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

Structure of the third catalytic domain of the protein disulfide isomerase ERp46

Irina E. Gulerez, Guennadi Kozlov, Angelika Rosenauer and Kalle Gehring*

Department of Biochemistry, Groupe de Recherche axé sur la Structure des Protéines, McGill University, 3649 Promenade Sir William Osler, Montréal, Québec H3G 0B1, Canada

Correspondence e-mail: kalle.gehring@mcgill.ca

Received 14 December 2011
Accepted 9 February 2012

PDB Reference: third catalytic domain of ERp46, 3uvt.

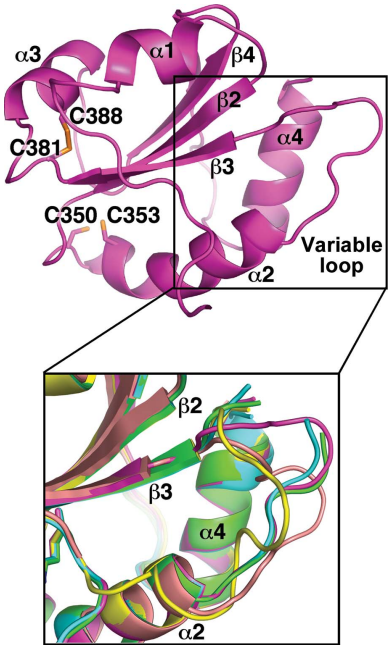
Protein disulfide isomerases are responsible for catalyzing the proper oxidation and isomerization of disulfide bonds of newly synthesized proteins in the endoplasmic reticulum. Here, the crystal structure of the third catalytic domain of protein disulfide isomerase ERp46 (also known as protein disulfide isomerase A5 and TXNDC5) was determined to 2.0 Å resolution. The structure shows a typical thioredoxin-like fold, but also identifies regions of high structural variability. In particular, the loop between helix $\alpha2$ and strand $\beta3$ adopts strikingly different conformations among the five chains of the asymmetric unit. Cys381 and Cys388 form a structural disulfide and its absence in one of the molecules leads to dramatic conformational changes. The tryptophan residue Trp349 of this molecule inserts into the cavity formed by helices $\alpha1$ and $\alpha3$ of a neighbouring molecule, potentially mimicking the interactions of ERp46 with misfolded substrates.

1. Introduction

In order to acquire the proper arrangement of disulfide bonds, newly synthesized membrane proteins and secreted proteins fold in the oxidative environment of the endoplasmic reticulum (ER) with the assistance of ER enzymes called protein disulfide isomerases (PDIs; Hatahet & Ruddock, 2009; Ellgaard & Ruddock, 2005; Maattanen *et al.*, 2006). PDI proteins are defined by their ER localization and the presence of one or more thioredoxin-like domains, which possess oxidoreductase activity mediated by a conserved catalytic C*xx*C motif. During catalysis, the N-terminal cysteine forms a mixed disulfide with the protein substrate, which is then resolved by the C-terminal cysteine to mediate release of the substrate (Walker & Gilbert, 1997). Several other residues have also been implicated in the catalytic activity of PDIs. A conserved glutamate positioned underneath the cysteines has been implicated in proton transfer during substrate release (Dyson *et al.*, 1997) and a conserved arginine modulates the pK_a of the C-terminal cysteine (Lappi *et al.*, 2004; Karala *et al.*, 2010). Most PDI-family members contain additional noncatalytic thioredoxin-like domains that do not contain the C*xx*C motif; some of these domains have been shown to be involved in substrate or co-chaperone recognition, while others may simply function as spacers for positioning catalytic elements (for a review, see Maattanen *et al.*, 2006).

Previous studies have shown that human ERp46 acts as a stress-survival factor (Sullivan *et al.*, 2003) and as a regulator of insulin content in pancreatic cells (Alberti *et al.*, 2009). It also specifically interacts with peroxiredoxin IV, a peroxidase in the ER (Jessop *et al.*, 2009). ERp46 is the only PDI composed of three catalytic domains with no noncatalytic domains. This suggests that the catalytic domains of ERp46 interact directly with substrates and chaperone partners. The NMR structure of the third catalytic domain of ERp46 has previously been determined (PDB entry 2diz; RIKEN Structural Genomics/Proteomics Initiative, unpublished work), but the basis for substrate recognition by ERp46 is still poorly understood.

