Decreased Interstrand H2–H1′ Distance in the GC-Rich Part of the Duplex d(CCTCAAACTCC)·d(GGAGTTTGAGG) in Solution at Low Temperature: Proton Nuclear Magnetic Resonance Investigation†

Nikolai B. Ulyanov,‡,§ Mukti H. Sarma,‡ Victor B. Zhurkin,§ and Ramaswamy H. Sarma*,‡

Institute of Biomolecular Stereodynamics, Department of Chemistry, State University of New York, Albany, New York 12222, and The Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, 117984 Moscow B-334, Russia

Received August 7, 1992; Revised Manuscript Received November 25, 1992

ABSTRACT: The non-self-complementary undecadeoxyribonucleotide duplex d(CCTCAAACTCC)·d(GGAGTTTGAGG) was studied by one- and two-dimensional NMR methods in solution at low and room temperatures. The width of the minor groove of the duplex was determined on the basis of the NOE's between adenine's H2 protons and H1′ protons from the complementary strand. In agreement with the previous reports, we found that the A_3·T_3 block forms a structure with a narrow minor groove at 5 °C, with the H2–H1′ interstrand distance decreasing in the 5′-to-3′ direction along the strand of adenines. Surprisingly, this distance is still short in the GC-rich part of the duplex downstream from the A-tract. This finding is interpreted in terms of pronounced buckle angles in the oligo(purine)·oligo(pyrimidine) blocks, which diminish the H2–H1′ interstrand distances. Both CA_n- and A_nC- junctions have distinct patterns of the proton chemical shifts, which suggests that both junctions may have some specific conformations in solution. Also, we report the temperature-driven changes in proton chemical shifts, which are significant in all parts of the duplex, except the 5′-ends of both strands. The structural interpretation for these changes is proposed, on the basis of the following notion: at low temperature the narrow minor groove is formed between the central A_3·T_3 trimer and the 3′-ends of both DNA strands, while the 5′-ends remain relatively exposed to the solvent.

The sequence-directed intrinsic bending, or curvature, of DNA was discovered 10 years ago (Marini et al., 1982). Since then, this phenomenon has been under intensive structural investigation due to its apparent connections with biologically important processes (Trifonov, 1985; Crothers et al., 1990; Hagerman, 1990). The strongest curvature of free DNA *in solution* is associated with repeating tracts of adenines (Diekmann, 1986; Koo et al., 1986); however, various other sequences were also shown to contribute to the bending of DNA (Shlyakhtenko et al., 1990; Milton et al., 1990). The narrow minor groove in A_n·T_n is one of the main structural distinctions between the A-tracts and the average B form for random nucleotide sequences (Alexeev et al., 1987; Lipanov & Chuprina, 1987). The ordered spine of hydration (Kopka et al., 1983) is believed to be one of the factors stabilizing a narrow minor groove in AT-rich sequences (Chuprina, 1987).

Steady progress in two-dimensional ^{1}H NMR[1] spectroscopy makes it a powerful tool for the investigation of DNA structure in solution, even despite its intrinsic difficulties connected with certain underdetermination of a structure (Metzler et al., 1990; Ulyanov et al., 1992). In particular, measurements of NOE's between the H2 proton of adenines and H1′ of a 3′-neighboring residue from the complementary strand proved to be very useful in the determination of a minor groove width for AT-rich sequences (Kintanar et al., 1987; Katahira et al., 1988, 1990a; Nadeau & Crothers, 1989; Chuprina et al., 1991). While the distance between these two protons is about 5 Å in standard B DNA, it can be below 4 Å in structures with a narrow minor groove. Among the sequences where this interstrand distance was shown to be short are A_n, A_nT_m, and (in certain cases) the step GA·TC in the context of GA_n·T_nC [see a survey by Chuprina et al. (1991)].

In this paper we report the assignments of proton NMR spectra and a qualitative structural analysis of a non-self-complementary DNA duplex (further, the ACT/AGT duplex) at low (5 °C) and room (25 °C) temperatures:

```
5'-C1  C2  T3  C4  A5  A6  A7  C8  T9  C10 C11
   G22 G21 A20 G19 T18 T17 T16 G15 A14 G13 G12-5'
```

We demonstrate for the first time that a decreased width of the minor groove can occur even in a GC-rich part of a DNA double helix in solution. Namely, the interstrand distance between H2A14 and H1′C10 is below 4 Å at low temperature, despite the fact that the step T9–C10·G13–A14 is surrounded by GC pairs.

MATERIALS AND METHODS

DNA Synthesis and Purification. The undecamers d(CCTCAAACTCC) and d(GGAGTTTGAGG) were synthesized on a DNA synthesizer (Applied Biosystems Model 380A) following the method of Matteuci and Caruthers (1981). The products were purified on a 1.1 × 50 cm column of Q-Sepharose (Pharmacia) with a linear gradient of 0.2–0.8 NaCl in 10 mM NaOH and further purified by several ethanol precipitations.

† This work was supported by a grant from the National Institutes of Health (GM29787) and by a contract from the National Foundation of Cancer Research. The high-field NMR experiments were performed at the NMR facility for Biomolecular Research located at the Francis Bitter National Magnet Laboratory, MIT. The NMR facility is supported by Grant RR00995 from the Division of Research Resources of the NIH and by the National Science Foundation under Contract C-6470.

‡ State University of New York, Albany.

§ Institute of Molecular Biology, Moscow.

[1] Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; DNA, deoxyribonucleic acid; 1D, one dimensional; 2D two dimensional; H-bond, hydrogen bond; R, purine; Y, pyrimidine.

