Structure of the Tetraheme Cytochrome from *Desulfovibrio desulfuricans* ATCC 27774: X-ray Diffraction and Electron Paramagnetic Resonance Studies†

José Morais,‡ P. Nuno Palma,§ Carlos Frazão,‡ Jorge Caldeira,§ Jean LeGall,|| Isabel Moura,§ José J. G. Moura,§ and Maria A. Carrondo*,‡,⊥

Instituto de Tecnologia Química e Biológica, Apt.127, 2780 Oeiras, Portugal, Departamento de Química and Centro de Química Fina e Biotecnologia, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2825 Monte de Caparica, Portugal, Department of Biochemistry and Molecular Biology, University of Georgia, Athens, Georgia 30602, and Instituto Superior Técnico, 1000 Lisboa, Portugal

Received April 17, 1995; Revised Manuscript Received July 6, 1995⊗

ABSTRACT: The three-dimensional X-ray structure of cytochrome c_3 from a sulfate reducing bacterium, *Desulfovibrio desulfuricans* ATCC 27774 (107 residues, 4 heme groups), has been determined by the method of molecular replacement [Frazão *et al.* (1994) *Acta Crystallogr. D50*, 233−236] and refined at 1.75 Å to an *R*-factor of 17.8%. When compared with the homologous proteins isolated from *Desulfovibrio gigas*, *Desulfovibrio vulgaris* Hildenborough, *Desulfovibrio vulgaris* Miyazaki F, and *Desulfomicrobium baculatus*, the general outlines of the structure are essentialy kept [heme−heme distances, heme−heme angles, His−His (axial heme ligands) dihedral angles, and the geometry of the conserved aromatic residues]. The three-dimensional structure of *D. desulfuricans* ATCC 27774 cytochrome c_3Dd was modeled on the basis of the crystal structures available and amino acid sequence comparisons within this homologous family of multiheme cytochromes [Palma *et al.* (1994) *Biochemistry 33*, 6394−6407]. This model is compared with the refined crystal structure now reported, in order to discuss the validity of structure prediction methods and critically evaluate the steps used to predict protein structures by homology modeling. The four heme midpoint redox potentials were determined by using deconvoluted electron paramagnetic resonance (EPR) redox titrations. Structural criteria (electrostatic potentials, heme ligand orientation, EPR *g* values, heme exposure, data from protein−protein interaction studies) are invoked to assign the redox potentials corresponding to each specific heme in the three-dimensional structure.

One of the most interesting proteins in the electron transport chains of sulfate-reducing bacteria is the tetraheme cytochrome c_3. It is found in all sulfate reducers belonging to *Desulfovibrio* species and other genera, such as *Desulfomicrobium* and *Desulfobulbus*. There are seven amino acid sequences and four tertiary structures known [Moura *et al.*, 1991; Palma *et al.*, 1994; Coutinho & Xavier, 1994; Czjzek *et al.* (1994) and references therein]. Although the tertiary structures indicate that the overall shape of these molecules has been conserved and that the four heme groups have a standard conformation, the amino acid sequences can exhibit as little as 20% homology.

The four hemes of cytochrome c_3 (covalently thioether-linked type *c* hemes, bishistidinyl axially ligated) have unusually low and different redox potentials. The remarkable stability of the four redox centres in a relatively small protein (*ca.* 13 000 Da) has permitted several detailed spectroscopic studies, comprising the application of multidimensional NMR,[1] EPR, and Mössbauer spectroscopies and electrochemistry (Coutinho *et al.*, 1992; Turner *et al.*, 1992; Guigliarelli *et al.*, 1990; Moura *et al.*, 1988; Ravi *et al.*, 1992; Moura *et al.*, 1991). The relative proximity and spatial arrangement of the hemes facilitate intramolecular electron transfer, and it was proven that they interact magnetically (Ravi *et al.*, 1992). Consequently, the midpoint redox potentials were shown to be dependent on the redox state of the protein and a model that takes into consideration redox interactions between the hemes has been explored. Also, the heme midpoint redox potentials, as well as some of their interacting potentials, are pH-dependent (Moura *et al.*, 1982; Santos *et al.*, 1983; Catarino *et al.*, 1991).

The complex electron transfer chain of sulfate-reducing bacteria has been discussed extensively but is still not completely established. Cytochrome c_3 has been postulated to be a coupling redox factor toward the hydrogen utilization system (hydrogenase). However, its physiological role is still under debate and the operating electron transfer chains involve multiple components. A large amount of information is now available concerning the structural and oxidation−reduction properties of a number of low molecular weight

† This work was supported by the European Union BIO2-CT942052 project, by JNICT Grant PBIC/C/BIO/1277/92, STRDA/C/CEN/538/2, and JNICT fellowship BD/2155/92-IF to J. M. and BD/1690/91-IF to P.N.P.

* To whom correspondence should be addressed.
‡ Instituto de Tecnologia Química e Biológica.
§ Universidade Nova de Lisboa.
|| University of Georgia.
⊥ Instituto Superior Técnico.

⊗ Abstract published in *Advance ACS Abstracts,* August 15, 1995.

[1] Abbreviations: Dmb, *Desulfomicrobium baculatus* Norway 4 [*Desulfovibrio desulfuricans* Norway 4, renamed *Desulfovibrio baculatus* Norway 4, was recently reclassified as *Desulfomicrobium baculatus* Norway 4 (Devereux, *et al.*, 1990)]; DvH, *Desulfovibrio vulgaris* Hildenborough; DvMF, *Desulfovibrio vulgaris* Miyazaki; Dg, *Desuifovibrio gigas* NCIB 9332; Dd, *Desulfovibrio desulfuricans* ATCC 27774; c_3, tetraheme cytochrome c_3; PEG, poly(ethylene glycol); esd, estimated standard deviation; rms, root mean square; SA, simulated annealing; MM, molecular mechanics; NMR, nuclear magnetic resonance; EPR, electron paramagnetic resonance. The numbering scheme adopted for hemes corresponds to the order of covalent heme binding to the polypeptide chain.

0006-2960/95/0434-12830$09.00/0

electron transfer proteins and enzymes involved in these systems. Many of these components are metalloproteins that have been crystallized and contain active sites with interesting spectroscopic properties. These prosthetic groups represent a fairly rigid, standard substructure, upon which the protein chains can only be folded in a very specific manner, thereby assuring a high degree of structural homology between similar proteins from different bacteria. The data available from several X-ray crystal structures of different members of these families of proteins, namely rubredoxins, ferredoxins and cytochromes c_3 [for references see Le Gall and Peck (1994)], support this observation. Several homology modeled structures have been described for these proteins and utilized to study the complexes between some of them (Stewart *et al.*, 1988, 1989; Palma *et al.*, 1994).

The refined crystallographic structure of the tetraheme cytochrome c_3Dd reported here brings more information that can be used to the detection of structural features responsible for the fine tuning of the properties of this electron carrier. Furthermore, the model previously proposed and derived from molecular modeling (Palma *et al.*, 1994) can now be compared with the crystal structure, allowing a discussion of the validity and limitations of homology modeling.

MATERIALS AND METHODS

Growth of Microorganisms and Protein Purification. Dd ATCC 27774 cells were grown in a medium containing nitrate rather than sulfate (as a terminal electron acceptor) in order to obtain a higher yield of biomass and c_3Dd was purified as described (Liu *et al.*, 1988).

Oxidation/Reduction Studies: EPR Measurements. Midpoint oxidation−reduction potentials of c_3Dd were determined by reductive mediated redox titrations followed by EPR. Potentiometric titrations were performed under anaerobic conditions in a redox cell slightly modified from the design of Dutton (1971), using 100 mM Tris/HCl (pH = 8.0) buffer and a protein concentration of 1 mM. Redox mediators were added at concentrations as described previously (Caldeira *et al.*, 1994). Samples taken at poised redox potentials were anaerobically transferred to the EPR tubes and frozen in liquid nitrogen for subsequent quantification. Spectra of the 26 aliquots obtained were recorded at 10 ± 1 K on a Bruker ESP 300 spectrometer.

Molecular Modeling. For the crystallographic structure determination, a first alignment between the c_3Dd and c_3DvMF sequences was accomplished, which accounted for five insertions and five deletions, and a three-dimensional model was manually built on a graphics workstation, using TOM/FRODO (Cambillau & Horjales, 1987; Roussel *et al.*, 1990). After replacement of all the different amino acids in c_3DvMF sequence with the corresponding ones in c_3Dd, a model of eight segments of polypeptide chain plus four heme groups was obtained. The temperature factor was set to 20.0 Å^2 for all atoms in the structure, while all new added atoms were given zero occupancy (Frazão *et al.*, 1994).

In parallel, computer calculations on a three-dimensional structure of c_3Dd (Palma *et al.*, 1994) were performed on a Silicon Graphics-Indigo 2 Workstation, using molecular modeling software SYBYL (Tripos Associates) and AMBER molecular mechanics force field (Weiner & Kollman, 1981).

Crystallographic Data. Crystals of cytochrome c_3 from *Desulfovibrio desulfuricans* ATCC 27774 belong to space group $P6_122$, with cell parameters $a = b = 61.84(4)$ Å and $c = 109.7(2)$ Å, $Z = 12$.

Table 1: Summary of Intensity Data Measurements

instrument	Nonius FAST area detector
X-ray source	synchrotron (SRS Daresbury, U.K.)
X-ray wavelength	0.89 Å
collimator	300 μm
detector distance	120 mm
detector tilt	−17°
exposure time/frame	15 s/0.1°
fully measured reflections	19 328
unique reflections	11 179
resolution range	$1.75 \leq d \leq 20.0$ Å
completeness	84.8%
$I > 3\sigma(I)$	81.9%
R_{merge} $(I)^a$	5.5%

^{a} $R_{merge}(I) = \sum (|I - \langle I \rangle|)/\sum I$.

