Volume 233, number 2, 319–325 FEB 05945 June 1988

Analysis of intrasugar interproton NOESY cross-peaks as an aid to determine sugar geometries in DNA fragments

K.V.R. Chary and Sandeep Modi

Chemical Physics Group, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Bombay 400005, India

Received 18 March 1988

A systematic analysis of the conformation of deoxyribofuranose rings in DNA fragments has been described using two-dimensional nuclear Overhauser effect spectroscopy (2D NOESY). The approach is based on the interpretation of the intrasugar proton-proton distances which can be estimated using a low-mixing-time pure-absorption mode ω_1-scaled NOESY spectrum. The experimental distances are compared with the theoretical values calculated as a function of pseudorotation phase angle (P) describing the sugar geometries. The approach can be used as a complementary aid to J couplings for establishing sugar conformations in individual nucleotide units of DNA fragments. Using this strategy on d-ACATCGATGT, we observed that individual nucleotides exhibit O4′-*endo* sugar pucker. The results rule out possibilities of the existence of a fast equilibrium (on the NMR time scale) between C2′-*endo* (or S-domain) and C3′-*endo* (or N-domain) sugar puckers.

Deoxyribofuranose; 2D NOESY; Pseudorotation phase angle

1. INTRODUCTION

With the advent of two-dimensional nuclear magnetic resonance (2D-NMR) techniques, it has become possible to obtain detailed structural information on DNA fragments [1–10]. These developments have led to new insights into the three-dimensional (3D) structures of DNA and its recognition by other biological molecules. The 3D structure of DNA is characterised by the sugar-phosphate backbone, sugar geometry and the glycosidic bond dihedral angle [11,12]. The sugar conformation is correlated with several other torsional angles and thus the secondary structure of DNA is dictated by the sugar puckers to a large extent. An elegant description of the conformation of a five-membered deoxyribofuranose ring is based on the pseudo-rotation concept [13]. Here, the torsional angles in the sugar ring are described in terms of pseudorotation phase angle (P) and maximum sugar pucker (T_m). Further, in view of the scatter of the observed sugar conformations, one often uses the terms N ($P = -90$ to $+90°$ covering conformations such as C3′-*endo*, C2′-*exo*, etc.) and S ($P = 90$ to $270°$ covering conformations such as C2′-*endo*, C3′-*exo*, etc.) to distinguish between the two families of energy-favoured conformations.

Previously, we have discussed approaches based on correlated spectroscopy (COSY) in 2D NMR for determination of sugar geometries in DNA fragments. Under the usual low-resolution conditions employed in COSY, the relative intensities of the intrasugar interproton cross-peaks can be used to assess qualitatively the magnitudes of the vicinal coupling constants involved. Such qualitative information about these coupling constants has been used to fix P in individual nucleotide units to a fair degree of precision [7,8,14–16]. More recently, we have described the use of the ω-scaled phase-sensitive COSY technique [17] for resolving the sugar geometry-dependent multiplet patterns of the COSY cross-peaks. This enabled us to quantify several intrasugar vicinal coupling constants, e.g.,

Correspondence address: K.V.R. Chary, Chemical Physics Group, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Bombay 400005, India

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/88/$3.50

$J(1'\text{-}2')$, $J(1'\text{-}2'')$ and $J(2''\text{-}3')$. Such information combined with the qualitative estimate of the other two couplings $J(2'\text{-}3')$ and $J(3'\text{-}4')$ obtained from the low-resolution magnitude-mode COSY spectrum has been used to fix P to a greater precision [9,10,18]. However, difficulties have been encountered when the linewidths are large and also when there is an extensive overlap of cross-peaks in the COSY spectra.

Here, we discuss a different approach to obtain details of the sugar geometries of individual nucleotide units in DNA fragments. This is based on the interpretation of quantitative intrasugar interproton distances derived from the pure-absorption mode ω_1-scaled NOESY spectrum recorded with a short mixing time (80 ms).

