

Solution Structure of *Desulfovibrio vulgaris* (Hildenborough) Ferrocyclochrome c_3 : Structural Basis for Functional Cooperativity

Ana C. Messias¹, Dieter H. W. Kastrau¹, Helena S. Costa¹
Jean LeGall^{1,2}, David L. Turner^{1,3}, Helena Santos¹ and António V. Xavier^{1*}

¹Instituto de Tecnologia
Química e Biológica
Universidade Nova de Lisboa
Rua da Quinta Grande, 6
Apartado 127, 2780 Oeiras
Portugal

²Department of Biochemistry
University of Georgia, Athens
GA 30602, USA

³Department of Chemistry
University of Southampton
Southampton, SO17 1BJ, UK

Desulfovibrio vulgaris cytochrome c_3 is a 14 kDa tetrahaem cytochrome that plays a central role in energy transduction. The three-dimensional structure of the ferrocyclochrome at pH 8.5 was solved through two-dimensional ¹H-NMR. The structures were calculated using a large amount of experimental information, which includes upper and lower distance limits as well as dihedral angle restraints. The analysis allows for fast-flipping aromatic residues and flexibility in the haem plane. The structure was determined using 2289 upper and 2390 lower distance limits, 63 restricted ranges for the ϕ torsion angle, 88 stereospecific assignments out of the 118 stereopairs with non-degenerate chemical shifts (74.6%), and 115 out of the 184 nuclear Overhauser effects to fast-flipping aromatic residues (62.5%), which were pseudo-stereospecifically assigned to one or the other side of the ring. The calculated NMR structures are very well defined, with an average root-mean-square deviation value relative to the mean coordinates of 0.35 Å for the backbone atoms and 0.70 Å for all heavy-atoms. Comparison of the NMR structures of the ferrocyclochrome at pH 8.5 with the available X-ray structure of the ferricytochrome at pH 5.5 reveals that the general fold of the molecule is very similar, but that there are some distinct differences. Calculation of ring current shifts for the residues with significantly different conformations confirms that the NMR structures represent better its solution structure in the reduced form. Some of the localised differences, such as a reorientation of Thr24, are thought to be state-dependent changes that involve alterations in hydrogen bond networks. An important rearrangement in the vicinity of the propionate groups of haem I and involving the covalent linkage of haem II suggests that this is the critical region for the functional cooperativities of this protein.

© 1998 Academic Press

Keywords: multihem cytochrome; NMR structure; ring current shifts; redox-Bohr effect; cooperativity

*Corresponding author

Introduction

The functional cooperativity between different regions of certain proteins is a fundamental prop-

erty for controlling and coordinating important chemical events in the living cell. Although the molecular basis for the fine regulation of several types of cooperativity mechanisms has been

Abbreviations used: DvHc₃, *Desulfovibrio vulgaris* (Hildenborough) cytochrome c_3 ; NMR, nuclear magnetic resonance; 2D, two-dimensional; 2D-¹H-NMR, two-dimensional proton NMR; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser spectroscopy; COSY, correlation spectroscopy; DQF-COSY, double-quantum filtered COSY; TOCSY, total correlation spectroscopy; TPPI, time-proportional phase incrementation; ppm, parts per million; DSS, 2,2-dimethyl-2-silapentane-5-sulfonate; MD, molecular dynamics; RMSD, root-mean-square deviation; lol, lower distance limit; upl, upper distance limit.

E-mail address of the corresponding author: xavier@itqb.unl.pt

successfully established (Perutz, 1990), including redox-linked regulation in enzyme catalysis (Williams *et al.*, 1997; Fülöp *et al.*, 1995; Fukuyama *et al.*, 1995; Chen *et al.*, 1994), no unequivocal structural basis for the cooperativity involved in electronic to protonic energy transduction has been established. This is still the case for the tetrahaem cytochromes c_3 from *Desulfovibrio* spp., even though considerable progress has been made in rationalising their thermodynamic properties (Santos *et al.*, 1984; Coletta *et al.*, 1991; Turner *et al.*, 1994, 1996; Park *et al.*, 1996; Louro *et al.*, 1996; Salgueiro *et al.*, 1997) and high-resolution X-ray structures are available for several cytochromes c_3 in the oxidised form (Higuchi *et al.*, 1984; Czjzek *et al.*, 1994; Matias *et al.*, 1996; Simões *et al.*, 1998).

Tetrahaem cytochrome c_3 from *Desulfovibrio vulgaris* (Hildenborough: DvHc₃) is a small (14 kDa; 107 amino acid residues) soluble protein whose crystal structure in the oxidised state has been determined by X-ray diffraction (Matias *et al.*, 1993; Simões *et al.*, 1998). Previous studies have shown that the four haem groups form a complex network of interacting redox centres, resulting in homotropic or redox cooperativity between their redox potentials (Turner *et al.*, 1996). The potentials are also dependent on solution pH, implying heterotropic or redox-Bohr cooperativity between their haem redox and protonation energies (Moura *et al.*, 1982; Turner *et al.*, 1996). Together, these properties give the protein the ability to transfer two electrons and two protons in a concerted step. When coupled to the oxidation of molecular hydrogen catalysed by hydrogenase, this mechanism results in transduction of electronic to protonic energy achieved in the absence of a membrane confinement through a mechanism described as a proton thruster (Louro *et al.*, 1996, 1997).

The present study is part of a project aiming to understand the structural basis for the network of cooperativities in this protein. It is necessary, therefore, to obtain high-resolution structures of at least the fully oxidised and the fully reduced states of the protein in different protonation states in order to identify conformational changes. NMR is a particularly suitable technique for obtaining these structures, since variation of experimental conditions (pH, oxidation state) is usually straightforward. Since the process of assignment and structure calculation is better established for diamagnetic proteins, we decided to determine the solution structure of the reduced state before proceeding to the oxidised state. A pH of 8.5 was chosen in order to ensure that the cytochrome is in the basic form because the pK_a of the redox-Bohr effect in the fully reduced state of DvHc₃ is 7.4 (Turner *et al.*, 1996). However, the high pH presents several problems for the assignment caused by the higher exchange rate of the labile protons.

Several strategies were employed in order to optimise and evaluate the quality of the structure.

Firstly, the measured NOEs were used to provide both upper distance limits (upls) and lower distance limits (lols). This ensures consistency between the input data and the proposed models, because the allowed range of distances between a pair of protons that give rise to a NOE is usually smaller than that allowed using only van der Waals radii as lols. If only upls are used, calculated structures may have protons in van der Waals contact even if no connectivity is found between them in the NOESY spectrum. The conversion of NOE intensities into distance restraints can never be exact, and is a possible source of distortions. In fact, no matter how precisely the intensities are measured, each pair of protons is affected differently by spin diffusion, additional relaxation mechanisms, and by mobility within the structure. Thus, the apparent cross-relaxation rates fall off more slowly with respect to distance than the limiting case of r^{-6} , which applies to an isolated pair of protons, and the exact function will be different for each pair of protons. Some allowance for spin diffusion may be introduced by reference to relaxation matrix calculations (Pfeiffer *et al.*, 1997), but these calculations are themselves approximate and address only part of the problem of calibration. Alternatively, the NOE intensities may be calibrated with respect to interproton distances in a self-consistent fashion, using a function that decreases more slowly than r^{-6} . Inevitably, some intensities will be larger than expected, so that the upls obtained will be too restrictive and the lols will be excessively loose, and *vice versa* if an intensity is anomalously low. Extreme violations of upl, lol and van der Waals constraints can be avoided by calibrating upls and lols separately, so that they are not excessively restrictive. Clearly, there is an arbitrary compromise to be made between losing definition in the structures (expressed as the RMSD of a family of solutions) and producing structures with large constraint violations, but any distortion introduced by the calibration procedure should be detected by independent evaluation of the structure. However, we find that the experimental errors of the measurements are the major source of anomalies in the calibration, and so relaxation of the constraints is the most appropriate treatment unless individual error bars can be obtained for the NOE intensities.

Secondly, the calculations allowed for flexibility in the haem plane, since it has been found that the haem is non-planar in all highly refined X-ray crystal structures of *c*-type cytochromes and may be described variously as being ruffled or saddle-shaped (Moore & Pettigrew, 1990). To account for this experimental observation, flexibility was incorporated in the haem plane by allowing the torsion angles about the eight covalent bonds linking the pyrrole rings to the meso carbon atoms to vary, while the pyrroles are kept flat and rigid. Finally, the distance restraints were modified to allow for fast-flipping aromatic residues. The procedure adopted is based on the fact that, although the ring

protons in aromatic residues are usually rendered equivalent by rapid 180° flips on the NMR time-scale, the orientations of the ring planes are effectively determined (Wareham *et al.*, 1995a,b). In this way, each NOE can be pseudo-stereospecifically assigned to the corresponding protons in a manner analogous to that used for diastereotopic pairs, and the distance restraints to the aromatic residues can be made more restrictive.

Finally, in addition to standard tests such as Ramachandran plots, the structures calculated using this methodology are subjected to rigorous evaluation with respect to the large observed ring current shifts.

Results and Discussion

Sequence-specific assignment

Although DvHc₃ (107 amino acid residues and four haem *c*) is quite large for analysis using only 2D- ^1H -NMR experiments, the extent of spectral overlap is reduced because the haem groups give rise to large ring current shifts for many of the residues. Therefore, a near-complete proton assignment was possible without recourse to isotopic labelling. Amino acid assignment was performed using the classical approach described by Wüthrich (1986), assisted by the use of ring current shifts calculated for the X-ray structure of the ferricytochrome (Simões *et al.*, 1998). Firstly, the spin-systems were identified through analysis of the

TOCSY and COSY spectra in $^2\text{H}_2\text{O}$ and H_2O at different temperatures. Next, sequence-specific assignment was performed by analysing the NOESY spectra and identifying connectivities between NH protons and between the NH and H^α protons of different spin-systems. The sequential connectivities involving NH, H^α and H^β protons observed in the NOESY spectrum recorded in water at 303 K with 100 ms mixing time are represented schematically in Figure 1. The sequential assignment was quite straightforward for some sections of the sequence because of the high number of sequential backbone connectivities. These include residues 1, 2, 4 to 26, 28 to 37, 42 to 53, 62 to 70, 75 to 87, 93 to 99 and 102 to 107. In these segments, all of the residues that have observable NH protons in the experimental conditions show at least one of the sequential connectivities to the NH, H^α or H^β of the preceding residue. For the Pro residues in these segments, NOEs were always observed at least between the Pro H^δ and the H^α and H^β of the previous residue, and between the NH of the following residue and the Pro H^α and H^β .

The assignment of the remaining residues was seriously complicated by the absence of some of the NH resonances at high pH. It was useful to refer to spectra recorded at lower pH (pH 7.3), since the NH exchange rate is reduced and thus it was possible to observe some of the missing NH signals. In fact, 85% (88 out of 103) of the NH protons were observed at pH 8.5, while 91% (94 out of

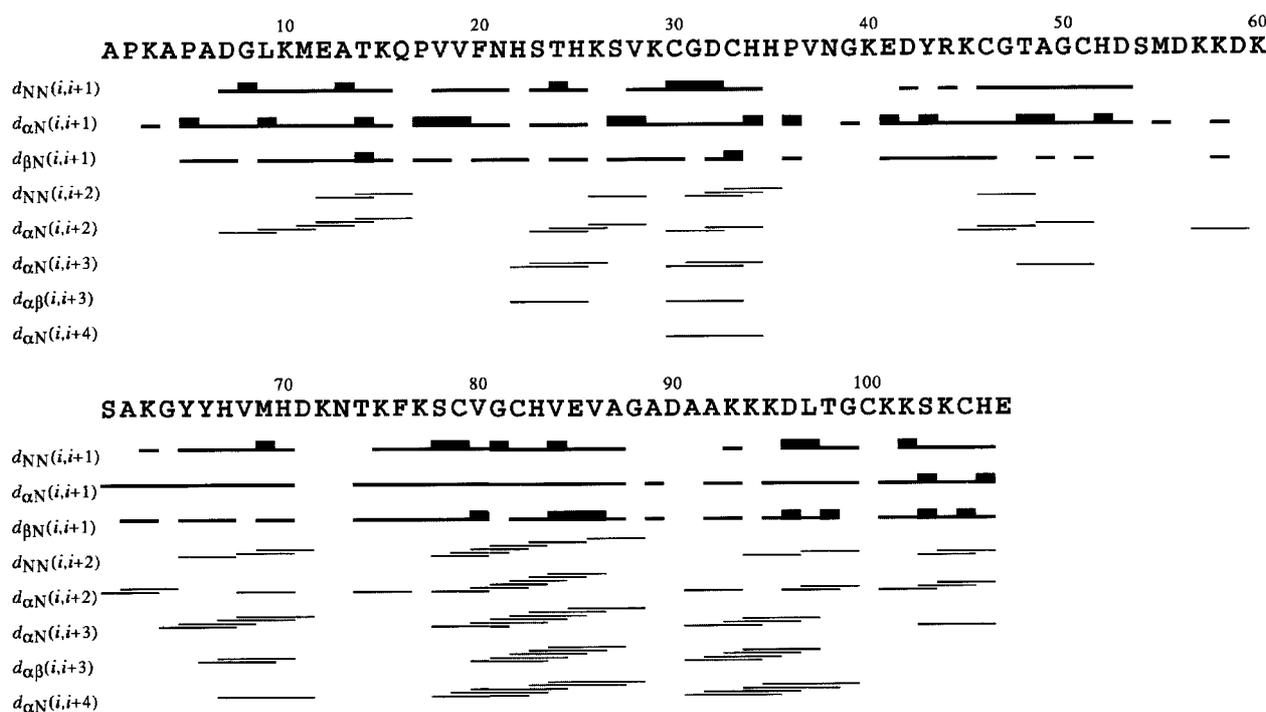


Figure 1. Sequential connectivities involving NH, H^α and H^β protons in fully reduced cytochrome c_3 from *Desulfovibrio vulgaris* (Hildenborough) observed in the NOESY spectrum recorded in water at 303 K with 100 ms mixing time. The line thickness is indicative of the NOE intensity.

