15 NH Backbone Dynamics of Protein GB1: Comparison of Order Parameters and Correlation Times Derived Using Various "Model-Free" Approaches

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · MARCH 2003

Impact Factor: 3.3 · DOI: 10.1021/jp022294b

3 AUTHORS, INCLUDING:

Djaudat Idiyatullin

Center for Magnetic Resonance Research Min…

74 PUBLICATIONS **614** CITATIONS

SEE PROFILE

^{15}NH Backbone Dynamics of Protein GB1: Comparison of Order Parameters and Correlation Times Derived Using Various "Model-Free" Approaches

Djaudat Idiyatullin, Vladimir A. Daragan, and Kevin H. Mayo*

Department of Biochemistry, Molecular Biology & Biophysics, University of Minnesota Health Science Center, 321 Church Street, Minneapolis, Minnesota 55455

Received: October 23, 2002; In Final Form: January 21, 2003

A new model-free approach that allows visualization of the distribution of motional correlation times is presented and is used to analyze ^{15}N NMR relaxation data [T_1, T_2, and {^{1}H}-^{15}N NOE], acquired at 500, 600, and 800 MHz, on the uniformly ^{15}N-enriched, 56 residue B1 domain from immunoglubulin protein G (GB1). Nanosecond time scale internal motions are found for all NHs of residues in protein GB1, a finding which is consistent with the concept of hierarchical internal motions in proteins first forwarded by Frauenfelder. Order parameters and overall tumbling correlation times derived using this approach are not influenced by the number of internal motional modes. Comparatively, use of the Lipari−Szabo, Clore et al., or LeMaster "model-free" methods yields underestimated order parameters and overestimated overall tumbling correlation times due to the presence of nanosecond time scale internal motions. This new model-free approach provides a straightforward way to derive more accurate order parameters and correlation times for overall tumbling and internal motions.

Introduction and Theory

For the last 2 decades, the Lipari and Szabo "model-free" approach has become the most popular way to analyze NMR relaxation data of macromolecules.[1,2] For this approach, the spectral density function $J(\omega)$ can be expressed as

$$J(\omega) = \frac{S^2\tau_o}{1+\omega^2\tau_o^2} + \frac{(1-S^2)\tau_i}{1+\omega^2\tau_i^2} \qquad (1)$$

where $\tau_i = \tau_o\tau_i^R/(\tau_o + \tau_i^R)$, τ_i^R is the correlation time for internal motions, τ_o is the correlation time for overall tumbling, and S^2 is the squared order parameter. The popularity of the Lipari−Szabo model arose because of its simplicity. Nevertheless, the method is not truly "model-free" since, during the derivation of eq 1, internal motions in the molecule were assumed to be governed by a single correlation time. In fact, there are only two instances for which eq 1 is valid: isotropic overall tumbling with internal rotational jumps between n equivalent states or between two nonequivalent states. Even a simple model with rotational jumps between three nonequivalent states (e.g., trans−gauch transitions) requires two internal motional correlation times. In proteins there are numerous internal motional correlation times that are required to fully describe various rotational fluctuations within potential wells and correlated rotational jumps within the backbone and side chains. Anisotropy of molecular overall tumbling makes the analysis even more complicated. For the most part, the Lipari−Szabo model provides a reasonable estimation of overall tumbling correlation times and order parameters, and it is most useful when only a limited amount of experimental data are available. However, even with a limited amount of data, Clore et al.[3] found that, in many cases, at least three Lorentzians are required to describe the spectral density function. With the current availability of high-quality NMR relaxation data acquired at multiple field strengths, the limitations of this approach become readily apparent.

Spectral density mapping[4] advanced the protein dynamics field; however, it too has limitations in that only a few points on the $J(\omega)$ curve, such as $J(0)$, $J(\omega_N)$, $J(\omega_H)$, are determined and these are difficult to interpret and nearly impossible to relate to parameters of molecular motion. In most cases, therefore, investigators still usually use the Lipari−Szabo[1,2] or Clore et al.[3] equations to interpret spectral density maps. The most useful "model-free" approach would allow derivation of motional correlation times and order parameters without assuming a priori the nature of the molecular motions involved. One way to achieve this is to use the sum of multiple Lorentzians to describe $J(\omega)$. This is a reasonable approach because it is valid for any Markovian type of motion. However, the problem here is not knowing the actual number of Lorentzians, i.e., the number of motional correlation times. A recent twist on the standard "model-free" approach, developed by LeMaster,[5] partially avoids this problem, but as will be shown later, is limited in its ability to analyze and differentiate nanosecond time scale motions.

Previously, we showed that a distribution of motional correlation times can be obtained without assuming a fixed number of correlation times by using the function[6]

$$F(\omega) = 2\omega J(\omega) \qquad (2)$$

This simplifies interpretation of NMR relaxation data because it presents the spectral density function in a more readily interpretable fashion. If, for example, two motional correlation times are well separated, $F(\omega)$ yields two maxima, each centered at some value of ω_i (the inverse of the correlation times, $1/\tau_i$) with each peak height being equal to the weighting coefficient

* To whom correspondence should be addressed. Tel: 612-625-9968. Fax: 612-624-5121. E-mail: mayox001@tc.umn.edu.

10.1021/jp022294b CCC: $25.00
Published on Web 02/21/2003

c_i in the Lorentzian expansion of the spectral density function. In general, $F(\omega)$ can be expressed as

$$F(\omega) = \frac{2c_0\omega\tau_o}{1+\omega^2\tau_o^2} + \frac{2c_1\omega\tau_1}{1+\omega^2\tau_1^2} + \ldots \quad (3)$$

where the number of correlation times must be as large as the number of availavle experimental parameters. $F(\omega)$ provides a simple representation of the motional correlation time distribution on the nanosecond to picosecond time scales. Furthermore, parameters derived from $F(\omega)$ are truly model free and practically do not depend on the number of motional modes.

