Structure-Free Validation of RDC and PRE Measurements of Disordered Proteins

Francisco Newby[1], Alfonso De Simone[2], Maho Yagi-Utsumi[1,3], Xavier Salvatella[1], Christopher M. Dobson[1] and Michele Vendruscolo[1,*]

[1]*Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, UK*

[2]*Department of Life Sciences, Imperial College London, South Kensington, London SW7 2AZ, UK*

[3]*Institute for Integrative Bioscience, National Institutes of Natural Sciences, 5-1 Higashiyama Myodaiji, Okazaki 444-8787, Japan*

[*]Correspondence to:
Prof. Michele Vendruscolo
E-mail: mv245@cam.ac.uk
Phone: +44 1223 763873
Mail: Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, UK

Abstract

Residual dipolar couplings (RDCs) and paramagnetic relaxation enhancements (PREs) have emerged as valuable parameters for defining the structures and dynamics of disordered proteins by nuclear magnetic resonance (NMR) spectroscopy. Procedures for their measurement, however, may lead to conformational perturbations because of the presence of the alignment media necessary for recording RDCs, or of the paramagnetic groups that must be introduced for measuring PREs. We discuss here experimental methods to quantify these effects by considering the case of the 40-residue isoform of the amyloid β peptide (Aβ40), which is associated with Alzheimer's disease. By carrying out RDC measurements over a range of concentrations of given alignment media, we show that perturbations arising from transient binding of Aβ40 can be identified, enabling appropriate corrections to be made. In addition, by using NMR experiments sensitive to long-range interactions we show that it is possible to identify relatively non-perturbing sites for attaching nitroxide radicals for PRE measurements. Thus, controlling the conformational perturbations introduced by RDC and PRE measurements should facilitate their use for the rigorous determination of the conformational properties of disordered proteins.

Introduction

Disordered states of proteins play crucial roles in the processes of protein folding and aggregation[1-5]. It is also increasingly clear that the presence of conformational disorder in native proteins is a widespread phenomenon of great significance, as intrinsically disordered proteins are closely involved in signalling and regulation, as well as in a range of medical conditions that include Alzheimer's and Parkinson's diseases[1-7]. Therefore, a better characterisation of protein disorder at the molecular level could offer many opportunities to advance our understanding of fundamental aspects of molecular biology as well as for the development of effective therapies for currently incurable diseases[1-7].

Disordered proteins rapidly interconvert between many structurally diverse conformations, and are best described as conformational ensembles. Nuclear magnetic resonance (NMR) spectroscopy has proved uniquely able to capture these conformational fluctuations at high resolution and even enable the determination of disordered protein ensembles[8-16]. Two important observables that have gained increasing prominence for this purpose are residual dipolar couplings (RDCs) between pairs of nuclei in partially aligned protein samples[13, 14, 17-19], and paramagnetic relaxation enhancements (PREs), which exploit the enhanced dipolar relaxation of nuclear spins in proximity of paramagnetic groups to obtain long-range distance information[9-11, 20-24]. These observables are particularly important as disordered proteins populate expanded conformations and yield few measurable short-range distances derived from nuclear Overhauser effects (NOEs), which are traditionally used to define the structures of ordered proteins by NMR spectroscopy.

As the use of RDC and PRE based methods is becoming widespread for the study of disordered proteins, it is important to investigate the perturbations that their measurements can introduce, which could potentially lead to biases in the analysis. As we will discuss, these perturbations are unlikely to be significant for folded proteins and hence their importance in the case of disordered proteins is just beginning to be recognised.

In order to measure RDCs, a weak alignment of the protein molecules is induced by the presence of a liquid crystalline alignment medium[25, 26], which could alter the protein conformations through long-range interactions or even by direct binding to the surfaces of the alignment media (nematogens)[27]. These perturbations are significantly more likely for

disordered than for native proteins as their conformations are not stabilised by large numbers of intramolecular contacts that are present in their ordered counterparts.

PRE measurements generally require tagging of the target protein with an extrinsic paramagnetic group, either through attachment of a chelating group that binds paramagnetic metal ions with high affinity or by direct conjugation of a stable nitroxide radical (commonly termed a 'spin label'). As such tags are typically at least 50% larger than naturally occurring amino acids, their presence may affect the structures and dynamics of proteins in the vicinity of the labelling site, and even lead to the formation of non-native long-range interactions.

For ordered proteins, the effects of molecular alignment and paramagnetic tagging can in principle be determined by reference to known high-resolution structures of the protein of interest. However, in the case of disordered proteins, structural information is generally not available, and indeed it is exactly for this purpose that RDCs or PREs are being measured in the first place. An alternative approach is to cross-validate the structures by showing that the different results obtained from subsets of RDC or PRE data are consistent with each other[12, 28-31]. Such validation can, however, be challenging because of the large number of degrees of freedom of disordered proteins. Given these limitations, we describe in this paper approaches to permit structure-free validation of RDC and PRE measurements, which can enable them to be applied more rigorously in the study of disordered proteins, other biomolecules that exhibit large structural fluctuations, or indeed folded proteins whose structures are unknown.

Materials and Methods

Preparation of Aβ40 NMR samples

Recombinant Aβ40 peptides with ammonium acetate counter ions were obtained from AlexoTech (Umea, Sweden), and handled on ice at all times. The peptide powder was solubilised in 10 mM NaOH at concentrations of 4-7 mg/ml and stored at -80 °C until required. In order to prepare NMR samples, Aβ40 stock solutions were diluted into sodium phosphate buffer (Na_2HPO_4/NaH_2PO_4, 50 mM) to give a final peptide concentration of 60-150 μM. The final buffer composition was 20 mM sodium phosphate, 0.5 mM EDTA, 0.02%

w/v NaN3 and contained 10% D_2O. HCl was pre-added to the sodium phosphate to neutralise NaOH in the Aβ40 stock solutions and, if necessary, the samples were titrated to pH 7.4 by addition of 100 mM NaOH. Final sample volumes were 160-200 μL and contained 10% D_2O. All samples were ultracentrifuged for 20 min at 8500×g to sediment any aggregates of peptide present before transferral to 3 mm MATCH^{TM} NMR tubes (Bruker BioSpin AG, Fällenden, Switzerland).

Aligned samples containing Pf1

Pf1 bacteriophage was obtained from Asla Biotech (Riga, Latvia). Buffer exchange was accomplished by precipitation in 1.25 M NaCl and ultracentrifugation for 1 h at 11300×g and 30 °C. The pellet was washed three times with de-ionised water and re-suspended in sodium phosphate buffer over 1-2 h at 4 °C. Aβ40 was then added (25-35 μL in 10 mM NaOH) to give a final buffer composition identical to that of the isotropic samples and final volume of 200 μL. From a comparison of ^{1}H-^{15}N HSQC spectra, Pf1 caused a slight increase in pH of the sample at high concentrations. In these cases, pH was adjusted by the addition of 100 mM HCl, which did not lead to aggregation of Aβ40 during the experiments. As solutions containing Pf1 had high viscosity, samples were often left overnight to ensure air bubbles were no longer present. Sample homogeneity in each case was confirmed by the appearance of a sharp doublet in the ^{2}H 1D spectrum. The quadrupolar splitting in Pf1 solutions was not observed to change particularly over the course of experiments.

Spin labelled derivatives

The (1-Oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl) methanethiosulfonate (MTSL) nitroxide radical, was obtained from Toronto Research Chemicals (Toronto, Canada) and dissolved by sonication in 100 mM sodium phosphate containing 6 M guanidium chloride (pH 8.0) to form a supersaturated solution. Undissolved spin labels were sedimented using a benchtop centrifuge and the supernatant used for subsequent reactions. Aβ40 cysteine mutants dissolved in 10 mM NaOH (as above) were first reduced *in situ* by an 8-fold excess of *tris*(2-carboxyethyl)phosphine (TCEP) for 15 min, ensuring the reaction pH remained above 7 by dissolving TCEP in NaOH. The saturated MTSL solution was then added to give a final sodium phosphate concentration of 60 mM, whereupon the reaction with MTSL was allowed to proceed for 15 min.

MALDI mass spectrometry confirmed the presence of the MTSL-labelled species as the major product in the reactions. However, in all mass spectra analysed, the peak corresponding to unlabelled peptide persisted at a fraction of approximately 10-20% of the total peptide present. An array of different pH values and reaction times were employed, but no significant differences could be detected. Following purification, we checked volumes of [^{1}H, ^{15}N]-HSQC peaks for residues near the labelling site with MTSL in its paramagnetic state. Because of the large PREs at these residues, peak volumes should be undetectably low[32], which was indeed found to be the case. These results indicate that yields were close to 100%, and suggest that peak areas in the mass spectra may be biased by a low mobility of the MTSL labelled Aβ40 species.

