Volume 11 Number 16 1983 Nucleic Acids Research

Proton NMR studies on the covalently linked RNA-DNA hybrid r(GCG)d(TATACGC). Assignment of proton resonances by application of the nuclear Overhauser effect

J.-R.Mellema, C.A.G.Haasnoot*, G.A. van der Marel, G.Wille, C.A.A. van Boeckel, J.H. van Boom and C.Altona

Gorlaeus Laboratories, State University of Leiden, PO Box 9502, 2300 RA Leiden, and *Faculty of Science, University of Nijmegen, Biophysical Department, Toernooiveld 6525 ED Nijmegen, The Netherlands

Received 15 June 1983; Accepted 21 July 1983

ABSTRACT:

Proton NMR spectra of a covalently linked self-complementary RNA•DNA hybrid, r(GCG)-d(TATACGC), are recorded in H_2O and D_2O. Imino proton resonances as well as the non-exchangeable base and H-1' resonances are unambiguously assigned by means of nuclear Overhauser effect measurements. Additional information was obtained by ^{31}P NMR and circular dichroism spectra. The RNA parts in the duplex attain full conformational purity and adopt the usual A-RNA conformation. The DNA residues opposite the RNA tract do not adopt an A-type structure completely. Their respective sugar rings still appear to possess a certain conformational freedom. The same holds true for the central d(-TATA-) sequence which forms a DNA•DNA duplex. There appears to be a structural break in this part: the first two residues, T(4) and A(5), are clearly influenced by the adjacent RNA structure, whereas residues T(6) and A(7) behave quite similar to what usually is found in DNA duplexes in aqueous solution.

INTRODUCTION: [1,2]

DNA replication is known to be a process of extremely high fidelity[3]. This fidelity requires very efficient control of the enzymes involved. Especially the stage of getting started (initiation) is prone to errors in the replica being formed. DNA dependent DNA polymerases are known to use several possible mechanisms of priming the biosynthesis of the new DNA chain [4,5]. Initiation of discontinuous replication of bacteriophage M13 DNA in *E.coli* was found to be inhibited by rifampicin, an inhibitor of RNA synthesis [6]. Covalently bound RNA-DNA hybrids (so called "Okazaki-fragments") were proposed and experimentally demonstrated as intermediates in DNA replication [7,8].

At present little detailed experimental evidence concerning the conformational consequences of covalent RNA-DNA hybrids appears available. For this reason we studied the structure in solution of a deliberately tailored fragment: r(GCG)-d(TATACGC). At the outset of our investigation the following picture was generally accepted. From X-ray studies [9,10] as well as from investigations carried out on solutions [11,12] it was believed that the RNA part in RNA•DNA hybrids dominates the DNA geometry, *i.e.* the complementary strand

adopts an A-DNA type structure. It remained a matter of debate whether or not this A-DNA structure is propagated into the DNA•DNA duplex part and, if so, how many nucleotidyl residues are involved. Selsing *et al.* studied block polymers of the general type $rC_idC_k \bullet dG_n$ by means of 300 MHz NMR [11] and CD [13]. It was found that the DNA•DNA part maintains a B conformation despite the adjacent A conformation of the RNA•DNA hybrid part. The NMR spectrum was not sufficiently resolved to provide information concerning the perturbation of the helix at the junction of the conformations. However, the close agreement between experimental and calculated CD spectra led to the conclusion that any disruption at the junction is probably only one or two base pairs in length. On the basis of these findings, a one-base-pair bend model of the DNA residues located at the junction was proposed [14] which involves S-type sugar pucker for the first deoxyribose unit (strand running 5' → 3'), *i.e.* T(4) in our compound, and N-type sugar conformation for its complementary residue, *i.e.* A(7).

In the present paper we describe a 500 MHz ^1H NMR study (and some additional ^{31}P NMR and CD data) on the title compound, including an assignment of the exchangeable imino and non-exchangeable base and 1' protons by means of the nuclear Overhauser effect (NOE). A relatively large molecule like the title compound poses several problems for the NMR spectroscopist. In the first place, larger molecular weights (> 3000) and concomittant conformational behaviour (e.g. rigidity; relatively slow exchange processes; slow overall tumbling rates) may result in severe line broadening due to short transversal relaxation times T_2 or coalescence. These phenomena often preclude detailed NMR studies of spin-spin coupled resonances of large biomolecules. Secondly, the large number of crowded resonances causes great difficulties with the traditional assignment procedures, e.g. the incremental method [15], the additivity method [16] and decoupling experiments (see for a discussion of various assignment strategies [17]). A magnetic resonance study of a nucleic acid decamer balances on the present border of possibilities for assigning non-exchangeable resonances [18,19]. In most published cases of assignments above the trimer level, knowledge of constituent oligomers was necessary to complete the picture [12,20-23]. It will be demonstrated in this study that nuclear Overhauser effect spectra provide an extremely powerful tool for the assignment of oligonucleotides. Direct connectivities between base, 1' and 2',2" protons within a given residue and between neighbouring residues enabled assignment of these signals in the hybrid decamer spectrum from only 17 NOE-difference spectra. The method presented here has several advantages above other published methods for oligonucleotide assignment:

- No constituent compounds are needed as in the case of incremental assignments [15,18].
- The assignment can be carried out at low temperatures where the molecule exists in the natural, and conformationally interesting, duplexed state. Recently, an assignment procedure based upon empirical chemical shift additivity parameters at high temperature was published [16]. Although this method appears to be highly useful in the case of smaller fragments, it is bound to fail for duplexes that remain largely intact at 70 °C or are rapidly exchanging between duplex and random coil structures at that temperature.
- The H-1' resonances can be assigned directly in a single experiment instead of using an intermediate step *via* the H-2' resonances [18].

The method described in this paper has already been successfully employed in our laboratories to assign the base and H-1' resonances of the DNA octamer d(GGC*C*GGCC) [24] (C*= m^{5}C). Independently from our work, an essentially similar approach led to recent assignments in ribonucleotides up to a hexamer [25,26].

Recently, a single-crystal study on the title compound was published [27] which showed that the molecule crystallized as a completely A-type double helix. However, Wang *et al.* [27] indicated some structural irregularities in the central -TATA- part. It will be demonstrated from the present proton NMR data that none of the DNA residues attains full conformational purity in solution, not even the ones opposing the RNA tract. The RNA residues, however, show 100 % N-type sugar pucker. Half of the central -TATA- sequence, residues T(4) and A(5), shows a tendency towards N-type sugar pucker (A-DNA type), whereas residues T(6) and A(7) show a predominant S-type sugar pucker as is usually found in B-DNA structures. It can be concluded from these findings that the influence of a small RNA tract on DNA conformation is different from that assumed earlier insofar its own strand is concerned and may extend to two nucleotidyl residues. In contrast, the influence on the complementary strand in the present hybrid molecule appears less outspoken than expected at the outset of this investigation.

MATERIALS AND METHODS:

The oligonucleotides under investigation were synthesized by a phosphotriester approach as has been described elsewhere [28-30]. After purification on a Sephadex G50 column the compounds were passed over a Dowex cation-exchange resin (Na^{+} form) to yield the sodium salts. NMR samples for observation of the imino-proton signals were prepared by dissolving 3-5 mg of the

purified oligonucleotide in 0.4 ml of a buffer solution containing 0.1 M NaCl, 5 mM sodium cacodylate (pH 7.2), 0.1 mM EDTA and 15:85 (v/v) D_2O/H_2O. Measurements of the non-exchangeable proton resonances were carried out on samples in 99.9 % D_2O obtained by three-fold lyophilization from D_2O. Salt concentrations were the same as for the H_2O samples and pD values were adjusted to 7.4 throughout. All samples were carefully degassed by repeated freezing/thawing under vacuum and transferred into 5-mm NMR samples tubes (Wilmad 528-PP or Stohler S-4) under a dry nitrogen atmosphere.

