Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The three-dimensional structure of catalase from *Enterococcus faecalis*

Kjell O. Håkansson,[a]* Myriam Brugna[b] and Lena Tasse[b]

[a]August Krogh Institute, Copenhagen University, Universitetsparken 13, DK-2100 Kbh O, Denmark, and [b]Department of Cell and Organism Biology, Lund University, Sölvegatan 35, SE-223 62 Lund, Sweden

Correspondence e-mail: kohakansson@aki.ku.dk

Received 10 March 2004
Accepted 17 May 2004

PDB Reference: catalase, 1si8, r1si8sf.

Enterococcus faecalis haem catalase was crystallized using lithium sulfate at neutral pH. The crystals belong to space group *R*3, with unit-cell parameters $a = b = 236.9$, $c = 198.1$ Å. The three-dimensional structure was determined by molecular replacement using a subunit of the *Proteus mirabilis* catalase structure. It was refined against 2.3 Å synchrotron data to a free *R* factor of 21.8%. Like other catalases, the *E. faecalis* catalase is a homotetramer with a fold and structure similar to those of its structurally closest relative *P. mirabilis*. The solvent structure in the active site is identical in the four subunits but differs from that found in other catalases. The structural consequences of the Ramachandran outlier Ser196 are discussed.

1. Introduction

Catalase (EC 1.11.1.6) catalyses the decomposition of hydrogen peroxide to water and molecular oxygen. Catalase protects the cell from oxidative damage caused by hydrogen peroxide. There are three types of catalases: prokaryotic Mn-catalases (Barynin *et al.*, 2001), bifunctional catalase–peroxidases and haem catalases. The three-dimensional structure of haem catalase is known from crystallographic studies of the beef liver (Murthy *et al.*, 1981; Reid *et al.*, 1981), *Penicillium vitale* (Vainshtein *et al.*, 1981), *Micrococcus lysodeikticus* (Murshudov *et al.*, 1992, 2002), *Escherichia coli* (Bravo *et al.*, 1999, 1995), *Proteus mirabilis* (Gouet *et al.*, 1995), *Saccharomyces cerevisiae* (Mate *et al.*, 1999), human erythrocyte (Ko *et al.*, 2000; Putnam *et al.*, 2000) and *Pseudomonas syringae* (Carpena *et al.*, 2003) enzymes. Haem catalase is a tetrameric enzyme with 222 point-group symmetry. Each subunit has a prosthetic protoporphyrin IX group that complexes an Fe^{III} ion. The iron ion is directly involved in catalysis, which occurs at a near-diffusion-limited turnover rate. Firstly, a molecule of H_2O_2 is reduced to water, resulting in a cationic haem radical and an $Fe^{IV}{=}O$ ion called compound I. The dismutation cycle is completed by the reaction of a second H_2O_2, resulting in the release of O_2 and water and regeneration of the enzyme in the resting Fe^{III} state. Some catalases also bind NADPH, which has been implicated in preventing the accumulation of an inactive partially oxidized dead-end form of the enzyme called compound II (Kirkman *et al.*, 1999).

The Gram-positive bacterium *Enterococcus faecalis* lacks the ability to synthesize porphyrin compounds, including haem groups. However, haem supplied in the growth medium can be utilized for the expression of haem-containing proteins, including the recently characterized KatA, a 4×54 kDa catalytically active haem catalase (Frankenberg *et al.*, 2002; Pugh & Knowles, 1983). In the absence of haem, no catalase

Table 1
Data and refinement statistics for *E. faecalis* catalase.

The space group was *R*3, with unit-cell parameters $a = b = 236.9$, $c = 198.1$ Å. There is one tetramer per asymmetric unit. Values in parentheses are for the highest resolution shell.

Data-set statistics	
Resolution (Å)	20.0–2.3 (2.41–2.30)
No. unique reflections	183487 (26811)
Completeness (%)	99.7 (99.8)
R_{sym} (%)	7.8 (25.6)
Multiplicity	3.6 (3.3)
$I/\sigma(I)$	7.7 (2.9)
Refinement statistics	
Resolution range (Å)	20–2.3
R (%)	20.2
R_{free} (%)	21.8
Protein atoms	15104
Solvent atoms	1409
R.m.s.d. bond lengths (Å)	0.0068
R.m.s.d. bond angles (°)	1.53
R.m.s.d. dihedral angles (°)	24.0
R.m.s.d. improper angles (°)	0.86
B_{ave} protein (Å^2)	28.2
B_{ave} solvent (Å^2)	32.4

polypeptide can be found in *E. faecalis*. It is not known whether the catalase is important when the cell grows in the presence of haem. However, the acquisition of haem and expression of catalase by *E. faecalis* might be of clinical importance, as the pathological invasion of live tissues often means a change from an intestinal anaerobic environment to an aerobic lifestyle. We report the crystallization and three-dimensional structure determination of *E. faecalis* haem catalase.

