ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/26806569

Quantitative Determination of the Conformational Properties of Partially Folded and Intrinsically Disordered Proteins Using NMR Dipolar Couplings

ARTICLE *in* STRUCTURE · SEPTEMBER 2009

Impact Factor: 5.62 · DOI: 10.1016/j.str.2009.08.001 · Source: PubMed

CITATIONS
102

READS
48

8 AUTHORS, INCLUDING:

Malene Ringkjobing Jensen
French National Centre for Scientific Research
77 PUBLICATIONS **1,885** CITATIONS
SEE PROFILE

Pau Bernado
French Institute of Health and Medical Rese…
79 PUBLICATIONS **2,869** CITATIONS
SEE PROFILE

All in-text references underlined in blue are linked to publications on ResearchGate, letting you access and read them immediately.

Available from: Pau Bernado
Retrieved on: 23 January 2016

Structure
Review

Quantitative Determination of the Conformational Properties of Partially Folded and Intrinsically Disordered Proteins Using NMR Dipolar Couplings

Malene Ringkjøbing Jensen,[1] Phineus R.L. Markwick,[1,5] Sebastian Meier,[2,6] Christian Griesinger,[3] Markus Zweckstetter,[3] Stephan Grzesiek,[2] Pau Bernadó,[1,4] and Martin Blackledge[1,*]
[1]Protein Dynamics and Flexibility, Institut de Biologie Structurale, UMR 5075 CEA-CNRS-UJF, 41 Rue Jules Horowitz, Grenoble 38027, France
[2]Department of NMR Based Structural Biology, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077, Goettingen, Germany
[3]Biozentrum, University of Basel, Klingelbergstrasse 50, 4056 Basel, Switzerland
[4]Present address: Institute for Research in Biomedicine. c/ Baldiri Reixac, 10. 08028-Barcelona, Spain
[5]Present address: University of California San Diego, Department of Chemistry and Biochemistry, 9500 Gilman Drive, Urey Hall, La Jolla, CA 92003-0365, USA
[6]Present address: Carlsberg Laboratory, Gamle Carlsberg Vej 10, 2500 Valby, Denmark
*Correspondence: martin.blackledge@ibs.fr
DOI 10.1016/j.str.2009.08.001

Intrinsically disordered proteins (IDPs) inhabit a conformational landscape that is too complex to be described by classical structural biology, posing an entirely new set of questions concerning the molecular understanding of functional biology. The characterization of the conformational properties of IDPs, and the elucidation of the role they play in molecular function, is therefore one of the major challenges remaining for modern structural biology. NMR is the technique of choice for studying this class of proteins, providing information about structure, flexibility, and interactions at atomic resolution even in completely disordered states. In particular, residual dipolar couplings (RDCs) have been shown to be uniquely sensitive and powerful tools for characterizing local and long-range structural behavior in disordered proteins. In this review we describe recent applications of RDCs to quantitatively describe the level of local structure and transient long-range order in IDPs involved in viral replication, neurodegenerative disease, and cancer.

Introduction

Over the last decade it has become evident that a significant fraction of proteins—over 40% of the human proteome—are not folded in their functional form (Uversky, 2002; Tompa, 2002; Fink, 2005). These intrinsically disordered proteins (IDPs) or regions (IDRs) have been shown to play key roles in a remarkable range of cellular processes, including signaling, cell cycle control, molecular recognition, transcription, and replication, as well as in the development of numerous human pathologies such as neurodegenerative disease and cancer, where this figure rises to 80%. Bioinformatics studies have identified 238 mostly regulatory and signaling functions from SWISS-PROTein that are likely associated with IDPs and IDRs as compared with 302 mostly enzymatic and transport functions associated with structured proteins. In addition, the functional classes associated with IDPs have been proposed to span a wider range of biological processes than the classes associated with structured proteins (Xie et al., 2007). IDPs necessarily fall outside the realm of classical structural biology due to their extreme structural flexibility, and as such pose an entirely new set of questions concerning the molecular understanding of functional biology. The traditional structure-function relationship, applied to stable, folded proteins, focuses on the precise determination of structural features in partner molecules, and underpins numerous successful structural genomics and proteomics programs (Aloy and Russell, 2004). A parallel and fundamentally different paradigm exists, however, wherein one or both of the partner proteins are not folded in the free form (Dyson and Wright, 2002; Fuxreiter et al. 2004). Many IDPs fold only upon binding, and the relationship between intrinsic conformational propensity and the structure adopted by the protein in its bound form represents a novel paradigm that adds an additional dimension to the characterization of protein interactions and their relation to function (Aloy and Russell, 2004; Dyson and Wright, 2002; Fuxreiter et al. 2004; Vacic et al., 2007; Vucetic et al., 2005). IDPs might even remain flexible in the bound form of the complex (Tompa and Fuxreiter, 2008). For these reasons, the development of novel methods to characterize the conformational behavior of IDPs, and eventually to solve important biological and medical questions involving these proteins, is an essential and extremely active field of research.

The determination of a single set of three-dimensional atomic coordinates, even if feasible, would have little meaning for a highly disordered protein. Rather the aim of a conformational description of IDPs must be to identify rules that define the conformational behavior of the chain in terms of probability, or more often, in terms of an explicit ensemble description of interconverting structures. In this and many other respects, the study of IDPs is closely related to the study of chemically or thermally denatured proteins, although subtle differences have become apparent concerning the physical behavior of denatured and natively disordered proteins. The conformational space available to IDPs is vast, and the mapping of this complex conformational energy landscape necessarily relies on the exploitation of complementary experimental techniques reporting on both short-range and long-range structural parameters. An additional level of complication is introduced by the dynamic nature of these conformational ensembles. The timescales that are involved in the interconversion of the members of an ensemble also need to be determined to develop a complete picture.

IDPs can be characterized using a range of complementary spectroscopic techniques, including circular dichroism (Uversky, 2002) and Raman (Syme et al., 2002; Maiti et al., 2004) and infrared spectroscopy (Denning et al., 2002), that report on local structural propensities averaged over the whole molecule. Although these techniques can be instructive in detecting overall tendencies, the disadvantage is that short stretches or low populations of local structure can be difficult to detect. The presence of long-range interactions between distant parts of the chain will affect the overall dimensions of the protein, and can be probed using techniques that are sensitive to the size of the chain, for example size-exclusion chromatography, dynamic light, X-ray or neutron scattering (Millett et al., 2002; Bernadó et al., 2007), or fluorescence correlation spectroscopy (Jeganathan et al., 2006; Schuler and Eaton, 2008). The identification of IDPs on the basis of all available experimental techniques has allowed for the development of powerful bioinformatics tools to predict the level of disorder on the basis of primary sequence (Dunker et al., 2008). These tools have recently been extended to encompass the prediction of protein folding upon interaction (Fuxreiter et al., 2004).

Nuclear magnetic resonance (NMR) spectroscopy reports on both local and long-range conformational behavior at atomic resolution on timescales varying over many orders of magnitude, and as such is probably the most powerful biophysical tool for studying IDPs (Dyson and Wright, 2004). The dynamic averaging properties of NMR observables are well understood, rendering their exploitation particularly appropriate for the development of an ensemble description of the unfolded state. Importantly the local motional properties of IDPs in solution allow for the use of multidimensional NMR experiments that compensate for the comparative spectral crowding experienced in the amide region of the proton spectrum, allowing assignment of ^{1}H, ^{15}N, and ^{13}C resonances from throughout the protein. In this respect the recent complete backbone resonance assignment of the full-length tau protein (440 amino acids) represents an inspiring demonstration of the power of NMR to study even the most intimidating members of the IDP family (Mukrasch et al., 2009). In this review we will describe recent advances in the study of highly disordered proteins using NMR spectroscopy, in combination with other biophysical techniques such as small-angle scattering, to develop explicit ensemble descriptions that can be used to understand the conformational behavior of unfolded proteins. In particular we will describe some recent applications of the use of residual dipolar couplings (RDCs) to quantitatively describe the level of local structure and transient long-range order in both intrinsically disordered and chemically denatured proteins.

NMR Spectroscopy of Intrinsically Disordered Proteins

Over the last 15 years, NMR spectroscopy has developed into a key technique for studying highly flexible systems, and in doing so has furnished a remarkable amount of important information on the unfolded state (Neri et al., 1992; Alexandrescu et al., 1994; Shortle, 1996; Schwalbe et al., 1997). Even the simplest measurement, such as chemical shift, depends on a population-weighted average over rapidly exchanging local conformations sampled by all molecules in the ensemble, on timescales up to the millisecond. Chemical shifts report essentially on the local physico-chemical environment of the nucleus of interest (Spera and Bax, 1991; Wishart et al., 1992), a characteristic that has contributed significantly to the success of chemical-shift-based structure determination approaches (Cavalli et al., 2007; Shen et al., 2008). Not surprisingly then, average chemical shifts measured from a broad conformational equilibrium can be interpreted in terms of local conformational propensity of the ensemble. Once an amino-acid-specific "random coil" shift, normally calibrated from short unstructured peptides, has been subtracted from the measured value, the so-called secondary chemical shift clearly identifies the presence of transient structure in flexible chains (Wishart et al., 1995; Schwarzinger et al., 2001; Wang and Jardetzky, 2002). For example, in the case of $^{\alpha}$C spins, successive positive secondary shifts can be interpreted in terms of populations of α-helical segments. One potential problem associated with this kind of approach concerns incorrect frequency referencing, which can result in systematic errors on the secondary shifts. In order to address this problem, the $^{\alpha}$C and $^{\beta}$C chemical shifts (that shift in opposing directions for α-helical segments) can be used simultaneously to estimate the level of secondary structure in disordered proteins (Wang et al., 2005; Marsh et al., 2006).

Three-bond scalar couplings that depend on backbone dihedral angles (Serrano, 1995; Smith et al., 1996) also represent a population-weighted average that can be interpreted in terms of conformational propensity. Here again random coil values have been measured in small peptides, and these can be compared to experimental values to identify transient local structure. In this case calibration depends on Karplus-type relationships that have been used to parameterize the analytical dependence on dihedral angle. One potential source of error concerns the influence of the rest of the chain, or at least near-neighbors, on the conformational preferences of the amino acid of interest. This so-called persistence length, beyond which the remainder of the chain can be considered to exert a negligible effect, concerns all approaches that are based on the interpretation of experimental measurements made in intact proteins in comparison to short peptides. The importance of the persistence length depends to an extent on the measured parameter and might vary over the protein, depending on local primary sequence. In order to study the importance of such affects, Schwalbe and coworkers recently used scalar coupling measurements to detect differences between the conformational sampling in short peptides in the context of a longer chain, and in isolation (Graf et al., 2007). In theory, more detailed information about local conformational sampling can be derived from interproton nuclear Overhauser enhancements (nOe) (Macura and Ernst, 1980). However, quantitative interpretation of nOe is complicated by the strong sensitivity of the interaction on the range of dynamic timescales commonly encountered in unfolded proteins. It is therefore difficult to extract precise information about the distance distribution function from these measurements. Although intrinsically more complex in terms of the physical averaging processes, ^{15}N spin relaxation has also been shown to provide information that correlates with local order (Klein-Seetharaman et al., 2002) provided it occurs on the microsecond to millisecond time scale.

Long-range information is more commonly extracted from dipolar relaxation between an unpaired electron, normally present due to an artificially introduced nitroxide group, and the observed spin (commonly termed paramagnetic relaxation enhancement or PRE) (Gillespie and Shortle, 1997). These measurements offer clear advantages over nOe measurements, because they rely on stronger interactions (the "gyromagnetic ratio of the electron spin is 660 times larger than that of the proton and enters quadratically in all formulae describing relaxation), and therefore provide longer-range information about distance distribution functions, or information about weaker populated transient contacts. The experimental data have either been interpreted in terms of average distance restraints between the unpaired electron and the observed spin, and then incorporated directly into a restrained molecular dynamics approach (Dedmon et al., 2005; Bertoncini et al., 2005a), or more recently in terms of probability distributions (Felitsky et al., 2008). Disadvantages of this kind of restraint include the production of the necessary number of cysteine-carrying mutants of the protein, and the possible influence that the non-native moiety might have on native long-range contacts. Nevertheless these measurements are extremely powerful, because they provide unambiguous evidence of the presence of fluctuating tertiary structure that can be very difficult to identify by any other technique.

Residual dipolar couplings, measured between pairs of nuclei in partially aligned proteins, are very sensitive probes of time and ensemble-averaged conformational equilibria exchanging on timescales up to the millisecond and can therefore be used to characterize both the structure and dynamics of unfolded proteins (Shortle and Ackerman, 2001; Louhivuori et al., 2003). Over the last ten years remarkable progress has been made in developing a clearer understanding of the nature of RDCs in the unfolded state, either using analytical random chain descriptions derived from polymer physics, or using explicit conformational ensemble descriptions of the protein. Below we will describe some recent results that demonstrate the extraordinary power of RDCs to describe the conformational behavior of intrinsically disordered proteins, and to correlate this to their function.

