# Effect of the Methyl Group on DNA Bending and Curvature: Structure of $d(GA_4U_4C)_2$ in Solution†

Kimiko Umemoto,‡ Mukti H. Sarma, Goutam Gupta, and Ramaswamy H. Sarma*
*Institute of Biomolecular Stereodynamics, Department of Chemistry, State University of New York, Albany, New York 12222*
*Received August 16, 1989; Revised Manuscript Received January 12, 1990*

ABSTRACT: NMR studies on $d(GA_4T_4C)_2$ and $d(GT_4A_4C)_2$ indicated two important factors that contribute to intrinsic DNA bending in polymers containing A/T tracts [Sarma, M. H., et al. (1988) *Biochemistry 27*, 3423; Gupta, G., et al. (1988) *Biochemistry 27*, 7909]. They are (i) propeller-twisted A·T pairs with associated bifurcated H bonds inside the A/T tract and (ii) the base sequence that joins the two neighboring A/T tracts. As an extension of our bending project, we carried out quantitative NMR studies on the decamer $d(GA_4U_4C)_2$, a structural analogue of $d(GA_4T_4C)_2$, to examine the effect of the methyl group on DNA bending. On the basis of quantitative NMR analysis, we arrive at the following results. (i) The decamer $d(GA_4U_4C)_2$ adopts the gross morphology of a right-handed B-DNA duplex with A and U nucleotides belonging to C2′-endo,anti domain. (ii) A·U pairs are propeller twisted and hence can result in an array of interstrand bifurcated H bonds involving N6 of A and O4 of U (one base pair apart) inside the A/U tract. (iii) The orientations of A and U with respect to the long axis of the molecule are different; as a result, at the A5-U6 sequence that joins the two A/U tracts, two neighboring frames of reference do not exactly coincide in space and a junction is created at A5-U6. (iv) Inside the A/U tract, intrastrand stacking is more compact (average separation between seccessive base planes being 3.2 Å) than at the A5-U6 junction, where average separation between the base planes of A5 and U6 is 3.6 Å. The NMR-derived structure of $d(GA_4U_4C)_2$ shows clearly identifiable differences from $d(GA_4T_4C)_2$, i.e., more compact intrastrand stacking inside the A/U tract of $d(GA_4U_4C)_2$ than in the A/T tract of $d(GA_4T_4C)_2$, and the A5-U6 junction of $d(GA_4U_4C)_2$ is more relaxed than the A5-T6 junction of $d(GA_4T_4C)_2$.

For past few years, our laboratory has been actively engaged in understanding the role of $A_n/T_n$ tracts in causing intrinsic DNA bending. In particular, we are interested in understanding the structural peculiarity of an $A_n/T_n$ tract in stereochemical details and also structural effects as caused by joining two neighboring $A_n/T_n$ tracts in a DNA molecule. Previously, we carried out extensive solution NMR studies on two decamers, i.e., $d(GA_4T_4C)_2$ and $d(GT_4A_4C)_2$. With the aid of a quantitative NOESY simulation technique, we demonstrated clear structural differences between $d(GA_4T_4C)_2$ and $d(GT_4A_4C)_2$ in solution: in the case of $d(GA_4T_4C)_2$ there was a structural discontinuity at the A5-T6 sequence, while for $d(GT_4A_4C)_2$ there was no discontinuity at T5-A6 (Sarma et al., 1988; Gupta et al., 1988). These observations led us to believe that the A-T junction formation as found in the decamer repeat $d(GA_4T_4C)_2$ can cause macroscopic bending in the polymer $d(GA_4T_4C)_{n>10}$, while smooth joining at the T-A sequence, as in $d(GT_4A_4C)_2$, results in a straight DNA for the polymer $d(GT_4A_4C)_{n>10}$. Subsequent to our findings from NMR data, Chen et al. (1988) performed high-resolution electrophoretic measurements on $d(GA_4T_4C)_2$ and $d(GT_4A_4C)_2$. They observed that $d(GA_4T_4C)_2$ was electrophoretically retarded compared to $d(GT_4A_4C)_2$. Thus, their results are consistent with our NMR analysis (Sarma et al., 1988; Gupta et al., 1988) and Hagerman's data on polymeric DNA (Hagerman, 1986).

NMR analysis of $d(GA_4T_4C)_2$ and $d(GT_4A_4C)_2$ also revealed a distinct structural property of the $A_n/T_n$ tracts in solution; i.e., inside the tract A·T pairs tend to be propeller twisted and this can result in a series of interstrand bifurcated H bonds in the major groove in a manner very similar to what has been observed in single crystals (Coll et al., 1987; Nelson et al., 1987). However, a clear distinction should be made between the crystal and NMR studies. Single-crystal studies provide a direct demonstration that the array of bifurcated H bonds is present inside the $A_n/T_n$ tracts. The presence of bifurcated H bonds is merely inferred from the NMR studies; i.e., 1D/2D NMR data are consistent with propeller-twisted A·T pairs, which in turn can give rise to an array of bifurcated H bonds inside the $A_n/T_n$ tracts.

In the past, it was proposed that the bulky methyl groups of thymine (i.e., the steric and hydrophobic effects) impart structural peculiarity in the $A_n/T_n$ tracts, and therefore, the methyl groups should play a key role in DNA bending (Olson et al., 1985; Jernigan et al., 1986). Such a proposal was in sharp disagreement with the electrophoretic measurements of Koo and Crothers (1987) on $d(GT_5GCCC)_n$ and $d(GU_5GCCC)_n$, which show that the replacement of T by U had no effect on the electrophoretic mobilities in the two polymers. In an attempt to resolve this question, we undertook structural studies of DNA oligomers containing A/U tracts. In this article, we report the results of our studies on the solution structure of $d(GA_4U_4C)_2$, the structural analogue of $d(GA_4T_4C)_2$.

## Materials and Methods

*DNA Synthesis and Purification.* The decamer $d(GA_4U_4C)_2$ was synthesized on a DNA synthesizer (Applied Biosystems Model 380 A) following the method of Matteuci et al. (1981). The product was purified on a 1.1 × 50 cm column of Q Sepharose (Pharmacia) with a linear gradient

†This research is supported by a grant from the National Institutes of Health (GM37812, GM29787) and by a contract from the National Foundation of Cancer Research. The high-field NMR experiments were performed at the NMR Facility for Biomolecular Research located at the F. Bitter National Magnet Laboratory, MIT. The NMR facility is supported by Grant RR00995 from the Division of Research Resources of the NIH and by the National Science Foundation under Contract C-670. This work is also supported by the U.S. Department of Energy.
‡Permanent address: Department of Chemistry, International Christian University, Mitaka, Tokyo 181, Japan.

0006-2960/90/0429-4714$02.50/0 

of 0.2–0.8 M NaCl in 10 mM NaOH (pH 12.0) and further purified by several ethanol precipitations.

*NMR Spectroscopy.* For both $H_2O$ and $D_2O$ samples, the DNA concentration was 2 mM in duplex, and the salt concentration was 100 mM in NaCl (pH 7.0 in 10 mM sodium phosphate buffer with 1 mM EDTA). One-dimensional NMR spectra of $d(GA_4U_4C)_2$ in $H_2O$ were recorded at 2–20 °C, with a time-shared long-pulse sequence. One-dimensional NOE spectra of $d(GA_4U_4C)_2$ in $H_2O$ were recorded for presaturation times $\tau_m$ = 200 and 100 ms (at 12 °C) and 50 ms (at 2 °C), with a relaxation delay RD = 1.2. NOESY spectra for $\tau_m$ = 50, 100, and 200 ms were collected at 12 °C in $D_2O$ with the pulse sequence $[RD-90°-\tau_1-90°-\tau_m-90°-Acq]_{NS}$.

The HDO signal was presaturated in the NOESY experiments. Two MINSY (mixing irradiation during a NOESY experiment) spectra, one with the decoupler irradiating at 2.2 ppm and the other at 2.6 ppm, were obtained for $\tau_m$ = 200 ms at 12 °C (Massefski & Redfield, 1988). In a MINSY experiment, a few protons are selectively saturated (during mixing) such that they are absent in the cross-relaxation process of the NOE buildup.

*NOESY Data Simulation.* The NOESY intensity between protons $i$ and $j$ is computed from

$$a_{ij} = [\exp(-R\tau_m)]_{ij} \quad (1)$$

where $R$ is the $(n \times n)$ relaxation matrix dependent upon the correlation time $\tau_c$ and interproton distances $r_{ij}$ [for details of the methodology and its application, see Keepers and James (1984), Broido et al. (1985), Gupta et al. (1988), and Sarma et al. (1988)]. We define a quantity $a^c_{ij}$ as

$$a^c_{ij} = \frac{a_{ij}}{a_{ii}} \times 100\% \quad (2)$$

for all $i$'s.

