Yagi, T., Vik, S. B., & Hatefi, Y. (1982) *Biochemistry 21*, 4777-4782.
Yang, C. S., & Huennekens, F. M. (1970) *Biochemistry 9*, 2127-2133.
Yang, X., & Trumpower, B. L. (1988) *J. Biol. Chem. 263*, 11962-11970.
Yu, C. A., & Yu, L. (1980) *Biochemistry 19*, 5715-5720.
Yu, C. A., & Yu, L. (1982) *Biochemistry 21*, 4096-4101.
Yu, C. A., Yu, L., & King, T. E. (1974) *J. Biol. Chem. 249*, 4905-4910.
Zhu, Q. S., Van Der Wal, H. N., Van Grondelle, R., & Berden, J. A. (1983) *Biochim. Biophys. Acta 725*, 121-130.

Three-Dimensional Structure of *p*-Cresol Methylhydroxylase (Flavocytochrome *c*) from *Pseudomonas putida* at 3.0-Å Resolution†,‡

F. Scott Mathews,* Zhi-wei Chen, and Henry D. Bellamy§
Department of Cell Biology and Physiology, Washington University School of Medicine, St. Louis, Missouri 63110
William S. McIntire
Molecular Biology Division, Veterans Administration Medical Center, San Francisco, California 94121, and Department of Biochemistry and Biophysics and Department of Anesthesia, University of California, San Francisco, California 94143
Received January 19, 1990; Revised Manuscript Received August 31, 1990

ABSTRACT: *p*-Cresol methylhydroxylase (PCMH) isolated from *Pseudomonas putida* is an $\alpha_2\beta_2$ tetramer of approximate subunit M_r 49 000 and 9000. It is a flavocytochrome *c* containing covalently bound FAD in the larger subunit and covalently bound heme in the smaller. Crystals in space group $P2_12_12_1$ with unit-cell parameters a = 140.3 Å, b = 130.6 Å, and c = 74.1 Å contain one full molecule per asymmetric unit and diffract anisotropically to about 2.8-Å resolution in two directions and to about 3.3-Å resolution in the third. An electron density map has been computed at a nominal resolution of 3.0 Å by use of area detector data from native crystals and from two derivatives. The phases were improved with the B. C. Wang solvent leveling procedure, and the map was averaged about the noncrystallographic 2-fold axis. The cytochrome subunit, whose amino acid sequence is known, has been fitted to the electron density on a graphics system. The course of the polypeptide chain of the flavoprotein subunit, whose sequence is mostly unknown, has been traced in a minimap and a model of polyalanine fitted to the electron density on the graphics system. The flavoprotein subunit consists of three domains in close contact. The N-terminal domain consists largely of β-structure and contains most of the FAD binding site. The second domain contains a seven-stranded antiparallel β-sheet of unusual topology connected by antiparallel α-helices on one side. The flavin ring lies at the juncture of the first two domains. The third domain lies against the first domain and helps cover the rest of the FAD chain. The cytochrome subunit resembles other small cytochromes such as *c*-551 and c_5 and fits into a depression on the surface of the large flavoprotein subunit. The flavin and heme planes are nearly perpendicular, the normals to the planes being approximately 65° apart. The two groups are separated by about 8 Å, the distance from one of the vinyl methylene carbon atoms of the heme to the 8α-methyl group of the flavin ring.

p-Cresol methylhydroxylase (PCMH)[1] is a flavocytochrome *c* found in the periplasmic space of certain pseudomonads. It catalyzes the first steps in the oxidation of *p*-cresol to *p*-hydroxybenzaldehyde (Hopper & Taylor, 1977). The product is then converted to *p*-hydroxybenzoic acid and subsequently undergoes ring fission in later catalytic steps. During catalysis, hydroxylation of *p*-cresol occurs by the abstraction of two hydrogen atoms by the flavoprotein after which the putative *p*-quinone methide intermediate is believed to be attacked by water to yield *p*-hydroxybenzyl alcohol (Hopper, 1976). The two electrons are then transferred one at a time to the heme on the cytochrome subunit and then to an acceptor protein in vivo (Hopper, 1978). The acceptor protein is believed to be an azurin (McIntire et al., 1985). The product of the reaction, *p*-hydroxybenzyl alcohol, can also serve as substrate for the enzyme in a dehydrogenation reaction yielding *p*-hydroxybenzaldehyde. In vitro, phenazine methosulfate can act as an artificial electron acceptor for the protein.

Seven forms of PCMH have been isolated from six bacterial sources (Koerbert et al., 1985; Bossert et al., 1989). A flavocytochrome *c* preferring 4-ethylphenol as its primary substrate, and which is closely related to PCMH, has also recently been characterized (Reeves et al., 1989). *Pseudomonas putida* strain NCIB 9869 produces two forms of PCMH (Keat & Hopper, 1978). One, form A, is carried on a plasmid and is expressed constitutively by cells grown on 3,5-xylenol. The other, form B, is chromosomally encoded and is induced by *p*-cresol. PCMH form A, the subject of this study, has a molecular weight of approximately 116000. It can be resolved

† This work has been supported by NSF Grants DMB-8816618 and DMB-8718741, USPHS Grant HL-16251, and the Veterans Administration.
‡ Crystallographic coordinates for PCMH have been submitted to the Brookhaven Protein Data Bank.
* Corresponding author.
§ Present address: Stanford Synchrotron Radiation Laboratory, SLAC Bin 69, P.O. Box 4349, Stanford, CA 94309.

[1] Abbreviations: PCMB, *p*-(chloromercuri)benzoate; PCMH, *p*-cresol methylhydroxylase; PEG, poly(ethylene glycol); rms, root mean square.

0006-2960/91/0430-238$02.50/0

Table I

(A) Multiwire Data Collection Statistics

data set[a]	max resolution (Å)	total no. of observations	no. of unique reflections	R_{sym}[b]	$\langle I/\sigma \rangle$[c] 3.5 Å	$\langle I/\sigma \rangle$[c] 3.0 Å	ΔF_{iso} (%)[d]
native	2.8	176049	32693	0.062	9.6	3.4	
Hg-1	2.8	179922	32348	0.087	5.3	1.7	21.0
Hg-2	2.8	54076	16839	0.066	7.8	2.0	19.3
Pt-1	2.8	176614	32433	0.063	7.2	2.3	13.7
Pt-2	3.0	101647	27859	0.088	3.4	1.8	19.7

(B) Overall Anisotropic Thermal Parameters (Å^2)

	B_{iso}[e]	B_{11}	B_{12}	B_{13}	B_{22}	B_{23}	B_{33}
native	24.6	16.0	6.0	−5.4	41.4	5.4	18.2
Hg-1	32.8	23.8	5.1	−5.1	53.4	3.2	29.9
Pt-1	26.8	18.9	4.1	−4.1	47.6	3.2	21.7

[a] Hg-1, Hg-2, and Pt-1 refer to the data sets collected from two PCMB and one K_2PtCl_4-soaked crystal collected at the UCSD area detector facility. Pt-2 refers to the combined data from four K_2PtCl_4-soaked crystals recorded at the Argonne area detector facility. [b] $R_{sym} = \sum_{hkl}\{\sum_{i=1}^{N}|\langle I(hkl)\rangle - I(hkl)_i|\}/\{\sum_{hkl}\sum_{i=1}^{N} I(hkl)_i\}$. [c] $\langle I/\sigma \rangle$ is the mean ratio of intensity to standard deviation recorded for the data set in a narrow resolution range centered at 3.5- or 3.0-Å resolution. For the Pt-2 data set this represents the average value from four crystals. [d] ΔF_{iso} is the mean isomorphous structure factor change. [e] Equivalent isotropic thermal parameter.

by isoelectric focusing into two components which differ by about 0.7 in p*I* (McIntire et al., 1984). One component, of subunit M_r 49000, contains FAD covalently bound through the 8α-position (McIntire et al., 1981) to tyrosine. The other, of M_r 9000, is a *c*-type cytochrome. The amino acid sequences of the complete cytochrome subunit, the first 56 residues of the flavoprotein subunit, and the FAD binding heptapeptide are known (McIntire et al., 1986).

