High-Accuracy Residual $^1\text{H}^\text{N}-^{13}\text{C}$ and $^1\text{H}^\text{N}-^1\text{H}^\text{N}$ Dipolar Couplings in Perdeuterated Proteins

Sebastian Meier, Daniel Häussinger, Pernille Jensen, Marco Rogowski, and Stephan Grzesiek*

Biozentrum, University of Basel, Basel 4056, Switzerland

Received September 29, 2002 ; E-mail: stephan.grzesiek@unibas.ch

Residual tensorial couplings of weakly aligned molecules have become an invaluable source of structural and dynamical information in high-resolution NMR.[1] Most commonly, this information is extracted from one-bond residual dipolar couplings (RDCs). The size of RDCs depends on the direction of the internuclear distance vector, its length r, its order parameter S, the magnetic moments of the nuclei, and the degree of alignment. Although in principle, long-range RDCs could define the molecular geometry more efficiently, their use has not been widespread, because the $1/r^3$ dependence of dipolar couplings reduces the signal size and therefore sensitivity and introduces ambiguities in cases where the internuclear distance is not determined by the covalent geometry. However, for small to medium-sized biomacromolecules, a significant part of the sensitivity problem is not caused by incoherent signal losses from relaxation during the transfer delays, but by coherent broadening from multiple residual dipolar couplings to protons.

In this communication we show that perdeuteration followed by reprotonation of labile hydrogen positions[2,3] greatly alleviates this problem for the observation of long-range RDCs from amide protons ($^1\text{H}^\text{N}$) to surrounding ^{13}C as well as $^1\text{H}^\text{N}$ nuclei. Very recently, Wu and Bax[4] have proposed a similar approach for the enhancement of $^1\text{H}^\text{N} \leftrightarrow {}^1\text{H}^\text{N}$ RDCs. For small perdeuterated proteins, a large number (up to 10 in protein G) of such RDCs to ^{13}C and $^1\text{H}^\text{N}$ can be observed from individual amide protons with high accuracy, thus defining individual internuclear distances to within few picometers.

In contrast to experiments for the determination of one- and two-bond $^1\text{H}^\text{N} \leftrightarrow {}^{13}\text{C}$ RDCs,[5] the long-range couplings were measured by a quantitative J-correlation experiment (Figure 1) that detects frequencies of both coupling partners. The three-dimensional (3D) experiment consists of two concatenated $^1\text{H}-^{13}\text{C}/^1\text{H}-^{15}\text{N}$ HSQCs. In brief, proton magnetization H^N_y present at point a is transferred by dipolar and scalar couplings ($K_{\text{HC}} = J_{\text{HC}} + D_{\text{HC}}$) into carbon antiphase magnetization $\text{H}^\text{N}_z\text{C}_y$ at point b. After a frequency labeling period t_1, the carbon antiphase magnetization is transferred into the usual $\text{H}^\text{N}_z\text{N}_y$ magnetization for the $^1\text{H}-^{13}\text{N}$ HSQC at point c. The transfer scheme gives rise to signals at frequency positions (ω_C, ω_N, ω_H) with intensities proportional to $\sin^2(2\pi K_{\text{HC}}T)$. A second, 2D reference experiment is carried out where the phase cycle for $\phi_{1,2,3}$ is changed for the selection of in-phase H^N_z magnetization at point b. The t_1-period is not sampled, and the experiments detects signals at frequency positions (ω_N, ω_H) with intensities proportional to $\cos^2(2\pi K_{\text{HC}}T)$. The coupling constant K_{HC} can then be extracted from the ratio of cross and reference peak intensities after a suitable correction for the 2- and 3D acquisition schemes has been applied. This procedure is conceptually similar to the 2D quantitative $^1\text{H}-^{13}\text{C}$-HMQC,[6] and further details are given in the Supporting Information.

