J|A|C|S
COMMUNICATIONS

Published on Web 08/27/2005

Efficient, Accurate Calculation of Rotational Diffusion and NMR Relaxation of Globular Proteins from Atomic-Level Structures and Approximate Hydrodynamic Calculations

Alvaro Ortega and Jose García de la Torre*

Departamento de Química Física, Facultad de Química, Universidad de Murcia, 30071 Murcia, Spain

Received May 11, 2005; E-mail: jgt@um.es

The sequence of residue-specific ratios of the longitudinal and transversal NMR relaxation times, T_1/T_2, embodies a considerable amount of valuable information about the structure of quasi-rigid macromolecules, such as globular proteins.[1−4] Using advanced hydrodynamics, García de la Torre[5] et al. recently provided a simple and accurate computer program, HYDRONMR, that enables the prediction of the relaxation times from the atomic structure taken from a PDB file. Even for situations when the NMR relaxation is not determined by overall rotational diffusion, Bernadó et al.[6,7] have shown how the program can be employed to obtain information about, for instance, chemical exchange, oligomerization, etc. The possibilities enabled by this methodology have recently been reviewed.[8] The full calculation of NMR relaxation for one structure takes a modest time, about 5 min, in a conventional PC. However, one can envision circumstances in which the calculation had to be repeated for a large number of structures, for instance, when the information embodied in the T_1/T_2 series is employed in a structural search or refinement[9−11] or folding predictions.[12] In these situations, a faster but yet accurate procedure would be most welcome.

In the atomic-level hydrodynamic modeling[13] that we use for predicting NMR relaxation from the rotational diffusion tensor, $\mathbf{D}_{rr}$, a primary hydrodynamic model (PHM) is constructed, replacing each non-hydrogen atom by a spherical element of radius a, which is expected to be larger than the van der Waals radius of the bare atoms, with some contribution due to hydration. In practice, it is floated as an adjustable parameter, although the range of acceptable values must be restricted (say, between 2.5 and 3.7 Å). Variations from one protein to another within this range may reflect effects of marginal importance, such as diversities in hydration, residual flexibility, etc. For a correct hydrodynamic evaluation, the PHM is in turn replaced by a rough shell model (RSM), composed of N small "minibeads" that represent the surface of the PHM (i.e., the solvent-accessible surface of the protein). This is done through an intermediate filling model (FM), in which the interior of the PHM is filled by N_{fill} closest packed beads; the internal beads that are surrounded by another 12 beads are removed to obtain the RSM. The greater suitability of shell modeling for hydrodynamic calculations was put forward by Bloomfield and Filson.[14] For more details on the shell-model calculation and its applications to proteins with atomic detail, see refs 13 and 15. The number of minibeads in the shell, N, should be sufficiently large (a few thousands, at least) to describe fine structural details. Unfortunately, the rigorous hydrodynamic treatment requires a computing time that is roughly proportional to N.[3] Thus, the physical rigor of this procedure obviously involves some computational cost.

A reduction in computational complexity can be achieved by introducing physical approximations, particularly in the description of hydrodynamic interaction. Indeed, this was done in the classical studies of Kirkwood[16] for simple, chainlike models and in those of Bloomfield et al.[14,17] for complex models of biological macromolecules. Their approach provided an expression for the translational diffusion coefficient, which is evaluated as a double sum, over all the beads in the model, of terms depending on the interbead distances. Later, García de la Torre et al.[18] presented a theory for $\mathbf{D}_{rr}$ using this double-sum approximation (DSA), a procedure that has the evident advantage of computational efficiency, because computing time in this procedure is proportional to N.[2] However, when applied to $\mathbf{D}_{rr}$, the DSA calculation introduces some appreciable bias into the results,[15,18] for which reason the rigorous procedure was preferred in the initial development of HYDRONMR. In the present work, we propose the atomic-level calculation of $\mathbf{D}_{rr}$ and NMR relaxation, replacing the rigorous procedure by the DSA; the results of the two procedures can then be compared in the search for some correction to the DSA results that would make them more accurate. This is done for a large sample of about 30 globular proteins (Supporting Information), which conform to the "quasi-rigid body" paradigm and have values of a in the acceptable range. One example is the outer surface protein A, OspA (1OSP),[19] which has a well-characterized structure, is remarkably rigid, and clearly shows the effects of rotational diffusion anisotropy with an appreciable variability in the $(T_1/T_2)_i$ sequence (see Supporting Information).

