JOURNAL OF MAGNETIC RESONANCE **131,** 159–162 (1998)
ARTICLE NO. MN971345

Direct Structure Refinement against Residual Dipolar Couplings in the Presence of Rhombicity of Unknown Magnitude

G. Marius Clore,*[,1] Angela M. Gronenborn,* and Nico Tjandra†

*Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Building 5, Bethesda, Maryland 20892-0520; and †Laboratory of Biophysical Chemistry, National Heart, Lung, and Blood Institute, National Institutes of Health, Building 3, Bethesda, Maryland 20892-0320

Received September 30, 1997; revised November 25, 1997

Residual dipolar couplings arising from small degrees of alignment of molecules in a magnetic field provide unique long-range structural information. The potential of this approach for structure refinement has recently been demonstrated for a protein–DNA complex in which the magnetic susceptibility tensor was axially symmetric. For most macromolecules and macromolecular complexes, however, axial symmetry cannot be assumed. Moreover, the presence of significant rhombicity will clearly affect the accuracy of the resulting coordinates. In this Communication we present a simple calculational strategy that makes use of simulated annealing refinement against the residual dipolar couplings in combination with a grid search, to simultaneously refine the structures and ascertain the magnitude of the axial and rhombic components of the tensor.

Key Words: **residual dipolar couplings; rhombicity; solution NMR structure determination; long-range restraints; simulated annealing.**

Residual dipolar couplings (*1–17*), induced by alignment of molecules with the magnetic field, can potentially provide important long-range restraints for structure determination. The long-range information content of the residual dipolar couplings is quite different from the short-range nature of the restraints derived from NOEs, *J* coupling constants, and chemical shifts, and can significantly improve the quality of macromolecular NMR structures (*16*). For example, in the case of a complex of the transcription factor GATA-1 with a 16-base-pair oligonucleotide, the addition of only 90 residual dipolar coupling restraints to the ~1500 NOE and ~300 torsion angle restraints resulted in a substantial improvement in the quality of the protein backbone, as judged by an approximately twofold reduction in the number of residues lying outside the most favored region of the Ramachadran ϕ,ψ plot (*16*). When residual dipolar couplings are induced by very small degrees of alignment arising from anisotropy of the molecular magnetic susceptibility (*6–16*), the use of residual dipolar coupling restraints for refinement of protein–DNA complexes, as opposed to proteins alone, is simplified as the magnetic susceptibility tensor is dominated by the DNA bases which are stacked on top of one another, resulting in axial symmetry (*16*). In the majority of cases, however, axial symmetry cannot be assumed *a priori.* This is particularly so when a dilute aqueous liquid crystalline medium of oriented bicelles is used to produce moderate degrees of alignment of biomolecules in a magnetic field while retaining the sensitivity, simplicity, and resolution of spectra recorded in an isotropic medium (*17*). Since the assumption of axial symmetry in the presence of significant rhombicity will clearly affect the accuracy of the resulting coordinates, it is important to devise a simple calculational strategy for determining the degree of rhombicity from the experimental residual dipolar coupling data without reference to a previously known structure.

The expression for the residual dipolar coupling $\delta(\theta,\phi)$ between two directly coupled nuclei can be simplified to the form (*7, 12*)

$$\delta(\theta,\phi) = D_a(3\cos^2\theta - 1) + \tfrac{3}{2} D_r(\sin^2\theta \cos 2\phi), \quad [1]$$

where D_a and D_r are the axial and rhombic components of the tensor given by $\frac{1}{3}[D_{zz} - (D_{xx} + D_{yy})/2]$ and $\frac{1}{3}(D_{xx} - D_{yy})$, respectively; θ is the angle between the interatomic vector and the z axis of the tensor; and ϕ is the angle which describes the position of the projection of the interatomic vector on the x–y plane, relative to the x axis. The tensor **D** may be the magnetic susceptibility tensor for molecules aligned in a magnetic field, the molecular alignment tensor for molecules aligned in anisotropic media such as liquid crystals, the electric field tensor for molecules aligned by an electric field, or the optical absorption tensor for molecules aligned by polarized light (*1–17*). Note that the terms D_a and D_r subsume various constants, including the gyromagnetic ratios of the two nuclei, the distance between the two nuclei, the generalized order parameter $\boldsymbol{S}$ for internal motion of the internuclear vector, the magnetic field strength, and the

[1] To whom correspondence should be addressed.

