Molecular Dynamics-Assisted Optimization of Protein NMR Relaxation Analysis

Janet S. Anderson, Griselda Hernández, and David M. LeMaster*

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: NMR relaxation analysis of the mobile residues in globular proteins is sensitive to the form of the experimentally fitted internal autocorrelation function, which is used to represent that motion. Different order parameter representations can precisely fit the same set of ^{15}N R_1, R_2, and heteronuclear NOE measurements while yielding significantly divergent predictions of the underlying autocorrelation functions, indicating the insufficiency of these experimental relaxation data for assessing which order parameter representation provides the most physically realistic predictions. Molecular dynamics simulations offer an unparalleled capability for discriminating among different order parameter representations to assess which representation can most

accurately model a wide range of physically realistic autocorrelation functions. Six currently utilized AMBER and CHARMM force fields were applied to calculate autocorrelation functions for the backbone H−N bond vectors of ubiquitin as an operational test set. An optimized time constant-constrained triexponential (TCCT) representation was shown to markedly outperform the widely used (S_f^2, τ_s, S^2) extended Lipari−Szabo representation and the more closely related (S_f^2, S_H^2, S_N^2) Larmor frequency-selective representation. Optimization of the TCCT representation at both 600 and 900 MHz ^1H converged to the same parameterization. The higher magnetic field yielded systematically larger deviations in the back-prediction of the autocorrelation functions for the mobile amides, indicating little added benefit from multiple field measurements in analyzing amides that lack slower (∼ms) exchange line-broadening effects. Experimental ^{15}N relaxation data efficiently distinguished among the different force fields with regard to their prediction of ubiquitin backbone conformational dynamics in the ps−ns time frame. While the earlier AMBER 99SB and CHARMM27 force fields underestimate the scale of backbone dynamics, which occur in this time frame, AMBER 14SB provided the most consistent predictions for the well-averaged highly mobile C-terminal residues of ubiquitin.
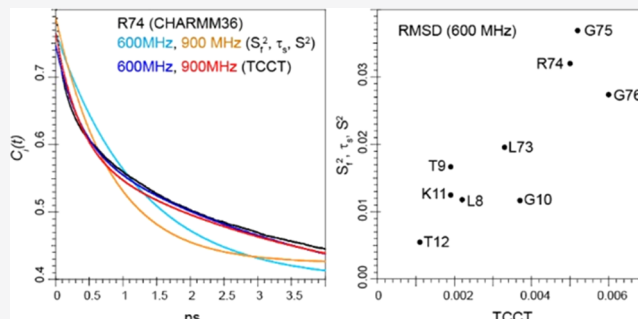
1. INTRODUCTION

Protein backbone dynamics in the ps−ns time frame, as characterized by NMR relaxation measurements, are generally analyzed in terms of the orientational autocorrelation functions for each amide ^1H−^{15}N bond vector and the associated approximately colinear ^{15}N chemical shift anisotropy tensor. In the absence of ∼ms dynamics that give rise to resonance line-broadening effects, the R_1, R_2, and heteronuclear NOE values are specified by the spectral density values at $J(\omega_{1H} + \omega_{15N})$, $J(\omega_{1H})$, $J(\omega_{1H} - \omega_{15N})$, $J(\omega_{15N})$, and $J(0)$, as determined by the Fourier transform of the corresponding bond vector autocorrelation functions. Back-prediction of the underlying physical autocorrelation functions from the experimental data involves the use of simplified mathematical representations of those autocorrelation functions that can be parameterized in terms of the measured NMR relaxation values. With only three experimental relaxation measurements being typically utilized for each amide ^{15}N site in the standard data set, at most, a three-parameter representation can be meaningfully fitted.

In the NMR relaxation analysis of many globular proteins, the assumption of statistically separable internal and global

dynamics appears to be robust.[1,2] When combined with either isotropic or modestly anisotropic global tumbling, the complete autocorrelation functions can be represented as the product of the individual internal autocorrelation functions and either a single global exponential or a set of bond-specific exponentials to model the global tumbling.[3,4] The order parameter-based analysis of such NMR data, originally introduced by Lipari and Szabo,[5] utilizes a single exponential representation of the internal autocorrelation function with the long time limit of the residual bond vector orientational order being represented as S^2, while the time constant τ_e is adjusted so as to match the integral under the modeled autocorrelation function with that of the physical curve. The value of that

integral corresponds to the conjugate spectral density value $J(0)$, which is predominantly determined by the R_2 measurement. To gain additional insight into more mobile segments of the backbone by more realistically modeling of the physical autocorrelation function, the widely used (S_f^2, τ_s, S^2) representation introduced a fast limit order parameter S_f^2, which reflects the degree of bond vector orientational disorder that occurs much more rapidly than the inverse of the highest Larmor frequency monitored.[6]

By construction, the (S_f^2, τ_s, S^2) representation cannot accurately model any autocorrelation function that has at least two significant distinguishable exponential components that occur in the time frame of the Larmor frequencies being monitored by the NMR experiment. On the other hand, given a reasonable modeling of the global tumbling behavior, physically realistic values for the R_1, R_2, and NOE measurements can generally be precisely fitted by adjusting the three parameters of the (S_f^2, τ_s, S^2) representation. In the effort to address the challenge posed by potentially generating a precise fit for the experimental relaxation data that corresponds to an autocorrelation function prediction that significantly deviates from the physical curve, NMR spectroscopists have commonly utilized the analysis of relaxation values measured at different magnetic field strengths in hopes of reducing the resulting ambiguities. Unfortunately, as is illustrated by the example of Arg 74 in the mobile C-terminal tail of ubiquitin, this is not a robust approach. When an internal bond vector autocorrelation function derived from CHARMM36 simulations is used to calculate ^{15}N relaxation values for a physically realistic global tumbling time, application of the (S_f^2, τ_s, S^2) representation at either 600 or 900 MHz ^{1}H field strengths cannot accurately back-predict the original function (Figure 1). Furthermore, not

Figure 1. Internal autocorrelation function for the amide H−N bond vector of Arg 74 in the C-terminal tail of ubiquitin. ^{15}N R_1, R_2, and NOE values were calculated, utilizing the internal autocorrelation function derived from 2 μs CHARMM36 simulations (green) and an isotropic global tumbling time of 4.10 ns. Those derived relaxation values are precisely fitted at 600 MHz (blue) and 900 MHz (red), utilizing the (S_f^2, τ_s, S^2) representation[6] with parameter sets (0.767, 1.252 ns, 0.398) and (0.790, 0.813 ns, 0.423), respectively.

only does the higher field back-prediction deviate more severely from the original MD-derived autocorrelation function but the averaging of the 600 and 900 MHz back-predicted curves also deviates more strongly than does the 600 MHz calculation alone. Furthermore, the two field-dependent back-predictions from the (S_f^2, τ_s, S^2) representation agree considerably more closely with each other than either curve agrees with the original autocorrelation function, thus indicating that the level of internal discrepancy between the two magnetic field-dependent estimates is only a weak indicator of their robustness. Similar observations apply for each of the nine residues of ubiquitin, which exhibit enhanced conformational dynamics on the ps−ns time frame.

In principle, the acquisition of relaxation data sets at multiple magnetic fields can provide the statistical opportunity to utilize order parameter representations with more than three parameters so as to enable a more accurate modeling of the underlying physical autocorrelation functions. However, in this regard, it may be noted that even in the context of a single field relaxation analysis, the most commonly invoked four-parameter unconstrained biexponential $(S_f^2, \tau_f, S_s^2, \tau_s)$ representation can underperform an alternative three-parameter expression that is based upon a simplified version of the time constant-constrained triexponential (TCCT) representation.[7] These results indicate the potential for further optimization of the form of the model autocorrelation function without expanding the number of adjustable parameters.

Molecular dynamics simulations provide the most realistic detailed modeling of such conformational dynamics. This implies that the degree to which a particular order parameter representation can be optimized to accurately fit a wide range of MD-derived internal autocorrelation functions provides the best available measure of its adequacy in representing the physical autocorrelation function. In turn, the degree to which such representations are unable to directly represent the physical internal bond vector autocorrelation function fundamentally limits their ability to indirectly back-predict those internal dynamics from experimental NMR relaxation data.[7,8] In the present study, molecular dynamics simulations applying six currently utilized CHARMM and AMBER-based force fields were carried out on ubiquitin, and the derived internal bond vector autocorrelation functions for nine mobile residues of the protein were used to generate a diverse set of such functional curves. This testbed was then be used to establish an optimized parameterization of the time constant-constrained triexponential representation by the back-calculation of NMR relaxation values that are directly determined from those molecular simulation trajectories and modeled global tumbling dynamics. When this optimized parameterization was then applied to experimental ^{15}N R_1, R_2, and heteronuclear NOE data from ubiquitin, it was apparent that these various AMBER and CHARMM force fields yielded NMR relaxation predictions of markedly differing quality, thus indicating a considerable potential utility for the use of NMR relaxation data in optimizing the conformational kinetics predicted from molecular dynamics force fields.

The order parameter representations considered herein are by no means the only mathematical formulations that have been proposed for representing the internal autocorrelation function for protein dynamics analysis.[9−12] Characterizing the quality of the predictive performance provided by an optimized time constant-constrained triexponential (TCCT) representation should provide a valuable benchmark for analyzing the
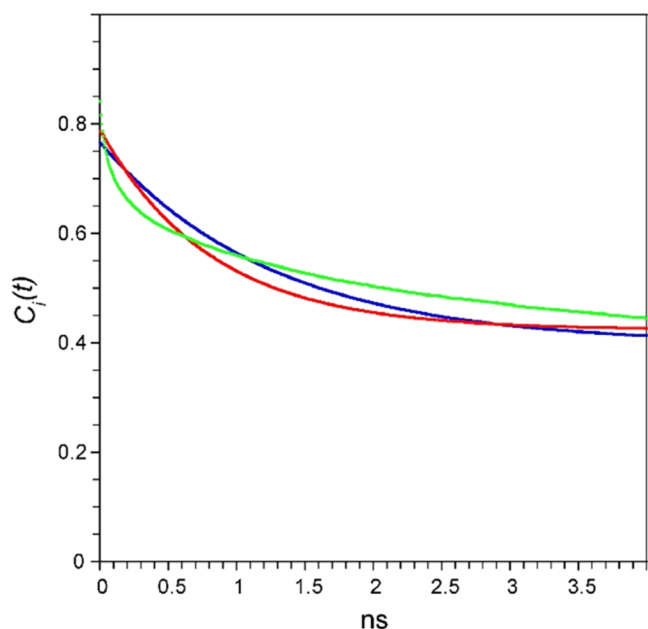
potential benefits that might arise from alternate model representations of the internal autocorrelation function used in NMR dynamics analyses.

