Solution Conformation of a Deoxynucleotide Containing Tandem G·A Mismatched Base Pairs and 3′-Overhanging Ends in d(GTGAACTT)$_2$†

Andrew Lane,*,‡ Stephen R. Martin,§ Susanne Ebel,∥ and Tom Brown∥

Laboratory for Molecular Structure and Division of Physical Biochemistry, National Institute for Medical Research, London NW7 1AA, U.K., and Edinburgh Centre for Molecular Recognition, Department of Chemistry, University of Edinburgh, King's Building, West Mains Road, Edinburgh EH9 3JJ, U.K.

Received June 17, 1992; Revised Manuscript Received September 21, 1992

ABSTRACT: We have used ^{31}P and ^{1}H NMR spectroscopy and circular dichroism to define the solution conformation of d(GTGAACTT)$_2$ which contains tandem G·A mismatched base pairs and 3′-overhanging TT ends. Measurements of coupling constants and NOE intensities show that the sugar puckers of the nucleotides are predominantly in the south domain (i.e., near C2′-endo) and that the glycosidic torsion angles are anti. The sequential NOE intensities indicate the presence of a right-handed helix. Analysis of the ^{31}P and ^{1}H NMR spectra of the duplex shows that the tandem mismatch forms a block in which there are unusual backbone torsion angles (i.e., in the B_{II} state), within an otherwise B-like structure. The chemical shift of the N1H of the mismatched guanosine and NOEs between the mismatched base pairs and their nearest neighbors are inconsistent with the imino pairing present in single A·G mismatches or in the X-ray structure of a tandem mismatch [Privé, G. G., et al. (1987) *Science 238*, 498–503] but the data are consistent with the amino pairing found by Li et al. (1991) [Li, Y., et al. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 26–30]. The strong base–base stacking both within the tandem G·A block and between the G·A mismatches and their other nearest neighbors offsets the intrinsic destabilizing effects of the mismatch. Further, the 3′-TT overhangs stack onto the ends of the helix and stabilize the duplex against fraying, which accounts for the observed increase in the melting temperature compared with the flush-ended duplex.

Isolated mismatched A·G base pairs generally destabilize DNA duplexes. For example, the melting temperature of d(CGCAAATTGGCG)$_2$ is about 25 °C lower than in the parent dodecamer d(CGCGAATTCGCG)$_2$ at pH 7 (Brown et al., 1990; Leonard et al., 1990). Several such A·G mismatches have been examined in detail both in the crystal state (Webster et al., 1990) and in solution (Brown et al., 1990; Leonard et al., 1990; Gao & Patel, 1988; Carbonnaux et al., 1991; Lane et al., 1991). The conformation of this mismatch is sensitive to the local sequence context and the solution conditions, such that A(anti)·G(anti), A(anti)·G(syn), and A(syn)·G(anti) pairings have been observed (Webster et al., 1990). The A(anti)·G(syn) state is stabilized by protonation on adenine N1 (Gao & Patel, 1988; Brown et al., 1990), and a slow transition between the G(anti) and G(syn) states occurs with an apparent pK of approximately 6 (Carbonnaux et al., 1991; Lane et al., 1991). The protonated state is thermodynamically more stable than the neutral state (Brown et al., 1990).

In contrast, the thermodynamics and conformational properties of tandem A·G mismatches are less well characterized. However, both in the crystal state and in solution, tandem A·G mismatches have been shown to adopt the A(anti)·G(anti) conformation, and the phosphate backbone adjacent to and within the mismatched sites has been reported to be in the less common B_{II} conformation (Kan et al., 1983; Privé et al., 1987; Nikonowicz & Gorenstein, 1990; Nikonowicz et al., 1991). These studies also showed that the A·G base pairing was in the "imino" form. More recently, Li et al. (1991) have shown that an alternative pairing can exist, in which the amino protons are involved in the hydrogen bonding. This pairing leads to a surprisingly compact B-like structure in which purines of one strand stack on purines of the complementary strand. The amino pairing has recently been confirmed in a different sequence, where it was shown that only the phosphate connecting the mismatched bases is in the B_{II} conformation (Chou et al., 1992).

In the preceding paper in this issue (Ebel et al., 1992), we have shown that tandem G·A mismatches adopt the imino pairing scheme when flanked by a 5′ purine and a 3′ pyrimidine, i.e., 5′ RGAY, whereas the amino pairing scheme is favored in sequences such as YGAR. Also, tandem G·A mismatches, unlike isolated mismatches, stabilize DNA duplexes compared with ones containing normal Watson–Crick base pairs. We report here a detailed analysis of the conformational features of d(GTGAACTT)$_2$ which contains a tandem G·A mismatch and single-stranded 3′-overhanging TT ends. These overhangs significantly stabilize the duplex compared with one that contains only flush ends.

MATERIALS AND METHODS

The oligonucleotides d(GTGAACTT), d(CCAAGATTGG)$_2$, and d(CCAAIATTGG)$_2$ were synthesized by the solid-phase phosphoramidite method and purified by reversed-phase HPLC as previously described (Brown & Brown, 1991).

The duplexes for NMR analysis were formed by dissolving 1–2 μmol in 1 mL of 0.1 M KCl and 0.01 M sodium phosphate, pH 7, containing 0.1 mM EDTA and annealing slowly from 70 °C. The solutions were lyophilized and redissolved in 0.6 mL of 100% D_2O.

† This work was supported by the MRC and the SERC under the Molecular Recognition Initiative.

‡ Laboratory for Molecular Structure, National Institute for Medical Research.

§ Division of Physical Biochemistry, National Institute for Medical Research.

∥ University of Edinburgh.

0006-2960/92/0431-12087$03.00/0 © 1992 American Chemical Society

^{1}H NMR spectra were recorded at 9.4 T and 11.75 T on Bruker AM spectrometers and at 14.1 T on a Varian Unity 600 spectrometer. Chemical shifts were referenced to internal 2,2′-dimethylsilapentane-5-sulfonate. ^{31}P NMR spectra were recorded at 4.7 T on a Bruker WM200 spectrometer and at 9.4 T on a Bruker AM400 spectrometer. ^{31}P NMR spectra were referenced to external methylene diphosphonate at 15 °C. At other temperatures, the spectra were referenced to the internal inorganic phosphate resonance. The ^{1}H spectra were assigned by NOESY[1] and two-quantum filtered COSY and the ^{31}P spectra by heteronuclear correlation spectroscopy as previously described (Forster & Lane, 1990). Spectra in ^1H_2O were recorded using the $1\bar{3}3\bar{1}$ composite pulse (Hore, 1983). Three-bond ^{1}H–^{1}H coupling constants were determined from resolution-enhanced one-dimensional spectra and by comparison of the cross-sections of a long mixing time NOESY spectrum recorded with a digital resolution of 1 Hz per point in F_2 and the DQF-COSY (0.8 Hz per point in F_2). The overall correlation time of the duplex was determined by measuring the cross-relaxation rate constant for the Cyt H6–H5 vector as previously described (Lane et al., 1986; Birchall & Lane, 1990). One-dimensional NOEs were measured using the method of Wagner and Wüthrich (1979). Recycle times of 8–10 s were used for Ade H2 protons to allow adequate time for relaxation.

Circular dichroism spectra were recorded on a JASCO Model J-600 spectropolarimeter using 2-mm cuvettes, at 10 °C. The DNA solutions were approximately 40 μM duplex in buffer containing 0.5 M KCl and 20 mM sodium phosphate, pH 7.4. Under these conditions, the oligonucleotides are in the duplex state. Measured intensities were converted to molar circular dichroic extinction coefficients using the concentration of base pairs.

Simulations of NOEs for given structures and optimizations of nucleotide conformations were performed using the program NUCFIT (Lane, 1990). Initial models were generated either within NUCFIT or using QUANTA (Polygen). No attempt has been made to model the detailed conformation of the phosphate backbone in the absence of sufficient information.

