Patel, D. J., & Shapiro, L. (1985) *Biochimie 67*, 887.
Patel, D. J., Shapiro, L., & Hare D. (1986) *Biopolymers 25*, 693.
Prive, G. G., Heinemann, U., Chandrasekaran, S., Kan, L.-S., Kopka, M. L., & Dickerson, R. E. (1987) *Science (Washington, D.C.) 238*, 437.
Sarma, M. H., Dhingra, M. M., Gupta, G., & Sarma, R. H. (1985) *Biochem. Biophys. Res. Commun. 131*, 269.
Shakked, Z., Rabinovich, D., Cruse, W. B. T., Salisbury, S. A., & Viswamitra, M. A. (1983) *J. Mol. Biol. 166*, 183.
Siedman, K., & Maciel, G. E. (1977) *J. Am. Chem. Soc. 99*, 3254.
Sklenar, V., & Bax, A. (1987) *J. Am. Chem. Soc. 109*, 2221.
Stone, M., Winkle, S. A., & Borer, P. N. (1986) *J. Biomol. Struct. Dyn. 3*, 767.
Suzuki, E., Pattabiraman, N., Zon, G., & James, T. L. (1986) *Biochemistry 25*, 6854.
Ughetto, G., Wang, A. H.-J., Quigley, G. J., van der Marel, G. A., van Boom, J. H., & Rich, A. (1985) *Nucleic Acids Res. 13*, 2305.
Wang, A. H.-J., Fujii, S., van Boom, J. H., van der Marel, G. A., van Boeckel, S. A. A., & Rich, A. (1982) *Nature (London) 299*, 601.
Winstein, S., Carter, P., Anet, F. A. L., & Bourn, A. J. R. (1965) *J. Am. Chem. Soc. 87*, 5247.
Zanatta, N., Borer, P. N., & Levy, G. C. (1987) *Recent Advances in Organic NMR Spectroscopy* (Lambert, J. B., & Rittner, R., Eds.) p 89, Norell, Landsville, NJ.
Zhou, N., Bianucci, A. M., Pattabiraman, N., & James, T. L. (1987) *Biochemistry 26*, 7905.

On the Question of DNA Bending: Two-Dimensional NMR Studies on $d(GTTTTAAAAC)_2$ in Solution†

Goutam Gupta, M. H. Sarma, and R. H. Sarma*
Institute of Biomolecular Stereodynamics, State University of New York, Albany, New York 12222
Received March 23, 1988; Revised Manuscript Received June 10, 1988

ABSTRACT: It is very well documented that the presence of an $A_n \cdot T_n$ tract causes intrinsic DNA bending. Hagerman demonstrated that the sequence in which the $A_n \cdot T_n$ tracts are joined plays a very crucial role in determining DNA bending. For example, Hagerman showed that the polymer with a repeat of $d(GA_4T_4C)_{n\geq 10}$ is bent but the polymer with a repeat of $d(GT_4A_4C)_{n\geq 10}$ is not bent [Hagerman, P. J. (1986) *Nature (London) 326*, 720–722]. Earlier we have shown that the decamer repeat $d(GA_4T_4C)_2$ is itself bent with a finite structural discontinuity at the A→T sequence [Sarma, M. H., Gupta, G., & Sarma, R. H. (1988) *Biochemistry 27*, 3423–3432]. In the present article, we summarize our studies on the decamer repeat $d(GT_4A_4C)_2$ structure in solution. By employment of 1D and 2D ^1H NMR studies at 500 MHz a complete sequential assignment has been made for the exchangeable and nonexchangeable protons belonging to the ten nucleotides. NOESY data were collected for $d(GT_4A_4C)_2$ at 17 °C in D_2O for three mixing times, 150, 100, and 50 ms. A quantitative NOESY simulation technique was employed to arrive at a structural model of $d(GT_4A_4C)_2$ in solution. Our detailed analyses revealed the following structural features: (i) The duplex adopts the gross morphology of a B-DNA. (ii) All the A·T pairs are propeller twisted ($\leq -15°$). (iii) Although both A and T nucleotides belong to the C2′-endo,anticonformational domain, there is a mild variation in the actual conformation of the A and T residues. (iv) Even though there is a subtle conformational difference in the A and T nucleotides, two structural frames of $T_4 \cdot A_4$ segments are joined at the T→A sequence in such a way that there is no finite discontinuity at the junction; i.e., two neighboring frames exactly coincide at the T→A junction. Thus, our studies on $d(GA_4T_4C)_2$ (Sarma et al., 1988) and on $d(GT_4A_4C)_2$ (this article) reveal the structural peculiarity of the $A_n \cdot T_n$ tract and the effect of A→T/T→A sequence in causing DNA bending.

Ever since the discovery of the bent kinetoplast DNA (Marini et al., 1982), a lot of research work has been devoted to understanding the structural features of DNA bending. It has been clearly demonstrated both by theory and by experiments that the presence of an $A_n \cdot T_n$ tract can cause DNA bending in solution (Levene & Crothers, 1983; Hagerman, 1984; Trifonov, 1985; Kitchin et al., 1986; Zahn & Blattner, 1987). The presence of an $A_n \cdot T_n$ tract in bent DNA suggests one of two possibilities: either (i) the $A_n \cdot T_n$ tract is intrinsically bent, or (ii) the $A_n \cdot T_n$ tract is structurally distinct from the neighboring sequence and the joining of two structurally dissimilar duplexes causes a junction and, hence, the DNA bending. In the literature there is considerable disagreement regarding the probable cause of bending in solution for DNA duplexes with $A_n \cdot T_n$ tracts. Trifonov and co-workers (Trifonov, 1985; Ulanovsky & Trifonov, 1987) show from theoretical and solution studies that the $A_n \cdot T_n$ tract is smoothly curved due to "AA wedge" formation and, thus, the DNAs containing them are bent while Levene and Crothers (1983) suggest that the junction between the $A_n \cdot T_n$ tract and other sequences is the probable cause of bending. Recently, crystal structures of two dodecamer DNA duplexes with internal $A_n \cdot T_n$ tracts have been reported. The structure of the dodecamer duplex

† This research is supported by a grant from the National Institutes of Health (GM29787) and by a contract from the National Foundation for 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.

0006-2960/88/0427-7909$01.50/0

d(CGCA$_6$GCG)·d(CGCT$_6$GCG) solved at 2.9-Å resolution reveals that the A$_6$·T$_6$ tract is essentially straight and the neighboring d(CGC) and d(GCG) sequences create junctions and may probably cause overall bending in the molecule (Nelson et al., 1987). The structure of the dodecamer d-(CGCA$_3$T$_3$GCG)$_2$ solved at 2.2-Å resolution show that the d(A$_3$T$_3$) sequence is straight and the neighboring d(CGC) and d(GCG) sequences create junctions and might result in overall bending of the molecule. The crystal data support the hypothesis of the "junction model"; i.e., the junction of A$_n$·T$_n$ and other sequences can cause bending. However, neither the solution data nor the crystal data, reported so far, can explain stereochemical details of the following observation. Hagerman (1986) observed that the polymer d(GA$_4$T$_4$C)$_{n\geq 10}$ behaves like a bent DNA while the polymer d(GT$_4$A$_4$C)$_{n\geq 10}$ behaves like a normal straight DNA. The replacement of terminal G (or C) has no effect on this observation which would indicate not only the presence of an A$_n$·T$_n$ tract but also how they are placed in sequence plays an important role in DNA bending; i.e., in one polymer [d(GA$_4$T$_4$C)$_{n\geq 10}$] the A$_4$·T$_4$ tracts are joined by an A→T sequence and the polymer is bent while in the other polymer [d(GT$_4$A$_4$C)$_{n\geq 10}$] the A$_4$·T$_4$ tracts are joined by a T→A sequence and the polymer is straight. The terminal G (or C) being of very little consequence in this phenomenon, the junctions between A$_n$·T$_n$ and other sequences (i.e., G→A, G→T, T→C, T→G, G→T, C→T, A→C, and A→G) should have little or no bearing on this observation. Ulanovsky and Trifonov (1987) explained the observation of Hagerman (1986) on the basis of a smoothly curved stretch of A$_4$·T$_4$—henceforth called the "curved DNA model". They argued that when two curved A$_4$·T$_4$ pieces are joined in the polymer d-(GA$_4$T$_4$C)$_{n\geq 10}$, the vector sum of curvature is additive resulting in overall bending in the polymer while in the polymer d-(GT$_4$A$_4$C)$_{n\geq 10}$ two curved A$_4$·T$_4$ pieces are so joined that two individual curvatures cancel each other and the polymer is straight as a result. We were interested in understanding the structural subtleties in the decamer repeats, i.e., d(GA$_4$T$_4$C)$_2$ and d(GT$_4$A$_4$C)$_2$, which finally give rise to the macroscopic properties of the corresponding polymers. We have shown from 2D NMR studies that the decamer duplex d(GA$_4$T$_4$C)$_2$ forms a bent DNA in solution (Sarma et al., 1988); in this model two structurally identical A$_4$·T$_4$ blocks are joined at the A→T junction in such a way that two neighboring structural frames do not coincide in space but make a finite angle to give an overall bending of about 10° in the decamer. In the present article, we demonstrate from 2D NMR studies in solution that in the decamer duplex d(GT$_4$A$_4$C)$_2$ two structurally identical A$_4$·T$_4$ blocks are joined at the T→A junction so that two neighboring frames of reference exactly coincide in space and the decamer is straight as a result. The combined results from our 2D NMR studies on d(GA$_4$T$_4$C)$_2$ and d(GT$_4$A$_4$C)$_2$ reveal unique structural properties of the A$_n$·T$_n$ tract in solution and also the role of A→T (or T→A) between two A$_n$·T$_n$ blocks in causing DNA bending.

