

# Solution Structure of $[d(A-T)_5]_2$ via Complete Relaxation Matrix Analysis of Two-Dimensional Nuclear Overhauser Effect Spectra and Molecular Mechanics Calculations: Evidence for a Hydration Tunnel<sup>†</sup>

Ei-ichiro Suzuki,<sup>‡</sup> Nagarajan Pattabiraman,<sup>‡</sup> Gerald Zon,<sup>§</sup> and Thomas L. James<sup>\*†</sup>

Departments of Pharmaceutical Chemistry and Radiology, University of California, San Francisco, California 94143, and The Center for Drugs and Biologics, Food and Drug Administration, Bethesda, Maryland 20205

Received March 17, 1986; Revised Manuscript Received June 12, 1986

**ABSTRACT:** Pure absorption phase proton two-dimensional nuclear Overhauser effect (2D NOE) spectra at 500 MHz have been obtained for  $[d(5'ATATATATAT3')]_2$  in deuterium oxide solution at several mixing times. The 100 nonexchangeable proton resonances have been assigned. The experimental 2D NOE spectra were compared with theoretical spectra calculated by using the complete relaxation matrix analysis method [Keepers, J. W., & James, T. L. (1984) *J. Magn. Reson.* 57, 404-426] and X-ray diffraction determined molecular coordinates of A, B, alternating B, left-handed B, C, D, and wrinkled D forms of DNA and of energy-minimized structures calculated from the most promising X-ray crystal structures by using the molecular mechanics program AMBER, in which all hydrogens, counterions, and hydration water molecules were included. The analysis of all features of the 2D NOE spectra played an important role in extracting the promising structures, and it was concluded that the wrinkled D form yields the best fit for the 2D NOE data of the A-T decamer. The molecular mechanics calculation indicated that this model structure, whose minor groove is comparatively deep and narrow, may be energetically more stable than the B form for alternating d(A-T) DNA. Interesting features of the structure include possible intra- and interchain sugar-phosphate attractions and a hydration tunnel inside the minor groove capable of accommodating three types of water molecules that aid in helix stabilization via hydrogen bonding. Counterions (sodium) serve to reduce interchain phosphate-phosphate repulsive effects.

A major goal for many years has been the determination of molecular structures in noncrystalline environments. High-resolution structures such as those derived from X-ray diffraction on crystals could not be remotely attained. Two-dimensional NMR techniques (Bax, 1982), in particular the two-dimensional nuclear Overhauser effect (2D NOE) experiment (Macura & Ernst, 1980), have the promise for yielding data that can provide high-resolution molecular structures in solution, especially if used with the distance geometry algorithm (Havel et al., 1983). High-resolution structures can be determined if one has a large number of structural constants, e.g., internuclear distances and bond torsion angles, to use in conjunction with the holonomic constraints of bond lengths, bond angles, and steric limitations.

The 2D NOE experiment has the potential for providing numerous internuclear distances. A number of elegant studies have utilized the technique for qualitative or semiquantitative insights into molecular structure, including oligonucleotides [e.g., Feigon et al. (1983), Frechet et al. (1983), Scheek et al. (1984), Broido et al. (1984), Lai et al. (1984), Wemmer et al. (1985), Hilbers et al. (1985), and Cavailles et al. (1985)]. Recent reviews (Kearns, 1984; James, 1984) have covered the use of NMR in general for studies of nucleic acid structure

and dynamics. We have been investigating the use of 2D NOE experiments for solution structure determination via numerous accurate internuclear distances with a complete relaxation matrix analysis theoretically (Keepers & James, 1984) and in a rigid molecule test case (Young & James, 1984). Unlike other 2D NOE methodologies, this approach takes into account all interproton interactions occurring in the molecule and is not limited by assumptions regarding the number or geometry of proton spins, spin diffusion, the details of molecular motions, or the range of mixing times utilized in the experiment.

The subject of this study is  $[d(5'ATATATATAT3')]_2$  in aqueous solution at low salt concentration (180 mM NaCl, 100 mM phosphate buffer, pH 7.0). There has been some controversy about the structure of poly(dA-dT) in solution with NMR evidence being cited for it being in a normal B-DNA conformation (Assa-Munt & Kearns, 1984; Borah et al., 1985) and for it being in a left-handed B-type conformation (Gupta et al., 1983). Indeed, structures suggested for poly(dA-dT) from X-ray diffraction have included A-DNA (Davies & Baldwin, 1963), B-DNA (Arnott & Hukins, 1972), alternating B-DNA (Klug et al., 1979), C-DNA (Marvin et al., 1961), right-handed D-DNA (Arnott & Selsing, 1974), left-handed D-DNA with Hoogsteen base pairing (Drew & Dickerson, 1982), and wrinkled D-DNA (Arnott et al., 1983).

We have used 2D NMR to assign the 100 nonexchangeable proton resonances of the A-T decamer duplex. With small molecules, we have been able to fit iteratively a series of four experimental 2D NOE spectra obtained at four mixing times with corresponding theoretical 2D NOE spectra simultaneously (Young & James, 1984). Computer limitations have so far made this infeasible for our studies of octamer duplexes (Broido et al., 1985; Jamin et al., 1985) and the present de-

<sup>†</sup> This work was supported by the National Science Foundation (Grant PCM 84-04198), the American Cancer Society (Grant NP-437), and the National Institutes of Health (Grant CA 27343) as well as by instrumentation grants from the National Institutes of Health (Grant RR 016688) and the National Science Foundation (DMB 84-06826).

\* Address correspondence to this author at the Department of Pharmaceutical Chemistry, UCSF.

<sup>‡</sup> University of California, San Francisco.

<sup>§</sup> Food and Drug Administration.

camer duplex, where four  $100 \times 100$  matrices would need to be fit simultaneously. Our present approach has been to compare the four experimental spectra with calculated spectra by using atomic coordinates from X-ray studies for A, B, alternating B, left-handed B, C, D, and wrinkled D forms of DNA. The most promising of these structures, namely B, D, and wrinkled D forms, were used with the molecular mechanics program AMBER (Weiner & Kollman, 1981) to generate energy-minimized structures. Some of the calculations entailed inclusion of ions and water molecules. Theoretical 2D NOE spectra for the energy refinement calculations, together with computer graphics display, have yielded additional insights into features of the molecular structure of the A-T decamer duplex in solution.

## MATERIALS AND METHODS

**Sample Preparations.** The self-complementary decamer [d(5'ATATATATAT3')]₂ was synthesized by the same methodology as described in a previous paper (Broido et al., 1984). An amount of 20 mg (ca. 400 absorbance units) of the nucleotide was dissolved in 0.4 mL of a phosphate buffer [100 mM phosphate, 180 mM NaCl, 0.2 mM ethylene glycol bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid (EGTA), pH 7.02]. The molecular weight (as duplex) is 6460, so the molar concentration is about 7 mM. This concentration is twice that used in the previous studies (Broido et al., 1985; Jamin et al., 1985). It was necessary for 2D NOE measurements because the signal line widths were broader (vide infra) and the consequent S/N ratio lower for the present decamer than for the octamers (Broido et al., 1985; Jamin et al., 1985). Nevertheless, the primary signal assignments were obtained for 3 mM solution. This sample was lyophilized several times from D₂O solution and finally dissolved in 99.996% D₂O (Stohler Isotope Chemicals) under dry nitrogen atmosphere to minimize the HDO peak.

**Running Temperature.** After the melting behavior of the duplex was investigated, the running temperature of 15 °C was chosen, which is sufficiently low to minimize fraying but not so low as to induce unnecessary signal broadening. At 6.5 °C or lower, serious broadening was observed, although the signal positions were almost the same as for poly(dA-dT) (Assa-Munt & Kearns, 1984), except for terminal residue signals. But 10.5 °C was also useful for resonance assignments.

**NMR Measurements.** All <sup>1</sup>H NMR spectra were obtained at 500 MHz on the GE/Nicolet GN500 equipped with an Oxford Instruments magnet and a Nicolet 1280 computer. Pure absorption 2D NOE spectra were acquired by using the pulse sequence (delay time-90°-t₁-90°-τₘ-90°-t₂)ₙ (Kumar et al., 1980) with alternate block accumulation, in which the phases of the three 90° pulses were cycled in a well-described sequence (States et al., 1982) to eliminate artifacts arising from pulse imperfections as well as single and double quantum coherences. The mixing times τₘ used were 50, 100, 250, and 400 ms. The typical 90° pulse width was 12 μs, and the delay time was 12 s. Magnetization transfer through zero quantum coherence was suppressed by using a slight randomization of the mixing time (Macura et al., 1981). The number of accumulations *n* was 32, and the carrier frequency was set to the HDO signal position. Four hundred free induction decays (FIDs) were collected in the *t*₁ dimension with 4K data points in the *t*₂ dimension stored in alternate blocks. The FIDs were apodized in the *t*₁ and *t*₂ dimensions by a 45° shifted sine bell function for resolution enhancement. After zero filling in both dimensions, the final data matrix size was 1K × 1K of real

points ( $\omega_1 \times \omega_2$ ) with digital resolution of 5.9 Hz per point, as the sweep width was 6024 Hz.

## THEORY

The theory for 2D NOE was developed in Ernst's laboratory (Macura & Ernst, 1980). Subsequently, we (Keepers & James, 1984) described a matrix method relating all interproton distances in a molecule to peak intensities in a theoretical 2D NOE spectrum, which can be iteratively compared with the experimental pure absorption 2D NOE spectrum.

