Published in final edited form as:
J Am Chem Soc. 2015 February 4; 137(4): 1432–1435. doi:10.1021/ja512593s.

Quantitative residue-specific protein backbone torsion angle dynamics from concerted measurement of ^3J couplings

Jung Ho Lee, **Fang Li**, **Alexander Grishaev**, and **Ad Bax**
Laboratory of Chemical Physics, NIDDK, National Institutes of Health, Bethesda, MD, 20892-0520, USA

Ad Bax: bax@nih.gov

Abstract

Three-bond $^3J_{C'C'}$ and $^3J_{HNH\alpha}$ couplings in peptides and proteins both are functions of the intervening backbone torsion angle φ. In well ordered regions, $^3J_{HNH\alpha}$ is tightly correlated with $^3J_{C'C'}$, but the presence of large φ angle fluctuations differentially affects the two types of couplings. Assuming the φ angles to follow a Gaussian distribution, the width of this distribution can be extracted from $^3J_{C'C'}$ and $^3J_{HNH\alpha}$, as demonstrated for the folded proteins ubiquitin and GB3. In intrinsically disordered proteins, slow transverse relaxation permits measurement of $^3J_{C'C'}$ and $^3J_{HNH}$ couplings at very high precision, and impact of factors other than the intervening torsion angle on ^3J will be minimal, making these couplings exceptionally valuable structural reporters. Analysis of α-synuclein yields rather homogeneous widths of 69±6° for the φ angle distributions, and $^3J_{C'C'}$ values that agree well with those of a recent maximum entropy analysis of chemical shifts, J couplings, and ^{1}H-^{1}H NOEs. Data are consistent with a modest (≤ 30%) population of the polyproline II region.

Solution NMR relaxation rates have long been used to study the amplitudes and time scales of backbone and sidechain dynamics in folded proteins.[1–5] Whereas longitudinal and transverse relaxation times together with heteronuclear NOE data can be used to probe both the amplitudes and rates of bond vector fluctuations, only motions faster than the rotational correlation time can be derived at good accuracy.[2] Time scales of motions much slower than the molecular tumbling time can be derived from relaxation dispersion measurements,[6,7] but usually the amplitude of these motions cannot be extracted from such data. Analysis of residual dipolar couplings (RDCs) acquired under three or more orthogonal alignment conditions can provide a quantitative measure for the width of the orientational distributions of any given bond vector, expressed as an order parameter,[8–11] and thereby complement the relaxation dispersion data. However, it often can be challenging to generate the requisite orthogonal alignments and to obtain the high RDC accuracy that is required when interpreting these in terms of dynamics. For example, substantial divergence in the

Corresponding Author: bax@nih.gov.

Notes
The authors declare no competing financial interests.

Supporting Information
Experimental procedures and additional data. This material is available free of charge via the Internet at http://pubs.acs.org.

magnitude of order parameters extracted from RDCs can be seen in various studies of ubiquitin, which has served as a model system for such analyses.[8,10,12–16]

Three-bond J couplings are related to the intervening dihedral angle, θ, by the classic Karplus equation:[17]

$$^{3}\mathrm{J}=A\cos^{2}\theta+B\cos\theta+C \quad (1)$$

where the "Karplus coefficients", *A*, *B*, and *C* depend on the nuclei involved and on the electronegativity of substituents, but are also impacted by intervening valence bond angles, and bond lengths.[18,19] For sidechains in proteins, where for many residue types separate $^{3}J_{H\alpha H\beta 2}$ and $^{3}J_{H\alpha H\beta 3}$ can be measured, the availability of two couplings together with the non-linear character of eq 1 allows χ_1 analysis in terms of rotamer distributions, thereby providing access to χ_1 dynamics integrated over the entire NMR time scale, from ps to ms.[20–22] Even in the absence of rotameric jumps, eq 1 is sensitive to θ fluctuations: Assuming a Gaussian θ distribution with standard deviation σ (in units of radians), the coefficients of eq 1 (for σ < ~1) to a very good approximation can be rewritten as: [23]

$$A'=A\exp(-2\sigma^{2}),\ B'=B\exp(-\sigma^{2}/2),\ C'=C+A[1-\exp(-2\sigma^{2})]/2 \quad (2)$$

