Solution Structure of Two Mismatches G·G and I·I in the K-*ras* Gene Context by Nuclear Magnetic Resonance and Molecular Dynamics†

Valerie Faibis,‡ Jean A. H. Cognet,§ Yves Boulard,‡ Lawrence C. Sowers,|| and G. Victor Fazakerley*,‡

CEA, Service de Biochimie et de Génétique Moléculaire, Bât. 142, Département de Biologie Cellulaire et Moléculaire, CEA-Saclay, 91191 Gif-sur-Yvette Cedex, France, Laboratoire de Physico-chimie Biomoléculaire et Cellulaire, CNRS URA 2056, Université Pierre et Marie Curie (T22), 4 Place Jussieu, 75252 Paris Cedex 05, France, and Division of Pediatrics, City of Hope National Medical Center, 1500 East Duarte Road, Duarte, California, 91010

Received April 11, 1996; Revised Manuscript Received September 5, 1996⊗

ABSTRACT: Two mismatches, G·G and I·I, have been incorporated at the central position of 5′d-(GCCACXAGCTC)·d(GAGCTXGTGGC) in order to carry out NMR and molecular dynamics studies. These duplexes constitute the sequence 29−39 of the K-*ras* gene coding for the glycine 12, a hot spot for mutation. The NMR spectra show that the duplexes are not greatly distorted by the introduction of the mismatches and their global conformation is that of a canonical B-form double helix. For the duplex containing the G·G mismatch, we propose for the major species, a type of pairing involving one hydrogen bond between the imino group of one central guanine and the carbonyl group of the opposite guanine. Both bases are in an *anti* conformation. Two conformations, with the same donor and acceptor pattern can coexist, one is obtained from the other by a 180° rotation about the pseudodyadic axis. Exchange between the two forms is observed by NMR at low temperature. A minor species involving hydrogen bonding between the guanine amino group and the carbonyl group of the guanine on the opposite strand may also exist as shown by the molecular dynamics calculations. For the I·I mismatch we observe the same major species, i.e., hydrogen bonding between an imino proton of one base and the carbonyl group of the base on the opposite strand with both bases in an *anti* conformation. Exchange between these two conformations is faster than for the G·G mismatch. Further, we observe that the I·I mismatch adopts a minor conformation, in which one or other of the bases is in the *syn* conformation.

The formation of stable mismatched base pairs can occur as a consequence of errors in genetic recombination and replication. If these mispairs are not corrected by the repair enzyme systems, they can lead to mutations, and thus the correction of the mismatched base pairs is a necessity to maintain the integrity of the genetic information. Nevertheless, all mismatches are not repaired with the same efficiency: purine−purine mismatches are among the most efficiently repaired whereas the pyrimidine−pyrimidine pairs are not well repaired (Fazakerley et al., 1986; Kramer et al., 1984). Nevertheless, G·G mismatches are generally better repaired than A·G mismatches. This may be because A·G mismatches can adopt conformations which are not recognized by the repair systems. Further, the efficiencies of repair have been found to be sequence dependent. Repair enzyme systems seem to be capable of recognizing mismatch structures or local flexibility.

Studies on oligonucleotides containing mispaired bases have been used to determine the structure of the mismatch and the hydrogen bonds involved in the pairing. For the G·G mismatch, Topal and Fresco (1976) suggested base pairing involving one guanine in a keto form and in a *syn* conformation and the other guanine in an enol form and in an *anti* conformation. A ^{31}P study (Roongta et al., 1990) found no evidence for a major modification of the backbone. Recent studies have only proposed mispairing in the major tautomer forms. A G(*syn*)·G(*anti*) mismatch involving symmetrical hydrogen bonding, N1H−O6, has been reported (Cognet et al., 1991) in which the G(*anti*) is between two pyrimidines in the sequence and the G(*syn*) between two purines. A different G(*syn*)·G(*anti*) conformation has been proposed (Castati et al., 1994) in which hydrogen bonding occurs between N1H−O6 and N2H−N7 where the G·G mispair was studied between two adjacent G·A mispairs. However, for a similar environment, in a study of the human centromere which contains the $(GGA)_2$ motif, it has been observed that the central guanines are not base paired but stack over each other (Chou et al., 1994). In the same study, replacement of the central G by I was shown to result in the same type of structure. A weakly hydrogen bonded structure with a G(*anti*)·G(*anti*)conformation has been proposed (Borden et al., 1992) as has a rapid equilibrium on a proton NMR time scale between G(*syn*)·G(*anti*) and G(*anti*)·G(*syn*) conformations (Lane & Peck, 1995) in similar sequence environments. The structure of an G·G mispair in a dodecamer has been solved by X-ray crystallography, and two G(*anti*)·G(*syn*) mispairs were observed (Skelly et al., 1993). One pairing contains bifurcated hydrogen bonds between GNH1 with GN7 and GO6 and the other GN7 with GNH1 and GN2H. Another study concerning a G·G mismatch in an oligoribonucleotide (Peterson et al., 1994) showed G(*syn*)·(*anti*) pairing. In an I·A mismatch, I(*anti*)·

† This work was supported in part by National Institutes of Health Grants GM41336 and CA33572. V.F. thanks Brüker Spectrospin for financial support.

* Author to whom correspondence should be addressed.
‡ CEA-Saclay.
§ CNRS URA 2056.
|| City of Hope National Medical Center.

⊗ Abstract published in *Advance ACS Abstracts,* November 1, 1996.

S0006-2960(96)00871-9 CCC: $12.00

A(*anti*) pairing has been reported (Uesugi et al., 1987), whereas for the I·G mismatch, I(*syn*)·G(*anti*) was observed (Oda et al., 1991) .

It would appear that the sequence environment around a G·G or I·I mismatch can strongly influence the conformation adopted by the mismatch for which the different conformations must differ very little in energy.

In this paper, we report the solution structure by NMR[1] and molecular dynamics of oligonucleotides containing G·G or I·I mismatches in the central position of the duplex. The sequence corresponds to the 29−39 sequence of the K-*ras* gene coding for Gly12, a hot spot for mutation. Both mismatches are incorporated in the position of the first base of the Gly 12 codon.

EXPERIMENTAL PROCEDURES

NMR Spectroscopy. The oligonucleotides corresponding to the sequence 29−39 of the K-*ras* gene were synthesized by a classical phosphotriester method (van der Marel et al.; 1981, Marrug et al., 1984). The pair of oligonucleotides was heated to 80 °C followed by slow cooling to form the duplexes. The sequences of the G·G and I·I duplexes are given below:

G1	C2	C3	A4	C5	**X6**	A7	G8	C9	T10	C11
C22	G21	G20	T19	G18	**X17**	T16	C15	G14	A13	G12

with X = G or I.

The duplexes were 3 mM in single-strand concentration dissolved in 10 mM phosphate buffer, 150 mM NaCl, and 0.1 mM EDTA. Chemical shifts were measured relative to the internal reference tetramethylammonium chloride, 3.18 ppm. NMR spectra were recorded on either a Bruker AMX500 or Bruker AMX600 spectrometer. NOESY spectra were recorded with mixing times of 40, 60, 80, 100, and 400 ms in D_2O and 250 ms in H_2O, in the phase sensitive mode (Bodenhausen et al., 1984). The residual HDO resonance was presatured during the relaxation delay for spectra in D_2O. After zero-filling, the data were multiplied by a 5−15° shifted sine bell function in both dimensions. Distance determination from the initial NOE build-up curves was done as described previously (Cuniasse et al., 1987). Additional 1D difference spectra in H_2O were recorded with presaturation times in the range of 200−400 ms. For the 1D and 2D spectra in H_2O, a jump and return sequence was used to suppress the water signal (Plateau & Guéron, 1982). The pulse maximum was at 15 ppm. TOCSY experiments were recorded in the phase sensitive mode (Davis & Bax, 1985) with a 60 ms MLEV-17 spin lock. 1D ^{31}P spectra were recorded between 1 and 36 °C.

Model Building, Molecular Dynamics, and Helical Parameters. All initial structures were generated from canonical B-DNA (Arnott et al., 1976). Homobase mismatches with non-equivalent hydrogen bonding can exist in two different conformations. One is obtained fron the other by a 180° rotation about the pseudodyadic axis. As for the A·A and T·T mispairs (Gervais et al., 1992), we refer to the two conformations of the G·G mispair: W↑ and W↓. In this notation, W stands for wobble pairing: W↑ with hydrogen bond G6NH1−G17O6 (Figure 4E) and W↓ with hydrogen bond G6O6−G17NH1 (Figure 4F). Identical notation is used for the I·I mismatch in Figure 4I and J. The W↑ refers to the guanine of the first strand, G6, that is the most displaced in the major groove, i.e., in the positive direction along the *x* axis of the EMBO convention (Dickerson, 1989). In the oligonucleotides with mismatches in W↑ or in W↓ conformations, all bases are in the *anti* conformation and all sugars are in the C2′ *endo* conformation. The reference frames are based on hydrogen bonding and were used for building all initial models and for measuring helical parameters (Cognet et al., 1995; Gervais et al., 1995). Note that the shear parameter is arbitrarily set to zero and that the stretch parameter corresponds to the interbase separation for the G·G and I·I mismatches. Energy minimization and molecular dynamics (MD) were carried out with the program AMBER (Weiner et al., 1984) as described previously (Boulard et al., 1995, Boulard et al., 1992; Cognet et al., 1990; Gervais et al., 1995). All MD runs were performed over 200 ps with restraints on all δ torsion angles to reinforce the C2′ *endo* conformation. MDI runs include all NMR distances used as restraints; MDII runs were computed without distance restraints. Helical parameters were computed as summarized before (Gervais et al., 1995) and as explained in detail previously (Cognet et al., 1995).

RESULTS AND DISCUSSION

We have recorded 1D proton spectra between pH 4.5 and 9 for both duplexes, and we did not observe any significant changes in the spectra. This indicates that there is no pH dependent conformational transition.

G·G Duplex

Nonexchangeable Protons. In order to determine the optimum experimental conditions for the duplex, we have measured 1D proton spectra, at pH 6, as a function of temperature. We observe a global melting temperature (T_m) of 57 °C for the duplex. We therefore chose to study the duplex at 37 °C.