A second data set, in the resolution range 20.0−1.75 Å, was collected at station 9.6 of the SRS, Syncrotron Radiation Source (Daresbury, England), also on a Enraf-Nonius FAST area detector diffractometer, using synchrotron radiation (λ = 0.89 Å) (Frazão *et al.*, 1994). These data were used for the crystallographic structural refinement. A summary of the intensity data measurement process is presented in Table 1.

Structure Refinement. The c_3Dd structure was solved by the molecular replacement method using, as search model, the coordinates of cytochrome c_3 from *Desulfovibrio vulgaris* Miyazaki F, refined to 1.8 Å with R = 19% (Higuchi *et al.*, 1984), retrieved from the Protein Data Bank (Bernstein *et al.*, 1977; Abola *et al.*, 1987). All molecular replacement calculations were performed with programs from the CCP4 suite (CCP4, 1994) and have been previously described in detail (Frazão *et al.*, 1994).

Refinement was performed with X-PLOR (Brünger, 1992), using the topology and parameters files toph19x.pro and param19x.pro, derived from CHARMM (Brooks *et al.*, 1983). The molecular replacement solution was submitted to rigid-body refinement, as described in Frazão *et al.* (1994), followed by the X-PLOR simulated annealing (SA) refinement protocol, which led to a structure model with an R = 25.9% for data within 8.0−2.3 Å. Inspection of Fourier maps with TOM/FRODO (Cambillau & Horjales, 1987; Roussel *et al.*, 1990), computed omitting windows of every five residues, indicated several places of readjustment. The most important was an out of register section from Leu64 to Val67, which was corrected by moving this section one residue toward the terminal C with readjustment of neighboring residues, from Lys59 to Val68. These changes allowed for a reinterpretation of the three-dimensional structural alignment of c_3Dd and c_3DvMF, now implying three insertions and three deletions in the amino acid sequence.

At this stage, a protocol consisting of 50 cycles of positional atomic refinement followed by the refinement of the temperature factors (overall, grouped and individual isotropic B values) and by 50 more cycles of positional atomic refinement was used henceforth. Using reflections from the range 8.0−1.9 Å, the R-factor was lowered to 23.0%. Resulting Fourier maps improved clearly, allowing the correction of several side chains and the inclusion of 40 water molecules. These were included only when corresponding density peaks were visible in both the $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ Fourier maps and if their positions allowed for hydrogen bonds (within distances 2.5−3.6 Å). All atoms that could not be seen in the $2|F_o| - |F_c|$ map at 0.7σ contour nor in the $|F_o| - |F_c|$ map at 1.0σ contour were considered as dummy atoms (*i.e.*, zero occupancy factor) and refined like that.

Table 2: Refinement Statistics for the Structure of c_3Dd Cytochrome

resolution limits	8.0−1.75 Å
initial *R*-factor[a]	51.7%
final *R*-factor[a]	17.8%
no. of reflections used[b]	10 914
no. of atoms	978
no. of solvent molecules	97
Rms Deviations from Ideality for X-PLOR Stereochemical Parameter Classes	
bond lengths (Å)	0.012
bond angles	2.5°
torsion angles[c]	24.8°
improper angles	2.0°

[a] *R*-factor = $100 \times \Sigma |F_o - F_c|/\Sigma|F_o|$. [b] With $F_o \geq 2\sigma(F_o)$. [c] The torsion angles were not restrained in the course of refinement.

Refinement with data to 1.75 Å decreased the *R*-factor to 19.7%. The quality of the resulting maps was now fairly good, except for five residues, Lys71−Lys75, where the electron density had been sparse from the beginning. Examination of the X-PLOR SA refined omit maps clearly suggested the fit of this segment into electron density following two possible paths in two alternate conformations (A and B). All atoms in those five residues were given 50% occupancy, except for atoms N71 and C75, which are common to the two conformations. Those still out of density were given a dummy status. Additionally, a total of 94 solvent molecules were modeled as waters.

The model was further refined, and near the end of the refinement process, a SA calculation was again performed to check for any residual model bias. A last examination of the whole molecule was carried out where solvent molecules were kept only if their temperature factors was below 50.0 Å^2, if they were clearly visible in the $2|F_o| - |F_c|$ map at 1.0σ contour, and if they showed a reasonable hydrogen bonding pattern.

A final *R*-factor of 17.8% was achieved for a molecule with 978 protein atoms and 97 water molecules. Table 2 lists a summary of the refinement statistics for the c_3Dd structure.

RESULTS

EPR Measurements. Heme Spectral Deconvolution and Determination of Heme Midpoint Redox Potentials. In Figure 1, we collect EPR spectra of c_3Dd poised at different redox potentials. The spectrum of the native protein is similar to the ones previously published and described for tetraheme cytochromes (Moura *et al.*, 1988; Guigliarelli *et al.*, 1990): all four hemes are paramagnetic ($S = {}^1/_2$). The redox titration reflects the variation of a complex EPR spectra composed of four components with distinct spectral features and different midpoint redox potentials. For example, in the early stages of the redox titration, the contribution of the EPR features associated with the most positive hemes decrease, well documented in Figure 1A−E. Subtraction of the spectra obtained during the redox titration at different redox potentials (*i.e.*, comparison of the fully oxidized and intermediate redox states) resolves each of the individual heme *g* values as shown in Figure 2 and enables us to obtain the intensity of the individual spectral lines (Moura *et al.*, 1988). Fifteen *g* values were determined and this was taken as an indication of the good quality of the obtained data set. The only exception is the g_{min} value of the heme with a g_{max} of 3.22, which could not be observed due to a high-field

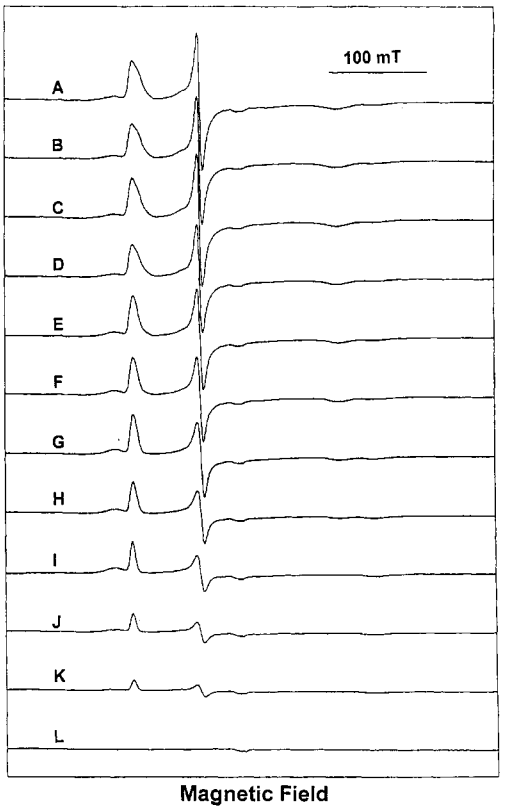

FIGURE 1: EPR redox titration of *D. desulfuricans* ATCC 27774 cytochrome c_3. Selected spectra from samples poised at different redox potentials: (A) + 93 mV, (B) −64 mV, (C) −86 mV, (D) −111 mV, (E) −148 mV, (F) −176 mV, (G) −191 mV, (H) −230 mV, (I) −305 mV, (J) −344 mV, (K) −381 mV, (L) −504 mV. EPR experimental conditions: microwave power 0.2 mW, temperature 10 ± 0.5 K, modulation amplitude 1.0 mT/point, sweep time 200 s and protein concentration 1 mM.

resonance (around 1.0 T) presumably too broad to be detected. This value could be estimated to be 0.63, according to the rule that indicates that the sum of the squares of the *g* values is 16 (Taylor, 1977).

The determination of the redox potential was done using the variation of the intensity of g_{max} and g_{min} features of each heme, since the g_{med} was found to be similar in all hemes. The midpoint redox potentials determined for the hemes of c_3Dd are indicated in Table 3 after the experimental data were fitted with $n = 1$ Nernst equations (see Figure 3).

The intensity of each heme component was estimated using areas rather than heights, since variations of the spectral line width occur during the redox titration. This phenomenon is demonstrated in Figure 2. A component at $g = 3.4$ is observed when the spectrum poised at −176 mV is substracted from the native one (+97 mV). This line width variation of the heme with $g = 3.22$ can be easily correlated with the reduction of heme with the second highest midpoint redox potential, indicating a magnetic interaction between these two hemes (later structurally assigned as hemes I and II). Spectral simulations were useful to determine the intensity of the contributing species at each poised redox potential and enabled us to estimate the mentioned line width variation.

Crystallography. The Final Crystallographic Model and Accuracy. Figure 4 presents a stereoscopic view of the c_3-Dd molecule, with the four heme groups in bold. The final model includes 107 amino acid residues, 4 heme groups, and 97 ordered solvent molecules, with one loop (from Lys71

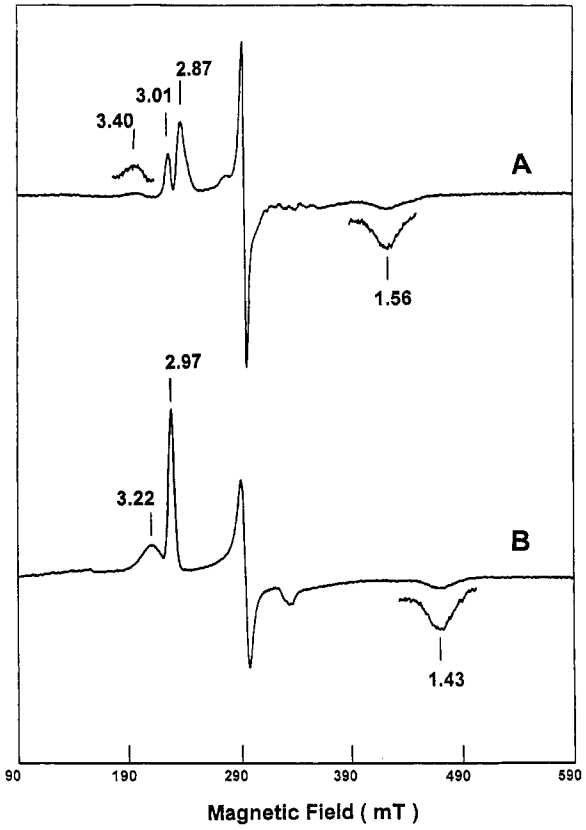

FIGURE 2: EPR spectra difference that demonstrate the individual heme spectral components: (A) Difference EPR spectra of *D. desulfuricans* ATCC 27774 cytochrome c_3 at +93 mV minus the one taken at −176 mV, where the two hemes of highest potential can be observed. Feature at $g = 3.4$ discussed in the text. (B) Spectra from the sample poised at −305 mV, where only the two hemes with more negative redox potential are observable and resolved.

