Analysis of Slow Interdomain Motion of Macromolecules Using NMR Relaxation Data

James L. Baber,[‡] Attila Szabo,[§] and Nico Tjandra*[,‡]

Contribution from the Laboratory of Biophysical Chemistry, Building 3, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892-0380, and Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892

Received December 6, 2000. Revised Manuscript Received February 19, 2001

Abstract: The interpretation of NMR relaxation data for macromolecules possessing slow interdomain motions is considered. It is shown how the "extended model-free approach" can be used to analyze ^{15}N backbone relaxation data acquired at three different field strengths for *Xenopus* Ca^{2+}-ligated calmodulin. This protein is comprised of two domains connected by two rigid helices joined by a flexible segment. It is possible to uniquely determine all "extended model-free" parameters without any a priori assumptions regarding their magnitudes by simultaneously least-squares fitting the relaxation data measured at two different magnetic fields. It is found that the two connecting helices (and consequently the domains) undergo slow motions relative to the conformation in which the two helices are parallel. The time scales and amplitudes of these "wobbling" motions are characterized by effective correlation times and squared-order parameters of approximately 3 ns and 0.7, respectively. These values are consistent with independent estimates indicating that this procedure provides a useful first-order description of complex internal motions in macromolecules despite neglecting the coupling of overall and interdomain motions.

Introduction

The analysis of NMR relaxation data of macromolecules that undergo only fast, small amplitude backbone motions is straightforward.[1,2] This is not the case, however, when more complex and, hence, more interesting internal motions are present. An example of such a system is *Xenopus* Ca^{2+}-ligated calmodulin, a 148 residue protein, which consists of two modules connected by a helical linker that is flexible in the middle.[3] Earlier results suggest that this flexibility permits the central helix to have temperature-dependent conformational variability in the crystalline environment,[4] and it also allows the two domains to adopt various relative orientations in solution.[5−7] Relative motion of the two domains is functionally important. Calmodulin (CaM) is involved in the transduction of extracellular Ca^{2+} signals into a cellular response.[8] One of the important signaling pathways involves the binding of CaM to various enzymes that all possess a common helical CaM target peptide. This helical peptide is essential for CaM recognition and regulation. The interaction between CaM and its target peptide is mainly hydrophobic in which the target peptide is clamped between the two domains of CaM.[9] The conformational change in CaM associated with peptide binding is clearly facilitated by flexibility of the central helix linker.

In this report, we consider how NMR relaxation data for such a system should be analyzed to estimate the time scale and amplitude of interdomain motions. At the outset, it is noted that, since the overall and interdomain motions are coupled, there is no simple, yet rigorous, procedure or formalism for analyzing the relaxation data for these types of systems. In this study, the simplest possible approach is adopted which nevertheless captures the main physical features of the problem. Specifically, we assume that the system can be described by the overall motion (which can be anisotropic) of a frame attached to the dynamically averaged structure of the macromolecule. Slow and fast internal motions occur relative to this frame. The correlation function that describes such internal motions is identical to that used in the "extended model-free approach" proposed by Clore et. al.[10] for proteins where the backbone dynamics of some residues needs to be described by at least two time scales. The slow internal motion in the present case is the motion of an entire domain relative to the frame that undergoes overall rotational diffusion.

This sort of approach has already been used to interpret backbone ^{15}N relaxation data of Ca^{2+}-free (Apo) calmodulin.[11] However, several simplifying assumptions were made due to

* To whom correspondence should be addressed: Building 3, Room 418, NIH, Bethesda, MD 20892-0380. Telephone (301) 402-3029. Fax (301) 402-3405. E-mail: nico@helix.nih.gov.

‡ Laboratory of Biophysical Chemistry, National Heart, Lung, and Blood Institute.

§ Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases.

(1) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4559.

(2) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4559−4570.

(3) Barbato, G.; Ikura, M.; Kay, L. E.; Pastor, R. W.; Bax, A. *Biochemistry* **1992**, *31*, 5269−5278.

(4) Wilson, M. A.; Brunger, A. T. *J. Mol. Biol.* **2000**, *301*, 1237−1256.

(5) Seaton, B. A.; Head, J. F.; Engelman, D. M.; Richards, F. M. *Biochemistry* **1985**, *24*, 6740−6743.

(6) Heidorn, D. B.; Trewhella, J. *Biochemistry* **1988**, *27*, 909−915.

(7) Matsushima, N.; Izumi, Y.; Matsuo, T.; Yoshino, H.; Ueki, T.; Miyake, Y. *J. Biochem.* **1989**, *105*, 883−887.

(8) Klee, C. B. *Interaction of Calmodulin with Ca^{2+} and Target Proteins*; Elsevier: Amsterdam, 1988; pp 35−46.

(9) Ikura, M.; Clore, G. M.; Gronenborn, A. M.; Zhu, G.; Klee, C. B.; Bax, A. *Science* **1992**, *256*, 632−638.

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

(11) Tjandra, N.; Kuboniwa, H.; Ren, H.; Bax, A. *Eur. J. Biochem.* **1995**, *230*, 1014−1024.

10.1021/ja0041876 CCC: $20.00
Published on Web 03/31/2001

the lack of sufficiently extensive data. These include assuming that all residues in a given module, independent of their orientation, have the same overall (isotropic) correlation time. In addition, S_f^2 for fast internal motions was fixed at a value of 0.85, and the corresponding fast effective correlation time was set equal to zero. In this report, we analyze ^{15}N relaxation data for Ca^{2+}-ligated (Holo) CaM without invoking any of the preceding assumptions. It will be shown that to accomplish this, it was necessary to acquire relaxation data over a larger range of field strengths. In the present case of Ca^{2+}-ligated CaM, one must simultaneously least-squares fit data acquired at a minimum of two fields (600 and 800 MHz) to obtain a unique fit of the data without any a priori assumptions regarding values of the extended model-free parameters. Any approximations such as zero fast effective correlation time, an underestimated ^{15}N CSA value, and isotropic rotational diffusion that were made in the previous study affect the field dependence of relaxation data and must be avoided. The availability of relaxation data from more than two fields allows us to assess the validity of using the extended model-free approach to describe slow interdomain motion in macromolecules which was not previously possible. A version of the extended model-free spectral density modified to account for axially symmetric overall rotational diffusion, that is more appropriate for CaM, is utilized in this study. Furthermore, the present work offers a more systematic analysis of slow interdomain motion and, therefore, suggests a better methodological approach for the relaxation data analysis of other macromolecules possessing such motions.

