Effect of Adenine Methylation on the Structure and Dynamics of TpA Steps in DNA: NMR Structure Determination of $[d(CGAGGTTTAAACCTCG)]_2$ and Its A9-Methylated Derivative at 750 MHz†

Jody Lingbeck,‡ Mark G. Kubinec,‡ Julie Miller,§ Brian R. Reid,|| Gary P. Drobny,§ and Michael A. Kennedy*,‡

Environmental Molecular Sciences Laboratory, P7-55, Richland, Washington 99352, and Departments of Chemistry and Biochemistry, University of Washington, Seattle, Washington 98195

Received June 16, 1995; Revised Manuscript Received September 5, 1995⊗

ABSTRACT: At TpA steps in DNA, the adenine base experiences exceptionally large amplitude (20°−50°) and slow (10 ms−1 μs) motion [Kennedy *et al.* (1993) *Biochemistry 32,* 8022−8035] which has been correlated with transitions between multiple conformational states [Lefevre *et al.* (1985) *FEBS Lett. 190,* 37−40]. The base dynamics can be detected in one-dimensional ^{1}H NMR spectra as excess line width of the aromatic proton resonances. The magnitude of the excess line width is temperature dependent and reaches a maximum at some temperature. In order to better understand the origin of the dynamics, we have studied the effect of N^6-methylation of the TpA adenine on both the line widths and its local structure. Here, solution-state 500 and 750 MHz ^{1}H NMR data collected on $[d(CGAGGTTTAAACCTCG)]_2$ show that the excess line width of theTpA adenine-H2 is diminished when the TpA adenine is N^6-methylated and that the line width no longer experiences a maximum as the temperature is varied. The resonances sharpen upon methylation because the amplitude of base motion is restricted due to steric effects and due to other structural changes at the TpA site. Additionally, both the TpA adenine-H8 and the exchangeable imino resonance of thymine at the TpA step were also found to have excess line width that is diminished upon N^6-methylation. In order to elucidate the structural features responsible for TpA base dynamics, solution-state NMR structures of $[d(CGAGGTTTAAACCTCG)]_2$ and its A9 N^6-methylated derivative were determined at 750 MHz. Comparison of the structures shows that poor base stacking at the TpA step may contribute to, or be the origin of, its base dynamics.

Conformational transitions localized at the TpA steps in DNA were first reported in the Pribnow box in the 20 base pair *trp* DNA promoter, CGTACTAGTTAACTAGTACG, by LeFevre *et al.* (1985). The dynamics were noticed as excess line width of the TpA adenine-H2 resonance and by nonmonotonic changes in the line width and chemical shift of the TpA adenine-H8 proton as the temperature was raised from 0 to 45 °C. In subsequent reports, the structure and conformational transitions for both the native *trp* promoter (Fazakerley *et al.*, 1984, 1985, 1987; Lefevre *et al.*, 1987, 1988) and a mutant *trp* promoter (Lane, 1989, 1991) were described. The conformational dynamics were interpreted as a population distribution over at least three conformational states. The two transitions were proposed to involve changes in the propeller twist of A11 and A12 for one transition and the helical twist and local pitch for the second transition. The thermodynamics of the transitions, including rates for transition between states I and II (200 s^{-1}) and between states II and III (250 s^{-1}), have been reported for the native promoter. The combined analysis of these data led to the proposition that the temperature dependence of conformational transitions was related to the initiation of transcription at the *trp* promoter.

We have reported conformational dynamics similar to those found in the *trp* promoter in three DNA sequences, CGAGGTTTAAACCTCG, GCTCCTTTAAAGGAGC, and GCCGTTAACGGC (Kennedy *et al.*, 1993). The sequences contain restriction sites of the form T_nA_n, e.g., the TTTAAA *Aha*III site. In each restriction site, enzymatic blunt-end cleavage of the phosphodiester backbone occurs precisely at the site of unusual base dynamics, i.e., at the TpA step. On the basis of our data, we proposed a model for the conformational dynamics in which the TpA adenine base and possibly the base-paired thymine undergo a continuous and possibly concerted large amplitude (20°−50°), slow (τ_c = 10 μs−1 ms) oscillation about the glycosidic bond in an otherwise static helix. Some structural features of the TpA site were inferred from our previous studies. For example, atypical NOESY cross-peak intensities in the aromatic to H1′ sequential walk region were observed at the TpA step where the (*n*)A9-H8 to (*n*)A9-H′ cross-peak was more intense than the (*n*)A9-H8 to (*n* − 1)T8-H1′ cross-peak (a pattern which is the reverse of that observed in B DNA and to all other cross-peak pairs in this sequence). We interpreted this cross-peak pattern as an indication of unusual base stacking at the TpA step. Furthermore, a striking discontinuity in

† This work was supported by National Institutes of Health Grant GM 32681. M.A.K. was supported by a Northwest College and University Association for Science Fellowship administered by Washington State University under DOE Grant DE-FG06-89ER-75522 and by an Alexander Hollaender Distinguished Postdoctoral Fellowship sponsored by the DOE Office of Health and Environmental Research and administered by the Oak Ridge Institute for Science and Education. The research performed by M.G.K. and J.L. was supported by the Associated Western Universities, Inc., Northwest Division under Grant DE-FG06-89ER-75522 with the U.S. Department of Energy.

* To whom correspondence should be addressed.

‡ Environmental Molecular Sciences Laboratory.

§ Department of Chemistry, University of Washington.

|| Department of Biochemistry, University of Washington.

⊗ Abstract published in *Advance ACS Abstracts,* December 15, 1995.

0006-2960/96/0435-719$12.00/0

the sugar conformations at the TpA step [T8 (east)-T9 (south)] was deduced on the basis of large differences in the intensity of H3′-H4′ DQF-COSY cross-peaks. There have been limited reports of DNA structures for sequences containing T_nA_n segments (Lefevre *et al*., 1987; Gupta *et al*., 1988; Kim & Reid, 1992; Kim *et al*., 1992). From these reports, the conclusions about the structures are inconsistent, varying from sequences in which half the sugars assume a "north" conformation (Lefevre *et al*., 1987), to structures in which no clear "junction" occurs at the TpA step (Gupta *et al*., 1988), to structures which have clear patterns of discontinuities in structure at the TpA step and all the sugars assume the "south" conformation (Kim & Reid, 1992; Kim *et al*., 1992). However, only in the work of Lefevre *et al.* (1987) were the structures solved and analyzed with knowledge of the base dynamics that occur at the TpA step. Hence, there is still no clear understanding of what structural features give rise to base dynamics and what structural features are characteristic of a TpA step in DNA.

We are currently interested in determining the structural origin of base dynamics at TpA steps in DNA and in establishing a complete understanding of when dynamics at TpA steps can be expected. Until recently the dynamics had only been observed in a single sequence context, i.e., TTAA, raising the possibility that the TpA dynamics were an anomaly associated with a single sequence context environment. One of our goals has been to determine if the base dynamics observed at TpA steps in DNA occur in more general sequence contexts. We have recently shown that the base dynamics occur when either a thymine precedes the TpA step or when an adenine follows the TpA step (McAteer *et al.*, 1995). In a more complete study of the sequence context effects on base dynamics at TpA steps, we have observed that the dynamics occur at the TpA step in all 16 possible sequence context environments (submitted for publication). Therefore, the base dynamics observed at TpA steps appear to represent a fundamental structural feature of DNA rather than a single anomaly of a particular sequence context. Therefore, it is important to probe the structural origin of TpA base dynamics. Here we report on how the dynamics and the structure surrounding the TpA step in CGAGGTTTAAACCTCG are affected when the TpA adenine is N^6-methylated. We show that the excess TpA adenine-H2 line width that is indicative of base dynamics is significantly diminished by N^6-methylation. We also report the structure of two DNA sequences, CGAGGTTTAAACCTCG and CGAGGTTT(mA)AAACCTCG, determined by NMR at 750 MHz. The structure surrounding the TpA step in the unmethylated sequence is distinct compared to the unmethylated sequence. These structural observations help to understand the unusual base dynamics at the TpA steps in DNA.

MATERIALS AND METHODS

Sample Preparation. The DNA oligomers used for structural studies (1) $[d(5'\text{-CGAGGTTTAAACCTCG-}3')]_2$ and (2) $[d(5'\text{-CGAGGTTT(mA)ACCTCG-}3')]_2$ (where mA indicates N^6-methylation) were synthesized on a 10 μmol scale. Other DNA oligonucleotides used for line width studies, (1) $[d(5'\text{-CGAGGTTTA(mA)ACCTCG-}3')]_2$, (2) $[d(5'\text{-CG AGGTTTAA(mA)CCTCG-}3')]_2$, (3) $[d(5'\text{-CGAGGTTT(mA)(mA)ACCTCG-}3')]_2$, (4) $[d(5'\text{-CGAGGTTTA(mA)(mA)CCTCG-}3')]_2$, (5) $(5'\text{-GCCGTTAACGGC-}3')_2$, (6) $[5'\text{-GCCGTT(mA)ACGGC-}3']_2$, (7) $[5'\text{-GCCGTTA(mA)CGGC-}3']_2$, (8) $[5'\text{-GCCGTTAA(mC)GGC-}3']_2$, (9) $[5'\text{-GCCGTT(mA)(mA)CGGC-}3']_2$, and (10) $[5'\text{-GCCGTTA(mA)(mC)GGC-}3']_2$, were synthesized on a 1 μmol scale. All DNA samples were prepared by the solid-state triester method using cyanoethyl phosphoramidites on an Applied Biosystems 380A DNA synthesizer. The samples were purified either by Sephadex G-50 column and homogeneous fractions, as monitored by gel electrophoresis, and were pooled and lyophilized or by reverse-phase HPLC and flow dialysis. The samples were dissolved either in 0.4 or 0.6 mL of phosphate buffer (pH 7.0) with a composition of 400 mM NaCl, 41 mM sodium phosphate, and 0.5 mM EDTA for NMR studies.

Melting Analysis. The duplex melting temperature of the 16 base pair DNA fragments was determined by measuring the absorbance at 260 nm in 20 mM phosphate buffer (pH = 7)−200 mM NaCl solution containing 0.7−1.0 OD of duplex DNA as a function of temperature between 20 and 90 °C. The buffer solution was degassed under vacuum using a water aspirator immediately before the DNA solution was prepared. A Hitachi 100-80 spectrophotometer was used with a dual water-jacketed cuvette holder employing a Haake Model KT2 refrigerated bath using ethylene glycol as the coolant solution. The temperature of the bath was manually stepped while silica gel dried N_2 gas purged the optical chamber during the experiment. The absorbance at 260 nm was recorded at 1 °C intervals after the thermocouple reading did not change for at least 45 s. On average, 3−5 min was required between steps before a temperature equilibrium was achieved. The melting curves were recorded two times for each sample, and the data from independent data sets were superimposable with excellent agreement.

NMR Spectroscopy. NMR experiments were performed on Varian Unity Plus spectrometers operating at 750 or 500 MHz, on a Bruker AM-500 spectrometer, or on a home-built 500 MHz NMR spectrometer (Gladden and Drobny, unpublished). One-dimensional spectra were all collected using a minimum of 8192 complex points with an acquisition time of 0.66 s and a spectral width of 4386 Hz for 100% D_2O solutions and 12 000 Hz for 10% D_2O/90% H_2O solutions at 500 MHz. At 750 MHz, the spectral width was 8000 Hz for D_2O samples and 18 000 Hz for 10% D_2O/90% H_2O samples. Each data set was processed using minimal baseline correction and no apodization. The field homogeneity was typically adjusted until the HDO resonance line width was no more than 3−4 Hz. For temperature studies, the temperature was changed and allowed to equilibrate for 7−10 min before data collection resumed. 2D-NOESY data sets used for structure determination were collected contiguously at 35 °C without removing the sample from the magnet for 50, 100, 150, and 200 ms mixing times. A total of 32 transients and 16 transients were collected per t_1 increment for CGAGGTTTAAACCTCG and its methylated derivative, respectively. Data were collected using the States method (States *et al.*, 1982) to generate hypercomplex data in the indirect dimension. A total of 2048 t_2 points and 800 t_1 increments were collected using 6 s of total recycle delay. The data were processed on an SGI Onyx using FELIX 2.30 with 90°-shifted sine bell function matched to the data size in both dimensions. The DQF-COSY experiment was collected with 32 transients in the hypercomplex mode using 2048 real points in t_2 and 800 t_1 increments.

Data were apodized using a 60°-shifted skewed sine squared function in both dimensions.

