Journal of Biomolecular NMR, **12:** 553–557, 1998.
KLUWER/ESCOM
© 1998 *Kluwer Academic Publishers. Printed in the Netherlands.*

PSEUDYANA for NMR structure calculation of paramagnetic metalloproteins using torsion angle molecular dynamics

Lucia Banci[a], Ivano Bertini[a,*], Mauro Andrea Cremonini[c], Giovanni Gori-Savellini[a], Claudio Luchinat[b], Kurt Wüthrich[d] and Peter Güntert[d]
[a]*Department of Chemistry, University of Florence, Via G. Capponi 7, I-50121 Florence, Italy;* [b]*Department of Soil Science and Plant Nutrition, University of Florence, P.le delle Cascine 28, I-50144 Florence, Italy;* [c]*Food Science and Technology Laboratory, University of Bologna, Via Ravennate 1020, I-47023 Cesena, Italy;* [d]*Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule-Hönggerberg, CH-8093 Zürich, Switzerland*

Received 28 April 1998; Accepted 5 August 1998

Key words: paramagnetic metalloproteins, pseudocontact shifts, solution structures, structure refinement

Abstract

The program DYANA, for calculation of solution structures of biomolecules with an algorithm based on simulated annealing by torsion angle dynamics, has been supplemented with a new routine, PSEUDYANA, that enables efficient use of pseudocontact shifts as additional constraints in structure calculations of paramagnetic metalloproteins. PSEUDYANA can determine the location of the metal ion inside the protein frame and allows to define a single tensor of magnetic susceptibility from a family of conformers. As an illustration, a PSEUDYANA structure calculation is provided for a metal-undecapeptide complex, where simulated pseudocontact shifts but no NOE restraints are used as conformational constraints.

Introduction

The NMR signals of protons located near the paramagnetic centers in metalloproteins are characterized by large linewidths, which may make their assignment difficult and cause a decrease of the intensity of the ^{1}H-^{1}H NOE connectivities (Bertini and Luchinat, 1996). To obtain well refined solution structures it is therefore useful to exploit, whenever possible, the hyperfine shifts as supplementary restraints (Banci et al., 1996). In this way, the metal ion itself is connected to the protein frame. Here, a new module of the recently introduced DYANA program (Güntert et al., 1997) is presented that allows efficient use of pseudocontact shifts as constraints in NMR structure calculations. We refer to the version of the program that includes this routine as PSEUDYANA. Due to the high efficiency of the torsion angle dynamics (TAD) algorithm used (Güntert et al., 1997), PSEUDYANA can be employed for structure calculations starting from random initial conformers. The presently introduced protocol for such calculations fully exploits the potentiality of pseudocontact shifts as structural constraints. It thus extends, and provides a rationale for, earlier approaches, in which the use of pseudocontact shift constraints was combined with that of NOEs using either distance geometry (Banci et al., 1996; Banci et al., 1997a), energy minimization and molecular dynamics (Banci et al., 1997b) or torsion angle dynamics (Banci et al., 1997c; Bentrop et al., 1997; Arnesano et al., 1998) algorithms. In addition, procedures for obtaining the magnetic susceptibility tensor parameters from a family of NMR conformers are discussed.

The origin of the pseudocontact shifts

A paramagnetic molecule is characterized by the presence of unpaired electrons which bear a magnetic moment. In an external magnetic field the latter par-

*To whom correspondence should be addressed.

tially orients itself in such a way as to give a non-zero time average, i.e. an induced magnetic moment. The effect of this induced magnetic moment on nuclei is that of creating a dipolar magnetic field which adds to the external magnetic field. A nucleus will sense this new magnetic field depending on its position within the dipolar field. This picture neglects other mechanisms of interaction between the unpaired electrons and the nuclei and is referred to as dipolar or through-space interaction (McConnell and Robertson, 1958; Kurland and McGarvey, 1970). Just like any dipolar interaction between two dipoles, the orientational average of this interaction is zero. However, the electron magnetic moment is constituted by a spin and an orbital contribution. Whereas the former is isotropic, the latter is anisotropic. In the presence of sizeable orbital contributions to the electronic magnetic moment, the induced magnetic moment changes in intensity upon molecular rotation in an external magnetic field and the magnetic susceptibility tensor associated with the molecule becomes anisotropic. Under these circumstances the above dipolar energy does not average to zero and the average magnetic field that is added to the external magnetic field (expressed in terms of chemical shifts) is (McConnell and Robertson, 1958; Kurland and McGarvey, 1970; Banci et al., 1996):

