Structural Characterization of *Paracoccus denitrificans* Cytochrome *c* Peroxidase and Assignment of the Low and High Potential Heme Sites†

Wei Hu,‡ Gonzalez Van Driessche,‡ Bart Devreese,‡ Celia F. Goodhew,§ Dermot F. McGinnity,§ Neil Saunders,§ Vilmos Fulop,|| Graham W. Pettigrew,§ and Jozef J. Van Beeumen*,‡

Department of Biochemistry, Physiology, and Microbiology, Laboratory of Protein Biochemistry and Protein Engineering, State University of Gent, B-9000 Gent, Belgium, Department of Molecular Biophysics, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K., and Department of Preclinical Veterinary Sciences, Royal (Dick) School of Veterinary Sciences, University of Edinburgh, Summerhall, Edinburgh EH9 1QH, U.K.

Received December 20, 1996; Revised Manuscript Received April 14, 1997⊗

ABSTRACT: The amino acid sequence of the diheme cytochrome *c* peroxidase from *Paracoccus denitrificans* has been determined as the result of sequence analysis of peptides generated by chemical and enzymatic cleavages of the apoprotein. The sequence shows 60% similarity to the cytochrome *c* peroxidase from *Pseudomonas aeruginosa*, 39% similarity to an open reading frame encoding a putative triheme *c*-type cytochrome in *Escherichia coli*, and remote similarity to the MauG proteins from two methylotrophic bacteria. It is proposed, on the basis of the pattern of conserved residues in the sequences, that a change in iron coordination in the N-terminal heme domain may accompany reduction to the active mixed valence state, a change which may be accompanied by conformational adjustments in the highly conserved interface between the N- and C-terminal domains. These conformational adjustments may also lead to the appearance of a second Ca^{2+} binding site in the mixed valence enzyme. The exposed edge of the heme in the C-terminal domain is surrounded by several different patterns of charged residues in the *Paracoccus* and *Pseudomonas* enzymes, and this is consistent with the interaction of the former with the highly positively charged front face of the donor cytochrome *c*-550.

In most living cells, hydrogen peroxide is toxic and therefore neutralized by the enzymes catalase or peroxidase. Catalase disproportionates peroxide into water and oxygen whereas peroxidase reduces peroxide to water using a variety of oxidizable substrates which, in the case of the cytochrome *c* peroxidases (CCP),[1] are the small monoheme *c*-type cytochromes. The bacterial members of the peroxidases can also oxidize the copper protein azurin (Soininen et al., 1972). Bacterial CCP's (which contain two covalently bound *c*-type hemes) so far have been isolated from *Pseudomonas aeruginosa* (Ronnberg & Ellfolk, 1979; Ellfolk et al., 1983; Foote et al., 1984), *Paracoccus denitrificans* (Goodhew et al., 1990; Pettigrew, 1991; Gilmour et al., 1993), *Rhodobacter capsulatus* (Hanlan et al., 1992), and *Nitrosomonas europeae* (Arciero & Hooper, 1994). The *Pseudomonas* CCP has been well characterized by different techniques, and its sequence has been redetermined recently (Ellfolk et al., 1991; Ridout et al., 1995; Samyn et al., 1995). The three-dimensional structure of the *Pseudomonas* CCP consists of two domains, each containing one heme (Fulop et al., 1995b). *Paracoccus* CCP is similar to *Pseudomonas* CCP in size, visible absorption spectrum, and cellular location. *Paracoccus* CCP likewise contains one high and one low potential heme, the former heme acting as a source of the second electron for the reduction of peroxide whereas the latter acts as a peroxidatic center (Gilmour et al., 1993). Reduction of the *Paracoccus* high potential heme results in a switch of the low potential heme to a high spin state, as shown by visible and NMR spectroscopy. The spin state switch of the low potential heme is accompanied by an increase of the redox potential of the high potential heme by 50 mV, an indication of heme−heme interaction. A bound Ca^{2+} is essential for the high spin state switch of the *Paracoccus* low potential heme (Gilmour et al., 1993, 1994, 1995). The *Paracoccus* enzyme is significantly different from that of *Pseudomonas* in the time course of high spin formation after reduction of the high potential heme, and in the requirement for bivalent cations.

In this paper, we present the covalent structure of the cytochrome *c* peroxidase from *Paracoccus denitrificans* (Pad CCP) and relate this to the known three-dimensional structure of the fully oxidized inactive form of *Ps. aeruginosa* CCP. Consequently, we propose an explanation for the Ca^{2+}-dependent activity of Pad CCP and for the change in heme properties associated with formation of the mixed valence form.

EXPERIMENTAL PROCEDURES

Starting Preparation of Cytochrome and Deheming. The cytochrome *c* peroxidase was prepared according to the procedure described by Pettigrew (1991). The hemes of the cytochromes were removed by dissolving the lyophilized cytochrome *c* preparation in an acidic solution of 8 M urea

† G.W.P. thanks the Wellcome Trust for financial support. D.F.M. thanks the Veterinary Faculty for a research studentship. J.J.V.B. is indebted to the Belgian Fund for Joint Basic Research (32001891) and the Flemish Government for a Concerted Research Action (12052293).

* Address correspondence to this author. Phone: 32 9 246 5109. Fax: 32 9 264 5338. E-mail: jozef.vanbeeumen@rug.ac.be.

‡ State University of Gent.

§ University of Edinburgh.

|| University of Oxford.

⊗ Abstract published in *Advance ACS Abstracts,* June 1, 1997.

[1] Abbreviations: CCP, cytochrome *c* peroxidase; Pad, *Paracoccus denitrificans*; Psa, *Pseudomonas aeruginosa*.

S0006-2960(96)03131-5 CCC: $14.00

containing $HgCl_2$. After incubation overnight at 37 °C, the hemes and salts were removed by gel filtration on a column of Sephadex G-25, fine (1.5 × 24 cm), equilibrated, and eluted with 5% formic acid.

Gel Electrophoresis. SDS−polyacrylamide gel electrophoresis was carried out on a 12% gel, as described by Laemmli (1970), in a SE 250 small electrophoresis system (Hoefer Scientific Instruments, San Francisco, CA). After electrophoresis, the proteins were transferred in a Transblot cell (Bio-Rad, Hercules, CA) onto two layers of PVDF membranes (Millipore, Bedford, MA) and stained with 0.015% Amido Black in 1% acetic acid. Destaining occurred with 2 volumes of water. The membranes were allowed to dry into air.

In Situ Enzymatic Digestion. In situ enzymatic digestion with trypsin was performed on the main bands (37 kDa) from four lanes as described by Aebersold et al. (1987). Bands from the stained PVDF membranes were cut into small pieces of about 2 × 2 mm and treated with 200 μL of PVP-40 in methanol and incubated at 37 °C for 30 min. Then, 200 μL of water was added, mixed with the solution, and discarded after about 5 min. The wet membranes were immediately immersed with 50 μL of 100 mM Tris-HCl buffer, pH 8.3, containing 1 μg of trypsin. After incubation for 5 h at 37 °C, the membranes were washed with 100 μL of 80% formic acid and four times with water. All the solutions were combined and dried in a Speedvac concentrator (Savant, Hicksville, NY).

Enzymatic Digestions and Chemical Cleavage. Twenty nanomoles of apoprotein was digested with *Staphylococcus aureus* V8 protease (Glu-C) in 50 mM phosphate buffer, pH 7.0, at 37 °C for 5 h using an E/S ratio of 1/35 (w/w). An enzymatic digestion with Lys-C endoproteinase from *Lysobacter enzymogenes* (Boehringer Mannheim, FRG) was carried out on 30 nmol of apoprotein in 0.1 M ammonium hydrogen carbonate, pH 8.0, for 4 h at 37 °C at an E/S ratio of 1/38.

