THE PROTEIN SOCIETY

A maximum entropy approach to the study of residue-specific backbone angle distributions in α-synuclein, an intrinsically disordered protein

Alexey B. Mantsyzov,[1] Alexander S. Maltsev,[1] Jinfa Ying,[1] Yang Shen,[1] Gerhard Hummer,[1,2] and Ad Bax[1]*

[1]Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892

[2]Max Planck Institute of Biophysics, Max-von-Laue-Straße 3, D-60438 Frankfurt am Main, Germany

Received 23 April 2014; Accepted 16 June 2014
DOI: 10.1002/pro.2511
Published online 28 June 2014 proteinscience.org

Abstract: α-Synuclein is an intrinsically disordered protein of 140 residues that switches to an α-helical conformation upon binding phospholipid membranes. We characterize its residue-specific backbone structure in free solution with a novel maximum entropy procedure that integrates an extensive set of NMR data. These data include intraresidue and sequential H^N—H^α and H^N—H^N NOEs, values for $^3J_{HNH\alpha}$, $^1J_{H\alpha C\alpha}$, $^2J_{C\alpha N}$, and $^1J_{C\alpha N}$, as well as chemical shifts of ^{15}N, $^{13}C^\alpha$, and $^{13}C'$ nuclei, which are sensitive to backbone torsion angles. Distributions of these torsion angles were identified that yield best agreement to the experimental data, while using an entropy term to minimize the deviation from statistical distributions seen in a large protein coil library. Results indicate that although at the individual residue level considerable deviations from the coil library distribution are seen, on average the fitted distributions agree fairly well with this library, yielding a moderate population (20–30%) of the PP_{II} region and a somewhat higher population of the potentially aggregation-prone β region (20–40%) than seen in the database. A generally lower population of the α_R region (10–20%) is found. Analysis of ^1H—^1H NOE data required consideration of the considerable backbone diffusion anisotropy of a disordered protein.

Keywords: diffusion anisotropy; intrinsically disordered proteins; Karplus curve; random coil; short-range NOE

Introduction

The study of the backbone conformational distribution of intrinsically disordered proteins (IDP) has attracted considerable interest in recent years,[1–6] building on extensive prior work that analyzed backbone torsion angle propensities in synthetic peptides.[7–12] The realization that many of the IDPs can form amyloids, which are closely linked to a wide range of diseases, has added impetus to their study.[13,14] α-Synuclein (aS) is a 140-residue IDP, with an N-terminal 100-residue region that is mostly

Abbreviations: aS, α-synuclein; IDP, intrinsically disordered protein; NOE, nuclear Overhauser enhancement; PRE, paramagnetic relaxation enhancement.

Additional Supporting Information may be found in the online version of this article.

Grant sponsors: National Institute of Diabetes and Digestive and Kidney Diseases, NIH.

*Correspondence to: Ad Bax, Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892. E-mail: bax@nih.gov

positively charged, and a highly acidic 40-residue C-terminal tail. The N-terminal region binds to negatively charged lipid membranes while adopting an α-helical conformation, whereas the C-terminus remains dynamically highly disordered.[15,16] In free solution, the backbone chemical shifts of the N-terminal region initially were interpreted as indicative of a weak α-helical propensity,[15] but subsequent analysis found this helical population to be rather low.[17] In fact, using refined random coil chemical shift reference values,[18] aS was shown to have chemical shifts that are considerably closer to random coil values than any other IDP studied to date.[19] This means that locally the aS backbone closely resembles a random coil, even though globally the presence of some long-range order is evidenced by compaction of the protein relative to an idealized random coil, as well as by long-range paramagnetic relaxation effects.[20–22] Presumably, this compaction results from weak, nonspecific electrostatic interaction between the protein's oppositely charged N- and C-terminal regions,[23,24] and is rather different in character from the type of compaction observed, for example, in a destabilized triple mutant of the protein Im7, which also exhibits close to random coil chemical shifts.[25]

A number of increasingly sophisticated methods have been developed to present an ensemble description for this IDP.[22,26,27] Effective programs have also been developed to predict structural propensity from NMR chemical shifts,[17,26,28] but these programs rely on relatively few experimental data and focus on secondary structure, i.e., cooperative formation of structural elements. A combination of the Flexible-Meccano[29] and ASTEROIDS[30] programs can generate ensembles of random coil structures,[27] and has been applied to the study of non-native states of hen lysozyme, using a large number of J couplings and chemical shifts, but no NOE data.[31] Alternatively, the program Ensemble[32–35] is widely used to select conformers from backbone models generated with the program TraDES.[36] A recently established database, pE-DB, for deposition and retrieval of such ensemble models and their underlying restraint reflects the widespread interest in gaining structural insights for IDPs and unfolded states of proteins in general.[37] However, although ensemble models of IDPs can be insightful for providing a pictorial view of the backbone conformations sampled by such a protein, it is important to realize that the majority of input parameters commonly used for selecting members from TraDES or Flexible–Meccano ensembles strictly report only on local backbone propensities. For example, homo- and heteronuclear J couplings simply reflect intervening torsion angles, and chemical shifts are dominated by the intraresidue backbone torsion angles, with only a minor effect related to torsion angles in the immediately adjacent residues. Residual dipolar couplings (RDCs), although commonly used as global parameters in well-ordered proteins, primarily report on the "extendedness" of the local backbone in disordered proteins.[38–41] Sources of true long range information can include NOEs, chemical shift perturbation, and paramagnetic relaxation enhancement (PRE). However, long-range NOEs in IDPs often are exceedingly weak and therefore difficult to identify unambiguously, and they only have been used rather infrequently to restrain IDP backbone ensembles.[34,42,43] Chemical shift perturbation (CSP), which reports on small effects of a conservative mutation made at one location in the sequence on chemical shifts of distant residues, can be an exquisitely sensitive method for identifying transient long-range contacts.[44] However, it lacks a quantitative relation to link the ensemble fraction and distance to the magnitude of the CSP, and therefore is primarily qualitative in nature. Similarly, PRE provides a very sensitive measure for transient long-range contacts in a disordered protein, but analogous to the NOE effect, its quantitative use is complicated by the fact that the PRE is a function not only of distance and population but also of the applicable spectral densities.[45–49] In the present study, we therefore limit ourselves to characterizing only the local backbone torsion angle populations that can be derived from NMR observables, where conformations sampled by a given residue can simply be described in terms of a moderate number of voxel populations in a two-dimensional Ramachandran map, rather than the exponentially larger space sampled by the full protein backbone. As experimental restraints, we use ^{1}H, ^{13}C, and ^{15}N chemical shifts, four different types of J couplings, and short-range NOEs.

In principle, the aforementioned problem with the unknown spectral densities in IDPs, needed for converting NOEs into distance restraints, can be addressed by integrating the NOE data into a molecular dynamics calculation, which restrains the distances and spectral density terms in the trajectory to be consistent with the NMR data.[43,50] In practice, however, generating trajectories of sufficient length, while minimizing their dependence on the parameterization of the force field used, is not yet computationally tractable with commonly available resources. Therefore, we here introduce an approximate, empirical method for correlating short-range NOE intensities to internuclear distances in IDPs. Molecular dynamics trajectories on a 40-residue fragment of aS quantitatively support the anisotropic rotational diffusion of the protein backbone, which underlies our empirical method for converting ^{1}H-^{1}H NOE intensities into internuclear distances.

While in the present study we collected all NMR data on full length aS, only strictly local parameters were measured at high precision. As a consequence,

analysis of structural preferences of the protein is restricted to local backbone angle propensities, and not intended to generate ensembles of realistic full chain models. Our analysis is therefore largely analogous to prior investigations of short, disordered peptides, which have been studied in great detail over the past few decades, both by NMR, infrared, and Raman spectroscopies,[6,10,11,51,52] as well as by SAXS.[53] Some of these prior results argue for a high fractional population of polyproline II (PP_{II}) conformations,[5,54] but other results contradict this conclusion.[6,53] For small peptide fragments of hen lysozyme, conformational propensity was found to depend on peptide length, indicative of a threshold in size needed to switch from random coil to a partially ordered structure.[10] By contrast, for aS we find that chemical shifts of peptide fragments fall very close to those of the full length protein, confirming the close similarity in backbone angles sampled by the full length protein and shorter peptide fragments of it. Our data therefore provide a detailed view of the propensities of the backbone chain in the context of a full length IDP, not impacted by end effects. At the same time, the large number of residues in aS for which detailed NMR measurements were made permits an evaluation of the statistical variation within a given residue type, and between different residues.

With only 10 observables available to define the ensemble of backbone torsion angles sampled by each residue on a ~100-voxel grid, experimental information is insufficient to uniquely define the angular distributions. However, by using a maximum entropy computational approach, we show that it is readily possible to define residue-specific backbone torsion angle distributions that deviate only moderately from those seen in a statistical coil library, while greatly improving the fit to experimental data over what the library populations would predict.

Results

Most past analyses of unstructured peptides in terms of conformational ensembles aimed at describing the system as a sum of a very limited number of discrete conformers, typically β, PP_{II}, and α_R, and sometimes also including α_L or turn conformations. Populations extracted, and agreement with NMR data, then depend strongly on the precise backbone torsion angle values chosen for each of these conformers. Here, we aim to describe the peptide as a summation of conformers that span all of the torsion angle space seen in coil library databases. To this extent, we have divided the Ramachandran ϕ/ψ map into a grid of $15^\circ \times 15^\circ$ voxels or cells, and for each cell the coil library[55] segment containing the residue whose ϕ/ψ angles are closest to the center of the cell is selected as the representative conformer. This procedure is carried out for each residue type, and only cells populated above a threshold (≥3 residues) are retained, while the number of conformers found for each cell in the coil library is stored separately for each residue type. On average, about 100 cells are populated above the threshold for each residue type, and the NMR-related parameters (interproton distances, r_{HH}, J-values, and chemical shifts) for the residue closest to the center are calculated. In principle, the distribution sampled by any residue in aS then can be determined by assigning weights, $w_k (\Sigma_k\ w_k = 1)$, to the cells such that the sum over the calculated NMR-parameters optimally matches the experimental data. The calculated NMR weight-averaged observable, $I_{calc}\ (q)$, for any given residue in the sequence is then given by:

$$I_{calc}(q) = \sum_{k=1}^{N_c} w_k I_k(q) \qquad (1)$$

where $I_k(q)$ is the value calculated for cell k, and q indicates the type of observable. N_c is the number of populated cells. The normalized residual in the total fit, averaged over all $N_q = 10$ restraint types, for any given residue is defined by:

$$\chi^2 = \frac{1}{N_q} \sum_{q=1}^{N_q} \left(I_{calc}(q) - I_{exp}(q)\right)^2 / \sigma^2(q) \qquad (2)$$

where $\sigma(q)$ is the estimated uncertainty in $I_{exp}(q)$ (SI Text). In practice, minimizing χ^2 is an ill-determined problem as there are $N_c \approx 100$ weights to be defined by 10 observations per residue. Therefore, it is desirable to regularize the weight calculations, and to search for the distribution of conformers that tries to follow the natural distribution found in the coil library as close as possible while also satisfying the experimental data. For this purpose we use the principle of maximum entropy, as described below.