Experimental Section

Expression and purification of *Xenopus* calmodulin was previously described.[9,11] The 220 μL NMR sample contained 1.6 mM calmodulin (uniformly ^{15}N-labeled), 8.0 mM $CaCl_2$, 100 mM KCl, 100 μM NaN_3, and 5% D_2O at pH* of 6.3. All experiments were performed at 35 °C on Bruker AMX360, AMX600, or DRX800 spectrometers. The DRX800 spectrometer was equipped with a shielded *x*,*y*,*z*-pulsed field gradient triple resonance 5 mm probehead. The AMX600 and AMX360 instruments were equipped with shielded *z*-pulsed field gradient triple resonance 5 mm probeheads. States-TPPI quadrature detection in t_1 [12] was used for all experiments. Acquisition times of 111, 96, and 62 ms for t_1 and 59, 83, 69 ms for t_2 were used in all 360, 600, and 800 MHz experiments, respectively. All data matrices were zero-filled to 512 × 2048 points and processed using a 60° squared sine-bell filter in both dimensions. Spectra were processed with NMRPipe[13] and analyzed with PIPP[14] software. ^{15}N and ^{1}H carrier frequencies were always set to 116.5 ppm and water frequency, respectively. Except for the 800 MHz T_1 data set, all experiments were performed once.

Previously described ^{15}N T_1 and $T_{1\rho}$ pulse schemes[15,16] were modified to include the Watergate scheme,[17] pulsed field gradients,[18] and a semiconstant time evolution period in t_1.[19] ^{15}N continuous spin-locking at 2.5 kHz was used for all $T_{1\rho}$ experiments. The T_1 experiments utilized the following relaxation delays: 8, 40, 88, 144, 208, 288, 408, and 504 ms for the 360 MHz sets; 8, 96, 200, 336, 488, 688, 840, and 1040 ms for the 600 MHz sets; 8, 232, 360, 600, 800, 900, 1144, and 1360 ms for the 800 MHz sets. The following delays were used for the $T_{1\rho}$ measurements: 8.2, 16.4, 32.8, 49.2, 65.6, 82.0, 106.6, and 139.4 ms for the 360 MHz set; 8.1, 16.2, 32.6, 48.9, 65.2, 82.5, 106.0, and 130.4 ms for the 600 MHz set; 8.2, 20.2, 35.1, 48.0, 68.2, 95.1, and 120.5 ms for the 800 MHz sets. To minimize the effects of sample heating or changes in NMR conditions, $T_{1\rho}$ data were collected in an interleaved fashion. Recycle delays of 1.0, 1.0, and 1.3 s and 96, 24, and 16 scans per t_1 point were used for acquisition of the 360, 600, and 800 MHz T_1 data sets, respectively. Recycle delays of 1.0, 1.0, and 1.3 s and 80, 48, and 24 scans per t_1 point were used for the 360, 600, and 800 MHz $T_{1\rho}$ measurements, respectively. Relaxation times were determined by fitting the delay dependent peak intensities to an exponential function using Powell's method of nonlinear optimization. T_2 values were calculated from corresponding $T_{1\rho}$, T_1, and chemical shift values, spin lock field strength, and ^{15}N carrier frequency in a straightforward manner described previously.[20,21]

The water flip-back NOE pulse sequence described by Grzesiek and Bax[22] was used for all ^{15}N−{^{1}H} NOE measurements. NOE values were calculated by taking the ratio of peak intensities from experiments performed with and without ^{1}H presaturation. The proton frequency was shifted off-resonance by ~3 MHz during the "presaturation" period for the unsaturated measurements. Saturated and unsaturated spectra were recorded in an interleaved manner. Sixty-four scans per t_1 point were used for data acquisition in both cases. The pulse train used for ^{1}H saturation utilized 162° pulses separated by 50 ms delays and was applied for a total of 2.2 and 3.8 s in the 600 and 800 MHz experiments, respectively. The large off-resonance shift in proton frequency and the ^{1}H saturation pulse flip-angle and their separation were determined empirically to minimize "Dante" off-resonance type effect on the reference spectrum. This effect depends strongly on the harmonic created by the flip-angle as well as the separation of the saturation pulses. Additional recovery delays of 1.0 and 1.3 s were used for the 600 and 800 MHz measurements, respectively. These recycle times are reasonably long, given ^{1}H T_1 values of ~1−1.5 s. Nevertheless, NOE values were corrected for the effects of incomplete ^{1}H magnetization recovery as previously described.[22]

Theory

^{15}N T_1, T_2, and ^{15}N−{^{1}H} NOE values are related to the spectral density $J(\omega)$, which is determined by the reorientational dynamics of the N−H bond vector, as[23]

$$1/T_1 = d^2[J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)] + c^2J(\omega_N)$$

$$1/T_2 = 0.5d^2[4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)] + (^1/_6)c^2[3J(\omega_N) + 4J(0)]$$

$$\mathrm{NOE} = 1 + (\gamma_H/\gamma_N)d^2[6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)]T_1 \quad (1)$$

where $d^2 = 0.1[(\gamma_N\gamma_H h)/(2\pi r_{NH}^3)]^2$ and $c^2 = (^2/_{15})[\omega_N^2(\sigma_{||} - \sigma_\perp)^2]$, h is Planck's constant, r_{NH} is the internuclear N−H distance (0.102 nm), γ_i is the gyromagnetic ratio of spin i, and $\sigma_{||}$ and $\sigma_\perp$ are components of the axially symmetric ^{15}N chemical shift anisotropy (CSA) tensor. A value of −172 ppm was used for $(\sigma_{||} - \sigma_\perp)$ as recent reports [21,24−26] demonstrate this number to be more appropriate than the conventional value of −160 ppm when the N−H bond length is 0.102 nm.

