Journal of Biomolecular NMR, **25:** 325–333, 2003.
KLUWER/ESCOM

Model selection for the interpretation of protein side chain methyl dynamics

Wing-Yiu Choy & Lewis E. Kay
The Protein Engineering Network Center of Excellence and Departments of Medical Genetics and Microbiology, Biochemistry, and Chemistry, University of Toronto, Toronto, Ontario, Canada M5S 1A8

Received 22 November 2002; Accepted 13 January 2003

Key words: backbone dynamics, ^{13}C spin relaxation, ^{2}H spin relaxation, methyl side chain dynamics, ^{15}N spin relaxation

Abstract

A number of different dynamics models are considered for fitting ^{13}C and ^{2}H side chain methyl relaxation rates. It is shown that in cases where nanosecond time scale dynamics are present the extended Lipari–Szabo model which is explicitly parameterized to include the effects of slow motions can produce wide distributions of fitting parameters even in cases where the errors are relatively small and large numbers of relaxation rates are considered. In contrast, fits of ^{15}N backbone dynamics using this model are far more robust. The origin of this difference is analyzed and can be explained by the different functional forms of the spectral density in these two cases. The utility of a number of models for the analysis of methyl side chain dynamics is presented.

Introduction

Much of structural biology is predicated on the determination of static three-dimensional structures of macromolecules and there is currently tremendous interest in generating pictures of as many proteins as possible (Edwards et al., 2000). The structures that have been produced have, in many cases, had a significant impact on our understanding of function and in the design of important pharmaceuticals. However, it is well recognized that proteins are not static entities and that dynamics are often critical for processes such as enzyme catalysis, molecular recognition, and ligand binding, for example (Brooks et al., 1988; Fersht, 1985). Thus, a complete understanding of protein function can only be achieved through a description of how structure changes with time and the relation between dynamics and function continues to be the focus of many biophysical studies.

Nuclear magnetic resonance (NMR) spectroscopy is a powerful technique for the study of protein dynamics (Ishima and Torchia, 2000; Kay, 1998; Palmer et al., 1996). Over the past decade a wide array of methods have been developed for probing protein motions on a site-specific basis spanning a broad range of time scales. The most frequently employed experiments measure backbone ^{15}N spin relaxation rates, providing information about the motions of amide ^{15}N-^{1}HN bond vectors. More recently relaxation of $^{13}C_{\alpha}$ (Engelke and Rüterjans, 1995; Yamazaki et al., 1994) and ^{13}CO (Cordier et al., 1996; Dayie and Wagner, 1997; Zeng et al., 1996) spins have been used as additional probes of backbone dynamics. The dynamic picture of a protein, however, cannot be completed without information about motion of side chains and several significant advancements have been made in the past few years in this regard. These include the development of both biosynthetic approaches to produce proteins with labeling patterns that are amenable to study and the design and optimization of new NMR experiments. For example, ^{13}C labeling strategies have been devised in which ^{13}C is enriched at alternate positions along a side chain (LeMaster and Kushlan, 1996), eliminating ^{13}C-^{13}C couplings which complicate measurement of ^{13}C spin relaxation rates (Lee et al., 1997; LeMaster and Kushlan, 1996). Other labeling strategies have focused on the production of $^{13}CHD_2$ methyls for subsequent analysis by ^{13}C re-

laxation (Ishima et al., 1999). Our laboratory has developed an approach for studying side chain dynamics in which ^{2}H spin relaxation rates are measured at methylene (^{13}CHD) and methyl positions ($^{13}CH_2D$) in proteins that are labeled uniformly with ^{13}C and fractionally with ^{2}H (Muhandiram et al., 1995; Yang et al., 1998). In principle, five relaxation rates can be obtained per site (Millet et al., 2002), providing an opportunity to characterize the dynamics in a detailed manner. The methodology has also been extended to studies of side chain dynamics of proteins in unfolded states (Choy and Kay, 2003; Muhandiram et al., 1997).

Once relaxation rates have been measured they can be interpreted using the Lipari-Szabo model (Lipari and Szabo, 1982a, b) to obtain, in the simplest case, an order parameter and a correlation time characterizing the amplitude and time scale of motions at a particular site, respectively. Clore et al. (1990) showed in the context of ^{15}N spin relaxation data that such a model could be readily extended to account for internal dynamics on a number of different time scales (extended Lipari–Szabo model) and this model is in widespread use today in the analysis of *both* backbone and side chain relaxation data. It is not clear however that a model that is well suited for fitting relaxation of a particular probe (say ^{15}N) in the protein backbone would be equally appropriate for the analysis of spin relaxation data recorded on a methyl ^{13}C or ^{2}H nucleus. For example, in a recent study of methyl containing side chain dynamics in protein L we found that fitting ^{2}H spin relaxation data derived from $^{13}CH_2D$ methyls to an extended Lipari–Szabo model used successfully in the analysis of ^{15}N relaxation rates was unstable in the sense that often wide ranges of dynamics parameters were obtained from Monte-Carlo-type error analyses (Millet et al., 2002; Skrynnikov et al., 2002). Excellent fits of the methyl data could be obtained, however, using simpler spectral density functions (see below). This has prompted us to investigate the utility of a number of different models for fitting side chain methyl dynamics data (either ^{2}H or ^{13}C) and, further, to explore what the inherent differences between side chain methyl and backbone amide relaxation data might be that lead to the suitability of a particular model in one case but not in another.

