Crystal structure of the dihaem cytochrome c_4 from *Pseudomonas stutzeri* determined at 2.2 Å resolution

Anders Kadziola and Sine Larsen*

Background: Cytochromes c_4 are dihaem cytochromes *c* found in a variety of bacteria. They are assumed to take part in the electron-transport systems associated with both aerobic and anaerobic respiration. The cytochrome c_4 proteins are located in the periplasm, predominantly bound to the inner membrane, and are able to transfer electrons between membrane-bound reduction systems and terminal oxidases. Alignment of cytochrome c_4 sequences from three bacteria, *Pseudomonas aeruginosa*, *Pseudomonas stutzeri* and *Azotobacter vinelandii*, suggests that these dihaem proteins are composed of two similar domains. Two distinctly different redox potentials have been measured for the *Ps. stutzeri* cytochrome c_4, however.

Results: The crystal structure of the dihaem cytochrome c_4 from *Ps. stutzeri* has been determined to 2.2 Å resolution by isomorphous replacement. The model, consisting of two entire cytochrome c_4 molecules and 138 water molecules in the asymmetric unit, was refined to an R value of 20.1% for all observations in the resolution range 8–2.2 Å. The molecule is organized in two cytochrome *c*-like domains that are related by a pseudo-twofold axis. The symmetry is virtually perfectly close to the twofold axis, which passes through a short hydrogen bond between the two haem propionic acid groups, connecting the redox centre of each domain. This haem–haem interaction is further stabilized by an extensive symmetrical hydrogen-bond network. The twofold symmetry is not present further away from the axis, however, and the cytochrome c_4 molecule can be considered to be a dipole with charged residues unevenly distributed between the two domains. The haem environments in the two domains show pronounced differences, mainly on the methionine side of the haem group.

Conclusions: The structure, in conjunction with sequence alignment, suggests that the cytochrome protein has evolved by duplication of a cytochrome *c* gene. The difference in charge distribution around each haem group in the two domains allows the haem group in the N-terminal domain to be associated with the lower redox potential of 241 mV and the C-terminal haem group with the higher potential of 328 mV. The molecular dipole characteristic of cytochrome c_4 is important for its interaction with, and recognition of, its redox partners. In cytochrome c_4, the hydrogen-bond network (between residues that are conserved in all known cytochrome c_4 subspecies) seems to provide an efficient pathway for an intramolecular electron transfer that can ensure cooperativity between the two redox centres. The C-pyrrole corners of the haem edges are potential sites for external electron exchange.

Address: Centre for Crystallographic Studies, Department of Chemistry, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark.

*Corresponding author.
E-mail: sine@xray.ki.ku.dk

Key words: cytochrome c_4, dihaem protein, electron transfer, haem group, molecular dipole, *Pseudomonas stutzeri*

Received: 23 September 1996
Revisions requested: 21 October 1996
Revisions received: 18 November 1996
Accepted: 22 November 1996

Electronic identifier: 0969-2126-005-00203

Structure 15 February 1997, 5:203–216

Introduction

The cytochrome c_4 (cytC4) proteins are dihaem class I cytochromes *c* (molecular mass approximately 20 kDa) found in a variety of bacteria [1,2]. Each haem group is covalently attached to the polypeptide chain via covalent bonds to two cysteines, and the haem iron is axially ligand-bound by a histidine and a methionine residue, forming a hexacoordinated low spin complex in its reduced state.

The cytC4 proteins, from three different bacteria, *Pseudomonas aeruginosa*, *Pseudomonas stutzeri* and *Azotobacter vinelandii*, have been characterized and sequenced [3,4]. Alignment of the three cytC4 sequences has suggested that these dihaem proteins are formed of two similar domains, although the sequence identity between the two domains is less pronounced (around 30%) than the overall sequence identity between different cytC4 proteins (around 80%).

The cytC4 proteins have been found tightly bound to the membrane, and it has been proposed that they play an important role in aerobic respiration of (Gram-negative) bacteria by acting as an electron conductor between oxidases and reductases in the membrane [5,6], although detailed information is not available about their exact function in the cell. Work several years ago [7,8] suggested that cytC4 interacts with an oxidase similar to cytochrome *o*.

In redox proteins with multiple redox centres, the redox centres may act independently or interact with each other to different degrees. The dihaem cytC4 provides a suitable system for studying cooperativity in the electron-transfer processes as each haem group acts as a redox centre. A recent analysis of the kinetics of the oxidation and reduction processes of cytC4 from *Ps. stutzeri* assigned the values of 241 mV and 328 mV for the reduction potentials of the two haem groups [9].

In view of the detailed structural knowledge that is available for the monohaem cytochromes *c*, it is surprising that only very little information was available about the three-dimensional structure of the apparently similar dihaem cytC4. The crystallization of cytC4 from *Ps. aeruginosa*, together with a low-resolution electron-density map based on two heavy-atom derivatives, has been published [10], but no report is given of a refined structure. Recently, the crystal structure of flavocytochrome *c* sulfide dehydrogenase (FCSD) from *Chromatium vinosum* has been determined [11]. This protein is a complex made up of an active flavine-containing subunit and a minor dihaem cytochrome *c* subunit (cytFCSD); the latter is about the size of cytC4 and has about 15% sequence identity to cytC4 [12]. FCSD is able to oxidize sulfide to free sulphur *in vitro*, whereas the apparently periplasmic location of FCSD seems incompatible with this function *in vivo* [13]. The recent structure determinations of cytochrome *c* oxidase [14,15] have similarly provided valuable information about what a potential oxidase partner of cytC4 might look like.

We report here the first three-dimensional structure to high resolution of the cytochrome c_4 from *Ps. stutzeri*, determined by the isomorphous replacement method. This structure determination has enabled us to make comparisons with the structures of monohaem cytochromes *c* and has given insight into the role that cytC4 may play in the electron-transfer processes.

Results and discussion

The crystal structure of cytC4 contains two crystallographically independent molecules, referred to here as A and B. In contrast to the majority of protein structures with more than one molecule in the asymmetric unit, the two molecules of cytC4 are not related to one another by simple point-group symmetry but, as will be discussed below in terms of crystal packing, their environments in the crystal are quite different. Comparison of the two molecules reveals that the regions that display the largest backbone root mean square (rms) deviations are also the more flexible (reflected in larger backbone B factors), especially in the loop and surface regions, where some different sidechain conformations and hydrogen bonds are seen between molecules A and B (Table 1). Apart from these small variations, the two molecules are virtually identical in the overall protein folding and protein–haem interactions.

Overall structure

The 190 amino acid residues of cytC4 from *Ps. stutzeri* are organized into two similar domains, each containing one haem group (Fig. 1). The two domains are related by a pseudo-twofold axis. Close to the axis, where two propionic acid groups (one from each haem group) interact in a highly stabilized hydrogen-bond pattern, the twofold symmetry is perfect. Further away from the pseudo-twofold axis, however, the symmetry gradually breaks down and the molecule can be considered to be a dipole with predominantly negatively charged sidechains (glutamate and aspartate) emerging from the N-terminal domain and predominantly positively charged sidechains (lysine and arginine) emerging from the C-terminal domain.

Secondary structure

Each domain of cytC4 adopts a cytochrome *c*-like fold with a predominantly helical secondary structure. On the basis of manual inspection of the whole structure, we have assigned nine α helices. Four of these are found in the N-terminal and five in the C-terminal domain, as shown by the ribbon representation produced with the program MOLSCRIPT [16] in Figure 1a. This figure shows the folding of cytC4 into two equally sized globular domains connected by an extended polypeptide chain, with one haem group bound to each domain. The interplanar angle between the two haem porphyrin rings is approximately 30°, and the corresponding Fe–Fe distance is 19.1 Å. The structure of cytC4 also reveals a characteristic pseudo-twofold symmetry that relates the two haem groups to one another, and relates four of the helices (A,B,C and E) in the N-terminal domain to the four helices in the C-terminal domain. One helix (D) in the C-terminal domain has no corresponding helix in the N-terminal domain; helix D contains one of the axial haem ligands, Met167. Helices C and E are of comparable size in the two domains, whereas helices A and B differ in size between domains. The fold of cytC4 is very similar to the dihaem cytFCSD, the crystal structure of which was recently reported [11].

The helical arrangement around the haem group in cytC4 is similar to that of other known structures of class I cytochromes *c*. Refined structures have been reported for cytochrome *c* from tuna [17,18], rice [19], yeast [20] and horse heart [21], along with refined structures of cytochrome c_2 from *Rhodospirillum rubrum* [22] and cytochrome *c*-551

Table 1

Hydrogen bonds assigned to the twofold symmetrical arrangement of residues in the centre of cytC4 and intramolecular salt-bridge interactions.*