Imino-proton spectra were recorded at 500 MHz on a Bruker WM-500 spectrometer using a "Redfield 214" weak pulse [31,32]. Data were collected using the alternate-delay-accumulation (ADA) technique. The combination of both techniques as well as the application of a time-shared version of the Redfield pulse [32,33], which was used to record NOE difference spectra in H_2O (Figure 2), was recently described by Haasnoot and Hilbers [34]. NOE experiments were carried out at 25 °C to assign the imino-proton resonances, employing a preirradiation pulse of 0.4 sec. The irradiation power was set to a level that just saturated the irradiated resonance completely in order to obtain an optimal S/N ratio. This point was checked by inspection of an on-resonance spectrum before subtraction from an off-resonance spectrum. Chemical shifts of the imino-proton resonances were measured relative to the solvent H_2O peak and converted to DSS reference by correcting for the H_2O to DSS chemical shift using the appropriate calibration curve [35,36].

NMR spectra of the non-exchangeable resonances were recorded at 500 MHz using the DASWEFT solvent suppression technique [37]. A DANTE pulse [38] is applied to create the 180° pulse in a [180°-delay-90°-acq.] sequence in order to selectively reverse the magnetization of the solvent peak. This method appeared to be a prerequisite for succesful measurements because in the case of the degassed samples of the hybrid decamer the longitudinal relaxation times T_1 of several of the solute resonances (in particular those of the ribose 1' protons) are of the same order of magnitude as T_1 of water. This causes a normal WEFT sequence [39,40] to fail. Resolution enhancement was carried out by a Lorentzian to Gaussian transformation as provided by the Bruker DISNMRP program prior to Fourier transformation. Free induction decays were accumulated on 8K datapoints and the FIDs were zero-filled to 32K in order to obtain a satisfactory digital resolution.

NOE experiments were carried out at 32 °C employing a preirradiation pulse of 1.0 sec. Notwithstanding this relatively long preirradiation pulse, no spin diffusion was detected as can be seen in Figs. 4 and 5. The on- and

off-resonance FIDs were subtracted for each series of 16 scans during accumulation. In this way a large number of scans (≥ 1000) can be accumulated to increase the signal to noise ratio significantly without invoking dynamic-range problems [41]. The NOE difference FIDs were processed using an exponential window before Fourier transformation. Chemical shifts of the non-exchangeable resonances are expressed relative to internal TMA reference (0.1 mM). The H-1' resonances were computer-simulated using the program LAME.

Phosphorus-31 {^{1}H} NMR spectra were recorded at 121.5 MHz on a Bruker WM-300 spectrometer at ambient temperature. The accumulation of 800-1000 scans proved satisfactory. Spectra were processed using Gaussian windows and zero-filling. Chemical shifts are indicated relative to TMPB which served as an external reference [42].

CD spectra were recorded at 2 °C on a CNRS-Roussel-Jouan III dichrographe (Jobin-Yvon, France), interfaced with an Apple-II computer. Ellipticities are given in $M^{-1}cm^{-1}$. Absorbancies were typically 0.80 indicating molar concentrations of ≈ $1.0x10^{-5}$. An estimated extinction coefficient of 7,856 per residue was used [43]. Samples contained 0.1 M NaCl, pH 7.2.

RESULTS AND DISCUSSION:

Imino-proton resonances.

Figure 1 shows the exchangeable imino-proton resonances of the hybrid together with those of the deoxy analogue. The spectra were recorded at 15 °C at

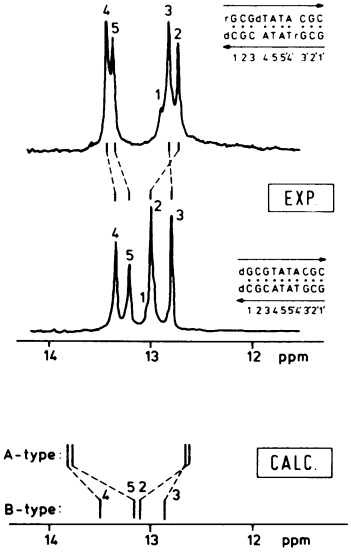


Figure 1. 500 MHz ^{1}H NMR spectra of the exchangeable imino-proton resonances of r(GCG)-d(TATACGC) and d(GCGTATACGC) recorded at 15 °C (DSS reference). Chemical shift calculations according to Patel and Tonelli [44] for this specific sequence are indicated by the vertical bars at the bottom.

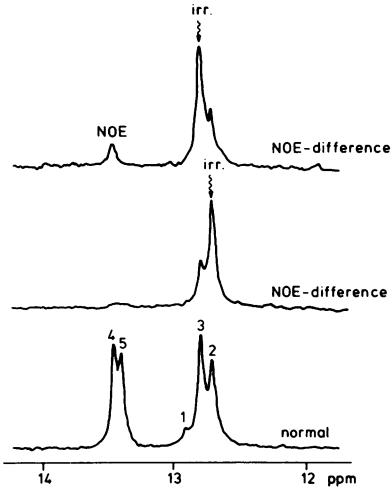


Figure 2. 500 MHz NOE difference spectra (DSS reference) of the imino-protons of r(GCG)-d(TATACGC). Irradiation of G•C base pairs 2 and 3 is shown.

which temperature the decamers occur mainly in the duplex form. CD melting studies on the deoxy compound yield a T_m of about 55 °C at 0.1 M NaCl (data not shown). At the usual concentration of the NMR samples one expects then a T_m of about 70 °C. A CD melting study of the hybrid unfortunately did not allow the determination of an accurate T_m value because the ellipticity extrema (see Fig. 9) shift significantly during melting. Some indication of the stability of the hybrid duplex was obtained from the temperature dependence of the line widths of the non-exchangeable base resonances when the temperature is increased. The greater number of these resonances start to broaden, reaching a maximum (coalescence) at around 55 °C, and become sharper again at higher temperatures. This coalescence phenomenon is ascribed to slow exchange between duplex (D) and random coil (C) forms. Although it is realized that the coalescence temperature depends upon the rate of the $D \rightleftarrows 2C$ reaction and T_m upon the equilibrium thermodynamics, it may be assumed that, to a rough approximation, the stability of the hybrid duplex does not differ greatly from that of the deoxy compound.

It can be seen in Fig. 1 that, even at 15 °C, G•C base-pair signal 1 suffers from exchange with the solvent (fraying ends). Assignment of G•C base-pair signals 2 and 3 of both compounds was initially based on their relative "melting" behaviour, *i.e.* at increasing temperature signal 2 collapses and disappears before that of G•C 3. These assignments were fully confirmed by NOE experiments designed in order to differentiate between the signals of A•T pairs 4 and 5. Figure 2 shows the crucial spectra. Irradiation of G•C resonance 2 did not produce NOE in the A•T region of either spectrum, whereas irradiation of

the G•C signal 3 resulted in a clear NOE of the signal designated 4 (see Fig. 2). It follows that the remaining signal 5 belongs to the central A•T pairs. Irradiation of base-pair signals 4 and 5 did also produce NOEs in the base-proton region (not shown). These NOEs are ascribed to the adenine H-2 protons (as indicated in Fig. 3), which are positioned close to the thymine N_3-H hydrogen-bonded imino proton in a Watson-Crick base pair.