Partial acid hydrolysis of 20 nmol of apoprotein was performed with 2% formic acid at 106 °C for 2 h.

Enzymatic Subdigestions. A subdigestion of 6 nmol of peptide K15 with 1 μg of trypsin (Boehringer Mannheim, FRG) was carried out in 50 mM ammonium hydrogen carbonate, pH 8.3, for 3 h at 37 °C. Two nanomoles of K17 was digested with Asp-N endoproteinase in 50 mM sodium phosphate, pH 7.0, at 37 °C for 22 h using an E/S ratio of 1/50 (w/w). Further subdigestion of peptide K17 D2 was performed with 0.5 μg of trypsin for 2 h at 37 °C in 50 mM ammonium hydrogen carbonate, pH 8.3.

Peptide Purification. The peptides obtained after *in situ* digestion or digestion on the apoprotein were separated by reversed-phase high-performance liquid chromatography on a Pep S C_2/C_{18} column (4.6 × 250 mm, Pharmacia, Uppsala, Sweden) or a TMS column (10 × 100 mm, Bio-Rad, Hercules, CA), respectively. The elution equipment consisted of an 870 three-headed piston pump, an 8800 system controller, a variable wavelength detector set at 220 nm (all parts from DuPont, Wilmington, DE), and a 7120 rheodyne injector equipped with a loop of 100 μL. The solvents used were 0.1% TFA in Milli-Q water (solvent A) and 0.1% TFA in 70% acetonitrile or 100% acetonitrile (solvent B) at a flow rate of 1 mL/min.

Peptides generated by subdigestions were purified on a PTC C_{18} column (2.1 × 220 mm, Brownlee, Perkin Elmer, Applied Biosystems Division, Foster City, CA) with a 140A solvent delivery system (Perkin Elmer, Applied Biosystems Division, Foster City, CA) equipped with a 1000S photodiode array detector (Perkin Elmer, Applied Biosystems Division, Foster City, CA). The solvents were 0.1% TFA in Milli-Q water and 0.1% TFA in 70% acetonitrile. They were pumped at a combined flow rate of 100 μL/min.

Amino-Terminal Sequence and Amino Acid Composition Analysis. Automated N-terminal sequence analysis of peptides was performed with an Applied Biosystems pulsed liquid 475A or 477A sequenator equipped with an on-line 120A PTH-amino acid analyzer (Perkin Elmer, Applied Biosystems Division, Foster City, CA).

Amino acid composition analysis was performed on a 420A derivatizer with an on-line 130A separation system (Perkin Elmer, Applied Biosystems Division, Foster City, CA). Gas-phase hydrolysis at 106 °C for 24 h was carried out in a borosilicate glass tube of 5 × 55 mm placed in a Pierce hydrolysis vial, using 6 N HCl as hydrolysis agent.

Mass Analysis. Plasma desorption mass spectra were collected on a BioIon 20K biopolymer mass analyzer (BioIon, Uppsala, Sweden) during 10^6 start counts of the ^{252}Cf source at an accelerating potential of 15 kV. Samples were applied in small volumes of 0.1% TFA in Milli-Q water and spin-dried on nitrocellulose-coated targets wetted with 5 μL of 0.1% TFA in methanol.

Electrospray mass spectrometry was performed on a Bio-Q quadrupole mass spectrometer equipped with an electrospray ionization source (Micromass, Altrincham, U.K.). Ten microliters of sample solution in 50% acetonitrile/1% acetic acid was injected manually in the 10 μL loop of the Rheodyne injector and pumped to the source at a flow rate of 5 μL/min. The solvent, 50% acetonitrile/1% acetic acid, was delivered by a 140A solvent delivery system. Scans of 12 s over the mass range of 400−1600 amu were collected over 2 min. Calibration of the scans was performed with 50 pmol of horse heart myoglobin.

RESULTS

The complete primary structure of Pad CCP has been determined as the result of sequence analysis of peptides generated by various enzymes and one chemical cleavage applied to the apoprotein or to the apoprotein in which the cysteines were derivatized by pyridylethylation. The result is given in Figure 1, although only the most important peptides necessary to deduce the sequence are shown.

An initial attempt to obtain N-terminal sequence information on electroblotted protein failed, suggesting that the N-terminal residue of the polypeptide chain was blocked for Edman degradation. The first peptides obtained were those from a tryptic digest on the electroblotted protein. The separation of the peptides is shown in Figure 2. The sequence information from all these peptides covered about half of the total polypeptide chain, mostly from the N-terminal region of the protein. Due to partial cleavage at a Lys−Gln bond, it was ascertained that fractions T12 (16−29) and T10 (30−49) were adjacent peptides.

Most of the other overlaps came from a digest of the apoprotein with Glu-C endoproteinase (Figure SI; see Supporting Information). The result was a set of large fragments of 43 residues (13−55), 68 residues (118−185), 27 residues (247−268 and 270−274), and 30 residues (305−325, 327, 329−336), respectively. The overlaps between these fragments and several not yet positioned tryptic and Glu-C peptides were established by sequence analysis of peptides

