Crystal structure of the di-haem cytochrome *c* peroxidase from *Pseudomonas aeruginosa*

Vilmos Fülöp[1]*, Christopher J Ridout[2], Colin Greenwood[2] and János Hajdu[1]

[1]Laboratory of Molecular Biophysics and Oxford Centre for Molecular Sciences, University of Oxford, South Parks Road, Oxford OX1 3QU, UK and [2]School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK

Background: Cytochrome *c* peroxidase from *Pseudomonas aeruginosa* (PsCCP) represents a new class of peroxidases which work without the need to create a semi-stable free radical for catalysis. The enzyme is located in the bacterial periplasm where its likely function is to provide protection against toxic peroxides. The soluble 323-residue single polypeptide chain contains two covalent *c*-type haems with very different properties: one of them is a low-potential (–330 mV) centre where hydrogen peroxide is reduced (the peroxidatic site); the other is a high-potential (+320 mV) centre which feeds electrons to the peroxidatic site from soluble electron-shuttle proteins such as cytochrome *c* and azurin.

Results: The crystal structure of the oxidized form of PsCCP has been determined to 2.4 Å resolution by multiple isomorphous replacement, and refined to an R-factor of 19.2%. PsCCP is organized into two domains, both of them containing a covalent *c*-haem in a structure reminiscent of class 1 cytochromes *c*. The domains are related by a quasi-twofold axis. The domain interface holds a newly discovered calcium-binding site with an unusual set of ligands.

Conclusions: The likely function of the calcium site is to maintain the structural integrity of the enzyme and/or to modulate electron transfer between the two haem domains. The low-potential haem has two histidine axial ligands (His55 and His71) and the high-potential haem is ligated by His201 and Met275. There are no polar residues at the peroxidatic site in the inactive oxidized enzyme. The structure suggests that, in the half-reduced functional form of the enzyme, the low-potential haem has to shed His71 in order to make the enzyme catalytically competent. This process is likely to trigger a reorganization of the active site, and may introduce new residues into the haem pocket.

Structure 15 November 1995, **3**:1225–1233

Key words: calcium-binding site, cytochromes, di-haem cytochrome *c* peroxidase, *Pseudomonas aeruginosa*, X-ray crystallography

Introduction

Cytochrome *c* peroxidase (EC 1.11.1.5) from *Pseudomonas aeruginosa* (PsCCP) catalyzes the oxidation of ferrocytochrome c_{551} and azurin(Cu^{I}) by peroxide in a reaction resembling that catalyzed by cytochrome *c* peroxidases from yeast (YCCP) and horseradish (HRP). The likely function of these enzymes is to protect against the build-up of toxic peroxides. The bacterial enzyme, however, differs from the yeast and horseradish peroxidases in that it does not require a stable free radical for catalysis and instead uses two haem *c* groups [1] covalently attached to a single polypeptide chain to store the two oxidizing equivalents in the high-energy high-oxidation (compound I) form of the enzyme. A number of studies have demonstrated substantial differences between the two haem centres of PsCCP. The midpoint potentials of the two haems lie 650 mV apart at +320 mV (high-potential haem) and –330 mV (low-potential haem) [2], the latter value being similar to the redox potential of other peroxidases. This indicates that the low-potential haem centre is the site of hydrogen peroxide reduction.

Whereas the ferric form of mono-haem peroxidases reacts rapidly with hydrogen peroxide to generate the relatively stable Fe^{IV} ferryl oxidation state, fully ferric PsCCP does not react with hydrogen peroxide. We have shown that PsCCP is potentiated for reaction with hydrogen peroxide when the high-potential haem group is reduced, an event which results in the low-potential haem shedding a histidine ligand to become high spin [3]. We have further characterized the activation and catalytic cycle spectroscopically [4] and kinetically [5]. Figure 1 shows the major steps in the reaction.

In order to store the two equivalents of oxidizing capacity available from peroxide, single-haem proteins generate an oxyferryl group and also use non-metal-based centres. In HRP, an electron is removed from the porphyrin ring to form a haem radical [6], whilst in YCCP a semi-stable radical is formed on a tryptophan residue [7–10]. The presence of two haem *c* groups per unit molecular weight in PsCCP obviates the need for stabilizing a radical species in the mechanism, and allows a faster than usual turnover for the enzyme. The nature of the low-potential haem pocket of PsCCP has been investigated using nanosecond laser photolysis of a range of ligands [11]. Three straight-chain isonitriles show about 50% geminate recombination with half-times of the order of 10 ns. *t*-Butyl isonitrile shows more and faster recombination. These results imply considerable freedom of movement within the active site for the smaller ligands.

Here we present the first three-dimensional structure of a di-haem cytochrome *c* peroxidase. The enzyme was isolated from *P. aeruginosa* and the structure was solved in its fully oxidized ferric–ferric form at 2.4 Å resolution.

*Corresponding author.

 ISSN 0969-2126

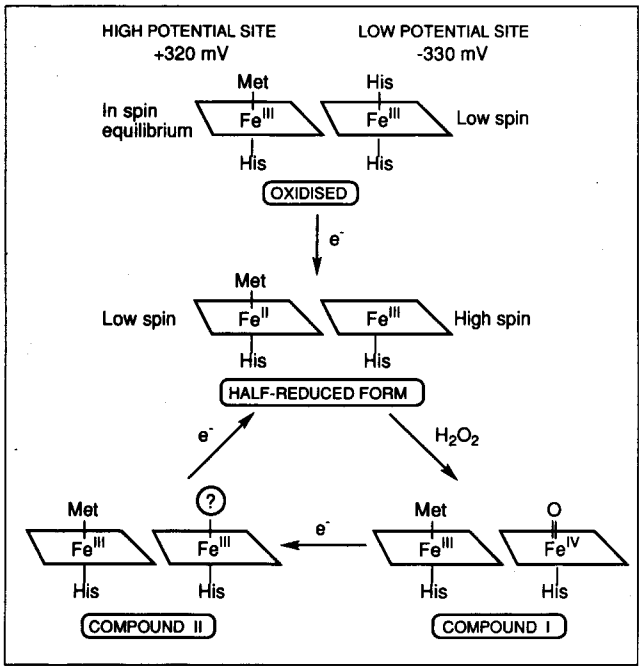

Fig. 1. Major steps in the activation and catalytic cycle of cytochrome *c* peroxidase from *P. aeruginosa*. The scheme is based on results from biochemical and spectroscopic studies [4]. The enzyme can accept electrons from both ferrocytochrome c_{551} or azurin(Cu^I).