As experimental restraints, we use the backbone $^{13}C^{\alpha}$, $^{13}C'$, and ^{15}N chemical shifts, four different types of J couplings, and intraresidue and sequential NOEs. Incorporation of chemical shift terms into Eq. (2) is closely analogous to procedures followed by others,[31,34,56,57] and is briefly summarized in the Materials and Methods section. Inclusion of the $^3J_{HNH\alpha}$, $^1J_{H\alpha C\alpha}$, $^2J_{C\alpha N}$, and $^1J_{C\alpha N}$ coupling information is also analogous to earlier work, but with some adaptations discussed below. Analysis of NOEs to derive local structural information in IDPs has only been used sparingly, in part because such NOEs tend to be relatively weak, but also because the anisotropy of the IDP chain motion makes it inappropriate to simply assign a uniform, isotropic rotational correlation time to the system. In principle, it is possible to use such

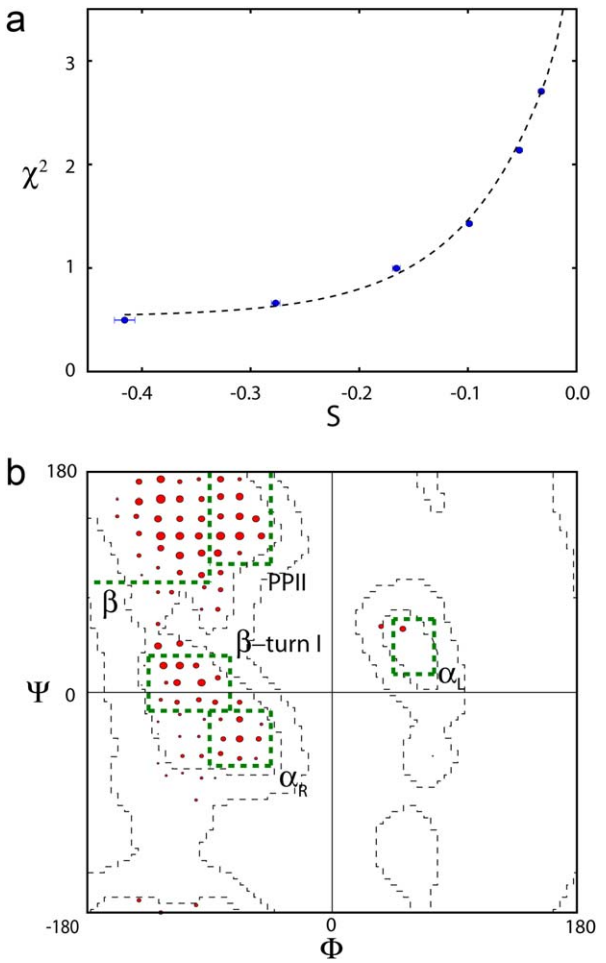

Figure 1. Example of distribution fitting of a residue's backbone angles to 10 independent experimental NMR parameters, shown for V40. (a) Plot of χ^2 versus S, for calculations carried out at θ values of 0.4, 0.8, 1.6, 3, 6, and 10 (left to right). $S = 0$ corresponds to a ϕ/ψ distribution that matches that of the coil database. (b) Backbone conformational distribution at $\theta = 0.8$. The population of each conformer is proportional to the area of the corresponding red circle. Green boxes mark secondary structure regions, β, PP_{II}, type I β-turn, α_R, and α_L. Results shown in both panels represent averages over eight simulated annealing runs. Conformer populations for all nonGly/Pro residues with complete sets of 10 NMR parameters are shown in Supporting Information Figure 2.

NOEs to restrain a dynamics trajectory that encodes both interproton distances as well as the applicable spectral density terms,[42,50] but in practice this proved impossible for a system as large as aS. In this work, we introduce an alternative approach to derive applicable spectral density terms, based on analysis of ^{15}N relaxation rates, while using an empirical orientational dependence of the interproton vector relative to the C^{α}—C^{α} chain direction to account for diffusion anisotropy of the chain. As discussed below, the latter is supported by unrestrained dynamics calculations on a long fragment of aS in a bath of explicit water.

Ramachandran Grid Populations Using Experimental Data and Maximum Entropy

To prevent overfitting of the experimental data while determining the populations of Ramachandran map voxels sampled by any given residue, we use the maximum entropy criterion. Implementation is in full analogy to its recent use for identifying ensembles that best agree with SAXS data.[58] The entropy term is defined by:

$$S = -\sum_{k=1}^{N_c} w_k \ln \frac{w_k}{w_k^{(0)}} \tag{3}$$

where $w_k^{(0)}$ is the reference weight for the conformer k, derived from the population of cell k in the coil library. In an effective free energy function,

$$G = \chi^2 - \theta S, \tag{4}$$

the parameter θ then controls the entropy impact on the effective energy, χ^2. The free energy function G is minimized by means of a simulated annealing algorithm.[58] Use of a high θ-value (e.g., $\theta = 50$) results in a distribution that closely follows the coil library distribution, whereas a small value (e.g., $\theta = 0.1$) can result in overfitting and poor convergence, with sometimes high populations of cells that fall close to the edges of the allowed region in the Ramachandran map.

The optimal value of θ varies somewhat from residue to residue, and can be selected by plotting χ^2 as a function of S [Fig. 1(a)]. When repeating calculations for increasing values of θ, the value that shows a modest increase (the larger of 0.25 or 25% of the χ^2 value obtained for $\theta = 0.4$) is selected to define the ensemble that best describes the experimental data [Fig. 1(b)].

J couplings for restraining ϕ/ψ space

Structural parameters that can be measured at high accuracy for aS include the $^3J_{HNH\alpha}$, $^1J_{C\alpha H\alpha}$, $^1J_{C\alpha N}$, and $^2J_{C\alpha N}$ values. The $^3J_{HNH\alpha}$ values in an IDP were previously measured at very high precision (<0.05 Hz)[19] and were shown to be quite insensitive to residue type, H-bonding effects, and structural variables other than ϕ, as evidenced by an RMSD of about 0.4 Hz between experimental values and those predicted by a Karplus curve when using an RDC-refined high resolution X-ray structure.[59,60] $^1J_{C\alpha H\alpha}$ couplings are sensitive to both ϕ and ψ, and residue-specific random coil values were reported previously.[61] $^1J_{C\alpha H\alpha}$ ranges from ~134 Hz in the α_L region to ~148 Hz for α_R.[62]

$^1J_{C\alpha N}$ and $^2J_{C\alpha N}$ also have a Karplus-like dependence on ψ.[10,63] Although their values vary by less than about 3 Hz across the entire range of

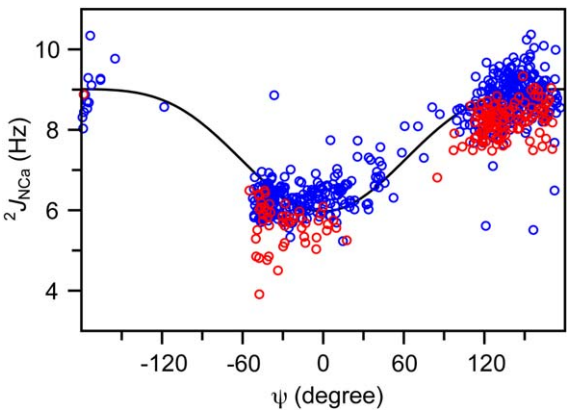

Figure 2. Plot of $^2J_{NC\alpha}$ values, previously reported by Schmidt *et al.*[66] for a set of six proteins of known structure, supplemented by values measured by us for protein GB3 (unpublished data), against the intervening torsion angle ψ, taken from the corresponding high resolution X-ray structure. Values are shown only for residues with backbone chemical shift values that are consistent with the X-ray structure, as judged by the program TALOS-N.[67] Red symbols correspond to Val, Ile, Thr, and Ser residues. Blue symbols are shown for all other residues. The solid line corresponds to $^2J_{NCa} = 8.15 - 1.51\ \cos(\psi) - 0.66\ \cos^2(\psi)$ Hz, where ψ is the torsion angle of the residue on which $^{13}C^{\alpha}$ resides.

ψ angles, they can be measured at very high precision[63,64] and have proven to be useful for characterizing structures of disordered peptides and proteins.[10,31] For $^1J_{NC\alpha}$, the Karplus equation parameterization of Wirmer and Schwalbe was used.[63,65] For $^2J_{NC\alpha}$, the Karplus equation parameterization of Ding and Gronenborn is commonly used.[63] However, we noticed a small but systematic residue-dependent offset between observed and best-fitted $^2J_{NC\alpha}$ data when generating ensembles of conformers that aimed to simultaneously fit all 10 NMR observables (4 types of J couplings, 3 chemical shifts, and 3 NOEs). The problem was most apparent for the β-branched Val, Ile, and Thr residues, as well as Ser. Evaluation of $^2J_{NC\alpha}$ values reported by Schmidt *et al.*[66] for a set of proteins of known structure confirms the presence of a small systematic difference between these four residue types versus all others (Fig. 2). We therefore used a slightly amended $^2J_{NC\alpha}$ Karplus equation:

$$^2J_{NCa} = -1.51\cos(\psi) - 0.66\cos^2(\psi) + C\ \text{Hz} \qquad (5)$$

where $C = 7.65$ for Val, Ile, Thr, and Ser, and $C = 8.15$ for the remaining residues.

