The Solution Structure of Oxidized *Escherichia coli* Cytochrome b_{562}[†,‡]

Fabio Arnesano, Lucia Banci, Ivano Bertini,* Jasmin Faraone-Mennella, and Antonio Rosato

Department of Chemistry and Centro di Risonanze Magnetiche, University of Florence, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Italy

Paul D. Barker and Alan R. Fersht

Centre for Protein Engineering, MRC Centre, Hills Road, Cambridge CB2 2QH, United Kingdom

Received November 24, 1998; Revised Manuscript Received March 24, 1999

ABSTRACT: The solution structure of the oxidized, paramagnetic form of cytochrome b_{562} from *Escherichia coli* (106 amino acids) is here reported as obtained from 1653 meaningful NOEs (from a total of 2051 unique NOEs), 33 $^3J_{HNH\alpha}$ values, and 339 pseudocontact shifts. The structure displays the typical four-helix bundle motif, and a disordered loop between helices α2 and α3, as found in the solid state. The solution structure has a conformation intermediate between the two independent solid-state molecules, although different orientations are observed for a few residues. The magnetic susceptibility tensor is similar to that of cytochrome *c*, which has the same ligands, although the anisotropy is somewhat smaller. This difference in the electronic structure is consistent with the thermal accessibility in cytochrome b_{562} of states with $S > 1/2$. The structure is also compared with the solution structure of the apoprotein, and some information on the role of the cofactor on the protein folding and mobility is obtained. Helix α4 seems to be the most sensitive to the chemical environment in terms of structure and mobility. The pK_a values affecting the hyperfine-shifted signals are also discussed. Quite intriguing is the comparison of the structure of cytochrome b_{562} with the available structures of cytochromes *c*′ which display a similar folding motif and similar pK_a values but very little sequence similarity.

An understanding of the role of cofactor binding to polypeptide in the folding of a protein is fundamental to our ability to design and build useful metalloproteins with defined electronic and catalytic properties. The four-helix bundle protein fold has been the subject of protein folding studies (*1*, *2*) and also the target for metalloprotein design (*3*, *4*). Nature has used this motif for proteins of diverse functions, including heme binding proteins (*5*, *6*) and dinuclear iron metalloproteins (*7*).

Cytochrome b_{562} (cyt b_{562} hereafter)[1] is a monomeric, four-helix bundle protein found in the periplasm of *Escherichia coli* (*8*). The protein of 106 residues binds noncovalently a single heme as prosthetic group. The iron ion is essentially low-spin, 6-coordinate, with methionine and histidine providing the axial ligands. The ease of overexpression of this protein in *E. coli* and the availability of structures for both the apo- and oxidized holoproteins have made this an excellent model system for the study of the role of cofactor binding in protein folding. As such, it is currently the subject of investigation in several laboratories, despite the fact that its function in vivo remains unknown. The properties of this protein are highly pH dependent, and the protein is tolerant of a variety of mutations without compromising the heme binding ability. Recent studies have investigated the linkage between the stability of the cytochrome b_{562} polypeptide and its affinity for heme (*9*−*11*). These have highlighted the large increase in protein stability afforded by binding of heme, especially in the reduced state (*10*), which has been utilized for electron-transfer-induced folding studies (*12*). Our interest in this protein is directed toward the understanding of the large differences in properties between it and the cytochrome *c* primes (cyt *c*′). The latter contain a *c*-type heme as the prosthetic group, the iron ion being high-spin, 5-coordinate. These cytochromes are present in photosynthetic species and are found as dimers in many species, although their 4-helical folds are highly homologous to that of cytochrome b_{562}.

The structure of the oxidized cyt b_{562} holoprotein has been solved by X-ray crystallography to high resolution (*13*). In that structure the protein is a dimer in the asymmetric unit with a significant region of intermolecular contact involving the C-terminal helix. This crystal packing may influence the structure of the protein in this important region of the polypeptide that provides one of the ligands (His 102) to the heme iron. The solution structure of the apoprotein has been determined using constraints from multidimensional NMR experiments (*14*). However, due to the nature of the protein in this state, only a moderate number of constraints were available for structure determination. Despite this, the structure of the apoprotein provides an interesting view of

† This work was supported by the European Community (TMR-LSF Contract ERBFMGECT950033, TMR Contract FMRX-CT98-0230), by Italian CNR (Biotecnologie e biologia molecolare, CTB CNR 95.02860.CT14), and by MURST ex 40%. P.D.B. thanks the BBSRC for an Advanced Fellowship.

‡ The average minimized structure is available at the Protein Data Bank: PDB entry 1QPU and RCSB entry RCSB009123.

* Corresponding author: Department of Chemistry and Centro di Risonanze Magnetiche, University of Florence, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Italy. Tel.: +39 055 4209272. Fax: +39 055 4209271. E-mail: bertini@CERM.UNIFI.IT.

[1] Abbreviations: cyt b_{562}, cytochrome b_{562}; RMSD, root-mean-square deviation; pcs, pseudocontact shifts.

10.1021/bi982785f CCC: $18.00
Published on Web 06/12/1999

the protein in a partially folded or molten globule state that may be relevant to the folding and assembly of the holoprotein.

Here we present the solution structure of the paramagnetic, oxidized holoprotein, under similar conditions to those used for the apoprotein, obtained by NMR spectroscopy. Indeed, a description of the solution structure is also required for detailed analyses of the folding of this protein. This paramagnetic system provides another test of the methodology required to assign NMR signals and also affords additional constraints for structural definition. The resulting structure is compared with both the crystal structure of the oxidized protein and the NMR structure of the apoprotein. This is the necessary starting point for detailed structural analysis of mutants designed to mimic the functional properties of cytochrome *c*′.

MATERIALS AND METHODS

Sample Isolation and Preparation. Unlabeled protein samples were prepared following the procedure described in (*15*). To prepare the ^{15}N-labeled samples, *E. coli* strain TG2 cells were transformed with the plasmid pPB10 (harboring the gene of *E. coli* cyt b_{562}; Barker, unpublished) and plated onto TY agarose containing 100 μg/mL ampicillin and 0.1 mg/mL glucose. Then 10 mL of 2 × TY media containing 100 μg/mL ampicillin was inoculated from a single colony, and cultured overnight at 37 °C with shaking. The inoculum was then transferred to a 2 L flask containing 1 L of minimal media with 50 μg/mL ampicillin. The culture was grown with shaking at 37 °C up to an OD_{600} of about 0.8, and then induced with 0.5 mM isopropyl β-D-thiogalactoside. The culture was harvested by centrifugation after 15 h from induction. The protein was extracted from the *E. coli* cells as previously described (*15*). Significant amounts of apoprotein were found to be present in the medium and were collected by concentrating the culture supernatant through ultrafiltration. The protein samples obtained from the cell lysis and from the culture media were combined, and the pH of the solution was adjusted to 7. The apoprotein was then saturated with exogenous heme (Sigma), and incubated at 4 °C for 24 h. From this solution, holo cyt b_{562} was purified as previously described (*15*). The minimal medium used for cultures consisted of M9 salts supplemented with $MgSO_4$, trace metal, and vitamin solutions. The nitrogen source was $^{15}NH_4Cl$ (1 g/L), and the carbon source was glucose (4 g/L).

The samples used for NMR spectroscopy were ca. 3 mM in protein, in 500 mM phosphate buffer.

NMR Spectroscopy. All NMR spectra except the TOCSY spectra were acquired on a Bruker Avance 800 spectrometer, operating at a proton nominal frequency of 800.13 MHz. The TOCSY experiments were recorded on a Bruker Avance 600 operating at 600.13 MHz. A triple resonance 5-mm probe has been used on both spectrometers. All the experiments described below were performed on a sample at pH 4.8. In addition, 2D NOESY experiments were replicated on samples at pH 7.0 and 9.2.

3D TOCSY-^{15}N HMQC (*16*) and NOESY-^{15}N HMQC (*17*) experiments were recorded in H_2O solution with 1024(^{1}H) × 100(^{15}N) × 512(^{1}H) data points. In these experiments, the delay between the ^{1}H 90° pulse following the mixing period and the first subsequent ^{15}N 90° pulse was set to 5.5 ms ($\approx 1/2 J_{NH}$), the mixing time was 80 ms for TOCSY-^{15}N HMQC and 100 ms for NOESY-^{15}N HMQC, and the acquisition was carried out by positioning the carrier frequency in the center of the amide proton region at 7.35 ppm. Spectral windows of 26, 6.5, and 12.0 ppm were used, respectively, for ^{15}N, direct ^{1}H dimension, and indirect ^{1}H dimension.

To detect connectivities among hyperfine-shifted signals, two 2D NOESY experiments (*18*, *19*) with spectral widths of 90 and 15 ppm in both frequency dimensions, with a recycle time of 450 ms, and with 40 ms mixing time were acquired. In addition, 2D NOESY (*18*, *19*) (either on the H_2O or on the D_2O samples) and TOCSY (*20*, *21*) spectra were acquired with a mixing time of 100 ms and a spin-lock time of 70 ms, respectively, and a recycle time of 900 ms. A DQF-COSY (*22*) experiment was acquired with a spectral width of 15 ppm in both dimensions. All data consisted of 4K data points in the acquisition dimension and of 1K experiments in the indirect dimension. Connectivities involving the hyperfine-shifted broad signals of the iron axial ligands were detected through 1D NOE's, using the reported methodology (*23*), by irradiating the signals at −8.1 and −21.1 ppm. Detection of broad signals was enhanced by acquiring spectra with the SuperWEFT pulse sequence (*24*, *25*), using very short recycle times.

The HNHA experiment (*26*) was performed to determine $^3J_{HN\alpha}$ coupling constants. These were used to obtain constraints for ϕ torsion angles. The spectrum was recorded as a 200(^{1}H) × 100(^{15}N) × 1024(^{1}H) data set using Pulsed Field Gradients (PFG) along the *z* axis. The mixing time was 100 ms. Spectral windows of 43, 13, and 13 ppm were used, respectively, for ^{15}N, direct ^{1}H dimension, and indirect ^{1}H dimension.

For all the experiments, quadrature detection in the indirect dimensions was performed in the TPPI mode (*19*), and water suppression was achieved through the WATERGATE sequence (*27*), except in the NOESY experiments performed to detect connectivties between paramagnetically shifted resonances, where presaturation was used.

All 3D and 2D spectra were collected at 298 K, processed using the standard Bruker software (XWINNMR), and analyzed on IBM RISC 6000 computers through the XEASY program (*28*).

