PROTEINS: Structure, Function, and Genetics 29:348–358 (1997)

Paramagnetic Relaxation as a Tool for Solution Structure Determination: *Clostridium pasteurianum* Ferredoxin as an Example

Ivano Bertini,[1]* Antonio Donaire,[1] Claudio Luchinat,[2] and Antonio Rosato[1]
[1]*Department of Chemistry, University of Florence, 50121 Firenze, Italy*
[2]*Department of Soil Science and Plant Nutrition, University of Florence, 50144 Firenze, Italy*

ABSTRACT **The possibility of using the relaxation properties of nuclei for solution structure determination of paramagnetic metalloproteins is critically evaluated. First of all, it is theoretically and experimentally demonstrated that magnetization recovery in nonselective inversion recovery experiments can be approximated to an exponential in both diamagnetic and paramagnetic systems. This permits the estimate of the contribution of paramagnetic relaxation when dominant or sizable. Then, it is shown that the averaging of paramagnetic relaxation rates due to cross relaxation is often tolerably small with respect to the use of paramagnetic relaxation rates as constraints for structural determination. Finally, a protocol is proposed to use such paramagnetic relaxation rates, which depend on the sixth power of the metal to resonating nucleus distance, as constraints for solution structure determination of proteins. As an example, the available solution structure of the oxidized ferredoxin from Clostridium pasteurianum has been significantly improved in resolution especially in the proximity of the metal ions by using 69 new constraints based on paramagnetic relaxation. Proteins 29:348–358, 1997.** © **1997 Wiley-Liss, Inc.**

Key words: ferredoxin; NMR; paramagnetic; relaxation; solution structure

INTRODUCTION

NMR is a tool which may lead to the determination of protein structures in solution, through the determination of proton-proton dipolar interactions and of internuclear scalar couplings.[1] In dealing with metalloproteins, a drawback in the determination of solution structures is represented by the metal itself, which generally does not provide information through NMR about its geometrical relationship with the donor atoms and the immediate neighbors. Sometimes magnetically active metal isotopes (e.g. ^{113}Cd, ^{119}Hg)[2–6] have been used to establish links from the metal to the protein frame.

If the metal is paramagnetic, it is more difficult to detect dipolar and scalar connectivities because of line broadening in the proximity of the paramagnetic center, which in turn affects the quality of the spectra. The degree of worsening depends on the nature of the metal ion.[7] In some favorable cases of proteins containing Fe_4S_4 clusters[8–12] or low spin iron(III) hemes,[13] the solution structure of the protein part has been obtained at a degree of resolution comparable to that of diamagnetic proteins of the same size, whereas the metal ion(s) has been anchored to the protein through links not experimentally detected.

On the other hand, the contact and pseudocontact shifts and the relaxation enhancements induced by the paramagnetic center are a potential source of independent structural information. In the case of reduced high potential iron-sulfur proteins and of oxidized Fe_4S_4-containing ferredoxins a relationship has been proposed between the contact shifts of cysteine β protons and the Fe-S-C-H dihedral angle.[14] In heme-containing proteins pseudocontact shifts have been shown to provide meaningful constraints.[15]

We demonstrate here the possibility of using the enhancement of nuclear relaxation rates due to the presence of unpaired electrons, which are mainly localized onto the metal ions, in order to obtain additional structural constraints involving the metal ion and the resonating nuclei. Nuclear relaxation enhancements have the same r^{-6} dependence as the NOE's,[16] and can be accommodated into any computer program devoted to the calculation of protein solution structures.

This paper consists of two parts. The first is devoted to a critical analysis of the results of non selective inversion recovery experiments. This analysis is not new in a theoretical sense, but allows us to

Abbreviations: NMR: nuclear magnetic resonance; NOE: nuclear Overhauser effect; CpFd: Clostridium Pasteurianum ferredoxin; NOESY: nuclear Overhauser effect spectroscopy; DG: distance geometry.

*Correspondence to: Ivano Bertini, Department of Chemistry, University of Florence, 50121 Firenze, Italy.
E-mail: bertini@risc1.lrm.fi.cnr.it
Received 7 November 1996; Accepted 11 November 1996

evaluate the sources and the extent of errors in determining the various parameters. The paramagnetic contribution to nuclear relaxation may contain structural information about the metal-nucleus distances which have always been pragmatically used.[17–20] However, statements in the pertinent and authoritative literature[21,22] underline the many pitfalls of such procedure, especially when applied to macromolecules. Indeed, cross-relaxation undermines the validity of the definition of T_1 itself and the use of paramagnetic relaxation as a structural constraint. A complete relaxation matrix analysis, which takes into account proton-proton and unpaired electron-proton interactions, is needed to evaluate deviations from simple theories.

The second part of the paper describes a procedure for using structural constraints obtained from paramagnetic relaxation for the determination of the solution structure of paramagnetic proteins. The use of paramagnetic relaxation rates for the determination of solution structures of proteins is conceptually different from their use for obtaining structural models of any kind of molecule, as done in the past. In the former case many constraints are needed to bring to a correct folding a randomly generated unfolded protein; in the latter case a structural model is generated which is consistent with one or more estimated interatomic distances. The structure, as determined by a family of structures, can be more or less defined but should not be wrong; a structural model can be somewhat different from the actual structure. It is proposed here that the constraints obtained from paramagnetic relaxation are used together with NOE or ^3J constraints to obtain a solution structure. As an example we take a protein whose solution structure has been recently solved by using only NOE's.[12] The protein is the oxidized ferredoxin from Clostridium pasteurianum (CpFd hereafter), which contains two $[Fe_4S_4]^{2+}$ clusters and is constituted by 55 amino acids. Owing to the many paramagnetic centers and the relatively small size of the protein, about 50% of the protons are affected by paramagnetism, and the number of detected NOE's is relatively low.[12] The inclusion of paramagnetic constraints based on nuclear relaxation significantly increases the number of constraints and leads to a sizable improvement in the overall definition of the structure.

MATERIALS AND METHODS

Sample preparation and NMR experiments

CpFd samples were prepared as previously described.[12] The sample concentrations were 3 mM in CpFd, 1 M NaCl, 50 mM phosphate buffer, pH 6.0. 1D inversion-recovery experiments (recycle delay-180°-τ-90°-AQ) at 45 different τ values ranging from 0.5 ms to 2 s and with 2 s of recycle delay were acquired. Different methods are reported in the literature to measure non-selective relaxation rates through 2D experiments.[23,24] In analogy to the procedure suggested by Bystrov and coworkers,[23] six Weft-NOESY experiments (recycle delay-180°-τ-NOESY) with 30 ms of NOESY mixing time τ_m and τ values of 10, 20, 40, 80, 160 and 300 ms were performed (1.4 s of recycle delay, 2K × 1K data points). The data were Fourier transformed to obtain a final 2K × 2K data matrix, by using a cosine squared function of apodization in both dimensions and zero-filling. The Weft-NOESY experiment is similar to the previously proposed SKEWSY experiment (Skewed Exchange Spectroscopy[25]). In both experiments, the intensity of a given cross peak is proportional to the residual magnetization after the 180°-τ perturbation of the proton whose diagonal signal lies in the same column of the cross peak itself,[25,26] provided the recycle delay is long enough.

