Published in final edited form as:
*Chemphyschem*. 2015 February 23; 16(3): 572–578. doi:10.1002/cphc.201402704.

# High Accuracy of Karplus Equations for Relating Three-Bond J Couplings to Protein Backbone Torsion Angles

**Fang Li**, **Jung Ho Lee**, **Alexander Grishaev**, **Jinfa Ying**, and **Ad Bax**
Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA

## Abstract

$^3J_{C'C'}$ and $^3J_{HNH\alpha}$ couplings are related to the intervening backbone torsion angle $\varphi$ by standard Karplus equations. Although these couplings are known to be affected by parameters other than $\varphi$, including H-bonding, valence angles and residue type, experimental results and quantum calculations indicate that the impact of these latter parameters is typically very small. The solution NMR structure of protein GB3, newly refined by using extensive sets of residual dipolar couplings (RDCs), yields 50–60% better Karplus equation agreement between φ angles and experimental $^3J_{C'C'}$ and $^3J_{HNH\alpha}$ values than does the high resolution X-ray structure. In intrinsically disordered proteins, $^3J_{C'C'}$ and $^3J_{HNH\alpha}$ couplings can be measured at even higher accuracy, and the impact of factors other than the intervening torsion angle on $^3J$ will be smaller than in folded proteins, making these couplings exceptionally valuable reporters on the ensemble of $\varphi$ angles sampled by each residue.

### Keywords

backbone torsion angle; DFT calculation; GB3; IDP; Karplus equation; protein structure

## 1. Introduction

Empirically parameterized Karplus relations[1] between a wide range of different types of three-bond *J* couplings and their corresponding intervening dihedral angles continue to be of exceptional value in structural analysis of organic compounds, including peptides and proteins.[2] Shortly after the initial introduction of the Karplus relation, multiple reports highlighted the dependence of $^3J_{HH}$ couplings on electronegativity of substituents and ring size for cyclic compounds, thereby limiting the precision at which these couplings can be interpreted.[3–6] In a subsequent response, Karplus pointed out that indeed $^3J$ couplings not only depend on the intervening torsion angle, but that theory also shows the coupling to be impacted by electronegativity of substituents, H-C-C valence angles, and bond lengths, and that “the person who attempts to estimate dihedral angles to an accuracy of one or two degrees does so at his own peril”.[7] Although the effect of substituent electronegativity has been accounted for by empirical adjustments to the Karplus equation,[8] the chemistry community largely has taken Karplus’ advice to heart and does not interpret these couplings to the level of precision at which they often can be measured.

Backbone $^{3}J_{HNH\alpha}$ couplings have been widely used for conformational analysis in peptides and proteins. With the exception of Gly residues, the electronegativity of the substituents is very similar for all residue types, suggesting that quantitative analysis of these parameters in terms of the intervening torsion angle φ may be quite robust. However, the observation of a relatively large value (0.85 Hz) for the root-mean-square difference (rmsd) between observed $^{3}J_{HNH\alpha}$ values and those predicted by a 1.4-Å X-ray structure of the small protein BPTI again pointed to the presence of alternate factors impacting these $^{3}J_{HNH\alpha}$ values.[9] Indeed, deviations from ideal covalent geometry in the protein backbone are well documented,[10] and variations in H-bonding are expected to impact the $H^{N}$ σ orbitals and thereby the $^{3}J_{HNH\alpha}$ value. Moreover, as pointed out by Brueschweiler and Case,[11, 12] non-uniformity in the amplitude of backbone dynamics along the protein backbone will differentially skew these experimental parameters towards a value of *ca* 5 Hz, again impacting the relation between $^{3}J_{HNH\alpha}$ and the time-averaged intervening dihedral angle. On the other hand, refinement of the 1.1-Å X-ray structure of the protein GB3 yielded a considerable drop from 0.76 to 0.43 Hz between measured $^{3}J_{HNH\alpha}$ values and those predicted by an optimized Karplus equation.[13] For the extensively studied protein ubiquitin, concerted analysis of $^{3}J_{HNH\alpha}$, $^{3}J_{HNC\beta}$, $^{3}J_{HNC'}$, $^{3}J_{C'H\alpha}$, $^{3}J_{C'C\beta}$, and $^{3}J_{C'C'}$, which all report on the backbone torsion angle φ, also indicated that uncertainty in the coordinates obtained from a 1.8-Å X-ray structure of this protein dominated the 0.73 Hz rmsd between observed and predicted $^{3}J_{HNH\alpha}$ values.[14]

Here, we report results which confirm that random uncertainty in the X-ray coordinates is the prime contributor to the rmsd between observed $^{3}J_{C'C'}$ values and those predicted by the Karplus equation. For $^{3}J_{HNH\alpha}$, we find that both uncertainties in the coordinates of the C and N backbone atoms as well as deviations from idealized in-peptide-plane positions of $H^{N}$ and from idealized tetrahedral geometry at $C^{\alpha}$ significantly impact the residual in the fit. Our results imply that experimental $^{3}J_{HNH\alpha}$ and $^{3}J_{C'C'}$ values can be interpreted in angular terms at a precision that approaches that of their measurements. This finding is of particular importance when defining or validating backbone angle distributions in intrinsically disordered proteins, where both $^{3}J_{HNH\alpha}$ and $^{3}J_{C'C'}$ are readily measured at an accuracy better than 0.1 Hz.

## 2. Results and Discussion

The 1.1-Å X-ray structure of GB3,[15] previously refined by solution NMR RDCs,[16] was further refined by newly measured $^{1}D_{C\alpha C\beta}$ RDCs under three different alignment orientations, as well as previously reported $^{1}H$-$^{1}H$ RDCs,[17] and $^{1}D_{C\alpha C'}$ values measured for a number of GB3 mutants that orient differentially in a Pf1 liquid crystal suspension.[18] In order to obtain optimal cross-validation results, we found it necessary to reduce by 5-fold the very tight force constants normally used for keeping the peptide bond planar. For details, see the Methods section. The root-mean-square deviation from Ω=180° in the final structures then was 5.0°, close to the values seen in the highest resolution protein X-ray structures.[10, 19, 20] The change in backbone coordinates between the newly calculated structure and the previous RDC-refined structure is quite small (0.37 Å), but nevertheless improves the fit of the $^{3}J_{HNH\alpha}$ and in particular the $^{3}J_{C'C'}$ couplings significantly (*vide infra*).