The structure of *p*-cresol methylhydroxylase was first determined at 6.0-Å resolution from X-ray data collected on a Picker diffractometer (Shamala et al., 1986). The crystals are orthorhombic, space group $P2_12_12_1$, with cell parameters a = 140.3 Å, b = 130.6 Å, and c = 74.1 Å. A four-site mercury [*p*-(chloromercuri)benzoate, PCMB] and a five-site platinum (K_2PtCl_4) heavy-atom derivative were used, with their anomalous scattering, to determine the X-ray phases. The iron positions were determined from an anomalous difference Fourier map. The asymmetric unit was shown to contain an $\alpha_2\beta_2$ tetramer with each pair of subunits related by a noncrystallographic 2-fold axis of symmetry.

The present paper describes the extension of the crystallographic analysis of *p*-cresol methylhydroxylase to a nominal resolution of 3.0 Å with X-ray data collected by area detector methods. Preliminary results at an early stage of the 3.0-Å resolution analysis have been described (Bellamy et al., 1987).

Materials and Methods

Crystals of native PCMH were prepared by the free interface diffusion method using poly(ethylene glycol) (PEG-8000, Fisher Scientific) as described previously (Shamala et al., 1985). Heavy-atom derivatives were prepared by soaking crystals in a solution containing the heavy-atom compound dissolved in 14% (w/v) PEG-8000, 0.05 M phosphate buffer (NaH_2PO_4:K_2HPO_4 = 40:60, ca. pH 7.0), and 0.1 M NaCl. For all the derivative crystals the soaking time was 2 days at 1 mM heavy-atom concentration at room temperature, except for one crystal (Pt-1, see below) for which the soaking time was 1 day at 0.5 mM concentration.

Data were collected to 2.8-Å resolution from four crystals at the area detector facility of the University of California at San Diego (Xuong et al., 1985). Full data sets were recorded from one crystal of the native protein and one of a platinum (Pt-1) and one of a mercury (Hg-1) derivative, each with approximately 3–4-fold redundancy. A partial data set from a second crystal of the mercury derivative (Hg-2) was also recorded, but data collection was stopped midway through because of excessive radiation damage. Additional data were recorded from four crystals of the platinum derivative at 3.0-Å resolution, on a Nicolet area detector at the Argonne National Laboratories, and merged into a single data set (Pt-2). The data collection statistics are summarized in Table IA.

The PCMH crystals were found to diffract anisotropically at higher resolution when examined on the area detector. Several preparations of crystals were tested, some growing with differing axial ratios, but all crystals examined displayed the same anisotropy. Diffraction from the native crystals extends to approximately 2.8-Å resolution in the *a* and *c* directions but only to approximately 3.3 Å in the *b* direction. The anisotropy of the diffraction was analyzed by the method of Sheriff and Hendrickson (1987). The overall anisotropic thermal parameters are shown in Table IB for the native and two of the derivative crystals.

The Friedel-related data were locally scaled (Matthews & Czerwinski, 1978) for all data recorded on the area detectors. The derivative data were then locally scaled to the native data. In this procedure, a local scale factor for a related pair of reflections is obtained from the sums of the intensities of the 124 reflections surrounding each point, provided at least 25 of these paired reflections were observed. In the case of data set Pt-2, the local scaling was carried out for each of the four crystals before the data were merged into a single data set. The mean changes in structure factor from the native crystal for each of the derivative area detector data sets are given in Table IA.

The native data recorded earlier at 6.0-Å resolution on a diffractometer (Shamala et al., 1986) were also locally scaled to the native area detector data and merged with it. The four low-resolution derivative data sets (Hg-a, Hg-b, Pt-a, and Pt-b) were locally scaled to the combined native data set and used in phase determination.

It was decided to limit the analysis to 3.0-Å resolution overall. With the area detector data sets, the positional and isotropic thermal parameters for the heavy-atom sites were refined. Starting with values from the 6.0-Å resolution results (Shamala et al., 1987), the parameters were refined by use of the origin-removed difference Patterson procedure HEAVY (Terwilliger et al., 1987). With the low-resolution derivative data sets the heavy-atom positional and occupancy parameters were further refined, to compensate for slight changes introduced by merging the native area detector and diffractometer data sets. The refined occupancies and thermal parameters are given in Table II.

Protein phases were computed by the multiple isomorphous replacement (MIR) method including anomalous scattering.

Table II: Relative Occupancies and Temperature Factors for the Eight Derivative Data Sets Used in MIR Phase Calculation and as Input to the Wang Phase Filtering Procedure[a]

derivative		site 1	site 2	site 3	site 4	site 5
PCMB (Hg-1)	K (rel)	3.38	2.92	2.42	2.86	
	B (Å^2)	74	55	54	74	
PCMB (Hg-2)	K (rel)	3.07	2.53	3.25	2.81	
	B (Å^2)	66	51	87	84	
PCMB (Hg-a)	K (rel)	2.84	2.53	2.17	1.71	
PCMB (Hg-b)	K (rel)	2.91	2.76	2.43	2.26	
K_2PtCl_4 (Pt-1)	K (rel)	1.48	1.72	2.43	2.22	1.48
	B (Å^2)	48	71	148	126	186
K_2PtCl_4 (Pt-2)	K (rel)	3.24	3.27	4.70	3.61	3.92
	B (Å^2)	64	79	129	89	320
K_2PtCl_4 (Pt-a)	K (rel)	3.71	3.32	2.96	3.27	1.58
K_2PtCl_4 (Pt-b)	K (rel)	3.42	3.12	2.76	2.87	1.45

[a] For the first two data sets (area detector data; see Table I) of each derivative the occupancies (K) and temperature factors (B) were refined to 3.0-Å resolution. For the last two crystals of each derivative (6.0-Å resolution diffractometer data) only the occupancies were refined while the temperature factors were held fixed at a value of 34. Sites 1 and 2 and sites 3 and 4 of each derivative are related by the noncrystallographic 2-fold symmetry of the molecule.

The presence of an anomalous scattering signal was verified for a given data set by examining anomalous difference Fourier maps with MIR phases without anomalous scattering. For all data sets except Pt-2, the highest peaks in the maps corresponded to the anomalous scattering centers. No anomalous data from Pt-2 or from Hg-a were used. The overall figure of merit was 0.56. The phase calculation statistics are summarized in Table III. The MIR phases, including anomalous scattering, were filtered by the Wang (1985) procedure. A solvent content of 55% was used in the calculations, a value about 3% lower than observed for the crystals (Shamala et al., 1986). Twelve cycles of phase filtering, three separate solvent masks being used, were carried out, during which the figure of merit rose from an initial value of 0.49 to a final value of 0.68.

The electron density was averaged about the local noncrystallographic 2-fold axis by the method of Bricogne (1976) and plotted as a minimap. The orientation of the noncrystallographic 2-fold axis was refined by least squares in order to match the positions of equivalent heavy atoms as described previously (Shamala et al., 1986). The rms deviation of equivalent positions was 0.46 Å. The path of the polypeptide chain in the cytochrome subunit was traced in the minimap, and the approximate positions of 74 α-carbon atoms were identified. The shape of the heme group in the density was clearly recognizable. No attempt was made to fit the side chains to the minimap density. The main chain of the flavoprotein subunit was also traced in the minimap. In this case a total of 462 α-carbon positions were located. Electron density for the FAD moiety could be easily identified. The approximate coordinates for the α-carbon atoms for both subunits were read from the minimap and entered into the computer for use on the molecular graphics system.

