Estimation of Dynamic Parameters from NMR Relaxation Data using the Lipari–Szabo Model-Free Approach and Bayesian Statistical Methods

Michael Andrec,*,† Gaetano T. Montelione,*,‡,1 and Ronald M. Levy†,1

*Center for Advanced Biotechnology and Medicine, †Department of Chemistry, Wright–Rieman Laboratories, and ‡Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08855-0939

E-mail: guy@nmrlab.cabm.rutgers.edu, ronlevy@lutece.rutgers.edu

In order to analyze NMR relaxation data in terms of parameters which describe internal motion, one must first obtain a description of the overall tumbling of the macromolecule in solution. Methods currently used to estimate these global parameters may not always provide reliable estimates of their values and uncertainties. In this paper, we present a general data analysis formalism based on products of Bayesian marginal probability densities which can be used to efficiently combine the information content from multiple experiments, such as R_1, R_2, and NOE data collected at multiple magnetic field strengths, or data from cross-correlation or rotating frame relaxation dispersion experiments. Our approach allows the estimation of global tumbling and internal dynamical parameters and their uncertainties without some of the assumptions which are made in the commonly-used methods for model-selection and global parameter estimation. Compared to an equivalent classical statistical approach, the Bayesian method not only is more computationally efficient, but also provides greater insight into the information content of the data. We demonstrate that this approach can be used to estimate both the isotropic rotational correlation time in the context of the original and "extended" Lipari–Szabo formalisms [Lipari & Szabo, *J. Am. Chem. Soc.* 1982, *104,* 4546; Clore *et al., J. Am. Chem. Soc.* 1990, *112,* 4989], as well as the rotational diffusion coefficients for axially symmetric anisotropic tumbling.   © 1999 Academic Press

Key Words: backbone dynamics; rotational correlation time; anisotropy; Monte Carlo; multiple fields.

INTRODUCTION

Dynamics play a significant role in the biological functions of proteins and other macromolecules (*1–3*), and the sensitivity of NMR to motions experienced by nuclear spins has made NMR a powerful tool for the study of macromolecular dynamics (*4, 5*). In particular, the dependence of relaxation rates on the spectral density $J(\omega)$ of the motion for various relaxation mechanisms is well known (*4, 6, 7*), allowing straightforward

prediction of the heteronuclear relaxation rates given a knowledge of $J(\omega)$. However, the inverse problem of learning about the motions from a knowledge of the relaxation rates is much more difficult (*8, 9*). Although it is possible to directly estimate $J(\omega)$ based on relaxation data (*10–12*), most analyses of NMR relaxation data assume some functional form for $J(\omega)$, the adjustable parameters of which have some intuitive physical meaning. The formalism proposed by Lipari and Szabo, known as the "model-free" approach since it was developed without any assumptions of a detailed physical model (*13, 14*), has proved to be extremely popular for the analysis of NMR relaxation data (*5, 15*) and will likely remain a *de facto* standard method because of its simplicity.

The original form of the Lipari–Szabo formalism contains an adjustable parameter for the overall tumbling of the macromolecule in solution as well as two parameters which describe the spatial restriction and timescale of the local dynamics of a given residue. It is clear that estimation of these parameters using the traditional three relaxation measurements (R_1, R_2, and NOE at one magnetic field strength) is dangerously close to being mathematically underdetermined, particularly for these nonlinear models. In practice, most current implementations of the Lipari–Szabo formalism estimate the global tumbling parameter independently at the start of the analysis, and then use that value to estimate the local dynamical parameters. These methods involve assumptions about the timescales of the internal motions and their distribution in the protein, which can lead to significant errors in estimates of their values and uncertainties. This is especially problematic as there exist strong correlations between the local and global parameter estimates, and this can lead to a propagation of errors in the estimation of the global parameters into the local parameters.

Furthermore, the Lipari–Szabo formalism has been extended in various ways, increasing the potential number of unknown local dynamical parameters from 2 to 5 (*7*). Clearly, it is not mathematically possible to fit all of these parameters using

---

[1] To whom correspondence should be addressed.

three relaxation measurements. Therefore, existing analysis methods select a subset of three or fewer of the possible parameters which adequately fit the data. Generally, only a single such subset is selected, even though the data may be equally well fit by different models with different ranges of motional parameters, or may in fact result from a complex motion requiring more than three motional parameters for its complete description.

One way in which these shortcomings can be avoided is by the use of more relaxation measurements. In fact, recent studies of protein dynamics by NMR relaxation are making increasing use of relaxation data collected at two magnetic field strengths (e.g., (16–20)). Since the implementations of the Lipari–Szabo formalism currently in use were developed when the measurement of only three data was routine, it is not clear that they are making the most efficient use of all of the available information in this larger amount of data. Furthermore, new experiments based on cross-correlated relaxation (21–26) and rotating-frame relaxation dispersion (27–30) are being developed and refined, and it would be desirable to have a general formalism which would allow the unified analysis of data from such experiments. Also, the need for more quantitatively reliable model-free parameter estimation has become imperative now that the interpretation of model-free parameters has moved beyond its original use as a qualitative description of backbone dynamics to more quantitative applications such as the estimation of thermodynamic parameters (31–33) and the analysis of changes in dynamics due to ligand binding and complex formation (5).

In this paper, we present a novel approach to the estimation of dynamical parameters based on products of Bayesian marginal probability densities which takes full advantage of the information content of relaxation data collected at multiple fields, is general enough to allow the incorporation of data from novel relaxation experiments currently being developed, and avoids the problems inherent in the "traditional" model-selection approaches currently in use. Previously, Jin et al. (34, 35) described how traditional analysis methods could seriously underestimate the uncertainties in the extracted model-free parameters, and proposed a graphical method for the analysis of NMR relaxation data. However, that analysis method could only be applied to the simplest form of the Lipari–Szabo formalism and assumed that the global tumbling correlation time was known a priori. The approach presented here can be viewed as a natural generalization of the graphical analysis method and allows the estimation of the global tumbling correlation time and the use of the full Lipari–Szabo formalism, while also retaining the ability to accurately characterize the uncertainties in the extracted model-free parameters.

THEORY

In the Lipari–Szabo model-free formalism, it is assumed that the total time-correlation function $C(t)$ is separable into a product of factors which depend only on the overall motion and the internal motion, respectively:

$$C(t) = C_O(t)C_I(t). \tag{1}$$

This is true if the overall tumbling is isotropic and is uncorrelated with the internal motion (13). If the overall tumbling is indeed isotropic, then $C_O(t)$ is given by

$$C_O(t) = \frac{1}{5} \exp\left(-\frac{t}{\tau_m}\right), \tag{2}$$

where τ_m is the rotational correlation time. Furthermore, it is assumed that the correlation function for the internal motion can be approximated by a single decaying exponential with a time constant of τ_e and decays to a value of S^2 as $t \to \infty$:

$$C_I(t) = S^2 + (1 - S^2)\exp\left(-\frac{t}{\tau_e}\right). \tag{3}$$

Thus, in the Lipari–Szabo formalism,

$$J(\omega) = 2 \int_0^\infty \cos(\omega t)C_O(t)C_I(t)dt$$

$$= \frac{2}{5}\left[\frac{S^2\tau_m}{1 + \omega^2\tau_m^2} + \frac{(1 - S^2)\tau}{1 + \omega^2\tau^2}\right], \tag{4}$$

where $\tau^{-1} = \tau_e^{-1} + \tau_m^{-1}$. The generalized order parameters S^2 can be interpreted as a measure of the spatial restriction of the internal motion, and the effective correlation time τ_e as a measure of the timescale of the internal motion (13). In the case of an internuclear vector oriented at an angle θ with respect to the symmetry axis of a molecule undergoing axially symmetric anisotropic diffusion with tumbling parameters $D_\parallel$ and $D_\perp$, Schurr et al. (36) have established conditions under which $J(\omega)$ can be written in the form

$$J(\omega) = \frac{2}{5} \sum_{j=0}^2 A_j\left[\frac{S^2\tau_j}{1 + \omega^2\tau_j^2} + \frac{(1 - S^2)\tau_j'}{1 + \omega^2\tau_j'^2}\right], \tag{5}$$

where $A_0 = \frac{1}{4}(3\cos^2\theta - 1)^2$, $A_1 = 3\cos^2\theta\sin^2\theta$, $A_2 = \frac{3}{4}\sin^4\theta$, $\tau_j^{-1} = (6 - j^2)D_\perp + j^2D_\parallel$, and $\tau_j'^{-1} = \tau_e^{-1} + \tau_j^{-1}$. This equation is an extension of the well-known expression for the spectral density of rigidly tumbling anisotropic ellipsoid (37).

The commonly measured relaxation data for a $^{15}N-^1H$ spin pair undergoing dipole–dipole and chemical shift anisotropy (CSA) relaxation in the absence of cross-correlation can then be computed using the well-known expressions

$$R_1 = \frac{d^2}{4} \left[J(\omega_H - \omega_N) + 3J(\omega_N) \right.$$

$$\left. + 6J(\omega_H + \omega_N) \right] + \frac{\Delta^2 \omega_N^2}{3} J(\omega_N) \qquad [6]$$

$$R_2 = \frac{d^2}{8} \left[4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) \right.$$

$$\left. + 6\,J(\omega_H + \omega_N) \right] + \frac{\Delta^2 \omega_N^2}{18}$$

$$\times \left[4J(0) + 3J(\omega_N) \right] + \omega_N^2 \Phi_{ex} \qquad [7]$$

and

$$NOE = 1 + \left(\frac{\gamma_H}{\gamma_N} \right) \frac{d^2}{4R_1} \left[6J(\omega_H + \omega_N) - J(\omega_H - \omega_N) \right],$$

$$[8]$$

where $d = \gamma_H \gamma_N h \mu_0 / 8\pi^2 r_{NH}^3$, Δ is the ^{15}N CSA, and Φ_{ex} is the field-independent portion of the contribution to R_2 due to chemical exchange. We should emphasize that the Bayesian formalism described below is in no way limited to the analysis of ^{15}N relaxation data, or to the Lipari–Szabo formalism, but can make use of other models for $J(\omega)$ and can be easily extended to make use of other relaxation data.