It is well known that protein structures are subject to relatively large amplitude internal dynamics,[21–23] which impacts $^3J$ couplings and therefore their fits to Karplus curves.[11]. For deriving the most accurate average torsion angles it is advantageous to calculate the structure as a single model that best fits all input restraints simultaneously, rather than as the average over an NMR ensemble calculation.[24] We therefore have chosen the latter approach for evaluating the relation between $^3J$ couplings and backbone torsion angles, while excluding the three residues (L12, D40, and G41) that were shown to have the highest amplitude internal dynamics as judged by $^{15}$N relaxation[25] and RDC measurements.[26]

## Fit of $^3J_{HNH\alpha}$ values to Karplus equation

High precision $^3J_{HNH\alpha}$ values, measured previously[13] using a multiple-quantum method,[27] that to a good approximation suppresses effects of cross-relaxation,[28] were used in this study. A fit of these couplings to the X-ray structure (PDB entry 1IGD)[15] to which hydrogens were added with the program REDUCE,[29] yields an rmsd of 0.76 Hz. This rmsd is minimally impacted by the refinement method used, as can be seen from the values obtained for the PDB_REDO structure [30] of GB3, obtained using an optimized automated refinement protocol. This result is not surprising, considering that the backbone coordinate rmsd between 1IGD and its PDB_REDO structure is only 0.046 Å (see Table S1 of the Supporting Information). A much better fit (0.43 Hz rmsd)[17] is observed for the previously RDC-refined model (PDB entry 2OED),[16] which drops to 0.34 Hz for the newly refined model, which includes many additional RDCs plus decreased peptide bond planarity restraints. The coordinates of a GB3 structure obtained from joint X-ray/NMR refinement, using the 1IGD structure factors and the 2OED NMR (RDC) restraints [31], yields rmsd values that are intermediate between 2OED and 1IGD. The relatively high rmsd value for this structure appears to result from the weak weights used for the NMR restraints, also reflected in a very low coordinate rmsd of 0.04 Å relative to 1IGD (see Table S1 of the Supporting Information).

As can be seen in Table 1, a significant fraction of the improvement in fit quality of the newly refined NMR structure over the original X-ray structure stems from the fact that the positions of the hydrogens in X-ray structures are unknown, and are added to the structure by model building using programs such as REDUCE[29] or MOLMOL.[32] If the $H^N$-N-$C^\alpha$-$H^\alpha$ dihedral angle is derived from the newly refined structure using the common assumption $\theta = \varphi - 60°$ the fit quality drops to a level that is not much better than the original X-ray structure (Table 1). As was found previously,[14, 24] and also can be seen from the fifth column in Table 1, most of this increase in rmsd is related to the out-of-peptide-plane angle $\Delta\varphi$, but small deviations of the $C^\alpha$-$H^\alpha$ vector from its idealized orientation also contribute a non-negligible amount. Indeed, when $H^\alpha$ is positioned in the plane that bisects the N-$C^\alpha$-$C^\beta$ and N-$C^\alpha$-C′ planes, i.e. $\theta_{HNH\alpha} = (\theta_{HNC\beta} + \theta_{HNC'} + 360°)/2$, the fit to the Karplus equation is slightly worse (0.37 *vs* 0.34 Hz).

Only in exceptional cases, such as the present study of GB3, are there sufficient experimental restraints to define deviations from idealized geometry for positioning hydrogens in a protein structure from NMR data. In practice, therefore when using $^3J_{HNH\alpha}$ couplings in structure determination, the 0.69 Hz value in Table 1 represents the limit of

precision at which $^{3}J_{HNH\alpha}$ can be interpreted in structural terms, even while inherently it is considerably more accurate at describing the $H^{N}$-N-$C^{\alpha}$-$H^{\alpha}$ dihedral angle.

### Fit of $^{3}J_{C'C'}$ values to Karplus equation

Measurement of $^{3}J_{C'C'}$ in isotopically uniformly enriched proteins has been demonstrated previously,[33] and was shown to be useful for sequential assignment in intrinsically disordered proteins.[34] Quantitative measurement of $^{3}J_{C'C'}$ was also shown feasible for rather slowly tumbling proteins ( $\tau_c \approx 12$ ns), even prior to the introduction of highly sensitive cryogenic probeheads.[35] Here, we used a slightly modified 3D NMR experiment that does not require the 2-spin approximation used in prior work, and which yields cross peak/diagonal peak intensity ratios between sequential amide groups that are proportional to $\tan^2(\pi J_{C'C'}T)$ (See Figure S1 of the Supporting Information). A projection of the 3D spectrum on the $^{15}$N-$^{15}$N plane (Figure 2) illustrates the excellent spectral resolution and sensitivity that can be obtained, enabling the measurement of a nearly complete set of $^{3}J_{C'C'}$ couplings (Table S2 of the Supporting Information). Spectra were recorded at 15 and 25 °C, and the pairwise rmsd between these two sets of measurements was 0.06 Hz, indicating a random error of only 0.03 Hz in their averaged values.