Computational methods

The Weft-NOESY build-up's were best fitted to an exponential recovery (Equation 1), as mentioned by Bystrov and coworkers,[23] with a rate constant indicated as ρ_I^{eff}

$$\langle I(t)\rangle = \langle I(\infty)\rangle - [\langle I(\infty)\rangle - \langle I(0)\rangle] \exp(-\rho_I^{eff} t) \qquad (1)$$

where $\langle I(\infty)\rangle$ is the equilibrium value of the magnetization, $\langle I(0)\rangle$ is the value of the magnetization immediately after the inversion pulse, and $\langle I(t)\rangle$ is the value of the magnetization at time t. The lower limits for the paramagnetic contributions to nuclear relaxation, ρ_I^{eff}(para), were determined by subtracting a value of 5 s^{-1} from ρ_I^{eff}. 5 s^{-1} would be the upper limit for proton relaxation rates if there were not paramagnetic contributions (see later).

To include distance constraints into structure calculations, it is necessary to relate ρ_I^{eff}(para) to the metal(s)-proton distance. As discussed elsewhere by us[27] and by La Mar and coworkers[28,29] we have assumed that

$$\rho_I(\text{para}) = K \sum_M r_{IM}^{-6} \qquad (2)$$

where r_{IM} is the distance of the proton I to the iron M and the sum is extended to all the iron atoms of the protein. ρ_I(para), is the theoretical paramagnetic contribution to nuclear relaxation rates, in a metal-centered approximation (see also Equation 11). It may differ from ρ_I^{eff}(para), which is experimentally determined, because of spin-diffusion effects (see discussion). Oh and Markley used a so-called "reduced distance,"[30] which represents a more approximate approach, to qualitatively relate distances and T_1 values.

The protocol for structure determination we want to propose in this paper relies on the use of ρ_I^{eff}(para) just like and in conjunction with NOE's, owing to the similarity of their dipolar origin. Both sets of con-

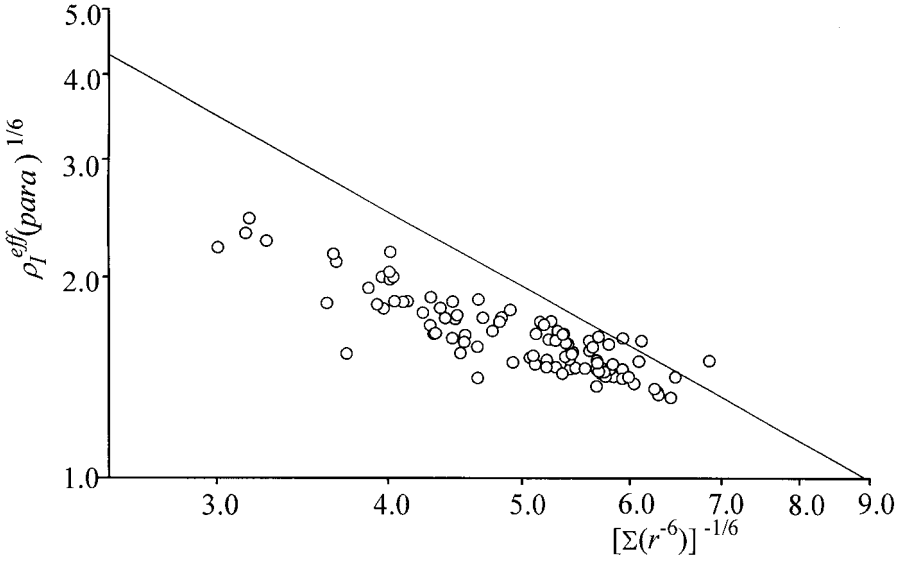

Fig. 1. Plot of the $\rho_I^{eff}(para)$ of *Cp*Fd protons versus $(\Sigma r_{ij}^{-6})^{-1/6}$, r_{ij} being the proton–iron distance and the sum being extended to the eight iron atoms of the protein. From the straight line the iron-nuclei upper distance limits can be obtained for each ρ_I^{eff}(para)

straints should be calibrated with the same procedure. We have used the same method as that proposed by Wütrich in the program CALIBA.[31] The initial upper limits are set by considering the experimental ρ_I^{eff}(para) of the Hα proton of a coordinating cysteine, and assuming that it corresponds to the longest possible distance between the coordinated Fe and the Hα proton plus 0.5 or 1.0 Å. When the structure starts to be defined after a first round of calculations (we have used the program DIANA,[31] see below), it is possible to decrease the upper limits. The final result of the calibration is shown in Figure 1.

The DG program DIANA[31] was modified in order to implement the use of constraints arising from the ρ_I^{eff}(para) values. In this new version, called PRETTIDIANA (Paramagnetic Relaxation Times to Improve DIANA), a further term has been added to the target function (i.e. a function of the discrepancy between an experimental limiting distance and the actual distance in the structure) to be minimized. Since the physical dependence on distance of both NOEs and ρ_I^{eff}(para) is the same (proportional to r^{-6}), this new term was defined similarly to that employed for NOE's constraints in the original program,[31] i.e. by adding to the target function T a term T_{para}:

$$T_{\text{para}} = \sum_i w_i \left(\frac{d_i^2 - b_i^{c^2}}{2d_i}\right)^2 \quad \begin{array}{l}\text{if } c = u \text{ and } d_i > b_i^c, \\ \text{or } c = l \text{ and } d_i < b_i^c\end{array}$$

$$T_{\text{para}} = 0 \qquad \text{otherwise} \qquad (3)$$

where c labels the type of constraint (u = upper distance limit, l = lower distance limit), i labels the constraints, w_i (>0) is the weight of the i-th constraint, b_i^c is the experimental distance constraint and d_i is give by:

$$d_i = \left(\sum_M r_{IM}^{-6}\right)^{-1/6} \qquad (4)$$

r_{IM} being the distance of the proton I to the iron M in the calculated structure and the sum being extended to all the iron atoms of the protein. Note that both d_i and b_i^c have the dimension of a distance, which makes T_{para} directly comparable to the contribution to the total target function arising from NOE's. The w_i values were taken equal to 1.0.

Simulated inversion-recovery experiments were calculated as explained in the next section through a modified version of the program CORMA,[32] version 5.0. An isotropic correlation time for all dipolar interactions was assumed, equal to the correlation time for the reorientation of the whole molecule.

RESULTS AND DISCUSSION

About non selective saturation recovery experiments

The 180°-τ-NOESY experiment, the 180° pulse being non selective, provides NOESY maps where the intensities of the cross peaks depend on τ. The τ values determine the extent of magnetization recovery of each proton depending on its relaxation characteristics. It can be shown through simulated experiments and other considerations[25,26] that the intensity of a cross peak is proportional to the intensity of the diagonal peak in the same column which, in turn, is proportional to its own recovery. The longitudinal relaxation rate of spin I, ρ_I, is defined, according to

the Bloch equations, as the rate constant for the return to equilibrium of a spin ensemble after a perturbation.[33] This definition implies that the spins of the ensemble interact with a lattice having infinite heat capacity. In practice, however, the spins under consideration do interact through cross relaxation with other spin ensembles, which constitute a part of the lattice not having infinite heat capacity. This interaction makes, in principle, the whole process not exponential. The recovery of longitudinal magnetization of spin I after a perturbation is given by:

$$-\frac{d\langle I_z(t)\rangle}{dt} = \rho_I[I_z(t) - I_z(\infty)] + \sum_J \sigma_{IJ}[J_z(t) - J_z(\infty)] \quad (5)$$

where σ_{IJ} are the cross-relaxation rate constants between spin I and all other spins J and ρ_I is the sum of all contributions to relaxation due to coupling with all other spins (neglecting for the moment other external sources of relaxation):

$$\rho_I = \sum_J \rho_{IJ} \quad (6)$$

The ρ-containing term in Equation 5 involves the time-dependent deviation from equilibrium magnetization ($\langle I_z(t)\rangle - \langle I_z(\infty)\rangle$) of spin I. The σ-containing terms are also time-dependent, because they involve the time-dependent deviation from equilibrium magnetization of the other spins. $\langle I_z(t)\rangle$ can be followed experimentally as a function of time through selective or non selective experiments, i.e. by either perturbing all spins, or only spin I with a soft pulse. The time dependencies are different in the two cases.[34] This has induced spectroscopists to talk in terms of "selective T_1" and "non selective T_1," by approximating part or all of the $\langle I_z(t)\rangle$ versus t curve to an exponential in the two cases.