2. THE APPROACH

The conformation of the deoxyribofuranose ring can be determined from a knowledge of the intrasugar interproton distances. These distances can be calculated theoretically for different values of P (the pseudorotation phase angle) by making use of bond distances, bond angles and dihedral angles as input parameters. The intrasugar ring dihedral angles used in these calculations have been chosen from the conformation wheel in the pseudorotational representation for D-deoxyribose [19]. Fig.1 shows the plots of all the intrasugar interproton distances as a function of P. Similar plots have been reported by Wuthrich [20]. However, we found several differences between the magnitudes of distances that he reported and those in fig.1. For example, in our theoretical calculations, the $1'-4'$ distance is found to vary between 2.8 and 4.1 Å as against Wuthrich's reported distances which vary between 2.2 and 3.8 Å. We have taken care to check the reliability of our simulations by model building. From these plots it is evident that $H1'-H3'$ and $H2'-H4'$ intrasugar proton-proton distances are always greater than 3.7 Å for all the values of P and will not give rise to NOESY cross-peaks of appreciable intensities at short mixing times. The $1-2''$ and $2'-3'$ interproton distances are the shortest and vary within a narrow range of 0.2 Å. These give rise to strong NOESY peaks which are very insensitive with respect to P. Similarly, $1'-2'$, $2''-3'$ and $3'-4'$ distances are also insensitive to P. However, the $1'-4'$ and $2''-4'$ distances show a strong dependence on sugar pucker. These distances vary between 2.8 and 5.0 Å and can be used for establishing the sugar ring conformation. Also, they can be used to assess the possible existence of a fast equilibrium on the NMR time scale, such as that involving C2′-*endo* (or S-domain) and C3′-*endo* (or N-domain). For example, if the conformation of the sugar happens to be C3′-*endo* ($P = 18°$) one would observe a very strong $2''\text{-}4'$ NOESY cross-peak. Likewise, if it were to be C2′-*endo* ($P = 162°$) one would see neither a $1'\text{-}4'$ nor a $2''\text{-}4'$ cross-peak.

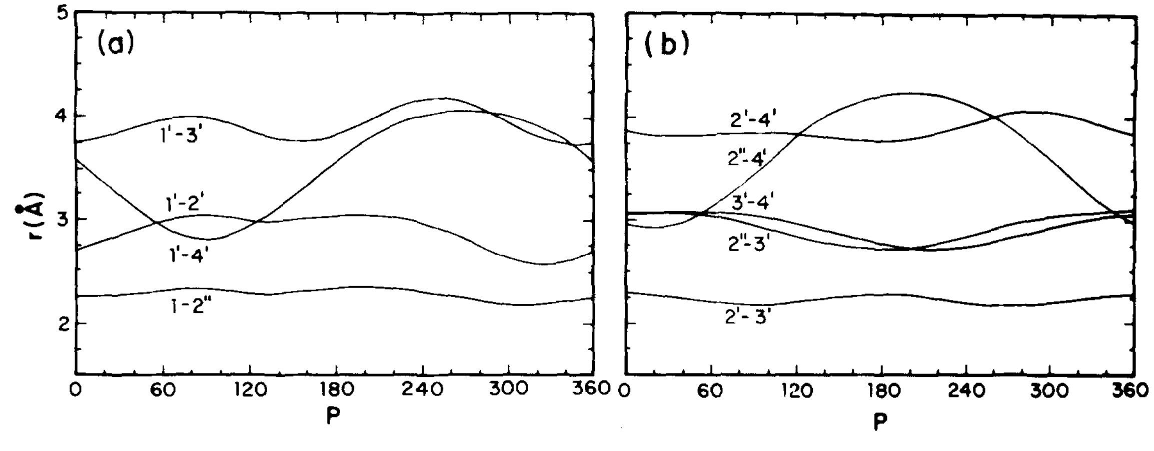

Fig.1. Calculated intrasugar ring interproton distances plotted as a function of pseudorotation phase angle (P). (a) $1'-2'$, $1'-2''$, $1'-3'$ and $1'-4'$ distances; (b) $2'-3'$, $2''-3'$, $2'-4'$, $2'-4''$ and $3'-4'$ distances. These have been calculated using a mononucleotide unit (dC) as a representative example. These have been calculated at regular intervals of 18° of the P value. For each input data set the sugar ring closure has been checked, which is an important factor in such simulations.