Constraints Used in Structure Calculations. The solution structure was determined for the major form of oxidized cyt b_{562}. The majority of the peaks used for the structure calculations (≈90%, see later) were integrated in the NOESY map acquired over the diamagnetic window at 298 K in H_2O by using the WATERGATE sequence for solvent suppression. Dipolar connectivities involving hyperfine-shifted protons were integrated in the NOESY spectrum acquired over the 90 ppm spectral window. Intensities of dipolar connectivities were converted into upper distance limits, to be used as input for structure calculations, by using the approach provided by the program CALIBA (*29*). Connectivities measured in NOESY maps acquired with different mixing times were calibrated independently. A scaling factor for the calibration parameters was estimated and found to be comparable with that obtained from the integration of well-defined and isolated peaks occurring in both spectra acquired at different mixing times. The calibration curves

were adjusted iteratively as the structure calculations proceeded. The proton−proton distance constraints derived from 1D NOE's were introduced as upper distance limits of 5.0 Å. Stereospecific assignments of diastereotopic protons were obtained using the program GLOMSA (*29*).

$^3J_{HNH\alpha}$ coupling constants were converted into constraints for the backbone torsion angle ϕ, by means of the appropriate Karplus curve (*26*). Ambiguity between different angles corresponding to the same value of the $^3J_{HNH\alpha}$ coupling constant were resolved by looking at structures obtained from PSEUDYANA (*30*) calculations run without the ambiguous torsion angle constraints. All angles were given a ±20° uncertainty.

Hydrogen bond constraints were introduced for backbone amide protons which were observed to be nonexchanging in D_2O solution and that were found to be within hydrogen bond distance and to have the correct orientation with respect to hydrogen bond acceptors in structural models obtained without inclusion of these constraints. The distance between the amide proton and the hydrogen bond acceptor was constrained to be in the 1.8−2.0 Å interval by inclusion of the corresponding upper and lower distance limits in structure calculations. In addition, upper and lower distance limits of 3.0 and 2.7 Å between the N and the acceptor atoms were included in calculations. No constraints were introduced for nonexchanging amide protons for which it was not possible to unambiguously identify the corresponding acceptor.

Pseudocontact shifts (pcs hereafter) were employed as additional constraints for the structure calculations (*31*−*33*). Pcs values were obtained by subtracting from the chemical shifts measured in the oxidized form of cyt b_{562} the diamagnetic chemical shifts calculated from a preliminary structure obtained with NOE and ^3J constraints with the program SHIFT (*34*) using the random coil chemical shifts reported by Merutka et al. (*35*). No pseudocontact shift constraints were used for exchangeable protons, as it is reported that SHIFT cannot calculate accurately their diamagnetic chemical shift (*34*). A tolerance of 10% of the estimated pcs was used in the PSEUDYANA calculations (*30*) (see later) with a minimum value of 0.5 ppm for all protons. The hyperfine shifts of the heme and of the axial heme ligands (Met 7 and His 102) were not included in the calculations since they experience a nonnegligible contact shift. A total of 339 pcs constraints were used.

Structure Calculations. The structure calculations were performed using PSEUDYANA (*30*), which is a modified version of the program DYANA (*36*) adapted to include pcs as additional restraints. A preliminary family of 20 conformers obtained using only the NOE constraints was used as an input model to the program FANTASIAN (*37, 38*) to provide the initial values of $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$. These calculations were performed for each of the 20 preliminary structures, and the average values of $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ were used as starting values in the PSEUDYANA structure calculations. The PSEUDYANA protocol needs only the initial values of the magnetic susceptibility anisotropy tensor and not its direction cosines; i.e., no specific initial orientation is provided for the tensor. A special residue, called LTNS, formed by dummy atoms is used to define the position of the metal center and the directions of the magnetic susceptibility tensor in the dihedral angle coordinate space. The dummy atoms of this residue have their van der Waals radii set to zero so that it can freely penetrate into the protein. An upper distance limit of 0.2 Å was set between the pseudoatom defining the metal ion and the iron atom. The location of this residue and the tensor orientation were optimized during the structure calculations. After each cycle of structure calculations, the magnetic anisotropy parameters were reevaluated and used as input for the following calculation until the final values did not deviate more than 5% from the initial ones. Three cycles were needed to reach convergence.

The heme group was included in the calculations by defining a new residue in the amino acid library according to the previously described procedure (*39*). The two axial ligands (Met 7 and His 102) are linked to the iron atom through upper distance limits of 2.35 and 2.10 Å with the sulfur atom of Met 7 and with the Nϵ2 of His 102, respectively. This approach does not impose any fixed orientation of the axial ligands with respect to the heme. In all structure calculations, the relative weight of all classes (NOE, pcs, and *J* couplings) of constraints was kept equal to the default PSEUDYANA values. A total of 200 random conformers were annealed in 10 000 steps using the above constraints. The 35 conformers with the lowest target function constitute the final family.

Restrained energy minimization was then applied to each member of the family using the AMBER 5.0 Package (*40*). The distance constraints were applied within the molecular mechanics and dynamics module of SANDER, and the pcs were included as constraints by means of the module PCSHIFTS (*38*). The force field parameters for the heme and its ligands, as well as the overall calculation procedure, were set up as previously reported for similar systems (*39, 41*). The initial values of the magnetic susceptibility anisotropy used as input in PSEUDOREM were the values on the final PSEUDYANA family. On the final PSEUDOREM family, the magnetic susceptibility anisotropy parameters deviate less than 5% from their initial values (i.e., before energy minimization).

The program CORMA (*42*), which is based on relaxation matrix calculations, was used to back-calculate the NOESY cross-peaks and to check the validity of the structure. It has been also used to locate a few more cross-peaks between already assigned resonances.

The quality of the structure was evaluated in terms of deviations from ideal bond lengths and bond angles and through Ramachandran plots obtained using the programs PROCHECK (*43*) and PROCHECK-NMR (*44*).

Structure calculations and analyses were performed on an IBM SP02 parallel computer.

RESULTS

Sequence-Specific Assignment. This study was performed on the oxidized form of cyt b_{562} at a protein concentration of 2−3 mM in 500 mM phosphate buffer, at pH 4.8. In these conditions, the protein does not experience aggregation, which occurs at concentrations larger than 4 mM (*45, 46*). Thus, the protein provides sharp hyperfine-shifted NMR signals (Figure 1). The protein exists in solution as two species, in which the orientation of the heme groups differs by a 180° rotation around the $\alpha-\gamma$ axis, in the ratio of ca. 7:3 (*46*) and referred to as forms A and B, respectively. Besides the heme substituents and the iron axial ligands,

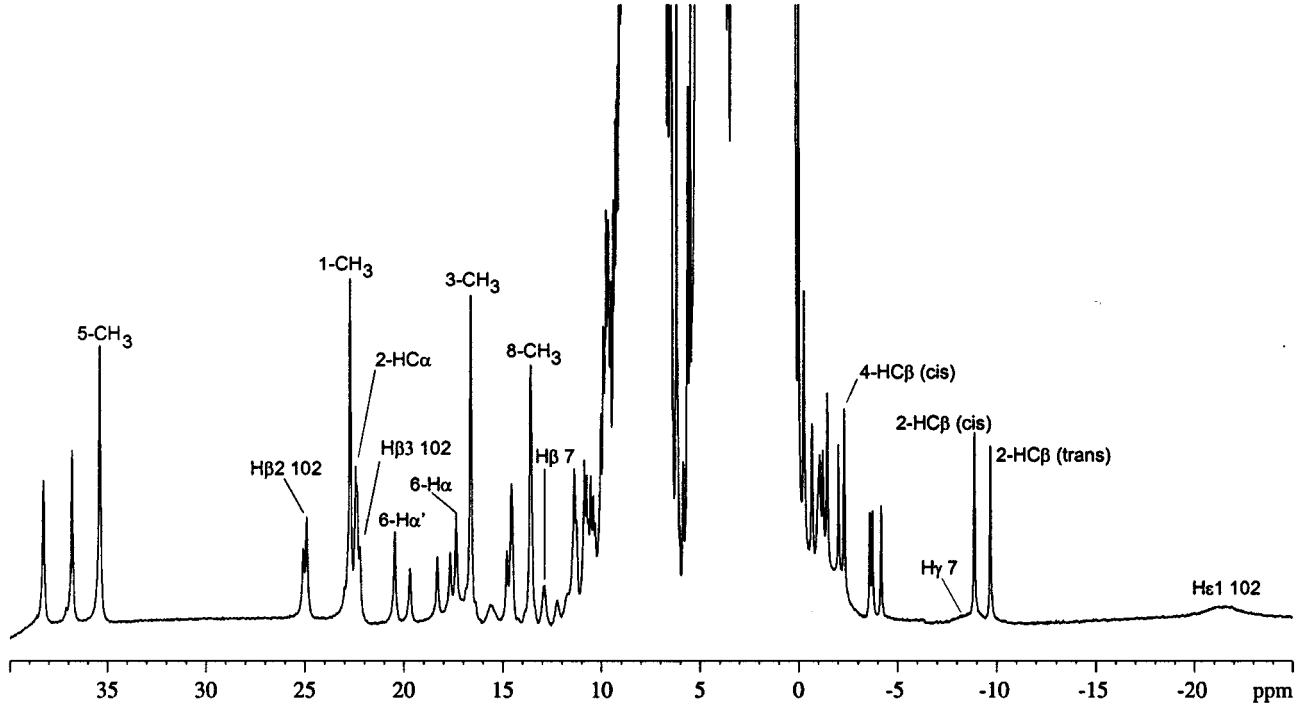

FIGURE 1: ^{1}H NMR 800 MHz spectrum of oxidized *E. coli* cytochrome b_{562} in 500 mM phosphate buffer, pH 4.8, at 298 K. The most shifted signals of the major form of the protein due to the protons of the heme and of the axial ligands of the iron are labeled.

proton resonances from these two species differ in chemical shift at most by 0.2 ppm. However, the presence of a minor form did not prevent assignment of the major form. For a large number of residues, mainly around the heme, the two forms can be resolved, thanks to the fact that the spectra were acquired at a magnetic field as high as 800 MHz, increasing the resolution. A ^{1}H−^{15}N HMQC spectrum is reported in Figure 2 which shows the level of resolution achieved. We will focus hereafter on form A.