103) were observable at pH 7.3. About 20% of the NH protons give rise to NOEs that are significantly stronger at pH 7.3 than at pH 8.5. A new set of sequential connectivities was observed between the residues in the fragments 58-62, 90-93 and 100-102. Some of the NH protons remained unobservable even at pH 7.3, and the assignment of these residues was based on connectivities to other assigned residues, in particular to the NH of the following residue where possible, and analysis of the preliminary calculated structures. This was the case for the remaining nine residues.

Apart from the propionate groups, the specific assignment of the four haem groups in the ferrocyclochrome has been published (Turner *et al.*, 1992). All of the resonances of the propionate protons have now been assigned, thus completing the assignment of the haem protons.

In total, 89% of all the protons in this protein were assigned. This corresponds to 95% of the protons after excluding exchangeable protons other than backbone NH. Of the total number of observed cross-peaks, 96% and 98% were assigned in the NOESY spectra recorded in H_2O and 2H_2O , respectively.

Solution structure determination

A total of 3307 assignments was obtained for cross-peaks in the H_2O and the 2H_2O NOESY spectra. Of this total, some could not be integrated because of excessive overlap and the measured integrals were converted into distance restraints, resulting in 2675 upls and 3114 lols. The number of lols is larger than the number of upls because it is impossible to obtain an upl from strongly overlapping NOEs, but the total volume of each cluster of overlapped peaks can be used to provide a lol for each of the peaks present in the cluster. Taking the volumes to be proportional to the inverse fourth power of the distance between the protons was found to be best suited for calibration using the program CALIBA. There is no sharp distinction between proton pairs that are "rigid" or "mobile", e.g. side-chains in the core of the protein may be less mobile than the backbone in exposed loops, and this function gave the best overall fit. Furthermore, we found no significant advantage in using additional conversion functions, other than for methyl groups. The calibration for upper and lower limits was made as restrictive as possible without giving rise to consistent and inexplicable violations larger than 0.2 Å. Even so, a few lols that involve partially saturated exchangeable protons were found by DIANA to be too restrictive, and for that reason they were not used in the structure calculations. Of all the calculated limits, 2289 upls and 2390 lols were found to be meaningful by DIANA; the remaining restraints were rejected either because they gave no additional structural restraint or because the distance

between the protons is fixed. A total of 69 $^3J_{HNH\alpha}$ values was obtained experimentally, of which 63 were found to provide additional restraints on the ϕ torsion angles by the program HABAS.

These restraints were used in the structure calculations with the program DYANA and each batch of calculated structures was checked for the existence of short distances (less than 2.5 Å) between protons for which no NOE had yet been measured. If the expected NOE was not visible in a clear region of the spectra, a lol of 2.5 Å was introduced. This procedure led to the introduction of a further 50 lols and ensures that the calculated structures are consistent with the complete spectra and not simply with the cross-peaks that were assigned and integrated at the outset.

From the analysis of calculated structures using the program GLOMSA, 88 stereospecific assignments were generated out of the 118 stereopairs with non-degenerate chemical shifts (74.6%) and 115 out of the 184 NOEs to the fast-flipping aromatic residues (62.5%) were pseudo-stereospecifically assigned to one or other side of the ring.

The number of NOE restraints per residue is shown in Figure 2 and detailed statistics of all experimental restraints derived from NMR data are given in Table 1. An average of 42 NOE restraints per amino acid residue (20 upls and 22 lols) and 389 per haem residue (188 upls and 202 lols) was used. It is commonly thought that at least 20 distance restraints per residue are necessary to obtain highly resolved NMR structures (Cavanagh *et al.*, 1996). Zhao & Jardetzky (1994) have demonstrated that, although the precision of NMR structures depends largely on the number of restraints used in the calculation, it is relatively insensitive to the quality of the data. However, the accuracy of NMR structures does depend on the accuracy and precision of the input data. In this work, therefore, the largest possible number of precisely defined restraints has been used.

Six hundred DYANA conformers were calculated with the experimental restraints described above plus an additional 76 upls to fix the haem geometry. Hydrogen bond restraints were not used in the structure calculations. Of the 600 calculated structures, the 20 structures with lowest target function values were selected as being representative of the solution structure of the protein. The atomic coordinates of these structures and the restraints used in structure calculation have been deposited at the Brookhaven Protein Data Bank and have been assigned the codes 1A2I and R1A2IMR, respectively.

A total of 96 hydrogen bonds was identified in the family of 20 DYANA conformers with the program WHAT IF (Vriend, 1990). Of these, 44 were present in at least 50% of the structures. The identified hydrogen bonds correlate, in general, with slowly exchanging protons in the 2H_2O sample.

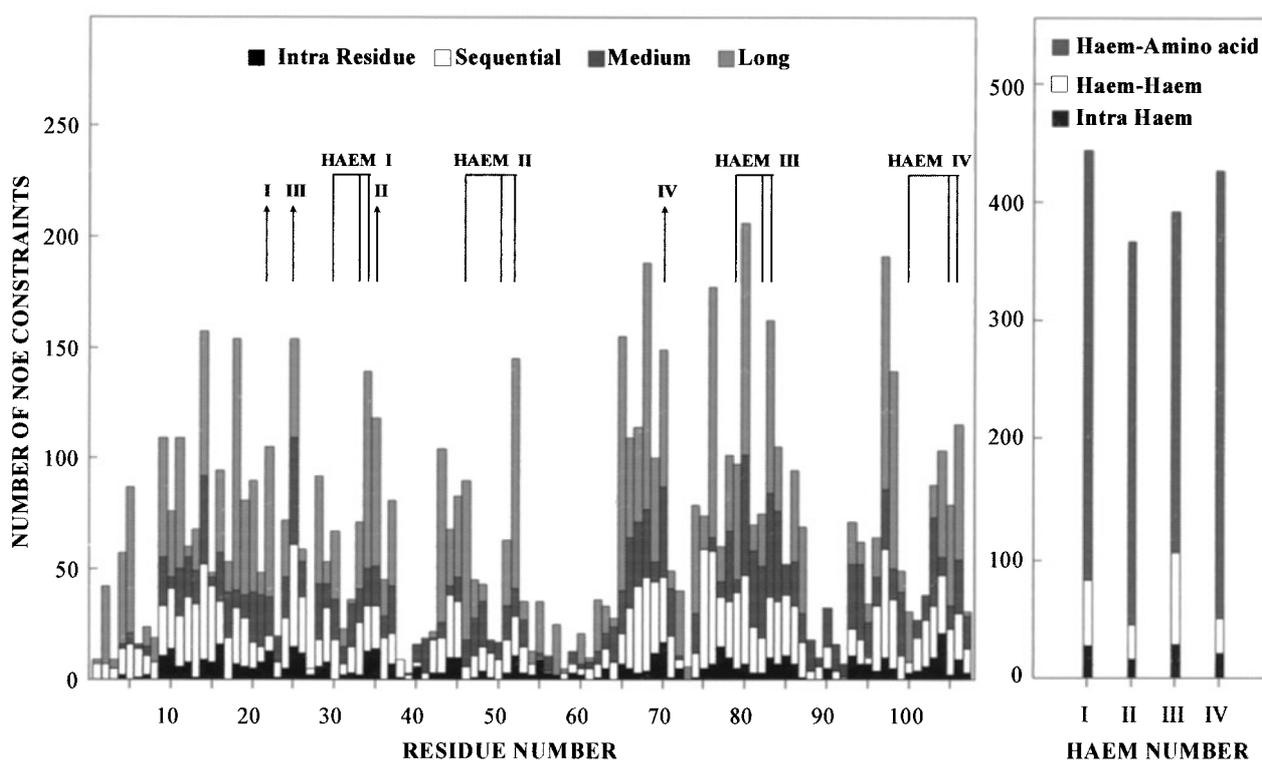


Figure 2. Number of NOE restraints (upls plus lols) per residue used in the structure calculation. Note that each interresidue restraint is included for both of the residues involved. The haem-binding sites are indicated, the sixth axial His ligand being marked by an arrow.

Quality analysis of the NMR structure models

Violation of structural restraints

The target function within the ensemble of the 20 best DYANA conformers ranges from 0.77 \AA^2 to 0.93 \AA^2 for the individual conformers (average value 0.89 \AA^2 , range 20.9% from the lowest value). The value of the target function for each structure is represented in Figure 3, together with the sum of the upls, lols, torsion angle and van der Waals vio-

lations as a function of the structure number. Note that the only term included in the DYANA force field is a simple van der Waals repulsion and, hence, this term alone represents the intrinsic energy of the structure. There is a steep increase in the target function between structures 3 and 4, but the increase is smooth thereafter, which shows that the convergence of the structures to the global minimum is not ideal, in spite of using torsion angle dynamics. Nevertheless, the sum of the violations of all types of restraints increases only slightly across the ensemble of selected structures, and there is no consistent difference between the violations found in the best and worst of these structures. The complete statistics of violated restraints are given in Table 2.

Table 1. Restraints used for structure calculation

Distance restraints	Upper distance limits	Lower distance limits
Amino acid residues		
Intra-residual	305	238
Sequential ($ i - j = 1$)	433	479
Medium-range ($2 \leq i - j < 5$)	360	415
Long-range ($ i - j \geq 5$)		
Amino acid-amino acid	440	476
Amino acid-haem	652	695
Subtotal	2190	2303
Haem residues		
Intra-residual	51	40
Haem-haem	48	47
Subtotal	99	87
Total	2289	2390
Phi torsion angle restraints		
Total	63	

Precision of the calculated family of structures

The structures taken to represent the solution structure were superimposed using all heavy atoms. These structures are closely similar, with an average RMSD relative to the mean coordinates of 0.35 \AA for the backbone atoms (N, C^α and C') and 0.70 \AA for all heavy-atoms (Table 2). The precision of the structures was evaluated using two different approaches: RMSDs per residue with respect to the mean structure for global precision (Figure 4) and angular order parameters for local precision (data not shown).

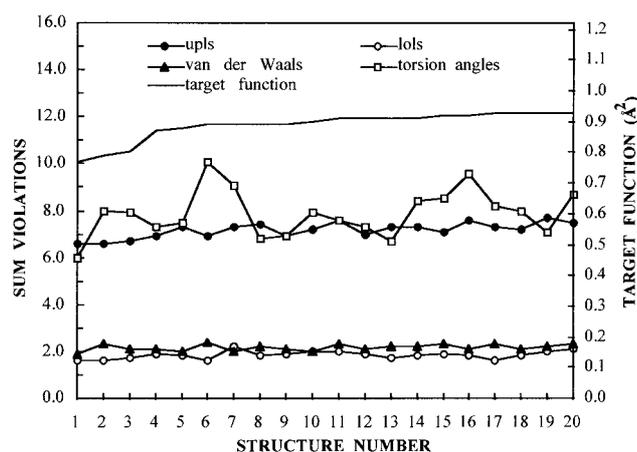


Figure 3. Target function (right-hand scale) and sums of the upls, lols, van der Waals (left-hand scale, Å), and torsion angle violations (left-hand scale, degrees) as a function of the structure number for the best 20 DYANA conformers.

A detailed analysis of the precision within the family of DYANA conformers reveals that several regions are very well defined, with backbone

RMSDs smaller than 0.3 Å and angular order parameters for ϕ and ψ torsion angles larger than 0.9. This is observed for fragments 4-36, 44-53, 63-87 and 93-106. These fragments correspond to regions with defined secondary structure elements (helices and β -sheets) or with haem-binding sites. The remaining residues are located in solvent-exposed loops or in the N and C terminus, and tend to be poorly defined, with large RMSDs and small angular order parameters for the backbone. In general, the number of NOE restraints for these residues is small (Figure 2) and the number of unobservable NH protons is high (Figure 4).

Interestingly, there are two residues in one of the poorly defined regions (Met55 and Lys58) for which the backbone RMSD is quite high but the angular order parameters for the backbone are also high. It appears that the local conformation of these residues is preserved despite forming part of a flexible loop. Methionine 55 is particularly notable, because all of the angular order parameters for the side-chain torsion angles are greater than 0.9. In another region, Lys40, although having a very high global RMSD, also maintains its backbone conformation, since the angular order parameters for the backbone torsion angles are very high.