On an MMS-X molecular graphics system (Barry et al., 1974) contours were generated within a 12-Å sphere centered on the known iron position. The graphics program NEWNIP (Lederer et al., 1981) was used. The orientation of the heme group in the density was determined by placing the heme group and essential ligands of cytochrome c-551 (Matsuura et al., 1982) onto the iron site of PCMH and rotating the rigid group to get the best fit to the density. The essential ligands consisted of residues 12–16 (Cys-Val-Ala-Cys-His) and residue 61 (Met) which contain the atoms which form covalent and coordinate bonds to the heme. One single orientation of the heme aggregate gave a superior fit to the density.

The main- and side-chain atoms of the cytochrome subunit were then fitted to the electron density on an IRIS graphics system using FRODO (Jones, 1985). This was done by placing large side chains ("marker groups") into prominent features of side-chain density and fitting the intervening density with the smaller side chains. In some cases it seemed necessary to omit a residue from the amino acid sequence in order to stay in register with the larger marker groups. However, subsequent refitting of the model allowed the omitted residues to be put back into the chain and a good fit to the density still to be maintained. In this manner all 78 residues of the cytochrome were placed in the density map. At this point three short helices were identified, and the stereochemistry of the model was regularized (Hermans & McQueen, 1974) with helical constraints.

Since its amino acid sequence is largely unknown, the flavoprotein subunit was fitted with a standard polyalanine backbone chain on the IRIS system using the α-carbon coordinates as guide points. The model was also regularized, but without secondary structural constraints. Throughout the chain, the tracing from the minimap was verified on the graphics system. The model for FAD was adjusted manually to give a good fit to the electron density. The adenosine moiety and the ribityl chain of the flavin could be fitted unambiguously. The orientation of the isoalloxazine ring could not be identified by shape alone. However, on one orientation there was a clear connection to the polypeptide at the 8α-methyl position, which is the site of covalent attachment of the flavin to the protein.

The flavin binding heptapeptide was fitted to the electron density, starting from the position of the tyrosine residue. Both directions along the polypeptide chain were tested with the

Table III: MIR Phase Calculation Statistics

	maximum resolution of shell (Å)						
	18.0	9.0	6.0	4.5	3.6	3.0	overall
no. of reflections	119	883	2418	4836	7776	11120	30752
av F_p[a]	30.6	32.5	21.3	20.7	17.5	10.5	16.1
av m[b]	0.87	0.92	0.89	0.74	0.55	0.38	0.56
phasing power[c]							
Hg-1	0.6	1.9	1.7	1.1	0.7	0.6	1.0
Hg-2			2.1	1.5	1.0	0.8	1.3
Pt-1	0.9	1.9	1.7	0.8	0.5	0.4	0.8
Pt-2	0.9	2.0	2.1	1.2	0.7	0.4	1.0
Hg-a[d]	2.2	2.3	2.0				2.1
Hg-b[d]	2.1	2.7	2.6				2.5
Pt-a[d]	1.9	2.1	1.6				1.7
Pt-b[d]	1.6	2.1	2.1				2.0

[a] F_p = native protein structure factor amplitude. [b] m = figure of merit. [c] Phasing power = $\langle f_H \rangle / \langle e_H \rangle$, where $\langle f_H \rangle$ is the rms-calculated heavy-atom structure factor and $\langle e_H \rangle$ is the rms isomorphous lack of closure error. [d] Hg-a, Hg-b, Pt-a, and Pt-b are the 6-Å diffractometer derivative data sets.

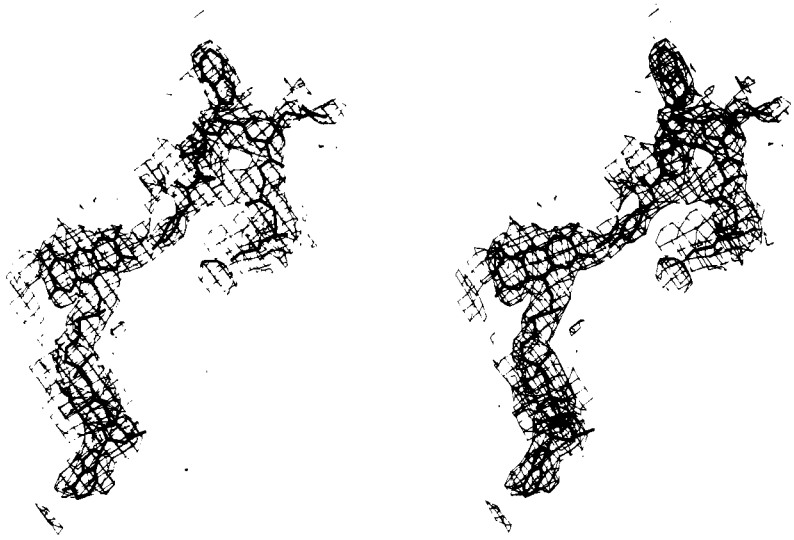

FIGURE 1: Skeletal model of FAD (lower half of diagram) and the covalently bound peptide Tyr-Asn-Trp-Arg-Gly-Gly-Gly (upper half of diagram) superimposed on the corresponding electron density.

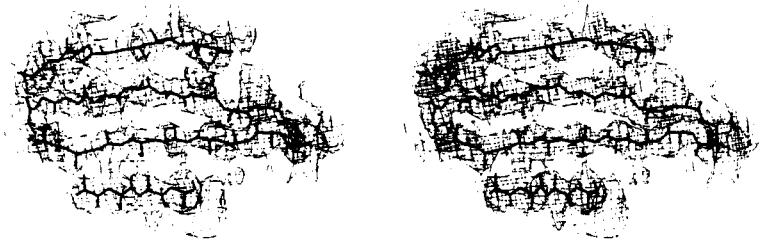

FIGURE 2: Electron density of parts of the first four β-strands (15, 12, 14, 13) of domain 2. The polyalanine backbone residues shown (starting from the upper left) are 313–310, 217–225, 289–281, and 246–253.

heptapeptide model, and one of these gave a clearly superior fit to the density. This procedure established the amino-terminal to carboxyl-terminal direction for the polypeptide chain. Examination of the disposition of side chains of several long helices indicated in most cases a directionality consistent with this chosen directions; i.e., most of them pointed toward the amino terminus. In the other cases no conclusions could be drawn. An attempt was also made to fit the amino-terminal 56 residues into the electron density. Even though the main-chain density was clear, the side-chain density was rather weak in this region so that it was not possible to match the side chains satisfactorily.

Results and Discussion

Phase Calculation and Electron Density Fitting. The anisotropic scattering of the PCMH crystals complicates the structure analysis since the resolution of the map will vary in different directions. It is also difficult to assess quantitatively the effective resolution in different directions on the basis of area detector images. The average I/σ for the native data, where I is the observed intensity and σ is its standard deviation, is about 3.4 at 3.0-Å resolution, and this resolution limit was selected for the current analysis. The phasing power (Table III) for the derivative crystals falls below unity at about 3.6-Å resolution, which is a more realistic estimate of the effective limit to the MIR phasing. However, the phase filtering procedure and 2-fold averaging serve to improve the resolution of the map.

The polypeptide chain could be traced with few apparent ambiguities. The relative reliability of the chain tracing is greatest in the cytochrome subunit since the sequence is known and the structure is homologous to other *c*-type cytochromes (see below). The chain tracing for the flavoprotein subunit is inherently less reliable since its sequence is unknown and the resolution is limited to a nominal value of 3.0 Å. When the polypeptide chain of the flavoprotein subunit was traced in the minimap, its pathway could be followed with confidence in most places. However, there were several places where the adjacent main-chain density was connected. Most of these were in domain 1, but two were in domain 2 (see below). In each case alternate pathways were tested but were found to lead to topological difficulties such as leaving some density unconnected. Thus the present chain tracing is the most self-consistent one and most likely to be correct. However, the structural results should be treated with some caution since errors in interpretation may still exist.