Here, we present the crystal structure of the third catalytic domain of human ERp46 determined at 2.0 Å resolution. The structure



reveals the crucial role of the Cys381–Cys388 disulfide for the structural stability of the ERp46 domains. Analysis of crystal contacts suggests a role for the tryptophan preceding the C*xx*C motif and the surface between helices $\alpha1$ and $\alpha3$ in substrate binding by ERp46.

2. Materials and methods

2.1. Cloning, protein expression and purification

The 12.2 kDa third domain of human ERp46 (residues 323–428 of Swiss-Prot entry Q8NBS9) was cloned into pGEX-6P-1 (GE Healthcare) and expressed in *Escherichia coli* BL21 (DE3) in rich (LB) medium as an N-terminal GST-tag fusion protein. Bacterial cultures were grown at 310 K with shaking until the absorbance measured at 600 nm reached 0.7. Protein expression was induced with 1 m*M* isopropyl β-D-1-thiogalactopyranoside and the cultures were left shaking overnight at 293 K. After centrifugation, the pellets were resuspended in PBS buffer (137 m*M* NaCl, 2.7 m*M* KCl, 10 m*M* Na_2HPO_4, 2 m*M* KH_2PO_4 pH 7.4) containing 1 m*M* phenylmethylsulfonyl fluoride, 0.1 mg ml^{-1} lysozyme and 0.1%(*v*/*v*) β-mercaptoethanol and lysed using sonication. Cell debris was removed by centrifugation and the fusion protein was bound to glutathione-Sepharose beads and washed with PBS. The GST tag was cleaved with PreScission Protease (GE Healthcare) overnight at 293 K. The resulting construct contained an additional N-terminal Gly-Pro-Leu-Gly-Ser sequence. The protein was additionally purified by size-exclusion chromatography using HPLC buffer (10 m*M* HEPES, 100 m*M* NaCl, 5 m*M* dithiothreitol pH 7.0). The protein eluted at a retention time corresponding to the monomeric form.

2.2. Crystallization, data collection and processing

Initial crystallization conditions were identified utilizing hanging-drop vapour diffusion using the Classics II screen (Qiagen). The best crystals of the third domain of ERp46 were obtained at 295 K by equilibrating a drop consisting of 0.6 µl ERp46 (6 mg ml^{-1}) in HPLC buffer mixed with 0.6 µl reservoir solution consisting of 0.2 *M* ammonium sulfate, 0.1 *M* bis-tris pH 5.5, 25%(*w*/*v*) PEG 3350 against 0.6 ml reservoir solution. The solution for cryoprotection consisted of the reservoir solution with the addition of 20%(*v*/*v*) glycerol. For data collection, crystals were picked up in a nylon loop and flash-cooled in an N_2 cold stream (Oxford Cryosystem). A native data set was collected using a single-wavelength (0.9770 Å) regime with an ADSC Quantum-210 CCD detector (Area Detector Systems Corp.) on beamline A1 at the Cornell High-Energy Synchrotron Source (CHESS; Table 1). Data processing and scaling were performed with *HKL*-2000 (Otwinowski & Minor, 1997).

2.3. Structure determination and refinement

The starting phases were obtained using molecular replacement with the ERp57 *a*-domain structure (PDB entry 3f8u; Dong *et al.*, 2009) as a search model in *Phaser* (McCoy *et al.*, 2007). The resulting model was extended manually with the help of the program *Coot* (Emsley & Cowtan, 2004) and was improved by several cycles of refinement using *REFMAC* (Murshudov *et al.*, 2011) followed by translation–libration–screw (TLS) refinement (Winn *et al.*, 2003). The final model has good stereochemistry, with no outliers in the Ramachandran plot computed using *PROCHECK* (Laskowski *et al.*, 1993). Figures were produced with *PyMOL* (http://www.pymol.org). The coordinates and structure factors have been deposited in the RCSB Protein Data Bank (PDB entry 3uvt).