In the event of an equilibrium between 50% C3′-*endo* and 50% C2′-*endo* sugar puckers, the 1′-4′ cross-peak would be completely absent and the 2″-4′ cross-peak would be of very low intensity in the NOESY spectrum. Lastly if the sugar pucker were centred around O4′-*endo*, it is clear from the plots that the 1′-4′ cross-peak would be stronger than the 1′-2′ cross-peak.

We demonstrate here the reliability and utility of this approach with a specific example in which we have determined the sugar puckers using both procedures.

3. EXPERIMENTAL

All ^1H spectra of d-ACATCGATGT were recorded on a Bruker AM-500 FT NMR spectrometer equipped with an Aspect 3000 computer. COSY data were recorded using a $(RD{-}90{-}t_1{-}\Delta{-}90{-}\Delta{-}t_2)_n$ pulse sequence, where Δ is the fixed delay (5 ms) used for the enhancement of the cross-peaks [21,23]. The time-domain data set consisted of 2048 and 512 data points along the t_2 and t_1 axis, respectively. The data set was zero-filled to 1024 along the t_1 axis and was multiplied by sine bell and sine-square bell window functions along the t_1 and t_2 axis, respectively, before 2D Fourier transformation.

All other 2D NMR spectra were recorded in the phase-sensitive mode using the time-proportional phase-incrementation method [24]. A phase-sensitive ω_1-scaled COSY spectrum [17,25] was recorded with a shift-scaling factor (α) of 0.5 and J-scaling factor (γ) of 1.5. The time-domain data set consisted of 2048 and 450 data points along the t_2 and t_1 axis, respectively. The data set was zero-filled to 4096 and 2048 along the t_2 and t_1 axes and was multiplied by sine-bells (shifted by $\pi/16$) along both axes before 2D Fourier transformation. The spectral width and offset were optimised to derive maximum possible digital resolution by allowing a part of the base proton region to fold up to 6.4 ppm without affecting the H1′ proton region. The most down-field H1′ proton is found to be at 6.3δ. Essentially, this resulted in the selection of the spectral offset at δ = 4.37 ppm and a spectral width of 3086 Hz along the ω_2 axis (δ ranging from 1.3 to 7.44 ppm).

Table 1

Experimentally observed and theoretically expected intrasugar interproton distances (in Å) in d-ACATCGATGT at 25°C

Nucleotide unit	1′–2′	1′–2″	1′–3′	1′–4′	2″–4′
Experimental values					
A1	o	o	a	2.6(2.9)	a
C2	2.9(3.2)	2.4(2.7)	a	o	a
A3	2.9(3.2)	2.4(2.7)	a	3.1(3.4)	a
T4	2.8(3.2)	2.4(2.7)	a	2.7(3.0)	a
C5	2.7(3.0)	2.3(2.6)	a	o	a
G6	3.0(3.3)	2.5(2.8)	a	2.7(3.0)	a
A7	2.6(3.0)	2.3(2.6)	a	o	a
T8	2.8(3.1)	2.4(2.8)	a	2.7(3.0)	a
G9	o.	o	a	2.9(3.2)	a
T10	o	o	a	2.6(2.9)	a
Theoretical values					
C3′-*endo*	2.8	2.3	3.8	3.7	2.9
O4′-*endo*	3.0	2.3	4.0	2.8	3.5
C1′-*exo*	3.0	2.3	3.8	3.0	3.0
C2′-*endo*	3.0	2.3	3.8	3.4	4.2

a, peak absent (proton–proton distance > 4 Å); o, peaks overlap. Experimental distances are estimated with cytosine H6–H5 and thymine H6–CH3 (in parentheses) distances as reference scales. The error in the estimation is ± 0.2 Å