The quantity of $a^c_{ij}$ is compared with the corresponding relative NOESY intensity $a^o_{ij}$ as observed from the 1D NOESY slice through the $i$th proton, i.e.

$$a^o_{ij} = \frac{\text{peak height at the } j\text{th NOE site}}{\text{peak height at the } i\text{th NOE site}} \quad (3)$$

To estimate the quantitative agreement, we compute the agreement index $R$ factor as

$$R = \frac{1}{n}\sum_{i,j\neq i}^{n} \frac{|a^o_{ij} - a^c_{ij}|}{a^o_{ij}} \quad (4)$$

for all $i$'s at each $\tau_m$.

*NOESY and MINSY Data Points for R-Factor Refinement.* 1D slices through the following signals are chosen for NOESY intensity measurements: H8/H6 (G1, A2, A3, C10, U6), H2/H5 (A2, A5, U6, U7), H1′ (A2, A3, A5), and H3′ (C10) (Table I). In most of these cases the NOE source peaks and/or NOE cross peaks were not overlapping. The pair H2′,H2″ is treated as nonoverlapping when both H2′ and H2″ did not show overlap with other signals. Nonoverlapping H2′,H2″ pairs are due to G1, A5, U6, U9, and C10 (Table I and Figure 5, parts E and F). For all these protons, NOESY slices could be taken through them or NOESY intensities can be measured to obtain the experimental peak heights $a^o_{ij}$ (due solely to one pairwise interaction) used in the structure refinement (and $R$-factor calculation) using the full-matrix NOESY simulation. As seen from Figure 5, NOESY slices through these protons at a given $\tau_m$ show 3–6 NOE sites; therefore, from each NOESY experiment we have ~70 independent NOESY intensities. In the present case, NOESY experiments are performed at three $\tau_m$ (i.e., 50, 100, and 200 ms). Thus, we have in all about 70 × 3 = 210 observed intensities for pairwise interactions from 50-, 100-, and 200-ms NOESY data. An additional 30 data points were obtained from the 200-ms MINSY data. Thus, in all we have 210 + 30 = 240 data points from NOESY and MINSY experiments. Out of these, only the NOESY data at 50 and 100 ms and the MINSY data at 200 ms are used for the NOESY simulation and associated $R$-factor refinement. In addition, the 1D NOE data ($\tau_m$ = 50, 100, and 200 ms) are used to obtain the gross morphology of $d(GA_4U_4C)_2$ (Gupta et al., 1988).

Table I: Chemical Shift Values (ppm) of Protons in the $d(GA_4U_4C)_2$ Duplex in Aqueous Solution at 12 °C with TSP as an Internal Standard

| | H8/H6 | H2/H5 | H1′ | H2′ | H2″ | H3′ | NH |
|---|---|---|---|---|---|---|---|
| G1 | 7.83 | | 5.49 | 2.42 | 2.65 | 4.81 | 12.75 |
| A2 | 8.14 | 7.36 | 5.71 | 2.68 | 2.79 | 5.02 | |
| A3 | 8.08 | 7.05 | 5.83 | 2.62 | 2.85 | 5.04 | |
| A4 | 8.00 | 7.02 | 5.93 | 2.56 | 2.95 | 5.04 | |
| A5 | 8.00 | 7.55 | 6.08 | 2.46 | 2.89 | 5.02 | |
| U6 | 7.28 | 4.84 | 5.96 | 2.05 | 2.57 | ~4.9 | 13.95 |
| U7 | 7.67 | 5.40 | 6.22 | 2.24 | 2.67 | ~4.9 | 14.24 |
| U8 | 7.70 | 5.52 | 6.22 | 2.24 | 2.67 | ~4.9 | 14.10 |
| U9 | 7.65 | 5.54 | 6.18 | 2.14 | 2.62 | ~4.9 | 14.10 |
| C10 | 7.49 | 5.49 | 6.20 | 2.28 | 2.24 | 4.60 | |

*Line-Shape Analysis for Overlapping Signals.* When two or more protons overlap at one NOE site, it is difficult to experimentally obtain the individual NOEs corresponding to each overlapping proton. Hence, under such a situation we resorted to a line-shape analysis in which the contribution of a proton (either the source or site of NOE) to the NOESY slice at a chemical shift value is computed as

$$F_{ij} = 2a^c_{ij}\frac{\exp(-\Delta x^2/0.9\Gamma^2)}{1 + \exp(-\Delta x^2/0.9\Gamma^2)} \quad (5)$$

where $a^c_{ij}$ is the calculated intensity of the NOE at $j$ due to $i$ by full-matrix NOESY simulation (expressions 1 and 2); $\Delta x = x - x_0$; $x_0$ is the chemical shift value of the $j$th proton; $x$ is the chemical shift value at which the contribution of the $j$th proton to the whole NOESY slice is computed; and $\Gamma$ is the half line width. $\Gamma$ is not a variable parameter for NOESY simulation (and related $R$-factor calculation). The estimate of $\Gamma$ for the overlapping peaks is obtained from the nonoverlapping peak of the same type, i.e., $\Gamma$ for overlapping H8/H6 is obtained from the corresponding value of the nonoverlapping H8/H6 and similarly for H5/H2, H1′, etc. The numbers 2 (in the numerator) and 0.9 (in the exponent) are the normalizing factors.

Expression 5 allows us to compute the theoretical NOESY slice at all chemical shift values and compare with the corresponding experimentally observed one.

For computation of theoretical NOESY intensities, an isotropic overall correlation time $\tau_c$ was used for all pairwise interactions. This approximation caused no problems in the simulation of NOESY data measured for $\tau_m$ = 50 ms. However, we have found that, especially in $d(GA_4U_4C)_2$, local spin diffusion plays a large part in determining the NOESY ($\tau_m$ = 200 ms) pattern, which could not be simulated satisfactorily (for H1′–H8/H6, H3′–H8/H6 NOEs) with a single isotropic correlation time. At $\tau_m$ = 100 ms, residual NOESY peaks were present for H8(A)–H3′(A). Such peaks can originate either due to spin diffusion H8–H2′, H2″–H3′ in B DNA or due to primary NOE H8–H3′ in A DNA. To resolve this situation, we conducted MINSY experiments in which H2′ and H2″ protons were selectively irradiated. If H8(A)–H3′ NOE is a primary effect in the NOESY experiment, it should show up as a cross peak in the MINSY experiment;

Table II: Helical Parameters of Cambridge Conventions for the NOESY-Derived Structures of $d(GA_4U_4C)_2$, $d(GA_4T_4C)_2$, and $d(GT_4A_4C)_2$[a]

| | tilt ($\tau$) (deg) | roll ($\rho$) (deg) | twist ($\Omega$) (deg) | slide ($D_y$) (Å) | rise ($D_z$) (Å) | Pr Tw[b] ($\omega$) (deg) | buckle (bu) (deg) |
|---|---|---|---|---|---|---|---|
| $d(GA_4U_4C)_2$ | | | | | | | |
| G1–C10 | | | | | | −11.7 | 11.0 |
| A2–U9 | −0.3 | −2.0 | 35.4 | −0.3 | 3.2 | −19.4 | 8.9 |
| A3–U8 | 1.5 | −2.3 | 35.3 | −0.3 | 3.2 | −19.3 | 9.1 |
| A4–U7 | 1.5 | −2.3 | 35.5 | −0.3 | 3.2 | −19.3 | 9.2 |
| A5–U6 | 1.5 | −2.3 | 35.8 | −0.3 | 3.2 | −19.3 | 9.3 |
| U6–A5 | 0.1 | −6.9 | 37.2 | −0.1 | 3.6 | −19.3 | −9.2 |
| U7–A4 | −1.5 | −2.4 | 35.8 | −0.3 | 3.2 | −19.3 | −9.2 |
| $d(GA_4T_4C)_2$ | | | | | | | |
| G1–C10 | | | | | | −16.2 | 5.8 |
| A2–T9 | 0.6 | −3.1 | 35.4 | −0.4 | 3.3 | −20.8 | 5.9 |
| A3–T8 | 1.4 | −2.8 | 35.0 | −0.3 | 3.3 | −20.7 | 6.2 |
| A4–T7 | 1.4 | −2.8 | 34.8 | −0.3 | 3.3 | −20.6 | 6.6 |
| A5–T6 | 1.4 | −2.8 | 35.0 | −0.3 | 3.3 | −20.5 | 6.9 |
| T6–A5 | 0.1 | −7.0 | 38.3 | −0.3 | 3.4 | −20.4 | −7.0 |
| T7–A4 | −1.4 | −2.8 | 35.5 | −0.3 | 3.3 | −20.5 | −6.9 |
| $d(GT_4A_4C)_2$ | | | | | | | |
| G1–C10 | | | | | | −16.1 | 5.4 |
| T2–A9 | −2.9 | −4.6 | 36.9 | 0.1 | 3.5 | −23.6 | 11.1 |
| T3–A8 | −2.3 | −1.5 | 36.0 | 0.0 | 3.4 | −23.6 | 11.2 |
| T4–A7 | −2.3 | −1.5 | 36.0 | 0.0 | 3.4 | −23.6 | 11.1 |
| T5–A6 | −2.3 | −1.5 | 36.0 | 0.0 | 3.4 | −23.6 | 11.1 |
| A6–T5 | 0.0 | 5.8 | 30.8 | −0.2 | 3.5 | −23.6 | −11.1 |
| A7–T4 | 2.3 | −1.5 | 36.0 | 0.0 | 3.4 | −23.6 | −11.1 |