2. Methods

E. faecalis catalase was purified as previously described (Frankenberg *et al.*, 2002). The protein was crystallized at a concentration of 10 mg ml^{-1} in 50 m*M* bis-tris pH 7.0, 1.6 *M* Li_2SO_4 by vapour equilibration in hanging drops. Crystals belong to space group *R*3, with unit-cell parameters $a = b = 236.9$, $c = 198.1$ Å, $\alpha = \beta = 90$, $\gamma = 120°$. Data were collected from crystals flash-frozen in 50% sucrose at MAX-lab synchrotron beam station I711, processed with *MOSFLM* (Leslie, 1992) and scaled with the *CCP*4 program package (Collaborative Computational Project, Number 4, 1994). The structure was solved by molecular replacement using a *P. mirabilis* catalase (Gouet *et al.*, 1995) monomer as search model and the *CCP*4 version of *AMoRe* (Navaza, 1994), yielding four solutions which were combined into a tetramer. With four catalase subunit molecules per asymmetric unit, the Matthews coefficient is 5.0 Å^3 Da^{-1}, corresponding to a solvent content of 75%. The structure was remodelled using the graphics program *O* (Jones *et al.*, 1991). The structure was refined and electron-density maps were calculated using *CNS* (Brünger *et al.*, 1998). Residues 4–477 were built in all four chains, but no residues could be visualized beyond these termini. Water molecules were automatically built using *ARP*/*wARP* (Perrakis *et al.*, 1999), but were inspected manually. Some water molecules were discarded, while others were rebuilt into two chloride ions and eight sulfate ions. Finally, 28 water molecules were defined manually. Non-crystallographic restraints of a value of 840 kJ mol^{-1} Å^{-2} were applied to residues 4–264, 285–411 and 416–471, with the exceptions of the side chains of residues 68, 72–73, 124, 159, 302, 387, 404, 426 and 433. The excluded parts displayed significant deviations from the non-crystallographic symmetry. The data and refinement statistics are shown in Table 1. One residue (see §3.4) falls outside the allowed regions in the Ramachandran plot (Fig. 1). All other residues are in most favoured regions (86.1%) or additional allowed regions (13.6%).

3. Results and discussion

3.1. Polypeptide structure

The structure is tetrameric with four identical subunits, each consisting of 478 amino acids (plus a hexahistidyl tag which was added in order to facilitate purification of the protein), as deduced from the gene sequence (Paulsen *et al.*, 2003) and confirmed by mass spectroscopy (M. Brugna & L. Hederstedt, unpublished data). The subunit labels *A*, *B*, *C* and *D* and the symmetry axes that relate them, *P*, *Q* and *R*, are the same as for the bovine liver enzyme (Murthy *et al.*, 1981; Reid *et al.*, 1981). The polypeptide chain was built from Gln4 to Gln477 for all subunits. There was no interpretable electron density for the three N-terminal residues or for Ala478 or the hexahistidyl sequence at the C-terminal end. The polypeptide chain of a single subunit can be divided into four regions: the N-terminal arm, the β-barrel, the wrapping region and the C-terminal helical region.

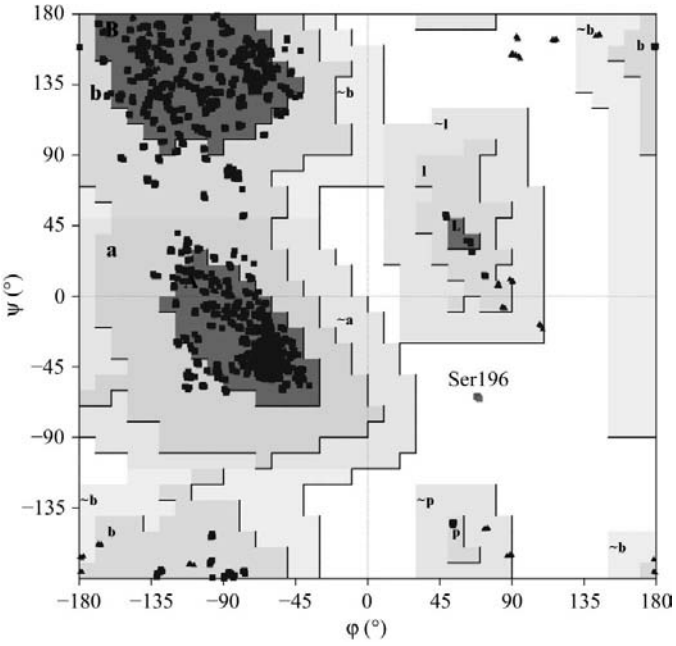

Figure 1
Ramanchandran plot of the *E. faecalis* catalase structure. The four Ser196 residues are the only non-glycine residues found outside the most favoured and additionally allowed regions (Laskowski *et al.*, 1993).

