PROTEINS: Structure, Function, and Genetics 29:68–76 (1997)

Pseudocontact Shifts as Constraints for Energy Minimization and Molecular Dynamics Calculations on Solution Structures of Paramagnetic Metalloproteins

Lucia Banci,[1] Ivano Bertini,[1]* Giovanni Gori Savellini,[1] Andrea Romagnoli,[1] Paola Turano,[1] Mauro A. Cremonini,[2] Claudio Luchinat,[2] and Harry B. Gray[3]
[1]*Department of Chemistry, University of Florence, Florence, Italy*
[2]*Institute of Agricultural Chemistry, University of Bologna, Bologna, Italy*
[3]*Beckman Institute, California Institute of Technology, Pasadena, California*

ABSTRACT **The pseudocontact shifts of NMR signals, which arise from the magnetic susceptibility anisotropy of paramagnetic molecules, have been used as structural constraints under the form of a pseudopotential in the SANDER module of the AMBER 4.1 molecular dynamics software package. With this procedure, restrained energy minimization (REM) and restrained molecular dynamics (RMD) calculations can be performed on structural models by using pseudocontact shifts. The structure of the cyanide adduct of the Met80Ala mutant of the yeast iso-1-cytochrome c has been used for successfully testing the calculations. For this protein, a family of structures is available, which was obtained by using NOE and pseudocontact shifts as constraints in a distance geometry program. The structures obtained by REM and RMD calculations with the inclusion of pseudocontact shifts are analyzed. Proteins 29:68–76, 1997.** © 1997 Wiley-Liss, Inc.

Key words: pseudocontact shifts; structure calculations; restrained energy minimization; molecular dynamics

INTRODUCTION

NMR spectroscopy can now be used successfully for determination of the solution structures of paramagnetic metalloproteins, obtaining structures with resolution as good as those obtained for diamagnetic proteins.[1–7] In this respect, NMR experiments tailored for the detection of connectivities between signals of nuclei sensing the paramagnetic center have been carefully designed[8–11]; and the computational methods for structure determination of proteins containing metal ions are also well developed.[12–16]

One or more paramagnetic centers in a protein produce a large spreading in the chemical shifts and an increase of the longitudinal and transverse relaxation rates.[7,17–23] The latter effect determines an increase in the threshold for the detection of cross-peaks between nuclei, which may be serious for nuclei close to the paramagnetic center, with the consequence that the weakest cross-peaks can be lost in the noise. Furthermore, no information is usually available a priori on the position of the metal ion, and its binding to the protein may be arbitrary and not strongly supported by experimental evidence.

On the other hand, the contributions to the chemical shifts[24–27] and to the relaxation rates[21,28] due to the paramagnetic metal ion contain important structural information as they are related to the distance between the metal center and the resonating nuclei.

The pseudocontact shifts are contributions to the hyperfine shifts of nuclei several bonds away from the metal ion and depend on the anisotropy of the magnetic susceptibility tensor within the susceptibility tensor axes. The relationship between pseudocontact shift (δ_{pc}) and coordinates of a proton within the molecular axes is well known[29]:

$$\delta^i_{pc} = \sum_j \frac{1}{12\pi r^3_{ij}}\left[\Delta\chi^j_{ax}(3n^2_{ij} - 1) + \frac{3}{2}\Delta\chi^j_{rh}(l^2_{ij} - m^2_{ij})\right] \quad (1)$$

where l_{ij}, m_{ij}, and n_{ij} are the direction cosines of the position vector of atom i with respect to the jth magnetic susceptibility tensor coordinate system and r_{ij} is the distance between the jth paramagnetic center and the proton i.

Paramagnetic metalloproteins are ideal for the measurement of many pseudocontact shifts as many protons are close to metal ion and sense its magnetic anisotropy. Therefore information on magnetic anisotropy and on the direction of the magnetic anisotropy tensor have been obtained every time a structural model was available.[30–33]

Williams and Xavier[25,34] used the pseudocontact shift in complexes of lanthanides in order to generate structural models consistent with the experimental

*Correspondence to: Ivano Bertini, Department of Chemistry, University of Florence, Via Gino Capponi, 7, 50121 Florence, Italy.
E-mail: bertini@risc1.lrm.fi.cnr.it
Received 20 September 1996; Accepted 23 January 1997

pseudocontact shifts. In this way they generated a family of structures just like a family of structures is nowadays generated with NOEs. Berg and coworkers[35] checked the orientation of some amino acids by using the pseudocontact shifts of a cobalt-substituted zinc finger peptide and the structure of the native zinc finger peptide obtained through NOEs. Gochin[36,37] was the first to calculate the magnetic susceptibility anisotropies by fitting pseudocontact shifts to a solid state structure and to minimize the deviations between calculated and experimental values by allowing structural variations, thus obtaining a "solution" structure. Finally, our laboratory used a solution structure, obtained through NOEs, to obtain the information on the magnetic anisotropy tensor and to minimize the deviations from both NOEs and pseudocontact shifts simultaneously.[15] In this way, for the first time, a solution structure consistent with both NOEs and pseudocontact shifts was obtained.

The use of pseudocontact shifts as further constraints, together with the NOEs and *J* coupling constraints, in the structure calculations of metalloproteins is of fundamental importance because i) it introduces the metal ion in the frame of the protein atoms, and ii) it provides further structural information (and therefore better defines the region) about the metal center.

