Solution structure of d-GAATTCGAATTC by 2D NMR

A new approach to determination of sugar geometries in DNA segments

R.V. Hosur, M. Ravikumar, K.V.R. Chary, Anu Sheth, Girjesh Govil, Tan Zu-Kun* and H. Todd Miles*

Chemical Physics Group, Tata Institute of Fundamental Research, Homi Bhabha Road, Bombay 400005, India and **National Institutes of Health, Bethesda, MD 20205, USA*

Received 28 May 1986

A new approach based on the correlated spectroscopy (COSY) in 2D NMR has been described for determination of sugar geometries in oligonucleotides. Under the usual low resolution conditions employed in COSY, the intensities of cross peaks depend on the magnitudes of coupling constants. There are five vicinal coupling constants in a deoxyribose ring which are sensitive to the sugar geometry. The presence, absence and rough comparison of relative intensities of COSY cross peaks arising from such coupling constants enable one to fix the sugar conformation to a fair degree of precision. The methodology has been applied to d-GAATTCGAATTC. It is observed that ten out of the twelve nucleotide units in this sequence exhibit a rare O1′-endo geometry. The *Eco*RI cleavage sites (between G and A) in the dodecanucleotide show an interesting variation in the conformation with the two sugars attached to the Gs acquiring a geometry between C2′-endo and C4′-endo.

2D NMR *Oligonucleotides solution structure* *Sugar geometry*

1. INTRODUCTION

The three-dimensional structure of nucleic acids is fixed by the sugar geometry, the glycosidic bond rotation χ and the conformation of the sugar-phosphate backbone [1–4]. The conformation of the deoxyribose ring thus plays a central role in the secondary structures of DNA. Models of the relatively better characterised DNA structures – A, B, and Z are based on C3′-endo (^{3}E), C2′-endo (^{2}E) or an alternation, such as sugar puckers. However, significant deviations from the above geometries have been detected in single crystal X-ray diffraction studies. An elegant and more general description of the five membered furanose ring is based on the pseudorotation concept [5]. Here, the torsional angles in the sugar ring are described in terms of the pseudorotation parameter P and the maximum sugar pucker (T_m). The correspondence between P and the ring conformations is indicated in fig.1. 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 = 90–270° covering conformations such as C2′-endo, C3′-exo, etc.) to distinguish between the two families of energy favored conformations.

Information on the solution structures of nucleic acids can be obtained using two-dimensional Fourier transform (2D FT) NMR techniques; strategies for structure determination using 2D NOESY techniques have been discussed earlier' [6–10]. In this paper we discuss new strategies which can be used to obtain details about the sugar ring conformations of individual nucleotide units in DNA segments. The ideas are demonstrated with d-GAATTCGAATTC as an example.

2. STRATEGIES

Conformation of deoxyribose ring can be deter-

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

mined by making use of the 3-bond coupling constants (^{3}J) between the various protons in the ring [1,2]. There are five ^3J values, namely H1′-H2′, H1′-H2″, H2′-H3′, H2″-H3′ and H3′-H4′, which are related to the relevant H-C-C-H dihedral angle ϕ, according to the relation [3].

$$J = 10.2\cos^2\phi - 0.8\cos\phi$$

The above dihedral angles are inter-dependent and their values can be calculated in terms of the two pseudorotation parameters, P and T_m. T_m is a constant for the d-ribose and thus various geometries can be expressed in terms of P.

Fig.1 shows the plots of the five coupling constants in a deoxyribose ring as a function of P. In these calculations, a value of 38° has been taken for T_m. Calculations have also been performed for different values of T_m. The effect of a larger T_m is to increase the amplitudes of the curves. The positions of maxima and minima are unaltered. It is clear from the curves that the values of coupling constants H1′-H2″ and H2′-H3′ vary within a narrow range of 6–10 Hz and are comparatively insensitive to the sugar geometry. On the other hand,

