Estimation of ^{31}P–^{1}H and ^{1}H–^{1}H Vicinal Coupling Constants along the DNA Backbone by 2D HELCO Measurements

K. V. R. Chary,[a] Vinit K. Rastogi,[a] Girjesh Govil[a] and H. Todd Miles[b]

[a] *Chemical Physics Group, Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay 400 005, India*
[b] *National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA*

An accurate method for simultaneous estimation of ^{31}P–^{1}H and ^{1}H–^{1}H vicinal coupling constants along the backbone of DNA fragments is described using the two-dimensional heteronuclear long-range correlation (2D HELCO) experiment.

Homonuclear and heteronuclear vicinal coupling constants (^{3}J) provide information about local conformations in proteins and nucleic acids.[1–3] For nucleic acids, $^{3}J(^{31}P–^{1}H)$ and $^{3}J(^{1}H–^{1}H)$ can be used to determine intervening torsion angels.[4] In large molecules, however, it is difficult to measure such coupling constants accurately.[5]

In earlier studies involving measurement of $J(^{31}P–^{1}H)$, ^{1}H-detected 2D $J(^{31}P–^{1}H)$ resolved spectroscopy has been used.[6] Measurements have also been carried out by simplication of complex multiplet structures in a 2D $^{31}P–^{1}H$ spectrum with the suppression of undesirable splitting of the individual cross-peaks.[7] These methods are only useful for extracting $^{3}J(^{31}P–H3')$ and are generally applicable to small nucleotides. One method for determining approximate values of $^{1}H–^{1}H$ coupling constants along the backbone of the DNA fragments is from the sum of the coupling constants[8] (ΣJ) which is obtained from antiphase absorptive spectra such as E COSY.[9]

We propose a more accurate method for simultaneous estimation of $^{31}P–^{1}H$ and $^{1}H–^{1}H$ vicinal coupling constants using HELCO which was proposed primarily[5] to correlate phosphorus and proton spins in nucleic acids. We have recorded the HELCO spectrum of a duplex dodecanucleotide d-(GGTACIAGTACC)$_2$, using the pulse sequence shown in Fig. 1. Fig. 2(*A*) shows the $^{31}P–H3'$ correlations. Assignments of the peaks are based on the sequential ($^{31}P–^{1}H$) correlation described earlier.[5] The cross-peaks observed at $(\omega_1, \omega_2) = \delta$ (^{31}P, ^{1}H) are inphase absorptive along the ω_1 axis and antiphase absorptive along ω_2. The latter is a direct result of the active $^{31}P–H3'$ coupling constants.

$J(^{31}P–^{1}H)$ can be estimated by simulation of HELCO spectra peaks. For such simulations, we have used the values of $^{3}J(H3'–H2')$ and $^{3}J(H3'–H2'')$, obtained by simulation of the characteristic multiplet structures of individual H1′–H2′ and H1′–H2″ cross-peaks in the E COSY spectrum.[9] As an example, the simulated cross-peaks arising from the coupling between H3′ and ^{31}P are shown along with the experimental ones in Fig. 2(*C*). We have been able to simulate all cross-peaks in the oligomer, except that due to G8. In this case, the $^{31}P–H3'$ cross-peak is weak and overlaps partially with that of G2. The *J* values are given in Table 1.

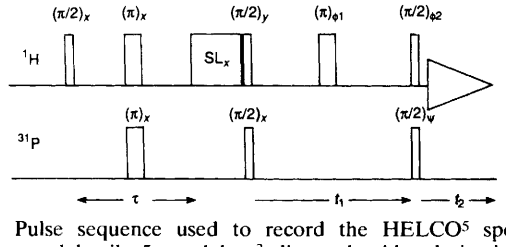

Fig. 1 Pulse sequence used to record the HELCO[5] spectrum. Experimental details: 5 mmol dm^{-3} oligonucleotide solution in 99.9% D_2O, 40 mmol dm^{-3} phosphate buffer; pH 7.0, temperature 32 °C; $\tau = 20$ ms, $t_{1,max} = 38.0$ ms, $t_{2,max} = 1.36$ s, recycle delay = 500 ms; no. of scans 256; time-domain data points 38 and 4096 along t_1 and t_2; total recording time *ca.* 4 h. The ^{1}H carrier frequency was kept on the water resonance (sweep width 1506 Hz). No presaturation was used. In ω_1, the carrier was in the centre of the ^{31}P chemical shifts of the oligonucleotide (sweep width 250 Hz). The data were multiplied with sine bell window functions shifted by $\pi/2$ and $\pi/8$ along t_1 and t_2, respectively, and zero-filled to 128 and 8096 data points along t_1 and t_2 prior to 2D-FT. The digital resolution along ω_1 and ω_2 corresponds to 6.5 and 0.36 Hz per pt, respectively. The spectrum was recorded on a Bruker AMX 500 spectrometer.

