Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structural features of peroxisomal catalase from the yeast *Hansenula polymorpha*

Esther Peña-Soler,[a] M. Cristina Vega,[a] Matthias Wilmanns[b] and Chris Williams[b]*

[a]Department of Structural and Quantitative Biology, Centro de Investigaciones Biológicas (CIB-CSIC), Ramiro de Maetzu, 28040 Madrid, Spain, and [b]Structural Biology Unit, European Molecular Biology Laboratory, Notkestrasse 85, 22603 Hamburg, Germany

Correspondence e-mail: c.williams@embl-hamburg.de

Received 21 March 2011
Accepted 10 June 2011

PDB Reference: peroxisomal catalase, 2xq1.

The reactive oxygen species hydrogen peroxide is a byproduct of the β-oxidation process that occurs in peroxisomes. Since reactive oxygen species can cause serious damage to biomolecules, a number of scavengers control their intracellular levels. One such scavenger that is present in the peroxisome is the oxidoreductase catalase. In this study, the crystal structure of heterologously expressed peroxisomal catalase from the thermotolerant yeast *Hansenula polymorpha* has been determined at 2.9 Å resolution. *H. polymorpha* catalase is a typical peroxisomal catalase; it is tetrameric and is highly similar to catalases from other organisms. However, its hydrogen peroxide-degrading activity is higher than those of a number of other catalases for which structural data are available. Structural superimpositions indicate that the nature of the major channel, the path for hydrogen peroxide to the active site, varies from those seen in other catalase structures, an observation that may account for the high activity of *H. polymorpha* catalase.

1. Introduction

Hansenula polymorpha (also known as *Pichia angusta*) is a thermotolerant methylotrophic yeast from the family Saccharomycetaceae. Since its first characterization (Levine & Cooney, 1973), *H. polymorpha* has been employed in a number of molecular biology and biotechnology applications, such as the cost-effective production of 1,3-propanediol (Hong *et al.*, 2011) and bioethanol (Hong *et al.*, 2010), the production of pharmaceutical products, including vaccines (Tregnaghi *et al.*, 2010), biosensors (Smutok *et al.*, 2007; Demkiv *et al.*, 2008) and antibiotics (Gidijala *et al.*, 2009), as a host system for the expression of heterologous proteins (Kensy *et al.*, 2009; Song *et al.*, 2010) and to improve general biotechnological processes (Kottmeier *et al.*, 2009). Additionally, *H. polymorpha* is a model organism for the study of peroxisomal function (Verduyn *et al.*, 1988; van Zutphen *et al.*, 2008; Bener Aksam *et al.*, 2008). Peroxisomes are single membrane-bound organelles that are found in all eukaryotic cells. Functionally, they compartmentalize metabolic enzymes involved in several cellular pathways, including fatty-acid β-oxidation and glyoxylate metabolism. However, peroxisome function depends greatly on the particular organism under investigation, the stage of development of the organism and its environment. In *H. polymorpha*, peroxisomes play a crucial role in the metabolism of unusual carbon and nitrogen sources (Saraya *et al.*, 2010).

One of the most common end products of all oxidative reactions taking place within the peroxisome is the reactive oxygen species hydrogen peroxide (H_2O_2). At low concentrations H_2O_2 regulates physiological processes such as cell proliferation, apoptosis, carbohydrate metabolism and platelet activation, but at high concentrations H_2O_2 is a toxic agent that is capable of damaging proteins, DNA and lipids. H_2O_2 levels within the cell therefore require tight regulation. Peroxisomal catalases play a crucial role in keeping the amount of H_2O_2 under control by decomposing it to water and molecular oxygen. While an absence of catalase (known as acatalasaemia or Takahara disease) is not fatal *per se* (Takahara & Miyamoto, 1948), numerous studies have suggested that deficiency in catalase can increase the risk of developing other pathologies (Fukuoka *et al.*, 2008; Takemoto *et al.*, 2009; Kikumoto *et al.*, 2010) and recent data have indicated that peroxisomal catalase may play an important role in ageing (Terlecky *et al.*, 2006).

Structural data are available for many catalases from different organisms and our understanding of how this important class of enzyme performs its function is considerable (Chelikani *et al.*, 2004). However, owing to the high level of conservation often seen between catalases from different organisms, it can be difficult to explain the observed differences between catalases based on amino-acid sequence alone. Therefore, in order to obtain further insight into structure–function relationships in catalase enzymes, we have performed biochemical, biophysical and structural studies on the previously uncharacterized peroxisomal catalase from *H. polymorpha*. We present data on the oligomeric state of the enzyme in solution and its hydrogen peroxide-degrading activity. Additionally, we compare our structure with those of other catalases, commenting on the overall fold, the haem environment and the major channel.

2. Materials and methods

2.1. Cloning and expression

The *H. polymorpha* catalase 1 gene was amplified from genomic DNA using the primers 5′-GCG**CCATGG**CAAACCCCCCTGTTTTC-3′ and 5′-GCG**AAGCTT**TATATTTTGGATGGAGAAGAAG-3′, which introduced an alanine residue instead of a serine at position 2 and also included *Nco*I and *Hind*III restriction sites, respectively (shown in bold). The resulting PCR product was digested with *Nco*I and *Hin*dIII and ligated into *Nco*I–*Hin*dIII-digested pETM11. The final construct contained an N-terminal His_6 tag followed by a tobacco etch virus (TEV) protease cleavage site. This vector was transformed into *Escherichia coli* BL21 (DE3) RIL (Stratagene) and transformants were grown at 310 K in Terrific Broth medium supplemented with 25 mg ml^{-1} kanamycin and 34 mg ml^{-1} chloramphenicol until an optical density at 600 nm of 1 was attained. Protein expression was then induced with 1 m*M* isopropyl β-D-1-thiogalactopyranoside for 16 h at 298 K. Cells were harvested by centrifugation (5500*g* at 277 K for 20 min) and the resulting pellets were stored at 253 K until required.