Results and discussion

As described above, the Lipari–Szabo model-free analysis is the most widely used approach for the translation of relaxation data to site-specific motional parameters characterizing fast internal dynamics in macromolecules (Lipari and Szabo, 1982a, b). The original approach assumes that the overall molecular tumbling is isotropic and independent of internal motions and that the internal motions can be described by a single exponential process. Under these assumptions, the correlation function adopts a two-exponent form and the corresponding spectral density (referred to in what follows as LS-2) is expressed as the sum of two Lorentzian terms:

$$J(\omega) = \frac{S^2\tau_r}{1+(\omega\tau_r)^2} + \frac{(1-S^2)\tau}{1+(\omega\tau)^2}, \qquad \frac{1}{\tau} = \frac{1}{\tau_r} + \frac{1}{\tau_f}, \tag{1}$$

where S is an order parameter describing the amplitude of the internal motions with correlation time τ_f and τ_r is the overall molecular tumbling time of the protein. When the dynamics are more complex with both fast (ps) and slower (perhaps ns) time scale motions, such as might be the case for a methyl group attached to a side chain, the correlation function can be approximated by $c(t) = \{\alpha S_f^2 + (1-\alpha S_f^2)\exp(-t/\tau_f)\}\{S_s^2 + (1-S_s^2)\exp(-t/\tau_s)\}\{\exp(-t/\tau_r)\}$, from which a spectral density function, LS-4, of the form,

$$J(\omega) = \alpha S_f^2 S_s^2 \frac{\tau_0}{1+(\omega\tau_0)^2} + \alpha S_f^2(1-S_s^2)\frac{\tau_1}{1+(\omega\tau_1)^2} + (1-\alpha S_f^2)S_s^2\frac{\tau_2}{1+(\omega\tau_2)^2} + (1-\alpha S_f^2)(1-S_s^2)\frac{\tau_3}{1+(\omega\tau_3)^2}, \tag{2}$$

$$\frac{1}{\tau_0} = \frac{1}{\tau_r}, \quad \frac{1}{\tau_1} = \frac{1}{\tau_r} + \frac{1}{\tau_s}, \quad \frac{1}{\tau_2} = \frac{1}{\tau_r} + \frac{1}{\tau_f}, \quad \frac{1}{\tau_3} = \frac{1}{\tau_r} + \frac{1}{\tau_s} + \frac{1}{\tau_f}$$

is obtained, where the slower internal motions are parameterized by S_s and τ_s (Skrynnikov et al., 2002). In Equation 2 the value of α is set to 1 to describe the relaxation of backbone ^{15}N or ^{13}C spins and to 1/9 for methyl ^{13}C or ^{2}H spins (assuming that the methyl group adopts ideal tetrahedral geometry). As with the simple LS-2 model (Equation 1) independence of overall and internal dynamics are assumed, as is independence of the fast and slow time scale internal motions. In the derivation of Equation 2 we have made no attempt to separate the contributions from rotation about the assumed methyl three-fold axis and fast (ps) time scale motions of this axis, since both processes occur with similar time constants. Note that in the limit where $\tau_f \ll \tau_s$ Equation 2 simplifies to the expression given in Clore et al. (1990). It is also of interest to note that in the infrequent case that methyl rotation is slow and on the same time scale as the slow internal

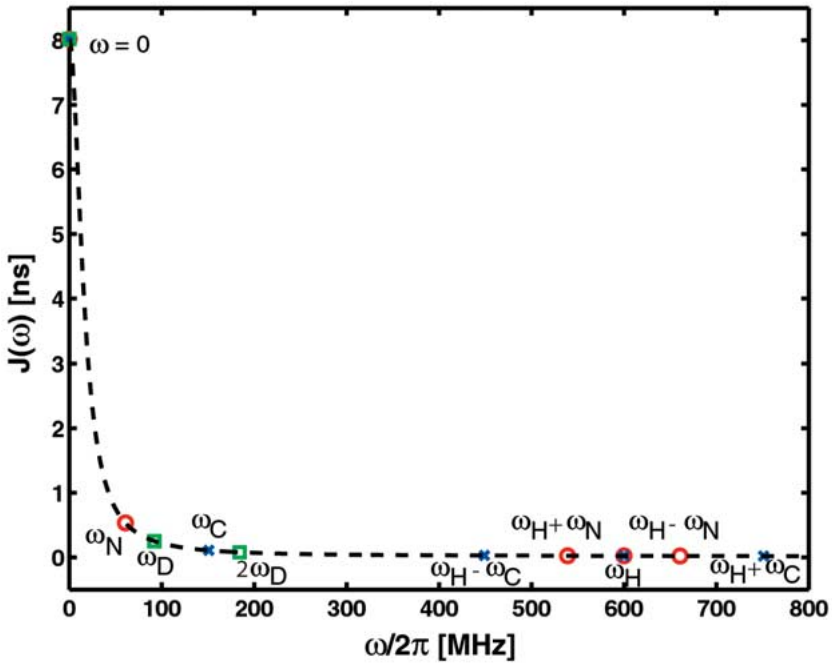

Figure 1. Plot of $J(\omega)$, using the spectral density form given by Equation 1, vs. frequency (dashed line). S^2, τ_r and τ_f are set to 0.8, 10 ns and 100 ps, respectively. The symbols (○), (□) and (×) are used to indicate the frequencies at which the spectral density is evaluated in expressions for ^{15}N, ^{2}H and ^{13}C relaxation at 14.1 T (600 MHz ^{1}H frequency).

dynamics Equation 2 is still valid with $\{\tau_f, S_f\}$ and $\{\tau_s, S_s\}$ interchanged.

