The crystal structure of chloroperoxidase: a heme peroxidase–cytochrome P450 functional hybrid

Munirathinam Sundaramoorthy[1], James Terner[2] and Thomas L Poulos[1]*

[1]Departments of Molecular Biology and Biochemistry and, Physiology and Biophysics, University of California, Irvine, CA 92717, USA and [2]Department of Chemistry, Virginia Commonwealth University, Richmond, VA 23284-2006, USA

Background: Chloroperoxidase (CPO) is a versatile heme-containing enzyme that exhibits peroxidase, catalase and cytochrome P450-like activities in addition to catalyzing halogenation reactions. The structure determination of CPO was undertaken to help elucidate those structural features that enable the enzyme to exhibit these multiple activities.

Results: Despite functional similarities with other heme enzymes, CPO folds into a novel tertiary structure dominated by eight helical segments. The catalytic base, required to cleave the peroxide O–O bond, is glutamic acid rather than histidine as in other peroxidases. CPO contains a hydrophobic patch above the heme that could be the binding site for substrates that undergo P450-like reactions. The crystal structure also shows extensive glycosylation with both N- and O-linked glycosyl chains.

Conclusions: The proximal side of the heme in CPO resembles cytochrome P450 because a cysteine residue serves as an axial heme ligand, whereas the distal side of the heme is 'peroxidase-like' in that polar residues form the peroxide-binding site. Access to the heme pocket is restricted to the distal face such that small organic substrates can interact with an iron-linked oxygen atom which accounts for the P450-like reactions catalyzed by chloroperoxidase.

Structure 15 December 1995, **3:**1367–1377

Key words: carbohydrate binding, halogenation, heme enzyme, peroxidase, thiolate ligand

Introduction

Halogenation of aliphatic molecules gives rise to a wide variety of natural products, some of which, like chloramphenicol, are of medicinal value [1]. Enzyme-catalyzed halogenation reactions most often involve oxidation of the halide ion and formation of a C–X bond. The heme enzyme, chloroperoxidase (CPO), from the fungus *Caldariomyces fumago*, participates in the production of the natural product caldariomycin (1,1-dichloro-2,5-dihydroxy cyclopentane) [2]. The overall catalytic cycle of CPO is characteristic of many heme peroxidases (Fig. 1a).

Although the primary biological function of CPO is chlorination [2], CPO also catalyzes other oxidative reactions characteristic of other heme peroxidases, catalase and cytochrome P450 (Fig. 1b) [3]. Owing to this diversity, CPO is the most versatile of the known heme enzymes and has been the subject of extensive investigations since the enzyme was first described by Morris and Hager in 1966 [2]. One of the main conclusions from these studies is that CPO appears to share properties with both P450s and heme peroxidases. CPO and P450s share various spectroscopic properties [4,5] which indicate that, like P450, one of the axial heme ligands in CPO derives from the sulfur atom of a cysteine residue. In contrast, heme peroxidases use histidine as a ligand. However, the distal heme pocket in CPO, which forms the peroxide-binding site, is predicted to be formed primarily by polar amino acids [6–8]. This is in sharp contrast with P450s in which the distal oxygen-binding site is lined primarily with non-polar groups. Hence, CPO is thought to have a P450-like proximal pocket and a peroxidase-like distal pocket.

Another important property of both peroxidases and P450s is accessibility to the heme group. In peroxidases, substrates are limited to electron-transfer reactions at the heme edge with restricted or no direct access to the $Fe^{4+}=O$ center in compound I [9]. In sharp contrast, the P450 heme edge is not accessible and substrates must bind directly adjacent to the $Fe^{4+}=O$ center for stereospecific hydroxylation [10]. CPO can do both, but whether this is due to multiple access routes to the heme remains unknown. The structure of this rather unusual and functionally diverse heme enzyme should help us

(a)

$$Fe^{3+}P + H_2O_2 \rightarrow \underset{\text{Compound I}}{Fe^{4+}=O\ P^{\cdot} + H_2O}$$

$$Fe^{4+}=O\ P^{\cdot} + \text{substrate} \rightarrow \underset{\text{Compound II}}{Fe^{4+}=O\ P} + \text{substrate}^{\cdot}$$

$$Fe^{4+}=O\ P + \text{substrate} \rightarrow Fe^{3+}\ P + \text{substrate}^{\cdot}$$

(b)

Halogenation	$RH + H_2O_2 + Cl^- + H^+ \rightarrow R\text{–}Cl + 2H_2O$
Dehydrogenation (heme peroxidases)	$2RH + H_2O_2 \rightarrow R\text{–}R + 2H_2O$
H_2O_2 decomposition (catalase)	$2H_2O_2 \rightarrow O_2 + 2H_2O$
Oxygen insertion (P450)	$R + H_2O_2 \rightarrow R\text{–}O + H_2O$

Fig. 1. (a) Overall catalytic cycle of heme peroxidases. In the first step, hydrogen peroxide removes one electron from the iron atom and one from the porphyrin (P) to give the oxyferryl ($Fe^{4+}=O$) center and a porphyrin π cation radical [33]. In the second step, compound I is reduced by a substrate molecule, and in the third and last step of the cycle, a second molecule of substrate is oxidized. **(b)** Reactions catalyzed by chloroperoxidase.

*Corresponding author.

 ISSN 0969-2126

to further understand the differences and similarities between peroxidases and P450s as well as to probe the structural basis for the wide range of activities exhibited by CPO. We have crystallized and solved the structure of CPO isozyme A in two space groups. The crystal structure shows many interesting features, some previously predicted and some hitherto unknown.

Results and discussion

Overall structure

The structure has been refined in two space groups, $P2_12_12_1$ and $C222_1$, at 1.9 Å and 2.1 Å respectively (Table 1). The root mean square deviation (rmsd) between the two models is 0.32 Å for backbone atoms and 0.4 Å for all protein atoms. The deviation of Cα atoms exceeds 1 Å for only four residues at the N terminus and two other residues on a surface loop. More than 90% of (ϕ,ψ) angles fall within the most favorable regions of Ramachandran plots [11]. Unless otherwise stated, the model described in this paper is the one that was refined in space group $P2_12_12_1$ because this crystal form diffracted to higher resolution and had better defined electron density for the carbohydrate structure.

Table 1. Model refinement statistics.

	$P2_12_12_1$	$C222_1$
No. of non-H atoms		
Protein	2316	2316
Heme	43	43
Carbohydrates	238	156
Cation	1	1
Solvent	190	195
Highest resolution	1.9 Å	2.1 Å
R-factor		
Resolution range	8.0–1.9 Å	8.0–2.1 Å
2σ cutoff	25 373 (18.2%)	23 137 (18.6%)
No σ cutoff	27 521 (19.2%)	25 061 (19.6%)
Resolution range	30.0–1.9 Å	30.0–2.1 Å
2σ cutoff	25 852 (20.9%)	23 681 (22.2%)
No σ cutoff	28 004 (21.9%)	25 607 (23.1%)
Free R-factor	22.4%	22.8%
Rmsd in bond lengths	0.009 Å	0.008 Å
Rmsd in bond angles	1.557°	1.488°
Mean coordinate error		
Luzzati plot [51]	0.20 Å	0.20–0.25 Å
SIGMAA [49]	0.22 Å	0.26 Å

Relative to the cDNA sequence [12], the crystal structure is missing 52 residues at the C terminus. The possibility that the C terminus is disordered can be ruled out because the electron density is well defined in both space groups and the models exhibit very low temperature (B) factors. Moreover, there is insufficient room to accommodate the extra 52 residues in either crystal from, and the amino acid analysis of the protein used for crystallization correlates with the crystal structure. The only other possibility is that the protein undergoes a maturation cleavage event resulting in a 299-residue catalytically active peptide and a 52-residue C-terminal fragment (LP Hager, personal communication). The original DNA sequence indicates that the complete CPO gene encodes a protein of 373 amino acid residues [12]. This 373-residue polypeptide chain is processed to a mature, secreted CPO enzyme containing 299 amino acid residues by two proteolytic cleavages: one removes a 20 amino acid signal peptide from the N terminus and the other removes a 52 amino acid peptide from the C terminus. CPO undergoes at least two other post-translational modifications. One of these is glycosylation (discussed later) and the second is cyclization of the N-terminal glutamic acid [13]. CPO is synthesized as a proenzyme with the proteolytic clip between Gln(−1) and Glu1. However, the crystal structure shows that residue −1 is still present as the cyclized pyroglutamic acid (designated residue number 0 to allow the previous numbering scheme to be retained).

