Ruf, W., Rehemtulla, A., & Edgington, T. S. (1991) J. Biol. Chem. 266, 2158–2166.
Schechter, I., & Berger, A. (1967) Biochem. Biophys. Res. Commun. 27, 157.
White, W. F., & Barlow, G. H. (1970) Methods Enzymol. 19, 665–672.
Zur, M., & Nemerson, Y. (1978) J. Biol. Chem. 253, 2203–2209.

Conformational Properties of the G·G Mismatch in d(CGCGAATTGGCG)$_2$ Determined by NMR†

Katherine L. B. Borden,‡ Terence C. Jenkins,§ Jane V. Skelly,§ Tom Brown,∥ and Andrew N. Lane*,‡

Laboratory for Molecular Structure, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, U.K., Cancer Research Campaign Biomolecular Structure Unit, Institute of Cancer Research, Sutton, Surrey SM2 5NG, U.K., and Department of Chemistry, University of Edinburgh, Edinburgh EH9 3JJ, U.K.

Received October 28, 1991; Revised Manuscript Received March 18, 1992

ABSTRACT: The conformational properties of the DNA duplex d(CGCGAATTGGCG)$_2$, which contains two noncomplementary G·G base pairs, have been examined in aqueous solution by ^{1}H and ^{31}P NMR as a function of temperature. The G·G mismatch is highly destabilizing, with a T_m value 35 K below that observed for the native *Eco*RI dodecamer. The dodecamer appears symmetric in the NMR spectra and exists largely as an average B-type DNA conformation. However, the ^{1}H and ^{31}P NMR spectra give evidence of considerable conformational heterogeneity at the mismatched nucleotides and their nearest neighbors, which increases with increasing temperature. There is no evidence for a significant population of the syn purine conformation. The imino protons of the mispaired bases G4 and G9 are degenerate, resonate at high field, and exchange readily with solvent. These results indicate that the mispaired bases are only weakly hydrogen-bonded and are only partially stacked into the helix. On raising the temperature, the duplex shows increasing exchange between two or more conformations originating from the mismatch sites. However, these additional conformations maintain their Watson–Crick hydrogen bonding. The increase in chemical exchange is consistent with a quasimelting process for which the G·G sites provide local nuclei. Extensive modeling studies by dynamic annealing have confirmed that the G(anti)·G(anti) conformation is favored and that the mispairs are poorly stacked within the helix. The results explain both the poor thermal stability and low hypochromicity of this duplex.

The limited fidelity of DNA polymerase is responsible for the relatively frequent [frequency of occurrence is $\sim 10^{-4}$; see review by Loeb and Kunkel (1982)] formation of noncomplementary base pairs during DNA synthesis. Such mismatches provide intermediates in the mutagenic pathway. To maintain the integrity of the genetic information mistakes that occur during replication must be repaired with high efficiency. The overall fidelity of replication is $\sim 10^{-10}$ errors per nucleotide (Fowler et al., 1974). Of the possible base mispairings, purine·purine and purine·pyrimidine pairs occur more frequently than pyrimidine·pyrimidine pairs (Topal & Fresco, 1976; Fersht & Knill-Jones, 1981). If uncorrected, purine·purine mismatches lead to transversions in the daughter strands, e.g., A·T → A·A$_{syn}$ → T·A; G·C → G·G$_{syn}$ → C·G (Topal & Fresco, 1976). Cellular repair mechanisms require effective recognition of the mismatch sites prior to base correction. Three major factors affect the detection of a mismatch base pair: (i) the structure of the mispairing, (ii) the nature of the adjacent base sequences, and (iii) the conformation of the DNA in the vicinity of the lesion. The relative repair efficiencies for base-pair mismatches are in the order A·C, G·G, G·T > A·A > A·G, C·C, C·T, T·T (Kramer et al., 1984). Further, the A·G repair efficiency is very sensitive to the base composition of the surrounding sequence, in particular the A·T versus G·C content, which is mirrored by the sensitivity of the mismatch conformation to the flanking sequences (Kramer et al., 1984; Dohet et al., 1985). There are also differences in repair efficiency within the purine·purine class, with G·G > A·A > A·G. Although the chemical structures of these mismatches are different, the ability of the cellular repair enzymes to detect these errors may also depend on the different conformations available to the DNA, for example, the possibility of hydrogen-bonding patterns in A·G pairings that are not available in G·G mismatches.

In the class of purine·purine mismatches, A·G mispairs have been studied extensively by both NMR and X-ray crystallography (Gao & Patel, 1988; Brown et al., 1990; Webster et al., 1990; Carbonnaux et al., 1991; Lane et al., 1991b), and NMR studies have also been reported for A·A (Arnold et al., 1987), AP·A (Fazakerley et al., 1986), and A·I mismatches (Uesugi et al., 1987). However, only comparatively little is known about the G·G mispairing. ^{31}P NMR studies (Roongta et al., 1990) and preliminary ^{1}H NMR experiments (Piotto & Gorenstein, 1991) have recently been reported for a G·G-mismatched dodecamer duplex. Topal and Fresco (1976) suggested that the G·G mismatch would pair using one base in the keto form and syn and the other in the enol form and anti. However, Cognet et al. (1991) have presented evidence

† This work was supported by the Cancer Research Campaign, the Medical Research Council, and the Science and Engineering Research Council of the U.K.
‡ National Institute for Medical Research.
§ Institute of Cancer Research.
∥ University of Edinburgh.

0006-2960/92/0431-5411$03.00/0

for a G(anti)·G(syn) mismatch in a non-self-complementary undecamer, in which both bases are in the keto form and the GN1 protons hydrogen bond with the GO6 carbonyl, forcing a large C1′–C1′ separation. Further, even assuming the normal tautomeric forms of the bases, there are several possible pairing schemes for G·G mismatches (Saenger, 1984; Cognet et al., 1991). It is also possible that the mismatch does not pair in a conventional sense, further complicating the conformational analysis.

We have undertaken ^{1}H and ^{31}P NMR solution studies and molecular modeling to determine the conformational properties of a G·G mismatch in the duplex d(CGCGAATTGGCG)$_2$ related to the native *Eco*RI 12-mer sequence, d-(CGCGAATTCGCG)$_2$, which has been extensively studied [e.g., Dickerson and Drew (1981), Fratini et al. (1982), Lane et al. (1991a), and Nerdal et al. (1988)].

Materials and Methods

Materials

The oligodeoxynucleotide dodecamers d-(CGCGAATTCGCG) and d(CGCGAATTGGCG) were prepared on an Applied Biosystems synthesizer using phosphoramidite chemistry and purified by serial reverse-phase (C8) and ion-exchange HPLC, as described previously (Leonard et al., 1990). For the NMR studies, the G·G-mispaired DNA was lyophilized and redissolved in 1 mL of 10 mM sodium phosphate, 100 mM KCl, and 0.5 mM EDTA. The DNA was annealed by heating to 355 K for 5 min and slow cooling to ambient temperature during several hours. The duplex DNA was then relyophilized and finally dissolved in 99.996% ^2H_2O for NMR studies. The concentration was approximately 0.8 mM in strands, pH 7.2 (pD*, uncorrected), and the concentration of salt was 20 mM sodium phosphate, 200 mM potassium chloride, and 1 mM EDTA. 4,4-Dimethylsilapentane-1-sulfonate (DSS, 0.1 mM) was used as an internal ^{1}H chemical shift reference.

Methods

Ultraviolet Denaturation Studies. Melting curves were measured at 260 nm on a Varian-Cary 219 spectrophotometer equipped with a Neslab ETP3/RTE4 circulating water heating/cooling accessory. Heating was applied at 1 K min^{-1} from 278 K until thermal denaturation was complete, as judged from the increase in optical absorption (A_{260}). The crude data were collected and processed as described previously (Jones et al., 1990). The oligonucleotides were dissolved in a buffer (pH 7.00) consisting of aqueous KCl (0.1 or 1.0 M), 10 mM sodium phosphate, and 0.5 mM EDTA to give a working solution 10 μM in strands. The concentration of the 12-mers is based on an extinction coefficient of 1.15×10^5 M^{-1} cm^{-1} in strands at 363 K and 260 nm for the native *Eco*RI dodecamer (Patel et al., 1982). All solutions were heated to 365 K for 15 min, then cooled to 295 K during 6 h, and finally cooled to 278 K during a further 12 h to ensure efficient annealing of the dodecamer duplexes.

Thermal denaturation temperatures (T_m) were determined at a relative absorbance change of 0.50 and are reported as the mean ± SEM of at least four determinations.

NMR Spectroscopy. One-dimensional NMR spectra were collected in ^2H_2O and ^1H_2O at 9.4 and 11.7 T, respectively, on Bruker AM spectrometers and at 14.1 T on a Varian Unity spectrometer at temperatures specified in the text. One-dimensional spectra in ^2H_2O were recorded using an acquisition time of 1.365 s, with presaturation of the residual solvent peak. Driven, truncated NOEs[1] were collected at 281, 298, and 303 K using the method of Wagner and Wüthrich (1979) with a recycle time of 7 s to ensure sufficient relaxation time for the A(H2) protons. Irradiation times ranged from 50 to 800 ms. Rotational correlation times were determined from cross-relaxation rate constants for the C(H5–H6) vectors (Lane et al., 1986). The cross-relaxation rate constant, σ, was determined by nonlinear regression to

$$\mathrm{NOE}(t) = \sigma/\rho[1 - \exp(-\rho/t)] \quad (1)$$

where ρ is the apparent spin–lattice relaxation constant and t is the irradiation time; at least four time points were recorded for each temperature. The apparent correlation time (τ_c) was calculated from

$$\sigma = \alpha/r^6[[6/(1 + \omega^2\tau^2)] - 1]\tau_c \quad (2)$$

where the C(H5)–C(H6) distance, r, is taken to be 0.246 nm and α is 56.92 Å^6 ns^{-2}.

