Sanchez-Ruiz, J. M., Lopez-Lacomba, J. L., Cortijo, M., & Mateo, P. L. (1988a) *Biochemistry 27*, 1648-1652.
Sanchez-Ruiz, J. M., Lopez-Lacomba, J. L., Mateo, P. L., Vilanova, M., Serra, M. A., & Aviles, F. X. (1988b) *Eur. J. Biochem. 175*, 225-230.
Saraste, M., Pentilla, T., & Wikstrom, M. (1981) *Eur. J. Biochem. 115*, 261-268.
Sebald, W., Weiss, H., & Jackl, G. (1972) *Eur. J. Biochem. 30*, 413-417.
Tanaka, M., Haniu, M., Yasunobu, K. T., Yu, C., Yu, L., Wei, Y., & King, T. E. (1979) *J. Biol. Chem. 254*, 3879-3885.
Wikstrom, M., & Sigel, E. (1979) in *Membrane Biochemistry* (Carafoli, E., & Semenza, G., Eds.) pp 82-91, Springer-Verlag, New York.
Yeates, T. O., Komiya, H., Rees, D. C., Allen, J. P., & Feher, G. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6438-6444.
Zhang, Y., Georgevich, G., & Capaldi, R. A. (1984) *Biochemistry 23*, 5616-5621.

Hairpin Structures in DNA Containing Arabinofuranosylcytosine. A Combination of Nuclear Magnetic Resonance and Molecular Dynamics†

Jane M. L. Pieters,‡ Erik de Vroom,‡,§ Gijs A. van der Marel,‡ Jacques H. van Boom,‡ Thea M. G. Koning,∥ Robert Kaptein,∥ and Cornelis Altona*,‡

Gorlaeus Laboratories, University of Leiden, P.O. Box 9502, 2300 RA Leiden, The Netherlands, and Department of Organic Chemistry, University of Utrecht, Padualaan 8, 3584 CH Utrecht, The Netherlands

Received May 3, 1989; Revised Manuscript Received August 21, 1989

ABSTRACT: Nuclear magnetic resonance (NMR) and model-building studies were carried out on the hairpin form of the octamer d(CG^{a}CTAGCG) (^{a}C = arabinofuranosylcytosine), referred to as the TA compound. The nonexchangeable protons of the TA compound were assigned by means of nuclear Overhauser effect spectroscopy (NOESY) and correlated spectroscopy (COSY). From a detailed analysis of the coupling data and of the NOESY spectra the following conclusions are reached: (i) The hairpin consists of a stem of three Watson-Crick type base pairs, and the two remaining residues, T(4) and dA(5), participate in a loop. (ii) All sugar rings show conformational flexibility although a strong preference for the S-type (C2′-endo) conformer is observed. (iii) The thymine does not stack upon the 3′ side of the stem as expected, but swings into the minor groove. (This folding principle of the loop involves an unusual α^t conformer in residue T(4).) (iv) At the 5′-3′ loop-stem junction a stacking discontinuity occurs as a consequence of a sharp turn in that part of the backbone, caused by the unusual β^+ and γ^t torsion angles in residue dG(6). (v) The A base slides over the 5′ side of the stem to stack upon the ^{a}C(3) residue at the 3′ side of the stem in an antiparallel fashion. On the basis of J couplings and a set of approximate proton-proton distances from NOE cross peaks, a model for the hairpin was constructed. This model was then refined by using an iterative relaxation matrix approach (IRMA) in combination with restrained molecular dynamics calculations. The resulting final model satisfactorily explains all the distance constraints.

Hairpin structures of RNA and DNA play an important role in biological systems (Sinden & Pettijohn, 1984; Weaver & DePamphilis, 1984; Sheflin & Kowalski, 1985). Particularly, hairpin structures in the anticodon regions of tRNAs and in the control mechanisms of gene expression (Sinden & Pettijohn, 1984; Weaver & DePamphilis, 1984; Sheflin & Kowalski, 1985) are of biological significance. Most of our knowledge with regard to the structural and kinetic, as well as the thermodynamic, aspects of hairpin structures has been obtained by systematic investigations of small synthetic oligonucleotides both in the crystalline state and in solution.

Recently, our NMR[1] studies on the hairpin form of the mismatched octamer d(m^{5}CGm^{5}CGTGm^{5}CG) (m^{5}C = 5-methylcytidine) revealed that this hairpin consists of a loop of only two residues, which bridges the minor groove (Orbons et al., 1987a,b; Altona et al., 1988a,b). A combined electrophoretic and spectroscopic study reported by Xodo et al. (1988) also showed evidence for hairpin structures of DNA compounds in which the loop is limited by two residues. This discovery stands in contrast to earlier intimations in the literature (Haasnoot et al., 1984, 1986; Hilbers et al., 1985), where it was reported that in DNA a four- to five-membered loop leads to the most stable hairpin fragment. The miniloop reported by Orbons et al. (1987b) shows some salient features. The O5′-C5′-C4′-C3′ torsion angle γ of residue dG(6) at the 5′-3′ loop-stem junction takes up the trans (γ^t) rotamer instead of the usual gauche plus (γ^+) conformer. As a consequence, the attached phosphate swings into the minor groove and the phosphate-phosphate distance across the minor groove decreases from the normal value of ca. 18 to 13.3 Å, which

† This research was supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organisation for Scientific Research (NWO). This is part 62 of the series "Nucleic Acid Constituents" from this laboratory; for part 61 see van den Hoogen et al. (1989).

* To whom correspondence should be addressed.
‡ University of Leiden.
§ Present address: Department of Chemistry, University of Virginia, Charlottesville, VA 22901.
∥ University of Utrecht.

[1] Abbreviations: ^{a}C, arabinofuranosylcytosine; m^{5}C, 5-methylcytidine; DSS, sodium 4,4-dimethyl-4-silapentanesulfonate; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser effect spectroscopy; NOE, nuclear Overhauser effect; COSY, correlated spectroscopy; IRMA, iterative relaxation matrix approach; MD, molecular dynamics.

0006-2960/90/0429-788$02.50/0

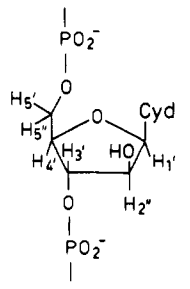


FIGURE 1: Structure of arabinofuranosylcytosine (^{a}C).

can be easily bridged by two nucleotides. The vertical base–base stacking is propagated into the loop from the 3′ side of the stem up to the second residue in the loop. A stacking discontinuity occurs between the 5′ side of the stem and the loop residue T(5).

In this paper we show that the octamer d(CG^{a}CTAGCG) adopts a hairpin structure under the experimental conditions chosen. Again, the stem of the hairpin is bridged by two residues. A detailed analysis of NMR coupling constant data and NOESY spectra reveals that the folding principle of the present hairpin differs significantly from that of the miniloop of the octamer d(m^{5}CGm^{5}CGTGm^{5}CG). A model of the hairpin structure is presented, which is extracted from the experimental structural parameters. The model was refined with the aid of restrained molecular dynamics calculations and molecular mechanics based on accurate proton–proton distances obtained from a relaxation matrix procedure to correct for the effect of spin diffusion on NOE intensities (Boelens et al., 1988, 1989).

Materials and Methods

The octamer d(CG^{a}CTAGCG) (^{a}C = arabinofuranosylcytosine, Figure 1), referred to as the TA compound, was synthesized via the hydroxybenzotriazole phosphotriester method (de Vroom et al., 1988). After purification, the compound was treated with a Dowex cation-exchange resin (Na^+ form) to yield the sodium salt.

NMR samples were lyophilized three times from D_2O (99.75%) and finally taken up in 0.4 mL of D_2O (99.95%). EDTA (0.1 mM) was added to neutralize paramagnetic contaminants, and tetramethylammonium chloride (Me_4NCl) was used as an internal reference; pH values were adjusted to 7.0–7.5 (meter reading). For most practical uses the Me_4NCl scale can be converted to the sodium 4,4-dimethyl-4-silapentanesulfonate (DSS) scale by addition of 3.18 ppm to the reported proton chemical shift. In order to observe exchangeable-proton resonances of the TA compound, two samples were prepared in a H_2O/D_2O (90/10) mixture; in this case DSS was used as reference compound. The nonexchangeable protons of the TA compound were studied at 2 and 5 mM oligonucleotide concentrations; the H_2O/D_2O (90/10) mixture samples contained 2 and 10 mM DNA. The NMR samples used in the nuclear Overhauser effect (NOE) experiments were carefully degassed and sealed under dry nitrogen.

NMR spectra were recorded on Bruker WM-300 and AM-500 spectrometers. The former instrument was interfaced with an ASPECT-2000 computer and the latter with an ASPECT-3000 computer. Suppression of the residual HDO peak was achieved by simple irradiation of the HDO resonance. Two-dimensional COSY and NOESY experiments were performed essentially as described elsewhere (Aue et al., 1976; Macura & Ernst, 1980; Wagner et al., 1984; Wider et al., 1984). In the case of the COSY experiment 512 2K (^{31}P-decoupled) spectra were collected. Before absolute-value Fourier transformation, the time-domain spectra were multiplied by a sine function and zero-filled to 2K (t_1) and 4K (t_2). In the NOESY experiments 450 2K spectra were collected. Before phase-sensitive Fourier transformation, both time-domain spectra were apodized with a sine function (phase shift $\pi/3$) and the t_1 domain was zero-filled twice. After Fourier transformation the ω_1 domain was corrected for the base line with the aid of a third-order polynomial. Mixing times were 75, 100, 150, 250, 325, 450, and 600 ms, respectively. All spectra from the mixing time series were processed with the same parameters. The observed cross peaks were integrated by simple summation of the intensity in an area around the center of the cross peak (Macura et al., 1981; Wagner & Wüthrich, 1979). Exchangeable-proton spectra were recorded at 500 MHz. The intensity of the H_2O signal was reduced by means of a time-shared long pulse in combination with a data-shift accumulation routine (Haasnoot & Hilbers, 1983; Roth et al., 1980).

^{31}P NMR spectra (broad band ^{1}H decoupled) were recorded on a Bruker WM-300 spectrometer at an operating frequency of 121.6 MHz.