In the case of anisotropic overall tumbling, eq 3 can be written as

$$F(\omega) = F_0(\omega) + F_i(\omega) = \sum_{k=0}^{m} \frac{2c_k\omega\tau_k}{1+\omega^2\tau_k^2} + \sum_{k=m+1}^{N} \frac{2c_k\omega\tau_k}{1+\omega^2\tau_k^2} \quad (4)$$

where c_k and τ_k with $k = 0, 1, \ldots, m$, are weighting coefficients and correlation times, respectively, that describe overall molecular tumbling. $F_0(\omega)$ and $F_i(\omega)$ are contributions to $F(\omega)$ from overall tumbling and internal motions, respectively. For symmetric top rotational diffusion with diffusion coefficients $D_{||}$ and $D_\perp$, $m = 2$, and c_k and τ_k can be expressed as

$$c_0 = S^2P_2^2(\cos(\theta)) \quad (5a)$$

$$c_1 = 3S^2\cos^2(\theta)\sin^2(\theta) \quad (5b)$$

$$c_2 = \frac{3}{4}S^2\sin^4(\theta) \quad (5c)$$

$$\tau_k = [6D_\perp + (D_{||} - D_\perp)k^2]^{-1} \quad (5d)$$

$P_2(\cos(\theta))$ is the second-order Legendre polynomial; θ is the angle between the motional vector and the symmetry axis. Correlation times τ_k with $k = m + 1, \ldots, N$ in eq 4 are combinations of internal motional and overall tumbling correlation times. The sums Σc_i are equal to

$$\sum_{k=0}^{m} c_i = S^2$$

$$\sum_{i=m+1}^{N} c_i = 1 - S^2 \quad (6)$$

To calculate $F(\omega)$, one needs to determine the coefficients c_i and correlation times τ_i for $i = 0, 1, \ldots, N$, with N being as large as possible for any given set of experimental data. For example, with three relaxation parameters [e.g., T_1, T_2, and NOE] acquired at three magnetic field strengths, nine fitting parameters, i.e., five Lorentzians in eq 3, can be determined by taking into account that $\Sigma c_i = 1$. Although it is practically impossible to obtain reliable values for c_i and τ_i for $N > 2$, linear combinations and other functions of c_i and τ_i over a given frequency range are very stable and can be determined accurately.[5,6] This is basically the key to obtaining the function $F(\omega)$ from experiment.

One way to calculate $F(\omega)$ is to use the Monte Carlo procedure and a simple step-by-step algorithm outlined below:

1. Randomly take five values of the internal motional correlation time τ_i, such that $t_1 < \tau_i < t_2$. t_1 should be about 100 ps because the shortest correlation time that can be determined by NMR is given by $1/(|\omega|_{NorC} + \omega_H)$. The value of t_1 can be smaller than 100 ps when contributions from very fast internal motions dominate the spectral density function, as in the case of methyl group rotations. To cover a range of possible errors and to account for the influence of rotational anisotropy, t_2 should be about 50−70% larger than the estimated overall correlation time τ_o for a given protein molecule. To estimate τ_o, the empirical equation can be used[7]

$$\tau_o = (9.18 \times 10^{-3}/T)\exp(2416/T)n_R^{0.93} \quad (7)$$

where n_R is the number of residues, and T is the temperature in Kelvin.

2. Using any minimization program, find appropriate values of c_i that best fit the experimental data, i.e., minimize the sum of the nine terms:

$$\chi^2 = \frac{1}{9}\Sigma(R_j - R_j^{calcd})^2/\sigma_i^2 \quad (8)$$

R_j are experimental parameters, R_j^{calcd} are calculated parameters, and σ_j are the experimental errors in determining R_j. If $\chi^2 < 1$, then store the set of c_i and τ_i. If $\chi^2 > 1$, then repeat steps 1 and 2. By repeating this process n times, one obtains n sets of c_i and τ_i.

3. For each kth set of c_i and τ_i, calculate $F(\omega)_k$ from eq 4. The probability P_k of finding the kth set of c_i and τ_i is[8]

$$P_k = \prod_{j=1}^{9} \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left[-\frac{(R_j - R_{jk}^{calcd})^2}{2\sigma_j^2}\right] \quad (9)$$

R_{jk}^{calcd} are calculated parameters for the kth set.

4. $5n$-weighted Lorentzians can now be used to describe $F(\omega)$. The average value of $F(\omega)$ and the standard deviation $\Delta(\omega)$ of $F(\omega)$ can be calculated as

$$F(\omega) = \Sigma p_k F(\omega)_k \quad (10a)$$

$$\Delta^2(\omega) = \Sigma p_k [F(\omega) - F(\omega)_k]^2 \quad (10b)$$

with $p_k = P_k/\Sigma P_k$.

A typical $F(\omega)$ curve is illustrated by the solid line in Figure 1. The maximum of the $F(\omega)$ curve $F(\omega)_{max}$ identifies $1/\tau_{max}$ on the ω axis. τ_{max} corresponds to low frequency (nanosecond time scale) motions, usually solely attributable to overall tumbling of the molecule if contributions from nanosecond time scale internal motions are small. The shoulders on the right or high-frequency side of $F(\omega)$ result from the presence of internal motions occurring at somewhat shorter correlation times. The tail of $F(\omega)$ (shortest correlation times) is prone to the greatest error because of experimental limitations, i.e., the limited range of accessible NMR spectrometer frequencies. Using modern NMR spectrometers, the smallest error in $F(\omega)$ is found over the frequency range up to $2\pi * (\omega_H + \omega_C) = (900 + 225)$ MHz = 7 GHz or 140 ps, which is the shortest correlation time that can be determined accurately. As mentioned above, this limit can be lower for groups undergoing very fast internal motions.

For isotropic overall tumbling and for well-separated correlation times of internal motions and overall tumbling, $F(\omega)_{max}$ is equal to the coefficient c_0 in eq 3. In this case, c_0 can be interpreted as the squared order parameter S^2 from the Lipari−Szabo model. However, when nanosecond time scale internal

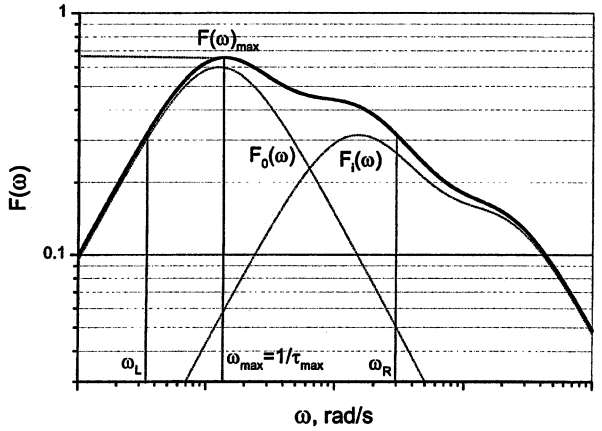

Figure 1. Example of an $F(\omega)$ function. A typical $F(\omega)$ curve (solid line) is illustrated. The curve was calculated using three Lorentzians with the following parameters: $\tau_o = 8$ ns, $\tau_1 = 0.7$ ns, $\tau_2 = 50$ ps, $c_o = 0.6$, $c_1 = 0.3$, and $c_2 = 0.1$. Results of deconvolution into $F_o(\omega)$ and $F_i(\omega)$ components (dotted lines) for overall tumbling and internal motional correlation time distributions, respectively, are also shown as labeled, along with τ_{max}, ω_L, and ω_R terms as discussed in the text.

