PROTEINS: Structure, Function, and Genetics 34:155–166 (1999)

Structure of Catalase HPII From *Escherichia coli* at 1.9 Å Resolution

Jerónimo Bravo,[1] Maria J. Mate,[1] Thomas Schneider,[2] Jack Switala,[3] Keith Wilson,[2,4] Peter C. Loewen,[3]* and Ignacio Fita[1]
[1]*Departamento de Biología Molecular y Celular, CID (C.S.I.C.), Barcelona, Spain*
[2]*European Molecular Biology Laboratory (EMBL), Hamburg, Germany*
[3]*Department of Microbiology, University of Manitoba, Winnipeg, Manitoba, Canada*
[4]*Chemistry Department, University of York, York, United Kingdom*

ABSTRACT **Catalase HPII from *Escherichia coli*, a homotetramer of subunits with 753 residues, is the largest known catalase. The structure of native HPII has been refined at 1.9 Å resolution using X-ray synchrotron data collected from crystals flash-cooled with liquid nitrogen. The crystallographic agreement factors R and R_{free} are respectively 16.6% and 21.0%. The asymmetric unit of the crystal contains a whole molecule that shows accurate 222-point group symmetry. The structure of the central part of the HPII subunit gives a root mean square deviation of 1.5 Å for 477 equivalencies with beef liver catalase. Most of the additional 276 residues of HPII are located in either an extended N-terminal arm or in a C-terminal domain organized with a flavodoxin-like topology. A small number of mostly hydrophilic interactions stabilize the relative orientation between the C-terminal domain and the core of the enzyme. The heme component of HPII is a *cis*-hydroxychlorin γ-spirolactone in an orientation that is flipped 180°with respect to the orientation of the heme found in beef liver catalase. The proximal ligand of the heme is Tyr415 which is joined by a covalent bond between its C^{β} atom and the N^{δ} atom of His392. Over 2,700 well-defined solvent molecules have been identified filling a complex network of cavities and channels formed inside the molecule. Two channels lead close to the distal side heme pocket of each subunit suggesting separate inlet and exhaust functions. The longest channel, that begins in an adjacent subunit, is over 50 Å in length, and the second channel is about 30 Å in length. A third channel reaching the heme proximal side may provide access for the substrate needed to catalyze the heme modification and His-Tyr bond formation. HPII does not bind NADPH and the equivalent region to the NADPH binding pocket of bovine catalase, partially occluded in HPII by residues 585–590, corresponds to the entrance to the second channel. The heme distal pocket contains two solvent molecules, and the one closer to the iron atom appears to exhibit high mobility or low occupancy compatible with weak coordination. Proteins 1999;34:155–166.**

Key words: bacterial catalase; heme-d; cryocrystallography; flavodoxin domain; channels

INTRODUCTION

Catalase (hydrogen peroxide:hydrogen peroxide oxidoreductase, EC:1.11.1.6) which is present in most eukaryotic and aerobic prokaryotic organisms can utilize hydrogen peroxide as both electron donor and electron acceptor in a dismutation reaction that yields water and molecular oxygen: $2\ H_2O_2 \rightarrow 2\ H_2O + O_2$.

To date, the structures of six heme-containing catalases have been determined including enzymes from bovine liver (BLC),[1, 2] the fungus *Penicillium vitale* (PVC),[3, 4] the yeast *Saccharomyces cerevisiae* (CatA),[5] and the bacteria *Micrococcus lysodeikticus* (MLC),[6] *Proteus mirabilis* (PMC),[7] and *Escherichia coli* (HPII).[8] All these enzymes have been found to be tetrameric and contain a core structure of about 470 residues with a closely related conformation.

E. coli catalase HPII, with 753 residues per subunit, is the largest known catalase and contains the additional sequence in a N-terminal extension and a C-terminal domain with a flavodoxin topology. Despite possessing an extra domain with a topology that is often associated with nucleotide binding, the larger size catalases, HPII and PVC, do not bind NADPH that has been found in a number of small subunit catalases, including BLC, PMC, and CatA. HPII and PVC also differ from other catalases in that they contain a heme *d*-like moiety rather than heme *b*. HPII, and presumably PVC, initially bind heme *b* and then catalyze its conversion to heme *d* utilizing H_2O_2 as substrate.[9] A recent analysis of the heme environment, done simultaneously for the refined crystal structures of the *P. vitale* and *E. coli* enzymes, revealed that the heme *d* cofactor in both enzymes is a spirolactone structure with a proximal location of the *cis*-oriented oxygens on ring III of the heme that is flipped in relation to the heme *b* in other

Grant sponsor: Direccion General de Investigacion Ciencia Y Technologia; Grant number: PB95–0218; Grant sponsor: The European Union through the HCMP to Large Installations Project; Grant number: CHGE-CT93–0040; Grant sponsor: The Natural Sciences and Engineering Research Council of Canada; Grant number: OGP9600.
*Correspondence to: Peter Loewen, Department of Microbiology, University of Manitoba, Winnipeg, MB R3T 2N2 Canada. E-mail: peter_loewen@umanitoba.ca
Received 2 June 1998; Accepted 22 September 1998

catalases.[10] In addition, an unique covalent bond between the C^{β} of the essential Tyr415 and the N^{δ} of His392 has been noted and a mechanism relating its formation to heme modification has been proposed.[11]

The structure of HPII determined at 2.8 Å resolution has been reported,[8] but the solvent structure was not included in the molecular model, the heme groups were not well-defined, and strict non-crystallographic symmetry was maintained among the four subunits in the asymmetric unit of the crystal. Furthermore, the accuracy of the HPII model was limited by both the quality and the resolution of the X-ray data. The present paper reports on the refinement and analysis of the HPII structure at 1.9 Å resolution using X-ray synchrotron data obtained from a single crystal flash-cooled with liquid nitrogen.

METHODS

Enzyme Purification

Cultures of *E. coli* strain UM255 *pro leu rpsL hsdMhsdR endl lacY katG2 katE::*Tn*10 recA*[12] transformed with plasmids pAMkatE72,[13] pH128A,[9] or pN201H[9] encoding HPII or its mutant variants His128Ala or Asn201His, respectively, were grown in Luria broth containing 10 g/L tryptone, 5 g/L yeast extract and 5 g/L NaCl. Growth of the mutant variants was for 22 h at 28°C and of the wild type HPII was for 16 h at 37°C with shaking. Cells were harvested and HPII was isolated as previously described.[21] DEAE cellulose DE52 (Whatman) replaced DEAE Sephadex.

Crystallization and Data Collection

Crystals of HPII, improved in size and diffraction capability with respect to the ones previously reported,[8] were obtained at 22° with a protein concentration of 15 mg/ml using the hanging drop vapor diffusion method over a reservoir solution containing 15–17% PEG 3350 (Carbowax), 1.6–1.7 M LiCl (Baker) and 0.1 M Tris- HCl pH 9.0. High-resolution data were collected at the EMBL DESY Hamburg, using an imaging plate detector (MarResearch 30-cm diameter, Hamburg, Germany). One single cryo-cooled crystal was used for data collection, transferring it to mother liquor containing 30% PEG 3350, and flash cooled directly into the nitrogen cryo-stream. Diffraction data were autoindexed and integrated using program DENZO, and merged using SCALEPACK[15] (Table I-A). Crystals belong to space group $P2_1$ with unit cell parameters a = 93.04 Å, b = 132.34 Å, c = 121.20 Å, and β = 109.3°. There is a reduction of about 3.8% in the unit cell volume with respect to crystals at room temperature. As previously reported, the crystals contain one tetramer per asymmetric unit and the crystal solvent content is approximately 40%.[8]

Refinement was carried out with the program XPLOR[16] using standard protocols. The starting model had been initially refined to 2.8 Å resolution maintaining strict non-crystallographic restraints.[8] The coordinates were adapted to the cryo-collected unit cell fractionalizing them using the option balloon of XPLOR. Six and one-half percent (6.5%) of the data were reserved for R_{free} monitoring. Non-crystallographic symmetry restraints were released at 2.5 Å and ordered water molecules were modeled either automatically or manually with TURBO-FRODO[17] and O.[18] Waters introduced corresponded to the strongest peaks found in difference Fourier maps that could make at least one hydrogen bond with atoms already in the model. The bulk solvent correction was only used in the later stages of refinement. Refinement was continued until no significant changes could be added to the molecular model (Table I-B). Calculation of channel accessibility using a molecular radius of 1.4 Å was carried out using VOIDOO.[19]