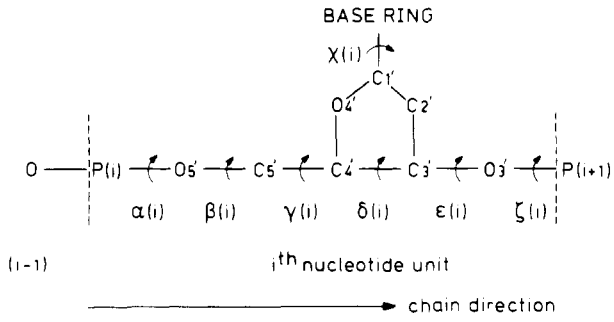
The hairpin models were generated with the aid of the program MacroModel (Copyright 1986 Columbia University, New York, NY) on an Evans and Sutherland PS 350 computer graphics system, interfaced to a VAX 11/750. The structures were energy-minimized with the all-atom version of the molecular mechanics program AMBER (Weiner et al., 1986).

Restrained molecular dynamics (MD) calculations on the present hairpin were performed with the GROMOS program as described before (Kaptein et al., 1985; van Gunsteren et al., 1985). To the potential energy function used in the conventional MD calculations was added the distance constraint function:

$$V_{dc} = \begin{cases} {}^1/_2K_{dc}(r_{ij} - r_{ij}^{l})^2 & 0 \le r_{ij} \le r_{ij}^{l} \\ 0 & r_{ij}^{l} \le r_{ij} \le r_{ij}^{u} \\ {}^1/_2K_{dc}(r_{ij} - r_{ij}^{u})^2 & r_{ij}^{u} \le r_{ij} \end{cases} \quad (1)$$

in which r_{ij} is the distance between the atoms i and j in the actual model and r_{ij}^{l} and r_{ij}^{u} are the lower and upper bound distance constraints between the atoms i and j, respectively. These upper and lower bound distance constraints were extracted from the series of 2D NOE spectra with the aid of an iterative relaxation matrix approach (IRMA) (Boelens et al., 1988, 1989), in which spin diffusion is fully taken into account. Each MD calculation started with 100 cycles steepest descent restrained-energy minimization on the model structure. The restrained MD calculations were performed under the following conditions: (i) the charges on the phosphates were neutralized; (ii) a united atom approach was used; (iii) solvent molecules were omitted; (iv) in order to keep the Watson–Crick type base pairs intact, a restraint was added on the hydrogen bonds (2.5 kJ·mol^{-1}·Å^{-2}). Further computational details are as follows: a dielectric permittivity, $\epsilon = 1$, was used; the nonbonded cutoff radius was set at 8 Å; the long-range cutoff radius was set at 12 Å; updates were carried out every 0.02 ps; a time step of 0.002 ps was used; the bond lengths were kept rigid during the simulation. Initial velocities were taken from a Maxwellian distribution at 300 K, with the overall translation and rotation of the molecule removed. During the MD runs, this temperature was maintained through the coupling with a thermal bath (Berendsen et al., 1984) at 300 K with a time constant of 0.01 ps in the equilibration period (0–5 ps) and 0.1 ps later on. The distance constraint

Chart I: Conformational Notation of the Nucleic Acid Backbone Torsion Angles



force constant, K_{dc}, was set to 40 kJ·mol^{-1}·Å^{-2}, which yields an energy contribution of $^1/_2kT$ at an excess distance of 0.25 Å at 300 K. The MD calculations covered a time span of 40 ps of which the last 30 ps were used for averaging. The calculation for the MD refinement and the IRMA procedure were carried out on a μVAX II. In the molecular mechanics calculations by means of the all-atom version of the AMBER program (Weiner et al., 1986), the fully anionic phosphate charges were partially shielded by the use of large "hydrated" Na^+ counterions (Seibel et al., 1985), placed at the bisectors of the PO_2^- groups. A similar distance constraint function (eq 1) was added to the potential energy function of the AMBER force field to carry out restrained molecular mechanics and MD calculations. Each molecular mechanics calculation consisted of a 200-cycle conjugate gradient (restrained) energy minimization, and a dielectric "constant" $\epsilon = r_{ij}/$Å was used.

The recommended IUPAC nomenclature for the torsion angles and the numbering of the residues is used throughout this work (IUPAC–IUB Nomenclature Commission, 1983, 1986) (Chart I).

Results and Discussion

The Exchangeable Protons. The low-field part of the ^{1}H NMR spectrum of the TA compound at various pH values is displayed in Figure 2. Four clearly resolved imino proton resonances are observed. Three of the protons resonate in a narrow spectral region, 13.4–13.1 ppm, which position corresponds to the position of imino protons that are hydrogen bonded in Watson–Crick type C·G base pairs. The remaining imino proton resonates at higher field, 11.7 ppm, the expected position of imino protons that are either shielded from the solvent or hydrogen bonded to C=O (Hilbers, 1979).

From dilution experiments (not shown) it is known that under these experimental conditions (2 mM DNA, 270 K, no added salt) the TA compound adopts a monomeric structure. Since single-helical species would not give rise to imino proton signals in the spectral region 13.4–13.1 ppm, this monomeric species can be identified unambiguously with a hairpin structure.

The length of the stem and the size of the loop of the hairpin are deduced from the following evidence. The three imino proton signals at low field are assigned to three Watson–Crick type C·G base pairs, in analogy with the imino proton resonances of the duplex structure at relatively high DNA concentration (10 mM) and salt conditions (100 mM NaCl) (not shown). It is concluded that these three C·G base pairs are fully formed at low temperature and form the stem of the hairpin. The loop of the hairpin comprises the two remaining residues.

Haasnoot et al. (1983a) have shown that one of the major properties of DNA hairpins—besides concentration independence of T_m—is given by the relatively fast exchange of

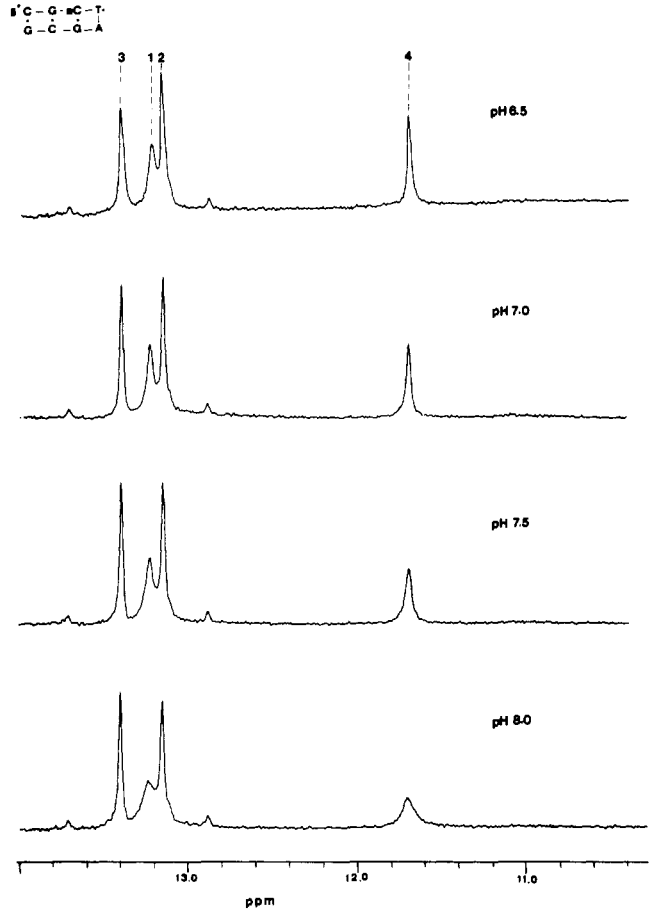


FIGURE 2: Imino proton spectrum (500 MHz) of the TA compound at various pH values; 2 mM, 270 K, no added salt.

the imino protons, which are located in the loop of the hairpin. In this case the exchange rate is confined to the exchange with the solvent protons, which means that the line width of the imino proton resonances is determined by the exchange rate with the solvent and not by the lifetime of the molecular species itself (Hilbers, 1979). Furthermore, imino protons in a complete duplex structure usually exhibit a more or less opening-limited exchange with the solvent. Thus, in the case of a solvent-accessible imino proton the line width will be affected by a change in pH, whereas in a base-paired duplex the line widths are largely independent of the acidity of the solvent at low temperature.

This different behavior is observed for the stem and loop imino protons upon a change of the pH as is illustrated in Figure 2. It is clearly seen that when the pH is increased from 6.5 to 8.0, the line width of the imino proton resonances of the terminal base pair (1) and the H3 signal of T(4) broaden, whereas the line widths of the imino proton resonances in the interior of the stem, peaks 2 and 3, appear unaffected by a small change of acidity of the solvent. In the case of terminal base pairs pH-dependent broadening is observed frequently and ascribed to fraying (Hilbers, 1979; Patel & Hilbers, 1975; Hilbers & Patel, 1975). However, the effect of pH on the line width of the loop imino proton appears less than that observed for other hairpin structures (Orbons et al., 1987a; Haasnoot et al., 1983a), for which imino proton resonances are observed as broad humps at pH 8.0. Moreover, for T(4) the resonance position is shifted downfield (0.65 ppm), compared to the H3(5) proton in the hairpin of d(m^{5}CGm^{5}CGTGm^{5}CG) (Orbons et al., 1987a), which resonates at 11.05 ppm, and to the non-base-paired imino protons of the thymine bases in the loop of the hairpins studied by Haasnoot et al. (1983a) and

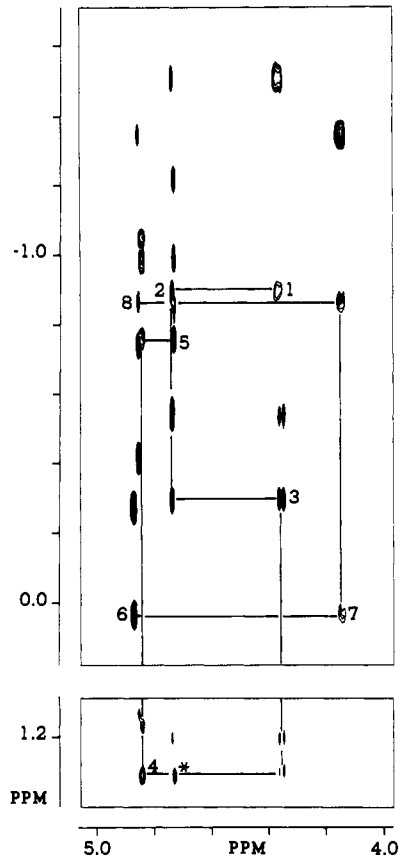


FIGURE 3: Part of the NOESY spectrum of the TA compound; 2 mM, 285 K, pH 7, τ_m = 600 ms. The region is shown in which H8/H6–H2′/H2″/Me5 connectivities are found. The abnormal NOE between $^{a}C(3)$ H2″ and dA(5) H8 is indicated with an asterisk.