The CPO structure was compared with 420 protein structures in the Protein Data Bank using a three-dimensional alignment algorithm (DALI [14]). A 30% sequence identity cutoff was applied. The comparison reveals that the CPO polypeptide folds into a novel tertiary structure (Fig. 2). The helical content (~50%) is similar to that found in other peroxidases and P450s, but the overall topology lacks any resemblance to these other heme enzymes. Nevertheless, there are important structural similarities near the active site. As with other heme peroxidases and other proteins, CPO folds into N-terminal and C-terminal domains with the heme sandwiched between the two domains. The fold comprises eight α-helical segments (A–H), three short 3_{10} helices (C′, D′ and G′) and a short antiparallel β pair (Table 2). Like other heme enzymes, CPO has a proximal helix (helix A) on the side of the heme containing the protein axial heme ligand, and a distal helix (helix F) that provides catalytic groups required for peroxide activation. Whereas in CPO the proximal helix A is approximately perpendicular to the heme plane, in both the peroxidases and P450s it is parallel to the heme plane. As anticipated from chemical modification studies [15], Cys29 provides the thiolate heme ligand. The axial ligand residue is in the N-terminal domain in CPO, whereas in other heme enzymes it resides in the C-terminal domain. The only other two cysteines, Cys79 and Cys87, form a disulfide bond (Fig. 2). Three prolines (Pro9, Pro230 and Pro292) adopt a *cis* conformation and one of these (Pro9) is in a classical *cis*-proline turn [16].

Proximal heme ligand

The polypeptide from residues 26–37 encompasses the cysteine ligand (Cys29) and provides a rigid scaffolding for the iron–sulfur interaction. The side chain of Asn33 forms a pair of hydrogen bonds with the backbone atoms of Ala27, while the side chains of Arg26 and Asp37 form hydrogen bonds to each other (Fig. 3). This conformation differs significantly from that present in P450cam in which the cysteine-ligand region is held together by an antiparallel β-pair hydrogen-bonding pattern. As with the P450 cysteine ligand, the Cys29 ligand in CPO is situated at the N terminus of the proximal helix. However,

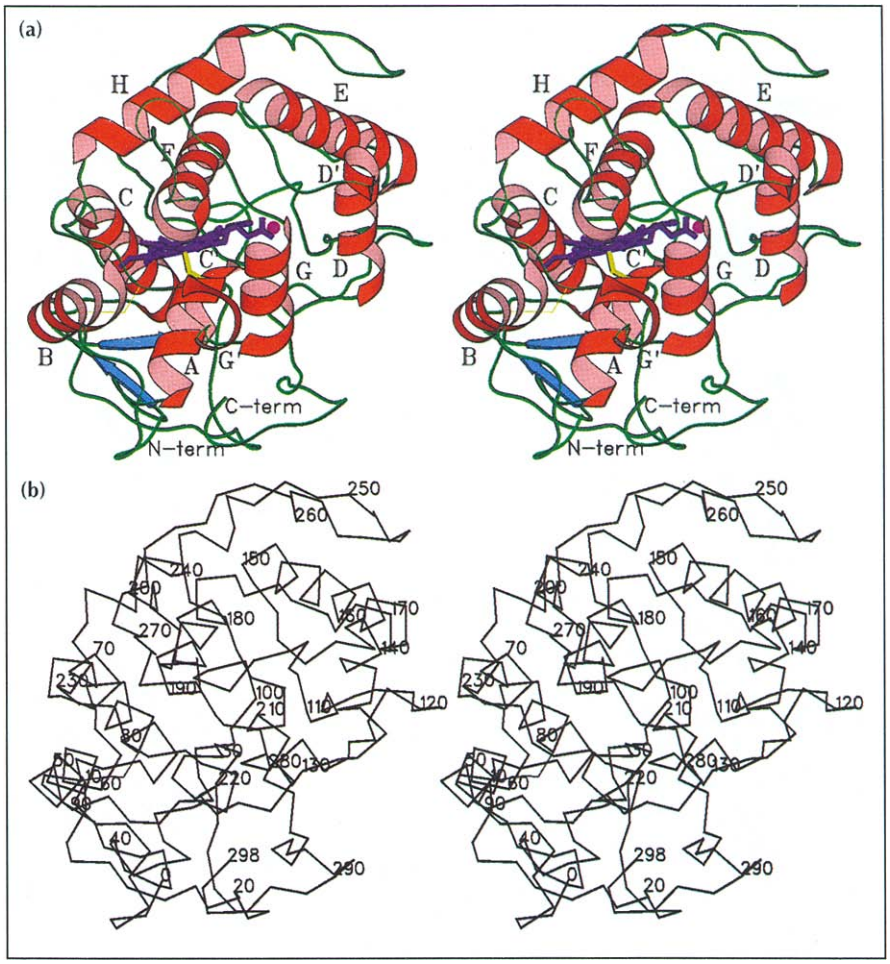

Fig. 2. **(a)** Stereo representation of the CPO molecule. α helices are shown in red and are labeled with upper-case letters from A–H and 3_{10} helices are labeled with primed letters. The β pair is shown as blue arrows. The cation is indicated as a pink sphere near a heme propionate. The N and C termini are only 5.6 Å apart and are bridged by a solvent molecule (not shown). **(b)** Stereo diagram of the Cα trace of CPO with every tenth Cα position labeled. (Figure made with SETOR [52].)

Table 2. Secondary structures.

Helix	Beginning	Ending	No. of residues
A	C29 O...N N33	A34 O...N H38	10
B	S49 O...N L54	P57 O...N M61	13
C	A64 O...N I68	E80 O...N G84	21
C′	N95 O...N A98		4
D	D131 O...N F135	Q136 O...N D140	10
D′	D140 O...N A143		4
E	D149 O...N M153	S164 O...N D168	20
F	K176 O...N Q180	I187 O...N V191	16
G	R206 O...N W210	W213 O...N E217	10
G′	Y221 O...N G244		4
H	E233 O...N V237	A243 O...N A247	15

Beginnings and endings of helices are defined by the first and last 1→4 hydrogen bonds, respectively. Primes indicate 3_{10} helices; β pair: 45–48, 91–94.

in CPO the proximal helix is located in the N-terminal half of the protein while in P450 the cysteine ligand (Cys357 in P450cam) is located near the C terminus. Moreover, as noted above, the proximal helix in CPO is perpendicular to the heme while in P450cam the proximal helix is parallel (Fig. 4a).

Even though the proximal helices are oriented very differently in CPO and P450, the conformation of the polypeptide backbone for a three-residue stretch surrounding the cysteine ligand is the same. The Pro-Cys-Pro backbone conformation in CPO superimposes on the X-Cys-X backbone of the various known P450 structures [10,17–20] with an rmsd of ~1 Å. Of particular interest is the very similar hydrogen-bonding arrangement between the cysteine sulfur atom and the neighboring peptide amide groups (Fig. 4a). In both CPO and P450cam, the cysteine ligand is at the beginning of the proximal helix. The N-terminal ends of helices often have amino acid side chains which serve to accept hydrogen bonds from peptide amide groups which, at the beginning of helices, have no peptide carbonyl oxygen hydrogen-bond partner [21]. The cysteine ligand in both CPO and P450cam serves this purpose. In CPO, the Cys29 sulfur is approximately 3.6 Å from the peptide amide groups of residues 31 and 32 while in P450cam, the sulfur of Cys357 is approximately 3.6 Å and 3.3 Å from the peptide amide groups of residues 358 and 359, respectively. This unusual amide–sulfur hydrogen bond may be of general significance in iron proteins that use cysteine as ligands [22]. For example, all four cysteine residues that serve as iron ligands in nitrogenase [23] form similar interactions at the N termini of helices. A superimposition of the Pro-Cys-Pro backbone in CPO with the four similar regions in nitrogenase gives rmsd values between 0. 66 Å and 0.87 Å (Fig. 4b). These similarities indicate that

