Kinetics and Energetics of Base-Pair Opening in 5′-d(CGCGAATTCGCG)-3′ and a Substituted Dodecamer Containing G·T Mismatches†

James G. Moe and Irina M. Russu*

Department of Molecular Biology and Biochemistry, Wesleyan University, Middletown, Connecticut 06459

Received February 6, 1992; Revised Manuscript Received June 9, 1992

ABSTRACT: Proton nuclear magnetic resonance (NMR) spectroscopy is used to characterize the kinetics and energetics of base-pair opening in the dodecamers 5′-d(CGCGAATTCGCG)-3′ and 5′-d(CGCGAATTTGCG)-3′. The latter dodecamer contains two symmetrical G·T mismatched base pairs. The exchange kinetics of imino protons is measured from resonance line widths and selective longitudinal relaxation times. For the G·T pair, the two imino protons (G-N1H and T-N3H) provide probes for the opening of each base in the mismatched pair. The lifetimes of individual base pairs in the closed state and the equilibrium constants for formation of the open state are obtained from the dependence of the exchange rates on the concentration of ammonia catalyst. The activation energies and standard enthalpy changes for base-pair opening are obtained from the temperature dependence of the lifetimes and equilibrium constants, respectively. The results indicate that the G·T mismatched pairs are kinetically and energetically destabilized relative to normal, Watson–Crick base pairs. The lifetimes of the G·T pairs are of the order of 1 ms or less, over the temperature range from 0 to 20 °C. The equilibrium constants for base-pair opening, at 20 °C, are increased up to 4000-fold, relative to those of normal base pairs. The energetic destabilization of the G·T base pairs is, at least in part, enthalpic in origin. The presence of the G·T mismatched base pairs destabilizes also neighboring base pairs. For the A·T base pairs adjacent to the mismatch sites, the lifetimes are decreased 3-fold, and the equilibrium constants for opening are increased 10-fold, as compared to the corresponding values in the absence of the mismatch. This destabilization is not accompanied by significant changes in the activation energy and the standard enthalpy change for base-pair opening. Lower standard enthalpy changes for opening are, however, observed for the central A·T base pairs, which are two bases removed from the mismatch site.

Noncomplementary base pairs (mismatches) occur with varying frequencies during DNA replication and genetic recombination or can be induced by chemical reactions. These errors are usually detected and eliminated by excision by DNA polymerase and by a postreplicative mismatch repair system of enzymes (Modrich, 1987). Recently, it has been shown that noncomplementary base pairs can be accommodated in the recognition sites of restriction endonucleases and the reduction in cleavage rate is dependent upon the nature and location of the mismatch in these sites (Petranovic et al., 1990; Thielking et al., 1990).

The presence of base-pair mismatches in DNA and RNA and the ability of enzymes to discriminate them from normal Watson–Crick base pairs have enhanced the interest in the conformation and dynamics of these structures. Among them, the G·T mismatch has been studied extensively (Kennard, 1987; Patel et al., 1987). This mismatch is incorporated with a high rate during DNA replication, (8×10^{-5} as compared, for example, to 2×10^{-7} for the C·T mismatch) but at the editing stage, it is proofread most efficiently (Fersht et al., 1982). G·T mismatches can be incorporated in the recognition sites of restriction endonucleases such as *Eco*RI (Thielking et al., 1990), *Sma*I, *Xma*I, *Ava*I, *Hpa*II, *Msp*I, *Nci*I, and *Nsp*III (Petranovic et al., 1990). In the case of *Eco*RI endonuclease, oligonucleotides containing a single G·T mismatch in various positions of the recognition site 5′-d(GAATTC)-3′ are better substrates than the corresponding *Eco*RI* oligonucleotides (Thielking et al., 1990).

The hydrogen bonding for the G·T mismatch, originally proposed by Crick (1966), is shown in Chart I. This wobble pairing by imino proton–carbonyl hydrogen bonds has been confirmed by X-ray diffraction (Kennard, 1985, 1987; Hunter et al., 1987) and nuclear magnetic resonance (NMR) spectroscopy (Patel et al., 1982a, 1987; Hare et al., 1986; Kalnik et al., 1988). The crystallographic structure of the dodecamer 5′-d(CGCGAATTTGCG)-3′ shows that the G·T pair is accommodated into the B-DNA helix with small changes in base stacking and sugar phosphate backbone at the mismatch site and no significant perturbations in the rest of the molecule. The structure of the related dodecamer 5′-d(CGTGAATTCGCG)-3′ obtained from a distance geometry analysis of 2D NMR data has confirmed this conclusion for the DNA in solution state (Hare et al., 1986).

In the present work, we have extended these structural studies of the G·T mismatched pair by a characterization of the kinetics and energetics of base-pair opening in the dodecamer 5′-d(CGCGAATTTGCG)-3′ containing symmetrical G·T pairs (designated GT dodecamer). The results are compared to those in the parent dodecamer, 5′-d(CGCGAATTCGCG)-3′, which contains normal, Watson–Crick base pairs (designated GC dodecamer). Extensive measurements on a large variety of DNA oligonucleotides have established that base pairs open one at a time, and in B-DNA, base-pair lifetimes are in the range of milliseconds at room temperature (Leroy et al., 1988a; Gueron et al., 1990; Braunlin & Bloomfield, 1988). The equilibrium constants for formation of the open state are small, namely, 10^{-5}–10^{-7} for A·T and G·C base pairs, respectively (Gueron et al., 1990). More recent studies have demonstrated the potential of the NMR proton exchange method for detecting and characterizing sequence- and ligand-

† This work was supported by a grant from the NSF (88-17589). J.G.M. was a recipient of a High-Technology scholarship from the State of Connecticut.

* To whom correspondence should be addressed.

0006-2960/92/0431-8421$03.00/0

Chart I: G·T Wobble Pair Observed By X-ray Diffraction in the Dodecamer 5′-d(CGCGAATTTGCG)-3′ (Kennard, 1987; Hunter et al., 1987)

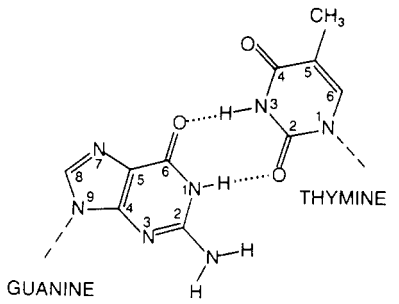

Chart II: Numbering of Base Pairs in the Two Dodecamers Investigated[a]

```
      1 2 3 4 5 6
5'-d(C G C G A A T T X G C G)-3'
3'- (G C G X T T A A G C G C)d-5'
          6 5 4 3 2 1
```

[a] X = C in GC dodecamer; X = T in GT dodecamer.

induced perturbations of the dynamics of base-pair opening (Leroy et al., 1988b, 1991). Sequence dependence of base-pair lifetimes has been observed for DNA oligonucleotides containing tracts of consecutive A·T base pairs (Leroy et al., 1988b; Moe & Russu, 1990). Transition from B- to Z-DNA strongly affects the exchange properties of imino protons and results in considerably longer lifetimes of the base pairs (Kochoyan et al., 1990).