A fit of the $^{3}J_{C'C'}$ values to the newly refined structure of GB3, calculated without using $^{3}J_{C'C'}$ or $^{3}J_{HNH\alpha}$ restraints, yields the following optimized Karplus equation:

$$^{3}J_{C'C'} = 1.61 \times \cos^2\phi - 0.93 \times \cos\phi + 0.55\ \text{Hz} \quad (1)$$

with an rmsd of 0.12 Hz between measured and predicted couplings (Fig. 3). For comparison, this rmsd increases from 0.12 to 0.29 Hz when using $\varphi$ angles taken from the high resolution X-ray structure (PDB entry 1IGD). A small, 4% drop in this rmsd value is obtained when using the PDB-REDO refinement[30] of the 1IGD structure factors. A slightly larger drop (~10%) is obtained for the joint X-ray/NMR refinement,[31] but a considerably larger decrease (~35%) is observed for the original 2OED NMR structure (Table 1). As mentioned above, the finding that the improvement in the $^{3}J$ fit of the joint X-ray/NMR refined structure over the original X-ray structure is modest presumably results from the weak weight assigned to the RDC restraints, also reflected in a minimal backbone coordinate rmsd relative to the 1IGD X-ray structure (see Table S1 of the Supporting Information). However, our finding that the joint X-ray/NMR refinement consistently yields $^{3}J_{C'C'}$ and $^{3}J_{HNH\alpha}$ rmsd values that are considerably higher than obtained for the 2OED NMR structure also points to true differences between the structures in the crystalline state and in free solution. Indeed, intermolecular H-bonding results in a continuous plane of β-sheet in the crystalline state,[15] and in all likelihood contributes to this modest difference in backbone structure.

The experimental $^{3}J_{C'C'}$ values in GB3 span a range of 2.46 Hz, which is 20-fold larger than the rmsd in their fit to the Karplus equation. This is nearly the same ratio as observed for $^{3}J_{HNH\alpha}$ (range of 7.2 Hz, versus 0.34 Hz rmsd). However, as discussed above, from a practical perspective the interpretation of $^{3}J_{HNH\alpha}$ couplings is limited to an uncertainty of 0.69 Hz, due to the fact that the deviations from idealized geometry for hydrogen positions

normally cannot be determined. Therefore, $^3J_{C'C'}$ couplings intrinsically have much higher restraining power when determining a protein structure from NMR data, in particular for small proteins where these couplings can be measured at very high accuracy.

## Factors other than ϕ impacting $^3J_{C'C'}$

Although the fit of the $^3J_{C'C'}$ couplings to the Karplus equation is very good, the residual we observe is nearly four times larger than the 0.03 Hz experimental measurement uncertainty. Although errors in the coordinates of the NMR-refined structure will contribute to the 0.12 Hz rmsd, the scatter does not correlate with the steepness of the curve, and therefore does not appear to be dominated by uncertainty in $\varphi$. Variations in the amplitude of $\varphi$-angle fluctuations among different residues would also scale the couplings differentially and thereby contribute to the scatter.[11] However, for the protein ubiquitin a very long (1-ms) molecular dynamics trajectory has been reported,[23] which shows that the amplitude of φ-angle fluctuations in the well ordered regions of the protein is quite homogeneous (rmsd of 13±3°), and insufficiently large to lead to significant differential scaling of the $^3J_{C'C'}$ couplings.[11] The same conclusion must apply for GB3, which also is a very well structured protein.

We therefore carried out quantum calculations to evaluate what other geometric factors can be expected to significantly impact $^3J_{C'C'}$. A tripeptide Gly-Ala-Gly was first built using the program Molden,[36] both in helical ($\varphi = -50,\ \psi = -50$) and extended ($\varphi = -120,\ \psi = 120$) backbone conformations. The tripeptide then was modified *in silico* to generate a smaller dipeptide analogue (Ac-Ala-NH-$CH_3$). Four water molecules were placed near the carbonyl and amide groups, in a linear hydrogen bond geometry (Fig. 4a). The water molecule that is hydrogen bonded to the C-terminal amide cap was rotated towards the N-methyl group within the peptide plane by 20° (i.e. making the C′–N–$O_{water}$ angle 140° instead of 120°) such that this water molecule does not significantly vary its interaction with nearby water molecules when changing backbone torsion angles.

The O–O or N–O distance were independently stepwise increased from 2.65 to 3.15 Å, except for the water H-bonded to the C-terminal N-methyl group, which was kept fixed at $d_{NO}$ = 2.85 Å. Full results are included in Tables S3–S5 of the Supporting Information. As can be seen from Table S4, the impact of removing H-bonds on $^3J_{C'C'}$ is to a first approximation additive, and therefore these H-bonds may be considered separately. The largest impact is observed when altering the distance to water-2 (Fig. 4), which accepts an H-bond from Ala-NH. For the extended backbone geometry, the computed $^3J_{C'C'}$ decreases from 1.36 to 1.25 Hz when $d_{NO}$ is increased from 2.65 to 3.15 Å, but $^3J_{C'C'}$ variation for the helical φ = −50° value is considerably smaller (Fig. 4b).

Variations in the N $C^{\alpha}$ C′ bond angle, $\tau$, also are known to occur in protein structures[37] and according to the DFT calculations can have a significant impact on $^3J_{C'C'}$, decreasing by ~0.15 Hz when $\tau$ is increased from 106.6 to 113.6°, both in extended and helical geometries (Fig. 4c). The DFT calculations also point to a non-negligible effect of amino acid type, both for $^3J_{C'C'}$ and for $^3J_{HNH\alpha}$ (Table S5 in the Supporting Information) and a substantial effect of the $\chi_1$ torsion angle on both $^3J_{C'C'}$ and $^3J_{HNH\alpha}$ is seen for Ser.

The above calculations confirm that variations in H-bonding and valence angles as well as residue type and their sidechain $\chi_1$ torsion angle are likely responsible for most of the residual scatter seen in Figures 1 and 3. Substantial outliers in these fits would be expected if, for example, an amide group were completely lacking an H-bond, or if a large deviation from ideality for $\tau$ would occur. The absence of such large outliers therefore indicates that both H-bonding and backbone valence angles are restricted to narrow ranges in GB3.