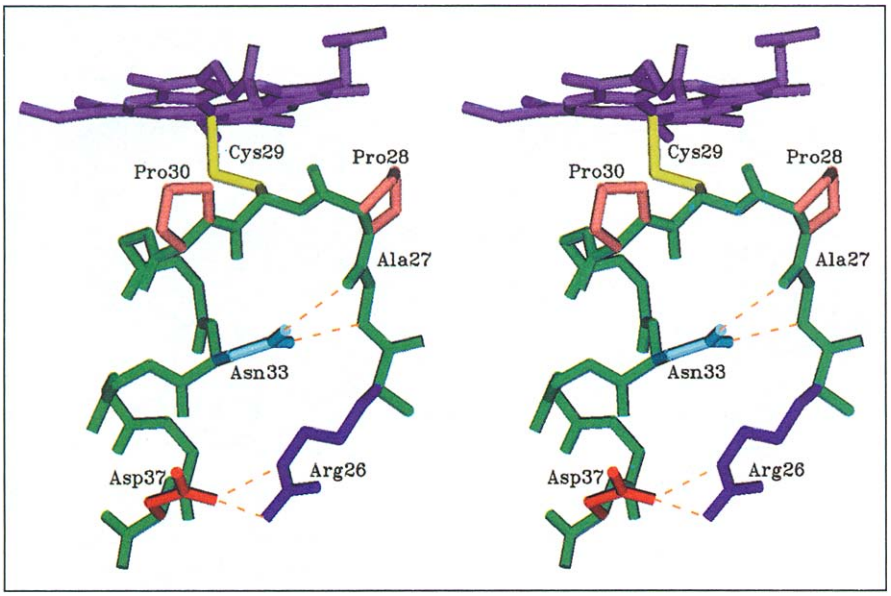

Fig. 3. Stereo diagram showing the hydrogen-bonding interactions that help to stabilize the cysteine-ligand loop. Key residues participating in hydrogen bonds are labeled. Dashed lines denote hydrogen bonds. (Figure made with SETOR [52].)

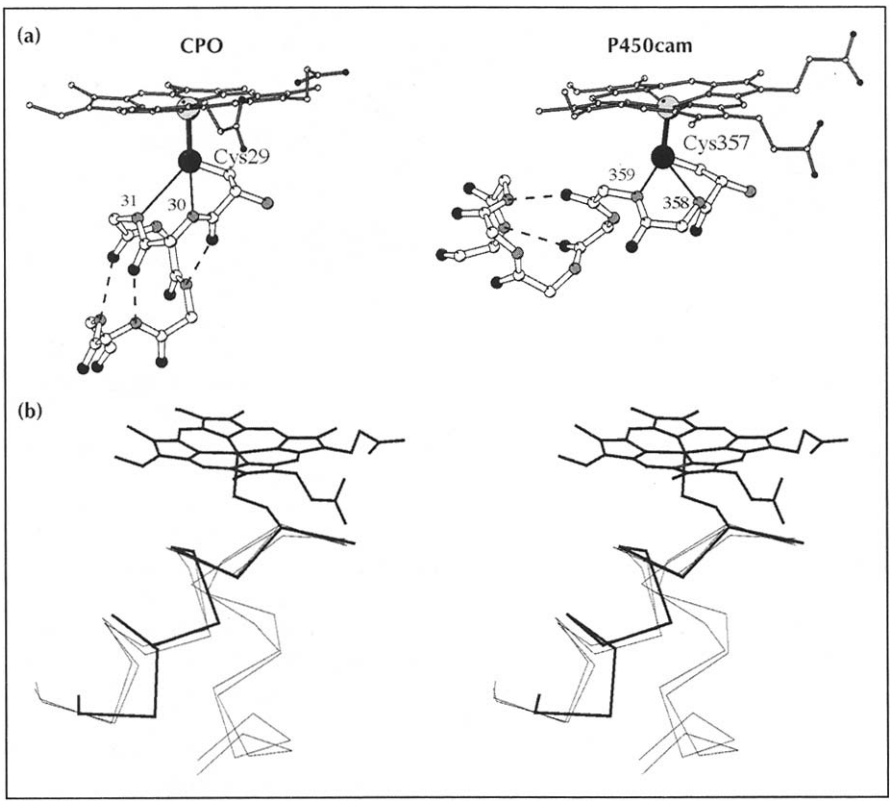

Fig. 4. **(a)** Conformation of the polypeptide near the cysteine ligand in CPO and P450cam. The dashed lines indicate helical hydrogen bonds while the thin solid lines indicate the hydrogen bonds between the cysteine sulfur and peptide amide groups. Note that in CPO the helix is perpendicular to the heme, whereas in P450 the helix is parallel to the heme. (Figure made with MOLSCRIPT [53].) **(b)** Stereoview showing the superimposition of the CPO cysteine-ligand region on three iron-ligating regions in nitrogenase. The CPO helix is shown as a thick line and the four helices of nitrogenase are shown as thin lines. The peptide amide–sulfur hydrogen bonds found in CPO and P450 also are present in the nitrogenase helices. (Figure made with SETOR [52].)

peptide amide–sulfur ligand hydrogen bonds may be a characteristic and important feature for metalloproteins that utilize cysteine residues to coordinate iron.

Heme environment

As with many heme proteins, the heme in CPO deviates from planarity and is bowl shaped (Table 3). The iron is displaced 0.14 Å below the mean plane of the porphyrin towards the cysteine ligand compared with 0.4 Å for high-spin P450cam. The iron–sulfur bond distance is 2.3 Å in both CPO crystal forms which is the same as the distance estimated by extended X-ray absorption fine structure analysis [24]. The opposite (distal) side of the heme is occupied by a solvent molecule which is positioned 3.4 Å from the iron atom and too far away to form a strong axial ligand. This is consistent with the high-spin spectroscopic properties of ferric CPO [25].

As with other heme proteins, the heme is held in place by both hydrophobic and hydrogen-bonding interactions. The environment around the heme propionates is more similar to peroxidases than P450s. In P450cam the propionate groups are surrounded by charged side chains. The propionates in CPO, however, interact primarily with backbone atoms. The loop between Glu104 and Ser110 provides the primary set of interactions (Fig. 5).

Table 3. Heme geometry.

Parameter	Distance (Å)	Angle (°)
Fe to Cys29 Sγ	2.30	
Fe to pyrrole N (average)	2.04	
Fe to pyrrole N plane	0.14	
Cys29 Sγ to pyrrole N plane	2.44	
Fe to porphyrin plane	0.36	
Pyrrole N plane to porphyrin plane		0.94
Pyrrole A plane to porphyrin plane		5.78
Pyrrole B plane to porphyrin plane		5.73
Pyrrole C plane to porphyrin plane		10.91
Pyrrole D plane to porphyrin plane		7.99
Cys29 Sγ–Fe bond vector to pyrrole N normal		177.37
Cβ–Sγ –Fe		110.54
Cys29 side-chain torsion angle		174.00

The most notable and unexpected feature involving heme interactions is a cation coordinated by one of the heme propionates. At the early stage of modeling water molecules, very strong difference density was observed near a heme propionate (Fig. 5). This was subsequently attributed to a cation when the difference density was still over 20 times above background with water included at this site, and because the site is surrounded by protein oxygen atoms and solvent molecules. Electron paramagnetic resonance studies indicated the presence of a Mn^{2+} ion [26] and this was later confirmed by X-ray fluorescence spectroscopy [1]. Hence, this has been tentatively modeled as a Mn^{2+}-binding site. In addition to the heme propionate, the cation is coordinated by five other protein and water ligands in an octahedral geometry (Fig. 5). The location and coordination geometry of the cation is remarkably similar to that of the Mn^{2+}-binding site in the crystal structure of manganese peroxidase [27]. In manganese peroxidase, the Mn^{2+} ion also coordinates with a heme propionate, but in this case the Mn^{2+} ion is a substrate [28]. While the exact role of Mn^{2+} ion in CPO has yet to be established, studies have shown that the activity of the enzyme is not altered by the presence or absence of Mn^{2+} (LP Hager, personal communication).

Distal heme pocket

Although CPO has a thiolate heme ligation similar to that in P450s, the opposite or distal face of the heme is expected to resemble the polar peroxidase active site rather than the more non-polar P450 active site [10]. Crystal structures of cytochrome *c* peroxidase (CcP) and catalase have revealed that both possess a distal histidine [9,29]: His52 in CcP is perpendicular to the heme plane and His74 in catalase is parallel. In both enzymes, the distal histidine works in concert with another polar residue (arginine in peroxidases and asparagine in catalase) in an acid–base catalytic mechanism required to cleave the peroxide O–O bond in the formation of compound I [30]. Various equilibrium binding studies of exogenous ligands and halides binding to the CPO heme iron have revealed the presence of a titratable residue that affects the binding. This residue has generally been considered to be a distal histidine [6–8].