Residual Dipolar Couplings

Dipolar couplings between two spins i and j depend on the geometry of the internuclear spin vector as follows (Emsley and Lindon, 1975):

$$D_{ij} = -\frac{\gamma_i \gamma_j \hbar \mu_0}{4\pi^2 r^3}\left\langle \frac{(3\cos^2\theta(t) - 1)}{2} \right\rangle = D_{\max}\langle P_2(\cos\theta(t))\rangle \qquad (1)$$

with

$$D_{\max} = -\frac{\gamma_i \gamma_j \hbar \mu_0}{4\pi^2 r^3} \qquad (2)$$

where θ is the angle of the internuclear vector relative to the static magnetic field. r is the internuclear distance, which is assumed constant in the case of covalently bound nuclei, and in all cases represents a vibrationally averaged distance. The angular parentheses define an average over all conformations exchanging on timescales faster than the millisecond. Dipolar couplings between covalently bound spins can be intrinsically very strong (around 11 kHz for an amide ^{15}N-^{1}H^{N} spin pair), nevertheless if all possible orientations θ are sampled with equal probability, as is the case in free solution, the value of the measured coupling averages very efficiently to zero. Residual couplings can be reintroduced by dissolving the protein in a weakly aligning medium such as lipid bicelles (Tjandra and Bax, 1997), filamentous phages (Torbet and Maret, 1979; Hansen et al., 1998; Clore et al., 1998), lyotropic ethylene glycol/alcohol phases (Rückert and Otting, 2000), and polyacrylamide gels that have been strained either laterally or longitudinally to produce anisotropic cavities (Sass et al., 2000; Tycko et al., 2000). In most of these cases, alignment results from a steric repulsion between the protein and the medium, whereas in the case of bacteriophage or charged forms of the other media, alignment results from a combination of electrostatic and steric interactions. In the case of electrostatic alignment the interpretation of RDCs in IDPs in terms of local structure is more complicated, although possible (Skora et al., 2006), making sterically aligning media the most commonly used for IDPs (see below).

RDCs measured in folded proteins provide information concerning the orientation of internuclear vectors connecting pairs of spins relative to a common alignment tensor. This tensor describes the net alignment of the protein relative to the magnetic field in terms of a second rank order matrix. This transformation assumes that the global alignment is not coupled to local fluctuations, such that RDCs for different spin pairs can be interpreted in terms of different orientations of the internuclear vectors relative to a common molecular frame. Equation 1 can be usefully recast to reflect this:

$$D = D_{\max} A_{zz}\left[P_2(\cos\vartheta) + \eta/2\sin^2\vartheta\cos 2\varphi\right] \qquad (3)$$

where A_{zz} is the longitudinal component of the alignment tensor, η is the rhombicity defined as $\eta = (A_{xx} - A_{yy})/A_{zz}$, and $\{\vartheta,\varphi\}$ are expressed as polar coordinates of the internuclear vector. Measured RDCs can then be interpreted in terms of different orientations of the internuclear vectors relative to the molecular frame. The correlation of angular order from distant parts of the molecule allows the determination of their average relative orientation, a type of information that is difficult to extract from isotropic solution state NMR. RDCs measured in globular folded proteins are used for structure determination (Bax, 2003; Prestegard et al., 2004; Blackledge, 2005), for the study of long-range order in extended molecules (Tjandra et al., 1997) and protein complexes (Clore, 2000; Ortega-Roldan et al., 2009). RDCs can also be very powerfully used for the characterization of local dynamics in proteins (Meiler et al., 2001; Clore and Schwieters, 2004; Tolman, 2002; Briggman and Tolman, 2003; Bernadó and Blackledge, 2004; Ulmer et al., 2004; Bouvignies et al., 2006; Lakomek et al., 2008; Salmon et al., 2009).

Interpretation of Residual Dipolar Couplings in Disordered Proteins

In the case of conformationally heterogeneous proteins such as IDPs, the alignment of all conformations of the molecule contributing to the time and ensemble average can be expected to vary significantly as a function of the shape and size of the individual conformation. In this case the RDC must be described in terms of the sum over the different time averages for all N molecules in the ensemble:

$$D = D_{\max}\frac{1}{N}\sum_{k=1}^{N}\frac{1}{t_{\max}}\int_{t=0}^{t_{\max}} P_2(\cos\theta_k(t))dt \qquad (4)$$

Assuming that each copy of the protein samples the conformational space of the ensemble (Meier et al., 2008), this can be further simplified to:

$$D = D_{\max}\frac{1}{t_{\max}}\int_{t=0}^{t_{\max}} P_2(\cos\theta(t))dt \qquad (5)$$

IDPs are highly flexible in solution, implying that the average in Equations 4 and 5 is potentially very complex and might not provide any useful information, or even that the dipolar coupling could be unmeasurably small in a truly unfolded state. Nevertheless, RDCs were measured in chemically denatured proteins, indicating that orientational sampling of internuclear vectors is not isotropic in these proteins (Figure 1).

$^1D_{NH}$ couplings were thus observed in partially aligned native and Δ131Δ mutants of Staphylococcal nuclease in 8M urea (Shortle and Ackerman, 2001), eglin C (Ohnishi et al., 2004), protein GB1 (Ding et al., 2004), apo-myoglobin (Mohana-Borges et al., 2004), and acyl-CoA binding protein (ACBP) (Fieber et al., 2004) under diverse denaturing conditions. A general distribution emerged in which the $^1D_{NH}$ couplings were found to have negative sign, with maximal values measured in the center of the protein, tapering off via a so-called bell-shaped distribution to zero at the extremities. Denatured proteins containing residual secondary structure, as for example identified from chemical shift measurements, showed deviations from the bell-shaped distributions. Examples of changes in magnitude and sign of $^1D_{NH}$ measurements were found for amide bonds in acid-denatured states of apo-myoglobin and ACBP and linked to raised helical propensities in regions that form helices in the native state. The observed change in sign was rationalized as follows: The average orientation of the amide bond vector present in an unfolded chain where the protein is preferentially aligned in a direction parallel to the magnetic field, for example in an elongated cavity, would be expected to be approximately orthogonal to the field, and therefore negative. In a helical element the bond vector would be aligned more or less parallel with the average chain direction and therefore the field (Figure 2). The angular averaging term $P_2(\cos\theta)$ would be expected to change sign between these two conditions.

Early work then used this apparent sensitivity to local structure to follow protein unfolding, either in a β-hairpin structure in the fibritin foldon domain, observing the diminution of the structure of the experimental RDC profile with increasing temperature (Meier et al., 2004), or a gradual decrease of RDCs with increasing temperature or decreasing salt concentration in α-helical ribonuclease S-peptide (Alexandrescu and Kammerer, 2003), or thermal unfolding of GB1 (Ding et al., 2004). RDCs were also used to probe amino acid conformational specificity in short peptides (Dames et al., 2006), and were then more generally applied to assess local and long-range structure in IDPs such as Tau protein (Sibille et al., 2006; Mukrasch et al., 2007a, 2007b) and α-synuclein (Bertoncini et al., 2005a; Bernadó et al., 2005b; Sung and Eliezer, 2007).

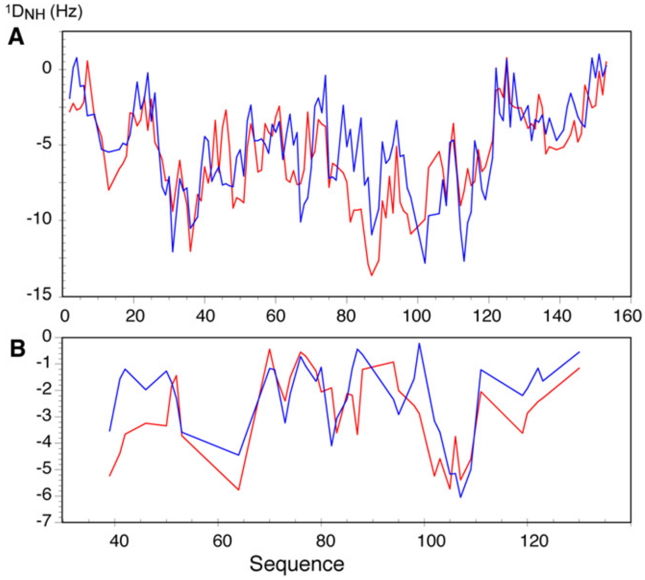


Figure 1. Experimental $^1D_{NH}$ Dipolar Couplings

Experimental $^1D_{NH}$ dipolar couplings (red) from the urea unfolded proteins (A) apo-myoglobin and (B) Staphylococcal nuclease Δ131Δ mutant aligned in radially squeezed polyacrylamide gel. RDCs were simulated using the explicit-ensemble Flexible-Meccano approach and are shown in blue for comparison. In each molecule all RDCs are multiplied by a common scaling factor to best reproduce the data. Copyright 2005, National Academy of Sciences, USA (Bernadó et al., 2005a).

The key to a more quantitative understanding of the time and ensemble average represented in Equation 5 was made by Annila and coworkers (Louhivuori et al., 2003; Fredriksson et al., 2004) who used polymer models to describe the unfolded protein as a series of connected segments of equal length experiencing restricted random walk. Integration of Equation 5 over available orientations of each segment formalizes the idea that in the presence of an obstacle, orientational sampling is more restricted in the center of the chain than at the termini, leading to non-vanishing RDCs, even when the torsion angles along the polymer chain can adopt random conformations. Segments in the center have more neighbors, and are therefore less flexible than those at the ends, rationalizing the experimentally observed bell-shaped distribution. This model has recently been extended, revised, and to an extent corrected, confirming the overall observations (Obolensky et al., 2007).

Such polymer-based models elegantly describe many aspects of the physical alignment of the unfolded polypeptide and present a relatively simple conceptual framework for the qualitative understanding of the experimental observations. However, the description of a natural amino acid sequence as a homopolymer is unrealistic and such analytical models cannot easily be adapted to interpret data from complex heteropolymeric systems such as proteins. An expected, site-specific dependence of RDCs was clearly predicted (Louhivuori et al., 2004) by

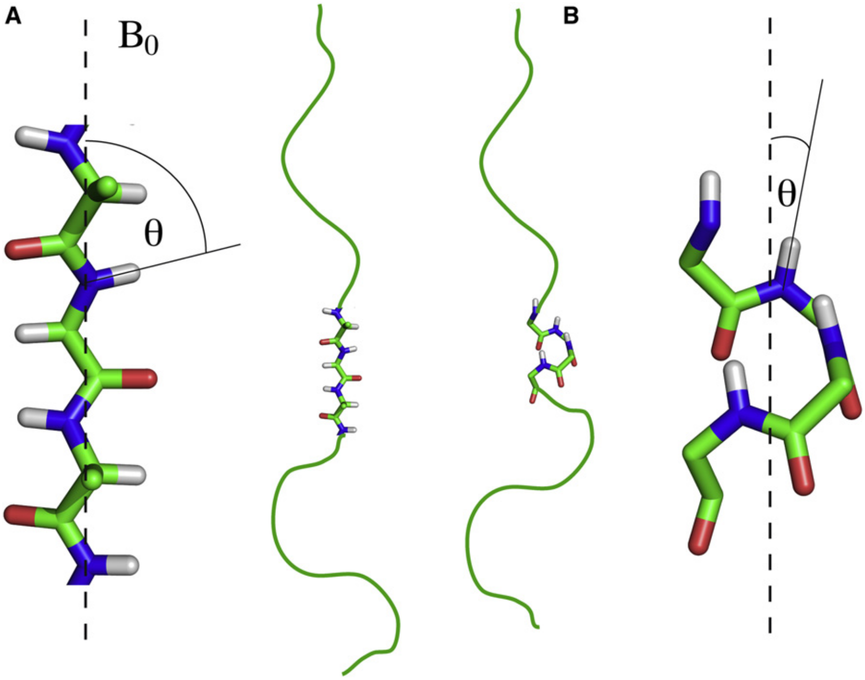


Figure 2. Figurative Representation of Effective Angular Averaging Properties of ^{15}N-^{1}H Vectors

Figurative representation of effective angular averaging properties of ^{15}N-^{1}H vectors in an unfolded protein dissolved in weakly aligning medium with the director along the magnetic field. Dipolar couplings measured for ^{15}N-^{1}H vectors in more extended conformations ($\theta \approx 90^\circ$), more commonly found in unfolded proteins, will have negative values (A), whereas those in helical or turn conformations align more or less parallel with the direction of the chain ($\theta \approx 0^\circ$) and will have larger positive values (B).

incorporation of glycine and proline residues into simulations of random homopolymers, which resulted in prediction of RDC profiles that increased or decreased due to the change in flexibility of the specific amino acids. The observation of significant site-to-site variation of experimental RDCs along an unfolded peptide chain compared with the overall bell-shaped prediction from the random flight models also underlined the need to introduce amino acid-specific conformational behavior into any interpretative model of RDCs measured in disordered proteins.