Table 3: Midpoint Redox Potentials[a] of Each Heme in *D. desulfuricans* ATCC 27774 Cytochrome c_3 and the Corresponding g Values

EPR g values			midpoint redox potential (mV) vs NHE (±15 mV)
g_x	g_y	g_z	
2.871	2.287	1.562	−140 (heme IV)
3.010	2.287	1.577	−260 (heme I)
3.221	2.290	(0.63)[b]	−370 (heme II)
2.965	2.287	1.430	−380 (heme III)

[a] Relative to the normal hydrogen electrode. [b] Calculated using the relation $\Sigma g_i^2 = 16$.

to Lys75) defined in two alternate conformations with 50% occupancy each. Some of the atoms were never located in either the $2|F_o| - |F_c|$ or the $|F_o| - |F_c|$ Fourier maps and therefore were kept as dummy atoms until the end of refinement (Table 4). The atomic coordinates of the final structural model have been already deposited in the Protein Data Bank, with accession entry 2CYR. The heme numbering scheme is the same adopted by Matias *et al.* (1993) in the hexagonal form of c_3DvH, corresponding to the order of covalent heme binding to the polypeptide chain.

The average error in the refined atomic coordinates, between 0.1 and 0.2 Å, is represented in the Luzatti plot (Luzatti, 1952) shown in Figure 5. The overall rms deviation from ideality for the final model is 0.012 Å for bond lengths and 2.5° for bond angles (see Table 2). A Ramachandran plot (Ramachandran & Sasisekharan, 1968), shown in Figure 6, indicates that all residues except Cys51 are within acceptable regions; this exception was also observed in

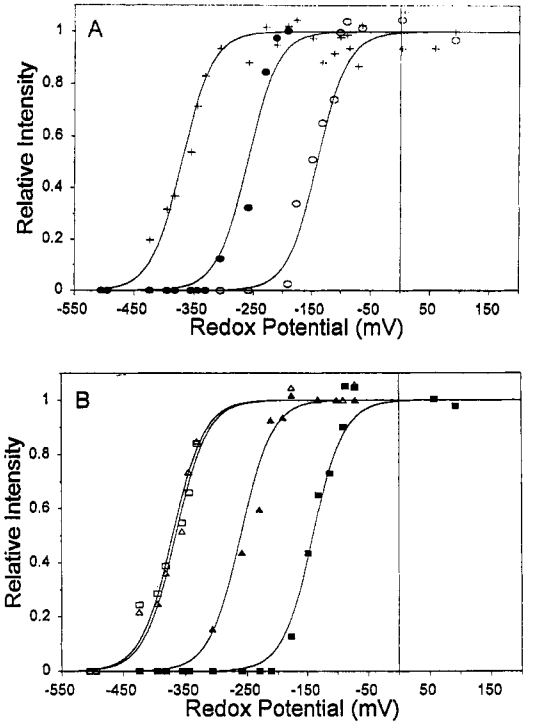

FIGURE 3: EPR intensity of each individual component measured from the experimental redox data shown in Figure 1 and manipulated as indicated in Figure 2. Nernst plot of the measured intensities at specific g values of the four hemes: heme 4, $g_{min} = 1.56$ (○) and g_{max} 2.87 (■); heme 3, $g_{min} = 1.58$ (●) and g_{max} 3.01 (▲); heme 2, g_{max} 3.22 (□); heme 1, $g_{min} = 1.43$ (+) and g_{max} 2.97 (△); the numbers are in the reduction order not acording to the a.a. sequence. Panel A indicates the Nernst plots based on the intensity of g_{min} features and panel B for g_{max} features.

cytochrome c_3 structures previously determined (Higuchi *et al.*, 1984; Morimoto *et al.*, 1991; Matias *et al.*, 1993; Czjzek *et al.*, 1994). Figure 7 presents a plot of the mean temperature factors for the main-chain and side-chain atoms of the polypeptide chain, with residues Lys71−Lys75 from alternate conformation A. The final B values for the solvent reange from 11.7 to 46.6, with an average value of 33.0 Å^2. The average B values for the heme groups are 11.6 Å^2 for heme I, 14.72 Å^2 for heme II, 11.29 Å^2 for heme III, and 10.74 Å^2 for heme IV.

The general quality of the refined model is illustrated in Figure 8, which depicts relevant sections of the molecule with the respective $2|F_o| - |F_c|$ electron density maps. In panels a and b, solvent molecules Wat1 and Wat3 hydrogen-bonded to His34 and His22, respectively, are represented. In panel c, a hydrogen bond between residues His22 and His25 is represented.

DISCUSSION

Molecular Structure Description. The secondary structure elements identified include a short right-handed 3_{10} helix (His22−Glu26), four segments of right-handed α-helix (Glu29−His34, Ser63−Ala70, Ser78−Lys90, and Leu93−Gly99), a short double-stranded antiparallel β-sheet (Val9−Gly13 and Lys16−Phe20), and six reverse turns of types I (His25−Val28, Lys45− Ser48, and Gly60−Ser63), I′ (Val37−Lys40), II (Ser48−Cys51), and III (Lys90− Leu93). The overall tertiary structure of the known cytochrome c_3 structures, hexagonal and orthorhombic DvH, DvMF, and Dmb, is essentially kept in the present molecule, with the four heme groups in very similar relative positions. In c_3-

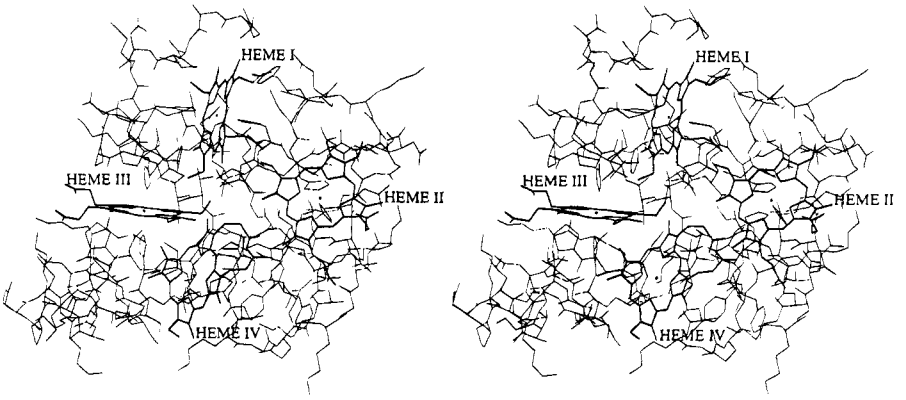

FIGURE 4: Stereo drawing of c_3Dd molecule, with the four heme groups in bold.

Table 4: Dummy Atoms in the Refined Structure of c_3Dd Cytochrome

residue no.	atoms
Gln 15	Oε1 and Nε2
Lys 40	Cγ through Nζ
Lys 59	Cδ through Nζ
Glu 61	Cγ through Oε2
Lys 71 A[a]	Nζ
Glu 73 A	Cβ through Oε2
Lys 75 A	Nζ
Lys 71 B[a]	Cε and Nζ
Glu 73 B	Cγ through Oε2
Lys 75 B	Cε and Nζ
Lys 94	Nζ
Lys 95	Nζ
Lys 102	Nζ

[a] A and B refer to the two alternate conformations presented by residues 71−75.

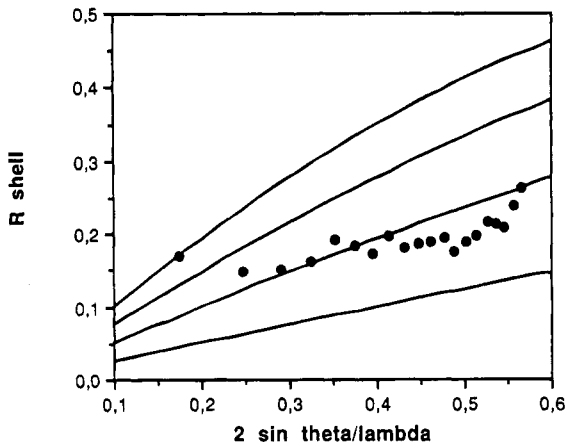

FIGURE 5: *R*-factor plot as a function of resolution for c_3Dd structure. The data up to 1.75 Å were divided in shells of equal volume in reciprocal space and the *R*-factor for each shell was plotted against the average resolution of the shell. The solid lines represent the predicted curves (Luzzati, 1952) for mean coordinate errors from 0.1 Å (bottom curve) to 0.4 Å (top curve).

