Analysis of NMR Relaxation Data of Biomolecules with Slow Domain Motions Using Wobble-in-a-Cone Approximation

Shou-Lin Chang and Nico Tjandra*

Laboratory of Biophysical Chemistry, Building 50
National Heart, Lung, and Blood Institute
National Institutes of Health, Bethesda, Maryland 20892

Received August 16, 2001

The knowledge of internal dynamics of biomolecules is required to understand the biological relevant processes such as protein folding, molecular recognition, and binding specificity. NMR relaxation processes provide windows to look into these internal motions on the picosecond to millisecond time scales. However, the validity of the interpretation of NMR relaxation data strongly depends on the motional models that are chosen. The model-independent Lipari−Szabo "model-free" approach has been most widely applied to analyze biomolecular relaxation data. The advantages of "model-free" approach[1] are its simplicity and its range of validity. It provides two dynamical parameters: the order parameter S^2, which describes the spatial restriction of the motion, and the effective correlation time τ_e, which is the time scale of the motion. Since in principle "model-free" is a single-exponential time-dependent Padé approximation[2] to the exact internal correlation function of the motional vector, in general the fitting to the experimental data is quite satisfactory. However, there are situations where "model-free" approach starts to break down and the interpretation of order parameter and correlation time is not very straightforward.[3] For proteins with several domains, the slow interdomain motions are easily underestimated or often ignored, unless special precaution is taken. These errors propagate to the estimation of parameters such as overall tumbling, anisotropy, and the relative domain orientation when relaxation data are to be used for retrieving structural information. In this report, we compare the extended "model-free" approach[4] with a mathematically more elaborate triple-exponential Padé approximation in the context of wobble-in-a-cone model and examine the range of the validity, and their application to the case of calmodulin, with the emphasis on slow internal motion on the nanosecond time scale.

The extended model-free approach was proposed to account for internal motions with components on two distinct time scales. An extra exponential term is introduced to the correlation function of the original "model-free" formalism. Assuming the overall tumbling of the molecule is axially symmetrical, the correlation function of the extended "model-free" approach is described by eq 1:

$$C(t) = \frac{1}{5}\sum_{m=-2}^{2} \exp\{-[6D_{xy} + m^2(D_z - D_{xy})]t\} \times [d_{m0}^{(2)}(\alpha)]^2 \times \{(1 - S_f^2)\exp(-t/\tau_f) + S_f^2[S_s^2 + (1 - S_s^2)\exp(-t/\tau_s)]\} \quad (1)$$

where S_f, τ_f, S_s, τ_s, D_z, D_{xy}, and $d_{mn}^{(2)}(\alpha)$ are the generalized order parameter and time constant for the fast motion, generalized order parameter, and time constant for the slow motion, parallel, and perpendicular components of the diffusion tensor, and reduced Wigner rotation matrix elements,[5] respectively. α is the angle between the motional vector of interest and the long axis of the diffusion tensor. In the context of wobble-in-a-cone, the correlation function can be better described by a triple-exponential Padé approximation as shown by Lipari and Szabo.[6] A fast motional component is incorporated into the original equation in the same manner as in the extended model-free approach:

$$C(t) = \frac{1}{5}\sum_{m=-2}^{2}\sum_{n=-2}^{2} \exp\{-[6D_{xy} + m^2(D_z - D_{xy})]t\} \times [d_{mn}^{(2)}(\alpha)]^2 G_n(t)$$

$$G_n(t) = (1 - S_f^2)\exp(-t/\tau_f)\delta_{n0} + S_f^2\{G_n(\infty) + [G_n(0) - G_n(\infty)]\exp(-t/\tau_{\mathrm{eff}}^{(n)})\}$$

$$\tau_{\mathrm{eff}}^{(n)} = \tau_n/[G_n(0) - G_n(\infty)] \quad (2)$$

In eq 2 $G_n(t)$, $G_n(0)$, $G_n(\infty)$, $\tau_{\mathrm{eff}}(n)$, and τ_n all can be expressed as closed-form functions of D_w, the diffusion constant in the cone, and δ, the semicone angle, as described previously.[6] δ_{n0} is the Kronecker delta. Throughout this work, an effective overall correlation time, τ_c, defined as $\tau_c = [2D_z + 4D_{xy}]^{-1}$ is presented and tabulated.

