J|A|C|S
COMMUNICATIONS

Published on Web 00/00/0000

Tracking Alignment from the Moment of Inertia Tensor (TRAMITE) of Biomolecules in Neutral Dilute Liquid Crystal Solutions

Hugo F. Azurmendi† and C. Allen Bush*

Department of Chemistry and Biochemistry, University of Maryland-Baltimore County, Baltimore, Maryland 21250

Received November 12, 2001

The use of dilute liquid crystals to induce a weak degree of molecular alignment with respect to a static magnetic field has been extended since its introduction for proteins by Tjandra and Bax in 1997.[1] The induced anisotropy permits measurement in high-resolution NMR spectra of residual dipolar couplings providing long-range structural information on the relative orientation of different parts of a biological molecule in aqueous solution.[2]

The alignment effect is thought to be the result of steric and electrostatic interactions.[3] Attempts to predict the dipolar couplings from physical models of the systems have been recently considered. In particular, when neutral media are used, the electrostatic contribution can be neglected, and predictions of dipolar couplings in proteins can be made from a shape-induced alignment model.[4] In this method the neutral bicelles are treated as an obstructive infinite wall, and all the allowed orientations of the target molecule are explicitly calculated as a function of the distance from the wall, which is varied from the minimum to the maximum distance of the center of mass to surface atoms. Another approach assumes the principal axes of the moment of inertia tensor of the molecule to be the same as those of the orientation tensor and then applies a minimization algorithm to determine the other two parameters of the orientation tensor.[5]

In this work we proposed a heuristic approach, which depends on the fact that the alignment phenomenon in neutral media just reflects the asymmetries in molecular shape, information already encoded in the moment of inertia tensor of the model. The simplest linear combination of eigenvalues of **I** consistent with these observations and with the properties of **A** is a very good approximation for the orientation tensor.[6]

The description of the average solute orientation with respect to the magnetic field is described by a tensor **A** with elements[7]

$$A_{ij} = \frac{1}{2}\langle 3\cos\theta_i \cos\theta_j - \delta_{ij}\rangle$$

$$(i, j = x, y, z; \quad \delta_{ij} = 1 \quad \text{for} \quad i = j, \quad \delta_{ij} = 0 \quad \text{for} \quad i \neq j) \tag{1}$$

where θ_i is the angle between the molecular eigenvector i and the magnetic field, and the $\langle\ \rangle$ brackets denote time or ensemble averaging. From the elements of **A**, the dipolar coupling D_{PQ} of two spin-$^1/_2$ nuclei P and Q, separated by a distance r_{PQ}, can be calculated by $D_{PQ} = -\mathbf{S}\mu_0\gamma_P\gamma_Q h/(8\pi^3\langle r_{PQ}{}^3\rangle)\sum A_{ij}\cos\phi_i^{PQ}\cos\phi_j^{PQ}$, where **S** is the Lipari−Szabo generalized order parameter,[8] ϕ_i^{PQ} is the angle between the P−Q internuclear vector and the ith molecular axis, and the rest of the constants have their usual meaning.

Although the moment of inertia tensor, **I**, differs from **A** in that the latter has a null trace, we assume that, in the absence of electrostatic effects,[3,4] their eigenvectors are parallel.[5] Consistent with the experimental observations in neutral media, we assume that the eigenvalues of **A** are proportional to the asymmetries of the solute in the planes defined by pairs of eigenvectors of **I**, but we assume no explicit correspondence between the axes. Specifically, we assume that $A_{ii} \propto (I_{ii} - I_{jj}) + (I_{ii} - I_{kk}) = 2I_{ii} - (I_{jj} + I_{kk})$. This assumption fulfills the traceless condition of **A** and leads to the definition:

$$A_{ii}' = I_{ii} - \frac{I_{jj} + I_{kk}}{2} \quad (i, j, k = 1,2,3)$$

By arranging these trial matrix elements from greater to lesser absolute values a tensor with the same properties as the orientation tensor can be built up. Because of the null trace condition on **A**, an analysis of the possible outcomes for A'_{ii} can be done by sorting the eigenvalues of **I** as $I_{33} > I_{22} > I_{11}$; then $A'_{33} > 0$, $A'_{11} < 0$, and $|A'_{22}| < |A'_{11}|$, A'_{33}. Therefore, we get two possible situations (depending on the actual values taken by each I_{ii}) defining a correspondence between the principal axes of **A** and **I**:

$$(i) \quad A_{33}' + A_{11}' = -A_{22}' > 0 \Longrightarrow A_{zz} = A_{33}', \quad A_{yy} = A_{11}', \quad A_{xx} = A_{22}'$$

$$(ii) \quad A_{33}' + A_{11}' = -A_{22}' < 0 \Longrightarrow A_{zz} = A_{11}', \quad A_{yy} = A_{33}', \quad A_{xx} = A_{22}'$$

At this point **A** is completely defined except for a scale factor, dependent on experimental conditions, reducing the number of variables needed to predict dipolar couplings from 5 to 1. We refer to this method as TRacking Alignment from Moment of Inertia TEnsors—TRAMITE.