IMINO PROTON EXCHANGE

The exchange of imino protons with solvent protons is believed to occur from an open state of the base pair. The conformational features of this state are not yet fully understood. To account for proton exchange data, it is generally assumed that, in the open state, the imino groups are fully accessible to solvent and the base-pairing hydrogen bonds are broken such that the imino proton becomes available for hydrogen bonding with water or catalyst molecules (Englander & Kallenbach, 1984). The exchange from the open state is catalyzed by proton acceptors such as added catalyst, OH^-, and the other base of the open base pair (Gueron et al., 1987). When the equilibrium constant for formation of the open state (K_{op}) is small and the exchange by added catalyst dominates, the overall exchange rate of the imino proton is

$$k_{ex} = k_{op}k_B[B]/(k_{cl} + k_B[B]) \quad (1)$$

where k_{op} and k_{cl} are the rate constants for opening and closing of the base pair, respectively ($K_{op} = k_{op}/k_{cl}$). An alternative form of eq 1 is (Leroy et al., 1988a)

$$\tau_{ex} = 1/k_{ex} = \tau_0 + D[B]^{-1} \quad (2)$$

where $\tau_0 = 1/k_{op}$ is the lifetime of the base pair in the closed state and D is a constant defined as $(K_{op}k_B)^{-1}$.

The lifetime of the base pair, τ_0, can be obtained as the limit of τ_{ex} at infinite concentration of catalyst. The determination of the equilibrium constant, K_{op}, from the dependence of the exchange rate on catalyst concentration requires the knowledge of the catalytic rate constant, k_B. In DNA, this rate is expected to be different from that in a free nucleoside for several reasons (Schurr et al., 1988; Gueron et al., 1990). First, the diffusion-controlled rate of association of the catalyst with the imino group should be affected by (i) the electrostatic potential of the DNA, (ii) the limited steric access of the catalyst to the imino proton in the open state of the base pair, and (iii) the lower diffusion coefficient of the imino group in DNA. Second, the electrostatic potential of the DNA affects the proton dissociation constant of the imino group.

According to a recent theoretical treatment (Benight et al., 1988), when ammonia is the catalyst and in 0.1 M NaCl, k_B can be corrected for electrostatic effects as

$$k_B = k_D'/(1 + 0.27 \times 10^{-\Delta pK}) \quad (3)$$

where k_D' is the rate of association of the catalyst with the imino group in the absence of any long-range electrostatic potentials and $\Delta pK = pK_c - pK_{NH}$, where pK_{NH} is the pK value of the imino group in DNA (which is predicted to be increased by 0.8 unit relative to that in a free nucleoside) and pK_c is the pK value of the catalyst. The association rate k_D' can be calculated at 25 °C as (Benight et al., 1988)

$$k_D' = \theta(1.2 \times 10^{10})[\Phi_c + \Phi_{NH}]/2 \quad (4)$$

where θ is a steric factor due to any constraints on the allowed solid angle of approach that are not alleviated by rotational diffusion and Φ_c and Φ_{NH} are the ratios of the diffusion coefficients of catalyst and imino group to that of NH_4^+, respectively ($\Phi_c = 1$ for ammonia, $\Phi_{NH} = 0.5$ for a free nucleoside, and $\Phi_{NH} \simeq 0$ for duplex DNA).

An experimental analysis of imino proton exchange induced by various catalysts in poly(rA)·poly(rU) and in the decamer 5′-d(CGCGATCGCG)-3′ indicates that the differences between exchange in free nucleosides and that in DNA can be attributed mainly to the lower diffusion coefficient of the imino group in DNA (Gueron et al., 1990). Certainly, this should be true for ammonia, which is small and uncharged. According to eq 4, due to the lower diffusion coefficient in DNA, the catalytic rate constant k_B in DNA should be reduced by a factor of 1.5 relative to its value in a free nucleoside.

EXPERIMENTAL PROCEDURES

Materials. Oligonucleotides were synthesized on an Applied Biosystems 381A automated DNA synthesizer using the solid-support phosphoramidite method. They were purified using reverse-phase high-pressure liquid chromatography (HPLC) on a Hamilton PRP-1 preparative column. The purification was run in 50 mM ethylenediamine/formate buffer at pH 7.5 with a gradient of 0–25% acetonitrile over 25 min. The DNA samples eluted at approximately 20 min.

The purified oligonucleotides were dialyzed extensively against 0.6 M NaCl to exchange the counterion to sodium. The samples were then dialyzed with four exchanges against 2 mM ammonia buffer containing 0.1 M NaCl, 2 mM EDTA, and 10% D_2O at pH 8.85 and at 20 °C. The dialysis buffer contained Chelex 100 (100–200 mesh, sodium form, Bio-Rad) to remove metal ions from the samples. Various concentrations of ammonia were obtained by titrating the samples with Chelex-treated stock solutions of either 0.4 or 2 M ammonia buffer containing 0.1 M NaCl and 2 mM EDTA at pH 8.85. The final DNA concentration was approximately 0.75 mM (duplex).

Methods. ^1H NMR experiments were carried out on a Varian VXR-400 NMR spectrometer operating at 400 MHz. The observation pulse was the 1–3–3–1 pulse (Hore, 1983). All spectra were referenced to 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) as an external reference. Longitudinal relaxation times, T_1, were measured using the selective saturation recovery method. The imino proton resonances

were saturated individually and 12–15 values for the recovery delay were used for each T_1 measurement. The T_1 values were obtained from exponential nonlinear least-squares fits of intensity as a function of recovery delay. For the imino proton resonances of G·C2 and G·C3 base pairs in the GT dodecamer (which overlap over the entire temperature range investigated), individual T_1 values were calculated by fitting the data to a sum of two exponentials. This analysis gave accurate T_1 values at intermediate concentrations of catalyst where the two T_1 values are significantly different (Craik et al., 1991). The T_1 values for these resonances are less accurate at low catalyst concentrations (where the two T_1 values are similar) and at high catalyst concentrations (where the T_1 values are very short and substantial broadening occurs). The transverse relaxation times, T_2, were calculated from the line widths at half-height ($\Delta\nu$) of the imino proton resonances: $\Delta\nu = 1/\pi T_2$. The errors in the line widths due to magnetic field inhomogeneity were estimated from four series of independent measurements on the same DNA, as a function of base catalyst concentration. They were found to be 10% or less of the measured values.