### Resolving degeneracies in the Karplus curve

Ignoring the small out-of-plane angles, $\Delta\varphi$, and the deviations from ideal tetrahedral geometry at $C^{\alpha}$, both $^{3}J_{C'C'}$ and $^{3}J_{HNH\alpha}$ report on the same backbone torsion angle, $\varphi$. However, the corresponding Karplus curves are shifted by 60° relative to one another (Fig. 5a). A plot of Karplus-predicted $^{3}J_{C'C'}$ *vs* $^{3}J_{HNH\alpha}$ values highlights that the $^{3}J_{HNH\alpha}$ degeneracy with respect to φ, around the extrema of the Karplus curve ($\varphi \approx -120° + N\times 90°$), is lifted by $^{3}J_{C'C'}$, and vice versa, degeneracies in the $^{3}J_{C'C'}$ values around $\varphi \approx 0°$, 180°, and ±73° are resolved by $^{3}J_{HNH\alpha}$ (Fig. 5b). The combination of these two sets of $^{3}J$ couplings therefore presents a very precise measure for the $\varphi$ angle in proteins, allowing this torsion angle to be determined at an accuracy that can exceed those of the best X-ray structures. The highly complementary nature of $^{3}J_{C'C'}$ and $^{3}J_{HNH\alpha}$, together with the high precision at which these values can be measured in flexible regions of proteins, also will make their joined use invaluable in defining the distribution of torsion angles sampled in intrinsically disordered proteins (IDPs) or large flexible loop regions.

## Methods

### NMR Measurements

GB3 $^{3}J_{C'C'}$ values were derived from two 3D HN(COCO)NH spectra, measured with the pulse sequence of Supporting Information Figure 1, using a 2.5 mM sample of uniformly $^{13}C/^{15}N$-enriched GB3 in 20 mM sodium phosphate, 50 mM NaCl, 0.05% w/v sodium azide, pH 6.6, 5% $D_2O$. The two spectra were recorded on a Bruker Avance-III 500 MHz NMR spectrometer equipped with a cryogenic z-gradient probehead, with sample temperatures set to 288 K and 298 K. The $^{13}C$ and $^{15}N$ carriers were set at 177 and 118 ppm, respectively. The time domain matrix consisted of 200* × 200* × 1024* complex data points, or acquisition times of 60.0 ms ($t_1$, $^{15}N$), 60.0 ms ($t_2$, $^{15}N$) and 146.3 ms ($t_3$, $^{1}H$), and approximately 88 hours of total measurement time, using 4 scans per free induction decay (FID), and a 1.5-s interscan delay. The time domain data were apodized with a 60°-shifted sine-bell window function in all dimensions, and zero-filled prior to Fourier transformation to yield high digital resolution. The spectra were processed and analyzed using the software package nmrPipe.[38]

### Structure calculations

An extensive set of previously reported backbone $^{1}D_{NH}$, $^{1}D_{C\alpha C'}$, $^{1}D_{NC'}$ and $^{1}D_{C\alpha H\alpha}$ RDCs measured under 5 alignment conditions for wild type GB3, and 6 sets of $^{1}D_{NH}$ values measured for a series of charge-perturbed mutants of GB3 aligned in liquid crystalline Pf1, was supplemented by 3 sets of $^{1}D_{C\alpha H\alpha}$ and $^{1}D_{C\alpha C\beta}$ RDCs for wild type GB3, aligned in Pf1,[39] polyethylene glycol,[40] and bicelles,[41] respectively.

The structure of GB3 was refined using the Xplor-NIH program[42] starting from the coordinates of the RCSB PDB deposition 2OED.[16] In addition to the N-H, $C^{\alpha}$-$H^{\alpha}$, C′-$C^{\alpha}$, and C′-N RDC restraints collected in five different alignment media, previously used to derive the 2OED model, the refinement protocol included previously reported N-H, $C^{\alpha}H^{\alpha}$, and C′-$C^{\alpha}$ RDCs measured in Pf1 for GB3 mutants K19A/D47K, K19E/D40N, K19E/K4A-N-$His_6$, K19E/K4A-C-$His_6$, K19A/T11K, and K19E/K4A.[18] The input RDCs also included N-H, $C^{\alpha}$-$H^{\alpha}$, and $H^{N}$-$H^{\alpha}$ RDCs measured for wild type GB3 in Pf1,[17] as well as newly collected N-H, $C^{\alpha}$-$H^{\alpha}$ and $C^{\alpha}$-$C^{\beta}$ RDCs from samples of wild-type GB3 in Pf1, bicelles and PEG, and N-H, $C^{\alpha}$-$H^{\alpha}$ RDCs from K4A/K19E/V42E-C$His_6$, K19A/V42E/D47K, and K4A/K19E/V42E mutants in Pf1. Fitting of variable-distance $H^{N}$-$H^{\alpha}$ RDCs was done using the xdip term in XPLOR[42] with a scaling constant of 11.13 with respect to the N-H RDCs, corresponding to the libration-corrected N-H bond length of 1.041Å. Unlike the 2OED deposition, no non-crystallographic symmetry restraint terms with respect to the 1IGD crystal structure were used in the refinement. Backbone/backbone hydrogen bonding geometries were enforced via a previously described database potential of mean force,[43] modified to include refinement against experimental through-H-bond $^{3h}J_{NC'}$ couplings.[44]

The empirical force field used in the refinement was slightly modified to better cope with a large number of input backbone RDC restraints. N-H and $C^{\alpha}$-$H^{\alpha}$ bond lengths were set to 1.02 and 1.09 Å, and improper torsion angles governing positioning of the $H^{N}$ and $H^{\alpha}$ atoms were softened from their default setting of 500 kcal $mol^{-1}$ $rad^{-2}$ to 50 kcal $mol^{-1}$ $rad^{-2}$. The empirical force field parameters used to describe the alignment tensor pseudo molecules were strengthened by introducing three additional bond and three additional angle restraint terms. In order to evaluate the effect of softening the empirical potential used to restrain the peptide bond angle Ω to 180°, the force constants for the corresponding improper terms were decreased by factors of 3, 5, and 10 with respect to their defaults settings of 500 kcal $mol^{-1}$ $rad^{-2}$, with 5-fold reduction yielding the best agreement between calculated structures and $^{3}J_{H\alpha HN}$ and $^{3}J_{C'C'}$ couplings in terms of rmsd observed in a Karplus equation fit. Note that these couplings were not used as input restraints, and that no $\varphi$ torsion angle restraints were used for structure refinement.

