Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

SAD phasing of a structure based on cocrystallized iodides using an in-house Cu *Kα* X-ray source: effects of data redundancy and completeness on structure solution

Manickam Yogavel, Jasmita Gill, Prakash Chandra Mishra and Amit Sharma*

Structural and Computational Biology Group, International Centre for Genetic Engineering and Biotechnology, New Delhi, India

Correspondence e-mail: asharma@icgeb.res.in

Received 4 May 2007
Accepted 14 June 2007

PDB Reference: superoxide dismutase, 2q2l, r2q2lsf.

Superoxide dismutase (SOD) from *Potentilla atrosanguinea* (Wall. ex. Lehm.) was crystallized using 20% PEG 3350 and 0.2 *M* ammonium iodide and diffraction data were collected to 2.36 Å resolution using an in-house Cu *Kα* X-ray source. Analyses show that data with a redundancy of 3.2 were sufficient to determine the structure by the SAD technique using the iodine anomalous signal. This redundancy is lower than that in previous cases in which protein structures were determined using iodines for phasing and in-house copper X-ray sources. Cocrystallization of proteins with halide salts such as ammonium iodide in combination with copper-anode X-ray radiation can therefore serve as a powerful and easy avenue for structure solution.

1. Introduction

To solve the crystal structure of a new protein it is necessary to have markers in the form of heavy atoms (Blundell & Johnson, 1976) that produce an anomalous diffraction signal suitable for phasing, thus allowing the structure to be determined by multi-wavelength anomalous diffraction (MAD) or single-wavelength anomalous diffraction (SAD) experiments. The heavy-atom anomalous scatterers are incorporated into the crystal by the preparation of selenium-labelled proteins (Hendrickson *et al.*, 1990; Hendrickson, 1997) or of heavy-atom derivatives by soaking the protein crystals in solutions containing salts of Hg, Pt, Au *etc*. (Blundell & Johnson, 1976; Petsko, 1985; Rould, 1997; Boggon & Shapiro, 2000; Islam *et al.*, 1998). As an alternative to heavy-atom soaking, heavy atoms that are naturally present in the protein crystal can also provide a suitable signal in some cases (*e.g.* metalloproteins).

Halides, being small monoatomic ions, are able to substitute for solvent water molecules on the protein surface and also do not show a strong preference for any specific coordination geometry. Iodides have been used as heavy atoms either by soaking the crystals in KI_3 salt (Sigler, 1970) or in *N*-iodosuccinimide (Brzozowski *et al.*, 1991), which iodinates tyrosine. In the recently proposed 'quick cryo-soaking' procedure crystals are soaked for a short duration (15 s to 1 min) in a cryoprotectant solution containing bromide or iodide salts (Dauter *et al.*, 2000) prior to crystal freezing. These approaches require relatively little preparative effort and may be particularly suitable for high-throughput crystallographic and structural genomics projects. An ingenious new method by Xie *et al.* (2004) allows the convenient production of iodine-containing proteins for SAD phasing for an in-house X-ray source. To date, only a small number of structures have been determined by the iodide SAD method using either in-house X-rays or synchrotron data collected to higher resolution at wavelengths ranging from 0.98 to 1.54 Å (Dauter *et al.*, 2000; Nagem *et al.*, 2001, 2005; Usón *et al.*, 2003; Xie *et al.*, 2004; Razeto *et al.*, 2004; Rumpel *et al.*, 2004; Hornberg *et al.*, 2005).

The present work describes X-ray data collected to 2.36 Å resolution using Cu *Kα* radiation that were sufficient to obtain phases using ten iodides for a crystal structure containing two SOD molecules of 15 kDa each in the asymmetric unit. The analysis shows that data that are only 3.2-fold redundant (180°, space group *C*2) were sufficient to yield experimental phases for iodide SAD.