The N-terminal arm, Gln4–Glu50, has relatively few intra-chain contacts but is intricately weaved into the other subunits of the tetramer, where two *P*-related N-terminal regions form a wedge inserted between the wrapping regions (see below) of the other two subunits. The N-terminal region is hydrophilic, despite its relatively buried location, and a majority of the residues are hydrogen bonded to water molecules either through their main chains or their side chains. This region is followed by the more compact domain structure comprised of residues Arg51–Asn300, forming an antiparallel β-barrel, although there are only two backbone hydrogen bonds between strand 4 and strand 5. The barrel has three long insertions after strands 4, 6 and 7. The haem group is juxtaposed perpendicularly against the outside of the barrel at strands 2, 3 and 4 with the wrapping region (see below) on the proximal side and the Thr129–Arg189 insertion between strands 4 and 5 on the distal side. The wrapping region, encompassing residues Pro301–Asp415, begins with a short α-helix followed by random coil structure passing the proximal side of the haem. The polypeptide chain then turns around to present the helix Leu330–His343 on the proximal side of the haem. The rest of this region consists mainly of random coil structure which interacts with the wrapping region of a *P*-related subunit, but is also exposed to the other subunits and to the solvent. The C-terminal region His416–Gln477 forms a bundle of three α-helices positioned between the β-barrel and the solvent. The polypeptide fold of a single subunit and the tetramer structure are displayed in Figs. 2(*a*) and 2(*b*), respectively.

