Correlation between $^{3h}J_{NC'}$ and Hydrogen Bond Length in Proteins

Gabriel Cornilescu, Benjamin E. Ramirez, M. Kirsten Frank, G. Marius Clore, Angela M. Gronenborn, and Ad Bax*

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

Received March 19, 1999. Revised Manuscript Received April 30, 1999

Abstract: Establishing a quantitative relationship between backbone−backbone hydrogen-bond (H-bond) length observed in protein crystal structures and recently observed $^{3h}J_{NC'}$ couplings across such bonds is limited by the coordinate precision of the X-ray structure. For an immunoglobulin binding domain of streptococcal protein G, very high-resolution X-ray structures are available. It is demonstrated that over the small range of N−O H-bond lengths (2.8−3.3 Å) for which $^{3h}J_{NC'}$ couplings are observable, the 32 measured $^{3h}J_{NC'}$ values can be fit to: $^{3h}J_{NC'} = -59000 \exp(-4R_{NO}) \pm 0.09$ Hz, or $R_{NO} = 2.75 - 0.25 \ln(-^{3h}J_{NC'}) \pm 0.06$ Å. Backbone amide to side-chain carboxyl hydrogen bonds were also investigated, and the measured $^{3h}J_{NC'}$ values tend to be smaller than expected from their crystallographically determined H-bond lengths. The sign of $^{3h}J_{NC'}$, determined from a zero-quantum/double-quantum experiment, is found to be the same as that of the $^{1}J_{NH}$ coupling, i.e., negative.

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] Recently, Dingley and Grzesiek demonstrated the presence of J couplings between the H bond donating and accepting ^{15}N nuclei in Watson−Crick RNA basepairs.[2] The size of these $^{2h}J_{NN}$ interactions, 6−7 Hz, is particularly surprising as the absolute values of J couplings to ^{15}N, transmitted through regular covalent bonds, tend to be quite small. For example, for peptides $^{1}J_{NC'} \approx -15$ Hz, $^{1}J_{NC\alpha} \approx -10$ Hz, and $|^{2}J_{NC\alpha}| \approx 5-9$ Hz.[3−5] Regular, three-bond J couplings from ^{13}C or ^{1}H to ^{15}N are invariably small (≤2 Hz),[6−8] and two-bond $^{2}J_{NC'}$ (≤1 Hz) and three-bond $^{3}J_{NN}$ values (≤0.3 Hz) are even smaller.[9,10] The presence of large $^{2h}J_{NN}$ interactions across H bonds was confirmed by Pervushin et al. in double stranded DNA.[11] They also discovered the presence of a smaller (2−4 Hz) J coupling between the imino hydrogen itself and the H bond accepting ^{15}N nucleus. These couplings provide direct experimental evidence for the presence of orbital overlap, even in regular, weak H bonds. Interestingly, a recent Compton X-ray scattering study of hexagonal ice also found evidence for a partial covalent character of hydrogen bonds.[13]

Very recently, the presence of $^{3h}J_{NC'}$ interactions across backbone−backbone H bonds in proteins was also reported.[14,15] Correlation of ^{13}C′ and ^{15}N chemical shifts through $^{3h}J_{NC'}$ is extremely useful for protein structure determination as it unambiguously identifies the carbonyl which accepts the H bond, and such $^{3h}J_{NC'}$ connectivities can be measured even in proteins as large as 30 kD.[16] Cordier and Grzesiek demonstrated the close correlation between $|^{3h}J_{NC'}|$ and the H^N chemical shift for the protein ubiquitin.[12] As the H^N shift is known to be correlated with the strength of the H bond,[17] this implies that $^{3h}J_{NC'}$ should also be a sensitive function of H-bond strength. However, a plot of $|^{3h}J_{NC'}|$ against the N−O distance,[14,15] obtained from ubiquitin's 1.8-Å crystal structure yields a rather poor correlation (data not shown). This can be understood, given the precision of the atomic coordinates in a 1.8-Å structure, which is commonly estimated to be 0.1−0.2 Å,[18] suggesting that the quality of the correlation is limited by the accuracy at which the atomic positions are known.

Here, we report data obtained for an immunoglobulin binding domain of streptococcal protein G, henceforth simply referred to as protein G. Resonance assignments and the NMR structure of this domain were reported previously,[19] and three high-resolution X-ray structures of highly homologous protein G domains are available, solved at resolutions of 1.92 Å (*R*-factor 20%, PDB code 1PGB),[20] and at 1.1 Å resolution (solved with isotropic *B*-factors, *R*-factor 19%, PDB code 1IGD;[21] also

