Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of dimeric cytochrome c_3 from *Desulfovibrio gigas* at 1.2 Å resolution

David Aragão,[a] Carlos Frazão,[a] Larry Sieker,[a,b] George M. Sheldrick,[c] Jean LeGall[a,d] and Maria Arménia Carrondo[a]*

[a]Instituto de Tecnologia Química e Biológica, P-2781-901 Oeiras, Portugal, [b]Department of Biological Structure, Box 357420, University of Washington, Seattle, WA 98195-7420, USA, [c]Lehrstuhl für Strukturchemie der Universität Göttingen, Tammannstrasse 4, D-37077 Göttingen, Germany, and [d]Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA

Correspondence e-mail: carrondo@itqb.unl.pt

Received 10 September 2002
Accepted 21 January 2003

PDB Reference: cytochrome c_3, 1gyo, r1gyosf.

The structure of dimeric cytochrome c_3 from the sulfate-reducing bacterium *Desulfovibrio gigas*, diDg, obtained by *ab initio* methods was further refined to 1.2 Å resolution, giving final reliability factors of $R_{\rm free}$ = 14.8% and R = 12.4%. This cytochrome is a dimer of tetraheme cytochrome c_3 molecules covalently linked by two solvent-accessible disulfide bridges, a characteristic unique to members of the cytochrome c_3 superfamily. Anisotropic analysis using the semi-rigid TLS method shows different behaviour for analogous loops in each monomer arising from their different packing environments. A detailed sequence and structural comparison with all other known cytochrome c_3 domains in single- and multi-domain cytochromes c_3 shows the presence of structurally conserved regions in this family, despite the high variability of the amino-acid sequence. An internal water molecule is conserved in a common structural arrangement in all c_3 tetraheme domains, indicating a probable electron-transfer pathway between hemes I and II. Unique features of diDg are an internal methionine residue close to heme I and to one of the axial ligands of heme III, where all other structures of the cytochrome c_3 superfamily have a phenylalanine, and a rather unusual C*XXX*CH heme-binding motif only found so far in this cytochrome.

1. Introduction

The structures of multiheme cytochromes from a wide variety of organisms have been determined; the cytochromes may perform a variety of functions, but they are generally related to electron-transfer processes. Structures containing a tetraheme *c* motif show that it can adopt different spatial arrangements, such as the heme cluster in cytochrome c_3 (Haser *et al.*, 1979) or other four-heme arrangements as found in the purple bacteria photosynthetic centre (Deisenhofer *et al.*, 1995), in cytochrome c_{554} (Iverson *et al.*, 1998) or in cytochrome *c* fumarate reductase (Bamford *et al.*, 1999; Taylor *et al.*, 1999). The heme arrangements found in the last two cytochromes are also present, at least partially (Iverson *et al.*, 1998; Bamford *et al.*, 1999), in the higher order multiheme cytochromes hydroxylamine oxidoreductase (Igarashi *et al.*, 1997) and cytochrome *c* nitrite reductase (Einsle *et al.*, 1999).

The most extensively studied class of tetraheme cytochromes is the c_3 superfamily (Bruschi, 1994). This class of cytochromes, defined by Ambler as type III (Ambler, 1980), has been found in sulfate- and sulfur-reducing bacteria, with all heme groups showing bis-histidinal axial ligation and relatively low reduction potentials (−50 to −400 mV). Crystal

structures of members of this superfamily include seven single-domain tetraheme cytochromes, two dimeric tetraheme cytochromes, two nine-heme cytochromes and one triheme cytochrome. The single-domain tetraheme cytochromes are those from *Desulfomicrobium* (*Dm.*) *norvegicum* (previously known as *Dm. baculatus* Norway 4; Dmn; Haser *et al.*, 1979), *Desulfovibrio* (*D.*) *vulgaris* Miyazaki (DvM; Higuchi, Bando *et al.*, 1981), *D. vulgaris* Hildenborough (DvH; Matias *et al.*, 1993), *D. desulfuricans* ATCC 27774 (Dd; Morais *et al.*, 1995), *D. gigas* (Dg; Matias *et al.*, 1996) and *D. desulfuricans* Essex 6 (DdE; Einsle *et al.*, 2001) and the acidic cytochrome c_3 from *D. africanus* (Daa; Norager *et al.*, 1999). Dimeric tetraheme cytochromes include the non-covalently bound tetraheme domains observed in *Dm. norvegicum* (diDmn; Czjzek *et al.*, 1996) and the bis-disulfide bridged tetraheme domains observed in *D. gigas* (diDg; Frazão *et al.*, 1999). A dimodular c_3 tandem on a single polypeptide chain has also been crystallographically characterized in the nine-heme cytochromes from *D. desulfuricans* (9Dd; Matias, Coelho *et al.*, 1999; Matias, Saraiva *et al.*, 1999) and *D. desulfuricans* Essex 6 (9DdE; Umhau *et al.*, 2001). Higher molecular mass (Hmc) cytochromes c_3, containing 16 heme groups, have been isolated from *D. vulgaris* Hildenborough (Higuchi *et al.*, 1987), *D. vulgaris* Miyazaki (Ogata *et al.*, 1993) and *D. gigas* (Chen *et al.*, 1994) and were predicted to be formed by the association of three tetraheme, one triheme and one single heme motif (Chen *et al.*, 1994). Another extension to the canonical tetraheme cluster is given by cytochrome c_7 from *Desulfuromonas acetoxidans* (Dsa; Czjzek *et al.*, 2001), a triheme cytochrome which resembles the other cytochromes c_3 in overall structure but has the region corresponding to heme II simply replaced by a straight polypeptide connection.

Dimeric cytochromes c_3 have been found in *D. gigas* (Bruschi *et al.*, 1969), *Dm. norvegicum* (Loutfi *et al.*, 1989) and *D. africanus* (Pieulle *et al.*, 1996). The two dimeric cytochromes of known sequence, diDmn (Bruschi *et al.*, 1994) and diDg (Bruschi *et al.*, 1996), show no particular primary structure similarity between themselves or with the monomeric cytochromes c_3 from the corresponding species (Bruschi *et al.*, 1996). In the case of *D. africanus* it was reported that the presence of a high salt concentration is required for stability reasons (Pieulle *et al.*, 1996), while oxidized diDg precipitates and crystallizes in low ionic strength media (Sieker *et al.*, 1986). While diDg has been shown to be specific for thiosulfate reductase (Bruschi *et al.*, 1969, 1977; Hatchikian *et al.*, 1972), diDmn has more recently been shown to be an efficient electron acceptor from hydrogenase *via* monomeric cytochrome c_3 (Aubert *et al.*, 2000). *D. gigas* is the only *Desulfovibrio* species where disulfide bridges have been described, in particular in ferredoxin II (Kissinger *et al.*, 1989) and in cytochrome diDg (Frazão *et al.*, 1999), and is the only bacterium described so far in which mono-domain, di-domain and Hmc cytochromes c_3 are all present (Bruschi *et al.*, 1996). Additionally, *D. gigas* only contains a periplasmic [NiFe] (Voordouw *et al.*, 1990), whereas other species may contain up to three different hydrogenases, namely [Fe] and [NiFeSe] hydrogenases.

A preliminary description of diDg structure determination and refinement to 1.4 Å has already been published (Frazão *et al.*, 1999). In this paper, we extend the refinement to 1.2 Å resolution and present a more detailed description of this cytochrome crystal structure, analyse the anisotropic details perceptible with the actual quasi-atomic resolution and compare the two independent monomers in the asymmetric unit as well as comparing them with all tetraheme cytochrome c_3 domains of this superfamily.