2. Methods

2.1. Crystallization, data collection and processing

SOD protein was crystallized using the hanging-drop vapour-diffusion method at 293 K. The drop contained 1 µl 10 mg ml^{-1} protein solution in buffer (25 m*M* Tris pH 8.0 and 25 m*M* NaCl) and 1 µl crystallization solution (20% PEG 3350 and 0.2 *M* ammonium iodide) and was equilibrated against 200 µl crystallization reservoir solution. A needle-shaped crystal was transferred for a short period (15–30 s) to mother liquor supplemented with 20% ethylene glycol (EG) and then frozen in a nitrogen-gas stream at 100 K. Data were collected using an in-house MicroMax-007 rotating-anode X-ray generator (Rigaku/MSC, 1.5418 Å wavelength) operated at 40 kV and 20 mA with Osmic mirrors (Vari Max HR). The slit size was set to 0.4/0.3. Images were recorded using a MAR345dtb imaging plate. The crystal was mounted in a random orientation and the inverse-beam technique to collect Bijovet pairs on the same frame was not used. Data were collected to 2.36 Å resolution using 1° oscillation steps over a total rotation of 360° with an exposure time of 3 min per frame. The crystal-to-detector distance was 220 mm. Data integration and scaling were carried out using the *HKL*-2000 suite (Otwinowski & Minor, 1997). The SOD crystals belong to the monoclinic space group *C*2 and contain two 15 kDa monomers per asymmetric unit, giving an estimated solvent content of ∼45% (Matthews, 1968). To establish the effects of data completeness, redundancy and resolution on structure determination using iodides, the master data set was split into four different φ ranges (90°, 180°, 270° and 360°; Table 1). The program *HKL2MAP* (Pape & Schneider, 2004) was used to analyse the anomalous signal-to-noise ratio. Substructure solution, phasing, density modification, preliminary model building, iterative model building and refinement were carried out using *AutoSol* and *AutoBuild* in *PHENIX* (Adams *et al.*, 2002).

Table 1
Summary of data-collection statistics, substructure solution, phasing and model building.

Data set	90°	180°	270°	360°
Unit-cell parameters (Å, °)	a = 89.14, b = 64.12, c = 63.62, γ = 90, β = 124.33, γ = 90			
Space group	*C*2			
Resolution (Å)	50–2.36			
R_{merge}	0.044 (0.122)	0.048 (0.130)	0.052 (0.145)	0.054 (0.149)
Completeness (%)	81.6 (30.9)	92.7 (56.3)	93.0 (62.9)	94.4 (63.6)
Mean redundancy	1.8	3.2	4.8	6.3
$I/\sigma(I)$	16.7 (4.4)	23.2 (4.9)	28.7 (5.7)	32.7 (6.0)
Mean $\langle\Delta F\rangle/\langle F\rangle$	0.073	0.069	0.066	0.063
Mean $\Delta F/\sigma(\Delta F)$	1.40	1.64	1.90	2.13
Bijvoet pairs	6228	8922	10442	10787
Lone Bijvoet mates	3889	2011	556	287
AutoSol				
No. of sites	5	10	9	10
Score	—	17.24	53.09	32.42
FOM	—	0.25	0.30	0.33
CC	—	0.30	0.33	0.40
Residues built	—	81	87	127
AutoBuild				
No. of residues built	151	179	200	194
No. of side chains built	—	35	48	91
Map CC	0.44	0.58	0.63	0.63
R_{work}	41	42	37	35
R_{free}	50	47	42	42
Electron-density correlation	—	0.52	0.58	0.62