The intensity of an auto (diagonal) or cross peak for any value of the experimental parameter τₘ, the mixing time, may be calculated from

$$a(\tau_m) = \chi \exp(-\lambda \tau_m) \chi^{-1} \quad (1)$$

where element *a*<sub>ij</sub> of the matrix **a** gives the 2D NOE cross-peak intensity for nuclei *i* and *j*,  $\chi$  is the matrix of eigenvectors of the relaxation rate matrix **R**, and  $\lambda$  is a diagonal matrix of eigenvalues, i.e., the solution to the system of equations comprising the rate matrix. The relaxation rate matrix **R** has diagonal elements

$$R_{ii} = \sum_j (W_0^{ij} + 2W_1^{ij} + W_2^{ij}) + R_{1i} \quad (2)$$

representing the direct relaxation contributions from all sources to spin *i*, and off-diagonal elements

$$R_{ij} = W_2^{ij} - W_0^{ij} \quad (3)$$

representing the cross-relaxation rate between spins *i* and *j*. *R*<sub>1i</sub> comprises all sources of relaxation other than proton-proton dipolar interactions. The zero-, single-, and double-quantum transition probabilities *W*<sub>0</sub><sup>ij</sup>, *W*<sub>1</sub><sup>ij</sup>, and *W*<sub>2</sub><sup>ij</sup> are proportional to *J*(ω)/*r*<sup>6</sup><sub>ij</sub>, where *J*(ω) is the spectral density for the molecular motion modulating the *ij* interaction and *r*<sub>ij</sub> is the distance between protons *i* and *j*.

A pure absorption 2D NOE spectrum will depend on choice of the mixing time τₘ used, molecular motions as manifest in the spectral densities, and internuclear distances between all protons, not just the two giving rise to a cross peak. Macura and Ernst (1980) showed that a pair of spins *i* and *j* can be considered as isolated if τₘ is sufficiently short, thereby enabling use of only a couple terms in a series expansion and obviating the need for the complete matrix solution outlined above. In practice it is difficult to isolate spins *i* and *j* completely from spin *k* if *r*<sub>ik</sub> ≤ *r*<sub>ij</sub>. Although an *ij* cross peak can still be observed and used qualitatively, the peak intensity will be modified, thus preventing quantitative internuclear distance determinations.

A theoretical investigation of the potential for 2D NOE spectral determinations of internuclear distances led to the following conclusions (Keepers & James, 1984; Young & James, 1984): (a) several pure absorption 2D NOE spectra at a series of mixing times should be obtained; (b) a practical upper limit on determination of distance is ≈5 Å; (c) distances should be determined to an accuracy of ≈10% if a good fit to all data (nonoverlapping peaks assumed) is obtained; and (d) isotropic motion with a single effective correlation time can be assumed for 10% distance accuracy even though the actual motion may be more complicated.

## RESULTS AND DISCUSSION

**Resonance Assignments.** The 500-MHz <sup>1</sup>H NMR spectrum of the nonexchangeable protons of the decamer duplex is shown in Figure 1. The line widths are somewhat broader than with the octamers in the previous studies (Broido et al., 1984; Jamin

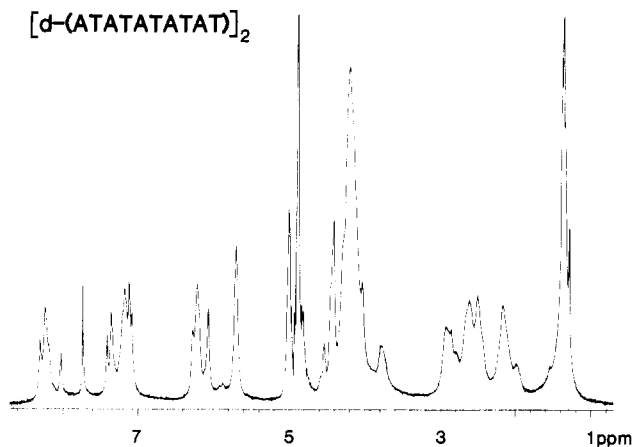


FIGURE 1: 500-MHz  $^1\text{H}$  NMR spectrum of  $[\text{d}-(5'\text{ATATATATAT}3')_2]$ . Sample conditions: 15  $^\circ\text{C}$ ; 7 mM; pH 7.0. The HDO signal was used as a chemical shift reference 4.89 ppm downfield of sodium 3-trimethyl[2,2,3,3- $^2\text{H}_4$ ]propionate.

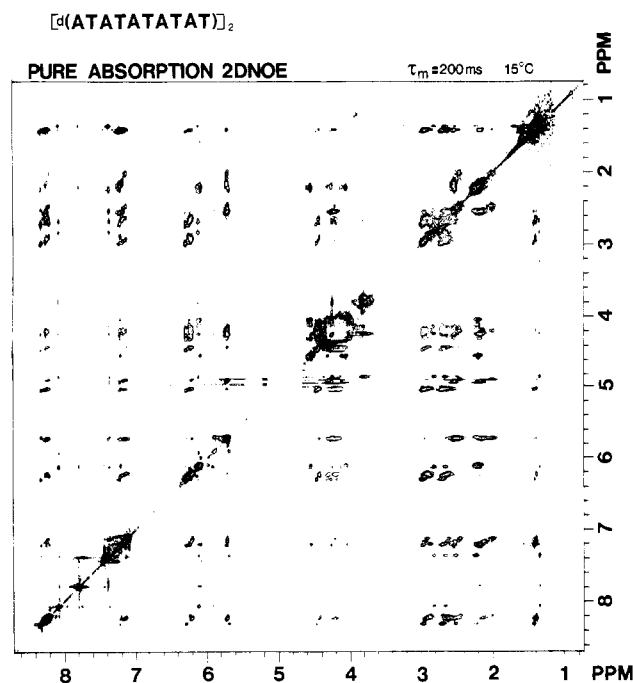


FIGURE 2: A pure absorption  $^1\text{H}$  2D NOE spectrum (500 MHz) of  $[\text{d}(5'\text{ATATATATAT}3')]_2$  at 15  $^\circ\text{C}$ .

et al., 1985), and signals are highly superimposed. But the base proton resonances (1.0–2.0 and 7.0–8.5 ppm) are a little better resolved than the sugar proton resonances (2.0–6.5 ppm). Peak positions and line widths were essentially independent of concentration in the range 3–7 mM.

Figure 2 shows the 500-MHz pure absorption phase 2D NOE spectrum of the decamer recorded with a 200-ms mixing time. Each cross peak observed in a 2D NOE spectrum correlates the chemical shifts of two nuclei between which cross-relaxation occurs during the mixing time due to the physical proximity of the nuclei and thus is useful for peak assignments. This is illustrated in Figure 3 for some of the dipolar connectivities in B-family DNA. With knowledge of some proton resonance assignments, others may be discerned by following the dipolar connectivities in a sequential manner. Since most nuclei have several cross peaks to other nuclei, there are cross-checks for the assignments; i.e., the whole pattern of 2D NOE peaks must be internally consistent. Deviation from the B-DNA family of structures is readily apparent. The 2D NOE spectra of DNA oligomers have previously been used

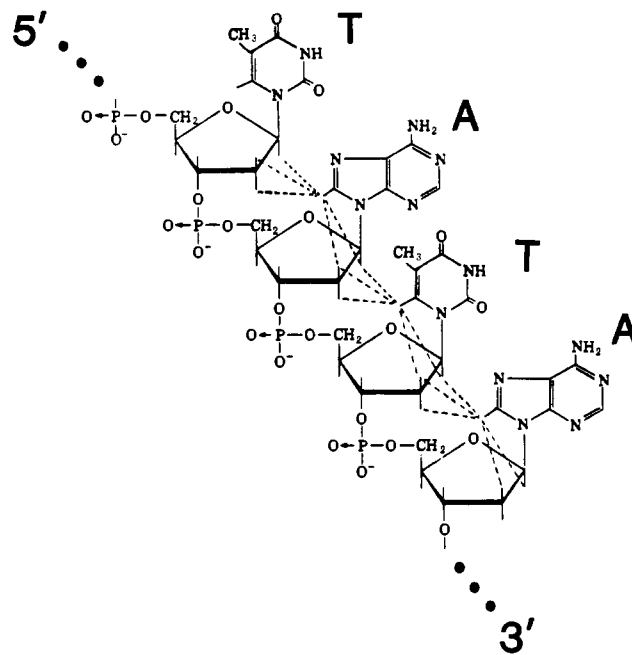
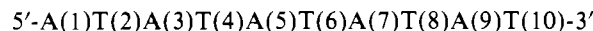


FIGURE 3: Some of the interproton dipolar connectivities (broken lines) enabling proton resonance assignments in B-family DNA, i.e., DNA in the base anti conformation.

in this manner to make resonance assignments (Feigon et al., 1983; Frechet et al., 1983; Scheek et al., 1984; Broido et al., 1984; Lai et al., 1984; Wemmer et al., 1985; Hilbers et al., 1985; Cavailles et al., 1985). We have used the same approach for assigning the 100 nonexchangeable proton resonances of  $[\text{d}(\text{A-T})_5]_2$  at 15  $^\circ\text{C}$ .

The numbering scheme used is



For example, the isolated peak at 3.78 ppm was assigned unequivocally to A(1)-H5',5'' (sugar protons of the 5'-terminal adenine residue) and correlated to its own A(1)-H8 (8.01 ppm) via a weak cross peak. A(1)-H2',2'' protons were assigned via cross peaks with A(1)-H8 and correlated to 3'-neighboring T(2)-H6 as A(1) has no 5'-neighboring thymine residue. A(3)-H8 was assigned to the 8.30-ppm peak from the cross peaks with T(2)-H2',2'' and correlated to its own A(3)-H2',2'' and so on. Thus, all the aromatic and H2',2'' protons were assigned reasonably. Methyl protons in thymine residues were assigned by their interactions with T-H6's. There was no assumption for the handedness, but only intrastrand interactions were assumed. After the above assignments, H1' protons could be assigned from the cross peaks with base protons (A-H8, T-H6). All T-H1' protons have almost the same chemical shift (5.7 ppm) except for the terminal one. Subsequently, H3' and H4' protons were assigned, except for T(2)-, T(4)-, T(6)-, and T(8)-H4' protons, which are superimposed with roughly assigned H5',5'' resonances.

