

Sequence-dependent structural variations in two right-handed alternating pyrimidine-purine DNA oligomers in solution determined by nuclear Overhauser enhancement measurements

G.Marius Clore* and Angela M.Gronenborn*

Division of Molecular Pharmacology, National Institute for Medical Research, Mill Hill, London NW7 1AA, UK

Communicated by F.Cramer

Received on 30 June 1983; revised on 22 August 1983

A 500 MHz ¹H-n.m.r. study on two right-handed self-complementary double-stranded alternating pyrimidine-purine oligodeoxyribonucleotides, 5'dCGTACG and 5'dACGCGCGT, is presented. Using the proton-proton nuclear Overhauser effect, proton resonances are assigned by a sequential method and a large number of interproton distances, both intra- and internucleotide, are determined (113 for 5'dCGTACG and 79 for 5'dACGCGCGT). The general procedure required to solve the three-dimensional solution structures of oligonucleotides from such distance data is outlined and applied to these two oligonucleotides. In the case of both oligonucleotides the overall solution structure is that of B DNA, namely a right-handed helix with a helical rise of ~3.3 Å, 10 bp per turn and the base pairs approximately perpendicular to the helix axis. In the case of 5'dCGTACG, subtle local structural variations associated with the pyrimidine and purine nucleotides are superimposed on the overall structure but the mononucleotide repeating unit is preserved. In contrast, 5'dACGCGCGT has a clear alternating structure with a dinucleotide repeat, alternation occurring in the local helical twist and the glycosidic bond, sugar pucker and phosphodiester backbone conformations. Key words: right-handed DNA/oligonucleotides/solution structure/dinucleotide repeat/n.m.r./nuclear Overhauser effect

Introduction

The classical view of B DNA derived from fibre diffraction is of a smooth regular right-handed double helix with a mononucleotide repeating unit, 10 bp per turn, a helical twist of 36°, a helical rise of 3.3 Å and a base tilt of ~0° (Neidle and Berman, 1983). With the advent of the single crystal structure of the self-complementary dodecamer 5'dCGCGAATTCGCG (Dickerson and Drew, 1981), these general features were confirmed but, in addition, sequence-specific local variations in nucleotide conformation, superimposed on the general B DNA structure, were shown to occur. More recently, fibre diffraction studies on alternating Pyr_pPur copolymers have demonstrated the existence of wrinkled B DNA with a dinucleotide repeating unit (Arnott *et al.*, 1983). These local variations in the structure of B DNA have generated considerable interest (Calladine, 1982; Dickerson, 1983) as they may play an important role in the recognition process involved in specific DNA-protein interactions.

Although X-ray diffraction provides extensive structural detail, the crystal and fibre structures of oligonucleotide crystals and DNA fibres are subject to crystal packing forces and local high ionic conditions (Rhodes, 1982), and these may well account for much of the local structure variations

observed. In this respect it should be noted that whereas the crystal structure of the self-complementary dodecamer 5'dCGCGAATTCGCG does not exhibit a dyad axis, the solution structure is clearly symmetric as judged by the observation of only a single set of resonances for each nucleotide (Sarma *et al.*, 1982; Patel *et al.*, 1983). It is therefore of interest to determine the three-dimensional solution structures of oligonucleotides under physiological conditions where intermolecular interactions do not distort the structure.

To date, the existence of local variations in the solution structure of DNA has been deduced from the finding of sequence-specific differential cleavage by the enzyme deoxyribonuclease I (Scheffler *et al.*, 1968; Lomonosoff *et al.*, 1981), and the n.m.r. observation of two ³¹P resonances and two C1', C2', C3' and C4' ¹³C resonances of equal intensity for alternating Pyr_pPur copolymers (Shindo *et al.*, 1979; Shindo, 1981; Cohen *et al.*, 1981; Chen and Cohen, 1983). The latter indicate some sort of regular alternating structure for the phosphate backbone and ribose rings. However, the precise nature of these local variations in the solution structure of DNA cannot be ascertained from the data.

Of all the physico-chemical techniques of conformational analysis applicable in solution, the proton-proton nuclear Overhauser effect (NOE) is potentially the most powerful as it can be used to demonstrate the proximity of two protons in space and determine their separation (Noggle and Schirmer, 1971). Given the interconvertibility of intramolecular distances, torsion angles and cartesian coordinates, providing the chirality of the structure is known (Crippen and Havel, 1978; Cohen and Sternberg, 1980; Braun *et al.*, 1981; Wüthrich *et al.*, 1982), this information can in principle be used to determine the complete three-dimensional structure of a molecule in solution at a resolution comparable to that attainable by X-ray crystallography. To date, the proton-proton NOE has principally been applied with considerable success to obtain proton resonance assignments and restricted structural information in the case of small proteins (Redfield and Gupta, 1971; Poulsen *et al.*, 1980; Wagner *et al.*, 1981), ligand-protein complexes (Gronenborn and Clore, 1982; Clore *et al.*, 1982; Clore and Gronenborn, 1983), tRNA (Johnston and Redfield, 1978; Roy and Redfield, 1983; Hare and Reid, 1982; Heerschap *et al.*, 1983) and oligonucleotides (Kan *et al.*, 1982; Reid *et al.*, 1983; Scheek *et al.*, 1983; Pardi *et al.*, 1983; Patel *et al.*, 1983). Moreover, the elucidation of the complete three-dimensional solution structures of a number of polypeptides and small proteins, including lipid-bound glucagon (Wider *et al.*, 1982), basic pancreatic trypsin inhibitor (Wagner and Wüthrich, 1982) and bull seminal plasma proteinase inhibitor IIA (Štrop *et al.*, 1983) are well in hand.