The amount of dynamics information that can be extracted from any relaxation experiment is limited by sampling of only a small number of discrete values of the spectral density function. The relaxation rates of different spins (^{15}N, ^{13}C, ^{2}H etc.), probe the spectral density function at different frequencies and thus the sensitivity of measurements to dynamic processes will vary depending on what nucleus is studied (Lee et al., 1999). Figure 1 shows the frequencies at which the spectral density function is sampled in ^{15}N (red), ^{13}C (blue) and ^{2}H (green) spin relaxation measurements. The commonly employed ^{15}N R_1, R_2 and steady state ^{1}H-^{15}N NOE measurements sample a wide range of frequencies extending from 0 to $\omega_H - \omega_N$. On the other hand, ^{2}H spin relaxation rates are dominated by the quadrupolar relaxation mechanism, which samples the spectral density function over a relatively more narrow range ($0 - 2\omega_D$). Moreover, in the case of side chain methyl relaxation rates the spectral density function includes a factor of 1/9 which accounts for the methyl three-fold rotation (see Equation 2) that is absent in the analysis of backbone ^{15}N data. Thus, spin-dependent differences in both the functional form and the frequency sampling of the spectral density suggest that a spectral density function that performs well in the analysis of ^{15}N backbone data may not be robust in applications involving side chain dynamics.

By means of example we have simulated methyl ^{13}C R_1, R_2 and steady state ^{1}H-^{13}C NOE relaxation data recorded at 400, 500, 600 and 800 MHz (^{1}H) frequencies using the spectral density function given in Equation 2 ($\alpha = 1/9$) with $\{S_f^2, \tau_f, S_s^2, \tau_s\} = \{0.8, 50\text{ ps}, 0.5, 1.0\text{ ns}\}$ and $\tau_r = 6$ ns. Gaussian noise with a standard deviation of 3% of the relaxation parameter in question (R_1, R_2) and 5% for the NOE was added to each of the relaxation parameters and data from all fields (400–800 MHz) were fit together using Equation 2, corresponding to the case where 12 pieces of relaxation data are obtained per methyl. In order to ensure that the global minimum was selected we have carried out a four-dimensional grid search in which S_f^2 and S_s^2 were varied from 0.0 to 1.0 in steps of 0.02, τ_f was varied from 40 to 100 ps with a grid size of 5 ps and τ_s was incremented from 50 ps to 6 ns in steps of 20 ps. Once a minimum was obtained the resulting parameter set was used as input to a simplex minimization procedure (Press et al., 1988). The process was repeated 1000 times and the distribution of fitting