RDCs in Highly Flexible Systems: Explicit Ensemble Models

Two very similar, more direct approaches to the interpretation of RDCs from unfolded proteins have therefore been proposed, and these rely on the development of explicit ensemble descriptions of the unfolded state (Jha et al., 2005; Bernadó et al., 2005a). Measured couplings are expressed in terms of a discrete average over RDCs predicted for all sampled conformers on the basis of the molecular shape, or on the basis of electrostatic charge distribution in the case of electrostatic alignment.

$$D = D_{\max}\frac{1}{M}\sum_{k=1}^{M} A_{k,zz}\left[P_2(\cos\vartheta_k) + \eta_k/2\sin^2\vartheta_k\cos 2\varphi_k\right] \tag{6}$$

RDCs are averaged over a sufficient number (M) of conformers to fully represent the available conformational sampling. Although this number may be of the order of many thousands for an unfolded strand of 50 amino acids in length, it has recently been demonstrated that convergence of RDCs toward experimental data can be achieved with smaller number of conformers if the protein is divided into small, uncoupled segments (Marsh et al., 2008), although this decoupling of distant regions in the chain might not represent the true nature of the conformational space (vide infra). These approaches explicitly account for the heteropolymeric nature of the peptide chain, and sampling amino-acid-specific {ϕ/ψ} propensities to construct the conformational ensemble (Jha et al., 2005; Bernadó et al., 2005a). The study from Bernadó et al. sampled conformations from an explicitly constructed coil library, comprising non-α-helical and non-β sheet conformations from 500 high-resolution crystal structures (Lovell et al., 2003). Additional sampling properties were included that account, for example, for amino acids preceding prolines. Rudimentary nonbonding considerations were accounted for between amino acid side chains by removing structures when a steric clash occurred between residue-specific spheres centered on the β-carbon atoms of each amino acid (α-proton in the case of glycines). Conformers were constructed by randomly sampling the amino acid specific coil library, and RDCs were predicted for each copy of the ensemble using shape-based alignment algorithms (Zweckstetter and Bax, 2000; Almond and Axelsen, 2002) and averaged over the entire ensemble.

This approach, termed Flexible-Meccano or FM, was initially applied to a two-domain viral protein, protein X, from Sendai virus phosphoprotein (Figure 3) comprising a disordered domain and a folded domain. $^{1}D_{NH}$ and $^{2}D_{C'NH}$ RDCs predicted using the FM approach are relatively well reproduced from throughout the protein both in amplitude and distribution. Note that in this particular system the RDCs from each copy of the protein, in both folded and unfolded domains, depend on the relative alignment of the two domains, constituting a quantitative test of the validity of the approach.

Using this approach it was possible to accurately reproduce the experimentally measured $^{1}D_{NH}$ couplings in the Δ131Δ mutant of staphylococcal nuclease (Shortle and Ackerman, 2001) simply on the basis of local conformational propensities, without residual tertiary fold as had initially been invoked. A number of further examples were also shown, for example the prediction of $^{1}D_{NH}$ RDCs measured in 8M urea unfolded apo-myoglobin (Mohana-Borges et al., 2004), establishing the statistical coil approach as a viable choice for predicting random coil RDCs that result directly from the conformational properties of the primary sequence and thereby constitute an unfolded "baseline." The absolute level of alignment is not accurately known in these simulations, so that all RDCs are finally scaled by the appropriate optimal scaling factor to reproduce the experimental data.

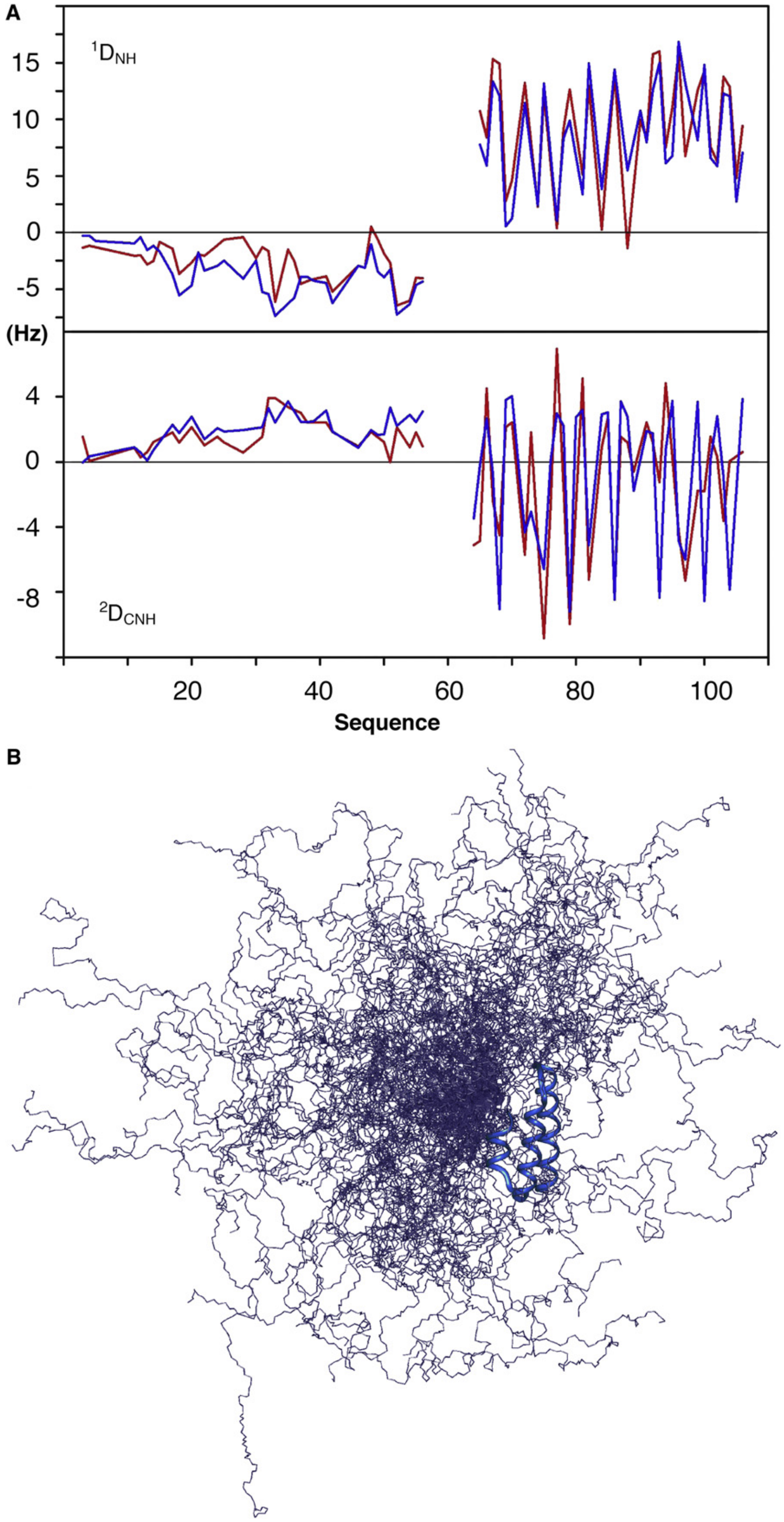


Figure 3. Experimental $^1D_{NH}$ and $^2D_{CNH}$ Dipolar Couplings from the Two-Domain Protein, PX, from Sendai Virus

(A) $^1D_{NH}$ and $^2D_{C'NH}$ RDCs are reasonably reproduced from throughout the protein using the explicit ensemble description Flexible-Meccano. Experimental values are shown in red. Copyright 2005, National Academy of Sciences, USA (Bernadó et al. 2005a).

(B) An ensemble representation of PX.

The observation that the distribution of experimental RDCs along the sequence is relatively well reproduced by explicit molecular descriptions that sample amino-acid specific statistical coil models is related to, and supported by, the observed correlation between amino-acid side-chain bulkiness and the amplitude of the $^{1}D_{NH}$ couplings (Cho et al., 2007).

RDCs Reveal Deviation from Random Coil Behavior

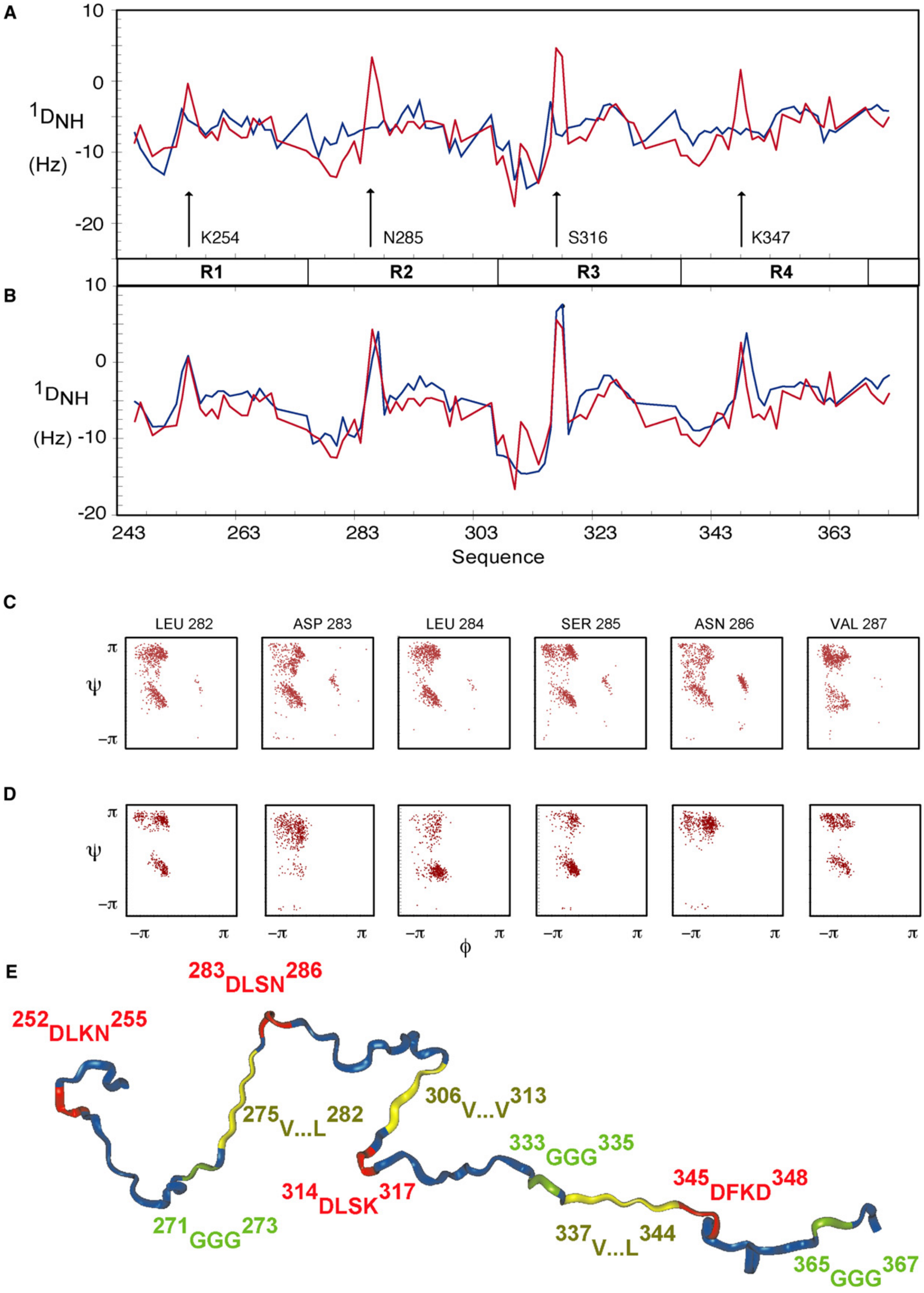
This statistical coil description of the unfolded state thus provides a straightforward method for calculating RDC profiles that would be expected if the protein behaved as a random coil, devoid of any specific or persistent local or long-range structure. Although the establishment of these approaches is clearly essential, the next step, the development of techniques whereby a departure from baseline values can then be interpreted in terms of specific local or long-range conformational behavior, is equally important and clearly more challenging. In particular, although the random coil sampling model can be used to detect transient local order, the quantitative description of conformational detail in these regions will require the development of additional techniques.