TABLE I. Data Collection and Structural Refinement Statistics

A. Data collection statistics	
Unique	213692
$\langle I/\sigma(I)\rangle$	9.21 (2.85)[a]
R_{merge} (%)[b]	8.9 (26.9)
Resolution (Å)	1.9
Completeness (%)	98.0 (97.0)
B. Structural refinement statistics	
Resolution range (Å)	18–1.9
Number of reflections (F > 0)	213117
R_{cryst} (%)[c]	16.6 (19.1)
R_{free} (%)[d]	21.0 (24.7)
Number of non-hydrogen atoms	
Protein	22984
Water	2766
Heme	176
R_{msd} from ideality	
Bond lengths (Å)	0.012
Bond angles (°)	2.6
Dihedrals (°)	24.4
Improper (°)	1.3
Average B factor (Å^2)	
Mainchain	8.8
Sidechain	9.7
Water	17.7

[a]Values in parentheses correspond to the highest resolution shell (1.94–1.90 Å).
[b]$R_{merge} = |I - \langle I\rangle|/\langle I\rangle$.
[c]$R_{cryst} = \| |F_{obs}| - |F_{calc}| \| / |F_{obs}|$.
[d]R_{free} is as for R_{cryst} but calculated for a test set comprising randomly chosen reflections not used in the refinement.

RESULTS AND DISCUSSION

Quality of the Refined HPII Structure

The crystal asymmetric unit of HPII contains four subunits with 3,012 amino acids.[8] The HPII model obtained in the present analysis, with the four subunits refined independently, corresponds to the coordinates of 22,984 protein atoms (2,908 amino acids), 2,766 solvent molecules and four heme *d* groups. The 26 N-terminal residues are disordered in the four subunits and have not been included in the final model which has crystallographic agreement R and R_{free} factors of 16.6% and 21.0%, respectively, for 213,117 reflections in the resolution shell between 1.9 Å and 18.0 Å. The averaged root mean square deviation obtained from the superimposition of all the C^{α} atoms of the four molecular subunits is 0.22 Å. This value decreases to 0.12 Å when 200 C^{α} atoms from the HPII C-terminal extension are not included in the superimposi-

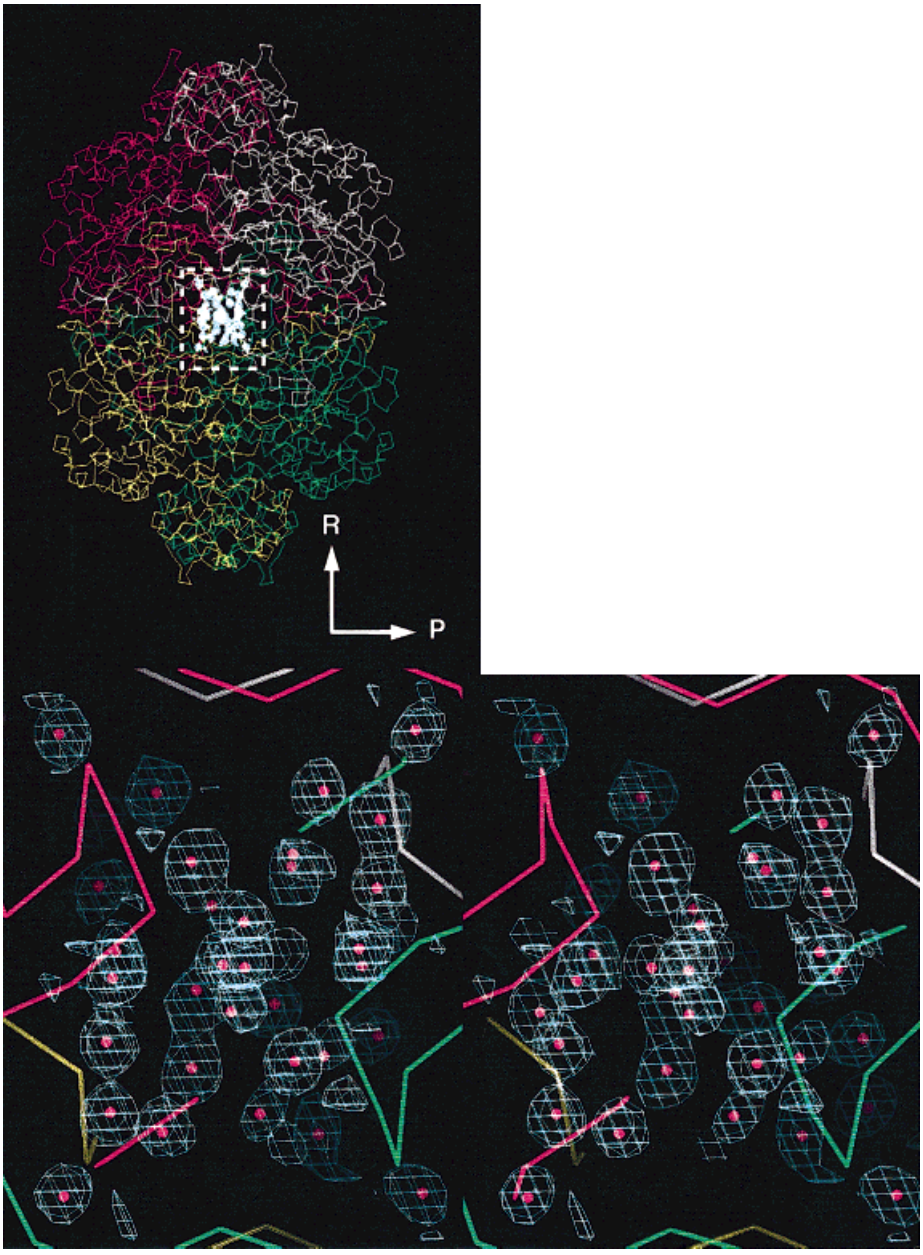

Fig. 1. Stereo view $2F_o$ - F_c electron density map showing the organization of water molecules in the large cavity in the center of catalase HPII. The location within the tetramer that is the focus of the stereo views is indicated in the box in the top part of the figure. Although over 2,700 water molecules were introduced into the structure independently, their organization follows the 222 molecular symmetry accurately, providing an unbiased control of the quality of the present model.

tion, and it rises to 0.37 Å when the C-terminal residues are superimposed alone. These values provide an upper limit to the error of the atomic positions determined and confirm that the molecule is organized with accurate 222 molecular symmetry presenting the largest deviations in the C-terminal region of HPII.

The final unaveraged electron-density maps show clear continuity in all four subunits over the complete length from Asp27 to the carboxy terminal Ala753 with the exception of a break in the main chain density in the region of Ile710 to Gln713 in one of the subunits. The quality of these final maps allowed a definition of the configuration of the modified heme *d* group[10] and the identification of a novel type of covalent bond between the side chains of His392 and the essential residue Tyr415.[11] The quality of the final maps is reflected in the accuracy of the solvent structure determined that also reflects the molecular symmetry. Most cavities within the HPII enzyme are filled with well-resolved solvent molecules including the large cavity located in the center of the molecular tetramer (Fig. 1). The detailed organization of solvent in HPII allows a deeper analysis of the molecular interactions providing