We have used pre-steady-state proton-proton NOE measurements to determine a large number of interproton distances, both intra- and internucleotide, for two self-complementary alternating Pyr_pPur oligodeoxyribonucleotides, 5'dCGTACG and 5'dACGCGCGT. From this distance data, the values of the glycosidic bond and backbone torsion angles and of the helical twist and rise are determined

*Reprint requests may be addressed to either author.

by model building, and the effects of sequence on local structural variations explored. Whereas the hexamer exists in a conventional B DNA form with a mononucleotide repeat and only minor base to base variations, the octamer exists in an alternating B DNA form with a dinucleotide repeating unit.

Results and Discussion

Resonance assignment and interproton distance determination

To assign resonances and determine interproton distances, we have made use exclusively of the proton-proton NOE which arises from cross-relaxation between protons close in space. In conventional one-dimensional n.m.r., the NOE experiment involves saturation of the resonance of proton *i* and observing the intensity of the other proton resonances. In the case of large molecules (mol. wt. ≥ 1000) with long correlation times ($\tau_c \geq 10^{-9}$ s) for which $\omega\tau_c > 1$, the NOEs observed are negative (Solomon, 1955). If the selective saturation pulse is applied for only a short time (typically ≤ 0.5 s for a molecule of mol. wt. ~ 4000) the NOE between two protons *i* and *j*, η_{ij} , will be given by

$$\eta_{ij} \sim \sigma_{ij}t \quad (1)$$

as the initial build-up rate of the NOE is equal to the cross-relaxation rate σ_{ij} between the two protons *i* and *j* (Wagner and Wüthrich, 1979; Dobson *et al.*, 1982). Quantitative distance information can then be obtained as σ_{ij} is inversely proportional to the sixth power of the distance r_{ij} between the two protons (Solomon, 1955). Consequently, the ratio of two interproton distances can be obtained from the equation

$$r_{ij}/r_{kl} = (\sigma_{kl}/\sigma_{ij})^{1/6} = (\eta_{kl}/\eta_{ij})^{1/6} \quad (2)$$

and providing an internal reference distance(s) is available, actual interproton distances can be calculated. In the case of DNA four such internal reference distances can be used: $r_{C(H6)_2C(H5)} = 2.46$ Å, $r_{T(H6)-T(CH_3)} = \sim 2.5$ Å, $r_{H2'-H2''} = 1.78$ Å and, in an AT base pair, $r_{T(N3H)-A(H2)} = 2.84$ Å.

Because $\eta_{ij} \propto 1/r_{ij}^6$, direct (first-order) NOEs are only detectable up to distances of ~ 5 Å, beyond which the effects fall

to less than -1% and become virtually undetectable. Bearing this in mind, a sequential assignment strategy for right-handed DNA, both B and A, can be defined as only a particular set of NOEs are to be expected, as illustrated in Figure 1.

All NOE measurements on the two oligonucleotides, 5'-dCGTACG and 5'-dACGCGCGT, were carried out under conditions where they were completely double-stranded, as judged by thermal denaturation studies (data not shown) monitoring the temperature dependence of the chemical shifts of the non-exchangeable base protons, (namely 5°C and 1 M KCl in the case of 5'-dCGTACG, and 10°C and 500 mM KCl in the case of 5'-dACGCGCGT). Under these conditions the c.d. spectra of these two oligonucleotides are characteristic of right-handed B DNA (Kuzmich *et al.*, 1982; Clore and Gronenborn, unpublished data). The selective irradiation pulse used in quantitative NOE measurements was applied for 0.3 s, and control experiments using different irradiation times for a few selected resonances indicated that Equation (1) was valid at this irradiation time. Systematic measurements were carried out by irradiating in turn all exchangeable and non-exchangeable proton resonances. On the basis of the scheme illustrated in Figure 1 all exchangeable and non-exchangeable base proton resonances and all sugar proton resonances (with the exception of the H5' resonances of T₃, C₅ and G₆ in the case of 5'-dCGTACG, and the H5' and H5'' resonances of C₂, G₃, C₄, G₅, C₆, G₇ and T₈ in the case of 5'-dACGCGCGT) were assigned for both oligonucleotides. Moreover, a large number of interproton distances, both intra- and internucleotide, were determined: a total of 113 in the case of 5'-dCGTACG and 79 in the case of 5'-dACGCGCGT. Assuming an error of ± 0.05 Å in the estimated values of the reference distances, the error in the values of these calculated distances is $\leq \pm 0.15$ Å. The complete set of assignments and interproton distances for the two oligonucleotides will be published elsewhere.

Figure 2 illustrates the application of the NOE to assignment and structure determination in the case of 5'-dCGTACG. Irradiation of the T₃(CH₃) resonance (Figure 2B) results in: (i) a direct intranucleotide NOE on the T₃(H6) resonance and indirect intranucleotide NOEs [via the T₃(H6) proton] on the T₃(H2') and T₃(H2'') resonances, and (ii) direct internucleotide NOEs to resonances of the 5' nucleotide, namely the G₂(H8), G₂(H1'), G₂(H2') and G₂(H2'') resonances. Irradiation of the T₃(H6)/C₅(H6) resonance (Figure 2C) results in a combination of NOEs: (i) intranucleotide NOEs between the T₃(H6) proton and the T₃(CH₃), T₃(H1'), T₃(H2') protons, and between the C₅(H6) proton and the C₅(H5), C₅(H1') and C₅(H2') protons; (ii) indirect intranucleotide NOEs between the T₃(H6) and C₅(H6) protons and their respective H2'' protons via the H2' protons (which occur owing to the very small separation (1.78 Å) between the H2' and H2'' protons of the same nucleotide); and (iii) direct internucleotide NOEs between the T₃(H6) proton and the G₂(H8), G₂(H1'), G₂(H2'), G₂(H2'') and A₄(H8) protons, and between the C₅(H6) proton and the A₄(H8), A₄(H1'), A₄(H2') and A₄(H2'') protons. Irradiation of the A₄(H8) resonance (Figure 2D) then results in (i) direct intranucleotide NOEs on the A₄(H1'), A₄(H2') and A₄(H3') resonances and an indirect intranucleotide NOE [via the A₄(H2') proton] on the A₄(H2'') resonance; (ii) direct internucleotide NOEs on resonances of the 5' nucleotide, namely the T₃(H6), T₃(CH₃), T₃(H1'), T₃(H2') and T₃(H2'') resonances; and (iii) direct internucleotide NOEs on