In a recent study of K18, a 130-amino-acid construct of the natively unfolded protein Tau that physiologically controls microtubule dynamics and stability and represents a significant fraction of the proteins found in tangles in Alzheimer's disease (Mandelkow and Mandelkow, 1998), $^{1}D_{NH}$ couplings were measured from throughout the chain (Mukrasch et al., 2007a). Local sign inversion of $^{1}D_{NH}$ RDCs was observed in four homologous repeat domains (R1–R4) containing the hexapeptide segments identified as the interaction sites of Tau with microtubules, as well as the sites involved in self-association, formation of paired helical filaments and eventual aggregation. This sign inversion was not reproduced by the statistical coil FM approach (Figure 4), and, following the logic presented previously, can be qualitatively interpreted as the presence of local helical or turn motifs. In this study the authors showed, via extensive simulation, that the simple observation of a single $^{1}D_{NH}$ RDC sign inversion within the sequence can have ambiguous origin. For example, left-handed helix backbone dihedral angle sampling of the neighboring amino acid can have a similar influence on a measured RDC at the site of interest as for example the presence of a right-handed helix at this site. In this case accelerated molecular dynamics (AMD) simulation, an approach that enhances access to rare conformational transitions and thereby extends the effective temporal range compared with standard MD simulation by many orders of magnitude (Markwick et al., 2007), was employed to predict the conformational behavior of pentadecapeptides centered on these turn regions. This revealed strong tendencies to form βI-turns for three repeats R1–R3. The results of this simulation were apparently validated by replacing intrinsic statistical coil backbone dihedral angles for the four amino acids involved in each turn region with the backbone dihedral sampling resulting from the AMD simulations (Figure 4). This model reproduced experimental RDC values closely, suggesting that the local βI-turn conformations and equally importantly their predicted populations were in very good agreement with the experimental RDC data. It is interesting to speculate on the importance of these turn conformations in the oligomerization process that precedes aggregation, and the absence of aggregation in the healthy form of the protein, particularly in the light of the observation, by solid-state NMR, that one of these regions is in the center of a long β sheet conformation in Alzheimer's-like paired helical filaments from the core domain of tau (Andronesi et al., 2008). These analyses allowed a detailed description of the conformational behavior of this protein form NMR data (Figure 4E). In this protein, no local ordering effects occurring on microsecond to millisecond timescales were observed by relaxation dispersion type of measurements (Klein-Seetharaman et al., 2002) and ^{15}N relaxation in the area of the turns was also inconspicuous.

Intriguingly, Flexible-Meccano also revealed the sensitivity of RDCs to the presence of transient long-range contacts in disordered proteins. This additional dependence should come as no surprise, when we recall that in the case of folded proteins RDCs report on orientations relative to a single alignment tensor, and that many partially disordered proteins lie somewhere between fully folded and completely random coil proteins, and might therefore exhibit characteristics associated with both extremes. Nevertheless the effect that transient long-range contacts would have on RDC profiles from disordered proteins remains difficult to predict in an intuitive way.

Experimental RDCs measured in a study of the IDP α-synuclein were not reproduced by a simple application of the random-coil FM approach, with significant and systematic variations in the RDC profile observed at the N and C termini of the protein (Figure 5) (Bernadó et al., 2005b). Although deviations compared with random coil RDCs in these regions might stem from local structure that is not predicted by the coil model, the potential relevance of an alternative explanation was also convincingly demonstrated. Ensembles were calculated that contained transient long-range contacts between different segments of the molecule. The 140-amino-acid molecule was arbitrarily divided into seven 20-amino-acid segments and RDCs were averaged for ensembles containing contacts of less than 15 Å between any residues present in different pairs of segments (Figure 5). A clear dependence was observed of the predicted RDCs on the presence of long-range contacts, even when weak and relatively nonspecific as in this case. RDCs are reinforced in the vicinity of the broad regions experiencing contacts, and quenched in the intervening regions. Importantly, although the RDCs are locally very different, the backbone conformational sampling of the amino acids that show increased RDCs compared to the completely unfolded ensemble is essentially identical in the presence and absence of the contacts. This dependence on transient contacts and fluctuating tertiary structure has important consequences for the recently proposed approaches that divide the unfolded protein into short uncorrelated segments to improve the efficiency of RDC prediction (Marsh et al., 2008), because such long-range effects would necessarily be absent from this kind of simulation.

The experimental RDCs were best reproduced in the presence of a long-range contact between the N- and C-terminal regions of α-synuclein (Figure 5). Interactions between the terminal regions have been detected using PREs (Dedmon et al., 2005 also saw evidence of contacts between the so-called NAC domain and the C-terminal) and have been shown to disappear


A
10
0
$^1D_{NH}$
(Hz)
-10
-20
K254
N285
S316
K347
R1
R2
R3
R4
B
10
0
$^1D_{NH}$
(Hz) -10
-20
243
263
283
303
323
343
363
Sequence
C
LEU 282
ASP 283
LEU 284
SER 285
ASN 286
VAL 287
π
ψ
−π
D
π
ψ
−π
−π
π
φ
E
^{252}DLKN^{255}
^{283}DLSN^{286}
^{275}V...L^{282}
^{306}V...V^{313}
^{333}GGG^{335}
^{345}DFKD^{348}
^{314}DLSK^{317}
^{271}GGG^{273}
^{337}V...L^{344}
^{365}GGG^{367}

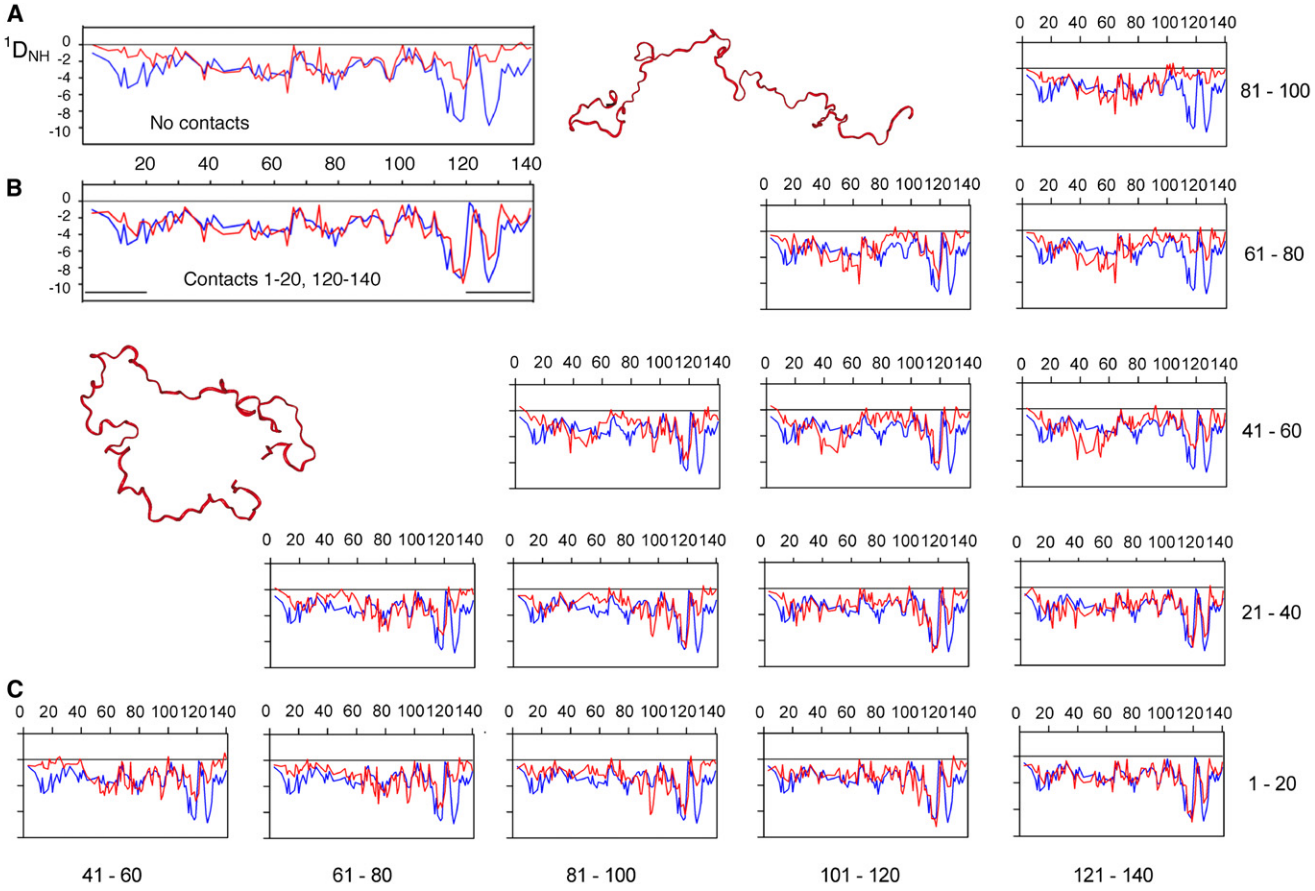


Figure 5. Sensitivity of RDCs to the Presence of Transient Long-Range Structure in IDPs

Experimental $^1D_{NH}$ RDCs were measured in α-synuclein aligned in bacteriophage (reprinted with permission from the *Journal of the American Chemical Society* [Bernadó et al. 2005b]). The RDC profile is reasonably well reproduced by FM except for the N- and C-terminal regions (A). The relevance of long-range interactions between different parts of the protein was systematically tested by dividing the 140-amino-acid chain into seven 20-residue strands (1-20, 21-40, etc.). The Flexible-Meccano procedure was repeated, and conformers were only accepted if a $^{\beta}C$ from one of the 20-residue domains was less than 15 Å from a $^{\beta}C$ from the other specified domain (C). The best reproduction is found when a contact between the N- and C-terminal regions is present (B). The presence of this contact is in agreement with paramagnetic relaxation enhancement measurements.

upon addition of denaturant, at high temperatures and upon polyamine binding. These conditions favor aggregation of α-synuclein in vitro, suggesting a role of the long-range interactions in α-synuclein against misfolding and aggregation (Bertoncini et al., 2005b). Importantly, in both the tau protein and in α-synuclein, no strong ordering effects occurring on microsecond to millisecond timescales were observed by relaxation dispersion type of measurements (Klein-Seetharaman et al., 2002). Thus, RDCs not only allow description of the conformational ensemble, but also suggest that interconversion between the ensemble members occur on the nanosecond to microsecond timescale.

Multiple RDCs for Accurate Characterization of Local Conformational Propensities

Although $^1D_{NH}$ RDCs provide sensitive probes of local conformational propensity, as well as evidence of the relevance of statistical coil models, their ambiguous interpretation in terms of local structural propensities highlights the need for complementary structural information. This can be provided in the form of additional RDCs between different pairs of nuclei on the protein backbone. Indeed Meier et al. have shown that long-range order can also be detected from RDCs by the measurement of $^1H^N$-$^1H^N$ RDCs in highly deuterated urea and acid denatured ubiquitin. $^1H^N$-$^1H^N$ RDCs suggested the presence of significantly

Figure 4. Identification of Stable Turn Conformations in the Four Repeat Domain of Tau Protein (K18) from RDCs and AMD Simulation

$^1D_{NH}$ RDCs measured in the 130 amino acid domain of Tau protein aligned in polyacrylamide gel (red). Experimental RDCs (blue) are reproduced reasonably well throughout the protein using the statistical coil model Flexible-Meccano (A). Four regions show significant deviations from expected behavior, with inversion of the sign of RDCs from highly homologous sequences in the protein. Backbone dihedral angle sampling from accelerated molecular dynamics simulation (D) of pentadecapeptides centered on the regions of interest differs from the statistical coil sampling (C). When incorporated into the Flexible-Meccano sampling, to replace the sampling from the coil database, the RDCs are better reproduced (B). (E) Ribbon diagram of K18 construct of Tau protein summarizing the conformational sampling propensities as derived from NMR data. The four strands identified in this study as containing turn propensities (252-255, 283-286, 314-317, 345-348) are shown in red the three GGG motifs (271-273, 333-335, 365-367) in green and the regions identified as having propensity toward β sheet conformations (274-283, 306-313, and 336-345) in yellow. Reprinted with permission from the *Journal of the American Chemical Society* (Mukrasch et al., 2007a).

