Characterization of the structure and redox behaviour of cytochrome c_3 from *Desulfovibrio baculatus* by ^{1}H-nuclear-magnetic-resonance spectroscopy

Isabel B. COUTINHO,* David L. TURNER,† Jean LeGALL‡ and António V. XAVIER*§

* Centro de Tecnologia Química e Biológica and Universidade Nova de Lisboa, Rua da Quinta Grande 6, Apartado 127, 2780 Oeiras, Portugal, † Department of Chemistry, University of Southampton, Southampton SO9 5NH, U.K., and ‡ Department of Biochemistry, University of Georgia, Athens, GA 30602, U.S.A.

Complete assignment of the aromatic and haem proton resonances in the cytochromes c_3 isolated from *Desulfovibrio baculatus* strains (Norway 4, DSM 1741) and (DSM 1743) was achieved using one- and two-dimensional ^{1}H n.m.r. Nuclear Overhauser enhancements observed between haem and aromatic resonances and between resonances due to different haems, together with the ring-current contributions to the chemical shifts of haem resonances, support the argument that the haem core architecture is conserved in the various cytochromes c_3, and that the X-ray structure of the *D. baculatus* cytochrome c_3 is erroneous. The relative orientation of the haems for both cytochromes was determined directly from n.m.r. data. The n.m.r. structures have a resolution of ~ 0.25 nm and are found to be in close agreement with the X-ray structure from *D. vulgaris* cytochrome c_3. The proton assignments were used to relate the highest potential to a specific haem in the three-dimensional structure by monitoring the chemical-shift variation of several haem resonances throughout redox titrations followed by ^{1}H n.m.r. The haem with highest redox potential is not the same as that in other cytochromes c_3.

INTRODUCTION

Cytochromes c_3 are tetrahaem proteins (molecular mass approx. 13 kDa, bis-histidine iron co-ordination) of very low redox potential present in all organisms belonging to the genus *Desulfovibrio* [1,2]. The periplasmic localization of cytochrome c_3, together with the observed activation in its presence of electron-transfer reactions catalysed by hydrogenase, points towards a common role of cytochrome c_3 as a specific coupling factor to this enzyme [3]; the cytochromes c_3 from *D. baculatus* (Norway 4 and DSM 1743) can also act as sulphur reductases [4]. The haem core architecture is conserved in all cytochrome c_3 structures available [5–9], presenting the same iron-to-iron distances, and, if the location of the substituent groups is neglected, the same relative haem plane orientations. When the substituents are taken into account, the orientations of the haems in the X-ray structure of cytochrome c_3 from *D. baculatus* (Norway 4) [5,6] are found to be quite different from those in the other X-ray structures, but it has recently been shown, using n.m.r. spectroscopy, that the X-ray structure is in error and that the correct orientation is the same as that found in the other proteins [10]. Thus the basic structure seems to be highly conserved, although the amino acid sequences of the various cytochromes c_3 can be quite different [9], with sequence similarity as low as 25% for the less phylogenetically related organisms *D. baculatus* and *D. vulgaris* [11].

^{1}H n.m.r. has been used extensively to probe the thermodynamic and kinetic, as well as structural, characteristics of the cytochromes c_3, as the four low-spin haems act as intrinsic probes both in the reduced (diamagnetic) and especially in the oxidized (paramagnetic) and intermediate redox states [12–19]. The redox properties have also been studied by other methods [20–22], including e.p.r. [23–25]. It was found that some of the physical properties of the cytochromes from *D. baculatus* differ from those observed for the proteins from *D. vulgaris* or *D. gigas*. In particular, for these last two proteins, the intermolecular electron-exchange rate is low on the n.m.r. time scale [15,16], and the four haems have redox potentials of the same order of magnitude, with a maximum difference of ~ 130 mV between the higher- and lower-potential haems [15,17], while in the *D. baculatus* cytochromes the intermolecular exchange is intermediate to fast [18] and the higher-potential haem has a redox potential considerably higher (more than 150 mV) than the second-higher-potential one [23–25].

The assignment of the redox potential corresponding to each specific haem in the three-dimensional (3D) structure has been obtained for the cytochromes c_3 from *D. vulgaris* (Hildenborough) [26] and from *D. gigas* (the following paper [27]) by using two-dimensional (2D) n.m.r. to monitor the haem methyl resonances, previously assigned in the fully reduced state, throughout the five individual redox stages. This strategy cannot be used for the cytochromes c_3 from *D. baculatus*, because the intermolecular electron transfer is faster on the n.m.r. time scale, so the haem methyl resonances broaden and/or shift smoothly from their

Abbreviations used: c_3E, cytochrome c_3 isolated from *Desulfovibrio baculatus* strain DSM 1743 (previously known as *Desulfovibrio desulfuricans* strain *ethylica*); c_3N, cytochrome c_3 isolated from *Desulfovibrio baculatus* strain DSM 1741 (previously known as *Desulfovibrio desulfuricans* strain Norway 4); TOCSY, total correlation spectroscopy; rmsd, root-mean-square deviation; *meso* protons, thioether protons, haem methyl groups and thioether methyl groups stand for the following haem substituents respectively: H5, H10, H15 and H20; H3^1 and H8^1; M2^1, M7^1, M12^1 and M18^1; and M3^2 and M8^2; any of these haem substituents, S, are assigned to a specific haem in the three-dimensional structure via the symbols (S:*n*), where *n*, in Roman numerals, refers to the haem numbering according to the position of its cysteine ligands in the amino acid sequence; 2D, two-dimensional; 3D, three-dimensional; n.O.e., nuclear Overhauser enhancement; NOESY, nuclear-Overhauser-enhancement spectroscopy; 1D, one-dimensional.

§ To whom correspondence should be addressed.

fully reduced to their fully oxidized positions, and no resonances due to intermediate stages (nor their connectivities) are detected by ^1H n.m.r. However, the redox behaviour of these systems can be partially monitored by measuring the shifts of specific resonances subjected to smaller extrinsic pseudocontact effects in fast exchange along redox titration processes, namely those resonances due to the resolved haem methine (*meso*) and thioether protons belonging to the higher-potential haem.

This work reports the complete assignment of the haem proton resonances of the cytochromes c_3 from both *D. baculatus* strains in the reduced state. N.m.r. is then used to establish an independent structure of the haem core, largely on the basis of calculation of chemical shifts. Whereas structures in solution are often determined by calculating paramagnetic shifts in n.m.r. (BURLESK [28]), or from nuclear Overhauser enhancements (n.O.e.s) (DISGEO [29]), these multihaem proteins are unusual insofar as the spectra are strongly dependent on the ring-current shifts generated from the haem macrocycles. The calculated shifts are used, together with long-range interhaem n.O.e.s, to generate the haem core structure in solution, which present similar overall features to the haem core structures established by X-ray diffraction for other cytochromes c_3 [7,8]. Finally, the variation of chemical shifts and linewidths observed for specific resonances as a function of the level of oxidation is used to establish the haem which has the highest redox potential.

MATERIALS AND METHODS

Sample preparation

Cytochromes c_3 were isolated from *D. baculatus* (Norway 4, DSM 1741), c_3N, and *D. baculatus* (DSM 1743), c_3E, as described in [30]. N.m.r. samples (protein concentration 2 mM in 0.2 M NaCl, pH 9.4–9.8, and temperature 305 K) were prepared by twice freeze-drying the required amount of desalted protein against ^2H_2O, redissolving in ^2H_2O and adding dry NaCl. The pH was adjusted with $NaO^2H/^2HCl$; values are not corrected for the isotope effect. Reduction was achieved enzymically in the presence of traces of hydrogenase from *D. gigas*, isolated as previously described in [31], by maintaining the sample under hydrogen pressure.