Table 2. Comparative analysis of the best 20 DYANA conformers calculated with both upls and lols and the best 16 conformers calculated without lols

Quantity	20 Conformers calculated with both upls and lols Average value \pm SD	16 Conformers calculated without lols Average value \pm SD
DYANA target function (\AA^2)	0.89 \pm 0.05	0.58 \pm 0.03
Residual distance restraint violations		
Upper distance limits		
Sum (\AA)	7.2 \pm 0.3	5.6 \pm 0.2
Maximum (\AA)	0.20 \pm 0.01	0.20 \pm 0.01
Number of violations ($>0.2 \text{\AA}$)	0 \pm 1	0
Number of consistent violations ($>0.2 \text{\AA}$)	0	0
Lower distance limits		
Sum (\AA)	1.8 \pm 0.2	–
Maximum (\AA)	0.17 \pm 0.02	–
Number of violations ($>0.2 \text{\AA}$)	0	–
Residual dihedral angle restraint violations		
Sum (\AA)	7.9 \pm 1.0	7.2 \pm 1.0
Maximum (\AA)	2.2 \pm 0.4	2.4 \pm 0.4
Number of violations ($>5^\circ$)	0	0
Residual van der Waals violations		
Sum (\AA)	2.2 \pm 0.1	1.9 \pm 0.1
Maximum (\AA)	0.10 \pm 0.01	0.11 \pm 0.02
Number of violations ($>0.2 \text{\AA}$)	0	0
Average RMSD to the mean structure		
Backbone	0.35 \pm 0.09	0.46 \pm 0.09
All heavy atoms	0.70 \pm 0.07	0.82 \pm 0.08
Ramachandran plot	% Residues ^a	% Residues
Residues in most favoured regions	67.0 (73.0)	63.7
Residues in additional allowed regions	27.8 (26.6)	30.7
Residues in generously allowed regions	2.9 (0.4)	2.0
Residues in disallowed regions	2.3 (0.0)	3.6

The number of conformers within the family calculated without lols was calculated such that the range of target function variation is 20.9%, the range found within the family calculated with both upls and lols.

^a In parentheses is indicated the percentage of residues with $S(\phi)$ and $S(\psi) > 0.9$.

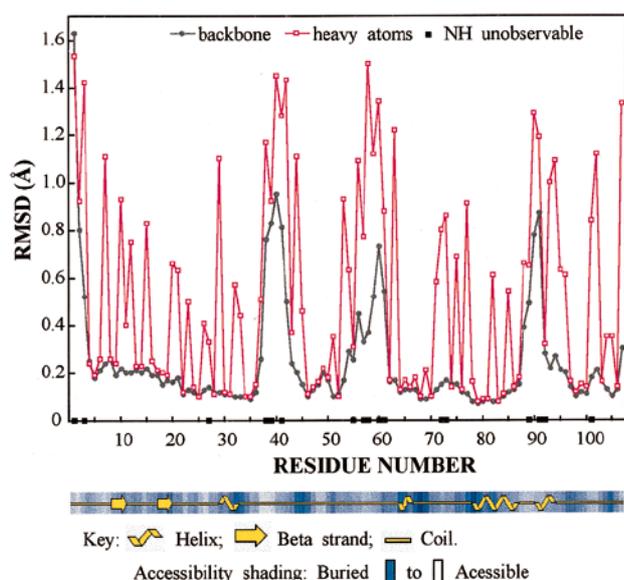


Figure 4. Average backbone and heavy-atom RMSD values per residue with respect to the mean structure of the best 20 DYANA conformers. Consensus secondary structure elements and average estimated accessibility per residue were calculated with the program PROCHECK-NMR (Laskowski *et al.*, 1996). Unobservable NH protons are represented by black squares.

Agreement with the experimental data

The agreement of the NMR structures with the experimental data may be evaluated by different approaches. Firstly, they may be analysed in terms of violation of structural restraints, as discussed above. Provided that the assignments are correct, a strictly limited number of small violations indicates a high level of accuracy. Secondly, a peak list may be produced for all the protons that are closer than some chosen value and compared with the experimental peak list. A value of 2.5 Å was chosen, which is small enough to allow for possible damping of the NOE intensities due to mobility or saturation transfer from the solvent. The agreement between the two peak lists was very good even with the initial structures, and the 50 new lols introduced to account for “invisible” NOEs gave further improvement (see Calculation of restraints in Materials and Methods). Thirdly, the calculated ring current shifts for the NMR structures may be compared with the observed ring current shifts to confirm assignments. Figure 5(a) illustrates the fact that the calculated ring current shifts for each proton in the best NMR structure agree very well with the observed ring current shifts. RMSD values were calculated between the observed ring current shifts of all protons of one particular residue and the ring current shifts calculated from the NMR structure for all residues of the same type; overall, the smallest RMSD was found for the corresponding residue in the structure. Furthermore, ring cur-

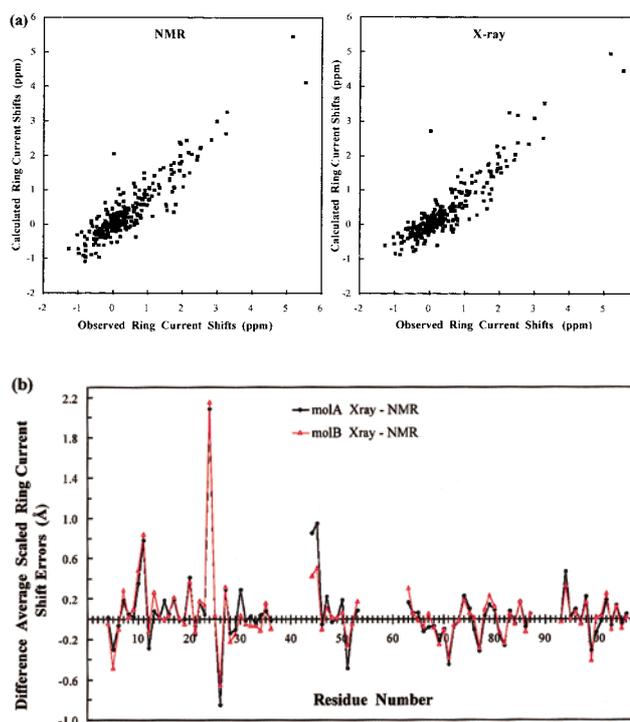


Figure 5. (a) Calculated ring current shifts for the best NMR structure and model A of the X-ray structure *versus* observed ring current shifts for each proton. (b) Difference in the average scaled ring current shift errors for each residue between the best NMR structure and molecules A and B of the X-ray structure. Only well-defined residues, or parts of the residues, were considered in these calculations, since the best NMR structure may not be representative for less well defined regions.

rent shift calculations can be used to provide an independent test of differences found between the conformations of some residues in NMR and in the X-ray structures (Figure 5(b)).

Ramachandran plots

Another criterion for the analysis of the quality of a structure is the interpretation of the deviation of the backbone torsion angles from ideality using a Ramachandran plot (Ramachandran *et al.*, 1963). The Ramachandran plot presented here (Figure 6) was produced by the program PROCHECK-NMR. The analysis performed by this program considers all non-Gly, non-Pro and non-terminal residues. According to this analysis, 67.0% of the residues are located in the most-favoured regions, 27.8% in additional allowed regions, 2.9% in generously allowed regions, and 2.3% in disallowed regions (Table 2). The introduction of flexibility in the ω peptide angle in the range suggested by MacArthur & Thornton (1996) did not improve the quality of the Ramachandran plot or reduce the number of residues in the disallowed regions

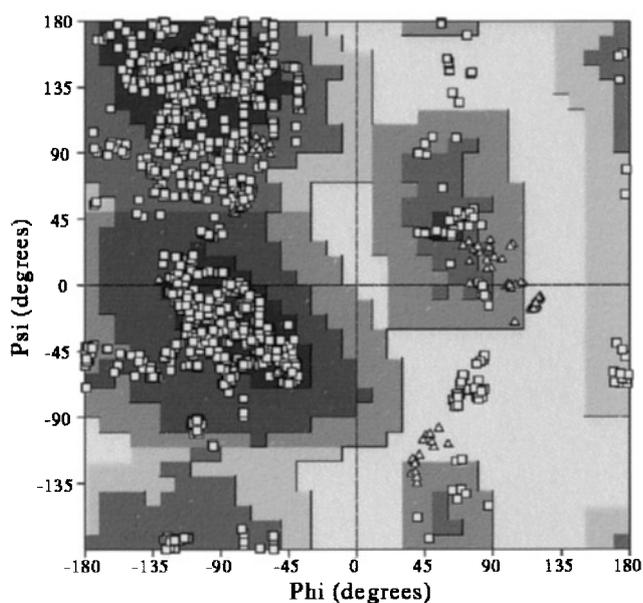


Figure 6. Ramachandran plot for the ensemble of the best 20 DYANA conformers. All the residues except the terminal residues are represented. The Gly residues are represented by triangles.

(results not shown). The residues falling in the disallowed regions do not have many NOE restraints, and for some of them the NH is not observable, which reduces the number of backbone restraints even further. The quality of the plot improves if only those residues with both $S(\phi)$ and $S(\psi)$ larger than 0.9 are included (Table 2), in which case no residue is found in disallowed regions and only 0.4% of these residues appear in the generously allowed regions.

It is noteworthy that there are a few residues that have very high angular order parameters and do not fall in the most-favourable regions. These residues are mainly haem ligands. A PROCHECK-NMR analysis of haem ligands reveals that 49.1% of these residues are located in the most-favoured regions, 49.4% in additional allowed regions, 1.6% in generously allowed regions, and none in disallowed regions. Similar behaviour was observed for these residues in the X-ray structure (Simões *et al.*, 1998), and we conclude that these local distortions of the backbone dihedral angles are a normal consequence of haem ligation (Gunasekaran *et al.*, 1996).

Tests of the methodology

Introduction of lower distance restraints

Three hundred structures were calculated with the same input as the structures calculated with flexible haem groups but excluding lower distance restraints (lols). The convergence of the structures calculated in this fashion is considerably better

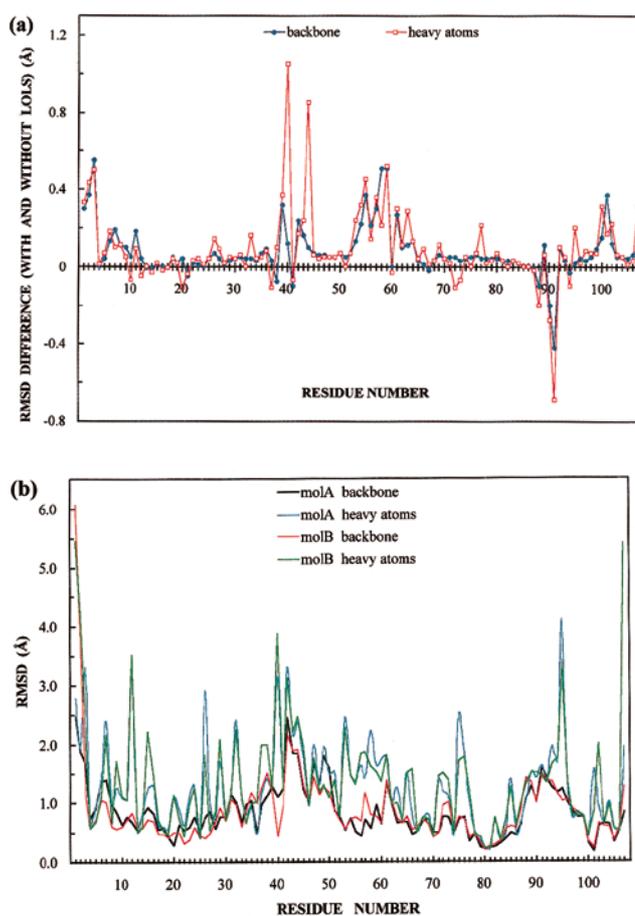


Figure 7. (a) Difference of the global RMSD per residue for the backbone atoms and for all heavy-atoms between the family calculated with only ups and the family calculated with both ups and lols. (b) RMSD per residue between the mean NMR structure and both X-ray structures.

than when lols are used, and 40,000 steps were sufficient in MD simulation. An ensemble of 16 structures was chosen for the comparison, since it has the same range of variation in the target function as the 20 best conformers calculated with lols, i.e. 20.9% (Table 2). As expected, the average target function value within the family calculated without lols is smaller than the corresponding value with lols.