All five A-H2 signal positions were determined, although each assignment is somewhat ambiguous. These assignments could be tentatively made from 2D NOE spectra at 10.5  $^\circ\text{C}$ . The two A-H2 resonances (7.08 and 7.11 ppm at 15  $^\circ\text{C}$ ) were completely separated at this temperature (6.95 and 7.00 ppm) from T-H6's and correlated with A(7)- and A(5)-H1', respectively. Hence, these signals can be assigned respectively to A(5)- and A(7)-H2 (if due to interstrand interactions of the base pair in a B-family structure) or A(7)- and A(5)-H2 (if due to intraresidue interactions). The lowest field resonance at 7.72 ppm should be assigned to A(1)-H2. The other two resonances at 7.35 and 7.40 ppm can be assigned to A(9)- and

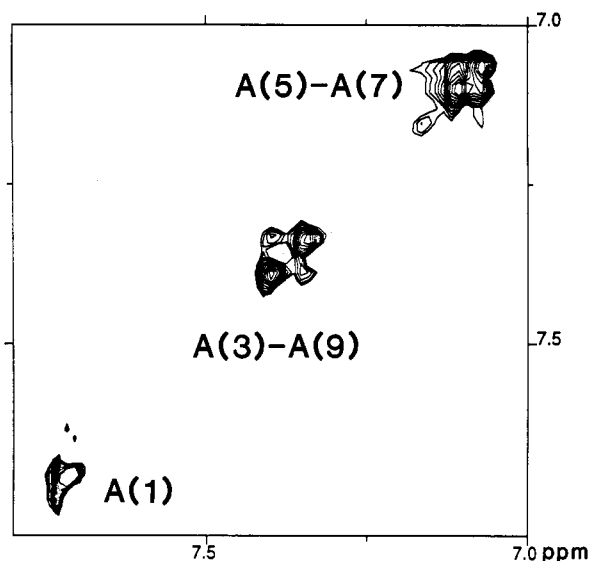


FIGURE 4: Expanded region of the 2D NOE spectrum with  $\tau_m = 400$  ms. The diagonal and cross peaks between A-H2 protons are shown. Table I gives assignments.

A(3)-H2. Following these assignments, it is evident that prominent cross peaks due to interstrand interactions between A(9)- and A(3)-H2 and between A(7)- and A(5)-H2 protons were observed (Figure 4); the latter is stronger, which may be due to the stability of the middle part of the duplex. At  $\tau_m$  of 400 ms, only these cross and diagonal peaks of A-H2 protons remain strong, although comparatively weak T-H6 cross peaks were also observed. This is an important observation (vide infra).

There are some weak unidentified cross and diagonal peaks remaining in the aromatic and H1' regions. These apparently belong to a second form of the decamer such as a hairpin structure. But, looking at the T-H1' region of the one-dimensional (1D) spectrum, the population of the second form is clearly less than 10% (Figure 1). So, this second structure was neglected in the following discussion.

Assignments for all 100 nonexchangeable protons are summarized in Table I. As mentioned above, the pure absorption 2D NOE experiment was extremely useful for assigning proton resonances in the complicated spectrum of a decamer with a simple repeat sequence of alternating A-T.

**Evaluation of Observed 2D NOE Peak Intensities.** The observed intensities were classified into nine grades: grade 0 (not observed) to grade 8 (the strongest). When the grade goes up by one, the intensity becomes twice as strong (geometrical progression with a factor of 2). As the cross-peak intensities between A(3)-H2 and A(9)-H12 protons were at most 5% of the diagonal peaks' intensities even at  $\tau_m$  of 400 ms, the diagonal peak of A-H2 was chosen as a nonchangeable

intensity standard (within 10% experimental error) and defined as the highest intensity value 1.0 (for convenience 1.024; grade 8). Hence, grade 7 means an intensity between 0.512 and 1.024, grade 6 is 0.256–0.512, grade 5 is 0.128–0.256, grade 4 is 0.064–0.128, grade 3 is 0.032–0.064, grade 2 is 0.016–0.032, and grade 1 is 0.004–0.016. This method of grading intensities is a conservative approach acknowledging the difficulties in obtaining cross-peak intensities with high accuracy. With this conservative approach, if the intensity difference between two signals is a factor of 2 or larger, it can be distinctly recognized as "different". As illustrated previously (Young & James, 1984; Keepers & James, 1984), the intensities in 2D NOE spectra are very sensitive to all the inter-proton distances.

The initial method for intensity evaluation was that of peak heights. Secondly, those were reconsidered by investigating peak volume and superimposition (Broide et al., 1985). The cross-peak intensity for each one of three methyl protons of thymine residues was obtained by dividing the observed intensity by 3.0. For the major part of our study, only the intensities of the central 5'-T(4)A(5)T(6)A(7)-3' part were analyzed as a very conservative approach to neglect end effects. We checked to see if this choice would affect our results; it seems that there is no serious difference from the 5'-T(2)A(3)T(4)A(5)-3' or 5'-T(6)A(7)T(8)A(9)-3' parts after analyzing the well-resolved base-H2',2'' interaction region. So, the intensity data are given without the residue numbers, just with the distinction of 5'- or 3'-neighbor interaction.

The 2D NOE peak intensities are reported in Table II for the four mixing times (50, 100, 250, 400 ms) employed. All of these cross-peak intensities are less than grade 5.

**Calculation of Theoretical 2D NOE Spectra via Complete Relaxation Matrix Analysis.** Although not required for the relaxation matrix approach (Keepers & James, 1984), isotropic motion of the decamer and dominance of the proton-proton dipolar relaxation mechanism were assumed to simplify the model calculations. Only protons in one strand were considered as the first approximation, and the effect of including both strands was checked for the final result of the best fit model.

Theoretical decay curves of A-H1' and T-H1' diagonal peaks with mixing time have been theoretically calculated for well-known X-ray structures of DNA and compared with experimental diagonal peak intensity data. The initial slopes of these decay curves give the commonly measured selective spin-lattice relaxation times. With the prevalent dipolar relaxation mechanism, proton relaxation will depend on the exact structure. So we have considered structural consequences on our selection of the effective isotropic correlation time. An isotropic correlation time ( $\tau_c$ ) of 3 ns was used as an initial value, which was obtained in the previous study of an octamer (Jamin et al., 1985). Fitting for regular B-form DNA was satisfactory, but it was not successful for any of the other

Table I: Chemical Shifts<sup>a</sup> (ppm) of the Nonexchangeable Protons in [d(A-T)<sub>5</sub>]<sub>2</sub>

	H8	H2	H6	CH <sub>3</sub>	H1'	H2'	H2''	H3'	H4'	H5'	H5''
A(1)	8.01 <sub>0</sub>	7.71 <sub>5</sub>			6.06 <sub>5</sub>	2.63 <sub>0</sub>	2.79 <sub>5</sub>	4.81 <sub>5</sub>	4.20 <sub>5</sub>	3.78 <sub>0</sub>	3.78 <sub>0</sub>
T(2)			7.33 <sub>5</sub>	1.28 <sub>0</sub>	5.68 <sub>0</sub>	2.19 <sub>0</sub>	2.49 <sub>0</sub>	4.90 <sub>0</sub>	(4.22 <sub>5</sub> )	4.22 <sub>0</sub>	4.12 <sub>5</sub> ) <sup>b</sup>
A(3)	8.29 <sub>5</sub>	7.40 <sub>0</sub> <sup>b</sup>			6.28 <sub>0</sub>	2.68 <sub>5</sub>	2.95 <sub>5</sub>	5.02 <sub>0</sub>	4.45 <sub>5</sub>	4.23 <sub>0</sub>	4.14 <sub>5</sub>
T(4)			7.19 <sub>5</sub>	1.38 <sub>5</sub>	5.71 <sub>0</sub>	2.13 <sub>5</sub>	2.52 <sub>5</sub>	4.90 <sub>0</sub>	(4.24 <sub>5</sub> )	4.30 <sub>5</sub>	4.17 <sub>0</sub> ) <sup>b</sup>
A(5)	8.23 <sub>5</sub>	7.11 <sub>0</sub> <sup>c</sup>			6.22 <sub>5</sub>	2.62 <sub>0</sub>	2.92 <sub>5</sub>	5.00 <sub>5</sub>	4.43 <sub>0</sub>		(4.23) <sup>d</sup>
T(6)			7.15 <sub>5</sub>	1.35 <sub>5</sub>	5.70 <sub>5</sub>	2.11 <sub>0</sub>	2.51 <sub>0</sub>	4.89 <sub>5</sub>	(4.22 <sub>5</sub> )	4.22 <sub>5</sub>	4.16 <sub>5</sub> ) <sup>b</sup>
A(7)	8.23 <sub>0</sub>	7.07 <sub>5</sub> <sup>c</sup>			6.19 <sub>5</sub>	2.60 <sub>5</sub>	2.90 <sub>5</sub>	5.00 <sub>5</sub>	4.41 <sub>5</sub>		(4.20) <sup>d</sup>
T(8)			7.11 <sub>0</sub>	1.36 <sub>0</sub>	5.70 <sub>5</sub>	1.98 <sub>5</sub>	2.45 <sub>0</sub>	4.87 <sub>5</sub>	(4.19 <sub>5</sub> )	4.28 <sub>5</sub>	4.16 <sub>0</sub> ) <sup>b</sup>
A(9)	8.17 <sub>5</sub>	7.34 <sub>5</sub> <sup>b</sup>			6.22 <sub>5</sub>	2.67 <sub>5</sub>	2.86 <sub>5</sub>	4.99 <sub>5</sub>	4.41 <sub>0</sub>		(4.20) <sup>d</sup>
T(10)			7.16 <sub>5</sub>	1.32 <sub>5</sub>	6.07 <sub>0</sub>	2.18 <sub>0</sub>	2.18 <sub>0</sub>	4.54 <sub>0</sub>	4.30 <sub>0</sub>	4.09 <sub>5</sub>	4.03 <sub>0</sub>

<sup>a</sup> Relative to the HDO resonance at 4.89 ppm from TSP; 15 °C; 7 mM as duplex; pH 7.02;  $\pm 0.005$  ppm. <sup>b</sup> It was not possible to distinguish between these resonances. <sup>c</sup> It was not possible to distinguish between these resonances. <sup>d</sup> Superimposed peaks.

