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

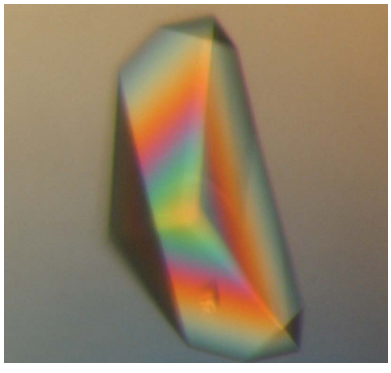
Purification, crystallization and preliminary crystallographic studies of the TLDc domain of oxidation resistance protein 2 from zebrafish

Husam M. A. B. Alsarraf,[a] Fabrice Laroche,[a,b] Herman Spaink,[a,b] Søren Thirup[a]* and Mickael Blaise[a]*

[a]CARB Centre, Department of Molecular Biology, Aarhus University, Gustav Wieds Vej 10c, 8000 Aarhus, Denmark, and [b]Institute of Biology, Leiden University, Leiden, The Netherlands

Correspondence e-mail: sth@mb.au.dk, mick@mb.au.dk

Received 1 June 2011
Accepted 12 July 2011



© 2011 International Union of Crystallography
All rights reserved

Cell metabolic processes are constantly producing reactive oxygen species (ROS), which have deleterious effects by triggering, for example, DNA damage. Numerous enzymes such as catalase, and small compounds such as vitamin C, provide protection against ROS. The TLDc domain of the human oxidation resistance protein has been shown to be able to protect DNA from oxidative stress; however, its mechanism of action is still not understood and no structural information is available on this domain. Structural information on the TLDc domain may therefore help in understanding exactly how it works. Here, the purification, crystallization and preliminary crystallographic studies of the TLDc domain from zebrafish are reported. Crystals belonging to the orthorhombic space group $P2_12_12$ were obtained and diffracted to 0.97 Å resolution. Selenomethionine-substituted protein could also be crystallized; these crystals diffracted to 1.1 Å resolution and the structure could be solved by SAD/MAD methods.

1. Introduction

Reactive oxygen species (ROS) are highly reactive molecules derived from oxygen. ROS are constantly produced in cells owing to metabolic processes. On accumulation of ROS, cells can be damaged, resulting, for example, in DNA mutagenesis or cell death (Georgiou, 2002; Grant *et al.*, 1998, 2000). In order to maintain the integrity of their cells, organisms have developed various strategies for protection against ROS. Many enzymes and compounds are involved in the processing and removal of ROS. Very well studied enzymes such as superoxide dismutase, catalase and glutathione peroxidase and small compounds such as vitamins C and E and glutathione play roles in protection against ROS (Halliwell, 1999; Kirkman & Gaetani, 2007). The discovery of new ROS-processing enzymes is of great importance in order to understand how cells protect themselves from oxidative stress. Recently, Elliott & Volkert (2004) and Durand *et al.* (2007) identified a new gene family in eukaryotes. In these studies, the authors identified two new genes encoding two proteins named oxidation resistance protein 1 (OXR1) and NCOA7 in humans and an OXR1 homologous gene in yeast. Upon stress response, these two genes have been shown to be up-regulated and the proteins have been shown to be involved in ROS protection. More recently, Jaramillo-Gutierrez *et al.* (2010) demonstrated the involvement of OXR1 in preventing oxidation damage in *Anopheles gambiae*. These studies have shown that the C-terminal domains of the OXR proteins seem to be responsible for the ROS protection. The C-terminal domain of the OXR proteins was first described in a bioinformatics study (Doerks *et al.*, 2002) searching for novel protein domains associated with nuclear function. In this study, it was shown that in eukaryotes the C-terminal domain of the OXR proteins is often associated with two other domains, the TBC domain (a GTPAse activator protein) and the LysM domain (a carbohydrate-binding domain), and it was proposed that the OXR protein C-terminal domain may also have catalytic activity. Therefore, the OXR C-terminal domain was renamed TLDc (**T**BC, **L**ysM **D**omain, **c**atalytic).

The exact function of the TLDc domain is unknown, but it has been shown that the TLDc domains of OXR1 and NCOA7 can reverse an oxidative mutator phenotype when expressed in an *Escherichia coli* strain (Elliott & Volkert, 2004; Durand *et al.*, 2007).