The method was tested with four oligosaccharides and three proteins, chosen to cover a wide range of sizes and shapes, whose structures were reported in the literature and available from the Protein Data Bank (PDB). In all cases the results were satisfactory except, as expected, when the dipolar coupling values were acquired in phage media where electrostatic effects are significant. We will illustrate the predictions by TRAMITE with some typical examples here, providing the remainder of the results in the Supporting Information.

Lewis^X is a well-studied rigid trisaccharide recently characterized by $^1D_{CH}$ measured in neutral bicelles.[9] In that work over 1000 model structures, from a Monte Carlo simulation, were individually oriented to obtain the best fitting to the experimental $^1D_{CH}$ values by means of a Powell algorithm. The same structures were analyzed with TRAMITE which produced similar results in a fraction of the time required previously. Figure 1A shows the correlation between measured and predicted $^1D_{CH}$ values for the best model of Lewis^X obtained by this procedure. Its glycosidic dihedral angles are in agreement with those reported.[9]

* To whom correspondence should be addressed. E-mail: bush@umbc.edu.
† Permanent address: Grupo de Biofísica, Departamento de Física, Universidad Nacional del Sur, (8000) Bahía Blanca, Argentina.

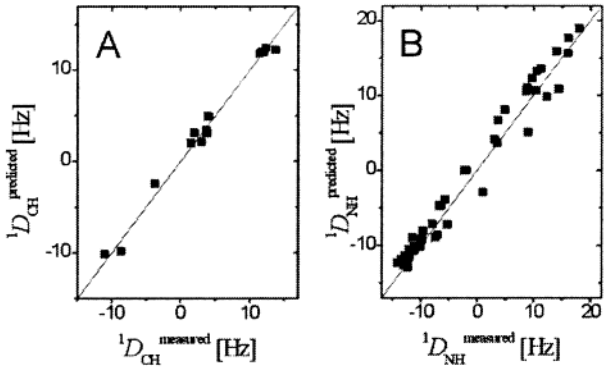

Figure 1. Correlations among experimental $^1D_{PQ}$ values and values predicted by TRAMITE. (A) Lewis^X β-methyl glycoside in a 7.5% w/v liquid crystalline bicelle medium.[9] (B) B1 domain of streptococcal protein G in a 5% w/v liquid crystalline bicelle medium.[10]

The Ig binding domain of streptococcal protein G was one of the molecules used by Zweckstetter and Bax[4] for testing of the shape-induced alignment method. Figure 1B shows the results by TRAMITE for one of the 32 models reported in the PDB from NMR experiments,[10] showing an agreement nearly as good as that obtained by the former method. The rhombicity value of 0.23 directly obtained from **I** was almost identical to the reported value of 0.20,[10] while the magnitude of ^1D_aNH was the same (-9.7 ± 0.1).

With a simple scheme for calculating **A** directly from a model, it is possible to analyze dipolar coupling data for a flexible oligosaccharide. As an example of a possible strategy to follow we show an analysis of the pentasaccharide LNF-2, that is known to have a rigid trisaccharide Lewis^a epitope.[9,11] Although a single conformation has been proposed for the disaccharide tail,[11] the TRAMITE scheme provides a method for critically testing this hypothesis. The models from a MC simulation of LNF-2[9] were analyzed in three different ways, using the merit function $\chi = \sqrt{\sum_i (^1D_i^{meas} - {}^1D_i^{pred})^2/N}$, by predicting the $^1D_{CH}$ values from (i) the individual **I**, (ii) averaging $\cos^2\theta_{ii}$ in eq 1 for structures 1 to j ($\langle\cos^2\theta_{ii}\rangle_{1-j}$), and (iii) same as (ii) discarding structures with individual χ values greater than 5. Figure 2A shows the results for these calculations. As can be observed, the predictions of $^1D_{CH}$ are distinctly improved when the average tensor is used, with further improvement obtained by filtering the structures in strong disagreement with the experimental values. Because this is an MC simulation without explicit solvent, it is not surprising that some low-energy models must be discarded. No similar improvement is seen with the rigid trisaccharides, but it is with LNF-3 (Supporting Information), suggesting flexibility for the lactose tails of both pentasaccharides. The correlations between measured and predicted $^1D_{CH}$ values for the best individual model (number 81) and the best average result (including 24 structures) are compared in Figure 2B. The representation of the 24 structures of LNF-2 superimposed on the central GlcNAc ring can be seen in part C of the Figure, where the rigid epitope at the left is clearly observed, while the tail region at right is relatively disordered.

