BIOCHEMISTRY
including biophysical chemistry & molecular biology

CURRENT TOPIC

pubs.acs.org/biochemistry

Nuclear Magnetic Resonance Provides a Quantitative Description of Protein Conformational Flexibility on Physiologically Important Time Scales

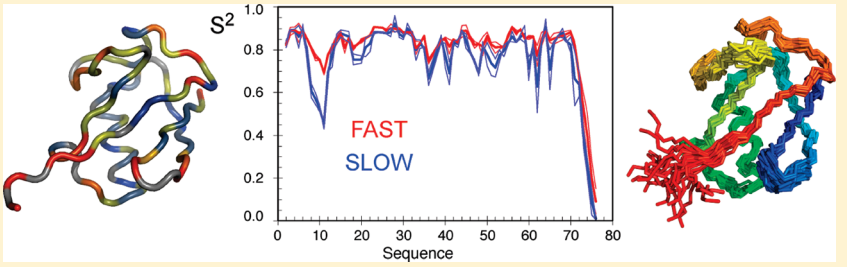
Loïc Salmon,[†] Guillaume Bouvignies,[‡] Phineus Markwick,[§] and Martin Blackledge*[,†]

[†]Protein Dynamics and Flexibility, Institut de Biologie Structurale Jean-Pierre Ebel, CEA, CNRS, UJF UMR 5075, 41 Rue Jules Horowitz, Grenoble 38027, France

[‡]Departments of Molecular Genetics, Biochemistry, and Chemistry, The University of Toronto, Toronto, Ontario, Canada M5S 1A8

[§]Department of Chemistry and Biochemistry, University of California at San Diego, 9500 Gilman Drive, Urey Hall, La Jolla, California 92003-0365, United States

ABSTRACT:



A complete description of biomolecular activity requires an understanding of the nature and the role of protein conformational dynamics. In recent years, novel nuclear magnetic resonance-based techniques that provide hitherto inaccessible detail concerning biomolecular motions occurring on physiologically important time scales have emerged. Residual dipolar couplings (RDCs) provide precise information about time- and ensemble-averaged structural and dynamic processes with correlation times up to the millisecond and thereby encode key information for understanding biological activity. In this review, we present the application of two very different approaches to the quantitative description of protein motion using RDCs. The first is purely analytical, describing backbone dynamics in terms of diffusive motions of each peptide plane, using extensive statistical analysis to validate the proposed dynamic modes. The second is based on restraint-free accelerated molecular dynamics simulation, providing statistically sampled free energy-weighted ensembles that describe conformational fluctuations occurring on time scales from pico- to milliseconds, at atomic resolution. Remarkably, the results from these two approaches converge closely in terms of distribution and absolute amplitude of motions, suggesting that this kind of combination of analytical and numerical models is now capable of providing a unified description of protein conformational dynamics in solution.

Protein structure determination has contributed immensely to our understanding of biology, providing three-dimensional atomic resolution descriptions of the molecular basis of physiologically important interactions between biochemically active molecules.[1] However, as we emerge from a decade of massive investment in structural genomics projects,[2] it is becoming increasingly clear that a complete description of biomolecular activity requires additional understanding of the nature and the role of the associated conformational dynamics of the protein. Proteins are intrinsically dynamic, exhibiting a flexibility that can be manifest in terms of large-scale reorganizations[3] or small-scale conformational fluctuations of backbone and side chain atoms about their mean conformation.[4] There is a growing body of evidence that shows that proteins intrinsically sample conformational flexibility that encodes their function, sampling different structures that may be required for interaction with different partners, emphasizing the necessity for the development of an atomic-resolution understanding of this flexibility.[5] An accurate description of the complex thermally accessible conformational space inherent to the protein at its physiological temperature is therefore essential to the understanding of its biomolecular function.

Despite the importance of molecular dynamics, three-dimensional protein structure determination, using for example X-ray diffraction or nuclear magnetic resonance (NMR) spectroscopy, has largely ignored protein motions, routinely representing rapidly exchanging conformational equilibria in terms of single static structures. NMR is, however, uniquely suited to characterization of protein structure and dynamics in solution, because of its atomic-resolution sensitivity to conformational fluctuations occurring on vastly different time scales, ranging from picoseconds to days. In fact, all NMR spectra report on a broad

Received: February 3, 2011
Revised: March 5, 2011
Published: March 09, 2011

conformational average, whose correct interpretation underpins NMR-based structural biology. Each measurable resonance represents an average of rapidly interchanging conformations whose difference in chemical shift is smaller than their interconversion rates, corresponding to an almost overwhelmingly vast dynamic average over the equilibrium ensemble. The accurate description of this ensemble average necessarily represents a major task for the biomolecular NMR spectroscopist. Unfortunately, despite major advances in the use of chemical shifts for the determination of protein structures,[6,7] the commonly measured population-weighted chemical shift itself cannot yet be used to describe this conformational ensemble. Other, so-called first-order interactions such as scalar and dipolar couplings, which also have a direct impact on the measured spectrum, can be used to characterize ensemble averaging over the same time scale and therefore provide the ideal tool with which to probe functionally important motions occurring in biomolecules.[8−10] In this review, we will present recent results that provide a consistent and quantitative description of the extent and nature of intrinsic dynamics in folded proteins in solution.

NMR AND DYNAMIC EQUILIBRIUM

NMR is of course routinely used to directly probe protein motions across a broad range of time scales, although such studies tend to be conducted in a manner independent of the structure determination step. The most common approach is to measure spin relaxation, describing the phenomenological mechanisms returning an excited nuclear spin state to its equilibrium value.[11] In the case of macromolecules in solution, ^{15}N and ^{13}C relaxation processes are dominated by the stochastic reorientational properties of physical interactions that induce random fields that relax the excited spin state back to equilibrium. In the case of ^{15}N relaxation, dipolar interactions between ^{15}N and ^{1}H^{N} spins dominate at magnetic fields below 20 T, so that measured relaxation rates report on the angular reorientational correlation function of internuclear bond vectors. Motions on time scales that are faster than the characteristic molecular rotational diffusion time constant τ_c (in the range of 5−30 ns for typical soluble proteins) are now routinely measured for nuclear spin pairs distributed throughout the protein backbone and side chains, providing information about the amplitude and frequencies of structural fluctuations.[12,13] Comparison of the dynamic behavior of proteins in the presence and absence of physiological cofactors can provide mechanistic insight into the role of fast motions in biological function. Dynamics occurring on time scales in the micro- to millisecond range are potentially of even greater interest because many biologically important processes, such as enzymatic catalysis, signal transduction, ligand binding, and allosteric regulation, are expected to occur on these time scales. For this reason over the past decade, there has been considerable development of techniques for accurately probing these slower time scale motions at atomic resolution.

One of the most powerful experimental developments in recent years exploits the fact that changes in the electronic environment of a nuclear spin on a time scale in the micro- to millisecond range can give rise to measurable chemical shift exchange phenomena that influence effective transverse spin relaxation rates. Rotating frame relaxation experiments[14,15] applied to the description of the dynamics and thermodynamics of conformational equilibria in small biopeptides used one-dimensional, homonuclear ^{1}H NMR.[16−19] In the past two decades, related methods have been extended to ^{15}N- and ^{13}C-labeled proteins[20,21] and are now routinely applied in probing molecular motions in the most challenging proteins in terms of size and complexity.[22−24] In addition to the accurate determination of time scales of conformational exchange, translation of measured exchange phenomena into effective chemical shifts and populations allows the extraction of structurally dependent parameters of the dynamically exchanging forms of the molecule, even in the case of very weakly populated substates.[25] This sensitivity to so-called "invisible states" has proven to be critical for providing structural and dynamic information about excited species that are often crucial for protein function, allowing, for example, the determination of atomic-resolution structures of protein folding intermediate states.[26]

By its nature, the nuclear Overhauser effect (NOE),[27] the standard tool for protein structure determination,[28] also contains important information defining disorder in the conformational ensemble. The NOE measures the efficiency of cross relaxation between two protons and depends on the motion of the spin pair and its spatial proximity. The standard usage of the NOE transforms the measured cross relaxation rates into a single structural constraint on the basis of the theoretical dependence of $1/r^6$ for a fixed spin pair bound to a molecule with a known rotational correlation time.[29] In fact, the measured NOE is also dependent on internal motions and on the characteristic correlation time of these motions.[30] Normally, the combination of a large number of such constraints is used to determine a single structure, such that possible deviations of individual rates from the simple assumption used to extract each distance tend to average incoherently and produce barely significant effects on the final structural model. Numerous approaches have, however, been proposed, to capture the essential motions present in solution on the basis of NOE. The majority are ensemble averaging methods relying on restrained molecular dynamics using multiple explicit copies of the protein,[31,32] or some kind of Monte Carlo sampling procedures.[33] While the modeling of ensemble-averaged experimental parameters with an explicit ensemble is obviously an appealing idea, the drawback when applied to the NOE is that the dynamic processes responsible for the measured experimental parameters are potentially more complex than a simple population-weighted average. For example, motions occurring on time scales faster than the overall correlation time of the protein can be averaged differently with respect to motions occurring on time scales slower than τ_c, for the same average internuclear distance and the same geometric parameters. The determination of a dynamic ensemble on the basis of NOEs is an underdetermined problem unless time scales and geometric distributions are known a priori. This statement holds, and is probably exacerbated, when NOEs are combined with other NMR parameters that are averaged on different time scales (vide infra). While great progress has been made over the past five years in addressing the correct statistical analysis of NOE data in terms of conformational models, in particular using inferential structure determination,[34] it is clear that the treatment of conformational flexibility remains of utmost importance in terms of both structural accuracy and understanding of biological function.

