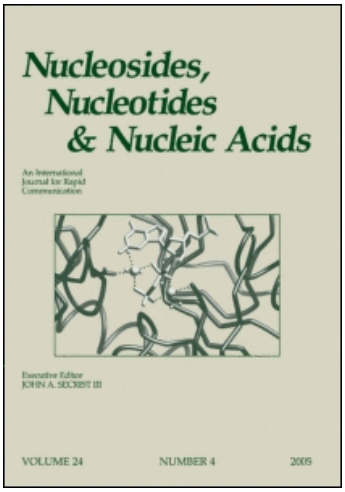


Nucleosides, Nucleotides and Nucleic Acids
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597286

Versatile Oligonucleotides: B DNA, Z DNA, and DNA Hairpins as Seen in Aqueous Solution by Two-Dimensional NMR

Cornells Altona[a]
[a] Gorlaeus Laboratory, University of Leiden, The Netherlands

To cite this Article Altona, Cornells(1987) 'Versatile Oligonucleotides: B DNA, Z DNA, and DNA Hairpins as Seen in Aqueous Solution by Two-Dimensional NMR', Nucleosides, Nucleotides and Nucleic Acids, 6: 1, 157 — 172
To link to this Article: DOI: 10.1080/07328318708056189
URL: http://dx.doi.org/10.1080/07328318708056189

NUCLEOSIDES & NUCLEOTIDES, 6(1&2), 157-172 (1987)

VERSATILE OLIGONUCLEOTIDES: B DNA, Z DNA, AND DNA HAIRPINS AS SEEN IN AQUEOUS SOLUTION BY TWO-DIMENSIONAL NMR

Cornelis Altona
Gorlaeus Laboratory, University of Leiden, The Netherlands

Abstract. Approaches to the study of DNA polymorphism by means of NMR spectroscopy are explored. A method is outlined whereby sequence-dependent variations of sugar pucker in the major conformer of DNA duplexes can be investigated in solution. The results are compared with predictions based upon X-ray studies. The formation of hairpin loops in self-complementary DNA oligomers appears more common than thought previously. Some 'rules of the game' are discussed.

Introduction. DNA is known to occur in a rich variety of three-dimensional forms; to name a few: A, B, Z duplexes, single-helical stacks, hairpin loops.[1,2] Information concerning the crystalline states of DNA is coming mainly from X-ray diffraction of single crystals,[2] whereas NMR, CD and a host of other techniques[1] are employed to probe the solution conformations. The advantages of X-ray crystallography are obvious: one obtains X, Y, Z coordinates of all atoms heavier than hydrogen and insights are gained into architectural details that are hard to come by otherwise. Nevertheless, there are some drawbacks. Even when a good crystal can be obtained (often a bottleneck), the diffraction data do not always allow a resolution good enough to satisfy the needs of demanding structural chemists. Moreover, without additional experiments one can never be certain that the molecular structure - gross or in detail - corresponds to that favoured in solution.

NMR techniques have different advantages: (i) the conformational type of a duplex (A, B, Z DNA) are recognized easily; (ii) structural changes, induced by variation of conditions (temperature, solvent, ionic strength) can be monitored continuously; (iii) thermodynamics of such changes can be determined; (iv) from J-couplings sets of torsion angles can be determined with high accuracy. The latter feature allows the study of local sequence-dependent variations of sugar ring and backbone. Some main disadvantages are: (i) as yet, atomic coordinates cannot

be obtained in an unambiguous way; (ii) signal line broadening, caused by slow tumbling and/or exchange, may prevent the determination of reliable coupling constants.

In the present paper some aspects and strategies of the detailed conformational analysis of DNAs by means of NMR will be discussed in the light of our recent studies of several oligomers: d(GGm^{5}Cm^{5}CGGCC),[3] d(GGCCGGCC),[4] d(m^{5}CGm^{5}CGm^{5}CG),[5,6] d(CCGAATTCGG),[7] d(CCGAm^{6}ATTCGG),[7] and d(CACATGTG).[8] The latest IUPAC-IUB recommendations[9] on nucleic acid nomenclature are adhered to.

Time scales of conformational changes in DNA. Many different conformational changes actually occur in DNA. Because an understanding of the time scales of interconversion that are specific for the different processes is essential for the correct interpretation of NMR data, some of the most important knowledge is briefly reviewed here, in order of decreasing rates and/or increasing life times.

(a) *Ring flip of the deoxyribose sugars.* The furanose ring in nucleosides, nucleotides,[10-13] DNA single helices,[14] and DNA duplexes[3-8,15] preferentially occurs in two distinct and relatively narrow conformational ranges, denoted N and S, Fig. 1, that are separated by a low energy barrier around P 90°. The barrier height of deoxyriboses is not known from experiment, but recent force-field calculations[16] have produced an estimate of 2.9 kcal/mol, irrespective of the attached base. If true, the rate constant of N $\rightleftharpoons$ S interconversion would be close to 0.5 ns^{-1} at 25 °C. However, in the present state of our knowledge a somewhat lower barrier - and thus a higher rate - cannot be excluded from consideration.

(b) *Single-strand stack $\rightleftharpoons$ unstack equilibration.* This also appears to be an ultrafast process, probably occurring on the ns time scale.[1]

(c) *Hairpin loop $\rightleftharpoons$ coil equilibration.* Much slower than the foregoing, but still fast on the NMR chemical shift time scale: a rate constant of 50 μs^{-1} at Tm has been estimated,[1] corresponding to a hairpin life time of about 2 ms.