The pseudocontact shifts were used for the paramagnetic oxidized protein Met80Ala cytochrome c–CN within the frame of the DIANA strategy for structure calculations. DIANA is a distance geometry calculation program[38,39] based on the fitting of the experimental constraints by varying the dihedral angles. We wrote a module, called PSEUDIANA,[15] which takes into account pseudocontact shifts as further constraints and which has been successfully applied. After distance geometry calculations, further energy minimization is usually appropriate. In the present case, the inclusion of the pseudocontact shifts together with the NOEs and the other experimental constraints in an energy minimization and molecular dynamics program, would be required. For this reason we have included in the SANDER module of the AMBER package[40,41] an integrated module (PCSHIFT) that uses the pseudocontact shifts as pseudopotential terms in the energy minimization and molecular dynamics calculations. SANDER already uses J coupling, NOESY volumes, and diamagnetic chemical shifts as constraints to perform restrained molecular dynamics calculations. The very same package can be used to include nuclear T_1 relaxation times in the MD simulations.[12]

In this way we have produced a complete and integrated package of programs that use the pseudocontact shifts for calculating and refining the solution structures of paramagnetic metalloproteins. As part of the presentation of this package, we discuss the general strategy of employing pseudocontact shifts for solution structure determinations.

MATERIALS AND METHODS

The programs have been written in standard FORTRAN 77, and all the tests were carried out on IBM 9076 SP2 parallel system. The family of solution structures of the Ala80Met mutant of yeast cytochrome c,[15] obtained with the PSEUDIANA program, was used in the tests.

Restrained Energy Minimization and Restrained MD Calculations

All the energy minimization calculations performed with the AMBER 4.1 package of programs, with or without the inclusion of the PCSHIFT module of SANDER (see the Results and Discussion section), were carried out by performing 2500 steepest descent minimization cycles and then applying the conjugate gradient algorithm until an RMS energy gradient lower than 0.20 kcal mol^{-1} $Å^{-1}$ was obtained. The NOE and H bond constraints were applied with $K_{NOE} = K_{H\ bond}$ of 32 kcal mol^{-1} $Å^{-2}$. The restrained MD calculations were performed, in vacuo, warming up the structures from 0 to 300 K and coupling them to a thermal bath at 300 K. The SHAKE algorithm[42] was used to constrain all bond lengths to their equilibrium values. A time step of 1.0 fs and a residue-based cutoff value of 10 Å were used for the evaluation of nonbonded interactions, and the pair list was updated every 10 MD time steps. The PCSHIFT-RMD calculations were performed for 48 ps on each structure. They reached an equilibrium conformation in terms of total energy and RMSD values within a few picoseconds after the systems reached 300 K. The last 24 ps of each trajectory were used to generate an average structure with the CARNAL module of AMBER. Each average structure was then energy-minimized. These structures constituted the PCSHIFT-RMD family and were then used for the analysis.

RESULTS AND DISCUSSION

Computational Methods

Calculation of the Magnetic Susceptibility Anisotropy Tensor

A modified version of the program FANTASIA[15] has been developed that is able to estimate the magnetic susceptibility anisotropy in the general case of *n* paramagnetic centers (FANTASIAN). This program calculates the magnetic susceptibility anisotropy tensor by fitting the experimental NMR pseudocontact shifts with respect to a starting structure. In the case of a series of starting structures, as it occurs for a family of NMR solution structures obtained through NOE, J coupling, and hydrogen bond constraints, the program calculates the average tensor of the superimposed structures. In other

words, a tensor is obtained by fitting the experimental shifts within the laboratory coordinates of the atoms with their indetermination, provided by the scattering of the family. Of course, the program can calculate a tensor for each member of the family.

Following the procedure of FANTASIA, FANTASIAN estimates the orientation and the axial ($\Delta\chi_{ax}$) and the equatorial ($\Delta\chi_{rh}$) anisotropies of the magnetic susceptibility tensor(s), by minimizing the error function F_{err} defined as:

$$F_{err} = \sum_i [\max (|\delta_c^i - \delta_o^i| - T_i, 0)]^2 \qquad (2)$$

where i runs over all the pseudocontact shift constraints, δ_o is the measured pseudocontact shift value, T is the tolerance assigned to that shift and δ_c is the calculated pseudocontact shift according to the Equation 1. The contribution due to the nucleus i to the error function F_{err} (Eq. 2) is equal to zero when the absolute value of the difference between its experimental and calculated pseudocontact shifts is smaller than the value chosen for the tolerance T_i. The minimization of F_{err} is achieved by fitting the experimental shifts to the five parameters ($\Delta\chi_{ax}$, $\Delta\chi_{rh}$, and the three directions of the principal axes of the χ tensor), which define the χ tensor of the molecule. In this version of the FANTASIAN program, the internal coordinate system of the $\Delta\chi$ tensor is not anchored to any atomic position.

We have implemented the nonlinear SIMPLEX algorithm[43] in FANTASIAN by allowing the minimization procedure to use different sets of starting parameters. In this way the search for the absolute minimum is more efficient and possible local minima are avoided. This algorithm is able to find the best solution and its equivalent symmetric values. In the case of n protons with degenerate shifts (e.g., methyl groups), the shift values have been treated by assuming that the measured paramagnetic shift is the mean shift value of those of the n protons. In this case the contribution to F_{err} is then calculated assuming that

$$\delta_c = \sum_i \frac{\delta_c^i}{n} \qquad (3)$$

where δ_c^i is the ith degenerate proton contribution.