RESIDUAL DIPOLAR COUPLINGS DIRECTLY PROBE CONFORMATIONAL SPACE

Over the past decade, residual dipolar couplings (RDCs) have emerged as powerful tools for studying protein structure and dynamics in solution. The dipolar or dipole−dipole interaction results from the coupling of a given magnetic moment with any other magnetic moment in its surroundings, and this coupling is given by[11]

$$D_{IS} = -\frac{\gamma_I \gamma_S \mu_0 h}{16\pi^3 r_{IS}^3}\langle P_2 \cos(\theta_{IS})\rangle \quad (1)$$

where γ is the gyromagnetic ratio for the two spins, r is the distance between the spins, μ_0 is the permeability of free space, and h is Planck's constant. Two extreme cases can be considered: the isotropic case, in which the interaction samples all orientations in free solution with equal probability, the measured value is averaged to zero, and the information content is lost, and the solid state case, in which the nearly complete absence of macroscopic motion leaves the dipolar interaction almost undiminished. In this case, the dipolar coupling between spins that are close in space (tens of kilohertz for covalently bound $^{15}N-^{1}H$ spins) can dominate the entire spectrum unless removed by artificial means. The idea of using a mesomorphic or liquid crystalline phase is to retain the essential geometric information content in a residual dipolar coupling that can be measured using the simplicity of liquid state NMR.[35,36]

RDCs are first and foremost highly sensitive structural constraints. For this reason, the majority of publications treating RDCs measured in biomolecules have assumed a so-called static description that supposes the absence of differential dynamics within the framework of the molecular or alignment frame. This approximation ignores the Boltzmann-weighted population of available conformational substates that one would expect to exist in such a complex macromolecule. However, a number of initial comparisons with data from model proteins for which high-resolution structures were already known indicated that a very good reproduction of experimental data could be obtained by invoking only a static structure. This suggested that the amount of dynamics present was either minimal or at the very least averaged in such a way that it remained in agreement with a static description. RDCs are now routinely incorporated into structure refinement protocols, leading to better-defined structural bundles. It is important to exercise some caution here; it is well-documented that a single RDC has a very broad angular degeneracy, such that refinement using only RDCs from $^{15}N-^{1}H^{N}$ pairs in a protein carries little additional structural information,[10] and it has recently been noted that even a peptide plane with multiple RDCs arranged in different directions can adopt 16 equally valid orientations.[37] When data from differently aligning media are combined, the intrinsic structural information becomes far more powerful,[38] even leading to the determination of the three-dimensional protein structure using only RDCs.[39,40]

More importantly, RDCs provide simultaneous information about time- and ensemble-averaged structural and dynamic processes occurring on time scales up to milliseconds and thereby encode key information for understanding biomolecular motions. Any measured interaction is modulated by motion occurring on time scales faster than the interaction magnitude expressed in frequency units. Typically, RDCs are in the range of tens of hertz, which corresponds to time scales of a few tens of milliseconds. As the chemical shift coalescence limit often occurs on times scales around the millisecond time scale, an RDC will be sensitive to all time scales up to the coalescence limit. This property confers an important role on RDCs, that of a motional probe.

The most obvious application of RDCs to the understanding of conformational flexibility in biomolecules is the consideration of domain motions within a larger macromolecular ensemble. This was first demonstrated in a study of magnetically aligned cyanometmyoglobin, in which the presence of diffusive motion of α-helices was invoked on the basis of RDCs and paramagnetic chemical shifts.[41] Following this pioneering study, a number of systems have been shown to exhibit domainlike dynamic behavior using similar approaches, for proteins[42−45] and RNA[46−49] or oligosaccharides.[50]

RDCs are also particularly sensitive to the structural details of local conformational dynamics, and this aspect will be the subject of the remainder of this review. The simple averaging properties of RDCs make them amenable to rigorous interpretation, and this means that assumptions made when analyzing protein dynamics from RDCs, as well as the robustness of the resulting dynamic description, can be statistically substantiated using self-consistency checks. This makes RDCs possibly even more powerful than standard spin relaxation experiments, where the potential for cross checking via independent sources of information is limited, with the additional advantage that they sample longer time scales. The disadvantage is of course that the time scale of any motions that are characterized can only be placed within an upper limit of the millisecond range. Nevertheless, comparison between dynamic amplitudes, normally expressed in terms of order parameters (S^2) extracted from RDCs and spin relaxation, can provide precise information about the presence of dynamics in the picosecond to millisecond range.

Diverse techniques have been proposed to characterize local protein backbone conformational dynamics from RDCs, and these can be classified into two generic approaches. The first directly determines dynamic amplitudes and anisotropies of bond vectors or structural motifs from multiple RDC measurements.[51−60] This kind of approach has the advantage of being independent of molecular models and relatively free from the bias incurred upon combination of potential energy force fields and experimental target functions. The disadvantages are that the motional parameters tend to be expressed in an abstract or mathematical form that is difficult to directly translate into a format comprehensible to the structural biologist. As with all fitting procedures, there is also the possibility that noise can be fitted into the motional parameters. The second, diametrically opposite approach, exploits the heavy artillery of molecular dynamics simulation to develop a picture of the protein ensemble that is in agreement with experimental data.[5,32,61−64] The implicit risk with the latter approach is that the statistical mechanical ensemble may not be properly described when the ensemble is driven into a conformational description that is in agreement with experimental data, and that it is not necessarily simple to detect a manifestation of this kind of artifact. The advantage, however, is evident: an explicit molecular description of the ensemble giving rise to experimental data and describing the conformational space sampled up to millisecond time scales holds immense potential for understanding the mechanistic basis of biomolecular interaction and function.

The first kind of approach, which one can term "analytical", attempts to determine the average spherical harmonics $\langle Y_m^2(\theta_i,\phi_i)\rangle$ that define the orientational sampling of each vector

relative to a common molecular alignment frame,[51,65] using the expression

$$D_i^j = A_a^j \frac{\gamma_I \gamma_S \mu_0 h}{8\pi^3 r_{IS}{}^3}\sqrt{\frac{4\pi}{5}}\left\{\langle Y_2^0(\theta_i,\phi_i)\rangle + \sqrt{\frac{3}{8}}\,R[\langle Y_2^2(\theta_i,\phi_i)\rangle + \langle Y_{-2}^2(\theta_i,\phi_i)\rangle]\right\} \quad (2)$$

where A_a is the amplitude and R the rhombicity of the alignment tensor. The measurement of RDCs in the presence of different alignment media changes the way the same dynamics averages the measured RDC. The spherical harmonics can be averaged in the different tensorial frames using Wigner rotation matrices and can therefore be mathematically determined from more than five independent experimental data sets. When considering a single dipolar interaction or internuclear vector, we can talk of "model-free" analytical approaches, because, as for relaxation data analysis, it describes motional averaging without invoking any kind of physical model.[13] This avoids bias due to an inevitably imperfect model but limits the possible interpretation of the results as most of the obtained parameters are purely mathematical and therefore sometimes difficult to interpret physically. Two different model-free approaches have been developed to extract dynamic information from RDCs.[51,65] Both are based on matrix descriptions of RDCs measured in different alignment media and determine the sampling characteristics defined by the spherical harmonic terms shown in eq 2. Although the details of the formalisms differ significantly, they lead to similar structural and dynamic information. Both authors noted, however, that the quantitative determination of $\langle Y_m^2(\theta_i,\phi_i)\rangle$ is dependent on an accurate determination of the alignment tensors (j), and in particular the amplitude term for each tensor, A_a^j. When only $^{15}N-{}^1H^N$ vectors are used in the analysis, it is virtually impossible to distinguish between changes in the overall alignment and overall changes in local dynamics, so that the authors have scaled the amount of dynamics to be lower than or equal to the levels determined from spin relaxation, measured on essentially the same bond vectors. This is justifiable because the amplitudes of motions occurring on all time scales up to the millisecond time scale must be equal to or greater than motions occurring up to the nanosecond time scale, which are sampled by spin relaxation. It is worth noting that different approaches to the treatment of overall rotational diffusion or the selection of the complexity of the models used to describe local motion can modulate the level of $^{15}N-{}^1H^N$ order parameters determined from spin relaxation measured at one magnetic field strength,[66] so that caution should possibly be exercised when taking these numbers as absolute references. Iteratively refined versions of this kind of interpretation of RDCs have been applied to the protein ubiquitin using data measured using many alignment media.[52,67−70] This resulted in the evocation of differing levels of slow dynamics, with pervasive slow motional order parameters ranging from 0.8 to 0.9, depending on the combination of data sets used in the analysis and the exact determination of the level of molecular alignment (vide infra). The origin of these disparities may also derive from the fact that the data derived from different alignment media are often of very different quality, leading to problems of relative weighting and potential overfitting to noise. In this respect, one of the most important advances was the application of the SECONDA analysis to a collection of more than 30 data sets.[71] This approach provides an entirely model-free criterion that can be used to identify data sets that are incompatible because of excessive noise or interaction with the alignment medium and allowed the larger data set to be pruned to a subset of data measured in 23 different alignment media.