2.2. Purification

The cells were lysed by pulse sonication (3 × 5 min at 40% amplitude, 5 s pulse/5 s rest) after resuspension in lysis buffer [50 m*M* Tris pH 7.5, 300 m*M* NaCl, 2 m*M* β-mercaptoethanol, 10% glycerol, 1 *M* urea, DNAse I, protease inhibitors (Sigma) and 1 mg ml^{-1} lysozyme] and cell debris was removed by centrifugation (20 000*g* at 277 K for 45 min). Lysates containing His_6-catalase were loaded onto Ni–NTA resin pre-equilibrated with buffer *A* (50 m*M* Tris pH 7.5, 300 m*M* NaCl, 2 m*M* β-mercaptoethanol, 10% glycerol). After extensive washing of the resin with buffer *A*, His_6-catalase was eluted with buffer *A* containing 500 m*M* imidazole. The fusion protein was then cleaved by incubation overnight at 277 K with His_6-TEV protease (50:1 catalase:TEV) with concomitant dialysis against 50 m*M* Tris pH 7.5, 140 m*M* NaCl, 10% glycerol, 2 m*M* β-mercaptoethanol and again loaded onto Ni–NTA (pre-equilibrated with buffer *A* supplemented with 25 m*M* imidazole). The His_6 tag, uncleaved fusion protein and the His_6-TEV protease were retained on the column, while pure catalase was present in the flowthrough. The protein was concentrated using a Vivaspin 20 (GE Healthcare, 30 kDa cutoff) and loaded onto a Superdex 200 16/60 column (GE Healthcare) equilibrated with 50 m*M* Tris pH 7.5, 300 m*M* NaCl, 1% glycerol, 2 m*M* β-mercaptoethanol. Elution fractions were analysed by SDS–PAGE. Protein concentration was measured using a NanoDrop (http://www.nanodrop.com).

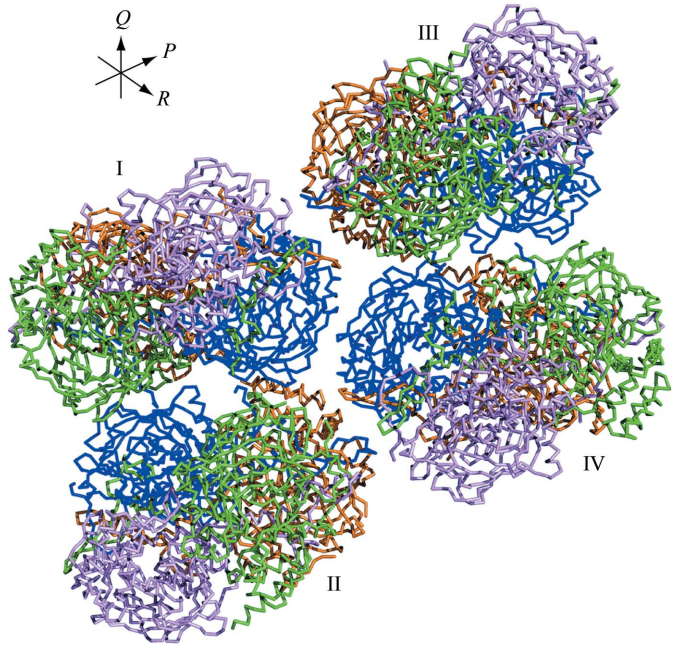

Figure 1
Arrangement of the four catalase tetramers (I–IV) in the asymmetric unit. Monomers (ribbon representation) are colour-coded according to their position in the tetramer, with chains *A*, *E*, *I* and *M* in green, chains *B*, *F*, *J* and *N* in blue, chains *C*, *G*, *K* and *O* in orange and chains *D*, *H*, *L* and *P* in pink.

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Data collection	
Resolution range (Å)	49.43–2.9 (3.06–2.9)
Unit-cell parameters (Å, °)	a = 132.5, b = 196.7, c = 170.8, $\alpha = \gamma$ = 90.0, β = 92.8
Space group	$P2_1$
Multiplicity	4.2 (4.2)
Reflections	805846
Unique reflections	193013
Completeness (%)	100 (99.9)
R_{merge}† (%)	12.0 (56.6)
$\langle I/\sigma(I)\rangle$	10.7 (2.5)
Refinement	
Resolution range (Å)	20.0–2.90
Reflections used	191418
R/R_{free}‡	19.0/22.3
No. of atoms in the asymmetric unit	
Protein	63129
Ligand	645
Solvent	65
R.m.s. deviations	
Bond lengths (Å)	0.014
Bond angles (°)	1.47
Average B factor (Å^2)	45.04
MolProbity validation	
Clashscore (all atom contacts)	4.2 [100th percentile]
Poor rotamers (%)	3.91
Ramachandran (%)	
Favoured	96.28
Allowed	3.59
Disallowed	0.13
C^β deviations > 0.25 Å	18
MolProbity score	1.90 [99th percentile]
Residues with bad bonds (%)	0
Residues with bad angles (%)	0.1

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an observation, $\langle I(hkl)\rangle$ is the mean value for that reflection and the summations are over all reflections. ‡ $R = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes, respectively. R_{free} was calculated with 976 reflections excluded from the refinement

2.3. Crystallization

Initial crystallization screening was performed at 292 K using the sitting-drop vapour-diffusion method in 96-well plates at the high-throughput crystallization facility at EMBL Hamburg (Mueller-Dieckmann, 2006). A number of brown needle-like crystals were obtained using the PEGs and PEGs II suites (Qiagen). Following optimization, diffracting crystals were obtained after 3 d using hanging-drop vapour diffusion with 1 µl protein solution at a concentration of 6 mg ml^{-1} mixed with 1 µl reservoir solution consisting of 20%(*w*/*v*) PEG 3350, 0.2 *M* magnesium formate equilibrated against 500 µl reservoir solution at 292 K. Prior to storage in liquid nitrogen, crystals were soaked in a solution of the reservoir condition containing an additional 20%(*v*/*v*) glycerol.

