J|A|C|S
ARTICLES
Published on Web 05/11/2002

Deuterium Spin Probes of Side-Chain Dynamics in Proteins. 2. Spectral Density Mapping and Identification of Nanosecond Time-Scale Side-Chain Motions

Nikolai R. Skrynnikov, Oscar Millet, and Lewis E. Kay*

Contribution from the Protein Engineering Network Centers of Excellence and Departments of Medical Genetics and Microbiology, Biochemistry, and Chemistry, University of Toronto, Toronto, Ontario, Canada M5S 1A8

Received November 7, 2001. Revised Manuscript Received January 29, 2002

Abstract: In the previous paper in this issue we have demonstrated that it is possible to measure the five different relaxation rates of a deuteron in $^{13}CH_2D$ methyl groups of ^{13}C-labeled, fractionally deuterated proteins. The extensive set of data acquired in these experiments provides an opportunity to investigate side-chain dynamics in proteins at a level of detail that heretofore was not possible. The data, acquired on the B1 domain of peptostreptococcal protein L, include 16 (9) relaxation measurements at 4 (2) different magnetic field strengths, 25 °C (5 °C). These data are shown to be self-consistent and are analyzed using a spectral density mapping procedure which allows extraction of values of the spectral density function at a number of frequencies with no assumptions about the underlying dynamics. Dynamics data from 31 of 35 methyls in the protein for which data could be obtained were well-fitted using the two-parameter Lipari−Szabo model (Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546). The data from the remaining 4 methyls can be fitted using a three-parameter version of the Lipari−Szabo model that takes into account, in a simple manner, additional nanosecond time-scale local dynamics. This interpretation is supported by analysis of a molecular dynamics trajectory where spectral density profiles calculated for side-chain methyl sites reflect the influence of slower (nanosecond) time-scale motions involving jumps between rotameric wells. A discussion of the minimum number of relaxation measurements that are necessary to extract the full complement of dynamics information is presented along with an interpretation of the extracted dynamics parameters.

Introduction

In the previous paper in this issue (referred to in what follows as paper 1), we presented NMR experiments for measuring the relaxation rates of deuterium double-quantum and antiphase transverse magnetization, along with the relaxation of deuterium quadrupolar order in $^{13}CH_2D$ methyl groups of ^{13}C-, fractionally ^{2}H-labeled proteins. In combination with previously published methodology for obtaining decay rates of longitudinal and in-phase ^{2}H transverse magnetization,[1] these experiments allow measurement of the five different quadrupolar relaxation rates associated with a single spin-1 particle. We have shown that the five rates measured for the B1 immunoglobulin binding domain of peptostreptococcal protein L [2] at a number of different temperatures and spectrometer fields are internally consistent. Satisfied that we can measure the five relaxation rates accurately, we now turn our attention to the interpretation of these data in terms of protein side-chain dynamics. The measurement of at least five rates per deuteron, and more if additional fields are employed, presents an opportunity to study side-chain dynamics in proteins in a way previously not possible. The equations describing the quadrupolar relaxation of the five modes, D_Z (longitudinal magnetization), D_+ (transverse in-phase magnetization), D_+^2 (double-quantum magnetization), $3D_Z^2 - 2$ (quadrupolar order), and $D_+D_Z + D_ZD_+$ (transverse antiphase magnetization), show that their decay rates depend on a spectral density function evaluated at only three distinct frequencies[3] (see eq 1, of paper 1). It therefore becomes possible to uniquely determine these spectral density values, via a process referred to as spectral density mapping, and ultimately to elucidate the fundamental features of the dynamics of the system.

Spectral density mapping[4,5] is an important tool in NMR studies of biomolecular dynamics since a set of measured relaxation rates can be interpreted directly without any a priori assumptions about the character of the underlying motions. The method, therefore, can be used in the study of both constrained backbone dynamics in folded proteins[6,7] or large-amplitude motions in partially or completely unfolded protein states.[8,9] Data recorded at multiple magnetic fields can be combined in a straightforward fashion to allow better sampling of the spectral

* Corresponding author. E-mail: kay@bloch.med.utoronto.ca.

(1) Muhandiram, D. R.; Yamazaki, T.; Sykes, B. D.; Kay, L. E. *J. Am. Chem. Soc.* **1995**, *117*, 11536−11544.
(2) Scalley, M. L.; Yi, Q.; Gu, H.; McCormack, A.; Yates, J. R.; Baker, D. *Biochemistry* **1997**, *36*, 3373−82.
(3) Jacobsen, J. P.; Bildsoe, H. K.; Schaumburg, K. *J. Magn. Reson.* **1976**, *23*, 153−164.
(4) Peng, J. W.; Wagner, G. *J. Magn. Reson.* **1992**, *98*, 308−332.
(5) Jaffe, D.; Vold, R. L.; Vold, R. R. *J. Magn. Reson.* **1982**, *46*, 496−502.
(6) Ishima, R.; Nagayama, K. *J. Magn. Reson., Ser. B* **1995**, *108*, 73−76.
(7) Farrow, N. A.; Zhang, O.; Szabo, A.; Torchia, D. A.; Kay, L. E. *J. Biomol. NMR* **1995**, *6*, 153−162.

density function. Fitting motional models directly to the spectral density function readily allows one to assess whether the data are overfitted (see below). Spectral density mapping has been routinely used to analyze relaxation of backbone ^{15}N[10−13] and ^{13}C[14,15] spins. However, despite a number of significant advances in methods for the measurement of side-chain relaxation rates,[1,16,17] the number of relaxation parameters measured at any particular site in the side chain has remained insufficient to support spectral density mapping analyses, with a few exceptions.[18]

Here we present for the first time the results of a comprehensive spectral density mapping study of side-chain methyl deuterons in proteins. Relaxation data in protein L have been recorded at spectrometer fields of 400, 500, 600, and 800 MHz, corresponding to deuterium resonance frequencies of 61, 77, 92, and 123 MHz, respectively, and spectral density profiles have been obtained for a total of 35 side-chain methyl groups in the protein. Three versions of the model-free approach of Lipari and Szabo[19,20] have been subsequently used to fit the spectral density profiles. For the majority of methyl-containing residues in protein L (31 methyls) adequate fits are obtained using the simple Lipari−Szabo model where a single order parameter and correlation time, S_f^2 and τ_f, respectively, describe the local dynamics with an overall tumbling correlation time, τ_R, obtained from ^{15}N spin relaxation measurements. However, the spectral density profiles of the remaining four methyls are clearly incompatible with this two-parameter Lipari−Szabo model. The spectral density profiles for these residues can be well−fit using a simple extension of the standard model, where τ_R is replaced with a variable effective correlation time, τ_c^{eff}. The correlation time τ_c^{eff} describes the combined effect of the overall tumbling and slow-time-scale (nanosecond) rotameric interconversions occurring in the individual side chain. Molecular dynamics data presented in the Appendix illustrate the applicability of the τ_c^{eff} model for treatments of side-chain dynamics.

The use of a comprehensive deuterium relaxation data set acquired on protein L allows the study of side-chain dynamics at a level of detail that was previously not feasible, including the identification of nanosecond-time-scale motions. The correct identification of such motion requires additional relaxation times beyond T_1 and $T_{1\rho}$ values recorded at a single field. The number of relaxation rates that are required for the analysis of side-chain dynamics is discussed along with the different models that can be used to extract the motional parameters.

(8) Farrow, N. A.; Zhang, O.; Forman-Kay, J. D.; Kay, L. E. *Biochemistry* **1995**, *34*, 868−878.
(9) Farrow, N. A.; Zhang, O.; Forman-Kay, J. D.; Kay, L. E. *Biochemistry* **1997**, *36*, 2390−2402.
(10) Constantine, K. L.; Friedrichs, M. S.; Wittekind, M.; Jamil, H.; Chu, C. H.; Parker, R. A.; Goldfarb, V.; Mueller, L.; Farmer, B. T. *Biochemistry* **1998**, *37*, 7965−7980.
(11) Bracken, C.; Carr, P. A.; Cavanagh, J.; Palmer, A. G., III. *J. Mol. Biol.* **1999**, *285*, 2133−2146.
(12) Kloiber, K.; Weiskirchen, R.; Krautler, B.; Bister, K.; Konrat, R. *J. Mol. Biol.* **1999**, *292*, 893−908.
(13) Viles, J. H.; Donne, D.; Kroon, G.; Prusiner, S. B.; Cohen, F. E.; Dyson, H. J.; Wright, P. E. *Biochemistry* **2001**, *40*, 2743−2753.
(14) Hill, R. B.; Racken, C.; DeGrado, W. F.; Palmer, A. G. *J. Am. Chem. Soc.* **2000**, *122*, 11610−11619.
(15) Atkinson, R. A.; Lefevre, J. F. *J. Biomol. NMR* **1999**, *13*, 83−88.
(16) Wand, A. J.; Urbauer, J. L.; McEvoy, R. P.; Bieber, R. J. *Biochemistry* **1996**, *35*, 6116−6125.
(17) LeMaster, D. M.; Kushlan, D. M. *J. Am. Chem. Soc.* **1996**, *118*, 9255−9264.
(18) Mayo, K. H.; Daragan, V. A.; Idiyatullin, D.; Nesmelova, I. *J. Magn. Reson.* **2000**, *146*, 188−95.
(19) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4559−4570.
(20) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4559.

Spectral Density Mapping and Self-Consistency of ^2H Relaxation Rates

An experimental data set for protein L measured at 25 °C has been obtained which is comprised of 16 deuterium relaxation parameters. These include $R^Q(D_Z)$, $R^Q(D_+)$, and $R^Q(D_+^2)$ measured at 400, 500, 600, and 800 MHz as well as $R^Q(3D_Z^2 - 2)$ and $R^Q(D_+D_Z + D_ZD_+)$ at 500 and 600 MHz. The relaxation rates obtained at any given field can be expressed via three spectral density values at the angular frequencies 0, ω_D, and $2\omega_D$ (see eq 1 of paper 1), where ω_D is the deuterium Larmor frequency.[3] The entire set of 16 measured relaxation rates, therefore, is a function of 8 spectral density values (considering that the zero frequency value is sampled at each field and that $2\omega_D$ at 400 MHz coincides with ω_D at 800 MHz). Determination of these 8 spectral densities from 16 measured relaxation rates constitutes the *spectral density mapping* procedure which, in contrast to many previous studies of backbone dynamics, is highly overdetermined in this case.