The assignment of the nonexchangeable protons resonances was obtained from analysis of a NOESY spectrum recorded with a 400 ms mixing time and a TOCSY spectrum in D_2O. The sequential NOE connectivities between aromatic and anomeric DNA protons at 37 °C are shown in Figure 1 (lower). Strong cross peaks, marked X, corresponding to the CH5−CH6 interactions, are observed which identified the cytosines. The H8 resonances of the 5′-terminal residues (G1 on the first strand and G12 on the second one) are characterized by only one cross peak, that with their own H1′. Starting from G1H8 at 7.92 ppm, we can follow the chain of connectivities to the C5H1′. In Figure 1 (lower) the cross peak G6H8−C5H1′ appears to be missing. This is due to the adjacent intense C3H6−C3H5 peak. This interaction is observed when the spectrum is not resolution enhanced. After the C5−G6 step, the chain of connectivities can be followed without any interruption.

For the second strand, the sequential NOEs were uninterrupted. Nevertheless, we observe only a weak cross peak between G17H8 and T16H1′ compared to the same type of interaction at other steps. A similar phenomenon has been observed for the TpG step in other duplexes (Oda et al., 1991;

[1] Abbreviations: NMR, nuclear magnetic resonance: NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; 1D, one-dimensional; 2D, two-dimensional; TOCSY, total correlation spectroscopy; MD, molecular dynamics.

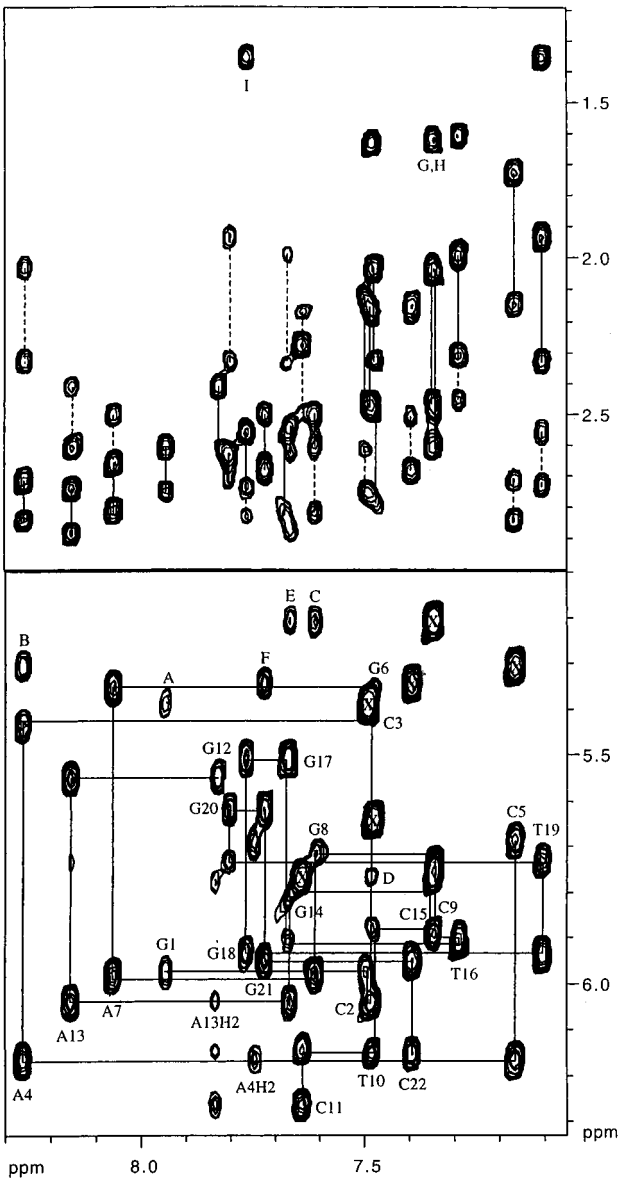

FIGURE 1: Expanded contour plot of the H6/H8−H1′/H5 (lower) and H6/H8−H2′/H2″/CH_3 (upper) regions of the NOESY spectrum (400 ms mixing time) of the G·G duplex in D_2O, 37 °C and pH 6. Cross peaks marked with an X correspond to CH6−CH5 interactions. Cross peaks A to F correspond to H6/H8−CH5 [(*n*)−(*n*+1)] interactions, and cross peaks G to I correspond to H6/H8−CH_3 [(*n*)−(*n*+1)] interactions.

Quignard et al., 1986). The H8/H6−H5 [(*n*)−(*n*+1)] interactions are observed for the A4−C5, G8−C9, G14−C15, and G21−C22 steps. We also observe AH2−AH1′ interactions which are characteristic of a right-handed B-DNA.

The H8/H6−H2′/H2″ region of the NOESY spectrum, shown in Figure 1 (upper), and the H1′−H2′/H2″ region (not shown) confirm the assignments. For the first strand, we observe normal intra- and interresidue interactions from G1 to C5. Similarly, the connectivities can be followed without interruption from C11 to A7 which shows a cross peak with the G6H2′. Almost all of the G6 protons are coincident with other resonances, and their chemical shifts were found from interactions with the neighboring residues. The chemical shift of the G6H8 proton was found from the A7H8−G6H8 cross peak in the aromatic−aromatic region, and the resonances of the G6H2′/H2″ were assigned unambiguously from

Table 1: Chemical Shifts of Nonexchangeable Protons at 37 °C and of Exchangeable Protons at 16 °C for the G·G Duplex, pH 6

	H8/6	H2/H5/CH_3	H1′	H2′	H2″	H3′	H4′	NH	NH2
G1	7.92		5.94	2.61	2.75	4.81	4.24		
C2	7.46	5.37	6.04	2.13	2.45	4.84	4.20		6.42/8.34
C3	7.48	5.63	5.41	2.03	2.30	4.83	4.08		6.81/8.55
A4	8.25	7.72	6.16	2.72	2.85	5.00	4.40		6.20/7.70
C5	7.14	5.30	5.67	1.73	2.15	4.68	4.24		6.41/8.09
G6	7.45		5.34	2.79	2.51	4.84	4.18	10.42	6.22[a]
A7	8.05	7.50	5.98	2.66	2.81	5.01	4.38		6.16/7.38
G8	7.59		5.70	2.51	2.57	4.89	4.36	12.82	
C9	7.33	5.28	5.91	2.05	2.59	4.68	4.13		6.50/8.07
T10	7.44	1.60	6.14	2.15	2.51	4.84	4.12	14.12	
C11	7.58	5.70	6.26	2.28	2.28	4.53	3.90		
G12	7.82		5.54	2.43	2.63	4.81	4.17		
A13	8.13	7.80	6.04	2.74	2.89	5.04	4.40		6.12/7.89
G14	7.63		5.79	2.56	2.60	4.94	4.40	12.87	
C15	7.32	5.19	5.90	2.05	2.46	4.68	4.21		6.50/8.07
T16	7.26	1.60	5.89	1.99	2.31	4.81	4.12	13.84	
G17	7.63		5.49	2.83	2.56	4.94	4.40	10.42	6.40[a]
G18	7.75		5.93	2.57	2.76	4.93	4.37	12.64	
T19	7.08	1.35	5.71	1.93	2.32	4.82	4.15	13.73	
G20	7.78		5.60	2.62	2.66	4.99	4.30	13.00	
G21	7.70		5.94	2.50	2.69	4.94	4.34	13.15	
C22	7.34	5.28	6.14	2.17	2.17	4.42	4.04		

[a] Could be exchanged.

the H1′−H2′/H2″ region. In Figure 1 (upper) we observe the cross peaks corresponding to the H8/6−CH3 [(*n*)−(*n*+1)] interactions, peaks G−I, characteristic of a B-DNA conformation.

In the H8/H6−H2′/H2″ region of the NOESY spectrum, Figure 1 (upper), we observe one of the G18H8−G17H2′/H2″ interactions and also one of the A7H8−G6H2′/H2″ interactions. The G6H8−C5H2′ interaction is not seen in Figure 1 (upper), although it is observed when the spectrum is plotted at lower levels. It is much weaker than corresponding interactions between other residues in the sequence. These cross peaks show that the mismatched bases G6 and G17 are stacked inside the helix.

Analysis of the H1′−H2′/H2″ region of a NOESY spectrum recorded with a short mixing time, 60 ms, gave the relative assignment of the H2′ and H2″ protons. The complete assignments of the H3′ and H4′ protons were obtained from the analysis of the TOCSY spectrum (not shown). The chemical shifts are given in Table 1. In Table 1, we observe an unusual inversion of the relative chemicals shifts for the H2′/H2″ protons of the mismatched bases G6 and G17. This phenomenon has, however, been observed for 3′-terminal bases or at chemically modified sites but is uncommon for a normal internal bases (Scheek et al., 1984; van de Ven & Hilbers, 1988). This H2′/H2″ inversion has been observed for certain mismatches (Borden et al., 1992; Carbonnaux et al., 1991; Kintanar et al., 1987; van de Ven & Hilbers, 1988). All of these oligonucleotides were found to have a local perturbation from B-DNA. Some authors suggested that this could be due to poor stacking of the mismatched bases with the adjacent bases (Borden et al., 1992; Greene et al., 1994). This result indicates that the G6 and the G17 bases are shifted out of the ring current field of their 3′ neighbors A7 and G18. Analysis of the relative intensities of the intraresidue H8/H6−H2′ cross peaks compared to those of the H8/H6−H3′ show that all sugar residues adopt a predominantly C2′-*endo* conformation.