$$\delta_i^{pc} = \sum_j \frac{1}{12\pi_{ij}^3}\left[\Delta\chi_{ax}^j\left(3n_{ij}^2 - 1\right) + \frac{3}{2}\Delta\chi_{rh}^j\left(l_{ij}^2 - m_{ij}^2\right)\right] \quad (1)$$

where l_{ij}, m_{ij}, and n_{ij} are the direction cosines of the position vector of atom *i* with respect to the *j*-th magnetic susceptibility tensor coordinate system, and r_{ij} is the distance between the *j*-th paramagnetic center and the proton *i*. This contribution to the chemical shift is called pseudocontact shift and is the only relevant term for the present analysis.

A minimization procedure (FANTASIAN (Banci et al., 1997b)) can be applied to determine the parameters $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$, the *l,m,n* direction cosines in Equation 1 and possibly the coordinates of the origin of the magnetic anisotropy tensor by starting from a structural model and by using Equation 1 (Williams et al., 1985; Emerson and La Mar, 1990; Veitch et al., 1990; Gao et al., 1991; Banci et al., 1992, 1995; Gochin and Roder, 1995; La Mar et al. 1995). If this procedure is applied to the results of a distance geometry calculation, we have to deal with a family of conformers, rather than with a single structure (Wüthrich, 1986). The problem of the $\Delta\chi$ parameters of a single structure with respect to that of a family of conformers will be discussed later. The magnetic susceptibility parameters may be obtained by fitting the experimental pseudocontact shifts with respect to a given structure (which could, for example, be a crystal structure or a preliminary NMR structure) (Keller and Wüthrich, 1972; Banci et al., 1992).

Pseudocontact shifts as a unique strategy to position the metal ion(s) with respect to the polypeptide atoms

The number of measureable pseudocontact shifts depends on the magnetic anisotropy of the metal ion: low spin iron (III) and high spin cobalt (II) are good probes (Wüthrich, 1970; Emerson and La Mar, 1990; Banci et al., 1992), lanthanides are even better (Lee and Sykes, 1983; Shelling et al. 1984). The number of pseudocontact shift constraints is small compared to that of conventional structural constraints from NOEs and J couplings. However, since pseudocontact shifts provide longer-range distance constraints (due to the r^{-3} dependence versus the r^{-6} dependence of NOEs), they represent a good check of the structure. They are particularly useful in the vicinity of the paramagnetic metal ion, where NOEs are usually less numerous because of increased relaxation. Pseudocontact shifts further allow to relate the position of the metal ion to those of the protein protons. Since metal ions are important parts of metalloproteins, their correct location in the protein frame is a major goal. Metal ions are included in the structure calculations as pseudoatoms with three unit vectors which define the molecular frame. A torsion angle dynamics approach capable of using pseudocontact shifts will locate the metal ion correctly.

From a family of conformers to a magnetic susceptibility tensor

Either at the beginning of the procedure or during the various steps it is necessary to extract the $\Delta\chi$ parameters and, as a consequence, the direction cosines. This is done with the already mentioned program FANTASIAN. The problem remains to define the χ tensor parameters for a family of conformers. In principle one may calculate a tensor for each member of the family by fitting the magnetic susceptibility tensor parameters ($\Delta\chi_{ax}$, $\Delta\chi_{rh}$, *l*, *m*, *n*) for a fixed

metal position. Alternatively, the coordinates of the paramagnetic center may also be treated as variable parameters. Another possibility is to superimpose the conformers and to calculate a single set of tensor parameters which best fits simultaneously all conformers. An average tensor during PSEUDYANA calculations has been used.

