$H \cdots N$ hydrogen bond lengths in double stranded DNA from internucleotide dipolar couplings

Zhengrong Wu^a, Akira Ono^b, Masatsune Kainosho^b & Ad Bax^a

^aLaboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520, U.S.A. ^bTokyo Metropolitan University, 1-1 Minami-ohsawa, Hachioji, Tokyo 192-0397, Japan

Received 8 January 2001; Accepted 13 February 2001

Key words: Dickerson dodecamer, dipolar coupling, E.COSY, hydrogen bond, liquid crystal, Watson-Crick

Abstract

The ratio of the internucleotide dipolar coupling and the corresponding one-bond imino ¹⁵N-¹H dipolar coupling provides a measure for the N···H/H-N distance ratio. Measurements were carried out for a dodecamer, $d(CGCGAATTCGCG)_2$, in which a C-G and an A-T basepair were uniformly enriched in ¹⁵N. When assuming H-bonds to be perfectly linear, dipolar data indicate time-averaged hydrogen bond lengths of 1.80 ± 0.03 Å for A-T and 1.86 ± 0.02 Å for C-G. When using H-bond orientations from high resolution X-ray data, H-bond lengths are about 0.1 Å shorter.

It has long been recognized that hydrogen bonds (Hbonds) are critical for stabilizing biomolecular structure. However, most experimental information related to individual H-bond interactions in macromolecules has been rather indirect. NMR has long played an important role in the study of H-bonding and is one of the few techniques that allow monitoring of individual Hbonds. In particular, the chemical shift of the hydrogen involved in the H-bond is known to be very sensitive to the H-bond strength, and effects of H-bonding on the one-bond J coupling between the donor atom and the hydrogen have also been reported (Becker, 1996). Recently, intermolecular J couplings across hydrogen bonds have also been observed (Dingley and Grzesiek, 1998; Pervushin et al., 1998a; Golubev et al., 1999; Wang et al., 1999). Both experimental results and theoretical calculations indicate that the magnitude of this through H-bond J coupling is closely related to the H-bond length (Cornilescu et al., 1999; Dingley et al., 1999; Benedict et al., 2000). Much of this work and related earlier observations have been reviewed by Grzesiek et al. (2001).

The present communication focuses on the measurement of direct dipolar couplings between the Hbond hydrogen and donor and acceptor ¹⁵N nuclei in a DNA dodecamer. Such dipolar couplings can be measured readily in macromolecules by dissolving them in a dilute aqueous liquid crystalline medium (Tjandra and Bax, 1997). If the H-bond were perfectly linear, the ratio of the two dipolar couplings simply scales as $\langle r_{HA}^{-3} \rangle / \langle r_{HD}^{-3} \rangle$, where r_{HA} is the distance between the hydrogen and the acceptor ^{15}N , and r_{HD} is the covalent bond length to the donating ¹⁵N. As we will show, the residual dipolar couplings can be measured at very high accuracy. However, determination of the value of $\langle r_{HA}^{-3} \rangle^{-1/3}$ is limited by the uncertainty in both the covalent imino N-H bond length and the deviation from linearity of the hydrogen bond. Nevertheless, our data are consistent with a shorter H-bond length for A-T basepairs relative to G-C, and yields values that are in good agreement with those observed in crystal structures.

The pulse scheme for measuring the across H-bond ${}^{1}\text{H}^{N} \dots {}^{15}\text{N}$ coupling is sketched in Figure 1. It is very similar to the sequences previously used for measuring the ${}^{15}\text{N}{}^{-15}\text{N}$ and ${}^{1}\text{H}^{N} \dots {}^{15}\text{N}$ J couplings in nucleic acids (Dingley and Grzesiek, 1998; Pervushin et al., 1998a), which in turn are very similar to the LRCC pulse scheme for measurement of long range ${}^{13}\text{C}{}^{-13}\text{C}$ J couplings (Bax et al., 1992). However, as will

Table 1. ¹⁵N-¹H couplings in d(CGCGAATTCGCG)₂^a

	$^{1}J_{NH}$	$^{1h}J_{\mathrm{NH}}$	$^{1}D_{NH}$	$^{1h}D_{\mathrm{NH}}$
G4-C21 ^b	-87.60 ± 0.05	$\begin{array}{c} 2.32\pm0.07\\ 1.84\pm0.07\end{array}$	-17.53 ± 0.07	-3.1 ± 0.1
A5-T20 ^b	-86.54 ± 0.05		-15.54 ± 0.07	-3.0 ± 0.1

 $^{\rm a}$ Couplings measured at 600 MHz, 25 °C, pH 6.8, 50 mM KCl, in the presence and absence of 28 mg/ml Pf1.

