, 1772–1779.
Shapiro, L., Fannon, A. M., Kwong, P. D., Thompson, A., Lehmann, M. S., Grübel, G., Legrand, J.-F., Als-Nielsen, J., Colman, D. R. & Hendrickson, W. A. (1995). *Nature (London)*, **374**, 327–337.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Talon, R., Kahn, R., Durá, M. A., Maury, O., Vellieux, F. M. D., Franzetti, B. & Girard, E. (2011). *J. Synchrotron Rad.* **18**, 74–78.
Usón, I. & Sheldrick, G. M. (1999). *Curr. Opin. Struct. Biol.* **9**, 643–648.
Vennila, K. N. & Velmurugan, D. (2011). *Acta Cryst.* F**67**, 1662–1665.
Wang, B.-C. (1985). *Methods Enzymol.* **115**, 90–112.
Weis, W. I., Kahn, R., Fourme, R., Drickamer, K. & Hendrickson, W. A. (1991). *Science*, **254**, 1608–1615.
Weiss, M. S. (2001). *J. Appl. Cryst.* **34**, 130–135.