2. METHODS

2.1. Molecular Dynamic Simulations. Coordinates from the 1.8 Å resolution X-ray structure (PDB code 1UBQ[13]) were modified by the addition of hydrogen atoms to the protein heavy atoms and crystallographically defined water molecules, and all carboxyl and aliphatic amine groups were modeled in the charged state. The protein coordinates were placed in a rectangular water box with no protein atoms within 10 Å of the boundary. The starting configuration for each trajectory was then generated by an independent placement of 13 sodium ions and 13 chloride ions added to establish electroneutrality and an ionic strength near 150 mM. Particle mesh Ewald summation was applied for long-range electrostatics using a 1 Å grid spacing. Trajectory frames were stored at 5 ps intervals, which was previously reported to yield a satisfactory sampling density for the NMR relaxation calculations.[2] Four 500 ns NVE trajectories were calculated at 298 K using CHARMM22/CMAP (denoted here as CHARMM27)[14,15] and CHARMM36[16,17] force fields, with CHARMM-modified TIP3P water utilizing the NAMD2 program.[18] Analogous simulation conditions were applied to sets of four 500 ns trajectories for the AMBER force fields ff99SB,[19] ff14SB,[20] ff15FB,[21] and ff15ipq,[22] utilizing the AMBER16 program with the exception that the TIP3P-FB water model,[23] rather than the standard TIP3P model, was used for the ff15FB and ff15ipq calculations.

2.2. Bond Vector Autocorrelation and Spectral Density Function Analysis. Following superposition of the trajectory frames upon the initial frame of the production run, the H−N bond vector internal autocorrelation functions and spectral density functions were calculated as previously described.[8] The internal autocorrelation functions were calculated out to 50 ns to help ensure adequate decay of the complete autocorrelation functions for calculations utilizing global rotational correlation times of up to 10 ns. In analyzing the effect of anisotropic global tumbling, the trajectory frames were superimposed upon the diffusion tensor that was derived from the experimental NMR relaxation data so as to determine the relative orientation of each H−N bond vector with respect to the principal axes of diffusion.

While in all cases, the calculation of NMR relaxation values utilizes the full range of the complete autocorrelation function, in statistical analysis of the quality of fit between the order parameter-derived autocorrelation functions and those obtained from the MD trajectories, the standard deviation was calculated starting from 30 ps in recognition of the fact that conformational dynamics that are substantially faster than the [1]H Larmor frequency cannot be adequately characterized by this method.[7]

To assess the influence of experimental error, 500 sets of model relaxation data were generated by adding a Gaussian sampling of experimental uncertainties to the observed R_1, R_2, and NOE values as well as to the bond-specific global correlation times. To define a 95% confidence limit for each residue, the 25 Monte Carlo samplings exhibiting the largest variance in their model R_1, R_2, NOE, and τ_M values were removed, and best-fit order parameters were calculated for the other 475 samplings. The corresponding internal autocorrelation functions were then calculated for these 475 samplings,

and the extremum values for each time point of the autocorrelation functions were determined.[7]

2.3. Expression and Purification of [U−²H,¹⁵N] Ubiquitin. An *E. coli*-optimized codon sequence for ubiquitin (Genscript) was cloned into the expression vector pET11a and transformed into BL21-STAR (DE3) (Invitrogen). Growth was carried out in deuterated minimal media containing 0.25% U−²H glycerol (Cambridge Isotopes, Inc.) at 37 °C until the cells reached an OD_{600} of 0.6 at which time the flasks were quickly chilled to 25 °C and induced overnight with 0.5 mM IPTG (Sigma Chemicals, Inc.). Following a freeze−thaw and lysozyme treatment in 50 mM Tris−acetate pH 8.0, the lysis supernatant was fractionated with 40 and 95% saturated ammonium sulfate precipitations. The final pellet was resuspended into the Tris−acetate buffer and passed through a Sephadex G-50 column equilibrated with the Tris−acetate buffer. After titrating to pH 5.0, the fractions were loaded onto an SP-Sepharose FF column equilibrated with 30 mM Tris−acetate pH 5.0. A salt gradient was run up to 0.4 M NaCl. The pooled fractions were back-exchanged after adding 6% D_2O and sufficient solid sodium carbonate to yield a pH of 10. After incubation for 3 h at 25 °C, the ubiquitin solution was adjusted to pH 5.0 and concentrated by centrifugal ultrafiltration. The sample was then exchanged into 25 mM sodium acetate-d_3 buffer (pH 5.00) containing 6% ^2H_2O and then concentrated to 0.5 mM protein.

2.4. NMR Spectroscopy and Analysis. HSQC-based ¹⁵N R_1, $R_{1\rho}$, and heteronuclear NOE measurements were carried out on ubiquitin at 25 °C, as described,[24,25] utilizing a 600 MHz Bruker Avance III spectrometer and a 900 MHz Bruker Avance spectrometer, both equipped with TXI gradient cryoprobes. The ¹⁵N spin lock field was calibrated for the $R_{1\rho}$ experiments by offset-dependent scalar coupling measurements.[26] The $R_{1\rho}$ experiments were repeated at multiple ¹⁵N offset frequencies (four at 600 MHz ¹H and five at 900 MHz ¹H) to minimize the offset correction required for the corresponding R_2 determination, as previously described.[27] A recycle time of 3 s was used for the R_1 and $R_{1\rho}$ experiments, and a 12 s presaturation pulse was used for the heteronuclear NOE experiments. The complete set of relaxation experiments was repeated for statistical analysis purposes. The median difference between the repeated experiments was used to define the minimal experimental uncertainty, while the resonances exhibiting larger differentials were assigned the correspondingly larger uncertainty. FELIX software (Felix NMR) was used for NMR data processing.

A prolate ellipsoidal diffusion tensor approximation has been found to be appropriate for ubiquitin.[28] Fifty comparatively rigid backbone amides were identified based upon applying an iterative 2.5 σ filter to the individual R_1, R_2, and heteronuclear NOE values. The program Modelfree (version 4.20) was used to derive the corresponding axially symmetric diffusion tensor, having an average global time of 4.10 ns and an axial ratio of 1.22.[29] While the 900 MHz data indicated close consistency in the diffusion tensor predictions when based upon either the three experimental measurements for each of these 50 amides or utilizing only their R_1 and R_2 data, the corresponding analysis of the 600 MHz data was further improved by utilizing a systematic increase in the heteronuclear NOE values of 0.01. The positions of the nitrogen atoms for these 50 residues were then used for the superpositioning of the individual trajectory frames. The effective global tumbling time for each N−H vector was then derived by averaging the instantaneous

rotational diffusion rates for the angle of the bond vector with respect to the principal axis.[30]

3. RESULTS AND DISCUSSION

3.1. Mathematical Characteristics of the (S_f^2,τ_s,S^2) Representation. In their original approach to model-free analysis of conformational dynamics using NMR relaxation data, Lipari and Szabo demonstrated that their (S^2,τ_e) representation could provide an exact estimate of the limit order parameter S^2 by adjusting the time constant for decay of the internal orientational order so as to match the integrals under the experimental and modeled autocorrelation functions as long as that internal motion occurred at rates substantially faster than the highest Larmor frequency sampled by the NMR experiment.[5] Under such circumstances, the S^2 order parameter quantifies the proportion of orientational order that is quenched by the global tumbling process. As noted above, the extended parameterization of the widely used (S_f^2,τ_s,S^2) representation can generally serve to exactly fit all three of the standardly measured relaxation values. Operationally, how is that fitting accomplished? As with the original Lipari−Szabo formulation, parameter optimization serves to match the integrals under the experimental and modeled autocorrelation functions, thus principally tying the prediction of $J(0)$ and S^2 to the R_2 value. In turn, the corresponding matching of the R_1 and NOE values must be predominantly accomplished by a trade-off between the selected values for τ_s and S_f^2.

Any physical or MD-simulated bond vector autocorrelation function can be expected to have some exponential components that are faster than and others that are slower than the single τ_s value that is optimized to fit the curve using the (S_f^2,τ_s,S^2) representation. Thus, as illustrated in Figure 1, the CHARMM36-derived curve initially decreases more rapidly than the order parameter-derived curves. It then crosses each order parameter curve and continues on with a less steep slope. Specifically, the autocorrelation function for Arg 74 crosses the 900 MHz order parameter curve at 0.620 ns and crosses the 600 MHz curve at 1.075 ns. The corresponding analysis of the other eight residues that lie within the mobile loop segment of Leu 8 to Thr 12 and the C-terminal tail segment of Leu 73 to Gly 76 demonstrated similar crossover times for both magnetic fields (Table 1), despite there being a considerable diversity in the dynamics of their bond vector autocorrelation functions. This pattern of similar magnetic field-dependent crossover times for the MD-derived and the

(S_f^2,τ_s,S^2) back-predicted autocorrelation functions was found to also hold for the trajectories derived from the other five force fields examined.

The positioning of these crossover times can be understood by consideration of the discrete Fourier transform that relates the autocorrelation functions to the corresponding spectral density components

$$J(\omega) = \sum 2C(t)\cos(\omega t)\Delta t \tag{1}$$

For the highest sampled frequencies, the incremental time-dependent contributions will be positive when ωt is less than $\pi/2$ and negative when ωt is between $\pi/2$ and $3\pi/2$. Drawing upon the results of the reduced spectral density analysis formalism,[31−33] the optimal effective ^{1}H frequency ω_{1H*} is approximately $0.87 \times \omega_{1H}$, which at a field strength of 600 MHz for ^{1}H yields values for $\pi/2$ and $3\pi/2$ of 0.48 and 1.44 ns, respectively. At the crossover points between the MD-derived and the (S_f^2,τ_s,S^2) back-predicted autocorrelation functions around 1.07 ns, the incremental time-dependent contributions to $J(\omega_{1H*})$ will be near their maximally negative values. In sharp contrast, with a Larmor frequency that is approximately 10-fold lower, the corresponding incremental time-dependent contributions to $J(\omega_{15N})$ will still have a $\cos(\omega t)$ value near 1. Since for the time points immediately preceding the crossover, $C(t)$ is systematically larger for the (S_f^2,τ_s,S^2) back-predicted autocorrelation function than for the MD-derived curve, its incremental time-dependent contribution to $J(\omega_{1H*})$ will be more negative. For time points above the crossover point, those relative contributions are reversed. Due to the positive value of $\cos(\omega t)$ at the ^{15}N frequency, that pattern of relative contributions is reversed for incremental time-dependent contributions to $J(\omega_{15N})$. As a result, the relative intensities of the $J(\omega_{1H*})$ and $J(\omega_{15N})$ values between the MD-derived analysis and the (S_f^2,τ_s,S^2) back-predicted autocorrelation functions can be systematically altered by comparatively modest changes in the crossover point generated by shifts in the S_f^2 and τ_s values of the order parameter analysis.