Due to differences between the alignment conditions of the individual samples from which the fitted RDC data were extracted, a total of 24 alignment tensors were used to fit the entire set of RDCs. The RDC data were found to fully cover the 5-dimensional space of alignment tensors as judged by SECONDA analysis.[45, 46] Therefore, RDC force constants were optimized by cross-validation, frequently resulting in values lower than those based solely on alignment tensor magnitudes and rhombicities. To account for the differences in static coupling magnitudes of various RDC types, force constant multipliers of 26, 75, 0.24, and 5.0 were used for C-C, C-N, C-H, and H-H RDCs. During structure refinement, both eigenvalues and orientations of the alignment tensors were adjusted every 0.1 ps by SVD-fitting. The structure refinement used Cartesian dynamics and consisted of 200 steps of conjugate gradient minimization, followed by a 2-ps high temperature stage at 2000 K, followed with a 100 ps simulated annealing schedule from 200 K to 0 K, and finishing with 200 steps of conjugate gradient minimization. During the entire procedure, non-bonded interactions were modeled by quartic repulsive-only energy terms, using a force constant of

4.0 kcal $mol^{-1}$ $A^{-2}$. All van der Waals radii are scaled by a factor of 0.81. The actual refinement script and all input tables as well as the coordinates of the refined model can be downloaded from http://spin.niddk.nih.gov/bax/structures/GB3_RDC_refinement_2014.zip

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

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 a KVSTA Fellowship and F. Li acknowledges financial support from the China Scholarship Council. This study utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, Maryland. (http://biowulf.nih.gov).

## References

1. Karplus M. J Chem Phys. 1959; 30:11.
2. Bystrov VF. Prog NMR Spectrosc. 1976; 10:41.
3. Williamson KL. J Am Chem Soc. 1963; 85:516.
4. Laszlo P, Schleyer PV. J Am Chem Soc. 1963; 85:2017.
5. Smith GV, Kriloff H. J Am Chem Soc. 1963; 85:2016.
6. Chapman OL. J Am Chem Soc. 1963; 85:2014.
7. Karplus M. J Am Chem Soc. 1963; 85:2870.
8. Haasnoot CAG, Deleeuw F, Altona C. Tetrahedron. 1980; 36:2783.
9. Pardi A, Billeter M, Wuthrich K. J Mol Biol. 1984; 180:741. [PubMed: 6084720]
10. MacArthur MW, Thornton JM. J Mol Biol. 1996; 264:1180. [PubMed: 9000639]
11. Bruschweiler R, Case DA. J Am Chem Soc. 1994; 116:11199.
12. Case DA, Scheurer C, Bruschweiler R. J Am Chem Soc. 2000; 122:10390.
13. Vogeli B, Ying JF, Grishaev A, Bax A. J Am Chem Soc. 2007; 129:9377. [PubMed: 17608477]
14. Wang AC, Bax A. J Am Chem Soc. 1996; 118:2483.
15. Derrick JP, Wigley DB. J Mol Biol. 1994; 243:906. [PubMed: 7966308]
16. Ulmer TS, Ramirez BE, Delaglio F, Bax A. J Am Chem Soc. 2003; 125:9179. [PubMed: 15369375]
17. Vogeli B, Yao L, Bax A. J Biomol NMR. 2008; 41:17. [PubMed: 18458825]
18. Yao L, Voegeli B, Ying JF, Bax A. J Am Chem Soc. 2008; 130:16518. [PubMed: 19049453]
19. Edison AS. Nature Structural Biology. 2001; 8:201.
20. Berkholz DS, Driggers CM, Shapovalov MV, Dunbrack RL Jr, Karplus PA. Proc Natl Acad Sci U S A. 2012; 109:449. [PubMed: 22198840]
21. Lindorff-Larsen K, Best RB, DePristo MA, Dobson CM, Vendruscolo M. Nature. 2005; 433:128. [PubMed: 15650731]
22. 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. [PubMed: 18556554]
23. Piana S, Lindorff-Larsen K, Shaw DE. Proc Natl Acad Sci U S A. 2013; 110:5915. [PubMed: 23503848]
24. Maltsev AS, Grishaev A, Roche J, Zasloff M, Bax A. J Am Chem Soc. 2014; 136:3752. [PubMed: 24568736]
25. Hall JB, Fushman D. J Biomol NMR. 2003; 27:261. [PubMed: 12975584]
26. Yao L, Vogeli B, Torchia DA, Bax A. J Phys Chem B. 2008; 112:6045. [PubMed: 18358021]