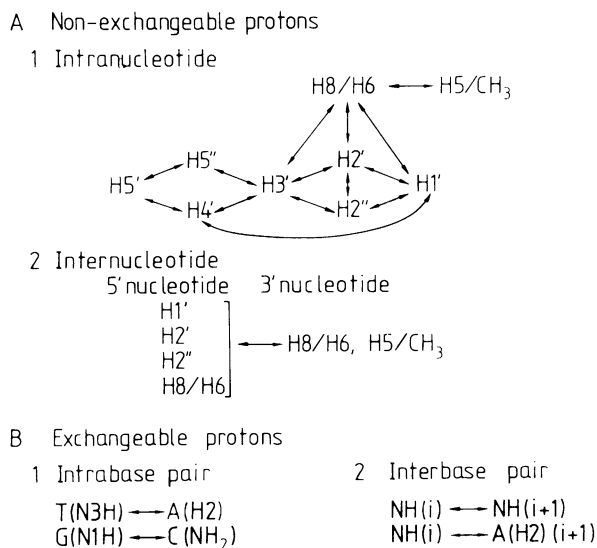


Fig. 1. Schematic illustration of the intra- and internucleotide interproton distances with values of ≤ 5 Å in right-handed DNA which form the basis of the sequential assignment procedure and structure determination by means of NOE measurements. These distance relationships are applicable to both B and A DNA (Arnott and Hukins, 1972).

resonances of the 3' nucleotide, namely the C₅(H5) and C₅(H6) resonances. A similar set of experiments for 5'-dACGCGGT is shown in Figure 3. It is evident from the examples shown in Figures 2 and 3 that the systematic application of NOE measurements can yield a complete set of assignments and an interproton distance matrix describing the solution structures of oligonucleotides.

Determination of the solution structures of oligonucleotides from the interproton distances determined by NOE measurements

Because of the limited degrees of freedom available in double-stranded DNA, the three-dimensional structure of short oligonucleotides can be solved by manual model building on the basis of the intra- and internucleotide interproton distances determined from NOE measurements. The sugar pucker conformation, defined in terms of the C4'-C3' bond torsion angle δ , and the glycosidic bond torsion angle χ are uniquely defined by the intranucleotide distances $r_{H1'-H2'}$, $r_{H1'-H2''}$, $r_{H2'-H3'}$, $r_{H2''-H3'}$, $r_{H1'-H8/H6}$, $r_{H2'-H8/H6}$ and $r_{H3'-H8/H6}$. In addition, the C5'-C4' bond torsion angle can also be uniquely defined providing one of the three distances $r_{H3'-H5''}$, $r_{H4'-H5''}$ and $r_{H4'-H5'}$ is determined. As an example, consider the nucleotide A₄ of 5'-dCGTACG: $r_{H1'-H2'}$, $r_{H1'-H2''}$, $r_{H2'-H3'}$, $r_{H2''-H3'}$, $r_{H1'-H4'}$, $r_{H2'-H8}$ and $r_{H3'-H8}$ are found to have values of 2.48, 2.24, 2.19, 2.48, 2.92, 2.15 and 3.21 Å, respectively, which yield values of χ

and δ of -110° and 120° , respectively; in addition, $r_{H3'-H5''}$, $r_{H4'-H5''}$ and $r_{H4'-H5'}$ have values of 2.61, 2.15 and 2.13 Å, respectively, yielding a value of 50° for γ . Once χ and δ are known for each nucleotide, the internucleotide interproton distances (see Figure 1) enable one to define the position of each base pair with respect to that of the adjacent base pairs on either side in terms of the helical rise, helical twist and base tilt. With all these structural parameters determined, the ranges of the P-O5'(α), O5'-C5'(β), C3'-O3'(ϵ) and O3'-P(ζ) backbone torsion angles can be restricted to a narrow range within $\pm 20^\circ$. Thus, in the case of the central AT base pairs of 5'-dCGTACG, the internucleotide interproton distances $r_{T_3(H6)-A_4(H8)}$, $r_{T_3(CH_3)-A_4(H8)}$, $r_{T_3(H1')-A_4(H8)}$, $r_{T_3(H2')-A_4(H8)}$ and $r_{T_3(H2'')-A_4(H8)}$ were found to have values of 3.69, 4.07, 3.16, 2.80 and 2.64 Å, respectively. Taken together with the values of χ and δ of -115° and 110° , respectively, determined for T₃ and those given above for A₄, the helical rise, helical twist and base tilt for the central AT base pairs are found to have values of 3.5 Å, $+40^\circ$ (i.e., right-handed) and $\sim 0^\circ$, respectively, with the backbone torsion angles α , β , ϵ and ζ lying in the g^- , t , t and g^- ranges, respectively. In addition, where at least two distances from the imino proton of one base pair to protons of the adjacent base pair [i.e., imino, amino and A(H2) protons] are known, the propeller twist and base roll angles can be defined. In the case of 5'-dCGTACG this was possible for the central AT