0006-2960/93/0432-6875$04.00/0

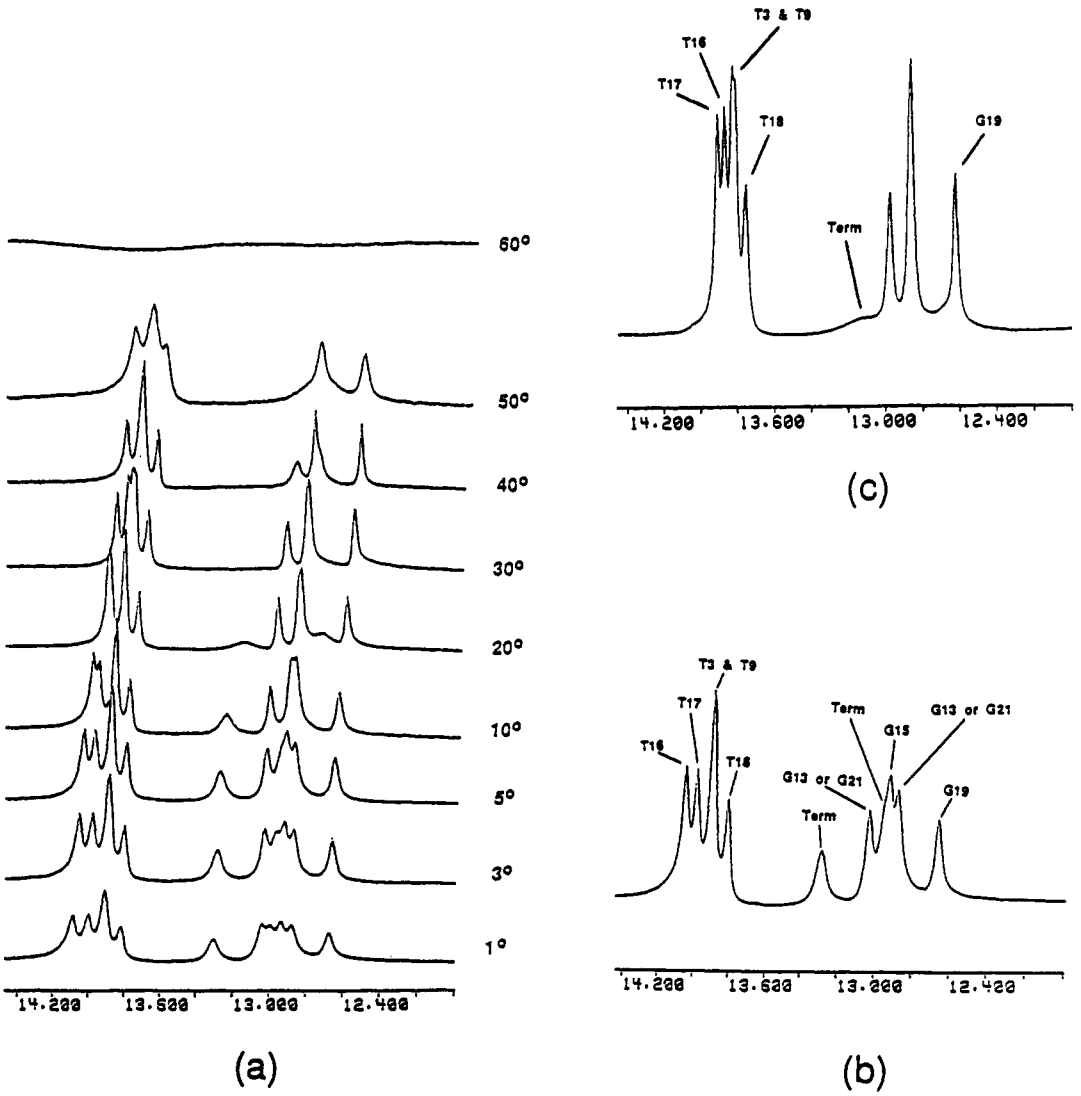


FIGURE 1: Imino proton NMR spectra of the ACT/AGT duplex in water. (a) Temperature dependence. (b and c) Results of the assignment of the imino protons at 5 and 25 °C, respectively. The imino protons in residues T3 and T9 are indistinguishable. No special attempts were made to assign imino protons in guanines; only those are shown which were assigned from the 1D NOE experiments with irradiation of the thymine's imino protons. "Term" denotes the resonances of imino protons in terminal guanines. Note that the positions of the imino protons of T16 and T17 were interchanged at the elevated temperature.

NMR Spectroscopy. For both H_2O and D_2O samples, the DNA concentration was 2.0 mM in duplex, and the salt concentration was 100 mM in NaCl (pH 7.0 in 10 mM sodium phosphate buffer with 1 mM EDTA). 1D NMR spectra of the ACT/AGT duplex in water were recorded for temperatures 1–60 °C, by the use of a time-shared long pulse sequence. NOE difference spectra of the ACT/AGT duplex at 5 and 25 °C in H_2O were recorded with mixing time τ_m = 200 ms for the imino protons with a relaxation delay RD = 1 s and 5000 transients (NS). All spectra were recorded at 500 MHz.

The COSY spectrum of the ACT/AGT duplex at 14 °C in D_2O was recorded with the pulse sequence $(RD–90°–t_1–90°–acq)_{NS}$ in the pure absorption mode (States et al., 1982) with RD = 1.2 s and a data matrix (t_2 = 2048 × t_1 = 256) for NS = 64. The NOESY spectra at τ_m = 50, 100, 150, and 300 ms were collected at 5 °C with RD = 1.5 s for NS = 64 and at 25 °C with RD = 5 s for NS = 16, with the pulse sequence $(RD–90°–t_1–90°–\tau_m–90°–acq)_{NS}$. All 2D spectra were Fourier transformed into a matrix (t_2 = 2048 × t_1 = 2048) with zero filling; the data were symmetrized.

RESULTS

Resonance Assignment at 5 °C. The temperature dependence of the ^1H NMR spectra in water (Figure 1a) shows that all AT pairs are very stable due to their location in the interior part of the ACT/AGT duplex. The duplex melts between 50 and 60 °C; resonances of the terminal GC pairs disappear at 30 °C. In order to assign H2 protons of adenines, we conducted 1D NOE experiments in water, irradiating resonances of iminoprotons in AT pairs (not shown). In the same experiments, we assigned the imino protons of thymines and that of some guanines (Figure 1b). The assignments of H2 protons were independently confirmed by 2D NOESY spectra in D_2O. Strong cross peaks H2A5–H2A6 and H2A6–H2A7 are shown in Figure 2a, and weak cross peaks H2A6–H1′T18, H2A7–H1′T17, and H2A14–H1′C10 are shown in Figure 2b. The rest of nonexchangeable protons (with the exception of H5′ and H5″) were assigned (Table I) with the use of various sequential connectivities in 2D NOESY spectra at high mixing times (300 and 150 ms). The procedure of the 2D NOESY-based assignment of proton NMR spectra of oligonucleotides is well established, so we will outline only some details of it.

Figure 2b shows the standard H6/H8–H1′ walk together with a similar H6/H8–H3′ walk at the mixing time of 300 ms. Note that at the experimental conditions used (low temperature and high mixing time) the process of spin diffusion is very effective, therefore, intra- and interresidue H6/H8–H3′ cross peaks are very strong. The only exceptions are the

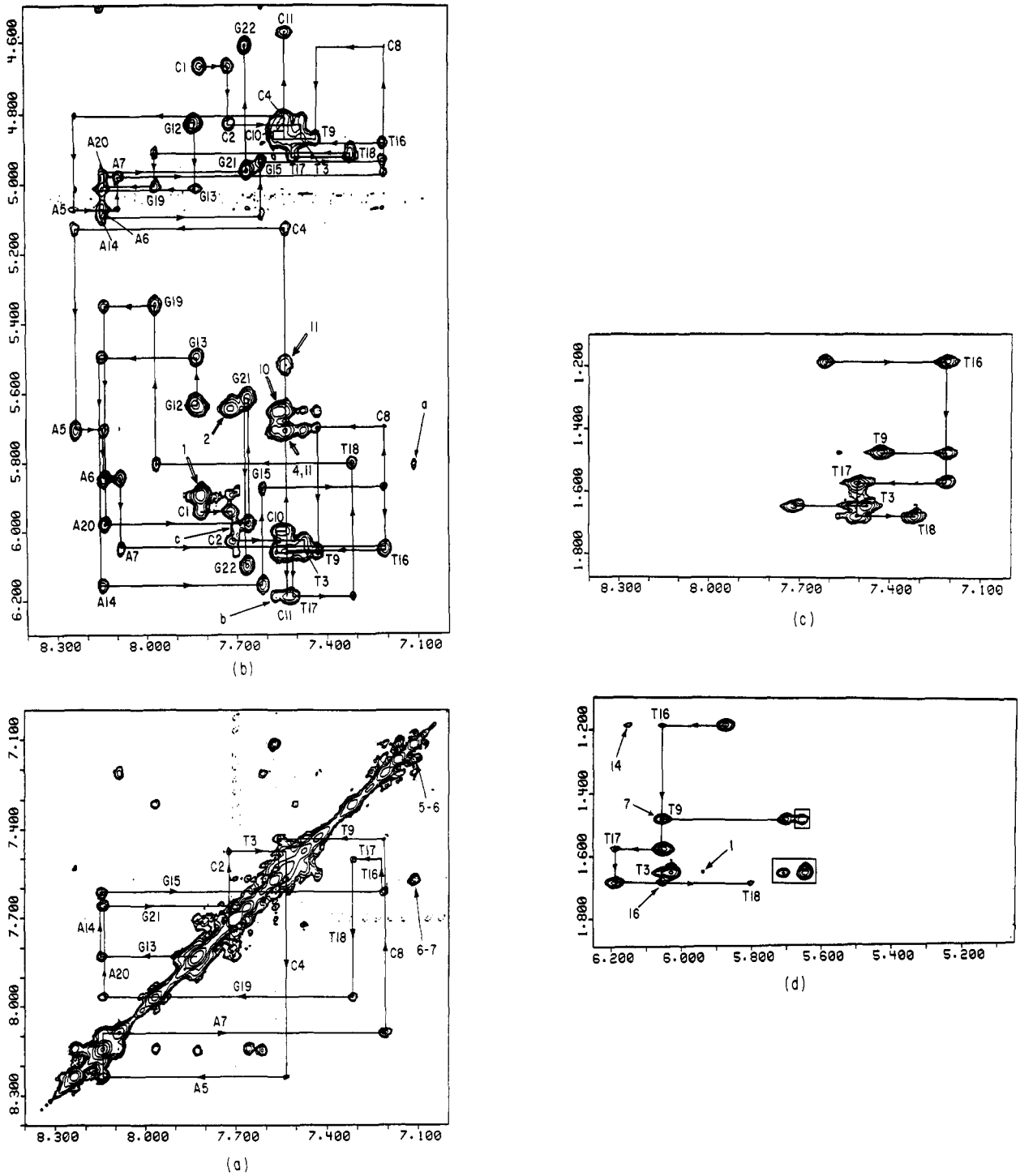


FIGURE 2: Representative portions of the 2D NOESY spectra of the ACT/AGT duplex in D_2O at 5 °C. (a, b, and d) The mixing time is 300 ms; (c) the mixing time is 100 ms. (a) The connectivities between H6/H8 protons in the neighboring residues. Those segments are labeled which connect two cross peaks between the same base proton and its two neighbors. "5–6" and "6–7" refer to the cross peaks H2A5–H2A6 and H2A6–H2A7, correspondingly. Only one of the two symmetric cross peaks was labeled for each interproton contact. (b) The H6/H8–H1′ and H6/H8–H3′ connectivities. The intraresidue cross peaks are labeled. The open numbered arrows indicate the intracytosine H6–H5 cross peaks. The thin arrows, which are marked by "a", "b", and "c", show the H2A6–H1′T18, H2A7–H1′T17, and H2A14–H1′C10 cross peaks, respectively. The interresidue cross peaks between H6/H8 and the next or the previous H5 are unmarked. (c) The H6/H8–methyl connectivities. The intrathymine H6–CH_3 cross peaks are labeled; the cross peaks between a methyl group and the previous H6/H8 proton are unlabeled. (d) The H1′–methyl connectivities. The connectivities are labeled similar to those in panel c. Note that the interresidue H1′–CH_3 cross peaks are much more intense than the intraresidue ones. The interresidue cross peaks between CH_3 and H5 protons are boxed. The numbered arrows with the numbers i show the cross peaks between H1′(i) and $CH_3(i+2)$.

cross peaks which involve H3′ of A5, A6, C4, and C8, the first two of them, probably, because their resonances are very close to the residual HDO peak (5.045 ppm). We found it very helpful to use both kinds of connectivities simultaneously for the assignment of base, H1′, and H3′ protons (still, the connectivities with H1′ protons are also very important for the assignment of H3′ protons). It is seen that some of the resonances are resolved reasonably well. On the other hand, some resonances (e.g., both base and H1′ protons in residues T3, T9, and C10) are heavily overlapped, which prevents their reliable assignment without additional sources of information. Fortunately, there is plenty of such information. Thymine's base protons are readily recognized on the basis of their cross peaks with methyl groups (Figure 2c). Moreover, the

Table I: Proton Chemical Shifts (ppm) of d(CCTCAAACTCC)·d(GGAGTTTGAGG) in Solution at 5 °C with TSP as an Internal Standard

	H6/H8	H2/H5 CH_3	H1′	H2′	H2″	H3′	H4′	imino
C1	7.81	5.89	5.94	2.32	2.56	4.67	4.12	
C2	7.72	5.64	6.02	2.21	2.54	4.83	4.22	
T3	7.48	1.65	6.07	2.15	2.49	4.87	4.19	13.90
C4	7.54	5.71	5.13	2.00	2.21	4.81	4.04	
A5	8.24	7.17	5.71	2.75	2.84	5.07	4.38	
A6	8.14	7.11	5.85	2.64	2.84	5.07	4.46	
A7	8.09	7.57	6.05	2.60	2.82	4.98	4.47	
C8	7.21	(5.05)[a]	5.70	1.93	2.45	4.62	4.20	
T9	7.43	1.48	6.06	2.16	2.54	4.87	(4.20)[a]	13.90
C10	7.56	5.65	6.00	2.16	2.46	4.84	4.15	
C11	7.53	5.71[b]	6.19	2.30	2.25	4.57	3.99	
G12	7.84		5.63	2.52	2.72	4.83	4.21	12.96/13.29
G13	7.83		5.50	2.69	2.74	5.02	4.35	12.89/13.04
A14	8.15	7.70	6.16	2.75	2.96	5.08	4.50	
G15	7.61		5.87	2.48	2.77	4.94	4.46	12.93
T16	7.21	1.19	6.06	2.17	2.61	4.88	4.31	14.05
T17	7.51	1.58	6.19	2.22	2.63	4.92	4.19	13.99
T18	7.32	1.68	5.81	2.00	2.38	4.91	4.13	13.81
G19	7.97		5.35	2.69	2.71	5.01	4.35	12.67
A20	8.14	7.70	5.98	2.66	2.86	4.96	4.44	
G21	7.66		5.61	2.52	2.66	4.97	4.35	12.89/13.04
G22	7.67		6.10	2.43	2.33	4.61	4.21	12.96/13.29

[a] Tentative assignments. [b] There are two species for H5C11, at 5.711 and 5.512 ppm (see text, under Results).

sequential connectivities H8G15 → MetT16 → H6T16 → MetT17 → H6T17 → MetT18 → H6T18 make it possible to assign these protons sequence-specifically, independently of the H6/H8–H1′ connectivities. Similar connectivities exist for H1′–Met cross peaks also (Figure 2d). The intraresidue H1′–Met cross peaks are much weaker than the interresidues ones; the former cross peaks exist mainly due to spin diffusion. It is interesting that under these conditions we have several examples of even longer pathways of spin diffusion, see, e.g., cross peaks between H1′(i) and Met(i+2) (shown by numbered arrows in Figure 2d; the numbers refer to the residues with participating H1′ protons). There are several possible pathways of magnetization transfer: H1′(i) → Met(i+1) → Met(i+2) (if residue i+1 is a thymine), H1′(i) → H6/H8(i+1) → Met(i+2), H1′(i) → H2″(i) → H6/H8(i+1) → Met(i+2), and some others.

The assignment of T3 and T9 residues, which are surrounded by cytosines, is more difficult, because the intensity of intraresidue H6–Met cross peaks in thymines is comparable to that of interresidue NOE's between H6 of 5′-flanking cytosines and the methyl group (Figure 2c). The cytosines were assigned by their intraresidue H6–H5 cross peaks (Figure 2b), and interresidue cross peaks H5(i)–Met(i–1), H5(i)–Met(i+1) (Figure 2d), and H5(i)–H1′(i–1) (not shown). The situation with H5 protons of cytosines is rather confusing. First, the cross peak H6C8–H5C8 is apparently missing from the NOESY spectra at 5 °C (Figure 2b). The only possible explanation for this is an overlapping of the H5C8 resonance with the residual HDO peak (5.05 ppm). Indeed, the cross peak is present in the NOESY spectrum at a higher temperature, after the HDO peak was moved upfield (Figure 3). Second, there are two species for the H5 in the terminal residue C11, the major one being at 5.71 ppm and the minor one at 5.51 ppm. The species at 5.71 ppm is identified by a cross peak with the H6C10 proton. The minor species disappears at higher temperatures (Figure 3), which suggests that it might be stabilized by an intermolecular interactions at low temperature. We are unaware of previous reports of multiple species for proton resonances in free Watson–Crick duplexes of DNA in solution. However, multiple species of base protons were previously observed for a complex of DNA

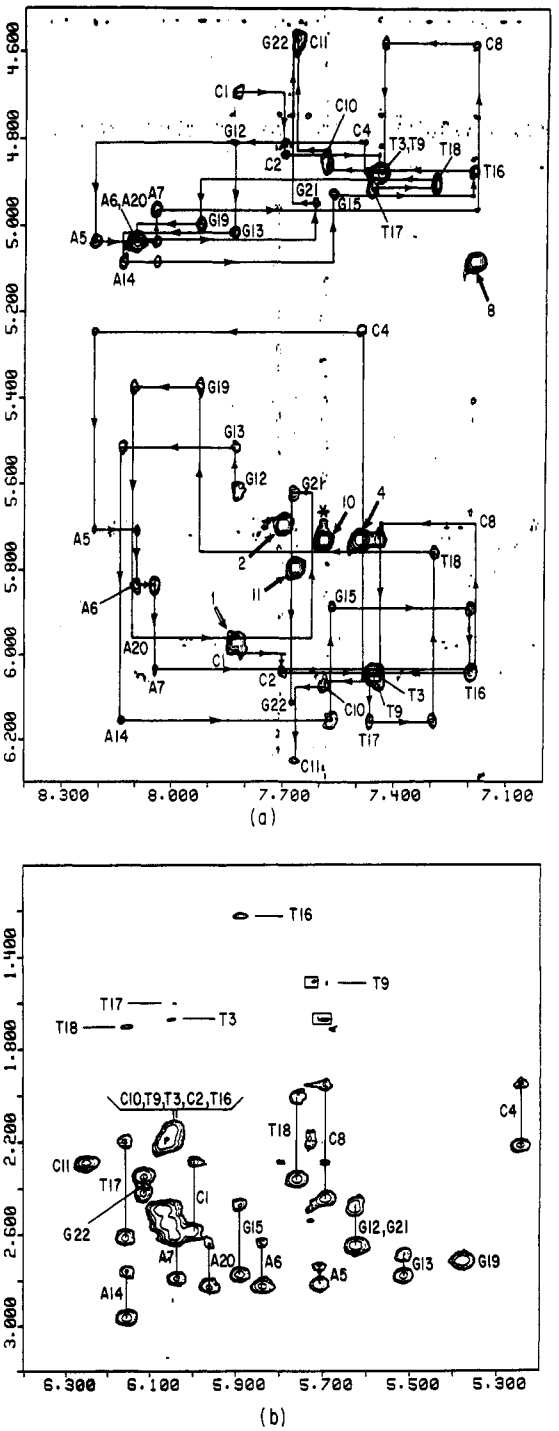


FIGURE 3: Portions of the 2D NOESY spectra of the ACT/AGT duplex in D_2O at 25 °C. The mixing time is 300 ms. (a) The H6/H8–H1′ and H6/H8–H3′ connectivities; see the legend to Figure 2b. The "star" symbol indicates the H2A7–H1′C8 cross peak. (b) The connectivities between H1′ and H2′/H2″ protons. Also, the cross peaks between H1′(i–1) and CH_3(i) are shown (compare with Figure 2d). The interresidue CH_3–H5 cross peaks are boxed. The cross peaks between H5 and H2′/H2″ protons are not marked.

decanucleotide with a minor groove binding drug (Leupin et al., 1986) and in A·A base pairs of a parallel self-complex of d(AAAAACCCCC) at acidic pH (Luo et al., 1992). Specific intermolecular end-to-end contacts, which presumably stabilize the multiple species of H5C11 in the ACT/AGT duplex at low temperature, were reported by us earlier for the *Bam*HI DNA hexamer at 7 °C (Ulyanov et al., 1992). The assignments of base protons in cytosines of the ACT/AGT duplex were confirmed by the COSY spectra at 14 °C (not shown).

Table II: Differences in Proton Chemical Shifts (ppm) of d(CCTCAAACTCC)·d(GGAGTTTGAGG) between 25 and 5 °C

	H6/H8	H2/H5 CH_3	H1′	H2′	H2″	H3′	Σ[a]	imino
C1	0.01	0.08	0.06	0.04	0.01	0.02	0.22	
C2	–0.01	0.05	0.02	0.00	0.00	0.00	0.08	
T3	–0.03	0.02	–0.01	0.00	0.00	0.00	0.04	–0.09
C4	–0.04	0.02	0.12	–0.05	0.00	–0.01	0.24	
A5	–0.02	0.02	0.00	–0.02	–0.02	–0.03	0.11	
A6	–0.05	0.04	–0.01	–0.01	–0.02	–0.03	0.16	
A7	–0.04	0.02	–0.01	0.03	–0.4	–0.2	0.16	
C8	–0.01	0.03	–0.01	0.02	–0.01	–0.04	0.12	
T9	–0.01	0.03	0.00	0.06	nd[b]	–0.01	>0.11	–0.09
C10	0.02	0.08	0.08	nd	nd	–0.01	>0.19	
C11	0.14	0.09	0.06	–0.01	0.04	–0.01	0.35	
G12	0.01		–0.01	–0.05	–0.07	–0.02	0.16	nd
G13	0.01		0.02	0.00	0.03	0.00	0.06	nd
A14	–0.01	0.02	0.00	0.00	–0.01	0.00	0.04	
G15	–0.04		0.02	–0.01	0.00	–0.01	0.08	nd
T16	–0.02	0.03	–0.02	–0.04	–0.02	–0.01	0.14	–0.19
T17	–0.04	0.02	–0.03	–0.02	–0.02	–0.01	0.14	–0.09
T18	–0.02	0.02	–0.05	0.00	–0.03	–0.01	0.13	–0.06
G19	–0.04		0.03	0.03	0.01	–0.01	0.12	–0.05
A20	–0.04	0.3	–0.02	–0.03	–0.04	0.08	0.24	
G21	–0.04		0.01	–0.04	–0.04	–0.02	0.15	nd
G22	0.01		0.01	–0.01	0.02	–0.01	0.06	nd

[a] A "sum index", which is calculated for each residue as a sum of absolute values of the temperature changes in chemical shifts for all assigned nonexchangeable protons. [b] Not determined.

Finally, the assignments for the H6/H8 protons are confirmed by the connectivities between neighboring base protons shown in Figure 2a. Almost all cross peaks are strong at these conditions; the exceptions are the cross peaks H6T16–H6T17, H8C8–H6T9, and H6C4–H8A5. The sequence of the ACT/AGT duplex is rather heterogeneous; therefore, many of these cross peaks are located far from the diagonal, which simplifies the assignment.

The resonances of the H2′, H2″, and H3′ protons were assigned mainly by their cross peaks with H1′ in the same residue (not shown). H4′ protons have cross peaks with H1′, and some of them have cross peaks with H6/H8 protons in the same residue. Also, in some steps there are interresidue H1′(i)–H4′(i+1) cross peaks. This makes it possible to assign the majority of H4′ protons in the ACT/AGT duplex. Besides, a convenient way to assign the H2′, H2″, H3′, and H4′ protons in the residues at the 5′-flank of a thymine is by their NOE's with the methyl group (not shown). The H6/H8–H2′/H2″ connectivities were not found to be a very useful in assigning H2′/H2″ because of the strong overlapping in this spectral region. However, the H6/H8–H2′/H2″ cross peaks were used in order to verify the assignments of base and H2′/H2″ protons. And again, the assignments of sugar protons were double-checked by the H3′–H2′/H2″, H3′–H4′, and H4′–H2′/H2″ cross peaks in those cases when these cross peaks are reasonably separated (not shown).

Resonance Assignment at 25 °C. The assignment of the proton NMR spectra of the ACT/AGT duplex at 25 °C has been done in a way similar to that at the low temperature. Figure 3 shows the results of the assignment of base protons and H1′, H3′, H2′, and H2″ protons. No attempts were made to assign H4′, H5′, and H5″ protons. Assignments of the imino protons are shown in Figure 1c. The differences in chemical shifts of proton resonances between 25 and 5 °C are given in the Table II. NOE's between various protons of the ACT/AGT duplex are usually less intense at room temperature than at 5 °C. Many of the base-to-H1′ cross peaks are missing (Figure 3a); the same is true for some other spectral regions. Partially, this effect can be explained by a decreased effective correlation time at these conditions compared to the low temperature. It leads to a less effective transfer of magnetization, both direct and indirect.

Measurement of Distances between H2 Protons of Adenines and H1′ Protons across the Minor Groove. In the ACT/AGT duplex there are five dinucleotide steps in which the H2–H1′ interstrand contact can be potentially observed: T3–C4·G19–A20, C4–A5·T18–G19, A5–A6·T17–T18, A6–A7·T16–T17, and T9–C10·G13–A14. The first two of them have NOE's below the noise level, and the other three are clearly seen at 5 °C (Figure 2b, thin arrows marked "a", "b", and "c").

The interproton distances for these three contacts were determined, using the equation $r = r_{ref}(noe_{ref}/noe)^{1/6}$, where r is the distance of interest, r_{ref} is a fixed reference distance, and noe and noe_{ref} are the corresponding NOE's. The cytosine H6–H5 distance (2.44 Å) was used as a reference distance. The NOE's were quantified as areas of the corresponding cross peaks in slices through H2 and H1′; the results obtained from the H2 → H1′ and H1′ → H2 NOE's were averaged. The well-resolved H6C2–H5C2 cross peak was used for the measurements of a reference NOE, noe_{ref}. Only the data collected at the lowest mixing time, 50 ms, were used for the calculations. There are no intervening protons between H2 and H1′; however, at higher mixing times the NOE between H6 and H5 of cytosine becomes affected by a spin diffusion and can no longer serve as a reference NOE. Even in the range of mixing time from 50 to 100 ms the H6–H5 NOE builds up nonlinearly for the DNA undecamer at 5 °C due to its high effective correlation time of 11 ns (Ulyanov et al., unpublished calculations).

The cross peak between H2A6 and H1′T18 coincides with the H2A6–H1′A6 cross peak at low temperature; also, the H2A14–H1′C10 and H2A20–H1′A20 cross peaks overlap (compare Table I with Figure 2b). The intraresidue distance H2–H1′ is about 4.5 Å for the C2′-endo/anti conformations of deoxyadenosine. The corresponding NOE's together with the reference H6–H5 NOE were calculated by the full-matrix NOESY simulation method (Gupta et al., 1988). After the appropriate normalization, the calculated intraresidue NOE's were subtracted from the observed magnitudes of the overlapped cross peaks. The outlined procedure leads to the following distances for the ACT/AGT duplex at 5 °C: H2A6–H1′T18 is 3.9 Å, H2A7–H1′T17 is 3.5 Å, and H2A14–H1′C10 is 3.7 Å.

In contrast, all of these contacts are absent from the spectra at 25 °C. There are two possible explanations for this absence. First, due to a decreased effective correlation time at room temperature, the rate of cross relaxation is decreased also. It is possible that the H2–H1′ distances across the groove are the same at both temperatures, but we cannot detect them at 25 °C with the experimental procedure used. However, we detected the cross peak between H2A7 and H1′C8 (marked by the "star" in Figure 3a); at the low temperature this cross peak is not seen because it coincides with the strong H6C10–H5C11 cross peak (Figure 2b). The corresponding distance, along the strand of DNA, is usually comparable with the distance across the groove (Chuprina et al., 1991). The second explanation is that the distances really become larger with the elevation of temperature. This possibility is favored by the fact that the structure of the ACT/AGT duplex at room temperature appears to be significantly different from that at 5 °C (see Discussion). Anyway, both possibilities cannot be ruled out; the question about the H2–H1′ distances in the ACT/AGT duplex at room temperature requires additional investigation.

DISCUSSION

The two strands of the ACT/AGT undecamer form a full duplex at both low and room temperature, although the terminal GC pairs are destabilized at 25 °C (Figure 1). The patterns of sequential connectivities in the NOESY spectra suggest that the conformation of the duplex belongs to the B family of forms. However, qualitative analysis of the NOE magnitudes and of proton chemical shifts indicates that the structure is not uniform. Below we discuss some of the irregularities in the structure of the ACT/AGT duplex.

Block of Three Consecutive Adenines. In agreement with previous investigations (Katahira et al., 1988; Nadeau & Crothers, 1989), we found that the run of adenines forms a special structure at the low temperature, which is characterized by a decreased distance between adenine's H2 and H1′ of a 3′-neighboring nucleotide from the complementary strand. Again in accordance with the previous reports, this distance decreases gradually in the 5′-to-3′ direction along the strand of adenines. While the corresponding cross peak is not observed at the 5′-end of the A tract (i.e., for the H2A5–H1′G19 contact), they are present for the other two adenines (A6 and A7). Although the precision of measurements of these distances is questionable, especially in the range of $r \sim 4$ Å, nevertheless, the qualitative tendencies are clear.

It was suggested that the cooperative formation of the "anomalous" narrow-grooved structure is responsible for the shortening of the interstrand H2–H1′ distances (Nadeau & Crothers, 1989; Chuprina et al., 1991). The reported sequences, which possess a decreased H2–H1′ interstrand distance in solution, are limited to the A_n and A_nT_m blocks with various numbers of adenines and thymines; also, the step GA·TC in the context of $GA_n{\cdot}T_nC$ may, in certain cases, have this "anomalous" conformation (Chuprina et al., 1991). In view of these reports, it was quite unexpected that we found a decreased H2–H1′ interstrand distance in the TC·GA step which is separated from the A_3 block by another GC pair. This finding extends the list of sequences where the conformation with a narrowed minor groove can be formed. It also puts forward a question of whether the presence of A-tract is necessary at all for the formation of a structure with a decreased H2–H1′ distance across the groove. On the one hand, it is possible that the shortening of the H2A14–H1′C10 distance in the ACT/AGT duplex at the low temperature is a consequence of the upstream-located block of three adenines. If so, it is interesting how far along the helix the A-tract can affect the DNA structure, and what are the restrictions for intervening sequences. Alternatively, it is conceivable that a relatively narrow minor groove may be formed without participation of A_n or A_nT_m blocks. A preliminary confirmation for the second alternative comes from the NMR investigations of DNA dodecamers containing the *Bam*HI site GGATCC in the middle of duplexes (Chuprina et al., 1991). The A_nT_m block in those dodecamers is degenerate, with $n = m = 1$; nevertheless, the H2–H1′ interstrand distance in the steps GA·TC is decreased to 4 Å. [Note that in our own NMR study of the *Bam*HI hexamer the minor groove was found to be wide (Ulyanov et al., 1992); one can explain this apparent discrepancy assuming that for a hexamer the minor groove is not completely formed yet.] Also, it is worthwhile to note that a decreased minor groove width was observed in the middle part of DNA decamer CCAACGTTGG in the crystal state (Grzeskowiak et al., 1991); however, a possible role of crystal packing forces in that phenomenon is unclear. Without any doubt, future structural NMR investigations of a sufficient number of DNA oligonucleotides in solution are necessary to clarify the sequence requirements

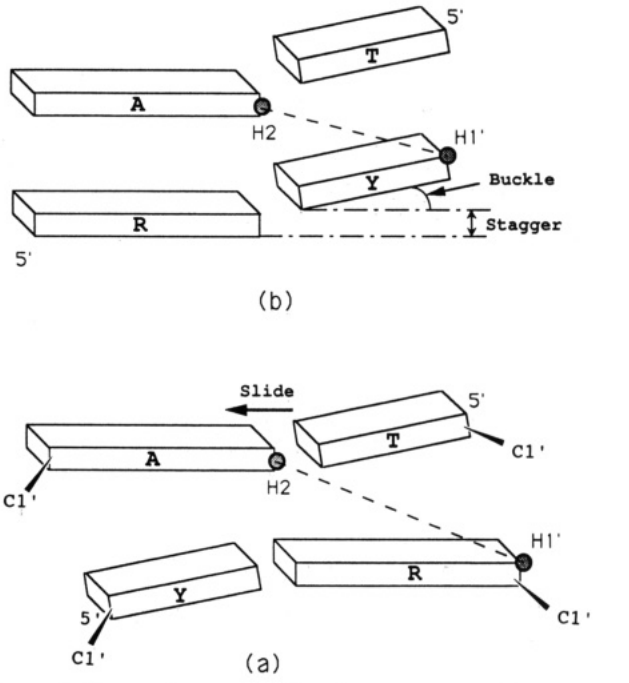


FIGURE 4: Schematic representation of complementary dinucleotides as viewed from the minor groove; sugar–phosphate backbones are omitted for clarity. (a) YpA·TpR base pair step. A positive parameter of slide in this dimer (the top base pair is shifted to the left) relieves the steric clash between purines, A and R, in the minor groove (Calladine, 1982); this slide increases the H2–H1′ interstrand distance (the dashed line). (b) RpA·TpY base pair step. Positive parameters of buckle and stagger decrease the H2–H1′ distance.

for the formation of structures with a narrow minor groove.

Sequence-Dependence of the Interstrand Distances between H1′ and Adenine H2 Protons at Low Temperature. There are several structural mechanisms which could contribute to the variation of the H1′–H2 distances across the minor groove of DNA. These interstrand distances can be observed only in two types of sequences: YA·TR, and RA·TY (Y = pyrimidine; R = purine). The structure of complementary dinucleotides YA·TR is determined, to a great extent, by a potentially unfavorable contact in the minor groove between purines from the opposite strands. Several mechanisms, which relieve this steric clash, were proposed, first by Calladine (1982), and then confirmed and elaborated by conformational calculations (Ulyanov & Zhurkin, 1984) and in single-crystal X-ray structures of DNA oligonucleotides (Prive et al., 1991; Balendiran & Sundaralingam, 1991). Those mechanisms are acting to increase the distance between two purines from the opposite strands, so that the H2–H1′ interstrand distance is also increased. The most important are the sliding motion (Figure 4a) and the rolling of base pairs toward the major groove. This explains why we have not detected the NOE between H2A5 and H1′G19, consistent with other published NMR data for YA·TR sequences where the corresponding distance was found to be not less than 4.5 Å (Chuprina et al., 1991).

The variation of the interstrand H2–H1′ distance in other (RA·TY) sequences can be qualitatively interpreted in the following way. According to the linear dichroism measurements of poly(dA)·poly(dT) (Edmondson & Johnson, 1985), thymines have a greater angle between their base normals and the helix axis than do adenines. On the basis of these data, Koo et al. (1986) hypothesized that AT pairs in $A_n{\cdot}T_n$ blocks have a pronounced buckle angle. Consistent with this model, the energetically optimal structures of oligo(R)·oligo(Y) runs, and $A_n{\cdot}T_n$ in particular, are indeed characterized by a strong buckle with a sign as shown schematically in Figure

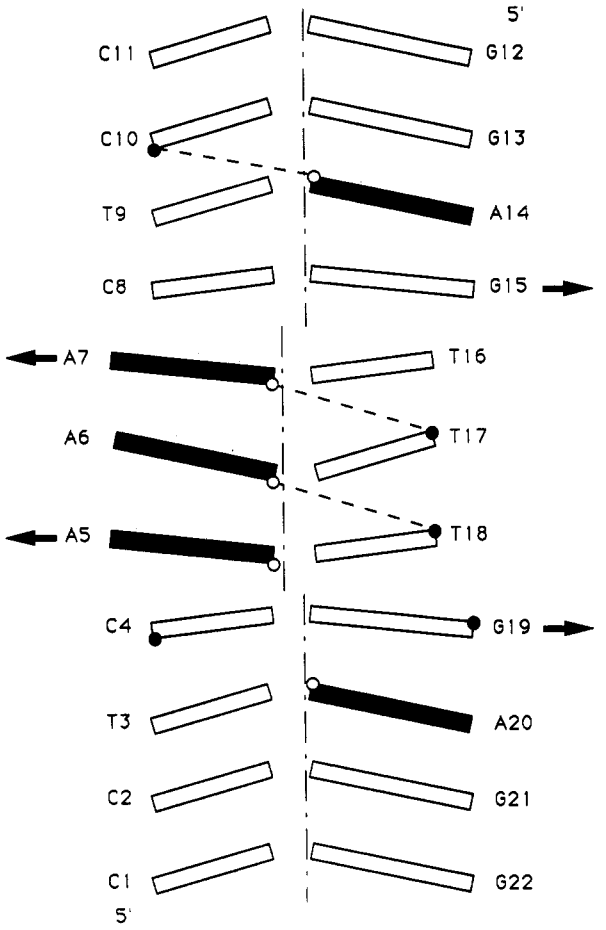


FIGURE 5: Schematic representation of the ACT/AGT duplex; view from the minor groove. Adenines are shown as filled boxes, H2 adenine protons are shown as white circles, and H1′ protons in the residues across the groove are shown as filled circles. Dashed lines show the short H2–H1′ interstrand distances. For adenines A5 and A20, the corresponding cross peaks were not observed. The difference between A5, A20, and other adenines (A6, A7, A14) can be explained by the difference in slide and buckle angle (see text). Vertical dash-dotted lines show the helical axes for oligo(R)·oligo(Y) runs. Note that the undecamer is divided into three such runs which are separated by Y–R and R–Y junctions. These junctions are characterized by opposite signs of slide (shown by arrows) and cup angles [for the definition of a cup, see Yanagi et al. (1991)]: negative cup, or "butterfly)(", for the C4–A5 step, and positive cup, or "rhombic structure ⟨⟩", for the A7–C8 step (Zhurkin et al., 1990).

4b (Edmondson, 1987; Zhurkin et al., 1990). Both pyrimidines, T and Y, rotate counterclockwise, and their right ends move up (Figure 4b), so that H1′(Y) proton is positioned closer to the corresponding H2 proton of adenine. According to this simple scheme, the interstrand H2–H1′ distance is decreased in RA·TY sequences inside R_n·Y_n blocks relative to the standard B form of DNA. It readily explains why we detected the H2A6–H1′T18, H2A7–H1′T17, and H2A14–H1′C10 cross peaks in the ACT/AGT duplex (Figure 5).

To understand the difference between the two GA·TC steps, G19–A20·T3–C4 and G13–A14·T9–C10, the flanking sequences have to be considered. (We recall that the interstrand cross peak is detected in the latter dimer, but not in the former one; see Results.) The 5′-flanks of the two GA dinucleotides are different; namely, the G19–A20·T3–C4 base pair step is localized next to Y_n–A_m junction, while G13–A14·T9–C10 is positioned inside a continuous run of purines. We suggest the following explanation of the difference between these two steps. As noted above, oligo(R)·oligo(Y) runs have a pronounced buckle angle. As a result, the Y_n–A_m and A_n–Y_m junctions differ in terms of the so-called cup angles (Yanagi et al., 1991): the Y_n–A_m junctions have a negative cup, or look like a "butterfly)(" (Figure 5), whereas the A_n–Y_m junctions have a positive cup, or resemble a "rhombic structure ⟨⟩" (Zhurkin et al., 1990). Furthermore, the magnitude of the buckle angle is considerably reduced in the both junctions compared to the central part of the oligopurine runs and to the free ends of a duplex (Zhurkin et al., 1990; see Figure 5). [These predictions are in accord with the 2D NMR data for *Bam*HI DNA hexamer GGATCC (Ulyanov et al., 1992).] "Flattening" of buckles in the vicinity of the Y_n–A_m junction implies that the left end of C4 in Figure 5 moves down significantly less compared to the left end of C10. Accordingly, the interstrand H2A20–H1′C4 distance is increased relative to the distance H2A14–H1′C10 in the center of oligopurine·oligopyrimidine run.

This "mechanistic" scheme neglects the differences between guanines and adenines; nevertheless, it explains the experimentally observed differences between H2A20–H1′C4 and H2A14–H1′C10 distances and between H2A7–H1′T17, H2A6–H1′T18, and H2A5–H1′G19. Also, the scheme is in accord with the statistical analysis performed by Chuprina et al. (1991), who found that YR sequences increase the interstrand H2–H1′ distance in the neighboring RA·TY dimer. However, this scheme does not explain the difference between the H2A6–H1′T18 and H2A20–H1′C4 distances (both are next to the pyrimidine–purine dimer, C4–A5·T18–G19; the first distance is 3.9 Å, whereas NOE for the second one is not observed). This difference may be connected to certain distinctions between guanines and adenines, or it may be a result of the different separation from the end of the duplex.

Our interpretation of the NMR results for the ACT/AGT duplex is in agreement with the X-ray single-crystal structures of DNA oligonucleotides. Indeed, a statistical analysis revealed a correlation between the H2–H1′ interstrand distances and parameters of slide and cup (Chuprina et al., 1991). In addition, we mention a parameter of stagger (Figure 4b), which may also decrease distance H2A–H1′Y. Positive stagger shown in Figure 4b would destabilize bifurcated H-bonds in the major groove (Nelson et al., 1987; Coll et al., 1987), but simultaneously it would decrease the distance N3A–O2T in the minor groove (Kopka et al., 1983), which is important for the spine of hydration in the minor groove (Chuprina, 1987).

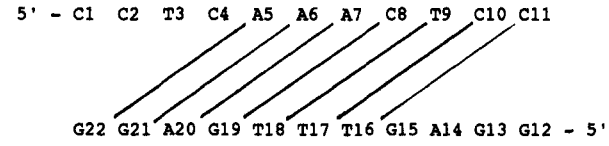
CA_n and A_nC Junctions. Two deoxycytidines, C4 and C8, at the 5′- and 3′-flanks of the block of three adenines, have some of their protons resonating at very unusual positions. First, the H1′C4 proton at the CA_n junction has a rare upfield position of its resonance (5.13 ppm at 5 °C and 5.24 ppm at 25 °C). Second, the H5C8 proton at the A_nC junction is also resonating at an unusually upfield position (about 5.05 and 5.08 ppm at low and room temperatures, respectively). To the best of our knowledge, the H1′ proton of C4 has the highest upfield resonance among the deoxycytidines in the DNA duplexes studied so far. However, similar tendencies of chemical shifts have been reported before. The deoxyguanosine's H1′ resonance can be shifted upfield as high as to 5.08–5.29 ppm, when situated at the 5′-flank of the A_n block (Kintanar et al., 1987; Katahira et al., 1990b). Among the deoxycytidines, the highest upfield-shifted position of H1′ was 5.40 ppm in the sequence CA_6 (Katahira et al., 1988). Also, there have been reports of high upfield positions of the H5 proton resonance in cytosines: 5.18 ppm (Kemming et al., 1987) and 5.08 ppm (Lefevre et al., 1987), both in the context of A_2C. In addition, the H5 proton of the junction cytosine resonates at 5.00 ppm in the DNA duplex A_5C_5·G_5T_5 at low temperature (our unpublished data). These unusual chemical shifts of the H1′ and H5 protons are indicative of the possibility

of some singular conformations for both junctions. We expect to obtain the details of these conformations after refinement of the solution NMR structure for the ACT/AGT duplex, which is in progress now; the preliminary results show that both dimers CA·TG and AC·GT have very pronounced parameters of slide which are of opposite sign (Figure 5).

Temperature-Driven Structural Transitions in the ACT/AGT Duplex. Many of the proton resonances in the ACT/AGT duplex are temperature-sensitive (Table II). The maximum changes occur in the terminal residues, especially in the C11, which has two different species at the low temperature. The large changes in chemical shifts of H1′ and H5 resonances of the penultimate C10 may be explained by its closeness to the terminal C11. Among the other residues, C4 and T18 experience the largest shifts in resonance positions of H1′ protons. It is possible to interpret these changes as an alteration of the interaction of sugars of C4 and T18 with bases of the next residues, A5 and G19. All these residues belong to the single pyrimidine–purine step in the ACT/AGT duplex, C4–A5·T18–G19. It is interesting to note in this respect that the pyrimidine–purine dinucleotides CA·TG are known as very labile sequences: they have a weak overlap between base pairs (Arnott, 1972), they are the most flexible dimers as revealed by NMR techniques (Cheung et al., 1984; Donlan & Lu, 1992) and Monte Carlo simulations of the DNA double helix (Zhurkin et al., 1991), and they can adopt unusual geometries for base pairs in single-crystal oligonucleotides (Timsit et al., 1991). The H1′C4 resonance changes linearly with an increase in temperature, and the changes are not completed yet at 25 °C (data not shown). This observation suggests that the temperature-driven alteration of the C4–A5·T18–G19 dimer is not an "all-or-nothing" transition between two distinct conformations, rather it is a gradual structural transition, similar to the well-known small unwinding of the double helix caused by temperature (Depew & Wang, 1975; Schyolkina et al., 1977). Alternatively, one can speculate that this temperature dependence of the H1′C4 chemical shift reflects a gradual change in equilibrium between several preexisting conformers, like those observed by Timsit et al. (1991) in X-ray structures at low and room temperatures.

Protons of the block A_3·T_3 constitute the next group of protons with temperature-dependent chemical shifts (Table II). One of the structural interpretations for these changes is connected with the possible alteration of the minor groove width in the region of A_3·T_3 with increasing temperature. Certain support for this interpretation comes from the consideration of the imino portion of the NMR spectrum of the ACT/AGT duplex. At the low temperature, the resonances of the imino protons in the AT pairs shift downfield in the 5′-to-3′ direction along the strand of adenines (Figure 1b). Such a sequence of the imino proton resonances is characteristic of the A-tract with a narrow minor groove (Nadeau & Crothers, 1989); this sequence is interrupted at 25 °C (Figure 1c), which is consistent with a disruption of a minor groove structure. However, we were not able to show unambiguously that the H2–H1′ interstrand distances increase at the room temperature (see Results).

Finally, it is interesting to compare the two pieces of the ACT/AGT duplex with the same nucleotide sequence: C2–T3–C4·G19–A20–G21 and C8–T9–C10·G13–A14–G15. Proton resonances in the residues G19, A20, and G21 move more significantly with increase in temperature than in G13, A14, and G15. On the other hand, the changes in the C2–T3–C4 triplet are smaller than in the C8–T9–C10 (with the exception of C4, which belongs to the pyrimidine–purine dimer); this is especially evident when looking at the sum index of changes for each residue (Table II). This seemingly paradoxical fact may also be attributed to the temperature-driven alteration of the minor groove in the region of the A_3·T_3 block. In B DNA, the shortest distances across the minor groove are observed for the residues in the complementary strands, which are shifted in the 3′-direction by three or four nucleotides relative to each other. This shift depends on whether the distance is measured between the sugars (shift of three nucleotides; Burkhoff & Tullius, 1987) or between the phosphates (shift of four nucleotides; Zhurkin et al., 1991). The closest distances between the phosphates are shown by the lines in the scheme below:



As evident from this scheme, the 3′-ends in both strands are juxtaposed to the A_3·T_3 block across the minor groove, whereas the 5′-ends in both chains do not participate in the formation of the groove and thus are more exposed to the solvent. If the increase in temperature causes the widening of the minor groove in the region of the central A_3·T_3 block, such a widening would affect not only the environment of the AT pairs but that of the 3′-ends as well. Such changes may influence the proton chemical shifts in these sequences, which is consistent with the data presented in the Table II.

SUMMARY

In this paper we report a qualitative and semiquantitative analysis of the 500-MHz 2D NOESY NMR spectra of the non-self-complementary mirror repeat undecamer:

```
5'-C1  C2  T3  C4  A5  A6  A7  C8  T9  C10 C11
     G22 G21 A20 G19 T18 T17 T16 G15 A14 G13 G12-5'
```

The oligomer is called ACT/AGT duplex for short. Information about the sequence dependence of the minor groove width in DNA double helixes is crucial for a sound understanding of such phenomena as protein recognition, drug binding, and DNA bending. The distance between adenine H2 in one strand and the H1′ of the nucleotide, located diagonally in the complementary strand (see slashes in the sequence above), is believed to be indicative of the minor groove width. A narrow minor groove is widely accepted to be characteristic of A-tracts, sequences with persistent repetitions of adenines. This is confirmed by numerous NMR studies of oligonucleotides in solution.

We have estimated the H2–H1′ interstrand distances in the ACT/AGT duplex from NOE data at low mixing time, using the two-spin approximation. In agreement with previous studies, we find the distances to be shorter in the AAA·TTT block of the ACT/AGT duplex than in the regular B-form of DNA. However, we also found that, at *low temperature*, this distance is still short in the GC-rich part of the duplex, two base pair steps, downstream of the AAA block. It is unclear yet whether the decrease in the distance between H2A14 and H1′C10 is a consequence of the upstream-located blocks of three adenines. However, we propose a simplified mechanistic scheme which explains qualitatively the observed trends in the interstrand distances in the ACT/AGT duplex, without making a distinction between adenines and guanines. This explanation makes use of a nonplanar, buckled geometry of purine·pyrimidine (R·Y) Watson–Crick base pairs. Such a geometry is energetically optimal for oligo(R)·oligo(Y)

sequences in DNA (Zhurkin et al., 1990) and is consistent with the linear dichroism data for poly(dA)·poly(dT) (Edmondson & Johnson, 1985). Our thesis that a narrow minor groove may be associated with both A·T- and G·C-rich regions is consistent with the recent findings that DNA bending is not limited to sequences with persistent repetitions of adenines (Shlyakhtenko et al., 1990; Milton et al., 1990).

We have detected significant changes in proton chemical shifts in the range of 5–25 °C in the ACT/AGT duplex, much before the melting of the duplex. The magnitude of these changes depends on the sequence and on nucleotide positions in the strands. The same sequences located on the 3′- and 5′-sides are differentially affected by increase in temperature. This is rationalized on the basis of differential solvent exposure for the same sequences at the 3′- and 5′-sides.

REFERENCES

Alexeev, D. G., Lipanov, A. A., & Skuratovsky, I. Ya. (1987) *Nature (London) 325*, 821–823.

Arnott, S. (1972) in *Purines: Theory and Experiment, Proceedings of the 4th Jerusalem Symposium* (Bergman, E. D., & Pullman, B., Eds.) pp 102–110, Israel Academy of Science and Humanities, Jerusalem.

Balendrian, K., & Sundaralingam, M. (1991) *J. Biomol. Struct. Dyn. 9*, 511–516.

Burkhoff, A. M., & Tullius, T. D. (1987) *Cell 48*, 935–943.

Calladine, C. R. (1982) *J. Mol. Biol. 161*, 343–352.

Cheung, S., Arndt, K., & Lu, P. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 3665–3669.

Chuprina, V. P. (1987) *Nucleic Acids Res. 15*, 293–311.

Chuprina, V. P., Lipanov, A. A., Fedoroff, O. Yu., Kim, S.-G., Kintanar, A., & Reid, B. R. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9087–9091.

Coll, M., Frederick, C. A., Wang, A. H.-J., & Rich, A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8385–9389.

Crothers, D. M., Haran, T. E., & Nadeau, J. G. (1990) *J. Biol. Chem. 265*, 7093–7096.

Depew, R. E., & Wang, J. C. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 4275–4279.

Diekmann, S. (1986) *FEBS Lett. 195*, 53–56.

Donlan, M. E., & Lu, P. (1992) *Nucleic Acids Res. 20*, 525–532.

Edmondson, S. P. (1987) *Biopolymers 26*, 1941–1956.

Edmondson, S. P., & Johnson, W. C., Jr. (1985) *Biopolymers 24*, 825–841.

Grzeskowiak, K., Yanagi, K., Prive, G. G., & Dickerson, R. E. (1991) *J. Biol. Chem. 266*, 8861–8883.

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

Hagerman, J. P. (1990) *Annu. Rev. Biochem. 59*, 755–781.

Katahira, M., Sugeta, H., Kyogoku, Y., Fujii, S., Fujisawa, R., & Tomita, K. (1988) *Nucleic Acids Res. 16*, 8619–8632.

Katahira, M., Sugeta, H., & Kyogoku, Y. (1990a) *Nucleic Acids Res. 18*, 613–618.

Katahira, M., Sugeta, H., Kyogoku, Y., & Fujii, S. (1990b) *Biochemistry 29*, 7214–7222.

Kemming, J., Boelens, R., Koning, T. M. G., Kaptein, R., van der Marel, G. A., & van Boom, J. H. (1987) *Eur. J. Biochem. 162*, 37–43.

Kintanar, A., Klevit, R. E., & Reid, B. R. (1987) *Nucleic Acids Res. 15*, 5845–5862.

Koo, H.-S., Wu, H.-M., & Crothers, D. M. (1986) *Nature (London) 320*, 501–506.

Kopka, M. L., Fratini, A. V., Drew, H. R., & Dickerson, R. E. (1983) *J. Mol. Biol. 163*, 129–146.

Lefevre, J.-F., Lane, A. N., & Jardetsky, O. (1987) *Biochemistry 26*, 5076–5090.

Leupin, W., Chazin, W. J., Hyberts, S., Denny, W. A., & Wüthrich, K. (1986) *Biochemistry 25*, 5902–5910.

Lipanov, A. A., & Chuprina, V. P. (1987) *Nucleic Acids Res. 15*, 5833–5844.

Luo, J., Sarma, M. H., Yuan, R., & Sarma, R. H. (1992) *FEBS Lett.* (in press).

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

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

Metzler, W. J., Wang, C., Kitchen, D. B., Levy, R. M., & Pardi, A. (1990) *J. Mol. Biol. 214*, 711–736.

Milton, D. L., Casper, M. L., Wills, N. M., & Gesteland, R. F. (1990) *Nucleic Acids Res. 18*, 817–820.

Nadeau, J. G., & Crothers, D. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 2622–2626.

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

Prive, G. G., Yanagi, K., & Dickerson, R. E. (1991) *J. Mol. Biol. 217*, 177–179.

Schyolkina, A. K., Minchenkova, L. E., & Ivanov, V. I. (1977) *Molek. Biol. (Engl. Transl.) 11*, 466–472.

Shlyakhtenko, L. S., Lyubchenko, Yu. L., Chernov, B. K., & Zhurkin, V. B. (1990) *Molek. Biol. (Engl. Transl.) 24*, 66–81.

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

Timsit, Y., Vilbois, E., & Moras, D. (1991) *Nature (London) 354*, 167–170.

Trifonov, E. N. (1985) *CRC Crit. Rev. Biochem. 19*, 89–106.

Ulyanov, N. B., & Zhurkin, V. B. (1984) *J. Biomol. Struct. Dyn. 2*, 361–385.

Ulyanov, N. B., Gorin, A. A., Zhurkin, V. B., Chen, B.-C., Sarma, M. H., & Sarma, R. H. (1992) *Biochemistry 31*, 3918–3930.

Yanagi, K., Prive, G. G., & Dickerson, R. E. (1991) *J. Mol. Biol. 217*, 201–214.

Zhurkin, V. B., Gorin, A. A., Charakhchyan, A. A., & Ulyanov, N. B. (1990) in *Theoretical Biochemistry and Molecular Biophysics* (Beveridge, D. L., & Lavery, R., Eds.) pp 411–431, Adenine, New York.

Zhurkin, V. B., Ulyanov, N. B., Gorin, A. A., & Jernigan, R. L. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 7046–7050.