In the cytochrome subunit the iron atom is located at the highest peak, and the three nearby sulfur atoms (two cysteines and a methionine) are also located in high density. In the flavoprotein subunit, the pyrophosphate group of FAD lies in a sausage of density corresponding to the highest electron density of that subunit. The electron density of the FAD group and the FAD binding heptapeptide is shown in Figure 1. In Figure 2 is shown the electron density for four antiparallel β-strands of the second domain of the flavoprotein subunit (see below) with the polyalanine backbone superimposed. The number of residues fitted to the electron density of the flavoprotein subunit, 462, is somewhat larger than the number of residues, approximately 442, determined by amino acid analysis (McIntire et al., 1986).

Quaternary Structure. The $\alpha_2\beta_2$ tetramer of PCMH is shown in Figure 3. The molecule is a prolate spheroid of approximate dimensions 110 Å × 75 Å × 60 Å. The flavoprotein subunits are tightly packed about the molecular 2-fold axis, sharing a common interface. The interface is roughly rectangular with an area of approximately 2000 Å^2. At the

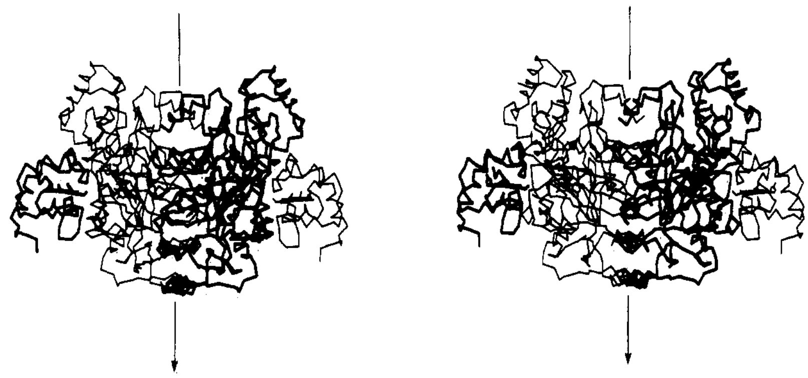

FIGURE 3: $\alpha_2\beta_2$ tetramer of PCMH with the molecular 2-fold axis vertical.

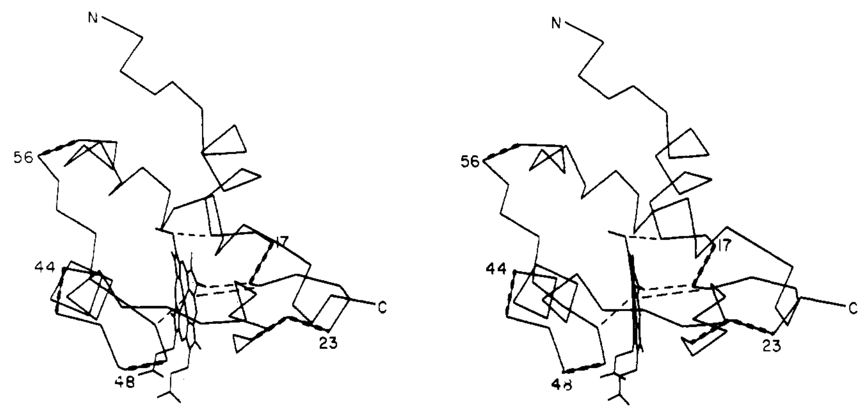

FIGURE 4: Backbone α-carbon tracing of the cytochrome subunit of PCMH in the "standard" orientation. The N- and C-termini are indicated. Bonds from the four heme ligands, Cys 15, Cys 18, His 19, and Met 50, are drawn as lightly dashed lines. The five polypeptide segments in close contact with the flavoprotein subunit, 17–18, 23–25, 44–45, 48–49, and 56–57, are heavily dashed. The numbering of the first α-carbon of each segment is also indicated.

interface, there is pairwise packing of domains 1 and 3 against domain 2. The cytochrome subunits lie on the periphery of the flavoprotein dimer and are nestled into depressions on its surface. The irons of the two heme groups are separated by about 76 Å while the center to center distance between flavin rings is about 43 Å.

Cytochrome Subunit. An α-carbon diagram of the cytochrome subunit is shown in Figure 4. It consists of a compact polypeptide chain which is essentially wrapped about the heme group. It shares, in common with all type I cytochromes (Ambler, 1982), a number of features characteristic of the "cytochrome *c* fold" (Almassy & Dickerson, 1978). The orientation shown in Figure 4 is the standard "front view" commonly used for this family. There are three helical segments formed by residues 8–14, 36–45, and 62–72 and three reverse turns, segments 2–5, 47–50, and 55–58. The remainder of the chain is in irregular or extended conformation. The amino acid sequence of the cytochrome subunit is shown in Figure 5.

The heme group is attached covalently through the vinyl groups at positions 2 and 4 to Cys 15 and Cys 18 of the cytochrome subunit, respectively. Coordinate bonds are also formed between the heme iron and both His 19 and Met 50. Even in the isolated subunit, most of the heme group appears to be shielded from solvent. Exceptions are the methyl and vinyl groups at position 3 and 4, the latter covalently attached to Cys 18, and the two propionate groups. The positions shown for these two propionates in Figure 4 are somewhat arbitrary since the electron density for them is weak. However, they are likely to extend somewhat into solution since the amino acid side chains in their vicinity are not close enough to fully shield them from solvent.

5 10 15 20
ASP SER GLN TRP GLY SER GLY LYS ASN LEU TYR ASP LYS VAL CYS GLY **HIS** **CYS** HIS LYS
(HELIX 1) HEME
25 30 35 40
PRO GLU **VAL** **GLY** **VAL** GLY PRO VAL LEU GLU GLY ARG GLY LEU PRO GLU ALA TYR ILE LYS
45 50 55 60
ASP ILE VAL **ARG** **ASN** GLY PHE **ARG** **ALA** MET PRO ALA PHE PRO ALA **SER** **TYR** VAL ASP ASP
(HELIX 2)
65 70 75
GLU SER LEU THR GLN VAL ALA GLU TYR LEU SER SER LEU PRO ALA PRO ALA GLN
(HELIX 3)

FIGURE 5: Amino acid sequence of the cytochrome subunit of PCMH. The three helical segments and four ligands to the heme are indicated. Also indicated, in underlined boldface, are the five segments in closest contact with the flavoprotein subunit.

There are 10 acidic side chains (Asp and Glu) and 7 basic ones (Lys and Arg). In addition, there are two propionate groups as well as the amino and carboxyl termini. This leads to a charge imbalance of 13 negative charges vs 8 positive charges at neutral pH, which is consistent with the p*I* of 4.55–4.70 observed in solution (McIntire et al., 1984). Two pairs of these side chains appear to form salt bridges, Lys 20 with Glu 30 and Arg 44 with Glu 61. The remaining charged side chains are distributed, somewhat unevenly, over the cytochrome surface. There is a slight concentration of acidic residues at the "back" of the molecule (defined in Figure 4) and an even mixture of acidic and basic side chains on the "left" and "right" sides. The "front" of the molecule, which