The reaction mixture was purified by size exclusion chromatography using an analytical Superdex 75 10/300 GL column (GE Healthcare, Amersham, UK) eluting with 4 mM sodium phosphate buffer containing 0.1 mM EDTA and 0.004% w/v NaN_3 (pH 7.5). Fractions containing Aβ40 were identified by UV absorbance at 280 nm, and peak areas were used to quantify the amount of peptide present using an extinction coefficient of 1595 cm^{-1}. The peptide fractions were lyophilised and redissolved in 9:1 H_2O:D_2O to given a final sodium phosphate concentration of 20 mM. Aβ40 concentrations in the NMR samples were 60-100 μM.

Validation experiments were performed with peptides for which MTSL had been reduced by addition of sodium ascorbate (400 μM). The reaction was allowed to proceed for 12 h at 5 °C and complete reduction was monitored by the increase in intensity of Aβ resonances in the ^{1}H NMR spectrum. Sodium ascorbate did not lead to any measurable changes in chemical shifts or peak intensities of Aβ40 wild-type at the concentration used for reduction of the spin labelled derivatives.

NMR spectroscopy

NMR spectra were recorded on Avance 500 MHz and 700 MHz spectrometers (Bruker BioSpin AG, Fällenden, Switzerland) equipped with four-channel *z*-gradient cryogenic probes. All experiments were recorded at 278 K. Data were processed using the NMRPipe[33] by application of shifted sine bell window functions and extensive zero filling prior to Fourier transformation. Spectra were analysed with NMRPipe and Sparky (T. D. Goddard and D. G.

Kneller, SPARKY 3, University of California, San Francisco). We found that data acquisition was substantially expedited by the use of a ^{15}N spectral width of 10.5 ppm. Well-resolved [^{1}H, ^{15}N]-HSQC spectra (**Figure S1**) could thus be recorded with only 32 complex points in the indirect dimension.

Residual dipolar couplings

^{1}H-^{15}N RDCs were derived from the difference in peak splittings between aligned and isotropic samples in 2D IPAP [^{1}H,^{15}N]-HSQC spectra[34]. Accurate peak positions were determined by fitting a Gaussian function to each peak within Sparky. $^{1}D_{NH}$ values were averaged over three distinct experiments in order to obtain residue-specific estimates of the standard errors of measurement.

The goodness-of-fit for each residue in the concentration dependence analysis (see Eq. 13) was determined by the reduced χ^2 parameter (χ^2_{red})

$$\chi^2_{red} = \frac{1}{\nu}\sum_{i=1}^{N}\frac{(D_{fit,i}-D_{exp,i})^2}{\sigma_i} \quad (1)$$

where $D_{fit,i}$ and $D_{exp,i}$ are the fitted and experimental RDC values, respectively, σ_i is the experimental error and the index i in the sum runs over the N concentrations used. The number of degrees of freedom (ν) is given by N-n-1 where n is the number of fitted parameters for each model. A value of χ^2_{red} close to unity indicates that the data are well fitted by the model given the experimental uncertainty.

The agreement between experimental RDCs and those corrected for the effects of alignment-induced perturbation was assessed by the RDC quality factor[35], which is defined as

$$Q = \sqrt{\frac{\sum_{k=1}^{N}(D_{corr,k}-D_{exp,k})^2}{\sum_{k=1}^{N}D_{corr,k}^2}} \quad (2)$$

where D_{exp} is the experimental RDC, D_{corr} is the corrected RDC corresponding to the unperturbed state, and the index k runs over all residues for which data were available. Values

of Q below 0.2 are generally thought to indicate good agreement between a predicted quantity (in this case D_{exp}) and its true value (in this case D_{corr}).

Spin relaxation

Backbone ^{15}N transverse relaxation rates (R_2) were measured with a sensitivity-enhanced CPMG pulse sequence that contained additional CPMG trains for heating compensation[36]. Backbone $^{1}H_N$ R_2 rates were measured using the HSQC-based pulse sequence proposed by Kay and co-workers[32, 37], employing a REBURP pulse (3 ms at 500 MHz, centred at 8.3 ppm) for selective refocussing of amide resonances during the first INEPT period. In both cases, 8-10 relaxation delays were recorded in interleaved fashion over the range 20-700 ms (^{15}N) or 13-223 ms ($^{1}H_N$). Decay curves were fitted with two-parameter single exponential functions in Sparky, and errors estimated by the standard deviation of decay constant values upon addition of Gaussian noise to the peak heights (Monte Carlo method).

Monitoring Aβ40 aggregation

^{1}H NMR experiments were used to assess the aggregation of Aβ40 peptides over time, as the signal observed in solution NMR comes entirely from the monomeric peptide[38]. To control for any systematic change in spectral intensity over time, which may be caused for example by changes in shims or water suppression, the time evolution of the integral over the Aβ40 methyl region (0.6-1.0 ppm) was compared to the integral of the EDTA resonance (3.65 ppm), which is not expected to vary over time.

Diffusion coefficients

Translational diffusion coefficients were measured using a stimulated echo pulse sequence[39] that included a 3-9-19 pulse train for effective water suppression[40]. Smooth square gradients were used throughout, with gradient pulse lengths of δ=4.5 ms and inter-gradient delays of Δ=100 ms. Spectra were acquired at 32 values of gradient strength (G) from 5% to 95% of the maximum value for the probe employed (nominally 53.5 G/cm). The decay of the integral over the aromatic region of the Aβ40 spectrum (6.9-7.4 ppm) was fitted using a modified form of the Stejskal-Tanner equation suitable for bipolar gradient experiments[39, 41]

$$S(G) = S_0 exp\left[-G^2\delta^2\gamma_H^2 D_t\left(\Delta - \frac{\delta}{3} - \frac{\tau}{2}\right)\right] \quad (3)$$

where S_0 is the signal when $G=0$, γ_H is the ^1H gyromagnetic ratio, D_t is the translational diffusion coefficient and τ is the delay between the two halves of the bipolar gradient (0.17 ms). Average values of D_t were determined based on three repeats of the experiment on the same sample.

Differences between the diffusion coefficients of the Aβ40 variants were assessed by statistical testing. In all cases we employed a significance level of $\alpha=0.05$. Variability was first assessed by a one-way analysis of variance (ANOVA) test. Upon finding significant variability ($p<\alpha$), we performed *post-hoc* pairwise comparisons of the paramagnetic variants with the wild-type peptide by unpaired two-tail t-tests. In spite of small sample sizes ($n=3$-4), the main the errors arise from spectral noise and therefore can be assumed to be normally distributed.

The translational diffusion coefficient is related to the hydrodynamic radius (R_H) by the Stokes-Einstein relation

$$R_H = \frac{k_B T}{6\pi\eta D_t} \quad (4)$$

where k_B is the Boltzmann constant, T is the temperature and η is the viscosity of the solution. As the temperature and viscosity of solutions containing different Aβ40 variants are identical, their sizes relative to Aβ40 wild-type can be calculated as $R_{H,mut} = D_{t,wt}/ D_{t,mut}$.

Chemical shift perturbations

[^{1}H, ^{15}N] weighted chemical shift differences (Δ_{NH}) between wild-type and spin labelled peptides were calculated as

$$\Delta_{NH} = \sqrt{\Delta\delta_{HN}^2 + \left(\Delta\delta_N \frac{\gamma_H}{\gamma_N}\right)^2} \quad (5)$$

where $\Delta\delta_{HN}$ and $\Delta\delta_N$ are the ^1H_N and ^{15}N chemical shift differences respectively and γ_H and γ_N are the ^1H_N and ^{15}N gyromagnetic ratios. Assignments for the spin labelled peptides were obtained from HNHA experiments (see Supplementary Information).

Paramagnetic relaxation enhancements

$^{1}H_N$ transverse PREs (Γ_2) were determined from [^{1}H,^{15}N]-HSQC spectra of Aβ40 V12C* in the paramagnetic and diamagnetic states according to the peak volume method proposed by Xue *et al.* [42]. Γ_2 was calculated for each residues according to

$$\Gamma_2 = \frac{1}{2\tau} ln\left(\frac{V_d}{V_p}\right) \quad (6)$$

where τ is the length of the INEPT periods ($1/2J_{NH}$). Volumes were determined by nlinLS Gaussian fitting to the peaks within NMRPipe. A REBURP pulse (1.714 ms at 700 MHz, centre 8.3 ppm) was used in the first INEPT period to refocus only $^{1}H_N$ spins and avoid intensity modulation by ^{3}JHNHα. To ensure equilibration of diamagnetic sample between scans and to avoid underestimation of Γ_2 as a result of longitudinal PRE effects, an inter-scan delay of 4.5 s was used. The effect on peak volumes of sample dilution by addition of sodium ascorbate was taken into account by the addition of a reference compound as suggested by Xue *et al.*[42], 100 μM [^{15}N] N-acetyl-glycine (Isotec, Old Bridge, NJ).

