Protein Science (1995), *4*:296–305. Cambridge University Press. Printed in the USA.

Protein structure refinement based on paramagnetic NMR shifts: Applications to wild-type and mutant forms of cytochrome *c*

MIRIAM GOCHIN[1,3] AND HEINRICH RODER[1,2]

[1] Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111
[2] Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania 19104

(RECEIVED October 18, 1994; ACCEPTED December 2, 1994)

Abstract

A new approach to NMR solution structure refinement is introduced that uses paramagnetic effects on nuclear chemical shifts as constraints in energy minimization or molecular dynamics calculations. Chemical shift differences between oxidized and reduced forms of horse cytochrome *c* for more than 300 protons were used as constraints to refine the structure of the wild-type protein in solution and to define the structural changes induced by a Leu 94 to Val mutation. A single round of constrained minimization, using the crystal structure as the starting point, converged to a low-energy structure with an RMS deviation between calculated and observed pseudocontact shifts of 0.045 ppm, 7.5-fold lower than the starting structure. At the same time, the procedure provided stereospecific assignments for more than 45 pairs of methylene protons and methyl groups. Structural changes caused by the mutation were determined to a precision of better than 0.3 Å. Structure determination based on dipolar paramagnetic (pseudocontact) shifts is applicable to molecules containing anisotropic paramagnetic centers with short electronic relaxation times, including numerous naturally occurring metalloproteins, as well as proteins or nucleic acids to which a paramagnetic metal ion or ligand may be attached. The long range of paramagnetic shift effects (up to 20 Å from the iron in the case of cytochrome *c*) provides global structural constraints, which, in conjunction with conventional NMR distance and dihedral angle constraints, will enhance the precision of NMR solution structure determination.

Keywords: energy minimization; hyperfine shift; metalloproteins; molecular dynamics; pseudocontact shift; solution structure

It has long been known that the effect of unpaired electrons on NMR spectral parameters such as chemical shifts and relaxation times is a valuable source of information on molecular structure (e.g., Mildvan & Cohn, 1970; Barry et al., 1971; Wüthrich, 1976; Jardetzky & Robert, 1981; Lee & Sykes, 1983; Geraldes, 1993; La Mar & de Ropp, 1993). The idea of using chemical shifts to determine structure is attractive because of their relative ease and accuracy of measurement. The theory of dipolar paramagnetic (pseudocontact) shifts is well understood (Kurland & McGarvey, 1970; Horrocks & Greenberg, 1973), and their structural interpretation is more straightforward than for diamagnetic shifts (e.g., Ösapay & Case, 1991; de Dios et al., 1993). Because of their relatively long-distance range, paramagnetic effects can provide global structural constraints that are especially valuable in conjunction with the nuclear Overhauser effect, which provides abundant short-range distance information. Nevertheless, there has been only limited use of paramagnetic effects for quantitative structure determination.

Electron–nuclear interactions involving a paramagnetic center with rapidly relaxing electron spins in an anisotropic ligand field give rise to large shifts in NMR frequencies. These paramagnetic shifts can be divided into two components, a scalar (through-bond) contact shift, which propagates no further than four or five bonds from the metal site, and a dipolar (through-space) pseudocontact shift. Using a polar coordinate system defined by the principal axis of the g-tensor, the pseudocontact shift ν_{pc} experienced by a nucleus at position (r, θ, ϕ) can be ex-

Reprint requests to: Heinrich Roder, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, Pennsylvania 19111; e-mail: h_roder@fccc.edu.

[3] Present address: Microbiology Department, School of Dentistry, University of the Pacific, San Francisco, California 94115.

Abbreviations: DQF-COSY, double quantum-filtered J-correlated spectroscopy; TOCSY, total correlation spectroscopy; TPPI, time-proportional phase.

pressed as follows (Kurland & McGarvey, 1970; Horrocks & Greenberg, 1973):

$$\nu_{pc} = \frac{\beta^2 S(S+1)}{9kTr^3}\left[g_{ax}(3\cos^2\theta - 1) + \tfrac{3}{2}g_{eq}(\sin^2\theta\cos 2\phi)\right] \tag{1}$$

where β is the Bohr magneton, S is the electron spin quantum number, T is the absolute temperature, and k is Boltzmann's constant. The axial and equatorial components of the anisotropic g-tensor are defined by:

$$g_{ax} = g_z^2 - (g_x^2 + g_y^2)/2; \qquad g_{eq} = g_x^2 - g_y^2. \tag{2}$$

Equation 1 applies to metal complexes with only one thermally populated multiplet with unpaired electron spin effectively localized at the metal ion. More general situations can be described by an analogous equation with two terms containing the principal components of the magnetic susceptibility tensor and identical geometric terms (Kurland & McGarvey, 1970; Bleany, 1972; Williams et al., 1985). Although Equation 1 permits a direct comparison with EPR data, the following analysis remains unchanged if one chooses the more general description in terms of the magnetic susceptibility tensor.

The graphic representation of Equation 1 in Figure 1 illustrates the pronounced angular and distance dependence of pseudocontact shifts that form the basis for their application as a structural tool. Because of the $1/r^3$ distance dependence, the effect is relatively long range, extending as far as 20 Å from the metal site. If the g-tensor and the nuclear coordinates are known, the pseudocontact term can be calculated exactly. However, the inverse problem is complicated by the degeneracy of $\nu_{pc}(r, \theta, \phi)$, as illustrated by the contours in Figure 1.

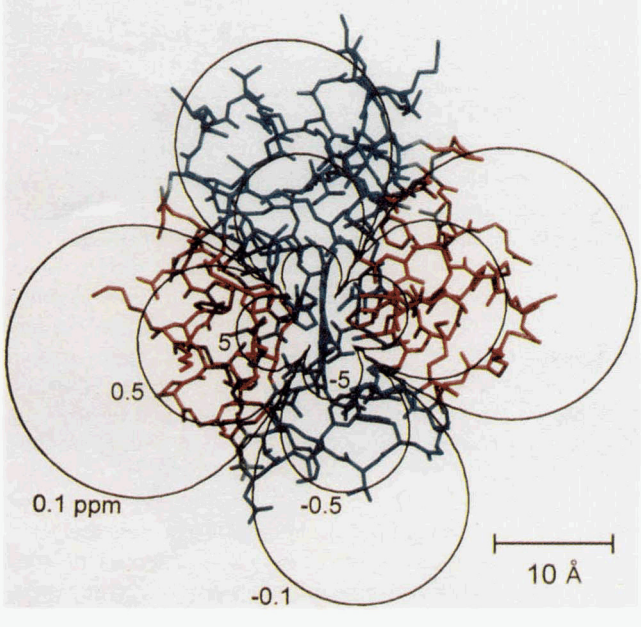

Fig. 1. Contours of constant pseudocontact shift (±5.0, ±0.5, ±0.1 ppm) calculated according to Equation 1 in the principal axis system of the heme, superimposed onto the structure of oxidized horse cytochrome *c* (Bushnell et al., 1990; PDB file 1hrc.pdb). Regions of positive shift are shown in red, negative in blue, and zero in gray. The heme *x*-axis is vertical and the *z*-axis is horizontal in the plane of the page.