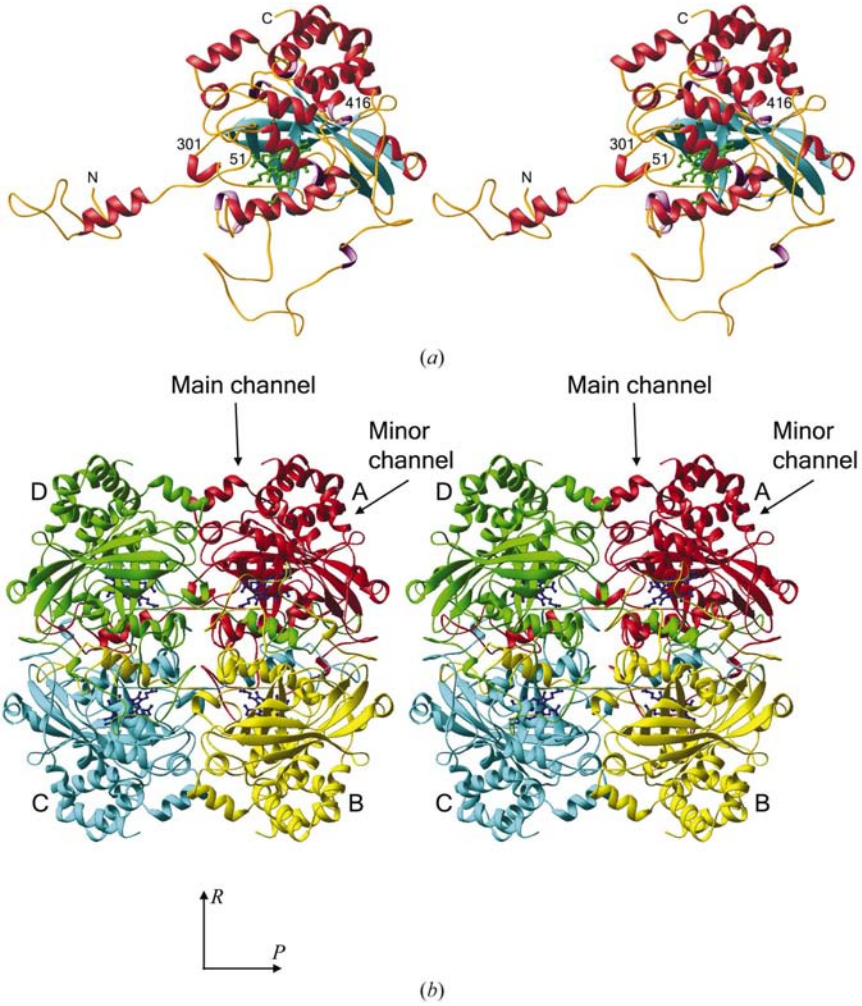

Figure 2
(*a*) The overall fold of one catalase subunit. β-Strands are shown in blue, α-helices in red, 3_{10}-helices in magenta and the haem group and the proximal ligand Tyr337 in green. The N- and C-termini, along with the positions of Arg51 (beginning of the β-barrel), Pro301 (beginning of the wrapping region) and His416 (beginning of the C-terminal region), are indicated. The Leu330–His343 helix with the proximal Fe ligand Tyr337 is seen below the haem group. (*b*) The catalase tetramer. The suggested entry sites (Fita & Rossmann, 1985; Gouet *et al.*, 1995) for subunit *A* are indicated. The orientations of the *P* and *R* axes are shown; the *Q* axis is perpendicular to the plane of the paper. Note how the N-termini of subunits *A* (shown from the same orientation as in *a*) and *B* penetrate into the wrapping regions of subunits *C* and *D*, to the left.

The conformations of the different subunits are essentially identical and were subject to non-crystallographic symmetry restraints, with the exception of three regions and a limited number of side chains. The three regions are Pro264—Lys285 (the β-barrel loop after strand 7), Ser411–Asp415 (just before the C-terminal region) and the C-terminus. The first two of these regions are close in space. Val269 displays the highest r.m.s. deviation of main-chain atom positions between residues in any two subunits, 1.6 Å. In two of the subunits (*C* and *D*), this region is involved in crystal packing. The highest main-chain r.m.s. deviation in the Ser411–Asp415 region is 0.5 Å (Asn413). The r.m.s. deviation for the symmetry-restrained regions were normally within 0.1 Å. The close conformational resemblance is in agreement with the structure of the other catalases and the fact that the nearest structurally known relative, *P. mirabilis* catalase, crystallizes with crystallographic symmetry.

The *E. faecalis* and *P. mirabilis* structures show identical backbone patterns for residues Gln4–Pro264, Phe276–Lys358, Glu392–Ser411, Asp415–Leu443 and Gln449–Gln477 (*E. faecalis* numbering). These regions superimpose with an r.m.s. deviation of C^{α} positions of 0.7 Å. Fig. 3 shows a sequence alignment of structurally investigated haem catalases. While the fold is similar throughout the group, deviations in loop regions are also evident (not shown). The weakest sequence similarity is found in the second part of the wrapping region and the C-terminal region. This part of the wrapping region, far from the active site, runs along the surface of the tetramer. The exact backbone pattern of this elongated random coil segment shows some variation in the group. However, mapping the conserved residues on the protein surface

(not shown) does not reveal any particularly conserved area, although the residues close to the Fe^{III} ion are conserved. The distribution of conserved residues is also relatively uniform in the sequence of the N-terminal arm, the β-barrel and the first part of the wrapping region.