Calculation of PREs from protein structures

The ^{1}H transverse relaxation enhancement in the presence of a paramagnetic species (Γ_2) is given by the Solomon-Bloembergen equation[43, 44]

$$\Gamma_2 = \frac{1}{15}\left(\frac{\mu_0 \gamma_H g_e \mu_B}{4\pi}\right) s_e(s_e+1)[4J(0) + 3J(\omega_H)] \quad (7)$$

where $J(\omega)$ is the generalised spectral density function for the electron-proton interaction, g_e is the electron g-factor, γ_H is the proton gyromagnetic ratio, s_e is the electron spin quantum number, $\omega_H/2\pi$ is the proton Larmor frequency, μ_B is the Bohr magneton and μ_0 is the permeability of vacuum. These physical constants are commonly merged into a single constant, K, which takes the value $1.23 \cdot 10^{16}$ Å^6 s^{-2} when a nitroxide radical is the paramagnetic group. For proteins containing a paramagnetic group attached by freely rotatable bonds, such as MTSL, the most appropriate form of the spectral density is[45]

$$J(\omega) = \langle r^{-6} \rangle \left[\frac{S^2\tau_c}{1+(\omega\tau_c)^2} + \frac{(1-S)^2\tau_t}{1+(\omega\tau_t)^2}\right] \quad (8)$$

where r is the electron-proton distance, angle brackets denote averaging over the distribution of distances, and S^2 is the generalised order parameter for the motion of the tagged side-chain. τ_c is the correlation time for the electron-proton interaction, which can be written as $(\tau_r^{-1}+\tau_s^{-1})^{-1}$, where τ_r is the rotational correlation time of the molecule, and τ_s is the electron spin longitudinal relaxation time. τ_t is the total correlation time, which can be written as $(\tau_c^{-1}+\tau_i^{-1})^{-1}$, where τ_i is the correlation time for the internal motion of the spin label. As τ_s is typically 100 times larger than τ_r for nitroxide radicals[46], τ_c can be approximated by τ_r. Therefore, the spectral density depends on only two correlation times, τ_r (overall tumbling) and τ_i (internal motions). This treatment assumes that the internal motions of the paramagnetic tag are uncoupled with the rotational tumbling of the macromolecule, with the latter expected to take place on longer timescales.

In order to evaluate the spectral density (Eq. 8), it is convenient to decompose S^2 into angular and radial parts

$$S^2 \approx S^2_{ang} S^2_{rad} \tag{9}$$

where the individual parts are given by

$$S^2_{ang} = \frac{4}{5\pi} \sum_{m-2}^{2} |\langle Y^2_m(\vartheta, \varphi) \rangle|^2 \tag{10}$$

and

$$S^2_{rad} = \frac{\langle r^{-3} \rangle^2}{\langle r^{-6} \rangle} \tag{11}$$

where Y^2_m are second-order spherical harmonics and (ϑ,φ) are the polar and azimuthal angles between the principal axes of the moment of inertia tensor of the molecule and the electron-proton vector for a given conformation of the side-chain.

To apply the above equations to calculate the PREs from explicit ensembles of Aβ40, we first generated a library of MTSL rotamers through molecular dynamics simulations of the spin

label attached at position 2 of the peptide. Clustering of the resulting structures led to 76 rotamers for MTSL, which we took to occur with equal likelihood.

For each backbone conformation in the ensemble, the allowed MTSL rotamers were obtained by first aligning the backbone Cα, C' and N atoms of residue 12 with those of the MTSL bearing side-chain. All MTSL conformations that produced a steric clash with the Aβ40 structure were then removed. Clashes were defined to occur when two atoms came within the sum of their van der Waals radii. Xplor-NIH values of the van der Waals radii were used in calculations, with a scaling factor of 0.75 as conventional in the final stages of structure refinement in this package[47]. Only backbone atoms (and Cβ) were used on Aβ40 as it is anticipated that side-chains could rearrange easily to accommodate MTSL, and in addition the Cβ atom was excluded from the residue at the labelling site. All atoms on the MTSL side-chain were included apart from Cβ, Hβ and Sγ. Following the removal of disallowed rotamers, Γ_2 was calculated for the backbone conformation using correlation time values of τ_r=3 ns[48] and τ_i =500 ps[22]. The PREs were then averaged over all backbone conformations according to

$$\Gamma_2 = \frac{1}{N_s}\sum_{i=1}^{N_s}\Gamma_{2,i} \qquad (12)$$

where N_s is the number of backbone conformations in the ensemble of interest. The separation of averaging over rotamer positions for a specific backbone conformations, and then backbone conformations, assumes that motion of the side-chain is faster than and decoupled from backbone conformational changes.

Results

RDC measurements: Possible perturbations from molecular alignment

The method that we propose for assessing the effects of molecular alignment on RDC measurements of disordered proteins is illustrated in **Figure 1**. The degree of compatibility of a given protein with a particular alignment medium is first assessed by consideration of the

pH, buffer and temperature requirements[49]. Electrostatic alignment media with opposite charge to the protein should be avoided because of the strong attractive forces that will result. More subtle perturbations, however, may arise at a local level around individual residues of charge opposite to that of the alignment medium. Working at high ionic strengths would reduce these perturbations, but also increase the correlations between RDCs measured in different alignment media, thus reducing the amount of information that can be extracted from them. Suitable alignment media should, in addition, enable high-resolution spectra to be recorded without extensive line broadening caused by binding[27] or by unresolved ^{1}H-^{1}H dipolar interactions[50]. Strong interactions can be mitigated by using lower concentrations of the alignment media or by screening electrostatic interactions with higher concentrations of salt[51, 52].

Following the optimisation of the conditions, the effects of perturbations introduced by the alignment process on the RDCs themselves are elucidated by measuring their values at different nematogen concentrations. In the absence of perturbations, RDC values (D) are expected to scale linearly with concentration ($D=\alpha c$), where c is the concentration and α is a scaling constant, as the overall anisotropy of the medium increases in proportion to the quantity of nematogen present. Where perturbations occur, provided that the fraction of perturbed molecules is below 5%, measured RDC values will show an effective quadratic concentration dependence (see Supplementary Information)

$$D = \alpha c + \beta' c^2 \tag{13}$$

where β' captures the magnitude of the perturbations introduced by the alignment media. Residues showing significant deviations from linear behaviour are identified as those showing an increase over 10% in the goodness-of-fit using the quadratic model versus a linear one. If many residues show high degrees of curvature (i.e. a large value of β'), this is an indication that the medium is significantly perturbing the conformational ensemble of the protein and that the medium or the conditions should be altered to record useful RDCs for structural studies. On the other hand, if only a few residues show high curvature, these can be removed and the remaining residues used for subsequent conformational analysis. This approach also enables the corrections needed for the RDCs to be determined, as the RDC value corresponding to the unperturbed state alone is αc.

Validity of RDCs for the Aβ40 peptide

We illustrate our method for the validation of RDCs in a study of Aβ40, a peptide whose aggregation from a disordered monomeric state to form amyloid assemblies is implicated in Alzheimer's disease[53]. Aβ40 has a net charge of about -2.4 at pH 7.4 (**Figure S2**) and therefore the negatively charged Pf1[50] alignment medium was selected for measurement of RDCs. RDCs from electrostatic alignment can provide valuable complementary information to those commonly recorded in steric media for disordered proteins, and therefore their use has the potential to expand the amount of RDC data available from these states[52].

At an ionic strength of 50 mM, use of Pf1 concentrations up to about 26 mg/ml yielded NMR spectra of similar resolution to unaligned samples. No NMR signals were broadened beyond detection and many showed no additional line broadening in the presence of Pf1 (**Figure S3**). As these results indicated that conditions were appropriate for RDC measurements, we determined the concentration dependence of the RDC values. Because of the high viscosity of the Pf1 solutions and difficulties in accurate concentration measurement, we used the ^{2}H quadrupolar splitting (Q_{cc}) as an *in situ* measure of Pf1 concentration. The two are directly proportional when Pf1 is fully aligned[51], a situation which was confirmed in the present study by the identical values of Q_{cc} for the lowest Pf1 concentration at two magnetic field strengths (11.7 and 16.4 T).

Goodness-of-fit tests for the quadratic and linear concentration dependence models are shown in **Figure 2a**, where a lower value χ^2_{red} indicates improved fit to the data for the quadratic model; since the statistics are normalised for the number of model parameters, the reductions in χ^2_{red} are not simply a result of increased model complexity. The numbers of, respectively, residues G25 and I31 show improved agreement with the quadratic model and the three examples of residues exhibiting pronounced curvature are shown in **Figure 2b**. The dashed lines in the plots show the RDC values of the unperturbed state, indicating that alignment-induced perturbations lead to measurable systematic deviations in RDC values.