NMR experiments

2D n.m.r. spectra of the reduced protein

2D n.m.r. spectra were obtained using a Bruker AMX 500 MHz spectrometer equipped with an Aspect X32 computer, where all data manipulation was carried out. 2D n.m.r. data were collected using standard NOESY (nuclear-Overhauser-enhancement spectroscopy) and total correlation spectroscopy (TOCSY) (MLEV sequence [32]) programs with 2 s water presaturation and 'TPPI' (time proportional phase incrementation [33]). The 90° pulse width was determined before each 2D data collection; the magnet resolution was adjusted until the (non-irradiated) water signal linewidth was lower than 4 Hz. Typical 2D experiments comprised 1024 points acquired in t_2 and 512 increments in t_1, and took 16–26 h according to the number of time-averaged scans (48–72). Fourier transforms (1024*1024 with zero-filling in the F1-dimension) were performed using various window functions (cosine, exponential, gaussian in F_2 and a simple cosine, with or without resolution enhancement, in F_1. One-dimensional (1D) spectra were obtained immediately before and after each 2D experiment to ensure that sample conditions were maintained. NOESY short-range and long-range connectivities were obtained using mixing times of 50 and 400 ms respectively.

The intensity of haem resonance/haem resonance and aromatic resonance/haem resonance cross-peaks in the NOESY data are classified as strong (S), medium (M) or weak (W), with a factor of ~ 2.5 between categories, determined by counting the number of geometrically spaced contour levels on each peak. Peaks may be resolved, non-determined (ND) when there is possible resonance overlap, or obscure (O) when in a crowded, non-resolved region. Typically, the data analysis was performed on several (F_1 and F_2) differentially transformed and/or treated spectra (with and without baseline correction).

Strategy used for the assignment of the NOESY intrahaem cross-peaks

The assignment follows the strategy previously established for 1D [34,35] and 2D [10,36] n.O.e. experiments (Figure 1). Strong intrahaem short-range cross-peaks in the *meso* proton/methyl group (Mi^1 and Mj^2) and thioether proton/methyl group regions of NOESY results obtained with a 50 ms mixing time allow for the initial separation of the *meso* protons H5 or H10 from the *meso* protons H20 and the *meso* protons H15 through the differences between the well-established linear networks H(5 or 10)/H(3^1 or 8^1)/M(7^1 or 12^1)/M(3^2 or 8^2) and H(3^1 or 8^1)/M(3^2 or 8^2), H20/$M2^1$ and $M18^1$, and H15/propionic protons. Thioether methyl groups $M3^2$ or $M8^2$ are unambiguously detected via their J-coupled thioether proton/thioether methyl group connectivities in the TOCSY spectra. Further assignment of the

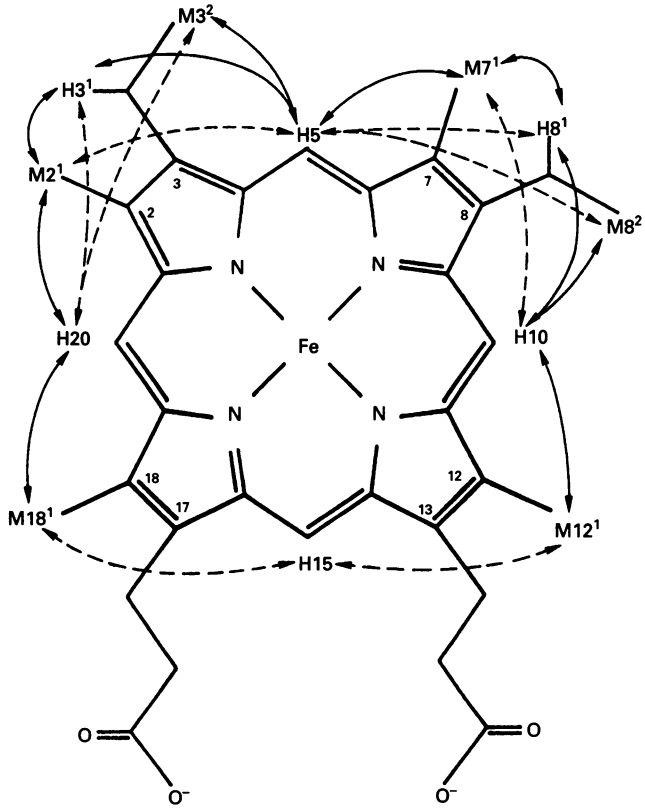

Figure 1 Diagram of haem *c* showing the n.O.e.s detected at short (50 ms) and long (400 ms) mixing times [10,36]

The four *meso* protons (H5, H10, H15 and H20), the four methyl groups ($M2^1$, $M7^1$, $M12^1$ and $M18^1$), the two thioether protons ($H3^1$ and $H8^1$) and the two thioether methyl protons ($M3^2$ and $M8^2$) are indicated [10,36]. Arrows with continuous lines represent short-range interactions, within the haem, and longer-range connectivities are represented by arrows with broken lines.

set of resonances corresponding to each specific haem is obtained via the longer-range (400 ms mixing time) strong H5/M2[1] and H5/M8[2]; H10/M7[1]; H20/M3[2] and H15/M18[1]; and H15/M12[1] cross-peaks in the *meso* proton/methyl group region, and confirmed through the H3[1]/M2[1], and H8[1]/M7[1] connectivities in the thioether proton/methyl group region. The short-range M2[1]/M3[2] and M7[1]/M8[2] cross-peaks which appear in the crowded methyl group/methyl group region are taken into account in the final assignment.

Assignment of the interhaem cross-peaks

Short-range (50 ms mixing time) interhaem connectivities in the *meso* proton/methyl group region are readily detected as cross-peaks connecting the same haem methyl or thioether methyl group to two different *meso* protons, one of the connectivities being intrahaem and the other interhaem [10,36]. Long-range interhaem connectivities cannot be separated *a priori* from the intrahaem signals; they are identified as the cross-peaks which remain after the assignment of the full set of resonances corresponding to each haem.

1D n.m.r. redox titration spectra

1D reoxidation results were obtained using Bruker CXP 300 MHz and Bruker AMX 500 MHz spectrometers. Protein redox titrations proceeded through the reoxidation of the fully reduced samples by injecting successive small volumes (5–20 μl) of air into the n.m.r. tube through serum caps after washing out the hydrogen with argon. Percentage oxidation values attributed to intermediate states were obtained through the summation of the normalized chemical-shift values measured for each haem methyl. Mi[1] resonances shifted to very low field by intrinsic paramagnetic effects, or by calibration of the sample inside the n.m.r. tube by visible spectroscopy. Visible spectra were obtained directly from the solution inside the n.m.r. tube using a procedure similar to that presented in [37]. Specially designed brass cell supports of 1 cm^2 square section with longitudinal 5 mm diameter cylindrical holes (where an n.m.r. tube can be fitted) and laterally aligned 2 mm slits co-linear with the spectrometer lightpath were used to collect visible spectra from the liquid film which sticks to the n.m.r. tube wall after tipping the tube upside down, using a similar support containing an empty n.m.r. tube as a reference. Total oxidation percentages are determined through the ratio of the difference of absorbances of the reduced protein band maxima at 553 nm and 522 nm and the difference of two isosbestic points at 551 and 650 nm; the ratios were calibrated using conventional spectroelectrochemical results.

Determination of the haem core architecture by n.m.r.

A computer program was written to construct an independent model for the haem core architecture from n.m.r. data alone. The haem core was defined by 18 parameters, the co-ordinates of the iron atoms and the Euler angles for the orientations of any three haems, in a system of axes defined by the fourth haem. A standard haem geometry was defined by averaging the co-ordinates obtained by superimposing the four haems from the structure of the cytochrome c_3 from *D. vulgaris* [7]. The orientations of the thioether bridges vary by only a few degrees, so that using the averaged geometry allows the data referring to the thioether protons and the thioether methyl groups to be included without introducing large systematic errors. The Marquardt method was used to optimize the 18 structural parameters with respect to distance constraints provided by the interhaem n.O.e.s and with respect to the ring-current shifts arising from the other haems, calculated as previously described [36].