Results and discussion

The primary structure of PsCCP was originally determined from cyanogen bromide fragments, and reported to contain 302 amino acids [12]. Electrospray ionization mass spectrometry subsequently revealed that the molecular weight was 2600 Da heavier than that calculated from the published primary amino acid sequence. Because of substantial peak profile broadening, this was first interpreted as indicative of protein glycosylation [13], although subsequent chemical analysis could not confirm this. Detailed electron density analysis suggested the presence of an additional loop consisting of approximately 20 amino acids close to one of the cyanogen bromide cleavage sites. Translation of the gene sequence [14] confirmed the insertion of 20 residues and revealed further minor errors in the original protein sequence (Fig. 2).

Architecture of the molecule

The crystal structure of the di-haem PsCCP was determined by multiple isomorphous replacement (MIR), and has been refined against diffraction data to 2.4 Å Bragg spacings (Table 1, Figs 3,4a). The protein molecules form dimers both in solution [13] and in the crystal (Fig. 4b), and the crystallographic asymmetric unit contains a monomer. The buried surface area between the monomers is 1277 $Å^2$ (8.3% of the surface area of the monomer). The subunit interactions are mainly hydrophobic, involving residues 39–66 and 303–321. The side of the high-potential haem domain opposite to the dimer interface is also hydrophobic but is exposed to the solvent. The high-potential haem pierces this surface and is accessible from the solvent (Fig. 4b).

The monomer is organized into two structurally distinct domains around the two haem centres (Figs 3 and 5a,b). A pseudo-twofold axis runs across the molecule, mapping the two haems on each other together with helices α2, α4, α7 and strand β3 onto α9, α10, α11 and strand β5, respectively. The surface area buried between the two domains is 1326 $Å^2$ (8.7% of the surface area of the monomer), similar to the buried area between the two monomers. The domain interface is hydrophobic and holds an unexpected calcium-binding site with unusual

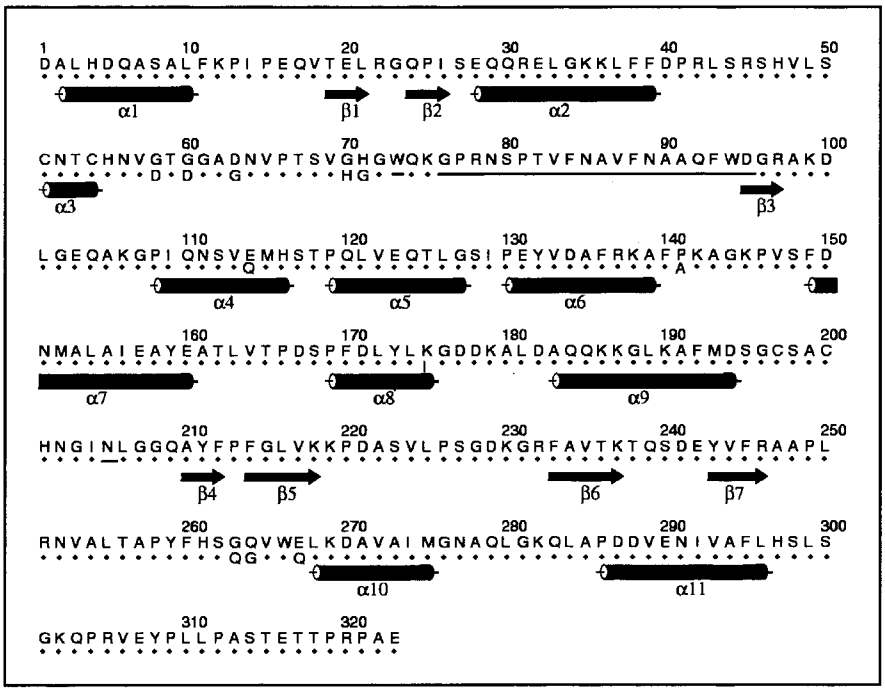

Fig. 2. Amino acid sequence and secondary structure in cytochrome *c* peroxidase from *P. aeruginosa*. The sequence deduced from the gene [14] is shown in the top row. Shown beneath are those residues which were different in the sequence determined from cyanogen bromide fragments [12]. Underlining denotes residues previously missing from the protein sequence.

Table 1. Data collection and phasing statistics.

Crystal	Resolution (Å)	Observations	Unique reflections	Completeness (%)	R_{merge}* (%)	R_{iso}† (%)	No. of heavy atom sites	Phasing power‡ acentric	Phasing power‡ centric	Cullis R§ acentric	Cullis R§ centric
Native	2.4	80795	20541	96	7.9						
PCMBS#	4.0	8621	3484	75	5.7	12.0	6	0.7	1.3	0.7	0.6
$Pb(Ac)_2$	4.0	5254	3438	74	6.0	10.2	2	1.2	1.0	0.8	0.7
$(Met)_3PbAc$	4.0	7174	3768	80	4.0	10.0	2	1.3	1.0	0.7	0.6

*$R_{merge}=\Sigma_j\Sigma_h|I_{h,j}-\langle I_h\rangle|/\Sigma_j\Sigma_h\langle I_h\rangle\times100$, where $I_{h,j}$ is the jth observation of reflection h, and $\langle I_h\rangle$ is the mean intensity of that reflection. †$R_{iso}=\Sigma||F_{PH}|-|F_P||/\Sigma|F_P|\times100$, where F_P and F_{PH} refer to the native and the derivative structure-factor amplitude. ‡Phasing power=rms ($|F_H|/E$), where $|F_H|$=heavy-atom structure-factor amplitude and E=rms lack of closure. §Cullis R=lack of closure/isomorphous difference. #PCMBS=*para*-chloromercuri-benzene-sulphonate.

ligation properties. Part of the interface contains those amino acids (residues 76–95) which were missing from the originally determined protein sequence, and includes one of the three protein ligands of the calcium (Asn79).

Domain structure

Both domains of PsCCP resemble the structure of class 1 *c*-type cytochromes for which many structures are known, including those of tuna cytochrome *c* and cytochrome c_{551} from *P. aeruginosa* [15] (Fig. 5). Both the three-dimensional arrangement and connectivity of the major helical elements are very similar: α2 and α9 correspond to the N-terminal helices, and α7 and α11 correspond to the C-terminal helices of class 1 *c*-type cytochromes. However, both domains are somewhat larger than usual cytochromes, containing additional α helices, β strands and loops.