The application of TRAMITE can be a useful tool for the prediction of residual dipolar couplings of biomolecules measured on neutral bicelles, as is shown in the examples analyzed here (along with those in the Supporting Information). The extension of the method to flexible molecules is simple and allows calculation of orientation tensors for ensembles of conformers of arbitrary size in extremely short times.

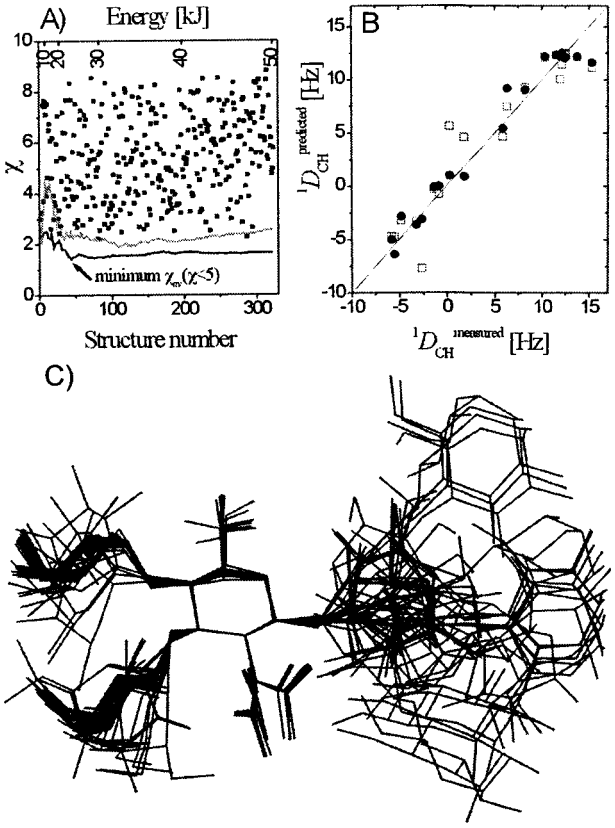

Figure 2. (A) χ values for predictions of $^1D_{CH}$ in LNF-2 made by TRAMITE for individual structures (●), using $\langle\cos^2\theta_{ii}\rangle$ for the structures 1 to j (with j varying from 1 to the total of structures) (gray line), and filtering the structures with $\chi_{indiv.} > 5$ (black line). (B) Correlations for model 81 (□) and for the filtered average structures where χ was minimum (●). (C) Superimposition of the 24 structures on the central GlcNAc residue that, averaged, best reproduce the $^1D_{CH}$ values.

Acknowledgment. This research was supported by NIH Grant GM-57211.

Supporting Information Available: A figure with the correlation of measured and predicted $^1D_{PQ}$ values by TRAMITE for: Lewis A β-methyl glycoside, LNF-3, ubiquitin, and cyanovirin-N (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111−1114.
(2) de Alba, E.; Baber, J. L.; Tjandra, N. *J. Am. Chem. Soc.* **1999**, *121*, 4282−4283.
(3) Losonczi, J. A.; Andrec, M.; Fisher, M. W. F.; Prestegard, J. H. *J. Magn. Reson.* **1999**, *138*, 334−342.
(4) Zweckstetter, M.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 3791−3792.
(5) Almond, A.; Bunkenborg, J.; Franch, T.; Gotfredsen, C. H.; Duus, J. Ø. *J. Am. Chem. Soc.* **2001**, *123*, 4792−4802.
(6) After this article was sent for consideration, a new method for predicting RDC based on the analytical solution for an ellipsoid was published: Fernandes, M. X.; Bernardó, P.; Pons, M.; García de la Torre, J. *J. Am. Chem. Soc.* **2001**, *123*, 12037−12047. The method seems to be also very promising, although a potential problem can arise in its application because the analytical solution is a non-linear function of experimental parameters. Furthermore, it requires the determination of the gyration tensor of the molecule that is apparently ill-defined for small proteins, indicating that would not be a good method for small molecules such as oligosaccharides.
(7) Saupe, A. *Angew. Chem., Int. Ed. Engl.* **1968**, *7*, 97−112.
(8) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4558.
(9) Azurmendi, H. F.; Martín-Pastor, M.; Bush, C. A. *Biopolymers* **2002**, *63*, 89−98.
(10) Clore, G. M.; Starich, M. R.; Gronenborn A. M. *J. Am. Chem. Soc.* **1998**, *120*, 10571−10572.
(11) Martín-Pastor, M.; Bush, C. A. *Biochemistry* **2000**, *39*, 4674−4683.

JA017524Z