[a] Parameters were calculated by the method of Bhattacharya & Bansal (1988). There are several algorithms for conformational analysis of DNA. The main differences between them arise from the choice of helix axes (local or global) or definitions of certain parameters; for example, "rise ($D_z$)" is defined as the vertical separation between C1′ atoms of successive bases (NEWHELIX; Dickerson, private communication, 1989), e.g., that between the centers of mass of successive base pairs (Soumpasis and Tung, private communication, 1989) or that between the midpoints of the C6...C8 vectors of successive base pairs (Bhattacharya & Bansal, 1988). To best describe the short, irregular nucleotide oligomers of the present system, and especially to point out the irregularities in the base-pair separations, we found the method of Bhattacharya and Bansal most suitable. [b] Propeller twist.

if not, no such cross peak should be seen in the MINSY experiment. We used the MINSY data ($\tau_m$ = 200 ms) for simulation, in addition to the NOESY data ($\tau_m$ = 50 and 100 ms). The NOESY data obtained for $\tau_m$ = 200 ms, therefore, were not used for the simulation purposes.

The theoretical MINSY intensities were computed by following the same procedure as for the NOESY, except that the irradiated protons were omitted from the list of the protons in the data set.

*NMR Constrained Molecular Modeling.* Molecular models for $d(GA_4U_4C)_2$ were constructed by performing the following steps. *Step 1.* The NMR data suggested that the $GA_4/U_4$ blocks on either sides of A5-U6 2-fold were conformationally equivalent. Therefore, the structure of the $GA_4/U_4C$ block was first constructed constrained to the NOE data. The variables used for A and G were as follows: one set of $\alpha$, $\beta$, $\gamma$, $\delta$, $\epsilon$, and $\zeta$ and so (C4′–C3′–C2′–C1′) for A, G, one $\chi$ for A2, A3, A4, A5, and one $\chi$ for G1 (i.e., 7 + 1 + 1 = 9 torsion angles for $GA_4$). One set of $\alpha$, $\beta$, $\gamma$, $\delta$, $\epsilon$, and $\zeta$ and $\sigma_0$(C4′–C3′–C2′–C1′) for U, C, one $\chi$ for U6, U7, U8, U9, and one $\chi$ for C10 (i.e., 7 + 1 + 1 = 9 torsion angles for $U_4C$). A total of 9 × 2 = 18 torsion angles are varied to construct the $GA_4/U_4C$ block such that the calculated NOE for the protons inside the block agrees with the corresponding observed NOE values. The base parameters (Table II) for this structural block are also treated as variables; they are base tilt, propeller twist, and base displacement (Gupta et al., 1988).

*Step 2.* Two identical $GA_4/U_4C$ structural blocks are then joined at A5-U6 constrained to the NOE distance criteria. For this purpose the torsion angles $\epsilon$, $\xi$, $\alpha$, and $\beta$ between A5-U6 are treated as variables.

The NMR model of $d(GA_4T_4C)_2$ (Sarma et al., 1990) is also constructed by using the methodology described above.

*Limitations.* All orders of NOE are incorporated in expression 1. Therefore, the full-matrix NOESY simulation in conjunction with linked-atom least-squares model building (Gupta et al., 1988) can be used to circumvent the problem of spin diffusion. However, the method is not free of shortcomings. One drawback occurs in situations in which there is an extensive overlap of peaks, which limits the number of observed intensities due to a single pairwise interaction. We addressed this problem by considering only NOE intensities due to a single pairwise interaction for the *R*-factor refinement. The line-shape analysis was then performed on the model (after *R*-factor refinement) to examine if the NOESY pattern due to overlapping protons compared with the corresponding observed ones. Another shortcoming of the method (Gupta et al., 1988) is that the structural solution is not unique—also inherent in many other techniques used to derive structures from the 2D NMR data. Therefore, the structure that we provide for $d(GA_4U_4C)_2$ is only one of several probable structures that fit the NMR data. In the following paper (Sarma et al., 1990), we elaborated this point and showed how one can narrow the number of possibilities by carrying out NOESY distance constrained energy minimization. But even after energy minimization we do not obtain a unique structural solution. It may be emphasized that the NOESY data of a DNA oligomer at a given $\tau_m$ provide information about sugar pucker, base orientation (from intranucleotide NOE), and base–base interaction (from internucleotide NOE). NOESY experiments at various $\tau_m$ and the full-matrix simulation permit a reliable estimate of these structural parameters (Gupta et al., 1988). However, missing from the data set is the knowledge about the inherently flexible phosphate backbone. It turns out that more than one conformation of the phosphate backbone is consistent with the experimentally (NOESY) observed sugar pucker, sugar–base orientation, and base–base interaction. Thus, due to the very nature of our experimental system, the structural solution that we derive is never a unique one. However, it may be stressed that the model we propose

from our analysis is a stereochemically acceptable and energetically stable structure. Any other structural solution that might agree with the observed NMR data ought to lie in the neighborhood of our proposed model.

In addition, it may be pointed out that 2D NMR experiments only provide information about short-range (2–5-Å) distances and do not give any direct information about long-range interactions. This means that from the quantitative analysis of 2D NMR data we have fairly reliable knowledge about sugar pucker, sugar–base orientation, base–base interactions, and local geometry (short-range interactions), but we have no direct information about the spatial relation between a residue and its third or fourth neighbor in sequence (long-range interaction).

## Results

*Assignment of the NH Protons.* The temperature dependence of the imino protons of G·C and A·U Watson–Crick pairs was monitored to study the duplex nature of the decamer $d(GA_4U_4C)_2$ in $H_2O$. One-dimensional NOE experiments in $H_2O$ were then conducted at 12 °C and these are shown in Figure 1, where four imino proton signals corresponding to the four N3–H (of A·U pairs) and one N1–H (of the terminal G·C pair) are observed in the low-field region of the control spectrum of $d(GA_4U_4C)_2$ (Figure 1A). From the combination of temperature dependence and the 1D NOE difference spectra (B–E), the following assignments were obtained.

(*i*) *Spectrum B*: The signal at 12.75 ppm was assigned as N1–H of the terminal G1·C10 pair since it was most sensitive to the temperature change. By irradiation of this peak, weak NOE signals are observed at 7.37 ppm and also at 14.08 ppm. They are identified as H2 and the imino proton, respectively, of the adjacent base pair, A2·U9. Weak NOE peaks also observed at 7.68–8.44 ppm are from amino $NH_2$ protons, which are present in all NOE difference spectra B–E.

(*ii*) *Spectrum C*: A single strong NOE signal was observed by irradiating at 13.88 ppm. Because of the 2-fold symmetry of $d(GA_4U_4C)_2$ about the molecular $(^{A5\text{-}U6}_{U6\text{-}A5})$ center, this is only expected when N3–H of U6·A5 is irradiated. Hence, the signal at 13.88 ppm belongs to N3–H and that at 7.55 ppm to H2, of the A5·U6 pair.

(*iii*) *Spectrum D*: An NOE is observed at 7.55 ppm (already assigned as H2 of A5) by irradiating at 14.24 ppm, indicating that this signal belongs to the N3–H of A4·U7. There is also a strong NOE signal at 7.02 ppm, identified as H2 of A4. Weaker signals at 7.05 and 7.37 ppm are from A3·U8 and A2·U9 base pairs. The NOE signal at 7.37 ppm is not observed at $\tau_m$ = 50 ms, suggesting that it occurs due to a second-order NOE from the distant H2 of A2·U9. Note that the NOE at H2 of A5 is as strong as that of A4, while that of A3 is much weaker. This observation is consistent with propeller-twisted (>15°) A·U pairs.

(*iv*) *Spectrum E*: The signal at 14.10 ppm contains two overlapping protons, which are now expected to be N3–H protons of A2·U9 and A3·U8 pairs. Irradiation of this peak results in strong NOE signals at 7.37 and 7.05 ppm, identified as H2 protons of A2 and A3, respectively. There are small but noticeable NOE peaks at 7.02 (partially overlapping with the peak at 7.05 ppm) and 7.54 ppm from the more distant H2 protons of A4 and A5.

The chemical shift values of the imino NH protons and those of the H2 protons are included in Table I.

