PCCP

PAPER

View Article Online
View Journal | View Issue

Published on 21 September 2015. Downloaded by NATIONAL INSTITUTES OF HEALTH on 26/10/2016 17:12:05.

Cite this: *Phys. Chem. Chem. Phys.*, 2016, **18**, 5759

Quantitative evaluation of positive ϕ angle propensity in flexible regions of proteins from three-bond *J* couplings†

Jung Ho Lee, Jinfa Ying and Ad Bax*

Received 31st July 2015,
Accepted 21st September 2015

DOI: 10.1039/c5cp04542h

www.rsc.org/pccp

$^{3}J_{HNH\alpha}$ and $^{3}J_{C'C'}$ couplings can be readily measured in isotopically enriched proteins and were shown to contain precise information on the backbone torsion angles, ϕ, sampled in disordered regions of proteins. However, quantitative interpretation of these couplings required the population of conformers with positive ϕ angles to be very small. Here, we demonstrate that this restriction can be removed by measurement of $^{3}J_{C'H\alpha}$ values. Even though the functional forms of the $^{3}J_{C'H\alpha}$ and $^{3}J_{HNH\alpha}$ Karplus equations are the same, large differences in their coefficients enable accurate determination of the fraction of time that positive ϕ angles are sampled. A four-dimensional triple resonance HACANH[C′] E.COSY experiment is introduced to simultaneously measure $^{3}J_{C'H\alpha}$ and $^{3}J_{HNC'}$ in the typically very congested spectra of disordered proteins. High resolution in these spectra is obtained by non-uniform sampling (in the 0.1–0.5% range). Application to the intrinsically disordered protein α-synuclein shows that while most residues have close-to-zero positive ϕ angle populations, up to 16% positive ϕ population is observed for Asn residues. Positive ϕ angle populations determined with the new approach agree closely with consensus values from protein coil libraries and prior analysis of a large set of other NMR parameters. The combination of $^{3}J_{HNC'}$ and $^{3}J_{C'C'}$ provides information about the amplitude of ϕ angle dynamics.

1 Introduction

A survey of experimentally determined protein structures contained in the protein structure databank (PDB) shows that, with the exception of Gly, the vast majority of residues adopt negative values for the backbone torsion angle, ϕ. Well-ordered residues with positive ϕ angles are mostly restricted to type I′, type II, and type II′ tight β-turns and helix-terminating C-cap motifs. Transient occupation of conformations where one or more residues adopt positive ϕ angles may be of high functional relevance, and has been implicated in substrate recognition.[1–3] Moreover, molecular dynamics simulations on folded proteins indicate that peptide plane flips associated with a change in the sign of ϕ are energetically feasible,[4,5] suggesting that transient formation of conformations with a positive ϕ backbone torsion angle could be more common than implied by the analysis of static PDB X-ray structures. Adoption of positive ϕ values has been proposed to be particularly relevant in intrinsically disordered proteins (IDPs) and intrinsically disordered regions (IDRs) of amyloidogenic proteins, where they may impact the kinetics of fibril formation.[6,7]

Analysis of chemical shifts and other NMR parameters of IDPs and IDRs suggests low but non-vanishing populations of positive ϕ angle conformers,[8–12] mostly following what is seen in statistical coil libraries extracted from the PDB.[13,14] Traditionally, the primary markers for positive ϕ conformations are $^{1}J_{C\alpha H\alpha}$, which is at a minimum in the α_L region of Ramachandran space,[15] and strong intraresidue H^N-H^α NOEs.[16] However, neither of these parameters provides reliable fractional populations of the positive ϕ conformations as the variations in their values associated with negative ϕ angles usually are large and not very well defined. The Karplus equation relating $^{3}J_{C'H\alpha}$ to ϕ indicates that positive ϕ angles, associated with a *trans* C'–N–C^α–H^α arrangement, have much larger $^{3}J_{C'H\alpha}$ values than those observed for negative ϕ angles.[17] Here, we demonstrate that $^{3}J_{C'H\alpha}$, in particular when used in conjunction with $^{3}J_{HNH\alpha}$ and $^{3}J_{C'C'}$, provides a precise quantitative measure for the fractional positive ϕ population of any given residue. We demonstrate that, even for the typically very poorly resolved spectral regions of disordered proteins, nearly complete sets of $^{3}J_{C'H\alpha}$ values can be obtained by recording the HACANH[C′] E.COSY spectrum in a four-dimensional (4D) fashion. This experiment then also yields $^{3}J_{HNC'}$, providing yet

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

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c5cp04542h

Published on 21 September 2015. Downloaded by NATIONAL INSTITUTES OF HEALTH on 26/10/2016 17:12:05.

another parameter to define the ϕ angle distribution. Relatively recently introduced non-uniform sampling (NUS) strategies permit the recording of 4D spectra at very high resolution and in relatively limited amounts of measurement time.[18–20] This approach is used in our study, and the uncertainty at which J couplings can be extracted from such NUS-reconstructed spectra is shown to scale approximately inversely with the square root of the sparsity of the acquired data. Thus, the accuracy depends inversely on the square root of measurement time, following the same rules as in fully sampled NMR spectra.

2 Experimental section

2.1 NMR sample preparation

Uniformly $^{13}C/^{15}N$-enriched, N-terminally acetylated[21] wild-type α-synuclein (aS) protein was expressed and purified as described previously.[22] The NMR sample contained 0.6 mM aS in 20 mM sodium phosphate, pH 6.0, 50 mM NaCl, and 5% D_2O. The NMR sample of the third IgG-binding domain of protein G (GB3, 1.2 mM, uniformly $^{13}C/^{15}N$-enriched) was prepared in 50 mM sodium phosphate, pH 6.5, 50 mM NaCl, and 5% D_2O. The NMR sample of uniformly $^{13}C/^{15}N$-enriched ubiquitin (2.8 mM) was prepared in 20 mM imidazole buffer, pH 6.0, containing 7% D_2O.

2.2 NMR measurements

The 4D HACANH[C′] E.COSY spectrum of α-synuclein was recorded using the pulse sequence of Fig. 1, using a Bruker Avance III 600 MHz spectrometer equipped with a TCI cryogenic probe and a z-axis gradient coil. Non-uniform sampling with a sparsity of 0.55% was employed in order to achieve high resolution and thereby the most reliable peak positions and J coupling measurements, while keeping the total measuring time restricted to about three days. The time domain matrix consisted of 80* × 62* × 250* × 1024* complex data points, or acquisition times of 112 ms (t_1, $^1H^\alpha$), 28 ms (t_2, $^{13}C^\alpha$), 179 ms (t_3, ^{15}N), and 102 ms (t_4, $^1H^N$). The $^1H^\alpha$, $^{13}C^\alpha$ and ^{15}N carrier frequencies were set at 4.2, 56, and 118 ppm, respectively and 4 scans per FID were recorded. Additionally, 3D HA(CA)NH[C′] E.COSY spectra were recorded for ubiquitin and GB3, again using the pulse scheme of Fig. 1, but keeping the $^{13}C^\alpha$ evolution time $t_2 = 0$. The 3D HN(COCO)NH pulse sequence[23] was used at 900 MHz using a 2.5% NUS scheme, in order to obtain additional $^3J_{C'C'}$ coupling values for α-synuclein residues that were previously inaccessible due to resonance overlap. The new 3D HN(COCO)NH spectrum was recorded with 200 ms acquisition times in both the t_1 (^{15}N) and t_2 (^{15}N) dimensions.

For NUS processing of the HACANH[C′] E.COSY and HN(COCO)NH spectra, we used the in-house written SMILE routine (Ying *et al.*, unpublished), incorporated in the data processing software package NMRPipe.[24] The digital resolution of the final 4D spectra was 1.4 Hz ($^1H^\alpha$, F_1), 8.7 Hz ($^{13}C^\alpha$, F_2), 1.4 Hz (^{15}N, F_3) and 2.4 Hz ($^1H^N$, F_4) for aS, and 1.0 (^{15}N, F_1), 1.0 (^{15}N, F_2) and 2.4 Hz ($^1H^N$, F_3) for the α-synuclein HN(COCO)NH spectrum. Peak picking and spectrum analysis was carried out with the programs NMRPipe and SPARKY.[25]

2.3 Coil database

For all comparisons with backbone torsion angle distributions expected in disordered regions of proteins we used the recently introduced coil library of Mantsyzov *et al.* which consists of all