Chemical shifts for restraining ϕ/ψ space

Backbone ^{13}C and ^{15}N chemical shifts depend on both ϕ and ψ, and here we use the empirically determined (ϕ,ψ)-dependence[68] of backbone ^{15}N, $^{13}C'$, and $^{13}C^{\alpha}$ shifts as additional restraints (Material and Methods section). ^1H chemical shifts, which show a weaker (ϕ,ψ)-dependence than ^{15}N and ^{13}C and can be substantially impacted by ring current effects, were not included in our analysis. Experimental chemical shifts used in the present study were taken from Maltsev *et al.*[19]

Backbone dynamics from ^{15}N relaxation

The backbone dynamics of aS in the absence of lipids has been evaluated previously,[69,70] but was repeated in our study to take advantage of gains in NMR sensitivity and resolution made over the past decade, and to probe the potential presence of conformational exchange contributions to the transverse relaxation rates. The ^{15}N relaxation data (Supporting Information Table I), recorded at 500 and 900 MHz ^1H frequency, permit mapping of the spectral densities and now show a smoother profile (Fig. 3), but the previously identified increased mobility for the fibril-implicated NAC region[69] remains evident in decreased ^{15}N-{^1H} NOE and ^{15}N R_2 rates (Supporting Information Table II). Comparison of the ^{15}N R_2 rates recorded at 11.7 and 23 Tesla shows no evidence for any slow or intermediate exchange contributions, confirming that the $J(0)$ spectral densities extracted from the relaxation rates by reduced spectral density mapping[71] are not contaminated by exchange effects. Remarkably, even though the high frequency (50 and 435 MHz) spectral densities are quite homogeneous across the entire sequence, $J(0)$ values vary by more than a factor of two? The lowest $J(0)$ values are found in the NAC region, with highest values for residues in, and immediately following the Pro[117]-Val-Asp-Pro region in the acidic C-terminal tail. Below, these $J(0)$ values will be used

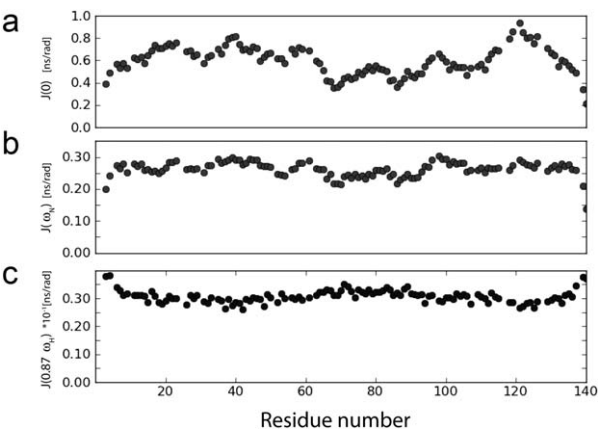

Figure 3. Spectral densities for backbone amide ^{15}N—^1H pairs in aS at 500 MHz, 15°C, obtained from reduced spectral density mapping of the relaxation rates listed in Supporting Information Table II. (a) $J(0)$, (b) $J(\omega_N)$, and (c) $J(0.87\omega_H)$, for $\omega_N = 50.6 \times 2\pi$ rad/s and $\omega_H = 499.5 \times 2\pi$ rad/s. The spectral density values are also listed in Supporting Information Table III.

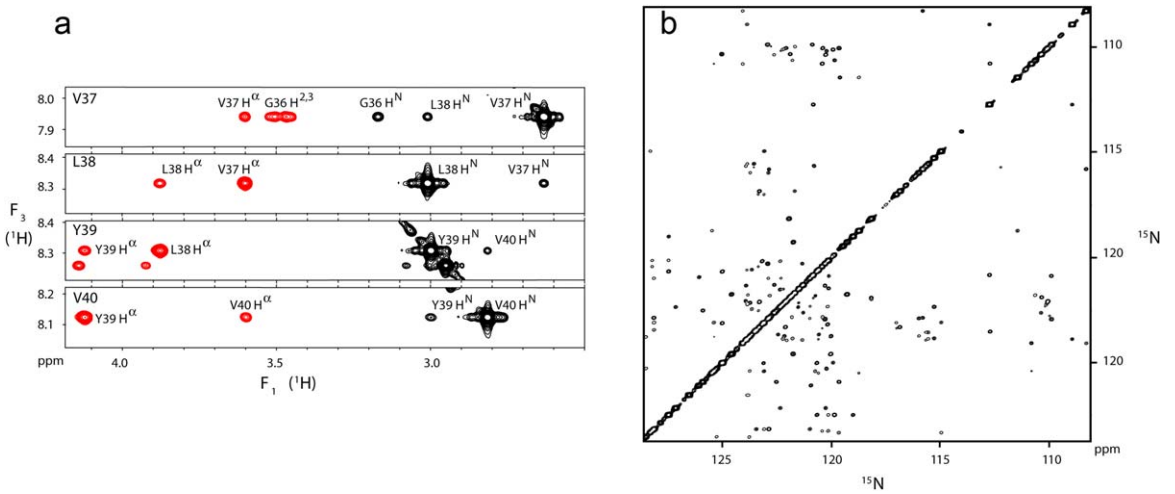

Figure 4. Examples of NOESY spectral data (100 ms NOE mixing; 15°C) recorded for aS. (a) Small regions of strips taken from the 900 MHz 3D ^{1}H—^{15}N—^{1}H NOESY-HSQC spectrum. To achieve improved digital resolution, a narrow F_1 spectral window (5.3 ppm) was used, resulting in aliasing and opposite signs of the amide signals (black contours) relative to the aliphatic signals (red). (b) Partial projection of the ^{15}N—^{15}N—^{1}H 3D NOESY spectrum[72] of aS (800 MHz) on the ^{15}N—^{15}N (F_1, F_2) plane, displaying H^N—H^N NOEs. The projection extends from 8.15 to 8.61 ppm in the ^{1}H dimension.

for extracting distance information from the ^{1}H—^{1}H NOE data.

Measurement of ^{1}H—^{1}H NOE data

Even though $J(0)$ values in IDPs are small relative to those of a globular protein the same size, an extensive set of intraresidue and sequential NOEs is readily observed for aS (Fig. 4). The intraresidue H^N—H^α distance solely depends on the backbone angle ϕ and varies little (2.9 ± 0.15 Å) in the most-populated, negative ϕ region of the Ramachandran map. ^{1}H—^{1}H NOE cross relaxation rates are proportional to the $\{J(0) - 6J(2\omega_H)\}$ spectral density difference.[73] Consequently, the corresponding NOEs are found to correlate closely with this term, which was derived independently from ^{15}N NMR relaxation (Fig. 5). H^N—H^α crossrelaxation rates measured at 600 and 900 MHz are very similar to one another, indicating that the impact of the $J(2\omega_H)$ term is negligible for the sequential H^α—H^N NOE [Fig. 6(b)]. However, the intraresidue H^N—H^α NOE is about 10% weaker at 600 MHz as compared to 900 MHz [Fig. 6(a)], pointing to a small contribution of the $6J(2\omega_H)$ spectral density term. However, even at 500 MHz ^{1}H frequency, the term $J(0.87\omega_H)$ is already more than ~15-fold smaller than $J(0)$ (Fig. 3), and $J(0.87\omega_H)$ rapidly decreases with increasing magnetic field strength. Therefore, at 900 MHz, the $6J(2\omega_H)$ term becomes much smaller than $J(0)$ and may be safely ignored. However, as discussed below, anisotropy of the overall chain dynamics and its internal motion potentially complicate the NOE analysis. In particular, the so-called γ-motions,[74] which correspond to peptide plane oscillations around the C^α_{i-1}—C^α_i chain direction, are expected to dominate the internal dynamics. Motions that correspond to sampling of backbone torsion angles in the β and PPII regions of Ramachandran space lack distinct energy barriers and are expected to be diffusion-limited in their rates, and to be of comparable amplitudes along the chain. Indeed, as can be seen from Figure 3(c), and also from inspection of $J(0.87\omega_H)$ at 900 MHz ^{1}H frequency (Supporting Information Table III), these high frequency spectral density terms show remarkably little residue-by-residue variation along the protein chain, pointing to quite homogeneous amplitudes and time scales of these γ-motions.

In contrast to the intraresidue H^N—H^α NOEs, sequential H^α—H^N NOEs show a considerable range of variation superimposed on the correlation with $J(0) - 6J(2\omega_H)$ [Fig. 5(c)]. In folded proteins, the sequential H^α—H^N distances exhibit a much larger spread than the intraresidue H^N—H^α pairs. The large variation in sequential H^α—H^N NOEs observed for residues of any given $J(0) - 6J(2\omega_H)$ value therefore indicates that they are significantly impacted by residue-by-residue variations in applicable distance distributions, i.e., by conformational propensities.

If motions of dipeptide units in an IDP were isotropic, it would be straightforward to extract the $<r_{HH}^{-6}>$ term applicable for the sequential H^α—H^N NOE from the ratio of the sequential and intraresidue H^N—H^α NOE, using the approximately invariant intraresidue H^N—H^α distance as an internal reference. Indeed, in folded proteins the ratio of the sequential and intraresidue H^α—H^N NOE intensities, $d_{\alpha N}(i-1,i)/d_{\alpha N}(i,i)$, is a sensitive measure for the ψ angle of residue $i-1$.[75] However, such an

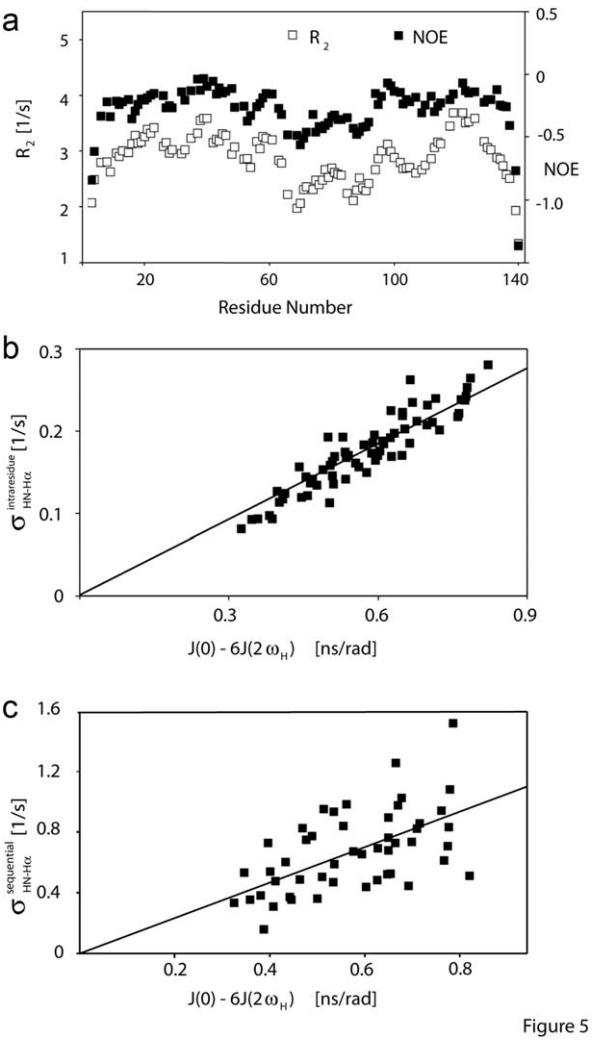

Figure 5

Figure 5. Variation in backbone dynamics of aS, reflected in substantial variations in (a) ^{15}N transverse relaxation rates, R_2 (open symbols), and ^{15}N—{^{1}H} NOE values (both at 500 MHz), and in the wide range of (b) intraresidue H^N—H^α cross relaxation rates, $\sigma_{HN—H\alpha}$ (at 900 MHz), which closely correlate with the spectral densities derived from ^{15}N-relaxation and (c) sequential H^α—H^N cross relaxation rates, which also scale with ^{15}N-relaxation derived spectral densities. Considerable scatter in (c) is indicative of residue-by-residue variation in the interproton distance distribution. Sample conditions: 0.2 m*M* ^{15}N-enriched aS; pH 6, 20 m*M* sodium phosphate, 288 K.

analysis assumes the same spectral densities to be applicable for the intraresidue and sequential H^α—H^N interaction, an assumption that is not compatible with the very high values of the $d_{\alpha N}(i-1, i)/d_{\alpha N}(i,i)$ ratios, in the 3–6 range, typically observed in IDPs.