Relationships between $^{3}J(H3'–^{31}P)$ and the intervening torsion angle, ϵ, have been proposed.[4,10,11] In several cases in Table 1, $^{3}J(H3'–^{31}P)$ clusters around 6 Hz, corresponding to ϵ of 0, 120, 200 and 280°. Potential-energy calculations have shown[10] that steric hindrance prevents nucleotides from acquiring conformations with $\epsilon < 160°$, while values around 200 and 270° correspond to minima in potential-energy surfaces. Thus, both the values in the above ranges correspond to acceptable solutions. While the usually encountered B_I conformation of DNA duplexes corresponds to $\epsilon \approx 200°$, the

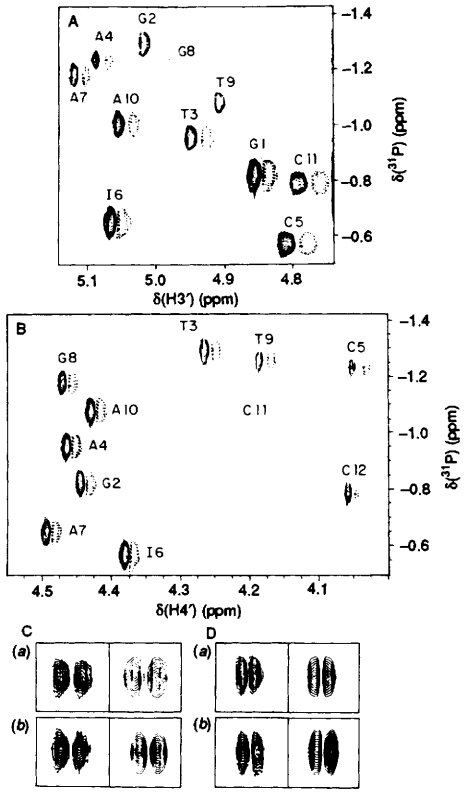

Fig. 2 (*A*) Selected region of HELCO spectrum of d-GGTACIAGTACC, showing the expected $^{31}P–H4'$ *J* correlations. (*B*) Selected region of the HELCO spectrum of d-GGTACIAGTACC, showing all the expected $^{31}P–H4'$ *J* correlations. (*C*) Simulated $^{3}J(^{31}P–H3')$ HELCO cross-peak multiplet structures for (*a*) T3 and (*b*) A10 units in comparison with experimental peaks. Simulations have been carried out using a software developed by us, to extract values of the heteronuclear ($^{31}P–H3'$) coupling constants. The coupling constants (in Hz) used in simulation are: (*a*) $^{3}J(H2'–H3') = 5.8$; $^{3}J(H2''–H3') = 2.7$; $^{3}J(H3'–H4') = 3.8$; and $^{3}J(^{31}P–H3') = 8.0$; (*b*) $^{3}J(H2'–H3') = 5.0$; $^{3}J(H2''–H3') = 2.5$; $^{3}J(H3'–H4') = 3.0$; and $^{3}J(^{31}P–H3') = 6.0$. (*D*) Simulated $^{31}P–H4'$ HELCO cross-peak multiplet structures for (*a*) T3 and (*b*) A10 units, along with the experimental ones. The *J* values (in Hz) used are: (*a*) $^{3}J(H4'–H5'/H5'') = 1.9, 3.1$; $^{3}J(H4'–H3') = 3.8$; and $^{3}J(^{31}P–H4') = 1.0$; (*b*) $^{3}J(H4'–H5'/H5'') = 2.3, 2.5$; $^{3}J(H4'–H3') = 3.0$; and $^{3}J(^{31}P–H4') = 2.6$.