As the $J(\omega_{1H*})$ and $J(\omega_{15N})$ values provide the dominant contributions to the heteronuclear NOE and R_1 values, respectively, the MD-generated relaxation values can be accurately fitted by shifts in the S_f^2 and τ_s values without accurately matching the MD-derived autocorrelation function. Indeed, in terms of the conformational dynamics that are represented by the instantaneous slope of the autocorrelation function, the sensitivity of this crossover adjustment process is maximized by likewise maximizing the difference in the slope between the two curves at the crossover point (Figure 1). As a result, at the crossover time point for the MD and (S_f^2,τ_s,S^2) back-predicted autocorrelation function, the discrepancy in the representation of the internal dynamics by these two curves reaches a maximum. As indicated in Table 1, the large systematic magnetic field-dependent difference in the crossover point between the original and back-predicted autocorrelation functions is accompanied by a comparably large shift in the predicted time constant for internal motion τ_s. While the back-predictions of the aggregate order parameters S^2 are reasonably consistent for the calculations carried out at 600 and 900 MHz, the τ_s values that were derived at 900 MHz were approximately 30% smaller than those deduced at 600 MHz for each of the nine mobile residues.

3.2. MD Trajectory-Dependent Optimization of the Time Constant-Constrained Triexponential Order Pa-

Table 1. Crossover in CHARMM36-Derived Internal Autocorrelation Functions

	600 MHz (S^2,τ_s,S_f^2)			900 MHz (S^2,τ_s,S_f^2)		
residue	crossover (ns)	τ_s	S^2	crossover (ns)	τ_s	S^2
Leu 8	1.150	1.275	0.743	0.640	0.905	0.753
Thr 9	1.100	1.241	0.598	0.640	0.907	0.611
Gly 10	1.060	1.070	0.562	0.645	0.764	0.570
Lys 11	1.085	1.223	0.611	0.665	0.858	0.620
Thr 12	1.090	1.261	0.774	0.670	0.939	0.779
Leu 73	1.050	1.353	0.520	0.655	0.971	0.537
Arg 74	1.075	1.254	0.398	0.620	0.814	0.423
Gly 75	1.075	1.088	0.268	0.625	0.704	0.294
Gly 76	1.030	0.817	0.151	0.660	0.599	0.163

rameter Representation. Another proposed approach to order parameter analysis was based upon the qualitative anticipation that since the NMR relaxation signals are most acutely sensitive to motion occurring at the nuclear Larmor frequencies, setting the time constants for the two exponential terms of the model autocorrelation function to the inverse of those frequencies should enhance the parameter responsiveness to the dynamics being modeled.[34] When this Larmor frequency-selective order parameter approach is applied to both ^{13}C and ^{15}N relaxation analysis of molecular dynamics-derived autocorrelation functions of protein G, for the great majority of the H−N and H−C bond vectors of the protein, this (S_f^2, S_H^2, S_X^2) representation back-predicted the initial MD-derived autocorrelation functions more accurately than did the widely used (S_f^2, τ_s, S^2) representation.[8] The only notable exception to this superior performance was for mobile C^β positions that were directly bonded to a comparatively immobile C^α so that the effectively single exponential dihedral angle rotational transition dynamics around that $C^\alpha−C^\beta$ bond dominated the predicted relaxation behavior.

From the perspective of analytical utility, the (S_f^2, S_H^2, S_X^2) representation offers the benefit that, to within the accuracy of the reduced spectral density approximation,[31−33] there exists an analytical relationship between the order parameters of the (S_f^2, S_H^2, S_X^2) representation and the corresponding NMR relaxation values for a given global correlation time.[8] The linear functional relationship between the order parameter values, the corresponding spectral density values, and the derived NMR relaxation values ensures a highly regularized one-to-one mapping among these values.[8] Furthermore, for positions exhibiting order parameter values below the rigid limit ($S_i^2 = 1$), the root-mean-square deviation of the experimental uncertainties in the NMR relaxation measurements was found to closely match the resultant rmsd on the (S_f^2, S_H^2, S_X^2) order parameters that were derived from the Monte Carlo analysis of those measurements.[7]

Three basic considerations have hindered the broader utilization of this Larmor frequency-selective order parameter approach as originally formulated. First is the question as to whether optimal sensitivity to dynamics is actually achieved by specifying the time constants of the representation in terms of the experimental Larmor frequencies. Some support for that assumption has been obtained by making use of the wider range of ps−ns dynamics that are exhibited by protein side-chain mobility. The ^{13}C relaxation analysis of the GB3 protein indicated that a systematic alteration of either of the two Larmor frequency-based time constants led to less robust modeling of the corresponding MD-derived internal autocorrelation functions.[8] Recently, a more formal approach has been proposed for defining an optimized set of "detectors" for the analysis of NMR relaxation in which the authors concluded that the previously described (S_f^2, S_H^2, S_N^2) formalism corresponds to a special case version of their analysis approach.[35,36] The second practical impediment was the absence of published analysis of how the Larmor frequency-selective order parameter approach could be effectively applied to multifield NMR relaxation data. As illustrated below, the third limitation arises from the recognition that constraining the lower-frequency time constant to the ^{13}C or ^{15}N Larmor frequency leads to degraded robustness when applied to more slowly tumbling proteins.

To address each of these considerations, the present study carried out sets of four 500 ns NVE simulations of ubiquitin using each of six force fields. The CHARMM27 (CHARMM22/CMAP)[14,15] and CHARMM36[16,17] calculations utilized the NAMD2 program.[18] The AMBER16 program[37] was applied to the ff99SB,[19] ff14SB,[20] ff15FB,[21] and ff15ipq[22] force fields. The initial set of studies examined how well the MD-derived internal autocorrelation functions could be back-calculated from NMR relaxation values that were obtained from these internal autocorrelation functions being multiplied by an exponential decay as would be expected for an isotropically tumbling protein.

Following superposition on the initial frame of the production run, internal autocorrelation functions were calculated for the backbone H−N bond vectors, as illustrated for the CHARMM36 and AMBER 14SB simulations of C-terminal residues Leu 73 to Gly 76 (Figure 2A,B) and the
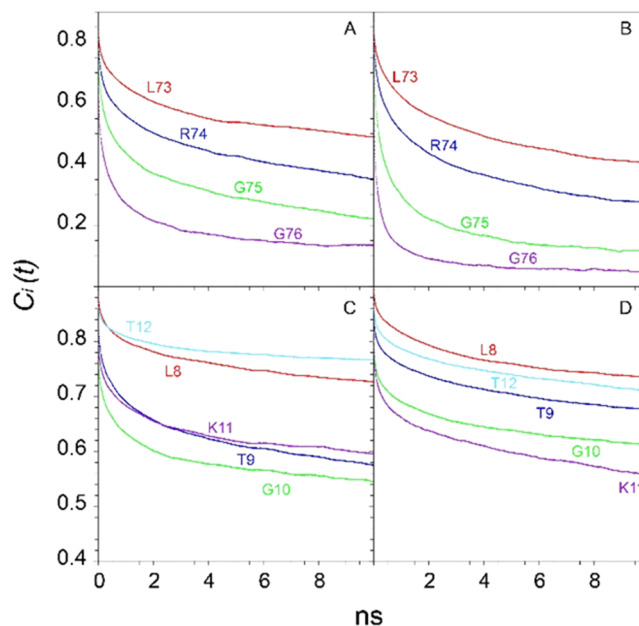

Figure 2. Internal bond vector autocorrelation functions for the C-terminal tail (Leu 73 to Gly 76) and mobile internal loop (Leu 8 to Thr 12) of ubiquitin as calculated from 2 μs of simulation with CHARMM36 (panels A and C) and AMBER 14SB (panels B and D).

other four force fields (Figure S1). For each of the four terminal residues, the AMBER 14SB force field predicted enhanced conformational dynamics, both in amplitude and in apparent rate, relative to CHARMM36. The predictions from the two other recent AMBER force fields more closely resemble those of AMBER 14SB, while in contrast, the earlier CHARMM27 and AMBER 99SB force fields calculated a lower degree of orientational disorder than seen for CHARMM36. A more strikingly disparate pattern of conformational dynamics was calculated for the flexible loop residues Leu 8 to Thr 12 in the CHARMM36 and AMBER 14SB simulations (Figure 2C,D) and the other four force fields examined (Figure S2). While the issue of completeness of conformational sampling will be considered below, these two sets of residues modeled across the six force fields exhibit a substantial range of calculated internal autocorrelation functions. As such, they should provide a useful test of robustness for the order parameter representations being examined.

As originally proposed, the Larmor frequency-selective representation of the internal autocorrelation function is

$$C_i(t) = S_f^2 \times \left[S_h^2 S_x^2 + (1 - S_h^2) \times \exp\left(-\frac{t}{\tau_h}\right) \right.$$

$$\left. + S_h^2 (1 - S_x^2) \times \exp\left(-\frac{t}{\tau_x}\right) \right] \tag{2}$$

where τ_h was set to the inverse of $0.87 \times \omega_{1H}$ and τ_x was set to the inverse of ω_{15N}. To illustrate the robustness of this representation as a function of the assumed global tumbling time τ_M, the internal autocorrelation functions derived from the CHARMM36 simulations for ubiquitin were used to calculate the corresponding R_1, R_2, and heteronuclear NOE values at 600 MHz ^{1}H for τ_M values from 3 to 10 ns. These calculated NMR relaxation values were then used to back-predict the original internal autocorrelation functions.