Donor atom	Position†	Acceptor atom	Position†	Distance (Å)		Comments
				A	B	
Lys31-N	Loop B^NC^N	Gln141-Oε1	Loop B^CC^C	2.74	2.96	Symmetrical H-bond network
Lys137-N	Loop B^CC^C	Gln35-Oε1	Loop B^NC^N	2.82	2.78	-
Gln35-Nε2	Loop B^NC^N	Lys137-O	LoopB^CC^C	3.04	3.03	-
Gln141-Nε2	Loop B^CC^C	Lys31-O	Loop B^NC^N	3.22	3.15	-
Gln35-Nε2	Loop B^NC^N	H_2O-O‡	205/219	2.87	2.89	-
Gln141-Nε2	Loop B^CC^C	H_2O-O‡	206/211	2.99	3.06	-
Tyr39-Oη	Helix C^N	Hem199-O2A		2.99	2.93	-
Tyr145-Oη	Helix C^C	Hem200-O2A		2.89	2.84	-
Lys42-Nζ	Helix C^N	Asp46-Oδ2	Helix C^N	2.72	2.80	-
Lys148-Nζ	Helix C^C	Asp152-Oδ2	Helix C^C	2.84	2.80	-
Lys42-Nζ	Helix C^N	Hem200-O1D		2.95	2.86	-
Lys148-Nζ	Helix C^C	Hem199-O1D		3.08	3.04	-
Gln43-N	Helix C^N	Tyr39-O	Helix C^N	2.86	2.89	-
Gln149-N	Helix C^C	Tyr145-O	Helix C^C	3.06	3.06	-
Gln43-Nε2	Helix C^N	Hem200-O2D		2.97	3.10	-
Gln149-Nε2	Helix C^C	Hem199-O2D		2.92	3.11	-
Asp46-N	Helix C^N	Lys42-O	Helix C^N	3.06	3.02	-
Asp152-N	Helix C^C	Lys148-O	Helix C^C	2.85	2.80	-
Arg61-Nη1	Loop C^NE^N	Gln43-Oε1	Helix C^N	2.84	2.89	-
Arg158-Nη1	Loop $C^CD^{(C)}$	Gln149-Oε1	Helix C^C	2.97	2.76	-
Arg61-Nη2	Loop C^NE^N	Tyr145-Oη	Helix C^C	2.97	2.86	-
Arg158-Nη2	Loop $C^CD^{(C)}$	Tyr39-Oη	Helix C^N	2.98	2.87	-
Arg61-Nη1	Loop C^NE^N	Hem199-O2A		2.92	2.98	-
Arg158-Nη1	Loop $C^CD^{(C)}$	Hem200-O2A		2.99	2.90	-
Arg61-Nη2	Loop C^NE^N	Hem199-O2D		2.87	2.78	-
Arg158-Nη2	Loop $C^CD^{(C)}$	Hem200-O2D		2.90	3.11	-
Hem199-O1A		Hem200-O1A		2.60	2.61	-
H_2O-O‡	205/219	Hem199-O1A		2.84	2.95	-
H_2O-O‡	206/211	Hem200-O1A		2.98	2.59	-
H_2O-O‡	205/219	Lys31-O	Loop B^NC^N	2.72	2.79	-
H_2O-O‡	206/211	Lys137-O	Loop B^CC^C	3.09	3.19	-
Lys10-Nζ	Loop A^NB^N	Asp77-Oδ1	Helix E^N	2.67	2.64	Structural stabilization
Lys10-Nζ	Loop A^NB^N	Asp80-Oδ2	Helix E^N	2.79	2.71	-
Lys111-Nζ	Loop A^CB^C	Asp178-Oδ2	Helix E^C	2.70	2.71	-
Arg168-Nε	Helix $D^{(C)}$	Asp161-Oδ1	Loop $C^CD^{(C)}$	2.81	2.77	-
Lys101B-Nζ	Helix A^C	Glu104B-Oε1	Helix A^C		2.76	Not in A molecule
Arg154B-Nε	Helix C^C	Glu155B-Oε2	Helix C^C		2.93	-
Lys177B-Nζ	Helix E^C	Glu180B-Oε2	Helix E^C		3.15	-

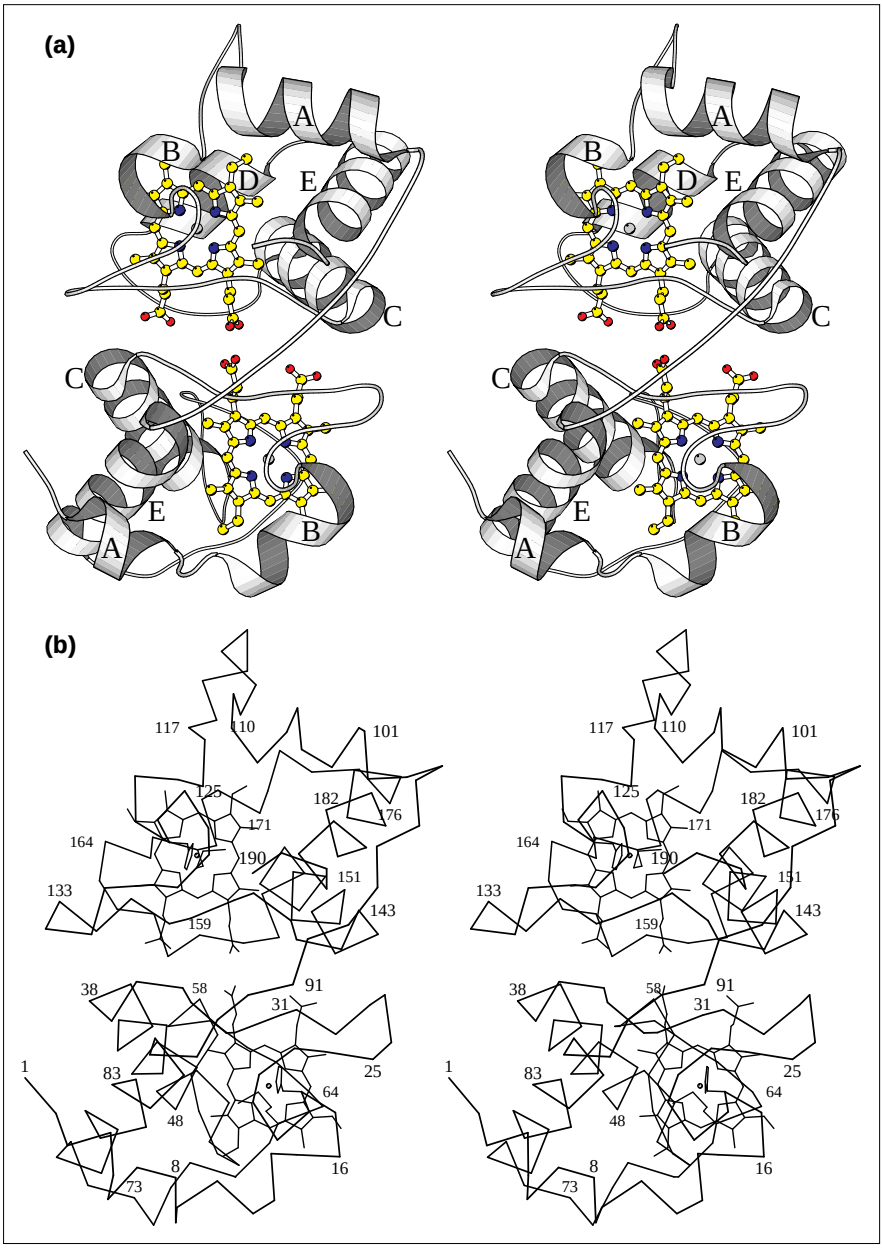
*Residues in bold are conserved throughout the cytochromes c_4 family (see Fig. 3). †Helix labels refer to Figure 1, with superscript letters denoting N- and C-terminal. ‡Numbers for water molecules refer to molecule A and B, respectively.

from *Ps. aeruginosa* [23]. The crystal structures of green alga cytochrome c_6 from *Monoraphidium braunii* [24] and *Chlamydomonas reinhardtii* [25] have also been reported recently, but the coordinates of these structures are not available at the present. The A-, C- and E- helix arrangement found in both domains of cytC4 has a correspondingly similarly arranged counterpart in the structures of cytochrome *c*, c_2 and *c*-551. This similarity includes the packing of helices relative to each other, their relative orientation with respect to the haem group and their sequential order. In all of the structures, the A helices and the E helices are oriented almost at right angles to one another. The A helices are situated on the histidine side of the porphyrin plane, whereas the C helices are typically on the methionine side (sometimes with minor parts on the histidine side) and E helices interconnect the two sides. It should be mentioned that cytochrome *c*-551 also contains a counterpart of helix B in cytC4. At the folding level, we conclude that the main differences between the dihaem class I cytC4 structure and its (distant) relatives, cytochromes *c*, c_2 and *c*-551, is found in the polypeptide segments close to the haem propionic acid groups. The parts of the sequence that bury the haem propionate groups in the monohaem cytochromes are missing or considerably changed in cytC4, creating the domain interface.

Twofold symmetry

The two domains of cytC4 are related by a characteristic pseudo-twofold symmetry, which is centred at a highly

Figure 1



Overall structure of *Ps. stutzeri* cytC4 in a stereo view along the pseudo-twofold axis. **(a)** Ribbon representation with haem groups shown as ball-and-stick models. Helices are assigned as follows. N-terminal domain: Asp3–Gln8, helix A; Val11–His18, helix B (with missing hydrogen bonds: Ala12-O–N-Ala16 and Val13-O–N-Cys17); Gly36–Ser51, helix C; Ser74–Ser87, helix E. C-terminal domain: Leu99–Gly110, helix A; Cys119–His123, helix B; His142–Gly156, helix C; Met165–Ala172, helix D; and Ser175–Gln187, helix E. Atoms are shown in standard colours. **(b)** Sequence-numbered Cα trace with haem groups and sidechains for axial Fe ligands. (Figure produced using the program MOLSCRIPT [16].)