Inclusion of the Pseudocontact Shift Constraints in Energy Minimization and Molecular Dynamics Simulations

This approach holds for a single structure as well for a family of structures. The latter case is the most common in the determination of solution structures. The solution structure can be obtained with any distance geometry procedure. As already mentioned, we have implemented the DIANA program to include pseudocontact shifts as constraints (PSEUDIANA),[15] which is a logical step before energy minimization. Any other distance geometry approach, including pseudocontact shifts, may be used. Alternatively, pseudocontact shift constraints can be introduced at this stage on any family of structures. The SANDER module of AMBER uses, among the first-order algorithms, the steepest descent and the conjugate gradient algorithms to minimize the potential energy. These algorithms calculate the direction of the atomic displacements in order to minimize the energy terms, by applying the forces obtained from the derivation of the potential function terms. The AMBER force field[44,45] contains bond terms (bendings, stretchings, and torsions), nonbonded terms (van der Waals interactions, H bond interactions, electrostatic terms), and pseudopotential terms (NMR constraints). To include the pseudocontact shift constraints during energy minimization and molecular dynamics calculations, we have added a pseudopotential term to the AMBER potential function:

$$U_{pc} = \sum_i K_{pc}[\max (|\delta_c^i - \delta_o^i| - T_i, 0)]^2 \qquad (4)$$

where K_{pc} is a force constant and i runs over all the measured shifts. The U_{pc} energy term was added to the complete AMBER potential function. To modulate the effect of adding a new term in the AMBER force field, we included the constant K_{pc} in U_{pc}. We emphasize that it is necessary to seek homogeneity among all the potential terms.

The force associated with these pseudocontact shift constraints, defined as the first derivative of the above potential, is evaluated for all the protons for which the calculated shift (δ_c) differs more than the tolerance (T) from the experimental shift (δ_o):

$$F_{pc}^i = K_{pc}\nabla[\max (|\delta_c^i - \delta_o^i| - T_i, 0)]^2 \qquad (5)$$

where ∇ stands for first derivative operator with respect to the cartesian coordinates and δ_i^c is evaluated from Equation 1 by considering that

$$\begin{pmatrix} l_{ij} \\ m_{ij} \\ n_{ij} \end{pmatrix} = A_j \begin{pmatrix} x_i/r_i \\ y_i/r_i \\ z_i/r_i \end{pmatrix} \qquad (6)$$

where

$$r_i = \sqrt{x_i^2 + y_i^2 + z_i^2} \qquad (7)$$

and A_j is the rotation matrix which converts the laboratory cartesian coordinate system (x, y, z) in a system of axes defined by the directions of the jth magnetic susceptibility anisotropy tensor. The present program, which takes into account the pseudo-

contact shifts, is able to minimize the deviation between experimental and calculated pseudocontact shifts, accompanied by a decrease in all the other energy terms, which is comparable to that obtained in similar unconstrained calculations. Then, the AMBER force vectors were updated adding the F_{pc} values every MD time step. Values for K_{pc} in the range of 10–100 kcal mol^{-1} ppm^{-2} are reasonable for energy minimization and molecular dynamics calculations.

We refer to the program, which takes into account the pseudocontact shifts in the complete force field used for energy minimization and molecular dynamics calculations, as the PCSHIFT program. It should be considered as a subroutine of the SANDER module of AMBER 4.1. The FANTASIAN program and the PCSHIFT module are available from the corresponding author (bertini@risc1.lrm.fi.cnr.it). The PSEUDIANA program is also available provided that license for the use of the DIANA program is obtained from Kurt Wüthrich, ETH, Zürich.

Evaluation of the Pseudocontact Shifts

One problem that arises in the whole approach is the choice of the diamagnetic reference to be subtracted from the experimental chemical shifts in order to obtain pseudocontact shifts. This problem is critical when small pseudocontact shifts are used. Indeed, pseudocontact shifts may be small because the nuclei may be close to the spatial regions where the pseudocontact shifts vanish. Therefore, pseudocontact shift values equal to zero can be critical constraints.

A straightforward way is that of calculating the diamagnetic contribution to the shifts through the appropriate module of Amber 4.1 (DSHIFT). Here the errors may be as large as 0.5 ppm for CH protons, and around 1 ppm for NH's.[46] When the assignment of the same protein in a diamagnetic state is available, its shifts can be used directly as the diamagnetic contribution. Of course, if there is a structural variation between the paramagnetic molecule and the diamagnetic reference, errors in the determination of the pseudocontact shifts can be introduced.

Finally, different tolerances (Eq. 5) can be assigned to different classes of protons. In any case, an a posteriori verification is needed in order to check that there are no systematic violations of the constraints.

PCSHIFT-REM and PCSHIFT-RMD Calculations

A K_{pc} constant of 100 kcal mol^{-1} ppm^{-2} was used during the PCSHIFT-REM calculations. In the PCSHIFT-RMD calculations, the K_{pc} constant was reduced to 25 kcal mol^{-1} ppm^{-2}. During all the PCSHIFT-RMD calculations, the rotational and translational motions about the center of mass were eliminated to keep constant the relative orientation of the magnetic susceptibility anisotropy tensor coordinate system with respect to the absolute cartesian coordinate system.

TABLE I. Average χ tensor anisotropy parameters calculated over the best 17 structures of the cyanide derivative of the Met80Ala mutant of yeast cytochrome c, after PSEUDIANA (A), PCSHIFT-REM (B), and PCSHIFT-RMD (C) calculations

Module	$\Delta\chi_{ax} \pm 3\sigma$ [a] ($m^3 \times 10^{32}$)	$\Delta\chi_{rh} \pm 3\sigma$ [a] ($m^3 \times 10^{32}$)	k_z [b,c,d]	k_x [c,d,e]
A	3.318 ± 0.042	0.710 ± 0.030	0.997	0.781
B	3.307 ± 0.010	0.734 ± 0.030	0.996	0.802
C	3.303 ± 0.045	0.733 ± 0.170	0.995	0.790

[a] 3σ values were evaluated from the $\Delta\chi$ parameters calculated for each structure of the family.
[b] k_z is the direction cosine of the z axis of the χ tensor with respect to the perpendicular to the average plane formed by the four heme pyrrole nitrogens.
[c] Only two of the three independent direction cosines are reported, as the almost perfect alignment of the z axis of the χ tensor along the perpendicular to the average plane formed by the four heme pyrrole nitrogens permits the definition of the k_x and k_y directions in the xy plane using only one direction cosine.
[d] The error is much lower than 10^{-3}.
[e] k_x is the direction cosine of the x axis of the χ tensor with respect to the projection of the average Fe-pyrrole III nitrogen direction on the average plane formed by the four heme pyrrole nitrogens.