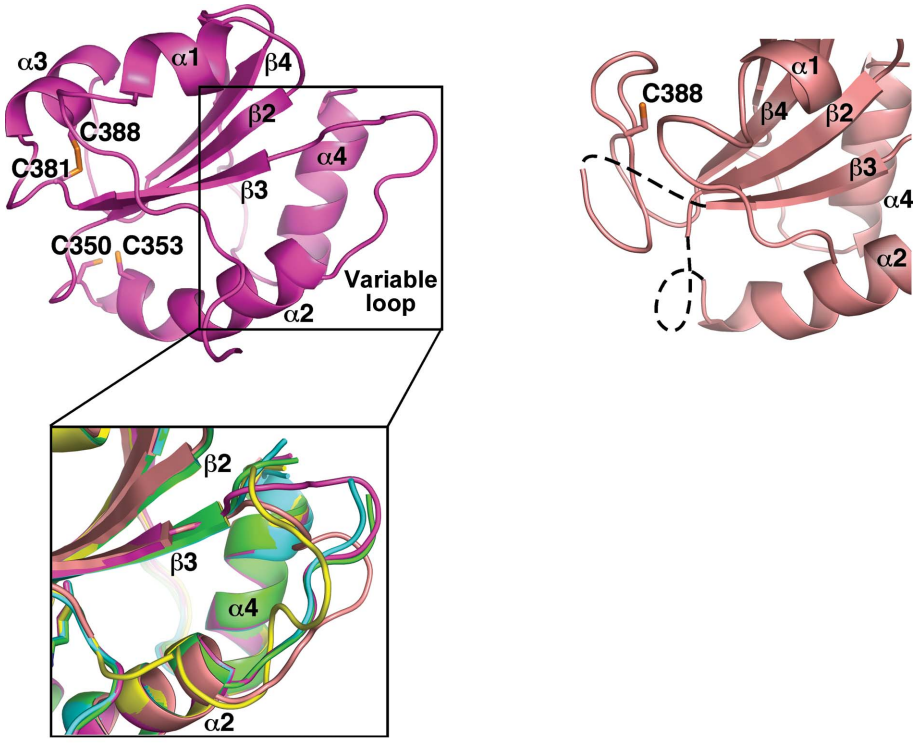


Figure 1
(*a*) The structure of the third catalytic domain of ERp46 shows a thioredoxin-like fold that contains the Cys381–Cys388 disulfide bridge. The insert displays an overlay of chains *A* (green), *B* (cyan), *C* (magenta), *D* (yellow) and *E* (coral) that reveals conformational heterogeneity in the $\alpha2$–$\beta3$ variable loop. (*b*) Chain *E* (top) shows large conformational changes resulting from loss of the Cys381–Cys388 structural disulfide. Helix $\alpha3$ is unfolded and the catalytic site cysteines Cys350 and Cys353 are disordered.

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Data collection	
X-ray wavelength (Å)	0.9770
Space group	$P2_1$
Unit-cell parameters (Å, °)	a = 68.11, b = 61.89, c = 71.00, β = 106.37
Resolution range (Å)	50.0–2.00 (2.03–2.00)
Unique reflections	36285
Completeness (%)	100.00 (100.00)
Multiplicity	4.2 (4.1)
$\langle I/\sigma(I)\rangle$	21.0 (3.3)
$R_{\rm merge}$†	0.067 (0.394)
Model refinement	
$R_{\rm work}/R_{\rm free}$‡	0.204/0.245
No. of non-H atoms	
Protein	4090
Water	366
Sulfate ions	25
Average B factor (Å^2)	
Protein	25.6
Water	41.6
Sulfate ions	40.3
R.m.s. deviations from ideal geometry	
Bond lengths (Å)	0.007
Bond angles (°)	1.03
Ramachandran plot	
Favoured (%)	93.5
Allowed (%)	6.5
Outliers (%)	0.0

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I(hkl)$ is the intensity of reflection hkl, $\sum_{hkl}$ is the sum over all reflections and $\sum_i$ is the sum over i measurements of reflection hkl. ‡ $R = \sum_{hkl}\big||F_{\rm obs}| - |F_{\rm calc}|\big|/\sum_{hkl}|F_{\rm obs}|$, where $R_{\rm free}$ is calculated for a randomly chosen 5% of reflections which were not used for structure refinement and $R_{\rm work}$ is calculated for the remaining reflections.

3. Results and discussion

In order to better understand catalysis and substrate recognition by ERp46, we crystallized the third catalytic domain of human ERp46 (residues 323–428) and determined its three-dimensional structure using molecular replacement to 2.0 Å resolution (Table 1). The crystal contained five molecules in the asymmetric unit, four of which are very similar and one which appears to be partially unfolded. Of the ERp46 residues in the construct, only Leu371 of chains *A* and *B* and Thr323, Alal347–His352 and Asp380–Thr382 of chain *E* are absent from the electron-density map.