To illustrate the spectral density mapping approach in some detail, we consider the case where the 5 independent ^2H relaxation rates have been measured at a single spectrometer field. In paper 1, we have shown that the measured relaxation rates $\Gamma_1^{exptl} = R(I_ZC_ZD_Z) - R(I_ZC_Z)$, $\Gamma_2^{exptl} = R(I_ZC_ZD_+) - R(I_ZC_Z)$, $\Gamma_3^{exptl} = R(I_ZC_Z\{2D_Z^2 - 1\}) - R(I_ZC_Z)$, and $\Gamma_4^{exptl} = R(I_ZC_Z\{D_+D_Z + D_ZD_+\}) - R(I_ZC_Z)$ approximate $R^Q(D_Z)$, $R^Q(D_+)$, $R^Q(3D_Z^2 - 2)$, and $R^Q(D_+D_Z + D_ZD_+)$ to better than 2% over a wide range of motional parameters. It has also been shown that the measured double-quantum relaxation rate contains nonnegligible contributions from dipolar interactions according to

$$\Gamma_5^{exptl} = R(I_ZC_ZD_+^2) - R(I_ZC_Z) - \sum_{k\neq ij} R^D_{DI^k} \approx R^Q(D_+^2) + 8d_{I^iD}J_{I^iD}(0) + 8d_{I^jD}J_{I^jD}(0) + 8d_{CD}J_{CD}(0)$$

(see eq 10 of paper 1) with the right and left hand sides of this equality agreeing to better than 0.5%. Recall that $\sum_{k\neq i,j} R^D_{DI^k}$ is the contribution to the relaxation of the methyl deuteron from all proton spins external to the methyl group, which can be accurately calculated from the measurement of $R(I_ZC_Z)$ (see paper 1), and $d_{qq'} = (1/10)(\mu_0/4\pi)^2(\hbar\gamma_q\gamma_{q'}/\langle r_{qq'}^3\rangle)^2$.

On the basis of eq 1 of the previous paper in this issue and the expression for Γ_5^{exptl}, we can write

$$\begin{bmatrix} \Gamma_1^{exptl}/\sigma_1 \\ \Gamma_2^{exptl}/\sigma_2 \\ \Gamma_3^{exptl}/\sigma_3 \\ \Gamma_4^{exptl}/\sigma_4 \\ \Gamma_5^{exptl}/\sigma_5 \end{bmatrix} = \left\{ \frac{1}{80}\left(\frac{e^2Qq}{\hbar}\right)^2 \begin{bmatrix} 0 & 6/\sigma_1 & 24/\sigma_1 \\ 9/\sigma_2 & 15/\sigma_2 & 6/\sigma_2 \\ 0 & 18/\sigma_3 & 0 \\ 9/\sigma_4 & 3/\sigma_4 & 6/\sigma_4 \\ 0 & 6/\sigma_5 & 12/\sigma_5 \end{bmatrix} + (9d_{ID} + 2d_{CD}) \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 4/\sigma_5 & 0 & 0 \end{bmatrix} \right\} \begin{pmatrix} J(0) \\ J(\omega_D) \\ J(2\omega_D) \end{pmatrix} \quad (1)$$

Note that the relatively small dipolar contributions to the relaxation of the ^2H double-quantum mode originating from the

$C-D$ and $I-D$ interactions ($\sim$5−10%) can, with good accuracy, be expressed with only the $J(0)$ spectral density term describing the dynamics of the $C-D$ bond vector since $J_{ID}(0) \approx (9/4)J_{CD}(0)$ (see paper 1). Hence, the set of three spectral densities in eq 1 is sufficient to describe both dominant quadrupolar contributions and secondary dipolar terms that contribute to double-quantum relaxation. Multiplication of the elements in eq 1 by $1/\sigma_p$, where σ_p is the average experimental error in the measurement of Γ_p^{exptl}, ensures that the biggest weight is given to those Γ_p^{exptl} values that have been measured with the highest accuracy.

Equation 1 assumes that the quadrupolar tensor of a deuterium spin residing in a methyl group is axially symmetric. An additional contribution from the small rhombic component of the quadrupolar tensor[21] is expected to remain below 2%, even though this term is largely unaffected by the fast spinning of the methyl group. Assuming axial symmetry we have recently established that the quadrupolar coupling constants of methyl ^{2}H spins are essentially site-independent and estimated the average e^2Qq/h value to be 167 kHz,[22] in good agreement with related solid-state studies.[23] This value is used in what follows to evaluate eq 1.

Equation 1 can be readily generalized to include measurements at several spectrometer fields as well as partial data sets in cases where less than 5 relaxation rates are measured per deuteron. In our case the size of the matrix corresponding to the one in eq 1 is 16×8 since 16 Γ_p^{exptl} ($p = 1...16$) values have been measured, which can be described in terms of 8 spectral densities, $j_q = J(\omega_q)$ ($q = 1...8$). Eight best-fit values of the spectral densities, j_q^{fit}, are obtained by solving this system using singular value decomposition.[24] These spectral densities can be subsequently substituted into an analogue of eq 1 to back-calculate relaxation rates, Γ_p^{fit}. The correlation between Γ_p^{exptl} and Γ_p^{fit} provides a measure of the consistency of the experimental data and can be viewed as a generalized form of the consistency relationships discussed in paper 1.

Figure 1 shows the correlations between experimental and fitted relaxation rates when all 16 relaxation rates, comprising the complete data set for protein L at 25 °C, are included in the analysis. Overall, the quality of correlations is excellent, with values of the systematic bias, δ calculated as the average $(\Gamma_p^{\text{exptl}} - \Gamma_p^{\text{fit}})/\Gamma_p^{\text{exptl}}$ ratio, indicated in each panel. For 10 out of 16 measured relaxation rates δ is below 1%, and for another three rates it does not exceed 2%. The worst systematic bias, which amounts to 5%, is obtained for the double-quantum relaxation rate measured at 400 MHz (third panel in the top row). More specifically, it appears that both this rate and the $R^Q(D_Z)$ rate measured at 400 MHz have a small degree of bias, with the discrepancy more evident for the double-quantum data set since it has a bigger experimental uncertainty and therefore is fitted more loosely. Since even a very small systematic bias in certain pieces of data may affect the outcome of the spectral density mapping procedure, we have implemented a jackknife algorithm[25] whereby the uncertainties in extracted values of the spectral density function and model-dependent motionalparameters are estimated from a series of calculations using subsets of the full experimental data set (see Materials and Methods for details of the procedure).

Model-Free Approaches to Analysis of Spectral Densities

We begin this section by reviewing the motional processes that can be important for spin relaxation of deuterium in the side-chain methyl positions of proteins. Four main dynamic processes can be identified, including (i) fast methyl spinning, (ii) fast side-chain motion (torsional librations, bond angle fluctuations, and fast rotameric transitions), (iii) slow side-chain motion (rotameric transitions, including concerted transitions, with characteristic times from hundreds of picoseconds to tens of nanoseconds), and (iv) overall tumbling (isotropic or anisotropic). In practice, fast motional processes (i) and (ii) cannot be separated in a reliable manner on the basis of the relaxation data presented in this paper, since both occur on similar time scales, $\sim$10−100 ps, that cannot be effectively probed by spectral density mapping of data measured at current spectrometer field strengths. Hence, a single characteristic correlation time, τ_f, must be used to characterize fast internal dynamics. Assuming ideal tetrahedral geometry for methyl groups, the combined order parameter from dynamic modes (i) and (ii) can be written as $1/9S_f^2$, where the factor of 1/9 originates from fast methyl spinning and S_f^2 reflects fluctuations of the methyl averaging axis. Note, however, that this identification is tentative since a single decay time, τ_f, is assumed for both processes.

In the presence of fast (i + ii) and slow (iii) local dynamics in addition to the overall tumbling (iv), the spectral density function can be modeled following the approach developed by Clore and co-workers.[26] Assuming that the three motional processes listed above are independent, the correlation function of interest, $g(t)$, can be written as $g(t) = \{(1/9)S_f^2 + (1 - (1/9)S_f^2)\exp(-t/\tau_f)\}\{S_s^2 + (1 - S_s^2)\exp(-t/\tau_s)\}\{\exp(-t/\tau_R)\}$. The resulting spectral density function is

$$J(\omega) = \frac{1}{9}S_f^2 S_s^2 \frac{\tau_0}{1+\omega^2\tau_0^2} + \frac{1}{9}S_f^2(1-S_s^2)\frac{\tau_1}{1+\omega^2\tau_1^2} + S_s^2\left(1-\frac{1}{9}S_f^2\right)\frac{\tau_2}{1+\omega^2\tau_2^2} + \left(1-\frac{1}{9}S_f^2\right)(1-S_s^2)\frac{\tau_3}{1+\omega^2\tau_3^2}$$

$$1/\tau_0 = 1/\tau_R$$

$$1/\tau_1 = (1/\tau_R) + (1/\tau_s) \qquad (2)$$

$$1/\tau_2 = (1/\tau_R) + (1/\tau_f)$$

$$1/\tau_3 = (1/\tau_R) + (1/\tau_s) + (1/\tau_f)$$

where the slower internal motions are parametrized by τ_s and S_s^2. Note that eq 2 assumes the independence of internal dynamics and overall tumbling; however, no assumptions about the relative magnitude of τ_s and τ_R are needed. (See reference 26 for further details).

Consider now the situation where the time scale of slow local dynamics, τ_s, is similar to the time scale of overall tumbling,

(21) Muller, C.; Schajor, W.; Zimmerman, H.; Haeberlen, U. *J. Magn. Reson.* **1984**, *56*, 235−346.
(22) Mittermaier, A.; Kay, L. E. *J. Am. Chem. Soc.* **1999**, *121*, 10608−10613.
(23) Burnett, L. H.; Muller, B. H. *J. Chem. Phys.* **1971**, *55*, 5829−5831.
(24) Press: W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes in C*; Cambridge University Press: Cambridge, U.K., 1988.
(25) Mosteller, F.; Tukey, J. W. *Data Analysis and Regression: A Second Course in Statistics*.; Addison-Wesley: Reading, MA, 1977.
(26) Clore, G. M.; Szabo, A.; Bax, A.; Kay, L. E.; Driscoll, P. C.; Gronenborn, A. M. *J. Am. Chem. Soc.* **1990**, *112*, 4989−4991.