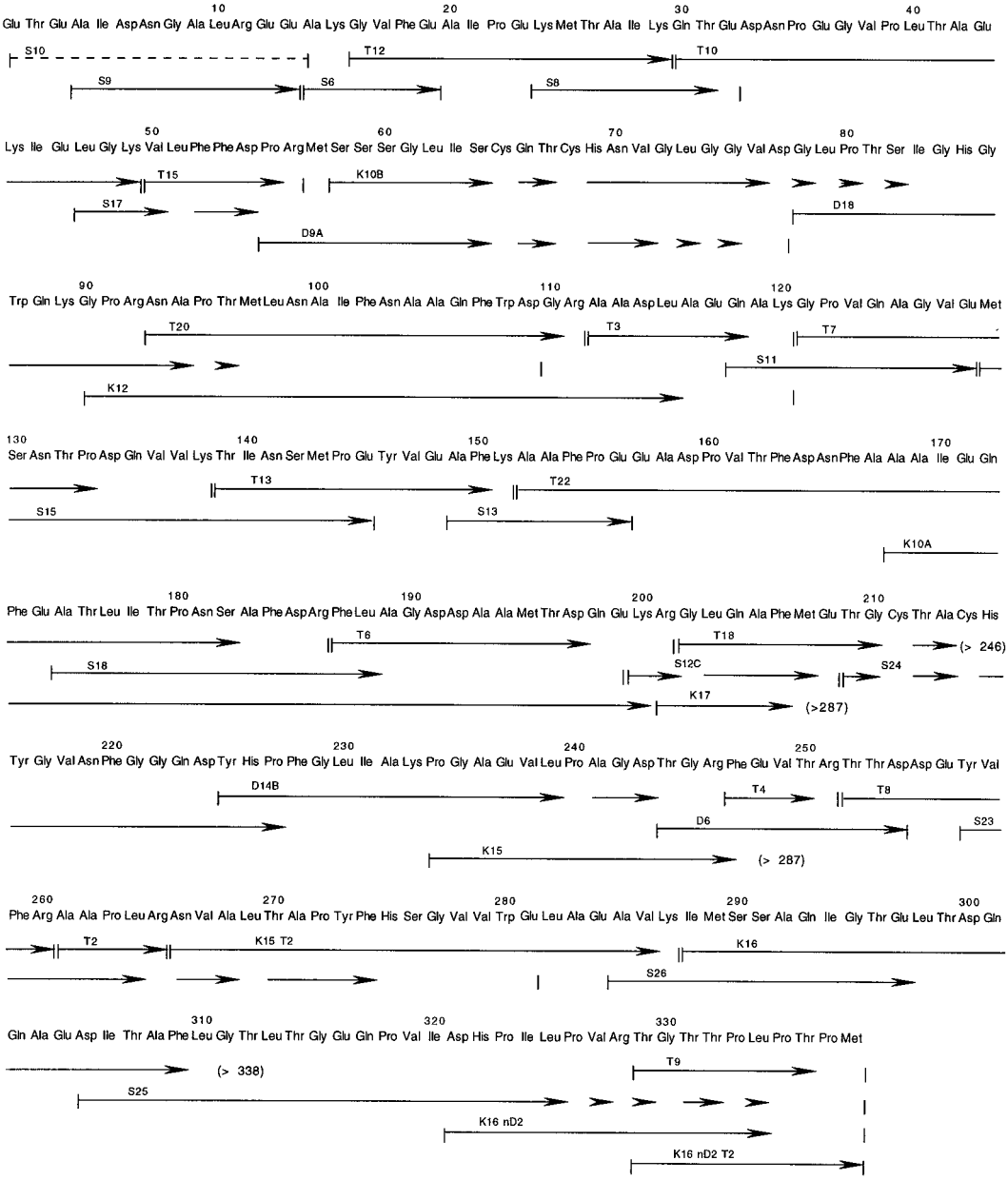

FIGURE 1: Amino acid sequence of cytochrome *c* peroxidase from *Pa. denitrificans.* Evidence for this sequence is obtained after Edman degradation of peptides generated by three enzymatic digestions with trypsin (T), Glu-C (S), and Lys-C (K) endoproteinase and one chemical cleavage with diluted formic acid (D). Amino acids which are identified unambiguously during Edman degradation are indicated with arrows. The start and end positions of peptides confirmed by mass spectrometric analysis are indicated by vertical lines or by "*>XXX*" of which *XXX* is the position where the peptide ends. The composition of the blocked peptide S10 was determined by amino acid analysis and therefore indicated with a dashed line.

from a Lys-C endoproteinase digest (Figure 3) and a partial acid hydrolysis with dilute formic acid (Figure SII). A subdigest of peptide K15 (Figure SIII) with trypsin was successful in establishing the overlap of S23 (257−291) with S26 (285−297)/K16 (288−338), resulting in structural information of the nearly complete C-terminal region of Pad CCP.

The exact sequence of the C-terminal region was deduced from a subdigest of Lys-C peptide K16 with N-Asp endoproteinase (Figure SIV). One of the two resulting fractions, K16nD2 (321−338), appeared to be the C-terminal peptide, but the sequence analysis failed to prove the last four residues. A further subdigest with trypsin also resulted in two peptides (Figure SV) of which K16nD2T2 (329−338) unambiguously proved the C-terminal sequence to be Leu-Pro-Thr-Pro-Met. This result was in agreement with the mass of 1016.0 Da measured by plasma desorption mass spectrometry (calculated mass: 1015.2 Da). The quantitative data of the sequence analysis of peptides from different enzymatic digestions and one chemical cleavage are given in Tables SI−SV (see Supporting Information).

The exact sequence of the first five residues of the polypeptide chain was deduced indirectly from the amino acid composition analysis of peptide S10 (1−12/13) (Table SVI). The peptide itself was resistant to Edman degradation. PDMS mass analysis of S9 (4−12/13) showed the presence of two peptides differing each by 129.4 Da (see Table SII), the mass of a glutamic acid residue. This confirmed that the C-terminal end of S9 was Arg-Glu-Glu and not Arg-Glu. In combination with the compositional analysis of S10 (which showed a similar C-terminal heterogeneity), and given the fact that an N-terminal glutamine residue can undergo cyclization during protein and/or peptide purification, it became clear that the N-terminal sequence of Pad CCP is

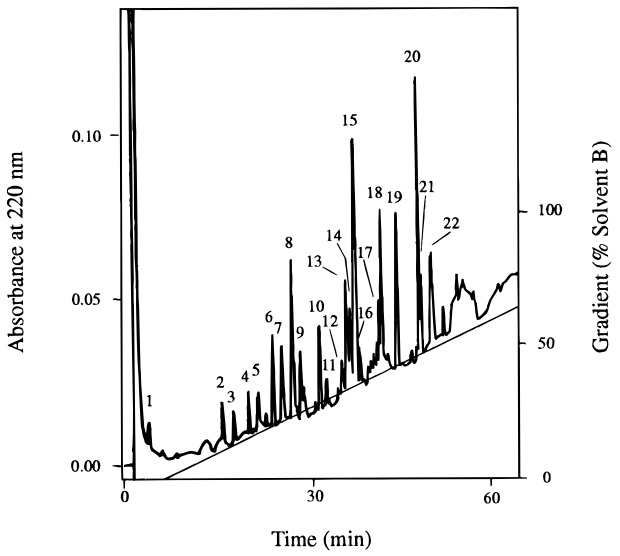

FIGURE 2: Reversed-phase high-performance liquid chromatography of peptides obtained after *in situ* tryptic digestion of electroblotted apoprotein. Conditions of the separation are described under Experimental Procedures.

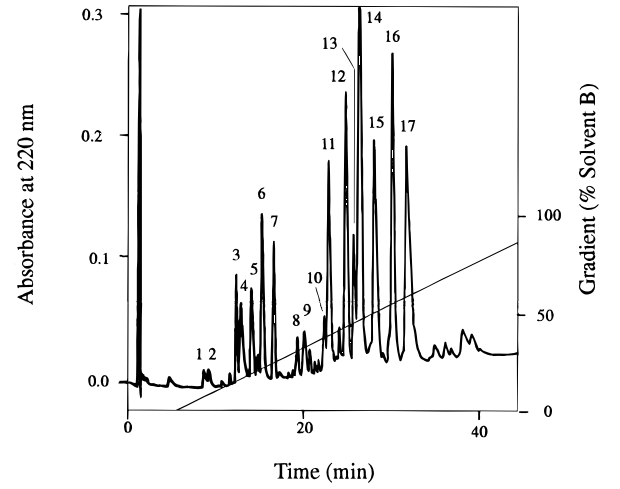

FIGURE 3: HPLC separation of peptides generated by digestion of apoprotein with Lys-C endoproteinase. Conditions of the separation are described under Experimental Procedures.

Gln-(Thr-Glu)-Ala and that the resistance to Edman degradation was due to the formation of pyroglutamic acid. This conclusion is in agreement with the fact that the mass difference between S10 and S9 (both forms) is 341.3 Da, which corresponds to the calculated mass of pGlu-(Thr-Glu) (341.4 Da). This sequence was confirmed by removal of the cyclized glutamine by pyrrolidonecarboxylic acid removing enzyme. Edman degradation of the remainder of the protein gave the sequence Thr-Glu-Ala-Ile-Asp-Asn.