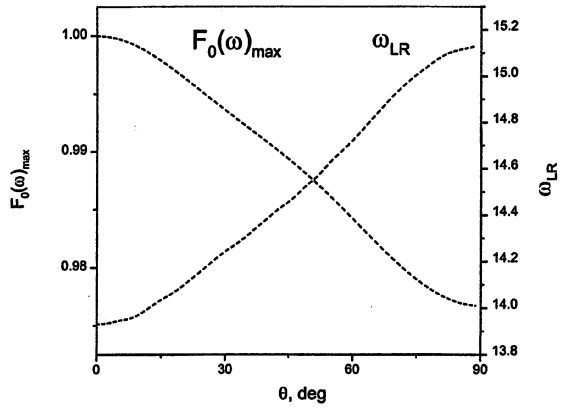

Figure 2. Calculated line width, ω_{LR}, and $F_0(\omega)_{max}$. The $F(\omega)$ line width at half-height ω_{LR} and the maximum value of $F_0(\omega)$, i.e., $F_0(\omega)_{max}$, have been calculated for the model of symmetric top rotational anisotropic diffusion and are plotted vs the angle θ that is formed between the motional vector and the symmetry axis of the molecule. The ratio of diffusion coefficients used in this calculation is equal to 2.

motions are present, overlapping Lorentzians lead to the inequality

$$F(\omega)_{max} > c_0 = S^2 \tag{11}$$

and contributed to the high frequency shoulder of $F(\omega)$. The line width of $F(\omega)$ ω_{LR} allows one to assess the presence of internal motions on the nanosecond time scale. ω_{LR} is determined from the ratio of frequencies:

$$\omega_{LR} = \omega_R/\omega_L \tag{12}$$

As indicated in Figure 1, ω_R and ω_L are the frequencies at

$$F(\omega_R) = 0.5\, F(\omega)_{max} \quad \text{for} \quad \omega_R > \omega_{max}$$

$$F(\omega_L) = 0.5\, F(\omega)_{max} \quad \text{for} \quad \omega_L < \omega_{max} \tag{13}$$

For tumbling of an isotropic molecule (single Lorentzian), ω_{LR}= 13.93 (ref 5). Therefore, when ω_{LR} is significantly greater than 13.93, it may be concluded that $F(\omega)$ is composed of at least two, and probably more, motional modes on the nanosecond time scale.

For overall tumbling of an anisotropic molecule, ω_{LR} will be somewhat greater, even in absence of internal motions, i.e., $S^2 = 1$. Rotational diffusion of a symmetric top with $D_{||}/D_{\perp} = 2$, illustrates this point. Figure 2 plots the dependence of ω_{LR} on the angle θ between some arbitrary motional vector and the symmetry axis of the molecule. It is apparent that for this model ω_{LR} remains less than 15.2 for all angles sampled. Note also that $F(\omega)_{max}$ is minimally affected, being reduced by only about 2.5% at $\theta = 90°$. It is also interesting that ω_{LR} depends linearly on $F(\omega)_{max}$ [$\omega_{LR} = 65.491 - 51.574\, F(\omega)_{max}$] with a correlation coefficient >0.99.

Because we have observed ω_{LR} to be often greater than 16 or so indicating that motional correlation time distributions for overall tumbling and internal motions overlap to various extents, it is important to be able to separate these motional modes as best possible. The remainder of this section is devoted to presenting a method to deconvolute $F(\omega)$ into components for overall tumbling $F_o(\omega)$ and internal motions $F_i(\omega)$ with the goal of increasing the accuracy in determining order parameters and motional correlation times. In general, deconvolution of $F(\omega)$ becomes less precise as internal motional and overall tumbling correlation times overlap more, and particularly when overall tumbling is anisotropic. Nevertheless, the goal of any deconvolution procedure is to separate, as best possible, correlation time distributions. Deconvolution can be aided by using various properties of $F(\omega)$.

First, consider $F_0(\omega)$. For anisotropic tumbling in the absence of internal motions, ω_{LR} for $F_o(\omega)$ cannot be larger than 15.2. Therefore, the largest correlation times, τ_i, which yield $F_0(\omega)$ with ω_{LR} less than 15.2, should be chosen. If the structure of the molecule is known, better separation of correlation times for internal motions and overall tumbling can be had by estimating the components and orientation of the rotational diffusion tensor [for example, by using the TENSOR-2 program[7]] and then calculating the correlation times τ_i and their coefficients c_i for use in determining values for ω_{LR} for individual motional vectors. For overall tumbling of an asymmetric top, for example, the number of terms that describe these motions is 5 and to calculate $F_0(\omega)$, only those terms with ω_{LR} < 15.2 should be used. For clarity, the following step-by-step deconvolution algorithm is provided:

1. Use the algorithm described above to obtain $5n$ sets of c_i and τ_i.

2. Sort these sets so that $\tau_{i+1} < \tau_i$

3. Select $m = 0, 1, 2 \ldots$, correlation times τ_i, and calculate ω_{LR} for the corresponding function $F_0(\omega)$ for all values of m.

4. Determine the maximum value of m that satisfies the condition $\omega_{LR} < 15.2$. Rather than using the value of 15.2, the number actually obtained from calculations using some model of anisotropic rotational diffusion can be used.

5. Determine the order parameter as $S^2 = c_0 + c_1 + \ldots + c_m$.

6. Calculate the function $F_i(\omega)$ for contributions from internal motions as

$$F_i(\omega) = F(\omega) - F_0(\omega)$$

7. Calculate the following parameters of $F_i(\omega)$: $F_i(\omega)_{max}$, ω^i_{LR}, τ^i_{max}. If $\omega^i_{LR} > 13.9$, then one can assume that the number of correlation times of internal motions is larger than 1. Parameters $F_i(\omega)_{max}$, ω^i_{LR}, and τ^i_{max} can serve as characteristics of the distribution of internal motional correlation times.

Figure 1 illustrates $F_o(\omega)$ and $F_i(\omega)$ components of $F(\omega)$ derived using this algorithm. These component curves are given as dotted lines in the figure.