The high accuracy of the method was demonstrated on weakly aligned, perdeuterated protein G for which a 1.1 Å crystal structure
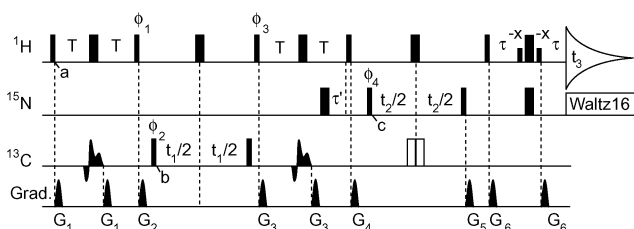

Figure 1. Pulse sequence of the quantitative HCN-HSQC. Narrow and wide pulses denote 90° and 180° flip angles, respectively, and unless indicated the phase is x. Further details are described in the Supporting Information.
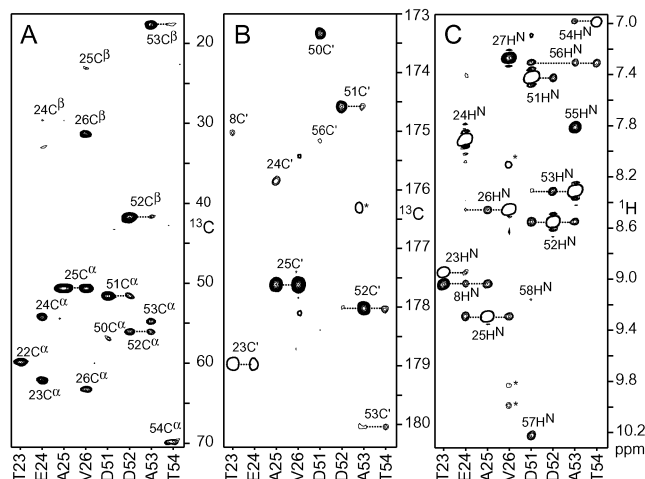

Figure 2. Cross sections extracted from 800 MHz 3D experiments for the detection of $^1\text{H}^\text{N}$ RDCs to aliphatic ^{13}C (A), carbonyl ^{13}C (B), and $^1\text{H}^\text{N}$ (C) nuclei. Negative intensities are shown as single thick contour lines. Asterisks indicate cross-peaks resulting from overlap of NH resonances or incomplete side chain deuteration. Experimental conditions: 2.2 mM $^{15}\text{N}/^{13}\text{C}/^2\text{H}$ protein G, pH 5.6, 98% $\text{H}_2\text{O}/2\%$ D_2O, 30 mg/mL Pf1,[9] 25 °C. (A) aliphatic HCN-HSQC. T = 10 ms; data matrixes consisted of 110* × 20* × 768* points; $t_{1,2,3,\text{max}}$ = 8.8, 24, 64 ms; t_{exp} = 32 h. (B) carbonyl HCN-HSQC. T = 10 ms; 50* × 20* × 768* data points; $t_{1,2,3,\text{max}}$ = 30, 24, 64 ms; t_{exp} = 14 h. (C) ^{15}N-edited $^1\text{H}-^1\text{H}$ COSY-HMQC. $^1\text{H}^\text{N} \leftrightarrow {}^1\text{H}^\text{N}$ transfer time = 47 ms; 72* (^1H) × 72* (^{15}N) × 768* ($^1\text{H}^\text{N}$) data points; $t_{1,2,3,\text{max}}$ = 23, 50, 83 ms; t_{exp} = 38 h.