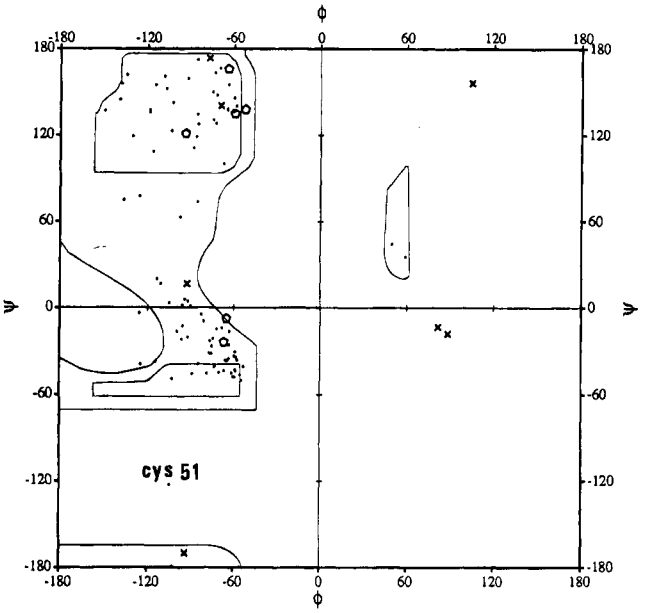

FIGURE 6: ϕ, ψ plot (Ramachandran & Sasisekharan, 1968) for c_3Dd molecule (residues 71−75 from alternate conformation A). Glycine residues are represented by crosses, proline residues by pentagons, other residues by dots. Drawing made with program PHIPSI (L. Tong, personal communication, 1988).

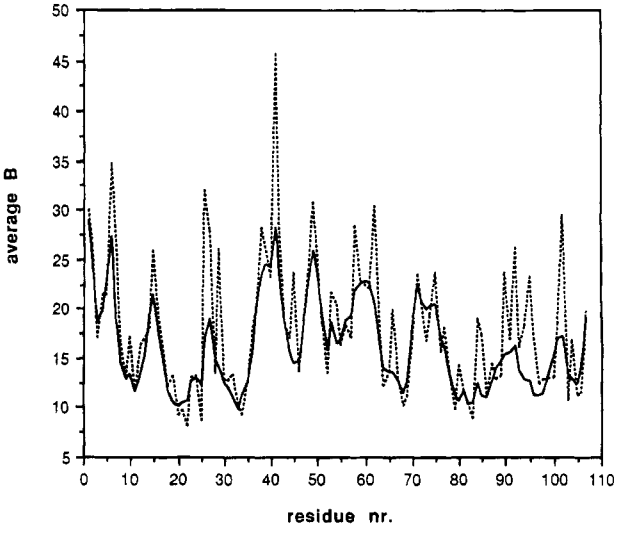

FIGURE 7: Plot of the main-chain (solid line) and side-chain (dotted line) average temperature factors per residue in c_3Dd molecule (residues 71−75 from alternate conformation A).

Dd molecule, heme I is covalently bound to Cys30 and Cys33, heme II to Cys46 and Cys51, heme III to Cys79 and Cys82, and heme IV to Cys100 and Cys105. His22 and His34, His35 and His52, His25 and His83, and His69 and His106 coordinate hemes I, II, III, and IV, respectively.

The intramolecular iron−iron and edge to edge heme distances and interplanar angles are listed in Table 5 and as noticed before (Matias *et al.*, 1993), they are distributed in two groups. Shorter distances occur between hemes I and II, I and III, and III and IV, with interplanar angles close to orthogonality, ranging from 85° to 87°, while longer distances occur between those hemes (I−IV, II−III, and II−IV) that form more acute interplanar angles, from 20.6 to 69.8°; a similar situation is observed in the relative position

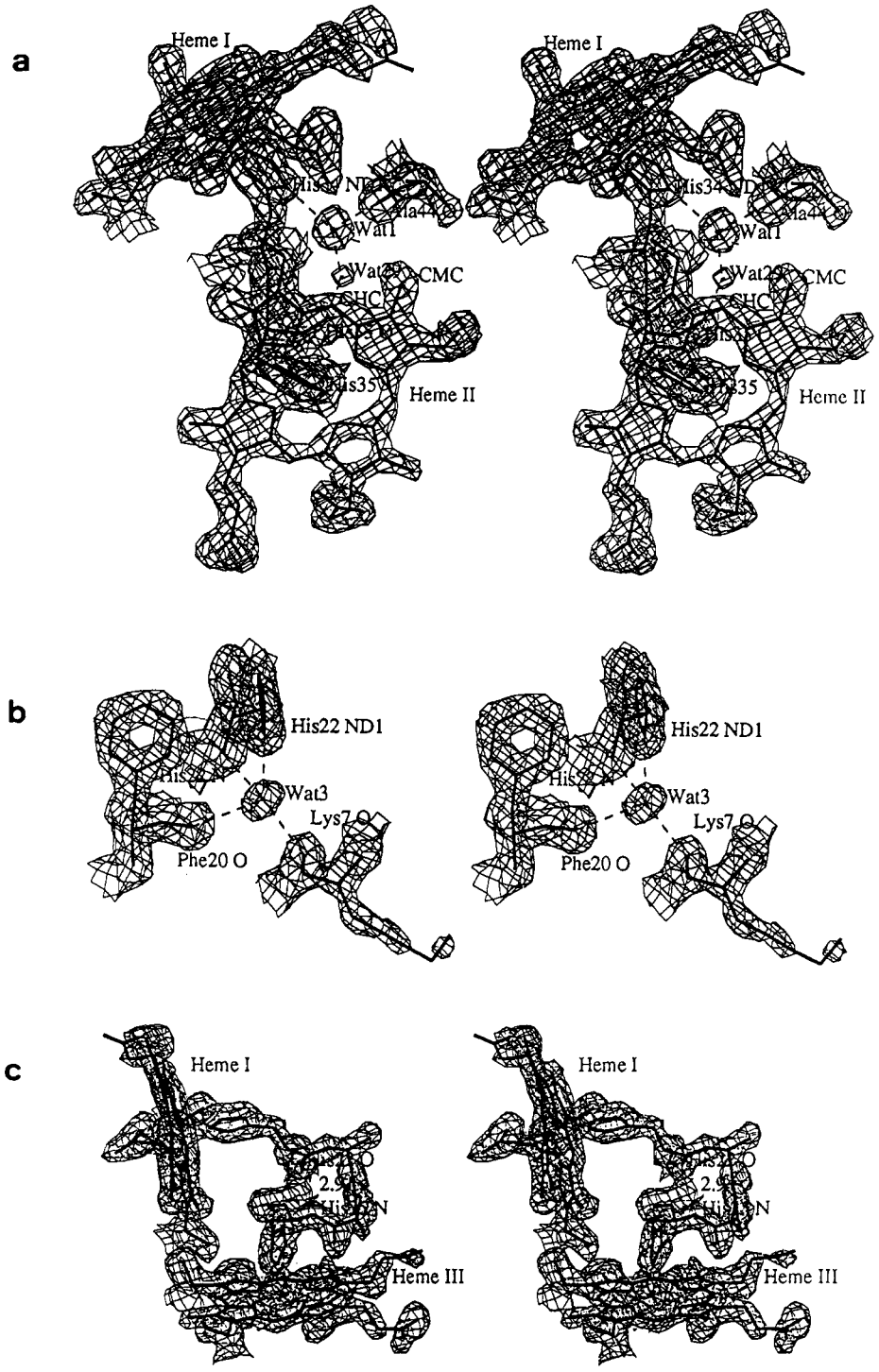

FIGURE 8: Stereo view of three sections of the final $2|F_o| - |F_c|$ electron density map with the refined model. Drawings made with TURBO (Roussel *et al.*, 1990). Maps drawn at 1.5σ contour. (a) Hydrogen bonding pattern of solvent molecule Wat1, with the corresponding residues involved. Hemes I and II are also represented, as well as a hydrogen bond between water molecule Wat29 and His35. Wat1 and Ala44 are within van der Waals distances to heme II (Wat1 O−CHC heme II, 3.70 Å; Ala44 O−CMC heme II, 3.60 Å). (b) Hydrogen-bonding pattern of solvent molecule Wat3. (c) Hydrogen bond between residues His22 and His25 coordinated to hemes I and III, respectively.

of the hemes in c_3Dmb molecule (Czjzek *et al.*, 1994). The position of the histidines that complete the iron coordination is characterized by the selected values presented in Table 6. This position is defined relative to the heme groups through the angle formed by the plane of the four heme nitrogen atoms with the Nϵ2−Fe vector (column 3 of Table 6) and the angle of tilt of the histidine rings relative to the same heme plane (column 4 of Table 6). These values are similar in the two structures, c_3Dd and c_3DvMF, and indicate no significant deviation from orthogonality between histidines and hemes. The fifth column in Table 6 lists, for each heme, the angle between the two histidine rings on both sides of the heme. Heme II shows a different environment from the other three hemes, with the two axial ligands making an angle 30° away from orthogonality. This observation is consistent with a similar situation for the same heme in c_3DvMF and in both molecules of c_3DvH (Matias *et al.*, 1993) and also in c_3Dmb, although in this case the angle is 40° away from orthogonality (Czjzek *et al.*, 1994).

c_3DvMF and c_3DvH structures present four aromatic residues close and approximately parallel to the imidazole ring of heme-coordinating histidines (Tyr43 to His34, Phe76

Table 5: Intramolecular Iron−Iron Distances and Heme−Heme Angles

heme[a]		angles (deg) [distance (Å)][b] I	II	III	IV
I	c_3 Dd		12.2 [6.4]	11.0 [5.8]	17.6 [9.6]
	c_3 DvMF		12.2 [6.5]	11.0 [5.8]	17.8 [9.9]
	c_3 DvH(A)[c]		12.4 [6.4]	11.1 [5.8]	17.7 [9.7]
II	c_3 Dd	87.0		16.1 [8.5]	16.6 [10.6]
	c_3 DvMF	89.7		15.8 [8.5]	16.4 [10.2]
	c_3 DvH(A)	88.3		16.0 [8.6]	16.6 [10.2]
III	c_3 Dd	85.4	59.7		12.2 [5.6]
	c_3 DvMF	81.6	59.7		12.0 [5.5]
	c_3 DvH(A)	81.0	59.8		11.9 [5.4]
IV	c_3 Dd	20.6	69.8	85.0	
	c_3 DvMF	22.0	71.0	82.0	
	c_3 DvH(A)	19.7	72.0	80.7	

[a] Heme planes are calculated for all atoms in the porphyrin ring. [b] Distances, in brackets refer to the nearest distance between porphyrin atoms. [c] Molecule A in the hexagonal form of c_3DvH (Matias *et al.*, 1993).