Although there are some similarities between our approach and that recently published by LeMaster,[5] there are significant differences. LeMaster also used multi-Lorentzian expansion of the spectral density function $J(\omega)$ with a large number of terms to derive order parameters S^2 and S_f^2 without determining specific values for the correlation times and their weighting parameters. The main difference from our method and that of LeMaster is the use of a graphical presentation of the distribution of correlation times using the $F(\omega)$ function and the deconvolution of $F(\omega)$ into components for overall tumbling $F_o(\omega)$ and internal motions $F_i(\omega)$. The $F(\omega)$ approach allows one to obtain better separation of these motional distributions, and consequently, more accurate motional parameters. LeMaster attempted to separate overall tumbling and internal motional correlation times by introducing a frequency cutoff ($b\omega$), which "represents the slowest internal motion monitored". This $b\omega$ frequency is determined at the point when the root-mean-square deviations (rmsd) of predicted S^2 and S_f^2 values become larger than the anticipated uncertainty in the experimental data. LeMaster also varied the number of Lorentzians that describes the internal motions and showed that if this number is larger than 4, the standard deviations of S^2 and S_f^2 are essentially invariant on the NMR relaxation data collected at a single spectrometer frequency. With limited experimental data, the LeMaster method, like the Lipari–Szabo model, is best suited for studying spherical molecules and requires that the overall tumbling time be determined from independent measurements. For anisotropic overall tumbling, the number and values of overall tumbling correlation times depend on the orientation of the motional vector within the molecular frame, which makes using the LeMaster approach difficult. If, however, NMR relaxation data are acquired at multiple spectrometer frequencies, overall tumbling correlation times can be determined during the minimization protocol. Using the LeMaster algorithm,[5] we found that the frequency $b\omega_N$ cannot be accurately determined because simultaneous variations of S^2 and τ_0 make rmsd of S^2 only slightly dependent on $b\omega_N$, such that the crucial selection of $b\omega_N$ becomes highly subjective, thereby introducing additional uncertainty or error in derived motional parameters. Even with the smallest estimated values of $b\omega_N$ (2.5 ns^{-1}), order parameters derived using the LeMaster approach are larger than those derived using the $F(\omega)$ deconvolution method. The $F(\omega)$ approach provides better separation of internal motional and overall tumbling correlation times and, therefore, more accurate order parameters and correlation times.

Materials and Methods

Protein Production. The 56-residue protein GB1 was produced as a recombinant protein as described by Barchi et al.[9] *E. coli* containing the expression system for GB1 were grown on M9 minimal media containing ^{15}N-ammonium, and GB1 was uniformly isotopically enriched in ^{15}N. The protein was purified by HPLC using a linear acetonitrile/water gradient, and purity was checked by MALDI-TOF mass spectrometry and analytical HPLC on a C18 Bondclone (Phenomenex) column. For NMR measurements, freeze-dried samples were dissolved in a H_2O/D_2O (95/5) mixture in 20 mM potassium phosphate. Protein concentration, determined from the dry weight of freeze-dried samples, was 10 mg/mL. The pH was adjusted to pH 5.25 by adding microliter quantities of NaOD or DCl.

NMR Relaxation Experiments. With uniformly ^{15}N-enriched GB1, spin–lattice (T_1), spin–spin (T_2) relaxation times, and $\{^1H\}$-^{15}N NOEs were measured at three Larmor precession frequencies (^1H frequencies of 500, 600, and 800 MHz) on Varian Inova 500, 600, and 800 NMR spectrometers equipped with triple-resonance probes. The temperature was 5 °C. Temperature calibration was performed by using the chemical shifts of resonances from methanol.

^{15}N spin–lattice and spin–spin relaxation rates were measured by using the HSQCSE sequence,[10] which employs pulsed field gradients for the coherence transfer pathway whereby magnetization passes from ^1H to ^{15}N and back again to ^1H for observation. The water flip-back method was used to minimize water saturation during the pulse sequence. The delay in the CPMG train was set to 0.625 ms. To attenuate cross-correlation between dipolar and chemical- shift anisotropy mechanisms during the relaxation period, an even number of 180° ^1H pulses with alternating phase were applied every 5 ms.[11,12] Spectra were recorded by using relaxation delays 20, 40, 60, 90, 160, 250 and 360 ms. During the relaxation delay, t_R, water magnetization affected by these ^1H 180° pulses decreases proportionally to $\exp(-t_R/T_{1\ \mathrm{water}})$. Consequently, in the presence H(N)- water proton exchange, conditions will be different at the beginning of the first INEPT in the pulse sequence for experiments acquired with different values of the relaxation delay and T_1 (or T_2) will be underestimated. To eliminate this effect, a compensation period (CP) with the same composite pulse as in the relaxation period was included prior to the recycle period at the beginning of the pulse sequence. The compensation period varies in length with the relaxation time t_R and is defined by $CP = t_{max} - t_R$, where t_{max} is the maximal relaxation time for a given set of experiments. By using this scheme, all experiments were heated equivalently and yielded the same average value for the saturation of water and protein resonances. Recovery of magnetization of nuclei from the protein depends only on the recycle period because ^1H 180° pulses prohibit recovery of magnetization during the compensation period. For all T_1 and T_2 experiments, the recycle period was 1.7 s, and all relaxation curves followed single-exponential decay.

Steady-state $\{^1H\}$-^{15}N NOEs were determined from 3 spectra recorded with proton broad band irradiation and in the absence of proton saturation with different water saturation conditions to compensate for H(N)–water proton exchange.[13] Saturation was achieved by application of 120° ^1H pulses applied every 5 ms (ref 14) during the 3 s recycle period.

The function $F(\omega)$ was calculated using the program FRELAN, which was developed in our lab and is available at www.nmr-relaxation.com. The number of correlation times, N, used to describe $F(\omega)$ was usually 5. Calculations with $N = 2$ (Lipari–Szabo), $N = 3$ (Clore et al.) and $N = 4$ were also performed. Compared to the statistical error χ^2 for $N = 2$, relative errors for $N = 3$, $N = 4$, and $N = 5$ are 0.03, 0.01, and 0.006, respectively.

Results and Discussion

^{15}N NMR relaxation data [R_1 (= $1/T_1$), R_2 (= $1/T_2$), and $\{^1H\}$-^{15}N NOE] on uniformly ^{15}N-enriched protein GB1 were acquired at three magnetic field strengths, ^1H frequencies of 500, 600, and 800 MHz. Relaxation data at 500 and 800 MHz are shown in Figure 3, along with secondary structure elements for the protein. Note that the N-terminus, turn 1 and loop, demonstrate smaller R_2 and NOE values, which is consistent with these segments of the protein being more mobile. These trends are less prominent in the R_1 data. To better follow the discussion

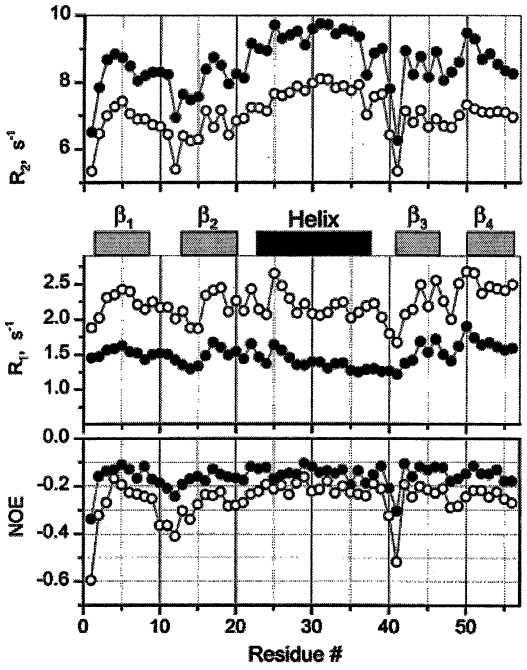

Figure 3. R_1, R_2 and NOE data for protein GB1. ^{15}N NMR relaxation data [R_1 (= $1/T_1$), R_2 (= $1/T_2$), and {^{1}H}-^{15}N NOE] on uniformly ^{15}N-enriched protein GB1 are plotted vs the GB1 residue number for data acquired at ^{1}H spectrometer frequencies of 500 MHz (filled-in circles) and 800 (open circles) MHz. Secondary structure elements for the protein are given between panels. The temperature for all measurements was 5 °C.