RESULTS

NMR Assignments

^{1}H NMR. The nonexchangeable protons of d(CCAAGATTGG)$_2$ and d(CCAAIATTGG)$_2$ were assigned using NOESY (data not shown). The assignments of the former agree with those reported by Nikonowicz and Gorenstein (1990). Similar chemical shifts were obtained for the inosine-substituted oligonucleotide, except for the modified base itself. However, although the melting temperature, T_m, for the IA oligonucleotide is similar to that of the GA decamer, both the ^{1}H and the ^{31}P NMR spectra of the inosine derivative indicate that there is a mixture of species present (see the preceding paper in this issue). We have not analyzed this complication further. We note, however, that the ^{31}P NMR spectra of the G·A mismatched decamer showed evidence of multiple conformations at low temperatures (Nikonowicz & Gorenstein, 1990).

The oligonucleotide d(GTGAACTT)$_2$ shows a cooperative melting transition at 313 K under the conditions of the NMR experiments (see below). In principle, there are two duplexes,

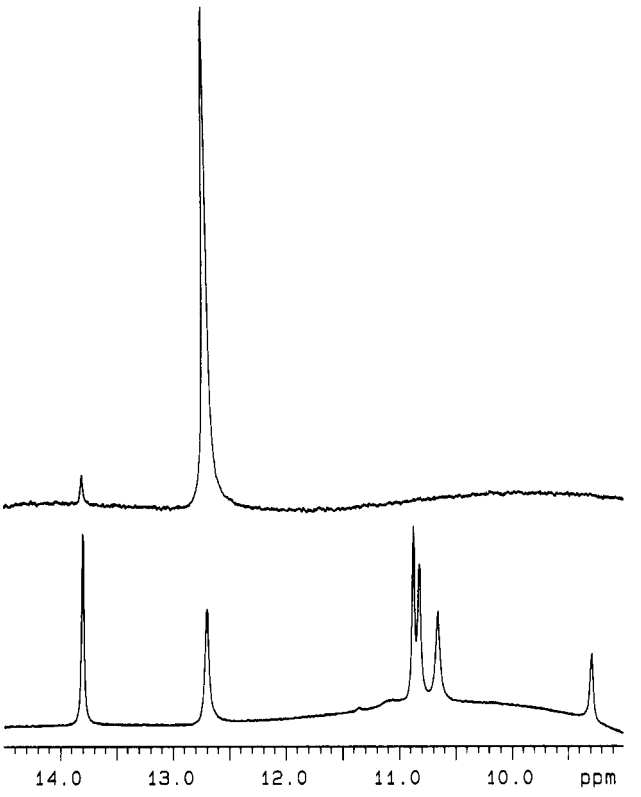

FIGURE 1: Imino proton spectrum of d(GTGAACTT)$_2$. Spectra in ^1H_2O were recorded at 14.1 T and 5 °C as described under Materials and Methods. (Lower spectrum) Relaxed spectrum showing the imino proton resonances. (Upper spectrum) NOE difference spectrum for 300-ms irradiation time. The irradiated peak is GC6 NH at 12.7 ppm.

each containing four Watson–Crick base pairs, that could form at low temperatures, namely

(I) GTGAACTT / TTCAAGTG (II) GTGAACTT / TTCAAGTG

These two possibilities can be readily distinguished by the NMR spectra of the imino protons, as duplex I will have two imino resonances corresponding to the AT and GC base pairs, whereas duplex II should have resonances for AT base pairs only. Figure 1 shows the ^{1}H NMR spectrum of the annealed oligonucleotide dissolved in ^1H_2O at 278 K. There are two resonances of equal intensity in the imino proton region of the spectrum corresponding to GC and AT base pairs. One-dimensional NOE spectroscopy gave NOEs from the peak at 12.7 ppm to the peak at 13.8 ppm, whereas irradiation of the peak at 13.8 ppm gave NOEs to the peaks at 12.7 ppm and a sharp peak at 7.65 ppm (not shown). This latter peak belongs to the H2 resonance of an adenine and must be A5. Hence, the peaks at 12.8 and 13.8 ppm are the imino protons of CG6 and AT5, respectively. Further, cross-strand NOEs (see below) are also consistent only with pairing I. The GC resonance does not broaden significantly between 10 and 25 °C nor does its intensity decrease on saturation of the solvent signal, indicating the presence of a stable, hydrogen-bonded base pair. Finally, the melting temperature of the duplex determined by ultraviolet absorbance is independent of pH in the range 5.5–8 (see the preceding paper), indicating that the duplex is not stabilized by protonation. This is in contrast to C·C base pairs which are stabilized by protonation (Brown et al., 1990) and also to isolated G·A mismatches (Gao & Patel, 1988; Brown et al., 1990; Carbonnaux et al., 1991; Lane et al., 1991).

[1] Abbreviations: NOE, nuclear Overhauser enhancement; DQF-COSY, double-quantum filtered correlated spectroscopy; CD, circular dichroism.

Table I: NMR Assignments[a] for d(GTGAACTT)$_2$

	chemical shift (ppm)									
base	H8/H6	H2/H5 Me	H1′	H2′	H2″	H3′	H4′	H5′/H5″	NH	Pi
G1	7.99		6.01	2.51	2.73	4.76	4.20	3.82	12.7	−17.74
T2	7.11	1.10	6.22	1.47	2.20	4.82	4.32	4.06	13.8	−18.33
G3	8.28		6.17	3.01	2.77	5.14	4.62	4.25/14	10.9	−15.80
A4	7.54	7.97	5.61	0.83	2.25	4.79	4.42	4.11		−17.57
A5	8.30	7.65	5.97	2.77	2.77	4.99	4.45	4.10		−17.56
C6	7.29	5.17	5.93	2.10	2.34	4.73	4.24	4.13		−17.60
T7	7.41	1.72	6.03	2.08	2.33	4.67	4.14	4.01	10.7/10.8	−17.69
T8	7.53	1.67	5.99	2.18	2.18	4.40	3.93	nd	10.8/10.7	

[a] Protons were assigned from NOESY and DQF-COSY spectra as described in the text. The phosphates were assigned using the ^{31}P–^{1}H shift correlation experiment. The ^{1}H chemical shifts are given for 283 K, whereas the ^{31}P shifts are given for 288 K.

There are three resonances between 10.5 and 11 ppm, which must arise from G3N1H and T7N3H, T8N3H, or very downfield shifted amino protons. Imino protons resonate in this region when they are not hydrogen bonded (Li et al., 1991; Williamson & Boxer, 1989) or when they participate in hydrogen bonds with carbonyl oxygen atoms (Cognet et al., 1991). In addition, there is a resonance of unit relative intensity at 9.3 ppm, which is from an amino proton (see below). Irradiation of the most downfield of the group of three resonances gives an NOE at 13.8 ppm (AT5 NH) and to resonances between 7 and 7.3 ppm. Irradiation of the other two resonances gave NOEs only through spillover of power. This suggests that the most downfield resonance, which is also the sharpest and least prone to exchange at higher temperatures, can be assigned to G3N1H, and the other two are the N3H of T7 and T8. The observation of the two thymidine imino protons is consistent with their being stacked on the end of the duplex (see below). Irradiation of the resonance at 9.3 ppm gave NOEs to AT5N3H and to a sharp resonance at 7.65 ppm, which is also obtained on irradiation of AT5N3H. Irradiation at 7.67 ppm gave NOEs at 13.8, 9.3, and 12.8 ppm (weak). The resonance at 7.67 ppm can therefore be assigned to the H2 of A5 (and see below). The proton resonating at 9.3 ppm has an NOE to a broad resonance at 6.6 ppm and is therefore an amino proton either of A5 or from the neighboring A·G mismatch. The assignments of the exchangeable proton resonances are given in Table I.

