

Identification of the Hydrogen Bonding Network in a Protein by Scalar Couplings

Gabriel Cornilescu, Jin-Shan Hu, and Ad Bax*

Laboratory of Chemical Physics
National Institute of Diabetes and
Digestive and Kidney Diseases
National Institutes of Health
Bethesda, Maryland 20892-0520

Received January 22, 1999

Revised Manuscript Received February 18, 1999

Hydrogen bonds (H bonds) are of key importance for stabilizing biomolecular structure, and for modulating the substrate binding specificity and reaction rate of virtually any enzymatic reaction.¹ In macromolecules, the presence of H bonds is indicated by the spatial proximity and relative arrangement of the atoms involved, after the structure has been solved by either crystallography or NMR. A variety of NMR spectroscopic parameters has also been used for characterizing these pivotal interactions. These include the effect of such bonds on isotropic chemical shifts² and chemical shift anisotropy,³ on the quadrupole coupling of ²H involved in a H bond,⁴ on protection of exchange of the labile hydrogen with solvent,⁵ and ¹H/²H fractionation.⁶ In a recent remarkable report, Dingley and Grzesiek⁷ were the first to demonstrate the presence of surprisingly large *J* couplings (6–7 Hz) between the H bond donating and accepting ¹⁵N nuclei in a Watson–Crick base pair in double-stranded RNA. This finding was confirmed by Pervushin et al.,⁸ and these authors additionally discovered the presence of a smaller (2–4 Hz) *J* couplings between the imino hydrogen itself and the H bond accepting ¹⁵N nucleus. These couplings confirm the interaction between the electronic orbitals of the atoms involved, and most importantly they identify unambiguously the pairs of atoms involved in a given H bond.

Although in organic chemistry the possibility of *J* couplings between atoms separated by less than the sum of the van der Waals radii has long been known, in particular for cases involving ¹⁹F,⁹ Summers et al. were the first to detect such couplings in a protein.¹⁰ They observed these interactions, ranging from 0.3 to 4 Hz, between the metal (Cd and Hg) in rubredoxin and backbone amide protons H-bonded to the S atoms ligating the metal. Remarkably, they also found *J* coupling between an alanyl methyl group, adjacent to such an S atom, and the metals.

(1) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer: New York, 1991.

(2) Wagner, G.; Pardi, A.; Wüthrich, K. *J. Am. Chem. Soc.* **1983**, *105*, 5948–5949. de Dios, A. C.; Pearson, J. G.; Oldfield, E. *Science* **1993**, *260*, 1491–1496.

(3) Berglund, B.; Vaughan, R. W. *J. Chem. Phys.* **1980**, *73*, 2037–2043. Tjandra, N.; Bax, A. *J. Am. Chem. Soc.* **1997**, *119*, 8076–8082. Tessari, M.; Vis, H.; Boelens, R.; Kaptein, R.; Vuister, G. W. *J. Am. Chem. Soc.* **1997**, *119*, 8985–8990.

(4) Soda, G.; Chiba, T. *J. Chem. Phys.* **1969**, *50*, 439–445. Hunt, M. J.; Mackay, A. L. *J. Magn. Reson.* **1976**, *22*, 295–301. LiWang, A. C.; Bax, A. *J. Magn. Reson.* **1997**, *127*, 54–64.

(5) Wagner, G. *Q. Rev. Biophys.* **1983**, *16*, 1–57. Englander, S. W.; Kallenbach, N. R. *Q. Rev. Biophys.* **1984**, *16*, 521–655.

(6) Loh, S. N.; Markley, J. L. *Biochemistry* **1994**, *33*, 1029–1036. Bowers, P. M.; Klevit, R. E. *Nat., Struct. Biol.* **1996**, *3*, 522–531. LiWang, A. C.; Bax, A. *J. Am. Chem. Soc.* **1996**, *118*, 12864–12865.

(7) Dingley, A. J.; Grzesiek, S. *J. Am. Chem. Soc.* **1998**, *120*, 8293–8297.

(8) Pervushin, K.; Ono, A.; Fernandez, C.; Szyperski, T.; Kainosho, M.; Wüthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14147–14151.

(9) Petrakis, L.; Sederholm, C. H. *J. Chem. Phys.* **1961**, *35*, 1243–1250. Brey, W. S.; Brey, M. L. In *Encyclopedia of Nuclear Magnetic Resonance*; Grant, D. M., Harris, R. K., Eds.; Wiley: Chichester, 1996; pp 2063–2071.

(10) Blake, P. R.; Adams, M. W. W.; Park, J.-B.; Summers, M. F. *J. Am. Chem. Soc.* **1992**, *114*, 6924–6925. Blake, P. R.; Lee, B.; Summers, M. F.; Adams, M. W. W.; Park, J.-B.; Zhou, Z. H.; Bax, A. *J. Biomol. NMR* **1992**, *2*, 527–533.