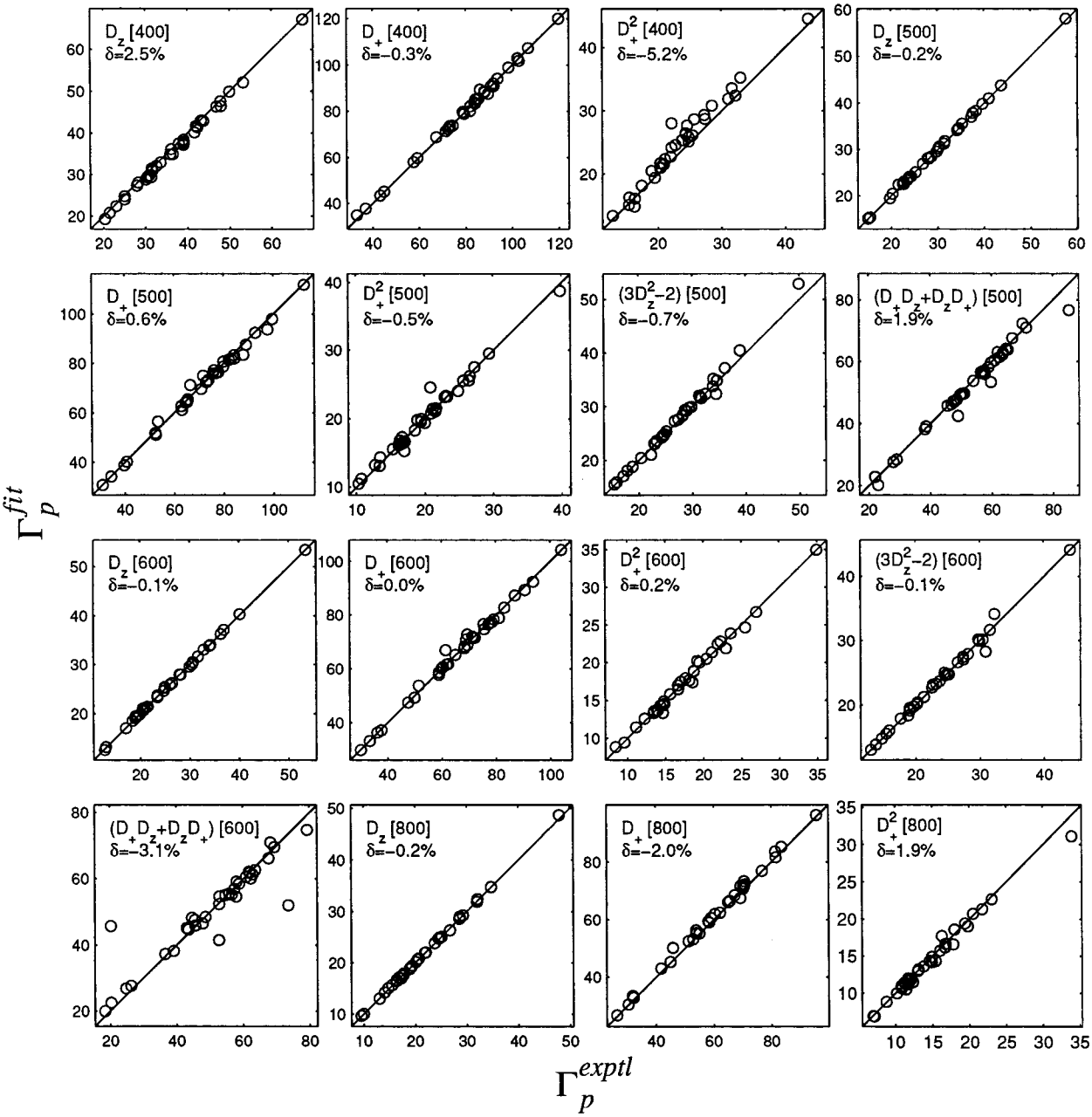

Figure 1. Correlations between experimental deuterium relaxation rates of protein L at 25 °C, Γ_p^{exptl}, horizontal axes, and the rates calculated using the best-fit spectral density values obtained from the spectral density mapping procedure, Γ_p^{fit}, vertical axis. A total of 16 rates have been interpreted using 8 independent spectral densities. The definitions of Γ_p are given in the text, and the respective deuterium spin operators and magnetic field strengths (expressed as proton resonance frequencies in MHz) are indicated in the top left-hand corners of each panel. The deviation δ, calculated as the average $(\Gamma_p^{exptl} - \Gamma_p^{fit})/\Gamma_p^{exptl}$ ratio, is also shown.

τ_R. The tail of the correlation function in this case is comprised of two exponentials with similar decay times, τ_0 and τ_1. Even a small amount of noise makes it exceedingly difficult to separate the contributions from two decay processes with similar rates. On the other hand, the product of the two can be well-approximated with a single exponential, leading to the simplified expression for $g(t)$, $g(t) \approx \{(1/9)S_f^2 + (1 - (1/9)S_f^2)\exp(-t/\tau_f)\}\{\exp(-t/\tau_c^{eff})\}$, and consequently for $J(\omega)$:

$$J(\omega) = \frac{1}{9}S_f^2\frac{\tau_c^{eff}}{1+\omega^2(\tau_c^{eff})^2} + \left(1-\frac{1}{9}S_f^2\right)\frac{\tau}{1+\omega^2\tau^2} \qquad (3)$$

$$1/\tau = (1/\tau_c^{eff}) + (1/\tau_f)$$

In this model τ_c^{eff} represents the combined effect of slow local dynamics and overall tumbling, which are assumed to proceed on a similar time scale (e.g. on a time scale of several nanoseconds). In the Appendix we present correlation functions calculated from a 50 ns molecular dynamics trajectory which lend support to the simplified model involving τ_c^{eff}. In particular, we show that slow time-scale (nanosecond) rotameric transitions observed in the MD trajectory for a number of side chains give rise to spectral densities consistent with eq 3. Note that in the limiting case, when $S_s^2 = 0$, eq 2 rigorously reduces to eq 3 with τ_c^{eff} and τ identified with τ_1 and τ_3, respectively. The order parameter S_s^2 is fully expected to be small in mobile side chains because of 3-fold rotameric transitions involving one or more dihedral angles along the side chain (i.e. for the same

reason that the fast methyl spinning gives rise to a low order parameter of 1/9).

Finally, if slow local dynamics is absent, then $\tau_c^{\text{eff}} = \tau_R$ and eq 3 is reduced in a trivial manner to the conventional Lipari−Szabo spectral density:

$$J(\omega) = \frac{1}{9}S_f^2\frac{\tau_R}{1+\omega^2\tau_R^2} + \left(1-\frac{1}{9}S_f^2\right)\frac{\tau}{1+\omega^2\tau^2} \qquad (4)$$

$$1/\tau = (1/\tau_R) + (1/\tau_f)$$

Expressions for $J(\omega)$, eqs 2−4, involve four $\{S_f^2,\tau_f,S_s^2,\tau_s\}$, three $\{S_f^2,\tau_f,\tau_c^{\text{eff}}\}$, and two $\{S_f^2,\tau_f\}$ fitting parameters, respectively, with the overall tumbling time τ_R known from ^{15}N relaxation measurements. In what follows these three models rooted in the work of Lipari and Szabo[19,20] are referred to as LS-4, -3, and -2, according to the number of fitting parameters contained therein. All models describe the dynamics of a vector that coincides with the C−H bond of the methyl group; these models can also be applied to the C^{methyl}−C (or C^{methyl}−S) bond if the factor (1/9) in the respective expressions is replaced with 1. The meaning of the best-fit order parameters S_f^2 extracted from fits using the LS-2 and -3 models in cases where slow side-chain dynamics is present is discussed below. Theoretical treatments which draw the connection between the model-free parameters and more fundamental quantities describing side-chain torsional dynamics have been presented in a number of recent papers.[27−30]

Spectral Density Mapping of Methyl Sites in Protein L

As discussed above, the extensive set of deuterium relaxation data obtained for protein L at 25 °C allows the determination of spectral densities $j_q = J(\omega_q)$ at eight distinct frequencies ω_q. As a first step, the spectral density profiles $J(\omega_q)$ were fitted using the conventional Lipari−Szabo model[19,20] adapted for the case of rapidly rotating methyl groups,[31] eq 4. The overall molecular tumbling time, τ_R, used in the analyses has been determined by standard ^{15}N relaxation measurements[32] (see Materials and Methods), with $\tau_R = 4.05$ ns for protein L at 25 °C.

Figure 2 shows values of $J(\omega_q)$ (circles in the plots) together with LS-2 best-fit curves (solid line) for a number of residues. The best-fit values of τ_f and S_f^2 are also indicated in each panel. Uncertainties in the fitted parameters have been estimated on the basis of a jackknife procedure in which a series of reduced data sets comprised of 14 relaxation rates per methyl group are fitted. To ensure that the magnitudes of the errors are not exaggerated, reduced data sets were constructed in such a manner that they remained overdetermined with respect to all spectral densities as described in detail in Materials and Methods. The error bars represent the entire range of $J(\omega_q)$ values that are obtained with the jackknife procedure.

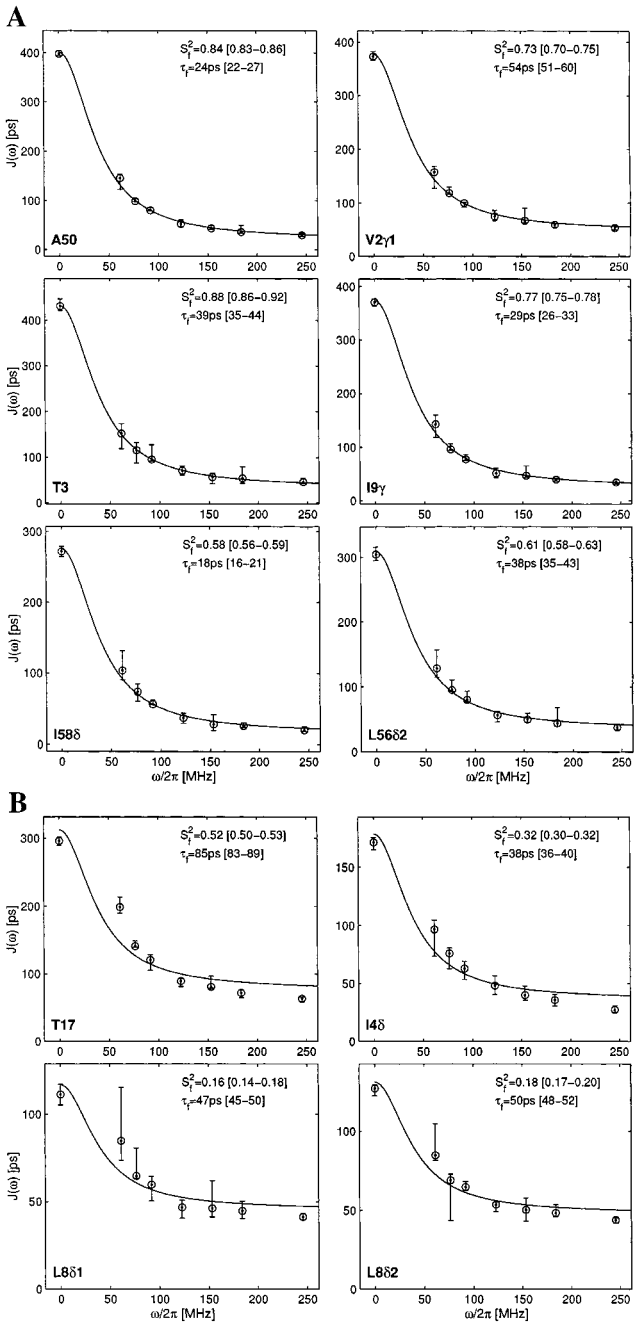

Figure 2. Results of the spectral density mapping procedure described in the text using the standard Lipari−Szabo spectral density function,[19,20] LS-2, for selected residues in protein L at 25 °C. The circles correspond to the 8 spectral density values derived from 16 experimentally measured deuterium relaxation rates and the curves have been generated by fitting these spectral densities to eq 4. The best-fit values of S_f^2 and τ_f along with their uncertainties are indicated in the upper right-hand corner of each panel. Part A (top six panels) shows the results for selected methyl sites that are fitted well with the LS-2 model (class A side chains). Part B (bottom four panels) shows the results for all methyl sites that are fitted poorly with the LS-2 model (class B side chains). A jackknife procedure[25] has been used to determine average spectral density values, associated error bars, and uncertainties in the fitted parameters, as described in Materials and Methods.