Assignments in the diamagnetic region were performed through 3D NOESY-^{15}N HMQC and TOCSY-^{15}N HMQC experiments. The dispersion of the resonances in the ^{15}N dimension in the 3D maps allowed us to resolve most of the connectivities in the fingerprint region, which were instead overlapped in 2D spectra. This overlap, involving mainly Hα protons, can been attributed to the large content of helical structure elements (see later). The first assigned residues were glycines, due to their specific ^{15}N chemical shifts, and threonines, alanines, and valines, as they have easily recognizable spin patterns. For some residues, the TOCSY connectivities, if not absent at all (around 10% of the residues), were restricted to backbone atoms or to the β protons while those of the rest of the side chain were not detectable. This is similar to what is observed for the apo form of cyt b_{562} (hereafter apo cyt b_{562}) (*47*). The missing resonances were assigned based on COSY spectra and/or on the analysis of the NOESY spectra in light of the structural models obtained without such assignments. The assignment spans all 106 residues, and is reported in detail in the Supporting Information.

In the present oxidized form of the protein, the occurrence of one unpaired electron on the iron atom, in the predominant $S = 1/2$ spin state, gives rise to paramagnetic shifted signals (*45*, *46*). These proton resonances were easily located in the 2D NOESY experiment recorded at shorter mixing times on the larger spectral window, optimized for the detection of connectivities among fast relaxing signals. This spectrum has been used to assign the heme resonances and some other paramagnetic shifted signals due to protons of residues in the proximity of the heme. The observed NOE pattern among heme resonances is such that it is possible to "walk" from propionate 6 to propionate 7 passing through all heme substituents, thus giving confidence in the present assignment. The γ-meso proton (i.e., the proton between the two propionates) could not be assigned. The assignment of the broad hyperfine-shifted signals from the axial iron ligands was achieved using 1D NOE experiments. One Hγ proton of the axial Met was assigned to the signal at −8.1 ppm, and the Hϵ1 proton of the histidine ring was assigned to the signal at −21.1 ppm. From 1D NOE experiments, where the latter signal has been saturated, and 1D SuperWEFT experiments, the resonance of Hδ1 at 8.1 ppm has been located. 1D NOE experiments at 276 K allowed us to detect and assign the resonance of the ϵCH$_3$ of Met 7 at −6.5 ppm. The assignment of the paramagnetically shifted signals essentially is in agreement to that already reported (*46*), except for the α-meso proton, which is assigned at 3.7 ppm instead of −1.1 ppm, as previously reported.

We have also followed the 1D NMR spectra with pH, and confirmed the previously available results (*45*). The pH dependence is consistent with the two reported pK_a's at 6.8 and 9.0 (*45*).

In total, 106 residues out of 106 were assigned; 85.1% and 85.0% of the expected ^{1}H and ^{15}N resonances, respectively, were assigned (Supporting Information).

Secondary Structure from NMR Data. The elements of secondary structure can already be identified by analyzing the pattern of assigned NOEs (*48*). Short- and medium-range NOEs were used to generate Figure 3. In general, β strands are expected to give strong $d_{\alpha N}$ intraresidue and sequential connectivities and weak d_{NN} connectivities, while helical structures can be identified by the high number of sequential

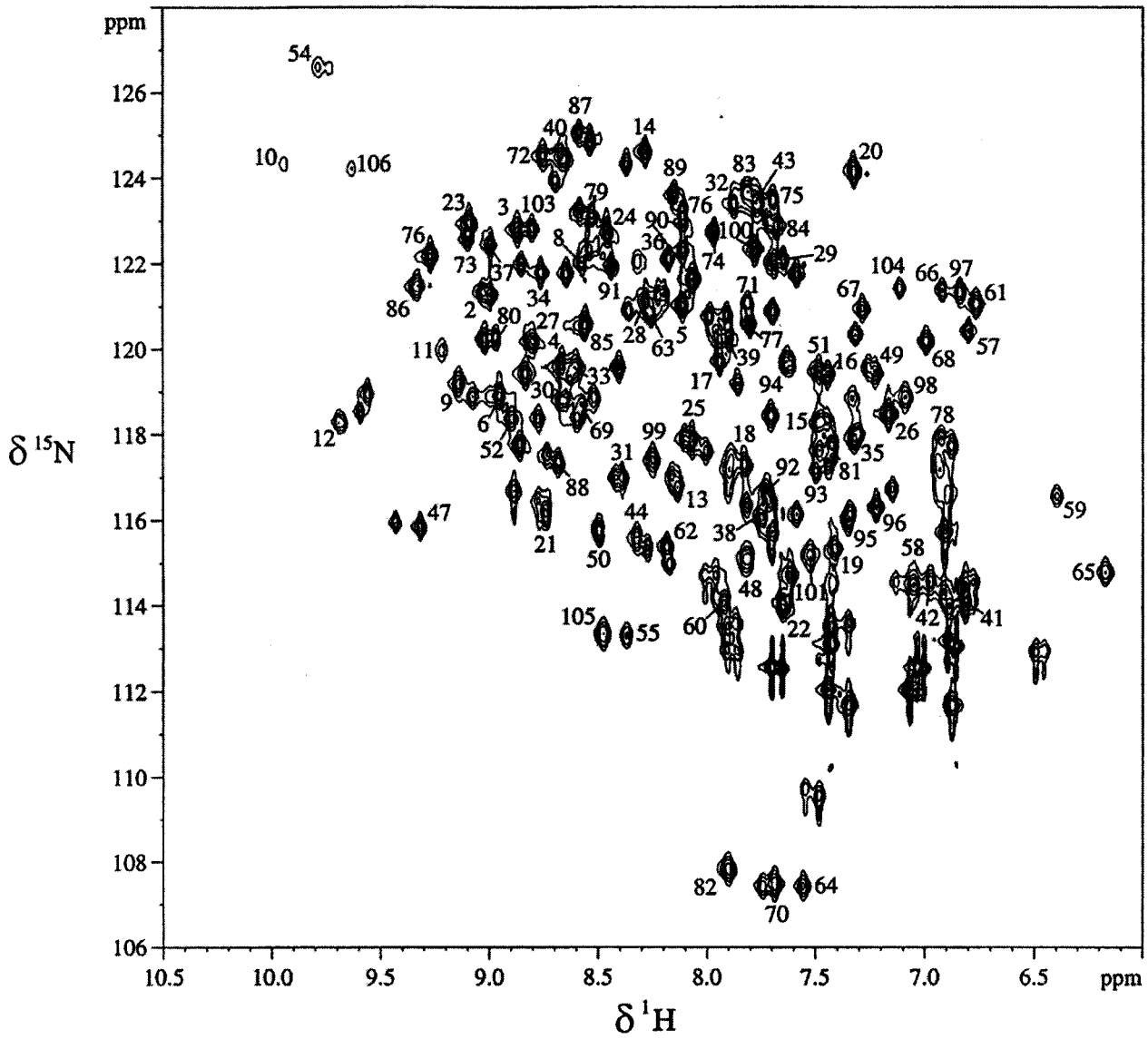

FIGURE 2: $^1H-^{15}N$ HMQC 800 MHz spectrum of oxidized *E. coli* cytochrome b_{562} in 500 mM phosphate buffer, pH 4.8, at 298 K. The peaks of the major form are labeled according to the residue assignment.

and medium-range connectivities such as $d_{NN}(i,i+1)$, $d_{NN}(i,i+2)$, $d_{\alpha N}(i,i+3)$, and $d_{\alpha N}(i,i+4)$, $d_{\alpha\beta}$ (i,i+3). Four elements of helical secondary structure can be predicted similarly to the characteristic secondary structure elements present in the cyt b_{562} crystal structure (*13, 49*). These are commonly referred to as helices α1, α2, α3, and α4, and in the present case they involve residues 3−21, 23−43, 55−83, and 84−104. Thus, about 80% of the protein residues are involved in α-helices. The region involving residues 45−55 constitutes a long loop between helices α2 and α3. Here, the experimental NOE pattern suggests the presence of a helical structure for residues 46−51.

Figure 3 also shows residues whose backbone amide protons were still observable, with full or partial intensity, after the sample had been dissolved in D_2O for 3 days. Most residues in the sequential stretches formed by residues 4−20 (besides residue 6, helix α1), 26−43 (besides residue 27, helix α2), and 57−80 (besides residues 60 and 69, helix α3) are nonexchanging, or display reduced intensity due to partial exchange. This is consistent with the presence of stable α-helices in these regions, in agreement with the observed NOE patterns. For the region 84−104, for which the experimental NOE pattern is indicative of the presence of an helical structure, only the amide protons of a few residues are not exchanging (88−91, 94−96, 99) or partially exchanging (84−85), indicating that this helix is more solvent-accessible than the other three. The longest stretch of exchanging backbone amide protons is that of residues 44−54 (which contains three prolines), whereas the second longest involves the C-terminus of the protein, spanning residues 100−106.

Solution Structure Calculations and Secondary Structure Analysis. A total of 3527 NOESY cross-peaks were assigned, integrated, and transformed into upper distance limits with the program CALIBA. These corresponded to 2051 upper distance limits, of which 1653 were found to be meaningful (nonmeaningful distance constraints are those which cannot be violated in any structure conformation and those involving proton pairs at fixed distance). The number of NOEs per residue is reported in Figure 4. No correction has been introduced to compensate for the fact that peaks between signals that are degenerate in the two forms are stronger than peaks between resolved signals due to protons at the same distance. Neglecting the correction determines an increase