The A5(H2)–A6(H2) separation was calculated from the cross-relaxation rate constant (eq 2). The appropriate correlation time is equal to that for end-over-end tumbling of the duplex, τ_L, which can be calculated from the correlation time of the cytosine vectors as described previously (Birchall & Lane, 1990; Lane et al., 1991a); thus, for the 12-mer, $\tau_L \approx 1.3\tau_c$.

Phase-sensitive NOESY spectra were obtained in ^2H_2O at 281, 288, 298, and 303 K using the time-proportional phase incrementation scheme (Marion & Wüthrich, 1983). A total of 512 t_1 increments of 2048 data points were collected with an acquisition time of 169 ms and a recycle time of 2 s. The mixing times were 200 ms at 303 and 298 K, 100 ms at 288 K, and 90 ms at 281 K. Sine-modulated data were recorded with proper setting of the receiver phase and dead-time to produce a flat base-plane (Frenkiel et al., 1990). The ROESY spectrum was acquired in a similar manner, using unmodulated output from the decoupler to supply the spin-lock field (Bauer et al., 1990, Bothner-By et al., 1984), which was 2 kHz. A phase-sensitive NOESY spectrum was also recorded at 14.1 T and 288 K using the method of States et al. (1982).

A NOESY spectrum in ^1H_2O was recorded at 281 K with a 200-ms mixing time and a acquisition period of 102 ms. The intense solvent signal was suppressed by presaturation. Under the conditions of these experiments, the rate of exchange of the nonterminal imino protons is sufficiently slow to enable their detection (Rajagopal et al., 1988). For both the ^1H_2O and ^2H_2O experiments, free-induction decays were zero-filled to 2048 points in F_1 and transformed with a 60°-shifted sine bell in both dimensions.

A TOCSY spectrum was recorded at 281 K using DIPSI (Shaka et al., 1988) to produce the isotropic mixing. The spin-lock field was 8.3 kHz, and the mixing time was 50 ms. The data were processed as for the NOESY experiment.

One-dimensional NMR spectra were measured in ^1H_2O at 281, 288, and 303 K at 14.1 T using the $1\bar{3}3\bar{1}$ composite pulse for excitation (Hore, 1983). Truncated NOEs were measured for irradiation of the imino protons for 100 ms; the acquisition period was typically 0.9 s.

^{31}P NMR spectra were recorded at several temperatures at 4.7 T on a Bruker WM spectrometer. One-dimensional spectra were recorded with an acquisition time of 1.024 s, proton decoupling using Waltz-16, and a relaxation delay of 1 s. Data were processed by zero-filling to 8192 complex points

[1] Abbreviations: NOE, nuclear Overhauser enhancement; NOESY, two-dimensional nuclear Overhauser enhancement spectroscopy; ROESY, two-dimensional rotating-frame Overhauser enhancement spectroscopy; DIPSI, decoupling in the presence of scalar interactions; TOCSY, total correlation spectroscopy.

and apodizing with a 2-Hz line-broadening exponential. Spectra were referenced at 303 K to external methylene diphosphonate. Spectra recorded at different temperatures were referenced to the internal phosphate resonance whose shift was determined at 303 K.

A ^{31}P–^{1}H shift correlation experiment was recorded at 303 K, with 128 increments of 1300 acquisitions recorded over spectral widths of 600 Hz in F_2 and 1600 Hz in F_1, and 40-ms refocusing delays. The data were processed by zero-filling to 2048 points in F_2 and 1024 points in F_1 and apodizing with a 60°-shifted sine-squared function in both dimensions.

Molecular Modeling and Energy Minimization. Simulations of NOESY experiments were performed for a variety of nucleotide conformations using the program NUCFIT as previously described (Lane et al., 1991b).

The modeling strategy adopted for the mispaired 12-mer is an adaptation of the approach used for an equivalent A·G-containing sequence (Lane et al., 1991b). Initial coordinates for the d(HO·CGCGAATTGGCG·OH) duplex were generated using the program GENHELIX2 (Jenkins, unpublished); an idealized global B-DNA conformation was selected on this basis of the derived nucleotide parameters (see below). Interactive molecular modeling was performed using the GEMINI 1.03 graphics package (available from the CRC Biomolecular Structure Unit at the above address) with visualization on a Silicon Graphics IRIS4D-20G platform. All energy calculations were carried out on an Alliant FX40/3 system.

Models for each of the four possible symmetric anti·anti, anti·syn, syn·anti, and syn·syn G4(16)·G21(9)-paired duplexes were initially prepared by rotation of the appropriate C1′-(sugar)–N(base) glycosidic torsion angles (χ), with values altered in tandem for corresponding nucleotides in each strand (i.e., G4 ≡ G16, and G9 ≡ G21). Thus, four 12-mer duplex models were generated for subsequent structure refinement and energy minimization. The two G·G mispairs in each duplex were treated as equivalent for modeling purposes, and no attempt was made to model the possible minor tautomeric forms of either guanosine (see later).

The energy was minimized at the all-atom level using the X-PLOR program (Brünger, 1990a), adapted for double-precision use with the Alliant FX40 array processor system (J. Jian-Sheng and R. E. Hubbard, University of York, U.K.). Rationalized molecular electrostatic potentials (MEP) were used for all DNA bases (Lane et al., 1991b).

The protocol used for structure refinement is that suggested by Brünger (1990b), using stepwise (i) 500 steps of conjugate-gradient (Powell mechanics) minimization to remove bad contacts, (ii) molecular dynamics with heating to 300 K during 1 ps (0.2-fs time step), (iii) 10-ps dynamics (0.5-fs time step) with coupling to a heat bath at 300 K (friction coefficient = 100 ps^{-1}) and sampling at 0.5-ps intervals, and finally (iv) unrestrained conjugate-gradient minimization of the rms-averaged snapshots to an ultimate rms gradient of ≤4 kJ mol^{-1} nm^{-1} (i.e., ≤0.10 kcal mol^{-1} Å^{-1}). Soft glycosidic torsion-angle constraints (force constant = 42 kJ mol^{-1}) were applied throughout the refinement for bases other than the G·G mispairs, such that χ(pyrimidine) and χ(purine) were −115 ± 10° and −110 ± 10°, respectively. The SHAKE algorithm was used in the dynamic annealing steps to maintain bond lengths. Explicit distance restraints corresponding to Watson–Crick base-paired geometry were not included, and no attempt was made to restrain either terminal base pairs or the DNA backbone.

The effects of solvent and counterions were simulated by using a distance-dependent dielectric constant with $\epsilon = r_{ij}$ for the Verlet dynamics in steps ii and iii, and $\epsilon = 4r_{ij}$ for steps i and iv. Nonbonded energy terms were included up to 0.90 nm, with switching between 0.75 and 0.80 nm, and a factor of 0.4 was used to damp 1,4 electrostatic interactions. The hydrogen-bonded interactions were switched on for heavy atom donor to acceptor distances between 0.55 and 0.65 nm, and terms up to 0.75 nm were included. Typically, each minimization required ~26 500 cpu seconds.

Table I: UV Spectrophotometric Melting Data for Dodecamer Duplexes[a]

	0.1 M KCl[b]		1.0 M KCl	
property	G·G	*Eco*RI	G·G	*Eco*RI
T_m (K)	292.6 ± 0.1	328.4 ± 0.1	307.0 ± 0.2	338.0 ± 0.2[c]
ΔT_m (K)	−35.8		−31.0	
ΔA_{260} (%)	11.7 ± 0.1	14.2 ± 0.3	14.1 ± 0.2	14.2 ± 0.2

[a] Denaturation of 12-mers (10 μM) was monitored at 260 nm using a heating rate of 1 K min^{-1} (see Methods). [b] Buffer: 10 mM phosphate, 0.5 mM EDTA, KCl as indicated, and $^{1}H_2O$, pH 7.00. [c] A value of 344.2 K has been reported for 40 μM 12-mer (Leonard et al., 1990) under comparable conditions.

RESULTS

Thermal Stability of the d(CGCGAATTGGCG)$_2$ Duplex. In general, base-pair mismatches are destabilizing, though the magnitude of the effect depends on the nature of both the mispairing and the flanking sequences (Leonard et al., 1990). Thermal denaturation temperatures (T_m) have been determined using UV spectrophotometry (Leonard et al., 1990) for many sequences of the type d(CGCXAATTYGCG)$_2$, but data for the 12-mer with X = Y = G have not been reported. We find that, compared with the native *Eco*RI 12-mer under NMR conditions (0.1 M KCl), the T_m is about 36 K lower using this method (Table I). Further, the hypochromicity at 260 nm is substantially lower for the G·G-mismatched 12-mer. At high concentrations of salt (1 M KCl), however, the ΔA_{260} values for each sequence are comparable, indicating lower cooperativity for the G·G-containing duplex, particularly at low concentrations of salt. The melting curves for the 12-mers are monophasic, and there is no evidence for early premelting transitions. The ΔT_m values determined at both concentrations of salt (Table I) indicate that G·G is among the most destabilizing of the possible mismatches in this sequence context [see Leonard et al. (1990)].