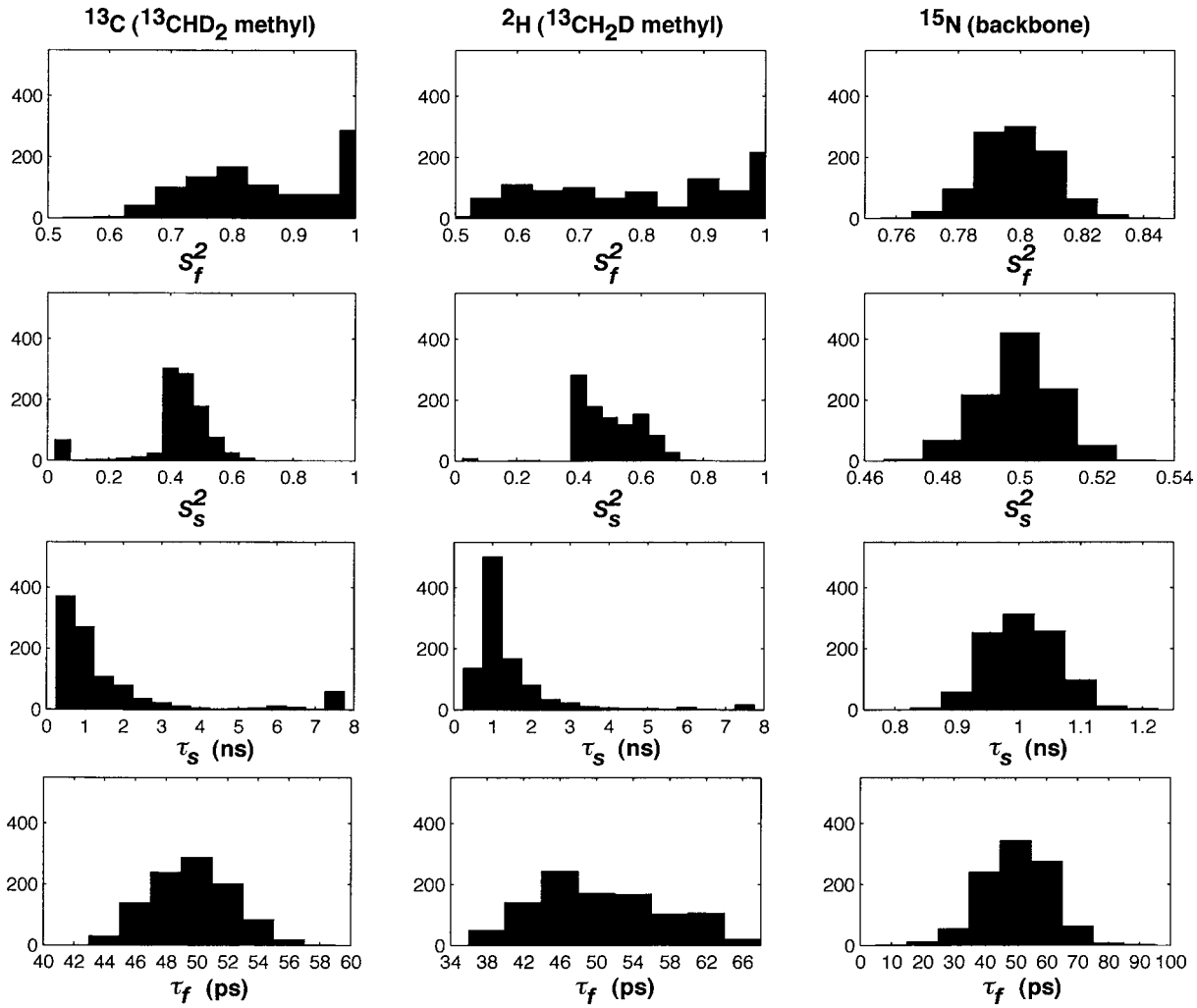

Figure 2. Histograms of S_f^2, S_s^2, τ_s and τ_f derived from fits of ^{13}C methyl (first column), ^{2}H methyl (second column) and ^{15}N backbone (third column) spin relaxation data. Relaxation parameters were simulated using the LS-4 spectral density function with $\alpha = 1$ and 1/9 for backbone and methyl data respectively and with input dynamics parameters of $\{S_f^2, \tau_f, S_s^2, \tau_s\} = \{0.8, 50\text{ ps}, 0.5, 1.0\text{ ns}\}$ and $\tau_r = 6$ ns. Errors were added to the relaxation data and the rates subsequently fit using the same LS-4 model, as described in Results and discussion, to give the distributions shown. All values of τ_s obtained from fits that are greater than 8 ns were included in the final histogram bar (between 7–8 ns) in the case of the ^{13}C and ^{2}H relaxation data. The methyl ^{13}C relaxation rates were simulated assuming a $^{13}CHD_2$ methyl, with auto-correlated dipolar contributions from both ^{1}H-^{13}C and ^{2}H-^{13}C interactions included, as well as contributions from chemical shift anisotropy, as described in Ishima and Torchia (Ishima et al., 2001). A value of 1.135 Å was used for the ^{1}H-^{13}C bond length along with an assumed axially symmetric chemical shift tensor with $\Delta\sigma = 25$ ppm. Methyl ^{2}H rates were simulated assuming a $^{13}CH_2D$ methyl, with equations describing the relaxation of the ^{2}H spin given in Millet et al. (2002). Backbone ^{15}N spin relaxation rates were calculated as described in Farrow et al. (1994) with $\Delta\sigma = -170$ ppm and $r_{HN} = 1.02$ Å.

parameters $\{S_f^2, \tau_f, S_s^2, \tau_s\}$ obtained in this manner is illustrated in the first column of Figure 2, indicating that quite a wide range of values can be obtained. A similar analysis was repeated for a set of simulated methyl ^{2}H spin relaxation rates including R_{D_z} (longitudinal), R_{D_+} (in-phase transverse), $R_{D_+D_z+D_zD_+}$ (anti-phase transverse), $R_{3D_z^2-2}$ (quadrupolar order) and $R_{D_+^2}$ (double quantum) using the dynamics parameters and four spectrometer fields listed above. All rates were assumed to have errors of 3% and in this case 20 rates (4 fields) were fit simultaneously. As with the ^{13}C data, a wide distribution of dynamics parameters were extracted (Figure 2, second column), despite the fact that the same model that was used to generate the data was also used to fit it. A further set of simulations with different values of input order parameters

and correlation times shows that the output parameter distribution widths are, in most cases, quite similar and are significantly wider than those observed from fits of ^{15}N data (see below).

In a final set of simulations, backbone ^{15}N spin relaxation data sets comprising R_1 (with 3% error), R_2 (3% error) and NOE (5% error) values at 400, 500, 600 and 800 MHz were calculated using the spectral density function of Equation 2 with α set to 1. The distributions of motional parameters obtained from fits of the relaxation data are shown in the third column of Figure 2. In contrast to what is observed for fits of methyl relaxation rates, narrow ranges of $\{S_f^2, S_s^2, \tau_s\}$ are obtained, although for τ_f the width of the distribution is somewhat larger.

The extra factor of 1/9 in expressions for $J(\omega)$ in the case of methyl (but not ^{15}N) relaxation scales down terms which depend on $\{S_f^2, S_s^2, \tau_s\}$ (see the first two terms in Equation 2). This can be understood by noting that for a broad range of dynamics parameters the sum of the last two terms of $J(\omega)$ in Equation 2 is of the same order as the sum of the first two for non-zero values of ω_D that are important in the relaxation expressions. For example, for $\omega_D/2\pi = 77$ MHz (^{1}H frequency of 500 MHz) and for $\{S_f^2, \tau_f, S_s^2, \tau_s\} = \{0.8, 50\text{ ps}, 0.5, 1.0\text{ ns}\}$ and $\tau_r = 6$ ns as above, the sum of the first (second) two terms in Equation 2 is 61.0 ps (44.1 ps). Note that in the limit that $\tau_f \ll \tau_s, \tau_r$ the sum of the second pair of terms is to excellent approximation given by $(1 - \frac{1}{9}S_f^2)\tau_f/(1 + (\omega\tau_f)^2)$, which is independent of S_s^2, τ_s and to good order also S_f^2. The effect of the second pair of terms in Equation 2, therefore, renders the $J(\omega)$ surface less sensitive to $\{S_f^2, S_s^2, \tau_s\}$ and conversely their extracted values are significantly more sensitive to errors in the relaxation data than is the case for ^{15}N measurements. This accounts for the differences in the widths of the distributions of S_f^2, S_s^2, τ_s obtained in fits of methyl and ^{15}N relaxation data, Figure 2. That the errors in extracted values of τ_f are in general smaller in the case of methyl data can be understood by noting that the last two terms in Equation 2 (which are the only ones sensitive to τ_f) contribute significantly more to $J(\omega)$ in the case of methyl relaxation (i.e., when the other terms are attenuated by 1/9) than for ^{15}N spin relaxation where the factor of 1/9 is absent.