Table 1 Coupling constants (Hz) in d-GGTACIAGTACC and the corresponding back-bone torsion angles (degrees)

	H3′–H4′	^{31}P–H3′	H4′–H5′/H5″	^{31}P–H4′	ϵ^a	γ^b
G1	1.0	7.0	*c*	*c*	204, 276	*c*
G2	2.0	6.0	2.0/2.8	1.0	200, 280	54
T3	3.8	8.0	1.9/3.1	1.0	209, 271	52, 60
A4	2.2	6.4	1.4/2.8	2.1	202, 278	*ca.* 60
C5	4.0	12.6	1.9/3.8	3.0	240	45, 67
I6	1.0	7.0	1.3/3.0	1.6	204, 27	*ca.* 60
A7	1.5	5.8	1.3/2.5	2.2	199, 281	*ca.* 60
G8	*c*	*c*	2.0/2.9	3.0	*c*	52
T9	6.0	9.5	2.0/3.7	2.6	217, 263	45, 67
A10	3.0	6.0	2.3/2.5	2.6	200, 280	56
C11	4.3	12.0	*c*	*c*	240	*c*
C12	*c*	*c*	*c*	*c*	*c*	*c*

^{a} ϵ estimated by making use of the equation:[4] $^3J_{HCOP} = 16.3\cos^2\phi - 4.6\cos\phi$, and the relation between ϕ (HCOP) and ϵ,[10] *i.e.* $\epsilon = 240 \pm |\phi|$. ^{b} estimated by making use of the following equation:[11] $^3J_{HH} = 13.7\cos^2\gamma - 0.73\cos\gamma + \Sigma_{i\chi i}[0.56 - 2.47\cos^2(Z_i\gamma + 16.9|\Delta\chi_i|)]$; $\Delta\chi_i = 1.3$ for O and 0.4 for C; Z_i = relative orientation factor: ± 1. ^{c} See text. The error in the estimation of *J* values from the simulation of cross-peak patterns is *ca.* 10%.

second range (*ca.* 270°) has been observed in the B_{II} conformation. From the coupling constant data, one cannot distinguish between the two acceptable conformations. Sklenar and Bax[7] stated that only the first conformation is acceptable.

For C5 and C11, where the experimentally observed *J* values are 12.6 and 12.0 Hz, the maximum *J* value, in the energetically accepted range of ϵ, is 11.7 Hz if the relationship proposed in refs. 4 and 10 is used, and 10.8 Hz if relationship proposed in ref. 11 is used. In both cases, the ϵ value corresponding to the maximum *J* value is 240°. Taking into account the experimental errors in the estimation of *J*, only the relationship proposed in ref. 4 is consistent with the observations. One of the two nucleotides which deviate from the usually observed ϵ values (*i.e.*, 200 or 270°) is in the mismatch region and the other is near the terminal end of the oligonucleotide. These simulations also enable us to estimate ^3J(H3′–H4′) which, in conjuction with ^3J(H1′–H2′), ^3J(H1′–H2″), ^3J(H3′–H2′) and ^3J(H3′–H2″) derived from the E COSY[9] spectrum, help to establish the sugar pucker and the backbone torsion angle δ.

The second part of the HELCO spectrum [Fig. 2(*B*)] contains the expected ^{31}P–H4′ cross-peaks. The antiphase character of the cross-peaks along the ω_2 axis is due to the long-range coupling $^4J(^{31}P–H4')$. These cross-peaks are modulated along the ω_2 axis by passive ^3J(H4′–H3′), ^3J(H4′–H5′) and ^3J(H4′–H5″) couplings. Knowing ^3J(H4′–H3′) (from the simulation of the ^{31}P–H3′ cross-peak), all the ^{31}P–H4′ cross-peaks (with the exception of C11 and C12), have been simulated [*e.g.* see insets in Fig. 2(*D*)]. This enables an estimate of hitherto inaccessible ^3J(H4′–H5′), ^3J(H4′–H5″) and $^3J(^{31}P–H4')$. For C11, the ^{31}P–H4′ cross-peak is weak and could not be used to estimate *J*. For C12, ^3J(H4′–H3′) could not be estimated from the ^{31}P–H3′ correlation, preventing simulation of the ^{31}P–H4′ cross-peak. The *J* values thus obtained are also given in Table 1. While $^4J(^{31}P–H4')$ reflects the proportion of the 'W' conformation along the P–O5′–C5′–C4′–H4′ coupling pathway, information about ^3J(H4′–H5′) and ^3J(H4′–H5″) is valuable in estimating γ. Since both the *J* values are <4 Hz, the conformation around the C4′–C5′ bond is g^+ (or *gg*). To obtain more accurate values of γ, we have used the proposed relation[11] between γ and ^3J(C4′–C5′, C5″). This gives unique values for G2, G8 and A10 (Table 1). Two values each are obtained for T3, C5 and T9, which range from 45 to 67°. For A4, I6 and A7, this relation does not yield a solution which simultaneously satisfies ^3J(H4′–H5′) and ^3J(H4′–H5″). Considering the limitations of the Karplus-type relations, the main conclusion is that the conformation around the C4′–C5′ bond corresponds to g^+ (or *gg*), as is generally observed in nucleic acids.