It should be made clear that the validity of Eq. [4] depends only on the assumptions mentioned above, namely, that the overall tumbling is isotropic and stochastically uncorrelated with the internal motion, and that the internal motion is well approximated by a single decaying exponential. Despite popular misconceptions, the derivation of Eq. [4] by Lipari and Szabo makes no assumption about the timescale of the internal motion, and for isotropic overall tumbling it is valid for *any* value of τ_e, including $\tau_e \geq \tau_m$, although the estimation of τ_e values in that regime becomes difficult and requires extremely precise data (*35*). The origin of this misconception appears to be the inclusion by Lipari and Szabo of simplifications of their main result (*13*, Eqs. [35]–[36]) which are only valid if $\tau_e \ll \tau_m$, as well as their discussion of the difficulties which arise if there are multiple internal motions on a timescale close to the overall tumbling. However, in the absence of such multiple internal motions, Eq. [4] is valid for any timescale motion as long as its correlation function can be well approximated by a single decaying exponential. The effect of slow motions in the case of anisotropic tumbling is less straightforward and will be considered below.

It has been observed that not all macromolecular NMR relaxation data can be fit well to the simple Lipari–Szabo spectral densities. In particular, it is sometimes necessary to account for contributions to the transverse relaxation rate R_2 due to chemical exchange effects by adding a phenomenological R_{ex} term to the predicted R_2. This effect can be put on

TABLE 1
Nomenclature for Subsets of Regression Variables

Model	Regression variables
1	S^2 ($\tau_e = 0$, $R_{ex} = 0$, $S_f^2 = 1$)
2	S^2, $\tau_e(R_{ex} = 0$, $S_f^2 = 1$)
3	S^2, $R_{ex}(\tau_e = 0$, $S_f^2 = 1$)
4	S^2, τ_e, $R_{ex}(S_f^2 = 1$)
5	S^2, τ_e, $S_f^2(R_{ex} = 0$)

somewhat firmer footing through its quadratic dependence on the spectrometer frequency, as indicated in Eq. [7], where $R_{ex} = \omega_N^2 \Phi_{ex}$ (*38, 39*), through the observation of the RF field-strength dependence of rotating-frame relaxation (*27–30*), or through the dependence of the apparent R_2 on the length of the CPMG delay (*20*). Furthermore, it is occasionally necessary to invoke motion on two widely-separated time scales to adequately fit the data, resulting in the so-called "extended model-free" spectral density of Clore *et al.* (*40*),

$$J(\omega) = \frac{2}{5} S_f^2 \left[\frac{S^2 \tau_m}{1 + \omega^2 \tau_m^2} + \frac{(1 - S^2)\tau}{1 + \omega^2 \tau^2} \right], \qquad [9]$$

where S^2 is the order parameter for the slow motion, S_f^2 is the order parameter for the fast motion, and τ_e is the effective correlation time for the slow motion (the fast motion is assumed to be in the extreme narrowing limit). It should be noted that Eq. [9] departs from conventional notation in order to emphasize its relationship with the original Lipari–Szabo spectral density of Eq. [4].

The current standard implementations of the Lipari–Szabo approach (*41–43*) share the following overall strategy: (1) the correlation time for overall tumbling (τ_m) is estimated from values of the ratio R_2/R_1 for a selected subset of the residues, (2) nonlinear least-squares fits to the observed relaxation data are performed using various subsets of the possible regression variables (Table 1), (3) model-selection criteria are used to decide which choice of regression variables is appropriate for each residue, and (4) the value of τ_m may be re-optimized using the selected models. The uncertainties in all of the estimated parameters are then determined by classical Monte Carlo error analysis. The use of the R_2/R_1 ratio to estimate the correlation time for overall tumbling in the context of the Lipari–Szabo formalism was first proposed by Kay *et al.* (*44*) in their analysis of ^{15}N relaxation data from staphylococcal nuclease. For the subset of residues which have motions only in the extreme narrowing limit (i.e. residues which satisfy model 1), the ratio R_2/R_1 will be independent of S^2 and therefore will depend only on τ_m. This is due to the fact that both R_1 and R_2 are linear functions of $J(\omega)$, and $J(\omega)$ is proportional to S^2 if $\tau_e = 0$ (Eqs. [4], [6], and [7]). Based on the experimentally measured NOE values for staphylococcal nuclease, Kay *et al.*
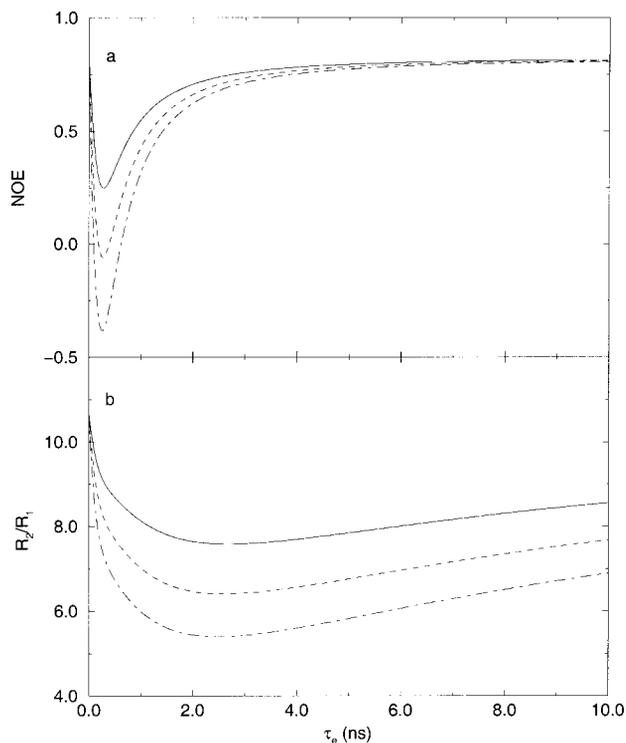
FIG. 1. The dependence of the NOE (a) and R_2/R_1 (b) on τ_e, calculated using Eqs. [4]–[7] assuming a spectrometer ^{1}H frequency of 600 MHz, an isotropic rotational correlation time of 10 ns, and S^2 values of 0.8 (solid line), 0.7 (dashed line), and 0.6 (dot-dashed line).

(44) pointed out that the second τ_e-containing term of Eq. [4] could be neglected after eliminating obvious outlier residues with anomalous R_2 values due to chemical exchange and residues undergoing complex dynamics near the N- and C-termini.

This assumption is often valid for a subset of the residues in a given protein. However, if it is assumed that the choice of model is unknown *a priori,* then the use of the R_2/R_1 ratio to estimate τ_m constitutes an *ad hoc* assumption, since we have no guarantee that removing the obvious outliers or residues with R_2/R_1 outside of one standard deviation of the mean will remove all nonmodel 1 residues. Indeed, in some systems there may be no residues which are accurately described by model 1. This problem has recently been recognized in the literature (45, 46). In particular, Yao *et al.* (46) have proposed an iterative scheme to determine a lower bound on the NOE value for residues to be used for τ_m determination, in an effort to exclude residues with significant τ_e contributions. However, it is clear from the dependence of the NOE on τ_e (Fig. 1a) that one obtains nearly identical NOE values both with $\tau_e \approx 0$ as well as with $\tau_e \geq 4$ ns (given $\tau_m = 10$ ns). Thus, one cannot definitively exclude the presence of large τ_e contributions on the basis of the magnitude of the NOE. Furthermore, it is clear that the presence of such large τ_e contributions will reduce the value of R_2/R_1, causing τ_m to be underestimated (see Fig. 1b).

Such inaccurate estimates of τ_m can have significant effects on the model selection, as has been observed in the analysis of protein ^{15}N relaxation data. In our experience, changes in τ_m of $<10\%$ are sufficient to cause the appearance or disappearance of R_{ex} terms, as well as other changes in the "best fit" model (e.g., Table 3 of Li and Montelione (47)). In particular, the underestimation of τ_m can lead to the appearance of inaccurate R_{ex} terms (see below).

The combined effect of τ_m uncertainty and model selection can be even more insidious. For example, if all τ_m values in a 10% range as described above were in fact consistent with the data, then one would not be able to decide conclusively whether the data for a given residue was better fit by model 1 or model 3, since different values of τ_m in that range would lead to a different inference about the presence of an exchange contribution. It is clear that in such a situation the uncertainty in τ_m would be propagated into an uncertainty in the model choice, and therefore the selection of a single "best-fit" model can result in the underestimation of the uncertainty in the model-free parameters. Such model-selection uncertainty is not considered in current analysis methods. This may be due to the fact that the concept of model uncertainty, although intuitively reasonable, does not arise naturally in classical statistics.