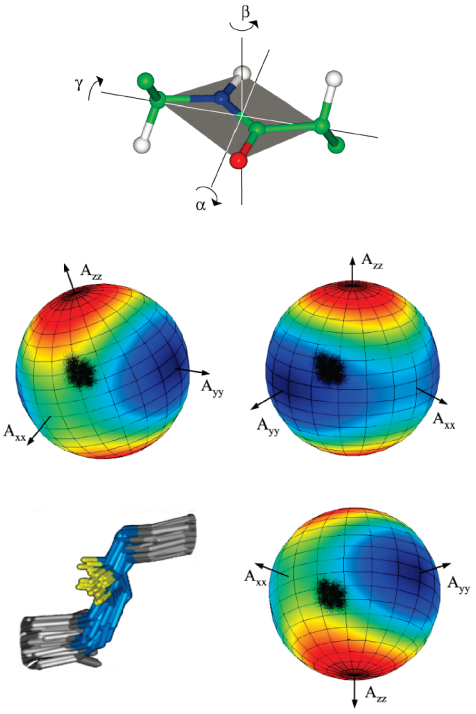


Figure 1. Gaussian axial fluctuation model of peptide plane reorientation (top) showing the three axes about which the amplitude of diffusive motions is determined in RDC-based GAF approaches. The four bottom panels illustrate how orientational sampling averages differently with respect to different molecular alignment tensors. The three spheres show the range of RDC values from large positive (red) to large negative (blue) for three differently oriented tensors, emanating from differently aligning media.

To overcome some of the difficulties of the model-free method, researchers have developed related approaches that combine the angular averaging properties of multiple bond vectors within a known structural motif. The first of these approaches proposed to determine effective local alignment tensors, giving a generalized degree of order for units of local structure comprising each C^α junction in the protein.[53] While this requires precise knowledge of one of the backbone dihedral angles in the protein (the ϕ angle) and is therefore applicable under only certain assumptions, the concept was a clear-sighted precursor to approaches that consider movement of peptide planes. These approaches were based on the analysis of multiple RDCs oriented in different directions in a single peptide plane and exploit the principle that anisotropic reorientational motion of the peptide plane will average differently oriented RDCs (Figure 1).[54] The apparent uniformity of the peptide unit throughout the chain, combined with the existence of a simple model for peptide plane motions, offers the possibility of characterizing the backbone dynamics of a protein using the multiple RDCs that can be measured in each plane of a ^{13}C- and ^{15}N-labeled protein. The most appropriate model for the description of peptide plane motions is the so-called Gaussian axial diffusion or GAF model, originally developed for the interpretation of spin relaxation data,[72−74] is a general extension of the so-called crankshaft motional model, and allows for diffusive motions around three orthogonal axes attached to each plane (Figure 1).

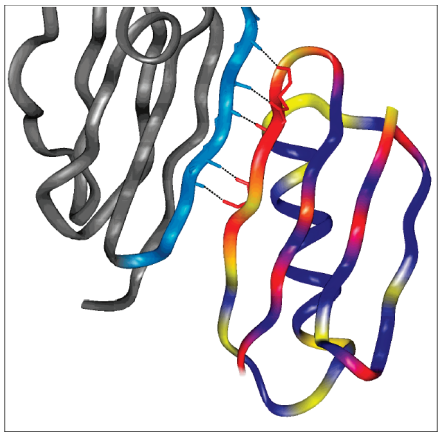


Figure 2. Illustration of the amplitude (red for a large amplitude and blue for a smaller amplitude) γ-motions present in the β-sheet of protein GB3. The motion across the sheet is found to be correlated, on the basis of trans-hydrogen bond scalar couplings, and reaches a maximal amplitude in the interaction site with Fab (colored gray and blue). The interaction is mediated via hydrogen bonds.

In an early work, it was shown that using a common amplitude for the γ-motion (Figure 1) for all the peptide planes in secondary structural elements of high-resolution protein structures for which $^{15}N-^{1}H^{N}$ RDCs had been measured led to a statistically significant improvement in the reproduction of data compared to a static analysis.[54] This was the first demonstration that anisotropic motions could be detected with statistical certainty and provided an important indication that this model was relevant to the averaging properties of backbone RDCs in proteins. It also explained previous observations that to simultaneously reproduce $^{15}N-^{1}H^{N}$ RDCs and $^{13}C^{\alpha}-^{13}C'$ RDCs using the same alignment tensor, an unusually long distance for the $^{15}N-^{1}H^{N}$ bond was necessary (1.04 Å) when using a static description.[75] This is simply an experimental manifestation of the additional motional averaging that this direction experiences compared to the $^{13}C^{\alpha}-^{13}C'$ direction. This is also the first indication that this anisotropy of peptide plane motion can provide the key to the determination of absolute levels of backbone dynamics in proteins that will be described in more detail below.

A complete three-dimensional GAF (3D-GAF) analysis of local motion was conducted using an extensive set of RDCs from the third immunoglobin binding domain of streptococcal protein G (GB3).[76] The average positions of the heavy atoms in each peptide plane were assumed to be identical to those present in a high-resolution X-ray structure, and the alignment tensors, which provide the reference frame relative to which the dynamics are determined, were defined using the RDCs associated with the least dynamic bond vectors identified from an initial analysis. Following the determination of alignment tensors, the coordinates of the protein were rotated into the principal axis system of each alignment tensor and the averaged coupling calculated as a function of σ_{α}, σ_{β}, and σ_{γ}, the amplitudes of motion about the three axes using

$$\frac{\langle D\rangle^{\text{3D-GAF}}}{D_{\max}} = \sqrt{\frac{4\pi}{5}}\left\{\langle Y_0^2(\theta_i,\phi_i)\rangle^{\text{3D-GAF}} + \sqrt{\frac{3}{8}}\,R[\langle Y_2^2(\theta_i,\phi_i)\rangle^{\text{3D-GAF}} + \langle Y_{-2}^2(\theta_i,\phi_i)\rangle^{\text{3D-GAF}}]\right\} \quad (3)$$

where

$$\langle Y_m^2(\theta_i,\phi_i)\rangle^{\text{3D-GAF}} = \sum_{l',l''=-2}^{2}\left\{e^{-l'^2\sigma_\gamma^2/2}e^{-i(l''-M)\phi_z}d_{l',M}(-\theta_z)d_{l'',l'}(\theta_z) \times \sum_{n',n''=-2}^{2}\left[e^{-n'^2\sigma_\beta^2/2}e^{-i(n''-l'')\phi_y}d_{n',l''}(-\theta_y)d_{n'',n'}(\theta_y) \times \sum_{m',m''=-2}^{2}e^{-m'^2\sigma_\alpha^2/2}e^{-i(m''-n'')\phi_x}d_{m',n''}(-\theta_x)d_{m'',m'}(\theta_x)Y_{2,m''}(\Omega)\right]\right\} \quad (4)$$

This study revealed an alternating pattern for the amplitude of the γ- and β-motions in the β-sheet, where the hydrophobic, buried residues were the least dynamic (Figure 2). This motion resembles the so-called "backrub" motion,[77] whereby concerted motion of adjacent backbone peptide planes results in an increased level of motion of the intervening side chain; in this case, these are the solvent-exposed sites. The amplitudes of motion increased from one edge of the sheet to the other, leading to maximal dynamic sampling in the interaction site, supporting conformational selection as a driving force for the interaction with its physiological partner, Fab.[78] The results were supported by extensive cross validation of so-called passive data that were not used in the analysis. Correlations between motions of peptide planes linked by hydrogen bonds across the β-sheet were revealed using $^{3}J_{N'C}$ trans-hydrogen bond scalar couplings. This observation provided a rare insight into one of the possible mechanisms by which proteins transfer information through space via correlated motions, in this case across a β-sheet, and has implications for our understanding of allosteric regulation. To address the level to which this approach is dependent on the precision of the structure used to represent the average coordinates in solution, the procedure was repeated using an NMR structure refined using many of the RDCs used in the dynamic analysis. Remarkably, the results were essentially identical, indicating again that the intrinsic dynamic information encoded in each average RDCs cannot be completely incorporated into a single average structure, and that the motional characteristics of the vectors involved can still be determined even after refinement against the RDCs. We also note that relaxation rates measured in the same protein in the solid state revealed a similar alternating pattern of ^{13}C relaxation rates in the β-strand, indicating that these motions are intrinsic to the fold of the protein.[79]

This analysis also suggested that one could go further than fitting dynamic amplitudes to RDCs on the basis of a known average structure and actually dispense entirely with the three-dimensional structural model from the start. This requires the simultaneous determination of the average structure and the associated dynamic modes and amplitudes describing motions about this mean. In this case, the tensor determination step was improved by allowing a one-dimensional GAF (1D-GAF) motion during the determination of the tensor parameter, during which the orientations of all planes were also simultaneously optimized.[60] Then a backbone structure was determined using an approach called *Dynamic Meccano*. This approach is based on the Meccano (Molecular Engineering Calculations using Coherent Association of Nonaveraged Orientations) method that was developed to sequentially orient all the protein peptide planes in the protein using only RDCs. Here the approach was

combined with analytically defined dynamic averaging using a GAF motion, leading to results very similar to those of the previous analysis determined using the high-resolution crystal structure. Perhaps surprisingly, the backbone coordinates determined using this approach were remarkably similar (backbone rmsd of <0.5 Å) to those of high-resolution X-ray crystallographic and NMR structures. This already suggests that protein dynamics are occurring in a pseudo-harmonic potential, allowing the structural constraints from NMR and diffraction to average to a similar conformation as the true dynamic mean conformation determined here. This was the first simultaneous determination of the high-resolution backbone structure and dynamics occurring on time scales up to the millisecond time scale that explicitly accounted for all dynamic fluctuations modulating the experimental data. These approaches were tested through cross validation of RDCs that were not used in the analysis (for example, couplings between α-carbons and protons), demonstrating for the first time that a dynamic description better reproduced independent data than an optimally applied static approach. Extensive simulations were conducted using molecular dynamics trajectories from which RDCs were calculated, indicating that the approach is robust when four sufficiently different alignment tensors are available.[80] This also highlights one of the key advantages of using a model that combines the orientational averaging properties of multiple RDCs within the same structural unit, that the minimum number of alignment media that are required is less experimentally demanding than for model-free approaches.