It is possible to calculate all σ_{IJ} and ρ_{IJ} terms from a starting structural model.[35] The equations to use for homonuclear I and J species are:[36]

$$\sigma_{IJ} = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\hbar^2\gamma_I^4}{10r_{IJ}^6}\left(\frac{6\tau_c}{1 + 4\omega_I^2\tau_c^2} - \tau_c\right) \quad (7)$$

$$\rho_{IJ} = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{\hbar^2\gamma_I^4}{10r_{IJ}^6}\left(\tau_c + \frac{3\tau_c}{1 + \omega_I^2\tau_c^2} + \frac{6\tau_c}{1 + 4\omega_I^2\tau_c^2}\right) \quad (8)$$

Here r_{IJ} is the internuclear I–J distance, γ_I is the nuclear magnetogyric ratio, ω_I the nuclear Larmor frequency in rad s^{-1}, and τ_c is the time constant for the reorientation of the two spins. For a rigid spherical molecule, τ_c equals the time constant for the rotation of the whole molecule, τ_r. The value of τ_r can, in principle, be estimated through the Stokes-Einstein equation:[37]

$$\tau_r = \frac{4\pi\eta a^3}{3kT} \quad (9)$$

where η is the microscopic viscosity of the solution, α is the radius of the molecule (assumed spherical), T is the absolute temperature and the other symbols have their usual meanings. Internal motions, e.g. methyl rotations and librations, decrease τ_c with respect to τ_r. This effect is larger, the larger the protein.[38]

The curves of $\langle I_z(t)\rangle$ versus t for all protein protons after a non selective 180°-τ perturbation can be obtained through integration of the complete set of N coupled differential Equations 5, which has been done by using the formalism of the relaxation matrix.[35,39] In the relaxation matrix formalism, the function $\langle I_z(t)\rangle$ is given by:

$$\langle \mathbf{I}_z(t)\rangle = \mathbf{I}_\infty - (\mathbf{I}_\infty - \mathbf{I}_0)\exp(-\mathbf{R}\tau) \quad (10)$$

where R is the relaxation matrix with the ρ values on the diagonal and the σ values off diagonal. The calculations have been performed on the N protons of the oxidized ferredoxin from Clostridium pasteurianum, using the structure previously determined through NMR with only NOE constraints[12] as input model. We have simulated inversion recovery curves for different τ_c values ranging from 0.1 to 10 ns. Then, such curves have been fitted to Equation 1. The quality of the fitting is expressed by its standard deviation, which gives an idea of the deviation from exponential behavior. The fitting of the curves always provided a rate constant, ρ_I^{eff}(dia), with small (lower than 10%) standard deviation. Therefore, such curves have a substantially exponential behavior.

In Figure 2A the calculated plot of $\langle I_z(t)\rangle$ versus t is shown for Gln25 Hϵ22 proton for different τ_c values in the range 0.1–10 ns. In the case of the shortest τ_c (0.1 ns), a ρ_I^{eff} still smaller than 5 s^{-1} is calculated. An increase of the value of τ_c corresponds to a decrease of the value of ρ_I^{eff}. From Stokes-Einstein equation, a correlation time for molecular reorientation of the protein under study of about 2.5 ns is predicted. Local motions may decrease τ_c values in some regions to values which are presumably still higher than 0.1 ns. Therefore a value of 5 s^{-1} constitutes a safe overestimate of the value of ρ_I^{eff}(dia). For larger proteins, smaller values may be considered, due to the decrease of ρ_I^{eff}(dia) upon increase of τ_c. The experimental magnetization recovery of the signal of Gln25 Hϵ22 versus the τ delay is reported in Figure 2B. The fitting of the curve yields a ρ_I^{eff} value of 3.97 ± 0.07 s^{-1}. The small error on the value of ρ_I^{eff} proves that the experimental build-up is effectively exponential, and fits well to Equation 1. This proton

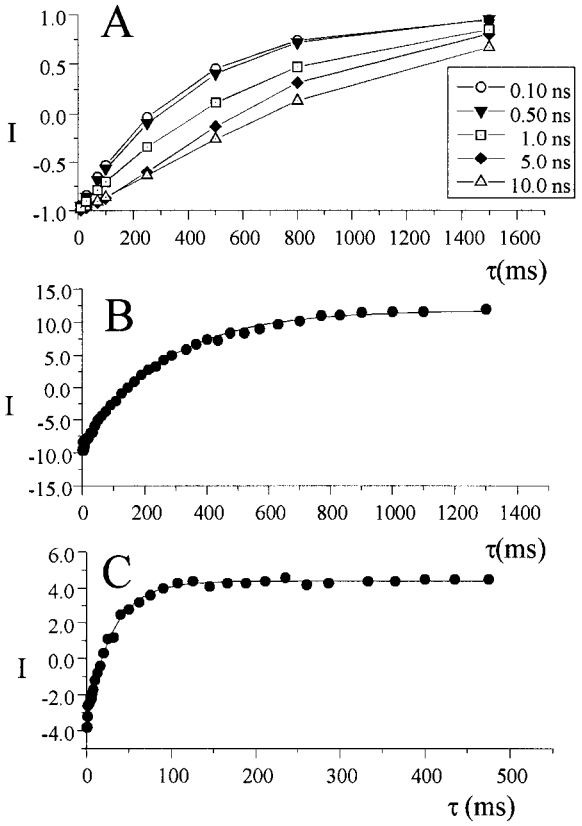

Fig. 2. **A:** Calculated magnetization recovery curves for Gln25 Hϵ22 of oxidized CpFd. The results for different values of τ_c ranging from 0.1 ns to 10 ns are shown. **B:** Experimental magnetization recovery obtained from a series of 1D inversion recovery experiments. The curve obtained through three parameters fitting according to Equation 1 is shown. **C:** Experimental magnetization recovery of Cys40 Hα of oxidized *Cp*Fd, together with the three parameters fitting curve.

is far from the polymetallic center (the distance from the closest iron ion is 11.0 Å), and therefore senses paramagnetic effects to a negligible extent.