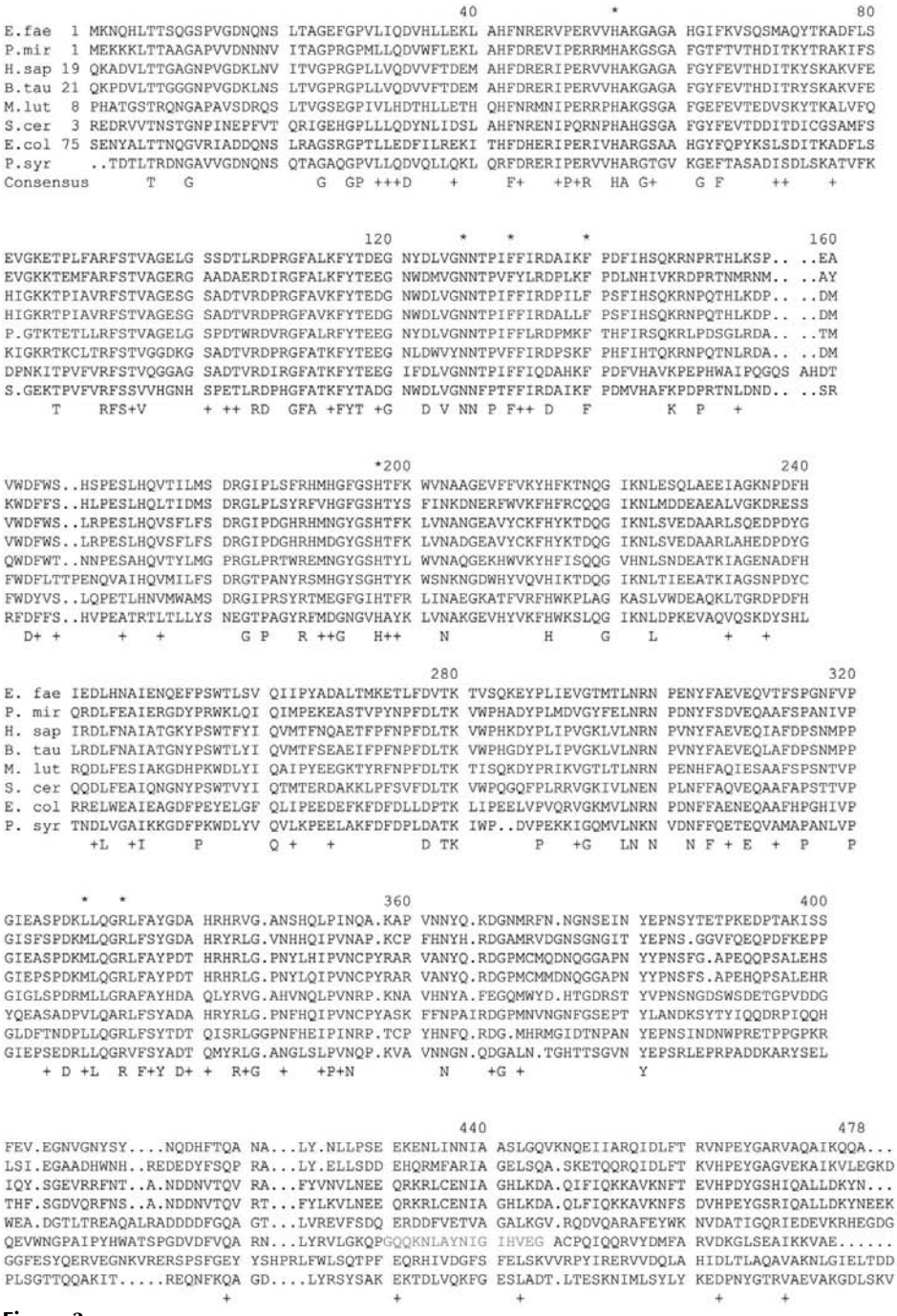

Figure 3
Sequence alignment of catalases of known structure and sequence, *i.e.* those from *E. faecalis*, *P. mirabilis*, *H. sapiens*, *B. taurus*, *M. lysodeikticus*, *S. cerevisiae*, *E. coli* and *P. syringae*. The alignment was performed with *MULTIALIGN* (Corpet, 1988) using the Blosum62 matrix. Only the parts of the sequences homologous to the *E. faecalis* sequence are shown and the numbers before and after the sequences indicate the number of the first residue and the number of residues following these parts, respectively. The numbers above the alignment show the *E. faecalis* residue numbers. The segment 431–446 (the second helix in the C-terminal region) in the *S. cerevisiae* sequence was adjusted manually after inspection of the structure and is shown in grey letters. Absence of sequence similarity and variations in backbone structure make the alignment in the 380–420 region uncertain. Some of the active-site residues discussed in the text are indicated by a '*' above the alignment. Consensus residues are shown in one-letter code if they are absolutely conserved and with a '+' if the residues in all sequences belong to the same one of the five groups (S, A, T), (D, E, N, Q, H), (K, R), (V, M, I, L) and (F, Y).

3.2. The haem group and its environment

The protoporphyrin IX molecule is bound with the same orientation as in *P. mirabilis* catalase, although there seems to be no steric hindrance to binding the haem in the alternative orientation, which would differ only in the positioning of the vinyl groups. The haem group orientation is similar in *M. lysodeikticus*, human erythrocyte, *S. cerevisiae* and beef liver catalase, but different in *E. coli* and *P. syringae* catalases. The porphyrin edge cornered by pyrrole rings II and III interacts with the β-barrel, from which it protrudes perpendicularly. The vinyl group of pyrrole ring II is in close contact with Ser196 of strand 5 and the important Asn127 (Loewen, 1996) of strand 4, while the methyl group interacts with Phe132 and His197. The methyl group of pyrrole ring III is in van der Waals contact with several atoms of the Gly110–Ala112 segment of strand 3. On the distal side, the imidazole group of the essential His54 (Loewen, 1996) is stacked parallel to pyrrole ring III. The carboxylate group of pyrrole ring III forms salt bridges with Arg51 of strand 1 and Arg 91 of strand 2, while the carboxylate of ring IV is ion-paired with Arg344. Finally, ring I is stacked parallel to Phe140 on the distal side and makes van der Waals contacts with Ala137. Part of the wrapping region, with the important residues Tyr337 and Arg333, faces the proximal side of the haem, while the Thr129–Arg189 segment dominates the distal side.