All these recent studies clearly indicate the involvement of the *OXR* gene family in ROS response. However, the precise mechanism of how the TLDc domain prevents oxidative damage is unknown and structural studies of this domain may therefore provide some information in this regard. In zebrafish, the *OXR2* gene has recently been identified as a homologue of the human *NCOA7* gene and two mRNA isoforms OXR2A and OXR2B have been identified (Laroche *et al.*, in preparation). Similar to the *NCOA7* gene, the *OXR2* gene encodes a protein possessing several domains: from the N-terminus to the C-terminus, a LysM domain described as a carbohydrate-binding domain, a GRAM domain, a coiled-coil domain and a TLDc domain. Each of these domains is separated by low-complexity regions according to the *SMART* server (Schultz *et al.*, 1998; Letunic *et al.*, 2009). The OXR2 TLDc domain shares 63 and 77% identity with the human NCOA7 and OXR1 TLDc domains, respectively. As part of a study investigating the structure and function of all LysM-containing proteins from zebrafish, we are studying the TLDc domain of OXR2. In this study, we present the purification, crystallization and preliminary crystallographic studies of the TLDc domain from the zebrafish OXR2 protein.

2. Materials and methods

2.1. Gene cloning, protein expression and purification

The zebrafish *OXR2* locus was identified by *BLAST* on the ZV8 zebrafish genome release using the human *NCOA7* sequence as a query. Zebrafish blastula-stage embryos were frozen in liquid nitrogen and crude RNA was obtained using the phenol/chloroform extraction method. RT-PCRs were carried out with SuperScript III High Fidelity (Invitrogen, Carlsbad, California, USA) using specific primer pairs flanking the UTR regions OXR2A (FW, GAGAAGAAGGACGGAAGACG; RV, ACAAGCGCGGTAACAGTAGC) and OXR2B (FW, CCAACATGTTCACCACCAAA; RV, OXR2A RV) (transcripts submitted to NCBI; Laroche *et al.*, in preparation). cDNA products were purified and cloned blunt into the PCR4 vector (Invitrogen, Carlsbad, California, USA). Subsequently, *OXR2* genes were amplified by PCR with Phusion High-Fidelity polymerase (Finnzymes) using primers flanked with Gateway system *att*B sequences and PCR products were recombined into pDONOR221 by BP reaction according to the standard Gateway cloning system procedure (Invitrogen, Carlsbad, California, USA). Subclonings into T7 N-terminal GST vector (pDEST15) and T7 N-terminal poly-His vector (pDEST17) were performed by LR reactions. The 3′ region of the *OXR2B* gene coding for the TLDc domain (amino acids 550–715) was further amplified and cloned using the ligation-independent cloning technique (Ek/LIC) into the pET-41 expression vector from Novagen. The gene was inserted in frame with GST, His and S tags coded by the expression vector. In order to remove this polypeptide tag, a sequence encoding a tobacco etch virus (TEV) protease cleavage site was inserted upstream of the *TLDc* gene. *E. coli* BL21 CodonPlus RIL cells (Agilent Technologies) were transformed with the recombinant plasmid. Protein expression was performed as follows. An overnight culture of 100 ml LB medium supplemented with kanamycin (50 µg ml^{-1}) and chloramphenicol (34 µg ml^{-1}) was used to incubate 4 l LB medium supplemented with the two antibiotics. The cells were then grown at 310 K to an optical density at 600 nm (OD_{600}) of 0.8 and expression was induced with 1 m*M* isopropyl *β*-D-1-thiogalactopyranoside (IPTG). Protein expression was continued overnight at 293 K. The cells were harvested and resuspended in 100 ml buffer *A* consisting of 50 m*M* Tris–HCl pH 8, 400 m*M* NaCl, 5 m*M* *β*-mercaptoethanol, 20 m*M* imidazole, 1 m*M* benzamidine and 1 m*M* phenylmethanesulfonyl fluoride (PMSF). The cells were disrupted by sonication and cell debris was removed by centrifugation at 30 000*g* for 45 min. The crude extract was incubated for 1 h at 277 K with Ni–NTA agarose beads previously equilibrated with buffer *A*. The beads were washed with 50 m*M* Tris–HCl pH 8, 1 *M* NaCl, 5 m*M* *β*-mercaptoethanol, 50 m*M* imidazole, 1 m*M* benzamidine and 1 m*M* PMSF to remove nonspecifically bound proteins. Elution was performed using buffer *E* (50 m*M* Tris–HCl pH 8, 400 m*M* NaCl, 5 m*M* *β*-mercaptoethanol, 500 m*M* imidazole, 1 m*M* benzamidine and 1 m*M* PMSF). The eluted protein was incubated with His-tagged TEV protease (as purified in our laboratory; see Supplementary Material[1]) in a 1:100(*w*:*w*) ratio; the cleavage reaction was performed during dialysis (dialysis-bag cutoff 12–15 kDa) against 1 l dialysis buffer (50 m*M* Tris–HCl pH 8, 100 m*M* NaCl and 5 m*M* *β*-mercaptoethanol) overnight at 277 K. After dialysis the proteins were spun down for 30 min at 30 000*g* and the supernatant was incubated again for 1 h with Ni–NTA agarose beads equilibrated with buffer *A*. The TLDc domain free from the tag was collected in the flowthrough. The TLDc domain was then concentrated to 5 mg ml^{-1} using a Vivaspin column (10 kDa cutoff), loaded onto a size-exclusion column (Superdex 75 10/300 GL, GE Healthcare) and eluted with a buffer consisting of 50 m*M* Tris–HCl pH 8, 100 m*M* NaCl and 5 m*M* *β*-mercaptoethanol. The predicted molecular weight (MW) of the TLDc domain was 18.75 kDa; as the elution peak of the protein on the size-exclusion chromatography was 12.8 ml