It is seen (Fig. 1) that the imino-proton spectrum of the hybrid decamer differs in several important aspects from that displayed by the DNA decamer, and this implies certain differences in the respective structures. On going from the DNA compound to the hybrid one notes that the resonance of G•C base pair 2 moves upfield by about 0.3 ppm, whereas that of 3 maintains its position. The A•T signals 4 and 5 move downfield, with that of 5 displaying the larger shift. These findings can be rationalized with the aid of shielding calculations [44]. The calculations predict shifts effects, similar to those observed, in going from the B- to the A-type double helix of this specific sequence, *viz.* downfield for the A•T resonances and upfield for the G•C signals, see the schematic drawing at the bottom of Fig. 1. The absolute values of the calculated imino-proton chemical shifts should not be regarded with too much confidence, but the calculated trends agree well with our experimental observations. It can be concluded from these imino-proton spectra that the helical structure of the hybrid differs from a normal B-type structure and that even A•T base pairs 5 and 5' have been affected to some extent by the RNA terminal tracts of the molecule. So far, nothing can be said concerning the conformational situation in the individual residues. This subject will be discussed in the following section.

Non-exchangeable proton resonances.

The non-exchangeable proton spectrum as measured in D_2O can be split up into two regions: (a) the 0-5 ppm region (Fig. 3) which contains the aromatic base-proton resonances, the anomeric H-1' resonances as well as the H-3', H-4', H-5' and H-5" signals, and (b) the 2 ppm region upfield from TMA reference which contains the thymine methyl signals and the H-2', H-2" resonances (Fig. 4). The two base resonances at 4.03 and 4.04 ppm (Fig. 3) belong to the adenine H-2(7) and H-2(5), respectively as follows from the obervation of NOEs at their positions upon irradiation of the A•T imino-proton signals 4 and 5 (see Fig. 1). Distinction can be made in the H-1' region (2.4-3.2 ppm) between the ribose H-1' resonances which appear as singlets or doublets (singlets in the present case because of the absence of observable $J_{1'2'}$ couplings) and the deoxyribose H-1' resonances which commonly appear as triplets or quartets because of couplings to H-2' and H-2". In order to connect the base resonances to

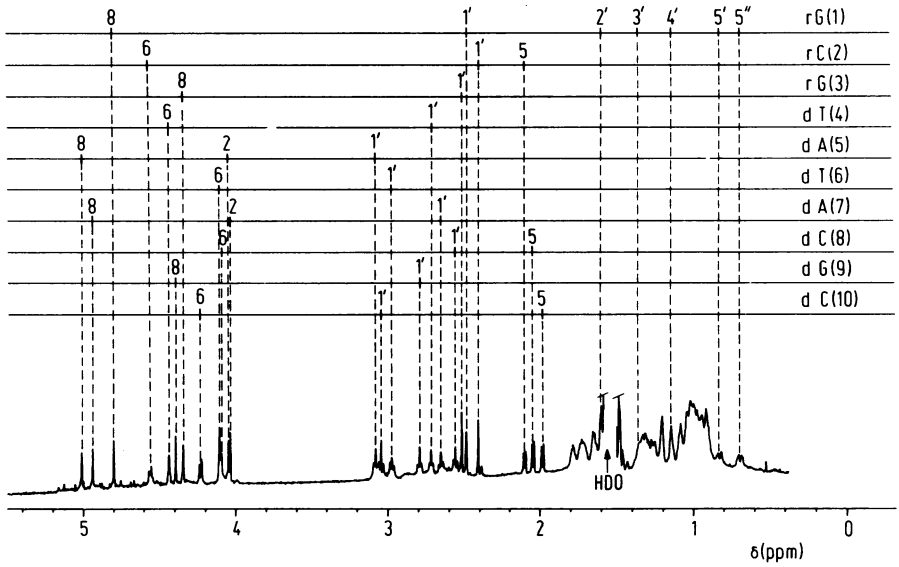


Figure 3. 0 - 5 ppm region (TMA reference) of the 500 MHz ^{1}H NMR spectrum of the non-exchangeable resonances of r(GCG)-d(TATACGC) as measured in D_2O at 32 °C. Sample conditions are as described in the Methods section. The assignment of the base and 1' protons and the rG(1) sugar protons is indicated by the staff notation.

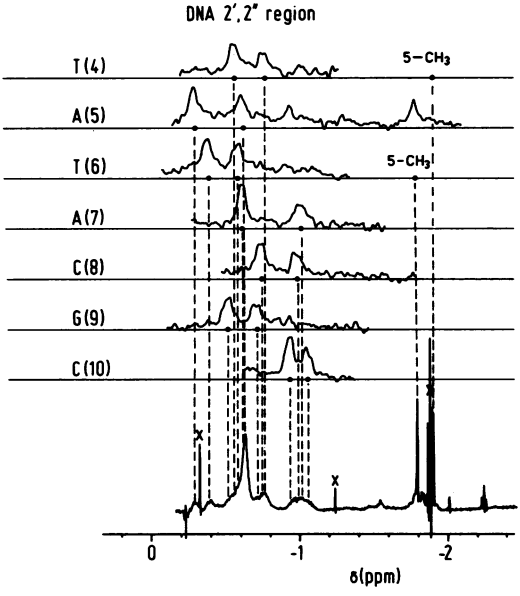
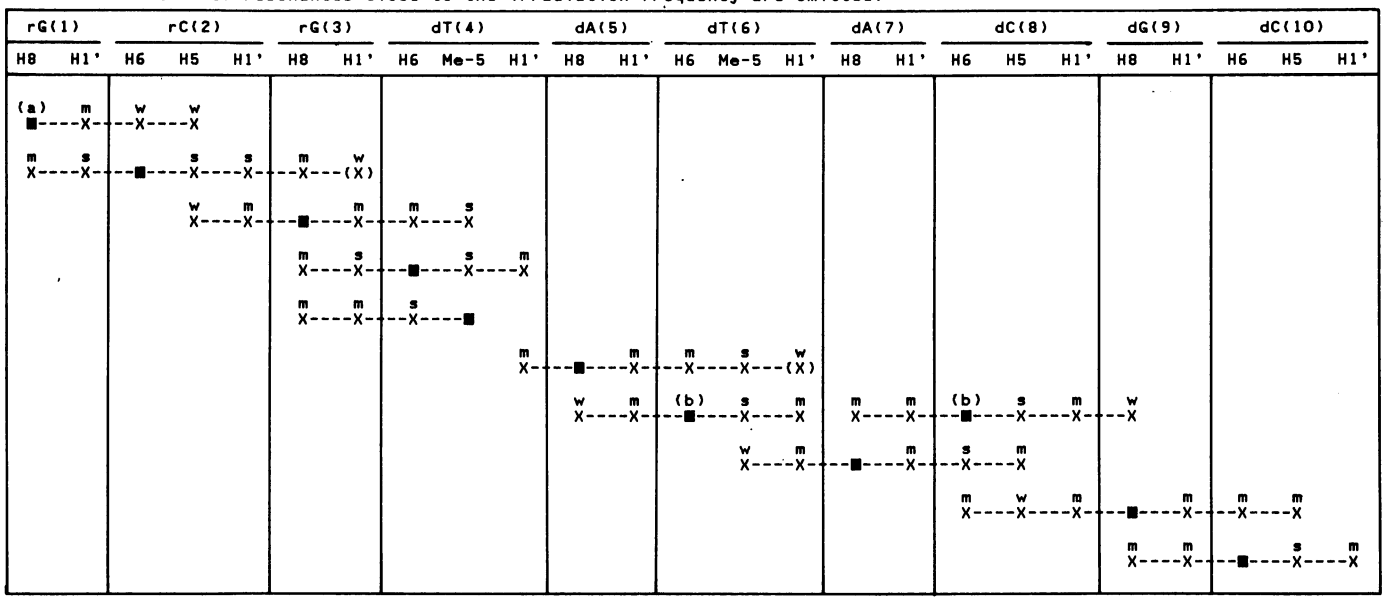


Figure 4. −2 - 0 ppm region (TMA reference) of the spectrum of Fig. 3 together with the NOE difference spectra obtained by irradiation of the deoxyribose H-1' signals.