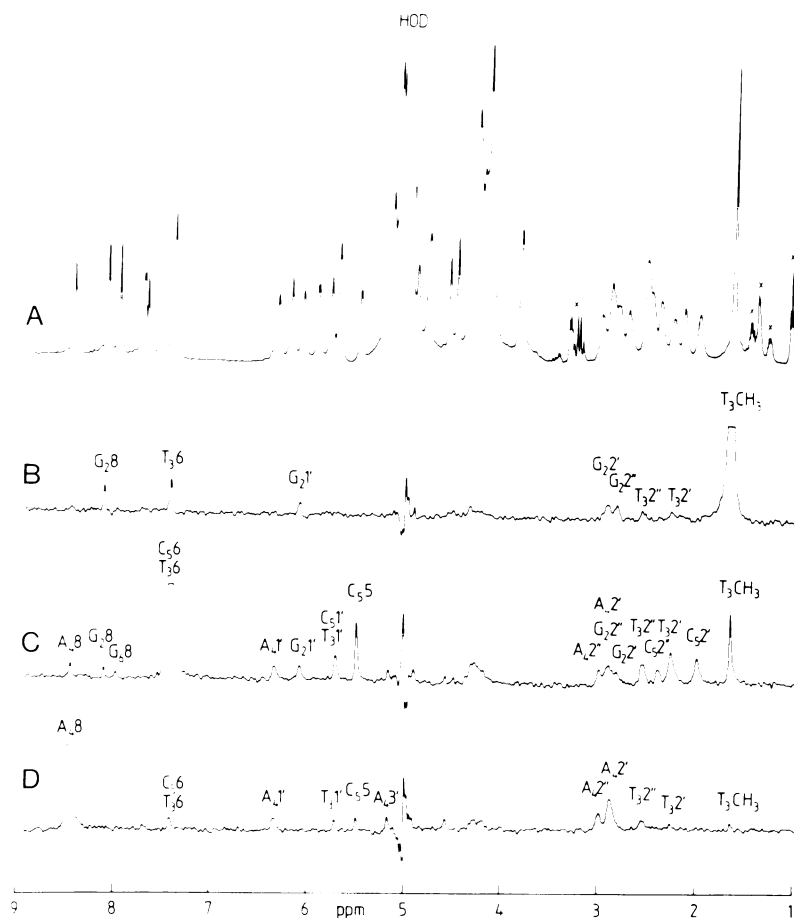


Fig. 2. Pre-steady state NOE measurements on double-stranded 5'-dCGTACG in 99.6% D₂O at 5°C. (A) The 500 MHz ¹H-n.m.r. spectrum; difference spectra (off-resonance minus on-resonance pre-irradiation) following pre-saturation for 0.3 s of the T₃(CH₃) (B), T₃(H6)/C₅(H6) (C) and A₄(H8) (D) resonances. The experimental conditions are: 0.35 mM double-stranded oligonucleotide in 99.96% D₂O, 1 M KCl, 50 mM potassium phosphate pH* 6.5 (meter reading uncorrected for the isotope effect on the glass electrode) and 0.1 mM EDTA. (The peaks labelled with an x are due to residual triethylammonium acetate and other impurities.)

base pairs as the distances $r_{T_3(H3)-G_2(H1)}$ and $r_{A_4(H2)-G_2(H1)}$ could be determined and were found to have values of 3.54 and 3.22 Å, respectively. These two distances are only compatible with a large propeller twist angle ($\sim 18^\circ$) and a small base roll angle ($\sim 0^\circ$).

The three-dimensional solution structures of 5'dCGTACG and 5'dACGCGCGT

Using the principles described in the previous section, the complete three-dimensional structures of double-stranded 5'dCGTACG and 5'dACGCGCGT were solved. The overall error between the experimentally determined interproton distances (113 in the case of 5'dCGTACG and 79 in the case of 5'dACGCGCGT) and those of the models are $\leq \pm 20\%$. The structural parameters for the individual nucleotides of the two oligonucleotides will be published elsewhere and only the broad outlines of the structures will be described here. These are summarised in Table I in terms of the mean purine and pyrimidine glycosidic bond and backbone torsion angles, the mean helical twist for the Pur_pPyr and Pyr_pPur steps, and the mean helical rise, base pairs per turn and base tilt, and are compared with those obtained by single crystal X-ray diffraction on B and A oligonucleotides (Dickerson and Drew, 1981; Shakked *et al.*, 1983) and by fibre diffraction on the wrinkled

B DNA form of poly d(GC) (Arnott *et al.*, 1983), and with those proposed for alternating B DNA by Klug *et al.* (1979).

The overall solution structure of both oligonucleotides is that of right-handed B DNA with an overall mean helical twist of $\sim 36^\circ$, 10 bp per turn and the bases approximately perpendicular to the helix axis. This result is entirely consistent with the c.d. spectra of these oligonucleotides under the same conditions as those used in the n.m.r. experiments (Kuzmich *et al.*, 1982; Clore and Gronenborn, unpublished data).

In the case of double-stranded 5'dCGTACG there are minor local structural variations superimposed on the overall B DNA structure but the mononucleotide repeating unit of classical B DNA is preserved. The differences in the glycosidic and C4'-C3' bond torsion angles between purine and pyrimidine nucleotides, though less marked, follow the same trend as that observed in the single crystal structure of the B DNA dodecamer 5'dCGCGAATTCGCG (Dickerson and Drew, 1981). In addition, the large propeller twist ($\sim 18^\circ$) for the central AT base pairs is associated with a small helical twist ($\sim 27^\circ$) for the Pur_pPyr step, thus relieving steric clash, exactly as predicted by Calladine's principles (Calladine, 1982; Dickerson, 1983) and previously found in the single crystal structure of the B DNA dodecamer (Dickerson and Drew, 1981).