While the power of RDCs to define both structure and dynamics to high resolution has been established, the quantitative determination of dynamics remains challenging. As described earlier, this problem has often been addressed by referencing or scaling order parameters to an external reference such as spin relaxation-derived order parameters. This issue is particularly important, first because very little is known about the absolute level of dynamic fluctuations present in proteins and second because RDCs are especially prone to errors of absolute dynamic amplitudes, because of the potential for a component of the dynamic averaging of RDCs to be absorbed into the estimated magnitudes of the alignment tensors. If we consider eq 2, it is apparent that an accurate characterization of the true level of local dynamics is impossible unless the properties of the alignment tensor are carefully determined. If the tensors are determined without considering motion, average components of both anisotropic and isotropic motional modes can be absorbed into A_a^j so that the effective molecular alignment appears lower. This reduces dynamic amplitudes extracted using any of the proposed approaches.

To address the question of how precisely one can determine the true level of structural fluctuations on time scales up to the millisecond time scale, it is important to consider a number of potential sources of artifacts. These include (1) the accurate estimation of the alignment of the protein, (2) the influence of the alignment medium on the internal dynamics of the protein, (3) the influence of noise in the experimental data, and (4) the dependence on the coordinates of any average structure used in the analysis. In principle, the approaches presented here can address points 1−3 specifically and remove point 4 in terms of global fold, but inherently depend on the precision, and generalizability, of the local structure in the peptide plane used to model each interpeptide unit in the protein. A recent study specifically attempted to determine the level of slow dynamics present in

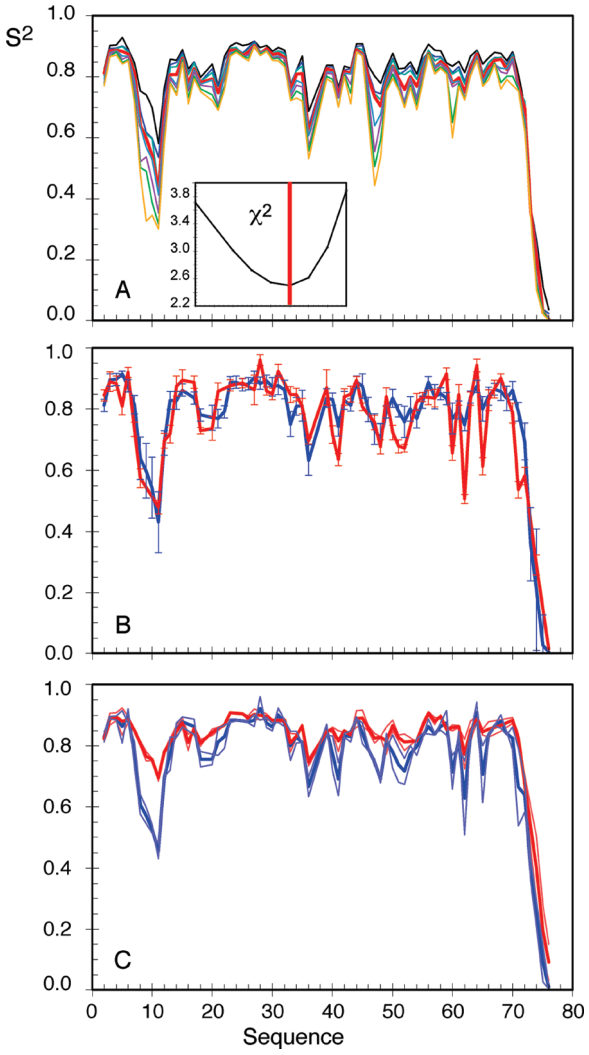


Figure 3. N−H^N S^2 order parameters for ubiquitin. (A) S^2 values derived from free energy-weighted ensembles derived from AMD simulations performed at increasing levels of acceleration (black represents the lowest level and orange the highest). The optimal level of acceleration, as determined from the χ^2 with respect to the experimental values (inset), is represented by the red line (also in the inset). (B) Comparison of AMD-derived (blue) and GAF-derived (red) S^2 values. (C) Comparison of fast motional S^2 values (the thin red lines show two independent experimental data sets measured 10 years apart, in different groups, and extracted using different analytical procedures; the thick red line shows the mean of the two values) and slow motional S^2 values (the thin blue line shows the two curves in panel B; the thicker blue line shows their mean).

ubiquitin by taking into account all of these sources of inaccuracy using a structure-free GAF approach (SF-GAF).[81]

Starting from the data sets that have been identified using the SECONDA self-consistency analysis, comprising couplings from 24 alignment media, the following function was minimized:

$$\chi^2[\{\theta,\phi,\psi\}_i,\{\mathbf{A}\}_j,\{S,\sigma_\alpha,\sigma_\beta,\sigma_\gamma\}_i] = \sum_{ij}(D_{i,j}^{\mathrm{exp}} - D_{i,j}^{\mathrm{calc}})^2/\delta_{i,j}^2 \quad (5)$$

where angles θ, ϕ, and ψ describe the mean orientation of plane i, **A** represents the alignment tensor for medium j, and σ_α, σ_β, σ_γ,

and S (order parameter) are the amplitudes of the GAF or axially symmetric motions that best fit the data for plane i. δ_{ij} is the estimated weighting of each RDC data set, determined using a robust scaling estimator, based on statistical analysis of the dispersion of the reproduction of experimental data. The average Cartesian coordinates of the heavy atoms in each peptide plane were taken from averages over ultra-high-resolution X-ray and neutron scattering crystalline protein structures (resolution of <0.8 Å), and the precise position of the amide proton was investigated by using diverse $^{15}N-^{1}H^{N}$ bond lengths along the $^{15}N-^{1}H^{N}$ vectorial direction. Because every peptide plane is treated separately from the others, this approach is "structure-free"; i.e., the 3D protein fold is neither required nor constructed, which is in contrast to previous RDC-based applications.

The initial procedure determines the optimal level of alignment relevant to the dominant motional mode (S or 1D-GAF α, β, or γ), which is then refined using indirect analysis of "free data sets" using the full 3D-GAF description, resulting in slightly higher tensors. Analysis of free data sets that were not used in a full 3D-GAF description shows a clear minimum that determines the level of molecular alignment as precisely as possible (simulation suggests a level of uncertainty of ~1%).

$^{15}N-^{1}H^{N}$ order parameters determined using the optimal alignment tensors with those determined using spin relaxation (two independent analyses) are shown in Figure 3. This comparison indicates the presence of increased dynamics in the loop regions of ubiquitin over longer time scales, but actually quite similar amplitude excursions in the regions of secondary structure of the molecule. The large amplitude motions that are apparent in the turn region of the N-terminal β-hairpin (residues 8−12) are invisible to spin relaxation because of the overall tumbling of the molecule (in the range of 4−5 ns) and are therefore slower than this limit. Significant dynamics also occur in the region of the turn of residues 62−65; one of the weakest detected hydrogen bonds is determined from $^{h3}J_{NC'}$ scalar coupling measurements[82] in which specific thermal instability has been found.[83] Planes 64 and 65 also show the largest amplitude (30°) α-motions in the protein, suggesting that the dynamics in this turn may be correlated. Note that the distribution of motions found using the SFGAF approach is very similar to the distribution found using the model-free approach applied to the same data set, although the absolute level of the motion is lower here, possibly because of the scaling applied to the data derived with the model-free approach.[68]

It is important to note here that the only additional dynamics that could be missed using the SFGAF approach would be an equi-amplitude, axially symmetric component that would be common to all vectors in the peptide plane for each peptide plane. In view of the well-determined anisotropy of the peptide plane, the presence of such additional motions is difficult to envisage, and in the absence of any experimental evidence, there is no reason to invoke the existence of such motions. To further address the robustness of the approach, data from each alignment medium were removed for 24 separate analyses, and two $N-H^{N}$ RDCs were randomly removed from each peptide plane; in both cases, RDCs were back-calculated using static and 3D-GAF models, using appropriate alignment tensors for the specific models (Figure 4A). In both tests, the reduced χ^2 was significantly lower for the SFGAF analysis compared to the static model, again demonstrating that the dynamics determined from the GAF analysis are required to correctly reproduce the data. Finally, the precision of the structural and dynamic description of

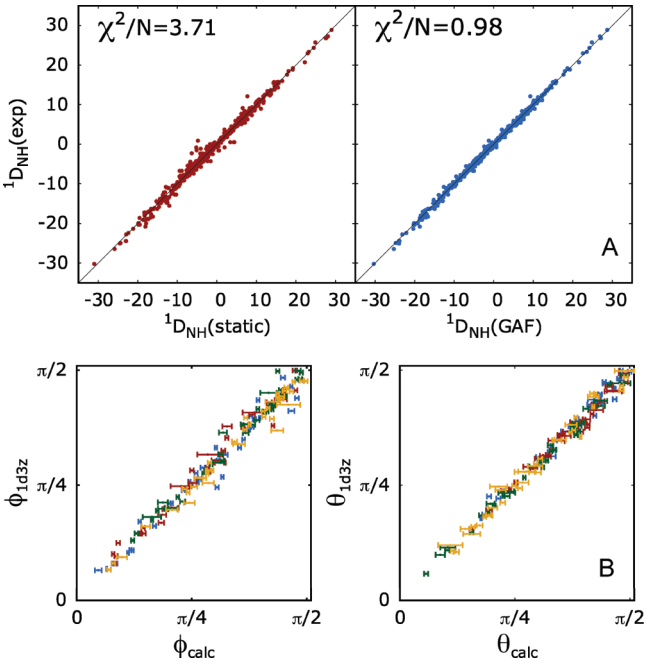


Figure 4. Statistical analysis of GAF interpretation. (A) Cross validation of the structure-free GAF analysis of local motions of ubiquitin. Two dipolar couplings were randomly removed from each peptide plane and not used in the analysis except for comparison. Fully optimized static and dynamic analyses are shown. (B) Comparison of the orientations of $N-H^{N}$ (blue), $C'-C^{\alpha}$ (red), $C'-H^{N}$ (green), and $C'-N$ (yellow) bonds in the high-resolution structure (Protein Data Bank entry 1d3z), compared to the mean orientations and their associated uncertainty from the SFGAF analysis. The error bars represent the distribution of values determined using Monte Carlo noise-based SFGAF simulations.

the SFGAF approach can be determined numerically by using Monte Carlo noise-based simulations, allowing the average orientations of internuclear bond vectors determined using this approach and their dispersion in angular space to be compared to those of the high-resolution NMR structure (Figure 4B). This is, to the best of our knowledge, the only case in which the intrinsic uncertainty in both the structure and the dynamics of the protein has been determined directly from experimental NMR data.

