Sequence Dependence of Base-Pair Opening in a DNA Dodecamer Containing the CACA/GTGT Sequence Motif†

Ewa Folta-Stogniew and Irina M. Russu*

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

Received May 2, 1994®

ABSTRACT: Proton nuclear magnetic resonance spectroscopy is used to characterize the kinetics and energetics of base-pair opening in two self-complementary DNA dodecamer duplexes: $[d(CGCACATGTGCG)]_2$ and $[d(CGCAGATCTGCG)]_2$. The first dodecamer contains two symmetrical CACA/GTGT motifs; in the second dodecamer, each motif is interrupted by a change of the central C·G base pair to a G·C base pair. The opening rates and the equilibrium constants for formation of the open state of each base pair are obtained from the dependence of the imino proton exchange rates on the concentration of ammonia catalyst. The results indicate that the opening rates of the central three base pairs in the CACA/GTGT motif are 3–8-fold larger than the corresponding ones in the CAGA/GTCT sequence. The activation enthalpies and entropies, and the standard enthalpy and entropy changes for formation of the open state, are obtained from the temperature dependence of the opening rates and equilibrium constants, respectively. The results reveal that enthalpy/entropy compensation exists, for all base pairs in both dodecamers, in activation as well as in the equilibria between closed and open states. As a result, the opening rates and equilibrium constants for opening are maintained, in both dodecamers, within a relatively narrow range of values. Nevertheless, large sequence-induced variations are observed for the activation enthalpies and the standard enthalpy changes for opening. The A·T base pair located between the C·G base pairs in the CACA/GTGT motif has a negative enthalpy change for formation of the activated state during opening. This is the first case in which a negative activation enthalpy is observed for opening of a Watson–Crick base pair in DNA. The activation enthalpies for opening of the C·G base pairs in the CACA/GTGT motif are increased by 5 and 9 kcal/mol relative to the corresponding ones in the CAGA/GTCT sequence. Sequence-induced variations are also observed for the standard enthalpy changes associated with the equilibria between closed and open states. For all base pairs in the CACA/GTGT motif, the standard enthalpy changes are 2–7 kcal/mol lower than the corresponding ones in the CAGA/GTCT sequence. Comparison of the activation enthalpies and standard enthalpy changes for opening suggests that a different opening pathway also exists for the central C·G base pair in the CACA/GTGT motif.

Alternating pyrimidine-purine sequences, $(CA)_n/(GT)_n$, are ubiquitous in eukaryotic genomes. They occur with high frequency in diverse genomes ranging from yeast to humans (Hamada et al., 1982). Their presence has been correlated to various biological functions. For example, they have been shown to be essential for the joining signals involved in variable (diversity) joining gene [V(D)J] recombination during immunoglobulin gene assembly (Hesse et al., 1989) and for transcriptional regulation of the mouse β-globin promoter (Cowie & Myers, 1988). The role of $(CA)_n/(GT)_n$ sequences in DNA–protein recognition has been demonstrated for the binding of catabolite gene activator protein to DNA (Schultz et al., 1991). The protein induces upon binding a 90° bend in the DNA, and this bend results almost entirely from two 40° kinks that occur between CA/GT base pairs. All these findings have enhanced the interest in defining the structural and dynamic features of $(CA)_n/(GT)_n$ sequences which may be responsible for their specific biological activities.

The existence of unique structural features at CA/GT sites has been first suggested by the ability of $(CA)_n/(GT)_n$ sequences to form left-handed Z-DNA in negatively supercoiled plasmids (Nordheim & Rich, 1983). DNA oligomers containing CA/GT repeats precisely positioned in the sequence also show abnormal gel retardation behavior. Their gel mobilities are higher than those predicted theoretically (Bolshoy et al., 1991) and do not decrease monotonically with DNA length (McNamara et al., 1990). These properties have been attributed to the formation of transient kinks in DNA at CA/GT sites (McNamara et al., 1990). CA/GT, CAC/GTG, and CACA/GTGT sequences have also been found to enhance DNA flexibility and to promote DNA bending in the complex of λ Cro protein with the O_R3 operator site (Lyubchenko et al., 1993). Previous relaxation studies by nuclear magnetic resonance (NMR)[1] spectroscopy have suggested that CAC/GTG sequences may act as molecular detents in the DNA structure (Lu et al., 1983; Cheung et al., 1984; Donlan & Lu, 1992).

High-resolution structures are presently available for several DNA fragments containing $(CA)_n/(GT)_n$ sequences. In one, the octamer $[d(GTGTACAC)]_2$ adopts an A-DNA conformation, and a novel interstrand bifurcated hydrogen bond is observed between the two purines in the CA/GT base pairs (Jain et al., 1989). In another, the CACA/GTGT sequence in $[d(ACCGGCGCCACA)]_2$ adopts, at low temperature, a frame-shifted structure with G·A and C·T interstrand hydrogen bonds (Timsit et al., 1991). Unique structural features

† This work was supported by the NSF (Grant 88-17589).

* Author to whom correspondence should be addressed.

® Abstract published in *Advance ACS Abstracts,* August 15, 1994.

[1] Abbreviations: NMR, nuclear magnetic resonance; 2D, two dimensional; NOESY, nuclear Overhauser effect spectroscopy; DQF-COSY, double quantum filtered correlated spectroscopy; T_1, longitudinal relaxation time.

0006-2960/94/0433-11016$04.50/0

Chart 1: Numbering of Bases in the Two Dodecamers Investigated[a]

```
          1  2  3  4  5  6  7  8  9 10 11 12
5'-d(C G C A X A T Y T G C G) - 3'
3'- (G C G T Y T A X A C G C)d- 5'
         12 11 10  9  8  7  6  5  4  3  2  1
```

[a] X = C and Y = G in the CACA/GTGT dodecamer. X = G and Y = C in the CAGA/GTCT dodecamer.

associated with the CAC/GTG sequence, e.g., base-pair tilt and displacement from the helix axis, have also been suggested by NMR studies of the [d(CGCACGCGCGTGCG)]$_2$ duplex (Patel et al., 1987).

In order to enhance our understanding of DNA dynamics at $(CA)_n/(GT)_n$ sequences, in the present work we have undertaken an investigation of the processes of base-pair opening in the dodecamer [d(CGCACATGTGCG)]$_2$. The dodecamer (henceforth abbreviated CACA/GTGT dodecamer) contains two symmetrically positioned CACA/GTGT motifs (Chart 1). In parallel, we have also characterized the processes of base-pair opening in the dodecamer [d(CGCAGATCTGCG)]$_2$ (henceforth abbreviated CAGA/GTCT dodecamer) in which the CACA/GTGT motif is interrupted by a simple change of the central C·G base pair to a G·C base pair.