RESULTS AND DISCUSSION

Assignment of the haem and aromatic resonances

1D and 2D n.m.r. spectra of haem resonances

The 1D spectra (−3 to 11 p.p.m.) obtained under standard conditions for the reduced cytochromes c_3N and c_3E are shown in Figure 2. The 16 one-proton *meso* resonances (numbered in italics according to increasing field) are detected between 11 and 8 p.p.m., and the eight thioether proton resonances, between 7 and 5 p.p.m., are identified via their thioether proton/thioether methyl connectivities in the TOCSY results for c_3N (Figure 3) and c_3E (not shown). The similarity between the n.m.r. 1D and 2D data obtained for both cytochrome c_3 samples in the reduced state show that c_3N and c_3E are structurally very close proteins. Besides their comparable 1D n.m.r. features, the two sets of TOCSY results present analogous cross-peak distributions, namely in the thioether proton/thioether methyl group region. Also the intra- and inter-haem connectivities present in the *meso* proton/thioether proton and *meso* proton/haem methyl and thioether methyl regions of NOESY results obtained at 50 ms mixing time for c_3N [10] and c_3E (not shown), as well as at 400 ms mixing time (Figure 4), show very similar patterns. The broadening effect detected in some of the *meso* and the thioether proton resonances (cf. Figure 2) and of the haem methyl and thioether methyl resonance connectivities (cf. Figures 3 and 4) is due to the intermolecular exchange with residual oxidized protein present in the sample, which is intermediate to fast on the n.m.r. time scale.

Assignment of the haem proton resonances according to type

The specific assignments were quite difficult to obtain, as there are overlapping *meso* proton, thioether proton, haem methyl and thioether methyl resonances in both sets of data. Furthermore, chemical-shift ambiguities are only partially lifted in the 2D spectra, as some connectivities in the *meso* proton/methyl group and thioether proton/methyl group regions are also superimposed. The complete assignment of unresolved haem resonances and/or of their intra- and inter-haem connectivities relied on the comparison of data obtained from the two proteins, as the overlapping resonances are slightly different in the two sets of NOESY results.

The various haem substituent resonances were detected initially through the NOESY intrahaem connectivities at 50 ms mixing time, following the methodology presented above [10,36]. The 16 *meso* proton resonances were classified as H5 or H10, H15 and H20 via their NOESY cross-peaks appearing in the *meso* proton/thioether proton and *meso* proton/methyl group spectral regions; at this stage, haem methyl resonances are only identifiable as M2[1] or M18[1], or as M7[1] or M12[1]. The *meso* proton ordering in the two proteins is slightly changed: *meso* proton *4*, a *meso* H15 in c_3E, switches to number *6* in c_3N. For the sake of uniformity, the numbering determined for c_3E will be used for both proteins when discussing the results.

Assignment of the set of resonances due to each haem

The assignment of the set of resonances due to the *meso* protons, thioether protons, haem methyl and thioether methyl substituents of each one of the four haems was based on the long-range

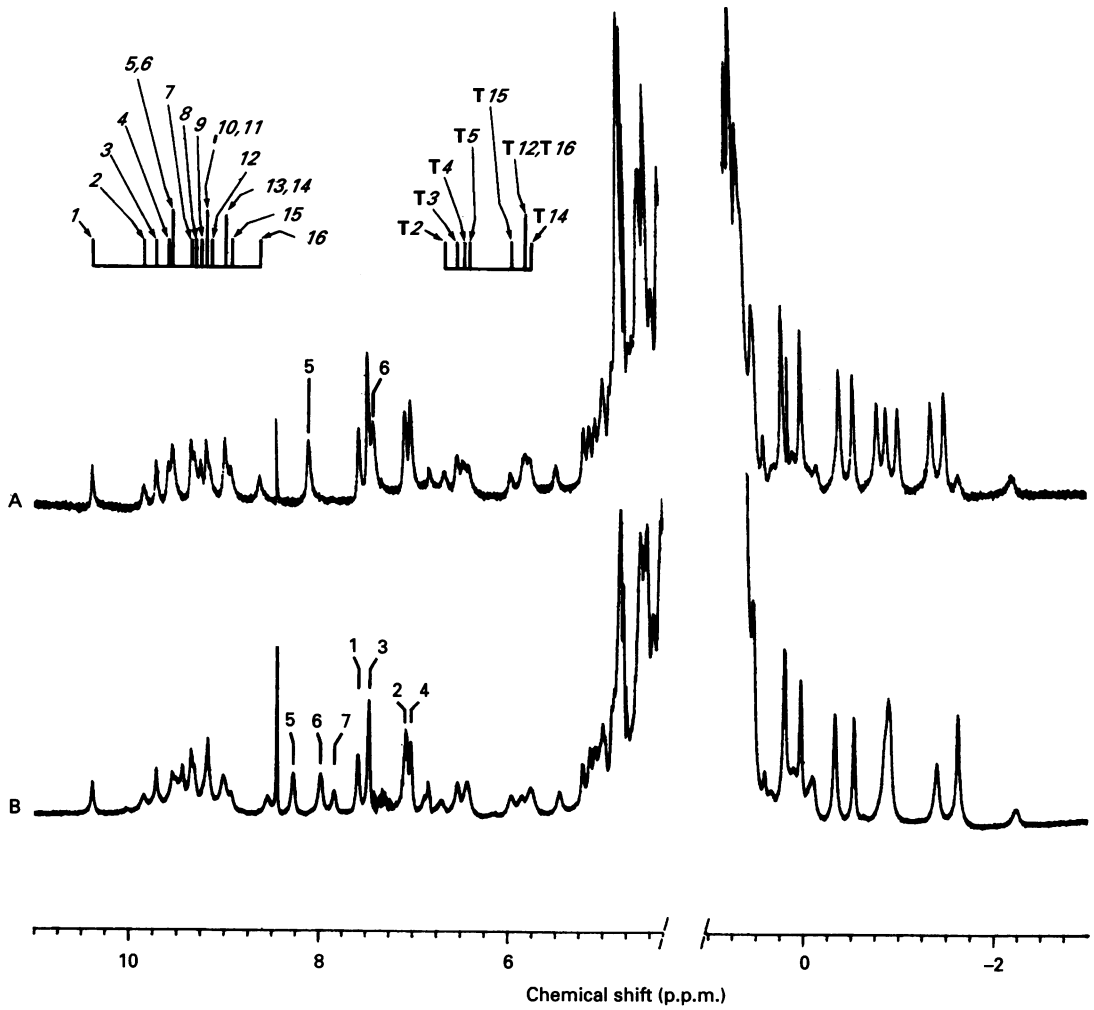

Figure 2 500 MHz 1D n.m.r. spectra (−3 to 11 p.p.m.) obtained for the cytochromes c_3N and c_3E in the reduced state

Trace A, c_3E, pH 9.0; trace B, c_3N, pH 9.5. The sample concentration was 2 mM and the temperature was 305 K. Chemical shifts were determined using the formate resonance (8.439 p.p.m.) as a secondary internal reference. *Meso* proton resonances are numbered in *italics* according to increasing field; thioether protons (T) are numbered according to the *meso* proton they connect to in the NOESY results obtained at 50 ms mixing time [10]. Aromatic proton resonances are indicated with numerals 1–7 (*not* italic); note that signals 5–7 from F8 in c_3N are replaced by signals 5 and 6 in c_3E.

intrahaem connectivity results obtained at 400 ms mixing time (Figure 4). The intrahaem connectivities were assigned by listing all strong connectivities in the *meso* proton/methyl group region which were not present in the NOESY data obtained at 50 ms mixing time and searching for the characteristic longer-distance correlation networks (cf. Figure 1). This process allowed for the separation of four self-consistent groups of haem resonances (50 ms mixing time haem methyl/thioether methyl intrahaem cross-peaks were taken into account at this stage), each due to one haem in the structure.

The full list of established intrahaem NOEs found in the c_3N and c_3E spectral data at 50 and 400 ms mixing time is presented in Table 1. Cross-peaks present in the NOESY results of both proteins which are not part of the above patterns were attributed to interhaem effects.