The nonexchangeable protons of d(GTGAACTT)$_2$ were assigned using a combination of NOESY and DQF-COSY. Figure 2 shows portions of a NOESY spectrum recorded in D_2O with a mixing time of 200 ms. Complete connectivities for the nucleotides could be obtained via the H8/6–H1′–H8/6 pathway and via the H2″ pathway (Reid, 1987; Wüthrich, 1986). This indicates that the molecule adopts a right-handed helical conformation. The assignments of the sugar resonances were confirmed by DQF-COSY (Figure 3), showing that G3 H2′ is downfield of G3H2″ and that H2′ of A4 is shifted unusually upfield. The ^{1}H assignments are given in Table I.

The chemical shifts given in Table I show some anomalies compared with normal duplex DNA. The H8 and H2′ of G3 are shifted unusually far downfield; indeed, the H2′ of G3 is downfield of its H2″, which is usually only found at the 3′ terminus of B-DNA (van de Ven & Hilbers, 1988). In contrast, the H8 and H2′ of A4 are shifted unusually far upfield. As G3 and A4 are the tandem mismatched bases, it is perhaps not surprising that their chemical shifts are unusual, as their stacking with the 5′ and 3′ base pairs is likely to be different from that in normal DNA. This pattern of shifts is distinctly different from that observed in single A·G mismatches such as d(GCCACAAGCTC)·d(GAGCTGTGGC) (Carbonnaux et al., 1991) and for the imino pairing in the tandem G·A mismatch in d(CCAAGATTGG)$_2$ (Nikonowicz & Gorenstein, 1990).

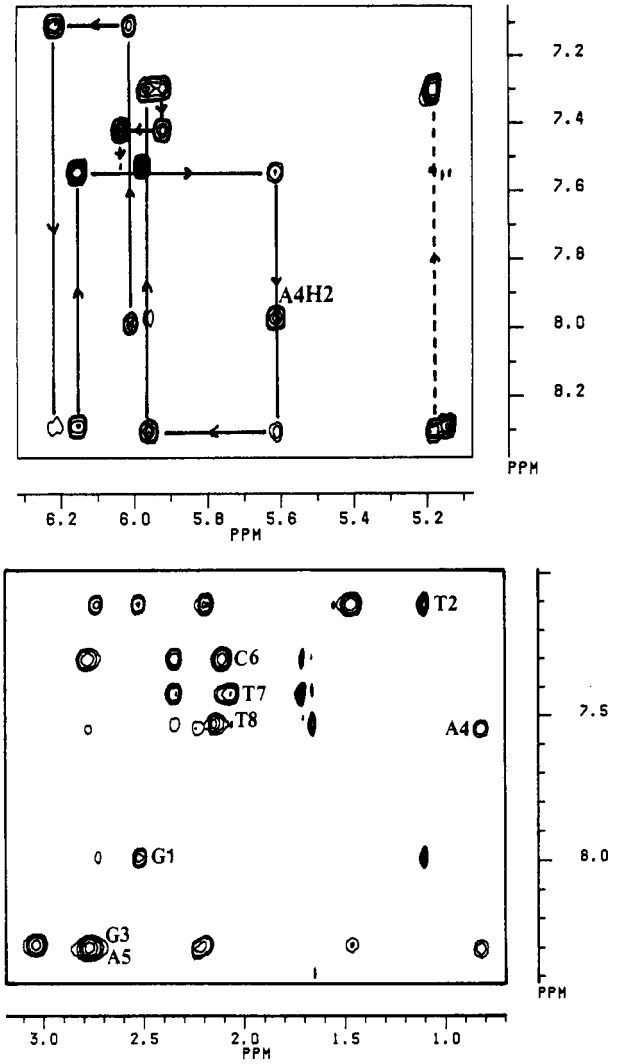

FIGURE 2: NOESY spectra of d(GTGAACTT)$_2$. The NOESY spectrum was recorded at 11.75 T at 10 °C using a mixing time of 200 ms. The data matrix was zero-filled from 2048 by 512 to 4096 by 2048 points and apodized by a 60°-shifted sine-squared function in both dimensions. (A, top) Base proton to H1′ connectivities. The dashed line shows the A5H8–C6H5 interaction. (B, bottom) Base proton to H2′/H2″ connectivities.

A rotational correlation time of 3.6 ± 0.3 ns has been determined at 10 °C for the G·A tandem mismatched octamer, from time-dependent NOEs for the Cyt 6 H6–H5 pair (Lane et al., 1986). This value is comparable to that obtained for a DNA hexamer (Lane & Forster, 1989) and a DNA octamer (Lane, 1991) under similar conditions, indicating that the molecule behaves hydrodynamically as a compact duplex and

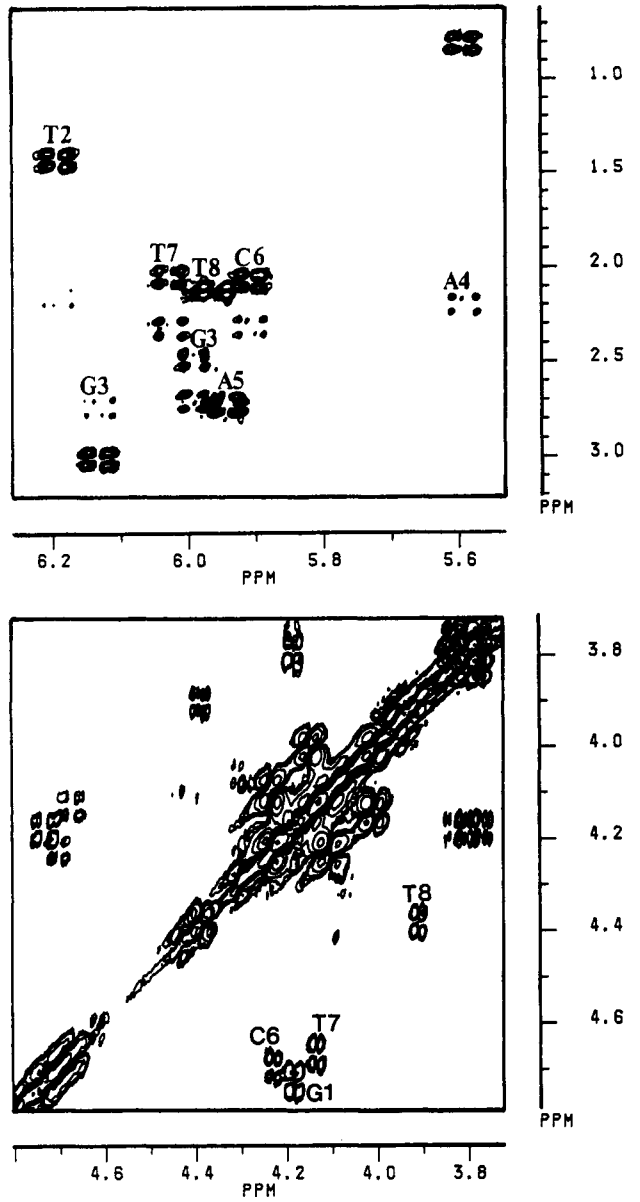

FIGURE 3: DQF-COSY spectrum of d(GTGAACTT)$_2$. The spectrum was recorded at 15 °C at 9.4 T. The data matrix was zero-filled from 4096 by 512 to 8192 by 2048 points and apodized with a sine-squared function shifted by 45 °C in F_2 and 60° in F_1. The spectral width was 3600 Hz in both dimensions. (A, top) H1′–H2′/H2″ region. (B, bottom) H3′–H4′ region. Note that no cross-peaks are present for T2, G3, A4, and A5.