In paramagnetic metalloproteins another term should be added to the sum of the ρ_{IJ} values (Equation 6), which is the contribution to nuclear relaxation due to the dipolar coupling with the unpaired electrons. This term is given by the well known Solomon equation:[16]

$$\rho_{IM} = \frac{2}{15}\left(\frac{\mu_0}{4\pi}\right)^2 \frac{\gamma_I^2 g_e^2 \mu_B^2 S(S+1)}{r_{IM}^6} \cdot \left(\frac{7\tau_s}{1+\omega_s^2\tau_s^2} + \frac{3\tau_s}{1+\omega_I^2\tau_s^2}\right) \quad (11)$$

where γ_I is the nuclear magnetogyric ratio, r_{IM} is the nucleus-metal distance, S is the electron spin angular momentum, and τ_s is the time constant for the nucleus-electron interaction. All other symbols have their usual meaning. Equation 11 is valid in the approximation of the electron being localized on the metal ion. This approximation holds as one gets a few bonds further from any iron ion.

The process of nuclear relaxation due to coupling with unpaired electrons is exponential because the electron magnetization is always at thermal equilibrium in absence of cross relaxation. Cross relaxation makes, in principle, non exponential the recovery of magnetization, but now this problem can be ignored because the increase of ρ_I makes σ still less important. In Figure 2C the experimental values of $\langle I_z \rangle$ versus the τ delay for the Hα of Cys40 in an inversion recovery experiment are reported. The result of the fitting to Equation 1 gives a ρ_I^{eff} value of 34.5 ± 0.9 s^{-1}. Just like the case of Gln25 Hϵ22, and despite the different origin of relaxation, the curve is very well fitted to Equation 1.

The problem of cross relaxation with respect to ρ_I^{eff}(para) is that it undermines the proportionality of ρ_I^{eff}(para) to r^{-6}. For the case of a two proton system, it has been demonstrated that when τ_c is large enough to make the cross relaxation rate between the two protons larger than the difference of their individual relaxation rates, the measured value of ρ_I^{eff}(para) is the same for both protons, and equal to the average of the two different paramagnetic relaxation rates.[21] This brings us to the question of Granot[21] and La Mar[22] on whether it is reasonable to extract from the experimental ρ_I^{eff}(para) the distances to the paramagnetic center in the presence of cross relaxation. The answer to this question comes from the comparison of ρ_I(para) calculated through Equation 2 (i.e. without considering cross relaxation effects), and ρ_I^{eff}(para) calculated by taking into account cross relaxation through the relaxation matrix analysis (Table I). The value of K in Equation 2 has been evaluated by comparing the experimental magnetization recovery curves with the calculated curves for protons which do not experience sizable cross relaxation effects. The value of K is $2.5 \cdot 10^5 \pm 0.5 \cdot 10^5$ $Å^6 s^{-1}$, which is, within its error, equal to the value obtained from a best fitting of the NOESY cross-peaks in a similar system.[27] For the relaxation analysis, the term obtained from Equation 2 for each proton was added to the corresponding diagonal term of the relaxation matrix, i.e. to the ρ_I of Equation 6.[27] Inversion recovery curves were then calculated, whose fittings to Equation 1 provided the calculated ρ_I^{eff} for all protons. The calculated ρ_I^{eff}(para) are obtained by subtracting from the above ρ_I^{eff} values the ρ_I^{eff}(dia) values, which were calculated through the same procedure, but without inclusion of paramagnetic effects. Since cross relaxation depends on τ_c (Equation 7), also these ρ_I^{eff}(para) values depend on τ_c. Besides some fast reorienting protons, a reasonable estimate of effective τ_c for the present protein is 0.8 ns, which also accounts for the cross-peak intensities in standard NOESY experiments. Smaller values of τ_c decrease the effect of cross

TABLE I. Theoretical ρ_I(para)* Values for Selected Protons in the Model Ferredoxin Structure, and Estimated $\rho_I^{eff}(para)$ Values Obtained From Fitting of Simulated Weft-NOESY Experiments†

Proton	Distance from closest iron ion (Å)	$\rho_I(para)$ (s^{-1}) (neglecting cross-relaxation)	$\rho_I^{eff}(para)$ (s^{-1}) (including cross-relaxation) $\tau_r = 0.8$ ns	$\rho_I^{eff}(para)$ (s^{-1}) (including cross-relaxation) $\tau_r = 2.5$ ns
HN Ala 1	8.3	1.50	1.46	1.71
Hϵ1 Tyr 2	6.4	7.1	7.17	7.08
HN Lys 3	6.8	6.8	6.70	6.80
Hα Ile 4	6.8	8.5	8.61	9.22
Hβ Val 9	6.9	5.6	5.70	6.22
Hα1 Gly 12	4.7	54.7	54.2	52.0
Hα2 Gly 12	5.8	13.9	14.7	16.7
Hβ2 Pro 19	6.7	6.0	6.22	6.79
Hγ3 Pro 19	5.0	31.0	31.1	32.6
Hϵ21 Gln 54	7.8	2.70	2.66	2.42
Hϵ22 Gln 54	9.2	1.10	1.29	1.75

*See Equation 2.
†$\tau_r = 0.8$ ns (effective τ_r derived from comparison with the experimental data) and $\tau_r = 2.5$ ns. The simulated data were fitted to Equation 1, in which we forced $\langle I(\infty)\rangle = -\langle I(0)\rangle = 1$.

relaxation and make the difference between ρ_I(para) and ρ_I^{eff}(para) smaller. The ρ_I(para) values calculated through Equation 2, which neglect cross relaxation, and the calculated ρ_I^{eff}(para), which take it into account, are compared in Table I for some selected protons. The effect of partial averaging of ρ_I^{eff}(para) due to the σ containing term is shown to be either negligible or as large as a few percent. Only in the case of geminal protons does the error reach 10%. However, average distances from a geminal pair in the absence of stereospecific assignment are routinely used for solution structure determination, in which case a 10% error is irrelevant. If the stereospecific assignment is known, the safest strategy would be to retain only the constraint relative to the faster relaxing proton (i.e. the one closer to the paramagnetic center(s)).

These results can be generalized also for larger proteins, for which a larger τ_c is expected to be operative. As already reported,[38,40] in macromolecules the internal motions of protein groups (e.g. methyl rotation and general movements of residues) contribute to proton relaxation, and their relative contribution is larger the larger the protein. For proteins of molecular weight up to about 30000, which is the present limit for the determination of structure of proteins, experimental proton ρ_I^{eff}'s are still around 1 s^{-1}.[38] Such values correspond to an effective τ_c value lower than 2.5 ns. In Table I the effect of cross relaxation on ρ_I^{eff}(para) for a value of τ_c of 2.5 ns is also shown. The averaging effect of ρ_I^{eff}(para) values due to cross relaxation is still acceptable (within 20%). Anyway, as it will be shown, the obtained ρ_I^{eff}(para) can be further analyzed if found to violate the resulting structure, and then discarded.

At this point it is important to underline that i) a ρ_I^{eff} can always be obtained from 2D non selective saturation recovery experiments with relatively small standard deviation; ii) the ρ_I^{eff}(para) values match satisfactorily well the r^{-6} dependence from the metals given by Equation 2, with little evidence of averaging due to cross-relaxation.

