Vol. 144, No. 1, 1987 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
April 14, 1987 Pages 26-34

TWO DIMENSIONAL NMR STUDIES ON THE SOLUTION STRUCTURE OF d-CTCGAGCTCGAG

Anu Sheth, M. Ravikumar, R.V. Hosur, Girjesh Govil,

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

and

Tan-Zu-Kun and H.T. Miles

National Institutes of Health, Bethesda, Maryland 20205

Received February 10, 1987

Two dimensional (2D) FT-NMR investigations have been carried out on the self-complementary dodecanucleotide d-CTCGAGCTCGAG, which has cleavage sites for the restriction enzyme Xho I (between C and T). The central TCG portion is also known to show a preference for DNAase activity. Complete resonance assignments have been obtained for the non-exchangeable sugar and base protons of the oligonucleotide. Information regarding sugar geometries, glycosidic torsion angles and other structural parameters has been obtained from the relative intensities of the cross peaks in the COSY and NOESY spectra. The results indicate that deoxyribose rings of C1 and C7 adopt a conformation different from the remaining sugars in the double helical oligonucleotide. The central TCG portion also exhibits variations in the backbone structure. The base stacking in the double helix shows interesting sequence dependent effects suggesting that the sequence effects are not localised to nearest neighbours but extended over longer stretches.

We have recently investigated by NMR, the structures of two oligonucleotides, d-GGATCCGGATCC [1,2] and d-GAATTCGAATTC [3] which have recoginition sites for restriction enzymes Bam H1 and Eco RI, respectively. In both cases, it has been observed that the cleavage sites of the restriction enzymes exhibit conformational deviations from the rest of the molecule. It is of interest therefore, to check if such conformational heterogeniety is a characteristic feature of DNA segments recognised by restriction enzymes. We have therefore undertaken investigations on a number of DNA segments conta-

Abbreviations: COSY , Correlated Spectroscopy;
NOESY , Nuclear Overhauser Effect Correlated Spectroscopy;
FT-NMR , Fourier Transform Nuclear Magnetic Resonance

0006-291X/87 $1.50

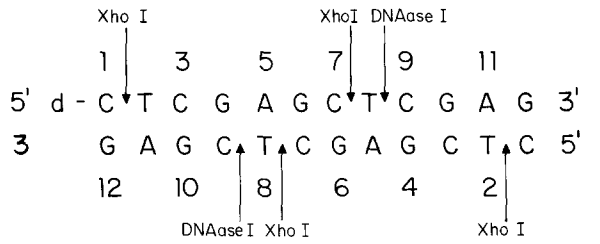


Fig. 1. The schematic representation of the self-complementary dodecanucleotide d-CTCGAGCTCGAG studied in this paper. The cleavage sites of Xho I and DNase I in the DNA segments are marked by arrows.

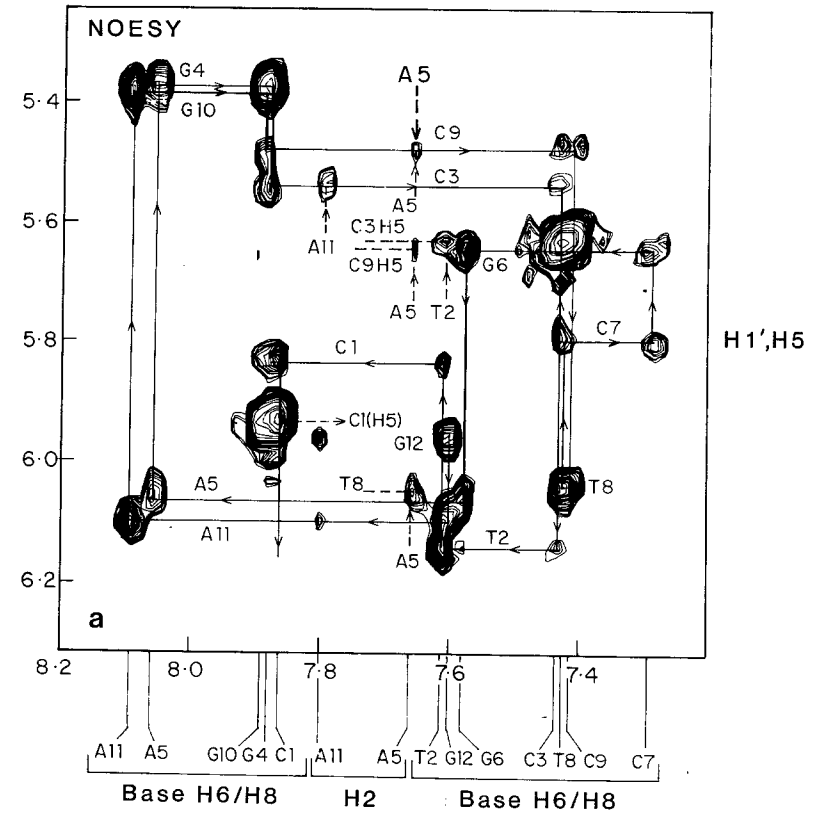
ining specific cleavage sites for restriction enzymes. In this paper we report, ^1H resonance assignments and some structural details on d-CTCGAGCTCGAG which has two cleavage sites (between C and T) for Xho I enzyme (Fig. 1).