The PSEUDYANA protocol

When the anisotropy tensor is known, pseudocontact shift constraints can be used together with NOE constraints to calculate the solution structure. PSEUDYANA allows to include pseudocontact shifts as constraints into the new DYANA package (Güntert et al., 1997). DYANA uses simulated annealing based on molecular dynamics in torsion angle space (Bae and Haug, 1987; Jain et al., 1997; Stein et al., 1997), which makes it a powerful method to search conformation space and to handle the problem of local minima. The program uses a fast recursive algorithm for solving the dynamical equations of motion (Jain et al., 1997). To include pseudocontact shift constraints into simulated annealing torsion angle dynamics, a supplementary pseudocontact shift term (t^{pc}) was added to the standard DYANA target function:

$$t^{pc} = K \sum_i \left[\max \left(|\delta^{pc}_{i_{calc}} - \delta^{pc}_{i_{obs}}| - T_i, 0 \right) \right]^2 \quad (2)$$

where $\delta^{pc}_{i_{calc}}$ and $\delta^{pc}_{i_{obs}}$ are, respectively, the experimental and the calculated (Equation 1) pseudocontact shift values of proton i, and T_i is the tolerance assigned to that proton. The routine for calculation of the derivative of the pseudocontact shift target function can be obtained from the authors as Supplementary Material.

Different tolerance values T for different protons can be set in PSEUDYANA. This is particularly useful when the diamagnetic reference shift is not experimentally available but is estimated on the basis of a preliminary structure. In this condition, it was found reasonable to use $T = 1.0$ ppm for amide protons and $T = 0.5$ ppm for all the other protons (Ösapay and Case, 1991). It could also be useful to assign a tolerance proportional to the absolute pseudocontact shift value. Indeed, protons experiencing larger pseudocontact shifts may be sensitive to local dynamics and to electron delocalization (breakdown of the point-dipole approximation). Furthermore, as the pseudocontact shifts depend on the reciprocal of the third power of the distance between the proton and the paramagnetic center, if the same tolerance value is used for all the pseudocontact shifts, protons near to the paramagnetic center are allowed to experience smaller movements than protons far from the metal ion. Indeed, a small movement of a proton near the metal ion sizeably alters the pseudocontact shift value of that proton, while the same movement of a proton far from the metal ion produces a much smaller change of the pseudocontact shift value for that atom. Using a tolerance proportional to the absolute value of the pseudocontact shifts, as described above, one avoids possible overrefinement of protons close to the paramagnetic center, which in general experience larger pseudocontact shift values than protons far from the paramagnetic center.

The metal ion and the relative anisotropy tensor are represented within the DYANA framework by a special residue that is connected to the end of the polypeptide chain through linker residues. Each of these has several torsional degrees of freedom and consists exclusively of pseudoatoms, which renders them highly flexible and 'invisible' to the steric repulsion. Linker residues are used in DYANA for the treatment of multicomponent systems in order to formally preserve the tree structure of dihedral angles (Güntert et al., 1997). The linker residues allow the metal ion to position itself and the anisotropy tensor to orient itself freely under the influence of the pseudocontact shifts. Several serially connected stretches of linker residues can be used if more than one metal ion is present. In this way the previous use of modified residues to which the metal ion is attached rigidly (Arseniev et al., 1988; Banci et al., 1996) is avoided. Only loose distance constraints to the donor atoms (some or all of which may be known on the basis of biochemical or relaxation data) are used to initially place the metal ion. The structure calculation then relies on the potential given by the set of pseudocontact shifts to correctly locate the metal within the coordination core. Moreover, the metal ion and the relative anisotropy tensor are free to orient themselves so as to give the lowest value of the target function, making it unnecessary to input the direction cosines of the principal axes of the magnetization into the PSEUDYANA calculations.

A simple test of PSEUDYANA

Using exclusively pseudocontact shifts for calculating a protein structure, without adding any distance constraints coming from NOE experiments, is highly unrealistic. However, it may be interesting to simu-

late the case of a system in which only pseudocontact shifts, a rough estimate of the metal position and the tensor are known.