(12) Marion, D.; Ikura, M.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1989**, *85*, 393−399.
(13) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277−293.
(14) Garrett, D. S.; Powers, R.; Gronenborn, A. M.; Clore, G. M. *J. Magn. Reson.* **1991**, *95*, 214−220.
(15) Kay, L. E.; Torchia, D. A.; Bax, A. *Biochemistry* **1989**, *28*, 8972−8979.
(16) Peng, J. W.; Thanabal, V.; Wagner, G. *J. Magn. Reson.* **1991**, *95*, 421−427.
(17) Piotto, M.; Saudek, V.; Sklenar, V. *J. Biomol. NMR* **1992**, *2*, 661−665.
(18) Bax, A.; Pochapsky, S. S. *J. Magn. Reson.* **1992**, *99*, 638−642.
(19) Grzesiek, S.; Bax, A. *J. Biomol. NMR* **1993**, *3*, 185−204.
(20) Davis, D. G.; Perlman, M. E.; London, R. E. *J. Magn. Reson. B* **1994**, *104*, 266−275.
(21) Tjandra, N.; Wingfield, P.; Stahl, S.; Bax, A. *J. Biomol. NMR* **1996**, *8*, 273−284.
(22) Grzesiek, S.; Bax, A. *J. Am. Chem. Soc.* **1993**, *115*, 12593−12594.
(23) Abragam, A. *The Principles of Nuclear Magnetic Resonance*; Clarendon Press: Oxford, 1961.
(24) Boyd, J.; Redfield, C. *J. Am. Chem. Soc.* **1999**, *121*, 7441−7442.
(25) Kroenke, C. D.; Rance, M.; Palmer, A. G. *J. Am. Chem. Soc.* **1999**, *121*, 10119−10125.
(26) Lee, A. L.; Wand, A. J. *J. Biomol. NMR* **1999**, *13*, 101−112.

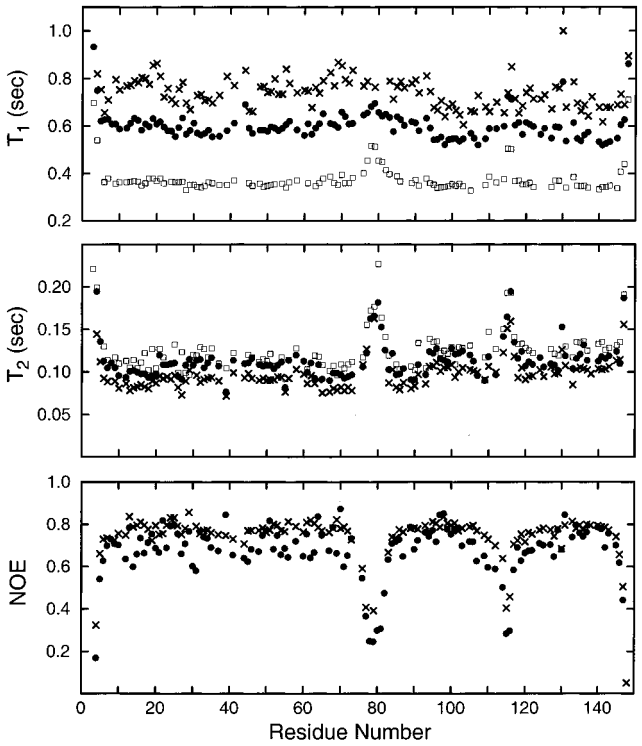

Figure 1. ^{15}N relaxation data for Ca^{2+}-ligated calmodulin measured at 35 °C and at ^{1}H frequencies of 360 MHz (open squares), 600 MHz (closed circles), and 800 MHz (X's).

For an isotropically reorienting macromolecule, the model-free spectral density is[1]

$$J(\omega) = [S^2\tau_c/(1 + (\omega\tau_c)^2) + (1 - S^2)\tau'^2/(1 + (\omega\tau')^2)] \quad (2)$$

where $1/\tau' = 1/\tau_c + 1/\tau$, τ_c is the overall rotational correlation time, and S^2 is the generalized order parameter characterizing the amplitude of internal motions on a time scale τ. If the isotropic overall and internal motions are independent, this spectral density is exact when all internal motions are in the extreme narrowing limit $[(\omega\tau_1)^2, (\omega\tau_2)^2, (\omega\tau_3)^2 ... \ll 1]$ irrespective of the physical nature of these internal motions (this is why the approach was called "model-free"). When the overall motion is anisotropic, the total correlation function (which determines the spectral density) is approximated by the product of the correlation functions of the anisotropic overall and internal motions. This is an approximation even when the internal and overall motions are uncoupled; however, it is a good approximation when the internal motions are axially symmetric and sufficiently fast.

When internal motions of significant amplitude occur on both fast (i.e., extreme narrowing) and slow time scales, there is no rigorous procedure that one can use to interpret NMR relaxation data. The simplest procedure that captures the essential physics is the "extended model-free" approach[10] in which the slow and fast motions have different correlation times (τ_f, τ_s) and order parameters (S_f, S_s). For an isotropically reorienting macromolecule where the internal and overall motions are independent, the "extended model-free" spectral density is

$$J(\omega) = S_f^2 S_s^2\tau_c/[1 + (\omega\tau_c)^2] + S_f^2(1 - S_s^2)\tau'_s/[1 + (\omega\tau'_s)^2] + (1 - S_f^2)\tau'_f/[1 + (\omega\tau'_f)^2] \quad (3)$$

where $1/\tau'_i = 1/\tau_i + 1/\tau_c$ with i = s or f.