EXPERIMENTAL PROCEDURES

Materials. The two dodecamers were synthesized using the solid-support phosphoramidite method and were purified by reverse-phase high-pressure liquid chromatography (in 50 mM ethylenediamine formate buffer at pH 7.5 with a gradient of 0–12.5% acetonitrile over 25 min). To exchange the counterion to sodium, the purified oligonucleotides were dialyzed extensively against 0.6 M NaCl. The DNA was then dialyzed with four exchanges against 2 mM ammonia buffer containing 100 mM NaCl and 2 mM EDTA in 90% H_2O/10% D_2O at pH 8.85 (at 20 °C). The dialysis buffer contained Chelex 100 (100–200 mesh, sodium form, Bio-Rad) to remove metal traces from the samples. Various concentrations of ammonia were obtained by titrating the samples with Chelex-treated stock solutions of either 0.4, 2, or 6 M ammonia buffer containing 100 mM NaCl and 2 mM EDTA at pH 8.85. The DNA concentration (duplex) was 5.1 mM for the CACA/GTGT dodecamer and 1.5 mM for the CAGA/GTCT dodecamer.

Methods. ^{1}H NMR experiments were carried out on a Varian VXR-400/54 NMR spectrometer operating at 400 MHz. All spectra were referenced to 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) as an external reference. The imino proton resonances were observed using the Jump-and-Return pulse (Plateau & Gueron, 1982). Longitudinal relaxation times (T_1) were measured by selective saturation of each imino proton followed by a variable recovery delay. For each measurement, 12 values for the recovery delay were used, and the T_1 values were calculated from exponential fits of intensity as a function of recovery delay. Phase-sensitive DQF-COSY spectra and NOESY spectra (at a mixing time of 100 ms) were obtained in 10 mM phosphate buffer in D_2O containing 100 mM NaCl and 2 mM EDTA at pH 7.0 as described (States et al., 1982; Piantini et al., 1982). Phase-sensitive NOESY spectra in aqueous ammonia solutions were obtained at a mixing time of 150 ms using a regular NOESY pulse sequence in which each pulse was replaced by a Jump-and-Return pulse (Otting et al., 1987; Michalczyk & Russu, 1993).

The longitudinal relaxation time (T_1) of each imino proton was measured as a function of the concentration of ammonia over the range from 0.002 to 1.5 M. The experimental T_1 values were fitted as a function of the inverse of the concentration of ammonia base, $[B]^{-1}$, to the equation (Moe & Russu, 1990):

$$T_1 = \frac{T_1^0(\tau_0 + D[B]^{-1})}{T_1^0 + \tau_0 + D[B]^{-1}} \quad (1)$$

where τ_0 is the lifetime of the base pair in the closed state (or the inverse of the rate constant for opening of the base pair, k_{op}); D is a constant defined as $(K_{op}k_B)^{-1}$ in which k_B is the catalytic rate constant for exchange of the imino proton from the open state of the base pair and K_{op} is the equilibrium constant for formation of the open state ($K_{op} = k_{op}/k_{cl}$, where k_{cl} is the rate constant for closing of the base pair); and T_1^0 is the longitudinal relaxation time corresponding to mechanisms other than exchange (e.g., proton–proton dipolar interactions). The fitted parameters were the base-pair lifetime, τ_0, the constant D, and the relaxation time, T_1^0, and their errors were taken as the standard deviations of the nonlinear least-squares fit (Johnson & Frasier, 1985). The concentration of ammonia base was calculated as the fraction $1/(1 + 10^{pK-pH})$ of the total ammonia concentration. The pH of the buffer was measured at each temperature of interest, and the change in the pK value of ammonia with temperature was calculated using a standard enthalpy of ionization of 12.4 kcal/mol (Weast, 1986).

The equilibrium constants for base-pair opening were obtained from the constant D as $K_{op} = (Dk_B)^{-1}$. The catalytic rate constant k_B for the open state of the base pair was reduced by a factor of 1.5 relative to its value in free nucleosides (namely, 2×10^8 M^{-1} s^{-1} for guanosine 2′,3′-cyclic monophosphate and thymidine; Gueron et al., 1990). This reduction corrects for the lower diffusion coefficient of the imino proton in DNA (Benight et al., 1988; Moe & Russu, 1992).

The activation enthalpy ($\Delta H^{\ddagger\circ}$) and activation entropy ($\Delta S^{\ddagger\circ}$) for base-pair opening were calculated on the basis of Eyring's absolute reaction rate theory from (Glasstone et al., 1941):

$$\ln(k_{op}/T) = \ln(\kappa k/h) + \Delta S^{\ddagger\circ}/R - \Delta H^{\ddagger\circ}/RT \quad (2)$$

where T is the absolute temperature, k is Boltzmann's constant, h is Planck's constant, and R is the gas constant. The transmission coefficient, κ, was assumed to have a value of 1.

The standard enthalpy change, ΔH°, and the standard entropy change, ΔS°, for base-pair opening were calculated from the Van't Hoff equation:

$$\ln(K_{op}) = -\Delta H^\circ/RT + \Delta S^\circ/R \quad (3)$$

The errors for ΔH° and ΔS°, and those for $\Delta H^{\ddagger\circ}$ and $\Delta S^{\ddagger\circ}$, represent standard deviations of the linear regression analysis for eqs 3 and 2, respectively.

RESULTS

The imino proton resonances of the two dodecamers are shown in Figure 1. Due to the symmetry of the base sequences, six resonances are expected for each dodecamer. At 15 °C, however, the imino proton resonance from the terminal C·G base pairs is not observed due to fraying at the ends of the duplexes. Assignments of the imino proton resonances in each dodecamer were obtained from NOESY experiments. The

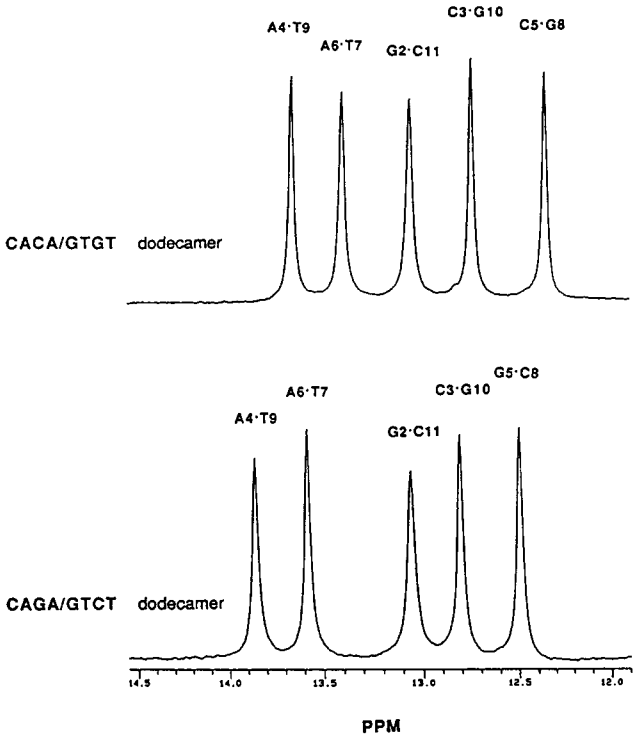

FIGURE 1: Imino proton resonances of the two dodecamers in 5 mM ammonia buffer containing 100 mM NaCl and 2 mM EDTA at pH 8.9 and at 15 °C. Assignments of resonances to individual base pairs are indicated according to the numbering of bases shown in Chart 1.