Materials and Methods

DNA Synthesis and Purification. The decamer d-(GT$_4$A$_4$C)$_2$ was synthesized on a DNA synthesizer (Applied Biosystems Model 380A) 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 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$_2$O and D$_2$O 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(GT$_4$A$_4$C)$_2$ in H$_2$O were recorded at temperatures of 5–30 °C, with a time-shared long pulse sequence. NOE difference spectra of d(GA$_4$T$_4$C)$_2$ at 17 °C in H$_2$O were recorded for presaturation times τ_m = 100 and 50 ms for the imino protons with a relaxation delay RD = 1.0 s and 5000 transients (NS).

The COSY spectrum of d(GT$_4$A$_4$C)$_2$ at 17 °C in D$_2$O was recorded with the pulse sequence [RD–90°–t_1–90°–Acq]$_{NS}$ in the pure absorption mode (States et al., 1982) with RD = 1.2 s and a data matrix (t_2 = 1024 × t_1 = 512) for NS = 64. NOESY spectra at τ_m = 150, 100, and 50 ms were collected under the same conditions but with the pulse sequence [RD–90°–t_1–90°–τ_m–90°–Acq]$_{NS}$. The HDO signal was presaturated in COSY and NOESY experiments.

Quantitative Structure Determination. Quantitative structure determination involves the following steps:

(Step I) Determination of a Qualitative Structure. Hydrogen-bonding pattern in the oligomer is derived from the 1D NMR data of the oligomer in H$_2$O. Nucleotide geometry (i.e., C2′-endo,anti/C3′-endo,anti) of the constituent residues is obtained from the NOE data in D$_2$O. This information is used to model the oligomer structure. For example, in the case of a duplex, the structural constraints used are (1) H-bonding, viz., Watson–Crick G·C/A·T pairing in a duplex; (2) helical constraints, i.e., a chain progression with n [number of repeating units per turn] and h [height per repeating unit] set at desired values (i.e., n = 10 and h = 3.40 Å for B-DNA and n = 11 and h = 2.56 Å for A-DNA); and (3) estimates of pairwise interproton distances from the primary NOE pattern from data obtained at the lowest possible mixing time.

For computational purpose, constraint 1 is realized by requiring that the C1′---C1′ vector in a base pair (see Figure 1) is about 10.6 Å and the C8(Pur)---C6(Pyr) vector in a base pair is about 9.3 Å and each glycosidic bond (i.e., N9–C1′/N1–C1′) makes angle of about 51° with the C1′---C1′ vector in a pair.

Helical constraints (constraint 2) are satisfied by requiring that three successive atoms of *a repeating unit* (i) in the chain maintain the following relations to *the same three atoms of the* next repeating unit (i + 1) in sequence:

$$X_{i+1} = X_i \cos t - Y_i \sin t$$
$$Y_{i+1} = Y_i \cos t + X_i \sin t \qquad (1)$$
$$Z_{i+1} = Z_i + h$$

where helical twist $t = 2\pi/n$, n = number of residues per turn, and h = height per residue.

Interproton distance constraints (constraint 3) are met by requiring that a given distance estimate (r_o) for a pairwise interaction from the NOE data and the corresponding distance (r_c) for the same proton pair in the structure can only be different by ±0.1 Å; i.e., $|r_o - r_c| \leq 0.1$ Å.

Molecular model building subject to the structural constraints is carried out with the backbone torsion angles (α, β, γ, δ, ϵ, ξ) and glycosyl torsion (χ) and the base parameters (see below) as variables. The torsion angle δ (for sugar pucker) and χ (for glycosyl torsion) are constrained within the conformational domain as revealed from the NOE data; for example

for C3′-endo,anti geometry (A-DNA)

$$70° \leq \delta \leq 100° \qquad 90° \leq \chi \leq 230°$$

for C2′-endo,anti geometry (B-DNA)

$$110° \leq \delta \leq 155° \qquad 230° \leq \chi \leq 270°$$

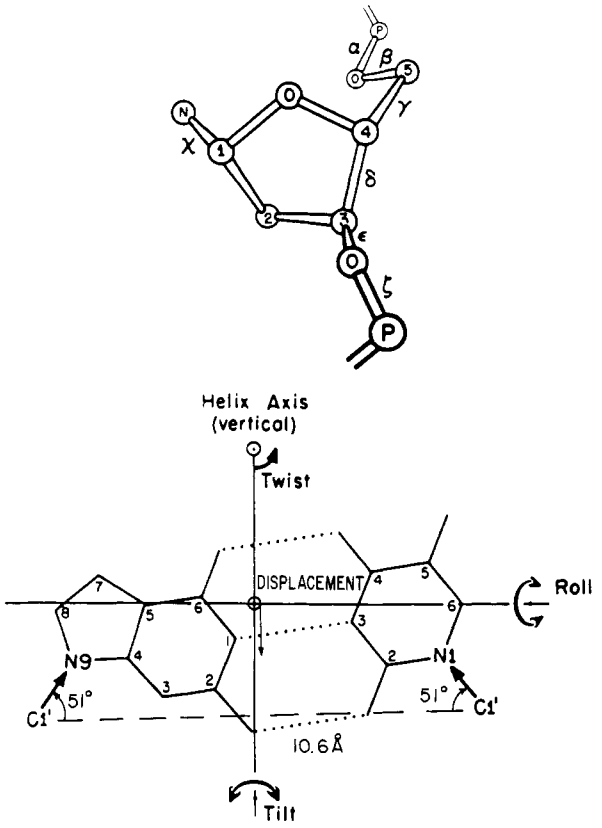


FIGURE 1: Nomenclature for backbone torsion angles and base parameters (Saenger, 1984). The glycosyl torsion for purines is χ-(O1′–C1′–N9–C4) and for pyrimidine it is χ(O1′–C1′–N1–C2).

The variables (Figure 1) are altered in each refinement cycle by a least-squares technique, and after each cycle it is checked whether the calculated structure converged to the desired structure [with all the constraints (1–3)]; if not, the refinement is continued until a convergence is reached.

Thus, at the end of the first step we obtain a stereochemically allowed structure which gives a qualitative agreement with the NOE data, obtained at the lowest possible mixing time.

(Step II) Multispin NOESY Computer Simulation. The qualitative structure in step I is derived to match the estimates of interprotons distances obtained from the primary NOE data for the lowest mixing time (τ_m) at which NOE experiments were conducted. The quantitative structure is required to predict theoretical NOESY intensities for all observed pairwise interactions at all τ_m.

The theoretical NOESY intensity is computed for a proton pair (i,j) from the expression

$$a_{ij} = [\exp(-R\tau_m)]_{ij} \tag{2}$$

where R is the (N × N) relaxation matrix dependent upon the correlation time τ_c and interproton distances r_{ij} [for all i,j with $i \neq j$]—for details of methodology and application, see Keepers and James (1984) and Broido et al. (1985).

Expression 2 also provides the intensity of the diagonal a_{ii} intensity for i's. We define a quantity a_{ij}^c as

$$a_{ij}^c = \frac{a_{ij}}{a_{ii}} \times 100 \tag{3}$$

for all i's.

The quantity a_{ij}^c is compared with the corresponding relative NOESY intensity as obtained from the 1D NOESY slice through the ith proton, i.e.

$$a_{ij}^o = \frac{\text{peak height at the } j\text{th NOE site}}{\text{peak height of the } i\text{th proton site}} \times 100 \tag{4}$$

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

$$R = \sum_{i,j \neq i} \frac{|a_{ij}^o - a_{ij}^c|}{a_{ij}^o} \tag{5}$$

for all i's at each τ_m. If for a given mode R (for each τ_m) is less than or equal to 0.15, the corresponding model is a quantitatively acceptable structure of the molecule. The R factor defined above has the same meaning as the R factor used by crystallographers (see later). The magnitude 0.15 for R as a reliability index for a close approximate of true structure is also taken from crystallography.

For computations of theoretical intensities (expressions 3 and 5 for a given model, an isotropic overall correlation time τ_c is chosen for all pairwise interactions. An estimate of τ_c is obtained by monitoring NOEs with decreasing τ_m for a proton pair separated by a fixed distance [viz., H5···H6 of C (2.5 Å), H2′···H2″ of sugar (1.8 Å), and H1′···H2″ of sugar (2.4 Å irrespective of any sugar pucker)]. In our computational procedure, the value of τ_c can also be treated as a variable and any desired value of τ_c can be introduced for any pairwise interaction.

The qualitative structure obtained after step I is used as a starting model for step II. In view of the fact that the qualitative structure matched the primary NOE data, it should be in the neighborhood of the final refined structure which gives the lowest agreement index R for all τ_m. Thus, in order to obtain the "final refined structure", the interproton distances in the "starting structure" are altered by use of a "grid-search technique" to obtain the lowest possible R. In the refinement analysis (i.e., minimization of R), the backbone torsion angles (i.e., α, β, γ, δ, ϵ, ξ), glycosyl torsion (χ), and base parameters (tilt, twist, base displacement, etc.) are used as variables in a least-squares refinement technique; the refinable variables are elastically bound by a high weight factor such that they remain close to their starting values.

On the Meaning of R Factor. In the quantitative procedure, the R factor is used in the same meaning as that which crystallographers use to define the agreement index R for a crystal structure. In making a quantitative judgment on the plausibility of a structure with the R factor as a deciding criterion, one must make sure that the number of data points (ND) outnumbers the number of variables (NV) used for structure generation. Let us take a self-complementary decamer duplex [viz., $d(GT_4A_4C)_2$ or $d(GA_4T_4C)_2$] for the purpose of illustration.