2.4. Data collection, structure determination and refinement

Diffraction data were collected on beamline ID14eh1 at the European Synchrotron Radiation Facility (ESRF), Grenoble using an ADSC Q210 CCD detector at 100 K. Data were processed using *XDS* (Kabsch, 2010) and *COMBAT* (Winn *et al.*, 2011) and scaled using *SCALA* (Evans, 2006). The structure was solved by molecular replacement using the structure of peroxisomal catalase from *Saccharomyces cerevisiae* (PDB entry 1a4e; Maté *et al.*, 1999) as the search model. Initially, three tetramers could be identified in the asymmetric unit using the *MOLREP* program (Vagin & Teplyakov, 2010), giving an R factor of 48.6% and score of 0.443. However, calculated $2mF_o - DF_c$ and $mF_o - DF_c$ electron-density maps showed interpretable density for a fourth tetramer. A closer look at the crystal packing in the *MOLREP* output, taking into consideration the positioning of the electron density of the fourth tetramer, suggested that the four tetramers in the asymmetric unit could be divided into two tetramer 'pairs'. The individual tetramers in a single pair (I–II and III–IV; Fig. 1) are related through a rotation about the noncrystallographic R axis followed by translation along the Q axis (Fig. 1). The second pair is related to the first through a translation along the P axis (Fig. 1). Therefore, we manually constructed a pair of tetramers (corresponding to tetramers I and II in Fig. 1) based on the original *MOLREP* output with *Coot* (Emsley *et al.*, 2010) and used this model as a search template in a second molecular-replacement run. This approach allowed us to identify all four tetramers in the asymmetric unit and resulted in an improved R factor (41.5%) and score (0.592) compared with the initial *MOLREP* run. Subsequently, rigid-body and restrained refinement steps were carried out with *REFMAC5* (Murshudov *et al.*, 1997). After an initial round of rigid-body refinement, the model had R and R_{free} values of 39.3% and a figure of merit (FOM) of 0.57. Owing to the large number of monomers in the asymmetric unit (16), noncrystallographic symmetry (NCS) restraints (five groups) were applied during

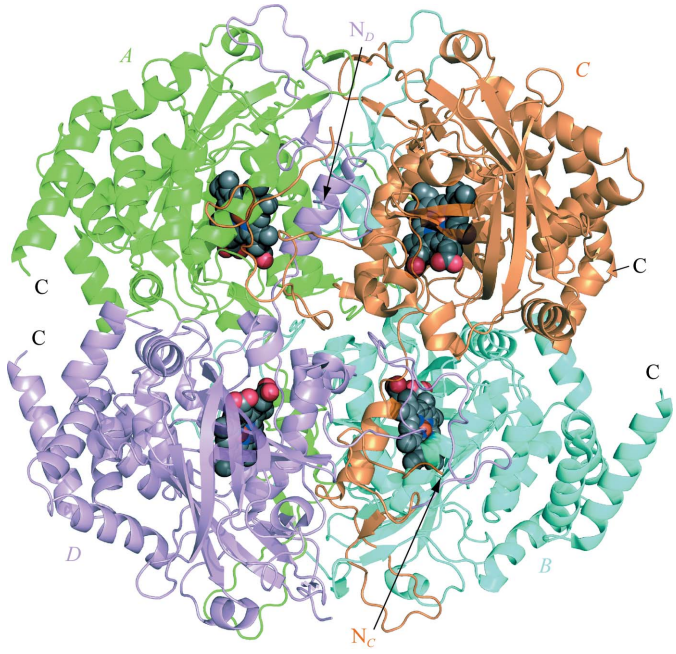

Figure 2
Structure of a tetramer (chains *A*–*D*) of *H. polymorpha* catalase. Each monomer coordinates an iron-bound haem group, depicted in sphere representation with C, O, N and Fe atoms coloured grey, red, blue and gold, respectively. The C-terminus of each monomer, together with the N-termini of chains *C* and *D* (N_C and N_D), are indicated. The N-termini of chains *A* and *B* are occluded in this orientation.

restrained refinement and, in the later stages, TLS parameters were used (Winn *et al.*, 2001). Tight NCS restraints were applied to residues 8–400 and 419–491 and medium restraints were applied to the more variable parts of the model (402–408, 414–417 and 493–501). For the R_{free} determination, 976 (0.5%) randomly selected reflections were omitted from the calculation. The R, R_{free} and FOM values for the final model are 19.0%, 22.3% and 0.860, respectively (Table 1). The quality of the final model was assessed with *MolProbity* (Chen *et al.*, 2010) and *PROCHECK* (Laskowski *et al.*, 1993) and interfaces were probed with *PISA* (Krissinel & Henrick, 2007). Structural representations were prepared using *PyMOL* (http://www.pymol.org). The coordinates and structure-factor data have been deposited in the Protein Data Bank (Berman *et al.*, 2003) with accession code 2xq1.

2.5. Catalase assay

Functional assays were carried out at 310 K in 50 m*M* potassium phosphate (KP_i) buffer pH 7.5, with catalase at a final concentration of 10.0 pg μl^{-1} and H_2O_2 at a concentration of 8.3 m*M*. The activity was followed by measuring the decrease in absorbance at 240 nm over a 5 min period, using an extinction coefficient of 43.6 M^{-1} cm^{-1} for H_2O_2 (Arnao *et al.*, 1990). A control experiment with KP_i buffer containing H_2O_2 was used to determine the rate of H_2O_2 degradation in