For tumbling times near the experimental value of 4.10 ns at 25 °C, the rmsd fits to the MD-derived internal autocorrelation functions exhibited rmsd fits near 0.01 or better, well within the level of experimental accuracy typically expected for protein NMR relaxation measurements (Figure 3A). On the
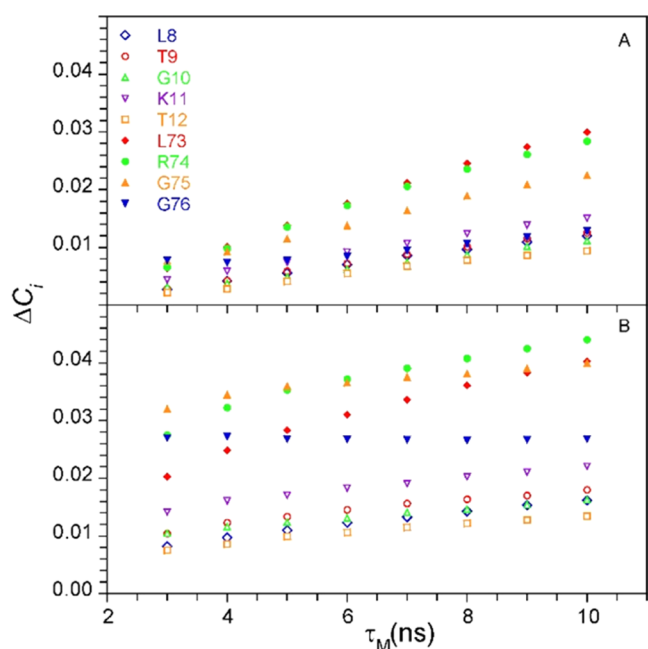

Figure 3. Standard deviations for the back-prediction of the six force field-derived internal autocorrelation functions at 600 MHz ^{1}H for the nine mobile residues of ubiquitin utilizing the (S_f^2, S_H^2, S_N^2) and (S_f^2, S^2, τ_s) representations (panels A and B, respectively), as averaged over the time interval between 30 ps and the assumed global tumbling time τ_M.

other hand, as the global tumbling time increased, the quality of these predictions deteriorated significantly, although they still remained superior to the corresponding predictions obtained from the (S_f^2, τ_s, S^2) representation (Figure 3B). These results clearly demonstrate that if the original Larmor frequency-selective representation is to be usefully generalized, the τ_x parameter of eq 2 must become a function of the experimentally accessible global tumbling time. Encouragingly, the increasingly large discrepancies for the Larmor frequency-based representation at longer tumbling times progress in a near-linear fashion, thus raising the prospect that a similar linear adjustment to the τ_x parameter might prove useful. In

the subsequent analysis, the τ_x parameter was optimized in the form of $a + b \times \tau_M$.

Given the considerable emphasis that has been placed upon the use of magnetic field dependence measurements in such protein NMR relaxation studies, it is warranted to clarify what practical benefits are to be anticipated from such measurements. Field-dependent measurements are clearly central to the characterization of NMR exchange dynamics that occur within the ~ms line-broadening regime. On the other hand, for residues of a protein that are known not to exhibit significant exchange line-broadening, if the relaxation analysis at one field can accurately back-predict the underlying internal autocorrelation function, the corresponding relaxation values at any other magnetic field strength are directly predictable. While close agreement in predictions from different field strengths is clearly necessary for an accurate representation of the underlying internal autocorrelation function, as the (S_f^2, τ_s, S^2) representation analysis of Arg 74 demonstrates (Figure 1), it is not a sufficient condition. This discrepancy regarding the lack of correlation between the magnitude of difference between field-dependent back-predictions and the magnitude of their differences from the underlying autocorrelation function demonstrates the importance of establishing an independent criterion for assessing the robustness with which a given model representation can accurately fit the underlying autocorrelation function.

While the original Larmor frequency-selective order parameter representation would imply the use of magnetic field-dependent time constants when analyzing NMR relaxation data collected at multiple field strengths, the underlying physical internal autocorrelation function is independent of the magnetic field being used to study it. Furthermore, the more precisely a given order parameter representation can provide a direct fit to a given autocorrelation function, the better that same representation can be expected to provide a precise back-prediction of that internal autocorrelation function from the relaxation values, which have been derived from that autocorrelation function.[8] However, as illustrated in the discussion of magnetic field-dependent analyses with the (S_f^2, τ_s, S^2) representation given above, the passage of spectral density values through eq 1 in carrying out the back-prediction of the autocorrelation function introduces a magnetic field-dependent weighting in the sampling of the internal autocorrelation function. To assess the potential benefits of utilizing magnetic field-dependent time constants in an optimized time constant-constrained triexponential representation, initial independent grid searches were carried out at magnetic field strengths corresponding to 600 and 900 MHz ^{1}H.

An additional consideration is whether the fast limit order parameter S_f^2 should be assigned a nonzero time constant. Our previous studies have indicated that, with the use of a fast limit order parameter, robust fits to MD-derived internal autocorrelation functions are generally only obtained for times of 30 ps and longer.[7,8] This earliest time point corresponds to an approximately 10-fold higher frequency than what is directly sampled in the experimental NMR experiment. As preliminary calculations indicated that comparatively little benefit is to be gained from the use of a third nonzero time constant, our analysis turned to the optimization of the τ_h and τ_x ($= a + b \times \tau_M$) factors in eq 2.

The MD-derived internal autocorrelation functions for the nine mobile residues using each of the six force fields were

used for back-prediction from the NMR relaxation values obtained from assuming global tumbling times spanning from 3 to 10 ns, assuming magnetic fields of 600 and 900 MHz ^{1}H. For the resultant 864 back-predictions of the MD-derived internal autocorrelation functions, the rmsd values between the back-predicted and original MD-derived curves were determined over the interval between 30 ps and the assumed global correlation time. The practical utility of that choice for the upper time limit is addressed in the later discussion of experimental uncertainty for bond vector-selective global tumbling times in anisotropic proteins. Following initially coarser grid searches over the three parameters used to specify τ_h and τ_x, the optimization searches for calculations at 600 and 900 MHz were carried out on 3×3 cubes in which the values for τ_h incremented in intervals of 0.01 ns, while the a and b components of τ_x were incremented by 0.1 ns and 0.01, respectively.

Strikingly, at both magnetic fields, the aggregate rmsd value for the 432 fits to each of the MD-derived internal autocorrelation functions for the nine individual mobile residues at each τ_M value reached a minimum at $\tau_h = 0.34$ ns and $\tau_x = 2.00 + 0.43 \times \tau_M$ (Figure 4). As reflected in the
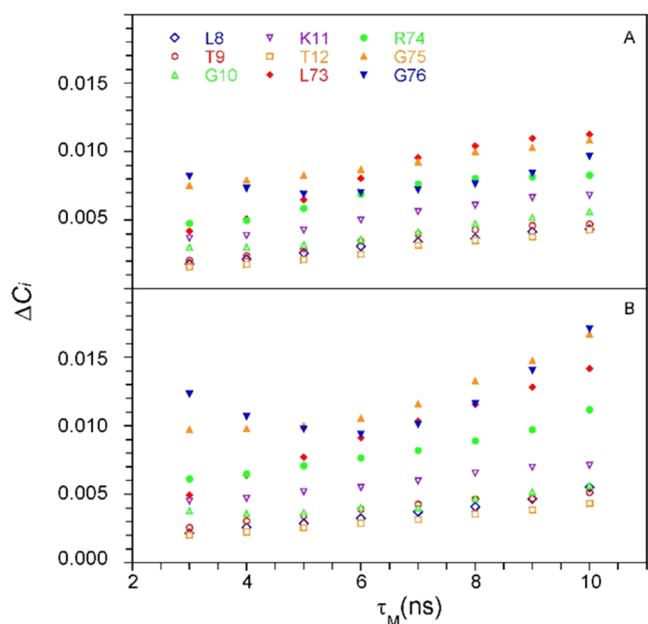

Figure 4. Standard deviations for the back-prediction of the six force field-derived internal autocorrelation functions at 600 MHz (panel A) and 900 MHz (panel B) for the nine mobile ubiquitin residues, utilizing the optimized time constant-constrained triexponential (TCCT) representation (S_f^2, S_h^2, S_x^2) as averaged over the time interval between 30 ps and the assumed global tumbling time τ_M.

expanded ordinate scale, the back-predictions for the nine mobile residues at 600 MHz with the optimized TCCT representation (Figure 4A) are 3- to 4-fold more accurate than those provided by the (S_f^2, τ_s, S^2) representation (Figure 3B). The maximal deviation for the back-predictions among all nine mobile residues for all force fields also reached a minimum near these same time constant values, with that value being consistently near 40% above the largest aggregate rmsd value observed at each τ_M value. To further test the insensitivity of these calculations to the magnetic field strength, analogous calculations that were carried out at 750 MHz for the local

neighborhood of these optimized values of τ_h and τ_x yielded a similar quality minimum with aggregate rmsd values and maximal deviations for the 432 fits to each of the six MD-derived internal autocorrelation functions that were intermediate between those obtained at 600 and 900 MHz.

It may be noted that the optimum value of 0.34 ns obtained for τ_h is approximately 10% above the τ_{H*} value obtained from the reduced spectral density analysis at 600 MHz ^{1}H. In addition, when τ_M is set to 3.4 ns, the optimized value for τ_x is equal to the τ_{15N} value of 2.62 ns at 600 MHz ^{1}H. As that τ_M value equals the global tumbling time of the previously cited GB3 protein, the Larmor frequency-selective relaxation analysis earlier applied to that protein closely matches the optimal parameterization deduced from the present study.[8] In the present study, the lower-case subscript τ_h has been applied to indicate its approximate correspondence to the inverse ^{1}H Larmor frequency. This particular set of time constant parameters has been determined with respect to nine mobile residues of one protein as assessed by six different force fields. The expectation of a broader utility for this parameterization is indicated by the fact that the analogous calculations over the 3×3 grid in which each of the three time parameters was varied by ±10% did not yield any aggregate rmsd values that exceeded the corresponding values of Figure 4 by more than 0.001 for both 600 and 900 MHz calculations. The comparatively shallow minimum suggests this parameterization to be near optimal for a wider range of conformational dynamics.

In considering the different sampling patterns generated by the two magnetic field strengths, particularly over the early time points of the internal autocorrelation functions, the derived S_f^2 values were compared for the optimized TCCT representation at 600 and 900 MHz. For nearly 80% of the total 3456 optimizations, the S_f^2 values obtained at the two field strengths agreed with each other to within ±0.003. Regarding the four more recent force fields considered, for the remainder of the differential S_f^2 values, the 900 MHz-derived values for S_f^2 were systematically larger. In particular, among the C-terminal four residues, the 900 MHz-derived S_f^2 values exceeded those for 600 MHz analysis by an average of nearly 0.03. For these more mobile residues, the back-predicted autocorrelation functions from the 900 MHz analysis more strongly diverged from the original MD-derived curves, particularly at slower tumbling times (Figure 4). Particularly for these more mobile residues, a simple averaging of the order parameters obtained from the two magnetic fields was found to systematically degrade the quality of the predictions relative to the values obtained from the 600 MHz analysis alone.