The $J(0)$ term, which dominates the ^{1}H—^{1}H NOE, is expected to depend strongly on the orientation of the ^{1}H—^{1}H vector relative to the chain direction. This anisotropy is impacted both by the shape of an extended chain, which even for short peptides results in distinct rotational diffusion anisotropy,[76] as well as by the differential sensitivity to internal motions, in particular the above noted γ-motions. In the absence of a detailed motional model of the polypeptide chain of an IDP, we here introduce an *ad hoc* functional form for the angular dependence of the $J(0)$ term applicable for the sequential $^1H^\alpha_{i-}$ $^1H^N_{i+1}$ NOE:

$$J(0)=J(0)_{\exp}\left\{1+1.5e^{-\left[\frac{\psi-120}{45}\right]^2}+1.5e^{-\left[\frac{\psi+240}{45}\right]^2}\right\} \quad (6)$$

where $J(0)_{exp}$ corresponds to the $J(0)$ spectral density derived from ^{15}N relaxation, and ψ is the backbone torsion angle of residue i. Equation (6) corresponds to an about 2.5-fold anisotropy, with the highest $J(0)$ value for extended conformers ($\psi \approx 120^\circ$) and near $J(0)_{exp}$ values for $\psi \approx -60^\circ$. Spectral densities for the intraresidue $^1H^N_i$—$^1H^\alpha_i$ and sequential $^1H^N_{i-1}$—$^1H^N_i$ vectors, which are at relatively large angles from the C^α_{i-1}—C^α_i vector, are approximated by $J(0)_{exp}$, the experimental $J(0)$ spectral density derived for $^{15}N_i$—^1H_i pairs that to a first

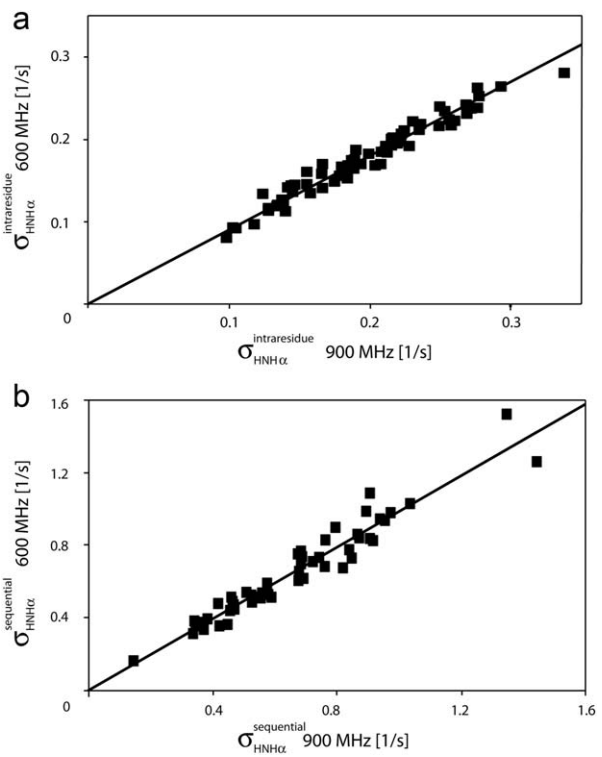

Figure 6. Correlation plot for the (a) intraresidue H^N_i—H^α_i and (b) sequential H^α_{i-1}—H^N_i crossrelaxation rates measured at 600 and 900 MHz ^{1}H frequency. For (b), which involves short interproton distances when the vector is approximately parallel to the C^α_{i-1}—C^α_i vector, the slope equals 0.99, indicating that the impact of the $6J(2\omega_H)$ term is negligible, considering that $J(2\omega_H)$ is estimated to be about two-fold smaller at 900 MHz compared to 600 MHz ^{1}H frequency. For the intraresidue H^N_i—H^α_i interaction, which makes a large angle with the $C^\alpha_{i-1} - C^\alpha_i$ vector, the slope is ~0.9, indicating that the $6J(2\omega_H)$ term is small (~10% of $J(0)$).

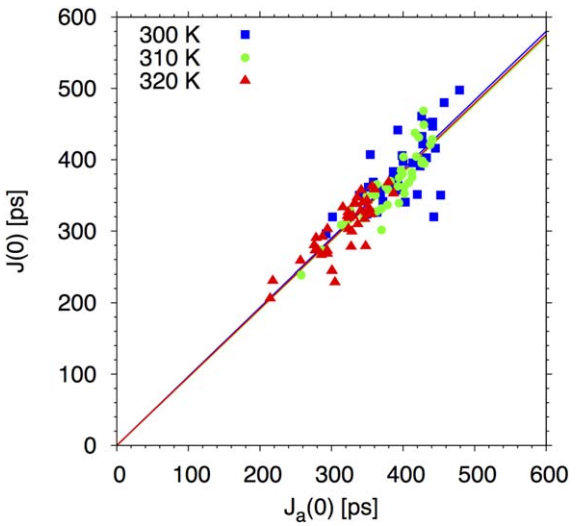

Figure 7. Scatter plot of the MD-derived total spectral density $J(0)$ (y-axis) and of the spectral density $J_a(0)$ for angular motion alone (x-axis) for sequential H^{α}—H^{N} couplings. The symbols show results for the individual residues at each of the three simulation temperatures (blue squares: 300 K; green circles: 310 K; red triangles: 320 K). The lines show least-squares straight-line fits of the form $J(0) = cJ_a(0)$ for each of the temperatures. The average slope of $c = 0.96 \pm 0.01$ (with the error determined by the bootstrap method) indicates that rotational dynamics dominates the relaxation.

approximation are orthogonal to C^{α}_{i-1}—C^{α}_{i}. Below, we demonstrate by evaluation of extended molecular dynamics trajectories of a polypeptide fragment of aS, that Eq. (1) provides a reasonable first order approximation for the impact of $J(0)$ anisotropy on the ^{1}H—^{1}H NOE buildup rates.

Dynamic corrections to ^{1}H—^{1}H NOEs in a random coil

Long (ca. 1 μs) molecular dynamics trajectories were generated in explicit water for a peptide fragment comprising the 40 N-terminal residues of aS, at three temperatures. Spectral densities were derived from these trajectories as detailed in the Material and Methods section. In Figure 7, we first compare the full spectral density $J(0)$ of sequential H^{α}—H^{N} dipole-dipole couplings, including both distance and angular fluctuations, to the spectral density $J_a(0)$ arising from angular relaxation alone. We find that the dynamic contribution to the full spectral density $J(0)$ is determined almost entirely by the rotational motion, $J(0) \approx J_a(0)$, at all temperatures. The reason is that the distance relaxations $C_r(t)$ of sequential H^{α}—H^{N} vectors are dominated by motions that are at least an order of magnitude slower than the angular motions in the MD simulations.

For computational purposes, we define the dynamic correction to the ratio of spectral densities of sequential and intraresidue H^{α}—H^{N} NOEs, normalized by the respective static $<r^{-6}>$ averages, as:

$$[J_{\mathrm{seq}}(0)/ < r^{-6} >_{\mathrm{seq}}]/[J_{\mathrm{intra}}(0)/ < r^{-6} >_{\mathrm{intra}},] \quad (7)$$

Figure 8 shows a scatter plot of this ratio, calculated for sequential (H^{α}_{i}—H^{N}_{i+1}) and intraresidue interactions (H^{α}_{i+1}—H^{N}_{i+1}) of each residue i for three temperatures, as a function of the weighted population w_i in the extended state evaluated along the simulation trajectory of duration T:

$$w_i(T) = \left\langle e^{-\left(\frac{\psi_i - 120}{45}\right)^2} + e^{-\left(\frac{\psi_i + 240}{45}\right)^2} \right\rangle_T \quad (8)$$

This is the same weight function used to correct the experimental $J(0)$ [Eq. (6)]. With the two nonoverlapping exponentials accounting for periodicity, w_i is effectively bounded between 0 and 1. As can be seen from Figure 8, the correction increases approximately linearly with the weight w_i at all three temperatures. The ratio of the corrections extrapolated to $w_i = 1$ and $w_i = 0$ equals about 2.5, fully consistent with the correction used for the experimental $J(0)$

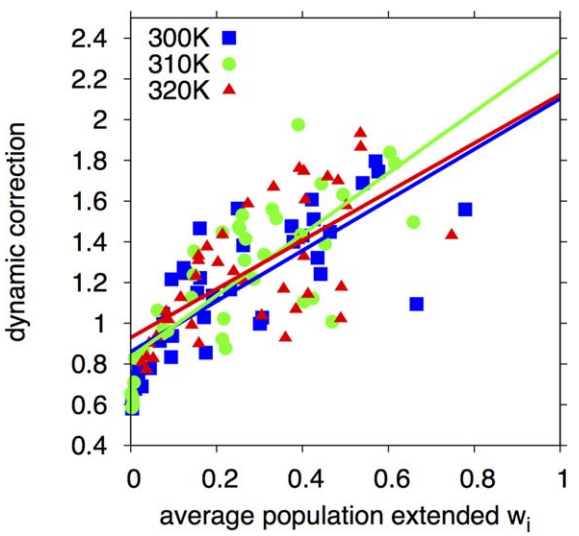

Figure 8. Dynamic correction [Eq. (7)] to $J(0)$ for sequential H^{α}—H^{N} couplings as a function of the weighted population [Eq. (8)] in the extended configuration. Results are shown for each residue at three temperatures (blue squares: 300 K; green circles: 310 K; red triangles: 320 K). The lines show least-squares straight-line fits for each of the temperatures. The ratio of the factors extrapolated to weights of 1 and 0 are 2.45, 2.80, and 2.28 at $T = 300$, 310, and 320 K, respectively. A global fit of all data gives a ratio of 2.51 ± 0.20, with the error determined by the bootstrap method. The simulation estimates for the correction are therefore fully consistent with the functional form of Eq. (6) and the maximum magnitude of 2.5 for the correction factor used in the analysis of the experimental NOE data.

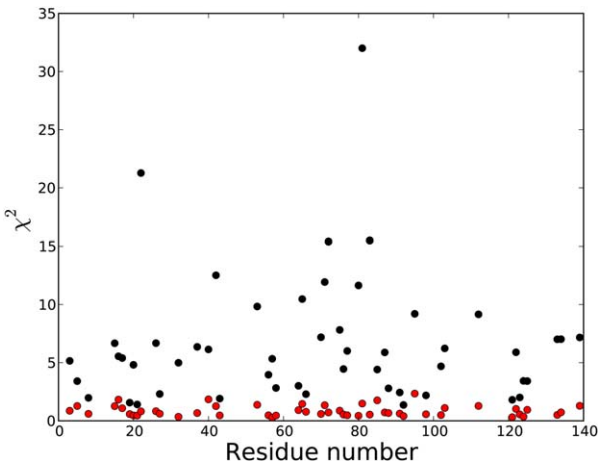

Figure 9. χ^2 as a function of residue number, obtained when using the coil database populations of conformers to predict the experimentally observed parameters (black symbols) and when using the optimized populations of Supporting Information Figure 2 (red).

values [cf Eq. (6)]. The reason why this correction can be extracted directly from the simulation trajectories is that the conformational dynamics associated with radial motions is slow as compared to angular motions, as discussed above. As a result of this time-scale separation, the correlation functions $C(t)$ (Material and Methods section) and thus, by definition, $J(0)$ are well approximated by weighted population averages.