Finally, in fitting either ^{13}C or ^{2}H data average χ^2 values ($\chi^2 = 1/N\,\Sigma_{i=1,N}(\text{Calc}_i - \text{Fit}_i)^2/\sigma_i^2$, where N is the number of relaxation rates/NOEs fit, Calc are the calculated rates/NOEs including errors, and Fit are the best fit rates/NOEs) of 0.68 ± 0.35 (^{13}C) and 0.81 ± 0.28 (^{2}H) are obtained (the 'errors' correspond to the standard deviation of χ^2 from the 1000 fits), with similar χ^2 values obtained in fits of the ^{15}N rates (0.67 ± 0.34). This indicates that the spread in extracted parameters noted for methyl data is not due to lack of convergence during minimization of the error function, but reflects the shallowness of the error surface itself.

In all of our studies of methyl side chain dynamics as probed by ^{2}H relaxation we have found that in cases where the simple Lipari–Szabo spectral density model does not fit the relaxation data (Equation 1 with S^2 replaced by $(1/9)S_f^2$) a slightly more complex model, LS-3, of the form,

$$J(\omega) = \frac{(1/9)S_f^2\tau_c^{\text{eff}}}{1 + (\omega\tau_c^{\text{eff}})^2} + \frac{(1 - (1/9)S_f^2)\tau}{1 + (\omega\tau)^2}, \qquad \frac{1}{\tau} = \frac{1}{\tau_c^{\text{eff}}} + \frac{1}{\tau_f} \tag{3}$$

does fit the data well (Millet et al., 2002; Skrynnikov et al., 2002). In Equation 3 τ_c^{eff} is a parameter that is obtained on a per-residue basis and represents the combined effects of slow local dynamics and overall tumbling. In a previous study we have also established that spectral density functions computed from a 50 ns molecular dynamics trajectory of an SH3 domain were as well fit by Equation 3 as by Equation 2 in cases where ns time scale dynamics were present (Skrynnikov et al., 2002). The results from both the experimental and computer generated data reflect the fact that the relaxation data are sensitive to the spectral density function evaluated at only a few frequencies (even in cases where large numbers of spectrometer fields are used). Fits of relaxation data effectively derive, therefore, a plateau value and time constant for the correlation function due to fast time scale motions along with a correlation time describing the decay from slower motions, which includes both internal and overall dynamics.

In order to illustrate the utility of the LS-3 model for fitting methyl dynamics data that involve motion on a number of time scales we have simulated error-free ^{13}C and ^{2}H relaxation parameters using the LS-4 form of $J(\omega)(\alpha = 1/9)$ at spectrometer fields of 400, 500, 600 and 800 MHz, as above, as a function of S_s^2, τ_s with $S_f^2 = 0.8$ and $\tau_f = 50$ ps. The rates were subsequently fit to a model of dynamics in which $J(\omega)$ is given by Equation 3 (LS-3), to obtain the LS-3 fitting parameters $S_f^2, \tau_f, \tau_c^{\text{eff}}$. Figure 3 illustrates