Structure Determination and Analysis. The structure determination was performed using molecular dynamics and mechanics with NMR restraints. NMR restraints were initially based on typical B DNA distances and iteratively refined by minimizing differences between experimental and back-calculated cross-peak volumes. A total of 510 total distance constraints were used for restrained molecular dynamics in CGAGGTTTAAACCTCG, and 446 total distance restraints were employed in the case of CGAGGTTT-(mA)AACCTCG. In each sequence, 40 loose hydrogen bond restraints (1.7−2.2 Å) were defined between base pairs. The remaining 400−460 restraints consist of about 15 restraints/residue defined on each strand of the DNA duplex consisting of the following types of restraints: (*n*)-H8/H6 to (*n*)-H1′, (*n*)-H8/H6 to (*n* − 1)-H1′, (*n*)-H8/H6 to (*n*)-H2′, (*n*)-H8/H6 to (*n*)-H2″, (*n*)-H8/H6 to (*n* − 1)-H2′, (*n*)-H8/H6 to (*n* − 1)-H2″, (*n*)-H2 to (*n*)-H1′, (*n*)-H2 to (*n* + 1)-H1′, (*n*)-H2 to (*m* + 1)H1′, (*n*)-H8 to (*n*)-H3′, (*n*)-H1′ to (*n*)-H4′, (*n*)-H1′ to (*n* + 1)-H5′ (where assignments were possible), and H8/H6/H2 to H8/H6/H2. The Biosym Discover 2.9/3.1 program was used with Amber potential functions for molecular dynamics/mechanics calculations. The BKMAT_1 program[1] was used for back-calculation. The BKMAT_1 program is based on a relaxation matrix calculation which (1) performs calculations assuming either isotropic or anisotropic diffusion, (2) accounts for partial recovery of spin magnetization, and (3) allows input of differential external relaxation or leakage rates. The details of the structure were analyzed using the Dials and Windows program (Ravishankar *et al.*, 1991) and plotted using Kaleidagraph (Abelbeck Software, 1993).

RESULTS

Effect of Methylation on Nonexchangeable Proton Resonances of the TpA Adenine. In DNA sequences that contain (T)TTAA(A) segments, the TpA adenine base has been shown to undergo an exceptionally large amplitude (20°−50°) and slow motion (10 ms−1 μs) which has not been found at ApT steps (Kennedy *et al.*, 1993). The base dynamics can be detected in the one-dimensional solution-state NMR spectrum as an excess line width of the nonexchangeable proton resonances. For example, in 5′-CGAGGTTTAAACCTCG-3′ at 750 MHz (Figure 1A), both the A9-H2 (15 Hz) and A9-H8 (13 Hz) resonances are broadened by motion in comparison to the resonances of the A3 residue (A3-H2 = 3 Hz). The base motion of the TpA adenine causes resonance broadening due to chemical exchange in the form of conformational dynamics, i.e., ring current fluctuations on an intermediate time scale with respect to the chemical shift differences between conformers (Kennedy *et al.*, 1993). Lefevre *et al.* (1985) originally found broad adenine-H2 resonances at the TpA step in a TTAA segment of the *trp* promoter sequence. In contrast, when the TpA adenine is methylated (Figure 1B), the A9-H2 (4Hz) resonance sharpens, indicating a significant change in its base dynamics. The H8 proton resonance of the TpA adenine, which also experiences broadening due to motion because of the rigid nature of the purine base, narrows upon methylation (Figure 1B).

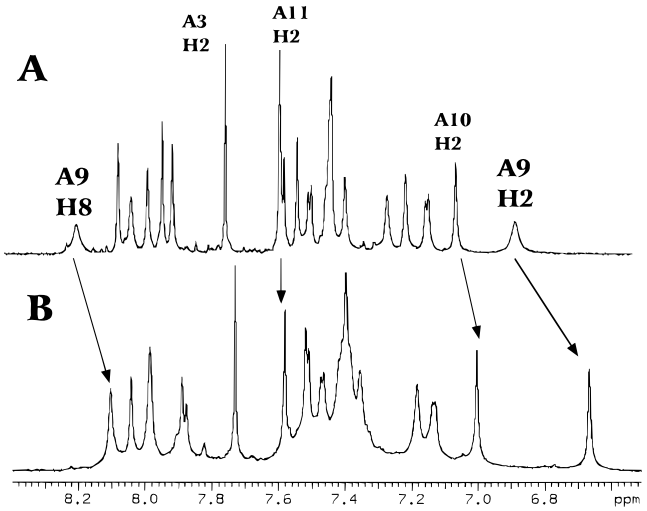

FIGURE 1: 750 MHz ^{1}H NMR spectra of (A) [d(5′-CGAGGTT-TAAACCACG-3′)]$_2$ and (B) [d(5′-CGAGGTTT(mA)AACCACG-3′)]$_2$ shown in the aromatic proton region. The (mA) refers to adenine methylated at the N^6 position. The data were collected at 35 °C in D_2O.

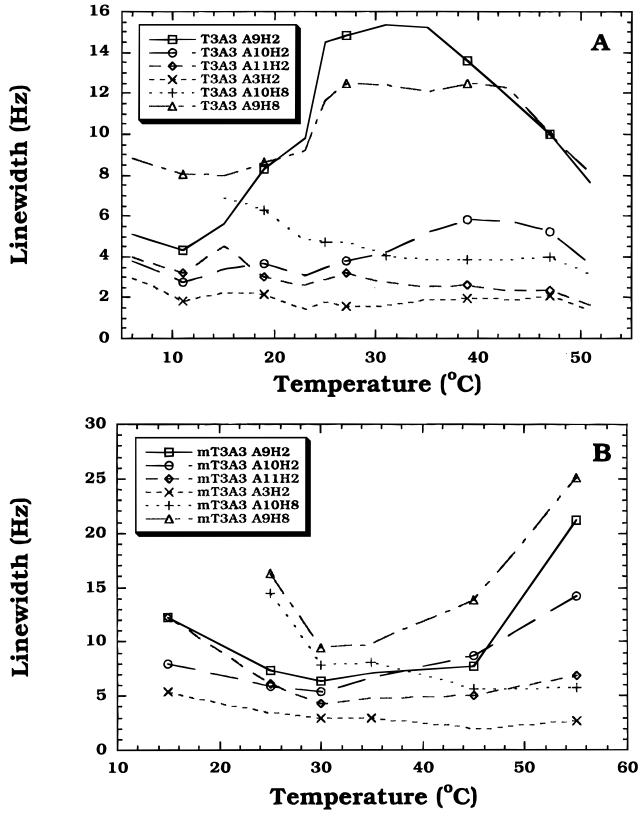

FIGURE 2: Selected H2 and H8 ^{1}H resonance line widths in (A) [d(5′-CGAGGTTTAAACCACG-3′)]$_2$ and (B) [d(5′-CGAGGTTT-(mA)AACCACG-3′)]$_2$ plotted as a function of temperature over a range from 5 to 55 °C. The (mA) refers to adenine methylated at the N^6 position. The data were collected at 750 MHz.

The conformational dynamics at the TpA step lead to proton line widths that are temperature dependent (LeFevre *et al.*, 1985). The A-H2 and A-H8 resonance line widths measured at 750 MHz in CGAGGTTTAAACCTCG and CGAGGTTT(mA)AATTCTG depend on temperature as shown in Figure 2. The A9-H2, A9-H8, and A10-H8 resonances in the unmethylated sequence reach a maximum

[1] The BKMAT_1 program was written by Leming Zhu at the University of Washington. It is a preceding version of the BIRDER program (Zhu & Reid, 1995).

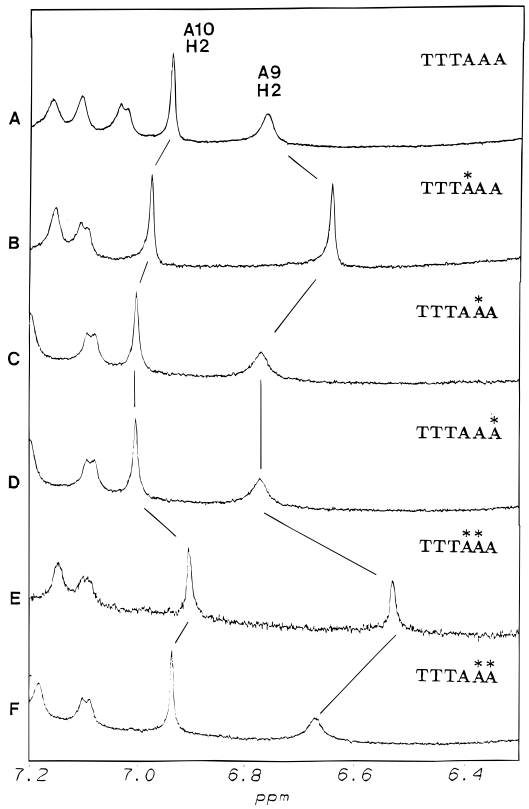

FIGURE 3: Region of the 500 MHz aromatic ^{1}H spectrum containing the A9-H2 and A10-H2 resonances measured at 31 °C in D_2O for (A) [d(5′-CGAGGTTTAAACCTCG-3′)]$_2$, (B) [d(5′-CGAGGTTT-(mA)AACCACG-3′)]$_2$, (C) [d(5′-CGAGGTTTA(mA)ACCTCG-3′)]$_2$, (D) [d(5′-CGAGGTTTAA(mA)CCTCG-3′)]$_2$, (E) [d(5′-CGAGGTTT(mA)(mA)ACCTCG-3′)]$_2$, and (F) [d(5′-CGA-GGTTTA(mA)(mA)CCTCG-3′)]$_2$. The (mA) refers to adenine methylated at the N^6 position.

line width at ≈35 °C (Figure 2A). In contrast, when methylated, the A9-H2 and A9-H8 resonance line widths decrease monotonically with increasing temperature like all the other A-H2 and A-H8 resonances until the duplex begins to melt (the T_m was found to be ∼65 °C for both sequences). On the basis of data collected at 500 and 750 MHz, the magnitude of the line width is field dependent with the A9-H2 reaching a value of ≈15 Hz at 750 MHz compared to only ≈10 Hz at 500 MHz. Since the broadening has been attributed to a modulation of ring current due to base dynamics (Kennedy *et al.*, 1993) and the span of resonance frequencies for the distribution of conformers increases linearly with increasing field strength, it is not surprising that the increase in excess line width due to base dynamics follows the ratio of the magnetic field strengths. The maximum line width occurs at the same temperature, ≈35 °C, at both fields.

In order to determine whether the line widths of the TpA adenine resonances were affected by methylation of other residues in the same sequence, the line widths of the TpA adenine resonances were monitored as the nearest neighbor bases were methylated in various positions. Figure 3 shows the A9-H2 and A10-H2 resonances for several methylated derivatives of 5′-(CGAGGTTTAAACCTCG)$_2$-3′. A summary of line widths is collected in Table 1. In Figure 3, it can be seen that the A9-H2 resonance sharpens only when the A9 adenine is methylated (Figure 3B,F). The same pattern of line narrowing upon adenine methylation was observed for the TpA adenine-H8 resonance. Methylation

Table 1: Summary of Line Widths for Resonances Affected by Adenine N^6-Methylation or Cytosine N^4-Methylation for Sequence 1 (C1-G2-A3-G4-G5-T6-T7-T8-A9-A10-A11-C12-C13-T14-C15-G16) and for Sequence 2 (C1-G2-T3-G4-T5-T6-A7-A8-C9-G10-A11-G12)[a]

	Sequence 1					
	U	M9	M10	M11	M9/M10	M10/M11
A9-H2	15	4	13	16	4	15
A10-H2	6	4	4	5	4	4
A9-H8	14	7	12	15	4	13
T8-N3H	27	20	20	29	20	21
	Sequence 2					
	U	M7	M8	M9	M7/M8	M8/M9
T6-N3H	25	17	30	25	17	21
T7-N3H	17	17	25	17	17	17

[a] The columns represent sequences that are either unmethylated (U) or methylated at the residue(s) indicated by M followed by the residue number.

significantly affects the excess line width for the aromatic proton resonances of the A9 base *only* when the A9 is N^6-methylated.[2]

Effect of Methylation on TpA Thymine Imino Resonances. In our previous study of CGAGGTTTAAACCTCG (Kennedy *et al.*, 1993), the line width of the T24(T8) thymine imino proton resonance was about 1.5 times broader than the other thymine imino resonances but the resonance persisted even at high temperatures (46 °C). Here, we observe that the T8 imino line width sharpens from 20 Hz (Figure 4A) to 14 Hz upon methylation (Figure 4B) in a fashion similar to that observed for nonexchangeable A-H2 and A-H8 resonances. Furthermore, the line width of the T8 imino resonance is temperature dependent and reaches a maximum at about 35 °C (Figure 4C), which is the same temperature at which the nonexchangeable A9 protons reach a maximum line width. When the restriction site is methylated in positions other that the TpA site, the pattern of methylation-induced line narrowing is slightly different when compared to that observed for the A9-H2 and A9-H8 resonances; i.e., the N3H imino resonance of T24(T8) sharpened not only when A9 adenine is methylated but also when A10 is methylated (see Table 1). The effect of line narrowing of exchangeable imino resonances upon methylation can be more easily seen in the sequence CGTGTTAACACG, where the TpA thymine N3H imino proton sharpens significantly upon methylation only when the TpA adenine is methylated (Figure 5 and Table 1). These data indicate that the line width of the TpA thymine imino resonance includes broadening due to the same base dynamics that lead to excess line width of the nonexchangeable resonances of the TpA steps in T_nA_n sequences.