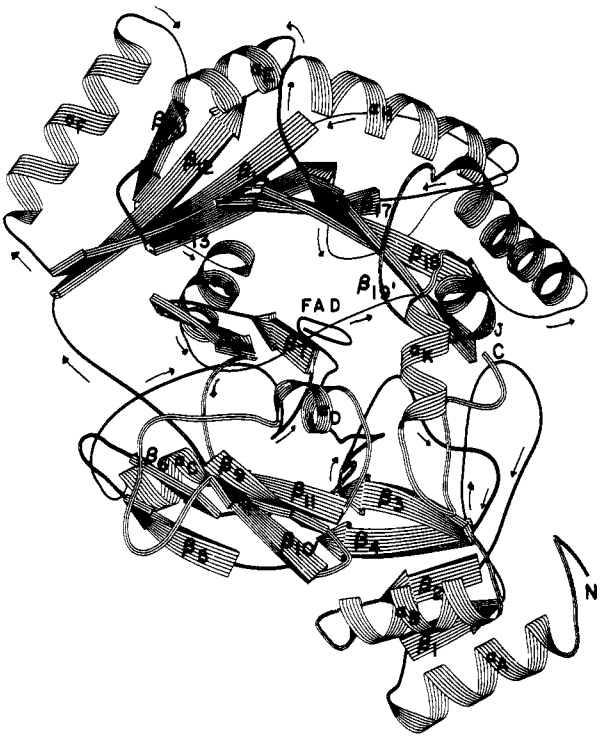

FIGURE 6: Cartoon of the flavoprotein subunit indicating elements of secondary structure. The positions of 11 helices (A–K) are indicated by spirals and of the 18 β-sheets by arrows. The N- and C-termini are also indicated. The FAD moiety is represented by a bent-paperclip-like figure. This picture was produced by the computer program RIBBON (Priestle, 1988).

faces the flavoprotein subunit (see below), is more or less neutral, except for Arg 44 and Arg 48 which seem to point back, away from the flavoprotein. The "bottom" of the molecule is also largely neutral, except for the two propionate groups and Arg 32, which might interact with the "rear" propionate.

There is one charged side chain, Asp 59, which appears to be located in the interior of the cytochrome. However, residues 56–59 were the most difficult to fit to the electron density map so the exact placement of Asp 59 may change after the protein structure is refined. Other regions in which difficulties in fitting were encountered were residues 21–23, where the side-chain densities were weak, and residues 3–4, where the density interpretation was somewhat ambiguous.

Flavoprotein Subunit. The overall conformation and secondary structure of the flavoprotein subunit is shown in Figure 6. It is folded into three domains (Figure 7). The N-terminal domain (domain 1, Figure 7a) consists of approximately the first 215 residues and envelopes most of the FAD cofactor. The central domain (domain 2, Figure 7b) consists of the next 208 residues and covers the flavin ring. The C-terminal domain (domain 3, Figure 7c) consists of the last 39 residues and lies against the N-terminal domain, covering the adenosine portion of the FAD.

The N-terminal domain contains four α-helices, A–D, and the first 11 β-strands. The first two helices are close to the N-terminus and are two to three turns in length (Figure 6). The other two are shorter and lie between extended chains. The 11 β-strands are divided into a four-stranded mixed β-sheet near the N-terminus (strands 1–4), a five-stranded antiparallel β-sheet (strands 5, 6, and 9–11), and a two-stranded β-ribbon (strands 7 and 8) which interrupts the five-stranded β-sheet. The remainder of the N-terminal domain consists of turns and extended chains. The ADP-ribose part of FAD is extended and lies in a groove in the N-termninal domain running parallel to several extended chains. This ADP-ribose-binding groove is covered by the small C-terminal domain, which contains two helices, J and K (Figure 6). The flavin ring protrudes from the body of domain 1. Its *re* face lies against elements of domain 1 while its *si* face is open to the interior of domain 2.

The pyrophosphate moiety of the FAD is surrounded by extended chains. The closest helix is D which is about 6 Å away with the pyrophosphate situated near the midpoint of the helix. Thus, in the case of PCMH, the pyrophosphate group is not located at the N-terminus of an α-helix.

The second domain is composed of a seven-stranded antiparallel β-sheet, strands 12–18, which forms a dome over the isoalloxazine ring of the flavin (Figure 6). The dome is covered on the outside by a side-by-side antiparallel arrangement of four α-helices, E, F, H, and I, and contains a fifth helix, G, on the inside. The site of covalent flavin attachment is at position 336 (using the present numbering scheme) and is located about four residues before the beginning of strand 16. There is another extended strand, 19′ (shown in Figure 6), which lies on the edge of the seven-stranded β-sheet and would form an eighth antiparallel strand except that it is somewhat oblique to the seventh strand (strand 16) and appears to form only one or two hydrogen bonds with it. However, for the purpose of the following discussion of topology it is useful to consider strand 19′ to be part of an eight-stranded antiparallel sheet.

The sequential order of the strands and helices in domain 2 is 12, E, 13, 14, F, 15, G, 16, H, 17, 18, I, 19′. The topological order of the strands in the β-sheet is 19′, 16, 18, 17, 13, 14, 12, 15 and of the helices in the side-by-side arrangement is I, H, E, F. In Richardson's topological notation (Richardson, 1977) the connectivity is $-2x$, $+1$, $+2x$, $-6x$, $+2x$, -1, $-2x$. The topology of this motif is shown schematically in Figure 8. The motif of the dome consists of two repeating structures each containing four antiparallel β-strands and two antiparallel α-helices, connected by an additional helix. The two structures are related by a 2-fold axis perpendicular to the sheet lying between strands 13 and 17. Each four-stranded structure is itself made up of a pair of interdigitating substructures each consisting of a βαβ structure. Each pair of βαβ structures is related topologically by a 2-fold axis normal to the sheet located between the innermost pair of interdigitated β-strands, i.e., between strands 12 and 14 and between strands 16 and 18 (Figure 8). Thus there appears to have been a structural duplication or possibly quadruplication in the evolutionary development of this domain.

Although rare, the basic four-stranded antiparallel motif found in the flavoprotein subunit is not unique. It has been observed in the eight-iron ferredoxin (Adman et al., 1973) and as part of the N-terminal domain of bacterial ribulose-1,5-bisphosphate carboxylase (Schneider et al., 1986). However, the duplication to the eight-stranded form, forming the FAD-containing dome-like structure of PCMH does appear to be unique.

Cytochrome–Flavoprotein Interactions. The common interface between the flavoprotein and cytochrome subunits seems quite extensive. The flavin group appears to be well shielded from solvent, even in the isolated flavoprotein subunit. However, the edge of pyrrole ring II of the heme, which is shielded from solvent in the complex, would probably be exposed in the isolated cytochrome subunit.

The cytochrome binding site on the flavoprotein comes mainly from domain 1 with some contribution from domain