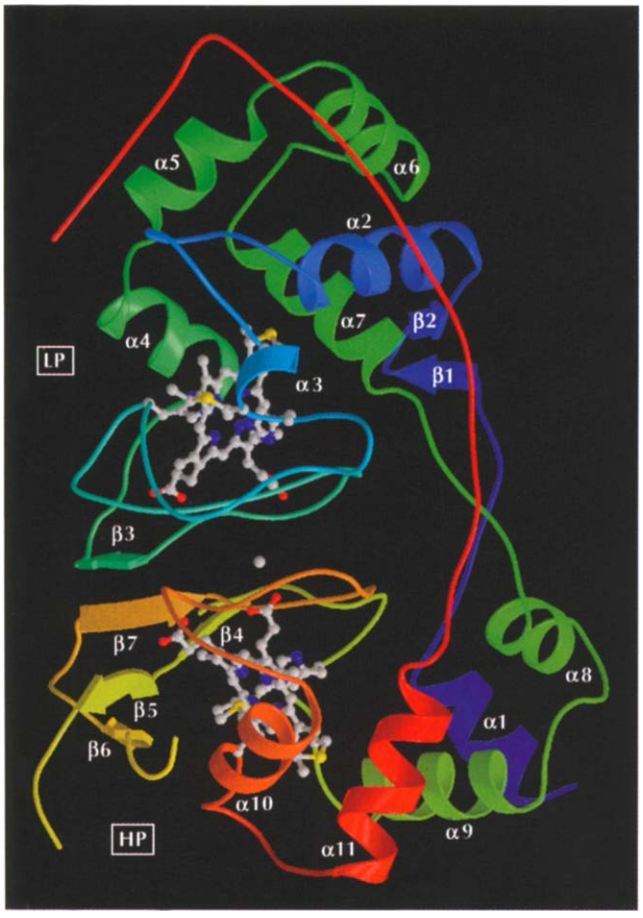

Fig. 3. The structure of cytochrome *c* peroxidase from *P. aeruginosa*. The ribbon diagram is colour-ramped blue to red from the N to the C terminus. The small grey sphere shows the location of the calcium ion and the haems are in ball-and-stick representation. LP and HP denote low-potential and high-potential haem domains, respectively.

Low-potential haem domain

Residues 17–164 form the N-terminal domain, containing the low-potential haem (Fig. 5a). Residues 51–55 correspond to the sequence fingerprint (Cys-X-Y-Cys-His) for the *c*-type cytochromes. The crystallized enzyme is in the fully oxidized state and shows that the low-potential haem iron has two axial ligands, in agreement with spectroscopic measurements [4]. Both ligands are histidines (His55 and His71; Figs 5a,6a), as in the *c*-domain of cytochrome cd_1 nitrite reductase [16]. The immediate environment of this haem is predominantly hydrophobic. In *c*-type cytochromes, a proline residue is positioned so that its peptide carbonyl forms a hydrogen bond with the hydrogen on the Nδ atom of the histidine coordinated to the haem iron and acts to orient the histidine ring. Both histidines of the low-potential haem interact via similar hydrogen bonds with a proline: His55 with Pro81, and His71 with Pro108 (Fig. 6a).

For catalysis to take place, one of the two histidine ligands must leave the low-potential haem iron to allow peroxide binding. His55 is unlikely to leave because it is buried and positioned next to Cys54 which covalently links the haem to the protein in a rigid structural unit. His71, on the other hand, sits on the tip of a flexible loop (residues 67–78) on the accessible surface of the molecule, not far from the crystallographic dimer interface (Fig. 4b). This arrangement suggests that His71 is likely to be the ligand that leaves the haem prior to catalysis. Activation of the enzyme might be coupled to domain movement controlled by changes in the spin and oxidation state of the two haems, as has been suggested for cytochrome cd_1 [16]. Further rearrangements may be required around the active site as there are no charged residues in the haem pocket. This is in contrast

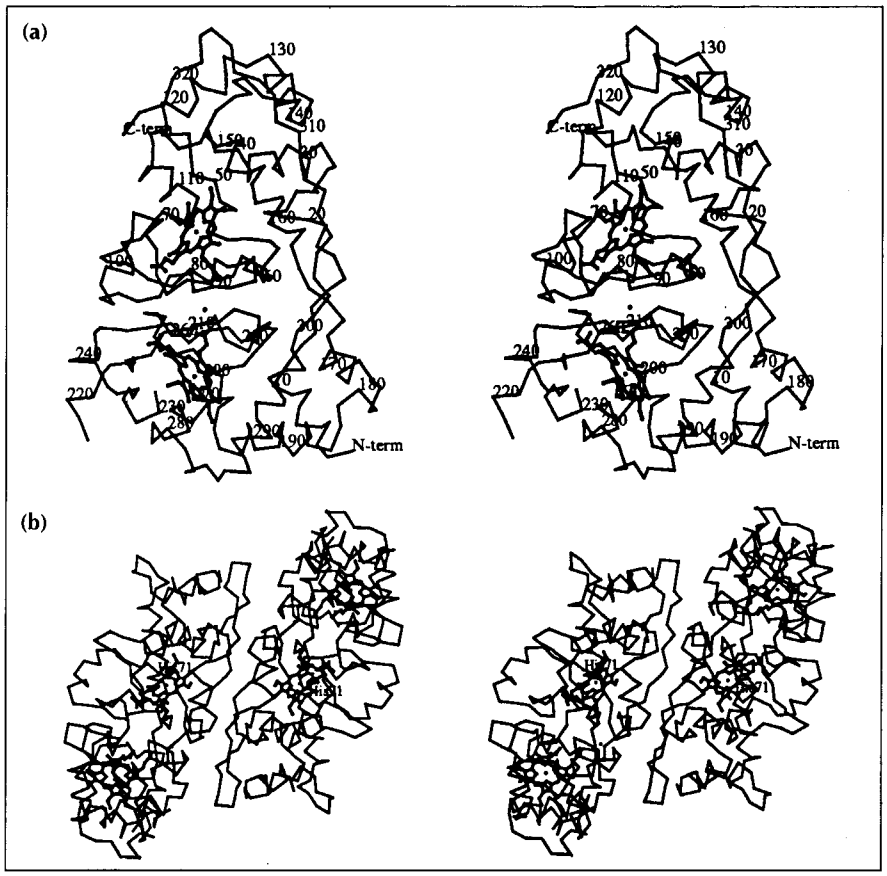

Fig. 4. The structure of the monomer and the dimer of PsCCP. **(a)** Stereo diagram of the Cα trace of PsCCP with every tenth Cα position labelled. The molecule is shown in the same view as in Figure 3. **(b)** Stereo diagram of the crystallographic dimer, looking down the twofold axis.

