Journal of Magnetic Resonance **143,** 229–232 (2000)
doi:10.1006/jmre.1999.2006, available online at http://www.idealibrary.com on IDEAL®

Simple and Accurate Determination of Global τ_R in Proteins Using ^{13}C or ^{15}N Relaxation Data

Joel Mispelter,[1] Nadia Izadi-Pruneyre, Eric Quiniou, and Elisabeth Adjadj

U350 INSERM, Institut Curie, Laboratoires R. Latarjet, F-91405 Orsay, France

Received September 1, 1999; revised December 10, 1999

In the study of protein dynamics by ^{13}C or ^{15}N relaxation measurements different models from the Lipari–Szabo formalism are used in order to determine the motion parameters. The global rotational correlation time τ_R of the molecule must be estimated prior to the analysis. In this Communication, the authors propose a new approach in determining an accurate value for τ_R in order to realize the best fit of R_2 for the whole sequence of the protein, regardless of the different type of motions atoms may experience. The method first determines the highly structured regions of the sequence. For each corresponding site, the Lipari–Szabo parameters are calculated for R_1 and NOE, using an arbitrary value for τ_R. The χ^2 for R_2, summed over the selected sites, shows a clear minimum, as a function of τ_R. This minimum is used to better estimate a proper value for τ_R. © 2000 Academic Press

Key Words: **rotational correlation time; relaxation parameters; protein dynamics; residues selection; factorial discriminant analysis.**

Recent advances in isotope labeling methods have enabled ^{13}C and ^{15}N nuclear spin relaxation to become widely used for studying internal motions in macromolecules (proteins). The relaxation data (R_1, R_2, NOE) are usually obtained at a single value of the static magnetic field, allowing a poor sampling of the spectral density function. The simplest way to modelize this function is by means of the Lipari–Szabo formalism (*1, 2*), that is, depending on only three parameters:

$$J(\omega) = \frac{2}{5}\left[\frac{S^2\tau_R}{1+(\omega\cdot\tau_R)^2} + \frac{(1-S^2)\tau_e}{1+(\omega\cdot\tau_e)^2}\right], \qquad [1]$$

where τ_R is the correlation time of the global motion (supposed isotropic) of the molecule in solution, τ_e is the effective correlation time of the internal motion, and S^2 represents the spatial restriction of the internal motion.

The determination of the rotational correlation time constitutes a primary requirement in the attempt at motion analysis in all molecular sites of the protein, starting from the experimental relaxation data obtained from labeled nuclei.

[1] To whom correspondence should be addressed at Institut Curie, Bat 112, Centre Universitaire, F-91405 Orsay, Cedex France. E-mail: joel.mispelter@curie.u-psud.fr.

The standard procedure for estimating this parameter is to use the R_2/R_1 ratio for selected sites in the protein sequence (*3*). This approach has been extensively discussed recently (*4–6*) and it is based on two principal ways of selecting residues.

One way requires that the selected sites should fulfill the extreme narrowing limit, which is a very restrictive condition. It assumes the second term of Eq. [1] to be negligible (*7*) (model 1 of LS formalism). The value for τ_R is then estimated from Eq. [2]:

$$\tau_R = \frac{1}{2\omega_X}\sqrt{\frac{6R_2}{R_1} - 7}. \qquad [2]$$

Due to the constraint imposed on the minimum value of τ_e, which is seldom fulfilled, τ_R is usually underestimated using this method (*4*). Besides, the proper sites might not be easily selectable a priori. Attempts have been made to improve the selection (*8*), based on NOE data, but nevertheless, the condition of independence of R_2/R_1 on S^2 still remains.

The second, less restrictive way to select residues consists of fitting simultaneously all three relaxation parameters with τ_R, S^2, and τ_e. The determination of τ_R in the latter approach has been carried out using the local site treatment of Shurr *et al.* (*9*) or using the globally linked approach of Dellwo and Wand (*10*). In this case the three LS parameters (τ_R, S^2, and τ_e) are calculated by minimizing the χ^2 function described by

$$\chi^2 = \sum_{\text{residues}}\left[\frac{(R_{1\text{exp}} - R_{1\text{calc}})^2}{\sigma^2(R_1)} + \frac{(\text{NOE}_{\text{exp}} - \text{NOE}_{\text{calc}})^2}{\sigma^2(\text{NOE})} + \frac{(R_{2\text{exp}} - R_{2\text{calc}})^2}{\sigma^2(R_2)}\right] \qquad [3]$$

in order to fit best, simultaneously, all three experimental relaxation parameters for each residue.