STATISTICAL BACKGROUND AND COMPUTATIONAL METHOD

As was alluded to in the Introduction, the problems arising from model selection and the estimation of the global tumbling parameters could be avoided by the measurement of a sufficient number of relaxation data and fitting to the most general form of the Lipari–Szabo model. In principle, such data analysis could proceed along either a classical or a Bayesian statistical route. For example, if one wished to treat the problem classically (i.e., where the data are considered to be random variables and the parameters are unknown states of nature (48)), one could easily obtain the maximum likelihood estimates for the local dynamical parameters for each residue conditional on some value for τ_m. If one assumes that the uncertainty in each data point is normally distributed, then this maximum likelihood estimate is simply the parameter values which minimize the weighted sum of squared residuals with respect to the data. One could then find the value of τ_m which simultaneously minimizes the total of the conditional sum of squared residuals for all residues using a one-dimensional optimization algorithm (49).

The uncertainties in the estimated dynamical parameters (including τ_m) could then be estimated by performing a classical Monte Carlo error analysis procedure, where one would repeat this optimization procedure for many random data sets generated using the assumed noise distribution. However, since τ_m is a *global* adjustable parameter, one can no longer perform the error analysis residue by residue. Instead, one would have to generate data sets corresponding to *all* of the relaxation data

for the entire protein. A large number of Monte Carlo samples would be required in order to adequately cover this very high dimensional parameter space, making a purely classical approach very computationally demanding.

A Bayesian version of this strategy is not only more computationally efficient, but also provides more insight into the information content of the data. Since Bayesian statistical methodology allows us to speak of the probability of a particular set of parameter values given the data (50, 51), we save a considerable amount of computational time by performing Monte Carlo sampling in the *parameter* space instead of the *data* space. Such sampling algorithms do not require repeated nonlinear optimizations, and therefore avoid convergence problems which sometimes arise with such algorithms. Furthermore, we can construct the overall solution based on the contributions from individual residues using the well-established methods of probability theory. This allows us to more easily assess the relative information content of the data for each residue and more easily identify outliers due to systematic errors in the data or model inadequacy.

A Bayesian version of the classical algorithm proposed above would begin by estimating the "local" joint probability density $P(S^2, \tau_e, R_{ex}, S_f^2, \tau_m | R_i)$ for all of the parameters (in this case for the "extended model-free" spectral density) based on the relaxation data R_i for residue i. By fitting to the most general form of this set of nested models, we avoid making any implicit model selections. Obviously, we must have at least as many data points as adjustable parameters if this local joint probability density is to provide an informative parameter estimate, but due to experimental uncertainties and the nonlinearities of the model function it is desirable that the fit be overdetermined. In this paper, we make use of R_1, R_2, and NOE data collected at multiple magnetic field strengths. In practice, we have found that six data points collected at sufficiently different fields (e.g., 400 and 600, or 500 and 800 MHz) provide good local parameter estimates, but that as few as five measurements (R_1, R_2, NOE, η_{xy}, and η_z) at one field strength (25) is also sufficient.

We can enforce the condition that τ_m is a global parameter *a posteriori* by "integrating out" all of the local parameters, and multiplying the resulting marginal probability densities $P(\tau_m | R_i)$ for all residues i. It is interesting to note that some earlier applications of the Lipari–Szabo formalism to protein dynamics have in fact proceeded in precisely this manner, by attempting to treat τ_m as an adjustable parameter on equal footing with S^2 and τ_e (52–55), while others have made use of the sum of squared residuals surface or other graphical methods (56, 57). This work extends these approaches by making full use of the information contained in the weighted sum of squared residuals surface, which is simply the negative of the logarithm of the Bayesian posterior probability under a uniform prior (see Eqs. [10]; [11] below). By recognizing this fact, we can bring to bear the full power of probability theory to make statistical inferences about the global and local parameters in a way that is both computationally efficient and provides considerable insight into the information content of the data.

In this paper, we test this strategy for the extraction of dynamical information from NMR relaxation data. For simplicity, let us first consider the original Lipari–Szabo spectral density (Eq. [4]). Using synthetic ^{15}N relaxation data, we generated Monte Carlo samples from the joint posterior probability density calculated using the Bayes theorem from the likelihood of the data and the prior probability of the parameters:

$$P(S^2, \tau_e, R_{ex}, \tau_m | R_i)$$
$$\propto P(R_i | S^2, \tau_e, R_{ex}, \tau_m) P(S^2, \tau_e, R_{ex}, \tau_m). \quad [10]$$

The likelihood of observing the data set R_i for a given residue i given that the underlying dynamic processes are described by the given values of S^2, τ_e, R_{ex}, and τ_m is given by

$$P(R_i | S^2, \tau_e, R_{ex}, \tau_m)$$
$$= \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi\sigma_{ij}^2}} \exp\left[\frac{-(R_{ij} - R_{ij}^{(calc)})^2}{2\sigma_{ij}^2}\right], \quad [11]$$

where n is the number of relaxation data per residue, R_{ij} is the jth "observed" relaxation value (i.e. R_1, R_2, or NOE) for the ith residue, $R_{ij}^{(calc)}$ is the jth relaxation parameter calculated using the given values of S^2, τ_e, R_{ex}, and τ_m, and σ_{ij} is the uncertainty in the jth "observed" relaxation value for the ith residue. The prior probability $P(S^2, \tau_e, R_{ex}, \tau_m)$ was taken to be equal to one in the region $0 < S^2 < 1$, $\tau_e \geq 0$, $R_{ex} \geq 0$, $\tau_m > 0$, and zero outside of this region. In principle, this prior could include information obtained from other sources, such as R_{ex} information obtained from rotating frame relaxation measurements (27–30). Equation [10] contains a proportionality sign since we have not normalized the posterior probability by dividing by the marginal likelihood of the data (49, 50). The Monte Carlo was performed using the Metropolis algorithm with an iteratively adjusted proposal distribution using the *XRambo* software package as described in detail by Andrec and Prestegard (58). For purposes of simulation, we have chosen σ_{ij} to be 5% of R_{ij} for all of the calculations described in the following sections. Based on these Monte Carlo samples, one can obtain an estimate of the marginal density of τ_m for the ith residue

$$P(\tau_m | R_i) = \int P(S^2, \tau_e, R_{ex}, \tau_m | R_i) dS^2 d\tau_e dR_{ex} \quad [12]$$

by estimating the one-dimensional probability density based only on the τ_m coordinates of the Monte Carlo samples (59, 60). Analogous procedures could be used for both axially-

TABLE 2
Parameters and Synthetic Data Used for the Determination of τ_{m} [a]

	Residue number					
	1	2	3	4	5	6
S^2	0.84	0.64	0.71	0.79	0.62	0.82
τ_{e} (ns)	1.0	2.3	0.0	0.5	2.4	1.5
R_{ex} (s^{-1}) [b]	0.0	0.0	1.0	0.5	3.1	1.2
R_1 (s^{-1}) @ 500 MHz	1.79	2.15	1.24	1.70	2.18	1.87
R_1 (s^{-1}) @ 800 MHz	1.10	1.49	0.65	1.02	1.52	1.18
R_2 (s^{-1}) @ 500 MHz	11.79	10.10	10.65	11.57	13.04	12.88
R_2 (s^{-1}) @ 800 MHz	14.50	12.29	14.47	14.86	20.02	17.43
NOE @ 500 MHz	0.549	0.602	0.805	0.276	0.605	0.616
NOE @ 800 MHz	0.695	0.771	0.855	0.425	0.774	0.754

[a] The synthetic R_1, R_2, and NOE relaxation data were calculated using Eqs. [4], [5], [6], and [7] with a τ_{m} of 10.0 ns and the values of S^2, τ_{e}, and R_{ex} listed in the table.

[b] We define R_{ex} to be the value of $\omega_{\mathrm{N}}^2\Phi_{\mathrm{ex}}$ at a proton frequency of 500 MHz.

symmetric anisotropic diffusion (Eq. [5]) or the "extended model-free" spectral density (Eq. [9]). In the former case, there are two global parameters ($D_{\parallel}$ and $D_{\perp}$), and therefore the required marginal densities are bivariate.

RESULTS AND DISCUSSION

1. Isotropic tumbling using the original Lipari–Szabo spectral density. Synthetic ^{15}N relaxation data at two field strengths (500 and 800 MHz) were generated using the parameters summarized in Table 2. In particular, since large τ_{e} values can severely limit the precision of Lipari–Szabo parameter estimates (*34, 35*), we have included τ_{e} values in the 1–2 ns range in order to test the robustness of our methodology in this regime. These data were then used to generate Monte Carlo samples from the "local" joint probability densities as described above. The resulting marginal probability densities $P(\tau_{\mathrm{m}}|R_i)$ are shown in Fig. 2. It is apparent from the widths of the distributions that these "local" τ_{m} estimates are of quite variable precision, with some residues (such as 1, 3, and 4) having fairly narrow distributions and thereby providing considerable information regarding the value of τ_{m}, while others (such as residue 5) are quite broad and provide very little τ_{m} information. This insight into the relative information content of the data for various residues with respect to the estimation of τ_{m} is something which would not be readily available from the classical parameter estimation procedure outlined above.

The resulting "local" τ_{m} distributions in several cases do not have their modes at the τ_{m} value (10.0 ns) used to generate the synthetic data, particularly for residues with substantial τ_{e} contributions (such as residue 2), although in all cases the "correct" τ_{m} value is near the maximum of the posterior probability density. The fact that the modes are not at 10.0 ns is not unexpected, since Eq. [4] represents a highly nonlinear model. Furthermore, the modes of $P(S^2, \tau_{\mathrm{e}}, R_{\mathrm{ex}}, \tau_{\mathrm{m}}|R_i)$ and $P(\tau_{\mathrm{m}}|R_i)$

need not occur at the same value of τ_{m} if the former contains significant nonlinear correlations (*58*). As expected, however, the $P(\tau_{\mathrm{m}}|R_i)$ do converge to sharp distributions centered at the "correct" value of 10.0 ns as the uncertainties σ_{ij} approach zero (data not shown).