The crystal structure reveals a somewhat different picture. The heme pocket in CPO is polar as predicted, but the acid–base catalytic group is not histidine. Instead, the side chain of Glu183 extends from the distal helix F and is positioned directly adjacent to the peroxide-binding site (Fig. 6). Histidine, however, is involved indirectly. His105 hydrogen bonds with Glu183 and the carbonyl oxygen of Asp106. This arrangement is quite different from that found in the heme peroxidases where both the catalytic histidine and arginine extend from the distal helix and can interact with the iron-linked peroxide. Glu183 in CPO is the only polar or charged group close enough to the peroxide-binding site to serve a direct catalytic role. The most probable role of His105 is to correctly orient Glu183 and, possibly, provide charge–charge interactions. It should be noted that His105 is only 3.5 Å above one of the propionate carboxylate groups (Fig. 6), thus increasing the pK_a of the histidine side chain. Also of note is the fact that the propionate with which His105 interacts is the same propionate that is a ligand for the cation. This intricate

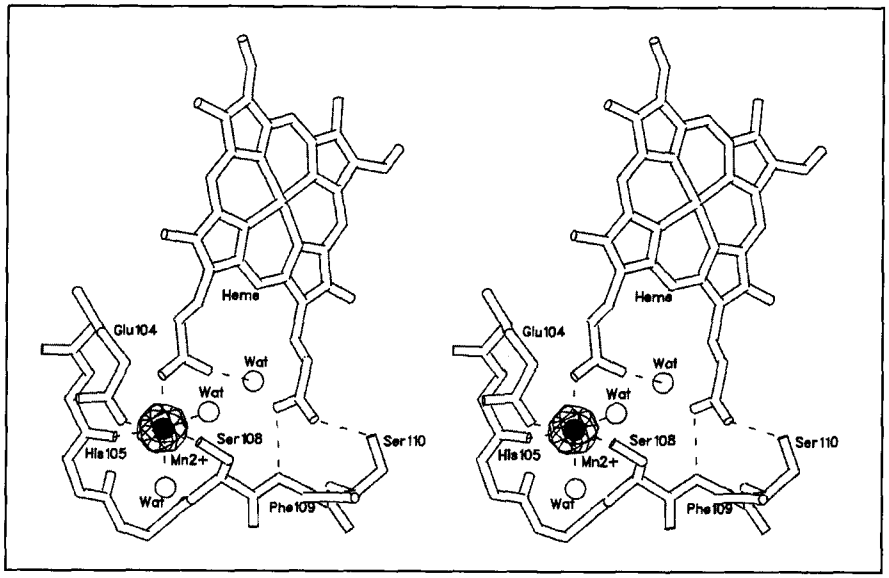

Fig. 5. Stereoview of the cation-binding site. The map is an F_o–F_c omit difference electron-density map contoured at 12 times above background level. Ligand interactions with the cation are indicated as dashed lines. (Ligands for Mn^{2+} and distances: heme propionate O2A, 2.33 Å; side-chain carboxyl Oε1–Glu104, 2.10 Å; carbonyl O–His105, 2.06 Å; side-chain Oγ–Ser108, 2.28 Å; water802, 2.33 Å; water803, 2.40 Å.) Also shown are the hydrogen-bond interactions of the heme and polypeptide. (Figure made with SETOR [52].)

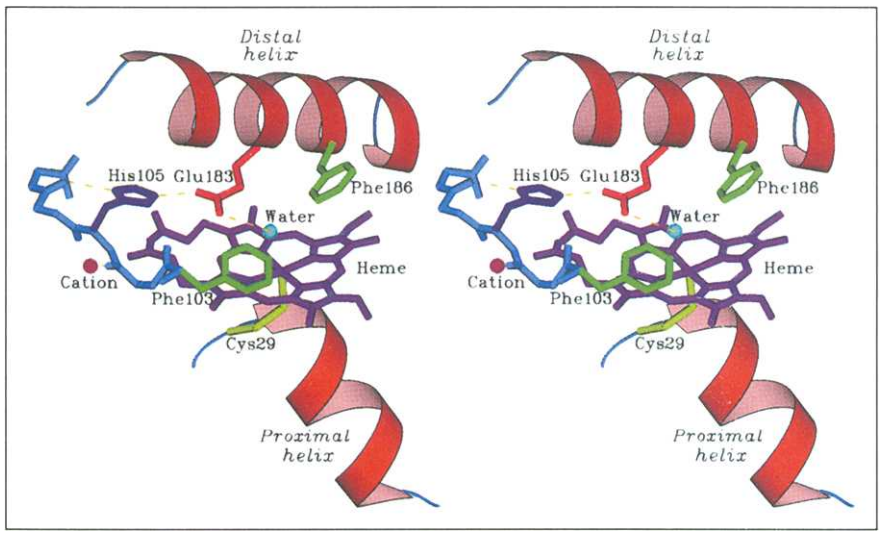

Fig. 6. Stereoview of the CPO active site. Hydrogen bonds involving the distal side water molecule, Glu183 (the proposed acid–base catalyst) and His105 are shown as dashed lines. The histidine is only 3.5 Å from the heme propionate that binds a cation. The proposed substrate-binding pocket is bracketed by Phe103 and Phe186. (Figure made with SETOR [52].)

set of electrostatic and hydrogen-bonding interactions involving side chains, the heme, and a cation may be important for catalysis and could imply a role for the cation other than a simply structural one.

Glu183 at the active site can help to explain some of the solution and pH properties of CPO. While ligands bind to the iron as the anion, CPO accepts weakly acidic ligands such as HCN (pK_a>3) into the active site in the neutral protonated form, whereas strong acids like HNO_3 (pK_a<0) bind preferentially at low pH as the anion [31]. The presence of the negatively charged carboxylate on Glu183 would be expected to repel entry of the anionic form of weak acids owing to unfavorable electrostatic interactions. Strong acids enter the active site in their anionic form only at low pH when Glu183 is protonated. The exact catalytic role of Glu183 is probably similar to that of the distal histidine in peroxidases. The carboxylate accepts a proton from the incoming peroxide and donates a proton to the leaving hydroxyl group:

$$Fe^{3+}\ GluCO_2^- + H_2O_2 \rightarrow Fe^{3+}\text{–}O\text{–}OH\ GluCO_2H \rightarrow$$
$$Fe^{3+}\text{–}O\ GluCO_2^- + H_2O$$

The Fe^{3+}–O species then oxidizes the porphyrin and iron to give compound I. In CcP, the crystal structure of compound I shows that the active-site arginine residue moves in to interact with and probably stabilize the oxygen of the Fe^{4+}=O center [32]. The electrostatic environment in CPO clearly must be different if Glu183, either as the anionic carboxylate or protonated carboxyl group, also interacts directly with the Fe^{4+}=O oxygen atom. In either case, the energetics should be quite different from those in other peroxidases, such as horseradish peroxidase, that form considerably more stable compound I porphyrin π-cation radical and Fe^{4+}=O centers than in CPO [33].