^b G4-C21 are indistinguishable from C9-G16 and the same applies for A5-T20 and T8-A17.



Figure 1. HNN pulse scheme for measurement of ^{1h}J_{NH}. In scheme A, the open composite $(90_x - 180_y - 90_x) 180^\circ$ pulse and adjacent selective ¹H water 180° flip-back pulses are not applied, whereas they are applied in scheme B. Narrow and wide pulses correspond to flip angles of 90° and 180° , respectively. Low amplitude corresponds to selective 90° (narrow) and 180° (wide) H₂O pulses. The ¹H carrier is switched to 13 ppm from just prior to the first ¹H composite 180° pulse until just after the second composite 180° in scheme B, and is at the H₂O frequency for all of the remainder. The shaped ¹H pulse in the middle of t1 evolution is of the G3 type (Emsley and Bodenhausen, 1989) and has a duration of 2 ms with its excitation profile centered at 7.5 ppm, and a >90% inversion bandwidth of 1.5 kHz. The four shaped ¹⁵N pulses are of the hyperbolic secant type (Silver et al., 1984) with a width of 1.1 ms each (for a near-perfect inversion bandwidth of ± 2.5 kHz). The ¹⁵N carrier is switched from 155 ppm (imino region) to 180 ppm at time point a, and back to 155 ppm at time point b. Delay durations: $\tau = 2.7$ ms; $\Delta = 34$ ms; $\delta = 1.2$ ms. Unless indicated, pulses are applied along the x axis. Phase cycling: $\phi_1 = 4(y, -x), 4(-y, x); \phi_2 = 2(-x, -y), 2(x, y)$ (scheme A), $\phi_2 = 2(225^\circ, 315^\circ), 2(45^\circ, 135^\circ)$ (scheme B); $\phi_3 = -x, -y, x, y$ (scheme A), $\phi_3 = 225^\circ, 315^\circ, 45^\circ, 135^\circ$ (scheme B); Receiver = 4(x,y),4(-x,-y). Quadrature detection in the t_1 dimension is accomplished by incrementing ϕ_1 and ϕ_2 in the regular States-TPPI manner. All gradients are sine-bell shaped, with durations of $G_{0,1,2,3,4} = 1, 0.5, 1, 0.5, 0.5$ ms, and directions $G_{0,1,2,3,4} = (xy), x, x, x, z.$

be described below, measurement of the actual couplings relies on the heteronuclear E.COSY principle (Montelione and Wagner, 1989).

The present scheme uses a ST2P-TROSY element (Pervushin et al., 1998a) to selectively transfer back only the slowly relaxing ¹⁵N doublet component, i.e., the component corresponding to ¹H^N in the $|\beta\rangle$ spin state. The experiment is executed twice: (A) without the ¹H 180° pulses immediately preceding and following the t₁ evolution period, and (B) with these

