Prediction of Sterically Induced Alignment in a Dilute Liquid Crystalline Phase: Aid to Protein Structure Determination by NMR

Markus Zweckstetter and Ad Bax*

*Laboratory of Chemical Physics
National Institute of Diabetes and Digestive
and Kidney Diseases, National Institutes of Health
Bethesda, Maryland 20892*

A tunable, weak degree of macromolecular alignment with respect to a static magnetic field can be obtained in a dilute, lyotropic liquid crystalline phase.[1] Provided the degree of solute alignment is sufficiently weak, its NMR spectrum retains the simplicity of the regular solution spectrum, but nevertheless permits measurement of dipole−dipole interactions.[1,2] So far, most dipolar coupling measurements have focused on the largest, one-bond heteronuclear dipolar couplings, $^1D_{NH}$ and $^1D_{CH}$. These residual dipolar couplings report on the average orientation of the bond vector with respect to the magnetic field, that is, with respect to the molecular alignment tensor.

Alignment of biological macromolecules in such a dilute nematic liquid crystalline phase can result from steric or electrostatic interactions with the nematogen, or both. Tjandra et al.[3] have recently demonstrated that in a bicelle medium,[4] consisting of large-diameter, disk-shaped zwitterionic phospholipid micelles,[4−6] principal axes of the molecular alignment tensor closely coincide with those of the rotational diffusion tensor. This confirms that in this nearly neutral medium, alignment is defined by the solute's shape. However, when adding a net charge to the bicelles, by doping them with either CTAB (+) or SDS (−),[7] the alignment tensor of a protein can be modified,[8] indicating that electrostatic interactions come into play. Indeed, for an oriented medium of negatively charged, rod-shaped viral particles,[9,10] or oriented purple membrane fragments,[11] electrostatic interactions often dominate alignment of solute proteins.

Here we demonstrate that both the magnitude and orientation of the steric component of the molecular alignment tensor can be accurately predicted from the solute's three-dimensional shape. Not only does this information provide the opportunity for structure validation, it may also distinguish between different degenerate solutions or provide input on flexibility in multidomain proteins.

Under the assumption that there are no attractive or long range repulsive interactions between the solute and the disk- or rod-shaped liquid crystal particles, solute alignment is dominated by the obstruction effect. Thus, the solute sample can be simulated as a collection of randomly oriented, uniformly distributed molecules, from which we remove the fraction that sterically clashes with the ordered array of liquid crystal particles. For example, for a disk-shaped nematogen and a rod-shaped solute molecule, a larger fraction of molecules oriented orthogonal to the disks will be obstructed than a fraction of molecules parallel to the disk surface, resulting in net ordering of the remaining, non-obstructed molecules. In practice, a systematic search of the solute orientations excluded by the obstruction effect, as a function of the distance from the nematogen is computationally more efficient, and this is the approach used below.

The average solute orientation with respect to the magnetic field is described by a second-rank tensor **A** (equivalent to the Saupe matrix if the liquid crystal director is parallel to the magnetic field) with a maximum of five independent elements.[12] The elements of this traceless tensor are

$$A_{ij} = {}^1\!/_2 \langle 3 \cos \theta_i \cos \theta_j - \delta_{ij} \rangle$$
$$(i,j = x, y, z; \delta_{ij} = 1 \text{ for } i = j, \delta_{ij} = 0 \text{ for } i \neq j)$$

where θ_i is the angle between the molecular axis i and the magnetic field, and the $\langle\ \rangle$ brackets denote time or ensemble averaging. The eigenvectors and eigenvalues of this real, symmetric matrix **A** correspond to the axes, the magnitude and the rhombicity of the molecular alignment tensor. For a pair of spin-$^1\!/_2$ nuclei, P and Q, separated by a distance r_{PQ}, the D_{PQ} dipolar coupling is related to the average orientation of the whole molecule by

$$D_{PQ} = -S \mu_o \gamma_P \gamma_Q h/(8\pi^3 \langle r_{PQ}{}^3 \rangle)) \sum A_{ij} \cos \phi_i{}^{PQ} \cos \phi_j{}^{PQ}$$

S is the Lipari−Szabo generalized order parameter, which scales D_{PQ} for the effect of fast librations of the internuclear vector,[13] γ_P and γ_Q are the gyromagnetic ratios, h is Planck's constant, μ_o is the magnetic permeability of vacuum, r_{PQ} is the internuclear distance, and $\phi_i{}^{PQ}$ is the angle between the P−Q internuclear vector and the ith molecular axis.