To investigate the conformation of the glycosidic bond, we have examined the H8/H6−H1′ interactions in the 60 ms NOESY spectrum (not shown). All of the bases, except that of G6, are clearly in an *anti* conformation. For the G6 base, the cross peak G6H8−G6H1′ overlaps with the C2H5−C2H6 and C3H6−C3H1′ interactions [see Figure 1 (lower)]. The distance between H5 and H6 of a cytosine is 2.5 Å, and the usual distance between H8 and H1′, in the same nucleotide, is 2.5 Å for a glycosidic angle in a *syn* conformation and 3.7 Å if it is *anti* (Patel et al., 1982a). Since the C3 is in an *anti* conformation, the volume due to the C3H6−C3H1′ cross peak should be small compared to that of the C2H5−C2H6 interaction. The total volume of the cross peak is very close to the average volume for the other CH5−CH6 interactions. Thus, the volume due to the G6H8−H1′ interaction must be weak. This indicates that the G6 base is in an *anti* range conformation. Further, in this NOESY spectrum, we clearly observe the G6H8−G6H2′ cross peak and also that of G17H8−G17H2′. This interaction for a *syn* nucleotide would be expected to be very weak or absent. Although the intensities of these cross peaks are lower than for other residues, this is due to the large cross peak widths, ca. twice that for the Watson−Crick pairs. This reflects an enhanced mobility of the mismatched pair.

The magnitude of the NOEs are indicative of a normal B-form DNA. A series of NOESY spectra with short mixing times was recorded to determine interproton distances for model building.

Exchangeable Protons. The downfield region of the 1D spectrum recorded at pH 6, 11°C, is shown Figure 2. At least eight resonances are present in the region 12−14 ppm corresponding to imino protons involved in a Watson−Crick hydrogen bond; we also observe another resonance at 10.42 ppm which integrates for two protons. From the chemical shifts, we can assign the resonances between 13.5 and 14 ppm to the three A·T base pairs and the ones between 12.5 and 13.2 ppm to five of the G·C base pairs. We would expect two more G·C imino protons from the terminal base pairs, but they are broadened by exchange with the solvent at this temperature.

The assignment of the resonances was obtained from the analysis of a NOESY spectrum recorded in H_2O with a mixing time of 250 ms at 16 °C. Two regions are represented in Figure 3: imino/imino interactions [Figure 3 (lower)] and imino/amino interactions [Figure 3 (upper)]. In this spectrum, the resonance at 10.42 ppm is strongly attenuated by exchange with the solvent. The protons were assigned from the following considerations: the A·T imino protons present very strong interactions with the AH2 resonances, the G·C imino protons give rise to interactions with the previously assigned CH5 protons, the pairs of amino resonances were assigned by examination of the amino/amino region of the spectrum (not shown), and the imino resonances show NOEs with their neighboring imino protons. The resonance at 13.73 ppm shows interactions with two imino protons of guanines at 13.00 and 12.64 ppm. It also shows a strong cross peak with an AH2 at 7.78 ppm and a cross peak with a CH5 at 5.63 ppm that has been assigned to the H5 of C3. The AH2 resonance must correspond to the H2 of A4, so the signal at 13.73 ppm must be that of the resonance of the imino proton T19 and the two imino−imino interactions are with the neighboring G20 and G18 imino protons. The signal at 13.00 ppm also shows an NOE with

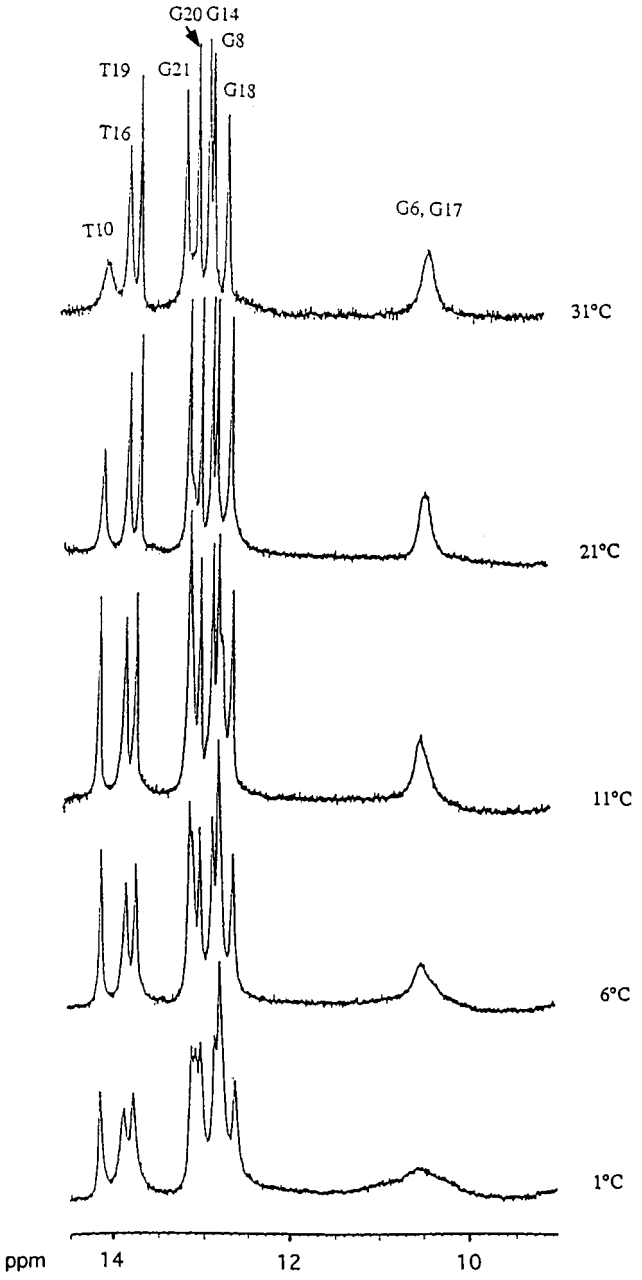

FIGURE 2: Temperature dependence of the imino protons of G·G duplex at 1−31 °C. The assignment is given on the spectrum at 31 °C .

the CH5 of C3 at 5.63 ppm, and it was thus assigned to the imino proton of G20. The other NOE, at 12.64 ppm, must correspond to the imino proton of G18. This is confirmed by its interaction with the C5H5 at 5.30 ppm. The imino proton of the T10 has been unambiguously assigned from its strong NOE at 7.80 ppm which has been assigned to the A13H2. The T10 imino proton shows an NOE at 12.87 ppm. At 16 °C, the imino protons of the terminal bases are in rapid exchange with the solvent and no NOE would be expected involving these bases. Taking this into account, we assigned the resonance at 12.87 ppm to the imino proton of the G14. The imino resonance at 13.15 ppm shows an NOE with the C2H5 and can be assigned to the G21 imino proton. The imino proton of T16 must correspond to the remaining resonance at 13.84 ppm. The strong NOE with the A7H2 confirms this assignment. We observe only one NOE from the T16 imino proton with a neighboring imino

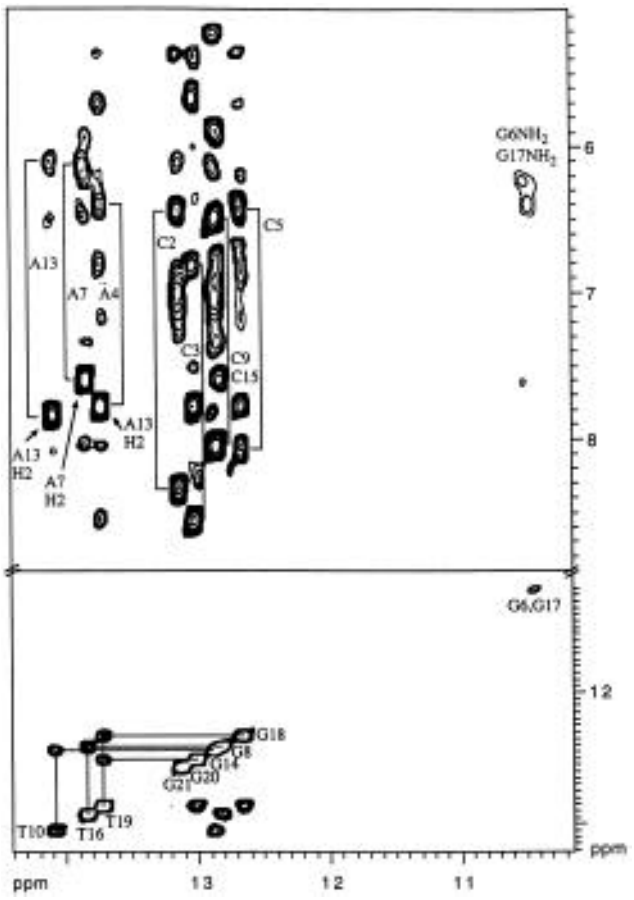

FIGURE 3: Expanded contour plot of the NOESY spectrum (250 ms mixing time) of the G·G duplex in H_2O, 11 °C and pH 6. The lower part shows the interactions between the imino protons. The upper part shows the interactions between the imino protons and the amino/H2/H5 protons. The protons of the amino group of cytidine or adenosine are connected by a continuous line.

proton, at 12.82 ppm, that must correspond to G8N1. There is still one signal not assigned at 10.42 ppm which must arise from the G6 and G17 imino protons. To probe further the environment of these protons, we have recorded 1D difference spectra. Presaturation of the signal at 10.42 ppm (not shown) gives rise to interactions with the T16 imino, G18 imino, and the A7H2 resonances, confirming the assignment we have proposed. In the NOESY spectrum, we observed interactions between the peak at 10.42 ppm with resonances at 6.22 and 6.40 ppm. At this temperature, the resonance at 6.40 could be assigned to an interaction with the non-hydrogen-bonded amino proton of the adjacent C5 although the absence of NOEs with the corresponding hydrogen-bonded proton would be surprizing. In a NOESY spectrum recorded at lower temperature (not shown), we observe that the two resonances at 6.22 and 6.40 move downfield relative to the adjacent amino proton resonances. These resonances must arise from the two G amino groups which are in rapid rotation on a proton NMR time scale. This indicates that they are not involved in stable interbase hydrogen bonding because in this case the amino protons would give rise to a broad and generally not observable resonance (Cognet et al., 1991; Sowers et al., 1987) but are probably hydrogen bonded to solvent. The complete assignments of nonexchangeable and exchangeable protons are given in Table 1.