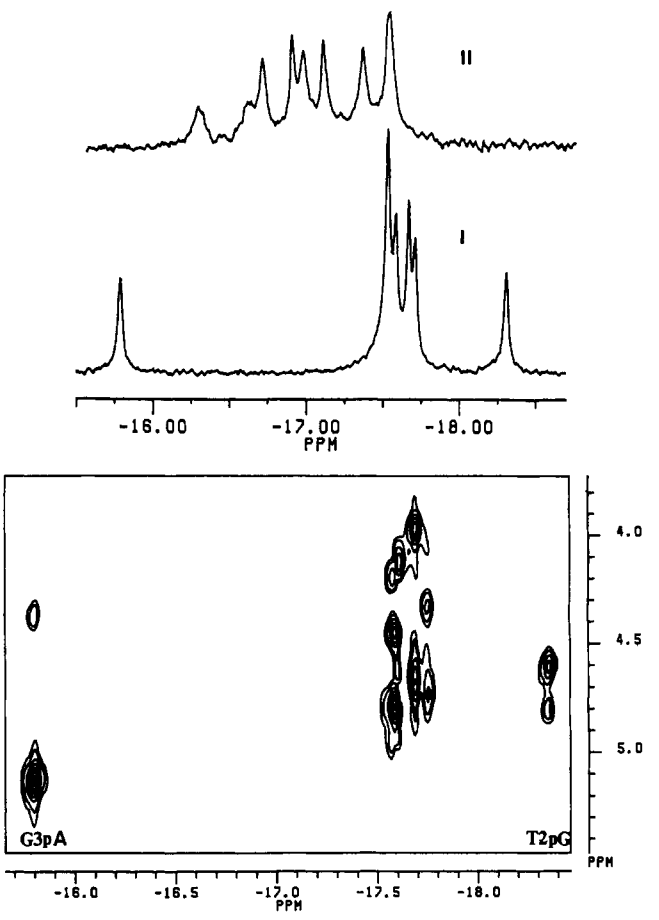

FIGURE 4: ^{31}P NMR spectra. NMR spectra were recorded at 4.7 T and 283 K with proton decoupling provided by WALTZ 16. (A, top) One-dimensional spectra. The spectra were recorded with an acquisition time of 2.024 s and a relaxation delay of 1 s. The free induction decays were zero-filled once and apodized with a 0.5-Hz line-broadening exponential. (I) d(CTGACTT)$_2$. (II) d(CCAAGATTGG)$_2$. (B, bottom) ^{31}P–^{1}H shift correlation spectrum of d(GTGACTT)$_2$. The acquisition times t_{2max} and t_{1max} were 0.5 and 0.051 s, respectively. The data table was filled with zeros to 2048 by 1024 points and apodized with a 60°-shifted sine-squared function in both dimensions before Fourier transformation.

does not aggregate significantly under these conditions.

One-dimensional NOEs were measured at irradiation times from 250 to 750 ms, using a long recycle time to allow relaxation of the adenine H2 protons. NOEs are observed between the resonances at 7.65 and 7.97 ppm. The former resonance was assigned to A5 H2 (see above); hence, A4 H2 must be at 7.97 ppm. A4 H2 also shows a strong NOE to A4H1′, which is only possible as a cross-strand interaction in the A·G mispairing (I). There is also a significant NOE to A5 H8, which implies an unusual stacking arrangement of the tandem G·A mismatched bases. Irradiation of the isolated A4H1′ resonance gave NOEs to A4H8, A5H8, A4H2, and A5 H2. The first two NOEs are normally found in B-like DNA, and the last is found in DNA in which there is a significant base-pair roll angle or propellor twist (Privé et al., 1987; Kintanar et al., 1987). The reverse experiment in which A4H1′ was irradiated confirmed the A4H1′–A4H2 NOE and agrees with the NOESY data (Figure 2). From the cross-relaxation rate constant σ, determined from the build-up rate of the one-dimensional NOEs and the rotational correlation time τ_c (see above), we calculate a distance r of approximately 4 Å using the formula $r = (-56.92\tau/\sigma)^{1/6}$. Clearly, the stacking of the tandem mismatch is of an unusual nature.

^{31}P NMR. ^{31}P NMR spectra are reporters of the phosphodiester backbone, and Gorenstein and co-workers have established correlations between ^{31}P chemical shifts, $^3J_{PH3'}$, and backbone conformation (Roongta et al., 1990). The ^{31}P NMR spectra of d(GTGAACTT)$_2$ and d(CCAAGATTGG)$_2$ are compared in Figure 4A. The phosphorus spectrum of the 8-mer duplex at low temperature has seven resonances, two of which are shifted greatly from the usual position found for B-DNA. In contrast, the spectrum of the decamer shows a much smaller dispersion. The latter spectrum agrees with that reported by Nikonowicz and Gorenstein (1990), who have assigned all of the ^{31}P NMR signals for the decamer. The ^{31}P NMR spectrum of the octamer was assigned using the ^{31}P–^{1}H shift correlation experiment (Figure 4B). Each of the seven phosphodiester resonances correlates with two sugar proton resonances, corresponding to H3′(i) and either H4′ or H5′/H5″ (i+1), which have been assigned (Table I). By matching the shifts of the H3′ and H4′, each phosphate can be assigned. For example, the phosphorus resonance at

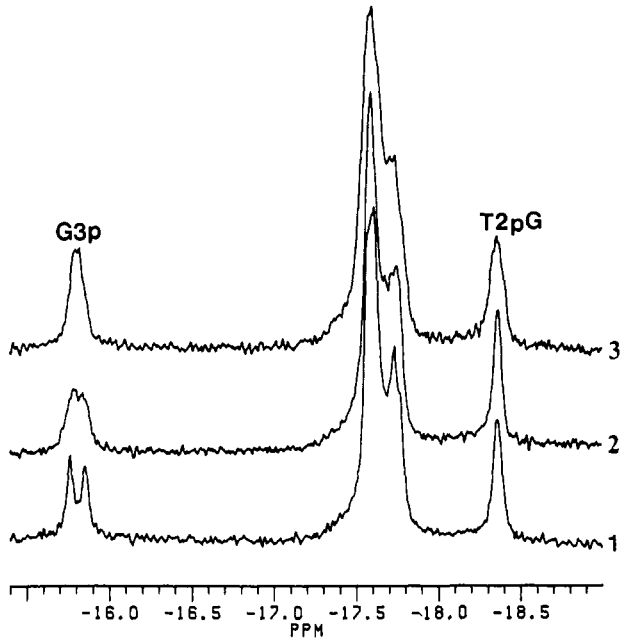

FIGURE 5: Proton–phosphorus coupling in d(GTGAACTT)$_2$. Spectra were recorded at 4.7 T and 15 °C. The spectral width was 700 Hz, and the acquisition time was 1.7 s. Free induction decays were apodized using a Lorentz–Gauss transformation (LB = −1, GB = 0.15). Selective proton decoupling was applied at 1, 4.38 ppm (A4H4′); 2, 4.67 ppm (T7 H3′); 3, 5.14 ppm (G3 H3′).

−15.8 ppm shows an intense cross-peak at 5.14 ppm, which has been assigned to G4H3′, and a weaker cross-peak at 4.42 ppm, which has been assigned to A4H4′ (Table I). Hence, the phosphorus resonance at −15.8 ppm is assigned to G3pA, the phosphate that connects the tandem A·G mismatches. Similarly, the resonance at −18.33 ppm can be assigned to T2pG, the phosphate immediately 5′ adjacent of the mismatches (cf. Table I). These unusual shifts are probably related to uncommon backbone torsion angles (Roongta et al., 1990). Further, the relative intensity of the P–H3′ cross-peaks suggests that there are also differences in the values of $^3J_{H3'P}$.