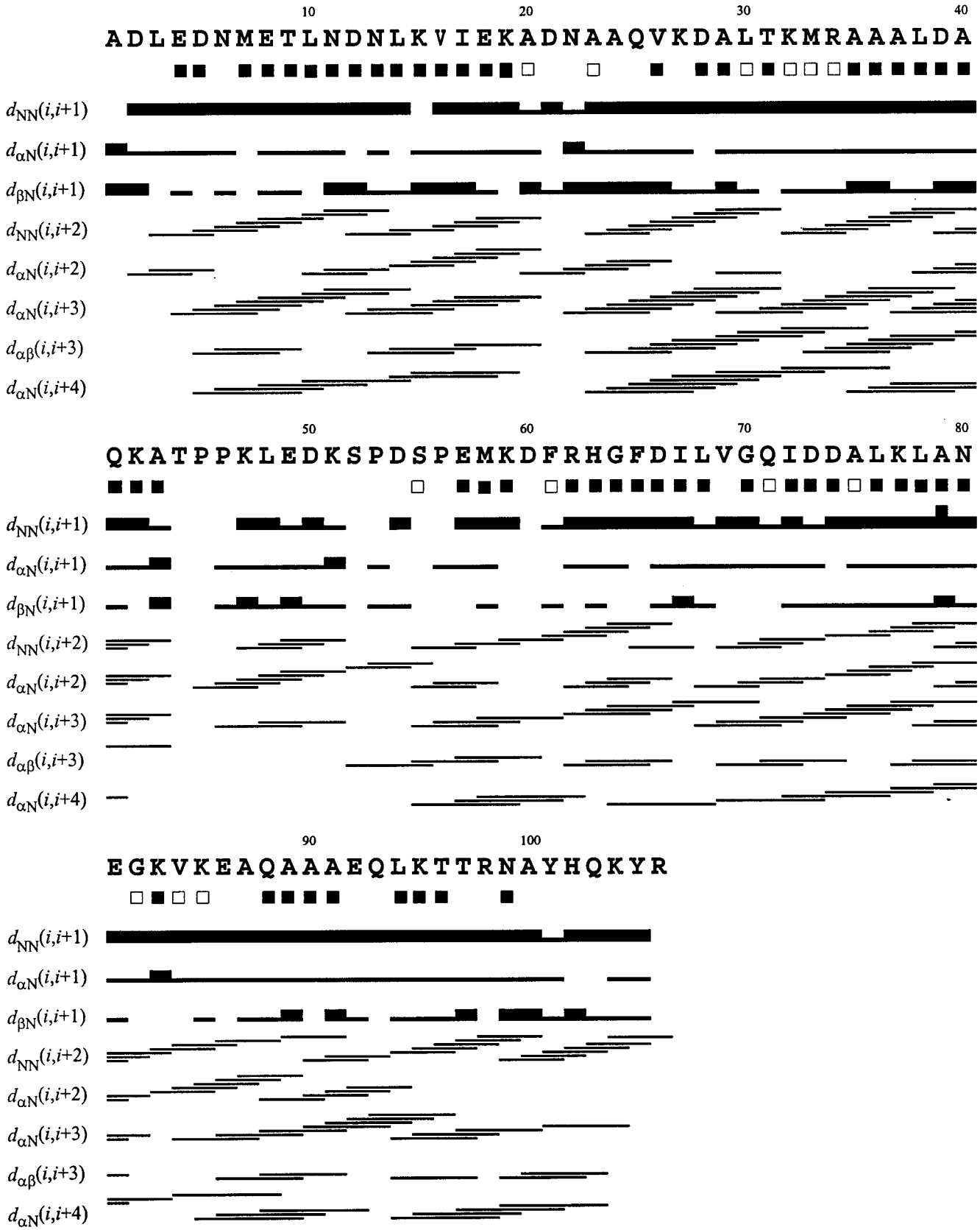

FIGURE 3: Schematic representation of the sequential and medium-range NOE connectivities involving NH, Hα, and Hβ protons for the oxidized form of cyt b_{562}. The thickness of the bar indicates the intensity of NOEs. Filled squares represent HN protons which are not exchanged after 3 days in D_2O; open squares represent HN protons which are only partially exchanged after 3 days in D_2O (see text for details).

of the upper distance limits derived from the peaks between resolved signals of 6−7%, which is well below the uncertainty on such distance limits due to the integration error. It should be noticed that despite the paramagnetism of the system, the heme protons provided 200 NOEs, while the axial ligands His 102 and Met 7 provided 30 and 24 NOEs,

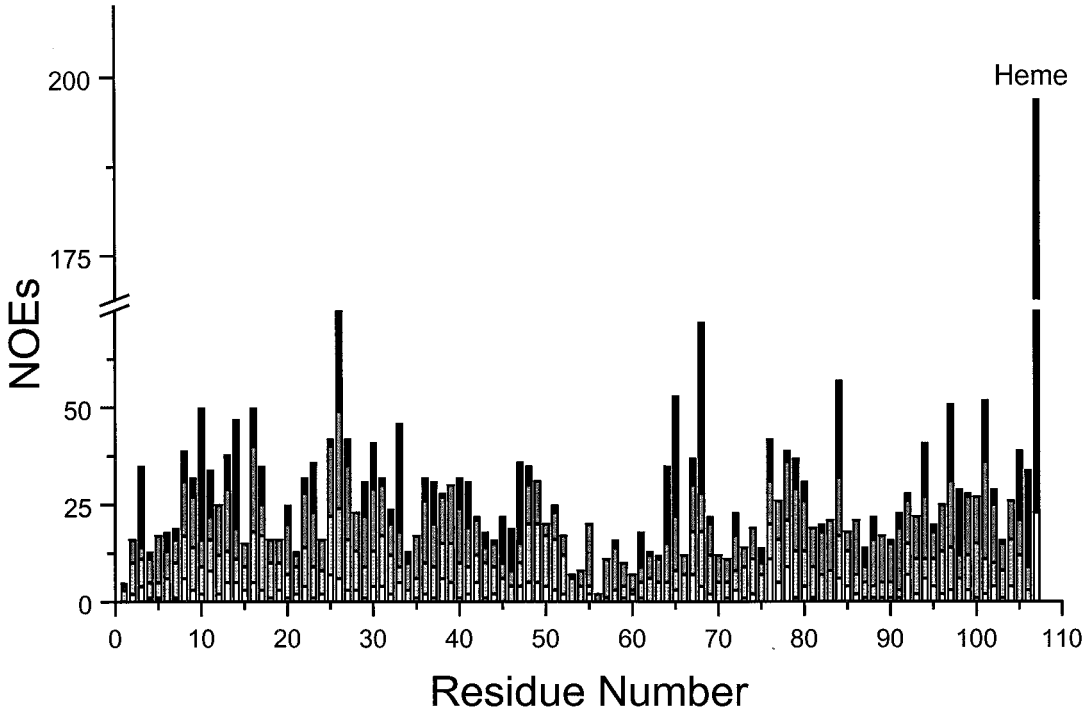

FIGURE 4: Number of meaningful NOE constraints per residue for oxidized *E. coli* cyt b_{562}. White, light gray, dark gray, and black bars indicate, respectively, intraresidue, sequential, medium-, and long-range connectivities. NOEs involving the heme moiety are indicated in the last column on the right.

respectively. Some of them have been obtained from 1D NOE. In addition to the above NOE constraints, 33 $^3J_{HNH\alpha}$ couplings obtained from the HNHA 3D spectrum were employed in structure calculations. Hydrogen bond constraints for 36 amide protons which fulfill the requirements described under Materials and Methods were used at later stages of structural calculations. A total of 34 proton pairs were stereospecifically assigned through the program GLOMSA. In particular, the Hβ protons of the axial histidine were assigned stereospecifically. All the constraints used for structure calculations and the obtained stereospecific assignments are reported in the Supporting Information.

The 35 conformers constituting the family obtained with the above constraints have an average target function of 0.78 $\pm$ 0.07 Å^2, and a RMSD from the mean structure for the backbone atoms of residues 3−105 of 0.97 $\pm$ 0.20 Å (see later for the analysis).

The structure was also calculated using pcs as further constraints (see Materials and Methods). The latter were used within the point dipole approximation, i.e., where the electron is considered localized on the iron atom. The relation between pcs and nuclear coordinates is (*50*):

$$\delta^{pcs} = \frac{1}{12\pi r_i^3}\left[\Delta\chi_{ax}(3n_i^2 - 1) + \frac{3}{2}\Delta\chi_{rh}(l_i^2 - m_i^2)\right] \quad (1)$$

where $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ are the axial and the rhombic anisotropies of the magnetic susceptibility tensor, r_i is the distance of nucleus i from the ion, and l_i, m_i, and n_i are the direction cosines of the position vector of atom i with respect to the orthogonal reference system formed by the principal axes of the magnetic susceptibility tensor.

The first use of pcs for solution structure determination dates back to 1996 (*37*), and a critical analysis is reported in (*30, 38, 51, 52*). Pcs are properly defined for the nuclei of the residues not directly bound to the metal ion. In this case, His and Met and the heme protons are not considered. Pcs are defined as the difference between the actual shift and the shifts experienced by the same nucleus in the absence of the unpaired electron. Two points should be kept in mind in this treatment: (*i*) the electron is indeed delocalized on the heme moiety and on the axial ligands, thus undermining the validity of eq 1 (see below); (*ii*) the shifts for the species without unpaired electrons (diamagnetic shifts) can only be estimated. As reported under Materials and Methods, point (*i*) is overcome by introducing a tolerance proportional to the actual pcs; larger pcs, which are experienced by protons close to the paramagnetic center and are more affected by electron delocalization, are considered to have a larger indetermination. As far as point (*ii*) is concerned, the diamagnetic chemical shift can be calculated from the structure obtained from NOEs or from the shifts of the reduced diamagnetic protein. In the latter case the oxidized and reduced proteins are assumed not to differ in structure so much to provide a difference in the diamagnetic chemical shift larger than the indetermination of pcs. Such indetermination is reflected in the tolerance given to the pcs in the structure determination protocol. In the present case, the diamagnetic chemical shifts were calculated with the program SHIFT (*34*) from the energy-minimized average structure obtained from NOEs. These calculated shifts were compared with those calculated from the X-ray structure. They compare well to each other, having a mean difference of 0.05 $\pm$ 0.23 ppm, and are reported in the Supporting Information. Since such calculations provide a considerable error, a tolerance of 0.5 ppm was given to all pcs. If some pcs were wrong they would be identified as they would induce an increase in the target function of the NOEs.

The new family of 35 structures obtained with the inclusion of pcs has no increase in the target function of the

Table 1: Calculation Statistics and Structural Analysis for the Final PSEUDOREM Family of Oxidized *E. coli* Cyt b_{562}

Total Number of NOE Constraints	
intraresidue	290
sequential	410
medium range[a]	523
long range	430
total	1653
RMS Violation per NOE Constraint (Average over the Family)	
intraresidue	0.0181 ± 0.0025 Å
sequential	0.0151 ± 0.0023 Å
medium range[a]	0.0213 ± 0.0018 Å
long range	0.0124 ± 0.0028 Å
total	0.0173 ± 0.0014 Å
Average Number of NOE Violations Greater than 0.05 Å per Structure	
intraresidue	10.4 ± 2.3
sequential	9.5 ± 2.6
medium range[a]	21.6 ± 3.2
long range	6.7 ± 2.1
total	48.2 ± 5.8
av no. of NOE violations larger than 0.3 Å	0.0 ± 0.0
av no. of ϕ dihedral angle violations larger than 5°	0.0 ± 0.0
largest residual NOE violation	0.25 Å
RMS violation/pseudocontact shift (av over family)	0.0059 ± 0.0006 ppm
av NOE and pcs penalty function	78 ± 11 kJ mol^{-1}
PROCHECK-NMR Analysis	
residues in most favorable regions[b]	84.8%
residues in allowed regions[b]	13.5%
residues in generously regions[b]	1.0%
residues in disallowed regions[b]	0.7%

[a] Medium-range distance constraints are those between residues (i,i+2), (i,i+3), (i,i+4) and (i,i+5). [b] As it results from the Ramachandran plot analysis.