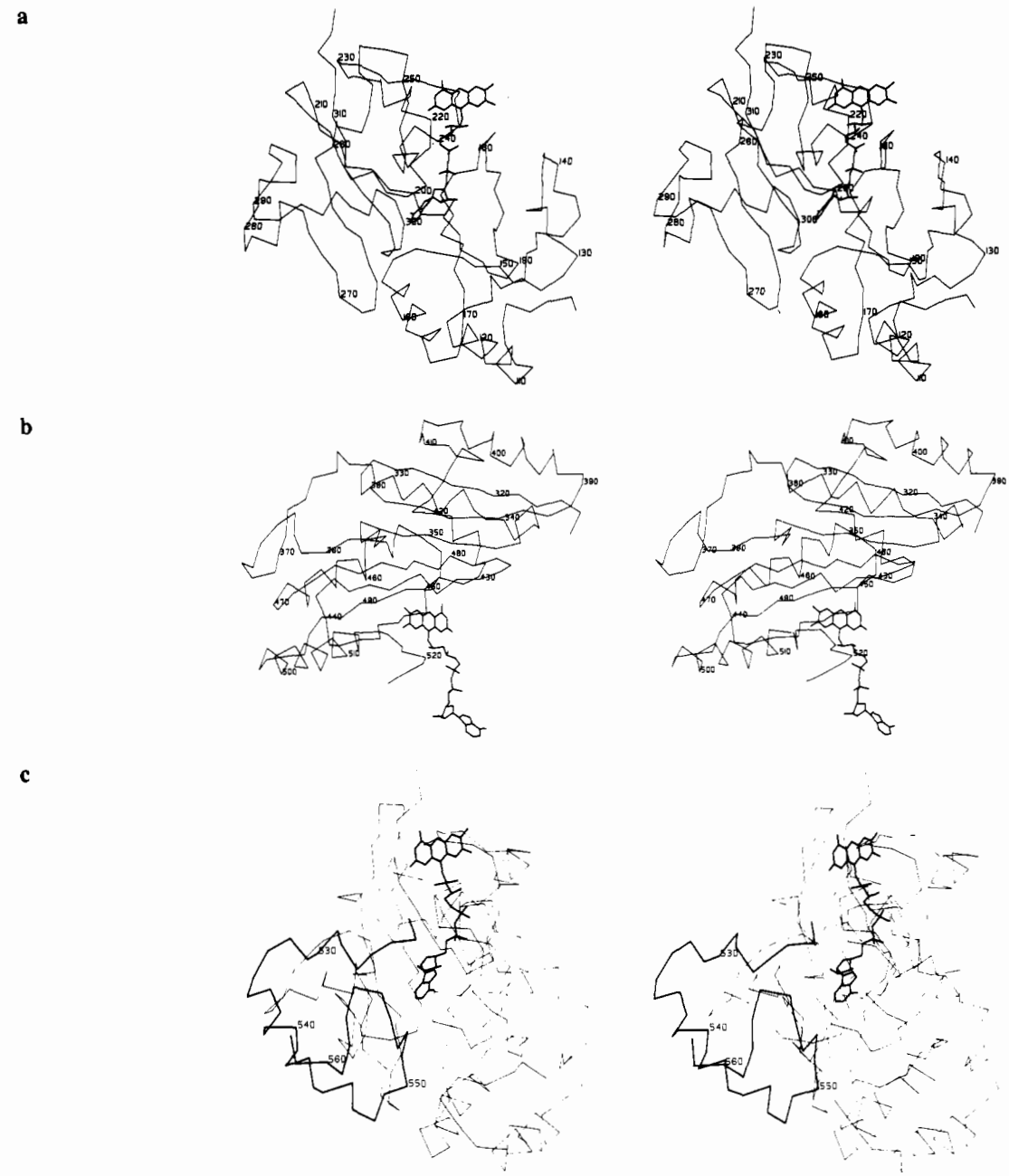

FIGURE 7: Backbone α-carbon tracing of the individual domains of the flavoprotein subunit. Every tenth α-carbon atom is numbered starting with 101 for the first residue for the flavoprotein subunit. The FAD skeleton is also included. (a) Domain 1, residues 101–315. (b) Domain 2, residues 316–523. (c) Domain 3, residues 524–562. Domain 3 lies against domain 1 which is shown, unnumbered, in light lines.

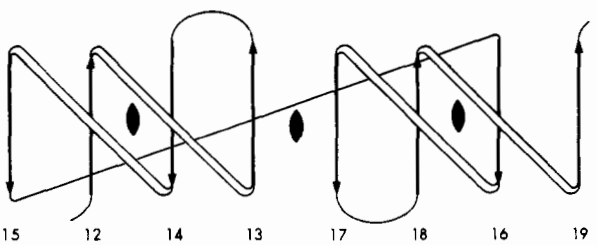

FIGURE 8: Topological diagram of domain 2 of PCMH viewed perpendicular to the β-sheet. The thick diagonal lines represent crossover connections above the page while the thin diagonal line indicates a crossover connection below the page. The topological 2-fold axes perpendicular to the page are indicated by the solid ellipsoids.

2. The portion of the cytochrome subunit situated at the interface with the flavoprotein consists mostly of the "front" face as viewed in Figure 4. The approximate chain segments of the cytochrome generally making up the interface region are 10–19, 23–25, and 44–57. The segments which make closest contact with the flavoprotein subunit, indicated in Figures 4 and 5, are 17–18, 23–25, 44–45, 48–49, and 56–57.

The relative orientations of the heme and flavin cofactors are shown in Figure 9. The heme iron is about 18 Å from the center of the flavin ring, and the closest approach of the two groups is about 8 Å. The benzenoid portion of the isoalloxazine ring is oriented toward the thioether-containing edge of the heme group. The flavin and heme planes are far from being coplanar, making an angle of about 65° to each other.

The most direct pathway for electron transfer from the flavin to the heme is via the benzene ring of the flavin to pyrrole ring II of the heme. Similar pathways were proposed for the electron transfer from flavodoxin to cytochrome *c* in the computer model of Simondson et al. (1982) and from flavodoxin to one of the hemes in the tetraheme cytochrome c_3 for the computer model of Stewart et al. (1988) in the respective flavodoxin–cytochrome complexes. In trimethylamine de-

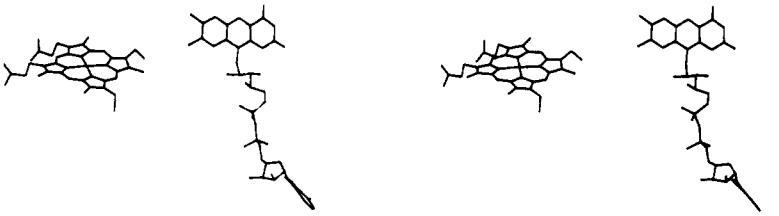

FIGURE 9: Stereoscopic view of the heme and flavin cofactors in PCMH.

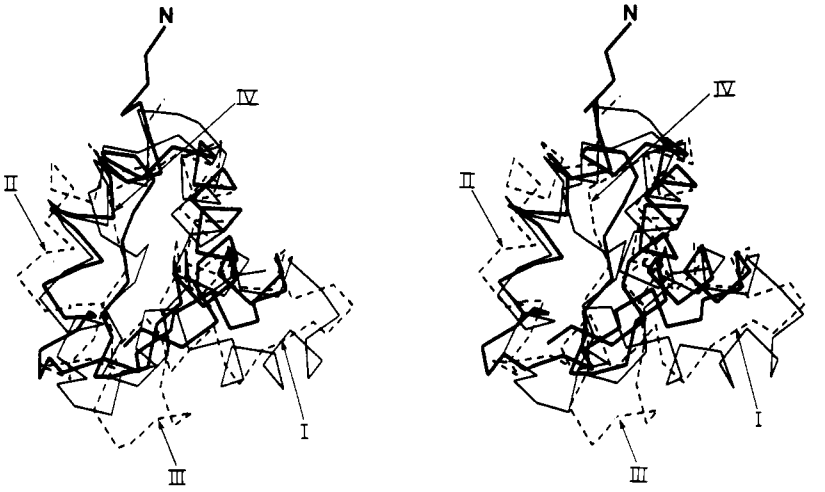

FIGURE 10: Comparison of the α-carbon backbones of PCMH (thick solid), cytochrome c_5 (thin solid), and cytochrome *c*-551 (dashed). The four regions of greatest structural differences, as discussed in the text, are indicated.

hydrogenase (Lim et al., 1986), an iron–sulfur flavoprotein, the benzene ring of the FMN cofactor is also the likely pathway for reduction of the endogenous iron–sulfur cluster since the 8α–methyl group is in contact with one of the cysteine ligands of the cluster.

The propionate groups of the heme remain partially exposed to solvent in the complex since that face of the cytochrome subunit is not made inaccessible by the interaction with the flavoprotein subunit. This area would be the most likely site for electron transfer to the azurin since here the heme is most accessible. This area might also serve as a good docking site for the azurin, which could interact with portions of the second domain of the flavoprotein subunit which is adjacent.