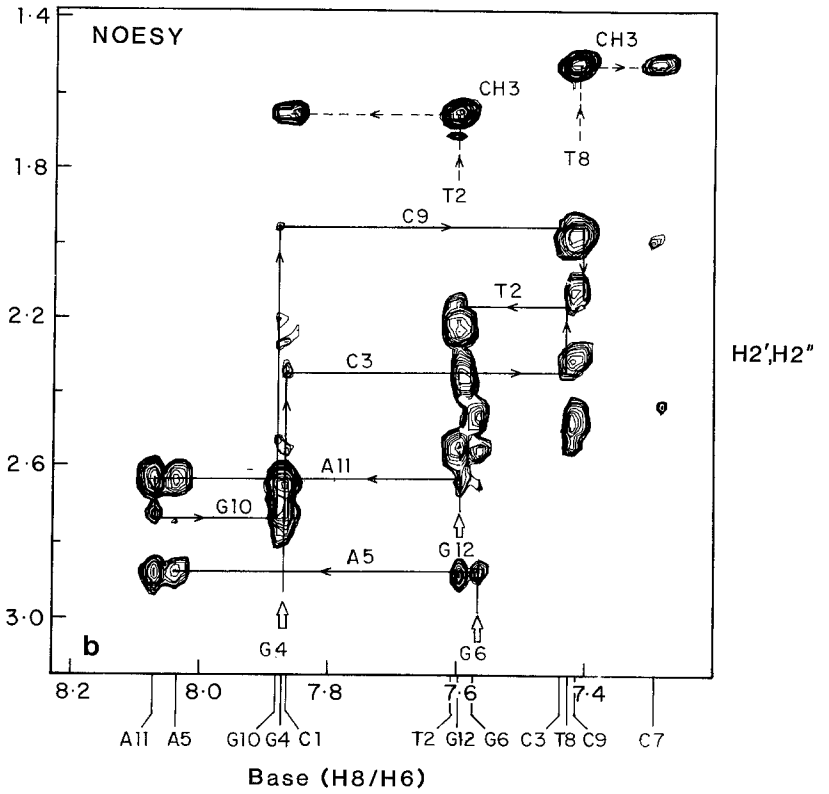
EXPERIMENTAL

Deoxy - (CTCGAGCTCGAG) was synthesised following procedures similar to those described previously [2]. 5 mg of the sample was dissolved in 0.5 ml of phosphate buffer (0.02M) having pH 7.2 and the total counterion concentration of 0.08 M. The solution was lyophilized and redissolved in D_2O twice and finally made up to 0.5 ml with 100% D_2O. ^1H NMR experiments were carried out using a Bruker AM 500 FT-NMR spectrometer operating at 500 MHz frequency for proton. Two dimensional COSY [4] and NOESY [5] experiments were carried out with 2048 data points along the t_2-axis and 512 data points along the t_1-axis. The COSY pulse sequence ($90°-t_1-90°-t_2$) was modified by introducing a small fixed delay of 5 msecs. after each $90°$ pulse, so as to enhance cross peak intensities in the two dimensional spectrum [6,7]. The NOESY sequence used was $90°-t_1-90°-\tau_m-90°t_2$ where τ_m is the mixing time. A relaxation delay of 1 sec was given in all 2D experiments. The time domain data were zero filled to 2048 points along the t_2-axis, and 1024 points along the t_1-axis and multiplied by sine-square bell and sine bell window functions prior to Fourier transformation along the t_2- and t_1-axis, respectively. Chemical shifts (Table 1) are expressed with respect to sodium 3-trimethylsilyl (2,2,3,3-^2H) propionate (TSP).

Table 1. Chemical Shifts of Non Exchangeable Protons in d-CTCGAGCTCGAG

Base	H6/H8	H1'	H2'	H2"	H3'	H4'	H5/H2/CH_3
C1	7.87	5.85	2.24	2.59	4.66	4.07	5.92
T2	6.62	6.15	2.23	2.55	4.89	4.24	1.66
C3	7.44	5.55	2.02	2.34	4.85	4.07	5.62
G4	7.88	5.38	2.65	2.72	4.97	4.28	
A5	8.06	6.05	2.64	2.91	5.09	4.40	7.66
G6	7.58	5.65	2.47	2.59	4.93	4.32	