2. Materials and methods

2.1. Protein purification and crystallization

diDg protein samples (Bruschi *et al.*, 1969) were concentrated in 1 *M* Tris–maleate pH 3.5. Three dialysis experiments were set up at pH values of 5.5, 6.0 and 6.5 at 277 K. The volume of the protein dialysis chamber used for these experiments was about 0.2 ml. Chambers were made from small mini test tubes (5 × 30 mm). The tubes were cut to about 15 mm in length and the ends were smoothed with emery paper and then fire-polished. Each chamber was almost filled with the extremely red protein solution at $\sim$15 mg ml^{-1} concentration and topped up with water (to prevent loss of protein as the membrane is put in place). The tube/chamber was covered with a dialysis membrane and held in place with a short segment of Tygon tubing. Part of the segment of Tygon tubing was pushed beyond the edge of the membrane to make a Tygon-to-glass seal in order to prevent the possibility of leakage past the membrane and to prevent the membrane from slipping off as the pressure increases. Each dialysis chamber was then placed into a 10 ml test tube containing 4 ml of 0.5 *M* Tris–maleate solution. The protein chamber was set up with the dialysis membrane at the top, so that the less dense outer solution could only exchange by diffusion processes. The reverse arrangement with the membrane at the bottom allows the more dense (concentrated) buffer to exchange across the membrane too rapidly by gravity. The test tubes were set up at an angle in order to prevent the possibility of an air bubble blocking the solution exchange. This also allowed better visualization of the progress of dialysis and crystal growth without disturbing the process of crystallization.

The ionic strength of the outer solution was reduced by incremental addition of distilled water to the 4 ml volume in order to reduce the outer solution buffer concentration to 0.2 *M*. This was allowed to stand until crystals could be seen to develop. Further reduction of the buffer concentration to 0.18 *M* and subsequently to 0.15 *M* was performed in order to continue the crystal-growth process. Over a period of one month, the ionic strength of the dialysing solution was reduced to a buffer concentration of 50 m*M*, whereupon the crystallizing solution became colourless, indicating that crystal growth had essentially finished. Cryoprotection was obtained by plunging the crystals into 28%(*v*/*v*) glycerol solution. Several diffraction data sets to a maximum resolution of 1.65 Å were obtained using an in-house Cu rotating-anode Enraf–Nonius FR571 generator operating at 4.5 kW and a

Table 1
Data-collection and diffraction statistics.

(*a*) Data-collection statistics.

Data set (crystal)	High (*A*)	Low (*B*)
Resolution limits (Å)	3.0–1.2	30–1.85
Detector distance (mm)	100	215
Oscillation range (°)	1.2	2.0
Total oscillation (°)	98.4	140
Exposure time (s)	600	600
Refined mosaicity (°)	0.487	0.232
λ (Å)	1.1037	1.1037

(*b*) Diffraction data statistics. Values in parentheses are for the outer resolution shell.

Space group	$P3_1$
Wavelength (Å)	1.1037
Unit-cell parameters (Å, °)	$a = b = 56.67$, $c = 94.17$, $\gamma = 120.00$
No. unique intensities	104411
Redundancy	4.9
Overall resolution (Å)	26.44–1.20 (1.21–1.20)
Completeness (%)	98.7 (95.6)
$R_{\rm merge}(I)$ (%)	9.7 (43.7)
$R_{\rm p.i.m.}$ (%)	3.7 (30.3)
$\langle I\rangle/\sigma\langle I\rangle$	16.3 (2.8)

Table 2
Refinement statistics.

Refinement and fitting stages	Resolution (Å)	Non-H protein atoms	Solvent (water + glycerol)	R†	$R_{\rm free}$†
Ab initio solution	1.20	1763	—	—	—
Fit on initial smoothed maps‡	1.40	1939	56 + 2	—	—
All isotropic	2.50	1939	56 + 2	25.3	30.7
Fe atoms anisotropic	1.50	2000	186 + 2	19.2	22.4
All anisotropic	1.40	2000	171 + 2	15.1	18.2
All anisotropic	1.20	1996	264 + 2	13.3	15.8
H atoms	1.20	1996	264 + 2	12.2	14.7
Final model, all data	1.20	1996	264 + 2	12.1	—

† R values calculated for $F > 4\sigma(F)$ data. ‡ The phases used for these maps were obtained using the set of initial atoms refined isotropically.

MAR Research 300 mm imaging plate. Two different cryo-cooled crystals at 110 K were used to obtain higher resolution data (1.2 Å) at beamline BW7B of the DESY synchrotron, EMBL Outstation, Hamburg with a MAR Research 345 mm imaging plate. Data were processed and merged using *DENZO*/*SCALEPACK* (Otwinowski & Minor, 1997) (see Table 1 for data-collection and diffraction data statistics).

2.2. Structure determination

The three-dimensional structure of diDg was determined *ab initio* using the program *SHELXD* (Uson & Sheldrick, 1999) with default settings, applying the 'half-baked' strategy (Sheldrick, 1998). The native single-wavelength diffraction data to 1.2 Å resolution were used and a model with eight Fe atoms, 18 S atoms and 2200 other second-period atoms was requested. The program uses two iterative loops: the first loop finds seed molecules, which are expanded in the second loop to the final *ab initio* models. For the first loop, the program generated initial sets of randomly distributed atoms, read in the 104 759 unique intensities and used the highest 7533 normalized observed structure factors [with $|E_o - \sigma(E_o)|/2 > 1.6$], producing 2 054 501 triple phase relationships. In this loop, each initial set of randomly distributed atoms was optimized by minimizing the figure of merit $\sum E_c^2(E_o^2 - 1)$ until a satisfying set of atoms was found, which led to a set of 65 atoms with CC = 18%. This seed was then used in the second loop in an iterative procedure of seven cycles of peak-list optimization (Sheldrick & Gould, 1995) and a model of 1763 point atoms was produced with CC = 74%. This *ab initio* model was smoothed by isotropic atomic displacement parameter (ADP) refinement during ten cycles of conjugate-gradient least squares using data to 1.4 Å. σ_A maps (Read, 1986) were produced using the smoothed model as a source of phases; a 1938 protein atom model of the di-tetraheme cytochrome (95% of the sequenced protein) plus 58 waters and two glycerol solvent molecules was easily modelled and fitted into the maps on a graphics workstation. One dimer was found in the asymmetric unit.

2.3. Structure refinement

The program *SHELXL* (Sheldrick & Schneider, 1997) was used for model refinement using 3.5% of the intensities for cross-validation distributed in thin shells of constant resolution in order to avoid bias from the non-crystallographic symmetry (NCS). NCS restraints imposed similar dihedral angles and similar isotropic ADPs between the two monomers of the asymmetric unit. The Engh and Huber (1991) stereochemical dictionary was used for the polypeptide moiety. Taking advantage of their internal symmetry and fourfold or eightfold redundancy, the geometrical parameters of the heme groups were restrained to be similar but without target values. The initial electron-density fitted model was refined in increasing resolution steps, beginning from 2.5 Å resolution data and refining isotropic ADPs. Anisotropic refinement was introduced for the Fe atoms with 1.5 Å resolution data and further generalized to all remaining atoms above 1.4 Å resolution. Anisotropic refinement was applied using default 'rigid-bond restraints' (DELU) and 'similar ADP restraints' (SIMU) on neighbouring atoms (Sheldrick & Schneider, 1997). Additionally, 'approximately isotropic restraints' (ISOR; Sheldrick & Schneider, 1997) were also used on atoms that might otherwise refine to non-positive definite displacement tensors, in particular for a few hydrophilic side chains and for solvent waters. At the end of the refinement, H atoms were introduced using the riding model. The structure was checked against σ_A maps (Read, 1986) using *XtalView* (McRee, 1992) and corrected and completed with discretely alternating side chains. Further solvent waters were added and the protonation assignment of basic residues was eventually corrected according to their immediate environment. See Table 2 for details of the progress of the refinement.

2.4. Anisotropic analysis

The anisotropic information available in the diDg high-resolution model was analysed using *ANISOANL* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). Semi-rigid groups were defined as sets of atoms with similar vibration and TLS (translation, libration and screw-rotation) parameters (Schomaker & Trueblood, 1968). The contribution of the rigid-body motion for each atom was calculated and averaged over the atoms in each residue, $U_{\rm iso}^{\rm calc}$, as well as the average of the three eigenvalues of the ADP tensors, $U_{\rm iso}$, and their radial and tangential components, $U_{\rm rad}$ and $U_{\rm tang}$, respectively.