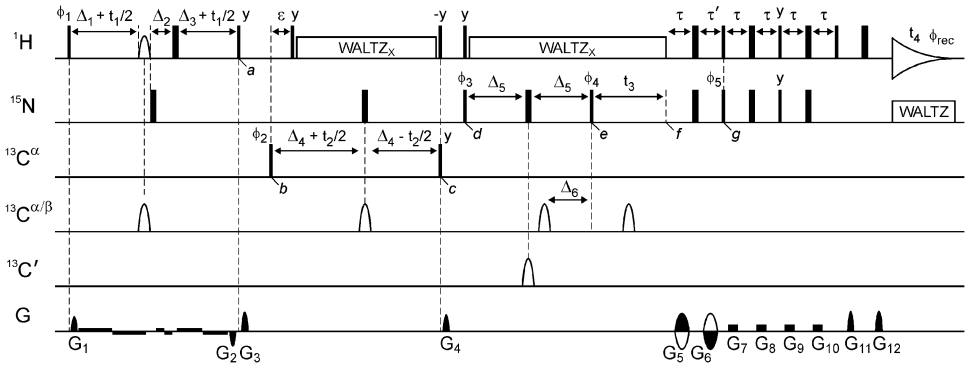

Fig. 1 Pulse scheme of the HACANH[C′] E.COSY experiment. Narrow and wide lines indicate non-selective 90° and 180° pulses, respectively, except for the 90° pulses on $^{13}C^\alpha$ that have their duration adjusted to have no excitation of $^{13}C'$. ^1H decoupling ($\gamma_H B_1$ = 3.1 kHz; phase x) is applied using the WALTZ16 scheme.[38] The carrier frequencies for ^1H, ^{15}N, $^{13}C^\alpha$, $^{13}C^{\alpha/\beta}$ and $^{13}C'$ are 4.7, 118, 56, 46 and 176 ppm, respectively. The $^1H^\alpha$ shaped pulse is of the ReBURP-type and has a duration of 3.5 ms. The $^{13}C'$ shaped pulse is of the hyperbolic-secant type and has a duration of 2 ms. The $^{13}C^\alpha$ shaped pulses are G3-shaped[53] with a duration of 350 μs, except for the pulse during $^{13}C^\alpha$ evolution (t_2) which is of the ReBURP-type[54] and has a duration of 420 μs. Note that the durations of all shaped pulses presented here are for measurements at 600 MHz ^1H frequency; these values scale inversely with field strength and should be adjusted accordingly when experiments are recorded at a field strength that differs from 14.1 T. Delays: ε = 2.5 ms, τ = 2.6 ms, $\tau' = \tau + t_3(0)$, Δ_1 = 1.7 ms − 0.92 × P_{rb}/2 where P_{rb} is the duration of the $^1H^\alpha$ ReBURP pulse and the scaling factor of 0.92 for $^1J_{C\alpha H\alpha}$ dephasing was measured experimentally for a pair of centered ^1H ReBURP and ^{13}C G3 shaped pulses with a duration ratio of 10 : 1, $\Delta_2 = \Delta_1 + \Delta_3$, Δ_3 = 0.17 ms, Δ_4 = 14 ms, Δ_5 = 15 ms and Δ_6 = 14 ms. All pulse phases are x, unless specified. Phase cycling: $\phi_1 = x$; $\phi_2 = x, -x$; $\phi_3 = x, x, -x, -x$; $\phi_4 = y, -y$; $\phi_5 = x$; and $\phi_{rec} = x, -x, -x, x$. The gradient pulses G_1–G_{12} (z axis) have strengths of 7, −7, 30.1, 28.7, 25.9, 25.9, 3.5, 3.5, 4.9, 4.9, 25.9 and 25.9 G cm^{-1}, with durations of 0.05, 0.05, 2, 1.3, 1, 2, 2.4, 2.4, 2.4, 2.4, 0.2 and 0.503 ms, respectively. Weak gradients (0.35 G cm^{-1}) with a polarity indicated in the figure were applied during $^1H^\alpha$ evolution (t_1) to minimize radiation damping of the solvent signal. Quadrature detection in the t_1 and t_2 dimensions was achieved by increasing ϕ_1 and ϕ_2, respectively, in the regular States-TPPI manner. Quadrature detection in the t_3 dimension is obtained by the Rance–Kay method,[37] inverting G_5, G_6 and ϕ_5.

 This journal is © the Owner Societies 2016

Published on 21 September 2015. Downloaded by NATIONAL INSTITUTES OF HEALTH on 26/10/2016 17:12:05.

longer (>3 residue) segments that lack intramolecular H-bonding in the PDB.[14] This library differs from the widely used Fitzkee coil library[13] mostly by the absence of tight β-turns and was developed as a reference for the MERA program,[14] which aims to provide a residue-specific ϕ/ψ distribution of IDPs and IDRs.

3 Measurement of $^3J_{C'H\alpha}$ from HACANH[C′] E.COSY

The exclusive-COSY (E.COSY) principle provides a convenient and accurate method for measurement of unresolved scalar coupling, J_{AB}, between a pair of nuclear spins A and B. E.COSY experiments require the presence of a third nuclear spin, C, that has a large, resolvable coupling to spin A and whose chemical shift can be correlated in a multi-dimensional NMR experiment with spin B.[26,27] Numerous E.COSY experiments have been introduced over the past few decades that permit the measurement of virtually every possible type of three-bond ^{1}H–^{1}H and heteronuclear J coupling in isotopically enriched proteins.[17,28–32] A condition for accurate measurement of the J_{AB} coupling from a B–C cross peak is that the spin-state of the passive spin, A, remains unperturbed. A reduction in the apparent coupling will result if the longitudinal relaxation time of spin A is insufficiently long relative to the delay durations used in the pulse scheme.[17,33] So, on the one hand it is important to select from the two spins, A and B, the spin with the longest T_1 relaxation time to function as the passive, unperturbed spin. On the other hand, as the J_{AB} displacement is read out from the relative displacement in resonance frequency of spin B when monitoring the components of the B–C cross peak, it is important that the line width of spin B is narrow. For measurement of $^3J_{C'H\alpha}$ by an E.COSY experiment, the choice is whether the ^{13}C′ or the ^{1}H$^\alpha$ should act as the passive spin. The ^{13}C′ T_1 relaxation times for IDPs at 600 MHz ^{1}H frequency and above are typically longer than 1 s, which is considerably longer than those of ^{1}H$^\alpha$ in ^{13}C-enriched proteins, thus favoring ^{13}C′ to be the passive spin. On the other hand, the transverse relaxation time of ^{13}C′ is longer than for ^{1}H$^\alpha$, favoring it to be the actively detected spin because it potentially offers higher frequency resolution. In our experiments, we have selected ^{13}C′ to be the passive spin, but enhance the spectral resolution of the ^{1}H$^\alpha$ spin by homonuclear decoupling it from ^{1}H^N and ^{1}H$^\beta$ in a band-selective manner.[34,35]

3.1 Description of the pulse scheme

We here used the HACANH[C′] E.COSY pulse scheme (Fig. 1) for measurement of the $^3J_{C'H\alpha}$ and $^3J_{HNC'}$ couplings. This triple resonance pulse scheme transfers magnetization in a unidirectional fashion from ^{1}H$^\alpha$ *via* ^{13}C$^\alpha$ and ^{15}N to ^{1}H^N and is similar in design to the early H(CA)NNH experiment,[36] but it lacks ^{13}C′ decoupling and it uses gradient-enhanced transfer from ^{15}N to ^{1}H for sensitivity enhancement and improved solvent signal suppression.[37] Other small but important differences will be discussed below. To minimize resonance overlap in crowded IDP spectra, we carry out the experiments in a 4D fashion, with evolution of ^{1}H$^\alpha$, ^{13}C$^\alpha$ and ^{15}N chemical shifts in the indirect dimensions, and direct detection of ^{1}H^N. The ^{13}C$^\alpha$ evolution period can be omitted if the measurement is carried out as a 3D experiment, potentially useful for small proteins such as ubiquitin and GB3, where resonance overlap is less of a concern. However, this will remove the possibility to also measure $^3J_{HNC'}$ from such a spectrum.

During the (t_1, ^{1}H$^\alpha$) evolution period, a ^{1}H$^\alpha$ band-selective pulse is paired with a non-selective ^{1}H 180° pulse (separated by the short Δ_2 delay, Fig. 1) to effectively invert all ^{1}H spins other than ^{1}H$^\alpha$, thereby removing ^{1}H–^{1}H J-dephasing during t_1 evolution and enhancing the attainable ^{1}H$^\alpha$ spectral resolution. Note that ^{1}H$^\alpha$–^{1}H$^\alpha$ J-dephasing for Ser and Thr residues is not refocused during t_1 because their ^{1}H$^\beta$ resonate within the bandwidth that is inverted by the ^{1}H$^\alpha$ band-selective pulse. Similarly, the geminal ^{1}H$^\alpha$–^{1}H$^\beta$ dephasing remains active for Gly residues. Evolution of $^1J_{C\alpha H\alpha}$ coupling during t_1 is adjusted to be active for ~3.4 ms, with most of the $^1J_{C\alpha H\alpha}$ dephasing actually taking place during the pair of simultaneously applied band-selective ^{1}H$^\alpha$ and ^{13}C$^\alpha$ pulses. As indicated in Fig. 1, very weak, antiphase gradients are also applied during the t_1 evolution period to prevent radiation damping of the ^1H_2O magnetization, which otherwise would rotate ^{1}H$^\alpha$ spins that resonate under the water resonance back to the z axis, making them effectively disappear from the spectrum. The antiphase G_1/G_2 gradient pulse pair serves to remove the effect of imperfections in the band-selective and non-selective 180° ^{1}H pulses, applied during t_1, while also eliminating any transverse ^{13}C magnetization. At time point a, the relevant, selected magnetization then is described by $2C_z^\alpha(i)H_z^\alpha(i)\cos(\delta_{H\alpha}t_1)\cos(\pi\,{}^2J_{H\alpha C'}t_1)\cos(\pi\,{}^3J_{H\alpha C'}t_1)$, where $\delta_{H\alpha}$ is the angular ^{1}H$^\alpha$ chemical shift frequency, and $^2J_{H\alpha C'}$ and $^3J_{H\alpha C'}$ are the two- and three-bond J couplings from ^{1}H$^\alpha$ to the intraresidue and preceding ^{13}C′, respectively.

During the subsequent (t_2, ^{13}C$^\alpha$) constant-time evolution period between time points b and c, $^1J_{C\alpha N}$ and $^2J_{C\alpha N}$ are active for the full constant-time duration $2\Delta_4$, and $^1J_{C\alpha H\alpha}$ rephasing is active for the duration ε, which is set to a compromise value (2.5 ms) that enables magnetization transfer to occur for both ^{13}C$^\alpha$ methine and methylene (Gly) groups. The constant-time duration is set to $2\Delta_4 \approx 1/{}^1J_{C\alpha C\beta} \approx 28$ ms such that the net effect of $^1J_{C\alpha C\beta}$ dephasing is eliminated. Broad-band ^{1}H-decoupling with the WALTZ16 modulation scheme[38] is used while care is taken that the water magnetization remains in a well-defined state, parallel to the decoupling radiofrequency field, thereby facilitating suppression of the intense water signal during the detection period. At time point c, following the ^{13}C$^\alpha$ evolution period, the $2C_z^\alpha(i)H_z^\alpha(i)$ terms from time point a are converted to $2C_z^\alpha(i)N_z(i)\sin(2\pi\,{}^1J_{C\alpha N}\Delta_4)\cos(2\pi\,{}^2J_{C\alpha N}\Delta_4) + 2C_z^\alpha(i)N_z(i+1)\sin(2\pi\,{}^2J_{C\alpha N}\Delta_4)\cos(2\pi\,{}^1J_{C\alpha N}\Delta_4)$. Undesired transverse magnetization terms are eliminated at this point by pulsed field gradient G_4. Between time points d and e, prior to (t_3, ^{15}N) evolution, a ^{12}C′ filter is applied to ensure removal of any signals from isotopomers where ^{15}N is bonded to a ^{12}C′ isotope (due to incomplete isotope labeling). In the absence of