Pramanik et al. (1988) (10–11 ppm). Note that the resonance position of a non-hydrogen-bonded imino proton of a thymine residue resonates at 11.2 ppm (Haasnoot et al., 1979). This indicates that the imino proton of the thymine base in the TA compound is located in a deshielding region of a neighboring base pair on the one hand and that this imino proton is more protected against fast exchange with solvent than imino protons located in loop regions known hitherto on the other hand.

Thus, it is concluded that the TA compound consists of a double-helical stem comprising three Watson–Crick type C·G base pairs and a loop of two residues.

Assignment of the Nonexchangeable Proton Resonances. The nonexchangeable proton resonances of the TA compound were assigned with the aid of well-established NOESY and COSY experiments (Scheek et al., 1984; Haasnoot et al., 1983b; Hare et al., 1983). The part of the NOESY spectrum (τ_m = 600 ms), in which one observes the connectivities between the base H8/H6 and the H2′/H2″ and the methyl protons, is shown in Figure 3. Note that the H2″ (deoxyribose numbering convention) of residue $^{a}C(3)$ resonates at lower field as a consequence of the extra OH group of the arabinose in residue 3. The discrimination between the H2′ and the H2″ signals follows from coupling constant analysis (Rinkel et al., 1986; Rinkel & Altona, 1987). This part of the 2D NOE spectrum contains NOE contacts which serve to deduce the sequential assignment of the ^{1}H NMR spectra of regularly stacked oligonucleotides.

It appears convenient to start the sequential assignment of the TA compound from the observed intranucleotide NOEs between H6(1) and H2′/H2″(1). The sequential assignment based upon inter- and intranucleotide NOEs between the base protons and the H2′/H2″ signals can be continued up to

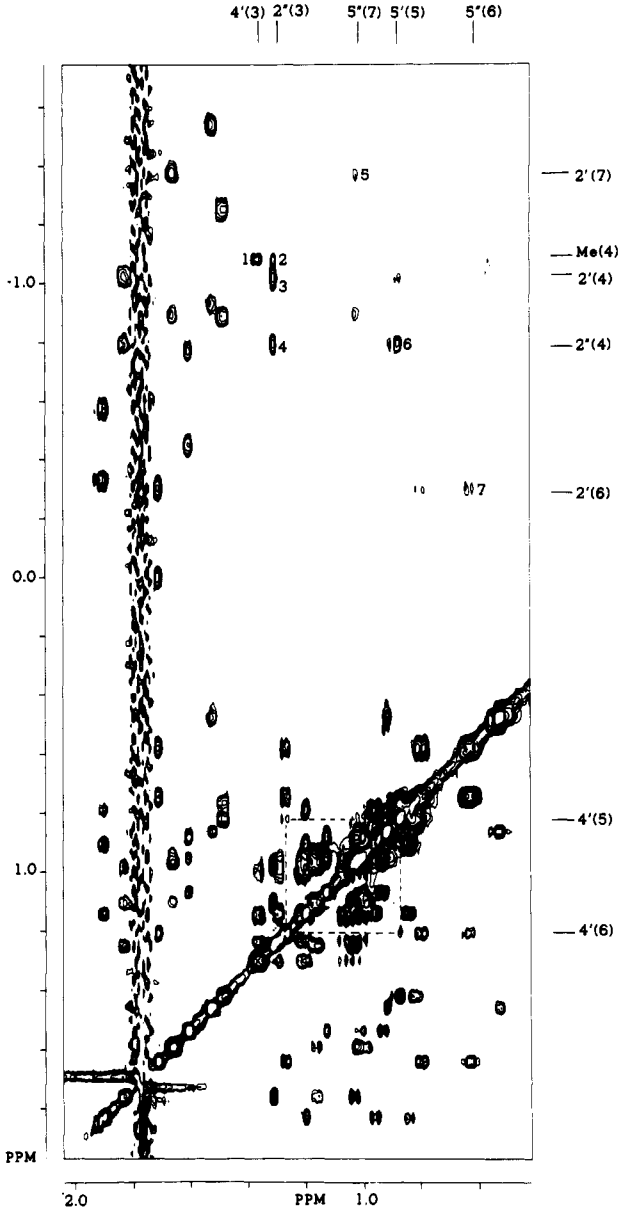


FIGURE 4: Part of the NOESY spectrum of the TA compound; 2 mM, 285 K, pH 7, τ_m = 600 ms.

residue dA(5) (solid lines in the figure). At this point no NOE correlation between the H2″ resonance of dA(5) and the H8 signal of dG(6) is seen. Moreover, neither the H2′(5)/H8(6) correlation nor the H1′(5)/H8(6) cross peak are observed. Starting at the 3′-terminal residue dG(8), the sequential assignment can be performed up to residue dG(6) in the same fashion. These results appear to suggest (but see below) that the vertical base–base stacking is propagated into the loop from the 3′ side of the stem up to dA(5), the second residue in the loop. In contrast, a stacking discontinuity is observed at the 5′ side of the stem between residues dA(5) and dG(6).

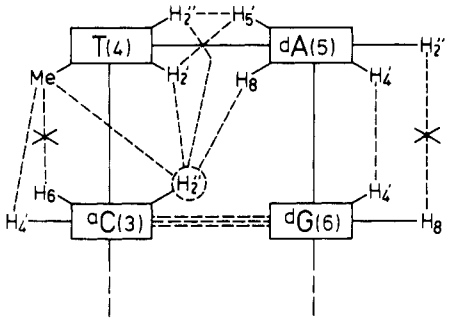
At first sight these results are quite similar to those observed in the hairpin structure of the octamer d-(m^5CGm^5CGTGm^5CG), referred to as the GT loop (Orbons et al., 1987a,b; Altona et al., 1988a,b). However, if one inspects the NOESY spectrum in more detail, some remarkable NOEs are observed, which are absent in the NOESY spectrum of the GT loop. First, in Figure 4, showing a part of the NOESY spectrum, one notes the abnormal interresidue NOE between the methyl resonance of residue T(4) and the H4′ signal of $^{a}C(3)$ and also surprising NOE connectivities between H2″ of residue $^{a}C(3)$ and CH_3, H2′, and H2″ resonances of

Table I: Chemical Shifts of the ^{1}H NMR Resonances of the TA Compound at 285 K[a]

residue	chemical shift								
	H6/H8	Me5/H5/H2	H1′	H2′	H2″	H3′	H4′	H5′	H5″
dC(1)	4.371	2.761	2.651	−1.509	−0.895	1.499	0.885	0.532	0.490
dG(2)	4.742		2.546	−0.543	−0.294	1.869	1.175	0.929	0.815
^{a}C(3)	4.359	1.891	2.668		1.285	1.259	1.341	1.174	1.012
T(4)	4.847	−1.051	3.363	−0.990	−0.755	1.798	1.282	1.131	1.005
dA(5)	4.736	4.910	2.608	−1.220	−0.858	1.460	0.897	0.866	0.793
dG(6)	4.877		3.181	−0.269	0.036	1.682	1.249	0.771	0.614
dC(7)	4.151	2.119	2.639	−1.347	−0.869	1.633	1.136	0.997	0.967
dG(8)	4.860		3.147	−0.416	−0.741	1.578	1.106	0.977	0.910

[a] Chemical shifts are given relative to Me_4NCl (ppm).

Chart II: Schematic Representation of the Major Present and Absent NOEs in the Loop–Stem Region of the TA Compound



T(4) (cross peaks 1–4 in Figure 4). These NOEs, which are absent in a regular right-handed stacking interaction, indicate that at the 3′–5′ loop–stem junction of the TA loop the regular stacking pattern appears quite distorted. This is confirmed by the absence of the H6(3)/Me5(4) NOE. The intense cross peak between H2″ of residue ^{a}C(3) and H8 of residue dA(5) (indicated by * in Figure 3) and the strong NOE observed between H2″(4) and H5′(5) (cross peak 6, Figure 4) suggest that residue dA(5) is involved in a kind of stacking interaction with the 3′ side of the stem, whereas T(4) more or less bulges out between dA(5) and ^{a}C(3). In Chart II the major present and absent NOEs in the loop–stem region are displayed.

Other atypical NOEs occur at the 5′ side of the stem. Strong intraresidue NOEs between H2′/H5″ (cross peak 7, Figure 4) and between H4′/H5′ of residue dG(6) combined with the relative weak NOE intensity of the H4′/H5″ cross peak imply that the torsion angle γ (O5′–C5′–C4′–C3′) of residue dG(6) is biased toward the γ^t rotamer (Figure 5). It should be noted that the H5′ and the H5″ resonances have been assigned according to their relative chemical shifts (Remin & Shugar, 1972). However, if the assignment of these protons is reversed, the torsion angle γ would prefer an uncommon γ^- conformer, which can be safely excluded since it would lead to an increase of the across-strand phosphate–phosphate distance. Finally, both the strong NOE connectivity between the H4′ resonances of residue dA(5) and dG(6) (broken lines in Figure 4) as well as the presence of the γ^t rotamer of residue dG(6) indicate the existence of a sharp turn in the backbone at the 5′–3′ loop–stem junction of the present hairpin structure, analogous to that found in the GT loop (Orbons et al., 1987a,b; Altona et al., 1988a,b).

In the stem of the hairpin an atypical intraresidue NOE between the H2′/H5″ of residue dC(7) is observed (cross peak 5, Figure 4). This indicates that the torsion angle γ(7) adopts the γ^t rotamer, at least during part of the time.

The cross peaks in the base-proton/H1′ region (not shown) reveal that all intraresidue base-proton/H1′ NOEs appear weak or even absent. This fact indicates that the bases of all residues adopt an anti conformation.

Table II: Chemical Shifts of the ^{31}P NMR Resonances of the TA Compound at 275 K[a]

residue	chemical shift	residue	chemical shift
dC(1)pdG(2)	−0.578	dA(5)pdG(6)	−0.763
dG(2)p^{a}C(3)	−0.991	dG(6)pdC(7)	−1.536
^{a}C(3)pT(4)	0.417	dC(7)pdG(8)	−0.315
T(4)pdA(5)	−1.766		

[a] Chemical shifts are relative to 85% phosphoric acid (ppm).