We propose a simpler method, in which the site selection is done using the general model 4 of the Lipari–Szabo formalism (*7*), accordingly fitting the relaxation data. In this way, for a given value for τ_R, one can obtain the values for S^2 and τ_e in

1090-7807/00 $35.00

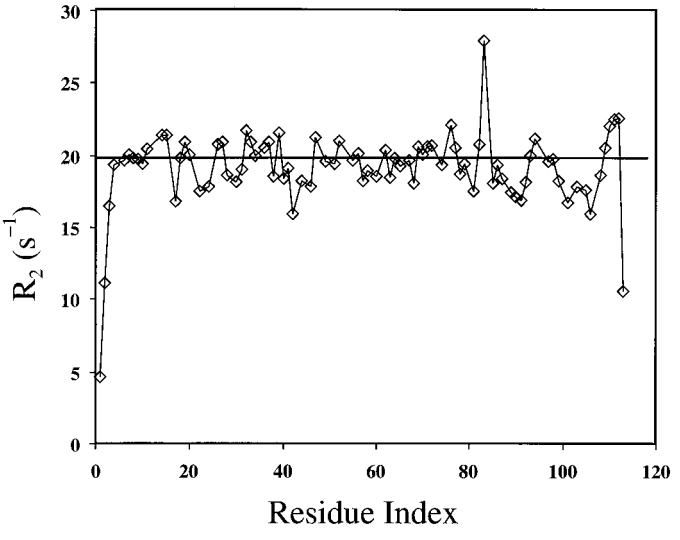

FIG. 1. C^{α} spin–spin relaxation rates as a function of residue index in the NCS protein.

Eq. [1] which fit exactly R_1 and NOE (*11*). Then, the value for τ_R is that obtained from the best fit of R_2 among the selected sites, using only the exact estimates of S^2 and τ_e from R_1 and NOE relaxation data. The extreme narrowing limit condition is no longer required but the selected sites should be submitted to fast and restrictive motion. They are usually located in the well-structured regions of the proteins and R_2 is a sensible parameter for making this choice.

In the case of very restrictive motions (low amplitude, $S^2 \approx 1$), the corresponding time scale of the correlation time is usually in the range of picoseconds. For these nuclei, the relaxation parameters R_1, R_2, and NOE can be fitted using Eq. [1] for $J(\omega)$ in the Lipari–Szabo formalism. For other nuclei submitted to a slow motion regime in the time scale of nanoseconds, it is necessary to apply an extension of the Lipari–Szabo formalism (*12*). In these cases the value of R_2 strongly diminishes. Moreover, there are cases where nuclei may have very slow motions contributing only to R_2 relaxation mechanisms and not to R_1 or NOE. For these specific residues, characterized by conformational changes on a time scale ranging in the domain of micro- to milliseconds, the R_2 values are increasing. Such an example is shown in Fig. 1 for the NCS protein (*13*). In the figure, one may easily distinguish two regions in the sequence, characterized by a practically constant R_2 value corresponding to highly structured domains in the protein. On the contrary, the edges of the sequence are characterized by decreased R_2 values, indicating the presence of nuclei with fast and high amplitude motions, in the time scale of nanoseconds. For these specific cases, where motions are modeled on two widely separated time scales, the Lipari–Szabo formalism proves to be inappropriate and the extended model-free approach of Clore *et al.* is invoked (*12*). Finally one may notice on the plot a residue (83) where the R_2 value is strongly increasing, indicating conformational changes affecting only the spin–spin relaxation rate.

The Lipari–Szabo formalism is concerned only with very fast and hindered motions. Hence, it is very appropriate to fit the dynamic parameters with R_1 and NOE that are insensible to slow motions. Assuming a known value for τ_R, it is possible to find unique solutions for S^2 and τ_e, exactly fitting the experimental data for R_1 and NOE (*11*).