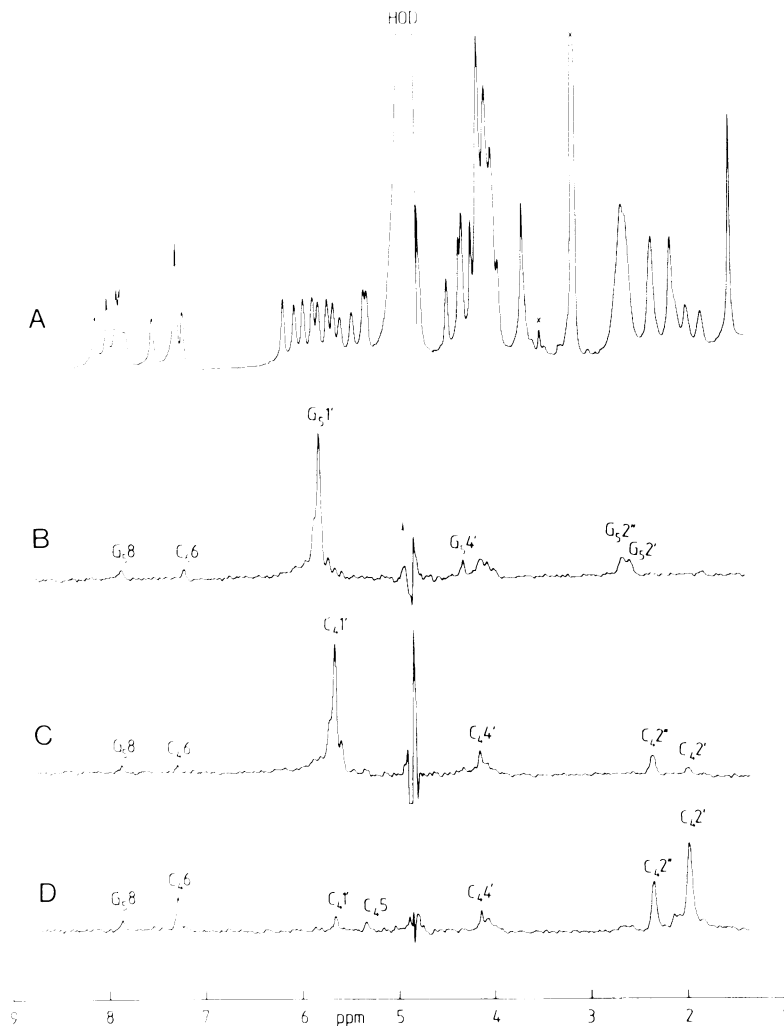


Fig. 3. Pre-steady-state NOE measurements on double-stranded 5'dACGCGCGT in 99.96% D₂O at 10°C. (A) The 500 MHz ¹H-n.m.r. spectrum; NOE difference spectra following pre-saturation for 0.3 s of the G₅(H1') (B), C₄(H1') (C) and C₄(H2') (D) resonances. The experimental conditions are the same as in Figure 2 except that the concentrations of oligonucleotide and KCl were 1 mM and 500 mM, respectively. (The peaks labelled with an x are due to residual triethylammonium acetate and other impurities.)

In contrast to 5'dCGTACG, double-stranded 5'dACGCGCGT exhibits a clear alternating structure with a dinucleotide repeating unit consisting of the following features (see Table I).

(i) The glycosidic bond and sugar pucker conformations lie in the high anti and O1'-endo ranges, respectively, in the case of the purine nucleotides, and in the anti and C1'-exo ranges in the case of the pyrimidine nucleotides. The variation in the sugar pucker can be deduced immediately from the NOE data shown in Figure 3B and C in which the G₅(H1') and C₄(H1') resonances, respectively, are irradiated: whereas the NOEs between the G₅(H1') protons and the G₅(H2') and G₅(H2'') protons are approximately equal, indicative of an N-type conformation (i.e., in the O1'-endo/C3'-endo range), the NOE between the C₄(H1') and C₄(H2'') protons is double the size of that between the C₄(H1') and C₄(H2') protons indicative of an S-type conformation (i.e., in the C1'-exo/C2'-endo

range).

(ii) The local helical twist is reduced for the Pur_pPyr steps (~30°) and increased for the Pyr_pPur steps (~42°) relative to an overall mean of ~36° corresponding to 10 bp per turn.

(iii) There is a clear alternation in the P-O5'(α), C3'-O3'(ε) and O3'-P(ζ) backbone torsion angles which lie in the *t*, *g*⁺/*t* and *g*⁻ ranges, respectively, in the case of the purine nucleotides, and in the *g*⁻, *g*⁺ and *g*⁺ ranges, respectively, in the case of the pyrimidine nucleotides.

(iv) As a result of the alternation in the backbone torsion angles, the line joining the two oxygen atoms of the O_L-P-O_R unit is perpendicular to the helix axis in the case of the Pyr_pPur steps and parallel to the helix axis in the case of the Pur_pPur steps, and the line joining two adjacent phosphorus atoms (which are separated by ~6.5 Å) is at ~25° from the helix axis in the case of a Pur_pPur sequence and at ~60° from the helix axis in the case of a Pyr_pPur sequence. This

Table I. Parameters describing the solution structures of double-stranded 5'dCGTACG and 5'dACGCGCGT determined by model building based on the inter-proton distances obtained from NOE measurements