The RMSD with respect to the mean structure decreases from 0.46 Å to 0.35 Å and from 0.82 Å to 0.70 Å for the backbone and heavy atoms, respectively, with the inclusion of lols. Figure 7(a) displays the difference of the global RMSD per residue between the two families of structures. The inclusion of lols gives the largest decrease in global RMSD in the N-terminal region (residues 1 to 3), residues 39, 40, 44, 54, 55, 58, 59, 101 and the C-terminal residue. The improvement in global precision is at least 0.3 Å for these residues and, in general, corresponds to an increase in the local pre-

Table 3. Comparative analysis of the best 50 DYANA conformers calculated with flat, rigid haem groups and the best ten DYANA conformers calculated with flexible haems

Quantity	50 Conformers with flat rigid haem Average value \pm SD	10 Conformers with flexible haem Average value \pm SD
DYANA target function (\AA^2)	1.58 \pm 0.06	0.86 \pm 0.05
Residual distance restraint violations		
Upper distance limits		
Sum (\AA)	10.1 \pm 0.3	7.0 \pm 0.3
Maximum (\AA)	0.32 \pm 0.01	0.19 \pm 0.01
Number of violations ($>0.2 \text{\AA}$)	5 \pm 1	0 \pm 1
Number of consistent violations ($>0.2 \text{\AA}$)	2	0
Lower distance limits		
Sum (\AA)	2.5 \pm 0.2	1.8 \pm 0.2
Maximum (\AA)	0.20 \pm 0.05	0.18 \pm 0.03
Number of violations ($>0.2 \text{\AA}$)	0 \pm 1	0 \pm 0
Number of consistent violations ($>0.2 \text{\AA}$)	0	0
Residual dihedral angle restraint violations		
Sum (\AA)	8.7 \pm 1.3	7.7 \pm 1.1
Maximum (\AA)	2.3 \pm 0.2	2.3 \pm 0.4
Number of violations ($>5^\circ$)	0	0
Residual van der Waals violations		
Sum (\AA)	3.0 \pm 0.2	2.1 \pm 0.1
Maximum (\AA)	0.11 \pm 0.01	0.10 \pm 0.01
Number of violations ($>0.2 \text{\AA}$)	0	0
Average RMSD to the mean structure		
Backbone	0.37 \pm 0.07	0.35 \pm 0.09
All heavy-atoms	0.76 \pm 0.07	0.67 \pm 0.07
Ramachandran plot		
Residues in most-favoured regions (%)	67.2	67.3
Residues in additional allowed regions (%)	27.6	28.2
Residues in generously allowed regions (%)	3.5	2.6
Residues in disallowed regions (%)	1.7	2.0

The number of conformers within each family was calculated such that the range of target function variation was the same

cision, measured by angular order parameters (results not shown). The largest decreases in the global RMSD occur for the heavy atoms of Lys40 (1.05 \AA) and Arg44 (0.85 \AA), whereas their backbone RMSD values remain essentially unaltered; these residues have a relatively small number of experimental restraints. However, for residues 88, 90 and 91, the effect of the inclusion of lols is to increase the global RMSD for both the backbone atoms and all heavy-atoms and, concomitantly, the angular order parameters decrease. Here, the lols seem to force the structure into several discrete conformations. The increased precision of the structure obtained with the addition of lols is to be expected, particularly in regions in which there are few restraints, but is not, in itself, a demonstration that the structure is more accurate. However, the use of lols clearly improves the definition of the structure and, hence, it can be evaluated more rigorously.

A total of 38 hydrogen bonds was identified in at least 50% of the conformers family calculated without lols, which is to be compared with 44 found when lols were included. In general, the additional hydrogen bonds correlate with the regions in which the inclusion of lols gives the largest improvement in precision. The structures calculated without lols included over 50 examples of short interproton distances that should have pro-

duced observable NOEs; in fact, no measurable intensity was found in these cases and so the structures obtained without lols are fundamentally incompatible with the experimental data. The dihedral angle distribution in the Ramachandran plot also improves when lols are included in the structure calculation (see Table 2).

In conclusion, the careful introduction of lols improves the quality of the structures in terms of agreement of the models with the input experimental data, also reducing the deviation of the distribution of dihedral angles from the ideal values, and increasing the global and local precision. It should be noted that this comparison may have underestimated the advantage of using lols, since several of the stereospecific assignments used for these calculations could not have been obtained without using lols to provide restraints for overlapping cross-peaks, including the degenerate signals of fast-flipping aromatic groups. However, these stereospecific assignments were used in the structure calculation without lols.

Introduction of flexible haem groups

A further 300 structures were calculated with flat, rigid haem groups in order to test the effect of introducing flexible haem groups into the structure calculation. A comparison of the most relevant

indicators of structure quality is presented in Table 3.

The convergence of the structures using flat, rigid haem groups is considerably better, presumably because of the reduced number of torsion angles. Although only 300 structures were calculated, the 50 best structures had target functions that covered a range of only 16.9%, which is the range covered by the ten best structures calculated using flexible haem groups. However, the increase in the absolute value of the minimum target function is much greater than might be expected for the loss of 32 variable torsion angles. This indicates that the distortion of the haem planes is significant.

Removing the flexibility of the haem groups results in an increase in the number of upl violations (Table 3). Whereas the structures calculated with flexible haem groups show no consistent upl violations larger than 0.2 Å, the structures with flat, rigid haem groups show two such violations, both involving haem protons.

The global precision of the structures calculated with flexible haem groups is similar to the precision of those calculated with rigid haem groups but the local precision is slightly worse if the haem molecules are kept rigid. In fact, the backbone angular order parameters of two residues and the side-chain angular order parameters of eight residues that were above 0.9 in the structures with flexible haem groups, dropped below 0.9 when the haem groups were kept flat (data not shown).

Despite the computational difficulties, the observation that haem groups are, in general, distorted in high-resolution X-ray structures shows that any solution structure of a haem protein that aspires to high resolution must account for flexibility in the

haem plane. Furthermore, the marked reduction in target functions obtained with flexible haem groups indicates that the data used here are sufficiently precise to be sensitive to such distortions, which is discussed below.

Description of the NMR structure models

Figure 8 depicts the backbone and the haem groups of the best 20 DYANA conformers of DvHc₃. Both the backbone and the haem molecules, including the propionate groups, are, in general, very well defined. The propionate groups with greater conformational variability in the ensemble are those most exposed to the solvent.

The N-terminal region (residues 1 to 3) is highly variable in the NMR structures, probably because of the high solvent-accessibility (Figure 4), which allows rapid amide proton exchange, thus decreasing the signal intensity and reducing the number of restraints. All the residues in this region present long-range NOEs. A two-stranded antiparallel β -sheet was found between the segments Leu9 to Met11 and Val18 to Phe20. The segment Leu9 to Phe20 forms a class 6:6 β -hairpin (Sibanda *et al.*, 1989). Although the loop of this hairpin contains six residues, it is well defined, probably due to the proximity of haem IV, which restricts the number of possible conformations. The ring proton resonances from Phe20 were broadened, approaching the very slow flipping regime in our experimental conditions, which is not surprising because the ring is inserted between the planes of haem I and the axial His25 of haem III. The segment defined by residues 21 to 28 forms a composite β -turn that is well defined, even though three of the residues

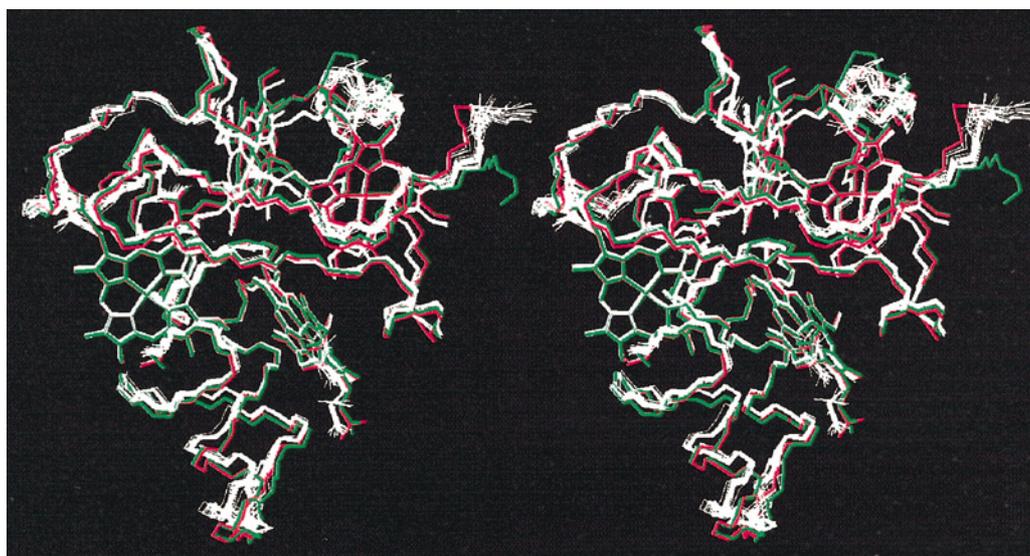


Figure 8. Stereo view of the backbone and haem groups of the 20 NMR structures of the ferrocyclochrome at pH 8.5 (white) and of the available X-ray structures of the ferricytochrome at pH 5.5 (model A red; model B green: Simões *et al.*, 1998). Superimposition was performed using all heavy-atoms. In this orientation haem II is at the top, haem III at the bottom, haem IV at the left and haem I at the right.

(Ser23, Lys26 and Ser27) have high solvent-accessibility (Figure 4). The restriction in the conformation of the residues in this loop is probably related to the fact that two of them, His22 and His25, are axial ligands of haem I and haem III, respectively. This loop is followed by an α -helix comprising residues 29 to 34 (helix I), for which all of the expected hydrogen bonds were identified. This helix includes the remaining three ligands of haem I (Cys30, Cys33 and His34), followed by His35, an axial ligand of haem II. Residues 36 to 40 are located in a region with high solvent-accessibility and, with the exception of Pro36, are not very well defined. This stretch forms a composite β -turn. A region of poor definition follows, ending at Arg44, which contains the least well defined aromatic residue, Tyr43. Residues 45 to 53 form another well-defined composite turn (Figure 4). This loop is located between haem I and haem II, close to the propionate groups of haem I and includes the other ligands of haem II (Cys46, Cys51 and His52). Residues 56 to 62 form a composite turn that is badly defined in the NMR structures because of a lack of restraints. This could result from high solvent-accessibility, which explains the large number of unobservable NH protons in this segment.

Segment 64 to 68 defines a 3_{10} -helix (helix II) for which the expected hydrogen bonds are observed. Residues 65 and 66 are tyrosine, which are fast-flipping on the NMR timescale but have well-defined orientations. The region defined by residues His67 to Lys77, which includes His70, one of the axial ligands of haem IV, is precisely defined and forms a composite turn. Residues Thr74, Lys75 and Phe76 are close to the propionate groups of haem II, and are stabilised by a network of hydrogen bonds to the haem propionate 17. Residues 78 to 88 form an α -helix (helix III), which is one of the best defined regions in this structure and includes the other ligands of haem III (Cys79, Cys82 and His83). A highly disordered, solvent-exposed loop follows that comprises residues 88 to 92. Residues 88 to 90 define a type IV β -turn, and residues 91 and 92 form a classic γ -turn. Segment 92-94 presents an α -helical conformation (helix IV) and the remainder of the sequence forms a long, well-defined turn that includes the other ligands of haem IV (Cys100, Cys105 and His106).

Comparison with X-ray structure models of the ferricytochrome

Comparison of the NMR structures of the ferrocyanochrome at pH 8.5 with the available X-ray structures of the ferricytochrome at pH 5.5 (Simões *et al.*, 1998) reveals that the general fold of the molecule as well as the relative position of the four haem groups is closely similar (Figure 8 and Table 4). The global RMSD values of the mean NMR structure to molecule A and molecule B of the unit cell of the crystal are 0.93 Å and 1.08 Å for

Table 4. Iron-iron distances (in Å) in the best DYANA structure and in the X-ray molecules

Haem		II	III	IV
I	NMR	11.9	10.8	17.6
	X-ray mol A	12.4	11.1	17.8
	X-ray mol B	12.4	11.1	17.8
II	NMR		15.3	15.8
	X-ray mol A		16.0	16.7
	X-ray mol B		16.2	16.6
III	NMR			12.0
	X-ray mol A			12.0
	X-ray mol B			12.0

the backbone and 1.47 Å and 1.59 Å for all heavy-atoms, respectively. Most of the consensus secondary structure elements identified in the NMR structures are common to the X-ray structures. Nevertheless, a detailed comparison of the NMR and X-ray structures reveals several differences (Figures 7(b), 8 and 9). Figures 7(b) and 8 show that the major differences in the backbone conformation occur in some of the most exposed and ill-defined regions: the N terminus (residues 1 to 3) and the loops comprising residues 37 to 43 and 88 to 92. However, a number of well-defined residues in the NMR structures exhibit conformations that are significantly different from both X-ray structures.

As a result of the high haem to polypeptide ratio of cytochrome c_3 , the majority of the amino acid protons experiences significant ring current shifts. The average deviation between the experimental and calculated ring current shifts therefore provides an independent method for testing the accuracy of the NMR structure models, although residues far from the haem groups, in particular those on the surface, have small shifts and this method is then less sensitive. It is most unusual to have such an opportunity, and this approach will be used extensively in this section.

The calculated ring current shifts for both the best NMR and X-ray structures are very similar to the observed ring current shifts (Figure 5(a)). Figure 5(b) depicts the difference in the average scaled ring current shift errors for each residue between the best NMR structure and molecules A and B of the X-ray structure. Only the well-defined residues, or parts of the residues, were considered, since the best NMR structure may not be representative for less well defined regions (see Precision of the calculated family of structures, above). This analysis shows that the overall agreement is slightly better for the NMR structure, but the essential purpose of these calculations is to evaluate localised regions in which the structures differ significantly.