Recently proposed analytical studies have used related approaches to determine the dynamic averaging of experimental RDCs in extended RNA molecules, to develop a quantitative description of domain motions in these molecules. The methods used to model the consensus structural elements and the combination of average spherical harmonics and reduced Wigner rotation matrices used to describe the sampled orientational space are theoretically very similar to the GAF and model-free approaches described here for the study of protein motions.[84−86]

To summarize, RDCs from ubiquitin measured in multiple alignment media were analyzed in terms of local backbone dynamics using a structure-free GAF-based approach. This method relies on only experimental RDCs, providing absolute alignment tensor information and internal motional modes and amplitudes from experimental data alone, and yields significant improvement in the reproduction of passive data over a model supposing no dynamics.

To address the physical reality of these results, in terms of the energy and stability of the protein, we have attempted to invert the problem. Rather than extracting motional modes and

amplitudes directly from the experimental data, as in the model-free or GAF-dependent approaches described above, we used a state-of-the-art molecular dynamics technology, in the complete absence of experimental constraints, to develop physically viable conformational ensembles. In this way, a comparison can be made between methods that necessarily fit to experimental data and methods that rely uniquely on a physically reasonable description of the potential energy surface of the protein. Classical molecular dynamics (MD) simulations provide access to motions on time scales that can be compared to measured experimental NMR data, and the correspondence of the accessible time scales with respect to spin relaxation data has led to the widespread use of NMR in combination with MD for the understanding of fast (picosecond to nanosecond) motions in globular proteins.[87] Similar approaches should be possible for the simulation of RDCs, and it has indeed been shown that increasing the length of a trajectory of ubiquitin resulted in increasingly well reproduced RDCs.[88] Nevertheless, despite increasing computational power, trajectories are still normally restricted to time scales of hundreds of nanoseconds, and while microsecond-long trajectories have been achieved, millisecond trajectories (the theoretically appropriate time scale for RDCs) are still not viable. In addition, long MD simulations provide only a single trajectory in phase space and do not prevent the problem of statistical mechanical sampling.

A popular alternative to performing long simulations is to implement time- or ensemble-averaged restraints, thereby driving a multiple-copy molecular ensemble into a description that reproduces the conformationally averaged RDCs.[5,32,61−64] Although this method is efficient for identifying conformational ensembles in agreement with experimental data, it is well understood that the addition of an arbitrary pseudopotential to a finely tuned physical force field will be expected to perturb the simulated dynamics in a nonpredictable manner, making further analysis of the resulting trajectories uncertain. Normally, such ensembles are of limited size in terms of the number of molecular copies, such that the relative free energy weighting of each member of the ensemble should also be somehow determined, a problem to which there may be no unique solution. The potential energy landscape of a folded protein is likely to be characterized by a high level of structure, resulting in a complex distribution of populations in conformational space, so that to accurately reproduce any conformationally averaged NMR observable, it would be necessary to include an accurate population analysis.[4] One additional drawback of many approaches employing ensemble averaging has been the tendency to mix restraints that are averaged differently or over different time scales. Without knowledge of the time scale of characteristic correlation times of the individual interactions, their interpretation in terms of interatomic distances may lead to artifact. In general, the combination of parameters that average differently over the ensemble and over different time scales cannot provide a useful description of the true ensemble, unless these specific averaging parameters are taken into consideration. Comparison of molecular dynamics simulations in the absence of conformational restraints is therefore a more attractive prospect, as long as one can either simulate enough sufficiently long trajectories or somehow overcome the sampling problem using alternative sampling procedures.

■ ACCELERATED MOLECULAR DYNAMICS AND RDCS

To address these questions, an entirely restraint-free approach was developed that aimed to provide a self-consistent structural dynamic representation of protein conformational sampling using a combination of state-of-the-art MD simulation and experimental RDC and scalar coupling data.[89] To sample conformational space efficiently, the accelerated molecular dynamics (AMD) approach was used.[90] The AMD approach involves adding a continuous non-negative bias potential to the potential energy surface of the protein to raise and flatten the potential energy landscape, thereby enhancing the rate of escape between low-energy conformational substates. The extent of acceleration (i.e., how aggressively the sampling of conformational space is enhanced) is determined by only two parameters for the entire system: the choice of the energy below which acceleration is applied and the acceleration parameter itself. The appropriate level of acceleration, and therefore conformational space, is directly determined by comparing the reproduction of millisecond-averaged experimentally measured dipolar and scalar couplings to those predicted from the different ensembles. Both experimental scalar and dipolar couplings represent a population-weighted average over all conformers contributing to the measurement and can therefore be directly treated using the same ensemble average. The method is used to describe conformational dynamics occurring on time scales over many orders of magnitude in ubiquitin and validated against experimental spin relaxation and RDC data sensitive to the diverse time scales.

With an increase in the level of acceleration, the simulation probes more conformational space. After the conformational space had been reweighted to yield the correct canonical Boltzmann distribution, a clustering protocol was implemented to identify low-energy conformational substates. A series of short 3 ns standard MD simulations were then seeded from the clusters obtained via AMD, to further sample the low-energy substates. Resulting ensembles represent free energy-weighted trajectories, sampling the conformational space explored by the AMD trajectories at the relevant acceleration level. An explicit solvent description was used in all simulations. This method represents an efficient equivalent to performing numerous long time scale MD simulations.

The next step is to identify which ensembles can best reproduce the experimental RDC data. At each increasing acceleration level, an increasingly large amount of conformational space was sampled. Ideally, an optimal sampling on the time scale relevant to the RDC data should exist. However, for each molecular ensemble, the optimal alignment tensor for a given alignment medium must be calculated. Bruschweiler and co-workers introduced a simple application of the singular-value decomposition (SVD) approach, to directly compare experimental RDCs with vectorial sampling from a given trajectory, that automatically provides the optimal alignment tensor, for a given molecular ensemble, obviating the need for calibration against external references or rescaling of order parameters.[88] The analysis was performed for available $N-H^N$ RDCs in 23 different alignment media. Using the optimized alignment tensors, theoretical RDCs were calculated for each molecular ensemble associated with a given acceleration level and the agreement between experiment and theory was monitored using the trajectory-averaged cumulative Q factor. A similar protocol was performed to calculate scalar J couplings.

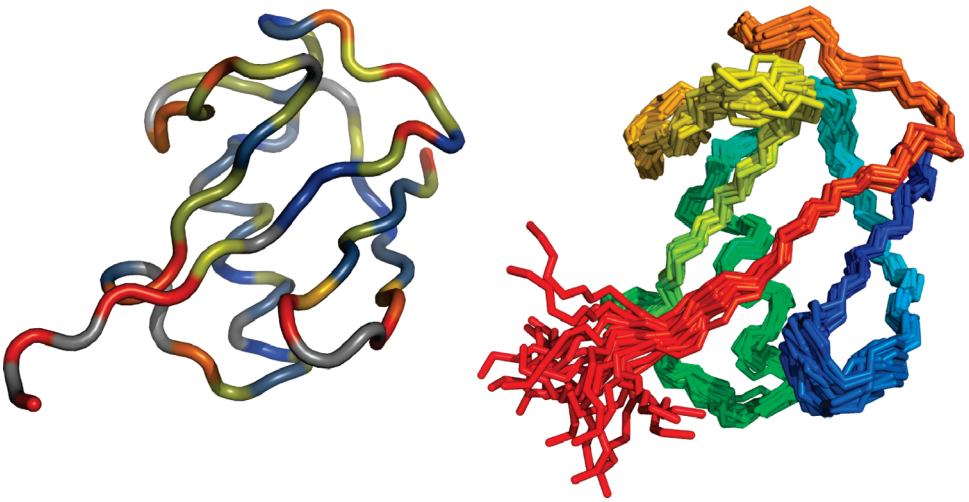


Figure 5. Ribbon diagram showing the distribution of amide vector S^2 values along the chain (left) as derived from the SFGAF analysis. Order parameters are represented from red (most mobile) to blue (least mobile). Sites at which order parameters were not determined because of a lack of data are colored gray. The right panel shows a representative AMD-derived ensemble from ubiquitin (colored from N- to C-terminus).