Previously,[5,7] we showed that the advantages of expressing the sequence dependence of $(T_1/T_2)_i$ are expressed by the quantity

$$\nabla_i = [(T_1/T_2)_i - \langle T_1/T_2\rangle]/\langle T_1/T_2\rangle \qquad (1)$$

where $\langle T_1/T_2\rangle$ is the average over all the residues (some outlier residues can be detected and disregarded[7]). Our present calculation shows that the rigorous and DSA values of $\langle T_1/T_2\rangle$ differ substantially, although they follow the same trend along the sequence; however, we find that the DSA values of ∇_i are practically coincident with the rigorous ones. This basic finding indicates that there is a simple relationship between the DSA and the rigorous rotational diffusion tensors, $\mathbf{D}_{rr}^{DSA}$ and $\mathbf{D}_{rr}^{rig}$; indeed, the eigenvectors are nearly identical, and the DSA-to-rigorous ratio is the same for the three eigenvalues. In short, this can be formulated as $\mathbf{D}_{rr}^{DSA} = Q\mathbf{D}_{rr}^{rig}$, where Q is a numeric factor that depends on the protein and the value of a in the PHM. The same coefficient relates the rotational diffusion coefficient, $D_r = \mathrm{Tr}(\mathbf{D}_{rr})/3$, which, in turn, determines the correlation time $\tau_c = 1/(6D_r)$, so that $Q = \tau_c^{rig}/\tau_c^{DSA}$ (values given in Supporting Information). It is known that the DSA is exact for a spherical bead/shell model, and its error increases with the axial ratio for ellipsoids.[14,17] For protein models, in addition to anisometry, an important factor will be the roughness of the surface. Both effects can be represented by a surface (S) to volume (V) ratio, expressed in dimensionless form as $S^{3/2}/V$. For the shell model, there is an equivalent term containing the number of beads in the filling model and in the shell model (the latter corrected by solvent exposure; see Supporting Information), namely

10.1021/ja053080l CCC: $30.25

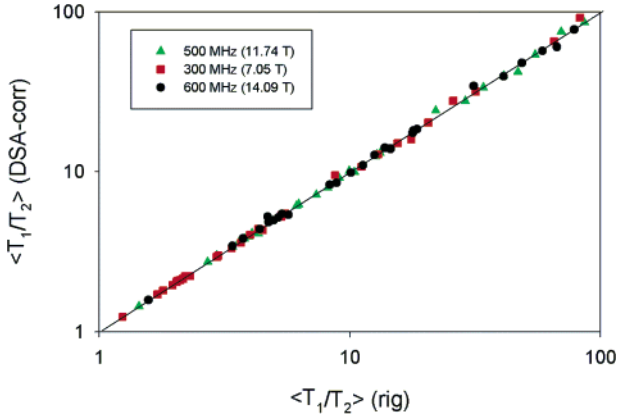

Figure 1. Comparison of the $\langle T_1/T_2 \rangle$ calculated $a = 2.5$ for 30 proteins with the two procedures for different magnetic fields.

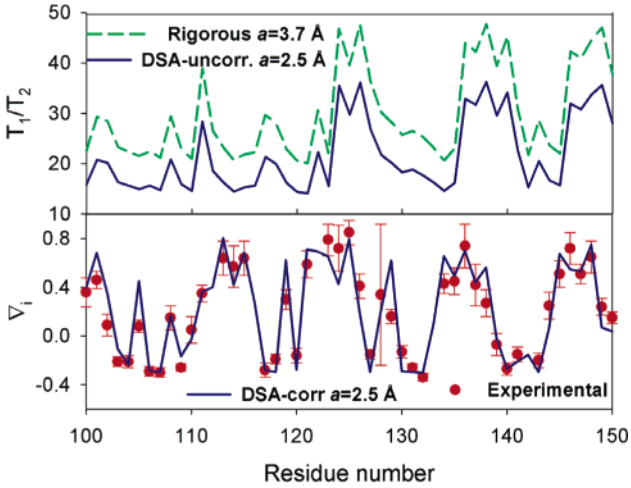

Figure 2. Results for OspA (600 MHz, 318 K) for residues 100−150. T_1/T_2 values calculated with different procedures and parameters. Experimental and DSA-corr values of ∇.

the ratio $\rho = N_{\text{shell,eff}}^{3/2}/N_{\text{fill}}$. With the values of Q determined for 30 proteins, each with five values of a between 2.5 and 3.7 Å, we find a good linear regression ($r^2 = 0.81$):

$$Q = 0.782 + 0.06363\rho \qquad (2)$$

In the new procedure, the bead/shell models are constructed as usual; from their numbers of minibeads we evaluate the ρ ratio, which, in turn, gives the Q factor from eq 2. The $\mathbf{D}_{\text{rr}}$ tensor is evaluated in the DSA (eqs 19−23 of García de la Torre et al.[18]) and corrected with the Q factor in eq 2 to obtain the corrected double-sum approximation result (DSA-corr):

$$\mathbf{D}_{\text{rr}}^{\text{DSA-corr}} = (1/Q)\mathbf{D}_{\text{rr}}^{\text{DSA}} \qquad (3)$$