Largest curvatures are observed at residues K16 and N27, which suggest that perturbations may arise from transient binding of Aβ40 to the negatively charged Pf1 surface (**Figure 2c**) as N27 is adjacent to the other lysine residue in the Aβ40 sequence (K28). To determine

whether or not binding was significant in this case, backbone ^{15}N transverse relaxation rates (R_2) were measured in the presence and absence of Pf1. R_2 values are expected to increase for residues involved in binding to Pf1 because of the very fast transverse relaxation of spins tumbling with the viral particle that has an estimated correlation time of the order 1-100 μs[54].

Figure 2d shows the changes in R_2 upon addition of about 26 mg/ml Pf1 to the solution containing the Aβ40 peptide. The largest increases occur between residues K16 and F20, indicating that binding to Pf1 occurs principally through this region. In contrast, increases in R_2 are not observed in the proximity of K28, indicating that this region does not bind to Pf1 directly. The perturbation of the RDC values between residues G25 and N27 may therefore arise due to stabilisation of turn conformations upon binding to Pf1, but with the residues at the head of the turn not directly in contact with the surface of Pf1. Turn formation has previously been shown to cause a change of sign in the RDC profiles[55], which is present here in the measured RDCs at residue G25, but not in the corrected RDCs corresponding to the unperturbed conformations alone, indicating that turn formation is unlikely (**Figure 2e**).

The effects of the systematic errors introduced by Pf1 on the RDC profile of Aβ40 can be seen by a comparison of corrected and measured data at Q_{cc}=15.8 Hz (**Figure 2e**). Overall, the correlation between the two sets of RDCs is high (r=0.97) and the RDC Q-factor (see Methods) has a value of 0.22. Nevertheless, significant deviations can be observed in several regions and five residues show errors of >1 Hz. Removal of these residues reduces the Q-factor to 0.13.

We also repeated the concentration dependence analysis for RDCs at an ionic strength of 200 mM and found that systematic errors were reduced (Q=0.13 with all residues included), particularly at residues K16 and N27 (**Figures 2f and S4**). These findings provide further support to the hypothesis that binding between the Aβ40 peptide and Pf1 is mediated by electrostatic interactions, and show that RDC concentration dependence measurements are able to identify whether conformational changes are taking place with high sensitivity.

PRE measurements: Possible perturbations from paramagnetic tagging

We next investigated a method for validation of paramagnetic derivatives to be used in PRE measurements. The method, which is outlined in **Figure 3**, involves detailed comparison of the tagged variant with the properties of the unmodified protein, with particular emphasis on determining the effect of paramagnetic tagging on long-range structure within the conformational ensemble. Candidate sites for tagging can be selected on the basis of their positions in the sequence (e.g. solvent exposure and location outside highly structured regions) and to match the physico-chemical properties of the final paramagnetic group. Since many nitroxide-based spin labels are hydrophobic, they should be substituted for large, non-polar side-chains. On the other hand, metal chelating moieties often contain amide or carboxylic acid functional groups, therefore similar polar side-chains should be selected for tagging[23].

Following successful production and tagging, each variant is compared to the unmodified protein using three criteria, which are considered in sequence by starting from the least sample intensive and time-consuming ones. The first step is to establish that native functions are preserved upon derivatisation, as a loss of these is indicative of changes in important structural features. Being sensitive to long-range distances (10-35 Å), PREs are expected to be strongly affected by changes in the compactness of the polypeptide chain. Therefore, the second step in the method is to assess the effect of tagging on the global dimensions of the protein, either through NMR diffusion or dynamic light scattering (DLS) measurements, which enable comparison of hydrodynamic radii (R_H), or small-angle X-ray scattering (SAXS), for comparison of radii of gyration (R_g). In the final stage, heteronuclear NMR experiments are used to probe the effect of paramagnetic tagging at the residue-specific level. Very useful parameters for the comparison are ^{15}N R_2 rates and, where possible, ^{1}H-^{15}N RDCs, both of which are sensitive to long-range structure within disordered proteins[22, 56]. Changes in these parameters should decrease further from the labelling site, with no changes observed in remote regions.

Validation of paramagnetic Aβ40 variants for PRE measurements

The use of the method is again illustrated in the study of the Aβ40 peptide, where we selected three positions for mutation of the native amino acid to cysteine and subsequent paramagnetic tagging by the thiol-specific nitroxide radical MTSL[57] (**Figure 4a**). Mutations were made outside the Aβ40 fibril core regions (grey arrows, **Figure 4a**), avoiding charged and aromatic residues. As a functional test, we monitored the aggregation rates of the tagged peptides (Aβ*) by ^{1}H NMR. Since in solution NMR, the Aβ signals have been shown to arise entirely from the monomeric peptide[38], a loss of signal corresponds to the conversion of monomeric Aβ40 into higher molecular weight species. Over the period studied, wild-type Aβ40 does not aggregate, a trend which is followed by the V12C* and S26C* variants (**Figure 4b**). However, the A2C* signal drops by 13%, indicating some aggregation for this peptide. An increase in the aggregation, by itself, does not necessarily imply an alteration of the monomeric conformational ensemble[58], as it may result from a variety of other microscopy processes contributing to the overall aggregation process[6, 59]. However, aggregation over this period does preclude accurate measurement of PREs, as intensity changes between paramagnetic and diamagnetic samples are no longer solely due to proximity to the paramagnetic group. As a result, the use of the A2C* mutant was not considered for further analysis.

To assess the effects of MTSL labelling on the overall dimensions of the Aβ40 peptide, diffusion coefficients were measured by NMR and used to determine the relative hydrodynamic radii (R_H) for each peptide relative to that of the wild-type (**Figure 4c**). The results exhibited significant variability, as also indicated by a one-way ANOVA test, which gave a *p*-value of 0.011. *Post-hoc* unpaired t-tests (two-tail, α=0.05, $n \geq 3$) indicated identical values for wild-type and V12C* (p=0.57), whereas wild-type and S26C* were significantly different (p=0.024). As alteration of the compactness of the S26C* peptide may affect PRE measurements, it was excluded from further analysis.

A final assessment of the suitability of the V12C* variant was performed using heteronuclear NMR experiments. The local effects caused by the introduction of a reduced MTSL spin label at position 12 were monitored using ^{15}N R_2 rates, which show the largest differences between the wild-type and V12C* variants in the region S8-F20 (**Figure 4d**), with higher values for

V12C* indicative of a decrease in segmental diffusion rate as a result of the introduction of the bulky spin label. These results were also confirmed by ^{1}H-^{15}N RDCs measured in Pf1 medium (**Figure 4e**), which show changes over a similar region (E11-E22). The V12C* derivative therefore satisfies our criteria as a valid model of wild-type Aβ40 for use in PRE measurements.

We also measured [^{1}H, ^{15}N] weighted chemical shift differences (Δ_{NH}) between wild-type and spin labelled peptides, which are typically used to assess the effects of paramagnetic tagging (**Figure 4f**). Changes are much smaller than those measured by relaxation rates and RDCs, though there are some small changes in the region F19-A21 detected for the V12C* variant. Δ_{NH} values for the S26C* variant are local and of similar magnitude to the V12C* despite the expansion of the ensemble observed for this peptide, which underlines the importance of the R_H measurements. For both peptides, ^{1}Hα chemical shifts and $^{3}J_{HNH\alpha}$ measurements showed even smaller changes with respect to Aβ40 wild-type (**Figure S5**); these parameters are therefore also of limited use for assessing the effect of paramagnetic tagging on the conformations of disordered proteins.

Discussion and Conclusions

RDCs and PREs have recently emerged as powerful NMR observables for the high-resolution characterisation of disordered proteins[9-11, 13, 14, 17, 18, 20-23]. These two types of data can be back-calculated readily from ensembles of structures and are particularly useful as they both contain long-range distance information, in contrast to other parameters such as chemical shifts, NOEs and scalar couplings. The combination of RDCs and PREs as structural restraints is also particularly advantageous as it makes use of their different dependence on the conformations of the protein[14]. This aspect is illustrated in **Figure 5**, where a schematic distribution of the radius of gyration in the conformational ensemble of a typical disordered protein is shown, along with bars indicating the contribution that each part of the distribution makes to measured RDC and PRE data. PREs, since they are highly sensitive to short distances due to their $\langle r^{-6} \rangle$ dependence, can provide important information on the nature of compact states within the ensemble. On the other hand, RDCs show a much weaker

dependence on the radius of gyration, but alignment magnitude is largest for conformers with high aspect ratio[60], making them more sensitive than PREs to extended conformations.