Williams et al. (1985) used Equation 1 to calculate pseudocontact shifts for tuna cytochrome *c* from crystal structure coordinates and found that small adjustments in EPR g-tensor parameters resulted in a significant improvement between predicted and observed shifts. Feng et al. (1990) applied an extended method, which included quantitative error analysis, to investigate structural differences between oxidized (Fe^{3+}) and reduced (Fe^{2+}) forms of horse cytochrome *c*. Their observations were consistent with the crystallographic analysis of tuna (Takano & Dickerson, 1981a, 1981b) and yeast cytochrome *c* (Berghuis & Brayer, 1992) in both oxidation states, which revealed small structural changes mainly in some of the more polar regions of the protein. Pseudocontact shift analysis of yeast iso-1 cytochrome *c* (Gao et al., 1991) and recent ^{13}C NMR results on horse cytochrome *c* (Turner & Williams, 1993) led to similar conclusions. Other proteins investigated by this approach include cytochrome b_5 (Veitch et al., 1990) and cytochrome c_{551} (Timkovich & Cai, 1993). Paramagnetic effects have also been used to obtain resonance assignments in conjunction with magnetization transfer methods to correlate resonances in diamagnetic and paramagnetic forms of the protein (Keller & Wüthrich, 1978; Feng & Roder, 1988; Emerson & La Mar, 1990; Feng et al., 1990, 1991; Veitch et al., 1990; Guiles et al., 1993).

In the case of a heme protein, such as cytochrome *c*, the chemical shift difference due to the change in iron oxidation state can be written as

$$\Delta\nu_{obs} = \nu_{ox} - \nu_{red} = \nu_c + \nu_{pc} + \nu_s, \tag{3}$$

where ν_{ox}, ν_{red} are the chemical shifts measured in oxidized (paramagnetic) and reduced (diamagnetic) forms, respectively; ν_c is a contact shift experienced only by protons of the heme and its direct ligands (Shulman et al., 1971; Wüthrich, 1976; Feng et al., 1990); and ν_s represents any remaining chemical shift differences due to changes in the diamagnetic properties between the reduced and oxidized forms. Possible contributions to ν_s are redox-related changes in structure and heme ring current effects. Previous studies on horse cytochrome *c* (Feng et al., 1990) have demonstrated that ν_s is substantial for main-chain protons of residues 39–43 and 50–60 as well as some side-chain protons in the vicinity of the axial ligands and 60's helix. Pseudocontact shifts calculated from a 1.9-Å crystal structure of oxidized horse cytochrome *c* (Bushnell et al., 1990) showed remarkable agreement with experimental shifts for the remainder of the protein backbone and for many methyl groups, indicating that ν_s is negligible for these regions. These observations are consistent with crystallographic analysis of two homologous proteins in both reduced and oxidized forms (Takano & Dickerson, 1981a, 1981b; Berghuis & Brayer, 1992). Another possible contribution to $\Delta\nu_{obs}$ can arise if the crystal structure used to calculate ν_{pc} differs from the actual structure in solution. We found that this uncertainty can be largely eliminated by using the paramagnetic shifts to calculate a refined solution structure, as demonstrated below for wild-type cytochrome *c* and the L94V variant.

Earlier studies relied on crystal structure coordinates for the calculation of pseudocontact shifts and were unable to account for any structural differences between crystal and solution. We show here that this limitation can be overcome by incorporating the measured paramagnetic shifts as constraints in a

molecular dynamics program for the structure refinement of metalloproteins in aqueous solution. Using horse cytochrome *c* as a test case, we found that a simple constrained energy minimization protocol was sufficient to obtain a low-energy structure consistent with more than 300 experimental constraints. In the case of a cytochrome *c* variant, the structure changes associated with a Leu to Val substitution at position 94 (L94V) were defined to within 0.3 Å, illustrating the potential of paramagnetic shifts as a quantitative structural tool.

Results

Refinement of the g-tensor

The magnitude and orientation of the g-tensor describing the electronic properties of the heme iron were optimized against a set of reference protons according to published procedures (Williams et al., 1985; Feng et al., 1990). Coordinate transformation from the heme coordinate system to the g-tensor principal axis system (Fig. 1) is effected by rotations around Euler angles (α, β, γ). The five independent variables, g_{ax}, g_{eq}, α, β, γ, were optimized by calculating the best fit of ν_{pc} (Equation 1) to $\Delta\nu_{obs}$ (Equation 3) for a set of reference protons, using a nonlinear least-squares program. Transformation from crystal coordinates to the heme principal axis system (cf. Fig. 1 in Feng et al., 1990) was obtained by using the center of mass of the 4 heme pyrrole nitrogen, Na, Nb, Nc, Nd as the origin, the unit vector along Nb–Nd as the *x*-axis, and the *y*-axis perpendicular to *x* in the plane defined by the vectors Nb–Nd and Na–Nc (atom labels as defined in the Brookhaven Protein Data Bank).

Pseudocontact shift constraints

The precision with which the position of a proton can be defined by pseudocontact shifts varies widely, depending on its distance and relative orientation with respect to the heme (Fig. 1). A quantitative error function can be obtained by calculating the maximum gradient of the pseudocontact shift tensor at a given position:

$$\frac{\delta}{\delta \mathrm{d}}\nu_{pc} = \left|\frac{\delta}{\delta r}\nu_{pc}\right| + \frac{1}{r}\left|\frac{\delta}{\delta\theta}\nu_{pc}\right| + \frac{1}{r\sin\theta}\left|\frac{\delta}{\delta\phi}\nu_{pc}\right| \qquad (4)$$

where δd is the displacement along the steepest gradient (Feng et al., 1990). To permit quantitative structure calculations, pseudocontact shift constraints were included as part of the energy function in the program XPLOR (Brünger, 1992). Thus, in addition to the conventional empirical energy functions defining bond lengths and angles, dihedral angles, and van der Waals interactions (electrostatic terms were omitted), we defined an explicit "energy" term for the pseudocontact shift:

$$E_{pc} = \Sigma_i f_i\left(\min\left[\,\left|\,|\Delta\nu^i_{obs} - \nu^i_{pc}| - 0.02\right|, 0\right]\right)^2, \qquad (5)$$

which allows for an error of 0.02 ppm in measuring the shift difference $\Delta\nu^i_{obs}$. The index *i* runs over all measured shifts and f_i is a force constant. Values of 300 kcal mol^{-1} ppm^{-2} for amide protons and 400 kcal mol^{-1} ppm^{-2} for all other protons were empirically found to be adequate. Energy minimization using XPLOR was carried out on a Silicon Graphics Iris 4D35, using a heme-based coordinate system with fixed pyrrole nitrogens. The contribution of the pseudocontact shift energy term to the gradient was calculated from partial derivatives of Equation 5 in the *x*, *y*, and *z* directions.

Pseudocontact shift measurements

Proton resonance assignments have previously been reported for wild-type cytochrome *c* in both oxidation states (Feng et al., 1989; Wand et al., 1989). In order to obtain a consistent set of experimental pseudocontact shifts, ν_{obs}, we recorded TOCSY spectra at 600 MHz for reduced and oxidized cytochrome *c* under matching conditions (20 °C, pH 5.7). Assignment of the L94V variant proved to be a relatively simple extension, using standard homonuclear 2D NMR methods (Wüthrich, 1986). Spectral perturbations caused by the mutation were relatively small and limited to residues in close proximity to the site of mutation. Most assignments were made by matching the patterns of cross peaks between amide and side-chain protons in TOCSY spectra of wild-type and mutant cytochrome *c* recorded under identical conditions. DQF-COSY spectra were useful for identifying some aliphatic spin systems.