Note that the empirically parameterized Karplus curves for $^{3}J_{C'C'}$ and $^{3}J_{HNH\alpha}$, using experimental values measured for GB3 and ubiquitin, respectively, and φ angles derived from the RDC-refined NMR structures, already include the effects of φ angle fluctuations,[24] i.e., these Karplus parameters correspond to *A*′, *B*′, and *C*′. A relatively short (1.5 ns) molecular dynamics trajectory of myoglobin pointed to rather homogeneous σ values of ~0.15 for residues engaged in secondary structure.[23] Similarly, a much longer room temperature molecular dynamics trajectory of ubiquitin[25] showed a narrow σ distribution (0.23±0.05) for the set of well-ordered residues, previously selected for calibration of the Karplus curve. Using this σ=0.23 value, inversion of eq 2 yields slightly modified Karplus coefficients for the static case (Table 1), which then constitute the starting point for evaluating $^{3}J_{C'C'}$ and $^{3}J_{HNH\alpha}$ in terms of dynamics. Note that for this modest σ amplitude, the best-fitted and dynamics-corrected curves are very close (Supporting Information (SI) Fig. S1). With a root-mean-square deviation (rmsd) of 0.12 Hz, the experimental $^{3}J_{C'C'}$ values follow these curves very closely, but $^{3}J_{HNH\alpha}$ values exhibit larger deviations (rmsd 0.65 Hz) when plotted against the backbone torsion angle φ (SI, Fig. S1B). The principal cause of this larger rmsd lies in deviations from idealized peptide plane and tetrahedral C^{α} geometries, which impact the Karplus equations when written in terms of φ, rather than the intervening H-N-C^{α}-H^{α} dihedral angle, θ. When N-H and C^{α}-H^{α} vector orientations are known accurately from residual dipolar coupling (RDC) measurements, the $^{3}J_{HNH\alpha}$ Karplus curve fit improves nearly two-fold (compare values with c and d superscripts in Table 1; SI Fig. S1C).

It is important to note that when only $^{3}J_{C'C'}$ or $^{3}J_{HNH\alpha}$ is available to define φ, the effect of φ angle fluctuations cannot be separated from a change in <φ>. However, with both couplings measured, in the absence of motion the values must fall on the outer purple

contour in Fig. 1 (σ = 0°), whereas for increasing values of σ the range of accessible $^3J_{C'C'}$ and $^3J_{HNH\alpha}$ values progressively decreases.

The combined impact of measurement error and of factors other than φ on the predicted ^3J value can be estimated from the observation that the root mean square difference (rmsd) between observed and best-fitted couplings is ~0.65 Hz for $^3J_{HNH\alpha}$ and ~0.12 Hz for $^3J_{C'C'}$. With the slight exception of Val-17 in ubiquitin (Fig. 1A), all experimental values then fall in the range accessible to such pairs of couplings, i.e. on or within the outer purple contour in Fig. 1, with the color of the contours (purple to red) marking increasing amplitudes of σ. Note that from such a graphic analysis, $^3J_{C'C'}$ and $^3J_{HNH\alpha}$ define both σ and <φ>.

First focusing on <φ>, values extracted from Fig. 1A agree closely with the newly refined static NMR structure (PDB entry 2MJB [16]; rmsd 8.2° for residues 2–74, incl. the residues identified as dynamic), but also when compared to dynamic ensembles previously derived from RDCs (PDB entries 2K39, 2KOX; rmsd ~6°)[10,15] (SI Fig. S2). Using a previously identified[16] ensemble of high resolution ubiquitin X-ray structures, even slightly better agreement is obtained for <φ> (rmsd 5.0°; Fig. 2A), whereas somewhat lower agreement (rmsd ~9°) is observed for ensembles or trajectories generated without RDCs [25,27] (SI Fig. S2D, E).

Interestingly, when extracting σ from Fig. 1A, the opposite trend is observed: The ensemble of X-ray structures systematically yields σ values that are too low (SI Fig. S3) and the highest rmsd (10.6°), together with a near-zero Pearson's correlation coefficient (R_p = 0.09); NMR-derived ensembles yield somewhat better agreement (rmsd 8–9°; R_p = 0.3–0.4); and closest agreement is observed for ensembles derived from molecular dynamics (rmsd 7.0°; R_p = 0.58; Fig. 2B). It is important to note, however, that the impact of these motions on the ^3J couplings scales with σ^2 (c.f. eq 2), which means that for small amplitude motions the σ value extracted from Fig. 1 has a large uncertainty. Inversely, for large amplitude motions, a good quantitative estimate of σ can be made. For example, residues R72–R74 in ubiquitin, as well as L12 in GB3, all highly disordered on the basis of ^{15}N relaxation studies, show considerably larger σ values than residues in secondary structure elements and good agreement with molecular dynamics results. A full set of σ and <φ> values, extracted from the $^3J_{HNH\alpha}$ and $^3J_{C'C'}$ data, is listed in SI Table S1.