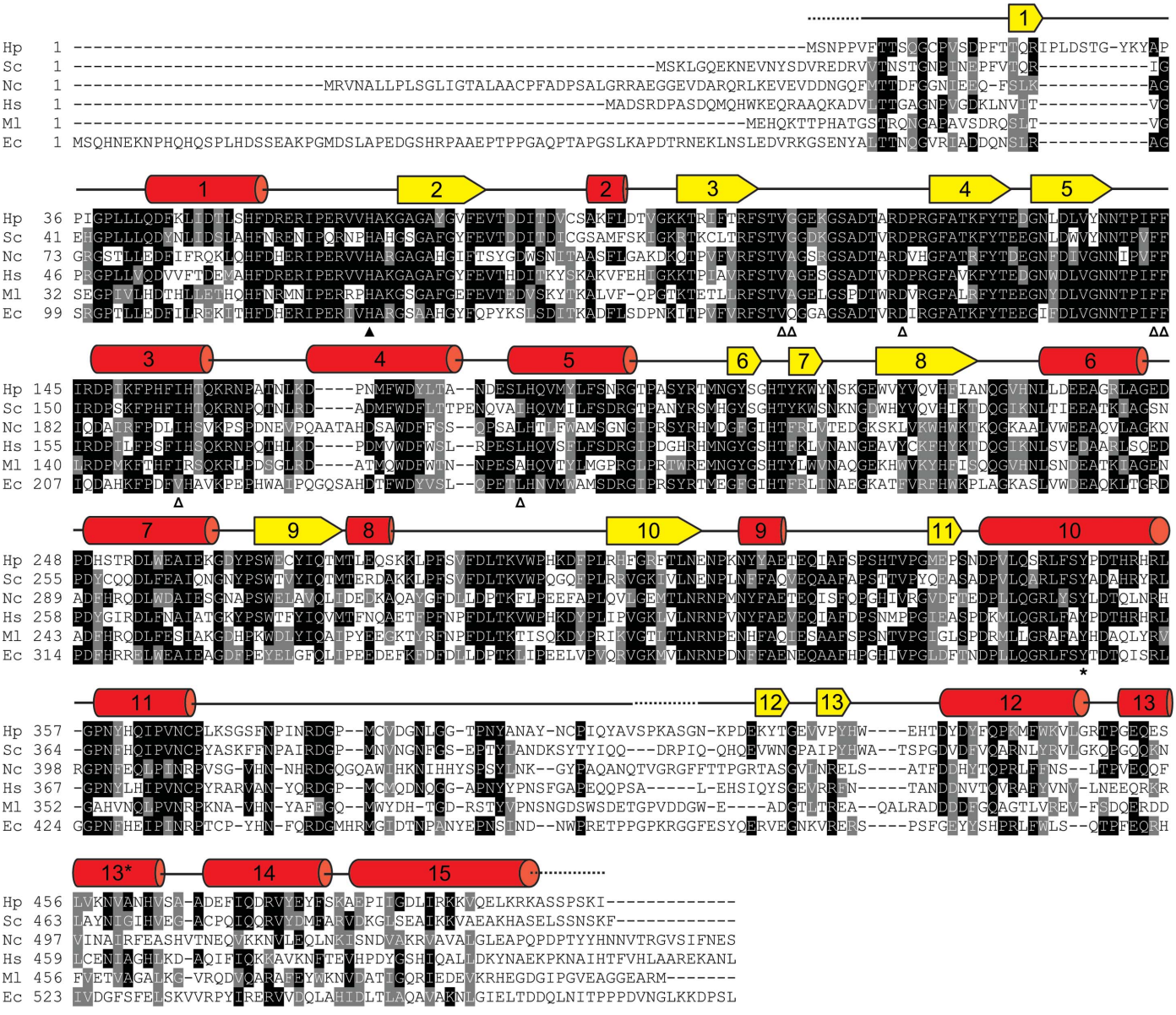

Figure 3
Sequence alignment of *H. polymorpha* catalase against a number of catalase proteins from different organisms; secondary-structure elements (helices in red, strands in yellow) are indicated above the alignment. The dotted lines indicate residues that are absent from the final model. Residues that line the major channel are indicated by open triangles (*cf.* Fig. 6). The active-site histidine is indicated by a solid triangle and Tyr348, the iron-coordinating residue, is indicated by an asterisk. Hp, *Hansenula polymorpha*; Sc, *Saccharomyces cerevisiae*; Hs, *Homo sapiens*; Nc, *Neurospora crassa*; Ml, *Micrococcus luteus*; Ec, *Escherichia coli*.

Table 2
H_2O_2-degrading activity of *H. polymorpha* catalase.

Organism	PDB code	Temperature (K)	Activity (μM H_2O_2 min^{-1} mg^{-1})	Identity to *H. polymorpha*† (%)	R.m.s.d.‡ (Å)
H. polymorpha	2xq1	310	$4.56 \times 10^5 \pm 4 \times 10^3$	N/A	N/A
S. cerevisiae	1a4e	310	1.16×10^5§	54.9	0.63
N. crassa	3ej6	298	1.02×10^5¶	27.3	0.76
H. sapiens	1dgf	310	2.74×10^5§	48.4	0.75
M. luteus	1gwe	310	1.09×10^5§	39.8	0.77
E. coli	1gge	310	2.07×10^5§	28.5	0.79

† As calculated by *LALIGN* (http://www.ch.embnet.org/software/LALIGN_form.html). ‡ Across all C^{α} atoms, as calculated by *RAPIDO* (Mosca & Schneider, 2008). § Switala & Loewen (2002). ¶ Jacob & Orme-Johnson (1979).

the absence of catalase. Specific activity is defined as the amount of hydrogen peroxide (micromoles) degraded per milligram of catalase per minute.

2.6. Miscellaneous

The use of static light scattering (SLS) coupled to the ÄKTA system has been described by Nettleship *et al.* (2008). Catalase sequences were compared using *ClustalW* (Thompson *et al.*, 1994) and structural superimposition and root-mean-square deviation (r.m.s.d.) calculations were made with *RAPIDO* (Mosca & Schneider, 2008). The structures used for r.m.s.d. calculations and structural superimpositions were PDB entries 1a4e (Maté *et al.*, 1999), 3ej6 (Díaz *et al.*, 2009), 1dgf (Putnam *et al.*, 2000), 1gwe (Murshudov *et al.*, 2002) and 1gge (Melik-Adamyan *et al.*, 2001) (see Table 2).

3. Results and discussions

3.1. Purification, oligomeric state determination and enzymatic assays

Owing to the iron-coordinating haem group, protein solutions of active tetrameric catalase display a typical yellow–brown colour, a characteristic that allowed us to monitor the presence of catalase during purification. Following the initial Ni–NTA purification step, samples showed two distinct peaks when subjected to size-exclusion chromatography, with the peak that eluted first, thought to be the tetrameric form,