Solution structure refinement by energy minimization with pseudocontact shift constraints

The α-protons of residues 7–13, 15, 20, 22, 25, 26, 30, 36, 61, 62, 64–70, 72, 73, 75, 76, 79, 81, and 85–101 were selected as reference protons for optimizing the electronic g-tensor. These protons were chosen on the basis of previous NMR (Feng et al., 1990) and crystallographic (Takano & Dickerson, 1981a, 1981b; Berghuis & Brayer, 1992) results, which indicated that these main-chain segments do not undergo significant redox-dependent structure changes. The optimized g-tensor parameters thus obtained for wild-type horse cytochrome *c* ($g_{ax} = 5.25$, $g_{eq} = -1.83$, $\alpha = 102.93$, $\beta = 12.74$, and $\gamma = 247.52$) and for the L94V variant ($g_{ax} = 5.19$, $g_{eq} = -2.09$, $\alpha = 104.48$, $\beta = 12.69$, and $\gamma = 245.64$) are very similar to those reported previously (Feng et al., 1990).

As a starting point for the solution structure refinement of wild-type cytochrome *c*, we used the crystal coordinates for oxidized horse cytochrome *c* reported by Bushnell et al. (1990). Alternatively, one could use for this purpose an NOE-based NMR solution structure, such as the recently reported structure for reduced horse cytochrome *c* (Qi et al., 1994). Initially, we applied 273 constraints corresponding to experimentally well-defined pseudocontact shifts ($\nu_c = 0$, $\nu_s \sim 0$). Protons in regions known to undergo oxidation-state-dependent structure changes (Takano & Dickerson, 1981a, 1981b; Feng et al., 1990) and those of the heme ligands (Hα of residues 14, 17, 39–42, 48–60; NH of 39–42, 49–61; Hβ of 4, 16–18, 39, 47, 50, 57–68, 80; Hγ of 57–58, 80, 85; Hδ of 57, 80, Hϵ of Met 80, and the ring protons of His 18) were not used in this calculation. By matching predicted with observed paramagnetic shifts in several cycles of constrained minimization, the pseudocontact shift calculation eventually resulted in stereospecific assignments for 45 pairs of glycine α protons, side-chain methylene and methyl groups, and aromatic ring protons. In the final minimization, a total of 309 constraints was used (cf. Table 1). For the L94V variant, an approximate starting structure was obtained by simply replacing the residue and side-chain atom names for residue 94 and de-

Table 1. *Effect of energy minimization in the presence and absence of pseudocontact shift constraints on energies and chemical shift differences for wild-type and L94V cytochrome c*

	Initial structure			Final structure			No. of constraints
	E(total)	E(shift)	Deviation[a]	E(total)	E(shift)	Deviation[a]	
Unconstrained							
Wild type (crystal)	4,956	0	0.343	154	0	0.257	0
L94V (model)	5,496	0	0.340	154	0	0.257	0
Constrained							
Wild type (solution)	18,847	14,043	0.343	371	64	0.045	309
L94V (solution)	20,695	15,198	0.340	373	55	0.045	294

[a] RMS deviation between observed and calculated pseudocontact shift, $[\Sigma(\Delta\nu_{obs} - \nu_{pc})^2/N]^{1/2}$, where N is the total number of protons (cf. Equations 1 and 3).

leting the γCH_2 group of Leu 94. Initial trial runs made use of 257 shift constraints, and the final minimization was based on 294 constraints (Table 1).

Because of the nonlinear form of dipolar chemical shifts (Equation 1), the pseudocontact shift energy surfaces vary in shape, depending on the location of a proton in the g-tensor field. A representative example is shown in Figure 2A. Interdependency of the x, y, and z coordinates helps to define atomic positions with a high degree of precision. Figure 2B shows the effect of pseudocontact shift constraints on the progress of energy minimization. Each run consisted of 500 cycles of minimization, which required ~2,000 s CPU time on a Silicon Graphics Iris 4D35 with negligible increase in time due to the additional energy constraint. The energy reached a nearly constant plateau after about 300 cycles in constrained minimization runs, compared to about 150 cycles in unconstrained runs. The reduction in pseudocontact shift "energy" during minimization is accompanied by a parallel decrease in other energy terms to final values near those observed in an unconstrained run (Table 1), indicating that satisfaction of pseudocontact constraints does not result in a large increase in other energy terms. For the wild-type protein, the decrease in energy during the course of minimization was monotonous. The L94V mutant, however, showed several sharp spikes in the pseudocontact term and the total energy (Fig. 3), corresponding to a temporary motion along a positive energy gradient, which probably reflects side-chain rearrangements in the vicinity of the mutation. In each case, the minimization routine was able to recover after these fluctuations and continued toward a lower energy minimum. In other cases, it may be necessary to use molecular dynamics or simulated annealing to move the system out of local minima. At a late stage of the process, the g-tensor was optimized again for the refined structure; another round of minimization produced negligible structure changes. The fact that very similar g-tensors were found for wild type and L94V further confirm that the refined structures are not very sensitive to the exact choice of g-tensor parameters.

The final structures were found to conform well with standard stereochemical criteria. No bond length violations and only two bond angle violations (−19° and −16°, both in the Trp 59 indole ring) were found. The latter appear to be caused by a computational artifact because they occurred irrespective of whether or not shift constraints were applied to the protons of Trp 59. Ramachandran plots before and after minimization showed no unusual backbone dihedral angles.

Constrained energy minimization resulted in dramatically improved agreement between calculated and measured pseudocontact shifts, as documented in Table 1 and Figure 3. Both for wild-type and L94V cytochrome c, the RMS deviation between $\Delta\nu_{obs}$ and ν_{pc} decreased by a factor 7.6 upon constrained minimization, reaching a final value of 0.045 ppm (Table 1). The most impressive reduction in shift discrepancies on minimization is observed for protons in the vicinity of the heme group, even though most of these were not included in the initial list of constraints. The ability to accurately predict pseudocontact shifts for protons in the immediate vicinity of the heme group indicates that the simple dipolar description (Equation 1) is still adequate, even at these short distances from the paramagnetic center (>4.5 Å), as long as there is no contact shift contribution. In this case, a small change in coordinates can cause a large change in pseudocontact shift. For example, for the NH of Cys 17, which forms a thioether linkage to the heme, the shift discrepancy $\nu_{pc} - \Delta\nu_{obs}$ was reduced from −0.59 to 0.01 on refinement. In the case of the Gly 29 α protons, the initially calculated shifts (−5.72/−1.50 ppm) did not permit a clear distinction between two alternative combinations of reduced and oxidized shifts (−8.08/−0.86, or −4.74/−4.19 ppm), whereas the shifts predicted by the structure after the first stage of refinement (−5.44/−4.26 ppm) agreed more closely with the second option above and thus provided stereospecific assignments for the Gly 29 α protons. For other residues (Phe 10, Glu 69, and Tyr 97), the predicted shifts could be consolidated with published assignments only after rearranging the assignments for some of the protons within these side chains. The following corrections were applied to published (Wand et al., 1989) assignments for the reduced form: Phe 10 H2 (6.98 ppm), H6 (7.18 ppm), H3 (7.17 ppm), H5 (6.12 ppm); E69 β1 (2.00 ppm), β2 (2.11 ppm); and for the oxidized form: Tyr 97 H2 (7.16 ppm), H6 (6.41 ppm), H3 (6.86 ppm), H5 (5.73 ppm). Alternative assignment options for several stereospecific pairs were checked, especially for residues near the site of mutation.