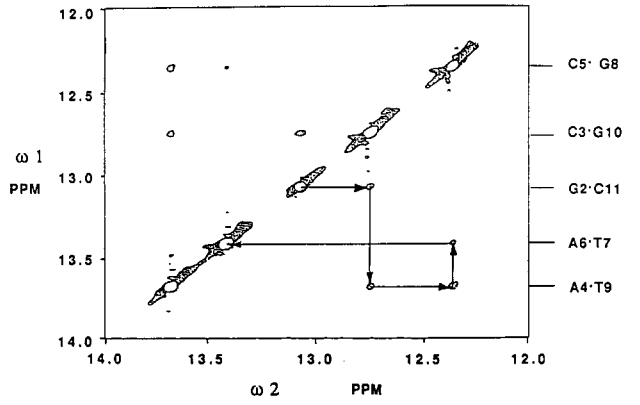

FIGURE 2: Imino proton resonance region of the NOESY spectrum of the CACA/GTGT dodecamer in 2 mM ammonia in 90% H_2O/10% D_2O with 100 mM NaCl and 2 mM EDTA at pH 8.9 and at 15 °C.

results are illustrated in Figure 2 for the CACA/GTGT dodecamer. As expected for a right-handed double-helical conformation (Boelens et al., 1985), all cross-peaks between imino protons in neighboring base pairs are observed. The connectivities can be traced from the G2·C11 to the A6·T7 base pair, as shown by arrows in Figure 2, and allowed sequential assignments of the corresponding protons. The assignments were confirmed by connectivities between thymine N3 and adenine C2 protons and guanine N1 and cytosine N4 protons (data not shown).

The processes of base-pair opening in the two dodecamers were characterized by measuring the selective T_1 values of the imino protons as a function of the concentration of ammonia catalyst. The dependence of T_1 values on ammonia base concentration is illustrated in Figure 3 for the imino proton in the A4·T9 base pair of the CACA/GTGT dodecamer. Lifetimes of the base pairs in the closed state, τ_0, were obtained by fitting experimental data, such as those in Figure 3, to eq 1. Representative values of the lifetimes are given in Table 1. At 15 °C, they range from 1 to 60 ms and are typical of lifetime values previously observed in a variety of DNA oligonucleotides (Leroy et al., 1988; Gueron et al., 1990; Moe & Russu, 1990, 1992).

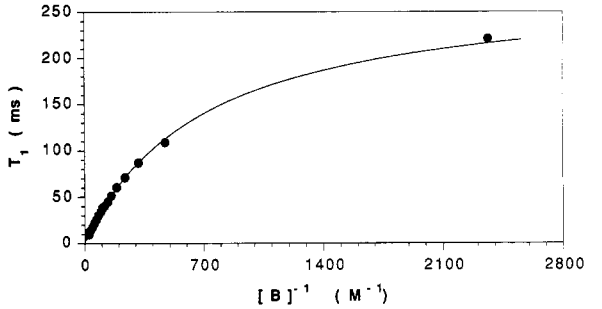

FIGURE 3: Dependence of the T_1 values of the imino proton in the A4·T9 base pair of the CACA/GTGT dodecamer on the inverse of the concentration of ammonia base at 15 °C. The curve corresponds to the nonlinear least-squares fit to eq 1.

Table 1: Base-Pair Lifetimes, τ_0 (ms), in the CACA/GTGT and CAGA/GTCT Dodecamers at 15 °C in Ammonia Buffer Containing 100 mM NaCl and 2 mM EDTA at pH 8.9

	base pair			
dodecamer	C3·G10	A4·T9	C5·G8/G5·C8	A6·T7
CACA/GTGT	12 ± 4	0.9 ± 0.7	24 ± 5	3 ± 1
CAGA/GTCT	12 ± 7	7 ± 1	62 ± 14	13 ± 1

Opening of the base pairs in the two dodecamers was investigated as a function of temperature, over the temperature range from 10 to 30 °C. The rationale for choosing this range of temperatures was 2-fold. First, this range is more that 30 °C lower than the melting temperatures of the two dodecamers [i.e., (76 ± 2) °C in 10 mM phosphate buffer containing 100 mM NaCl and 2 mM EDTA at pH 7.0]; thus, strand dissociation is not expected to contribute to the measured exchange rates of imino protons. Second, over this temperature range, the effects of fraying are confined to the first two base pairs at the ends of the duplexes and thus do not affect the opening of base pairs in the sequences of interest to this work, namely, CACA/GTGT and CAGA/GTCT. The temperature dependence of the opening rates ($k_{op} = 1/\tau_0$) for the base pairs in these sequences is shown in Figure 4. The corresponding activation parameters are given in Table 2. Similarly, Figure 5 and Table 3 give the temperature dependence of the equilibrium constants for opening, K_{op}, and the standard enthalpy and entropy changes for formation of the open state in each of these base pairs.

The solution conformation of the CACA/GTGT dodecamer was characterized qualitatively from NOESY and DQF-COSY experiments in 10 mM deuterated phosphate buffer containing 100 mM NaCl and 2 mM EDTA at pH 7.0. The use of phosphate buffer is advantageous for these experiments since it allows the residual water resonance to be greatly reduced by repeated lyophilization. This improves significantly the observation of sugar protons, such as H3′ and H4′. In contrast, in ammonia buffer, the large residual water proton resonance cannot be reduced by repeated lyophilization, and observation of some sugar protons is compromised. Assignments of all base and sugar protons (except H5′ and H5″) were obtained from NOESY experiments and confirmed, for scalar-coupled protons, from DQF-COSY experiments. The results are illustrated in Figure 6B,C which shows the H8/H6 to H2′/H2″ region of the NOESY spectrum and the H1′ to H2′/H2″ region of the DQF-COSY spectrum, respectively.