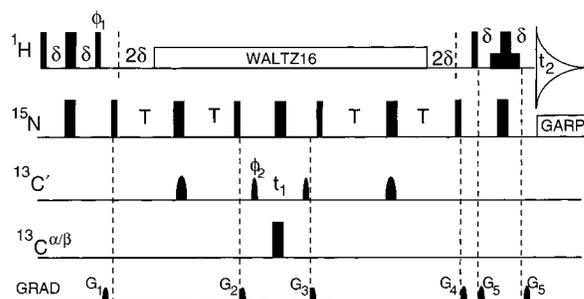


Figure 1. Pulse scheme of the 2D H(N)CO experiment, used to detect through-H bond *J* connectivities. Narrow and wide pulses correspond to flip angles of 90° and 180°, respectively. All pulses phases are *x*, unless specified. ¹³C' pulses have the shape of the center lobe of a sin *x/x* function, and durations of 150 μs. Shaped ¹³C' pulses at the midpoint of the 2*T* periods are 180°; pulses bracketing the ¹³C' evolution period are of lower power and correspond to 90°. Water suppression is accomplished by means of a WATERGATE sequence in the final reverse INEPT transfer. Composite ¹H decoupling was applied using WALTZ16 modulation, ¹⁵N decoupling during *t*₂ used GARP modulation. The ¹³C^{α/β} 180° pulse duration (applied at 50 ppm) was adjusted such that it has a null at the ¹³C' frequency. Delay durations: δ = 2.65 ms; *T* = 16.6, 50, or 66.6 ms (see text). Phase cycling: φ₁ = *y*, −*y*; φ₂ = 2(*x*), 2(−*x*); Receiver = *x*, 2(−*x*), *x*. Gradients are sinebell shaped, with peak amplitudes of 30 G/cm, and durations *G*_{1,2,3,4,5} = 1.1, 0.5, 0.6, 0.7, 0.4 ms, and directions *xyz*, *y, x*, −*x*, and *z*, respectively. For proteins with substantial overlap in the ¹H^N dimension, it may be advantageous to extend the experiment to a 3D scheme, where the second 2*T* period is converted into a constant-time ¹⁵N evolution period, and Rance–Kay coherence selection¹⁶ can be used for retaining sensitivity comparable to the 2D experiment.

Here, we demonstrate for the protein ubiquitin that small *J* couplings can be observed across H bonds between the donating ¹⁵N atom and the accepting carbonyl/carboxyl ¹³C. The scheme used for detecting this through-H-bond *J* connectivity (^h*J*_{NC'}) is essentially the regular CT-HNCO experiment,¹¹ optimized for the detection of these small ¹⁵N–¹³C' couplings (Figure 1). The magnitude of ^h*J*_{NC'} is then determined through the principle of quantitative analysis of cross-peak intensity.¹² In brief, the scheme of Figure 1 is executed three times: (A) with the dephasing time, 2*T*, tuned to 1/(2^h*J*_{NC'}), (B) with 2*T* = 3/(2^h*J*_{NC'}), (C) with 2*T* = 4/(2^h*J*_{NC'}). Values of ^h*J*_{NC'} in polypeptides are quite homogeneous (15 ± 1 Hz),¹³ and much larger than the ^h*J*_{NC'} which are the focus of this study. The signal intensity observed for a correlation between N and C' with coupling *J*_{NC'} is proportionate to exp(−4*T*/*T*₂) sin²(2π*J*_{NC'}*T*) Π_k cos²(2π*J*_{Nk}*T*), where the product extends over all other carbonyl/carboxyl carbons coupled to the ¹⁵N of interest with coupling constants *J*_{Nk}, and *T*₂ is the ¹⁵N transverse relaxation time. The dephasing times, 2*T*, used in this study have a maximum duration of 133.2 ms, and cos²(2π*J*_{Nk}*T*) ≈ 1 for *J*_{Nk} ≤ 1 Hz. Hence, a good approximation for *T*₂ is obtained from the ratio between experiments A and B. To a good approximation, therefore, the magnitude of the small *J* coupling between N and *k* can be derived from

$$\sin^2(2\pi J_{Nk} T) = I^C \exp(1/J_{NC'} T_2) / I^B \quad (1)$$

where *I*^C is the intensity of the (weak) correlation between ¹⁵N and *k*, observed in experiment C, and *I*^B is the intensity of the

(11) Grzesiek, S.; Bax, A. *J. Magn. Reson.* **1992**, *96*, 432–440.