Table 6: Selected Angles for the Coordination Geometry of Ligated Histidine Residues

		heme pyrrole N plane[a] and vector Nε2−Fe	His plane[b]	His plane normal and His plane normal
		Heme I		
His22	c_3 Dd	89.0	87.0	9.8
	c_3 DvMF	88.6	86.5	5.0
His34	c_3 Dd	88.6	88.5	
	c_3 DvMF	88.3	86.4	
		Heme II		
His35	c_3 Dd	87.9	86.4	60.2
	c_3 DvMF	86.9	83.6	63.7
His52	c_3 Dd	86.6	89.3	
	c_3 DvMF	87.0	89.4	
		Heme III		
His25	c_3 Dd	88.0	82.8	5.3
	c_3 DvMF	82.9	78.2	16.9
His83	c_3 Dd	88.5	87.4	
	c_3 DvMF	87.3	86.8	
		Heme IV		
His69[c]	c_3 Dd	88.8	88.1	1.4
	c_3 DvMF	87.5	89.3	6.3
His106	c_3 Dd	88.6	88.0	
	c_3 DvMF	86.7	88.5	

[a] Least-squares plane of the four heme nitrogen atoms. [b] Least-squares plane of histidine imidazole. [c] His70 in c_3DvMF.

to His35, Phe20 to His25, and Tyr66 to His70), as pointed out by Higuchi *et al.* (1984) and Matias *et al.* (1993), while in c_3Dmb structure only one of those arrangements was found (Phe34 is approximately parallel to His39) by Czjzek *et al.* (1994). In the present molecule, two aromatic residues are approximately parallel to heme axial ligands, namely Phe20 (to His25) and Tyr65 (to His69). In the first case, the position of the aromatic residue seems to be strictly conserved in the four structures (Phe20 is also conserved in all known cytochrome c_3 sequences), while in the second case, Tyr65 is the aromatic residue structurally conserved corresponding to Tyr66 in c_3DvMF and c_3DvH. As for the remaining two aromatic residues, Tyr43 is also present in the c_3Dd molecule in the environment of heme I, although the aromatic ring is not parallel to the imidazole ring of His34, while Phe76 has been replaced by His76 in the c_3Dd sequence and structure.

Thus, in c_3Dd, two aromatic residues are located in the vicinity of heme I: Phe20 is parallel to the heme plane and close to the iron-coordinated His22 (Phe20−CHD heme I, 3.41 Å; Phe20−Nδ1 His22, 3.61 Å), and Tyr43 is close to His34, the other heme axial ligand (Tyr43−CHB heme I, 3.82 Å; Tyr43−Cγ His34, 3.63 Å). Phe20 is also close to heme III and approximately parallel to one of its axial ligands, His25 (Phe20−C2B heme III, 3.59 Å; Phe20−Cε1 His25, 3.44 Å). In the heme II environment, the imidazole ring of His76 is close to the heme and parallel to the ring of axial ligand His35 (His76−C1A heme II, 3.55 Å; His76−Cγ His35, 3.66 Å), and in the environment of heme IV, Tyr65 is approximately parallel to, and forms one hydrogen bond with, the imidazole ring of axial His69 (Tyr65−C4A heme IV, 3.51 Å; O Tyr65−Nδ1 His69, 2.77 Å).

Hydrogen Bonds. Although the polypeptide chain of c_3-Dd is only 44% homologous to that of c_3DvMF and 48% homologous to that of c_3DvH, the intramolecular hydrogen bonding patterns are similar, contributing to the conservation of the secondary structure elements, with a few exceptions characteristic of each structure.

The vicinity of the hemes, where water molecules are bound through hydrogen bonds, is also very much conserved in the three molecules. A list of the hydrogen bonds involving the histidines coordinated to the iron atoms and the additional bonds linking these water molecules to other atoms in nearby residues is presented in Table 7. It is quite interesting to note that, in all three structures, three water molecules were cocrystallized in very similar positions, forming strong hydrogen bonds with Nδ1 from His22 and His34 around heme I and with Nδ1 from His106 coordinated to heme IV. Furthermore, these water molecules donate their hydrogens to form other hydrogen bonds, giving rise to similar geometries in the three structures. Panels a and b of Figure 8 illustrate, in the case of c_3Dd structure, the hydrogen bonding pattern of water molecules Wat1 and Wat3. Although most of the heme propionate groups are exposed to solvent and, therefore, solvated by water molecules, some of them are oriented toward atoms of the polypeptide chain, establishing hydrogen bonds with them, as in the case of heme II O1A and O1D (2.89 Å to N Lys75 and 3.05 Å to Nε2 His76, respectively), heme III O2D (2.83 Å to Nζ Lys90, a salt bridge) and heme IV O1A (2.84 Å to N Gln15, 2.89 Å to N Lys16, 3.11 Å to N Ser14) and O2A (2.91 Å to N Ser14, 2.96 Å to Oγ Ser14). Three water molecules hydrogen-bonded to heme propionate groups also establish hydrogen bonds to main-chain atoms. Additionally, there are three other significant hydrogen bonds which bring heme I in contact with two other hemes. Two cysteines covalently bound to hemes I and II, respectively, are hydrogen-bonded to an axial histidine of heme II and to a propionate group of heme I (O Cys33−N His35, 2.92 Å; N Cys46−O2D, 3.03 Å), while two main-chain atoms of axial histidines coordinated to hemes I and III, respectively, are also linked by a hydrogen bond (O His22−N His25, 2.91 Å) (see Figure 8c).

All these links formed between heme groups and surrounding amino acid residues contribute to keep the hemes well fixed in their relative positions and also may provide possible pathways for heme−heme and heme−protein electron transfer.

Crystal Packing. Crystals of c_3Dd contain 12 symmetry-related molecules in the unit cell. As Matias *et al.* (1993) reported for the c_3DvH hexagonal structure, c_3Dd crystal structure also presents a large solvent channel along the hexagonal screw axis around which the molecules pack in a helical way, while the packing around the other axes is more compact. Symmetry-related molecules 5 [$y, y - x, z + 5/6$],

Table 7: Hydrogen Bonds Involving Ligated Histidine Residues

(A) Distances for Nδ1 Atoms (Å)

c_3 Dd			c_3 DvMF[a]		c_3 DvH(A)[a,b]		c_3 Dmb[a]		
				Heme I					
His22	O (Wat3)	2.8	O (Wat29)	2.9	O (Wat19)	2.8	O (Wat126)	2.7	His36
His34	O (Wat1)	2.7	O (Wat8)	3.0	O (Wat21)	2.6	O (Wat129)	2.8	His48
				Heme II					
His35	O (Leu36)	2.8	O (Pro36)	2.9	O (Pro36)	3.1	O (Wat349)	3.0	His49
His52	O (Glu61)	2.7	O (Ala62)	2.5	O (Ala62)	2.5	O (Ile81)	2.7	His67
				Heme III					
His25	O (Pro21)	2.7	O (Asn21)	2.7	O (Asn21)	2.7	O (Pro35)	2.9	His39
His83	O (Leu97)	2.8	O (Leu97)	2.9	O (Leu97)	2.7	O (Pro108)	2.8	His96
				Heme IV					
His69[c]	O (Tyr65)	2.8	O (Tyr66)	2.9	O (Tyr66)	2.8	O (Wat266)	2.9	His89
His106	O (Wat39)	2.9	O (Wat30)	2.9	O (Wat2)	2.9	O (Lys30)	2.8	His115

(B) Further Hydrogen-Bond Distances for the Water Molecules Listed Above (Å)

c_3 Dd			c_3 DvMF[a]			c_3 DvH(A)[a,b]			c_3 Dmb[a]		
O (Wat3)	O (Lys7)	2.7	O (Wat29)	O (Ala6)	2.6	O (Wat19)	O (Ala6)	2.7	O (Wat126)	O (Gly5)	2.8
	O (Phe20)	3.0		O (Phe20)	2.9		O (Phe20)	2.9		O (Phe34)	2.8
O (Wat1)	O (Ala44)	2.8	O (Wat8)	O (Asn42)	2.8	O (Wat21)	O (Asp42)	3.0	O (Wat129)	O (Lys59)	2.7
	O (Wat29)	2.7		O (Gln44)	2.6		O (Arg44)	2.7		O (Wat142)	2.7
O (Wat39)	O (Thr17)	3.6	O (Wat30)	O (Pro17)	2.7	O (Wat2)	O (Pro17)	3.2	O (Wat266)	Oε2 (Glu85)	3.0
	O (Wat28)	3.5		Oε1 (Gln16)	3.5		Oε1 (Gln16)	3.2		O (Wat265)	3.3

[a] Residue numbering and distances for c_3DvMF, c_3DvH(A), and c_3Dmb structures were taken from PDB entries 2CDV, 1CTH, and 2CY3, respectively. [b] Molecule A in the hexagonal form of c_3DvH (Matias *et al.*, 1993). [c] His70 in c_3DvMF and c_3DvH.

6 [$x - y$, x, $z + 1/6$], 8 [$-x$, $y - x$, $-z + 2/3$] and 10 [$-y$, $-x$, $-z + 5/6$] are the closest ones to molecule 1 [x, y, z]. There are five intermolecular hydrogen bonds between the polypeptide chain of molecules 5, 6 and 10 and molecule 1, three of which are main-chain−side-chain contacts (Oε1 Glu10−N Glu29 [6], 3.13 Å; O Thr17−Nζ Lys85 [6], 2.91 Å; O Ala88−Nζ Lys104 [5], 3.02 Å) and two are side-chain−side-chain contacts (Nδ2 Asn6−Oε1 Glu26 [10], 3.18 Å; Nζ Lys12−Oγ1 Thr32 [6], 2.81 Å); two salt bridges are also formed between molecules 5 and 6 and molecule 1 (Nζ Lys12−Oε1 Glu29 [6], 3.00 Å, and Oε2 Glu29−Nζ Lys45 [5], 3.54 Å). Molecule 8 is hydrogen-bonded to molecule 1 via water molecules only. Two of these water molecules, each one hydrogen-bonded to a heme in a different molecule, make an intermolecular hydrogen bridge connecting, undirectly, heme II of molecule 1 to heme IV of molecule 8: O2D heme I−water50−water84 [8]−O Cys100 [8]−heme IV [8].