Substrate-binding pocket

In heme peroxidases the heme edge is available for substrate interactions but direct access to the Fe^{4+}=O center is restricted. The opposite is true in P450s, in which the heme edge is not accessible but a substrate pocket exists directly adjacent to the Fe^{4+}=O center. The crystal structures of two P450–substrate complexes [10,20] with very different specificities clearly show that the substrate is surrounded by protein groups required for holding the

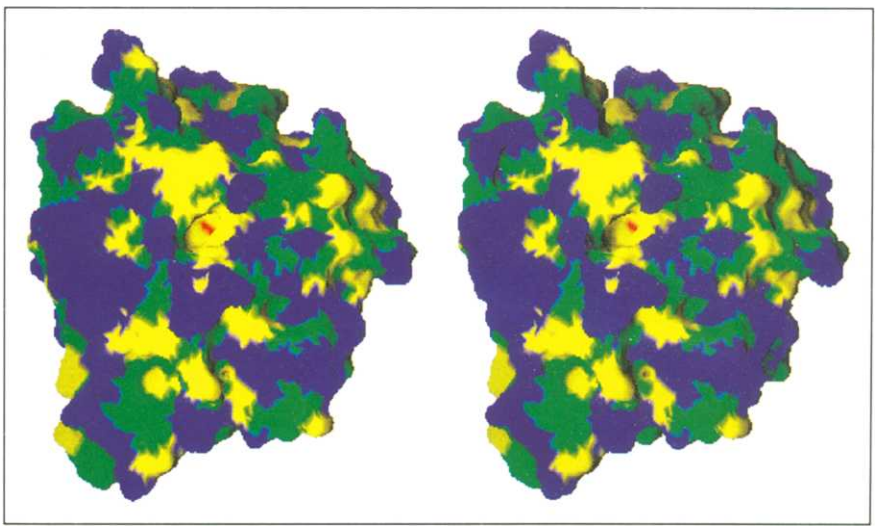

Fig. 7. Stereoview of the van der Waals surface of the substrate-binding region. Colour coding is as follows: blue, charged/polar side chains; green, backbone atoms; yellow, hydrophobic side chains. The red region at the bottom of the proposed substrate-binding site is the catalytic Glu183 which is situated just above the heme. (Figure made with GRASP [54].)

substrate in place for stereospecific hydroxylation. Because the heme edge is not connected to the molecular surface via a channel as it is in other peroxidases, CPO more closely resembles P450 in this respect. However, there is a small opening above the heme which could allow access to the Fe^{4+}=O center in compound I (Fig. 7). This channel, which connects the surface to the heme distal side, is formed by the C, F and H helices and loops 263–269 and 101–110. The channel is mostly hydrophobic near the heme and wider and more polar at the top. A mannose carbohydrate bound to Thr238 is located at the opening of the channel, but would not prevent substrate entry. Glu69 and Glu266 are also located near the entrance of the channel but these residues appear to be quite flexible, as indicated by their high temperature factors.

The side chains of Phe103, Val182 and Phe186 are situated at the bottom of this opening, and are available for interactions with organic substrates. One such substrate is dimethylaniline which can undergo a CPO-catalyzed oxidative demethylation [34]. Docking of dimethylaniline into the CPO access channel results in hydrophobic contacts with the side chains of Phe103, Leu70, Phe186, Val182 and Ala267. However, much of the substrate remains exposed to solvent and there is no indication, such as a disordered loop nearby, of an open/close motion that might result in a more buried substrate. Such a mechanism has been proposed to occur in P450s on the basis of energy minimization and molecular dynamics simulations [19,35].

Nevertheless, there are indications that the bottom of the pocket can adjust to substrate binding. Phe103, one of the residues at the bottom of the channel expected to contact substrates, appears to be flexible. This flexibility is intimated by a difference in the heme–phenyl ring distances in the two crystal structures. The phenyl ring is closer to the heme in the $P2_12_12_1$ form by 0.2 Å relative to the $C222_1$ form. Furthermore, the temperature factors are higher for this residue compared with the other residues in the region. A difference Fourier map calculated with one of the heavy-metal derivative data sets in the $C222_1$ form showed strong positive difference density above the heme iron on the distal side. This density is consistent with a diatomic molecule bound to the heme iron, probably a contaminant in the heavy-atom reagent. The model refined with this heavy-metal derivative data set in $C222_1$ form showed a larger deviation in the Phe103–heme distance (0.4 Å). The flexibility is also manifested in the slight rearrangement of solvent molecules in the access channel. The residual electron density in the channel close to Phe103, which was modeled as water molecules, is stronger and continuous, extending to the phenyl ring in the $C222_1$ form. The presence of 310 mM NaCl in the crystallization medium used to grow crystals of this form might be responsible for the differences in heme–phenyl ring distances as crystal contacts are not involved in this region in either space group. This opening appears to provide a somewhat flexible hydrophobic patch that will only partially bury an aromatic substrate the size of dimethylaniline. This may account for the rather broad range of specificities exhibited by CPO in various P450-like reactions. By contrast, in P450 numerous specific protein–substrate interactions, sequestered from bulk solvent, rigidly orient the substrate for stereospecific hydroxylation.

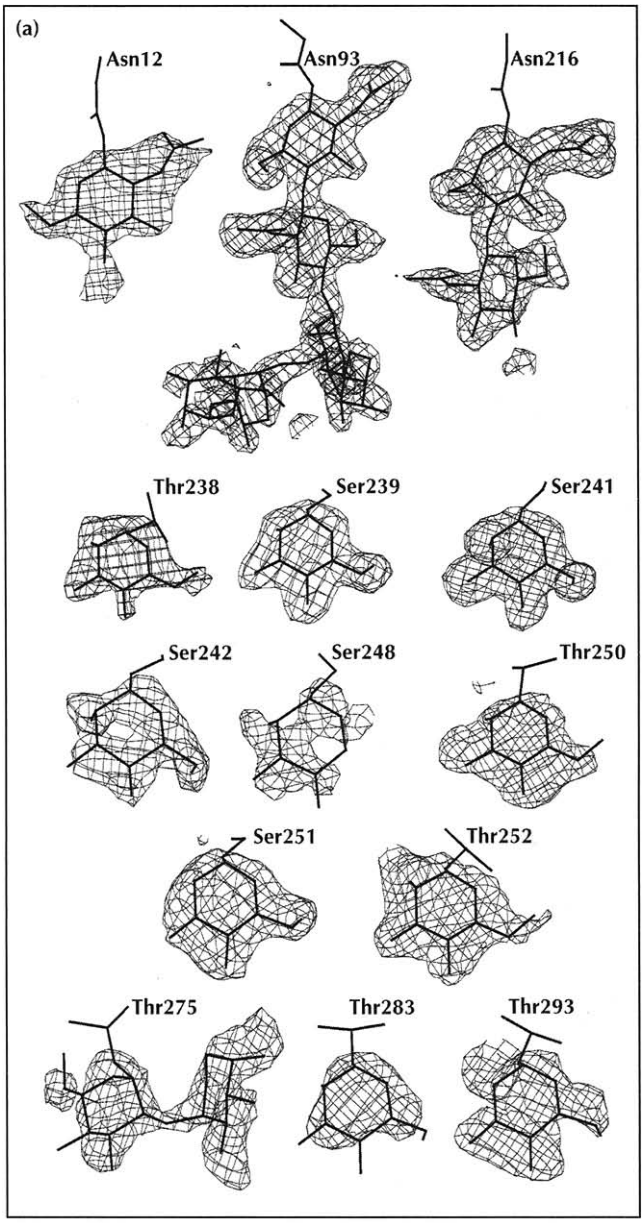

Fig. 8. **(a)** F_o–F_c omit electron-density maps contoured at 2.5σ for the 14 carbohydrate attachment sites found in the $P2_12_12_1$ crystal form. **(b)** The structure of the most extensively branched site which is attached to Asn93.

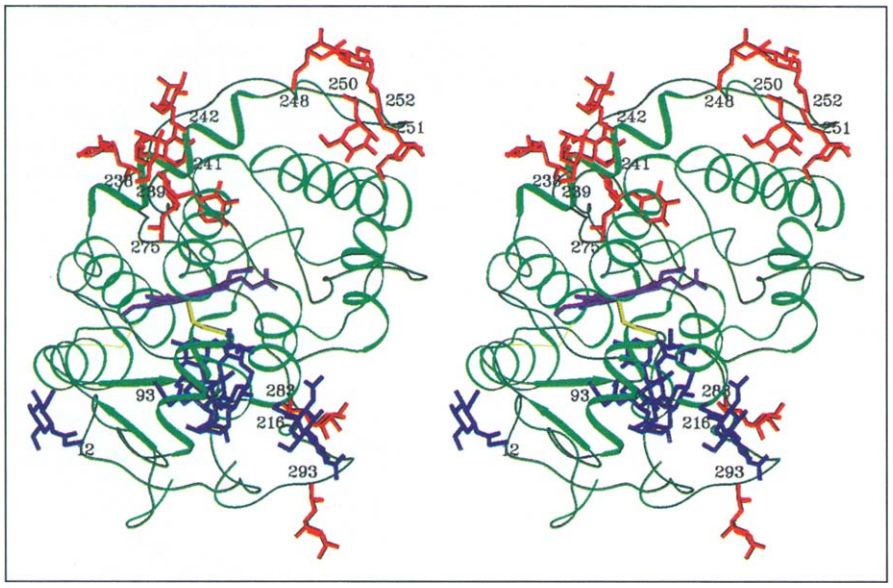

Fig. 9. Stereoview showing the backbone of CPO and the location of all observed carbohydrate units. The *O*-glycosylation sites are shown in red and the *N*-glycosylation sites in blue. (Figure made with SETOR [52].)