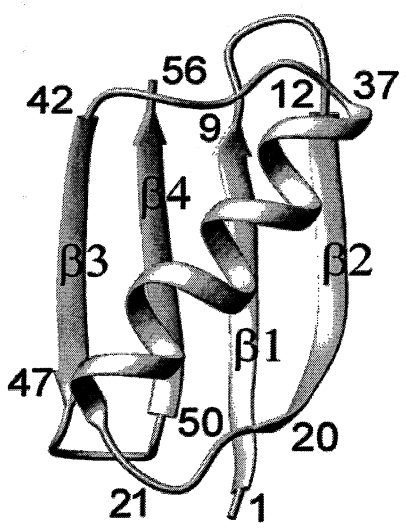

Figure 4. Overall fold for GB1. The overall fold of GB1 is illustrated as given by Gronenborn et al.[15]

concerning structural dynamics in the protein, the overall fold of GB1 (ref 15) is illustrated in Figure 4.

Motional Correlation Time Distributions. Figure 5 plots $F(\omega)$ vs ω for NH groups of five residues in protein GB1: M1 (β-strand 1), T11 (turn 1), A24 (α-helix), G41 (loop), and T55 (β-strand 4). Similar appearing curves were obtained for all NH groups of residues throughout the protein. Perhaps the most striking feature in these curves is the presence of at least two motional correlation time distributions (multiple Lorentzians) on the nanosecond time scale, one for overall tumbling and the other(s) for internal motions. This is most evident with distribu-

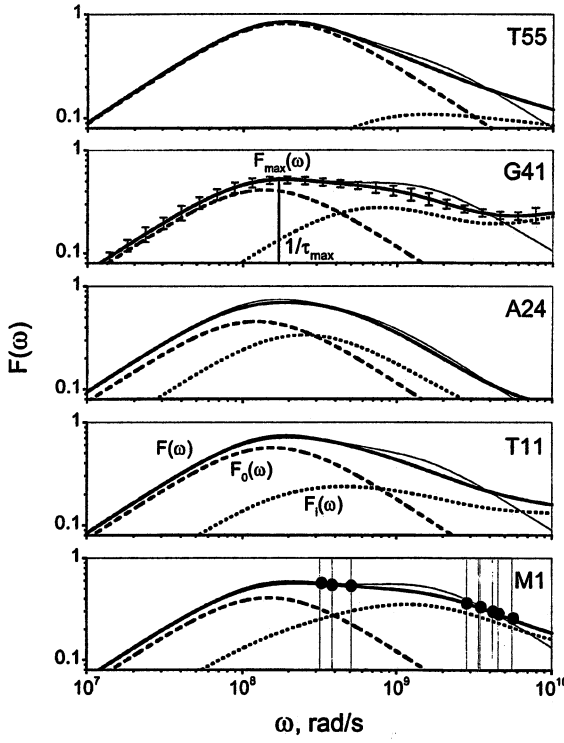

Figure 5. Plots of $F(\omega)$, $F_o(\omega)$, and $F_i(\omega)$ curves for five residues in GB1. Because ω is inversely related to the correlation time, $F(\omega)$ gives the distribution of motional correlation times for NH fluctuations over the nanosecond to picosecond range. $F(\omega)$ motional correlation time distributions are exemplified with five NH groups: M1 (β-strand 1), T11 (turn 1), A24 (α-helix), G41 (loop), and T55 (β-strand 4) through the structure of GB1. $F(\omega)$ curves (thick solid lines) were deconvoluted into $F_o(\omega)$ (dashed lines) and $F_i(\omega)$ (dotted lines) component curves as discussed in the text. Thin solid lines give the summation of three Lorentzians resulting from the analysis using the Clore et al. model free approach. Data used were acquired at 5 °C.

tions for M1 and G41, whereas those for the other three NHs shown are only broadened on their high-frequency side. An initial structural interpretation for this is that internal motional correlation time distributions for NHs of more mobile residues, M1 at the N-terminus and G41 in the loop connecting the α-helix and β-strand-3, are shifted to shorter correlation times relative to NHs of residues in more 'rigid' structural elements. This is what one would expect given the trends in the relaxation data noted above and lends confidence to the new approach.

To semiquantify contributions from nanosecond time scale internal motions in these correlation time distributions ω_{LR}, the width of $F(\omega)$ at half-height, was used as discussed in the Theory Section. A single Lorentzian associated with overall tumbling of a spherical molecule has a ω_{LR} value of 13.93 (ref 6). Protein GB1, however, is ellipsoid in shape with a $D_\parallel/D_\perp$ ratio of 1.8. Because of this, ω_{LR} also depends on the orientation of a particular N−H vector within the molecular frame of folded GB1. Using the program TENSOR II (ref 7) and the coordinates for GB1 averaged over 60 NMR-derived structures from the PDB database [access code: GB1], values of ω_{LR} for N−H vectors in GB1 were calculated and found to range from 14 to 15.5 as illustrated in Figure 6A. NHs of the α-helix (residues 22−37), for example, are mostly oriented parallel to the long axis of the molecule, yielding smaller values of ω_{LR}, whereas those in most other segments of the protein are oriented more perpendicular to this axis, resulting in larger values of ω_{LR}. Experimentally determined values for ω_{LR} are plotted vs residue

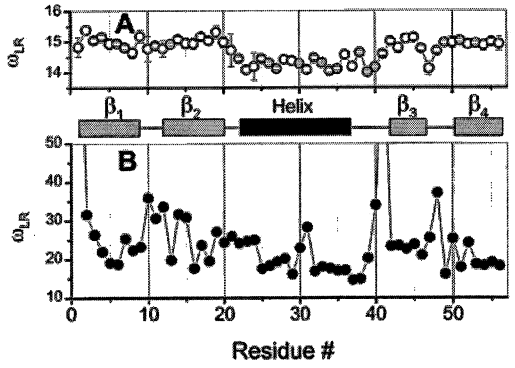

Figure 6. ω_{LR} plotted vs residue number from GB1. Calculated values of ω_{LR} are plotted for NH vectors in GB1 (A) and for experimentally determined values of ω_{LR} (B). Values of ω_{LR} for N−H vectors in GB1 were calculated using the program TENSOR II[7] and the coordinates for GB1 averaged over 60 NMR-derived structures from the PDB (accession code GB1). Experimentally determined values of ω_{LR} are plotted for 5 °C. For calculated values of ω_{LR}, the scale is more expanded to illustrate the effect through the sequence. Secondary structure elements in GB1 are shown between panels.