Comparisons with Other Proteins. (1) Flavoprotein Subunit. There are five distinct structural classes of flavoproteins which have been observed. Two of these, the flavodoxins (Ludwig et al., 1982) and the $\beta_8\alpha_8$ barrel enzymes such as flavocytochrome b_2 (Xia et al., 1987), glycolate oxidase (Lindqvist, 1989), and trimethylamine dehydrogenase (Lim et al., 1986), bind FMN. The other three classes are FAD binding enzymes and are represented by glutathione reductase (Karplus & Schulz, 1987), by acyl-CoA dehydrogenase (Kim & Wu, 1988), and by PCMH, respectively. The glutathione reductase class binds the flavin in a characteristic domain similar to a Rossmann fold (Rossmann et al., 1975) containing a parallel β-sheet with α-helices on one side and an antiparallel three-stranded β-sheet on the other. In the acyl-CoA dehydrogenase the FAD is bound to a highly helical domain with the flavin ring close to a strictly β-domain. Thus, the conformation of the flavoprotein subunit of PCMH is quite different from that of other flavoproteins of known structure.

In all flavoproteins of known structure except acyl-CoA dehydrogenase and PCMH the phosphate or pyrophosphate group of the flavin is located at the N-terminal end of an α-helix. The positive dipole of the helix (Hol et al., 1978) is thought to stabilize the binding of the negatively charged phosphate groups. This binding feature has been found for most other dinucleotide binding proteins (Wierenga et al., 1985), whether or not positively charged side chains are present nearby. In PCMH the FAD is bound covalently, which might be thought to compensate for the lack of stabilization by the missing helix dipole. However, in acyl-CoA dehydrogenase, which also lacks this helix dipole feature (J. Kim, personal communication), the flavin is bound noncovalently. The energetic need for covalent attachment of the flavin in PCMH is also questionable. For example, the flavin binding sites in both flavocytochrome b_2 and trimethylamine dehydrogenase are nearly identical, in spite of the fact that the flavin is noncovalently attached in the former but covalently attached in the latter.

(2) Cytochrome Subunit. The cytochrome subunit of PCMH clearly belongs to the small bacterial subclass of cytochromes (Dickerson, 1980). These are acidic proteins with p*I* below 5.0 and contain about 80 or 90 amino acids. The main differences between this class and the eukaryotic class of cytochrome *c* are a major deletion in the former at the "bottom" of the molecule near the propionic acids and their acidic nature relative to the highly basic eukaryotic class. Other deletions with respect to the eukaryotic class occur as well.

There are two small bacterial cytochromes whose structures have been refined to 2.5-Å resolution or better, cytochrome c_5 from *Azotobacter vinlandii* (Carter et al., 1985) and cytochrome *c*-551 from *Pseudomonas aeruginosa* (Matsuura et al., 1982). Cytochrome c_5 contains 83 residues and has a p*I* of 4.4. Cytochrome *c*-551 contains 82 residues and has a p*I* of 4.7. The PCMH cytochrome subunit is the smallest at 78 residues and has a p*I* of 4.55–4.70.

The backbone α-carbon structures of cytochromes c_5 and *c*-551 have been manually aligned and compared with that of the PCMH cytochrome subunit. The comparison of the three backbones is shown in Figure 10, and an alignment of their amino acid sequences is shown schematically in Figure 11.

The major difference in structure is the large deletion in

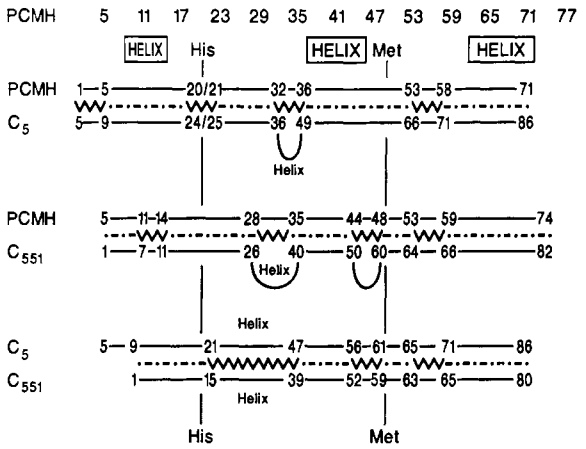

FIGURE 11: Pairwise comparison of the schematic sequences of the PCMH cytochrome subunit, cytochrome c_5, and cytochrome *c*-551, indicating regions of structural similarity (dot–dash line) and dissimilarity (jagged line). The approximate positions of the three helices common to the three proteins and the His and Met axial heme ligands are indicated. Insertions in cytochrome c_5 and cytochrome *c*-551 and additional helical segments with respect to PCMH are indicated.

PCMH in the center of the sequences (region I, Figure 10), between the two heme ligands His 19 and Met 50. In PCMH there are only 30 residues between these two ligands compared with 39 in c_5 and 43 in *c*-551. In this segment there is only one helix in PCMH compared with two helices in each of the other two cytochromes (Figure 11). Thus, the second of the four helices in cytochromes c_5 and *c*-551 is missing in PCMH. The major effect of this deletion is to leave the heme propionates even more exposed to solvent in PCMH. The remaining three helices are present in all three structures at corresponding locations and match quite well, except for helix 1 in *c*-551 which differs somewhat in position compared to the other two cytochromes (region II, Figure 10).

When cytochromes c_5 and *c*-551 are compared (Carter et al., 1985), there are three regions of difference between them (Figure 11). One is the position of helix 2, which is missing in PCMH, as described above. The second occurs just before the Met ligand (region III, Figure 10). In cytochrome *c*-551, this segment serves to cover one of the heme propionates. In this respect cytochrome c_5 matches PCMH closely. The third difference occurs four or five residues past the methionine and consists of a three to six residue segment which differs in all three proteins (region IV, Figure 10).

It is interesting to try to correlate the structural differences among the three cytochromes with their physiological functions. Of the five sites in PCMH which make closest contact with the flavoprotein subunit, the first two, at 17–18 and 23–25 (see Figure 4), are quite similar in structure in the three proteins. The next two sites, 44–45 and 48–49, are quite similar to cytochrome c_5 but differ from cytochrome *c*-551. The fifth site, 56–57, is different in all three proteins (region IV, Figure 10), as mentioned above. This suggests that the interaction sites of cytochrome c_5 and *c*-551 with their reductases occur at these corresponding positions and that the interaction of c_5 with its (unknown) redox partner is geometrically more similar to that of PCMH than is that of cytochrome *c*-551.

The most likely interaction site of PCMH with its naturally occurring reoxidant, thought to be a relatively basic azurin (p*I* 7.25; Causer, 1985), is in the region of the propionic acid groups as discussed earlier. This is probably the region where the basic horse heart cytochrome *c* and the cationic dyes phenazine methosulfate, phenazine ethosulfate, and Wurster blue interact to accept electrons from reduced PCMH (McIntire et al., 1987). This region shows the greatest difference between the PCMH cytochrome and the other two cytochromes (region I, Figure 10). In the latter two, this segment helps cover this edge of the heme but in PCMH it seems to leave the edge more exposed to solvent. Thus, the interactions with the reoxidants may differ considerably among the three proteins. Since cytochrome *c*-551 is monomeric, the oxidation and reduction might occur through the same face; this is not possible for PCMH in the natural complex. This difference may apply to cytochrome c_5 as well, although it may be bound to a membrane in vivo (Carter et al., 1985).

The distribution of charged groups among the three small bacterial cytochromes is also quite similar. In all three, the patch of negative charges at the "back" (defined in Figure 4) is maintained, as is the more even distribution of positive and negative charges on the "left" and "right" sides, the presence of basic groups at the "bottom" near the propionic acid groups, and the generally neutral character of the "front" of the cytochrome surrounding the exposed heme face. This supports the idea that the interaction of cytochrome c_5 and *c*-551 with their redox partners occurs at the "front" face and is consistent with the geometry of the cytochrome c_5 dimer interaction observed in the crystalline state (Carter et al., 1985). More detailed comparisons of conserved basic residues around the outer edge of the exposed face must await refinement of the PCMH structure.