pulses included. As described by Dingley and Grzesiek, magnetization is transferred from ¹H^N to the H-bond-accepting ¹⁵N nucleus by the first half of the pulse sequence (preceding t_1), then evolves as $A^{-15}N_1$ magnetization for A-T basepairs (or C-15N3 for G-C), before being transferred back to ¹H^N by the second half of the sequence. In the absence of the open 180° ¹H pulses (experiment A), only $A^{-15}N_1$ (and $C^{-15}N_3$) t_1 evolution with ¹H^N in the $|\beta\rangle$ spin state contributes to the observed signal, because only the ${}^{1}H^{N} = |\beta\rangle$ ¹⁵N doublet component is converted to ¹H transverse magnetization by the TROSY element. In experiment B, the ¹H^N spin state is flipped just prior to and just following the t_1 evolution period, so only A-¹⁵N₁ (and C-¹⁵N₃) t₁ evolution with ¹H^N in the $|\alpha\rangle$ spin state is observed. In principle, the 180° pulse preceding the t_1 evolution period could be omitted, but the ¹⁵N doublet component with ${}^{1}H^{N} = |\alpha\rangle$ is weaker due to faster transverse relaxation and an opposite sign contribution from the ¹⁵N Boltzmann magnetization (Pervushin et al., 1998b). Therefore it is more efficient to switch the two doublet components by application of the ¹H 180° pulse. When this 180° pulse is applied $1/(8J_{NH})$ prior to the $90^{\circ}_{\phi 2}$ pulse, the undesired ^{15}N doublet component $({}^{1}H^{N} = |\beta\rangle)$ is converted into two-spin ${}^{15}N$ - ${}^{15}N$ coherence by the $90^{\circ}_{\phi2}$ pulse, in an $S^{3}E$ manner (Meissner et al., 1997), and is removed by ϕ_2 phase cycling. This ensures that all observed evolution during t₁ corresponds to the ¹H^N = $|\alpha\rangle$ spin state, even if spin state inversion by the composite 180° ¹H pulse following t1 were imperfect. In order to narrow the F1 line width, ³J_{NH} coupling to the amino protons and ${}^{2}J_{NH}$ to A-C2H are decoupled by the application of a 2-ms selective G3 pulse at the midpoint of t_1 . This pulse is only applied for t₁ durations that exceed this pulse width.

The measurements are demonstrated for a sample of the dodecamer d(CGC<u>GAATTC</u>GCG)₂, where the underlined nucleotides are labeled uniformly with ¹⁵N, with a level of enrichment varying between 79

Table 2. Orientation of hydrogen bond vectors in the frame of the principal alignment tensor in three experimentally determined structures of $d(CGCGAATTCGCG)_2^a$

Stru	icture	Basepair	θ	φ	θ'	ϕ'
1BI	NA ^b	A5-T20	102	55	76	122
1B1	NA	T8-A17	84	132	108	38
1B1	NA	G4-C21	95	93	106	81
1B1	NA	C9-G16	87	81	77	95
355	D ^c	A5-T20	104	56	73	130
355	D	T8-A17	80	131	105	46
355	D	G4-C21	92	95	102	83
355	D	C9-G16	86	82	77	101
1D	UF ^d	T8-A17	85	131	105	39
1D	UF	G4-C21	88	95	102	77

^aThe angles θ and ϕ are defined in Figure 3. Angles in degrees after best fitting the alignment tensor to the experimental structures, using a set of ${}^{13}C{}^{-1}H$ and ${}^{15}N{}^{-1}H$ dipolar couplings measured for a sample of d(<u>CGCGAATTCGCG</u>)₂ in Pf1, where underlined nucleotides are enriched uniformly in ${}^{13}C$ and ${}^{15}N$. ^bFrom Dickerson and Drew (1981).

^cFrom Shui et al. (1998).

^dFrom Tjandra et al. (2000). Because of the symmetric input restraints used when calculating this structure, A5-T20 and A8-T17 basepairs are indistinguishable, and the same for G4-C21 and C9-G16. Note that θ' in this NMR structure results primarily from H-bond model restraints, as no experimental interstrand NOEs or dipolar couplings were used to derive this structure.

Table 3. Hydrogen bond lengths derived from dipolar couplings, assuming orientations as defined in Table 2

	Linear	1BNA	355D	1DUF
A-T G-C	$\begin{array}{c} 1.80\pm0.02\\ 1.86\pm0.02\end{array}$	1.78 ^a /1.62 ^b 1.72 ^c /1.76 ^d	1.73/1.73 1.77/1.77	1.68 1.68
^a A5-T20. ^b T8-A17.		^c G4-C21. ^d C9-G16.		

and 88% for the four labeled nucleotides. Its preparation has been described previously (Kojima et al., 2000). Two 1-mM NMR samples were used, each dissolved in 300 μ l H₂O/D₂O (93/7%), both containing 10 mM phosphate buffer (pH 6.8) and 50 mM KCl, and one of the samples also containing 28 mg/ml Pf1 phage (ASLA http://130.237.129.141//asla/asla-phage.htm).