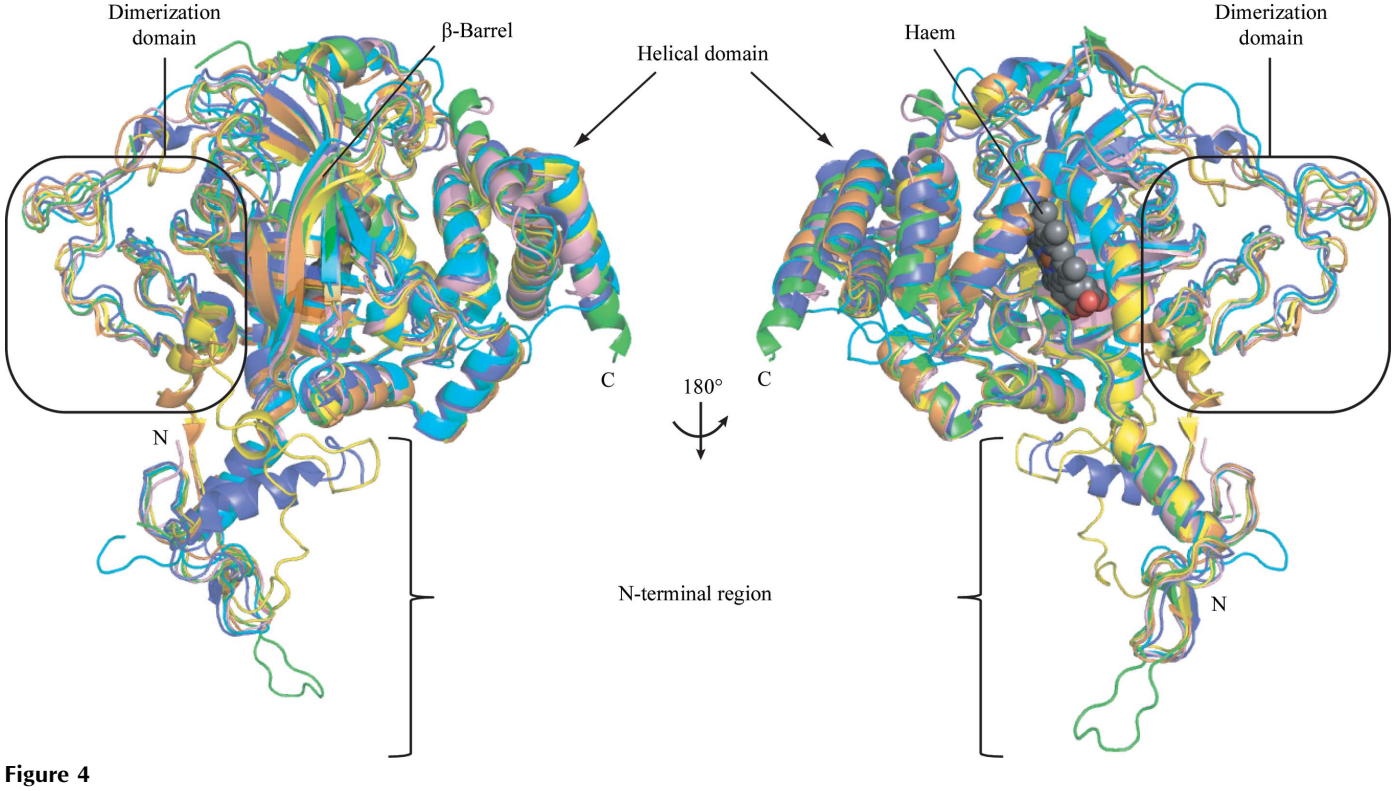

Figure 4
Structural superimposition of monomers of *H. polymorpha* (green), *S. cerevisiae* (pink), *N. crassa* (orange), *H. sapiens* (blue), *M. luteus* (cyan) and *E. coli* (yellow) catalase, indicating the β-barrel, N-terminal region, helical domain and dimerization domain (see text for details). For clarity, the C-terminal domains of the *N. crassa* and *E. coli* structures have been omitted. The haem group from the *H. polymorpha* structure is shown in sphere representation, with colours as in Fig. 2.

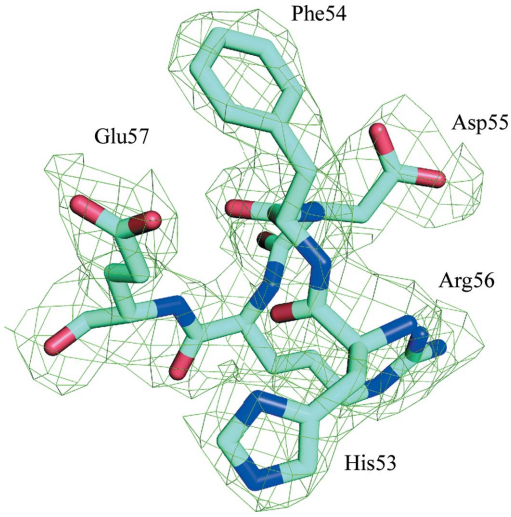

Figure 5
$mF_o - DF_c$ difference electron-density map calculated by omitting residues 53–57 from chain *J* during the refinement process. The map, which is contoured at 2.5σ, is superimposed onto the final model.