*Assignment of the Proton System H8/H6, H2/H5, H1′, H2′, H2″, and H3′.* A combination of NOESY data for $\tau_m$ = 50, 100, and 200 ms was used to arrive at the sequential assignment of the proton system H8/H6, H2/H5, H1′, H2′, H2″, and H3′ belonging to the 10 nucleotides in $d(GA_4U_4C)_2$ at 12 °C. By studying NOEs as a function of mixing time, primary NOEs can be isolated from secondary effects. Figure 2 compares the H8/H6 vs H5/H1′ NOESY cross sections for $\tau_m$ = 50 ms (A) and 100 ms (B). Note that, for $\tau_m$ = 100 ms, both the intra- and internucleotide H8/H6–H1′ cross peaks for all 10 nucleotides are observable, as well as the H8/H6–H5 and some of the H8–H3′ cross peaks. When $\tau_m$ is reduced to 50 ms, most of the H8/H6–H1′ and H8–H3′ cross peaks weaken in intensity. The five strong cross peaks remaining in spectrum A ($\tau_m$ = 50 ms) are easily identified as the H6–H5 cross peaks from the five pyrimidine bases in the molecule; four uracils and one cytosine. Note that these five cross peaks correspond to the 1D NMR signals that have broader line widths than the rest of the signals in the H8/H6 region. The extra line broadening in these H6 protons of the pyrimidine bases in the 1D spectrum comes from the H6–H5 coupling, unresolved at this temperature. Since the peaks at 7.37 and 7.55 ppm are from H2 protons of A2 and A5 (already

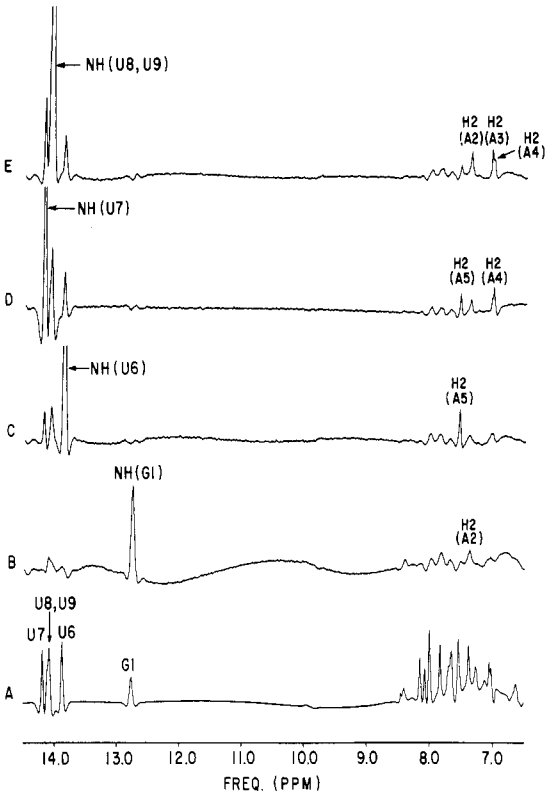



FIGURE 1: One-dimensional NOE difference spectra of the d-$(GA_4U_4C)_2$ duplex in $H_2O$ at 12 °C. Each of the imino proton signals was presaturated for $\tau_m$ = 200 ms. (A) The control spectrum shows the four imino proton signals and the signals in the base/amino proton region. (B) The signal at 12.75 ppm of the G1·C10 pair is irradiated and the weak NOE signals at 14.08 and 7.37 ppm are identified as N3–H(U9) and H2(A2) of the adjacent base pair, A2·U9. (C) A strong NOE signal is observed at 7.55 ppm, by irradiating the signal at 13.88 ppm of the A5·U6 pair, and identified as H2(A5). Because of the 2-fold symmetry of the duplex, only one signal is observed for the A5·U6 and U6·A5 pairs. (D) The signal at 14.24 ppm, assigned as N3–H of A4·U7, is irradiated and two strong NOE peaks are observed at 7.02 ppm from H2(A4) and 7.55 ppm from H2(A5). There are also two weak NOE signals observed at 7.05 ppm (from A3·U8) and 7.37 ppm (from A2·U9). The signal at 7.37 ppm is not observed at $\tau_m$ = 50 ms, indicating that it occurs due to a second-order NOE. Thus, irradiation of N3–H of the *i*th A·U pairs gives strong NOE at H2 of the $(i-1)$th A·U pair but not at H2 of the $(i+1)$th pair. (E) The signal at 14.10 ppm, containing signals from N3–H protons of both A2·U9 and A3·U8 pairs, was irradiated, resulting in strong NOE signals at 7.37 ppm (from A2·U9) and 7.05 ppm (from A3·U8). Small secondary NOEs are also observed at 7.02 and 7.54 ppm from the H2 protons of A4 and A5, respectively.

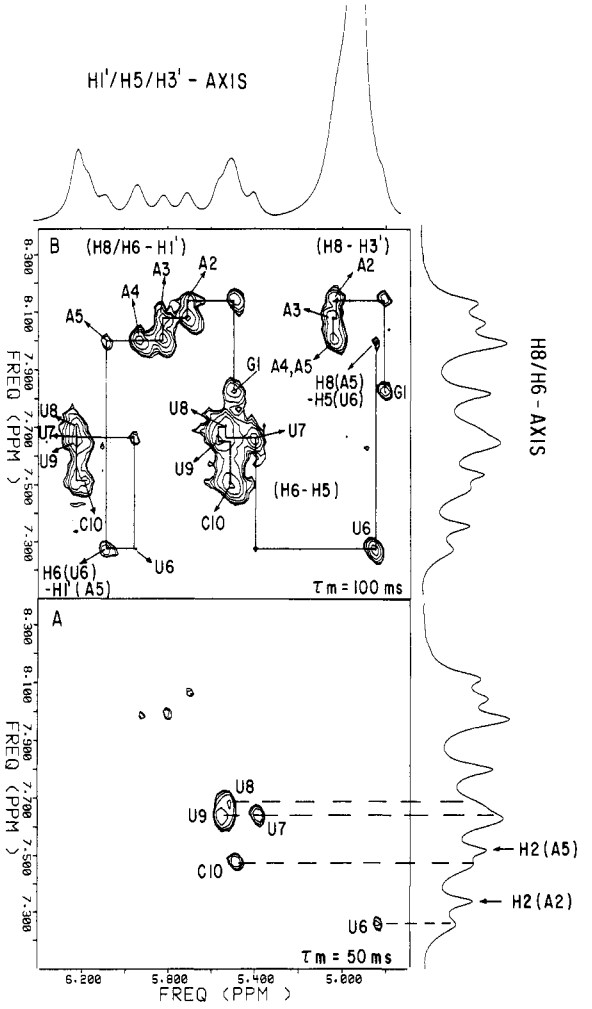



FIGURE 2: NOESY cross sections for the H8/H6 vs H5/H1′ region for (A) $\tau_m$ = 50 ms and (B) $\tau_m$ = 100 ms. In (A), five strong cross peaks are observed corresponding to the H6($i$)–H5($i$) cross peaks of the five pyrimidines, U6, U7, U8, U9, and C10. Note that these five peaks correspond to the 1D NMR signals and that these have broader line widths than the rest of the signals in the H8/H6 region. Since the peaks at 7.37 and 7.55 ppm have been identified as H2 protons of A2 and A5 (see Figure 1), the remaining four peaks at 7.8–8.2 ppm in the 1D spectrum are from the five purine bases in the molecule. The three small cross peaks found in this region are of H8($i$)–H1′($i$ − 1) from A2, A3, and A4. (B) At $\tau_m$ = 100 ms, all the cross peaks of H6($i$)–H5($i$ − 1) become observable in addition to the intranucleotide H6($i$)–H5($i$). Among them, only the peak at 7.26, 4.84 ppm has a connectivity with an H8 of A at 8.00, 4.84 ppm. Thus, this cross peak is assigned as H8(A5)–H6(U6), from which the H6–H5 cross-connectivity route is established as shown. Similarly, the H8/H6–H1′ cross connectivity is found by noting the cross peak at 7.26, 6.08 ppm, which connects H6(U6) with H1′(A5). (B) also contains H8–H3′ cross peaks for the five purine bases, while no such cross peaks are found for pyrimidines in the cross sections.

assigned in the previous section), the remaining four peaks at 7.8–8.2 ppm region are from the five purine bases in the molecule.

For $\tau_m$ = 100 ms, we observe all the cross peaks of H6–H5, not only for intranucleotide but also for internucleotide NOEs. Among the five H6–H5 cross peaks, only the one at 7.26, 4.84 ppm has a connectivity with an H8 of A, i.e., at 8.00, 4.84 ppm. Thus, this cross peak is assigned as H8(A5)–H5(U6), belonging to the bases that are located at the purine–pyrimidine junction. Starting from H6(U6), the H6–H5 cross-connectivity route was established as shown in Figure 2B.