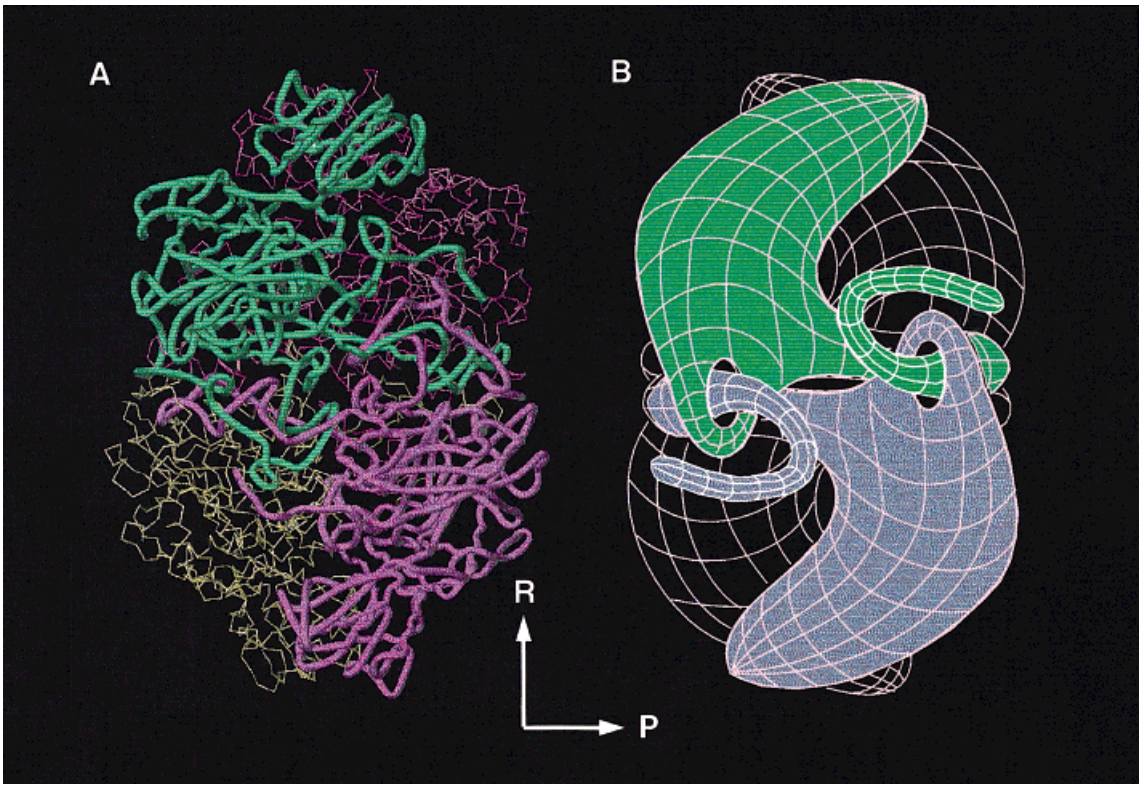

Fig. 2. The HPII tetramer is shown with each of the four subunits represented in a different color for clarity. The extent and complexity of the intersubunit interactions (**A**) appears to require a close dependency between subunit folding and subunit association. The schematic representation of HPII in (**B**) highlights the interweaving or knot-like structure between Q-related subunits by insertion of the amino terminal arm of one subunit through the loop formed by residues from the wrapping domain of the opposite subunit.

information about possible catalytic mechanisms and about the movement of substrate and products.

Both the main-chain torsional angles and their differences among crystallographically independent subunits are good indicators of the quality of the model.[20] Ramachandran plots (data not shown) confirm that only two residues, Ile274 and His739, are outside the energetically favorable regions in each of the four subunits.[8] Differences between C^{α} dihedral angles of different subunits larger than 25° are found only around a small number of residues, mostly from the C-terminal domain, which are involved in crystal contacts in at least one subunit (Gly596, Glu610, Gly645, Asp712, and Gly726). Atomic disorder is similar within each of the four crystallographically independent subunits, as derived from the atomic isotropic temperature factors B (data not shown). As in the averaged HPII model, the highest mobility is found in the exposed regions, particularly in the amino terminal region where the 26 N-terminal residues could not be positioned in any of the four crystallographically independent subunits. However, the strong fluctuations in B factors reported for the carboxy terminal domain of the averaged HPII structure[8] have mostly disappeared in the present unrestrained model, confirming that the disorder observed in this region was partly due to conformational differences among the four subunits. The temperature factor distribution still shows the largest deviations among the carboxy terminal domains of different subunits, probably because most crystal contacts are located in these regions.

His449 is the only internal residue not involved in crystal contacts whose disposition in different subunits appears to depart from molecular symmetry. His449 is located very close to a molecular two-fold axis and interacts with the His449 in the related subunit adopting two alternative configurations.

Structure of HPII

The general folding pattern of the present HPII model (Fig. 2A) retains the essential features of the averaged structure determined at 2.8 Å[8] but shows many changes in details. HPII appears as a compact molecule with maximum dimensions of about 100 Å, 80 Å, and 150 Å along the three molecular binary axes referred to as P, Q, and R respectively,[1] and a waist of about 80 Å in diameter in the R = 0 plane. The HPII subunit conformation (Fig. 3A) can be described, following the nomenclature used for BLC[1] and PVC,[21] as being composed of five different regions, including (1) the amino terminal arm (residues Met1 to Val127), (2) an antiparallel eight-stranded β-barrel (residues His128 to Ala390), (3) an extended wrapping loop (residues Phe391 to Tyr504), (4) a helical domain (residues Tyr505 to Ile564), and (5) a carboxy-terminal, flavodoxin-like domain (residues Gly600 to Ala753). Residues Glu565 to Lys599 form an extended hinge that joins the helical

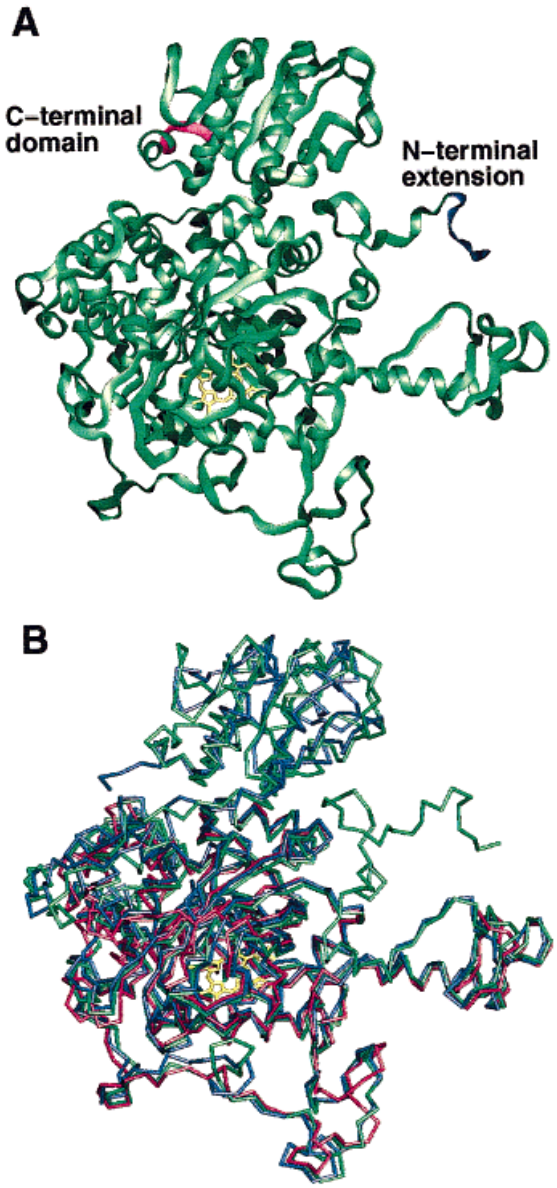

Fig. 3 (**A**) The HPII subunit with the C-terminal domain and N-terminal extension indicated. The N-terminal five residues are colored blue and the C-terminal five residues are colored red to facilitate identification. The heme d is colored yellow. (**B**) Superimposition of the C^{α} tracings of one subunit of BLC (red), PVC (blue) and HPII (green). The conservation of structure among enzymes from mammalian, fungal, and bacterial sources is evident. The main differences among the structures are found in the amino terminal arms and in the C-terminal domain (absent in BLC). The poor superimposition between the C-terminal domains of PVC and HPII is mostly due to differences in the location of this domain relative to the core of the enzymes.

and carboxy-terminal domains (Fig. 4). The extra C-terminal domain increases the maximum dimension with respect to BLC only along the R axis.

The HPII model (Fig. 4) includes 18 helices and 16 β-sheets. Helices range in size from one turn (α8 and α14) to more than five turns (α11 and α18). Despite lacking proline residues the essential helix, α9, presents a disrupted hydrogen bonding pattern in the vicinity of the proximal ligand, Tyr415, where the C^{β} atom forms a covalent bond with the side chain of His 392.[11] Three proline residues, Pro230, Pro462, and Pro470, exist in the less stable *cis* conformation, but, of these three, only Pro462, in the middle of the wrapping loop, is well preserved, both in sequence and conformation, among the catalases.