Since τ_{m} is a global parameter, it must have the same value for all residues. We can impose this constraint by taking the product of $P(\tau_{\mathrm{m}}|R_i)$ over all N residues:

$$P(\tau_{\mathrm{m}}|R) \propto \prod_{i=1}^{N} P(\tau_{\mathrm{m}}|R_i). \qquad [13]$$

The probability density of τ_{m} obtained by multiplying the kernel density estimate points of Fig. 2 is shown in Fig. 3. Thus, $P(\tau_{\mathrm{m}}|R)$ constitutes the estimate of the value of τ_{m} and its uncertainty given the experimental data. Again, because of the nonlinearities in the model, the mode of the resulting distribution need not occur at 10.0 ns. However, the results do indicate that a τ_{m} of 10.0 ns is highly probable.

Although this is not a feature of the synthetic data used here, it should be recognized that systematic error in the relaxation measurements, more complex internal motions, or rotational anisotropy could result in $P(\tau_{\mathrm{m}}|R_i)$ distributions which are mutually inconsistent. Such inconsistency could be used as an indicator of such effects and is another example of insight into the data that could not have been easily obtained using classical parameter estimation methods. In practice, one could confirm that an inconsistency in τ_{m} is due to anisotropy by observing a correlation between the τ_{m} values and the relative orientations of the amide bond vectors as determined from suitable X-ray crystal or NMR solution structures (*36*). However, it is also possible that systematic errors in relaxation measurements or ill-determined relative amide bond orientations may not result
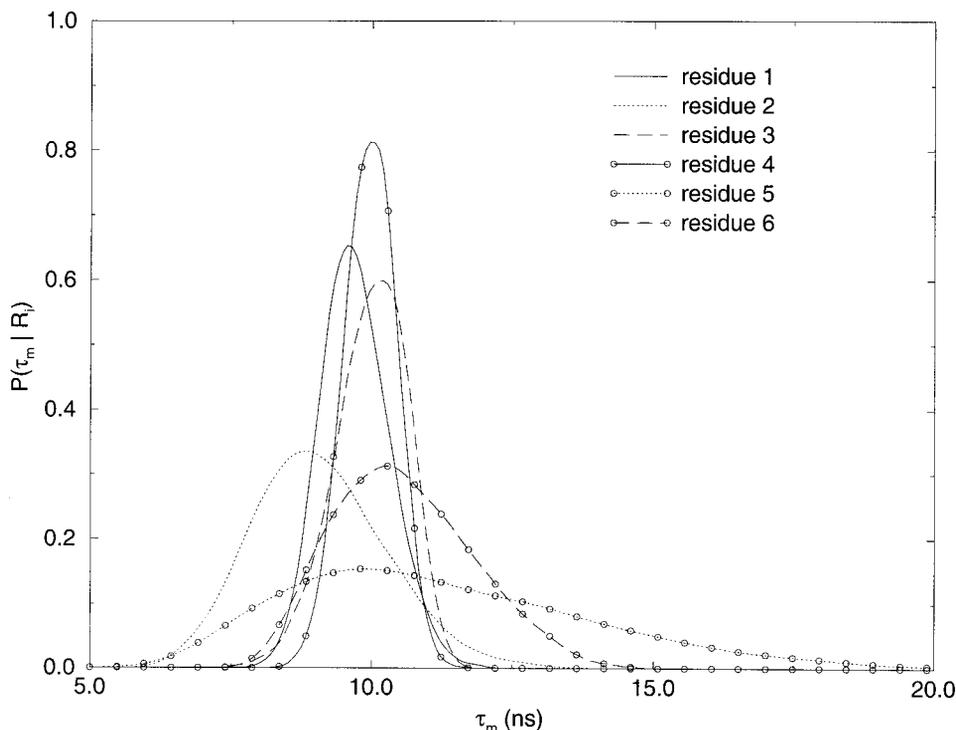
FIG. 2. The "local" marginal probability densities $P(\tau_m|R_i)$ calculated using the relaxation data R_i for each of the six residues shown in Table 2. The curves were calculated using a univariate Gaussian kernel density estimator (59) from Monte Carlo samples using the "maximal smoothing" window width (60). A total of 120,000 Monte Carlo samples were generated using *XRambo,* of which every eighth sample was used for the probability density estimate.

in a detectable deviation in the local τ_m estimates. In such cases, it may still be possible to detect such errors through a residual analysis in which the observed data is compared to the distribution of data values back-calculated from $P(S^2, \tau_e, R_{ex}, \tau_m|R_i)$.

Once the local distributions for the tumbling parameter have been constructed, they can then be used to determine the
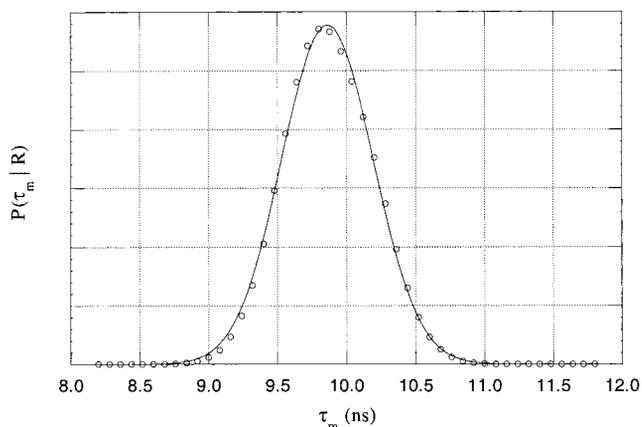

FIG. 3. The circles represent the products of $P(\tau_m|R_i)$ (Fig. 2) for $i = 1, \ldots, 6$ evaluated at τ_m increments of 0.08 ns. The solid line is the least-squares-fit Gaussian, having a mean of 9.86 ns and a standard deviation of 0.33 ns.

probability density of S^2, τ_e, and R_{ex} for each residue. This is done by generating Monte Carlo samples from the "global" joint density

$$P(S^2, \tau_e, R_{ex}, \tau_m|R)$$
$$\propto P(S^2, \tau_e, R_{ex}, \tau_m|R_i)[\prod_{j \neq i} P(\tau_m|R_j)] \quad [14]$$

for the local parameters of residue i, where the quantity in square brackets represents the information about τ_m contained in the other N-1 residues. Since model selection in the traditional analysis procedure consists of determining whether there is a significant τ_e or R_{ex} contribution, we have plotted the results as bivariate marginal densities of τ_e and R_{ex} in Fig. 4. In all cases, the probability densities of the local parameters are quite well-determined. The precision of the R_{ex} estimate could be further improved by collecting additional R_2 data at a third field strength, if that is practical. It is clear that for residues 4, 5, and 6 most of the probability mass is located well away from either the $\tau_e = 0$ or the $R_{ex} = 0$ axes, confirming that they are well fit by model 4 ($\tau_e \neq 0$ and $R_{ex} \neq 0$). Residues 1 and 2, on the other hand, have most of the probability mass lying at the $R_{ex} = 0$ axis, but well separated from the $\tau_e = 0$ axis, confirming that they can be fit to model 2 ($\tau_e \neq 0$ and $R_{ex} = 0$). Similarly, residue 3 has most of its probability mass lying at
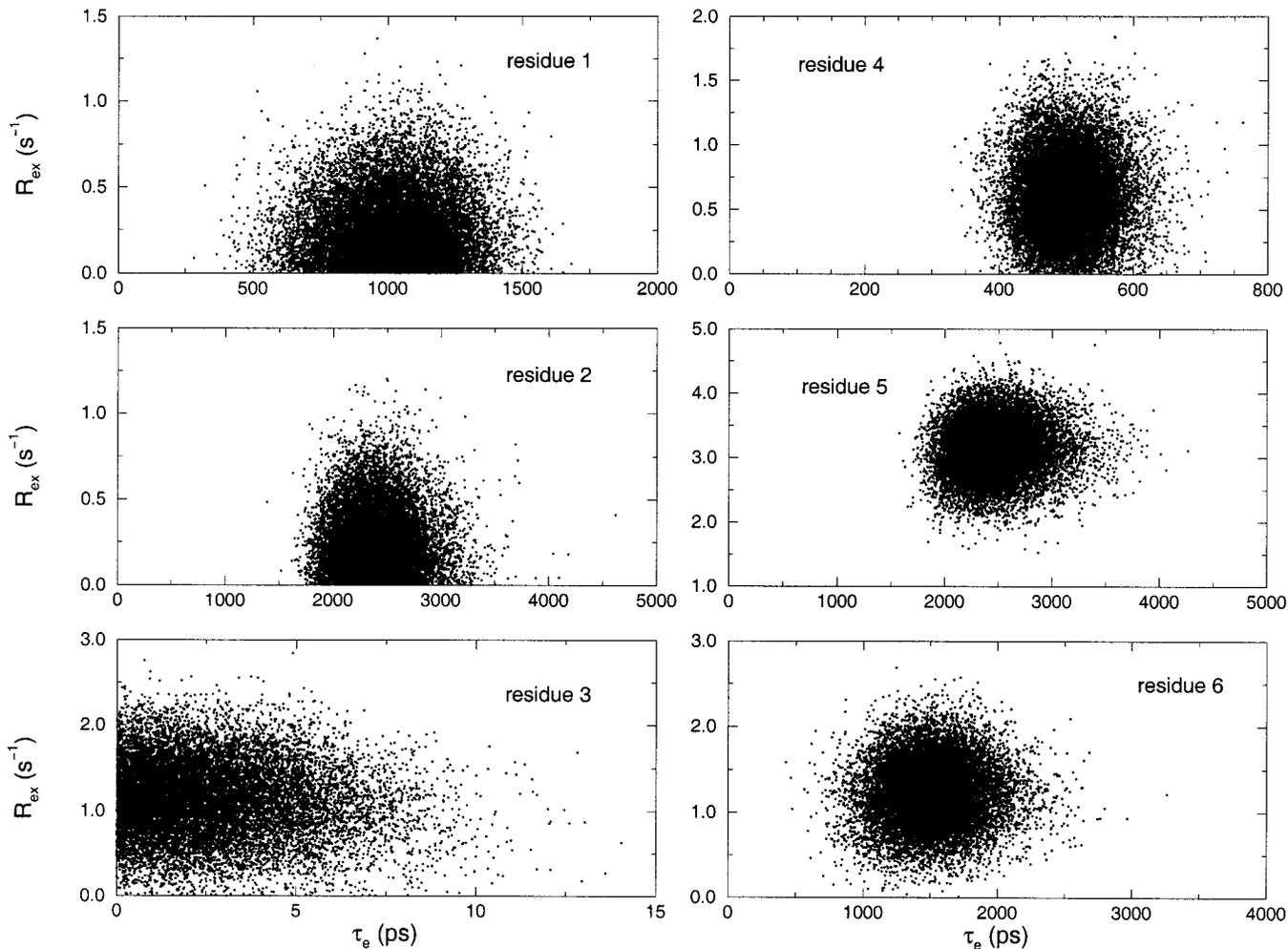
FIG. 4. Monte Carlo samples generated from the probability density of Eq. [14] using *XRambo* for each of the six residues of Table 2. A total of 120,000 Monte Carlo samples were generated, of which every eighth sample is shown. Each panel corresponds to a projection of the Monte Carlo samples onto the τ_e–R_{ex} plane and is a representation of the marginal probability density $P(\tau_e, R_{ex}|R)$.