Overall, it can be said that the sequence as given in Figure 1 is solidly proven for all peptide bonds, except for regions 75−85 and 283−289. However, given the fact that, in all those cases where detailed mass analysis of peptides was performed, the data corroborated the proposed sequence (Tables SI−SV), we are convinced that this sequence proposal is error free. The amino acid composition of the holoprotein expressed as moles of amino acid per 2 moles of heme shows a very good agreement with the proposed sequence composition (Table SVII). The mass of the holoprotein measured by ESMS was 37512.27 Da, in very good agreement with the calculated sequence mass of 37512.79 Da, including two hemes (Figure 4A). The apoprotein gave a mass of 36679.49 Da (Figure 4B). This value is compatible with a calculated mass of 36680.91 Da for the apoprotein of which each heme binding sequence (Cys-X-Y-Cys-His) will carry a Hg atom bound after the heme removal process. Such a mercury adduct has already been described for several apocytochromes obtained after heme removal of the native protein with acidic $HgCl_2$ (Samyn et al., 1994). Pad CCP does not contain cysteines other than those of the heme binding sites. Further confirmation that the whole protein sequence has been determined comes from digestion of the holoprotein with subtilisin under native conditions (McGinnity et al., 1996). A single peptide bond is cleaved and the two peptides have masses of 27950.9 and 9581.38 Da. These masses closely correspond to those calculated for pGlu1−Thr250 (27950.93 Da) and Arg251−Met338 (9579.87 Da).

DISCUSSION

Sequence Alignment and Sequence Similarity. The amino acid sequences of the cytochrome *c* peroxidases from *Pa. denitrificans* (Pad CCP) and *Ps. aeruginosa* (Psa CCP) (Ridout et al., 1995; Samyn et al., 1995) are compared in Figure 5 along with the open reading frame (f465) reported in the *E. coli* genome (Sofia et al., 1994). The sequence similarity is shown in the matrix of Figure 6. The 60% similarity between the Pad and Psa CCP sequences is high in comparison with the low levels of similarity seen in the small cytochromes *c* and in S1 oligonucleotide rRNA maps (Moore & Pettigrew, 1990a). The *E. coli* open reading frame f465 also shows a high degree of homology (approximately 40%) although it has an N-terminal extension that has no counterpart in the other sequences. This extension contains a putative methionine ligand, positioned 61 residues to the C-terminal side of a third heme site (boxed in Figure 5). It may be that this is the result of a gene fusion involving a small class I cytochrome donor, a possibility which has precedents in the case of the caa3 type oxidases (Sone & Yangita, 1982). No protein product of this open reading frame has yet been identified in *E. coli*, and indeed, *c*-type cytochromes only appear in this bacterium under quite specialized growth conditions (Iobbi-Nivol et al., 1994).

The partial sequence of a cytochrome *c* peroxidase from *N. europaea* has been determined (Arciero & Hooper, 1994). Over the 47 residues covered by the two heme peptides, there is respectively 49% and 51% similarity to the Pad CCP and Psa CCP sequences and 38% similarity to the *E. coli* orf. Although the sequence information for *N. europaea* CCP is limited, these data suggest that the divergence of the *Nitrosomonas* gene is more recent than that of the *E. coli* orf.

Although a case can be made for the inclusion of the *E. coli* orf and the *Nitrosomonas* CCP in a new class of bacterial cytochrome *c* peroxidases, the situation is less clear for the putative MauG proteins from *Methylobacter extorquens* and *Methylophilus methylotrophus* (Chistoserdov et al., 1994a,b). Their alignment requires greater use of gaps, and the percentage similarities are in the range 23−35% (Figure 6). Chistoserdov et al. (1994a,b) have suggested that the MauG proteins are indeed peroxidases but ones for which the oxidizable substrate is not cytochrome *c* but an intermediate in the synthesis of the prosthetic tryptophan tryptophylquinone (TTQ) group of methylamine dehydrogenase. It therefore seems best to view the MauG proteins as members of a

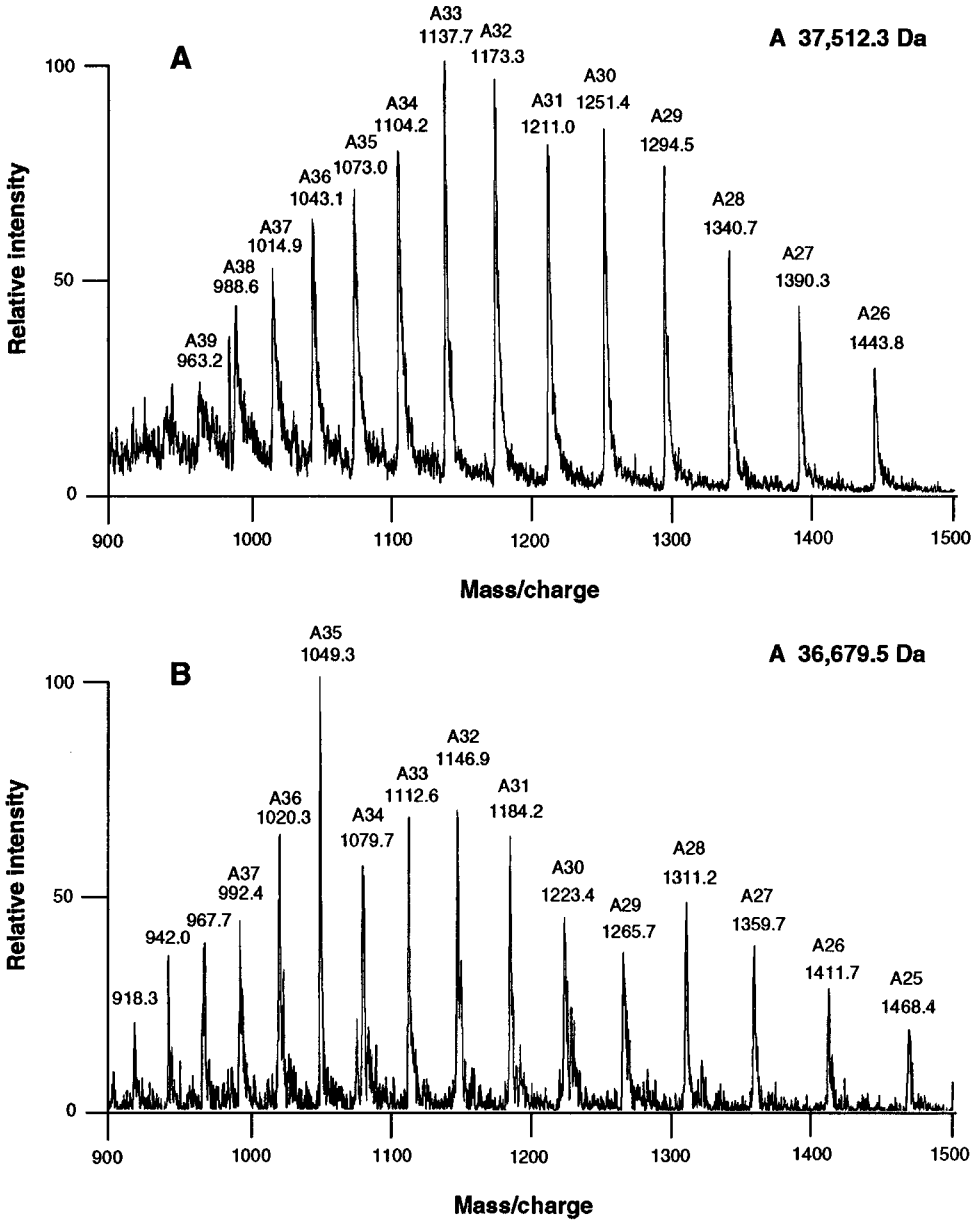

FIGURE 4: Electrospray ionization mass spectra of holoprotein (A) and apoprotein (B) of cytochrome *c* peroxidase from *Pa. denitrificans*. The number at the top of each peak represents the number of positive charges for the particular *m/z* peak.

related but distinct family. This is consistent with the recent report of a MauG gene sequence in *Pa. denitrificans* (van der Palen et al., 1995) which is even more remote from the cytochrome *c* peroxidase sequences than are its MauG relatives from *M. extorquens* and *M. methylotrophus*. This suggests that MauG and cytochrome *c* peroxidase are paralogous rather than orthologous gene products that have diverged to perform different but related functions.