For main-chain protons outside the immediate vicinity of the heme group, the effect of minimization on pseudocontact shifts was small (Fig. 3), indicating that, in terms of backbone con-

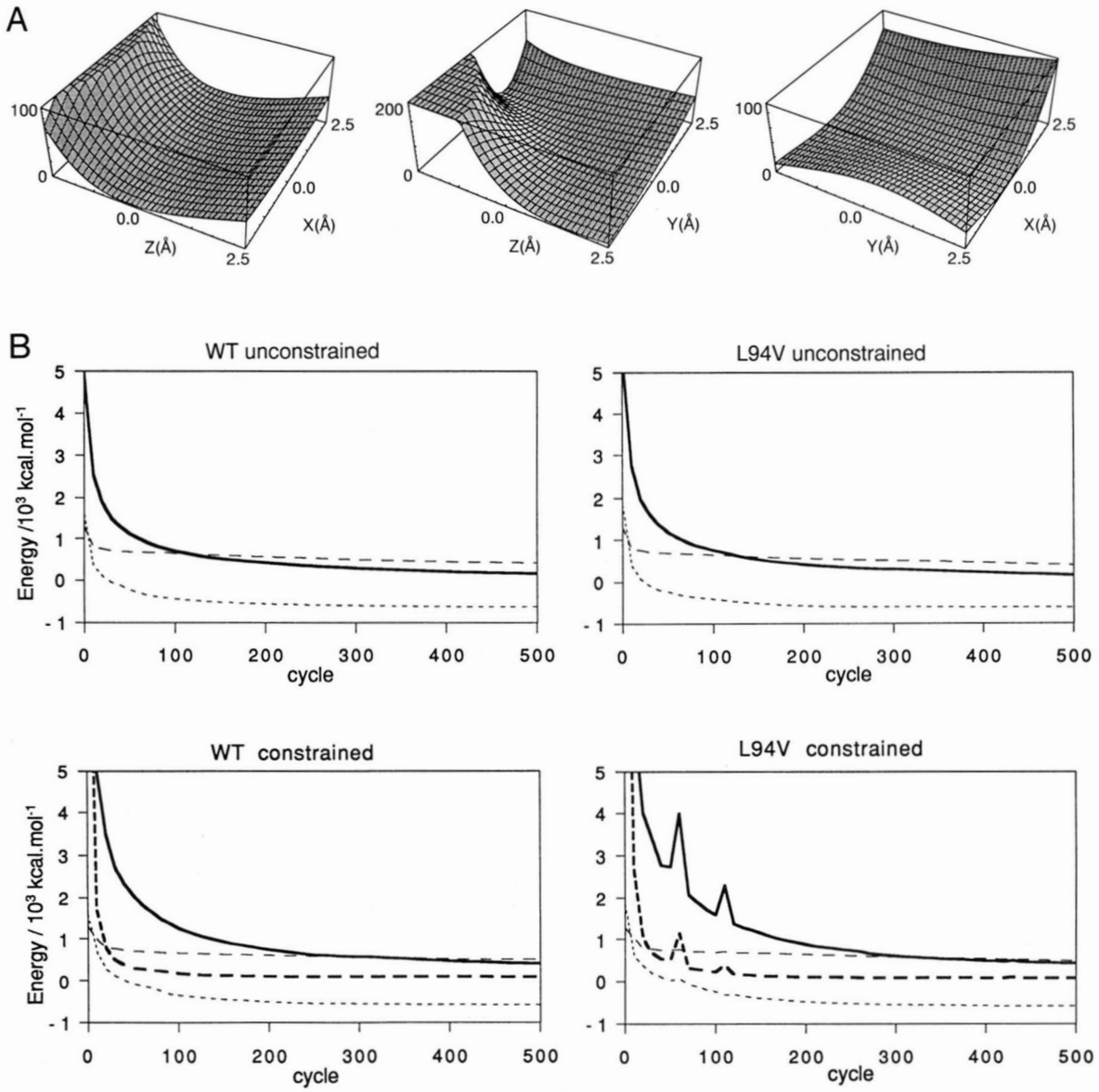

Fig. 2. A: Surface of the pseudocontact energy function for one of the Leu 94 Cδ methyl groups in wild-type cytochrome *c*. Three views of the surface are shown spanning ±2.5 Å around the optimal position in the directions indicated. The *z*-axis is the value of the pseudocontact shift energy E_{pc} in kcal/mol. **B:** Progress curves of XPLOR energy minimization runs in the absence (top) and presence (bottom) of pseudocontact shift constraints carried out separately for wild-type horse cytochrome *c* (left) and the L94V variant (right). Curves shown represent: ———, total energy; – – – –, van der Waals energy; — — — —, dihedral angle energy; and — — — —, pseudocontact shift energy. These final minimization runs included 147 main-chain and 162 side-chain constraints for the wild type, and 136 main-chain and 158 side-chain constraints for L94V.

formation, the crystal structure is an accurate representation of the structure in solution. However, for some side-chain protons, pseudocontact shifts calculated from the crystal coordinates did not match experimental shifts very well, and minor rearrangement of these side chains during constrained minimization resulted in a significantly improved fit.

Figure 4 illustrates the maximum precision with which proton positions can be defined on the basis of pseudocontact shift constraints, calculated according to Equation 4, assuming that the experimental chemical shifts are reproduced by the model within 0.045 ppm, the final RMS chemical shift deviation obtained after constrained minimization (Table 1). In this case, the positional uncertainty for many protons can be as small as ±0.1 Å, and about 90% of the residues are defined to within 0.5 Å, whereas only a few residues in peripheral regions of the structure are poorly constrained (within 2 Å). This surprisingly high degree of precision in average positions for some protons is due to the steep gradients of the pseudocontact shift tensor in the vicinity of the heme. A comparison of the different structures in terms of RMS differences is given in Table 2. Energy minimization of the X-ray coordinates (Bushnell et al., 1990), using the standard energy terms of XPLOR, results in small