the $\tau_e = 0$ axis, but away from the $R_{ex} = 0$ axis, confirming that it is well fit by model 3 ($\tau_e = 0$ and $R_{ex} \neq 0$). However, it is also clear that in all cases there is significant probability density away from the $\tau_e = 0$ and $R_{ex} = 0$ axes. Therefore, rather than stating that a certain residue (e.g. residue 1) is "model 2," one may wish to simply report that R_{ex} is most likely less than 0.75 s^{-1} and that τ_e lies between 0.5 and 1.5 ns. Such upper and lower bounds can be estimated visually or determined automatically from the Monte Carlo samples using nonparametric statistical methods (58, 61). Alternatively, one could also describe the results in terms of a mixture of traditional models, for example, 80% probability of model 2 and 20% probability of model 4. Although one could proceed as in the traditional approach and reestimate the local parameters while holding τ_e and/or R_{ex} fixed at zero, this is not necessary, and may underestimate the true uncertainties by eliminating model-choice uncertainty.

2. Anisotropic axially symmetric tumbling. It has been noted in the literature that the presence of anisotropic tumbling can have insidious effects on the extracted motional parameters if the data are fit using an isotropic model (36, 62–64). One method that has been proposed for the identification of such effects involves estimating a "local" τ_m independently for each residue using the spectral density of Eq. [4] (36, 63), as we have done in Fig. 2. However, our approach can be easily extended to directly estimate $D_\parallel$ and $D_\perp$ using Eq. [5] if we assume that the diffusion tensor orientation is known. It should be noted that the derivation of Eq. [5] by Schurr *et al.* (36) assumes that the internal motions occur on a significantly faster time scale than the overall tumbling. Although a rigorous factorization of the correlation function along the lines of Eq. [1] for anisotropic tumbling is not possible in general (13, 36), we have found that Eq. [5] is able to estimate S^2 to within 5% accuracy from synthetic data generated with the anisotropic

<div align="center">

TABLE 3

Parameters and Synthetic Data Used for the Determination of $D_\parallel$ and $D_\perp{}^a$

</div>

	Residue number					
	1	2	3	4	5	6
S^2	0.84	0.64	0.71	0.79	0.62	0.82
τ_e (ns)	1.0	2.3	0.0	0.5	2.4	1.5
R_{ex} (s^{-1})^b	0.0	0.0	1.0	0.5	3.1	1.2
θ	30°	70°	50°	40°	35°	60°
R_1 (s^{-1}) @ 500 MHz	1.18	1.86	0.81	1.16	1.77	1.44
R_1 (s^{-1}) @ 800 MHz	0.76	1.31	0.42	0.73	1.27	0.94
R_2 (s^{-1}) @ 500 MHz	20.84	14.10	16.31	18.93	19.36	18.55
R_2 (s^{-1}) @ 800 MHz	26.34	17.52	21.89	24.51	28.28	24.87
NOE @ 500 MHz	0.438	0.594	0.814	0.047	0.596	0.578
NOE @ 800 MHz	0.638	0.772	0.858	0.263	0.776	0.739

^{a} The synthetic R_1, R_2, and NOE data were calculated using Eqs. [5], [6], [7], and [8] with $D_\parallel = 0.016$ ns^{-1} and $D_\perp = 0.008$ ns^{-1} and the values of S^2, τ_e, R_{ex}, and θ listed in the table.

^{b} We define R_{ex} to be the value of $\omega_N^2\Phi_{ex}$ at a proton frequency of 500 MHz.

diffusion-in-a-cone model (*13, 65, 66*) with internal motional correlation times in the range of $0 \leq \tau_{cone} \leq (3\,D_\parallel)^{-1}$ in the limit $D_\parallel \gg D_\perp$ (data not shown). These results show that internal motional parameters extracted using Eq. [5] can be meaningful even if τ_e is not much smaller than the correlation times for the overall tumbling. Therefore, as in the isotropic example above, we have chosen to use τ_e values of 1–2 ns in order to validate our formalism in this regime.

We proceeded in an analogous manner, using synthetic data for a protein with an anisotropic diffusion tensor (Table 3) to compute Monte Carlo samples from the joint density $P(S^2, \tau_e, R_{ex}, D_\parallel, D_\perp | R_i, \theta)$. The resulting bivariate marginal densities $P(D_\parallel, D_\perp | R_i, \theta)$ are shown in Fig. 5. If we impose the

constraint that $D_\parallel$ and $D_\perp$ must be equal for all residues, the resulting global probability density $P(D_\parallel, D_\perp | R, \theta)$ is shown in Fig. 6. It is clear that although the information content in the data for each residue is insufficient to estimate $D_\parallel$ and $D_\perp$ with adequate precision, the superposition of the information from all residues affords a reasonably precise estimate which is also accurate when compared to the values of $D_\parallel$ and $D_\perp$ used to generate the synthetic data (0.016 ns^{-1} and 0.008 ns^{-1}, respectively). It is interesting to note that the estimates of $D_\parallel$ and $D_\perp$
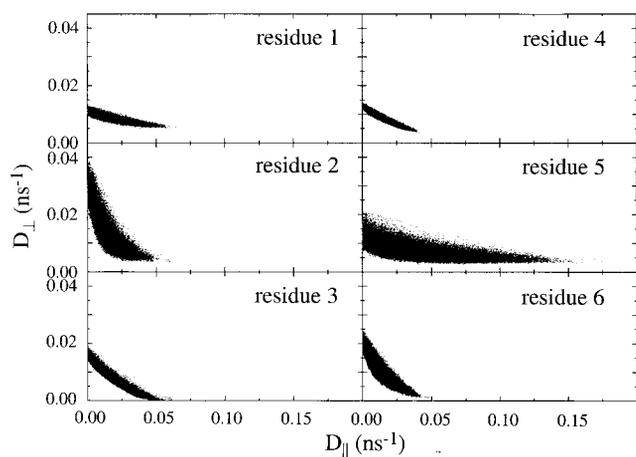
FIG. 5. Monte Carlo samples generated from the probability density $P(S^2, \tau_e, R_{ex}, D_\parallel, D_\perp | R_i, \theta)$ using *XRambo* and relaxation data R_i for each of the six residues shown in Table 3. A total of 400,000 Monte Carlo samples were generated for each residue, of which every eighth sample is shown. Each panel corresponds to a projection of the Monte Carlo samples onto the $D_\parallel$–$D_\perp$ plane and is a representation of the marginal probability density $P(D_\parallel, D_\perp | R_i, \theta)$.
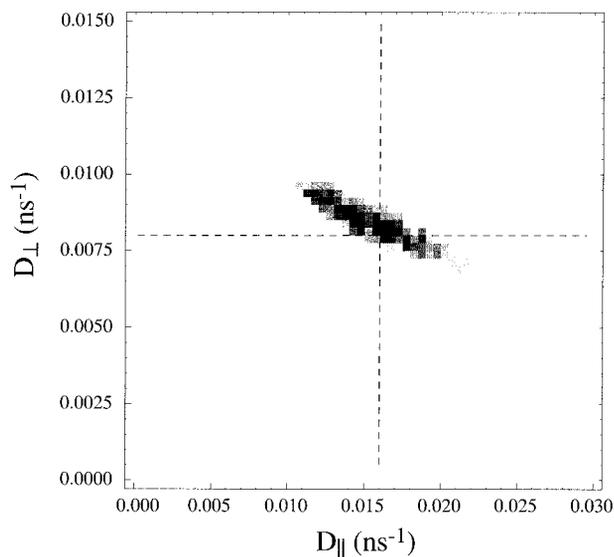
FIG. 6. A density plot of the two-dimensional histogram obtained by binning the Monte Carlo samples from $P(S^2, \tau_e, R_{ex}, D_\parallel, D_\perp | R_i, \theta)$ shown in Fig. 5 and multiplying the resulting bin counts for $i = 1, \ldots, 6$. The result is an approximate representation of the global parameter estimate $P(D_\parallel, D_\perp | R, \theta)$. The dotted lines denote the values of $D_\parallel$ and $D_\perp$ used to generate the synthetic data (Table 2), and the solid gray line corresponds to $D_\parallel = D_\perp$ (i.e., isotropic tumbling).
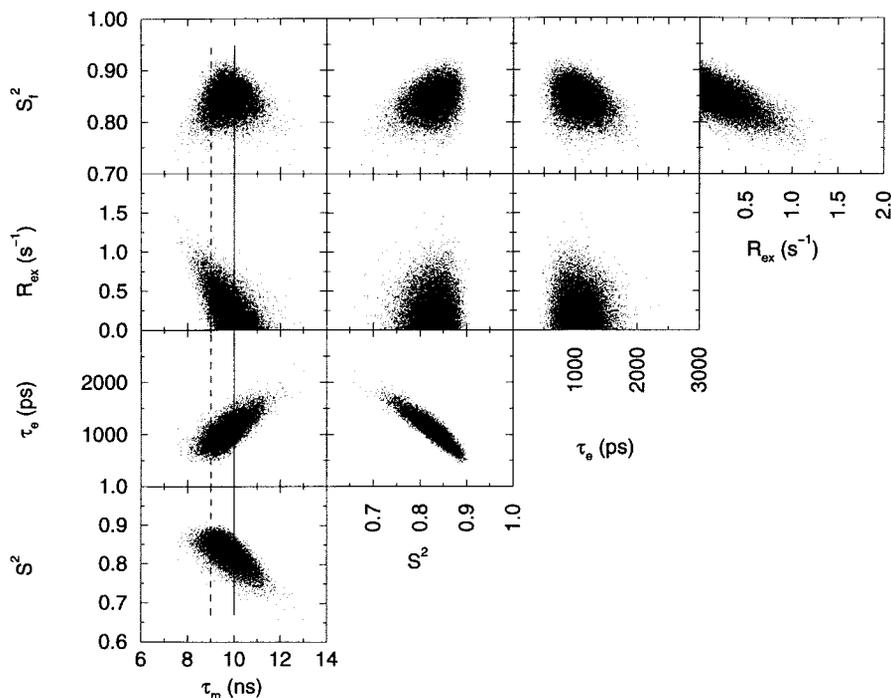
FIG. 7. Monte Carlo samples generated from the probability density $P(S^2, \tau_e, R_{ex}, S_f^2, \tau_m | R)$ with *XRambo* using synthetic data for a single hypothetical residue. R_1, R_2, and NOE data were generated at spectrometer ^{1}H frequencies of 400, 600, and 800 MHz using the parameters $S^2 = 0.84$, $\tau_e = 1.0$ ns, $R_{ex} = 0.0$ s^{-1}, $S_f^2 = 0.87$, and $\tau_m = 10.0$ ns ($R_1 = 2.08, 1.26, 0.95$ s^{-1}; $R_2 = 9.92, 10.85, 12.61$ s^{-1}; and NOE $= 0.497, 0.601, 0.695$ for 400, 600, and 800 MHz, respectively). A 5% relative error in each relaxation value was assumed. A total of 120,000 Monte Carlo samples were generated, of which every eighth sample is shown. Each panel corresponds to a projection of the full set of Monte Carlo samples onto a plane corresponding to each pair of parameters.