Intrinsically disordered proteins typically have very favorable NMR relaxation properties, allowing measurement of $^3J_{C'C'}$ and $^3J_{HNH\alpha}$ couplings at very high precision.[24] Moreover, the average effect of factors such as H-bonding and distortion of valence angles on the ^3J couplings will be far more uniform than in well-ordered proteins. Therefore, the combination of $^3J_{C'C'}$ and $^3J_{HNH\alpha}$ couplings is particularly well suited for defining the residue-specific <φ> and σ values. However, transient switches to the α_L region of Ramachandran space (φ > 0°) result in a very large deviation from the approximation of a Gaussian distribution on which eq 2 is based. Fortunately, $^3J_{C'C'}$ and $^3J_{HNH\alpha}$ couplings for the α_L region fall close to typical random coil values (Fig. 3), and small populations of α_L therefore do not strongly impact $^3J_{C'C'}$ or $^3J_{HNH\alpha}$. Thus, if the α_L population is small (≤ 15%), like we found to be the case for all non-Gly residues in α-synuclein,[28] the contribution from the α_L population to the ^3J couplings may simply be ignored, and σ then

reports on the standard deviation of φ angle fluctuations in the φ <0° region. Indeed, when predicting $^{3}J_{C'C'}$ and $^{3}J_{HNH\alpha}$ couplings for the previously derived φ/ψ ensembles of α-synuclein, the effect of ignoring the positive-φ conformers is minimal (rmsd of 0.03 Hz for $^{3}J_{HNH\alpha}$ and 0.02 Hz for $^{3}J_{C'C'}$; SI Fig. S4).

With an rmsd of only 2.7°, values for σ derived from the graphic analysis of Fig. 3 agree well with the φ/ψ ensembles derived previously from ^{1}H-^{1}H NOEs, $^{1}J_{C\alpha H\alpha}$, $^{1}J_{NC\alpha}$, $^{2}J_{NC\alpha}$, $^{3}J_{HNH\alpha}$, and $^{13}C^{\alpha}$, $^{13}C'$ and ^{15}N chemical shifts (Fig. 4B). This close agreement may be attributed to the unusually large number of restraints (10) per residue that was previously used to derive this ensemble. The α-synuclein σ values are remarkably homogeneous, with the smallest values observed for Leu residues (σ=24.5±0.6°; N=3), and the largest values for Ala and Ser residues (σ=31.3±1.7°; N=14), or full width at half maximum (FWHM) values of 58 and 74°, respectively. When comparing the average φ angles derived from the graphic analysis with those of the prior maximum entropy (ME) analysis (Fig. 4A) the correlation (R_P = 0.93) between <φ> derived from $^{3}J_{C'C'}$ and $^{3}J_{HNH\alpha}$ and <φ> angles from the ME ensemble (which used and $^{3}J_{HNH\alpha}$ as one of 10 restraints per residue) is also good and considerably higher than between $^{3}J_{C'C'}$ and $^{3}J_{HNH\alpha}$ values themselves (R_P = 0.74), but a small systematic difference is also observed. This systematic difference can be attributed to the assumption of a Gaussian φ angle distribution, implicit in eq 2 and in the graphic analysis, whereas the true distribution is strongly skewed. Indeed, for Ala residues, which exhibited a highly asymmetric distribution in the prior analysis (SI Fig. S5), the difference in <φ> obtained from the two methods is largest (Fig. 4A). When comparing the $<^{3}J_{HNH\alpha}>$ values computed for the previously derived ensembles with couplings predicted from eqs 1 and 2, using σ and <φ> values derived from the ensembles (while ignoring contributions to the ^{3}J couplings from these positive-φ conformers), the non-Gaussian distribution of the Ala φ angles results in a similar, modest but systematic difference (SI Fig. S6). By contrast, for β-branched residues, which show a nearly Gaussian φ distribution, extracted <φ> values closely agree with the prior results (Fig. 4A). Similarly, for ubiquitin, where the angular φ ranges are much narrower, the Gaussian approximation is also perfectly valid (SI Fig. S7).