In applying these optimized τ_h and τ_x time constant values to the full set of 72 backbone amide bond vectors, in nearly all instances, the rmsd fits to the individual sets of three relaxation values during the back-prediction of the internal autocorrelation functions for all six force fields were within 0.0008. Given the cubic spacing of 0.001 for the S_f^2, S_h^2, and S_x^2 parameter grid searches, $\sqrt{3}/2$ times that value corresponds to the maximum anticipated statistical error. As noted above, this close correspondence in the magnitude of the "noise" levels for both the order parameters and relaxation rates reflects the direct linear relationship that exists between the spectral density values, the relaxation rates, and the S_f^2, S_h^2, and S_x^2 order parameters. The only notable exception to these effectively "perfect" fits was observed for the highly mobile Gly 76, particularly for longer τ_M values, where for a subset of the force

fields, the S_f^2, S_h^2, and S_x^2 parameterization were unable to fully match the derived relaxation values.

A comparison was made between the accuracy of the back-predicted internal autocorrelation functions obtained with the (S_f^2, τ_s, S^2) and optimized TCCT representations applied to the 63 less mobile backbone amide bond vectors in ubiquitin, as assessed at 600 MHz ^{1}H. While the magnitude of those deviations was reduced, relative to the nine highly mobile residues, the median ratio of the rmsd values for the (S_f^2, τ_s, S^2) representation relative to those obtained for the optimized TCCT representation was 3.8 for the six force fields and eight τ_M values, quite similar to the results obtained for the more mobile residues (Figures 3B and 4A). To assess the practical significance of these discrepancies in the back-predictions, the fraction of discrepancies in excess of 0.002 and 0.005 was analyzed for each of the six force fields where 0.002 provides an approximate upper bound for the effective statistical noise level and 0.005 serves as a lower-bound estimate for when such a level of discrepancy might begin to perturb the analysis of experimental data. In all cases, the TCCT representation markedly outperforms the (S_f^2, τ_s, S^2) representation (Table 2).

Table 2. Autocorrelation Function Back-Calculation for Less Mobile Residues

	TCCT		StS	
	% > 0.002[a]	% > 0.005	% > 0.002	% > 0.005
CHARMM27	3.8	1.4	30.0	4.4
CHARMM36	4.4	0.6	32.7	8.3
AMBER 99SB	11.7	2.8	74.4	30.2
AMBER 14SB	1.4	0.8	41.7	5.2
AMBER 15FB	13.1	0.6	76.6	31.7
AMBER 15ipq	8.5	1.6	67.3	21.2

[a]rmsd between the original and back-predicted MD-derived curves.

The AMBER 14SB force field exhibits substantially more accurate back-predictions for both representations than does the 99SB force field from which it was derived. In the other instances, the change in performance between earlier and later versions of the two families of force fields is considerably more modest. These results further reinforce the previous conclusion drawn from the ^{13}C relaxation dynamics analysis of the GB3 protein that the comparative benefit of the (S_f^2, τ_s, S^2) representation is largely limited to nearly pure single mode dynamical processes such as those exhibited by a rotatable C^β side chain that is attached to a comparatively rigid C^α position in the protein backbone.[7] For the more complex dynamical behavior exhibited by the protein backbone, the use of the more robust TCCT representation is clearly justified.

Having optimized the values of τ_h and τ_x under the assumption that τ_f is set to zero, the impact of setting τ_f to a nonzero value was examined. From a functional perspective, this implies that the internal autocorrelation function is set to 1.0 at $t = 0$ and then decreases to S_f^2 with an exponential time constant of τ_f. In the simpler analysis considered up to this point, that decrease of the internal autocorrelation function to the value of S_f^2 is assumed to be arbitrarily fast, relative to the NMR Larmor frequencies sampled. In carrying out a (τ_f, τ_h) grid search analysis for τ_h values in the neighborhood of 0.34 ns, the τ_f parameter was sampled at increasing values. When evaluated against the MD-derived autocorrelation functions as described above, the quality of the back-predictions was appreciably degraded even for comparatively small values of τ_f.

This behavior reflected the fact that nonzero values of τ_f give rise to systematically lower optimized values for S_f^2 and systematically higher values for S_h^2. As a substantial fraction of the relatively immobile amides yield S_h^2 predictions near 1.0 when τ_f is set to zero, for many of these positions, the S_h^2 prediction became pegged at 1.0 as the value of τ_f was increased. Indeed, nearly a third of all amides yielded S_h^2 predictions of 1.0 when τ_f was increased to 10 ps. As the value of S_h^2 was not allowed to further increase for these residues, the optimizations for these residues were no longer able to precisely reproduce the input NMR relaxation parameters. Based upon these observations, the technically simpler assumption of two nonzero time constants appears to provide the most robust representation.

3.3. Errors in Back-Prediction of Internal Autocorrelation Functions Arising from Uncertainties in Global Tumbling Dynamics.

NMR relaxation analysis based upon the use of superimposed MD trajectory frames to deduce internal bond vector correlation functions that are then combined with experimentally determined global tumbling dynamics to predict relaxation rates is conceptually redundant. Unfortunately, such a circuitous path of analysis is warranted by the acute sensitivity that the NMR relaxation behavior of globular proteins exhibits toward the global tumbling dynamics, which severely restricts the utility of the comparatively inaccurate molecular dynamics predictions of the global rotational correlation functions. For global tumbling times up to at least 10 ns, the precision of the back-predictions for the MD-derived internal autocorrelation functions described above compared favorably with the noise level that can be expected from good quality experimental NMR relaxation data. Crucial to this quality of back-prediction is the assumption of a precisely known global tumbling effect. Even in the most favorable case of a protein that exhibits essentially isotropic tumbling behavior, there still remains a degree of experimental uncertainty in determining the correct tumbling rate. Considerably more challenging is the common circumstance of anisotropic tumbling in which each H–N bond vector of the protein will be characterized by bond-specific rotational dynamics. For the highly rigid peptide units of a protein that tumbles as a prolate or oblate ellipsoid, projection of the principal axis of global rotation onto the crystal structure of the protein, combined with the anisotropy factor characterizing the relative rotation rates around the major axis and the two minor axes, is sufficient to determine the global tumbling dynamics for each H–N bond vector. On the other hand, for peptide positions that exhibit appreciable internal conformational dynamics, the distribution of the orientation of those bond vectors with respect to the principal axis of the diffusion tensor must be known, and MD simulations provide the most robust approach to assessing that distribution. For each such bond vector in each trajectory frame, the relative orientation with respect to the principal axis predicts the instantaneous rotational diffusion rates, which are then averaged to yield an effective global tumbling time.[30] While the accurate determination of the effective internal orientation of a N–H bond vector is more challenging for mobile residues, the resultant uncertainty in the predicted relaxation behavior is typically decreased by the orientational averaging of the bond vector, which tends to reduce the magnitude of the anisotropy effect. Indeed, molecular dynamics studies have indicated that the magnitude of deviation from the isotropic relaxation behavior roughly scales

with the aggregate order parameter S^2 for that bond.[8] The errors in back-predicting the internal autocorrelation function that arise from errors in the estimated global correlation time typically become more severe for the longer time points of the autocorrelation function curve. This consideration has guided the choice of assessing the back-predicted internal autocorrelation functions based upon rmsd values calculated up to the τ_M value being analyzed.

While the global correlation function for an anisotropically tumbling protein formally requires a multiexponential representation,[30] the approximation of a bond-specific single exponential decay is quite robust.[3,4] Specifically, in the context of ellipsoidal rotational diffusion, for ordinate values of the autocorrelation function between 1.0 and 0.01, the single exponential approximation conforms to the exact expression to within an rmsd of 0.01 for all angles when the axial ratio D_z/D_x lies between 0.4 and 2.1.[7]

If the assumed global tumbling time constant for a given bond vector is in error, it can be readily anticipated that the back-prediction of the internal autocorrelation function will give rise to a compensating error in the order parameter corresponding to the long time constant τ_x. To assess how large of an error is introduced into the back-predicted internal autocorrelation function as a result of an inaccurate bond-specific τ_M value, calculations were made under conditions that roughly approximate the experimental τ_M value for this protein. The NMR relaxation values predicted from the MD-derived internal autocorrelation functions and an isotropic 4.00 ns global correlation time as discussed above were used to back-predict those internal autocorrelation functions, assuming global tumbling times of 4.04 and 3.96 ns. For a τ_M of 4.00 ns, the optimized τ_x value $(2.00 + 0.43 \times \tau_M)$ is 3.72 ns. The distortion in the back-predicted internal autocorrelation function when a 1% error in the time constant for global tumbling is introduced is most easily analyzed for the comparatively rigid amide positions for which the internal motion has largely decayed to a limit S^2 value before the significant onset of decay effects from global tumbling.

As discussed below, from the experimental measurements on ubiquitin, a set of 50 backbone amide positions were identified as being suitably ordered for use in determining the global diffusion tensor. When an isotropic global correlation time of 4.04 ns was used to back-predict the internal correlation functions for this set of residues with NMR relaxation values that were derived assuming a τ_M value of 4.00 ns, the resultant average rmsd error in fitting the original MD-derived internal autocorrelation functions was 0.007, consistent with the original 1% error in the assumed tumbling time being scaled down by approximately the aggregate S^2 values for those residues. On the other hand, when an isotropic global correlation time of 3.96 ns was used to back-predict from these same NMR relaxation values, the resultant average rmsd error in fitting the original MD-derived internal autocorrelation functions was reduced to only 0.004. Similar to the analysis of nonzero τ_f values discussed above, for many residues, the increase in the S_x^2 parameter needed to compensate for the artificially shortened τ_M value had risen to a maximum of 1.00 so that further compensation for the foreshortened τ_M value was blocked. In this circumstance, the inability to fully respond to an erroneous τ_M value serves to limit how far the optimization procedure can deviate from the correct order parameter values. The residues that exhibited reduced errors in the back-predicted internal autocorrelation functions for a τ_M

of 3.96 ns also exhibited appreciable degradation in the quality of their optimal fits to the relaxation values used for the back-predictions. For the rest of the N−H bond vectors, which exhibited sufficient mobility so as not to "peg" the order parameter values in the 3.96 ns analysis, the scaling of the errors in the back-predicted internal autocorrelation functions relative to the error in the global tumbling time varied roughly proportionate to the aggregate S^2 values, as seen under the assumption of τ_M equal to 4.04 ns. Among the residues with limited internal mobility, the inability of the optimized (S_f^2, S_b^2, S_x^2) order parameters to precisely back-predict the NMR relaxation parameters provides a strong indication of a significant error in either those relaxation parameters or in the global tumbling parameters used to analyze them.