(1) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer: New York, 1991.
(2) Dingley, A. J.; Grzesiek, S. *J. Am. Chem. Soc.* **1998**, *120*, 8293−8297.
(3) Bystrov, V. F. *Prog. Nucl. Magn. Reson. Spectrosc.* **1976**, *10*, 41−81.
(4) Delaglio, F.; Torchia, D. A.; Bax, A. *J. Biomol. NMR* **1991**, *1*, 439−446.
(5) Juranic, N.; Ilich, P. K.; Macura, S. *J. Am. Chem. Soc.* **1995**, *117*, 405−410.
(6) Vuister, G. W.; Wang A. C.; Bax A. *J. Am. Chem. Soc.* **1993**, *115*, 5334−5335.
(7) Wang, A. C.; Bax, A. *J. Am. Chem. Soc.* **1995**, *117*, 1810−1813.
(8) Hu, J.-S.; Bax, A. *J. Biomol. NMR.* **1997**, *9*, 323−328.
(9) Theis, K.; Dingley, A. J.; Hoffmann, A.; Omichinski, J. G.; Grzesiek, S. *J. Biomol. NMR* **1997**, *10*, 403−408.
(10) Lohr, F.; Rüterjans, H. *J. Magn. Reson.* **1998**, *132*, 130−137.
(11) Pervushin, K.; Ono, A.; Fernandez, C.; Szyperski, T.; Kainosho, M.; Wüthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14147−14151.
(12) Reference deleted in proof.
(13) Isaacs, E. D.; Shukla, A.; Platzman, P. M.; Hamann, D. R.; Barbiellini, B.; Tulk, C. A. *Phys. Rev. Lett.* **1999**, *82*, 600−603.
(14) Cordier, F.; Grzesiek, S. *J. Am. Chem. Soc.* **1999**, *121*, 1601−1602.
(15) Cornilescu, G.; Hu, J.-S.; Bax, A. *J. Am. Chem. Soc.* **1999**, *121*, 2949−2950.
(16) Wang, Y.-X.; Jacob, J.; Cordier, F.; Wingfield, P.; Stahl, S. J.; Lee-Huang, S.; Torchia, D. A.; Grzesiek, S.; Bax, A. *J. Biomol. NMR*, in press.
(17) Wagner, G.; Pardi, A.; Wüthrich, K. *J. Am. Chem. Soc.* **1983**, *105*, 5948−5949.
(18) Rhodes, G. *Crystallography made crystal clear*; Academic Press: San Diego, 1993; p 160.
(19) Gronenborn, A. M.; Filpula, D. R.; Essig, N. Z.; Achari, A.; Whitlow, M.; Wingfield, P. T.; Clore, G. M. *Science* **1991**, *253*, 657−661.

10.1021/ja9909024
Published on Web 06/22/1999

refined with anisotropic *B*-factors, PDB code 2IGD[22]). Through-H-bond *J* correlations were observed for all regular backbone−backbone H bonds present in the crystal structures, and measured $^{3h}J_{NC'}$ values are found to decrease exponentially with increasing H-bond length.

Experimental Section

Experiments were carried out at 25 °C on 600 and 750 MHz Bruker DMX-model spectrometers, equipped with triple resonance, 3-axis pulsed field gradient probeheads. The NMR sample contained 9 mg of protein G in 300 μL of H_2O/D_2O (90:10), 50 mM sodium phosphate, pH 5.6. Measurement of $^{3h}J_{NC'}$ were carried out at 600 MHz, using the quantitative *J* correlation scheme described previously.[14,15] Both the reference spectrum and the through-H-bond correlation spectrum were recorded at 600 MHz as $128^* \times 1024^*$ data matrixes, with acquisition times of 51 ms (t_1, $^{13}C'$) and 52 ms (t_2, $^1H^N$), with 16 scans per complex t_1 increment for the reference spectrum (1 h), and 600 scans (36 h) for the through-H-bond spectrum. Both reference and through-H-bond spectra were recorded using $^{15}N-\{^{13}C'\}$ dephasing delays of 133.2 ms, but for the reference spectrum the 180° $^{13}C'$ pulses were shifted to positions 16.6 ms prior to the midpoint of the dephasing delays, resulting in effective $J_{NC'}$ dephasing delays of 100 ms.[14] Measurements were repeated at pH 4.0, using 400 scans (24 h), but otherwise identical parameters.

The absolute value of $^{3h}J_{NC'}$ is obtained from the ratio of the intensity in a through-$^{3h}J_{NC'}$ 2D H(N)CO correlation spectrum, relative to that of a reference spectrum in which only the much larger $^1J_{NC'}$ gives rise to a correlation, as described previously.[14,15]

The sign of $^{3h}J_{NC'}$ was determined using the zero-quantum−double-quantum (ZQ/DQ) pulse scheme shown in Figure 1. The ZQ/DQ spectrum was recorded at 600 MHz, using 800 scans per complex t_1 increment (40 h), with acquisition times of 30 ms (t_1, $^{15}N-^{13}C'$ ZQ/DQ) and 52 ms (t_2, $^1H^N$), and repeated at 750 MHz using 800 scans per complex t_1 increment (40 h), with acquisition times of 25.6 ms (t_1, $^{15}N-^{13}C'$ ZQ/DQ) and 60 ms (t_2, $^1H^N$), both at pH 5.6. Peak picking in this spectrum was carried out both by means of contour averaging,[23] using the program PIPP,[24] and by polynomial interpolation, using the program NMRPipe.[25]

For determining the H-bond angles, hydrogens were added to the crystal structures with the program X-PLOR,[26] using idealized covalent geometry (i.e., the amide hydrogen in the $C'_{i-1}-N_i-C^\alpha_i$ plane, with the N−H vector on the line bisecting the $C'_{i-1}-N_i-C^\alpha_i$ angle, and a 1.02 Å N−H bond length), and reported values (and N−O distances) represent the averages over the three X-ray structures.