Table II: Observed Intensities of 2D NOE Spectra of the A-T Decamer Duplex at 50-, 100-, 250-, and 400-ms Mixing Times<sup>a</sup>

mixing time (ms)		to											
		T-Me	T-H2'	T-H2''	A-H2'	A-H2''	T-H4'	A-H4'	T-H3'	A-H3'	T-H1'	A-H1'	T-H6
50	from												
	A-H8	2 <sup>c</sup>	1 <sup>b</sup>	2 <sup>b</sup>	3	2	0	0	0	1	2 <sup>b</sup>	1	0
	T-H6	2	4	2	2 <sup>b</sup>	3 <sup>b</sup>	0	0	1	0	1	2 <sup>b</sup>	
	A-H1'	0	0	0	3	4	0	2	0	1	0		
	T-H1'	0	3	4	0	0	2	0	1	0			
	A-H3'	0	0	0	3	2	0	3	<i>d</i>				
	T-H3'	0	3	2	0	0	<i>d</i>	0					
	A-H4'	0	0	0	1	1	<i>d</i>						
	T-H4'	0	1	1	0	0							
	A-H2''	1 <sup>c</sup>	0	0	5								
	A-H2'	1 <sup>c</sup>	0	1									
	T-H2''	0	5										
	T-H2'	0											
100	from												
	A-H8	2 <sup>c</sup>	2 <sup>b</sup>	3 <sup>b</sup>	4	3	0	1	1 <sup>b</sup>	2	2 <sup>b</sup>	2	1 <sup>c</sup>
	T-H6	3	4	3	3 <sup>b</sup>	3 <sup>b</sup>	1	0	2	1 <sup>b</sup>	2	3 <sup>b</sup>	
	A-H1'	1 <sup>c</sup>	0	1	4	5	1 <sup>c</sup>	3	1	2	1		
	T-H1'	0	2	4	0	0	2	0	2	1			
	A-H3'	1 <sup>c</sup>	0	0	4	3	0	4	<i>d</i>				
	T-H3'	0	3	3	0	0	<i>d</i>	0					
	A-H4'	0	0	0	1	2	<i>d</i>						
	T-H4'	0	1	1	0	0							
	A-H2''	2 <sup>c</sup>	0	1	5								
	A-H2'	2 <sup>c</sup>	0	1									
	T-H2''	1	5										
	T-H2'	0											
250	from												
	A-H8	3 <sup>c</sup>	2 <sup>b</sup>	3 <sup>b</sup>	3	3	<i>d</i>	2	2 <sup>b</sup>	3	3 <sup>b</sup>	3	1, <sup>b</sup> 2 <sup>c</sup>
	T-H6	3	4	3	3 <sup>b</sup>	3 <sup>b</sup>	<i>d</i>	2	2	2 <sup>b</sup>	3	3 <sup>b</sup>	
	A-H1'	2 <sup>c</sup>	1	2	3	4	<i>d</i>	3	2	3	2		
	T-H1'	1	3	4	1	1	<i>d</i>	1	3	2			
	A-H3'	2	1	2	4	4	<i>d</i>	4	<i>d</i>				
	T-H3'	1	3	3	1	1	<i>d</i>	1					
	A-H4'	1	0	1	3	3	<i>d</i>						
	T-H4'	<i>d</i>	<i>d</i>	<i>d</i>	<i>d</i>	<i>d</i>							
	A-H2''	3	0	2	4								
	A-H2'	3	1	2									
	T-H2''	2	4										
	T-H2'	1											
400	from												
	A-H8	4 <sup>c</sup>	2 <sup>b</sup>	3 <sup>b</sup>	3	3	<i>d</i>	3	3 <sup>b</sup>	3	3 <sup>b</sup>	3	2, <sup>b</sup> 3 <sup>c</sup>
	T-H6	4	4	3	3 <sup>b</sup>	3 <sup>b</sup>	<i>d</i>	3	3	3 <sup>b</sup>	3	3 <sup>b</sup>	
	A-H1'	3 <sup>c</sup>	2	3	3	4	<i>d</i>	3	3	3	3		
	T-H1'	2	4	4	2	2	<i>d</i>	2	4	3			
	A-H3'	3	2	3	4	4	<i>d</i>	4	<i>d</i>				
	T-H3'	2	3	3	2	2	<i>d</i>	1					
	A-H4'	2	1	2	3	3	<i>d</i>						
	T-H4'	<i>d</i>	<i>d</i>	<i>d</i>	<i>d</i>	<i>d</i>							
	A-H2''	3	1	3	4								
	A-H2'	3	2	3									
	T-H2''	2	4										
	T-H2'	2											

<sup>a</sup> Only A(5) and T(6) cross-peak intensities are shown here. Other interior nucleotide residues had essentially the same intensities. See text for intensity grading scale. <sup>b</sup> Cross-peak intensity "from" a nucleotide proton "to" a proton on its 5'-neighbor. <sup>c</sup> Cross-peak intensity "from" a nucleotide proton "to" a proton on its 3'-neighbor. <sup>d</sup> Not resolved.

models, for which the decay rates were too slow with this correlation time. The reason why the calculated diagonal peak decay for regular B form is so fast is because the distances between H1' and its 3'-neighboring H-5',5'' are quite short in B form. But the observed cross peaks between H1' and the 3'-neighboring H-5',5'' are definitely no stronger than grade 3, but are probably grade 1, while the calculated cross-peak intensities with H1' for both T- and A-H5',5'' should be at grade 5 for B form (at  $\tau_m = 100$  ms). These intensities can only be estimated but not evaluated exactly because of superimposition with the H4' peak. Regardless, a 3-ns correlation time is clearly too short.

From nonselective  $T_1$  and Hahn spin-echo  $T_2$  relaxation time measurements of the relatively well-resolved A(9)-H2 proton (3.6 and 0.011 s, respectively) and TH1' proton (2.3 and

0.0085 s, respectively),  $\tau_c$  values of 7.0 and 6.0 ns were calculated by using the equation (simple and rigid molecular motion assumed), which was derived from the well-known formula [e.g., Woessner (1962)]

$$\tau_c = 2\omega^{-1}(3T_2/T_1)^{-1/2}$$

which holds when  $\omega\tau_c \gg 1$ ;  $\omega$  is the Larmor frequency. In the octamer case (Jamin et al., 1985), for A-H8,  $T_2 = 0.024$  s and  $T_1 = 1.6$  s, yielding  $\tau_c \approx 3.0$  ns, which is just the same as the value used in that study.

Hence, a  $\tau_c$  value of 7 ns was used for the 2D NOE calculations; A-H1' diagonal peak decay curves calculated for several DNA structures are shown in Figure 5. It can be seen in the figure that all of the calculated curves effectively fit the experimental observations. The value of 7 ns as the isotropic

Table III: Calculated 2D NOE Intensities at 100-ms Mixing Time for Several DNA Models<sup>a</sup>

model		to											
		T-Me	T-H2'	T-H2''	A-H2'	A-H2''	T-H4'	A-H4'	T-H3'	A-H3'	T-H1'	A-H1'	T-H6
B	from												
	A-H8	3 <sup>c</sup>	3 <sup>b</sup>	4 <sup>b</sup>	4	3	0	1	1 <sup>b</sup>	2	3 <sup>b</sup>	2	1 <sup>b,c</sup>
	T-H6	3	5	4	3 <sup>b</sup>	3 <sup>b</sup>	1	0	3	1 <sup>b</sup>	2	3 <sup>b</sup>	
	A-H1'	2 <sup>c</sup>	2 <sup>c</sup>	1 <sup>c</sup>	3	3	3 <sup>c</sup>	1	1 <sup>c</sup>	1	0		
	T-H1'	0	3	4	1 <sup>c</sup>	1 <sup>c</sup>	1	3 <sup>c</sup>	2	1 <sup>c</sup>			
	A-H3'	1 <sup>c</sup>	0	1 <sup>c</sup>	4	3	0	4	0				
	T-H3'	0	4	4	0	1 <sup>b</sup>	4	0					
	A-H4'	0	0	1 <sup>b</sup>	1	1	0						
	T-H4'	0	1	1	0	1 <sup>b</sup>							
	A-H2''	3 <sup>c</sup>	2 <sup>c</sup>	1 <sup>b,c</sup>	5								
	A-H2'	3 <sup>c</sup>	1 <sup>b,c</sup>	1 <sup>b,c</sup>									
	T-H2''	1	5										
	T-H2'	1											
alternating B	from												
	A-H8	1 <sup>c</sup>	3 <sup>b</sup>	4 <sup>b</sup>	2	1	0	0	1 <sup>b</sup>	1	3 <sup>b</sup>	1	2 <sup>b</sup>
	T-H6	3	5	5	2 <sup>b</sup>	2 <sup>b</sup>	1	0	2	1 <sup>b</sup>	2	0	
	A-H1'	0	0	0	4	4	0	3	0	2	0		
	T-H1'	0	3	4	0	0	3	0	2	0			
	A-H3'	1 <sup>c</sup>	0	0	4	4	0	3	0				
	T-H3'	0	3	3	0	1 <sup>b</sup>	4	0					
	A-H4'	0	0	0	2	2	0						
	T-H4'	0	1	1	1 <sup>b</sup>	1 <sup>b</sup>							
	A-H2''	1 <sup>c</sup>	1 <sup>c</sup>	1 <sup>c</sup>	5								
	A-H2'	2 <sup>c</sup>	1 <sup>c</sup>	1 <sup>c</sup>									
	T-H2''	1	5										
	T-H2'	2											
wrinkled D	from												
	A-H8	1 <sup>c</sup>	2 <sup>b</sup>	3 <sup>b</sup>	4	3	0	1	1 <sup>b</sup>	2	1 <sup>b</sup>	2	1 <sup>c</sup>
	T-H6	3	4	4	3 <sup>b</sup>	4 <sup>b</sup>	1	0	2	1 <sup>b</sup>	2	3 <sup>b</sup>	
	A-H1'	1 <sup>c</sup>	1 <sup>c</sup>	0	3	4	1 <sup>c</sup>	2	0	2	1 <sup>c</sup>		
	T-H1'	0	3	4	0	0	3	0	2	0			
	A-H3'	1 <sup>c</sup>	0	0	4	4	0	4	0				
	T-H3'	0	4	4	0	0	4	0					
	A-H4'	0	0	0	1	1	0						
	T-H4'	0	1	1	0	0							
	A-H2''	3 <sup>c</sup>	2 <sup>c</sup>	1 <sup>c</sup>	5								
	A-H2'	3 <sup>c</sup>	1 <sup>c</sup>	1 <sup>b</sup>									
	T-H2''	1	5										
	T-H2'	1											
D	from												
	A-H8	2 <sup>c</sup>	3 <sup>b</sup>	4 <sup>b</sup>	4	3	0	1	1 <sup>b</sup>	2	3 <sup>b</sup>	2	1 <sup>b</sup>
	T-H6	3	5	4	2 <sup>b</sup>	3 <sup>b</sup>	1	0	3	1 <sup>b</sup>	2	2 <sup>b</sup>	
	A-H1'	2 <sup>c</sup>	1 <sup>c</sup>	1 <sup>c</sup>	3	4	1 <sup>c</sup>	1	0	2	0		
	T-H1'	0	3	4	1 <sup>c</sup>	0	1	1 <sup>c</sup>	2				
	A-H3'	2 <sup>c</sup>	0	0	4	3	0	4	0				
	T-H3'	0	4	4	0	0	4	0					
	A-H4'	0	0	0	1	1	0						
	T-H4'	0	1	1	0	0							
	A-H2''	4 <sup>c</sup>	1 <sup>c</sup>	1 <sup>c</sup>	5								
	A-H2'	4 <sup>c</sup>	1 <sup>c</sup>	1 <sup>b</sup>									
	T-H2''	1	5										
	T-H2'	1											