3.4. Assessing the Conformational Sampling of the Ubiquitin C-Terminal Tail. The MD-derived internal autocorrelation functions obtained in this study generally exhibit the smooth monotonically decreasing slopes to be expected of a satisfactory statistical sampling of the relevant conformational space. To provide a more rigorous test for the adequate sampling needed to justify comparing these calculations against experimental NMR data, similar autocorrelation functions were derived from four 500 ns AMBER 14SB simulations[38] initiated from the ubiquitin crystal structures 1TBE[38] and 2G45.[39] In each of these crystal structures, the C-terminal tail interacts in a larger protein complex that results in a conformation that differs markedly both from each other and from the 1UBQ crystal structure that was used as the initial model in this study. The resultant internal bond vector autocorrelation functions from the four terminal residues were compared against those predicted using the 1UBQ model (Figure S3). The first half of each production run (250 ns) was excluded from the bond vector correlation function calculations to allow for the initial residual conformational memory of the intermolecular interactions within the crystal form to dissipate. The corresponding exclusion of the first half of the 1UBQ production runs had minimal effect upon the resulting internal autocorrelation functions. The reasonably similar internal autocorrelation functions for each of the four C-terminal residues that were obtained from the three starting protein conformations support the assumption that the simulations of this study provide a usefully extensive sampling of the energetically accessible conformational space for these residues, thus providing justification for their application to the direct prediction of experimentally determined NMR relaxation parameters.

3.5. Analysis of Experimentally Determined Internal Bond Vector Autocorrelation Functions. To test the practical utility of this approach in deriving an experimentally based estimate of the internal bond vector autocorrelation functions for the backbone amides of ubiquitin, R_1, $R_{1\rho}$, and heteronuclear relaxation measurements were carried out on a 0.5 mM ^{2}H,^{15}N labeled sample using optimized pulse sequences.[24,25] The pH of the sample was adjusted to 5.00 to minimize the self-aggregation effects that have been observed for ubiquitin at neutral pH.[40,41] An iterative 2.5 σ filter was applied to each of the ^{15}N R_1, R_2, and NOE sets of measurements on ubiquitin that identified 50 residues exhibiting limited internal dynamics. The relaxation data for these residues were fitted to a prolate global diffusion tensor having an isotropic τ_M of 4.10 ns and an axial asymmetry of 1.22 at 25 °C.[29] After aligning the principal axis of experimentally determined global diffusion tensor along the

Z-axis, the trajectory frames for each of the molecular dynamics simulations were superimposed, utilizing the nitrogen atoms of the 50 relatively immobile residues. The bond vector angles with respect to the Z-axis were then used to predict effective global tumbling times for each bond vector.

Initial estimates of the S_f^2, S_h^2, and S_x^2 values that are consistent with the experimental NMR relaxation values were directly calculated by applying the reduced spectral density approximation.[8] Localized grid search optimizations were then carried out utilizing the complete relaxation equations. As anticipated, these experimental data identified the residue segments 8−12 and the C-terminal tail as the principal regions of enhanced backbone conformational dynamics here indicated by a color coding of the autocorrelation function as assessed at the τ_M value of 4.10 ns (Figure 5).
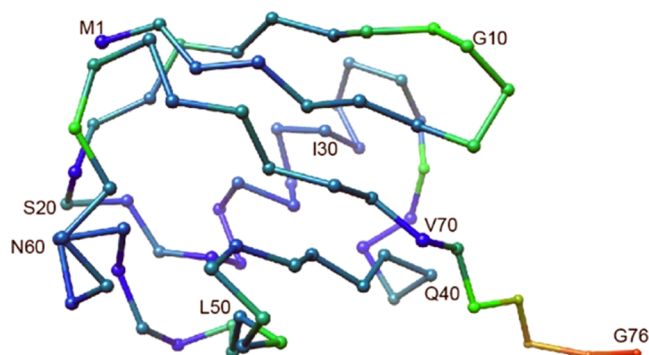

Figure 5. Experimentally determined values of the internal bond vector autocorrelation functions of the ubiquitin backbone at 4.10 ns are displayed as rainbow coloring, with dark blue set to 1.0, green set to 0.69, and red set to 0.0. Residues lacking experimental measurements are set to dark blue.

Consistent with the results obtained from the MD-derived autocorrelation functions, for each of the last four residues of the protein, the fast limit order parameter S_f^2 was found to be higher for the experimental NMR relaxation data at 900 MHz than for the experimental data at 600 MHz. Over the interval between 30 ps and 4.10 ns, the rmsd between the bond vector autocorrelation functions that were back-predicted from the data at the two different field strengths were 0.011, 0.014, 0.021, and 0.021 for Leu 73 to Gly 76 (Figure 6). To further assess the consistency between the internal autocorrelation functions that were back-predicted from experimental data collected at the two magnetic fields as well as for the corresponding results predicted by the various force field trajectories, Monte Carlo simulations were carried out for each amide N−H bond vector in which not only the experimental ^{15}N R_1, R_2, and NOE values but the predicted bond-specific global correlation times were varied. To facilitate qualitative comparisons between residues exhibiting significantly different dynamical behavior, the Gaussian-distributed perturbations on the ^{15}N R_1, R_2, NOE, and τ_M values were scaled by the physically realistic levels of 1%, 1%, 0.01, and 1%, respectively. The 95% confidence limits for each autocorrelation function were obtained from a set of 475 Monte Carlo simulations drawn from 500 random samplings of the four variable parameters from which the 25 samplings with the largest rmsd values for those parameters were removed. The extrema of those 475 curves at each time point of the autocorrelation function were then selected. Given the direct linear relation-
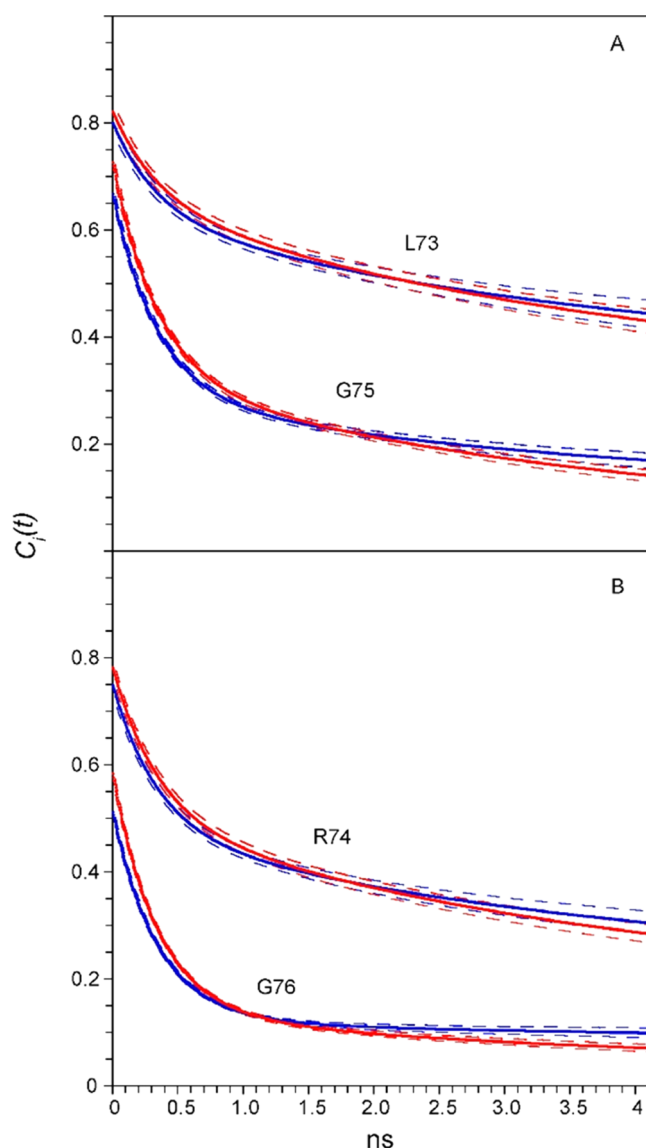

Figure 6. Experimentally determined internal bond vector autocorrelation functions for the C-terminal tail of a ubiquitin sample having an average τ_M of 4.10 ns and an ellipsoidal asymmetry of 1.22 utilizing the optimized TCCT representation, with data from 600 MHz indicated in blue and the 900 MHz data indicated in red. The Monte Carlo-derived 95% confidence limits (dashes) were determined, assuming uncertainties of 1%, 1%, 0.01, and 1% for the R_1, R_2, NOE, and bond-specific τ_M values, respectively.

ship between the spectral density functions and the order parameters in the TCCT representation, the bounds of these confidence limits will change proportionate to a change in the experimental uncertainties.[7,8]

For Leu 73 and Arg 74, the 95% confidence limits for the experimental data collected at the two magnetic field strengths are overlapping across essentially the entire portion of the autocorrelation function being analyzed (Figure 6). This indicates that the effects of the actual experimental errors and the systematic differences arising from back-prediction at the two different magnetic field strengths lie within the assumed experimental uncertainties. The 95% confidence limits for Gly 75 and Gly 76 at the two magnetic field strengths modestly diverge from each other, in part, due to the larger field strength-dependent divergence of the back-

predictions for these two more mobile residues, as discussed above. As the residual orientational order nears zero, the relative bandwidth of the percentage-based 95% confidence limits decreases as well.