C7	7.30	5.81	2.00	2.46	4.63	4.17	5.14
T8	7.43	6.05	2.15	2.50	4.85	4.18	1.52
C9	7.41	5.49	1.97	2.31	4.83	4.05	5.64
G10	7.89	5.39	2.65	2.72	4.97	4.28	
A11	8.10	6.10	2.64	2.91	5.02	4.40	7.80
G12	7.60	5.96	2.38	2.25	4.60	4.14	


NOESY
5·4
5·6
5·8
6·0
6·2
G4
G10
A5
C9
C3
A5
C3H5
A11
C9H5
G6
A5 T2
C7
C1
C1(H5)
G12
A5
T8
T8
A11
T2
A5
H1',H5
a
8·2
8·0
7·8
7·6
7·4
A11 A5 G10 G4 C1 A11 A5 T2 G12 G6 C3 T8 C9 C7
Base H6/H8
H2
Base H6/H8


NOESY
1·4
1·8
2·2
2·6
3·0
CH3
CH3
T8
T2
C9
T2
C3
A11
G10
A5
G12
G4
G6
H2',H2"
b
8·2
8·0
7·8
7·6
7·4
A11 A5 G10 G4 C1 T2 G12 G6 C3 T8 C9 C7
Base (H8/H6)

RESONANCE ASSIGNMENTS

Individual assignments of ^1H resonance have been obtained using 2D NMR techniques following the strategies described earlier [1-2,8-11]. First, the COSY and NOESY spectra are used to identify the subsets of J-coupled protons corresponding to the individual sugar residues. The cytosine and the thymine base protons are identified using J(H5,H6) and J(CH_3,H6) correlations. In the next stage, using the NOE's between base protons (H6,H8,H5,CH_3) and the sugar resonances H1',H2' and H2", these sets are assigned to individual units. Four different types of sequential correlations have been observed namely, $(base)_i \longrightarrow (base)_{i\pm1}(d_1)$, $(base)_i \longrightarrow (H1')_{i-1}(d_2)$ $(base)_i \longrightarrow (H2",H2')_{i-1}$ (d_3) and $(base)_i \longrightarrow (H3')_{i-1}(d_4)$. As an illustration Fig. 2 shows the d_2,d_3 and some d_1 connectivities. Stereospecific assignment of the H2',H2" protons has been obtained from the NOESY spectrum by monitoring the intensities of H1'-(H2',H2") cross peaks (Fig. 3). H1'-H2" cross peaks are always stronger than H1'-H2' peaks.