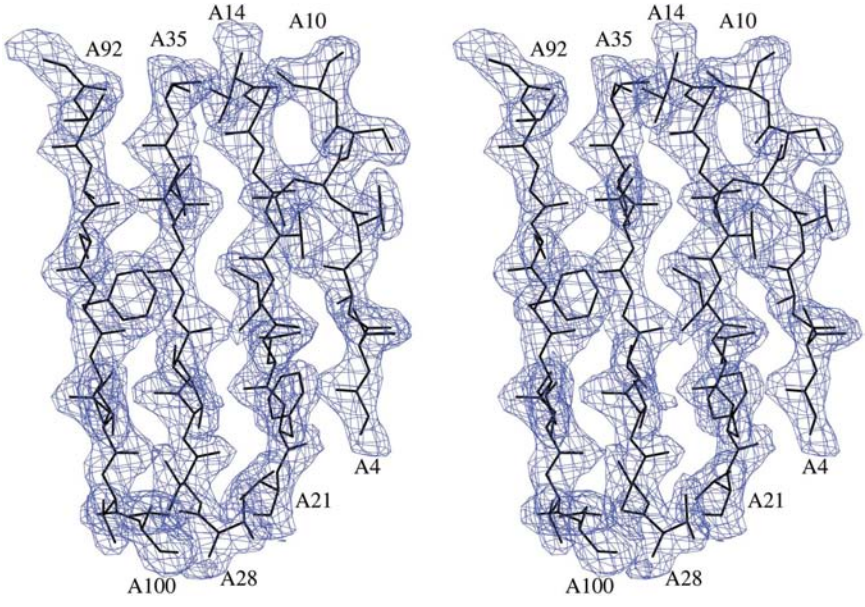

Figure 1
Electron-density map at 2.36 Å resolution after refinement with *REFMAC*. The final $2F_o - F_c$ map is contoured at the 1.5σ level.

3. Results and discussion

3.1. Analysis of the data quality and anomalous signal

Data collected at 100 K to a maximum resolution of 2.36 Å had an overall redundancy of 6.3. Friedel-related reflections were treated as equivalent when scaling the data and computing statistics. The overall $\langle I\rangle/\langle\sigma(I)\rangle$ value for the 50–2.36 Å resolution range is 32.7 and is 6.0 in the highest resolution shell (2.44–2.36 Å). The 360° data set is obviously the best as evidenced by it having the highest value of $I/\sigma(I)$, which is two times higher than that for the 90° data set (16.7). There is no major fluctuation in the R_{merge} value (4.4–5.4%) throughout the φ rotation ranges (Table 1 and Supplementary Table 1[1]). The theoretical anomalous signal ($\langle\Delta F\rangle/\langle F\rangle$) is estimated to be ∼2.5% based on four iodides in the 30 kDa protein in the asymmetric unit (Hendrickson & Ogata, 1997). Analysis of the anomalous signal using *HKL2MAP* (Pape & Schneider, 2004) revealed an anomalous signal-to-noise ratio [$\Delta F/\sigma(\Delta F)$] of >1.29 for all data to 2.36 Å. The Bijovet ratio ($\langle|\Delta F|\rangle/\langle F\rangle$) is shown in Table 1. The complete 360° data set clearly showed significant anomalous signal throughout the entire resolution range to 2.36 Å; $\Delta F/\sigma(\Delta F)$ was 4.4 in the low-resolution shells and 1.29 in the outermost shell (2.6–2.36 Å) (Table 2). From this analysis it was clear that the iodides present in the crystallization liquor were potentially bound to some regions of the protein and were of possible utility in phase determination. A section of the representative fragment of the final model superposed with the final $2F_o - F_c$ map contoured at the 1.5σ level is shown in Fig. 1.

3.2. Substructure determination, phasing and model building

The anomalous signal was analyzed to be significant (*i.e.* the signal-to-noise ratio value was

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: DZ5111). Services for accessing this material are described at the back of the journal.

Table 2
Statistics of the anomalous signal-to-noise ratio $[\Delta F/\sigma(\Delta F)]$ against resolution for different φ rotation ranges.

	Resolution in Å										
Total φ rotation (redundancy)	∞–8	8–6	6–5	5–4	4–3.6	3.6–3.4	3.4–3.2	3.2–3.0	3.0–2.8	2.8–2.6	2.6–2.36
90° (1.8)	2.35	2.06	1.65	1.35	1.19	1.07	1.10	1.06	1.18	1.21	1.13
180° (3.2)	3.11	2.54	2.01	1.62	1.42	1.22	1.25	1.20	1.27	1.24	1.19
270° (4.8)	3.84	3.03	2.43	1.92	1.61	1.43	1.43	1.31	1.37	1.28	1.24
360° (6.3)	4.41	3.45	2.80	2.22	1.82	1.63	1.59	1.42	1.47	1.34	1.29