Evaluation of residue-specific ϕ/ψ distributions

For aS, complete sets of all 10 aforementioned parameters ($^3J_{\mathrm{HNH}\alpha}$, $^1J_{\mathrm{C}\alpha\mathrm{H}\alpha}$, $^1J_{\mathrm{C}\alpha\mathrm{N}}$ and $^2J_{\mathrm{C}\alpha\mathrm{N}}$, $\delta_{15\mathrm{N}}$, $\delta_{13\mathrm{C}\alpha}$, $\delta_{13\mathrm{C}\beta}$, $d_{\alpha\mathrm{N}}$ $(i, i+1)$, $d_{\alpha\mathrm{N}}(i,i)$, and $d_{\mathrm{NN}}(i,I+1)$) were obtained for 52 residues, not including 18 Gly and 5 Pro residues. Using Eqs. (2) and (4), and the σ-values defined in the Experimental Section, these values were used to derive lowest free energy ϕ/ψ distributions by means of a simulated annealing protocol, while weakly restraining these distributions to not deviate radically from those seen in the coil database. The weight, θ, used to conform with the statistical database, and implemented through the principle of maximum entropy [cf. Eq. (4)], was varied by approximately doubling θ in successive simulated annealing runs. The distribution that shows a modest increase in total χ^2 when doubling θ (the larger of 0.25 or 25% of the χ^2 value obtained for $\theta = 0.4$) is selected as representative of the ensemble that best describes the experimental data. For most residues, $\theta = 0.4$ or 0.8 yielded optimal results, and fits to the experimental data that for many residues were one to two orders of magnitude better than obtained for coil populations of Ramachandran space (Fig. 9). On the other hand, coil database ϕ/ψ distributions were approximately compatible with experimental data for several residues too, incl. A19, K21, and T92. Distributions of ϕ/ψ for all 52 residues, analogous to those shown in Figure 1(b), are included in the Supporting Information.

Discussion

To compare our results with prior studies of peptides, we group our cells into five regions, β, α_{R}, $\mathrm{PP_{II}}$, type I β-turn, and α_{L} [Fig. 1(b)]. For nearly all residues, the $\mathrm{PP_{II}}$ population falls in the 20–30% range, closely following the statistical coil populations (white bars in Fig. 10), but the β-region is more highly populated in our aS results than in the database for Ala, Asp, Asn, Gln, Glu, Lys, and Thr residues. The small standard deviations seen in Figure 10 indicate that residues of a given type tend to yield similar results (Supporting Information Fig. 3), while in a number of cases deviating significantly from the populations seen in the statistical coil library. Population of the α_{R} region, on average, is lower than seen in the coil library, in agreement with what was found in many of the peptide studies.[7–11] Our results indicate that the type I β-turn region is somewhat less populated than seen in the coil library, but its 10–20% population is nevertheless significant. Neither chemical shifts nor J couplings yield a unique signature for this region, but the short distances between sequential $\mathrm{H^N}$ atoms (2.1–2.4 Å) give rise to substantial amide–amide NOEs, permitting unambiguous identification of their presence. The only other region with short $\mathrm{H^N}$—$\mathrm{H^N}$ distances is α_{R} (~2.8 Å), and a very high population would be required to satisfy the experimentally observed NOE intensity, incompatible with the other observed NMR parameters. In contrast to the type 1 β-turns reported recently in Tau protein,[77] another IDP, in aS the propensity for this turn type is relatively flat across the protein

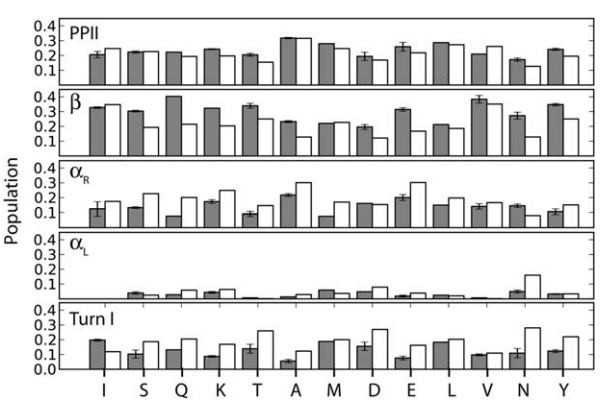

Figure 10. Average populations of the five regions marked in Figure 3(b): β, $\mathrm{PP_{II}}$, type I β-turn, α_{R}, and α_{L}, by residue type as observed in aS (gray bars), with the corresponding population in the coil library of Fitzkee *et al.*[55] shown in white. Residue-specific values are presented in Supporting Information Fig. S3.

sequence. Analogous to the coil library, our data show no evidence for significant population of the α_L region, which carries as its most distinct NMR features a very small $^1J_{H\alpha C\alpha}$ and a strong intraresidue H^N—H^α NOE. Even Asn residues, which show an elevated α_L presence in the coil library (~15%), exhibit low α_L occupancy (5%) in aS. The latter observation also highlights that our results are not unduly biased by the entropy term, which skews our populations towards those of the coil library.

However, our analysis also shows that, on a courser scale, the distribution of backbone torsion angles in aS does not differ drastically from that seen in the statistical coil library obtained from X-ray structures. For nearly one-third of the residues, a reasonable fit ($\chi^2 \leq 3$) to the experimental data (Fig. 9) is obtained when using the ϕ/ψ distribution seen in the coil library. The absence of a dominant PP_{II} population is clearly evidenced by our results, even for Ala17, Ala19, and Ala91, which are part of stretches of three consecutive Ala residues with complete NMR data. On the other hand, our results also indicate that some deviations from coil library populations are present in aS (Fig. 10, Supporting Information Fig. 3) with, on average, the β-region being populated somewhat higher than in the coil library, an effect most pronounced for residues in the fibril-implicated NAC region (residues 61–95). The variation in backbone dynamics along the sequence is quite pronounced, and considerably larger than seen, for example, in fully denatured ubiquitin.[78] In particular, the low $J(0)$ spectral density observed for the NAC region shows it to be considerably more flexible than the remainder of aS (Fig. 3A).

Material and Methods

NMR experiments

3D ^{1}H—^{15}N—^{1}H NOESY-HSQC and ^{15}N—^{15}N—^{1}H HSQC-NOESY-HSQC spectra were recorded with a mixing time of 100 ms at 15°C on a Bruker AV-III 900 MHz spectrometer equipped with a single axis gradient TCI cryogenic probe. The sample contained 0.35 m*M* ^{15}N-enriched aS and 20 m*M* sodium phosphate at pH 6. The two spectra consisted of 1024* (F_3, ^{1}H, 94.7 ms) × 120* (F_2, ^{15}N, 62.2 ms) × 230* (F_1, ^{1}H, 52.7 ms) and 1024* (F_3, ^{1}H, 94.7 ms) × 126* (F_2, ^{15}N, 64.5 ms) × 126* (F_1, ^{15}N, 64.5 ms) complex data points, respectively. After the NOE mixing period, a sensitivity-enhanced, gradient-selected HSQC was employed.[79] To maximize digital resolution, a small spectral width of 4.84 ppm in F_1 was used in the ^{1}H—^{15}N—^{1}H NOESY-HSQC experiment (resulting in aliasing and opposite phases of the amide resonances) and both spectra were recorded using two scans per FID and a two-step phase cycling scheme. In the ^{1}H—^{15}N—^{1}H experiment, the ^{1}H 90° pulses before and after t_1 evolution were phase cycled to yield States-TPPI quadrature detection, and a two-step phase cycling was applied to the ^{15}N 90° pulse prior to t_2, in addition to quadrature detection by gradient-enhanced coherence selection,[79] thereby better canceling the residual solvent signal. The absence of a phase cycle for axial peak suppression in the F_1 dimension resulted in a strong peak at the edge of the indirect ^{1}H dimension (F_1). However, the narrow spectral width was chosen such that the axial artifact, largely removed during the data processing, did not interfere with the cross peaks or folded amide diagonal. Similarly, the first ^{15}N 90° pulse in the first HSQC segment of the ^{15}N—^{15}N—^{1}H HSQC-NOESY-HSQC experiments, was alternated in successive scans, to also yield a two-step phase cycle. 3D NOESY spectra with similar digital resolution and acquisition parameters were also recorded on Bruker AV-II 600 MHz and AV-III 800 MHz spectrometers, both equipped with a cryogenic probe. The NOESY data collected at multiple magnetic fields allowed us to estimate the relative contribution to the NOE intensities of the field-dependent $J(2\omega_H)$ term relative to $J(0)$, but only the better resolved and most accurate 900 MHz data were used as input restraints for conformer selection by the simulated annealing protocol.

Longitudinal ^{15}N R_1 relaxation rates and steady-state heteronuclear ^{15}N—{^{1}H} NOE at 15°C were measured at 900 MHz, using methods described by Lakomek *et al.*[80] The sample contained 0.15 m*M* perdeuterated ^{15}N-enriched aS and 20 m*M* sodium phosphate at pH 6. For the R_1 and ^{15}N—{^{1}H} NOE measurements at 900 MHz, the TROSY readout was used.[80] Six 2D experiments were interleaved, with variable T_1 delays of 0, 200, 400, 600, 800, and 1000 ms, each consisting of 2048* (F_2, ^{1}H, 206.5 ms) × 256* (F_1, ^{15}N, 116.7 ms) complex data points. Using an interscan delay of 1.5 s and 8 scans per FID, the total data recording time was 13 h. For the NOE measurement, two experiments with and without the proton saturation pulse train were interleaved, each consisting of 2048* (F_2, ^{1}H, 206.5 ms) × 310* (F_1, ^{15}N, 141.4 ms) complex data points. In the NOE experiment, after the initial ^{15}N 90° pulse for removing the so-called BEST TROSY effect,[81] the water was first presaturated using an RF field strength of 89 Hz for 1 s, followed by a train of nonselective 180° ^{1}H pulses centered on the amide proton resonances for 7 s, while a total delay of 8 s was used in the reference experiment. The total data collection time was 34 h using 12 scans per FID.

R_1 and ^{15}N—{^{1}H} NOE were also measured on a Bruker AV-III 500 MHz spectrometer with a cryoprobe using the same sample, methods, and comparable acquisition parameters as described above. In addition, transverse ^{15}N R_2 relaxation times were determined at 500 MHz ^{1}H frequency using a $R_{1\rho}$

experiment[80] with a spin lock RF power of 1.4 kHz. Six interleaved 2D spectra with spin lock durations of 10, 20, 80, 160, 270, and 380 ms were recorded, each consisting of 1536* (F_2, ^{1}H, 219.6 ms) × 220* (F_1, ^{15}N, 180.4 ms) complex data points. A long interscan delay of 5 s was used to reduce the amplifier duty cycle and RF heating effects. With 8 scans per FID, the total data recording time was 35 h. The ^{15}N carrier was positioned at 118 ppm, and the R_2 rates were corrected for the off-resonance tilted field, using the relation $R_2 = R_{1\rho}/\sin^2\theta - R_1/\tan^2\theta$ with $\tan\theta = \omega_1/\Omega$, where $R_{1\rho}$ is the directly measured decay rate of the $R_{1\rho}$ experiment, ω_1 is the spin-lock RF field strength and Ω the offset from the ^{15}N carrier.