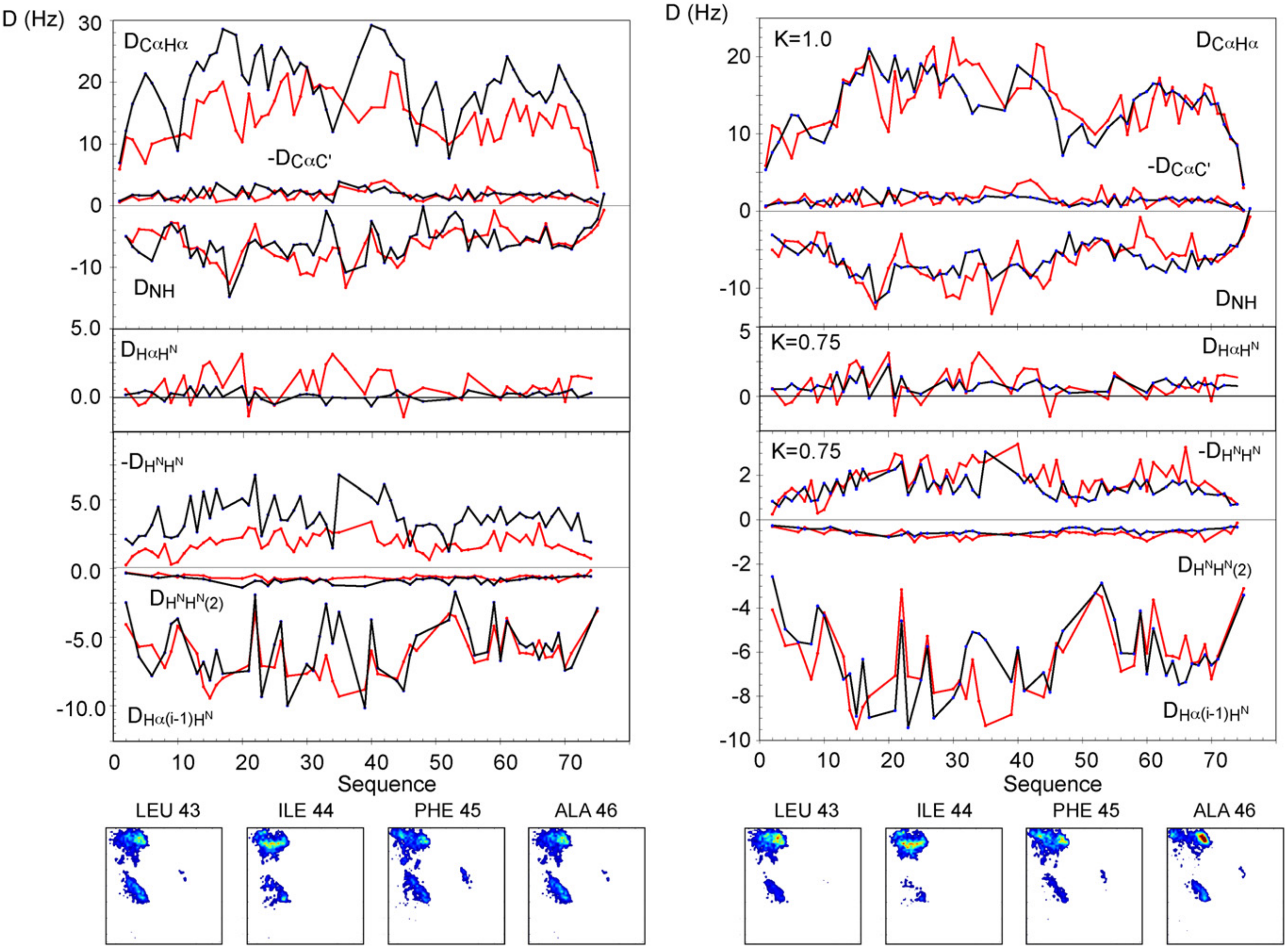


Figure 6. Conformational Sampling in Urea Unfolded Proteins

$^1D_{NH}$, $^1D_{C\alpha H\alpha}$ and $^1D_{C\alpha C'}$, $^1H^N_i$-$^1H^\alpha_i$, $^1H^N_i$-$^1H^\alpha_{i-1}$ were measured in ubiquitin at pH 2 and 8M urea, and $^1H^N_i$-$^1H^N_{i+1}$ and $^1H^N_i$-$^1H^N_{i+2}$ measured under the same conditions in perdeuterated ubiquitin and compared with expected couplings from the standard statistical coil database (black). All couplings are scaled using scaling factors appropriate for the $^1D_{NH}$ coupling (left). The general disagreement between experimental and simulated RDCs appears to stem from the nature of the statistical coil model, which, when modified to reflect enhanced sampling in the more extended regions of Ramachandran space (right), provides a better overall reproduction of the RDCs. In this case RDCs between covalently bound spins are scaled using scaling factors appropriate for the $^1D_{NH}$ coupling, whereas all ^1H-^1H are scaled using the best scaling for $^1H^N_i$-$^1H^\alpha_{i-1}$ couplings. Four sample Ramachandran plots are shown to illustrate this enhanced sampling. Reprinted with permission from the *Journal of the American Chemical Society* (Meier et al., 2007b).

populated (around 20%) native-like local structure in the N-terminal β-hairpin (Meier et al., 2007a). In the same study of ubiquitin, up to seven RDCs per peptide unit were measured, including $^1H^N$-$^1H^N$ and $^1H^N$-$^1H^\alpha$ RDCs. The FM approach was initially used to predict one-bond RDCs ($^1D_{NH}$, $^1D_{C\alpha H\alpha}$, and $^1D_{C\alpha C'}$) from the unfolded chain. Although the profiles of the experimental $^1D_{NH}$ and $^1D_{C\alpha H\alpha}$ RDCs were reasonably well reproduced by simulation, a significantly different scaling factor was required to reproduce the different RDC types. Further simulation suggested that the standard statistical coil distribution of dihedral angles was inappropriate for the description of RDCs from urea-unfolded proteins (Figure 6). Refinement of the conformational sampling distribution invoking a generally higher propensity for extended conformations {$50° < \psi < 180°$} achieved simultaneous reproduction of the different types of RDCs ($^1D_{NH}$, $^1D_{C\alpha H\alpha}$, and $^1D_{C\alpha C'}$). This observation was verified by a comparison of calculated and experimental interproton RDCs $^1H^N_i$-$^1H^\alpha_i$ and $^1H^N_i$-$^1H^\alpha_{i-1}$, and interamide proton RDCs measured using quantitative J-coupling approaches (Meier et al., 2003) in perdeuterated ubiquitin ($^1H^N_i$-$^1H^N_{i+1}$ and $^1H^N_i$-$^1H^N_{i+2}$). We note that an additional scaling was required for all ^1H-^1H RDCs, possibly due to increased mobility not accounted for in the statistical coil model.

It is therefore necessary to evoke an increased population of more extended conformations in order to reproduce RDCs measured in urea-denatured proteins than appears appropriate for IDPs. This supports the proposition that urea denaturation extends the unfolded amino acid chain, an observation that would be in agreement with local binding of urea to the polypeptide chain, inducing restricted, and more extended sampling of backbone dihedral angles (Meier et al., 2007b, although clearly does not disprove the side-chain solvation model. The suggestion that conformational sampling that is appropriate for urea-denatured proteins is not adapted to the behavior of IDPs is further supported by the observation that the modified sampling determined to be relevant for urea denatured ubiquitin strongly violates RDCs measured in the protein PX, for which the standard database is found to be suitable (see above and compare Figure 7 with Figure 3). Analysis of $^3J_{HNH\alpha}$ scalar couplings measured under the same conditions indicated that although the ψ angle was more extended, the ϕ dihedral angle appeared to span both polyproline II and extended β-regions such that neither dominated the additional sampling of extended conformations.

Quantitative Analysis of Local Conformational Propensities from RDCs: Application to Sendai Virus N_{TAIL}

A key aspect of the intricate relationship between structural dynamics and biological function in IDPs is the observed capacity of members of this family to undergo a disorder-to-order transition on interaction with physiological partners, where molecular recognition is accompanied by local folding into a characteristic three-dimensional conformation (Vucetic et al., 2005; Sickmeier et al., 2007; Tompa and Fuxreiter, 2008; Vacic et al., 2007). These processes can exhibit high specificity but low affinity, with rapid dispersal due to high k_{on} and k_{off} rates and might be promiscuous, allowing binding to multiple partners via conformational plasticity in the molecular recognition element. Protein interactions that fall into this category are prevalent, but fall outside the range of classical structure-based approaches. In order to develop an understanding of the physical basis of induced folding upon binding, an accurate description of the conformational behavior of the prerecognition, free form of the protein is required.

The inherent flexibility of IDPs has hindered detailed atomic resolution characterization of the prerecognition state. Recently the dynamics of peptide folding upon interaction have been studied using rotating frame relaxation, identifying the formation of initial encounter complexes via weak, nonspecific interactions that facilitate the formation of a partially folded state upon binding (Sugase et al., 2007). Such observations support the previously proposed "fly-casting" mechanism that provides a theoretical framework for speeding up molecular recognition processes via the folding funnel (Shoemaker et al., 2000).

As we have seen, RDCs are sensitive probes of local conformational sampling in the unfolded state, and as such can contribute significantly to our understanding of the extent to which regions of a protein that play a role in binding and function are preconfigured prior to interaction. We have also seen that existing methods for the interpretation of RDCs in terms of local conformational propensities had thus far remained qualitative. In a recent study a major step was taken toward the quantitative, and eventually insightful analysis of local structure from RDCs. FM was used to study the structural properties of the partially ordered molecular recognition element of the C-terminal domain, N_{TAIL}, of Sendai virus nucleoprotein (Jensen et al., 2008). Replication and transcription of the viral RNA are initiated by an interaction between N_{TAIL} and the C-terminal three-helix bundle domain, PX, of the phosphoprotein P (Blanchard et al., 2004). The molecular recognition element was found, from chemical shift and prediction based on primary sequence, to present a nascent α-helix in free solution that further folds upon interaction with PX via a negative patch on the surface of PX (Houben et al., 2007).

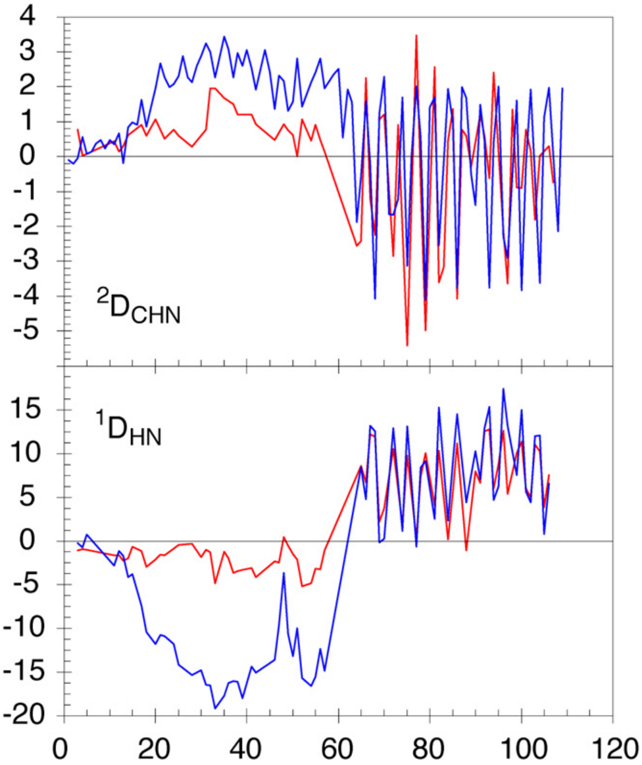


Figure 7. Experimental $^1D_{NH}$ and $^2D_{CNH}$ Dipolar Couplings from the Two-Domain Protein, PX, from Sendai Virus

$^1D_{NH}$ and $^2D_{C'NH}$ RDCs were calculated using the modified statistical coil model that was found to be most appropriate for urea-denatured proteins. Compared with the analysis (Figure 3) using the standard database, the data reproduction is significantly worse. In particular the RDCs are too large in the unfolded domain compared with the folded domain. This result supports the notion that conformational sampling in intrinsically unfolded and denatured proteins is significantly different.