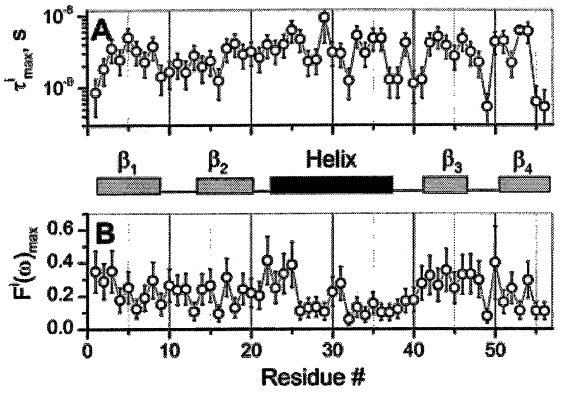

Figure 7. $F_i(\omega)_{max}$ and τ^i_{max} values plotted vs the residue number. τ^i_{max} (A) and $F_i(\omega)_{max}$ (B) values are plotted vs the GB1 residue number. τ^i_{max} represents the inverse of the frequency ω at $F_i(\omega)_{max}$. Errors are exemplified by vertical bars. Secondary structure elements in GB1 are shown between parts.

number in Figure 6B. Note the different scales for ω_{LR} used in Figures 6A and 6B. Overall, ω_{LR} is much larger than expected for overall tumbling alone (Figure 6A). This indicates that for most backbone NHs in the protein, there are significant contributions from internal motions on the nanosecond time scale (Figure 6B). As noted above for M1 and G41 (Figure 5), a larger value of ω_{LR} indicates that the internal motional correlation time distribution is shifted to higher frequency (shorter correlation time), consistent with relatively greater internal mobility. The largest values of ω_{LR} are associated with NHs located at the N-terminus, turn 1 into β-strand 2, the loop at residues D40 and G41, and residue A48 within the final turn.

Overall and internal motional correlation time distributions, $F_o(\omega)$ and $F_i(\omega)$, were derived from $F(\omega)$ curves as described in the Theory Section. $F_i(\omega)$ curves for residues M1, T11, A24, G41, and T55 are illustrated in Figure 5 with dotted lines, whereas $F(\omega)$ curves are given by solid lines and $F_o(\omega)$ by dashed lines. Error corridors for $F(\omega)$ are exemplified with vertical error bars for G41, and experimental frequencies are indicated for M1 by solid circles and vertical lines. Note that $F(\omega) = F_o(\omega) + F_i(\omega)$. The breadth of these distributions indicates that they are composed of multiple Lorentzians. In other words, one internal motional correlation time, not surprisingly, is insufficient to fully describe NH rotational fluctuations, which occur on the nanosceond to picosecond time scales. Nanosecond time scale internal motions have been observed by few authors. Orechov et al.[16] reported that the reason for this is that detection of nanosecond time scale internal motions is difficult when one uses NMR relaxation data acquired at a single spectrometer frequency and when the overall tumbling correlation time is estimated from fitting these limited experimental data. Orechov et al. analyzed NMR relaxation data collected at two frequencies and found significant contributions from nanosecond time scale internal motions.

Each $F_i(\omega)$ distribution (Figure 5) can be characterized by the position of its maximum τ^i_{max}, which are plotted vs residue number in Figure 7A. Notice that for most NHs of residues in more structured, hydrogen-bonded regions (β-strands 1, 2, and 3 and the helix), τ^i_{max} falls between 2 and 5 ns, whereas for other, more mobile NHs, τ^i_{max} lies closer to 1 ns or less. Because these distributions are broad, and in cases rather flat, this parameter can be deceptive, and one should always also view the entire $F_i(\omega)$ distribution. Relative contributions from these internal motions to $F(\omega)$ are given, to a first approximation, by $F_i(\omega)_{max}$. In general, $F_i(\omega)_{max}$ falls between 0.1 and 0.3 (Figure 7B), indicating that internal motions occurring on the nanosecond time scale contribute about 10−30% to the spectral density function. Some sense of their contribution on the picosecond time scale is indicated by the tail on $F_i(\omega)$ (Figure 5), which tapers off to less than 0.1 in most cases. In other words, there is a 10% or less contribution from picosecond time scale internal motions to the spectral density function. For more mobile residues, e.g., M1 (N-terminus) and G41 (post-helix loop), contributions on the picosecond time scale are greater, an observation that is not unexpected.

Analysis Using Other Model Free Approaches. For comparison to these parameters derived using this new model free approach, ^{15}NH NMR relaxation data were also analyzed using the Lipari and Szabo,[1,2] Clore et al.,[3] and LeMaster[5] approaches. For Lipari and Szabo and Clore et al. methods, a Monte Carlo minimization of the χ^2 function given by eq 8 was performed separately for each residue to take into account the dependence of the overall tumbling correlation time on the orientation of individual NH bonds within the molecular frame. Of the numerous solutions were generated, those with $\chi^2 < 1$ were considered for further weighted-averaging as described in the Theory Section. LeMaster's method, which was also used with 4 internal motional correlation times, yielded a limit for the lowest frequency for the internal motional correlation time of 2.5 ns^{-1}. This limit falls near the sequence-averaged value for τ^i_{max} that was found using the $F(\omega)$ approach, indicating that the LeMaster approach is not accounting for much of distribution of internal motional correlation times.

Order parameters S^2, derived using the Lipari−Szabo, Clore et al., and $F(\omega)$ methods, are plotted in Figure 8A (Lipari and Szabo compared to the $F(\omega)$ approach) and 8B (Clore et al. compared to the $F(\omega)$ approach). Because the LeMaster method gave results similar to those derived using the Clore et al. approach, they are not shown here. In both parts A and B, darkened bars are S^2 values derived using the $F(\omega)$ approach. With the Lipari−Szabo method, S^2 values are generally much larger than those derived using the $F(\omega)$ approach, as well as using the Clore et al. method. This may be explained simply because the Lipari− Szabo model uses only two Lorentzians and nanosecond time scale internal motions contribute significantly to c_o (=S^2). This means that the Lipari−Szabo approach