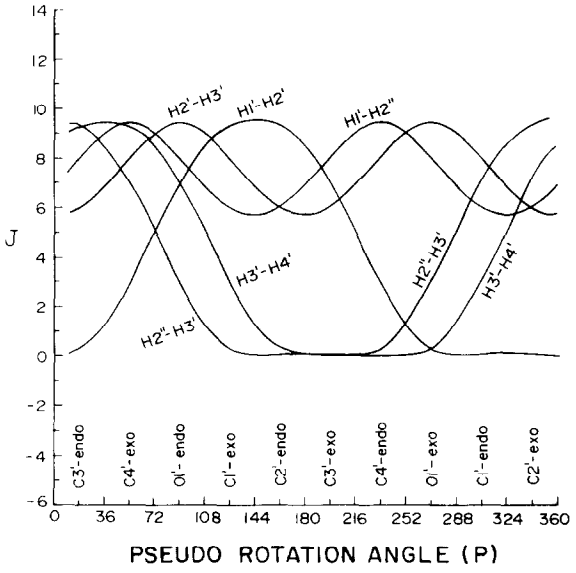

Fig.1. Plots showing the dependence of the 3-bond H1′-H2′, H1′-H2″, H2′-H3′, H2″-H3′ and H3′-H4′ coupling constants on the pseudorotational variable P. The positions of the classical sugar puckers are also indicated in the figure [5].

values of H2″-H3′, H3′-H4′ and H1′-H2′ coupling constants vary in the range 0–10 Hz and can be utilised with greater advantage in fixing the domains of sugar geometries.

In the case of oligonucleotides, measurement of ^3J from the one-dimensional spectrum becomes difficult because of extensive overlap of the multiplets corresponding to the sugar protons. Two-dimensional J-resolved spectroscopy, in which the J information is spread along the second frequency (ω_1) axis, while the usual axis (ω_2) contains exclusively the chemical shift information, can be helpful. However, this technique suffers from sensitivity problems due to T_2^* relaxation during the evolution period (t_1). For oligonucleotides dissolved in aqueous solutions, T_2^* values are generally less than 0.3 s, while typical values for t_1^{max} lie in the range 0.5–1 s. Consequently, magnetization decays substantially before the detection period starts leading to poor S/N ratio. Moreover, when the chemical shifts of the two protons are identical, the measured J values cannot be assigned unambiguously. Further, even if the H1′ protons are well resolved, assignment of coupling constants H1′-H2′ and H1′-H2″ is difficult, since an unambiguous assignment of the coupling constants requires clear resolution in the H2′, H2″ region as well. These problems render the J-resolved experiment less useful.

A possible solution to these problems is to obtain information through the COSY technique which has higher sensitivity and where the chemical shift information is retained along both the axes. Therefore, the chemical shift separation on either of the two axes is sufficient to separate the cross peaks. For example, even when the H1′ protons of two nucleotides have identical chemical shifts, they can be identified if the chemical shifts of the corresponding H2′/H2″ protons are different. To measure J values precisely from the COSY spectrum, experiments have to be performed with very high resolution, so that the multiplet components can be resolved in the individual cross peaks. This places demanding requirements in terms of storage space and instrument time. However, we show below that a careful analysis of even the normal low resultion COSY spectrum depicting J connectivities between H1′, H2′, H2″, H3′ and H4′ protons can yield valuable information about sugar geometry.

The basic idea is that, under the usual low resolution conditions for recording COSY spectra, the intensities of the cross peaks depend directly on the magnitudes of the coupling constants. This is due to the fact that the components of cross peaks have anti-phase character and tend to cancel each other when the resolution is not enough to resolve the J separation between them. It may be mentioned here that for a given J value, the cancellation depends on two factors: (i) T_2^* or the linewidths; (ii) the digitial resolution along the ω_1 axis. In 2D spectroscopy, the second factor is more important, since an attempt to improve resolution places serious demands on the expensive instrument time. The above aspects have been discussed along with our recent proposals dealing with J scaling [11] and SUPERCOSY [12]. In this paper, we make use of the dependence of intensities of the cross peaks on J values for the determination of sugar geometries in oligonucleotides.