On average, Ala residues have the least negative <φ> value. Some studies have concluded that Ala residues in random coil peptides favor the polyproline II (PPII) region of Ramachandran space, centered at (φ,ψ) ≈ (−75°, 160°), which at first sight appears consistent with the <φ> value derived from the graphic analysis. However, a high population of PPII disagrees with the observation that Ala residues also show among the largest σ values. In contrast to the very small $^{3}J_{C'C'}$ coupling of 0.25 Hz, previously reported for the center residue in the Ala_3 tripeptide,[29] which drives the conformation towards the PPII region, $^{3}J_{C'C'}$ values for Ala residues in α-synuclein are all considerably larger (0.73±0.03 Hz; N=11). This includes A18, which is flanked by two Ala residues, and shows a $^{3}J_{HNH\alpha}$ coupling of 5.22 Hz, close to the 5.68 Hz observed for the center residue in Ala_3. The very small $^{3}J_{C'C'}$ value observed in the previous study contributed to a high fitted population of PPII for this residue (92%), much higher than seen in molecular dynamics simulations,[29] and was excluded in recent work that relied on NMR observables to calibrate molecular dynamics force fields.[30] Indeed, inspection of the residue-specific (φ,ψ) distributions

derived for the highly disordered α-synuclein protein finds a modest PPII population for Ala residues (~30%), only slightly higher than for other residue types.[28] The β-branched residues (Val, Ile, and Thr) show the most negative <φ> values (Fig. 3, Fig. 4A), consistent with their propensity to be found in β-sheet secondary structure, but their $\sigma \approx 30°$ again point to a highly dynamic distribution.

Our results indicate that $^3J_{C'C'}$ and $^3J_{HNH\alpha}$ together not only provide remarkably accurate information on the average value of φ backbone torsion angles, but also yield straightforward access to the amplitude of motions. The approach will be particularly useful for highly dynamic and intrinsically disordered proteins, where the dynamic angular distributions are difficult to assess by other experimental methods. Importantly, for such dynamic systems the $^3J_{C'C'}$ and $^3J_{HNH\alpha}$ measurement will be most accurate, whereas the effect of other factors impacting these couplings, such as H-bonding will be more uniform than in highly ordered proteins. The large set of highly precise $^3J_{C'C'}$ and $^3J_{HNH\alpha}$ values presented here (SI Table S2) may also serve as useful benchmarks when evaluating force fields used for molecular dynamics studies of disordered linear peptides.[29,30] At least in principle, other φ dependent couplings for which the Karplus curves are 60° phase-shifted relative to $^3J_{C'C'}$ and $^3J_{HNH\alpha}$ (i.e. $^3J_{HNC\beta}$, $^3J_{C'C\beta}$) could be added to the analysis. However, they have larger intrinsic scatter relative to their Karplus curves[31] and therefore may prove to be less restraining.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank J. Ying, J.L. Baber, and A. Mantsyzov for technical support. 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. Lee is the recipient of the KVSTA Fellowship and F. Li acknowledges financial support from the China Scholarship Council.

References

1. Nelson DJ, Opella SJ, Jardetzky O. Biochemistry. 1976; 15:5552–5560. [PubMed: 999828]
2. Lipari G, Szabo A. J Am Chem Soc. 1982; 104:4546–4559.
3. Nirmala NR, Wagner G. J Am Chem Soc. 1988; 110:7557–7558.
4. Kay LE, Torchia DA, Bax A. Biochemistry. 1989; 28:8972–8979. [PubMed: 2690953]
5. Dellwo MJ, Wand AJ. J Am Chem Soc. 1989; 111:4571–4578.
6. Palmer AG. Chem Rev. 2004; 104:3623–3640. [PubMed: 15303831]
7. Mittermaier A, Kay LE. Science. 2006; 312:224–228. [PubMed: 16614210]
8. Peti W, Meiler J, Bruschweiler R, Griesinger C. J Am Chem Soc. 2002; 124:5822–5833. [PubMed: 12010057]
9. Tolman JR, Al-Hashimi HM, Kay LE, Prestegard JH. J Am Chem Soc. 2001; 123:1416–1424. [PubMed: 11456715]
10. Lange OF, Lakomek NA, Fares C, Schroder GF, Walter KFA, Becker S, Meiler J, Grubmuller H, Griesinger C, de Groot BL. Science. 2008; 320:1471–1475. [PubMed: 18556554]
11. Tolman JR, Ruan K. Chem Rev. 2006; 106:1720–1736. [PubMed: 16683751]
12. Tolman JR. J Am Chem Soc. 2002; 124:12020–12030. [PubMed: 12358549]