Results

NMR data were recorded on a protein sequence which differs by two substitutions (T1M, T2Q) from the 1PGB structure, and by eight substitutions (T1M, T2Q, V6I, I7L, K19E, E24A, A29V, V42E) and an N-terminal deletion from the 1IGD and 2IGD crystal structures. The backbone conformation of the 1PGB and 1IGD structures differ by 0.26 Å, confirming their close structural similarity. Measurement of backbone N−H and $C^\alpha-H^\alpha$ dipolar couplings for protein G, both in bicelle[27,28] and rod-shaped viral[29,30] media, indicates that the structure in solution is in excellent agreement with the structures observed in the crystalline state. Quantitatively, this agreement is expressed by a quality factor, *Q*,[31] defined as

$$Q = [\Sigma_{i=1,\cdots,N}(D_{ABi}^{meas} - D_{ABi}^{pred})^2/N]^{1/2}/D^{RMS} \quad (1)$$

where the summation extends over all *N* residues for which dipolar couplings between nuclei *A* and *B* (D_{ABi}) could be measured. D^{RMS} refers to the root-mean-square (rms) value of the measured D_{ABi} values, and D_{ABi}^{pred} is the value predicted for the D_{ABi} coupling when using a molecular alignment tensor obtained from best fitting all dipolar couplings simultaneously to the X-ray structure.[32] *Q* values (averaged for N−H and $C^\alpha-H^\alpha$) are 14% for the 1IGD structure, 16% for the 2IGD structure, and 18% for the 1PGB structure. Pairwise rms differences in N−O H-bond lengths were 0.05 Å (1IGD vs 2IGD), 0.12 Å (1IGD vs 1PGB), and 0.10 Å (2IGD vs 1PGB). Even though the measured dipolar couplings agree slightly less well with the 1PGB structure than with the 1IGD and 2IGD structures, the 1PGB structure was solved completely independently of the 1IGD and 2IGD structures and differs by only two rather than eight substitutions from the protein used in this study. Therefore, we use the average of the N−O distances observed in these

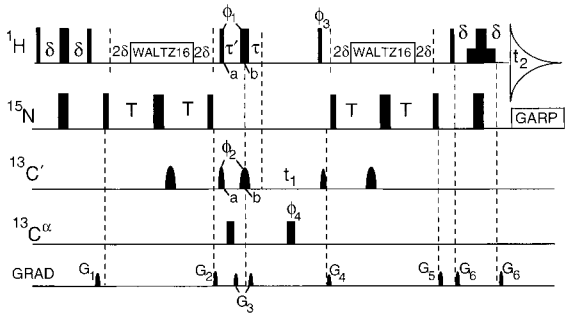

Figure 1. Pulse scheme of the 2D ZQ/DQ H(N)CO experiment, used to detect the sign of $^3J_{NC'}$. $^1H-^{13}C$ zero- and double-quantum coherence is generated at the beginning of the t_1 evolution period (time point *a*). At time point *b*, 180° ^1H and $^{13}C'$ pulses are applied to eliminate the effect of evolution during the finite duration of the ^1H, $^{13}C^\alpha$, and $^{13}C'$ pulses. A $^{13}C^\alpha$ Bloch-Siegert compensation pulse is applied at time point *a* in order to eliminate a phase error from the $^{13}C^\alpha$ 180° decoupling pulse, applied at the midpoint of t_1.[39] Narrow and wide pulses correspond to flip angles of 90° and 180°, respectively. All pulses phases are *x*, unless specified. $^{13}C'$ pulses have the shape of the center lobe of a (sin*x*)/*x* function, and durations of 150 μs. Shaped $^{13}C'$ pulses at the midpoint of the 2T periods are 180°; pulses bracketing the $^{13}C'$ 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, ^{15}N decoupling during t_2 used GARP modulation. Delay durations: $\delta = 2.65$ ms; $T = 66.6$ ms; $\tau' = 500$ μs; $\tau = \tau' - PW_{180}$, where PW_{180} is the duration of the $180°_{\phi 4}$ pulse (applied at 56 ppm), the duration of which is adjusted to yield a null at the $^{13}C'$ frequency. The midpoints of the 90° and 180° pulses, applied with phase ϕ_1 and ϕ_3, coincide with the midpoints of the shaped $^{13}C'$ pulses. Phase cycling: $\phi_1 = x,-x$; $\phi_2 = 4(x),4(-x)$; $\phi_3 = 2(x),2(-x)$; $\phi_4 = 8(x)$, $8(-x)$; Receiver = $x,2(-x),x,-x,2(x),-x$. Gradients are sine bell shaped, with peak amplitudes of 30 G/cm, and durations G1,2,3,4,5,6 = 1.1, 0.5, 0.28, 0.6, 0.42, 0.4 ms, and directions *z*, *x*, *z*, *y*, *xy*, and *z*, respectively.