(12) Bax, A.; Max, D.; Zax, D. *J. Am. Chem. Soc.* **1992**, *114*, 6923–6925. Bax, A.; Vuister, G. W.; Grzesiek, S.; Delaglio, F.; Wang, A. C.; Tschudin, R.; Zhu, G. *Methods Enzymol.* **1994**, *239*, 79–105.

(13) Delaglio, F.; Torchia, D. A.; Bax, A. *J. Biomol. NMR* **1991**, *1*, 439–446. Juranic, N.; Likic, V. A.; Prendergast, F. G.; Macura, S. *J. Am. Chem. Soc.* **1996**, *118*, 7859–7860.

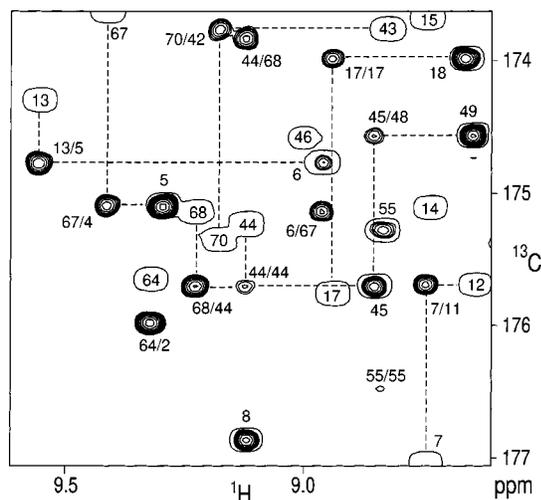


Figure 2. Superposition of small regions of the 600 MHz 2D H(N)CO spectra of U- $^{13}\text{C}/^{15}\text{N}$ ubiquitin, recorded with dephasing delays optimized for one bond connectivity (single contour, $2T = 1/2^1 J_{\text{NC}'} = 33.2$ ms), and optimized for small couplings (multiple contours, $2T = 2/1^1 J_{\text{NC}'} = 133.2$ ms). Correlations are marked by the residue number of the amide, followed by the number of the residue containing the carbonyl J -coupled to this amide. For one-bond $^1J_{\text{NC}'}$ connectivities to the preceding residue, only the amide is labeled. Both spectra result from $128^* \times 1024^*$ data matrixes, with acquisition times of 51 ms (t_1) and 60 ms (t_2). The reference spectrum was recorded with 32 scans per complex t_1 increment (total time 1.4 h); the spectrum with $2T = 133.2$ ms was recorded with 400 scans (total time 20 h).

one-bond N-C' correlation in experiment B. Implicit use of the $\cos^2(2\pi J_{\text{NC}'} T)$ 1 assumption in eq 1 results in a small (<10%) underestimate of the $J_{\text{NC}'}$ value derived from eq 1, which can be corrected for if $^1J_{\text{NC}'}$ is accurately known. In practice, through H-bond $^hJ_{\text{NC}'}$ couplings are very small, and such corrections are mostly smaller than the random uncertainty in the measurement.

The experiment is demonstrated for a 5 mg sample of uniformly $^{15}\text{N}/^{13}\text{C}$ -labeled ubiquitin (VLI Research, Malvern, PA), dissolved in 250 μL $\text{H}_2\text{O}/\text{D}_2\text{O}$ (15:1), pH 6.5. Experiments were carried out at 25 $^\circ\text{C}$, on a Bruker DMX600 spectrometer equipped with a triple resonance three-axis pulsed field gradient probehead. Figure 2 shows a small region of the 2D $\text{H}^{\text{N}}-^{13}\text{C}'$ correlation map, obtained with the pulse sequence of Figure 1. One-bond correlations, observed in the H(N)CO spectrum recorded with $2T = 33.2$ ms, are represented by a single contour. The 2D H(N)CO spectrum obtained with $2T = 2/1^1 J_{\text{NC}'} = 133.2$ ms is superimposed, and correlations observed in this spectrum are displayed in the regular multiple-contour manner. For these longer de- and rephasing durations, many of the correlations resulting from $^1J_{\text{NC}'}$ couplings are vanishingly weak, whereas for others (5, 6, 8, 18, 45, 49, 55) one-bond correlations to the preceding carbonyl result from small deviations from the $^1J_{\text{NC}'} = 15$ Hz condition. Numerous new correlations can be seen for the long dephasing times, mostly corresponding to J connectivity through hydrogen bonds. In particular, all backbone-backbone H bonds observed in the crystal structure¹⁴ of ubiquitin's five-stranded β -sheet result in strong correlations (two are obscured by resonance overlap). Several weak intraresidue correlations, marked 17/17, 44/44, and 55/55 in Figure 2, result from $^2J_{\text{NC}'}$ couplings, but the vast majority

of these intraresidue connectivities falls below the signal-to-noise threshold, indicating $|^2J_{\text{NC}'}| < 0.25$ Hz.