Returning to the initial illustration of the magnetic field-dependent back-prediction of the CHARMM36-derived internal autocorrelation function for Arg 74 utilizing the (S_f^2, τ_s, S^2) representation (Figure 1), the analogous back-predictions utilizing the experimental 600 and 900 MHz NMR relaxation data yielded internal autocorrelation functions, which exhibited an rmsd of 0.017 over the interval 30 ps to 4.10 ns (Figure 7). This field-dependent difference is only
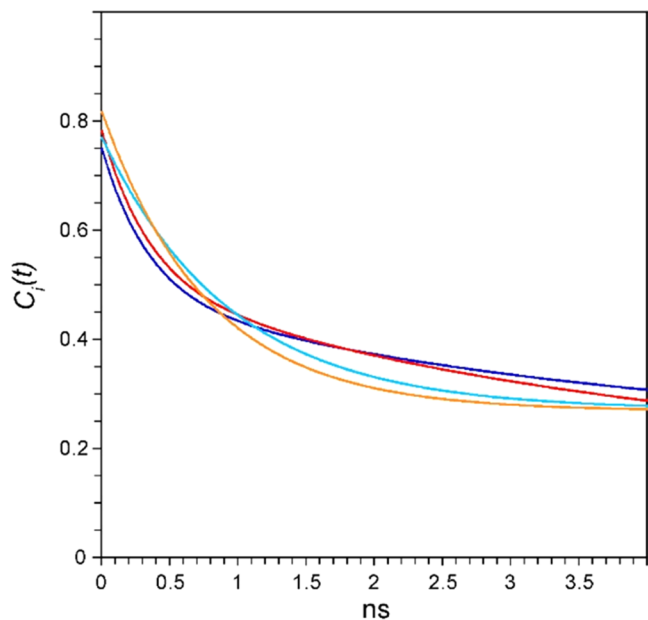

Figure 7. Experimentally determined internal bond vector autocorrelation function for the amide H−N bond vector of Arg 74 utilizing the optimized TCCT representation, with 600 MHz data indicated in blue and the 900 MHz data indicated in red as well as the (S_f^2, τ_s, S^2) representation with 600 MHz data indicated in cyan and the 900 MHz data indicated in orange.

modestly larger than the rmsd of 0.014 obtained between the two corresponding back-predictions utilizing the optimized TCCT representation. On the other hand, back-prediction from the experimental 600 MHz data utilizing these two different order parameter representations yields an rmsd of 0.039, while the analogous comparison for the 900 MHz data yields an rmsd of 0.042. As the differences in the predictions generated by the two different order parameter presentations are nearly three times as large as the field strength-dependent differences obtained for each representation, these results affirm the conclusion drawn from the MD-derived autocorrelation analysis that the degree of similarity between the back-predictions derived from magnetic field-dependent measurements does not necessarily provide a useful measure of their accuracy.

The internal consistency in the analysis of the MD-derived internal bond vector autocorrelation functions with the optimized TCCT representation provides a firm basis for confidence in the capability of that representation to extract usefully reliable estimations of the physical autocorrelation functions underlying experimental protein NMR relaxation data. Given the markedly varying predictions of these

autocorrelation functions provided by the six force fields examined here (Figures 2, S1, and S2), it is appropriate to consider how useful this analysis of experimental data may prove to be in characterizing the quality of the dynamics predictions provided by such force fields. To further pursue such comparisons, it is necessary to examine in more detail how the bond vector orientation information represented in the molecular dynamics simulation is quantitatively correlated with the observed NMR relaxation data. The two crucial parameters to be specified for such backbone amide bond vectors are the ^{15}N chemical shift anisotropy (CSA) and the effective N−H bond length. The smaller of the two principal contributions to ^{15}N NMR relaxation arises from the asymmetrical electron distribution surrounding the ^{15}N nucleus. Although evidence for residue-specific ^{15}N CSA values has been reported,[42−44] the approximation of a uniform value appears to be reasonably robust in conventional protein NMR relaxation studies. Given the $1/r^6$ dependence of the dipole−dipole interaction between the ^{1}H and ^{15}N nuclei, a highly precise estimate of the bond length is required. This estimate is complicated by the fact that radial fluctuations of the ^{1}H and ^{15}N nuclei serve to shorten that effective distance, while angular fluctuations cause the effective bond length to increase. As the covalent bond preferentially suppresses the magnitude of the radial fluctuations, the effective bond length for dipolar relaxation effects is lengthened relative to crystallographically determined values.[45] Further complicating the quantitative analysis of the relaxation interactions is the fact that classical molecular dynamics simulations disregard quantum mechanical dynamics and suppress the high-frequency classical vibrational/librational motions by freezing the H−X bond lengths during the calculations.[46]

Each of the order parameter representations analyzed in this study makes use of a fast limit order parameter S_f^2, which directly scales the corresponding spectral density values. Since a change in the assumed effective bond length can be precisely compensated for by a corresponding shift in the S_f^2 value, in practical terms, the bond length parameter is optimized against the order parameter values that are deduced from the analysis of molecular dynamics trajectories of the comparatively immobile amides of the protein structure. The conceptual utility of the effective bond length parameter is somewhat compromised by the fact that different force fields predict different average values for the degree of orientational disorder of the amide bond vectors over the dynamical range that is used to deduce the order parameters for the comparatively immobilized residues. As shown in Table 3, the average value of the internal bond vector autocorrelation functions at $t = 30$ ps for the 50 motionally restricted residues used to analyze the global tumbling dynamics varies by a factor of 4.0%, while at t

Table 3. Average Internal Autocorrelation Function Values[a]

	AC_i (30 ps)	AC_i (4.10 ns)
600 MHz exp (1.030 Å)	0.892	0.880
CHARMM27	0.895	0.881
CHARMM36	0.891	0.873
AMBER 99SB	0.874	0.844
AMBER 14SB	0.886	0.871
AMBER 15FB	0.860	0.828
AMBER 15ipq	0.874	0.849

[a]50 motionally restricted residues.

= 4.10 ns, that variation among these six force fields increases to 6.0%. As this range of dynamics is far slower than the simulated motions that might be perturbed by the freezing of H−X bond lengths during the molecular dynamics calculations, scaling of the predicted NMR relaxation values based upon a force field-specific effective bond length can potentially be misleading. In the present analysis, the value of 1.030 Å for the effective bond length has been selected as it yielded similar average predictions for these motionally restricted amides, as those obtained with the two CHARMM force fields and the widely used AMBER 14SB force field.

Assuming this value of 1.030 Å for the effective bond length, the internal autocorrelation functions derived from the six force fields were compared to the corresponding autocorrelation functions back-predicted from the ^{15}N relaxation data obtained at 600 MHz (Table 4), reflecting the more robust

Table 4. Fit to Back-Predicted Internal Autocorrelation Functions[a]

	Leu 73	Arg 74	Gly 75	Gly 76
CHARMM27	0.076	0.066	0.140	0.172
CHARMM36	0.087	0.124	0.148	0.103
AMBER 99SB	0.178	0.280	0.277	0.228
AMBER 14SB	0.045	0.062	0.015	0.028
AMBER 15FB	0.106	0.083	0.021	0.069
AMBER 15ipq	0.076	0.074	0.010	0.015

[a]rmsd between MD-derived and back-prediction from 600 MHz NMR data.

performance that is obtained from lower magnetic field strength data for these mobile amide positions. The AMBER 14SB force field yielded the closest predictions for both Leu 73 and Arg 74, while its reasonably precise predictions for Gly 75 and Gly 76 were modestly bested by the AMBER 15ipq force field. In stark contrast, both CHARMM force fields and the earlier AMBER 99SB force field predict a considerably lower level of conformational dynamics for the C-terminal tail. For all four residues, the AMBER 15FB force field yielded autocorrelation function predictions that were of a quality intermediate between the comparatively accurate predictions from the AMBER 14SB and 15ipq force fields and the overly rigid predictions from the three other force fields. Quantitative estimation of the statistical significance of the rmsd values of Table 4 would require a precise characterization of the statistical robustness of the various force field simulations. That said, the marked variations among the autocorrelation functions derived from these force fields and the apparent ability to back-predict the physical autocorrelation functions to a precision comparable to the statistical uncertainty of the experimental NMR relaxation measurements indicate considerable potential for utilizing such measurements as a means of quantitatively assessing the accuracy with which specific force fields can model the dynamics of protein backbone motion.

4. CONCLUSIONS

Modeling analysis based upon protein molecular dynamics simulations provides a powerful means for optimizing the mathematical representation of the internal bond vector autocorrelation functions used to interpret ^{15}N NMR relaxation data for backbone amide positions. Optimization of the time constant-constrained triexponential (TCCT) representation was obtained at both 600 and 900 MHz ^1H

when the shortest time constant τ_f is set to zero, the τ_h time constant is set to 0.34 ns, and the global tumbling time-dependent τ_x time constant is set to 2.00 ns + 0.43 × τ_M. As averaged across six CHARMM and AMBER molecular force fields at 600 MHz ^1H, the MD-derived autocorrelation functions can be back-predicted from the corresponding R_1, R_2, and heteronuclear NOE NMR relaxation values to within 1.0% over the global tumbling times of 3−7 ns for each of the nine mobile residues of ubiquitin. The modest deterioration of those back-predictions at slower tumbling rates suggests the potential for useful applicability for this order parameter representation at τ_M values even beyond 10 ns. The optimized TCCT representation is somewhat less robust for the NMR analysis of relaxation data collected at 900 MHz ^1H. With an optimal performance for τ_M values near 5−6 ns, the increasingly severe systematic deviations in the back-predicted autocorrelation functions that occur at slower tumbling times utilizing these higher field data will likely limit the practical value of such data. Any fractional averaging of the order parameters obtained at 600 and 900 MHz ^1H resulted in a reduced accuracy as compared to the predictions obtained at the lower field. That observation illustrates the fact that the practical benefit derived from NMR relaxation measurements obtained at multiple magnetic fields for analyzing protein motion in the ps−ns time frame is dependent upon the robustness with which the mathematical representation of the autocorrelation function is capable of modeling the underlying dynamics at each of those magnetic field strengths. As the differences among the internal autocorrelation functions derived from the various force field calculations far exceed the estimated uncertainties associated with the experimentally derived autocorrelation functions, the described approach to NMR relaxation analyses appears to offer considerable promise for helping to guide the further optimization of dynamical modeling in the time frame surrounding the ^1H and ^{15}N Larmor frequencies.

■ ASSOCIATED CONTENT

ⓈI Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.1c01165.

Force field-derived autocorrelation functions for Leu 73 to Gly 76 of ubiquitin (Figure S1); force field-derived autocorrelation functions for Leu 8 to Thr 12 of ubiquitin (Figure S2); and starting structure-dependent autocorrelation functions for ubiquitin C-terminal residues (Figure S3) (PDF)

(PDF)

■ AUTHOR INFORMATION

Corresponding Author

David M. LeMaster − *Wadsworth Center, New York State Department of Health, Albany, New York 12201, United States;* ⓘ orcid.org/0000-0003-2810-9465; Phone: (518) 474-6396; Email: david.lemaster@health.ny.gov; Fax: (518)473-2895

Authors

Janet S. Anderson − *Department of Chemistry, Union College, Schenectady, New York 12308, United States*

Griselda Hernández − *Wadsworth Center, New York State Department of Health, Albany, New York 12201, United States;* ⊙ orcid.org/0000-0002-7557-8806

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jctc.1c01165

Notes
The authors declare no competing financial interest.
NMR relaxation values are available in the BioMagResBank database under the accession number 51174.