The structure shows a typical thioredoxin-like fold containing a five-stranded central β-sheet with a $\beta1$–$\beta3$–$\beta2$–$\beta4$–$\beta5$ arrangement with all strands parallel except for strand $\beta4$. The β-sheet is flanked on both sides by two α-helices which are parallel to each other (Fig. 1*a*). Comparison of the chains in the asymmetric unit shows high structural similarity between chains *A* to *D*, with r.m.s.d.s varying between 0.27 and 0.33 Å for 90–95 C$^\alpha$ atoms. Chain *E* is less similar, with an r.m.s.d. of 0.73 Å to chain *A* for 82 C$^\alpha$ atoms. Overlay of chain *A* with the solution NMR structure results in an r.m.s.d. of 0.83 Å over 102 C$^\alpha$ atoms. Importantly, the overlays also identify regions of high structural plasticity in the domain. In particular, the loop between helix $\alpha2$ and strand $\beta3$ adopts strikingly different conformations in chains *D* and *E* in comparison with the other three chains of the asymmetric unit. This is a result of different crystal contacts involving this loop (Fig. 1*a*). Residue Leu371 at the tip of the loop is absent from the electron-density map in chains *A* and *B*. Interestingly, the conformational flexibility in this loop is much more pronounced in the crystal structure than in the NMR structure of this domain (PDB entry 2diz).

The conserved catalytic C*xx*C motif is located at the N-terminus of the long $\alpha2$ helix, with the catalytic cysteines reduced in all five molecules. In the vicinity of the catalytic site, chains *A*, *B*, *C* and *D* each contain a disulfide bond between Cys381 and Cys388 connecting helix $\alpha3$ to the preceding loop (Fig. 1*a*). This disulfide is absent in chain *E*, which leads to striking conformational changes (Fig. 1*b*). In