The longitudinal relaxation time, T_1, is related to the exchange rate by

$$1/T_1 = k_{ex} + 1/T_1^\circ \quad (5)$$

or, based on eq 2

$$T_1 = T_1^\circ(\tau_0 + D[B]^{-1})/(T_1^\circ + \tau_0 + D[B]^{-1}) \quad (6)$$

where T_1° is the longitudinal relaxation time corresponding to mechanisms other than exchange (e.g., proton–proton dipolar interactions). The equations for T_2 are similar.

The concentration of base catalyst was calculated as $[B] = 1/(1 + 10^{(pK-pH)})$. The p$K$ values of ammonia (pK = 9.24 at 25 °C) were corrected for the temperature dependence using a standard enthalpy of ionization of 12.4 kcal/mol (Weast, 1986). The pH of the buffer was measured directly at all temperatures of interest.

The experimental values for T_1 and T_2 were fitted to eq 6 using a nonlinear least-squares algorithm (Johnson & Frasier, 1985) from which the base-pair lifetime (τ_0), the constant D, and the relaxation times T_1° or T_2° were obtained. The errors reported in the paper for these parameters represent the standard deviations of the nonlinear least-squares fit. The exchange data for several imino proton resonances were also analyzed as k_{ex} as a function of catalyst concentration (eq 1) and τ_{ex} as a function of the inverse of the catalyst concentration (eq 2). The exchange rates k_{ex} (or the exchange times τ_{ex}) were calculated using eq 5 from the experimental relaxation times and the relaxation times T_1° or T_2° obtained from the nonlinear least-squares fit. The results of the latter two methods were, within experimental errors, the same as those obtained by fitting experimental T_1 and T_2 values to eq 6.

The equilibrium constants for base-pair opening were calculated from the constant D as $K_{op} = 1/Dk_B$. The catalytic rate constant k_B was reduced by a factor of 1.5 relative to the value measured for free nucleosides [namely, 2×10^8 M^{-1} s^{-1} for guanosine 2′:3′-cyclic monophosphate and thymidine (Gueron et al., 1990)] as described in the previous section.

EXPERIMENTAL RESULTS

Base-pair opening in the two dodecamers was characterized by use of the imino proton resonances (N3H in thymine and N1H in guanine) shown in Figure 1. The imino resonances of the GC dodecamer have been previously assigned by Patel and co-workers (Patel et al., 1982b). In the GT dodecamer,

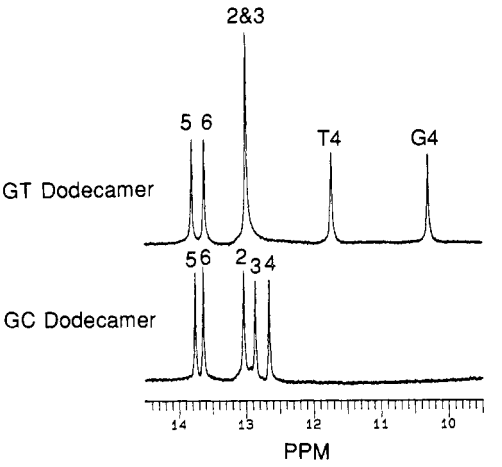

FIGURE 1: Imino proton resonances of the dodecamers investigated in 2 mM ammonia buffer containing 100 mM NaCl and 2 mM EDTA at pH 9.0 and at 15 °C. Assignments of resonances to individual protons are indicated according to the numbering of base pairs shown in Chart II.

the imino proton resonances from G·C and A·T base pairs occur between 12.5 and 14.5 ppm whereas those from the wobble G·T pair occur at 10.3 and 11.7 ppm. They were assigned in the present work by standard 1D NOE methods. The results are illustrated in Figure 2 for the assignments of the two imino proton resonances of the wobble G·T pair. The two imino proton resonances show strong NOEs between themselves as well as NOEs to the neighboring G·C3 and A·T5 base pairs (Figure 2B and C). The resonance at 10.3 ppm also gives a NOE to the A5-H2 proton resonance at 7.25 ppm (Figure 2C). This result assigns the 10.3 ppm resonance to G4-N1H since, according to the crystallographic structure, the G4-N1H is closer to A5-H2 than T4-N3H is (Kennard, 1987). The rest of the imino proton resonances in the GT dodecamer occur at the same positions as the corresponding ones in the GC dodecamer except for the resonance of the G·C3 base pair, which is shifted downfield by 0.1 ppm and thus overlaps with the imino proton resonance of G·C2 (Figure 1). Our assignments of the imino protons of the G·T pair are in perfect agreement with those obtained previously for related dodecamers containing G·T mismatches, namely, 5′-d(CGT-GAATTCGCG)-3′ (Patel et al., 1982a; Hare et al., 1986) and 5′-d(CGCGAGCTTGCG)-3′ (Kalnik et al., 1988).

In order to characterize the base-pair opening in the GC and GT dodecamers, we have measured the line width (for the G4 and T4 imino proton resonances in the GT dodecamer) and the selective T_1 (for all the other imino proton resonances) as a function of the concentration of base catalyst. The exchange of the imino proton of the G·C1 pair could not be measured since its resonance broadens out at temperatures above 5 °C due to fraying at the ends of the duplex. Representative experimental curves are shown in Figure 3 for the G-N1 protons of the mismatched pair in the GT dodecamer and of the G·C4 base pair in the GC dodecamer. An example of the results for base-pair lifetimes is given in Table I. For G·C2, the lifetimes are greatly affected by fraying at the ends (lifetimes of <5 ms), and because of this, they are not reported. As shown in Table I, the experimental errors of the lifetime measurements for A·T base pairs are 10% or less. In contrast, for G·C base pairs, the errors are much larger. This is probably due to the fact that, for base pairs with small dissociation constants such as these, the slope of the relaxation time vs the inverse of the catalyst concentration is greater and, thus, extrapolation to high base catalyst concentration is less accurate (Figure 3A). The larger errors