This extension was originally proposed to describe backbone dynamics of certain residues in loops that undergo fast librational motions as well as slower motion due to dihedral transitions. In this case, the assumption that overall and internal motions are uncoupled is a good one. When the slow internal motion is the reorientation of an entire domain, such an approximation is certainly not rigorous, but is difficult to avoid.

When the overall motion is anisotropic and the internal motions occur both on the fast and slow time scales, the problem becomes even more complex. The simplest approximation is to multiply the correlation function describing the anisotropic overall motion by the correlation function describing internal motion in the "extended model-free" approach. When the overall motion is axially symmetric,[27] this leads to the spectral density

$$J(\omega) = \sum_{k=1}^{3} A_k\{S_f^2 S_s^2[\tau_k/(1 + (\omega\tau_k)^2)] + S_f^2(1 - S_s^2)[\tau_{s,k}/(1 + (\omega\tau_{s,k})^2)] + (1 - S_f^2)[\tau_{f,k}/(1 + (\omega\tau_{f,k})^2)]\} \quad (4)$$

with:

$$A_1 = 0.75 \sin^4 \alpha, \quad A_2 = 3(\sin^2 \alpha)(\cos^2 \alpha), \quad \text{and } A_3 = (1.5 \cos^2 \alpha - 0.5)^2$$

$$\tau_1 = (4D_{||} + 2D_\perp)^{-1}, \quad \tau_2 = (D_{||} + 5D_\perp)^{-1}, \quad \text{and } \tau_3 = (6D_\perp)^{-1}$$

$$1/\tau_{i,k} = 1/\tau_i + 1/\tau_k \quad \text{where } i = \text{s or f} \quad \text{and } k = 1, 2, 3$$

where α is the angle between the N−H bond vector and the long axis of the rotational diffusion tensor and $D_{||}$ and $D_\perp$ are the parallel and perpendicular components, respectively, of this tensor. Throughout this work, "effective" overall correlation times calculated for "axially symmetric" models are discussed and tabulated. This parameter is expressed as $\tau_{c,eff} = (4D_\perp + 2D_{||})^{-1}$.

Results

Ca^{2+}-ligated *Xenopus* CaM T_1, T_2, and ^{15}N−{^{1}H} NOE values measured at 35 °C and ^{1}H field strengths of 360, 600, and 800 MHz are presented in Figure 1. NOE values were not measured at 360 MHz. The sensitivity of the ^{15}N−{^{1}H} NOE experiment is inherently low, necessitating an unreasonable amount of instrument time at a ^{1}H field strength of 360 MHz. Error estimates were determined by consideration of the signal and root-mean-square (rms) spectral noise intensities except for the 800 MHz T_1 data where the error was taken to be half the average pairwise rms between the two sets of measurements divided by the average T_1 value.[28,29] Errors determined in this manner were 1.3, 1.4, and 2.1% for T_1 data and 5.1, 2.4, and 3.4% for T_2 data at 360, 600, and 800 MHz, respectively. Error estimates for the NOE data were 5.3 and 2.0% at 600 and 800 MHz, respectively.

Residues undergoing large amplitude, long time scale (several hundred ps) internal motions identified by low NOE values will be eliminated from the following analyses.[3] Ca^{2+}-ligated CaM residues with NOE <0.6 for the 600 MHz data or <0.7 for the 800 MHz data were selected for exclusion. Two different NOE cutoffs were required because of the expected field dependence.

(27) Woessner, D. E. *J. Chem. Phys.* **1962**, *36*, 647−654.
(28) Kamath, U.; Shriver, J. W. *J. Biol. Chem* **1989**, *264*, 5586−5592.
(29) Tjandra, N.; Feller, S. E.; Pastor, R. W.; Bax, A. *J. Am. Chem. Soc.* **1995**, *117*, 12562−12566.

Table 1

domain	^{1}H B_0 (MHz)	τ_c (ns)	S^2	τ (ps)	*E*/*N*
N	600	7.0	0.87	67	7.4
N	800	6.7	0.93	71	6.3
N	360 and 600	9.4	0.67	1200	15.6
N	360 and 800	9.1	0.70	930	14.7
N	600 and 800	6.9	0.89	46	9.7
N	360, 600, and 800	9.2	0.73	1300	18.0
C	600	6.5	0.87	50	11.8
C	800	6.0	0.92	59	7.5
C	360 and 600	8.9	0.64	1300	19.6
C	360 and 800	8.4	0.66	980	18.0
C	600 and 800	6.3	0.89	42	12.7
C	360, 600, and 800	8.6	0.70	1400	21.2

[a] Results of least-squares fitting Ca^{2+}-ligated calmodulin NMR relaxation data (Table 1) using the simple model-free spectral density given in eq 2. S^2 and τ were determined globally. Data for some residues were excluded from these calculations on the basis of low ^{15}N−{^{1}H} NOE values or chemical exchange (see text).

Chemical exchange terms were not considered in the following analyses. Consequently, residues undergoing chemical exchange must also be excluded. Residues involved in chemical exchange were identified by consideration of T_1/T_2 ratios[3] or pairwise ratios of T_2 values at different field strengths.[11] A cutoff value of 1.5 times standard deviation from their average ratios were used in both criteria. Residues 12, 19, 27, 28, 29, 39, 55, 63, 99, 101, 107, 109, 118, 120, 127, 129, and 136 were selected for exclusion by this latter criteria. Prolines and residues with weak or overlapped peak intensities are also discounted. Data for the following number of residues remained available for analysis after the preceding considerations: 43 (N) and 29 (C) at 360 MHz, 49 (N) and 34 (C) at 600 MHz, and 48 (N) and 35 (C) at 800 MHz. The N-domain, helical linker, and C-domain comprise residues 1−72, 73−83, and 84−148, respectively. A target error function that is composed of the difference between measured and calculated relaxation rates was used throughout this study.[30] Powell's method of nonlinear optimization was used to determine all parameter values tabulated in this report. A grid search of values was performed prior to the Powell optimizations so that additional possible minima would not be missed.