13. Briggman KB, Tolman JR. J Am Chem Soc. 2003; 125:10164–10165. [PubMed: 12926926]
14. Lakomek NA, Carlomagno T, Becker S, Griesinger C, Meiler J. J Biomol NMR. 2006; 34:101–115. [PubMed: 16518697]
15. Salmon L, Bouvignies G, Markwick P, Lakomek N, Showalter S, Li DW, Walter K, Griesinger C, Bruschweiler R, Blackledge M. Angew Chem Int Ed. 2009; 48:4154–4157.
16. Maltsev AS, Grishaev A, Roche J, Zasloff M, Bax A. J Am Chem Soc. 2014; 136:3752–3755. [PubMed: 24568736]
17. Karplus M. J Chem Phys. 1959; 30:11–15.
18. Karplus M. J Am Chem Soc. 1963; 85:2870–2871.
19. Haasnoot CAG, Deleeuw F, Altona C. Tetrahedron. 1980; 36:2783–2792.
20. Pachler KGR. Spectrochimica Acta. 1964; 20:581–587.
21. Dzakula Z, Westler WM, Edison AS, Markley JL. J Am Chem Soc. 1992; 114:6195–6199.
22. Nagayama K, Wuthrich K. Eur J Biochem. 1981; 115:653–657. [PubMed: 6165586]
23. Bruschweiler R, Case DA. J Am Chem Soc. 1994; 116:11199–11200.
24. Li, F.; Lee, JH.; Grishaev, A.; Ying, J.; Bax, A. ChemPhysChem. 2014. http://dx.doi.org/10.1002/cphc.201402704
25. Piana S, Lindorff-Larsen K, Shaw DE. Proc Natl Acad Sci U S A. 2013; 110:5915–5920. [PubMed: 23503848]
26. Vogeli B, Ying JF, Grishaev A, Bax A. J Am Chem Soc. 2007; 129:9377–9385. [PubMed: 17608477]
27. Lindorff-Larsen K, Best RB, DePristo MA, Dobson CM, Vendruscolo M. Nature. 2005; 433:128–132. [PubMed: 15650731]
28. Mantsyzov AB, Maltsev AS, Ying J, Shen Y, Hummer G, Bax A. Protein Sci. 2014; 23:1275–90. [PubMed: 24976112]
29. Graf J, Nguyen PH, Stock G, Schwalbe H. J Am Chem Soc. 2007; 129:1179–1189. [PubMed: 17263399]
30. Best RB, Zheng W, Mittal J. J Chem Theory Comput. 2014; 10:5113–5124. [PubMed: 25400522]
31. Schmidt JM, Blumel M, Lohr F, Ruterjans H. J Biomol NMR. 1999; 14:1–12. [PubMed: 21136331]

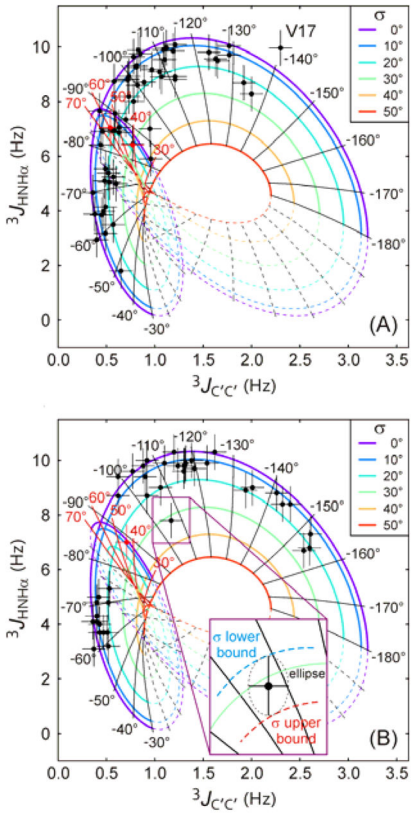

Figure 1.