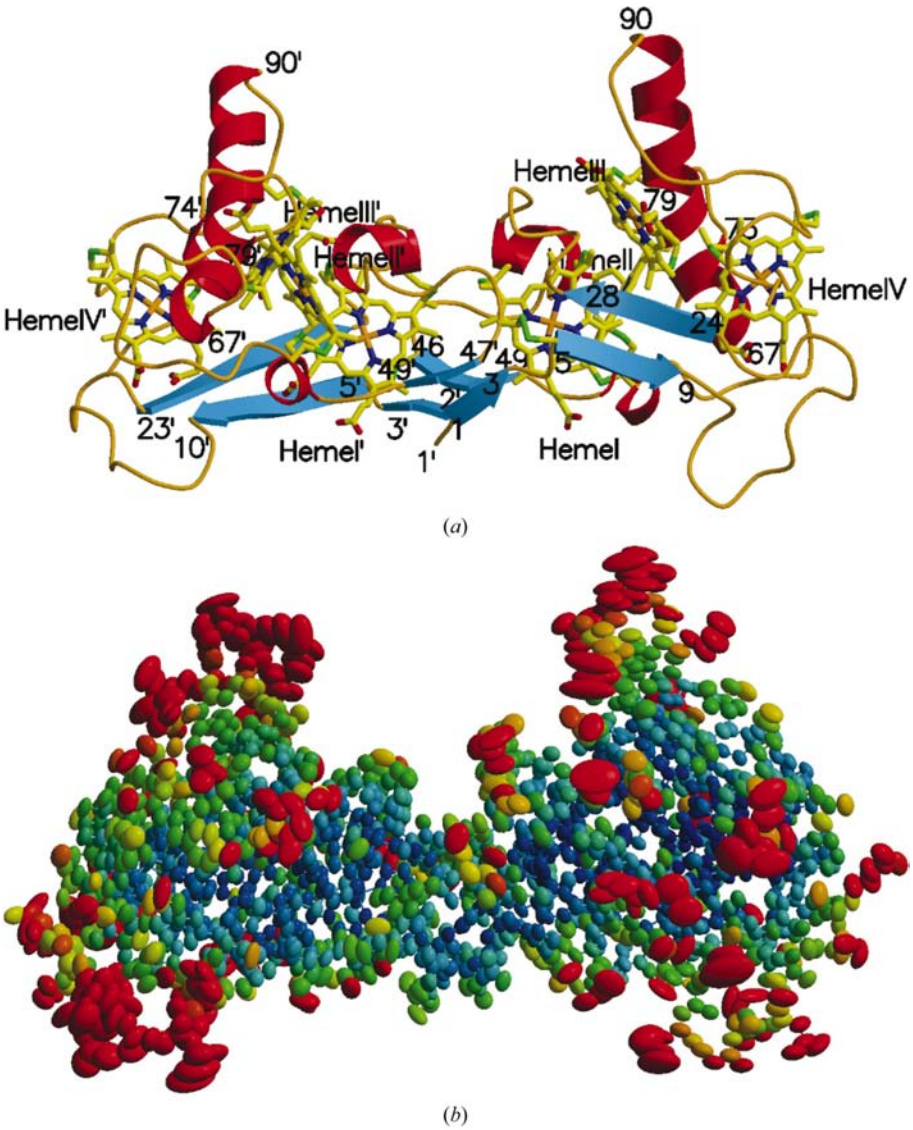


Figure 1
(*a*) Cartoon representation of the diDg dimer. The two monomers *A* (right) and *B* (left, numbered with primes) are related by a twofold axis which lies in the centre of the figure, approximately vertical in the plane of the paper. Secondary-structure elements, delimited by corresponding residue numbers in each monomer, are coloured blue for the β-chains and red for the α-helices; the remaining random-coil segments are orange. Stick bonds are shown for the heme groups and cysteine side chains, including inter-monomer disulfide bonds. (*b*) Ellipsoidal ADP representation of the diDg dimer. Using the same dimer orientation as in (*a*), the atomic anisotropic behaviour of diDg is depicted with 50% probability ellipsoids coloured from dark blue (10 Å^2) to dark red (30 Å^2). Note the gradual increase in thermal vibration from the core to the surface, the different magnitude of displacement of the loops of the β-hairpin motif of each monomer and the main libration mode perpendicular to the polypeptide path, *e.g.* around residue 90.

3. Results and discussion

3.1. Crystallization

Crystals of diDg were initially obtained by Sieker *et al.* (1986) using the salting-out technique (Blundell & Johnson, 1976). This cytochrome is quite insoluble at physiological pH, requiring 1 *M* phosphate or 4 *M* Tris–HCl to obtain a reasonable amount of protein in solution (Sieker *et al.*, 1986). To some extent, this insolubility explains the paucity of studies on this multi-heme protein. However, it was subsequently discovered that at pH 3.5 the cytochrome is not only stable but is also much more soluble (Sieker, unpublished results). Crystals grew as slightly elongated dark red trigonal bipyramids (Sieker *et al.*, 1986; Sieker, 1995) independently of the initial pH of the media. Diffraction data from three crystals grown at pH 6.0 revealed merohedral twinning diffraction statistics (C. Frazão, unpublished results). All measured crystals from media at the other pH values, pH 5.5 and 6.5, did not show any twinning characteristics.

3.2. *Ab initio* determination

Structure determination was straightforward using the *ab initio* option in the *SHELXD* program, which required no human intervention and worked from default values. It also established the total number of molecules in the asymmetric unit as one homodimer (corresponding to 61% solvent content) instead of two (corresponding to 23% solvent content), which was the *a priori* expected value given the very high resolution of the diffraction data. The relatively high number of heavier atoms, including eight Fe and 20 S atoms among ~2000 other second-period atoms in the dimer, certainly helped in the 'seeding' procedure, as the majority of those heavier atoms are associated with heme groups, which were the fragments of the structure with the lowest ADP values. The initial seed in the *ab initio* procedure was obtained using the highest normalized structure factors, a procedure that scales up the data with increasing reso-

lution. These data mainly arise from scattering from atoms with a higher number of electrons and lower ADP values, which in the present case represent an important fraction of the whole structure.

3.3. Structure description and refinement

The diDg homodimer structure consists of two crystallographically independent monomers linked by two disulfide bridges involving Cys5 and Cys46 of each monomer (Fig. 1*a*). The monomers display the characteristic tetraheme cytochrome c_3 fold, with a set of four heme groups wrapped by a polypeptide chain which is covalently linked to the hemes by pairs of thioether bridges. A pair of histidine residues completes the octahedral coordination of each heme iron. Each monomer contains as secondary-structure elements a hairpin motif consisting of an antiparallel β-sheet (residues 6–9 and 24–27 for monomer *A*) connected by a loop of 14 residues and four helices, of which two are short helical segments (residues 35–40 and 51–55) and the other two are somewhat longer (residues 67–73 and 79–90). These last segments form a long helical motif broken by an insertion of five residues and kinked by ∼28°. Additionally, the N-terminal regions of each monomer (residues 1/2–4) form two incipient two-stranded antiparallel β-sheets with a small strand of the other monomer (residues 46–49; Frazão *et al.*, 1999) (see Fig. 1*a*). The fold of each diDg monomer and its secondary-structure elements is characteristic of the cytochrome c_3 family (Higuchi, Kusunoki *et al.*, 1981), but includes two additional inter-monomeric small β-sheets involving each N-terminus and a β-chain of the other monomer immediately after the unique inter-monomeric disulfide bonds.