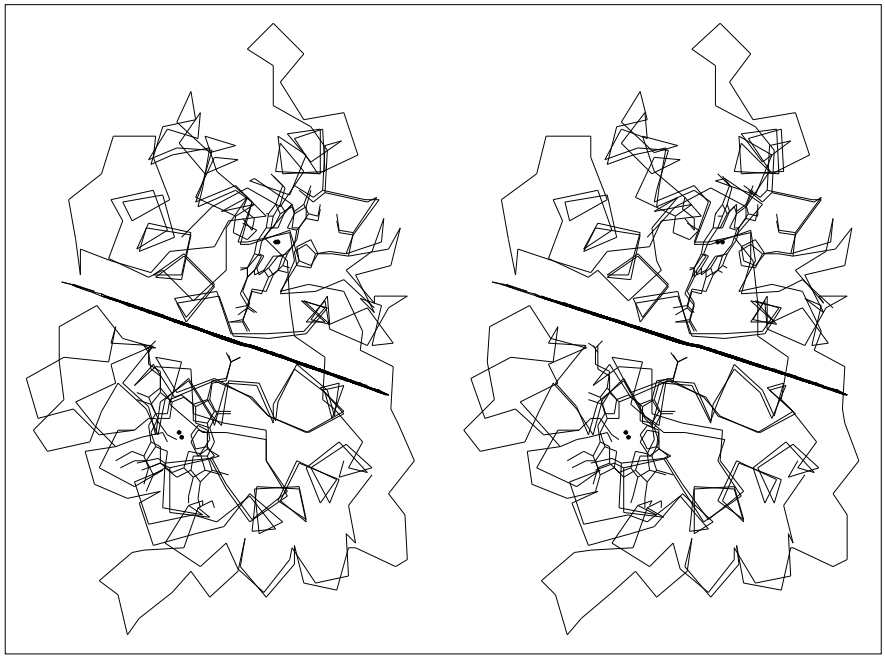
stabilized propionic acid–propionate hydrogen bond. This symmetry is not surprising if one considers the double haem content, with its associated consensus residues and the large degree of identity between the two halves of the sequence.

In order to analyse this symmetry, Cα tracings of two identical molecules were superimposed head-to-tail. This superposition was refined with a least-squares procedure in the graphics display program TURBO-FRODO [26], using all Cα–Cα pairs with distances less than 2 Å. Finally, the pairwise corresponding coordinates were averaged to give points defining the pseudo-twofold axis. The result is shown in Figure 2, in which it can be seen that, apart from diversity of some loops and the extended chain connecting the two domains, the Cα-backbone and especially helices A, B, C and E superimpose quite well. The most flagrant violation of the twofold symmetry is associated with sequences at positions before and after the methionine residues that are axial ligands to the haem irons, whereas the chains on the histidine side show a better match.

The head-to-tail superposition of cytC4 was used to align the sequences of the two domains. In a similar manner, we

Figure 2

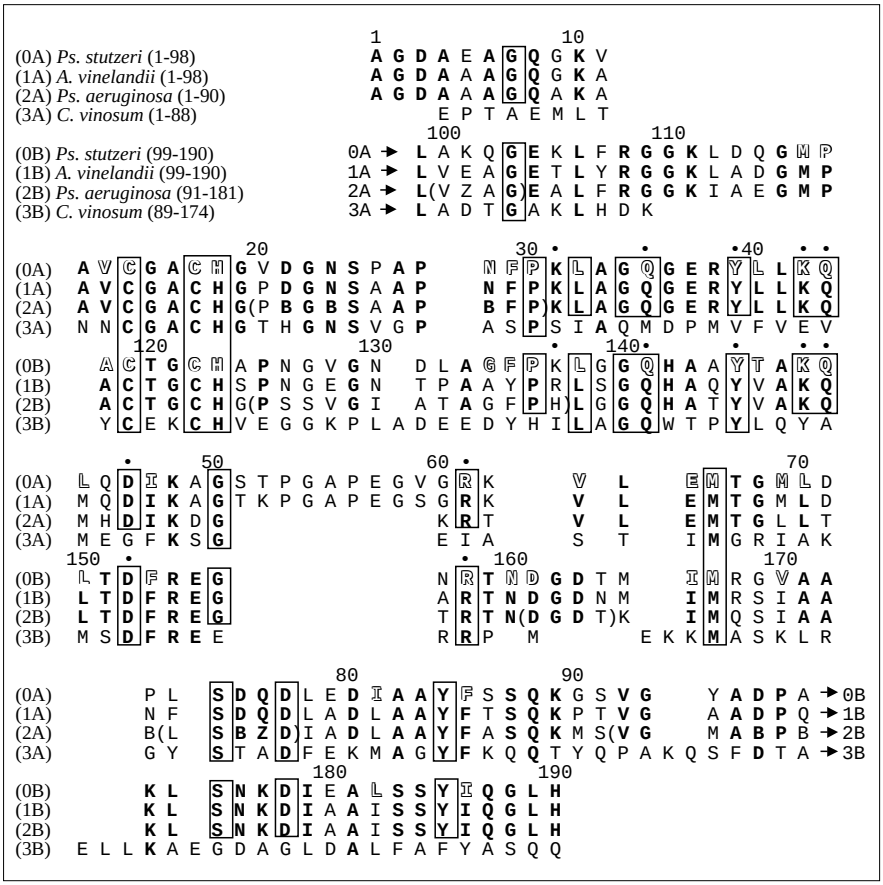
Head-to-tail Cα superposition of two identical cytC4 molecules, including haem groups and axial ligands shown with the pseudo-twofold axis. (Figure produced using the program TURBO-FRODO [26].)



used the superposition of cytC4 with cytFCSD to align their sequences. The results of the alignments are shown in Figure 3. The three known sequences of cytC4, from *Ps. stutzeri*, *A. vinelandii* and *Ps. aeruginosa*, could be aligned easily as they have 123 identical residues, of which 38 are shared with *C. vinosum* cytFCSD according to this alignment. The two domains of cytC4, which have very similar folding, show surprisingly little sequence homology, with only 18 identical residues in the two domains. Here, it should be noted that our interdomain alignment in Figure 3, based on the three-dimensional cytC4 structure, differs considerably from the previous alignments that were made exclusively from the sequences [2].

Beyond the Cα level of comparison, a list of residues — Cys14/Cys119, Cys17/Cys122, His18/His123, Phe29/Phe135, Pro30/Pro136, Lys31/Lys137 (only backbone), Leu32/Leu138, Gly34/Gly140, Gln35/Gln141, Tyr39/Tyr145, Lys42/Lys148, Gln43/Gln149, Leu44/Leu150, Asp46/Asp152 and Arg61/Arg158 — are structurally conserved (i.e. main and sidechains superpose almost perfectly) when twofold symmetry is applied around the pseudo-twofold axis. The remaining sequentially conserved pairs — Gly7/Gly103, Gly50/Gly156, Ser74/Ser175, Asp77/Asp178 and Tyr84/Tyr185 — are found further away from the pseudo-twofold axis, where this symmetry is less pronounced. The most surprising result of the structural alignment schematized in Figure 3 is the Arg61/Arg158 correspondence, which could not be inferred from previous sequence alignments of cytochromes c_4 [2], probably because of the poor interdomain sequence homology in this region. The interdomain structurally conserved residues can be classified into three groups: the well known cytochrome *c* fingerprint — Cys14/Cys119, Cys17/Cys122 and His18/His123; nonpolar residues, the sidechains of which interact with the haem groups — Phe29/Phe135, Pro30/Pro136, Leu32/Leu138 situated on the histidine side and Leu44/Leu150 in the bottom of the haem cleft; and residues involved in the symmetrical hydrogen-bond network, associated with the two propionic acid groups connected by a hydrogen bond, the midpoint of which contains the pseudo-twofold axis — Lys31/Lys137 (only backbone), Gln35/Gln141, Tyr39/Tyr145, Lys42/Lys148, Gln43/Gln149, Asp46/Asp152 and Arg61/Arg158. Two of the nonpolar haem-interacting residues, Pro30/Pro136 and Leu32/Leu138, are structurally conserved in cytochrome *c* and c_2, whereas only the proline residue is present in cytochrome c_5 [27] and *c*-551. In cytFCSD (Fig. 3), Pro30 is conserved in the N-terminal domain, whereas Leu32 is conserved in the C-terminal. In the structures that contain this conserved proline residue, the proline oxygen atom accepts a proton from Nδ of the axial histidine forming a hydrogen bond that stabilizes the orientation of this histidine as a ligand [17]. Pro30 may also play a role in making the axial histidine more nucleophilic, thereby lowering the redox potential [28]. In conclusion, the part of the structure comprising the histidine side of the porphyrin-ring plane seems to be the best-conserved region throughout class I cytochromes *c*. The pseudo-twofold symmetry of cytC4, which is almost perfectly close to the symmetry axis, suggests that this protein has originated as a result of gene duplication. This duplication of a common ancestral gene

Figure 3



Interdomain alignment of the three known cytC4 sequences (*Ps. stutzeri*, *A. vinelandii*, *Ps. aeruginosa*) with cytFCSD (dihaem subunit of flavocytochrome *c* from *C. vinosum*) [3,4,12,13]. The interdomain alignment is based on the *Ps. stutzeri* cytC4 structure (this paper), which in turn is used together with the cytFCSD structure [11] to align the cytC4 sequences to cytFCSD. It should be noted that the interdomain alignment of cytC4 is substantially different from previous alignments (e.g. [2]). Conserved residues throughout cytC4 are in bold and additional interdomain identities are framed. Where appropriate, this notation is extended to cytFCSD. An additional coding is applied to *Ps. stutzeri* cytC4: haem-interacting residues (residues with an atom closer than 4.5 Å to a haem group; see Figs 4,6) are highlighted and residues participating in the twofold symmetrical hydrogen-bond network (Fig. 4) are marked with a dot.

has been followed by evolution to give the two haem-group redox centres different properties. The evolution has mainly resulted in changes of the regions around the axial methionines and the peripheral-loop regions.