CPO still exhibits some selectivity with respect to substrate size and stereochemistry in the oxidation of aromatic molecules. Substrates with bulky substituents are not oxidized while, for example, only the L isomer of *N*-acetyl tyrosine is oxidized [36]. The size-dependence is relatively easy to rationalize as the hydrophobic patch is small and cannot accommodate very large aromatic molecules. The stereoselectivity is a more intriguing problem. The modeling we have carried out so far indicates that simple steric considerations govern the stereoselectivity of CPO. We hope to be able to test these hypotheses by diffusing selected substrates into the CPO crystals.

A discussion of CPO substrates would not be complete without considering where and if halides (specifically chloride) bind. The $C222_1$ crystal form grows in the presence of 310 mM NaCl. If Cl^- bound tightly and specifically to some site on the enzyme, we might anticipate that one of the solvent molecules modeled as water would be situated in extra electron density corresponding to the much larger (18 electrons) Cl^- ion. We found no such peak although the $C222_1$ form grown in high NaCl buffers does contain some extraneous but not very well defined electron density near the proposed substrate-binding site. Further predictions from the crystal structure on where and how an active chlorinating intermediate is formed are complicated by the complex and uncertain chemistry involved. For example, the precise identification of the chlorinating species and which enzyme intermediate is involved is an unsolved problem with CPO [3]. It is unclear whether chlorination takes place while substrates are bound to CPO or whether the chlorination is an enzyme-free reaction involving HOCl as the active agent, although free HOCl is favored [3]. There are also suggestions that it is the oxy-complex of CPO and not compound I which is the active intermediate [3]. Radical mechanisms have been proposed where both the organic substrate and Cl^- are oxidized in separate steps with the radical ultimately chlorinated by Cl_2 [3]. The structure does not allow us to choose between these various mechanisms but at the very least, the resting enzyme appears not to have a specific high-affinity site for Cl^-.

Glycosylation

The two major isozymes of CPO, A and B, have the same amino acid composition but different carbohydrate composition. Form A, used in these studies, is more heavily glycosylated and carbohydrate accounts for about 19% of the total molecular weight [13]. The crystal structure shows 14 glycosylation sites, with a total of 21 sugar molecules in the $P2_12_12_1$ form (Figs 8,9; Table 4), whereas the $C222_1$ form has only 13 carbohydrate groups. Because the content, but not the sequence, of carbohydrate groups is known in CPO, the following discussion on carbohydrate structure is based on other similar glycoproteins [37] and the size and shape of electron density envelopes in the higher-resolution $P2_12_12_1$ form.

Table 4. Carbohydrate-binding sites.

Glycosylation site	Carbohydrate
N-glycosyl chains	
Asn12	β-GlcNAc
Asn93	$(\beta\text{-GlcNAc})_2(\beta\text{-Man})(\alpha\text{-Xyl})_2(\alpha\text{-Man})$
Asn216	$(\beta\text{-GlcNAc})_2$
O-glycosyl chains	
Thr238	α-Man
Ser239	α-Man
Ser241	α-Man
Ser242	α-Man
Ser248	α-Xyl
Thr250	α-Man
Ser251	α-Man
Thr252	α-Man
Thr275	(α-Man)(α-Arb)
Thr283	α-Man
Thr293	α-Man

Abbreviations: GlcNAc, *N*-acetylglucosamine; Man, mannose; Xyl, Xylose; Arb, Arabinose.

The three *N*-linked glycosyl chains are more complex than the eleven *O*-linked chains, with a minimum of two *N*-acetylglucosamine (GlcNAc) molecules in β1→4 linkages (Fig. 8). At the first site, Asn12, the electron density is very clear for one GlcNAc molecule. Although the density extends beyond the first GlcNAc, a second molecule could not be modeled as the density was too weak — probably because the molecule is disordered. The second site, at Asn93, has two molecules each of GlcNAc, mannose and xylose. The second GlcNAc is β1→4 linked to a mannose where the chain branches with α1→6 and α1→3 linked xyloses and one of these is α1→3 linked to a mannose (Fig. 8). A similar pattern of branching is observed in the crystal structure of glucoamylase at pH 4 [37], which is classified as a high-mannose type [38]. The third site, at Asn216, has only two clearly defined GlcNAc molecules. The amino acid sequence at all the three sites conforms to the pattern Asn-X-Ser/Thr.

The C-terminal domain is rich in serine and threonine residues. All of the observed *O*-glycosylation sites are located in this domain. With the exception of Thr275, which is glycosylated with an α1→2 linked mannose–arabinose dimer, only a single mannose unit was visible at each of the *O*-linked glycosyl positions (Figs 8,9). Weak residual electron density indicates a second sugar at sites Thr283 and Thr293 but these sugars were not included in the model.

Biological implications

Chloroperoxidase (CPO) is the most diverse of the known heme enzymes owing to its ability to catalyze peroxidase, catalase and P450-like reactions. It has long been anticipated that CPO would contain a pocket on the proximal side of the heme similar to that found in P450s and a distal pocket resembling the peroxide-binding site of peroxidases. The crystal structure is consistent with these expectations, although the active-site acid–base catalytic group is identified as a glutamate and not a histidine as expected. The very close similarity between CPO, P450 and other sulfur–iron proteins in the hydrogen-bonding pattern involving the cysteine heme ligand and nearby peptide amide groups indicates an important role for these hydrogen bonds in sulfur–iron ligation. The sulfur not only hydrogen bonds with peptide amide groups but is also situated at the N-terminal or positive dipole end of an α helix. Hence, the cysteine sulfur atom is surrounded by, and very likely stabilized by, a positive electrostatic environment. This is in direct contrast to the case of the proximal histidine ligand in peroxidases where the ligand is situated at the C terminus of an α helix and hydrogen bonds with an aspartate. It appears that in these iron–sulfur proteins, the goal is to diminish the negative charge on the sulfur ligand, while in peroxidases the goal is to increase the negative charge on the ligand. In peroxidases it has been argued that the reason for this arrangement is to decrease the redox potential of the heme iron. Using the same argument for sulfur–iron proteins, the local electropositive environment may be necessary to increase the heme iron redox potential.

Another structural feature of heme enzymes that governs biological activity is the accessibility of the active site. It is generally believed that P450s and peroxidases utilize a similar Fe^{4+}=O oxidizing center and that access to this center helps to determine the type of activity that is exhibited. Heme peroxidases have restricted access to the Fe^{4+}=O center; hence they accept electrons from reducing substrates via the heme edge. This leads to the production of substrate radicals. In contrast, P450s have substrate pockets directly adjacent to the Fe^{4+}=O center allowing for stereoselective hydroxylation of substrates. CPO more closely resembles P450s in that a small opening above the heme allows access to the Fe^{4+}=O center. Nevertheless, CPO also exhibits peroxide-degrading activities typical of peroxidases and catalase. Part of the P450-like activity may be controlled by the cysteine ligand as well as accessibility to the Fe^{4+}=O center. The peroxidase-like activity is determined, in part, by the relative stability of CPO compound I compared with P450: the peroxide-binding site is more polar in CPO. The additional stability of CPO compound I allows substrates that cannot approach close enough for direct oxidation by the Fe^{4+}=O center the time to bind and transfer an electron. Much of this is speculative, but the availability of several new P450 and peroxidase structures, and now the CPO structure, should help to provide a much better correlation between known biological activities and structure.

Materials and methods

Protein purification and crystallization

The enzyme was purified from *Caldariomyces fumago* as described by Blanke *et al.* [39]. Details of crystallization and data collection have been reported [40]. Briefly, two orthorhombic crystal forms, one primitive and the other centric, have been grown using polyethylene glycol (PEG) 6000 as the precipitant in the presence of 10 mM potassium phosphate buffer at pH 6.0 by the hanging drop vapor diffusion method at room temperature (~20°C). The primitive form grows at higher concentrations of PEG (30% w/v) than the centric form (20% PEG), but the latter requires the presence of 310 mM NaCl. Both crystal forms, each with a single molecule in the asymmetric unit, were used in the structure determination: $C222_1$ (a=58.43 Å, b=152.76 Å, c=102.29 Å, Z=8) and $P2_12_12_1$ (a=58.77Å, b=71.36 Å, c=91.73 Å, Z=4). The $C222_1$ form, grown using seeding procedures, diffracts to a resolution of about 2.1 Å, while the spontaneously grown $P2_12_12_1$ form diffracts to 1.9 Å.