NOEs (0.74 ± 0.13 Å^2) with respect to the family obtained without pcs and essentially the same RMSD (0.93 ± 0.20 Å for residues 3−105). In addition, the contribution to the target function of the pcs constraints is small (ca. 0.05 Å^2). The consistency of the two sets of constraints indicates that the assignment of protons is correct and the structure is reliable. The obtained family was then subjected to a refinement with the program AMBER as already reported (*38*). The PSEUDOREM family has an NOE and pcs penalty function of 96 ± 17 kJ mol^{-1}, corresponding to an average PSEUDYANA NOE and pcs target function of 0.8 Å^2.

The average violation per NOE constraint is 0.0173 ± 0.0014 Å (Table 1). This family is taken as representing the solution structure. A superposition of the backbone atoms of the family is shown in Figure 5. Table 1 reports calculation statistics and the structure quality analysis for the final family. These low values of the target function indicate that the obtained family is in very good agreement with the experimental constraints. The average backbone RMSD from the mean structure for residues 3−105 is 0.93 ± 0.21 Å, whereas the heavy atoms RMSD is 1.39 ± 0.20 Å. These values drop to 0.42 ± 0.09 Å and 0.96 ± 0.09 Å, respectively, when the disordered region formed by residues 52−61 is excluded. The RMSD values per residue of the final family from the mean structure are shown in Figure 6. The largest backbone RMSD values are obtained for residues 52−61; the rest of the protein is, on the contrary, very well-defined, with local RMSD maxima, as usual, for the N- and C-terminal residues.

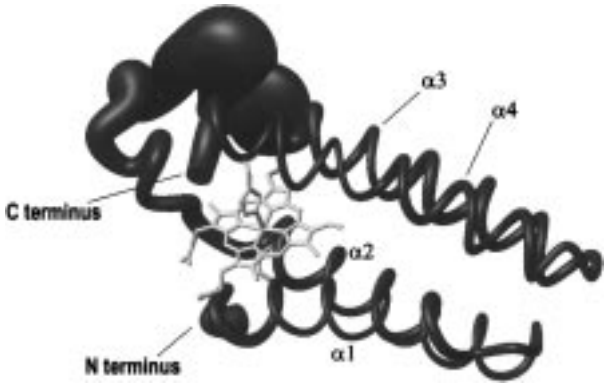

FIGURE 5: Display of the backbone atoms for the solution structure of oxidized *E. coli* cyt b_{562} as a tube with variable radius, proportional to the backbone RMSD of each residue. The figure was generated with the program MOLMOL (*66*). The heme moiety is also shown.

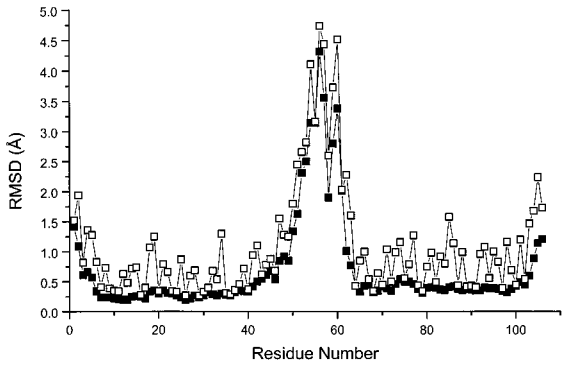

FIGURE 6: RMSD per residue from the mean structure in oxidized *E. coli* cyt b_{562} for the backbone (filled squares) and all heavy atoms (open squares) of the PSEUDOREM family of 40 conformers of the oxidized form of cyt b_{562}.

The high RMSD of residues 52−61 is due to the paucity of NOEs in this region (Figure 4). In particular, almost no long-range NOEs were obtained for residues 51−60; the highest RMSD is obtained for residue 56, which has the lowest number of NOE constraints (Figure 4). Residues 50−55 constitute the C-terminal part of the loop connecting helices α1 and α2. This loop points away from the core of the protein, making it intrinsically poor in the number of observable NOEs. In addition, this region contains two prolines, which break up the sequential connectivities. This region is also quite disordered in the X-ray structure (see later).

The final family of conformers was analyzed in terms of secondary structure with PROCHECK-NMR (*44*). According to this program, helical structures were found for residues 3−18 (α1), 22−41 (α2), 45−49 (3_{10}), 62−82 (α3), and 83−105 (α4); and in the energy-minimized mean structure, residues 3−18 (α1), 22−41 (α2), 45−49 (3_{10}), 55−82 (α3, with a break at residue 61), and 83−106 (α4). This result is similar to that obtained from the analysis of NOE patterns (see above). Only helix α3 is sizably shorter than what is expected on the basis of the sequential NOEs. From inspection of Figure 6, it can be noticed that three of the four α-helices are very well-defined, with an average backbone RMSD lower than 0.5 Å, whereas the N-terminal part of helix α3 (residues 55−60) is rather poorly defined, due to the low number of NOE constraints in this region

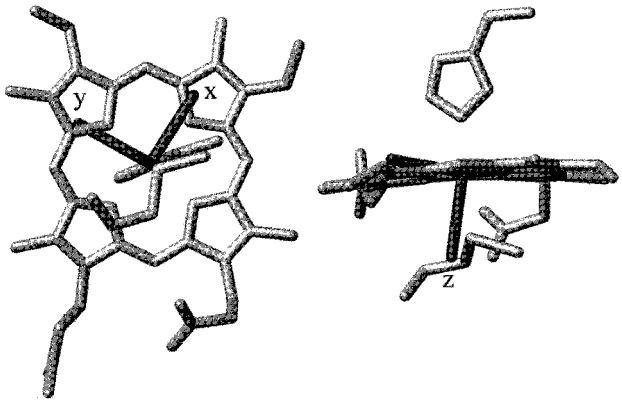

FIGURE 7: Display of the orientation of the magnetic susceptibility anisotropy axes in oxidized *E. coli* cyt b_{562} with respect to the heme moiety.

(Figure 4). Indeed, despite the poor definition, the NOE pattern indicates that segment 55−60 is part of the α3 helix.

The present family of conformers has ca. 85% of the residues in the most favored regions of the Ramachandran plot, and ca. 13% in the allowed regions. No residues were in the disallowed regions. In the energy-minimized average structure, 88% of the residues are in the most favored regions of the Ramachandran plot, 11% of the residues are in the allowed regions, and only 1 residue is in a generously allowed region.

Magnetic Anisotropy Tensor and Contact Shift Pattern. The final $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ values are $(1.60 \pm 0.05) \times 10^{-32}$ m^3 and $(-0.69 \pm 0.08) \times 10^{-32}$ m^3, respectively. The *z* axis of the magnetic susceptibility tensor deviates slightly from the normal to the heme plane, making with it an angle of about 20°. The *y* axis makes an angle of 27° with the Fe−N(pyrrole I) direction. The arrangement of the tensor with respect to the heme moiety and the ligands is shown in Figure 7. It is known that the directions of the magnetic susceptibility tensor axes are determined by the relative arrangement of the side chains of the axial ligands (*53*). Indeed, when a single axial ligand, giving rise to an asymmetric π interaction with the metal (e.g., His), is bound to the iron ion, the *y* axis of the magnetic susceptibility tensor rotates counterclockwise when the His plane rotates clockwise, by essential the same angle (*37*, *54*−*56*). The *y* axis coincides with the normal to the His plane when the latter is aligned along a N−N direction. The same would hold if a methionine were the single axial ligand, with the nodal plane of the lone pair of the sulfur taking the place of the His plane. When, as in the present case, there are two axial ligands, the two effects sum up, scaled by the relative strength of their π interaction with the iron ion. In the present case, the His plane makes an angle of about 30° with the N(pyrrole II)−N(pyrrole IV) direction, whereas the nodal plane of the lone pair of the sulfur atom of the Met makes an angle of about −5° with the same direction. By simply averaging the two rotations, one would expect the *y* direction of the magnetic susceptibility anisotropy tensor to make an angle of 12.5° with the N(pyrrole I)−N(pyrrole III) direction, which is in good agreement with the observed value of 27°, confirming that the π interactions with the iron ion of the two axial ligands are of comparable strength (*57*). The deviation of the *z* axis from the normal to the heme plane is rather small, consistent with the fact that both Fe−Nϵ2 and Fe−Sδ bonds are essentially orthogonal to the heme plane (within 2−3°). This holds both in the X-ray and in the present solution structure. Similar deviations had already been reported for other *c*-type cytochromes (*41*, *58*).

Once the magnetic susceptibility anisotropy tensor is available, it is possible to evaluate through eq 1 the pcs contribution to the total hyperfine shift experienced by nuclei of the heme moiety and of the axial ligands. The diamagnetic values for the heme protons, which are necessary to determine the hyperfine shifts, are taken from the previously reported shifts of reduced, diamagnetic cyt b_{562} (*15*). The diamagnetic values of the axial ligands were instead taken from average values measured in diamagnetic *c*-type cytochromes (*59*, *60*), and are in excellent agreement with preliminary assignments of reduced cyt b_{562} obtained in this lab (not shown). When the pcs are subtracted from the total hyperfine shifts, the contact shifts are obtained. The results are shown in Table 2. The pattern of contact shifts observed for the methyls is in good agreement with what is expected on the basis of simple Hückel calculations and the present arrangement of the ligands (*57*). The most striking feature observed in the present system is the fairly large difference between the contact shifts of methyls 5 and 1, which are instead expected to be equal.

DISCUSSION

The solution structure of oxidized *E. coli* cyt b_{562} consists of four long α helices, packed together in an antiparallel fashion (Figure 5). This arrangement corresponds to the well-known four-helix bundle topology. As can be seen from Figure 6, the largest values of the RMSD are observed for residues 51−61. Local maxima of the RMSD are also observed for the N- and C-terminal parts of the protein, as it is customary for solution structures of proteins. The heme cofactor is located between helices α1 and α4.