Spectra as a Function of Temperature. We have recorded NMR spectra between 1 and 36 °C. In Figure 2 is shown the imino region of 1D spectra as a function of temperature. Decreasing the temperature from 21 °C, we observe a specific line broadening of the imino protons of the mismatch. At 1 °C, the resonance at 10.42 ppm had almost disappeared. Although all of the imino proton resonances become broader at 1 °C due to aggregation, the line width of the G6 and G17 imino protons can only be explained by a conformational exchange. At 1 °C this is in the intermediate exchange time scale range. Above 21 °C, the 10.42 ppm resonance begins to broaden due to exchange with solvent. We note that the imino resonance of one of the base pairs, T10·A13, has a similar line width at this temperature. We have examined the cross peaks of interactions in and around the mismatched base pairs by recording NOESY spectra in D_2O at 37, 19, and 6 °C. At 19 °C we observe that the cross peaks of the following interactions, G6H8−G6H2′H2″, G17H8−H1′, G17H8−T16H1′, and G17H8−H2′H2″, are considerably broadened. At lower temperature, 6 °C, the latter two are no longer detectable. The others broaden further, and significant broadening is also observed for A7H8−G6H1′, G18H8−G18H1′, and G18H8−G17H1′. All the other interactions observed in the duplex are still observed showing only a common small broadening due to aggregation. In agreement with the imino proton spectra, it is clear that at low temperature we observe a slowing down of an conformational equilibrium involving the central part of the duplex.

^{31}P NMR Spectra. ^{31}P spectra have been recorded at different temperatures (not shown) to investigate the backbone chain (Gorenstein, 1984; Privé et al., 1987). At all temperatures, the chemical shift dispersion is small, ca. 0.9 ppm, which is typical for a B_I conformation. No ^{31}P resonance was found to be shifted significantly downfield as would be expected for a B_{II} conformation (Chou et al., 1992).

G·G Pairing. Several G·G pairing schemes have been proposed in the literature from NMR and X-ray studies: some involved two hydrogen bonds (Castati et al., 1994; Cognet et al., 1991; Peterson et al., 1994; Skelly et al., 1993), but other authors proposed that the mispaired bases are only weakly hydrogen bonded (Borden et al., 1992; Chou et al., 1994).

The chemical shift of the GN1H resonance at 10.45 ppm is unusual. It is not characteristic of an imino proton bonded to nitrogen. This indicates that the GN1H is either hydrogen bonded to an oxygen atom, carbonyl, or solvent or non-hydrogen bonded (Carbonnaux et al., 1990; Patel et al., 1982b; Tibanyenda et al., 1984; Woodson & Crothers, 1988).

There are two possibilities of base pairing which could account for our NMR data: the imino protons could be either hydrogen bonded to water or in an NH-carbonyl bond. In the first case, we would expect to observe rapid disappearance of the imino mismatch resonances by exchange with the solvent with increasing the temperature. The imino proton must be protected to some extent from exchange as they are clearly observed at 31 °C at which temperature the imino proton resonances of the terminal base pairs have disappeared and that of the penultimate A·T base pair is very broad. Although we are not able to measure the individual line widths of the two G imino protons due to overlap, it is probable that they are similar. This suggests that they are

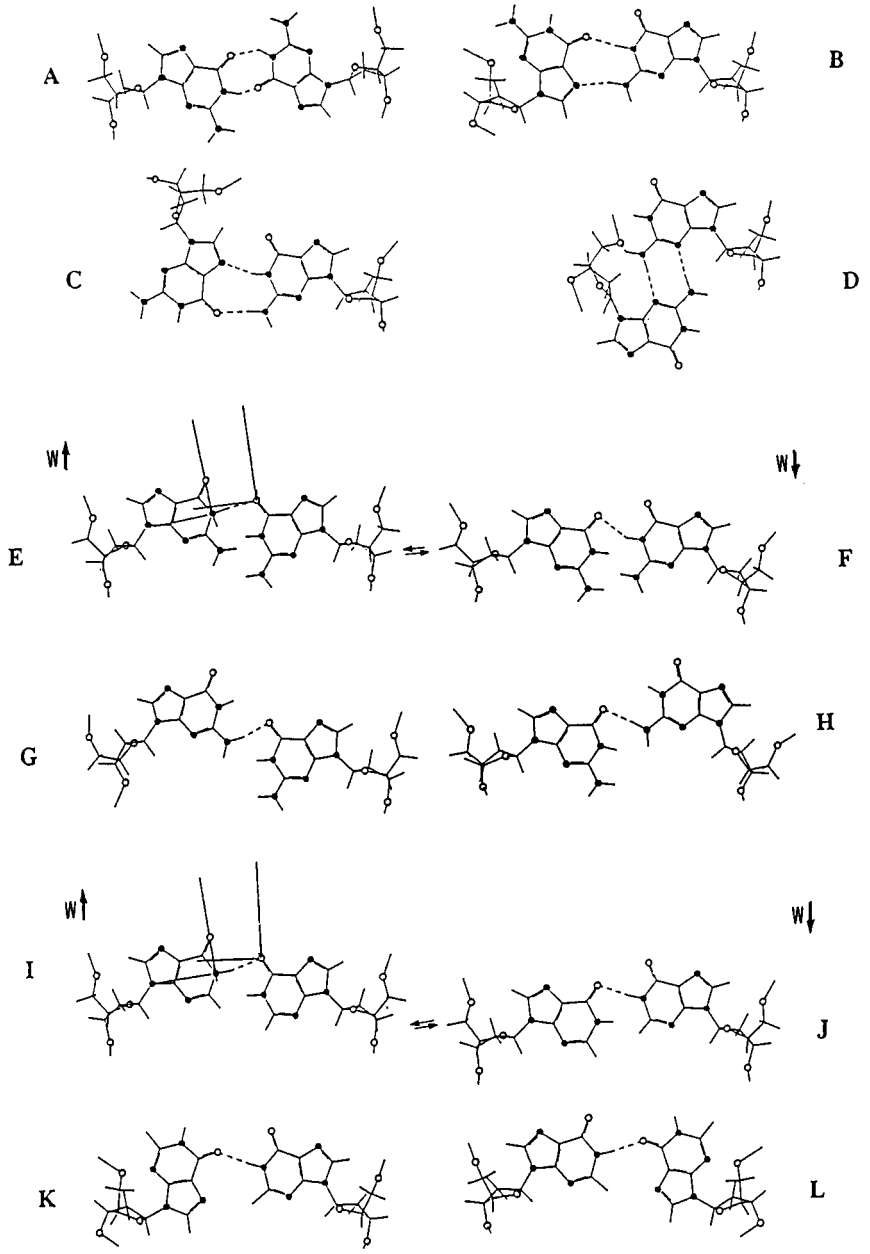

FIGURE 4: (A−H) G·G pairs. The four possible base-pairing schemes with two hydrogen bonds (A−D) where the bases are (A) *anti/syn*, (B) *syn/anti*, (C) *anti/anti*, and (D) *syn/anti*. The two base-pairing schemes E−F and G−H are shown with their two possible conformations, W↑ and W↓, where both bases are in an *anti* conformation and possess a single hydrogen bond. Figures I−L correspond to the proposed pairings and conformations for the I·I mismatch with a single hydrogen bond. The pairing scheme I−J is isomorphic to the E−F scheme for G·G. In pairing scheme K−L, the conformation of the inosines is (K) *syn:anti* and (L) *anti/syn*.

in a time averaged equivalent environment. We can exclude structures where one imino proton is hydrogen bonded and inside the helix while the other exposed in either the major or the minor groove but not necessarily with a rapid equilibrium between two such forms. At 31 °C, the imino resonance still integrates for two protons. On lowering the temperature from 21 °C, Figure 2, we observe specific broadening of the 10.42 ppm resonance which cannot be attributed to aggregation. This must arise from the slowing down of the conformational exchange of the G·G pair which is in the intermediate exchange range at 1 °C.

In order to determine the conformation of the G·G base pair, we have constructed different hydrogen bonding patterns for two guanines as shown in Figure 4A−H. There are only four possible base pairings with two hydrogen bonds (Figure 4A−D). Due to homobase symmetry, each scheme corresponds to two possible conformations. In the schemes A, B, and D, one of the guanines must be in the *syn* conformation, while in scheme C, both guanines are in the *anti* conformation. Because both bases were found in an *anti* conformation, patterns A, B, and D can be excluded. For model C, the G·G pairing features two kinds of hydrogen bonding. One imino proton is bonded to a nitrogen while the other is hydrogen bonded to a carbonyl group. We would expect to find different chemical shifts for the mismatched imino protons, one in the classical range of 12.5−13.5 ppm and the other in the range of 10−12 ppm, corresponding to the two types of hydrogen bond. However, we observe that the two imino protons of the mismatched base pair are both involved in a NH···O hydrogen bond, and thus model C does not fit the NMR data.

The two schemes E and F present bases in an *anti* conformation and possess a single hydrogen bond. The pairing involves the imino proton of one guanine and the carbonyl group of the other one. Both pairings are consistent with NMR data. The second pairing is deduced from the first by "shearing" one of the guanines relative to the other. An exchange between these two conformations would explain the change in line width of the imino protons in the temperature range 1−21 °C. Further, each imino proton is protected from exchange only part of the time in agreement with the enhanced solvent exchange observed for these protons. Previous studies on the same sequence containing an A·A mispair (Gervais et al., 1995) showed an exchange between two conformations in which the pairing involves a single hydrogen bond. Models G and H are derived from models E and F, respectively, by further increasing the shearing. In these models neither imino proton is involved in hydrogen bonding and thus we would not expect them to be major species, but their transient appearance cannot be excluded.

I·I Duplex

Nonexchangeable Protons. As for the G·G duplex we have recorded 1D proton spectra, at pH 6, as a function of temperature. We observed a global T_m of 57 °C, the same value as for the G·G duplex.