The structure chosen for the test comprised a small 11-residue fragment taken from the NMR solution structure of the fragment TR1C of Ce_2^{3+}-calmodulin (Bentrop et al., 1997). Residues 20–31, forming the first EF site of calmodulin, were taken from the average structure of the family of 23 conformers representing the protein in solution. The position of the metal and the orientation of the tensor were the ones reported for Ce_2^{3+}-calmodulin (Bentrop et al., 1997); these parameters were used for calculating a set of 82 pseudocontact shifts, corresponding to 82 single protons. In addition, the metal was constrained to be in the 2-3 Å range from the following atoms: OD2-Asp^{20}, OD2-Asp^{22}, OD2-Asp^{24}, O-Thr^{28} and OE2-Glu^{31}. During the calculations, zero tolerance was used for pseudocontact shifts and an overall weight of $K = 1.0$ Å^2 ppm^{-2} was given to the pseudocontact contribution.

Two calculations were made for comparison purposes, starting from 430 randomly generated structures and using only pseudocontact shifts as structural constraints. In both cases, initial estimates of the $\Delta\chi$ parameters equal to those used to generate the pseudocontact shifts were given. The first run used the standard simulated annealing protocol of the program DYANA (Güntert et al., 1997) with 10 000 torsion angle dynamic steps. To conformers with target function values between 2 and 10 Å^2 at this stage of calculation, the same protocol was applied again with the initial high temperature value reduced by a factor of 160. Starting from 430 randomized structures PSEUDYANA yielded 11 conformers with a target function value lower than 2 Å^2, which we used as the cutoff for accepting a conformer. The number of accepted structures is rather low compared to that obtained in a typical NOE-based structure determination (Güntert et al. 1997). This is, however, not surprising because the information content of the pseudocontact shift constraints is both lower and more difficult to exploit than that of an extensive set of NOE upper distance bonds. The average error in pseudocontact shifts is about 0.05 ppm. This family of structures showed an average rmsd from the mean structure of 0.47 Å, 0.74 Å, and 0.15 Å for the backbone, for all heavy atoms, and for the metal ion, respectively. The rmsd among the dummy atoms representing the metal was 0.20 Å, which is another indication of the correct functioning of the routine. The average structure obtained from this family has an rmsd of 0.63 Å for the backbone, of 0.90 Å for the heavy atoms and of 0.20 Å for the metal ion, with respect to the structure used to generate the constraints. The $\Delta\chi$ values recalculated on the final family were within $\pm$ 5% from the initial values, thus providing an estimate of the accuracy of the obtained tensor parameters. Calculations were repeated using initial $\Delta\chi$ values differing by 50% from the correct ones, and no significant worsening of the convergence occurred. Of course, further cycles were needed to generate a new family with reestimated $\Delta\chi$ values.

The second calculation was done using the variable target function method (Güntert and Wüthrich, 1991), that is also contained in DYANA, with 1000 minimization steps at each intermediate minimization level, followed by 3000 steps at the final level including all constraints, thus resembling the conditions used for producing the solution structure of M80A cyano cytochrome c with the previous PSEUDIANA routine (Banci et al., 1996). In contrast to the calculation using torsion angle dynamics, this protocol did not yield any structure with target function value lower than 10 Å^2, which underlines the power of our new torsion angle dynamics approach.

Conclusions and outlook

The presently described extension of the program package DYANA for the use of pseudocontact shifts as supplementary constraints in NMR structure calculations of paramagnetic metalloproteins is a powerful novel approach for this type of investigation. Compared to previous work based on the program DIANA, PSEUDYANA has the important advantage that pseudocontact shifts can be included as supplementary constraints into the input from the very start of the NMR structure calculations, rather than only during a final stage of structure refinement, provided that estimates of $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ in Equation 1 are available. The accuracy of the starting estimates of the $\Delta\chi$ values was not crucial for a correct behavior of the procedure. The efficient functioning of the PSEUDYANA approach is documented with the test calculation presented in this paper as well as by practical applications with several structure determinations of paramagnetic metalloproteins (Banci et al., 1997c; Bentrop et al., 1997; Arnesano et al. 1998).

References

Abe, H., Braun, W., Noguti, T. and Go, N. (1984) *Comput. Chem.* **8**, 239–247.

Arnesano, F., Banci, L., Bertini, I. and Felli, I.C. (1998) *Biochemistry*, **37**, 173–184.