A. Mean glycosidic bond and main chain torsion angles

	χ	α	β	γ	δ	ε	ζ
Double-stranded 5'dCGTACG							
Pyr	-113 ± 6° (anti)	<i>g</i> ⁻	<i>t</i>	63 ± 6° (<i>g</i> ⁺)	113 ± 5° (C1'-exo)	<i>t</i>	<i>g</i> ⁻
Pur	-108 ± 6° (anti)	<i>g</i> ⁻	<i>t</i>	53 ± 6° (<i>g</i> ⁺)	118 ± 6° (C1'-exo)	<i>t</i>	<i>g</i> ⁻
Double-stranded 5'dACGCGCGT							
Pyr	-100 ± 5° (anti)	<i>g</i> ⁻	<i>t</i>	~55° (<i>g</i> ⁺)	123 ± 8° (C1'-exo)	<i>g</i> ⁺	<i>g</i> ⁺
Pur	-70 ± 12° (high anti)	<i>t</i>	<i>t</i>	~55° (<i>g</i> ⁺)	105 ± 6° (O1'-endo)	<i>g</i> ⁺ / <i>t</i>	<i>g</i> ⁻
B DNA ^a _{crystal}							
Pyr	-124 ± 8° (anti)	<i>g</i> ⁻	<i>t</i>	56 ± 6° (<i>g</i> ⁺)	117 ± 19° (C1'-exo)	<i>t</i>	<i>g</i> ⁻
Pur	-110 ± 14° (anti)	<i>g</i> ⁻	<i>t</i>	53 ± 9° (<i>g</i> ⁺)	129 ± 20° (C1'-exo)	<i>t</i>	<i>g</i> ⁻
A DNA ^b _{crystal}							
Pyr	-155 ± 3° (low anti)	<i>g</i> ⁻	<i>t</i>	47 ± 12° (<i>g</i> ⁺)	88 ± 4° (C3'-endo)	<i>t</i>	<i>g</i> ⁻
Pur	-166 ± 9° (low anti)	<i>g</i> ⁻	<i>t</i>	60 ± 12° (<i>g</i> ⁺)	88 ± 3° (C3'-endo)	<i>t</i>	<i>g</i> ⁻
Wrinkled B DNA ^c							
Pyr	-107 ± 3° (anti)	<i>g</i> ⁻	<i>t</i>	47 ± 4° (<i>g</i> ⁺)	143 ± 2° (C2'-endo)	<i>g</i> ⁻	<i>t</i>
Pur	-86 ± 3° (high anti)	<i>g</i> ⁻	<i>t</i>	22 ± 3° (<i>g</i> ⁺)	148 ± 2° (C2'-endo)	<i>t</i>	<i>t</i>
Alternating B DNA ^d							
Pyr	-104° (anti)	<i>g</i> ⁻	<i>t</i>	65° (<i>g</i> ⁺)	145° (C2'-endo)	<i>t</i>	<i>t</i>
Pur	-147° (low anti)	<i>g</i> ⁻	<i>t</i>	59° (<i>g</i> ⁺)	99° (O1'-endo)	<i>t</i>	<i>g</i> ⁻

B. Further structural parameters

	Double-stranded 5'dCGTACG	Double-stranded 5'dACGCGCGT	B DNA _{crystal} ^a	A DNA _{crystal} ^b	Wrinkled B DNA ^c	Alternating B DNA ^d
Helical sense	right	right	right	right	right	right
Mean helical rise	3.4 ± 0.2 Å	~3.3 Å	3.33 ± 0.13 Å	3.03 ± 0.39 Å	3.36 Å	3.3 Å
Mean local helical twist ^e						
Pur _p Pyr	27°	30°	39 ± 6°	33 ± 0.7°	30°	31°
Pyr _p Pur	39°	42°	35 ± 3°	30 ± 0.3°	42°	41°
Mean base pairs per turn	10.3	~10	10.1	11.2	10	10
Base tilt	~0°	~0°	4°	18°	~0°	~0°
Repeating unit	mononucleotide	dinucleotide	mononucleotide	mononucleotide	dinucleotide	dinucleotide
Ratio of groove widths (major/minor)	>1	>	>1	<<1	>1	>1
Ratio of groove depths (major/minor)	~1	~1	~1	>>1	~1	~1

The main chain torsion angles are defined P α O5' β C5' γ C4' δ C3' ε O3' ζ P and the glycosidic bond torsion angles by χ_{pur} = O1'-C1'-N9-C4 and χ_{pyr} = O1'-C1'-N1-C2. A positive torsion angle indicates a clockwise rotation of the more distant atom. The ranges *g*⁺, *t* and *g*⁻ are defined by the angles 60 ± 60°, 180 ± 60° and -60 ± 60°, respectively.

^aFrom the crystal structure of the B DNA dodecamer 5'dCGCGAATTCGCG (Dickerson and Drew, 1981).

^bFrom the crystal structure of the A DNA octamer 5'dGGTATACC (Shakke *et al.*, 1983).

^cFrom the fibre diffraction data on the wrinkled B DNA form of poly d(GC) (Arnott *et al.*, 1983).

^dFrom the model of alternating B DNA proposed by Klug *et al.* (1979) and the crystal structure of the mini 2-bp duplex formed by 5'dATAT (Viswamitra *et al.*, 1982).

^eOnly Pur_pPyr and Pyr_pPur steps have been considered in calculating the mean local helical twists for the crystal structures of the B DNA dodecamer and A DNA octamer.

latter feature gives a slight zig-zag appearance to the backbone.

A skeletal model of the alternating structure of 5'dACGCGCGT is shown in Figure 4.

Concluding remarks

Here we have outlined the procedures required to determine the complete three-dimensional structures of oligonucleotides in solution from proton-proton NOE data, and applied these to two right-handed B type self-complementary oligodeoxyribonucleotides, 5'dCGTACG and 5'dACGCGCGT. We have shown that local structural variations do occur in B DNA in solution, and these may be of a rather subtle nature as in the case of 5'dCGTACG where the mononucleotide repeating unit is preserved, or of a more profound nature as in the case of 5'dACGCGCGT which has a clear alternating structure with a dinucleotide repeat.

The local structural variations found in the solution structure of 5'dCGTACG are similar, though less marked, than in the crystal structure of the B DNA dodecamer 5'dCGCGAATTCGCG (Dickerson and Drew, 1981) and this difference may in part be due to the absence of intermolecular interactions, such as crystal packing forces, in solution.