Scheme 1. Proton nuclear magnetic resonance assignment of base and H1' resonances of the RNA-DNA hybrid decamer r(GCG)-d(TATACGC) by NOE-difference spectroscopy at 500 MHz. Black squares (■) indicate the irradiated resonances and crosses (X) indicate observed NOE's (s=strong, m=moderate, w=weak). Some non-nearest-neighbour NOE's due to co-irradiation of resonances close to the irradiation frequency are omitted.



(a) This experiment also yields NOE's on the 5'-terminal H5' and H5" resonances.
(b) Both resonances H6(6) and H6(8) display identical chemical shifts and are therefore irradiated simultaneously.

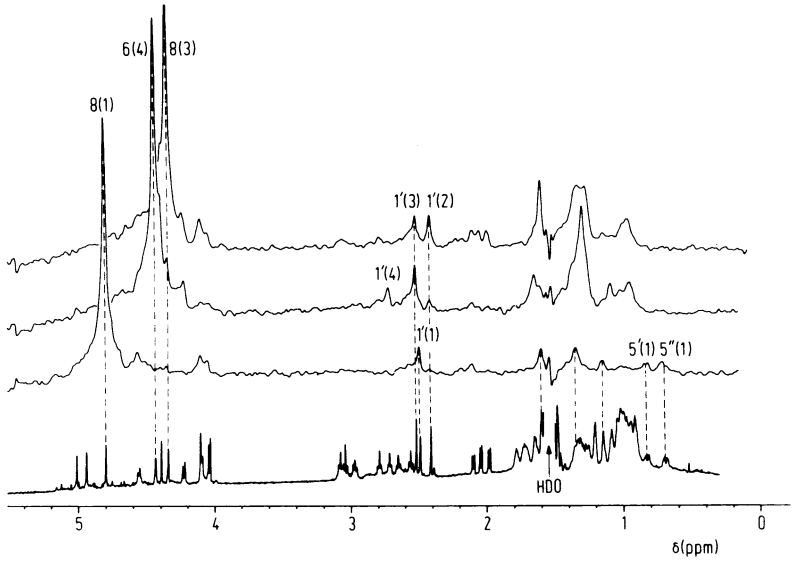


Figure 5. NOE difference spectra (TMA reference) for the irradiation of H-8(1), H-6(4) and H-8(3). The spectrum of Fig. 3 is reproduced at the bottom for comparison.

their associated H-1' signals, ten NOE difference spectra were recorded. A summary of these NOE difference spectra obtained by irradiation of the base resonances is given in Scheme 1. Typical spectra are presented in Fig. 5. This method proved to be highly informative, because under the conditions described above (Methods section) magnetization transfer not merely occurs within a given nucleotide residue (intraresidue NOE) but also to adjacent residues (interresidue NOE). Irradiation of a pyrimidine H-6 or a purine H-8 signal generally yields NOEs on the adjacent bases and on two H-1' signals. Scheme 1 shows that all ten experiments can be linked. The chemically established sequence then leads to a complete and unambiguous assignment of the base and H-1' resonances of the hybrid decamer. At this point several interesting observations concerning NOEs in the present molecule can be made.

In a regular helical structure (A- or B-type helix) an aromatic base can be considered as "sandwiched" between the associated sugar and the sugar of the 5'-neighbouring residue. Counting the residues of a given oligonucleotide in 5' → 3' direction, it is therefore likely that saturation of an H-6 or H-8 of the n-th residue yields NOEs to the 1' protons of residues n and (n-1). Similar phenomena were recently observed by Petersheim and Turner in an NOE study on several RNA oligomers [25]. A confirmation of this "rule" is provided by the

experiment in which H-8(1) was irradiated. This experiment (Fig. 5) constitutes an additional handle for assignment because it shows NOEs to the easily recognizable 5'-terminal protons H-5' and H-5", and to only one ribose H-1' signal. Every other experiment (scheme 1) yielded two medium to strong NOEs in the H-1' region. In addition, two weak H-1' NOEs (rG(3) and dT(6)) are detected (Scheme 1) that on first sight appear to violate the n → (n-1) rule. These are probably artifacts caused by tailing of the frequency distribution to H-6(4) and H-8(7), respectively.

The interbase NOEs exhibit a different pattern compared to the base-sugar NOEs. In general, H-8 → H-6 and H-6 → H-8 NOEs occur in both directions, n-1 and n+1. This is observed for dG-dC, dC-dG, dA-dC, dA-dT, rG-dT, rC-rG, and rG-rC sequences in the present hybrid. The only exception noted concerns both dT-dA sequences. Under the conditions employed no NOE is observed upon H-8(5) when H-6(4) is irradiated and *vice versa,* Scheme 1. Similarly, irradiation of H-8(7) has no effect upon H-6(6). The reverse experiment is not conclusive, unfortunately, because near isochronicity between H-6(6) and H-6(8) precludes irradiation of either resonance without spill-over to the other one.

From our earlier study on d(GGC*C*GGCC) [24] (C*=m^{5}C) it is noted that the above list of interbase NOEs can be extended by H-6 → H-6 in dC-dC. In contrast, no H-8 → H-8 NOE was detected for dG-dG [24]. Without doubt, a list of strong, medium, weak, and absent interbase NOEs in DNA and RNA fragments will be completed in the near future and such a list will prove to be an invaluable aid in carrying out spectral assignments in complex nucleic acid oligomers using 2D NOE spectroscopy.

In order to determine the positions of the H-2' and H-2" resonances a second series of NOE difference spectra was recorded under identical conditions as were the previous series. Saturation of the successive deoxyribose H-1' signals yielded all the H-2' and H-2" resonances separately (shown in Fig. 4), whereas the convential spectrum showed only several "humps" (bottom spectrum in Fig. 4) due to severe overlap. However, these NOE difference spectra are not sufficiently resolved to enable distinction between H-2' and H-2" signals on the basis of their respective line widths. Chemical shifts of the proton resonances assigned so far are given in Table 1.