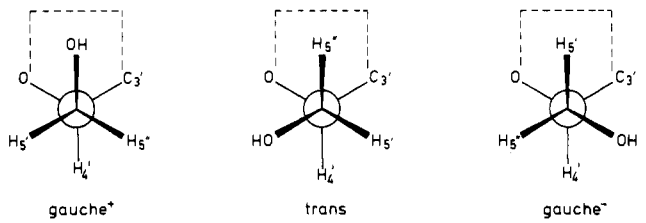


FIGURE 5: Newman projections of the rotamers about the C5′–C4′ bond (γ).

It is concluded from the observation of the relatively strong interresidue base(n)/H2″ (n − 1) NOEs and intraresidue base(n)/H2′(n) NOEs in the stem region of the hairpin that the stem forms a B-DNA type helix. The final assignment of the ^{1}H NMR spectrum of the TA compound is summarized in Table I.

^{31}P NMR Spectroscopy. ^{31}P NMR chemical shifts are frequently used to obtain information about the conformational preference of the phosphodiester torsion angles O3′–P (ζ) and P–O5′ (α) (Lerner et al., 1984). The chemical shift data of the phosphorus resonances of d($Cp_2Gp_3{}^aCp_4Tp_5Ap_6Gp_7Cp_8G$) are listed in Table II. It should be noted that the proton-decoupled ^{31}P NMR spectrum of d(CG^{a}CTAGCG) has been published elsewhere (Altona et al., 1988b). The ^{31}P NMR spectrum was completely assigned by means of selective heteronuclear decoupling experiments. Four ^{31}P signals (2, 3, 6, and 8) resonate in a narrow range that coincides with observations of other B-DNA type structures (Patel et al., 1982a,b). However, P(5) and P(7) are shifted upfield by ≈1.0 and ≈0.8 ppm, respectively, whereas P(4) resonates at much lower field (0.42 ppm). A downfield shift of a phosphorus signal is usually associated with conformational changes about the phosphodiester torsion angles α and/or ζ (Lerner et al., 1984); i.e., a gauche/gauche combination changes into a gauche/trans or trans/gauche combination. An upfield shift of a phosphorus resonance is usually ascribed to alteration in solvatation and/or hydrogen bonding of the phosphate group (Lerner et al., 1984).

The peculiar shifts of P(4) and P(5) could correspond with the altered backbone conformation of the region between bases 3 and 5 as deduced from NOEs discussed in the previous section. The origin of the upfield shift of P(7), which is located in the stem of the hairpin between dG(6) and dC(7), remains still unclear at this stage.

Table III: Coupling Constants of the TA Compound at 285 K

residue	coupling constants (Hz)						
	1′2′	1′2″	2′3′	2″3′	3′4′	4′5′	4′5″
dC(1)	8.4	6.1	6.3	2.5	2.7		
dG(2)	10.6	4.3	4.7	1.2	1.2[a]		
^{a}C(3)		2.7				2.3	10.7
T(4)	9.4	5.3	4.7	1.7	1.3[a]		
dA(5)	9.0	5.8	5.5	1.2	1.2[a]		
dG(6)	8.7	6.6	4.8	1.9	1.3[a]	2.3	9.5
dC(7)	8.7	6.0	7.8	3.3	4.0		
dG(8)	8.2	6.5	6.2	2.3	2.3		

[a] Calculated value.

Table IV: Pseudorotation Parameters, Torsion Angle δ (C5′–C4′–C3′–O3′), and Population S-Type Conformer (x_S) of the TA Compound (285 K)

residue	P_S (deg)	Φ_S	δ_S	x_S (%)	rms[a] (Hz)
dC(1)	151	31	136	80	0.08
dG(2)	154	42	143	96	0.20
^{a}C(3)				≥85	
T(4)	170	39	151	91	0.12
dA(5)	170	32	146	94	0.07
dG(6)	195	40	162	91	0.05
dC(7)[b]	122	33	119	80	0.53
dG(8)	163	29	140	82	0.05

[a] Root-mean-square deviation between observed and calculated coupling constants. [b] The relatively high rms value probably indicates an unusual flexibility of the sugar ring of dC(7).

Coupling Constant Analysis. Accurate ^{1}H–^{1}H coupling constants of the TA compound at 285 K have been derived from a 500-MHz computer-simulated ^{1}H NMR spectrum. Unfortunately, severe overlap in the H4′/H5′/H5″ spectral region prevents determination of the $J_{4'5'}$ and $J_{4'5''}$ coupling constants of most residues. Only the H4′ resonances of residues ^{a}C(3) and dG(6) were sufficiently resolved to obtain accurate $J_{4'5'}$ and $J_{4'5''}$ couplings. The proton–proton coupling data are collected in Table III.

The coupling constants of the sugar protons—H1′ to H4′—were interpreted, with the aid of an improved version (Donders, 1989; van Wijk et al., unpublished results) of the program PSEUROT (de Leeuw & Altona, 1983a,b; Haasnoot et al., 1981), in terms of P (phase angle of pseudorotation), Φ_m (amplitude), and population of the major S-type (C2′-endo) conformer; see Table IV. Note that, during the least-squares fit of the coupling constants, the minor N-type conformation was fixed at $P_N = 19°$ and $\Phi_m = 37°$.

Inspection of Table IV reveals that all sugar rings strongly prefer an S-type conformer ($x_S \geq 80\%$) but still enjoy some degree of conformational freedom. The expected relatively large amount of conformational flexibility for the 3′-terminal nucleotide dG(8), as observed in single-stranded as well as in double-helical species (Mellema et al., 1984; Orbons & Altona, 1986; Rinkel et al., 1987), appears rather more restricted in the present hairpin ($x_S = 82\%$). Both loop residues show strong preference for the S-type sugar ($x_S = 91$–94%). This stands in contrast with the conformational flexibility of the sugar ring of residue dG(4) in the GT loop (Orbons et al., 1987b), which displays only a moderate preference for the S-type sugar (70%).

Unfortunately, the P_S and Φ_m values of residue ^{a}C(3) could not be determined unambiguously. Extensive overlap in the region 1.1–0.5 ppm prevents accurate determination of the $J_{2''3'}$ and $J_{3'4'}$ coupling constants of the arabinonucleoside. However, both couplings must be small, because no cross peaks between H2″/H3′ and H3′/H4′ were observed in the COSY spectrum. The population S-type conformer can be approximated from the $J_{1'2''}$ coupling constant. The S-type sugar is characterized by a relatively small $J_{1'2''}$ coupling constant. The limiting $J_{1'2''}$ couplings are predicted as ≈7.1 Hz for pure N and ≈2.0–2.6 Hz for pure S conformers (Donders, 1989; van Wijk et al., unpublished results). Thus, it may be safely assumed that the population S-type sugar of the ^{a}C(3) residue exceeds 85%.

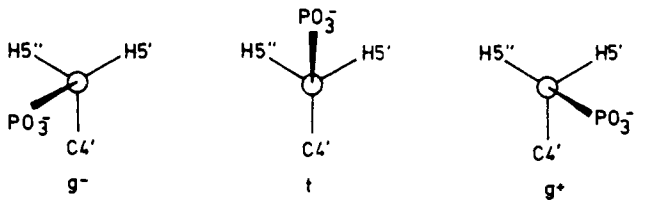


FIGURE 6: Newman projections of the rotamers about the O5′–C5′ bond (β).

The backbone torsion angle δ (C5′–C4′–C3′–O3′) is directly related to the pseudorotation parameters P and Φ_m (Altona, 1982) by eq 2. The values of δ_S of the TA compound, collected in Table IV, range from 140° to 147°, with two exceptions, i.e., dG(6) and dC(7). The δ value of residue dG(6), 162°, seems rather large compared to findings in regular B-DNA duplexes. However, the increase of the torsion angle δ(6) causes an appreciable sliding of the G(6) base with respect to the 5′ neighboring A base. This results in a better overlap of the A base with the 3′ side of the stem. The small δ value of residue dC(7) (119°), in the stem of the hairpin, agrees with earlier observations of the d(m^{5}C) residues in the core of the B-DNA duplex of d(m^{5}CG)$_3$ (Orbons & Altona, 1986).

$$\delta = 120.6 + 1.1\Phi_m \cos(P + 145.2) \quad (2)$$

The torsion angle γ (O5′–C5′–C4′–C3′) is monitored by the proton–proton coupling constants $J_{4'5'}$ and $J_{4'5''}$ (Figure 5). The population of the γ^+ rotamer can be calculated by means of an approximate sum rule from the sum of the $J_{4'5'}$ and $J_{4'5''}$ couplings, Σ (Altona, 1982):

$$P\gamma^+ = (13.3 - \Sigma)/9.7 \quad (3)$$

Unfortunately, only the sum values for residue ^{a}C(3) and dG(6) could be obtained. In a regular right-handed stacking pattern the torsion angle γ adopts a γ^+ rotamer (Altona, 1982) and a small value for the sum is observed. However, an anomalous behavior of this torsion angle is observed between the residues dA(5) and dG(6), the 5′–3′ loop–stem junction, and between dG(2) and ^{a}C(3). For both residues ^{a}C(3) and dG(6) a γ^t rotamer is preferred, 97% and 84%, respectively. This conformer at the 5′–3′ loop–stem junction implies a sharp turn in this part of the backbone. As a consequence P(6) swings into the minor groove and the cross-chain distance between P(4) and P(6) decreases by approximately 5 Å to ≈13 Å. This sharp turn in the backbone accords with the 5′–3′ loop–stem junction previously deduced for the GT loop (Orbons et al., 1987b).

At first sight, the pronounced preference of the torsion angle γ of residue ^{a}C(3) for the γ^t rotamer (97%) is not expected, because the ^{a}C(3) residue forms part of the double-helical stem of the hairpin. However, in the S-type arabinose ring the γ^+ rotamer appears disfavored because of steric repulsion between the phosphate group and the O2′ atom (Doornbos et al., 1983). This repulsion is absent in the γ^t conformer.

The torsion angle β (P–O5′–C5′–C4′) is monitored by the proton–proton coupling constants $J_{5'P}$ and $J_{5''P}$ (Figure 6). In a regular right-handed stack the torsion angle β adopts a trans rotamer (Altona, 1982), which implies small values for the $J_{5'P}$ and $J_{5''P}$ couplings. In the present hairpin the $J_{5''P}$ coupling of residue dG(6) is rather large, 11.5 Hz, whereas the $J_{5'P}$ coupling remains relatively small. This indicates that the preference for a β^t conformer of dG(6) is strongly diminished

and the rotamer population is biased toward the β^+ conformer. Note that the experimental value for the $J_{5''P}$ coupling does not imply conformational purity about β.