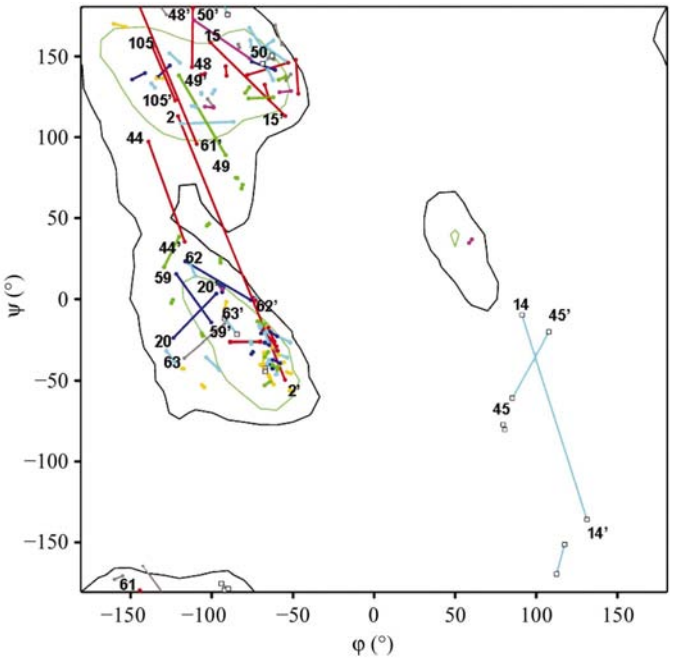


Figure 2
NCS Kleywegt plot of the diDg dimer. Ramachandran diagram of the two monomers with segments linking the φ–ψ coordinates of pairs of homologous residues. Pairs of residues with rather different Ramachandran coordinates are numbered, using primes for residues of monomer *B*. The N-terminus of monomer *A*, in particular the side chain of Leu1*A*, is located in an inter-monomer pocket and crosses the non-crystallographic twofold rotation axis, which implies that Leu1*B* must disobey the non-crystallographic symmetry relationship (or it would clash with the Leu1*A* side chain). It is therefore rotated towards the exterior of the dimer, establishing hydrophobic packing interactions with the aliphatic part of the Arg62*B* side chain. These interactions produce peptide-conformation differences in residues 2 and 62 as well as in the neighbouring residues 59, 61 and 63. Additionally, Phe49*A* fits ∼2 Å deeper than its counterpart Phe49*B* in the altered pocket, which induces different Ramachandran coordinates for residues 48, 49 and 50. While the β-hairpin loop of monomer *B* is immersed in the solvent media, in monomer *A* it is involved in packing contacts, resulting in different Ramachandran coordinates for loop residues 14, 15 and 20 and for the contacting residue 44 and its neighbour 45.

The model contains 106 amino-acid residues in each chain. Three residues are missing in each monomer at the C-terminus (Ala107-Gln108-Lys109) as they are not visible in electron-density maps, a feature not overly surprising given their mainly hydrophilic nature. Some hydrophilic side chains at the protein surface (10 in monomer *A* and 17 in monomer *B*) are only partially visible in 1.5σ electron-density maps, but were nevertheless modelled based on local residual density. The structure also contains 17 residues, a heme propionic group and a glycerol solvent molecule modelled in discretely alternating conformations, with a total of 63 discretely disordered atoms. The final model was refined anisotropically for all atoms except H atoms, which were included using the riding model. The final reliability factors of R = 12.4% and $R_{\rm free}$ = 14.8% were within the usual values for such a high-resolution crystallographic structure. *PROCHECK* (Laskowski *et al.*, 1993) indicated that no residues were located in disallowed or generously allowed regions of the Ramachandran plot; 19% of the residues were located in additional allowed regions and the remaining residues were located in the most favoured regions.

3.4. Packing analysis

A comparison of the Ramachandran plots of the two monomers with *SHELXPRO* (Sheldrick & Schneider, 1997) indicates that over 10% of the residues show discrepancies in peptide-chain torsional angles that may be explained by different molecular-packing contacts in the two independent monomers (see Fig. 2). Contact regions involve hemes IV and IV′ through loops 15–20, 59–63 and 98–105 from neighbouring monomers. Other packing interactions involve hemes II and II′ flanked by loop 42–57 and the disulfide bridges 5–46. Owing to the pseudo-symmetrical nature of the non-crystallographic twofold axis, monomer *A* establishes more packing interactions than monomer *B* (165 compared with 129 van der

Waals contacts to 4.0 Å and 21 compared with 17 hydrogen bonds to 3.6 Å, respectively).

3.5. Anisotropic analysis

An inspection of the anisotropic behaviour of the diDg dimer (see Fig. 1*b* and Fig. 3) shows that in each monomer the residues have a fairly flat mean ADP distribution, except for the β-hairpin loop region (residues 11–21) and the C-tip of the long α-helix motif and its adjacent random-coiled segment (residues 85–95), as well as the C-terminal regions, which tend to disorder in the solvent media. A comparison of mean magnitudes of displacement per residue, $U_{\rm iso}$, and mean TLS components per residue, $U_{\rm iso}^{\rm calc}$, shows that the vibrational movement of the β-hairpin loop 11–21 in monomer *B* is significantly higher than the rigid-body model for the monomer would predict, while in monomer *A* such additional vibration seems to freeze out at the middle of the loop. This different behaviour of the loops can be explained by their different packing environments in each monomer (see Fig. 3*a*). Additionally, the decomposition of ADPs into their radial and tangential components shows that the libration in the labile regions (residues 11–21 and 85–95) is mainly tangential, except for the solvent-surrounded β-hairpin loop of monomer *B*, where the radial component is of a similar magnitude (see Fig. 3*b*). In the case of the helix tip and the coiled segment where residues 85–95 turn back to the protein core, the residues lack radial movement but show tangential vibration owing to the stiffness of the adjacent helical structure.

3.6. Comparison of c_3 structures

A superposition of the available cytochrome c_3 'modules' from both monomeric and multimeric or multidomain cytochromes c_3 was undertaken in order to detect structural invariants in this superfamily. This analysis is relevant given the fact that the sequence identity in the cytochrome c_3 superfamily is rather low, as indicated for each pair in Table 3.

A crystal structural model from each cytochrome c_3 was retrieved from the PDB (Berman *et al.*, 2000) and the individual domains were superimposed with *MODELLER* (Sali & Blundell, 1993) using 1.75 Å as the radius of the spheres defining structurally homologous C^{α} atoms (see Table 3). Although the superposition procedure was based on the minimization of C^{α} distances, Table 3 shows that their pairwise root-mean-square distances (r.m.s.d.s) are systematically larger than the r.m.s.d.s calculated for the heme Fe-atom positions. The exception of the three-heme Dsa case (last column of Table 3) was because of a deficient overlap of heme I that distorts the entry in Table 3; heme I has an Fe r.m.s.d. of 2.61 Å, deviating strongly from the values of 0.75 and 0.94 Å for the Fe atoms of hemes III and IV, respectively. Nevertheless, Table 3 clearly indicates that the heme core is the most conserved structural feature of the family.

In order to obtain an overview of the relationships among cytochrome c_3 molecules/domains, *CLUSTALX* (Thompson *et al.*, 1997) was used to obtain a distance matrix from the three-dimensional alignment, which is displayed with *NJPLOT* (Perriere & Gouy, 1996) in Fig. 4. The obtained phyletic tree gives an idea of the (dis)similarities between the compared