(20) Gallagher, T., Alexander, P.; Bryan, P.; Gilliland, G. L. *Biochemistry* **1995**, *33*, 4721−4729.
(21) Derrick, J. P.; Wigley, D. B. *J. Mol. Biol.* **1994**, *243*, 906−9181.
(22) Butterworth: S.; Lamzin, V. S.; Wigley, D. B.; Derrick, J. P.; Wilson, K. S. *Protein Brookhaven Databank entry* 2IGD.
(23) Wang, A. C.; Bax, A. *J. Am. Chem. Soc.* **1996**, *118*, 2483−2494.
(24) Garrett, D. S.; Powers, R.; Gronenborn, A. M.; Clore, G. M. *J. Magn. Reson.* **1991**, *95*, 214−220.
(25) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR.* **1995**, *6*, 277−293.
(26) Brünger, A. T. *X-PLOR, A system for X-ray crystallography and NMR*; Yale University Press: New Haven, 1987.
(27) Sanders, C. R.; Schwonek, J. P. *Biochemistry* **1992**, *31*, 8898−8905.
(28) Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111−1114.
(29) Clore, G. M.; Starich, M. R.; Gronenborn, A. M. *J. Am. Chem. Soc.* **1998**, *120*, 10571−10572.
(30) Hansen, M. R.; Rance, M.; Pardi, A. *J. Am. Chem. Soc.* **1998**, *120*, 11210−11211.
(31) Cornilescu, G. C.; Marquardt, J. L.; Ottiger, M.; Bax, A. *J. Am. Chem. Soc.* **1998**, *120*, 6836−6837.
(32) Tjandra, N.; Grzesiek, S.; Bax, A. *J. Am. Chem. Soc.* **1996**, *118*, 6264−6272.

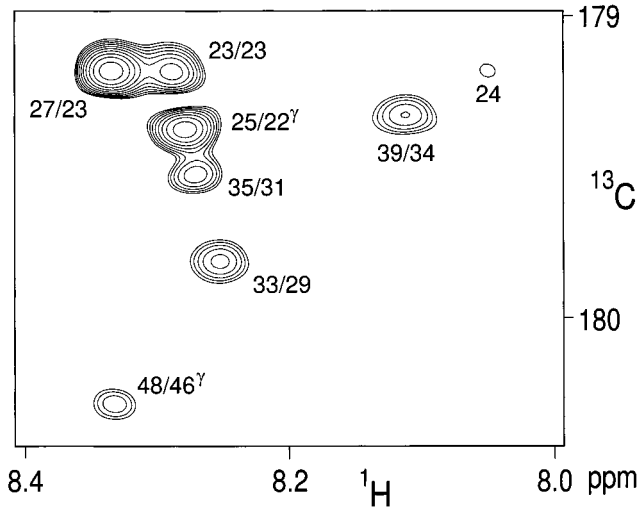

Figure 2. Small section of the 600 MHz 2D H(N)CO spectra of U-^{13}C/^{15}N protein G, recorded with de/rephasing delays of 133 ms, optimized for detection of small couplings and suppression of 15 Hz ($^1J_{NC'}$) interactions. Correlations are marked by the residue number of the amide, followed by the number of the residue containing the carbonyl *J*-coupled to this amide. Correlations to Asp side-chain carboxylate carbons are marked with superscript γ. The one-bond $^1J_{NC'}$ connectivity between the amide of Ala^{24} and the ^{13}C′ of Ala^{23} is labeled by the residue number of the amide. The spectrum results from a $128^* \times 1024^*$ data matrixes, with acquisition times of 51 ms (t_1) and 60 ms (t_2). The spectrum was recorded with 400 scans per complex increment (total time 24 h).