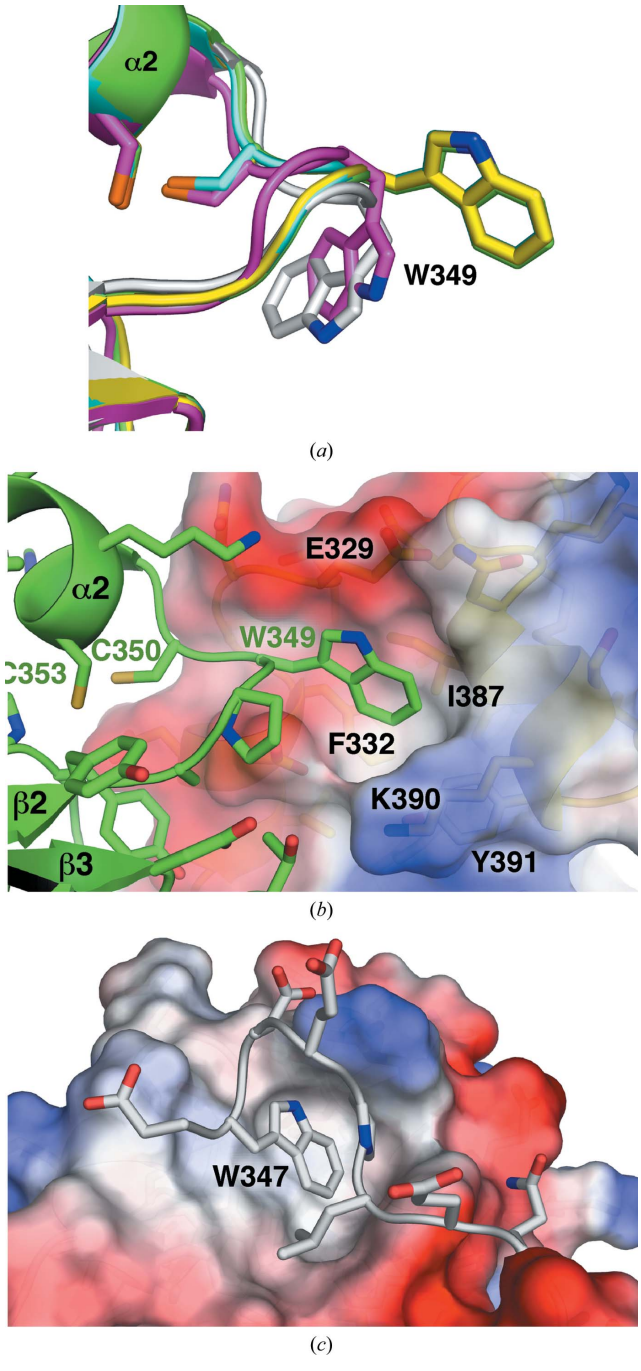


Figure 2
Crystal contacts identify putative substrate-recognition elements. (*a*) The side chain of Trp349 adopts a markedly different orientation in the different chains. In chain *C* (magenta) and the solution structure (PDB entry 2diz; grey) the tryptophan is backed against the catalytic cysteine Cys350, while in chains *A* (green) and *D* (yellow) the tryptophan is turned out and makes interdomain crystal contacts. (*b*) Crystal packing at the interface between chains *A* and *D* is mediated in part by the insertion of tryptophan Trp349 of chain *A* into a cavity formed by Phe332, Ile387, Glu329, Lys390 and Trp391 of chain *D*. Similar crystal contacts occur between Trp349 of chains *B* and *D* and neighbouring molecules across a crystallographic axis. (*c*) Trp347 of the b'–a' linker of PDIA1 inserts into the substrate-binding site of the b' domain. The domain is shown in a similar orientation to that in (*b*).

particular, the α3 helix unfolds and adopts a more extended irregular conformation. Overall, this region is characterized by significantly higher *B* factors (55–60 Å^2) than the rest of the domain (30–35 Å^2). It is noteworthy that these changes also result in the catalytic site of this domain being disordered in chain *E*. These results clearly show the important structural role of the Cys381–Cys388 disulfide in stabilizing the domain.

The tryptophan residue Trp349 preceding the C*xx*C motif adopts a very different conformation in chain *C* relative to the other chains in the asymmetric unit (Fig. 2*a*). Analysis of intermolecular contacts in the crystal shows that this tryptophan plays a crucial role in crystal lattice formation. In three of the five chains the side chain of this residue is inserted into a pocket formed by the side chains of Phe332, Ile387, Glu329, Lys390 and Tyr391 between helices α1 and α3 of another ERp46 domain (Fig. 2*b*). Previously, this tryptophan was suggested to be involved in the binding to substrate proteins based on its proximity to the catalytic cysteines (Kozlov *et al.*, 2009). In the structure of the PDI-family member ERp57 in complex with tapasin, the residue at the position of Trp349 makes important intermolecular contacts with tapasin (Dong *et al.*, 2009). These intermolecular contacts involving Trp349 are reminiscent of the *b*′*x* structure of human PDIA1, the archetypal member of the disulfide isomerase family (Fig. 2*c*). In PDIA1 a tryptophan residue of the linker connecting domains *b*′ and *a*′ is inserted into the substrate-binding site positioned between helices α1 and α3 of the noncatalytic *b*′ domain (Nguyen *et al.*, 2008). In the absence of noncatalytic domains with substrate-binding sites, the interactions observed in the ERp46 crystal suggest that the third catalytic domain of ERp46 may have some ability to bind substrates *via* the small hydrophobic pocket on the α1–α3 surface or through the exposure of residue Trp349.

An unusual feature of the ERp46 structure is the presence of a buried lysine residue Lys344 close to the second cysteine of the C*xx*C motif (Fig. 3*a*). This lysine forms a salt bridge with Glu378 on the adjacent β-strand. This arrangement is opposite to that observed in other PDI-family members, in which a glutamate is positioned underneath the catalytic site and is thought to participate in catalysis by accepting a proton from the second cysteine (Fig. 3*b*). Another residue that is likely to be implicated in catalysis is the conserved arginine residue Arg415 (Fig. 3*a*). Mutagenesis of equivalent arginines in the catalytic domains of PDIA1 affects activity by modulating the pK_a values of the second cysteine in the C*xx*C motifs (Karala *et al.*, 2010). The proposed mechanism of this modulation relies on transient polar contacts of the arginine with the cysteine side chain (Lappi *et al.*, 2004). Future studies should address the role of these residues in modulating the reactivity of the C*xx*C motifs in ERp46.