In the segment defined by residues 7 to 20, the backbone RMSD between the best NMR and X-ray structures is relatively small, with the exception of Asp7 (Figure 7(b)). The β -hairpin comprising residues 9 to 20 found in the NMR structures exists in

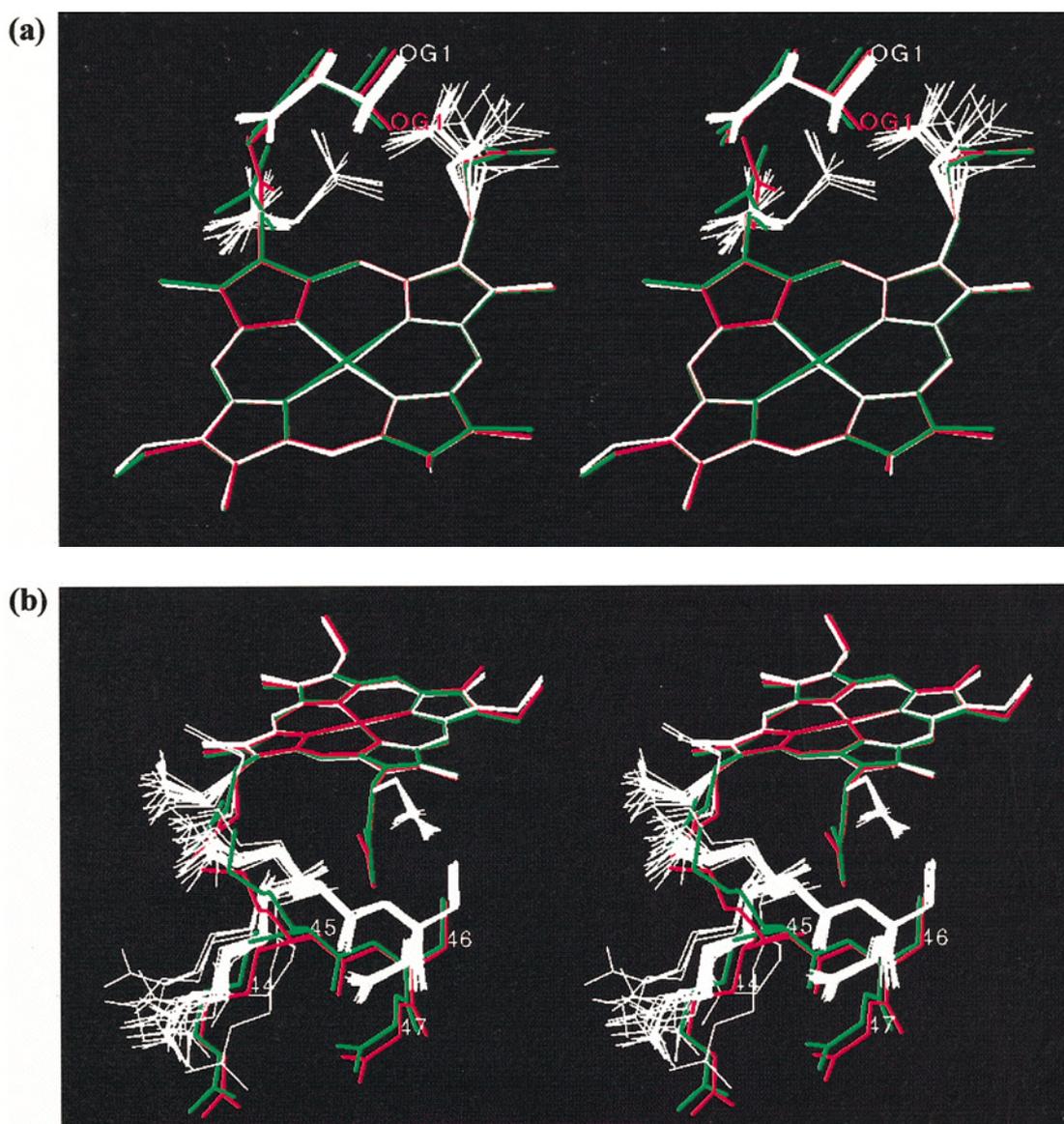


Figure 9. Localised differences between the NMR structures of the ferrocytochrome at pH 8.5 (white) and the X-ray structures of the ferricytochrome at pH 5.5 (model A red; model B green). (a) Stereo view of Thr24 and haem III. In the NMR structures the Thr24 side-chain is rotated by approximately 120° in relation to both X-ray structures. The hydroxyl oxygen atom of Thr24 is indicated in the NMR structure (white) and model A of the X-ray structure (red). (b) Stereo view of Arg44, Lys45, Cys46, Gly47 and haem I. In the NMR structures these amino acid residues are systematically closer to haem I and the propionate groups have different conformations. Superimposition was performed using all the carbon atoms of the haem macrocycle but excluding those of the substituents.

both X-ray models. However, the β -turn comprising residues 14 to 17 is of type VIII in the X-ray structures, whilst in the NMR structures it is of type IV. In the NMR structures, the normal pattern of hydrogen bonding in the β -hairpin is disturbed by the lack of a hydrogen bond between Leu9 NH and Phe20 CO. As a result, Phe20 forms a typical type G1 β -bulge with the dipeptide 8-9. This β -bulge is not observed in the X-ray models. Overall, the ring current shifts of the residues in the segment 7-20 calculated from the NMR structure, in particular those of the residues defining the β -hairpin, agree much better with the observed ring cur-

rent shifts than do those calculated from the X-ray structures.

Concerning the region defined by residues 21 to 28, the backbone RMSD between the best NMR and the X-ray structures is very small. A composite β -turn was identified in the segment 21-25 of the NMR structure but a 3_{10} -helix is found in both X-ray structures. Residue Thr24 is particularly interesting because its side-chain orientation differs by approximately 120° between the NMR and X-ray structures (Figure 9(a)). In the conformation observed in the NMR structures, it is not possible to form three of the hydrogen bonds that involve Thr24 in both X-ray

models: between NH and O γ^1 of Thr24, Thr24 H γ^1 and the carboxylate group of haem III propionate 17, and between His25 H δ^1 and Thr24 O γ^1 . Also, the Thr24 hydroxyl proton was not observed in the NMR spectra, which indicates that it is in fast exchange with the solvent protons and is not hydrogen bonded. Assuming that the X-ray models are an accurate representation of the conformation of Thr24 in the oxidised cytochrome, then the reduced form is destabilised by the loss of three hydrogen bonds. It is possible that the formation/breakage of hydrogen bonds involving either propionate 17 or His25 (an axial ligand of haem III) may be important in controlling the redox potential of haem III in this protein. Oxidation state-dependent conformational changes that involve propionate groups and realignment of the hydrogen bonds formed with it have been observed in several other cytochromes (Moore *et al.*, 1986; Berghuis & Brayer, 1992; Banci *et al.*, 1997a,b). In relation to the ring current shifts of the residue 21 to 28 stretch, two residues have large differences between the NMR and the X-ray structures, Thr24 and Lys26. The calculated ring current shifts are in much better agreement with the Thr24 side-chain conformation in the NMR structure (Figure 5(b)). By contrast, the average scaled ring current shift errors are smaller for the conformation of Lys26 in the X-ray structures, but this residue has large B-factors in both X-ray models and the terminal atoms of the side-chain are not observed in molecule A, so the improved agreement with the X-ray structures may be fortuitous.

The RMSD between the mean NMR structure and both X-ray structures increases slightly in the segment from residue 29 to 36 (Figure 7(b)). The difference is apparent also in the secondary structure elements: residues 29 to 34 form an α -helix in the NMR structures, with an inverse γ -turn comprising residues 33 to 35, while residues 29 to 33 define a 3_{10} -helix with a type IV β -turn for residues 32 to 35 in the X-ray structures. Although the differences are small, the calculated ring current shifts suggest that molecule B in the X-ray structure provides the most accurate representation of this segment.

Analysis of the region comprising residues 44 to 53 reveals a systematic difference between the NMR and the X-ray structures, since the backbone RMSD is generally high (Figure 7(b)). The backbone of residues 44 to 47 is, on average, 2.5 Å closer to haem I in the NMR models of the ferrocyanochrome (Figure 9(b)), and the average scaled ring current shift error is significantly smaller for the NMR structure of the reduced state (Figure 5(b)). This rearrangement does not affect the distance between propionate 17 of haem I and the Lys45 amino group, which remain close to each other. Furthermore, the hydrogen bond between Cys46 NH and carboxylate oxygen atoms of propionate 13 of haem I, which is present in both X-ray structures, is preserved in the NMR

structure through a change in the conformation of the propionate. These conformational modifications are discussed further in relation to the functional cooperativities of DvHc $_3$.

The backbone of residues 63 to 72 in the NMR structures is well defined and has a small RMSD with respect to the X-ray structures. Nevertheless, helix II in the NMR structures, comprising residues 64 to 68, is of the type 3_{10} , whilst in the X-ray structures a longer α -helix includes residues 64 to 71. The calculated ring current shifts show an increasing preference for the X-ray structures, which could be associated with a small change in the relative position of haem II (Table 4) and deserves further discussion.

The haem groups are very well defined in the NMR structures (Figures 8 and 10), although a few of the propionate groups show some conformational disorder that is, in general, related to higher solvent exposure. The iron-iron distances found in the NMR and X-ray structures are similar, but there is a clear tendency to shorter distances in the NMR structures (Table 4). This may be viewed as a movement of haem II and, to a lesser extent, of haem I towards haem III and haem IV, since the distance between these last two is unaltered. Although this could be a real phenomenon, the localised deviations in calculated ring current shifts of residues between haem II and haem IV (65 to 71) suggest that this represents a small distortion in the NMR structure. It is not surprising that the position of haem II in the NMR structures is subject to distortion, since its binding sequence is flanked by two loops that, together with the haem edge from the second Cys residue to propionate 13, are highly exposed to the solvent (Figure 4). Furthermore, on the other edge of the haem, although it has a large number of restraints to the nearby residues, several of the cross-peaks are overlapped, so that some of the stereospecific assignments could not be obtained. In particular, the nearby His67 H δ^2 and H ϵ , Val68 H α and γ protons and Met69 NH, β and ϵ proton resonances are strongly overlapped and, although in some cases it was possible to obtain a upl, this could lead to cumulative errors. Furthermore, the stereospecific assignment of the Val68 methyl γ_1 and γ_2 protons and the corresponding restraints were obtained using spectra at lower temperature, where the methyl groups are non-degenerate. It was also impossible to assign stereospecifically the His67 β protons, Met69 γ protons and Asp71 β protons because they were degenerate, which weakens the restraints to the nearby protons. A distortion of this type may also be viewed as an anisotropic compression of the structure and this might reveal a fundamental limitation in the methodology that would not be apparent in less well resolved structures. Both simulated annealing and direct minimisation produce structures at an implicit temperature of absolute zero, in which even the orientations of the methyl groups are frozen and, hence, the structures are likely to be con-

tracted. It should be noted that minimisation after annealing had a negligible effect on the Fe-Fe distances. It is impractical to generate time averages of restrained molecular dynamics for every structure, but it may be possible to overcome the problem by modifying the form of the repulsive van der Waals interactions. In general, there may also be problems if structural water molecules are not included, but there are only two of these found by X-ray diffraction in the hydrophobic core of DvHc₃.

The region comprising residues 73 to 88, which includes the binding site of haem III, is very well defined in the NMR and the X-ray structures. The backbone conformation of these residues is very similar in the two structures (RMSD < 1 Å), both of which form α -helices involving residues 78 to 88.

The ring current shift calculations show no systematic preference for either structure.

The final stretch of the sequence includes the binding site of haem IV. The RMSD between the X-ray and NMR structures is somewhat larger for the main chain of residues 93 to 97 than in the other well-defined regions and a helix spans residues 90 to 99 in the X-ray structure but includes only residues 92 to 94 in the consensus NMR structure, though this helix extends to residues 90 to 97 in some of the individual structures. The charged residues Lys93 and Lys104 move towards propionate groups of haem III and, again, the observed ring current shifts of these residues are better predicted using the NMR structure conformation than the X-ray structure conformation.

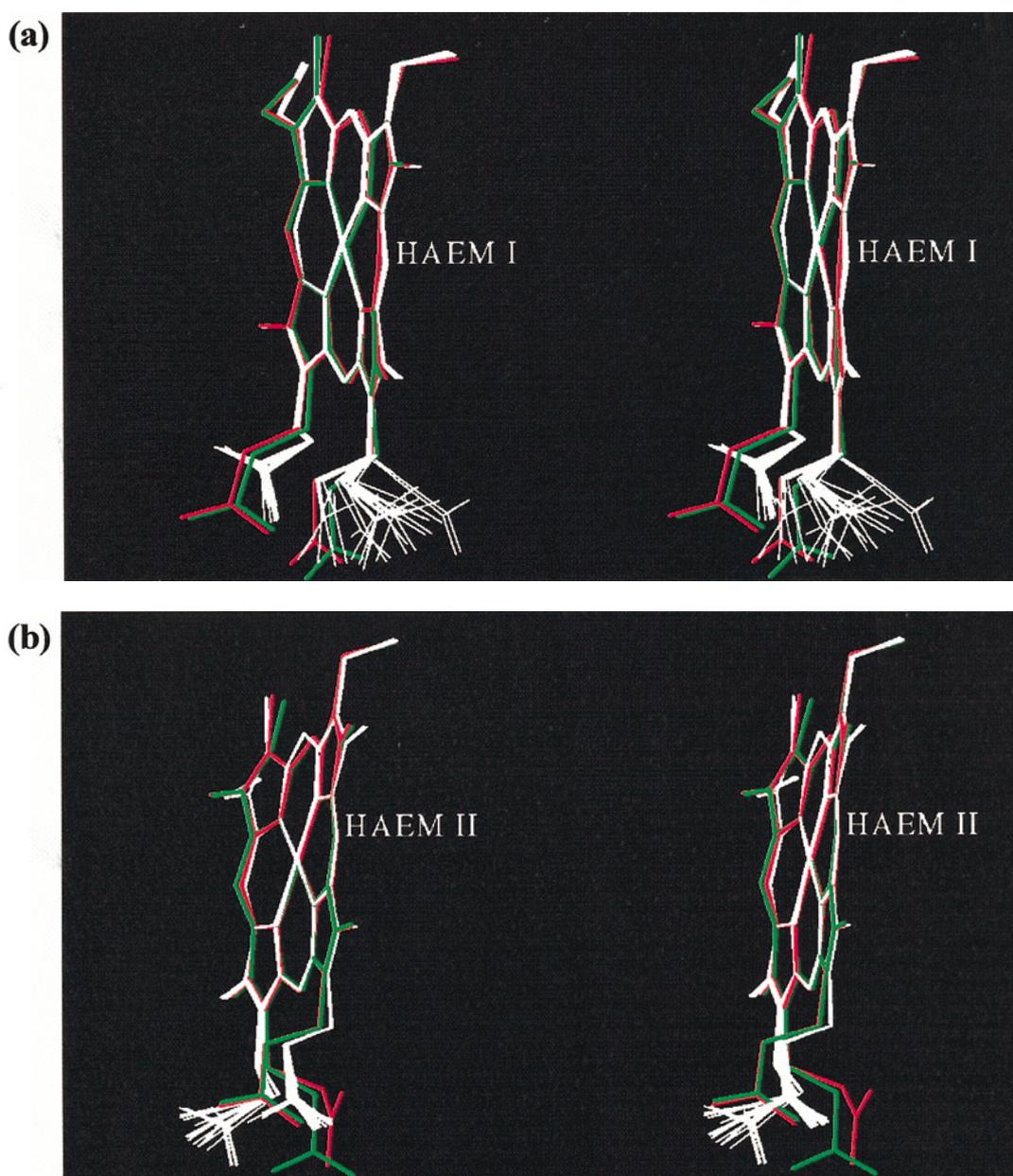


Figure 10(a-b) (legend opposite)