An unique feature of the quaternary organization of catalases is the slipping of the N-terminal arm of the reference subunit through a loop composed of residues from the wrapping loop of the Q-related subunit (Fig. 2B). Although this is a common feature in all catalase structures, it is particularly striking in HPII, where over 90 residues of the N-terminal arm are inserted through the narrow loop. The extent of the intricate pattern of intersubunit interactions, repeated four times in each tetramer, might be responsible for the enhanced stability of HPII at extreme pHs, at 70°C and in SDS, or urea.[22] By contrast, the quaternary structure of PMC, which presents only a short N-terminal extension,[7] is fragile at high dilutions.[23]

A quantitative analysis, in terms of the atomic contact areas, done with the program GRASP,[24] with a probe atom radius of 1.4 Å, indicates that the subunit surface between P-related dimers is 3,050 $Å^2$, about half of the Q- and R-related dimer contact surface of 5,520 and 5,494 $Å^2$ respectively. This was also the case in BLC, although the contact surfaces reported were smaller than in HPII. The contact area between the C-terminal domains of R-related subunits is only 360 $Å^2$. Most of the inter-subunit interactions generated by the molecular P, Q, and R two-fold axes involve residues from the amino terminal arm and from the wrapping loop. As in BLC, the amino terminal arm is dominant in the interactions of the R-related dimers, while the wrapping loop contributes largely to the interactions between Q-related dimers. Both the N-terminal and the wrapping loop, with a small content of secondary elements and a non-globular shape, are likely to experience important rearrangements during the oligomerization of subunits.

The superimposition of C^{α} atoms of HPII with BLC, PMC, and PVC, reveals 477, 471, and 632 equivalencies with root mean square deviations of 1.5 Å, 1.5Å, and 1.9 Å, respectively (Fig. 3B). The larger number of equivalencies between HPII and PVC is a result not only of the presence of a C-terminal domain in both molecules, but also of the close structural similarities of the N-terminal regions. In fact, when the C-terminal domains are not considered, comparison of HPII with PVC shows 525 equivalencies with a root mean square deviation of only 1.1 Å. The structural equivalencies of the C^{α} atoms used to align the catalase sequences reveals that structural similarity among catalases is more extensive than sequence similarity.

N-Terminal Extension and the C-Terminal Domain Structure

The N-terminal 26 residues are not visible in any of the four structurally-independent HPII subunits determined, although N-terminal protein sequence analysis has confirmed the presence of the N-terminal residues in the molecule. Therefore, this region must be sufficiently flex-

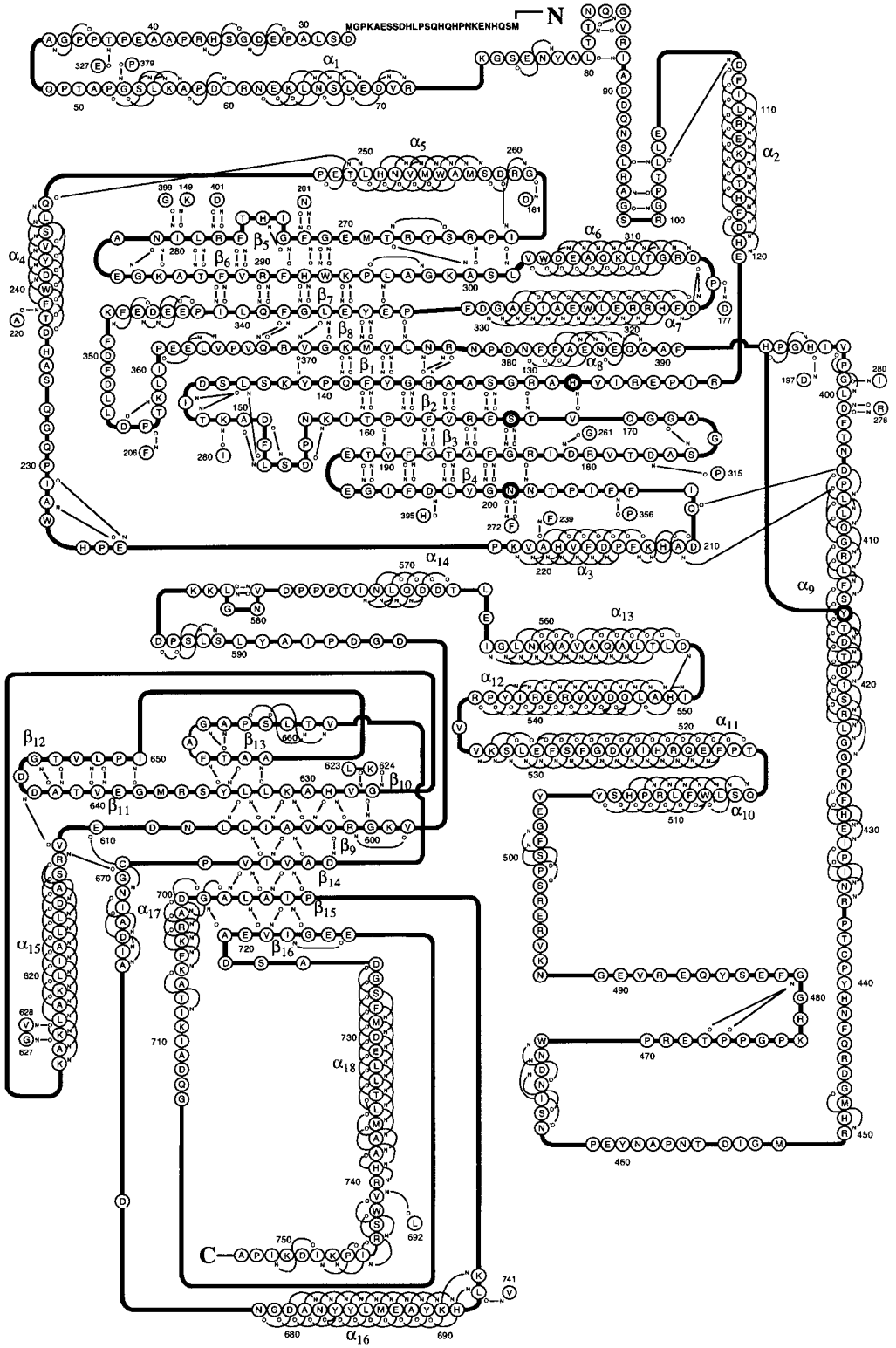

Fig. 4. Main-chain hydrogen-bonding network for one HPII subunit. The nomenclature of the secondary structural elements is also indicated. Essential or important catalytic residues are indicated with thicker circles including the important residue Ser167, that interacts with the essential His128. The covalent bond between residues His392 and Tyr415 is also indicated.

TABLE II. Intra-Subunit Contacts Between the Core of HPII and the N-Terminal Extension or the C-Terminal Doman

Core residue and atom	N/C-terminal residue and atom	Distance (Å)	Type of interaction
A. N-terminal extension			
G136 O	N62 $N^{\delta 2}$	3.02	H bond
E327 O	T43 N	3.23	H bond
E327 $O^{\epsilon 1}$	A51 N	3.02	H bond
E327 $O^{\epsilon 2}$	A51 O	3.27	H-bond
P379 O	G53 N	2.96	H bond
D380 $O^{\delta 2}$	S54 N	2.90	H bond
D380 $O^{\delta 2}$	S54 O^{γ}	2.75	H bond
D380 $O^{\delta 1}$	R72 N^{ϵ}	2.89	H-bond
P432 O	K64 N^{ζ}	2.71	H bond
I433 O	K64 N^{ζ}	3.26	H bond
R435 O	K64 N^{ζ}	2.89	H bond
R435 $N^{\eta 1}$	D70 $O^{\delta 1}$	3.02	Ionic
R435 $N^{\eta 2}$	D70 $O^{\delta 2}$	2.79	Ionic
R435 $N^{\eta 2}$	S67 O	3.15	H-bond
B. C-terminal domain			
R264 $N^{\eta 2}$	D663 $O^{\delta 1}$	2.84	Ionic
R264 N^{ϵ}	D663 $O^{\delta 2}$	2.97	Ionic
S301 O	H739 N^{ϵ}	2.81	H bond
D305 $O^{\delta 1}$	K690 N^{ζ}	2.74	Ionic
D305 $O^{\delta 1}$	H691 $N^{\epsilon 2}$	2.91	Ionic
E306 $O^{\epsilon 1}$	K693 N^{ζ}	2.86	Ionic
E306 $O^{\epsilon 2}$	R740 $N^{\eta 1}$	2.80	Ionic
E306 $O^{\epsilon 1}$	R740 $N^{\eta 2}$	2.89	Ionic
K309 N^{ζ}	Y683 O^{η}	3.00	H bond
K309 N^{ζ}	E687 $O^{\epsilon 2}$	2.81	Ionic
D330 $O^{\delta 1}$	H629 $N^{\epsilon 2}$	2.90	Ionic
F518	W742		Hydrophobic

ible to adopt a number of different conformations. The next 40 residues exist in an extended structure with only a short α-helical section, α1 (Fig. 4). There is no biochemical function ascribed to the N-terminal extension, but it has a very high proline content (8 of 28 residues between Pro31 and Pro58), supporting the idea that it may contribute to the stabilization of the quaternary structure[25, 26] by supplementing the limited number of intramolecular contacts between the N-terminal extension and the core of the enzyme (Table II). It is this part of the subunit that is threaded through the loop formed by the wrapping loop (Fig. 2).