(1IGD)[7] exists. Due to the limited ^{13}C RF strength, two HCN-HSQC experiments were carried out, optimized separately (Supporting Information) for the detection of $^1\text{H}^\text{N}$ RDCs to aliphatic and carbonyl ^{13}C nuclei. Figure 2A,B shows cross sections from these 3D experiments extracted at the $^1\text{H}^\text{N}-^{15}\text{N}$ frequency positions of residues T23−V26 and D51−T54. Clearly visible are a large number of correlations to $^{13}\text{C}^\alpha$, $^{13}\text{C}^\beta$, and $^{13}\text{C}'$ nuclei of the same and the preceding amino acid; in some cases, the correlations even extend across the hydrogen bond (T23H$^\text{N} \leftrightarrow$ Y8C′, D51H$^\text{N} \leftrightarrow$ T56C′), to the second next residue (T54H$^\text{N} \leftrightarrow$ D52C′), or to $^{13}\text{C}^\gamma$ resonances (data not shown). In addition to these long-range $^1\text{H}^\text{N} \leftrightarrow {}^{13}\text{C}$ RDCs, a large number of $^1\text{H}^\text{N} \leftrightarrow \text{H}^\text{N}$ RDCs (Figure 2C)
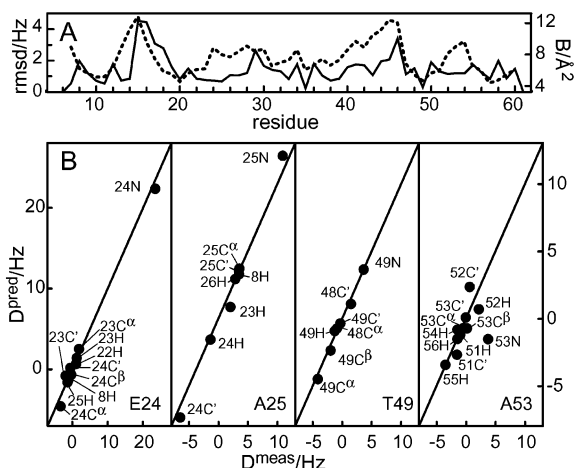
Figure 3. Accuracy of RDCs to $^1H^N$ in protein G. (A) rms deviations between measured and predicted RDCs per single residue (solid line) and crystallographic B-factor (dotted line). (B) comparison of measured and predicted RDCs for residues E24, A25, T49, A53.

could be detected by a water flip-back 3D ^{15}N-edited $^1H-^1H$ COSY-HMQC,[8] which is conventionally used for the quantification of scalar $^3J_{HNH\alpha}$ couplings. The observed $^1H^N \leftrightarrow ^1H^N$ correlations span distances of up to 7.2 Å.

To separate dipolar and scalar parts of the coupling constants, a second set of quantitative HCN-HSQCs was carried out on an isotropic sample of protein G. The dipolar couplings D_{HC} were then calculated as the difference of the two measurements. This procedure is not necessary for K_{HNHN}, since J_{HNHN} is negligible. Quantitative J-correlation is insensitive to the sign of the couplings constants. Therefore the method yields two possibilities for the value of D when $|J| < |D|$. In principle, more sophisticated (and less sensitive) techniques[4] could be used to resolve some of these ambiguities. In the present case, the ambiguities were resolved according to the value predicted from the crystal structure and an alignment tensor that was determined independently from one-bond $^1H^N \leftrightarrow ^{15}N$ RDCs, assuming an N−H distance of 1.02 Å.

The number and accuracy of the determined dipolar couplings is very high. For a total of 99 $^1H^N \leftrightarrow ^{13}C^{aliphatic}$, 85 $^1H^N \leftrightarrow ^{13}C'$, and 75 $^1H^N \leftrightarrow ^1H^N$ RDCs the rms deviations between predicted and measured values are 0.8, 1.1, 1.1 Hz, respectively (Supporting Information). Including the 55 $^1H^N \leftrightarrow ^{15}N$ RDCs, the total number of couplings per amide proton is 5.8 for residues T7−E61. The deviations are not equally distributed along the primary sequence (Figure 3A), but are strongest for the loop/turn regions around residues T16 and G46 located between strands $\beta 1$ and $\beta 2$ and between helix $\alpha 1$ and $\beta 3$. In other parts of the sequence, the deviations are very small, for example, 0.7, 0.7, and 0.3 Hz for E24, A25, and T49, respectively (Figure 3B).