Proton Resonance Assignments. Two-dimensional NOESY data collected at 750 MHz ^{1}H frequency allowed assignments of all protons (except for some H5′ and H5″) in both sequences. Sequential assignments in the aromatic to H1′ and assignments in the aromatic to H2′/H2″ for CGAGGTTTAAACCTCG and CGAGGTTT(mA)AAC-CTCG are shown in Figures 6 and 7, respectively. The chemical shift assignments for d(CGAGGTTT(mA)AAC-CTCG) are summarized in Table 2. Figure 8 shows how

[2] *Aha*III restriction digests of the methylated sequences showed that methylation prevented cleavage in all methylated sequences.

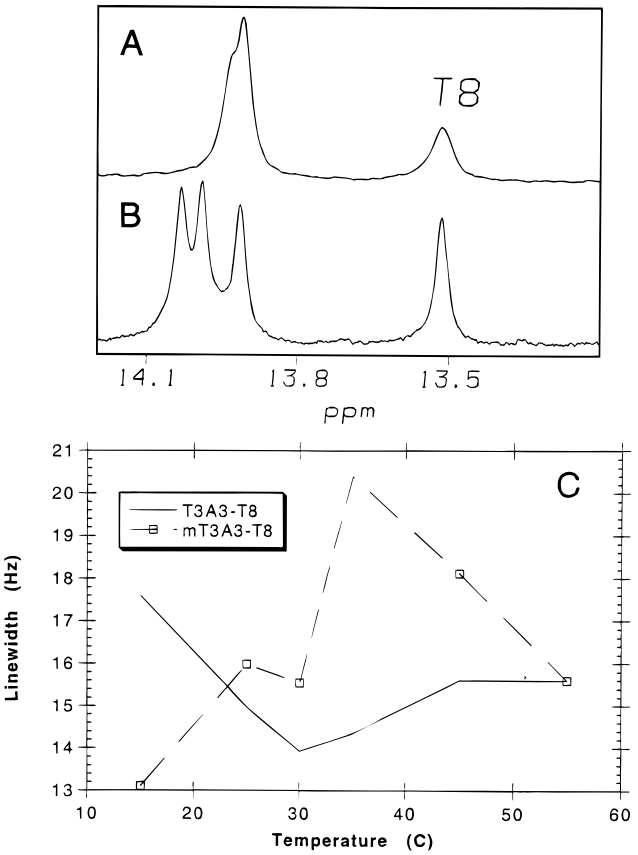

FIGURE 4: Thymine imino proton spectrum collected at 500 MHz in 90% H_2O/10% D_2O for (A) [d(5′-CGAGGTTTAAATTCTG-3′)]$_2$ and (B) [d(5′-CGAGGTTT(mA)AATTCTG-3′)]$_2$. In (C) the line width of the T8 resonance in each sequence is plotted as a function of temperature.

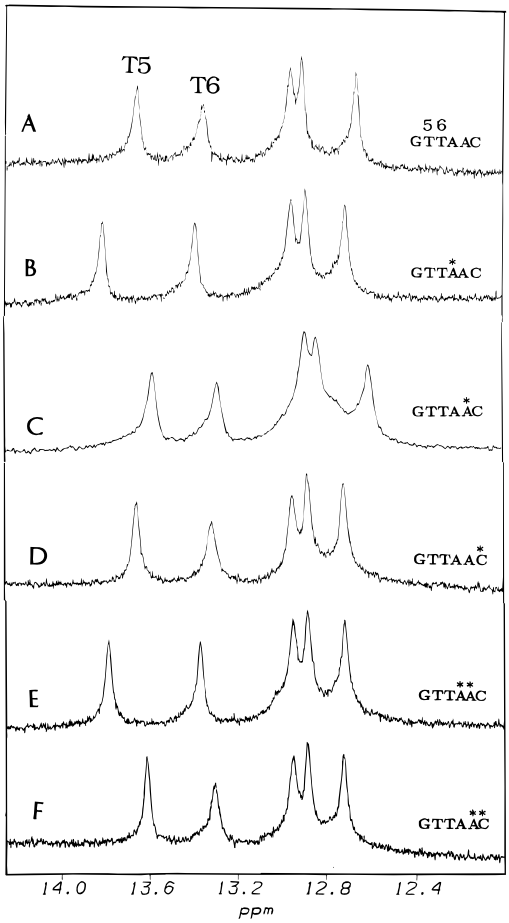

FIGURE 5: 500 MHz imino ^{1}H spectrum measured at 31 °C in 10% D_2O/90%H_2O for (A) [d(5′-GCCGTTAACGGC-3′)]$_2$, (B) [d(5′-GCCGTT(mA)ACGGC-3′)]$_2$, (C) [d(5′-GCCGTTA(mA)CGGC-3′)]$_2$, (D) [d(5′-GCCGTTAACGGC-3′)]$_2$, (E) [d(5′-GCCGTT(mA)-(mA)CGGC-3′)]$_2$, and (F) [d(5′-GCCGTTA(mA)(mC)GGC-3′)]$_2$. The T5 and T6 thymine imino resonances are labeled in (A). The (mA) indicates N^6-methylation of adenine. The (mC) indicates that cytosine has been methylated at the N^4 position.

the proton chemical shift differs between the methylated and unmethylated sequences for all assigned protons. As might be expected, the largest chemical shift differences are concentrated around the TpA site. Chemical shift changes could arise from a change in the intrinsic ring current of the methylated adenine, from structural changes surrounding the modified base, or from a combination of both mechanisms.

Distinct NMR Observations at the TpA Step. The first indication of unusual base stacking at the TpA step can be inferred from the chemical shifts of the protons of the TpA adenine. In Figure 1A, the TpA A9-H2 proton is shielded by ~1 ppm compared to the A3-H2. At the same time, the A9-H8 is distinctly the most deshielded of all adenine-H8 resonances. The chemical shift of the A10-H2 and A11-H2 resonances becomes progressively more shielded, approaching the value of the A3-H2 as the adenine residue becomes more removed from the TpA step. The fact that the H8 and H2 protons, which reside on opposite sides of the purine ring, experience chemical shifts in opposite directions indicates that the average conformation of the purine ring generates a ring current that spans the "magic" angle of the $P_2 \cos\theta$ dependence of the ring current contribution to the chemical shift. This same effect can also be seen as distinct shielding of the thymine imino resonance at the TpA step as shown in Figure 4 where the TpA N3H resonance is distinctly shielded relative to all the other thymine imino resonances. Distinct shielding of the TpA thymine imino proton is consistent with the shielding of the adenine H2 proton since the two protons are proximal to one another in the TpA base pair. Upon methylation, the chemical shifts of the H2, H8, and N3H protons in the T·A base pair do not change significantly even though the line widths all sharpen. This indicates that the average conformation of the adenine bicyclic ring relative to the surrounding nucleotides is not significantly changed even though the dynamics *are* significantly affected.

In the aromatic to H1′ region of the NOESY spectrum which is used to help constrain the glycosidic torsion angle, rise, and sugar conformations, evidence for unusual stacking at the TpA step can be seen from the spectra of both the native and methylated sequences. Figure 9 shows a part of this region in the NOESY spectrum that includes the A9 cross-peaks and the cross-peaks of several surrounding residues. In the unmethylated sequence, the (*n*)-H8/H6 to (*n*)-H1′ cross-peaks are shown for residues 3, 9, 10, 11, and 16. This distance should be relatively invariant from residue to residue since it only varies from ~3.6 to ~3.9 Å over the range of glycosidic torsion angles ($\chi = -120°$ to $-180°$) typically found in B DNA. In Figure 9A, it can be seen that the A9 cross-peak (designated 9 in the Figure 9) is slightly weaker than the other intraresidue cross-peaks. In B-type DNA, the distance from the (*n*)-aromatic proton to the sugar H1′ proton of the previous residue is typically much shorter than that to its own sugar H1′ proton. Therefore, the cross-peak to the previous H1′ is usually much stronger

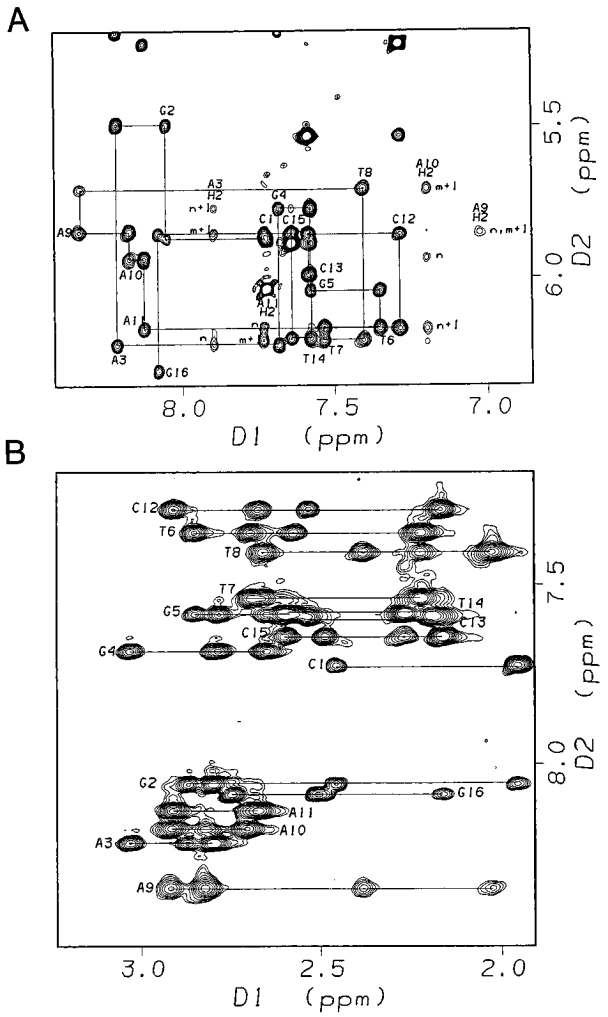

FIGURE 6: 750 MHz ^{1}H 2D-NOESY spectrum collected at 35 °C for [d(5′-CGAGGTTTAAACCACG-3′)]$_2$ in the (A) aromatic to H1′ and (B) aromatic to H2′/H2″ regions. In (A), the sequential NOE cross-peaks are connected by a solid line. The intraresidue cross-peaks are indicated by a number following the residue name. Interresidue cross-peaks are unlabeled. In (B), the four cross-peaks expected for each aromatic proton frequency are connected by a solid line.

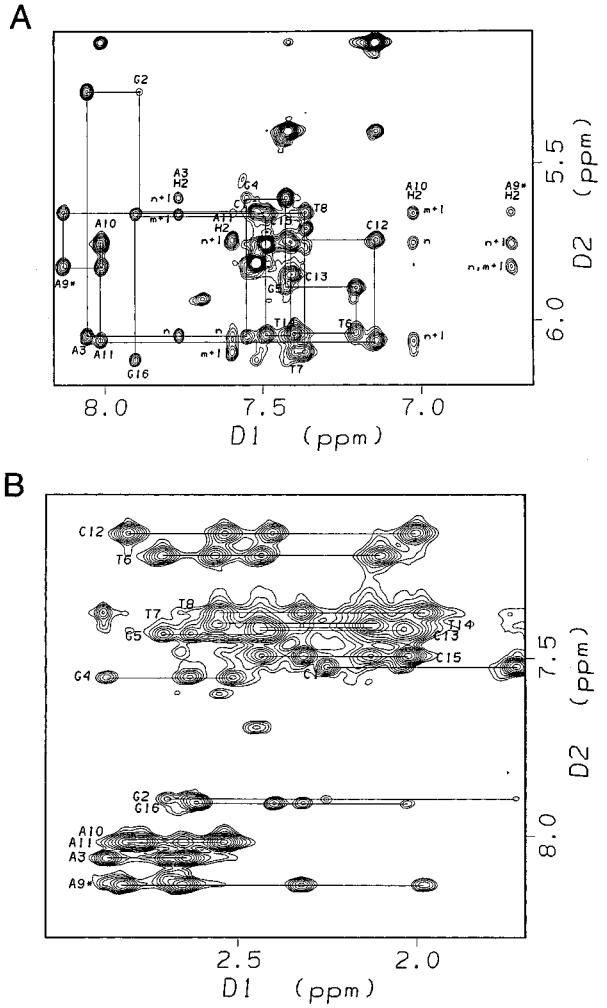

FIGURE 7: 750 MHz ^{1}H 2D-NOESY spectrum collected at 35 °C for [d(5′-CGAGGTTT(mA)AACCACG-3′)]$_2$ in the (A) aromatic to H1′ and (B) aromatic to H2′/H2″ regions. Sequential NOE cross-peaks are connected by a solid line. The intraresidue cross-peaks are indicated by a number following the residue name. Interresidue cross-peaks are unlabeled. In (B), the four cross-peaks expected for each aromatic proton frequency are connected by a solid line. The mA or asterisk indicates N^6-methylation of adenine.

than to its own H1′. This pattern exists for all the residues in CGAGGTTTAAACTCG with the exception of the TpA adenine. The anomaly at the TpA can be seen in Figure 9A. The normal pattern can be seen by inspecting the 10 and 9-10, 11 and 10-11, and 3 and 2-3 pairs of cross-peaks. However, the A9-H8 to T8-H1′ cross-peak is considerably weaker than its corresponding intraresidue cross-peak and is excessively weaker than the other (n) to (n − 1) cross-peaks shown in the Figure 9. Upon methylation, the same general pattern of cross-peak intensities can be seen as shown in Figure 9B. However, even though the A9-H8 to T8-H1′ cross-peak becomes significantly stronger, it remains weaker than the A9 intraresidue cross-peak. Concomitantly, the intraresidue cross-peak volume for A9 becomes more in line with the intensity observed for the other residues.