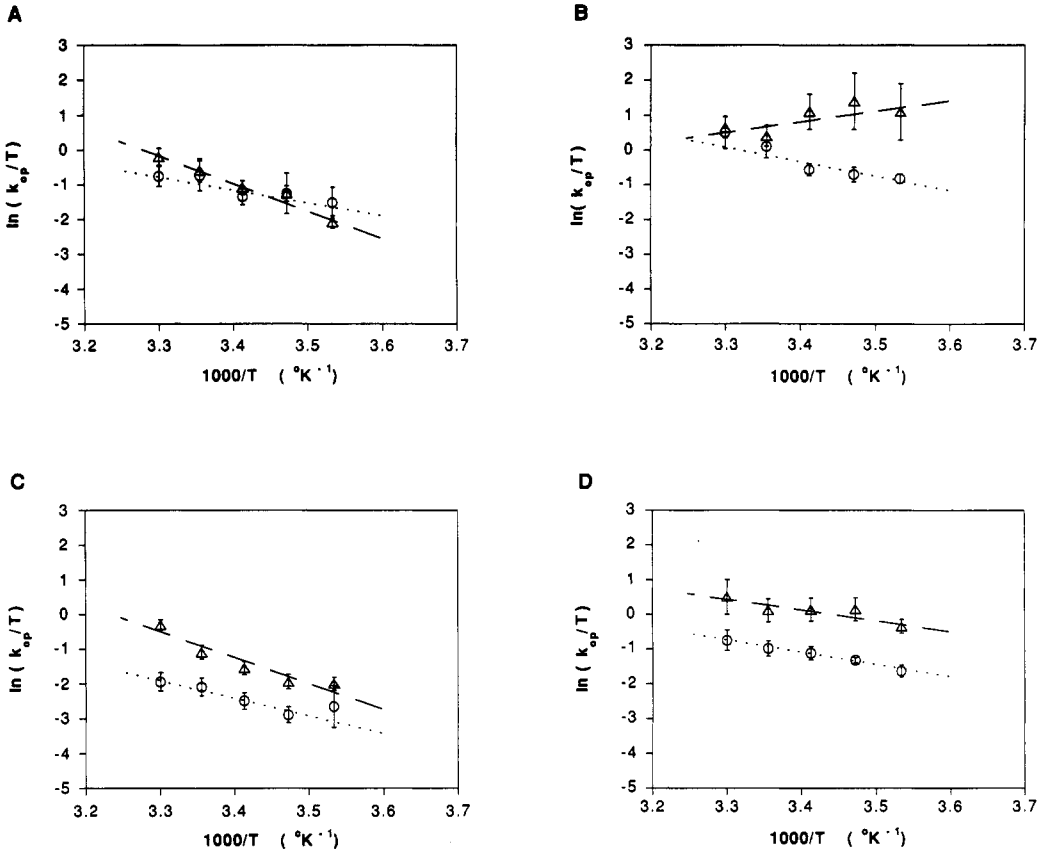

FIGURE 4: Temperature dependence of the opening rates for C3·G10 (panel A), A4·T9 (panel B), C5·G8/G5·C8 (panel C), and A6·T7 (panel D) base pairs in the CACA/GTGT dodecamer (triangles) and CAGA/GTCT dodecamer (circles).

Table 2: Activation Parameters, $\Delta H^{\ddagger\circ}$ (kcal/mol) and $\Delta S^{\ddagger\circ}$ [cal/(mol·K)], for Base-Pair Opening in the CACA/GTGT and CAGA/GTCT Dodecamers

	$\Delta H^{\ddagger\circ}$		$\Delta S^{\ddagger\circ}$	
base pair	CACA/GTGT dodecamer	CAGA/GTCT dodecamer	CACA/GTGT dodecamer	CAGA/GTCT dodecamer
C3·G10	16 ± 2	7 ± 4	5 ± 8	−25 ± 14
A4·T9	−6 ± 6	8 ± 2	−66 ± 19	−19 ± 9
C5·G8/G5·C8	15 ± 2	10 ± 3	1 ± 6	−18 ± 12
A6·T7	6 ± 3	7 ± 2	−26 ± 11	−25 ± 8

In order to establish the relevance of the 2D NMR structural analysis in phosphate buffer to our studies of base-pair opening, we have also obtained a NOESY spectrum of the CACA/GTGT dodecamer in the same solvent conditions as those used in the measurements of imino proton exchange (i.e., 2 mM ammonia buffer with 100 mM NaCl and 2 mM EDTA in 90% H_2O/10% D_2O at pH 8.9). The H8/H6 to H2′/H2″ region of this spectrum is shown in Figure 6A.

DISCUSSION

Solution Conformation of the CACA/GTGT Dodecamer. Our 2D NMR results indicate that the CACA/GTGT dodecamer adopts in solution a B-type conformation. For all bases (except possibly the terminal C1), the interresidual H8/H6–H2′ NOESY cross-peaks are less intense than the intraresidual H8/H6–H2′ cross-peaks. For example, in Figure 6B, the cross-peak of A6–H8 (8.20 ppm) to the H2′ proton in the same nucleotide (2.62 ppm) is stronger than that to the H2′ proton in the neighboring nucleotide (2.05 ppm). These results indicate that the interresidual distances H8/H6–H2′ are longer than the intraresidual ones, a property specific to the B-type DNA conformation (Van de Ven & Hilbers, 1988). Further evidence for the B-type conformation of the CACA/GTGT dodecamer is provided by the DQF-COSY spectra. We have measured the $^3J_{H1'-H2'}$ and $^3J_{H1'-H2''}$ coupling constants for each nucleotide in the dodecamer except G2 and G10 for which the DQF-COSY cross-peaks partially overlap (Figure 6C). For all (except G12), the $^3J_{H1'-H2''}$ values (8.6–9.6 Hz) were found to be larger than the $^3J_{H1'-H2'}$ values (5.1–6.0 Hz). Moreover, no H2″–H3′ cross-peaks were observed, except for the terminal C1 and G12. These two findings limit the deoxyribose pseudorotation angles, P, to values between 110° and 180° (Hosur et al., 1986). This range includes the C2′-endo conformation (P = 162°) characteristic of B-DNA but excludes the C3′-endo conformation (P = 18°) characteristic of A-DNA. Variations in the sugar conformation along the dodecamer were inferred from an analysis of the intensities of H3′–H4′ DQF-COSY cross-peaks. For C5, C11, and T9, the H3′–H4′ cross-peaks are very strong, indicating a value of P close to 126°, i.e., C1′-exo conformation of deoxyribose. On the other hand, for C3, A4, A6, and G8, the H3′–H4′ cross-peaks are slightly weaker, suggesting that the P values for these nucleotides range from 120° to 160°, i.e., between C1′-exo and C2′-endo conformation. For T7, no H3′–H4′ cross-peak is observed, consistent with a C2′-endo conformation ($P \sim 160°$). All these observations indicate that the nucleotides in the CACA/GTGT dodecamer (except for the terminal ones) adopt a sugar conformation in the range of the C2′-endo conformation of canonical B-DNA and that deviations toward the C1′-exo conformation exist for the C5 and T9 nucleotides in the CACA/GTGT sequence.