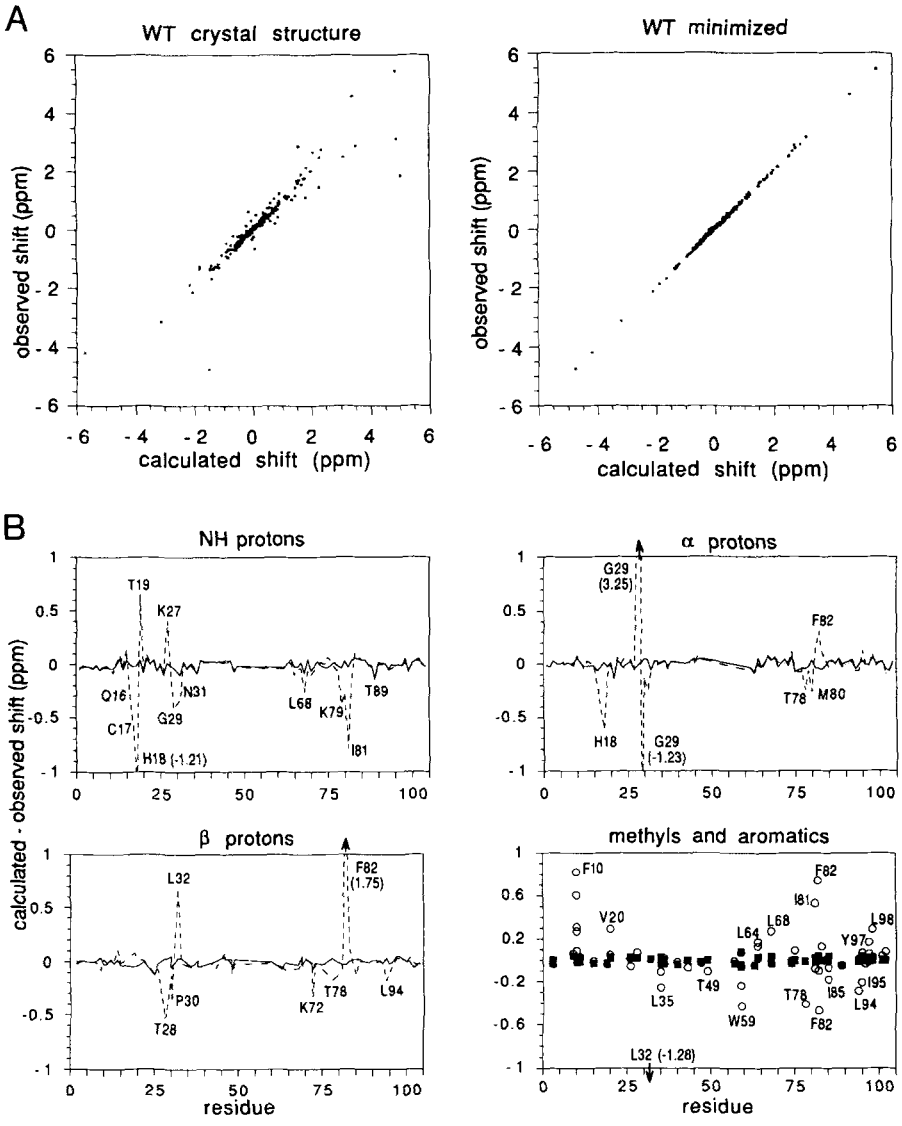

Fig. 3. A: Plot of experimental versus calculated pseudocontact shifts for all assigned protons in wild-type cytochrome *c* before and after constrained minimization. **B:** Plots of the difference between observed and calculated chemical shifts before (– – –, ○) and after (——, ■) constrained minimization, for NH, α, β, methyl, and aromatic protons. For methyls and rapidly rotating aromatic rings, shift values were computed by averaging the calculated values for all equivalent protons. A scale of +1 to −1 ppm was used; arrows indicate larger shift differences found for the unminimized structure. Residues 39–42 and 50–60 were omitted.

structural rearrangements (0.60 Å RMS difference), suggesting relaxation of the structure due to some strain in the crystallographic model. However, these changes could also reflect imperfections in the empirical energy terms, or the fact that these calculations were performed without inclusion of bulk water other than crystallographically defined water molecules. On the other hand, we found that the structure obtained after constrained minimization is very similar to the relaxed structure (0.31 Å RMS difference), indicating that energy minimization does not introduce serious artifacts.

Structural changes associated with the L94V mutation

Leu 94 is located in the C-terminal α-helix (residues 87–102) of cytochrome *c* at the interface with the N-terminal α-helix (residues 3–14). In the wild-type structure, the side chain of Leu 94 is involved in van der Waals contacts with main-chain atoms of Gly 6 and Phe 10 and the side chain of Ile 9 (Bushnell et al., 1990). A variant with a valine substitution at position 94 was prepared in collaboration with Fred Sherman (University of Rochester) as part of an ongoing mutagenesis study on the contribution of these helix-helix packing interactions to protein stability (cf. Fredericks & Pielak, 1993) and their role in folding. Our interest in this region of cytochrome *c* is based on earlier studies on the folding mechanism of cytochrome *c* by pulsed hydrogen exchange and NMR methods, which showed that association of the N- and C-terminal helices is an essential event at an early stage of the folding process (Roder et al., 1988; Elöve & Roder, 1991; Wu et al., 1993; Roder & Elöve, 1994).

The structural changes associated with the Leu to Val substitution at position 94 are illustrated in Figure 5. Fortunately, residue 94 lies in a highly sensitive region of the g-tensor field, approximately 11 Å from the heme iron, so that the structure changes associated with this mutation can be very well defined by using pseudocontact shift constraints. For many residues near the site of mutation, the positional uncertainty is less than 0.3 Å, and some protons, including the methyl groups of both Leu 94 and Val 94, can be placed with a precision better than 0.1 Å (Fig. 5A). Val 94 Cβ is within 5 Å of parts of residues Leu 68 (Cδ1, Cδ2), Arg 91 (Cγ1), Leu 98 (Cδ1, Cδ2), Ile 95 (N),

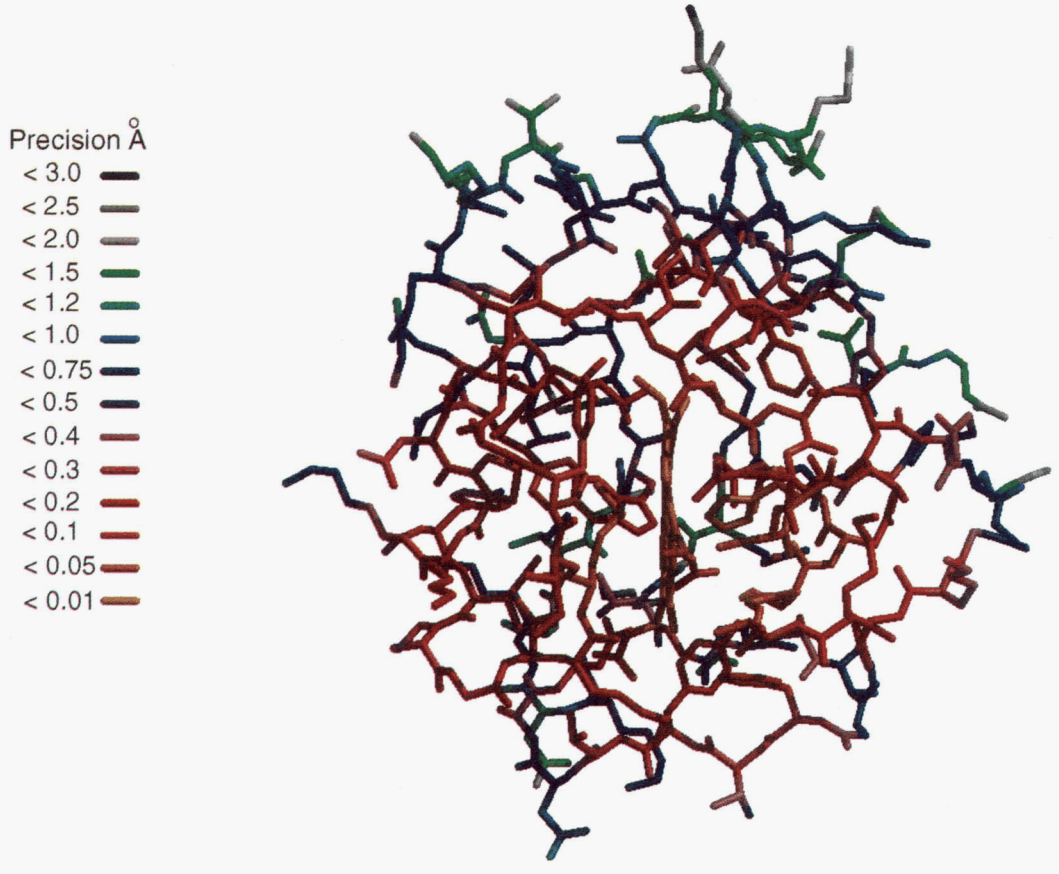

Fig. 4. Illustration of the theoretical precision to which protons of wild-type cytochrome *c* can be determined by pseudocontact shift constraints. The precision is calculated according to Equation 4, assuming that any residual chemical shift differences, $\Delta\nu_{obs} - \nu_{pc}$, are within ±0.045 ppm. Precision is color coded as indicated by the key to the left.