Structure refinement

In the present system the ρ_I^{eff}(dia) are 5 s^{-1} or less, as calculated and indeed found for protons far from the iron ions. The ρ_I^{eff}(para) values are found, and calculated from Equation 2, to range from 200 s^{-1} to virtually zero. We can thus safely decide that ρ_I^{eff} values larger than 10 s^{-1} are determined for the most part by paramagnetic relaxation, and from these values we can subtract the maximum ρ_I^{eff}(dia) value of 5 s^{-1}. If ρ_I^{eff}(dia) is smaller than 5 s^{-1}, the nucleus-to-proton distance is slightly overestimated, but this is not a problem when using upper distance limits. Indeed, ρ_I^{eff}(para) as well as the NOE values are not directly used to obtain distances, but are used in a calibration protocol to obtain upper distance limits between atoms. Such upper distance limits are very loose at the beginning, and then tightened up to make them consistent with the resulting structures. For example, in the very first round of calculations one can assume that the ρ_I^{eff}(para) of the slowest relaxing Hα proton among those of the cysteines coordinating the iron ions corresponds to the maximum possible distance between the coordinated iron atom and the Hα proton plus 0.5 or 1.0 Å. If some constraints are seriously not consistent with the structure, they should be reanalyzed or discarded. This has not been the case in the present investigation. One should also keep in mind that loose upper limits are necessary to take the mobility of the protein into account. In Figure 1 the ρ_I^{eff}(para) values are plotted versus the $(\Sigma r^{-6})^{-1/6}$ (where the sum is extended to the eight iron atoms

TABLE II. Results of the Fitting of the Six Weft-NOESY Experiments

Proton	ρ_I^{eff} (s^{-1})	Standard deviation (%)	$\rho_I^{eff}(para)$ (s^{-1})	Upper distance limit (Å)
Ala-1 Hα	47.4	13.4	43.4	4.86
Tyr-2 HN	59.5	20.0	55.9	4.66
Tyr-2 Hα	24.7	16.4	20.7	5.50
Tyr-2 Hβ3	19.7	19.9	15.7	5.75
Tyr-2 Hβ2	21.8	18.0	17.8	5.57
Lys-3 HN	15.8	17.2	11.8	5.96
Lys-3 Hα	15.4	24.3	11.4	6.20
Ile-4 Hα	28.5	16.6	24.5	5.35
Ile-4 Hβ	22.3	17.2	18.3	5.61
Ile-4 Hγ12	39.3	24.6	35.3	5.00
Ala-5 HN	18.3	12.2	14.3	5.84
Cys-8 HN	34.1	14.7	30.1	5.14
Val-9 Hα	29.6	12.3	25.6	5.31
Val-9 Hβ	18.2	8.5	14.2	5.85
Cys-11 Hα	36.3	11.3	32.3	5.11
Gly-12 Hα1	47.0	18.1	43.0	4.87
Gly-12 Hα2	28.4	21.3	24.4	5.35
Cys-14 Hα	49.2	12.5	45.2	4.83
Ala-15 HN	42.7	17.3	38.7	4.93
Ala-15 Hα	18.6	12.6	14.6	5.82
Ser-16 HN	14.7	14.5	10.7	6.13
Glu-17 HN	16.9	19.6	12.9	5.87
Glu-17 Hβ3	18.3	12.5	14.3	5.83
Glu-17 Hβ2	20.3	13.6	16.3	5.65
Glu-17 Hγ2	17.4	20.1	13.4	5.90
Glu-17 Hγ3	19.2	23.2	15.2	5.78
Cys-18 HN	22.0	24.2	18.0	5.63
Pro-19 Hα	21.1	21.4	17.1	5.38
Pro-19 Hβ3	21.0	21.6	17.0	5.66
Pro-19 Hβ2	24.8	22.0	20.8	5.49
Asn-21 Hα	25.0	16.1	21.0	5.48
Ile-23 HN	32.3	8.5	28.3	5.13
Ile-23 Hα	26.3	16.9	22.3	5.18
Ile-23 Hβ	25.3	13.1	21.3	5.47
Ser-24 HN	22.7	19.1	18.7	5.59
Ser-24 Hα	14.4	16.7	10.4	6.16
Ile-29 Hα	18.4	12.2	14.4	5.62
Phe-30 HN	42.1	14.7	38.1	4.73
Phe-30 Hβ2	29.7	16.2	25.7	5.32
Phe-30 Hζ	18.6	14.8	14.6	5.76
Val-31 HN	16.6	11.3	12.6	5.71
Ile-32 HN	19.6	17.0	15.9	5.74
Ile-32 Hα	34.1	15.8	30.1	5.05
Ile-32 Hβ	26.7	12.2	22.7	5.41
Ile-32 Hγ12	35.7	14.6	31.7	5.12
Asp-33 HN	21.3	18.5	17.3	5.64
Ala-34 Hα	18.2	14.3	14.2	5.78
Thr-36 Hβ	17.4	9.6	13.4	5.88
Cys-37 HN	29.6	19.7	25.6	5.31
Ile-38 Hα	34.4	8.8	30.4	5.15
Ile-38 Hβ	20.1	9.3	16.1	5.73
Asp-39 Hβ2	17.7	17.9	13.7	5.72
Cys-40 Hα	36.1	13.2	32.1	5.11
Gly-41 Hα1	45.1	11.0	41.1	4.91
Gly-41 Hα2	41.2	11.2	37.2	4.97
Cys-43 Hα	47.4	19.7	43.4	4.92
Ala-44 HN	47.6	5.9	43.6	4.73
Ala-44 Hα	32.7	9.9	28.7	5.21
Asn-45 HN	17.8	16.7	13.8	5.87
Val-46 HN	16.8	7.8	12.8	5.92
Val-46 Hβ	19.2	7.2	15.2	5.78
Pro-48 Hα	27.4	14.2	23.4	5.33
Val-49 Hβ	23.8	15.3	19.8	5.55
Gly-50 HN	28.8	8.2	24.8	5.37
Gly-50 Hα2	19.6	17.9	15.6	5.76
Gly-50 Hα1	16.7	17.2	12.7	5.94
Pro-52 Hα	28.3	9.0	24.3	5.36
Pro-52 Hβ3	50.9	14.7	46.9	4.77
Val-53 HN	15.1	16.9	11.1	6.05

TABLE III. Summary of Parameters and Results Relative to the DG Families of PRETTIDIANA and DIANA*

	PRETTIDIANA	DIANA
Number of constraints (except ρ_I^{eff}(para))	456	456
ρ_I^{eff}(para) constraints	69	—
Average target function (Å^2)	2.12 ± 0.19	0.53 ± 0.16
Average ρ_I^{eff}(para) constraints contribution to the target function (Å^2)	0.70 ± 0.20	—
RMSD from the average structure, backbone atoms (Å)	0.40 ± 0.04	0.51 ± 0.10
RMSD from the average structure, heavy atoms (Å)	0.76 ± 0.04	0.81 ± 0.09

*See Ref. 12.

and the distance is averaged over the final family of structures), on a logarithmic scale. The scattering of these points is very low, indicating a high degree of agreement between the experimental constraints (ρ_I^{eff}(para) values) and the obtained structures, as well as confirming the validity of the proposed method. Following the analogous procedure described for obtaining distance limits from NOE's constraints with the program CALIBA,[31] a straight line with slope −1 can be drawn in such a way that most of the ρ_I^{eff}(para) values remain below it. From this line the ρ_I^{eff}(para) values are converted into the upper limits for proton-iron distances. It should be pointed out that, at variance with the approach used in CALIBA, the slope of the line is fixed, owing to the r^{-6} dependence of ρ_I^{eff}(para), and that, in practice, adjusting the line means adjusting the value of K in Equation 2 (where ρ_I^{eff}(para) replaces ρ_I(para)).