Numerous AMD simulations performed at different acceleration levels confirm that the method can efficiently and accurately sample conformational space in globular proteins. Comparison of the order parameters at different acceleration levels (Figure 3) reveals that even at the most aggressive acceleration level employed in this work, the observed enhanced conformational space sampling is restricted to well-defined regions of the protein, effectively surface loops involved in molecular recognition. The optimal AMD molecular ensemble (Figure 5) results in excellent reproduction of RDCs and scalar J couplings, reporting on averages up to the millisecond time scale. Importantly, it is also possible to separate time scales in this way and to compare to appropriate ensemble-averaged parameters: fast motional order parameters derived from spin relaxation data, reporting on picosecond to nanosecond motions, were calculated over all of the different short MD simulations performed at each acceleration level. The resulting agreement demonstrates that both fast and slow motions are described well by this ensemble and confirms the dependence of fast dynamics on the details of the local conformational substates of the protein backbone that are sampled on time scales slower than the correlation time of the protein. This observation was also made for protein GB3 in a similar study.[91]

Interestingly, the average backbone structure of the optimal molecular ensemble compares very closely with that of the experimentally refined 1d3z structure[92] (using RDCs form two alignment media and many other experimental restraints, including more than 2500 NOE restraints), indicating that although these ensembles sample more conformational space, they appear to be distributed about a mean that resembles the experimentally determined time- and ensemble-averaged structure. Principal component projections of the conformational sampling reveal that the restraint-free AMD/SVD approach described here samples essentially the same conformational space as a previously published ensemble, determined using ensemble-averaged restrained molecular dynamics simulation.[5] This remarkable result is all the more surprising considering that no local structural restraints are applied, and the conformational space is only defined via global agreement with the entire data set. The AMD approach further refines the available conformational space following free energy weighting, therefore providing a more realistic structural dynamic representation of the system. The free energy-weighted AMD ensemble also reproduces the experimental observables to a substantially greater degree of accuracy than a control set of 5 ns standard MD simulations and provides better results compared to the static X-ray crystal structure (PDB entry 1ubq).[93]

Using enhanced sampling from a biased potential molecular dynamics simulation, we are thus able to define a self-consistent representative molecular ensemble for solution state protein conformational dynamics. The theoretical incapacity of NOE-based approaches to accurately describe local motions without a precise knowledge of the associated time scales is avoided by simply not introducing these complex phenomena into the analysis. By contrast, RDCs provide an extremely powerful theoretical and analytical framework with which to address the true level of disorder present in biomolecules at equilibrium.[49,51,60,61,65,71,86,89,94−96] The problems of statistical mechanical sampling associated with the incorporation of additional terms into a hybrid potential energy force field and thereby perturbing the simulated dynamics are avoided by using only restraint-free trajectories seeded at different points of conformational space sampled by the accelerated MD. The accuracy of the resulting RDCs is hardly compromised by this procedure, with a similar level of reproduction compared to state-of-the-art single-structure or restrained-ensemble approaches. Most importantly, the level of acceleration, and therefore the extent of conformational sampling, is the only variable parameter in this entire procedure, and this is determined directly from the measured couplings, with no direct fitting to experimental data.

■ ANALYTICAL AND NUMERICAL APPROACHES CONVERGE TO A UNIFIED DESCRIPTION OF PROTEIN MOTION

We are now in a position to compare two entirely different approaches to the extraction of quantitative dynamic parameters from RDCs measured in solution. In one case, analytical expressions are used to determine, in a least-squares fitting approach, the level and nature of local dynamics along the backbone of the

protein. This method requires no structural model, is entirely independent of any physical assumptions except for the average geometry of each peptide plane, and determines the level of overall alignment, the local dynamics, and the average conformation, in a single algorithmic procedure. In the second approach, the experimental data are removed entirely, and recent advances in potential energy force field development and, more importantly, enhanced sampling of conformational space are exploited in the development of a statistical mechanical description of the rapidly exchanging structural ensemble that is responsible for the measured NMR spectrum. The vastly differing approaches presented here can be directly compared via, for example, their $N-H^N$ bond vector order parameters (Figure 3). The results are quite remarkable, showing a level of agreement concerning both the level and detail of the local motions that justified the care that was taken over the treatment of the quantitative level of dynamics in both cases, and strongly substantiate the results found in both cases. We note that this unified description of the backbone dynamics of ubiquitin is in contrast to previous attempts to reconcile ensemble descriptions of proteins and model-free approaches, where significant additional scaling parameters were necessary to provide agreement, such that the published ensemble did not represent the analytically determined level of dynamics.[5]

CONCLUSIONS

A detailed description of protein flexibility is essential for improving our understanding of the role that structural dynamics plays in controlling biological processes. Despite the high degree of sensitivity of NMR data to molecular motion, and possibly because of the very complexity of this dependence, the faithful incorporation of dynamic fluctuations into a structural model of biomolecules has proven to be an almost insurmountable task. The development of experimental and analytical methods for the measurement and interpretation of RDCs and their combination with state-of-the-art molecular dynamics-based approaches have provided a powerful arsenal of techniques for resolving this fundamental question.

Recently, purely analytical and purely simulation-based approaches have been applied in the determination of protein motions using RDCs. The agreement is excellent, providing a unified structural dynamic representation of the statistical mechanical properties of a protein in solution and describing dynamic fluctuations on time scales ranging over 6 orders of magnitude. In addition to describing the conformational fluctuations that as noted by Lange et al.[5] sample the functional states of the protein already in free solution, this ensemble description describes the origin of the high-resolution solution state NMR spectrum of the protein, and one can be only confident that this in itself will provide the basis for improving our understanding of molecular interaction, stability, and folding. So far, these techniques have been applied to only small, very well-behaved proteins that can be aligned in multiple different alignment media, providing essential biophysical information about intrinsic protein flexibility in solution. However, the recent demonstration that GAF-based techniques can be quantitatively applied using a restricted number of different alignment conditions[80] and that AMD-type approaches can almost certainly be used in combination with yet fewer data sets, signifies that RDCs are potentially extremely powerful tools for studying the local dynamics of larger systems and protein−protein complexes, examples of which are currently underway in our laboratory.

AUTHOR INFORMATION

Corresponding Author

*E-mail: martin.blackledge@ibs.fr. Telephone: +33 4 38789554. Fax: +33 4 38785494.

ABBREVIATIONS

NMR, nuclear magnetic resonance; RDC, residual dipolar coupling; MD, molecular dynamics; AMD, accelerated molecular dynamics; GAF, Gaussian axial fluctuation; NOE, nuclear Overhauser effect.

REFERENCES

(1) Russell, R. B., Alber, F., Aloy, P., Davis, F. P., Korkin, D., Pichaud, M., Topf, M., and Sali, A. (2004) A structural perspective on protein-protein interactions. *Curr. Opin. Struct. Biol. 14*, 313–324.

(2) Chandonia, J., and Brenner, S. (2006) The impact of structural genomics: Expectations and outcomes. *Science 311*, 347–351.

(3) Grant, B., Gorfe, A., and McCammon, J. (2010) Large conformational changes in proteins: Signaling and other functions. *Curr. Opin. Struct. Biol. 20*, 142–147.

(4) Frauenfelder, H., Sligar, S., and Wolynes, P. (1991) The Energy Landscapes and Motions of Proteins. *Science 254*, 1598–1603.

(5) Lange, O. F., Lakomek, N., Farès, C., Schröder, G. F., Walter, K. F. A., Becker, S., Meiler, J., Grubmüller, H., Griesinger, C., and de Groot, B. L. (2008) Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science 320*, 1471–1475.

(6) Cavalli, A., Salvatella, X., Dobson, C., and Vendruscolo, M. (2007) Protein structure determination from NMR chemical shifts. *Proc. Natl. Acad. Sci. U.S.A. 104*, 9615–9620.

(7) Shen, Y., Lange, O., Delaglio, F., Rossi, P., Aramini, J., Liu, G., Eletsky, A., Wu, Y., Singarapu, K., Lemak, A., Ignatchenko, A., Arrowsmith, C., Szyperski, T., Montelione, G., Baker, D., and Bax, A. (2008) Consistent blind protein structure generation from NMR chemical shift data. *Proc. Natl. Acad. Sci. U.S.A. 105*, 4685–4690.

(8) Tolman, J., and Ruan, K. (2006) NMR residual dipolar couplings as probes of biomolecular dynamics. *Chem. Rev. 106*, 1720–1736.

(9) Blackledge, M. (2005) Recent progress in the study of biomolecular structure and dynamics in solution from residual dipolar couplings. *Prog. NMR Spectrosc. 46*, 23–61.

(10) Bax, A., and Grishaev, A. (2005) Weak alignment NMR: A hawk-eyed view of biomolecular structure. *Curr. Opin. Struct. Biol. 15*, 563–570.

(11) Abragam, A. (1994) *The Principles of Nuclear Magnetism Reprint*, Clarendon Press, Gloucestershire, U.K.

(12) Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc. 104*, 4546–4559.

(13) Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J. Am. Chem. Soc. 104*, 4559–4570.

(14) Deverell, C., Morgan, R., and Strange, J. (1970) Studies of chemical exchange by nuclear magnetic resonance relaxation in the rotating frame. *Mol. Phys. 18*, 553–573.

(15) Bleich, H., and Glasel, J. (1978) Rotating frame spin-lattice relaxation experiments and the problem of intra-molecular motions of peptides in solution. *Biopolymers 17*, 2445–2457.

(16) Kopple, K., Bhandary, K., Kartha, G., Wang, Y., and Parameswaran, K. (1986) Conformations of cyclic octapeptides. 3. Cyclo-(D-Ala-Gly-Pro-Phe)$_2$: Conformations in crystals and a $T_{1\rho}$ examination of internal mobility in solution. *J. Am. Chem. Soc. 108*, 4637–4642.

(17) Kopple, K., Wang, Y., Cheng, A., and Bhandary, K. (1988) Conformations of cyclic octapeptides. 5. Crystal-structure of Cyclo-(Cys-Gly-Pro-Phe)$_2$ and rotating frame relaxation NMR-studies of internal mobility in cyclic octapeptides. *J. Am. Chem. Soc. 110*, 4168–4176.