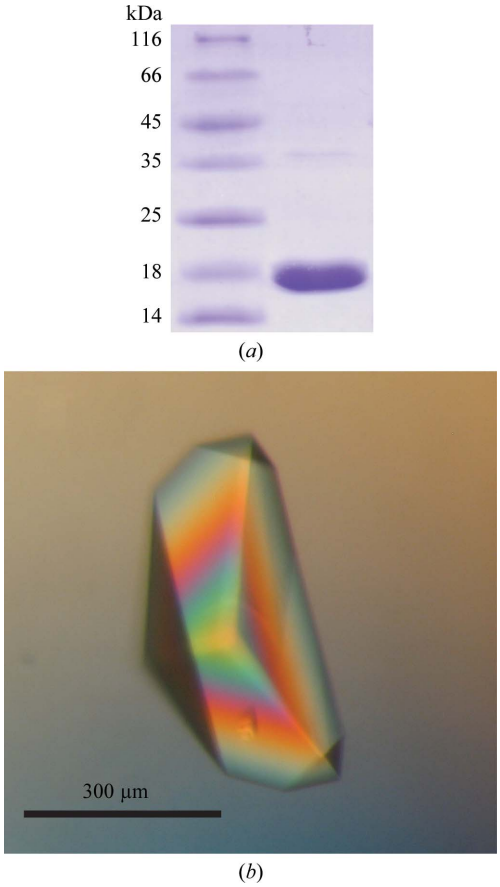


Figure 1
(*a*) SDS–PAGE (15%) of the TLDc domain after the last step of purification. 15 µg protein was loaded onto the gel. (*b*) TLDc-domain crystal.

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: FW5320).

Table 1
Data-collection statistics.

The table reports the data-collection statistics for the native and selenomethionine-substituted crystals. R_{meas} and $R_{mrgd\text{-}F}$, defined according to Diederichs & Karplus (1997), are quality measures of the individual intensity observations and of the reduced structure-factor amplitudes. Values in parentheses are for the last resolution shell. The completeness of the anomalous data and $R_{anomalous}$ were calculated with the *SCALA* program (Evans, 2011).