Published on 21 September 2015. Downloaded by NATIONAL INSTITUTES OF HEALTH on 26/10/2016 17:12:05.

such a filter, we found that in the ^{15}N dimension of the final spectrum, the ^{15}N–{^{12}C′} signal approximately overlaps with the downfield component of the ^{15}N–{^{13}C′} doublet,[39] causing an artificial reduction in the apparent $^{3}J_{C'H\alpha}$ E.COSY splitting. A hyperbolic-secant-shaped ^{13}C′ pulse is used during this ^{12}C′ filter, such that complete ^{13}C′ inversion is not adversely impacted by RF inhomogeneity.[40] During the ^{12}C′ filter, a $^{13}C^{\alpha}$ pulse is applied which rephases a significant fraction of the $2C_z^{\alpha}(i)N_y(i)$ + $2C_z^{\alpha}(i)N_y(i+1)$ terms present at time point d to yield $2C_z'(i-1)N_y(i)$ and $2C_z'(i)N_y(i+1)$ at point e. Therefore, at time point e, (t_3, ^{15}N) evolution starts from antiphase ^{15}N–{^{13}C′} magnetization, which is in-phase with respect to $^{1}H^{N}$. Following t_3 evolution (time point f) ^{15}N magnetization is converted to the requisite antiphase $2N_yH_z$ terms during the duration $\tau + \tau'$, which also serves as the gradient-encoding delay (by oppositely signed gradient pulses G_5 and G_6), followed by standard gradient-enhanced transfer to $^{1}H^{N}$,[37] which is detected during the time t_4. Note that ^{13}C′ decoupling is absent during all chemical shift evolution periods, leading to the desired E.COSY effect. However, since both $^{13}C'(i)$ and $^{13}C'(i-1)$ play a role in the evolution of the various terms, discussed above, the 4D E.COSY multiplet pattern is somewhat more complex than in typical E.COSY applications. Below, we will therefore briefly discuss the fine structure observed in these 4D multiplets.

3.2 *J* couplings contained in the HACANH[C′] E.COSY spectrum

The 4D HACANH[C′] E.COSY spectrum contains a rich array of *J* coupling information involving the backbone carbonyl nuclei. Next to the "active" $^{1}J_{NC'}$ coupling, which yields the requisite well-resolved antiphase E.COSY splitting in the ^{15}N dimension and which contains potentially useful information on H-bonding,[41] the spectrum yields $^{1}J_{C'C\alpha}$, $^{2}J_{C'N}$, $^{3}J_{HNC'}$, $^{2}J_{C'H\alpha}$, and the main purpose of our study: $^{3}J_{C'H\alpha}$. The values of $^{1}J_{C'C\alpha}$, $^{2}J_{C'N}$, and $^{2}J_{C'H\alpha}$ show relatively little variation with structure, and so far have not yet found much use in terms of protein structure analysis.[42–44] Both $^{3}J_{HNC'}$ and $^{3}J_{C'H\alpha}$ correlate with ϕ *via* well-defined Karplus equations and are of most immediate structural value. Below, we therefore briefly illustrate how these couplings are manifested in the 4D spectrum.

3.3 Measurement of $^{3}J_{C'H\alpha}$ and $^{2}J_{C'H\alpha}$

As discussed above, two types of magnetization transfer take place from $^{13}C^{\alpha}$ to ^{15}N: intraresidue, mediated by $^{1}J_{C\alpha N}$ and sequential, to ^{15}N of residue $i+1$, mediated by $^{2}J_{C\alpha N}$. Fig. 2a concerns the intraresidue correlation observed for K102, and illustrates the relevant magnetization transfer pathway. Note that in the ^{15}N dimension, the cross peak is split by the presence of $^{1}J_{C'N}$ to the $^{13}C'(i-1)$ of the preceding residue. The splitting is antiphase due to the presence of the ^{12}C′ filter (between time points d and e in Fig. 1), and the displacement in the $^{1}H^{\alpha}$ dimension corresponds to the desired $^{3}J_{C'H\alpha}$ coupling. The spin state of the ^{13}C′ is inverted between $^{1}H^{\alpha}$ and ^{15}N evolution, giving rise to the negative slope of the doublets in Fig. 2a. However, after accounting for this sign inversion, and knowing that the sign of $^{1}J_{C'N}$ is negative (due to the negative magnetogyric ratio of ^{15}N), this slope corresponds to a positive value for the $^{3}J_{C'H\alpha}(i-1, i)$ coupling. The $H^{\alpha}(i) \rightarrow H^{N}(i)$ cross peak appears twice, as it is split in the $^{13}C^{\alpha}$ dimension by $^{13}C'(i)$, therefore providing duplicate measurements for each $^{3}J_{C'H\alpha}$ coupling (Fig. 2a). Comparison of such pairs of measured $^{3}J_{C'H\alpha}$ couplings provides a measure for their reproducibility, and averaging their values is used to reduce their random uncertainty by $\sqrt{2}$.

Analogous to the $^{1}J_{C\alpha N}$ transfer pathway discussed above, transfer *via* $^{2}J_{C\alpha N}$ yields a $H^{\alpha}(i) \rightarrow H^{N}(i+1)$ cross peak (Fig. 2b). Here, the $^{15}N(i+1)$ and $^{13}C^{\alpha}(i)$ frequencies of the cross peak depend on the spin state of $^{13}C'(i)$ only, yielding a single measurement of $^{2}J_{C'H\alpha}$ when extracting this value from the difference in $^{1}H^{\alpha}$ frequency of the two multiplet components. An upfield $^{1}H^{\alpha}$ displacement together with a downfield $^{13}C^{\alpha}$ displacement for the red *versus* the blue multiplet component in Fig. 2b points to opposite signs for $^{1}J_{C\alpha C'}$ and $^{2}J_{C'H\alpha}$. Note that $^{2}J_{C'H\alpha}$ can also be obtained from Fig. 2a, from the relative $^{1}H^{\alpha}$ displacement of the duplicate doublets shown in its two panels.

3.4 Measurement of $^{3}J_{HNC'}$

As mentioned above, the intraresidue $H^{\alpha}(i) \rightarrow H^{N}(i)$ cross peak is split in the $^{13}C^{\alpha}$ dimension by $^{1}J_{C\alpha C'}$, with the upfield and downfield components corresponding to ^{13}C′ in the $|\alpha>$ and $|\beta>$ spin states, respectively. Therefore, when viewing $^{13}C^{\alpha}$–$^{1}H^{N}$ cross-sections of the 4D spectrum taken through the cross peak of Fig. 2a, an E.COSY tilt reflecting the intraresidue $^{3}J_{HNC'}$ is seen, with the $^{1}J_{C\alpha C'}$ splitting in the $^{13}C^{\alpha}$ dimension, and $^{3}J_{HNC'}$ in the $^{1}H^{N}$ dimension (Fig. 2c). Again this coupling is obtained in duplicate by viewing the two cross-sections obtained at $\delta_N \pm {}^{1}J_{NC'}/2$ (Fig. 2c).

As with most 4D NMR experiments, the HACANH[C′] E.COSY spectrum can also be recorded in fewer dimensions, and here this can be done most conveniently by not incrementing the duration the $^{13}C^{\alpha}$ evolution period (*i.e.*, keeping $t_2 = 0$). However, in practice we find that overlap between the sequential and intraresidue $H^{\alpha} \rightarrow H^{N}$ cross peaks is rather common, in particular in disordered proteins. In folded proteins, the 3D version of the experiment is a viable option, although reducing the dimensionality of the experiment in this manner removes the ability to measure $^{3}J_{HNC'}$ from the spectrum.

3.5 Reproducibility of the measured $^{3}J_{C'H\alpha}$ and $^{3}J_{HNC'}$ values

The random measurement uncertainty, σ, of a peak position in a frequency domain spectrum is directly proportional to the resonance line width, LW, and inversely related to the signal-to-noise ratio, SNR. To a fair approximation, it is given by:[45]

$$\sigma = LW/(2SNR) \quad (1)$$

The uncertainty in the *J* splitting requires propagation of the uncertainty in both of its components and is therefore $\sqrt{2}$ larger. For time domain signals that have not fully decayed, as typically applies in the indirect dimensions of multi-dimensional NMR experiments, the attainable line width is approximately given by 1/AT, where AT is the length of the acquisition time in

 This journal is © the Owner Societies 2016

Published on 21 September 2015. Downloaded by NATIONAL INSTITUTES OF HEALTH on 26/10/2016 17:12:05.