The absorption mode ω_1-scaled NOESY [17,25] spectrum of the decamer was recorded with 80 ms mixing time. Since all the components of individual cross-peaks are absorptive the α and β parameters were chosen as 0.5 and 0.6, respectively, to enhance the sensitivity and resolution of the spectra. The time-domain data set was collected with 1024 and 256 data points along the t_2 and t_1 dimensions. Such a data set was zero-filled to 2048 and 1024 data points along the t_2 and t_1 axes and multiplied by cosine-bell window functions before 2D Fourier transformation. Interproton distances from the ω_1-scaled

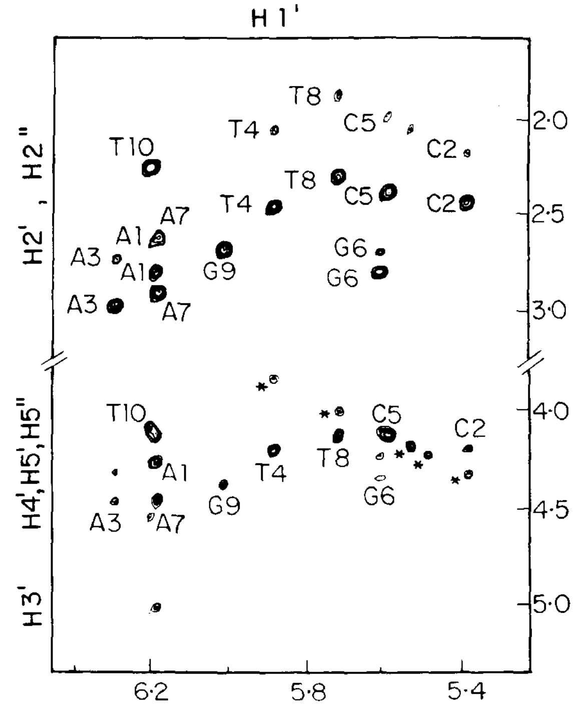

Fig.2. Expansion of selected regions of 500 MHz pure-absorption mode ω_1-scaled NOESY spectrum of d-ACATCGATGT in ^2H_2O at 25°C recorded with a mixing time of 80 ms. This figure contains all the expected NOESY cross-peaks from the H1′ proton to H2′, H2″ and H4′ protons. These portions were used to estimate the intrasugar interproton distances.

NOESY spectrum were obtained following the procedures described in [9,10].

All the above-mentioned NMR spectra were recorded in ^2H_2O solution at 25°C.

4. SUGAR RING CONFORMATION IN d-ACATCGATGT

Table 1 shows the theoretically expected and experimentally observed values of intrasugar proton-proton distances. Fig.2 shows a portion of the pure-absorption mode NOESY spectrum of decamer recorded with 80 ms mixing time. This depicts the NOESY cross-peaks from H1′ to H2′, H2″, H3′ and H4′ protons. The resonance assignments in this and subsequent figures have been obtained following well-established procedures, which will form the subject matter of a detailed paper. Qualitative assessment of these NOEs reveal that H1′-H2″ cross-peaks are stronger compared to other NOESY cross-peaks. Also, in most of the nucleotide units H1′-H4′ NOESY cross-peaks are stronger than those of H1′-H2′. Another important feature is the complete absence of H1′-H3′ peaks for all nucleotide units.

The experimentally estimated interproton distances which are given in table 1 are consistent with such qualitative assessment. The H1′–H2″ distances are found to be in the range 2.3–2.8 Å. The H1′–H2′ and H1′–H4′ distances are found to be in the range 2.6–3.1 and 2.6–3.4 Å, respectively. Further, the H1′–H3′ and H2″–H4′ distances are found to be greater than 4 Å, since they have not appeared in the NOESY spectrum.