(18) Blackledge, M. J., Brüschweiler, R., Griesinger, C., Schmidt, J. M., Xu, P., and Ernst, R. R. (1993) Conformational backbone dynamics of the cyclic decapeptide antamanide. Application of a new multiconformational search algorithm based on NMR data. *Biochemistry 32*, 10960–10974.

(19) Lane, A., Bauer, C., and Frenkiel, T. (1993) Determination of conformational transition rates in the Trp promoter by H-1-NMR rotating frame and cross-relaxation rate measurements. *Eur. Biophys. J. 21*, 425–431.

(20) Palmer, A. (2004) NMR characterization of the dynamics of biomacromolecules. *Chem. Rev. 104*, 3623–3640.

(21) Tollinger, M., Skrynnikov, N., Mulder, F., Forman-Kay, J., and Kay, L. (2001) Slow dynamics in folded and unfolded states of an SH3 domain. *J. Am. Chem. Soc. 123*, 11341–11352.

(22) Sprangers, R., and Kay, L. E. (2007) Probing supramolecular structure from measurement of methyl ^{1}H-^{13}C residual dipolar couplings. *J. Am. Chem. Soc. 129*, 12668–12669.

(23) Mittermaier, A., and Kay, L. E. (2006) New tools provide new insights in NMR studies of protein dynamics. *Science 312*, 224–228.

(24) Henzler-Wildman, K., Lei, M., Thai, V., Kerns, S., Karplus, M., and Kern, D. (2007) A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature 450*, 913–916.

(25) Hansen, D. F., Vallurupalli, P., Lundström, P., Neudecker, P., and Kay, L. E. (2008) Probing chemical shifts of invisible states of proteins with relaxation dispersion NMR spectroscopy: How well can we do?. *J. Am. Chem. Soc. 130*, 2667–2675.

(26) Korzhnev, D. M., Religa, T. L., Banachewicz, W., Fersht, A. R., and Kay, L. E. (2010) A transient and low-populated protein-folding intermediate at atomic resolution. *Science 329*, 1312–1316.

(27) Macura, S., and Ernst, R. (1980) Elucidation of cross-relaxation in liquids by two-dimensional NMR spectroscopy. *Mol. Phys. 41*, 95–117.

(28) Neuhaus, D., and Williamson, M. P. (2000) *The Nuclear Overhauser Effect*, John Wiley & Sons, New York.

(29) Wüthrich, K. (1986) *Nuclear Magnetic Resonance of Proteins and Nucleic Acids*, Wiley-Blackwell, New York.

(30) Brüschweiler, R., Roux, B., Blackledge, M., Griesinger, C., Karplus, M., and Ernst, R. (1992) Influence of rapid intramolecular motion on NMR cross-relaxation rates: A moleuclar dynamics study of Antamanide in solution. *J. Am. Chem. Soc. 114*, 2289–2302.

(31) Bonvin, A., Rullmann, J., Lamerichs, R., Boelens, R., and Kaptain, R. (1993) Ensemble iterative relaxation matrix approach: A new NMR refinement protocol applied to the solution structure of Crambin. *Proteins 15*, 385–400.

(32) Lindorff-Larsen, K., Best, R., DePristo, M., Dobson, C., and Vendruscolo, M. (2005) Simultaneous determination of protein structure and dynamics. *Nature 433*, 128–132.

(33) Brüschweiler, R., Blackledge, M., and Ernst, R. R. (1991) Multi-conformational peptide dynamics derived from NMR data: A new search algorithm and its application to antamanide. *J. Biomol. NMR 1*, 3–11.

(34) Rieping, W., Habeck, M., and Nilges, M. (2005) Inferential structure determination. *Science 309*, 303–306.

(35) Tjandra, N., and Bax, A. (1997) Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science 278*, 1111–1114.

(36) Prestegard, J. H., al-Hashimi, H. M., and Tolman, J. R. (2000) NMR structures of biomolecules using field oriented media and residual dipolar couplings. *Q. Rev. Biophys. 33*, 371–424.

(37) Hus, J., Salmon, L., Bouvignies, G., Lotze, J., Blackledge, M., and Brüschweiler, R. (2008) 16-fold degeneracy of peptide plane orientations from residual dipolar couplings: Analytical treatment and implications for protein structure determination. *J. Am. Chem. Soc. 130*, 15927–15937.

(38) Al-Hashimi, H. M., Valafar, H., Terrell, M., Zartler, E. R., Eidsness, M. K., and Prestegard, J. H. (2000) Variation of molecular alignment as a means of resolving orientational ambiguities in protein structures from dipolar couplings. *J. Magn. Reson. 143*, 402–406.

(39) Hus, J. C., Marion, D., and Blackledge, M. (2000) De novo determination of protein structure by NMR using orientational and long-range order restraints. *J. Mol. Biol. 298*, 927–936.

(40) Béraud, S., Bersch, B., Brutscher, B., Gans, P., Barras, F., and Blackledge, M. (2002) Direct structure determination using residual dipolar couplings: Reaction-site conformation of methionine sulfoxide reductase in solution. *J. Am. Chem. Soc. 124*, 13709–13715.

(41) Tolman, J. R., Flanagan, J. M., Kennedy, M. A., and Prestegard, J. H. (1997) NMR evidence for slow collective motions in cyanometmyoglobin. *Nat. Struct. Biol. 4*, 292–297.

(42) Fischer, M. W., Losonczi, J. A., Weaver, J. L., and Prestegard, J. H. (1999) Domain orientation and dynamics in multidomain proteins from residual dipolar couplings. *Biochemistry 38*, 9013–9022.

(43) Braddock, D. T., Cai, M., Baber, J. L., Huang, Y., and Clore, G. M. (2001) Rapid identification of medium- to large-scale interdomain motion in modular proteins using dipolar couplings. *J. Am. Chem. Soc. 123*, 8634–8635.

(44) Skrynnikov, N. R., Goto, N. K., Yang, D., Choy, W. Y., Tolman, J. R., Mueller, G. A., and Kay, L. E. (2000) Orienting domains in proteins using dipolar couplings measured by liquid-state NMR: Differences in solution and crystal forms of maltodextrin binding protein loaded with β-cyclodextrin. *J. Mol. Biol. 295*, 1265–1273.

(45) Ulmer, T. S., Soelaiman, S., Li, S., Klee, C. B., Tang, W., and Bax, A. (2003) Calcium dependence of the interaction between calmodulin and anthrax edema factor. *J. Biol. Chem. 278*, 29261–29266.

(46) Sibille, N., Pardi, A., Simorre, J. P., and Blackledge, M. (2001) Refinement of local and long-range structural order in theophylline-binding RNA using ^{13}C-^{1}H residual dipolar couplings and restrained molecular dynamics. *J. Am. Chem. Soc. 123*, 12135–12146.

(47) Al-Hashimi, H. M., Gosser, Y., Gorin, A., Hu, W., Majumdar, A., and Patel, D. J. (2002) Concerted motions in HIV-1 TAR RNA may allow access to bound state conformations: RNA dynamics from NMR residual dipolar couplings. *J. Mol. Biol. 315*, 95–102.

(48) Zhang, Q., Sun, X., Watt, E., and Al-Hashimi, H. (2006) Resolving the motional modes that code for RNA adaptation. *Science 311*, 653–656.

(49) Zhang, Q., Stelzer, A. C., Fisher, C. K., and Al-Hashimi, H. M. (2007) Visualizing spatially correlated dynamics that directs RNA conformational transitions. *Nature 450*, 1263–1267.

(50) Tian, F., Al-Hashimi, H. M., Craighead, J. L., and Prestegard, J. H. (2001) Conformational analysis of a flexible oligosaccharide using residual dipolar couplings. *J. Am. Chem. Soc. 123*, 485–492.

(51) Tolman, J. (2002) A novel approach to the retrieval of structural and dynamic information from residual dipolar couplings using several oriented media in biomolecular NMR spectroscopy. *J. Am. Chem. Soc. 124*, 12020–12030.

(52) Peti, W., Meiler, J., Brüschweiler, R., and Griesinger, C. (2002) Model-free analysis of protein backbone motion from residual dipolar couplings. *J. Am. Chem. Soc. 124*, 5822–5833.

(53) Tolman, J. R., Al-Hashimi, H. M., Kay, L. E., and Prestegard, J. H. (2001) Structural and dynamic analysis of residual dipolar coupling data for proteins. *J. Am. Chem. Soc. 123*, 1416–1424.

(54) Bernadó, P., and Blackledge, M. (2004) Anisotropic small amplitude peptide plane dynamics in proteins from residual dipolar couplings. *J. Am. Chem. Soc. 126*, 4907–4920.

(55) Meiler, J., Prompers, J. J., Peti, W., Griesinger, C., and Brüschweiler, R. (2001) Model-free approach to the dynamic interpretation of residual dipolar couplings in globular proteins. *J. Am. Chem. Soc. 123*, 6098–6107.

(56) Briggman, K., and Tolman, J. (2003) De Novo determination of bond orientations and order parameters from residual dipolar couplings with high accuracy. *J. Am. Chem. Soc. 125*, 10164–10165.

(57) Bernadó, P., and Blackledge, M. (2004) Local dynamic amplitudes on the protein backbone from dipolar couplings: Toward the

elucidation of slower motions in biomolecules. *J. Am. Chem. Soc. 126*, 7760–7761.

(58) Bouvignies, G., Meier, S., Grzesiek, S., and Blackledge, M. (2006) Ultrahigh-resolution backbone structure of perdeuterated protein GB1 using residual dipolar couplings from two alignment media. *Angew. Chem., Int. Ed. 45*, 8166–8169.