■ REFERENCES

(1) Wong, V.; Case, D. A. Evaluating Rotational Diffusion from Protein MD Simulations. *J. Phys. Chem. B* **2008**, *112*, 6013−6024.

(2) Anderson, J. S.; LeMaster, D. M. Rotational Velocity Rescaling of Molecular Dynamics Trajectories for Direct Prediction of Protein NMR Relaxation. *Biophys. Chem.* **2012**, *168−169*, 28−39.

(3) Brüschweiler, R.; Liao, X.; Wright, P. E. Long-Range Motional Restrictions in a Multidomain Zinc-Finger Protein from Anisotropic Tumbling. *Science* **1995**, *268*, 886−889.

(4) Lee, L. K.; Rance, M.; Chazin, W. J.; Palmer, A. G. I. Rotational Diffusion Anisotropy of Proteins from Simultaneous Analysis of ^{15}N and ^{13}C Nuclear Spin Relaxation. *J. Biomol. NMR* **1997**, *9*, 287−298.

(5) Lipari, G.; Szabo, A. Model-Free Approach to the Interpretation of Nuclear Magnetic Resonance Relaxation in Macromolecules. 1. Theory and Range of Validity. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4559.

(6) Clore, G. M.; Szabo, A.; Bax, A.; Kay, L. E.; Driscoll, P. C.; Gronenborn, A. M. Deviations from the Simple 2-Parameter Model-free Approach to the Interpretation of N-15 Nuclear Relaxation of Proteins. *J. Am. Chem. Soc.* **1990**, *112*, 4989−4991.

(7) Anderson, J. S.; Hernández, G.; LeMaster, D. M. ^{13}C NMR Relaxation Analysis of Protein GB3 for the Assessment of Side Chain Dynamics Predictions by Current AMBER and CHARMM Force Fields. *J. Chem. Theory Comput.* **2020**, *16*, 2896−2913.

(8) Anderson, J. S.; Hernández, G.; LeMaster, D. M. Prediction of Bond Vector Autocorrelation Functions from Larmor Frequency-Selective Order Parameter Analysis of NMR Relaxation Data. *J. Chem. Theory Comput.* **2017**, *13*, 3276−3289.

(9) Peyron, M.; Pierens, G. K.; Lucas, A. J.; Hall, L. D.; Stewart, R. C. The modified stretched-exponential model for characterization of NMR relaxation in porous media. *J. Magn. Reson., Ser. A* **1996**, *118*, 214−220.

(10) Buevich, A. V.; Shinde, U. P.; Inouye, M.; Baum, J. Backbone dynamics of the natively unfolded pro-peptide of subtilisin by heteronuclear NMR relaxation studies. *J. Biomol. NMR* **2001**, *20*, 233−249.

(11) Abergel, D.; Volpato, A.; Coutant, E. P.; Polimeno, A. On the reliability of NMR relaxation data analyses: a Markov chain Monte Carlo approach. *J. Magn. Reson.* **2014**, *246*, 94−103.

(12) Hsu, A.; Ferrage, F.; Palmer, A. G. Analysis of NMR Spin-Relaxation Data Using an Inverse Gaussian Distribution Function. *Biophys. J.* **2018**, *115*, 2301−2309.

(13) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. Structure of ubiquitin refined at 1.8 Å resolution. *J. Mol. Biol.* **1987**, *194*, 531−544.

(14) MacKerell, A. D.; Feig, M.; Brooks, C. L. Improved treatment of the protein backbone in empirical force fields. *J. Am. Chem. Soc.* **2004**, *126*, 698−699.

(15) MacKerell, A. D.; Feig, M.; Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics. *J. Comput. Chem.* **2004**, *25*, 1400−1415.

(16) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. D.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the backbone ϕ,ψ and Side-Chain χ_1 and χ_2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257−3273.

(17) Huang, J.; MacKerell, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **2013**, *34*, 2135−2145.

(18) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Taikhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781−1802.

(19) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712−725.

(20) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696−3713.

(21) Wang, L. P.; McKiernan, K. A.; Gomes, J.; Beauchamp, K. A.; Head-Gordon, T.; Rice, J. E.; Swope, W. C.; Martinez, T. J.; Pande, V. S. Building a More Predictive Protein Force Field: A Systematic and Reproducible Route to AMBER-FB15. *J. Phys. Chem. B* **2017**, *121*, 4023−4039.

(22) Debiec, K. T.; Cerutti, D. S.; Baker, L. R.; Gronenborn, A. M.; Case, D. A.; Chong, L. T. Further along the Road Less Traveled: AMBER ff15ipq, an Original Protein Force Field Built on a Self-Consistent Physical Model. *J. Chem. Theory Comput.* **2016**, *12*, 3926−3947.

(23) Wang, L. P.; Martinez, T. J.; Pande, V. S. Building Force Fields: An Automatic, Systematic, and Reproducible Approach. *J. Phys. Chem. Lett.* **2014**, *5*, 1885−1891.

(24) Lakomek, N. A.; Ying, J.; Bax, A. Measurement of ^{15}N relaxation rates in perdeuterated proteins by TROSY-based methods. *J. Biomol. NMR* **2012**, *53*, 209−221.

(25) Ferrage, F.; Reichel, A.; Battacharya, S.; Cowburn, D.; Ghose, R. On the measurement of ^{15}N-{^{1}H} nuclear Overhauser effects. 2. Effects of the saturation scheme and water signal suppression. *J. Magn. Reson.* **2010**, *207*, 294−303.

(26) Palmer, A. G.; Kroenke, C. D.; Loria, J. P. Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods Enzymol.* **2001**, *339*, 204−238.

(27) Hernández, G.; LeMaster, D. M. Quantifying protein dynamics in the ps-ns time regime by NMR relaxation. *J. Biomol. NMR* **2016**, *66*, 163−174.

(28) Tjandra, N.; Feller, S. E.; Pastor, R. W.; Bax, A. Rotational diffusion anisotropy of human ubiquitin from 15N NMR relaxation. *J. Am. Chem. Soc.* **1995**, *117*, 12562−12566.

(29) Mandel, A. M.; Akke, M.; Palmer, A. G. Backbone dynamics of *Escherichia coli* ribonuclease HI: Correlations with structure and function in an active enzyme. *J. Mol. Biol.* **1995**, *246*, 144−163.

(30) Woessner, D. E. Nuclear Spin Relaxation in Ellipsoids Undergoing Rotational Brownian Motion. *J. Chem. Phys.* **1962**, *37*, 647−654.

(31) Farrow, N. A.; Zhang, O.; Forman-Kay, J. D.; Kay, L. E. Comparison of the backbone dynamics of a folded and an unfolded SH3 domain existing in equilibrium in aqueous buffer. *Biochemistry* **1995**, *34*, 868−878.

(32) Ishima, R.; Nagayama, K. Protein Backbone Dynamics Revealed by Quasi Spectral Density-Function Analysis of Amide N-15 Nuclei. *Biochemistry* **1995**, *34*, 3162−3171.

(33) Farrow, N. A.; Zhang, O.; Szabo, A.; Torchia, D. A.; Kay, L. E. Spectral Density Function Mapping using ^{15}N Relaxation Data Exclusively. *J. Biomol. NMR* **1995**, *6*, 153−162.

(34) LeMaster, D. M. Larmor Frequency Selective Model Free Analysis of Protein NMR Relaxation. *J. Biomol. NMR* **1995**, *6*, 366−374.

(35) Smith, A. A.; Ernst, M.; Meier, B. H. Optimized "detectors" for dynamics analysis in solid-state NMR. *J. Chem. Phys.* **2018**, *148*, No. 045104.

(36) Smith, A. A.; Ernst, M.; Meier, B. H.; Ferrage, F. Reducing bias in the analysis of solution-state NMR data with dynamics detectors. *J. Chem. Phys.* **2019**, *151*, No. 034102.

(37) Case, D. A.; Betz, R. M.; Cerutti, D. S.; T E Cheatham, I.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Luchko, T.; Luo, R.; Madej, B.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Botello-Smith, W. M.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Xiao, L.; Kollman, P. A. *AMBER 2016*; University of California: San Francisco, CA, 2016.

(38) Cook, W. J.; Jeffrey, L. C.; Kasperek, E.; Pickart, C. M. Structure of tetraubiquitin shows how multiubiquitin chains can be formed. *J. Mol. Biol.* **1994**, *236*, 601−609.

(39) Reyes-Turcu, F. E.; Horton, J. R.; Mullally, J. E.; Heroux, A.; Cheng, X.; Wilkinson, K. D. The Ubiquitin Binding Domain ZnF UBP Recognizes the C-Terminal Diglycine Motif of Unanchored Ubiquitin. *Cell* **2006**, *124*, 1197−1208.

(40) Liu, Z.; Zhang, W. P.; Xing, Q.; Ren, X.; Liu, M.; Tang, C. Noncovalent Dimerization of Ubiquitin. *Angew. Chem., Int. Ed.* **2012**, *51*, 469−472.

(41) Shine, A.; Shenoy, J.; Jayan, P.; Jiji, A. C.; Vijayan, V. Residual Dipolar-Coupling-Based Conformational Comparison of Noncovalent Ubiquitin Homodimer with Covalently Linked Diubiquitin. *ChemPhysChem* **2020**, *21*, 888−894.

(42) Damberg, P.; Jarvet, J.; Gräslund, A. Limited Variations in ^{15}N CSA Magnitudes and Orientations in Ubiquitin are Revealed by Joint Analysis of Longitudinal and Transverse NMR Relaxation. *J. Am. Chem. Soc.* **2005**, *127*, 1995−2005.

(43) Hall, J. B.; Fushman, D. Variability of the ^{15}N Chemical Shielding Tensors in the B3 Domain of Protein G from ^{15}N Relaxation Measurements at Several Fields. Implications for Backbone Order Parameters. *J. Am. Chem. Soc.* **2006**, *128*, 7855−7870.

(44) Kroenke, C. D.; Rance, M.; A G Palmer, I. Variability of the ^{15}N chemical shift anisotropy in *Escherichia coli* Ribonuclease H in solution. *J. Am. Chem. Soc.* **1999**, *121*, 10119−101125.

(45) LeMaster, D. M.; Kay, L. E.; Brunger, A. T.; Prestegard, J. H. Protein Dynamics and Distance Determination by NOE Measurements. *FEBS Lett.* **1988**, *236*, 71−76.

(46) Case, D. A. Calculations of NMR dipolar coupling strengths in model peptides. *J. Biomol. NMR* **1999**, *15*, 95−102.