The alternating solution structure of 5'dACGCGCGT is in principle similar to that of wrinkled B DNA found in fibres of poly d(GC) (Arnott *et al.*, 1983) and of the model of alter-

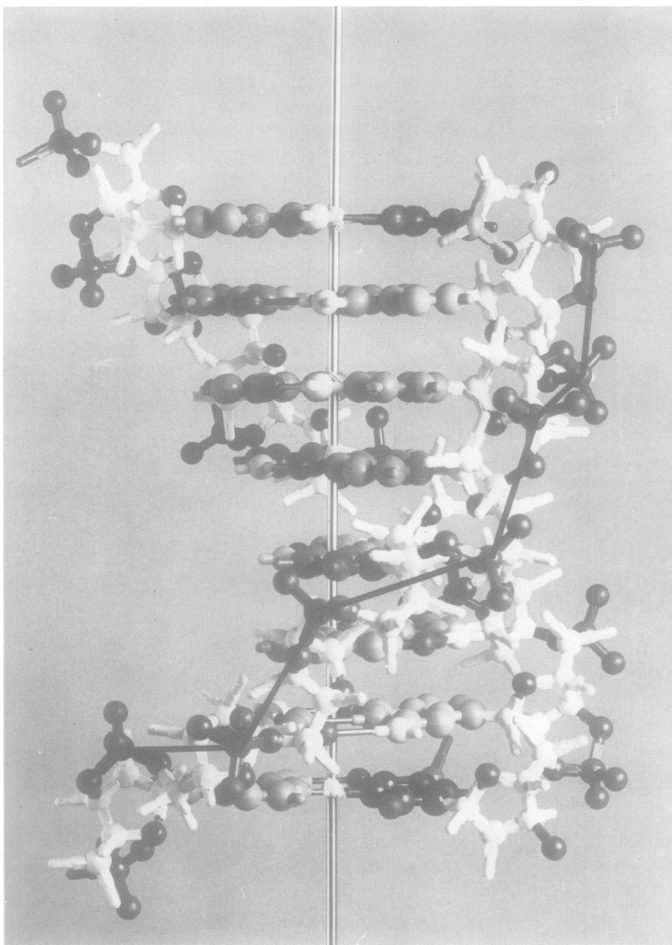


Fig. 4. Skeletal model of the alternating B type solution structure of 5'dACGCGCGT. The phosphorus atoms are connected up by straight lines to highlight the dinucleotide repeat and the zig-zag appearance of the backbone.

nating B DNA proposed by Klug *et al.* (1979). However, there are major structural differences between these three structures (see Table I). Whether the present structure of 5'dACGCGCGT represents a common solution variant cannot be ascertained. It may possibly represent an early intermediate state along the B to Z transition as 5'dACGCGCGT adopts a Z conformation in high salt (>4 M NaCl) as judged by c.d. spectroscopy (Clore and Gronenborn, unpublished data).

The existence of both subtle and more dramatic local sequence-specific variations in the solution structure of B DNA, as illustrated by the structures of the two oligonucleotides presented here, can clearly have a major influence on specific protein-DNA interactions, and it is not difficult to visualise that the zig-zag distribution of the phosphorus atoms around a B DNA helix, as in the case of the solution structure of 5'dACGCGCGT, can present a specific multifaceted array of negative charges to a potentially interacting protein surface.

Materials and methods

5'dCGTACG and 5'dACGCGCGT were prepared by the phosphotriester method as described by Seliger *et al.* (1982) and purified by reverse phase h.p.l.c. using a Walters μ -Bondapak C₁₈ column.

All n.m.r. spectra were recorded on a Bruker AM500 spectrometer. Spectra in D₂O were recorded with a 90° observation pulse (pulse length 9 μ s), an acquisition time of 0.5 s (8196 data points and an 8.2 KHz spectral width) and an interpulse delay of 1 s. Spectra in H₂O were recorded using the time shared hard 1-1 observation pulse for water resonance suppression (Clore *et al.*, 1983) with an acquisition time of 0.366 s (8196 data points and a spectral width of 12.2 KHz) and an interpulse delay of 1 s. The NOEs were observed directly by collecting the difference free induction decay (FID), by interleaving 16 transients after saturation for 0.3 s of a given resonance, with 16 transients of off-resonance irradiation (also applied for 0.3 s), negating the memory between 16 transient cycles. Prior to Fourier transformation the reference and difference FIDs were multiplied by an exponential equivalent to a line broadening of 2 and 4 Hz, respectively. Chemical shifts are expressed relative to 4,4-dimethylsilapentane-1-sulfonate.

The relative error $\Delta x/x$ for the measured NOE values is $\leq \pm 0.15$ which yields an error in the values of the calculated distances of $\leq \pm 0.15$ Å (assuming an error of ± 0.05 Å in the estimated values of the internal reference distances).

Model building was carried out manually using Nicholson skeletal models at a scale of 1 cm to 1 Å.

Acknowledgements

This work was supported by the Medical Research Council (GMC and AMG) and the Lister Institute of Preventive Medicine (GMC). GMC is a Lister Institute Research Fellow. The n.m.r. spectra were recorded on the AM500 spectrometer of the Medical Research Council Biomedical N.M.R. Centre at the National Institute for Medical Research. This paper is dedicated to Professor Friedrich Cramer, in honour of his sixtieth birthday.

References

- Arnott, S. and Hukins, D.W.L. (1972) *Biochem. Biophys. Res. Commun.*, **47**, 1504-1509.
- Arnott, S., Chandrasekaran, R., Puigjaner, L.C., Walker, J.K., Hall, I.H., Birdsall, D.L. and Ratcliff, R.L. (1983) *Nucleic Acids Res.*, **11**, 1457-1474.
- Braun, W., Bösch, C., Brown, C.R., Go, N. and Wüthrich, K. (1981) *Biochim. Biophys. Acta*, **667**, 377-396.
- Calladine, C.R. (1982) *J. Mol. Biol.*, **161**, 343-352.
- Chen, C.-W. and Cohen, J.S. (1983) *Biopolymers*, **22**, 879-893.
- Clore, G.M., Gronenborn, A.M., Mitchinson, C. and Green, N.M. (1982) *Eur. J. Biochem.*, **128**, 113-114.
- Clore, G.M. and Gronenborn, A.M. (1983) *J. Magn. Resonance*, **53**, 423-442.
- Clore, G.M., Kimber, B.J. and Gronenborn, A.M. (1983) *J. Magn. Resonance*, **54**, 170-173.
- Cohen, F.E. and Sternberg, M.J.E. (1980) *J. Mol. Biol.*, **138**, 321-333.
- Cohen, J.S., Wooten, J.B. and Chatterjee, C.L. (1981) *Biochemistry (Wash.)*, **20**, 3049-3055.
- Crippen, G.M. and Havel, T.F. (1978) *Acta Crystallogr.*, **A34**, 282-284.