<sup>a</sup> Only A(5) and T(6) cross-peak intensities are shown here. Other interior nucleotide residues had essentially the same intensities. See text for intensity grading scale. <sup>b</sup> Cross-peak intensity "from" a nucleotide proton "to" a proton on its 5'-neighbor. <sup>c</sup> Cross-peak intensity "from" a nucleotide proton "to" a proton on its 3'-neighbor.

correlation time was used in subsequent calculations of the 2D NOE spectra reported here, although a limited number of calculations using 3 ns exhibited only slight variations from the 7-ns spectra and did not influence the choice of molecular model which best fits the experimental data.

The theoretical 2D NOE spectra were calculated by using atomic coordinates available from X-ray diffraction studies of DNA in A (Arnott & Hukins, 1972), B (Arnott & Hukins, 1973), alternating B (Klug et al., 1979), left-handed B (Gupta et al., 1980), C (Marvin et al., 1961), D (Arnott & Selsing, 1974), wrinkled D (Arnott et al., 1983), and Z [derived from [d-(CGCGCG)]<sub>2</sub> (Wang et al., 1979)] forms. As these molecular coordinates generally do not include protons, protons were placed in the structures by using standard bond lengths

and angles, yielding a table containing all proton coordinates for a given molecular model. All of the interproton distances for theoretical 2D NOE spectral calculations (see Theory) are determined from the proton coordinates of the model.

The calculated intensities were classified in the same way as the observed intensities. The results are given for a mixing time of 100 ms in Table III for the case of regular B, alternating B, regular D, and wrinkled D (wD) form models, which come closer to fitting the experimental data than other models tried (vide infra).

**Comparison between Experimental and Theoretical 2D NOE Spectra.** Comparison was made between observed and theoretical 2D NOE intensity matrices of the decamer. For convenience, a total difference index (TDI) between two in-



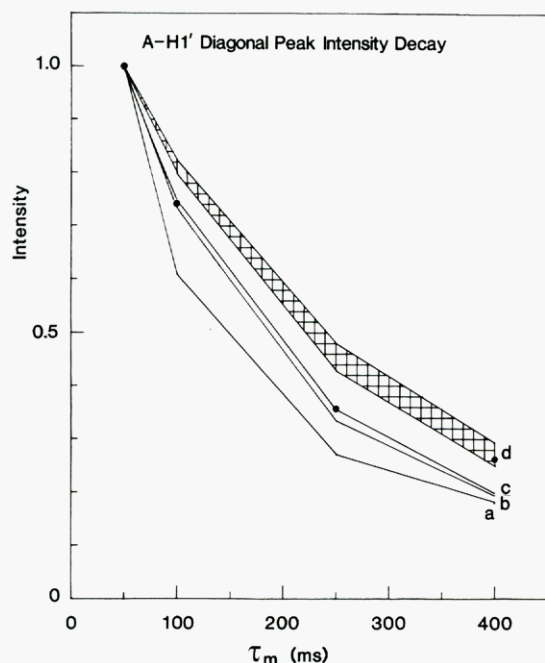


FIGURE 5: Theoretical decay curves for the A(3)-H1 proton for several models compared with observed data (●): (a) regular B; (b) left-handed B; (c) D; (d) all seven other models (A, C, wrinkled D, alternating B, energy-minimized B, energy-minimized D, energy-minimized wD) within this region.

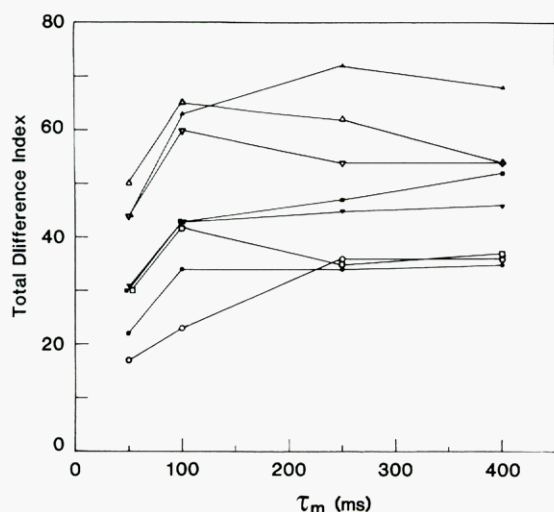


FIGURE 6: Total difference index (TDI) between calculated and observed 2D NOE spectra as a function of mixing time for well-known DNA structural models: A (Δ); B (□); alternating B (■); left-handed B (▽); C (▼); D (●); wD (○); Z (▲) form DNA models.

tensity matrices (one experimental and the other theoretical) was introduced:

$$TDI(\alpha, \beta, \tau_m) = \sum_{i,j} |I_{ij}(\alpha) - I_{ij}(\beta)|$$

$I_{ij}$  is the intensity grade for an observed or calculated 2D NOE cross peak occurring between nuclei  $i$  and  $j$ , and  $\alpha$  and  $\beta$  refer to observed or calculated 2D NOE spectra, respectively. The TDI may depend on  $\tau_m$ . Of course, a lower value for TDI is indicative of a better fit of the structural model to the experimental data. For example, Table IV shows how the TDI value was obtained for the case of the wD model at  $\tau_m = 50$  ms.

The TDI values are reported in Table V and plotted in Figure 6 for all well-known molecular models at the four mixing times. From these, it was recognized that Z, A, C, left-handed B, and alternating B models can be clearly excluded and that the wD form is the best one, although regular

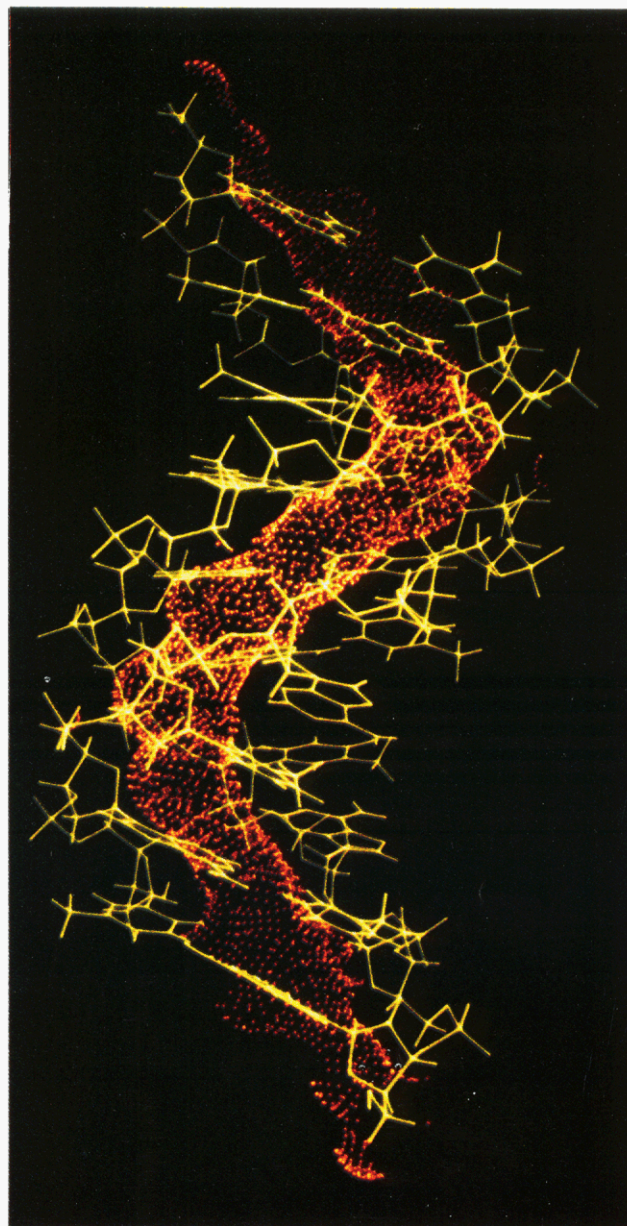


FIGURE 7:  $[d(5'ATATATATAT3')]_2$  fragment depicted as wrinkled D-form DNA with a hydration tunnel illustrated by the red molecular surface representation.

B and D form models are also possible. With left-handed B, Z, and A form models, the TDI values are always quite high and indicate little coincidence between experimental and theoretical 2D NOE spectral patterns. The TDI values of C and alternating B are almost the same level as that of B form at the shorter mixing times (50, 100 ms) but are much worse for the longer mixing times (250, 400 ms). The TDI values of regular D and wD forms in contrast with regular B form are much better for the shorter mixing times and are comparable for the longer mixing times.