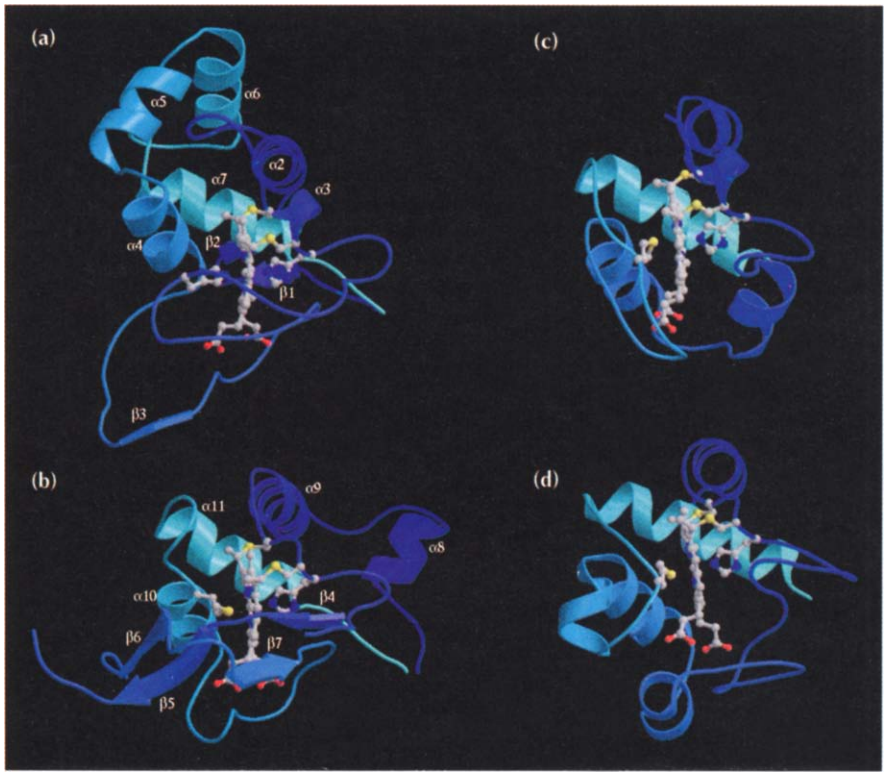

Fig. 5. Comparison of the fold of the *c*-type cytochrome domains of PsCCP with the fold of typical class 1 cytochromes *c*. **(a)** High-potential haem domain of PsCCP. **(b)** Low-potential haem domain of PsCCP. **(c)** Cytochrome c_{551} from *P. aeruginosa* (PDB [23] entry 351C). **(d)** Tuna cytochrome *c* (PDB entry 3CYT). The cytochromes are colour-ramped dark blue to light blue from the N to the C terminus. The PsCCP domains are coloured similarly. The haems are shown in ball-and-stick representation.

to YCCP where Arg48 and His52 are in the haem pocket and play a role in catalysis [7,10,17–20]. Once His71 has been removed as a ligand to the low-potential haem iron it may then serve to polarize the oxygen–oxygen bond of the peroxide molecule and may also permit the entry of other residues to the haem pocket. The inactive oxidized enzyme form and compound II are isoelectronic but the iron ligands are likely to be different (Fig. 1). We postulate that rebinding of His71 to the haem iron is a slow process.

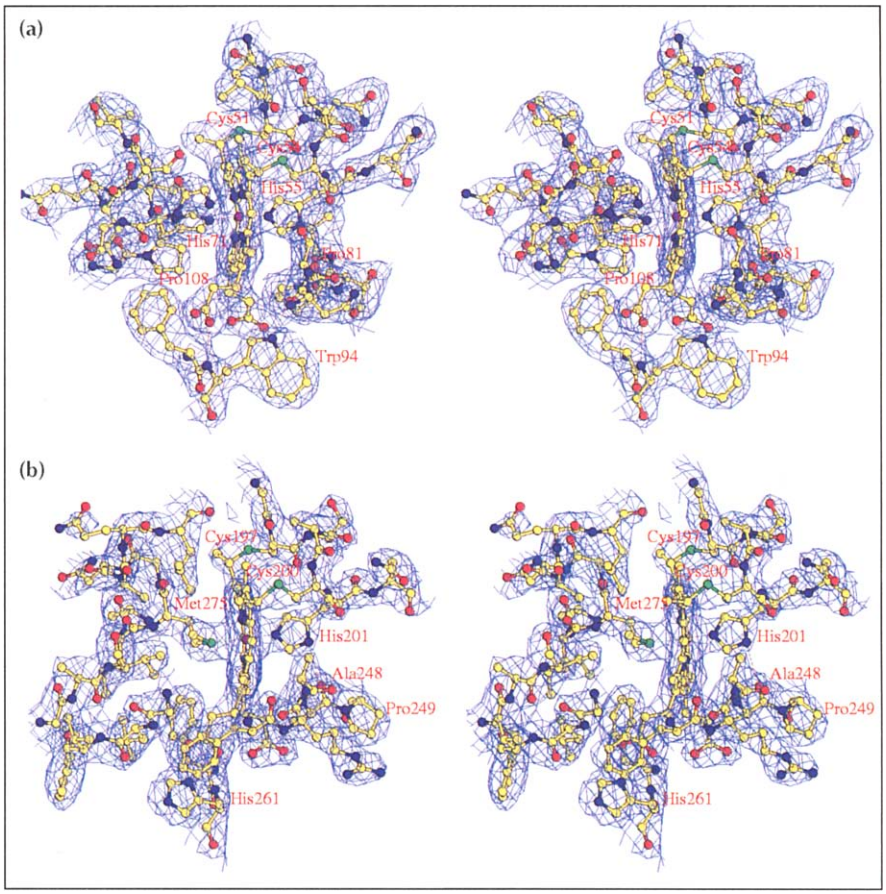

Fig. 6. Stereoview of the final electron-density map around the two haems. SIGMAA-weighted [45] $2mF_o$–DF_c electron density using phases from the final model is contoured at 1σ level, where σ represents the root mean square electron density for the unit cell. Contours more than 1.5 Å from any of the displayed atoms have been removed for clarity. **(a)** The low-potential haem is covalently attached to Cys51 and Cys54. The His55 Nε2–Fε and the His71 Nε2–Fε bond distances in the refined model are 1.96 Å and 1.98 Å, respectively. The hydrogens on the Nδ atoms of His55 and His71 are hydrogen bonded with the carbonyl oxygens of Pro81 and Pro108, respectively. Trp94 lies between the low-potential and high-potential haem centres and is believed to be involved in electron transfer between them (see also Fig. 7b). **(b)** The high-potential haem is covalently bound to Cys197 and Cys200. The His201 Nε2–Fe distance is 1.97 Å, and the Met275 Sδ–Fe distance in the refined model is 2.30 Å. The hydrogen on the Nδ atom of His201 is hydrogen bonded with the carbonyl oxygen of Ala248. This interaction is reminiscent of the His–Pro interactions at the low-potential haem (see also text). His261 is believed to be involved in electron transfer between the two haems (see Fig. 7b).