As discussed above, the intensity ratio in the spectra recorded with 133.2 and 100 ms de/rephasing periods can be used directly to obtain a quantitative measure for the magnitude of the small $J_{\text{NC}'}$ couplings giving rise to the through-H-bond correlations. Values observed for through-H-bond couplings range from 0.3 Hz for R72H $^{\text{N}}$ -Q40C' to 0.8 Hz for E64H $^{\text{N}}$ -Q2C'. For the β -sheet H bonds we find an average $^hJ_{\text{NC}'}$ value of 0.56 ± 0.10 Hz. For ubiquitin's α -helix, all but one of its H bonds (K33H $^{\text{N}}$ -K29C') present in the crystal (and NMR) structure give rise to observable cross-peaks. For the missing correlation, $^hJ_{\text{NC}'}$ falls below the detection limit of ~ 0.25 Hz. Excluding this unobservably small $^hJ_{\text{NC}'}$ coupling, the average for the α -helical $^hJ_{\text{NC}'}$ values is 0.37 ± 0.15 Hz. None of the 3_{10} H bonds observed in the crystal and NMR structures, which tend to be longer than β -sheet H bonds, give rise to observable correlations. In contrast, a number of irregular types of H bonds, not part of helix or sheet, give rise to substantial correlations (e.g., S57H $^{\text{N}}$ -P19C' and I23H $^{\text{N}}$ -R54C'). All of these correspond to short (<2 \AA) H $^{\text{N}}$ -O distances in the X-ray structure.

Detection of the weak $^hJ_{\text{NC}'}$ connectivities relies on the use of long de- and rephasing intervals. Although for small proteins such as ubiquitin this does not present a significant problem, for larger proteins the long intervals during which ^{15}N magnetization is relaxing at the transverse relaxation rate result in much lower sensitivity. One way to alleviate this problem is to modify the pulse scheme in such a way that only the most slowly relaxing ^{15}N doublet component is conserved. As shown by Pervushin et al., transverse relaxation times for the downfield ^{15}N doublet component remain long in larger, perdeuterated proteins in which the exchangeable hydrogens are protonated.¹⁵ A pulse scheme suitable for such measurements is presented in the Supporting Information, together with the full 2D H(N)CO spectrum. An additional advantage of this latter method is that H bonds involving side chain NH_2 groups are also visible with this method, although none were observed in ubiquitin. Note that such correlations are eliminated with the scheme of Figure 1, where the $^{15}\text{N}-\{^1\text{H}\}$ rephasing delays are optimized for NH and not NH_2 groups. For smaller, protonated proteins such as ubiquitin, the scheme of Figure 1 offers about 50% higher sensitivity, however.

Experimental observation of J coupling through H bonds is invaluable in NMR structure determination, where unambiguous identification of the H bond accepting carbonyl can be particularly difficult. Moreover, the presence of $^hJ_{\text{NC}'}$ coupling is of conceptual interest as it provides unequivocal evidence for the partial covalent character of even regular, weak H bonds in proteins. Data obtained on ubiquitin indicate that only H bonds for which the H $^{\text{N}}$ -O distance is less than 2.2 \AA give rise to an observable $^hJ_{\text{NC}'}$ connectivity, and $^hJ_{\text{NC}'}$ may prove to be a useful parameter for characterizing the strength of H bonds. In this case, $^hJ_{\text{NC}'}$ will be an ideal parameter for studying such interactions in the active sites of enzymes, where H bond strengths frequently are of key interest.

Acknowledgment. Work by G.C. is in partial fulfillment for the Ph.D. degree at the University of Maryland, College Park. We thank Frank Delaglio, John Marquardt, Marcel Ottiger, Marius Clore, and Angela Gronenborn for useful discussions. J.-S.H. is supported by a postdoctoral fellowship from the Cancer Research Institute.

Supporting Information Available: One table with the $^hJ_{\text{NC}'}$ couplings observed in ubiquitin; one pulse sequence for recording 2D H(N)CO spectra on larger, perdeuterated proteins, and one 2D H(N)CO spectrum, recorded with this pulse scheme for ubiquitin (PDF). This material is available free of charge via the Internet at <http://pubs.acs.org>.

(14) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1987**, *194*, 531-544.

(15) Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 12366-12371. Pervushin, K.; Wider, G.; Wüthrich, K. *J. Biomol. NMR* **1998**, *12*, 345-348.

(16) Kay, L. E.; Keifer, P.; Saareinen, T. *J. Am. Chem. Soc.* **1992**, *114*, 4, 10663-10665.