Specific assignment of the haem resonances within the 3D structure

The assignment of each group of *intra*-correlation haem resonances to a specific haem moiety in the X-ray structure, where the haems were numbered according to the position of their cysteine ligands in the amino acid sequence, was based on the NOESY short-range *inter*haem connectivities assigned for c_3N using 50 ms mixing time [10] (c_3E *inter*haem cross-peaks are identical), and confirmed via the NOESY long-range (400 ms mixing time) *inter*haem connectivities and the haem resonance/aromatic resonance ones (see below). In both proteins, two *inter*haem patterns are present in the *meso* proton/methyl group region [10]: (i) the cross-correlations of *meso* proton *5*, identified as *meso* proton 10 of haem IV (H10:IV), with haem methyl *1*, ($M2^1$:III), and of *meso* proton *1* (H20:III) with thioether methyl *5* ($M8^2$:IV); and (ii) the cross-correlations of *meso* proton *2* (H5:III) with thioether methyl *16* ($M8^2$:I), and of *meso* proton *16* (H10:I) with thioether methyl *2* ($M3^2$:III). Such *inter*haem connectivities show that the edges of haems I and IV, which contain the substituents H10 and $M8^2$, are the ones oriented towards the interior of the haem core, close to the correlated haem III substituents. As demonstrated previously [10], these results are not compatible with the haem core architecture established for c_3N by X-ray crystallography [5,6],

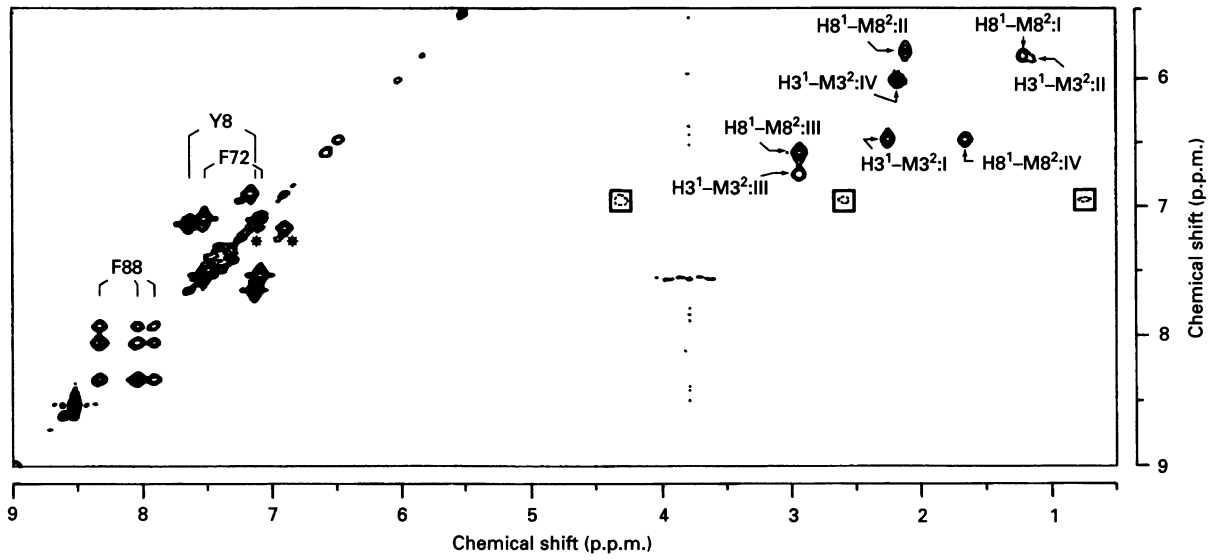

Figure 3 Part of the 500 MHz TOCSY (40 ms mixing time) spectrum obtained for c_3N in the reduced state

The sample concentration was 2 mM, the temperature 305 K and the pH 9.5. Cross-peaks arising from the eight thioether proton/thioether methyl group couplings are clearly visible, as well as the patterns due to the tyrosine Y8 and two of the phenylalanine residues (F88 and F72). The boxes mark the positions of four cross-peaks connecting a broad resonance in the aromatic region, which are visible at lower contour levels. The pattern is assigned to the aromatic sidechain of F34. The peaks marked '*' are attributed to an impurity.

but can be rationalized by assuming that the haem core architecture determined for the cytochromes c_3 from *D. vulgaris* [7] and *D. gigas* [8] is conserved in c_3N and c_3E (see below).

Interhaem connectivities thus allowed for the identification of *meso* proton *5* as H10:IV, *meso* proton *16* as H10:I and *mesos* protons *1* and *2* as *mesos* H20:III and H5:III respectively, and automatically assign the remaining intrahaem correlated substituent resonances of haems I, III and IV; the haem signals due to haem II, being the only set left at this stage, are then assigned by exclusion. The chemical shifts of the various substituents of each haem in the 3D structure of c_3N and c_3E are presented in Table 2.

The interhaem connectivities detected at longer (400 ms) mixing time (Figure 4) agree with the proposed assignment, as they correspond to effects predicted assuming the conservation of the structure of the haem core. Thus the cross-peaks *meso* proton *16*/thioether methyl *12* and *meso* proton *5*/thioether methyl *2* are identified as the correlations H10:I-M3²:II (distance of 0.435 nm), and H10:IV-M3²:III (distance of 0.530 nm); other interhaem connectivities are obscure and/or indeterminate, or appear in one data set only. Several interhaem cross-peaks are picked out in Figure 4.

Assignment of the aromatic resonances

The assignment of the aromatic resonances to specific residues in the structure was performed by observing the c_3N and c_3E TOCSY and NOESY spectra and confirmed by monitoring their behaviour along the redox titration monitored by 1D n.m.r.

Resonances in the aromatic region of the 1D spectra of c_3N and c_3E (Figure 2) are assigned via the observed TOCSY cross-peak patterns in the aromatic region for c_3N (Figure 3) and for c_3E (results not shown). These include in both samples a pattern for a rapidly flipping tyrosine (peaks 1 and 2) and a phenylalanine (peaks 3 and 4) where the *para* proton has an identical chemical shift to either the *ortho* or the *meta* protons. A lower-field phenylananine pattern in the c_3N spectrum (peaks 5, 6 and 7) is absent in the c_3E data, and instead a tyrosine pattern is present (peaks 5 and 6).

The amino acid sequence of c_3N [38] includes three phenylalanine (positions 34, 72 and 88) and one tyrosine (position 8) residues; the sequence partially determined for c_3E is identical, although its overall amino acid composition differs somewhat from that of c_3N [39]. In the 1D and TOCSY spectra of fully reduced c_3N, the only tyrosine pattern is obviously assigned to Y8. The two phenylalanine patterns detected are due to the residues F72 or F88. Indeed, the resonances due to F34, the structurally conserved residue in all cytochromes c_3 studied and which is in close contact with two haems [17], are shifted out of the aromatic region, as is also observed in the n.m.r. data obtained for the cytochromes c_3 from *D vulgaris* [26] and *D. gigas* [27]. Apart from being shifted into the main envelope of signals, these resonances are broadened by slow flipping. In fact, there is a series of weak cross-peaks in the TOCSY spectra of c_3N (Figure 3), connecting a broad resonance in the aromatic region to others in the range 2–4 p.p.m. This is compatible with a partial TOCSY pattern arising from a rotationally frozen phenylalanine side chain, and the frequencies are tentatively assigned to F34. The connected chemical shifts are in broad agreement with the values calculated considering the ring-current shifts generated by the haem core for the F34 aromatic protons, which predict aromatic resonances at 7.5, 6.3, 2.8, 0.9 and 0 p.p.m. This pattern is conserved in the c_3E TOCSY (results not shown), together with the Y8 (peaks 1 and 2) and the anomalous phenylalanine (peaks 3 and 4) cross-peaks; as already stated, the lower-field phenylalanine pattern is missing, being replaced by a tyrosine pattern (peaks 5 and 6).