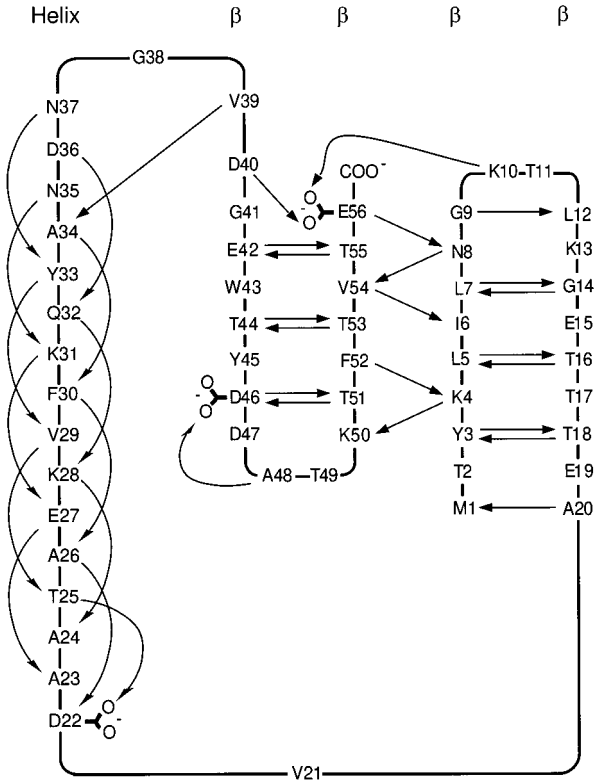

Figure 3. Topology of protein G, with observed $^{3h}J_{NC'}$ correlations marked by thin arrows.

three structures when correlating the H-bond lengths with the observed *J* couplings. This also yields a somewhat tighter correlation between $^{3h}J_{NC'}$ and H-bond length, compared to using any of the individual structures (data not shown).

Magnitude of $^{3h}J_{NC'}$. Figure 2 shows a small region of the $^{3h}J_{NC'}$ correlation spectrum, comprising correlations that mainly involve helical residues. The ^{15}N−{^{13}C′} dephasing delay is set to $2/^1J_{NC'}$, assuming $^1J_{NC'} = 15$ Hz.[14,15] Next to the through-H-bond connectivities of interest, labeled by the residue numbers of both amide proton and carbonyl, a weak sequential $^1J_{NC'}$ connectivity between Ala^{24}−H^N and Ala^{23}−C′ (labeled 24) results from the fact that $^1J_{NC'}$ for this peptide bond deviates from 15 Hz. The complete spectrum includes backbone−backbone correlations for 36 H bonds, all showing N−O H-bond lengths less than 3.3 Å in the crystal structures. These interactions are schematically shown by arrows in the backbone diagram of protein G (Figure 3). Correlations to Glu and Asp side-chain carboxyl carbons are also observed, but one of the five such interactions observed in the X-ray structures is obscured by resonance overlap, and its presence or absence in solution could not be established. In addition to the through-H-bond and the one-bond sequential connectivities, several intraresidue connectivities are also observed (e.g., Ala^{23}−H^N to Ala^{23}−C′ in Figure 2), resulting from nonzero values of $^2J_{NC'}$. Table 1 summarizes the measured $^{3h}J_{NC'}$ values, together with the corresponding N−O distances, and the N−H··O and H··O=C angles.

Sign of $^{3h}J_{NC'}$. Quantitative *J* correlation yields the absolute value of the $^{3h}J_{NC'}$ values only. The sign of these couplings can be measured by zero-quantum/double-quantum (ZQ/DQ) spectroscopy, and the pulse scheme used here is shown in Figure 1. A detailed description of this general type of experiment has been presented by Rexroth et al.[33] and Otting et al.,[34] and will not be repeated here. The present pulse scheme generates ^{13}C′−^{1}H^N zero- and double-quantum coherence at time point *a* in the pulse sequence. To ensure that the phase of the signals in the final 2D spectrum is purely absorptive in the F_1 dimension, 180° ^{1}H and ^{13}C′ refocusing pulses have been inserted at the start of the t_1 evolution period. These pulses compensate for evolution during the finite ^{1}H and ^{13}C′ pulse widths and during the ^{13}C$^\alpha$ decoupling pulse. The detected signal is modulated by the ^{13}C′−^{1}H^N zero- and double-quantum frequencies. The zero-quantum coherence evolves at $\delta_H - \delta_{C'} \pm (^1J_{NH} - {}^{3h}J_{NC'})/2$, whereas the double-quantum coherence evolves at $\delta_H + \delta_{C'} \pm (^1J_{NH} + {}^{3h}J_{NC'})/2$. The difference in splitting, $2\ {}^{3h}J_{NC'}$, yields both the sign and the magnitude of $^{3h}J_{NC'}$.

Figure 4 shows a small region of the ZQ/DQ spectrum of protein G, recorded at 750 MHz. As the intensity of the through-H-bond correlation in such a spectrum is spread over four resonance lines, the experiment is inherently four times lower in signal-to-noise compared to the H(N)CO spectrum of Figure 2. Because $^{3h}J_{NC'}$ is derived from the very small ($\leq$1.6 Hz) difference in splitting of the zero- and double-quantum signals, reliable measurement of the sign of $^{3h}J_{NC'}$ requires high signal-to-noise. The spectrum shown corresponds to a 40-h experiment at 750 MHz. This spectrum was recorded following a measurement at 600 MHz, which resulted in insufficient signal-to-noise to determine the values of $^{3h}J_{NC'}$ at sufficient accuracy. The inset in Figure 4 shows the signal-to-noise level for the most-downfield shifted amide signal of Phe^{52}. Values for $^{3h}J_{NC'}$ were measured from this 2D spectrum for the nine H bonds with $|^{3h}J_{NC'}| \geq 0.5$ Hz which yield nonoverlapped correlations in the ZQ/DQ spectrum. Peak picking based on contour averaging yields an average $^{3h}J_{NC'}$ value of -0.47 ± 0.9 Hz, whereas peak picking based on polynomial interpolation yields an average $^{3h}J_{NC'}$ value of -0.42 ± 0.7 Hz. The uncertainty in the mean is