High-potential haem domain

Residues 165–302 form the C-terminal domain, containing the high-potential haem (Fig. 5b), the site of electron entry from donors [4]. The haem is attached to residues Cys197 and Cys200, and has histidine (His201) and methionine (Met275) ligands (Figs 5b,6b), similar to class 1 cytochromes *c* (His18 and Met80 in tuna cytochrome *c*). The environment of the haem pocket is predominantly hydrophobic. The position of Pro249 is conserved in cytochromes (Pro30 in tuna cytochrome *c*). The direction of the chain containing this residue is opposite to that found in *c*-type cytochromes, putting the main-chain carbonyl of the adjacent alanine residue (Ala248) within hydrogen-bonding distance of the hydrogen on the Nδ atom of His201. This interaction maintains the correct orientation of the histidine ring to the haem iron. Residues 223–228 are disordered and were not included in the final model. Although these residues are the target for proteolytic cleavage by *Pseudomonas* elastase [21], we believe that the molecules are intact in the crystal because dissolved PsCCP crystals produce a single band when analyzed by SDS-PAGE. Clipped molecules lose their activity and do not form crystals.

Domain–domain interactions

The two domains are connected by three lengths of extended chain (Fig. 3): residues 1–19, 160–169 and 296–323 (see Fig. 2). In addition to the connecting polypeptide chains, the domains within a molecule interact in two different ways. First, the interface region involves many hydrogen bonds between the two domains, with β3 from the low-potential haem domain forming a three-stranded antiparallel sheet with β7 and β4 from the high-potential haem domain. Second, and unexpectedly, the interface holds a calcium site.

Metal-binding site

The available evidence suggests that the bound ion is calcium and its location in the overall structure is shown in Figure 3, with details of the binding site in Figure 7a. Crystals of the enzyme were obtained from a low ionic strength buffer without the explicit addition of calcium [13] and no calcium was added during protein purification. By modelling the full complement of Ca^{2+} electrons, F_o–F_c electron-density maps became flat in this region. Moreover, F_o–F_c omit maps exhibit very strong difference density at the presumed calcium site, similar to the phase-improved MIR maps used for model building. Finally, all the ligands are oxygen atoms and there are seven of them in a pentagonal bipyramidal arrangement, reminiscent of calcium sites in other proteins.

What is rather unusual, but not completely unprecedented, is the absence of any negatively charged residue coordinated to the calcium. An ASSAM search [22] of the Brookhaven Protein Data Bank [23] showed a hexa-coordinated calcium with somewhat similar ligation in human phospholipase A_2 [24]. In the case of PsCCP, the seven ligands are the amide oxygen of Asn79, the main-chain carbonyls of Thr256 and Pro258 and four water

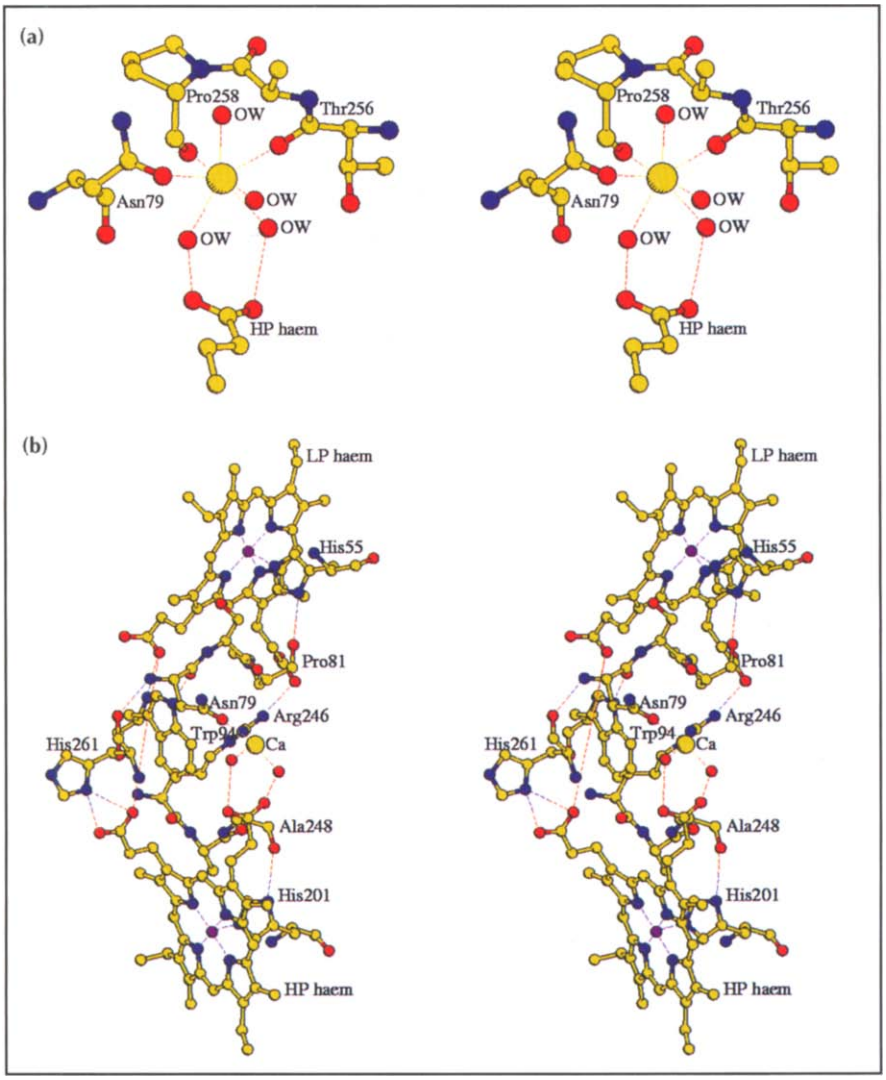

Fig. 7. The calcium-binding site and possible routes for electron transfer in PsCCP. **(a)** Stereoview of the calcium site. The ligand distances to the calcium ion are: Asn79 Nδ1 2.29 Å; Thr256 carbonyl O 2.28 Å; Pro258 carbonyl O 2.20 Å; and for the four water (OW) molecules ~2.29 Å. One of the propionates of the high-potential haem is hydrogen bonded to two bound water molecules. **(b)** Stereoview of the residues most likely to be involved in the electron transfer between the two haems.

molecules. The water molecules could be clearly located in difference density maps. The nearest carboxylate to the calcium ion is one of the propionates of the high-potential haem centre whose oxygen atoms form hydrogen bonds to two bound water ligands (Figs 6a,7a). Although the significance of calcium to PsCCP activity has not been reported, it is known that calcium ions are important for activity in HRP [25] and in the di-haem *Paracoccus* cytochrome *c* peroxidase [26]. They also play an important role in maintaining the integrity of the active site of lignin peroxidase [27]. The calcium ion in PsCCP is located at the interface between the two haem domains. The likely function of this site is to maintain the structural integrity of the enzyme and/or to permit allosteric regulation through the fine-tuning of the haem redox potentials across the interface. The lack of compensating negative charges around the calcium ion may have a functional significance in modulating electron transfer rates between the two haem domains.