While UV-absorbance melting profiles are useful chiefly for characterizing the global thermodynamic properties of a DNA duplex, the NMR experiment can, in contrast, provide information about individual groups. We have therefore recorded both ^{1}H and ^{31}P NMR spectra of the mismatched 12-mer as a function of temperature. Figure 1A shows representative ^{1}H spectra at several temperatures. Although there are some changes in the range 283–303 K, the greatest change occurs between 303 and 313 K, where the spectrum approaches that of the single-stranded state. Similarly, the ^{31}P NMR spectra (Figure 1B) show a relatively wide chemical shift dispersion (0.7 ppm) for the duplex state at low temperature. By contrast, at temperatures >303 K the spectra change radically and the chemical shift dispersion ultimately collapses to ~0.3 ppm, which is more typical of the single-stranded state (Patel et al., 1982). These data indicate that the duplex-to-strand transition occurs at ~313 K under the NMR conditions used. The single resonance at high field ($\delta \approx -16.9$ ppm) is broad between 279 and 285 K, but sharpens markedly as the temperature is increased to 303 K. This behavior is indicative of a low-temperature exchange process that affects the phosphodiester backbone of one nucleotide. The nature of this phosphate is discussed below.

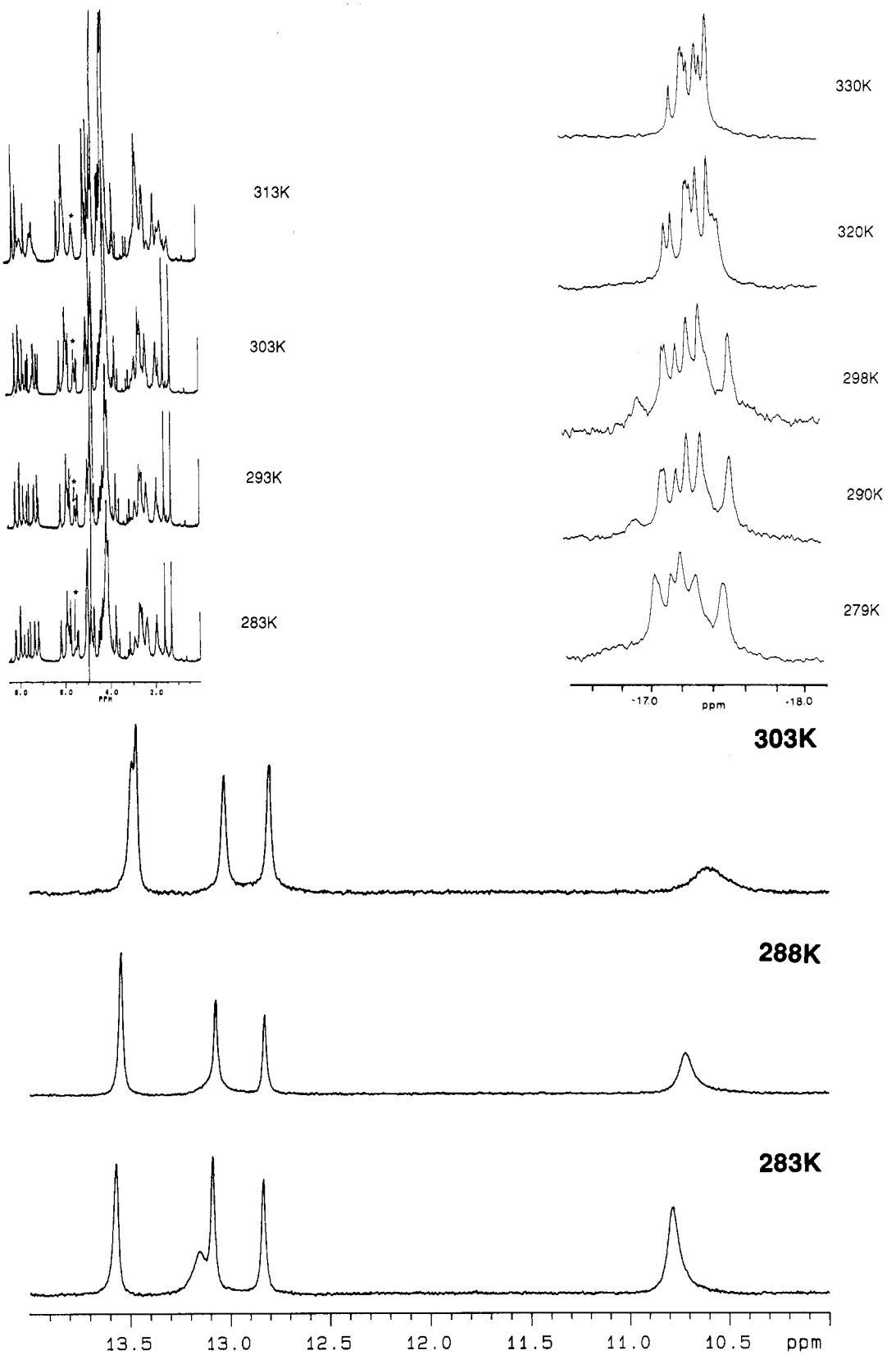

FIGURE 1: Temperature dependence of the NMR spectra for the G·G-mispaired 12-mer duplex. Spectra were recorded as described under Methods. (A, top left) ^1H NMR spectra in ^2H_2O solution at 283–313 K; the asterisk indicates the G9(H1′) resonance (see text). (B, top right) ^{31}P NMR spectra at 279–330 K. (C, bottom) ^1H NMR spectra at 281–303 K for the imino protons.

A more detailed examination of the temperature dependence of both individual chemical shifts and the overall shift dispersion (not shown) shows a global T_m of ca. 318 K, which should be compared with the value of 345 ± 2 K obtained for the *Eco*RI 12-mer under comparable NMR conditions (Patel et al., 1982). This result is in agreement with the optical studies performed at much lower DNA concentrations (see above). Under the conditions employed for our NMR experiments, we conclude that ≤5% of the G·G-mispaired DNA is single-stranded at 303 K.

Imino protons resonate between 13 and 15 ppm when they participate in hydrogen bonding and move several ppm upfield

Table II: Assignments for Nonexchangeable Protons of d(CGCGAATTGGCG)$_2$ at 281 and 303 K[a]

base	H6/H8	Me/H2/5	H1′	H2′	H2″	H3′	H4′
C1	7.63/7.61	5.85/5.87	5.71/5.75	1.96/1.95	2.45/2.37	4.65/4.67	4.04/4.05
G2	7.97/7.94		5.83/5.85	2.55/2.60	2.65/2.70	4.95/4.93	4.31/4.27
C3	7.33/7.29	5.42/5.41	5.76/5.75	1.85/1.85	2.30/2.30	4.75/4.74	4.12/4.15
G4	7.98/7.94		5.82/5.85	2.67/2.68		4.97/4.93	4.35/4.34
A5	8.17/8.12	7.17/7.17	5.86/5.88	2.70/3.01	2.81/3.01	5.05/5.03	4.43/4.40
A6	8.19/8.15	7.57/7.54	6.16/6.15	2.64/2.61	2.91/2.88	5.05/5.00	4.43/4.47
T7	7.14/7.09	1.30/1.30	5.85/5.89	2.01/1.96	2.60/2.51	4.80/4.77	4.15/4.12
T8	7.18/7.17	1.60/1.59	5.90/5.89	1.90/1.88	2.52/2.37	4.84/4.92	4.15/4.12
G9	7.78/7.72		5.51/5.51	2.84/2.87	2.62/2.64	4.97/4.92	4.28/4.27
G10	7.80/7.77		5.88/5.87	2.62/2.59	2.67/2.70	4.97/4.96	4.50/4.41
C11	7.33/7.29	5.37/5.38	5.69/5.77	1.93/1.90	2.33/2.30	4.82/4.80	4.12/4.09
G12	7.95/7.92		6.14/6.17	2.63/2.61	2.34/2.35	4.67/4.65	4.17/4.17

[a] Proton assignments at 281/303 K are based on the connectivities observed in the NOESY and TOCSY spectra (see text). Chemical shifts (δ, ppm) are referenced to internal DSS. Buffer: 20 mM phosphate, 0.2 M KCl, 1.0 mM EDTA, and ^2H_2O, pH 7.2.

in the single-stranded state. Further, non-hydrogen-bonded imino protons are often broadened by increased exchange with solvent. Hence, the imino protons are useful reporters of the duplex-to-strand transition. Figure 1C shows the imino protons resonances at several temperatures. There are five resonances at low temperatures. The four resonances between 13.6 and 12.8 ppm arise from the Watson–Crick hydrogen-bonded imino protons; the presence of only four resonances implies that the molecule is symmetric on average, i.e., two AT base pairs and three GC base pairs. The broader resonance at 10.8 ppm, which contains two protons, must arise from the G4 and G9 N1H. On raising the temperature, this high-field resonance shifts and broadens. Saturation of the solvent resonance causes complete loss of this signal (see below), but not of the Watson–Crick hydrogen-bonded protons. Hence, both G4 and G9 N1H exchange readily with the solvent, suggesting that they are only weakly hydrogen-bonded. In contrast, with the exception of the broader resonance at 13.15 ppm, the temperature dependence of the shifts of the imino proton resonances is small, and they do not broaden significantly between 281 and 303 K. This behavior indicates that, for bases other than G4 and G9, the Watson–Crick hydrogen bonding remains intact up to at least 303 K with the oligomer present largely as a duplex, supporting our conclusions based on ^{31}P chemical shifts (see above). The disappearance of the one resonance at 13.15 ppm can be attributed to "end fraying".