The standard Lipari−Szabo (LS-2) model produces good fits for a total of 31 out of 35 methyl groups in the protein and for these methyls *F*-test statistics[25] do not justify the use of more complex models (average probability of chance improvement, *p*, greater than 85%). In these cases the best-fit curves are inside or, on several occasions, just outside the span of the error bars

(27) Bremi, T.; Brüschweiler, R.; Ernst, R. R. *J. Am. Chem. Soc.* **1997**, *119*, 4272−4284.
(28) Mikhailov, D.; Daragan, V. A.; Mayo, K. H. *J. Biomol. NMR* **1995**, *5*, 397−410.
(29) Mikhailov, D. V.; Washington, L.; Voloshin, A. M.; Daragan, V. A.; Mayo, K. H. *Biopolymers* **1999**, *49*, 373−383.
(30) Wong, K. B.; Daggett, V. *Biochemistry* **1998**, *37*, 11182−11192.
(31) Kay, L. E.; Torchia, D. A. *J. Magn. Reson.* **1991**, *95*, 536−547.
(32) 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.; Kay, L. E. *Biochemistry* **1994**, *33*, 5984−6003.

for all $J(\omega_q)$ points. This is illustrated in Figure 2A (top six panels) for selected methyl groups, one from each methyl type in protein L. In contrast, four methyl groups—T17, I4δ, L8δ1, and L8δ2—clearly cannot be fitted well with the LS-2 model, Figure 2B (bottom four panels). In a preliminary discussion of these results, it should be pointed out that three of the four deviant methyl groups belong to residues with long side chains (δ positions of Leu and Ile). In contrast, spectral density values from all nine Ala methyls can be well-fit with the LS-2 model. This suggests that the reason for the unusual behavior of the 4 methyl groups in Figure 2B lies in the large-amplitude motions that can occur in longer side chains. In what follows, those methyls (31 of 35) that are well-fit by the LS-2 model are referred to as *class A* methyls, while the remaining 4 are referred to as *class B* (fit with the LS-3 model but not the LS-2 model; see below).

The separation of all methyl sites into two groups, as illustrated in Figures 2A and 2B, is not entirely clear-cut. Some of the residues from the first group show a slight trend similar to what is observed in the second group, indicating possible side-chain mobility (see for example Leu 56 in Figure 2A). The effect is, however, below the uncertainty level indicated by the error bars.

As a next step, experimental spectral densities were fitted to the more complex LS-3 and LS-4 models that include the effects of slow local dynamics. To compare the performance of the various models, we have selected two methyl groups, including one from a class A side chain (Ala 50, first panel in Figure 2A, fitted well with the LS-2 model) and another from a class B side chain with suspected nanosecond time-scale internal dynamics (Thr 17, first panel in Figure 2B, fitted poorly with the LS-2 model). The results are shown in Figure 3.

For Ala 50, the use of the more complicated LS-3 and LS-4 models shows no statistically significant improvement over fits obtained using the simple LS-2 model (left column in Figure 3). Indeed, probabilities that the improved fits from the LS-3 and LS-4 models are due to chance are greater than 85%, and the more complex models can therefore be rejected.[25] A correlation time of $\tau_c^{\text{eff}} = 3.77$ ns is obtained with the LS-3 model, with the range of uncertainty extending from 3.6 to 4.2 ns. This interval encompasses the ^{15}N-derived value employed in the LS-2 fit, $\tau_R = 4.05$ ns. Likewise, the order parameters S_f^2 obtained with LS-2 and LS-3 models are consistent within their respective uncertainty intervals. In contrast, application of the four-parameter LS-4 model to the Ala 50 data clearly leads to overfitting (lower left panel in Figure 3). The extracted value of the order parameter S_s^2 has a range of uncertainty from 0.0 to 1.0. Note that both order parameters, S_f^2 and S_s^2, have been constrained in the LS-4 fitting procedure in an attempt to improve the stability of fitting, while no such constraints were needed for either LS-2 or -3 models.

Representative fits of spectral densities from a class B side chain using LS-2, -3, and -4 are illustrated for Thr 17 in the right column of Figure 3. In this case, application of the LS-3 model leads to significant improvement relative to the LS-2 model, with a less than 1% probability that the improvement is obtained by chance. The correlation time, $\tau_c^{\text{eff}} = 2.40$ [2.3–2.6] ns, is distinctly different from $\tau_R = 4.05$ ns, indicating the presence of nanosecond-time-scale motions in the side chain. While the LS-3 model is a clear improvement, a small

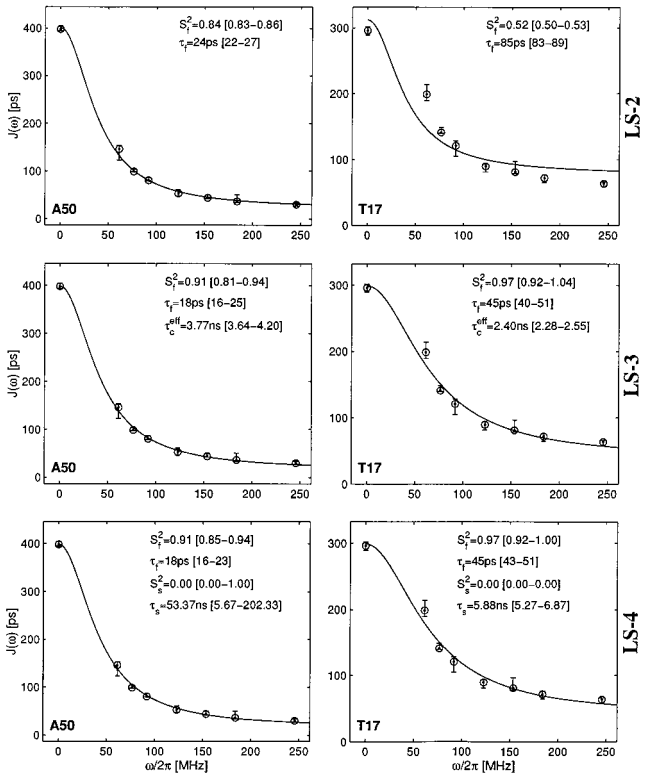

Figure 3. Fits of $J(\omega)$ values extracted from spectral density mapping of 16 experimental relaxation rates for two representative residues in protein L at 25 °C obtained with three different versions of the Lipari–Szabo model, LS-2, LS-3, and LS-4 (top, middle, and bottom rows, respectively). Data for the residues Ala 50 from class A and Thr 17 from class B (left and right columns, respectively) are presented.

discrepancy between the best-fit function and the experimental data remains. This is not surprising since the model-free approach provides only an approximate description of what must certainly be complex side-chain dynamics in many cases. In fact, similar discrepancies are observed when the LS-3 model is used to fit spectral density values derived from MD simulations (results not shown). The order parameter obtained from the LS-2 model, $S_f^2 = 0.52$, can only be viewed as a rough measure of side-chain mobility because of the poor quality of the fit in this case. This value is also too low if one assumes that S_f^2 reflects only small-amplitude fluctuations in the relatively short threonine side chain. On the other hand, the value of 0.97 from the LS-3 model appears to be an overestimate which can be likely attributed to the approximate character of the model, as discussed above. While fitting with the LS-3 model constitutes a substantial improvement over the fit obtained with LS-2 ($p < 0.01$; see above), no further improvement is obtained when the LS-4 model is used, bottom right panel in Figure 3, as the χ^2 residuals obtained from fitting the spectral density profile with the LS-3 and LS-4 models are equal.

The trends illustrated in Figure 3 for Ala 50 and Thr 17 have been also observed for other class A and class B side chains, respectively. The LS-3 fitting of the spectral densities for class A side chains results in an average τ_c^{eff} of 3.7 ns, which is slightly lower than $\tau_R = 4.05$ ns. In contrast, an average τ_c^{eff} of 2.5 ns is obtained for the class B side chains, with an average range of uncertainty between 2.3 and 3.1 ns.

Application of the LS-4 model to the analysis of class A side chains in protein L invariably leads to overfitting. Values of

the order parameter S_s^2 obtained in this manner have uncertainties ranging from 0.0 to 1.0 for 20 methyl sites with the remaining 11 sites also showing excessively large variations. Instability of fits with the LS-4 model have also been seen in the analysis of MD-based simulated data (not shown). For the four class B side chains $S_s^2 = 0.0$, with correlation times τ_s ranging from 5.5 to 7.9 ns. It is interesting that a recent solid-state NMR study of dynamics of Met side chains in the streptomyces subtilisin inhibitor complexed to subtilisin[33] showed that these side chains had correlation times in the range of 10 ns, similar to what has been observed here. Noting that (i) for $S_s^2 = 0$ the LS-4 model is rigorously reduced to the LS-3 model with $1/\tau_c^{eff} = (1/\tau_R) + (1/\tau_s)$, that (ii) the more complex LS-4 model shows no improvement relative to the LS-3 model in terms of data fitting, and that (iii) fits with LS-4 tend to be unstable, we will not consider the LS-4 model any further. As an alternative to the LS-4 model used here, the more conventional four-parameter model due to Clore and co-workers[26] has been tested, with results very similar to the LS-4 model.