The B-type conformation of the CACA/GTGT dodecamer is maintained under the solvent conditions used for imino proton exchange measurements. As shown in Figure 6A, in

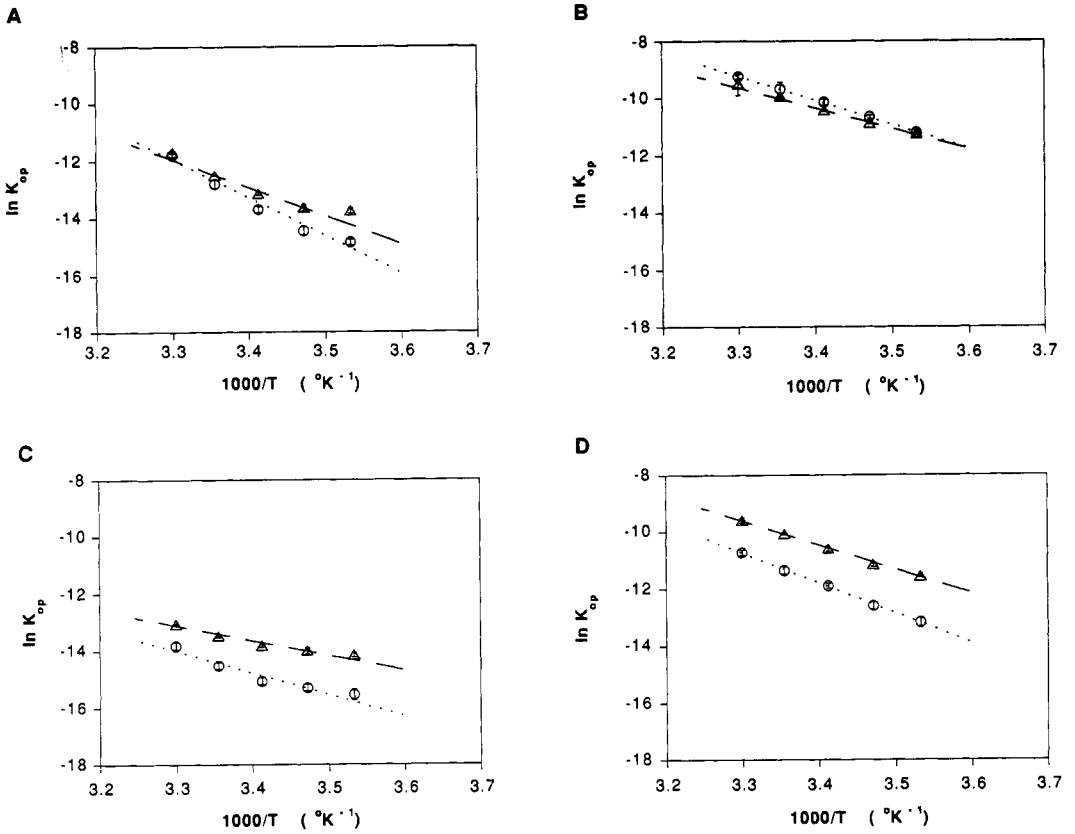

FIGURE 5: Temperature dependence of the equilibrium constants for opening for C3·G10 (panel A), A4·T9 (panel B), C5·G8/G5·C8 (panel C), and A6·T7 (panel D) base pairs in the CACA/GTGT dodecamer (triangles) and CAGA/GTCT dodecamer (circles).

Table 3: Thermodynamic Parameters, $\Delta H°$ (kcal/mol) and $\Delta S°$ [cal/(mol·K)], for Base-Pair Opening in the CACA/GTGT and CAGA/GTCT Dodecamers[a]

	$\Delta H°$		$\Delta S°$	
base pair	CACA/GTGT dodecamer	CAGA/GTGT dodecamer	CACA/GTGT dodecamer	CAGA/GTCT dodecamer
C3·G10	19.5 ± 0.7	26 ± 1	41 ± 2	63 ± 4
	16.9	*16.9*		
A4·T9	14.2 ± 0.6	16.6 ± 0.9	28 ± 2	37 ± 3
	12.3	*13.6*		
C5·G8/G5·C8	10.4 ± 0.6	17 ± 2	8 ± 2	28 ± 6
	12.3	*13.4*		
A6·T7	16.9 ± 0.6	21 ± 2	36 ± 2	47 ± 6
	14.4	*14.2*		

[a] The numbers in italics are the stacking enthalpies derived from thermal transition data (Breslauer et al., 1986).

ammonia buffer, the NOESY cross-peaks H8/H6–H2′/H2″ are very similar to those in phosphate buffer (Figure 6B). Hence, the nature of the buffer does not affect significantly the conformation of this DNA molecule, and in ammonia buffer, the dodecamer remains in a B-type conformation.

Kinetics of Base-Pair Opening. Inspection of the results in Table 1 shows that, at 15 °C, the lifetimes of the three central base pairs in the CACA/GTGT dodecamer are shorter than the corresponding ones in the CAGA/GTCT dodecamer. These changes correspond to 3–8-fold increases in the opening rates, indicating that the central base pairs in the CACA/GTGT sequence are kinetically destabilized. Further insight into the origin of this sequence-induced destabilization is provided by the dependence of the opening rates on temperature (Figure 4A–D). Unusual behavior is exhibited by the A·T base pair in the fourth position of the CACA/GTGT dodecamer (A4·T9): the opening rate for this base pair decreases upon increasing the temperature from 10 to 30 °C. This temperature dependence corresponds to a negative value of the activation enthalpy (Table 2). In contrast, when placed in the sequence context CAGA/GTCT, the same A·T base pair has a positive activation enthalpy for opening. Positive activation enthalpies are also observed for all the other base pairs in the two dodecamers (Table 2). Moreover, for all DNA oligonucleotides studied thus far by this and other laboratories, the activation enthalpies for base-pair opening have been found to be positive (Leroy et al., 1988; Moe & Russu, 1992). Thus, the kinetics of opening of the A·T base pair situated between C·G base pairs in the sequence CACA/GTGT is unique. Nevertheless, the kinetic behavior of this base pair is similar to that of a mismatched base pair, such as the G·T base pair previously investigated by this laboratory. We have found that, in the duplex $[d(CGCGAATTTGCG)]_2$, the opening rates of the two mismatched G·T pairs are of the order of 10^3 s^{-1} and do not change significantly upon a change in temperature from 0 to 20 °C (Moe & Russu, 1992). This finding indicated that the activation enthalpy for opening of mismatched G·T base pairs is close to zero and, as a result, these pairs are kinetically destabilized. The results obtained in the present work suggest that the activated state for opening of the A4·T9 base pair in the CACA/GTGT dodecamer is similar to that involved in opening mismatched G·T pairs. In both cases, the enthalpy of the activated state is the same as, or lower than, that of the closed state, and the dominant energetic contribution to the formation of the activated state is entropic in origin.