1090-7807/98 $25.00

medium permeability. (It is worth pointing out that since D_a and D_r scale with $\boldsymbol{S}$ and not $\boldsymbol{S}^2$, the assumption of a uniform $\boldsymbol{S}$ value introduces a negligible error of at most a few percent in the dipolar coupling providing $\boldsymbol{S}^2 \geq 0.6$, particularly when one considers that $\boldsymbol{S}^2$ values in structured regions of a protein typically fall in the range 0.85 ± 0.05 (*16*)).

The geometric content of the residual dipolar couplings can be incorporated into the simulated annealing protocol (*18*) used for structure determination by minimizing the term E_{dipolar} (*16*):

$$E_{\text{dipolar}} = k_{\text{dipolar}}(\delta_{\text{calc}} - \delta_{\text{obs}})^2, \qquad [2]$$

where k_{dipolar} is a force constant, and δ_{calc} and δ_{obs} are the observed and calculated values of the residual dipolar couplings, respectively. E_{dipolar} is evaluated by calculating the θ and ϕ angles between the appropriate bond vectors (e.g., N–H, Cα–H, or Cα–C′) and an external arbitrary axis system. As the orientation of the axis system is not known *a priori* it is allowed to float during the simulated annealing calculations. For the purposes of the calculations, the external axis is represented by an artificial tetraatomic molecule comprising four atoms, X, Y, Z, and O, with three mutually perpendicular bonds, X–O, Y–O, and Z–O, representing the x, y, and z axes of the tensor, respectively. The final value of the force constant k_{dipolar} is chosen such that the agreement between observed and calculated values of the residual dipolar couplings is approximately equal to the experimental error.

In order to make use of Eqs. [1] and [2] for structure refinement, the values of D_a and the rhombicity R (defined as D_r/D_a) must be determined directly from the experimental data. The theoretical dependence of δ on the angles θ and ϕ for values of D_a and R of 1 Hz and 0.4, respectively, is illustrated in Fig. 1A. Figure 1B shows the values of δ (with an added random error of ± 0.1 Hz) as a function of the angle θ calculated for the N–H vectors of the 259-residue N-terminal domain of Enzyme I (EIN) (*19*) using the same values of D_a and R with the axis system of the tensor arbitrarily chosen to be represented by the coordinate molecular frame. A ribbon diagram of EIN and the axes of the tensor are depicted in Fig. 1C. The minimum value of the residual dipolar coupling, δ_{min}, occurs at $\theta = \phi = 90°$, such that D_a is given by

$$D_a = -\delta_{\text{min}}/(1 + 1.5R). \qquad [3]$$

Experimentally, a reliable value of δ_{min} is obtained by taking the average of the smallest residual dipolar couplings such that the standard deviation of the estimated δ_{min} value is equal to the measurement error. The maximum value of the residual dipolar coupling, δ_{max}, which occurs at $\theta = 0°$, is given by $2D_a$. Unfortunately, a reliable estimate of δ_{max} is more difficult to obtain from the experimental data, since