NMR Assignment of Nonexchangeable Protons. The ^1H NMR spectra were assigned at 281, 288, and 303 K using both NOESY and TOCSY information (Clore & Gronenborn, 1983; Hare et al., 1983; Patel et al., 1984; Scheek et al., 1983). Figure 2A shows a portion of the NOESY spectrum at 288 K, with the sequential H6/H8–H1′ assignments. The characteristic internucleotide NOEs for G2(H8)–C3(H5) and G10(H8)–C11(H5) were convenient anchor points for the sequential assignments in the H1′–H6/H8 proton pathway. Most of the H1′ and H6/H8 resonances could be readily found. There is only one missing connectivity in the base proton to the H1′ region, between C3H1′ and G4H8. This connectivity was found at 281 K but not at 288 and 303 K. The sequential NOE A6(H8)–T7(Me) provides an anchor point for the assignment by the H2′/H2″ to H6/H8 pathway (Figure 2B). There were some ambiguities in the H2′/H2″ pathway, due to missing connectivities and degeneracies involving the H8 of guanine residues. In particular, G9 H8 and G10 H8 have similar shifts, and there are missing sequential connectivities C3(H2″)–G4(H8) and T8(H2″)–G9(H8). To assign the sugar protons of these residues, we made use of the correlations from H1′ in both the TOCSY and NOESY experiments (not shown). Indeed, the NOESY experiment allows an unambiguous discrimination between H2′ and H2″, as the H2″–H1′ NOE is always larger than that for H2′–H1′. Further, an NOE connecting A5(H2)–A6(H2), and the interstrand A6(H2)–T8(H1′) NOE, allowed the A5(H2) and A6(H2) resonances to be assigned. The proton assignments for the 12-mer at 281 and 303 K are given in Table II.

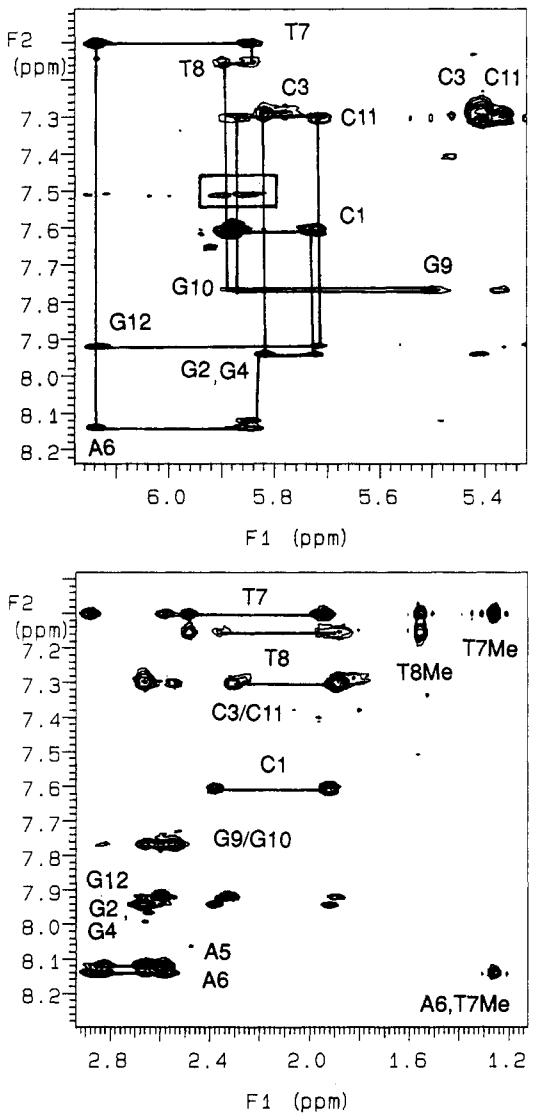

FIGURE 2: Partial NOESY spectrum of d(CGCGAATTGGCG)$_2$ at 288 K. The NOESY spectrum was acquired and processed as described under Methods with a mixing time of 150 ms on a Varian Unity spectrometer at 14.1 T. The continuous lines show the sequential assignments (A, top) H6/H8 to H1′ and (B, bottom) base to H2′/H2″.

Table II shows that, for both G9 and G12, the H2′ proton resonates downfield of H2″. While this is common for a

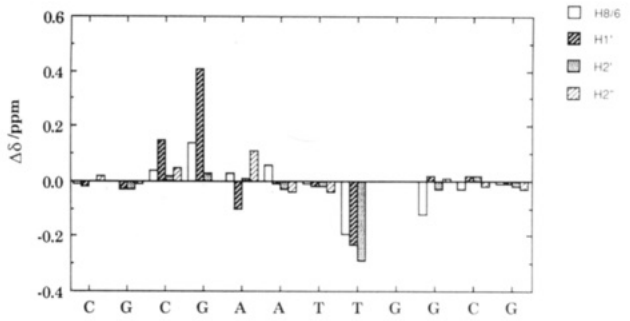

FIGURE 3: Proton chemical shift differences between the G·G-mispaired and native [d(CGCGAATTCGCG)$_2$] 12-mer duplexes. Chemical shifts for the *Eco*RI duplex were taken from Lane et al. (1991a). No information is given for G9 as this corresponds to a cytosine in the native duplex.

Table III: Assignments for Imino Protons of d(CGCGAATGGCG)$_2$[a]

base pair	283 K	288 K	303 K
C1·G12	13.18	*b*	*b*
G2·C11	13.09	13.05	13.03
C3·G10	12.85	12.83	12.80
G4·G9	10.78	10.72	10.63
A5·T8	13.56	13.53	13.48/13.47
A6·T7	13.56	13.53	13.47/13.48

[a] The assignments were obtained from one-dimensional NOE and NOESY experiments in ^1H_2O (see text). [b] The peak was absent at the higher temperatures. Chemical shifts (δ, ppm) are referenced to internal DSS. Buffer: 20 mM phosphate, 0.2 M KCl, 1.0 mM EDTA, and $^2H_2O/^1H_2O$, pH 7.2.

3′-terminal base, it is unusual for an internal base and suggests that the local magnetic environment for G9 is more like that of a terminal nucleotide. This would be consistent with poor base stacking of G9 with G10. It is instructive to compare the chemical shifts of the mispaired duplex with those in the parent *Eco*RI dodecamer, recently reported at 303 K (Lane et al., 1991a). The chemical shift differences determined for the two sequences are shown in Figure 3. Ignoring the altered base at position 9, it is clear that the major differences are localized at the mismatched base G4 and the neighbors nearest to the G·G site as a consequence of the altered ring currents from the mismatch. The absence of significant chemical shift differences elsewhere suggests that any major conformational rearrangements are not propagated along the duplex.

Assignment of Exchangeable Protons. The imino proton of CG1 was assigned by noting the exchange broadening due to fraying (see above). Other protons were assigned by one-dimensional NOE and NOESY methods in ^1H_2O. The two A·T imino protons, which are degenerate at most temperatures, were readily assigned by both their chemical shift and the observation of NOEs to the assigned H2 resonances. The G·C protons were assigned by observation of NOEs to C(H5). The broad peak at 10.8 ppm (Figure 1C) assigned to the G4/9 imino protons (see above) disappears when the solvent signal is saturated, indicating an exchange rate constant on the order of $1/T_1$. The other imino protons do not disappear by transfer of saturation under these conditions. Exchangeable proton assignments are given in Table III. Because both G4 and G9N1H are observed, base-pairing schemes based on enol forms of the bases can be ruled out. Hence the mechanism proposed by Topal and Fresco (1976) is not operative for this sequence.

Assignment of Phosphorus Resonances. The ^{31}P resonances at 303 K were assigned using the ^{31}P–^1H heteronuclear shift correlation experiment (not shown), the relative areas of the peaks as in Figure 1B, and by reference to the assignments of the parent duplex (Ott & Eckstein, 1985). Owing to degeneracies of H4′ and H3′ resonances (cf. Table II), the assignment of some phosphates from the heteronuclear shift correlation experiment was ambiguous, e.g., C3pG4, T7pT8, and G10pC11. In other cases, e.g., C11pG12, the ambiguity could be resolved by assuming that its shift in the mismatched duplex should be very similar to that in the parent dodecamer. This is because shifts of residues remote from the mismatch site in general are similar to those in the parent dodecamer (cf. Figure 3). These assignments are marked with an asterisk in Table IV. The remaining ambiguities are those involved in the mismatch site (G4pA5 and G9pG10). However, from the heteronuclear shift correlation experiment either G4pA5 or G9pG10 must be in the peak at low field (which contains one phosphate resonance).

Table IV: Phosphorus Assignments of d(CGCGAATTGGCG)$_2$ at 303 K[a]

position	chemical shift (ppm)
C1pG2	−17.14
G2pC3	−17.22
C3pG4	−17.14/−17.32/−17.55
G4pA5/G9PG10	−16.86
A5pA6	−17.33
A6pT7	−17.22
T7pT8	−17.14/−17.32/−17.55
T8pG9	−17.31
G10pC11	−17.14/−17.32/−17.55
C11pG12	−17.03*

[a] The assignments were obtained from ^{31}P,^1H heteronuclear correlation experiments, and by reference to the native dodecamer (Ott & Eckstein, 1985) as described in the text. An asterisk denotes ambiguities resolved by reference to the native dodecamer (see text). The chemical shifts were referenced to external methylene diphosphonate.

Most of the assigned shifts are consistent with those of in the native dodecamer (Ott & Eckstein, 1985), with the exception of G4pA5/G9pG10 and A6pT7. The latter seems to be particularly sensitive to the nature of mismatches in this kind of duplex; Roongta et al. (1990) have reported variations of 0.45 ppm. Also, as phosphate shifts reflect backbone conformation, the shifts for the G4pA5/G9pG10 indicate a significantly asymmetric conformation at the mismatch site (Roongta et al., 1990; and see below).

Conformation at Low Temperatures. The NOESY and one-dimensional NMR experiments show that the 12-mer is a duplex up to at least 303 K, and this is confirmed by measurements of the apparent rotational correlation time, τ_c, for the C(H6)–(H5) vectors. The apparent correlation time of 6.1 ± 0.5 ns at 281 K determined from the NOE time courses should be compared with 3.3 ± 0.3 ns reported for the *Eco*RI 12-mer at 303 K (Lane et al., 1991a). Assuming a simple dependence on solvent viscosity (Birchall & Lane, 1990), the correlation time for the G·G mismatched 12-mer would be approximately 3.2 ns at 303 K. This indicates that the oligomer behaves as a rod-like duplex with dimensions typical for B-DNA.