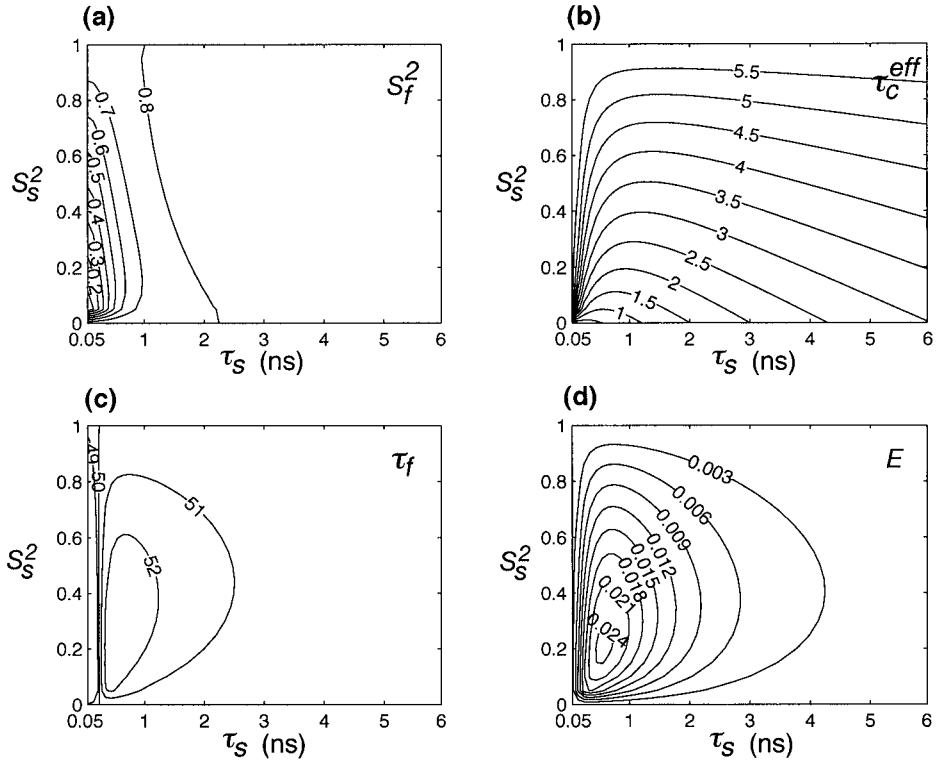

Figure 3. Values of S_f^2 (a), τ_c^{eff} (b) and τ_f (c) extracted from fits of simulated ^{13}C relaxation parameters (T_1, T_2 and NOE, no error) of a methyl $^{13}CHD_2$ group as a function of input values S_s^2 and τ_s. Relaxation parameters were calculated using the LS-4 model with $S_f^2 = 0.80$, $\tau_f = 50$ ps, $\tau_r = 6$ ns and τ_s, S_s^2 values indicated along the x and y axes respectively, and fit using the LS-3 model. (d) E value obtained in the fits [$E^2 = 1/N\,\Sigma_{i=1,N}(\text{Calc}_i - \text{Fit}_i)^2/\text{Calc}_i^2$, where N is the number of relaxation rates/NOEs fit and the Calc (Fit) rates/NOEs are those obtained using the LS-4 (LS-3) model].

results obtained from fits of the simulated ^{13}C relaxation data along with an error parameter, $E(E^2 = 1/N\,\Sigma_{i=1,N}(\text{Calc}_i - \text{Fit}_i)^2/\text{Calc}_i^2)$, showing the goodness of fit, as a function of the values of τ_s and S_s^2 used in the simulations. It is clear that for τ_s values less than about 1 ns the extracted values of S_f^2 are smaller than the values input in the LS-4 model (Figure 3a). As τ_s becomes small, 'fast' and 'slow' internal motions become inseparable and the order parameter that is extracted from the LS-3 model approaches $S_f^2 \times S_s^2$. Hence the deviations between S_f^2 values obtained from fits using the LS-3 form of the spectral density function and input values become large. In contrast, in the limit that $S_s^2 \rightarrow 0, 1$ or as τ_s becomes on the order of τ_r (or larger) the LS-3 and LS-4 models are essentially indistinguishable and S_f^2 values extracted are in good agreement with those that are input. Values of the correlation time, τ_c^{eff}, are affected by internal dynamics and are therefore always smaller than τ_r, Figure 3b. Inspection of Equation 2 shows that in the limit that $S_s^2 \rightarrow 0, 1/\tau_c^{\text{eff}} \rightarrow 1/\tau_r + 1/\tau_s$, while for $S_s^2 \rightarrow 1, 1/\tau_c^{\text{eff}} \rightarrow 1/\tau_r$. The simulations establish further that the extracted value of τ_f is quite insensitive to the details of the dynamics model, Figure 3c. Finally, Figure 3d shows that good fits of the relaxation data (generated using the LS-4 model) are obtained (LS-3) with E values that are under 0.025.

Figure 4 shows the results from the corresponding analysis using simulated ^{2}H methyl spin relaxation data. In general, the extracted values of S_f^2, τ_f and τ_c^{eff} depend in a similar way on input τ_s and S_s^2 for both ^{13}C and ^{2}H relaxation data. As described above, values of S_f^2 obtained using the LS-3 model converge to the input values as $\tau_s \rightarrow \tau_r$. However, the convergence occurs more rapidly in the case of ^{13}C data (compare Figures 3a, 4a) since ^{13}C relaxation depends on spectral density terms evaluated at higher frequencies than ^{2}H (Figure 1). These high frequency spectral density terms are particularly sensitive to fast time scale dynamics (Lee et al., 1999) and therefore help