Figure 2 shows small regions of the four HNN spectra. The spectra are recorded in the absence (A,B) and presence (C,D) of 28 mg/ml Pf1 phage, using the scheme of Figure 1, selecting (A,C) the ${}^{1}\text{H}^{N} = |\beta\rangle$ spin state (no 180° ${}^{1}\text{H}$ pulses) and (B,D) the ${}^{1}\text{H}^{N} = |\alpha\rangle$ spin state. The spectra show a larger ${}^{1}\text{H}^{N}$ linewidth in the presence of phage, which is caused by unresolved dipolar couplings, primarily to the nearby



Figure 2. T-8 imino region of the HNN correlation spectra recorded at 600 MHz, 25 °C, for 1 mM d(CGC<u>GAATTC</u>GCG)₂ in (A, B) the isotropic phase and (C,D) the liquid crystalline phase. Spectra A and C are recorded with scheme A of Figure 1, whereas B and D are collected using scheme B. The vertical displacement between the two T8-¹⁵N₃-{¹H} doublet components in A and B corresponds to ¹J_{NH} and to ¹J_{NH} +¹ D_{NH} in spectra C and D. Similarly, the displacement of the A5-¹⁵N₁ resonance yields ^{1h}J_{NH} (A,B) and ^{1h}J_{NH} + ^{1h}D_{NH}(C,D). Note that the sign of ^{1h}D_{NH} is opposite to that of ¹J_{NH}, whereas the sign of ^{1h}D_{NH} is the same as that of ¹J_{NH}. The spectra were acquired as 700* × 1024* data matrices, corresponding to acquisition times of 52.5 ms (t₁) and 68.2 ms (t₂), and zero-filled to 4096* × 2048* prior to Fourier transformation. Spectra A and B (and also C and D) were acquired in an interleaved mode, with a measuring time of 45 h per pair of spectra.

protons on the adenine base. The upfield ¹⁵N region of the spectra shows the downfield (A,C) and upfield (B,D) components of the T7 imino ¹⁵N₃-{¹H} doublet. Even at the thermal noise level (about six times lower than the lowest contour level shown), no trace of the upfield ¹⁵N₃-¹H component is present in spectra A and C, and vice versa, no trace of the downfield component is present in B and D. This demonstrates that the ¹H^N = $|\beta\rangle$ and ¹H^N = $|\alpha\rangle$ spectra are very well separated. This is critical when measuring the relative displacement of the downfield A-¹⁵N₁resonance, which corresponds to the through-H-bond J coupling, ^{1h}J_{NH}, in the absence of phage, and to ^{1h}J_{NH}+^{1h}D_{NH} in the presence of phage.

The relative vertical displacements of the $A^{-15}N_1$ resonance (at 202 ppm) in the absence (Figure 2A,B) and presence of phage (Figure 2C,D) are of opposite sign, indicating that ${}^{1h}J_{NH}$ and ${}^{1h}D_{NH}$ are of opposite sign. Clearly, ${}^{1h}D_{NH}$ has the same sign as the one-bond ${}^{1}J_{NH}$ and ${}^{1}D_{NH}$ couplings, i.e., negative, and ${}^{1h}J_{NH}$ is positive. The couplings measured from these spectra are listed in Table 1. Uncertainties in these splittings are estimated both from reproducibility in repeated



Figure 3. Diagram showing the angles θ and ϕ , defining the orientation of the T imino N₃H bond vector relative to the molecular alignment tensor, and the angles θ' and ϕ' , defining the orientation of the A-N₁···H hydrogen bond. The *z* axis corresponds to the largest component of the nearly axially symmetric diagonalized alignment tensor (D_a^{NH} = +16.1 Hz), and the *x* axis is the principal axis with the most negative dipolar coupling (taking into account the negative sign of γ_N). It nearly coincides with the *x*-axis in the bicelle medium (Tjandra et al., 2000), but the rhombicity is much smaller (R = 0.05 in Pf1; 0.26 in bicelles). The alignment tensor orientation was determined as described in footnote *a* to Table 2. Its magnitude is scaled according to the size of the ¹⁵N-¹H one-bond dipolar couplings in the two samples, in order to correct for slightly different Pf1 concentrations.

experiments, and from the relation that the random uncertainty in a peak position corresponds approximately to half its line width, divided by the signal-to-noise ratio (Kontaxis et al., 2000).