Refinement of a Solution Structure

The PCSHIFT module has been tested on the NMR solution structure family of the cyanide adduct of the oxidized Met80Ala mutant of yeast cytochrome c.[15] In this system, the methionine axial ligand has been replaced by an alanine residue and the open coordination site is occupied by cyanide. The low-spin ferric ion has a relatively large magnetic anisotropy, which produces relatively large pseudocontact shifts. The structures of the family used for testing this refinement procedure were those obtained by distance geometry calculations, with the inclusion of pseudocontact shifts through the PSEUDIANA program; these structures have already been reported in the literature.[15] The same 315 pseudocontact shifts from the previous procedure were used both for calculating the magnetic susceptibility anisotropy tensor and as structural constraints. We performed two calculations: one in which the tolerance (Eq. 5) was maintained equal to 0.05 ppm, as done previously in PSEUDIANA, with mainly the aim of checking the methodology, and the other with tolerance set to different values (0.3 ppm for the nonexchangeable protons and 0.5 ppm for the exchangeable protons) related to the reliability of the shifts of the diamagnetic reference, in order to check how the refinement is affected by this parameter.

The choice of the tolerance, as already outlined in the previous section "Evaluation of the Pseudocon-

TABLE II. Range of Values for the Initial (i) and Final (f) Energy Contributions Calculated for the Best 17 Structures of the Cyanide Adduct of the Oxidized Met80Ala Mutant of Yeast Cytochrome c*

Module	E_p(i) (kcal mol^{-1} × 10^{-3})	E_p(f) (kcal mol^{-1} × 10^{-3})	E_{NOE}(i) (kcal mol^{-1})	T_{NOE}(i) (nm^2)	E_{NOE}(f) (kcal mol^{-1})	T_{NOE}(f) (nm^2)	E_{pc}(i) (kcal mol^{-1})	T_{pc}(i) (nm^2)	E_{pc}(f) (kcal mol^{-1})	T_{pc}(f) (nm^2)
A	—	—	—	0.001 ÷ 0.002	—	0.008 ÷ 0.010	—	0.202 ÷ 0.234	—	0.022 ÷ 0.024
B	1 ÷ 1000	−1.3 ÷ −1.6	18.3 ÷ 73.3	0.008 ÷ 0.010	16.5 ÷ 22.7	0.005 ÷ 0.007	206.5 ÷ 243.1	0.021 ÷ 0.024	1100 ÷ 2200	0.110 ÷ 0.220
C	1 ÷ 1000	−1.1 ÷ −1.4	18.3 ÷ 73.3	0.008 ÷ 0.010	21.9 ÷ 29.3	0.007 ÷ 0.009	206.5 ÷ 243.1	0.021 ÷ 0.024	121.3 ÷ 134.4	0.012 ÷ 0.013
D	−1.1 ÷ −1.4	−1.8 ÷ −1.9	21.9 ÷ 29.3	0.007 ÷ 0.009	22.6 ÷ 31.6	0.007 ÷ 0.010	121.3 ÷ 134.4	0.012 ÷ 0.013	128.7 ÷ 158.2	0.013 ÷ 0.016

*(i), initial; (f), final; E_p, potential energy; E_{NOE}, NOE distance constraints; E_{pc}, pseudocontact shift constraints; T_{NOE}, NOE target function; T_{pc}, pseudocontact shift target function; A, PSEUDIANA; B, REM; C, PCSHIFT-REM; D, PCSHIFT-RMD.
Data are reported for the total potential energy (E_p), the NOE distance constraints (E_{NOE}) and the pseudocontact shift constraints (E_{pc}) penalty functions. Values for the NOE (T_{NOE}) and pseudocontact shift (T_{pc}) target functions are also reported. Data are reported for the PSEUDIANA, REM without the inclusion of pseudocontact shift constraints, the PCSHIFT-REM, and the PCSHIFT-RMD calculations.

tact Shifts," depends on the confidence of the actual data. The evaluation of this is outside the scope of this paper. It is interesting to note, however, that a larger tolerance leads to an increase in the RMSD of the family (even at the level of PSEUDIANA calculations, besides at the level of PCSHIFT-REM and PCSHIFT-RMD), although the average structure is not appreciably affected, as well as the tensor parameters. As far as the latter is concerned, the orientation remains absolutely the same in the two sets of calculations, while the size of the anisotropies varies of around 2%.

Magnetic Susceptibility Anisotropy Tensor Calculations

As in the previous calculations,[15] the $\Delta\chi$ tensor was calculated by fitting the five parameters with respect to the experimental constraints over all the 17 structures with the best agreement with the experimental NOEs and pseudocontact shifts. The tensor parameters were fitted simultaneously over all these 17 structures which had first been superimposed through a best fitting of all the heavy atom coordinates. The five independent parameters were obtained minimizing the F_{err} function (Eq. 2) by best fitting δ_c to Equation 1. In Table I the magnetic susceptibility anisotropy tensor parameters for this starting family together with those obtained after the PCSHIFT-REM and PCSHIFT-RMD calculations are given. These values are comparable with those already reported.[15] The distribution of the values of the calculated pseudocontact shift, as a function of colors, is shown in Figure 1. The figure illustrates how far from the metal center meaningful pseudocontact shift constraints can be obtained.