There are ten different nucleotides and each nucleotide can be defined by six backbone torsion angles (α, β, γ, δ, ϵ, ξ) and a glycosyl torsion (χ). Hence, for ten residues we need 10 × 7 = 70 variables. To define the location of the base pairs with respect to the backbone, we need there base parameters, i.e., tilt, twist, and base displacement. For a decamer with a center of symmetry, this could vary from a minimum of 3 to a maximum of 15. Thus, in order to have a complete description of a self-complementary decamer duplex, we need NV = 73–85 refinable variables.

The following is a list of possible NOE sites for a given τ_m from a nucleotide stacked in a B-DNA geometry:

NOE from	NOE sites
H8/H6(i)	H1′(i), H2′(i), H2″(i), H3′(i), H5/CH_3(i), H1′(i − 1), H2′(i − 1), H2″(i − 1), H5/CH_3(i + 1) (about 9)
NH(G/T_i)	H2(A_i), H2(A_{i+1}) (about 2)
H2(A_i)	H2(A_{i-1}), H2(A_{i+1}) (about 2)
H1′(i)	H2′(i), H2″(i), H8/H6(i), H8/H6(i + 1) (about 4)
H2′(i)	H2″(i), H1′(i), H8/H6(i), H8/H6(i + 1) (about 4)
H2″(i)	H2′(i), H1′(i), H8/H6(i), H8/H6(i + 1) (about 4)
H3′(i)	H2′(i), H2″(i), H8/H6(i) (about 3)
H5/CH_3(i)	H6(i), H8/H6(i + 1) (about 2)

Therefore, for a τ_m from each nucleotide we can have a maximum of about 30 NOESY intensities. However, in practice overlap of H2′ and H2″ of the same and different residues and overlap of H8/H6, H1′, and H3′ of different residues limit the number of observed (distinct) NOESY intensities for a residue to about 20. Remember that the absence of NOESY intensity is treated as a zero value in the intensity [viz., H8/H6(i)---H3′(i) in a C2′-endo,anti geometry]. In summary, for a decamer duplex for a τ_m we can expect about 10 × 20 = 200 (ND) nonoverlapping data points, which outnumber the number of refinable parameters (NV = 73–85).

Results and Discussion

Assignment of the NH Protons. The temperature dependence of the imino protons (NH) of G·C and A·T Watson–Crick pairs was monitored to study the duplex nature of the decamer d(GT_4A_4C)$_2$ in H_2O at 100 mM NaCl (pH 7.0). It was observed that the decamer remained in the duplex state up to 30 °C. We chose 17 °C for performing the 1D NOE and 2D NOESY/COSY experiments because at this temperature both exchangeable NH (in H_2O) and nonexchangeable base protons H8/H6, H2/H5/CH_3 (in D_2O) were best resolved. Figure 2 shows the low-field region of the 500-MHz ^{1}H NMR spectrum of d(GT_4A_4C)$_2$ in H_2O at 17 °C. Notice four signals from N3–H (of A·T) and one from N1–H (of G·C)—as expected from a self-complementary double helix (see the inset of Figure 2). The N1–H (G1) signal at 12.80 ppm from the terminal G1·C10 pair was most sensitive to the temperature change. We performed 1D NOE experiments with five NH signals as irradiation sites for τ_m = 100 and 50 ms. With N1–H of G1·C10 as a marker and by monitoring the NOEs at τ_m = 100 and 50 ms, the following assignments were obtained from the spectra in Figure 2:

(*i*) *Spectrum B.* From N1–H of G1·C10 at 12.80 ppm for τ_m = 50 ms, an NOE is observed at 14.20 ppm identified as N3–H of T2·A9—the next neighbor in sequence. An NOE is also observed at 8.00 ppm—the NH_2 of C10. When τ_m is raised to 100 ms, an NOE is observed in the base region (H8/H6, H2) at 7.75 ppm which is assigned to H2A of the T2·A9 pair.

(*ii*) *Spectrum C.* With N3–H of T2·A9 at 14.20 ppm for τ_m = 50 ms as the site of NOE, there is a strong NOE, as expected, at H2A of T2·A9 at 7.75 ppm. There is also a strong NOE in the NH region at 14.00 ppm—identified as belonging to N3–H of the T3·A8 pair. There is a small but noticeable NOE at 13.85 ppm probably occurring due to a second-order NOE, i.e., N3–H(T2·A9)$\overset{3.4\text{ Å}}{\cdots}$N3–H(T3·A8)$\overset{3.4\text{ Å}}{\cdots}$N3–H(T4·A7). Hence the signal at 13.85 ppm should belong to N3–H of T4·A7.

(*iii*) *Spectrum D.* From N3–H of T3·A8 at 14.00 ppm for τ_m = 50 ms, there are two NOE sites in the NH region, one at 13.85 ppm beloning to N3–H of T4·A7 and the other to N3–H of T2·A9 at 14.20 ppm. There are two NOE sites in the H2A region, the stronger one at 7.11 ppm beloning to H2A of T3·A8 and the weaker one at 7.75 ppm, i.e., H2A of the T2·A9 pair.

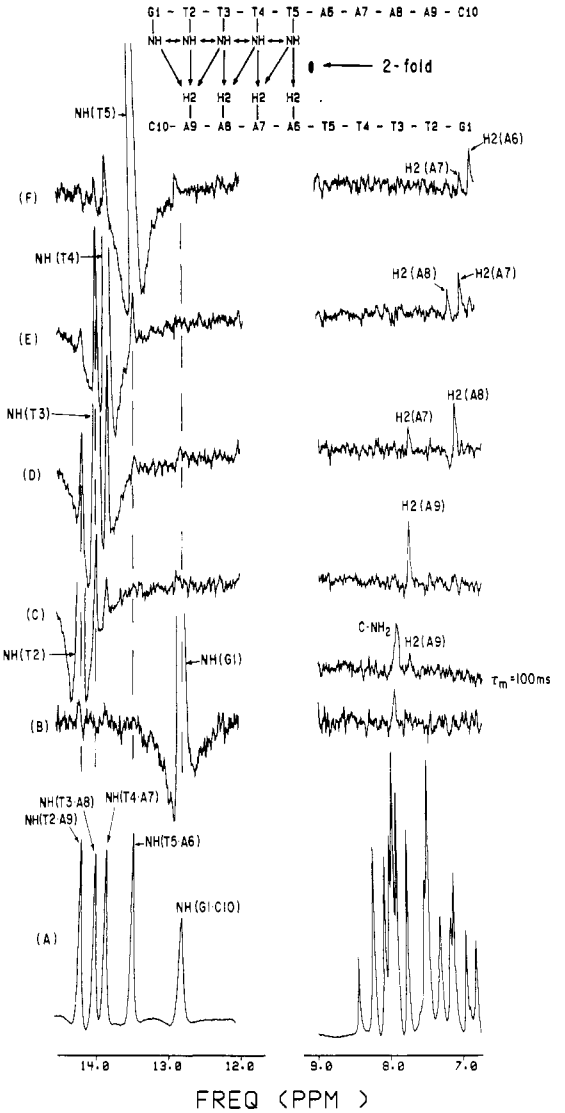


FIGURE 2: 1D NMR spectra of d(GT_4A_4C)$_2$ in H_2O at 17 °C. The spectra were recorded with a time-shared long-pulse sequence. DNA concentration was 2 mM in duplex in 10 mM sodium phosphate buffer, 1 mM EDTA, and 100 mM NaCl, pH 7.0. (A) Control spectrum at 17 °C showing the exchangeable NH signals and the signals in the base region H8/H6/H2 etc. (B) 1D NOE spectrum in which N1–H(G1·C10) is irradiated for τ_m = 50 ms; N3–H(T2·A9) appears as one site of NOE at 14.20 ppm. NOE is also seen at 8.00 ppm from the NH_2 of C10. Whtn τ_m is raised to 100 ms, the NOE shows up at H2A(T2·A9) at 7.75 ppm. (C) 1D NOE spectrum for τ_m = 50 ms in which N3–H(T2·A9) is irradiated. NH(T3·A8) at 14.00 ppm and H2A(T2·A8) at 7.75 ppm are the strong sites of NOE. There is also a noticeable NOE at 13.85 ppm (N3–H of T4·A7) probably occurring as N3–H(T2·A9)$\overset{3.4\text{ Å}}{\cdots}$N3–H(T3·A8)$\overset{3.4\text{ Å}}{\cdots}$N3–H(T4·A7). (D) 1D NOE spectrum for τ_m = 50 ms in which N3–H(T3·A8) is irradiated. N3–H(T4·A7) at 13.85 ppm, N3–H(T2·A9) at 14.20 ppm, H2A(T2·A9) at 7.75 ppm, and H2A(T3·A8) at 7.11 ppm are the sites of NOE. (E) 1D NOE spectrum for τ_m = 50 ms in which N3–H(T4·A7) is irradiated. N3–H(T5·A6) at 13.48 ppm, N3–H(T3·A8) at 14.00 ppm, H2A(T4·A7) at 6.94 ppm, and H2A(T3·A8) at 7.11 ppm are the NOE sites. (F) 1D NOE spectrum for τ_m = 50 ms in which N3–H(T5·A6) is irradiated. N3–H(T4·A7) at 13.85 ppm, H2A(T4·A7) at 6.94 ppm, and H2A(T5·A6) at 6.80 ppm are the prominent sites of NOE. A residual NOE at N3–H(T3·A8) at 14.00 ppm originated as N3–H(T5·A6)$\overset{3.4\text{ Å}}{\cdots}$N3–H(T4·A7)$\overset{3.4\text{ Å}}{\cdots}$N3–H(T3·A8). The inset gives a summary of the primary NOE pattern. Note that, *N3–H of the ith T·A pair shows NOE at H2A of the (i − 1)th T·A pair* but not at H2A of the (i + 1)th T·A pair. For example, N3–H(T3·A8) shows NOE at H2A(T2·A9) but not at H2A(T4·A7). This is a consequence of a propeller twist (≥15°) in the A·T pairs. When propeller-twisted A·T pairs are stacked in the duplex d(GT_4A_4C)$_2$, one expects a shorter distance (about 3.3 Å) for NH(T3·A8) and H2A(T2·A9) and a larger distance (about 4.8 Å) for NH(T3·A8) and H2(T4·A7).