The presence of the CACA/GTGT motif does not affect significantly the activation parameters for opening of the A·T base pair at the 3′-end of this sequence (Table 2). The opening rates for the A6·T7 base pair are increased up to ~4-fold in the CACA/GTGT dodecamer as compared to the CAGA/

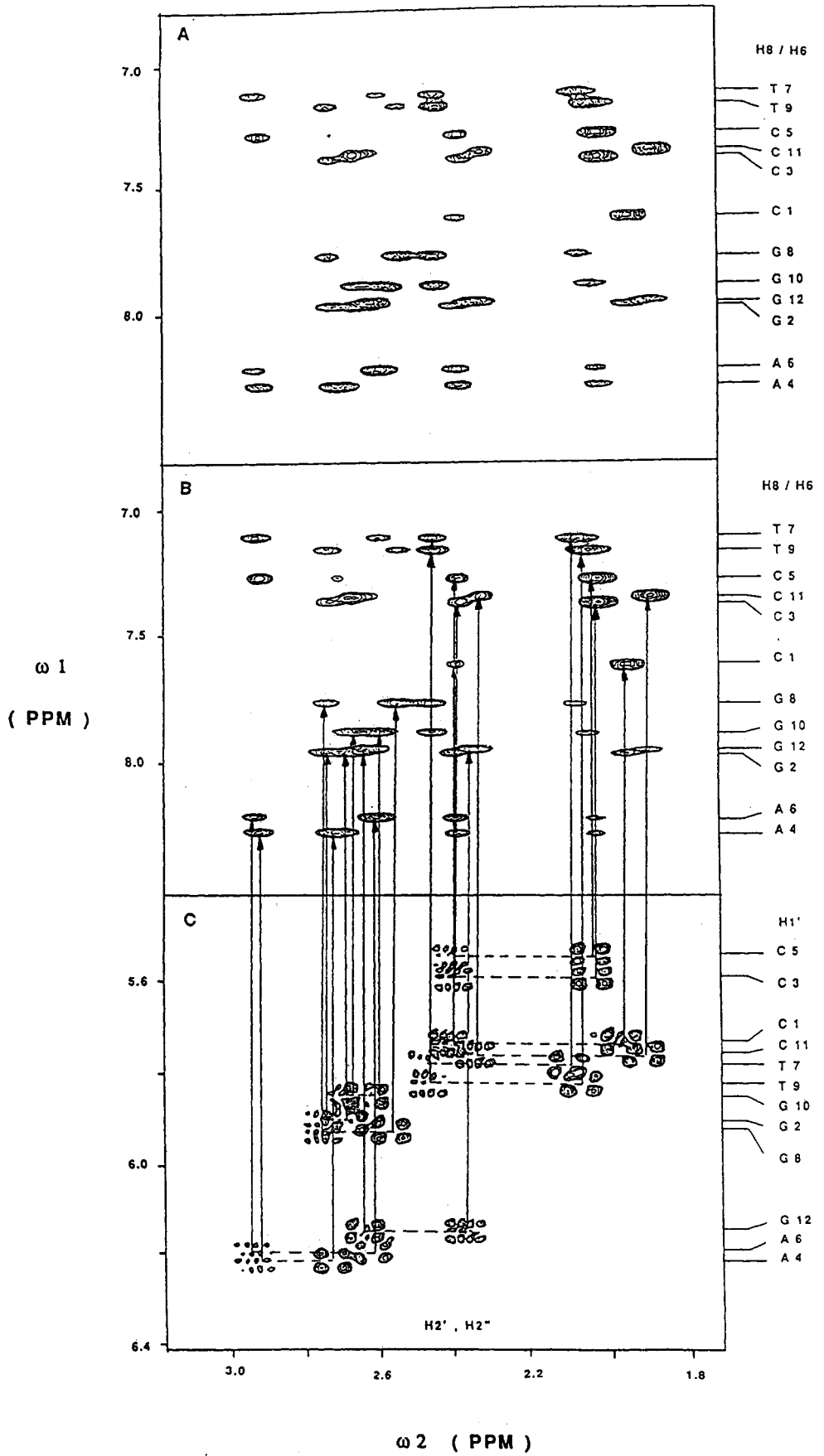

FIGURE 6: Expanded regions of the NOESY and DQF-COSY spectra of the CACA/GTGT dodecamer at 15 °C. (A) H8/H6 to H2′/H2″ region of the NOESY spectrum in 2 mM ammonia in 90% H_2O/10% D_2O with 100 mM NaCl and 2 mM EDTA at pH 8.9. (B) H8/H6 to H2′/H2″ region of the NOESY spectrum in 10 mM phosphate buffer in D_2O with 100 mM NaCl and 2 mM EDTA at pH 7.0. (C) H1′ to H2′/H2″ region of the DQF-COSY spectrum in 10 mM phosphate buffer in D_2O with 100 mM NaCl and 2 mM EDTA at pH 7.0. DQF-COSY cross-peaks H1′–H2′ and H1′–H2″ are connected by horizontal dashed lines; the vertical lines connect these DQF-COSY cross-peaks to the NOESY cross-peaks H8/H6–H2′ and H8/H6–H2″, respectively, within the same nucleotide.

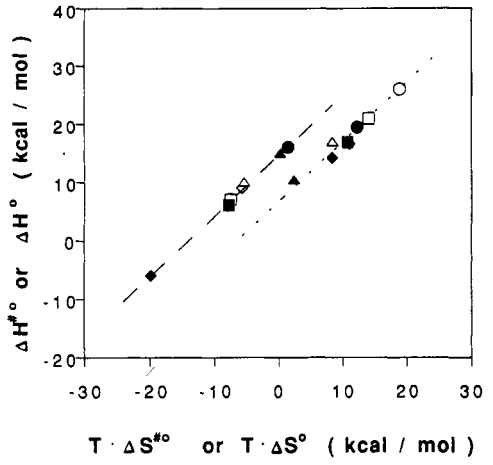

FIGURE 7: Enthalpy–entropy compensation for base-pair opening in the two dodecamers investigated. Dashed line: $\Delta H^{\ddagger\circ}$ versus $T\Delta S^{\ddagger\circ}$. Dotted line: ΔH° versus $T\Delta S^{\circ}$ (T = 300 K). Circles, C3·G10; diamond, A4·T9; triangles, C5·G8 or G5·C8; squares, A6·T7. Filled symbols are for the CACA/GTGT dodecamer, and open symbols are for the CAGA/GTCT dodecamer.

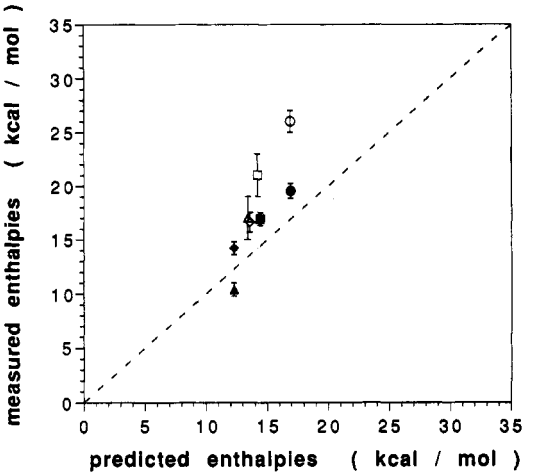

FIGURE 8: Comparison of the measured standard enthalpy changes for opening to the stacking enthalpies derived from thermal transition data (Breslauer et al., 1986). Circles, C3·G10; diamond, A4·T9; triangles, C5·G8 or G5·C8; squares, A6·T7. Filled symbols are for the CACA/GTGT dodecamer, and open symbols are for the CAGA/GTCT dodecamer.

GTCT dodecamer (Table 1). These differences correspond to changes in the activation parameters of less than 1 kcal/mol and thus fall within the experimental uncertainties for determination of activation enthalpies and entropies (Table 2).