To measure RDCs or PREs, a protein must either be exposed to a non-native environment, namely the alignment medium, or subjected to non-native covalent modification, such as attachment of a nitroxide spin label. In an effort to assess the effects of these changes, we have described experimental methods that can be applied in order to validate the measurement of RDCs and PREs. We anticipate that they will be particularly useful for the study of disordered proteins and other highly dynamic protein states, as they do not rely on prior structural knowledge. Furthermore, they can be applied independently for each medium or tagging site and do not rely on potentially ambiguous cross-validation of multiple data sets, which is a limitation of current structure-free RDC validation protocols[61].

The RDC validation method that we have discussed enables data to be corrected for the effects of the alignment medium, yielding the RDCs corresponding in principle to the unperturbed state. If RDCs from multiple bond vectors are required and several concentration dependence series are too time consuming, data from residues identified as invalid from the analysis of ^{1}H-^{15}N RDCs can be removed from all sets, yielding valid couplings from multiple bond vectors. This method is applicable to RDCs recorded in any alignment medium provided that: (i) the protein can be aligned over a sufficient range of concentrations such that RDC variations are detectable above measurement errors; (ii) line broadening is not too severe at high concentrations that it precludes accurate RDC measurements; (iii) the order parameter of the nematogen remains constant over this concentration range. The latter can be verified by linear variation of the ^{2}H quadrupolar splitting of the D_2O resonance over the concentration range studied. Based on a survey of currently available media, these criteria are satisfied by most systems used for weak alignment of proteins. Provided the alignment medium is amenable to dilution, the concentration-dependence series can be obtained by preparation of a single protein sample at high nematogen concentration, followed by successive dilution. However, for samples in anisotropic gels, it is necessary to prepare a separate protein sample at each gel concentration. Strained polyacrylamide gels, which are commonly used for recording RDCs of disordered proteins, may also provide additional complications as the alignment of the samples depends on both the concentration of acrylamide and the degree of compression applied to the gel. The former may be subject to uncertainties if the gel is only partially rehydrated after drying, and the latter has been found to vary substantially in our

hands as evidenced by about 30% variability in the ^{2}H quadrupolar splitting for samples containing identical acrylamide concentrations.

The PRE validation method represents a stringent quality control of possible sites for paramagnetic tagging, which should lead to increased confidence in the reliability of PRE data from disordered proteins. Application of the method to Aβ40 shows that substantial changes can result from spin labelling and highlights the importance of validation using measurements that report on changes in long-range structure. It should be noted, however, that as none of the techniques used have comparable sensitivity to PREs for detecting the presence of low population states[23], the method that we have presented cannot rule out small population shifts caused by the paramagnetic groups. These changes may nevertheless manifest themselves in the PREs, especially if the conformations favoured are more compact than in the unmodified protein. It is not clear at the present time how this limitation can be overcome experimentally and it is therefore inevitable that PRE measurements must be subjected to consistency checks with other experimental data or with PREs from different labelling sites.

In conclusion, we anticipate that the quantitative assessment of the systematic errors associated with RDC and PRE measurements using the strategies discussed in this work will expand the range of applicability of these NMR parameters for the structural characterisation at high resolution of disordered states of proteins that are increasingly recognised as being of considerable biological interest and significance.

Supporting Information

Methods, supporting results and discussion, and four supporting figures. The supporting materials may be accessed free of charge online at "http://pubs.acs.org".

Funding information

This work was partly supported by the JSPS/MEXT Grants in Aid for Scientific Research on Innovation Areas (25102008 to M.Y.U.), Grant-in-Aid for Scientific Research (A) (15H02491 to M.Y.U.), and Young Scientists (B) (15K21680 to M.Y.U.). M.Y.U. is a recipient of the Naito Foundation Grant for Studying Overseas.

References

1. Wright, P. E., and Dyson, H. J. (2015) Intrinsically disordered proteins in cellular signalling and regulation, *Nat. Rev. Mol. Cell Biol. 16*, 18-29.
2. van der Lee, R., Buljan, M., Lang, B., Weatheritt, R. J., Daughdrill, G. W., Dunker, A. K., Fuxreiter, M., Gough, J., Gsponer, J., and Jones, D. T. (2014) Classification of intrinsically disordered regions and proteins, *Chem. Rev. 114*, 6589-6631.
3. Uversky, V. N. (2013) A decade and a half of protein intrinsic disorder: biology still waits for physics, *Protein Sci. 22*, 693-724.
4. Tompa, P. (2012) Intrinsically disordered proteins: a 10-year recap, *Trends Bioch. Sci. 37*, 509-516.
5. Habchi, J., Tompa, P., Longhi, S., and Uversky, V. N. (2014) Introducing protein intrinsic disorder, *Chem. Rev. 114*, 6561-6588.
6. Knowles, T. P., Vendruscolo, M., and Dobson, C. M. (2014) The amyloid state and its association with protein misfolding diseases, *Nat. Rev. Mol. Cell Biol. 15*, 384-396.
7. Heller, G. T., Sormanni, P., and Vendruscolo, M. (2015) Targeting disordered proteins with small molecules using entropy, *Trends Bioch. Sci. 40*, 491-496.
8. Varadi, M., Kosol, S., Lebrun, P., Valentini, E., Blackledge, M., Dunker, A. K., Felli, I. C., Forman-Kay, J. D., Kriwacki, R. W., and Pierattelli, R. (2014) pE-DB: a database of structural ensembles of intrinsically disordered and of unfolded proteins, *Nucl. Acids Res. 42*, D326-D335.
9. Lindorff-Larsen, K., Kristjansdottir, S., Teilum, K., Fieber, W., Dobson, C. M., Poulsen, F. M., and Vendruscolo, M. (2004) Determination of an ensemble of structures representing the denatured state of the bovine acyl-coenzyme a binding protein, *J. Am. Chem. Soc. 126*, 3291-3299.
10. Dedmon, M. M., Lindorff-Larsen, K., Christodoulou, J., Vendruscolo, M., and Dobson, C. M. (2005) Mapping long-range interactions in α-synuclein using spin-label NMR and ensemble molecular dynamics simulations, *J. Am. Chem. Soc. 127*, 476-477.
11. Bertoncini, C. W., Jung, Y.-S., Fernandez, C. O., Hoyer, W., Griesinger, C., Jovin, T. M., and Zweckstetter, M. (2005) Release of long-range tertiary interactions potentiates aggregation of natively unstructured α-synuclein, *Proc. Natl. Acad. Sci. USA 102*, 1430-1435.

12. Mittag, T., and Forman-Kay, J. D. (2007) Atomic-level characterization of disordered protein ensembles, *Curr. Op. Struct. Biol. 17*, 3-14.
13. Nodet, G., Salmon, L. c., Ozenne, V., Meier, S., Jensen, M. R., and Blackledge, M. (2009) Quantitative description of backbone conformational sampling of unfolded proteins at amino acid resolution from NMR residual dipolar couplings, *J. Am. Chem. Soc. 131*, 17908-17918.
14. Huang, J.-r., and Grzesiek, S. (2009) Ensemble calculations of unstructured proteins constrained by RDC and PRE data: a case study of urea-denatured ubiquitin, *J. Am. Chem. Soc. 132*, 694-705.
15. 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.
16. Camilloni, C., and Vendruscolo, M. (2014) Statistical mechanics of the denatured state of a protein using replica-averaged metadynamics, *J. Am. Chem. Soc. 136*, 8982-8991.
17. Jensen, M. R., Markwick, P. R., Meier, S., Griesinger, C., Zweckstetter, M., Grzesiek, S., Bernado, P., and Blackledge, M. (2009) Quantitative determination of the conformational properties of partially folded and intrinsically disordered proteins using NMR dipolar couplings, *Structure 17*, 1169-1185.
18. Esteban-Martín, S., Fenwick, R. B., and Salvatella, X. (2010) Refinement of ensembles describing unstructured proteins using NMR residual dipolar couplings, *J. Am. Chem. Soc. 132*, 4626-4632.
19. Camilloni, C., and Vendruscolo, M. (2014) A Tensor-Free Method for the Structural and Dynamical Refinement of Proteins using Residual Dipolar Couplings, *J. Phys. Chem. B 119*, 653-661.
20. Allison, J. R., Varnai, P., Dobson, C. M., and Vendruscolo, M. (2009) Determination of the free energy landscape of α-synuclein using spin label nuclear magnetic resonance measurements, *J. Am. Chem. Soc. 131*, 18314-18326.
21. 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.

