Analysis of Sugar Puckers and Glycosidic Torsion Angles in a DNA G-Tetrad Structure by Heteronuclear Three-Bond *J* Couplings

Guang Zhu,† David Live,*,‡ and Ad Bax§

Department of Electrical Engineering
University of Maryland, College Park, Maryland 20874
Cellular Biochemistry and Biophysics Program
Memorial Sloan-Kettering Cancer Center
New York, New York 10021
Laboratory of Chemical Physics
NIDDK/NIH, Bethesda, Maryland 20892
Received May 23, 1994

As no homonuclear three-bond ^{1}H couplings are available to define the glycosidic torsion angle, determination of this important variable in oligonucleotide structure has relied entirely on interpretation of NOE data. Accurate determination of sugar conformation in oligonucleotides remains difficult even when in addition to NOEs the important information contained in homonuclear ^{1}H–^{1}H *J* couplings[1,2] can be obtained. Here we demonstrate that heteronuclear three-bond ^{1}H–^{13}C *J* couplings between sugar protons and either sugar or base carbons can readily be measured, providing important supplemental information for determining nucleic acid structure.[2-8] The approach used in the present work is based on the concept of quantitative *J* correlation[9,10] and is applicable even when the size of the heteronuclear *J* coupling is significantly smaller than the line widths of the pertinent resonances. In view of the current limitations with ^{13}C labeling DNA, the viability of an experiment such as described below, which is applicable at natural abundance, is very appealing. This strategy can be extended to measuring heteronuclear couplings in a range of other molecules, including RNA, where the reduced ^{1}H spectral dispersion and C3′ *endo* sugar pucker frequently make measurement of ^{1}H–^{1}H couplings more difficult. Of course, the ability to isotopically enrich RNA and the ongoing development of techniques for labeling DNA can greatly increase the sensitivity of the long range J_{CH} coupling measurement.

Below we report the measurement of $^3J_{CH}$ couplings for the unlabeled DNA oligomer, d(GGTCGG). Under the conditions studied, this fragment adopts a G-tetrad structure.[11-12] It has been postulated that such structures are found at the ends of chromosomes, and variation in their length with age controls stability of the chromosome and longevity of the cell. A variety of possible molecular conformations consistent with four G residues hydrogen bonded in a plane have been postulated, e.g., chains parallel or antiparallel and bases *syn* or *anti* relative to the sugar.[11-21] In the model system studied here, the G1 and G5 previously were found to be *syn* and G2 and G6 *anti*.[11,12]

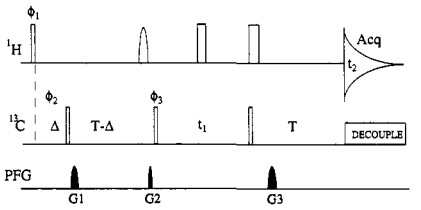

Figure 1. Pulse sequence of the band-selective HMBC experiment. Narrow and wide bars represent 90° and 180° pulses, respectively. The shaped 180° ^{1}H pulse immediately preceding pulsed field gradient *G*2 has an amplitude profile corresponding to the central lobe of a sinc pulse and a duration of 1–2 ms. Gradient pulses are sine-bell shaped with a peak amplitude of 25 G/cm and have the following durations: *G*1 = 650 μs, *G*2 = 150 μs, *G*3 = 500 μs. The effect of ^{1}H–^{1}H *J* modulation during the de- and rephasing delays, *T*, is eliminated by the band-selective ^{1}H inversion just prior to t_1, and spectra are acquired in the 2D absorptive mode. The first 90° ^{13}C pulse is optional and serves to suppress one-bond ^{1}H–^{13}C correlations, which is useful when measuring H1′–C4′ correlations. Phase cycling is as follows: $\phi_1 = x,y,-x,-y$; $\phi_2 = 8(x),8(-x)$; $\phi_3 = 4(x),4(-x)$; receiver = $x,-y,-x,y,-x,y,x,-y$. Quadrature in the t_1 dimension is obtained by alternating ϕ_3 in the usual States-TPPI manner. Delay durations: T = 34 ms; Δ = 3.1 ms. For the reference spectrum, data are collected with only the first four steps of the phase cycle. The second spectrum for each t_1 is not collected, and zeros are inserted instead.