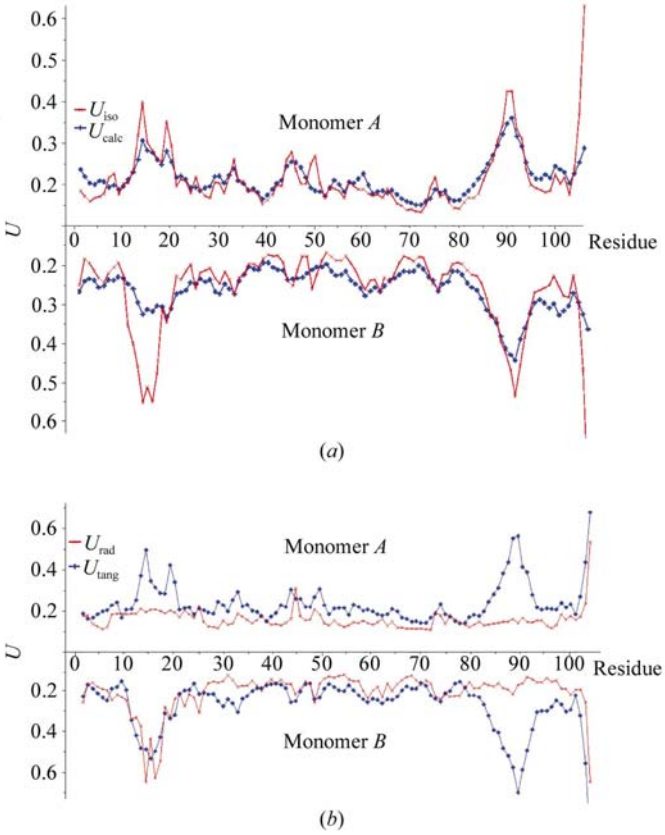


Figure 3
(*a*) Mean atomic displacement parameters for residues. The observed mean displacement magnitudes of the residues (blue crosses, $U_{\rm iso}$) show three regions in each monomer with higher vibrational mobility, corresponding to the β-hairpin loop (residues 11–21), the end of the long helical motif and the adjacent coiled segment 85–95) and the C-terminal region. A calculation using *ANISOANL* of the mean residue TLS components (red crosses, $U_{\rm iso}^{\rm calc}$), assuming each monomer to be a rigid body, shows that they are systematically overestimated when compared with the corresponding global displacement magnitudes $U_{\rm iso}$. Nevertheless, this analysis reveals that the 11–21 β-hairpin loops have remarkably different behaviours in the two monomers. The different behaviours are explained by the different loop environments; in monomer *B* the loop is only surrounded by solvent, while it has packing contacts in monomer *A*. These involve a salt bridge between the δ-guanido group of Arg20*A* and a heme II′ propionic group of a symmetry-related dimer, together with the hydrogen bond between His18*A* NE2 and the O2 atom of a glycerol molecule that in turn is linked *via* hydrogen bonds to the carboxylic group of Asp44*B* of a symmetry-related molecule. Additionally, the loop in *A* has two additional hydrogen bonds from the main-chain amine N atom of Glu15 and the carbonyl O atom of Asp16 to Pro12 O and Arg59 NH1, respectively, which explain the rigidness of the middle section of this loop. (*b*) Rigid-body vibration of diDg. Decomposition of the residue mean anisotropy parameters into rigid-body tangential (blue crosses, $U_{\rm tang}$) and radial (red crosses, $U_{\rm rad}$) components for monomers *A* and *B*. In monomer *B*, the β-hairpin 11–21 loop has both tangential and radial movements contributing to the observed local high displacement parameters as this loop has no space restrictions in its solvent-surrounding neighbourhood. In monomer *A*, the steric interactions of this loop mainly restrict radial movements; namely, the packing contacts of His18*A* and Arg20*A* obviate radial movements and the main-chain hydrogen bonds of Glu15*A* and Asp16*A* hinder tangential movements in the middle of the loop. The 85–95 region in both monomers only shows rigid-body tangential vibrations, which is a consequence of the sterical restraints of the neighbouring helical motif.

Table 3
Comparison of cytochrome c_3 domains.

Values given above the diagonal (from top left to bottom right) are positional root-mean-square distances for C^{α} atoms (and for heme Fe atoms) (Å). Values on the diagonal are numbers of residues. Values below the diagonal are numbers of three-dimensional homologous residues within spheres of 1.75 Å radius (and their percentage sequence identities). Dmn, *Dm. norvegicum* (PDB code 2cy3; Czjzek *et al.*, 1996); DvM, *D. vulgaris* Miyazaki (2cdv; Higuchi, Bando *et al.*, 1981); DvH, *D. vulgaris* Hildenborough, chain *A* (2cth; Simoes *et al.*, 1998); Dg, *D. gigas* (1wad; Matias *et al.*, 1996); Dd, *D. desulfuricans* ATCC 27774 (3cyr; Simoes *et al.*, 1998); DdE, *D. desulfuricans* Essex 6 (1i77; Einsle *et al.*, 2001); diDmn, *Dm. norvegicum* di-tetraheme (1czj; Czjzek *et al.*, 1996); diDg, *D. gigas* di-tetraheme, chain *B* (1gyo); 9Dd_1, *D. desulfuricans* ATCC 27774 first domain of nine-heme c_3, chain *A* (19hc; Matias, Coelho *et al.*, 1999); 9DdE1, *D. desulfuricans* Essex 6 first domain of nine-heme c_3 (1duw; Umhau *et al.*, 2001); 9Dd_2, *D. desulfuricans* ATCC 27774 second domain of nine-heme c_3, chain *A* (19hc; Matias, Coelho *et al.*, 1999); 9DdE2, *D. desulfuricans* Essex 6 second domain of nine-heme c_3 (1duw; Umhau *et al.*, 2001); Daa, *D. africanus* acidic oxidized c_3 (3cao; Norager *et al.*, 1999); Dsa, *Desulfuromonas acetoxidans* three-heme c_3 (1hh5; Czjzek *et al.*, 2001).

	Dmn	DvM	DvH	Dg	Dd	DdE	diDmn	diDg	9Dd_1	9DdE1	9Dd_2	9DdE2	Daa	Dsa
Dmn	118	1.30 (0.89)	1.25 (0.80)	1.34 (0.96)	1.36 (1.06)	1.34 (0.88)	1.28 (0.80)	1.33 (0.68)	1.56 (0.73	1.61 (0.75)	1.47 (0.92)	1.48 (0.92)	1.26 (0.69)	1.38 (1.69)
DvM	81 (29)	107	0.52 (0.23)	0.74 (0.33)	1.06 (0.48)	1.00 (0.42)	1.29 (0.54)	1.27 (0.81)	1.49 (0.91)	1.52 (0.89)	1.63 (1.33)	1.64 (1.36)	1.54 (1.20)	1.60 (1.69)
DvH	81 (30)	107 (87)	107	0.78 (0.30)	1.06 (0.53)	1.00 (0.48)	1.26 (0.51)	1.28 (0.78)	1.55 (0.89)	1.54 (0.89)	1.62 (1.24)	1.63 (1.29)	1.51 (1.08)	1.59 (1.69)
Dg	83 (30)	100 (50)	100 (50)	111	1.16 (0.65)	1.14 (0.64)	1.35 (0.53)	1.40 (0.88)	1.55 (0.97)	1.57 (0.98)	1.59 (1.32)	1.62 (1.36)	1.57 (1.16)	1.68 (1.65)
Dd	73 (29)	97 (38)	97 (42)	96 (37)	106	0.52 (0.24)	1.30 (0.86)	1.39 (0.94)	1.41 (1.09)	1.49 (1.10)	1.63 (1.45)	1.64 (1.48)	1.57 (1.39)	1.63 (1.93)
DdE	74 (34)	98 (38)	98 (43)	97 (36)	106 (82)	107	1.40 (0.80)	1.34 (0.78)	1.49 (0.97)	1.57 (0.97)	1.62 (1.31)	1.62 (1.31)	1.50 (1.23)	1.67 (1.89)
diDmn	85 (28)	76 (29)	76 (28)	81 (25)	78 (26)	80 (28)	109	1.34 (0.71)	1.28 (0.63)	1.21 (0.64)	1.52 (1.13)	1.55 (1.19)	1.56 (1.03)	1.33 (1.55)
diDg	82 (33)	73 (27)	73 (26)	75 (25)	73 (27)	74 (25)	91 (26)	106	1.45 (0.55)	1.44 (0.54)	1.54 (0.92)	1.48 (0.95)	1.34 (0.71)	1.63 (1.65)
9Dd_1	70 (21)	70 (27)	71 (27)	70 (21)	69 (25)	71 (22)	76 (27)	70 (20)	134	0.57 (0.21)	1.32 (0.87)	1.29 (0.92)	1.56 (0.77)	1.46 (1.26)
9DdE1	71 (20)	70 (26)	70 (26)	70 (20)	70 (23)	72 (21)	74 (25)	70 (19)	131 (84)	131	1.30 (0.92)	1.33 (0.97)	1.54 (0.77)	1.43 (1.19)
9Dd_2	69 (22)	72 (22)	72 (22)	72 (21)	69 (24)	69 (24)	71 (21)	69 (26)	67 (26)	67 (21)	126	0.54 (0.16)	1.46 (0.72)	1.50 (1.83)
9DdE2	69 (20)	72 (21)	72 (21)	72 (20)	69 (24)	69 (23)	71 (22)	68 (25)	66 (21)	67 (20)	125 (75)	126	1.43 (0.75)	1.54 (1.90)
Daa	68 (21)	66 (22)	66 (21)	64 (19)	60 (20)	61 (20)	66 (21)	68 (23)	64 (22)	64 (25)	73 (24)	72 (25)	102	1.54 (1.45)
Dsa	47 (25)	47 (19)	47 (21)	47 (21)	44 (22)	45 (22)	46 (28)	49 (25)	46 (16)	46 (15)	52 (21)	52 (21)	51 (22)	68

cytochrome structures, based primarily on the fit of their three-dimensional structures (Johnson, Sali *et al.*, 1990; Johnson, Sutcliffe *et al.*, 1990). One should take in consideration that diDg is unique in containing solvent-accessible disulfide bridges and in showing high specificity towards thiosulfate reductase (Bruschi *et al.*, 1969, 1977; Hatchikian *et al.*, 1972).