$$D_w(1 - S_{\mathrm{cone}}^2)\tau_s = x^2(1+x)^2\{\ln[(1+x)/2] + (1-x)/2\}/[2(x-1)] + (1-x)(6 + 8x - x^2 - 12x^3 - 7x^4)/24$$

$$S_{\mathrm{cone}} = \frac{1}{2}x(1+x), \quad x = \cos(\delta) \quad (3)$$

Assuming wobble-in-a-cone is the slower component of the internal motion, for a fair comparison between the two different approximations, the correlation time τ_s in the extended model-free approach has to be translated to D_w according to eq 3[2], which is a closed-form function of δ. The $\tau_{\mathrm{eff}}^{(n)}$ in the triple-exponential approach is then calculated through D_w and δ. The deviation of the extended model-free from the triple-exponential approximation is most significant when the semicone angle δ becomes larger. The percentage differences between the relaxation rate $^1H-^{15}N$ R_1, R_2, and NOE of the extended model-free and triple-exponential approximation are plotted against the slow internal motion correlation time τ_s as shown in Figure 1. At $\tau_s = 3$ ns, for semicone angles from 0° to 90°, the absolute percentage differences reach a maximum of roughly 9, 6, and 18% for R_1, R_2, and NOE at semicone angles of 73°, 78°, and 70°, respectively. The percentage difference for NOE becomes very sensitive when τ_s close to 1 ns where the NOE values are close to zero.

A set of synthetic relaxation data generated from the triple-exponential Padé approximation using different semicone angle values from 0° to 90° are fitted with extended model-free approach to assess the possible errors that may be introduced when data are fitted by a mathematically more simplified approximation. Note that both approaches have the same number of unknowns to fit. When the semicone angle is equal to 0°, the two approaches are identical. As the semicone angle increases, the values of τ_s and δ tend to be underestimated (or S_s^2 is overestimated), while

(1) (a) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4559. (b) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4559−4570.
(2) Lipari, G.; Szabo, A. *Biophys. J.* **1980**, *30*, 489−506.
(3) Korzhnev, D. M.; Orekhov, V. Y.; Arseniev, A. S. *J. Magn. Reson.* **1997**, *127*, 184−191.
(4) Clore, G. M.; Szabo, A.; Bax, A.; Kay, L. E.; Driscoll, P. C.; Gronenborn, A. M. *J. Am. Chem. Soc.* **1990**, *112*, 4989−4991.
(5) Brink, D. M.; Satchler, G. R., *Angular Momentum*, 2nd. ed.; Oxford University: Oxford, 1961.
(6) Lipari, G.; Szabo, A. *J. Chem. Phys.* **1981**, *75*, 2971−2976.

10.1021/ja016862x CCC: $20.00
Published on Web 10/27/2001

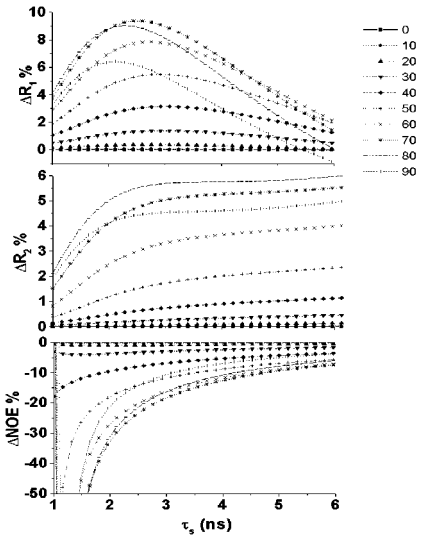

Figure 1. The percentage difference for R_1, R_2, and NOE between the extended Lipari−Szabo model-free and the triple-exponential Padé approximation. The simulation is done at 600 MHz spectrometer frequency by varying τ_s from 1 to 6 ns and δ from 0° to 90°. Both dipolar and chemical shift anisotropy relaxation are included. The τ_c is kept at 10 ns, D_z/D_{xy} at 1.5, τ_f at 10 ps, S_f^2 at 0.9, and α at 30°.

Table 1. Fitting of the Target Relaxation Data Generated from Triple-Exponential Padé Approach by Extended Model-Free Approach (EX) When Semicone Angle Is Fixed at 30° and 50°

	$\delta = 30°$		$\delta = 50°$	
	target	EX	target	EX
τ_c (ns)	10.0	9.7	10.0	7.1
D_z/D_{xy}	1.50	1.48	1.50	1.80
α	20°	20°	20°	19°
S_f^2	0.90	0.90	0.90	0.91
τ_f (ps)	10	16	10	30
S_s^2 (δ)	0.65 (30°)	0.69	0.28 (50°)	0.43
τ_s (ns)	3.0	2.6	3.0	2.3

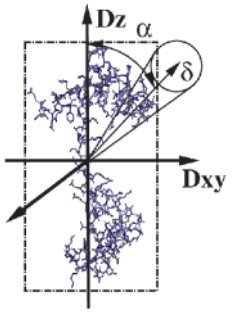

Figure 2. Schematic plot of calmodulin. The coordinates are from the structure of Ca^{2+}-ligated calmodulin.[10] Each domain moves as a rigid body. The cone is the proposed slow internal motion. The domain motion is deduced from the average property of individual vectors in each domain.

the values of τ_f and S_f^2 tend to be overestimated. Since τ_c, D_z/D_{xy}, S_s^2, and τ_s compensate each other (the χ-square surface corresponding to these parameters is rather flat), significant variations can be observed in all of these parameters. Table 1 shows the results when semicone angle is set to 30° and 50°.