The HMBC scheme used in the present work is sketched in Figure 1. Homonuclear ^{1}H *J* dephasing during the lengthy J_{CH} de-rephasing delays, *T*, which normally has a detrimental effect on the sensitivity of the HMBC experiment, is refocused by a band-selective 180° pulse. The quantitative J_{CH} information results from comparing the long range correlation intensity, I_{CH}, with the intensity, I_{ref}, observed for the same proton in a reference spectrum which is recorded with the same pulse sequence but selects magnetization that is not transferred to ^{13}C (Figure 1).[10] For a given ^{1}H resonance, the ratio of peak intensities can be expressed as:

$$I_{CH}/I_{ref} = A(N/N_{ref}) \sin[\pi J_{CH}(T-\Delta)] \sin(\pi J_{CH}T) \sin^2(\alpha) \quad (1)$$

where *A* is the natural abundance of ^{13}C (A = 0.01108), *N* and N_{ref} are the numbers of scans in the ^{1}H–^{13}C correlation and reference spectra, respectively, and α is the effective flip angle of the ^{13}C pulses. The probe used here gives $\sin^2(\alpha)$ = 0.88 for a nominal 90° pulse. Considering that the reference spectrum is far more sensitive than the HMBC spectrum, the uncertainty in the J_{CH} value derived from eq 1 may be attributed solely to the noise in the HMBC spectrum. The most accurate determination of J_{CH} is made when the second derivative of I_{CH} with respect to *T* is zero:

$$\frac{d^2}{dT^2} \sin[\pi J_{CH}(T-\Delta)] \sin(\pi J_{CH}T) \exp(-2T/T_2) = 0 \quad (2)$$

For long values of the ^{1}H transverse relaxation time, T_2, the optimum *T* duration therefore equals $1/(4J_{CH}) - \Delta/2$. In the limit where the ^{1}H T_2 is much shorter than $1/(4J_{CH})$, the optimum *T* duration becomes $\sim T_2$. In the present study, a *T* delay of 34 ms was used. The intensity ratio of eq 1 is determined from the scaling factor needed to minimize the difference between the reference peak and the long range correlation. Uncertainties in the measurements are estimated from the range of scaling factors that do not show intensity in the difference spectrum above the

† University of Maryland.
‡ Memorial Sloan-Kettering Cancer Center.
§ NIDDK/NIH.

(1) Altona, C *Recl. Trav. Chim. Pays-Bas* **1982**, *101*, 413–433. Majumdar, A.; Hosur, R. V. *Prog. Nucl. Magn. Reson. Spectrosc.* **1992**, *24*, 109–158.
(2) Davies, D. B. *Prog. Nucl. Magn. Reson. Spectrosc.* **1978**, *12*, 135–225.
(3) Poppe, L.; Stuike-Prill, R.; Meyer, B.; van Halbeek, H. *J. Biomol. NMR* **1992**, *2*, 109–136.
(4) Delbaere, L. T. J.; James, M. N. G.; Lemieux, R. U. *J. Am. Chem. Soc.* **1972**, *95*, 7866–7868.
(5) Dea, P.; Kreishman, G. P.; Schweizer, M. P.; Witkowski, J. T.; Nunlist, R.; Bramwell, M. In *Proceedings of the 1st International Conference on Stable Isotopes in Chemistry, Biology and Medicine*; Klein, P. D., Ed.; NTIS: Springfield, VA, 1973; pp 84–88.
(6) Davies, D. B.; Rajani, P.; MacCross, M.; Danyluk, S. S. *Magn. Reson. Chem.* **1985**, *23*, 72–77.
(7) Schmieder, P.; Ippel, J. H.; van den Elst, H.; van der Marel, B. A.; van Boom, J. H.; Altona, C.; Kessler, H. *Nucleic Acids Res.* **1992**, *20*, 4747–4751.
(8) Hines, J. H.; Varani, G.; Landry, S. M.; Tinoco, I. *J. Am. Chem. Soc.* **1993**, *115*, 11002–11003.
(9) Bax, A.; Max, D.; Zax, D. *J. Am. Chem. Soc.* **1992**, *114*, 6923–6925.
(10) Zhu, G.; Bax, A. *J. Magn. Reson.* **1993**, *104A*, 353–357.
(11) Wang, Y.; de los Santos, C.; Gao, X.; Greene, K.; Live, D.; Patel, D. J. *J. Mol. Biol.* **1991**, *222*, 819–832.
(12) Wang, Y.; Patel, D. J. Unpublished results.
(13) Williamson, J. R.; Rughuraman, M. K.; Cech, T. R. *Cell* **1989**, *59*, 871–880.
(14) Blackburn, E. H. *Nature* **1991**, *350*, 569–573.
(15) Henderson, E.; Hardin, C. C.; Walk, S. K.; Tinoco, I., Jr.; Blackburn, E. H. *Cell* **1987**, *51*, 899–908.
(16) Smith, F. W.; Feigon, J. *Nature* **1992**, *356*, 164–168.
(17) Kay, C. H.; Zhang, X.; Rutliff, R.; Moyzis, R.; Rich, A. *Nature* **1992**, *356*, 126–131.
(18) Wang, Y.; Patel, D. J. *Biochemistry* **1992**, *31*, 8112–8119.
(19) Smith, F. W.; Feigon, J. *Biochemistry* **1993**, *32*, 8682–8692.
(20) Wang, Y.; Patel, D. J. *Structure* **1993**, *1*, 263–383.
(21) Wang, Y.; Patel, D. J. *J. Mol. Biol.* **1993**, *234*, 1171–1183.