are not independent, but are approximately linearly correlated. Furthermore, in this particular case the uncertainties in $D_{\parallel}$ and $D_{\perp}$ are large enough that one cannot unambiguously eliminate the possibility of isotropic tumbling (Fig. 6), despite the fact that the "true" anisotropy is quite large ($D_{\parallel}/D_{\perp} = 2$). The interpretation of relaxation data in terms of Bayesian probability densities allows the use of products of marginal densities to efficiently estimate global $D_{\parallel}$ and $D_{\perp}$ values. This represents a significant improvement over classical statistical methods and circumvents the requirement for criteria which attempt to exclude residues which have internal motion outside of the extreme narrowing limit (*25, 62, 64, 67*).

3. Isotropic tumbling using the "extended model-free" spectral density. Although the application of our approach to the "extended model-free" spectral density (*40*) is in principle identical to the examples shown above, problems can arise in practice because one cannot unambiguously estimate values for both S^2 and S_f^2 as τ_e approaches zero, since S_f^2 loses physical meaning in that limit. This is not a substantial problem, however, since $P(\tau_m | R_i)$ can still be correctly estimated if the model is reparametrized so that one estimates the product $S^2 S_f^2$. It should be noted that this is not a shortcoming of our formalism, but rather is a reflection of the underlying physics. To demonstrate the feasibility of our approach for handling the extended model, we show the results obtained for the estima-

tion of $P(S^2, \tau_e, R_{ex}, S_f^2, \tau_m | R_i)$ for one set of synthetic data generated using a τ_e value of 1.0 ns (at a τ_m of 10 ns) (Fig. 7). Since in this case τ_e is sufficiently greater than zero, it is not necessary to reparameterize the model. The resulting marginal density of τ_m (which corresponds to the width of the "cloud" in the left-most column of plots in Fig. 7) is quite narrow, considering that it is derived from relaxation data for only one residue. A typical protein will likely have a number of residues which will allow similarly narrow τ_m estimates. As before, if the resulting $P(\tau_m | R_i)$ distributions have modes at nearly equal values of τ_m, then the global τ_m estimate $P(\tau_m | R)$ will be even narrower. Correspondingly, if the modes do not occur at nearly equal values of τ_m, then this may be evidence for systematic experimental error or the violation of the assumption of a single global dynamic parameter.

The projections of the five-dimensional probability distribution onto planes corresponding to each of the 10 possible orthogonal two-dimensional "viewpoints" shown in Fig. 7 provides a clear way to visualize the correlations among the model-free parameters. First of all, it is clear that τ_m and R_{ex} are linearly correlated, resulting in small nonzero R_{ex} contributions at smaller values of τ_m. This is a clear example of the model uncertainty and correlations among the local and global parameter estimates discussed above. Furthermore, it is also apparent that τ_e and S^2 show a positive and negative linear
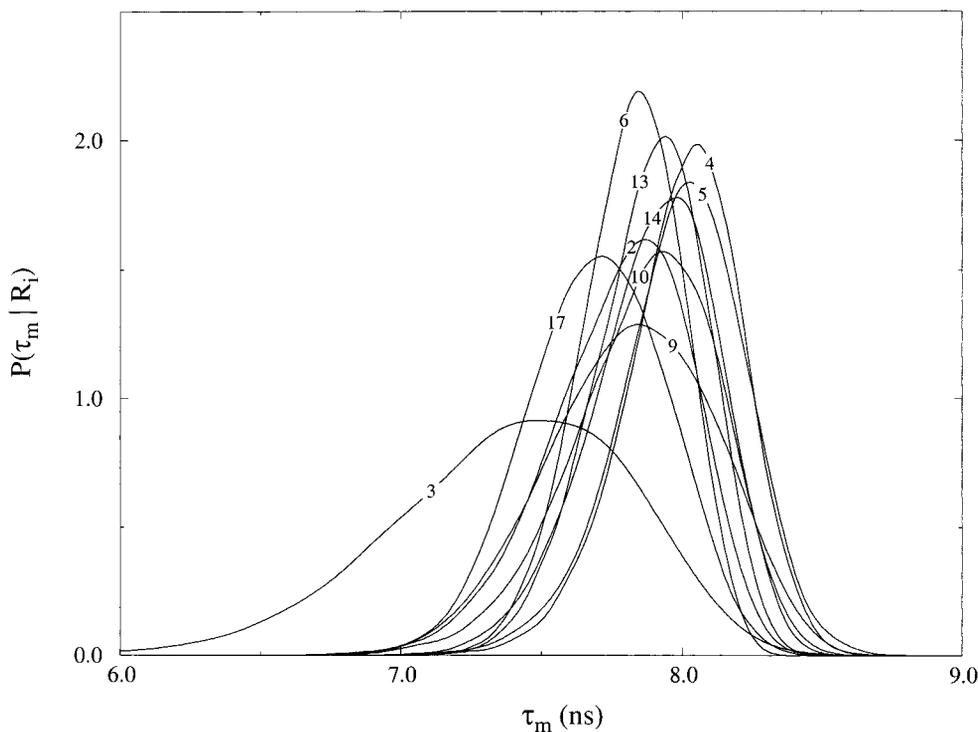
FIG. 8. The local marginal probability densities $P(\tau_m|R_i)$ calculated using the relaxation data R_i for helix 1 (residues 2–18) of the protein S100B($\beta\beta$) (K. G. Inman and D. J. Weber, unpublished data). A total of 19,500 Monte Carlo samples were generated for each residue using a "slicing" algorithm (*71*), and the curves were calculated as in Fig. 2. Data for residues 7, 8, 11, 16, and 18 were unavailable due to spectral overlap, and the joint distributions $P(S^2, \tau_e, R_{ex}, \tau_m|R_i)$ for residues 12 and 15 were unbounded with respect to τ_e and were eliminated from the analysis.

correlation, respectively, with τ_m. Thus, by underestimating the uncertainty in τ_m, either by artificially fixing its value at the value determined from R_2/R_1 or by forcing $R_{ex} = 0$ by choosing regression model 5, the true uncertainty in τ_e and S^2 would be significantly underestimated. This is particularly apparent if we consider the dotted line at $\tau_m = 9.0$ ns. In this case, we cannot obtain a good fit to the data unless we allow a small but finite R_{ex} contribution of 0.5 s^{-1}. Had we forced $R_{ex} = 0$ by choosing regression model 5, we would have been misled into a smaller uncertainty for τ_m, since τ_m values less than ≈ 9.0 ns are inconsistent with $R_{ex} = 0$. Also, we would have underestimated the uncertainties in S^2 and τ_e by partially eliminating values of $\tau_e < 1.0$ ns and $S^2 > 0.85$ due to the correlations among the dynamical parameters.