greater than 1.5) to 3.0 Å (Table 2). The anomalous scatterers were located in *AutoSol* using a 3.0 Å resolution cutoff. More than nine heavy-atom sites were obtained for the various φ rotation ranges with the exception of the 90° data set. The figures of merit (FOMs) and correlation coefficients (CC) were >0.25 and >0.3 for the 180°, 270° and 360° range data sets. Preliminary model building was carried out using *AutoSol* in *PHENIX*. For the 360° data set, 42% of the residues in the asymmetric unit were built using *AutoSol*, while 28% were built for the 180° and 270° data sets. Although the asymmetric unit contains two protomers, noncrystallographic symmetry (NCS) averaging was not performed as we wanted to test whether the iodides were sufficient to phase a structure consisting of 304 residues in the asymmetric unit. The partial models obtained from *AutoSol* in *PHENIX* were fed into *AutoBuild* for iterative model building and refinement. The data set with mean redundancy 3.2 [180° data set, $\langle\Delta F\rangle/\sigma(\Delta F) = 1.64$] was the minimum data that provided suitable starting phases for substructure determination for the combination of iodides and an in-house Cu $K\alpha$ X-ray source. The overall FOM was 0.25 for the 180° data set and the experimental map is clearly interpretable. When compared with previously solved protein structures using iodide phasing (Evans & Bricogne, 2002; Usón *et al.*, 2003), the 180° data set with redundancy of 3.2 shows the use of twofold to fourfold fewer data as being sufficient for phasing and automated model building. For example, in the case of the iodinated trigonal form of T4 lysozyme, suitable initial phases were only found for a data set with a mean redundancy of ∼6.8 and $\langle\Delta F\rangle/\sigma(\Delta F)$ = 2.6 (Xie *et al.*, 2004). Furthermore, in the present case the 180°, 270° and 360° data sets produced a model map CC of ∼0.6 after iterative model building and refinement. For the 360° data set, more than 64% of the residues and 30% of side chains were built automatically using *AutoBuild*. The electron-density correlation between the experimental maps obtained after density modification and automatic model building with the final refined model was 0.52, 0.58 and 0.62 for the 180°, 270° and 360° data sets, respectively. In Fig. 2, a representative section of the final refined model together with the experimental maps after automatic model building are shown.

The atomic model was refined using *REFMAC* (Murshudov *et al.*, 1997) to crystallographic $R_{\rm work}$ and $R_{\rm free}$ values of 16.9% and 23.7%, respectively. The final model consists of 304 residues, two Zn^{2+} ions, ten iodides and 203 water molecules. The occupancy values of the bound iodides were refined using *SHELXL* (Sheldrick & Schneider, 1997) and were also manually adjusted based on temperature factor and