From these observations, we can draw the following conclusions:

(i) The absence of H2″-H4′ NOESY cross-peaks rules out the possibility of the C3′-*endo* conformation for which the interproton distances would be 2.8 Å and should have given rise to a very strong peak (NOE).

(ii) The estimated H1′–H4′ distances, which are in the range 2.6–3.1 Å, indicate that the sugar puckers in all but C2, C5 and A7 nucleotide units

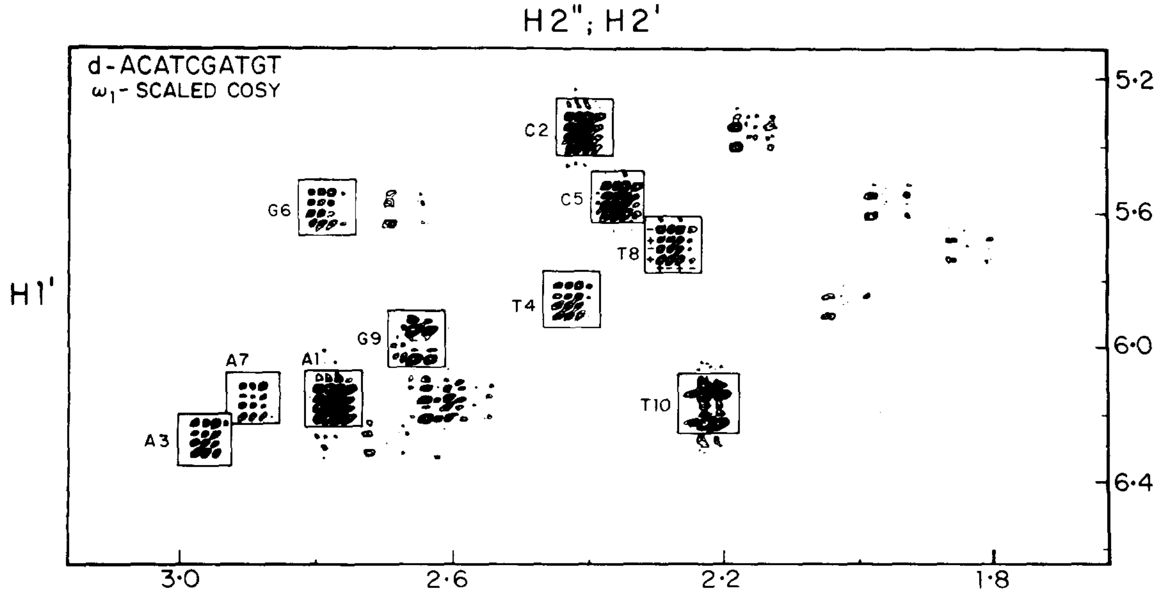

Fig.3. Selected region of 500 MHz ω_1-scaled phase-sensitive COSY spectrum of d-ACATCGATGT recorded in ^2H_2O at 25°C. This figure covers all the expected H1′-H2′ and H1′-H2″ cross-peaks. Boxed portions in this figure show the characteristic phase-sensitive multiplet patterns of H1′-H2″ cross-peaks belonging to different nucleotide units. These portions helped in the estimation of $J(1'\text{-}2')$, $J(1'\text{-}2'')$ and $J(2'\text{-}2'')$ values. Unboxed portions depict all the expected H1′-H2′ cross-peak multiple patterns. Although these could not be resolved into individual components, they could be used to estimate the value of $[J(2'\text{-}1') + J(2'\text{-}2'') + J(2'\text{-}3')]$. The digital resolution along the ω_1 and ω_2 axes is 1.5 Hz/pt.

are centred around O4′-*endo*. In the case of C2, C5 and A7 this distance could not be estimated because of overlapping resonances.