Both approaches described above are applied to the case of calmodulin for comparison. Calmodulin has two distinct domains as shown in the schematic plot of Figure 2. Each has significant domain motions on the nanosecond time scale.[8] This is consistent with the correlated disorder observed in the recent X-ray study of the molecule.[7] Relaxation data R_1, R_2, and NOE acquired at spectrometer frequency 800, 600, and 360 MHz are used for dynamical analysis. Residues that are highly mobile and classified as parallel residues, that is, residues with N−H bond vectors parallel to the long axis of the rotational diffusion tensor, are excluded from the analysis.[8] The results are shown in Table 2. For both extended model-free approach and triple-exponential approximations, with NOE data from 360 MHz, the normalized χ-square is smaller than that from previous studies.[8] Note that the time constant of the domain motion becomes slower and its amplitude becomes larger than that from the previous study when NOE data at 360 MHz are included. NOE data at low field are important for estimating the slow motional component since they sample the motional fluctuation at lower frequency. Comparing to the model-free approach where the slow domain motion is ignored, the improvement in fitting is statistically significant as justified by the *F*-test[9] with the probability of chance improvement *P* much less than 0.001 for both approaches. The results in Table 2 show that for calmodulin, the triple-exponential approximation does not offer a better fitting to the experimental data. The error is similar in both approximations. One possibility is that the amplitude of the slow motion is not large enough for the two models to have statistically significant difference in fitting the NMR relaxation data. However, the trend that slow motion is underestimated and that fast motion is overestimated in extended model-free approach can be observed as we have noticed in the simulation. Another possibility is that the complicated domain motions of calmodulin might not follow the exact wobble-in-a-cone model; thus, a mathematically more elaborate approach does not seem to provide a better fitting.

Table 2. Dynamical Parameters for the N- and C-domain of Calmodulin Obtained by Fitting the Relaxation Data R_1, R_2, and NOE at 800, 600, and 360 MHz Spectrometer Frequency Using Lipari−Szabo Model-Free (LS), Extended Model-Free (EX), and Triple-Exponential Padé Approximation (TE) Approach[a]

	LS		EX		TE	
	N	C	N	C	N	C
τ_c (ns)	6.4	6.4	8.7	8.7	8.9	8.9
D_z/D_{xy}	1.72	1.72	1.54	1.54	1.55	1.55
θ (deg)[a]	68	51	64	67°	64	67
φ (deg)	95	133	96	138	96	138
S_f^2	0.84	0.85	0.87	0.85	0.87	0.86
τ_f (ps)	31	33	13	13	10	7
S_s^2			0.72	0.61	0.70	0.59
(δ) (deg)			(27)	(32)	(28)	(33)
τ_s (ns)			2.6	3.4	2.8	3.6
E/*N*	15.7		3.7		3.7	

[a] θ, φ: the angles for the long diffusion axis D_z with respect to the molecular frame; *E*/*N*: normalized χ-square against all the relaxation data used in the analysis.[8]

The wobble-in-a-cone model is potentially suitable for describing the slow internal motion of nucleic acid fragments and multidomain proteins.[11] Our results demonstrate that for motions with small amplitudes, the extended Lipari−Szabo model-free approach is sufficient. However, for motions with a larger cone angle, the triple-exponential approximation has to be applied to avoid significant deviation from the reality. It is also worth noting that NOE data at low field are important to extract such information.

Acknowledgment. We thank Attila Szabo for many useful discussions.

JA016862X

(7) Wilson, M. A.; Brunger, A. T. **2000**, *J. Mol. Biol. 301*, 1237−1256.
(8) Baber, J. L.; Szabo, A.; Tjandra, N. **2001**, *J. Am Chem. Soc. 123*, 3953−3959.
(9) Bevington, P. R.; Robinson, D. K. *Data Reduction and Error Analysis for Physical Sciences*, 2nd ed.; McGraw-Hill: New York, 1992.
(10) Babu, Y. S.; Bugg, C. E.; Cook, W. *J. Mol. Biol*. **1988**, *204*, 191−204.
(11) Braddock, D. T.; Cai, M.; Baber, J. L.; Huang, Y.; Clore G. M. *J. Am. Chem. Soc*. **2001**, *123*, 8634−8635.