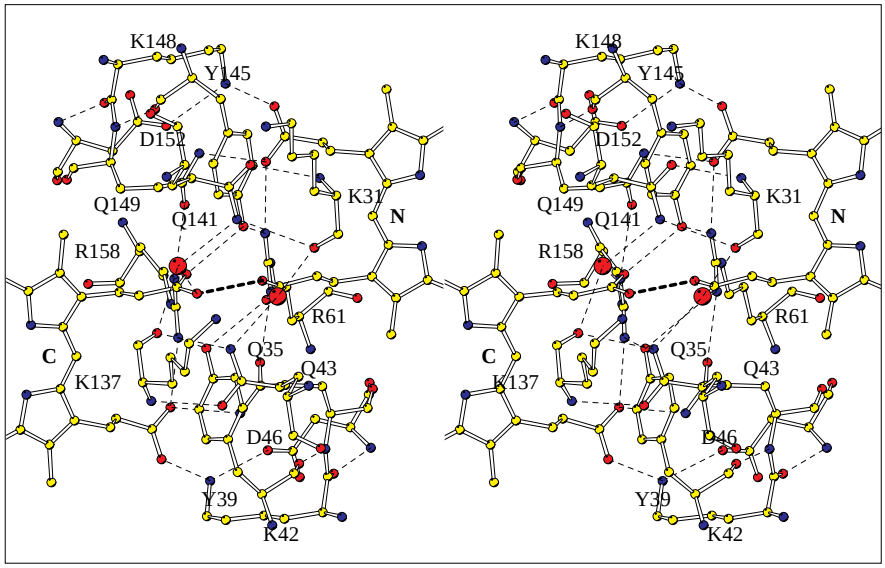
A twofold symmetrical hydrogen-bond network

The intramolecular pseudo-twofold axis passes through the centre of the hydrogen bond connecting the haem groups from each domain. Their interaction through the carboxy groups, with a corresponding O–O distance of 2.6 Å (Table 1), implies that at least one of the propionate groups must be protonated. As mentioned above, the residues Lys31/Lys137 (only backbone), Gln35/Gln141, Tyr39/Tyr145, Lys42/Lys148, Gln43/Gln149, Asp46/Asp152 and Arg61/Arg158 are also part in a hydrogen-bonded network, which includes the haem propionic acid groups and two water molecules. The residues involved in this network, which are all conserved in cytochromes c_4, are marked with dots in Figure 3. Within experimental errors, this arrangement has twofold symmetry.

We have used common knowledge about hydrogen-bond donors, acceptors and their corresponding donor–acceptor distances/angles to postulate a plausible hydrogen-bond pattern. The symmetrical network is illustrated as a stereo pair in Figure 4, and Table 1 contains its relevant interatomic distances. The assignment is as follows; Lys42/Lys148 is hydrogen bonded to the opposite domain haem carboxylate group, not taking part in the haem–haem interaction, and to Asp46/Asp152. Gln43/Gln149 plays a similar role by serving as a hydrogen-bond donor to the other oxygen atom of the propionate group as well as accepting a proton from Arg61/Arg158. Arg61/Arg158 makes hydrogen bonds with both haem carboxylate groups and donates a proton to the opposite domain Tyr145/Tyr39 Oη. Backbone atoms (N and O) of Lys31/Lys137 form a cyclic hydrogen-bonded system with the opposite domain sidechain atoms (Oε1 and Nε2) of Gln141/Gln35. Two equivalent water molecules serve as hydrogen-bond acceptors for the other Nε2 hydrogen atom of Gln35/Gln141. These water molecules donate their protons to the backbone O atom of Lys31/Lys137 and to the interacting haem carboxylate groups.

Different models for electron transport have been proposed. All models agree that the electron-transport rate is dependent upon three factors: the change in free energy

Figure 4



Twofold symmetrical hydrogen-bond network at the propionic acid–propionate interface in the centre of the cytC4 molecule. Residues involved in this network are marked with a dot in Figure 3 and hydrogen-bond distances are listed in Table 1; atoms are shown in standard colours. (Figure produced using the program MOLSCRIPT [16].)

(the difference between reactant and product state); the reorganization energy (the energy required to distort the nuclear configuration of product state into the reactant state without electron transfer), as proposed by Marcus and Sutin [29]; and an electronic coupling (between reactant and product wavefunctions) through the intervening medium between donor and acceptor. The models disagree on how the electronic coupling through the medium takes place. One proposal is that the protein-mediated electronic coupling depends only on the distance between donor and acceptor and the nature of the intervening medium, regardless of its structure (i.e. an organic glass) [30]. An alternative structure-dependent model is proposed in the 'tunneling-pathway' model [31], in which contributions from σ bonds, hydrogen bonds and through-space jumps are considered for a particular pathway in order to account for electron transport rates associated with the pathway. The latter hypothesis has been experimentally confirmed for cytochrome *c* and azurin by measuring rates along presumed predominating electron transport pathways when site-directed mutations have been made to components of each pathway [32,33]. We have therefore chosen to interpret the structure of cytC4 in terms of this hypothesis.

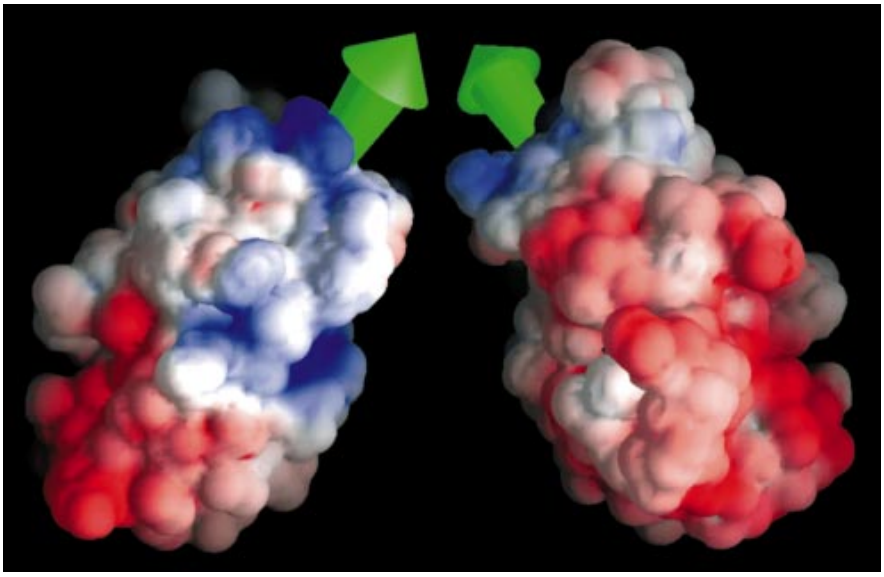
The symmetrical complex hydrogen-bond network appears to be designed to assist intramolecular electron transport. In support of this hypothesis, the attention is drawn to several nonpolar haem-interacting sidechains that are in connection with the network: Pro30/Pro136 and Leu32/Leu138 through the backbone of Lys31/Lys137; Leu44/Leu150 and Ile47/Phe153 through Gln43/Gln149 and Asp46/Asp152; and Leu40/Thr146 through Tyr39/Tyr145. The twofold symmetry will provide two equivalent electron-transport pathways between the haem groups, which will ensure a constructive interference of the electron waves as a supplement to a direct transfer through the interpropionate hydrogen bond. A linear proton displacement in this hydrogen bond (heavy dashed line in Fig. 4), which appears to fulfil the criteria of a low barrier hydrogen bond [34], seems energetically favourable and may facilitate electron transfer between haem groups.

The twofold symmetrical hydrogen-bond arrangement seems to be unique for cytochromes c_4, as only half of the innermost residues that define this network — Gln141, Tyr145, Asp152 and Arg158 — are structurally conserved in the C-terminal domain of cytFCSD.

Salt bridges and free-surface charges

The content of charged residues (asparagine, glutamine, lysine and arginine) in *Ps. stutzeri* cytC4 is similar to that of other cytochrome proteins. Between the two domains, however, the charges are unevenly distributed. The N-terminal domain (residues 1–95) contains 12 aspartate or lutamate residues (of which 8 are conserved among cytochromes c_4) and 8(7) lysine or arginine residues, whereas the C-terminal domain (residues 96–190) contains 10(6) aspartate or glutamate and 11(7) lysine or arginines. Compared in this way, the difference in number of charged residues between domains does not seem that flagrant. A better overview is obtained in Figure 5, in which the colour-coded electrostatic potential [35] is mapped at the surface of the molecule. This figure clearly shows that the negatively charged residues tend to cluster in the N-terminal domain, furthest away from the pseudo-twofold axis. At the opposite end of the

Figure 5



The electrostatic potential mapped onto the molecular surface with direction of the dipole moment shown (green arrows). Blue represents positive charge distribution and red negative charge. The view is along the pseudo-twofold axis: front view (left, as Fig. 1) and back view (right). (Figure produced using the program GRASP [35].)

molecule, positively charged residues cluster in the extreme end of the C-terminal domain, giving rise to an overall dipole moment of the molecule. The negatively and positively charged residues located at the extremes of the molecule comprise Asp3, Glu5, Asp75, Asp77, Glu79 and Asp80 (partly balanced by Lys10 and Lys48) in the N-terminal domain, and Lys101, Lys105, Arg108, Lys111, Lys173, Lys177 (partly balanced by Glu104, Asp113 and Asp178) in the C-terminal domain. The interdomain region surrounding the propionate interface contains a belt of positively charged residues — Lys31, Arg38, Lys42, Arg61, Lys62, Lys89, Lys137, Lys148 and Arg158 — which is partly balanced by Asp46, Glu57, Glu65, Asp131 and Asp152. One would expect the negatively charged propionate groups to be stabilized by the surrounding excess of positive charge, which would result in an overall effect of raising both redox potentials, consistent with the mainly nonpolar porphyrin protein interactions [36].

The salt bridges between charged sidechains and haem propionate groups are listed in Table 1. Because of the different packing environments of the two molecules in the asymmetric unit, the B molecule has three additional intramolecular salt-bridge interactions, which are not present for molecule A and involve surface residues. The majority of the 12 intramolecular salt-bridge interactions common to both molecules are between residues conserved among cytochromes c_4.

