Model-Dependent Artifacts in Long-Range Structure Determination Using Orientational Restraints Derived from NMR Relaxation

Jean-Christophe Hus, Dominique Marion, and
Martin Blackledge*

*Institut de Biologie Structurale
Jean-Pierre Ebel C.N.R.S.-C.E.A.
41, Avenue des Martyrs, 38027 Grenoble Cedex, France*

*Received September 14, 1998
Revised Manuscript Received January 19, 1999*

The study of heteronuclear relaxation is now established as an important probe for local and global dynamics in macromolecules.[1] It has been shown that the amplitude and orientation of the principal components of the global rotational diffusion tensor can be determined from the angular dependence of the R_2/R_1 ratio for spins in similarly rigid regions of macromolecules of known three-dimensional structure,[2] and that this ratio can provide a novel long-range constraint for NMR structure determination.[3]

In many studies of the global dynamics of molecules of known structure, an axially symmetric tensor was either assumed or shown, using statistical significance tests, to provide an adequate description of the rotational diffusion.[4] For the purposes of structure determination the axially symmetric model is also intuitively attractive as the relevant spectral density function provides a nondegenerate correspondence of R_2/R_1 and vectorial orientation with respect to the unique axis of the tensor.[5] Recently, structure determination algorithms have been presented using this model, and an associated restricted description of biomolecular form has been proposed to be of general applicability. [3,6]

It has recently been pointed out that two orthogonal solutions are present if the axially symmetric model is used to describe a fully anisotropic system, representing the prolate and oblate approximations to the real tensor.[7] In this communication we analyze the relevance of using either of these simplified tensorial models in the structure calculation algorithm for a molecule with a fully asymmetric diffusion tensor, and characterize the artifacts resulting from this approximation. This study is of additional interest as it has also been noted that the inappropriate use of an analogous approximation for structure determination using residual dipolar coupling may affect the accuracy of resulting coordinates.[8]

We have implemented relaxation rate constraints in the program Discover (MSI), incorporating the explicit target potential[9] into the classical potential energy function. A model system was
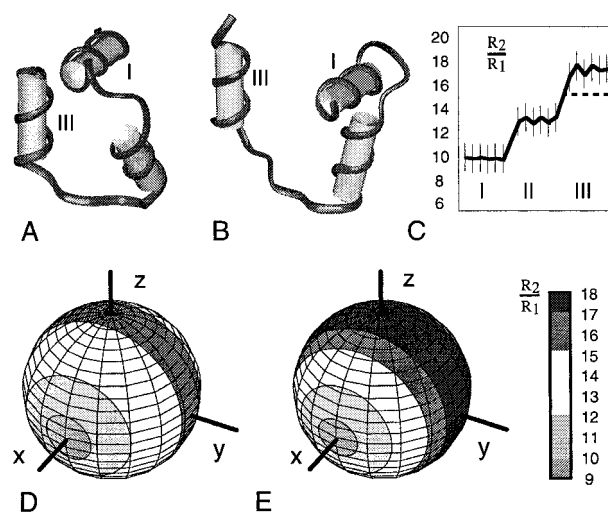

Figure 1. Structure determination using an oblate tensor. Typical structures resulting from simulated annealing calculations using an oblate tensor and relaxation ratios from the fully anisotropic diffusion tensorial model (Figure 3A). Tensorial values used are (A) $\{D_{||} = 10\ \mu s^{-1}, D_{\perp} = 17.5\ \mu s^{-1}\}$ and (B) $\{D_{||} = 7.5\ \mu s^{-1}, D_{\perp} = 17.5\ \mu s^{-1}\}$. (C) Relaxation ratios for these structures (A, dashed line; B, solid line) and model ratios (values centered on error bars). Vectorial orientation required with respect to these oblate tensors (D for tensor **A**; E for tensor **B**). Actual orientation is shown along the axes. The scale shows the available relaxation rates from the fully anisotropic tensor.[10]