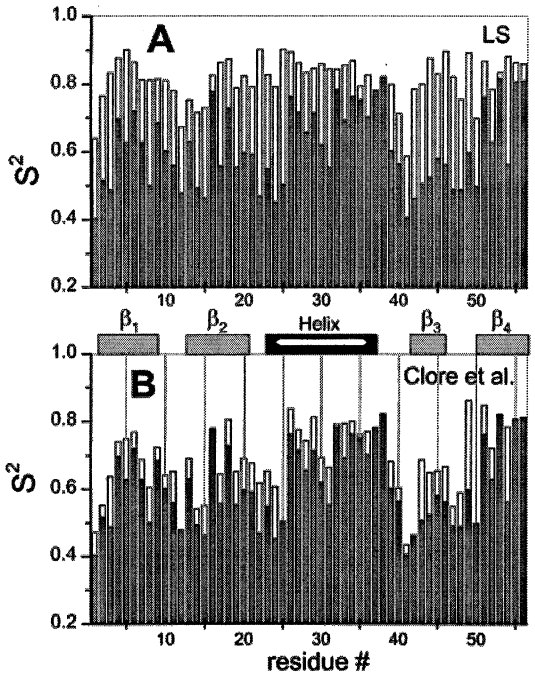

Figure 8. Order parameters obtained using three "model-free" approaches. S^2 values, derived using the Lipari–Szabo, Clore et al., and $F(\omega)$ approaches, are shown. Part A compares S^2 from the Lipari–Szabo (LS) and $F(\omega)$ approaches, and part B compares S^2 from the Clore et al. and $F(\omega)$ approaches. In both parts, darkened bars give S^2 values derived using the $F(\omega)$ approach, and open bars give S^2 values derived using either the Lipari–Szabo or Clore et al. approach. Secondary structure elements in GB1 are shown between panels.

effectively yields overestimated S^2 values. Use of a third Lorentzian in the Clore et al. approach alleviates some of this problem, but nonetheless gives less accurate order parameters. This inaccuracy is reflected in the sum of the three Lorentzians shown as a thin line shown in Figure 5 on top of the $F(\omega)$ functions. Notice that these summations do not fall precisely on the $F(\omega)$ curve and at various points, they can deviate by up to about 0.1 or more. Use of the $F(\omega)$ approach avoids computation of individual coefficients and yields an order parameter that has been determined by subtracting all components due to internal motions without requiring the use of a set number of Lorentzians. This is advantageous because it produces more accurate motional parameters.

As shown in Figure 9, overall tumbling times τ_o determined using the $F(\omega)$ approach and that of Clore et al. also compare more favorably to each other than to those determined using the Lipari–Szabo model. In addition, τ_o values from the Lipari–Szabo model are mostly smaller, giving a sequence-averaged value of 5.5 ns, compared to 6.5 ns from the $F(\omega)$ approach, 6.4 ns from the Clore et al. approach and 6.2 ns from the LeMaster approach. τ_0 values derived using a semiempirical eq 7 (8 ns) suggest that the actual τ_0 value should be at the upper end of this 5.5 ns to 6.5 ns range. The $F(\omega)$ method does give the largest value of $\langle\tau_0\rangle$. Another interesting observation is to compare τ_{max} values obtained at the maxima of $F(\omega)$ curves to τ_o values derived from the Lipari–Szabo model. They are nearly superimposable. This indicates two things: (1) τ_{max} essentially yields τ_o values as derived from the Lipari–Szabo approach, and (2) τ_o values derived using the Lipari–Szabo approach are underestimated for the same reason that S^2 values are overestimated. τ_{max} will be the same as the actual τ_o value only for spherical molecules and in the absence nanosecond

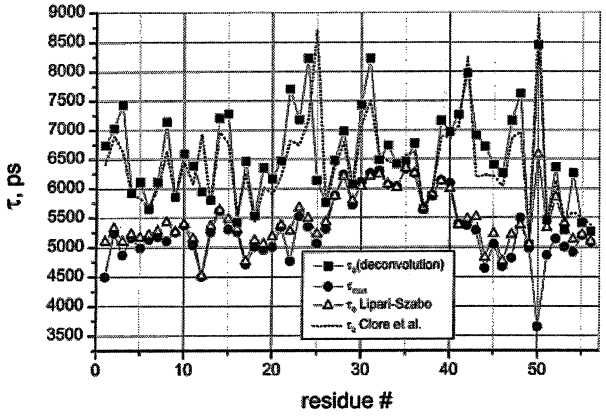

Figure 9. Correlation times obtained using three "model-free" approaches. Overall tumbling correlation times τ_o derived using the Lipari–Szabo (open triangles) and Clore et al. (dashed line) model free approaches, along with τ_o values obtained using the new approach (solid squares), are shown. In addition, τ_{max} values, read off the ω axis at $F(\omega)_{max}$, are shown as filled-in circles. Data used were acquired at 5 °C.

time scale internal motions. When internal motions on the nanosecond time scale are present, τ_{max} will be smaller than the actual τ_o value because the sum of Lorentzians tends to shift $F(\omega)_{max}$ to higher frequency ($1/\omega$). Comparison of overall tumbling correlation times determined by deconvoluting $F(\omega)$ with respective correlation times calculated using the program TENSOR-II (ref 7) shows a linear dependence with a correlation coefficient of 0.74.

Although average τ_o values derived using the $F(\omega)$, Clore et al. and LeMaster approaches compare favorably, average S^2 values do not. From the Lipari–Szabo, Clore et al., LeMaster and $F(\omega)$ approaches, $\langle S^2\rangle$ values are 0.82, 0.68, 0.7, and 0.62, respectively. While the Clore et al. and LeMaster methods yield nearly the same values of 0.68 and 0.7, the Lipari–Szabo model gives an $\langle S^2\rangle$ value (0.82) that is much greater. Because the $F(\omega)$ method, which gives the lowest value for $\langle S^2\rangle$, gives the best separation of internal motional and overall tumbling correlation times, the $\langle S^2\rangle$ value of 0.62 derived using the $F(\omega)$ method should be the most accurate from these four approaches. Furthermore, it should be mentioned that although overall trends in S^2 values through the sequence using the Clore et al., LeMaster, and $F(\omega)$ approaches are similar, the actual values can be quite different for reasons discussed above. This, in turn, makes having the most accurate S^2 values crucial for deriving conformational entropy[21–23] and interpreting dynamics changes for protein–ligand or protein–protein interactions. For example, an order parameter derived using the Lipari–Szabo, Clore et al., or LeMaster methods may be found to increase upon protein complex formation, suggesting greater motional restriction and a decrease in entropy. However, if the internal motional correlation time distribution were actually shifted more toward that of overall tumbling and not corrected for, then this might be an erroneous interpretation that has more to do with shifting motional distributions than with motional restrictions and changes in conformational entropy.