Unfortunately, the $J_{5'P}$ and $J_{5''P}$ coupling constants of residue ^{a}C(3) could not be determined experimentally. From model-building studies and structure refinement (next section), it is predicted that in the case of a γ^t rotamer and a "regular" stacking interaction between residues dG(2)–^{a}C(3) [observed from NOE connectivities dG(2)/^{a}C(3)] the sequential torsion angles $\zeta/\alpha/\beta$ adopt a $\zeta^t/\alpha^+/\beta^-$ combination, which means a large value for the $J_{5'P}$ coupling constant.

Model-Building Studies and Structure Refinement. It appears desirable to construct a model for the TA compound, which contains all major structural and conformational features as obtained from our NMR studies. The observed NMR results thus far can be summarized as follows: (i) the hairpin consists of a double-helical stem of three Watson–Crick type C·G base pairs, and the two remaining residues form the loop; (ii) the second residue in the loop, dA(5), stacks upon the 3′ side of the stem, whereas residue T(4) is not involved in any stacking interaction; (iii) the 5′–3′ loop–stem junction is formed by a sharp turn in the backbone by means of an unusual γ^t and β^+ rotamer in residue dG(6); (iv) torsion angle γ of residue ^{a}C(3) adopts the γ^t rotamer; (v) short interproton distances are present between H4′(3)/Me5(4), H2″(3)/Me5(4), H2″(3)/H2′(4), H2″(3)/H2″(4), H2″(3)/H8(5), and H2″-(4)/H5′(5); (vi) all sugar rings prefer to adopt the S-type conformation; (vii) all glycosyl torsion angles χ occur in the anti range. A preliminary model, which embodied most of the above features, was constructed with the aid of computer program MacroModel (Copyright 1986 Columbia University, New York, NY) and energy-minimized by means of the all-atom version of the molecular mechanics program AMBER (Weiner et al., 1986).

The resulting crude hairpin structure was used as a starting structural model in a refinement procedure based upon 2D NOE data in combination with restrained molecular dynamics (MD) calculations. Distance constraints were obtained from a set of 2D NOE spectra taken at various mixing times (75–600 ms). The NOE buildup curves of some resonances of selected protons in the loop region of the TA compound are displayed in Figure 7. In the case of indirect magnetization transfer via other protons the cross peak intensities are not directly related to the cross-relaxation rate of the two protons. To extract accurate proton–proton distances from the 2D NOE spectra, we have used an iterative relaxation matrix approach (IRMA), in which the effect of spin diffusion is corrected for (Boelens et al., 1988, 1989). The IRMA procedure is based upon back-transformation of a mixed NOE matrix (a combination of experimental NOEs and calculated ones from a starting structural model). The buildup curves of the H5/H6 cytosine cross peaks were used for calibration purposes (distance 2.45 Å). For this hairpin structure a set of 130 distance constraints (upper and lower distance limits) was obtained. In the first cycle of the structure refinement the upper and lower bound distance constraints were relaxed by 5% to allow for other sources of errors in the distance determination. In the second cycle the upper and lower bound distances were directly taken from IRMA. The final structure, obtained after two cycles, hardly differs from the structure, obtained after the first cycle [atomic root-mean-square (rms) difference 0.4 Å]; thus rapid convergence is obtained. Moreover, the restrained energy term becomes 241 kJ/mol, corresponding to an average amplitude of the violations of 0.3 Å. Clearly, the final structure calculated with the GROMOS force field fulfils

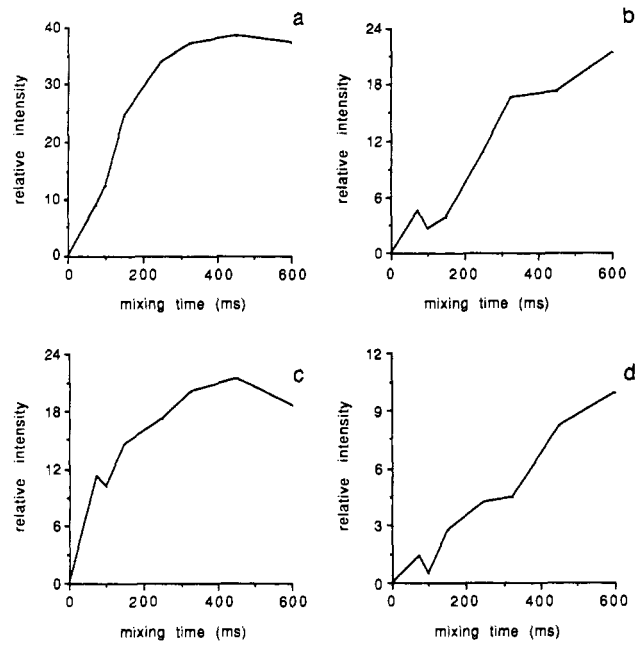


FIGURE 7: NOE buildup curves for selected cross peaks. The different curves correspond to the contacts H2″(3)/H2′(4), 2.1–2.4 Å (a); H2″(3)/H2″(4), 2.6–3.0 Å (b); H2″(4)/H5′(5), 2.3–2.5 Å (c); and H2″(3)/H8(5), 3.2–3.5 Å (d); compare Chart II. Note that these NOEs are incompatible with a duplex model.

the NOE constraints in a satisfactory way.

Our earlier model-building studies on the GT loop were based upon NMR data, molecular mechanics, and MD by means of the AMBER program. For the purposes of comparison, therefore, the final structure of the TA compound was energy-minimized by means of the all-atom version of the molecular mechanics program AMBER (Weiner et al., 1986), with and without the application of distance constraints. First, the results of the GROMOS (united atom) and AMBER calculations were compared, revealing an atomic rms difference of about 0.8 Å. The spread in the conformational space for the two AMBER structures appears only 0.34 Å, and moreover, the restrained-energy term drops to only 113 kJ/mol. In the second place, an unrestrained molecular dynamics calculation with the aid of the AMBER force field shows that the peculiar folding of the hairpin is retained during an 80-ps MD run. For these reasons we prefer to discuss the AMBER (restrained) energy-minimized structure as our final model. All upper and lower bound distance constraints obtained for the TA loop are listed in Table V. In our final model 85% of the constrained distances show distance violations $\leq$0.2 Å. Clearly, it can be concluded that our final (AMBER) model fulfils the NOE constraints in a satisfactory way.

The preferred hairpin structure for the TA compound, using the AMBER force field, is shown in Figure 8. The backbone torsion angles, sugar-ring pseudorotation parameters, and glycosyl torsion angles are listed in Table VI. From Figure 8 it is clearly seen that this hairpin model deviates substantially from hairpin models obtained hitherto (Orbons et al., 1987a; Haasnoot et al., 1983a; Hare & Reid, 1986). All hairpin loops studied until now show a retention of the vertical base–base stacking in the loop region. In contrast, in the TA compound the T base dips into the minor groove and stacks neither upon the ^{a}C(3) residue nor upon dA(5). Besides, the A base is involved in a stacking interaction at the 3′ side of the stem. Inspection of Table VI reveals some interesting features: (i) The double-helical stem of the model structure shows atypical torsion angles in the backbone between dG(2)–^{a}C(3) and dG(6)–dC(7). The torsion angle γ in both cytosine residues

Table V: Distances from NOE Data for the TA Compound As Derived from the IRMA Procedure[a]

	intranucleotide sugar–sugar distance (H8/H6 to sugar) (Å)						
residue	base/H1′	base/H2′	base/H2″	base/H3′	base/H5′	base/H5″	base/H5
dC(1)	3.2–3.4 (0.2)	2.4–2.5	3.6–3.8 (0.1)	3.6–5.0			2.4–2.6
dG(2)	4.0–4.7 (0.1)	2.1–2.2	2.5–2.8	4.1–5.0			
^{a}C(3)	2.8–3.1 (0.5)						2.4–2.7
T(4)	3.9–4.0 (0.1)	2.5–2.6 (0.2)	2.7–3.3 (0.3)	3.4–3.9 (0.1)	3.0–3.4		
dA(5)	3.7–3.9	3.0–3.2 (0.3)	3.1–3.4 (0.6)		3.0–3.1	3.4–4.1	
dG(6)	4.1–4.3 (0.3)	2.0–2.1	2.4–2.7 (0.5)	3.2–3.3 (0.6)	3.0–4.2 (0.1)	2.5–2.8	
dC(7)	3.4–3.9	2.2–2.3	3.0–3.4 (0.3)	2.6–3.2 (0.5)			2.4–2.5
dG(8)	3.7–4.3	2.2–2.3 (0.1)	3.9–4.3	3.4–3.5 (0.1)			

	internucleotide base–sugar and base–base distance ($n/n-1$) (Å)				
residue	base/H1′	base/H2′	base/H2″	base/H3′	H5/base
dC(1)					
dG(2)		2.9–3.1 (0.1)	3.5–3.6 (0.1)	3.7–4.1 (0.1)	
^{a}C(3)	3.3–3.7 (0.1)	3.2–3.7 (0.3)	2.5–2.6 (0.2)		3.4–3.6 (0.1)
T(4)	3.1–3.2		2.6–3.1 (0.1)		
dA(5)	5.0–5.4 (0.1)	3.1–3.4	2.5–2.8	3.0–3.4	
dG(6)					
dC(7)	2.9–3.2		2.8–3.1 (0.1)		3.9–4.2
dG(8)		3.0–3.4	3.0–3.2 (0.3)		