created consisting of three orthogonal helices, present in a molecule with a fully anisotropic tensor.[10] For the purposes of illustration, we have placed the helices collinear with the principal components of the tensor—these represent the limiting values of the available relaxation rates for such a system (C and N terminal helices), and all intermediary values. Theoretical ^{15}N relaxation rates were calculated for this system and used as simulated experimental values in all structure calculations. For ease of presentation a single long-range distance was used in addition to short-range helical constraints. This is an unrealistically underdetermined system, but one which illustrates the general characteristics of the use of orientational constraints in structure calculations. Initially we assume that the component values have been estimated,[11] and that only the orientation of the vectors is allowed to float during the calculation. Subsequent fitting of the diffusion tensor to resulting structures would in this case allow refinement of the components $D_{||}$ and $D_{\perp}$. All calculations were performed with a routine simulated annealing protocol[12] starting from randomized Cartesian coordinates.

Initially we used the most appropriate oblate approximation to the real tensor—this assumes that $D_{||} = D_{xx}$ and $D_{\perp} = (D_{yy} + D_{zz})/2$ (Figure 1). While vectors aligned with the unique axis of the diffusion tensor are well-oriented and reproduce the experimental values, this is not the case for the other two helices—in fact the highest experimental values can never be attained with this tensor. Despite the correct positioning of the vectors, there is a high residual energy term in the structure calculation. Intermediate values (helix II) cannot be satisfied if they adopt

* To whom correspondence should be addressed.

(1) Palmer, A. G. *Curr. Opin. Struct. Biol.* **1997**, *7*, 732−737. Kay, L. E. *Nat. Struct. Biol.* **1998**, *5*, 513−517.

(2) Barbato, G.; Ikura, M.; Kay, L. E.; Pastor, R. W.; Bax, A. *Biochemistry* **1992**, *31*, 5269−5278. Brüschweiler, R.; Liao, X.; Wright, P. E. *Science* **1995**, *268*, 886−889.

(3) Tjandra, N.; Garrett, D. S.; Gronenborn, A. M.; Bax, A.; Clore, G. M. *Nat. Struct. Biol.* **1997**, *4*, 443−449. Clore, G. M.; Gronenborn, A. M. *Proc. Natl. Acad. Sci.* **1998**, *95*, 5891−5898.

(4) Tjandra, N.; Feller, S. E.; Pastor R. W.; Bax, A. *J. Am. Chem. Soc.* **1995**, *117*, 12562−12566.

(5) Woessner, D. E. *J. Chem. Phys.* **1962**, *37*, 647−654.

(6) The authors[3] propose that the rhombicity $\eta = (3/2)(D_{yy} - D_{xx})/(D_{zz} - (D_{xx} + D_{yy})/2)$ does not exceed 0.4 in macromolecules of significant axial anisotropy (defined as $2D_{zz}/(D_{xx} + D_{yy}) > 1.5$, which assumes a prolate form), and that the dispersion in R_2/R_1 cannot be explained by an oblate angular dependence. However, in the case of fully anisotropic tensors, similar angular dispersions with respect to both D_{zz} and D_{xx} can exist in biomolecules, even for small anisotropy. The cytochrome c_2[7] has $\eta = 0.7$ and $2D_{zz}/(D_{xx} + D_{yy}) = 1.23$, while the cytochrome c' has $\eta = 0.76$ and $2D_{zz}/(D_{xx} + D_{yy}) = 1.39$ (unpublished results). The model system treated here has $\eta = 1.0$ and $2D_{zz}/(D_{xx} + D_{yy}) = 1.6$.

(7) Blackledge, M.; Cordier, F.; Dosset, P.; Marion, D. *J. Am. Chem. Soc.* **1998**, *120*, 4538−4539. Cordier, F.; Caffrey, M.; Brutscher, B.; Cusanovich, M.; Marion, D.; Blackledge, M. *J. Mol. Biol.* **1998**, *281*, 341−361.