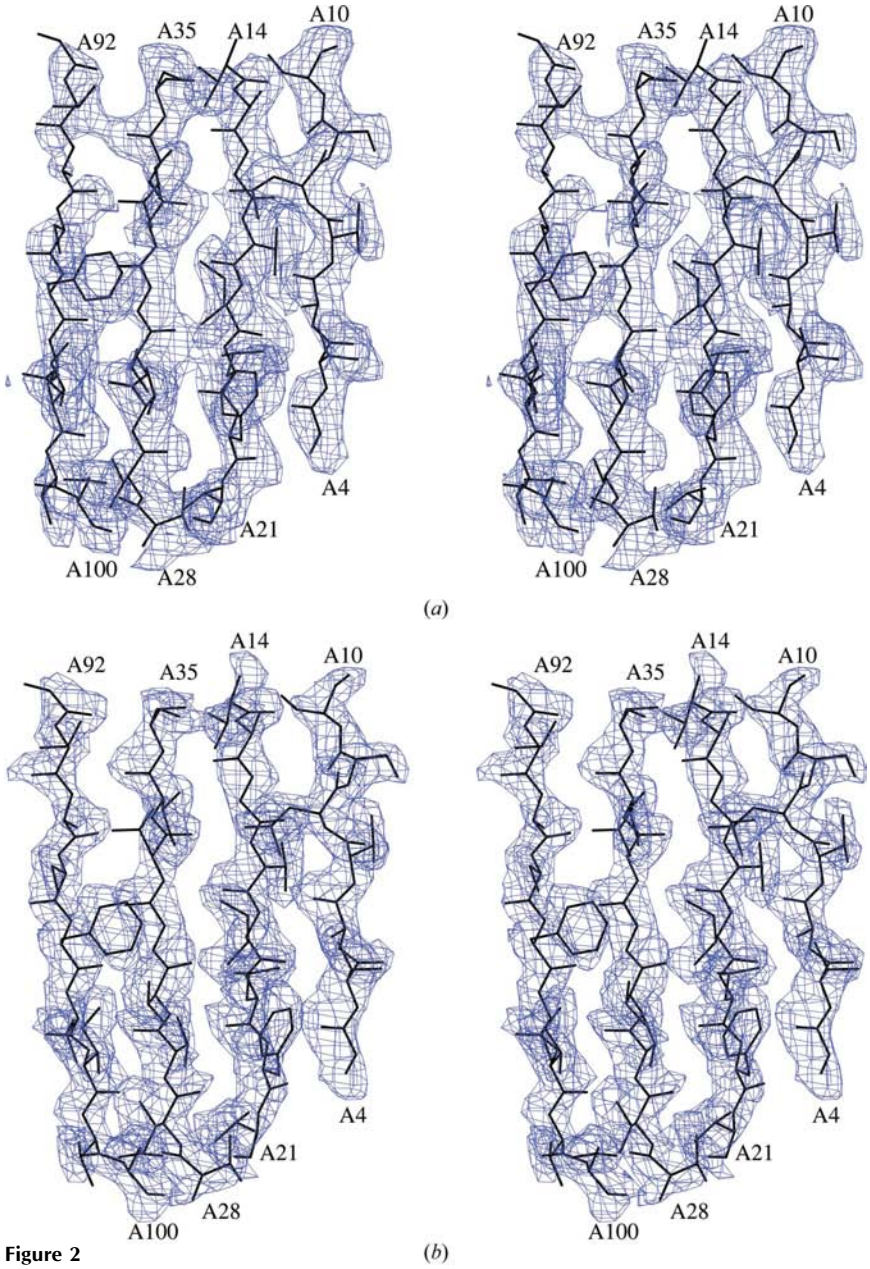

Figure 2
Experimental stereo electron-density maps after automatic model building for the same region as shown in Fig. 1. Maps were obtained from *AutoBuild* in *PHENIX* at 2.36 Å resolution based on phases resulting from the (*a*) 180°, (*b*) 270° and (*c*) 360° oscillation data sets, respectively. The maps are contoured at the 1σ level.

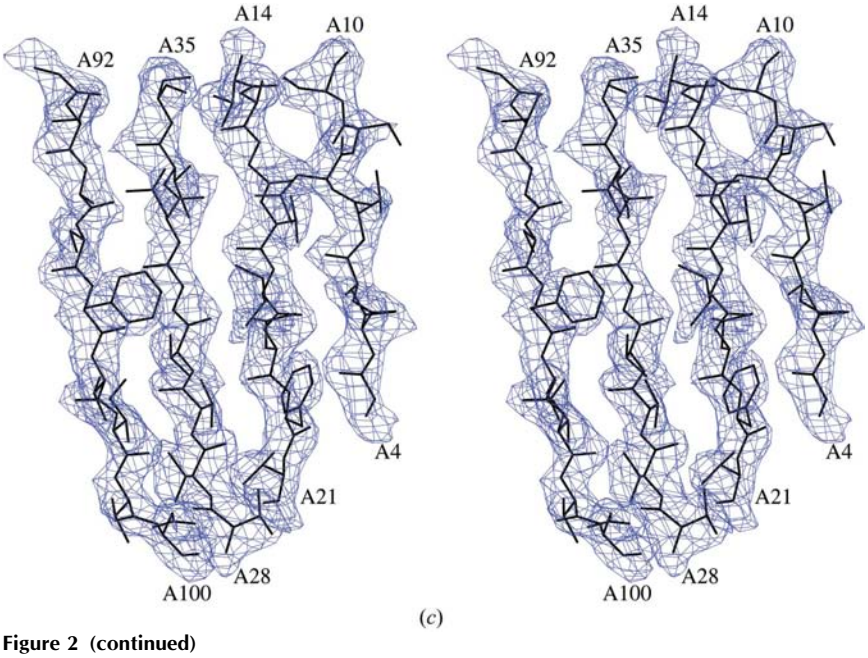

Figure 2 (continued)

electron density. Occupancy values for the iodide sites range from 0.25 to 0.80. The average temperature factors for protein atoms, Zn^{2+} ions, iodide sites and water molecules are 23, 24, 29 and 26 Å^2, respectively. There is no Cu ion in the present SOD structure and this may be a consequence of overexpression in *E. coli*, which may lack sufficient amounts of the copper-loading protein necessary for transferring coppers to SOD.