Finally, NMR relaxation is calculated as in the original procedure. As a first test of the accuracy of the new method, the root-mean-square (rms) value of percent difference between the corrected DSA and rigorous values of the correlation time, τ_c, for the 30 proteins is 2.3−4.3% for the range of $a = 2.5$−3.7 Å. As the DSA (without correction) already gives the correct ∇_i, it is clear that the residue-specific ratios can be evaluated correctly as $(T_1/T_2)_i = \langle T_1/T_2 \rangle(\nabla_i + 1)$, if the average $\langle T_1/T_2 \rangle$ is accurate. This is demonstrated by our calculations for ^{15}N NMR relaxation of the sample of 30 proteins, which give an rms percent difference in $\langle T_1/T_2 \rangle$ of a few percent; for instance, with $a = 3.1$ Å, we find differences of 2.9, 3.3, and 3.9% for magnetic fields of 7.05, 11.74, and 14.09 T, respectively. Such small differences are acceptable as they are within the range of typical experimental errors. Figure 1 shows the coincidence of the DSA-corr and rigorous results for $\langle T_1/T_2 \rangle$; the agreement is very good for all the fields and all the values of a.

Figure 2 illustrates the good performance of the DSA in the calculation of NMR relaxation data. The advantage of the ∇ representation is evident, as it suppress the constant offset between calculations with different procedures and/or values the a parameter; the ∇ values are practically independent of these aspects. Finally, the T_1/T_2 values of the 270 residues of OspA can be predicted with the DSA-corr with the same accuracy as with the rigorous calculation (Supporting Information).

While a full NMR relaxation calculation with the original procedure takes 290 s in a PC with a 2.8 GHz Pentium, the new procedure takes just 0.9 s. This 300-fold increase in performance can make the new procedure a useful alternative for methodologies that require intensive use of NMR relaxation predictions.

Acknowledgment. We are grateful to Prof. M. Pons and Dr. P. Bernadó for valuable information and encouragement. This work was supported by a grant to J.G.T. (BQU2003-04517) and a predoctoral fellowship to A.O. from Ministerio de Educación y Ciencia.

Supporting Information Available: Tables and additional figures with detailed results. The new procedure has been implemented as an optional working mode of HYDRONMR, which can be downloaded from http://leonardo.fcu.um.es/macromol. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Kay, L.; Torchia, D.; Bax, A. *Biochemistry* **1989**, *28*, 8972−8979.
(2) Korzhnev, D. M.; Billeter, M.; Arseniev, A. S.; Orekhov, V. Y. *Prog. Nucl. Magn. Reson. Spectrosc.* **2001**, *39*, 197−266.
(3) Brüschweiler, R. *Curr. Opin. Struct. Biol.* **2003**, *13*, 175−183.
(4) Palmer, A. G. *Chem. Rev.* **2004**, *104*, 3623−3640.
(5) García de la Torre, J.; Huertas, M. L.; Carrasco, B. *J. Magn. Reson.* **2000**, *147*, 138−146.
(6) Bernadó, P.; Akerud, T.; García de la Torre, J.; Akke, M.; Pons, M. *J. Am. Chem. Soc.* **2003**, *125*, 916−923.
(7) Bernadó, P.; García de la Torre, J.; Pons, M. *J. Biomol. NMR* **2002**, *23*, 139−150.
(8) García de la Torre, J.; Bernadó, P.; Pons, M. *Methods Enzymol.* **2005**, *394*, 419−430.
(9) Tjandra, N.; Garret, D. S.; Gronenborn, A. M.; Bax, A.; Clore, G. M. *Nat. Struct. Biol.* **1997**, *4*, 443−449.
(10) Hus, J. C.; Marion, D.; Blackledge, M. *J. Am. Chem. Soc.* **1999**, *121*, 2311−2312.
(11) Wu, H.; Blackledge, M.; Maciejewski, M. W.; Mullen, G. P.; King, S. M. *Biochemistry* **2003**, *42*, 57−71.
(12) Annila, A.; Aitio, H.; Thilin, E.; Drakenberg, T. *J. Biomol. NMR* **1999**, *14*, 223−230.
(13) García de la Torre, J.; Huertas, M. L.; Carrasco, B. *Biophys. J.* **2000**, *78*, 719−730.
(14) Filson, D. P.; Bloomfield, V. A. *Biochemistry* **1967**, *6*, 1650−1658.
(15) Carrasco, B.; García de la Torre, J. *Biophys. J.* **1999**, *76*, 3044−3057.
(16) Kirkwood, J. G. *J. Polym. Sci.* **1954**, *12*, 1−16.
(17) Bloomfield, V. A.; Dalton, W. O.; van Holde, K. E. *Biopolymers* **1967**, *5*, 135−148.
(18) García de la Torre, J.; López Martínez, M. C.; García Molina, J. J. *Macromolecules* **1987**, *20*, 661−666.
(19) Pawley, N. H.; Koide, S.; Nicholson, L. K. *J. Mol. Biol.* **2002**, *324*, 991−1002.

JA053080L