STRUCTURE OF CTCGAGCTCGAG

The NOESY spectrum provides vital clues on the overall structural profile of the dodecanucleotide. For example, we observe a number of interstrand NOE's (Fig. 2a); between the H2 proton of A5 and the H1' protons of C9 and T8, between H2 of A5 and H5 of C9, and between H2 of A11 and H1' of C3. This proves that the A-T base pairs are locked in hydrogen bonded

---

Fig. 2a. The region (ω_1 = 7 - 8.2 ppm; ω_2 = 5 - 6.4 ppm) in the NOESY spectrum. The spectrum shows complete intrastrand d_2 sequential connectivities involving base (H6/H8) protons and the H1' protons. The base protons are identified along the bottom of the figure, and the H1' protons along the horizontal lines. It may be noted that starting from G12, one is able to jump to the preceding nucleotides sequentially without any breaks till C1, as expected from d2 distances in a right handed DNA. In the same section of the spectrum one also observes, an intrastrand base-base (d_1) connectivity (dashed lines) involving H5 of C3 and H6 of the adjacent T2. Finally, four interstrand cross peaks are observed. These involve (i) H2 of A5 and H1' of C9, (ii) H2 of A5 and H1' of T8, (iii) H2 of A11 and H1' of C3 and (iv) H2 of A5 and H5 of C9 (depicted by dashed lines). These cross-peaks confirm the existence of A----T base pairing at all the four sites in the double helical DNA schematically shown in Fig. 1.

Fig. 2b. Sequential connectivities (d_3) involving base (H6/H8) protons and sugar H2'/H2" protons. In this region one also observes cross peaks from the thymine (T2 & T8) methyl protons to the respective H6 protons and to the H6 proton of the preceding cytosine (d_1 connectivity, shown by dashed lines).

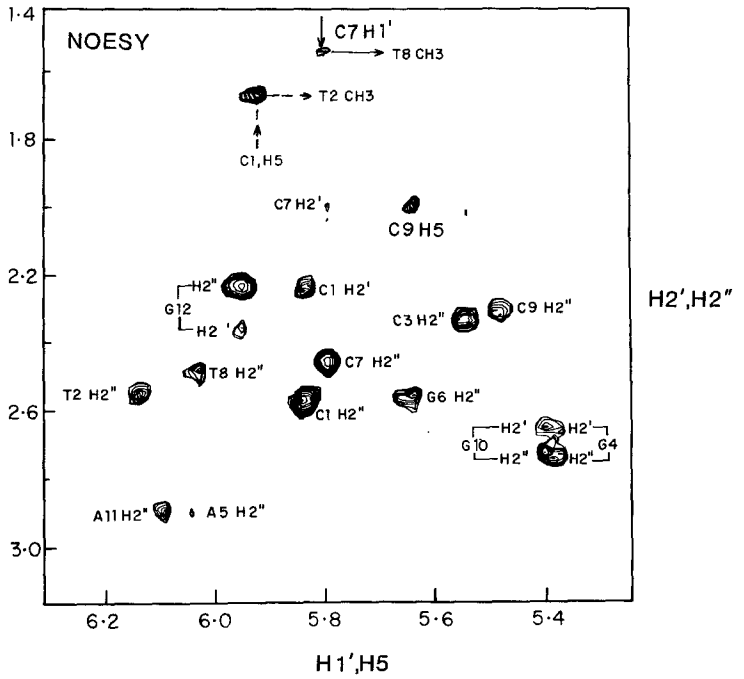


Fig. 3. Expansion of the H1'-(H2', H2") of the NOESY spectrum of d-CTCGAGCTCGAG. The mixing time is 300 sec. and the digital resolution along the two axes is 7.9 Hz. It is expected that H1' - H2" cross-peak for a nucleotide is more intense than the correspoding H1' - H2' peak. In fact, in seven cases (T2, C3, A5, G6, T8, C9 and A11), the H1' - H2' peaks are not visible at all. In this domain of the NOESY spectrum, one also observes the connectivities involving thymine methyl protons and the H1' (d_2) or H5 (d_1) proton on the adjacent cytosine nucleotide in the same strand (depicted by dashed lines).

net-work and the molecule is in the double helical form under the present experimental conditions. Further, the observed sequential connectivities indicate that the molecule adopts on overall right handed helical structure. Finally, since all the cross peaks in theNOESY spectrum have been assigned, only one type of structure is present in aqueous solution and coexistence of single stranded structures is excluded.