The NOE difference spectrum obtained by saturation of H-1'(5) (Fig. 4) shows three NOEs instead of the two expected in the H-2',2" region. This phenomenon was traced to the near chemical shift equivalence of H-1'(5) and H-1'(10). The third unexpected NOE indeed corresponds to one of the NOEs obtained by irradiation of H-1'(10) (Fig. 4). Inspection of the H-2',2" regions

Table 1. Chemical shifts of base, H-1', H-2' and H-2" resonances of the hybrid decamer r(GCG)-d(TATACGC) recorded at 32 °C by 500 MHz ^{1}H NMR.

proton	rG(1)	rC(2)	rG(3)	dT(4)	dA(5)	dT(6)	dA(7)	dC(8)	dG(9)	dC(10)
	ppm(TMA)									
H-8	4.795	-	4.339	-	5.008	-	4.935	-	4.388	-
H-2	-	-	-	-	4.042	-	4.029	-	-	-
H-6	-	4.554	-	4.431	-	4.101	-	4.094	-	4.225
H-5/Me-5	-	2.094	-	-1.897	-	-1.789	-	2.039	-	1.977
H-1'	2.481	2.404	2.511	2.711	3.074	2.967	2.647	2.554	2.78	3.037
H-2' (a)	-1.60	-1.60	-1.60	-0.55	-0.28	-0.37	-0.60	-0.73	-0.51	-0.93
H-2" (a)	-	-	-	-0.75	-0.60	-0.57	-1.00	-0.97	-0.70	-1.05

a) H-2' and H-2" chemical shifts are determined from NOE difference spectra (Fig. 4) which do not display sufficient resolution to allow a stereochemical assignment.

of the previous series of NOE difference spectra (irradiations of the base resonances, Scheme 1) reveals four NOEs in most cases (not shown). Comparison of these spectra with the second series of NOE difference spectra (Fig. 4) leads to the conclusion that saturation of a purine H-8 or a pyrimidine H-6 of residue number n resulted in a superposition of NOEs on the H-2' and H-2" resonances of residues n and (n-1) in the DNA helix. In this way the second series of NOE difference spectra (Fig. 4) provided an additional confirmation of the base proton and H-1' assignments.

Sugar conformations.

Figure 6 shows the expanded H-1' region together with a computer simulation. The shapes of the H-1' signals are of great interest because they serve as a probe for the preferred type of sugar pucker and, indirectly, for the type of helix present. It is generally accepted that A-type helices are characterized by N-type (C3'-endo) and B-type helices by S-type (C2'-endo) sugar conformations. N-type and S-type furanose rings display different torsion angles about C1'-C2' and are therefore characterized by different spin-spin proton coupling constants along this bond. Moreover, the coupling pattern allows clear distinction between ribose and deoxyribose sugars. A ribose H-1' is coupled only to H-2' and shows up as a singlet or doublet with a separation ($J_{1'2'}$) varying between ≈ 1 Hz for pure N to ≈ 7.8 Hz for pure S [45-47]. The absence of ob-

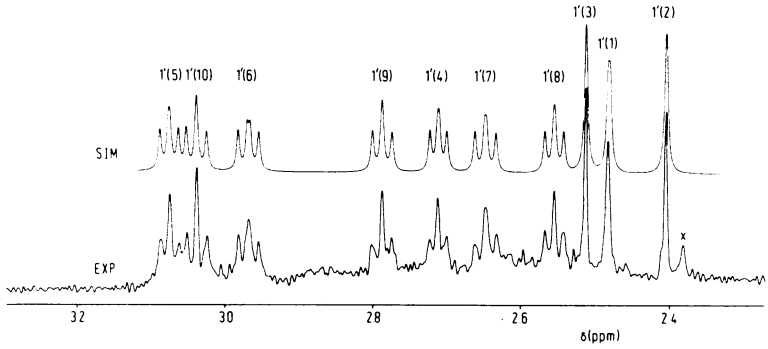


Figure 6. Expansion of the H-1' region of the spectrum shown in Fig. 3 (EXP) together with a computer simulation (SIM).

servable splitting ($J_{1'2'}$<1.5 Hz) of the three ribose H-1' resonances (Fig. 6) indicates a virtually pure N-type sugar pucker and it can be concluded, therefore, that the r(GCG-) sequence in the hybrid adopts the expected A-type helix. The middle singlet of the three, assigned to rG(1), appears somewhat broadened relative to the other two and this increased linewidth is probably caused by the existence of conformational freedom of the ribose at the 5'-terminus of the molecule. In contrast to the ribose 1' proton, a deoxyribose H-1' is coupled to both H-2' and H-2" and displays a triplet or quartet signal. Current best values of the total width of the H-1' signal $\Sigma J_{1'} = J_{1'2'} + J_{1'2"}$ predicted [47] are of the order of 9.2 - 11.1 Hz for pure N-type and 15.3 - 15.8 Hz for pure S-type deoxyribose rings. However, $\Sigma J_{1'}$ is less indicative for the conformational situation in the deoxyriboses than are the individual couplings $J_{1'2'}$ and $J_{1'2"}$.

The shapes of the seven deoxyribose H-1' resonances (Fig. 6) show similarity, but also subtle differences in their triplet-like patterns. An effort was made to reproduce these signals by computer simulation, in which special attention was paid to the relative height and width of the central peak of each triplet. These simulations indicated that in most cases the two different coupling constants could be determined separately within an accuracy of ≈ 0.3 Hz. The couplings thus obtained are listed in Table 2.

Strictly speaking, an unambiguous experimental assignment of $J_{1'2'}$ and $J_{1'2"}$ would require analysis of the H-2' and H-2" signals in terms of $J_{1'2'}$, $J_{2'3'}$ and $J_{1'2"}$, $J_{2"3'}$, respectively [47]. Unfortunately, the present spectra are unsuited for this purpose, partly because of serious overlap and partly because of the relatively short T_2 values of these methylene protons [Sanderson, M.R. and

Table 2. Pseudorotation analysis of the hybrid decamer r(GCG)-d(TATACGC). Values between brackets are constraints. Rms values are zero in all cases, *i.e.* the couplings are exactly repoduced.

residue	$J_{1'2'}$	$J_{1'2''}$	P_N	Φ_N	P_S	Φ_S	X_N
rG(1)	1.5	-					1.00
rC(2)	1.0	-					1.00
rG(3)	0.5	-					1.00
dT(4)	5.1	6.4	<0>	39.6	<153>	<38>	.59
dA(5)	5.6	6.7	<0>	35.0	<153>	<35>	.53
dT(6)	8.0	6.0	<0>	33.6	<153>	<37>	.27
dA(7)	7.9	6.4	6	38.2	<153>	32.2	.24
dC(8)	5.8	6.9	<9>	34.3	<153>	<35>	.52
dG(9)	7.0	6.3	<0>	35.0	<153>	<37>	.37
dC(10)	7.1	6.6	8	36.2	<153>	31.5	.33

Altona, C. unpublished results], which give rise to line broadening. Therefore, in each case both possible assignments of $J_{1'2'}$ and $J_{1'2''}$ were taken into consideration in the course of the pseudorotation analysis. It should be noted that these couplings are mutually correlated [46-48] through the phase angle P, the puckering amplitude Φ_m and the N/S ratio. In first instance, the pure O4'-endo ($P \approx 90°$) conformation, which appears indicated for some pyrimidine residues in d(CGCGAATTCGCG) from single-crystal X-ray studies [49], can be ruled out for the present hybrid molecule, because conformations of this type would produce a $\Sigma J_{1'}$ value of about 16 Hz or more [46], contrary to experiment.

Of course, knowledge of two couplings along one bond (C1'-C2') of the five-membered rings precludes a complete pseudorotation analysis. However, a search for possible solutions can be carried out under the reasonable assumption that the N and S geometries fall within the ranges indicated by ^{1}H NMR studies on many other DNA fragments [47], *viz.* P_N lies between 0° and 9°, P_S between 144° and 162°, and Φ_m of both conformers between 32° and 38°. A recent NMR study on the covalent RNA-DNA hybrid d(CG)r(CG)d(CG), which enabled the separate determination of four couplings along two bonds (C1'-C2' and C2'-C3') of the deoxyribose rings revealed that indeed no significant deviations from these assumed geometries occur [50].