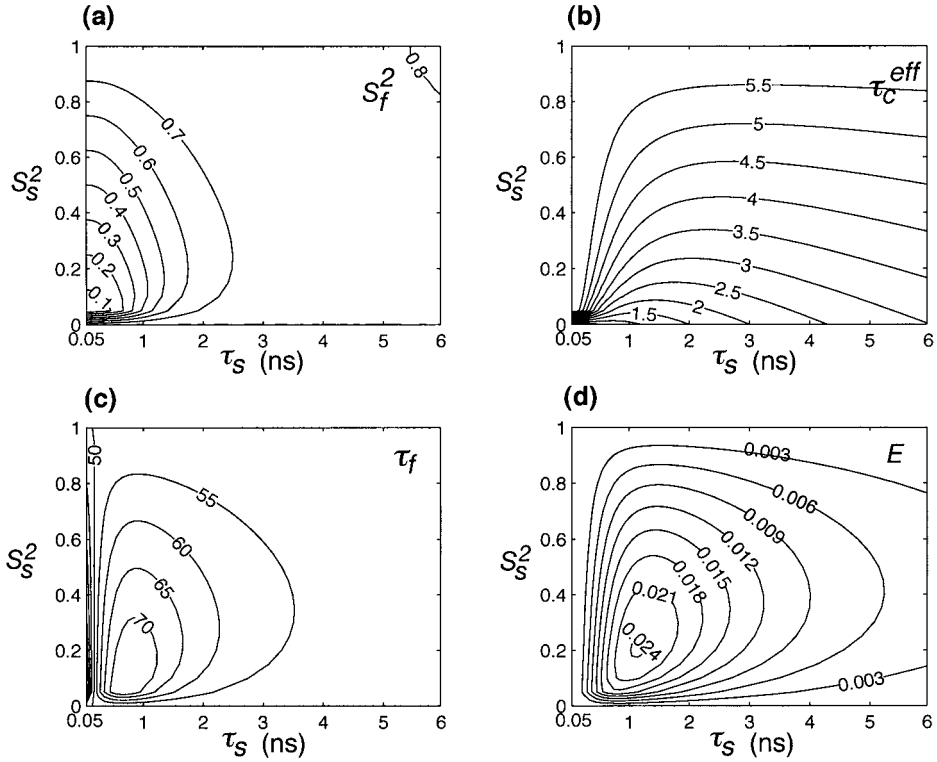

Figure 4. Values of S_f^2 (a), τ_c^{eff} (b), τ_f (c), and goodness of fit parameter, E, extracted from LS-3 fits of simulated ^{2}H relaxation rates (R_{D_z}, R_{D_+}, $R_{D_+D_z+D_zD_+}$, $R_{3D_z^2-2}$ and $R_{D_+^2}$, no error) of a methyl $^{13}CH_2D$ group generated using the LS-4 model ($S_f^2 = 0.80$, $\tau_f = 50$ ps and $\tau_r = 6$ ns) as a function of input values S_s^2 and τ_s (see text for details).

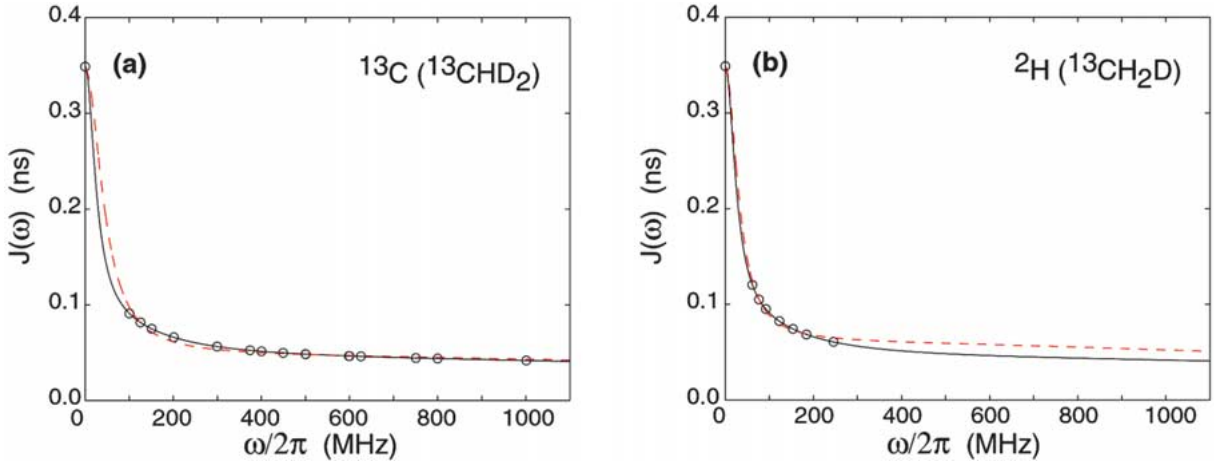

Figure 5. Comparison of the LS-4 spectral density functions (solid black lines) used to generate ^{13}C (a) and ^{2}H (b) methyl relaxation rates ($\{S_f^2, \tau_f, S_s^2, \tau_s\} = \{0.8, 50$ ps, 0.5, 1.0 ns$\}$ and $\tau_r = 6$ ns) with the LS-3 spectral density (dashed red line) that best fits the data. The black circles indicate the values of the LS-4 spectral density function that contribute to the spin relaxation parameters obtained in measurements performed at 400, 500, 600 and 800 MHz, see text.

force convergence to the correct value of S_f^2. Figure 5 illustrates this clearly, with LS-4 spectral density functions (black solid lines) calculated using {0.8, 50 ps, 0.5, 1 ns} for $\{S_f^2, \tau_f, S_s^2, \tau_s\}$ and $\tau_r = 6$ ns. The black circles correspond to those values of the spectral density function that contribute to spin relaxation parameters obtained in measurements performed at 400, 500, 600 and 800 MHz. The best fit LS-3 spectral density functions (dashed red lines) are shown. It is quite clear that the tail of the spectral density function is better fit by the LS-3 model in the case of ^{13}C data, which is due to the fact that there is data in this region to be fit in the first place.

As a final comparison of the LS-3 and -4 models for the analysis of methyl relaxation data we return to the case highlighted above where ^{13}C and ^{2}H relaxation rates have been calculated using $\{S_f^2, \tau_f, S_s^2, \tau_s\} = \{0.8, 50 \text{ ps}, 0.5, 1 \text{ ns}\}$ and $\tau_r = 6$ ns. One thousand data sets have been generated with small errors added to the rates (3%) and NOE values (5%) (see above). F-test statistical analyses showed that the LS-4 model fit the rates better than the LS-3 approach at the 95% confidence level for less than 8% of the data sets. That so few trials were better fit using the LS-4 form of the spectral density is quite remarkable given that the data was generated using this model in the first place! Notably, all cases involving fitting of ^{15}N data using the LS-4 model were superior to those obtained with the LS-3 form of $J(\omega)$ (95% confidence level). Further simulations with different input dynamics parameters have established that the results obtained for the specific case described above are quite general.

In summary, the present study suggests that the LS-4 model is appropriate for the analysis of ^{15}N spin relaxation data (in cases where ns time scale local motions are present) and that robust dynamics parameters can be extracted given the level of experimental errors that are currently reported in most studies. In contrast, in many cases fits of side chain relaxation rates to the LS-4 model produce significantly wider distributions of motional parameters than might be anticipated on the basis of the errors in the input data, even in cases where extensive data sets are used. This has been observed through simulations reported here, even for small errors, and in the case of applications to experimental data, described previously (Millet et al., 2002; Skrynnikov et al., 2002). The origin of the differences between the distributions in dynamics parameters obtained in studies involving backbone and side chain methyl relaxation has been explored. In cases where the LS-2 model fails, the LS-3 spectral density function provides fits of the relaxation data that, in general, reproduce the data as well as the LS-4 model, and the extracted parameters are significantly more stable to error.