(33) Rexroth, A.; Schmidt, P.; Szalma, S.; Geppert, T.; Schwalbe, H.; Griesinger, C. *J. Am. Chem. Soc.* **1995**, *117*, 10389−10390.

(34) Otting, G.; Messerle, B. A.; Soler, L. P. *J. Am. Chem. Soc.* **1997**, *119*, 5425−5434.

Table 1. $^{3h}J_{NC'}$ Couplings in Protein G, and Corresponding H-Bond Length and Angles[a]

H^N	C′	$^{3h}J_{NC'}$ (Hz)	R_{NO} (Å)	α(NHO) (deg)	α(HOC) (deg)
Y3	T18	−0.51	2.88	170	172
K4	K50	−0.42	3.04	167	155
L5	T16	−0.70	2.88	167	154
L7	G14	−0.68	2.85	174	152
N8	V54	−0.70	2.87	163	170
G9	L12	−0.33	2.94	135	159
G14	L7	−0.24	3.10	158	144
T16	L5	−0.38	3.02	157	145
T18	Y3	−0.41	2.99	156	150
A20	M1	−0.51	2.97	161	154
A26	D22	−0.18	3.24	162	149
E27	A23	−0.54	2.90	157	148
K28	A24	−0.13	3.14	154	140
V29	T25	−0.21	3.17	162	147
F30	A26	−0.64	2.91	161	154
K31	E27	−0.72	2.84	161	152
Q32	K28	−0.19	3.06	159	145
Y33	V29	−0.27	2.98	161	144
A34	F30	−0.49	2.91	157	148
N35	K31	−0.31	2.93	163	147
D36	Q32	−0.60	2.82	155	151
N37	Y33	−0.19	3.16	137	137
G38	N35	>−0.1	3.17	159	105
V39	A34	−0.34	2.90	158	149
E42	T55	−0.43	2.95	155	152
T44	T53	−0.53	2.95	158	146
D46	T51	−0.36	2.95	165	153
T49	D46	>−0.1	3.34	145	107
T51	D46	−0.22	3.26	157	145
F52	K4	−0.70	2.80	171	163
T53	T44	−0.61	2.92	158	141
V54	I6	−0.39	3.01	164	168
T55	E42	−0.51	2.99	162	152
E56	N8	−0.33	3.03	165	148
K10	E56C^{δ}	−0.38	2.76	173	147
T25	D22C^{γ}	−0.47	3.02	147	144
D40	E56C^{δ}	−0.10	2.90	167	121
A48	D46C^{γ}	−0.19	2.76	136	147

[a] $^{3h}J_{NC'}$ were measured at pH 5.6, 25 °C. R_{NO} and α values are averaged over crystal structures at pH 4.0 (1PGB) and 4.8 (1IGD and 2IGD).

$\sqrt{9}$ times smaller, i.e., less than 0.3 Hz. Since the average absolute value of these nine couplings is 0.6 Hz, these results indicate that the sign of $^{3h}J_{NC'}$ is negative. The measurement at 600 MHz yielded lower signal-to-noise and permitted measurement of only seven splittings, with an average value of −0.9 Hz and an uncertainty in the mean value of ±1.2 Hz. Considering that the absolute value of $^{3h}J_{NC'}$ decreases monotonically with increasing H-bond length (see below), all $^{3h}J_{NC'}$ couplings must have the same sign, i.e., they must be negative.

Sign of $^{2}J_{NC'}$. The ZQ/DQ spectrum indicates that intraresidue $^{2}J_{NC'}$ and sequential $^{1}J_{NC'}$ values are also negative, as are the intraresidue couplings between the backbone ^{15}N and side-chain $^{13}C^{\gamma}$ in Asp and Asn residues. In all cases, the negative sign of the ^{15}N magnetogyric ratio is taken into account.