Simple Analysis

Results from least-squares fits of the data for individual domains at various combinations of field strengths using the simple model-free spectral density given in eq 2 are presented in Table 1. Unique results for fits of 360 MHz data only were not obtained as a consequence of the lack of NOE information at this field strength. The S^2 and τ values shown in Table 1 were determined globally (i.e.. no residue specific determination of these parameters was performed). The results in Table 1 clearly indicate that data from a single field will produce acceptable fits. Even the extracted parameters for internal motions (S^2 and τ) from data at 600 MHz are quite reasonable for a typical rigid protein. The overall rotational correlation time tends to be slightly faster than what one would expect for a protein of this size. In general, however, it is very difficult to obtain a reasonable estimate for the overall correlation time independent from NMR relaxation rates; therefore, data from a single field will not necessarily provide any indication that additional internal motions exist.

Problems in the simple analysis only arise when data from different fields are fit simultaneously. For both domains, the minimum error increases substantially upon global fitting of the data acquired at two different fields. Simultaneously fitting data for all three fields leads to an even larger error, unusually small S^2 values, and exceptionally large τ values. Furthermore, the model-free parameters should not be field-dependent. However, Table 1 indicates a significant field-dependent variation in the values of these parameters.

Korzhnev et. al.[31] found that such behavior is typical of proteins containing global slow internal motion. These authors fitted simulated data generated from the extended Lipari−Szabo spectral density with the original, simple model-free spectral density (eq 2) using an error function dependent on differences between calculated and observed T_1/T_2 ratios and found that the calculated overall correlation time decreases with increasing field strength. They noted that chemical exchange contributions would have the opposite field dependence and is also too large to be accounted for by additional rotational diffusion anisotropy. Fitting the present Ca^{2+}-ligated CaM relaxation data using the standard Lipari−Szabo analysis (eq 2) with $\tau = 0$ and the T_1/T_2-dependent error function also results in calculated overall correlation times that significantly decrease with increasing field strength (results not shown).

When data from both domains and all three field strengths were least-squares fit simultaneously using the model-free spectral density in eq 2, the global minimum occurs when $\tau_c = 9.0$ ns, $S^2 = 0.72$, and $\tau = 1.4$ ns. The fact that τ is so slow immediately indicates that the model-free approach is inapplicable and that some other motion must be present in addition to the usual fast librational motions. Thus, even though the model-free spectral density does a poor job of fitting the data and the resulting parameters are not necessarily meaningful, this kind of simple analysis is useful because it indicates that the NMR data does contain information about slow internal motions.

Another feature of the results shown in Table 1 is that overall correlation times calculated for the C-terminal domain are shorter than those determined for the N-terminal module. This is not unexpected as the raw relaxation data in Figure 1 shows that the average T_1 values measured at 600 and 800 MHz are clearly larger for the N-domain than the C-domain, whereas average T_2 values are smaller for the N-domain than the C-domain at all field strengths. This indication that overall tumbling time of the C-domain is shorter than that of the N-domain suggests the presence of motion on a time scale different from those of the overall tumbling and fast internal librational motions.

Extended Model-Free Analysis

The evidence presented thus far suggests that slow global internal motions are present in Ca^{2+}-ligated CaM. Residues in this central helix have low NOE values (Figure 1), indicating large amplitude, several hundred picosecond time scale "fast" internal motion in this region. Thus, the residues in the central helix linking the two modules are flexible, indicating that this region could serve as a "hinge" for relative motion between the two modules.[3] A ribbon diagram of Ca^{2+}-ligated CaM is shown in Figure 2. The dashed lines indicate proposed motion of the individual modules with cone semi angles found in our subsequent analysis.

As noted in the Introduction, the "extended model-free" approach is not rigorous because the anisotropic overall tumbling motion is coupled to the relative interdomain motions. As a first step, only relaxation data for backbone ^{15}N atoms in the

(30) Dellwo, M. J.; Wand, A. J. *J. Am. Chem. Soc.* **1989**, *111*, 4571−4578.

(31) Korzhnev, D. M.; Orekhov, V. Y.; Arseniev, A. S. *J. Magn. Reson.* **1997**, *127*, 184−191.

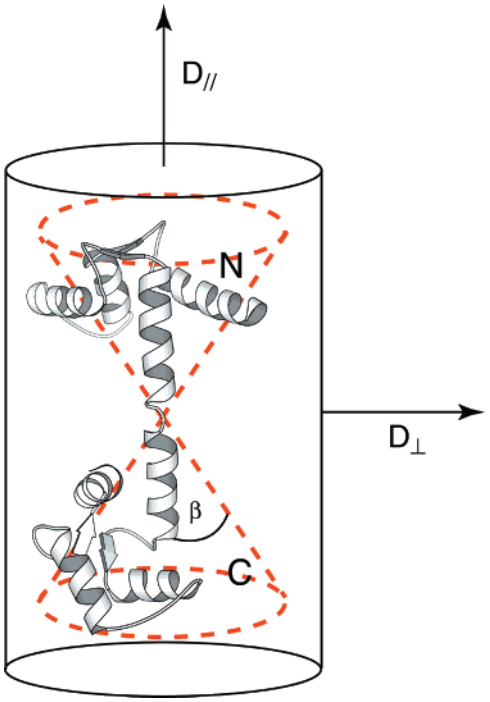

Figure 2. A ribbon diagram of Ca^{2+}-ligated calmodulin generated from the coordinates of Babu et. al.'s structure[34] (PDB accession number 3CLN) using the program Molscript.[36] The dashed lines indicate the proposed slow internal motion. The semi-cone angle β is derived from the internal slow order parameter S_s^2 calculated in the extended model-free analysis. $D_{||}$ and $D_\perp$ represent the components of an axially symmetric diffusion tensor.