Despite the structural analogy between the C-terminal domains of PVC and HPII, the HPII domain is modified around the anti-parallel strands β11 and β12 that are absent in PVC. This domain likely folds independently of the core of the enzyme as only 11 hydrogen or ionic bonds and one hydrophobic interaction, the stacking between Phe518 and Tyr742, are formed between the two regions (Table II and Fig. 5). The side chains involved in direct interactions are largely grouped in one area of the interface while the remainder of the interface is filled with a matrix of well-defined solvent molecules that mediate the interactions between the two parts of the enzyme (Fig. 5). Despite the small number of direct interactions between the core and the C-terminal domain, removal of the domain destabilizes HPII to the extent that no active enzyme accumulates in the cell.[26] The essential role of the domain may therefore lie in facilitating the folding and oligomerization process. The flavodoxin-like topology of the C-terminal domain has lead to conjecture that it might be associated with the binding of nucleotides, but no experimental evidence in support of nucleotide binding in this domain has been reported in either PVC or HPII. Certainly there is no evidence of a nucleotide in the electron density maps of either enzyme. As will be discussed below, the water-filled interface between the C-terminal domain and the core forms part of the channel leading to the active site in the adjacent subunit and may have a role in controlling substrate access.

Ramachandran plots have shown that His739 is one of two residues that has a conformation lying outside of the range of energetically favorable torsional angles. This unusual conformation may be explained in part by the participation of the imidazole ring of His739 in a hydrogen bond network including the side chain of Arg601, the carboxyl oxygen of Ser301, and two water molecules (Fig. 6 and Table II). The location of His739 in the interface between the C-terminal domain and the enzyme core may explain its unfavorable torsional angles.

Channels Leading to the Active Site

Two major channels can be defined in each subunit that provide access to the deeply-buried active sites of HPII (Fig. 7). One channel that leads to the heme distal pocket is related to the major channel defined in BLC[27] and PMC,[7] but it is 20 Å longer (55 Å in length) because of the involvement of the C-terminal domain of the R-related subunit. An entering substrate would pass along the interface between the core and the C-terminal domain of the R-related subunit until it reaches a bend where it enters the part of the channel that is equivalent to the channel in BLC and PMC (Fig. 7A). The integration of the C-terminal domain into the channel raises the possibility that the structure may have a role in some other, as yet undefined, function in which the channel links a catalytic site on the flavodoxin-like C-terminal domain with the main active site of the catalase. This extended channel may also be a major factor contributing to the high selectivity towards hydrogen peroxide observed in HPII as compared to small subunit catalases. The portion of the channel that opens into the heme distal side pocket has a small diameter and is lined mostly with hydrophobic residues which would further limit the size and shape of compounds reaching the active site and might also help to pre-orient the substrate. In fact this lower part of the channel appears interrupted around Val169 indicating that it is at the lower limit of accessibility in the representation (Fig 7A). Very small movements in side chains that have been observed in mutant variants of HPII are sufficient to open the channel at this point (unpublished data). The structure of this channel is further complicated by an interconnection with the same channel in the R-related subunit, potentially allowing the exchange of substrate

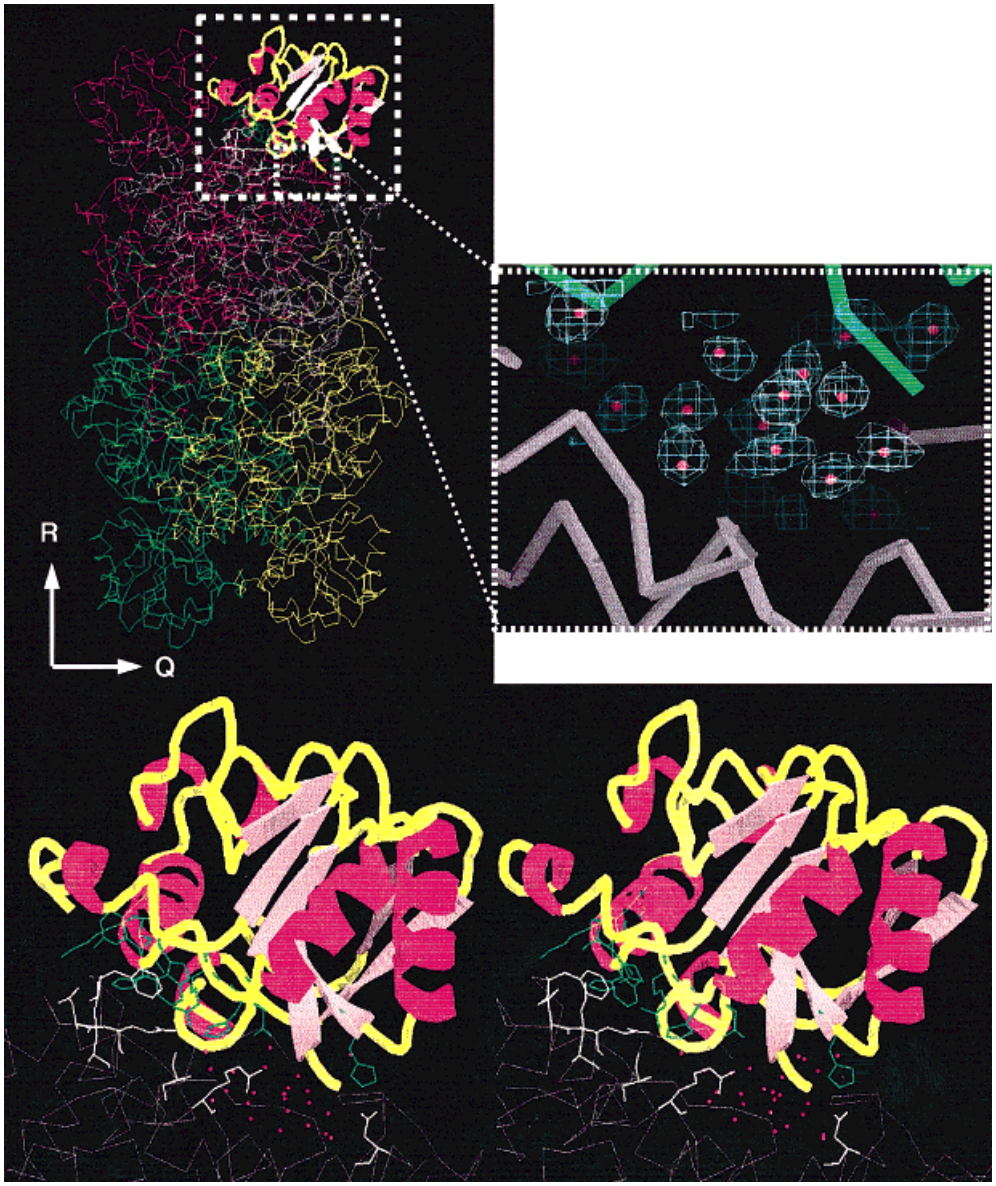

Fig. 5. Structure of the C-terminal domain of HPII including interactions between this domain and the core of the same subunit. The lower panel shows a stereo view of the C-terminal domain. The residues involved in the 12 interactions between the domain and the core (Table II) are also shown. The side panel displays most of the solvent molecules that fill the interface between the C-terminal domain and the core of HPII. The corresponding ($2F_o$ - F_c) electron density map is also shown. The location of the two regions within the tetramer are indicated in boxes in the upper panel.

molecules between them, and by a branch that leads to a cavity situated close to the heme propionate groups.