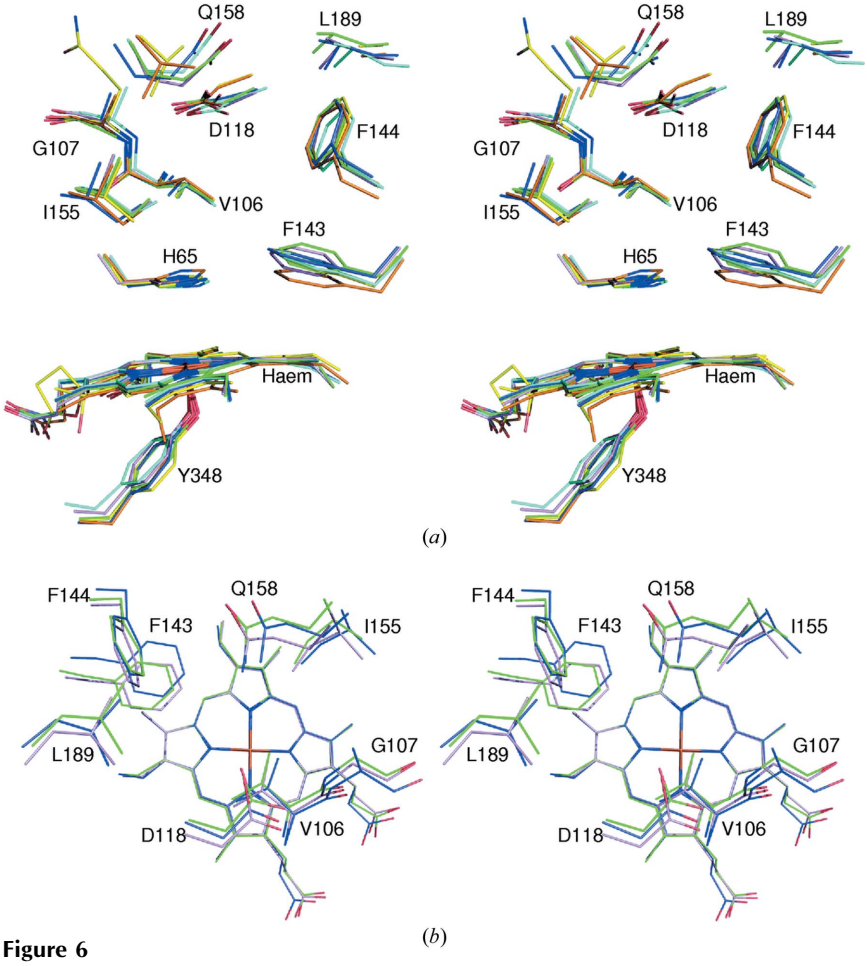

Figure 6
Stereoview of the haem environment and the lower region of the major channel in *H. polymorpha* catalase. (*a*) Lateral view of the major channels of *H. polymorpha*, *S. cerevisiae*, *N. crassa*, *M. luteus* and *E. coli* catalases, with residues (numbering for *H. polymorpha*, *cf.* Fig. 3) coloured as in Fig. 4. O and N atoms are shown in red and blue, respectively. (*b*) View along the major channel towards the haem group, showing for clarity only the *H. polymorpha*, *S. cerevisiae* and *H. sapiens* structures.

exhibiting an intense brown colour. The second peak, which was presumed to be the monomeric form of catalase, was colourless and additionally exhibited no H_2O_2-degrading activity (see below). Performing SLS analysis on pooled fractions of the first peak, we determined the molecular weight of the protein to be 239.5 kDa ± 4%, which corresponds well to the calculated molecular weight of tetrameric catalase (231.4 kDa). Additionally, the purified protein was able to degrade H_2O_2 (Table 2), indicating that we had indeed purified the active tetrameric enzyme. In comparison with catalases from a number of other organisms (for which structural data are available) we observed that the H_2O_2-degrading activity of our enzyme was higher than those of catalases from *Saccharomyces cerevisiae*, *Neurospora crassa*, *Homo sapiens*, *Micrococcus luteus* and *E. coli* (Table 2).

3.2. Overall structure

To gain further insight into the high level of activity seen for *H. polymorpha* catalase, we have solved the three-dimensional structure of the enzyme at 2.9 Å resolution (Figs. 1 and 2; Table 1). Overall, the quality of the model is good, with 96.28% of the residues in the favoured regions and an additional 3.59% of the residues in the allowed regions of the Ramachandran plot. A total of ten residues can be found in the disallowed region. Of these ten, four are present in extreme N-terminal regions (Pro5 from chains *A*, *I* and *M* and Asn3 from chain *J*), where the density is generally of poor quality. The other six residues correspond to Asp91 in six separate chains (*C*, *E*, *I*, *J*, *N* and *P*). The geometry of this residue, which is present at the C-terminus of α2, may be influenced by its hydrogen bond to the initiating residue of β-strand 3, Lys95, an arrangement that could restrict its orientation. The equivalent residues in the other chains (*A*, *B*, *D*, *F*, *G*, *H*, *K*, *L*, *M* and *O*) are all found in the generously allowed region of the Ramachandran plot, suggesting a general trend for Asp91 to adopt a less than favourable geometry. The assembly state of *H. polymorpha* catalase, as observed in all other haem-coordinating catalases, is tetrameric with accurate 222 molecular symmetry (Fig. 2). The asymmetric unit contains a total of 16 catalase molecules, which are in turn arranged into four tetramers (Fig. 1). This observation could suggest that *H. polymorpha* catalase forms higher order complexes in solution. However, analysis of the intratetramer interfaces with *PISA* (Krissinel & Henrick, 2007) indicates that these interfaces are involved in crystal contacts and the only significant interfaces

are found between the individual subunits of each tetramer. Additionally, we saw no evidence of higher order complexes when using SLS coupled with gel filtration.

3.3. Description of the monomer

Interpretable density is visible for nearly all main-chain atoms between Val6 and Lys500 in each chain, with the exception of six residues that are present in the loop between α12 and β12 (Fig. 3). The high sequence identity shared between different catalases (Table 2) is translated into high structural conservation (Figs. 3 and 4). R.m.s.d. values calculated across all C^{α} atoms, comparing monomers of *H. polymorpha* with those of *S. cerevisiae*, *N. crassa*, *H. sapiens*, *M. luteus* and *E. coli*, can be seen in Table 2. *H. polymorpha* catalase is a typical small clade 3-type catalase (Sicking *et al.*, 2008) and contains a central β-barrel formed by β-strands 2–10, a 70-residue N-terminal arm which runs along the interface between two monomers, a helical domain lying on one side of the β-barrel and comprising α-helices 3–7 and 12–15 (the last of which corresponds to the C-terminus), and a 'dimerization' domain on the other side formed by the loop region between α11 and β12 which embeds itself into the neighbouring monomer (Figs. 2 and 4). To assess potential model bias introduced because of the high sequence and structural homology shared with the molecular-replacement model used, we systematically deleted a number of regions from the model and performed further refinement steps. In all cases we observed a good correlation between the calculated $mF_o - DF_c$ electron-density maps and our final model (Fig. 5).

3.4. Haem environment and the major channel

All *H. polymorpha* catalase monomers coordinate a bent iron-bound type-b haem group sitting in a pocket formed by β-strands 2–4 on one side and α-helices 3 and 4 on the other and closed by the loop between α2 and β2. The central Fe atom, which is coordinated by the four pyrrole N atoms from the haem group, is also apically bound by a fifth ligand, Tyr348 (Fig. 6). The sixth coordination site, which is the site at which H_2O_2 binds, does not exhibit any interpretable density in any of the chains in the asymmetric unit. The haem group, the orientation of which is typical of small clade 3-type catalases (Sicking *et al.*, 2008), lies at the bottom of the major channel, a path of around 25 Å that leads from the surface of the enzyme to the distal pocket of the haem environment in which the active-site histidine (His65) is present (Fig. 6). This channel, which becomes constricted around 12 Å above the haem group, plays a crucial role in regulating access to the active site, allowing only small molecules such as H_2O_2 to enter the distal pocket (Maté *et al.*, 1999; Chelikani *et al.*, 2003). This is illustrated by studies on the *S. cerevisiae* enzyme. Val111, which is invariant in catalases (Val106 in *H. polymorpha*; Fig. 3) restricts access to the distal pocket (Fig. 6). Mutating this residue to an alanine decreases the ability of the protein to degrade H_2O_2, while resulting in a concomitant gain in non-specific activity (Maté *et al.*, 1999). Such an increase in the size of the channel is very likely to disturb a network of water molecules present in the channel, the function of which is thought to be to gauge the size of substrates through a 'molecular ruler' system (Putnam *et al.*, 2000).