The haem group is distorted and deviates from planar geometry, similar to the devia-

tions from planarity observed in the other available catalase structures. There are several planar groups with delocalized π-electron systems that are stacked more or less parallel to the plane of the haem and close to the edges: Arg333 on the proximal side and Phe132, Phe140, Asn127 and His54 on the distal side. These are shown in Fig. 4 together with the proximal Fe ligand Tyr337.

The Fe^{III} ion is coordinated to the four pyrrole N atoms of the porphyrin group and to the phenolic side chain of Tyr337. The iron–oxygen distance is 1.8 Å, with a subunit variation of 1.76–1.88 Å. Tyr337 is also hydrogen bonded to Arg333, but neither the iron nor the Arg333 N atoms are in the plane of the phenolic side chain. This is in contrast to the general hydrogen-bonding pattern of tyrosine side chains (Ippolito *et al.*, 1990). The coordination geometry suggests that the hydroxyl group is sp^3-hybridized, which could be expected to enhance the electronegativity of the O atom.

3.3. Solvent structure

The water structure is very similar in all four subunits. Eight water molecules are found within a sphere of 10 Å around the Fe ion. These were compared after superimposing the haem group of different subunits. The highest deviation between corresponding water molecule positions in different subunits ranges from 0.03 to 0.33 Å, although the water molecules themselves were not used in the superimposition. The highest r.m.s. deviation of atomic positions for this set of eight water molecules between any two subunits is 0.21 Å.

In total, the model contains 1195 water molecules, four haem groups with Fe^{3+} ions, eight sulfate ions and two chloride ions per tetramer. The two chloride ions are found on the *P* axis, at the same positions as the sulfate ions found in *S. cerevisiae* catalase (Mate *et al.*, 1999). There are two sulfate ions per subunit. The more strongly bound (average $B = 40$ Å^2) of the two is hydrogen bonded to the backbone amino groups of residues Gly15 and Asp16, with similar orientations in all four subunits. The weaker of the two sulfate ions (average $B = 56$ Å^2) is bound in the same position as the 2′ phosphate of the adenine moiety of NADPH in *P. mirabilis* catalase. Some broad but weak density is found close to the adenine base-binding site. However, the other parts of the NADPH position are void of density and no NADPH could be modelled into the structure (NADPH was never added to the crystallization medium). The S-atom positions differ by only 0.11–0.31 Å between the subunits for this sulfate ion, although the orientations of the O atoms (which are difficult to determine exactly at 2.3 Å resolution) display larger variations. In general, one sulfate oxygen is hydrogen bonded to Lys216, a second to His192 and Arg182 and a third to Arg182. In the absence of NADPH, *P. mirabilis* catalase also has a sulfate ion in this position and the *M. lysodeikticus* enzyme has a sulfate ion in a nearby position, although Arg182 is the only sulfate-binding residue conserved between the two enzymes. All these enzymes were crystallized in excess sulfate and probably do not bind sulfate under physiological conditions. In the absence of supplied NADPH, bovine liver and human erythrocyte catalase bind NADPH to two of the subunits, while structures of *P. mirabilis* catalase are known both with and without NADPH. Although most of the residues that interact with NADPH are conserved in *E. faecalis* and *P. mirabilis* catalases, residue 284 is a glutamine in *E. faecalis* catalase but a histidine in *P. mirabilis* catalase. It has been suggested that a histidine at residue position 284 is required for strong NADPH binding (Gouet *et al.*, 1995).

The water structure at the haem site differs between the catalases. *S. cerevisiae* catalase has an azide ion as an Fe^{3+} ligand on the distal side. One of the azide atoms, N1, coordinates to the iron and the N3 on the other side of the azide hydrogen bonds to the homologues of His54 (His70 in *S. cerevisiae* catalase) and Asn127 (Asn143 in *S. cerevisiae* catalase). *E. faecalis* catalase does not have any water coordinated directly to the iron, but has a water molecule in the position corresponding to the azide N3 in *S. cervisae* catalase. The Fe^{3+}–water distance ranges from 4.0 Å in subunits *B* and *D* to 4.3 Å in subunit *A*. The water–Fe^{3+}–pyrrole N angles vary between 53 and 125° in subunit *A*, with similar values in the other subunits. The situation is similar in *M. lysodeikticus* and human erythrocyte catalase, while *E. coli* and *P. syringae* catalase have water molecules in both the N1 position, coordinated to the iron, and in the N3 position. *P. mirabilis* catalase has a single water between these two positions, corresponding to the position of the azide N2 atom in *S. cerevisiae* catalase. Beef liver catalase, finally, does not have a defined water molecule in any of these positions.