Conservation of α-Helices and the Class I Fold. As emphasized by Fulop et al. (1995b), both the N- and C-terminal heme domains have the fold of the class I *c*-type cytochromes, although with some variations on the theme, in different regions of the molecules. The class I patterns of the two domains are shown in Figure 7c,d and compared with the patterns known for tuna cytochrome *c* and for *Pseudomonas* cytochrome *c*-551 (Figure 7a,b).

Both the N- and C-terminal domains of the cytochrome *c* peroxidases differ from the standard class I fold in that the right side loop traverses across the front of the heme before returning to the right side. The loop carries histidine 71 which acts as a sixth coordination ligand of the iron in place of the conventional methionine.

The two domains also differ in the pattern of chain following the left rear α-helix. In the standard class I fold, the chain forms a left-side loop and then passes upward to bring the methionine iron ligand into place followed by the C-terminal α-helix. Using the SOPM helix prediction suite assembled by Geourjon and Deleage (1994), it is clear that the α-helices are conserved features of the three sequences shown in Figure 5. Not only are most of the actual α-helices in the *Ps. aeruginosa* structure correctly predicted, but also the same pattern of helix prediction is essentially obtained for the Pad CCP and *E. coli* orf (broken bars in Figure 5). The only regions where the helix is predicted but not found in the Psa CCP structure are 84−104 and 242−246, which correspond to the highly conserved stretches 88−98 and 241−263. The significance of this will be discussed in the following section.

The Heme Ligands in Cytochrome c Peroxidase in the Oxidized and Mixed Valence States. In cytochrome *c* peroxidase from *Ps. aeruginosa*, Fulop et al. (1995b) have shown that the heme of the N-terminal domain is coordinated

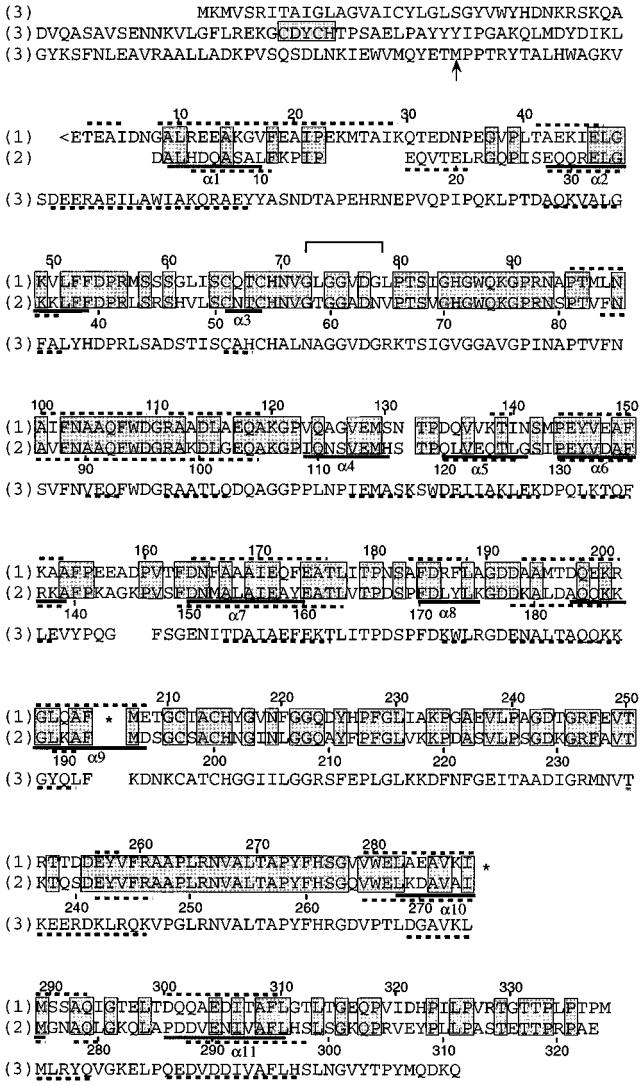

FIGURE 5: Comparison of the amino acid sequences of the bacterial cytochrome *c* peroxidases from (1) *Pa. denitrificans* as determined in this paper, (2) *Ps. aeruginosa*, and (3) open reading frame 465 in the *E. coli* genome. The *E. coli* open reading frame encodes an N-terminal extension that has no counterpart in the other two sequences. An additional possible heme binding site is boxed, and a possible methionine ligand is indicated by an arrow. The asterisks indicate regions where the MauG protein sequences (not shown) are longer than the sequences shown here. Identical residues in the amino acid sequences from *Pa. denitrificans* and *Ps. aeruginosa* are boxed and shaded. The actual α-helices present in the X-ray structure of the *Ps. aeruginosa* enzyme are indicated by a solid line below that sequence with the helix numbering (α1−α11) of Fulop et al. (1995b). α-Helices predicted by the SOPM method of Geourjon and Deleage (1994) for the *Pa. denitrificans*, *Ps. aeruginosa*, and *E. coli* sequences are shown as broken lines. Predictions of three or more residues are included. The bracketed region indicates a sequence which is related to sequences in the *β*-roll proteins involved in Ca^{2+} binding.

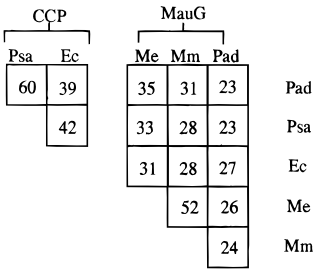

FIGURE 6: Matrix of percent identities of the amino acid sequences of the bacterial cytochrome *c* peroxidases and relatives. Psa and Pad CCP, cytochrome *c* peroxidases from *Ps. aeruginosa* and *Pa. denitrificans*, respectively; Ec, open reading frame 465 from the *E. coli* genome; Me, Mm, and Pad MauG, MauG proteins from *M. extorquens*, *M. methylotrophus*, and *Pa. denitrificans*, respectively. The number of identities in a pairwise comparison was obtained by summing (a) the number of identical residues and (b) the number of identical gap regions (gaps were scored as one point of comparison no matter the size).