Gly 6 (N, Cα, C), and Ile 9 (Cδ1, Cγ). Overall, the mutation induces only minor changes in backbone conformation (0.14 Å RMS difference). The largest displacement is found in the C-terminal helix, where the main chain between residues 93 and 95 is displaced by 0.5 ± 0.1 Å toward the N-terminal helix, filling some of the space created by the loss of a methylene group on mutation. Val 94 Cβ is displaced by 0.74 ± 0.10 Å relative to the original leucine side chain, changing the orientation of the new valine side chain. The Val 94 γ1 methyl points between the original Leu 94 γ and δ1 methyl groups and the Val 94 γ2 methyl points out toward the δ2 methyl of the original leucine. Several adjacent groups move toward Val 94, including the Leu 68 δ1 (by 0.4 ± 0.04 Å), Asp 93 (0.4 ± 0.4 Å), Ile 85 (0.30 ± 0.07 Å), and Phe 10 (by 0.25 ± 0.02 Å), partially compensating for the decrease in side-chain volume caused by the leucine to valine substitution. Nevertheless, a small cavity remains between the Val 94 methyl groups and the ring of Phe 10. The Leu 98 side chain moves away from the site of mutation by 0.3 ± 0.06 Å at the Cγ to avoid contact with Val 94 Cγ1. Structural changes for more distant residues are minimal, with one

Table 2. *Comparison of cytochrome c structures before and after constrained energy minimization*[a]

	Crystal structure	Wild type minimized		L94V minimized	
		Unconstrained	Constrained	Unconstrained	Constrained
Crystal structure	0	0.53	0.48	0.54	0.52
Wild type minimized unconstrained	0.60	0	0.28	0.07	0.26
Wild type minimized constrained	0.58	0.31	0	0.30	0.12
L94 minimized unconstrained	0.60	0.08	0.33	0	0.28
L94V minimized constrained	0.62	0.30	0.14	0.32	0

[a] RMS deviations for heavy atoms of the backbone are shown above the diagonal, and those for all heavy atoms except side chain of residue 94 are shown below the diagonal.

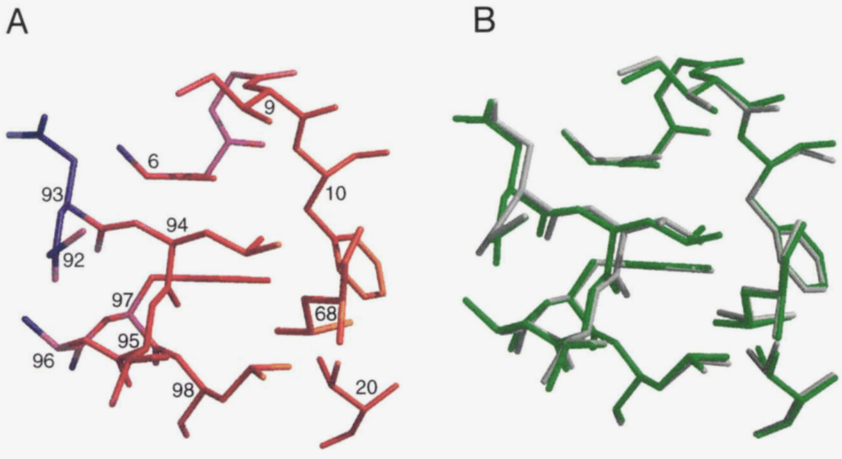

Fig. 5. Structural changes of cytochrome *c* induced by mutation of Leu 94 to Val (L94V). **A:** Color-coded map of the precision to which protons in the region of mutation are determined by the pseudocontact shift tensor as in Figure 4, but assuming an uncertainty of ±0.03 ppm in $\Delta\nu_{obs} - \nu_{pc}$. For the residues shown, the RMS difference between calculated and observed shift is 0.03 ppm for side-chain protons, and 0.045 ppm for main-chain protons. **B:** Comparison of constrained and minimized structures of wild-type (green) and L94V (gray) cytochrome *c* in the vicinity of residue 94.

notable exception. The Val 20 γ1 methyl shifts by 0.54 ± 0.05 Å. Given the 9.6-Å distance of Val 20 from the site of mutation, this must be a secondary effect propagated through some of the intervening residues, such as Phe 10 or Leu 98. The limited structural rearrangements found here are typical of the effect of similar conservative amino acid changes in other proteins, such as T4 lysozyme (Eriksson et al., 1992).

Discussion

Assessment of paramagnetic shifts for quantitative structure determination

The applications presented here demonstrate the power of using paramagnetic effects on proton chemical shifts as structural constraints in the refinement of protein structures in solution. In favorable regions relative to the paramagnetic center, it is possible to determine nuclear positions at an unprecedented level of precision, often less than 0.1 Å. The pseudocontact effect has a much longer range (as far as 20 Å in the case of cytochrome *c*) than the NOE, which is limited to proton–proton distances of less than about 5 Å. Thus, by including pseudocontact shift constraints, it should be possible to avoid the potential bias introduced by the short-range nature of NOE constraints (Wagner et al., 1992; Clore et al., 1993; Nakai et al., 1993), which should lead to NMR structures with enhanced global precision. Although the method requires two sets of resonance assignments, one for the paramagnetic form of the molecule and another for a suitable diamagnetic reference state, this is still more straightforward and far less time consuming than the laborious task of extracting distance constraints from NOE spectra. Furthermore, a simple iterative procedure can be used to confirm assignments and obtain new ones, including stereospecific assignments. Chemical shifts can inherently be measured more accurately than NOE intensities. Moreover, in refining structures on the basis of pseudocontact shifts, one directly fits the experimentally observed parameter, which depends only on the position of one nucleus independent of all others, in contrast to the distances derived from NOE intensities, which are model dependent and highly coupled. Thus, inclusion of pseudocontact shifts enhances the precision of a solution structure and permits a more direct assessment of its accuracy. Energy minimization or molecular dynamics calculations with pseudocontact shift constraints are computationally efficient and robust.

The method is, of course, limited to molecular systems with suitable paramagnetic centers, typically tightly bound metal ions with rapidly relaxing electron spins. Another potential problem is the requirement for a diamagnetic reference state with essentially identical structure. In the case of a redox protein like cytochrome *c*, this means that the method cannot be applied to parts of the protein that undergo significant structural changes upon reduction or oxidation. Fortunately, most regions of the cytochrome *c* structure, including the site of our mutation at the interface between the N- and C-terminal helices, are highly conserved in the two oxidation states.