A closer look at the *H. polymorpha* major channel reveals that the residues lining the channel are well conserved (Figs. 3 and 6). However, subtle differences in the positioning of the Asp118–Gln158–Leu189 triad, residues whose equivalents in the other structures form the mouth of the constriction 12 Å above the haem group (Fig. 6), block the channel leading

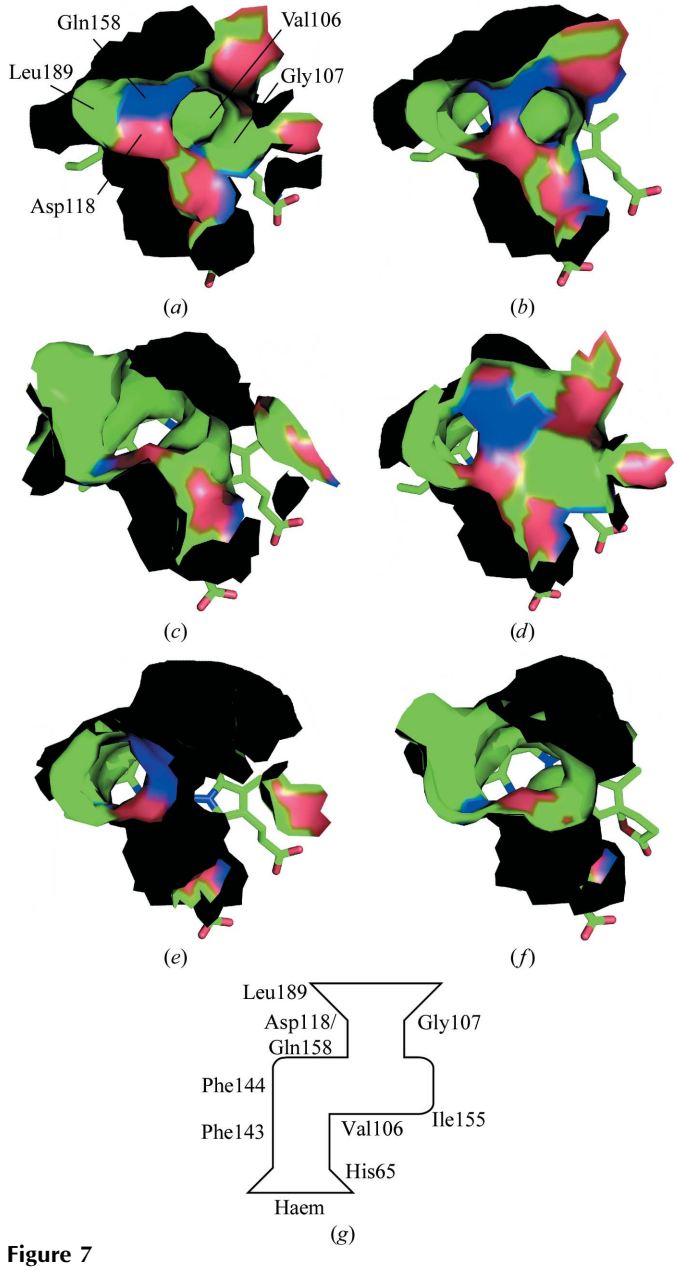

Figure 7
Surface representation (orientation as in Fig. 6*b*) of the major channel of (*a*) *H. polymorpha*, (*b*) *S. cerevisiae*, (*c*) *N. crassa*, (*d*) *H. sapiens*, (*e*) *M. luteus* and (*f*) *E. coli* catalase, showing the haem group in stick representation. C atoms are coloured green, O atoms red and N atoms blue. Black represents the inner side of the surface representation. (*g*) Schematic representation of the lower region of the major channel in *H. polymorpha* catalase, indicating the residues that form the walls of the channel (*cf.* Fig. 6).

towards the haem group (Fig. 7). Instead, a separate entrance to the active site can be seen caused by the glycine residue adjacent to Val111 (Figs. 6 and 7). In the other structures depicted the side chain of the equivalent residue occludes this entrance, except interestingly in the *S. cerevisiae* structure, in which a bifurcation of the channel is caused by the presence of a glycine at this position (Figs. 3, 6 and 7). Consequently, the entrance to the distal pocket in *H. polymorpha* catalase appears to be no longer directly above the haem group, as is seen in the other structures (Fig. 7), and the final section of the channel is S-shaped rather than straight (Fig. 7*g*), a situation that could restrict access to the distal pocket and consequently affect the activity of *H. polymorpha* catalase.

4. Concluding remarks

In this study, we have been able to express and purify oligomeric cofactor-bound *H. polymorpha* catalase from a heterologous source and have solved the structure of the enzyme at 2.9 Å resolution. Although the resolution of our structure does not allow an accurate description of the major channel, comparisons between *H. polymorpha* catalase and a number of other catalase structures suggest that the nature of the substrate pathway is more restrictive in *H. polymorpha* catalase. Data on other catalases indicate that restricting the path towards the active site results in an enhancement in activity (Maté *et al.*, 1999; Chelikani *et al.*, 2003). Whether these variations play a defining role in the high level of H_2O_2-degrading activity exhibited by *H. polymorpha* catalase, however, remains to be seen.

The authors thank Ida van der Klei for providing *H. polymorpha* genomic DNA and Santosh Panjikar and Matthew Groves for helpful advice and critically reading the manuscript. This project was partially supported by a Rubicon Fellowship from the Netherlands Organization for Scientific Research (NWO) awarded to CW (825.08.023) and performed under the PhD program 'Structure and Function of Proteins' from the Universitat Autònoma de Barcelona (UAB).