As it turned out, only one of the two possible assignments of $J_{1'2'}$ and $J_{1'2''}$ led to a perfect solution of the pseudorotation problem (Table 2) within the constraints mentioned above (except T(4), *vide infra*), whereas reversal of the assignment did not produce meaningful results. Further analysis of the solution space by means of program PSEUROT [51] indicated that the acceptable range of solutions is fairly narrow. Most important is the finding that the calculated percentage of N-type conformer did not vary significantly (± 3 %).

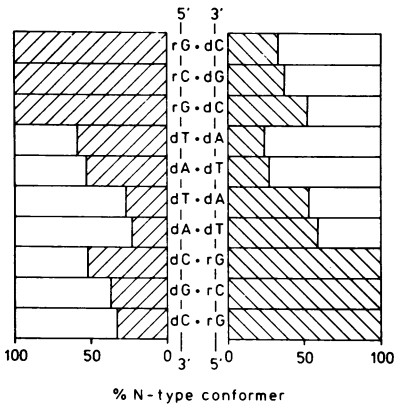


Figure 7. Schematic drawing of the occurrence of N-type conformer at 32 °C (hatched areas) in the various nucleotidyl units of the hybrid molecule r(GCG)-d(TATACGC).

The couplings of residue T(4), with an estimated accuracy of ≈ 0.2 Hz, deviate slightly from the normal pattern. The pseudorotation analysis indicates a relatively large amplitude of pucker ϕ_m (38-40°) for the sugar ring in this particular residue. One notes that residue T(4) plays a unique role in the molecule, being at the transition point between the RNA chain and the DNA chain. It is of interest to note that a comparable relatively large amplitude of pucker at an RNA-DNA junction is also observed at residue dC(5) in the covalent RNA-DNA hybrid d(CG)r(CG)d(CG) [Haasnoot, C.A.G. *et al.* , unpublished observations].

The pseudorotation analysis shows clearly that none of the sets of $J_{1'2'}$ and $J_{1'2''}$ couplings can be interpreted in terms of 100 % pure N- or S-type deoxyribose conformers. Instead, each individual residue appears to display a specific N to S ratio, depending upon its position along the DNA chain, Fig. 7 and Table 2. This surprising fact appears to rule out the possibility of the existence of a mixture of pure A- and B-type duplexes, in equilibrium through more or less complete opening and closing reactions, because if that were the case one would expect a constant N/S ratio for all residues at a given temperature. Instead, one finds for example (at 32 °C) 59 % and 53 % N-type conformer for the central residues dT(4) and dA(5), respectively, whereas the adjoining central residues dT(6) and dA(7) display only 27 % and 24 % N-type sugars, respectively. Perhaps these data can be explained by assuming the existence of three or more duplexes displaying mixed sugar puckers, but such an explanation appears somewhat forced. It seems more likely that the deoxyriboses along the intact duplex are able to flip between N and S conformers, the equilibrium constant depending upon local conformational demands. Seen in this light, Fig. 7 reveals several interesting features. Up till now it is usually taken for granted that a

DNA strand, hybridized with an RNA strand, completely reverts to the RNA structure. In our case, the residues dC(10), dG(9) and dC(8), which face the RNA residues rG(1), rC(2) and rG(3), do not show the expected N-type conformational purity that is associated with an A-type duplex. Residue dC(8) displays 52 % N-type sugar ring, which means a significantly greater N population than usually encountered in DNA duplexes at 32°C (0-20 %), but nevertheless remains far below 100 % purity. Residues dG(9) and dC(10) show even less N-type populations, 37 % and 33 % respectively, but this could be due to end effects. It is well to note that Zimmerman and Pheiffer [52], from a fiber X-ray diffraction study of the RNA•DNA hybrid poly(rA)•poly(dT) under conditions of high humidity, proposed a structural model for this duplex in solution which features a poly(dT) strand with S-type sugar combined with an N-type poly(rA) chain. Earlier CD studies [53] on the polymers {r(GU)•d(AC)} and {r(AC)•d(GT)} already indicated that the conformations of these hybrids in aqueous solution are not identical to the pure RNA conformation. ^{1}H NMR spectroscopy now for the first time permits a residue-for-residue insight into the conformational behaviour of the sugar rings along the chain. Before general rules can be formulated, a systematic NMR study of RNA•DNA hybrids is called for.

The central d(-TATA-) part of the molecule also deserves some comment. Fig. 7 shows a surprising break in behaviour between residues dA(6) and dT(7). It can be surmised that dT(4) and dA(5) are affected by the neighbouring RNA sequence and are forced to adopt more N-type deoxyribose conformation than would be the case in an analogous DNA sequence. This conformational transmission effect apparently dies out after two DNA residues, or anyway at a dA-dT sequence. Further work is clearly necessary before the roles, if any, of the RNA and DNA base sequences in determining the magnitude and the span of the conformational transmission effect can be delineated. At the moment it suffices to conclude that modifications are needed in the model proposed by Selsing *et al.* [14] for a covalent RNA-DNA junction, which model predicts a striking discontinuity in sugar conformation after the first DNA residue at the junction. The present work indicates a more gradual transition from A-RNA to B-DNA, and also the unexpected presence of conformational mobility in the B-DNA chain. The latter feature implies the presence of stacking interactions of "mixed" type (N-S) at the RNA-DNA junction as well as N-S and S-N further down along the DNA chain. The existence of such mixed stacks comes as no surprise, since similar conformational freedom is known to exist already on the dimer level, e.g. in rA-dA [54] and in dA-dA [55].

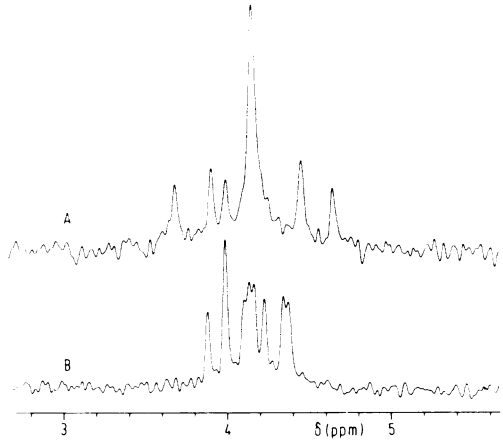


Figure 8. ^{31}P spectra (121.5 MHz, recorded at ambient temperature) of the hybrid decamer r(GCG)-d(TATACGC) (trace A) and the deoxy analogue d(GCGTATACGC) (trace B). The same samples were used as the ones described for the ^{1}H NMR experiments. Reference TMPB.

Phosphorus-31 spectra.

Figure 8 shows the ^{31}P spectra of the hybrid decamer and its deoxy analogue, recorded at ambient temperature (≈ 20 °C). The spectrum of the DNA compound (bottom trace) consists of eight well-resolved resonances within a chemical shift range of 0.5 ppm. One of these lines has a two-fold intensity and represents two isochronous resonances. The resonances in the hybrid spectrum (upper trace) extend over 1 ppm and show larger linewidths than do the DNA resonances. Six lines are resolved, which implies that four resonances have virtually identical chemical shifts. No attempts have been made so far to assign the individual resonances.