Figure 2 also contains the H8/H6 vs H1′/H3′ region. Since NOEs from H8/H6 to H3's are weak compared to those from H8/H6 to H2′ and H2″ protons (see Figure 3), and totally

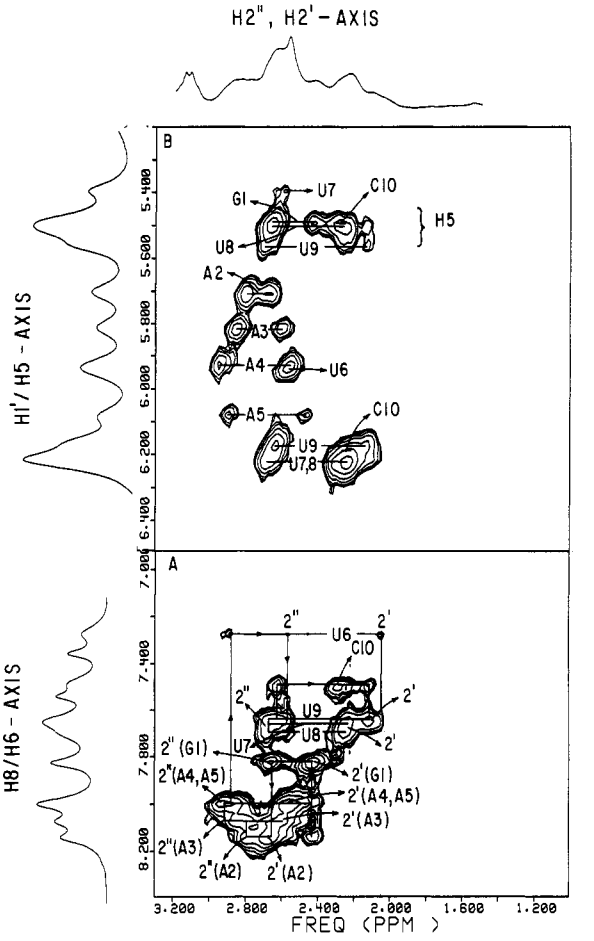



FIGURE 3: Assignment of the spin system H2′ and H2″ from NOESY cross sections of d($GA_4U_4C$)$_2$ observed at 12 °C in $D_2O$ for $\tau_m$ = 100 ms. (A) The cross-connectivity route involving H2″($i$ − 1)–H8/H6($i$)–H2′($i$) found here is what is expected for a B-DNA duplex with residues in average C2′-endo,anti geometry. (B) NOESY cross section for H2″($i$)–H8/H6($i$)–H2′($i$) cross peaks. Since H8/H6 and H1′/H5 chemical shifts have been assigned (Figure 2), (A) together with (B) results in the sequential assignment of the protons H2′ and H2″.

absent from $\tau_m$ = 50 ms NOESY spectrum, we conclude that the nucleotide geometry of each of the 10 residues belongs to average C2′-endo,anti conformation as in a B DNA. It follows then that one expects the NOE pattern H1′($i$ − 1)$\xrightarrow{3.0\text{–}3.8\ \text{Å}}$H8/H6($i$)$\xrightarrow{3.6\text{–}4.0\ \text{Å}}$H1′($i$) in the H8/H6 vs H1′ cross section. Here again, the H6(U6)–H1′(A5) cross peak at 7.28, 6.08 ppm can be used as a marker, since this is the only cross peak connecting a pyrimidine base proton with a purine H1′. Thus the cross-connectivity route, established as shown in Figure 2B, gives the sequential assignment of H8/H6, H5, and H1′ of the 10 nucleotides. Figure 2B shows H8–H3′ cross peaks for the five purine bases, while no such cross peaks were found for U's in the NOESY cross sections.

For a B-DNA structure, NOE cross connectivities are expected for H2″($i$ − 1)$\xrightarrow{2.1\text{–}2.5\ \text{Å}}$H8/H6($i$)$\xrightarrow{2.1\text{–}2.5\ \text{Å}}$H2′($i$) in the H8/H6 vs H2′,H2″ NOESY cross section, which is shown in Figure 3A for $\tau_m$ = 100 ms. Note that except for G1, U6, and C10, cross peaks are crowded and overlapping with each other. This is because at $\tau_m$ = 100 ms we also observe H2′($i$ − 1)$\xrightarrow{3.5\text{–}3.9\ \text{Å}}$H8/H6($i$)$\xrightarrow{3.5\text{–}3.9\ \text{Å}}$H2″($i$) cross peaks closely spaced together. They are more clearly resolved in Figure 3B, which displays the H2′($i$)$\xrightarrow{3.1\ \text{Å}}$H1′($i$)$\xrightarrow{2.4\ \text{Å}}$H2″($i$) and H2′($i$ − 1)$\xrightarrow{3.2\text{–}3.4\ \text{Å}}$H5($i$)$\xrightarrow{2.7\text{–}3.2\ \text{Å}}$H2″($i$ − 1) cross peaks.

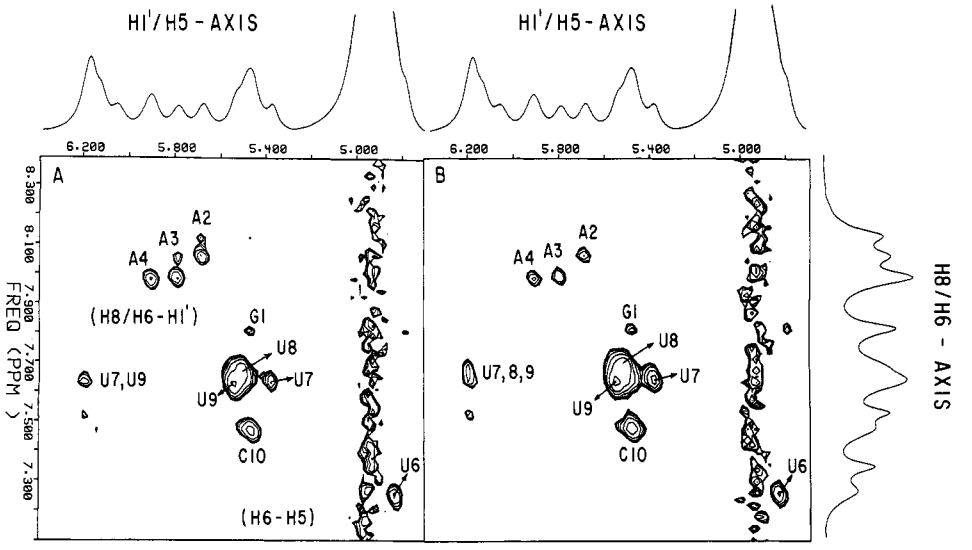



FIGURE 4: MINSY cross sections for H8/H6–H1′/H5 regions observed for $d(GA_4U_4C)_2$ at 12 °C in $D_2O$ for $\tau_m$ = 200 ms. (A) The decoupler is placed at 2.2 ppm, where H2′ of pyrimidines and H2″ of C10 are irradiated. Note that even at this long mixing time, 200 ms, most of the cross peaks of pyrimidines, except those of intranucleotide H6–H5, and many of the intranucleotide H8/H6–H1′ peaks have disappeared. (B) Here, the decoupler is placed at 2.6 ppm, irradiating H2′ of adenines and H2″ of uracils and guanine. All the H8($i$)–H1′($i$) peaks of adenines have disappeared.

All the H8/H6 and H1′/H5 chemical shifts were assigned in Figure 2 and Figure 3A together with B results in the sequential assignment of the protons H2′ and H2″. The assignment of the proton system (H8/H6, H2/H5, H1′, H2′, H2″, H3′) of the 10 residues is listed in Table I.

*Structural Features of $d(GA_4U_4C)_2$ As Observed in NMR Data.* (i) H8/H6–H3′ NOE cross peaks are consistently weaker than those of H8/H6–H2′,H2″, indicating that all 10 nucleotides of $d(GA_4U_4C)_2$ take C2′-endo,anti conformations in solution. That is, like $d(GA_4T_4C)_2$, $d(GA_4U_4C)_2$ belongs to the B-DNA family (Sarma et al., 1988).

(ii) As in the case of $d(GA_4T_4C)_2$, $d(GA_4U_4C)_2$ has a large propeller twist in the A·U pair as suggested by 1D NOE experiment in $H_2O$ (Figure 1). This observation is also supported by the fact that internucleotide H2″($i$ − 1)–H8/H6 NOEs are weaker than corresponding intranucleotide H8/H6–H2′($i$) NOEs.

(iii) Parts A and B of Figure 2 show that the internucleotide NOE involving H8/H6($i$)–H1′($i$ − 1) cross peaks is generally stronger than the corresponding intranucleotide NOE involving H8/H6($i$)–H1′($i$).