Dependence of $^{3h}J_{NC'}$ on R_{NO}. Figure 5 shows the correlation between the 32 measured backbone−backbone $^{3h}J_{NC'}$ values and H-bond length, averaged over the 1IGD, 2IGD, and 1PGB structures. In the limit of weak H bonding, electron orbital overlap decreases exponentially with increasing H-bond length. Best fitting of the data to an exponentially decaying function (Figure 5) yields

$$^{3h}J_{NC'} = -59000 \exp(-4R_{NO}) \pm 0.09 \text{ Hz}. \qquad (2a)$$

The inverse of this equation yields

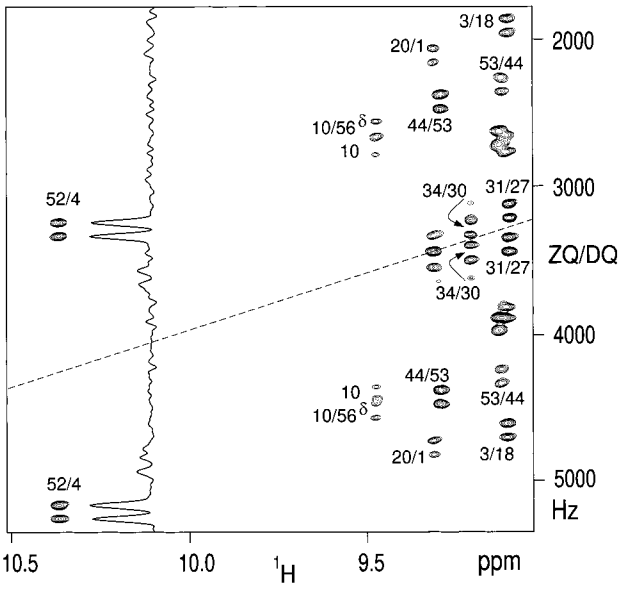

Figure 4. Section of the 750 MHz 2D ($^{13}C'-^{1}H^{N}$) ZQ/DQ HNCO spectrum of protein G, recorded with the pulse scheme of Figure 2. The dashed line corresponds to the proton offset relative to the ^{1}H carrier, and a zero-quantum resonance is displaced to lower frequency if the $^{13}C'$ resonates downfield of the $^{13}C'$ carrier. Splittings in the ZQ/DQ dimension correspond to $^{1}J_{NH} - {}^{3h}J_{NC'}$ (ZQ) and $^{1}J_{NH} + {}^{3h}J_{NC'}$ (DQ). The spectrum was recorded as a 128* × 768* data matrix, with 800 scans per complex t_1 increment (total time 40 h).

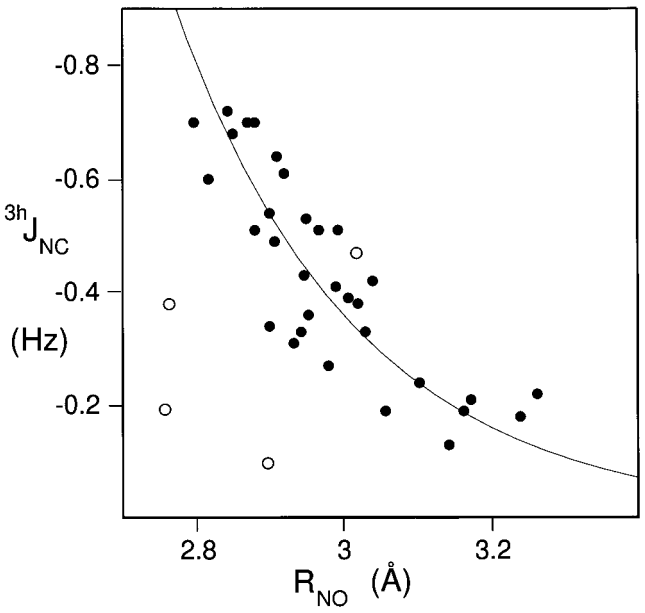

Figure 5. Correlation between the observed $^{3h}J_{NC'}$ values and H-bond lengths in protein G, averaged over three crystal structures (1IGD, 2IGD, and 1PGB), to which protons were added with X-PLOR, assuming $r_{NH} = 1.02$ Å. The curve is the best fit exponential, −59000 exp(−4R_{NO}) Hz, where R_{NO} is the H bond length in Å. Open circles correspond to side-chain carboxyl groups, H-bonded to backbone amides, and were not used in the fit. The crystallographic temperature factors for the H bond accepting atoms are nearly 2-fold larger than for most backbone carbonyls.

$$R_{NO} = 2.75 - 0.25 \ln(-{}^{3h}J_{NC'}) \pm 0.06 \text{ Å}. \qquad (2b)$$

The rms difference between R_{NO} values derived from eq 2b and the crystallographic data equals 0.06 Å. This rms spread presumably overestimates the true uncertainty in R_{NO} derived from eq 2b, as it also includes the effect of the uncertainty in

the atomic position in the X-ray structure and possible real differences between the R_{NO} distances in solution and in the crystalline state. Also, orbital overlap and thereby the $^{3h}J_{NC'}$ coupling is expected to be sensitive to the N−H··O and H··O═C angles. As can be seen from Table 1, for the backbone−backbone H bonds in protein G, these angles are confined to a relatively narrow range.