As intensities were roughly classified, it is thought that the TDI value of the true structure should be no more than 20. To get better fits, it may be necessary to account for local motions rather than assume the motion is simply isotropic (James, 1984; Keepers & James, 1984). So, the wD form may be the best model for alternating d(A-T) DNA from the various proposed X-ray structures in low salt solution, as the TDI value of this model at shorter mixing times is smaller than those of the others. Of course, we recognize the possibility that the solution structure is not completely fit by any of the



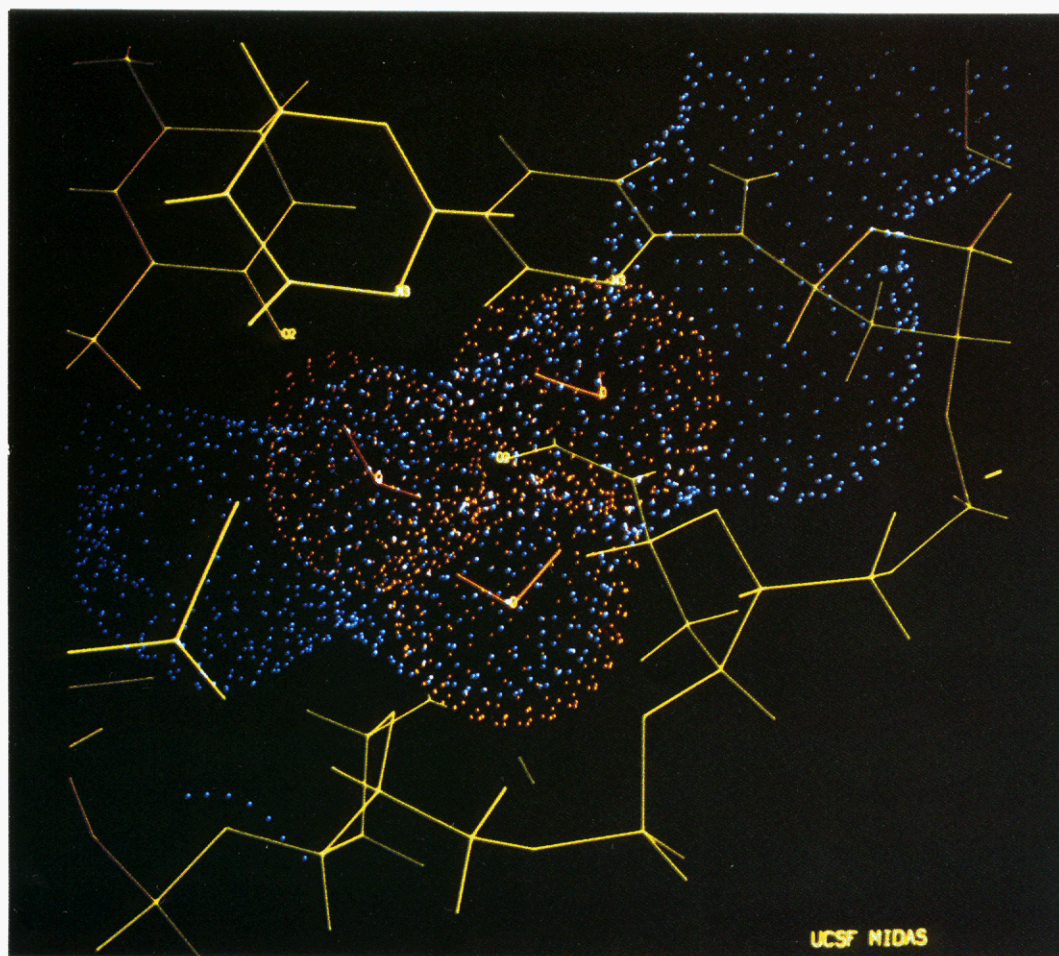


FIGURE 8: A portion of the A-T decamer as wrinkled D form with the molecular surface of the hydration tunnel colored blue and with three water molecules with van der Waals radii in red.

X-ray structures. So we will use the most promising X-ray structures and manipulate them in a reasonable manner to see if further insights into the solution structure may be obtained. We note that if a correlation time of 3 ns is used, the TDI values for wrinkled D, regular D, and regular B forms come slightly closer. For comparison, regular B and D form models were also included as starting models in the following molecular mechanics calculations using AMBER, which gives us refined structural coordinates following energy minimization on the starting structure. Both D forms belong to the B family of DNA structures, but fairly large differences exist in the calculated spectral patterns, so we should be able to distinguish between them and between any energy-refined structures.

**Molecular Mechanics Calculations and Energy-Minimized Structures.** Molecular mechanics calculations were carried out on the double-helical structures of regular B, D, and wD models with ten residues per strand with the program AMBER, Assisted Model Building with Energy Refinement (Weiner & Kollman, 1981; Weiner et al., 1984a). The partial atomic charges used in our calculations were taken from Singh and Kollman (1984), and the various constants to evaluate the energy were from Weiner et al. (1984b). The structures were refined until the root mean square gradient was less than 0.1 kcal/Å. The initial energy refinements were performed without explicit inclusion of counterions or solvent by using a distance-dependent dielectric constant to mimic some medium effects as described previously (Kollman et al., 1981). The energy-minimized structures were displayed and compared on color graphics display PS2 by using the program MIDAS (Gallo et al., 1985). However, the TDI values between the

energy-minimized models' theoretical spectra and observed spectra are all around 33 (at  $\tau_m = 100$  ms) for each of the energy-minimized structures, in which the narrow minor grooves of D and wD models became wider. Furthermore, no energy-minimized structure yielded a lower TDI value than did the original wD form. Interestingly, the TDI values among these calculated theoretical spectra themselves are quite small (under 10), and the backbone torsion angles (structural parameters) are also very similar. In other words, energy minimization of B, D, and wD forms yielded nearly the same structure regardless of the initial structure. A most important point is that the original wD form is closer to these energy-minimized models in all parameters except for the  $\zeta(O3'-P)$  value than any of the other X-ray structures; the  $\beta(O5'-C5')$ ,  $\gamma(C5'-C4')$ , and  $\epsilon(C3'-O3')$  parameters are especially close. This also suggests that the most stable form of alternating d(A-T) in solution is similar to the wD form.

By using molecular (solvent-accessible) surface analysis and examining the structures via computer graphics, we have considered why the original wD form is more favorable in solution. It was found for wD-DNA that possible interstrand hydrophobic interactions exist between 4' atoms (carbon and hydrogen) on opposing strands and between 5' atoms on opposing strands. An additional feature of the wD structure is that the minor groove is sufficiently narrow and deep that it effectively forms a tunnel that can accommodate water molecules (Figure 7). After placement of two water molecules (between N3 and N3 of opposing strand adenines and between O2 and O2 of opposing strand thymines), a cavity of almost the same size as that of one more water molecule remains



Table IV: Comparison of Calculated and Observed 2D NOE Peak Intensities for the Wrinkled D Form of DNA at 50-ms Mixing Time<sup>a</sup>

	to											
	T-Me	T-H2'	T-H2''	A-H2'	A-H2''	T-H4'	A-H4'	T-H3'	A-H3'	T-H1'	A-H1'	T-H6
(A) Calculated												
from												
A-H8	1 <sup>c</sup>	1 <sup>b</sup>	3 <sup>b</sup>	4	2	0	0	0	1	1 <sup>b</sup>	1	0
T-H6	3	4	3	2 <sup>b</sup>	4 <sup>b</sup>	0	0	1	0	1	2 <sup>b</sup>	
A-H1'	1 <sup>c</sup>	0	0	3	4	0	1	0	1	0		
T-H1'	0	3	4	0	0	2	0	1	0			
A-H3'	0	0	0	3	3	0	3	0				
T-H3'	0	3	3	0	0	3	0					
A-H4'	0	0	0	1	1	0						
T-H4'	0	1	1	0	0							
A-H2''	3 <sup>c</sup>	1 <sup>c</sup>	0	5								
A-H2'	2 <sup>c</sup>	0	0									
T-H2''	0	5										
T-H2'	0											
(B) Difference between Observed and Calculated Cross-Peak Intensities												
from												
A-H8	1	0	-1	-1	0	0	0	0	0	0	1	0
T-H6	-1	0	-1	0	-1	0	0	0	0	0	0	
A-H1'	-1	0	0	0	0	0	1	0	0	0		
T-H1'	0	-1	0	0	0	-1	0	0	0			
A-H3'	0	0	0	0	-1	0	0	<i>d</i>				
T-H3'	0	0	-1	0	0	<i>d</i>	0					
A-H4'	0	0	0	0	0	<i>d</i>						
T-H4'	0	0	0	0	0							
A-H2''	-2	-1	0	0								
A-H2'	-1	0	1									
T-H2''	0	0										
T-H2'	0											

$$\text{TDI} = [1](15) + [2](1) = 17$$

<sup>a</sup>Only A(5) and T(6) cross-peak intensities are shown here. Other interior nucleotide residues had essentially the same intensities. See text for intensity grading scale. <sup>b</sup>Cross-peak intensity "from" a nucleotide proton "to" a proton on its 5'-neighbor. <sup>c</sup>Cross-peak intensity "from" a nucleotide proton "to" a proton on its 3'-neighbor. <sup>d</sup>Not resolved.

Table V: Total Difference Index (TDI) Values as a Function of Mixing Time for Several DNA Models

model <sup>a</sup>	mixing time (ms)			
	50	100	250	400
regular A	50	65	62	54
regular B	30	42	35	37
alternating B	30	43	47	52
left-handed B	44	63	72	68
regular C	31	43	45	46
regular D	22	34	34	36
wrinkled D	17	23	36	35
Z	44	60	54	54

<sup>a</sup>See the text for references to descriptions of these DNA structural forms.

between those two water molecules in the minor groove. Figure 8 shows the placement of the water molecules. Such a hydration tunnel was not found with the regular D (minor groove too shallow) or B form (minor groove too large).