Table I: Chemical Shift Values (ppm) of Various Protons in the $d(GT_4A_4C)_2$ Duplex in Solution at 17 °C with TSP as an Internal Standard

	H8	H2/H5/CH_3	H1′	H2′	H2″	H3′	NH
G1	7.93		5.98	2.76	2.84	4.81	12.80
T2	7.51	1.33	6.24	2.29	2.66	4.95	14.20
T3	7.55	1.62	6.22	2.19	2.66	4.93	14.00
T4	7.53	1.65	6.10	2.15	2.60	4.94	13.85
T5	7.35	1.72	5.57	1.92	2.29	4.87	13.48
A6	8.27	6.80	5.67	2.74	2.81	5.00	
A7	8.11	6.94	5.73	2.58	2.79	5.03	
A8	8.04	7.11	5.86	2.55	2.84	5.03	
A9	8.00	7.75	6.06	2.44	2.81	4.96	
C10	7.20	5.07	5.94	2.11	2.11	4.93	

(*iv*) *Spectrum E.* When N3–H of T4·A7 at 13.85 ppm is the site of irradiation, for τ_m = 50 ms, there are two strong NOEs in the NH region. One at 14.00 ppm already assigned to N3–H of T3·A8 and the other at 13.48 ppm belonging to the remaining N3–H of the T5·A6 pair. There are two noe sites in the H2A region—the stronger one at 6.94 ppm should belong to H2A of T4·A7 while the weaker one is at H2A of T3·A7 at 7.11 ppm.

(*v*) *Spectrum F.* One-dimensional NOE spectrum for τ_m = 50 ms in which N3–H of T5·A6 at 13.48 ppm is irradiated. A strong NOE is observed at one neighbor N3–H of T4·A7 at 13.85 ppm. The other neighbor of T5·A6 is A6·T5 (see the inset of Figure 2) which is 2-fold related to T5·A6, and thus the N3–H of T5·A5 is chemically equivalent to N3–H of the A6·T5 pair. There are two NOE signals in the H2A region: the stronger one at 6.80 ppm beloning to H2A of T5·A6 while the weaker one at 6.94 ppm belonging to H2A of the T4·A7 pair. There is a small residual NOE at 14.00, i.e., N3–H of T3·A8 probably occurring as N3–H(T5·A6)$^{3.4\ \text{Å}}$---N3–H(T4·A7)$^{3.4\ \text{Å}}$---N3–H(T3·A8).

The inset of Figure 2 gives the primary NOE pattern from the NH signals in the duplex $d(GT_4A_4C)_2$ at 17 °C as follows:

(i) N3–H of a T·A pair shows NOEs at the N3–H signals of both the 3′ and 5′ neighbors.

(ii) N3–H of the *i*th T·A pair shows a stronger NOE at H2A of the same pair (distance H2A---N3–H is about 2.8 Å in an A·T pair) and a weaker NOE at H2A of the $(i - 1)$th T·A pair (see the inset of Figure 2)—hence, the distance between N3–H of the *i*th T·A pair and H2A of the $(i - 1)$th T·A pair should be above 3.0 Å.

(iii) NOE is observed from N3–H of the *i*th T·A pair to H2A of the $(i - 1)$th T·A pair but not to H2A of the $(i + 1)$th T·A pair. This is a direct consequence of propeller-twisted (15°) A·T pairs. When A·T pair is propeller twisted the distance between N3–H of the *i*th T·A pair and H2A of the $(i - 1)$th T·A pair could be about 3.3–3.6 Å while the distance between N3–H of the *i*th T·A pair and H2A of the $(i + 1)$th T·A pair can be beyond 4.8 Å. Therefore, the data in Figure 1 result in the sequential assignment of the five NH signals and H2A of four T·A pairs (see Table I). The data also provide experimental proof that the T·A pairs in $d(GT_4A_4C)_2$ in solution are propellar twisted.

Assignment of the Proton System H8/H6, H2/H5/CH_3, H1′, H2′, H2″, H3′. A combination of NOESY data for τ_m = 150, 100, and 50 ms was used to arrive at the sequential assignment of the proton system H8/H6, H2/H5/CH_3, H1′, H2′, H2″, H3′ belonging to the ten nucleotides in $d(GT_4A_4C)_2$ at 17 °C. From an examination of primary NOEs by monitoring NOEs as a function of τ_m's, we were able to conclude that each of the ten nucleotides belongs to average C2′-endo,anti geometry as in a right-handed B-DNA because there were strong primary NOEs from the H8/H6 to the H2′, H2″ region but very weak or almost zero NOEs to the H3′ region [details of methodology outlined in Gupta et al. (1985, 1987) and Sarma et al. (1985, 1986, 1987a,b)]. When nucleotides with C2′-endo,anti geometry are stacked in the gross morphology of a B-DNA duplex, one expects the NOE pattern H1′($i - 1$)$^{3.0–3.8\ \text{Å}}$H8/H6(i)$^{3.6–4.0\ \text{Å}}$H1′(i) in the H8/H6 vs H1′ NOESY cross section. Figure 3A shows the NOESY H8/H6 vs H1′ cross section for τ_m = 100 ms. At this mixing time one observes an NOESY cross section for proton pairs at a distance less than 4.0 Å. To establish the H1′(i – 1) → H8/H6(i) → H1′(i) connectivity route, one needs a marker (i.e., a known H1′ or H8/H6). H6(C10) was located at 7.20 ppm because COSY data (not shown) revealed a cross-peak at (7.20, 5.07) ppm as expected for H6---H5(C10)—only C present in the oligomer. Having found C10 as a marker, the cross-connectivity route was easily established in Figure 3A. Thus, Figure 3A gives the sequential assignment of H8/H6, H1′ of the ten nucleotides. The four CH_3(T) signals were located as strong NOESY cross-peaks from H6(T) in the region of 1.8–1.3 ppm. Figure 3B shows the H1′ vs H2′,H2″ NOESY cross section for τ_m = 100 ms. Note that for each of the ten H1′ there are two NOESY cross-peaks: one corresponding to H1′(i)---H2′(i) and the other due to H1′(i)---H2″(i) except when H2′ and H2″ of a given residue are too close in chemical shift—for example, G1 and C10 (see Table I). The NOESY cross-peaks for τ_m = 100 ms are shown for H2′---H2″ of a given sugar in Figure 3C. Note that except for G1 and C10, the remaining eight H2′,H2″ pairs are sequentially located in Figure 3C. H2′ and H2″ of a sugar were chemically distinguished from the H1′ vs H2′,H2″ COSY cross section (data not shown). In this cross section H1′---H2′ appears as a stronger peak than the H1′---H2″ cross-peak because for a C2′-endo sugar $J_{1'2'} \sim 10$ Hz and $J_{1'2''} \sim 5$ Hz. From the H3′ vs H2′,H2″ cross section, H3′ of the ten nucleotides are sequentially assigned. This results in complete sequential assignment of the proton system H8/H6, H2/H5/CH_3, H1′, H2′, H2″, H3′ belonging to the ten residues of $d(GT_4A_4C)_2$ at 17 °C. See Table I for the complete list of the chemical shift values.

Structural Details. NMR data discussed, so far, indicate that $d(GT_4A_4C)_2$ in solution adopts the gross morphology of a B-DNA duplex with all the ten nucleotides belonging to average C2′-endo,anti geometry. A few more structural details are revealed when one analyzes the following data.

Figure 4 compares the H8/H6 vs H1′ NOESY cross sections for τ_m = 150, 100, and 50 ms. Note that for τ_m = 150 ms, one observes both the intranucleotide H8/H6(i)---H1′(i) cross-peaks for all the ten nucleotides and all the ten internucleotide H8/H6(i)---H1′(i – 1) cross-peaks. On reduction of τ_m to 100 ms, a few cross-peaks begin to weaken in intensity, namely, H1′(T5)---H8(A6), H1′(T6)---H6(T6), and H1′(T4)---H6(T5). For τ_m = 50 ms only four cross-peaks are present in the H8/H6 vs H1′ cross-section; they are H8-

(A7)---H1′(A6), H8(A8)---H1′(A7), H8(A9)---H1′(A8), and the four overlapping cross-peaks H1′(T2)---H6(T3), H1′(T3)---H6(T4), and H1′(T2,T3)---H6(T2,T3). For C2′-endo,anti geometry, the H8(i)---H1′(i) distance is 3.7–4.0 Å. Therefore, the absence of intranucleotide H8---H1′ cross-peaks suggests that, for τ_m = 50 ms, NOESY peaks are seen for a proton pair less than 3.5 Å apart. This suggests that the distance H8(A7)---H1′(A6) is ≤3.5 Å. By similar argument, the absence of cross-peaks H6(T4)---H1′(T4), H6(T5)---H1′(T4), H6(T5)---H1′(T5), and H8(A6)---H1′(T5) suggests that the corresponding distances are beyond 3.5 Å. It may be noted that H1′ of all four T's and C show a weak NOESY cross-peak at 7.53, 6.23 ppm which is present only because four weak peaks are clustered at that position. Hence, we conclude that A's in the $d(GT_4A_4C)_2$ duplex are stacked in such a way that internucleotide H1′($i-1$)---H8(i) distances are ≤3.5 Å while intranucleotide H1′(i)---H8(i) distances are above 3.5 Å. We also note that T's in the duplex are so stacked that both intranucleotide H6(i)---H1′(i) and internucleotide H6(i)---H1′($i-1$) distances are above 3.5 Å. This observation suggests the possibility that in the $d(GT_4A_4C)_2$ duplex the T_4·A_4 tract adopts a structure in which the A chain could be conformationally mildly different from the T chain even though both A and T nucleotides belong to the same C2′-endo,anti conformational domain. Similar structural features were also detected in the A_4·T_4 tract of the $d(GA_4T_4C)_2$ duplex (Sarma et al., 1988).