The shortest intermolecular edge-to-edge heme distances are 9.64 Å for heme I−heme I [10], 11.24 Å for heme II−heme IV [8], 14.42 Å for heme I−heme III [6], and 14.56 Å for heme I−heme I [9] [$x - y$, $-y$, $-z$]. Shortest intermolecular iron−iron distances occur between molecule 1 and molecules 8, 10, 6, and 5, in this order (Fe II−Fe IV [8], 17.01 Å; Fe I−Fe I [10], 17.32 Å; Fe I−Fe III [6], 20.17 Å; Fe III−Fe IV [5], 22.14 Å), and are shorter than those reported for other cytochrome c_3 structures.

There are no significant differences in the packing of the two alternate conformations of loop Lys71−Lys75.

Solvent Accessibility and Heme Exposure. The relative percentage of solvent accessibility was calculated for the hemes of three cytochrome c_3 molecules, c_3Dd, c_3DvMF, and c_3Dg (P. Matias, personal communication), with X-PLOR, using the algorithm by Lee and Richards (1971). The calculation was done as described in Matias *et al.* (1993), using a water probe of 1.4 Å radius and including the calculated hydrogen atoms (Table 8). While heme II is the most exposed in c_3DvMF, with a similar situation in the two molecules of the hexagonal c_3DvH (Matias *et al.*, 1993) and in c_3Dmb (Czjzek *et al.*, 1994), the values now reported for the other two cytochromes, c_3Dd and c_3Dg, have a different order, with heme III being the most exposed.

Table 8: Relative Percentage of Solvent Accessibility for the Heme Groups

heme	c_3 Dd	c_3 DvMF	c_3 Dg[a]
I	11.2	15.0	12.0
II	12.4	16.2	16.4
III	19.0	15.0	19.5
IV	12.2	14.6	16.6

[a] The recently refined coordinates file used in this calculation was kindly provided by P. Matias.

As previously reported for the c_3Dmb structure (Czjzek *et al.*, 1994), two of the covalently bound cysteines present higher solvent exposure than the others. In the c_3Dd molecule, they are Cys30 and Cys100, which are attached to hemes I and IV, respectively. Cys100, with a relative solvent accessibility of 20%, is the most exposed one, while Cys30 has only 8% of its surface exposed to solvent. Cysteines attaching hemes II and III are even less exposed.

As stated above, most of the heme propionate groups are accessible to solvent and, therefore, are able to form hydrogen bonds with water molecules; however, the exposure area is different from heme to heme and between propionate groups of the same heme. The propionate groups of heme III have the highest solvent accessibility, which is in accordance with the high value of exposure calculated for the same heme, while those of heme IV are completely buried inside the protein, making tight hydrogen bonds with the surrounding amino acid residues. The propionate groups of hemes I and II present similar areas of solvent exposure and are hydrogen-bonded both to water molecules and to protein atoms.

Evaluation of Homology Modeling for Protein Structure Prediction: Modeled versus Crystal Structure. In this section, we present a critical evaluation of our previously

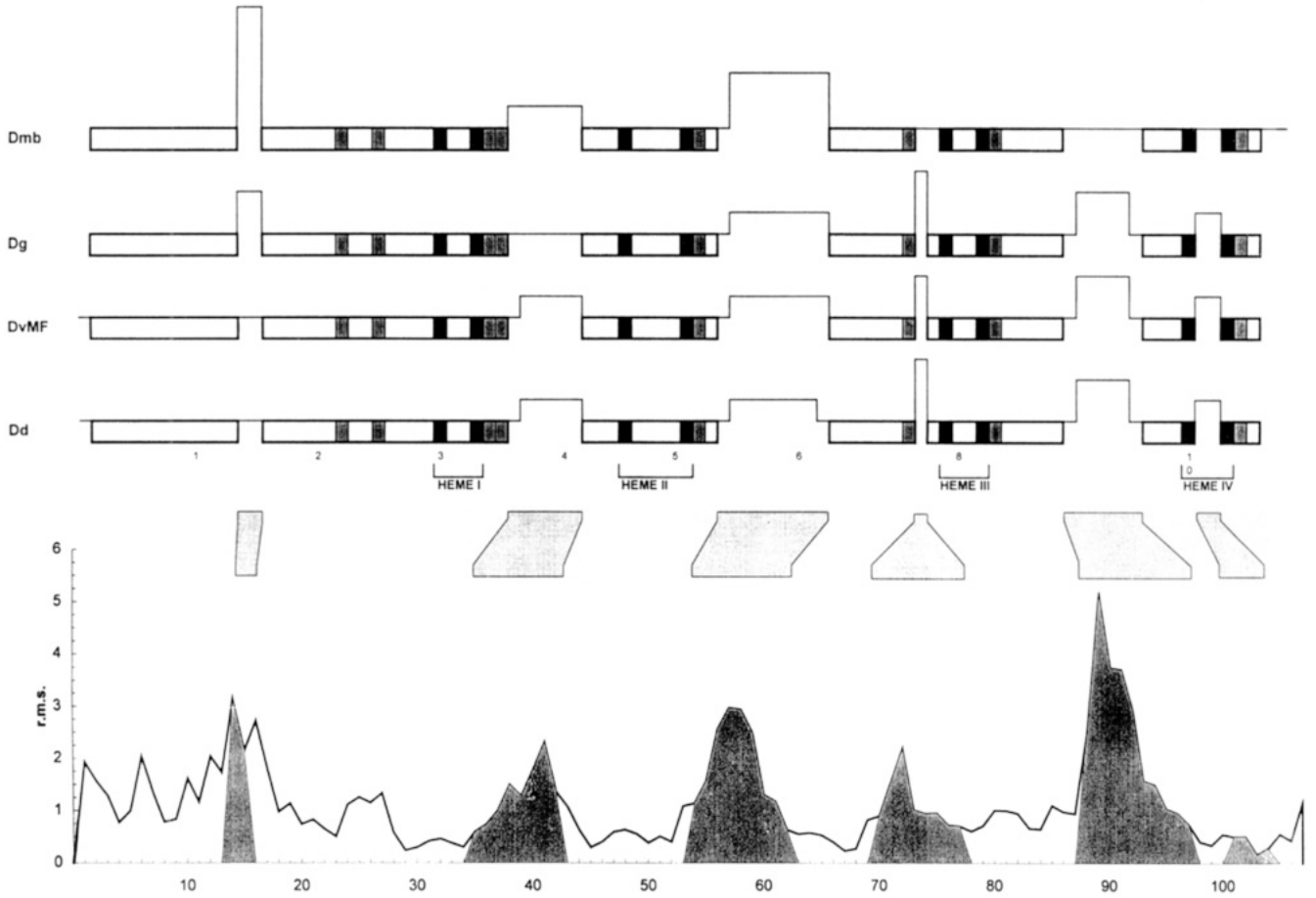

FIGURE 9: (Top) Proposed schematic alignment of the amino acid sequences of four *Desulfovibrio* type cytochromes c_3 (Palma *et al.*, 1994). This alignment was done on the basis of the structural (and not amino acid) homology, after best superposition of the atomic coordinates of the three known structures. The horizontal rectangles at the base of each sequence represent conserved structural domains, with the structurally variable regions shown as thin vertical loops. The amino acids are omited for clarity, except the cysteines (in black) and histidines (in gray) responsible for binding the hemes. (Bottom) Corresponding plot of rms versus amino acid number. The rms include the main-chain atoms, after best fitting of the two structures using the backbone atoms, and the shaded regions correspond to the variable loops shown in the sequence alignment.

predicted structural model (Palma *et al.*, 1994) as compared to the structure of c_3Dd now reported and described. The prediction was performed by a homology modeling methodology, which is based on the observation that functionally related proteins usually share the same class of tertiary structures, in particular, around the functional domains (Brown *et al.*, 1969; Sweet, 1986; Lesk *et al.*, 1986). Figure 9 (top) shows a proposed schematic alignment of the amino acid sequences of four cytochromes c_3 (Palma *et al.*, 1994). This alignment was done on the basis of the structural (and not amino acid) homology, after best superposition of the atomic coordinates of the three known structures. Despite the differences in size (107−118 amino acids) of these proteins, it is obvious, from the inspection of this figure, that they all possess a set of well-conserved structural domains (horizontal rectangles at the base of each sequence), with the variable regions located in outstanding loops of variable length (thin vertical loops). It should be noted that with this sort of analysis the global structural homology can be higher (*ca.* 60−90%) than the sequence homology (*ca.* 30−50%). Indeed, in some of the structurally well-conserved regions the amino acid homology can be very low. Once aligned, all the necessary mutations were introduced in the starting model and this mutated structure was then refined by energy minimization using the AMBER (Weiner & Kollman, 1981) molecular mechanics (MM) force field (Palma *et al.*, 1994).

Figure 10 compares the main-chain folding of the predicted structural model of c_3Dd, with the data obtained by X-ray diffraction. It is the best fit between the backbone atoms of both structures and represents the conserved regions (see Figure 9, top) as heavy lines, while the variable loops are shown as solid (X-ray) or dashed (homology model) thin lines. Despite the observed differences, the global folding of the two structures is clearly similar. It is, however, interesting to note that the larger differences are consistently located in what we called the variable loops. This can also be seen qualitatively in Figure 9 (bottom), where the rms differences between both structures are plotted along the amino acid sequence. In this graph, the shaded regions correspond to the variable loops shown in the sequence alignment and thin (solid and dashed) lines in Figure 10. The average rms difference for the main-chain atoms is 1.54 Å, being 1.20 Å for the regions that were considered conserved and 1.99 Å for the more variable segments. These values can be compared with those obtained from the structure determination of cytochrome c_3DvH refined to 1.9 Å, $R = 19.6\%$ using the same program, XPLOR (Matias *et al.*, 1993). The main-chain rms differences between the two independent c_3DvH molecules (excluding the five amino acids at both termini) found in the crystal range from 0.089 to 1.237 Å, with an average of 0.490 Å.