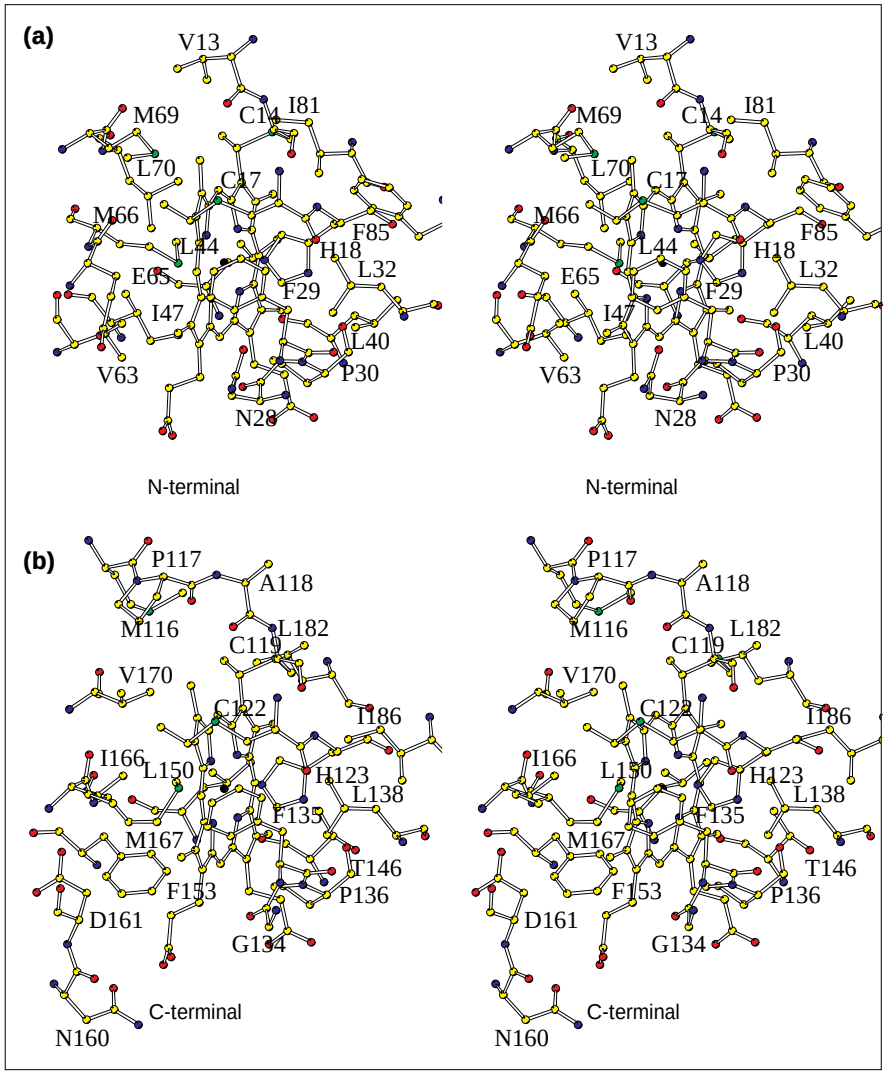
Environment of the haem groups

Spectroscopic characterization of the purified cytC4 used for crystallization showed that it was the reduced form, but the protein is known to be readily oxidized in solution. As no reducing agent was used for crystallization, we believe that the haem groups in the crystal structure contain Fe(III). Previous structural investigations of oxidized and reduced cytochrome *c* have shown that the changes in oxidation state are accompanied by only small structural changes [18]. The following overall description of the haem environment is, therefore, assumed to be valid for cytC4 in its different oxidation states.

Although iron atoms in the two domains have the same ligands, His18/His123 and Met66/Met167, there are many noteworthy differences between the haem environments. These variations do not lead to significant variations in the coordinating distances, however. The four Fe–His-Nε2 distances are in the range 1.9 to 2.0 Å, and the four Fe–Met-Sδ distances are between 2.2 and 2.4 Å. Likewise, the ligands have similar low B values in the two domains. In order to study the haem environments, we have examined the structure for non-bonding contacts between the haem groups and the rest of the protein. Residues were selected that had at least one atom closer than 4.5 Å to any atom of the haem groups. These residues are highlighted in Figure 3. Among these, residues which participate in the symmetrical hydrogen-bond network (marked with a dot in Fig. 3) have already been described. The remaining residues, which only interact with the haem group in their respective domains, are shown in Figure 6. As expected for a hydrophobic haem pocket, these residues are all nonpolar apart from Asn28 and Glu65 in the N-terminal domain and Thr146, Asn160 and Asp161 in the C-terminal domain. The sidechains of Asn28 and Asp161 are pointing away from the haem group, however, the Asn160 sidechain is close to the haem propionate and Thr146 interacts with the porphyrin

Figure 6

Haem groups with neighbouring residues (residues with an atom closer than 4.5 Å to either of the haem groups) not involved in the symmetrical hydrogen-bond network (see Fig. 4). (a) N-terminal haem group. (b) C-terminal haem group. Atoms are shown in standard colours. (Figure produced using the program MOLSCRIPT [16].)



ring through its Cγ atom. More unusual is the presence of the Glu65 carboxylate group close to the N-terminal haem porphyrin. The shortest interaction here is 3.5 Å between Glu65-Cδ and the methyl carbon attached to pyrrole ring D. The carboxylate group of Glu65 does not take part in any hydrogen bond with the rest of the molecule but appears to be stabilized by the accessible solvent. Another unusual interaction takes place between Val13-O and the methyl carbon atom at pyrrole ring C in the N-terminal haem group; these atoms are only 3.0 Å apart. The corresponding distance in the C-terminal domain between Ala118-O and the methyl carbon atom of the haem group is 3.5 Å. Consistent with these observations, the carbonyl group of Val13 is not involved in any classical hydrogen bond, whereas the carbonyl group of Ala118 is hydrogen bonded to a water molecule that forms a bridge between Gly121-N and Ala118-O. This suggests that the haem group and the carbonyl group of Val13 are interacting through a C-H–O hydrogen bond [37]. It is increasingly apparent that this type of interaction is important in biological systems [38].

The analysis of the pseudo-twofold symmetry, which relates the two domains, clearly showed that it is well preserved on the histidine side of the porphyrin ring system. Figure 6 illustrates the previously mentioned twofold-related pairs of residues in the two domains — Phe29/Phe135, Pro30/Pro136, Leu32/Leu138 and Leu44/Leu150. However, an interesting difference is found in the sequence segments following His18 and His123: His18 is followed by Gly–Val–Asp, whereas Ala–Pro–Asn follows His123. Ala124 is the only residue found in the disallowed region of the Ramachandran plot and is adjacent to the less flexible Pro125 residue. This suggests that His123 is in a more strained conformation than His18. A more detailed inspection of the coordination

for the two histidine residues seems to reveal small systematic variations, although experimental accuracy limits our certainty in these observations. Relative to His18, the C-terminal His123 imidazole ring is moved slightly away from a completely axial position; this means that its Nε2 lone pair does not point directly towards Fe. Both coordinating histidine residues are hydrogen bonded to conserved prolines (Pro30 and Pro136), but in the C-terminal domain, the hydrogen bond between His123-Nδ1 and Pro136-O appears to be slightly longer than the equivalent hydrogen bond in the N-terminal domain. This may have the effect of making the imidazole ring system less nucleophilic.

More pronounced differences are observed with respect to the coordinating methionine. The regions around the two methionine residues Met66 and Met167 are not related by the pseudo-twofold symmetry. In the C-terminal domain, Met167 is found in the D helix, which does not have an analogue in the N-terminal domain. Here, the coordinating Met66 is the first residue in a β turn. The stereo pairs in Figure 6 display some of the differences between the two domains on the methionine side of the porphyrin-ring system. Most noteworthy is the difference in conformation of the methionine sidechains and the different positions of the Cα backbone of these residues. Despite the differences on the methionine side and the extra loop in the C-terminal domain, with the haem-interacting residues Met116 and Pro117 (Figs 3,6), the solvent-accessible haem areas calculated to be 63 Å^2 (N-terminal) and 57 Å^2 (C-terminal) are very similar.

Making the assumption that the difference in redox potentials originates from the difference in charge distributions around the haem groups, we used the structural information to assign the observed redox potentials of 241 and 328 mV to each haem group [9]. The previous description demonstrates clearly that the two haem groups have significantly different environments which, together with the dipole arrangement of charges, account for the observed difference in redox potentials. We noticed that the cytC4 molecule appears as a dipole with excess negative charges in the N-terminal domain and excess positive charges in the C-terminal domain (Fig. 5). This was further substantiated by the close contacts between the carboxylate group of Glu65 and the N-terminal haem group (Fig. 6). The redox potentials of haem groups in proteins are influenced both by the surface charges and by buried charges within the protein [39,40]. For cytC4, both factors would tend to lower the redox potential of the N-terminal haem group. Therefore, we conclude that the lower redox potential of 241 mV is associated with the N-terminal haem group and the higher potential of 328 mV with the C-terminal haem group.

Crystal packing

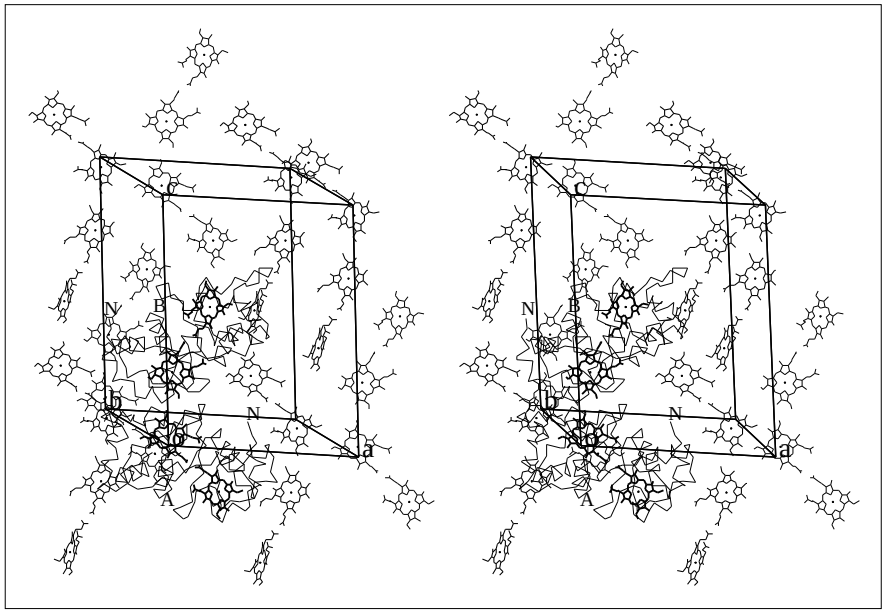
Cytochromes c_4 take part in bacterial electron-transfer processes that must involve interactions between cytC4 and its redox partners through the formation of protein complexes. To ensure optimal and controlled electron transfer, a well defined spatial relationship must exist between the redox centres in these protein complexes. The formation of protein crystals is also based on specific interactions between protein molecules. We therefore examined the crystal packing for interactions between protein molecules, which may mimic the type of interactions that cytC4 can make with its redox partners.