In the aromatic to H2′, H2″ region of the NOESY spectrum which can also be used to help constrain the glycosidic torsion angle, rise, and sugar conformations, another anomaly in cross-peak intensities expected for B-type DNA can be observed at the TpA step. Figure 10 shows traces taken at the frequencies of three aromatic protons indicated to the right of each trace for (A) the unmethylated sequence and (B) the methylated sequence. Traces were selected for residues in which all four expected cross-peaks were baseline resolved. The expected pattern of cross-peak

Table 2: Proton Chemical Shift Assignments for [(dCGAGGTTTmAAACCTCG)]$_2$ at 35 °C in D_2O Determined at 750 MHz ^{1}H Frequency

base	H8/H6	H1′	H2′	H2″	H3′	H4′	H2	H5	methyl
C1	7.54	5.66	1.76	2.29	4.64	4.13		5.81	
G2	7.91	5.32	2.66	2.73	4.96	4.26			
A3	8.08	6.07	2.66	2.90	5.06	4.42	7.76		
G4	7.56	5.62	2.52	2.64	4.95	4.35			
G5	7.44	5.91	2.46	2.72	4.80	4.39			
T6	7.22	6.04	2.12	2.58	4.84	4.28			1.18
T7	7.43	6.10	2.16	2.57	4.87	4.16			1.57
T8	7.40	5.70	2.02	2.37	4.88	4.13			1.62
A9*	8.14	5.86	2.65	2.85	5.02	4.39	6.67		
A10	8.02	5.77	2.55	2.79	5.01	4.42	7.03		
A11	6.07	6.07	2.55	2.82	4.95	4.43	7.62		
C12	7.16	5.76	2.03	2.42	4.67	4.18		5.09	
C13	7.44	5.86	2.06	2.47	4.75	4.14		5.38	
T14	7.43	6.06	2.13	2.46	4.87	4.15			
C15	7.51	5.67	2.03	2.35	4.84	4.12		5.74	
G16	7.92	6.13	2.37	2.61	4.86	4.18			

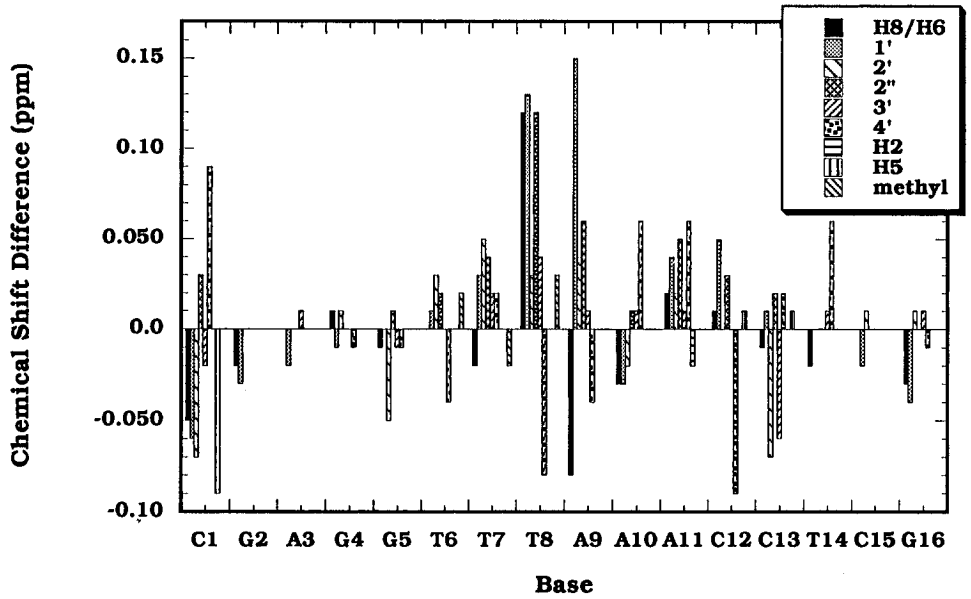

FIGURE 8: Chemical shift differences between [d(5′-CGAGGTTT(mA)AACCACG-3′)]$_2$ and [d(5′-CGAGGTTTAAACCACG-3′)]$_2$ for all but the H5′ and H5″ protons measured at 35 °C. Values were calculated by subtracting the chemical shift of the unmethylated sequence from that of the methylated sequence. The column-filling pattern for different proton types is indicated in the legend inset into the figure.

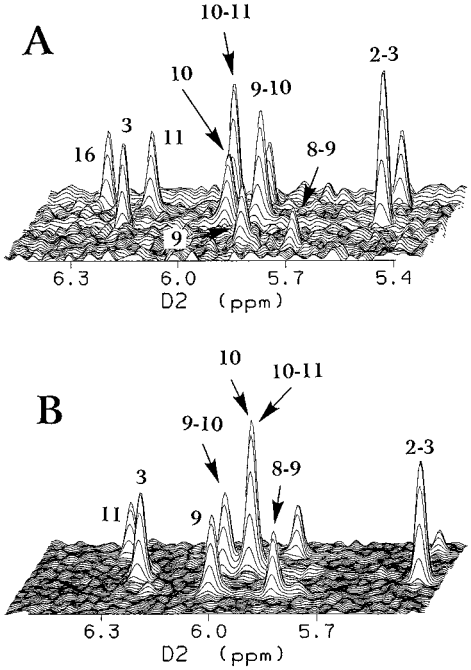

FIGURE 9: Stack plot of the aromatic to H1′ in the region containing the cross-peaks for the residues at the TpA step in (A) the CGAGGTTTAAACCTCG and (B) the CGAGGTTT(mA)AACCTCG sequences. Single numbers are used to label the intraresidue cross-peaks while interresidue cross-peaks are indicated by the number of each residue separated by a dash.

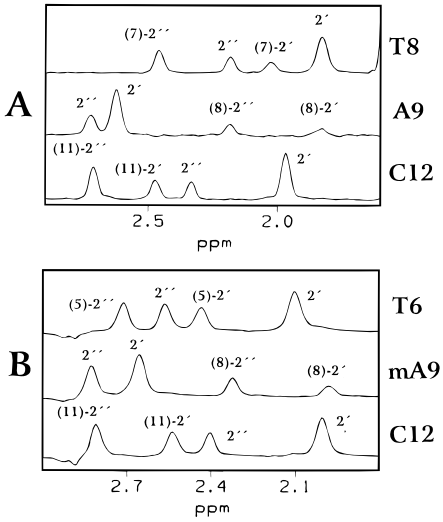

FIGURE 10: Selected traces from the 150 ms ^{1}H 2D-NOESY spectrum collected at 750 MHz. Traces were taken at the resonance frequency of the aromatic protons indicated at the right and shown in the H2′, H2″ region of the spectrum in (A) CGAGGTTTAAACCTCG and (B) CGAGGTTT(mA)AACCTCG. The (mA) indicates adenine is methylated at the N^6 position.

intensities between the aromatic proton and its own and previous sugar protons is (*n*)-H8/H6 to (*n*)-H2′ > (*n*)-H8/H6 to (*n* − 1)-H2″ > (*n*)-H8/H6 to (*n*)H2″ (*n*)-H8/H6 to (*n* −1)-H2′. This pattern can be seen in Figure 10A in the unmethylated sequence at the traces taken for T8 and the C12 residues. This pattern is clearly violated for the A9 residue in which the A9-H8 to T8-H2″ cross-peak is weaker than the A9-H8 to A9-H2″ cross-peak. In the methylated sequence, the expected pattern persists for the T6 and C12 residues but is once again violated at the A9 position. However, the intensity of the A9-H8 to T8-H2″ cross-peak relative to its own H2″ is stronger than in the unmethylated sequence.

The sugar conformations are normally constrained using distance information taken from the H1′ to H4′ region of the NOESY spectrum and from coupling constant data, sometimes taken from DQF-COSY cross-peak intensities in the H3′ to H4′ region. Figure 11 shows the NOESY spectrum in the H1′ to H4′ region for the (A) unmethylated and (B) methylated sequences. There is a significant variation in cross-peak intensities, which reflects that the distance can vary by ∼0.8 Å over the range of pseudorotation angles found in B-type DNA with stronger cross-peaks indicating *P* values in the lower range (around 120°−140°) and weaker cross-peaks indicating *P* values in the higher range (around 140°−160°). In Figure 11A, a large difference in the cross-peak intensities can be seen between the T8 and A9 residues. In the unmethylated sequence (Figure 11B)

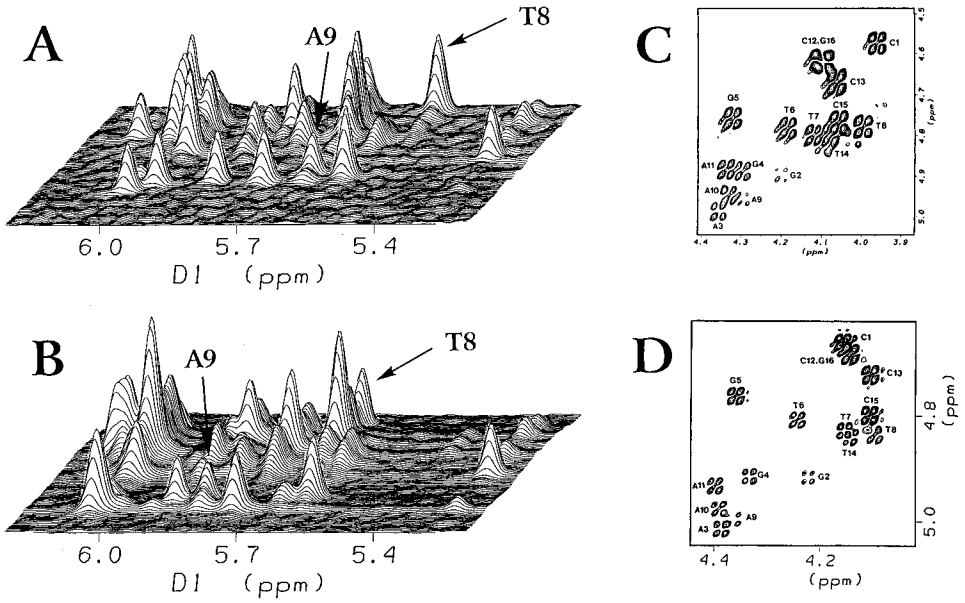

FIGURE 11: Stack plot of the H1′ to H4′ region of the 150 ms ^{1}H 2D-NOESY collected at 750 MHz for (A) CGAGGTTTAAACCTCG and (B) CGAGGTTT(mA)AACCTCG. The intraresidue cross-peaks for the T8 and A9 residues are indicated by arrows in each plot. In (C), DQF-COSY spectra are shown in the H3′ to H4′ region for CGAGGTTTAAACCTCG and in (D), for CGAGGTTT(mA)AACCTCG.

Table 3: External Relaxation Rates Measured for CGAGGTTTAAACCTCG (abbreviated TTTAAA) and CGAGGTTT(mA)AAACCTCG [abbreviated TTT(mA)AA][a]

	TTTAAA	TTT(mA)AA
H8/H6	0.0	0.0
H1′	0.0	0.0
H2′/H2″	0.7	0.4
H3′	0.0	0.0
H4′	0.4	0.4
H5′/H5″	0.0	0.4
H2	0.0	0.0
H5	0.5	0.4

[a] The value shown in each column was applied uniformly for all residues to the type of proton indicated to the left.

Table 4: Final Energy Parameters[a]

	T_3A_3 (B)	T_3A_3 (A)	mT_3A_3 (B)	mT_3A_3 (A)
total energy	−110.74	−131.95	−88.13	−83.16
bond energy	29.60	26.50	30.31	29.81
angle energy	141.94	138.16	168.62	175.87
dihedral energy	365.08	358.28	359.23	358.32
H-bond energy	−25.48	−25.68	−25.28	−25.29
nonbond energy	−305.24	−312.48	−325.36	−326.34
Coulomb energy	−316.88	−316.88	−295.75	−295.60
NOE forcing potential	14.58	23.17	3.33	7.28

[a] The columns indicate the final values for structures refined from B-form (B) or A-form (A) starting structures for CGAGGTTTAAACCTCG (T_3A_3) and CGAGGTTT(mA)AACCTCG (mT_3A_3). The units of energy are in kcal/mol.