These results show one intrinsic limitation of the homology modeling methodology: the energy minimization methods,

FIGURE 10: Stereo plot comparing the main chain folding of the predicted structural model of *D. desulfuricans* ATCC 27774 cytochrome c_3, with its determined X-ray structure. It is the best fit between the backbone atoms of both structures and represents the conserved regions (see Figure 9, top) as heavy lines, while the variable loops are shown as solid (X-ray) or dashed (homology model) thin lines. The hemes of the X-ray structure are shown as white squares.

used to refine the structure after mutation, can improve local structural conformations, but they usually tend to keep the starting global structure, not changing it much, due to entrapment in local energy minima. So, its success is highly dependent on the choice and availability of an existing similar structure to start with and on the correct alignment of their sequences.

Besides, the variable loops, which are pointing outward, are probably more flexible than those segments that are conserved between the several cytochromes c_3, since these contain all 16 anchoring amino acids (Cys, His) that keep the four hemes in position and so are highly constrained.

While the global folding prediction can be considered reasonably achieved, the conformation of the amino acid side chains is less correct as compared with the X-ray structure (rms = 2.45 Å). This is especially true for the surface amino acids, with the side chains pointing outward and relatively free to rotate. This problem may arise, in part, by the fact that the energy minimization procedure was performed in vacuum, without considering the solvent effect.

Finally, it is important to note that our model also failed in predicting correctly the conformation in the vicinity of the heme groups (not shown). While the position and relative orientations of the porphyrin planes are correctly modeled, the same is not true for the bond distances between the iron and the imidazole nitrogens, as well as the relative orientations of the histidine planes. This observation brings up a second limitation of homology modeling for 3D structure prediction, especially when applied to metalloproteins. The use of molecular mechanics for structure refinement is strongly dependent on the force field parameters used to model interatomic interactions. While several sets of such parameters are available and largely tested for the polypeptide chain, in the commercially available MM programs, the lack of fine-tuned parameters for the organometallic moieties involving transition metals is still limiting the application of many computational methods to the study of these type of proteins. This matter requires intense investigation in the coming years.

It is common knowledge that the prediction of the complete tertiary structure of even a small protein, starting from its amino acid sequence, is still a very difficult task, far from being considered a standard procedure. It cannot yet substitute for an experimental structure determination, either by X-ray diffraction or multidimensional NMR. However, the usefulness of such predictive models must be evaluated in the light of their application. In our case, we believe that our structural model was useful for the purpose it was developed for, *i.e.*, the calculation of the electrostatic potential distribution around the surface of the protein (Palma *et al.*, 1994), since it is essentially dependent on the global folding and the relative position of the amino acids contributing to that potential.

Structural Assignments of Heme Redox Potentials. The modulation of the redox potentials is the result of the complex interplay of many different structural and electronic contributions (Moore *et al.*, 1986). In this section we shall present a tentative structural assignment of the measured redox potentials to the four hemes of c_3Dd, based on the available structural data, as well as spectroscopic and biochemical information. Although none of the criteria that will be listed below is sufficiently global *per se* to assign unequivocally all the redox potentials, we believe that the arguments presented here support our tentative assignment.

Heme II. From Table 3, it can bee seen that there is one heme for which a g_{max} with a significantly high value is determined, which also corresponds to the heme with the second most negative redox potential (−370 mV). This is a relevant spectroscopic result which can be correlated with measurable structural information, since the g_{max} value is known to be maximum when the His/His planes are perpendicular and minimum in the case of a parallel orientation. This correlation is well-established with respect to the theory of low-spin iron EPR spectra and the heme molecular orbitals (Palmer, 1985; Walker *et al.*, 1986). From the analysis of the outcoming structure, it can be seen that the angles between the pairs of axial histidine planes associated to the four hemes are not identical. In particular, for heme II, the value of this angle (60.2°) is very distinct from the other three (1.4°, 5.3°, and 9.8°), so in what concerns the EPR potentiometric data, this heme should be responsible for the second more negative redox potential.

The extent of heme exposure to solvent has also been considered as a factor contributing to the modulation of the redox potential (Wagen, 1978), with the percent exposure inversely related to the potential. However, this factor may

not be a determinant for the assignment of the redox potential to heme II, since its percent solvent accessibility is not distinguishable from those of hemes I and IV (see Table 8).

Heme IV. In all the reported studies on the interaction between cytochromes c_3 and different acidic redox partners, such as ferredoxin (Moura *et al.*, 1977; Xavier & Moura, 1978; Xavier *et al.*, 1979; Guerlesquin *et al.*, 1984, 1985, 1987; Capeillère-Blandin *et al.*, 1986; Cambillau *et al.*, 1988, Park *et al.*, 1991), rubredoxin (Moura *et al.*, 1980; Stewart *et al.*, 1989), flavodoxin (Moura *et al.*, 1980; Stewart *et al.*, 1988; Palma *et al.*, 1994), and even a macroinorganic anion (Mus-Veteau *et al.*, 1992), the same heme seems to be involved in the interaction process. The identification of this heme as heme IV has been made in several cases by cross-linking studies (Dolla & Bruschi, 1988; Dolla *et al.*, 1991) and extended NMR experiments (Guerlesquin *et al.*, 1985b; Park *et al.*, 1991; Mus-Veteau *et al.*, 1992). In addition, a set of experiments involving cycles of reduction and reoxidation of different cytochromes c_3 followed by NMR (Moura *et al.*, 1977, 1982; Turner *et al.*, 1994; Salgueiro *et al.*, 1992; Piçarra-Pereira *et al.*, 1993; Coletta *et al.*, 1991) demonstrated that the heme directly involved in the above protein−protein interaction (heme IV) is the one possessing the highest redox potential value.

It can be argued that the electrostatic potential created by the rest of the protein on the surroundings of each heme, may influence its redox potential, as well. A more positive electrostatic potential would help in stabilizing an additional electron in the heme, hence facilitating its reduction and increasing the redox potential. It is interesting to note that for all the tested cytochromes c_3, DvMF, Dg and Dd, the heme assigned to the more positive redox potential is also the one surrounded by a patch of highest positive electrostatic potential (Palma *et al.*, 1994). This heme is, for all three proteins, heme IV and in c_3Dd the electrostatic potential is also clearly more positive around heme IV. In the case of c_3Dmb, however, heme III has been assigned to the more positive redox potential (Coutinho *et al*, 1993), but in this structure, this is also the heme embedded in the more positive electrostatic potential.

Finally, this assignment is still reinforced by the fact that the heme of higher redox potential is the one associated with the lowest g_{max} value (2.87 in Table 3), which in turn may be related to heme IV, since it is the one with the lowest angle between the associated axial histidine planes (1.4°; see Table 6).

Hemes I and III. The assignment of redox potentials to hemes I and III according to the above arguments is more ambiguous. Inspection of the structure shows that the angles between the planes of the correspondent axial histidines are not significantly different, the same being true for the remaining two g_{max} values (2.97 and 3.01). Calculation of the electrostatic potential around these two hemes does not give conclusive results either, since they are also similar (Palma *et al.*, 1994). However, there is one property that clearly distinguishes the two hemes. The percent of solvent accessibility for heme III is much higher than for any other heme, which would make its redox potential, in the absence of other strong arguments in the contrary, the lowest one (−380 mV).

In view of the above discussion, we propose the following assignment of the measured redox potential values to all the four hemes of the structure: heme III (−380 mV), heme II (−370 mV), heme I (−260 mV), and heme IV (−140 mV).

Table 9: Intramolecular Electron Transfer in Tetraheme Cytochromes

c_3DvH	III → II → I → IV	Turner *et al.*, 1994 Salgueiro *et al.*, 1992
c_3DvH	II → I → III → IV	Matias *et al.*, 1993
c_3Dmb	II → I → IV → III	Coutinho *et al.*, 1993 Gayda *et al.*, 1988
c_3Dg	I → II → III → IV	Coletta *et al.*, 1991 Piçarra-Pereira *et al.*, 1993
c_3Dd	III → II → I → IV	this work

CONCLUSIONS

Table 9 compiles some information on structural−redox-based correlations and hypothetical pathways for different cytochromes c_3. NMR and EPR have provided a body of evidence for these assignments (see references indicated in the table). It is remarkable that, while the overall core architecture is maintained, inversions of the redox potential order are observed, indicating that facility for electron tranfer and the design of electron transfer pathways are controlled by a large number of parameters as we have previously discussed, such as heme exposure to solvent, effect of amino acids sourrounding the heme structure, charge distributions in the molecule, and heme ligand geometry.

In most cases, heme IV is the most positive center. The patch of positive charges at this site of the molecule is well-conserved betwen different cytochromes c_3, independently of the isoelectric point, thus favoring the interaction with flavodoxin, rubredoxin, and ferredoxin. Also, when interacting with large molecules such as hydrogenase, heme IV is directly affected by the binding process, indicating again that the positive lysine patch is involved in the molecular recognition step (J. J. G. Moura, unpublished results).

A remarkable and interesting feature is the conservation of a crystallographic water molecule within different cytochromes c_3 (Table 7). This solvent molecule is situated and held in an also conserved hydrogen-bond pattern, between hemes I and II, allowing for a possible pathway for electron transfer between the two hemes. This proposal is supported by the experimental observation that the oxidoreduction processes of these two hemes seems to be coupled. In fact, as mentioned above, a change in line width is observed for the g_{max} EPR signal of heme II during the reduction of heme I. These two observations are in accordance with the fact that, for all the assignments summarized in Table 9, hemes I and II are always located in adjacent positions in the ordering of the redox potentials. This further supports the side-by-side positioning of those two hemes in our proposed assignment.