Haem position and environment

The two haem groups are located in the hydrophobic pockets of the two domains. They are perpendicular to each other and separated by an iron–iron distance of 20.9 Å. One of the propionates of one haem points towards that of the second. The shortest distance between the two haems is via these propionates (9.6 Å). The closest iron–iron distance between symmetry-related monomers in the dimer is 26.7 Å.

Electron transfer between groups separated by more than 10 Å is possible in proteins [28]. An electron-tunnelling theory has been developed to explain these transfers [29–31]. The theory supports the suggestion that transfer of electrons from the donor to the acceptor may involve several jumps between the orbitals of the separating atoms. Jumps through covalently or hydrogen bonded atoms are more likely than between non-bonded atoms. The ease of electron transfer is governed by the chemical properties of the intervening residues and the medium in the following order: aromatic amino acids > non-polar aliphatic amino acids > polar amino acids > water [32].

Within the framework of this theory, we suggest four possible routes for electron transfer from the high-potential haem to the low-potential haem (Fig. 7b). The first path runs from His201 of the high-potential haem via

the protein backbone (between residues Ala248–Arg246) to reach the low-potential haem propionate. The second path is from one of the high-potential haem propionates via the calcium ion and the protein backbone (residues Asn79–Pro81) to reach His55 of the low-potential haem. In the third path, His261 connects the other propionate of the high-potential haem with Asn79 to join the second path. The shortest distance between the two haems (9.6 Å) is via the two haem propionates and runs through Trp94 which lies in the same plane as the high-potential haem, perpendicular to the low-potential haem. The position of this tryptophan is conserved among the mitochondrial cytochromes *c* (e.g. Trp59 in tuna cytochrome *c*) and in most class 1 cytochromes *c* [15]. A tryptophan residue has been proposed to act as an electron transfer mediator between putidaredoxin and cytochrome P-450$_{cam}$ [33]. Similarly, the shortest straight line between the two haems in the YCCP–cytochrome *c* complex runs through Trp191 [34]. In the case of PsCCP, the shortest and most direct route of electron transfer is via Trp94 (Fig. 7b). This is therefore likely to be the major line of communication between the two haems. The peptide nitrogen of Trp94 is within hydrogen-bonding distance of the propionate of the low-potential haem. Note that the orientation of the haems in the YCCP–cytochrome *c* complex and in PsCCP is different.

Relationships with other proteins

The structure of PsCCP is similar to that of the membrane-associated di-haem cytochrome c_4. Cytochrome c_4 of *P. aeruginosa* is also a two-domain molecule, each domain containing a covalently bound *c*-type haem and resembling the class 1 cytochrome fold (see [15]). The two haems in cytochrome c_4 also differ spectroscopically and possess different redox properties. In both PsCCP and cytochrome c_4, the two domains are connected by a long polypeptide chain and related by a rotation of ~180°. Another protein of similar domain structure and symmetry is the cytochrome subunit of flavocytochrome *c* sulfide dehydrogenase [35]. The two haems in this cytochrome are linked by hydrogen bonding through a pair of propionic acids. All these di-haem cytochrome *c* molecules might have evolved from mono-haem cytochromes through gene duplication accompanied by gene fusion. Although the cytochrome domains of these molecules exhibit class 1 structure, there is no similarity between their sequences.

Biological implications

Cytochrome *c* peroxidase from *Pseudomonas aeruginosa* (PsCCP) is located in the bacterial periplasm where its likely function is to provide protection against toxic peroxides. The protein contains two haem groups covalently attached to a single polypeptide chain. The chain is folded into two very similar domains whose overall structure is reminiscent of class 1 cytochromes *c*. The redox potentials of the haems are markedly different.

The high-potential haem domain is the site of entry for electrons from soluble electron-shuttle proteins such as cytochrome *c* or azurin. The other domain contains the low-potential haem. This is the site of peroxide break down (the peroxidatic site). The domains are related by a quasi-twofold axis, and the domain interface holds a newly discovered calcium-binding site. The likely function of the calcium site is to maintain the structural integrity of the enzyme and/or to modulate electron transfer between the two haem domains. The calcium-binding site exhibits unusual ligation properties with no compensating negatively charged ligands and shows some similarities to the calcium site in human phospholipase A_2. It is also known that calcium plays an important, if as yet ill-defined, role in other peroxidases.

In contrast to single-haem cytochrome *c* peroxidases, PsCCP works without the need for the creation of a stable free radical during the catalytic cycle. It thus represents a new class of peroxidases. The presence of two redox-active haem centres in PsCCP allows the second oxidizing equivalent to be stored in the high-potential haem which switches between Fe^{II} and Fe^{III} forms during the reaction, while the low-potential haem (i.e. the site of peroxide break down) alternates between Fe^{III} and Fe^{IV}. This arrangement obviates the need for storing (and stabilizing) a radical in the protein. This property may lead to a longer-lived protein which is likely to be less affected by mutations than radical proteins.

The structure described in this paper is that of the inactive, fully ferric form of the enzyme, which does not react with hydrogen peroxide. For catalysis to begin, a half-reduced functional form of PsCCP needs to be created. In the inactive fully ferric form of the enzyme, the low-potential haem has two histidine axial ligands (His55 and His71) and the high-potential haem is ligated by His201 and Met275. There are no polar residues at the peroxidatic, low-potential haem site in the inactive enzyme. The structure suggests that in the half-reduced functional form of the enzyme, the low-potential haem has to shed His71 in order to create a functional peroxidatic site and make the enzyme catalytically competent. This process is likely to trigger a reorganization of the active site, and may introduce new residues into the haem pocket.

Both domains of PsCCP are folded in a manner typical of class 1 cytochromes *c*, possibly indicating that the molecule arose through gene duplication. However, no similarity exists between the amino acid sequences of the two domains, supporting the generally accepted notion that the three-dimensional structures of proteins evolve more slowly than their amino acid sequences.

Although the enzyme shows some similarity to the membrane-bound di-haem cytochrome c_4 of *P. aeruginosa* and to the cytochrome subunit of flavocytochrome *c* sulfide dehydrogenase, the orientation of the haem groups and the likely electron transfer paths for these proteins are very different.

Materials and methods

PsCCP was crystallized as described by Fülöp *et al.* [13]. Crystals belong to the trigonal space group $P3_121$ with cell dimensions a=b=113.9 Å, c=72.0 Å, and there is one molecule in the asymmetric unit. The solvent content of the crystals is 61% by volume. Crystals diffract to about 2.4 Å Bragg spacings at room temperature using synchrotron radiation.