A computationally fast method for deriving **A** averages all individual alignment matrices, **A′**, calculated for each non-obstructed position and orientation of the solute. For a bicelle, its surface is approximated by an infinite wall located in the yz plane at $x = 0$. The center of gravity of the solute molecule is then moved in 0.5 Å steps from r_{min} to r_{max} along the x axis, where r_{min} and r_{max} are the minimum and maximum distance from its center to a surface atom,[14] increased by the radius of the corresponding surface atom. At each step, a uniform distribution of different solute orientations (spaced in ∼10° intervals) is sampled, and **A′** matrices are co-added for all allowed members of this distribution, that is, all orientations where none of its atoms has an x coordinate smaller than the atomic radius. At distances larger than r_{max} there is no obstruction, and the average **A′** value at each step therefore is zero. The net alignment tensor **A** equals the average over all **A′** values, for all permitted orientations at distances $r < d/(2V_f)$, where d is the bilayer thickness and V_f is the sample volume fraction (typically 3−6%) of dimyristoyl phosphatidylcholine (DMPC), which makes up the flat surface of the bicelle disks. Finally, **A** is multiplied by 0.8, to account for incomplete bicelle alignment relative to the magnetic field. Alignment caused by the bicelle rims, which slightly decreases **A** but has a negligible effect on its principal axes orientation or rhombicity, is neglected. The entire calculation typically takes less than 1 s.

For the cylindrically shaped filamentous phage based liquid crystals, calculation of **A** is fully analogous, but the center of the

(1) Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111−1114.
(2) Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 9279−9283.
(3) de Alba, E.; Baber, J. L.; Tjandra, N. *J. Am. Chem. Soc.* **1999**, *121*, 4282−4283.
(4) Sanders, C. R.; Schwonek, J. P. *Biochemistry* **1992**, *31*, 8898−8905.
(5) Vold, R. R.; Prosser, R. S. *J. Magn. Reson. Ser. B* **1996**, *113*, 267−271.
(6) Ottiger, M.; Bax, A. *J. Biomol. NMR* **1998**, *12*, 361−372.
(7) Losonczi, J. A.; Prestegard, J. H. *J. Biomol. NMR* **1998**, *12*, 447−451.
(8) Ramirez, B. E.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 9106−9107.
(9) Clore, G. M.; Starich, M. R.; Gronenborn, A. M. *J. Am. Chem. Soc.* **1998**, *120*, 10571−10572.
(10) Hansen, M. R.; Mueller, L.; Pardi, A. *Nat. Struct. Biol.* **1998**, *5*, 1065−1074.
(11) Koenig, B. W.; Hu, J. S.; Ottiger, M.; Bose, S.; Hendler, R. W.; Bax, A. *J. Am. Chem. Soc.* **1999**, *121*, 1385−1386.
(12) Saupe, A. *Angew. Chem., Int. Ed. Engl.* **1968**, *7*, 97−112.
(13) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4558.
(14) Eisenhaber, F.; Argos, P. *J. Comput. Chem.* **1993**, *14*, 1272−1280.
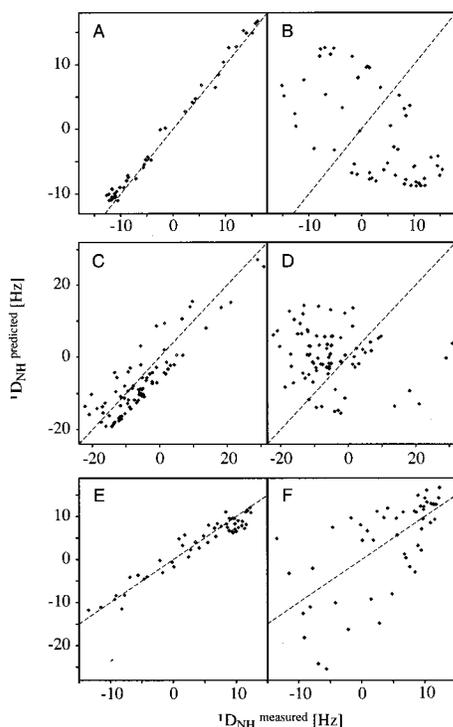
Figure 1. Correlations between experimental $^1D_{NH}$ values and values calculated from shape-predicted alignment tensors, for different protein−liquid crystal systems. (A) Protein G domain in a 5% w/v liquid crystalline bicelle medium of 3:1 DMPC/DHPC, and (B) in the nematic phase of bacteriophage *f*d (28 mg/mL). Cyanovirin-N in 4% w/v bicelles, using (C) the monomeric NMR structure and (D) the domain-swapped dimer X-ray structure. Apo-S100B in 5% w/v bicelles using (E) the NMR structure of the symmetric homodimeric protein apo-S100B, and (f) the structure of its monomeric subunit for predicting **A**. Dashed lines correspond to $y = x$. The increased scatter in panel C reflects a small difference between the predicted and best-fitted alignment tensors, presumably resulting from transient, weak electrostatic interaction with the not-quite-neutral bicelles.