Systematic errors in D_{HC} arise from imperfections of RF pulses. These were estimated smaller than 3% on the basis of a determination of $^1J_{CH}$ on a sample of ^{13}C-acetate using the first part of the HCN-HSQC sequence. Similar estimates can be made for D_{HH}.[8] Additional systematic errors of less than 2% for D_{HC} (D_{HH}) arise from incomplete ^{13}C-labeling (amide protonation). On the basis of these estimates, the total systematic error should be smaller than

4%, whereas the statistical error was determined as smaller than 0.14 Hz from a repetition of the experiments. For the present alignment tensor ($D_{NH,max} = 25.4$ Hz), such errors indicate that a single H−C distance of 2 Å can be determined to a precision that is higher than 0.03 Å.

Many of the deviations observed in Figure 3A are larger than these experimental errors and must be the result of an inaccurate description of the solution state by the crystal coordinates. A large part of these deviations result from short-range RDCs (e.g., Figure 3B, A53, and Supporting Information) where slight variations in structure lead to large changes in expected RDC values. A good correlation is observed between the deviations of the RDC values and the crystalline B-factor (Figure 3A). This indicates that disorder may be a main cause for the deviations in the loop regions and that the description by a single structure (with uniform order parameters) may not be justified.

In contrast to NOEs, which can also be obtained over large distances in perdeuterated proteins, RDCs do not suffer from noncoherent transfer processes such as spin diffusion or other nonlinear effects.[10,4] Therefore, their interpretation in terms of structure and dynamics is straightforward. Clearly, the present analysis depended on the availability of a high-resolution structure to determine the orientation tensor and to test the accuracy of the measured RDCs. In principle, the large number of observable long- and short-range RDCs should overdetermine the problem of an a priori structure calculation. This overdetermination can be further augmented by upper limits in cases where no long-range RDCs are detected and by the use of additional alignment media. Efforts are underway to explore this potential for the fast and accurate determination of small protein structures. Further possibilities exist to adapt the proposed experiments to somewhat larger proteins by TROSY techniques. Finally, for cases where initial structural information is available, the current example shows that the large number of accurate RDCs to single protons can reveal additional, unknown details of proton position and dynamics.

Supporting Information Available: Details of the HCN-HSQC experiment and Figures showing correlations between measured and predicted $H^N \leftrightarrow ^{13}C$, $^1H^N \leftrightarrow ^1H^N$, and $^1H^N \leftrightarrow ^{15}N$ RDCs (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

 (1) Tolman, J. R. *Curr. Opin. Struct. Biol.* **2001**, *11*, 532−539.
 (2) LeMaster, D. M.; Richards, F. M. *Biochemistry* **1988**, *27*, 142−150.
 (3) Torchia, D. A.; Sparks, S. W.; Bax, A. *J. Am. Chem. Soc.* **1988**, *110*, 2320−2321.
 (4) Wu, Z.; Bax, A. *J. Am. Chem. Soc.* **2002**, *124*, 9672−9673.
 (5) Yang, D.; Venters, R. A.; Mueller, G.; Choy, W. Y.; Kay, L. E. *J. Biomol. NMR* **1999**, *14*, 333−343.
 (6) Zhu, G.; Bax, A. *J. Magn. Reson. Ser. A* **1993**, *104*, 353−357.
 (7) Derrick, J. P.; Wigley, D. B. *J. Mol. Biol.* **1994**, *243*, 906−918.
 (8) Kuboniwa, H.; Grzesiek, S.; Delaglio, F.; Bax, A. *J. Biomol. NMR* **1994**, *4*, 871−878.
 (9) Hansen, M. R.; Mueller, L.; Pardi, A. *Nat. Struct. Biol.* **1998**, *5*, 1065−1074.
(10) Tian, F.; Fowler, C. A.; Zartler, E. R.; Jenney, F. A., Jr.; Adams, M. W.; Prestegard, J. H. *J. Biomol. NMR* **2000**, *18*, 23−31.

JA028740Q