(8) Tjandra, N.; Omichinski, J. G.; Gronenborn, A. M.; Clore, G. M.; Bax, A. *Nat. Struct. Biol.* **1997**, *4*, 732−738.
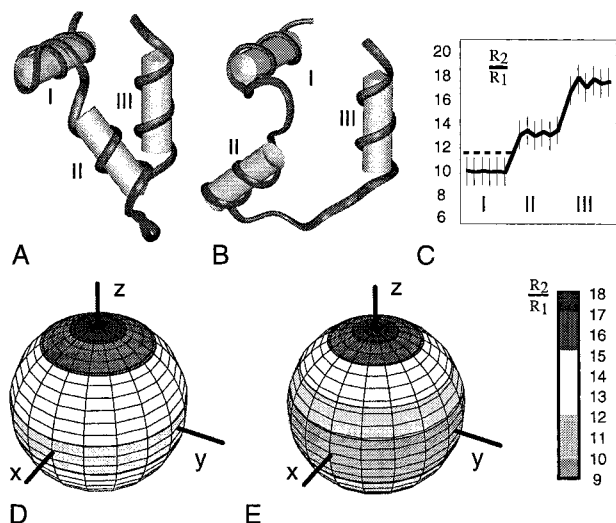
(9) $E_{aniso} = k[(R_{2meas}/R_{1meas}) - (R_{2calc}/R_{1calc})]^2/\sigma^2$. In the calculations shown here $k = 1.0$ kcal·mol^{-1}. The uncertainty σ is fixed at a value of 1.0 for the model calculation.

(10) All values were calculated assuming a proton resonance frequency of 600 MHz. The model tensor has the following components: ($D_{xx} = 10\ \mu s^{-1}$, $D_{yy} = 15\ \mu s^{-1}$, $D_{zz} = 20\ \mu s^{-1}$.

(11) Clore, G. M.; Gronenborn, A. M.; Szabo, A.; Tjandra, N. *J. Am. Chem. Soc.* **1998**, *120*, 4889−4890.

(12) Blackledge, M.; Medvedeva, S.; Poncin, M.; Guerlesquin, F.; Bruschi, M.; Marion, D. *J. Mol. Biol.* **1995**, *245*, 661−681.

Figure 2. Structure determination using a prolate tensor. Typical structures from tensors: (A) $\{D_{||} = 20\ \mu s^{-1},\ D_{\perp} = 12.5\ \mu s^{-1}\}$ and (B) $\{D_{||} = 23.0\ \mu s^{-1},\ D_{\perp} = 12.5\ \mu s^{-1}\}$. (C) Relaxation ratios for these structures (A, dashed line; B, solid line) and model ratios (values centered on error bars). Vectorial orientation required with respect to these prolate tensors (D for tensor A, E for tensor B). Scale and axes as for Figure 1.

the position of the true model, orthogonal to helices I and III, and must adopt a position between the $D_{||}$ axis and the $D_{\perp}$ plane.

An oblate tensor exists that allows all experimental data to be fulfilled; this gives a low target energy, but is foreshortened along the unique component axis to compensate for the higher relaxation rates, and is therefore no longer physically reasonable.

This effect is illustrated generally in Figure 1D,E, showing the relaxation rates from the model system and their respective orientations relative to the tensor in the two oblate models. The actual solution from which the data are derived, shown along the axes, is excluded in all cases. In the flattened model (E), all values from the real tensor are allowed, but the required orientation of intermediary relaxation rates is further from the correct model.

The same phenomenon is observed when a prolate axially symmetric model is used to describe the diffusion tensor (Figure 2). As in the oblate case, the tensor which fulfills all experimental values is nonphysical and in this case is extended to compensate for the lowest relaxation rates.

These observations have important consequences: Any calculation designed to determine both structure and tensor using either prolate or oblate models for this tensorial system will inevitably converge to the potential energy minima, even in the absence of noise, and therefore evoke the incorrect physical models represented by a highly prolate ($D_{||}/D_{\perp} = 1.84$) or highly oblate ($D_{||}/D_{\perp} = 0.43$) tensor. A more general study using a uniform distribution of vectors shows the same artifactual drift to nonphysical tensors,[13] and studies using experimental relaxation data of macromolecules combined with nOe data show similar characteristics.[13] Note that previously proposed fitting of experimental R_2/R_1 to the expected distribution from an axial tensor will also evoke tensors close to the minimum energy descriptions in both prolate and oblate cases.[11]

We have repeated the calculations allowing the eigenvalues of an axially symmetric diffusion tensor to float throughout the structure calculation, giving identical results.[14] The problem is therefore of a general nature, irrespective of the method used to parametrize an axially symmetric tensor from data emanating from a fully anisotropic tensor.