There are only two backbone−backbone H bonds (G38−N35 and T49−D46) which do not yield observable $^{3h}J_{NC'}$ connectivities, indicating $|^{3h}J_{NC'}| \leq 0.1$ Hz. Both are among the longest H bonds in protein G (R_{NO} values of 3.17 and 3.34 Å), but they also have rather sharp H··O═C angles of 105 and 107°, which fall far outside the range observed for those of the other H bonds. On the basis of eq 2a, $^{3h}J_{NC'}$ couplings of 0.18 and 0.09 Hz are predicted for these H bonds. It is therefore not possible to determine unequivocally whether the absence of these $^{3h}J_{NC'}$ connectivities results from the unusual H··O═C angles. It is interesting to note that ab initio calculations also do not show a significant dependence of the bond energy on the H··O═C and N−H··O angles for intermolecular amide H bonds.[35] This is in contrast to a strong angular dependence calculated for strong, low-barrier H bonds.[36]

$^{3h}J_{NC'}$ connectivity is observed for four of the five backbone amide to side chain carboxyl groups observed in the crystal structure (Table 1). The absence or presence of a correlation between Tyr^{45}−N and Asp^{47}−C^{γ}, corresponding to a H bond in the X-ray structures, could not be determined since the corresponding cross-peak is obscured by overlap. At 2.76 Å, the N−O distances for two of the observed backbone to side-chain correlations correspond to the shortest H bonds in the crystal structures, but the $^{3h}J_{NC'}$ couplings are relatively small. The corresponding points (marked "o") fall well off the curve in Figure 5. This could either be due to the different nature of the H bond accepting group or be caused by structural fluctuations, where the H bond is present only part of the time. Crystallographic temperature factors for these side-chain carboxylates in the 1IGD structure are about 1.7-fold higher than for the backbone carbonyls, suggesting that structural fluctuations may cause the decrease in $^{3h}J_{NC'}$ for these interactions. It is also interesting to note that the NMR measurements are conducted at pH 5.6, whereas the crystal structures were determined at pH 4.0 (1PGB) and 4.8 (1IGD). When the pH of the NMR sample is lowered to 4.0, the $^{3h}J_{NC}$ for the Lys^{10}−H^N to Glu^{56}−$O^{\epsilon 1}$ decreases from 0.38 to 0.26 Hz. This indicates a weakening of this H bond, which is accompanied by a 0.6 ppm upfield shift of Lys^{10}−H^N, and results from partial protonation of the side-chain carboxylate. Similarly, the Asp^{40}−H^N, H-bonded to Glu^{56}−$O^{\epsilon 2}$ in the X-ray structures, shifts upfield by 0.35 ppm, suggesting that it is also weaker at lower pH. At the lower pH, the $^{3h}J_{NC'}$ connectivity for this H bond falls below the detection threshold of 0.15 Hz, but none of the backbone−backbone $^{3h}J_{NC'}$ couplings differ significantly from the values in Table 1.

Concluding Remarks

Except for one correlation partially obscured by overlap (Ile^6−Phe^{52}), $^{3h}J_{NC'}$ couplings were measured for all regular backbone−backbone H bonds with an N−O H-bond length less than 3.3 Å in the three crystal structures of protein G. The close correlation between $^{3h}J_{NC'}$ and N−O internuclear distance (eq 2b) should be useful as a tight restraint during protein structure determination. The empirical correlation of eq 2 also presents a stringent test for evaluating the results of ab initio calculations for these couplings at different levels of theory.

It is expected that eq 2 only applies to the limit of weak H bonding, i.e., to cases where the pK values of the H bond donating and accepting groups are quite different, and the donor−acceptor distance is relatively large. For strong, low-barrier H bonds[37] with extensive orbital overlap, the correlation between $^{3h}J_{NC'}$ and H-bond length is expected to be considerably more complex.[38]

Acknowledgment. Work by G.C. and M.K.F. is in partial fulfillment for the Ph.D. degree at the University of Maryland, College Park. B.R. is the recipient of a NSF postdoctoral fellowship. We thank Attila Szabo for useful discussions.

JA9909024

(35) Adalsteinsson, H.; Maulitz, A. H.; Bruice, T. C. *J. Am. Chem. Soc.* **1996**, *118*, 7689−7693.

(36) Smallwood, C. J.; McAllister, M. A. *J. Am. Chem. Soc.* **1997**, *119*, 11277−11281.

(37) Cleland, W. W.; Kreevoy, M. M. *Science*, **1994**, *264*, 1887−1890.

(38) Globulev, N. S.; Shenderovich, I. G.; Smirnov, S. N.; Denisov, G. S.; Limbach, H. H. *Chem. Eur. J.* **1999**, *5*, 492−497; Shenderovich, I. G.; Smirnov, S. N.; Denisov, G. S.; Gindin, V. A.; Globulev, N. S.; Dunger, A. Rebke, R.; Kirpekar, S.; Malkina, O. L.; Limbach, H. H. *Ber. Bunsen-Ges. Phys. Chem.* **1998**, *102*, 422−428.

(39) Vuister, G. W.; Bax, A. *J. Magn. Reson.* **1992**, *98*, 428−435.