0002-7863/94/1516-8370$04.50/0

Table 1. Proton–Carbon Coupling Constants[a] Measured in d(GGTCGG)

	C_4/C_6–$H_{1'}$	C_8/C_6–$H_{1'}$	$C_{1'}$–$H_{3'}$	$C_{1'}$–$H_{4'}$	$C_{5'}$–$H_{3'}$	$C_{3'}$–$H_{1'}$
G1	5.4	3.5[b]	7.8	2.7	2.3	<1.5
G2	2.0	4.4	7.2	*c*	2.8	<1.2
T3	1.7[b]	3.4	6.7[b]	3.6[b]	2.7	<1.2
C4	≤1.1	2.4	7.3	*c*	2.5	<1.1
G5	5.6	3.2	7.8	<1.7	3.1	<1.4
G6	1.8[b]	4.5	*d*	*c*	*d*	<0.9

[a] Errors are estimated to be ±0.3 Hz unless indicated otherwise. [b] Error is estimated at ±0.4 Hz. [c] $H_{4'}$ resonates outside the spectral window. [d] $H_{3'}$ overlaps with the HDO resonance.

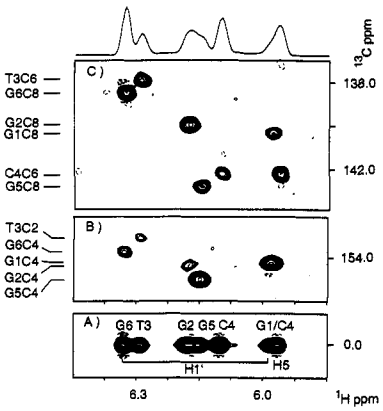

Figure 2. Regions of (A) the reference and (B,C) the 2D band-selective HMBC spectra, showing cross peaks between H1′ and base carbons for a sample of d(GGTCGG), 19 mM monomer, in 0.5 mL of 0.1 M NaCl, 10 mM sodium phosphate in D_2O, at 20 °C, recorded on a Bruker AMX-600 spectrometer. The pH of the buffered solution before drying and redissolving in D_2O was 6.8. The ^1H carrier was positioned at 6.05 ppm. The spectrum results from a 224* × 256* data matrix (*n* refers to *n* complex data points) with acquisition times of 9 (t_1) and 64 ms (t_2). Using 384 transients per complex t_1 increment, the accumulation time was ~40 h. The reference spectrum was acquired with four transients per real t_1 increment in a total time of ~25 min. Peaks are labeled by residue and atom number.

thermal noise. For the common case where $J_{CH}T \ll 0.5$, it is seen from eq 1 that the uncertainty in the derived $^3J_{CH}$ value depends on the square root of the error in I_{CH}/I_{ref}, i.e., a 4-fold lower sample concentration results only in a doubling of the error in $^3J_{CH}$.

The $^3J_{CH}$ couplings reported in Table 1 have been measured from two separate experiments, one with the selective 180° pulse in the H1′ region at 6 ppm, and one at 5 ppm, in the H3′ and H4′ region. Cross peaks arising from the three-bond couplings between H1′ and base carbons are illustrated in Figure 2. The ^{13}C assignments are based on ^1H assignments made previously by Wang and Patel[12] and one-bond ^1H–^{13}C HMQC correlation. The intensity of the H5–C6 *J* correlation for the only cytidine in d(GGTCGG), also visible in Figure 2, corresponds to a *J* coupling of 4.7 Hz, in good agreement with a 4.4 Hz value measured directly in free cytidine and a 4.5 Hz value reported by Schmieder et al.[7] using a multiplet fitting procedure.[22] Procedures such as multiplet fitting and ECOSY[8] have been of limited use, however, for measuring ^1H–^{13}C vicinal couplings in sugars or between H1′ and base carbons. In addition to the couplings reported in Table 1, H1′–H2′ and H1′–H2″ couplings have been measured from a PE COSY experiment,[23] providing an independent measure of the C1′–C2′ torsion angle. The $^3J_{H1'H2'}$ values for the four unique G residues range from 10.9 to 12.2 Hz, indicative of H1′–C1′–C2′–H2′ dihedral angles near 180°, corresponding to a pseudorotation angle, *P*, of 144° and a C2′ *endo* sugar pucker. A deoxyribose with this *P* gives rise to a torsion angle between H1′ and C3′ of about 80° and between H3′ and C1′ of about −160°. An estimate of values expected for $^3J_{H1'C3'}$ and $^3J_{H3'C1'}$ can be made using a heteronuclear Karplus equation which also accounts for electronegativity effects.[24] This equation predicts $^3J_{H1'C3'} \approx 0.8$ Hz and $^3J_{H3'C1'} \approx 8.3$ Hz, assuming the substituent effects of N9 and O4′ are additive. Both values are consistent with the measured values reported in Table 1. Based on the $^3J_{H1'C3'}$ and $^3J_{H3'C1'}$ values (Table 1) and $J_{H1'H2'}$ couplings of 9.9 (T3) and 10.4 Hz (C4), the sugar puckers for the T and C residues are also in the C2′ *endo* range. These couplings indicate, however, that the conformation of these sugars is less extreme than that of the Gs.