PCSHIFT-REM Calculations

All the 39 structures obtained by the PSEUDIANA calculations[15] were energy minimized, using the present modified version of the program AMBER, by including the 1541 available NOEs,[3] 31 H bond and 315 pseudocontact shifts as structural constraints. As previously reported, we selected the 17 structures with the lowest total target function values which constituted the PCSHIFT-REM family. In Table II the initial and final energy values for the total potential energy (E_p), the NOE and the pseudocontact shift penalty (E_{NOE} and E_{pc}, respectively) and target (T_{NOE} and T_{pc}, respectively) functions, are given. The same energy values are reported for the REM calculations performed without the inclusion of pseudocontact shift constraints. In this case the E_{pc} values are only calculated but this term was not included in the energy minimization calculations. As it can be seen from Table II, the inclusion of the pseudocontact shifts constraints minimizes the deviation between the experimental and calculated pseudocontact shifts (see the T_{pc} and E_{pc} columns in Table II) and also produces a decrease in all the other energy terms. The energy values for the final families, obtained by restrained energy minimization calculations with and without the inclusion of the pseudocontact constraints, are comparable. However, in the case of REM calculations the values of E_{pc} for the final family are always more than 10 times larger than those obtained with PCSHIFT-REM calculations. In Table III the RMSD values, with respect to the mean structure, are reported for the starting family (PSEUDIANA family) and for the family after energy minimization. The inclusion of pseudocontact shift constraints allows us to refine (see the lower RMSD value in Table III) the structures obtained with the PSEUDIANA program. Indeed, PCSHIFT-REM improves the RMSD values (Table III), while the REM calculation without the

TABLE III. RMSD Values for the Families of the Cyanide Adduct of the Oxidized Met80Ala Mutant of Cytochrome c*

Module	BB (nm)	HA (nm)
A	0.040	0.086
B	0.040	0.087
C	0.036	0.085
D	0.044	0.093

*BB, backbone; HA, heavy atoms; A, PSEUDIANA; B, REM; C, PCSHIFT-REM; D, PCSHIFT-RMD.
Values based on the averaged structure, obtained after PSEUDIANA, REM without the inclusion of pseudocontact shift constraints, PCSHIFT-REM, and PCSHIFT-RMD calculations. Values are reported for backbone and all heavy atoms.

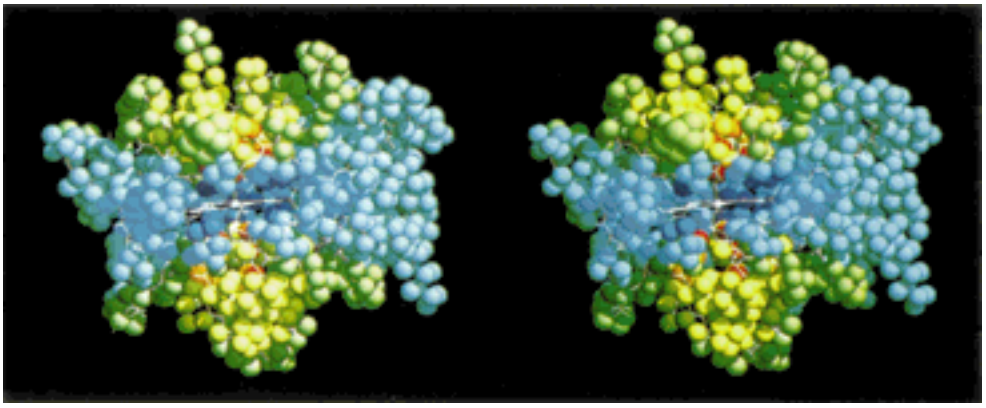

Fig. 1. Pseudocontact shifts calculated for the average minimized structure of the Met80Ala mutant of the *yeast* cytochrome c. Color codes: red $>$ 1.0 ppm, yellow $>$ 0.05 ppm, 0 ppm $<$ green $<$ 0.05 ppm; dark blue $<$ $-$1.0 ppm, blue $<$ $-$0.05 ppm, 0 ppm $>$ light blue $>$ $-$0.05 ppm. The green and light blue regions have absolute values of pseudocontact shift smaller than the tolerance.

inclusion of pseudocontact constraints does not improve the RMSD of the family (Table III). The NOE target function, despite the inclusion of new different constraints, decreased from 0.008 ÷ 0.010 nm^2 to 0.007–0.009 nm^2 after the PCSHIFT-REM calculations (see Table II). In Figure 2 the calculated pseudocontact shifts values after REM without the inclusion of pseudocontact shift constraints (A) and after PCSHIFT-REM (B) calculations versus the experimental values are plotted. It is of interest that the PCSHIFT-REM calculations produce an improvement of all the properties of the family, that is, RMSD, total energy, and agreement with the experimental constraints.

The refinement of a structure with the use of the PCSHIFT-REM (and PCSHIFT-RMD) module is completely general. A sizable improvement of the quality of the structure and of the agreement with the experimental constraints can be obtained by starting from a less refined structure, that is, one for which no PSEUDIANA calculations have been performed. This means that these calculations can also be successfully performed for the refinement of structures determined by crystallographic methods.

PCSHIFT-RMD Calculations

A 42-ps molecular dynamics trajectory was performed on the 17 structures of the PCSHIFT-REM family. The last 24 ps of the trajectory were used to generate a PCSHIFT-RMD family as described in the Materials and Methods section. In Figure 2C the calculated pseudocontact shifts values for the PCSHIFT-RMD families versus the experimental values are plotted. The PCSHIFT module maintains good agreement between calculated and experimental pseudocontact shift data, also in MD calculations (see the NOE and pseudocontact shift target functions columns in Table II). Furthermore, PCSHIFT-RMD calculations allow us to reach lower potential energy values with respect to the minimized structures (Table II, row 4). The RMSD values for the PCSHIFT-RMD family are higher than those obtained after PCSHIFT-REM calculations. While the total energy of the molecule decreases, the NOE and pseudocontact shift target functions remain the same or show a slight increase. In Figure 3 the number of NMR constraints per residue (A) and the RMSD value per residue (B), after PCSHIFT-REM and PCSHIFT-RMD calculations, are plotted. This figure shows that the higher RMSD values of some of the residues are essentially due to the lack of NMR constraints. Giving kinetic energy to the molecule produces some spreading of the structures in the family together with an increase in the target functions. Even if the effects are minor and mostly due to changes where the constraints are fewer, they could indicate that the PCSHIFT-REM family was slightly overrefined.