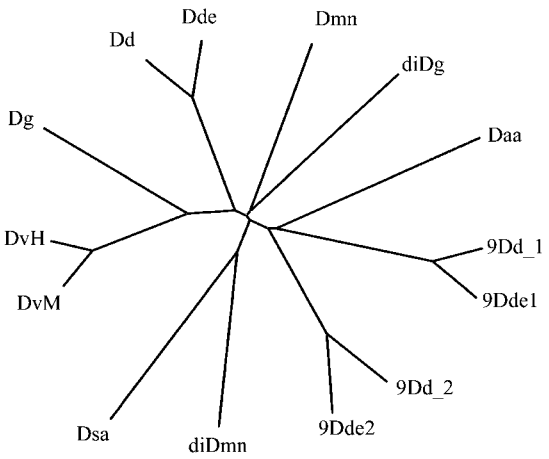


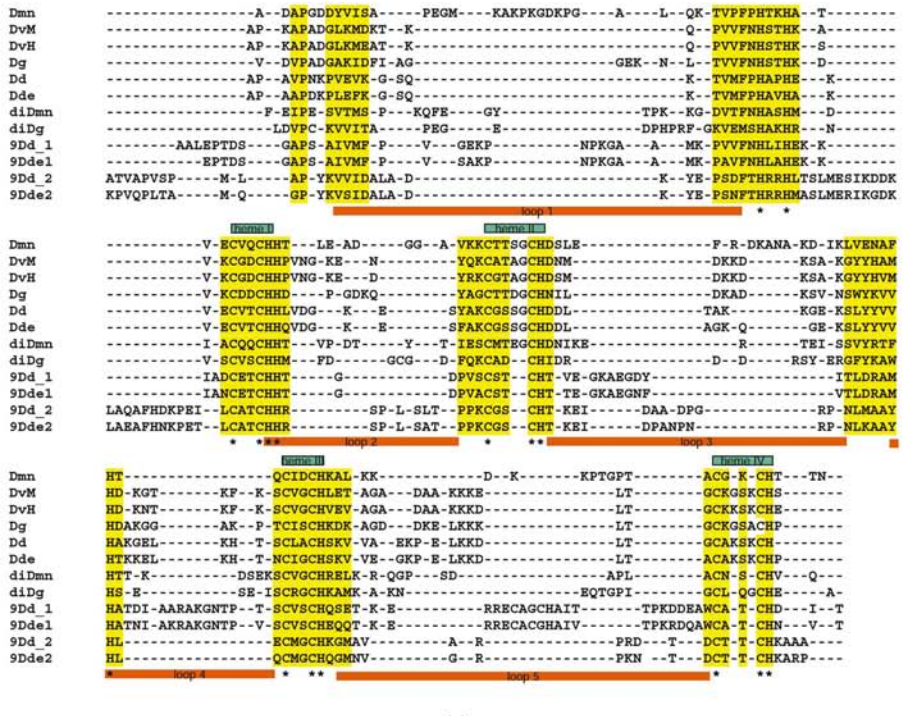
Figure 4
Three-dimensional phyletic relationships among single c_3 domains. The three-dimensional sequence alignment of c_3 domains was used by *CLUSTALX* (Thompson *et al.*, 1997) to produce a distance matrix represented by *NJPLOT* (Perriere & Gouy, 1996). Monomeric cytochromes c_3 (DvM, DvH, Dg, Dd and DdE) that are believed to be electron partners of hydrogenase cluster together, as well as the two pairs of domains from the nine-heme cytochrome of *D. desulfuricans* (9Dd_1 and 9DdE1 or 9Dd_2 and 9DdE2). Strikingly, the two dimeric cytochrome c_3 structures diDmn and diDg fall clearly apart; this is either a consequence of the distance between the genera *Desulfovibrio* and *Desulfomicrobium* or may possibly arise from their different physiological functions.

In the ensemble of 14 structures under analysis there are two in which all of the heme-binding residues do not overlap. The histidine sixth iron ligand of heme III in Daa (Norager *et al.*, 1999) and the binding motif Cys-Asp-Ala-Cys-His of heme I in Dsa do not align with the corresponding heme-binding residues of the remaining structures. These two proteins are known to differ from the rest of the family, as Daa has a different physiological function (Norager *et al.*, 1999) and Dsa is a three-heme cytochrome where heme II is not present (Czjzek *et al.*, 2001). The alignment of the remaining 12 structures is presented in Fig. 5(*a*) and shows that hemes I and III are linked to the polypeptide by the binding motif C*XX*CH, which is conserved in all known structures. Heme II is linked by two types of binding motifs, either C*XX*CH in diDg, in the multidomain cytochromes c_3 and in Daa or C*XXXX*CH in monomeric c_3 cytochromes and in diDmn. The second cysteine of this heme II binding motif C*XXXX*CH is always outside the allowed regions of the Ramachadran plot. This observation can be correlated (Simoes *et al.*, 1998) with the formation of a conserved hydrogen bond between the amino group of this Cys and the carbonyl O atom of the second *X* amino acid of the bridging motif (bonding distances in the seven structures are in the range 2.74–2.90 Å). In the case of heme IV, not only are the above-mentioned types of binding motifs C*XX*CH and C*XXXX*CH present, but also the rather unusual C*XXX*CH motif only found so far in diDg. In the cases where the heme IV binding motif is of the type C*XXXX*CH, no such hydrogen bond is observed and thus all residues in this motif are in allowed regions of the Ramachandran plot.

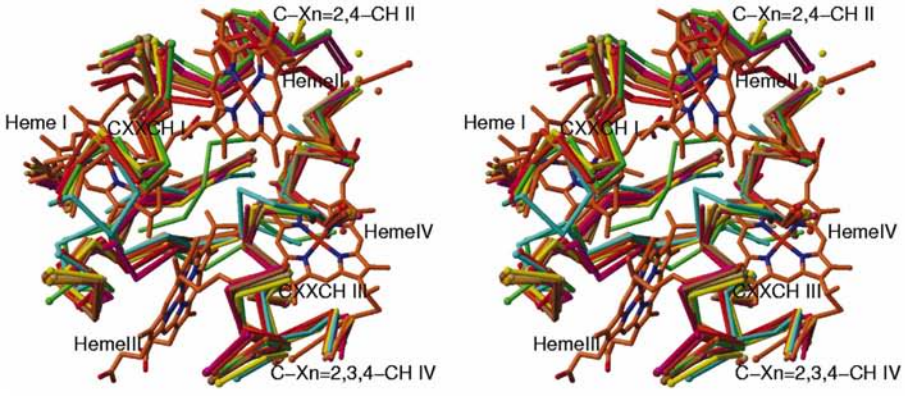
According to Higuchi, the cytochrome c_3 fold is built up of a set of packed hemes wrapped by the peptide chain, described as five variable loops interleaved by four inner-core peptide segments and the two chain termini (Higuchi, Kusunoki *et al.*,

1981). The three-dimensional sequence alignment resulting from the global comparison of c_3 structures (see Figs. 5*a* and 5*b*) shows the family's structurally conserved regions, SCR, where C^{α} atoms consistently superimpose, interleaved by structurally variable regions that correspond essentially to the five variable loops described by Higuchi, Kusunoki *et al.* (1981). A new loop is visible between loops 1 and 2 of Higuchi's description, corresponding to the insertion found in the second domain of the nine-heme cytochromes 9Dd and 9DdE, which contains the histidine sixth iron ligand of the extra heme group found between the two c_3 domains in these multi-heme cytochromes (see Fig. 5*a*).