Phosphorus chemical shifts have been rationalized on the basis of the position of the conformational equilibrium between a *gauche/gauche* conformer around the phosphodiester bonds in the regular stacked state *versus* a blend of *gauche/trans* and *gauche/gauche* conformers in the destacked states [42,56]. However, solvent accessibility as well as other factors such as the base-base sequence and the presence or absence of the 2'OH group may also play a role in determining the ^{31}P shift [42,57]. Whatever the case may be, the spectra displayed (Fig. 8) indicate considerable differences in structure; the hybrid appears to feature greater variability around its phosphodiester bonds along the various units than does the pure DNA duplex. It is tempting to assume that this variability is related to the greater conformational variability detected for the

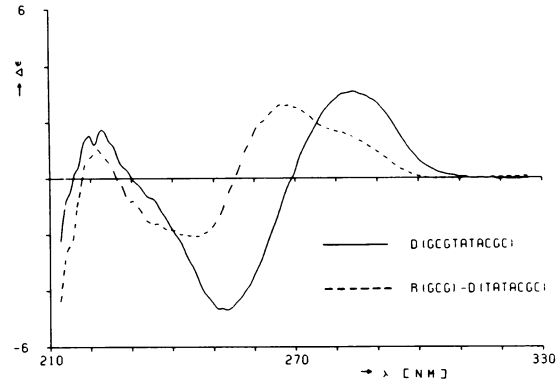


Figure 9. CD spectra recorded at 2 °C of d(GCGTATACGC) (solid line) and r(GCG)-d(TATACGC) (dashed line).

deoxyriboses in the hybrid compared to that occurring in the DNA.

CD spectra.

A comparison between the CD spectra of the hybrid molecule and of the DNA analogue, recorded at 2 °C, is given in Figure 9. The long-wavelength extremum observed in the deoxy compound (solid line) shifts strongly towards shorter wavelength in the hybrid (dashed line) which behaviour appears characteristic for RNA•DNA hybrid duplexes [53,58]. As was mentioned before, Gray and Ratliff [53] demonstrated that the CD spectra of two synthetic hybrids, poly{r(AC)•d(GT)} and poly{d(AC)•r(GU)}, differed from each other and also from the spectra of the corresponding DNA•DNA and RNA•RNA duplexes. The hybrid spectra, taken in aqueous solution, showed qualitative similarity to the RNA•RNA spectrum, but differed quantitatively. The CD spectrum of our covalent hybrid again shows some qualitative similarity to that of poly{d(AC)•r(GU)}, but features a much weaker first positive band. A quantitative comparison between the CD of the DNA and of the hybrid appears out of question, because of the presence of mixed stacks, *vide supra.* Olsthoorn *et al.* [59] found large differences between the CD of S-S and S-N stacks in $(dA)_n$. Clearly, if various mixed stacks display different characteristics, depending on the sugar pucker and on the base-base sequence, the quantitative interpretation of the shape of CD spectra in terms of A-type or B-type duplex conformation remains at best an extremely hazardous procedure. The numerous required parameters are simply unknown. Nevertheless, it may be said that the observed hypsochromic shift of the first maximum in the hybrid compared to its position in the DNA•DNA duplex agrees qualitatively with the presence of more

N-type sugar in the former as compared to the latter. In this sense the CD spectrum of the hybrid agrees with the results of the NMR analysis.

CONCLUSION:

The influence of a small tract of RNA upon the conformational behaviour of a DNA duplex was investigated. Simple and straightforward one-dimensional NOE experiments allowed a complete assignment of all base proton and 1' proton resonances. The conformations of the ribose and deoxyribose rings then could be studied on a residue-per-residue basis from the individual $J_{1'2'}$ and $J_{1'2''}$ couplings. It was found that the RNA part attains 100 % A-type conformational purity but does not induce pure A-DNA type conformation into the remainder of the molecule, although distinct conformational transmission is detected on the first two residues at the RNA-DNA junction, dT(4) and dA(5). The facing residue dC(8) is also significantly affected. Taken together, there is no evidence for a sharp conformational discontinuity at the RNA-DNA junction as required by the model postulated by Selsing *et al.* [14].

On first sight the present results disagree with the single-crystal X-ray investigation of the title compound. Wang *et al.* [27] report an overall A-type structure for the hybrid. It should be remembered, however, that the crystals were obtained in the presence of 2-methyl-2,4-pentanediol, *i.e.* under conditions of lower water content than in the NMR study. The importance of the relative humidity on the conformational behaviour of poly{d(AT)•d(AT)} was recently stressed by Mahendrasingam *et al.* [60]. These authors observed the various known DNA fibre diffraction patterns (A, B, B', C, D) upon variation of water and sodium chloride content. An NMR investigation of the hybrid under dehydrating conditions might prove informative on this point.

Finally, we conclude that a better understanding of the structural behaviour of RNA-DNA covalent hybrids requires a detailed insight into local conformational effects. The flexibility of the sugar rings to adopt either N and S conformation may well play a role in the control of transcription and replication.

ACKNOWLEDGEMENTS:

This research was supported by the Netherlands Foundation of Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO).

NMR spectra were recorded at the Dutch National 500-200 MHz HF NMR facility at Nijmegen and on the 300 MHz spectrometer in the Department of Chemistry at Leiden. We wish to thank Ing. P.A.W. van Dael, Ing. W. Guijt and Mr. C. Erkelens for technical assistance. We are indebted to Marjan Pieters for her contribution to part of this work.

A preliminary account of part of the present work has appeared elsewhere [61]. Note that in Fig. 5 of [61] the assignment of base-pair signals 2 and 3 of the

hybrid decamer were inadvertently interchanged.

REFERENCES:

1. This paper is part 33 in the series "Nucleic Acid Constituents" from this laboratory. For part 32 see: Doornbos, J., Barascut, J.-L., Lazrek, H., Imbach, J.-L., van Westrenen, J., Visser, G.M., van Boom, J.H. and Altona, C. (1983) *Nucl. Acids. Res.* 11, 4583-4600
2. Abbreviations used:
NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; CD, circular dichroism; WEFT, water eliminated Fourier transform; TMA, tetramethylammonium chloride; DSS, 3-(trimethylsilyl)-1-propanesulphonate [Na^+]; TMPB, tetramethylphosphonium bromide; EDTA, ethylenediaminetetraacetic acid; FID, free induction decay.
3. Fowler, R.G., Degnen, G.E. and Cox, E.C. (1974) *Mol. Gen. Genet. 133,* 179-191.
4. Kornberg, A. (1980) *DNA Replication,* pp. 347-413. W.H. Freeman and Company, San Francisco.
5. Luder, A. and Mosig, G. (1982) *Proc. Natl. Acad. Sci. USA 79,* 1101-1105.
6. Brutlag, D., Schekman, R. and Kornberg, A. (1971) *Proc. Natl. Acad. Sci. USA 68,* 2826-2829.
7. Sugino, A., Hirose, S. and Okazaki, R. (1972) *Proc. Natl. Acad. Sci. USA 69,* 1863-1867.
8. Ogawa, T. and Okazaki, T. (1980) *Ann. Rev. Biochem. 49,* 421-457.
9. Milman, G., Langridge, R. and Chamberlin, M.J. (1967) *Proc. Natl. Acad. Sci. USA 57,* 1804-1810.
10. Arnott, S., Chandrasekaran, R. and Selsing, E. (1975) in *"Structure and Conformation of Nucleic Acids and Protein-Nucleic Acid Interactions",* Sundaralingam, M., Ed., pp. 577-596, University Park Press, Baltimore, MD.
11. Selsing, E., Wells, R.D., Early, T.A. and Kearns, D.R. (1978) *Nature 21,* 249-250.
12. Pardi, A., Martin, F.H. and Tinoco, I. (1981) *Biochemistry 20,* 3986-3996.
13. Selsing, E. and Wells, R.D. (1979) *J. Biol. Chem. 254,* 5410-5416.
14. Selsing, E., Wells, R.D., Alden, C.J. and Arnott, S. (1979) *J. Biol. Chem. 254,* 5417-5422.
15. Cheng, D.M., Kan, L.-S., Leutzinger, E.E., Jayaraman, K., Miller, P.S. and Ts'o, P.O.P. (1982) *Biochemistry 21,* 621-630.
16. Bell, R.A., Alkema, D., Coddington, J.M., Hader, P.A., Hughes, D.W. and Neilson, T. (1983) *Nucl. Acids Res. 11,* 1143-1149.
17. Jardetzky, O. and Roberts, G.C.K. (1981) *"NMR in Molecular Biology".* Academic Press, New York.
18. Kan, L.-S., Cheng, D.M., Jayaraman, K., Leutzinger, E.E., Miller, P.S. and Ts'o, P.O.P. (1982) *Biochemistry 21,* 6723-6732.
19. Feigon, J., Wright, J.M., Leupin, W., Denny, W.A., and Kearns, D.R. (1982) *J. Am. Chem. Soc. 104,* 5540-5541.
20. Tran-Dinh, S., Neumann, J.-M., Huynh-Dinh, T., Igolen, J. and Kan, S.K. (1982) *Org. Magn. Reson. 18,* 148-152.
21. Tran-Dinh, S., Neumann, J.-M., Huynh-Dinh, T., Genissel, B. Igolen, J. and Simonnot, G. (1982) *Eur. J. Biochem. 124,* 415-425.
22. Stone, M.P., Johnson, D.L. and Borer, P.N. (1981) *Biochemistry 20,* 3604-3610.
23. Mellema, J.-R., Haasnoot, C.A.G., van Boom, J.H. and Altona, C. (1981) *Biochim. Biophys. Acta 655,* 256-264.