The ρ_I^{eff}(para) values of 69 protons were used in calculations together with all the previous NOE's. The values of the relaxation rates as well as the standard deviations are given in Table II. Despite it is not important, we may note that in all cases these standard deviations are lower than 25%. For comparison purposes, we may remember that the NOE's are often affected by an error of 50% or larger. We have thus translated these 69 ρ_I^{eff}(para) values into upper distance limits as described above, and the latter were introduced into PRETTIDIANA calculations (see experimental section). Their values are also given in Table II. All the other constraints (upper distance limits arising from NOESY and 1D NOE experiments, restraints derived from the cluster topology, from the ϕ dihedral angles and χ2 dihedral angles of the coordinated cysteines) were the same as in the previous published work.[12] DG calculations were started from 500 random structures and 4 cycles of calculations were carried out, according to

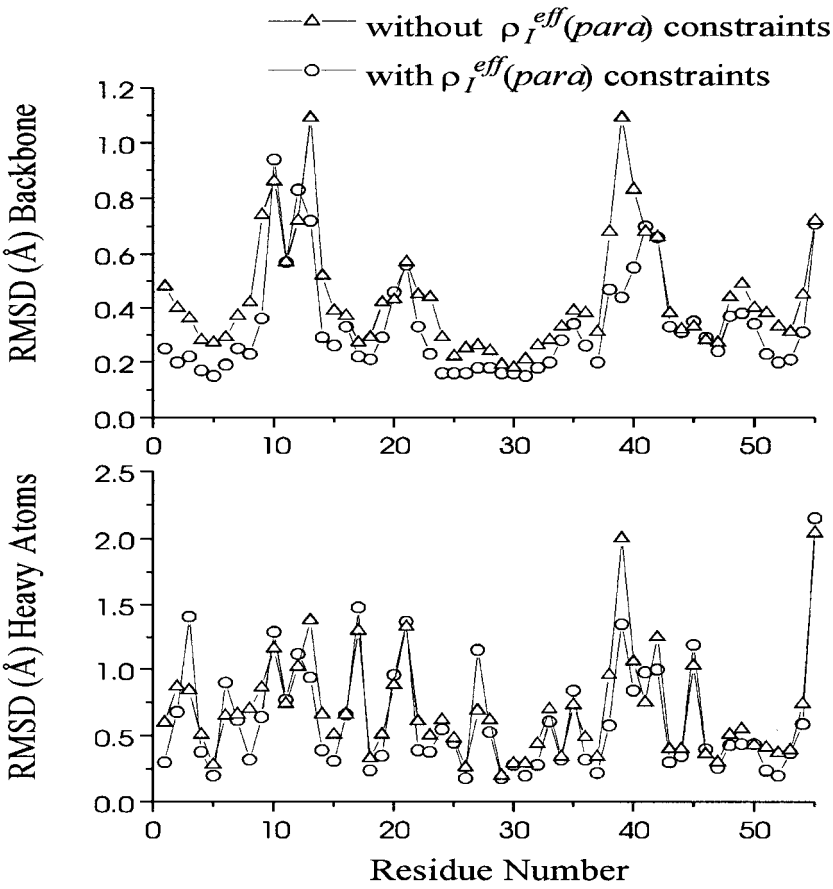

Fig. 3. RMSD from the average structure per residue for the backbone (*upper*) and heavy atoms (*lower*) for DIANA and PRETTIDIANA families.

the REDAC[31] strategy. The resulting best sixteen structures, all of them with target functions lower than 2.5 Å^2, showed an RMSD from the average structure (calculated on residues 3–53) of 0.40 Å for the backbone atoms and 0.76 Å for the heavy atoms. For the analogous published family, calculated only with conventional constraints, the RMSD values at this stage were of 0.51 Å and 0.81 Å, respectively.[12] Despite the contribution of the NOE constraints to the target function is larger in the PRETTIDIANA than in the DIANA family (Table III), no consistent violations are found after introduction of the ρ_I^{eff}(para) constraints. Since the aim of the present study is to assess the possibility of using relaxation constraints in the calculation of the solution structure of a paramagnetic metalloprotein, proton-proton distance constraints were taken from our previous work[12] without any modifications. This may be the reason for the increase of the contribution of the NOE constraints to the target function. When starting the calculation of the solution structure of a paramagnetic metalloprotein from scratch, NOE's should be calibrated in the presence of all other constraints. The constraints used for the calculations and their results are summarized in Table III.

In Figure 3 the RMSD per residue for the family previously published[12] and those for the PRETTIDIANA family are shown. By looking at the RMSD per residue relative to the backbone atoms (Figure 3A), it can be immediately appreciated that the overall definition of the protein folding greatly improves upon introduction of ρ_I^{eff}(para) constraints. This is particularly striking for the region close to cluster II (i.e. for residues 35–45). This region was poorly defined in the old family because of the small number of NOE constraints arising from protons in the neighborhoods of cluster II. The determination of ρ_I^{eff}(para) for some protons in this region gives an additional set of important constraints that univocally determine the orientation of the backbone. For the regions close to cluster I (residues 8–20), the effect of the inclusion of new constraints is still sizable, but less impressive. This is probably due to the fact that there are fewer distance limits in this region as well as fewer ρ_I^{eff}(para) constraints. Even the definition of the regions furthest from the paramagnetic centers (e.g. residues 20–30) is somehow slightly improved upon introduction of the new constraints. In Figure 4 a stereoview of both the