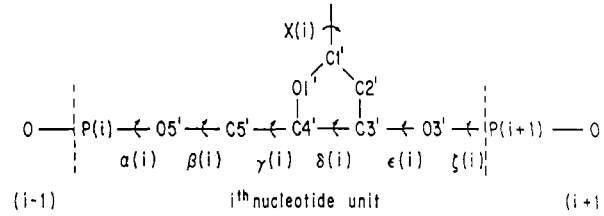
Figure 5 compares the H8/H6 vs H2′,H2″ NOESY cross section for τ_m = 150, 100, and 50 ms. For a B-DNA structure, in these cross sections one expects the following two cross-connectivity routes: H2″($i-1$) $\overset{2.2\text{ Å}}{\cdots}$ H8/H6(i) $\overset{2.2\text{ Å}}{\cdots}$ H2′(i) $\overset{1.8\text{ Å}}{\cdots}$ H2″(i)---... and H2′($i-1$) $\overset{3.7\text{ Å}}{\cdots}$ H8/H6(i) $\overset{2.2\text{ Å}}{\cdots}$ H2′(i) $\overset{1.8\text{ Å}}{\underset{\text{fixed}}{\cdots}}$ H2″(i)---.... Note that, for τ_m = 150 ms, internucleotide NOESY cross-peaks are observed for H2′($i-1$)---H8/H6(i) and H2″($i-1$)---H8/H6(i). On a gradual lowering of the τ_m to 50 ms, only the peaks for H2″($i-1$)---H8/H6(i) prevail, and they form a connectivity route involving H2″($i-1$) → H8/H6(i) → H2′(i) → H2″(i) → ...—a typical signature of a B-DNA conformation. Also, shown in Figure 5 is the connectivity route marked by broken lines involving H8(G1) → CH_3(T2) → H6(T2) → CH_3(T3) → H6(T3) → CH_3(T4) → H6(T4) → CH_3(T5) → H6(T5). Note that CH_3 of T3, T4, and T5 are very close in chemical shift.

For τ_m = 150 ms three small but detectable NOESY cross-peaks were observed (data not shown) for H2(A9)---H2(A8), H2(A8)---H2(A7), and H2(A7)---H2(A6). These cross-peaks were no longer present for τ_m = 100 and 50 ms. This indicated that the distances involving H2(A_i)---H2(A_{i+1}) should be about 4.0 Å. The relaxation behavior of the H2A protons is quite the same for the NOESY experiments at τ_m = 150 and 100 ms because we keep RD fixed at 1.2 s. Therefore, the fact that we observe H2(A_i)---H2(A_{i+1}) at τ_m = 150 ms but not at τ_m = 100 ms implies that the presence or absence of these cross-peaks is governed by the distance effect and not by the relaxation effect.

Table II lists the observed NOESY peak heights for τ_m = 150, 100, and 50 ms due to the base protons H8/H6. Note that for four A's and four T's intranucleotide NOEs H2′(i)---H8/H6(i) are always stronger than the corresponding internucleotide NOEs H2″($i-1$)---H8/H6(i). This pattern of NOE is inconsistent with a conventional fiber B-DNA model of Arnott et al. (1980)—see Sarma et al. (1988) for details. It may also be pointed out that in the A_4·T_4 segment in $d(GT_4A_4C)_2$ and in $d(GA_4T_4C)_2$ (Sarma et al., 1988) neither A nor T adopts C3′-endo sugar puckers; thus, the possibility of a heternomous B-DNA model of Arnott (Arnott et al., 1983) for A_4·T_4 is totally ruled out. It may also be added that contrary to the claim of Arnott et al. (1983), Alexeev et al. (1987) have shown that the fiber model of poly[d(A)]·poly[d(T)] is not heternomous but very close to the solution structure we proposed in which both A and T have C2′-endo pucker (Sarma et al., 1985).

The pattern of NOEs in D_2O involving H8/H6(i)---H2′(i) and H8/H6(i)---H2″($i-1$) and 1D NOEs in H_2O involving N3–H of the ith T·A pair to H2A of the ($i-1$)th T·A pair are consistent with propeller-twisted (≤−15°) T·A pairs.

Structural Refinement through NOESY Simulation. (A) Starting Model. The starting model was generated by a model building procedure in which the backbone torsion angles and glycosyl torsion angles [for nomenclature, see Saenger (1984)] of each of the ten nucleotides were used as variables, i.e.



The model-building procedure utilized the following structural constraints as evident from the NMR data (Figures 2–5, Table II):

(i) Watson–Crick G·C and A·T pairing exists.

(ii) C2′-endo,anti geometry of all the ten nucleotides was present. However, variations within this domain were allowed; i.e., δ varied within 120°–155° and χ within 230°–270°.

(iii) A·T pairs are propeller twisted as explained in Figure 2.

(iv) Nucleotides on two chains are conformationally equivalent because there is no resonance doubling of the proton signals; viz., T2 on one chain has the same conformation as T2 on the other chain and so on.

(v) Although both A and T belong to the same C2′-endo,anti domain, there is a mild conformational difference in the A chain from the T chain (Figure 4). But two structural blocks on either side of T5→A6 are conformationally equivalent; i.e., d(G1-T2-T3-T4-T5)·d(A6-A7-A8-A9-C10) and d(A6-A7-A8-A9-C10)·d(G1-T2-T3-T4-T5) do have the same structures.

(vi) For all the A nucleotides the internucleotide H8(i)---H2″($i-1$) NOE is weaker in intensity than the corresponding intranucleotide H8(i)---H2′(i) NOE. The same trend is observed for H6 of T. The pattern of NOE observed for all H6 of T's and all H8 of A's is very similar indicating that there is no detectable structural discontinuity in the molecule at any nucleotide (Table II). It may be recalled that for the oligomer $d(GA_4T_4C)_2$ in solution a finite discontinuity was detected from the primary NOE data (Sarma et al., 1988).

In the starting B-DNA model, the average torsion angles for A and G were α = 320°, β = 145°, γ = 40°, δ = 140°, ϵ = 225°, ξ = 210°, and χ = 260°. The average torsion angles for T and C were α = 335°, β = 165°, γ = 46°, δ = 120°, ϵ = 182°, ξ = 247°, and χ = 245°. In this model A·T pairs had very little tilt (∼−2°) but had a large propeller twist (15°). In the starting model two structural blocks, i.e., d(G1-T2-T3-T4-T5)·d(A6-A7-A8-A9-C10) and d(A6-A7-A8-A9-C10)·d(G1-T2-T3-T4-T5), were joined in such a way that two neighboring structural frames exactly coincided in space. Our intention was to examine whether refinement against the ob-