(d) *Transient base-pair opening.* Considered a moderately fast process. In the lower temperature range life times are typically 1-20 ms, with A.T. < G.C.[17] Life times of terminal base pairs are < 1 ms. Only 1% of base pairs are open at a given time.[1]

(e) *Duplex $\rightleftharpoons$ 2 coil equilibration.* The life time of the duplex is strongly temperature dependent. For small self-complementary DNAs, 8-10 base pairs long, typical life times vary from about 150 ms at low temperature to 1.0 ms at Tm. Exchange line broadening is usually observed in

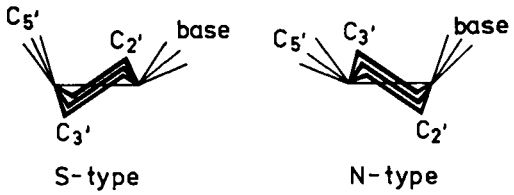


FIG. 1. N- and S-type conformations of furanose rings.[10,11] Within each type a large variation of geometry is theoretically possible, but actual variations are relatively limited.[10-12]

the transition region, but cases of actual doubling of proton signals appear to be rare.

(f) *B-type duplex ⇌ 2 hairpin loops.* This is a slow process; a doubling of proton signals normally occurs.[18-20]

(g) *B DNA ⇌ Z DNA.* A very slow process; life times are usually counted in seconds or minutes.[21]

For all conformationally fast processes - relative to the time scale of the experiment - the ground state equilibrium populations p can be inferred from any suitable measured physical or chemical property Q_{obs}, provided that the properties for the individual conformers Q_I, Q_{II} .. either can be predicted *a priori* or calculated from the experimental data. For the sugar equilibrium, eq. (1) applies:

$$J_{obs} = p(N)\ J_N + p(S)\ J_S \qquad (1)$$

In cases of a slow equilibrium process, integrated signal intensities directly reveal the population ratios.

NMR resonance assignment. The general principles and strategy have been described in detail by several authors and need not be reiterated here.[15] Suffice it to say that the 2D Nuclear Overhauser Spectroscopy (NOESY) experiment gives rise to cross peaks indicative of short (< 5 Å) through-space interproton contacts, both intra- and interresidue. A-RNA and B-DNA type duplexes allow a continuous sequential assignment, because H8/H6, located on residue n in a right-handed stack, is sandwiched between H1', H2', H2" of its own residue and H1', H2'. H2" of residue n-1 'above' on the 5'-neighbouring side. Thus, except at the 5'-terminal residue, each H8 or H6 displays two connectivities to H1' resonances. The same is true for the base proton x H2', H2" cross peaks. The relative NOESY cross-peak intensities allow one to make a quick distiction between A- and B-type duplexes in oligomers.[22] Thus far, no A-type duplexes have been encountered

in aqueous DNA solutions. Moreover, the (cis) H1'...H2'' distance is always shorter than the (trans) H1'...H2' distance and the former NOE should be the most intense of the two. The validity of this useful discrimination is indeed corroborated by the observed coupling constants, see below.

For the Y-R steps in Z-DNA type duplexes a different NOESY pattern analysis applies.[5] Here, one uses the connectivities between H8(n) and H4', H5', H5'' of (n-1).

The cross peaks in the correlated spectroscopy (COSY) experiment betray the existence of scalar J couplings. The normal ^{1}H COSY plot of a DNA yields the n sets (n is the number of magnetically distinct residues present) of seven signals that belong to individual sugars. Various useful extensions of the COSY technique have been described recently: homonuclear double-RELAY,[23,24] 'Exclusive' or E-COSY,[25] Z-COSY,[26] 1D-relayed COSY,[27] and 'differences and sums of traces within COSY spectra' (DISCO).[28,29] Several of these new techniques hold great promise for the accurate determination of couplings or sums of couplings from 'difficult' NMR spectra and hopefully open the way for an even more intimate analysis of sugar ring and backbone in DNA than is possible today, see the next section.

Conformational analysis of duplexes: a strategy. From the stage where the various sets of spin systems of a given DNA fragment have been sequenced (NOESY) and identified (COSY), several approaches are open. The strategy currently developed by the Leiden group is summarized here.

1. Melting profiles. These are constructed for selected resonances, preferably at two or more nucleotide concentrations. These chemical shift profiles are computer processed to yield T_m, ΔH° and ΔS° of the transition of interest as monitored *locally* by each individual proton. This has the advantage over melting profiles obtained from UV or CD spectra, because the latter measure only the overall process. In this way it has been shown repeatedly[3,4,7] that the melting of short DNA duplexes (up to 10-12 base pairs) occurs in an all-or-none fashion, with the possible exception of the terminal bases. Although the T_m values found for the 5'-terminal residues appear normal compared to residues located in the core, the 3'-terminals sometimes behave in a different manner, especially in the case of a pyrimidine residue. For example, the temperature profiles of H5, H6 and H1' of dC(8) in d(GGm^{5}Cm^{5}CGGCC)^3 and in its non-methylated parent[4] suggest that dC(8) is still only partially stacked at temperatures much below T_m of the core. Thus, the shift profiles may be taken to indicate that the 3'-terminal pyrimidine is only gradually engaged in Watson-Crick base pairing with its partner.