Table 5: Root Mean Square Deviation (RMSD) Values between Final Structures[a]

	T_3A_3 (B)	T_3A_3 (A)	mT_3A_3 (B)	mT_3A_3 (A)	B DNA	A DNA
T_3A_3 (B)	0.31				1.91	5.11
T_3A_3 (A)	0.75	0.41			2.12	4.78
mT_3A_3 (B)	1.63	1.53	0.26		2.24	4.57
mT_3A_3 (A)	1.77	1.50	0.92	0.36	2.28	4.52

[a] The column and row designations indicate the sequences CGAGGTTTAAACCTCG (T_3A_3) and CGAGGTTT(mA)AACCTCG (mT_3A_3). The (B) or (A) indicates that the structure was refined starting from B-type DNA or A-type DNA, respectively. The numbers indicated the average value taken from superimposing the three lowest energy structures from each column or row.

the difference between the T8 and A9 cross-peak intensities is diminished. The DQF-COSY cross-peak intensities also report on the value of *P*, where the more intense values indicate values of *P* in the lower range and weak cross-peak intensities indicate *P* values in the higher range (Kim & Reid, 1992). In Figure 11C, a striking difference in the T8 and A9 cross-peak intensities can be seen, indicating a discontinuity in sugar conformations at the TpA step. Once again, the difference between the T8 and A9 cross-peak intensitites is diminished upon A9-methylation as shown in Figure 11D.

External Relaxation Rates. External relaxation or leakage rates were determined as described by Zhu *et al*. (1995). Briefly, the leakage rates were adjusted until the calculated percent recoveries matched the experimental recoveries, based on 1D spectra collected with several different total recovery times between 0.1 and 20 s. Table 3 summarizes the external rates determined for both CGAGGTTTAAACCTCG and CGAGGTTT(mA)AAACCTCG.

Anisotropic Diffusion Constants. Isotropic tumbling correlation times or anisotropic diffusion constants were required for NOESY back-calculation. The experimental NMR correlation time was determined to be 6.2 ns in CGAGGTTT(mA)AACCTCG using the ratio of the resolved cytosine H5 autopeaks to corresponding cytosine H5-H6 cross-peak volumes and then extrapolating the fitted curves to a vanishing mixing time. All back-calculated spectra were generated with calculations accounting for anisotropic diffusion. The anisotropic diffusion constants, calculated using the hydrodynamic theory of Tirado and de la Torre (1980), were found to be $1/D_\perp = 11.4$ ns and $1/D_{||} = 3.8$ ns (including correction for end effects). Back-calculated spectra generated with anisotropic diffusion coefficients calculated ignoring end effects showed no significant differences.

Structure Refinement and Back-Calculation of NOESY Spectra. The structures of CGAGGTTTAAACCTCG and CGAGGTTT(mA)AACCTCG were determined by using

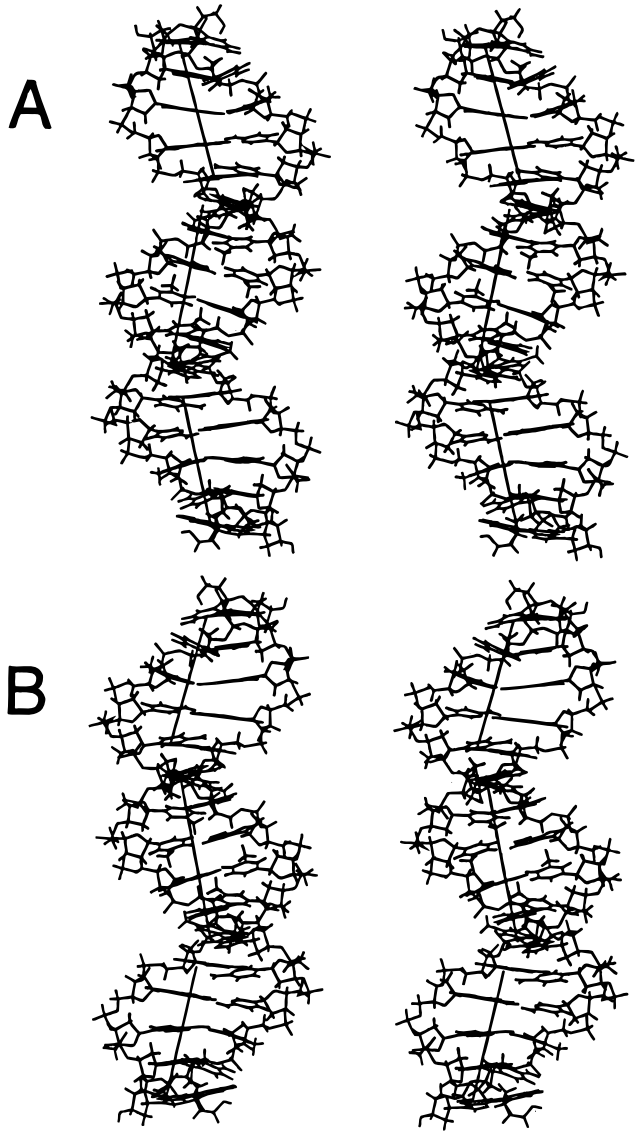

FIGURE 12: Cross-eyed stereo presentation of one structure from the final family of refined structures for the CGAGGTTTAAAC-CTCG sequence looking into (A) the major groove and (B) the minor groove at the central TpA base pair. Three distinct local helix axes are depicted by a solid line running along the center of the helix.

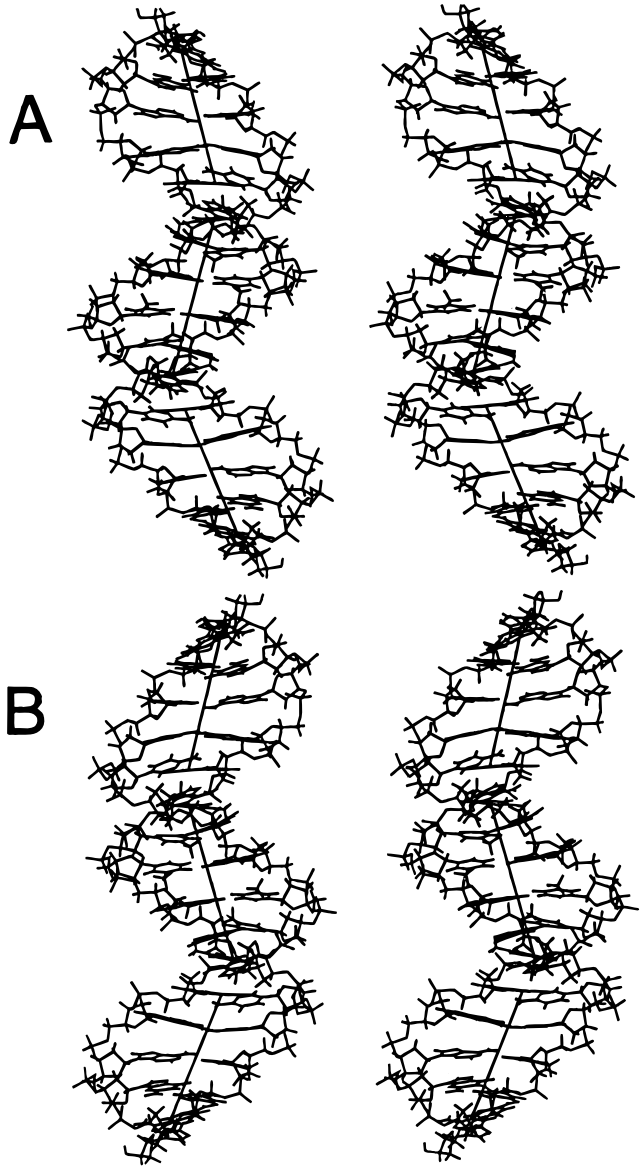

FIGURE 13: Cross-eyed stereo presentation of one structure from the final family of refined structures for the CGAGGTTT(mA)-AACCTCG sequence looking into (A) the major groove and (B) the minor groove at the central TpA base pair. Three distinct local helix axes are depicted by a solid line running along the center of the helix.

iterative process restrained molecular dynamics and distance bounds adjustments based on a comparison of back-calculated NOESY spectra with data collected at 50, 100, 150, and 200 ms as described above. A complete set of figures showing stack plots of the H8/H6 to H1′, H8/H6 to 2′/2″, H1′ to H4′, and H8/H6 to H3′/H4′ regions at all mixing times is available for both sequences as supporting information. The intraresidue aromatic to H1′ distances all fell between the 3.6 and 4.0 Å range expected for B DNA. The interresidue aromatic to H1′ distances varied with sequence depending on the glycosidic torsion angle. Intraresidue H8/H6 to H2′ distances fell in the range from 2.0 to 2.4 Å. Interresidue H8/H6 distances fell into the range between 3.5 and 4.2 Å. Intraresidue H8/H6 to H2′ distances fell into the range between 3.4 and 4.0 Å. Interresidue distances fell into the range between 2.2 and 3.0 Å. The H1′ to H4′ distances were found to vary between 2.6 and 3.3 Å, depending on the pseudorotation angle. The energy terms for structures refined from B-form and A-form starting structures are collected in Table 4. The root mean square deviations (RMSD) between final structures are summarized in Table 5. The RMSDs were below 0.75 for the unmethylated sequence and below 0.95 for the methylated sequence. The residual factors (*R* factors) for the unmethylated and methylated sequences were 0.25 and 0.11, respectively.

Analysis and Comparison of the NMR-Based Models for the CGAGGTTTAAACCTCG and CGAGGTTT(mA)AAC-CTCG Structures. The structures of CGAGGTTTAAAC-CTCG and CGAGGTTT(mA)AACCTCG are shown in Figures 12 and 13. The two structures are straight in a global sense; i.e., there is no detectable overall bend in either sequence. However, in both sequences three distinct helix axes can be defined where one axis is contained in the central TTTAAA segment and the two remaining axes are defined by the first five and last five base pairs. The angle between the terminal and central axes is about 35°. While no change in global structure is observed after methylation, some local changes in structure are clearly evident and, in some cases,

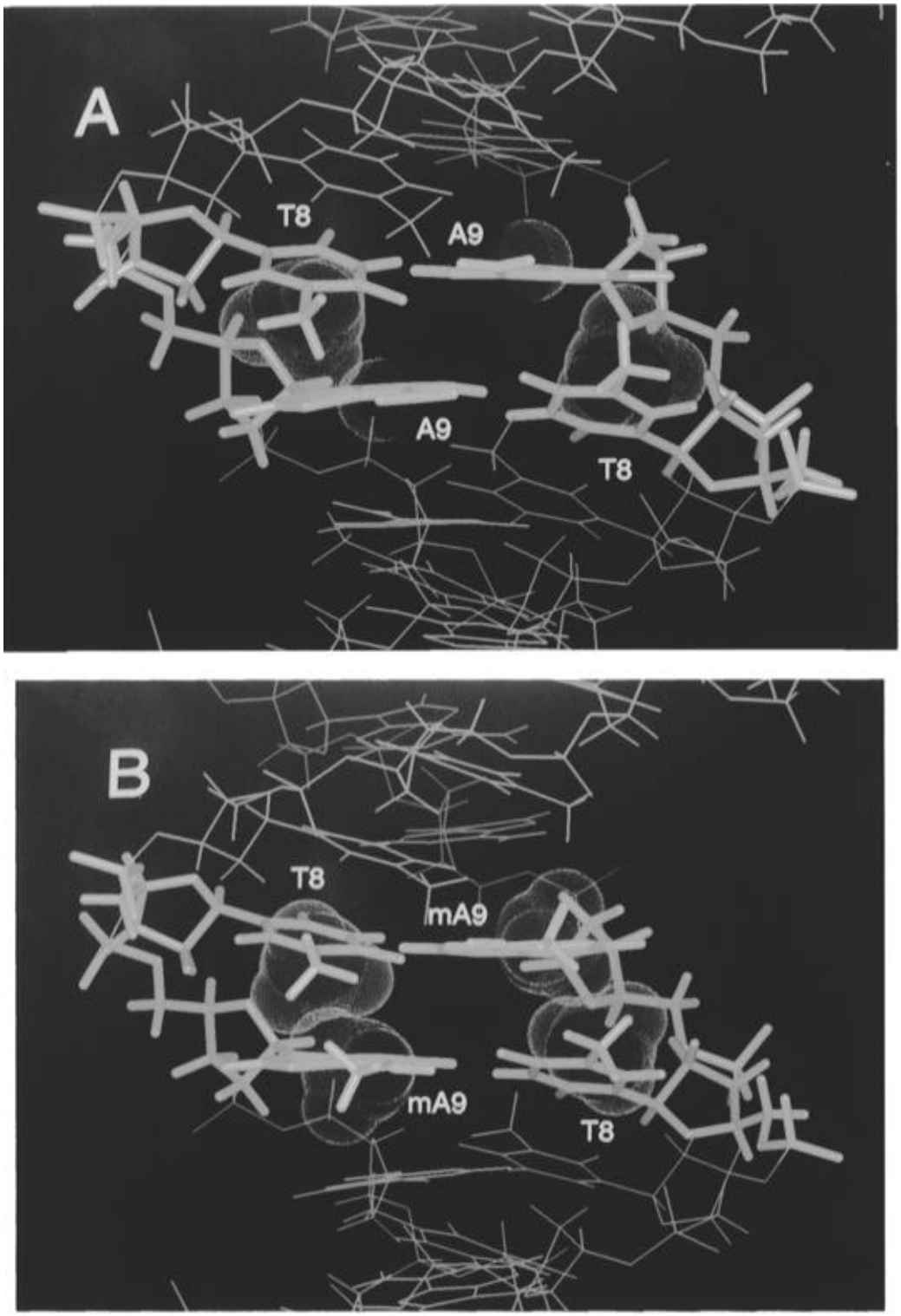

FIGURE 14: Base pairing at the T8-A9 step (A) looking perpendicular to the major helix axis in CGAGGTTTAAACCTCG and (B) looking perpendicular to the major helix axis in CGAGGTTT(mA)AACCTCG. Van der Waals surfaces are shown for the non-hydrogen-bonded adenine amino protons in (A) (violet), for thymine methyl groups in both (A) and (B) (blue), and for the methyl group at the N^6 position of adenine in (B) (violet).