NOESY connectivity signals between haem resonances and aromatic resonances were found with growing intensities in the 50-, 200- and 400-ms-mixing-time NOESY results. According to the X-ray structure for c_3N, partially reproduced in Figure 5(a), F72 is removed (more than 1 nm) from the vicinity of any haem, so no NOESY correlations between F72 and haem resonances are expected. However, both Y8 and specially F88 are closer to

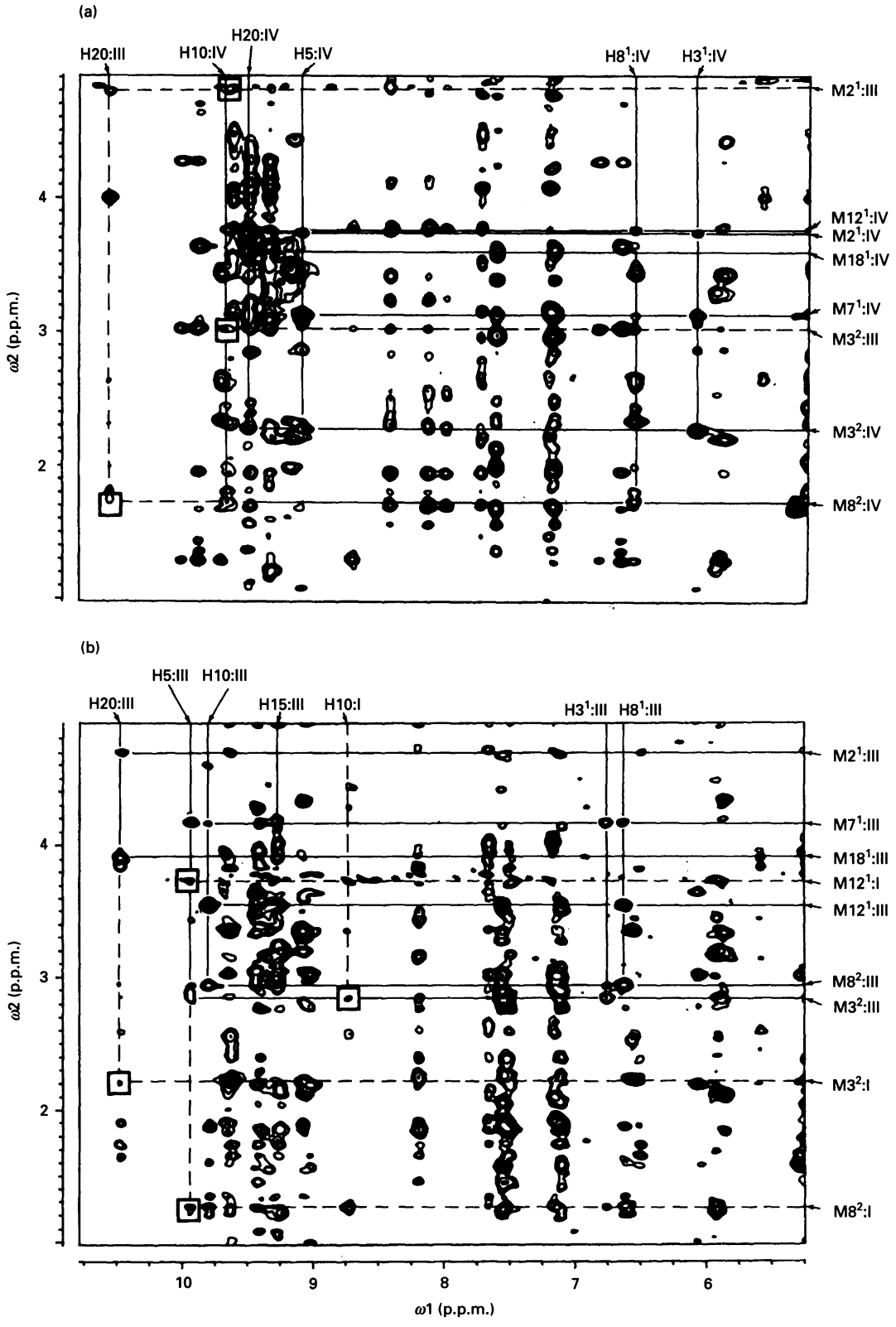

Figure 4 500 MHz NOESY (400 ms mixing time) spectra obtained for the cytochromes c_3N and c_3E in the reduced state

(**a**) c_3N, pH 9.5; (**b**) c_3E, pH 9.3. The sample concentration was 2 mM and the temperature was 305 K. Intrahaem connectivity networks are indicated with full lines; interhaem cross-peaks are enclosed by boxes, and the connected resonances are identified with broken lines. Spectrum (**a**) shows the intrahaem network for haem IV and the interhaem cross-peaks between haems III and IV, and spectrum (**b**) shows the intrahaem network for haem III and the interhaem cross-peaks between haems I and III. $\omega 1$ and $\omega 2$ are the frequencies of evolution and detection respectively.

Table 1 Grouping of the haem resonances in c_3N and c_3E

The intrahaem connectivities obtained at 50 and 400 ms mixing time were used for the assignment; S, ND, M etc. are defined in the text

Mixing time (ms)	Connectivity	Haem... I	II	III	IV
50	$M2^1$–$M3^2$	S	S	ND	S
	$M7^1$–$M8^2$	S	S	S	ND
400	H20–$M3^2$	S	M/ND	W†	S
	H5–$M2^1$	ND	ND	M	S
	H5–$M8^2$	S	ND	S*	W/ND†
	H10–$M7^1$	M/ND	S	S	S/ND
	H15–$M18^1$	ND	ND	S	ND
	H15–$M12^1$	ND	ND	ND	ND
	$H3^1$–$M2^1$	ND	M*	W	S
	$H8^1$–$M7^1$	ND	M*	S	W

* Present in some c_3E plots only.
† Present in some c_3N plots only.

Table 2 Chemical shifts determined for the haem substituents in c_3N and c_3E

	Chemical shift (p.p.m.)							
	Haem I		Haem II		Haem III		Haem IV	
	c3E	c3N	c3E	c3N	c3E	c3N	c3E	c3N
H5	9.52	9.61	9.10	9.21	9.82	9.91	8.88	9.00
H10	8.51	8.63	8.95	9.05	9.67	9.79	9.47	9.60
H15	9.43	9.67	9.29	9.34	9.15	9.25	9.29	9.41
H20	8.99	9.05	9.15	9.30	10.38	10.45	9.31	9.41
$M2^1$	3.26	3.33	3.05	3.12	4.58	4.69	3.53	3.60
$M7^1$	3.22	3.79	3.11	3.17	4.06	4.16	2.89	2.98
$M12^1$	3.58	3.70	3.23	3.33	3.42	3.51	3.55	3.33
$M18^1$	3.26	3.33	3.42	3.51	3.77	3.88	3.38	3.45
$H3^1$	6.40	6.49	5.73	5.90	6.64	6.65	5.90	6.04
$H8^1$	5.73	5.90	5.71	5.85	6.48	6.60	6.37	6.49
$M3^2$	2.12	2.18	1.01	1.28	2.80	2.83	2.04	2.17
$M8^2$	1.07	1.20	1.98	2.11	2.82	2.92	1.59	1.67

one or more haem moieties. Thus the NOESY aromatic resonance/haem resonance connectivities were interpreted using the distances presented in Table 3. Those distances are measured between the protons of F88 and Y8 in the X-ray structure of the cytochrome from *D. baculatus* (Norway 4) and the *meso* protons and C-atoms of haem methyl and thioether methyl substituents from the haem core in the X-ray structure of the cytochrome from *D. vulgaris*, after carefully superimposing the four iron atoms in both structures [root-mean-square deviation (rmsd) 0.06 nm]. There were cross-peaks connecting haem methyl *5*, ($M12^1$:IV), thioether methyl *2* ($M3^2$:III), and haem methyl *16* ($M12^1$:I), as well as haem methyl *2* ($M7^1$:III) and haem methyl *10* ($M18^1$:II), with resonances due to the lower-field phenylalanine pattern in c_3N, which was thus assigned to F88. Similar connectivities occur between the same haem resonances and the lower-field tyrosine pattern in c_3E, assigning it to a tyrosine residue structurally similar to F88. The remaining phenylalanine pattern is then due to F72. One cross-peak between Y8 and *meso* proton *4*, already detectable at 50 ms mixing time, identifies this proton as H15:I; this correlation was also predicted

Table 3 Distances between aromatic and haem protons

The distances (in nm) were measured between the aromatic protons from residues F88 and Y8 in the c_3N X-ray structure and the *meso* protons and C-atoms of haem methyl and thioether methyl substituents in the haem core of *D. vulgaris* after superimposing the four iron atoms in the two structures. The right-hand column shows the cross-peaks observed at 400 ms mixing time. Note that protons 2 and 6, and 3 and 5, are rendered equivalent by rapid 180° rotation of the ring.