2. Interpretation of coupling constant data. In aqueous solutions of B-type DNA the sugars are often found to be engaged in a fast conformational equilibrium: S(major) $\rightleftarrows$ N(minor).[3,4,6] This is also true for the DNA part of RNA-DNA covalent hybrids[30,31] and for the pyrimidine residues in Z-DNA.[6] This means that a complete analysis of conformational behaviour of DNAs should not be limited *a priori* to a description in terms of a single sugar structure (as is often done) but must be carried out at least in terms of the relative populations and of the geometry of S(major). Such an analysis rests upon the correct interpretation of observed proton spin-spin coupling data. The experimental data either may consist of (preferably complete) sets of individual couplings ($J_{1'2'}$, $J_{1'2''}$, $J_{2'3'}$, $J_{2''3'}$, $J_{3'4'}$), determined in the usual fashion by computer simulation of 1D spectra or, alternatively, of sets containing couplings and sums of couplings: $J_{1'2'}$, $J_{1'2''}$, $\Sigma_{2'}$ (= $J_{1'2'} + J_{2'3'} + J_{2'2''}$), $\Sigma_{2''}$ (= $J_{1'2''} + J_{2''3'} + J_{2'2''}$), and $\Sigma_{3'}$ (= $J_{2'3'} + J_{2''3'} + J_{3'4'}$, ^{31}P decoupled). These sums correspond to the distance (Hz) between the outer peaks of the $H_{2'}$, $H_{2''}$, and $H_{3'}$ {^{31}P} signal multiplets, respectively. A set of two individual J's and three Σ values carries the same information as is contained in the set of the five individual J's. The advantage of the use of Σ values lies in the fact that these sums in many cases can be obtained by simple first-order measurement of the distance between well-defined peaks in 1D or 2D NMR spectra without recourse to time-consuming computer simulations.

At this point it is important to note that the sets of individual couplings (and their sums) can be predicted *a priori* with confidence for any desired sugar geometry (< $\pm$ 0.5 Hz for individual J's). Because of ring closure conditions the five endocyclic torsion angles in a five-membered ring, and thus also the corresponding exocyclic CH-CH torsion angles, are mutually related. Conformational analysis of the individual deoxysugars in a stretch of DNA thus boils down to the attainment of an *optimum match* between couplings or Σ values measured along different carbon-carbon bonds and corresponding predicted ones.

The prediction of couplings is based upon three key elements: (i) the concept of pseudorotation,[10,11] by which any sugar conformer to a good approximation is described by two parameters, P and Φ_m; (ii) a set of empirical equations, derived from X-ray and neutron-diffraction data, which correlate proton-proton torsion angles with the endocyclic torsions;[12,32,33] (iii) the empirical generalization of the classical three-parameter Karplus equation by the introduction of three additional terms - based upon quantum-chemical studies[34,35] - which account for the effect of electronegativi-

ty and orientation of substituents upon vicinal couplings in saturated fragments. The generalized equation was parametrized with the aid of 315 experimental J's.[36] Further details can be found in the references cited.

When a complete set of individual couplings is available, the desired optimum match can be achieved with the aid of a non-linear iterative least-squares minimization, program PSEUROT.[37,38] Alternatively, a graphical method[8] serves well when sums of couplings are available, especially in cases where one or more couplings are lacking from the set. This method will be described briefly below. It turns out that the geometry of the S(major) conformer, and thus the important backbone angle δ of double helices in solution, now routinely can be determined with an accuracy at least equal to that of a high-resolution X-ray structure analysis.

Tables of predicted coupling constants of deoxyriboses along the pseudorotation itinerary (at constant Φ_m 35°), corrected for the through-space effect,[39] have been presented elsewhere.[8,14,15] A plot of predicted J's and Σ's *vs* pseudorotation angle P for two values of the puckering amplitude Φ_m (35°, 40°) is shown in Fig. 2A. Experimental results from 24 residues in three different duplexes, after extrapolation to pure S, are also displayed. Points derived from each individual sugar are located on a vertical line, *i.e.* at constant P. The excellent fit lends confidence to the present procedure. The points should actually be visualized as lying on a surface, the Φ_m parameter perpendicular to the plane of the paper. However, Φ_m does not vary much from one residue to another (37° ± 3°).

3. Conformational populations. Examination of Fig. 2A shows that two of the experimentally accessible quantities, Σ_1, and $\Sigma_{2''}$, are quite sensitive to the N/S population ratio and much less so to local geometry variations within the S-type range. If we take into account the finding[10,12] that N-type sugars tend to occupy a relatively narrow P_N range, this means that predicted limiting couplings can be used with confidence in eq. (1). In practice, the population ratio is best approximated from $\Sigma_{2''}$, eq. (2).

$$p(S) = (31.5 - \Sigma_{2''})/10.9 \quad \text{or} \quad p(N) = (\Sigma_{2''} - 20.6)/10.9 \qquad (2)$$

Eq. 2 holds for the normally encountered ranges P_N -9°--+18°, P_S 126°--180°, Φ_m 34°–40°, and $J_{2',2''}$ = -14.0 Hz. The present numbers differ slightly from those published earlier,[15] since Φ_m variation is now taken into account.

Populations p(S) found for individual sugars in three B-DNA duplexes are shown in Table 1. It is seen that only three residues out of 22 attain a state of conformational purity (≥ 95% S), and these three are all dG. Purines in general tend to attain higher S-type purity (∿ 90%) compared to