It is clear from the above discussion that depending upon the sugar geometry, certain cross peaks in the COSY spectrum will be more prominent than others and the peaks corresponding to low J values may even be absent. Fig.2 shows typical spectra expected for C3′-endo, C2′-endo and O1′-endo geometries. The C3′-endo sugar geometry gives rise to strong H1′-H2″ peaks ($J \sim 10$ Hz) but extremely weak H1′-H2′ cross peaks ($J \sim 0$ Hz; in most cases these peaks may be absent). The H3′-H4′ ($J \sim 10$ Hz) cross peaks will also be strong. In the case of C2′-endo sugar geometry, the H1′-H2′ cross peak would be slightly stronger compared to H1′-H2″ cross peak and the H3′-H4′ cross peak would be weak or absent ($J \sim 0$ Hz). For O1′-endo geometry, all the cross peaks except H2″-H3′ will be strong. Thus, even at a rough level of quantification of cross peak intensities, important conclusions about the sugar geometry can be made.

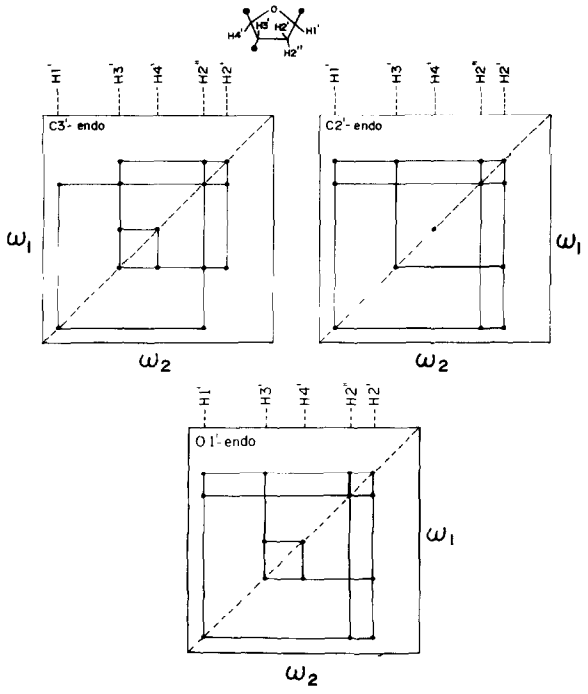

Fig.2. Schematic COSY spectra of d-ribose for 3 different geometries.

The above discussion with respect to the COSY is also applicable to correlated spectroscopy with shift scaling (COSS), a technique described recently [13]. This technique achieves shift scaling by a factor α and gives a better resolution along the ω_1 axis of the 2D spectrum. It can be used whenever problems of peak overlap are present in the COSY spectrum. In such spectra, the cross peaks do not show prominent multiplicity differences and thus it is easier to compare the intensities. However, it should be remembered that in the COSS spectrum, the phase characteristics of the cross peak components are modified by an extent which depends on J.

The above strategy for determination of sugar geometry requires that the H2′ and H2″ protons should be properly identified in the COSY spectrum. Some authors have used the multiplet patterns of the cross peaks to identify these protons [14]. Since the multiplet pattern depends on the coupling constants involved, which in turn depend on the sugar geometry, we feel that multiplicity is not a good criterion for the identification of the H2′ and H2″ protons. We feel that the intensity of NOESY cross peaks H1′-H2′ and H1′-H2″ at short mixing times is a better criterion and multiplicity can then be a good internal check for the sugar geometry. H1′-H2″ cross peaks are expected to be stronger than the H1′-H2′ cross peaks in the NOESY spectrum, since the distance between the former set of protons is much shorter than that between the latter.

An alternative 2D technique for identification of

sugar geometries applicable only for right handed DNA is based on the sequential (H8/H6)→(H2′,H2″)$_{i-1}$ cross peaks in the NOESY spectrum. If the $(i-1)$th nucleotide has a C2′-endo geometry, both the base to H2′ and H2″ sequential connectivities are observable. For C3′-endo, only the (H8/H6)$_i$−(H2′)$_{i-1}$ connectivity is observed. This criterion can be used to decide whether the conformation is close to ^{3}E (N domain) or to ^{2}E (S domain).