If hydrogen bonds were perfectly linear, their length follows directly from $r_{N...H} = ({}^{1}D_{NH}/{}^{1h}D_{NH})^{1/3}$ ×r_{NH}. Assuming an N-H bond length of 1.04 Å (Dingley et al., 1999), this yields $r_{N...H} = 1.80$ Å for A-T and $r_{N...H} = 1.86$ Å for G-C. However, the model of a perfectly linear hydrogen bond is clearly an oversimplification. Analysis of the orientation of the N-H and N···H vectors in two crystal structures and a recently refined solution structure of this dodecamer yields $N-H \cdots N$ angle estimates that vary in the range from 159 to 177°. Using the orientation of the N···H hydrogen bond vectors from these structures relative to the alignment tensor, listed in Table 2, slightly different r_{N...H} values are obtained. Comparison of the resulting $r_{N...H}$ lengths, reported in Table 3, indicates that the uncertainty in this distance derived from the present set of experiments is dominated by the uncertainty in the structure of this dodecamer.

The H-bond lengths obtained for the linear H-bond model are in closest agreement with those from a statistical analysis of three very high resolution B-DNA crystal structures, which yielded $r_{N \cdots N} = 2.92 \pm 0.05$ Å for C-G and $r_{N...N} = 2.81 \pm 0.05$ Å for A-T (Dingley et al., 1999). Subtracting 1.04 Å for the covalent N-H bond, these numbers are in near-perfect agreement with the above reported r_{N···H} values. However, when using the orientations of the various H-bonds with respect to the alignment tensor, as defined by the angles θ , ϕ , θ' and ϕ' (Figure 3, Table 2), we find considerably shorter $r_{N...H}$ values (Table 3). These shorter values result primarily from the larger deviation from 90° with respect to the helix axis (which is collinear with the z axis of the alignment tensor) for θ' than for θ . The across H-bond dipolar coupling is largest when $\theta' = 90^\circ$, or inversely a given dipolar coupling would yield the longest $r_{N..H}$ for $\theta' = 90^{\circ}$, and a very short distance when θ' approaches the magic angle of 54.7°.

The above analysis assumes that the librational and vibrational corrections are the same for the covalent N-H bond and for the $H \cdot \cdot \cdot N$ interaction. This is unlikely to be correct. However, although one may expect the relative movement of the imino proton and the Hbond-accepting N to be larger than for the covalent imino N-H pair, this does not necessarily translate into larger fluctuations in θ' than θ because the $r_{N\cdots H}$ distance is nearly 80% longer than r_{NH}. It is also possible that the shorter distances observed in our study (Table 3) compared to crystallographic data result from relatively large amplitude fluctuations of r_{N...H}. As the dipolar coupling measures $\langle r_{N\cdots H}^{-3} \rangle$, the $r_{N\cdots H}$ obtained from $r_{N\cdots H} = \langle r_{N\cdots H}^{-3} \rangle^{-1/3}$ results in a shortening relative to $\langle r_{N\cdots H} \rangle$. It is likely that analysis of molecular dynamics trajectories and additional relaxation and cross-correlated relaxation data will be able to provide insight into how large the effects from these motions on the dipolar coupling will be.

Our data present the first report of residual dipolar coupling across hydrogen bonds in a nucleic acid. Data are in fair agreement with crystallographic data, and indicate a shorter H-bond length for A-T than for G-C. However, when using the crystallographic or NMR structure as a reference, H-bond distances are found to be about 0.1 Å shorter than observed in high resolution crystal structures. The reason for this slight discrepancy is likely due to dynamic averaging of the $\langle r_{N\cdots H}^{-3} \rangle$ dipolar interaction, or it may be that H-bonds are slightly more linear than seen in experimental structures. The ability to obtain such precise information from dipolar couplings suggests that these couplings are well suited to study the critical H-bond interactions in the active sites of enzymes.