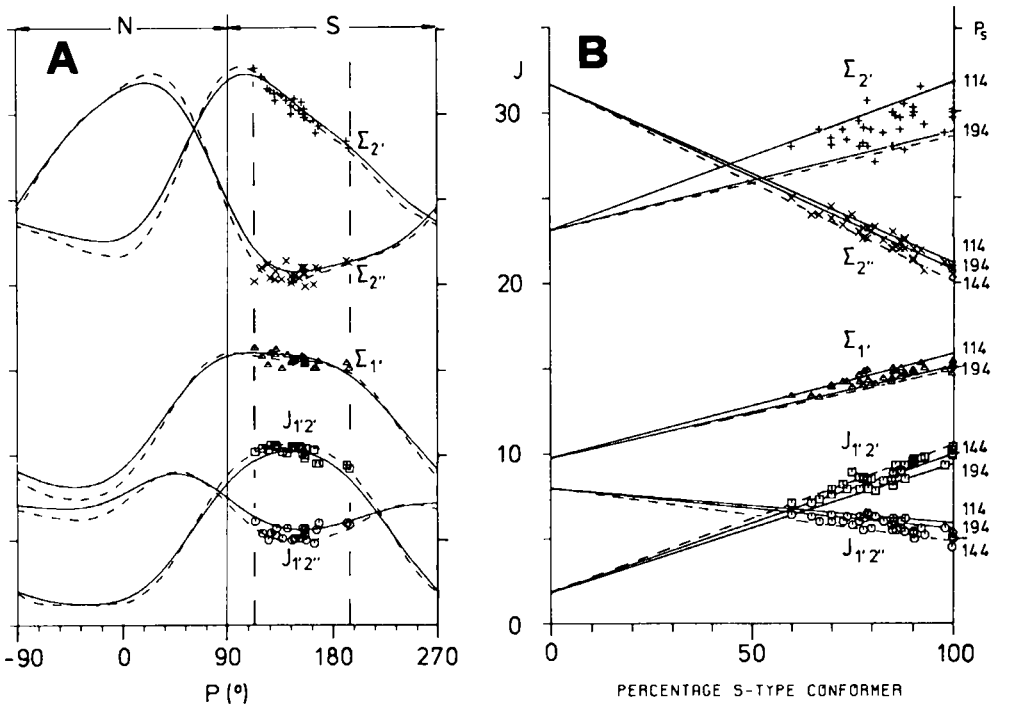


FIG. 2A. Predicted vicinal couplings and sums of couplings (see text) *vs* pseudorotation phase angle P in deoxyriboses. Solid line Φ_m 35°, dashed line Φ_m 40°. Data points are taken from studies of 24 residues in intact B-DNA duplexes. Vertical scale in Hz.

FIG. 2B. Band widths of expectation values of vicinal couplings and sums of couplings in deoxyriboses *vs* populations of S-type conformer, see text. Solid line Φ_m 35°, dashed line Φ_m 40°. Data points are taken from studies of 40 residues in intact B-DNA duplexes.

TABLE 1. Populations p(S) (%) for individual sugars in B-DNA duplexes in aqueous solution (t 27-30 °C)

Sequence	m^5C	G	m^5C	G	m^5C	G			Ref.
p(S)	83	100	90	98	77	74			6
Sequence	G	G	C	C	G	G	C	C	
p(S)	90	85	75	84	100	80	65	65	4
Sequence	C	A	C	A	T	G	T	G	
p(S)	70	93	88	78	92	85	79	79	8

pyrimidines (∿ 80%), although the ranges overlap. A certain dependence on the sequence seems to be present, but further work appears necessary to settle this point.

4. Pseudorotation analysis. From Fig. 2A it follows that $\Sigma_{2'}$ (and $\Sigma_{3'}$, not shown, but see Table 1 in ref. 15) should constitute excellent markers for local P variations within the S-type range. Thus, once the S/N ratio is known for a given nucleotide, analysis of $\Sigma_{2'}$ and $\Sigma_{3'}$ should pinpoint the location of the sugar on the pseudorotational pathway. Fig. 2B serves as an aid to this purpose. The straight lines connect expectation values of $J_{1'2'}$, $J_{1'2''}$, $\Sigma_{1'}$, $\Sigma_{2'}$, and $\Sigma_{2''}$ for an equilibrium between a chosen N-type conformer (P_N 9°, Φ_N 35°) on the left-hand side and several possible S-type conformers (P_S 114°–194°, Φ_S 35°–40°) on the right-hand side. The data points concern 40 individual sugars in a collection of DNA duplexes in aqueous solution, measured in our laboratory. Points derived from each sugar are located on a vertical line, *i.e.* at constant p(S). It is gratifying to note that the great majority of the experimental data indeed is found to lie within the predicted band widths. The four independent variables appear mutually consistent, as they should be. This means that the observed $\Sigma_{2'}$ (and $\Sigma_{3'}$) values may serve to determine the phase angle P_S.[8] Comparison with Fig. 2A reveals that any significant contribution of high-N conformers (P_N > 18°) should be readily revealed by scatter in the $\Sigma_{1'}$ and $\Sigma_{2'}$ plots upon increasing N-type population and this is not observed.

5. Sugar variation in intact B-DNA duplexes. Fig. 3A displays the distribution of P_S values of 24 sugars in three B-DNA duplexes in aqueous solution. Such a distribution is not entirely unexpected, since X-ray studies of 18 S-type deoxyribofuranosides[12] reveal a similar picture. Moreover, a recent rerefinement[40] of the crystal structure data of the 12-mer. d(CGCGAATTCGCG)[41,42] has yielded a comparable, though slightly wider, P_S distribution. In particular, four out of 24 residues in the rerefined 12-mer structure occur in the region P_S 72° ($^0_4$T) – 108° ($^0_1$T), which region remains empty in the present solution studies. This apparent discrepancy can be resolved as follows. Various authors[40,41] have noted a peculiar packing of the helix ends in the crystal and interduplex hydrogen bonding. This finding, combined with the lack of symmetry seen for formally equivalent residues in the two halves of the crystalline duplex, strongly suggests that sugar pucker indeed is affected significantly by packing forces. If true, NMR should be the method of choice in studies concerned with sugar pucker variations.