Table 1 shows that the present PSEUDOREM structure is in very good agreement with the experimental constraints, having no violation larger than 0.3 Å. The relatively low average RMS violation of the long-range NOE constraints ensures that this class of constraints, which is the most important in defining the overall fold of the protein, has not been calibrated too tightly.

Comparison with the X-ray Structure. The overall fold of the present solution structure is essentially identical to that observed in the solid state (*13*). The X-ray structure of cyt b_{562} (PDB entry 256B) showed that two nonequivalent molecules are present in the unit cell. These two molecules are slightly different, with a mean backbone RMSD for the superposition of residues 3−51 and 62−105 of 0.26 Å. The RMSD for the same region between the energy-minimized mean solution structure and the two solid state conformers is 0.94 Å for molecule A and 1.02 for molecule B.

The largest differences between the two solid-state molecules are observed for the region 50−57, which is a portion of the long loop between helices α2 and α3, and for the C-terminal helix α4. The first region is highly disordered in the crystal structure, and indeed only a very weak electron density could be observed for residues 45−59 at a resolution lower than 2 Å (*13*). In the solution structure these residues constitute the region with the highest RMSD values (Figure 6). It is thus conceivable that this loop spans a number of

Table 2: Separation of Contact and Pseudocontact Contributions to the Hyperfine Shift in Oxidized *E. coli* cyt b_{562} at 298 K

atom name	residue name	chemical shift, oxidized species (ppm)	chemical shift, reduced species (ppm)	hyperfine shift (ppm)	calculated pseudocontact shift (ppm)	contact shift (ppm)
HN	Met	10.7	6.8	3.9	1.3 ± 0.3	2.6
Hα	Met	4.1	1.2	2.9	5.7 ± 1.2	−2.8
Hβ2/Hβ3	Met	12.8	0.7	13.5	−0.5 ± 2	14.0
Hγ2/Hγ3	Met	−8	−2	−6.0	3 ± 3	−9
ϵ-CH_3	Met	−6.5[a]	−3.2	−3.3	6.5 ± 4.5	−9.8
HN	His	10.5	6.0	4.5	0.8 ± 0.2	3.7
Hα	His	NA[b]			2.1 ± 0.4	NA
Hδ1	His	8.1	0.2	7.9	6.0 ± 0.5	1.9
Hδ2	His	NA			2.8 ± 1.6	NA
Hϵ1	His	−21.1	−1.6	−19.5	13 ± 5	−32.5
8-CH_3	heme	13.6	3.44	10.2	−2.4 ± 0.7	12.6
meso-δH	heme	11.4	9.91	1.5	−2.5 ± 0.7	4.0
1-CH_3	heme	22.7	3.68	19.0	−0.3 ± 0.3	19.3
2-Hα	heme	22.4	8.45	14.0	−0.8 ± 0.2	14.8
2-Hβ (trans)	heme	−8.9	5.63	−14.5	−0.25 ± 0.04	−14.3
2-Hβ (cis)	heme	−9.7	5.94	−15.6	−0.10 ± 0.04	−15.5
meso-αH	heme	3.7	9.40	−5.7	−4.40 ± 0.25	−1.3
3-CH_3	heme	16.6	2.69	13.9	−2.7 ± 0.1	16.6
4-Hα	heme	8.4	8.29	−0.11	−1.7 ± 0.2	1.6
4- Hβ (trans)	heme	−0.3	4.30	−4.6	−1.5 ± 0.2	−3.1
4- Hβ (cis)	heme	−2.4	5.68	−8.1	−0.96 ± 0.07	−7.1
meso-βH	heme	1.5	9.98	−8.5	−3.2 ± 0.5	−5.3
5-CH_3	heme	35.3	3.78	31.5	−0.5 ± 0.4	32.0

[a] At 276 K. [b] NA, not assigned.

Table 3: Backbone RMSD (Å) between the Mean Oxidized Cyt b_{562} Solution Structure and the X-ray and Apoprotein Structures, Calculated by Superimposing the Backbone Atoms of Each Element of the Secondary Structure Individually

residues	X-ray	apoprotein
4−20 (α1)	0.48	1.54
23−43 (α2)	0.47	0.83
45−55 (loop)	0.81	2.76
62−83 (α3)[a]	0.41	1.16
84−105 (α4)	0.83	3.85

[a] The N-terminal part of the helix, which is disordered in the present structure, has been omitted.

alternative conformations both in the crystal and in solution, in thermodynamic equilibrium. Backbone amide hydrogen exchangeability indicates that residues 44−54, which are located just before helix α3, are solvent-accessible whereas residues 57−59 (constituting the N-terminal portion of helix α3) are not (Figure 3). This suggests that the first half of region 50−60 is more exposed than the second half. Helix α4 is well-ordered in the crystal structure (i.e., electron density could be observed for this region of the molecule), but its RMSD between the two molecules of the cell is larger than in the rest of the protein. Helix α4 is also very well-defined in the NMR structure, having an average RMSD from the mean of 0.4 Å. There is a significant deviation between the backbone conformation of helix α4 observed in solution and in the two X-ray structures, with an average RMSD of ca. 1.0 Å with respect to both solid-state molecules (when the molecules are superimposed on the backbone atoms of residues 3−51 and 62−105). Backbone amide proton exchange shows that this region is accessible to the solvent. All of these findings are consistent with the idea of helix α4 being structurally sensitive to the chemical environment. Table 3 reports a comparison of each individual secondary structure element at the time. It is readily observed that the largest local (i.e., neglecting changes due to different relative orientations of the helices) conformational rearrangements occurring when passing from the crystal to the solution are localized in helix α4.

The conformations of several side chains of charged residues (Asp, Glu, Arg, Lys) and of Leu 76 differ sizably in the two solid-state molecules. Unsurprisingly, these side chains also show a marked displacement when comparing the crystal and solution structures. In the latter, most of them are found in conformations comparably different from those observed in each molecule of the crystal. Exceptions are observed for Glu8 and for the neutral side chain of Leu 76, which are much closer to the molecule indicated as B in the crystal structure than to that indicated as A. The side chains whose conformation is the same in the two crystal conformers, but which changes when passing to solution, are again involving charged residues. Among these, Lys 42 is experiencing a large change. This residue is located at the surface of the molecule, and forms a salt bridge in the solid-state structure with Asp 39. In solution, this salt bridge is not present, as the two side chains are extending in solution.

The two molecules present in the crystal cell are in close contact with each other in region 99−103. The superposition of this region for the solution and crystal structures reveals that its local conformation is maintained when passing from the solid state to the solution.

The conformation of propionate 7 in the X-ray structure is different in the two molecules (*13*). In solution, its conformation is rather well-defined, and it is similar to that observed in the molecule identified as B in the crystal structure (Figure 8). It is noteworthy that in this conformation the propionate is not forming any hydrogen bond to surrounding residues either in solution or in the crystal. In the A molecule of the X-ray structure, propionate 7 forms hydrogen bonds with the carbonyl oxygen of Glu 4, and with the carboxylate of Glu 8. In the present structure, the carboxylate of Glu 8 is at 8 Å from the carboxylate of propionate 7. The carbonyl oxygen of Glu 4 is at 6 Å from the propionate.

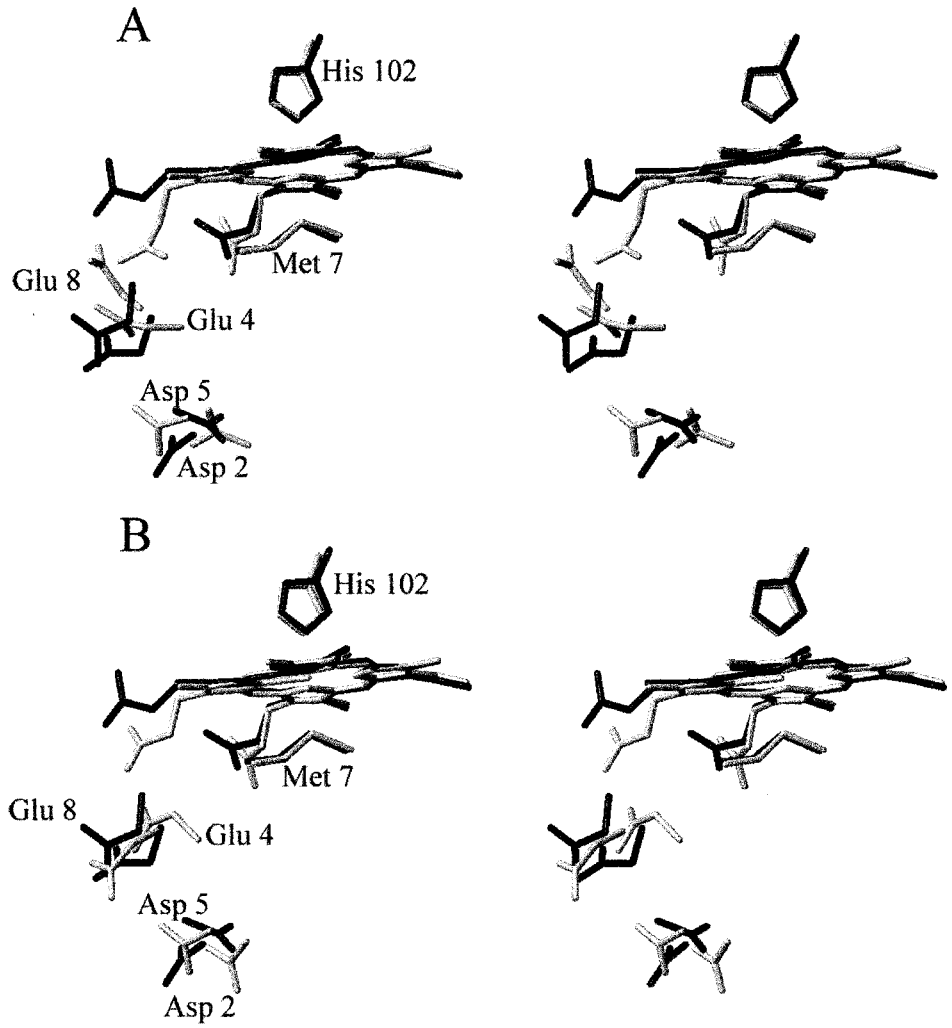

FIGURE 8: Comparison of the conformation of the heme moiety and of some selected residues (Asp 2, Glu 4, Asp 5, Met 7, Glu 8, His 102) in the present solution structure (black) of *E. coli* cyt b_{562} and in the molecules indicated as A (panel A) and B (panel B) in the crystal structure (gray).