Significant variations in the kinetics of opening are observed for the two C·G base pairs in the CACA/GTGT sequence. As shown in Table 2, for both base pairs, the activation enthalpies in the CACA/GTGT dodecamer are significantly larger than those in the CAGA/GTCT dodecamer. As for the A4·T9 base pair, these values suggest that the activated states during opening of the two C·G base pairs in the CACA/GTGT sequence are different from those in the CAGA/GTCT sequence. Thus, it appears that the whole CAC/GTG triplet exhibits unique kinetics for base-pair opening.

The larger activation enthalpies for opening of C·G base pairs in the CACA/GTGT dodecamer may be expected to decrease substantially the opening rates for these base pairs. As shown in Table 1 and in Figure 4A,C, this is not the case: over the temperature range investigated, the opening rates for the C·G base pairs in the two dodecamers differ, at most, by a factor of 5. This effect can be readily understood by noting that the increases in activation enthalpies are almost completely compensated by increases in activation entropies (Table 2). For all base pairs investigated, the activation enthalpies are linearly related to the activation entropies (Figure 7, dashed line). This includes the A4·T9 base pair in the CACA/GTGT dodecamer which, as a result of its unique kinetic behavior (i.e., $\Delta H^{\ddagger\circ} < 0$), falls at the extreme of lowest activation enthalpy/entropy on the graph.

Equilibria between Open and Closed States of Base Pairs. The data summarized in Figure 5 indicate that, for some base pairs in the CACA/GTGT motif (i.e., C3·G10, C5·G8, and A6·T7), the equilibrium constants for opening are larger than the corresponding ones in the CAGA/GTCT sequence. These differences are temperature dependent and are enhanced over the lower range of temperature investigated. The A4·T9 base pair is an exception to this trend, and its equilibrium constants for opening in the two dodecamers are similar (Figure 5B).

The standard enthalpy changes for opening of all base pairs in the CACA/GTGT motif are lower than those in the CAGA/GTCT sequence (Table 3). This enthalpic destabilization may originate from changes in base stacking interactions in the two molecules. To address this possibility, we have compared the standard enthalpy changes measured in the present work to the stacking enthalpies predicted by the near-neighbor analysis of thermal transition data (Breslauer et al., 1986). The latter are included in italics in Table 3 and are compared to our experimental values in Figure 8. Clearly, the measured values of the standard enthalpy changes show an overall, qualitative correlation to the stacking enthalpy changes. However, for most base pairs, the experimental ΔH° values exceed the theoretically predicted values. Moreover, the differences in the measured ΔH° values between the two dodecamers (Table 3) are larger than the differences predicted by simple stacking energies. This observation is best illustrated by the results for the C3·G10 base pair. This base pair is two positions removed from the site of the C·G → G·C change and, thus, should not be affected by differences in the near-neighbor stacking energies. In contrast, the measured standard enthalpy changes for opening of this base pair differ in the two dodecamers by more than 6 kcal/mol. These results add further support to the suggestion that the open state of the base pair monitored by imino proton exchange is not the same as the state observed in a melting transition (Benight et al., 1988). Clearly, the variations in the energetics of base-pair opening cannot be fully explained, at least in the two sequences investigated here, by simple stacking interaction rules and probably reflect more subtle alterations induced in DNA by the base sequence.

Comparison of the activation enthalpies with the standard enthalpy changes for opening (Tables 2 and 3) shows that, for the majority of base pairs investigated, the standard enthalpy changes are larger than the activation enthalpies. Thus, for most base pairs, transition from the activated to the open state involves a further increase in enthalpy. Depending on the exact location of the base pair in the sequence, this increase ranges from 3 to 20 kcal/mol. The only exception to this trend is the C5·G8 base pair in the CACA/GTGT dodecamer. For this base pair, the transition from the activated to the open state is accompanied by a decrease in enthalpy (~5 kcal/mol). This observation supports our suggestion that the nature of the activated/open state for the central C·G base pair in the CACA/GTGT motif differs from that of other base pairs.

The standard enthalpy changes for opening show, in both dodecamers, a significant dependence on the sequence context

of the base pair (Table 3). This dependence parallels that of the standard entropy changes, and enthalpy/entropy compensation is observed for the equilibria between open and closed forms (Figure 7, dotted line). As for the activation process, the entropy changes during opening compensate, in part, the large variations in standard enthalpies between various base pairs and between the two dodecamers and, thus, reduce the net differences in equilibrium constants for the open state. Hence, the presence of enthalpy/entropy compensation, in the activation steps as well as in the equilibria between open and closed states of the base pairs, attenuates the variations in the corresponding opening parameters such that these variations may be overlooked experimentally. On the other hand, as we have shown in the present work, significant sequence-induced variations in base-pair opening can be detected and analyzed by monitoring the temperature dependence of the base-pair opening processes.

Longitudinal Relaxation of Imino Protons in the CACA/GTGT Dodecamer. The NMR relaxation properties of imino protons in the CAC/GTG sequence have been previously investigated by Lu and co-workers for a variety of DNA oligonucleotides including the lactose operon operator of *Escherichia coli* and several of its fragments, a sequence from the mouse heavy chain immunoglobulin enhancer and a sequence from the critical core of the SV40 enhancer (Lu et al., 1983; Cheung et al., 1984; Donlan & Lu, 1992). For all these DNA fragments, Lu and co-workers have found that the T_1^{-1} rate of the imino proton in the A·T base pair of the CAC/GTG triplet is enhanced relative to the rates of the other imino protons in the molecule. These results have demonstrated the existence of a dynamic heterogeneity at the CAC/GTG sequence. The origin of this unique relaxation behavior has not been, however, fully understood. One limitation arised from the fact that the measurements of Lu and co-workers were carried out in the presence of phosphate buffer which is not an efficient catalyst for imino proton exchange (Gueron et al., 1990). Thus, the observed variations in the T_1^{-1} rates could not be directly related to the kinetics of base-pair opening (Donlan & Lu, 1992).

In the present work, we have confirmed that the same relaxation behavior is present in the CACA/GTGT dodecamer (Figure 9A) and is absent in the CAGA/GTCT dodecamer (Figure 9B). The selective T_1^{-1} rates of the imino proton in the A4·T9 base pair of the CACA/GTGT dodecamer, measured in phosphate buffer, are consistently higher than those of the other imino protons in the molecule. Over the temperature range from 10 to 30 °C, these T_1^{-1} values are within experimental error of the $(T_1^0)^{-1}$ values obtained from ammonia-catalyzed exchange experiments (eq 1).