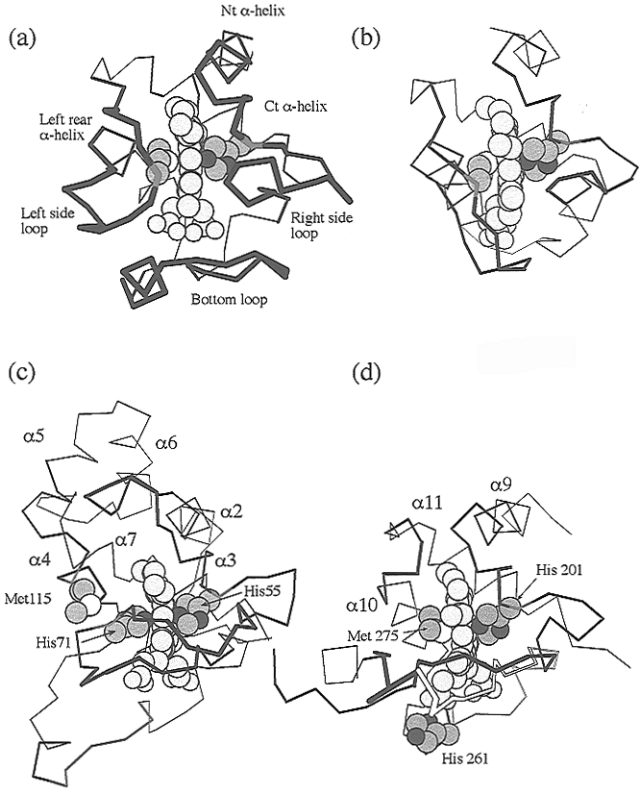

FIGURE 7: Class I folds of (a) tuna cytochrome *c*, (b) *Ps. aeruginosa* cytochrome *c*-551, (c) the N-terminal domain of Psa CCP, and (d) the C-terminal domain of Psa CCP. Coordinates for tuna ferrocytochrome *c* (Takano & Dickerson, 1981) and *Ps. aeruginosa* cytochrome *c*-551 (Matsuura et al., 1982) were obtained from the Brookhaven Protein Data Bank. Structures were viewed using MacImdad software (Molecular Application Group, Palo Alto, CA). In each case, the chain is traced by connecting α-carbon positions, and key residues are shown space-filled. The positions of the heme ligands on the α-chain are indicated in red. A low atomic radius (0.7 Å) was used to enhance clarity of the diagram. In the case of the Nt domain of CCP, a highly conserved segment of sequence (residues 88−98) is shown in green (see also Figure 8). Similarly, residues 241−263 in the C-terminal domain are shown in yellow (see also Figure 8). The heme groups are shown side-on at the front face of the structures with propionate groups pointing downward. At the far left of the Ct domain (d), the free chain ends are due to residues missing from the X-ray structure determination.

by the proximal histidine 55 (69 in the Pad CCP) and the distal histidine 71 (85 in Pad CCP). For the latter to occupy the distal side, the loop of the chain following the heme which usually forms the proximal side in class I cytochromes must traverse across the heme edge and then return (Figure 7). The heme of the C-terminal domain is coordinated by the proximal histidine 201 (215 in Pad CCP) and the distal methionine 275 (289 in Pad CCP). This is a more conventional coordination structure for a class I cytochrome *c*.

Three of these four amino acids (His55, His201, and Met275) are conserved not only in the two known cytochrome *c* peroxidases but also in the *E. coli* f465 open reading frame and in the MauG proteins. The fourth (His71) is not found in either the f465 or the MauG proteins, and

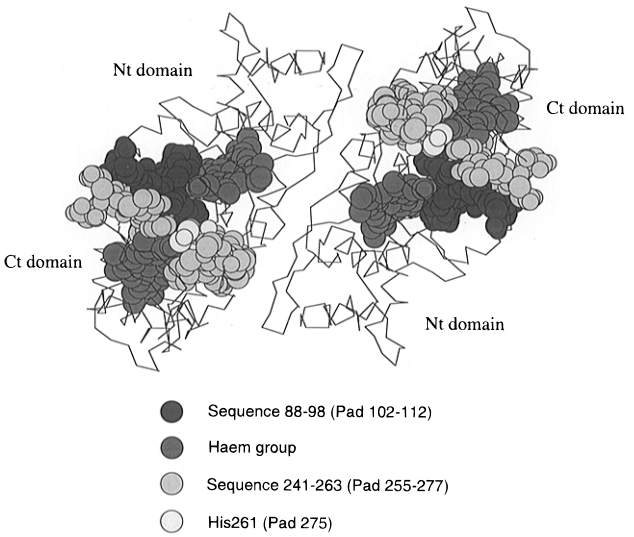

FIGURE 8: Highly conserved regions in the cytochrome *c* peroxidases. The view is of the crystallographic dimer of the *Ps. aeruginosa* cytochrome *c* peroxidase looking down the 2-fold axis. Connections are made between the α-carbon positions. The two heme groups of each monomer are indicated on the gray scale, and their orientation is such that the exposed edges of the N-terminal hemes point toward us and the exposed edges of the C-terminal hemes point to the right and left. In each monomer also sequences 241−263 (Pad 255−277) and 88−98 (Pad 102−112) are represented as space-filled regions. Histidine 261 (Pad 275), which is easily modified in the oxidized enzyme (see text), is shown as the least shaded residue and faces toward us.

there is no alternative conserved histidine in their N-terminal domains. In this connection it should be remembered that the oxidized form which has been crystallized represents a dead-end conformation into which the protein relaxes to protect the catalytic site in the absence of reductant (McGinnity et al., 1996). Thus the conformation of the oxidized form may not be so evolutionarily constrained as that of the functional mixed valence enzyme.

A residue that is conserved over the broader group of sequences is methionine 115 (129 in Pad CCP). It is located a very typical 60 residues C-terminal to the heme binding site that is characteristic of the class I cytochromes. Also, this residue is located on the left rear α-helix (Figure 7c) in a position equivalent to the coordinating methionine of the C-terminal domain and close to His71. We suggest that this residue is the distal ligand of the N-terminal heme domain in the *E. coli* and MauG proteins and must also be considered a candidate for the distal ligand in all the proteins in their mixed valence state. If that were the case, the roles of the peroxidatic center and electron-transferring center tentatively assigned to the N- and C-terminal hemes, respectively (Fulop et al., 1995b), may require reassessment once the structure of the mixed valence enzyme is known. The possibility of a switch in the heme properties associated with reduction has already been noted (Prazeres et al., 1995).

Highly Conserved Regions. In Figure 5, the most highly conserved regions of the sequences are the 11-residue segment Psa 88−98 (Pad 102−112) and the 23-residue segment Psa 241−263 (Pad 255−277). Even in the more remote *E. coli* sequence, most of these regions remain strongly conserved. These segments are located in the N- and C-terminal domains, respectively, in which they form the base of their respective heme crevices (Figure 7). Figure 8 shows the relationship between the two domains and their conserved regions as part of the whole molecule. Clearly, the conservatism is associated with the fact that it is at the base of their hemes that the two domains interact. Indeed, as Fulop et al. (1995b) have pointed out, a small stretch of β-strand in the N-terminal domain (Psa 95−97) is hydrogen bonded to a partner strand in the C-terminal domain (Psa 243−246).

In conventional monoheme class I cytochromes such as in mitochondrial cytochrome *c*, the polypeptide segments which form the base of the heme crevice and the left-rear α-helix are among the most variable in the molecule (Moore & Pettigrew, 1990). Thus, in the cytochrome *c* peroxidases, the constraints of domain−domain interactions may impose a structural conservatism on a region which is relatively unconstrained in the simple monoheme proteins.

We have mentioned above the possibility that some rearrangement of the protein occurs during the transition to the mixed valence state so that Met115 (129 in Pad CCP) becomes the coordinating ligand of the N-terminal domain. Of interest in this respect is that an α-helix is predicted in both the highly conserved regions (Psa 88−98, Psa 241−263) in all three sequences of Figure 5 and yet no α-helix is observed in the X-ray structure. Perhaps these conserved regions are constrained in the observed conformation in the oxidized state but switch to a new conformation including an α-helix in the mixed valence enzyme. A conformational rearrangement in this region upon reduction of the enzyme is consistent with the results of modification of histidine 261 (275 in Pad CCP). This residue, which appears shaded and exposed in Figure 8 is very easily modified at a low excess of the reagent diethyl pyrocarbonate in the oxidized enzyme but is much more difficult to modify in the mixed valence enzyme (McGinnity et al., 1996).

Ca^{2+} Binding. We have shown (Gilmour et al., 1993, 1994) that Ca^{2+} is required for the switch to the active high-spin state at the low potential heme of Pad CCP. Also, the enzyme loses activity on dilution and was therefore proposed to be in a dimer−monomer equilibrium with the dimer as the active form. A more detailed analysis of Ca^{2+} binding showed that at least two types of binding site were present (Gilmour et al., 1995). Site I is occupied in the oxidized enzyme while site II is of low affinity in the oxidized state but becomes high affinity in the mixed valence state in a transition associated with dimerization. Although no requirement for Ca^{2+} had previously been shown for the Psa CCP, Fulop et al. (1995b) found one bound Ca^{2+} per monomer in a buried coordination shell provided by four water molecules and oxygen atoms associated with residues N79, T256, and P258 (Figure 9). These also occur in Pad CCP (and indeed in the *E. coli* orf), and we therefore presume that this Ca^{2+} site is conserved. We propose that it represents site I.