Outlook

Because of their distinct physical properties, paramagnetic shift constraints and NOE distance constraints are highly complementary, and their combined application will be especially powerful. Pseudocontact shift constraints can easily be incorporated into constrained molecular dynamics and simulated annealing protocols, along with NOE distance bounds, which will make it possible to determine more precise solution structures. Because the method was found to be quite insensitive to the choice of g-tensor parameters, we expect that a low-resolution NOE-based model will be sufficient as a starting structure for an initial estimate of the g-tensor parameters and a first round of pseudocontact shift refinement, followed by several iterations of g-tensor optimization and constrained molecular dynamics. It will also be interesting to explore other nuclei, especially ^{13}C (Turner & Williams, 1993), which will provide many additional constraints and may permit refinement of larger molecules. Moreover, ^{13}C chemical shifts are expected to be less sensitive than proton shifts to redox-dependent structural changes, especially if these involve changes in hydrogen bonding.

Molecules amenable to structural analysis by this method include a large number of metalloproteins containing intrinsic paramagnetic centers with favorable electronic properties, such as low-spin forms of Fe^{3+} and high-spin forms of Co^{2+}. In other cases, it may be possible to replace a diamagnetic metal, such as Mg^{2+}, Ca^{2+}, or Zn^{2+}, with a suitable paramagnetic analogue, such as Co^{2+} or certain lanthanides (e.g., Shelling

et al., 1985; Michael et al., 1992). Replacement of divalent cations by lanthanides is facilitated by their similar atomic radii and the tight binding of the substituted metal ions (Bertini & Luchinat, 1986). The possibility of introducing new metal binding sites into proteins is suggested by several recent studies, including the work of Toma et al. (1991), who used site-directed mutagenesis to transfer a calcium binding loop from thermolysin to a homologous protease, as well as several reports of model proteins designed to bind metals (Regan & Clarke, 1990; Handel et al., 1993) or heme groups (Robertson et al., 1994). Thus, we anticipate that paramagnetic shifts in conjunction with molecular modeling approaches will find more widespread application for the investigation of macromolecular structures in solution.

Materials and methods

NMR sample preparation

Horse cytochrome *c* was purchased from Sigma Chemical Co. (Type VI). A variant of horse cytochrome *c* with a Leu to Val substitution at position 94 (L94V cyt *c*) was generously provided for these preliminary structural studies by Fred Sherman at the University of Rochester Medical Center. The mutant was constructed by standard oligonucleotide-directed mutagenesis techniques (Kunkel et al., 1987), based on a previously described synthetic gene encoding wild-type horse cytochrome *c* (Hickey et al., 1991), and expressed in yeast. Further details on the preparation and biophysical characterization of L94V cyt *c* will be reported elsewhere (G.A. Elöve, W. Colón, P. Wakem, F. Sherman, & H. Roder, in prep.). NMR samples of 1.5 mM wild-type or mutant cytochrome *c* were prepared in 90% H_2O/10% D_2O solutions containing 50 mM potassium phosphate and 50 mM acetic acid d_4 buffer at pH 5.7. To ensure complete oxidation, the samples were treated with small amounts of ferricyanide, which was subsequently removed by gel filtration (Sephadex G25 spinning columns). Following NMR analysis in the oxidized form, the samples were reduced by adding ~2 mM dithionite.

NMR spectroscopy

All proton NMR spectra were recorded at 20 °C on a Bruker AM-600 spectrometer. For DQF-COSY (Shaka & Freeman, 1983) and TOCSY (Braunschweiler & Ernst, 1983) spectra, 64 scans of 2,048 complex data points were acquired in TPPI mode (Marion & Wüthrich, 1983) for 512 increments (5 μs to 23 ms), corresponding to a spectral width of 11,111 Hz in both dimensions. The water resonance was suppressed by selective irradiation during a 1.5-s delay preceding the first pulse. A 65-ms MLEV-17 mixing sequence (Bax & Davies, 1985) was used for TOCSY experiments. The program Felix 2.3 (Biosym Technologies, San Diego, California) was used for NMR data processing on a Silicon Graphics Iris workstation (4D-35). The final data matrices consisted of 2,048 × 2,048 data points (5.4 Hz/point).

Acknowledgments

We are indebted to Fred Sherman and his colleagues for providing the L94V variant of horse cytochrome *c*. We thank Jiang Jiangsheng and Axel Brünger for their help in incorporating additional routines into XPLOR. We are grateful to Paul Laub for his assistance in the general use of XPLOR and Gülnur Elöve for help with NMR and discussions. We thank G.D. Brayer for providing coordinates for the crystal structure of horse cytochrome *c* (Bushnell et al., 1990). Molecular graphics images were produced using the MidasPlus software system from the Computer Graphics Laboratory, University of California, San Francisco (supported by NCRR grant RR-01081). This work was supported by grants from the National Institutes of Health (R01 GM35926 and CA06927) and an appropriation from the Commonwealth of Pennsylvania to the Institute for Cancer Research.

References

Barry CD, North ACT, Glasel JA, Williams RJP, Xavier AV. 1971. Quantitative determination of mononucleotide conformations in solutions using lanthanide ion shift and broadening NMR probes. *Nature 232*: 236–245.

Bax A, Davies DG. 1985. MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *J Magn Reson 65*:355–360.

Berghuis AM, Brayer GD. 1992. Oxidation state-dependent conformational changes in cytochrome *c*. *J Mol Biol 223*:959–976.

Bertini I, Luchinat C. 1986. *NMR of paramagnetic molecules in biological systems*. Menlo Park, California: Benjamin/Cummings Publishing.

Bleany B. 1972. Nuclear magnetic resonance shifts in solution due to lanthanide ions. *J Magn Reson 8*:91–100.

Braunschweiler L, Ernst RR. 1983. Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy. *J Magn Reson 53*: 521–528.

Brünger AT. 1992. *X-PLOR version 3.0*. New Haven, Connecticut: Yale University.

Bushnell G, Louie G, Brayer GD. 1990. High-resolution three-dimensional structure of horse heart cytochrome *c*. *J Mol Biol 214*:585–595.

Clore GM, Robien MA, Gronenborn A. 1993. Exploring the limits of precision and accuracy of protein structures determined by nuclear magnetic resonance spectroscopy. *J Mol Biol 231*:82–102.

de Dios AC, Pearson JG, Oldfield E. 1993. Secondary and tertiary structural effects on protein NMR chemical shifts: An ab initio approach. *Science 260*:1491–1496.

Elöve GA, Roder H. 1991. Structure and stability of cytochrome *c* folding intermediates. *ACS Symp Ser 470*:50–63.

Emerson SD, La Mar GN. 1990. NMR determination of the orientation of the magnetic susceptibility tensor in cyanometmyoglobin: A new probe of steric tilt of bound ligand. *Biochemistry 29*:1556–1566.

Eriksson AE, Baase WA, Zhang XJ, Heinz DW, Blaber M, Baldwin EP, Matthews BW. 1992. Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science 255*:178–183.

Feng Y, Roder H. 1988. Relayed magnetization transfer by isotropic mixing in exchanging systems. *J Magn Reson 78*:597–602.

Feng Y, Roder H, Englander SW, Wand AJ, DiStefano DL. 1989. Proton resonance assignments of horse ferricytochrome *c*. *Biochemistry 28*: 195–203.

Feng Y, Roder H, Englander SW. 1990. Redox-dependent structure changes and hyperfine nuclear magnetic resonance shifts in cytochrome *c*. *Biochemistry 29*:3494–3504.

Feng Y, Wand AJ, Roder H, Englander SW. 1991. Chemical exchange in two dimensions in the 1H NMR assignment of cytochrome *c*. *Biophys J 59*:323.

Fredericks ZL, Pielak GJ. 1993. Exploring the interface between the N- and C-terminal helices of cytochrome *c* by random mutagenesis within the C-terminal helix. *Biochemistry 32*:929–936.