A second channel, approximately 30 Å in length, that is structurally related to the minor channel noted in BLC adjacent to the NADPH binding pocket, can be proposed for HPII. The route of this second channel starts with a large funnel-shaped opening near the segment containing residues 590 to 595 and leads past an isolated cavity near Arg260 (Fig. 7B) to the vicinity of Asn201 close to the distal side heme pocket. The existence of two channels that provide access to the distal side active site suggests an inlet/outlet model in which separate channels are used for substrate access and product exhaust. Such a structure would avoid the interference of substrate and products moving in opposite directions along a single channel, and may, in part, explain the very fast turn-over rates of catalases (10^5 to 10^6 per second). An extended cavity in the vicinity of the essential Tyr415 on the proximal side of the heme can also be proposed to be a third channel, although it presents some discontinuities and is not as clearly defined as the other two. Access for substrate that is required to catalyze the heme modification and His-Tyr joining reactions[11] might be provided by this path.

The cavity near Arg260 (Fig. 7B) arises in part because of the absence of NADPH[28] which is present in BLC.[29, 30] A comparison of the region of HPII equivalent to the NADPH

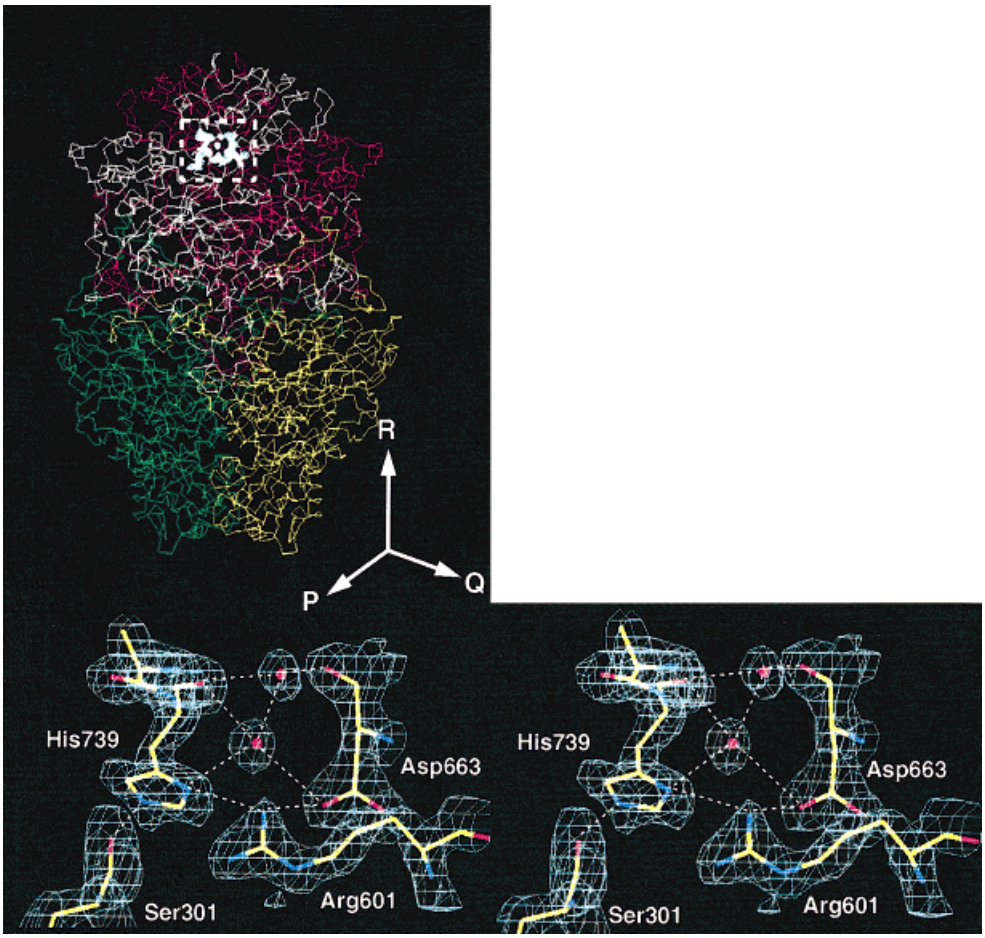

Fig. 6. The lower panel shows a stereo view of the ($2F_o$ - F_c) electron density map in the vicinity of His739, which has uncommon main chain conformational angles (see text). The molecular model and the intricate network of hydrogen bonds involving His739, Ser301, Arg601 and two water molecules are also shown. The location of this region within the tetramer, in the interface between the C-terminal domain and the core of the enzyme, is indicated with a box in the upper panel.

binding pocket of BLC reveals that two key residues interacting with NADPH in BLC, Arg202 and Lys236, are conserved as Arg260 and Lys294 in HPII, but two other residues in BLC, Asp212 and His304 are changed to Glu270 and Glu362 in HPII. These changes might be enough to explain the lack of NADPH binding in HPII, but of even greater significance is the intrusion of a segment (residues 585 to 590) of the extended hinge region of HPII into the cavity. Imposition of the NADPH from BLC into the corresponding space in HPII results in the NADPH occupying some of the same space as the 585–590 segment (Fig. 8). This steric overlap would clearly prevent NADPH binding. Mutants of HPII, including Glu270Asp and Glu362His which change HPII residues to the BLC sequence, did not promote nucleotide binding[26] confirming the dominant importance of the hinge region in preventing NADPH binding.

Active Site Environment

Both the position of the heme and its environment are strongly conserved among catalases, including three amino acids which appear to be directly involved in catalysis,[31] a histidine and an asparagine in the heme distal side pocket and a tyrosine in the proximal side. In HPII these residues are His128, Asn201, and Tyr415 (Fig. 9). A serine residue (Ser167 in HPII) is hydrogen bonded to the N^{δ} of the essential histidine and might facilitate the enzymatic mechanism. In spite of these similarities with other catalases, three important differences have been observed in the heme environment of HPII . (1) The heme is modified to *cis*-hydroxychlorin γ-spirolactone heme *d* in a self-catalyzed reaction.[9, 10, 32] (2) The heme is inverted with respect to the heme in BLC.[8] (3) The C^{β} of Tyr415 is covalently bound to N^{δ} of His392, a modification so far unique to HPII.[11]

The imidazole ring of the essential histidine, His128, is within 10° of being parallel to the plane of the heme group and is situated at a mean distance of 3.5 Å above the pyrrole ring IV of the heme *d*. This stacking of the histidine over the heme group is different from the location of the distal imidazole groups in globins, cytochromes and peroxidases.[33] The positions of the N^{δ} and N^{ϵ} atoms of the His128 imidazole ring were assigned on the basis of possible hydrogen bonding of N^{δ} with the hydroxyl group of the