Using the methodologies described previously, [1-2,11-12] we have used COSY and NOESY spectra to obtain detailed information regarding the sugar geometry, glycosidic torsion angle and base-base stacking along the sequence of the molecule.

Sugar Geometries: Information about the sugar ring geometry is obtained from 2D COSY spectrum [12].

Fig. 4 shows expansions of all the domains of the COSY spectrum of the dodecanucleotide which are important for fixing the conformations of the deoxyribose rings. Several sugar residues show common features but a few show variations from the general pattern. One of the most striking observation is the complete absence of the H2"-H3' cross peaks in the COSY spectrum for ten of the twelve sugar residues (Fig. 4b). This implies that the value of J(H2"-H3') is small and it can be concluded that the conformation of deoxyribose rings of these ten nucleotides is restricted to the S domain of the pseudorotational wheel (P = 90 to 240°). The intensity patterns of the various cross peaks in Fig. 4 are consistent with the following conclusions (see ref.12).

(a) C7 is close to O1'-endo.

(b) T2, C3, A5, G6, T8, C9 and A11 have sugar geometries close to C1'-exo.

(c) G4 and G10 have geometries in the range C2'-endo to C4'-endo.

(d) C1 and G12 exhibit sugar geometries in the range O1'-endo to C4'-exo.

Glycosidic Dihedral Angle (χ): Information about χ can be obtained from the NOESY spectrum by monitoring the intensities of (H8/H6)---(H1') and (H8/H6) ---(H2',H2") cross peaks.

The behaviour of these cross peaks can be generally summarised as follows: (H8/H6) - H1' cross peaks are generally intense and those belonging to purines look relatively more intense than the ones belonging to the pyrimidines (Fig. 2a). (H8/H6) -- (H2', H2") (Fig. 2b) cross peaks, unfortunately show too much overlap in several cases and it is difficult to estimate their relative intensities. Although from these spectra it is not possible to fix the χ angle with accuracy, a qualitative conclusion that the χ angles are in high anti domain seems reasonable. A detailed analysis with NOE build-up is necessary for more precise determination.

Base-Base Stacking: The sequential d_1 connectivities carry direct information about base stacking patterns. These NOEs are not much complicated by spin diffusion processes, since for these, spin-diffusion paths are long and not very efficient. Thus even a long mixing time NOESY can be used to derive structural information to a fair degree of accuracy. In d-CTCGAGCTCGAG, d_1 connectivities are observed in the stretches C1-T2-C3, A5-G6-C7-T8 and A11-G12. Of these, all except A5-G6 and A11-G12 are depicted by dashed