References

Arnao, M. B., Acosta, M., del Río, J. A. & García-Cánovas, F. (1990). *Biochim. Biophys. Acta*, **1038**, 85–89.
Bener Aksam, E., Jungwirth, H., Kohlwein, S. D., Ring, J., Madeo, F., Veenhuis, M. & van der Klei, I. J. (2008). *Free Radic. Biol. Med.* **45**, 1115–1124.
Berman, H., Henrick, K. & Nakamura, H. (2003). *Nature Struct. Biol.* **10**, 980.
Chelikani, P., Carpena, X., Fita, I. & Loewen, P. C. (2003). *J. Biol. Chem.* **278**, 31290–31296.
Chelikani, P., Fita, I. & Loewen, P. C. (2004). *Cell. Mol. Life Sci.* **61**, 192–208.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Demkiv, O., Smutok, O., Paryzhak, S., Gayda, G., Sultanov, Y., Guschin, D., Shkil, H., Schuhmann, W. & Gonchar, M. (2008). *Talanta*, **76**, 837–846.
Díaz, A., Valdés, V. J., Rudiño-Piñera, E., Horjales, E. & Hansberg, W. (2009). *J. Mol. Biol.* **386**, 218–232.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Fukuoka, N., Sugiyama, H., Inoue, T., Kikumoto, Y., Takiue, K., Morinaga, H., Nakao, K., Maeshima, Y., Asanuma, M., Wang, D.-H., Ogino, K., Masuoka, N. & Makino, H. (2008). *Am. J. Nephrol.* **28**, 661–668.
Gidijala, L., Kiel, J. A., Douma, R. D., Seifar, R. M., van Gulik, W. M., Bovenberg, R. A., Veenhuis, M. & van der Klei, I. J. (2009). *PLoS One*, **4**, e8317.
Hong, W.-K., Kim, C.-H., Heo, S.-Y., Luo, L. H., Oh, B.-R., Rairakhwada, D. & Seo, J.-W. (2011). *Bioprocess Biosyst. Eng.* **34**, 231-326
Hong, W.-K., Kim, C.-H., Heo, S.-Y., Luo, L. H., Oh, B.-R. & Seo, J.-W. (2010). *Biotechnol. Lett.* **32**, 1077–1082.
Jacob, G. S. & Orme-Johnson, W. H. (1979). *Biochemistry*, **18**, 2967–2975.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kensy, F., Engelbrecht, C. & Büchs, J. (2009). *Microb. Cell Fact.* **8**, 68.
Kikumoto, Y., Sugiyama, H., Inoue, T., Morinaga, H., Takiue, K., Kitagawa, M., Fukuoka, N., Saeki, M., Maeshima, Y., Wang, D.-H., Ogino, K., Masuoka, N. & Makino, H. (2010). *Biochim. Biophys. Acta*, **1802**, 240–246.
Kottmeier, K., Muller, C., Huber, R. & Buchs, J. (2009). *Appl. Microbiol. Biotechnol.* **86**, 93–101.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Laskowski, R. A., Moss, D. S. & Thornton, J. M. (1993). *J. Mol. Biol.* **231**, 1049–1067.
Levine, D. W. & Cooney, C. L. (1973). *Appl. Microbiol.* **26**, 982–990.
Maté, M. J., Zamocky, M., Nykyri, L. M., Herzog, C., Alzari, P. M., Betzel, C., Koller, F. & Fita, I. (1999). *J. Mol. Biol.* **286**, 135–149.
Melik-Adamyan, W., Bravo, J., Carpena, X., Switala, J., Materé, M. J., Fita, I. & Loewen, P. C. (2001). *Proteins*, **44**, 270–281.
Mosca, R. & Schneider, T. R. (2008). *Nucleic Acids Res.* **36**, W42–W46.
Mueller-Dieckmann, J. (2006). *Acta Cryst.* D**62**, 1446–1452.
Murshudov, G. N., Grebenko, A. I., Brannigan, J. A., Antson, A. A., Barynin, V. V., Dodson, G. G., Dauter, Z., Wilson, K. S. & Melik-Adamyan, W. R. (2002). *Acta Cryst.* D**58**, 1972–1982.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Nettleship, J. E., Brown, J., Groves, M. R. & Geerlof, A. (2008). *Methods Mol. Biol.* **426**, 299–318.
Putnam, C. D., Arvai, A. S., Bourne, Y. & Tainer, J. A. (2000). *J. Mol. Biol.* **296**, 295–309.
Saraya, R., Veenhuis, M. & van der Klei, I. J. (2010). *FEBS J.* **277**, 3279–3288.
Sicking, W., Korth, H. G., de Groot, H. & Sustmann, R. (2008). *J. Am. Chem. Soc.* **130**, 7345–7356.
Smutok, O., Dmytruk, K., Gonchar, M., Sibirny, A. & Schuhmann, W. (2007). *Biosens. Bioelectron.* **23**, 599–605.
Song, H.-L., Niu, Z.-D., Qian, W.-D., Wang, H. & Qiu, B.-S. (2010). *Biotechnol. Lett.* **32**, 1473–1479.
Switala, J. & Loewen, P. C. (2002). *Arch. Biochem. Biophys.* **401**, 145–154.
Takahara, S. & Miyamoto, H. (1948). *Okayama Igakkai Zasshi*, **60**, 90.
Takemoto, K., Tanaka, M., Iwata, H., Nishihara, R., Ishihara, K., Wang, D.-H., Ogino, K., Taniuchi, K. & Masuoka, N. (2009). *Clin. Chim. Acta*, **407**, 43–46.
Terlecky, S. R., Koepke, J. I. & Walton, P. A. (2006). *Biochim. Biophys. Acta*, **1763**, 1749–1754.
Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Tregnaghi, M. W., Voelker, R., Santos-Lima, E. & Zambrano, B. (2010). *Vaccine*, **28**, 3595–3601.

Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Verduyn, C., Giuseppin, M. L., Scheffers, W. A. & van Dijken, J. P. (1988). *Appl. Environ. Microbiol.* **54**, 2086–2090.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). *Acta Cryst.* D**57**, 122–133.
Zutphen, T. van, van der Klei, I. J. & Kiel, J. A. (2008). *Methods Enzymol.* **451**, 197–215.