4. Conclusion

The structure determination of SOD reveals that protein cocrystallized with a lower concentration of ammonium iodide (0.2 *M*) can produce sufficient phasing power to allow automatic structure determination. The 180° data set with a redundancy of 3.2 was sufficient for phasing and for automated structure determination of a 30 kDa protein. In the present study, the halide concentration used was five to ten times lower than those used for bromide or iodide salts in fast-soaking experiments and was very effective for the iodide SAD technique. The benefit of using iodide from crystallization liquor, or in many cases even substituting similar salts by iodide salts in the crystallization liquor, is that the cocrystals do not require any further manipulation (aside from cryocooling) for data collection. Indeed, iodide-based crystallization screens should be promoted as they are likely to reduce the time spent on heavy-atom searches. Iodides bind to protein surfaces in a nonspecific manner without altering protein conformation and offer a considerable anomalous signal at a Cu $K\alpha$ X-ray source. These qualities warrant the inclusion of iodides in crystallization buffers in an effort to speed up the ability to solve the crystal structures of novel protein folds even further.

AS is an International Wellcome Trust Senior Research Fellow in Biomedical Sciences. The X-ray facility at ICGEB is funded by the Wellcome Trust. The authors thank Drs Ahuja, Kumar, Bhardwaj and Dutt of the Institute of Himalayan Bioresource Technology, Palampur, HP 176061, India for help in cloning and purification.

References

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger. T. C. (2002). *Acta Cryst.* D**58**, 1948–954.
Blundell, T. & Johnson, L. N. (1976). *Protein Crystallography*. New York: Academic Press.
Boggon, T. J. & Shapiro, L. (2000). *Structure*, **8**, R143–R149.
Brzozowski, A. M., Derewenda, U., Derewenda, Z. S., Dodson, G. G., Lawson, D. M., Turkenburg, J. P., Bjorkling, F., Huge-Jensen, B., Patkar, S. A. & Thim, L. (1991). *Nature (London)*, **351**, 491–494.
Dauter, Z., Dauter, M. & Rajashankar, K. R. (2000). *Acta Cryst.* D**56**, 232–237.
Evans, G. & Bricogne, G. (2002). *Acta Cryst.* D**58**, 976–991.
Hendrickson, W. A. (1997). *J. Synchrotron Rad.* **6**, 845–851.
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.
Hornberg, A., Hultdin, U. W., Olofsson, A. & Sauer-Eriksson, A. E. (2005). *Biochemistry*, **44**, 9290–9299.
Islam, S. A., Carvin, D., Sternberg, M. J. E. & Blundell, T. L. (1998). *Acta Cryst.* D**54**, 1199–1206.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Nagem, R. A. P., Ambrosio, A. L. B., Rojas, A. L., Navarro, M. V. A. S., Golubev, A. M., Garratt, R. C. & Polikarpov, I. (2005). *Acta Cryst.* D**61**, 1022–1030.
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.
Pape, T. & Schneider, T. R. (2004). *J. Appl. Cryst.* **37**, 843–844.
Petsko, G. A. (1985). *Methods Enzymol.* **114**, 147–156.
Razeto, A., Ramakrishnan, V., Litterst, C. M., Giller, K., Griesinger, C., Carlomagno, T., Lakomek, N., Heimburg, T., Lodrini, M., Pfitzner, E. & Becker, S. (2004). *J. Mol. Biol.* **336**, 319–329.
Rould, M. A. (1997). *Methods Enzymol.* **276**, 461–472.
Rumpel, S., Razeto, A., Pillar, C. M., Vijayan, V., Taylor, A., Giller, K., Gilmore, M. S., Becker, S. & Zweckstetter, M. (2004). *EMBO J.* **23**, 3632–3642.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
Sigler, P. B. (1970). *Biochemistry*, **9**, 3609–3617.
Usón, I., Schmidt, B., von Bulow, R., Grimme, S., von Figura, K., Dauter, M., Rajashankar, K. R., Dauter, Z. & Sheldrick, G. M. (2003). *Acta Cryst.* D**59**, 57–66.
Xie, J., Wang, L., Brock, A., Spraggon, G. & Schultz, P. G. (2004). *Nature Biotechnol.* **22**, 1297–1301.