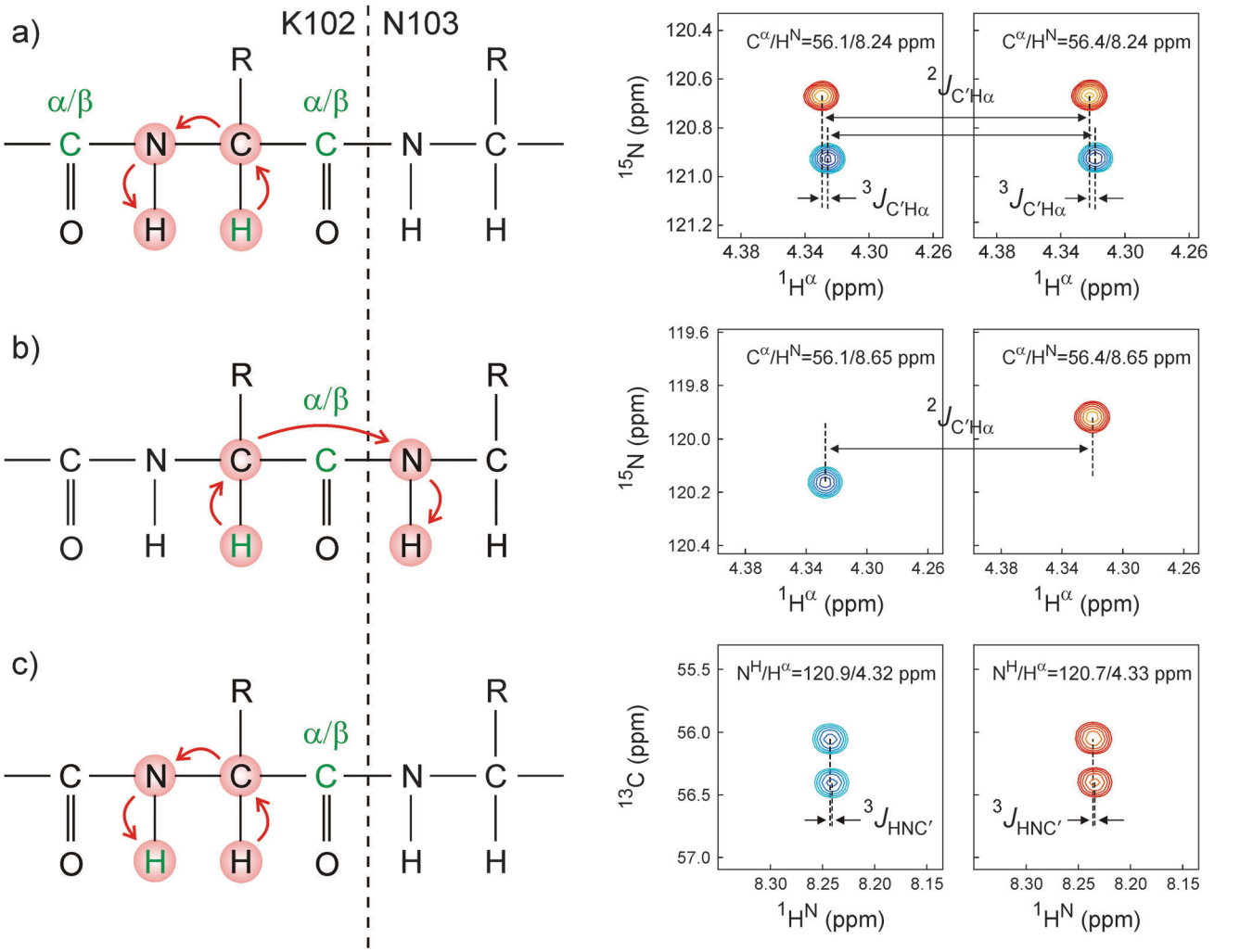

Fig. 2 Analysis of the 4D HACANH[C′] E.COSY spectrum, shown for couplings involving residue K102 in aS. The left hand side of each panel denotes the magnetization transfer pathway, with active and passive nuclei whose pairwise J coupling is measured shown in green. The right hand side of each panel shows small cross sectional regions through the 4D NMR spectrum, illustrating the measurement of various J couplings. (a) Measurement of $^{3}J_{C'H\alpha}$ and $^{2}J_{C'H\alpha}$ from ^{15}N–$^{1}H^{\alpha}$ cross sections. Two independent measurements are shown, separated in the orthogonal $^{13}C^{\alpha}$ dimension by $^{1}J_{C\alpha C'}$. Relative $^{1}H^{\alpha}$ displacement within each panel corresponds to $^{3}J_{C'H\alpha}$ relative $^{1}H^{\alpha}$ displacement between the two panels corresponds to $^{2}J_{C'H\alpha}$. (b) Alternative measurement of $^{2}J_{C'H\alpha}$ from the sequential $^{13}C^{\alpha}$–$^{1}H^{\alpha}$ (K102) to ^{15}N–$^{1}H^{N}$ (N103) cross peak. Relative displacement in the $^{1}H^{\alpha}$ dimension corresponds to $^{2}J_{C'H\alpha}$. ^{15}N displacement corresponds to $^{1}J_{NC'}$, and $^{13}C^{\alpha}$ displacement to $^{1}J_{C\alpha C'}$. (c) Measurement of $^{3}J_{HNC'}$ from a $^{13}C^{\alpha}$–$^{1}H^{N}$ cross section. Relative $^{1}H^{N}$ displacement within each panel equals $^{3}J_{HNC'}$ and relative $^{13}C^{\alpha}$ displacement equals $^{1}J_{C\alpha C'}$. ^{15}N displacement between the two panels equals $^{1}J_{NC'}$; relative $^{1}H^{N}$ displacement between the two panels equals $^{2}J_{HNC'}$.

the dimension considered. This consideration indicates that it is important to sample the signals in the dimension where the coupling is measured for as long as practically feasible (up to *ca.* $2 \times T_2$, where T_2 is the decay constant of the signal), and that it is desirable to minimize the apparent decay rate through the removal of ^{1}H–^{1}H J splittings, as is done in the scheme of Fig. 1 by band-selective homonuclear decoupling in the $^{1}H^{\alpha}$ dimension. Band-selective homonuclear decoupling in the detected $^{1}H^{N}$ dimension of the spectrum would be advantageous for enhancing $^{1}H^{N}$ resolution and thereby the accuracy of $^{3}J_{HNC'}$. However, this mode of decoupling in the directly detected dimension is technically more challenging in terms of water signal suppression on cryogenic probes, which typically have only a single-axis pulsed field gradient accessory, and therefore was not used in our experiments. Note also that $^{1}H^{N}$ is only split by $^{1}H^{\alpha}$, making homonuclear decoupling less advantageous than it is for $^{1}H^{\alpha}$, where the couplings to both $^{1}H^{N}$ and $^{1}H^{\beta}$ are removed. Overall, the line width in the $^{1}H^{N}$ dimension remains somewhat larger than in the homonuclear decoupled $^{1}H^{\alpha}$ dimension, and with the applicable SNR (*cf.* eqn (1)) being the same for defining the peak position precision, the random error in the $^{3}J_{HNC'}$ will be correspondingly larger than for $^{3}J_{C'H\alpha}$.

Each intraresidue $H^{\alpha} \rightarrow H^{N}$ cross peak consists of four multiplet components, corresponding to the splitting by $^{13}C'(i)$ and $^{13}C'(i-1)$ (Fig. 2a and c) whereas each sequential $H^{\alpha}(i) \rightarrow H^{N}(i+1)$ cross peak is split into doublets by $^{13}C'(i)$ (Fig. 2b). Therefore, the number of resonances in the 4D spectrum approximately equals $6N$, where N is the number of

Published on 21 September 2015. Downloaded by NATIONAL INSTITUTES OF HEALTH on 26/10/2016 17:12:05.

residues in the protein. This number typically is small relative to the total number of time domain data points sampled in the indirect frequency dimensions, meaning that the HACANH[C′] E.COSY spectrum is relatively sparse and well suited for the use of non-uniform sampling (NUS) techniques.[18–20,46–48] Here, spectral reconstruction was carried out by the in-house written SMILE routine of the NMRPipe program[24] which integrates accurate spectral reconstruction with spectral resolution enhancement by using a method that effectively mimics multi-dimensional linear prediction.[49]

The availability of duplicate measurements for $^3J_{\mathrm{HNC'}}$ and $^3J_{\mathrm{C'H\alpha}}$ makes evaluation of their random uncertainty straightforward, and permits evaluation of how closely this random error approaches the theoretical limit set by eqn (1). As can be seen in Fig. 3a, the pairwise root-mean-square deviation (RMSD) between the two sets of $^3J_{\mathrm{C'H\alpha}}$ is 0.11 Hz, and averaging the two sets of values therefore reduces the random uncertainty to 0.06 Hz.

The average signal-to-noise ratio for intra-residue cross peaks is *ca* 120 : 1, and line widths are approximately 10 Hz ($^1H^{\alpha}$ of non-T/S/G), *ca.* 15–20 Hz for T/S/G; 25 Hz ($^{13}C^{\alpha}$), 5 Hz (^{15}N) and 15–20 Hz ($^1H^N$). Eqn (1) can then be used to predict precisions of *ca.* 0.1 and 0.06 Hz for $^3J_{\mathrm{HNC'}}$ and $^3J_{\mathrm{C'H\alpha}}$, respectively. As can be seen from the pairwise RMSD values observed for the duplicate sets of $^3J_{\mathrm{H\alpha C'}}$ and $^3J_{\mathrm{HNC'}}$ (Fig. 3), these experimental precisions fall close to their theoretical limit, indicating that the NUS reconstruction process used in our analysis faithfully reproduces peak positions.

The processed NMR spectrum containing the $^3J_{\mathrm{H\alpha C'}}$ and $^3J_{\mathrm{HNC'}}$ splittings were generated from a 4D NUS time domain matrix that included measured data for only 0.55% of its elements, *i.e.*, with 99.45% of them initially set to zero and subsequently adjusted during the iterative NUS processing procedure to values that approach those that would have been sampled if measurements for these time points had been made. As mentioned above, with only *ca.* 800 resonances in the HACANH[C′] E.COSY spectrum of aS, the total number of time domain data points sampled in the indirect dimensions is far larger than this number of 800 resonances, making the NUS reconstruction feasible. It is therefore interesting to test