propagate their effect by several base pairs (see below). We have previously suggested that poor base stacking may in part account for the dynamics observed at the TpA step. The base stacking at the TpA step is shown in Figure 14. Due to the self-complementary nature of the sequence, the T8·A9 base pair of one strand is stacked with the T8·A9 base pair of the complementary strand. Evidence of poor base stacking can be seen looking perpendicular to the major axis of the DNA at the T8·A25 base pair stacked with A9·T24 shown in Figure 14A. A large negative propeller twist is observed at the TpA base pair ($\sim -46°$), producing a large angle between the planes of the bases of T8 and A9 residues in addition to a unusually large rise. Upon methylation, the propeller twist is significantly reduced ($\sim -25°$), causing the planes of the T8 and A9 bases to be much closer to parallel (Figure 14B), and the rise is diminished. Inspection of the van der Waals surfaces of the the amino proton and methyl groups reveals that there is no steric conflict between the

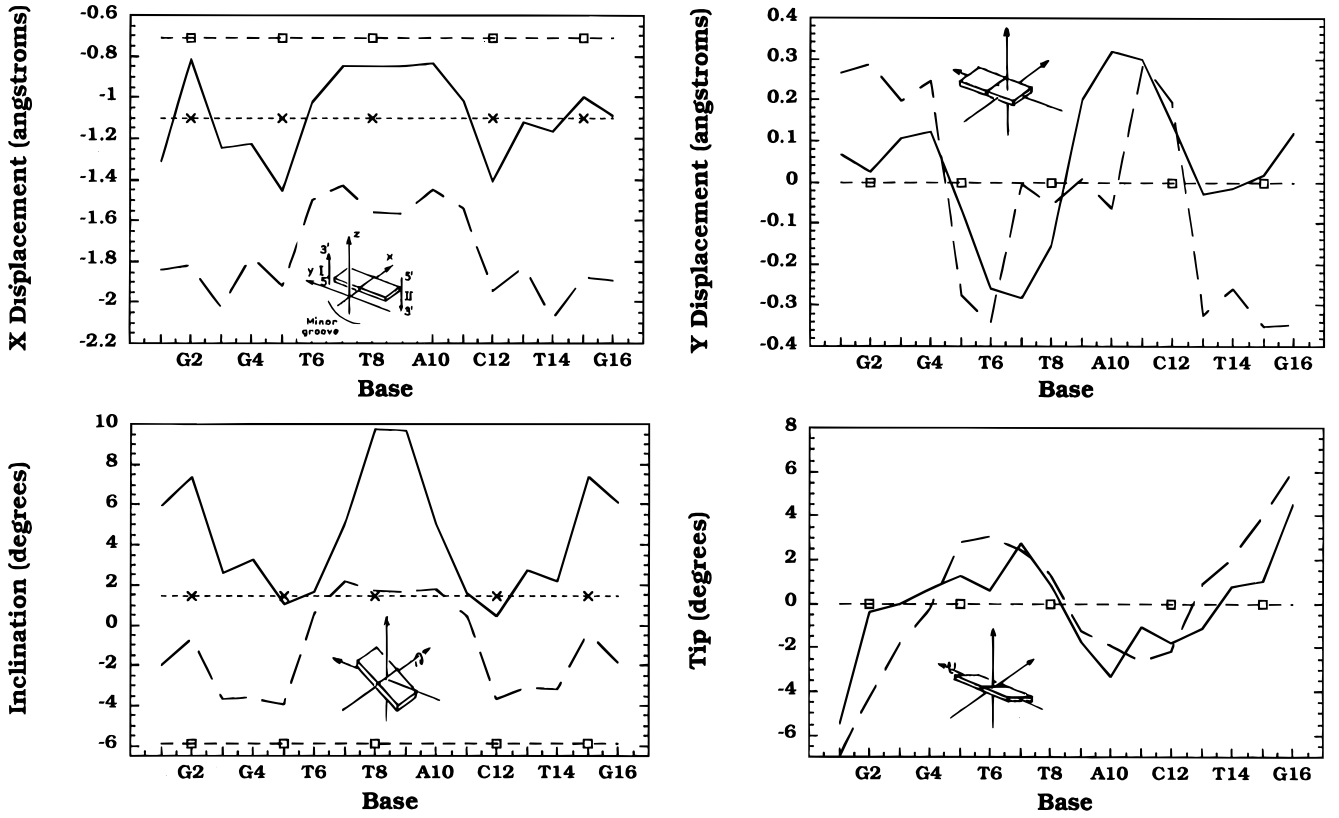

FIGURE 15: Axis-base/base pair parameters plotted as a function of residue number. The solid line indicates the value for CGAGGTT-TAAACCTCG. The dashed line indicates values for A9-methylated CGAGGTTT(mA)AACCTCG. Straight lines indicate the values found in regular B (squares) and B′ (crosses) taken from Arnott *et al.* (1973). In this figure and Figures 16−18, the schematic representation of the helix parameters was adapted from Arnott *et al.* (1973).

C5-methyl group of T8 and the exocyclic 6-amino proton of A9 (Figure 14A) in the unmethylated sequence. However, upon methylation, the surfaces of the T8-thymine methyl and A9 aminomethyl group make contact with each other. Thus, the improved base stacking upon methylation is in part a result of the minimization of the steric interaction of the four compactly organized methyl groups at the center of the sequence. Minimal rearrangement of TpA base stacking occurs upon methylation parallel to the helix axis.

Helix Parameters. Data summarizing the orientation of bases and base pairs relative to the helix axis are plotted in Figure 15. The *y* displacement and tip for the two structures superimpose quite well. While the *y* displacement changes from slightly negative (−0.3 Å) to slightly positive (−0.3 Å) in passing through the TpA step, the tip angle goes from being slightly positive (+3°) to slightly negative (−3°). The *x* displacement and inclination reach a maximum at the TpA step in both sequences, and the pattern, as a function of residue number, maps quite well; however, the absolute values are slightly displaced when comparing the two structures. The intrabase pair parameters are plotted in Figure 16. The shear and stretch map exceedingly well for the two structures. The shear changes linearly from −1 to +1 Å in the region defined by the central axis. The stretch reaches a maximum value of ∼−0.2 to −0.5 Å precisely at the site of the junctions between the three distinct helix axes. The stagger, buckle, propeller twist, and opening (Figure 16A−D) all reveal differences between unmethylated and methylated structures either surrounding or at the TpA step. For example, in the unmethylated sequence, the propeller twist reaches a maximum negative value of −46° at the TpA step, whereas, in the methylated sequence, the propeller twist experiences a local maximum of −26°. The buckle increases from −10° in the unmethylated sequence to +7° in the methylated sequence. The opening reaches a maximum negative value in both sequences at the two junctions between the three defined helix axes and experiences a maximum value of +10° in the unmethylated sequence compared to a local minimum value of ∼+6° when methylated. The intrabase/base pair parameters reveal values that are distinct when compared to B DNA and yet common to both sequences at the two junctions between distinct helix axes and values that differ when the structure surrounding the TpA step in methylated and unmethylated sequences is compared. The interbase/base pair parameters are shown in Figure 17. The shift, slide, roll, and twist all map well between the two structures and have values that are similar to those of B DNA. The slide reaches a maximum negative value (−1 Å) at the two junctions between the three axes. The rise shows the most dramatic change upon methylation, going from ∼4.1 Å (unmethylated) to ∼3 Å (methylated). The tilt changes from −4° to +3° at the A10 position. Finally, the axis junction parameters (Figure 18) all map well for the two structures. However, clear patterns of deviation from B DNA exist, e.g., the variation of negative to positive *y* displacement and the positive to negative trends in the tip angle.

Sugar Conformations. Figure 19A shows how the pseudorotation angle varies as a function of residue in each sequence. Two patterns are apparent. In the unmethylated sequence, the sugars are initially near the "south" conformation and gradually shift toward an "east" conformation. Then at the TpA step, the conformation abruptly returns to a south conformation and then, once again, moves toward the east conformation. In this sequence, it appears that the gradual

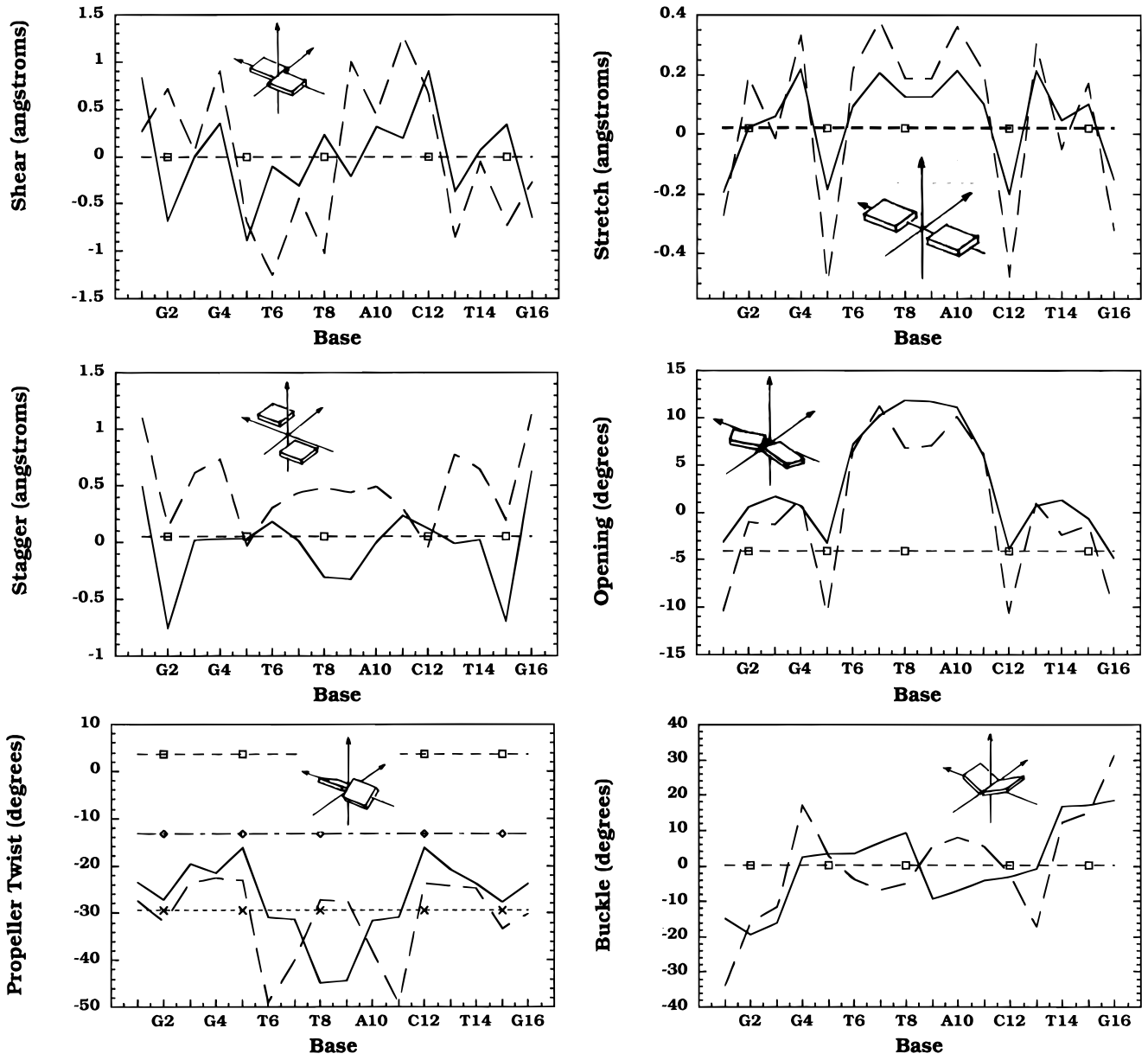

FIGURE 16: Intrabase pair parameters plotted as a function of residue number. The dashed line indicates values for A9-methylated CGAGGTTT(mA)AACCTCG. In each plot, the straight line with square points indicates the value found in a regular form of B DNA. The three straight lines shown in the plot of propeller twist refer to B DNA [squares from Arnott *et al.* (1973) and diamonds from Arnott *et al.* (1980)] and H or B′ DNA (crosses), i.e., A·T-rich DNA from Arnott *et al.* (1983).

shift from south to east occurs in polypurine regions, i.e., G2-A3-G4-G5 and A9-A10-A11. Note that even though the structures were determined using only distance contraints from NOE data, the pattern of sugar conformations indicated by the magnitude of the coupling constants inferred from the H3′-H4′ DQF-COSY cross-peak intensities shown in Figure 11C,D is consistent with distance-restrained structures.