Data collection	Native	Se peak	Se remote	Se inflection
Beamline	MAX-lab I911-3	SLS X06DA	SLS X06DA	SLS X06DA
Wavelength (Å)	1.0	0.978	0.970	0.980
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit-cell parameters				
a (Å)	69.11	69.34	69.41	69.48
b (Å)	65.60	65.65	65.70	65.80
c (Å)	36.27	36.39	36.42	36.45
Resolution (Å)	50–0.97 (0.98–0.97)	50–1.1 (1.16–1.10)	50–1.1 (1.16–1.10)	50–1.1 (1.16–1.10)
R_{meas} (%)	5.8 (38.1)	6.3 (46.1)	6.6 (56.6)	6.3 (66.0)
$R_{mrgd\text{-}F}$ (%)	4.5 (44.0)	9.1 (52.8)	10.2 (62.3)	10.7 (74.7)
$\langle I/\sigma(I)\rangle$	25.6 (3.9)	19.9 (3.2)	18.2 (2.6)	18.75 (2.2)
Completeness (%)	99.8 (96.9)	98.8 (91.8)	99.2 (94.5)	98.6 (90.3)
Multiplicity	8.94 (3.61)	6.74 (3.39)	6.79 (3.59)	6.68 (3.30)
Anomalous completeness (%)	—	98.2 (87.9)	98.6 (90.8)	97.9 (85.8)
$R_{anomalous}$ (%)	—	6.1 (20.9)	5.8 (24.1)	2.6 (26.6)

we therefore concluded that the protein behaves as a monomer in solution. Following this protocol, 3 mg pure protein (Fig. 1*a*) was obtained from 1 l culture.

2.2. Expression protocol for selenomethionine-derivatized TLDc

Selenomethionine-substituted protein was expressed in the methionine-auxotrophic *E. coli* B834 (DE3) strain (Novagen). A 20 ml overnight culture was used to incubate 2 l LB medium containing kanamycin (50 µg ml^{-1}) and chloramphenicol (34 µg ml^{-1}); the cells were grown at 310 K until the OD_{600} reached 0.8. The cells were then spun down at 12 000*g* and washed twice in M9 medium (20 m*M* $Na_2HPO_4{\cdot}12H_2O$, 22 m*M* KH_2PO_4, 8 m*M* NaCl, 18 m*M* NH_4Cl). The cells were resuspended in M9 medium and used to inoculate 6 × 2 l flasks containing M9 medium supplemented with trace elements, 0.4% glycerol, 1 m*M* $MgSO_4$, 1 m*M* $CaCl_2$, 4 µ*M* biotin, 3 µ*M* thiamine and the two antibiotics. The culture was grown for 1 h at 310 K to remove any trace of methionine and was then supplemented with 50 mg l^{-1} selenomethionine for 3 h, at which point protein expression was induced with 1 m*M* IPTG; the culture was continued overnight at 293 K. The selenomethione-substituted protein was purified using the same protocol as used for the native protein.

2.3. Crystallization

After the last step of purification, the protein was concentrated to 8.7 mg ml^{-1} in 50 m*M* Tris–HCl pH 8, 100 m*M* NaCl and 5 m*M* β-mercaptoethanol. The protein was centrifuged for 30 min at 30 000*g* prior to crystallization experiments. Crystallization screening was performed using the Index and PEGRx 1 and 2 crystallization screens from Hampton Research. The sitting drops were set up at 277 K by mixing 1.5 µl reservoir solution with 1.5 µl TLDc protein solution at 8.7 mg ml^{-1}. After 24 h crystals appeared in condition Nos. 3, 4, 5, 6, 20, 21 and 23 from the Index screen and several weeks later in condition Nos. 27 and 29 from the Index screen and condition No. 39 from the PEGRx 1 screen. The crystals from condition No. 5 of the Index screen, consisting of 100 m*M* Na HEPES pH 7.5 and 2 *M* ammonium sulfate, diffracted to 2 Å resolution on a rotating-anode generator after cryoprotection by a quick soak in a solution consisting of 100 m*M* Na HEPES pH 7.5, 2 *M* ammonium sulfate and 10% glycerol. We therefore optimized the crystallization around this condition. Crystal optimization was performed by varying the ammonium sulfate concentration from 1.4 to 2.4 *M* and by varying the pH value from 7.5 to 8.5. The best diffracting crystals were obtained in 100 m*M* Na HEPES pH 8, 1.8 *M* ammonium sulfate by mixing 1.5 µl reservoir solution with 1.5 µl protein solution at 8.7 mg ml^{-1}. The selenomethionine-substituted protein was crystallized under similar conditions but at a concentration of 7.0 mg ml^{-1}.