22. Salmon, L. c., Nodet, G., Ozenne, V., Yin, G., Jensen, M. R., Zweckstetter, M., and Blackledge, M. (2010) NMR characterization of long-range order in intrinsically disordered proteins, *J. Am. Chem. Soc. 132*, 8407-8418.
23. Clore, G. M., and Iwahara, J. (2009) Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes, *Chem. Rev. 109*, 4108-4139.
24. Fawzi, N. L., Ying, J., Torchia, D. A., and Clore, G. M. (2010) Kinetics of amyloid β monomer-to-oligomer exchange by NMR relaxation, *J. Am. Chem. Soc. 132*, 9948-9951.
25. 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.
26. Tolman, J., Flanagan, J., Kennedy, M. A., and Prestegard, J. (1995) Nuclear magnetic dipole interactions in field-oriented proteins: information for structure determination in solution, *Proc. Natl. Acad. Sci. USA 92*, 9279-9283.
27. Ojennus, D. D., Mitton-Fry, R. M., and Wuttke, D. S. (1999) Induced alignment and measurement of dipolar couplings of an SH2 domain through direct binding with filamentous phage, *J. Biomol. NMR 14*, 175-179.
28. Goto, N. K., Skrynnikov, N. R., Dahlquist, F. W., and Kay, L. E. (2001) What is the average conformation of bacteriophage T4 lysozyme in solution? A domain orientation study using dipolar couplings measured by solution NMR, *J. Mol. Biol. 308*, 745-764.
29. Lukin, J. A., Kontaxis, G., Simplaceanu, V., Yuan, Y., Bax, A., and Ho, C. (2003) Quaternary structure of hemoglobin in solution, *Proc. Natl. Acad. Sci. USA 100*, 517-520.
30. Meier, S., Blackledge, M., and Grzesiek, S. (2008) Conformational distributions of unfolded polypeptides from novel NMR techniques, *J. Chem. Phys. 128*, 052204.
31. Vendruscolo, M. (2007) Determination of conformationally heterogeneous states of proteins, *Curr. Op. Struct. Biol. 17*, 15-20.
32. Iwahara, J., Tang, C., and Clore, G. M. (2007) Practical aspects of 1 H transverse paramagnetic relaxation enhancement measurements on macromolecules, *J. Mag. Res. 184*, 185-195.

33. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR 6*, 277-293.
34. Ottiger, M., Delaglio, F., and Bax, A. (1998) Measurement ofJand Dipolar Couplings from Simplified Two-Dimensional NMR Spectra, *J. Mag. Res. 131*, 373-378.
35. Cornilescu, G., Marquardt, J. L., Ottiger, M., and Bax, A. (1998) Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase, *J. Am. Chem. Soc. 120*, 6836-6837.
36. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and a phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation, *Biochemistry 33*, 5984-6003.
37. Donaldson, L. W., Skrynnikov, N. R., Choy, W.-Y., Muhandiram, D. R., Sarkar, B., Forman-Kay, J. D., and Kay, L. E. (2001) Structural characterization of proteins with an attached ATCUN motif by paramagnetic relaxation enhancement NMR spectroscopy, *J. Am. Chem. Soc. 123*, 9843-9847.
38. Hou, L., Shao, H., Zhang, Y., Li, H., Menon, N. K., Neuhaus, E. B., Brewer, J. M., Byeon, I.-J. L., Ray, D. G., and Vitek, M. P. (2004) Solution NMR studies of the Aβ (1-40) and Aβ (1-42) peptides establish that the Met35 oxidation state affects the mechanism of amyloid formation, *J. Am. Chem. Soc. 126*, 1992-2005.
39. Wu, D., Chen, A., and Johnson, C. (1995) Flow imaging by means of 1D pulsed-field-gradient NMR with application to electroosmotic flow, *J. Mag. Res. A 115*, 123-126.
40. Sklenar, V., Piotto, M., Leppik, R., and Saudek, V. (1993) Gradient-tailored water suppression for 1 H-15 N HSQC experiments optimized to retain full sensitivity, *J. Mag. Res. A 102*, 241-245.
41. Stejskal, E., and Tanner, J. (1965) Spin diffusion measurements: spin echoes in the presence of a time‐dependent field gradient, *J. Chem. Phys. 42*, 288-292.
42. Xue, Y., Podkorytov, I. S., Rao, D. K., Benjamin, N., Sun, H., and Skrynnikov, N. R. (2009) Paramagnetic relaxation enhancements in unfolded proteins: Theory and application to drkN SH3 domain, *Protein Sci. 18*, 1401-1424.
43. Solomon, I. (1955) Relaxation processes in a system of two spins, *Phys. Rev. 99*, 559.
44. Bloembergen, N., and Morgan, L. (1961) Proton relaxation times in paramagnetic solutions. Effects of electron spin relaxation, *J. Chem. Phys. 34*, 842-850.

45. Iwahara, J., Schwieters, C. D., and Clore, G. M. (2004) Ensemble approach for NMR structure refinement against 1H paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule, *J. Am. Chem. Soc. 126*, 5879-5896.

46. 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.

47. Schwieters, C. D., Kuszewski, J. J., and Clore, G. M. (2006) Using Xplor–NIH for NMR molecular structure determination, *Prog. Nucl. Magn. Reson. Spectrosc. 48*, 47-62.

48. Danielsson, J., Andersson, A., Jarvet, J., and Gräslund, A. (2006) 15N relaxation study of the amyloid β‐peptide: structural propensities and persistence length, *Mag. Res. Chem. 44*, S114-S121.

49. Chen, K., and Tjandra, N. (2012) The use of residual dipolar coupling in studying proteins by NMR, *Top. Curr. Chem. 326*, 47-67.

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

51. Zweckstetter, M., and Bax, A. (2001) Characterization of molecular alignment in aqueous suspensions of Pf1 bacteriophage, *J. Biomol. NMR 20*, 365-377.

52. 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. 45*, 7012-7015.

53. Haass, C., and Selkoe, D. J. (2007) Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid β-peptide, *Nat. Rev. Mol. Cell Biol. 8*, 101-112.

54. Tsang, P., and Opella, S. (1986) Pf1 virus particle dynamics, *Biopolymers 25*, 1859-1864.

55. Mukrasch, M. D., Markwick, P., Biernat, J., von Bergen, M., Bernadó, P., Griesinger, C., Mandelkow, E., Zweckstetter, M., and 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.

56. 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.

57. Berliner, L. J., Grunwald, J., Hankovszky, H. O., and Hideg, K. (1982) A novel reversible thiol-specific spin label: papain active site labeling and inhibition, *Anal. Biochem. 119*, 450-455.

58. Camilloni, C., and Vendruscolo, M. (2013) A relationship between the aggregation rates of α-synuclein variants and the β-sheet populations in their monomeric forms, *J. Phys. Chem. B 117*, 10737-10741.

59. Knowles, T. P., Waudby, C. A., Devlin, G. L., Cohen, S. I., Aguzzi, A., Vendruscolo, M., Terentjev, E. M., Welland, M. E., and Dobson, C. M. (2009) An analytical solution to the kinetics of breakable filament assembly, *Science 326*, 1533-1537.

60. 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.

61. Hus, J.-C., and Brüschweiler, R. (2002) Principal component method for assessing structural heterogeneity across multiple alignment media, *J. Biomol. NMR 24*, 123-132.

62. Hansen, M. R., Hanson, P., and Pardi, A. (2000) Filamentous bacteriophage for aligning RNA, DNA, and proteins for measurement of nuclear magnetic resonance dipolar coupling interactions, *Methods Enzymol. 317*, 220.

63. Paravastu, A. K., Leapman, R. D., Yau, W.-M., and Tycko, R. (2008) Molecular structural basis for polymorphism in Alzheimer's β-amyloid fibrils, *Proc. Natl. Acad. Sci. USA 105*, 18349-18354.