The measurement of the $^3J_{\mathrm{HNH}\alpha}$ couplings at 900 MHz and 15°C has been described previously.[19] The $^2J_{\mathrm{C}\alpha\mathrm{N}}$ couplings were determined from a 2D high resolution TROSY spectrum without the $^{13}\mathrm{C}^\alpha$ decoupling pulse but with a selective $^{13}\mathrm{C}'$ and an IBURP $^{13}\mathrm{C}^\beta$ decoupling pulse (covering a bandwidth from 35 to 15 ppm), applied in the indirect ^{15}N dimension. The spectrum was recorded on a Bruker AV III 800 MHz spectrometer at 15°C using a sample of 0.34 m*M* perdeuterated and ^{13}C/^{15}N-enriched aS, with acquisition times of 213 (^{1}H, t_2) and 1064 (^{15}N, t_1) ms (2048* × 2048* data matrix). Using an interscan delay of 1.2 s and 16 scans per FID, the measurement time was 36.5 h. At this digital resolution, the couplings of ^{15}N to both the intraresidue and preceding $^{13}\mathrm{C}_\alpha$ were well resolved for nearly all cross peaks, resulting in a doublet of doublet splitting pattern from which two measurements of $^2J_{\mathrm{C}\alpha\mathrm{N}}$ and $^1J_{\mathrm{C}\alpha\mathrm{N}}$ were made for each amide, and the averaged value is used and reported in this study.

The $^1J_{\mathrm{C}\alpha\mathrm{H}\alpha}$ couplings were measured from the doublet splitting in the ^{13}C dimension of a 3D TROSY-HN(CO)CA spectrum, recorded without application of the $^1\mathrm{H}^\alpha$ decoupling pulse during the 28 ms constant-time ^{13}C evolution period, which effectively eliminates the $^1J_{\mathrm{C}\alpha\mathrm{C}\beta}$ splitting for improved spectral resolution. The 3D data matrix consisted of 700* × 30* × 86* complex points for acquisition times of 97.1, 195, and 28 ms in the ^{1}H, ^{15}N, and ^{13}C dimensions, respectively. Note that the ^{15}N dimension was recorded using 30 complex data points only, but with a narrow spectral window of 2.53 ppm to achieve high digital resolution. The ^{15}N acquisition time of 195 ms, much longer than the $^1J_{\mathrm{NCO}}$ refocusing delay of ~25 ms, was realized by using the mixed-time (MT) evolution approach.[82] The total data recording time was 35 h, using a 1 s interscan delay and 8 scans per FID.

All the NMR data were processed using the NMRPipe software[83] and analyzed in NMRDraw and Sparky.[84]

Selection of conformers from the coil library

The coil library used in our study was originally compiled by Fitzkee *et al.*[55] and represents a database of fragments of protein X-ray structures that do not adopt regular α-helical or β-strand secondary structure. From this full coil library, available online, we selected fragments using the following criteria:

- sequence identity between fragments ≤20%,
- resolution of the X-ray structure is 1.6 Å or better,
- refinement factor, *R*, of the X-ray structure ≤0.25.

Fragments were selected from 2093 different protein chains. Selected fragments have been additionally filtered: Only fragments with a length of the coiled region of three or more amino acids were retained. To exclude any influence of adjacent fragments with secondary structure on the distribution, we further excluded the most N- and C-terminal residue for each selected coiled fragment. The number of residues for each type remaining after these procedures were applied is listed in Supporting Information Table IV.

The populations of each of the Ramachandran map voxels that optimally fit the experimental data while minimizing the effective free energy function of Eq. (4) were determined by using a simulated annealing protocol.[58] For each value of θ, and for each residue, the procedure was repeated 5 times, using different random seeds. Calculations were performed on a workstation with two 6-core Intel 2.67 GHz Xeon X5650 processors, using 12 MB of Cache memory per processor, 12 GB of DDR3-1333 ECC-registered memory, and using Hyper-threading. On average, 10,000 steps of simulated annealing minimization took 63 min per residue using one thread.

The minimum size of the voxels used, 15° × 15° was determined by two factors: First, the statistical coil library approach used in our study requires significant population of each voxel in order to remove the impact of statistical fluctuations of the database voxel population, when using the entropy factor to deviate minimally from the database distribution. Second, convergence of the simulated annealing protocol rapidly decreases when using voxel sizes significantly smaller than 15° × 15°.

Calculation of synthetic NMR parameters for the database structures

Calculation of the J-coupling constants was carried out using Karplus-type equations, using the torsion angles taken from the database. $^1J_{\mathrm{C}\alpha\mathrm{H}\alpha}$ values were calculated using the parameterization of Vuister *et al.*[61] with residue-type specific random coil values. The random coil value of Ala was decreased by 1.7 to 142 Hz, as both DFT calculations and our experimental results were not compatible with this value being much higher than that of other residues. For $^3J_{\mathrm{HNH}\alpha}$ the Karplus parameterization of Vogeli *et al.*

was used using the 'rigid model' parameterization. For the small protein GB3, these parameters yielded an RMSD of 0.42 Hz between observed and calculated $^3J_{HNH\alpha}$ values.[59] Following subsequent RDC refinement of the H^N positions,[85] this RMSD decreased to 0.34 Hz. Considering that this RMSD includes the effect of residual uncertainty in ϕ, measurement error in $^3J_{HNH\alpha}$, and potential impact of H-bonding, residue type, and conformational effects other than the ϕ torsion angle, the impact of these last three factors must be considerably less than 0.34 Hz. For $^1J_{NC\alpha}$, the Karplus equation parameterization of Wirmer and Schwalbe was used.[63,65] For $^2J_{NC\alpha}$, the Karplus equation parameterization of Ding and Gronenborn was used,[63] but the value of the constant factor (7.85 Hz) was increased by 0.3 to 8.15 Hz for all residues except Ser and the β-branched Val, Ile, and Thr residues, for which the value was decreased to 7.65 Hz. Without this adjustment, small systematic discrepancies between our experimental values and those resulting from the simulated annealing search remain. Evaluation of $^2J_{NC\alpha}$ values reported by Schmidt *et al.*[66] for a set of proteins of known structure confirmed the presence of this small systematic difference between these two pools of residues.

Prediction of the chemical shifts was carried out using random coil chemical shifts and neighboring residue corrections taken from the program SPARTA.[68] Residue-specific ϕ/ψ-dependence of the secondary chemical shift was extracted from the SPARTA+ database,[86] and the predicted chemical shift was calculated as the sum of the random coil chemical shift, the secondary chemical shift, and the correction for neighboring residues.

The cross relaxation rates for the aS ensemble were calculated from the weighted average of the rates calculated for the corresponding grid points. For each grid point, the ^{1}H-^{1}H cross relaxation rate at 900 MHz was calculated using the equation:

$$\sigma_{HH} = D^2 J(0) r_{HH}^{-6} \quad (9)$$

where r_{HH} is the interproton distance, $J(0)$ the spectral density at zero frequency, $D = h\mu_0\gamma_H^2/(16\pi^2)$ with μ_0 being the permeability of vacuum, γ_H the proton gyromagnetic ratio and h being Planck's constant. We neglected the impact of the $6J(2\omega_H)$ term because this term is very small compared to $J(0)$ (Supporting Information Table SIII). The term $J(2\omega_H)$ at 600 MHz ^{1}H frequency is estimated by propagation from the $J(0.87*900*10^6*2\pi)$ value determined from reduced spectral density mapping.[71]

Weighting of constraints

For calculating the χ^2 value [Eq. (2) main text] the following error values $\sigma(q)$ were used:

Chemical shifts [ppm]: ^{15}N 1.28; $^{13}C'$ 0.4; $^{13}C^\alpha$ 0.4.

J couplings [Hz]: $^1J_{C\alpha H\alpha}$ 0.35; $^2J_{C\alpha N}$ 0.2; $^1J_{C\alpha N}$ 0.2; $^3J_{HNH\alpha}$ 0.15.

^{1}H—^{1}H cross relaxation rate σ_{HH} [%]: 15.

The $J(0)$ spectral density function for vectors parallel to the long axis, as applies for the sequential H_i^α—H_{i+1}^N vector when $\psi = 120°$, i.e., when the internuclear distance is shortest, is given by $J(0) = {}^2/_5\ \tau_\perp$, whereas for vectors orthogonal to the chain direction one has $J(0) = {}^2/_5[{}^1/_4\tau_\perp + {}^3/_4\tau_{//}]$. Here, $\tau_\perp = (6D_\perp)^{-1}$ and $\tau_{//} = (2D_\perp + 4D_{//})^{-1}$, and $D_\perp$ and $D_{//}$ are the rotational diffusion coefficients orthogonal and parallel to the long axis, respectively. When diffusion anisotropy is large, $J(0)$ is dominated by $D_\perp$, both for vectors parallel and orthogonal to the long axis. Therefore, the large degree of scatter observed for the plot of sequential H^α—H^N cross relaxation rate versus $J(0)$ cannot be dominated by variations in $D_\perp$ and instead must be attributed to residue-specific differences in the $<r_{HN-H\alpha}^{-6}>$ distance distribution function.

As can be seen in Figure 3, the residue-by-residue variation in $J(0)$ for the N—H vector, orthogonal to the chain direction, is very small for adjacent residues, indicating that $D_\perp$ does not vary rapidly as a function of residue number. While the correlation between σ_{HH}^{intra} and $J(0)$ is tight [Fig. 5(b)], the large degree of scatter seen when plotting σ_{HH}^{seq} versus $J(0) - 6J(2w_H)$ [Fig. 5(c)] strongly indicates that the variation in σ_{HH}^{seq} is dominated by differences in H—H distance, not by differences in the effective correlation time. We also note that a high degree of diffusion anisotropy necessarily involves a significant number of adjacent residues, and therefore cannot vary rapidly from one residue to the next. The observation that sequential H_i^α—H_{i+1}^N cross relaxation rates show sharp differences between adjacent residues therefore confirms that these variations are dominated by differences in the applicable $<r_{HN-H\alpha}^{-6}>$ distance distribution.

For the intraresidue H^N—H^α interaction the NOE buildup rate quantitatively agrees with the $J(0)$ spectral density derived from ^{15}N relaxation, assuming an $<r^{-6}>$ distance distribution that follows that of the statistical coil library (variations in $<r^{-6}>^{-1/6}$ caused by our derived changes in Ramachandran map population relative to this library are minor for intraresidue H^N—H^α interactions), confirming that the rigid limit applies for the intraresidue H—N—C^α—H^α unit.