(iii) The estimated H1′–H4′ distances are less than or almost equal to the H1′–H2′ distances (table 1). This is also consistent with the earlier observation (ii).

(iv) The above observations rule out the possibility of C2′-*endo* sugar pucker for which the H1′–H4′ distance is 3.4 Å and greater than the H1′–H2′ distance which is 3.0 Å. In the present case, the estimated H1′–H4′ distances are less than those of H1′–H2′ for each individual nucleotide unit, except a few where there were problems due to overlap (table 1).

In studies on the solution conformation of non-helical mono- and dinucleotides, the NMR data have often been interpreted in terms of a dynamic equilibrium between C2′-*endo* (or S-domain) and C3′-*endo* (or N-domain) sugar puckers [11]. Here, the estimated H1′–H4′ distance (2.6–3.1 Å) cannot be explained by a time average behaviour between C3′-*endo* and C2′-*endo* conformations. In any such situation, the H1′–H4′ distance would be greater than 3.6 Å and would give rise to a very weak NOESY, while the experimentally estimated distance lies in the range 2.9–3.1 Å.

Thus, it seems that the molecular dynamics of the DNA segment d-ACATCGATGT is confined to a narrow region of the sugar pucker (in the S domain) centred around O4′-*endo*. Although in some of the nucleotide units certain intrasugar interproton distance estimates could not be carried out because of the overlap problems due to overlap, it is evident even in these cases (from qualitative assessment) that the sugar puckers are centred around O4′-*endo*.

To substantiate these observations we have performed an independent check using the ω_1-scaled phase-sensitive COSY technique. This technique has been used earlier for resolving the sugar-geometry-dependent characteristic multiplet patterns of the intrasugar COSY cross-peaks, mainly the H1′-H2′ and H1′-H2″ cross-peaks (fig.3). By making use of fig.3 we could quantify *J*(1′-2′), *J*(1′-2″), *J*(2′-3′) and *J*(2″-3′) within an error of ± 0.3 Hz following methods described elsewhere