Concluding Remarks

Here we have reported an analysis of the use of pseudocontact shifts for the refinement of solution structures of paramagnetic metalloproteins. In a previous paper, we included the pseudocontact shifts in a distance-geometry algorithm.[15] Now, we have included them as constraints in structure calculations which also involve energy terms; and we have written a module which can be integrated into molecular mechanics and molecular dynamics calculations (PCSHIFT-REM and PCSHIFT-RMD module).

We have applied these calculations to a system, the cyanide derivative of oxidized Met80Ala cyt c, on which PSEUDIANA calculations had already been performed.[15] With these calculations, and in particular with PCSHIFT-REM, a sizable improvement of the quality of the structure is obtained. This improvement is represented by a decrease in the RMSD

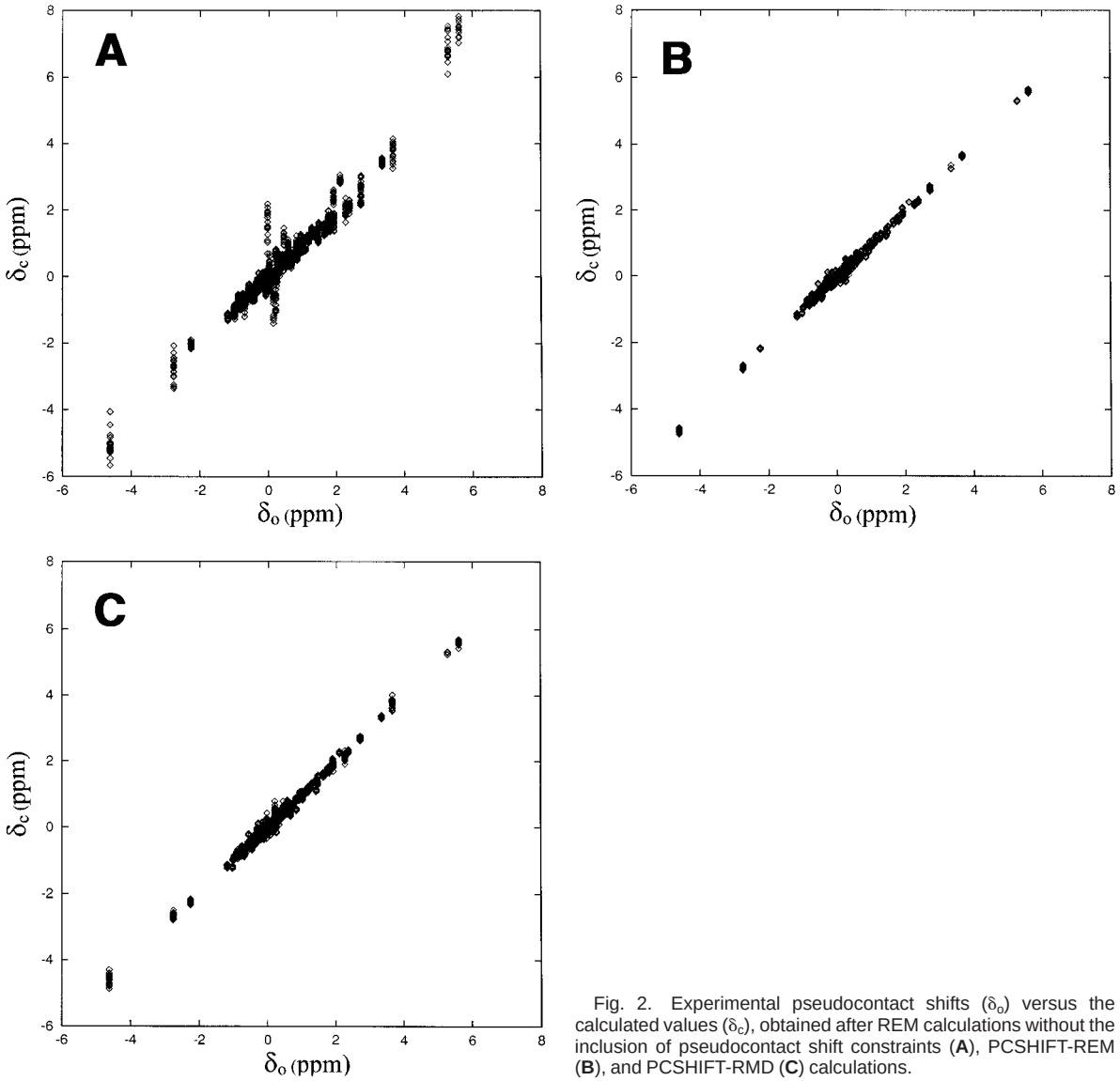

Fig. 2. Experimental pseudocontact shifts (δ_o) versus the calculated values (δ_c), obtained after REM calculations without the inclusion of pseudocontact shift constraints (**A**), PCSHIFT-REM (**B**), and PCSHIFT-RMD (**C**) calculations.

values within the family of structures and a decrease in the violations for both the NOE and the pseudocontact shift constraints. Therefore, the pseudopotential which describes the constraints due to the pseudocontact shifts is able to better define the structure.