Our results also indicated that, in a fairly rigid structural situation, where a polypeptide folds around large molecular mass prosthetic groups and where a high degree of sequence homology is observed, prediction of structural features by homology modeling may be a valuable help in situations where a crystal structure is not available. Continued investigations of these electron transfer proteins through the combination of efforts in molecular modeling and structure prediction, in spectroscopic and enzymatic analysis, and in crystallization and determination of X-ray structures will lead to a much better understanding of their role and the mechanism of electron transfer between flavoproteins, heme, and non-heme iron proteins.

ACKNOWLEDGMENT

We are grateful to Dr. P. Rizkallah and to Dr. M. Papiz for the support during the X-ray diffraction data collection and processing at station 9.6 of the SRS (Daresbury, England), to Professor J. Van Beeumen and Dr. H. Demol for the amino acid sequence of c_3Dd prior to publication and to Dr. Cristina Costa for the purified protein. We thank the UGA Fermentation Plant for the growth of the bacteria used in this work and M. Y. Liu for preparing bacterial extracts.

REFERENCES

Abola, E. E. Bernstein, F. C., Bryant, S. H., Koetzle, T. F., & Weng, J. (1987) Protein Data Bank, in *Crystallographic Databases—Information Content, Software Systems, Scientific Applications* (Allen, F. H., Bergerhoff, G., & Sievers, R., Eds.) pp 107−132, Data Commission of The International Union of Crystallography, Bonn, Cambridge, Chester.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983) *J. Comput. Chem. 4*, 187−217.

Brünger, A. T. (1992) *X-PLOR. A System for Crystallography and NMR*. Version 3.0, The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, New Have, CT.

Caldeira, J., Palma, P. N., Regalla, M., Lampreia, J., Calvete, J., Schäfer, W., LeGall, J., Moura, I., & Moura, J. J. G. (1994) *Eur. J. Biochem. 220*, 987−995.

Cambillau, C. & Horjales, E. (1987) *J. Mol. Graphics 5*, 175−177.

Cambillau, C., Frey, M., Mosse, J., Guerlesquin, F., & Bruschi, M. (1988) *Proteins: Struct., Funct., Genet. 4*, 63−70.

Capeillere-Blandin, C., Guerlesquin, F., & Bruschi, M. (1986) *Biochim. Biophys. Acta 848*, 279−293.

Catarino, T., Coletta, M., LeGall, J., & Xavier, A. V. (1991) *Eur. J. Biochem. 202*, 1107−1115.

CCP4 -Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr. D50*, 760−763.

Coletta, M., Catarino, T., LeGall, J., & Xavier, A. V. (1991) *Eur. J. Biochem. 202*, 1101−1106.

Coutinho, I., Turner, D. L., Santos, H., LeGall, J., & Xavier, A. V. (1993) *Biochem. J. 294*, 899−908.

Czjzek, M., Payan, F., Guerlesquin, F., Bruschi, M., & Haser, R. (1994) *J. Mol. Biol. 243*, 653−667.

Devereux, R., He, S. H., Orkland, S., Atahl, D. A., LeGall, J., & Whitman, W. B. (1990) *J. Bacteriol. 172*, 3609−3619.

Dolla, A., & Bruschi, M. (1988) *Biochim. Biophys. Acta 932*, 26−32.

Dolla, A., Leroy, G., Guerlesquin, F., & Bruschi, M. (1991) *Biochim. Biophys. Acta 1058*, 171−177.

Dutton, P. N. (1971) *Biochim Biophys. Acta 226*, 63−80.

Frazão, C., Morais, J., Matias, P. M., & Carrondo, M. A. (1994) *Acta Crystallogr. D50*, 233−236.

Gayda, J. P., Yagi, T., & Bertrand, P. (1987) *FEBS Lett. 217*, 57−61.

Gayda, J. P., Benosman, H. Bertrand, P., More, C., & Asso, M. (1988) *Eur. J. Biochem. 177*, 199−206.

Guerlesquin, F., Bruschi, M., & Bovier-Lapierre, G. (1984) *Biochimie 66*, 93−99.

Guerlesquin, F., Noailly, M., & Bruschi, M. (1985) *Biochem. Biophys. Res. Commun. 130*, 1102−1108.

Guerlesquin, F., Sari, J. C., & Bruschi, M. (1987) *Biochemistry 26*, 7438−7443.

Guigliarelli, B., Bertrand, P., More, C., Haser, R., & Gayda, J. P. (1990) *J. Mol. Biol. 216*, 161−166.

Higuchi, Y., Kusunoki, M., Matsuura, Y., Yasuoka, N., & Kakudo, M. (1984) *J. Mol. Biol. 172*, 109−139.

Lee, B., & Richards, F. M. (1971) *J. Mol. Biol. 55*, 379−400.

Le Gall, J., & Peck, H. D., Jr., Eds. (1994) *Methods in Enzymology*, Vol. 243, Academic Press, New York.

Lesk, A. M., & Chothia, C. H. (1986) *Philos. Trans. R. Soc. London, A 317*, 345.

Lesk, A. M., Levitt, M., & Chothia, C. H. (1986) *Protein Eng. 1*, 77.

Liu, M.-C., Costa, C., Coutinho, I., Moura, J. J. G., Moura, I., Xavier, A. V., & LeGall, J. (1988) *J. Bacteriol. 170*, 5545−5551.

Luzatti, V. (1952) *Acta Crystallogr. 5*, 802−810.

Matias, P. M., Frazão, C., Morais, J., Coll, M., & Carrondo, M. A. (1993) *J. Mol. Biol. 234*, 680−699.

Moore, G. R., Pettigrew, G. W., & Rogers, N. K. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 4998−4999.

Morimoto, Y., Tani, T., Okumura, H., Higuchi, Y., & Yasuoka, N. (1991) *J. Biochem. (Tokyo) 110*, 532−540.

Moura, I., Moura, J. J. G., Santos, M. H., & Xavier, A. V. (1980) *Cienc. Biol. (Portugal) 5*, 195−197.

Moura, I., Teixeira, M., Huhyn, B. H., LeGall, J., & Moura, J. J. G. (1988) *Eur. J. Biochem. 176*, 365−369.

Moura, J. J. G., Xavier, A. V., Hatchikian, E. C., & LeGall, J. (1977) *FEBS Lett. 89*, 177−179.

Moura, J. J. G., Santos, H., Moura, I., LeGall, J., Moore, G. R., Williams, R. J. P., & Xavier, A. V. (1982) *Eur. J. Biochem. 127*, 151−155.

Moura, J. J. G., Costa, C., Liu, M.-Y., Moura, I., & LeGall, J. (1991) *Biochim. Biophys. Acta 1058*, 61−66.

Mus-Veteau, I., Chottard, G., Doris, L., Guerlesquin, F., & Bruschi, M. (1992) *Biochim. Biophys. Acta 1102*, 353−359.

Palma, P. N., Moura, I., LeGall, J., Van Beeumen, J., Wampler, J. E., & Moura, J. J. G. (1994) *Biochemistry 33*, 6394−6407.

Palmer, G. (1985) *Biochem. Soc. Trans. 13*, 548−560.

Park, J.-S., Kano, K., Morimoto, Y., Higuchi, Y., Yasuoka, N., Ogata, M., Niki, K., & Akutsu, H. (1991) *J. Biomol. NMR 1*, 271−282.

Pflugrath, J. W., & Messerschmidt, A. (1989) *MADNES, Munich Area Detector New EEC System*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, and Max-Planck-Institut für Biochemie, Martinsried, Germany.

Piçarra-Pereira, M. A., Turner, D. L., LeGall, J., & Xavier, A. V. (1993) *Biochem. J. 294*, 909−915.

Ramachandran, G. N., & Sasisekharan, V. (1968) *Adv. Protein. Chem. 23*, 283−437.

Ravi, N., Moura, I., Costa, C., Teixeira, M., LeGall, J., Moura , J. J. G., & Huynh, B. H. (1992) *Eur. J. Biochem. 204*, 4489−4496.

Roussel, A., & Cambillau, C. (1989) TURBO-FRODO molecular modelling package, in *Silicon Graphics Geometry Partner Directory* (Silicon Graphics, Eds.) pp 77−78, Silicon Graphics, Mountain View, CA.

Roussel, A., Fontecilla-Camps, J. C., & Cambillau, C. (1990) *Acta Crystallogr. A46*, C66−67.

Salgueiro, C., Turner, D. L., LeGall, J., & Xavier, A. V. (1992) *FEBS Lett. 314*, 155−158.

Santos, H., Moura, J. J. G., Moura, I., LeGall, J., & Xavier, A. V. (1993) *Eur. J. Biochem. 141*, 283−296.

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

Stewart, D. E., LeGall, J., Moura, I., Moura, J. J. G., Peck, H. D., Jr., Xavier, A. V., Weiner, P. K., & Wampler, J. E. (1989) *Eur. J. Biochem. 185*, 695−700.

Sweet, R. M. (1986) *Biopolymers 25*, 1565

Taylor, C. P. S. (1977) *Biochim. Biophys. Acta 491*, 137−149.

Turner, D. L., Salgueiro, C. A., LeGall, J., & Xavier, A. V. (1992) *Eur. J. Biochem. 210*, 931−936.

Turner, D. L., Salgueiro, C. A., Catarino, T., LeGall, J., & Xavier, A. V. (1994) *Biochim. Biophys. Acta 1187*, 232−235.

Wagen, E. S. (1978) *Nature 275*, 73−74.

Walker, F. A., Huynh, B. H., Scheidt, W. R., & Osvath, S. R. (1986) *J. Am. Chem. Soc. 108*, 5288−5297.

Weiner, P. K., & Kollman, P. A. (1981) *J. Comput. Chem. 2*, 287−303.

Xavier, A. V., & Moura, J. J. G. (1978) *Biochimie 60*, 327−338.

Xavier, A. V., Moura, J. J. G., LeGall, J., & DerVartanian, D. V. (1979) *Biochimie (Paris) 61*, 689−695.

BI950863I