The crystals of *Ps. stutzeri* cytC4 contain two molecules in the asymmetric unit. The two molecules are not related by simple non-crystallographic symmetry elements. The B molecule can be generated from the A molecule by a 148.9° rotation about the unit vector (–0.1765, 0.1730, 0.9690) followed by a (20.544 Å, 22.659 Å, –32.060 Å) translation; the numbers refer to the conventional orthogonalization scheme (first axis parallel to the crystallographic a axis and third axis parallel to the crystallographic c* axis). This is in contrast to the behaviour observed in most crystals containing more than one molecule in the asymmetric unit, in which the molecules are related by non-crystallographic symmetry elements parallel or orthogonal to well defined crystallographic directions [41].

The two molecules in the cytC4 crystals are in a considerably different packing environments. Molecules of type A are connected by salt bridges from Asp3 and Glu5 in the N-terminal domain to Lys101 and Lys105 in the C-terminal domain of a molecule translated along the a axis, and by salt bridges between Arg168 and Glu104 from two molecules related by the crystallographic twofold screw axis. B molecules are not interconnected but interact strongly with A molecules through hydrophobic contacts in the direction of the c axis, as shown in Figure 7. The A and B molecules alternate, such that the C-terminal domain of A pack with the N-terminal domain of the B and, in turn, the C-terminal domain of B pack with the N-terminal domain of A. The strongest intermolecular interaction appears to be between the C-terminal domain of A and the N-terminal domain of B, which brings the respective haem groups very close together in an almost coplanar mode with a corresponding Fe–Fe distance of 16.0 Å. In addition, this intermolecular interaction, defined by interatomic distances less than 4.0 Å, involves residues Pro117, Gly121, Cys122, Asn130, Leu132, Ala133, Phe135, Asp163, Thr164 and Ile166 from the A molecule, and residues Ala12, Val13, Ala16, Cys17, Pro25, Ala26, Phe29, Leu64, Glu65, Thr67 and Met69 from the B molecule. The other intermolecular interaction between the C-terminal domain of the B molecule and the N-terminal domain of the A molecule also brings the respective haem groups into close contact. The Fe–Fe distance is here 16.6 Å, but in this case the interplanar angle between the two haem groups is ~70°. This interaction (interatomic distances less than 4.0 Å) brings the residues Leu112, Gly115, Gly121, Cys122, Asn130,

Figure 7

Packing arrangement in the unit cell, shown by Cα tracings of the two molecules in the asymmetric unit labelled, A and B, with symmetry-related haem groups. Labels N indicate the N termini. Lower case label o denotes the origin, and a, b and c denote the directions of the unit cell axes, respectively.(Figure produced using the program TURBO-FRODO [26].)



Leu132, Ala133 and Phe135 from the B molecule close to the residues Val13, Ala16, Ala26, Pro27, Asn28, Phe29, Glu65 and Met69 from the A molecule.

The packing interactions can be visualized in terms of dipole–dipole interactions in the direction of the a axis and tight hydrophobic van der Waals-type interactions directed along the c axis. As a result of this packing arrangement, chains of interacting haem groups are found parallel to the c axis, as shown in Figure 7. The shortest contact, between the C-terminal haem group of molecule A and the N-terminal haem group of molecule B, where these are coplanar, is 3.9Å between methyl carbon atoms Hem200A-CBC and Hem199B-CMC of the haem groups. In the other contact, between the C-terminal haem group of B and the N-terminal haem group of A translated along the c axis, the shortest distance is 5.2Å between atoms Hem200B-CMC and Hem199A-CBC. Interestingly, the intermolecular hydrophobic contacts involve the exposed surface at the C-pyrrole corner of the haem group. Similar hydrophobic interactions have been observed in the crystals of *C. reinhardtii* cytochrome c_6 [25], where this oligomerization is proposed to have a functional role. In the FCSD structure [11], it is the same region of the dihaem cytochrome *c* subunit that interacts with the flavine-containing subunit of the complex.

Biological implications

Class I cytochrome *c* proteins function as important links between oxidases and reductases at the oxidizing end of the electron-transport system associated with both aerobic and anaerobic respiration of bacteria or of mammalian mitochondria. Cytochromes *c* are located in the periplasm, where they function through interactions with electron-transport chains located in the membrane to deliver electrons to cytoplasmic sites [42]. Cytochromes *c* may act either as diffusible links from reductase to terminal oxidase systems or as an integral part of either of these systems [1]. Soluble cytochrome c_4 (cytC4) isolated from cells of *Pseudomonas aeruginosa*, *Pseudomonas stutzeri* and *Azotobacter vinelandii* was found in the periplasm, but most cytC4 was found in the membrane, presumably bound to the periplasmic face [5,6].

As predicted from previous sequence alignments, the dihaem cytC4 from *Ps. stutzeri* folds into two similar cytochrome *c*-like domains, but on the basis of the three-dimensional structure, substantial adjustments have been made for the interdomain sequence alignment. The protein contains two isolated redox centres connected by a short hydrogen bond between the propionic acid groups of the haem moieties, and a well defined symmetrical system of hydrogen bonds that provide an efficient pathway for intramolecular electron transfer. A unique characteristic of cytC4 is its asymmetric distribution of charges. The N-terminal domain contains an excess of negatively charged residues and the C-terminal domain an equivalent excess of positively charged residues, such that the protein behaves as a molecular dipole. The detailed structural knowledge of the haem–protein interactions, together with the difference in distribution of the surrounding charges, makes it possible to assign the measured redox potentials of 241 and 328 mV to each haem group. The N-terminal haem group, with excess

of negative charges corresponds to the lower reduction potential. The crystal structure of cytC4 allows us to study properties of a protein with multiple redox centres, and it provides a basis for subsequent prediction of possible electron-transfer pathways.

The characteristic molecular dipole of cytochrome c_4 is important for electrostatic interactions, and the subsequent recognition of potential redox partners of cytC4. Although only limited knowledge is available on the specific function of cytC4, it seems plausible that its function in electron-transfer processes resembles that of the monohaem class I cytochromes *c*. Investigations have verified that lysine residues on the surface of cytochrome *c*, especially around the exposed haem edge, are important for cytC4's reaction with its physiological redox partners cytochrome c_1 and cytochrome *c* oxidase [43]. The excess of positively charged residues in the C-terminal domain of cytC4 may enable it to interact with an oxidase of a type similar to cytochrome *c* oxidase [14,15], which is known to contain a pocket with negatively charged residues; simultaneously or sequentially, the N-terminal domain is free to interact with a reductase. The structure of the flavocytochrome *c* sulfide dehydrogenase complex [11] is an example in which the N-terminal domain of the dihaem cytochrome *c* subunit (with a fold very similar to that of cytC4) interacts with the flavine-containing enzymatically active subunit. On the basis of the above example and the strong hydrophobic edge-to-edge haem interactions seen in the crystal packing of cytC4, the C-pyrrole corners of both haem groups seem to be possible sites for external electron exchange to redox partners.

The structure-based sequence homology between the two domains of cytC4 suggests that this protein has evolved by duplication of an ancestral cytochrome *c* gene. The haem–haem propionic acid–propionate hydrogen bond surrounded by a complex twofold symmetrical hydrogen-bond network indicates that cooperativity between two redox centres is an integral part of cytC4 function. The double haem content is suitable to assist two-electron transfer processes, similar to the *in vitro* oxidation of sulfide to free sulphur as carried out by flavocytochrome *c* sulfide dehydrogenase [11].

Materials and methods

Crystals and derivatives

The crystallization of *Ps. stutzeri* cytochrome c_4 by vapour diffusion has previously been described [44] (space group $P2_1$ with cell parameters, a = 49.5 Å, b = 58.6 Å, c = 63.5 Å, β = 97.0° and two molecules in the asymmetric unit). One isomorphous heavy-atom derivative (U) was prepared by soaking with 1 mM $UO_2(NO_3)_2$; 1 μl of a 10 mM $UO_2(NO_3)_2$ solution was added to a 10 μl drop containing crystals and left soaking for one week prior to the data collection.

Data collection

X-ray diffraction data were collected for one native and one derivative ($UO_2(NO_3)_2$) crystal. In both cases, the crystals were cooled to 1°C during data collection with the low-temperature device provided by Molecular Structure Corporation. An R-AXIS II imaging-plate system was used to measure the intensities, employing graphite monochromatized Cu-Kα radiation from a Rigaku Rotaflex RU 200 rotating anode operating at 50 kV and 180 mA. Integrated intensities were obtained with DENZO from the HKL suite [45] and data reduction was performed with the programs ROTAVATA and AGROVATA from the CCP4 suite (except where otherwise noted, names of programs used for computations refer to the CCP4 suite [46]). Additional details about data collection and reduction are given in Table 2 (see also [44]).