The similarity of the chemical shifts for protons on residues at least one base removed from the mismatch sites suggests that the conformation in these regions may be similar to that in the *Eco*RI 12-mer. Examination of cross-sections of the NOESY spectra through the H6 and H8 protons reveals that the intranucleotide NOEs are large for H6/H8–H2′ (i.e., of equal or greater intensity than the H6–H5 cross-peaks of the cytosine residues), and small for H6/H8 to H1′, H3′, and H2″ (cf. Figure 4). This pattern of NOEs is characteristic for nucleotides where the glycosidic torsion angle (χ) is anti (i.e., $\chi \approx -110°$), and the sugar pucker is in the "south" domain. Further, the internucleotide NOEs, except for the steps C3–G4 and T8–G9, are indicative of a right-handed helix. Additional

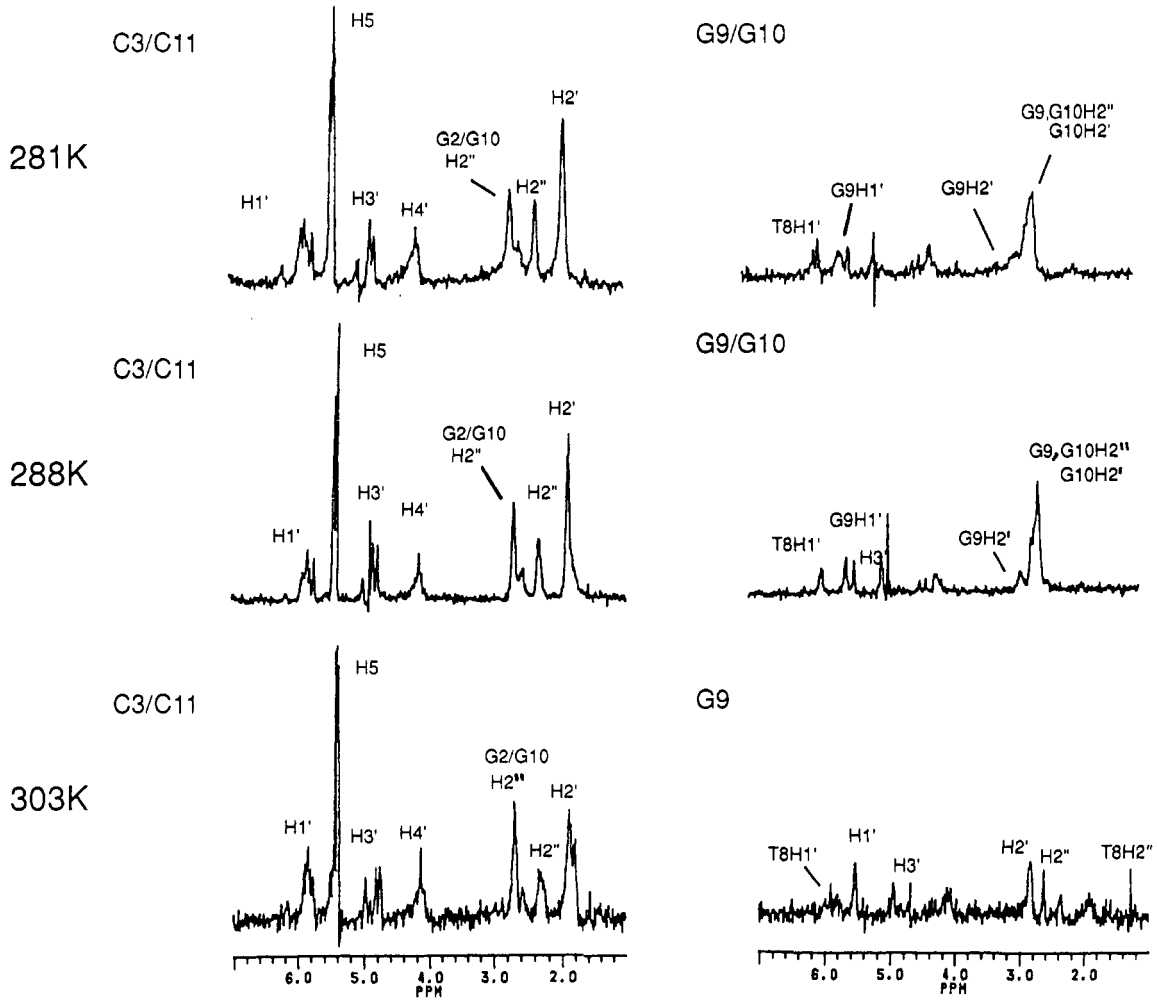

FIGURE 4: Cross-sections from NOESY spectra for the G·G duplex at 281 K (t_m = 90 ms) (top), 288 K (t_m = 100 ms) (middle), and 303 K (t_m = 200 ms) (bottom). (Left) Rows through the C3/11(H6) protons. (Right) Rows through the G9/10(H8) protons. The spectra are scaled identically. See Methods for details of the data collection and processing.

NOEs, including G2(H1′)–C3(H5) and G10(H1′)–C11(H5), are also consistent with a right-handed screw at these steps.

The NOESY spectra show a substantial intrastrand NOE between A5(H2) and A6(H2) and a cross-strand NOE between A6(H2) and T8(H1′). In addition, we have used time-dependent NOEs (see Methods) to determine the A5-(H2)–A6(H2) separation at low temperature; we obtained r = 0.3 ± 0.02 nm. This relatively short separation (in B-DNA the distance is approximately 0.38 nm) and the significant cross-strand NOE are consistent with a large differential propeller twist for the A5(17)·T20(8) to A6(18)·T19(7) base-pair steps. Hence, we conclude that much of the duplex adopts a typical B-like DNA conformation in solution.

However, there are anomalies in the NOEs at the steps C3–G4 and T8–G9. As Figure 4 reveals, the cross-sections through the G9(H8) resonance show nontypical intranucleotide NOEs. In particular, all of the NOEs are relatively weak, compared with those of the other nucleotides. The lower signal-to-noise ratio of the cross-sections at 303 K arises from the presence of multiple conformations in slow exchange (see below), which has the effect of reducing the effective concentration of the observed species. Although the G9H8 and G10H8 resonances overlap at low temperatures, their sugar protons are well-resolved from one another (see Table II). A row through the G9/G10 H8 resonance will give a correct relative intensity for G9H1′ and an overestimate for G10H1′ (which will contain contributions from both intra- and internucleotide NOEs). Similarly, the G9H8–G9H2′ cross-peak will be overestimated due to the (small) contribution from the internucleotide NOE. We have measured the relative intensities for the cross-peaks, normalized to the cytosine H6–H5 cross-peak intensities, and compare them with those calculated for different nucleotide conformations assuming a correlation time of 6 ns and a mixing time of 90 ms, corresponding to the conditions at 281 K (cf. Figure 4). We find that the intensities for most nucleotides are in the order CH6–CH5 ≈ H2′ > H2″ > H3′ ≈ H1′, with the H2′ cross-peak approximately 8 times as strong as the H1′ cross-peak. The relative intensities for G9 are in the order CH6–CH5 > H2′ > H1′. In the syn conformation, the H8–H1′ cross-peak intensity is expected to give the largest intranucleotide base–sugar intensity, nearly as strong as the CH6–CH5 cross-peak, and very low intensities to H2′ and H2″. The observed NOE intensities for G9 are inconsistent with the syn conformation but are consistent with a χ(anti) value near −150°. However, the NOEs could be affected by averaging about the glycosidic bond (see below), though the nucleotide must spend only a small fraction of its time near χ = −110° (where the H8–H2′ NOE is large). There is no evidence to suggest that the sugar pucker of G9 is C3′-endo, as no H2″–H4′ NOE was observed. Also, the G9(H8–H3′) NOE is small; a relatively large NOE would be expected for a sugar in the C3′-endo state. We also note that the G9(H1′–H2″) NOE is similar to that for other nucleotides. Therefore the small NOEs observed from H8 are not due to exchange processes that compete with magnetization transfer to either H1′ or H2″.

G2 and G4 also have overlapping resonances. However, the relative intensity of the G2/G4H8–H1′ is not significantly larger than any other base to H1′ cross-peak intensities. As the observed cross-peak will in this instance be the sum of two intranucleotide interactions, the intensity for either must be smaller than the observed value. Because the intensity is much smaller than that of the Cyt H6–H5, we conclude that G4 must also be anti. However, because of the overlap of the other sugar protons, it is not possible to determine the conformation with greater precision.

The lack of H2″(i–1)–H8(i) NOEs for G4 and G9 also suggests a conformation different from canonical B-DNA at the mismatch sites. However, the degeneracy of the H2′/H2″ protons of G9 and G10 prevents even a qualitative determination of the conformation at this step. The chemical shift of G4pA5/G9pG10 is unusually downfield-shifted, consistent with a nonstandard backbone angle at this step (Nikonowicz & Gorenstein, 1990; Roongta et al., 1990). In addition to the unusual shifts of the H2′ and H2″ of G9 and the lack of hydrogen bonding of the G4/9 imino protons, it seems likely that the G·G bases are only partially stacked into the helix and that these nucleotides have considerable flexibility. This behavior would explain both the weak NOE connectivities and the different backbone properties at the mismatch sites and would be consistent with a considerable helical bulge. Further, Gorenstein and co-workers have argued that the ^{31}P chemical shift is strongly influenced by the backbone torsion angles and is a measure of the relative populations of the B_I and B_{II} conformations (Gorenstein et al., 1988; Nikonowicz & Gorenstein, 1990; Roongta et al., 1990). On this basis, the downfield-shifted phosphate resonance for G4pA5 would reflect a relatively high population of the less common B_{II} state.