Based on these observations, it is clear that the traditional model-selection strategy can lead to a serious underestimation of the uncertainties in the dynamical parameters due to the neglect of ambiguities in model choice. By fitting to the most general model, one can avoid explicit model selection and thereby avoid this problem. Furthermore, the concept of model uncertainty arises quite naturally in Bayesian statistics, and we can use the concept of model uncertainty to describe the results of such a general fit. In fact, Bayesian methods allow the extension of this concept to situations in which the models are not nested and no such "most general model" exists (*51*).

4. Application to experimental data. In order to demonstrate the applicability of our method using experimental data, we have performed a local τ_m analysis for the first helix of the protein S100B($\beta\beta$). S100B($\beta\beta$) is a dimeric Ca^{2+}-binding protein (91 residues/monomer) which has been implicated in the neuropathologies of Down's syndrome and Alzheimer's disease (*68*), and the structure of which has been determined by NMR methods (*69*). R_1, R_2, and NOE data obtained at 400 and 600 MHz (K. G. Inman and D. J. Weber, unpublished data) were analyzed as described above using the original Lipari–Szabo spectral density including R_{ex} (Eqs. [4] and [12]). The calculation required approximately 60 s of CPU time per residue on an SGI R10000 computer. The results for the τ_m distributions in helix 1 of S100B($\beta\beta$) are shown in Fig. 8. All remaining local dynamical parameter estimates were within physically reasonable ranges ($S^2 \approx 0.8$–0.95, $\tau_e < 300$ ps, $R_{ex} < 1.5$ s^{-1} at 400 MHz). With the possible exception of residues 3 and 17, the marginal τ_m distributions are quite consistent, with an apparent τ_m of approximately 8 ns, which is reasonable for a 182-residue protein. It should be noted that this apparent τ_m need not correspond to a global τ_m, since S100B($\beta\beta$) may have nonnegligible anisotropy. However, even for a significantly anisotropic protein, the apparent τ_m values in a given helix should be nearly equal, since all of the amide vectors point in approximately the same direction (*64*).

The small deviations that are seen, such as those for residues 3 and 17, may be due to experimental errors in the relaxation data or more complex dynamics at the ends of the helix. Similar consistent behavior is observed for the other helices in the protein. A detailed analysis of the global and local dynamics of S100B($\beta\beta$) is currently under way using the methods described here, as well as using traditional methods based on model selection. The results of these analyses will be presented elsewhere.

CONCLUSIONS

We have shown that the use of a Bayesian statistical approach based on the product of marginal densities for the estimation of dynamical parameters from relaxation data is straightforward, powerful, and avoids the additional assumptions which are necessary in the standard procedures. In part, such assumptions were made necessary by the lack of sufficient data to fit the most general form of the Lipari–Szabo spectral density. In this paper, we have expanded the data set by including the R_1, R_2, and NOE data collected using multiple magnetic field strengths. The collection of relaxation data at multiple spectrometer frequencies does represent an additional investment of time and resources for the spectroscopist. We believe, however, that such an investment will have substantial benefits for the detailed quantitative interpretation in terms of Lipari–Szabo dynamical parameters. It should be noted that the use of multiple field strengths is *not* a requirement for the application of this formalism. Rather, the product of marginal densities formalism provides a natural and consistent means for combining the information contained in data obtained from multiple sources or methods, such as rotating-frame or cross-correlated relaxation measurements. Furthermore, we could also incorporate relaxation data from other nuclei, such as $^{13}C^{\alpha}$ relaxation, in order to further improve the global parameter estimates (*70*). This formalism in fact could be used for any nonlinear parameter estimation problem with a high-dimensional parameter space that can be separated into low-dimensional subspaces which are independent except for a small number of global parameters.

We have shown that by collecting more data the traditional model selection and global parameter estimation methods and their inherent assumptions can be avoided, allowing the reliable estimation of the uncertainties in the global tumbling parameters. Although one could construct classical statistical parameter estimation methods which avoid these assumptions, we show that a Bayesian method is computationally efficient and provides more insight into the information content of the data. The reliable estimation of the uncertainties in the global parameters is crucial, since there are significant correlations among the global and local parameters, and the estimation of the uncertainties in the local parameters are necessary for quantitative applications of the Lipari–Szabo formalism. Work is in progress to apply the product of marginal densities formalism described here to more experimental data and to implement these methods in a practical, user-friendly software package for use by the NMR spectroscopy community.

REFERENCES

1. J. A. McCammon and S. Harvey, "Dynamics of Proteins and Nucleic Acids," Cambridge Univ. Press, Cambridge, UK (1987).

2. C. L. Brooks, M. Karplus, and B. M. Pettit, "Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics," Advances in Chemical Physics, Vol. 71, Wiley, New York (1988).

3. O. Jardetzky, Protein dynamics and conformational transitions in allosteric proteins, *Prog. Biophys. Mol. Biol.* **65**, 171–219 (1996).

4. G. Wagner, S. Hyberts, and J. W. Peng, Study of protein dynamics by NMR, *in* "NMR of Proteins" (G. M. Clore and A. M. Gronenborn, Eds.), pp. 220–257, Macmillan, London (1993).

5. A. G. Palmer, III, Probing molecular motions by NMR, *Curr. Opinion Struct. Biol.* **7**, 732–737 (1997).

6. A. Abragam, "Principles of Nuclear Magnetism," Oxford Press, 1961.

7. A. G. Palmer, III, J. Williams, and A. McDermott, Nuclear magnetic resonance studies of biopolymer dynamics, *J. Phys. Chem.* **100**, 13,293–13,310 (1996).

8. R. M. Levy, M. Karplus, and P. G. Wolynes, NMR relaxation parameters in molecules with internal motion: Exact Langevin trajectory results compared with simplified relaxation models, *J. Am. Chem. Soc.* **103**, 5998–6011 (1981).

9. R. M. Levy and J. Keepers, Computer simulations of protein dynamics: Theory and experiment, *Comm. Mol. Cell. Biophys.* **3**, 273–295 (1986).

10. J. W. Peng and G. Wagner, Mapping of spectral density functions using heteronuclear NMR relaxation measurements, *J. Magn. Reson.* **98**, 308–332 (1992).

11. R. Ishima and K. Nagayama, Protein backbone dynamics revealed by quasi spectral density function analysis of amide N-15 nuclei, *Biochemistry* **34**, 3162–3171 (1995).

12. N. A. Farrow, O. Zhang, A. Szabo, D. A. Torchia, and L. E. Kay, Spectral density function mapping using ^{15}N relaxation data exclusively, *J. Biomol. NMR* **6**, 153–162 (1995).

13. 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).

14. 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).

15. K. T. Dayie, G. Wagner, and J.-F. Lefèvre, Theory and practice of nuclear spin relaxation in proteins, *Annu. Rev. Phys. Chem.* **47**, 243–282 (1996).

16. J. J. Barchi, Jr., B. Grasberger, A. M. Gronenborn, and G. M. Clore, Investigation of the backbone dynamics of the IgG-binding domain of streptococcal protein G by heteronuclear two-dimensional ^{1}H–^{15}N nuclear magnetic resonance spectroscopy, *Protein Sci.* **3**, 15–21 (1994).

17. D. Fushman, S. Cahill, and D. Cowburn, The main-chain dynamics of the dynamin pleckstrin homology (PH) domain in solution: Analysis of ^{15}N relaxation with monomer/dimer equilibrium, *J. Mol. Biol.* **266**, 173–194 (1997).

18. K.-B. Wong, A. R. Fersht, and S. M. V. Freund, NMR ^{15}N relaxation and structural studies reveal slow conformational exchange in barstar C40/82A, *J. Mol. Biol.* **268**, 494–511 (1997).

19. C. Baldellon, J.-R. Alattia, M.-P. Strub, T. Pauls, M. W. Berchtold, A. Cavé, and A. Padilla, ^{15}N NMR relaxation studies of calcium-loaded parvalbumin show tight dynamics compared to those of other EF-hand proteins, *Biochemistry* **37**, 9964–9975 (1998).

20. A. E. Meekhof, S. J. Hamill, V. L. Arcus, J. Clarke, and S. M. V. Freund, The dependence of chemical exchange on boundary selection in a fibronectin type III domain from human tenascin, *J. Mol. Biol.* **282**, 181–194 (1998).

21. M. Ernst and R. R. Ernst, Heteronuclear dipolar cross-correlated cross relaxation for the investigation of side-chain motions, *J. Magn. Reson. A* **110**, 202–213 (1994).

22. M. W. F. Fischer, L. Zeng, Y. Pang, W. Hu, A. Majumdar, and E. R. P. Zuiderweg, Experimental characterization of models for backbone picosecond dynamics in proteins: Quantification of NMR auto- and cross-correlation relaxation mechanisms involving different nuclei of the peptide plane, *J. Am. Chem. Soc.* **119**, 12629–12642 (1997).

23. D. Yang, A. Mittermaier, Y.-K. Mok, and L. E. Kay, A study of protein side-chain dynamics from new ^{2}H auto-correlation and ^{13}C cross-correlation NMR experiments: Application to the N-terminal SH3 domain from drk, *J. Mol. Biol.* **276**, 939–954 (1998).