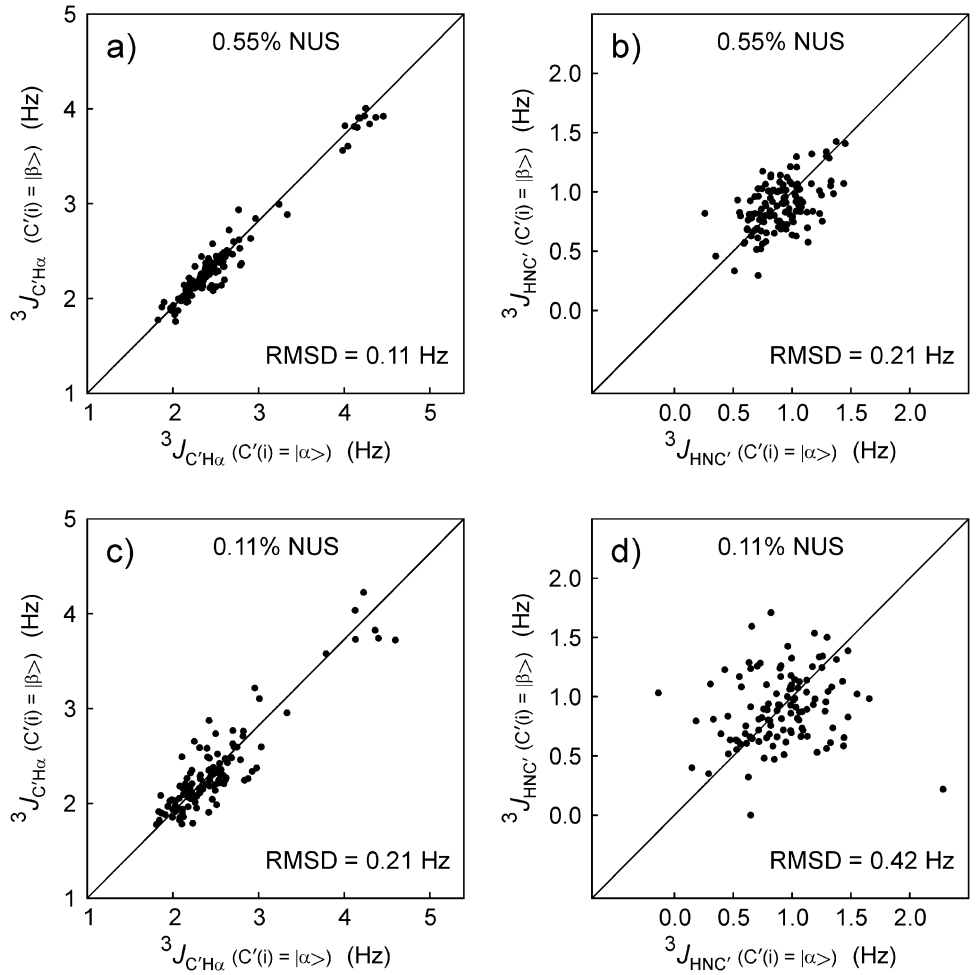

Fig. 3 Reproducibility of the duplicate measurements of ^3J couplings from the 4D NUS HACANH[C′] E.COSY spectrum of α-synuclein. Comparisons of the duplicate $^3J_{\mathrm{C'H\alpha}}$ measurements (see right panels of Fig. 2a) obtained when (a) using 0.55% NUS sampling and (c) 0.11% NUS sampling. Analogous comparisons of the two $^3J_{\mathrm{HNC'}}$ measurements (see right two panels of Fig. 2c) obtained when using (b) 0.55% and (d) 0.11% NUS sampling.

 This journal is © the Owner Societies 2016

Published on 21 September 2015. Downloaded by NATIONAL INSTITUTES OF HEALTH on 26/10/2016 17:12:05.

whether a similar quality spectrum can be reconstructed from even fewer sampled data. For this purpose, we also reconstructed a 0.11% sampled data set, representing 20% of the same data used above for spectral reconstruction, and consequently yielding $\sqrt{5}$ lower SNR. As can be seen in Fig. 3c and d, the reproducibility of the measured J couplings then also decreases by a factor of *ca.* $\sqrt{5}$, indicating that no additional uncertainty in extracted peak position is added by the NUS processing over what is expected on the basis of simple signal-to-noise considerations, even for a sparsity as low as 0.11%.

4 Karplus equation for $^3J_{C'H\alpha}$

An empirically parameterized $^3J_{C'H\alpha}$ Karplus equations was introduced on the basis of values measured for ubiquitin using a 3D E.COSY method conceptually analogous to the 4D experiment of Fig. 1.[32] Comparison of the previously measured and new ubiquitin data shows pairwise RMSD values of 0.29 Hz relative to the earlier measurements. Considering that this RMSD is more than $\sqrt{2}$ higher than the estimated uncertainty in the newly acquired data, we conclude that the newer data is of higher accuracy than the earlier data, and we will only use the new data for any fitting and analysis.

Best fits of the Karplus equation to the experimental $^3J_{C'H\alpha}$ values and the coordinates of the NMR-determined structure of ubiquitin (PDB entry 2MJB)[50] and the RDC-refined structure of GB3 (PDB entry 2OED)[23] are shown in Fig. 4 (dashed line).

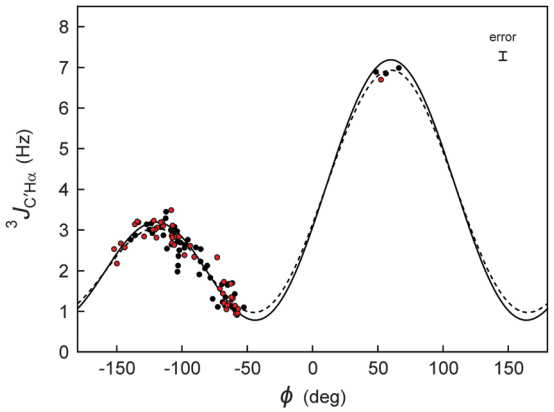

Fig. 4 Plot of experimental $^3J_{C'H\alpha}$ coupling values against the backbone torsion angle ϕ, for ubiquitin (black symbols) and GB3 (red). The backbone ϕ values are derived from RDC-refined structures of GB3 (PDB entry 2OED) and ubiquitin (PDB entry 2MJB), both obtained without $^3J_{C'H\alpha}$ data. The superimposed best-fit Karplus equation, shown as a dashed curve, is given by $^3J_{C'H\alpha} = 3.76 \times \cos^2(\phi - 60^\circ) + 1.95 \times \cos(\phi - 60^\circ) + 1.23$. After factoring out the effect of Gaussian ϕ angle fluctuations ($\sigma = 0.226$),[51,52] the rigid limit Karplus equations (solid line) is given by: $^3J_{C'H\alpha} = 4.17 \times \cos^2(\phi - 60^\circ) + 2.00 \times \cos(\phi - 60^\circ) + 1.02$. Residues with elevated backbone dynamics (L12, D40, and G41 for GB3; T9-K11, D32-G35, G47, D52, and R72-G76 for ubiquitin) are excluded from the plot and the fitting process. The error bar (top right) corresponds to the random error (±1 standard deviation) based on the reproducibility of the $^3J_{C'H\alpha}$ measurements. The considerably larger RMSD of 0.25 Hz relative to the best-fitted Karplus curve (dashed line) reflects the impact of factors other than ϕ on the $^3J_{C'H\alpha}$ value.

The RMSD between observed and Karplus-predicted $^3J_{C'H\alpha}$ values equals 0.25 Hz. As expected on the basis of the higher measurement precision reached in the current study, this value is somewhat lower than obtained previously.[17] However, it remains considerably higher than the random uncertainties in the measurement. This latter observation indicates that, similar to what was previously noted for the $^3J_{C'C'}$ coupling,[23] parameters other than ϕ can also influence these couplings. These other factors, which may include valence angle distortion of, for example, the τ angle (N–C$^\alpha$–C′) and/or H-bonding of amide and carbonyl groups, are difficult to account for in experimental structures, and therefore limit the extent to which agreement between experimental $^3J_{C'H\alpha}$ and predicted values can be enforced during the calculation of structures of folded proteins. In IDPs and IDRs, the effects of factors other than ϕ on the ^3J couplings are expected to average to roughly zero, and should permit interpretation of these couplings in a more quantitative fashion, which then can approach the experimental uncertainty in the measurement of these ^3J couplings.[14]

The best-fit Karplus curve, described above, was derived using a single, static representation of the protein structure. However, it is well recognized that proteins are subject to modest amplitude ϕ-angle fluctuations, previously estimated to have an RMS amplitude of *ca.* 13° in regions of well-ordered structure.[51] When factoring out the effect of these backbone fluctuations,[52] slightly different, ''rigid-limit'' Karplus parameters are obtained that should be applicable when refining a structure in terms of an ensemble of perfectly static conformers (solid line in Fig. 4), and this latter curve is therefore used when describing an IDP or IDR in terms of an ensemble representation.[14]

Tables containing the experimental $^3J_{C'H\alpha}$ and $^2J_{C'H\alpha}$ values for α-synuclein, ubiquitin, and GB3, together with the α-synuclein $^3J_{HNC'}$ values are included as ESI.†

5 Fractional population of positive ϕ angles from ^3J couplings

Under the assumption that all ϕ angles are negative, we previously were able to extract two parameters, the average $\langle\phi^-\rangle$ and its standard deviation σ, from two experimental observables: $^3J_{HNH\alpha}$ and $^3J_{C'C'}$.[51] Although a small population, P^+, of conformers with positive ϕ values in IDPs was shown to have minimal impact on the extracted $\langle\phi^-\rangle$ and σ values, the actual value of P^+ could not be determined from that data. As we will show below, the large difference between $^3J_{C'H\alpha}$ coupling values for negative and positive values of ϕ now allows determination of P^+. To this extent, we name the $^3J_{C'H\alpha}$ values that would correspond to the fractions of conformers with only negative and only positive ϕ angles $^3J_{C'H\alpha}(-)$ and $^3J_{C'H\alpha}(+)$, respectively. Considering that the positive ϕ angles sampled in the PDB, and in particular in the coil library,[13,14] cluster relatively tightly around $\langle\phi^+\rangle$ values, this permits us to make an amino-acid-type-specific estimate for the average, $\langle^3J_{C'H\alpha}(+)\rangle$, by using the positive-$\phi$ conformers of the coil library in concert

Published on 21 September 2015. Downloaded by NATIONAL INSTITUTES OF HEALTH on 26/10/2016 17:12:05.

with the rigid-limit Karplus equation. Note that the effect of dynamic averaging is included when calculating $\langle {}^3J_{C'H\alpha}(+)\rangle$ over the coil library distribution of positive ϕ conformers, and that it therefore is the rigid-limit parameterization of the Karplus curve that needs to be used for predicting $\langle {}^3J_{C'H\alpha}(+)\rangle$.