33.5-Å radius (r_{cyl}) phage cylinder is placed at $x = y = 0$, and obstruction occurs when $x^2 + y^2 < r_{cyl}^2$ for any of the solutes surface atoms. For solutes small relative to the diameter of the phage, alignment tensors predicted in phage and bicelle media are very similar.

Figure 1 shows the results obtained for three different proteins. The alignment tensors, **A,** calculated for the Ig binding domain of Streptococcal protein G from its 1.0 Å crystal structure[15] are very similar for phage and bicelle media (Supporting Information). However, whereas in neutral bicelles the **A** tensor predicts dipolar couplings in perfect agreement with observed ones (Figure 1A), values in *f*d phage medium are very different (Figure 1B). This indicates that protein alignment in *f*d phage is not dominated by obstruction, and presumably contains a large contribution from electrostatic attraction between negatively charged *f*d and positive surface groups on the protein G domain.

As the shape of a monomeric protein and its homodimer are typically quite different, dipolar couplings can readily distinguish

between the two cases. For cyanovirin-N, a potent HIV-inactivating 101-residue protein, a domain-swapped dimer is observed in the crystalline state,[16] whereas solution NMR yielded a monomer.[17] Figure 1C shows that the dipolar couplings calculated from the **A** tensor predicted for the monomer are in very good agreement with experimental data, whereas those for the dimer are not (Figure 1D).

In contrast, $^1D_{NH}$ dipolar couplings measured for apo-S100B are only consistent with the **A** tensor predicted for the homodimeric form (Figure 1E).[18] Very different couplings are predicted for the monomer (Figure 1F), confirming that the protein is a homodimer in solution.

Prediction of the alignment tensor is most difficult for nearly spherical proteins, where small changes in shape have relatively larger effects on the alignment tensor. However, for most proteins asymmetry in their shape is defined at sufficient accuracy by typical NMR quality structures. For example, predicted alignment tensors for all 60 members of the ensemble of protein G domain NMR structures cluster closely around those in the X-ray structure (Supporting Information).

The obstruction model used for predicting macromolecular alignment yields a quantitative correlation between the solute's shape and its alignment. This offers a number of new opportunities. For example, it can be used to validate the correctness of a derived structure, or it may be used directly in the structure determination protocol to ensure agreement. Alternatively, it can be used to evaluate multiple-conformer models for flexible proteins: For a protein rapidly interconverting between multiple shapes, the measured dipolar couplings must agree with the ensemble-weighted average of couplings predicted for these shapes. The obstruction model works well for proteins in a suspension of nearly neutral aligned phospholipid bicelle particles. For a negatively charged nematogen, such as filamentous phage, the model is also expected to work well, provided the solute does not carry positive charges at its surface and a small correction for the radius of the negatively charged solute surface atoms is made. For cases where the nematogen and solute carry oppositely charged surface groups, it is expected that quantitative prediction of **A** will be much more complex.

The program for calculating the alignment tensor can be downloaded from http://spin.niddk.nih.gov/bax.

Supporting Information Available: One table containing the alignment tensor orientation and magnitude, calculated for the six cases of Figure 1, and one figure showing alignment tensor orientation and magnitude for all 60 NMR structures of the protein G domain (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0000908

(16) Yang, F.; Bewley, C. A.; Louis, J. M.; Gustafson, K. R.; Boyd, M. R.; Gronenborn, A. M.; Clore, G. M.; Wlodawer, A. *J. Mol. Biol.* **1999**, *288*, 403−412.
(17) Bewley, C. A.; Gustafson, K. R.; Boyd, M. R.; Covell, D. G.; Bax, A.; Clore, G. M.; Gronenborn, A. M. *Nat. Struct. Biol.* **1998**, *5*, 571−578.
(18) Drohat, A. C.; Tjandra, N.; Baldisseri, D. M.; Weber, D. J. *Protein Sci.* **1999**, *8*, 800−809.

(15) Derrick, J. P.; Wigley, D. B. *J. Mol. Biol.* **1994**, *243*, 906−918.