In the case of PCSHIFT-RMD calculations, the RMSD of the family increases, even if the agreement with the experimental constraints remains essentially the same. The RMSDs increase in the protein regions where few constraints are available: the application of kinetic energy where the constraints are lacking produces a spreading due to possible multiple conformations with similar energy.

The pseudocontact shift constraints have a different nature with respect to the NOE or J constraints, as they depend on the reciprocal of the third power of the metal–proton distance. Furthermore, in most cases they are many and are more sizable in those regions of the molecule where the other conventional constraints could be less, that is, close to the metal center. These properties make pseudocontact shifts absolutely essential for the detailed description and the refinement of structures of paramagnetic molecules, up to the level of molecular dynamics calculations.

The programs we have written and their critical use represent a valid contribution to the refinement of structures of paramagnetic molecules which display magnetic anisotropy and therefore experience pseudocontact shifts.

As already noted, pseudocontact shifts are a powerful tool to link the metal ion to the protein atoms and as such provide a further step in the solution structure determination of paramagnetic metalloproteins. Furthermore they define better (smaller RMSD) the regions around the metal ion. The correct-

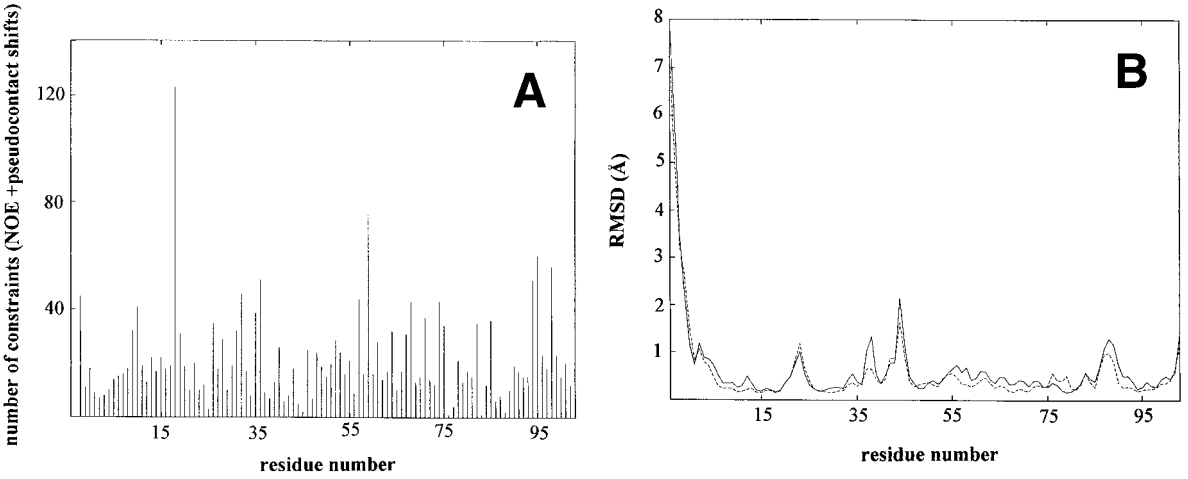

Fig. 3. (**A**) Number of NMR constraints per residue and (**B**) RMSD value per residue after PCSHIFT-REM (broken line) and PCSHIFT-RMD (solid line) calculations.

ness of the use of pseudocontact shift is shown by its consistency with NOEs, in the sense that the introduction of about 300 pseudocontact shifts does not essentially alter the sum of the quadratic deviations of about 1500 meaningful NOEs.

Acknowledgment

This work was supported by the E.U. BIOTECH program (BIO2-CT94/2052) and by CNR-Comitato Nazionale Biotecnologia e Biologia Molecolare (95.02860.CT14). A.R. acknowledges a study contract with IBM-SEMEA.

REFERENCES

1. 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.
2. 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.
3. 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.
4. 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.
5. Muskett, F.W., Kelly, G.P., Whitford, D. The solution structure of bovine ferricytochrome b/55 determined using heteronuclear NMR methods. J. Mol. Biol. 258:172–189, 1996.
6. Bentrop, D., Bertini, I., Capozzi, F., Dikiy, A., Eltis, L.D., Luchinat, C. The three dimensional structure of the reduced C77S mutant of the *Chromatium vinosum* high potential iron-sulfur protein through NMR: Comparison with the solution structure of the wild-type protein. Biochemistry 35:5928–5936, 1996.
7. Bertini, I., Luchinat, C., Rosato, A. The solution structure of paramagnetic metalloproteins. Prog. Biophys. Mol. Biol. 1996.
8. Banci, L., Bertini, I., Luchinat, C. 2D NMR spectra of paramagnetic systems. Methods Enzymol. 239:485–514, 1994.
9. Unger, S.W., Lecomte, J.T.J., La Mar, G.N. The utility of the nuclear Overhauser effect for peak assignment and structure elucidation in paramagnetic proteins. J. Magn. Reson. 64:521–526, 1985.
10. Bertini, I., Luchinat, C., Piccioli, M., Tarchi, D. COSY spectra of paramagnetic macromolecules, observability, scalar effects, cross correlation effects, relaxation allowed coherence transfer. Concepts Magn. Reson. 6:307–335, 1994.
11. Bertini, I., Dikiy, A., Luchinat, C., Piccioli, M., Tarchi, D. NOE-NOESY: A further tool in NMR of paramagnetic metalloproteins. J. Magn. Reson. Ser. B 103:278–283, 1994.
12. Bertini, I., Donaire, A., Luchinat, C., Rosato, A. Paramagnetic relaxation as a tool for solution structure determination: *Clostridium pasterianum* ferredoxin as an example. Proteins 1996.
13. Bertini, I., Donaire, A., Felli, I.C., Luchinat, C., Rosato, A. From NOESY cross-peaks to structural constraints in a paramagnetic metalloprotein. Magn. Reson. Chem. 34:948–950, 1996.
14. 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.
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. Banci, L., Pierattelli, R. 3D structure of HiPIPs in solution through NMR and molecular dynamics study. NATO ASI Ser. 457:281–296, 1995.
17. La Mar, G.N., Horrocks, W.D.W., Holm, R.H. "NMR of Paramagnetic Molecules." New York: Academic Press, 1973.
18. Bertini, I., Turano, P. The hyperfine coupling. NATO ASI Ser. 457:29–54, 1994.
19. Banci, L. Nuclear relaxation in paramagnetic metalloproteins. Biol. Magn. Reson. 12:79–111, 1993.
20. Bertini, I., Turano, P., Vila, A.J. NMR of paramagnetic metalloproteins. Chem. Rev. 93:2833–2932, 1993.
21. Banci, L., Bertini, I., Luchinat, C. "Nuclear and Electron Relaxation: The Magnetic Nucleus-Unpaired Electron Coupling in Solution." Weinheim: VCH, 1991.