Conclusions

The structure of PCMH reveals a number of details within an electron transfer complex and demonstrates that electron transfer can occur over an 8-Å distance in a naturally occurring complex. The heme and flavin cofactors need not be coplanar since the dihedral angle observed between them is 65°. Electron transfer thorough the benzenoid ring of the flavin seems to be a preferred pathway, since the ring is closest to the electron acceptor group in this and other flavin systems (Lim et al., 1986; Simondson et al., 1982; Stewart et al., 1988).

The structure also demonstrates that the "front" face of the cytochrome subunit is the one region which interacts with electron donor proteins, as predicted from modeling and solution studies [see Mathews (1985) for review]. The similarity of the PCMH cytochrome subunit to cytochrome c_5 and *c*-551 shows that they are closely related, especially with respect to interaction with electron donors. The biggest structural difference is the presence of only three helices in PCMH compared with four in the other two cytochromes.

The analysis of the structure of PCMH in its present state provides little detailed information on the enzymatic activity or on the chemical groups involved in catalysis. The orientation of the flavin ring, with its *re* face close to a backbone chain, suggests that the dome interior of domain 2 contains the substrate binding site and that the *si* face is involved in catalysis. Furthermore, the side chains which participate in catalysis are probably also contained in the dome domain.

The binding site for FAD in the flavoprotein subunit is different from any previously observed binding site. The four β-strand plus helix motif associated with most FAD and other dinucleotide binding proteins (Birktoft & Banaszak, 1984; Wierenga et al., 1985) is absent in PCMH. Finally, a unique structural motif consisting of eight antiparallel strands and four antiparallel crossover helices has been found in the isoalloxazine binding domain of PCMH.

Acknowledgments

We thank J. Richardson and C.-I. Branden for helpful discussions.

Registry No. PCMH, 66772-07-4; FAD, 146-14-5; heme, 14875-96-8.

REFERENCES

Adman, E. T., Sieker, L. C., & Jensen, L. H. (1973) *J. Biol. Chem. 248*, 3987–3996.

Almassy, R. J., & Dickerson, R. E. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 2674–2678.

Ambler, R. P. (1982) in *From Cyclotrons to Cytochromes* (Robinson, A. R., & Kaplan, N. O., Eds.) pp 263–280, Academic Press, New York.

Barry, C. D., Bosshard, H. E., Ellis, R. A., & Marshall, G. R. (1974) *Fed. Proc. 33*, 2368–2372.

Bellamy, H. D., Mathews, F. S., McIntire, W. S., & Singer, T. P. (1987) in *Flavins and Flavoproteins* (McCormick, D. B., & Edmondson, D. E., Eds.) pp 673–676, Walter de Gruyter, Berlin.

Birktoft, J. J., & Banaszak, L. J. (1984) *Protein Pept. Rev. 4*, 1–46.

Bossert, I. D., Whited, G., Gibson, D. T., & Young, L. Y. (1989) *J. Bacteriol. 171*, 2956–2962.

Bricogne, G. (1976) *Acta Crystallogr. A32*, 832–847.

Carter, D. C., Melis, K. A., O'Donnell, S. E., Burgess, B. A., Furey, W. F., Jr., Wang, B. C., & Stout, C. D. (1985) *J. Mol. Biol. 184*, 279–295.

Causer, M. J. (1985) Ph.D. Thesis, University College of Wales, Aberystwyth.

Dickerson, R. E. (1980) *Sci. Am. 242*, 98–110.

Hermans, J., Jr., & McQueen, J. E. (1974) *Acta Crystallogr. A30*, 730–739.

Hol, W. G. J., van Duijnen, P. T., & Berendsen, H. J. C. (1978) *Nature 273*, 443–446.

Hopper, D. J. (1978) *Biochem. J. 175*, 345–347.

Hopper, D. J. (1983) *FEBS Lett. 161*, 100–102.

Hopper, D. J., & Taylor, D. G. (1977) *Biochem. J. 167*, 155–162.

Jones, T. A. (1985) *Methods Enzymol. 115*, 157–171.

Karplus, P. A., & Schulz, G. E. (1987) *J. Mol. Biol. 195*, 701–729.

Keat, M. H., & Hopper, D. J. (1978) *Biochem. J. 175*, 649–658.

Kim, J.-J. P., & Wu, J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6677–6681.

Koerber, S. C., McIntire, W., Bohmont, C. W., & Singer, T. P. (1985) *Biochemistry 24*, 5276–5280.

Lederer, F., Glatigny, A., Bethge, P. H., Bellamy, H. D., & Mathews, F. S. (1981) *J. Mol. Biol. 148*, 427–448.

Lim, L. W., Shamala, N., Mathews, F. S., Steenkamp, D. J., Hamlin, R., & Xuong, Ng. (1986) *J. Biol. Chem. 261*, 15140–15146.

Lindqvist, Y. (1989) *J. Mol. Biol. 209*, 151–166.

Ludwig, M. L., Pattridge, K. A., Smith, W. W., Jensen, L. H., & Watenpaugh, K. (1982) in *Flavins and Flavoproteins* (Massey, V., & Williams, C. H., Jr., Eds.) pp 19–27, Elsevier/North-Holland, New York.

Mathews, F. S. (1985) *Prog. Biophys. Mol. Biol. 45*, 1–56.

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

Matthews, B. W., & Czerwinski, E. W. (1978) *Acta Crystallogr. A31*, 480–487.

McIntire, W. S., Edmondson, D. E., Hopper, D. J., & Singer, T. P. (1981) *Biochemistry 20*, 3068–3075.

McIntire, W. S., Koerber, S. C., Bohmont, C. W., & Singer, T. P. (1984) in *Flavins and Flavoproteins* (Bray, R. C., Engel, P. C., & Mayhew, S. G., Eds.) pp 643–654, DeGruyter, New York.

McIntire, W., Hopper, D. J., & Singer, T. P. (1985) *Biochem. J. 228*, 325–335.

McIntire, W., Singer, T. P., Smith, A. J., & Mathews, F. S. (1986) *Biochemistry 25*, 5975–5981.

McIntire, W. S., Hopper, D. J., & Singer, T. P. (1987) *Biochemistry 26*, 4107–4117.

Priestle, J. (1988) *J. Appl. Crystallogr. 21*, 572–576.

Reeves, C. D., Carver, M. A., & Hopper, D. J. (1989) *Biochem. J. 263*, 431–437.

Richardson, S. J. (1977) *Nature 268*, 495–500.

Rossmann, M. G., Liljas, A., Branden, C.-I., & Banaszak, L. J. (1975) *Enzymes (3rd Ed.) 11*, 61–102.

Schneider, G., Lindqvist, Y., Branded, C.-I., & Lorimer, G. (1986) *EMBO J. 5*, 3409–3415.

Shamala, N., Lim, L. W., Mathews, F. S., McIntire, W., Singer, T. P., & Hopper, D. J. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 4626–4630.

Sheriff, S., & Hendrickson, W. A. (1987) *Acta Crystallogr. A43*, 118–121.

Simondsen, R. P., Weber, P. C., Salemme, F. R., & Tollin, G. (1982) *Biochemistry 21*, 6366–6375.

Stewart, D. E., LeGall, J., Moura, I., Moura, J. J. G., Peck, H. D., Xavier, A. V., Weiner, P. K., & Wampler, J. E. (1988) *Biochemistry 27*, 2444–2450.

Terwilliger, T. C., Kim, S. H., & Eisenberg, D. (1987) *Acta Crystallogr. A43*, 1–5.

Wang, B. C. (1985) *Methods Enzymol. 115*, 90–112.

Wierenga, R. K., De Maeyer, M. C., & Hol, W. G. J. (1985) *Biochemistry 24*, 1346–1357.

Xia, Z., Shamala, N., Bethge, P. H., Lim, L. W., Bellamy, H. D., Xuong, N., Lederer, F., & Mathews, F. S. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 2629–2633.

Xuong, N. H., Nielsen, C., Hamlin, R., & Anderson, D. (1985) *J. Appl. Crystallogr. 18*, 342–350.