	intranucleotide sugar–sugar distance (Å)						
residue	H1′/H2′	H1′/H2″	H1′/H3′	H1′/H4′	H2′/H3′	H2″/H3′	H3′/H4′
dC(1)	2.9–3.2	2.2–2.4		3.0–3.4	2.3–2.5	3.0–3.2 (0.1)	2.9–3.3
dG(2)	2.7–2.8 (0.2)	2.3–2.4 (0.1)		2.7–2.9	2.2–2.3	2.2–2.6 (0.1)	2.9–3.0
^{a}C(3)		2.3–2.4		2.8–2.9 (0.1)			
T(4)	2.7–2.9 (0.1)	2.3–2.5		3.0–3.3	2.1–2.3	2.5–2.6	2.7–2.9
dA(5)	2.6–2.7 (0.3)	2.4–2.5	2.9–3.0 (0.6)	2.6–2.7 (0.3)	2.2–2.3	2.6–2.7	
dG(6)	2.5–2.7 (0.2)	2.3–2.4	2.8–2.9 (0.7)	3.0–3.2 (0.2)	2.2–2.4	2.4–2.6	2.6–2.8
dC(7)	2.8–3.0	2.4–2.5		2.6–2.7 (0.4)	2.2–2.4	2.5–2.8	3.0–3.4 (0.1)
dG(8)	2.8–2.9	2.3–2.4	3.6–3.8 (0.1)	3.3–3.4 (0.3)	2.4–2.6 (0.1)	2.6–2.9	2.8–3.0

	intranucleotide sugar–sugar distance (Å)					
residue	H2′/H5″	H2″/H5″	H3′/H5′	H3′/H5″	H4′/H5′	H4′/H5″
dC(1)						
dG(2)			2.9–3.0 (0.5)	2.5–2.8	2.5–2.8	2.3–2.4
^{a}C(3)						
T(4)			2.9–3.4 (0.2)	2.7–2.8 (0.1)	2.4–2.5	2.2–2.3 (0.1)
dA(5)				2.6–2.7		
dG(6)	2.5–2.7 (0.2)		2.4–2.6 (0.1)	2.5–2.8	2.4–2.5	2.7–2.9 (0.1)
dC(7)	2.7–2.9 (0.2)	3.3–4.5	2.3–2.6 (0.1)	3.1–3.3 (0.1)		
dG(8)			3.3–4.4	2.6–2.8		2.4–2.6

protons	distance (Å)	protons	distance (Å)	protons	distance (Å)
H1′(3)–Me5(4)	2.7–3.7	H2″(3)–Me5(4)	3.5–4.5	H2′(4)–H5′(5)	2.8–3.3 (0.6)
H2″(3)–H2′(4)	2.1–2.4	H2″(3)–H8(5)	3.2–3.5	H2″(4)–H5′(5)	2.3–2.5 (0.1)
H2″(3)–H2″(4)	2.6–3.0 (0.4)	H4′(3)–Me5(4)	2.7–3.7	H4′(5)–H4′(6)	3.2–3.3 (0.1)

[a] See text. Violations ≥0.1 Å are indicated in parentheses. Bars indicate violations of the lower bound distances.

Table VI: Backbone Torsion Angles of the Energy-Minimized Hairpin Structure for the TA Compound, by Means of the AMBER Force Field and Distance Constraints

residue	α	β	γ	δ	ϵ	ζ	χ	P	Φ_m
dC(1)		177	61	126	207	293	−145	133	35
dG(2)	259	177	51	132	198	176	−87	138	46
^{a}C(3)	91	276	191	134	207	277	−153	149	38
T(4)	188	161	52	145	222	279	−132	161	38
dA(5)	302	193	47	148	280	119	−146	166	39
dG(6)	203	77	157	155	189	230	−72	184	39
dC(7)	79	220	236	139	190	276	−130	155	35
dG(8)	276	178	51	126	190		−115	131	38

adopts a γ^t rotamer. The preference for the γ^t rotamer of residue ^{a}C(3) has been discussed already (previous section). However, residue dC(7) is a regular deoxynucleotide and is expected to adopt the normal γ^+ conformer. The observed NOE [H2′(7)/H5″(7)] however clearly indicates a γ^t at this position in the backbone. The reason for this peculiar conformational preference of residue dC(7) still remains unclear at this stage. (ii) The thymine base swings into the minor groove by means of a rotation about the torsion angle $\alpha(4)$ from α^- to α^t. All the other torsion angles, $\epsilon(3)$ to $\delta(4)$, retain their regular conformational preference. This rotation possibly is responsible for the ^{31}P downfield shift of P(4) (vide supra).

Besides torsion angles it appears useful to introduce two more parameters to gain more insight into the structural details of the model. (i) The local helical twist angle Ω (Dickerson, 1989; Lavery & Sklenar, 1988) between successive bases in a strand is defined as the angle between the projection of the successive C1′–N vectors onto the plane perpendicular to the helix axis. A clockwise rotation of the lower vector corresponds to a positive value of Ω. This definition yields twist angles that are virtually identical with the t_g angles defined by Fratini et al. (1982). (ii) The second parameter θ represents the angle between the normals of the successive base planes in a strand. No further distinction is made between roll (θ_r) and tilt (θ_t) defined by Fratini et al. (1982). To calculate the helix axis and the base-plane normals, the programs AHELIX (written by

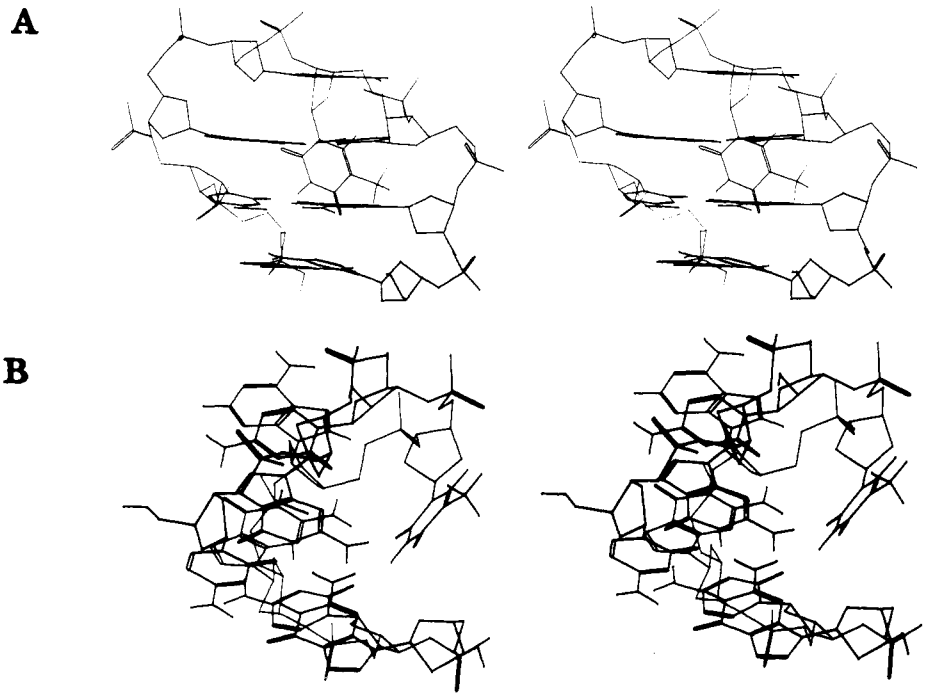


FIGURE 8: Stereo representation of the preferred hairpin structure obtained for the TA compound energy minimized with the aid of the AMBER force field: (A) view toward the major groove of the structure; (B) view into the minor groove of the structure.

Table VII: Local Helical Twist Angle Ω and Angles between Normals of the Successive Base Planes (θ) in the TA Compound[a]

stack	Ω (deg)	θ
dC(1)–dG(2)	38	9.3
dG(2)–^{a}C(3)	31	6.0
^{a}C(3)–dA(5)	117	7.4
dA(5)–dG(6)	44	7.7
dG(6)–dC(7)	30	8.5
dC(7)–dG(8)	37	3.2
^{a}C(3)–T(4)		73.6
T(4)–dA(5)		66.1

[a] Energy-minimized structure by means of the AMBER force field and distance constraints.

Table VIII: Hydrogen-Bond Angles and Distances in Hydrogen-Bonding Interactions in the TA Compound[a]

residues	hydrogen bond	angle (deg)	distance (Å)
dC(1)–dG(8)	N4–H4···O6	172	1.86
	N3···H1–N1	176	1.80
	O2···H2–N2	163	1.85
dG(2)–dC(7)	N4–H4···O6	163	1.85
	N3···H1–N1	170	1.84
	O2···H2–N2	166	1.84
^{a}C(3)–dG(6)	N4–H4···O6	174	1.82
	N3···H1–N1	172	1.82
	O2···H2–N2	165	1.81
T(4)–dG(2)	O4···H2–N2	140	1.98
T(4)–dG(6)	O2···H2–N2	138	1.93
T(4)–dC(7)	N3–H3···O2	152	1.94
dA(5)–^{a}C(3)	N7···H–O2′	167	1.75
	N2–H2···O2′	163	1.94

[a] Energy-minimized structure by means of the AMBER force field and distance constraints.

J. Rosenberg) and BROLL (written by R. E. Dickerson) were used.

The twist angles Ω in the stem region of the present hairpin range from 30° to 37° (Table VII). These values correspond closely to the range recorded for the X-ray structure of the dodecamer d(CGCGAATTCGCG) studied by Dickerson and Drew (1981). In the calculation of the Ω values in the loop–stem region the T base is left out of consideration, because the T base is not involved in any stacking interaction. The interaction between ^{a}C(3) and dA(5) shows an antiparallel stacking pattern, which corresponds to a large Ω value, 117°. The twist angle in the dA(5)–dG(6) step remains in the normal range of B-DNA, 44°. Table VII shows the changes between the base-plane normals of the successive bases in one strand. The calculated θ values in the stem region range from 3° to 9°, which again agrees with the θ values measured for the dodecamer d(CGCGAATTCGCG) (Dickerson & Drew, 1981). However, in the loop region the T(4) and A(5) bases occur more or less perpendicular to each other, $\theta = 66°$. The same feature is observed between ^{a}C(3) and T(4), $\theta = 74°$, whereas the θ values for the bases ^{a}C(3)–dA(5) and dA(5)–dG(6) remain similar to those observed in the stem region, $\theta \approx 8°$.