Residue protons	Haem protons	Distance (nm)	Cross-peaks at 400 ms
Tyr^8			
2	H15:I	0.415	S
3	H15:I	0.195	S
2	$M12^1$:I	0.693	M
3	$M12^1$:I	0.626	W*
Phe^{88}			
2	$M12^1$:I	0.465	S†
3	$M12^1$:I	0.665	S†
2	H5:II	0.569	
3	H5:II	0.670	
2	H20:II	0.551	
2	$M2^1$:II	0.361	S/ND
3	$M2^1$:II	0.584	ND
2	$M3^2$:II	0.416	ND
3	$M3^2$:II	0.624	ND
2	$M3^2$:III	0.495	S
3	$M3^2$:III	0.708	W
2	H10:IV	0.527	
3	H10:IV	0.657	
3	H15:IV	0.579	M
2	$M12^1$:IV	0.393	S†
3	$M12^1$:IV	0.469	S†
4	$M12^1$:IV	0.616	

* Observed in c_3E results only.
† Non-determined in the c_3N results.

for c_3N, as the *meso* H15 position is unchanged upon the haem I rotation. The connectivity signal between F88 and either of the two *meso* protons H15:II or H15:IV is attributed to the connectivity F88(3)–H15:IV. Some predicted connectivities, such as F88–H20:II and F88–$M3^2$:II, are absent, and unexpected cross-peaks connecting F88–$M7^1$:III and F88–$M18^1$:II resonances occur in the spectra. Such anomalies are considered acceptable bearing in mind the approximate nature of the atomic co-ordinates. Thus it appears that the positions of the aromatic groups in the X-ray structure of c_3N are roughly correct with respect to the haem iron co-ordinates, confirming the haem assignment reported above.

Again, it is noteworthy that the haem resonance–aromatic resonance connectivities predicted using distance data extracted directly from the co-ordinates from c_3N, without rotating haems I and IV, are incompatible with the observed experimental results: cross-peaks should exist between F88 (2) and $M18^1$:I, separated by 0.453 nm, and between F88 (protons 2, 3 and 4) and $M18^1$:IV, separated by 0.474 nm, 0.385 nm and 0.470 nm respectively, which are not present in the NOESY results, the connectivities between F88 and $M12^1$:IV and $M12^1$:I being detected instead.

Independent determination of the haem core architecture by n.m.r.

In the preliminary report of n.m.r. studies on cytochrome c_3 from *D. baculatus* [10], only two interhaem n.O.e.s were needed to

demonstrate that the X-ray structure [5,6] is fundamentally incorrect and that the n.m.r. data are compatible with the orientation of the four haems found in the *D. vulgaris* cytochrome c_3 [7]. The complete assignment of the haem proton resonances now reveals a total of six interhaem n.O.e.s and allows a more detailed comparison of structures.

If the orientations of the thioether bridges and propionates are neglected, the haem core architecture can be defined by 18 parameters, as described in the Materials and methods section. The interhaem n.O.e.s provide approximate distance constraints, but their number is limited by resonance overlap and line-broadening by fast exchange with residual oxidized molecules, and there are too few of them to define the haem orientations. The constraints are also imprecise because of the effects of spin diffusion; comparison of the n.O.e. intensities with data from the c_3 from *D. vulgaris* [36] shows that the six interproton distances may be anywhere in the range 0.35–0.60 nm. However, the study of the *D. vulgaris* c_3 [38] also showed that the haem protons are subject to large ring-current shifts arising from neighbouring haems, and that the effect of other aromatic groups, including the axial histidine ligands, is relatively small. Thus the differences between the observed chemical shifts and reference shifts (taken from horse cytochrome *c* [38]) provide a further 48 structure-dependent values. Although the chemical-shift data are less specific than the n.O.e.s, it was possible to determine the structural parameters by using the two types of information together. The relative weight given to n.O.e.s and chemical shifts in the global minimization was determined by adding the squares of the deviations from the range 0.35–0.60 nm in units of 0.1 nm to the squares of the deviations of the ring-current shifts in units of p.p.m.

A test of the program using n.m.r. data obtained for *D. vulgaris* gave a best fit with a deviation of 2×10^{-4} nm^2 for 20 distance constraints and an rmsd of 0.27 p.p.m. for the shifts. This may be compared with an rmsd of 0.32 p.p.m. for shifts calculated directly from the X-ray co-ordinates and including all of the aromatic groups. Small differences were found between this n.m.r. structure and the X-ray structure, with an rmsd of 0.18 nm for the iron co-ordinates and an rmsd of 21° for the Euler angles, and these differences may be regarded as a measure of the effect of the systematic errors described above. The overall rmsd of the porphyrin heavy atoms obtained after optimizing the superposition of the iron atoms in the two structures was 0.233 nm, so we are satisfied that the method is capable of producing a unique structure with a precision for atomic co-ordinates of approx. 0.25 nm.

The best-fit structure obtained for the *D. baculatus* data gave rms errors of 7×10^{-4} nm^2 for the distance constraints and 0.26 p.p.m. for the chemical shifts. This is illustrated in Figure 5 together with the haems and the aromatic side chains from the X-ray structures of *D. baculatus* and *D. vulgaris* (Brookhaven files 1cy3.pdb and 2cdv.pdb). Comparison with the *D. vulgaris* X-ray structure shows an rmsd for the Euler angles of 32° and of 0.13 nm for the iron co-ordinates. The rmsd for the porphyrin heavy atoms was 0.297 nm, with the largest deviation, 0.417 nm, coming from haem II, which is unsurprising, since the haem II position was determined using a single distance constraint. With the exception of haem II, therefore, the structure derived independently by using all of the n.m.r. data fits the X-ray structure of the counterpart protein from *D. vulgaris* within the expected accuracy of 0.25 nm. By contrast, the overall deviation from the *D. baculatus* X-ray structure was 0.449 nm, with individual contributions of 0.505, 0.436, 0.272, and 0.536 nm from haems I–IV, respectively. Thus, apart from the relatively poor definition of the orientation of haem II, we find that the deviations for haems I and IV have twice the expected values.

These deviations may be compared with the values 0.494, 0.075, 0.041 and 0.488 nm obtained for the four haems by direct comparison of the two crystal structures and indicate clearly that the solution structure of *D. baculatus* cytochrome c_3 has haems I and IV in a similar orientation to that found in the crystal structure of *D. vulgaris* cytochrome c_3, rather than the rotated orientation of the X-ray structure of the protein from *D. baculatus*.

It is important to note that the n.m.r. structure does not depend on specific assignment of haem protons within the protein structure; it only requires that the proton resonances be assigned to four different haems. Finding the optimum superimposition of the iron atoms in the n.m.r. structure on those of the crystal structure implicitly assigns the resonances to specific haems. It is therefore completely independent of the specific assignment discussed above with respect to interhaem and haem–aromatic interproton distances in the crystal structure, and the two are in complete agreement.