27. Rexroth A, Schmidt P, Szalma S, Geppert T, Schwalbe H, Griesinger C. J Am Chem Soc. 1995; 117:10389.
28. Harbison GS. J Am Chem Soc. 1993; 115:3026.
29. Word JM, Lovell SC, Richardson JS, Richardson DC. J Mol Biol. 1999; 285:1735. [PubMed: 9917408]
30. Joosten RP, Salzemann J, Bloch V, Stockinger H, Berglund AC, Blanchet C, Bongcam-Rudloff E, Combet C, Da Costa AL, Deleage G, Diarena M, Fabbretti R, Fettahi G, Flegel V, Gisel A, Kasam V, Kervinen T, Korpelainen E, Mattila K, Pagni M, Reichstadt M, Breton V, Tickle IJ, Vriend G. J Appl Crystallogr. 2009; 42:376. [PubMed: 22477769]
31. Rinaldelli M, Ravera E, Calderone V, Parigi G, Murshudov GN, Luchinat C. Acta Crystallogr Sect D-Biol Crystallogr. 2014; 70:958. [PubMed: 24699641]
32. Koradi R, Billeter M, Wuthrich K. J Mol Graph. 1996; 14:51. [PubMed: 8744573]
33. Hu JS, Bax A. J Am Chem Soc. 1996; 118:8170.
34. Bermel W, Bertini I, Felli IC, Lee YM, Luchinat C, Pierattelli R. J Am Chem Soc. 2006; 128:3918. [PubMed: 16551093]
35. Grzesiek S, Bax A. J Biomol NMR. 1997; 9:207. [PubMed: 9090134]
36. Schaftenaar G, Noordik JH. J Comput-Aided Mol Des. 2000; 14:233. [PubMed: 10756478]
37. Laskowski RA, Moss DS, Thornton JM. J Mol Biol. 1993; 231:1049. [PubMed: 8515464]
38. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. J Biomol NMR. 1995; 6:277. [PubMed: 8520220]
39. Hansen MR, Mueller L, Pardi A. Nature Struct Biol. 1998; 5:1065. [PubMed: 9846877]
40. Ruckert M, Otting G. J Am Chem Soc. 2000; 122:7793.
41. Ottiger M, Bax A. J Biomol NMR. 1998; 12:361. [PubMed: 9835045]
42. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM. J Magn Reson. 2003; 160:65. [PubMed: 12565051]
43. Grishaev A, Bax A. J Am Chem Soc. 2004; 126:7281. [PubMed: 15186165]
44. Cornilescu G, Ramirez BE, Frank MK, Clore GM, Gronenborn AM, Bax A. J Am Chem Soc. 1999; 121:6275.
45. Hus JC, Bruschweiler R. J Biomol NMR. 2002; 24:123. [PubMed: 12495028]
46. Tolman JR. J Am Chem Soc. 2002; 124:12020. [PubMed: 12358549]

## Highlights

- A convenient, sensitive, and accurate method for $^{3}J_{C'C'}$ measurement is presented
- Karplus equation fits for $^{3}J_{C'C'}$ and $^{3}J_{HNH\alpha}$ in proteins are limited by the accuracy of these structures
- Best $^{3}J_{HNH\alpha}$ Karplus equation fit is obtained when using the experimentally derived H-N-$C^{\alpha}$-$H^{\alpha}$ dihedral angle, rather than $\varphi$
- $^{3}J_{C'C'}$ values in proteins are reliable $\varphi$ reporters, minimally impacted by residue type and variations in H-bonding.

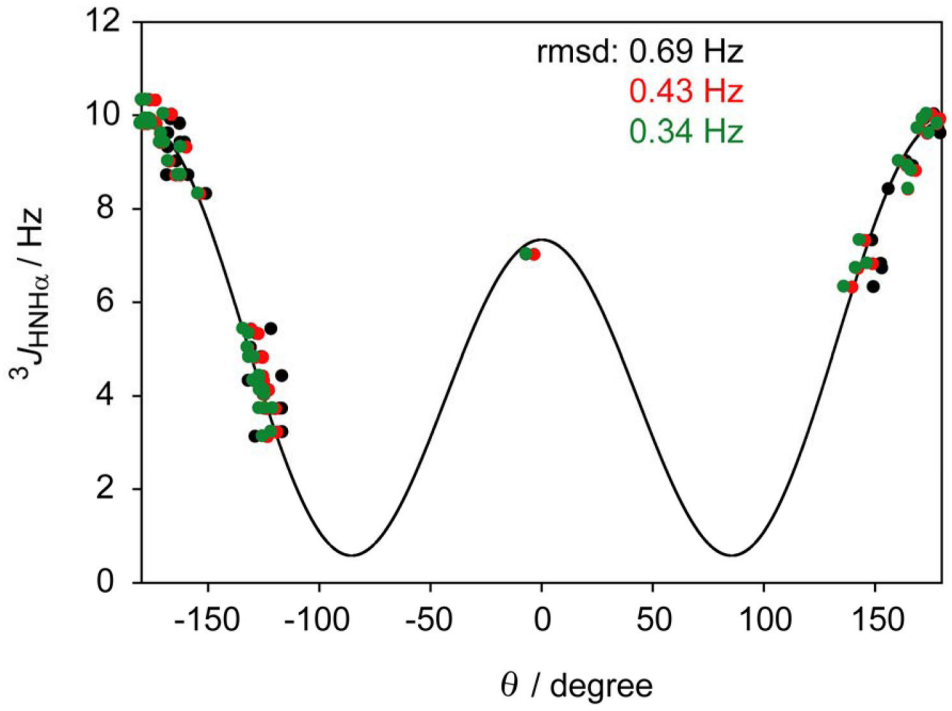


**Figure 1.**

Plot of $^3J_{HNH\alpha}$ values measured for GB3 against the H-N-$C^\alpha$-$H^\alpha$ dihedral angle, $\theta$ (Note: $\theta \approx \varphi - 60°$). Dihedral angles derived from the coordinates of the RDC-refined NMR structure with amide protons in their RDC-optimized positions (red data points; $\theta = \varphi - 60° + \Delta\varphi$), with $H^N$ atoms moved to their idealized in-peptide plane positions (black data points; $\theta = \varphi - 60°$), or with both $H^N$ and $H^\alpha$ atoms in their RDC-optimized positions (green data points; $\theta$ being the H-N-$C^\alpha$-$H^\alpha$ dihedral angle).