We note, however, that the precise functional form of the applicable spectral densities for a random coil is difficult to establish because the coupling between internal motions (rotations about ϕ and ψ) cannot be separated from the overall rotational diffusion. Although in principle it may be possible to carry out molecular dynamics simulations on short

peptides, where the force field is iteratively adjusted to reach agreement with the $^1H—^1H$ cross-relaxation rates and ^{15}N relaxation rates observed experimentally, such an analysis goes well beyond the scope of our current study. However, modern force fields, such as AMBER 99SB*, have been calibrated to yield ϕ/ψ populations that reflect those seen in experiment.[87] Ratios of the applicable spectral densities for sequential and intraresidue $H^{\alpha}—H^{N}$ interactions then may be derived from a molecular dynamics simulation carried out in a large box of explicit water, as described below.

Molecular dynamics simulations

To evaluate the impact of anisotropic diffusion on sequential and intraresidue $H^{\alpha}—H^{N}$ NOEs, we carried out molecular dynamics (MD) simulations of 40-amino-acid N-terminal aS fragment in water, using all-atom explicit-solvent. The peptide was solvated in a box of 17,223 TIP3P water molecules,[88] 9 chloride ions, and 7 sodium ions, resulting in a simulation system of 52,284 atoms. The simulations were performed using GROMACS 4.5.5,[89] with the AMBER99SB*-ILDN force field[87,90–92] that incorporates corrections for the secondary-structure preference[87] and for side-chain dihedrals.[92] In three independent runs, constant temperatures of 300, 310, and 320 K were maintained by means of a Langevin thermostat using a time step of 0.002 ps and a 1/ps friction coefficient. The pressure was held constant at 1 bar using the Parrinello–Rahman barostat[93] with a 1 ps time constant, acting isotropically in the rhombic dodecahedron simulation cell. Long-range electrostatics interactions were treated with particle-mesh Ewald summation,[94] using a cubic-spline interpolation, a ~1.2 Å mesh width, and a 10 Å cutoff for real-space nonbonded interactions. Bond lengths were constrained. Each of the three runs comprised an initial equilibration period of at least 0.19 μs, followed by production runs of 0.82, 0.9, and 0.92 μs at 300, 310, and 320 K.

Calculation of spectral densities

To determine the spectral densities,[50] we calculated the following correlation function:

$$C(\tau)=\langle P_2(\cos(\theta_{t,t+\tau})/r^3(t)r^3(t+\tau)\rangle/\langle r^{-6}\rangle \qquad (10)$$

where $\boldsymbol{r}(t)$ is a proton–proton distance vector depending on time t, $r(t) = |\boldsymbol{r}(t)|$ is the corresponding distance, and

$$\cos(\theta_{t,t+\tau})=r(t)\cdot r(t+\tau)/r(t)r(t+\tau). \qquad (11)$$

is the cosine of the angle between the corresponding unit vectors at times t and $t + \tau$. $C(\tau)$ was then fitted to a sum of two exponentials over the range of 1 to 1000 ps,

$$C(\tau)=a_1e^{-t/t1}+ a_2e^{-t/t2}. \qquad (12)$$

Fits to three exponentials were also performed, with the results being essentially unchanged. From the exponential fits, we obtained the spectral density

$$J(\omega)=2\left\{a_1t_1/\left[1+ (\omega t_1)^2\right]+a_2t_2/\left[1+ (\omega t_2)^2\right]\right\}/5. \qquad (13)$$

We thus have $J(0) = 2(a_1t_1 + a_2t_2)/5$. To separate the contributions to the spectral density arising from radial and angular motions, the following correlation functions were also calculated:

$$C_r(\tau) =\langle r^{-3}(t)r^{-3}(t+\tau)\rangle \qquad (14)$$

and

$$C_a(\tau) =\langle P_2(\cos(\theta_{t,t+\tau})\rangle. \qquad (15)$$

For $C_a(t)$, exponential fits were again used to extract the corresponding spectral density $J_a(\omega)$.

References

1. Sickmeier M, Hamilton JA, LeGall T, Vacic V, Cortese MS, Tantos A, Szabo B, Tompa P, Chen J, Uversky VN, Obradovic Z, Dunker AK (2007) DisProt: the database of disordered proteins. Nucleic Acid Res 35:D786–D793.
2. Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradovic Z (2002) Intrinsic disorder and protein function. Biochemistry 41:6573–6582.
3. Csizmok V, Felli IC, Tompa P, Banci L, Bertini I (2008) Structural and dynamic characterization of intrinsically disordered human securin by NMR spectroscopy. J Am Chem Soc 130:16873–16879.
4. Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Biol 6: 197–208.
5. Shi ZS, Chen K, Liu ZG, Kallenbach NR (2006) Conformation of the backbone in unfolded proteins. Chem Rev 106:1877–1897.
6. Makowska J, Rodziewicz-Motowidlo S, Baginska K, Vila JA, Liwo A, Chmurzynski L, Scheraga HA (2006) Polyproline II conformation is one of many local conformational states and is not an overall conformation of unfolded peptides and proteins. Proc Natl Acad Sci U S A 103:1744–1749.
7. Dyson HJ, Wright PE (1991) Defining solution conformations of small linear peptides. Ann Rev Biophys Biophys Chem 20:519–538.
8. Baldwin RL, Rose GD (1999) Is protein folding hierarchic? I. Local structure and peptide folding. Trends Biochem Sci 24:26–33.
9. Smith LJ, Bolin KA, Schwalbe H, MacArthur MW, Thornton JM, Dobson CM (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.
10. Graf J, Nguyen PH, Stock G, 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.
11. Hagarman A, Measey TJ, Mathieu D, Schwalbe H, Schweitzer-Stenner R (2010) Intrinsic propensities of amino acid residues in GxG peptides inferred from amide I ' band profiles and NMR scalar coupling constants. J Am Chem Soc 132:540–551.
12. Long HW, Tycko R (1998) Biopolymer conformational distributions from solid-state NMR: alpha-helix and 3(10)-helix contents of a helical peptide. J Am Chem Soc 120:7039–7048.
13. Chiti F, Dobson CM (2006). Protein misfolding, functional amyloid, and human disease. Ann Rev Biochem 75:333–366.
14. Dyson HJ, Wright PE (2004) Unfolded proteins and protein folding studied by NMR. Chem Rev 104:3607–3622.
15. Eliezer D, Kutluay E, Bussell R, Browne G (2001) Conformational properties of alpha-synuclein in its free and lipid-associated states. J Mol Biol 307:1061–1073.
16. Bodner CR, Dobson CM, Bax A (2009) Multiple tight phospholipid-binding modes of alpha-synuclein revealed by solution NMR spectroscopy. J Mol Biol 390: 775–790.
17. Camilloni C, De Simone A, Vranken WF, Vendruscolo M (2012) Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts. Biochemistry 51: 2224–2231.
18. Kjaergaard M, Brander S, Poulsen FM (2011) Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. J Biomol NMR 49:139–149.
19. Maltsev AS, Ying JF, Bax A (2012) Impact of N-terminal acetylation of α-synuclein on its random coil and lipid binding properties. Biochemistry 51:5004–5013.
20. Bertoncini CW, Jung YS, Fernandez CO, Hoyer W, Griesinger C, Jovin TM, Zweckstetter M (2005) Release of long-range tertiary interactions potentiates aggregation of natively unstructured alpha-synuclein. Proc Natl Acad Sci U S A 102:1430–1435.
21. Bernado P, Bertoncini CW, Griesinger C, Zweckstetter M, Blackledge M (2005) Defining long-range order and local disorder in native alpha-synuclein using residual dipolar couplings. J Am Chem Soc 127:7968–17969.
22. Allison JR, Varnai P, Dobson CM, Vendruscolo M (2009) Determination of the free energy landscape of alpha-synuclein using spin label nuclear magnetic resonance measurements. J Am Chem Soc 131:18314–18326.
23. Paleologou KE, Schmid AW, Rospigliosi CC, Kim HY, Lamberto GR, Fredenburg RA, Lansbury PT, Fernandez CO, Eliezer D, Zweckstetter M, Lashuel HA (2008) Phosphorylation at Ser-129 but not the phosphomimics S129E/D inhibits the fibrillation of alpha-synuclein. J Biol Chem 283:16895–16905.
24. Wu KP, Weinstock DS, Narayanan C, Levy RM, Baum J (2009) Structural reorganization of alpha-synuclein at low pH observed by NMR and REMD simulations. J Mol Biol 391:784–796.
25. Pashley CL, Morgan GJ, Kalverda AP, Thompson GS, Kleanthous C, Radford SE (2012) Conformational properties of the unfolded state of Im7 in nondenaturing conditions. J Mol Biol 416:300–318.
26. Marsh JA, Singh VK, Jia ZC, Forman-Kay JD (2006) Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation. Protein Sci 15:2795–2804.
27. Salmon L, Nodet G, Ozenne V, Yin GW, Jensen MR, Zweckstetter M, Blackledge M (2010) NMR characterization of long-range order in intrinsically disordered proteins. J Am Chem Soc 132:8407–8418.
28. Tamiola K, Mulder FAA (2012) Using NMR chemical shifts to calculate the propensity for structural order and disorder in proteins. Biochem Soc Trans 40:1014–1020.
29. Ozenne V, Bauer F, Salmon L, Huang J-r, Jensen MR, Segard S, Bernado P, Charavay C, Blackledge M (2012) Flexible-meccano: a tool for the generation of explicit ensemble descriptions of intrinsically disordered proteins and their associated experimental observables. Bioinformatics 28:1463–1470.
30. Nodet G, Salmon L, Ozenne V, Meier S, Jensen MR, 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.
31. Sziegat F, Silvers R, Haehnke M, Jensen MR, Blackledge M, Wirmer-Bartoschek J, Schwalbe H (2012) Disentangling the coil: modulation of conformational and dynamic properties by site-directed mutation in the non-native state of hen egg white lysozyme. Biochemistry 51:3361–3372.
32. Choy WY, Forman-Kay JD (2001) Calculation of ensembles of structures representing the unfolded state of an SH3 domain. J Mol Biol 308:1011–1032.
33. Krzeminski M, Marsh JA, Neale C, Choy W-Y, Forman-Kay JD (2013) Characterization of disordered proteins with ENSEMBLE. Bioinformatics 29:398–399.
34. Marsh JA, Forman-Kay JD (2012) Ensemble modeling of protein disordered states: experimental restraint contributions and validation. Proteins 80:556–572.
35. Marsh JA, Teichmann SA, Forman-Kay JD (2012) Probing the diverse landscape of protein flexibility and binding. Curr Opin Struct Biol 22:643–650.
36. Feldman HJ, Hogue CWV (2000) A fast method to sample real protein conformational space. Proteins 39:112–131.
37. Varadi M, Kosol S, Lebrun P, Valentini E, Blackledge M, Dunker AK, Felli IC, Forman-Kay JD, Kriwacki RW, Pierattelli R, Sussman J, Svergun DI, Uversky VN, Vendruscolo M, Wishart D, Wright PE, Tompa P (2014) pE-DB: a database of structural ensembles of intrinsically disordered and of unfolded proteins. Nucleic Acids Res 42:D326–D335.
38. Louhivuori M, Paakkonen K, Fredriksson K, Permi P, Lounila J, Annila A (2003) On the origin of residual dipolar couplings from denatured proteins. J Am Chem Soc 125:15647–15650.
39. Dames SA, Aregger R, Vajpai N, Bernado P, Blackledge M, Grzesiek S (2006) Residual dipolar couplings in short peptides reveal systematic conformational preferences of individual amino acids. J Am Chem Soc 128: 13508–13514.
40. Meier S, Blackledge M, Grzesiek S (2008) Conformational distributions of unfolded polypeptides from novel NMR techniques. J Chem Phys 128:052204.
41. Obolensky OI, Schlepckow K, Schwalbe H, Solov'yov AV (2007) Theoretical framework for NMR residual dipolar couplings in unfolded proteins. J Biomol NMR 39:1–16.
42. Fawzi NL, Phillips AH, Ruscio JZ, Doucleff M, Wemmer DE, Head-Gordon T (2008) Structure and