As an additional point of interest we have considered how anisotropic tumbling[34−36] would affect the outcome of the analysis described above. The parameters of the diffusion tensor for protein L have been determined on the basis of backbone ^{15}N relaxation data[32] (see Materials and Methods). Using these parameters, the effect of anisotropic tumbling on ^{2}H relaxation in each methyl group was estimated using the so-called "quadric approximation".[34,36] Assuming that the structure of the protein in solution is rigid and identical to the X-ray structure, the correlation time τ_R^{quad} can be calculated for each individual methyl group as a function of the orientation of the $C^{methyl}-C$ axis with respect to the diffusion frame; see eq 14 of Lee et al.[36] Subsequently, residue-specific values of τ_R^{quad} were used instead of the generic value $\tau_R = 4.05$ ns in fitting the spectral density data with the LS-2 model.[19,20]

We have found that the range of τ_R^{quad} values for methyl groups in protein L is rather narrow, 3.8−4.3 ns (for 23 methyls τ_R^{quad} is confined to an even more narrow interval extending from 3.9 to 4.1 ns). As a result, using τ_R^{quad} in place of τ_R in the LS-2 model leads to only very small changes in the parameters extracted from fits of the spectral densities. For example, for the extreme case of Ala 61 use of $\tau_R^{quad} = 3.8$ ns results in an increase in S_f^2 from 0.60 to 0.65. In general, it should be noted that using τ_R^{quad} is appropriate for Ala residues. In contrast, for long side chains that likely display some degree of conformational disorder, this approach is of questionable utility.

An analysis similar to that described above has also been carried out on the relaxation data recorded for protein L at 5 °C. In this case a data set comprised of $R^Q(D_Z)$, $R^Q(D_+)$, $R^Q(D_+^2)$, and $R^Q(3D_Z^2 - 2)$ measured at 500 and 600 MHz as well as $R^Q(D_+D_Z + D_ZD_+)$ obtained at 600 MHz is available. The spectral density mapping procedure in this case amounts to extracting five spectral densities from nine experimentally measured relaxation rates. Since the data set does not include measurements at 400 and 800 MHz and since the correlation times increase with the drop in temperature, sampling of the spectral density profiles is considerably less extensive than in the study carried out at 25 °C. At 5 °C, spectral densities obtained for all of the methyl-containing side chains can be adequately described using the LS-2 model, although fits of the spectral densities for T17, I4δ, L8δ1, and L8δ2 show the same trend as seen in Figure 2B (see Supporting Information).

In this context it is important to discuss the requirements presented by the increasingly complex models of motion in terms of the requisite size of experimental data sets. In some cases, it may not be possible to measure more than ^{2}H T_1^{-1} and $T_{1\rho}^{-1}$ values at a single magnetic field strength (this issue becomes particularly relevant for applications involving proteins of approximately 150 residues or more since the experiments that measure ^{2}H double-quantum, quadrupolar order and anti-phase transverse relaxation rates are considerably less sensitive than the schemes for obtaining T_1^{-1} and $T_{1\rho}^{-1}$ values, as described in paper 1). Two deuterium relaxation rates provide a minimum data set that supports analysis with the LS-2 form of spectral density function. The limitations of LS-2 analyses that arise from nanosecond time-scale local dynamics are further discussed below.

The minimum data set necessary for interpretation using the LS-3 model consists of three deuterium relaxation rates measured at a single field. This data set should include at least one of the transverse rates, $T_{1\rho}^{-1} = R^Q(D_+)$ or $R^Q(D_+D_Z + D_ZD_+)$, so that the spectral density $J(0)$ can be accessed. Furthermore, it is important that the dispersion region of the $J(\omega)$ profile be sampled by the relaxation data, i.e., that $\omega_D\tau_c^{eff} \sim 1$. Conversely, if for example $\omega_D\tau_c^{eff} >> 1$, then both $J(\omega_D)$ and $J(2\omega_D)$ fall on the plateau of the spectral density profile and effectively represent a single spectral density value, $J(\omega_D) \approx J(2\omega_D)$, so that use of the LS-3 model leads to overfitting of the data. It is worth noting that the most favorable conditions for studies of nanosecond side-chain dynamics are when $\omega_D\tau_s \sim 1$ and $\tau_s \ll \tau_R$. In practice, one should exercise caution using minimum data sets comprised of three relaxation rates since the spectral density mapping procedure in this situation is not overdetermined. We have found that using five relaxation rates at a single field instead of three can greatly improve the accuracy of the analyses in some cases.

To attempt data fits using the LS-4 model, a set of deuterium relaxation rates must be measured at no less than two different magnetic field strengths. In principle, the combination of T_1^{-1} and $T_{1\rho}^{-1}$ measurements at two fields related by a factor of 2, e.g. 400 and 800 MHz, provides enough data to support a four-parameter fitting. In our experience the significant sensitivity drop in the 400 MHz measurements (approximately 2-fold relative to 500 MHz) is problematic, although data sets recorded at lower fields are particularly useful since the $J(\omega_D)$ values obtained are more likely to lie in the dispersion portion of the spectral density profile. In any event, as described above, it is important to ensure that the spectral density values obtained will not be overfitted using the chosen model. The spectral density mapping approach provides a nice visual check that this is indeed the case.

In summary, in the present work we have tried to be comprehensive in our study of side-chain dynamics and we have therefore carried out measurements at as many magnetic fields as possible. The results from the small globular protein, protein L, suggest that in many cases measurements at a single field

(33) Tamura, A.; Matsushita, M.; Naito, A.; Kojima, S.; Miura, K. I.; Akasaka, K. *Protein Sci.* **1996**, *5*, 127−39.
(34) Brüschweiler, R.; Liao, X.; Wright, P. E. *Science* **1995**, *268*, 886−889.
(35) Tjandra, N.; Feller, S. E.; Pastor, R. W.; Bax, A. *J. Am. Chem. Soc.* **1995**, *117*, 12562−12566.
(36) Lee, L. K.; Rance, M.; Chazin, W. J.; Palmer, A. G. *J. Biomol. NMR* **1997**, *9*, 287−298.

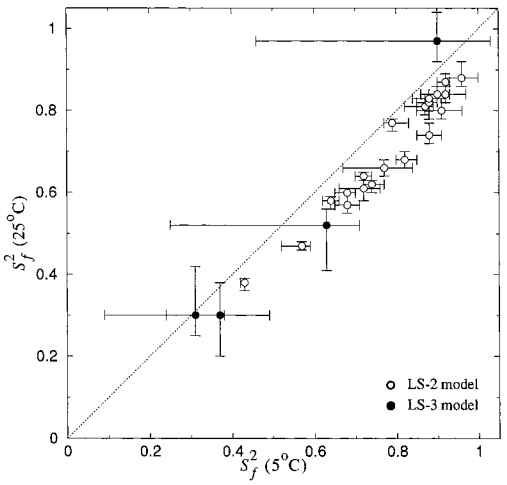

Figure 4. Correlation between the methyl order parameters S_f^2 determined from experimental relaxation data recorded on protein L at 5 °C (horizontal axis) and 25 °C (vertical axis). S_f^2 values were obtained from fitting of spectral densities using the LS-3 model (four methyl sites from class B, closed circles) or the LS-2 model (other methyl sites, open circles). Note the substantial uncertainty in the LS-3-derived values of S_f^2 at 5 °C which is due to the limited sampling of the spectral density profiles at this temperature (see Supporting Information).

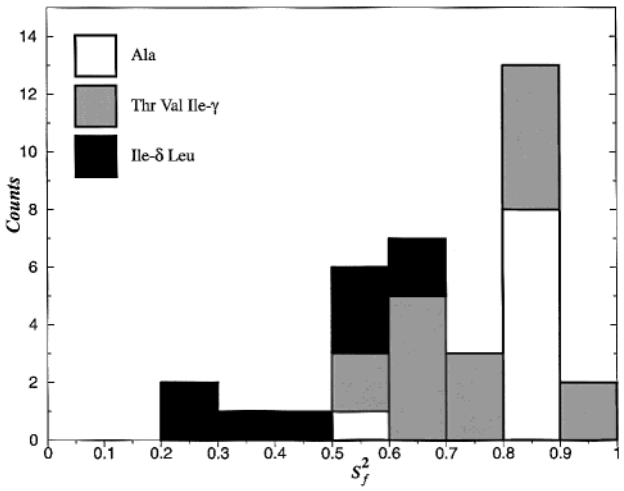

Figure 5. Distribution of S_f^2 values extracted from fits of spectral density values obtained from 16 experimental relaxation rates measured on protein L at 25 °C. The bars in the histogram are color coded according to the number of dihedral angles between the backbone Cα and the methyl carbon.

may be sufficient to extract the complement of information available from relaxation experiments. For applications to other macromolecules, including unfolded proteins, the question of the minimum number of field strengths necessary will have to be explored further.

Interpretation of the Fast Time-Scale Order Parameters, S_f^2

The fast motion order parameter, S_f^2, is an important measure of internal mobility in proteins providing insight into molecular recognition events[10,37−40] and protein stability,[41,42] for example. Using deuterium relaxation data from protein L, order parameters were obtained at two temperatures using either LS-2 or LS-3 spectral density models. An excellent correlation is observed between S_f^2(5°C) and S_f^2(25°C), Figure 4, with the low-temperature values on average higher by 12%, consistent with what has been observed by Lee and Wand in their study of methyl dynamics in calmodulin.[41]

The statistical distribution of S_f^2 values from protein L at 25 °C is shown in Figure 5. As expected, high-order parameters are observed in methyls attached to short side chains, while low-order parameters are usually noted in the long side chains (on average, $S_f^2 = 0.81$ for Ala and 0.56 for Leu side chains, for example). The distribution appears to have more than one maximum, reminiscent of the trimodal pattern described by Lee and Wand.[41] In general it has proven difficult to correlate the order parameters, S_f^2, with certain unique characteristic features of the local environment such as the number of side-chain contacts,[43] solvent exposure, or secondary structure, as has been already pointed out by Mittermaier et al.[44] This difficulty is exemplified by the comparison between S_f^2 values for L8 and L56 where very different order parameters (0.30 and 0.61, respectively; the values for $\delta 1$ and $\delta 2$ sites happen to be equal in both cases) are obtained despite the fact that both residues are located in β-strands, have no solvent exposure, and form multiple contacts with other side chains. It has been previously pointed out that S_f^2 may depend on the conformational state of the side chain and the orientation of the averaging axis as well as on the rigidity of the local environment.[37,45]

Analysis of order parameters in terms of molecular properties is predicated on their correct interpretation. It is therefore of interest to evaluate how nanosecond-time-scale local dynamics affect the values of S_f^2 extracted from the LS-2 model. To address this issue, a series of numerical simulations were carried out where synthetic input data were generated using the LS-4 model, eq 2, which explicitly includes nanosecond-time-scale dynamics. In the first case, Figure 6A, the input data consisted of simulated T_1^{-1} and $T_{1\rho}^{-1}$ rates (600 MHz). These two rates were subsequently fitted with eqs 1a,c from paper 1 using the LS-2 model, and S_f^2, τ_f values were recovered. The contour plot in Figure 6A shows the deviation between the fitted order parameter value, S_f^2 (LS-2), and the target value $S_f^2 = 0.8$ as a function of the two slow (nanosecond) motion parameters used to generate the input data, τ_s and S_s^2. A similar plot, Figure 6B, has been produced using input data consisting of $J(0)$, $J(\omega_D)$, and $J(2\omega_D)$ values calculated from eq 2. Both plots demonstrate that the presence of nanosecond-time-scale side-chain dynamics may lead to a dramatic underestimation of S_f^2 if the LS-2 model is used to interpret the data, consistent with the analyses of the experimental data in the previous section. Clearly in this situation S_f^2 determined with the standard Lipari−Szabo (LS-2) model cannot be considered as a measure of the amplitude of fast motion, but should rather be viewed as an empirical parameter that absorbs the effects of both fast and slow local dynamics.