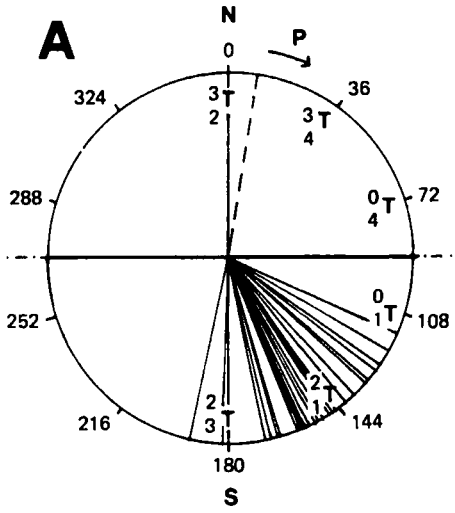


FIG. 3A. Conformational wheel representing the distribution of sugar geometries (P values) of 24 residues in intact B-DNA duplexes determined in aqueous solution by means of NMR. The dashed line represents the minor N-type conformer.

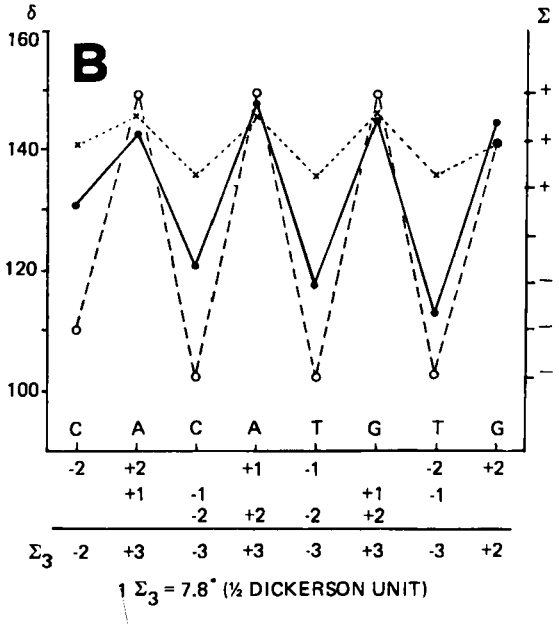


FIG. 3B. The sequence-dependent variation of backbone angle δ of the major S-type conformation of the d(CACATGTG) duplex in aqueous solution (filled circles). Crosses indicate predicted values for the corresponding single helix,[43,45] open circles indicate predicted values from the Σ_3 sum function[42] shown at the bottom of the figure.

6. Backbone angle δ. The important backbone angle δ (C5'-C4'-C3'-O3') is directly related to the geometry adopted by the sugar ring eq. (3).

$$\delta = 120.6 + 1.1\ \Phi_m \cos(P + 145.2) \qquad (3)$$

Eq. (3) is based upon an empirical data set[12] in conjunction with an extended pseudorotation equation.[13] The variation of δ in single-helical stacks is not large (± 5° from δ(mean) 141°) and appears predictable[15,43] from a simple sum function Σ_δ, although cross-chain steric clashes are absent. It has been suggested that sliding of bases is involved in order to maximize base-base overlap when the 5' neighbour is a purine.[15,43]

The behaviour of δ in double-helical B DNA appears more complicated. Fig. 3B shows that the δ angles of the 8-mer duplex d(CACATGTG) vary from 113° to 148°, δ(mean) 133°.[8] The pyrimidines in the core of the 8-mer occupy the lower part of the δ range (113°-121°), whereas the purine nucleotide sugars cluster in the higher part (143°-148°). On first sight, these phenomena seem to accord with predictions from the Calladine/Dickerson Σ_3 sum function,[42,44] but a closer look reveals certain peculiarities. The crystal structure data[41] of the 12-mer d(CGCGAATTCGCG) led Calladine[44] to suggest that base pairs can slide along their long axis in order to relieve cross-chain clashes between purine bases produced by propeller twist. Such sliding supposedly causes an increase of δ at the purine side of the base pair concomitant with a decrease of the same magnitude at the pyrimidine site (anticomplementaity principle). For a given base pair, the difference Δδ = δ(strand I) − δ(strand II) is predicted[42] by assigning unit values of +1, -1 to the sugars at each R-Y step; at each Y-R step the values are doubled and inverted in sign: -2, +2. Additivity of these numbers is assumed, see the bottom of Fig. 3B. According to Dickerson,[42] each unit of Σ_3 corresponds to 15.6° of Δδ.

By means of NMR each δ can be monitored with reasonable precision (± 5°) and there is no *a priori* reason to use Δδ. Therefore, Fig. 3B displays δ angles deduced for the 8-mer duplex, together with those predicted for the corresponding single helix (Σ_δ function) and predictions according to the Σ_3 sum function. For the latter the unit value is taken equal to one half the Δδ unit (7.8°) and δ(mean) 126° is taken from Westhof's re-refinement.[40]

It is interesting to note (Fig. 3B) that the δ angles obtained from NMR data follow the predicted *trend* for this alternating Y-R sequence. However, the *magnitude* of unit δ change amounts to about 4.5° and thus remains significantly smaller than that deduced from the X-ray data (7.8°). Moreover, the measured and predicted points for all purine residues are clus-

tered closely (δ(exp) 145° ± 3°, Σ_δ 145°, Σ_3 149°). Comparing the δ angles of the duplex purines to the overall δ(mean) found for 61 *single helical* residues[15,45] (141°), one sees only a small increase (+ 4°). In contrast, the pyrimidine sugars in the core of the duplex experience a spectacular decrease of δ angles (NMR -24°, Σ_3 predicted -38°) compared to the single helix. In our opinion, this finding throws some doubts upon the validity of the anticomplementarity principle. The pyrimidines appear to have a much greater capacity for δ change than the purines.