One possible location in the sequence for site II is the segment G(72)LGGVDGL(79) (Pad CCP) which corresponds well to the consensus sequence GGXGXDXq (where q is a hydrophobic residue) found in the β-roll proteins such as alkaline protease (Baumann et al., 1993). In alkaline proteases, this sequence is situated at the end of a β-strand region and contributes one-half of the coordination for a Ca^{2+} ion, the other half being provided by a second sequence of this type on a neighboring turn to form a β-roll.

Since there are no further consensus patterns of this type detectable in the cytochrome *c* peroxidase sequences, it is possible that Ca^{2+} is held at the interface of a dimer in which each monomer contribues to the coordination. However, the

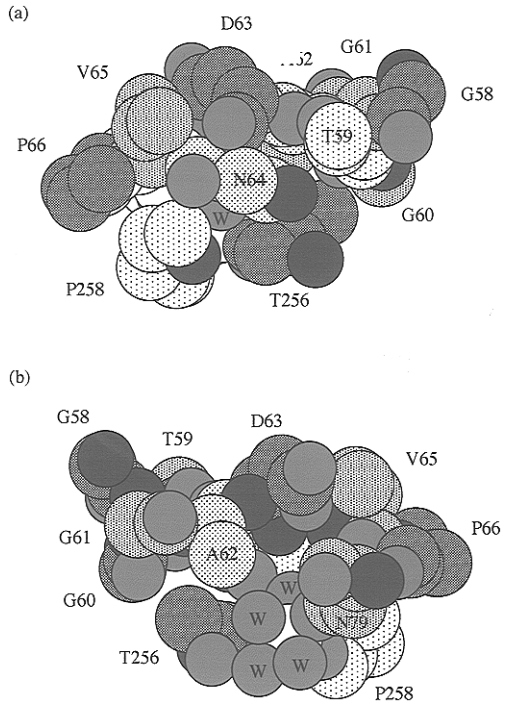

FIGURE 9: Calcium binding site of *Ps. aeruginosa* CCP. The figure shows the segment G58−P66, N79, T256, P258, and four bound water molecules (W) around the Ca^{2+} which is colored yellow. The carbonyl oxygen and water molecules are colored red, and amide nitrogen is in blue. View a is showing a region of the surface of one monomer within the dimer interface and is looking inward to the just visible calcium ion. View b is showing the same residues but in an orientation rotated 180° perpendicular to the plane of the paper. Here we are seeing the coordinating oxygens of the bound Ca^{2+}.

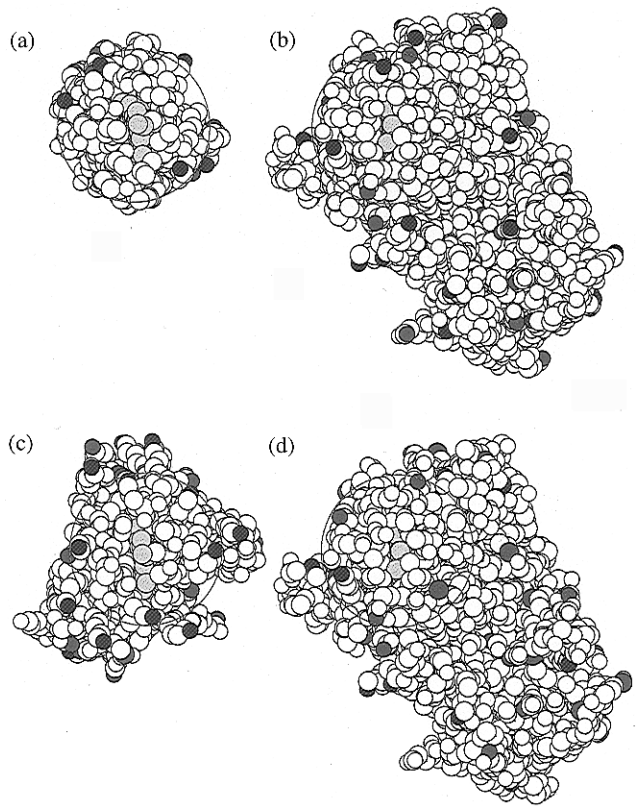

FIGURE 10: Charged surface of the cytochrome *c* peroxidases and their redox partners: (a) *Ps. aeruginosa* cytochrome *c*-551 [coordinates of Matsuura et al. (1982)], (b) *Ps. aeruginosa* CCP (coordinates of Fulop), (c) *Pa. denitrificans* cytochrome *c*-550 [coordinates of Benning et al. (1994)], and (d) *Pa. denitrificans* CCP (based on the sequence reported in this paper applied to the coordinates for the *Ps. aeruginosa* enzyme). In all cases, positive charges are shown blue and are placed on lysine NZ or arginine NH1; negative charges are shown in red and are placed on Asp OD1 or Glu OE1 (this is an arbitrary choice of one of the two carboxylate oxygens). In the case of the *Pa. denitrificans* structure, charges were placed according to the sequence on a position on the *Ps. aeruginosa* side chains most closely corresponding to that for a Glu, Lys, Asp, or Arg charge. A circle of diameter 28 Å is shown on each structure which approximately corresponds to the face of cytochrome *c*-551. Hemes are shown in yellow.

X-ray structure shows us that this is not so. Despite a careful search, Fulop et al. (1995b) found no Ca^{2+} site other than site I. Segment G58−P66 (72−80 in Pad CCP) as well as the known Ca^{2+} coordination ligands are included in Figure 9. The structure of the fully oxidized Psa CCP demonstrates that (1) some of the groups (such as the side chain of D63 (D77 in Pad CCP) expected to bind to a Ca^{2+} are pointing inward rather than outward at the dimer interface, (2) the sequence segments in the two monomers do not lie close together in the dimer (not shown), and (3) although the segment lies very close to the residues known to be coordinating the site I Ca^{2+}, they are clearly too far away to be participating in the coordination themselves. However, we have noted above that site II is only high affinity in the mixed valence state and that there may well be conformational changes in the transition to that state. Perhaps there are rearrangements at the dimer interface in the mixed valence enzyme so that the β-roll-like sequences are brought close together.

The potential Ca^{2+} binding sequence immediately following the proximal histidine ligand (His55) of the N-terminal domain recalls the Ca^{2+} binding to lignin peroxidase (Poulos et al., 1993). In this enzyme, serine 177 contributes its carbonyl and hydroxyl groups to the coordination of a Ca^{2+} and lies adjacent to the proximal heme ligand His176. The remainder of the Ca^{2+} coordination sphere is formed by a type I reverse turn from Asp194 to Ile199.