(iv) Weak interstrand NOE was observed for $\tau_m$ = 100 ms between H2(A4) and H1′(U8), and between H2(A5) and H1′(U7), suggesting that the distances between these protons are ~4.0 Å. Since these protons are located facing the minor groove, these distances roughly indicate the size of the minor groove.

(v) Cross-peak intensities, including these NOESY cross peaks mentioned above, are generally larger for $d(GA_4U_4C)_2$ when compared with those of $d(GA_4T_4C)_2$. This is particularly marked for H8/H6–H1′ cross peaks. For intranucleotide H8($i$)–H1′($i$), the distance usually is ~3.9 Å and at this distance one expects NOE cross-peak intensity for $\tau_m$ = 100 ms to be not more than 2%. In the case of $d(GA_4U_4C)_2$, this value is 5–8%, which is unusually large. This observation, together with high NOE intensities for H8/H6($i$)–H3′($i$) and H2($i$)–H2($i$ + 1) cross peaks as mentioned above, suggests that the $d(GA_4U_4C)_2$ molecule is compact, and as a result, spin diffusion plays a more dominant role for $d(GA_4U_4C)_2$ at $\tau_m$ = 100 and 200 ms than in the case of $d(GA_4T_4C)_2$ under the same conditions; in other words, observed H8($i$)–H1′($i$) NOEs should have originated from diffusion via H8($i$)$\xrightarrow{2.2\text{ Å}}$H2′($i$)$\xrightarrow{3.0\text{ Å}}$H1′($i$). For the same reason, we believe that the residual NOE for H8–H3′ at $\tau_m$ = 100 ms (Figure 2B) originated via diffusion pathway H8($i$)$\xrightarrow{2.2\text{ Å}}$H2′($i$)$\xrightarrow{2.4\text{ Å}}$H3′($i$) or H8($i$)$\xrightarrow{2.4\text{ Å}}$H2″($i$ − 1)$\xrightarrow{2.9\text{ Å}}$H3′($i$ − 1). In order to check this point, we conducted a MINSY experiment, in which either H2′ or H2″ was irradiated.

Parts A and B of Figure 4 show examples of MINSY cross sections for the H8/H6–H1′/H5 region for $\tau_m$ = 200 ms. Note that even at this long mixing time, most of the secondary cross peaks [H8–H3′, intranucleotide H8($i$)–H1′($i$)] are absent. When a decoupler is placed at 2.2 ppm (Figure 4A), the H2′ proton of the pyrimidine bases and H2″ of C10 are being irradiated. Hence, one expects the H6($i$)–H1′($i$ − 1) cross peaks of these nucleotides to be affected, while most of the purine H8($i$)–H1′($i$) and H8($i$)–H1′($i$ − 1) cross peaks are undisturbed. We also find that H6(U6)–H1′(A5) and H6(U7)–H1′(U6) cross peaks have disappeared. When the decoupler irradiates at 2.6 ppm (Figure 4B), H2′ of A's and H2″ of U's are being saturated. Note that, in addition to the above peaks, all the H8($i$)–H1′($i$) peaks of A's have disappeared. This implies that the NOE between H8($i$) and H1′($i$) is of a secondary nature due to spin diffusion via H2′. Again the effect is not obvious for U8-C10 due to overlaps; however, their H6($i$)–H1′($i$) cross peaks are likely affected. As will be discussed more quantitatively in the next section, irradiation of H2′ and/or H2″ protons effectively disrupted the diffusion pathways to reduce secondary effects.

*Structural Refinement through NOESY Simulations.* Structural models of $d(GA_4U_4C)_2$ were constructed by a least-squares refinement technique subject to the following distance constraints as revealed by the primary NOESY data: (i) Watson–Crick G·C and A·U pairing is present; (ii) its overall structure is a B DNA having C2′-endo,anti conformation; (iii) A·U pairs are propeller twisted; (iv) there is a 2-fold symmetry at the molecular center, $\binom{A5\text{-}U6}{U6\text{-}A5}$, about which the two structural blocks on both sides are conformationally equivalent to each other; (v) H8/H6($i$)–H2′($i$) distances are shorter than the corresponding H8/H6($i$)–H2″($i$ − 1) distances; (vi) H8/H6($i$)–H1′($i$ − 1) distances are shorter than the corresponding H8/H6($i$)–H1′($i$) distances; and (vii) although (A,G) and (U,C) belong to the same C2′-endo,anti

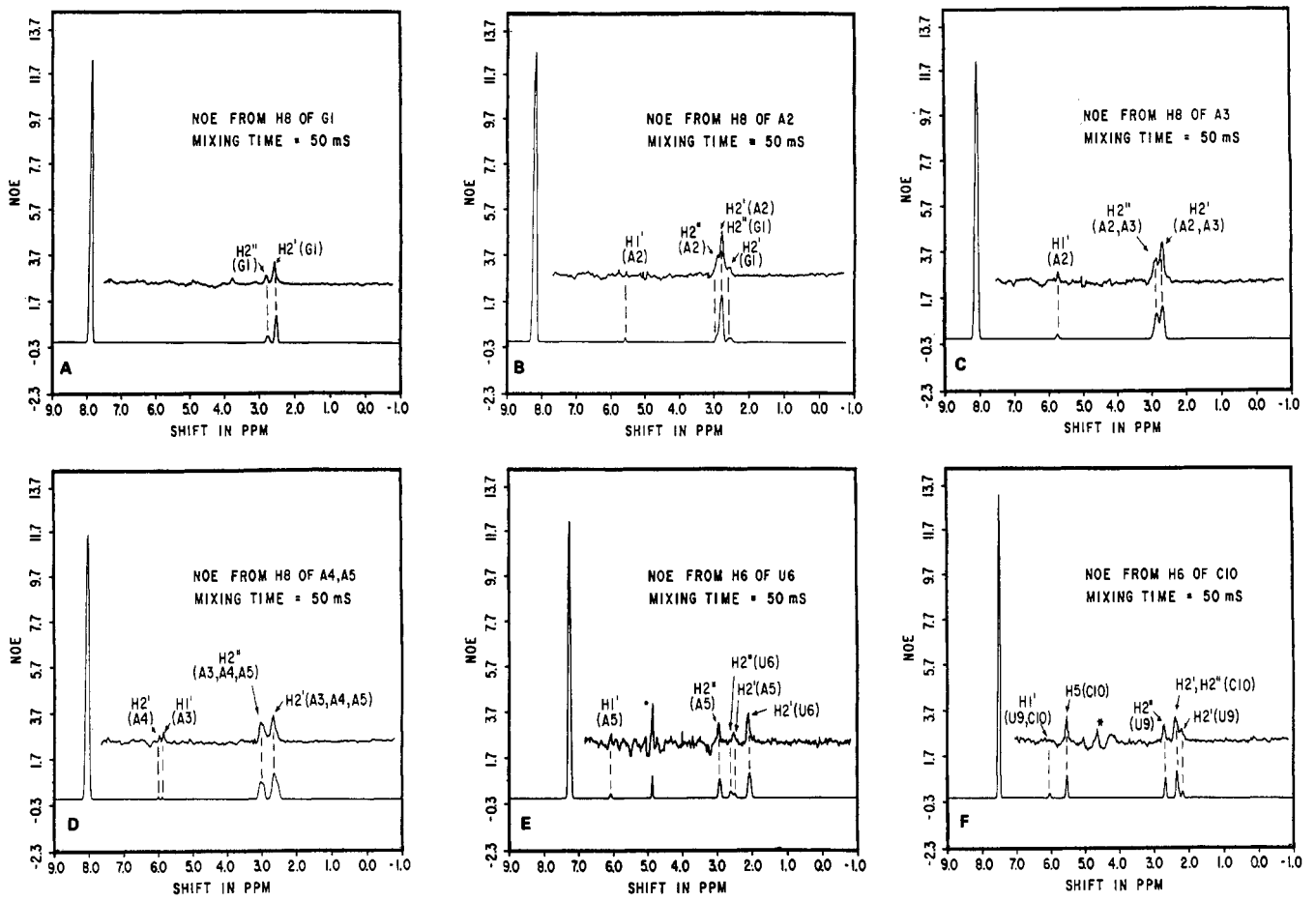



FIGURE 5: Examples illustrating the NOESY data simulation. Theoretically constructed NOESY slices are compared with the corresponding observed ones. The structural parameters have been adjusted for best fits. The intense peaks at the base region (7.2–8.2 ppm) are the peaks from which the NOESY slices have been obtained. Because of overlapping, the H2′,H2″ regions of B–D and F contain four to six resonance signals. The vertical axis indicates NOE intensities in arbitrary units. The peaks with * contain noise from HDO.