The nonexchangeable protons were assigned from NOESY spectra (recorded at 400 and 60 ms) and TOCSY spectrum (recorded at 60 ms). NOESY spectra of the I·I duplex show only few differences relative to the spectra of the G·G duplex. The differences are limited to the environment of the mismatched base pairs. The H8/H6−H1′/H5 region of the NOESY spectra recorded at 400 ms and at 42 °C is shown in Figure 5 (lower). For the first strand, the chain of connectivities can be followed without interruption, except for the step I6−C5 as the C5H1′ and I6H1′ have the same chemical shift. However, at 7°C, these H1′ resonances are resolved (not shown) and the chain of connectivities can be followed. For the second strand, the chain of connectivities can also be followed clearly except for the I17−T16 step where the cross peak I17H8−T16H1′ is missing or very weak. It is observed in the NOESY spectrum at 7 °C although much weaker than the same interaction between Watson−Crick base pairs. From the H8/H6−H8/H6 region we observe the usual B-DNA interactions, in particular two cross peaks I17H8−T16H6 and I6H8−C5H6, peaks C and E, respectively, which confirm the stacking of the I17 and I6 bases inside the helix. We also observe a cross peak between I6H2 and I17H2 (not shown).

Examination of the H8/6−H2′/H2″ region of the same NOESY spectrum shows normal B-DNA cross peaks for all H8/6 protons, Figure 5 (upper). The spectrum shows the normal I6H8−I6H2′/H2″, I6H8−C5H2′/H2″, I17H8−I17H2′/H2″, and I17H8−T16H2′/H2″ cross peaks. However, we observe unusual interactions from the H2 protons of both inosines: the I6H2 and the I17H2 protons exhibit strong NOEs with H2′/H2″ protons of the 5′-flanking C5 and T16 residues. These NOEs are not possible for a regular B-DNA and can only arise when the inosines I6 and I17 are in a *syn* conformation for at least part of the time (Oda et al., 1991).

From the H8/H6−H8/H6 region of the 400 ms NOESY we observe four cross peaks involving inosine protons, peaks

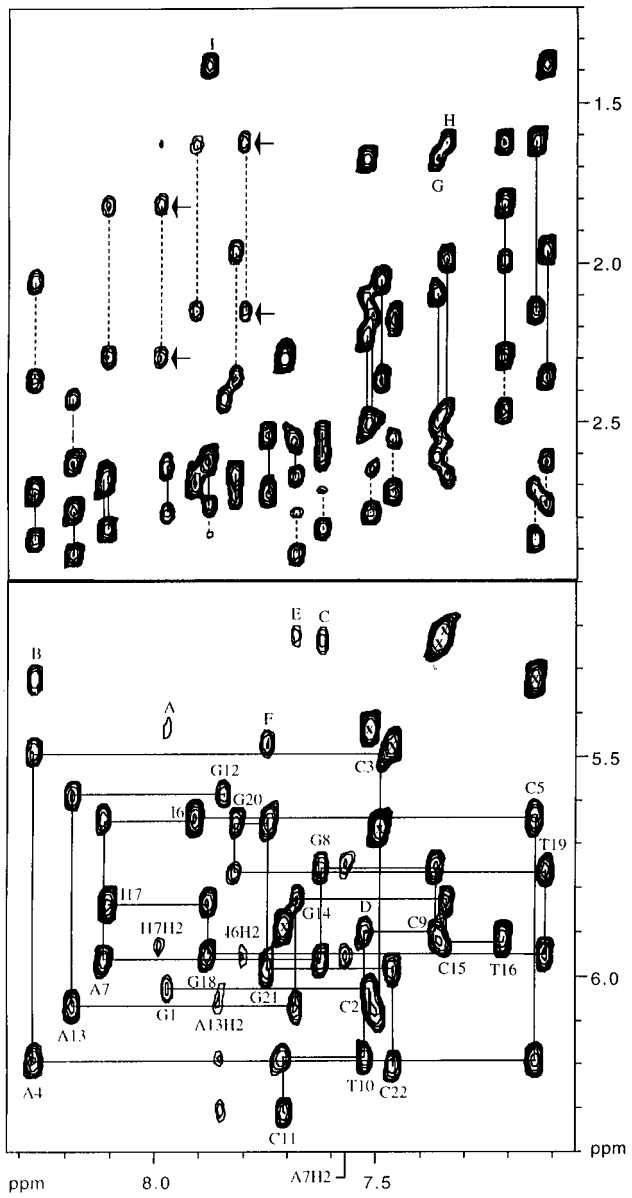

FIGURE 5: Expanded contour plot of the H6/H8−H1′/H5 (lower) and H6/H8−H2′/H2″/CH_3 (upper) regions of the NOESY spectrum (400 ms mixing time) of the I·I duplex in D_2O, 42 °C and pH 6. Cross peaks marked with an X correspond to CH6−CH5 interactions. Cross peaks A to F correspond to H6/H8−CH5 [(*n*)−(*n*+1)] interactions, and cross peaks G to I to H6/H8−CH_3 [(*n*)−(*n*+1)] interactions. Unusual cross peaks for a B DNA are indicated by double arrows.

C and E, which are consistent with an *anti* conformation for the I6 and I17 bases, but also C5H6−I6H2 and T16H6−I17H2 interactions, peaks H and D, which are typical of a *syn* conformation. Given that the A7·T16 and C5·G18 base pairs have a classical B-DNA conformation, the NOEs we observe demonstrate that more than one conformation for the mismatched bases must exist in solution.

Observation of the H8/H6−H2′/H2″ interactions in the NOESY spectrum recorded at 60 ms mixing time (not shown) confirms our results. The spectrum shows I6H8−C5H2′/H2″, I17H8−T16H2′/H2″, I6H2−C5H2′/H2″ and I17H2−T16H2′/H2″ cross peaks that could not exist for a single conformation of the mismatched bases. The latter two observed in this NOESY spectrum cannot be due to spin diffusion within a B-form conformation. From analysis of this spectrum we find that all the non-mismatched bases are

in an *anti* conformation and the sugars are predominantly C2′-*endo*. To investigate the conformation of the I6 base, we have examined the volume of the I6H8−I6H1′ cross peak and also that relative to the interaction with I6H2′/H2″. The I6H8−I6H1′ cross peak is coincident with the I6H8−C5H1′ cross peak at this temperature. Integration of this cross peak shows that it is ca. 50% greater than that of the sum of the inter- and intraresidue H8/H6−H1′ cross peaks for the nonterminal Watson−Crick bases. A predominantly *syn* conformation can thus be excluded. The I6H8 shows a much larger NOE to its own H2′ than that to its own H1′ which indicates a predominatly *anti* conformation; nevertheless the ratio determined is not typical for an pure *anti* conformation. It is not possible to make the same measurements for the I17 base since the A7H8−A7H2′/H2″ cross peaks overlap with the I17H8−I17H2′/H2″ cross peaks. We have measured the NOE build-up rate and the interproton distance for the I17H8−I17H1′. We calculate a distance of 3.4 Å for this interaction. For a typical *syn* interaction we would expect a distance of 2.5 Å and for a typical *anti* conformation a distance of 3.7 Å. From model building studies and molecular dynamics calculations starting form the a *syn*/*anti* or *anti*/*syn* conformation, we observe that the distance between the H2 of the inosine which is *syn* and the H2′ and H2″ of the 5′ residue is on average ca. 2.4 Å for the pairing shown in Figure 4K,L. From integration of these cross peaks relative to the intra- and interresidue H8/H6−H2′/H2″ cross peaks for the Watson−Crick base pairs in the 60 ms NOESY spectrum, we can estimate the population of the *syn*/*anti* and *anti*/*syn* conformations to be 5−10% at 42 °C. An additional *syn*/*anti* base pairing scheme could exist, that corresponding to the conformation of the G·G in Figure 4A. However, in this conformation, the IH2 proton is ca. 1 Å further away from the H2′/H2″ protons of the 5′ residue. A much higher population of *syn* conformations would be required to account for the observed interactions with these protons. This would also be reflected in the IH8−IH1′ intraresidue cross peaks which are not consistent with a high *syn* population. From these results, it seems clear that the mismatched bases exist predominantly in an *anti*/*anti* conformation, however, a minor contribution from a *syn*/*anti* and *anti*/*syn* conformation appears necessary to explain all the data. We also note that the atypical interactions I6H2−C5H2′/H2″ and I17H2−T16H2′/H2″ are more pronounced at low temperature.

The complete assignment of the H3′ and H4′ protons was obtained from analysis of the TOCSY spectrum (not shown) and the chemical shifts are given in Table 2. A series of NOESY spectra with short mixing times were recorded to determine interproton distances for model building.

Exchangeable Protons. We have recorded 1D spectra, at pH 6, of the duplex as a function of temperature (not shown). The spectrum at 17 °C is shown Figure 6 (upper). In the low-field region, 12.5−14.2 ppm, the eight resonances, which integrate for ten protons, correspond to the imino protons of the Watson−Crick base pairs. This is confirmed, and their assignment is given by a NOESY spectrum recorded at this temperature with a mixing time of 250 ms (not shown). The broad resonance at 12.27 ppm corresponds to the two imino protons of the I·I base pair which overlap. The 1D spectra as a function of temperature show that the apparent line width of this resonance does not significantly change between 1 and 23 °C. Above this temperature, it broadens but it is still visible in the spectrum recorded at 48 °C, which is very similar to that observed above for the G·G base pair. On the other hand, the low-temperature spectra indicate a more rapid conformational exchange than observed for the G·G mismatch. Presaturation of the 12.27 ppm resonance, Figure 6 (lower), results in NOEs to the H2 protons of the I·I base pair and a weak NOE to A7H2. NOEs are also observed to the neighboring imino protons, although that to the T16 is very weak. This resonance is broad, indicating that the A7·-T16 base pair is destabilized by the adjacent mispair as has been observed for this base pair with other central mispairs which showed only minor conformational exchange. It cannot, however, be excluded that the major conformational exchange here, *syn*−*anti*, influences the line width of the T16 imino proton. The NOEs observed show that the inosine imino protons are in the interior of the helix. The chemical shift is characteristic of inosine imino protons hydrogen bonded either to a carbonyl group or to solvent (Carbonnaux et al., 1991; Oda et al., 1991) whereas when bound to nitrogen a chemical shift of ca. 15 ppm has been reported (Uesugi et al., 1987). We do not observe separate resonances of the imino protons corresponding to *syn*/*anti* or *anti*/*syn* conformations. Either they are not visible due to rapid exchange with the solvent as they are exposed in the major groove of the helix, Figure 4K,L, or there is a rapid exchange with the imino proton resonances of the *anti*/*anti* conformation. The complete assignment of nonexchangeable and exchangeable protons are given in Table 2.