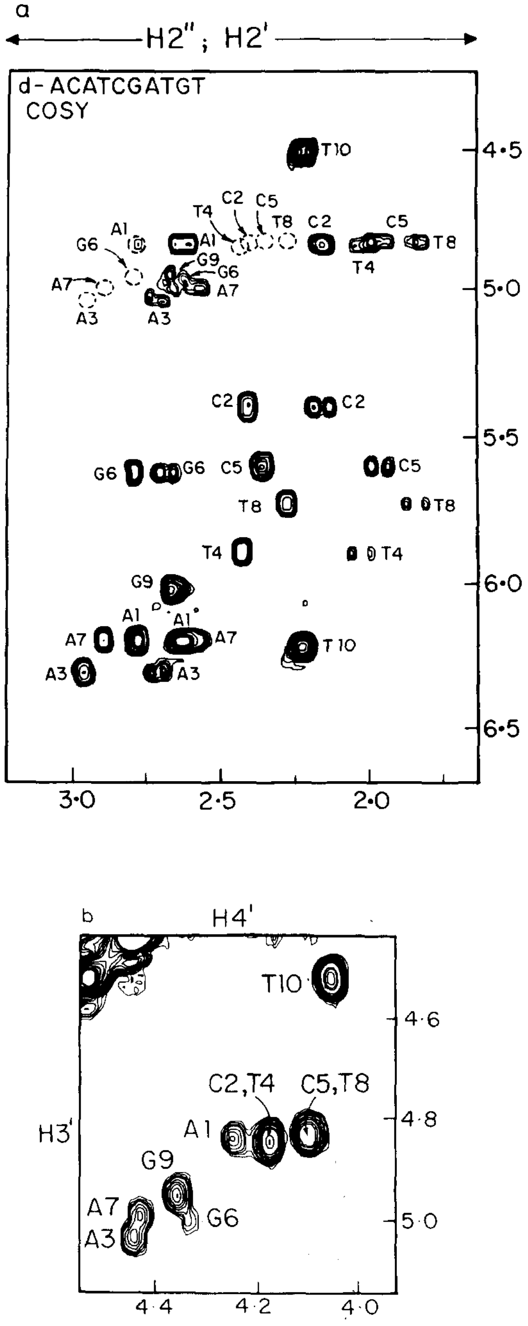

Fig.4. Expansion of selected portion of 500 MHz proton COSY spectrum of d-ACATCGATGT in ^2H_2O at 25°C. Panel (a) contains all the expected *J*-coupling correlations between 1′-2′, 1′-2″, 2′-3′ and 2″-3′. The absent and very weak 2″-3′ cross-peaks *J*(2″-3′) are shown with circles in their expected region. Panel (b) contains all the expected *J*-coupling correlations between 3′-4′.

Table 2

Experimentally observed and theoretically expected values of scalar coupling constants (J in Hz) in the sugar rings of d-ACATCGATGT at 25°C

Residue	1′-2′	1′-2″	2′-3′	2″-3′	3′-4′	2′-2″
(a) Experimental values						
A1	8.4	7.0	8.9	w; 3	m	−14.5
C2	8.3	6.6	8.2	a; 3	m	−14.2
A3	7.9	6.9	8.5	a; 3	m	−15.2
T4	8.2	7.2	9.6	a; 3	m	−14.3
C5	7.7	7.3	7.8	a; 3	m	−14.7
G6	8.9	6.9	8.8	a; 3	m	−14.0
A7	7.8	7.0	8.5	a; 3	m	−14.3
T8	7.3	7.1	9.1	a; 3	m	−14.4
G9	o	o	o	o	s	o
T10	o	o	o	o	s	o
(b) Theoretical values for selected conformations						
C3′-*endo*	0.0	7.2	6.0	9.5	9.0	−14.0
O4′-*endo*	7.0	8.0	9.5	3.0	7.0	−14.0
C1′-*exo*	9.0	6.0	8.5	0.0	3.0	−14.0
C2′-*endo*	9.0	6.0	6.0	0.0	0.0	−14.0

Estimates of J(3′-4′) are based on the intensity of the respective cross-peaks in the low-resolution COSY spectrum. a, peak absent (very small J); m, medium peak (moderately large J); s, strong (large J); o, peaks overlap; w, weak peak (small J). The theoretical values have been calculated using a Karplus-type relation [11]. The accuracy of experimental coupling constants measured is ± 0.3 Hz

[9,10]. Such information about the coupling constants combined with the qualitative estimate of the J(3′-4′) obtained from the low-resolution COSY spectrum (fig.4) has been used to identify the P values in all the individual nucleotide units. Table 2 lists the experimentally observed and theoretically expected values of vicinal coupling constants in the sugar rings of d-ACATCGATGT. Comparison of the theoretical and experimental coupling constants reveals that all the sugars belonging to A1-T8 lie between O4′-*endo* (P = 90°) and C1′-*exo* (P = 126°). In the case of G9 and T10 the situation is complicated by the near equivalence of the H2′ and H2″ protons. However, the general behaviour of the intensity patterns leads us to believe that the conformation in these two cases is in the range P = 70–90° [8–10].

Thus, we have demonstrated that experimentally estimated intrasugar interproton distances could be used to fix the sugar geometries in individual nucleotide units of DNA fragments.

Acknowledgements: The facilities provided by the 500 MHz FT-NMR National Facility, supported by the Department of Science and Technology, Government of India, are gratefully acknowledged. We thank Dr H.T. Miles (National Institutes of Health, Bethesda, MD) for the sample of oligonucleotide. We thank Professor Girjesh Govil and Dr R.V. Hosur of the Chemical Physics Group (Tata Institute of Fundamental Research, Bombay) for helpful discussions.