When the LS-3 model is used to fit the three spectral densities calculated from eq 2, as described above, the accuracy of extracted S_f^2 values clearly improves relative to fits with the

(37) Kay, L. E.; Muhandiram, D. R.; Farrow, N. A.; Aubin, Y.; Forman-Kay, J. D. *Biochemistry* **1996**, *35*, 361−368.
(38) Kay, L. E.; Muhandiram, D. R.; Wolf, G.; Shoelson, S. E.; Forman-Kay, J. D. *Nat. Struct. Biol.* **1998**, *5*, 156−163.
(39) Lee, A. L.; Kinnear, S. A.; Wand, A. J. *Nat. Struct. Biol.* **2000**, *7*, 72−77.
(40) Ishima, R.; Louis, J. M.; Torchia, D. A. *J. Mol. Biol.* **2001**, *305*, 515−21.
(41) Lee, A. L.; Wand, A. J. *Nature* **2001**, *411*, 501−504.
(42) Desjarlais, J. R.; Handel, T. M. *J. Mol. Biol.* **1999**, *290*, 305−318.
(43) Kim, D. E.; Fisher, C.; Baker, D. *J. Mol. Biol.* **2000**, *298*, 971−84.
(44) Mittermaier, A.; Kay, L. E.; Forman-Kay, J. D. *J. Biomol. NMR* **1999**, *13*, 181−185.
(45) Yang, D.; Mittermaier, A.; Mok, Y. K.; Kay, L. E. *J. Mol. Biol.* **1998**, *276*, 939−954.

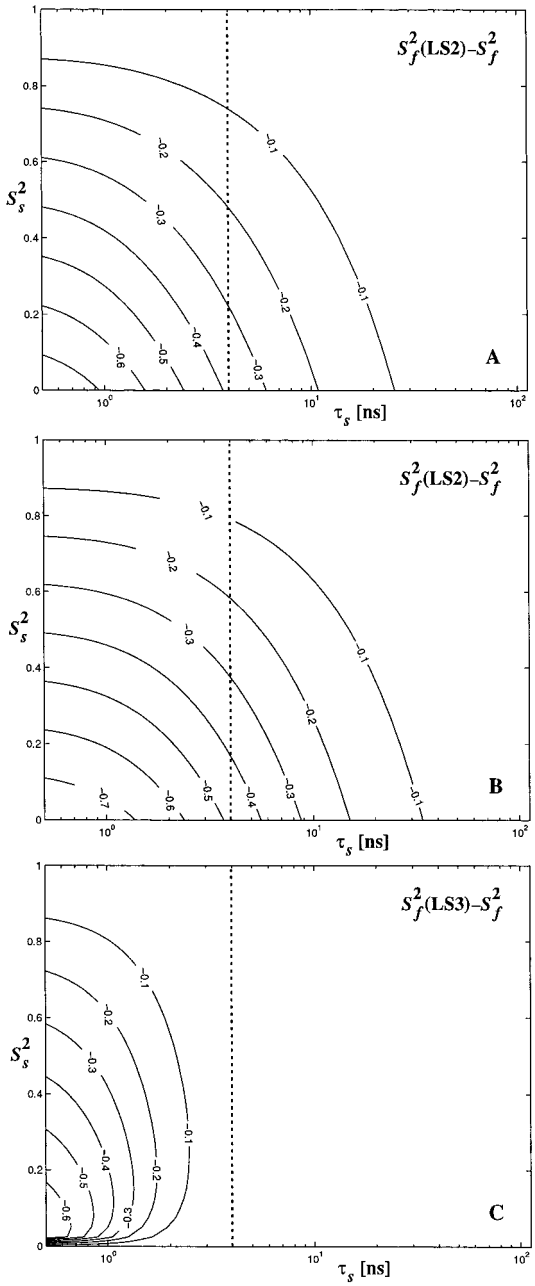

Figure 6. Errors in the order parameter S_f^2 extracted from fitting simulated input data to the LS-2 (A, B) and LS-3 (C) models. The input data were generated using the LS-4 spectral density function (eq 2) with the following parameters: $\tau_R = 4.0$ ns, $\tau_f = 50$ ps, $S_f^2 = 0.8$, proton frequency 600 MHz, with τ_s and S_s^2 varied as indicated by the labels on the axes of the contour plot. The input data consisted of the two simulated quadrupolar relaxation rates $\{T_1^{-1}, T_{1\rho}^{-1}\}$ (A) or three values of the spectral density function $\{J(0),J(\omega_D),J(2\omega_D)\}$ (B, C). At each point in the contourplot, the difference S_f^2(LS-2) − S_f^2 (or S_f^2(LS-3) − S_f^2) is plotted, where S_f^2 is the target value of the fast motion order parameter (0.8) and S_f^2(LS-2) (S_f^2(LS-3)) is the best-fit value extracted by fitting of the simulated data with the LS-2 (LS-3) model. The point where $\tau_s = \tau_R$ is indicated by a vertical dashed line.

LS-2 model. Accurate S_f^2 values are recovered for a wide range of correlation times τ_s, $\tau_s > 2.5$ ns, Figure 6C. In comparison, even very slow local dynamics with $\tau_s \sim 25$ ns may compromise the extracted S_f^2 values in the case of the LS-2 model, Figure 6B.

Finally, we wish to mention an alternative interpretation of the LS-2-based order parameter S_f^2(LS-2). It can be viewed as *either* an approximation of the original quantity, S_f^2, *or* as an approximation of the product, $S_f^2S_s^2$. For example, in the limiting case in which $\tau_s \rightarrow \tau_f$, "slow" and "fast" local dynamics become indistinguishable and S_f^2(LS-2) $\approx S_f^2S_s^2$. To illustrate the quality of this dual interpretation of S_f^2(LS-2), we have plotted the minimum of the two deviations, $|S_f^2$(LS-2) − $S_f^2|$ and $|S_f^2$(LS-2) − $S_f^2S_s^2|$, in Figure 7A. While the interpretation of S_f^2(LS-2) in terms of either S_f^2 or $S_f^2S_s^2$ holds reasonably well in both the low and high end of the τ_s range, it is less accurate when τ_s becomes comparable to τ_R (dashed vertical line in the plot). In general, this dual interpretation of S_f^2(LS-2) is useful since it relates S_f^2(LS-2) to the amplitudes of local motions, irrespective of their time scales. Figure 7B shows the analogous plot for S_f^2(LS-3) extracted using the LS-3 model. Finally, note that the distinction between fast and slow time scales is somewhat arbitrary. In the present example we have set $\tau_f = 50$ ps and $\tau_s \geq 0.5$ ns so that at least a factor of 10 separates τ_f from τ_s.

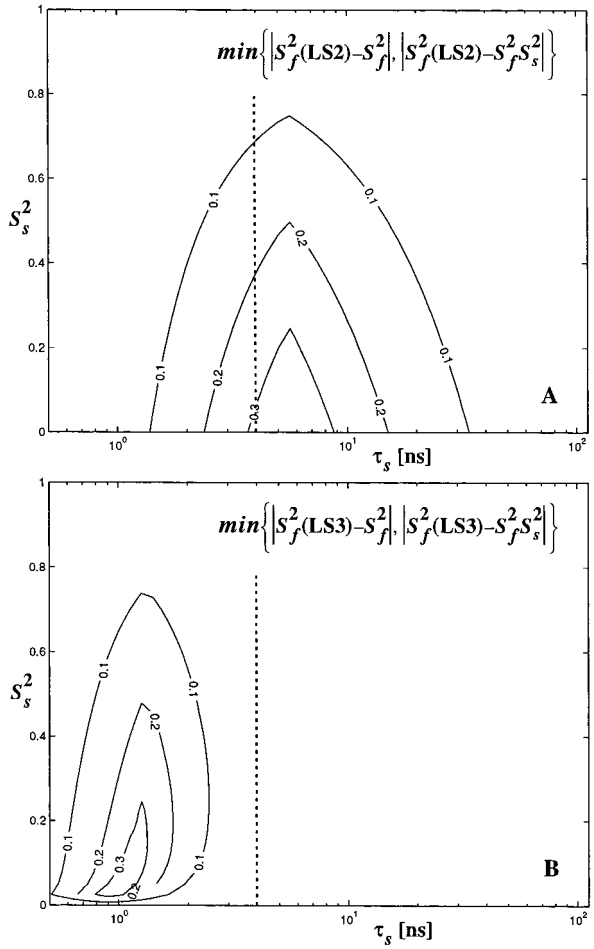

Figure 7. Order parameters S_f^2(LS-2) (A) and S_f^2(LS-3) (B) as a measure of amplitude of local dynamics, including fast and slow dynamics, that approximates either S_f^2 or $S_f^2S_s^2$. The simulated input data are the same as in Figure 6B,C. In part A, at each point in the contour plot the minimum of the two quantities $|S_f^2$(LS-2) − $S_f^2|$ and $|S_f^2$(LS-2) − $S_f^2S_s^2|$ is plotted, where S_f^2 and S_s^2 represent the values used to generate the input data and S_f^2(LS-2) is the best-fit value obtained from fitting the synthetic input data with the LS-2 model. Part B displays the same measure for the LS-3 model.