Closely similar results (not shown) have been obtained[6] for the alternating 6-mer d(m^5CGm^5CGm^5CG), while parallel NMR researches[7] on the 10-mer d(CCGAATTCGG) indicate that the present Σ_3 function does not account well for the experimental data. Nevertheless, the combined NMR results decidedly favour the concept of local variations of geometry in an intact B-DNA duplex, which concept was pioneered by Dickerson and his colleagues.[41,42,44]

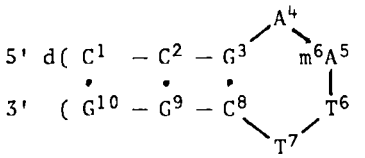
Hairpin loops. Any self-complementary oligomeric stretch of DNA in principle can occur as a duplex dimer or, alternatively, in the form of a hairpin-loop monomer if the stretch is not too short. The following questions will be briefly discussed: (i) what are the rules for the detection of such hairpins in NMR spectra, and (ii) what are the conditions that promote hairpin stability? In our laboratory,[7,20] as well as elsewhere,[18,19] hairpin-loop species of self-complementary DNA oligomers have been observed in NMR and some identification rules have been established. The equilibrium:

$$\text{dimer duplex} \rightleftharpoons 2 \text{ monomer hairpins}$$

is slow on the NMR time scale and signals of the two species are observed side by side. The overall equilibrium situation is in fact tripartite, the monomeric hairpin occurring in fast equilibrium with the monomeric 'random coil' form. This means that the chemical shift *vs* temperature profiles of the total monomeric species (hairpin ⇌ coil) display a sigmoidal shape. Such profiles can be analysed in the usual manner[7] to yield the thermodynamics of hairpin melting; T_m, $\Delta H°$ and $\Delta S°$. Of course, T_m of the monomeric species is independent of the nucleotide concentration C_M (provided the ionic strength is kept constant), whereas the T_m of the dimeric species increases with C_M according to the mass law.[1] It was also often noted[7,20,46] that the relative population of the dimer increases with increasing ionic strength (at constant C_M and temperature) at the expense of the monomer.

The lengths of the stem and loop regions can be gauged from various chemical shift criteria.[7] (A) the non-exchangeable protons of residues in the stem region - except those of the inner pair next to the loop - reso-

nate at the 'correct' position, 'correct' meaning close to the corresponding duplex resonances. Many non-exchangeable protons in the loop region display 'anomalous' chemical shifts.



For example, the base protons H6/H8/H2 of residues C^1, C^2, G^9, and G^{10} in the stem of the hairpin 10-mer, depicted above, resonate within 0.1 ppm of the corresponding duplex signals, whereas those of G^3 and C^8 differ by + 0.34 and - 0.21 ppm, respectively. These differences are partly ascribed to a change in stacking pattern and partly to the disruption of continuous base pairing in the loop region. The base protons located in the loop display varying differences, up to -0.53 ppm (H_2 of A^4). (B) The imino protons of the stem resonate at the 'correct' positions, whereas those of the loop display an upfield shift amounting to several ppm.[7,20,47] Moreover, it was found[20,48] that the non-base-paired imino protons of the loop display a marked excess line broadening when the pH is raised from 6 to 8. Last but not least, a loop structure may be revealed by a discontinuity in the NOESY connectivity pattern.

An example of the consideration discussed above is provided[7] by the thymine methyl protons of the N^6-methylated 10-mer d(CCGAm^6ATTCGG), Fig.4. The 10 mM sample at 22 °C displays two methyl signals, as expected. Upon dilution to 2 mM two new sharp signals appear and these dominate the spectrum at 0.5 mM. Similar drastic changes are seen for the base-proton resonances. The non-methylated parent 10-mer displays similar phenomena, although the hairpin populations remain markedly smaller. At this point it should be reiterated that the ratio of the integrated signal intensities reveals the ratio duplex/monomer at each temperature, whereas the ratio hairpin/coil comes from analysis of the chemical shift *vs* T profiles. The results of such a three-state population analysis are displayed in Fig. 5. At a given C_M, the fraction hairpin first increases upon increasing temperature and then passes a maximum. Careful adjustment of the temperature is required in order to achieve the optimum relative population of hairpin loop, especially at low concentration.

The thermodynamic parameters of the three-state equilibrium: random coil (RC)/hairpin (HP)/double helix (DH) of the N^6-methylated 10-mer and of its parent are shown in Fig. 6. N^6-methylation is seen to destabilize the duplex by about 0.5 kcal/mol of m^6A, whereas the hairpin stability is

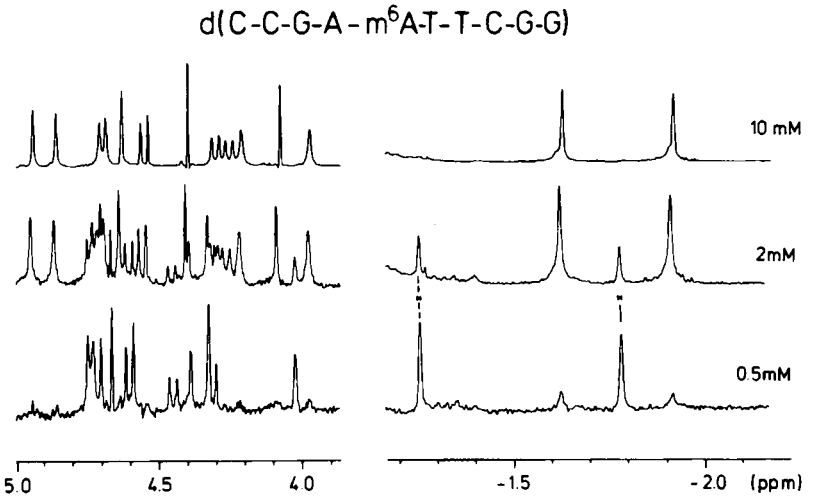


FIG. 4. NMR spectra of a self-complementary 10-mer that folds into a hairpin-loop structure upon dilution. Base proton resonances at the left-hand side, thymine methyl resonances at the right-hand side. Reference of chemical shift scale is Me_4NCl, 3.18 ppm upfield from DSS.