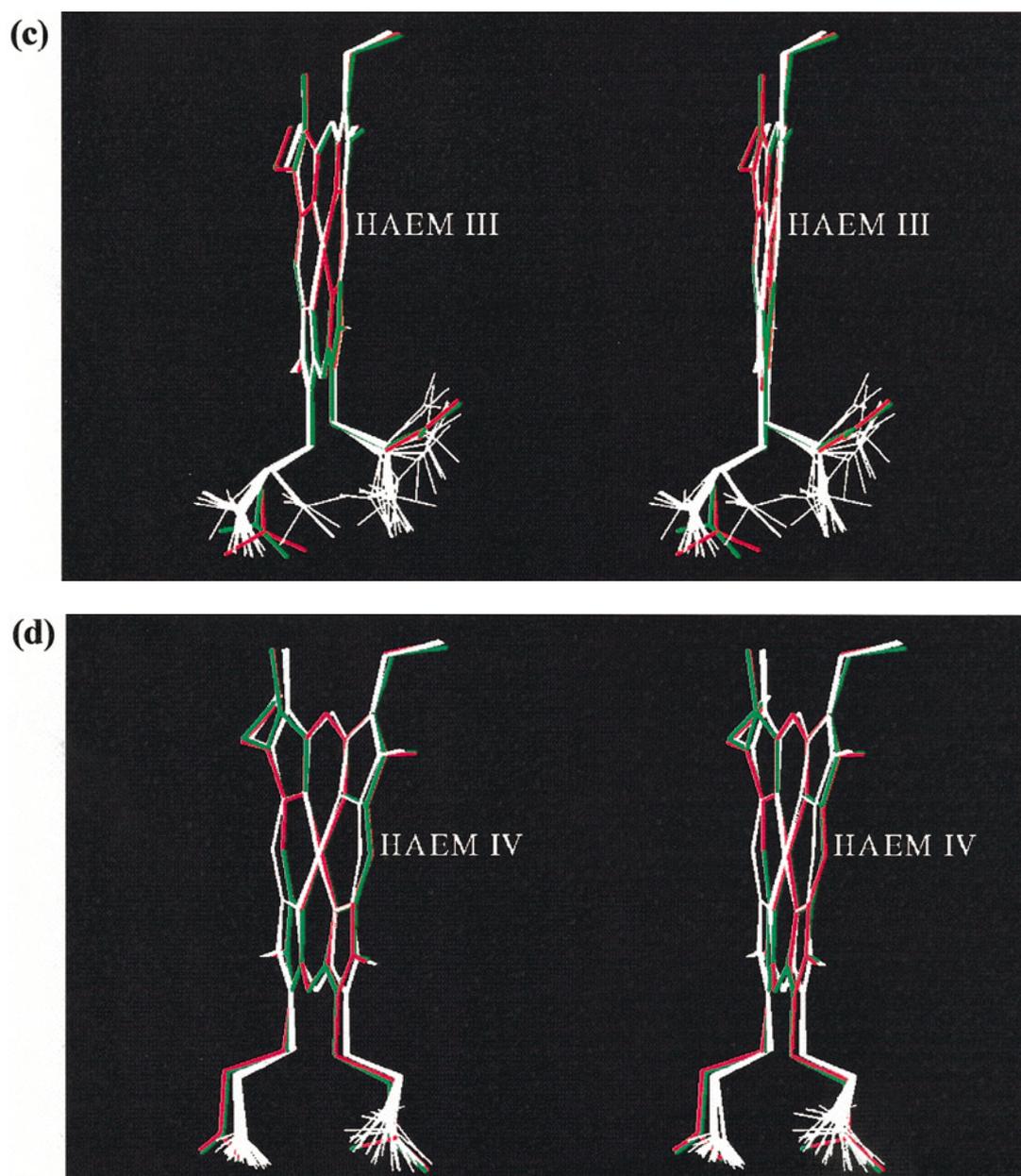


Figure 10. Haem puckering in NMR structures of the ferrocyanochrome at pH 8.5 (white) and the X-ray structures of the ferricytochrome at pH 5.5 (model A red; model B green). (a) Haem I; (b) haem II; (c) haem III; and (d) haem IV. Superimposition was performed using all the carbon atoms of the haem macrocycle but excluding those of the substituents.

Finally, it is significant that the haem puckers in the NMR structures agree quite well with those in the X-ray structures (Figure 10). This is the first example, to our knowledge, of a solution structure calculated with flat, or randomly distorted, haem groups as the starting point and, therefore, there was no bias.

Structural basis for the functional cooperativity of DvHc₃

The pK_a of the redox-Bohr effect of DvHc₃ changes from 7.4 in the fully reduced state to 5.3 in

the fully oxidised state and its magnitude indicates that the protonation centre is in the region of the propionate groups of haem I (Turner *et al.*, 1994, 1996). An analogous situation was observed for the homologous cytochrome from *D. vulgaris* (Miyazaki F: Park *et al.*, 1996; Salgueiro *et al.*, 1997). Recent theoretical studies (Soares *et al.*, 1997, 1998) also suggested that a redox-linked conformational change of haem I propionate 13 is the most likely candidate for the redox-Bohr effect, and that it could even be responsible for the positive cooperativity observed between haem I and haem II (Turner *et al.*, 1996).

As mentioned before when comparing the X-ray structure of the ferricytochrome with the NMR structure of the ferrocyclochrome, there is a localised movement of the backbone of residues 44 to 47 towards the propionate edge of haem I (Figure 9(b)). Spectra obtained at pH 7.6 and pH 7.3 show that Cys46 NH and the protons from both propionate groups of haem I experience chemical shift variations in the pH range close to the pK_a of the reduced state. Also, the decrease of the Cys46 NH chemical shift (0.17 ppm) is consistent with the breakage of its hydrogen bond to propionate 13 (Wüthrich, 1986). These observations clearly support the previous suggestions that the propionate substituents of haem I are directly involved in the cooperativity mechanism of DvHc₃. Indeed, segment 44-47 includes Lys45, which forms a salt-bridge with propionate 17, and Cys46, whose NH is hydrogen bonded to propionate 13. A concerted movement of these two residues could explain the coupled two-proton transfer step (Louro *et al.*, 1996) observed for the redox-Bohr cooperativity. Furthermore, since Cys46 is covalently bound to haem II, this movement could also be responsible for the positive cooperativity between haem I and haem II. The architecture of the region that includes haem I (with its two propionate groups), Lys45, Cys46 and haem II should then constitute the basic structural motif for the functional network of cooperativities governing the $2e^-/2H^+$ concerted step crucial for the energy transduction achieved by this protein (Louro *et al.*, 1997).

Concluding Remarks

Extensive assignment of the proton spectra of the fully reduced DvHc₃ at pH 8.5 (95% of the protons, i.e. all non-exchangeable protons and the exchangeable NH protons of the backbone) and careful quantitative analysis of the experimental NMR data enabled us to determine a high-quality structure of the protein in solution.

This work shows that the inclusion of lower distance restraints has significant advantages in the utilisation of information from overlapped or degenerate NOESY cross-peaks, which would otherwise be discarded. Furthermore, this facilitates the process of pseudo-stereospecific assignment used for fast-flipping aromatic groups.

A detailed comparison with the available X-ray structures of the ferricytochrome was performed, showing that it is possible to determine such detailed features as haem plane distortion with good accuracy. However, since the method of determining structures in solution relies on short-range interactions, distances on the scale of the protein diameters are susceptible to cumulative errors (cf. distance between haem II and haem IV in Table 4). This effect might not be significant for a structure of lower resolution, but its observation here should allow the methodology to be improved still further. Comparison of calculated

ring current shift errors showed that the localised structural differences found between the X-ray and the NMR structures very probably reflect genuine differences between the true crystal structure of the fully oxidised form of DvHc₃ at pH 5.5 and the true solution structure of its fully reduced form at pH 8.5. There is evidence that these differences are related to the redox-Bohr and redox cooperativity effects observed in DvHc₃: this study is an important step towards elucidating its mechanism.

Materials and Methods

Sample preparation

Cytochrome c_3 from *D. vulgaris* (Hildenborough) was purified as described (LeGall *et al.*, 1971). For NMR experiments in H₂O, the protein was freeze-dried and redissolved in 92% H₂O/8% ²H₂O. For NMR experiments in ²H₂O, the protein was freeze-dried once from ²H₂O (after incubation at 323 K for two hours) and redissolved in ²H₂O (99.96 atom %). Reduction of the protein was achieved by reaction with gaseous hydrogen in the presence of catalytic amounts of hydrogenase from *D. gigas* (LeGall *et al.*, 1982). The pH was adjusted in an anaerobic chamber (Mbraun MB 150 GI) by addition of 0.1 M NaO²H or ²HCl for ²H₂O samples, or by addition of 0.1 M NaOH or HCl for H₂O samples. The pH was monitored using a glass electrode (Ingold) inserted directly in the NMR tube. The pH values determined are direct meter readings without correction for the isotope effect. The final protein concentration was approximately 2 mM. An antibiotic cocktail (70 μM ampicillin, 50 μM kanamycin, 50 μM chloramphenicol) was added in order to prevent bacterial growth.

NMR spectroscopy

¹H-NMR spectra were obtained on a Bruker AMX-500 spectrometer equipped with a 5 mm inverse detection probe head with internal B₀ gradient coil. The sample temperature was controlled using a Eurotherm 818 temperature control unit with a B-CU 05 cooling unit.

One-dimensional spectra (32 k data points per free induction decay) were acquired using an approximate 45° flip angle, a one second recycle delay and water presaturation with a sweep-width of 16 ppm (8 kHz) or 60 ppm (30 kHz). Spectra acquired with larger bandwidth were used to confirm that the protein was fully reduced. Typically, 128 scans were added for each spectrum and processed with an exponential window function with 1 to 10 Hz line-broadening prior to Fourier transformation.

Assignments were performed using spectra of samples at pH 8.5 acquired at 303 K, but additional spectra were obtained at 283 K and 293 K to help resolve peak overlap and at pH 7.3 (303 K) to check sequential assignment in the solvent-exposed loops. All 2D spectra were acquired in the phase-sensitive mode by the States-TPPI method (Marion *et al.*, 1989) collecting 4096(t_2) × 1024(t_1) data points to cover a sweep-width of 8 kHz, with 32 scans per increment. NOESY spectra were acquired with mixing times of 40, 50, 70 and 100 ms (Jeener *et al.*, 1979; Brown *et al.*, 1988). NOESY spectra of the ²H₂O sample were recorded

with standard pulse sequences with continuous low-power water presaturation during the relaxation delay (at least 1.3 seconds) and the mixing time. NOESY spectra of the H_2O sample were recorded using a sequence with presaturation of the water resonance by a composite 180° inverse pulse followed by a SCUBA sequence to facilitate recovery of potentially saturated α -protons (Brown *et al.*, 1988). Total correlation spectra were acquired using the clean TOCSY pulse sequence (Briand & Ernst, 1991) with spin-lock times of 40 and 60 ms. COSY (Marion & Wüthrich, 1983) and DQF-COSY spectra were also recorded (Rance *et al.*, 1983; Derome & Williamson, 1990). Raw data were multiplied in the F_2 dimension by a Gaussian function with line-broadening of -5 Hz and gaussian broadening of 0.05, and by a pure cosine-squared function in the F_1 dimension, except for the COSY spectra that were multiplied by a pure sine-squared function in both dimensions. The 2D spectra were typically processed to a final size of $2 \text{ k} \times 1 \text{ k}$ data points, except COSY spectra used for assignments, which were processed with $1 \text{ k} \times 1 \text{ k}$ data points, and the NOESY spectra used for volume integration, which were processed with $2 \text{ k} \times 2 \text{ k}$ data points. The COSY spectrum used to obtain the $^3J_{\text{HNH}\alpha}$ coupling constants was processed with $8 \text{ k} \times 2 \text{ k}$ data points. Polynomial baseline corrections were applied in both dimensions of each spectrum. Data were processed using XWIN-NMR software (Bruker, Rheinstetten).