Gao Y, Boyd J, Pielak GJ, Williams RJP. 1991. Comparison of reduced and oxidized yeast iso-1-cytochrome *c* using proton paramagnetic shifts. *Biochemistry 30*:1928–1934.

Geraldes CFGC. 1993. Lanthanide shift reagents. *Methods Enzymol 227*:43–78.

Guiles RD, Basus VJ, Sarma S, Malpure S, Fox KM, Kuntz D, Waskell L. 1993. Novel heteronuclear methods of assignment transfer from a diamagnetic to a paramagnetic protein: Application to rat cytochrome b_5. *Biochemistry 32*:8329–8340.

Handel TM, Williams SA, DeGrado WF. 1993. Metal ion-dependent modulation of the dynamics of a designed protein. *Science 261*:879–885.

Hickey DR, Berghuis AM, Lafond G, Jaeger JA, Cardillo TS, McLendon D, Das G, Sherman F, Brayer GD, McLendon G. 1991. Enhanced thermodynamic stabilities of yeast iso-1-cytochromes *c* with amino acid replacements at positions 52 and 102. *J Biol Chem 266*:11686–11694.

Horrocks WD Jr, Greenberg S. 1973. Evaluation of dipolar nuclear mag-

netic resonance shifts in low-spin hemin systems: Ferricytochrome *c* and metmyoglobin cyanide. *Biochim Biophys Acta 322*:38–44.

Jardetzky O, Roberts GCK. 1981. *NMR in molecular biology*. Orlando: Academic Press.

Keller RM, Wüthrich K. 1978. Evolutionary change of the heme c electronic structure: Ferricytochrome *c*-551 from *Pseudomonas aeruginosa* and horse heart ferricytochrome *c*. *Biochem Biophys Res Commun 83*:1132–1139.

Kunkel TA, Roberts JD, Zakour RA. 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol 154*:367–383.

Kurland RJ, McGarvey BR. 1970. Isotropic NMR shifts in transition metal complexes: The calculation of the fermi contact and pseudocontact terms. *J Magn Reson 2*:286–301.

La Mar GN, de Ropp JS. 1993. NMR methodology for paramagnetic proteins. *Biol Magn Reson 12*:1–111.

Lee L, Sykes BD. 1983. Use of lanthanide-induced nuclear magnetic resonance shifts for determination of protein structure in solution: EF calcium binding site of carp parvalbumin. *Biochemistry 22*:4373–4379.

Marion D, Wüthrich K. 1983. Application of phase sensitive two-dimensional correlated spectroscopy for measurements of ^1H–^1H spin-spin coupling constants in proteins. *Biochem Biophys Res Commun 113*:967–974.

Michael SF, Kilfoil VJ, Schmidt MH, Amann BT, Berg J. 1992. Metal binding and folding properties of a minimalist Cys 2 His 2 zinc finger peptide. *Proc Natl Acad Sci USA 89*:4796–4800.

Mildvan AS, Cohn M. 1970. Aspects of enzyme mechanisms studied by nuclear spin relaxation induced by paramagnetic probes. *Adv Enzymol 33*:1–70.

Nakai T, Kidera A, Nakamura H. 1993. Intrinsic nature of the three-dimensional structure of proteins as determined by distance geometry with good sampling properties. *J Biomol NMR 3*:19–40.

Ösapay K, Case DA. 1991. A new analysis of proton chemical shifts in proteins. *J Am Chem Soc 113*:9436–9444.

Qi PX, DiStefano DL, Wand AJ. 1994. Solution structure of horse heart ferrocytochrome *c* determined by high-resolution NMR and restrained simulated annealing. *Biochemistry 33*:6408–6417.

Regan L, Clarke ND. 1990. A tetrahedral zinc II-binding site introduced into a designed protein. *Biochemistry 29*:10878–10883.

Robertson DE, Fari RS, Moser CC, Urbauer JL, Mulholland SE, Pidikiti R, Lear JD, Wand AJ, DeGrado WF, Dutton PL. 1994. Design and synthesis of multi-haem proteins. *Nature 368*:425–432.

Roder H, Elöve GA. 1994. Early stages of protein folding. In: Pain RH, ed. *Mechanisms of protein folding*. Oxford, UK: Oxford University Press. pp 26–55.

Roder H, Elöve GA, Englander SW. 1988. Structural characterization of folding intermediates in cytochrome by H-exchange labelling and proton NMR. *Nature 35*:700–704.

Shaka AJ, Freeman R. 1983. Simplification of NMR spectra by filtration through multiple-quantum coherence. *J Magn Reson 51*:169–173.

Shelling JG, Hofmann T, Sykes BD. 1985. Proton nuclear magnetic resonance studies of the interaction of the lanthanide ions ytterbium and lutetium with apo- and calcium-saturated porcine intestinal calcium binding protein. *Biochemistry 24*:2332–2338.

Shulman RG, Glarum SH, Karplus M. 1971. Electronic structure of cyanide complexes and hemes and heme proteins. *J Mol Biol 57*:93–115.

Takano T, Dickerson RE. 1981a. Conformational change of cytochrome *c* I ferrocytochrome *c* structure refined at 15 Å resolution. *J Mol Biol 153*:79–94.

Takano T, Dickerson RE. 1981b. Conformational change in cytochrome *c* II ferricytochrome *c* refinement at 18 Å resolution and comparison with the ferrocytochrome structure. *J Mol Biol 153*:95–115.

Timkovich R, Cai M. 1993. Investigation of the structure of oxidized *Pseudomonas aeruginosa* cytochrome *c*-551 by NMR: Comparison of observed paramagnetic shifts and calculated pseudocontact shifts. *Biochemistry 32*:11516–11523.

Toma S, Campagnoli S, Margarit I, Gianna R, Grandi G, Bolognes, M, De Filippis V, Fontana A. 1991. Crafting of a calcium binding loop of thermolysin to *Bacillus subtilis* neutral protease. *Biochemistry 30*:97–106.

Turner DL, Williams RJP. 1993. H- and C-NMR investigation of redox-state-dependent and temperature-dependent conformation changes in horse cytochrome *c*. *Eur J Biochem 211*:563–568.

Veitch NC, Whitford D, Williams RJP. 1990. An analysis of pseudocontact shifts and their relationship to structural features of the redox states of cytochrome *b5*. *FEBS Lett 269*:297–304.

Wagner G, Hyberts SG, Havel TF. 1992. NMR structure determination in solution: A critique and comparison with X-ray crystallography. *Annu Rev Biomol Struct 21*:167–198.

Wand AJ, DiStefano DL, Feng Y, Roder H, Englander SW. 1989. Proton resonance assignments of horse ferrocytochrome *c*. *Biochemistry 28*: 186–194.

Williams G, Claydon NJ, Moore GR, Williams RJP. 1985. Comparison of the solution and crystal structures of mitochondrial cytochrome *c*. *J Mol Biol 183*:447–460.

Wu LC, Laub PB, Elöve GA, Carey J, Roder H. 1993. A noncovalent peptide complex as a model for an early folding intermediate of cytochrome *c*. *Biochemistry 32*:10271–10276.

Wüthrich K. 1976. *NMR in biological research: Peptides and proteins*. Amsterdam/New York: North-Holland/American Elsevier.

Wüthrich K. 1986. *NMR of proteins and nucleic acids*. New York: Wiley-Interscience.