We can suggest a new approach to determine τ_R from the R_2 fitting, in these highly structured domains of the sequence, where R_2 tends to be practically constant. This method is based only on LS formalism, taking into account the characteristics of the internal motion of the nuclei, without any further assumption on S^2 or τ_e.

From the R_1 and NOE fittings at an arbitrary chosen τ_R value, one could reconstruct the R_2 values at each site. The comparison with experimental data is done by calculating the corresponding χ^2 function,

$$\chi^2 = \sum_{\text{residues}} \frac{(R_{2\text{exp}} - R_{2\text{calc}})^2}{\sigma^2(R_2)}, \qquad [4]$$

where the sum is done over the selected sites. Note that, here, $R_{1\text{calc}}$ and NOE_{calc} equal exactly $R_{1\text{exp}}$ and NOE_{exp}, respectively, so that Eq. [3] reduces to Eq. [4].

The sites involved in the summation of Eq. [4] can be straightforwardly selected by examination of R_2 (Fig. 1), but a more objective method is provided by multivariate statistics analysis. For example, a factorial discriminant analysis (*14*) performed on R_1, R_2, and NOE experimental data allow us to clearly distinguish three regions (Fig. 2), of motionally similar sites. Residues inside region 2 have all three relaxation parameters fitting the LS formalism while those inside regions 1 and 3 are submitted to a slow motion regime (extension of LS) and exchange broadening, respectively.

The dependence of the χ^2 function over τ_R presents a clear minimum (Fig. 3), which gives the best value for the rotational correlation time. The error on the estimation of τ_R may be

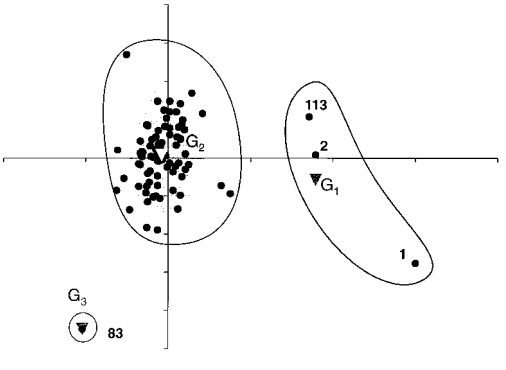

FIG. 2. Factorial discriminant analysis on experimental relaxation data (R_1, R_2, NOE) used for determining motional similarities among residues.

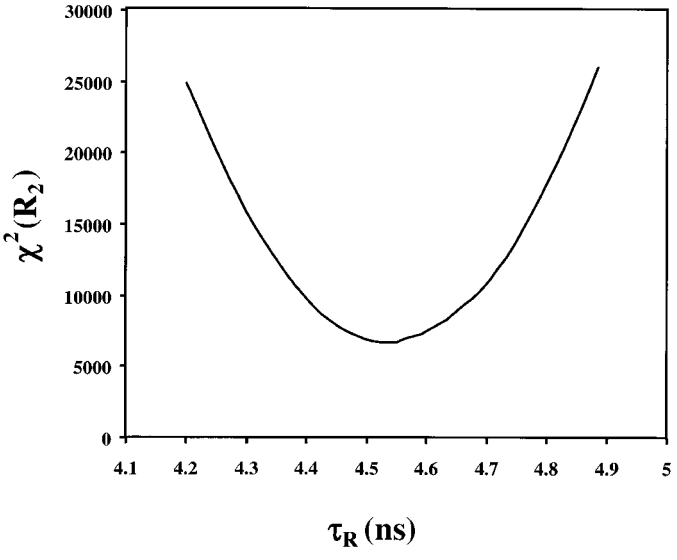

FIG. 3. $\chi^2(R_2)$ calculated from the set (S^2, τ_e) values, fitting R_1 and NOE data for the arbitrary chosen value of τ_R (the LS formalism, Eq. [1]), as a function of τ_R.