3.4. The Ramachandran outlier Ser196

As in all other catalases of known structure, Ser196 is found in a non-allowed region of the Ramachandran plot, with $\varphi = 69$, $\psi = -64$° (Fig. 1). This residue is conserved in *P. mirabilis*, *M. lysodeikticus*, human erythrocyte and bovine liver catalase, but the corresponding residues in *E. coli* and *P. syringae* catalases are Ile274 and Val219, respectively. In *S. cerevisiae* this residue is a glycine (Gly214) with similar Ramachandran values, which however are not prohibited for a glycine residue.

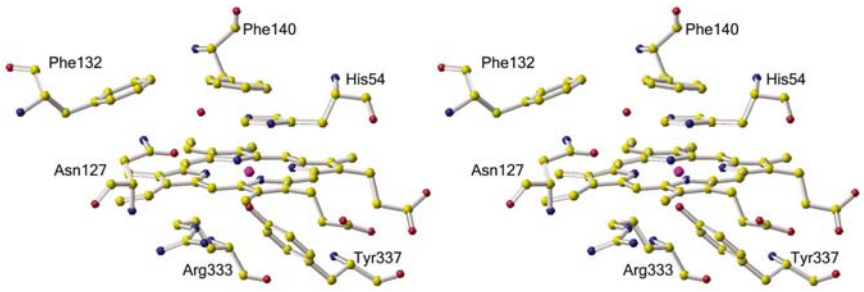

Figure 4
The active site of *E. faecalis* catalase, showing the proximal ligand Tyr337 and its hydrogen-bond partner Arg333, together with a number of planar side-chain groups stacked on or around the haem group.

Ramachandran outliers are known for some groups of hydrolytic enzymes, including the α/β-hydrolases, where the particular conformation has been suggested to direct the electrophilic amino group towards the oxyanion site (Hakansson, 2002), but the situation is very different in the catalases. It has been suggested that the Ser196 side chain in catalase is part of the path of electron transfer from NADPH to the haem (Gouet *et al.*, 1995), but the unfavourable Ramachandran values seem to depend neither on the side chain nor on whether or not the catalase binds NADPH. While we can discuss the structural consequences of the unusual conformation of Ser196, it is not clear why similar interactions could not have been created with a combination of favourable Ramachandran values instead. The most obvious consequences of the unusual conformation of Ser196, situated on strand 5 of the β-barrel, are the introduction of a bulge into the β-strand and the placement of His197 into a position where it can hydrogen bond to Arg333. Furthermore, flipping of the peptide bond between Gly195 and Ser196 breaks up the backbone interactions with the adjacent strands and the carbonyl group of Gly195 is now directed towards the active site. The carbonyl group of Gly195 and the carbonyl group of Val125 both hydrogen bond to a water molecule. This water molecule, which is found in all catalase structures except bovine liver catalase, can therefore be expected to be a strong hydrogen-bond acceptor towards Arg333. Thus, the Arg333 $N^{\eta1}$ atom is stabilized by two hydrogen bonds (from His197 and the water molecule), while $N^{\eta2}$ hydrogen bonds to the proximal ligand Tyr337. Another water molecule is found to be hydrogen bonded to the Tyr213 N atom, replacing the β-strand interaction between Gly195 O and Tyr213 N that was lost as a result of the unusual conformation of Gly195–Ser196. The conformation is stabilized by a hydrogen bond between the carbonyl of Ser196 and the side chain of Lys212. This is a conserved feature of all structurally known catalases, except *S. cerevisiae* catalase, in which the homologous glycine residue is not a Ramachandran outlier. In *S. cerevisiae* catalase, the lysine is replaced by Gln230 and a similar but longer hydrogen bond. The hydrogen-bond network around Ser196 is illustrated schematically in Fig. 5.