Glycosidic Torsion Angles. These angles are found to fall in a range correlated with a 2′-endo or south sugar conformation (Figure 19B) with pyrimidine nucleotides (Saenger, 1984) in all residues. The χ angle indicates that all nucleotides assume an anti conformation; however, the value falls from $\sim-115°$ to $\sim-136°$ in the T6-T7-T8 segment, abruptly returns to $\sim-110°$ at A9, and then gradually dips below $-130°$ in the C12-T13-C14-C15 segment. It appears that the χ tends toward the edge of the anti range near the "high syn" conformation in the polypurine segments and returns to a conformation more in the center of the anti range in polypyrimidine regions. Both the T8 and A9 χ angles are nearer the high syn edge of the anti range compared to the values measured in the methylated sequence.

DISCUSSION

We and others have now observed nucleotide base dynamics which are apparently unique to TpA steps in DNA. The sequence studied here contains the T_3A_3 *Aha*III restriction site where enzymatic cleavage of the phosphodiester backbone occurs precisely at the TpA step. Our goal was to determine whether the large amplitude, slow base dynamics that occur at TpA steps in DNA could be understood by examining the structure surrounding the TpA step. By chemically modifying the TpA adenine in such a way that eliminates all evidence for base dynamics and determining the structural changes local to the TpA site, we hoped to perceive what structural features are characteristic of the TpA step and what features might give rise to the base dynamics.

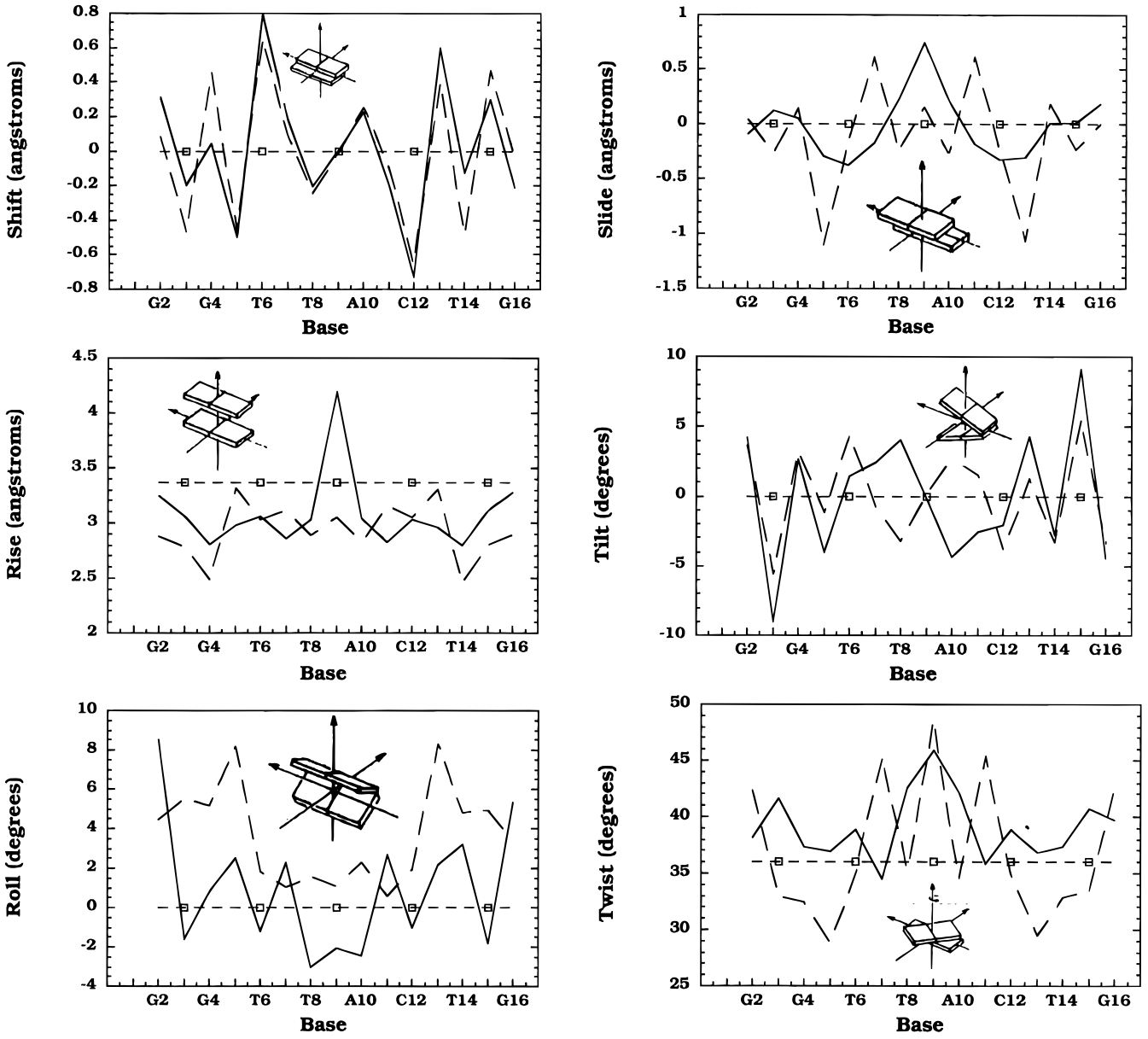

FIGURE 17: Interbase/base pair parameters plotted as a function of residue number. The dashed line indicates values for A9-methylated CGAGGTTT(mA)AACCTCG. Straight lines indicate B DNA values taken from Arnott *et al.* (1973).

The methylation of the TpA adenine was shown to restrict its base motion to the point where any residual motion is undetectable.

We report that the thymine imino proton resonances at the TpA steps in DNA also contain excess line width in the unmethylated sequences that is about one-third the value of the measured line width. The mechanism for broadening of TpA thymine imino resonances is understandable considering the local structure. The T24(T8) N3H imino proton is close to the A9-H2 of the complementary residue in a conventional Watson−Crick base pair. Because of the unusual base stacking pattern, the T24(T8) imino proton is chemically shielded by 0.3−0.8 ppm relative to other thymine imino protons, indicating that it experiences a considerable ring current field in the conformation found at TpA steps (Kennedy *et al.*, 1993; Kim & Reid, 1992). Therefore, the A9 adenine motion in the A9·T24 base pair broadens the T8 imino proton of the T8·A25 base pair. The imino resonance of the TpA thymine, which is base paired to and stacked with the mobile adenine, sharpens upon methylation in a fashion similar to that observed for nonexchangeable TpA adenine protons. The sharpening of the A9(A25)-H2, A9(A25)-H8, and T24(T8)-N3H resonances in the methylated sequence can be interpreted as resulting from a diminished contribution to the line width due to conformational dynamics caused by either diminished rate or amplitude of the dynamics. Additionally, reduced solvent exchange due to solvent inaccessibility must also considered. The most likely scenario, however, is reduced amplitude of motion. The sharpening of the resonance upon methylation allows the determination of the magnitude of the excess line width due to chemical exchange in the form of conformational dynamics. Leroy *et al.* (1988) have already shown that the base pair lifetime at a TpA step (1−6 ms) in a T_nA_n segment is about five times shorter than that at the following A·T base pairs (5−30 ms). Since the base pair lifetimes at the TpA steps occur in the 1−10 ms time scale, the measurements were based on line width measurements using the relationship $k_{ex} = 1/\tau_{ex} = \pi\Delta$, where Δ is the excess line width after the addition of catalyst. However, since the

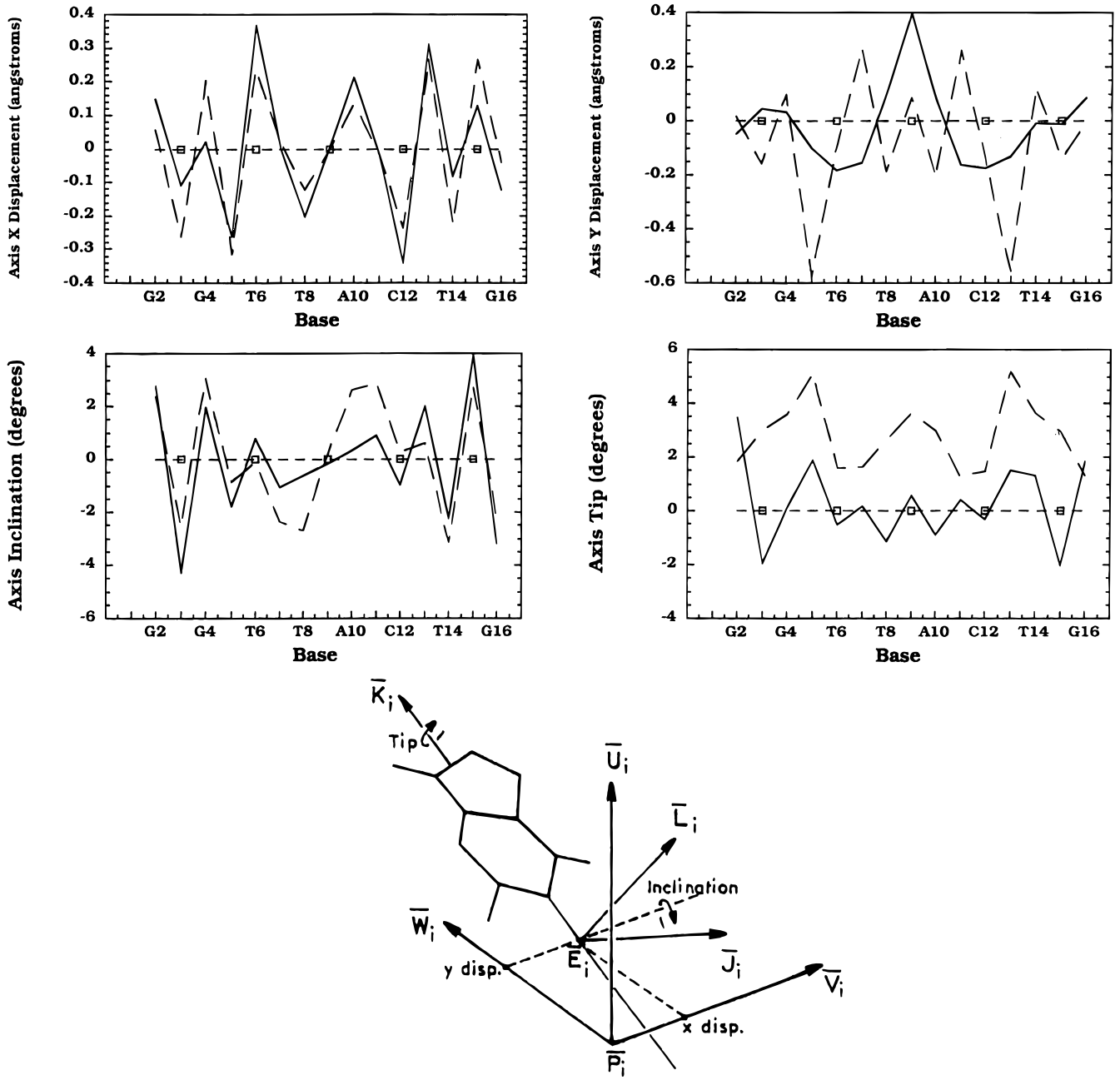

FIGURE 18: Axis junction parameters plotted as a function of residue number. The dashed line indicates values for A9-methylated CGAGGTTT(mA)AACCTCG. Straight lines indicate B DNA values taken from Arnott *et al.* (1973). The schematic diagram was also adapted from Arnott *et al.* (1973).

base pair lifetime is determined from the value of the intercept of the slope of the line produced when the exchange time τ_{ex} (which only depends on the excess line width after addition of catalyst) is plotted as a function of the concentration of added catalyst, neither the slope nor the intercept should be affected by the contribution from conformational dynamics to the imino line widths in the absence of added catalyst. Therefore, the values of base pair lifetimes determined by this method should be unnaffected by the intrinsic base dynamics at the TpA step. It is important to note, however, that the structural changes accompanying base pair opening could also be a source of modulation of the ring currents giving rise to broadening of nonexchangeable adenine H2 and H8 resonances.