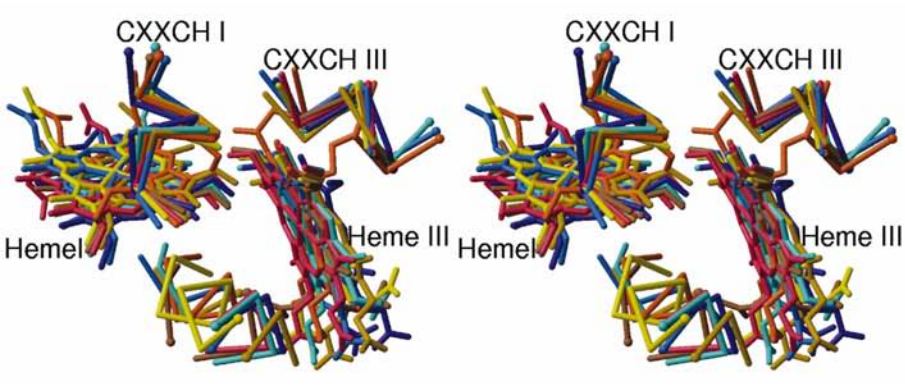
The few secondary-structure elements consistently found in the family are included within the set of SCRs. In particular, the two chains of the β-hairpin motif belong to the first two SCRs and the helical motif with a kink in its middle is found in the SCRs



(*a*)



(*b*)



(*c*)

Figure 5
(*a*) Sequence alignment of c_3 structures. The structurally invariant regions (yellow), involving almost 60 residues where three-dimensionally homologous C^{α} atoms overlap, consist of a segment near the N-terminal region and zones around the 16 heme-binding side chains (marked below the sequence alignment with an asterisk). The heme-binding motifs are mapped above the sequence alignment in green and the variable loops (as defined by Higuchi, Kusunoki *et al.*, 1981) are mapped below in orange, interleaving the conserved regions. (*b*) Stereo representation of c_3 structurally conserved regions, showing the structural fit of C^{α} traces in the regions conserved among c_3 domains. Additionally, the heme groups of diDg are represented as sticks. The C^{α} traces of Daa and Dsa (not shown in the sequence alignments) are represented as green and cyan traces, respectively. Spheres correspond to the C-termini of each region. (*c*) Stereo representation of the di-heme elbow motif. A C^{α}-atom superposition of C*XX*CH motifs, linked to hemes I and III of diDg (coral), with those from di-heme elbow sets of hydroxylamine oxidoreductase (hemes II/III, V/VI and VII/VIII in light, medium and dark blue, respectively; Igarashi *et al.*, 1997), cytochrome c_{554} (hemes III/IV, pink; Iverson *et al.*, 2001), cytochrome *c* nitrite reductase (hemes II/III and IV/V in light and dark brown, respectively; Einsle *et al.*, 1999) and flavocytochrome fumarate reductase (hemes III/IV in yellow; Taylor *et al.*, 1999) shows this heme arrangement to be widespread among multiheme protein families. While the C*XX*CH motif is well conserved, the peptide segments around the other axial His are not conserved and could not be used for fitting purposes.

flanking loop 4, which corresponds to the insertion between the two helical segments. It is interesting to note that the only absolutely conserved residues are the cytochrome-linking residues (cysteines and histidines) and a proline in the N-terminal region.

The three-dimensional arrangement of hemes I and III and their covalently bound residues is also found in several other multiheme protein families, *e.g.* hydroxylamine oxidoreductase (Igarashi *et al.*, 1997), cytochrome c_{554} (Iverson *et al.*, 1998), cytochrome *c* nitrite reductase (Einsle *et al.*, 1999) and flavocytochrome fumarate reductase (Taylor *et al.*, 1999; Bamford *et al.*, 1999; see Fig. 5*c*). These findings support the idea that heme–heme interactions of van der Waals and hydrophobic nature contribute to maintaining the cluster geometry and are an important element determining the protein topology.

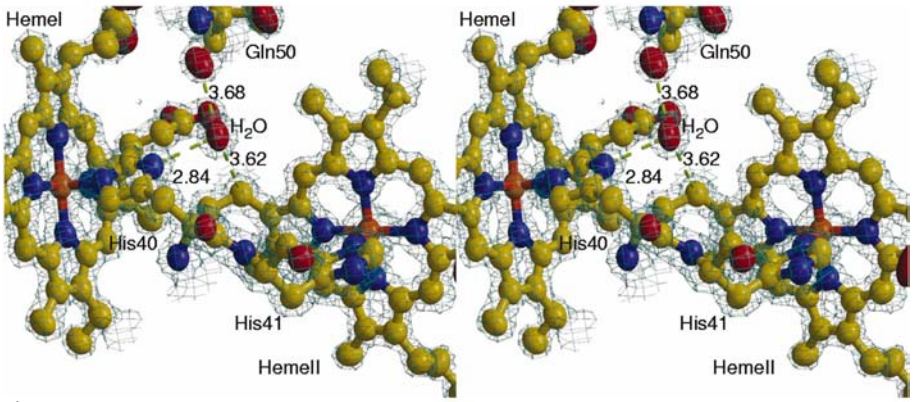


Figure 6
Electron density around the conserved water molecule in the cytochrome c_3 superfamily. σ_A map of diDg showing the conserved water molecule, its neighbouring hemes I and II and the hydrogen bonds to the sixth heme I axial ligand His40 and to the conserved carbonyl O atom of Gln50, as well as the closest contact with heme II *via* the CAB atom. Atoms are represented using 35% probability ellipsoids; the electron-density map is contoured at 3σ (1.5 e Å^{-3}).

In diDg, the Met26 side chain fills the space that is occupied by a phenyl ring in all other structures. This residue is surrounded by heme I, by its fifth axial histidine ligand, and by heme III. These residues are part of the second SCR described above. The otherwise conserved Phe residue has been suggested to play an important role in electron transfer between hemes I and III (Pierrot *et al.*, 1982) of monomeric c_3 cytochromes. Recent mutational studies on DvH indicated that Phe plays an important structural role in the heart of the cytochrome hydrophobic core (Saraiva *et al.*, 1996; Dolla *et al.*, 1999) and that only minor effects could be detected in the redox properties of the studied mutants. Additionally, a role in the variation of a cooperativity effect between native and mutated DvH with hydrogenase cannot be excluded (Dolla *et al.*, 1999).

A comparison of crystal structures from monomeric c_3 cytochromes found two conserved solvent water molecules that form hydrogen bonds with the axially ligated histidines at each side of heme I (Morais *et al.*, 1995). Using the present wider set of crystal structures of c_3 domains, it is found that one of these water molecules appears (Matias *et al.*, 1993) consistently, with the exception of Dsa, in these structures between heme I and heme II. The absence of this water in Dsa is not surprising since this cytochrome lacks heme II.

Fig. 6 represents the environment of this water molecule in diDg cytochrome. In all cases it is hydrogen bonded to ND1 of the sixth axial histidine ligand of heme I and to the structurally conserved carbonyl O atom of a residue of the SCR containing the heme II binding motif; Gln50 in the case of diDg. Additionally, it is also in close contact with heme II atoms (within 3.38–3.73 Å for the present set of structures). This finding is consistent with an earlier proposal considering this internal water molecule to be part of the electron-transfer pathway between hemes I and II (Matias *et al.*, 1993).

Thanks are due to Fundação para a Ciência e a Tecnologia (PRAXIS/PCNA/C/BIO/0089/96 and PRAXIS/P/BIO/12080/1998). Access to the EMBL Hamburg facility was supported by the EU programme 'Access to Research Infrastructure Action of the Improving Human Potential Programme' (HPRI-1999-CT-00017). Figures were produced using *Xtal-View* (McRee, 1992), *Bobscript* (Esnouf, 1997) and *Raster3D* (Merritt & Bacon, 1997).