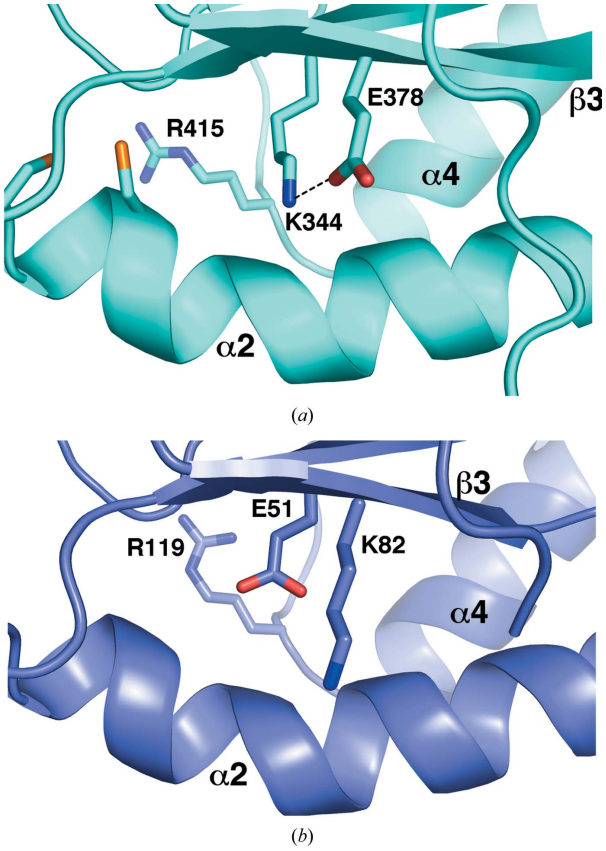


Figure 3
A conserved charged pair is reversed in ERp46 relative to other PDI-domain structures. (*a*) Lys344 is buried at the base of the catalytic site and makes a salt bridge with Glu378 on the adjacent β-strand. Based on homology to other PDI domains, Arg415 may be important for the disulfide-isomerization activity of ERp46. (*b*) The arrangement is opposite in protein disulfide isomerase ERp57 (PDB entry 3f8u), in which Glu51 is buried.

This work was funded by Canadian Institutes of Health Research grant MOP-81277. Data acquisition took place at the Macromolecular Diffraction (MacCHESS) facility at the Cornell High Energy Synchrotron Source (CHESS). CHESS is supported by the NSF and NIH/NIGMS *via* NSF award DMR-0225180; the MacCHESS resource is supported by NIH/NCRR award RR-01646.

References

Alberti, A., Karamessinis, P., Peroulis, M., Kypreou, K., Kavvadas, P., Pagakis, S., Politis, P. K. & Charonis, A. (2009). *Am. J. Physiol. Endocrinol. Metab.* **297**, E812–E821.
Dong, G., Wearsch, P. A., Peaper, D. R., Cresswell, P. & Reinisch, K. M. (2009). *Immunity*, **30**, 21–32.
Dyson, H. J., Jeng, M. F., Tennant, L. L., Slaby, I., Lindell, M., Cui, D. S., Kuprin, S. & Holmgren, A. (1997). *Biochemistry*, **36**, 2622–2636.
Ellgaard, L. & Ruddock, L. W. (2005). *EMBO Rep.* **6**, 28–32.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Hatahet, F. & Ruddock, L. W. (2009). *Antioxid. Redox Signal.* **11**, 2807–2850.
Jessop, C. E., Watkins, R. H., Simmons, J. J., Tasab, M. & Bulleid, N. J. (2009). *J. Cell Sci.* **122**, 4287–4295.
Karala, A. R., Lappi, A. K. & Ruddock, L. W. (2010). *J. Mol. Biol.* **396**, 883–892.
Kozlov, G., Määttänen, P., Schrag, J. D., Hura, G. L., Gabrielli, L., Cygler, M., Thomas, D. Y. & Gehring, K. (2009). *Structure*, **17**, 651–659.
Lappi, A. K., Lensink, M. F., Alanen, H. I., Salo, K. E., Lobell, M., Juffer, A. H. & Ruddock, L. W. (2004). *J. Mol. Biol.* **335**, 283–295.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Maattanen, P., Kozlov, G., Gehring, K. & Thomas, D. Y. (2006). *Biochem. Cell Biol.* **84**, 881–889.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Nguyen, V. D., Wallis, K., Howard, M. J., Haapalainen, A. M., Salo, K. E., Saaranen, M. J., Sidhu, A., Wierenga, R. K., Freedman, R. B., Ruddock, L. W. & Williamson, R. A. (2008). *J. Mol. Biol.* **383**, 1144–1155.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Sullivan, D. C., Huminiecki, L., Moore, J. W., Boyle, J. J., Poulsom, R., Creamer, D., Barker, J. & Bicknell, R. (2003). *J. Biol. Chem.* **278**, 47079–47088.
Walker, K. W. & Gilbert, H. F. (1997). *J. Biol. Chem.* **272**, 8845–8848.
Winn, M. D., Murshudov, G. N. & Papiz, M. Z. (2003). *Methods Enzymol.* **374**, 300–321.