Conclusion

Deuterium relaxation rates measured using the three new relaxation experiments described in paper 1 in concert with T_1^{-1} and $T_{1\rho}^{-1}$ rates measured using two previously described experiments[1] can be analyzed with spectral density mapping

procedures to characterize side-chain dynamics in proteins. Remarkably, the analysis is strongly overdetermined with respect to the spectral densities of interest (for example, measurements at the four magnetic fields—400, 500, 600, and 800 MHz—result in up to 20 experimentally determined relaxation rates which depend on 8 distinct spectral densities). In this work we take full advantage of the deuterium spin-1 probe that (i) relaxes almost completely via the quadrupolar mechanism[1,46] and (ii) has five unique relaxation rates[3] that can be measured to probe dynamics.

Analysis of relaxation data measured for the protein L domain shows that in the majority of cases the relaxation-active dynamics of methyl-bearing side chains is limited to fast, mainly small-amplitude, torsional fluctuations. The order parameters associated with these fast librations are expected to be relatively high ($\geq \sim 0.5$). In contrast, unusually low order parameters of ca. 0.1−0.3 likely point to the occurrence of nanosecond-time-scale rotameric transitions. For example, in the case of Leu side chains only two (χ_1, χ_2) rotamers, (−60°, +180°) and (+180°, +60°), account for 88% of all conformations encountered in the structural database.[47] The transition between these two states, which likely entails only very modest rearrangement of the protein interior, may well have a strong effect on methyl relaxation for this residue.[48,49] In these cases S_f^2 values obtained by fitting the data to the LS-2 model[19,20] should be viewed as an empirical (and qualitative) measure of side-chain mobility.

We suggest that whenever $R^Q(D_Z)$ and $R^Q(D_+)$ deuterium relaxation rates only are available they should be interpreted using the standard Lipari−Szabo (LS-2) model. The extracted value of S_f^2 should be considered a generalized measure of side-chain mobility that, in addition to fast picosecond dynamics, may also reflect the presence of nanosecond-time-scale motions. The latter motions are likely to be present if low values for S_f^2 ($\leq \sim 0.3$) are extracted from fits of the data. If a more extensive relaxation data set is available, then it is possible to consider additional models (such as the LS-3 model), allowing the separation of fast- (picosecond) and slower-time-scale (nanosecond) internal motions so that S_f^2 retains its original physical meaning as the amplitude of the fast internal motion. Application of the LS-3 variant of the Lipari−Szabo model also permits determination of the correlation time τ_c^{eff} which reflects nanosecond-time-scale internal dynamics of individual side chains. The present approach to the analysis of ^{2}H relaxation data is proving useful in studies of other proteins, including a number of thermodynamically (de)stabilized mutants of protein L, currently underway in our group.

It is worth pointing out that side-chain relaxation and coupling constant measurements[50] provide complementary information. Indeed, it has been demonstrated that in the side chains with very low order parameters the observed scalar coupling constants reflect averaging over several rotameric states.[37] The relaxation studies, however, are advantageous in that they can also identify the time scale of rotameric jumps, as demonstrated in this work.

Spectral density mapping based on comprehensive sets of five deuterium relaxation rates measured at multiple magnetic fields provides a clear and precise signature of local side-chain dynamics in proteins. The information that is available through these studies is expected to provide an important contribution to the emerging picture of side-chain dynamics and ultimately increase our understanding of protein folding, stability, ligand binding, and molecular recognition in a variety of systems.

Materials and Methods

Spectral Density Mapping. An experimental data set has been recorded for protein L at 25 °C comprised of 16 relaxation rates, including $R^Q(D_Z)$, $R^Q(D_+)$, and $R^Q(D_+^2)$ measured at 400, 500, 600, and 800 MHz as well as $R^Q(3D_Z^2 - 2)$ and $R^Q(D_+D_Z + D_ZD_+)$ at 500 and 600 MHz. To properly account for errors, a jackknife procedure[25] has been implemented as follows. From the original set containing 16 relaxation rates, all possible subsets comprised of 14 rates have been generated. Subsequently, those subsets have been excluded that contained only one relaxation parameter measured at 400 or at 800 MHz. This ensures that all remaining partial data sets (114 subsets) are overdetermined with respect to all spectral densities, j_q. As a next step, spectral density mapping was performed on each of these partial data sets so that a set of j_q values is obtained by solving the analogue of eq 1. The eight $j_q = J(\omega_q)$ values obtained for a given partial data set were fitted using one of the theoretical models for the spectral density function, eqs 2−4. By means of example, when the LS-2 model[19,20] was used, spectral densities were fitted with the two-parameter function given by eq 4 to generate a pair of best-fit values, $\{S_f^2,\tau_f\}$. For each of the 114 partial data sets an array of 8 spectral densities, $J(\omega_q)$, and a pair of best-fit values, $\{S_f^2,\tau_f\}$, are obtained. The minimum and the maximum S_f^2 and τ_f values from the list are reported as the error limits throughout the paper. Average values of each of the spectral densities, $J(\omega_q)$, were calculated from the 114 values (shown as circles in Figures 2 and 3). In addition, minimum and maximum $J(\omega_q)$ values were extracted, defining the errors associated with each spectral density value (error bars in Figures 2 and 3). Average $J(\omega_q)$ values were subsequently fitted using one of eqs 2−4, and the best-fit values obtained in this analysis (e.g. S_f^2 and τ_f in the case of fits with the LS-2 model) are reported. All fittings were based on a simplex minimization routine,[24] which was repeated several times starting from randomized initial conditions to ensure that good convergence is achieved. The protocol described above was also applied to the analysis of relaxation data from protein L collected at 5 °C. In this case the complete data set is comprised of 9 pieces of data ($R^Q(D_Z)$, $R^Q(D_+)$, $R^Q(D_+^2)$, and $R^Q(3D_Z^2 - 2)$ measured at 500 and 600 MHz as well as $R^Q(D_+D_Z + D_ZD_+)$ obtained at 600 MHz) so that values of the spectral density function at 5 frequencies can be determined. Partial data sets of 7 rates, including at least 3 rates measured at 500 MHz, were employed in this situation to estimate uncertainties.

We have also carried out a series of Monte Carlo simulations[51] to assess how random error in the experimental rates, Γ_p^{expt}, affects the extracted spectral densities. These errors have been shown to have a relatively small impact on uncertainties in the resulting spectral densities, $J(\omega)$. The major source of uncertainty derives from very small (see Figure 1) but systematic errors in the relaxation data, and it has been addressed using the jackknife procedure described above.

Overall Molecular Tumbling. Parameters of rotational diffusion for protein L were obtained from analysis of ^{15}N relaxation data (T_1, T_2, and ^{1}H−^{15}N NOE measured at 600 MHz at 5 and 25 °C) using the program R2R1.[36] Values of τ_R ($=1/6D_{iso}$) = 4.05 and 8.01 ns at 25 and 5 °C, respectively, were obtained, with $D_\parallel/D_\perp = 1.43$ (25 °C) and 1.46 (5 °C). A more complex treatment which includes fits with a fully asymmetric diffusion tensor was not warranted. Following the determination of the diffusion tensor, we analyzed T_1, T_2, and NOE values on a per-residue basis[19,52] to identify the sites potentially affected by

(46) Yang, D.; Kay, L. E. *J. Magn. Reson., Ser. B* **1996**, *110*, 213−218.
(47) Ponder, J. W.; Richards, F. M. *J. Mol. Biol.* **1987**, *193*, 775−791.
(48) Nicholson, L. K.; Kay, L. E.; Baldisseri, D. M.; Arango, J.; Young, P. E.; Bax, A.; Torchia, D. A. *Biochemistry* **1992**, *31*, 5253−5263.
(49) LeMaster, D. M. *J. Am. Chem. Soc.* **1999**, *121*, 1726−1742.
(50) Bax, A.; Vuister, G. W.; Grzesiek, S.; Delaglio, F.; Wang, A. C.; Tschudin, R.; Zhu, G. *Methods Enzymol.* **1994**, *239*, 79−105.
(51) Kamith, U.; Shriver, J. W. *J. Biol. Chem.* **1989**, *264*, 5586−5592.
(52) Lipari, G.; Szabo, A. *J. Chem. Phys.* **1981**, *75*, 2971−2976.

slow conformational exchange. The data from three residues most likely to be affected by exchange were subsequently removed from the data set, and the diffusion tensor was recalculated. The refined diffusion parameters are virtually identical to the original ones, indicating that conformational exchange is not a factor.

Molecular Dynamics. A 50 ns molecular dynamics trajectory of the drkN SH3 domain was kindly provided to us by Drs. M. Philippopoulos, J. D. Forman-Kay, and R. Pomes of the Hospital for Sick Children, Toronto. Details of the simulation will be published by these authors shortly. Briefly, the MD simulation was performed using version 6 of the AMBER suite of programs[53] and the AMBER PARM96 force-field.[54,55] The lowest-energy member of the NMR-derived ensemble of structures of the salt-stabilized drkN SH3 domain was used as the starting point for the simulation. The molecule was immersed in a periodic box of water molecules of dimensions 56 × 49.6 × 46.5 Å, big enough to ensure that the minimum distance between any protein atom and the box sides was 10 Å, resulting in a total of 4007 added water molecules. After a 50 ps heating and equilibration stage, the system was subjected to 50 ns dynamics at constant temperature and volume with coordinates saved every 1 ps. All structures generated by molecular dynamics were initially fitted to the starting structure so that the effects of overall motion were removed. Autocorrelation functions used in the present analysis were calculated from MD snapshots (one every 2 ps) from the final 48 ns of the trajectory according to the relation $g(t) = 1.5\langle\cos^2\theta(t)\rangle - 0.5$, where $\theta(t)$ is the angle between the vector of interest at time t_0 and the same vector at time $t_0 + t$ and angular brackets imply averaging over t_0. The value of t extends from 0 to 24 ns to allow sufficient averaging. The values of $g(t)$ were subsequently multiplied by $\exp(-t/\tau_R)$ to account for overall molecular tumbling. The resulting profiles were fitted and the spectral densities were evaluated as described by Bremi et al.[27] The sets of $J(\omega_q)$ values generated in this manner effectively "mimic" the output of the spectral density mapping procedure applied to the experimental datasets.

Acknowledgment. N.R.S and O.M. made equal contributions to this work. The authors are extremely grateful to Drs. M. Philippopoulos, J. D. Forman-Kay, and R. Pomes (Hospital for Sick Children, University of Toronto) for kindly making their 50 ns molecular dynamics trajectory of the drkN SH3 domain available prior to publication. The authors also wish to thank Drs. A. Szabo and D. A. Torchia, NIH (Bethesda, MD), for helpful discussions during the course of data analysis and Dr. Ranjith Muhandiram (Toronto) for help with recording the data. Funding from the Ministerio de Educacion y Cultura (MEC) (O.M.) and the Canadian Institutes of Health Research (CIHR) in the form of a Centennial Fellowship (N.R.S.) is gratefully acknowledged. This work is supported by grants from the Natural Sciences and Engineering Research Council of Canada and the CIHR. L.E.K. is a foreign investigator of the Howard Hughes Medical Research Institute.