Arseniev, A., Schultze, P., Wörgötter, E., Braun, W., Wagner, G., Vasak, M., Kägi, J.H. and Wüthrich, K. (1988) *J. Mol. Biol.*, **201**, 637–657.

Bae, D.S. and Haug, E.J. (1987) *Mech. Struct. Mech.*, **15**, 359–382.

Banci, L., Dugad, L.B., La Mar, G.N., Keating, K.A., Luchinat, C. and Pierattelli, R. (1992) *Biophys. J.*, **63**, 530–543.

Banci, L., Bertini, I., Pierattelli, R., Tien, M. and Vila, A.J. (1995) *J. Am. Chem. Soc.*, **117**, 8659–8667.

Banci, L., Bertini, I., Bren, K.L., Cremonini, M.A., Gray, H.B., Luchinat, C. and Turano, P. (1996) *JBIC*, **1**, 117–126.

Banci, L., Bertini, I., Bren, K.L., Gray, H.B., Sompornpisut, P. and Turano, P. (1997a) *Biochemistry*, **36**, 8992–9001.

Banci, L., Bertini, I., Gori Savellini, G., Romagnoli, A., Turano, P., Cremonini, M.A., Luchinat, C. and Gray, H.B. (1997b) *Proteins Struct. Funct. Genet.*, **29**, 68–76.

Banci, L., Bertini, I., Gray, H.B., Luchinat, C., Reddig, T., Rosato, A. and Turano, P. (1997c) *Biochemistry*, **36**, 9867–9877.

Bentrop, D., Bertini, I., Cremonini, M.A., Forsén, S., Luchinat, C. and Malmendal, A. (1997) *Biochemistry*, **36**, 11605–11618.

Bertini, I. and Luchinat, C. (1996) *NMR of paramagnetic substances*, Coord. Chem. Rev. Vol. 150, Elsevier, Amsterdam, The Netherlands.

Emerson, S.D. and La Mar, G.N. (1990) *Biochemistry*, **29**, 1556–1566.

Gao, Y., Boyd, J., Pielak, G.J. and Williams, R.J.P. (1991) *Biochemistry*, **30**, 1928–1934.

Gochin, M. and Roder, H. (1995) *Protein Sci.*, **4**, 296–305.

Güntert, P., Mumenthaler, C. and Wüthrich, K. (1997) *J. Mol. Biol.*, **273**, 283–298.

Güntert, P. and Wüthrich, K. (1991) *J. Biomol. NMR*, **1**, 447–456.

Jain, A., Vaidehi, N. and Rodriguez, G. (1997) *J. Comput. Phys.*, **106**, 258–268.

Keller, R.M. and Wüthrich, K. (1972) *Biochim. Biophys. Acta*, **285**, 326–336.

Kurland, R.J. and McGarvey, B.R. (1970) *J. Magn. Reson.*, **2**, 286–301.

La Mar, G.N., Chen, Z.G., Vyas, K. and McPherson, A.D. (1995) *J. Am. Chem. Soc.*, **117**, 411–419.

Lee, L. and Sykes, B.D. (1983) *Biochemistry*, **22**, 4366–4373.

McConnell, H.M. and Robertson, R.E. (1958) *J. Chem. Phys.*, **29**, 1361–1365.

Ösapay, K. and Case, D.A. (1991) *J. Am. Chem. Soc.*, **113**, 9436–9444.

Shelling, J.G., Bjorson, M.E., Hodges, R.S., Taneja, A.K. and Sykes, B.D. (1984) *J. Magn. Reson.*, **57**, 99–114.

Stein, E.G., Rice, L.M. and Brünger, A.T. (1997) *J. Magn. Reson.*, **124**, 154–164.

Veitch, N.C., Whitford, D. and Williams, R.J.P. (1990) *FEBS Lett.*, **269**, 297–304.

Williams, G., Clayden, N.J., Moore, G.R. and Williams, R.J.P. (1985) *J. Mol. Biol.*, **183**, 447–460.

Wüthrich, K. (1970) *Struct. Bonding*, **8**, 53–121.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY, U.S.A.