$^1D_{HN}$, $^1D_{C\alpha C'}$, $^2D_{HNC'}$, and $^1D_{C\alpha H\alpha}$ RDCs were measured from N_{TAIL} aligned in liquid crystalline ethylene glycol/alcohol phase. Not unexpectedly, the helical region exhibits strongly positive $^1D_{HN}$ RDCs, confirming the presence of a helical motif, whereas the disordered strands predominately have negative RDCs (Figure 8). In order to interpret all experimental RDCs in terms of quantitative conformational sampling, all possible combinations and populations of continuous helical segments, from a minimum of 4 amino acids to a maximum of 20, from throughout the molecular recognition element segment were systematically combined (Figure 8), including the possibility of an unfolded state in equilibrium with the helices. In this case the effective RDC is given by:

$$D_{ij,eff} = \sum_{k=1,n} p_k D_{ij}^k + \left(1 - \sum_{k=1,n} p_k\right) D_{ij}^U. \tag{7}$$

Here, p_k represents the populations of the n helical conformers, for which D_{ij}^k are the individual predicted couplings between nuclei i and j, and D_{ij}^U are the couplings from the

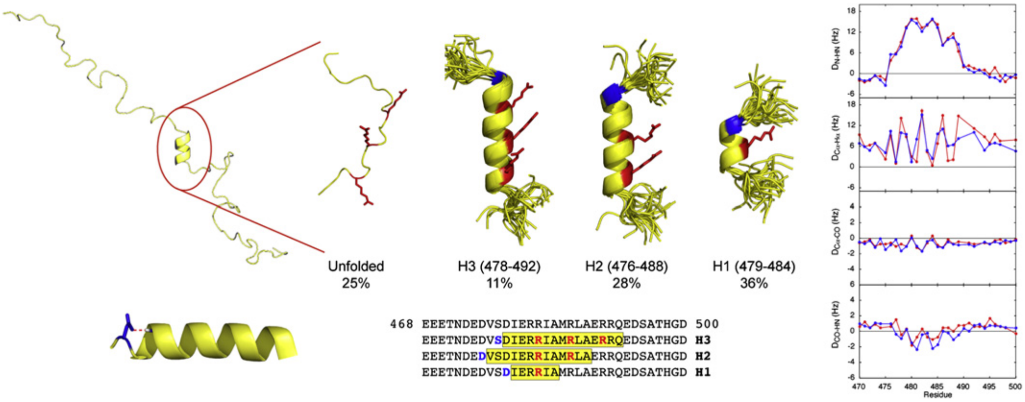


Figure 8. Determination of the Conformational Equilibrium in the Molecular Recognition Element of N_{TAIL} in Solution

Explicit structural ensembles were simulated using specific helical elements of all possible combinations and populations of continuous helical segments, from a minimum of 4 amino acids to a maximum of 20 in the range 476-495. Ensemble equilibria comprising combinations of increasing numbers of conformers (n = 0, 1, 2, 3, 4) of the 153 helical conformations were compared to the experimental data. In each case the population of each member of the ensemble was optimized. The four conformations are presented as: a single structure, representing the 25 ± 4% unfolded conformers, the shortest helical element, comprising six amino acids 479-484, populated at a level of 36 ± 3%, 476-488 populated at 28 ± 1% and a longer stretch 478-492 populated to a level of 11 ± 1%. The molecular recognition site arginines are shown in red. Twenty randomly selected conformers are shown for each of the helical segments to illustrate the directionality of the adjacent chains projected from the helix caps. Reproduction of experimental data (red) is shown compared with simulation (blue) in the molecular recognition element on the right. Each of the helices is found to be preceded by an amino acid capable of forming an N-capping interaction that can stabilize the formation of helices in flexible peptides (shown in blue on the ribbon and primary sequence). Reprinted with permission from the *Journal of the American Chemical Society* (Jensen et al., 2008).

unfolded state. These effective couplings are compared with experimental data using the expression:

$$\chi^2 = \sum \left(D_{ij,eff} - D_{ij,\exp}\right)^2 / \sigma_{ij}^2 \tag{8}$$

where σ represents the uncertainty on the experimental coupling.

A minimum ensemble representation of the molecular recognition element of N_{TAIL} was obtained from all available RDCs, in terms of interconverting conformational states using statistical F-tests to test the significance in the improvement in data fitting upon addition of a new population. The results clearly indicate that rather than fraying randomly, the molecular recognition sequence of N_{TAIL} preferentially populates three specific helical conformers. The two highest-populated conformers were found to differ by one helical turn in length at both termini, both enclosing the recognition site amino acids (Figure 8). Remarkably, the three interconverting helical conformers were all found to be stabilized by so-called N-capping interactions via hydrogen bonds between the side chain of the N-capping amino acid (normally aspartic acids or serine) and the backbone amide in position two or three in the helical elements. The preferential helices are stabilized by these motifs, suggesting that the favored conformations are encoded in the primary sequence of the molecular recognition element. This provides clear detail of the molecular basis of nascent helix formation in partially folded chains, with additional implications for understanding the early steps of protein folding. Possibly equally importantly, the direction in which the disordered strands adjacent to the helix are projected is selectively controlled as a result of these stabilizing interactions. A mechanism by which the partially folded form of the protein could project the unfolded strands in the most functionally useful direction to achieve efficient fly-casting interactions is thereby identified (Shoemaker et al., 2000).

The origin of the periodicity of the $^1D_{HN}$ couplings within helical elements (Figure 8), in addition to that exhibited in the $^1D_{C\alpha H\alpha}$, and to some extent $^1D_{C\alpha C'}$, and $^2D_{HNC'}$ couplings, is not immediately obvious. If one assumes that the helix is not deformed, the periodicity should only occur if the effective orientation of the vectors on either side of the helix differs relative to the magnetic field, resulting in an effective tilt of the main axis of the helical element with respect to this axis. In disordered proteins, the effective tilt of the helix relative to the alignment axis is determined by the directionality of the unfolded chains projected from the helix termini. The amplitude and phase of the dipolar wave have indeed been shown to depend in a predictable and analytical way on helix length (Jensen and Blackledge, 2008), in theory obviating the need for construction of explicit ensembles for all helical lengths (Figure 9). This dependence has been formalized for use as an alternative to the computationally onerous explicit ensemble construction, for the interpretation of RDCs measured in helical elements of partially folded chains.

RDCs Provide the Key to a Description of Conformational Sampling in the Disordered Transactivation Domain of Human Tumour Suppressor p53

RDCs offer the possibility for quantitative description of local structural detail and as such provide powerful probes with which to map the structural and dynamic properties of IDPs in solution. It is however evident that a full understanding of the vast conformational space available to these proteins requires experimental data from as many complementary biophysical techniques as possible in order to understand the nature of the unfolded state. An example of the combination of dipolar couplings with

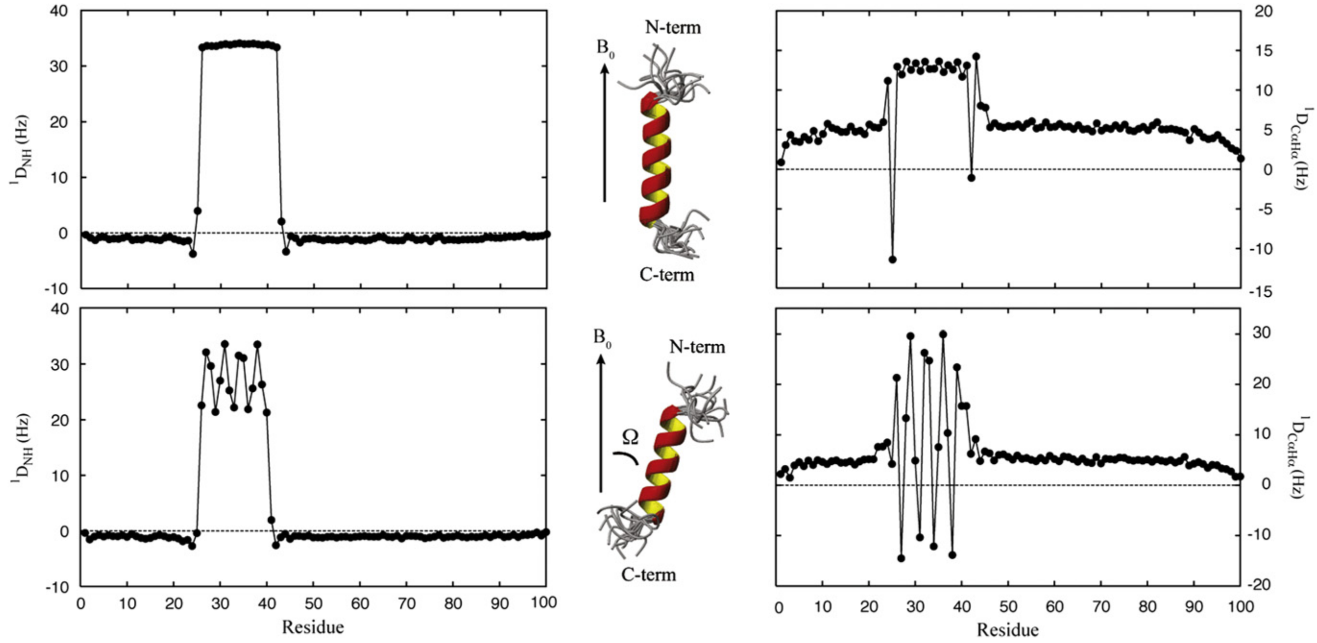


Figure 9. Residual Dipolar Couplings

RDCs (right: $^{1}D_{NH}$, left: $^{1}D_{CaHa}$) simulated for a 100-amino-acid poly-alanine chain with helices at 26-41 (top) and 26-39 (bottom). The RDCs were averaged over 50K conformers created by the Flexible-Meccano protocol. The appearance of the RDCs within the helical elements is strongly correlated to the directionality of the disordered chains projected from the helix caps. If the disordered chains are projected in the same direction (26-41), the dipolar oscillations are small because the effective helix orientation is close to parallel to the field (Ω = 1.3°). If the chains are projected in opposing directions (26-39), large oscillations are observed due to a large induced effective tilt of the helix (Ω = 18°). The amplitude and phase of the dipolar waves, therefore, depend in a predictable and analytical way on helix length. Reprinted with permission from the *Journal of the American Chemical Society* (Jensen and Blackledge, 2008).

complementary biophysical techniques was illustrated in the development of the first explicit ensemble description of the human tumor suppressor p53 (Wells et al., 2008). This protein plays a vital role in maintaining the integrity of the human genome, controlling apoptosis, cell-cycle arrest, and DNA repair. P53 exists as a homotetramer, with folded tetramerization and core domains that are linked via and flanked by intrinsically disordered domains at the N and C termini. The structure and quaternary geometry of the folded domains of p53 have been studied using NMR and crystallography, but no explicit model of the unfolded domain has been obtained in the context of the entire protein. Wells et al. have recently combined RDCs, AMD, and small-angle scattering to study the intrinsically disordered N-terminal transactivation domain of p53 in isolation, in the full-length form of the protein bound to DNA and in the unbound form.

This study clearly recognized differential flexibility in different regions of the N-terminal disordered domain of p53 (Figure 10). A proline-rich region, attached to the folded core domain, exhibited enhanced stiffness relative to the transactivation domain, thereby effectively projecting the MDM2 interaction site away from the surface of the protein. MDM2 is an important negative regulator of p53, and multiple RDCs measured from throughout the N-terminal domain, revealed the presence of a single helix turn at the MDM2 interaction site. Similar approaches to those applied for the study of N_{TAIL} quantified the population of the helix to be approximately 30%, consistent with AMD calculations, and supporting earlier suggestion of a nascent helix that fully folds upon interaction with MDM2. Importantly this structural motif is again preceded by an aspartic acid, as in Sendai virus N_{TAIL} and the beta turns present in Tau K18, again suggesting that the nascent structure is prepared and stabilized via N-capping interactions. Possibly more remarkably, the dynamic properties of the disordered N-terminal domain allow the measurement of RDCs from the entire disordered domain in the presence of the full-length tetrameric protein, both in the free and the DNA-complexed forms (a particle of molecular mass of 240 kDa). The results indicated that local conformational sampling of the N-terminal domain is remarkably similar in both the full-length protein and in isolation, and were validated in the isolated and intact forms against small-angle scattering data.