(13) See Supporting Information.
(14) Calculations were performed as shown here except that the values of $D_{||}$ and $D_{\perp}$ were free to float throughout the simulated annealing protocol. The final ensemble contained minimum energy structures with prolate and oblate tensor eigenvalues $\{D_{||} = 23.4\ \mu s^{-1},\ D_{\perp} = 12.4\ \mu s^{-1}\}$ and $\{D_{||} = 7.86\ \mu s^{-1},\ D_{\perp} = 17.5\ \mu s^{-1}\}$. respectively.
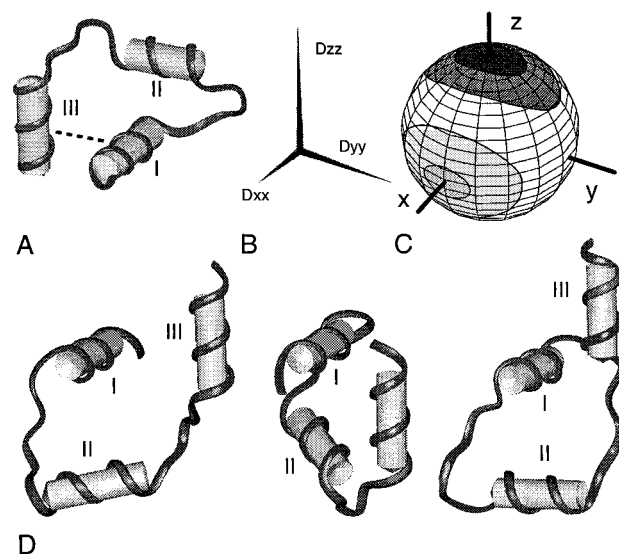

Figure 3. Structure determination using the fully anisotropic tensor description. Model structure (A) and tensor orientation (B) from which relaxation values were calculated. The dashed line represents the 10 Å long range distance included in the calculation. Available orientations for vectors using the fully anisotropic diffusion tensor (C). Actual orientation of the model is shown along the axes. Scale as in Figure 1. (D). Equally valid structures showing three orthogonal helices (right) and two orthogonal helices and the nonorthogonal middle helix (left, center).

Not surprisingly the model data can be reproduced if the correct tensor is used in the calculation (Figure 3). The directional dependence of the intermediary values of R_2/R_1 is however highly degenerate as vectors can adopt a large range of orientations to fulfill the correct values. This degeneracy may be partially raised by including relaxation rates defining a second vector in the same peptide plane or, as has recently been shown, by optimizing the structure with respect to a favorably oriented second tensor,[15] for example the tensor available from residual dipolar coupling measurements.[8]

In conclusion, the implications of using the axially symmetric model to describe the rotational diffusion tensor of a system of unknown structure appear to be potentially quite serious if this tensor is in reality rhombic. If used in structure determination algorithms, this approximation, of relevance for the study of rotational diffusion of a molecule of known structure, can lead to the incorrect alignment of vectors. More importantly, any procedure designed to determine structure and simultaneously evaluate rotational diffusion will provide an equally precise fit to either of the two inappropriate models represented by the geometrically orthogonal flattened oblate or extended prolate energy minima. Both final structure and rotational diffusion characterization are therefore strongly and nontrivially dependent on the initially chosen tensorial model. As the differentiation of experimental R_2/R_1 distributions from axially symmetric or fully anisotropic tensors[11] is unlikely to be statistically significant given the undersampling of angular space available from relaxation rates measured in biomolecules, it appears to be very important to use the fully anisotropic tensorial description in all ab initio structure calculations.

Supporting Information Available: Structure determination using 99 vectors with a uniform distribution of relaxation rates and structure calculations of cytochrome c_2 (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA983265A

(15) Ramirez, B. E.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 9106−9107.