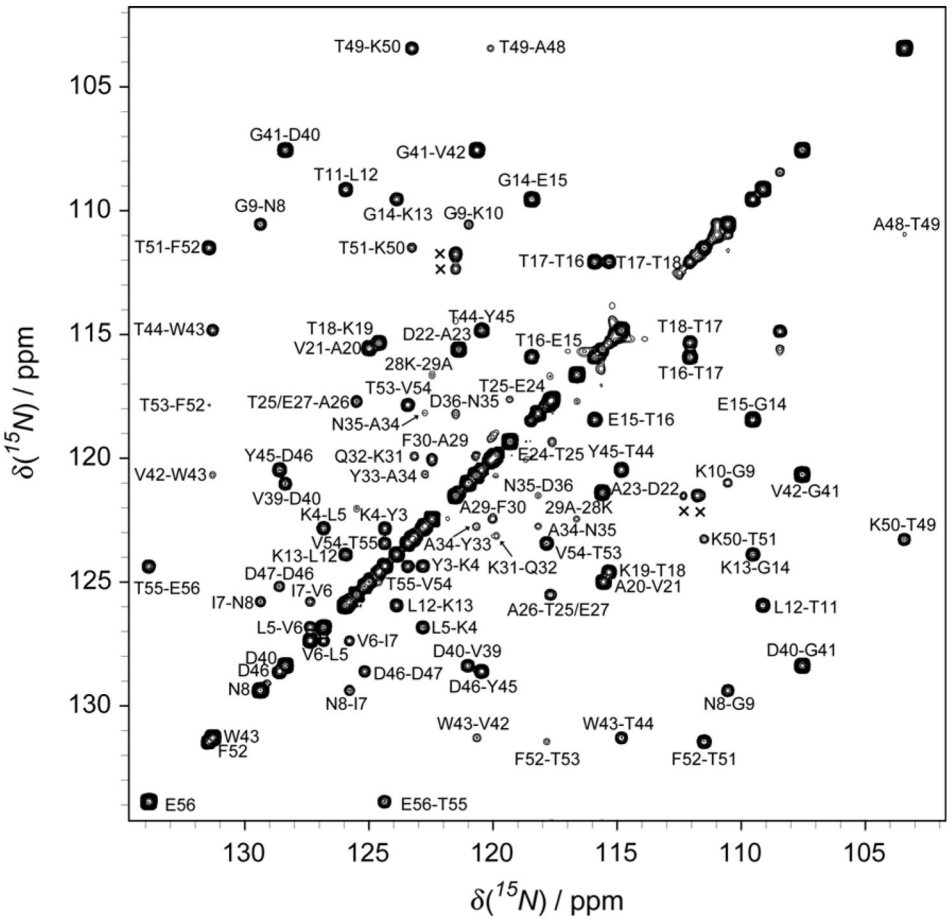


**Figure 2.**

Projection of the 500 MHz 3D HN(COCO)NH spectrum of GB3 on the $^{15}$N-$^{15}$N ($F_1$, $F_2$) plane, recorded at 298 K. The projection extended from 10.5 to 7.0 ppm in the $H^N$ ($F_3$) dimension. Cross peak to diagonal peak ratios correspond to $\tan^2(\pi J_{C'C'}T)$, with the $^{13}$C′-$^{13}$C′ de-/re-phasing interval $T$ set to 100 ms. Cross peaks are labeled, and peaks marked × show the correlations between N35 $^{15}N^{\delta}$ (with and without $^2$H isotope shift) and D36 $^{15}$N.

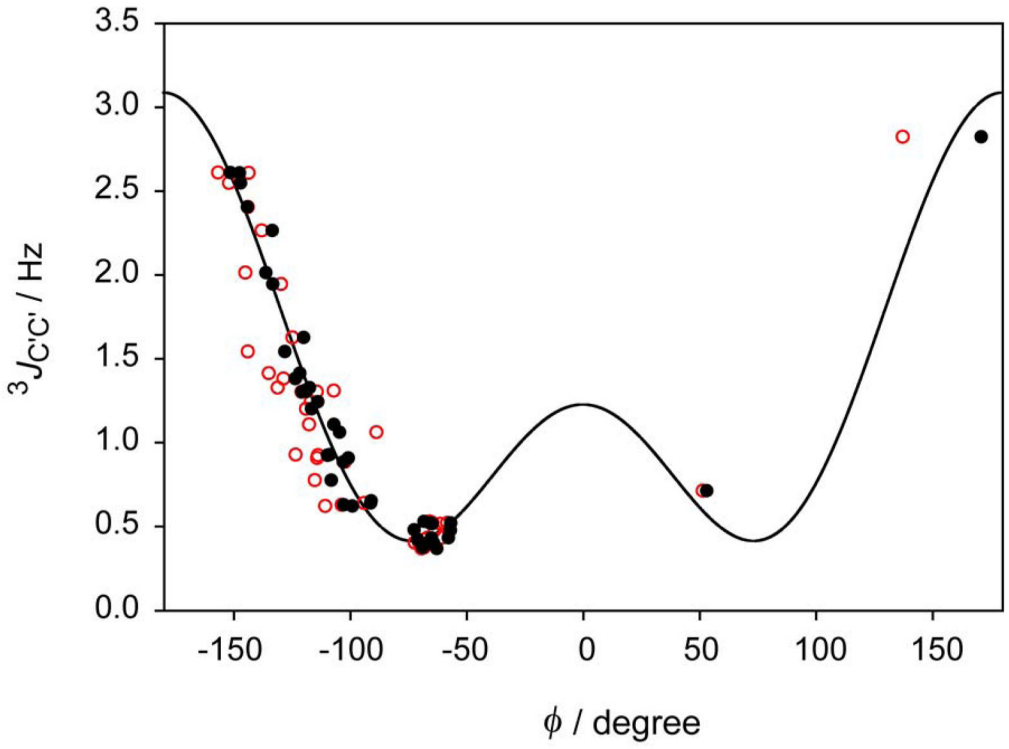


**Figure 3.**

Plot of experimental $^3J_{C'C'}$ values against the backbone torsion angle $\varphi$. The solid line corresponds to the Karplus equation (eq 1). Red data points correspond to $\varphi$ angles extracted from the X-ray structure (PDB entry 1IGD)[15]; black data points correspond to the RDC-refined GB3 structure. Dynamically disordered residues (L12, D40, and G41, previously identified on the basis of $^{15}$N relaxation order parameters $S^2$ <0.7, axially symmetric diffusion model,[25] are excluded from the plot. The rmsd between observed and best-fitted $^3J_{C'C'}$ values is 0.12 Hz.