2.4. Data collection and processing

Prior to data collection, the crystals were cryoprotected by soaking them overnight at 277 K in a solution consisting of 100 m*M* Na HEPES pH 8, 1.8 *M* ammonium sulfate and 10% glycerol. The crystals were then flash-frozen in liquid nitrogen. Two native data sets were collected on the I911-3 beamline at MAX-lab, Lund, Sweden. An initial high-resolution data set was collected with a total of 250 images using the following settings: an exposure time of 10 s, 1° rotation range per image, a wavelength of 1.0 Å and a crystal-to-detector distance of 60 mm. A second data set at lower resolution was collected from the same crystal covering 100 images using the same settings except that the exposure time was reduced to 2 s per image and the crystal-to-detector distance was increased to 100 mm. The two data sets were integrated, scaled and merged with *XDS* (Kabsch, 2010). A multiwavelength anomalous dispersion data set was measured at three wavelengths from a single crystal of the seleno-

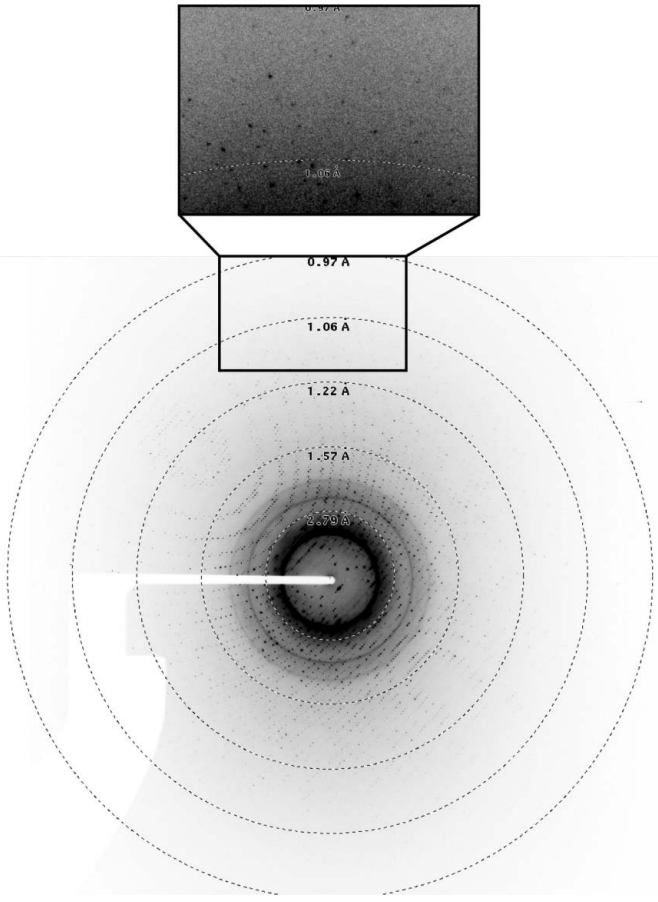


Figure 2
Diffraction image of a native TLDc crystal. The enlargement shows diffraction spots extending to 0.97 Å.

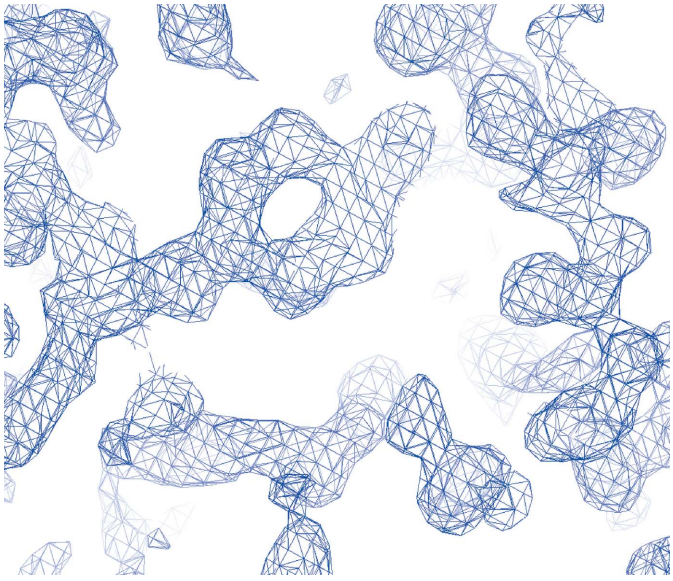


Figure 3
Experimental electron-density map calculated to 1.1 Å resolution and contoured at 1σ. This figure was prepared with *Coot* (Emsley *et al.*, 2010).