Finally, it is important to emphasize that although we have assumed in the above analysis that the side chain dynamics are described by the LS-4 model, it is likely that a correct description is in fact more complex; the LS-4 model is, however, the simplest that considers dynamics on both fast and slow time scales. In reality, in cases where ns time scale motions are present, the values of correlation times and order parameters extracted from fits using any model are likely to depend in complex ways on the actual dynamics parameters. Our studies of side chain dynamics in Protein L (Millet et al., 2002; Skrynnikov et al., 2002), the N-terminal SH_3 domain of drk (Millet et al., 2002; Skrynnikov et al., 2002) and the fyn SH_3 domain (T. Mittermaier, in preparation) suggest that 10–20% of the residues have contributions from ns motions and the relaxation data in these cases cannot be fit using the LS-2 model. For these residues a more complete description will have to await the development of additional experiments, allowing the spectral density function to be probed at an increased number of frequencies.

Acknowledgements

W.-Y.C. is a recipient of a Senior Research Fellowship from the Canadian Institutes of Health Research (CIHR). This research was supported by a grant from the CIHR. L.E.K. holds a Canada Research Chair in Biochemistry.

References

Brooks, C.L., Karplus, M. and Pettitt, B.M. (1988) *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics*, John Wiley and Sons, New York, NY.

Choy, W.Y., Shortle, D. and Kay, L.E. (2003) *J. Am. Chem. Soc.*, **125**, 1748–1758.

Clore, G.M., Szabo, A., Bax, A., Kay, L.E., Driscoll, P.C. and Gronenborn, A.M. (1990) *J. Am. Chem. Soc.*, **112**, 4989–4991.

Cordier, F., Brutscher, B. and Marion, D. (1996) *J. Biomol. NMR*, **7**, 163–168.

Dayie, K.T. and Wagner, G. (1997) *J. Am. Chem. Soc.*, **119**, 7797–7806.

Edwards, A.M., Arrowsmith, C.H., Christendat, D., Dharamsi, A., Friesen, J.D., Greenblatt, J.F. and Vedadi, M. (2000) *Nat. Struct. Biol. Suppl.*, **7**, 970–972.

Engelke, J. and Rüterjans, H. (1995) *J. Biomol. NMR*, **5**, 173–182.
Farrow, N.A., Muhandiram, R., Singer, A.U., Pascal, S.M., Kay, C.M., Gish, G., Shoelson, S.E., Pawson, T., Forman-Kay, J.D. and Kay, L.E. (1994) *Biochemistry*, **33**, 5984–6003.
Fersht, A. (1985) *Enzyme Structure and Mechanism*, 2nd edn., Freeman and Co., New York.
Ishima, R. and Torchia, D.A. (2000) *Nat. Struct. Biol.*, **7**, 740–743.
Ishima, R., Louis, J.M. and Torchia, D.A. (1999) *J. Am. Chem. Soc.*, **121**, 11589–11590.
Ishima, R., Petkova, A.P., Louis, J.M. and Torchia, D.A. (2001) *J. Am. Chem. Soc.*, **123**, 6164–6171.
Kay, L.E. (1998) *Nat. Struct. Biol. NMR Suppl.*, **5**, 513–516.
Lee, A.L., Flynn, P.F. and Wand, A.J. (1999) *J. Am. Chem. Soc.*, **121**, 2891–2902.
Lee, A.L., Urbauer, J.L. and Wand, A.J. (1997) *J. Biomol. NMR*, **9**, 437–440.
LeMaster, D.M. and Kushlan, D.M. (1996) *J. Am. Chem. Soc.*, **118**, 9255–9264.
Lipari, G. and Szabo, A. (1982a) *J. Am. Chem. Soc.*, **104**, 4546–4559.
Lipari, G. and Szabo, A. (1982b) *J. Am. Chem. Soc.*, **104**, 4559–4570.
Millet, O., Muhandiram, D.R., Skrynnikov, N.R. and Kay, L.E. (2002) *J. Am. Chem. Soc.*, **124**, 6439–6448.
Muhandiram, D.R., Johnson, P.E., Yang, D., Zhang, O., McIntosh, L.P. and Kay, L.E. (1997) *J. Biomol. NMR*, **10**, 283–288.
Muhandiram, D.R., Yamazaki, T., Sykes, B.D. and Kay, L.E. (1995) *J. Am. Chem. Soc.*, **117**, 11536–11544.
Palmer, A.G., Williams, J. and McDermott, A. (1996) *J. Phys. Chem.*, **100**, 13293–13310.
Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1988) *Numerical Recipes in C*, Cambridge University Press, Cambridge.
Skrynnikov, N.R., Millet, O. and Kay, L.E. (2002) *J. Am. Chem. Soc.*, **124**, 6449–6460.
Yamazaki, T., Muhandiram, R. and Kay, L.E. (1994) *J. Am. Chem. Soc.*, **116**, 8266–8278.
Yang, D., Mittermaier, A., Mok, Y.K. and Kay, L.E. (1998) *J. Mol. Biol.*, **276**, 939–954.
Zeng, L., Fischer, M.W.F. and Zuiderweg, E.R.P. (1996) *J. Biomol. NMR*, **7**, 157–162.