Effects of Increased Temperature. The qualitative temperature dependence of the NMR spectra in terms of global melting was discussed above. Although there appears to be one dominant average conformation at temperatures lower than ca. 288 K, the G9(H1′) resonance at this temperature is nevertheless broader than those for other resolved H1′ protons, suggesting the presence of some chemical exchange. This conclusion is supported by the observation that the downfield G4pA5 phosphate resonance also shows pronounced exchange broadening at low temperature but becomes sharp by 303 K, a temperature which is below the global T_m (Figure 1B). We have therefore examined the conformational behavior in greater detail by recording NOESY spectra at 298 and 303 K, where there are new cross-peaks in both the base–base and H1′–H1′ regions (Figure 5A,B). In the latter case, these additional cross-peaks are likely to arise from exchange processes. To determine whether the peaks arise from exchange or from cross-relaxation, we have recorded a ROESY spectrum at 303 K with a mixing time of 200 ms. The appropriate segments of the spectrum are shown in Figure 5C,D. Many of the cross-peaks are negative and must therefore arise from chemical exchange. While there are clearly many cross-peaks, in all cases these connect to a frequency observed at low temperatures. Hence, we conclude that there must be at least two conformations present in slow equilibrium exchange at 303 K. The rate of interconversion must be of the order $1/t_{mix}$, or $\sim 5\ s^{-1}$. Further, the relative areas for C3(H6) in the two states are approximately 1:1 at 303 K, based on the one-dimensional spectrum. This ratio far exceeds the proportion of single-stranded state present at 303 K implied by the spectra in Figure 1.

The absence of a large effect on the shifts of the imino protons suggests that the conformational differences may be relatively minor. At 303 K, many of the resonances on residues C3–C11 are doubled (Figure 5), suggesting widespread conformational heterogeneity in the duplex adjacent to the mismatch site. However, at 298 K there are fewer exchange peaks, and those that are present involve resonances for either the mispaired bases or their nearest neighbors. Unfortunately, the NOEs can then no longer be quantified as the effects of exchange and unknown populations make the intensities quantitatively uninterpretable. Thus, we are not able to determine the nature of the alternative conformation. Nevertheless, the flexibility of G9 (and probably G4) increases as the temperature is raised, such that at least two conformations in slow exchange become populated above 298 K. This is probably related to poorer base stacking with adjacent bases. On further raising the temperature, alternative conformations become populated for nucleotides which are more distant from G4/G9. Hence, it appears that the G·G mismatch is intrinsically flexible and can act as a nucleus for melting from within the DNA duplex. This in some part explains the reduced thermal stability of the G·G-mispaired 12-mer.

Table V: Calculated Energies for the G·G-Mispaired 12-mer Duplexes[a]

mispair	E_{vdW}	E_{elec}	E_{HB}	E_{total}
G(anti)·G(anti)	−1481	−1296	−351	**−810**
G(anti)·G(syn)	−1466	−1312	−348	−786
G(syn)·G(anti)	−1477	−1312	−370	−802
G(syn)·G(syn)	−1421	−1348	−342	−794

[a] Structures for the possible G·G base-pair conformations involving anti·anti, anti·syn, syn·anti, and syn·syn duplexes were refined using X-PLOR as described. All energies are given in kilojoules per mole for duplexes containing two symmetrically disposed G4·G21 and G9·G16 base pairs. [b] The numbering scheme used above is as follows:

d5′ –(C1 G2 C3 **G4** A5 A6 T7 T8 **G9** G10 C11 G12)
d3′ –(G24 C23 C22 **G21** T20 T19 A18 A17 **G16** C15 G14 C13)

Changes in the apparent propeller twist occur as the temperature is increased. At 288 K there is an intrastrand NOE between A6H2 and T7H1′ and a cross-strand NOE between A6H2 and T8H1′ (cf. Figure 2A), whereas at 303 K only the intrastrand NOE between A6H2 and T7H1′ was observed and at 281 K only the cross-strand NOE between A6H2 and T8H1′ is present. The propeller twists probably decrease with increasing temperature, as has been observed in other oligonucleotides (Lane, 1991).

Molecular Modeling Studies. The G·G-mismatched duplex was investigated by using a simulated isothermal dynamic annealing procedure. The X-PLOR protocol used (see Methods) is similar to that adopted for the corresponding A·G mismatch (Lane et al., 1991b) but differs by using dynamic averaging of samples taken at 0.5-ps intervals during 10 ps of temperature-coupled annealing at 300 K. Thermal equilibration of the system is achieved within our 1-ps heating period, as judged from component energy terms and the temperature. We note that, in the absence of counterions and solvent molecules, a simulation period of 10 ps is sufficient to ensure that atomic coordinates converged satisfactorily to those of the rms-averaged snapshots accumulated after each 0.5 ps. Prolonged heating of a 12-mer duplex has recently been reported to induce a transition from B-DNA to A-DNA during 30 ps (Boehncke et al., 1991). In contrast, improved structures are achieved following only very protracted dynamics simulations if time-averaged distance restraints are employed (Pearlman & Kollman, 1991).

Table V shows the total energies and component energy terms calculated for each of the four model 12-mers containing symmetrically disposed G·G base-pairs following 10-ps dy-

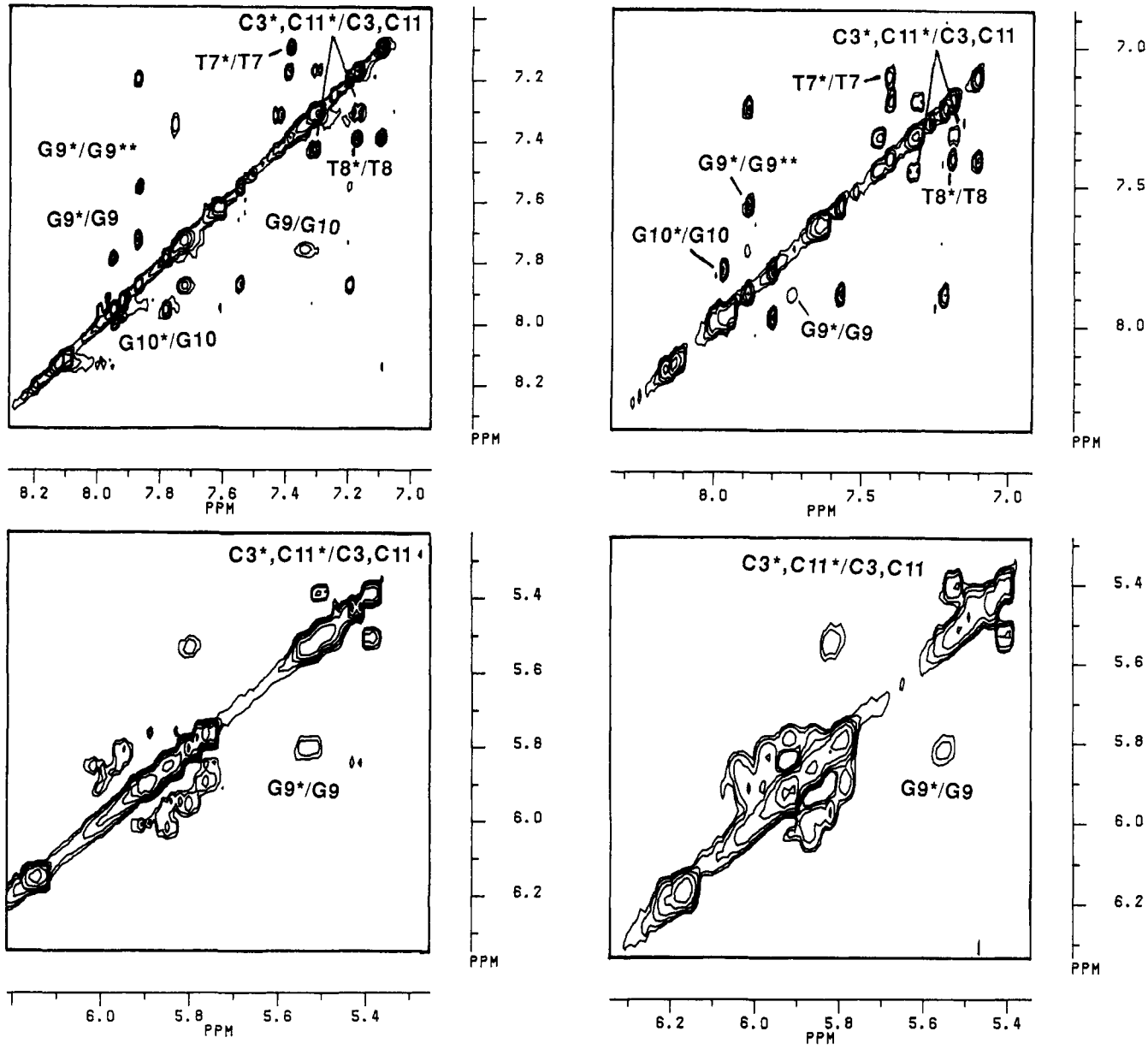

FIGURE 5: Chemical exchange in d(CGCGAATTGGCG)$_2$. NOESY (left) and ROESY (right) spectra at 303 K were recorded at 9.4 T using a 200-ms mixing time in both experiments, as described under Methods. The base–base region is shown in the top two panels and the H1′–H1′ region in the bottom two panels. Negative contours only are displayed.

namic annealing at 300 K. The G(anti)·G(anti) alignment of the mispair is clearly of lower energy than the other possible conformations, largely as a consequence of superior van der Waals nonbonded interactions (E_{vdW}) rather than electrostatic (E_{elec}) or hydrogen-bonded (E_{HB}) terms, but is only 8 kJ mol^{-1} more stable than the G(syn)·G(anti) pairing. Interestingly, the order of anti·anti > syn·anti > anti·syn ≈ syn·syn for total energy is similar to that determined for the equivalent A·G-mismatched d(CGCAAATTGCGC) duplex at high pH (Lane et al., 1991b). However, the energy range separating the four conformations is only 24 kJ mol^{-1}, rather than the 75 kJ mol^{-1} span calculated for A·G, suggesting that this G·G-mispaired sequence is inherently less stable in conformational terms and that several different conformations might become populated under slightly different conditions.