- Dickerson, R.E. (1983) *J. Mol. Biol.*, **166**, 419-441.
- Dickerson, R.E. and Drew, H.R. (1981) *J. Mol. Biol.*, **149**, 761-786.
- Dobson, C.M., Olejniczak, E.T., Poulsen, F.M. and Ratcliffe, R.G. (1982) *J. Magn. Resonance*, **48**, 87-110.
- Gronenborn, A.M. and Clore, G.M. (1982) *Biochemistry (Wash.)*, **21**, 4040-4048.
- Hare, D.R. and Reid, B.R. (1982) *Biochemistry (Wash.)*, **21**, 5129-5135.
- Heerschap, A., Haasnoot, C.A.G. and Hilbers, C.W. (1983) *Nucleic Acids Res.*, **11**, 4501-4520.
- Johnston, P.D. and Redfield, A.G. (1978) *Nucleic Acids Res.*, **5**, 3913-3927.
- Kan, L.S., Cheng, O.M., Jayaraman, K., Leutzinger, E.E., Miller, P.S. and Ts'o, P.O.P. (1982) *Biochemistry (Wash.)*, **21**, 6723-6732.
- Klug, A., Jack, A., Viswamitra, M.A., Kennard, O., Shakked, Z. and Steitz, T.A. (1979) *J. Mol. Biol.*, **131**, 669-680.
- Kuzmich, S., Marky, L.A. and Jones, R.A. (1982) *Nucleic Acids Res.*, **10**, 6265-6271.
- Lomonosoff, G.P., Butler, P.J.G. and Klug, A. (1981) *J. Mol. Biol.*, **149**, 745-760.
- Neidle, S. and Berman, H.M. (1983) *Prog. Biophys. Mol. Biol.*, **41**, 43-66.
- Noggle, J.H. and Schirmer, R.E. (1971) *The Nuclear Overhauser Effect — Chemical Applications*, published by Academic Press, NY.
- Pardi, A., Walker, R., Rapoport, H., Wider, G. and Wüthrich, K. (1983) *J. Am. Chem. Soc.*, **105**, 1652-1653.
- Patel, D.J., Kozlowski, S.A. and Bhatt, S. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3908-3912.
- Poulsen, F.M., Hosch, J.C. and Dobson, C.J. (1980) *Biochemistry (Wash.)*, **19**, 2596-2607.
- Redfield, A.G. and Gupta, R.K. (1971) *Cold Spring Harbor Symp. Quant. Biol.*, **36**, 405-419.
- Reid, D.G., Salisbury, S.A., Bellard, S., Shakked, Z. and Williams, D.H. (1983) *Biochemistry (Wash.)*, **22**, 2012-2025.
- Rhodes, D. (1982) in Neidle, S. (ed.), *Topics in Nucleic Acid Structures Part 2*, MacMillan Press, London, pp. 287-304.
- Roy, S. and Redfield, A.G. (1983) *Biochemistry (Wash.)*, **22**, 1386-1390.
- Sarma, R.H., Warner, B.G. and Mitra, C.K. (1982) in Sarma, R.H. (ed.), *Biomolecular Stereodynamics, Vol. 1*, Ademine Press, NY, pp. 89-98.
- Scheek, R.M., Russo, N., Boelers, R. and Kaptein, R. (1983) *J. Am. Chem. Soc.*, **105**, 2914-2916.
- Seliger, H., Klein, S., Narang, C.K., Seeman-Preisling, B., Eiband, J. and Havel, N. (1982) in Gassen, G.H. and Lang, A. (eds.), *Chemical and Enzymatic Synthesis of Gene Fragments: A Laboratory Manual*, Verlag Chemie, Mannheim, pp. 81-96.
- Shakked, Z., Rabinovich, D., Kennard, O., Cruse, W.B.T., Salisbury, S.A. and Viswamitra, M.A. (1983) *J. Mol. Biol.*, **166**, 183-201.
- Scheffler, I.E., Elson, E.L. and Baldwin, R.L. (1968) *J. Mol. Biol.*, **36**, 291-304.
- Shindo, H. (1981) *Eur. J. Biochem.*, **120**, 309-312.
- Shindo, H., Simpson, R.T. and Cohen, J.S. (1979) *J. Biol. Chem.*, **254**, 8125-8128.
- Solomon, I. (1955) *Phys. Rev.*, **90**, 559-565.
- Štrop, P., Wider, G. and Wüthrich, K. (1983) *J. Mol. Biol.*, **166**, 641-667.
- Viswamitra, M.A., Shakked, Z., Jones, P.G., Sheldrick, G.M., Salisbury, S.A. and Kennard, O. (1982) *Biopolymers*, **21**, 513-533.
- Wagner, G. and Wüthrich, K. (1979) *J. Magn. Resonance*, **33**, 675-680.
- Wagner, G. and Wüthrich, K. (1982) *J. Mol. Biol.*, **155**, 347-366.
- Wagner, G., Kumar, A. and Wüthrich, K. (1981) *Eur. J. Biochem.*, **114**, 375-384.
- Wider, G., Lee, K.H. and Wüthrich, K. (1982) *J. Mol. Biol.*, **155**, 367-388.
- Wüthrich, K., Wider, G., Wagner, G. and Braun, W. (1982) *J. Mol. Biol.*, **155**, 311-319.