obtained from a Monte-Carlo simulation. This method has been tested and compared (Fig. 4) with the R_2/R_1 method (*15*), using data obtained for the NCS protein, totally enriched in ^{13}C, measuring relaxation rates of Cα, at 500 MHz. Three different selections of residues were used. The first one, giving the smaller χ^2, was restricted to a few residues belonging to the most structured regions of the protein. A second selection was extended to practically all of the protein backbone, except for the few residues at the edges and residue 83 which clearly exhibits an exchange contribution to R_2. This selection corresponds to all residues contained inside region 2 of Fig. 2. It gives the largest χ^2, due to the larger number of terms in the summation (Eq. [4]). Finally a third selection was tested. It includes the residues inside the gray region (Fig. 2) of the factorial discrimant analysis map. Figure 4 clearly shows that the τ_R values obtained by the R_2/R_1 method are systematically lower than those obtained by the method proposed in this paper, as expected (for none of the selected residues is the extreme narrowing limit fulfilled). Furthermore, it must be emphasized that the τ_R values thus obtained do not depend on the residue selection that is different from the R_2/R_1 case which gives τ_R values dependent on the selected sites. As a result, the calculated value for τ_R using the sites inside region 2 of the FDA map is 4.53 ns ($\sigma = 0.07$) for the NCS protein at 35°C.

We have also applied this method in order to obtain τ_R for the NCS protein for four temperatures (35, 40, 45, and 50°C). The variation of determined correlation times τ_R in function of η/T (Fig. 5) shows the expected linear dependence predicted by the Stokes–Einstein equation:

$$\tau_R = \frac{4\pi\eta_w r_H^3}{3kT}. \qquad [5]$$

Finally, a global exchange process or an anisotropic motion may be a drawback of this approach but it has the advantages

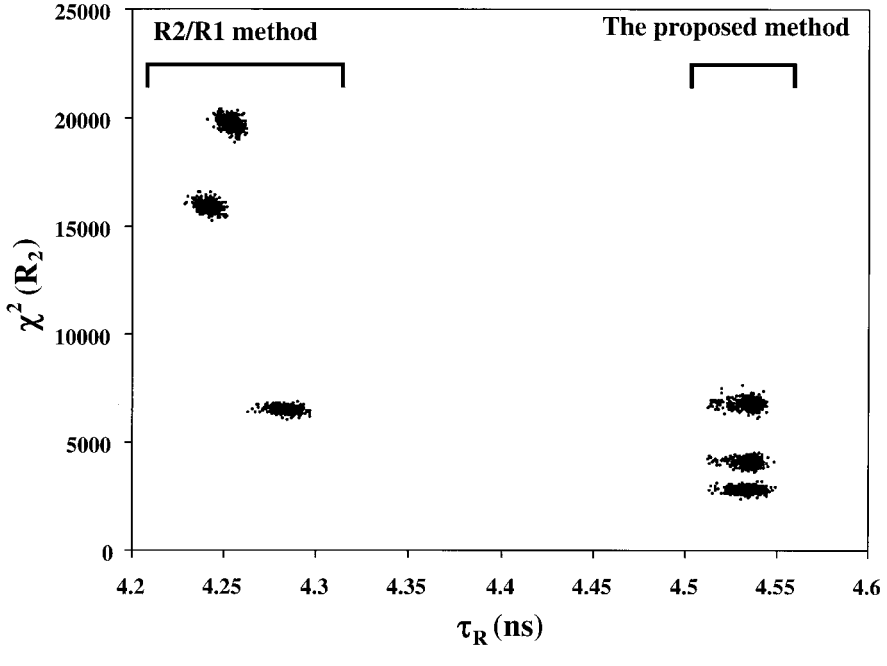

FIG. 4. $\chi^2(R_2)$ as a function of τ_R in a Monte-Carlo simulation. To the left, τ_R is calculated from Eq. [2] (R_2/R_1 method) using three different residue selections. To the right is the result obtained with the proposed method, using the same residue selections.