We have used selective ^{1}H–^{31}P decoupling experiments to obtain estimates of $^3J_{H3'P}$ for the two anomalously shifted phosphates. Figure 5 compares the ^{1}H-coupled and uncoupled ^{31}P NMR spectra. Clearly the apparent width of G3pA is greater than that of T2pG in the ^{1}H-coupled spectrum. By stepping through the frequencies corresponding to H3′, H4′, and H5′/H5″ (see Table I), it was found that decoupling the H3′ leads to only a modest simplification of the coupled spectrum, whereas decoupling the H4′ leads to the appearance of a clear doublet for G3pA, whose splitting is approximately equal to $^3J_{H3'P}$ = 7.50 Hz. Irradiation of H4′ will also cause partial decoupling from H5′ and H5″ by spillover of power, as these resonances are close to H4′ at 200 MHz (see Table I). $^3J_{H3'P}$ for T2pG found by this procedure was <3 Hz. We note that $^3J_{H3'P}$ determined in a similar way and also by two-dimensional NMR methods on a nogalomycin–DNA hexamer in which there is a distorted backbone due to intercalation of the antibiotic (Searle & Lane, 1992) showed that the most downfield-shifted phosphate has a coupling constant significantly larger than the others and was attributed to the presence of the B_{II} phosphate conformation. The present result also qualitatively supports the correlation of chemical shift and coupling constant made by Roongta et al. (1990). Using the parameterization of the Karplus equation given by Roongta et al. (1990), the coupling constants of 7.5 and 3 Hz are consistent with torsion angles ϵ = −90°/−150° and −180° to −220°, respectively. These angles also correlate with the chemical shifts and are consistent with the B_{II} and B_I states, respectively. We conclude that G3pA4 is in the B_{II} state, in agreement with the results reported recently by Chou et al. (1992) for a different DNA sequence containing a tandem G·A mismatch. As discussed in some detail by Privé et al. (1987), the B_I and B_{II} states differ in the torsion angles ϵ,ζ, which are (t,g$^-$) for B_I and (g$^-$,g) for B_{II}.

For the decamer d(CCAAGATTGG)$_2$, both the 5′-adjacent phosphodiester and the intervening phosphodiester resonances were downfield of the remaining resonances (Nikonowicz & Gorenstein, 1990). The magnitude of the shift was, however, substantially smaller than in the present octamer and in the sequence reported by Chou et al. (1992). This confirms that the backbone conformation adjacent to the mismatch site is dependent on the sequence context. It is notable that, in the tandem A·G mismatch studied by Privé et al. (1987), both the 5′-adjacent phosphate and the connecting phosphate were in the B_{II} conformation.

Melting Behavior

While the ^{1}H NMR spectra consist of narrow resonances at 15 °C or below, increasing the temperature to 25 °C or above leads to extensive broadening of many of the resonances (not shown), indicating the onset of melting. We have made use of the resolved T2pG and G3pA ^{31}P resonances to monitor the melting of the duplex under the conditions of the NMR experiments.

Figure 6A shows ^{31}P NMR spectra at several temperatures. The two highly shifted ^{31}P resonances of T2pG and C3pA first broaden and move toward the center of the spectrum on increasing the temperature, showing the onset of melting. The resonances become sharp again at high temperatures, behavior that is typical of intermediate exchange, with a dissociation rate constant at 40 °C comparable to the chemical shift difference, i.e., in the range 300–400 s^{-1}. The dependence of the shifts on temperature is presented in Figure 6B and shows a strongly cooperative transition for both T2pG and G3pA. The T_m under these conditions is about 40 °C, which is higher than that observed optically (see the preceding paper) and reflects the much higher concentration of strands used in the NMR experiments.

DNA Conformation

Circular Dichroism. Near ultraviolet circular dichroism is sensitive to the relative orientations and separations of the dipoles of the bases and therefore to the conformation. Average B-DNA typically shows a conservative CD spectrum with extrema at 278 and 245 nm. We have recorded CD spectra of several oligonucleotides that contain mismatched base pairs. Typical spectra are presented in Figure 7. The parameters that describe these and other spectra are summarized in Table II.

d(GTGAACTT)$_2$ has a conservative CD spectrum in the near UV. However, in comparison to the octamer d(CACTAGTG)$_2$ known from NMR studies to be in the B-conformation (Lane, 1991), the intensity per base pair of the CD is about a factor of 4 higher (see Figure 7 and Table II). This is additional evidence for strong base stacking in an orientation that maximizes the excitation coupling between the bases. In contrast, the CD spectrum of the decamer is inverted (Figure 7), despite being in the B-family of conformations in solution (Nikonowicz et al., 1991). Regardless of the origin of this effect, it is clear that the conformation of the decamer G·A mismatch is significantly different from that of the octamer

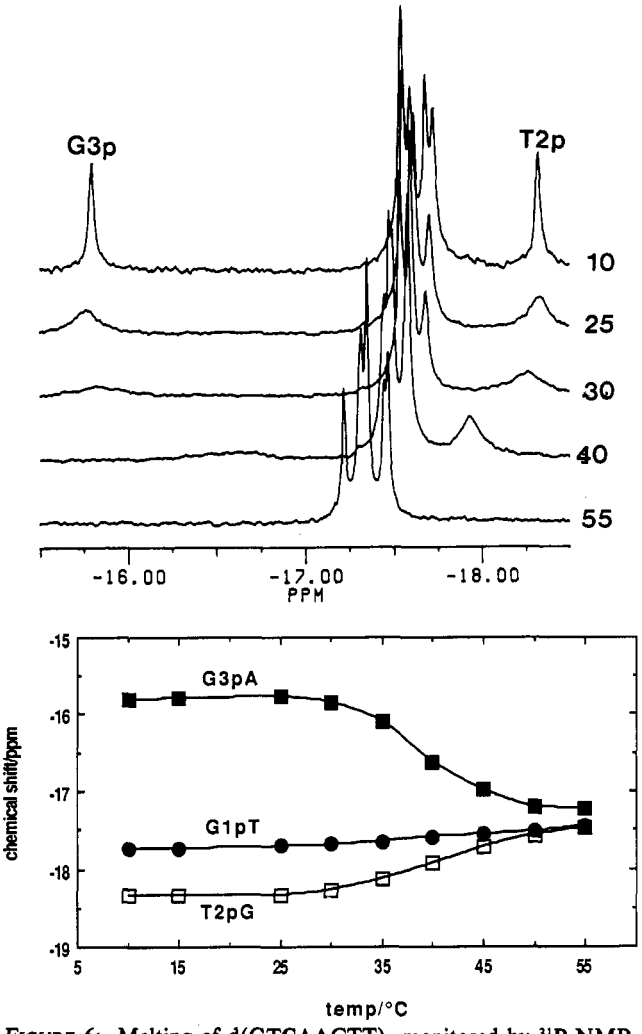

FIGURE 6: Melting of d(GTGAACTT)$_2$ monitored by ^{31}P NMR. (A, top) Dependence of the spectra on temperature. (B, bottom) Variation of the shifts of T2pG and G3pA with temperature.

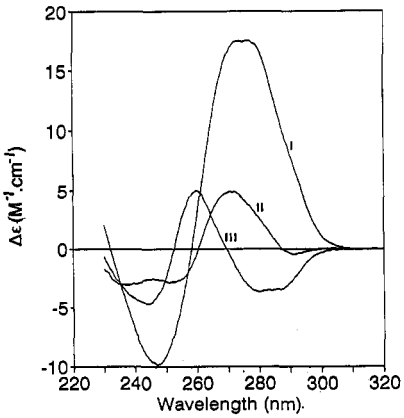

FIGURE 7: Circular dichroism spectra of oligonucleotides. Spectra were recorded at 15 °C as described under Materials and Methods. Intensities are normalized per base pair. Note that, for d(GTGAACTT)$_2$, the number of base pairs was taken to be eight, as the 3′-overhanging nucleotides appear to be stacked (see text). (I) d(GTGAACTT)$_2$, (II) d(CACTAGTG)$_2$, (III) d(CCAAGATTGG)$_2$.