Propionate 6 is less defined than propionate 7 in the present structure, but its conformation is close to that observed in the crystal (which is very similar in the two molecules) (Figure 8). In the crystal structure, propionate 6 formed hydrogen bonds only with surrounding water molecules. Also in solution, no hydrogen bonds to the protein were observed.

Comparison with the Apoprotein Structure. The solution structure of apo cyt b_{562} (PDB entry 1APC) (*14*) displays large deviations from both the X-ray and the present NMR structures of the holoprotein. Indeed, the RMSD between the present mean solution structure and the mean structure of apo cyt b_{562} is 3.3 Å (for the superposition of residues 3−51 and 62−105), whereas the RMSD between the latter and the X-ray structure of the holoprotein is 3.1 Å. A superposition of the present structure and that of apo cyt b_{562} is shown in Figure 9. The largest differences between the present structure and the structure of the apoprotein are observed for helix α1 (mean backbone RMSD for residues 3−23 of 3.1 Å), the loop (3−7 Å), and the last part of helix α4 (mean backbone RMSD for residues 82−104 of 4.5 Å). Helix α2 is the most similar between the two solution structures of the apo- and holoprotein. The above is true also for comparison between the X-ray structure of the holoprotein and the solution structure of the apoprotein (*14*). It is interesting to observe that when the apo- and holoprotein solution structures are superimposed on one element of the secondary structure at a time (i.e., by best-fitting the coordinates of the backbone atoms of the residues involved in each helix or loop), helix α4 displays the largest RMSD values among the four helices (Table 3). This shows that this helix undergoes extensive local conformational rearrangements when the heme cofactor is removed from the protein.

Comparison with Cytochromes c′. The crystal structures of several cytochromes *c′* from different organisms are available (*Rhodopseudomonas palustris*, PDB entry 1A7V; *Chromatium vinosum*, PDB entry 1BBH; *Alcaligenes denitrificans*, PDB entry 1CGN; *Alcaligenes sp.*, PDB entry 1CGO; *Rhodobacter capsulatus*, PDB entry 1CPQ, 1RCP; *Rhodobacter capsulatus* strain St. Louis, PDB entry 1CPR; *Rhodocyclus gelatinosus*, PDB entry 1JAF; *Rhodospirillum molischianum*, PDB entry 2CCY). The sequence identity between cyt b_{562} and the various cytochromes *c′* is rather low, ranging from 10% to 20%. Among the cytochromes *c′*, the sequence similarity is somewhat higher, with 20−50% amino acid identity (besides the *Alcaligenes denitrificans* and *Alcaligenes sp.* proteins, whose sequences differ only at four positions). The only strictly conserved residue in the sequences of cyt b_{562} and of the above listed cytochromes *c′*

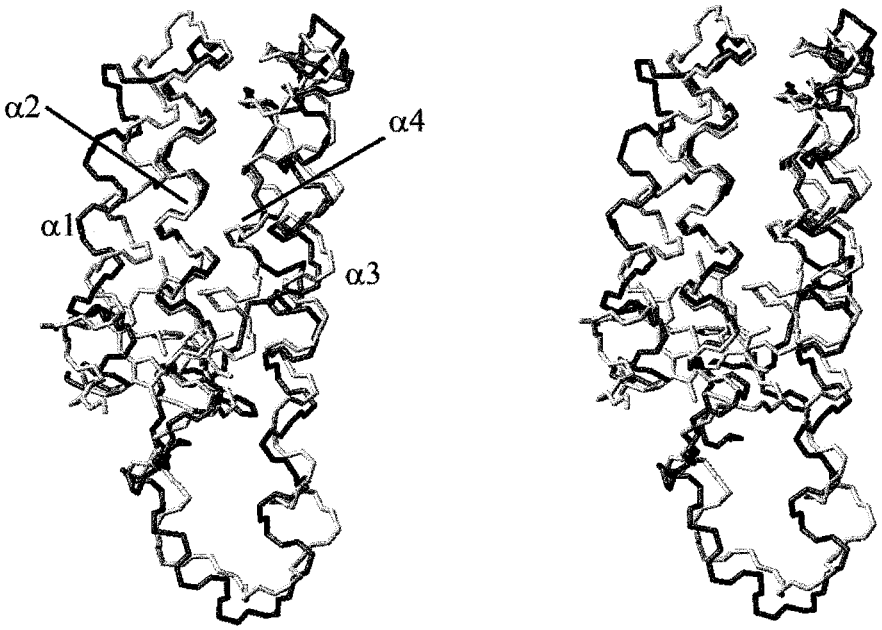

FIGURE 9: Superposition of the backbone atoms for the present solution structure of *E. coli* holo cyt b_{562} (gray) and the solution structure of the apoprotein (black) (*14*). The side chains of Met 7 and His 102 are also shown.

is His102 (cyt b_{562} numbering). Conservative substitutions are found at positions 10, 17, 89, 105, and 106. Five to nine additional residues with respect to the sequence of cyt b_{562} are present at the N-terminus of the cytochrome *c*′ sequences, and two to four at the C-terminus. Most of the insertions and deletions occur in the loops connecting the four helices.

Despite this low homology, the two classes of proteins display an impressively similar folding, with the typical four-helix bundle motif. The largest RMSD values between the cyt b_{562} solution structure and the above-mentioned cytochrome *c*′ structures are observed in the loop regions, and at the termini of the protein. If these regions are discarded, the average RMSD values for the superposition of the backbone atoms of the above structures are in the 1.5−2.5 Å range, indicating that the global fold is substantially maintained. Also the orientations of the heme cofactor and of the plane of the axial histidine in cyt b_{562} are the same as in the cytochrome *c*′ structures.

pH-Dependent Properties of the Solution Structure. Previous studies had shown that the hyperfine-shifted signals in the ^{1}H NMR spectra change shifts with pH, indicating the presence of two titrable groups with pK_a's of 6.8 and 9.0 (*45*). The reduction potentials were also found to be pH dependent, with the presence of three acid−base equilibria with pK_a's differing in the oxidized with respect to the reduced form (*45*, *61*). While two groups have the same pK_a's in the oxidized form as those obtained from the NMR studies (*45*), the third has a value below 5. On the basis of the X-ray structure (and of other spectroscopic measurements, e.g., EPR), these pK_a's have been attributed to the ionization of heme propionate 7 (pK_a = 6.8) and of the iron-bound histidine (pK_a = 9.0), and possibly to the ionization of a carboxylate group for the pK_a below 5.

The two former pK_as are relatively different from those expected for the corresponding residues. The increase in the pK_a of propionate 7 was attributed to its involvement in a hydrogen bond network. The pK_a of the axial histidine, on the contrary, is relatively low, which was proposed to be due, besides to the coordination to iron(III), to the vicinity of a positively charged group. Arg 106, which in the 1.4 Å resolution X-ray structure was found at 3.7 Å from the His, can stabilize the negative charge of the histidinate anion.

The pH dependence of the hyperfine-shifted signals, measured in the present study, confirmed the previous results, showing that by changing the pH from 4.8 to 7 the most affected hyperfine-shifted resonances are those of 8-CH_3 and propionate 7. From the analysis of the 2D NOESY map recorded at pH 7, it appears that the folding of the protein is completely maintained at this pH. All the α-helices are still present and encompass essentially the same protein segments as at pH 4.8. In particular, helix α4, which was found to be disrupted in the apoprotein, is completely maintained. The region around propionate 7 shows changes both in terms of variation in the shift resonances and in the pattern and intensity of the NOEs. These structural changes are consistent with propionate 7 either being the ionizable group with the pK_a of 6.8 or involved in an acid−base network. Indeed, propionate 7, as suggested from the X-ray structure and confirmed by the present solution structure, is close to several negatively charged groups which would increase one apparent pK_a. These groups are Glu 4 (whose carboxylate oxygen atoms are at 6 Å), Glu 8 (at 7.5 Å), and, of course, propionate 6. The carboxylate moieties of Asp 2 and Asp 5 are at 5 and 7 Å from the carboxylate of Glu 4, and therefore capable of affecting one another. These titratable groups must be interacting electrostatically, with at least one of the carboxylates having a high pK_a. Preliminary results suggest that the electrostatic properties in this region of the protein are little perturbed by single mutation of these acidic residues or esterification of the heme propionate groups (Barker, Fersht, and J. K. Rice, unpublished observations).

This interpretation is further supported by comparison with the behavior of cytochromes *c*′. The *R. gelatinosus* and *C. vinosum* proteins experience also three pK_a values at about 5, 6.5, and 9.5, which have been attributed, on the basis of the analysis of the NMR data (*62*) in light of the available

X-ray structures(*63*), to the ionization of a carboxylate group, of a propionate group, and of the axially bound histidine, respectively. In the *R. molischianum* (*64*) and *R. palustris* (*65*) cytochromes *c*′, only the two higher pK_a's were observed. The deviation of the ionization constants from the normal values for the specific moiety has been interpreted as due to the vicinity of charged groups, as proposed for cyt b_{562}.

It is quite relevant that these proteins, which adopt a similar fold despite their low homology, also exhibit a very similar ionization behavior, with residues with similar properties surrounding the ionizable groups. These observations might point to a similar biological function of these two classes of proteins.

CONCLUSIONS

In this paper the solution structure of oxidized cyt b_{562} from *E. coli* was solved. Despite the paramagnetism of the molecule, the quality of the structure is satisfactory, as was observed for similar proteins (*52*), and allows a meaningful comparison with the X-ray structure of the same protein, and with the solution structure of the apoprotein. It appears that the four-helix bundle structure is maintained, and that the loop between helices α2 and α3 is very disordered. Helix α4 is the most sensitive to the chemical environment in terms of both its structural and its dynamic properties. Propionate 7 and the acidic groups surrounding it in solution are more similar to the molecule indicated as B in the crystal structure, than to the A molecule. Despite the very low sequence similarity, the overall fold and the conformation of the iron-bound histidine are found to be essentially the same in the present protein and in cytochromes *c*′, as already noted (*13*). The X-ray structure of cyt b_{562} shows a large protein−protein interface which may lead to consideration of the system as a dimer. In solution, there is a large tendency to dimerization as the protein concentration increases (*46*). The known cytochromes *c*′, except for the one from *R. palustris*, are also always dimers even at low protein concentration, but the protein−protein interface is more extensive, and the residues involved in the interaction are in different regions of the protein with respect to cyt b_{562}. The high structural similarity and the low homology in the primary sequence are interesting features for the investigation of protein folding and folding prediction.