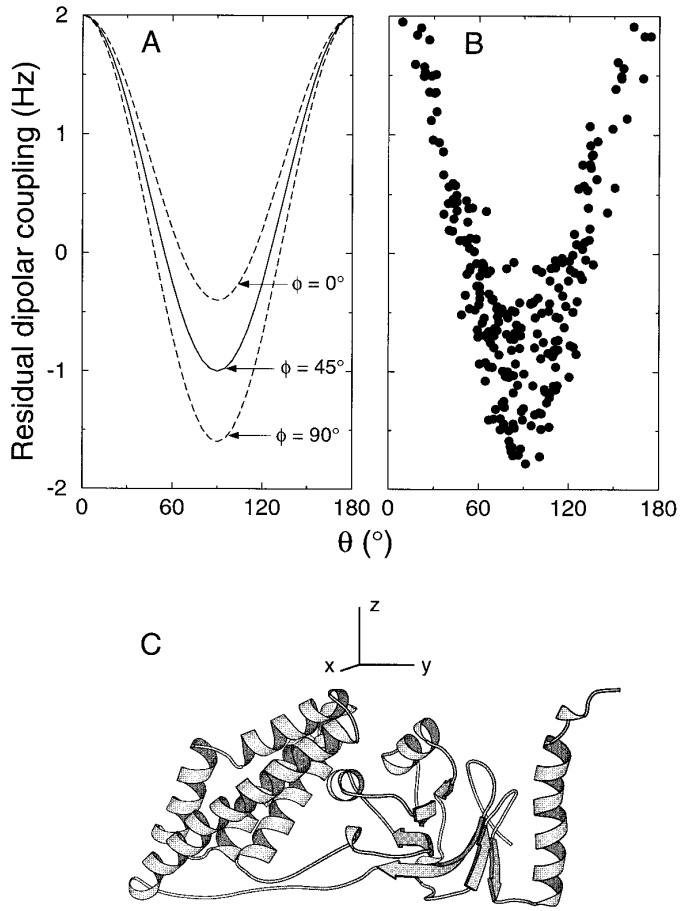

FIG. 1. Effect of the rhombicity R of the magnetic susceptibility tensor on residual dipolar couplings. (A) Theoretical dependence on the angles θ and ϕ. The solid line corresponds to the fully symmetric case and the dashed lines correspond to the fully asymmetric case for a value of 1 Hz for D_a and a rhombicity ($R = D_r/D_a$) of 0.4. (B) Distribution of the residual ^{1}H–^{15}N dipolar couplings as a function of the angle θ calculated for the N-terminal domain of Enzyme I (259 residues) with a random error of ± 0.1 Hz, and values of 1 Hz and 0.4 for D_a and R, respectively. (C) Ribbon diagram of the N-terminal domain of Enzyme I (*19*) together with the coordinate frame of the magnetic susceptibility tensor employed to calculate the ^{1}H–^{15}N residual dipolar couplings in (B). The random error of ± 0.1 Hz in the model dipolar couplings was chosen to represent typical experimental error in the residual dipolar couplings extracted from a series of J-modulated ^{1}H–^{15}N HSQC spectra (*15, 16*).

the probability of finding a bond vector that makes an angle θ to the z axis of the tensor is proportional to $\sin\theta$, and hence the probability of finding a bond vector with $\theta \sim 0°$ is very low (*16*). Consequently, the value of δ_{max} and hence the value of D_a can be underestimated by 15–20%. Nevertheless, the observed value of δ_{max} can still be used to obtain an upper limit for the value of R given by $[-2\delta_{\text{min}}(\text{obs})/\delta_{\text{max}}(\text{obs}) - 1]/1.5$.

Since δ_{min} can be accurately determined experimentally (for $R < 0.6$), but D_a cannot be obtained independently of R (unless an accurate value of δ_{max} is known), the strategy

we use involves calculating a series of structure ensembles for different estimates of R. (Note that the rhombicity reaches a maximum value of $\frac{2}{3}$ when $D_{zz} = -D_{yy} \propto 1$ and $D_{xx} = 0$; at this point the z and y axes are interchangeable so that the probability of finding a N–H vector perpendicular to the z axis is the same as finding one parallel to the z axis.) We illustrate this approach with model calculations on EIN. The restraints employed comprise the experimental NOE (3048) and torsion angle (543) restraints (*19*), in addition to three different sets of model $^1H-^{15}N$ dipolar coupling constant restraints. The latter were calculated for rhombicities of 0.1, 0.2, and 0.4 from the original NMR structure (*19*) with an added random error of ±0.1 Hz in the residual dipolar couplings and a value of 1 Hz for D_a. Ten structures were calculated by simulated annealing (*18*) for each estimated value of R which was stepped up in increments of 0.1 unit and the results are summarized in Fig. 2. It should be noted that since the orientation of the tensor with respect to the protein is not known *a priori* it is essential to slowly increase the value of the force constant $k_{dipolar}$ during the course of the simulated annealing calculation. (We used an initial value of 0.001 kcal · Hz^{-2} and a final value of 20 kcal · Hz^{-2} for $k_{dipolar}$.) In the case of all three sets of calculations the dependence of the rms difference between target and calculated dipolar couplings on the estimated value of R (R_{est}) shows a minimum when R_{est} is approximately equal to the target value of R (R_{target}). The same dependence on R_{est} is also observed for the total energy of the target function, reflecting not only the agreement between target and calculated dipolar couplings but also small changes in the agreement between target and calculated values of the other terms in the target function (i.e., experimental distance and torsion angle restraints, covalent geometry, and nonbonded contacts). The precision of the backbone atomic coordinates, for this particular set of model calculations, is independent of R_{est} and has a value of ~0.70 ± 0.05 Å. The accuracy of the coordinates (defined as the rms difference between the mean coordinate positions calculated for each ensemble and the target structure which was used to generate the model dipolar coupling data), however, does show a small dependence on R_{est} and is highest when $R_{est} \sim R_{target}$. Thus, for example, in the calculations carried out with residual dipolar coupling constants calculated for $R_{target} = 0.4$, the accuracy of the mean coordinates for $R_{est} = 0$ is 0.61 Å but reaches a minimum of 0.4 Å for $R_{est} \sim 0.4$. The minimum, however, is shallow, so that errors of ±0.2 in R_{est} will have little impact on coordinate accuracy.