Comparison to Other Studies on GB1. There are two other ^{15}N NMR relaxation studies on protein GB1 (refs 9 and 17). In both cases, relaxation data were analyzed using four variations of the Lipari–Szabo approach[1,2] as well as the Clore et al. model[3] for NHs of some residues where it was required to better explain the data. Overall, generalized order parameters S^2 run high compared to those derived using our new model free

approach. In Barchi et al.,[9] where data were acquired at 26 °C, a few S^2 values are greater than 0.9 and most are between 0.8 and 0.9. In Seewald et al.,[17] where data were acquired in 10 °C increments from 0 to 50 °C, order parameters lie mostly between 0.7 and 0.8 for an interpolated temperature of 25 °C. At 5 °C where our data were acquired, Seewald et al.[17] give S^2 values mostly between 0.8 and 0.9 compared to mostly between 0.6 and 0.7 in our study. There are two main reasons for discrepancies in the magnitude of derived order parameters among these three studies: (1) differences in the amount of data acquired and (2) the model used to analyze the data. In Seewald et al.,[17] ^{15}N relaxation data were acquired at only one magnetic field strength (500 MHz), and in Barchi et al.,[9] data were acquired at two magnetic field strengths (500 and 600 MHz). Here, we used three magnetic field strengths (500, 600, and 800 MHz). However, even if we use our data for two magnetic field strengths, $F(\omega)$ curves are essentially the same; therefore, observed differences among these studies have more to do with differences in the model free approach used to analyze the data. Here, we have already demonstrated that use of the Lipari–Szabo model results in overestimated order parameters, which is supported by the comparison made with data from the Barchi et al. and Seewald et al. studies.

For a few residues in each of these previous studies, a second time scale of nanosecond internal motions was required to adequately fit the relaxation data. However, the set of residues exhibiting such motions was not consistent between studies, nor between temperatures in the Seewald et al. study. This is most likely attributable to the fact that τ_s values (correlation times for nanosecond time scale internal motions) from the Clore et al. approach can be derived with limited precision.[18,19] Unlike these two previous studies, we find that NHs of all residues exhibit nanosecond time scale internal motions. Adjusting for contributions from nanosecond time scale internal motions using the $F(\omega)$ approach allows S^2 values to be determined more accurately. Derivation of nanosecond time scale internal motions and accurate order parameters is crucial to a better, more detailed understanding of protein dynamics[20] and to the determination of more accurate residual entropy contributions using derived order parameters.[21–23]

According to S^2 values derived using $F(\omega)$ approach (Figure 8), the least internally mobile NHs are associated with residues within the C-terminal part of the α-helix (Q32 to N37), which is structurally consistent with these NHs usually forming a network of stronger hydrogen bonds. The G38 NH, which is hydrogen bonded to carbonyls of A34 and N35, is also highly motionally restricted, one of the most restricted in the protein. On the other hand, the first turn of the helix, D22 to A24, is somewhat more flexible, consistent with the structural fact that the first N-terminal turn in a helix either has yet to form hydrogen bonds or has formed only weak ones. In addition, NHs of residues in β-strands 1 and 4 (the two internal strands in the four-stranded β-sheet) are, on average, less internally mobile than those from external β-strands 2 and 3. These moto-structural correlations were not as evident in the other dynamics studies on protein GB1.

Conclusions

Here, we have demonstrated that the $F(\omega)$ approach, which uses a large number of correlation times to analyze NMR relaxation data, provides a straightforward way to visualize internal motional correlation time distributions, is more consistent with theoretical considerations, and allows derivation of more accurate order parameters. Researchers should be cautioned that use of other model free approaches may yield overestimated order parameters and underestimated overall tumbling correlation times when internal motions are present on the same nanosecond time scale as overall tumbling. Nanosecond time scale internal motions are found for all NHs of residues in protein GB1, suggesting that all motional vectors in proteins may likely exhibit similar nanosecond time scale motions. This hypothesis is consistent with the concept of heirarchical internal motions in proteins first forwarded by Frauenfelder et al.[20]

Acknowledgment. This work was supported by a research grant from the National Institutes of Health (NIH, GM-58005) and benefitted from use of the high field NMR facility at the University of Minnesota. We thank Judy Haseman for preparing ^{15}N-enriched samples of protein GB1.

References and Notes

(1) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546.
(2) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4559.
(3) Clore, G. M.; Szabo, A.; Bax, A.; Kay, L. E.; Driscoll, P. C.; Gronenborn, A. M *J. Am. Chem. Soc.* **1990**, *112*, 4989.
(4) Peng, J. W.; Wagner, G. *J. Magn. Reson.* **1992**, *98*, 308.
(5) LeMaster, D. M. *J. Am. Chem. Soc.* **1999**, *121*, 1726.
(6) Idiyatullin, D.; Daragan, V. A.; Mayo, K. H. *J. Magn. Reson.* **2001**, *152*, 132.
(7) Daragan, V. A.; Mayo, K. H. *Progress NMR Spectrosc.* **1997**, *32*, 63.
(8) Andrec, M.; Montelione, G. T.; Levy, R. M. *J. Magn. Reson.* **1999**, *139*, 408.
(9) Barchi, J. J.; Grasberger, B.; Gronenborn, A. M.; Clore, G. M. *Protein Sci.* **1994**, *3*, 15.
(10) Farrow, N. A.; Zhang, O.; Szabo, A.; Torchia, D. A.; Kay, L. E. *J. Biomol. NMR* **1995**, *6*, 153.
(11) Kay, L. E.; Nicholson, L. K.; Delaglio, F.; Bax, A.; Torchia, D. A. *J. Magn. Reson.* **1992**, *97*, 359.
(12) Palmer, A. G. III; Skelton, N. J.; Chazin, W. J.; Wright, P. E.; Rance, M. *Mol. Phys.* **1992**, *75*, 699.
(13) Idiyatullin, D.; Daragan, V. A.; Mayo, K. H. *J. Magn. Reson.* **2001**, *153*, 138.
(14) Markley, J. L.; Horsley, W. J.; Klein, M. P. *J. Chem. Phys.* **1971**, *52*, 3604.
(15) Gronenborn, A. M.; Filpula, D. R.; Essig, N. Z.; Achari, A.; Whitlow, M.; Wingfield, P. T.; Clore, G. M. *Science* **1991**, *253*, 657.
(16) Orekhov, V. Y.; Korzhnev, D. M.; Pervushin, K. V.; Hoffmann, E.; Arseniev A. S. *J. Biomol. Struct.* **1999**, *17*, 157.
(17) Seewald, M. J.; Pichumani, K.; Stowell, C.; Tibbals, B. V.; Regan, L.; Stone, M. J. *Protein Sci.* **2000**, *9*, 1177.
(18) Stone, M. J.; Chandrasekhar, K.; Holmgren, A.; Wright, P. E.; Dyson, H. J. *Biochemistry* **1993**, *32*, 426.
(19) Jin, D.; Andrec, M.; Montelione, G. T.; Levy, R. M. *J. Biomol. NMR* **1998**, *12*, 471.
(20) Frauenfelder, H.; Sligar, S. G.; Wolynes, P. G. *Science* **1991**, *254*, 1598.
(21) Akke, M.; Bruschweiler, R.; Palmer, A. G., III. *J. Am. Chem. Soc.* **1993**, *115*, 9832.
(22) Yang, D.; Kay, L. E. *J. Mol. Biol.* **1996**, *263*, 369.
(23) Lee, A. L.; Wand, A. J. *Nature* **2001**, *411*, 501.