Data collection and preparation of heavy-metal derivatives

Native data to 2.4 Å resolution were collected using 1.00 Å radiation at station 9.5 (18 cm diameter MAR Research image plate detector) at the SERC Synchrotron Radiation Source, Daresbury, UK. Data were processed using DENZO and its companion program SCALEPACK [36]. Heavy-metal derivatives were prepared by soaking crystals in solutions of heavy-metal compounds in a synthetic mother liquor overnight. Soaked crystals very often cracked or suffered from non-isomorphism with a typical 2–3 Å shrinkage in the crystallographic c direction. None of these crystals diffracted X-rays beyond 4 Å resolution. All potential derivative data were collected on a Siemens (Xentronix) area detector using CuKα radiation supplied by a Rigaku rotating anode generator operating at 60 kV, 60 mA. Derivative data were integrated with XDS [37], and further processing was done with the CCP4 package [38].

Phase determination

Three out of about 40 different soaks produced useful derivatives. The major heavy-metal sites for each derivative were located by analysis of isomorphous difference Patterson maps; subsites were found by difference Fourier syntheses. Heavy-atom parameters were refined using the program MLPHARE [39]. The mean figure of merit for acentric/centric phases to 4 Å was 0.54/0.72. Table 1 summarizes the data collection and phasing statistics. The MIR phases were improved by solvent-flattening and histogram-matching techniques using the CCP4 suite of programs (program DM [40]), and extended to a resolution of 3.5 Å in an iterative process.

Model building and crystallographic refinement

Skeletonized maps [41] were calculated for chain tracing. Although the phase-improved map at 3.5 Å resolution showed good connectivity, it lacked details in many places. Starting from the two haem-binding Cys-X-Y-Cys-His motifs, it was possible to fit about 80% of all protein atoms to the density using the program O [42]. The partial model was subjected to refinement with X-PLOR [43]. The crystallographic R-factor for the initial model was 44.5% and was reduced to 27.5% (free R-factor 33.9%, [44]) in one round of energy minimization and simulated annealing. Further model building and refinement were pursued with alternate cycles of manual refitting using O and simulated annealing with X-PLOR. When the free R-factor did not decrease any further, water molecules were added to the atomic model at the positions of large positive peaks (>3.5σ) in the difference electron density, only at places where the resulting water molecule fell into an appropriate hydrogen-bonding environment. At the final stages, restrained isotropic temperature factor refinements were carried out for each individual atom. Introduction of water molecules in each of the refinement steps resulted in a decrease in both conventional and free R-values. The final model contains 317 residues (residues 223–228 are missing in the present structure), two *c* haems, one calcium ion, and 73 solvent molecules. The crystallographic R-factor is 19.2% in the resolution range 8.0–2.4 Å (defined as $R_{cryst}=\Sigma||F_{obs}|-|F_{calc}||/\Sigma|F_{obs}|\times100$), using 18821 reflections with $F_{obs}>2\sigma(F_{obs})$. The free R-value [44] is 25.0% based on 748 randomly selected reflections (4% of the total) with $F_{obs}>2\sigma(F_{obs})$. The R_{cryst} and free R-values for all observations within this resolution range (19223 and 764 unique reflections) are 19.6% and 25.3%, respectively. The rms deviation in bond lengths, bond angles and torsion angles from standard values is 0.01 Å, 2.0° and 24.0°, respectively. The mean coordinate error is 0.39 Å based on the SIGMAA method [45]. There is only one outlier in the Ramachandran plot, Phe260 (ϕ=71°,ψ=159°), which is located in the hydrophobic environment of the high-potential haem centre. Coordinates have been submitted to the Brookhaven Protein Data Bank.

Preparation of figures

Figure 2 was produced with ALSCRIPT [46]. Figures 3 and 5 were drawn by MOLSCRIPT [47] modified by R Esnouf and rendered with Raster3D [48]. Figures 4, 6 and 7 were prepared using XOBJECTS (MEM Noble, unpublished program).

Acknowledgements: We thank A Thompson and M Nobbs for help in protein purification. We are grateful to PJ Artymiuk for the ASSAM search of the Brookhaven Protein Data Bank, to MEM Noble for help in figure preparation and to the staff of SRS Daresbury Laboratory. This work was supported by the MRC and the BBSRC.

References

1. Ellfolk, N. & Soininen, R. (1971). *Ps.* cytochrome *c* peroxidase III. The size and shape of the enzyme molecule. *Acta Chem. Scand.* **25**, 1535–1540.
2. Ellfolk, N., Rönnberg, M., Aasa, R., Andreasson, L.-E. & Vanngard, T. (1983). Properties and function of the two hemes in *Pseudomonas* cytochrome *c* peroxidase. *Biochim. Biophys. Acta* **743**, 23–30.
3. Foote, N., Peterson, J., Gadsby, P.M.A., Greenwood, C. & Thomson, A.J. (1985). Redox linked spin state changes in the di-haem cytochrome c_{551} peroxidase from *Ps. aeruginosa*. *Biochem. J.* **230**, 227–237.
4. Greenwood, C., Foote, N., Gadsby, P.M.A. & Thomson, A.J. (1988). A di-haem cytochrome *c* peroxidase (*Pseudomonas aeruginosa*): its activation and catalytic cycle. *Chemica Scripta* **28A**, 79–84.
5. Foote, N., Turner, R., Brittain, T. & Greenwood, C. (1992). A quantitative model for the mechanism of action of cytochrome *c* peroxidase of *Ps. aeruginosa*. *Biochem. J.* **283**, 839–843.
6. Dolphin, D., Forman, A., Borg, D.C., Fajer, J. & Felton, R.A. (1971). Compounds I of catalase and horseradish peroxidase: π cation radicals. *Proc. Nat. Acad. Sci. USA* **68**, 614–618.
7. Edwards, S.L., Xuong, N., Hamlin, R.C. & Kraut, J. (1987). Crystal structure of cytochrome *c* peroxidase compound I. *Biochemistry* **26**, 1503–1511.
8. Mauro, J.M., *et al.*, & Kraut, J. (1988). Tryptophan-191→phenylalanine, a proximal-side mutation in yeast cytochrome *c* peroxidase that strongly affects the kinetics of ferrocytochrome *c* oxidation. *Biochemistry* **27**, 6243–6256.
9. Sivaraja, M., Goodin, D.B., Smith M. & Hoffman B.M. (1989). Identification by ENDOR of Trp191 as the free-radical site in cytochrome c peroxidase compound ES. *Science* **245**, 738–740.
10. Fülöp, V., *et al.*, & Edwards S.L. (1994). Laue diffraction study on the structure of cytochrome *c* peroxidase compound I. *Structure* **2**, 201–208.
11. Greenwood, C. & Gibson, Q.H. (1989). Ligand binding to cytochrome *c* peroxidase from *Ps. aeruginosa*. *J. Biol. Chem.* **264**, 19022–19027.
12. Rönnberg, M., Kalkkinen, N. & Ellfolk, N. (1989). The primary structure of *Pseudomonas* cytochrome *c* peroxidase. *FEBS Lett.* **250**, 175–178.
13. Fülöp, V., Little, R., Thompson, A., Greenwood, C. & Hajdu, J. (1993). Crystallization and preliminary X-ray analysis of the di-haem cytochrome *c* peroxidase from *Pseudomonas aeruginosa*. *J. Mol. Biol.* **232**, 1208–1210.