dynamics of the A ss(21–30) peptide from the interplay of NMR experiments and molecular simulations. J Am Chem Soc 130:6145–6158.
43. Ball KA, Phillips AH, Nerenberg PS, Fawzi NL, Wemmer DE, Head-Gordon T (2011) Homogeneous and heterogeneous tertiary structure ensembles of amyloid-beta peptides. Biochemistry 50:7612–7628.
44. Bruun SW, Iesmantavicius V, Danielsson J, Poulsen FM (2010) Cooperative formation of native-like tertiary contacts in the ensemble of unfolded states of a four-helix protein. Proc Natl Acad Sci U S A 107:13306–13311.
45. Marsh JA, Neale C, Jack FE, Choy W-Y, Lee AY, Crowhurst KA, Forman-Kay JD (2007) Improved structural characterizations of the drkN SH3 domain unfolded state suggest a compact ensemble with native-like and non-native structure. J Mol Biol 367: 1494–1510.
46. Gillespie JR, Shortle D (1997) Characterization of long-range structure in the denatured state of staphylococcal nuclease. II. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. J Mol Biol 268:170–184.
47. Dedmon MM, Lindorff-Larsen K, Christodoulou J, Vendruscolo M, Dobson CM (2005) Mapping long-range interactions in alpha-synuclein using spin-label NMR and ensemble molecular dynamics simulations. J Am Chem Soc 127:476–477.
48. Francis CJ, Lindorff-Larsen K, Best RB, Vendruscolo M (2006) Characterization of the residual structure in the unfolded state of the Delta 131 Delta fragment of staphylococcal nuclease. Proteins 65:145–152.
49. Esteban-Martin S, Silvestre-Ryan J, Bertoncini CW, Salvatella X (YEAR) Identification of fibril-Like tertiary contacts in soluble monomeric alpha-synuclein. Biophys J 105:1192–1198.
50. Peter C, Daura X, van Gunsteren WF (2001) Calculation of NMR-relaxation parameters for flexible molecules from molecular dynamics simulations. J Biomol NMR 20:297–310.
51. Eker F, Griebenow K, Cao XL, Nafie LA, Schweitzer-Stenner R (2004) Preferred peptide backbone conformations in the unfolded state revealed by the structure analysis of alanine-based (AXA) tripeptides in aqueous solution. Proc Natl Acad Sci U S A 101:10054–10059.
52. Schweitzer-Stenner R (2009) Distribution of conformations sampled by the central amino acid residue in tripeptides inferred from amide I band profiles and NMR scalar coupling constants. J Phys Chem B 113:2922–2932.
53. Zagrovic B, Lipfert J, Sorin EJ, Millett IS, van Gunsteren WF, Doniach S, Pande VS (2005) Unusual compactness of a polyproline type II structure. Proc Natl Acad Sci U S A 102:11698–11703.
54. Shi ZS, Chen K, Liu ZG, Ng A, Bracken WC, Kallenbach NR (2005) Polyproline II propensities from GGXGG peptides reveal an anticorrelation with beta-sheet scales. Proc Natl Acad Sci U S A 102:17964–17968.
55. Fitzkee NC, Fleming PJ, Rose GD (2005) The protein coil library: a structural database of nonhelix, nonstrand fragments derived from the PDB. Proteins 58: 852–854.
56. Ozenne V, Schneider R, Yao M, Huang J-R, Salmon L, Zweckstetter M, Jensen MR, Blackledge M (2012) Mapping the potential energy landscape of intrinsically disordered proteins at amino acid resolution. J Am Chem Soc 134:15138–15148.
57. Kragelj J, Ozenne V, Blackledge M, Jensen MR (2013) Conformational propensities of intrinsically disordered proteins from NMR chemical shifts. ChemPhysChem 14:3034–3045.
58. Rozycki B, Kim YC, Hummer G (2011) SAXS ensemble refinement of ESCRT-III CHMP3 conformational transitions. Structure 19:109–116.
59. Vogeli B, Ying JF, Grishaev A, Bax A (2007) Limits on variations in protein backbone dynamics from precise measurements of scalar couplings. J Am Chem Soc 129:9377–9385.
60. Maltsev AS, Grishaev A, Roche J, Zasloff M, Bax A (2014) Improved cross validation of a static ubiquitin structure derived from high precision residual dipolar couplings measured in a drug-based liquid crystalline phase. J Am Chem Soc 136:3752–3755.
61. Vuister GW, Delaglio F, Bax A (1993) The use of $^{1}J_{C\alpha H\alpha}$ coupling constants as a probe for protein backbone conformation. J Biomol NMR 3:67–80.
62. Vuister GW, Delaglio F, Bax A (1992) An empirical correlation between 1J(CaHa) and protein backbone conformation. J Am Chem Soc 114:9674–9675.
63. Ding KY, Gronenborn AM (2004) Protein backbone H-1(N)-C-13(alpha) and N-15-C-13(alpha) residual dipolar and J couplings: new constraints for NMR structure determination. J Am Chem Soc 126:6232–6233.
64. Delaglio F, Torchia DA, Bax A (1991) Measurement of 15N-13C J couplings in staphylococcal nuclease. J Biomol NMR 1:439–446.
65. Wirmer J, Schwalbe H (2002) Angular dependence of (1)J(N-i,C-alpha i) and (2)J(N-i,C alpha(i-1)) coupling constants measured in J-modulated HSQCs. J Biomol NMR 23:47–55.
66. Schmidt JM, Hua Y, Loehr F (2010) Correlation of (2)J couplings with protein secondary structure. Proteins 78:1544–1562.
67. Shen Y, Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. J Biomol NMR 56:227–241.
68. Shen Y, Bax A (2007) Protein backbone chemical shifts predicted from searching a database for torsion angle and sequence homology. J Biomol NMR 38:289–302.
69. Bussell R, Eliezer D (2001) Residual structure and dynamics in Parkinson's disease-associated mutants of alpha-synuclein. J Biol Chem 276:45996–46003.
70. Wu K-P, Baum J (YEAR) Backbone assignment and dynamics of human alpha-synuclein in viscous 2 M glucose solution. Biomol NMR Assign 5:43–46.
71. Farrow NA, Zhang OW, Szabo A, Torchia DA, Kay LE (1995) Spectral density-function mapping using N-15 relaxation data exclusively. J Biomol NMR 6:153–162.
72. Ikura M, Bax A, Clore GM, Gronenborn AM (1990) Detection of nuclear Overhauser effects between degenerate amide proton resonances by heteronuclear three-dimensional nuclear magnetic resonance spectroscopy. J Am Chem Soc 112:9020–9022.
73. Cavanagh J, Fairbrother WJ, Palmer AG, Rance M, Skelton N (2007) Protein NMR spectroscopy: principles and practice, 2nd ed. Burlington, MA: Elsevier Academic Press.
74. Lienin SF, Bremi T, Brutscher B, Bruschweiler R, Ernst RR (1998) Anisotropic intramolecular backbone dynamics of ubiquitin characterized by NMR relaxation and MD computer simulation. J Am Chem Soc 120: 9870–9879.
75. Gagne SM, Tsuda S, Li MX, Chandra M, Smillie LB, Sykes BD (1994) Quantification of the calcium-induced secondary structural changes in the regulatory domain of troponin-C. Protein Sci 3:1961–1974.
76. Ying J, Roche J, Bax A (2014) Homonuclear decoupling for enhancing resolution and sensitivity in NOE and

RDC measurements of peptides and proteins. J Magn Reson 241:97–102.
77. Mukrasch MD, Markwick P, Biernat J, von Bergen M, Bernado P, Griesinger C, Mandelkow E, Zweckstetter M, Blackledge M (2007) Highly populated turn conformations in natively unfolded Tau protein identified from residual dipolar couplings and molecular simulation. J Am Chem Soc 129:5235–5243.
78. Wirmer J, Peti W, Schwalbe H (2006) Motional properties of unfolded ubiquitin: a model for a random coil protein. J Biomol NMR 35:175–186.
79. Kay LE, Keifer P, Saarinen T (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. J Am Chem Soc 114:10663–10665.
80. Lakomek NA, Ying JF, Bax A (2012) Measurement of 15N relaxation rates in perdeuterated proteins by TROSY-based methods J Biomol NMR 53:209–221.
81. Favier A, Brutscher B (2011) Recovering lost magnetization: polarization enhancement in biomolecular NMR. J Biomol NMR 49:9–15.
82. Ying JF, Chill JH, Louis JM, Bax A (2007) Mixed-time parallel evolution in multiple quantum NMR experiments: sensitivity and resolution enhancement in heteronuclear NMR. J Biomol NMR 37:195–204.
83. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRpipe—a multidimensional spectral processing system based on Unix pipes. J Biomol NMR 6:277–293.
84. Goddard TD, Kneller DG (2008) Sparky 3, University of California, San Francisco.
85. Yao L, Vogeli B, Torchia DA, Bax A (2008) Simultaneous NMR study of protein structure and dynamics using conservative mutagenesis. J Phys Chem B 112:6045–6056.
86. Shen Y, Bax A (2010) SPARTA plus: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. J Biomol NMR 48:13–22.
87. Best RB, Hummer G (2009) Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. J Phys Chem B 113:9004–9015.
88. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79: 926–935.
89. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 4:435–447.
90. Wang JM, Cieplak P, Kollman PA (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? J Comput Chem 21:1049–1074.
91. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C (2006) Comparison of multiple amber force fields and development of improved protein backbone parameters. Proteins 65:712–725.
92. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw DE (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins 78:1950–1958.
93. Parrinello M, Rahman A (1980) Crystal structure and pair potentials—a molecular dynamics study. Phys Rev Lett 45:1196–1199.
94. Darden T, York D, Pedersen L (1993) Particle Mesh Ewald—an N.log(N) method for Ewald sums in large systems. J Chem Phys 98:10089–10092.