We considered why these features could not be maintained in the energy minimization from the starting wrinkled D structure; it was suggested by group analysis for each residue (sugar, phosphate, and base) that phosphate-phosphate repulsion forces a widening of the minor groove during energy minimization. Subsequently, 17 water molecules (for the middle part of the decamer, two per base pair) were put into the minor groove (vide supra), and the counterion distribution program (Pattabiraman et al., 1984; S. Devarajan and N. Pattabiraman, FORTRAN Program To Calculate Iteratively Ionic Positions around DNA, unpublished results), in which the ions were placed iteratively by calculating the electrostatic potential, was used for obtaining adequate ion locations. AMBER calculations were carried out for wD and, for comparison, regular B models containing the water molecules and

ions. The result of this calculation was that the energy-minimized wD form at -2177 kcal is somewhat more stable than the energy-minimized B form (-2148 kcal). The calculated 2D NOE spectral features of the energy-minimized wD model with ions and water is almost the same as the original one. On the contrary, the features of the energy-minimized B model differ from the original and are rather similar to those of the wD model. The structural parameters (backbone and glycosidic torsion angles) of both energy-minimized models are also very close to those of the wD X-ray structure, in contrast with regular B-DNA. This observation is the same as with the initial energy refinement described above, but the characteristics of the minor groove (deep and narrow) were well maintained during structure refinement. These structures and the initial (original) structures are illustrated in Figure 9 as stereo diagrams; their structural parameters are shown in Figure 10 and Table VI. The following aspects of the structures were found by group analysis: (1) the intra- and interstrand phosphate-phosphate repulsive effect is larger for the wD than the B structures, but these effects are fully reduced by sodium ions; (2) intra- and interstrand phosphate-sugar attractive interactions and intraresidue sugar interactions are better for the wD form; and (3) the role of hydration water and hydrophobic interactions are not entirely clear from these static calculations, but the manner in which the water molecules fit so well into the hydration tunnel of wD-DNA with formation of stabilizing hydrogen bonds is certainly appealing. Elucidation of the potential role of these features awaits dynamics calculations in the aqueous medium.

It can be concluded that the wD form of DNA may be energetically more stable than the B form and is the most promising model. The detailed features of the molecular mechanics calculation will be published elsewhere.

*On the Structure of Poly(dA-dT).* There are three reported

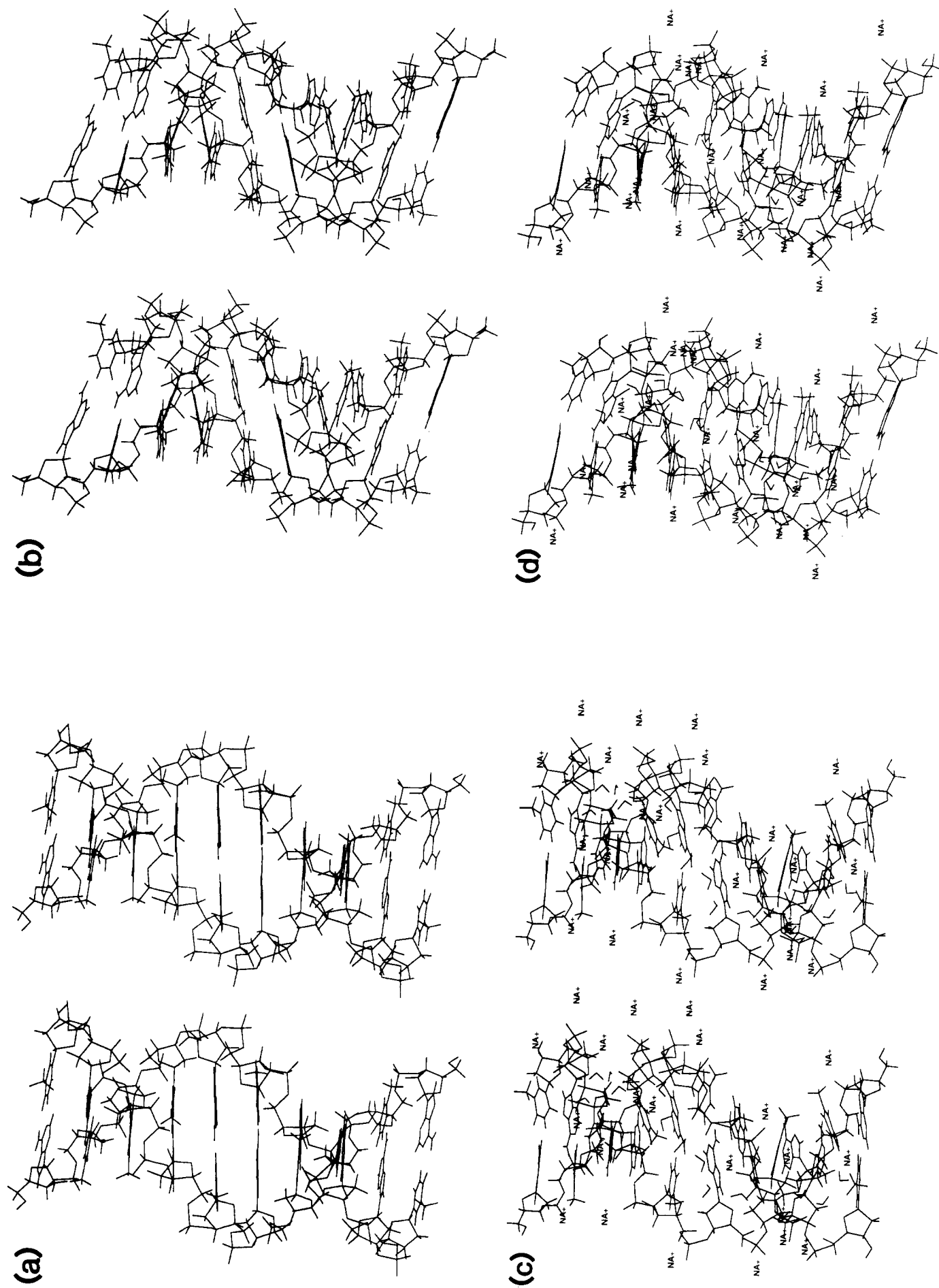


FIGURE 9: Stereo diagrams for B and wD models: (a) original B; (b) original wD; (c) energy-minimized wD; (d) energy-minimized B. Parts c and d include sodium ion and water molecules.

Table VI: Backbone and Glycosyl Torsion Angles for Several DNA Models

model	structure parameters (deg) <sup>a</sup>							<i>n</i> <sup>d</sup>	<i>h</i> (nm) <sup>e</sup>	reference
	$\alpha$	$\beta$	$\gamma$	$\delta$	$\epsilon$	$\zeta$	$\chi$			
regular B	314	214	36	156	155	264	82	10.0	0.337	Arnott & Hukins (1973)
energy-minimized B	290	163	64	122	187	258	55	9.9	0.336	this work <sup>f</sup>
alternating B (A-T) <sup>b,c</sup>	292	172	65	144	192	278	73	10.0	0.340	Klug et al. (1979)
alternating B (T-A) <sup>c</sup>	300	151	59	99	200	226	33			
regular C	315	143	48	141	211	212	83	9.3	0.332	Marvin et al. (1961)
regular D	298	208	69	156	141	260	82	8.0	0.303	Arnott & Selsing (1974)
wrinkled D (A-T) <sup>c</sup>	315	142	55	142	203	215	65	8.0	0.302	Arnott et al. (1983)
wrinkled D (T-A) <sup>c</sup>	288	143	69	154	227	207	78			
energy-minimized wrinkled D (A-T) <sup>c</sup>	287	161	61	107	207	255	55	8.7	0.299	this work <sup>f</sup>
energy-minimized wrinkled D (T-A) <sup>c</sup>	290	151	70	146	219	221	74			

<sup>a</sup> Definition of torsion angles are the same as in Seeman et al. (1976) except for  $\chi$ , which is defined as O4'-C1'-N1-C6(T) and O4'-C1'-N9-C6(A) in this study. <sup>b</sup> Only this is C3'-endo; others are C2'-endo. <sup>c</sup> Residues A and T (including phosphates) have different parameter values. <sup>d</sup> Base pairs per turn. <sup>e</sup> Rise per base pair. <sup>f</sup> Averaged for all residues.

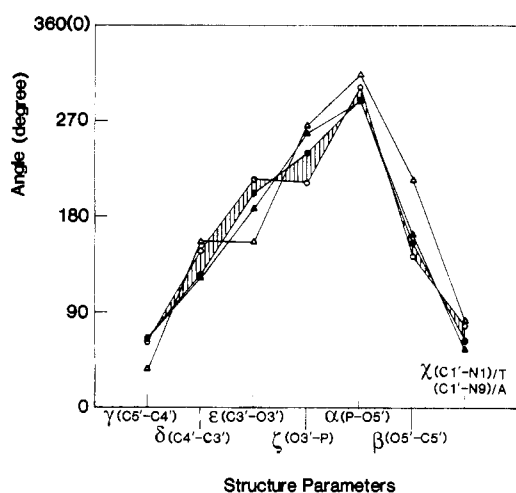


FIGURE 10: Structural parameters (backbone torsion angles) of the original and energy-minimized B- and wD-form DNA models, averaged for all residues: original wD (○), original regular B (△), energy-minimized wD (●), and energy-minimized B models (▲).

studies on the structure of double-helical poly(dA-dT) using NMR methods (Gupta et al., 1983; Assa-Munt & Kearns, 1984; Borah et al., 1985). The conclusions of two studies (Assa-Munt & Kearns, 1984; Borah et al., 1985) are the same, i.e., that poly(dA-dT) exists in the B form. The other study concluded that poly(dA-dT) is in the left-handed B form (Gupta et al., 1983). As pointed out previously (Assa-Munt & Kearns, 1984; Borah et al., 1985), this notion is easily rejected as it is based only on the observed weak interaction between T-H6 and T-H2', H2'', ignoring the many other interactions. All experimental NOE features are important and useful; of course, 2D NOE should preferably be used for such a structure study. Regular B-, D-, and wD-DNA structures are in the B family of structures and are consequently similar if not identical. As shown by the present studies, the 2D NOE spectra of B, D, and wD forms of DNA are also similar, but they can be distinguished. Assa-Munt and Kearns (1984) and Borah et al. (1985) each used poly(dA-dT) approximately 50 bp long under similar experimental conditions. Our use of the decamer with narrower line widths provides us with some advantage. But our use of the entire relaxation matrix and subsequent molecular mechanics calculations lead us to distinguish among the B family of structures, with the wrinkled D form being best supported by the NMR results and the molecular mechanics calculations.