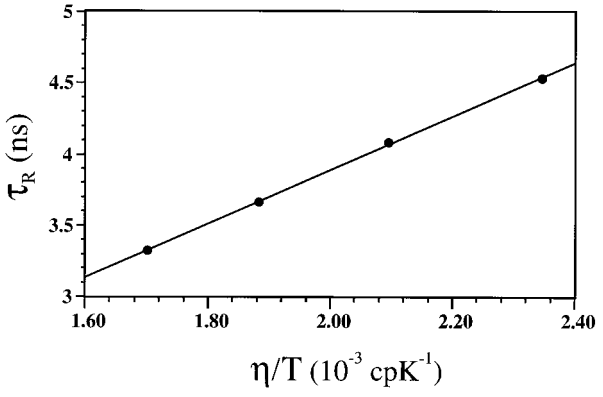

FIG. 5. τ_R dependence on the (η/T) term in the Stokes–Einstein equation as determined experimentally in NCS.

of rapid determination of motion parameters from measurements at a single value of the static magnetic field. It can also determine the proper τ_R value using the minimum hypothesis, fitting the best possible motion parameters inside the LS formalism—model 4 (*7*)—for the highly structured residues of the protein.

ACKNOWLEDGMENT

N.I.P. was supported by a postdoctoral fellowship from the Ligue Nationale Contre le Cancer (France).

REFERENCES

1. G. Lipari and A. Szabo, Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity, *J. Am. Chem. Soc.* **104,** 4546–4559 (1982).
2. G. Lipari and A. Szabo, Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results, *J. Am. Chem. Soc.* **104,** 4559–4570 (1982).
3. L. E. Kay, D. A. Torchia, and A. Bax, Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: Applications to staphylococcal nuclease, *Biochemistry* **28,** 8972–8979 (1989).
4. M. Andrec, G. T. Montelione, and R. M. Levy, Estimation of dynamic parameters from NMR relaxation data using the Lipari–Szabo model-free approach and bayesian statistical methods, *J. Magn. Reson.* **139,** 408–421 (1999).
5. A. L. Lee and A. J. Wand, Assessing potential bias in the determination of rotational correlation times of proteins by NMR relaxation, *J. Biomol. NMR* **13,** 101–112 (1999).
6. M. T. McMahon and E. Oldfield, Determination of order parameters and correlation times in proteins: A comparison between Bayesian, Monte Carlo and simple graphical methods. *J. Biomol. NMR* **13,** 133–137 (1999).
7. A. M. Mandel, M. Akke, and A. G. Palmer III, Backbone dynamics of *Escherichia coli* ribonuclease HI: Correlations with structure and function in an active enzyme. *J. Mol. Biol.* **246,** 144–163 (1995).
8. S. Yao, M. G. Hinds, and R. S. Norton, Improved estimation of protein rotational correlation times from ^{15}N relaxation measurements, *J. Magn. Reson.* **131,** 347–350 (1998).
9. J. M. Shurr, H. P. Babcock, and B. S. Fujimoto, A test of the model-free formulas: Effects of anisotropic rotational diffusion and dimerization, *J. Magn. Reson. B* **105,** 211–224 (1994).
10. M. J. Dellwo and A. J. Wand, Model-independent and model-dependent analysis of the global and internal dynamics of cyclosporin A, *J. Am. Chem. Soc.* **111,** 4571–4578 (1989).
11. J. Mispelter, C. Lefevre, E. Adjadj, E. Quiniou, and V. Favaudon, Internal motion of apo-neocarzinostatin as studied by ^{13}C NMR methine relaxation at natural abundance. *J. Biomol. NMR* **5,** 233–244 (1995).
12. G. M. Clore, A. Szabo, A. Bax, L. E. Kay, P. C. Driscoll, and A. Gronenborn, Deviations from the simple two-parameters model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *J. Am. Chem. Soc.* **112,** 4989–4991 (1990).
13. E. Adjadj, E. Quiniou, J. Mispelter, V. Favaudon, and J. M. Lhoste, Three-dimensional solution structure of apo-neocarzinostatin from *Streptomyces carzinostaticus* determined by NMR spectroscopy, *Eur. J. Biochem.* **203,** 505–511 (1992).
14. I. T. Jolliffe, "Principal Component Analysis," Springer, New York (1986).
15. N. A. Farrow, R. Muhandiram, A. U. Singer, S. M. Pascal, C. M. Kay, G. Gish, S. E. Shoelson, T. Pawson, J. D. Forman-Kay, and L. E. Kay, Backbone dynamics of a free and a phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation, *Biochemistry* **33,** 5984–6003 (1994).