Conclusions

RDCs offer remarkably sensitive and agile probes for the study of local structural propensity in intrinsically disordered proteins. As we have seen from the examples presented here from our laboratory, the development of appropriate methods to interpret measured couplings in terms of conformational behavior is evolving very rapidly, as befits a nascent field of research. The combination of appropriate ensemble descriptions has been shown to allow the extraction of unique and important information on the conformational propensities of IDPs. The future success of the technique depends on the establishment of robust approaches that can be used to unambiguously identify structural properties of IDPs accurately and directly from experimental data. These techniques are then destined to make significant and original contributions to our understanding of

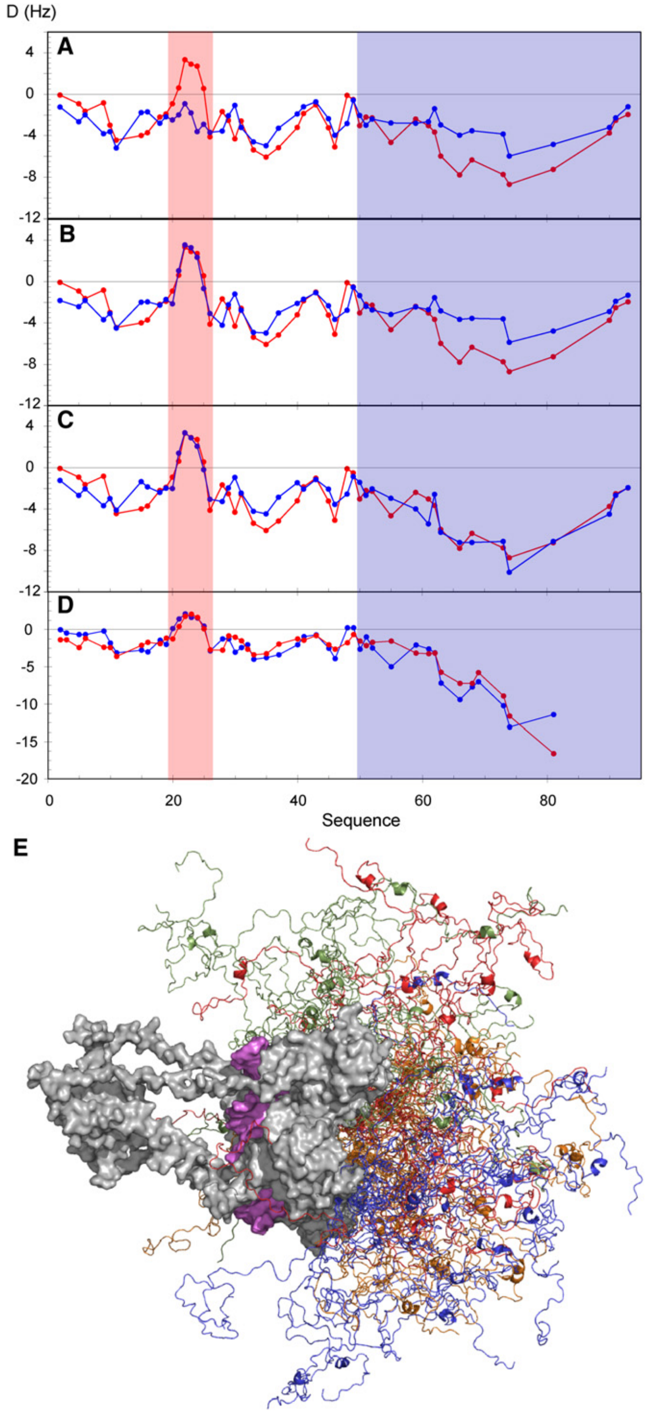


Figure 10. Defining the Conformational Behavior of p53 N-Terminal Intrinsically Unfolded Domain Using RDCs

(A) $^1D_{NH}$ RDCs from p53(1-93) compared with amino-acid-specific statistical coil predictions from FM. All simulated values are scaled by the same prefactor to best reproduce the experimental data.

(B) $^1D_{NH}$ RDCs from p53(1-93) compared with predicted values from amino-acid-specific statistical coil predictions including the presence of a single-turn helix at amino acids 22-24, populated at a level of 30% (red shading). Helical values were centered on the conformations present in the X-ray crystal structure of the α-helix formed when [1-93] binds to the ubiquitin ligase MDM2.

(C) Conformational sampling as in (B) with an increased level of polyproline II sampling for each amino acid in the region [58-91]-(blue shading). All simulated values are scaled by the same prefactor to best reproduce the experimental data.

(D) Comparison of experimental NH^N RDCs from the 1-93 region of intact p53[1-393]:DNA (red) with predictions from Flexible-Meccano using the full-length intact model of DNA bound form (blue). Conformational sampling of the unfolded domain as in (C).

(E) N-terminal domain ensemble in our model with one representative full-length p53 molecule included for illustration. The flexible C-terminal is not shown for reasons of clarity. N-terminal domains from the four different monomers are shown in different colors for clarity. Twenty copies are shown for each monomer.

Copyright 2008, National Academy of Sciences, USA (Wells et al., 2008).

the relationship between conformational preferences in the disordered state, the timescale of their interconversions, and their role in molecular function and malfunction, and to provide fundamentally important tools for studying the dynamic structural biology that is characteristic of intrinsically disordered proteins.

REFERENCES

Alexandrescu, A.T., Abeygunawardana, C., and Shortle, D. (1994). Structure and dynamics of a denatured 131-residue fragment of staphylococcal nuclease: a heteronuclear NMR study. Biochemistry *33*, 1063–1072.

Alexandrescu, A.T., and Kammerer, R.A. (2003). Structure and disorder in the ribonuclease S-peptide probed by NMR residual dipolar couplings. Protein Sci. *12*, 2132–2140.

Aloy, P., and Russell, R.B. (2004). Ten thousand interactions for the molecular biologist. Nat. Biotechnol. *22*, 1317–1321.

Almond, A., and Axelsen, J.B. (2002). Physical interpretation of residual dipolar couplings in neutral aligned media. J. Am. Chem. Soc. *124*, 9986–9987.

Andronesi, O.C., von Bergen, M., Biernat, J., Seidel, K., Griesinger, C., Mandelkow, E., and Baldus, M. (2008). Characterization of Alzheimer's-like paired helical filaments from the core domain of tau protein using solid-state NMR spectroscopy. J. Am. Chem. Soc. *130*, 5922–5928.

Bax, A. (2003). Weak alignment offers new NMR opportunities to study protein structure and dynamics. Protein Sci. *12*, 1–18.

Bernadó, P., and Blackledge, M. (2004). Local dynamic amplitudes on the protein backbone from dipolar couplings: towards the elucidation of slower motions in biomolecules. J. Am. Chem. Soc. *126*, 7760–7761.

Bernadó, P., Blanchard, L., Timmins, P., Marion, D., Ruigrok, R.W., and Blackledge, M. (2005a). A structural model for unfolded proteins from residual dipolar couplings and small-angle x-ray scattering. Proc. Natl. Acad. Sci. USA *102*, 17002–17007.

Bernadó, P., Bertoncini, C.W., Griesinger, C., Zweckstetter, M., and Blackledge, M. (2005b). Defining long-range order and local disorder in native α-synuclein using residual dipolar couplings. J. Am. Chem. Soc. *127*, 17968–17969.

Bernadó, P., Mylonas, E., Petoukhov, M.V., Blackledge, M., and Svergun, D.I. (2007). Structural characterization of flexible proteins using small-angle X-ray scattering. J. Am. Chem. Soc. *129*, 5656–5664.

Bertoncini, C.W., Jung, Y.S., Fernandez, C.O., Hoyer, W., Griesinger, C., Jovin, T.M., and Zweckstetter, M. (2005a). Release of long-range tertiary interactions potentiates aggregation of natively unstructured α-synuclein. Proc. Natl. Acad. Sci. USA *102*, 1430–1435.

Bertoncini, C.W., Fernandez, C.O., Griesinger, C., Jovin, T., and Zweckstetter, M. (2005b). Familial mutants of alpha-synuclein with increased neurotoxicity have a destabilized conformation. J. Biol. Chem. *280*, 30649–30652.

Blackledge, M. (2005). Recent advances in the use of residual dipolar couplings for the study of biomolecular structure and dynamics in solution. Prog. Nucl. Magn. Reson. Spectrosc. *46*, 23–61.

Blanchard, L., Tabouriech, N., Blackledge, M., Timmins, P., Burmeister, W.P., Ruigrok, R.W.H., and Marion, D. (2004). Structure and dynamics of the nucleocapsid-binding domain of the Sendai virus phosphoprotein in solution. Virology *319*, 201–211.

Bouvignies, G., Markwick, P.R.L., 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.

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

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

Cho, M.-K., Kim, H.-Y., Bernado, P., Fernandez, C.O., Blackledge, M., and Zweckstetter, M. (2007). Amino acid bulkiness defines the local conformations and dynamics of natively unfolded α-synuclein and tau. J. Am. Chem. Soc. *129*, 3032–3033.

Clore, G.M. (2000). Accurate and rapid docking of protein–protein complexes on the basis of intermolecular nuclear Overhauser enhancement data and dipolar couplings by rigid body minimization. Proc. Natl. Acad. Sci. USA *97*, 9021–9025.

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.

Clore, G.M., Starich, M.R., and Gronenborn, A.M. (1998). Measurement of residual dipolar couplings of macromolecules aligned in the nematic phase of a colloidal suspension of rod-shaped viruses. J. Am. Chem. Soc. *120*, 10571–10572.

Dames, S.A., Aregger, R., Vajpai, N., Bernadó, P., Blackledge, M., and Grzesiek, S. (2006). Residual dipolar couplings in short peptides reveal systematic conformational preferences of individual amino acids. J. Am. Chem. Soc. *128*, 13508–13514.

Dedmon, M.M., Lindorff-Larsen, K., Christodoulou, J., Vendruscolo, M., and Dobson, C.M. (2005). Mapping long-range interactions in alpha-synuclein using spin-label NMR and ensemble molecular dynamics simulations. J. Am. Chem. Soc. *127*, 476–477.

Denning, D.P., Uversky, V., Patel, S.S., Fink, A.L., and Rexach, M. (2002). The *Saccharomyces cerevisiae* nucleoporin Nup2p is a natively unfolded protein. J. Biol. Chem. *277*, 33447–33455.

Ding, K., Louis, J.M., and Gronenborn, A.M. (2004). Insights into conformation and dynamics of protein GB1 during folding and unfolding by NMR. J. Mol. Biol. *335*, 1299–1307.

Dunker, A.K., Silman, I., Uversky, V.N., and Sussman, J. (2008). Function and structure of inherently disordered proteins. Curr. Opin. Struct. Biol. *18*, 756–764.

Dyson, H.J., and Wright, P.E. (2002). Coupling of folding and binding for unstructured proteins. Curr. Opin. Struct. Biol. *12*, 54–60.

Dyson, H.J., and Wright, P.E. (2004). Intrinsically unstructured proteins and their functions. Chem. Rev. *104*, 3607–3622.

Emsley, J.W., and Lindon, J.C. (1975). NMR Spectroscopy Using Liquid Crystal Solvents (Oxford: Pergamon Press).

Felitsky, D.J., Lietzow, M.A., Dyson, H.J., and Wright, P.E. (2008). Modeling transient collapsed states of an unfolded protein to provide insights into early folding events. Proc. Natl. Acad. Sci. USA *105*, 6278–6283.

Fieber, W., Kristjansdottir, S., and Poulsen, F.M. (2004). Short-range, long-range and transition state interactions in the denatured state of ACBP from residual dipolar couplings. J. Mol. Biol. *339*, 1191–1199.

Fink, A.L. (2005). Natively unfolded proteins. Curr. Opin. Struct. Biol. *15*, 35–41.

Fredriksson, K., Louhivuori, M., Permi, P., and Annila, A. (2004). On the interpretation of residual dipolar couplings as reporters of molecular dynamics. J. Am. Chem. Soc. *126*, 12646–12650.

Fuxreiter, M., Simon, I., Friedrich, P., and Tompa, P. (2004). Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. J. Mol. Biol. *338*, 1015–1026.

Gillespie, J.R., and Shortle, D. (1997). Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels. J. Mol. Biol. *268*, 158–169.

Graf, J., Nguyen, P.H., Stock, G., and Schwalbe, H. (2007). Structure and dynamics of the homologous series of alanine peptides: a joint molecular dynamics/NMR study. J. Am. Chem. Soc. *129*, 1179–1189.

Hansen, M.R., Mueller, L., and Pardi, A. (1998). Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. Nat. Struct. Biol. *5*, 1065–1074.

Houben, K., Marion, D., Tarbouriech, N., Ruigrok, R.W.H., and Blanchard, L. (2007). Interaction of the C-terminal domains of sendai virus N and P proteins: Comparison of polymerase-nucleocapsid interactions within the paramyxovirus family. J. Virol. *81*, 6807–6816.

Jeganathan, S., von Bergen, M., Brutlach, H., Steinhoff, H.J., and Mandelkow, E. (2006). Global hairpin folding of tau in solution. Biochemistry *45*, 2283–2293.

Jensen, M.R., Houben, K., Lescop, E., Blanchard, L., Ruigrok, R.W.H., 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.

Jensen, M.R., and Blackledge, M. (2008). On the origin of NMR dipolar waves in transient helical elements of partially folded proteins. J. Am. Chem. Soc. *130*, 11266–11267.

Jha, A.K., Colubri, A., Freed, K.F., and Sosnick, T.R. (2005). Statistical coil model of the unfolded state: Resolving the reconciliation problem. Proc. Natl. Acad. Sci. USA *102*, 13099–13104.

Klein-Seetharaman, J., Oikawa, M., Grimshaw, S.B., Wirmer, J., Duchardt, E., Ueda, T., Imoto, T., Smith, L.J., Dobson, C.M., and Schwalbe, H. (2002). Long-range interactions within a nonnative protein. Science *295*, 1719–1722.

Lakomek, N.A., Walter, K.F., 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.

Louhivuori, M., Pääkkönen, K., Fredriksson, K., Permi, P., Lounila, J., and Annila, A. (2003). On the origin of residual dipolar couplings from denatured proteins. J. Am. Chem. Soc. *125*, 15647–15650.

Louhivuori, M., Fredriksson, K., Paakkonen, K., Permi, P., and Annila, A. (2004). Alignment of chain-like molecules. J. Biomol. NMR *29*, 517–524.