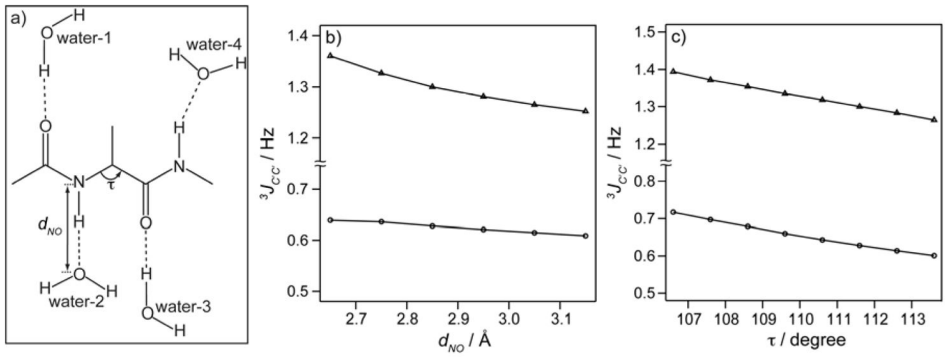


**Figure 4.**

Effect of H-bonding on $^3J_{C'C'}$ as obtained from DFT calculations. (a) The dipeptide analog Ac-Ala-NH-$CH_3$ H-bonded to four water molecules, used in the DFT calculations. (b) $^3J_{C'C}$ as a function of $d_{NO}$, the distance between the O of water-2 and the Ala N atom, for α-helical (circles) and β-sheet (triangles) backbone geometry. (c) $^3J_{C'C'}$ as a function of the $\tau$ angle, for α-helical (circles) and β-sheet (triangles).

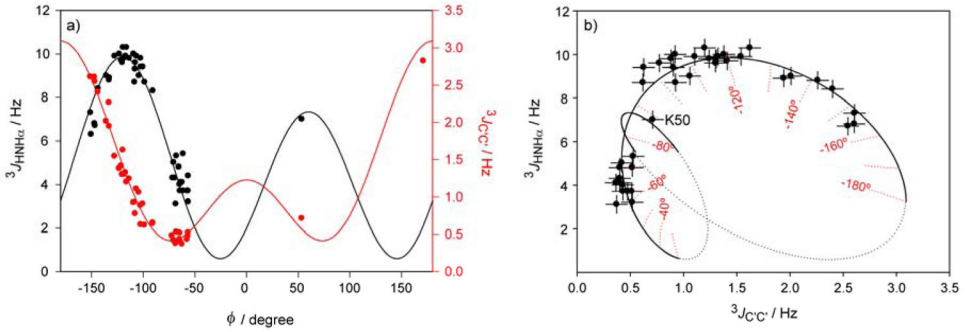


**Figure 5.**

Resolving ambiguities in the Karplus relation for determining the backbone torsion angle $\varphi$ in proteins. (a) Karplus relations between $^{3}J_{C'C'}$ (red) or $^{3}J_{HNH\alpha}$(black) values and the backbone torsion angle, $\varphi$, assuming the out-of-plane angle $\Delta\varphi = 0°$. (b) Plot of $^{3}J_{C'C'}$ *vs* $^{3}J_{HNH\alpha}$ values in GB3, again assuming $\Delta\varphi = 0°$. Only residues that are well-ordered ($\boldsymbol{S}^2 \geq 0.7$) on the basis of prior $^{15}$N relaxation studies[25] are shown. The solid line corresponds to the relation between the predicted values in the most favored region of the Ramachandran plot ($-180° < \varphi < -30°$; $30° < \varphi < 90°$), with the dotted section corresponding to sparsely populated or forbidden $\varphi$ angles. For calculating this curve, $\Delta\varphi = 0°$. Error bars correspond to ±0.4 Hz for $^{3}J_{HNH\alpha}$ and to ±0.1 Hz for $^{3}J_{C'C'}$. Residue K50 adopts a positive $\varphi$ angle ($\varphi = 53°$).

**Table 1**

Rmsd between measured $^3J$ couplings and values predicted by the Karplus equations.[a]

| | $^3J_{HNH\alpha}$ | | | | $^3J_{C'C'}$ |
|---|---|---|---|---|---|
| | **angle used for Karplus equation** | | | | |
| **Model** | $\theta_{HNH\alpha}$ | $(\theta_{HNC\beta} + \theta_{HNC'} + 360°)/2$ | $(\varphi - 60°)$ [c] | $(\varphi - 60° + \Delta\varphi)$[d] | $\varphi$ |
| 1IGD | 0.76 | 0.71 | 0.74 | 0.72 | 0.29 |
| REDO[e] | 0.76 | 0.74 | 0.76 | 0.69 | 0.28 |
| Xray/NMR[f] | 0.62 | 0.71 | 0.68 | 0.63 | 0.26 |
| 2OED | 0.43 | 0.48 | 0.58 | 0.51 | 0.19 |
| New | 0.34 [b] | 0.37 [b] | 0.69 | 0.43 | 0.12 |

[a] In units of Hz, excl. residues L12, D40, and G41. Predicted values are derived from the previously parameterized Karplus equation,[13] $^3J_{HNH\alpha} = 7.97 \times \cos^2\theta - 1.26 \times \cos\theta + 0.63$ Hz, and the newly parameterized $^3J_{C'C'} = 1.61 \times \cos^2\varphi - 0.93 \times \cos\varphi + 0.55$ Hz.

[b] The rmsd between $\theta_{HNH\alpha}$ and $(\theta_{HNC\beta} + \theta_{HNC'} + 360°)/2$ in the newly refined model is 1.2°.

[c] Using $\theta = \varphi - 60°$, where $\varphi$ is the $C'_{i-1}$-$N_i$ -$C^{\alpha}_i$ -$C'_i$ dihedral angle.

[d] Using $\theta = \varphi - 60° + \Delta\varphi$, where $\Delta\varphi$ is the out-of-plane angle observed in the newly refined NMR structure.

[e] X-ray structure calculated from the 1IGD structure factors using PDB_REDO.[30]

[f] Structure calculated using joint 1IGD X-ray structure factors and 2OED NMR restraints. Coordinates taken from the supporting information of Rinaldelli et al.[31]