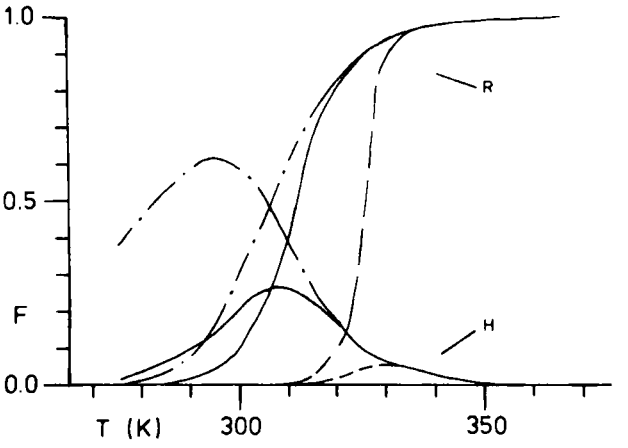


FIG. 5. Fractional equilibrium populations (F) of hairpin (H) and random coil (R) *vs* temperature of d(CCGAm^{6}ATTCGG) at three concentrations: --- 10 mM, — 2 mM, — · — 0.5 mM. For the sake of clarity the fraction duplex has been omitted from the figure.

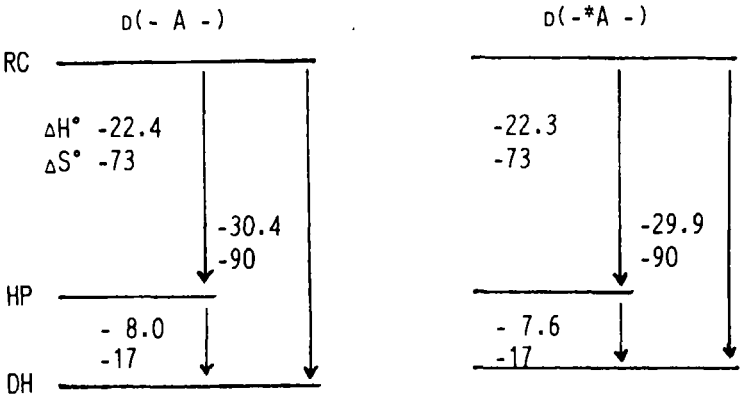


FIG. 6. Thermodynamic parameters (kcal/mol single strand) for the 10-mers d(CCGAATTCGG) (left) and d(CCGAm^{6}ATTCGG) (right); [Na$^+$] 8mM; RC random coil, HP hairpin-loop, DH double helix.

hardly affected. These results confirm and extend earlier conclusions by Engel and von Hippel.[49]

Summarizing, it now appears quite likely that any circumstance that locally destabilizes the center part of a self-complementary stretch of double-helical DNA, such as N^6-methylation,[7] mismatches,[20] platination, covalent binding of carcinogens like N-acetoxy-N-2-acetylaminofluorene,[50] and so forth, will promote the formation of a hairpin loop (or cruciform structure), especially at low and moderate ionic strengths.

Acknowledgement. This research was supported by the Netherlands Foundation for Chemical Research (S.O.N.) with financial aid from the Netherlands Organization for the Advancement of Pure Research (Z.W.O.). Mr L.J. Rinkel and mr L.P.M. Orbons kindly supplied much of the material used in this paper. The syntheses were carried out by Prof. J.H. van Boom and Dr G.A. van der Marel.

REFERENCES

1. Cantor,C.R.; Schimmel,P.R. *Biophysical Chemistry*, (1980) Vol.I-III, Freeman & Co, San Francisco.
2. Saenger,W. "Principles of Nucleic Acid Structure", Springer Verlag, Heidelberg, 1984.
3. Sanderson,M.R.; Mellema,J.-R.; van der Marel,G.A.; Wille,G.; van Boom,J.H.; Altona,C. *Nucleic Acids Res.* (1983) 11, 3333-3346.
4. Rinkel,L.J.; Sanderson,M.R.; van der Marel,G.A.; van Boom,J.H.; Altona,C. *Eur. J. Biochem.* (1986) 159, 85-93.
5. Orbons,L.P.M.; van der Marel, G.A.; van Boom,J.H.; Altona,C. *Eur.J. Biochem.* (1986) in press.
6. Orbons,L.P.M.; Altona,C. *Eur. J. Biochem.* (1986) in press.
7. Rinkel,L.J.; van der Marel, G.A.; van Boom, J.H.; Altona,C. submitted for publication.
8. Rinkel,L.J. *et.al.* Manuscript in preparation.