Lovell, S.C., Davis, I.W., Arendall, W.B., III, de Bakker, P.I.W., Word, J.M., Prisant, M.G., Richardson, J.S., and Richardson, D.C. (2003). Structure validation by C alpha geometry: phi,psi and C beta deviation. Proteins *50*, 437–450.

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

Maiti, N.C., Apetri, M.M., Zagorski, M.G., Carey, P.R., and Anderson, V.E. (2004). Raman spectroscopic characterization of secondary structure in natively unfolded proteins: α-synuclein. J. Am. Chem. Soc. *126*, 2399–2408.

Mandelkow, E.-M., and Mandelkow, E. (1998). Tau in Alzheimer's disease. Trends Cell Biol. *8*, 425–427.

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.

Marsh, J.A., Singh, V.K., Jia, Z.C., and Forman-Kay, J.D. (2006). Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: Implications for fibrillation. Protein Sci. *15*, 2795–2804.

Marsh, J.A.,, Baker, J.M.R., Tollinger, M., and Forman-Kay, J.D. (2008). Calculation of residual dipolar couplings from disordered state ensembles using local alignment. J. Am. Chem. Soc. *130*, 7804–7805.

Meier, S., Häussinger, D., Jensen, P., Rogowski, M., and Grzesiek, S. (2003). High-accuracy residual 1HN-13C and 1HN-1HN dipolar couplings in perdeuterated proteins. J. Am. Chem. Soc. *125*, 44–45.

Meier, S., Guthe, S., Kiefhaber, T., and Grzesiek, S. (2004). Foldon, the natural trimerization domain of T4 fibritin, dissociates into a monomeric A-state form containing a stable beta-hairpin: Atomic details of trimer dissociation and local beta-hairpin stability from residual dipolar couplings. J. Mol. Biol. *344*, 1051–1069.

Meier, S., Strohmeier, M., Blackledge, M., and Grzesiek, S. (2007a). Direct observation of dipolar couplings and hydrogen bonds across a β-hairpin in 8 M urea. J. Am. Chem. Soc. *129*, 754–755.

Meier, S., Grzesiek, S., and Blackledge, M. (2007b). Mapping the conformational landscape of urea-denatured ubiquitin using residual dipolar couplings. J. Am. Chem. Soc. *129*, 9799–9807.

Meier, S., Blackledge, M., and Grzesiek, S. (2008). Conformational distributions of unfolded polypeptides from novel NMR techniques. J. Chem. Phys. *128*, 052204.

Meiler, J., Prompers, J., 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.

Millett, I.S., Doniach, S., and Plaxco, K.W. (2002). Toward a taxonomy of the denatured state: small angle scattering studies of unfolded proteins. Adv. Protein Chem. *62*, 241–262.

Mohana-Borges, R., Goto, N.K., Kroon, G.J., Dyson, H.J., and Wright, P.E. (2004). Structural characterization of unfolded states of apomyoglobin using residual dipolar couplings. J. Mol. Biol. *340*, 1131–1142.

Mukrasch, M.D., Markwick, P., Biernat, J., Bergen, M.V., Bernadó, P., Griesinger, C., Mandelkow, E., Zweckstetter, M., and Blackledge, M. (2007a). Highly populated turn conformations in natively unfolded tau protein identified from residual dipolar couplings and molecular simulation. J. Am. Chem. Soc. *129*, 5235–5243.

Mukrasch, M.D., von Bergen, M., Biernat, J., Fischer, D., Griesinger, C., Mandelkow, E., and Zweckstetter, M. (2007b). The "jaws" of the Tau-microtubule interaction. J. Biol. Chem. *282*, 12230–12239.

Mukrasch, M.D., Bibow, S., Korukottu, J., Jeganathan, S., Biernat, J., Griesinger, C., and Mandelkow, E. (2009). Structural polymorphism of 441-residue tau at single residue resolution. PLoS Biol. *7*, 399–414.

Neri, D., Billeter, M., Wider, G., and Wuthrich, K. (1992). NMR determination of residual structure in a urea-denatured protein, the 434 repressor. Science *257*, 1559–1563.

Obolensky, O.I., Schlepckow, K., Schwalbe, H., and Solov'yov, A.V. (2007). Theoretical framework for NMR residual dipolar couplings in unfolded proteins. J. Biomol. NMR *39*, 1–16.

Ohnishi, S., Lee, A.L., Edgell, M.H., and Shortle, D. (2004). Direct demonstration of structural similarity between native and denatured eglin C. Biochemistry *43*, 4064–4070.

Ortega-Roldan, J.L., Jensen, M.R., Brutscher, B., Azuaga, A.I., Blackledge, M., and van Nuland, N.A.J. (2009). Accurate characterization of weak macromolecular interactions by titration of NMR residual dipolar couplings: application to the CD2AP SH3-C:Ubiquitin complex. Nucleic Acids Res. *37*, e70.

Prestegard, J.H., Bougault, C.M., and Kishore, A.I. (2004). Residual dipolar couplings in structure determination of biomolecules. Chem. Rev. *104*, 3519–3540.

Rückert, M., and Otting, G. (2000). Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. J. Am. Chem. Soc. *122*, 7793–7797.

Salmon, L., Bouvignies, G., Markwick, P., Lakomek, N., Showalter, S., Li, D.W., 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. Engl. *48*, 4154–4157.

Sass, H.J., Musco, G., Stahl, S.J., Wingfield, P.T., and Grzesiek, S. (2000). Solution NMR of proteins within polyacrylamide gels: Diffusional properties and residual alignment by mechanical stress or embedding of oriented purple membranes. J. Biomol. NMR *18*, 303–309.

Schwalbe, H., Fiebig, K.M., Buck, M., Jones, J.A., Grimshaw, S.B., Spencer, A., Glaser, S.J., Smith, L.J., and Dobson, C.M. (1997). Structural and dynamical properties of a denatured protein. heteronuclear 3D NMR experiments and theoretical simulations of lysozyme in 8M Urea. Biochemistry *36*, 8977–8991.

Schuler, B., and Eaton, W. (2008). Protein folding studied by single molecule FRET. Curr. Opin. Struct. Biol. *18*, 16–26.

Schwarzinger, S., Kroon, G.J.A., Foss, T.R., Chung, J., Wright, P.E., and Dyson, H.J. (2001). Sequence-dependent correction of random coil NMR chemical shifts. J. Am. Chem. Soc. *123*, 2970–2978.

Serrano, L. (1995). Comparison between the φ distribution of the amino acids in the protein database and NMR data indicates that amino acids have various φ propensities in the random coil conformation. J. Mol. Biol. *254*, 322–333.

Shen, Y., Lange, O., Delaglio, F., Rossi, P., Aramini, J.M., Liu, G., Eletsky, A., Wu, Y., Singarapu, K.K., Lemak, A., et al. (2008). Consistent blind protein structure generation from NMR chemical shift data. Proc. Natl. Acad. Sci. USA *105*, 4685–4690.

Shoemaker, B.A., Portman, J.J., and Wolynes, P.G. (2000). Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. Proc. Natl. Acad. Sci. USA *97*, 8868–8873.

Shortle, D. (1996). The denatured state (the other half of the folding equation) and its role in protein. stability. FASEB J. *10*, 27–34.

Shortle, S., and Ackerman, M.S. (2001). Persistence of native-like topology in a denatured protein in 8 M urea. Science *293*, 487–489.

Sibille, N., Sillen, A., Leroy, A., Wieruszeski, J.M., Mulloy, B., Landrieu, I., and Lippens, G. (2006). NMR investigation of the interaction between the neuronal protein tau and the microtubules. Biochemistry *45*, 12560–12572.

Sickmeier, M., Hamilton, J.A., LeGall, T., Vacic, V., Cortese, M.S., Tantos, A., Szabo, B., Tompa, P., Chen, J., Uversky, V.N., et al. (2007). DisProt: the database of disordered proteins. Nucleic Acids Res. *35*, D786–D793.

Skora, L., Cho, M.K., Kim, H.Y., Becker, S., Fernandez, C.O., Blackledge, M., and Zweckstetter, M. (2006). Charge-induced molecular alignment of intrinsically disordered proteins. Angew. Chem. Int. Ed. Engl. *45*, 7012–7015.

Smith, L.J., Bolin, K.A., Schwalbe, H., MacArthur, M.W., Thornton, J.M., and Dobson, C.M. (1996). Analysis of main chain torsion angles in proteins: prediction of NMR coupling constants for native and random coil conformations. J. Mol. Biol. *255*, 494–506.

Spera, S., and Bax, A. (1991). Empirical correlation between protein backbone conformation and C-alpha and C-beta C-13 NMR chemical shifts. J. Am. Chem. Soc. *113*, 5490–5492.

Sugase, K., Dyson, H.J., and Wright, P.E. (2007). Mechanism of coupled folding and binding of an intrinsically unstructured protein. Nature *447*, 1021–1025.

Sung, Y.H., and Eliezer, D. (2007). Residual structure, backbone dynamics, and interactions within the synuclein family. J. Mol. Biol. *372*, 689–707.

Syme, C.D., Blanch, E.W., Holt, C., Jakes, R., Goedert, M., Hecht, L., and Barron, L.D. (2002). A Raman optical activity study of rheomorphism in caseins, synucleins and tau. New insight into the structure and behaviour of natively unfolded proteins. Eur. J. Biochem. *269*, 148–156.

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.

Tjandra, N., Omichinski, J., Gronenborn, A.M., Clore, G.M., and Bax, A. (1997). Use of dipolar 1H-15N and 1H-13C couplings in the structure determination of magnetically oriented macromolecules in solution. Nat. Struct. Biol. *4*, 732–738.

Tolman, J.R. (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.

Tompa, P. (2002). Intrinsically unstructured proteins. Trends Biochem. Sci. *27*, 527–533.

Tompa, P., and Fuxreiter, M. (2008). Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. Trends Biochem. Sci. *33*, 2–8.

Torbet, J., and Maret, G. (1979). Fibres of highly oriented PF1 bacteriophage produced in a strong magnetic field. J. Mol. Biol. *134*, 843–845.

Tycko, R., Blanco, F.J., and Ishii, Y. (2000). Alignment of biopolymers in strained gels: a new way to create detectable dipole–dipole couplings in high-resolution biomolecular NMR. J. Am. Chem. Soc. *122*, 9340–9341.

Ulmer, T.S., Ramirez, B.E., Delaglio, F., and Bax, A. (2004). Evaluation of backbone proton positions and dynamics in a small protein by liquid crystal NMR spectroscopy. J. Am. Chem. Soc. *125*, 9179–9191.

Uversky, V.N. (2002). Natively unfolded proteins: A point where biology waits for physics. Protein Sci. *11*, 739–756.

Vacic, V., Oldfield, C.J., Mohan, A., Radivojac, P., Cortese, M.S., Uversky, V.N., and Dunker, A.K. (2007). Characterization of molecular recognition features, MoRFs, and their binding partners. J. Proteome Res. *6*, 2351–2366.

Vucetic, S., Obradovic, Z., Vacic, V., Radivojac, P., Peng, K., Iakoucheva, L.M., Cortese, M.S., Lawson, J.D., Brown, C.J., Sikes, J.G., et al. (2005). DisProt: a database of protein disorder. Bioinformatics *21*, 137–140.

Wang, L., Eghbalnia, H.R., Bahrami, A., and Markley, J.L. (2005). Linear analysis of carbon-13 chemical shift differences and its application to the detection and correction of errors in referencing and spin system identifications. J. Biomol. NMR *32*, 13–22.

Wang, Y., and Jardetzky, O. (2002). Probability-based protein secondary structure identification using combined NMR chemical-shift data. Protein Sci. *11*, 852–861.

Wells, M., Tidow, H., Rutherford, T.J., Markwick, P., Jensen, M.R., Mylonas, E., Svergun, D.I., Blackledge, M., and Fersht, A.R. (2008). Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. Proc. Natl. Acad. Sci. USA *105*, 5762–5767.

Wishart, D.S., Sykes, B.D., and Richards, F. (1992). The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. Biochemistry *31*, 1647–1651.

Wishart, D.S., Bigam, C.G., Holm, A., Hodges, R.S., and Sykes, B.D. (1995). H-1, C-13 and N-15 random coil NMR shifts of the common amino acids. 1. Investigations of nearest neighbour effects. J. Biomol. NMR *5*, 67–81.

Xie, H., Vucetic, S., Iakoucheva, L.M., Oldfield, C.J., Dunker, A.K., Uversky, V.N., and Obradovic, Z. (2007). Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. J. Proteome Res. *6*, 1882–1898.

Zweckstetter, M., and Bax, A. (2000). Prediction of sterically induced alignment in a dilute liquid crystalline phase: Aid to protein structure determination by NMR. J. Am. Chem. Soc. *122*, 3791–3792.