G·A mismatch. In addition, we have recorded the CD spectrum of the dodecamer d(CGCAAATTGGCG)$_2$, which contains two isolated A·G mismatches (Lane et al., 1991) (not shown). The shape of the spectrum is similar to that of B-DNA, though with some differences in the normalized intensities (cf. Table II). It is notable that, according to the NMR data, all of these oligonucleotides are right-handed

Table II: Circular Dichroism[a] of Oligonucleotides

sequence	λ_{max} (nm)	$\Delta\epsilon$ (M^{-1} cm^{-1})	helix type
B-DNA[b]	278	5	B
	245	−6	
A-DNA[b]	260	19	A
	240	−1	
	210	−14	
d(CACTAGTG)$_2$	245	−2.5	B
	272	4.5	
dGTGAACTT	275	17.5	B/LZW
	247	−10	
dCCAAGATTGG	280	−3.5	B/PD
	260	5	
	245	−5	
dCGCAAATTGGCG[c]	280	2.1	B/PD
	250	−7.8	
dCGCAAATTGGCG[d]	282	3.1	B
	250	−7.7	

[a] OD spectra were recorded as described under Materials and Methods. Helix type is taken from the literature. LZW, Li–Zon–Wilson base pairing; PD Privé–Dickerson base pairing. The parameters for B- and A-DNA are taken from Ivanov et al. (1974). [b] From Ivanov et al. (1974). [c] G(anti)·A(anti). [d] G(syn)·A(anti).

double helices, with all glycosidic torsion angles anti and sugar conformations in the C2′-endo domain (except for the A·G mismatched dodecamer at low pH, see Table II).

Nucleotide Conformations from ^{1}H NMR. Three-bond coupling constants were determined in some cases from one-dimensional NMR spectra, and from cross-sections of NOESY and DQF-COSY spectra, as shown in Table III. The coupling constants are expected to be overestimated from the two-quantum COSY data (Wüthrich, 1986) and to be underestimated from the NOESY spectrum. A comparison of the results from the two experiments then leads to better determination of the coupling constants (Lane, 1991). Generally, the values of $\Sigma_{1'}$ determined from the NOESY and COSY experiments are very similar, indicating that the resonance line widths are not dominating the apparent splittings. The coupling constants $\Sigma_{1'}$ for the terminal G1 and C6 are smaller than for the internal residues, consistent either with significant averaging between C2′-endo and C3′-endo states (Rinkel & Altona, 1987) or with a conformation nearer to C3′-exo. Further, it was found that the H2″–H3′ and H3′–H4′ cross-peaks in the DQF-COSY spectrum were absent for residues T2, G3, A4, and A5. These cross-peaks are observed only when the coupling constants exceed ca. 2 Hz. Only if the sugar conformation exists for a significant fraction of the time (ca. >25%) near C3′-endo (Rinkel & Altona, 1987) do the coupling constants become this large. Conformational flexibility in oligomers of this size is perhaps to be expected. Assuming that the sugars exists as a blend of north (near C3′-endo) and south (near C2′-endo) conformers, the values of $\Sigma_{1'}$ can be used to calculate the fraction of the "south" conformation, f_s (Rinkel & Altona, 1987). Residues G1, A4, C6, T7, and T8 have $f_s < 0.75$, whereas the other nucleotides show a higher degree of conformational purity. The values of $^3J_{2''3'}$ and $^3J_{3'4'}$ for T2, G3, and A5 would be 2–3 Hz, calculated assuming the two-state equilibrium and the data given in Table III, whereas the coupling constants calculated for the other residues are 3–4 Hz (cf. Table III) and would be expected to give rise to significant cross-peak intensity in the DQF-COSY experiment, as observed (see above and Figure 3B). The data are consistent with a conformational mix of C3′-endo with C2′-endo, in which the fractions of the C3′-endo state are greatest for residues at the ends of the duplex, especially the unpaired T7 and T8.

Table III: Coupling Constants and Glycosidic Torsion Angles[a] in d(GTGAACTT)$_2$

nucleotide	$\Sigma_{1'}$(N) (Hz)	$\Sigma_{1'}$(C) (Hz)	$\Sigma_{2'}$(C) (Hz)	$\Sigma_{2''}$(C) (Hz)	f_s	P_s (deg)	$^3J_{3'4'c}$ (Hz)	$-\chi$ (deg)
G1	13.5	13.5	27	nd	0.62	170	4	125
T2	15	14.7	29.5	nd	0.84	135	4	110
G3	15	15	29	nd	0.88	162	2.2	95
A4	14	14.5	28	22.5	0.74	170	2.7	130
A5	14.6	15	nd	nd	0.84	(S)	(2.5)	(anti)[b]
C6	13.7	13.6	28	23	0.63	144	4.3	110
								114[c]
T7	14	14	28	24	0.71	155	3.6	120
T8	14	nd	nd	nd	0.71		≈4	125[d]

[a] Coupling constants were determined from one- and two-dimensional spectra. N refers to NOESY and C to DQF-COSY. f_s (fraction of the south conformation) was calculated according to Rinkel and Altona (1987). Glycosidic torsion angles (χ) were determined using measurements of cross-peak intensities as described in the text. $^3J_{3'4'c}$ is the coupling calculated using the values of f_s and P_s in the table, and $P_N = 9°$. [b] Not all NOEs were observed owing to overlap. [c] Value determined from one-dimensional NOEs. [d] Probable averaging about χ.

The intensities of the NOEs between the H8/H6 of the bases and sugar protons are all in the order H2′ > H2″ > H1′ ≈ H3′, which is characteristic of nucleotides in which the dominant sugar pucker is in the south domain and the glycosidic torsion angle is anti. From one-dimensional NOE build-up curves, the glycosidic torsion angle for Cyt 6 was found to be −114 ± 4°. The glycosidic torsion angle determined from the NOESY spectra was −110 ± 5° for this residue. The glycosidic torsion angles for the other residues, based on NOESY cross-peak intensities, are also given in Table III. All values are in the anti range, though there is greater variation than is commonly observed for Watson–Crick duplexes. The angles determined for the overhanging nucleotides in particular are probably subject to significant error owing to the likely greater conformational flexibility of these nucleotides (see above). Nevertheless, both T7 and T8 must spend a substantial fraction of the time with χ near −110° to produce such relatively large intranucleotide NOEs (Lane, 1990). The H8–H2′ NOE for A4 is weaker than for the other nucleotides (cf. Figure 2B). Fitting the data indicates that the glycosidic torsion angle of A4 is more negative than is typical of purines. Although both G3 and A4 are anti, the conformations of the nucleotides in this mismatched base pair are somewhat different from typical B-DNA structures found in solution.

Structure of the Mismatched Base Pairs. The chemical shift of G3N1H is typical of imino protons that do not participate in a hydrogen bond with a nitrogen acceptor (Li et al., 1991; Cognet et al., 1991). This would argue against the G(anti)·A(anti) structure found in the crystal state (Privé et al., 1987) and by NMR (Kan et al., 1983; Gao & Patel, 1988; Carbonnaux et al., 1991; Lane et al., 1991; Nikonowicz & Gorenstein, 1990; Nikonowicz et al., 1991). Further, any structures involving a syn conformation of either mismatched nucleotide can be ruled out on the considerations given above. The structure of the mismatch site is clearly different from the usual B structure, even though the nucleotides are each in the B range of conformations. The chemical shifts of both G3 and A4 are unusual; the H8 of G3 is shifted downfield, and its H2′ resonance is downfield of its H2″. This is commonly observed for 3′-terminal bases and indicates that the sugar moiety of G3 lies farther out of the ring-current field of the 3′ neighbor than in typical B-DNA (van de Ven & Hilbers, 1988). In contrast, the H8 of A4 is unusually upfield shifted, as are the H1′, H2′, and H2″ resonances. This suggests that A4 has moved into the ring-current field of its nearest neighbors. The presence of an NOE between A4H2 and A5 H2 indicates that A4H2 is in the minor groove and stacks onto A5. This is confirmed by the presence of significant NOEs A4H1′, H2′, and H2″ to A5 H8. Similarly, G3 stacks onto T2 in a right-handed helical sense as there are significant NOEs between G3H8 and T2H6, H1′, H2′, and H2″ (see Figure 2B). In contrast, the NOEs from G3H2′/H2″ to A4H8 are weak, while that between G3H1′ and A4H8 is strong (Figure 2). A significant NOE between A5NH$_2$ and G3 H1′ and a weak NOE between G3N1H and G3 H8 are not expected in B-like structures. Many of these observations are similar to those made by Li et al. (1991). The detailed consideration of all of the data leads us to conclude that the mismatches adopt the amino pairing scheme as first shown by Li et al. (1991).