Table 5. Data collection and phase refinement statistics.

Data set	Resolution (Å)	Reflections		Complete (%)*	R_{sym} †	R_{iso} ‡	No. of sites	R_{Cullis} §	Phasing power# (2.7 Å)
		measured	unique						
Native	1.90	152222	29366	94 (75)	0.082				
$GdCl_3$	2.20	98201	19976	97 (86)	0.070	0.119	1	0.674	1.23
K_2PtCl_4	2.40	86687	15414	99 (94)	0.106	0.147	3	0.638	1.21
HMPS**	2.50	67701	13397	96 (76)	0.105	0.115	3	0.650	0.89
$GdCl_3$ anomalous	2.20	100333	36571	96 (79)	0.054		1		1.22
Native anomalous	2.67	92484	14455	69 (48)	0.039		1		1.68

*The overall completeness is given, with the completeness in the highest resolution shell shown in parentheses. †$R_{sym}=\Sigma_h\Sigma_i|<I(h)>-I(h)_i|/\Sigma_h\Sigma_i I(h)_i$, where I(h) is the intensity of reflection h, Σ_h is the sum over all reflections and Σ_i is the sum over all i measurements of reflection h. ‡$R_{iso}=\Sigma_h|F_{PH}-F_P|/\Sigma_h F_P$, where F_{PH} and F_P are the native and derivative structure-factor amplitudes, respectively. §$R_{Cullis}=\Sigma_h||F_{PH}\pm F_P|-F_H|/\Sigma_h|F_{PH}\pm F_P|$, where F_H is the calculated structure-factor amplitude of the heavy-atom structure. #Phasing power=F_H/E, where E is the estimated lack of closure error. **HMPS, hydroxymercuricphenylsulfonic acid.

Data collection

The details of data collection for native crystals of both space groups and heavy-atom derivatives of the $C222_1$ space group have been reported [40]. Heavy-atom derivatives have been prepared since for the $P2_12_12_1$ form and the data have been collected on a Siemens area detector using a rotating anode X-ray source equipped with focussing optics. The data were processed and scaled using XENGEN [41]. Native anomalous data were collected by swinging the crystal about the b* axis alternately 180° apart at 60° intervals until the complete 360° was swept and scaled without merging Bijvoet pairs. Derivative anomalous data were collected without optimizing for recording Bijvoet pairs as described above, but simply scaling the data without merging the Bijvoet pairs. Native and derivative data sets were collected from single crystals at ambient temperatures. Typically, the derivatives were prepared by soaking crystals in 20 µl artificial mother liquor (30% PEG 6000, with no buffer) containing ~1–4 mM heavy-metal compound for 1–2 days. The data collection details are provided in Table 5.

Structure solution

The heme iron position was located using the native anomalous difference Patterson map and the heavy-atom positions by isomorphous difference Patterson maps and were confirmed by the cross-difference Fourier maps. Multiple isomorphous replacement (MIR) phases were calculated to 2.7 Å in both space groups. The combined figures of merit for the $C222_1$ and $P2_12_12_1$ forms were 0.68 and 0.69, respectively. All the calculations up to this point were performed using PHASES [42]. The phasing statistics are provided in Table 5.

Initially, a spherical mask was constructed around the asymmetric unit from the solvent-flattened map of the $C222_1$ form using MAMA [43] and was later edited to cover the asymmetric unit as completely as possible using bones (or bone images) as a guide in O [44]. The non-crystallographic relation between the two space groups was established by the least-squares superimposition of five common heavy-atom sites and refined using the electron density correlation of the asymmetric units in the two space groups using MAVE [43]. The MIR maps were averaged and solvent flattened followed by histogram matching, and the phases were extended to 2.1 Å using MAGICSQUASH [45] in the CCP4 suite of programs [46]. The averaged maps were of excellent quality and clearly showed the backbone density and many bulky side chains. The polypeptide from residue 5–254 and residue 281–298 was built and the remaining portions of the backbone density were fitted with a polyserine model using FRODO [47].

Refinement

The model was refined at this stage by simulated annealing using X-PLOR [48] to an R-factor of about 0.31 in both space groups. The calculated phases from refined partial models were combined with the MIR phases using SIGMAA [49] in CCP4 and a new round of averaging, solvent flattening, histogram matching and phase extension was performed using a new mask obtained from the model. The maps calculated from the new phase sets showed considerable improvement in the previously poorly defined regions and the remaining portions of the chain could be traced with ease. This exercise was repeated until the complete polypeptide could be traced and all the side chains had been fitted. Once the complete polypeptide was built, no more map averaging between the two space groups was necessary. The models were further improved using omit maps calculated by excluding the regions under examination from refinement and phase calculations, and rebuilding those regions. During this stage of model building, electron density for many carbohydrates sites became clear and these sites were included in the model in a stepwise manner once the density could be interpreted beyond doubt. Ordered solvent molecules were included at this stage in several steps based on the following three criteria: first, there should be 3σ F_o-F_c density and 1σ $2F_o-F_c$ density; second, the solvent peak should be within hydrogen-bonding distance of protein atoms or other solvent atoms, and third, the B-factors should be less than 60 $Å^2$. At the final stages, conventional positional and individual B-factor refinements were carried out using X-PLOR. The quality of the models was assessed using PROCHECK [50].

Atomic coordinates have been submitted to the Brookhaven Protein Data Bank.

Acknowledgements: We would like to thank J Matthew Mauro who first obtained crystals of CPO in our lab, David J Schuller for assistance in using MAGICSQUASH, Liisa Holms (EMBL, Heidelberg, FRG) for carrying out the DALI structural comparison of CPO with available structures in the PDB and Lowell Hager for invaluable advice during the preparation of this manuscript. This work was supported by NSF Grant MCB 9405128 (TLP) and NIH grant GM 34443 (JT).