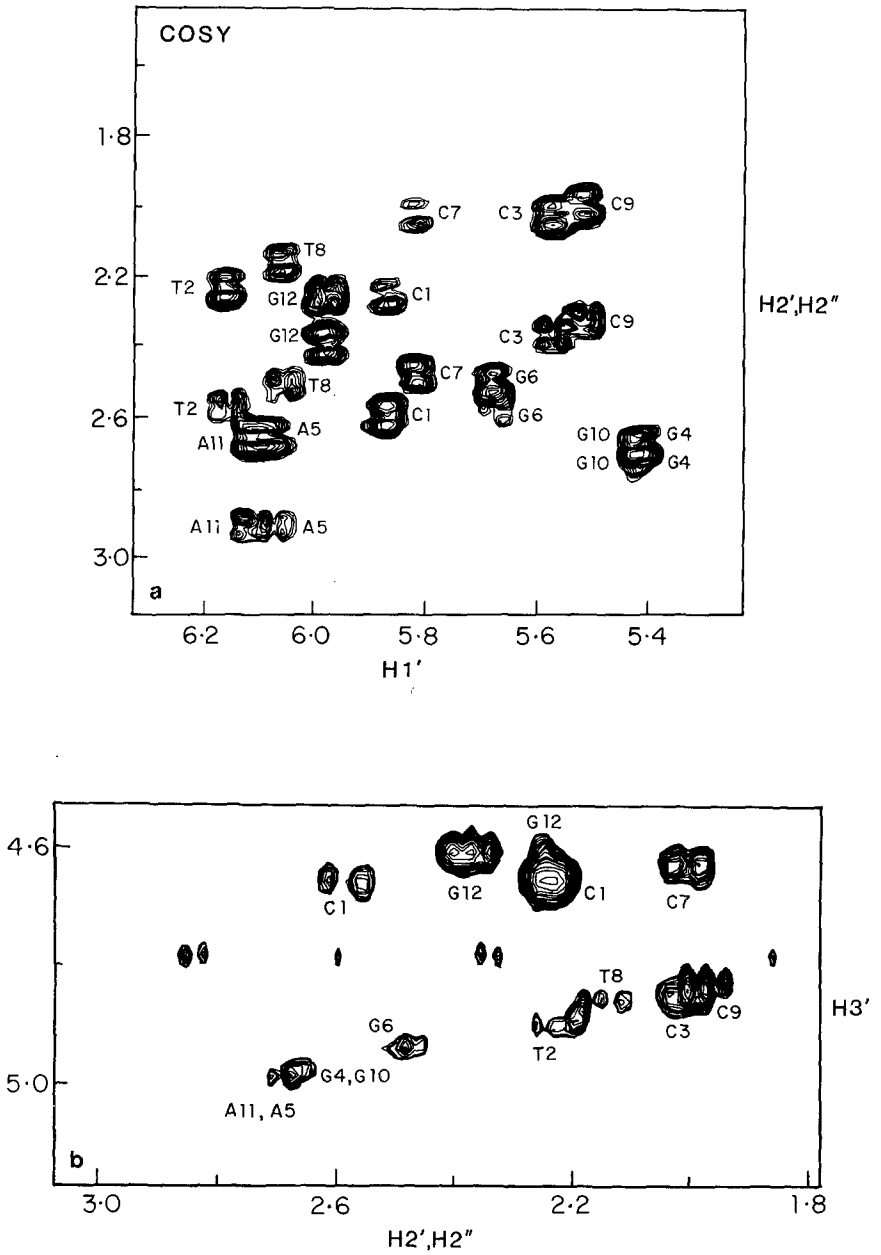


Fig. 4. Expanded sections of the 500 MHz COSY spectrum of d-CTCGAGCTCGAG at 25°C in D_2O at pH 7.2. The digital resolution along both axes is 7.9 Hz/pt.

a. H1' - (H2', H2") cross peaks.

b. H2' - H3' cross peaks; in this region one also sees two H2" - H3' cross peaks corresponding to C1 and G12. In the former case the H2' - H3' peak is upfield with respect to the H2" - H3' cross peak, whereas in the latter case the situation is reversed. All the other ten H2" - H3' cross peaks are absent.

c. H3' - H4' cross peaks which are observed for ten out of twelve nucleotides. The peaks corresponding to G4 and G10 are not observed.

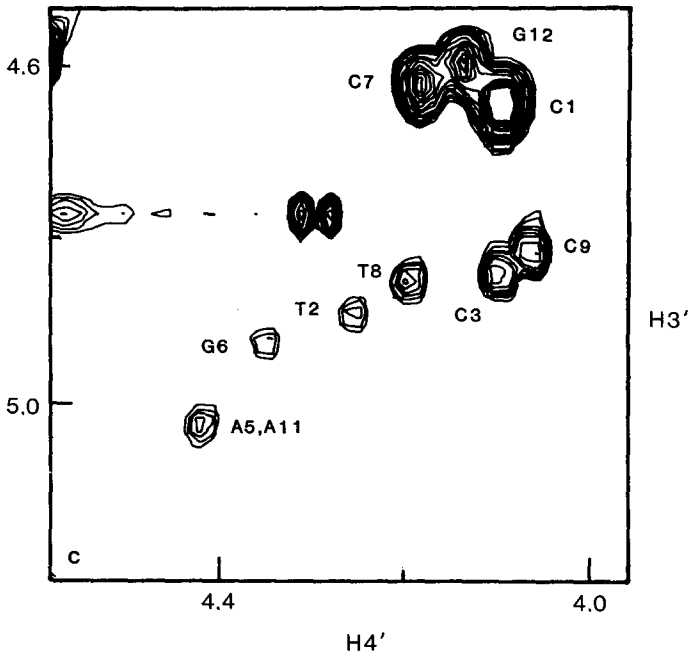


Figure 4. continued.