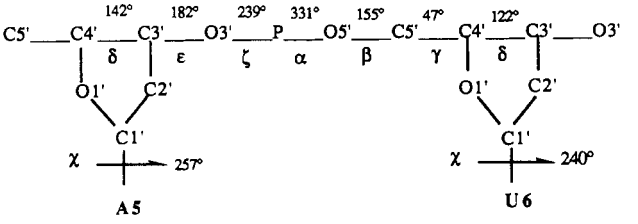
domain, the conformations of (A,G) and (U,C) are not exactly identical.

Inherent flexibility of the polynucleotide backbone allows several models of $d(GA_4U_4C)_2$ to roughly satisfy the distance criteria mentioned above. Quantitative NOESY simulations were performed on all these models. The model that showed the best agreement with the NOESY (at $\tau_m$ = 50 and 100 ms) and the MINSY (at $\tau_m$ = 200 ms) data is given as one of the probable time-average NMR models of $d(GA_4U_4C)_2$. The agreement of the final model with respect to the experimental NOE data is discussed below.

Parts A–F of Figure 5 compare the simulated and the observed NOESY slices through the base protons H8/H6 for $\tau_m$ = 50 ms. Note that the agreement is quite satisfactory. When $\tau_m$ is increased to 100 ms, some of the observed NOE intensities for H8/H6($i$)–H1′($i$) and H8/H6($i$)–H3′($i$) cross peaks tend to become large compared with the calculated values. Since agreement for MINSY is still satisfactory even at $\tau_m$ = 200 ms, this implies that the main source of the disagreement comes from the unusually efficient spin diffusion pathways via H2′ and/or H2″, the protons that are located close to H8/H6 protons. Thus, as far as the primary NOEs are concerned, our structure reflects all the major features of the NOE patterns observed by NMR.

The structural parameters of the B-DNA model thus obtained for $d(GA_4U_4C)_2$ are calculated by the method of Bhattacharya and Bansal (1988) and listed in Table II together with the corresponding parameters for $d(GA_4T_4C)_2$ and $d(GT_4A_4C)_2$ for comparison. The main-chain and glycosyl torsion angles of $d(GA_4U_4C)_2$ are as follows. For A and G: $\alpha$ = 328°, $\beta$ = 182°, $\gamma$ = 32°, $\delta$ = 142°, $\epsilon$ = 172°, $\xi$ = 250°, and $\chi$ = 257°. For U and C: $\alpha$ = 319°, $\beta$ = 173°, $\gamma$ = 45°, $\delta$ = 127°, $\epsilon$ = 178°, $\xi$ = 253°, and $\chi$ = 240°. Note that both A·U and G·C pairs have very little base tilt. The propeller twist in the A·U pairs is 19° while that in G·C pairs is 12°. The glycosyl torsion and propeller twist are reduced for G and C compared with those of A and U in order to accommodate three hydrogen bonds between G and C. It is also noted that purine nucleotides are conformationally different from the pyrimidine nucleotides and base pairs of A and U have different orientations with respect to the long axis of the molecule.

At the A5-U6 junction, two structural blocks $d(GA_4)$·$d(U_4C)$ and $d(U_4C)$·$d(GA_4)$ are connected constrained to the observed NOESY distance criteria. This resulted in a stereochemically allowed A5-U6 junction with a conformation as shown:




By use of these torsion angles, the first block is in effect shifted with respect to the second block by $\Delta x$ = 0.28 Å, $\Delta y$ = 0.45 Å, and $\Delta z$ = 0.54 Å. That is, rise ($D_z$) between A5 and U6 at the junction is 3.6 Å (Table II). Because of the large propellar twist in A·U pairs, such adjustment was necessary

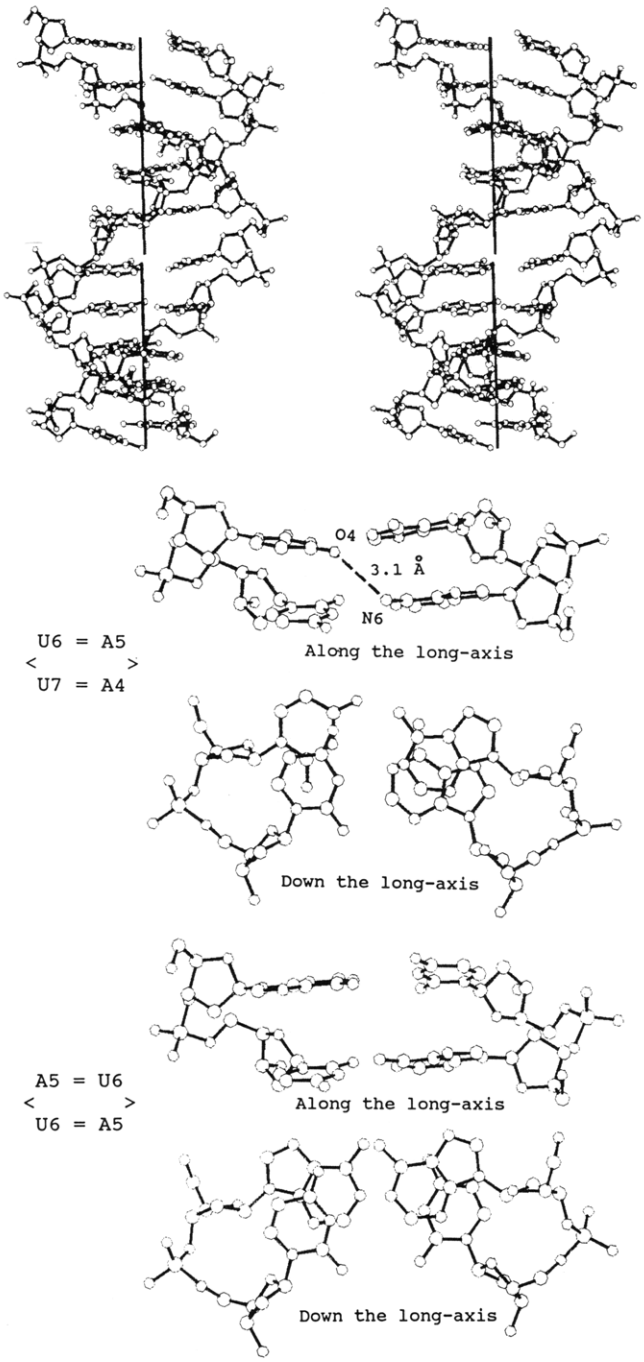


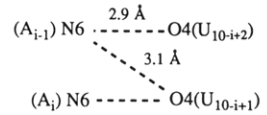

FIGURE 6: (A, top) Stereoview of "right-handed B-DNA junction model" for $d(GA_4U_4C)_2$ as determined from quantitative NOESY simulation. In this view the structure is projected along the long axis of the molecule. Propeller-twisted A·U pairs are clearly seen in the diagram. (B, bottom) Two stacking arrangements in $d(GA_4U_4C)_2$ are shown: one for the sequence $\binom{U6\text{-}A5}{U7\text{-}A4}$ inside the A/U tract, in two mutually perpendicular projections, and the other for the sequence $\binom{A5\text{-}U6}{U6\text{-}A5}$ at the junction, also in two mutually perpendicular projections. Note that inside the A/U tract a bifurcated hydrogen bond is formed between O4(U6) and N6(A4) in the $\binom{U6\text{-}A5}{U7\text{-}A4}$ stack. Such bifurcated hydrogen bonds are also presented between O4(U7) and N6(A3), and between O4(U8) and N6(A2). Inside the A/U tract there is a substantial geometric overlap between the purine rings, but pyrimidines show little geometric overlap. The average separation between two successive intrastrand base planes in the A/U tract is 3.3 Å. At the A5-U6 junction, there is an overlap between purine and pyrimidine rings. The two strands show identical geometric overlaps in the A5-U6 sequence. Note that average intrastrand base-plane separations between A5 and U6 are 3.6 Å.

to avoid stereochemically unacceptable contacts in the neighboring base atoms in the junction. Figure 6A shows the stereopairs of the final model for $d(GA_4U_4C)_2$ in solution, and Figure 6B, the stacking arrangements inside the $A_n/T_n$ tract and at the A5-A6 junction.

## Discussion

The result of the quantitative NOESY simulation, together with that of the 1D NOE experiment, indicates that the gross structure of $d(GA_4U_4C)_2$ is quite similar to that of $d(GA_4T_4C)_2$ (Sarma et al., 1988). Both belong to a B-DNA family with C2′-endo,anti conformation. Like A·T pairs in $d(GA_4T_4C)_2$, A·U pairs in $d(GA_4U_4C)_2$ have large propeller twist (~19°). This results in two unique structural features in the $A_n/U_n$ tract (as in the $A_n/T_n$ tract):

(i) A series of interstrand bifurcated hydrogen bonds between adenine N6 and thymine O4 as shown:




(ii) A narrow minor groove with the shortest P–P distances in the range 3.6–4.0 Å. This feature is consistent with the observation of weak NOESY cross peaks at $\tau_m$ = 100 ms between H2(A5)–H1′(U7) and H2(A4)–H1′(U8) across the strand.