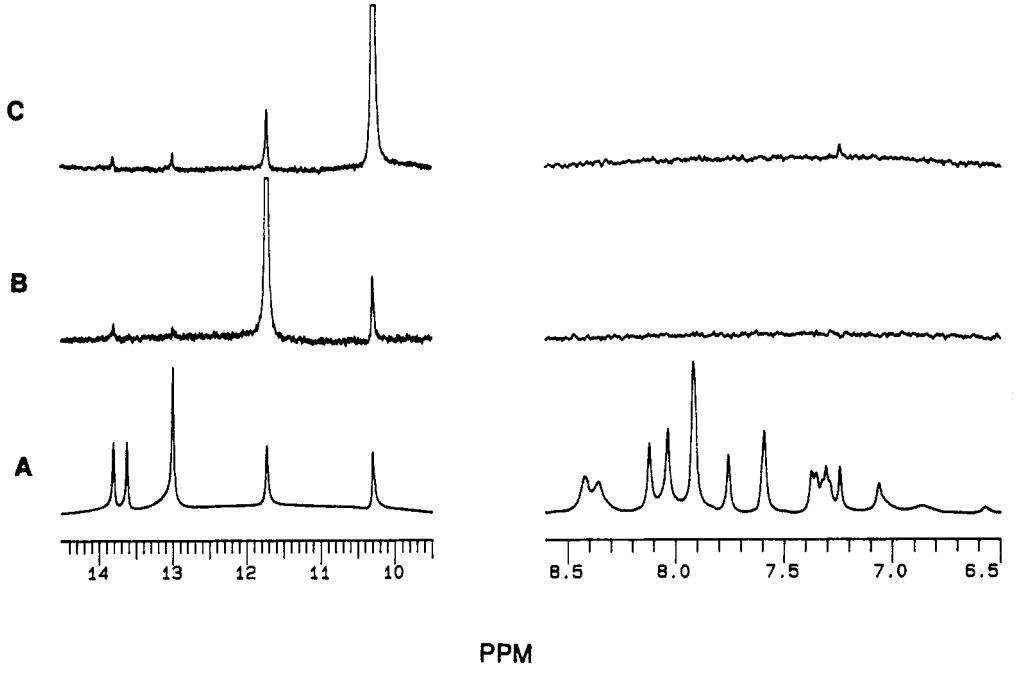

FIGURE 2: Assignments of the G-N1H and T-N3H resonances from the G·T pair in the GT dodecamer: (A) control spectrum; (B and C) NOE difference spectra with an irradiation time of 200 ms. The samples were in 10 mM phosphate buffer containing 0.1 M NaCl, 2 mM EDTA, and 10% D_2O at pH 7.0.

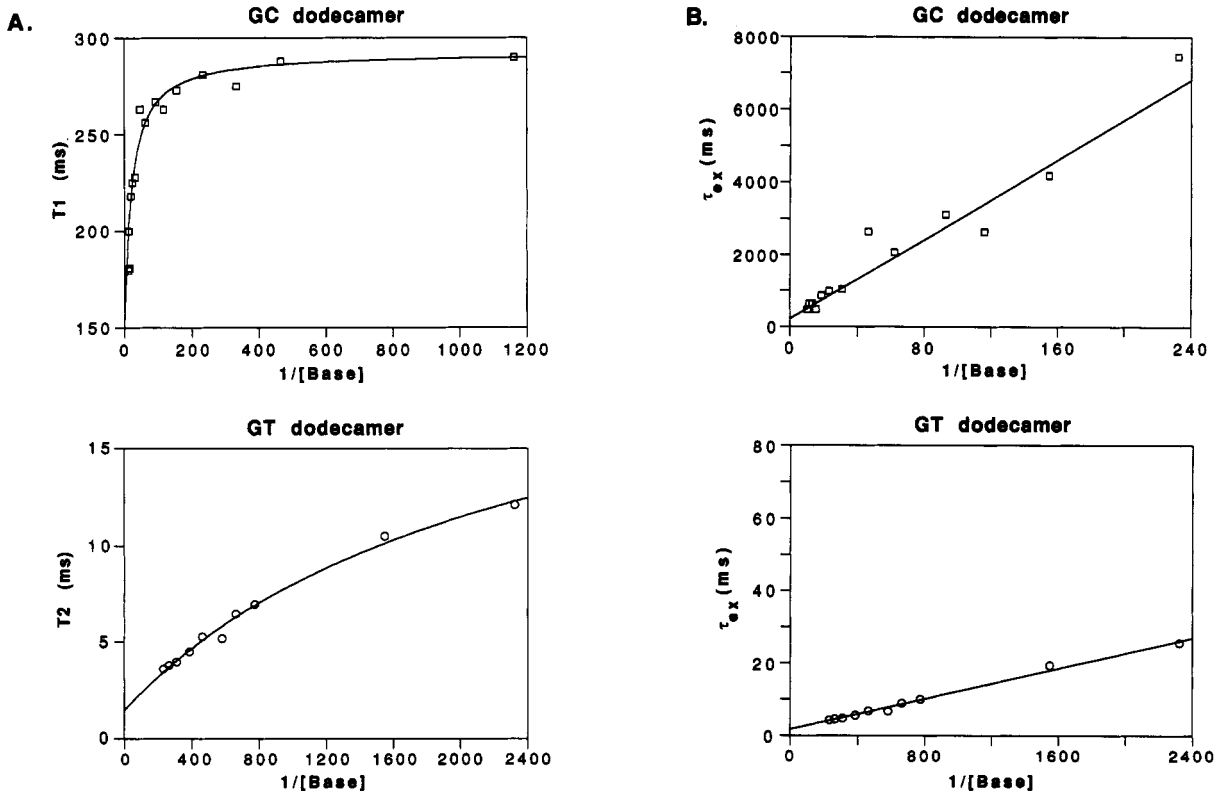

FIGURE 3: Relaxation times (A) and exchange times (B) as a function of the inverse of ammonia base concentration for G-N1H in the G·C4 base pair of the GC dodecamer and G-N1H in the G·T base pair of the GT dodecamer at 15 °C. The curves in (A) correspond to nonlinear least-squares fits to eq 6.

for these base pairs are maintained when lifetimes are calculated from exchange rates (eq 1) or from exchange times (eq 2). The lifetimes of the G·T pair obtained from the line width of the G4 imino proton resonance are all around 1 ms or less. This range represents the lower limit for the applicability of the NMR proton exchange method (Gueron et al., 1990). For the same base pair, the data obtained from the line width of the T4 imino proton resonance extrapolate to negative values at infinite concentration of catalyst. All these values are within experimental errors of zero (for example, see footnote in Table I). A similar behavior has been observed in several other DNA oligomers (Plum & Bloomfield, 1990).

The exchange measurements were carried out as a function of temperature from 0 to 20 °C for the GT dodecamer and from 10 to 30 °C for the GC dodecamer. The choice of these temperature ranges was dictated by the melting temperatures of the two dodecamers (i.e., 56 °C for the GT dodecamer in

Table I: Base-Pair Lifetimes (ms) in the GC and GT Dodecamers at 15 °C in Ammonia Buffer Containing 100 mM NaCl and 2 mM EDTA at pH 9.00

	base pair			
dodecamer	G·C3	G·C4/G·T4	A·T5	A·T6
GC	48 ± 29	299 ± 117	21 ± 2	71 ± 5
GT	31 ± 6[a]	1.5 ± 0.6[b]	7 ± 1	65 ± 4

[a] Value obtained from the overlapped resonances of G·C2 and G·C3 using a sum of two exponentials. [b] Value obtained from the exchange of the G-N1 proton; the value obtained from the exchange of the T-N3 proton is −0.4 ± 0.4 ms.

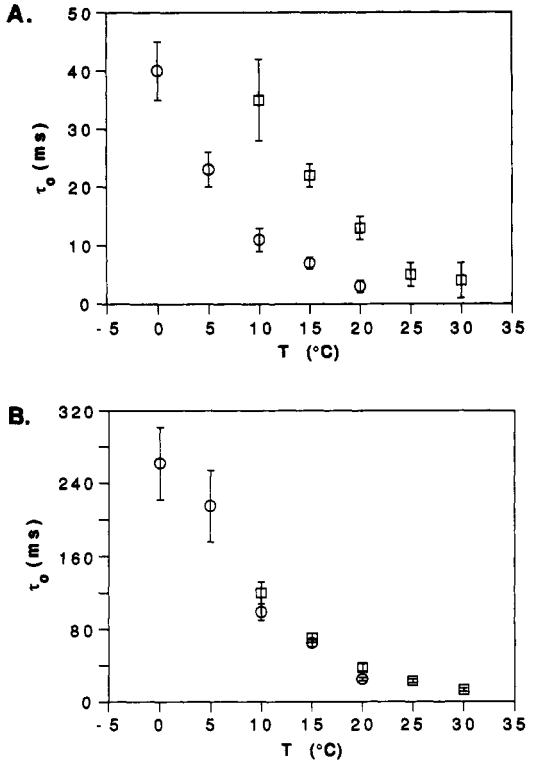

FIGURE 4: Dependence on temperature of the lifetimes of base pairs A·T5 (A) and A·T6 (B): (□) GC dodecamer; (○) GT dodecamer.

0.5 M ammonia buffer/0.1 M NaCl/2 mM EDTA at pH 8.85 and 80 °C for the GC dodecamer in 10 mM phosphate buffer/0.3 M NaCl/2 mM EDTA at pH 7.0, as measured from the melting transitions observed in the NMR spectra). At the highest temperatures used in this study, the two dodecamers are more than 30 °C below their melting temperatures and, thus, the contribution of strand dissociation to the exchange can be neglected (Braulin & Bloomfield, 1988; Leroy et al., 1988a). The dependence of the base-pair lifetimes on temperature is illustrated in Figure 4 for the A·T base pairs in the two dodecamers.

Activation energies for base-pair opening, E_a, were obtained from the Arrhenius equation: $\ln \tau_0 = -\ln(A) + E_a/RT$, where A is the frequency factor, T is the absolute temperature, and R is the gas constant. The results for the activation energies are summarized in Table II. For the mismatched G·T pair, the lifetime measured from exchange of the G4 imino proton remained of the order of 1 ms or less over the entire temperature range. This result suggests that the activation energy for opening the G·T pair is small; however, since the lifetimes are at or below the lower limit for the applicability of the NMR method, the determination of activation energies from lifetimes is not reliable for the G·T pair.

The equilibrium constants for base-pair opening are illustrated in Table III. The enthalpies, $\Delta H°$, for base-pair opening

Table II: Activation Energies (kcal/mol) for Base-Pair Opening in the GC and GT Dodecamers

	base pair			
dodecamer	G·C3	G·C4/G·T4	A·T5	A·T6
GC	21 ± 5	19 ± 4	18 ± 3	19 ± 1
GT	19 ± 6	*a*	19 ± 2	19 ± 1

[a] Not determined (see text).

were calculated on the basis of the dependence of equilibrium constants on temperature according to the van't Hoff equation: $\ln K_{op} = -\Delta H°/RT + \Delta S°/R$ (Figure 5). The results are summarized in Table IV.

DISCUSSION

Imino Proton Exchange in the Mismatched G·T Pair. The G·T mismatched pair contains two hydrogen-bonded imino protons (N1H in G and N3H in T) which can be individually observed in the NMR spectra (Figure 1). The exchange of these protons with water protons is base-catalyzed (Figure 3) and is greatly slowed down relative to that in free nucleosides. For example, at 15 °C and an ammonia base concentration of 1.3 mM, the exchange rate of the G4 imino proton is 100 s^{-1}. For free guanosine 2′:3′-cyclic monophosphate, at the same temperature and ammonia concentration, the exchange rate of the N1H proton would be 2.6×10^5 s^{-1} [based on a value of the catalytic rate constant of 2×10^8 M^{-1} s^{-1} (Gueron et al., 1990)]. Thus, the exchange of the imino proton is slowed down by a factor of 2.6×10^3 in the DNA duplex as compared to that in a free nucleoside. This fact strongly suggests that the exchange of the imino proton takes place from an open state of the G·T mismatched base pair.

The crystallographic structure of the GT dodecamer indicates that the G and T bases in the mismatched pair are displaced toward the minor and the major grooves, respectively (Kennard, 1987; Hunter et al., 1987). In this conformation, the T-N3 proton forms a hydrogen bond with G-O6 and the G-N1 proton forms a hydrogen bond with T-O2 (Chart I). The displacement of the bases alters the symmetry of the base pair with respect to the glycosidic bonds. The angle between the glycosidic bond C1′–N1 of T and the vector joining the C1′ carbons of the two sugar residues increases to ~70° (from its value of 50° in a Watson–Crick A·T base pair). Similarily, the angle between the glycosidic bond C1′–N9 of G and the C1′–C1′ vector decreases to ~43° (from its value of 51° in a Watson–Crick G·C base pair). These structural observations suggest that the formation of the open states (i.e., states operative in proton exchange) for the G·T pair may be different from that for a normal pair. The process should involve breakage of the hydrogen bonds, but the unstacking of the bases could be less than that in a normal base pair (assuming opening toward the major groove for the T and toward the minor groove for the G). This mechanism would result in lifetimes for the G·T pair in the closed (nonexchanging) state much shorter than the corresponding ones in normal base pairs, as observed experimentally (Table I). Furthermore, this different opening process should be expected to have a larger equilibrium constant, K_{op}, and a lower enthalpy change (vide infra).

Influence of the G·T Pair on the Kinetics of Opening of Neighboring Base Pairs. The presence of the G·T pair decreases the lifetime of the neighboring A·T5 base pair (Table I), and this decrease is maintained at all temperatures investigated (Figure 4A). No significant kinetic destabilization is observed for the 5′-neighboring base pair (G·C3) or

Table III: Equilibrium Constants for Base-Pair Opening in the GC and GT Dodecamers at 20 °C in Ammonia Buffer Containing 100 mM NaCl and 2 mM EDTA at pH 8.87

	base pair			
dodecamer	G·C3	G·C4/G·T4	A·T5	A·T6
GC	$(1.0 \pm 0.1) \times 10^{-6}$	$(3.0 \pm 0.6) \times 10^{-7}$	$(5.8 \pm 0.3) \times 10^{-6}$	$(4.0 \pm 0.4) \times 10^{-6}$
GT	$(6 \pm 1) \times 10^{-6}$	$(7 \pm 2) \times 10^{-4}$[a]	$(5.0 \pm 0.3) \times 10^{-5}$	$(1.0 \pm 0.1) \times 10^{-5}$

[a] Value obtained from the exchange of the G-N1 proton; the value obtained from the exchange of the T-N3 proton is $(1.2 \pm 0.2) \times 10^{-3}$.

Table IV: Standard Enthalpy Changes $\Delta H°$ (kcal/mol) for Base-Pair Opening in the GC and GT dodecamers

	base pair			
dodecamer	G·C3	G·C4/G·T4	A·T5	A·T6
GC	25 ± 1	17 ± 3	16 ± 1	22 ± 1
GT	19 ± 4	10 ± 2[a]	17 ± 1	16 ± 3

[a] Value obtained from the exchange of the G-N1 proton; the value obtained from the exchange of the T-N3 proton is 12 ± 2 kcal/mol.

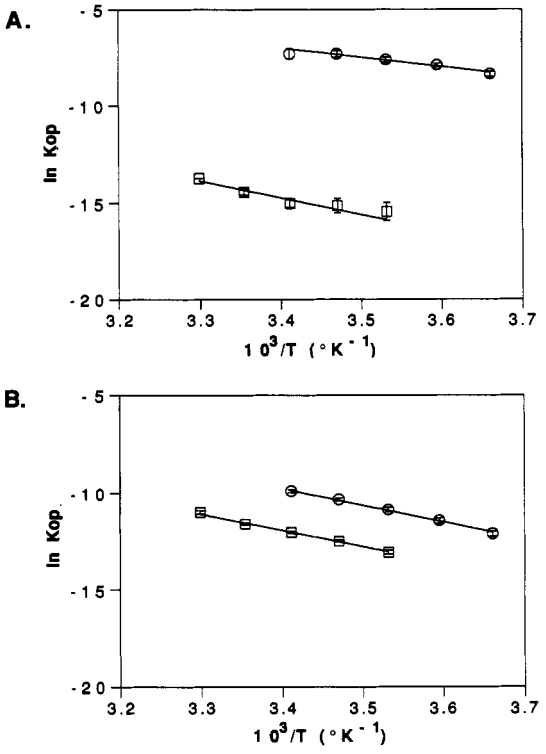

FIGURE 5: Van't Hoff plots for the opening of G·C4 and G·T4 base pairs (A) and A·T5 base pairs (B): (□) GC dodecamer; (○) GT dodecamer.

for base pairs removed from the site of the mismatch (Table I and Figure 4B). Hence, the dynamic effects of the G·T pair are localized in close vicinity of the mismatch site. This finding is consistent with the crystallographic structure, which shows that conformational differences between normal B-DNA and the GT dodecamer are small and affect solely the local environment of the mismatch site (Kennard, 1987; Hunter, et al., 1987).

Previous studies have demonstrated that, in DNA oligonucleotides containing tracts of four or more consecutive A·T base pairs, the lifetimes of the central A·T pairs are anomalously long (Leroy et al., 1988b; Moe & Russu, 1990). For example, in the GC dodecamer, the lifetime of the A·T6 base pair at 15 °C is increased more than 3-fold relative to that of the A·T5 base pair (Table I). This property is maintained in the GT dodecamer over the entire temperature range investigated (Figure 4).

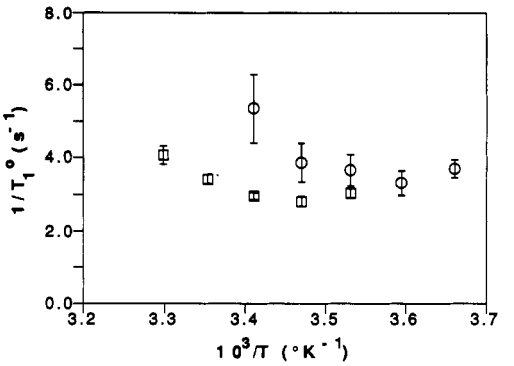

FIGURE 6: Dependence of the exchange-independent longitudinal relaxation rate, $1/T_1°$, on temperature for the imino proton of the A·T5 base pair: (□) GC dodecamer; (○) GT dodecamer. The $T_1°$ values were obtained by fitting the experimental T_1 values to eq 6.

The activation energies for opening of the normal base pairs in the G·T dodecamer, including A·T5, are the same as the corresponding ones in the GC dodecamer (Table II). This finding suggests that the decrease in the lifetime of the A·T5 pair in the GT dodecamer (Figure 4A) is due to a change in the entropy of activation for the formation of the open state.

Imino proton exchange kinetics in the presence of a G·T mismatch has been investigated earlier using NMR for the dodecamer 5′-d(CGTGAATTCGCG)-3′ (Pardi et al., 1982; Patel et al., 1984). Base-pair lifetimes, measured by saturation recovery in 0.1 M phosphate buffer, indicated a destabilization by a factor of 2 for the base pair 3′-adjacent to the mismatch site and no effect on the lifetimes of distant base pairs. These results are in qualitative agreement with those obtained in the present paper (Table I). In these earlier studies, it was also observed that, for the central A·T base pairs, the slopes of the saturation recovery rates vs the inverse of temperature were greatly increased in the G·T-containing dodecamer. This result yielded higher activation energies for opening of these base pairs in the presence of the mismatch (namely, 37 and 48 kcal/mol as compared to 14 and 15 kcal/mol in the dodecamer containing G·C pairs). The differences between these previous results and those obtained here can be explained by several facts. First, the previous observations were made at temperatures higher than 20 °C, a range in which proton exchange could involve, in addition to single base-pair opening, cooperative opening of several base pairs (including strand dissociation). Given the thermal destabilization induced by the mismatch, the latter pathway is more probable at elevated temperatures (Pardi et al., 1982). Second, in the presence of phosphate buffer, which is an inefficient catalyst for imino proton exchange, the exchange-independent relaxation (eq 5) makes a significant contribution to the observed saturation recovery rates. We have found that the exchange-independent relaxation times, $T_1°$, for the central A·T base pairs have a stronger temperature dependence in the GT dodecamer than in the GC dodecamer. This fact is illustrated in Figure 6 for the A·T5 base pairs. Thus, the previously observed temperature dependence originates, in part, from the dependence on temperature of the relaxation times $T_1°$.

Energetics of Base-Pair Opening. In the GC dodecamer, the equilibrium constants for base-pair opening, K_{op}, at 20 °C range from $\sim 6 \times 10^{-6}$ to $\sim 3 \times 10^{-7}$ (Table III). These values correspond to standard free energy changes for opening, $\Delta G°$, from 7.0 to 8.7 kcal/mol. The equilibrium constants are consistently lower for G·C than for A·T base pairs, over the entire temperature range investigated. These results are in good agreement with values obtained previously for other DNA oligonucleotides of various base sequences (Gueron et al., 1990).

The equilibrium constants for the G·T mismatched pair are increased 100–4000-fold relative to those of normal base pairs in the GC dodecamer (Table III). Thus, the formation of the open state(s) in the G·T pair is energetically more favorable than in a normal base pair (for example, the formation of the open state of the T-N3 proton at 20 °C is more favorable by 3.1 kcal/mol than that of the A·T5 base pair in the GC dodecamer). The presence of the mismatched G·T pair destabilizes the other base pairs in the GT dodecamer. The largest effects are observed for A·T5 and G·C3 base pairs for which the standard free energy changes for opening are decreased by 1.3 and 1.0 kcal/mol, respectively. Hence, as for the kinetic effects described above, this destabilization is mostly confined to the base pairs adjacent to the mismatch site.

The dependence of the equilibrium constants K_{op} on temperature can be used to calculate the standard enthalpy changes, $\Delta H°$, for base-pair opening. The validity of this approach is supported by experimental results as well as theoretical considerations. Leroy et al. (1991) have found that the rate constants for ammonia-catalyzed exchange in free nucleosides are independent of temperature over the range from 15 to 60 °C. This result suggests that the temperature dependence of the imino proton exchange rates in DNA reflects the dependence on temperature of the equilibrium constants for base-pair opening. A similar conclusion can be reached based on a theoretical analysis of the catalytic rate constants k_B in DNA. According to eq 3, the temperature dependence of k_B should result from the temperature dependence of the pK values of interest and of the diffusion-controlled rates of association, k_D' (the effects of the DNA electrostatic potential have a weak temperature dependence and thus are assumed temperature-independent). The dependence on temperature of the rates k_D' can be estimated by assuming a linear dependence on T/η_T where η_T is the solvent viscosity at the temperature of interest (T) (Figure 7). The pK values of the catalyst and imino groups at various temperatures can be calculated on the basis of their standard enthalpies of ionization, $\Delta H°$ [12.4 kcal/mol for ammonia, 7.7 kcal/mol for the imino group in T, and 3.6 kcal/mol for the imino group in G (Izatt et al., 1971)]. The resulting temperature dependence of the factor $1 + 0.27 \times 10^{-\Delta pK}$ (eq 3) and of the catalytic rate constants are shown in Figure 7. Clearly, the theory predicts that, for ammonia-catalyzed exchange, the increase in the rate of association with temperature is largely compensated by the changes in pK values and, thus, the catalytic rate constant is nearly temperature-independent. This theoretical observation supports the idea that the temperature dependence of the constant D in eq 6 reflects mostly the dependence on temperature of the equilibrium constant for formation of the open state.

The standard enthalpy changes, $\Delta H°$, for base-pair opening in the GC dodecamer range from 16 to 25 kcal/mol (Table IV). These values are in good agreement with those obtained previously for DNA oligonucleotides of various sequences

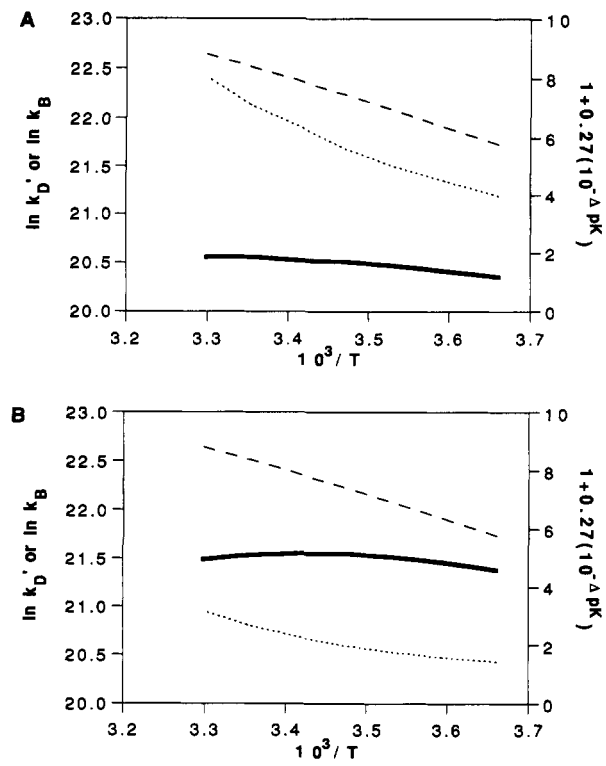

FIGURE 7: Theoretical analysis of the temperature dependence of the rate constant for ammonia-catalyzed exchange of T-N3H (A) and G-N1H (B): (- - -) diffusion-controlled association constant between catalyst and imino group in DNA, k_D'; (···) $1 + 0.27 \times 10^{-\Delta pK}$; (—) catalytic rate constant, k_B (for details, see text).

(Gueron et al., 1990) and for poly(dA)·poly(dT) and $[d(AT)]_{15}$·$[d(AT)]_{15}$ (Plum & Bloomfield, 1990) as well as with the molecular modeling of base-pair opening carried out by Ramstein and Lavery (1988). The enthalpy changes do not show any correlation to the nature of the base pair but have some dependence on base sequence; namely, $\Delta H°$ for the central A·T6 base pair in the A_2T_2 tract is larger than that for the A·T5 base pair. The enthalpy changes calculated on the basis of the temperature dependence of the equilibrium constants K_{op} (Table IV) are, within experimental errors, the same as the activation energies obtained from the temperature dependence of the lifetimes (Table II). This fact implies that the activation energies for closing of the base pairs are close to zero and, thus, no stabilizing interactions operate in the open state(s).

In the GT dodecamer, the enthalpy change for opening of the mismatched pair is lower (Table IV). This decrease is consistent with our suggestion that the formation of the open states of the T-N3 and G-N1 protons of the G·T pair involves less unstacking of the bases. For the other base pairs, the enthalpy changes are, within experimental errors, the same as those in the GC dodecamer except for the central A·T6 base pair. The enthalpy change for opening of this base pair is lower by 6 kcal/mol than the corresponding value in the GC dodecamer (Table IV). This finding suggests that the presence of the mismatch next to the A_2T_2 tract affects the energetics of base-pair opening within the tract.

CONCLUSIONS

The G·T mismatched base pair characterized in the present work is kinetically and energetically destabilized. The lifetime of the base pair in the closed state is greatly reduced relative to those of normal Watson–Crick base pairs (1 ms or less over the temperature range from 0 to 20 °C). This kinetic

destabilization is accompanied by an increase in the equilibrium constant for formation of the open state(s). For example, at 20 °C, ~0.1% of the mismatched base pairs are in the open, solvent-accessible state. The energetic basis of this destabilization consists, in part, of a lower enthalpy change for the formation of the open state.

The base pairs next to the mismatch site have shorter lifetimes and larger equilibrium constants for opening than the corresponding ones in the GC dodecamer. This local, kinetic and energetic destabilization does not result from changes in the activation energies or the enthalpies for formation of the open state. This suggests that the effects of the G·T mismatched base pair on adjacent base pairs have an entropic origin.

The presence of the mismatch also affects the energetics of opening of the central A·T base pairs, which are two bases removed from the mismatch site. The enthalpy change for opening of these base pairs is lower than in the GC dodecamer, whereas the corresponding free energy change is approximately the same. This result suggests that entropic effects compensate for the lower enthalpy change involved in the opening of these base pairs in the GT dodecamer.

REFERENCES

Benight, A. S., Schurr, J. M., Flynn, P. F., Reid, B. R., & Wemmer, D. E. (1988) *J. Mol. Biol. 200*, 377–399.

Braunlin, W. H., & Bloomfield, V. A. (1988) *Biochemistry 27*, 1184–1191.

Craik, D. J., Higgins, K. A., & Shehan, B. P. (1991) *J. Magn. Reson. 91*, 261–272.

Crick, F. H. C. (1966) *J. Mol. Biol. 19*, 548–555.

Englander, S. W., & Kallenbach, N. R. (1984) *Q. Rev. Biophys. 16*, 521–655.

Fersht, A. R., Knill-Jones, J. W., & Tsui, W.-C. (1982) *J. Mol. Biol. 156*, 37–51.

Gueron, M., Kochoyan, M., & Leroy, J.-L. (1987) *Nature (London) 328*, 89–92.

Gueron, M., Charretier, E., Hagerhorst, J., Kochoyan, M., Leroy, J.-L., & Moraillon, A. (1990) in *Structure and Methods* (Sarma, R. H., & Sarma, M. H., Eds.) Vol. 3, pp 113–137, Adenine Press, Guilderland, NY.

Hare, D., Shapiro, L., & Patel, D. J. (1986) *Biochemistry 25*, 7445–7456.

Hore, P. J. (1983) *J. Magn. Reson. 55*, 283–300.

Hunter, W. N., Brown, T., Kneale, G., Anand, N. N., Rabinovich, D., & Kennard, O. (1987) *J. Biol. Chem. 262*, 9962–9970.

Izatt, R. M., Christensen, J. J., & Rytting, J. H. (1971) *Chem. Rev. 71*, 439–481.

Johnson, M. L., & Frasier, S. G. (1985) *Methods Enzymol. 117*, 301–342.

Kalnik, M. W., Kouchakdjian, M., Li, B. F. L., Swann, P. F., & Patel, D. J. (1988) *Biochemistry 27*, 108–115.

Kennard, O. (1985) *J. Biomol. Struct. Dyn. 3*, 205–226.

Kennard, O. (1987) in *Nucleic Acids and Molecular Biology* (Eckstein, F., & Lilley, D. M. J., Eds.) Vol. 1, pp 25–52, Springer-Verlag, Berlin, Heidelberg.

Kochoyan, M., Leroy, J. L., & Gueron, M. (1990) *Biochemistry 29*, 4799–4805.

Leroy, J. L., Kochoyan, M., Huynh-Dinh, T., & Gueron, M. (1988a) *J. Mol. Biol. 200*, 223–238.

Leroy, J.-L., Charretier, E., Kochoyan, M., & Gueron, M. (1988b) *Biochemistry 27*, 8894–8898.

Leroy, J. L., Gao, X., Gueron, M., & Patel, D. J. (1991) *Biochemistry 30*, 5653–5661.

Modrich, P. (1987) *Annu. Rev. Biochem. 56*, 435–466.

Moe, J. G., & Russu, I. M. (1990) *Nucleic Acids Res. 18*, 821–827.

Pardi, A., Morden, K. M., Patel, D. J., & Tinoco, I., Jr. (1982) *Biochemistry 21*, 6567–6574.

Patel, D. J., Kozlowski, S. A., Marky, L. A., Rice, J. A., Broka, C., Dallas, J., Itakura, K., & Breslauer, K. J. (1982a) *Biochemistry 21*, 437–444.

Patel, D. J., Pardi, A., & Itakura, K. (1982b) *Science 216*, 581–590.

Patel, D. J., Kozlowski, S. A., Ikuta, S., & Itakura, K. (1984) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 43*, 2663–2670.

Patel, D. J., Shapiro, L., & Hare, D. (1987) in *Nucleic Acids and Molecular Biology* (Eckstein, F., & Lilley, D. M. J., Eds.) Vol. 1, pp 70–84, Springer-Verlag, Berlin, Heidelberg.

Petranovic, M., Petranovic, D., Dohet, C., Brooks, P., & Radman, M. (1990) *Nucleic Acids Res. 18*, 2159–2162.

Plum, G. E., & Bloomfield, V. A. (1990) *Biochemistry 29*, 5934–5940.

Ramstein, J., & Lavery, R. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 7231–7235.

Schurr, J. M., Benight, A. S., Wemmer, D. E., Flynn, P. F., & Reid, B. R. (1988) in *Ordering and Organization in Ionic Solutions* (Ise, N., & Sogami, I., Eds.) pp 212–222, World Scientific, Singapore.

Thielking, V., Alves, J., Fliess, A., Maass, G., & Pingoud, A. (1990) *Biochemistry 29*, 4682–4691.

Weast, R. C., Ed. (1986) *CRC Handbook of Chemistry and Physics*, 67th ed., CRC Press, Boca Raton, FL.

Registry No. 5′-d(CGCGAATTCGCG)-3′, 77889-82-8; 5′-d(CGCGAATTTGCG)-3′, 123270-12-2; guanine, 73-40-5; thymine, 65-71-4; adenine, 73-24-5; cytosine, 71-30-7.