5.1 Iterative procedure for determination of P^+

Since ${}^3J_{HNH\alpha}$ and ${}^3J_{C'C'}$ coupling values for averaged negative ϕ angles in coil regions ($\langle\phi^-\rangle$, typically between -80 and -100°) are similar to the coupling values averaged for positive ϕ angles, experimental ${}^3J_{HNH\alpha}$ and ${}^3J_{C'C'}$ couplings are not much impacted by a small fraction of conformers with positive ϕ values. Therefore, to a first approximation, they may be ignored when calculating residue-specific $\langle\phi^-\rangle$ and σ values.[51] These $\langle\phi^-\rangle$ and σ values then can be used to calculate ${}^3J_{C'H\alpha}(-)$ from the dynamics-corrected[52] form of the ${}^3J_{C'H\alpha}$ Karplus equation (see also Section 6, eqn (4)). Next, P^+ for each residue can be derived from

$$ {}^3J_{C'H\alpha} = P^+\langle {}^3J_{C'H\alpha}(+)\rangle + (1 - P^+){}^3J_{C'H\alpha}(-), \qquad (2) $$

where ${}^3J_{C'H\alpha}$ is the experimentally observed value. In a subsequent iteration, this P^+ value can be used to recalculate the ${}^3J_{HNH\alpha}(-)$ and ${}^3J_{C'C'}(-)$ values, yielding an improved estimate for $\langle\phi^-\rangle$ and σ, and an improved ${}^3J_{C'H\alpha}(-)$ value. This then allows a better estimate of P^+ in this second iteration. More iterations could be carried out, but in practice the procedure converges rapidly (Fig. 5a) and a single iteration suffices. With an RMSD of only 2.5%, excellent agreement is seen (Fig. 5b) between the P^+ values derived in this manner and those previously obtained by MERA analysis[14] of a large set of NMR input parameters (incl. ${}^{13}C$ and ${}^{15}N$ chemical shifts, ${}^1H$–${}^1H$ NOEs, and five different types of ${}^1J$, ${}^2J$ and ${}^3J$ couplings).

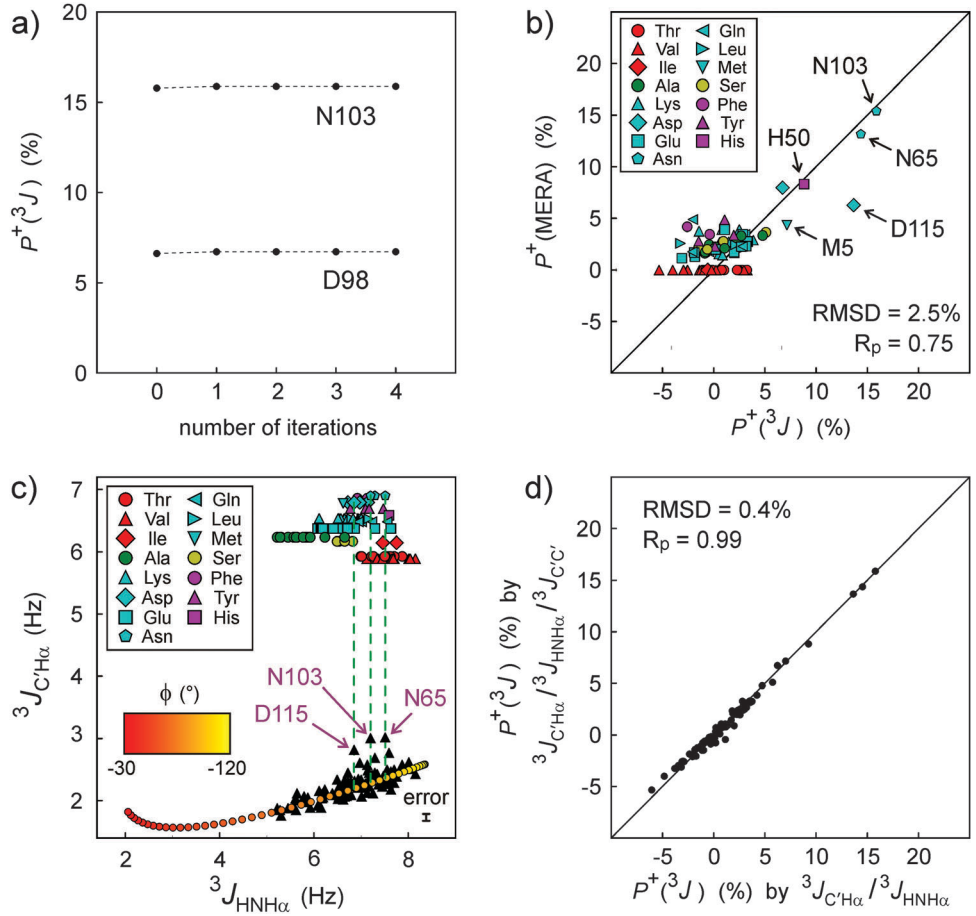

Fig. 5 Deriving the fractional population of positive ϕ angles, P^+, in IDPs and IDRs. (a) P^+ for α-synuclein residues N103 and D98 during successive iterations of ${}^3J_{HNH\alpha}$, ${}^3J_{C'C'}$, and ${}^3J_{C'H\alpha}$ analysis (see main text). (b) Comparison of the P^+ derived from iterative ${}^3J$ analysis (${}^3J_{C'H\alpha}$, ${}^3J_{HNH'}$, and ${}^3J_{C'C'}$) with P^+ obtained with the MERA webserver[14] without using ${}^3J_{C'H\alpha}$ as input parameter. Different residue types are presented by the symbols shown in the inset. Full results are listed in ESI,† Table S5. (c) Extracting P^+ from only ${}^3J_{C'H\alpha}$ and ${}^3J_{HNH\alpha}$ by interpolation. Red-to-yellow color shaded points correspond to the calculated correlation between ${}^3J_{C'H\alpha}$ and ${}^3J_{HNH\alpha}$ dynamic Karplus equations (eqn (4)), assuming $\sigma = 29.5^\circ$ (average IDP σ value previously derived for α-synuclein[51]). Shaped symbols mark the residue-type specific $\langle {}^3J_{C'H\alpha}\rangle$ values calculated from the Mantsyzov coil library for residues with $\phi > 0$.[14] Pairs of ${}^3J_{HNH\alpha}$ and ${}^3J_{C'H\alpha}$ values are shown as black filled triangles, all in the ${}^3J_{C'H\alpha}$ = 2–3 Hz range. Residues with ${}^3J_{C'H\alpha}$ that are significantly higher than expected from ${}^3J_{HNH\alpha}$ when assuming $P^+ = 0$ are marked by arrows and residue number. The error bar (bottom right) corresponds to the random error (±1 standard deviation) in the ${}^3J_{C'H\alpha}$ measurements. The random error in ${}^3J_{HNH\alpha}$ (±0.05 Hz) is too small for display purposes. (d) Comparison of P^+ calculated by iterative analysis of ${}^3J_{C'H\alpha}$, ${}^3J_{HNH\alpha}$ and ${}^3J_{C'C'}$ with those of simplified analysis (panel c method) of ${}^3J_{C'H\alpha}$ and ${}^3J_{HNH\alpha}$.

 This journal is © the Owner Societies 2016

Published on 21 September 2015. Downloaded by NATIONAL INSTITUTES OF HEALTH on 26/10/2016 17:12:05.

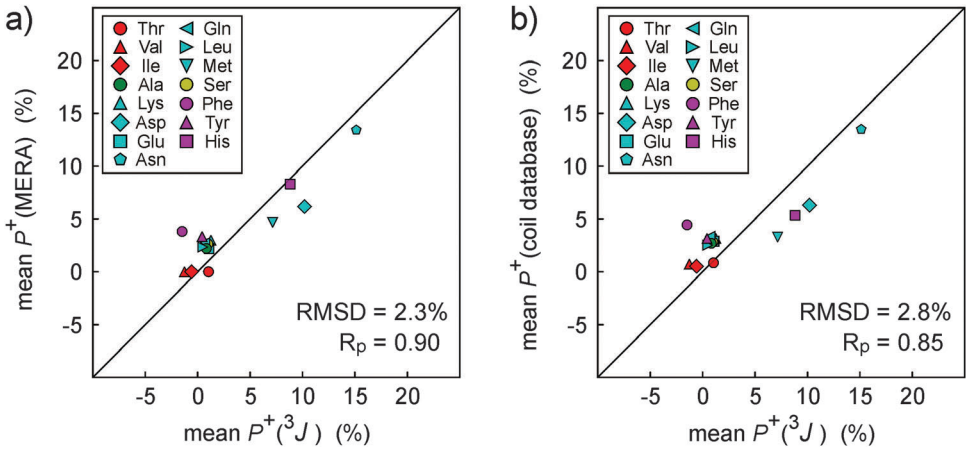

Fig. 6 Positive ϕ angle propensity in α-synuclein for different residue types. (a) Comparison of P^+ obtained from MERA analysis (lacking $^3J_{C'H\alpha}$ as input) with results of iterative analysis of $^3J_{C'H\alpha}$, $^3J_{HNH\alpha}$ and $^3J_{C'C'}$. (b) Comparison of P^+ obtained from iterative analysis of $^3J_{C'H\alpha}$, $^3J_{HNH\alpha}$ and $^3J_{C'C'}$ with residue-type averaged P^+ values of the Mantsyzov coil library.[14]