Structure determination

The structure determination was achieved by a combination of phase information from the single isomorphous derivative and the anomalous dispersion differences due to Fe in the native data. The positions of the four independent Fe atoms were derived from the native anomalous difference Patterson map as described, [44]. The program SCALEIT was used to scale derivative data to native. A relative overall scale and an isotropic B factor were manually adjusted (over-riding the default option) to obtain unity scale factors. Based on 17 721 common reflections, the fractional isomorphous difference was: R_{iso} = 16.2%. An isomorphous difference Patterson map could be interpreted in terms of two heavy-atom sites, consistent with the anomalous difference Patterson map. Positions of the two U sites were related to the same origin as the four Fe atoms by the use of cross-Fourier syntheses. Two phase sets (one for each possible enantiomorph) were generated with MLPHARE [46,47] using native anomalous data and the known Fe positions. In order to perform these computations, the native dataset was duplicated in order to simulate a native–derivative relationship. In this way, Fe atoms could be treated as heavy-atom sites with zero real occupancy and with an anomalous occupancy different from zero. Coordinates, anomalous occupancies and isotropic B factors were refined and used for phase calculations (4261 anomalous reflections between 10 and 3 Å with a phasing power of 1.37). Derivative–native isomorphous differences could then be phased to give either two peaks (correct enantiomorph) or two holes (wrong enantiomorph) at positions consistent with the isomorphous difference Patterson map.

By comparing the native and derivative anomalous difference Patterson maps, it was clear that scattering contributions from the anomalous Fe atoms should be included together with the anomalous U atoms in order to give a full heavy-atom description of the derivative anomalous signal. The heavy-atom model (see Table 2) was refined and phases calculated with MLPHARE using all data and refining all refinable parameters. This refinement gave an unrealistically high variation of anomalous occupancies for Fe in the range from 0.721 to 1.072 with isotropic B factors between 7.9 and 34.3 $Å^2$. The problem of not being able to constrain the anomalous Fe occupancies to be equal in MLPHARE was circumvented by means of various adjustments in the heavy-atom model, which did not, however, lead to significant improvements in either phasing statistics or maps — so the final heavy-atom model was based on the default MLPHARE run. Phasing statistics are given in Table 2.

Phases were improved with the program SQUASH [48] simultaneously using Sayre's equation, histogram-matching and solvent-flattening options, assuming 40% solvent in the crystal.

Model building and refinement

The map calculated with SQUASH phases was of a quality that allowed an almost complete assignment of the two molecules in the asymmetric unit. First, a BONES representation of the map was calculated with the program O [49]. The molecules could be easily recognized with their characteristic helical structural elements. The four haem groups could be positioned from the known Fe positions and unambiguously oriented in the map based on their covalently bound cysteine sulphur atoms, which appeared clearly in a supplementary experimental map based on anomalous differences in the native data and SQUASH phases shifted

Table 2

Data collection and statistics.

	Native	U derivative*
Diffraction limit (Å)	2.2	2.2
Completeness (%)	98.1 (78.9)†	97.4 (75.4)
Completeness outer shell (%)‡	96.6 (72.8)	96.6 (70.4)
R_{sym} (%)§	3.3	3.1
R_{sym} outer shell (%)	12.1	7.7
I/σ (I)	13.5	12.2
I/σ (I) outer shell	5.7	6.5
Multiplicity	3.8	3.7
Number of sites#	- (4)	2 (6)
Number of phased reflections**	- (9392)	17705 (10616)
R_{cullis}††	- (0.71)	0.83 (0.66)
Phasing power‡‡	- (1.0)	0.9 (1.1)

*1 mM $UO_2(NO_3)_2$. †Quantities in parentheses are for anomalous data. ‡Outer resolution shell = 2.32–2.20 Å. §$R_{sym} = \Sigma|I - \langle I \rangle| / \Sigma I$, where I = observed intensity and $\langle I \rangle$ = average intensity obtained from multiple observations of symmetry-related reflections. #Four anomalous Fe sites and two isomorphous/anomalous U sites. **A total of 18000 reflections were phased with a mean figure of merit of 0.481. ††The original R_{cullis} definition for centric reflections is: $\Sigma||F_{PH} \pm F_P| - |F_H|| / \Sigma|F_{PH} - F_P|$. Here we have used the MLPHARE [46] definition, extended to include acentric reflections and anomalous data. ‡‡Phasing power is defined as $\langle F_H \rangle / \langle$ Lack of closure $\rangle$.

by –90°, as shown in Figure 8. The consensus sequences Cys14–Gly–Ala–Cys–His at the N-terminal haem and Cys119–Thr–Gly–Cys–His at the C-terminal haem could be distinguished in the map. This distinction was confirmed by the presence of the Met116 sulphur atom in the anomalous SQUASH map. From these anchor points, most of the amino acid sequence could be assigned, again using the anomalous SQUASH map as a valuable tool to confirm the methionine positions.

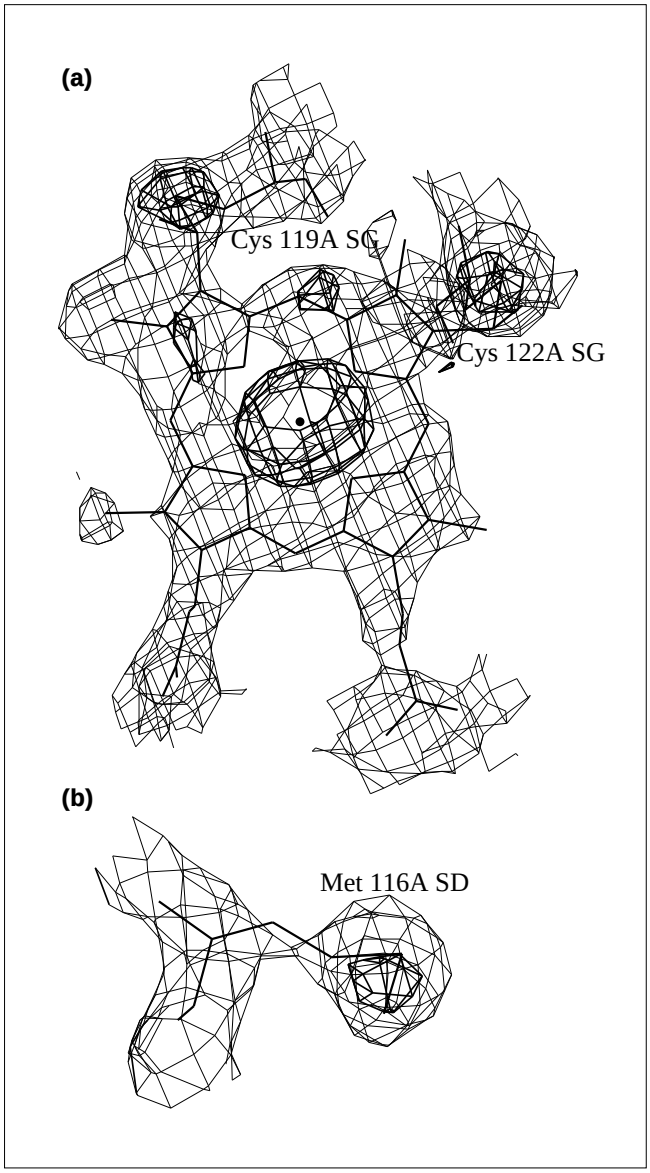
The graphics display program TURBO-FRODO [26] was used for construction and refitting of the model. Missing parts were built into the structure and corrections were made by means of subsequent $2F_{obs}–F_{calc}$ and $F_{obs}–F_{calc}$ maps based on reflections between 8 and 2.2 Å resolution and model phases after refinement. In addition, residual density peaks were modelled as water molecules if they had good hydrogen-bond geometry, their B factors refined to values smaller than 50 $Å^2$ and gave back clear $2F_{obs}–F_{calc}$ density at a 1σ cutoff level.

The model was refined by simulated annealing and neighbour-restrained individual isotropic temperature-factor optimization with X-PLOR [50], using the Engh and Huber [51] structural parameter set. Structural parameters for the haem group with ligands were taken from the 1.5 Å resolution structure of cytochrome *c* [17]. Non-crystallographic symmetry restraints were not applied. The current model contains all 2×190 amino-acid residues with 2×2 haem groups and 138 water molecules.

Quality of the model

The present model was refined using 90% of all observed reflections between 8 and 2.2 Å resolution. The resulting R factor is 20.1% for all (17609) reflections in this resolution range with an R_{free} value of 26.7% for the 10% test set. The rms deviations in bond lengths and bond angles from standard values are 0.011 Å and 1.7°, respectively. A Ramachandran plot calculated with PROCHECK [52] revealed all but one non-glycine residue in the 'most favoured' (91.9%) or 'additional allowed regions' (7.4%). Ala124 is found in the forbidden region with (ϕ,ψ) close to (70°,150°). This residue is situated between the axial haem ligands His123 and Pro125, in a region where the electron density for both molecules is well defined.

Figure 8



Experimental maps after phase improvement with SQUASH [48], superposed on selected parts of the refined structure. Fine line map: best-Fourier synthesis to 2.5 Å resolution with native F_{obs} and SQUASH-phases contoured at a 1σ level. Heavy line map: best-Fourier synthesis to 2.5 Å resolution with native ΔF_{ano} and SQUASH-phases shifted by –90° contoured at a 2.5σ level. (Figure produced using the program TURBO-FRODO [26].)

The geometry of the two molecules in the asymmetric unit is virtually identical. Some differences in sidechains are observed for surface residues due to different packing environments for the two molecules. The rms difference between molecules A and B is 1.00 Å for all atoms and 0.47 Å for backbone atoms only.

Accession numbers

The atomic coordinates and observed structure factor amplitudes for cytochrome c_4 from *Ps. stutzeri* have been deposited at the Brookhaven Protein Data Bank with accession code 1ETP.

Acknowledgements

We thank Jens-Jakob Karlsson and Jens Ulstrup from the Danish Technical University for valuable discussions and for providing purified protein. The support from The Danish National Research Foundation is gratefully acknowledged.