The Charged Surface and Interaction with Donor Cytochromes. The class I cytochromes *c* are diffusable carriers shuttling between donor and acceptor redox centers which are often part of fixed proteins in the membrane. In the particular case of mitochondrial cytochrome *c*, electron transfer takes place at an exposed heme edge surrounded by a pattern of surface positive charge that allows specific interaction with the redox partner. Similarly, Williams et al. (1995) have argued that the positively charged hydrophobic patches around a partly exposed redox center may be a common feature of the periplasmic electron shuttle proteins of bacteria. The electron donor to Pad CCP is cytochrome *c*-550 (Pettigrew, 1991), a close relative of mitochondrial cytochromes *c*. This cytochrome carries a net negative charge but has a highly skewed distribution of lysines on the molecular surface and a positively charged front face (Pettigrew, 1991; Williams et al., 1995) (Figure 10c). The electron donor to Psa CCP is a cytochrome *c*-551 (Soininen & Ellfolk, 1972) which is a less close relative of the mitochondrial cytochromes *c*. Cytochrome *c*-551 also carries a net negative charge and also excludes negative residues from the front face (Figure 10a). However, there is very little cross-reactivity of these cytochromes with the nonphysiological peroxidase (Pettigrew, 1991). A reasonable hypothesis is that fruitful collision of these donor cytochromes with their respective peroxidases involves interaction of the front face of the donor cytochrome with an

equivalent but complementarily charged heme face of the peroxidase. We would expect the surface charge features of the two peroxidases to differ in order to accommodate their specificity for their physiological donor. One edge of the heme of the C-terminal domain is indeed exposed at the molecular surface, and the surrounding surface charge topographies of the Psa CCP and the Pad CCP are shown in Figure 10b, and d. The predominantly negative features of the Pad CCP are consistent with its binding the very positive front face of the Pad cytochrome *c*-550. The heme of the N-terminal domain also has an exposed edge (not shown), but the surrounding surface does not have the expected differences in charge that we see around the C-terminal heme edge. We conclude therefore that the C-terminal heme is the likely site of electron entry. This is consistent with its His-Met coordination and the expected high potential that we associate with such a coordination (Moore & Pettigrew, 1990b).

SUPPORTING INFORMATION AVAILABLE

Five figures showing HPLC chromatograms of peptides under various experimental conditions and seven tables giving sequence and mass data and amino acid composition analyses of peptides discussed under Results (20 pages). Ordering information is given on any current masthead page.

REFERENCES

Aebersold, R. H., Leavitt, J., Saavedra, R. A., Hood, L. E., & Kent, S. H. B. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6970−6974.

Arciero, D. M., & Hooper, A. B. (1994) *J. Biol. Chem. 269*, 11878−11886.

Baumann, U., Wu, S., Flaherty, K. M., & McKay, D. B. (1993) *EMBO J. 12*, 3357−3364.

Benning, M. M., Meyer, T. E., & Holden, H. M. (1994) *Arch. Biochem. Biophys. 310*, 460−466.

Chistoserdov, A. Y., Chistoserdova, L. V., McIntire, W. S., & Lidstrom, M. E. (1994a) *J. Bacteriol. 176*, 4052−4065.

Chistoserdov, A. Y., McIntire, W. S., Mathews, F. S., & Lidstrom, M. E. (1994b) *J. Bacteriol. 176*, 4073−4080.

Ellfolk, N., Ronnberg, M., Aasa, R., Andreasson, L. E., & Vanngard, T. (1983) *Biochim. Biophys, Acta 743*, 23−30.

Ellfolk, N., Ronnberg, M., & Osterlund, K. (1991) *Biochim. Biophys. Acta 1080*, 68−77.

Foote, N., Peterson, J., Gadsby, P. M. A., Greenwood, C., & Thomson, A. J. (1984) *Biochem. J. 223*, 369−378.

Fulop, V., Moir, J. W. B., Ferguson, S. J., & Hajdu, J. (1995a) *Cell 81*, 369−377.

Fulop, V., Ridout, C. J., Greenwood, C., & Hajdu, J. (1995b) *Structure 3*, 1225−1233.

Geourjon, C., & Deleage, G. (1994) *Protein Eng. 7*, 157−164.

Gilmour, R., Goodhew, C. F., Pettigrew, G. W., Prazeres, S., Moura, I., & Moura, J. J. (1993) *Biochem. J. 294*, 745−752.

Gilmour, R., Goodhew, C. F., Pettigrew, G. W., Prazeres, S., Moura, J. J., & Moura, I. (1994) *Biochem. J. 300*, 907−914.

Gilmour, R., Prazeres, S., McGinnity, D. F., Goodhew, C. F., Moura, J. J. G., Moura, I., & Pettigrew, G. W. (1995) *Eur. J. Biochem. 234*, 878−886.

Goodhew, C. F., Wilson, I. B., Hunter, D. J., & Pettigrew, G. W. (1990) *Biochem. J. 271*, 707−712.

Hanlan, S. P., Hold, R. A., & McEwan, A. G. (1992) *FEMS Lett. 97*, 283−288.

Iobbi-Nivol, C., Crooke, H., Griffiths, L., Grove, J., Hussain, H., Pommier, J., Mejean, V., & Cole, J. A. (1994) *FEMS Microbiol. Lett. 119*, 89−94.

Laemmli, U. K. (1970) *Nature 227*, 680−685.

Matsuura, Y., Takano, T., & Dickerson, R. E. (1982) *J. Mol. Biol. 156*, 389−409.

McGinnity, D. F., Devreese, B., Prazeres, S., Van Beeumen, J., Moura, I., Moura, J. J. G., & Pettigrew, G. W. (1996) *J. Biol. Chem. 271*, 11126−11123.

Moore, G. R., & Pettigrew, G. W. (1990a) in *Cytochromes c. Evolutionary, Structural and Physicochemical Aspects*, pp 115−159, Springer-Verlag, Berlin, Germany.

Moore, G. R., & Pettigrew, G. W. (1990b) in *Cytochromes c. Evolutionary, Structural and Physicochemical Aspects*, pp 27−113, Springer-Verlag, Berlin, Germany.

Pettigrew, G. W. (1991) *Biochim. Biophys. Acta 1058*, 25−27.

Poulos, T. L., Edwards, S. L., Wariishi, H., & Gold, M. H. (1993) *J. Biol. Chem. 268*, 4429−4440.

Prazeres, S., Moura, J. J. G., Moura, I., Gilmour, R., Goodhew, C. F., Pettigrew, G. W., Ravi, N., & Huynh, B. H. (1995) *J. Biol. Chem. 270*, 24264−24269.

Ridout, C. J., James, R., & Greenwood, C. (1995) *FEBS Lett. 365*, 152−154.

Ronnberg, M., & Ellfolk, N. (1979) *Biochim. Biophys. Acta 581*, 325−333.

Samyn, B., Berks, B. C., Page, M. D., Ferguson, S. J., & Van Beeumen, J. J. (1994) *Eur. J. Biochem. 219*, 585−594.

Samyn, B., Van Craenenbroeck, K., De Smet, L., Vandenberghe, I., Pettigrew, G., & Van Beeumen, J. (1995) *FEBS Lett. 377*, 145−149.

Sofia, H. J., Burland, V., Daniels, D. L., Plunkett, G., & Blattner, F. R. (1994) *Nucleic Acids Res. 22*, 2576−2586.

Soininen, R., & Ellfolk, N. (1972) *Acta Chem. Scand. 26*, 861−872.

Sone, N., & Yangita, Y. (1982) *Biochim. Biophys. Acta 682*, 216−226.

Takano, T., & Dickerson, R. E. (1981) *J. Mol. Biol. 153*, 79−94.

van der Palen, C. J., Slotboom, D. J., Jongejan, L., Reijnders, W. N., Harms, N., Duine, J. A., & van Spanning, R. J. (1995) *Eur. J. Biochem. 230*, 860−871.

Williams, P. A., Fulop, V., Leung, Y. C., Chan, C., Moir, J. W. B., Howlett, G., Ferguson, S. J., Radford, S. E., & Hajdu, J. (1995) *Nat. Struct. Biol. 2*, 975−982.

BI963131E