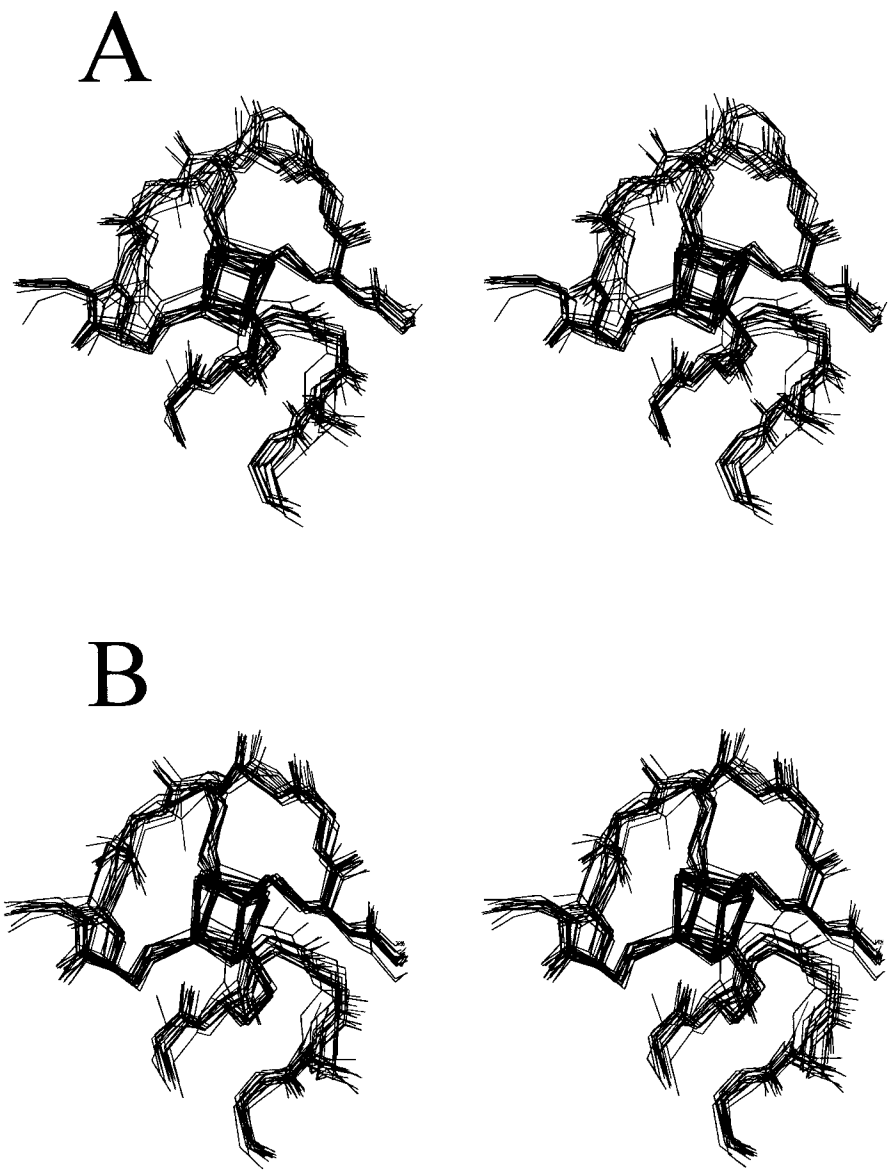

Fig. 4. A stereoview of DIANA (**A**) and PRETTIDIANA (**B**) families in the region close to cluster II. Only backbone and cysteine atoms are shown.

previously published and PRETTIDIANA family for the region close to cluster II is shown. The increase in the resolution of the backbone can clearly be observed. It can be observed that the definition of the polymetallic center is improved as well. The new constraints provide experimental links between the iron atoms and the protein protons, thus increasing the extent of information available both for the polypeptide chain and the polymetallic center. The improvement of the definition of the latter with respect to the previously published family is a logical consequence of this. Indeed, the ρ_I^{eff}(para) constraints are the only source of experimental information on the polymetallic center.

CONCLUDING REMARKS

We have shown that paramagnetic metalloproteins offer the possibility of using constraints that are different than the traditional ones, that is NOE's and scalar couplings, through the measurement of relaxation times. The example chosen for the elucidation of both the theoretical background and the practical implementation of the method is quite appropriate in that eight paramagnetic centers (the eight iron ions) are present in a small protein frame. The use of ρ_I^{eff}(para) as metal-proton distance constraints proves robust and useful. It should be noticed that the distance constraints based on

ρ_I^{eff}(para) are effective at a larger distance than those based on NOE's, because the electron magnetic moment is larger than that of protons.

A comment must be devoted to the problem of spin delocalization over the ligand atoms. This effect may, in principle, introduce a bias when extracting the metal-nucleus distance from the paramagnetic contribution to nuclear relaxation rate. This bias rapidly falls off as one moves away from the metal center(s). In the present test case, constraints relative to the Hβ protons of the cysteines coordinating the polymetallic center were not included, as it has been shown that ligand-centered relaxation is not negligible for these protons.[41] However, ρ_I^{eff}(para) constraints for the Hα protons of the same cysteines could be successfully used. The presence and the analysis of consistent violations would reveal whether a ρ_I^{eff}(para) constraint is seriously affected by ligand-centered effects. In such a case, the constraint may be discarded, or simply loosened to account for the ligand-centered effects.

We would like to stress the point that these new experimental constraints provide structural information about the metal center(s) itself, thus overcoming the main drawback of NMR as applied to paramagnetic metalloproteins, that is the lack of directly detectable signals from the metal center or even (in unfavorable cases) from neighboring nuclei. Also heteronuclei, such as ^{13}C, ^{15}N, ^{31}P, etc., can provide direct information on distance constraints to the metal, even when they lie in close proximity to the metal ion. Since heteronuclei in general have low magnetogyric ratios, they can be detected even at small metal-nucleus distances. Furthermore, the small magnetogyric ratio makes cross relaxation problems even less important.

Finally, the smaller the molecule, the smaller the cross relaxation and the smaller its averaging effect among different ρ_I^{eff}(para) values. This means that the quality of ρ_I^{eff}(para) constraints is higher. In small metal-polypeptide or metal-nucleic acid complexes it is common to have a smaller number of NOE constraints than is needed to generate a structure. In this case the use of ρ_I^{eff}(para) constraints may be decisive.

ACKNOWLEDGMENTS

We gratefully acknowledge J.A. Cowan for carefully reading the manuscript and for many valuable suggestions, and M.A. Cremonini for helping us in the implementation of PRETTIDIANA. A.D. thanks the European Union for a post-doctoral fellowship (Programme Training and Mobility of Researchers).