While both the strategies discussed above can be used in practice, the one based on COSY has several advantages. There are no complications such as spin diffusion which hamper the interpretation of NOESY spectra. The sugar geometries can be identified with higher precision than just classifying them in the N or S domain. The method is independent of the overall DNA structure. Finally, sugar geometries can be associated with individual spin systems and resonance assignment is not a prerequisite. Notwithstanding the relative merits, both COSY and NOESY strategies can provide important information and should be used as complementary tools.

3. EXPERIMENTAL

Oligonucleotides were synthesised as described elsewhere [7]. All NMR experiments were carried out on a Bruker AM-500 FT-NMR spectrometer operating at 500 MHz for ^{1}H. COSY and NOESY experiments were performed with 512 t_1 and 2048 t_2 points. The time domain data were zero filled to 1024 points along the t_1 axis. The data were multiplied by sine square bell and sine bell window functions along the t_2 and t_1 axis, respectively, prior to respective Fourier transformations. Digital resolution in all the spectra is 7.5 Hz/point along both ω_1 and ω_2 axes. Absolute value spectra have been presented.

4. SUGAR RING GEOMETRIES IN d-GAATTCGAATTC

Following the principles discussed above, we have determined sugar geometries in several oligonucleotides namely d-CGCGCGCGCGCG, d-GAATTCGAATTC and d-GAATTCCCGAATTC. Here, we illustrate the methodology using the double-helical dodecamer (I) as an example.

	1	2	3	4	5	6	7	8	9	10	11	12	
5′−	G	A	A	T	T	C	G	A	A	T	T	C	−3′
3′−	C	T	T	A	A	G	C	T	T	A	A	G	−5′

(I)

In the first step, the H2′ and H2″ protons have to be distinguished using the NOESY spectrum. Fig.3 shows the H1′-H2′ and H1′-H2″ cross peak region of the NOESY spectrum of I. It may be mentioned that the complete assignments in this and subsequent figures have been obtained following the sequential assignment procedures. It is not our intention to give these details here, since they will form the subject matter of a detailed paper. In fig.3, one observes cross peaks corresponding to both H1′-H2′ and H1′-H2″ for all residues except C12, where due to the small chemical shift difference between H2′ and H2″, the two cross peaks overlap. For each pair one cross peak is stronger than the other. The stronger peak is attributed to H1′-H2″ correlation and the weaker peak to H1′-H2′ NOE. The H2″ proton is always found to resonate downfield of H2′ proton.

Fig.4 shows a section of the COSY spectrum of I which contains the H1′-H2′ and H1′-H2″ cross peaks. Fig.5 shows sections corresponding to the (H2′, H2″)-H3′ and H4′-H3′ cross peak

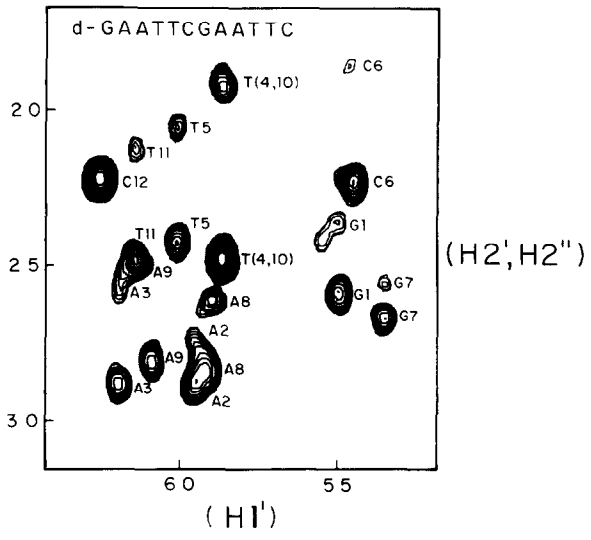

Fig.3. A section of NOESY spectrum of I showing H1′-H2′ and H1′-H2″ cross peaks. Mixing time 300 ms.