**A-H2 Interactions with A-H2.** Cross peaks between A(9)- and A(3)- and between A(7)- and A(5)-H2 protons were clearly observed (vide supra, Figure 4). Similar cross peaks

have been observed in the case of poly(dA-dT) (Assa-Munt & Kearns, 1984), but the residue numbers may be specified in this study, revealing the interaction to be between the A-H2 proton in the 5'-TA-3' sequence of one strand and that of the other strand of the duplex. It is suggested that this observation is related to the somewhat weaker (longer distance) interaction observed between A-H8 and its 5'-neighboring T-H2'. At the same time, it may be related to a stronger interaction (shorter distance) between A-H8 and its 3'-neighboring T-Me. Pertinent to this point, the stacking energy calculated for the energy-minimized wD model is lower in 5'-AT-3' (-17.4 kcal) than in the 5'-TA-3' (-14.2 kcal) sequence; this alternation was not found for the energy-minimized B form.

These points are understandable by the general idea of positive propeller twisting. A-H8 is located on the major groove side, while A-H2 is on the minor groove side. The sequence 5'-pyrimidine-purine-3' leads to clashes in the minor groove (A-H2 side), with the consequence that the A-H2-A-H2 distance becomes shorter and the A-H8-T-H2' distance becomes longer. The sequence 5'-purine-pyrimidine-3' leads to clashes in the major groove (A-H8 side) leading to longer A-H2-A-H2 distances and shorter A-H8-T-Me distances. This is the extreme case of positive propeller twist, which in reality will probably be reduced as Calladine prescribed (Calladine, 1982). In the present case, it may not be reduced substantially though due to the stabilization by hydration water molecules bridging between adenine N3 atoms on opposing strands in the minor groove, which would make the A-H2-A-H2 distance shorter in the 5'-TA-3' sequence region. This is predicted for the wrinkled D form.

**Hydrophobic Inter-Sugar Interaction.** Some interstrand cross peaks may be expected for sugar protons of the wD model. Major interactions are between A- or T-H4' of one strand and that of another, and between H5',5'' protons. But these are difficult to observe as H5',5'' and T-H4' proton resonances are located in a crowded region of the spectrum (ca. 4.2 ppm), and all A-H4' chemical shifts are nearly the same except for the terminal.

## CONCLUSIONS

In conclusion, it is strongly suggested that [d(A-T)<sub>5</sub>]<sub>2</sub> exists in low salt solution as a *wrinkled D form*, which belongs to the B family but is quite distinct from regular B-DNA, by studying pure absorption proton 2D NOE spectra as a function of mixing time with theoretical 2D NOE spectral calculations carried out on the basis of the complete relaxation matrix method and molecular energy minimization using AMBER including all hydrogens on the DNA, counterions, and water molecules. In this study, the importance of analyzing all

features of the 2D NOE spectra for extracting a possible structure of DNA was emphasized, and the total difference index between the experimental and theoretical 2D NOE spectra was used. Features, such as *phosphate-sugar attractive forces, inter-sugar interactions, a hydration tunnel, and the role of ions*, which were suggested from the molecular mechanics calculations and computer graphics, may be significant for alternating d(A-T) sequences in DNA.

#### ACKNOWLEDGMENTS

We thank Dr. Vladimir J. Basus for assistance in obtaining NMR data, Drs. Sadasivam Manogran and Ruud Scheek for the NMR data processing program arrangement, and Dr. Gregory B. Young for initial aid in theoretical calculations of 2D NOE spectra. We acknowledge our grateful use of the UCSF Computer Graphics Laboratory (Dr. R. Langridge, Director; supported by NIH Grant RR 01081).

**Registry No.** d(5'ATATATATAT3'), 85240-23-9; Na, 7440-23-5.

#### REFERENCES

- Arnott, S., & Hukins, D. W. L. (1972) *Biochem. Biophys. Res. Commun.* **47**, 1504-1509.
- Arnott, S., & Hukins, D. W. L. (1973) *J. Mol. Biol.* **81**, 93-105.
- Arnott, S., & Selsing, E. (1974) *J. Mol. Biol.* **88**, 509-521.
- Arnott, S., Chandrasekaran, R., Puigjaner, L. C., Walker, J. K., Hall, I. H., Birdsall, D. L., & Ratliff, R. L. (1983) *Nucleic Acids Res.* **11**, 1457-1474.
- Assa-Munt, N., & Kearns, D. R. (1984) *Biochemistry* **23**, 791-796.
- Bax, A. (1982) *Two-Dimensional Nuclear Magnetic Resonance in Liquids*, Reidel Publishing, Dordrecht, The Netherlands.
- Borah, B., Cohen, J. S., & Bax, A. (1985) *Biopolymers* **24**, 747-765.
- Broido, M. S., Zon, G., & James, T. L. (1984) *Biochem. Biophys. Res. Commun.* **119**, 663-670.
- Broido, M. S., James, T. L., Zon, G., & Keepers, J. W. (1985) *Eur. J. Biochem.* **150**, 117-128.
- Calladine, C. R. (1982) *J. Mol. Biol.* **161**, 343-352.
- Cavaillès, J. A., Neumann, J.-M., Tran-Dinh, S., Huynh-Dinh, T., d'Estaintot, B. L., & Igolen, J. (1985) *Eur. J. Biochem.* **147**, 183-190.
- Davies, D. R., & Baldwin, R. L. (1963) *J. Mol. Biol.* **6**, 251-255.
- Drew, H. R., & Dickerson, R. E. (1982) *EMBO J.* **1**, 663-667.
- Feigon, J., Leupin, W., Denny, W. A., & Kearns, D. R. (1983) *Biochemistry* **22**, 5930-5942, 5943-5951.
- Frechet, D., Cheng, D. M., Kan, L.-S., & Ts'o, P. O. P. (1983) *Biochemistry* **22**, 5194-5200.
- Gallo, L., Huang, C., Ferrin, T., & Langridge, R. (1985) *User's Manual, Molecular Interactive Display and Simulation (MIDAS)*, Computer Graphics Laboratory, University of California, San Francisco.
- Gupta, G., Bansal, M., & Sasisekharan, V. (1980) *J. Biol. Macromol.* **2**, 368-380.
- Gupta, G., Sarma, M. H., Dhingra, M. M., Sarma, R. H., Rajagopalan, M., & Sasisekharan, V. (1983) *J. Biomol. Struct. Dyn.* **1**, 395-416.
- Havel, T. F., Kuntz, I. D., & Crippen, G. M. (1983) *Bull. Math. Biol.* **45**, 665-720.
- Hilbers, C. W., Haasnoot, C. A. G., de Bruin, S. H., Joordens, J. J. M., van der Marel, G. A., & van Boom, J. H. (1985) *Biochimie* **67**, 685-695.
- James, T. L. (1984) in *Phosphorus-31 NMR: Principles and Applications* (Gorenstein, D. G., Ed.) pp 349-400, Academic, Orlando, FL.
- Jamin, N., James, T. L., & Zon, G. (1985) *Eur. J. Biochem.* **152**, 157-166.
- Kearns, D. R. (1984) *CRC Crit. Rev. Biochem.* **15**, 237-290.
- Keepers, J. W., & James, T. L. (1984) *J. Magn. Reson.* **57**, 404-426.
- Klug, A., Jack, A., Viswamitra, M. A., Kennard, O., Shakked, Z., & Steitz, T. A. (1979) *J. Mol. Biol.* **131**, 669-680.
- Kollman, P. A., Weiner, P. K., & Dearing, A. (1981) *Biopolymers* **20**, 2583-2621.
- Kumar, A., Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun.* **95**, 1-6.
- Lai, K., Shah, D. O., DeRose, E., & Gorenstein, D. G. (1984) *Biochem. Biophys. Res. Commun.* **121**, 1021-1026.
- Macura, S., & Ernst, R. R. (1980) *Mol. Phys.* **41**, 95-117.
- Macura, S., Huang, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson.* **43**, 259-281.
- Marvin, D. A., Spencer, M., Wilkins, M. H. F., & Hamilton, L. D. (1961) *J. Mol. Biol.* **3**, 547-565.
- Pattabiraman, N., Langridge, R., & Kollman, P. A. (1984) *J. Biomol. Struct. Dyn.* **1**, 1525-1533.
- Scheek, R. M., Boelens, R., Russo, N., van Boom, J. H., & Kaptein, R. (1984) *Biochemistry* **23**, 1371-1376.
- Seeman, N. C., Rosenberg, J. M., Suddath, F. L., Park Kim, J. J., & Rich, A. (1976) *J. Mol. Biol.* **104**, 142-143.
- Singh, U. C., & Kollman, P. A. (1984) *J. Comput. Chem.* **5**, 129-149.
- States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson.* **48**, 286-292.
- Wang, A. H.-J., Quigley, G. J., Kolpak, F. J., Crawford, J. L., van der Marel, G., & Rich, A. (1979) *Nature (London)* **282**, 680-686.
- Weiner, P. K., & Kollman, P. A. (1981) *J. Comput. Chem.* **2**, 287-303.
- Weiner, P. K., Singh, U. C., Kollman, P. A., Caldwell, J., & Case, D. A. (1984a) *A Molecular Mechanics and Dynamics Program*, University of California, San Francisco.
- Weiner, S., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., & Weiner, P. K. (1984b) *J. Am. Chem. Soc.* **106**, 765-784.
- Wemmer, D. E., Chou, S. H., Hare, D. R., & Reid, B. R. (1985) *Nucleic Acids Res.* **13**, 3755-3772.
- Woessner, D. E. (1962) *J. Chem. Phys.* **36**, 1-4.
- Young, G. B., & James, T. L. (1984) *J. Am. Chem. Soc.* **106**, 7986-7988.