The lowest-energy G(anti)·G(anti) structure determined for the mispair is shown in Figure 6A. This structure represents one of a class of conformations that may be present in solution. Helix analysis of the duplex shows an average helical rise of 0.34 ± 0.06 nm and a mean helical rotation of 37 ± 9° (9.7 bp/turn), compared with 0.35 ± 0.05 nm and 37 ± 5° (9.7 bp/turn) for the *Eco*RI 12-mer (Lane et al., 1991a). The duplex thus adopts a general B-like DNA conformation, but there is considerable localized perturbation of the helix (Figure 6A). Examination of the structure reveals that all sugar puckers are C2′-endo, except those for G4 and G16, which approach a C1′-exo twist form in the southern pucker domain. The base stacking is not disrupted, other than at the two G·G sites, and there appear to be only weak induced rotation or translation effects upon the neighboring bases. However, it is clear that, according to the Cambridge convention, both G4 and G16 are perturbed by significant displacement (D_x) and shift (D_y) translations toward the major groove. The two G·G base pairs are thus aligned in a parallel-slipped arrangement which results in poor stacking of G4 and G16 with their adjacent bases (Figure 6B). Such an effect may account for the low hypochromicity found during UV melting (Table I) and the interruption of the sequential NOE patterns at steps involving these bases. The translation effects for G4 and G16 are equivalent by symmetry and result in localized bulges of the backbone of each strand. The mean C1′–C1′ cross-strand separations for the G·G pairs are 1.23 nm, compared with an

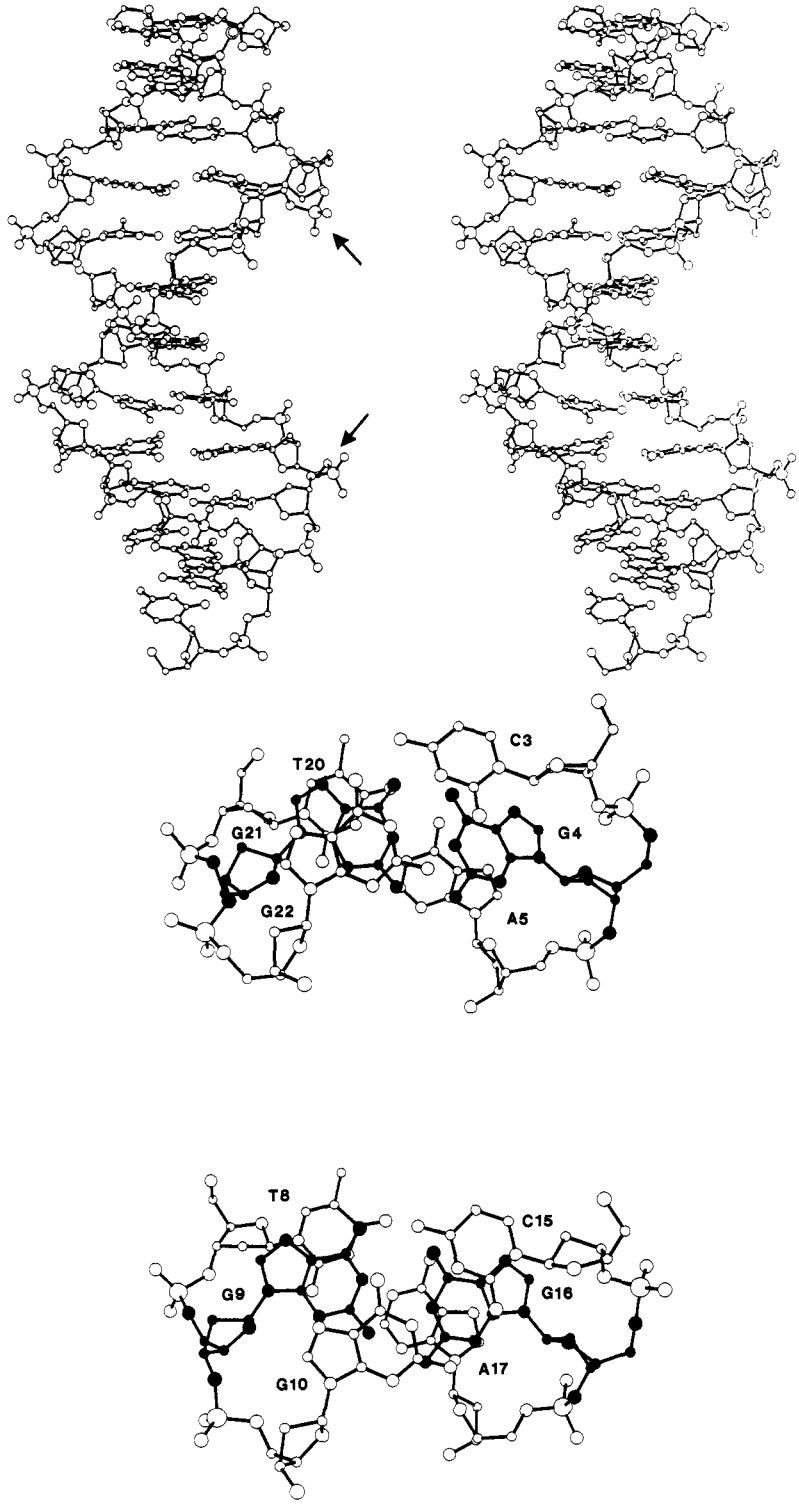

FIGURE 6: Views of the G·G-mismatched duplex generated by using PLUTO following X-PLOR refinement. (A, top) Stereoviews of the 12-mer showing the major groove of the core (H atoms have been removed for clarity). The arrows at G4pA5 and G16pA17 indicate the bulges of the DNA backbone toward the major groove. (B, bottom) Views of the C3·G22, G4·G21, and A5·T20 base pairs (left) and of the T8·A17, G9·G16, and G10·C15 base-pairs (right) through the helical axis showing the stacking of the bases. Note the poorly stacked arrangements for G4 and G16.

average of 1.04 nm for the remaining base pairs. The DNA bulges result in a minor groove which widens at the mismatch sites, such that the effective width is increased by ~20% compared to the normal minor groove width expected for a B-type DNA (Saenger, 1984). This widening effect is confined to nearest-neighbor bases and not propagated throughout the stack.

The DNA backbone in the vicinity of C3pG4pA5 and the corresponding nucleotides on the complementary strand have altered backbone torsion angles which reflect the helical translations of G4 and G16. Inspection of the dynamic progress during simulated annealing showed that the phosphodiester backbones in these regions, and particularly the G4pA5 and G16pA17 phosphate groups, experience the largest dynamic excursions. However, this loosening appears to be focused on only 2–3 bases of each strand. The two G·G pairs are only weakly hydrogen-bonded with mean G4(O6)···G21(H1) and G16(O6)···G9(H1) separations of 0.23 nm, but the

achieved geometries are poor with mean subtended O6···H1–N1 angles of only 128°, respectively. Further hydrogen bonding is absent, largely as a consequence of the large displacements for G4 and G16, and neither N3/N7 nor the exocyclic NH_2 groups are involved. Favorable hydrogen-bonded arrangements are formed for the higher-energy G·G conformations, but in each case the helical disruption (largely as induced buckle, roll, or twist effects) was considerably greater than for the low-energy G(anti)·G(anti) conformation.

Discussion

Several structural features can be deduced from the NMR information obtained for the G·G-containing sequence: (i) the two purines are not strongly hydrogen-bonded, (ii) the base pairs bulge out from the DNA helix, (iii) G9 has an unusual χ angle which is more typical for the anti conformation in A-form DNA, although the sugar appears to remain largely in the C2′-endo domain, and (iv) unusual phosphate backbone torsion angles are probably involved at the mismatch site. These conclusions are supported by our molecular modeling studies.

Topal and Fresco (1976) predicted that the chemical structure of the bases for a G·G mispairing would be G-(enol,imino,anti)·G(syn). This particular structure is ruled out on the basis of the imino proton spectra, which show that both mispaired bases are in the keto tautomeric form. Other base pairings are possible in the normal tautomeric forms, which are difficult to distinguish on the basis of the NMR spectra of the exchangeable protons. We find that at low temperature the major conformer is G(*anti*)·G(*anti*), for which strongly hydrogen-bonded structures are not possible. Molecular modeling is also consistent with only weak hydrogen bonding, and relatively poor base–base stacking of the mismatched pair, which may account for the observed low stability of this duplex. However, we cannot rule out a small fraction of the syn state, either at low temperature or in the alternative high-temperature conformations. A detailed conformational analysis is not possible since at least two conformations, with unknown relative populations, are present at high temperature. Further, at low temperatures where one average structure appears to be present, there is considerable spectral overlap. The quantitative information available from the NMR experiments is therefore insufficient to calculated structures ab initio. The conformational flexibility associated with the mismatch site is also in agreement with the energy calculations, such that different conformations are likely to form under different conditions, including a G(syn)·G(anti) pair (Cognet et al., 1991).