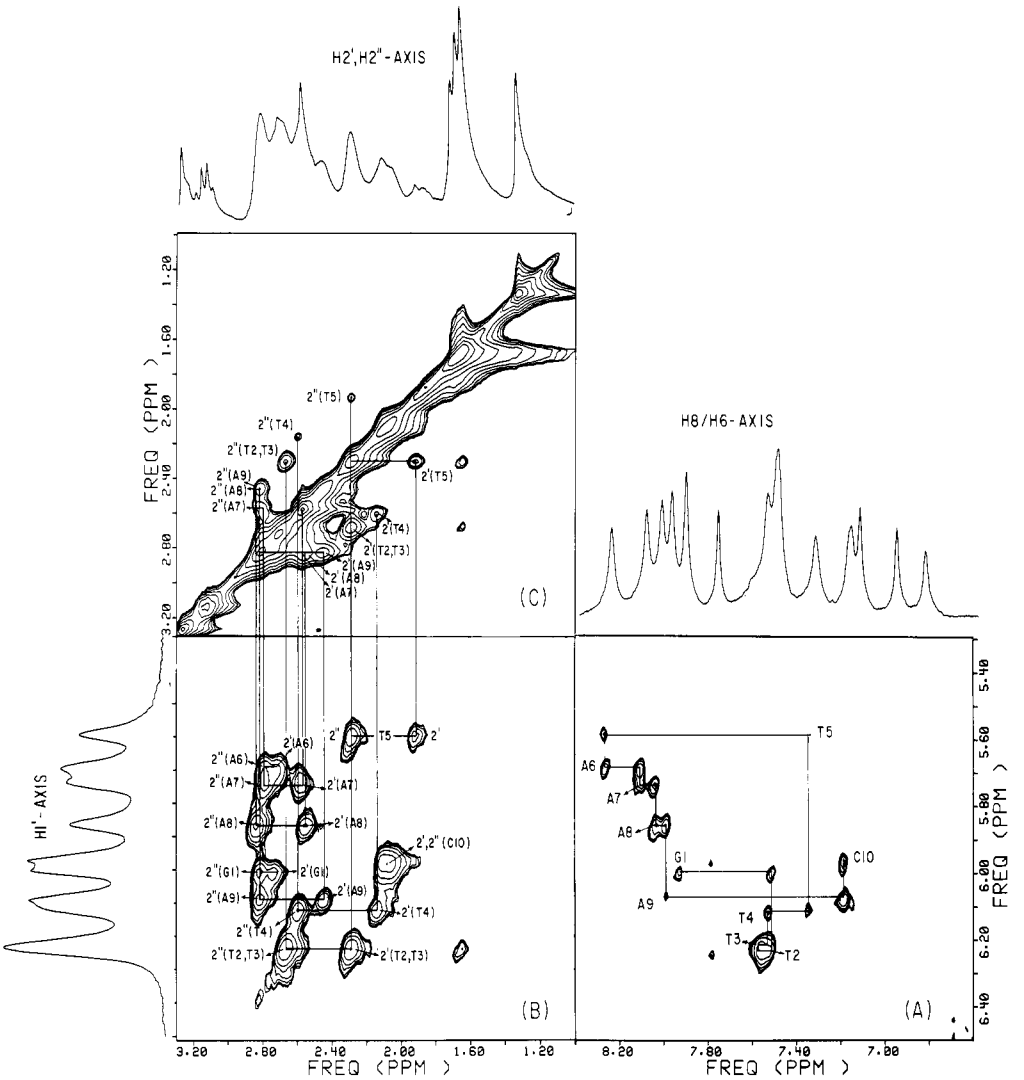


FIGURE 3: Assignment of the spin system H8/H6, H5/CH_3, H1′, H2′, H2″ from the relevant NOESY cross sections of $d(GT_4A_4C)_2$ at 17 °C in D_2O. Other experimental conditions are the same as in Figure 1. (A) Cross-connectivity route involving H8/H6---H1′ NOESY cross-peaks. In a B-DNA structure, the intranucleotide H8/H6(i)---H1′(i) distance is about 3.8 Å while the internucleotide H8/H6(i)---H1′(i − 1) distance is 3.0–4.0 Å. In the oligomer d(G1-T2-T3-T4-T5-A6-A7-A8-A9-C10), H6(C10) is easily identified from the presence of a strong COSY/NOESY cross-peak involving H6(C10)---H5(C10). With H6(C10) as a marker in the connectivity route, the protons H8/H6 and H1′ for the ten nucleotides are assigned from the H8/H6---H1′ cross section. (B) Cross-section H1′---H2′,H2″ shows the NOESY cross-peaks involving H1′---H2′ and H1′---H2″. For nucleotides in a B-DNA geometry (i.e., C2′-endo sugar), H1′---H2″ ~ 2.3 Å and H1′---H2′ ~ 3.1 Å. From this and (A) we arrive at the sequential assignment of H8/H6, H1′, and H2′,H2″. (C) H2′,H2″---H2′,H2″ NOESY cross section to reconfirm the identification of H2′,H2″ of the same sugar. H3's of the ten residues are obtained from the H2′,H2″---H3′ NOESY cross section. CH_3 groups of T are identified from the presence of strong NOESY peaks at 1.3–1.8 ppm from H6 of T. H5(C10) is located at 5.07 ppm as a strong NOESY peak from H6(C10) at 7.20 ppm. See Table I for complete listing of chemical shift values (in ppm) of the spin system H8/H6, H2/H5/CH_3, H1′, H2′, H2″, H3′ of ten nucleotides.

served NOESY data resulted in any structure drastically different from the starting model—for example, one with a bend at the T5→A6 sequence. It may be pointed out that the fact that A and T chains are mildly conformationally different does not necessarily imply that there should be any structural discontinuity when two A_4·T_4 blocks are joined at A→T or T→A sequence. In other words, two such A_4·T_4 blocks (with the A chain being conformationally different from the T chain) could be joined at A→T or T→A sequence with or without any finite discontinuity.

(*B*) *Refined Model.* NOESY intensities for the starting model were computed for τ_m = 150, 100, and 50 ms with expression 2 under Materials and Methods. For two overlapping protons the individual contributions of each protons were added to compare with the observed NOESY data. For NOESY intensity calculation, an isotropic overall τ_c was used; an estimate of τ_c = 4 ns was obtained by monitoring H6(C10)---H5(C10) and H6(T2)---CH_3(T2) NOEs as a function of τ_m. For the starting model the calculated intensities for τ_m = 150, 100, and 50 ms were compared with the corresponding observed NOESY peak heights for the protons H8/H6, H2/H5/CH_3, H1′, H2′, H2″, and H3′. Then the discrepancies in the calculated and observed values were examined for each cross-peak. For a proton pair, if the observed intensity was smaller (larger) than the calculated one, then in the actual structure the corresponding pairwise interproton distance had to be longer (shorter) than the distance in the trial starting structure. Thus, necessary distance adjustments were made in the starting structure, and NOESY intensities were computed for the altered structure and compared with the observed intensities for all τ_m's, i.e., 150, 100, and 50 ms. Such refinement cycles continued until a satisfactory agreement was reached between the calculated and observed intensities for all pairwise interactions involving H8/H6, H2/H5/CH_3, H1′, H2′, H2″, and H3′ for τ_m's = 150, 100, and 50 ms. Table II lists the calculated percent NOE intensities

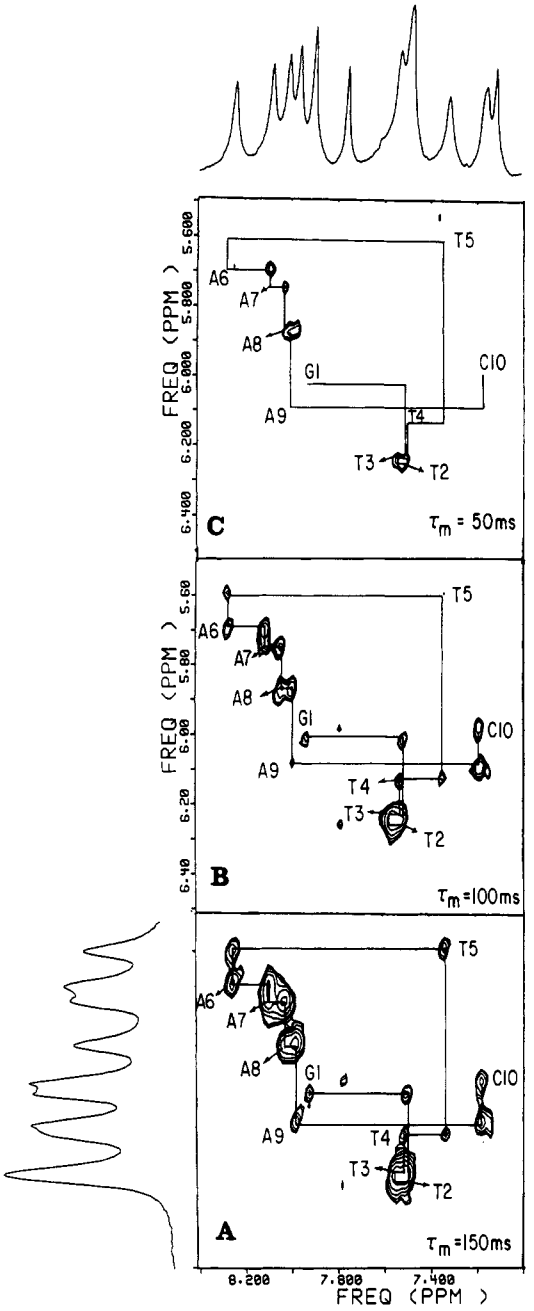


FIGURE 4: H8/H6---H1′ NOESY cross sections for mixing times (τ_m's) of (A) 150, (B) 100, and (C) 50 ms. Note that with reduction of τ_m internucleotide NOEs involving H8/H6(i)---H1′(i − 1) prevail and intranucleotide NOEs involving H8/H6(i)---H1(i) become too weak to be observed in this cross section.