rigid regions of the central helix that flank the flexible middle joint were considered. This was done to circumvent the complication due to anisotropic overall motion. The N−H bond vectors associated with these atoms point along the N- and C-terminal central helix axes. It is assumed that these bond vectors undergo fast librational motions described by S_f^2 and τ_f and slow interdomain motions characterized by S_s^2 and τ_s relative to a frame in which the long axis of the diffusion tensor is approximately parallel to both the N- and C-terminal central helices. The relaxation for such atoms is determined only by the motion of the axes of the central helices and not by rotation of the N−H bond vectors about these axes. If the overall and internal motions are assumed to be independent, then the spectral density is described by the "extended model-free" function shown in eq 3 where τ_c is the correlation time for the overall motion of the vector passing through the axes of both the N- and C-terminal central helices (i.e., $\tau_c = {}^1/_{(6D\perp)}$). It is interesting to note that in the limit where S_s^2 goes to zero, which corresponds to the case where the two domains undergo independent and unrestricted motion, eq 3 is still applicable when τ_s is identified as the isotropic overall correlation time of the individual domain. The residues selected for this analysis all have largest T_1/T_2 ratios, suggesting that their bond vectors are indeed approximately parallel to the long axis of the diffusion tensor. Note that one must be cautious when analyzing data with high T_1/T_2 values as chemical exchange may be present.[3] Eight residues of the N-terminal module and 4, 5, and 5 residues of the C-terminal module satisfy these criteria from the 360, 600, and 800 MHz data sets, respectively. The actual residues are listed in the table footnotes. Henceforth, such residues are called "parallel helix residues".

Relaxation data for the central helix residues acquired at three different fields were least-squares fit with the spectral density shown in eq 3 assuming that τ_c is the same for both domains. A global minimum was found for the "extended model-free" parameters presented in Table 2. Incidentally fitting data from 600 and 800 MHz alone for the "parallel helix residues" did not yield a global minimum. It is emphasized that all of the values shown in Table 2 were obtained by a Powell optimization and not forced to assume values consistent with our physical intuition. Thus, it is significant that all parameters derived result in reasonable values. Values of 0.87 and 20 ps for S_f^2 and τ_f, respectively, are typical for fast backbone librational motion in an otherwise rigid protein. If motion in a cone model[32,33] is used to interpret the order parameter characterizing the interdomain motion, the value of 0.7 for S_s^2 implies that the rigid linking helices of both the N- and C-terminal modules can "wobble" independently in cones with a semi-angle of 27°. This is close to the maximum that is sterically allowed. The time scale for this wobbling motion is also physically reasonable. One can estimate the effective correlation time for this motion as follows. If we ignore the friction on the helix and assume that the joint connecting the N and C domains is completely flexible, the wobbling rotational diffusion coefficient is approximately given by $D_W = D_{Trans}/L^2$ where D_{Trans} is the translation diffusion coefficient of a single module ($\sim k_BT/6\pi\eta R$ where k_B is Boltzmann's constant, T is temperature, η is viscosity, and R is the effective radius of the module) and L is the distance from the flexible joint to the center of the domain. D_W is equal to 4.5×10^{-7} s^{-1} for the quantities relevant to the present system; $T = 308$ K, $\eta = 0.719$ cp, $R = 14$ Å, and $L = 24$ Å for the C terminal domain. Using this value for D_W and values of 0.71 and 27° for S_s^2 and the cone semi-angle, respectively, τ_s is estimated to be 1.6 ns from eq A4 of Lipari and Szabo.[1] This estimate is sensitive to the cone semi-angle. An increase of this angle from 27° to 30° results in a change of the estimated τ_s from 1.6 to 1.9 ns. For the N-terminal domain

Table 2

mod.	^{1}H B_0 (MHz)	$\tau_{c,eff}$ (ns)	$D_{||}/D_\perp$	S_f^2	τ_f (ps)	S_s^2	τ_s (ns)	θ (deg)	ϕ (deg)	E/N
				(a)						
N	360, 600, and 800	10.1	–	0.87	21	0.71	3.3	–	–	
C	360, 600, and 800	10.1	–	0.87	20	0.70	2.9	–	–	3.6
				(b)						
N	600 and 800	8.4	1.6	0.88	9	0.74	2.2	66	95	
C	600 and 800	8.4	1.6	0.86	11	0.63	3.0	66	138	4.1
N	360, 600, and 800	8.6	1.6	0.88	10	0.72	2.3	64	97	
C	360, 600, and 800	8.6	1.6	0.86	12	0.62	3.3	66	139	3.8

[a] Results from extended model-free analysis of NMR relaxation data acquired at 35 °C for Ca^{2+}-ligated calmodulin. These values were calculated using a ^{15}N CSA value of −172 ppm and N−H bond length of 1.02 Å. E/N refers to the total error divided by the number of data values used in the respective calculation. θ and ϕ describe the orientation of the long axis of the rotational diffusion tensor relative to the PDB coordinates. All parameter values listed were determined globally. The second column identifies the sets of data being simultaneously fit. (a) Results obtained from fitting "parallel residues" (see text) using the spectral density shown in eq 3. Residues 65, 67−73, 85−87, 89, and 90 for the 600 and 800 MHz data sets and 65, 67−73, 85, 86, 89, and 90 for the 360 MHz data set were deemed "parallel" residues. τ_c was required to be the same for both modules. (b) Results calculated using the "axially symmetric" extended model spectral density (eq 4) and "domain residues" with NOE values above the cutoff and not identified as undergoing chemical exchange. $\tau_{c,eff}$ and $D_{||}/D_\perp$ were required to be the same for both domains.