The overall stability of the mismatched sequence is considerably less than for the native *Eco*RI 12-mer, as seen by the difference in T_m determined using both NMR and spectrophotometric techniques. It has been suggested (Patel et al., 1982; Ott & Eckstein, 1985) that the A/T stretch of the native dodecamer melts before the G/C region in a biphasic melting process at salt concentrations similar to those used here. However, in the temperature range we have examined, the A/T region appears not to undergo a base-pair melting transition, as is evident from the comparable line widths of the A/T imino protons at the three temperatures. The salt-dependent hypochromicity found for the G·G sequence (Table I) can be explained in terms of poor stacking within the DNA, in accord with our modeling results.

Roongta and co-workers (1990) have recently published ^{31}P NMR data on a number of mispaired sequences, including a different G·G mismatch, d(CGGCAATTGGCG)$_2$, where the G·G pairing influenced the chemical shifts of both the nearest and next-nearest neighbors. Thus, in general, it can be assumed that the conformation adopted by the G·G mispair tends to cause large conformational perturbations which are propagated through the helix. It has been pointed out (Roongta et al., 1990) that, like most other mismatches, the G·G (and the G·A, G·U, and A·C) mispaired oligomers do not show isolated downfield resonances in the ^{31}P NMR spectrum. This contrasts with our observation that a downfield-shifted resonance is observed at some temperatures (Figure 1B). However, the line width and shift of this resonance are strongly temperature-dependent and reflect the conformational averaging that must occur in the vicinity of the G·G mispairing. This feature could therefore be easily missed.

The three possible purine·purine mismatches are repaired with very different efficiencies, in the order G·G > A·A > A·G (Kramer et al., 1984). The A·G mispair can adopt a variety of conformations depending on the nature of the flanking sequence and the conditions (Gao & Patel 1988; Brown et al., 1990; Lane et al., 1991b; Webster et al., 1990), although the predominant conformation appears to be A(anti)·G(anti) under physiological conditions. Further, the A·A base-pairing in d(CCCAGGG)$_2$ is primarily A(anti)·A(anti) (Arnold et al., 1987), and its stability and stacking properties appear to be very similar to those we describe for the G·G mismatch. Other purine·purine mispairs also appear to adopt an anti·anti conformation in solution (Fazakerley et al., 1987; Uesugi et al., 1987).

One major difference between the A·G pairing and the homopurine mismatches is that hydrogen bonds can be formed in the heteropurine mismatch in a way that preserves the structure of the base pair similar to that of the Watson–Crick base pair. Analogous hydrogen bonding in the G·G mismatch either requires use of unusual tautomers or destabilizing conformational rearrangements. This hydrogen bonding in A·G mismatches may be sufficient to stabilize the base pair, afford better stacking with its neighbors, and hence more resemble a typical B-like DNA. The greatly decreased stability of the G·G mismatch, and the resulting poor base stacking, leads to a more pronounced bulge and a larger-sized lesion that might be detected by the error-correcting enzymes. However, this does not explain why G·T and A·C purine·pyrimidine mispairs are repaired more efficiently than A·A, suggesting that factors other than local conformational effects are also implicated.

Acknowledgments

We thank C. J. Bauer for helpful discussions, especially with regard to the ROESY experiment. We thank Prof. R. E. Dickerson for the NEWHEL91 program used for helical analysis.

Registry No. d(CGCGAATTGGCG)$_2$, 123270-11-1; guanine, 73-40-5.

References

Arnold, F. H., Wolk, S., Cruz, P., & Tinoco, I. (1987) *Biochemistry 26*, 4068–4075.

Bauer, C. J., Frenkiel, T. A., & Lane, A. N. (1990) *J. Magn. Reson. 87*, 144–152.

Birchall, A. J., & Lane, A. N. (1990) *Eur. Biophys. J. 19*, 73–78.

Boehncke, K., Nonella, M., Schulten, K., & Wang, A. H.-J. (1991) *Biochemistry 30*, 5465–5475.

Bothner-By, A. A., Stevens, R. L., Lee, J. T., Warren, C. D., & Jeanloz, R. W. (1984) *J. Am. Chem. Soc. 106*, 811–815.

Brown, T., Leonard, G. A., Booth, E. D., & Kneale, G. (1990) *J. Mol. Biol. 212*, 437–440.

Brünger, A. T. (1990a) X-PLOR *Manual*, version 2.1, Yale University, New Haven, CT.
Brünger, A. T. (1990b) in *Molecular Dynamics: Applications in Molecular Biology* (Goodfellow, J. M., Ed.) pp 137–178, Macmillan Press, London.
Carbonnaux, C., van der Marel, G. A., van Boom, J. H., Guschlbauer, W., & Fazakerley, G. V. (1991) *Biochemistry 30*, 5449–5458.
Clore, G. M., & Gronenborn, A. M. (1983) *EMBO J. 2*, 2109–2115.
Cognet, J. A. H., Gabarro-Arpa, J., Le Bret, M., van der Marel, G. A., van Boom, J. H., & Fazakerley, G. V. (1991) *Nucleic Acids Res. 24*, 6771–6779.
Dickerson, R. E., & Drew, H. R. (1981) *J. Mol. Biol. 149*, 761–786.
Dohet, C., Wagner, R., & Radman, M. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 503–505.
Fazakerley, G. V., Quignard, E., Woisard, A., Guschlbauer, W., van der Marel, G. A., van Boom, J. M., Jones, M., & Radman, M. (1986) *EMBO J. 5*, 3697–3703.
Fazakerley, G. V., Sowers, C., Erijta, R., Kaplan, B. E., & Goodman, M. F. (1987) *Biochemistry 26*, 5640–5641.
Fersht, A. R., & Knill-Jones, J. W. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4251–4255.
Fowler, R. G., Degen, G. E., & Cox, E. C. (1974) *Mol. Gen. Genet. 133*, 179–191.
Fratini, A. V., Kopka, M. L., Drew, H. R., & Dickerson, R. E. (1982) *J. Biol. Chem. 257*, 14686–14707.
Frenkiel, T. A., Bauer, C. J., Carr, M. D., Birdsall, B. B., & Feeney, J. (1990) *J. Magn. Reson. 90*, 420–425.
Gao, X., & Patel, D. J. (1988) *J. Am. Chem. Soc. 110*, 5178–5182.
Gorenstein, D. G., Schroeder, S. A., Fu, J. M., Metz, J. T., Roongta, V., & Jones, C. R. (1988) *Biochemistry 27*, 7223–7237.
Hare, D. R., Wemmer, D. E., Chou, S.-H., Drobny, G., & Reid, B. R. (1983) *J. Mol. Biol. 171*, 319–336.
Hore, P. J. (1983) *J. Magn. Reson. 55*, 283–300.
Jones, G. B., Davey, C. L., Jenkins, T. C., Kamal, A., Kneale, G. G., Neidle, S., Webster, G. D., & Thurston, D. E. (1990) *Anti-Cancer Drug Des. 5*, 249–264.
Kramer, B., Kramer, W., & Fritz, H. J. (1984) *Cell 38*, 879–887.
Lane, A. N. (1991) *Biochem. J. 273*, 383–391.
Lane, A. N., Lefèvre, J.-F., & Jardetzky, O. (1986) *J. Magn. Reson. 66*, 201–218.
Lane, A. N., Jenkins, T. C., Brown, T., & Neidle, S. (1991a) *Biochemistry 30*, 1372–1385.
Lane, A. N., Jenkins, T. C., Brown, D., & Brown, T. (1991b) *Biochem. J. 279*, 269–281.
Leonard, G. A., Booth, E. D., & Brown, T. (1990) *Nucleic Acids Res. 18*, 5617–5623.
Loeb, L. A., & Kunkel, T. A. (1982) *Annu. Rev. Biochem. 51*, 429–457.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 124*, 774–783.
Nerdal, W., Hare, D. R., & Reid, B. R. (1989) *Biochemistry 28*, 10008–10021.
Nikonowicz, E. P., & Gorenstein, D. G. (1990) *Biochemistry 29*, 8845–8858.
Ott, J. E., & Eckstein, F. (1985) *Biochemistry 24*, 2530–2535.
Patel, D. J., Kozlowski, S. A., Marky, L. A., Broka, C., Rice, J. A., Itakura, K., & Breslauer, K. J. (1982) *Biochemistry 21*, 428–436.
Patel, D. J., Kozlowski, S. A., Ikuta, S., & Itakura, K. (1984) *Biochemistry 23*, 3207–3217.
Pearlman, D. A., & Kollman, P. A. (1991) *J. Mol. Biol. 220*, 457–479.
Piotto, M. E., & Gorenstein, D. G. (1991) *J. Am. Chem. Soc. 113*, 1438–1440.
Rajagopal, P., Gilbert, D. E., van der Marel, G. A., Boom, J. H., & Feigon, J. (1988) *J. Magn. Reson. 78*, 526–537.
Roongta, V. A., Jones, C. R., & Gorenstein, D. G. (1990) *Biochemistry 29*, 5245–5258.
Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, Berlin.
Scheek, R. M., Russo, N., Boelens, R., & Kaptein, R. (1983) *J. Am. Chem. Soc. 105*, 2914–2916.
Searle, M. S., Hall, J. G., Denny, W. A., & Wakelin, L. P. G. (1988) *Biochemistry 27*, 4340–4349.
Shaka, A. J., Lee, C. J., & Pines, A. (1988) *J. Magn. Reson. 77*, 274–293.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
Topal, M. D., & Fresco, J. R. (1976) *Nature 263*, 285–293.
Uesugi, S., Oda, Y., Ikehara, M., Kawase, Y., & Ohtsuka, E. (1987) *J. Biol. Chem. 262*, 6965–6968.
Wagner, G., & Wüthrich, K. (1979) *J. Magn. Reson. 33*, 675–680.
Webster, G. D., Sanderson, M. R., Skelly, J. V., Neidle, S., Swann, P., Li, B. F., & Tickle, I. J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 6693–6697.