This completely different folding principle of the TA compound needs to be studied in more detail. In order to investigate the behavior of the thymine base, we analyzed the possibilities of hydrogen bonding of the thymine in the minor groove. A thymine has three positions for hydrogen bonding, viz., two potential hydrogen-bond acceptors, O2 and O4, and one hydrogen-bond donor H3. The stem of the hairpin consists of three Watson–Crick type C·G base pairs. In the case of C·G or G·C base pairs the hydrogen-bonding positions in the minor groove nearly overlap (Dickerson et al., 1987). They consist of two hydrogen-bond acceptors, cytosine O2 and guanine N3, and one hydrogen-bond donor, the NH_2 of guanine (Dickerson et al., 1987). In Table VIII the distances between hydrogen-bonding atoms together with the hydrogen-bond angles in the model are listed. One notes that, besides the nine regular hydrogen-bonding interactions in the stem (three Watson–Crick type C·G base pairs), five other unexpected H-bonds are present in the model. It appears that the thymine is rather strongly bonded in the minor groove by means of three hydrogen bonds. The amino groups of dG(2) and dG(6) serve as hydrogen-bond donors for the O4 and O2 of the thymine, respectively, whereas the H3 of T(4) is involved in a hydrogen-bonding interaction with O2 of residue dC(7) (Figure 9). The remaining (three-center) hydrogen bonds are formed between the 2′ OH group of ^{a}C(3) at the 3′ side

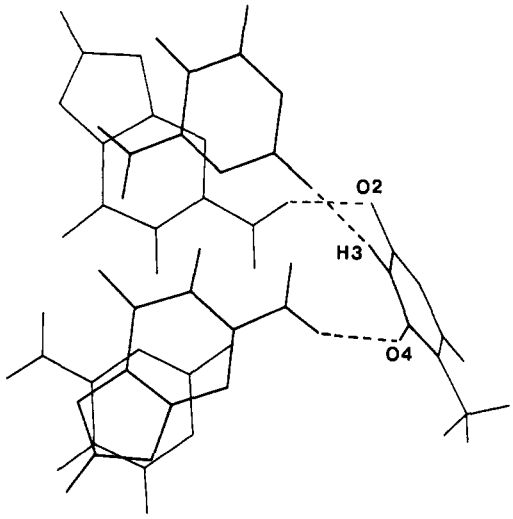


FIGURE 9: Hydrogen-bonding interactions in the minor groove of the thymine (energy-minimized structure by means of the AMBER force field).

of the stem and the adenine. The NH_2 group of the adenine serves as donor to O2′, and the N7 is the acceptor of a hydrogen bond from the hydroxyl group. The hydrogen bonds in the stem show the expected length and the hydrogen-bond angles correspond well to the results from the refinement of the X-ray data of the dodecamer d(CGCGAATTCGCG) (Orbons et al., 1987b). The hydrogen bonds that involve the bases in the loop differ only slightly. It should be noted that the GROMOS-minimized model differs in some details from the AMBER-minimized one, particularly with respect to the type and number of hydrogen-bonding interactions of the thymine in the minor groove. However, the general folding of this miniloop appears independent of the chosen force field.

The model building described above essentially rests upon information obtained from NMR spectroscopy, i.e., torsion angles and NOE data, in combination with restrained MD and energy minimization. It should be mentioned that at no stage was explicit use made of chemical shift considerations. In the chemical shift analysis (Table I) strong upfield shifts of the sugar protons of residue dA(5) were observed, particularly H2′ and H2″ protons, whereas H2(5) resonates at an atypical low-field position. In the model the adenine base stacks upon $^{a}C(3)$. This means that the sugar-ring protons of residue dA(5) are located inside the shielding region of dG(6), which explains the upfield shifts of the sugar protons, at least qualitatively. The geometrical position of H2(5) gives rise neither to shielding nor to deshielding. Thus, the resonance position should hardly differ from the chemical shift in the random-coil form, as is indeed observed. The base protons as well as the sugar protons of residue T(4), particularly H6 and H1′, show pronounced downfield shifts compared to the random-coil form. This feature can be qualitatively explained from the present model. The thymine occurs more or less perpendicular to the base pairs located in the stem. This means that the protons of residue T(4) are located inside the deshielding region of the base pairs $^{a}C(3)$·dG(6) and dG(2)·dC(7). We conclude that the chemical shift data support the present model.

Finally, a comparison of the TA compound and the GT loop now appears in order. Orbons et al. (1987b) have shown that in the GT loop the vertical base–base stacking is retained in the loop region in a more or less regular fashion, propagated from the 3′ side of the stem. At the 5′ side of the stem the loop is closed by a sharp turn in the backbone, which involves unusual torsion angles β^{+} and γ^{t} in residue dG(6). This results in a stacking discontinuity in the 5′–3′ loop–stem junction. At the 5′ side of the stem of the TA compound the loop is closed in a similar fashion as the GT loop; however, at the 3′ side the structures differ completely. The thymine swings into the minor groove, whereas the second residue in the loop, dA(5), stacks upon the 3′ side of the stem. This finding appears in contradiction with the suggestion, proposed by Haasnoot et al. (1983a, 1984), that propagation of the base stacking from the 3′ side of the stem into the loop contributes significantly to the hairpin stability. In the present model two relatively weak stacking interactions, i.e., $^{a}C(3)$–T(4) and T(4)–dA(5), have disappeared, which loss appears compensated for by a 1–3 stacking pattern and various hydrogen-bonding interactions. Thus, it seems likely that the particular folding of a hairpin structure depends upon the base sequence of the molecule. The nature of the bases that participate in the loop as well as those in the stem region adjacent to the loop appears of major importance.

CONCLUSIONS

The hairpin form of the octamer d(CG^{a}CTAGCG) consists of a stem of three Watson–Crick type C·G base pairs and a loop formed by the two remaining residues. Accurate analysis of the coupling data and the NOESY spectra in combination with restrained molecular dynamics calculations reveals that the folding principle of this miniloop differs substantially from that of the miniloop of the octamer d(m^{5}CGm^{5}CGTGm^{5}CG) studied previously by Orbons et al. (1987b). The stem of the TA loop shows regular B-characteristics, perhaps excepting the torsion angles γ of residues $^{a}C(3)$ and dC(7) that are assumed to adopt a γ^{t} rotamer. The sugar rings are not conformationally pure, although a strong preference for the S-type conformer is indicated from the coupling constant analyses. At the 5′ side of the stem the loop features a sharp turn in the backbone which involves unusual torsion angles β^{+} and γ^{t} in residue dG(6). This results in a stacking discontinuity at the 5′–3′ loop–stem junction. The stacking interaction at the 3′ side of the stem is not propagated into the loop as expected from earlier results, but the thymine dips into the minor groove instead. This structure is accomplished by rotation about the torsion angle α of residue T(4), which adopts the α^{t} conformer. Besides, the A base slides over the 5′ side of the stem to stack upon the cytosine at the 3′ side of the stem. The interesting features of this site of the molecule are monitored by atypical NOE connectivities between the loop residues and the 3′ side of the stem. Relatively large upfield shifts of H2′(5) and H2″(5) and large downfield shifts of the H6 and H1′ signals of residue T(4) appear to constitute important markers for the peculiar folding pattern of the hairpin structure reported in this paper.

ACKNOWLEDGMENTS

Spectra were recorded on the 500-MHz spectrometer in the Department of Chemistry at Utrecht and on the 300-MHz NMR facility in the Department of Chemistry at Leiden. We thank C. Erkelens and Dr. R. Boelens for technical assistance.

Registry No. d(CG^{a}CTAGCG), 124042-28-0.

REFERENCES

Altona, C. (1982) *Recl. Trav. Chim. Pays-Bas 101*, 413–433.

Altona, C., van Beuzekom, A. A., & Orbons, L. P. M. (1988a) in *Structure and Expression, Vol. 2: DNA and Its Drug Complexes* (Sarma, R. M., & Sarma, M. H., Eds.) pp 49–71, Adenine Press, Guilderland, New York.

Altona, C., van Beuzekom, A. A., Orbons, L. P. M., & Pieters, J. M. L. (1988b) in *Biological and Artificial Intelligence*

Systems (Clementi, E., & Chin, S., Eds.) pp 93–124, ESCOM Science Publishers, Leiden.

Aue, W. P., Bartholdi, E., Ernst, R. R., & Wüthrich, K. (1976) *J. Chem. Phys. 64*, 2229–2246.

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., & Hoek, J. R. (1984) *J. Chem. Phys. 81*, 3684–3690.

Boelens, R., Koning, T. M. G., & Kaptein, R. (1988) *J. Mol. Struct. 173*, 299–311.

Boelens, R., Koning, T. M. G., van der Marel, G. A., van Boom, J. H., & Kaptein, R. (1989) *J. Magn. Reson. 82*, 290–308.

de Leeuw, F. A. A. M., & Altona, C. (1982) *J. Chem. Soc., Perkin Trans. 2*, 375–384.

de Leeuw, F. A. A. M., & Altona, C. (1983a) *J. Comput. Chem. 4*, 428–437.

de Leeuw, F. A. A. M. & Altona, C. (1983b) *Quantum Chemistry Program Exchange*, No. 463.

de Vroom, E., Roelen, H. C. P. F., Saris, C. P., Budding, T. N. W., van der Marel, G. A., & van Boom, J. H. (1988) *Nucleic Acids Res. 16*, 2987–3003.

Dickerson, R. E. (1989) *EMBO J. 8*, 1–4.

Dickerson, R. E., & Drew, H. R. (1981) *J. Mol. Biol. 149*, 761–786.

Dickerson, R. E., Kopka, M. L., & Pjura, P. E. (1987) in *DNA–Ligand Interactions* (Guschlbauer, W., & Saenger, W., Eds.) pp 45–62, Plenum Publishing Corp., New York.

Donders, L. A. (1989) Ph.D. Thesis, University of Leiden.

Doornbos, J., Barascut, J.-L., Lezrek, H., Imbach, J.-L., Westrenen, J., Visser, G. M., van Boom, J. H., & Altona, C. (1983) *Nucleic Acids Res. 11*, 4583–4600.

Fratini, A. V., Kopka, M. L., Drew, H. R., & Dickerson, R. E. (1982) *J. Biol. Chem. 257*, 14686–14707.

Haasnoot, C. A. G., & Hilbers, C. W. (1983) *Biopolymers 22*, 1259–1266.

Haasnoot, C. A. G., den Hartog, J. H. J., de Rooy, J. F. H., van Boom, J. H., & Altona, C. (1979) *Nature (London) 281*, 235–236.

Haasnoot, C. A. G., de Leeuw, F. A. A. M., de Leeuw, H. P. M., & Altona, C. (1981) *Org. Magn. Reson. 15*, 43–52.

Haasnoot, C. A. G., de Bruin, S. H., Berendsen, R. G., Janssen, H. G. J. H., Binnendijk, T. J. J., Hilbers, C. W., van der Marel, G. A., & van Boom, J. H. (1983a) *J. Biomol. Struct. Dyn. 1*, 115–129.

Haasnoot, C. A. G., Westerink, H. P., van der Marel, G. A., & van Boom, J. H. (1983b) *J. Biomol. Struct. Dyn. 1*, 131–149.