Table 2: Chemical Shifts of Nonexchangeable Protons at 42 °C and of Exchangeable Protons at 17 °C for the I·I Duplex, pH 6

	H8/6	H2/H5/CH_3	H1′	H2′	H2″	H3′	H4′	NH	NH2
G1	7.96		6.01	2.64	2.77	4.85			
C2	7.50	5.42	6.07	2.15	2.49	4.86	4.19		6.49/8.39
C3	7.48	5.64	5.47	2.05	2.36	4.84	4.10		6.82/8.57
A4	8.26	7.74	6.18	2.72	2.86	5.03	4.40		6.21/7.71
C5	7.12	5.31	5.63	1.62	2.13	4.72	4.22		6.59/8.08
I6	7.90	7.79	5.62	2.66	2.70	4.95	4.23	12.27	
A7	8.10	7.56	5.95	2.68	2.83	5.02	4.38		6.13/7.40
G8	7.61		5.73	2.53	2.60	4.93	4.37	12.80	
C9	7.35	5.22	5.89	2.08	2.48	4.71	4.21		6.52/8.07
T10	7.51	1.61	6.17	2.22	2.53	4.89	4.19	14.15	
C11	7.70	5.88	6.29	2.29	2.29	4.57	4.05		
G12	7.83		5.57	2.42	2.62	4.82	4.16		
A13	8.17	7.84	6.05	2.77	2.91	5.06	4.42		615/790
G14	7.67		5.82	2.55	2.66	4.98	4.42	12.90	
C15	7.33	5.20	5.90	1.98	2.46	4.71	4.26		6.50/8.03
T16	7.20	1.61	5.91	1.81	2.29	4.82	4.14	13.62	
I17	8.09	7.98	5.82	2.63	2.84	4.96	4.38	12.27	
G18	7.87		5.94	2.62	2.74	4.98	4.39	12.61	
T19	7.10	1.37	5.75	1.95	2.35	4.85	4.14	13.77	
G20	7.80		5.64	2.66	2.73	4.97	4.32	13.03	
G21	7.73		5.97	2.54	2.70	4.95	4.35	13.13	
C22	7.45	5.45	6.19	2.17	2.17	4.50	4.05		

^{31}P NMR Spectra. In order to investigate the phosphodiester backbone conformation, we have recorded ^{31}P spectra (not shown). No ^{31}P resonance was found to be shifted significantly downfield as would be expected for a B_{II} conformation.

I·I Pairing. The majority of the data obtained on the I·I system strongly resembles those data for the G·G system. In both cases the imino protons of the mismatched pair are still visible in the range 40−50 °C. The inosine imino protons must be protected to some extent from exchange with the solvent as they exchange more slowly than those of the terminal base pairs. For the predominant species, our data

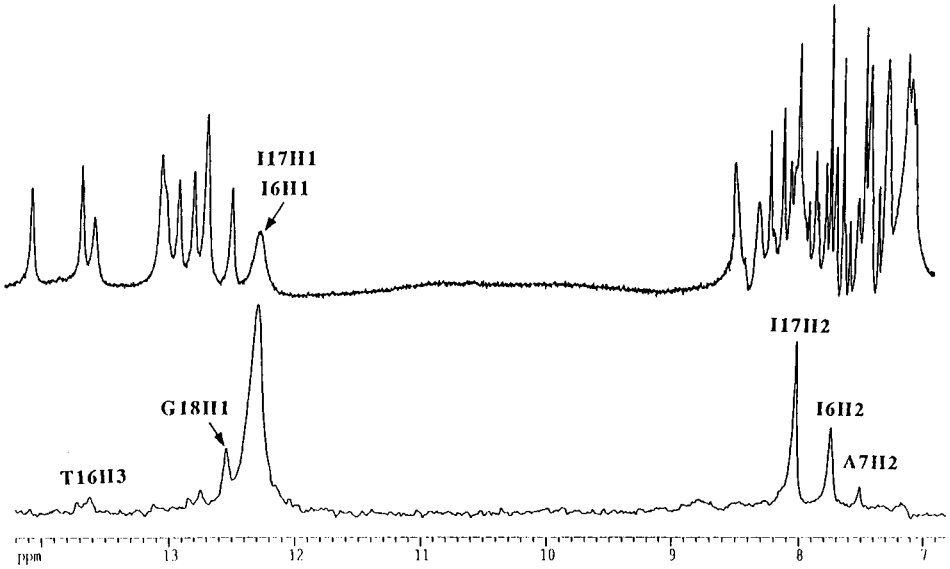

FIGURE 6: 1D spectrum (upper) of the I·I duplex in H_2O, 17 °C and pH 6 (lower). Difference spectrum after presaturation for 0.3 s of the resonances I6NH1 and I17NH1.

fit for an equilibrium between two structures involving one hydrogen bond in an *anti*/*anti* conformation, Figure 4I,J. However, the unusual interactions observed with the IH2 protons, Figure 5, do not fit these models and must arise from a minor *anti*/*syn* and *syn*/*anti* equilibrium, Figure 4K,L. As no separate resonances are observed, exchange between *syn* and *anti* conformations must be rapid on a proton NMR time scale.

Molecular Modeling Studies. In order to compare the molecular modeling studies described here with those that were performed on the same oligonucleotide but with different central mismatch base pairs (Boulard et al., 1995; Cognet et al., 1995; Gervais et al., 1995) all the molecular mechanics and dynamics have been computed as previously described. The starting conformations, W↑ and W↓, of the G·G mismatch were first energy minimized. From each of the structures obtained, two types of MD runs were performed: MDI, with distance restraints derived from NOE build-up measurements and MDII, without distance restraints. The distance restraint force constant used was chosen to be as low as possible but yet sufficient to keep average interproton distances within the distance range determined from the NMR measurements (Cognet et al., 1995). The exchange between the two forms W↑ and W↓ is observed, independently of the starting structure, in the MDI or MDII runs. A trajectory plot of the shear parameter is shown in Figure 7A. The shear oscillates between −0.8 and −3.5 Å, and Figure 7B presents the corresponding histogram of the shear distribution. These observed fluctuations for the shear are large compared to those of other translation parameters such as the stretch and the stagger (not shown). These two plots show the oscillations back and forth between the two conformations W↑ and W↓ of Figure 4E,F. Guanine features the largest number of hydrogen bond donors or acceptors of all natural bases: one other base pairing is observed as a minor form in MDI and MDII. This is shown in Figure 4G. It has a shear parameter of ca. 1.5 Å and appears transiently as shown in Figure 7A,B. The reversed structure Figure 4H is not observed.

The time variations of the internuclear distances G6NH1−G17O6 and G6O6−G17NH1 during the MD runs are represented in Figure 7C,D. For each potential donor−acceptor pair the distance oscillates between a close contact, ca. 2.4 Å corresponding to hydrogen bonding, and a large separation, ca. 4 Å. These two conformations account for most of the structures observed in the MD runs as the transition between the two is fast. These variations are very well correlated with the variation of the shear parameter, cf. Figure 7A. The time variations of G6HN2−G1706 in Figure 7E illustrate the formation of the minor form shown in Figure 4G. The time variations of the distance G6HN2−G17HN2 are presented in Figure 7F. That of G6H1−G17H1 are nearly identical and are not shown. Apart from small deviations due to the appearance of the minor form Figure 4G, these distances correspond to close contacts and their remarkable steadyness shows that, within fluctuation limits, these pairs of atoms behave as virtual hinges for the lateral displacement between W↑ and W↓. This accounts for the minimal distortion of the backbone that occurs during the transition as characterized below.

We have computed, on the basis of the shear trajectory plots, the time segments of all the MDI runs where the mismatch is in the W↑ or in the W↓ conformation in order to determine the average torsion angles and wedge parameters for each conformation. In the W↑ and the W↓ conformations, the guanine bases of the G·G mismatch are close to the geometrical position that they would occupy in a G·C or C·G Watson−Crick base pair. As a result the deformation of the sugar phosphate remains minimal, and the torsion angle values are close to those of standard B-DNA. The largest distorsions are observed for the neighboring base pairs of the mismatch, C5·G18 and A7·T16, as shown in Table 3. The parameter values that lie outside the threshold values centered about reference values (Gervais et al., 1995) are marked with an asterisk. The torsion angles δ, ϵ, and ζ of C5 and T16 and β of G6 and G17 showed the largest deviations as well as the largest fluctuations. For the bases adjacent to the mismatch, the most significant deformations are thus found for those 5′ of the G residues in a pseudodyadic manner relative to the helical axis passing through the G·G base pair. The wedge parameters of the G·G mismatch are remarkable in two respects. The shear of W↑ is not zero