REFERENCES

1. Wüthrich, K. "NMR of Proteins and Nucleic Acids." New York: Wiley, 1986.
2. Otvos, J.D., Engeseth, H.R., Wehrli, S. Multiple-Quantum ^{113}Cd-^{1}H correlation spectroscopy as a probe of metal coordination environments in metalloproteins. J. Magn. Reson. 61:579–584, 1985.
3. Live, D.H., Christopher, L.K., Cowburn, D., Markley, J.L. Identification of proton NMR signals from the metal ligands in cadmium-substituted plastocyanin via two-dimensional multiple-quantum detection in the absence of explicitly resolved ^{1}H-^{113}Cd coupling. J. Am. Chem. Soc. 107:3043–3045, 1985.
4. Coleman, J.E. Cadmium-113 nuclear magnetic resonance applied to metalloproteins. Methods Enzymol. 227:16–43, 1993.
5. Pan, T., Coleman, J.E. GAL4 transcription factor is not a zinc finger but forms a $Zn(II)_2Cys_6$ binuclear cluster. Proc. Natl. Acad. Sci. U.S.A. 87:2077–2081, 1990.
6. Utschig, L.M., Bryson, J.W., O'Halloran, T.V. Mercury-199 NMR of the metal receptor site in MerR and its protein-DNA complex. Science 268:380–385, 1995.
7. Banci, L., Bertini, I., Luchinat, C. "Nuclear and Electron Relaxation: The Magnetic Nucleus-Unpaired Electron Coupling in Solution," Weinheim: VCH, 1991.
8. Banci, L., Bertini, I., Eltis, L.D., Felli, I.C., Kastrau, D.H.W., Luchinat, C., Piccioli, M., Pierattelli, R., Smith, M. The three-dimensional structure in solution of the paramagnetic protein high-potential iron-sulfur protein I from *Ectothiorhodospira halophila* through nuclear magnetic resonance. Eur. J. Biochem. 225:715–725, 1994.
9. Banci, L., Bertini, I., Dikiy, A., Kastrau, D.H.W., Luchinat, C., Sompornpisut, P. The tridimensional solution structure of the reduced high potential iron sulfur protein *Chromatium vinosum* through NMR. Biochemistry 34:206–219, 1995.
10. Bertini, I., Dikiy, A., Kastrau, D.H.W., Luchinat, C., Sompornpisut, P. The three dimensional solution structure of the oxidized HiPIP from *Chromatium vinosum* through NMR: comparative analysis with the solution structure of the reduced species. Biochemistry 34:9851–9858, 1995.
11. Bertini, I., Eltis, L.D., Felli, I.C., Kastrau, D.H.W., Luchinat, C., Piccioli, M. The solution structure of oxidized HiPIP I from *Ectothiorhodospira halophila:* can NMR probe rearrangements associated to electron transfer processes? Chem Eur J 1:598–607, 1995.
12. Bertini, I., Donaire, A., Feinberg, B.A., Luchinat, C., Piccioli, M., Yuan, H. Solution structure of the oxidized 2[Fe_4S_4] ferredoxin from *Clostridium pasteurianum*. Eur. J. Biochem. 232:192–205, 1995.
13. Banci, L., Bertini, I., Bren, K.L., Gray, H.B., Sompornpisut, P., Turano, P. The three dimensional solution structure of the cyanide adduct of *Saccharomyces cerevisiae* Met80Ala-iso-1-cytochrome c. Identification of ligand-residue interactions in the distal heme cavity. Biochemistry 34:11385–11398, 1995.
14. Bertini, I., Capozzi, F., Luchinat, C., Piccioli, M., Vila, A.J. The Fe_4S_4 centers in ferredoxins studied through proton and carbon hyperfine coupling. Sequence specific assignments of cysteines in ferredoxins from *Clostridium acidi urici* and *Clostridium pasteurianum*. J. Am. Chem. Soc. 116:651–660, 1994.
15. Banci, L., Bertini, I., Bren, K.L., Cremonini, M.A., Gray, H.B., Luchinat, C., Turano, P. The use of pseudocontact shifts to refine the solution structures of paramagnetic metalloproteins: Met80Ala cyano-cyt c as an example. JBIC 1:117–126, 1996.
16. Solomon, I. Relaxation processes in a system of two spins. Phys. Rev. 99:559–565, 1955.
17. Yu, L., Meadows, R.P., Wagner, R., Fesik, S.W., NMR studies of the FK506 binding protein bound to a spin-labeled ascomycin analog. J. Magn. Reson. Ser. B 104:77–80, 1994.
18. Kosen, P.A. Spin labeling of proteins. Methods Enzymol. 177:86–121, 1989.
19. Frederick, A.F., Kay, L.E., Prestegard, J.H. Localization of divalent ion sites in acyl carrier protein using relaxation perturbed 2D NMR. FEBS Lett. 238:43–48, 1988.
20. Dunham, S.U., Lippard, S.J. Long-range distance constraints in platinated nucleotides: structure determination of the 5′ orientational isomer of *cis*-[$Pt(NH_3)$(4-amino-

TEMPO)[d(GpG)]]$^{+}$ from combined paramagnetic and diamagnetic NMR constraints with molecular modeling. J. Am. Chem. Soc. 107:10702–10712, 1995.
21. Granot, J. Paramagnetic relaxation in dipolar-coupled homonuclear spin systems. J. Magn. Reson. 49:257–270, 1982.
22. La Mar, G.N. and de Ropp, J.S. NMR Methodology for Paramagnetic Proteins. In: "Biological Magnetic Resonance," Vol. 12. Berliner, L.J., Reuben, J. (eds.) New York: Plenum, 1993:1–78.
23. Arseniev, A., Sobol, A.G., Bystrov, V.F. T_1 Relaxation measurements by two-dimensional NMR spectroscopy. J. Magn. Reson. 70:427–435, 1986.
24. Kay, L.E., Prestegard, J.H. Spin-lattice relaxation rates of coupled spins from 2D accordion spectroscopy. J. Magn. Reson. 77:599–605, 1988.
25. Bremer, J., Mendz, G.L., Moore, W.J. Skewed Exchange spectroscopy: two-dimensional method for the measurment of cross relaxation in 1H NMR spectroscopy. J. Am. Chem. Soc. 106:4691–4696, 1984.
26. Zhu, L., Reid, B.R. An improved NOESY simulation program for partially relaxed spectra: BIRDER. J. Magn. Reson. Ser. B 106:227–235, 1995.
27. Bertini, I., Felli, I.C., Luchinat, C., Rosato, A. A complete relaxation matrix refinement of the solution structure of a paramagnetic metalloprotein: reduced HiPIP I from *E. halophila*. Proteins 24:158–164, 1996.
28. Gorst, C.M., Yeh, Y.-H., Teng, Q., Calzolai, L., Zhou, Z.-H., Adams, M.W.W., La Mar, G.N. ^{1}H NMR investigation of the paramagnetic cluster environment in *Pyrococcus furiosus* three-iron ferredoxin: sequence-specific assignment of ligated cysteines independent of tertiary structure. Biochemistry 34:600–610, 1995.
29. Donaire, A., Zhou, Z.-H., Adams, M.W.W., La Mar, G.N. ^{1}H NMR investigation of the secondary structure, tertiary contacts and cluster environment of the four-iron ferredoxin from the hyperthermophilic archaeon *Thermococcus litoralis*. J. Biomol. NMR 7:35–47, 1996.
30. Oh, B.-H., Markley, J.L. Multinuclear magnetic resonance studies of the 2Fe-2S* ferredoxin from Anabaena species strain PCC 7210. 3. Detection and characterization of hyperfine-shifted nitrogen-15 and hydrogen-1 resonances of the oxidized form. Biochemistry 29:4012–4017, 1990.
31. Güntert, P., Braun, W., Wüthrich, K. Efficient computation of three-dimensional protein structures in solution from Nuclear Magnetic Resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. J. Mol. Biol. 217:517–530, 1991.
32. Eccles, C., Güntert, P., Billeter, M., Wüthrich, K. Efficient analysis of protein 2D NMR spectra using the software package EASY. J. Biomol. NMR 1:111–130, 1991.
33. Ernst, R.R., Bodenhausen, G., Wokaun, A. "Principles of Nuclear Magnetic Resonance in One and Two Dimensions," London: Oxford University Press, 1987.
34. Neuhaus, D., Williamson, M. "The Nuclear Overhauser Effect in Structural and Conformational Analysis." New York: VCH, 1989.
35. Keepers, J.W., James, T.L. A theoretical study of distance determinations from NMR. Two-dimensional nuclear Overhauser effect spectra. J. Magn. Reson. 57:404–426, 1984.
36. Noggle, J.H., Schirmer, R.E. "The Nuclear Overhauser Effect." New York: Academic Press, 1971.
37. Tanford, C. "Physical Chemistry of Macromolecules." New York: 1985.
38. Ishima, R., Shibata, S., Akasaka, A. General features of proton longitudinal relaxation in proteins in solution. J. Magn. Reson. 91:455–465, 1991.
39. Boelens, R., Koning, T.M.G., Van der Marel, G.A., Van Boom, J.H., Kaptein, R. Iterative procedure for structure determination from proton-proton NOEs using a full relaxation matrix approach: application to a DNA octamer. J. Magn. Reson. 82:290–308, 1989.
40. Chen, W., Zhu, X.-H., Avison, J.A., Shulman, R.G. Nuclear magnetic resonance relaxation of glycogen H1 in solution. Biochemistry 32:9417–9422, 1993.
41. Ciurli, S., Cremonini, M.A., Kofod, P., Luchinat, C. ^{1}H NMR of high potential iron-sulfur protein from the purple non-sulfur bacterium *Rhodoferax fermentans*. Eur. J. Biochem. 236:405–411, 1996.