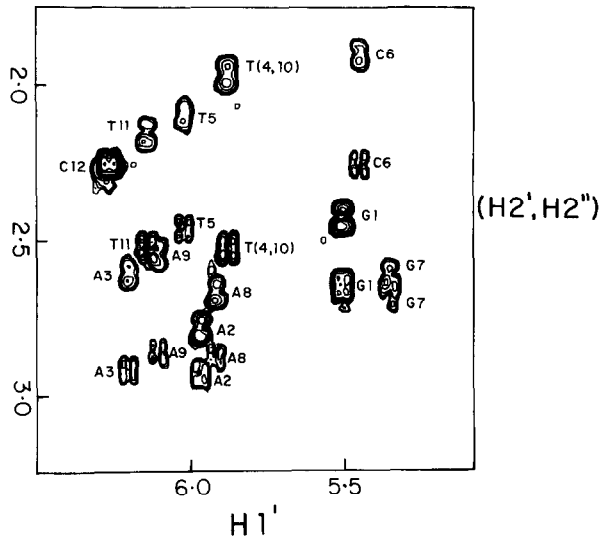

Fig.4. A section of the 500 MHz COSY spectrum of d-GAATTCGAATTC showing H1′-H2′, H1′-H2″ cross peaks.

regions. We shall first discuss the general features of the 2D spectra of residues A2, A3, T4, T5, C6, A8, A9, T10 and T11 which are very similar.

(i) For each residue, two distinct cross peaks corresponding to H1′-H2′ and H1′-H2″ are observed in the COSY spectrum (fig.4). Both cross peaks have similar intensities. In every case, the downfield peak (due to H1′-H2″) exhibits a finer multiplet pattern compared to the upfield peak (due to H1′-H2′). The presence of H1′-H2′ peaks rules out sugar conformations with $P > 240°$ and $P < 30°$ (fig.1).

(ii) In fig.5, all the H2″-H3′ cross peaks are absent. This shows that the value of J(H2″-H3′) is very small. This observation is consistent with the difference in multiplet structures in the (H1′-H2′, H2″) domain i.e. H1′-H2″ cross peaks show a better resolution compared to H1′-H2′ peaks. It also restrict the sugar conformation to the domain $P = 90$ to 250° (fig.1).

(ii) The H4′-H3′ peaks are observed for these nucleotides. It may be noticed from fig.1 that the near zero value of J(H2″-H3′) and finite value of J(H3′-H4′) restrict the P value to a very narrow region of the S domain of sugar pucker i.e. between C1′-exo and O1′-endo. The near equal intensities of the H1′-H2′ and H1′-H2″ cross peaks further imply that the conformation is closer to O1′-endo rather than C1′-exo.

The H2′ and H2″ protons of C12 have almost equal shifts. This is expected, since the C12 lies at the 3′-end which does not have a phosphate group. This results in a single COSY cross peak for the H1′-H2′ and H1′-H2″ pairs. The intensity of this peak is much higher than other cross peaks in fig.4 suggesting substantial contributions from both the H1′-H2′ and H1′-H2″ correlations. Further, the presence of substantial intensity for the H4′-H3′ COSY cross peak indicates that in this case also the geometry is close to O1′-endo.