In other studies of methylated DNA, N^6-methylation has been shown to slow down the single-strand−duplex exchange rate in DNA sequences containing the 5′-GATC-3′ *dam* methylation site (Fazakerley *et al.*, 1984, 1985, 1987; Quignard *et al.*, 1985). In those rather short DNA sequences, the single- and double-strand forms were found to be in slow exchange even at 29 °C. In all the studies of the *dam* methylation site, the structural consequences of methylation on the DNA duplex have been minimal. The methylated sequences studied here exist in the duplex form at 35 °C, with no sign of resonances in slow exchange. Rinkel *et al.* (1987) have studied 5′-CCGA(mA)TTCGG-3′, where A5 was N^6-methylated. On the basis of the sharpening of imino resonances in the methylated sequence, they concluded that N^6-methylation slowed down duplex−single-strand exchange. At the same time, they claimed that the thermodynamic stability of the methylated duplex was reduced compared to that of the unmethylated sequence since the T_m was reduced from 332 to 325 K. The slower duplex−single-strand exchange (which might imply increased duplex

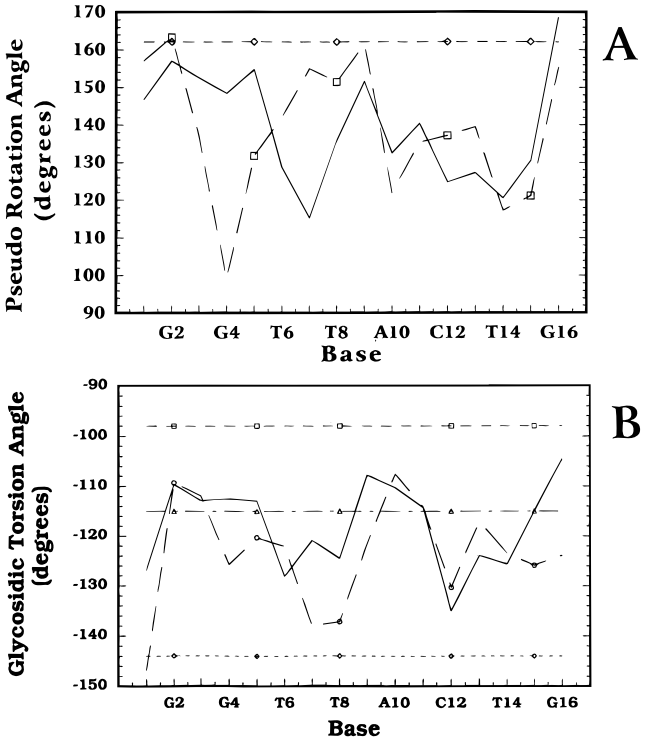

FIGURE 19: (A) Pseudorotation angle plotted as a function of residue number. (B) Glycosidic torsion angle plotted as a function of residue number. The solid and dashed lines indicate values for A9-methylated CGAGGTTT(mA)AACCTCG. In (B), the straight lines with triangle and diamond points indicate the upper and lower limit for the glycosidic torsion angle when the sugar conformation is found in the 2′-endo or south conformation. The straight line with square points indicates the value found in canonical B DNA taken from Arnott *et al*. (1973).

stability) and the reduced duplex stability (based on a reduced duplex melting temperature) present an apparent contradiction. In the 5′-CCGA(mA)TTCGG-3′ sequence studied by Rinkel *et al.* (1987), no anomalous motion was detected at the ApT step in contrast to the TpA step in T_nA_n sequences. Therefore, the interpretation of sharpened imino resonances in the methylated derivative can be directly attributed to a decreased rate of imino exchange with the solvent. However, the direct correlation of a decreased imino exchange rate with a slower base pair opening rate is not so straightforward. The sharpened imino resonances might reflect a reduced access of the exchangeable imino protons to solvent rather than a decreased rate of base pair opening. Conceivably, there are two pathways for solvent exchange with the thymine imino proton at the ApT step that may be altered by adenine N^6-methylation. First, due to the hydrophobic nature of the relatively bulky methyl group at the A5 and A15 residues, water may be simply excluded from access to the T16 and T6 imino protons. Second, the normally solvent-accessible (non-hydrogen-bonded) and exchangeable N^6-amino proton has been replaced by a methyl group. Consequently, even if increased base pair opening does occur, the remaining exchangeable amino proton may not exchange as readily with solvent. The bulky nature of the methyl group prevents the 2-fold rotation of the amino group about the C6−N6 bond (Engel & von Hippel, 1974), and the remaining exchangeable proton experiences reduced exposure to the solvent. Therefore, if proton exchange of the exocyclic 6-amino proton with solvent in any way facilitates imino exchange with solvent in a normal A·T base pair, then adenine N^6-methylation may well retard solvent exchange with the imino proton. Hence, in the 5′-CGAGGTT-TAAACCTCG-3′ sequence studied here, the sharpening of the T8(T24) imino resonance when the A9 adenine is methylated probably reflects (1) reduced motional broadening and (2) reduced solvent access to the imino protons. The sharpening of the T8(T24) imino resonance upon methylation of the $(m + 1)$-A10(A26) residue provides additional evidence that reduced solvent access may affect imino exchange in methylated sequences since the A9-H2 remains dynamically broadened.

The comparison of the structures of the unmethylated and methylated sequences reveals that they are are globally quite similar but that significant differences exist in the local structure and dynamics surrounding the TpA step. The overall structures of both sequences are straight; however, three distinct local helix axes can be defined in a common manner in both sequences. Poor base stacking is the most evident characteristic structural feature of the TpA step and is defined by a large propeller twist, unusually large rise, negative buckle, large positive opening, and negative tilt. This is not surprising since it has been shown that the TpA base stacking is the highest energy pair of all possible combinations (Rein *et al.*, 1978). N^6-Methylation of the TpA adenine improves the base stacking reflected by the return of several structural parameters surrounding the TpA step to more B DNA-like values. Also, the striking discontinuity in sugar conformations observed at the TpA step is greatly diminished upon methylation. The motionally broadened A9-H2 and A9-H8 resonances were shown to reach a maximum line width at ∼35 °C. At temperatures above and below the line-width maximum, these resonances sharpen considerably, which implies that either the rate or amplitude of the motion is temperature dependent. The DNA helix is known to unwind with increasing temperature (Jardetsky & Roberts, 1981). Our current feeling is that at low temperatures the structure surrounding the TpA step is such that the base dynamics are restricted. As the temperature is raised and the helix unwinds, the local conformation changes in such a way that the conformational dynamics are enabled and occur with an amplitude and time scale that give rise to the observed effects on the TpA resonances. As the temperature is raised even further, either the rate of the motion becomes fast on the chemical shift time scale or the amplitude of the motion again becomes restricted due to the assumption of a new and restrictive local conformation. Comparison of NOESY data collected at 15 and 45 °C did not reveal significant changes in the spectrum, indicating that the NMR experiments may be insensitive to the type of structural change that takes place or that the average structure does not change significantly and that the rate of TpA base dynamics shifts the time scale of the motion away from intermediate exchange. Comparison of these structures with previously published structures which also contain T_nA_n segments reveals mixed observations. In both structures described here, all sugars assume a south conformation in contrast to earlier reports of the structure of the *trp* operator where many sugars were found in a north conformation. Also, here we observe several parameters that reach extreme values at the TpA step and which might be interpreted as defining a junction in contrast to the description of the GT_4A_4C sequence described by Gupta *et al.* (1988) where no evidence for a junction was observed. Some structural

features such as increased rise, the pattern of buckle, and pseudorotation angles were found to be consistent with earlier reports of Kim and Reid (1992) and Kim *et al*. (1992).

In summary, we have shown that the evidence for base dynamics present at TpA steps in DNA can be eliminated by methylation of the TpA adenine as indicated by the loss of excess line width for three different types of protons. Elsewhere, we have shown that the conformational dynamics occur at the TpA step in all 16 possible immediate sequence contexts. The intrinsic conformational dynamics localized at TpA steps in DNA serve as one example of sequence-dependent structure and dynamics in DNA. Such sequence-dependent DNA structure and dynamics may play an important role in DNA−protein interactions. For example, in the case of transcription factors and restriction enzymes, sequence-dependent variations in structure and base dynamics may signal sequence information to the protein required for binding. Subtle variations in secondary DNA structure and dynamics may also impact the chemistry that takes place on the surface of the DNA molecule, e.g., the deposition of drugs and carcinogenic chemicals in DNA. Finally, these same structural variations may influence the profile of DNA damage and the subsequent efficiency of DNA repair nearby TpA steps in DNA.

ACKNOWLEDGMENT

We thank Kate McAteer for critical reading of the manuscript.

SUPPORTING INFORMATION AVAILABLE

Stack plots of experimental and back-calculated data plotted at 50, 100, 150, and 200 ms mixing times (14 pages). Ordering information is given on any current masthead page.

REFERENCES

Arnott, S., & Hukins, D. W. L. (1973) *J. Mol. Biol. 81*, 93−105.
Arnott, S., Chandrasekaran, R., Birdsall, D. L., Leslie, A. G. W., & Ratliff, R. L. (1980) *Nature 283*, 743−746.
Arnott, S., Chandrasekaran, R., Hall, H., & Puigjaner, L. C. (1983) *Nucleic Acids Res. 11*, 4141−4155.
Engel, J. D., & von Hippel, P. H. (1974) *Biochemistry 13*, 4143−4158.
Fazakerley, G. V., Teoule, R., Guy, A., & Guschlbauer, W. (1984) *FEBS Lett. 176*, 449−452.
Fazakerley, G. V., Teoule, R., Guy, A., Fritzsche, H., & Guschlbauer, W. (1985) *Biochemistry 24*, 4540−4548.
Fazakerley, G. V., Quignard, E., Teoule, R., Guy, A., & Guschlbauer, W. (1987) *Eur. J. Biochem. 167*, 397−404.
Gupta, G., Sarma, M. H., & Sarma, R. H. (1988) *Biochemistry 27*, 7909−7919.
Jardetsky, O., & Roberts, G. C. K. (1981) *NMR in Molecular Biology*, Chapter 13, Academic Press, New York.
Kennedy, M. A., Nuutero, S. T., Davis, J. T., Drobny, G. P., & Reid, B. R. (1993) *Biochemistry 32*, 8022−8035.
Kim, S.-G., & Reid, B. R. (1992) *Biochemistry 31*, 12103−12116.
Kim, S.-G., Lin, L.-J., & Reid, B. R. (1992) *Biochemistry 31*, 3564−3574.
Lane, A. N. (1989) *Biochem. J. 259*, 715−724.
Lane, A. N. (1991) *Biochem. J. 273*, 383−391.
Lefevre, J.-F., Lane, A., & Jardetzky, O. (1985) *FEBS Lett. 190*, 37−40.
Lefevre, J.-F., Lane, A., & Jardetzky, O. (1987) *Biochemistry 26*, 5076−5090.
Lefevre, J.-F., Lane, A., & Jardetzky, O. (1988) *Biochemistry 27*, 1086−1094.
Leroy, J. L., Kochoyan, M., Huynh-Dinh, T., & Gueron, M. (1988) *J. Mol. Biol. 200*, 223−238.
McAteer, K., Ellis, P. D., & Kennedy, M. A. (1995) *Nucleic Acids Res. 23*, 3962−3966.
Quignard, E., Fazakerley, G. V., Teoule, R., Guy, A., & Guschlbauer, W. (1985) *Eur. J. Biochem. 152*, 99−105.
Ravishankar, G., Swaminatham, S., Beveridge, D. L., Levery, R., & Sklenar, H. (1989) *J. Biomol. Struct. Dyn. 6*, 669.
Rein, R., Ornstein, R. L., & MacElroy, R. D. (1978) *Proc. Indian Acad. Sci. 87*, 135−145.
Rinkel, L. J., van der Marel, G. A., van Boom, J. H., & Altona, C. (1987) *Eur. J. Biochem. 163*, 275−286.
Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer Verlag, New York.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286−292.
Tirado, M. M., & de la Torre, G. J. (1980) *J. Chem. Phys. 73*, 1986−1993.
Zhu, L., & Reid, B. R. (1995) *J. Magn. Reson., Ser. B 106*, 227−235.
Zhu, L., Salazar, M., & Reid, B. R. (1995) *Biochemistry 34*, 2372−2380.

BI951364K