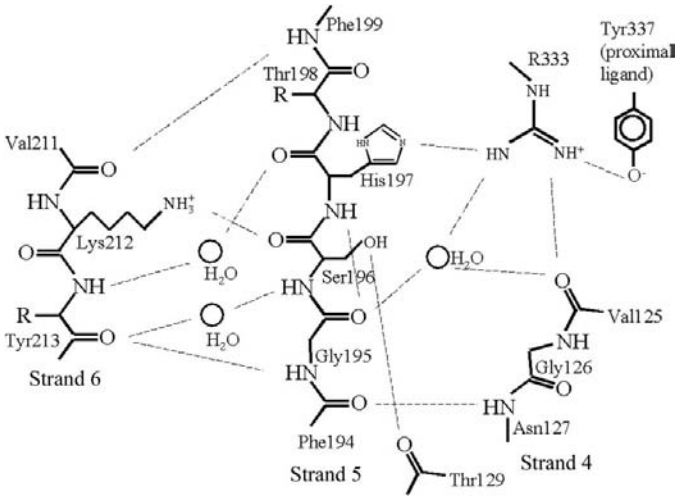

Figure 5
Schematic representation of selected hydrogen bonds around the Ramachandran outlier Ser196. The figure attempts to illustrate how the unusual main-chain conformation affects the interactions between the Phe194–Phe199 segment and the neighbouring β-strand and their hydrogen bonds to the important residue Arg333. The O^{γ} atom of Ser196 makes a van der Waals contact with the haem group and is hydrogen bonded to the backbone carbonyl group of Thr129 and to a water molecule. The unusual conformation of Ser196 contributes two hydrogen bonds to Arg333: one from His197 and another from a water molecule bound to the carbonyl oxygen of Gly195.

Professor Lars Hederstedt is thanked for his support and initiation of the project. The Carlsberg foundation (KOH) and the Swedish Research Council (621-2001-3125 to LH) are thanked for financial support. Myriam Brugna is the recipient of a EU Marie Curie long-term fellowship (contract HPMF-CT-2000-00918). Professor Sine Larsen is thanked for sharing synchrotron beam time. The data were collected at the MAX-lab beamstation I711 under the supervision of Dr Yngve Cerenius.

References

Barynin, V. V., Whittaker, M. M., Antonyuk, S. V., Lamzin, V. S., Harrison, P. M., Artymiuk, P. J. & Whittaker, J. W. (2001). *Structure*, **9**, 725–738.
Bravo, J., Mate, M. J., Schneider, T., Switala, J., Wilson, K., Loewen, P. C. & Fita, I. (1999). *Proteins*, **34**, 155–166.
Bravo, J., Verdaguer, N., Tormo, J., Betzel, C., Switala, J., Loewen, P. C. & Fita, I. (1995). *Structure*, **3**, 491–502.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Carpena, X., Soriano, M., Klotz, M. G., Duckworth, H. W., Donald, L. J., Melik-Adamyan, W., Fita, I. & Loewen, P. C. (2003). *Proteins*, **50**, 423–436.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Corpet, F. (1988). *Nucleic Acids Res.* **16**, 10881–10890.
Fita, I. & Rossmann, M. G. (1985). *J. Mol. Biol.* **185**, 21–37.
Frankenberg, L., Brugna, M. & Hederstedt, L. (2002). *J. Bacteriol.* **184**, 6351–6356.
Gouet, P., Jouve, H. M. & Dideberg, O. (1995). *J. Mol. Biol.* **249**, 933–954.
Hakansson, K. (2002). *Int. J. Biol. Macromol.* **30**, 273–277.
Ippolito, J. A., Alexander, R. S. & Christianson, D. W. (1990). *J. Mol. Biol.* **215**, 457–471.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kirkman, H. N., Rolfo, M., Ferraris, A. M. & Gaetani, G. F. (1999). *J. Biol. Chem.* **274**, 13908–13914.
Ko, T. P., Safo, M. K., Musayev, F. N., Di Salvo, M. L., Wang, C., Wu, S. H. & Abraham, D. J. (2000). *Acta Cryst.* D**56**, 241–245.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACMB Newsl. Protein Crystallogr.* **26**, 27–33.
Loewen, P. (1996). *Gene*, **179**, 39–44.
Mate, 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.
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., Melik-Adamyan, W. R., Grebenko, A. I., Barynin, V. V., Vagin, A. A., Vainshtein, B. K., Dauter, Z. & Wilson, K. S. (1992). *FEBS Lett.* **312**, 127–131.
Murthy, M. R., Reid, T. J. III, Sicignano, A., Tanaka, N. & Rossmann, M. G. (1981). *J. Mol. Biol.* **152**, 465–499.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Paulsen, I. T. *et al.* (2003). *Science*, **299**, 2071–2074.
Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Pugh, S. Y. & Knowles, C. J. (1983). *Arch. Microbiol.* **136**, 60–63.
Putnam, C. D., Arvai, A. S., Bourne, Y. & Tainer, J. A. (2000). *J. Mol. Biol.* **296**, 295–309.
Reid, T. J. III, Murthy, M. R., Sicignano, A., Tanaka, N., Musick, W. D. & Rossmann, M. G. (1981). *Proc. Natl Acad. Sci. USA*, **78**, 4767–4771.
Vainshtein, B. K., Melik-Adamyan, W. R., Barynin, V. V., Vagin, A. A. & Grebenko, A. I. (1981). *Nature (London)*, **293**, 411–412.