methionine-substituted protein on the X06DA beamline of the Swiss Light Source, Villigen, Switzerland. For each data set, 1440 images were collected with a rotation range of 0.25° and a crystal-to-detector distance of 100 mm. After performing a fluorescence scan, data were collected at wavelengths corresponding to the selenium absorption peak, inflection and remote positions. The data were collected in the following order: firstly at the edge, followed by the inflection and finally the remote. Moreover, to avoid radiation damage the crystal was translated between each data collection and the transmission of the beam was set to 10%. These three data sets were also processed with *XDS* (Kabsch, 2010). Data statistics for the native and selenomethionine-derivative data are presented in Table 1.

3. Results and discussion

The TLDc domain of the zebrafish OXR2 protein could be crystallized. After optimization, native and selenomethionine-substituted crystals both reached about 0.5 mm in the largest dimension (Fig. 1*b*). The native TLDc crystals diffracted to 0.97 Å resolution (Fig. 2). They belonged to the orthorhombic space group $P2_12_12$, with unit-cell parameters a = 69.11, b = 65.62, c = 36.27 Å. The Matthews coefficient (V_M) of the crystal was 2.22 Å^3 Da^{-1}, with 44.7% of the unit cell occupied by solvent, assuming one TLDc domain (molecular weight 18.75 kDa) per asymmetric unit.

The TLDc domain does not present any sequence similarity to any structure deposited in the PDB; we therefore produced selenomethionine-substituted protein in order to solve the phase problem. Crystals of the selenomethionine-substituted protein diffracted to 1.1 Å resolution, with anomalous signal extending to 1.1 Å (Table 1). We were able to solve the phase problem by multiple-wavelength anomalous dispersion and single-wavelength anomalous dispersion using the *AutoSol* software from the *PHENIX* package (Adams *et al.*, 2010). Six Se sites were located with a figure of merit of 0.57. An interpretable electron-density map was obtained (Fig. 3). After performing density modification with the program *RESOLVE* (Terwilliger, 2003), we rebuilt and refined the structure. The presence of one molecule per asymmetric unit was confirmed. This model was further used to perform molecular replacement on the native data set using the program *Phaser* from the *PHENIX* package. The structure was rebuilt with *Coot* (Emsley *et al.*, 2010) and refined with *PHENIX* and will be described elsewhere.

We thank the CARB Centre, the Danish National Research Foundation and Danscatt for financial support. We thank J. Stougaard for constant support and fruitful discussions. We would also like to thank the staff at SLS and MAX-lab for support. We also thank N. S. Laursen, R. Kidemose and C. Olesen for help with data collection.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.
Doerks, T., Copley, R. R., Schultz, J., Ponting, C. P. & Bork, P. (2002). *Genome Res.* **12**, 47–56.
Durand, M., Kolpak, A., Farrell, T., Elliott, N. A., Shao, W., Brown, M. & Volkert, M. R. (2007). *BMC Cell Biol.* **8**, 13.
Elliott, N. A. & Volkert, M. R. (2004). *Mol. Cell. Biol.* **24**, 3180–3187.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Evans, P. R. (2011). *Acta Cryst.* D**67**, 282–292.
Georgiou, G. (2002). *Cell*, **111**, 607–610.
Grant, C. M., Luikenhuis, S., Beckhouse, A., Soderbergh, M. & Dawes, I. W. (2000). *Biochim. Biophys. Acta*, **1490**, 33–42.
Grant, C. M., Perrone, G. & Dawes, I. W. (1998). *Biochem. Biophys. Res. Commun.* **253**, 893–898.
Halliwell, B. (1999). *Free Radic. Res.* **31**, 261–272.
Jaramillo-Gutierrez, G., Molina-Cruz, A., Kumar, S. & Barillas-Mury, C. (2010). *PLoS One*, **5**, e11168.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 133–144.
Kirkman, H. N. & Gaetani, G. F. (2007). *Trends Biochem. Sci.* **32**, 44–50.
Letunic, I., Doerks, T. & Bork, P. (2009). *Nucleic Acids Res.* **37**, D229–D232.
Schultz, J., Milpetz, F., Bork, P. & Ponting, C. P. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 5857–5864.
Terwilliger, T. C. (2003). *Methods Enzymol.* **374**, 22–37.