Haasnoot, C. A. G., de Bruin, S. H., & Hilbers, C. W. (1985) *Proc. Int. Symp. Biomol. Struct. Interact., Suppl. J. Biosci. 8*, 767–780.

Haasnoot, C A. G., Blommers, M. J. J., & Hilbers, C. W. (1987) *Springer Series in Biophysics, Vol. I: Structure and Dynamics and Function* (Ehrenberg, A., Rigler, R., Graslund, A., & Nilssen, L., Eds.) pp 212–216, Springer-Verlag, New York.

Hare, D. R., Wemmer, D. E., Chou, S.-H., Drobny, G., & Reid, B. R. (1983) *J. Mol. Biol. 171*, 319–336.

Hare, R. H., & Reid, B. R. (1986) *Biochemistry 25*, 5341–5350.

Hilbers, C. W. (1979) in *Biological Applications of Magnetic Resonance* (Shulman, R. G., Ed.) pp 1–43, Academic Press, New York.

Hilbers, C. W., & Patel, D. J. (1975) *Biochemistry 14*, 2656–2660.

Hilbers, C. W., Haasnoot, C. A. G., de Bruin, S. H., Joordens, J. J. M., van der Marel, G. A., & van Boom, J. H. (1985) *Biochimie 67*, 685–695.

IUPAC–IUB Nomenclature Commission (1983) *Eur. J. Biochem. 131*, 9–15.

IUPAC–IUB Nomenclature Commission (1986) *J. Biol. Chem. 261*, 13–17.

Kaptein, R., Zuiderweg, E. R. P., Scheek, R. M., Boelens, R., & van Gunsteren, W. F. (1985) *J. Mol. Biol. 182*, 179–182.

Lavery, R., & Sklenar, H. (1988) *J. Biomol. Struct. Dyn. 6*, 63–91.

Lerner, D. B., Becktel, W. J., Everett, R., Goodman, M., & Kearns, D. R. (1984) *Biopolymers 23*, 2157–2172.

Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95–117.

Macura, S., Huang, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson. 43*, 259–281.

Mellema, J.-R., Pieters, J. M. L., van der Marel, G. A., van Boom, J. H., & Altona, C. (1984) *Eur. J. Biochem. 143*, 285–301.

Müller, U. R., & Fitch, W. M. (1982) *Nature (London) 298*, 582–585.

Orbons, L. P. M., & Altona, C. (1986) *Eur. J. Biochem. 160*, 141–148.

Orbons, L. P. M., van der Marel, G. A., van Boom, J. H., & Altona, C. (1987a) *J. Biomol. Struct. Dyn. 4*, 939–963.

Orbons, L. P. M., van Beuzekom, A. A., & Altona, C. (1987b) *J. Biomol. Struct. Dyn. 4*, 965–987.

Patel, D. J., & Hilbers, C. W. (1975) *Biochemistry 14*, 2651–2656.

Patel, D. J., Kozlowski, S. A., Marky, L. A., Broka, C., Rice, J. A., Itakura, K., & Breslauer, K. J. (1982a) *Biochemistry 21*, 428–436.

Patel, D. J., Kozlowski, S. A., Marky, L. A., Rice, J. A., Broka, C., Itakura, K., & Breslauer, K. J. (1982b) *Biochemistry 21*, 451–455.

Pramanik, P., Kanhouwa, N., & Kan, L.-S. (1988) *Biochemistry 27*, 3024–3031.

Remin, M., & Shugar, D. (1972) *Biochem. Biophys. Res. Commun. 48*, 636–643.

Rinkel, L. J., & Altona, C. (1987) *J. Biomol. Struct. Dyn. 4*, 621–648.

Rinkel, L. J., Sanderson, M. R., van der Marel, G. A., van Boom, J. H., & Altona, C. (1986) *Eur. J. Biochem. 159*, 85–93.

Rinkel, L. J., van der Marel, G. A., van Boom, J. H., & Altona, C. (1987) *Eur. J. Biochem. 166*, 87–101.

Roth, K., Kimber, B. J., & Feeney, J. (1980) *J. Magn. Reson. 41*, 302–309.

Scheek, R. M., Boelens, R., Russo, N., van Boom, J. H., & Kaptein, R. (1984) *Biochemistry 23*, 1371–1376.

Seibel, G. L., Singh, U. C., & Kollman, P. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 6537.

Sheflin, L. G., & Kowalski, D. (1985) *Nucleic Acids Res. 13*, 6137–6154.

Sinden, R. R., & Pettijohn, D. E. (1984) *J. Biol. Chem. 259*, 6593–6600.

van den Hoogen, Y. Th., Hilgersom, C. M. A., Brozda, D., Lesiak, K., Torrence, P. F., & Altona, C. (1989) *Eur. J. Biochem. 182*, 629–637.

van Gunsteren, W. F., Boelens, R., Kaptein, R., Scheek, R. M., & Zuiderweg, E. R. P. (1985) in *Molecular Dynamics and Protein Structure* (Hermans, J., Ed.) p 82, Polycrystal Book Service, Western Spring, IL.

Wagner, G., & Wüthrich, K. (1979) *J. Magn. Reson. 33*, 675–680.

Wagner, G., Kumar, A., & Wüthrich, K. (1984) *Eur. J. Biochem. 114*, 375-384.
Weaver, D. T., & DePamphilis, M. L. (1984) *J. Mol. Biol. 180*, 961-986.
Weiner, S. J., Kollman, P. A., Nguyen, D. T., & Case, D. A. (1986) *J. Comput. Chem. 7*, 230-235.
Wider, G., Macura, S., Kumar, A., Ernst, R. R., & Wüthrich, K. (1984) *J. Magn. Reson. 56*, 207-234.

Characterization of the Overall and Internal Dynamics of Short Oligonucleotides by Depolarized Dynamic Light Scattering and NMR Relaxation Measurements†

Wolfgang Eimer, James R. Williamson,‡ Steven G. Boxer, and R. Pecora*
Department of Chemistry, Stanford University, Stanford, California 94305-5080
Received July 10, 1989; Revised Manuscript Received September 1, 1989

ABSTRACT: The dynamics of three synthetic oligonucleotides $d(CG)_4$, $d(CG)_6$, and d(CGCGTTGTTCGCG) of different length and shape were studied in solution by depolarized dynamic light scattering (DDLS) and time-resolved nuclear Overhauser effect cross-relaxation measurements. For cylindrically symmetric molecules the DDLS spectrum is dominated by the rotation of the main symmetry axis of the cylinder. The experimental correlation times describe the rotation of the oligonucleotides under hydrodynamic stick boundary conditions. It is shown that the hydrodynamic theory of Tirado and Garcia de la Torre gives good predictions of the rotational diffusion coefficients of cylindrically symmetric molecules of the small axial ratios studied here. These relations are used to calculate the solution dimensions of the DNA fragments from measured correlation times. The hydrodynamic diameter of the octamer and dodecamer is 20.5 ± 1.0 Å, assuming a rise per base of 3.4 Å. The tridecamer, d(CGCGTTGTTCGCG), adopts a hairpin structure with nearly spherical dimensions and a diameter of 23.0 ± 2.0 Å. The DDLS relaxation measurements provide a powerful method for distinguishing between different conformations of the oligonucleotides (e.g., DNA double-helix versus hairpin structure). Furthermore, the rotational correlation times are a very sensitive probe of the length of different fragments. The NMR results reflect the anisotropic motion of the molecules as well as the amount of local internal motion present. The experimental correlation time from NMR is determined by the rotation of both the short and long axes of the oligonucleotide. The contributions of the various correlation times are given by the relative orientation of the internuclear vector with respect to the main axis and the correlation function for internal motion, if it is assumed that the librational motions decay much faster than the overall reorientational motion. Our results are compared with those from NMR relaxation measurements on other short oligonucleotides with lengths of up to 20 base pairs. Various dynamic models for the reorientation of the internuclear vector are applied to their interpretation. The extent of local internal motion is apparently independent of the length and base sequence of the DNA fragments and gives a significant contribution to the effective correlation time measured by the cross-relaxation rate.

The biological function of DNA and RNA is most likely determined not only by sequence-dependent properties and static conformation but also by its dynamics. DNA is a semiflexible macromolecule and, due to its length, exhibits a complex dynamic behavior. The local internal motions of DNA have become the subject of increasing interest in recent years. Fast fluctuations in the conformation of DNA may play an important role in transcription and the DNA-protein recognition process. The conformational fluctuations occur on a time scale between 10^{-6} and 10^{-12} s and include long-range twisting and bending motions as well as very local fluctuations of small segments about an equilibrium position. Experiments using various types of probes, mainly time-resolved fluorescence polarization anisotropy (FPA) (Genest & Wahl, 1978; Barkley & Zimm, 1979; Millar et al., 1981; Magde et al., 1983), ESR (Bobst et al., 1984; Kao & Bobst, 1985), and NMR[1] relaxation experiments of different nuclei (Early & Kearns, 1979; Bolton & James, 1980a,b; Hogan & Jardetzky, 1980; Mirau & Kearns, 1983; Assa-Munt et al., 1983; Mirau et al., 1985), have been performed to characterize the nanosecond to subnanosecond motions. The fluorescence depolarization measurements show a very rapid initial decay which could be partially resolved by using a picosecond laser setup (Magde et al., 1983). The very fast decay was attributed to the "wobbling" of the intercalated dye molecules (Barkley & Zimm, 1979) with a root mean square fluctuation of about ±21°. ESR and NMR studies confirmed the existence of librational motions for the bases and other segments in polynucleotides. The time scale and especially the amplitude for the relaxation processes have been the subject of controversy (Hård, 1987; Schurr & Fujimoto, 1988). For an unambiguous interpretation of NMR and ESR results, it is necessary to provide a precise theory that accounts for the rather complex dynamics of the long DNA fragments, where it is expected that different relaxation modes couple. When collective twisting and bending motions are neglected, the experimental

† This work was supported by a grant from the National Institutes of Health (R01 GM27738) to S.B. and a grant from the National Science Foundation (CHE-88-14641) to R.P. Further support came from the NSF-MRL program through the Center for Materials Research at Stanford University.
‡ Present address: Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80309.

[1] Abbreviations: NMR, nuclear magnetic resonance; DDLS, depolarized dynamic light scattering; NOE, nuclear Overhauser enhancement.

0006-2960/90/0429-799$02.50/0