(59) Yao, L., Vögeli, B., Torchia, D. A., and Bax, A. (2008) Simultaneous NMR study of protein structure and dynamics using conservative mutagenesis. *J. Phys. Chem. B 112*, 6045–6056.

(60) Bouvignies, G., Markwick, P., Brüschweiler, R., and Blackledge, M. (2006) Simultaneous determination of protein backbone structure and dynamics from residual dipolar couplings. *J. Am. Chem. Soc. 128*, 15100–15101.

(61) Clore, G. M., and Schwieters, C. D. (2004) How much backbone motion in ubiquitin is required to account for dipolar coupling data measured in multiple alignment media as assessed by independent cross-validation?. *J. Am. Chem. Soc. 126*, 2923–2938.

(62) Clore, G. M., and Schwieters, C. D. (2004) Amplitudes of protein backbone dynamics and correlated motions in a small α/β protein: Correspondence of dipolar coupling and heteronuclear relaxation measurements. *Biochemistry 43*, 10678–10691.

(63) De Simone, A., Richter, B., Salvatella, X., and Vendruscolo, M. (2009) Toward an accurate determination of free energy landscapes in solution states of proteins. *J. Am. Chem. Soc. 131*, 3810–3811.

(64) Hess, B., and Scheek, R. M. (2003) Orientation restraints in molecular dynamics simulations using time and ensemble averaging. *J. Magn. Reson. 164*, 19–27.

(65) Meiler, J., Prompers, J. J., Peti, W., Griesinger, C., and Brüschweiler, R. (2001) Model-free approach to the dynamic interpretation of residual dipolar couplings in globular proteins. *J. Am. Chem. Soc. 123*, 6098–6107.

(66) Tsan, P., Hus, J., Caffrey, M., Marion, D., and Blackledge, M. (2000) Rotational diffusion anisotropy and local backbone dynamics of carbon monoxide-bound *Rhodobacter capsulatus* cytochrome c. *J. Am. Chem. Soc. 122*, 5603–5612.

(67) Lakomek, N., Lange, O. F., Walter, K. F. A., Farès, C., Egger, D., Lunkenheimer, P., Meiler, J., Grubmüller, H., Becker, S., de Groot, B. L., and Griesinger, C. (2008) Residual dipolar couplings as a tool to study molecular recognition of ubiquitin. *Biochem. Soc. Trans. 36*, 1433–1437.

(68) Lakomek, N., Walter, K. F. A., Farès, C., Lange, O. F., de Groot, B. L., Grubmüller, H., Brüschweiler, R., Munk, A., Becker, S., Meiler, J., and Griesinger, C. (2008) Self-consistent residual dipolar coupling based model-free analysis for the robust determination of nanosecond to microsecond protein dynamics. *J. Biomol. NMR 41*, 139–155.

(69) Lakomek, N. A., Carlomagno, T., Becker, S., Griesinger, C., and Meiler, J. (2006) A thorough dynamic interpretation of residual dipolar couplings in ubiquitin. *J. Biomol. NMR 34*, 101–115.

(70) Meiler, J., Peti, W., and Griesinger, C. (2003) Dipolar couplings in multiple alignments suggest α helical motion in ubiquitin. *J. Am. Chem. Soc. 125*, 8072–8073.

(71) Hus, J., Peti, W., Griesinger, C., and Brüschweiler, R. (2003) Self-consistency analysis of dipolar couplings in multiple alignments of ubiquitin. *J. Am. Chem. Soc. 125*, 5596–5597.

(72) Lienin, S. F., and Brüschweiler, R. (2000) Characterization of collective and anisotropic reorientational protein dynamics. *Phys. Rev. Lett. 84*, 5439–5442.

(73) Bremi, T., and Bruschweiler, R. (1997) Locally anisotropic internal polypeptide backbone dynamics by NMR relaxation. *J. Am. Chem. Soc. 119*, 6672–6673.

(74) Bruschweiler, R., and Wright, P. (1994) NMR order parameters of biomolecules: A new analytical representation and application to the Gaussian Axial Fluctuation model. *J. Am. Chem. Soc. 116*, 8426–8427.

(75) Ulmer, T. S., Ramirez, B. E., Delaglio, F., and Bax, A. (2003) Evaluation of backbone proton positions and dynamics in a small protein by liquid crystal NMR spectroscopy. *J. Am. Chem. Soc. 125*, 9179–9191.

(76) Bouvignies, G., Bernadó, P., Meier, S., Cho, K., Grzesiek, S., Brüschweiler, R., and Blackledge, M. (2005) Identification of slow correlated motions in proteins using residual dipolar and hydrogen-bond scalar couplings. *Proc. Natl. Acad. Sci. U.S.A. 102*, 13885–13890.

(77) Davis, I. W., Arendall, W. B., Richardson, D. C., and Richardson, J. S. (2006) The backrub motion: How protein backbone shrugs when a sidechain dances. *Structure 14*, 265–274.

(78) Derrick, J. P., and Wigley, D. B. (1994) The third IgG-binding domain from streptococcal protein G. An analysis by X-ray crystallography of the structure alone and in a complex with Fab. *J. Mol. Biol. 243*, 906–918.

(79) Lewandowski, J. R., Sein, J., Sass, H. J., Grzesiek, S., Blackledge, M., and Emsley, L. (2010) Measurement of site-specific ^{13}C spin-lattice relaxation in a crystalline protein. *J. Am. Chem. Soc. 132*, 8252–8254.

(80) Bouvignies, G., Markwick, P. R. L., and Blackledge, M. (2008) Characterization of protein dynamics from residual dipolar couplings using the three dimensional Gaussian axial fluctuation model. *Proteins 71*, 353–363.

(81) Salmon, L., Bouvignies, G., Markwick, P., Lakomek, N., Showalter, S., Li, D., Walter, K., Griesinger, C., Brüschweiler, R., and Blackledge, M. (2009) Protein conformational flexibility from structure-free analysis of NMR dipolar couplings: quantitative and absolute determination of backbone motion in ubiquitin. *Angew. Chem., Int. Ed. 48*, 4154–4157.

(82) Cordier, F., and Grzesiek, S. (2004) Quantitative comparison of the hydrogen bond network of A-state and native ubiquitin by hydrogen bond scalar couplings. *Biochemistry 43*, 11295–11301.

(83) Cordier, F., and Grzesiek, S. (2002) Temperature-dependence of protein hydrogen bond properties as studied by high-resolution NMR. *J. Mol. Biol. 317*, 739–752.

(84) Fisher, C. K., and Al-Hashimi, H. M. (2009) Approximate reconstruction of continuous spatially complex domain motions by multialignment NMR residual dipolar couplings. *J. Phys. Chem. B 113*, 6173–6176.

(85) Fisher, C. K., Zhang, Q., Stelzer, A., and Al-Hashimi, H. M. (2008) Ultrahigh resolution characterization of domain motions and correlations by multialignment and multireference residual dipolar coupling NMR. *J. Phys. Chem. B 112*, 16815–16822.

(86) Zhang, Q., and Al-Hashimi, H. M. (2008) Extending the NMR spatial resolution limit for RNA by motional couplings. *Nat. Methods 5*, 243–245.

(87) Brüschweiler, R. (2003) New approaches to the dynamic interpretation and prediction of NMR relaxation data from proteins. *Curr. Opin. Struct. Biol. 13*, 175–183.

(88) Showalter, S. A., and Brüschweiler, R. (2007) Quantitative molecular ensemble interpretation of NMR dipolar couplings without restraints. *J. Am. Chem. Soc. 129*, 4158–4159.

(89) Markwick, P. R. L., Bouvignies, G., Salmon, L., McCammon, J. A., Nilges, M., and Blackledge, M. (2009) Toward a unified representation of protein structural dynamics in solution. *J. Am. Chem. Soc. 131*, 16968–16975.

(90) Hamelberg, D., Mongan, J., and McCammon, J. A. (2004) Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules. *J. Chem. Phys. 120*, 11919–11929.

(91) Markwick, P. R. L., Bouvignies, G., and Blackledge, M. (2007) Exploring Multiple Timescale Motions in Protein GB3 using Accelerated Molecular Dynamics and NMR. *J. Am. Chem. Soc. 129*, 4724–4730.

(92) Cornilescu, G., Marquardt, J., Ottiger, M., and Bax, A. (1998) Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *J. Am. Chem. Soc. 120*, 6836–6837.

(93) Vijay-Kumar, S., Bugg, C. E., Wilkinson, K. D., Vierstra, R. D., Hatfield, P. M., and Cook, W. J. (1987) Comparison of the three-dimensional structures of human, yeast, and oat ubiquitin. *J. Biol. Chem. 262*, 6396–6399.

(94) Jensen, M. R., Markwick, P. R., Meier, S., Griesinger, C., Zweckstetter, M., Grzesiek, S., Bernadó, P., and Blackledge, M. (2009) Quantitative determination of the conformational properties of partially folded and intrinsically disordered proteins using NMR dipolar couplings. *Structure 17*, 1169–1185.

(95) Nodet, G., Salmon, L., Ozenne, V., Meier, S., Jensen, M. R., and Blackledge, M. (2009) Quantitative description of backbone conformational sampling of unfolded proteins at amino acid resolution from NMR residual dipolar couplings. *J. Am. Chem. Soc. 131*, 17908–17918.

(96) Jensen, M. R., Houben, K., Lescop, E., Blanchard, L., Ruigrok, R. W., and Blackledge, M. (2008) Quantitative conformational analysis of partially folded proteins from residual dipolar couplings: Application to the molecular recognition element of Sendai virus nucleoprotein. *J. Am. Chem. Soc. 130*, 8055–8061.