The relationship between the glycosidic bond torsion angle and H1′-to-base carbon couplings previously has been analyzed for bicyclic analogs of uridine.[2,6] Davies's work shows that these heteronuclear couplings can be used to distinguish *syn* and *anti* base arrangements, with $^3J_{H1'C2}$ smaller than $^3J_{H1'C6}$ for the *anti* configuration and the opposite for the *syn* case.[6] Although Davies's Karplus equation[6] was parametrized for U, our data indicate that this parametrization also provides reasonable results for G, as both $^3J_{H1'C4}$ and $^3J_{H1'C8}$ are found to be simultaneously compatible with a single glycosidic torsion angle. For G1 and G5, both of which were found to be *syn* on the basis of NOE data, the average $^3J_{H1'C4}$ values (5.5 Hz) are larger than $^3J_{H1'C8}$ (3.3 Hz). These values are compatible with two values of the C4–N9–C1′–O4′ glycosidic torsion angle, *ca.* 87° and *ca.* 33°, both in the *syn* range. For G2 and G6, the $^3J_{H1'C8}$ and $^3J_{H1'C4}$ values are ~4.4 and ~1.9 Hz, respectively, corresponding to a glycosidic torsion angle of either *ca.* −85° or *ca.* −155°, both in the *anti* region. The T and C residues have $^3J_{H1'C6}$ values that are larger than $^3J_{H1'C2}$ (Table 1), consistent with *anti* configurations.

Although no detailed structure is yet available for d(GGTCGG), a high-quality structure recently has been reported[25] for d(GGTTGGTGTGGTTGG), with features similar to the molecule studied here in terms of sequence and structure. It contains two stacked G tetrads with alternating *syn* and *anti* bases. In d(GGTTGGTGTGGTTGG), the individual glycosidic angles of the four *syn* Gs in the six lowest energy structures have an overall range of 45–120°, and the *anti* Gs range between −100° and −160°. These NOE derived glycosidic torsion angles are complementary to the tighter angular constraints derived from *J* couplings, as they can resolve the ambiguity associated with the multivalued character of the Karplus equation.

Our results demonstrate that quantitative HMBC correlation provides an effective means for measuring a number of important heteronuclear $^3J_{CH}$ couplings and torsion angles in oligonucleotides. The *J* couplings measured for d(GGTCGG) indicate that Davies's parametrization of the Karplus equation, which was based on cyclic uridine analogs, is also applicable to purine nucleotides. The heteronuclear *J* couplings provide valuable conformational constraints in understanding the structure of d(GGTCGG) in particular. In conjunction with NOE data, this can lead to a better definition of solution structures of nucleic acids by NMR in general.

Acknowledgment. We thank K. L. Greene for recording HMQC spectra on d(GGTCGG) and Stephan Grzesiek for valuable assistance. This work was supported by NIH Grants GM 49707 (G.Z.) and DK 38676 (D.L.), Cancer Center Support Grant NCI-P30-CA-08478 (MSKCC), and the Intramural AIDS-Directed Antiviral Program of the Office of the Director of the National Institutes of Health (A.B.).

Supplementary Material Available: Figure showing H3′/H4′–^{13}C long range connectivities in d(GGTCGG) (2 pages). This material is contained in many libraries on microfiche, immediately follows this article in the microfilm version of the journal, and can be ordered from the ACS; see any current masthead page for ordering information.

(22) Keeler, J.; Neuhaus, D.; Tittman, J. J. *Chem. Phys. Lett.* **1988**, *146*, 545–548.
(23) Bax, A.; Lerner, L. *J. Magn. Reson.* **1988**, *79*, 429–438.
(24) van Beuzekom, A. A.; de Leeuw, F. A. A. M.; Altona, C. *Magn. Reson. Chem.* **1990**, *28*, 68–74.
(25) Schultz, P.; Macaya, R. F.; Feigon, J. *J. Mol. Biol.* **1994**, *235*, 1532–1547.