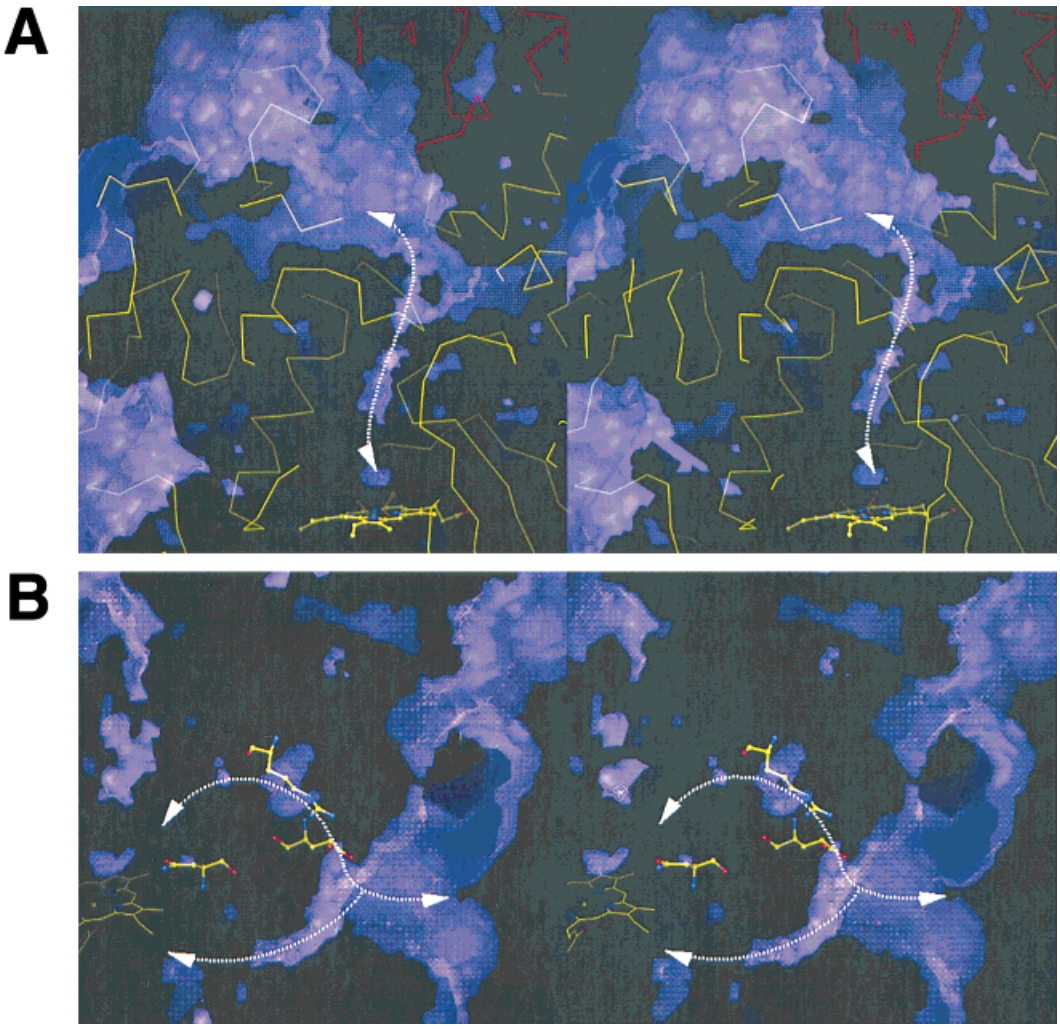

Fig. 7. Stereo views showing the location of proposed channels leading to the active site of HPII. Solvent accessibility to demonstrate the channel locations was calculated using the program VOIDOO. (**A**) Channel leading perpendicularly to the distal side heme pocket. The channel appears to be interrupted at two locations indicating that the channel is at the lower limit for substrate accessibility. (**B**) A bifurcated channel leading laterally into the distal side heme pocket. The upper branch of the channel enters the active site near Asn201 (shown). Arg260 and Glu270, which form an ionic interaction near the channel, are shown for reference.

neighboring Ser167, and the requirement of accessibility of the atom closest to the heme Fe. Besides helping to select the orientation of the imidazole ring, the interaction with Ser167 should also favor one of the two tautomeric forms of the histidine.[33] N^{δ} of His128 is also at a hydrogen-bond distance from the carbonyl oxygen atom of Thr168. This alternate hydrogen bond is not coplanar with the imidazole ring and it may facilitate small reorientations of the His128 ring, such as those observed in the complex with azide (unpublished data). The conservation among catalases of the essential His128 in a conformation of left-handed helix may be due, at least in part, to the strong restraints imposed by the close interactions observed with both Arg165 and Thr168 that are highly conserved in the 81 catalase sequences currently available.

The positions of the N^{δ} and O^{δ} atoms of the essential Asn147 in BLC were chosen on the basis of the proposed mechanism for interactions with the substrate.[27] These positions appear to be confirmed in the refined structures of MLC[33] and in this work for Asn201 of HPII. However, the options for hydrogen bonding of the O^{δ} of Asn201 are not satisfactory, suggesting that the side chain may be in an energetically unfavorable position, providing an explanation for its catalytic role. A number of catalase HPII mutants for this residue show a large diversity of catalytic activities[9] and are currently being structurally investigated.

Water molecules were not detected in the heme distal pocket in the electron density map of BLC, but are defined in the crystal structures of PMC at or very close the positions modeled for BLC.[7] In HPII, the electron density clearly supports the presence of two water molecules in the distal side heme pocket, but with some disorder. Both water molecules can form a hydrogen bond with N^{δ} of His128, and the second water molecule can also be hydrogen bonded to the N^{δ} of Asn201 (Fig. 9). There does not appear to be a direct coordination of water with the iron because the closest water molecule is about 3.0 Å distant.

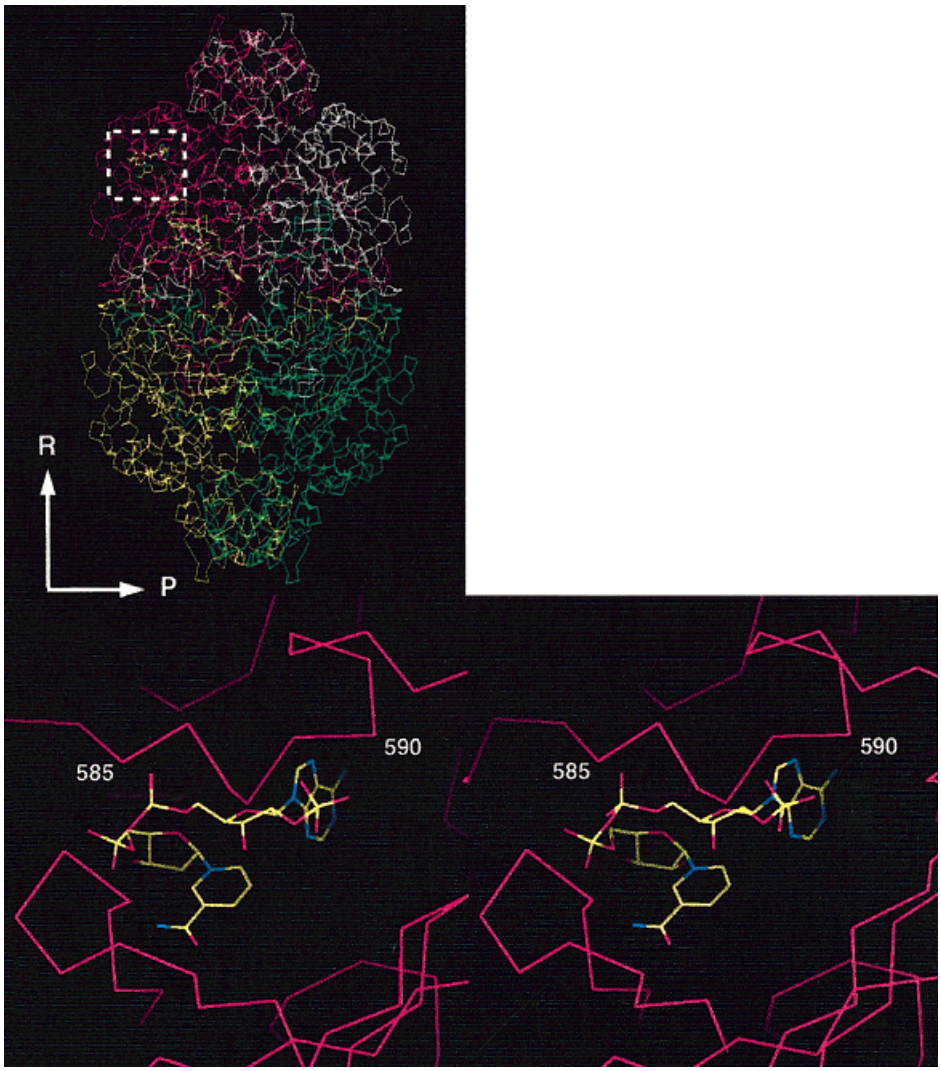

Fig. 8. Stereo view of the superposition of the NADPH molecule found in BLC onto HPII which does not bind NADPH. Only the C^{α} trace of HPII and the NADPH from BLC are shown for clarity. The HPII segment between residues 585 and 590 partially occludes the NADPH binding pocket found in the BLC structure, hindering the binding of nucleotides to HPII (see the text). The portion of the tetramer that is the focus of the stereo view in the lower panels is indicated by the box in the upper panel.