14. Ridout, C.J., James, R. & Greenwood, C. (1995). Nucleotide sequence encoding the di-haem cytochrome c_{551} peroxidase from *Pseudomonas aeruginosa*. *FEBS Lett.* **365**, 152–154.
15. Moore, G.R. & Pettigrew, G.W. (1990). *Cytochromes c: Evolutionary, Structural and Physicochemical Aspects*. Springer-Verlag, Berlin.
16. Fülöp, V., Moir, J.W.B., Ferguson, S.J. & Hajdu, J. (1995). The anatomy of a bifunctional enzyme: structural basis for reduction of oxygen to water and synthesis of nitric oxide by cytochrome cd_1. *Cell* **81**, 369–377.
17. Poulos, T.L. & Kraut, J. (1980). The stereochemistry of peroxidase catalysis. *J. Biol. Chem.* **255**, 8199–8205.
18. Finzel, B.C., Poulos, T.L. & Kraut, J. (1984). The crystal structure of yeast cytochrome *c* peroxidase refined at 1.7 Å resolution. *J. Biol. Chem.* **259**, 13027–13036.
19. Erman, J.E., Vitello, L.B., Miller, M.A., Shaw, A., Brown, K.A. & Kraut, J. (1993). Histidine 52 is a critical residue for rapid formation of cytochrome *c* peroxidase compound I. *Biochemistry* **32**, 9798–9806.
20. Vitello, L.B., Erman, J.E., Miller, M.A., Wang, J. & Kraut, J. (1993). The effect of arginine 48 replacement on the reaction of cytochrome *c* peroxidase and hydrogen peroxide. *Biochemistry* **32**, 9807–9818.
21. Ellfolk, N., Rönnberg, M. & Osterlund, K. (1991). Structural and functional features of *Pseudomonas* cytochrome *c* peroxidase. *Biochim. Biophys. Acta* **1080**, 68–77.
22. Artymiuk, P.J., Poirrette, A.R., Grindley, H.M., Rice D.W. & Willett, P. (1994). A graph-theoretic approach to the identification of three-dimensional patterns of amino acid side-chains in protein structures. *J. Mol. Biol.* **243**, 327–344.
23. Bernstein, F.C., *et al.*, & Tasumi, M. (1977). The protein data bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535–542.
24. Scott, D.L., White, S.P., Browning, J.L., Rosa, J.J., Gelb, M.H. & Sigler, P.B. (1991). Structures of free and inhibited human secretory phospholipase A_2 from inflammatory exudate. *Science* **254**, 1007–1010.
25. Haschke, R.H. & Friedhoff, J.M. (1978). Calcium related properties of horseradish peroxidase. *Biochem. Biophys. Res. Commun.* **80**, 1039–1042.
26. Gilmour, R., Goodhew, C.F., Pettigrew, G.W., Prazeres, S., Moura, J.J.G. & Moura, I. (1994). The kinetics of the oxidation of cytochrome *c* by *Paracoccus* cytochrome *c* peroxidase. *Biochem. J.* **300**, 907–914.
27. Poulos, T.L., Edwards, S.L., Wariishi, H. & Gold, M.H. (1993). Crystallographic refinement of lignin peroxidase at 2 Å. *J. Biol. Chem.* **268**, 4429–4440.
28. Moser, C.C., Keske, J.M., Warncke, K., Farid, R.S. & Dutton, P.L. (1992). Nature of biological electron transfer. *Nature* **355**, 796–802.
29. Marcus, R.A. & Sutin, N. (1985). Electron transfers in chemistry and biology. *Biochim. Biophys. Acta* **811**, 265–322.
30. Canters, G.W. & van de Kamp, M. (1992). Protein-mediated electron transfer. *Curr. Opin. Struct. Biol.* **2**, 859–869.
31. Beratan, D.N., Onuchic, J.N., Winkler, J.R. & Gray H.B. (1992). Electron-tunneling pathways in proteins. *Science* **258**, 1740–1741.
32. Mayo, S.L., Ellis, W.R., Crutchley, R.J. & Gray, H.B. (1986). Long-range electron transfer in heme proteins. *Science* **233**, 948–952.
33. Baldwin, J.E., Morris, G.M. & Richards, W.G. (1991). Electron transport in cytochromes P-450 by covalent switching. *Proc. R. Soc. Lond. B* **245**, 43–51.
34. Pelletier, H. & Kraut, J. (1992). Crystal structure of a complex between electron transfer partners, cytochrome *c* peroxidase and cytochrome *c*. *Science* **256**, 1748–1755.
35. Chen, Z., *et al.*, & Mathews, F.S. (1994). The structure of flavocytochrome *c* sulfide dehydrogenase from a purple phototropic bacterium. *Science* **266**, 430–432.
36. Otwinowski, Z. (1993). Oscillation data reduction program. In *Data Collection and Processing*. (Sawyer, L., Isaacs, N.W. & Bailey S., eds), pp. 55–62, SERC Daresbury Laboratory, Warrington, UK.
37. Kabsch, W. (1988). Evaluation of single crystal X-ray diffraction data from a position sensitive detector. *J. Appl. Cryst.* **21**, 916–924.
38. Collaborative Computational Project, Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst. D* **50**, 760–763.
39. Otwinowski, Z. (1991). Maximum likelihood refinement of heavy atom parameters. In *Isomorphous Replacement and Anomalous Scattering*. (Wolf, W., Evans, P.R. & Leslie, A.G.W., eds), pp. 80–86, SERC Daresbury Laboratory, Warrington, UK.
40. Cowtan, K. (1994). 'DM': an automated procedure for phase improvement by density modification. In *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography*. **31**, 34–38.
41. Jones, T.A. & Thirup, S. (1986). Using known substructures in protein model-building and crystallography. *EMBO J.* **5**, 819–822.
42. 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.
43. Brünger, A.T. (1992). *X-PLOR: Version 3.1. A System for Protein Crystallography and NMR*. Yale University Press, New Haven, CT.
44. Brünger, A.T. (1992). The free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472–474.
45. Read, R.J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Cryst. A* **42**, 140–149.
46. Barton, G.J. (1993) ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng.* **6**, 37–46.
47. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946–950.
48. Merritt, E.A. & Murphy, M.E.P. (1994). Raster3D Version 2.0. A program for photorealistic molecular graphics. *Acta Cryst. D* **50**, 869–873.

Received: **17 Jul 1995**; revisions requested: **8 Aug 1995**; revisions received: **4 Sep 1995**. Accepted: **14 Sep 1995**.