Appendix: Molecular Dynamics

A 50 ns molecular dynamics trajectory of the folded, wild-type drkN SH3 domain (Philippopoulos, M.; Forman-Kay, J. D.; Pomes, R. Manuscript in preparation; see Materials and Methods) is used to illustrate some general features of side-chain dynamics in globular proteins that are relevant in the context of the present deuterium relaxation study. Note that, unlike the situation experimentally where the wild-type drkN SH3 domain exists in equilibrium between folded and unfolded states under near physiological conditions,[56] "in the computer" the protein is folded. Our goal is not to provide a comprehensive interpretation of the MD data or to compare the MD results with NMR relaxation data, which has been the subject of many stimulating studies in the past.[57−60] Rather we use the molecular dynamics data to gain insight into the dynamic processes that can occur in protein side chains and to demonstrate that MD-based spectral density profiles show the same trends as observed experimentally.

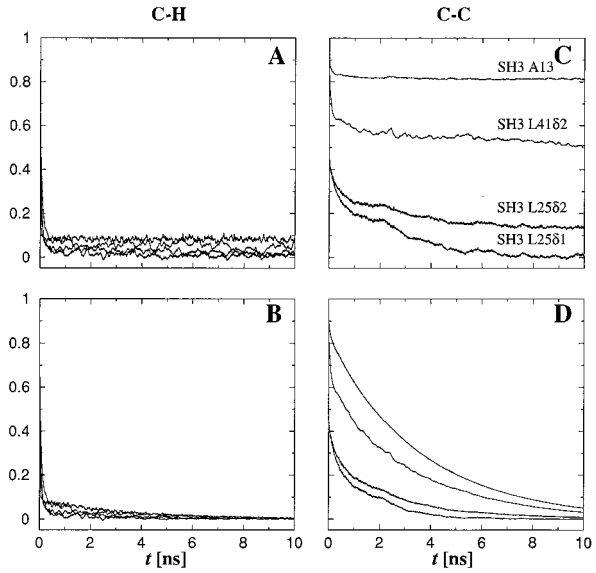

Figure 8. Autocorrelation functions for the axially symmetric second-rank interactions along the C−H (A,B) and C−C (C,D) bonds in the drkN SH3 domain obtained from analysis of a molecular dynamics simulation of the protein. The effects of molecular tumbling are removed (A,C) or, alternatively, included (B,D), as described in Materials and Methods. An overall tumbling correlation time $\tau_R = 3.61$ ns, as experimentally determined for the T22G drkN SH3 domain at 25 °C, has been used. Four methyl sites that illustrate the entire range of dynamic behavior observed for methyl-containing side chains in the trajectory are indicated in panel C.

Correlation functions have been computed for the second-rank axially symmetric interactions along the methyl C−H and the C−C bonds from the MD trajectory, with representative examples shown in Figure 8. In panels A and B correlation functions for four C−H vectors (methyl sites identified in panel C) are illustrated. In A only the effects of internal dynamics are included, with overall tumbling removed by standard processing of the MD trajectory. In B overall tumbling is reintroduced via multiplication of the correlation function by the factor $\exp(-t/\tau_R)$. Correlation functions of the sort illustrated in Figure 8B are directly related to the deuterium quadrupolar relaxation rates under investigation, since to excellent approximation the quadrupolar interaction is axially symmetric with the symmetry axis along the C−D bond.

The correlation functions in Figure 8A,B experience a sharp drop from 1 to (1/9) as a consequence of fast methyl spinning.

(53) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. A. *Comput. Phys. Commun.* **1995**, *91*, 1−41.
(54) Beachy, M. D.; Chasman, D.; Murphy, R. B.; Halgren, T. A.; Friesner, F. A. *J. Am. Chem. Soc.* **1997**, *119*, 5908−5920.
(55) Kollman, P. A.; Dixon, R.; Cornell, W.; Fox, T.; Chipot, C.; Pohorille, A. *Comput. Simul. Biomol. Syst.* **1997**, 83−96.
(56) Zhang, O.; Forman-Kay, J. D. *Biochemistry* **1995**, *34*, 6784−6794.
(57) Lipari, G.; Szabo, A.; Levy, R. M. *Nature* **1982**, *300*, 197−198.
(58) Olejniczak, E. T.; Dobson, C. M.; Karplus, M.; Levy, R. M. *J. Am. Chem. Soc.* **1984**, *114*, 1923−1930.
(59) Palmer, A. G.; Case, D. A. *J. Am. Chem. Soc.* **1992**, *114*, 9059−9067.
(60) Chatfield, D. C.; Szabo, A.; Brooks, B. R. *J. Am. Chem. Soc.* **1998**, 5301−5311.

As a result of this down-scaling, it is difficult to analyze the trailing tails of these functions which are of critical importance in the determination of the relaxation properties of the system. To circumvent this problem, we have also computed the correlation functions for $C^{methyl}-C$ vectors, Figure 8C,D. These functions are sensitive to the same dynamic processes as C−H bonds with the exception of the fast methyl spinning, which makes them more suitable for visual analysis. Similar to the C−H correlations illustrated in A and B, the curves in panel D are derived from those in panel C by applying the multiplier $\exp(-t/\tau_R)$.

The four correlation functions considered in Figure 8 illustrate the diversity of local dynamics that have been observed for the methyl-containing side chains in the simulation. The upper profile in Figure 8C, drkN SH3 Ala 13, displays a steep drop followed by a plateau at approximately 0.85. The local motion in this case is limited to rapid small-amplitude fluctuations mainly involving torsional angles, typical of backbone dynamics. The next curve, derived from the coordinates of the side chain of drkN SH3 Leu 41δ2, reaches a plateau value of ∼0.60. The lower plateau level results from a cumulative effect involving fast fluctuations of several torsional angles along the side chain. The curve for drkN SH3 Leu 25δ2 drops rapidly to 0.40 and then undergoes a slow descent before leveling off at ∼0.15. Its counterpart, drkN SH3 Leu 25δ1, shows a similar dependence with a plateau very close to zero. In these cases, the steep decay in the correlation functions is the result of both low-amplitude fluctuations and fast rotameric transitions. In turn, slowly decaying components of the correlation functions, which are especially important for spin relaxation, are influenced by relatively infrequent rotameric transitions that occur on the time scale of several nanoseconds.

In Figure 8D overall tumbling has been included, and it is noteworthy that, despite the relative diversity in local dynamics, the correlation functions shown in the panel are, to excellent approximation, biexponential. Specifically, these functions are comprised of fast- and slow-decaying components. The characteristic times of the slow decay, determined from a biexponential fitting procedure applied to the curves in Figure 8D, are 3.6, 3.3, 2.5, and 1.6 ns for the drkN SH3 Ala 13, Leu 41δ2, Leu 25δ2, and Leu 25δ1 sites, respectively. Note that only for Ala is the slow decay of the correlation function determined entirely by overall tumbling, $\tau_R = 3.61$ ns. For all other residues, the effective decay times are substantially shorter, reflecting the effect of nanosecond-time-scale internal motion (i.e., side-chain isomerizations). The use of the single effective decay time, τ_c^{eff}, to describe the slowly decaying tail of the correlation function is the basis of the LS-3 model which is considered in this paper, eq 3. The MD data, therefore, lend support to use of the LS-3 model.

Correlation functions such as those shown in Figure 8B have been used to evaluate spectral densities, $J(\omega_q)$, at frequencies

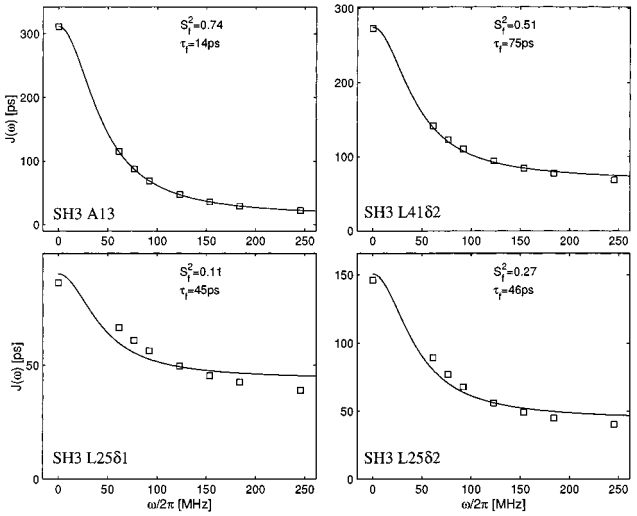

Figure 9. Fits of the spectral densities obtained from the MD trajectory of the drkN SH3 domain using the standard Lipari−Szabo (LS-2) model. The four methyl sites illustrated in Figure 8 are selected. The procedure used to extract the correlation functions from MD coordinates and evaluate the corresponding spectral densities is explained in Materials and Methods.

ω_q which match those accessed in the experimental study of protein L described above (see Materials and Methods). The simulated spectral densities were subsequently fitted with the Lipari−Szabo (LS-2) model[19,20] in the same manner as the experimental spectral densities discussed above. The results of the fitting procedure are presented in Figure 9 for the four methyl sites whose correlation functions are shown in Figure 8. For two out of four methyl groups, drkN SH3 Ala 13 and drkN SH3 Leu 41δ2, the simulated $J(\omega_q)$ data are well-fitted with the LS-2 model (top panels in Figure 9). In contrast, the fits are clearly unsatisfactory for the two other methyl groups, drkN SH3 Leu 25δ2 and, especially, drkN SH3 Leu 25δ1 (bottom panels in Figure 9), where the LS-2 model largely fails because it does not account for nanosecond-time-scale rotameric jumps that take place in these side chains.

Overall, the trends observed in fits of MD-based spectral densities are very similar to those encountered in the analysis of experimental data (cf. Figures 9 and 2). In both cases, the majority of residues can be fit fairly well using the standard LS-2 model, while a small fraction of residues show significant deviations from predicted behavior. This is strong evidence that the misfits in Figure 2B reflect the complexity of actual side-chain motions and do not arise from experimental artifacts.

Supporting Information Available: Tables of dynamics parameters for protein L at 25 and 5 °C and a figure with spectral density maps of a number of residues of protein L at 5 °C (analogue of Figure 2 from the text). This material is available free of charge via the Internet at http://acs.pubs.org.

JA012498Q