Molecular Conformation. Although continuous sequential NOEs are observed along the base proton–H1′ pathway, there are two steps at which the base–H2′/H2″ pathway produces only weak NOEs. These are the G3–A4 and T7–T8 steps (cf. Figure 2). However, the T2–G3 and A4–A5 steps have substantial sequential NOEs, indicating the presence of helical stacking, whereas the NOEs from one mismatched base pair to the next are weak. The weak NOEs connecting T7 and T8 can probably be attributed to significant fraying of T8, though the connectivities from C6 to T7 indicate that T7 must be stacked on C6 most of the time. This conclusion is supported by the observation that the G1–C6 imino proton shows no evidence of fraying as the temperature is raised from 10 to 20 °C or on saturation of the solvent resonance. This is in contrast to the NH resonance of terminal base pairs in oligonucleotides containing flush ends, where the imino proton is frequently very broad or not observable (Patel & Hilbers, 1975).

We have therefore undertaken a modeling study of the gross conformation of the tandem mismatch starting from the structure observed in the crystal state (Privé et al., 1987) and from that proposed by Li et al. (1991).

Imino Base Pairing (Privé et al., 1987)

Using the nucleotide parameters given in Table III, we have calculated the expected NOEs in the mismatched duplex as a function of the helical parameters. In addition, the distance C3′(i)–C5′(i+1) and the Lennard-Jones energies are calculated, which can be used to filter out "impossible" structures. Consideration of the G·A/G·A block has shown that both the G and A residues each need to slide (D_y) (Dieckmann, 1989) about 1 Å from the helix axis to achieve an acceptable stereochemistry. Similarly any shift (D_x) needs to be the same for both bases and relatively small (<1 Å). Steric clashes are then avoided by significant propellor twists of the base pairs. This model predicts essentially normal helical NOEs, including those for G–A. On detailed analysis of the observed NOE intensities and the patterns of chemical shifts, we find that the imino pairing model is not present in this case. This conclusion agrees with that made on the basis of

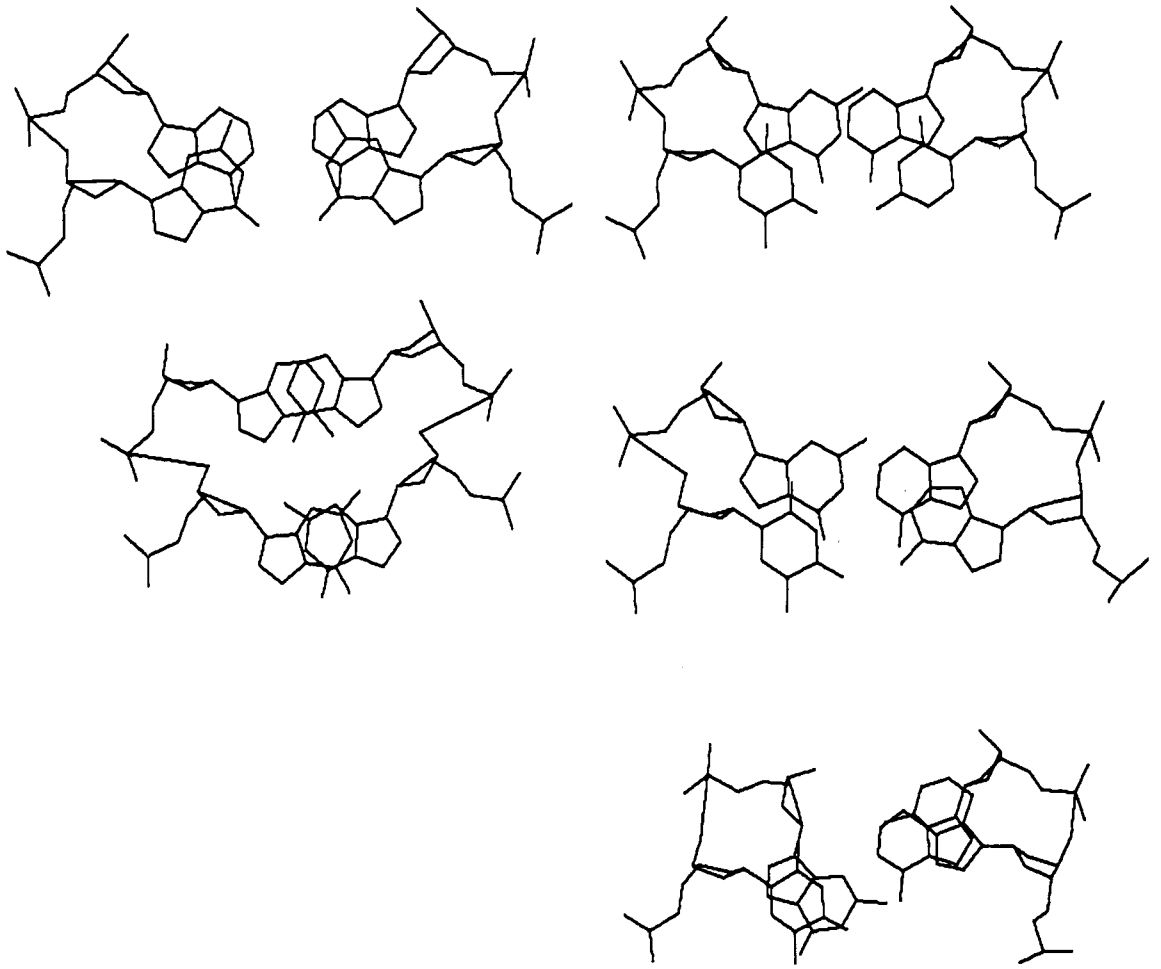

FIGURE 8: Comparison of base-pair overlaps in the amino and imino base pairings. Structures were generated as described in the text, using the nucleotide conformations given in Table III for the amino- and imino-paired mismatches. The positions of the phosphates are arbitrary. Protons have been omitted. (A, left) GA/AG stack: upper, imino-paired bases; lower, amino-paired bases. (B, right) TG/AA stack: top, B-DNA (with GC base pair); center, imino pairing; bottom, amino pairing.

the effect of inosine substitution on helix stability (see the preceding paper).

Amino Pairing (Li et al., 1991)

The amino pairing of G to A can be achieved from a Watson–Crick conformation by translation along the x-axis (D_x) of the adenosine nucleotide by about 7 Å and a translation along the y-axis (D_y) of about 4 Å. The stacking of the next base pair A on A and G on G requires a small negative twist angle of the 3′ nucleotide and a shift in the opposite direction. To place the G·A/A·G block inside a normal B helix requires a positive shift and a small positive slide with respect to the axis of the surrounding B-helix. A large twist (approximately twice the normal value found in B-DNA) has to be incorporated to connect the G·A block to the base pairs above and below. Note that if the helical twist angle is taken as the angle between vectors connecting the C1′ of the mismatched base, then the twist angle for G–A is large and positive (approximately 70°), as described by Chou et al. (1992). Structures of this kind were built such that the distance between C3′(i) and C5′(i+1) did not exceed 5 Å, the tandem G·A block was symmetrically placed within the normal B-helix, and negative Lennard-Jones energies were similar to those calculated for a B-DNA helix. Further, the tandem G·A block was constructed to retain symmetry of the two strands, as required for this sequence. Under these conditions, all of the NOEs can be qualitatively satisfied, which is not true for the Privé structure. For example, the NOEs predicted for A4H8 to G3 H2′, H2″, and H3′ are very weak, in contrast to the Privé structure. Similarly, the NOEs predicted for G3H8 to T2H2′ and H2″ are relatively large (comparable to those expected in B-DNA), as is observed (cf. Figure 2). However, because of the symmetry of the sequence and spectral overlap, we have insufficient independent data to determine all of the backbone and helical parameters reliably. Nevertheless, the approximate stacking of successive base pairs can be determined. The stacking of the base pairs for B-DNA, the structure of Privé et al. (1987), and the amino-paired structure are shown in Figure 8. Remaining base pairs are essentially normal B-type stacks. Particularly noticeable is not only the stacking of G·A over A·G but also the stacking of this block onto the adjacent base pairs. This stacking is more extensive than in canonical B-DNA.