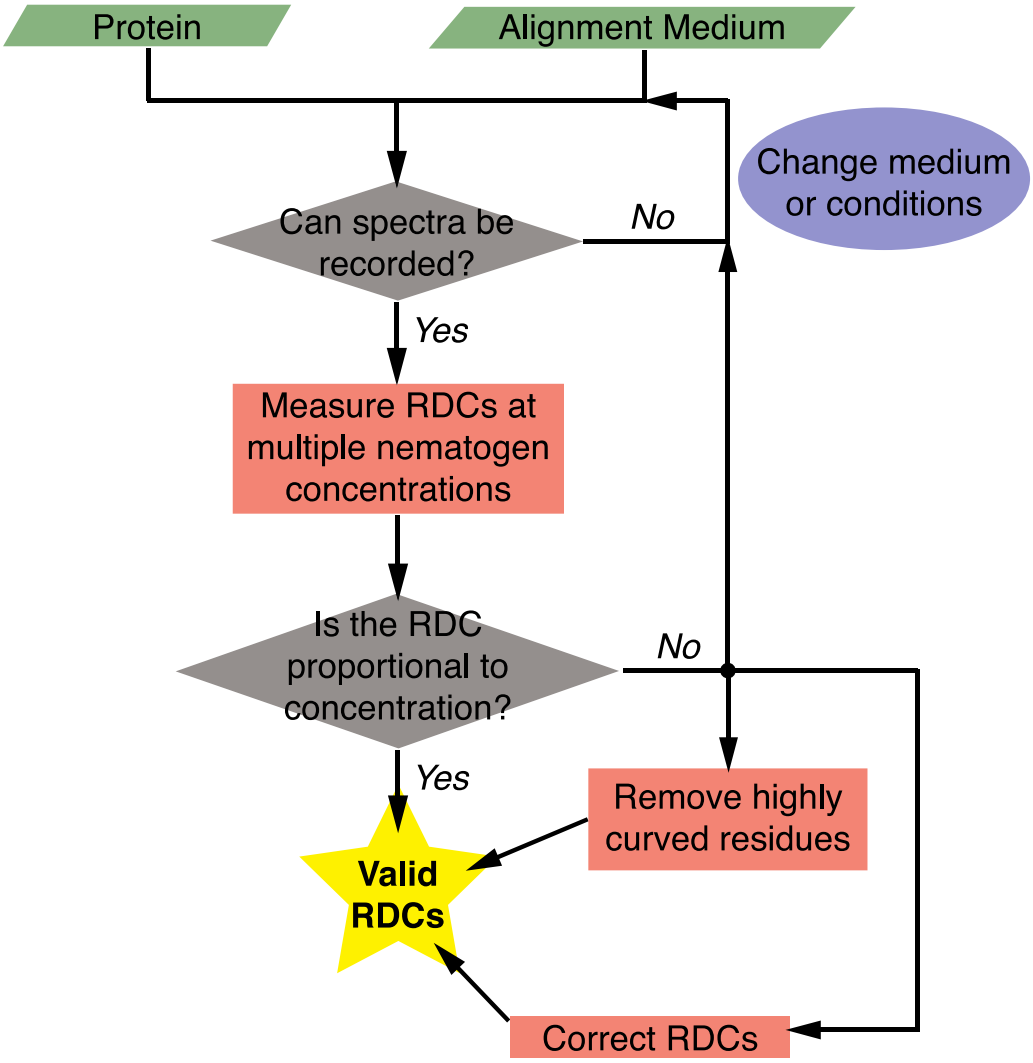

Figure 1. Schematic illustration of the method discussed in this work for the measurement of RDCs for disordered proteins. Possible interactions between the protein and the alignment medium (nematogen) are first assessed by measuring NMR spectra. Widespread peak disappearance and sample degradation are indicative of strong interactions. In this case the solution conditions should be altered or a different medium employed. If NMR spectra of sufficient quality can be recorded, ^{1}H-^{15}N RDCs should be measured over as wide a range of concentrations of the nematogen as practically possible. In the absence of perturbations caused by the alignment medium the RDC values should be proportional to the concentration. Therefore RDCs are invalid for residues (denoted as 'curved' in the flow diagram) for which significant deviations from linearity are observed. Such residues can be discarded from further analysis, or their RDCs corrected according to the procedure described in the text. If RDCs from many residues are subject to perturbations, it may be necessary to change alignment medium or conditions.

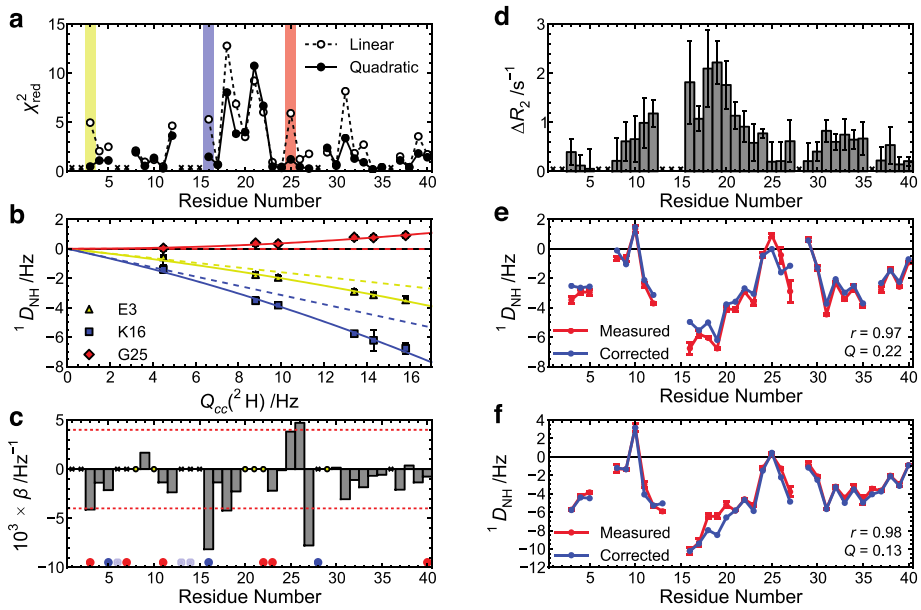

Figure 2. Concentration dependence of RDCs measured for the Aβ40 peptide. This analysis reveals a slight perturbation of the Aβ40 conformational ensemble induced by Pf1 under the conditions used in this study. **(a)** Reduced χ^2 values for fits of the RDC data to linear and quadratic models (see text), showing improvements upon addition of the quadratic term that are not merely the result of increased model complexity; coloured bars refer to the three residues in panel (b). **(b)** ^{1}H-^{15}N RDC values for three residues as a function of the ^{2}H quadrupolar splitting, $Q_{cc}(^2H)$, which is proportional to the concentration of Pf1[51, 62]. Solid lines are quadratic fits to the data; dashed lines are the linear term from the fit, corresponding to the RDC values of the unperturbed state. **(c)** Errors in the RDC values given by the coefficient of the quadratic term in Eq. (13). Since we used quadrupolar splittings instead of concentrations, the curvature is indicated by β, rather than by β' as in Eq. (13). Negatively and positively charged amino acids are denoted by red and blue circles, respectively, and histidine residues are shown by pale blue circles that indicate their degree of protonation at the given pH. Red dashed lines indicate the value of β above which residues are excluded from analysis because they would be affected by an error of over 1 Hz at a quadrupolar splitting of 15.8 Hz. Six residues did not show significant improvement in χ^2_{red} upon addition of a quadratic term (yellow circles). **(d)** Changes in ^{15}N transverse relaxation rates ΔR_2 upon addition of Pf1. Residues for which no data are available are marked with crosses (×). **(e)** Comparison of measured RDCs at Q_{cc}=15.8 Hz and corrected RDCs obtained from the linear terms of the quadratic fits shown in (b). **(f)** Comparison of measured and corrected RDCs from a second set of concentration dependence experiments performed in in the presence of an additional 150 mM NaCl at Q_{cc} =18.4 Hz.