REFERENCES

[1] Feigon, J., Wright, J.M., Leupin, W., Denny, W.A. and Kearns, D.R. (1982) J. Am. Chem. Soc. 104, 5540–5541.

[2] Feigon, J., Denny, W.A., Leupin, W. and Kearns, D.R. (1983) Biochemistry 22, 5943–5951.

[3] Frechet, D., Cheng, D.M., Kan, L.S. and Ts'O, P.O.P. (1983) Biochemistry 22, 5194–5200.

[4] Hare, D.R., Wemmer, D.E., Chu, S.H., Drobny, G. and Reid, B.R. (1983) J. Mol. Biol. 171, 319–336.

[5] Hosur, R.V., Kumar, M.R., Roy, K.B., Zu-Kun, T., Miles, H.T. and Govil, G. (1985) in: Magnetic Resonances in Biology and Medicine (Govil, G. et al. eds) pp.243–260, Tata McGraw-Hill, New Delhi.

[6] Clore, G.M. and Gronenborn, A.M. (1985) EMBO J. 4, 829–835.

[7] Hosur, R.V., RaviKumar, M., Chary, K.V.R., Sheth, A., Govil, G., Zu-Kun, T. and Miles, H.T. (1986) FEBS Lett. 205, 71–76.

[8] Chary, K.V.R., Hosur, R.V., Govil, G., Zun-Kun, T. and Miles, H.T. (1987) Biochemistry 26, 1315–1322.

[9] Chary, K.V.R., Hosur, R.V., Govil, G. and Miles, H.T. (1988) Biochemistry, in press.

[10] Hosur, R.V., Chary, K.V.R., Sheth, A., Govil, G. and Miles, H.T. (1988) J. Biosci., in press.

[11] Govil, G. and Hosur, R.V. (1982) Conformation of Biological Molecules: New Results from NMR, Springer, Heidelberg.

[12] Saenger, W. (1984) Principle of Nucleic Acid Structure, Springer, New York.

[13] Altona, C. and Sunderlingam, M. (1972) J. Am. Chem. Soc. 94, 8205–8212.

[14] Hosur, R.V., Sheth, A., Chary, K.V.R., Ravi-Kumar, M., Govil, G., Zu-Kun, T. and Miles, H.T. (1986) Biochem. Biophys. Res. Commun. 139, 1222–1232.

[15] Sheth, A., RaviKumar, M., Hosur, R.V., Govil, G., Zu-Kun, T. and Miles, H.T. (1987) Biochem. Biophys. Res. Commun. 144, 26–35.

[16] Sheth, A., RaviKumar, M., Hosur, R.V., Govil, G., Zu-Kun, T., Roy, K.B. and Miles, H.T. (1987) Biopolymers 26, 1301–1313.

[17] Hosur, R.V., Chary, K.V.R., Majumdar, A. and Govil, G. (1988) Life Sci. Adv. Biochem., in press.

[18] Chazin, W.J., Wuthrich, K., Hyberts, S., Rance, M., Danny, W.A. and Leupin, W. (1986) J. Mol. Biol. 19, 439–453.

[19] Saran, A., Perahia, D. and Pullman, B. (1973) Theor. Chim. Acta 30, 31.

[20] Wuthrich, K. (1986) NMR of Proteins and Nucleic Acids, Wiley, New York.

[21] AnilKumar, Hosur, R.V. and Chandrasekhar, K. (1984) J. Magn. Reson. 60, 143–148.

[22] Hosur, R.V., Chary, K.V.R., Kumar, A. and Govil, G. (1985) J. Magn. Reson. 62, 123–127.

[23] Wynants, C. and Van Binst, G. (1984) Biopolymers 23, 1799–1804.

[24] Marion, D. and Wuthrich, K. (1983) Biochem. Biophys. Res. Commun. 113, 967–974.

[25] Hosur, R.V., Chary, K.V.R. and RaviKumar, M. (1985) Chem. Phys. Lett. 116, 105–108.