In conclusion, we have shown that providing a sufficient number of NOE restraints are available to generate a well-defined polypeptide fold in the absence of residual dipolar coupling restraints (*20*), it is possible to use a grid search approach toward dipolar coupling constant refinement that permits the values of D_a and R to be reliably obtained. Thus, we anticipate, particularly with the advent of dilute aqueous liquid crystalline media to induce moderate degrees of alignment of macromolecules while preserving resolution and sensitivity (*17*), that dipolar couplings can be employed for macromolecular structure refinement in a relatively straightforward manner, thereby providing unique long-range restraints that promise to significantly increase the attainable accuracy of biomolecular NMR structures.

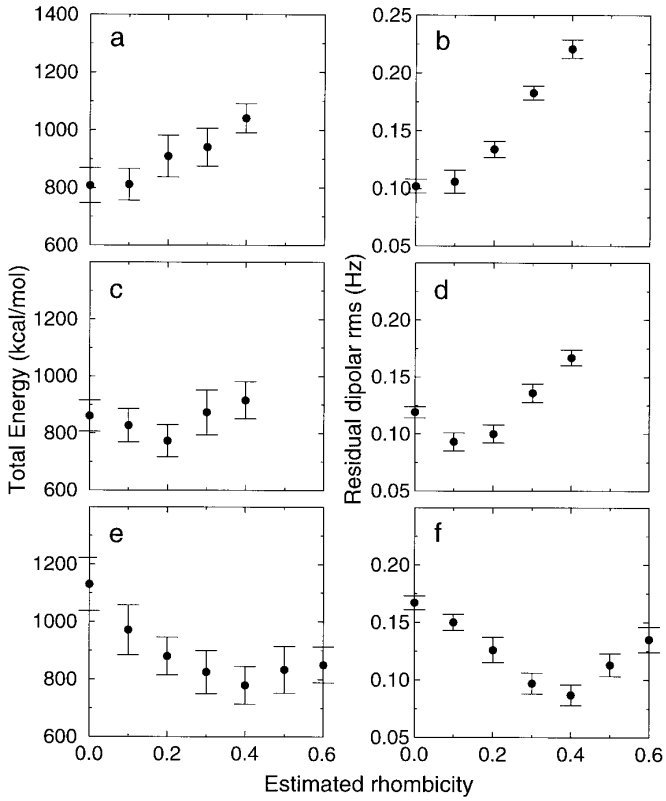

FIG. 2. Dependence of the total energy of the target function and the rms difference between target and calculated values of the residual dipolar couplings on the estimated rhombicity (R_{est}) used in the simulated annealing calculations (*18*) upon refinement of the N-terminal domain of Enzyme I against model $^1H-^{15}N$ residual dipolar coupling data with rhombicities of 0.1 (a and b), 0.2 (b and c), and 0.4 (d and e). All the model $^1H-^{15}N$ residual dipolar coupling data were calculated with a random error of ±0.1 Hz and a value of 1 Hz for D_a. The solid circles and vertical bars represent the mean values and standard deviations for the 10 simulated annealing structures calculated for each value of R_{est}. The target function includes terms for 3048 NOE and 543 torsion angle restraints (*19*), covalent geometry, and nonbonded contacts, in addition to the 258 N–H residual dipolar coupling restraints. The structures were calculated by simulated annealing (*18*) using the programs XPLOR (*21*) or CNS (*22*) modified to incorporate a pseudo-potential for the residual dipolar coupling restraints.

ACKNOWLEDGMENTS

We thank Ad Bax and Dan Garrett for useful discussions. The source code for dipolar coupling constant refinement can be obtained upon request from G.M.C. This work was supported by the AIDS Targeted Antiviral Program of the Office of the Director of the National Institutes of Health (to G.M.C. and A.M.G.).