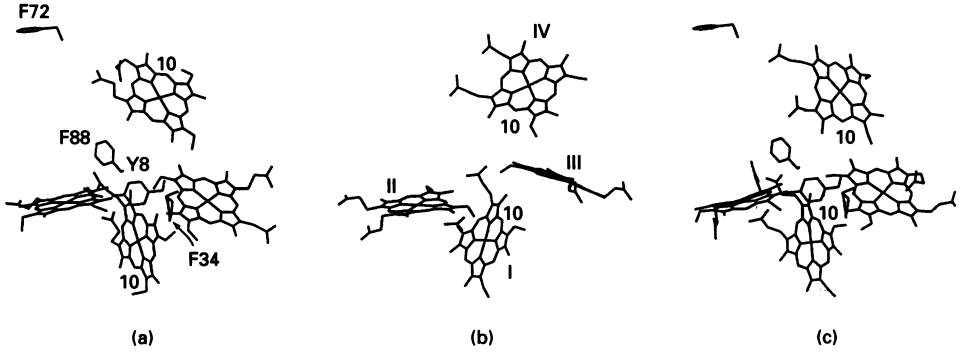

Figure 5 Orientation of the four haems in cytochrome c_3

Structures were obtained from: (**a**) the X-ray structure of the protein from *D. baculatus* (Norway 4), c_3N; (**b**) calculations based on the n.m.r. data from c_3N (see the text); (**c**) the X-ray structure of the protein from *D. vulgaris*, c3V. The phenylalanine and tyrosine side chains of c_3N are included in that structure, and the same side chains are shown superimposed on the structure of c3V. The three structures have the same overall orientations, determined by the position of the four iron atoms, and the haems are numbered according to the position of their cysteine ligands in the protein sequence. The positions of the H10 *meso* protons of haems I and IV are indicated in each case.

Assignment of the higher-potential haem in the 3D structure

Previous e.p.r. studies [23–25] have shown that the both c_3N and c_3E have three lower-potential haems, with relatively similar redox potentials ($e_1 = -350$ mV, $e_2 = -325$ mV and $e_3 = -280$ mV), and the potential of the fourth one is more than 100 mV higher than any of the other three.

The assignment of the highest redox potential to a specific haem in the 3D structure relies on the behaviour of some of the previously identified *meso* proton and thioether proton resonances observed throughout the redox titration process from 0 to 75% sample oxidation (Figure 6). Under these conditions, the haem resonances of the higher-potential haem are expected to be shifted exclusively via the (small) extrinsic pseudocontact effects created by the reoxidation of the neighbouring low-potential haem, becoming undetectable because of exchange broadening due to their own isotropic effects at percentage oxidation values higher than 75%. The lower-potential haem resonances should broaden from the very start of the reoxidation procedure.

The shift variation detected for the *meso* and thioether proton resonances with the redox titration process are presented in Figure 6. The well-resolved *meso* proton *1*, H20:III, as well as other *meso* protons from haem III, *mesos 3*, H10:III, and *11*, H15:III, can be monitored up to 70% oxidation. The resonance due to thioether proton T*3*, H8^1:III, is also detected during the early stages until ~65% oxidation, but at higher values it cannot be assigned unambiguously, as it becomes superimposed on other resonances. The remaining *meso* and thioether proton resonances were shifted and broadened from the start of the reoxidation procedure. These include the protons H5:III and H3^1:III, which are oriented towards the interior of the haem core and therefore subjected to strong interhaem pseudocontact interactions due to the oxidation of the lower-potential haems.

These observations establish haem III as the higher-potential haem in the cytochromes c_3 isolated from *D. baculatus*. This identification is different from previous suggestions based either on structural considerations [21,24] or on conclusions deriving from data obtained using electrochemical methods on samples subjected to chemical modification [19,41]. However, they agree with one of the two possible solutions (the most favoured one) proposed by Gayda and co-workers based on redox titration results followed by single-crystal e.p.r. studies [25].

Conclusions

The two strains, Norway 4 and DSM 1743, of *D. baculatus* are phylogenetically closely related, as indicated by their 16 S ribosomal RNA sequences [11]. In addition, they are atypical of the genus both in terms of their peculiar morphology and in lacking the proteins desulfoviridin (a taxonomical marker) and flavodoxin.

The close phylogenetic relationship between the *D. baculatus* organisms, substantiated by the identical first 20 N-terminal amino acids and by the overall similar amino acid composition of c3N and c3E [39], is reflected in the striking similarities found in the n.m.r. spectra and redox titration patterns of both proteins. These patterns, on the other hand, are quite different from the data obtained for other cytochromes c_3 in the reduced state [26,27,36]. 1D and 2D spectra allowed for the assignment of the aromatic resonances and showed one amino acid change, that of F88 in c_3N by a spatially isomorphic tyrosine residue in c_3E. These assignments are confirmed by the early redox titration results followed by 1D n.m.r. (cf. Figure 6). Resonances 5, 6 and 7 in c_3N and 5, 6 in c_3E, assigned in c_3N to F88, are strongly shifted during the titration, so they must be close to at least one haem, while resonances 3 and 4, assigned to the spatially distant F72, remain practically unshifted. Other substitutions or

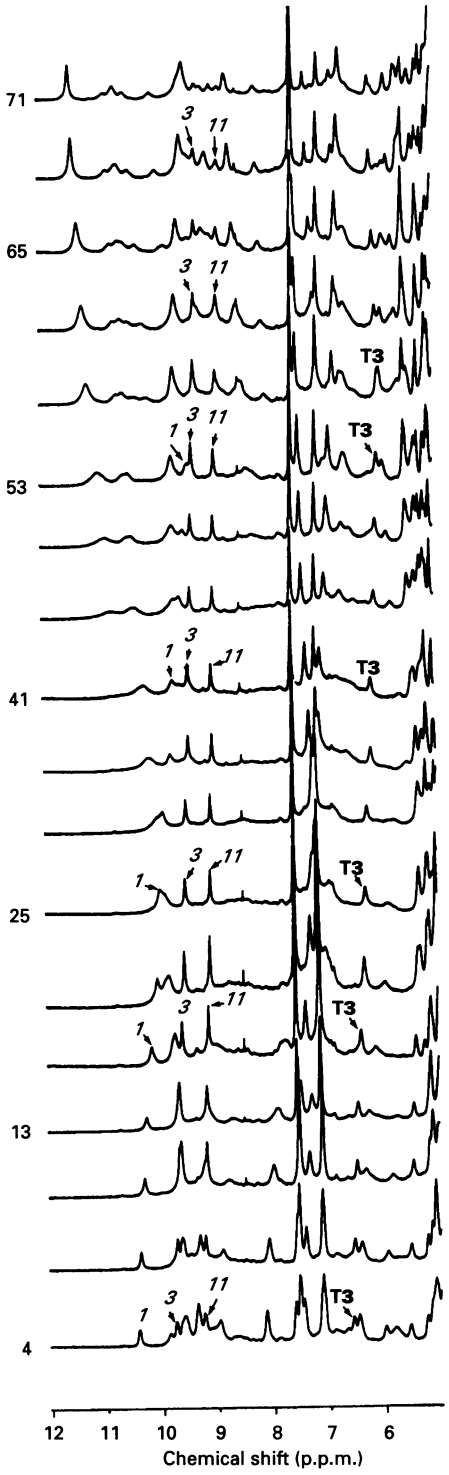

Figure 6 Redox titration of cytochrome c_3E followed by 1D ^{1}H n.m.r. (500 MHz) in the *meso* and aromatic regions

The sample concentration was 2 mM, the pH 7.8 and the temperature 305 K. The spectra are ordered according to increasing percentage oxidation of the sample, which is indicated on the left side of the corresponding spectrum. The bottom spectrum corresponds to the reduced and the top one to the partially (71%) oxidized sample. *1*, H20:III; *3*, H10:III; *11*, H15:III; **T3**, H8^1:III.

deletions, involving non-resolved resonances such as those due to aliphatic side chains, were not searched for and are by no means impossible, but their effect is apparently negligible in the most noticeable n.m.r. features.