Plots of $^3J_{HNH\alpha}$ versus $^3J_{C'C'}$ for (A) ubiquitin and (B) GB3. Colored contours correspond to the correlation between their respective Karplus curves for σ values of 0–50°, with contours corresponding to φ angles outside of the most populated region of the Ramachandran map being dashed. Radial "spokes" correspond to the <φ> angles marked in the figure, with red spokes corresponding to the α_L region. Red data points correspond to residues with φ > 0 in the reference structure. The inset in (B) shows how the asymmetric error bar for σ is determined. The minor semi-axis of the ellipse corresponds to the $^3J_{C'C'}$ rmsd (0.12 Hz) and major semi-axis to the $^3J_{HNH\alpha}$ rmsd (0.65 Hz). Note that σ relates to the full width at half maximum (FWHM) of the Gaussian distribution according to: FWHM = 2.35σ.

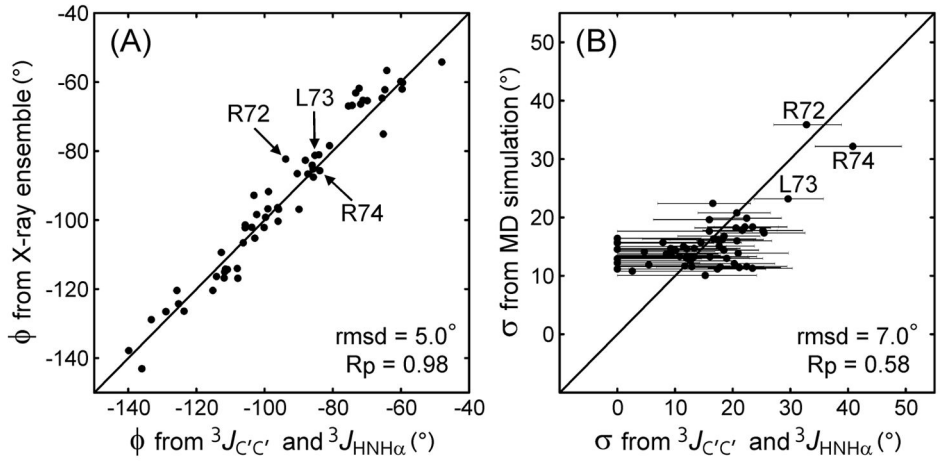

Figure 2.

Comparison of (A) <φ> values for ubiquitin extracted from Fig. 1A with averaged <φ> values from 15 high resolution (≤ 1.8 Å) X-ray structures, listed in the SI of Maltsev et al. [16] and (B) σ values extracted from Fig. 1A with those obtained from the 1-ms molecular dynamics simulation of Piana et al. (2013). Only ensemble conformers with φ < 0° were included in calculating <φ> and σ. Correlations for distributions of other ensembles are presented in SI Figs. S2 and S3.

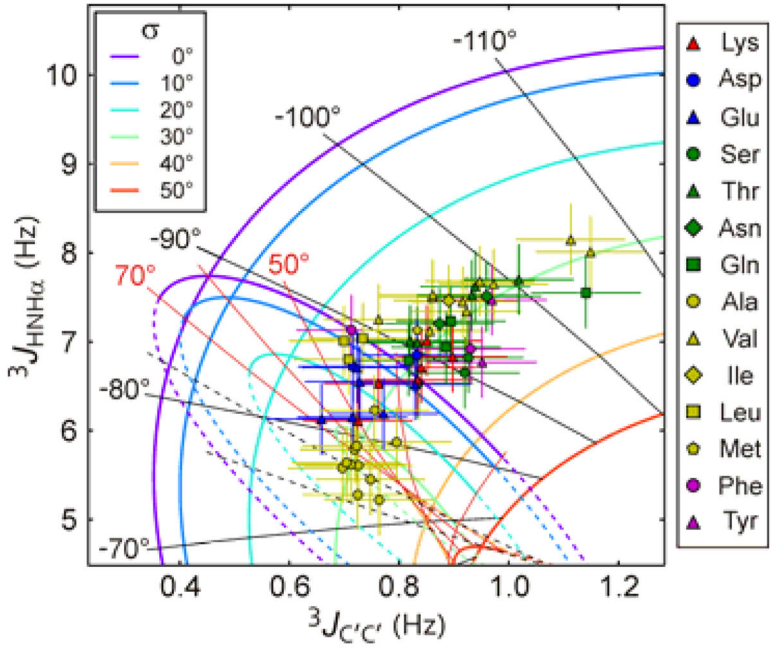

Figure 3.