lines in Figs. 2 and 3. The NOEs between A5 H8 and G6 H8 and between A11 H8 and G12 H8 protons occur in a different region of the NOESY spectrum which is not shown here. The fact that d_1 connectivities are absent in the stretches C3-G4-A5 and T8-C9-G10-A11, indicates that the base-base stacking in the major portion of the molecule deviates from that normally observed in the B DNA models. This conclusion is also supported by the observed d_3 connectivities. Examination of the d_1 patterns vis a vis the sequence of the molecule reveals an interesting sequence effect in the structure of the molecule. It is seen that the stretch

-GAGCTCGAG-

of the molecule has a symmetry in the base sequence around the thymine in the centre. Interestingly the d_1 connectivities are exactly complementary in the two halves; connectivities are observed in the stretch AGCT but not in the TCGA stretch. Similarly -GA- has no d_1 connectivity whereas -AG- at the 3' end shows a d_1 connectivity. The -CTC- stretch in the centre looks different as compared to the -CTC- fragment at the 5' - end; the latter shows both the d_1 connectivities. These observations suggest that the sequ-

ence effects on the structure are not limited to just the nearest neighbours but extend over a longer stretch of the sequence. In the present case, the relevant sequence is five units long. This provides a possible explanation as to why the restriction enzymes require longer stretches of DNA for specific recognition, rather than just the structure around the cleavage site.

ACKNOWLEDGEMENTS

The help provided by the 500 MHz FT-NMR National Facility supported by the Department of Science and Technology, Government of India, is gratefully acknowledged.

REFERENCES

1. Hosur, R.V., Ravikumar, M., Roy, K.B., Tan-Zu-Kun, Miles, H.T. and Govil, G. (1985) 'Magnetic Resonance in Biology and Medicine', Tata McGraw-Hill Publishing Company Limited, New Delhi, 243-260.
2. Ravikumar, M., Hosur, R.V., Roy, K.B., Tan-Zu-Kun, Miles, H.T. and Govil, G. (1985), Biochemistry, 24, 7703-7711.
3. Chary, K.V.R., Hosur, R.V., Govil, G. and Tan-Zu-Kun and Miles,H.T. (1987) Biochemistry (in press).
4. Jeener, J. (1971) Ampere International Summer School, Basko, Polje, Yugoslavia.
5. Jeener, J., Meier, B.H., Bachmann, P. and Ernst, R.R., J. Chem. Phys. (1979) 71, 4546-4553.
6. Anil Kumar, Hosur, R.V. and Chandrasekhar, K. (1984) J. Magn. Reson. 60, 143-148.
7. Hosur, R.V., Chary, K.V.R., Anil Kumar and Govil, G. (1985), J. Magn. Reson. 62, 123-127.
8. Hare, D.R., Wemmer, D.L., Chou, S.H., Drobny, G. and Ried, B.R. (1983), J. Mol. Biol. 171, 319-336.
9. Scheek, R.M., Boelens, R., Russo, N., van Boom, J.H. and Kaptein, R. (1984) Biochemistry 23, 1371-1376.
10. Gronenborn, A.M. and Glore, G.M., Prog. NMR spectrc. (1985), 17,1-32.
11. Hosur, R.V. (1986), Curr. Sci. 55, 597-605.
12. Hosur, R.V., Ravikumar, M., Chary, K.V.R., Sheth, A., Govil, G., Tan-Zu-Kun and Miles, H.T. (1986) FEBS Letts. 205, 71-76.