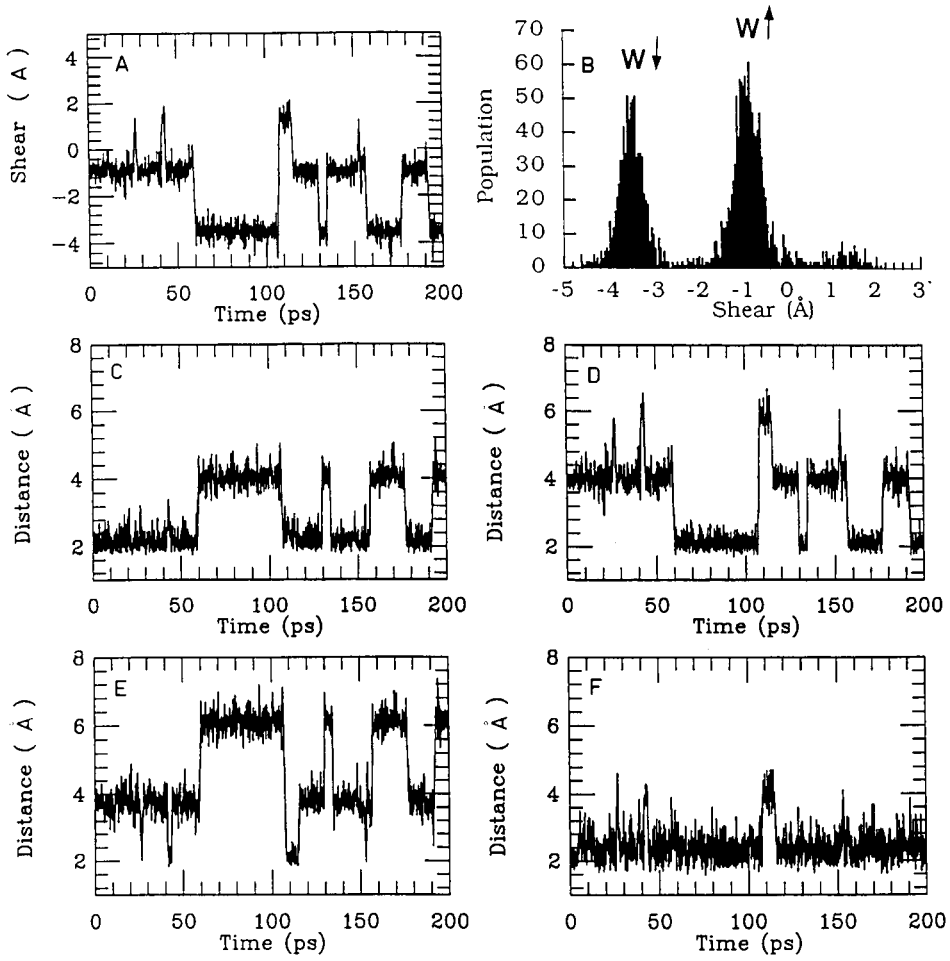

FIGURE 7: G·G duplex, MDI runs. (A) Trajectory plot of the shear, (B) corresponding shear population distribution, the vertical scale is in arbitrary units, (C) distance G6NH1−G17O6, (D) distance G6O6−G17NH1, (E) distance G6NH2−G17O6, and (F) distance G6NH2−G17NH2.

Table 3: Torsion Angles and Standard Deviations (in deg) and Wedge Parameters for the 200 ps MDI Runs[a]

	base	pucker		amp		α		β		γ		δ		ϵ		ζ		χ	
W↑	C5	123*	19*	41	5	−74	10	176	8	56	8	120	14	−150*	31*	−131*	34*	−124	12
W↑	G6	163*	17	39	5	−73	12	161*	17*	65	10	143	8	−175	8	−106	12	−92	12
W↑	T16	121*	19*	39	6	−73	10	−179	9	57	8	118*	13	−169	11	−112	23*	−118	10
W↑	G17	151	14	42	5	−72	10	−167	10	64	9	138	9	−176	8	−101	10	−107	9
W↓	C5	125*	17	41	5	−74	10	176	8	56	8	121	13	−155*	25*	−132*	27*	−123	12
W↓	G6	161*	13	42	4	−77	13	166	13	62	10	144	8	172	8	−99	9	−94*	10
W↓	T16	126*	17	40	6	−73	10	−177	9	56	8	121	12	−170	14	−111	22*	−116	11
W↓	G17	147	15	41	5	−70	10	164*	11	66	8	135	10	178	8	−105	10	−109	9

	pair	buckle *X*		prop *Y*		open *Z*		shear *X*		stretch *Y*		stagger *Z*	
G−GW↑	C5−G18	−6	8	0*	7	−2	5	0.0	0.6	0.2	0.2	0.9*	0.4
G−GW↑	G6−G17	5	11	−17	7	2	7	−0.8*	0.6	2.8	0.2	−0.7*	0.4
G−GW↓	C5−G18	−3	8	−1*	7	−1	5	−0.1	0.5	0.2	0.2	1.0*	0.4
G−GW↓	G6−G17	11*	9	−15*	7	−1	6	−3.5*	0.3	2.8*	0.2	−0.8*	0.4

[a] The table includes residues at and adjacent to the mismatch site d(C5 G6 A7)·d(T16 G17 G18). The values which lie outside the normal range are marked with an asterisk.

as it would be if the hydrogen bond G6H1−G17O6 was perfectly formed. The average shear values, −0.8 and −3.5 Å, reflect in part the steric hindrance introduced by the two imino protons and the two amino groups of the G·G mismatch. These values were found closer to normal in the W↑ and the W↓ (−0.4 and −4.7 Å respectively) conformations of the I·I mismatch, Figure 4I,J. Another consequence of this is reflected in the stagger parameter, or *z* translational parameter, with a large negative value at the mismatch, that is compensated by a large positive value at the neighboring base pair C5·G18. The propeller twist, opening, stretch, and stagger parameters are pseudodyadic, i.e., unchanged upon 180° rotation, and remain very similar in W↑ and W↓ conformations as shown in Table 3.

We observe that the conversion of W↑ to W↓ occurs systematically along the same pathway. Figure 8 presents stereoviews of W↑ and W↓ and of an intermediate conformation observed during MDI. Each view is taken at ap-

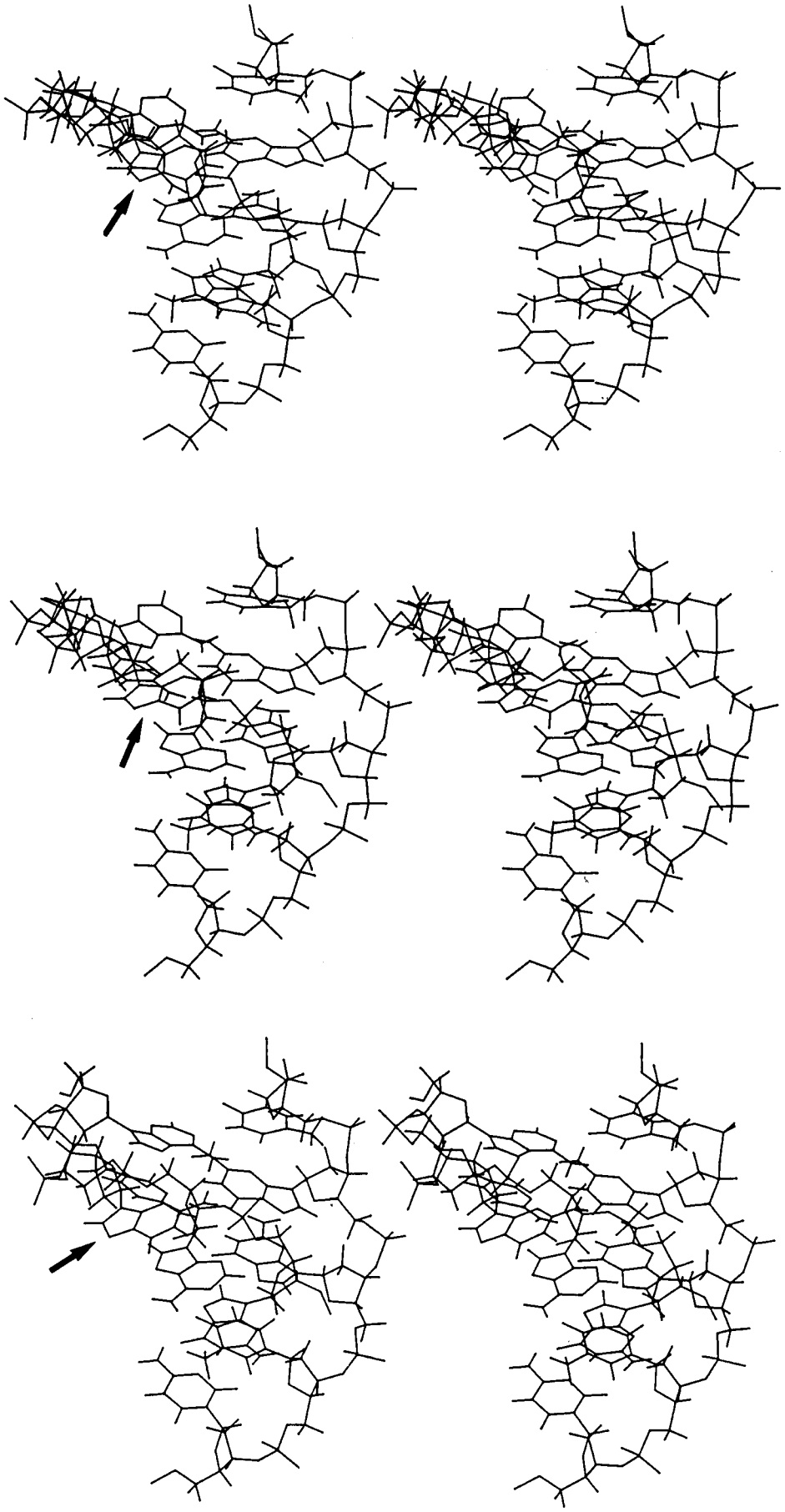

FIGURE 8: Stereoviews of the central part of the G·G duplex showing the conformation W↓ (top) and W↑ (bottom), and an intermediate conformation observed during the exchange (center) observed during MDI runs. Arrows point to the central G·G mispair.

proximately 1 ps interval. In Figure 8, the central G base of the second strand (front) moves in the 5′ direction below the opposite G base (back). This observation is partly accounted for by the right-handedness and negative propeller twist of the DNA. It may also be partly due to possible hydrogen bonding between the amino group of C5 and the carbonyl of G17. Although the exchange involves large lateral displacement, no major distortion of the sugar phosphate chain is observed. The time scale of the exchange as observed by NMR and MD runs is greatly different. The latter were performed *in vacuo* which may have a strong influence upon the rate of interconversion.

The molecular dynamics runs for the oligonucleotide with the mismatch I·I are very similar to those observed with mismatch G·G. The conformation W↑ appears somewhat less favored than W↓. As we have observed exchange between *anti* and *syn* conformations, distance restraints could not be incorporated and only MDII calculations were performed.