(32) Lipari, G.; Szabo, A. *Biophys. J.* **1980**, *30*, 489−506.
(33) Kinoshita, K.; Kawato, S.; Ikegami, A. *Biophys. J.* **1977**, *20*, 289−305.

R is 16 Å. This slightly larger R results in an increase in τ_s by 14%. In addition a 20% smaller L value for the N-terminal domain will result in a decrease of 36% in τ_s. Given the crudeness of the model, this estimate is not unreasonable, and it provides a lower limit for the expected slow internal time scale.

We now analyze the relaxation of residues excluding the "parallel helix residues" using the axially symmetric extended model-free spectral density given in eq 4. These residues are designated as "domain residues". We assume that $D_{\|}$ and $D_{\perp}$ (i.e., $\tau_{c,eff}$ and $D_{\|}/D_{\perp}$ in Table 2) are the same for both domains, but we vary the orientation as specified by polar angles θ and ϕ of the long axis of the diffusion tensor relative to the PDB coordinate system in the two domain. Order parameters S_s^2 and S_f^2 and internal correlation times τ_s and τ_f are assumed to be the same for all residues in each domain. This is a good approximation for S_s^2 and τ_s as all residues in a given domain likely experience similar slow internal motion. Since residues with exceptionally low S_f^2 and long τ_f values as indicated by low NOE values were not used in these calculations, the residue specific variation in S_f^2 for the remaining rigid regions of the protein is relatively small (standard deviation of 5% for ubiquitin[29]), justifying the use of global values for S_f^2.

Parameter values derived from fits described above are shown in Table 2. The spectral density (eq 4) used to obtain these values accounts for axially symmetric rotational diffusion of the entire macromolecule. Relative orientations of N−H bond vectors were obtained from the X-ray structure reported by Babu et. al. (PDB accession number 3CLN).[34] The significant difference in ϕ values calculated for the two modules indicates a difference in the average relative orientation of the two domains between the crystal and solution structures. (Recall that ϕ is the azimuthal polar angle that defines the orientation of the long axis of the diffusion tensor relative to the PDB coordinate frame.) The crystal structure samples but one relative orientation of the two domains, whereas the relative interdomain motion in solution samples many possible orientations. No significant difference between individual domain structures of the crystalline and solution samples are anticipated.

The global and fast internal diffusion parameters (S_f^2, τ_f) extracted simultaneously from data collected at 600 and 800 MHz for "domain residues" agree quite well with those derived using the "parallel helix residues" described above. It is also encouraging that the rotational diffusion rate ($D_{\perp}$) calculated from the "parallel helix residues" is remarkably close (within ~2%) to the value calculated for the "domain residues". However, the parameters for the slow internal motion are different in the two cases. In particular, the slow-order parameter for the C-terminal domain decreased from 0.70 to 0.63, while the N-terminal domain stayed unchanged. The differences found for slow internal motion parameters indicate that the two domains have different slow internal motions. A lower internal order parameter for the C-terminal relative to the N-terminal domain suggests that the former is more flexible.

Since fitting two sets of data did yield a unique minimum, one can evaluate the effect on the error and on the parameters of including data from one additional field strength. It can be seen from Table 2 that the addition of data collected at 360 MHz to that acquired at 600 and 800 MHz has no significant effect on the extracted parameter values, as is to be expected if the analysis is meaningful. In addition the fitting error of all three data sets is marginally lower than when one fits the 600 and 800 MHz alone, indicating that the parameters calculated from the high fields agree with the data from the 360 MHz. This again confirms that the minimum found is quite stable and the results are consistent for all magnetic fields.

Discussion

A method for globally estimating extended model-free parameters for a protein possessing significant slow interdomain motion has been presented. For the Ca^{2+}-ligated CaM system, it was found that backbone ^{15}N relaxation data from at least two different fields are necessary to uniquely estimate values for all of these parameters. This is important as previous studies[10,11] employing the "extended model-free" approach used a priori knowledge or assumptions about two of the parameters.

Analyzing the relaxation data of only those residues whose N−H bond vectors are parallel to the long axis of the rotational diffusion tensor ("parallel helix residues") is useful because only relative interdomain "bending" motions are relevant in such an analysis. However, relative interdomain "twisting" motions will affect relaxation rates for the "domain residues". They can lower the value of the slow internal order parameter, S_s^2, determined from an "extended model-free" fit of relaxation data for these residues. That is, one expects S_s^2 values calculated from fits of data for "domain residues" to be less than or equal to those determined from fits of data for "parallel helix residues" as a result of "twisting" motions that occur around the helix axis. Indeed S_s^2 calculated for the C-terminal "domain residues" is smaller than the S_s^2 determined for the corresponding "parallel helix residues". For the N-terminal domain values for these parameters are essentially the same. This indicates that the C-terminal domain is more flexible because in addition to the wobbling motion it also undergoes twisting motion around its central helix axis. However, the variation in slow internal correlation time is much harder to rationalize since its value contains contributions from both bending and twisting motions.

Recent high-resolution X-ray study of Ca^{2+}-ligated CaM showed the presence of correlated disorder of several residues in the crystalline environment.[4] A careful analysis of the disorder revealed that displacement in the central helix was least constrained in the orthogonal direction to the helix axis. This finding is consistent with our work which shows that the helix axis "wobbles". Furthermore, a rigid domain displacement analysis of the above disorder showed highly anisotropic librations with different magnitudes for the N- and C-terminal domains. Even though the magnitudes of these domain displacements are much smaller than what occurs in solution, the qualitative type of domain motions observed in the crystalline state is very similar to our results. A 3 ns molecular dynamics (MD) study of Ca^{2+}-ligated CaM showed large domain reorientation.[35] This domain rearrangement primarily consists of a change from the conformation of CaM found in the crystal structure[34] to an orientation close to that observed in the CaM−peptide complex.[9] This process occurs in the first 1.5 ns of the MD simulation after which the protein seems to stabilize in one conformation with relatively small interdomain fluctuation. The fact that the "wobbling" motions found in this paper were not seen in the simulation is not surprising because the entire length of the simulation was equal to (rather than much greater than) the time scale of these motions.