In the present system, at least one high spin state for the Fe^{3+} ion is thermally accessible at room temperature (*46*), at variance with the case of cytochromes *c*, which have the same iron ligands. However, at the present resolution no structural strains, nor structural variation with respect to cytochromes *c*, possibly explaining this behavior could be observed. This implies that some small variations in the ligand field (e.g., arising from a different distortion of the heme moiety) are responsible for the different electronic structure. The χ tensor anisotropy and the direction of the magnetic susceptibility tensor axes are in fair agreement with the arrangement of the iron axial ligands. The magnetic anisotropy of the present system is somewhat smaller than that of cytochromes *c*.

Finally, the pH-dependent structural properties of cyt b_{562} have been investigated, showing that heme propionate 7 undergoes a rearrangement when the pH is raised from 4.8 to 7.0.

ACKNOWLEDGMENT

We thank Karoly von Glos for assistance in the preparation of labeled protein.

SUPPORTING INFORMATION AVAILABLE

The ^{1}H chemical shifts, the experimental NOE intensities used for the structure calculations, the list of stereospecific assignments, the pseudocontact shift constraints, the torsion angle constraints, the hydrogen bond constraints, and the diamagnetic chemical shifts calculated from both the X-ray structure and the average solution structure obtained without pseudocontact shift constraints are listed (44 pages). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Kamtekar, S., and Hecht, M. H. (1995) *FASEB J. 9*, 1013−1022.
2. Schafmeister, C. E., LaPorte, S. L., Miercke, L. J., and Stroud, R. M. (1997) *Nat. Struct. Biol. 4*, 1039−1046.
3. Rojas, N. R. L., Kamtekar, S., Simons, C. T., McLean, J. E., Vogel, K. M., Spiro, T. G., Farid, R. S., and Hecht, M. H. (1997) *Protein Sci. 6*, 2512−2524.
4. Rau, H. K., DeJonge, N., and Haehnel, W. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 11526−11531.
5. Lawson, D. M., Artymiuk, P. J., Yendall, S. J., Smith, J. M., Livingstone, J. C., Treffry, A., Luzzago, A., Levi, S., Arosio, P., Cesareni, G., Thomas, C. D., Shaw, W. V., and Harrison, P. M. (1991) *Nature 349*, 541−544.
6. Finzel, B. C., Weber, P. C., Hardman, K. D., and Salemme, F. R. (1985) *J. Mol. Biol. 186*, 627−643.
7. Holmes, M. A., Trong, I. L., Turley, S., Sieker, L. C., and Stenkamp, R. E. (1991) *J. Mol. Biol. 218*, 583−593.
8. Itagaki, E., and Hager, L. P. (1966) *J. Biol. Chem. 241*, 3687−3695.
9. Robinson, C. R., Liu, Y., Thomson, A. J., Sturtevant, J. M., and Sligar, S. G. (1997) *Biochemistry 36*, 16141−16146.
10. Fisher, M. T. (1991) *Biochemistry 20*, 10012−10018.
11. Robinson, C. R., Liu, Y., O'Brien, R., Sligar, S. G., and Sturtevant, J. M. (1998) *Protein Sci. 7*, 961−965.
12. Wittung-Stafshede, P., Gray, H. B., and Winkler, J. R. (1997) *J. Am. Chem. Soc. 119*, 9562−9563.
13. Hamada, K., Bethge, P. H., and Mathews, F. S. (1995) *J. Mol. Biol. 247*, 947−962.
14. Feng, Y. Q., Sligar, S. G., and Wand, A. J. (1994) *Nat. Struct. Biol. 1*, 30−35.
15. Barker, P. D., Nerou, E. P., Freund, S. M. V., and Fearnley, I. M. (1995) *Biochemistry 34*, 15191−15203.
16. Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., and Bax, A. (1989) *J. Am. Chem. Soc. 111*, 1515−1517.
17. Kay, L. E., Marion, D., and Bax, A. (1989) *J. Magn. Reson. 84*, 72−84.
18. Macura, S., Wüthrich, K., and Ernst, R. R. (1982) *J. Magn. Reson. 47*, 351−357.
19. Marion, D., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967−974.
20. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson. 63*, 207−213.
21. Griesinger, C., Otting, G., Wüthrich, K., and Ernst, R. R. (1988) *J. Am. Chem. Soc. 110*, 7870−7872.
22. Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479−485.
23. Banci, L., Bertini, I., Luchinat, C., Piccioli, M., Scozzafava, A., and Turano, P. (1989) *Inorg. Chem. 28*, 4650−4656.

24. Inubushi, T., and Becker, E. D. (1983) *J. Magn. Reson. 51*, 128−133.
25. Rigby, S. E. J., Burch, A. M., and Moore, G. R. (1991) *Magn. Reson. Chem. 29*, 1036−1039.
26. Vuister, G. W., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 7772−7777.
27. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−666.
28. Eccles, C., Güntert, P., Billeter, M., and Wüthrich, K. (1991) *J. Biomol. NMR 1*, 111−130.
29. Güntert, P., Braun, W., and Wüthrich, K. (1991) *J. Mol. Biol. 217*, 517−530.
30. Banci, L., Bertini, I., Cremonini, M. A., Gori Savellini, G., Luchinat, C., Wüthrich, K., and Guentert, P. (1998) *J. Biomol. NMR 12*, 553−557.
31. Bertini, I., and Luchinat, C. (1986) in *NMR of paramagnetic molecules in biological systems*, Benjamin/Cummings, Menlo Park, CA.
32. Bertini, I., and Turano, P. (1994) in *NMR of paramagnetic macromolecules. NATO ASI Series* (La Mar, G. N., Ed.) Kluwer Academic, Dordrecht.
33. Bertini, I., and Luchinat, C. (1996) in *NMR of paramagnetic substances, Coord. Chem. Rev. 150*, Elsevier, Amsterdam.
34. Ösapay, K., and Case, D. A. (1991) *J. Am. Chem. Soc. 113*, 9436−9444.
35. Merutka, G., Dyson, H. J., and Wright, P. E. (1995) *J. Biomol. NMR 5*, 14−24.
36. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) *J. Mol. Biol. 273*, 283−298.
37. Banci, L., Bertini, I., Bren, K. L., Cremonini, M. A., Gray, H. B., Luchinat, C., and Turano, P. (1996) *JBIC 1*, 117−126.
38. Banci, L., Bertini, I., Gori Savellini, G., Romagnoli, A., Turano, P., Cremonini, M. A., Luchinat, C., and Gray, H. B. (1997) *Proteins: Struct., Funct., Genet. 29*, 68−76.
39. Arnesano, F., Banci, L., Bertini, I., and Felli, I. C. (1998) *Biochemistry 37*, 173−184.
40. Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., Ferguson, D. M., Seibel, G. L., Singh, U. C., Weiner, P. K., and Kollman, P. A. (1997) in *AMBER 5.0*, University of California, San Francisco.
41. Banci, L., Bertini, I., Gray, H. B., Luchinat, C., Reddig, T., Rosato, A., and Turano, P. (1997) *Biochemistry 36*, 9867−9877.
42. Borgias, B., Thomas, P. D., and James, T. L. (1989) in *COmplete Relaxation Matrix Analysis (CORMA)*, University of California, San Francisco, CA.
43. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
44. Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR 8*, 477−486.
45. Moore, G. R., Williams, R. J., Peterson, J., Thomson, A. J., and Mathews, A. J. (1985) *Biochim. Biophys. Acta 829*, 83−96.
46. Wu, J. Z., La Mar, G. N., Yu, L. P., Lee, K. B., Walker, F. A., Chiu, M. L., and Sligar, S. G. (1991) *Biochemistry 30*, 2156−2165.
47. Feng, Y., Wand, A. J., and Sligar, S. G. (1991) *Biochemistry 30*, 7711−7717.
48. Wüthrich, K. (1986) in *NMR of Proteins and Nucleic Acids*, Wiley, New York.
49. Lederer, F., Glatigny, A., Bethge, P. H., Bellamy, H. D., and Mathews, F. S. (1981) *J. Mol. Biol. 148*, 427.
50. Kurland, R. J., and McGarvey, B. R. (1970) *J. Magn. Reson. 2*, 286−301.
51. Bertini, I., Luchinat, C., and Rosato, A. (1996) *Prog. Biophys. Mol. Biol. 66*, 43−80.
52. Banci, L., Bertini, I., Luchinat, C., and Turano, P. (2000) *Handbook of Porphyrins* (in press).
53. Shulman, R. G., Glarum, S. H., and Karplus, M. (1971) *J. Mol. Biol. 57*, 93−115.
54. Turner, D. L. (1995) *Eur. J. Biochem. 227*, 829−837.
55. Banci, L., Rosato, A., and Turano, P. (1996) *JBIC 1*, 364−367.
56. Shokhirev, N. V., and Walker, F. A. (1998) *J. Am. Chem. Soc. 120*, 981−990
57. Shokhirev, N. V., and Walker, F. A. (1998) *JBIC* (in press).
58. Banci, L., Bertini, I., Bren, K. L., Gray, H. B., Sompornpisut, P., and Turano, P. (1997) *Biochemistry 36*, 8992−9001.
59. Wand, A. J., Di Stefano, D. L., Feng, Y. Q., Roder, H., and Englander, S. W. (1989) *Biochemistry 28*, 186−194.
60. Gao, Y., Boyd, J., Williams, R. J. P., and Pielak, G. J. (1990) *Biochemistry 29*, 6994−7003.
61. Barker, P. D., Butler, J. L., de Oliveira, P., Hill, H. A. O., and Hunt, N. I. (1996) *Inorg. Chim. Acta 250*, 71−77.
62. Banci, L., Bertini, I., Turano, P., and Vicens Oliver, M. (1992) *Eur. J. Biochem. 204*, 107−112.
63. Bertini, I., Gori, G., Luchinat, C., and Vila, A. J. (1993) *Biochemistry 32*, 776−783.
64. Jackson, J. T., La Mar, G. N., and Bartsch, R. G. (1983) *J. Biol. Chem. 258*, 1799−1805.
65. La Mar, G. N., Jackson, J. T., Dugad, L. B., Cusanovich, M. A., and Bartsch, R. G. (1990) *J. Biol. Chem. 265*, 16173−16180.
66. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.

BI982785F