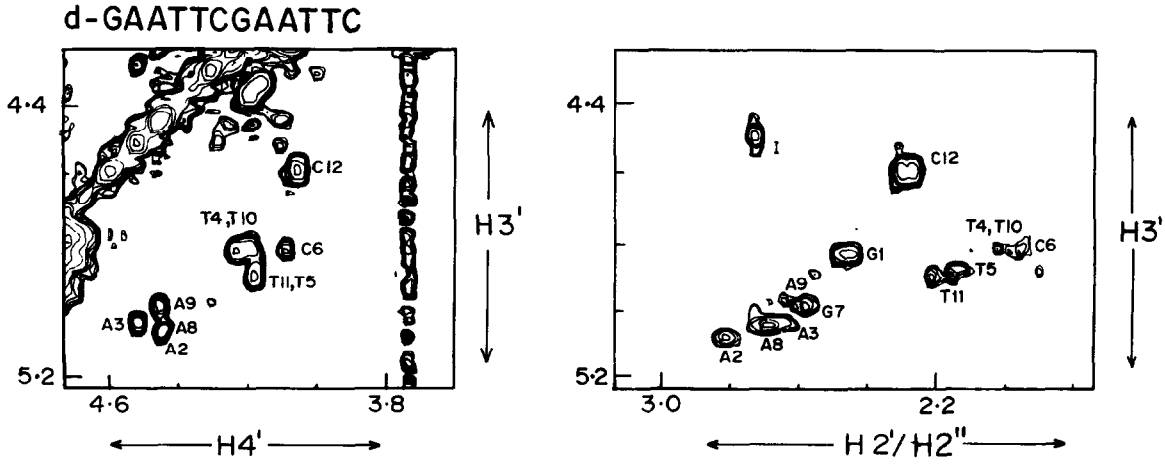

Fig.5. Sections of 500 MHz COSY spectrum of d-GAATTCGAATTC showing (H2′, H2″)-H3′ (right) and H3′-H4′ cross peak (left) regions. All the twelve H2′-H3′ connectivities are seen but the H2″-H3′ cross peaks are absent.

The behaviour of G1 and G7 is significantly different. In these cases, we do not observe H4′–H3′ cross peaks. Further, we observe sequential NOESY connectivities from H8 of A2 to both H2′ and H2″ of G1 and from H8 of A8 to both H2′ and H2″ of G7. Besides, in these two cases as well, we do not observe H2″–H3′ connectivities. This implies that the conformation of G1 and G7 lies in the domain P = 160 to 240° i.e. in the S domain, between C2′-endo and C4′-endo.

We have thus demonstrated that even with a low resolution (~7–8 Hz/point) COSY spectrum, one can obtain valuable and fairly accurate information about the sugar geometry. A very high resolution (~0.5–1 Hz/point) spectrum will of course enable measurement of coupling constants and thus improve upon the precision of the conclusions. However, we do not expect any gross changes in the conclusions, since the present analysis is based on a simultaneous use of five different coupling constants, and even an approximate knowledge of the five values will yield highly reliable results.

It is important to note that in d-GAATTCGAATTC, all the nucleotide units except G1 and G7 show a rare O1′-endo geometry. G1 and G7 exhibit a conformation between C2′-endo and C4′-endo. The interesting feature about the variation in sugar geometry is that it occurs at exactly the cleavage sites of the restriction enzyme *Eco*RI, which cleaves between the G and A units.

ACKNOWLEDGEMENT

The facilities provided by 500 MHz FT-NMR National Facility are gratefully acknowledged.

REFERENCES

[1] Govil, G. and Hosur, R.V. (1982) Conformation of Biological Molecules: New Results from NMR, Springer, Heidelberg.
[2] Saenger, W. (1984) Principles of Nucleic Acid Structure, Springer, New York.
[3] Davies, D.B. (1985) Prog. Nucl. Magn. Reson. Spectrosc. 12, 135–225.
[4] Guschlbauer, W. (1976) Nucleic Acid Structures, Springer, Heidelberg.
[5] Altona, C. and Sunderalingam, M. (1972) J. Am. Chem. Soc. 94, 8205–8212.
[6] Hosur, R.V., Ravikumar, M., Roy, K.B., Tan Zu-Kun, Miles, H.T. and Govil, G. (1985) in: Magnetic Resonance in Biology and Medicine (Govil, G. et al. eds) pp. 243–260, Tata McGraw-Hill, New Delhi.
[7] Ravikumar, M., Hosur, R.V., Roy, K.B., Miles, H.T. and Govil, G. (1985) Biochemistry 24, 7703–7711.
[8] Clore, G.M. and Gronenborn, A.M. (1985) FEBS Lett. 179, 187–198.
[9] Gronenborn, A.M. and Clore, G.M. (1985) Prog. Nucl. Magn. Reson. Spectrosc. 17, 1–32.
[10] Kearns, D.R. (1984) Biochemistry 23, 791–793.
[11] Hosur, R.V., Chary, K.V.R. and Ravkumar, M. (1985) Chem. Phys. Lett. 116, 105.
[12] Anil Kumar, Hosur, R.V. and Chandrasekhar, K. (1984) J. Magn. Reson. 60, 143–148.
[13] Hosur, R.V., Ravikumar, M. and Anu Sheth (1985) J. Magn. Res. 65, 375–381.
[14] Hare, D.R., Wemmer, D.E., Chou, S.H., Drobny, G. and Reid, B.R. (1983) J. Mol. Biol. 171, 319–336.