22. Bertini, I., Luchinat, C. "NMR of Paramagnetic Molecules in Biological Systems." Menlo Park, CA: Benjamin/Cummings, 1986.
23. Bertini, I., Luchinat, C. NMR of paramagnetic substances. Coord. Chem. Rev. 1–300, 1996.
24. Campbell, I.D., Dobson, C.M., Williams, R.J.P., Xavier, A.V. The determination of the structure of proteins in solution: Lysozyme. Ann. N.Y. Acad. Sci. 222:163, 1973.
25. Barry, C.D., Glasel, J.A., Williams, R.J.P., Xavier, A.V. Quantitative determination of conformations of flexible molecules in solution using lanthanide ions as nuclear magnetic resonance probes: Application to adenosine-5′-monophosphate. J. Mol. Biol. 84:471–490, 1974.
26. Emerson, S.D., La Mar, G.N. Solution structural characterization of cyanometmyoglobin: resonance assignment of heme cavity residues by two-dimensional NMR. Biochemistry 29:1545–1556, 1990.
27. Gao, Y., Boyd, J., Pielak, G.J., Williams, R.J.P. Proton nuclear magnetic resonance as a probe of differences in structure between the C102T and F82S,C102T variants of iso-1-cytochrome c from the yeast *Saccharomyces cerevisiae*. Biochemistry 30:7033–7040, 1991.
28. Solomon, I. Relaxation processes in a system of two spins. Phys. Rev. 99:559–565, 1955.
29. McConnell, H.M., Robertson, R.E. Isotropic nuclear resonance shifts. J. Chem. Phys. 29:1361–1365, 1958.
30. Williams, G., Clayden, N.J., Moore, G.R., Williams, R.J.P. Comparison of the solution and crystal structures of mitochondrial cytochrome c: Analysis of paramagnetic shifts in the nuclear magnetic resonance spectrum of ferricytochrome c. J. Mol. Biol. 183:447–460, 1985.
31. Lee, L., Sykes, B.D. Use of lanthanide-induced nuclear magnetic resonance shifts for determination of protein in solution: EF calcium binding site of carp parvalbumin. Biochemistry 22:4366–4373, 1983.
32. Emerson, S.D., La Mar, G.N. 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, 1990.
33. Banci, L., Dugad, L.B., La Mar, G.N., Keating, K.A., Luchinat, C., Pierattelli, R. ^{1}H Nuclear Magnetic Resonance investigation of cobalt(II) substituted carbonic anhydrase. Biophys. J. 63:530–543, 1992.
34. Barry, C.D., Martin, D.R., Williams, R.J.P., Xavier, A.V. Quantitative determination of the conformation of cyclic 3′,5′-adenosine monophosphate in solution using lanthanide ions as nuclear magnetic resonance probes. J. Mol. Biol. 84:491–502, 1974.
35. Harper, L.V., Amann, B.T., Vinson, V.K., Berg, J.M. NMR studies of a cobalt-substituted zinc finger peptide. J. Am. Chem. Soc. 115:2577–2580, 1993.
36. Gochin, M., Roder, H. Protein structure refinement based on paramagnetic NMR shifts: Applications to wild-type and mutants forms of cytochrome c. Prot. Sci. 4:296–305, 1995.
37. Gochin, M., Roder, H. Use of pseudocontact shifts as a structural constraint for macromolecules in solution. Bull. Magn. Reson. 17:1–4, 1995.
38. Güntert, P., Wüthrich, K. Improved efficiency of protein structure calculations from NMR data using the program DIANA with redundant dihedral angle constraints. J. Biomol. NMR 1:447–456, 1991.
39. 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.
40. Weiner, S.J., Kollman, P.A., Nguyen, D.T., Case, D.A. AMBER: Assisted model building with energy refinement: A general program for modeling molecules and their interactions. J. Comp. Chem. 7:287–303, 1986.
41. Pearlman, D.A., Case, D.A., Caldwell, G.C., et al. AMBER 4.0, University of California, San Francisco, 1991.
42. van Gunsteren, W.F., Berendsen, H.J.C. Algorithms for macromolecular dynamics and constraint dynamics. Mol. Phys. 34:1311–1327, 1977.
43. Nelder, J.A., Mead, R. A simplex method for function minimization. Comput. J. 7:308–313, 1965.
44. Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S. Jr. A new force field for molecular mechanical simulation of nucleic acids and proteins. J. Am. Chem. Soc. 106:765–784, 1984.
45. Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., Kollman, P.A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J. Am. Chem. Soc. 117:5179–5197, 1995.
46. Ösapay, K., Case, D.A. A New Analysis of proton chemical shifts in proteins. J. Am. Chem. Soc. 113:9436–9444, 1991.