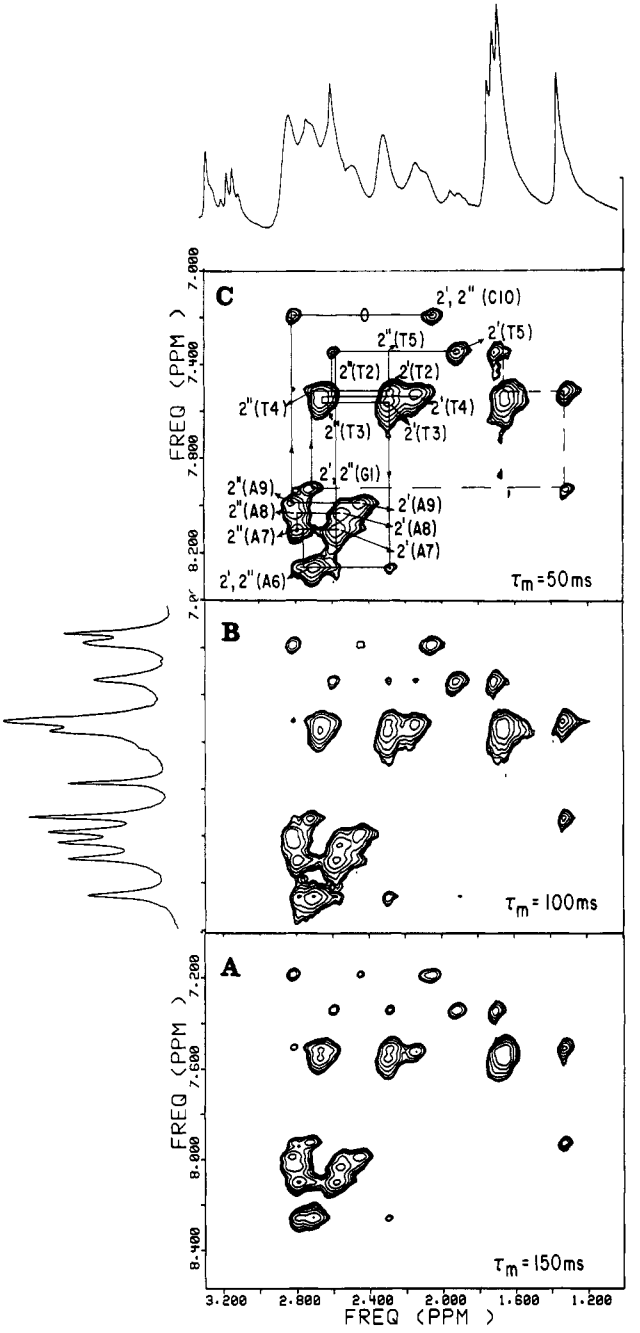


FIGURE 5: H8/H6---H2′,H2″ NOESY cross sections for τ_m's of (A) 150, (B) 100, and (C) 50 ms. The cross sections at τ_m = 100 and 50 ms essentially reflect the primary NOE pattern: a cross-connectivity route involving NOE pathway H2″(i − 1) $\overset{2\text{–}3.0\ \text{Å}}{\cdots}$ H8/H6(i) $\overset{2.0\text{–}2.5\ \text{Å}}{\cdots}$ H2′(i) $\overset{1.8\ \text{Å}}{\cdots}$ H2″(i) indicates on a qualitative basis a gross morphology of a B-DNA for d(GT_4A_4C)$_2$ in solution.

for a few representative protons in one refined B-DNA model for d(GT_4A_4C)$_2$ in solution. Note that, for τ_m = 150, 100, and 50 ms, there is a satisfactory agreement between the calculated and corresponding observed values.

The R factors (as defined under Materials and Methods) for all observed intensities were 0.13, 0.14, and 0.15 for τ_m = 150, 100, and 50 ms, respectively. Agreement between the observed and calculated NOESY intensities did not require any finite discontinuity at the T5→A6 junction; in other words, the identical structural blocks, i.e., d(G1-T2-T3-T4-T5)·d(A6-A7-A8-A9-C10) and d(A6-A7-A8-A9-C10)·d(G1-T2-T3-T4-T5), are connected in such a way that two local frames exactly coincide in space. Even though the A chain is structurally slightly different from the T chain, the flexibility in the T5→A6 junction allows two neighboring structural frames to be coincident. The conformation of the T5→A6 junction is

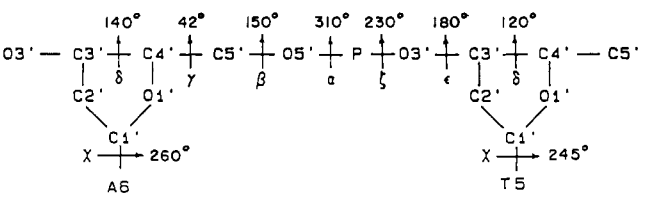


Figure 6 shows the stereo pairs of the final B-DNA model for d(GT_4A_4C)$_2$ in solution. Note that the structure is essentially shown (as marked in the straight axis) to be straight.

Table II: Representative Observed Percent NOEs for Mixing Times of 150, 100, and 50 ms and Those Calculated for the Refined Straight B-DNA Model of $d(GT_4A_4C)_2$[a]

		NOE from H8(G1) at τ of					NOE from H6(T3) at τ of		
NOE to		50 ms	100 ms	150 ms	NOE to		50 ms	100 ms	150 ms
H2′(G1)	obsd	3	6	11	H1′(T2,T3)	obsd		4	6
	cald	5	8	14		cald		2	4
H2″(G1)	obsd	1	2	4	H2″(T2,T3)	obsd	8	17	24
	cald	1	3	5		cald	5.5	15	22
H1′(G1)	obsd		1	1.5	H2′(T2,T3)	obsd	10	21	28
	cald		1	1.5		cald	9	19	26
CH3(T2)	obsd	2.5	4	8	CH3(T3,T4)	obsd	9	18	29
	cald	2	3	7		cald	9	17	28

		NOE from H6(T2) at τ of					NOE from H6(T4) at τ of		
NOE to		50 ms	100 ms	150 ms	NOE to		50 ms	100 ms	150 ms
H1′(G1)	obsd		2	3	H1′(T3)	obsd		2	3
	cald		1	2		cald		1	2.5
H1′(T2)	obsd		2	3	H1′(T4)	obsd		2	3
	cald		1	2		cald		1	2
H2″(T2,G1)	obsd	6	16	23	H2″(T3,T4)	obsd	6	16	23
	cald	5.5	15	22		cald	5.5	15	22
H2′(G1)	obsd	2	4	7	H2′(T3)	obsd	3	5	9
	cald	1	2	5		cald	1	3	6
H2′(T2)	obsd	8	16	24	H2′(T4)	obsd	8	16	24
	cald	8	16	23		cald	8	16	23
CH3(T2)	obsd	6	12	19	CH3(T3,T4)	obsd	9	18	29
	cald	6	12	20		cald	9	18	28
CH3(T3)	obsd	3	6	9					
	cald	2	5	8					

		NOE from H6(T5) at τ of					NOE from H8(A7) at τ of		
NOE to		50 ms	100 ms	150 ms	NOE to		50 ms	100 ms	150 ms
H1′(T4)	obsd			3	H1′(A6)	obsd	1	2	4
	cald			2		cald	0.5	2	3.5
H1′(T5)	obsd			2	H1′(A7)	obsd		1	3
	cald			2		cald		1	2.5
H2′(T4)	obsd	3	5	9	H2′(A6)	obsd	2	4	6
	cald	1.5	4	6.5		cald	0.5	3	5
H2′(T5)	obsd	9	16	22	H2′(A7)	obsd	7	12	15
	cald	9	16	23		cald	6.5	12	16
H2″(T4)	obsd	4.5	7	11	H2″(A6,A7)/H2′(A6)	obsd	5	10	15
	cald	5	9	13		cald	3.5	13	19
H2″(T5)	obsd	3	6	12					
	cald	2	5	10					
CH3(T5)	obsd	6	13	20					
	cald	6	12	20					

		NOE from H8(A6) at τ of					NOE from H8(A8) at τ of		
NOE to		50 ms	100 ms	150 ms	NOE to		50 ms	100 ms	150 ms
H1′(T5)	obsd			2	H1′(A7)	obsd	1	2.5	4
	cald			1.5		cald	0.5	2	3.5
H1′(A6)	obsd		1	2.5	H1′(A8)	obsd		1	2.5
	cald		0.5	1.5		cald		1	1.5
H2′(T5)	obsd	1	2	4	H2′(A7,A8)	obsd	9	13	18
	cald	0.5	2	4		cald	7	14	20
H2″(T5)	obsd	3	5	8	H2″(A7,A8)	obsd	5	9	13
	cald	1.5	4	8		cald	6	11	17
H2′(A6)	obsd	8	13	16					
	cald	6.5	12	16					
H2″(A6)	obsd	4	8	11					
	cald	3	6	8					

		NOE from H8(A9) at τ of					NOE from H6(C10) at τ of		
NOE to		50 ms	100 ms	150 ms	NOE to		50 ms	100 ms	150 ms
H1′(A8)	obsd	1	2.5	4	H1′(A9)	obsd		2.5	4
	cald	0.5	2	3.5		cald		2.5	4.5
H1′(A9)	obsd		1	3	H1′(C10)	obsd		1	3
	cald		1	2.5		cald		1	2
H2′(A8)	obsd	3	6	9	H2′(A9)	obsd	1	3	6
	cald	1	3	5		cald	1	3	6
H2″(A8,A9)	obsd	6	10	14	H2″(A9)	obsd	3.5	6	10
	cald	5	11	17		cald	5	9	13
H2′(A9)	obsd	6	11	15	H2′,H2″(C10)	obsd	6	10	14
	cald	6.5	12	16		cald	7	15	19
					H5(C10)	obsd	6	13	18
						cald	5	11	17

[a] The correlation time is set at 4 ns. The percent NOE is defined as % NOE = (peak height at the site of NOE)/(peak height through which the slice was taken) × 100. Two overlapping sites of NOE are put on the same row, viz., H2″ of T2, T3, etc.