As a first approach, the faster intermolecular electron-exchange rates found for the cytochromes from *D. baculatus* can be rationalized taking into account the different repulsive electrostatic potentials between protein molecules. In fact, the cytochrome c_3 from *D. baculatus* (Norway 4) has an ionizable side-chain amino acid composition (18 lysine-plus-arginine against 20 glutamate-plus-aspartate residues) which ensures that the formal charge in a wide pH range is smaller than that occurring in the cytochromes c_3 from *D. vulgaris* (22 lysine-plus-arginine-plus-histidine against 17-glutamate-plus-aspartate residues) and *D. gigas* (17 lysine-plus-arginine against 22 glutamate-plus-aspartate residues). Assuming at this stage that the repulsive potential is dominant in terms of the preferred intermolecular pathways for electron transfer, such features broadly account for the distinctive kinetic properties found for the cytochromes from *D. baculatus*.

The identification of haem III as the higher-potential haem agrees with the assignment obtained by Gayda et al. from single-crystal e.p.r. [25]. Thus haem III is the lowest-potential haem in the cytochrome c_3 from *D. vulgaris* [26], but the highest-potential one in the cytochrome from *D. baculatus*. In structural terms, a possible explanation for the considerable increase of the haem III redox potential in c_3N is the presence of a positive charge due to the amine group of Lys-99 less than 0.8 nm (nitrogen atom) from the iron atom of haem III (and less 0.5 nm from the carbon C10:III, on the haem edge). Such a charged group is absent in the cytochrome from *D. vulgaris*. Further comparison of haem redox potentials in both proteins is currently underway.

REFERENCES

1 Postgate, J. R. (1984) The Sulphate-Reducing Bacteria (2nd edn.), pp. 75–77, Cambridge University Press, Cambridge
2 Fauque, G., LeGall, J. and Barton, J. L. (1991) in Variations of Autotropic Life (Shively, J. M. & Barton, L. L., eds.), pp. 304–305, Academic Press, New York
3 Peck, H. D. Jr and LeGall, J. (1982) in Sulphur Bacteria (Postgate, J. R. & Kelly, D. P., eds.), pp. 13–36, The Royal Society, London
4 Fauque, G. F., Henri, D. and LeGall, J. (1979) Arch. Microbiol. **121**, 262–264
5 Haser, P., Pierrot, M., Frey, M., Payan, F., Astier, J.-P., Brusch, M. & LeGall, J. (1979) Nature (London) **282**, 806–810
6 Pierrot, M., Haser, R., Fey, M., Payan, F. and Astier, J-P. (1982) J. Biol. Chem. **257**, 14341–14348
7 Higushi, Y., Kusonoki, M., Matsuura, Y., Yasuoka, N. and Kakudo, M. (1984) J. Biol. Chem. **172**, 109–139
8 Kissinger, C. (1989) Ph.D. Thesis, Washington University
9 Higushi, Y., Kusonoki, K., Yusuoka, N., Kakudo, M. and Yagi, T. (1981) J. Biochem. (Tokyo) **90**, 1715–1723
10 Coutinho, I. B., Turner, D. L., LeGall, J. and Xavier, A. V. (1992) Eur. J. Biochem. **209**, 329–333
11 Deveraux, R., He, S.-H., Doyle, C. L., Oackland, S., Stahl, D. A., LeGall, J. and Whitman, W. B. (1990) J. Bacteriol. **172**, 3609–3619
12 Dobson, C. M., Hoyle, N. J., Geraldes, C. F., Bruschi, M., LeGall, J., Wright, P. E. and Williams, R. J. P (1974) Nature (London) **249**, 425–429
13 Moura, I., Moura, J. J. G., Santos, H. and Xavier, A. V. (1980) Cienc. Biol. (Coimbra) **5**, 189–191
14 Moura, J. J. G., Santos, H., Moura, I., LeGall, J., Moore, G. R., Williams, R. J. P. and Xavier, A. V. (1982) Eur. J. Biochem. **127**, 151–155
15 Santos, H., Moura, J. J. G., Moura, I., LeGall, J. and Xavier, A. V. (1984) Eur. J. Biochem. **141**, 283–296
16 Fan, K., Akutsu, H., Kyoguku, V. and Niki, K. (1990) Biochemistry **29**, 2257–2263
17 Su-Park, J., Kano, K., Niki, K. and Akutsu, H. (1991) FEBS Lett. **285**, 149–151
18 Guerlesquin, F., Bruschi, M. and Wuthrich, K. (1986) Biochem. Biophys. Acta **830**, 296–303
19 Dolla, A., Guerlesquin, F., Bruschi, M. and Haser, R. (1991) J. Mol. Recognition **4**, 27–33
20 Bianco, P. and Haladjian, Y. (1991) Biochim. Biophys. Acta **26**, 1001–1004
21 Bruschi, M., Loufti, M., Bianco, P. and Haladjian, J. (1984) Biochem. Biophys. Res. Commun. **120**, 384–359
22 Niki, K., Kobayashi, Y. and Matsuda, H. (1984) J. Electroanal. Chem. **178**, 333–341
23 Gayda, J-P., Berosman, H., Bertrand, P., More, L. and Asso, M. (1988) Eur. J. Biochem. **177**, 199–206
24 Moura, I., Teixeira, M., Huyinh, B. H., LeGall, J. and Moura, J. J. G. (1988) Eur. J. Biochem. **176**, 365–369
25 Guigliarelli, B., Bertrand, P., More, L., Haser, R. and Gayda, J. P. (1990) J. Mol. Biol. **216**, 161–166
26 Salgueiro, C., Turner, D. L., Santos, H., LeGall, J. and Xavier, A. V. (1992) FEBS Lett. **314**, 155–158
27 Piçarra-Pereira, M. A., Turner, D. L., LeGall, J. and Xavier, A. V. (1993) Biochem. J. **294**, 909–915
28 Barry, C. D., Glasel, J. A., North, A. C. T., Williams, R. J. P. and Xavier, A. V. (1971) Nature (London) **232**, 236–245
29 Havel, T. F. and Wuthrich, K. (1984) Bull. Math. Biol. **46**, 673–698
30 Bruschi, M., Hatchikian, C. E., Golovleva, L. A. and LeGall, J. (1977) J. Bacteriol. **129**, 30–38
31 LeGall, J., Ljunghdhal, P. O., Moura, I., Peck, H. D., Jr., Xavier, A. V., Moura, J. J. G., Teixeira, M., Huynh, B. H. & DerVartanian, D. V. (1982) Biochem. Biophys. Res. Commun. **106**, 610–616
32 Bax, A. and Davis, D. G. (1985) J. Magn. Reson. **63**, 355–360
33 Marion, D. S., Wuthrich, K. (1982) Biochem. Biophys. Res. Commun. **95**, 1–6
34 Keller, R. M. and Wuthrich, K. (1978) Biochim. Biophys. Acta **533**, 195–208
35 Senn, H., Keller, R. M. and Wuthrich, K. (1980) Biochem. Biophys. Res. Commun. **92**, 1362–1369
36 Turner, D. L., Salgueiro, C., LeGall, J. and Xavier, A. V. (1992) Eur. J. Biochem. **210**, 931–936
37 Kinura, W., Nakajima, S., Niki, K. and Inokuchi, H. (1985) Bull. Chem. Soc. Jpn. **58**, 1010–1012
38 Bruschi, M. (1981) Biochem. Biophys. Acta **671**, 219–224
39 Fauque, G. (1985) Thése au Doctorat d'Etat, Université de Technologie, Compiègne
40 Dolla, A., Cambillau, C., Bianco, P., Halladjian, J. and Brushi, M. (1987) Biochem. Biophys. Res. Commun. **147**, 812–829
41 Dolla, A., Leroy, G., Guerlesquin, F. and Bruschi, M. (1991) Biochem. Biophys. Acta **1058**, 171–177

Received 18 January 1993/13 May 1993; accepted 18 May 1993