C1′−C1′ Distances and λ6 and λ17 Angles for G·G and I·I. For the two conformations, the pairing G·G results in an average C1′−C1′ distance of 12.8 Å, i.e., ca. 2.0 Å larger than for Watson−Crick base pairs. The corresponding value is 12.4 Å for I·I. Since W↑ and W↓ are pseudodyadic, we would expect that the angles λ6 and λ17 are interchanged in the two conformations. This is true for the average values but differences are observed in the amplitude of the fluctuations.

We are able to observe the exchange between the two conformations by molecular dynamics computations. The NMR results indicate the presence of an exchange process but do not allow us to identify the different conformations independently as only a time average spectrum is observed, exchange being in the fast regime. The lifetime of the different intermediate states is probably short such that the NMR data correspond to the time average of the limiting structures. The interproton distances for the nonexchangeable protons which can be measured for the central three base pairs do not change significantly between the structures W↑ and W↓. The NMR data do not give any information concerning the relative population of the two structures, and MD runs suggest that W↑ and W↓ are about equally favored as shown in Figure 7B, but this should be treated with caution as longer MD runs might change this.

Comparison of G·G Structures in Different Sequence Contexts. The nature of the neighboring bases will certainly play an important role through stacking energy interactions. Only one case of a G·G mismatch has been reported (Cognet et al., 1991) in which the sequence context is pur-G-pur on one strand. This residue was found in an *anti* conformation, while that between the pyrimidines was in a *syn* conformation. Symmetrical imino-carbonyl hydrogen bonding was observed. The choice of which strand conserves the G *anti* is probably governed by more favorable stacking interactions. This is the only report of this type of pairing.

All other studies, including the present one, have focused on the pyr-G-pur context although with different conclusions. For the same sequence containing CGA, an NMR study found *anti−anti* conformations for the G·G pair (Borden et al., 1992) whereas an X-ray study found *anti−syn* conformations (Skelly et al., 1993) but with different hydrogen bonding to that in the pur-G-pur environment. It is known that the difference in energy between the two forms is very small and that different environmental conditions may be sufficient to influence the structure which is adopted.

Our present study also focuses on the CGA context and while we also observe *anti−anti* conformations for the G·G pair we clearly observe a conformational equilibrium at low temperature which was not observed previously (Borden et al., 1992). Recently (Lane et al., 1995) reported a study on the G·G pair also in the CGA context and proposed a rapid *anti−syn* exchange between a variety of structures.

Thus an additional factor must also be involved in determining the final structure. In the crystal structure study it was proposed that base pairing conformations may also be determined by the flanking sequence by influencing the extent of base stacking of the immediately adjacent bases with those of the mispair. This is in agreement with observations that repair efficiency can be sequence dependent.

Conclusion. We find for both mismatches that the bases are stacked in the helix, with a single hydrogen bond between the bases, and induce only minor structural changes for the backbone conformation. Nevertheless, some differences are observed between the G·G and I·I mispairs.

The substitution of inosine for guanosine in oligonucleotide sequences has been used in order to determine whether the guanosine amino group is involved in hydrogen bonding within a mismatched base pair (Chou et al., 1994; Li et al., 1991). These amino protons are generally not observable in NMR spectra and their implication or not in hydrogen bonding is not easily established. In this study we observe that the global T_m of the two duplexes studied is identical within experimental error which suggests that the G amino group does not participate in hydrogen bonding. Further, the major species observed for the two duplexes is the same: an *anti/anti* conformation with a single hydrogen bond between the imino proton of one base and the carbonyl group of the other. Exchange between the two pseudodyadic conformations approaches the intermediate exchange rate at low temperature for the G·G mispair, as was observed for the A·A mismatch in the same sequence context (Gervais et al., 1995), but this is not observed for the I·I mismatch. An amino group, either in the major or the minor groove, appears to slow down the exchange possibly for steric reasons. The two pseudodyadic conformations are approximately equally populated for both mispairs. From the MD runs exchange between the two conformations does not require base pair opening and follows the same pathway independent of the starting structure.

We clearly observe a minor species for the I·I mismatch corresponding to a structure in which one of the inosines is in the *syn* conformation. Although the population of this species is low, it is clearly identified by the IH2−H2′/H2″ (n−1) interactions. Generally *syn* conformations are revealed by strong H8/H6−H1′ interactions, but these are not very sensitive to the presence of a minor species. Suprisingly, separate resonances are not observed for the *anti/anti* and *anti/syn* conformations which are in rapid exchange on a proton NMR time scale. Although rapid *syn−anti* exchange has been previously proposed for the G·G mismatch (Lane & Peck, 1995), this study shows unambiguously that this exchange is possible and that the conformations can be in rapid exchange. Thus, we cannot exclude a minor contribution from a *syn/anti* and or *anti/syn* conformation for the

G·G mismatch in this sequence although there is no direct evidence for this.

Clearly the mismatch repair system must be able to recognize a large variety of base pairings different from those of Watson−Crick pairs. For a given mispair three factors can affect the conformation of the mispair: (1) the pH, protonation. or deprotonation of a base, 2) the adjacent sequence, as we have observed in another sequence that a G·G mismatch adopts a *syn/anti* conformation with no evidence for an exchange with an *anti/syn* conformation (Cognet et al., 1991); and (3) the temperature. Interpretation of mispair repair pathways or recognition of these mispairs requires studies as a function of these three parameters for which the biochemical data is not yet available.

ACKNOWLEDGMENT

We are most grateful to INSERM-SC5 for allowing us to compute the inosine charge distributions and to Mr R. Attias for help with the CCDB. We thank Professor P. A. Kollman for providing us with AMBER.

SUPPORTING INFORMATION AVAILABLE

One figure showing an expanded contour plot of part of the H6/H8−H6/H8 region of the NOESY spectrum of the I·I duplex in D_2O at 42 °C, and one table giving C1′−C1′ distances for the base pairs G6−G17 and I6−I17 and values of angles formed between the C1′−C1′ vector and the glycosidic bonds (λ) (2 pages). Ordering information is given on any current masthead page.

REFERENCES

Arnott, S., Smith, P. J. C., & Chandrasekaran, R. (1976) in *Atomic coordinates and molecular conformations for DNA−DNA, RNA−DNA and DNA−RNA helices.* (Fasman G. D., Ed.) Vol. 2, CRC Press, Cleveland, OH.

Bodenhausen, G., Kogler, H., & Ernst, R. R. (1984) *J. Magn. Reson. 58*, 370−388.

Borden, K. L., Jenkins, T. C., Skelly, J. V., Brown, T., & Lane, A. N. (1992) *Biochemistry 31*, 5411−5422.

Boulard, Y., Cognet, J. A. H., Gabarro-Arpa, J., Le Bret, M., Sowers, L. C., & Fazakerley, G. V. (1992) *Nucleic Acids Res. 20*, 1933−1941.

Boulard, Y., Cognet, J. A. H., Gabarro-Arpa, J., Le Bret, M., Carbonnaux, C., & Fazarkerley, G. V. (1995) *J. Mol. Biol. 246*, 194−208.

Carbonnaux, C., Fazakerley, G. V., & Sowers, L. C. (1990) *Nucleic Acids Res. 18*, 4075−4081.

Carbonnaux, C., van der Marel, G. A., van Boom, J. H., Guschlbauer, W., & Fazakerley, G. V. (1991) *Biochemistry 30*, 5449−5458.

Castati, P., Gupta, G., Garcia, A. E., Radliff, R., Hong, L., Yau, P., Moyzis, R. K., & Bradbury, E. M. (1994) *Biochemistry 33*, 3819−3830.

Chou, S.-H., Cheng, J.-W., Fedoroff, O. Y., Chuprina, V. P., & Reid, B. R. (1992) *J. Am. Chem. Soc. 114*, 3115−3117.

Chou, S.-H., Zhu, L., & Reid, B. R. (1994) *J. Mol. Biol. 244*, 259−268.

Cognet, J. A. H., Gabarro-Arpa, J., Cuniasse, P., Fazakerley, G. V., & Le Bret, M. (1990) *J. Biomol. Struct. Dyn. 7*, 1095−1115.

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. 19*, 6771−6779.

Cognet, J. A. H., Boulard, Y., & Fazakerley, G. V. (1995) *J. Mol. Biol. 246*, 209−226.

Cuniasse, P., Sowers, L. C., Eritja, R., Kaplan, B., Goodman, M. F., Cognet, J. A. H., & Le Bret, M. (1987) *Nucleic Acids Res. 15*, 8003−8022.

Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2820−2821.

Dickerson, R. E. (1989) *Nucleic Acids Res. 17*, 1797−1803.

Fazakerley, G. V., Quignard, E., Woisard, A., Guschlbauer, W., van der Marel, G. A., van Boom, J. H., Jones, M., & Radman, M. (1986) *EMBO J. 5*, 3697−3703.

Gervais, V., Guy, A., Teoule, R., & Fazakerley, G. V. (1992) *Nucleic Acids Res. 20*, 6455−6460.

Gervais, V., Cognet, J. A. H., Le Bret, M., Sowers, L. C., & Fazakerley, G. V. (1995) *Eur. J. Biochem. 228*, 279−290.

Gorenstein, D. G. (1984) in *Phosphorus 31 NMR: Principles and Applications* (Gorenstein, D. G., Ed.) pp 37−53, Academic Press, New York.

Greene, K. L., Jones, R. L., Robinson, H., Wang, A. H.-J., & Wilson, W. D. (1994) *Biochemistry 33*, 1053−1062.

Kintanar, A., Klevit, R. E., & Reid, B. R. (1987) *Nucleic Acids Res. 15*, 5845−5862.

Kramer, B., Kramer, W., & Fritz, H. J. (1984) *Cell 38*, 879−887.

Lane, A. N., & Peck, B. (1995) *Eur. J. Biochem. 230*, 1073−1087.

Li, Y., Zon, G., & Wilson, W. D. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 26−30.

Marrug, J. E., Tromp, M., Ihurani, P., Hoyng, C. F., van der Marel, G. A., & van Boom, J. H. (1984) *Tetrahedron 40*, 73−78.

Oda, