Takagi, T., & Isemura, T. (1962) *J. Biochem. (Tokyo) 52*, 314-323.
Takagi, T., & Toda, H. (1962) *J. Biochem. (Tokyo) 52*, 16-27.
Takagi, T., & Isemura, T. (1965) *J. Biochem. (Tokyo) 57*, 89-95.
Takahashi, K., & Sturtevant, J. M. (1981) *Biochemistry 20*, 6185-6190.
Toda, H., & Akabori, S. (1963) *J. Biochem. (Tokyo) 53*, 102-110.
Toda, H., Kondo, K., & Narita, K. (1982) *Proc. Jpn. Acad., Ser. B, 58B*, 208-212.
Vallee, B. L., Stein, E. A., Summerwell, W. N., & Fischer, E. H. (1959) *J. Biol. Chem. 234*, 2901-2905.
Velicelebi, G., & Sturtevant, J. M. (1979) *Biochemistry 18*, 1180-1186.

Structure, Dynamics, and Thermodynamics of Mismatched DNA Oligonucleotide Duplexes $d(CCCAGGG)_2$ and $d(CCCTGGG)_2$†

Frances H. Arnold,‡ Steven Wolk, Phillip Cruz,§ and Ignacio Tinoco, Jr.*

Department of Chemistry and Laboratory of Chemical Biodynamics, University of California, Berkeley, Berkeley, California 94720

Received October 23, 1986; Revised Manuscript Received February 20, 1987

ABSTRACT: The structures and hydrogen exchange properties of the mismatched DNA oligonucleotide duplexes $d(CCCAGGG)_2$ and $d(CCCTGGG)_2$ have been studied by high-resolution nuclear magnetic resonance. Both the adenine–adenine and thymine–thymine mismatches are intercalated in the duplexes. The structures of these self-complementary duplexes are symmetric, with the two strands in equivalent positions. The evidence indicates that these mismatches are not stably hydrogen bonded. The mismatched bases in both duplexes are in the anti conformation. The mismatched thymine nucleotide in $d(CCCTGGG)_2$ is intercalated in the duplex with very little distortion of the bases or sugar–phosphate backbone. In contrast, the bases of the adenine–adenine mismatch in $d(CCCAGGG)_2$ must tilt and push apart to reduce the overlap of the amino groups. The thermodynamic data show that the T·T mismatch is less destabilizing than the A·A mismatch when flanked by C·G base pairs in this sequence, in contrast to their approximately equal stabilities when flanked by A·T base pairs in the sequence d(CAAAXAAAG·CTTTYTTTG) where X and Y = A, C, G, and T [Aboul-ela, F., Koh, D., & Tinoco, I., Jr. (1985) *Nucleic Acids Res. 13*, 4811]. Although the mechanism cannot be determined conclusively from the limited data obtained, exchange of the imino protons with solvent in these destabilized heteroduplexes appears to occur by a cooperative mechanism in which half the helix dissociates.

When an incorrect base is incorporated during DNA synthesis, in the great majority of cases it will be recognized and excised. Failure to repair a misincorporated base will result in a mutation when the newly synthesized strand is replicated. Thus, DNA that contains non-Watson–Crick base pairs is an intermediate in mutagenesis.

The probability that a mismatched base will be recognized and repaired does not correlate with the thermodynamic stability of the mismatch (Tinoco et al., 1987). The proofreading or repair enzymes are apparently capable of recognizing some aspect of the mismatch structure or motions. Several oligonucleotides containing non-Watson–Crick base pairs and extra nucleotides have been studied by NMR in order to determine the approximate orientation of the bases and the hydrogen exchange behavior in the destabilized molecules. By nuclear Overhauser effect (NOE) experiments, it was found that the extra cytosine in $d(CA_3CA_3G{\cdot}CT_6G)$ is outside the helix (Morden et al., 1983), whereas the extra A in self-complementary d(CGCAGAATTCGCG) appears to be incorporated inside the duplex structure (Patel et al., 1982). Patel and co-workers (Patel et al., 1984a) have studied the mismatches G·T, G·A, C·T, and A·C in the duplexes $d(CGTGAATTCGCG)_2$, $d(CGAGAATTCGCG)_2$, $d(CGCGAATTCTCG)_2$, and $d(CGCGAATTCACG)_2$. In all cases, the mismatches were incorporated inside the helix. Imino hydrogens at and adjacent to the mismatch site were found to exchange with solvent by a high activation energy mechanism in which more than one base pair became accessible to the solvent at one time.

The structures of thymine–thymine and adenine–adenine mismatches have not been determined, although there has been some evidence that T·T is also incorporated in the helix. Haasnoot and co-workers (Cornelis et al., 1979) studied the self-complementary duplex $d(ATCCTATTAGGAT)_2$, in which a T·T mismatch occurs at base pair 7. They compared the chemical shift of the mismatched thymine imino proton at 10.4 ppm to that of d(TTTT) and thymidine 5′-(*O*-methylphosphate) at 11.2 ppm, from which they inferred that the mismatched T was shielded in the duplex and therefore intercalated in the helix. They also observed that the imino protons near the T·T mismatch melted at lower temperatures

†This work was supported by National Institutes of Health Grant GM 10840 and by the U.S. Department of Energy Office of Energy Research, under Contract AT03-82-ER60090.
‡Present address: Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125.
§Present address: Department of Biological Chemistry, Wright State University, Dayton, OH 45435.

0006-2960/87/0426-4068$01.50/0

than the neighboring CG regions. Similar conclusions were drawn by Summers et al. (1985) about the T·T mismatch in d(CGCGATTCGCG)$_2$.

Kennard and co-workers (Kneale et al., 1985; Kennard, 1985; Hunter et al., 1986) have determined the crystal structures of DNA duplexes containing A·G, T·G, and C·A mismatches. The A·G mismatch has the adenine in a syn conformation; this was not found for the A·G mismatches studied by NMR (Kan et al., 1983; Patel et al., 1984b). It is obviously important to study mismatches both in solution and in the crystal to establish the generality of any structural conclusions.

We have studied the A·A and T·T mismatches in the self-complementary oligonucleotides d(CCCAGGG)$_2$ and d-(CCCTGGG)$_2$. Thermodynamic parameters have been measured for these mismatch-containing molecules, as well as for the perfect duplexes d(CCCGGG)$_2$ and d-(CCCAGGG·CCCTGGG), in a continuing effort of this laboratory to characterize the thermodynamics of heteroduplexes. In a previous study (Aboul-ela et al., 1985), the thermodynamic parameters were measured for the complete set of oligonucleotides with the sequence d(CAAAX-AAAG·CTTTYTTTG), where X and Y = A, C, G, and T. In the present molecules, the mismatches are flanked by C·G base pairs. Thus, we can compare the thermodynamics of the A·A and T·T mismatches with C·G nearest neighbors to those of the same mismatches but with A·T nearest neighbors in order to investigate the sequence dependence of the stability of a mismatch.

We have used NMR to determine the gross features of the structure at the mismatch site and to investigate how the incorporation of two noncomplementary bases perturbs the motions that lead to imino hydrogen exchange. The data presented here provide a good starting point for understanding some of the features of DNA mismatches and for obtaining higher resolution structures. We are in the process of obtaining more detailed structures of these two molecules by a combination of NMR–NOE experiments and molecular mechanics calculations.

Materials and Methods

The DNA strands d(CCCAGGG), d(CCCTGGG), and d(CCCGGG) were synthesized by the solid-phase phosphoramidite method and purified by reverse-phase high-performance liquid chromatography (HPLC). Typical conditions for the purification were a 20-min gradient of 5–15% acetonitrile in 0.1 M triethylammonium acetate, flow rate 1 mL/min, on an analytical C-18 column (Axxiom). Samples for optical studies contained 0.01 M sodium phosphate, pH 7, 1 M NaCl, and 0.1 mM ethylenediaminetetraacetic acid (EDTA). Optical melting behavior was measured as described previously (Borer et al., 1974), and the data were analyzed to obtain thermodynamic parameters as described by Aboul-ela et al. (1985). NMR samples were typically 3–5 mM in single strands of DNA dissolved in 350 μL of buffer containing 0.01 M sodium phosphate, pH 7, 0.2 M NaCl, and 1 mM EDTA, with (trimethylsilyl)propionate (TSP) as a reference. The samples were lyophilized several times from 99.8% D_2O, and the final sample was prepared under nitrogen with 100.00% D_2O. NMR samples for observation of exchangeable protons were prepared in buffer with 90% H_2O/10% D_2O.

NMR spectra were obtained on a Bruker AM 500-MHz spectrometer. Spectra in H_2O were measured with a 1–3–3–1 pulse sequence for solvent suppression (Hore, 1983). Imino exchange rates were measured by the saturation recovery technique, with a 1–3–3–1 observation pulse and a saturation time of 1 s. The recovery of the magnetization was measured at 15 different delays between saturation and observation, with interleaving of the free induction decays (FID's) every 32 scans. One-dimensional NOE's in H_2O were measured with a 500-ms decoupling pulse on and off resonance with a 1–3–3–1 observe pulse sequence and interleaving every 32 scans. Two-dimensional phase-sensitive NOESY spectra were measured with the TPPI pulse sequence (Bodenhausen et al., 1984) and 1K data points in the t_2 dimension and typically 350 data points zero filled to 1K in t_1. A delay of 2 s was used between scans, and 64 scans were taken at each t_1 for a total acquisition time of approximately 20 h. A mixing time of 500 ms was chosen as the best compromise between maximizing the signal-to-noise and minimizing the appearance of new cross-peaks resulting from spin–diffusion. The mixing time of 500 ms was necessary in order to observe the NOE's, since in a 7-mer the NOE buildup rate is quite small relative to oligonucleotides 12–20 base pairs long.

Table I: Thermodynamic Parameters of Duplex Formation for Mismatches and Perfect Duplexes

duplex	ΔH° (kcal/mol)[a]	ΔS° (eu)[b]	ΔG°_{25} (kcal/mol)[c]	T_m (°C)[d]
CCCTGGG GGGTCCC	−60	−179	−6.7	31
CCCAGGG GGGACCC	−60	−180	−5.9	27
CCCAGGG GGGTCCC	−63	−177	−10.0	48
CCCGGG GGGCCC	−62	−176	−9.0	43

[a] Estimated precision in ΔH° is ±5 kcal/mol. [b] Estimated precision in ΔS° is ±20 eu [cal/(deg·mol)]. [c] Estimated precision in ΔG°_{25} is ±0.3 kcal/mol. [d] Estimated precision in T_m is ±2 °C; 400 μM strand concentration for self-complementary molecules; 100 μM strand concentration for non-self-complementary molecules.

Table II: Effects of Neighboring Sequence on the Free Energy Cost ($\Delta\Delta G^\circ_{25}$) of Replacing a Watson–Crick A·T Base Pair with Mismatched Base Pairs A·A and T·T[a]

	parent sequence		
mismatch	C-A-G G-T-C	T-T-T A-A-A[b]	A-T-A T-A-T[b]
A·A	+4.1	+4.6	+3.5
T·T	+3.3	+4.6	+3.5

[a] Mismatches replace the central A·T base pair in the parent sequence. [b] Values obtained from Aboul-ela et al. (1985).

Results

Optical Melting Behavior. The thermodynamic parameters of duplex formation for the four molecules studied are listed in Table I. The ΔH° values for the single-strand to double-strand transition in all the oligonucleotides studied are nearly equal, within the limits of experimental error. This indicates that the base stacking in the mismatched duplexes is not greatly perturbed relative to that of the perfect helix. The free energy cost of replacing the A·T base pair at position 4 in the perfect duplex with a T·T mismatch is 3.3 kcal/mol at 25 °C. Replacing the A·T pair with an A·A mismatch has an even greater destabilizing effect: $\Delta\Delta G^\circ_{25}$ = 4.1 kcal/mol.

That mismatch stability depends on the neighboring base pair sequence is clear when these results are compared to those of Aboul-ela et al. (1985) for mismatches surrounded by A·T base pairs. Table II shows the free energy cost of replacing an A·T base pair in a perfect duplex with A·A and T·T mis-

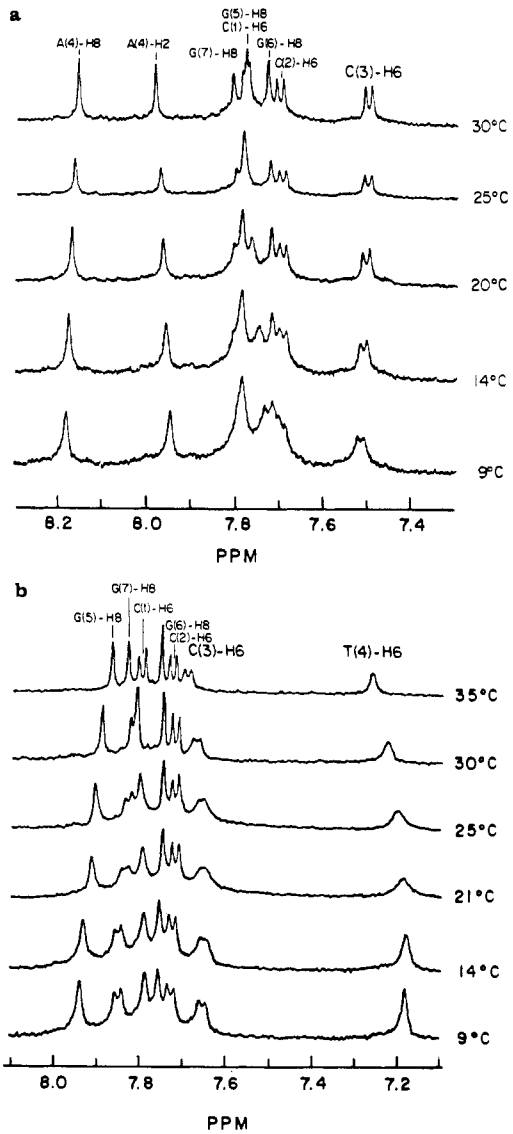


FIGURE 1: (a) Aromatic region of 500-MHz proton spectrum of $d(CCCAGGG)_2$. Conditions: 3 mM single strands in D_2O, 0.01 M sodium phosphate, 0.2 M NaCl, and 1 mM EDTA, pH 7. (b) Aromatic region of 500-MHz proton spectrum of $d(CCCTGGG)_2$. Conditions: 3 mM single strands in D_2O, 0.01 M sodium phosphate, 0.2 M NaCl, and 1 mM EDTA, pH 7.

matches flanked by A·T and C·G base pairs. The T·T mismatch is most destabilizing when it replaces the central A·T base pair in the sequence $\frac{\text{A-T-A}}{\text{T-A-T}}$, while it is the least destabilizing when it replaces the A·T in $\frac{\text{C-A-G}}{\text{G-T-C}}$ to give $\frac{\text{C-T-G}}{\text{G-T-C}}$. In contrast, the A·A mismatch is most destabilizing when it replaces the central A·T base pair in $\frac{\text{A-A-A}}{\text{T-T-T}}$ and least destabilizing when placed in $\frac{\text{T-A-T}}{\text{A-T-A}}$. The A·A and T·T mismatches destabilize a perfect duplex equally when the neighboring base pairs are A·T and T·A. Between C·G base pairs, however, the A·A mismatch causes a greater change in the free energy than does T·T.

NMR. The resonances were assigned by 2-D NOE techniques, as described further on. One-dimensional proton NMR spectra of the aromatic protons of $d(CCCAGGG)_2$ and $d(CCCTGGG)_2$ are shown at various temperatures in Figure 1. Single peaks for the mismatched bases are seen at all temperatures. Thus, it is clear that $d(CCCAGGG)_2$ duplex formation results in a structure that has retained its symmetry; the two strands are in equivalent positions in the duplex. The aromatic resonances in $d(CCCAGGG)_2$ all broaden simultaneously as the temperature is lowered, a result of the increased viscosity of the solvent. The behavior of the aromatic resonances in $d(CCCTGGG)_2$ is somewhat different (Figure 1b). At approximately 20 °C, well below the melting temperature in the NMR sample, $T_m \sim 40$ °C, the resonances of T(4)-H6 and C(3)-H6 become significantly broader than the other aromatic resonances. As the temperature decreases further to 10 °C, the C-H6 and T-H6 once again become narrow. At even lower temperature all the peaks broaden simultaneously (not shown).

In Figure 2 the chemical shifts of the aromatic resonances of the two molecules are plotted as a function of temperature. It was possible to follow each of the aromatic resonances through the melting transition from duplex to single-stranded forms. The H6 of the mismatched thymine in $d(CCCTGGG)_2$ and the H2 of the adenine in $d(CCCAGGG)_2$ both experience large (0.2 ppm) upfield shifts upon duplex formation (panels a and b of Figure 2). The H8 of the adenine, on the other hand, shows practically no change in chemical shift between 5 and 70 °C. The chemical shifts for the three cytosine H6's and the three guanine H8's of the two molecules are shown in panels c and d of Figure 2. The behavior of these resonances is qualitatively very similar in both molecules. The guanine H8 nearest the mismatch site (G5-H8) undergoes a downfield shift as the temperature is lowered. In contrast, the G-H8's on the terminal two base pairs move to higher field at low temperatures. The cytosine H6's of the base pairs nearest the mismatch site, base pairs 2 and 3, both move upfield with lower temperature, while the H6 of the terminal cytosine shows a downfield change in its chemical shift.

NMR Spectra in H_2O. The far-downfield portions of the spectra in H_2O showing the imino resonances in the two oligomers are shown in Figure 3. These resonances were assigned by 1-D NOE experiments. Saturation of the imino peak at 12.85 ppm in $d(CCCAGGG)_2$ results in an NOE at the H2 of adenine at 7.9 ppm and to the imino resonance at 13.1 ppm, as shown in Figure 3a. Thus, the resonance at 12.85 ppm belongs to the imino proton of base pair 3, and the peak at 13.1 ppm is assigned to the imino proton of base pair 2. Furthermore, the A-H2 proton must be less than 4 Å away from the imino on the neighboring base pair, which places the mismatched adenines inside the duplex structure.

The imino proton of the thymine base in $d(CCCTGGG)_2$ at 10.4 ppm appears far upfield from the rest of the hydrogen-bonded iminos (Figure 3B). This is identical with the chemical shift observed for the T·T mismatches in d-(ATCCTATTAGGAT)$_2$ (Cornelis et al., 1979) and d-$(CGCGATTCGCG)_2$ (Summers et al., 1985). Saturation of the resonance at 10.4 ppm results in an NOE at 12.75 ppm, which is assigned to the guanine of base pair 3. Thus, the thymine imino is less than 4 Å from the neighboring base pair imino, which confirms that the mismatched thymines are intercalated in the duplex. The remaining two imino protons at the ends of both duplexes can be assigned by observing the melting behavior (not shown); as the temperature is raised, the terminal base pairs melt and disappear from the NMR spectrum first.

2-D NMR. Two-dimensional NOESY spectra were taken of both oligomers to aid with the assignment of the resonances and to obtain structural information about the mismatches. The aromatic to H1′ region of the NOESY spectrum of d-$(CCCAGGG)_2$ is shown in Figure 4. The connectivity pattern of the NOE's in B-form DNA [the H6 or H8 of the base shows NOE's to the H1′ of its own sugar and to the H1′ of the *n* –

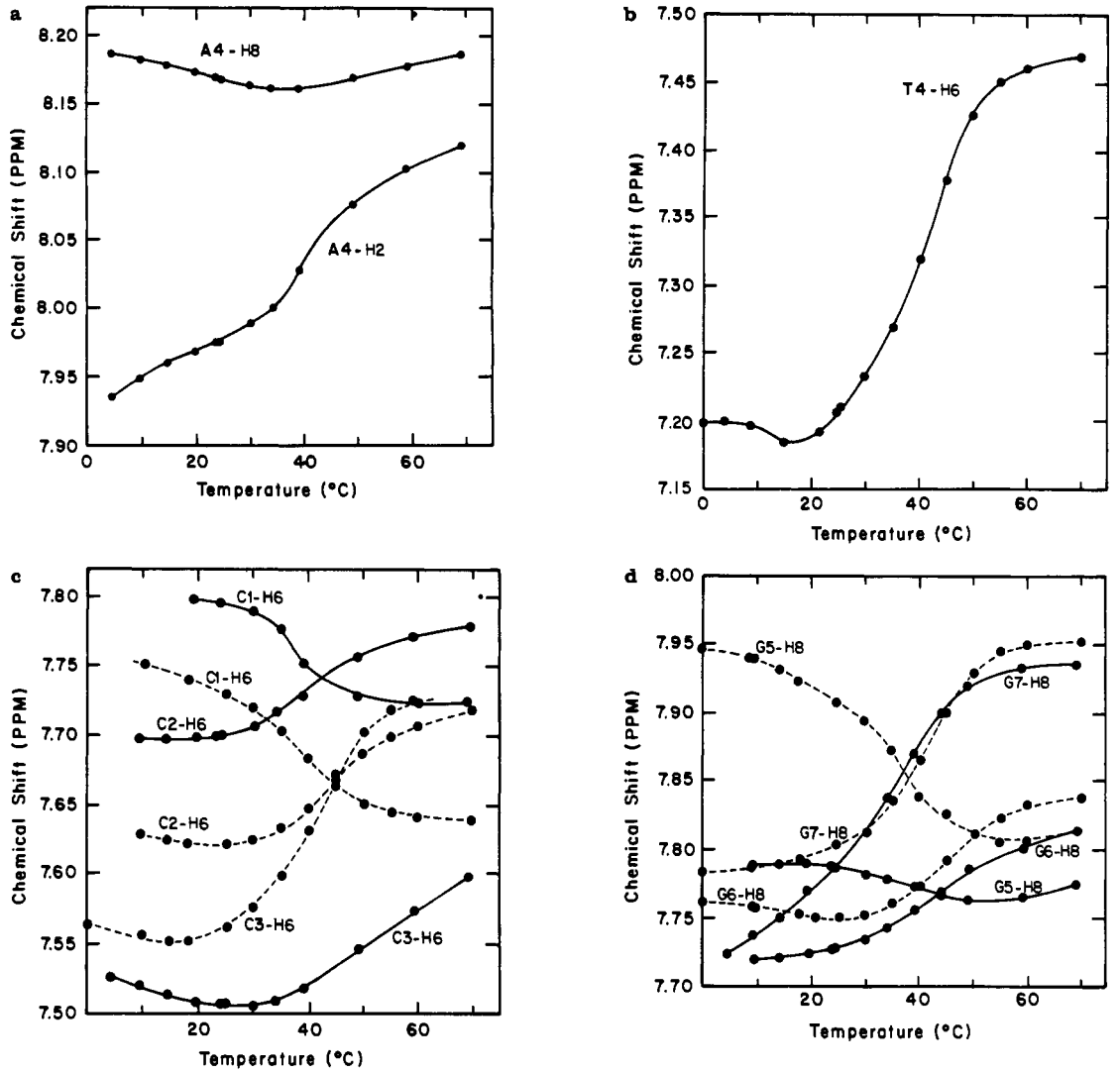


FIGURE 2: (a) Chemical shifts of A-H2 and A-H8 vs. temperature in d(CCCAGGG). Conditions: 3 mM single strands in D_2O, 0.01 M sodium phosphate, 0.2 M NaCl, and 1 mM EDTA, pH 7. (b) Chemical shift of T-H6 vs. temperature in d(CCCTGGG). Conditions: 3 mM single strands in D_2O, 0.01 M sodium phosphate, 0.2 M NaCl, and 1 mM EDTA, pH 7. (c) Chemical shifts of C-H6 in d(CCCAGGG) (—) and d(CCCTGGG) (--). Conditions: 3 mM single strands in D_2O, 0.01 M sodium phosphate, 0.2 M NaCl, and 1 mM EDTA, pH 7. (d) Chemical shifts of G-H8 in d(CCCAGGG) (—) and d(CCCTGGG) (--). Conditions: 3 mM single strands in D_2O, 0.01 M sodium phosphate, 0.2 M NaCl, and 1 mM EDTA, pH 7.

1 sugar (Hare et al., 1983)] is maintained throughout for both molecules (not shown here for CCCTGGG). Introduction of a base–base mismatch has not perturbed this connectivity of intra- and internucleotide NOE's. The assignments of the aromatic, H1′, H2′, and H2″ protons, obtained from the NOESY spectra by standard techniques (Hare et al., 1983), are listed in Table III.

Approximate sugar conformations can be determined by inspecting the splitting patterns of H1′ resonances and the cross sections of NOESY spectra showing NOE's from aromatic to sugar protons (J. F. Lefevre, A. Lane, and O. Jardetzky, unpublished results). Two H1′ resonances are well resolved in $d(CCCAGGG)_2$, G(7)-H1′ and G(5)-H1′. Both these resonances are triplets with J = 7–8 Hz, as expected for a conformation that is an average of predominantly C2′ endo, with some contribution from C3′ endo (J. F. Lefevre, A. Lane, and O. Jardetzky, unpublished results). For each of the nucleotides in $d(CCCAGGG)_2$ and $d(CCCTGGG)_2$, the NOE from an aromatic proton to its own sugar 2′ is larger than the NOE's to its sugar 3′ and to the 2″ of the $n-1$ sugar. This is the pattern expected for sugars with C2′-endo conformations—much larger aromatic-to-3′ NOE's are expected for C3′-endo conformations.

Table III: Resonance Assignments for $d(CCCTGGG)_2$ (20 °C) and $d(CCCAGGG)_2$ (20 °C)

			1 2 3 4 5 6 7 C C C T G G G					
base	C-H6	G-H8	T-H6	C-H5	H1′	H2″	H2′	T-CH_3
1	7.84			5.99	6.04	2.58	2.24	
2	7.73			5.76	6.04	2.52	2.30	
3	7.66			5.74	6.15	2.56	2.22	
4			7.19		5.78	2.08	1.49	1.74
5		7.92			5.47	2.74	2.74	
6		7.76			5.81	2.73	2.61	
7		7.80			6.19	2.52	2.40	

			1 2 3 4 5 6 7 C C C A G G G					
base	C-H6	G-H8	A-H8	A-H2	C-H5	H1′	H2″	H2′
1	7.81				5.92	5.97	2.56	2.22
2	7.71				5.74	5.99	2.47	2.23
3	7.53				5.81	5.84	2.16	1.93
4			8.18	7.96		5.69	2.64	2.64
5		7.81				5.51	2.64	2.64
6		7.74				5.77	2.72	2.60
7		7.77				6.15	2.53	2.39

When the sugar conformation is C2′ endo, NOE's can be observed from a given aromatic proton to the H2′ and H2″

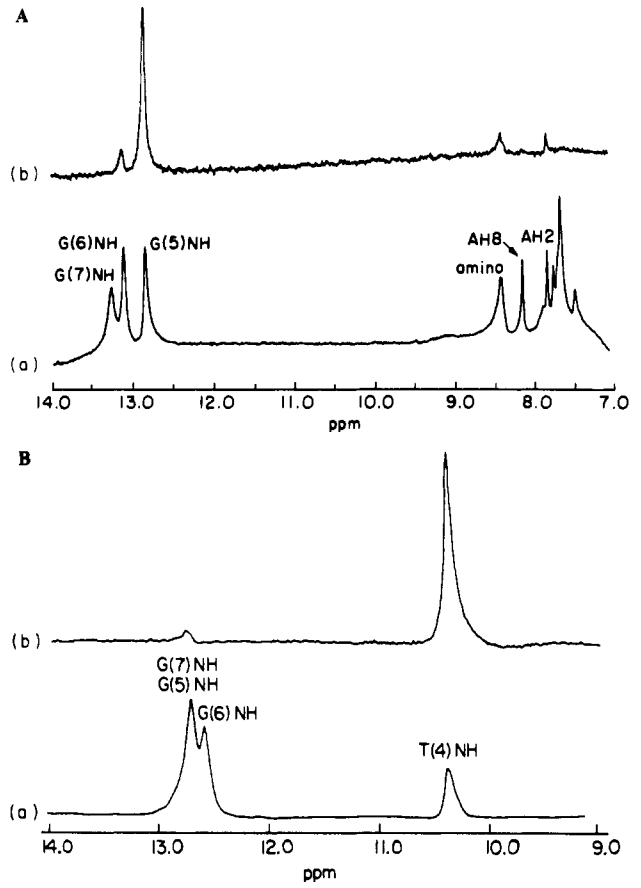


FIGURE 3: (A) Control (a) and NOE difference spectra (b) for $d(CCCAGGG)_2$: imino and aromatic region. Saturation of the CG imino resonance at 12.85 ppm results in an NOE to A-H2 at 7.9 ppm and G(5)-NH1 at 13.1 ppm. Saturation pulse was 0.5 s, with 1-3-3-1 observe sequence for water signal suppression. (B) Control (a) and NOE difference spectra (b) for imino region of $d(CCCTGGG)_2$ (1.9 °C). Saturation of the thymine imino at 10.4 ppm results in an NOE to G(5)-NH1 at 12.75 ppm. Saturation pulse was 0.5 s, with 1-3-3-1 observe sequence for water signal suppression.

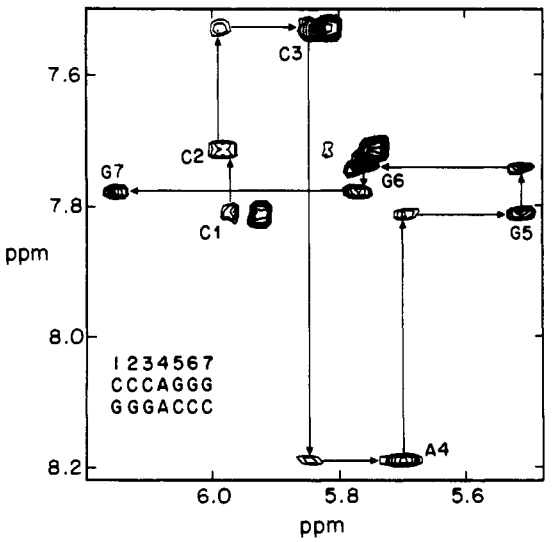


FIGURE 4: Aromatic-H1′ region of 500-MHz 2-D NOESY spectrum for $d(CCCAGGG)_2$ (20 °C). Conditions: 1K data points in t_2, 350 points zero-filled to 1K in t_1; 64 scans for each t_1, with 2-s delay between scans; mixing time 500 ms.

of its own sugar and to the H2″ of the $n-1$ sugar. A very different pattern is expected for C3′ endo (Haasnoot et al., 1984). The NOE's characteristic of C2′ endo are in fact observed throughout both mismatched molecules. It is interesting to note, however, that at the mismatched adenine of $d(CCCAGGG)_2$, the NOE from A(4)-H8 to the H2″ of C(3) is much weaker than the NOE's from all the other aromatics to their $n-1$ H2″ protons (Figure 5). Apparently the mismatched adenine has taken on an orientation in the duplex that pulls A-H8 away from the C(3) sugar.

Table IV: Imino Proton Lifetimes: Observed Lifetimes and Solvent Exchange Lifetimes Corrected for Magnetic Dipolar Relaxation[a]

$d(CCCAGGG)_2$					
base pair 2			base pair 3		
T (°C)	τ_{obsd} (ms)	τ_{ex} (ms)	T (°C)	τ_{obsd} (ms)	τ_{ex} (ms)
−1.1	86		−1.1	82	
3.7	61	94	3.7	60	92
6.4	41	52	7.5	35	42
8.3	29	33	9.6	23	26
10.3	19	21	10.4	20	22
11.5	16	17	11.6	17	18
12.4	13	14	12.4	14	15
E_a = 35 ± 2 kcal/mol			E_a = 33 ± 2 kcal/mol		
$d(CCCTGGG)_2$					
base pair 3			thymine		
T (°C)	τ_{obsd} (ms)	τ_{ex} (ms)	T (°C)	τ_{obsd} (ms)	τ_{ex} (ms)
−2.1	142				
−0.6	156		−3.1	183	
0.8	153		0.7	128	
6.5	95	240	3.3	97	184
8.1	77	150	8.1	48	61
10.4	54	82	13.1	19	21
13.4	30	37	14.6	22	24
15.7	20	23	15.8	15	15
16.0	20	23	16.0	16	17
16.6	17	19	16.6	13	14
17.3	15	16	17.3	9	9
E_a = 37 ± 2 kcal/mol			E_a = 31 ± 2 kcal/mol		
$d(CCCGGG)_2$					
base pair 2			base pair 3		
T (°C)	τ_{obsd} (ms)	τ_{ex} (ms)	T (°C)	τ_{obsd} (ms)	τ_{ex} (ms)
−0.3	310		−0.2	366	
2.5	335		3.8	411	
5.6	301		7.5	329	
9.3	272		11.6	334	
12.2	224	544	14.6	311	
15.7	129	195	17.7	219	504
18.6	99	133	21.7	97	129
21.7	53	61	24.7	51	59
24.7	41	45			
E_a = 29 ± 3 kcal/mol					

[a] Lifetimes were measured in 0.01 M sodium phosphate, pH 7, buffer–0.2 M NaCl.

Measurement of Hydrogen Exchange Lifetimes. The lifetimes of the imino protons of $d(CCCAGGG)_2$, $d(CCCTGGG)_2$, and $d(CCCGGG)_2$ were measured by the saturation recovery technique, in which a resonance is saturated and the recovery of its magnetization is observed. The recoveries of the imino protons of the two internal C·G base pairs in $d(CCCAGGG)_2$, the thymine and C·G base pair 3 in $d(CCCTGGG)_2$, and the C·G base pairs 2 and 3 of $d(CCCGGG)_2$ were studied. The lifetimes observed at various temperatures are listed in Table IV. At low temperature, where exchange with solvent is very slow, the lifetime is determined by the magnetic dipolar relaxation time T_1(magnetic), whose small temperature dependence is mainly a result of changes in the solvent viscosity. At higher temperatures (still well below T_m), exchange with the solvent becomes important and significantly decreases the imino-hydrogen NMR

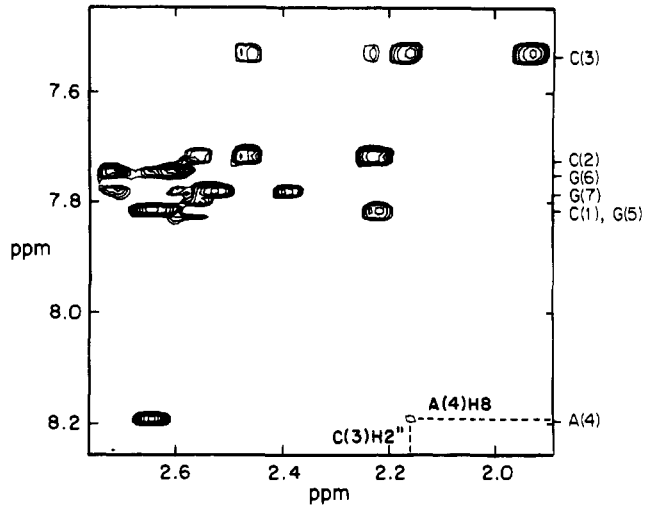


FIGURE 5: Aromatic-H2',H2" region of 500-MHz 2-D NOESY spectrum for d(CCCAGGG)$_2$ (20 °C). NOE between A(4)-H8 and C(3)-H2" is unusually weak. Conditions: 1K data points in t_2, 350 points zero-filled to 1K in t_1; 64 scans for each t_1, with 2-s delay between scans; mixing time 500 ms.

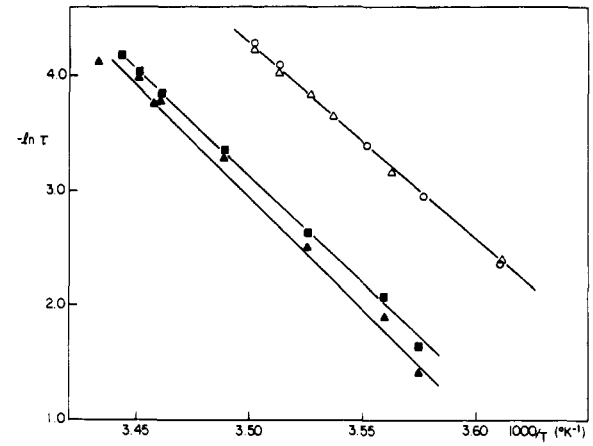


FIGURE 6: Arrhenius plot for imino hydrogen exchange in d(CCCAGGG)$_2$ and d(CCCTGGG)$_2$. (△) CG base pair 2; (○) CG base pair 3; (■) thymine imino; (▲) CG base pair 3. Hydrogen exchange rates measured by saturation recovery with 0.5-s saturation pulse and 1-3-3-1 observe sequence for water signal suppression. Each point represents 15 measurements at different delays between saturation and observation, with typically 300 scans at each delay.

relaxation lifetime. When exchange with solvent dominates, the recovery is described by a single exponential, and the activation energy is indicative of the mechanism by which the proton becomes accessible to solvent. The imino protons of d(CCCTGGG)$_2$ have longer lifetimes than the iminos in d(CCCAGGG)$_2$, whereas the lifetimes of the iminos in both the mismatched molecules are significantly shorter than those in the perfect parent duplex d(CCCGGG)$_2$. The magnetic dipolar effect is evident at low temperature for d(CCCTGGG)$_2$ and d(CCCGGG)$_2$. In contrast, for d(CCCAGGG)$_2$, exchange with solvent appears to dominate at all temperatures studied. The lifetimes of the imino protons of C·G base pair 3 and the mismatched thymine in d(CCCTGGG)$_2$ differ significantly from one another. The two internal C·G base pairs studied in d(CCCAGGG)$_2$ have nearly identical values; the mismatched adenine, of course, has no imino proton to study.

The solvent exchange lifetime can be obtained by correcting the measured relaxation lifetimes for the dipolar contribution:

$$1/\tau(\text{observed}) = 1/T_1(\text{magnetic}) + 1/\tau(\text{exchange})$$

Over the small temperature range studied (~20 °C), the magnetic dipolar contribution should change very little. This is not true, however, if significant aggregation is taking place at low temperatures, as appears to be the case for these oligonucleotides at the high concentrations in the NMR samples.

The corrected solvent exchange lifetimes for d(CCCAGGG)$_2$ and d(CCCTGGG)$_2$ were obtained by assuming a dipolar relaxation contribution of 0.33 s at 20 °C, which is that expected for a seven base pair DNA fragment (Early et al., 1981). It was assumed that the magnetic dipolar relaxation rate increased linearly as the temperature decreased, until it reached the rate measured at 0 °C for d(CCCTGGG)$_2$. The relaxation lifetimes of the imino protons at 0 °C were 0.16 s for the C·G imino of base pair 3 and 0.19 s for the thymine imino. Similarly, the magnetic dipolar contribution for d(CCCGGG)$_2$ imino relaxation was assumed to be 0.39 s at 20 °C, decreasing to 0.37 s at 0 °C. Small errors in the magnetic dipolar correction will not greatly affect the calculated exchange lifetimes at higher temperature, nor will it significantly affect the activation energy. The effect is the greatest for the C·G base pair in d(CCCGGG)$_2$, which has the smallest relaxation rate. The corrected exchange rates are shown in Table IV.

Activation energies for exchange with solvent can be obtained from the Arrhenius plot of ln τ(exchange) vs. $1/T$ (Figure 6). The exchange rates and the activation energies ($E_a \sim 34$ kcal/mol) measured for two C·G base pairs in d(CCCAGGG)$_2$ are practically identical. In d(CCCTGGG)$_2$ the exchange lifetimes of the imino proton of C·G base pair 3 are greater than the lifetimes of the imino proton of the thymine mismatch. The activation energies for the imino protons are similar: $E_a = 37$ kcal/mol for the imino of base pair 3, and $E_a = 32$ kcal/mol for the mismatch thymine imino.

Discussion

Most of the aromatic proton resonances in the two mismatched duplexes undergo an upfield shift relative to the single strands (Figure 2). This is the expected effect of the ring current shielding caused by increased stacking of the bases in the double helix. However, the G(5)-H8 resonance shows a downfield shift upon duplex formation in d(CCCTGGG)$_2$ and is essentially unchanged in d(CCCAGGG)$_2$. Thus G(5) is apparently less stacked in the duplexes than in the single strands. Its base-paired partner C(3) shows a normal upfield shift for its H6 in both duplexes, with a large shift in d(CCCTGGG)$_2$. It appears that the base pair on either side of the mismatch is rotated so as to increase the shielding of the cytosine and to decrease the shielding of the guanine in the duplex form.

The large adenine–adenine mismatch cannot be incorporated in a B-form duplex without some adjustment of the bases and the sugar–phosphate backbone. The very small change in the chemical shift of A(4)-H8 in d(CCCAGGG)$_2$ implies that this proton is not well shielded in the duplex form. The A-H2 proton, on the other hand, undergoes a large upfield change in chemical shift upon duplex formation. The position of the adenine bases in the mismatched duplex should explain this shielding of the A-H2 as well as the lack of shielding at A-H8 and the strong NOE between the imino proton of the neighboring 3' guanine and A-H2.

The adenine–adenine mismatch can be incorporated in a B-type helix if the bases can pull apart and propeller twist. A slight rotation about the glycosidic bond to increase the twist, accompanied almost certainly by distortion of the sugar–phosphate backbone, would allow a symmetric A·A base pair to be inserted in the duplex without overlap of the adenine amino groups. By inspecting a model of B-form d-

$(CCCAGGG)_2$, one can see that decreasing the glycosidic torsion angle slightly brings A-H2 closer to the G imino proton (as compared to the A-H2 position in standard B form with an A·T base pair). As a result of this rotation, the distances between A(4)-H8 and the H2′ and H2″ protons of its own sugar decrease, while the separation from the H2′ and H2″ protons on the C(3) nucleotide increases. This is consistent with the results of the NOESY experiment, in which only a very weak NOE is observed between A(4)-H8 and C(3)-H2″. Much stronger NOE's are seen from all the other aromatic protons to the H2″ protons on the $n-1$ nucleotides. The basic sugar conformation is 2′ endo throughout the duplex, including the mismatch site, as judged by the NOE patterns from the aromatic to the sugar protons and the H1′ splitting patterns. It is possible that the large adenine–adenine base pair is accomodated inside the helix by a combination of the increased propeller twist and by pushing the backbone out to make more room for the bases. This would cause A-H8 to move farther out from the helix axis, reducing the shielding arising from the neighboring bases, consistent with the chemical shift behavior of the A-H8 resonance.

The evidence presented here indicates that there is no long-lived adenine–adenine mismatch structure with hydrogen bonding, which would almost certainly break the symmetry of the base pair. If a hydrogen-bonded structure exists, then it must be exchanging quickly on the NMR time scale.

In $d(CCCTGGG)_2$, the NOE observed from the thymine imino proton to the neighboring C·G base pair imino (Figure 3B) shows that the mismatch is intercalated in the helix. The upfield change in chemical shift of the T-H6 aromatic proton (Figure 2b) upon duplex formation and the relatively upfield chemical shift (10.4 ppm) of the thymine imino resonance in the duplex also support this conclusion. If the thymine mismatch position is close to that expected for thymine in an A·T base pair in B-DNA, the distance between the closest proton of T-CH3 and the H6 of cytosine at position 3 is only 2.7 Å. A reasonably large cross-peak between T(4)-CH3 and C(3)-H6 is in fact seen in the 2-D NOESY experiment (not shown).

The short T·T base pair can be incorporated in B-form DNA without rotation of the bases, although hydrogen bonding between the bases is not possible in this orientation. A possible hydrogen-bonded T·T base pair has been drawn (Aboul-ela et al., 1985) in which both T's are in the anti conformation and there are two hydrogen bonds between the imino NH3 and O4 and between NH3 and O2. It is possible that the broadening of the aromatic resonances observed between 10 and 20 °C is a result of exchange between two or more base conformations. In this interpretation the alternate forms disappear below 10 °C, where the resonances of C(3)-H6 and T(4)-H6 once again become narrow.

It is worthwhile pointing out that although the NMR results show that the sugar conformation is predominantly C2′ endo, the data do not allow us to determine the detailed orientation of the bases. Crystals of $d(GGGGCCCC)_2$ and other molecules with GpG sequences have A-form structures (McCall et al., 1985, 1986). It is possible that the orientation of the bases in $d(CCCAGGG)_2$ and $d(CCCTGGG)_2$ is more closely related to A-form DNA than B-form DNA and that the overall structures contain characteristics of each.

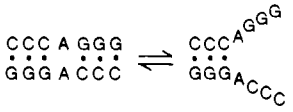
Hydrogen Exchange. Exchange of the imino proton for a solvent proton occurs from some "open" state in which the proton is presumably not base pair hydrogen-bonded and is accessible to the solvent. Imino protons in a stable duplex have long lifetimes relative to single strands. It is believed that these lifetimes reflect the rate of opening of a single base pair. For exchange that is limited by the opening of the base pairs, the exchange rates should exhibit no dependence on the catalyst concentration. The activation energies for solvent exchange by single base-pair opening are approximately 16 kcal/mol for A·T base pairs in poly(dA-dT) (Assa-munt et al., 1984) and 29 kcal/mol for C·G base pairs in poly(dG)·poly(dC) (Mirau & Kearns, 1984).

Hydrogen exchange is somewhat complicated in smaller oligonucleotides and in duplexes that have been destabilized by the incorporation of base–base mismatches. In such cases end fraying and dissociation of the helix into single strands become important contributors to exchange of the imino protons with solvent. If the rate of dissociation of the duplex is fast compared to the rate of single base pair opening, then hydrogen exchange rates will reflect the rate of duplex dissociation instead of single base pair opening. In this case, the exchange process is cooperative, and all neighboring imino protons in the dissociated unit should have identical lifetimes. Furthermore, the lifetimes would correlate with the kinetics of helix dissociation. No such correlation is expected if single base pair opening is the dominant mechanism.

The exchange lifetimes in $d(CCCGGG)_2$ are typical of an exchange mechanism dominated by kinetic fraying. The imino hydrogens of the central base pairs are the longest lived, and the lifetimes decrease nearer the ends. The activation energy measured for exchange at base pair 2 (E_a = 29 kcal/mol) is similar to that found for exchange in poly(dG)·poly(dC).

Due to the low stability of the duplexes studied, imino proton lifetimes were not measured at different catalyst concentrations. Without these additional measurements, it cannot be said conclusively that the imino exchange is open-limited. As a result, interpretation of the lifetimes and activation energies in terms of helix opening is conjecture. The observed behavior, however, shows some interesting features, and one possible interpretation of the results will be discussed. Additional data would be required to confirm these hypotheses.

$d(CCCAGGG)_2$, the least stable of the oligomers studied, has the shortest imino lifetimes, and the lifetimes of base pairs 2 and 3 are identical at all temperatures studied. (The lifetime of base pair 1 was not measured since this imino is at a frayed end and melts at very low temperature.) It is likely that the mechanism for exchange involves a cooperative dissociation of some or all of the duplex. The measured activation energy, $E_a \sim 34$ kcal/mol, is significantly smaller, however, than the enthalpy of the single-strand–duplex transition given in Table I ($\Delta H° \sim 60$ kcal/mol). Since this oligonucleotide contains a mismatch in the center, with weak or no hydrogen bonding to the cross strand at that position, a reasonable interpretation of the data is that exchange occurs from an open state in which only half the molecule is dissociating. This would require a correspondingly smaller activation energy.



The exchange lifetimes of the imino protons in $d(CCCTGGG)_2$ are longer than those observed in $d(CCCAGGG)_2$, and the activation energies for exchange are slightly larger (Figure 6 and Table IV). Unlike $d(CCCAGGG)_2$, however, the two iminos observed, in base pair 3 and the mismatch T, have different exchange rates. If the mechanism were half-helix opening, the lifetime of the T would be half that of the C·G base pair, since every time the helix opened the T imino would exchange, while only half of the

neighboring C·G's would exchange. This is close to what is observed in $d(CCCTGGG)_2$, where the T imino exchange lifetimes are significantly smaller than the C·G imino lifetimes at all temperatures (Table IV). The activation energies for exchange of the two imino protons in $d(CCCTGGG)_2$ differ slightly. Exchange from the thymine imino may also occur by an independent, low activation energy process such as single base pair opening. The latter hypothesis is difficult to quantitate, since exchange by single base pair opening of a mismatched base pair has not yet been reported.

Conclusion

From the data provided by the optical absorption and NMR experiments described here, we have been able to determine thermodynamics, structural features, and the solvent exchange properties of the adenine–adenine and thymine–thymine mismatches in the self-complementary DNA duplexes $d(CCCAGGG)_2$ and $d(CCCTGGG)_2$. Both oligonucleotides are significantly destabilized compared to the fully Watson–Crick base-paired duplex. The destabilization that results from a particular mismatch pair appears to depend on the neighboring sequence.

Although the mismatches destabilize the duplex thermodynamically, this is not accompanied by gross perturbations of the duplex geometry. NOE connectivity patterns between base protons and the sugar H1′ on its own nucleotide (n) and the preceding ($n - 1$) nucleotide are maintained through the mismatch site. Additional NOE's measured between mismatch base protons and protons on neighboring bases show that both mismatches are intercalated in the duplexes. All the sugar conformations are predominantly 2′-endo, including those at the mismatch sites. The T·T mismatch within a certain temperature range appears to be exchanging between two or more conformations, neither of which leads to a large perturbation in the duplex structure at the mismatch site. Model building indicates that the mismatched adenine appears to reduce its overlap with the cross-strand adenine amino group by increased propeller twist, which is accomplished by a decrease in the glycosidic torsion angle. The low-resolution structures that have been obtained here are being used as the starting point for determining the detailed structures of the two molecules by a combined molecular mechanics–NMR approach.

Although the data are incomplete, a reasonable interpretation of the imino hydrogen exchange behavior in the mismatched duplexes is that exchange is taking place by a cooperative mechanism in which half the helix is dissociating to allow simultaneous exchange of the exposed protons. The two guanine iminos observed in $d(CCCAGGG)_2$ have identical lifetimes and activation energies roughly equal to half the enthalpy of helix dissociation. In $d(CCCTGGG)_2$ the lifetime of the thymine imino is smaller than that of the neighboring C·G base pair, although they have very similar activation energies, consistent with half-helix opening. Exchange lifetimes correlate with thermodynamic stability of the duplex, as is expected for a helix opening mechanism. The exchange rates in the perfect parent duplex $d(CCCGGG)_2$ are much slower than in the mismatched duplexes and do not show this cooperative behavior.

Registry No. $d(CCCAGGG)_2$, 108347-07-5; $d(CCCTGGG)_2$, 108347-08-6; d(CCCAGGG)·d(GGGTCCC), 108347-10-0; $d(CCCGGG)_2$, 81994-17-4; adenine, 73-24-5; thymine, 65-71-4.

References

Aboul-ela, F., Koh, D., & Tinoco, I., Jr. (1985) *Nucleic Acids Res. 13*, 4811.

Assa-Munt, N., Granot, J., Behling, R. W., & Kearns, D. R. (1984) *Biochemistry 23*, 944–955.

Bodenhausen, G., Kogler, H., & Ernst, R. R. (1984) *J. Magn. Reson. 58*, 370.

Borer, P. N., Dengler, B., Tinoco, I., Jr., & Uhlenbeck, O. C. (1974) *J. Mol. Biol. 86*, 843–853.

Cornelis, A. G., Haasnoot, J. H. J., den Hartog, J. F., de Rooij, M., van Boom, J. H., & Cornelis, A. (1979) *Nature (London) 281*, 235.

Early, T. A., Kearns, D. R., Hillen, W., & Wells, R. D. (1981) *Biochemistry 20*, 3756–3764.

Haasnoot, C. A. G., Westerink, H. P., van der Marel, G. A., & van Boom, J. H. (1984) *J. Biomol. Struct. Dyn. 2*, 345.

Hare, D. R., Wemmer, D. E., Chou, S.-H., Drobny, G., & Reid, B. R. (1983) *J. Mol. Biol. 171*, 319.

Hore, P. J. (1983) *J. Magn. Reson. 54*, 539–542.

Hunter, W. N., Brown, T., Anand, N. N., & Kennard, O. (1986) *Nature (London) 320*, 552.

Kan, L., Chandrasegaran, O., Pulford, S. M., & Miller, P. S. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4263.

Kennard, O. (1985) *J. Biol. Struct. Dyn. 3*, 205.

Kneale, G., Brown, T., & Kennard, O. (1985) *J. Mol. Biol. 186*, 805.

McCall, N. J., Brown, T., & Kennard, O. (1985) *J. Mol. Biol. 183*, 385.

McCall, M., Brown, T., Hunter, W. N., & Kennard, O. (1986) *Nature (London) 322*, 661.

Mirau, P. A., & Kearns, D. R. (1984) *J. Mol. Biol. 177*, 207.

Morden, K. M., Chu, Y. G., Martin, F. H., & Tinoco, I., Jr. (1983) *Biochemistry 22*, 5557.

Patel, D. J., Kozlowski, S. A., Marky, L. A., Rice, J. A., Broka, C., Dallas, J., Itakura, K., & Breslauer, K. J. (1982) *Biochemistry 21*, 437.

Patel, D. J., Kozlowski, S. A., Ikuta, S., & Itakura, K. (1984a) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 43*, 2663.

Patel, D. J., Kozlowski, S. A., Ikuta, S., & Itakura, K. (1984b) *Biochemistry 23*, 3207.

Summers, M. F., Byrd, R. A., Gallo, K. A., Samson, C. J., Zon, G., & Egan, W. (1985) *Nucleic Acids Res. 13*, 6375.

Tinoco, I., Jr., Wolk, S., Arnold, F. H., & Aboul-ela, F. (1987) in *Structure and Dynamics of Biopolymers* (Nicolini, C., Ed.) Plenum, New York (in press).