References

1. Pettigrew, G.W. & Moore, G.R. (1987). *Cytochromes c - Biological Aspects*. Springer-Verlag, Berlin Heidelberg.
2. Moore, G.R. & Pettigrew, G.W. (1990). *Cytochromes c - Evolutionary, Structural and Physicochemical Aspects*. Springer-Verlag, Berlin Heidelberg.
3. Ambler, R.P., Daniel, M., Melis, K. & Stout, C.D. (1984). The amino acid sequence of the dihaem cytochrome c_4 from bacterium *Azotobacter vinelandii*. *Biochem. J.* **222**, 217–227.
4. Christensen, H.E.M. (1994). Cloning and characterization of the gene encoding cytochrome c_4 from *Pseudomonas stutzeri*. *Gene* **144**, 139–140.
5. Pettigrew, G.W. & Brown, K.R. (1988). Free and membrane-bound forms of bacterial cytochrome c_4. *Biochem. J.* **252**, 427–435.
6. Hunter, D.J.B., Brown, K.R. & Pettigrew, G.W. (1989). The role of cytochrome c_4 in bacterial respiration. *Biochem J.* **262**, 233–240.
7. Matsushita, K., Shinagawa, E., Adachi, O. & Ameyama, M. (1982). *o*-type cytochrome oxidase in the membrane of aerobically grown *Pseudomonas aeruginosa*. *FEBS Lett.* **139**, 255–258.
8. Yang, T. (1986). Biochemical and biophysical properties of cytochrome *o* of *Azotobacter vinelandii*. *Biochim. Biophys. Acta* **848**, 342–351.
9. Conrad, L.S., Karlsson, J.J. & Ulstrup, J. (1995). Electron transfer and spectral α-band properties of the di-heme protein cytochrome c_4 from *Pseudomonas stutzeri*. *Eur. J. Biochem.* **231**, 133–141.
10. Sawyer, L., Jones, C.L., Damas, A.M., Harding, M.M., Gould, R. & Ambler, R.P. (1981). Cytochrome c_4 from *Pseudomonas aeruginosa*. *J. Mol. Biol.* **153**, 831–835.
11. Chen, Z., *et al.*, & Mathews, F.S. (1994). The structure of flavocytochrome *c* sulfide dehydrogenase from a purple phototropic bacterium. *Science* **266**, 430–432.
12. Van Beeumen, J.J., *et al.*, & Cusanovich, M.A. (1991). Covalent structure of the diheme cytochrome subunit and amino-terminal sequence of the flavoprotein subunit of flavocytochrome *c* from *Chromatium vinosum*. *J. Biol. Chem.* **266**, 12921–12931.
13. Dolota, M.M., Van Beeumen, J.J., Ambler, R.P., Meyer, T.E. & Cusanovich, M.A. (1993). Nucleotide sequence of the heme subunit of flavocytochrome *c* from the purple phototropic bacterium, *Chromatium vinosum*. *J. Biol. Chem.* **268**, 14426–14431.
14. Tsukihara, T., *et al.*, & Yoshikawa, S. (1995). Structures of metal sites of oxidized bovine heart cytochrome *c* oxidase at 2.8 Å. *Science* **269**, 1069–1074.
15. Iwata, S., Ostermeier, C., Ludwig, B. & Michel, H. (1995). Structure at 2.8 Å resolution of cytochrome *c* oxidase from *Paracoccus denitrificans*. *Nature* **376**, 660–669.
16. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946–950.
17. Takano, T. & Dickerson, R.E. (1981). Conformation change of cytochrome *c*: ferrocytochrome *c* structure refined at 1.5 Å resolution. *J. Mol. Biol.* **153**, 79–94.
18. Takano, T. & Dickerson, R.E. (1981). Conformation change of cytochrome *c*: ferrocytochrome *c* refinement at 1.8 Å and comparison with the ferrocytochrome structure. *J. Mol. Biol.* **153**, 95–115.
19. Ochi, H, *et al.*, & Morita, Y. (1983). Structure of rice ferrocytochrome *c* at 2.0 Å resolution. *J. Mol. Biol.* **166**, 407–418.
20. Louie, G.V. & Brayer, G.D. (1990). High resolution refinement of yeast iso-1-cytochrome *c* and comparisons with other eukaryotic cytochromes *c*. *J. Mol. Biol.* **214**, 527–555.
21. Bushnell, G.W., Louie, G.V. & Brayer, G.D. (1990). High resolution three-dimensional structure of horse heart cytochrome *c*. *J. Mol. Biol.* **214**, 585–595.
22. Salemme, F.R., Freer, S.T., Xuong, N.H., Alden, R.A. & Kraut, J. (1973). The structure of oxidized cytochrome c_2 of *Rhodospirillum rubrum*. *J. Biol. Chem.* **248**, 3910–3921.
23. Matsuura, Y., Takano, T. & Dickerson, R.E. (1982). Structure of cytochrome *c*-551 from *Ps. aeruginosa* refined at 1.6 Å resolution and comparison of the two redox forms. *J. Mol. Biol.* **156**, 389–409.
24. Frazão, C., *et al.*, & Sheldrick, G.M. (1995). *Ab initio* determination of the crystal structure of cytochrome c_6 and comparison with plastocyanin. *Structure* **3**, 1159–1169.
25. Kerfeld, C.A., Anwar, H.P., Interrante, R., Merchant, S. & Yeates, T.O. (1995). The structure of chloroplast cytochrome c_6 at 1.9 Å resolution: evidence for functional oligomerization. *J. Mol. Biol.* **250**, 627–647.
26. Roussel, A. & Cambillau, C. (1992). TURBO-FRODO, Biographics, LCCMB, Marseille, France.
27. Carter, D.C., *et al.*, & Stout, C.D. (1985). Crystal structure of *Azotobacter* cytochrome c_5 at 2.5 Å resolution. *J. Mol. Biol.* **184**, 279–295.
28. Valentine, J.S., Sheridan, R.P., Allen, L.C. & Kahn, P.C. (1979). Coupling between oxidation state and hydrogen bond formation in heme proteins. *Proc. Natl. Acad. Sci. USA* **76**, 1009–1013.
29. Marcus, R.A. & Sutin, N. (1985). Electron transfer in chemistry and biology. *Biochim. Biophys. Acta* **811**, 265–322.
30. Moser, C.C., Keske, J.M., Warncke, K., Farid, R.S. & Dutton, P.L. (1992). Nature of biological electron transfer. *Nature* **355**, 796–802.
31. Beratan, D.N., Onuchic, J.N., Winkler, J.R. & Gray, H.B. (1992). Electron-tunneling pathways in proteins. *Science* **258**, 1740–1741.
32. Karpishin, T.B., Grinstaff, M.W., Komar-Panicucci, S., McLendon, G. & Gray, H.B. (1994). Electron transfer in cytochrome *c* depends upon the structure of the intervening medium. *Structure* **2**, 415–422.
33. Farver, O., *et al.*, & Pecht, I. (1996). Structure–function correlation of intramolecular electron transfer in wild type and single-site mutated azurins. *Chem. Phys.* **204**, 271–277.
34. Cleland, W.W. & Kreevoy, M.M. (1994). Low barrier hydrogen bonds and enzyme catalysis. *Science* **264**, 1887–1890.
35. Nicholls, A., Sharp, K.A. & Honig B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296.
36. Kassner, R.J. (1972). Effects of nonpolar environments on the redox potentials of heme complexes. *Proc. Natl. Acad. Sci. USA* **69**, 2263–2267.
37. Desiraju, G.R (1991). The C-H...O hydrogen bond in crystals. *Accounts Chem. Res.* **24**, 290–296.
38. Jeffrey, G.A. & Saenger, W. (1991). *Hydrogen bonding in biological structures*. Springer-Verlag, Hiedelberg, Berlin.
39. Rees, D.C. (1985). Electrostatic influence on energetics of electron transfer reactions. *Proc. Nat. Acad. Sci. USA* **82**, 3082–3085.
40. Rogers, N.K. & Moore, G.R. (1988). On the energetics of conformational changes and pH dependent redox behaviour of electron transfer proteins. *FEBS Lett.* **228**, 69–73.
41. Wang, X. & Janin, J. (1993). Orientation of non-crystallographic symmetry axes in protein crystals. *Acta Cryst. D* **49**, 505–512.
42. Wood, P.M. (1983). Why do c-type cytochromes exist? *FEBS Lett.* **164**, 223–226.
43. Smith, H.T., Ahmed, A.J. & Millett, F. (1981). Electrostatic interaction of cytochrome *c* with cytochrome c_1 and cytochrome oxidase. *J. Biol. Chem.* **256**, 4984–4990.
44. Kadziola, A., Larsen, S., Christensen, H.M., Karlsson, J.-J. & Ulstrup, J. (1995). Crystallization and preliminary crystallographic investigations of cytochrome c_4 from *Pseudomonas stutzeri*. *Acta Cryst. D* **51**, 1071–1073.
45. Gewirth, D. (1994). *The HKL manual: an oscillation data processing suite for macromolecular crystallography*. Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT, USA.
46. Collaborative Computational Project No.4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst. D* **50**, 760–763.
47. Otwinowski, Z. (1991). Maximum likelihood refinement of heavy atom parameters. In *Proceedings of the CCP4 Study Weekend: Isomorphous Replacement and Anomalous Scattering*. (Wolf, W., Evans, P.R. & Leslie A.G.W., eds), pp. 80–86, SERC Daresbury Laboratory, Warrington, UK.
48. Zhang, K.Y.J. (1993). SQUASH: combining constraints for macromolecular phase refinement end extension. *Acta Cryst. D* **49**, 213–222.
49. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst. A* **47**, 110–119.
50. Brünger, A.T. (1991). *X-PLOR, Version 3.1. A System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT, USA.
51. Engh, R.A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Cryst. A* **47**, 392–400.
52. Laskowski, R.A., MacArthur, M.V., Moss, D.S. & Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291.