REFERENCES

1. L. C. Snyder, Analysis of nuclear magnetic resonance spectra of molecules in liquid-crystal solvents, *J. Chem. Phys.* **43,** 4041–4050 (1965).
2. R. E. J. Sears and E. L. Hahn, Upper limits to electric-field-induced nuclear magnetic dipole–dipole couplings in polar liquids, *J. Chem. Phys.* **45,** 2753–2769 (1966).
3. J. D. Macomber, N. S. Ham, and J. S. Waugh, Upper limit to the electric-field effect on the NMR spectrum of nitromethane, *J. Chem. Phys.* **46,** 2855–2856 (1967).
4. A. D. Buckingham and K. A. McLaughlan, High resolution nuclear magnetic resonance in partially oriented molecules, *Prog. NMR Spectrosc.* **2,** 63–109 (1967).
5. J. W. Emsley and J. C. Lindon, ''NMR Spectroscopy Using Liquid Crystal Solvents,'' Pergamon Press, New York (1975).
6. J. Lohman and C. MacLean, Magnetic field induced alignment effects in ^{2}H NMR spectra, *Chem. Phys. Lett* **58,** 483–486 (1978).
7. C. Gayathri, A. A. Bothner-By, P. C. M. van Zijl, and C. MacLean, Dipolar magnetic field effects in NMR spectra of liquids, *Chem. Phys. Lett.* **87,** 192–196 (1982).
8. P. C. M. van Zijl, B. H. Ruessink, J. Bulthuis, and C. MacLean, NMR of partially aligned liquids: Magnetic susceptibility anisotropies and dielectric properties, *Acc. Chem. Res.* **17,** 172–180 (1984).
9. A. A. Bothner-By, C. Gayathri, P. C. M. van Zijl, C. MacLean, J.-J. Lai, and K. M. Smith, High field orientation effects in the high resolution proton NMR spectra of diverse porphyrins, *Magn. Reson. Chem.* **23,** 935–938 (1985).
10. E. W. Bastiaan, C. MacLean, P. C. M. Van Zijl, and A. A. Bothner-By, High resolution NMR of liquids and gases: Effects of magnetic-field-induced molecular alignment, *Annu. Rep. NMR Spectrosc.* **9,** 35–77 (1987).
11. K. L. Facchine, S. W. Staley, P. C. M. van Zijl, P. K. Mishra, and A. A. Bothner-By, Torsional angles in vinyl arenes determined by high field NMR spectroscopy, *J. Am. Chem. Soc.* **110,** 4900–4905 (1989).
12. A. A. Bothner-By, Magnetic field induced alignment of molecules, *in* ''Encyclopedia of Nuclear Magnetic Resonance'' (D. M. Grant and R. K. Harris, Eds.), pp. 2932–2938, Wiley, Chichester (1995).
13. J. R. Tolman, J. M. Flanagan, M. A. Kennedy, and J. H. Prestegard, Nuclear magnetic dipolar interactions in field-oriented proteins: Information for structure determination in solution, *Proc. Natl. Acad. Sci. USA* **92,** 9279–9283 (1995).
14. H. C. King, K. Y. Wang, I. Goljer, and P. H. Bolton, Magnetic alignment of duplex and quadruplex DNAs, *J. Magn. Reson. B* **109,** 323–325 (1995).
15. N. Tjandra, S. Grzesiek, and A. Bax, Magnetic field dependence of nitrogen-proton *J* splittings in ^{15}N-enriched human ubiquitin resulting from relxation interference and residual dipolar coupling, *J. Am. Chem. Soc.* **118,** 6264–6272 (1996).
16. N. Tjandra, J. G. Omichinski, A. M. Gronenborn, G. M. Clore, and A. Bax, Use of dipolar ^{1}H–^{15}N and ^{1}H–^{13}C couplings in the structure determination of magnetically oriented macromolecules in solution, *Nature Struct. Biol.* **4,** 732–728 (1997).
17. N. Tjandra and A. Bax, Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science* **278,** 1111–1114 (1997).
18. M. Nilges, A. M. Gronenborn, A. T. Brünger, and G. M. Clore, Determination of three-dimensional structures of proteins by simulated annealing with interproton distance restraints: Application to crambin, potato carboxypeptidase inhibitor and barley serine proteinase inhibitor 2, *Protein Eng.* **2,** 27–38 (1988).
19. D. S. Garrett, Y.-J. Seok, D.-I. Liao, A. Peterkofsky, A. M. Gronenborn, and G. M. Clore, Solution structure of the 30 kDa N-terminal domain of enzyme I of the *Escherichia coli* phospho*enol*pyruvate:sugar phosphotransferase system by multidimensional NMR. *Biochemistry* **36,** 2517–2530 (1996).
20. G. M. Core, M. A. Robien, and A. M. Gronenborn, Exploring the limits of precision and accuracy of protein structures determined by nuclear magnetic resonance spectroscopy, *J. Mol. Biol.* **231,** 81–102 (1993).
21. A. T. Brünger, ''XPLOR Manual Version 3.1,'' Yale University, New Haven, CT (1993).
22. A. T. Brünger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kinstleve, J.-S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, G. Warren, Crystallography and NMR System (CNS): a new software suite for macromolecular structure determination. *Acta Crystallog. Series D,* in press (1998).