Proton chemical shifts are referenced to the resonances of the methyl groups in DSS designated at 0.0 ppm.

Assignment and integration

The software package XEASY (Version 1.2; ETH, Zürich; Bartels *et al.*, 1995) was used for assignment and volume integration of NOESY cross-peaks. All NOEs were measured at 303 K in the 100 ms NOESY spectra in H_2O and $^2\text{H}_2\text{O}$ at pH 8.5. The NOESY spectra at lower pH or at lower temperature were used to confirm sequential assignments. The NOESY spectrum at lower temperature was also used to determine the ratio of individual contributions to some NOEs that were completely overlapped at 303 K but were resolved at 283 K or at 293 K.

Integration was performed with the routines available in XEASY, either by manual integration for isolated peaks or with line-shape integration for overlapped peaks. In order to achieve good quantitative estimates of peak volumes, the baseline was determined around each individual peak (or cluster of peaks) and used to correct the peak volume (or the volume for a cluster of peaks).

Calculation of restraints

Volumes of NOESY cross-peaks that involve exchangeable protons in the spectrum of the H_2O sample were measured on one side of the spectrum only: several of these peaks were found to be weaker on the side of the diagonal for which the exchangeable proton frequency appears in F_1 because the evolving signal may be reduced by saturation transfer from the solvent. Before calibration, the volumes of cross-peaks assigned to exchangeable protons were also corrected for the percentage of $^2\text{H}_2\text{O}$ present. Signals from H^α protons close to the H_2O frequency may be reduced by saturation and so these NOEs were taken from the $^2\text{H}_2\text{O}$ spectrum only,

except of course for those involving NH protons. A few of the NH- H^α NOEs may be distorted on both sides of the diagonal but no consistent violation was found for any NH- H^α restraint.

Distances for non-exchangeable protons were calculated using the average of the volumes of the NOESY cross-peaks for each available assignment in the spectra obtained with H_2O and $^2\text{H}_2\text{O}$ as solvent. A single scaling factor for the ratio of the volumes in these spectra was obtained by dividing the sum of the volumes of all assigned and integrated peaks involving non-exchangeable protons in water by the sum of the corresponding volumes in the $^2\text{H}_2\text{O}$ spectrum. Furthermore, the volumes of the peaks assigned to protons separated by fixed distances, and all intrahaem cross-peaks except those involving the propionate groups, were excluded.

NOESY cross-peak intensities were converted into upls and lols using the program CALIBA (Güntert *et al.*, 1991) with the modification that pseudo-atom corrections (Wüthrich *et al.*, 1983) are set negative for lols. The atoms and the pseudo-atoms of the non-standard fragments used in the calculations (see the next section for details) were introduced into the program. Separate calibration curves were used for upls and lols in order to avoid excessive restriction, such that a given volume would be converted into an upl roughly 50% greater than the corresponding lol. Two separate calibration classes were defined both for upls and lols: one for NOEs involving non-methyl protons and another for NOEs that involved one or more methyl groups.

The experimental distance restraints calculated by CALIBA for non-stereospecifically assigned protons were modified using the earlier program DIANA Version 2.1 (Güntert, 1992). The experimental distance restraints were then used as input to generate protein conformers by using the new program for restrained molecular dynamics with simulated annealing, DYANA Version 1.4 (Güntert *et al.*, 1997).

Initially, only upls were used in structure calculation, with the distances calibrated approximately by using the volumes of NOE cross-peaks between protons separated by fixed distances. The preliminary structures were used to evaluate the scaling factors for the volume to distance conversion for each class by plotting the logarithm of the volumes against the logarithm of the distances in the structures. This procedure was repeated each time a new batch of structures was calculated. When the structures had been refined using upls alone, lols were introduced with scaling factors determined in a fashion similar to that used for upls. However, in order to avoid distorting the structures through excessively restrictive lols for protons in flexible regions of the protein, or solvent-exposed exchangeable protons, the complete set of lols was kept quite loose.

Where possible, ϕ torsion angles were calculated from $^3J_{\text{HNH}\alpha}$ coupling constants using the program HABAS, with the following parameterisation (Güntert *et al.*, 1989):

$$^3J_{\text{HNH}\alpha}(\phi) = 6.4 \cos^2(\phi - 60) - 1.4 \cos(\phi - 60) + 1.9 \quad (1)$$

Stereospecific assignments were obtained with the aid of the program GLOMSA (Güntert *et al.*, 1991).

Definition of the non-standard fragments used in structure calculations

The standard topology files were modified for fast-flipping aromatic residues and a haem residue was built as an extension of the second Cys residue in each haem-binding site, in such a way as to allow flexibility in the haem plane (Turner *et al.*, 1998).

The fast-flipping aromatic residues are identical with the standard library fragments, except that four extra pseudo-atoms are defined, HQD1, HQD2, HQE1 and HQE2, which have coordinates identical with HD1, HD2, HE1 and HE2, respectively. Initially, all but one of the cross-peaks to the ring protons were assigned arbitrarily to one side of the ring using the original atom labels, which, since they are not stereospecifically assigned, automatically generate loosened upl restraints, whereas the lols are applied directly to the protons in each equivalent pair. One strong NOE was chosen that was assigned to the respective pseudo-atom, for which the distance restraints are used without adjustment. Since the proton giving rise to the strong NOE to the ring protons must be close to just one side of the ring, this effectively defines the labelling with respect to a static ring orientation. The remaining NOEs were then assigned individually to the corresponding pseudo-protons in the course of structure refinement in a manner analogous to that used for diastereotopic pairs.

The haem groups were included in the structure calculations through the addition of an artificial amino acid residue to the residue library which consists of a Cys residue whose S atom is connected to the haem CAC atom (IUPAC C8¹). The haem topology was constructed using a zeta-matrix with the AMBER force-field parameters for a haem with two axial ligands (Cornell *et al.*, 1995). The zeta-matrix was built up starting from the iron atom and moving towards each of the pyrroles and its substituents so as to minimise build-up of cumulative errors. The porphyrin model structure thus generated for the fragment library is perfectly flat. A special covalent bond is defined to link the haem to the first Cys residue in each haem-binding site. The haem iron atom is represented by a pseudo-atom that is included in both His ligands and the hexacoordinate geometry of the iron was achieved by imposing upls of 2.80 Å between the haem pyrrole nitrogen atoms and the N^{ε2} atom of the axial His residues, 2.10 Å between the pyrrole nitrogens and the iron pseudo-atoms, and 0.1 Å between the two pseudo-atoms of the two His ligands. The haem is able to flex through torsion angles about the bonds that link the pyrroles to the meso carbon atoms, the pyrroles themselves being kept flat and rigid. The macrocycle is closed by a special covalent bond between CHC and C1C (IUPAC C5 and C6, respectively) with pseudo-atoms and 0.1 Å upl restraints. In total, 19 extra upls are defined in this design of the haem and it has sufficient flexibility to reproduce the full range of haem distortions found in X-ray structures.

Restrained torsion angle dynamics with simulated annealing

The default parameters for simulated annealing in the DYANA program required modification, since the structure calculation is very demanding because of the intricacy of cross-linking imposed by the haem groups in this

protein, even though a large number of restraints was used in the calculation (an average of 42 restraints per residue). In fact, the four bulky haem residues in the core of the protein and their definition, with eight extra torsion angles to allow for flexibility and 19 extra upls with added weight to keep the geometry of the haem within a reasonable range, make it difficult to achieve convergence in the structure calculations. The number of steps for the initial conjugate gradient minimisation of the random structure was increased to 500 to remove strong overlaps and the time step used for the MD simulation was half of the default value (Kirkpatrick *et al.*, 1983). The structures were annealed with a total of 80,000 steps, of which 3000 were used for the high-temperature MD. The weighting of van der Waals repulsions was then increased to 2.0 for a further 4000 steps of conjugate-gradient minimisation, followed by MD simulation at constant temperature and another 2000 steps of conjugate-gradient minimisation.

Since the pH of the sample (pH 8.5) was quite high when compared with the pK_a of carboxylic groups in model peptides (Bashford & Karplus, 1990) and haem propionate groups (Moore & Pettigrew, 1990), they were treated as deprotonated forms.

Structure analysis

The structure analysis, including Ramachandran plots, deviation from ideal structural parameters and solvent accessibility was performed with both PROCHECK Version 3.4.4 (Morris *et al.*, 1992; Laskowski *et al.*, 1993) and PROCHECK-NMR programs Version 3.4.4 (Laskowski *et al.*, 1996). Superimposition, visual inspection and drawing of the various families was accomplished with the program MOLMOL Version 2.4 (Koradi *et al.*, 1996). Identification of possible hydrogen bonds in the DYANA family of conformers and in available X-ray structures of the ferricytochrome (Simões *et al.*, 1998) was performed with the program WHAT IF Version 5.0 using the default parameters (Vriend, 1990). Since the hydrogen atoms are not visible in X-ray models, positioning of hydrogen atoms in these models was accomplished by optimising hydrogen bond networks (Hooft *et al.*, 1996). Identification and classification of the consensus secondary structure elements in the NMR structures ensemble was accomplished with the program PROMOTIF Version 2.0 (Hutchinson & Thornton, 1996).

Ring current shift calculations

The proton ring current shifts were calculated using the method of Johnson & Bovey (1958) with parameters for the eight-loop model of the haem taken from Cross & Wright (1985). Elliptic integrals were calculated using the subroutine CEL (Press *et al.*, 1989).

Ring current shifts were calculated for the best NMR structure and for the two molecules of the unit cell of the available X-ray structure of the ferricytochrome (Simões *et al.*, 1998). Ring current shifts were determined as the difference between the observed chemical shifts and random coil chemical shifts (Wishart *et al.*, 1995). This approximation holds fairly well (range ± 0.2 ppm) in a diamagnetic protein for protons not involved in hydrogen bonds, which may cause a large deviation from the random coil shift. For this reason, exchangeable protons were excluded from the calculations.

For each proton, the scaled error for the ring current shifts was calculated as:

$$\text{scaled error} = \sqrt{\frac{(\delta_{\text{calc}} - \delta_{\text{obs}})^2}{\text{grad}^2 + \Delta^2}} \quad (2)$$

where δ_{calc} is the calculated ring current shift, δ_{obs} is the ring current shift determined experimentally, grad is the local ring current shift gradient (in ppm/Å), and Δ is a term that allows for uncertainty in the reference shifts for each proton, set at 0.2 ppm. The average scaled error for each residue was calculated as the arithmetic average of the scaled errors of the protons of the residue.

Acknowledgements

We acknowledge Professor Lucia Banci of the University of Florence for her contribution in the initial stage of this work. We thank Ricardo Louro for his contribution and the acquisition of the preliminary spectra at the EC Large Scale Facility in Florence. The authors are grateful to Joaquim Mendes for useful suggestions and for help with computer programming. This work was supported by PRAXIS/PCNA/BIO/0074/96 and EC contracts BIO2-CT94-2052 and ERB CHRX-CT94-0540. A.C.M. is supported by a PRAXIS XXI program BD/2709/94 fellowship.