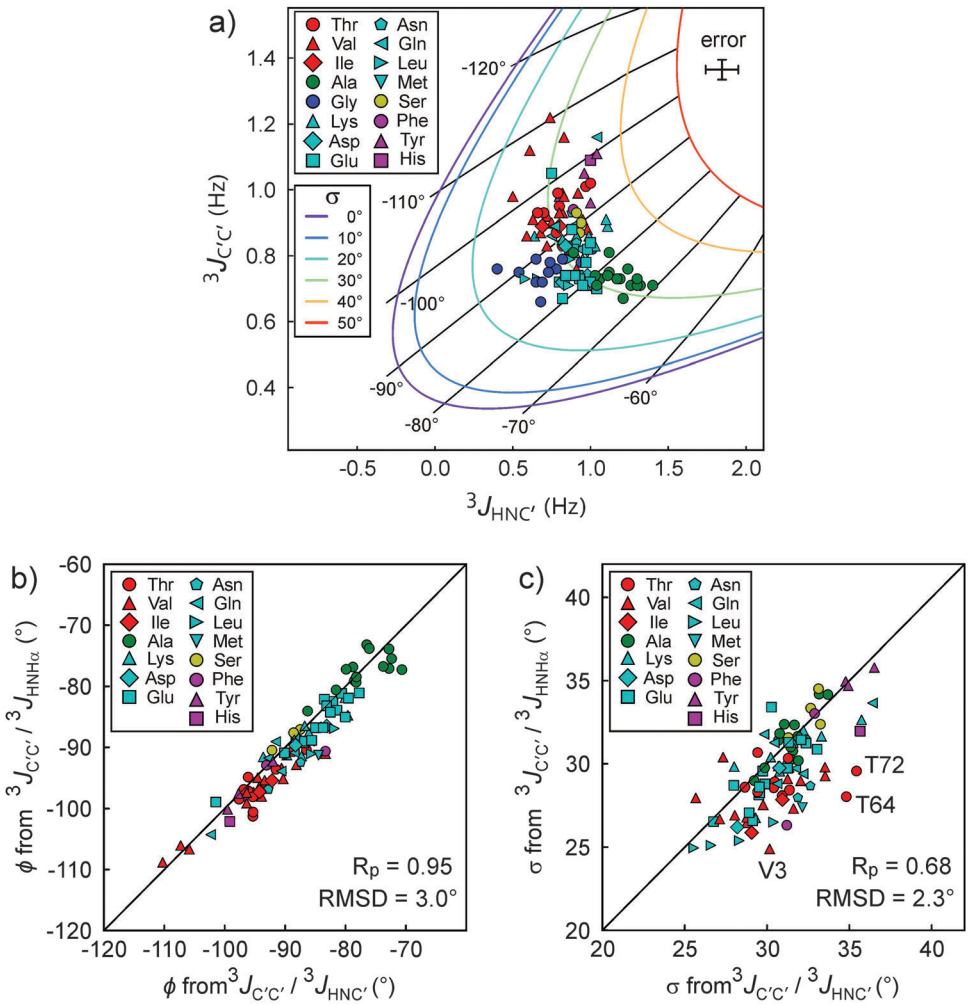

Fig. 7 Extracting the amplitude of ϕ angle fluctuations from $^3J_{C'C'}$ and $^3J_{HNC'}$. (a) Plot of $^3J_{C'C'}$ *versus* $^3J_{HNC'}$ for α-synuclein. The rigid limit (after factoring out σ = 13° motions) $^3J_{HNC'}$ Karplus parameterization[17] was used: $^3J_{HNC'}$ = 4.78 × $\cos^2(\phi)$ + 0.86 × $\cos(\phi)$ − 0.23 Hz. Solid lines show correlations between $^3J_{C'C'}$ and $^3J_{HNC'}$ predicted by the dynamics-corrected Karplus equations for ϕ angles (ϕ values marked by black "radial spokes"), with the different colors representing the amplitude of σ. The error bars (top right) corresponds to the random error (±1 standard deviation) of the $^3J_{C'H\alpha}$ and $^3J_{C'C'}$ measurements. (b) ϕ and (c) σ values that are extracted from the $^3J_{C'C'}$ and $^3J_{HNC'}$ analysis are compared to those from $^3J_{C'C'}$ and $^3J_{HNH\alpha}$ analysis, reported previously.[51]

Published on 21 September 2015. Downloaded by NATIONAL INSTITUTES OF HEALTH on 26/10/2016 17:12:05.

5.2 Direct determination of P^+ from $^3J_{HNH\alpha}$ and $^3J_{C'H\alpha}$

As an alternative to the above iterative method, a good approximation for P^+ can be obtained directly from the experimental $^3J_{HNH\alpha}$ and $^3J_{C'H\alpha}$ values without requiring knowledge of $^3J_{C'C'}$: The fact that to a good approximation $^3J_{HNH\alpha}$ and $^3J_{C'H\alpha}$ are linearly correlated over the range from $-60^\circ < \phi < -120^\circ$ (Fig. 5c) permits a direct prediction of $^3J_{C'H\alpha}$ ($^3J^{pred,neg}_{C'H\alpha}$) under the assumption that only this range of ϕ is populated. In practice, higher values may be observed, and P^+ can be calculated from

$$P^+ = ({}^3J^{obs}_{C'H\alpha} - {}^3J^{pred,neg}_{C'H\alpha})/({}^3J^{pred,pos}_{C'H\alpha} - {}^3J^{pred,neg}_{C'H\alpha}) \quad (3)$$

where $^3J^{pred,pos}_{C'H\alpha}$ is the residue-type-specific predicted value for positive coil angles, spanning a narrow range from 6 to 7 Hz for the different residue types (Fig. 5c; ESI,† Table S4). In practice, the iterative analysis of $^3J_{C'H\alpha}$, $^3J_{C'C'}$ and $^3J_{HNH\alpha}$ and the simple interpolation method of Fig. 5c yield very similar P^+ values (Fig. 5d).

The newly acquired $^3J_{C'H\alpha}$ data allows us to revisit the propensities to adopt positive ϕ angles for different residue types in the IDP α-synuclein. We find that P^+ values, averaged by residue type, agree very closely with those obtained from MERA analysis (Fig. 6a), which did not use $^3J_{C'H\alpha}$ as an input parameter. The only slight outlier in the correlation concerns Phe residues, which show a weak propensity (~4%) for positive ϕ angles, both in the MERA results and in the coil database (Fig. 6b), whereas $^3J_{C'H\alpha}$ points to a vanishingly low population. However, with $^3J_{C'H\alpha}$ data available for only two Phe residues in α-synuclein, the statistical uncertainty in the Phe $\langle P^+ \rangle$ values obtained for α-synuclein is rather large.

6 Amplitudes of ϕ angle fluctuations from ^3J couplings

Assuming that $^3J_{HNH\alpha}$ and $^3J_{C'C'}$ couplings only depend on the intervening backbone torsion angle, ϕ, the amplitude of ϕ angle fluctuations (assumed to follow a Gaussian distribution) can be extracted from such a pair of measurements.[51] This approach relies on the motion-corrected Karplus equation,[52]

$$^3J = A'\cos^2\theta + B'\cos\theta + C' \quad (4)$$

where $A' = A\exp(-2\sigma^2)$, $B' = B\exp(-\sigma^2/2)$, and $C' = C + A[1 - \exp(-2\sigma^2)]/2$, with A, B, and C being the rigid limit Karplus coefficients, θ is the averaged intervening dihedral angle, and σ is its standard deviation.

Our current study measures $^3J_{C'H\alpha}$ and $^3J_{HNC'}$ and this pair of couplings, or any other pairwise combination of the four sets of $^3J_{C'H\alpha}$, $^3J_{HNC'}$, $^3J_{C'C'}$, or $^3J_{HNH\alpha}$ can be used for analysis of the amplitude of σ, provided that their respective Karplus equations are not linearly related. Previously, we used the combination of $^3J_{HNH\alpha}$ and $^3J_{C'C'}$ to extract residue-specific σ values for α-synuclein.[51] Here, we demonstrate that σ values can be extracted from pairs of $^3J_{C'C'}$ and $^3J_{HNC'}$ values (Fig. 7). Colored lines in Fig. 7a correlate the $^3J_{HNC'}$ and $^3J_{C'C'}$ values predicted by their respective Karplus equations for different values of $\langle\phi\rangle$ (marked by black, radial spokes), with different colors corresponding to different amplitudes of σ. As can be seen from the figure, the plot of $^3J_{HNC'}$ *versus* $^3J_{C'C'}$ values for α-synuclein places the σ values of all residues in the 25–35° range. With an RMSD of only 3°, the residue-specific $\langle\phi\rangle$ values extracted from this figure agree closely with the corresponding $\langle\phi\rangle$ values previously obtained from $^3J_{HNH\alpha}$ and $^3J_{C'C'}$ analysis (Fig. 7b), and similarly, with an RMSD of only 2.3°, the σ values also agree very well with our prior study (Fig. 7c).[51]

7 Concluding remarks

We have shown that $^3J_{C'H\alpha}$ and $^3J_{HNC'}$ can be measured simultaneously and at high precision by using a 4D NMR pulse sequence. Only when such experiments are carried out in conjunction with very sparse non-uniform sampling does it become experimentally feasible to use long acquisition times in all four dimensions. These long acquisition times are required for obtaining the very high spectral resolution that is needed for precise measurement of very small J couplings. Although many of the widely used spectral analysis software packages are not yet fully optimized for convenient analysis of the extremely large 4D matrices, there is no conceptual or technical problem in obtaining resonance positions at the same accuracy as in spectra of lower dimensionality. Our analysis shows that the uncertainty in the extracted 4D peak positions is primarily determined by the intrinsic signal to noise ratio in the acquired time domain data. Indeed, the uncertainty in peak positions when using only a fraction of this data for NUS reconstruction is found to scale approximately with the inverse square root of the size of this fraction.

The large difference in $^3J_{C'H\alpha}$ values for residues with negative and positive ϕ angles makes this coupling particularly useful for identifying fractional populations of positive ϕ angles in disordered regions of a protein. In particular, when used in conjunction with $^3J_{HNH\alpha}$, the uncertainty in populations of positive ϕ conformers can be detected with uncertainties of only a few percent. Positive ϕ populations determined in this manner for α-synuclein agree well with those from the recent MERA analysis[14] which used far more experimental data, but not $^3J_{C'H\alpha}$. Results obtained for α-synuclein are also in very close agreement with positive ϕ propensities seen in a recently introduced large coil library, extracted from non-H-bonded residues in the PDB, suggesting that the new method is not only precise but also accurate.

As previously shown for the combination of $^3J_{HNH\alpha}$ and $^3J_{C'C'}$, the newly measured $^3J_{C'H\alpha}$ and $^3J_{HNC'}$ values can be combined with either $^3J_{HNH\alpha}$ or $^3J_{C'C'}$ to extract the average of the negative ϕ angles sampled by a given IDP or IDR residue, and the standard deviation relative to this average value. Comparison of $\langle\phi^-\rangle$ and σ values derived from $^3J_{C'C'}$ and $^3J_{HNC'}$ agree very well with the earlier analysis and confirm that such an analysis is quite robust.