Acknowledgements

We thank Nico Tjandra and Dennis Torchia for useful discussions, and Akira M. Ono and Yu Oogo for help with the synthesis of the ¹⁵N-labeled oligonucleotide. This work is supported in part by a grant from CREST of JST.

Note added in proof

After submission of this manuscript, a paper appeared describing a pulse sequence for measurement of transhydrogen-bond J_{NH} couplings that is closely related to the one used in the present paper (Yan, X.Z., Kong, X.M., Xia, Y.L., Sze, K.H. and Zhu, G. (2000) *J. Magn. Reson.*, **147**, 357–360).

References

- Bax, A., Max, D. and Zax, D. (1992) J. Am. Chem. Soc., 114, 6923–6925.
- Becker, E.D. (Ed.) (1996) Hydrogen bonding, Encyclopedia of Nuclear Magnetic Resonance, Wiley, Chichester.
- Benedict, H., Shenderovich, I.G., Malkina, O.L., Malkin, V.G., Denisov, G.S., Golubev, N.S. and Limbach, H.H. (2000) J. Am. Chem. Soc., 122, 1979–1988.
- Cornilescu, G., Ramirez, B.E., Frank, M.K., Clore, G.M., Gronenborn, A.M. and Bax, A. (1999) J. Am. Chem. Soc., 121, 6275–6279.

- Dickerson, R.E. and Drew, H.R. (1981) J. Mol. Biol., 149, 761–786. Dingley, A.J. and Grzesiek, S. (1998) J. Am. Chem. Soc., 120, 8293– 8297.
- Dingley, A.J., Masse, J.E., Peterson, R.D., Barfield, M., Feigon, J. and Grzesiek, S. (1999) J. Am. Chem. Soc., 121, 6019–6027.
- Emsley, L. and Bodenhausen, G. (1989) J. Magn. Reson., 82, 211-221.
- Golubev, N.S., Shenderovich, I.G., Smirnov, S.N., Denisov, G.S. and Limbach, H.H. (1999) *Chemistry – a European Journal*, 5, 492–497.
- Grzesiek, S., Cordier, F. and Dingley, A.J. (2001) Methods Enzymol., in press.
- Kojima, C., Ono, A. and Kainosho, M. (2000) J. Biomol. NMR, 18, 269–277.
- Kontaxis, G., Clore, G.M. and Bax, A. (2000) J. Magn. Reson., 143, 184–196.
- Meissner, A., Duus, J.O. and Sørensen, O.W. (1997) J. Magn. Reson., 128, 92–97.
- Montelione, G.T. and Wagner, G. (1989) J. Am. Chem. Soc., 111, 5474–5475.
- Pervushin, K., Ono, A., Fernandez, C., Szyperski, T., Kainosho, M. and Wüthrich, K. (1998a) *Proc. Natl. Acad. Sci. USA*, 95, 14147–14151.
- Pervushin, K.V., Wider, G. and Wüthrich, K. (1998b) J. Biomol. NMR, 12, 345–348.
- Shui, X.Q., McFail-Isom, L., Hu, G.G. and Williams, L.D. (1998) Biochemistry, 37, 8341–8355.
- Silver, M.S., Joseph, R.I. and Hoult, D.I. (1984) *Nature*, **310**, 681–683.
- Tjandra, N. and Bax, A. (1997) Science, 278, 1111-1114.
- Tjandra, N., Tate, S., Ono, A., Kainosho, M. and Bax, A. (2000) *J. Am. Chem. Soc.*, **122**, 6190–6200.
- Wang, Y.X., Jacob, J., Cordier, F., Wingfield, P., Stahl, S.J., Lee-Huang, S., Torchia, D., Grzesiek, S. and Bax, A. (1999) J. Biomol. NMR, 14, 181–184.