Plots of $^3J_{HNH\alpha}$ versus $^3J_{C'C'}$ for α-synuclein (0.6 mM; pH 6.0; 50 mM NaCl, 288 K). Only $^3J_{C'C'}$ values for which both diagonal ^{15}N-^1H correlations and both crosspeaks were well resolved in the 3D HN(COCO)NH spectrum [24] are used for the figure. Radial "spokes" correspond to the marked <φ> values, with positive values marked in red. The list of $^3J_{HNH\alpha}$ and $^3J_{C'C'}$ values used for generating the plot is included as SI Table S2.

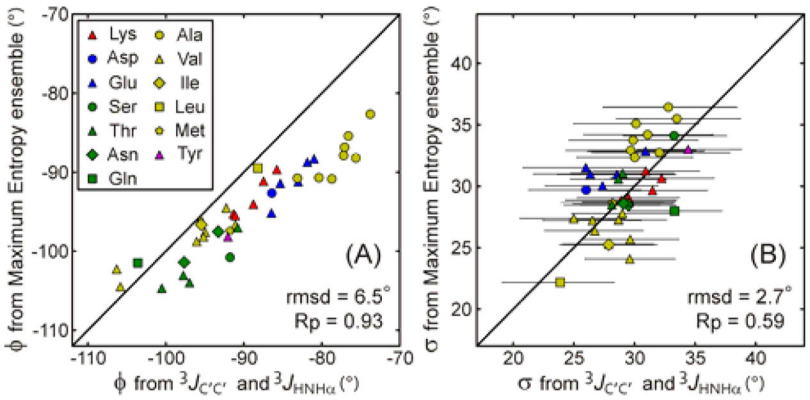

Figure 4.

Values for (A) <φ> and (B) σ derived from the graphic analysis of Fig. 3 versus the values extracted from the φ/ψ ensembles derived previously from ^{1}H-^{1}H NOEs, $^1J_{C\alpha\,H\alpha}$, $^1J_{NC\alpha}$, $^2J_{NC\alpha}$, $^3J_{HNH\alpha}$, and $^{13}C^{\alpha}$, $^{13}C'$ and ^{15}N chemical shifts.[26] Only ensemble conformers with φ < 0° were included in calculating <φ> and σ. The error bars in (B) are based on an uncertainty of 0.4 Hz in $^3J_{HNH\alpha}$ and 0.1 Hz in $^3J_{C'C'}$, which likely overestimates their actual uncertainties, because impact of variations in H-bonding and valence angles will be highly averaged in IDPs.

Table 1

Dynamics-corrected Karplus equation coefficients for $^{3}J_{HNH\alpha}$ and $^{3}J_{C'C'}$.[a]

	A (Hz)	*B* (Hz)	*C* (Hz)	Rmsd UBQ[b]	Rmsd UBQ[b]	Rmsd GB3[b]	Rmsd GB3[b]
$^{3}J_{HNH\alpha}$	8.83	−1.29	0.20	0.43[c]	0.61[d]	0.34[c]	0.69[d]
$^{3}J_{C'C'}$	1.78	−0.95	0.46	0.12		0.12	

[a] Karplus coefficients back-calculated from the best-fitted A', B' and C' coefficients, using eq 2, and assuming σ=0.226 to factor out the effect of dynamics. These are the coefficients to be used when extracting dynamics from $^{3}J_{HNH\alpha}$ and $^{3}J_{C'C'}$. A', B' and C' values correspond to the Karplus coefficients of Vogeli[26] (for $^{3}J_{HNH\alpha}$) and Li[24] (for $^{3}J_{C'C'}$).

[b] Rmsd when using the eq 2 correction to *A*, *B*, and *C*, with σ=0.226, excl. residues T7-K11, D32-G35, A46, G47, D52, and V70-G76 for ubiquitin, and residues L12, D40, and G41 for GB3.

[c] Using the RDC-derived H-N-C^{α}-H^{α} dihedral angle.

[d] Using $\theta = \varphi - 60°$.