24. J. Engelke and H. Rüterjans, Dynamics of beta-CH and beta-CH2 groups of amino acid side chains in proteins, *J. Biomol. NMR* **11**, 165–183 (1998).

25. C. D. Kroenke, J. P. Loria, L. K. Lee, M. Rance, and A. G. Palmer, III, Longitudinal and transverse ^{1}H-^{15}N dipolar/^{15}N chemical shift anisotropy relaxation interference: Unambiguous determination of rotational diffusion tensors and chemical exchange effects in biological macromolecules, *J. Am. Chem. Soc.* **120**, 7905–7915 (1998).

26. B. Brutscher, N. R. Skrynnikov, T. Bremi, R. Brüschweiler, and R. R. Ernst, Quantitative investigation of dipole-CSA cross-correlated relaxation by ZQ/DQ spectroscopy, *J. Magn. Reson.* **130**, 346–351 (1998).

27. T. Szyperski, P. Luginbühl, G. Otting, P. Güntert, and K. Wüthrich, Protein dynamics studied by rotating frame ^{15}N spin relaxation times, *J. Biomol. NMR* **3**, 151–164 (1993).

28. S. Zinn-Justin, P. Berthault, M. Guenneugues, and H. Desvaux, Off-resonance rf fields in heteronuclear NMR: application to the study of slow motions, *J. Biomol. NMR* **10**, 363–372 (1997).

29. M. Akke, J. Liu, J. Cavanagh, H. P. Erickson, and A. G. Palmer, III, Pervasive conformational fluctuations on microsecond time scales in a fibronectin type III domain, *Nature Struct. Biol.* **5**, 55–59 (1998).

30. F. A. A. Mulder, R. A. de Graaf, R. Kaptein, and R. Boelens, An off-resonance rotating frame relaxation experiment for the investigation of macromolecular dynamics using adiabatic rotations, *J. Magn. Reson.* **131**, 351–357 (1998).

31. M. Akke, R. Brüschweiler, and A. G. Palmer, III, NMR order parameters and free energy: an analytical approach and its application to cooperative Ca^{2+} binding by calbindin D$_{9k}$, *J. Am. Chem. Soc.* **115**, 9832–9833 (1993).

32. D. Yang and L. E. Kay, Contributions to conformational entropy arising from bond vector fluctuations measured from NMR-derived order parameters: application to protein folding, *J. Mol. Biol.* **263**, 369–382 (1996).

33. D. Yang, Y.-K. Mok, J. D. Forman-Kay, N. A. Farrow, and L. E. Kay, Contributions to protein entropy and heat capacity from bond vector motions measured by NMR spin relaxation, *J. Mol. Biol.* **272**, 790–804 (1997).

34. D. Jin, F. Figueirido, G. T. Montelione, and R. M. Levy, Impact of the precision in NMR relaxation measurements on the interpretation of protein dynamics, *J. Am. Chem. Soc.* **119**, 6923–6924 (1997).

35. D. Jin, M. Andrec, G. T. Montelione, and R. M. Levy, Propagation of experimental uncertainties using the Lipari–Szabo model-free analysis of protein dynamics, *J. Biomol. NMR* **12**, 471–492 (1998).

36. J. M. Schurr, 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).

37. D. E. Woessner, Nuclear spin relaxation in ellipsoids undergoing rotational Brownian motion, *J. Chem. Phys.* **37**, 647–654 (1962).

38. J. W. Peng and G. Wagner, Frequency spectrum of NH bonds in eglin c from spectral density mapping at multiple fields, *Biochemistry* **34**, 16733–16752 (1995).

39. I. Q. H. Phan, J. Boyd, and I. D. Campbell, Dynamic studies of a fibronectin type I module pair at three frequencies: anisotropic modelling and direct determination of conformational exchange, *J. Biomol. NMR* **8**, 369–378 (1996).

40. G. M. Clore, A. Szabo, A. Bax, L. E. Kay, P. C. Driscoll, and A. M. Gronenborn, Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins, *J. Am. Chem. Soc.* **112**, 4989–4991 (1990).

41. 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).

42. V. Y. Orekhov, D. E. Nolde, A. P. Golovanov, D. M. Korzhnev, and A. S. Arseniev, Processing of heteronuclear NMR relaxation data with the new software DASHA, *Appl. Magn. Reson.* **9**, 581–588 (1995).

43. 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).

44. 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).

45. D. M. Korzhnev, V. Y. Orekhov, and A. S. Arseniev, Model-free approach beyond the borders of its applicability, *J. Magn. Reson.* **127**, 184–191 (1997).

46. 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).

47. Y.-C. Li and G. T. Montelione, Human type-α transforming growth factor undergoes slow conformational exchange between multiple backbone conformations as characterized by nitrogen-15 relaxation measurements, *Biochemistry* **34**, 2408–2423 (1995).

48. J. C. Kiefer, "Introduction to Statistical Inference," Springer-Verlag, New York (1987).

49. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, "Numerical Recipes in C: The Art of Scientific Computing," Cambridge Univ. Press, Cambridge, UK (1992).

50. H. Jeffreys, "Theory of Probability," Oxford Univ. Press, London (1961).

51. D. S. Sivia, "Data Analysis: A Bayesian Tutorial," Oxford Univ. Press, Oxford (1996).

52. G. D. Henry, J. H. Weiner, and B. D. Sykes, Backbone dynamics of a model membrane protein: ^{13}C NMR spectroscopy of alanine methyl groups in detergent-solubilized M13 coat protein, *Biochemistry,* **25,** 590–598 (1986).

53. A. J. Weaver, M. D. Kemple, and F. G. Prendergast, Tryptophan sidechain dynamics in hydrophobic oligopeptides determined by the use of ^{13}C nuclear magnetic resonance spectroscopy, *Biophys. J.* **54,** 1–15 (1988).

54. 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).

55. N. Tjandra, H. Kuboniwa, H. Ren, and A. Bax, Rotational dynamics of calcium-free calmodulin studied by ^{15}N-NMR relaxation measurements, *Eur. J. Biochem.* **230,** 1014–1024 (1995).

56. D. Fushman, R. Weisemann, H. Thüring, and H. Rüterjans, Backbone dynamics of ribonuclease T1 and its complex with 2'GMP studied by two-dimensional heteronuclear NMR spectroscopy, *J. Biomol. NMR* **4,** 61–78 (1994).

57. A. J. Wand, J. L. Urbauer, R. P. McEvoy, and R. J. Bieber, Internal dynamics of human ubiquitin revealed by ^{13}C-relaxation studies of randomly fractionally labeled protein, *Biochemistry* **35,** 6116–6125 (1996).

58. M. Andrec and J. H. Prestegard, A Metropolis Monte Carlo implementation of Bayesian time-domain parameter estimation: Application to coupling constant estimation from antiphase multiplets, *J. Magn. Reson.* **130,** 217–232 (1998).

59. A. J. Izenman, Recent developments in nonparametric density estimation, *J. Am. Stat. Assoc.* **86,** 205–224 (1991).

60. G. R. Terrell, The maximal smoothing principle in density estimation, *J. Am. Stat. Assoc.* **85,** 470–477 (1990).

61. J. Besag, P. Green, D. Higdon, and K. Mengersen, Bayesian computation and stochastic systems, *Stat. Sci.* **10,** 3–41 (1995).

62. N. Tjandra, P. Wingfield, S. Stahl, and A. Bax, Anisotropic rotational diffusion of perdeuterated HIV protease from ^{15}N NMR relaxation measurements at two magnetic fields, *J. Biomol. NMR* **8,** 273–284 (1996).

63. P. Luginbühl, K. V. Pervushin, H. Iwai, and K. Wüthrich, Anisotropic molecular rotational diffusion in ^{15}N spin relaxation studies of protein mobility, *Biochemistry* **36,** 7305–7312 (1997).

64. G. Barbato, M. Ikura, L. E. Kay, R. W. Pastor, and A. Bax, Backbone dynamics of calmodulin studied by ^{15}N relaxation using inverse detected two-dimensional NMR spectroscopy: The central helix is flexible, *Biochemistry* **31,** 5269–5278 (1992).

65. G. Lipari and A. Szabo, Padé approximants to correlation functions for restricted rotational diffusion, *J. Chem. Phys.* **75,** 2971–2976 (1981).

66. G. Lipari and A. Szabo, Nuclear magnetic resonance relaxation in nucleic acid fragments: Models for internal motion, *Biochemistry* **20,** 6250–6256 (1981).

67. N. Tjandra, S. E. Feller, R. W. Pastor, and A. Bax, Rotational diffusion anisotropy of human ubiquitin from ^{15}N NMR relaxation, *J. Am. Chem. Soc.* **117,** 12562–12566 (1995).

68. B. W. Schäfer and C. W. Heizmann, The S100 family of EF-hand calcium-binding proteins: Functions and pathology, *TIBS* **21,** 134–140 (1996).

69. A. C. Drohat, J. C. Amburgey, F. Abildgaard, M. R. Starich, D. Baldisseri, and D. J. Weber, Solution structure of rat Apo-S100B($\beta\beta$) as determined by NMR spectroscopy, *Biochemistry* **35,** 11577–11588 (1996).

70. L. K. Lee, M. Rance, W. J. Chazin, and A. G. Palmer, III, Rotational diffusion anisotropy of proteins from simultaneous analysis of ^{15}N and ^{13}C$^{\alpha}$ nuclear spin relaxation, *J. Biomol. NMR* **9,** 287–298 (1997).

71. R. M. Neal, "Markov Chain Monte Carlo Methods Based on 'Slicing' the Density Function," Department of Statistics, University of Toronto, Technical Report 9722 (1997).