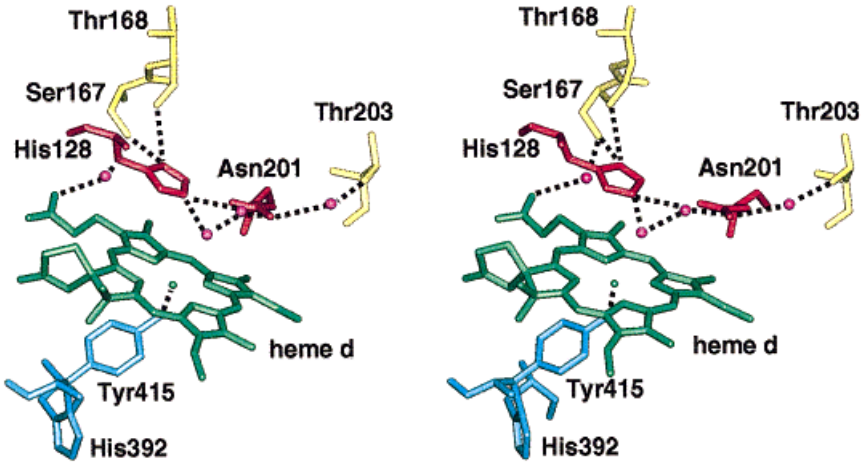

Fig. 9. Diagrammatic representation showing the key elements in the heme environment of HPII. The essential or important residues, His128, Ser167, Asn201, Thr203, His392, and Tyr415 are indicated. Solvent molecules bound to these residues are also shown. The water molecule closest to the iron appears to exhibit high mobility, probably with two main dispositions, compatible with a weak coordination. The heme component is cis-hydroxychlorin ϵ-spirolactone in an orientation that is flipped 180° with respect to the orientation of the heme in BLC.

This water molecule appears disordered or with partial occupancy which may indicate that it can easily exchange position, and this is consistent with the concept that they have to be removed in order to allow rapid replacement by the substrate peroxide. A third water molecule, hydrogen-bonded to the side chains of Asn201, Thr203, Asn252 and to the main chain oxygen of Ile205 might also be considered as being in the heme distal pocket (Fig. 9). The locations of these waters is similar to those found in PMC and are consistent with the model proposed for H_2O_2 binding in BLC involving the side chains of His128 and Asn201 and the heme iron. H_2O_2 would effectively replace the two waters bound between His128 and Asn201 allowing a direct interaction with the heme iron. Further work characterizing the structures of several mutant variants of HPII should lead to a better understanding of water and substrate flux in the region of the HPII active site.

ACKNOWLEDGEMENTS

Work in Barcelona was funded by the Direccion General de Investigacion Ciencia Y Technologia (PB95–0218) and by the European Union through the HCMP to Large Installations Project (contract CHGE-CT93–0040). Work in Winnipeg was supported by grant OGP9600 from the Natural Sciences and Engineering Research Council of Canada. Thanks are due to Dr. W. Melik-Adamyan and to Dr. F. Koller for helpful discussions. The HPII refined coordinates have been deposited in the Protein Data Bank and can also be obtained from the authors.

REFERENCES

1. Murthy MRN, Reid TJ, Sicignano A, Tanaka N, Rossmann, M. Structure of beef liver catalase. J Mol Biol 1981;152:465–499.
2. Fita I, Silva AM, Murthy MRN, Rossmann MG. The refined structure of beef liver catalase at 2.5 Å resolution. Acta Crystallogr B 1986;42:497–515.
3. Vainshtein BK, Melik-Adamyan WR, Barynin VV, Vagin AA, Grebenko AI. Three-dimensional structure of the enzyme catalase. Nature 1981;293:411–412.
4. Vainshtein BK, Melik-Adamyan WR, Barynin VV, et al. Three-dimensional structure of catalase from *Penicillium vitale* at 2.0 Å resolution. J Mol Biol 1986;188:49–61.
5. Berthet S, Nykyri LM, Bravo J, et al. Crystallization and preliminary structural analysis of catalase A from *Saccharomyces cerevisiae*. Protein Sci. 1997;6:481–483.
6. Murshudov GN, Melik-Adamyan WR, Grebenko AI, et al. Three-dimensional structure of catalase from *Micrococcus lysodeikticus* at 1.5 Å resolution. FEBS Letts 1992;312:127–131.
7. Gouet P, Jouve HM, Dideberg O. Crystal structure of *Proteus mirabilis* PR catalase with and without bound NADPH. J Mol Biol 1995;249:933–954.
8. Bravo J, Verdaguer N, Tormo J, Betzel C, Switala J, Loewen PC, Fita I. Crystal structure of catalase HPII from *Escherichia coli*. Structure 1995;3:491–502.
9. Loewen PC, Switala J, von Ossowski I, et al. Catalase HPII of *Escherichia coli* catalyzes the conversion of protoheme to *cis*-heme *d*. Biochemistry 1993;32:10159–10164.
10. Murshudov GN, Grebenko AI, Barynin V, et al. Structure of the heme *d* of *Penicillium vitale* and *Escherichia coli* catalases. J Biol Chem 1996;271:8863–8868.
11. Bravo J, Fita I, Ferrer JC, et al. Identification of a novel bond between a histidine and the essential tyrosine in catalase HPII of *Escherichia coli*. Protein Sci. 1997;6:1016–1023.
12. Mulvey MR, Sorby PA, Triggs-Raine BL, Loewen PC. Cloning and physical characterization of *katE* and *katF* required for catalase HPII expression in *Escherichia coli*. Gene 1988;73:337–345.
13. Von Ossowski I, Mulvey MR, Leco PA, Borys A, Loewen PC. Nucleotide sequence of *Escherichia coli katE*, which encodes catalase HPII. J Bacteriol 1991;173:514–520.
14. Loewen PC, Switala J. Purification and characterization of catalase HPII from *Escherichia coli*. K12. Biochem. Cell Biol 1986;64: 638–646.
15. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Meth. Enzymol 1996;276:307–326.
16. Brünger AT. XPLOR Manual.Version 3. New Haven, CT: Yale University; 1992. 405 p.
17. Roussel A, Cambillau C. TURBO-FRODO, Manual. Mountain View, CA: Silicon Graphics Directory, Silicon Graphics; 1993. 170 p.
18. Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for building protein models in electron density maps. Acta Crystallogr. 1990;47:110–119.
19. Kleywegt GJ, Jones TA. Detection, delineation, measurement and display of molecules. Acta Crystallogr D 1994;50:178–185.
20. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: A program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283–291.
21. Melik-Adamyan WR, Barynin VV, Vagin AA, et al. Comparison of beef liver and *Penicillium vitale* catalases. J Mol Biol 1986;188: 63–72.
22. Sevinc S, Ens W, Loewen PC. The cysteines of catalase HPII of *Escherichia coli*, including Cys438 which is blocked, do not have a catalytic role. Eur J Biochem 1995;230:127–132.
23. Jouve HM, Tessier S, Pelmont J. Purification and properties of the *Proteus mirabilis* catalase. Can J Biochem Cell Biol 1983;61:8–14.
24. Nichols A, Bharadwaj R, Honig B. GRASP-Graphical representation and analysis of surface properties. Biophys J 1993;64:A116.
25. Bergdoll M, Remy M-H, Cagon C, Masson J-M, Dumas P. Proline-dependent oligomerization with arm exchange. Structure 1997;5: 391–401.
26. Sevinc MS, Switala J, Bravo J, Fita I, Loewen PC. Truncation and heme pocket mutations reduce production of functional catalase HPII in *Escherichia coli*. Protein Eng. 1998;11:549–555.
27. Fita I, Rossmann MG. The active center of catalase. J Mol Biol 1985;185:21–37.
28. Hillar A, Nicholls P, Switala J, Loewen PC. NADPH binding and control of catalase compound II formation: Comparison of bovine, yeast, and *Escherichia coli* enzymes. Biochem J 1994;300:531–539.
29. Fita I, Rossmann MG. The NADPH binding site on beef liver catalase. Proc. Natl. Acad. Sci. USA 1985;82:1604–1608.
30. Kirkman HN, Gaetani GF. Catalase: a tetrameric enzyme with four tightly bound molecules of NADPH. Proc. Natl. Acad. Sci. USA. 1984;81:4343–4347.
31. Schonbaum GR, Chance B. Catalase. In: Boyer PD, editor. The enzymes, Vol 13. New York: Academic Press; 1976. p 363–408.
32. Chiu JT, Loewen PC, Switala J, Gennis RB, Timkovich R. Proposed structure for the prosthetic group of the catalase HPII from *Escherichia coli*. J Am Chem Soc 1989;111:7046–7050.
33. Bravo J, Fita I, Gouet P, Jouve HM, Melik-Adamyan W, Murshudov GN. Structure of catalases. In: Scandalios JG, editor. Oxidative stress and the molecular biology of antioxidant defenses. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 1997. p 407–445.