It has been hypothesized, mostly based on molecular dynamics simulations, that formation of α-pleated sheet

 This journal is © the Owner Societies 2016

Published on 21 September 2015. Downloaded by NATIONAL INSTITUTES OF HEALTH on 26/10/2016 17:12:05.

structures may be a common conformational transition intermediate on the pathway to amyloid formation.[6] The requisite Ramachandran map flips from $\alpha_R\alpha_L$ to $\beta\beta$ for pairs of adjacent residues is commonly observed when comparing X-ray crystal structures of closely related proteins,[7] providing some experimental support that such transitions are energetically feasible. α-Synuclein clearly falls in the category of amyloidogenic proteins, but its intrinsic disorder previously made it difficult to assess the presence of elevated transient population of α-pleated sheet backbone conformations. Analysis of our $^3J_{C'H\alpha}$ data finds no evidence for elevated population of the positive ϕ angles associated with α-pleated sheet, and propensities of its residues to adopt positive ϕ angles closely correlate with the residue-type specific propensities seen in the coil library. We note, however, that this observation does not preclude the formation of α-pleated sheet structures as an intermediate, but simply that if present the population of such structures must be very low.

For folded proteins, factors other than the intervening torsion angle ϕ are found to have a non-negligible impact on $^3J_{C'H\alpha}$. Although this limits the accuracy at which ϕ angles can be extracted from $^3J_{C'H\alpha}$, it also offers new opportunities to gain access to these additional factors, which include deviations from idealized geometry, H-bonding, and χ_1 angles. We are currently exploring the feasibility of developing quantitative relations that extend the Karplus equations to take these additional factors into account.

Acknowledgements

We thank F. Li for providing GB3 and ubiquitin samples, and J. L. Baber, Y. Shen, and A. Grishaev for technical support. NUS reconstruction utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, Md. (http://biowulf.nih.gov). This work was supported by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases and by the Intramural Antiviral Target Program of the Office of the Director, NIH. J.H.L. is the recipient of a KVSTA Fellowship.

References

1 C. Ramakrishnan, V. S. Dani and T. Ramasarma, *Protein Eng.*, 2002, **15**, 783–798.
2 C. E. Fitzgerald, S. B. Patel, J. W. Becker, P. M. Cameron, D. Zaller, V. B. Pikounis, S. J. O'Keefe and G. Scapin, *Nat. Struct. Biol.*, 2003, **10**, 764–769.
3 V. S. Honndorf, N. Coudevylle, S. Laufer, S. Becker and C. Griesinger, *Angew. Chem., Int. Ed.*, 2008, **47**, 3548–3551.
4 A. R. Fadel, D. Q. Jin, G. T. Montelione and R. M. Levy, *J. Biomol. NMR*, 1995, **6**, 221–226.
5 S. Piana, K. Lindorff-Larsen and D. E. Shaw, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 5915–5920.
6 R. S. Armen, M. L. DeMarco, D. O. V. Alonso and V. Daggett, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 11622–11627.
7 E. J. Milner-White, J. D. Watson, G. Qi and S. Hayward, *Structure*, 2006, **14**, 1369–1376.
8 J. A. Marsh, C. Neale, F. E. Jack, W.-Y. Choy, A. Y. Lee, K. A. Crowhurst and J. D. Forman-Kay, *J. Mol. Biol.*, 2007, **367**, 1494–1510.
9 J. Kragelj, V. Ozenne, M. Blackledge and M. R. Jensen, *ChemPhysChem*, 2013, **14**, 3034–3045.
10 M. R. Jensen, L. Salmon, G. Nodet and M. Blackledge, *J. Am. Chem. Soc.*, 2010, **132**, 1270–1272.
11 L. Salmon, G. Bouvignies, P. Markwick, N. Lakomek, S. Showalter, D. W. Li, K. Walter, C. Griesinger, R. Bruschweiler and M. Blackledge, *Angew. Chem., Int. Ed.*, 2009, **48**, 4154–4157.
12 M. Krzeminski, J. A. Marsh, C. Neale, W.-Y. Choy and J. D. Forman-Kay, *Bioinformatics*, 2013, **29**, 398–399.
13 N. C. Fitzkee, P. J. Fleming and G. D. Rose, *Proteins*, 2005, **58**, 852–854.
14 A. B. Mantsyzov, Y. Shen, J. H. Lee, G. Hummer and A. Bax, *J. Biomol. NMR*, 2015, **63**, 85–95.
15 G. W. Vuister, F. Delaglio and A. Bax, *J. Am. Chem. Soc.*, 1992, **114**, 9674–9675.
16 A. B. Mantsyzov, A. S. Maltsev, J. Ying, Y. Shen, G. Hummer and A. Bax, *Protein Sci.*, 2014, **23**, 1275–1290.
17 A. C. Wang and A. Bax, *J. Am. Chem. Soc.*, 1996, **118**, 2483–2494.
18 D. Rovnyak, D. P. Frueh, M. Sastry, Z. Y. J. Sun, A. S. Stern, J. C. Hoch and G. Wagner, *J. Magn. Reson.*, 2004, **170**, 15–21.
19 V. Y. Orekhov, P. V. Dubovskii, H. Yamada, K. Akasaka and A. S. Arseniev, *J. Biomol. NMR*, 2000, **17**, 257–263.
20 W. Bermel, I. Bertini, I. C. Felli, L. Gonnelli, W. Kozminski, A. Piai, R. Pierattelli and J. Stanek, *J. Biomol. NMR*, 2012, **53**, 293–301.
21 M. Johnson, A. T. Coulton, M. A. Geeves and D. P. Mulvihill, *PLoS One*, 2010, **5**, e15801.
22 A. S. Maltsev, J. F. Ying and A. Bax, *Biochemistry*, 2012, **51**, 5004–5013.
23 F. Li, J. H. Lee, A. Grishaev, J. Ying and A. Bax, *ChemPhysChem*, 2015, **16**, 572–578.
24 F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer and A. Bax, *J. Biomol. NMR*, 1995, **6**, 277–293.
25 T. D. Goddard and D. G. Kneller, *University of California*, San Francisco, 2008.
26 C. Griesinger, O. W. Sorensen and R. R. Ernst, *J. Am. Chem. Soc.*, 1985, **107**, 6394–6396.
27 C. Griesinger, O. W. Sørensen and R. R. Ernst, *J. Magn. Reson.*, 1987, **75**, 474–492.
28 G. T. Montelione and G. Wagner, *J. Am. Chem. Soc.*, 1989, **111**, 5474–5475.
29 H. B. Olsen, S. Ludvigsen and O. W. Sørensen, *J. Magn. Reson., Ser. A*, 1993, **105**, 321–322.
30 R. Weisemann, H. Rüterjans, H. Schwalbe, J. Schleucher, W. Bermel and C. Griesinger, *J. Biomol. NMR*, 1994, **4**, 231–240.
31 A. Meissner, T. Schulte-Herbruggen and O. W. Sorensen, *J. Am. Chem. Soc.*, 1998, **120**, 3803–3804.
32 A. C. Wang and A. Bax, *J. Am. Chem. Soc.*, 1995, **117**, 1810–1813.

Published on 21 September 2015. Downloaded by NATIONAL INSTITUTES OF HEALTH on 26/10/2016 17:12:05.

33 T. J. Norwood and K. Jones, *J. Magn. Reson., Ser. A*, 1993, **104**, 106–110.
34 A. Hammarström and G. Otting, *J. Am. Chem. Soc.*, 1994, **116**, 8847–8848.
35 J. Ying, J. Roche and A. Bax, *J. Magn. Reson.*, 2014, **241**, 97–102.
36 L. E. Kay, M. Ikura and A. Bax, *J. Magn. Reson.*, 1991, **91**, 84–92.
37 L. E. Kay, P. Keifer and T. Saarinen, *J. Am. Chem. Soc.*, 1992, **114**, 10663–10665.
38 A. J. Shaka, J. Keler and R. Freeman, *J. Magn. Reson.*, 1983, **53**, 313–340.
39 M. Ikura, M. Krinks, D. A. Torchia and A. Bax, *FEBS Lett.*, 1990, **266**, 155–158.
40 M. S. Silver, R. I. Joseph and D. I. Hoult, *Nature*, 1984, **310**, 681–683.
41 N. Juranic, P. K. Ilich and S. Macura, *J. Am. Chem. Soc.*, 1995, **117**, 405–410.
42 J. M. Schmidt, S. Zhou, M. L. Rowe, M. J. Howard, R. A. Williamson and F. Loehr, *Proteins: Struct., Funct., Bioinf.*, 2011, **79**, 428–443.
43 J. M. Schmidt, Y. Hua and F. Loehr, *Proteins: Struct., Funct., Bioinf.*, 2010, **78**, 1544–1562.
44 J. M. Schmidt, M. J. Howard, M. Maestre-Martinez, C. S. Perez and F. Loehr, *Magn. Reson. Chem.*, 2009, **47**, 16–30.
45 G. Kontaxis, G. M. Clore and A. Bax, *J. Magn. Reson.*, 2000, **143**, 184–196.
46 K. Kazimierczuk, J. Stanek, A. Zawadzka-Kazimierczuk and W. Kozminski, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2010, **57**, 420–434.
47 S. G. Hyberts, A. G. Milbradt, A. B. Wagner, H. Arthanari and G. Wagner, *J. Biomol. NMR*, 2012, **52**, 315–327.
48 B. E. Coggins, J. W. Werner-Allen, A. Yan and P. Zhou, *J. Am. Chem. Soc.*, 2012, **134**, 18619–18630.
49 G. A. Zhu and A. Bax, *J. Magn. Reson.*, 1992, **98**, 192–199.
50 A. S. Maltsev, A. Grishaev, J. Roche, M. Zasloff and A. Bax, *J. Am. Chem. Soc.*, 2014, **136**, 3752–3755.
51 J. H. Lee, F. Li, A. Grishaev and A. Bax, *J. Am. Chem. Soc.*, 2015, **137**, 1432–1435.
52 R. Bruschweiler and D. A. Case, *J. Am. Chem. Soc.*, 1994, **116**, 11199–11200.
53 L. Emsley and G. Bodenhausen, *Chem. Phys. Lett.*, 1990, **165**, 469–476.
54 H. Geen and R. Freeman, *J. Magn. Reson.*, 1991, **93**, 93–141.

 This journal is © the Owner Societies 2016