9. IUPAC-IUB Joint Commission on Biochemical Nomenclature, *Eur. J. Biochem.* (1982) 131, 9-15.
10. Altona,C.; Sundaralingam,M. *J.Am.Chem.Soc.* (1972) 94, 8205-8212.
11. Altona,C.; Sundaralingam,M. *J.Am.Chem.Soc.* (1973) 95, 2333-2344.
12. de Leeuw, H.P.M.; Haasnoot, C.A.G.; Altona, C. *Isr. J. Chem.* (1980) 20, 108-126.
13. de Leeuw,F.A.A.M.; van Kampen,P.N.; Altona,C.; Diez,E.; Esteban,A.L. *J. Mol. Struct.* (1984) 125, 67-88.
14. Altona,C. *Recl. Trav. Chim. Pays-Bas* (1982) 101, 413-433.
15. Altona,C. *J. Mol. Struct.* (1986) 141, 109-125.
16. Pearlman,D.A.; Kim, S.-H. *J. Biomol. Struct. Dyns.* (1985) 3, 99-125.
17. Leroy,J.L. (1986) private communication.
18. Wemmer,D.E.; Chou,S.H.; Hare,D.R.; Reid,B.R. *Nucleic Acids Res.* (1985) 13, 3755-3772.
19. Summers,M.F.; Byrd,R.A.; Gallo,K.A.; Samson,C.J.; Zon,G; Egan,W. *Nucleic Acids Res.* (1985) 13, 6375-6386.
20. Orbons,L.P.M.; van der Marel,G.A.; van Boom,J.H.; Altona,C. *Nucleic Acids Res.* (1986) 14, 4187-4196.
21. Pohl,F.M.; Jovin,T.M. *J. Mol. Biol.* (1972) 67, 375-396.
22. Haasnoot,C.A.G.; Westerink,H.P.; van der Marel,G.A.; van Boom,J.H. *J. Biomol. Struct. Dyns.* (1984) 2, 345-360.
23. Eich,G.; Bodenhausen,G.; Ernst,R.R. *J.Am.Chem.Soc.* (1982) 104, 3731-3733.
24. Bax,A.; Drobny,G. *J. Magn. Reson.* (1985) 61, 306-320.
25. Griesinger,C.; Sørensen,O.W.; Ernst,R.R. *J.Am.Chem.Soc.* (1985) 107, 6394-6396.
26. Oschkinat,H.; Pfändler,P.; Bodenhausen,G. Abstr. 8th Eur.NMR.Conf. no 83 (1986).
27. Bermel,W.; Kessler,H.; Griesinger,C.; Oschkinat,H. Bruker Report 22-25 (1986).
28. Kessler,H.; Müller,A.; Oschkinat,H. *Magn.Reson.Chem.* (1985) 23, 844-852.
29. Oschkinat,H; Freeman,R. *J.Magn.Reson.* (1984) 60, 164-169.
30. Mellema,J.-R.; Haasnoot,C.A.G.; van der Marel,G.A.; Wille,G.; van Boeckel,C.A.A.; van Boom,J.H.; Altona,C.; *Nucleic Acids Res.* (1983) 11, 5717-5738.
31. Haasnoot,C.A.G.; Westerink,H.P.; van der Marel,G.A.; van Boom,J.H. *J.Biomol.Struct.Dyns.* (1983) 1, 131-149.
32. Haasnoot,C.A.G.; de Leeuw,F.A.A.M.; de Leeuw,H.P.M.; Altona,C. *Org. Magn.Reson.* (1981) 15, 43-52.
33. de Leeuw,F.A.A.M.; Altona,C. *J.Chem.Soc.Perkin II* (1982) 375-384.
34. Pachler,K.G.R. *Tetrahedron* (1971) 27, 187-199.
35. de Leeuw,F.A.A.M.; Haasnoot,C.A.G.; Altona,C. *J.Am.Chem.Soc.* (1984) 106, 2299-2306.
36. Haasnoot,C.A.G.; de Leeuw,F.A.A.M.; Altona,C. *Tetrahedron* (1980) 36, 2783-2792.
37. de Leeuw,F.A.A.M.; Altona,C. *J.Comput.Chem.* (1983) 4, 428-437.
38. de Leeuw,F.A.A.M.; Altona,C. Quant.Chem.Progr.Exch. no 463, Indiana University at Bloomington (1983).
39. de Leeuw,F.A.A.M.; van Beuzekom,A.A.; Altona,c. *J.Comput.Chem.* (1983) 4, 438-448.
40. Westhof,E., Private communication.
41. Dickerson,R.E.; Drew,H.R. *J.Mol.Biol.* (1981) 149, 761-786.
42. Dickerson,R.E. *J.Mol.Biol.* (1983) 166, 419-441.
43. Mellema,J.-R.; Pieters,J.M.L.; van der Marel,G.A.; van Boom,J.H.; Haasnoot,C.A.G.; Altona,C. *Eur.J.Biochem.* (1984) 143, 285-301.
44. Calladine,C.R. *J.Mol.Biol.* (1982) 161, 343-352.
45. Mellema,J.-R. Thesis, Leiden (1984).
46. Marky,L.A.; Blumenfeld,K.S.; Kozlowski,S.; Breslauer,K. *Biopolymers* (1983) 22, 1247-1257.

47. Haasnoot,C.A.G.; den Hartog,J.H.J.; de Rooÿ,J.F.M., van Boom,J.H., Altona,C. *Nucleic Acids Res.* (1980) 8, 169-181.
48. Haasnoot,C.A.G.; de Bruin,S.H.; Berendsen,R.G.; Jansen,H.G.J.M.; Binnendijk,T.J.J.; Hilbers,C.W.; van der Marel,G.A.; van Boom,J.H. *J.Biomol.Struct.Dyns.* (1983) 1, 115-129.
49. Engel,J.D.; von Hippel,P.H. *J.Biol.Chem.* (1978) 253, 927-934.
50. Fuchs,R.P.P.; Schwarz,N.; Daune,M.P. *Environ.Health Perspect.* (1983) 42, 135-140.