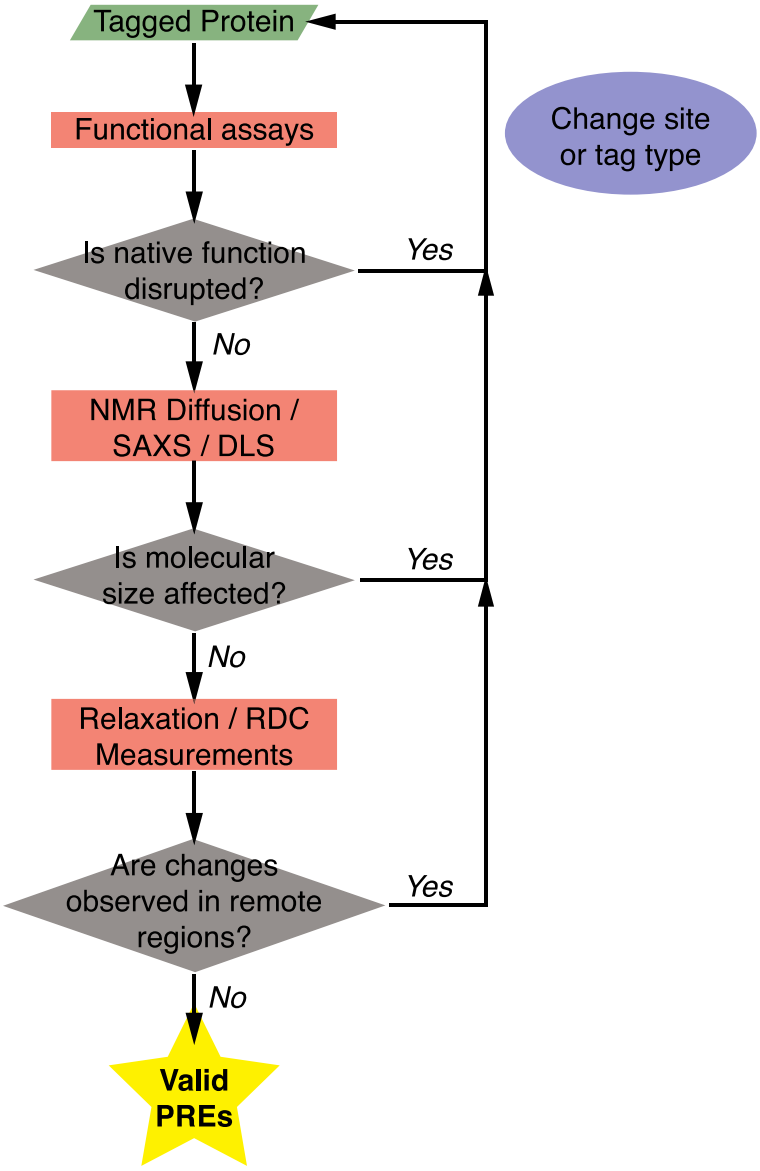

Figure 3. Schematic illustration of the method discussed in this work for the measurement of PREs for disordered proteins. Following the production and paramagnetic tagging of the protein, the first stage is to perform functional assays to ensure the presence of the tag does not abrogate native function. If this is not the case, it is then particularly important to ascertain the effect of spin labelling on the molecular dimensions via the measurement of the radius of gyration (R_g) or the hydrodynamic radius (R_H). Variants that successfully pass through this stage should be produced with appropriate isotopic enrichment and tested for deviations in ^{15}N transverse relaxation rates (R_2) or, better still, ^{1}H-^{15}N RDCs. These measurements give an indication as to whether the spin label affects long range contacts within the protein, which would preclude the use of that variant in PRE measurements.

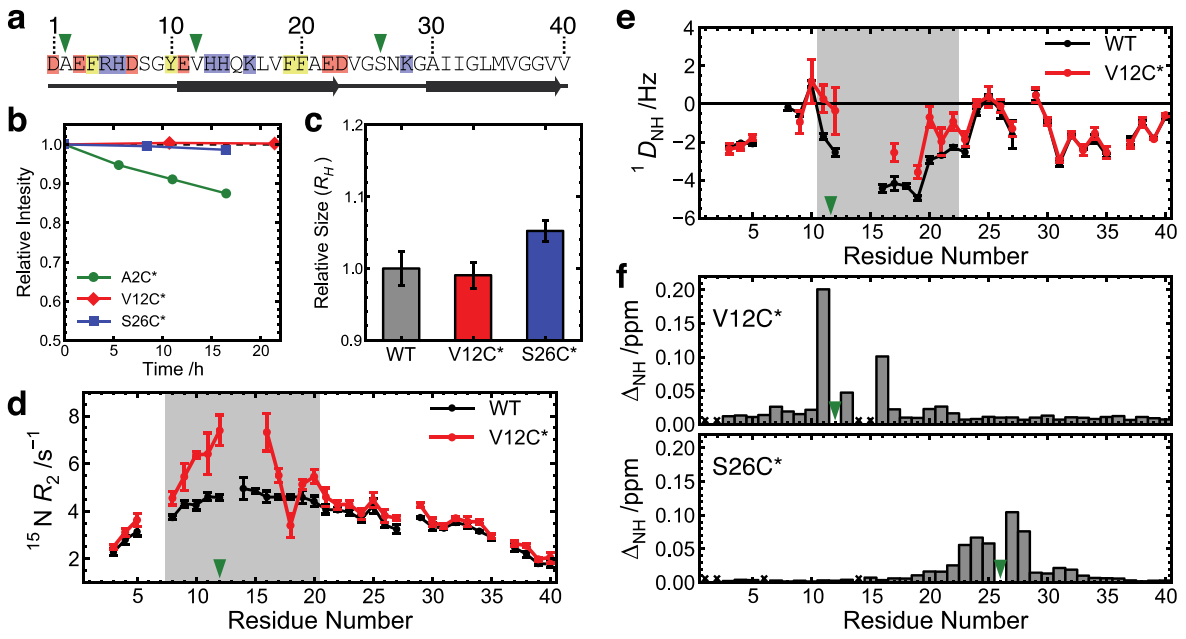

Figure 4. Validation of the paramagnetic derivatives of the Aβ40 peptide for PRE experiments. (a) Sequence of wild-type Aβ40 showing sites chosen for cysteine mutation and spin labelling (green triangles). Charged (red and blue) and aromatic (yellow) residues were excluded, as were those in core region of Aβ40 amyloid fibrils[63] (grey arrows). **(b)** Stability of spin labelled Aβ variants monitored by the integral over the methyl region of the ^{1}H NMR spectrum. The A2C* variant aggregates during the 16 h period required to perform PRE experiments. **(c)** Relative hydrodynamic radii of the different Aβ variants determined from NMR diffusion measurements, showing significant increase in the case of the S26C* variant. **(d)** ^{15}N transverse relaxation rates compared between wild-type Aβ40 and the V12C* variant. Major differences are observed around the site of labelling (green triangle). **(e)** ^{1}H-^{15}N RDCs for wild-type Aβ40 and the V12C* variant also exclusively show differences in proximity of the spin labelling site (grey shaded region). **(f)** Both relaxation rates and RDCs are much more sensitive to spin labelling than [^{1}H, ^{15}N] weighted chemical shift perturbations (Δ_{NH}), which are typically used to assess the effects of spin labelling. Δ_{NH} values (below 0.03) are also comparable to noise and highly localised for the S26C* variant, highlighting the importance of measurement of the global dimensions of the protein for validation.

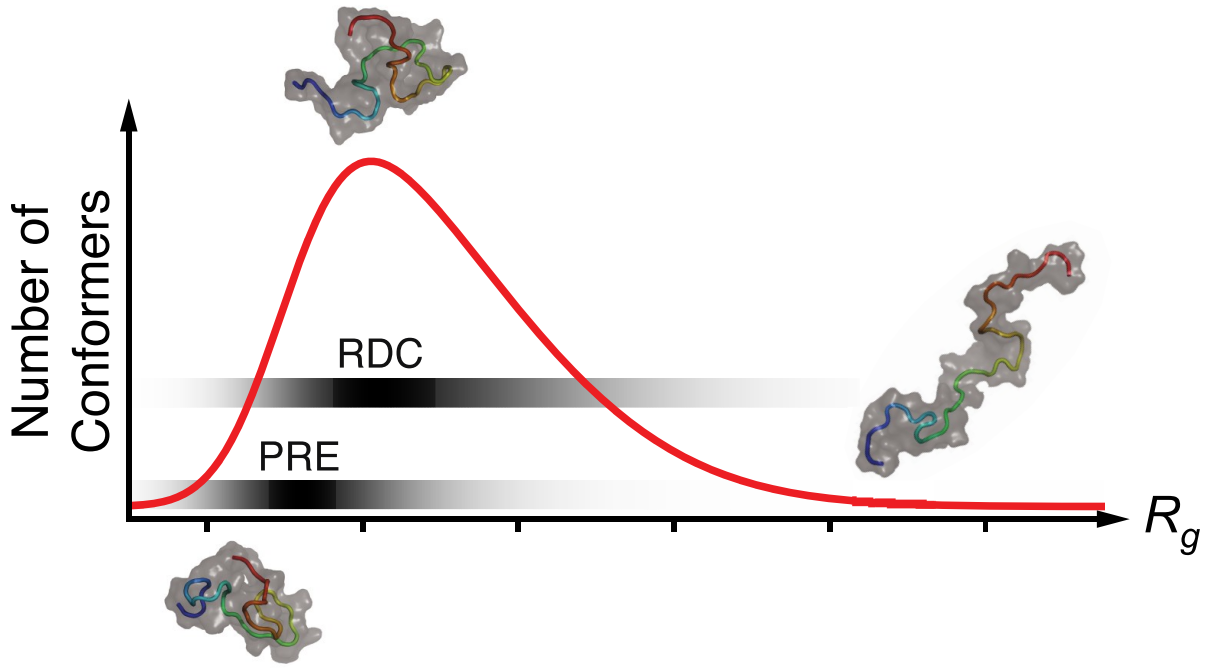

Figure 5. Schematic illustration of a typical distribution of the radius of gyration R_g for a disordered protein. Representative conformers of a typical disordered protein are shown as ribbon diagrams. Grey shaded bars indicate the weights of the contributions that conformers with a particular R_g value make to the observed NMR signal. These weights depend on the magnitude of the RDCs or PREs expected for a conformer with that value of R_g and the number of conformers with that size in the ensemble. PREs are particularly sensitive to compact species[14], whereas RDCs have a weaker size dependence, but are largest for conformations with a high aspect ratio[60]. As a result, the two techniques provide a complementary coverage of the conformational space of a protein.

ToC Graphics