There are six torsion angles in the DNA backbone which determine its 3D structure. The constraints on ϵ and γ obtained from the HELCO measurements are therefore valuable. In conjunction with the intrastrand–internucleotide distances from the homonuclear NOESY spectrum, one may be able to derive ranges of all the torsion angles. When coupled with distance geometry algorithms such as TANDY2S,[12] the possible families of conformations which are consistent with the NMR data and interstrand hydrogen-bonding network can be further restricted.

The facilities provided by the National Facility for High Field NMR supported by the Department of Science and Technology, Government of India, and located at T.I.F.R., Bombay are gratefully acknowledged. We thank Mr R. Ajay Kumar for the use of his graphics software for spectral simulation.

Received, 23rd August, 1993; Com. 3/5083A

References

1 R. R. Ernst, *Angew. Chem., Int. Ed. Engl.*, 1992, **31**, 805.
2 K. Wuthrich, *NMR of Proteins and Nucleic Acids*, Wiley, New York, 1986; *Science*, 1989, **243**, 45.
3 D. Neuhaus, G. Wagner, M. Vasak, J. H. R. Kagi and K. Wuthrich, *Eur. J. Biochem.*, 1985, **151**, 257; G. T. Motelione and G. Wagner, *J. Am. Chem. Soc.*, 1989, **111**, 5474; D. Neri, G. Otting and K. Wuthrich, *J. Am. Chem. Soc.*, 1990, **112**, 3663; O. W. Sorensen, *J. Magn. Reson.*, 1990, **87**, 183; P. Schmieder, V. Thanabal, L. P. McIntosh, F. W. Dahlquist and G. Wagner, *J. Am. Chem. Soc.*, 1991, **113**, 6223; A. S. Edison, W. M. Westler and J. L. Markley, *J. Magn. Reson.*, 1991, **92**, 434; K. V. R. Chary, G. Otting and K. Wuthrich, *J. Magn. Reson.*, 1992, **93**, 218.
4 B. J. Blackburn, R. D. Lapper and I. C. P. Smith, *J. Am. Chem. Soc.*, 1973, **95**, 2873.
5 K. V. R. Chary, V. K. Rastogi and G. Govil, *J. Magn. Reson.*, 1993, **B102**, 81.
6 V. Sklenar and A. Bax, *J. Am. Chem. Soc.*, 1987, **109**, 7525.
7 V. Sklenar, H. Miyashiro, G. Zon, H. T. Miles and A. Bax, *FEBS Lett.*, 1986, **208**, 94; D. Williamson and A. Bax, *J. Magn. Reson.*, 1988, **76**, 174; A. Bax and R. Freeman, *J. Magn. Reson.*, 1981, **44**, 542; H. Kessler, C. Griesinger, J. Zarbock and H. R. Loosli, *J. Magn. Reson.*, 1984, **57**, 331; J. M. Fu, S. Schroeder, C. R. Jones, R. Santini and D. G. Gorenstein, *J. Magn. Reson.*, 1988, **7**, 577.
8 C. Altona, *Recl. Trav. Chim. Pays Bas.*, 1982, **101**, 413; S. G. Kim, L. J. Lin and B. R. Reid, *Biochemistry*, 1992, **31**, 3564.
9 C. Griesinger, O. W. Sorensen and R. R. Ernst, *J. Chem. Phys.*, 1986, **85**, 6837.
10 R. V. Hosur, K. V. R. Chary, A. Saran, G. Govil and H. T. Miles, *Biopolymers*, 1990, **29**, 953.
11 P. P. Lankhorst, C. A. G. Haasnoot, C. Erkelens and C. Altona, *J. Biomol. Struct. Dynam.*, 1984, **1**, 1387; F. J. M. van der Ven and C. W. Hilbers, *Eur. J. Biochem.*, 1988, **178**, 1.
12 R. Ajay Kumar, R. V. Hosur and G. Govil, *J. Biomol. NMR*, 1991, **1**, 363.