DISCUSSION

The stacking energy of an A·G base pair on a G·A base pair in an optimal geometry would be expected to be more favorable than the stacking of purine–pyrimidine base pairs (Saenger, 1984; Rinkel et al., 1987). While the stacking of one G·A mismatch with its neighbor is strong in both the amino and imino pairing, it is notable that the stacking of the tandem G·A block with its nearest neighbors is actually more extensive than ordinary base pair stacking in B-DNA (Figure 8A). Further, the bulging in the amino pairing is less pronounced than in the imino-paired mismatch (cf. Figure 8B), in which conformational rearrangements are propagated two base pairs

from the mismatch sites (Privé et al., 1987; Nikonowicz & Gorenstein, 1990). The improved stacking may overcome the unfavorable effects of the bulge created by any mismatched purines and partly explain why the amino pairing gives the most stable duplexes. The tandem amino-paired G·A mismatches appear to act as quasi-independent blocks, as they can easily be slotted into a normal B-like duplex. This may account for the observed stability of a dodecamer containing 6 G·A mismatches of this kind (see the preceding paper). Further, the 2-amino group of guanine can hydrogen bond to two water molecules in the single-stranded slate, but in the imino-paired structure only one of the protons can hydrogen bond to water, resulting in a net loss of one hydrogen bond. In contrast, in the amino-paired structure, one amino proton is involved in a hydrogen bond with the adenine base, and the other is free to hydrogen bond to water, with no net change in the number of hydrogen bonds.

The 3′-overhanging ends significantly stabilize the duplex state; the melting temperature increases about 5 °C (see the preceding paper). According to Senior et al. (1988), 5′ overhangs are more stabilizing than 3′ overhangs. The stabilization was attributed to an enthalpic effect from the overhanging bases stacking on the helix. This stacking of the single-stranded overhangs would decrease the fraying or tendency to melt from the ends, as observed in the spectra of the imino protons (see above). It may also account in part for the high frequency of pairing of "sticky" ends of restriction fragments. A similar result was obtained recently in RNA duplexes, where a single 3′-overhanging base caused a substantial stabilization of the duplex (Santa Lucia et al., 1990); the stabilization clearly depends on the overall structure of the helix.

ACKNOWLEDGMENT

We thank Dr. K. Borden for helpful comments on the manuscript.

REFERENCES

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.
Birchall, A. J., & Lane, A. N. (1990) *Eur. Biophys. J. 19*, 73–78.
Brown, D. J. S., & Brown, T. (1991) in *Oligonucleotides and Analogues, A Practical Approach* (Eckstein, F., Ed.) pp 1–24, IRL Press, Oxford.
Brown, T., Booth, E. D., & Leonard, G. A. (1990a) in *Molecular Mechanisms in Bioorganic Processes* (Bleasdale, G., & Golding, B. T., Eds.) Royal Society of Chemistry Press, London.
Brown, T., Leonard, G. A., Booth, E. D., & Kneale, G. (1990b) *J. Mol. Biol. 212*, 437–440.
Carbonnaux, C., van der Marel, G. A., van Boom, J. H., Guschlbauer, W., & Fazakerly, G. V. (1991) *Biochemistry 30*, 5449–5458.
Chou, S.-H., Cheng, J.-W., Fedorff, O. Y., Chuprina, V. P., & Reid, B. R. (1992) *J. Am. Chem. Soc. 114*, 3114–3115.
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.
Dieckmann, S. (1989) *EMBO J. 8*, 1–4.
Ebel, S., Lane, A. N., & Brown, T. (1992) *Biochemistry* (preceding paper in this issue).
Forster, M. J., & Lane, A. N. (1990) *Eur. Biophys. J. 18*, 347–355.
Gao, X., & Patel, D. J. (1988) *J. Am. Chem. Soc. 110*, 5178–5182.
Hore, P. J. (1983) *J. Magn. Reson. 55*, 283–300.
Ivanov, V. I., Minchenkova, L. E., Minyat, E. E., Frank-Kamenetskii, M. D., & Schyolkina, A. K. (1974) *J. Mol. Biol. 87*, 817–833.
Kan, L.-S., Chandrasegaran, S., Pulford, S. M., & Miller, P. S. (1983) *Proc. Natl. Acad. Sci., U.S.A. 80*, 4263–4265.
Kintanar, A., Klevit, R. E., & Reid, B. R. (1987) *Nucleic Acids Res. 14*, 5845–5862.
Lane, A. N. (1990) *Biochim. Biophys. Acta 1049*, 189–204.
Lane, A. N. (1991) *Carbohydr. Res. 221*, 123–144.
Lane, A. N., & Forster, M. J. (1989) *Eur. Biophys. J. 17*, 221–232.
Lane, A. N., Lefévre, J.-F., & Jardetzky, O. (1986) *J. Magn. Reson. 66*, 201–218.
Lane, A. N., Jenkins, T. C., Brown, D. J. S., & Brown, T. (1991) *Biochem. J. 279*, 269–281.
Leonard, G. A., Booth, E. D., & Brown, T. (1990) *Nucleic Acids Res. 18*, 5617–5623.
Li, Y., Zon, G., & Wilson, W. D. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 26–30.
Neuhaus, D., Wagner, G., Vasak, M., Kägi, J. H. R., & Wüthrich, K. (1985) *Eur. J. Biochem. 151*, 257–273.
Nikonowicz, E. P., & Gorenstein, D. G. (1990) *Biochemistry 29*, 8845–8858.
Nikonowicz, E. P., Meadows, R. P., Fagan, P., & Gorenstein, D. G. (1991) *Biochemistry 30*, 1323–1334.
Patel, D. J., & Hilbers, C. W. (1975) *Biochemistry 14*, 2651–2656.
Privé, G. G., Heinemann, U., Chandrasegaran, S., Kan, L.-S., Kopka, M. L., & Dickerson, R. E. (1987) *Science 238*, 498–503.
Reid, B. R. (1987) *Q. Rev. Biophys. 20*, 1–34.
Rinkel, L. J. & Altona, C. (1987) *J. Biomol. Struct. Dyn. 4*, 621–649.
Rinkel, L. J., van der Marel, G. A., van Boom, J. H., & Altona, C. R. (1987) *Eur. J. Biochem. 166*, 87–101.
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.
Santa Lucia, J., Kierzek, R., & Turner, D. H. (1990) *Biochemistry 29*, 8813–8819.
Senior, M., Jones, R. A., & Breslauer, K. J. (1988) *Biochemistry 27*, 3879–3885.
Searle, M. S., & Lane, A. N. (1992) *FEBS Lett. 297*, 292–296.
Schmitz, U., Zon, G., & James, T. L. (1990) *Biochemistry 29*, 2357–2368.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
van de Ven, F. J. M., & Hilbers, C. W. (1988) *Nucleic Acids Res. 13*, 5713–5726.
Wagner, G., & Wüthrich, K. (1979) *J. Magn. Reson. 33*, 675–680.
Webster, G. D., Sanderson, M. R., Skelly, J. V., Neidle, S., Swann, P. F., Li, B. F., & Tickle, I. J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 6693–6697.
Williamson, J. R., & Boxer, S. G. (1989) *Biochemistry 28*, 2819–2831.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.

Registry No. d(GTGAACTT), 144156-86-5; 5′-TMP, 365-07-1; guanine, 73-40-5; adenine, 73-24-5.