Slow internal motions on a time scale less than about 1.5 ns significantly reduce NOE values, making it easier to directly

(34) Babu, Y. S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1988**, *204*, 191−204.

(35) Wriggers, W.; Mehler, E.; Pitici, F.; Weinstein, H.; Schulten, K. *Biophys. J.* **1998**, *74*, 1622−1639.

(36) Kraulis, P. J. *J. Appl. Crystallogr.* **1991**, *24*, 945−950.

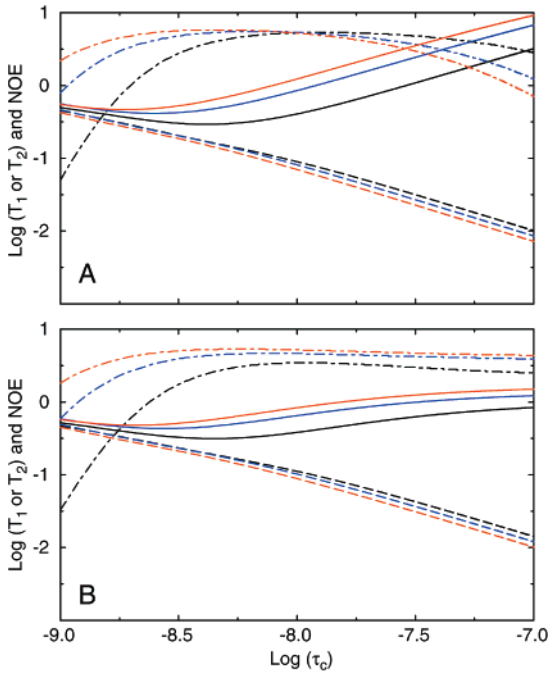

Figure 3. Plots of log(T_1), log(T_2), and ^{15}N−{^{1}H} NOE represented as solid, dash, and dot−dash lines, respectively versus log(τ_c) at ^{1}H frequencies of 360 MHz (black), 600 MHz (blue), and 800 MHz (red). Plotted values were calculated using (A) the simple (eq 1) and (B) the extended (eq 2) model-free spectral densities. $S_f^2 = 0.85$ and $\tau_f = 30$ ps for all curves presented. Values of 0.7 and 3.0 ns for S_s^2 and τ_s, respectively, were used to generate the data presented in (B). A N−H bond length of 1.02 Å and ^{15}N CSA ($\sigma_{||} - \sigma_{\perp}$) of −172 ppm were utilized.

assess the physical situation. The effect of slow internal motion on NOE values decreases with increasing τ_s. Nevertheless, for a system such as Ca^{2+}-ligated CaM with a τ_s of about 3 ns, NOE values still contain important information about this type of motion. It can be shown from model data synthesized from the Ca^{2+}-ligated CaM results that it is primarily the simultaneous fitting of T_1 and NOE values acquired at different field strengths that lead to a unique minimum. Simulated relaxation data calculated at three different field strengths from the simple and the extended model-free spectral densities are presented in Figure 3. The slow internal motion parameter values used to generate the data presented in Figure 3b were taken from the "parallel helix residues" results (Table 2). The dependence of T_1 and NOE values on field strength is different when slow internal motions are present. While T_2 values also depend on the magnitudes of the model parameters, differences between T_2 values derived at different field strengths exhibit less dependence than T_1 or NOE on any given set of reasonable parameters. The change in heteronuclear NOE field dependence in the presence of slow internal motion is particularly significant. The dependence, for large molecules, is an exact opposite of the one observed when no slow internal motion is present. In the case of Ca^{2+}-ligated CaM the average NOE at 600 MHz is 8% lower than at 800 MHz, thus showing a behavior closer to the one presented in Figure 3b than 3a. Similarly, the field dependence for T_1 is less steep when slow internal motion is present. A reduction of as much as a factor of 2 in the slope for T_1 field dependence can be expected for a large protein possessing internal motion of roughly 3 ns. The inclusion of T_2 values from more than one field is nevertheless critical for determination of the overall diffusion parameters as well as to facilitate the identification of residues undergoing chemical exchange. Consequently, to properly describe large interdomain motions in proteins ones should use relaxation data from as many fields as possible.

Analysis of relaxation data from one single field using the simple model-free approach fails to detect the presence of any slow internal motion. The results of such a fit are misleadingly reasonable. The model-free analysis is still useful in detecting slow internal motion when data from more than one field are available. For the system studied here an "extended model-free" analysis on the relaxation data from two fields (600 and 800 MHz) yields slow internal motion parameters that are unique and physically meaningful. Moreover, the addition of data at 360 MHz has no effect on the parameters. This is encouraging and a little surprising, considering the simplicity of the spectral density used to describe such complicated motions. Different orientation of the N−H bonds in the diffusion frame effectively provides probes that are sensitive to different types of slow internal motions. The choice of "parallel helix residues" simplifies the initial analysis of the slow internal motion because they are only sensitive to the "bending" motions. The combination of structural and dynamic information inherent in NMR relaxation data can potentially be utilized to decompose various types of internal motions. So far, no coupling between the overall and slow internal motion has been considered. Even though extended model-free spectral density used in this study accurately describes the data at three fields, it is not clear how such coupling would affect our fitted parameters and their interpretation. Nevertheless this paper represents a significant first step toward quantifying complex interdomain motions in macromolecules.

Acknowledgment. We thank Ad Bax for many useful discussions and for providing us with the relaxation data of Ca^{2+}-ligated calmodulin at 360 and 600 MHz, and Dennis Torchia for critical comments on this manuscript.

JA0041876