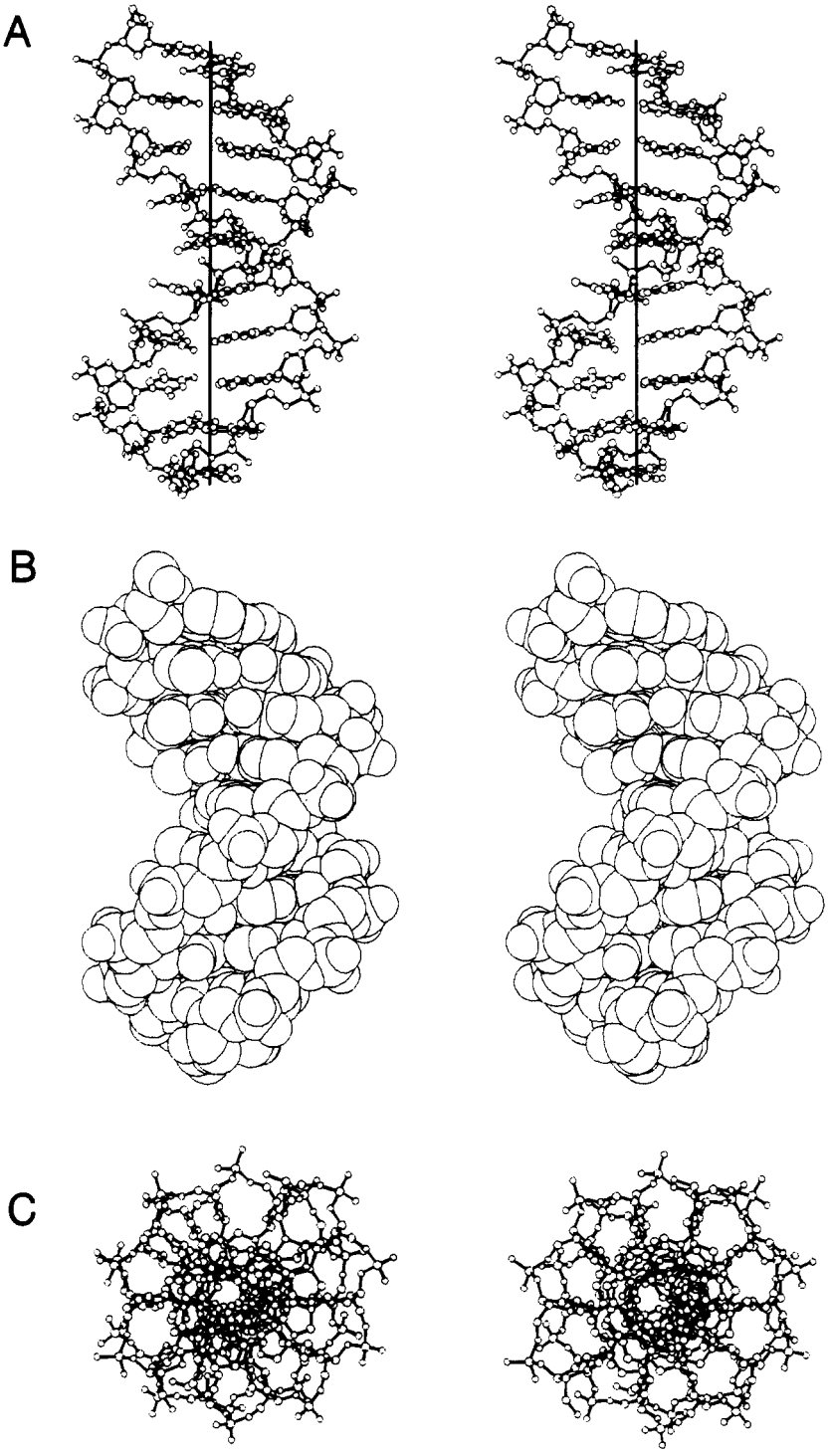


FIGURE 6: Stereo pairs of the straight B-DNA model for $d(GT_4A_4C)_2$: (A) ball and stick model along the helix axis; (B) space-filling model along the helix axis; (C) ball and stick model down the helix axis.

Electrophoretic Studies on d(GAAAATTTTC)$_2$ and d(GTTTTAAAAC)$_2$. Electrophoresis of the synthetic decamers $d(GAAAATTTTC)_2$ and $d(GTTTTAAAAC)_2$ was performed on a 28% polyacrylamide gel at 4 °C and 15 V/cm for 48 h. The electropherogram clearly revealed that the mobility of $d(GAAAATTTTC)_2$ was retarded in the gel compared to that of $d(GTTTTAAAAC)_2$. In view of the fact that the two oligomers have identical molecular weights and charge distribution, this observation is best rationalized on the basis of a bent structure for $d(GAAAATTTTC)_2$ and a straight structure for $d(GTTTTAAAAC)_2$. These electrophoretic studies essentially confirm the conclusions from NMR studies reported in this paper and a previous one (Sarma et al., 1988).

ACKNOWLEDGMENTS

We gratefully acknowledge the assistance of Virginia E. Dollar in the preparation of the manuscript. We thank Professor N. R. Kallenbach for running the cold electrophoresis of $d(GAAAATTTTC)_2$ and $d(GTTTTAAAAC)_2$.

SUPPLEMENTARY MATERIAL AVAILABLE

Fundamental expressions associated with the rate matrix simulation of NOESY data and calculation of R factor and a flow chart of the modus operandi of the programs as well as various input data such as NOESY data not present in Table II and interproton distances (8 pages). Ordering in-

formation is given on any current masthead page.

Registry No. $d(GT_4A_4C)_2$, 116054-40-1.

References

Alexeev, D. G., Lipanov, A. A., & Skuratovskii, I. Y. (1987) *J. Biomol. Struct. Dyn. 4*, 989–1012.

Arnott, S., Chandrasekaran, R., Birdsall, D. L., Leslie, A. G. W., & Ratliff, R. L. (1980) *Nature (London) 283*, 743–745.

Arnott, S., Chandrasekaran, R., Hall, I. H., Puigjaner, L. C. (1983) *Nucleic Acids Res. 11*, 4141–4152.

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

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

Gupta, G., Sarma, M. H., & Sarma, R. H. (1985) *J. Mol. Biol. 186*, 463–469.

Gupta, G., Sarma, M. H., Sarma, R. H., Bald, R., Engelke, U., Oei, S. L., Gessner, R., & Erdmann, V. A. (1987) *Biochemistry 26*, 7715–7723.

Hagerman, P. J. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 4632–4636.

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

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

Kitchin, P., Klein, V. A., Ryan, K. A., Gann, K. L., Rauch, C., Kang, D. S., Wells, R. D., & Englund, P. T. (1986) *J. Biol. Chem. 261*, 11301–11310.

Levene, S. D., & Crothers, D. M. (1983) *J. Biomol. Struct. Dyn. 1*, 429–435.

Marini, J. C., Levene, S. D., Crothers, D. M., & Englund, P. T. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 7664–7668.

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.

Saenger, W. (1984) in *Principles of Nucleic Acid Structures*, Springer-Verlag, New York.

Sarma, M. H., Gupta, G., & Sarma, R. H. (1985) *J. Biomol. Struct. Dyn. 2*, 1057–1084.

Sarma, M. H., Gupta, G., & Sarma, R. H. (1986) *Biochemistry 25*, 3659–3665.

Sarma, M. H., Gupta, G., & Sarma, R. H. (1987a) *J. Biomol. Struct. Dyn. 1*, 1423–1455.

Sarma, M. H., Gupta, G., Sarma, R. H., Bald, R., Engelke, U., Oei, S. L., Gessner, R., & Erdmann, V. A. (1987b) *Biochemistry 26*, 7707–7714.

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

States, D. J., Haberkorn, R. H., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.

Trifonov, E. N. (1986) *CRC Crit. Rev. Biochem. 19*, 89–106.

Ulanovsky, L. E., & Trifonov, E. N. (1987) *Nature (London) 326*, 720–722.

Zahn, K., & Blattner, F. R. (1987) *Science (Washington, D.C.) 236*, 416–422.

Fluorescence Measurement of the Kinetics of DNA Injection by Bacteriophage λ into Liposomes†

Steven L. Novick‡ and John D. Baldeschwieler*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Received March 16, 1988; Revised Manuscript Received June 23, 1988

ABSTRACT: Bacteriophage λ attaches to Gram-negative bacteria using the outer membrane protein LamB as its receptor. Subsequently, DNA is injected by the bacteriophage into the host cell for replication and expression. The mechanism of DNA injection, however, is poorly understood. In order to begin to characterize DNA injection, a quantitative kinetic assay to detect injection into reconstituted LamB liposomes is described. The technique involves monitoring the increase in fluorescence of liposome-encapsulated ethidium bromide, which occurs as DNA enters the aqueous compartment of the vesicles. The data indicate that injection is several times faster than indicated by earlier studies and is complete within 1 min. Such assays which allow direct observation of this process are necessary first steps toward a mechanistic understanding.

Bacteriophage λ is an important tool in molecular biology and biochemistry due to its ease of laboratory manipulation. Little is known, however, about the molecular mechanisms by which it functions. Bacteriophage λ is a temperate, double-stranded DNA-containing phage which infects Gram-negative bacteria, using the outer membrane maltose porin LamB as its receptor (Szmelcman & Hofnung, 1975). LamB exists in the membrane as an integral trimer of 47-kDa subunits (Schwartz, 1975, 1983). In addition to its role in bacteriophage attachment, LamB is also the channel through which phage DNA passes as it enters the host cell (Roessner & Ihler, 1986).

Some studies have been conducted to investigate the phage–receptor interaction in vitro and in well-characterized model membranes such as phospholipid vesicles. These studies focused on phage λ interactions with LamB extracted from *Escherichia coli* K-12. The complexes are formed reversibly (Schwartz, 1975) and require addition of chloroform to trigger irreversible binding and DNA injection (Schwartz, 1975; Mackay & Bode, 1976a; Zgaga et al., 1973). λ host range mutants (λh) which bind LamB from *E. coli* K-12 irreversibly and inject their DNA in the absence of chloroform have been identified (Randall-Hazelbauer & Schwartz, 1973). In ad-

† This investigation was supported by ARO Grant DAAG-29-83-K-0128, by NIH Biomedical Research Support Grant RR07003 awarded by the Division of Research Resources, and by gifts from Amersham and Monsanto.

* To whom correspondence should be addressed.

‡ Recipient of an NSF predoctoral research fellowship and a CIT institute fellowship.

0006-2960/88/0427-7919$01.50/0