References

- Banci, L., Bertini, I., Bren, K. L., Gray, H. B., Sompornpisut, P. & Turano, P. (1997a). Solution structure of oxidized *Saccharomyces cerevisiae* iso-1-cytochrome *c*. *Biochemistry*, **36**, 8992–9001.
- Banci, L., Bertini, I., Gray, H. B., Luchinat, C., Reddig, T., Rosato, A. & Turano, P. (1997b). Solution structure of oxidized horse heart cytochrome *c*. *Biochemistry*, **36**, 9867–9877.
- Bartels, C., Xia, T., Billeter, M., Güntert, P. & Wüthrich, K. (1995). The program XEASY for computer-supported NMR spectral-analysis of biological macromolecules. *J. Biomol. NMR*, **6**, 1–10.
- Bashford, D. & Karplus, M. (1990). pK_a s of ionizable groups in proteins: atomic detail from a continuum electrostatic model. *Biochemistry*, **29**, 10219–10225.
- Berghuis, A. M. & Brayer, G. D. (1992). Oxidation state-dependent conformational changes in cytochrome *c*. *J. Mol. Biol.* **223**, 959–976.
- Briand, J. & Ernst, R. R. (1991). Computer-optimized homonuclear TOCSY experiments with suppression of cross relaxation. *Chem. Phys. Letters*, **185**, 276–285.
- Brown, S. C., Weber, P. L. & Mueller, L. (1988). Towards complete ^1H NMR spectra in proteins. *J. Magn. Reson.* **77**, 166–169.
- Cavanagh, J., Fairbrother, W. J., Palmer, A. G., III & Skelton, N. J. (1996). Sequential assignment and structure calculations. In *Protein NMR Spectroscopy: Principles and Practice*, p. 551, Academic Press, San Diego.
- Chen, Z., Koh, M., Driessche, G. V., Beeumen, J. J. V., Bartsch, R. G., Meyer, T. E., Cusanovich, M. A. & Mathews, F. S. (1994). The structure of flavocytochrome *c* sulfide dehydrogenase from a purple phototrophic bacterium. *Science*, **266**, 430–432.
- Coletta, M., Catarino, T., LeGall, J. & Xavier, A. V. (1991). A thermodynamic model for the cooperative functional properties of the tetraheme cytochrome c_3 from *Desulfovibrio gigas*. *Eur. J. Biochem.* **202**, 1101–1106.
- Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Jr, Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W. & Kollman, P. A. (1995). A second generation force field for the simulation of proteins, nucleic acids and organic molecules. *J. Am. Chem. Soc.* **117**, 5179–5197.
- Cross, K. J. & Wright, P. E. (1985). Calibration of ring current models for the haem ring. *J. Magn. Reson.* **64**, 220–231.
- Czjzek, M., Payan, F., Guerlesquin, F., Bruschi, M. & Haser, R. (1994). Crystal structure of cytochrome c_3 from *Desulfovibrio desulfuricans* Norway at 1.7 Å resolution. *J. Mol. Biol.* **243**, 653–667.
- Derome, A. E. & Williamson, M. P. (1990). Rapid-pulsing artifacts in double-quantum-filtered COSY. *J. Magn. Reson.* **88**, 177–185.
- Fukuyama, H., Kunishima, N., Amada, F., Kubota, T. & Matsubara, H. (1995). Crystal structures of cyanide- and triiodide-bound forms of *Arthromyces ramosus* peroxidase at different pH values. Perturbations of active site residues and their implication in enzyme catalysis. *J. Biol. Chem.* **37**, 21884–21892.
- Fülöp, V., Ridout, C. J., Greenwood, C. & Hadju, J. (1995). Crystal structure of the di-haem cytochrome *c* peroxidase from *Pseudomonas aeruginosa*. *Structure*, **3**, 1225–1233.
- Güntert, P. (1992). DIANA Program Version 2.1, Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule-Hoenggerberg, CH-8093, Zürich, Switzerland.
- Güntert, P., Braun, W., Billeter, M. & Wüthrich, K. (1989). Automated stereospecific ^1H NMR assignments and their impact on the precision of protein structure determination in solution. *J. Am. Chem. Soc.* **111**, 3997–4004.
- Güntert, P., Braun, W. & Wüthrich, K. (1991). Efficient computation of three-dimensional protein structures in solution from nuclear magnetic resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J. Mol. Biol.* **217**, 517–530.
- Güntert, P., Mumenthaler, C. & Wüthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* **273**, 283–298.
- Gunasekaran, K., Ramakrishnan, C. & Balaram, P. (1996). Disallowed Ramachandran conformations of amino acid residues in protein structures. *J. Mol. Biol.* **264**, 191–198.
- Higuchi, Y., Kusunoki, M., Matsuura, Y., Yasuoka, N. & Kakudo, M. (1984). Refined structure of cytochrome c_3 at 1.8 Å resolution. *J. Mol. Biol.* **172**, 109–139.
- Hooft, R. W. W., Sander, C. & Vriend, G. (1996). Positioning hydrogen atoms by optimizing hydrogen-bond networks in protein structures. *Proteins: Struct. Funct. Genet.* **26**, 363–376.
- Hutchinson, E. G. & Thornton, J. M. (1996). PROMOTIF: a program to identify and analyse structural motifs in proteins. *Protein Sci.* **5**, 212–220.
- Jeener, J., Meier, B. H., Bachmann, P. & Ernst, R. R. (1979). Investigation of exchange processes by two-dimensional n.m.r. spectroscopy. *J. Chem. Phys.* **71**, 4546–4553.

- Johnson, C. E. & Bovey, F. A. (1958). Calculation of nuclear magnetic resonance spectra of aromatic hydrocarbons. *J. Chem. Phys.* **29**, 1012–1014.
- Kirkpatrick, S., Gelatt, C. D., Jr & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, **220**, 671–680.
- Koradi, R., Billeter, M. & Wüthrich, K. (1996). MOL-MOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51–55.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
- Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R. & Thornton, J. M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR*, **8**, 477–486.
- LeGall, J., Bruschi, H. M. & DerVartanian, D. V. (1971). Electron paramagnetic resonance and light absorption studies on *c*-type cytochromes of the anaerobic sulfate reducer *Desulfovibrio*. *Biochim. Biophys. Acta*, **234**, 499–512.
- LeGall, J., Ljungdahl, P., Moura, I., Peck, H. D., Jr, Xavier, A. V., Moura, J. J. G., Teixeira, M., Huynh, B. H. & DerVartanian, D. V. (1982). The presence of redox-sensitive nickel in the periplasmic hydrogenase from *Desulfovibrio gigas*. *Biochem. Biophys. Res. Commun.* **106**, 610–616.
- Louro, R. O., Catarino, T., Salgueiro, C. A., LeGall, J. & Xavier, A. V. (1996). Redox-Bohr effect in the tetrahaem cytochrome c_3 from *Desulfovibrio vulgaris*: a model for energy transduction mechanisms. *J. Biol. Inorg. Chem.* **1**, 34–38.
- Louro, R. O., Catarino, T., Salgueiro, C. A., LeGall, J. & Xavier, A. V. (1997). Redox-Bohr effect in electron/proton energy transduction: cytochrome c_3 coupled to hydrogenase works as a 'proton thruster' in *Desulfovibrio vulgaris*. *J. Biol. Inorg. Chem.* **2**, 488–491.
- MacArthur, M. W. & Thornton, J. M. (1996). Deviations from planarity of the peptide bond in peptides and proteins. *J. Mol. Biol.* **264**, 1180–1195.
- Marion, D. & Wüthrich, K. (1983). Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurement of ^1H - ^1H spin-spin coupling constants in proteins. *Biochem. Biophys. Res. Commun.* **113**, 967–974.
- Marion, D., Ikura, M., Tschudin, R. & Bax, A. (1989). Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins. *J. Magn. Reson.* **85**, 393–399.
- Matias, P. M., Frazão, C., Morais, J., Coll, M. & Carrondo, M. A. (1993). Structure analysis of cytochrome c_3 from *Desulfovibrio vulgaris* Hildenborough at 1.9 Å resolution. *J. Mol. Biol.* **234**, 680–699.
- Matias, P. M., Morais, J., Coelho, R., Carrondo, M. A., Wilson, K., Dauter, Z. & Sieker, L. (1996). Cytochrome c_3 from *Desulfovibrio gigas*: crystal structure at 1.8 Å resolution and evidence for a specific calcium-binding site. *Protein Sci.* **5**, 1342–1354.
- Moore, G. R. & Pettigrew, G. W. (1990). Stereochemical and physicochemical properties of hemes: the structures of class I cytochromes *c*. In *Cytochromes c. Evolutionary, Structural and Physicochemical Aspects*, pp. 11–12; 185–189, Springer-Verlag, Berlin.
- Moore, G. R., Leitch, F. A., Pettigrew, G. W., Rogers, N. K. & Williams, G. (1986). The importance of electrostatic interactions involving buried groups in determining the structure and properties of metalloproteins. In *Frontiers in Bioinorganic Chemistry* (Xavier, A. V., ed.), pp. 494–506, VCH Verlagsgesellschaft, Weinheim.
- Morris, A. L., MacArthur, M. W., Hutchinson, E. G. & Thornton, J. M. (1992). Stereochemical quality of protein structure coordinates. *Proteins: Struct. Funct. Genet.* **12**, 345–364.
- Moura, J. J. G., Santos, H., Moura, I., LeGall, J., Moore, G. R., Williams, R. J. P. & Xavier, A. V. (1982). NMR redox studies of *Desulfovibrio vulgaris* cytochrome c_3 . Electron transfer mechanisms. *Eur. J. Biochem.* **127**, 151–155.
- Park, J. S., Ohmura, T., Kano, K., Sagara, T., Niki, K., Kyogoku, Y. & Akutsu, H. (1996). Regulation of the redox order of four hemes by pH in cytochrome c_3 from *D. vulgaris* Miyazaki F. *Biochim. Biophys. Acta*, **1293**, 45–54.
- Perutz, M. (1990). In *Mechanisms of Cooperativity and Allosteric Regulation in Proteins*, Cambridge University Press, Cambridge, UK.
- Pfeiffer, S., Karimi-Nejad, Y. & Rüterjans, H. (1997). Limits of NMR structure determination using variable target function calculations: ribonuclease T_1 , a case study. *J. Mol. Biol.* **266**, 400–423.
- Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. (1989). Special functions. *Numerical Recipes*, pp. 183–188, Cambridge University Press, Cambridge, U.K.
- Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**, 95–99.
- Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R. & Wüthrich, K. (1983). Improved spectral resolution in COSY ^1H NMR spectra of proteins via double quantum filtering. *Biochem. Biophys. Res. Commun.* **117**, 479–485.
- Salgueiro, C. A., Turner, D. L., LeGall, J. & Xavier, A. V. (1997). Reevaluation of the redox and redox-Bohr cooperativity in tetrahaem *Desulfovibrio vulgaris* (Miyazaki F) cytochrome c_3 . *J. Biol. Inorg. Chem.* **2**, 343–349.
- Santos, H., Moura, J. J. G., Moura, I., LeGall, J. & Xavier, A. V. (1984). NMR studies of electron transfer mechanisms in a protein with interacting redox centres: *Desulfovibrio gigas* cytochrome c_3 . *Eur. J. Biochem.* **141**, 283–296.
- Sibanda, B. L., Blundell, T. L. & Thornton, J. M. (1989). Conformation of β -hairpins in protein structures. A systematic classification with applications to modelling by homology, electron density fitting and protein engineering. *J. Mol. Biol.* **206**, 759–777.
- Simões, P., Matias, P. M., Morais, J., Wilson, K., Dauter, Z. & Carrondo, M. A. (1998). Refinement of the three-dimensional structures of cytochromes c_3 from *Desulfovibrio vulgaris* Hildenborough at 1.67 Å resolution and from *Desulfovibrio desulfuricans* ATCC 27774 at 1.6 Å resolution. *Inorg. Chim. Acta.* **273**, 213–224.
- Soares, C. M., Martel, P. J. & Carrondo, M. A. (1997). Theoretical studies on the redox-Bohr effect in cytochrome c_3 from *Desulfovibrio vulgaris* Hildenborough. *J. Biol. Inorg. Chem.* **2**, 714–727.
- Soares, C. M., Martel, P. J., Mendes, J. & Carrondo, M. A. (1998). Molecular dynamics simulation of cytochrome c_3 . Studying the reduction processes using free energy calculations. *Biophys. J.* **74**, 1708–1721.

- Turner, D. L., Salgueiro, C. A., LeGall, J. & Xavier, A. V. (1992). Structural studies of *Desulfovibrio vulgaris* ferrocyclochrome c_3 by two-dimensional NMR. *Eur. J. Biochem.* **210**, 931–936.
- Turner, D. L., Salgueiro, C. A., Catarino, T., LeGall, J. & Xavier, A. V. (1994). Homotropic and heterotropic cooperativity in the tetrahaem cytochrome c_3 from *Desulfovibrio vulgaris*. *Biochim. Biophys. Acta*, **1187**, 232–235.
- Turner, D. L., Salgueiro, C. A., Catarino, T., LeGall, J. & Xavier, A. V. (1996). NMR studies of cooperativity in the tetrahaem cytochrome c_3 from *Desulfovibrio vulgaris*. *Eur. J. Biochem.* **241**, 723–731.
- Turner, D. L., Brennan, L., Chamberlin, S. G., Louro, R. O. & Xavier, A. V. (1998). Determination of solution structures of paramagnetic proteins by NMR. *Eur. Biophys. J.* **27**, 367–375.
- Vriend, G. (1990). WHAT IF: a molecular modelling and drug design program. *J. Mol. Graph.* **8**, 52–56.
- Wareham, R. S., Kilburn, J. D., Rees, N. H., Turner, D. L., Leach, A. R. & Holmes, D. S. (1995a). Synthesis and solution conformation of a c_2 symmetric macrobicyclic. *Tetrahedron Letters*, **36**, 3047–3050.
- Wareham, R. S., Kilburn, J. D., Turner, D. L., Rees, N. H. & Holmes, D. S. (1995b). Homeomorphic isomerism in a peptidic macrobicyclic. *Angew. Chem. Int. Ed. Engl.* **34**, 2660–2662.
- Williams, P. A., Fülöp, V., Garman, E. F., Saunders, N. F. W., Ferguson, S. J. & Hadju, J. (1997). Haem-ligand switching during catalysis in crystals of a nitrogen-cycle enzyme. *Nature*, **389**, 406–412.
- Wishart, D. S., Bigam, C., Holm, A., Holdges, R. & Sykes, B. (1995). ^1H , ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbour effects. *J. Biomol. NMR*, **5**, 67–81.
- Wüthrich, K. (1986). NMR spectra of proteins and nucleic acids in solution: sequence-specific resonance assignments in proteins. In *NMR of Proteins and Nucleic Acids*, pp. 130–161; 30–31, John Wiley & Sons, New York.
- Wüthrich, K., Billeter, M. & Braun, W. (1983). Pseudo-structures for the 20 amino acids for use in studies of protein conformations by measurements of intramolecular proton-proton distance constraints with nuclear magnetic resonance. *J. Mol. Biol.* **169**, 949–961.
- Zhao, D. & Jardetzky, O. (1994). An assessment of the precision and accuracy of protein structures determined by NMR: dependence on distance errors. *J. Mol. Biol.* **239**, 601–607.

Edited by P. E. Wright

(Received 27 January 1998; received in revised form 21 May 1998; accepted 27 May 1998)



<http://www.hbuk.co.uk/jmb>

Supplementary material comprising one Table and three Figures is available from JMB Online