References

Ambler, R. P. (1980). In *From Cyclotrons to Cytochromes*, edited by A. B. Robinson & N. O. Kaplan. London/New York: Academic Press.
Aubert, C., Brugna, M., Dolla, A., Bruschi, M. & Giudici-Orticoni, M. T. (2000). *Biochim. Biophys. Acta*, **1476**, 85–92.
Bamford, V., Dobbin, P. S., Richardson, D. J. & Hemmings, A. M. (1999). *Nature Struct. Biol.* **6**, 1104–1107.
Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
Blundell, T. L. & Johnson, L. N. (1976). *Protein Crystallography*. London: Academic Press.
Bruschi, M. (1994). *Methods Enzymol.* **243**, 140–155.
Bruschi, M., Hatchikian, E. C., Golovleva, L. A. & LeGall, J. (1977). *J. Bacteriol.* **129**, 30–38.
Bruschi, M., LeGall, J., Hatchikian, C. E. & Dubourdieu, M. (1969). *Bull. Soc. Fr. Physiol. Végét.* **15**, 381–390.
Bruschi, M., Leroy, G., Bonicel, J., Campese, D. & Dolla, A. (1996). *Biochem. J.* **320**, 933–938.
Bruschi, M., Leroy, G., Guerlesquin, F. & Bonicel, J. (1994). *Biochim. Biophys. Acta*, **1205**, 123–131.
Chen, L., Pereira, M. M., Teixeira, M., Xavier, A. V. & Le Gall, J. (1994). *FEBS Lett.* **347**, 295–299.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Czjzek, M., Arnoux, P., Haser, R. & Shepard, W. (2001). *Acta Cryst.* D**57**, 670–678.
Czjzek, M., Guerlesquin, F., Bruschi, M. & Haser, R. (1996). *Structure*, **4**, 395–404.
Deisenhofer, J., Epp, O., Sinning, I. & Michel, H. (1995). *J. Mol. Biol.* **246**, 429–457.

Dolla, A., Arnoux, P., Protasevich, I., Lobachov, V., Brugna, M., Giudici-Orticoni, M. T., Haser, R., Czjzek, M., Makarov, A. & Bruschi, M. (1999). *Biochemistry*, **38**, 33–41.
Einsle, O., Foerster, S., Mann, K., Fritz, G., Messerschmidt, A. & Kroneck, P. M. H. (2001). *Eur. J. Biochem.* **268**, 3028–3035.
Einsle, O., Messerschmidt, A., Stach, P., Bourenkov, G. P., Bartunik, H. D., Huber, R. & Kroneck, P. M. (1999). *Nature (London)*, **400**, 476–480.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 132–134.
Frazão, C., Sieker, L., Sheldrick, G., Lamzin, V., LeGall, J. & Carrondo, M. A. (1999). *J. Biol. Inorg. Chem.* **4**, 162–165.
Haser, R., Pierrot, M., Frey, M., Payan, F., Astier, J. P., Bruschi, M. & Le Gall, J. (1979). *Nature (London)*, **282**, 806–810.
Hatchikian, E. C., Le Gall, J., Bruschi, M. & Dubourdieu, M. (1972). *Biochim. Biophys. Acta*, **258**, 701–708.
Higuchi, Y., Bando, S., Kusunoki, M., Matsuura, Y., Yasuoka, N., Kakudo, M., Yamanaka, T., Yagi, T. & Inokuchi, H. (1981). *J. Biochem. (Tokyo)*, **89**, 1659–1662.
Higuchi, Y., Inaka, K., Yasuoka, N. & Yagi, T. (1987). *Biochim. Biophys. Acta*, **911**, 341–348.
Higuchi, Y., Kusunoki, M., Yasuoka, N., Kakudo, M. & Yagi, T. (1981). *J. Biochem. (Tokyo)*, **90**, 1715–1723.
Igarashi, N., Moriyama, H., Fujiwara, T., Fukumori, Y. & Tanaka, N. (1997). *Nature Struct. Biol.* **4**, 276–284.
Iverson, T. M., Arciero, D. M., Hooper, A. B. & Rees, D. C. (2001). *J. Biol. Inorg. Chem.* **6**, 390–397.
Iverson, T. M., Arciero, D. M., Hsu, B. T., Logan, M. S., Hooper, A. B. & Rees, D. C. (1998). *Nature Struct. Biol.* **5**, 1005–1012.
Johnson, M. S., Sali, A. & Blundell, T. L. (1990). *Methods Enzymol.* **183**, 670–690.
Johnson, M. S., Sutcliffe, M. J. & Blundell, T. L. (1990). *J. Mol. Evol.* **30**, 43–59.
Kissinger, C. R., Adman, E. T., Sieker, L. C., Jensen, L. H. & LeGall, J. (1989). *FEBS Lett.* **244**, 447–450.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Loutfi, M., Guerlesquin, F., Bianco, P., Haladjian, J. & Bruschi, M. (1989). *Biochem. Biophys. Res. Commun.* **159**, 670–676.
McRee, D. E. (1992). *J. Mol. Graph.* **10**, 44–46.
Matias, P. M., Coelho, R., Pereira, I. A., Coelho, A. V., Thompson, A. W., Sieker, L. C., Gall, J. L. & Carrondo, M. A. (1999). *Structure Fold Des.* **7**, 119–130.
Matias, P. M., Frazão, C., Morais, J., Coll, M. & Carrondo, M. A. (1993). *J. Mol. Biol.* **234**, 680–699.
Matias, P. M., Morais, J., Coelho, R., Carrondo, M. A., Wilson, K., Dauter, Z. & Sieker, L. (1996). *Protein Sci.* **5**, 1342–1354.
Matias, P. M., Saraiva, L. M., Soares, C. M., Coelho, A. V., LeGall, J. & Carrondo, M. A. (1999). *J. Biol. Inorg. Chem.* **4**, 478–494.
Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.
Morais, J., Palma, P. N., Frazão, C., Caldeira, J., LeGall, J., Moura, I., Moura, J. J. & Carrondo, M. A. (1995). *Biochemistry*, **34**, 12830–12841.
Norager, S., Legrand, P., Pieulle, L., Hatchikian, C. & Roth, M. (1999). *J. Mol. Biol.* **290**, 881–902.
Ogata, M., Kiuchi, N. & Yagi, T. (1993). *Biochimie*, **75**, 977–983.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Perriere, G. & Gouy, M. (1996). *Biochimie*, **78**, 364–369.
Pierrot, M., Haser, R., Frey, M., Payan, F. & Astier, J. P. (1982). *J. Biol. Chem.* **257**, 14341–14348.
Pieulle, L., Haladjian, J., Bonicel, J. & Hatchikian, E. C. (1996). *Biochim. Biophys. Acta*, **1273**, 51–61.
Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.
Sali, A. & Blundell, T. L. (1993). *J. Mol. Biol.* **234**, 779–815.
Saraiva, L. M., Salgueiro, C. A., LeGall, J., van Dongen, W. & Xavier, A. V. (1996). *J. Biol. Inorg. Chem.* **1**, 542–550.
Schomaker, V. & Trueblood, K. N. (1968). *Acta Cryst.* B**24**, 63–76.
Sheldrick, G. M. (1998). *Direct Methods for Solving Macromolecular Structures*, edited by S. Fortier, pp. 401–411. Dordrecht: Kluwer Academic Publishers.
Sheldrick, G. M. & Gould, R. O. (1995). *Acta Cryst.* B**51**, 423–431.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
Sieker, L. C. (1995). *Colóquio/Ciências, Revista de Cultura Científica*, Vol. 16, pp. 37–48. Lisbon: Fundaçao Calouste Gulbenkian.
Sieker, L. C., Jensen, L. H. & LeGall, J. (1986). *FEBS Lett.* **209**, 261–264.
Simoes, P., Matias, P. M., Morais, J., Wilson, K., Dauter, Z. & Carrondo, M. A. (1998). *Inorg. Chim. Acta*, **273**, 213–224.
Taylor, P., Pealing, S. L., Reid, G. A., Chapman, S. K. & Walkinshaw, M. D. (1999). *Nature Struct. Biol.* **6**, 1108–1112.
Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). *Nucleic Acids Res.* **25**, 4876–4882.
Umhau, S., Fritz, G., Diederichs, K., Breed, J., Welte, W. & Kroneck, P. M. (2001). *Biochemistry*, **40**, 1308–1316.
Uson, I. & Sheldrick, G. M. (1999). *Curr. Opin. Struct. Biol.* **9**, 643–648.
Voordouw, G., Niviere, V., Ferris, F. G., Fedorak, P. M. & Westlake, D. W. S. (1990). *Appl. Environ. Microbiol.* **56**, 3748–3754.