Imino proton relaxation in phosphate buffer (or in the presence of low concentrations of ammonia) is expected to result from proton–proton dipolar interactions and from exchange catalyzed by a proton acceptor intrinsic to the DNA (such as an acceptor on the other base of the open base pair; Gueron et al., 1987). The latter process (also termed exchange in the absence of added catalyst) involves the same open state of the base pair as that detected in catalyst-dependent exchange. Thus, the two processes, i.e., catalyst-dependent exchange and exchange in the absence of added catalyst, should be characterized by the same kinetic parameters for base-pair opening. The results obtained in the present work indicate that, in the CACA/GTGT dodecamer at temperatures higher than 30 °C, the opening rate of the A4·T9 base pair should be increasingly lower than that of the A6·T7 base pair (Figure 4B,D; Tables 1 and 2). This temperature dependence is

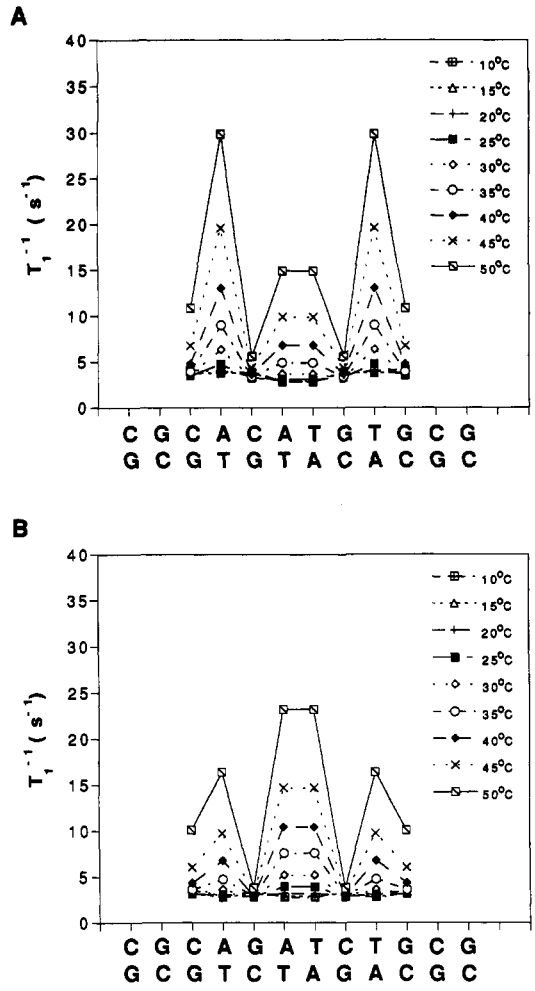

FIGURE 9: Temperature dependence of the T_1^{-1} rates of imino protons in the CACA/GTGT dodecamer (panel A) and CAGA/GCTC dodecamer (panel B) in 10 mM phosphate buffer with 100 mM NaCl and 2 mM EDTA at pH 7.0.

opposite to that observed for the T_1^{-1} rates and suggests that the faster relaxation of the A4·T9 imino proton in the absence of added catalyst (Figure 9A) does not reflect an enhancement in the opening rates for this base pair. It is more likely that the larger relaxation rates associated with the A4·T9 base pair originate from unique structural features and/or motional properties at this site of the CACA/GTGT motif.

CONCLUSIONS

The results obtained in the present work demonstrate that unique features exist for the opening of base pairs in the CAC/GTG triplet of the CACA/GTGT motif. Most unusual is the behavior of the central A·T base pair in the triplet. In contrast to all other DNA Watson–Crick base pairs investigated thus far, formation of the activated state during opening of this base pair is enthalpically favored. This property could contribute to the ability of A·T base pairs in CAC/GTG contexts to enhance DNA flexibility. On the other hand, for the two C·G base pairs in the CAC/GTG triplet, formation of the activated state involves larger enthalpy changes than those observed for the C·G base pairs in the CAG/GTC sequence. Moreover, for the C·G base pair at the 3′-end of the CAC/GTG triplet, the transition from the activated to the open state is accompanied by a decrease in enthalpy. This behavior is different from those of all other base pairs in the two dodecamers investigated.

The sequence-induced variations in the kinetics and energetics of base-pair opening in the CACA/GTGT motif do not

appear to result from gross conformational changes in this sequence. As we have shown, the overall solution structure of the CACA/GTGT dodecamer corresponds to a B-type conformation. It is much more likely that the observed variations in base-pair opening reflect subtle structural features associated with the CACA/GTGT motif. Such features, and their influence on the pathway(s) and energetics of base-pair opening, are currently being investigated in our laboratory by NMR and molecular dynamics methods.

REFERENCES

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

Boelens, R., Scheek, R. M., Dijkstra, K., & Kaptein, R. (1985) *J. Magn. Reson. 62*, 378–386.

Bolshoy, A., McNamara, P., Harrington, R. E., & Trifonov, E. N. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 2312–2316.

Breslauer, K. J., Frank, R., Blocker, H., & Marky, L. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3746–3750.

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

Cowie, A., & Myers, R. M. (1988) *Mol. Cell. Biol. 8*, 3122–3128.

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

Glasstone, S., Laidler, K. J., & Eyring, H. (1941) *The theory of rate processes. The kinetics of chemical reactions, viscosity, diffusion and electrochemical phenomena*, McGraw-Hill, New York and London.

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

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

Hamada, H., Petrino, M. G., & Kakunaga, T. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 6465–6469.

Hesse, J. E., Lieber, M. R., Mizuuchi, K., & Gellert, M. (1989) *Genes Dev. 3*, 1053–1061.

Hosur, R. V., Ravikumar, M., Chary, K. V. R., Sheth, A., Govil, G., Zu-Kun, T., & Miles, H. T. (1986) *FEBS Lett. 205*, 71–76.

Jain, S., Zon, G., & Sundaralingam, M. (1989) *Biochemistry 28*, 2360–2364.

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

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

Lu, P., Cheung, S., & Arndt, K. (1983) *J. Biomol. Struct. Dyn. 1*, 509–521.

Lyubchenko, Y. L., Shlyakhtenko, L. S., Appella, E., & Harrington, R. E. (1993) *Biochemistry 32*, 4121–4127.

McNamara, P. T., Bolshoy, A., Trifonov, E. N., & Harrington, R. E. (1990) *J. Biomol. Struct. Dyn. 8*, 529–538.

Michalczyk, R., & Russu, I. M. (1993) *FEBS Lett. 331*, 217–222.

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

Moe, J. G., & Russu, I. M. (1992) *Biochemistry 31*, 8421–8428.

Nordheim, A., & Rich, A. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 1821–1825.

Otting, G., Grutter, R., Leupin, W., Minganti, C., Ganesh, K. N., Sproat, B. S., Gait, M. J., & Wuthrich, K. (1987) *Eur. J. Biochem. 166*, 215–220.

Patel, D. J., Shapiro, L., & Hare, D. (1987) in *Unusual DNA Structures* (Wells, R. D., & Harvey, S. C., Eds.) pp 115–161, Springer-Verlag, New York.

Piantini, U., Sorensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800–6801.

Plateau, P., & Gueron, M. (1982) *J. Am. Chem. Soc. 104*, 7310–7311.

Schultz, S. C., Shields, G. C., & Steitz, T. A. (1991) *Science 253*, 1001–1007.

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

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

Van de Ven, F. J. M., & Hilbers, C. W. (1988) *Eur. J. Biochem. 178*, 1–38.

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