24. Sanderson, M.R., Mellema, J.-R., van der Marel, G.A., Wille, G., van Boom, J.H. and Altona, C. (1983) *Nucl. Acids Res. 11,* 3333-3346.
25. Petersheim, M. and Turner, D.H. (1983) *Biochemistry 22,* 264-268.
26. Petersheim, M. and Turner, D.H. (1983) *Biochemistry 22,* 269-277.
27. Wang, A.J.-H., Fujii, S., van Boom, J.H., van der Marel, G.A., van Boeckel, C.A.A. and Rich, A. (1982) *Nature 299,* 601-604.
28. van Boeckel, C.A.A., van der Marel, G.A., Wille, G., Westerink, H., Wang, A.J.-H., Mellema, J.-R., Altona, C. and van Boom. J.H. (1981) *Recl. Trav. Chim. Pays-Bas 100,* 389-390.
29. de Rooy, J.F.M., Wille, G., van Deursen, P.H., Serdijn, J. and van Boom, J.H. (1979) *Recl. Trav. Chim. Pays-Bas 98,* 537-548.
30. van der Marel, G.A., van Boeckel, C.A.A., Wille, G. and van Boom, J.H. (1981) *Tetrahedron Lett. 22,* 3887-3890.
31. Redfield, A.G., Kunz, S.D. and Ralph, E.K. (1975) *J. Magn. Reson. 19,* 114-117.
32. Redfield, A.G., and Kunz, S.D. (1981) in *"NMR and Biochemistry"*, Opella, S.J. and Lu, P., Eds., pp. 225-239, Dekker, New York.
33. Wright, J.M., Feigon, J., Denny, W., Leupin, W. and Kearns, D.R. (1981) *J. Magn. Reson. 45,* 514-519.
34. Haasnoot, C.A.G. and Hilbers, C.W. (1983) *Biopolymers 22,* 1259-1266.
35. Hartel, A.J., Lankhorst, P.P. and Altona,C. (1982) *Eur. J. Biochem. 129,* 343-357.
36. Orbons, L.P.M., Haasnoot, C.A.G. and Hilbers, C.W. (1982) unpublished observations.
37. Haasnoot, C.A.G. (1983) *J. Magn. Reson. 52,* 153-158.
38. Morris, G.A. and Freeman, R. (1978) *J. Magn. Reson. 29,* 433-462.
39. Patt, S.L. and Sykes, B.D. (1972) *J. Chem. Phys. 56,* 3182-3184.
40. Benz, F.W., Feeney, J. and Roberts, G.C.K. (1972) *J. Magn. Reson. 8,* 114-121.
41. Roth, K., Kimbler, B.J. and Feeney, J. (1980) *J. Magn. Reson. 41,* 302-309.
42. Haasnoot, C.A.G. and Altona, C. (1979) *Nucl. Acids Res. 6,* 1135-1149.
43. van Boom, J.H. and de Rooy, J.F.M. (1977) *J. Chromatogr. 131,* 169-177; de Rooy, J.F.M. (1979) Thesis, Leiden.
44. Patel, D.J. and Tonelli, A.E. (1974) *Biopolymers 13,* 1943-1964.
45. Haasnoot, C.A.G., de Leeuw, F.A.A.M., de Leeuw, H.P.M. and Altona, C. (1981) *Org. Magn. Reson. 15,* 43-52.
46. de Leeuw, F.A.A.M. and Altona, C. (1982) *J.C.S. Perkin II,* 375-384.
47. Altona, C. (1982) *Recl. Trav. Chim. Pays-Bas 101,* 413-433.
48. Altona, C. and Sundaralingam, M. (1973) *J. Am. Chem. Soc. 95,* 2333-2344.
49. Fratini, A.V., Kopka, M.L., Drew, H.R. and Dickerson, R.E. (1982) *J. Biol. Chem. 257,* 14686-14707.
50. van Boom, J.H., Westerink, H., van der Marel, G.A. and Haasnoot, C.A.G. (1983) to be published.
51. de Leeuw, F.A.A.M. and Altona, C. (1983) *J. Comp. Chem.* in the press.
52. Zimmerman, S.B. and Pheiffer, B.H. (1981) *Proc. Natl. Acad. Sci. USA 78,* 78-82.
53. Gray, D.M. and Ratliff, R.L. (1975) *Biopolymers 14,* 487-498.
54. Olsthoorn, C.S.M., Doornbos, J., de Leeuw, H.P.M. and Altona, C. (1982) *Eur. J. Biochem. 125,* 367-382.
55. Olsthoorn, C.S.M., Bostelaar, L.J., van Boom, J.H. and Altona, C. (1980) *Eur. J. Biochem. 112,* 95-110.
56. Gorenstein, D.G. (1981) *Ann. Rev. Biophys. Bioeng. 10,* 355-386.

57. Altona, C. (1982) in *"Structural Molecular Biology"*, Davies, D.B., Saenger, W. and Danyluk, S.S., Eds., pp. 161-214 and references therein. NATO Advanced Study Institute Series. Series A, Life sciences; Vol.45. Plenum Press, New York.
58. Tunis, M.-J.B. and Hearst, J.E. (1968) *Biopolymers 6*, 1217-1223.
59. Olsthoorn, C.S.M., Bostelaar, L.J., de Rooy, J.F.M., van Boom, J.H. and Altona, C. (1981) *Eur. J. Biochem. 115*, 309-321.
60. Mahendrasingam, A., Rhodes, N.J., Goodwin, D.C., Nave, C., Pigram, W.J. and Fuller, W. (1983) *Nature 301*, 535-537.
61. van Boom, J.H., van der Marel, G.A., Westerink, H., van Boeckel, C.A.A., Mellema, J.-R., Altona, C., Hilbers, C.W., Haasnoot, C.A.G., de Bruin, S.H. and Berendsen, R.G. (1983) *Cold Spring Harbor Symp. Quant. Biol. 47*, 403-409.