References

1. Neidleman, S.L. & Geigert, J. (1992). *Biohalogenation: Principles, Basic Roles, and Applications.* Ellis Horwood, Chichester.
2. Morris, D.R. & Hager, L.P. (1966). Chloroperoxidase. I. Isolation and properties of the crystalline glycoprotein. *J. Biol. Chem.* **241**, 1763–1768.
3. Griffin, B.W. (1991). Chloroperoxidase: a review. In *Peroxidases in Chemistry and Biology.* (Everse, J., Everse, K.E. & Grisham, M.B., eds), vol. **II**, pp. 85–137, CRC Press, Boca Raton, FL.
4. Dawson, J.H. (1988). Probing structure–function relations in heme-containing oxygenases and peroxidases. *Science* **240**, 433–439.
5. Dawson, J.H. & Sono, M. (1987). Cytochrome P450 and chloroperoxidase: thiolate-ligated heme enzymes. Spectroscopic determination of their active site structures and mechanistic implications of thiolate ligation. *Chem. Rev.* **87**, 1255–1276.
6. Blanke, S.R. & Hager, L.P. (1989). Chemical modification of chloroperoxidase with diethylpyrocarbonate. *J. Biol. Chem.* **265**, 12454–12461.
7. Samokyszyn, V.M. & Ortiz de Montellano, P.R. (1991). Topology of the chloroperoxidase active site: regiospecificity of heme modification by phenylhydrazine and sodium azide. *Biochemistry* **30**, 11646–11653.
8. Dugad, L.B., Wang, X., Wang, C.-C., Lukat, G.S. & Goff, H.M. (1992). Proton nuclear Overhauser effect study of the heme active site structure of chloroperoxidase. *Biochemistry* **31**, 1651–1655.
9. Finzel, B.C., Poulos, T.L. & Kraut, J. (1984). Crystal structure of yeast cytochrome *c* peroxidase refined at 1.7 Å resolution. *J. Biol. Chem.* **259**, 13027–13036.
10. Poulos, T.L., Finzel, B.C. & Howard, A.J. (1987). High-resolution crystal structure of cytochrome P450cam. *J. Mol. Biol.* **195**, 687–700.
11. Ramachandran, G.N., Ramakrishnan, C. & Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**, 95–99.
12. Fang, G-H., Axley, M.J., Nuell, M. & Hager, L.P. (1986). Cloning and sequencing of chloroperoxidase cDNA. *Nucleic Acids Res.* **14**, 8061–8071.
13. Keningsberg, P., Fang, G.-H. & Hager, L.P. (1987). Post-translational modifications of chloroperoxidase from *Caldariomyces fumago. Arch. Biochem. Biophys.* **254**, 409–415.
14. Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123–138.
15. Blanke, S.R. & Hager, L.P. (1988). Identification of the fifth axial heme ligand of chloroperoxidase. *J. Biol. Chem.* **263**, 18739–18743.
16. Richardson, J. (1981). The anatomy and taxonomy of protein structure. *Adv. Protein Chem.* **34**, 167–339.
17. Ravichandran, K.G., Boddupalli, S.S., Hasemann, C.A., Peterson, J.A. & Deisenhofer, J. (1993). Crystal structure of hemoprotein domain of P450BM-3, a prototype for microsomal P450's. *Science* **261**, 731–736.
18. Hasemann, C.A., Ravichandran, K.G., Peterson, J.A. & Deisenhofer, J. (1995). Crystal structure and refinement of cytochrome P450terp at 2.3 Å resolution. *J. Mol. Biol.* **236**, 1169–1185.
19. Li, H. & Poulos, T.L. (1995). Modelling protein–substrate interactions in the heme domain of cytochrome P450BM-3. *Acta Cryst. D* **51**, 21–32.
20. Cupp-Vickery, J.R. & Poulos, T.L. (1995). Structure of cytochrome P450eryF involved in erythromycin biosynthesis. *Nat. Struct. Biol.* **2**, 144–153.
21. Richardson, J.S. & Richardson, D.C. (1988). Amino acid preferences for specific locations at the ends of alpha helices. *Science* **240**, 1648–1652.
22. Adman, E., Watenpaugh, K.D. & Jensen, L.H. (1975). NH–S hydrogen bonds in *Peptococcus aerogenes* ferredoxin, *Clostridium pasteurianum* rubredoxin, and *Chromatium* high potential iron protein. *Proc. Natl. Acad. Sci. USA* **72**, 4854–4858.
23. Georgiadis, M.M., Komiya, H., Chakrabarti, P., Woo, D., Kornuc, J.J. & Rees, D.C. (1992). Crystallographic structure of the nitrogenase iron protein from *Azotobacter vinelandii. Science* **257**, 1653–1659.
24. Cramer, S.P., Dawson, J.H., Hodgson, K.O. & Hager, L.P. (1978). Studies on the ferric forms of cytochrome P450 and chloroperoxidase by extended X-ray absorption fine structure. Characterization of the Fe–N and Fe–S distance. *J. Am. Chem. Soc.* **100**, 7282–7290.
25. Hollenberg, P.F. & Hager, L.P. (1973). The P450 nature of the carbon monoxide complex of ferrous chloroperoxidase. *J. Biol. Chem.* **248**, 2630–2062.
26. Thomas, J.A. (1968). Studies on the mechanism of the halogenation reactions catalyzed by chloroperoxidase [PhD Thesis]. University of Illinois.
27. Sundaramoorthy, M., Kishi, K., Gold, M.H. & Poulos, T.L. (1994). The crystal structure of manganese peroxidase from *Phanerochaete chrysosporium* at 2.06 Å resolution. *J. Biol. Chem.* **269**, 32759–32767.
28. Kuwahara, M., Glenn, J.K., Morgan, M.A. & Gold, M.H. (1984). Separation and characterization of two extracellular H_2O_2-dependent oxidases from ligninolytic cultures of *Phanerochaete chrysosporium. FEBS Lett.* **169**, 247–250.
29. Murthy, M.R.N., Reid, T.J., Sicignano, A., Tanaka, N. & Rossmann, M.G. (1981). Structure of beef liver catalase. *J. Mol. Biol.* **152**, 465–499.
30. Poulos, T.L. (1987). Heme enzyme crystal structures. *Adv. Inorg. Biochem.* **88**, 1–36.
31. Sono, M., Dawson, J.H., Hall, K. & Hager, L.P. (1986). Ligand and halide binding properties of chloroperoxidase: peroxidase-type active site heme environment with cytochrome P450 type endogenous axial ligand and spectroscopic properties. *Biochemistry* **25**, 347–356.
32. Edwards, S.L., Nguyen, H.X., Hamlin, R.C. & Kraut, J. (1987). Crystal structure of cytochrome *c* peroxidase compound I. *Biochemistry* **26**, 1503–1511.
33. Rutter, R. & Hager, L.P. (1982). The detection of two electron paramagnetic resonance radical signals associated with chloroperoxidase compound I. *J. Biol. Chem.* **257**, 7958–7961.
34. Kedderis, G.L. & Hollenberg, P.F. (1983). Steady state kinetics of chloroperoxidase-catalyzed *N*-demethylation reactions. *J. Biol. Chem.* **258**, 12413–12419.
35. Paulsen, M.D. & Ornstein, R.L. (1995). Dramatic differences in the motions of the mouth of open and closed cytochrome P450BM-3 by molecular dynamics simulations. *Proteins* **21**, 237–243.
36. Casella, L., Poli, S., Gullotti, M., Selvaggini, C., Beringhelli, T. & Marchesini, A. (1994). The chloroperoxidase-catalyzed oxidation of phenols. Mechanism, selectivity, and characterization of enzyme–substrate complexes. *Biochemistry* **33**, 6377–6386.
37. Aleshin, A.E., Hoffman, C., Firsov, L.M. & Honzatko, R.B. (1994). Refined crystal structures of glucoamylase from *Aspergillus awamori* var. X100. *J. Mol. Biol.* **238**, 575–591.
38. Einsphar, H.M. (1989). Protein–carbohydrate interactions in biological systems. *Trans. Am. Cryst. Assoc.* **25**, 1–22.
39. Blanke, S.R., Yi, S. & Hager, L.P (1989). Development of semi-continuous and continuous flow bioreactors for the high level production of chloroperoxidase. *Biotechnol. Lett.* **11**, 769–774.
40. Sundaramoorthy, M., Mauro, J.M., Sullivan, A.M., Terner, J. & Poulos, T.L. (1995). Preliminary crystallographic analysis of chloroperoxidase from *Caldariomyces fumago. Acta Cryst. D* **51**, 842–844.
41. Howard, A.J., Gilliland, G.L., Finzel, B.F., Poulos, T.L., Olendorf, D.H. & Salemme, F.R. (1987). The use of an imaging proportional counter in macromolecular crystallography. *J. Appl. Cryst.* **20**, 383–387.
42. Furey, W. & Swaminathan, S. (1990). PHASES — a program package for the processing and analysis of diffraction data for macromolecules. American Crystallography Association Meeting Abstracts **18**, 73.
43. Kleywegt, G.J. & Jones, T.A. (1994). Halloween...masks and bones. In *From First Map to Final Model.* (Bailey, S., Hubbard, R. & Waller, D., eds), pp. 59–66, SERC Daresbury Laboratory, Warrington, UK.
44. Jones, T.A., Zou, J.-Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst. A* **47**, 110–119.
45. Schuller, D.J.M. (1996). MAGICSQUASH: More versatile noncrystallographic averaging with multiple constraints. *Acta Cryst. D*, in press.
46. Collaborative Computational Project Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst. D* **50**, 760–763.
47. Jones, T.A. (1978). A graphics model building and refinement system for macromolecules. *J. Appl. Cryst.* **11**, 268–272.
48. Brünger, A.T. (1992). *X-PLOR Manual, Version 3.1: A System for X-ray Crystallography and NMR.* Yale University Press, New Haven, CT.
49. Read, R.A. (1986). Improved coefficients for map calculation using partial structures with errors. *Acta Cryst. A* **42**, 140–149.
50. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291.
51. Luzzati, V. (1952). Traitement statistique des erreurs dans la determination des structures cristallines. *Acta Cryst.* **5**, 802–810.
52. Evans, S.V. (1993). SETOR: hardware-lighted three dimensional solid model representation of macromolecules. *J. Mol. Graphics* **11**, 134–138.
53. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structure. *J. Appl. Cryst.* **24**, 946–950.
54. Nicholls, A., Sharp, K.A. & Honig, B. (1991). Protein folding and association; insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296.

Received: **19 Sep 1995**; revisions requested: **10 Oct 1995**; revisions received: **27 Oct 1995**. Accepted: **30 Oct 1995**.