When the stacking patterns of successive A·T pairs in $d(GA_4T_4C)_2$ and successive A·U pairs in $d(GA_4U_4C)_2$ are compared, we note a subtle difference. The vertical separation of the successive A·U pairs along the long axis is 3.2 Å while that of the A·T pairs was 3.3 Å in $d(GA_4T_4C)_2$. A more compact stacking of A·U pairs in $d(GA_4U_4C)_2$ was required to seek agreement with respect to the observed NOESY peaks for H1′($i$ − 1)–H8($i$) in adenine residues. As observed in Figure 6B, the resulting arrangement provides a substantial geometric overlap between the purine rings. It is possible that the absence of bulky methyl groups in $d(GA_4U_4C)_2$ allows successive A·U pairs to have more compact vertical stacking compared to corresponding A·T pairs in $d(GA_4T_4C)_2$.

Such compact structure accounts very well for the unusually large isotropic correlation time, 6.0 ns, compared with 4.0 ns for both $d(GA_4T_4C)_2$ and $d(GT_4A_4C)_2$. It also agrees with the large local spin diffusion found at long mixing times. The MINSY experiment conducted in this case was quite effective in blocking the diffusion pathways via H2′/H2″, such as H8/H6–H2′/H2″–H1′/H3′, thereby showing that H8/H6–H1′/H3′ NOEs are by spin diffusion effects.

Like the $A_4/T_4$ tract, the long axis of the $A_4/U_4$ tract remains essentially straight. However, the A5-U6 junction where two $A_4/U_4$ blocks are joined together creates a structural discontinuity. That is, because of the difference in orientation of adenine and uracil bases, two $A_4/U_4$ blocks cannot be joined together with the kind of conformational adjustment as described under Results, thereby increasing rise ($D_z$) from 3.2 to 3.6 Å between A5 and U6. The overall effect is to have two very compact and tightly stacked $d(GA_4)$·$d(U_4C)$ blocks joined together at a comparatively loose, flexible junction. As in the case of $d(GA_4T_4C)_2$ (Sarma et al., 1988), the present NMR data are interpreted in terms of straight tract as opposed to smoothly curved A/T(U) tracts. As catagorically mentioned in our previous paper (Sarma et al., 1988), the solution NMR data cannot discriminate two structural models for $d(GA_4T_4C)_2/d(GA_4U_4C)_2$, i.e., a smoothly curved and a junction model [for details refer to the supplementary material of Sarma et al. (1988)]. However, it may be pointed out that single-crystal data show that A/T tracts are straight and not intrinsically bent (Nelson et al., 1987; Coll et al., 1987; DiGabriele et al., 1989).

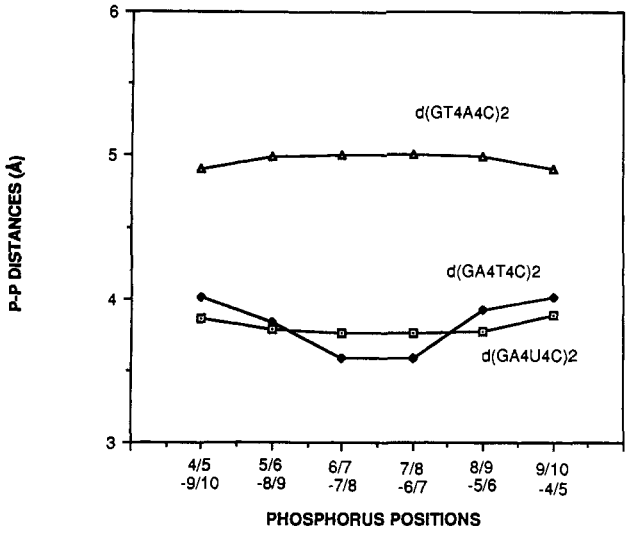



FIGURE 7: The P–P distance, decreased by 5.8 Å to approximate two van der Waals phosphate group radii, plotted for $d(GA_4U_4C)_2$, $d(GA_4T_4C)_2$, and $d(GA_4U_4C)_2$. Two successive nucleoside residues (e.g., 4/5) identify each phosphorus that connects them. Note that the width of the minor groove in $d(GT_4A_4C)_2$ is ~1 Å larger than those for $d(GA_4U_4C)_2$ and $d(GA_4T_4C)_2$. Also note that inside the A/T tract there is a gradual narrowing of the minor groove in $d(GA_4T_4C)_2$ in the 5′- to 3′-direction, while for $d(GT_4A_4C)_2$ the width of the minor groove is uniform all through the chain length.

The P–P distances across the minor groove are plotted in Figure 7 for the three oligomers studied. They reveal two important features: (i) The width of the minor groove in $d(GT_4A_4C)_2$ is ~1 Å larger than those for $d(GA_4U_4C)_2$ and $d(GA_4T_4C)_2$. (ii) Even inside the A/T tract, there is a gradual narrowing of the minor groove in $d(GA_4T_4C)_2$ in a 5′- to 3′-direction, while for $d(GT_4A_4C)_2$, the width of the minor groove is uniform all through the chain length. This feature in our model is consistent with the results of hydroxyl radical cleavage experiments reported by Tullius and Burkhoff (1988). The cleavage patterns observed for poly $d(GA_4T_4C)_2$ and poly $d(GT_4A_4C)_2$ are distinctly different: for poly $d(GA_4T_4C)_2$, the sinusoidal cutting pattern is consistent with the gradual narrowing of the minor groove in the 5′- to 3′- direction, while for poly $d(GT_4A_4C)_2$, the normal cutting pattern suggests a uniform minor groove width all through the chain length including the A·T and G·C regions.

In all the three decamer duplexes we examined, $d(GA_4U_4C)_2$, $d(GA_4T_4C)_2$, and $d(GT_4A_4C)_2$, we consistently found a large propeller twist of A·T or A·U base pairs and resulting close stacking of bases and small minor grooves. Similar features were also pointed out from crystallographic studies of DNA oligomers (Nelson et al., 1987; Coll et al., 1987; DiGabriele et al., 1989). It now seems clear that $A_n/T_n$ has a quite rigid and straight structure, so that bending is possible only at junctions with other nucleotides.

## ACKNOWLEDGMENTS

K.U. thanks the International Christian University, Tokyo, Japan, for granting a sabbatical leave. We thank Dr. Manju Bansal for providing us with the parameters listed in Table II.

## REFERENCES

Bhattacharya, D., & Bansal, M. (1988) *J. Biomol. Struct. Dyn. 6*, 93–104.

Broido, M. S., James, T. L., Zon, G., & Keepers, J. W. (1985) *Eur. J. Biochem. 150*, 117–128.

Chen, J.-H., Seeman, N. C., & Kallenbach, N. R. (1988) *Nucleic Acids Res. 16*, 6803–6812.

Coll, M., Frederick, C. A. Wang, A. H.-J., & Rich, A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8385–8389.

DiGabriele, A. D., Sanderson, M. R., & Steitz, T. A. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 1816–1820.

Gupta, G., Sarma, M. H., & Sarma, R. H. (1988) *Biochemistry 27*, 7909–7919.

Hagerman, P. J. (1986) *Nature (London) 321*, 1149–1150.

Jernigan, R. L., Sarai, A., Ting, K.-L., & Nussinov, R. (1986) *J. Biomol. Struct. Dyn. 4*, 41–48.

Keeper, J. W., & James, T. L. (1984) *J. Magn. Reson. 57*, 404–429.

Koo, H.-S., & Crothers, D. M. (1987) *Biochemistry 26*, 3745–3748.

Massefski, W., Jr., & Redfield, A. G. (1988) *J. Magn. Reson. 78*, 150–155.

Matteuci, M. D., & Caruthers, M. H. (1981) *J. Am. Chem. Soc. 103*, 3185–3191.

Nelson, H. C. M., Finch, J. T., Luisi, B. F., & Klug, A. (1987) *Nature (London) 330*, 221–226.

Olson, W. K., Srinivasan, A. R., Cueto, M. A., Torres, R., Maroun, R. C., Cicariello, J., & Nauss, J. L. (1986) *Biomolecular Stereodynamics IV* (Sarma, R. H., & Sarma M. H., Eds.) pp 75–100, Adenine Press, New York.

Sarma, M. H., Gupta, G., & Sarma, R. H. (1988) *Biochemistry 27*, 3423–3432.

Sarma, M. H., Gupta, G., Garcia, A. E., Umemoto, K., & Sarma, R. H. (1990) *Biochemistry* (following paper in this issue).

Tullius, T. D., & Burkhoff, A. M. (1988) *Structure & Expression* (Olson, W. K., Sarma, M. H., Sarma, R. H., & Sundaralingam, M., Eds.) Vol. 3, pp 77–85, Adenine Press, New York.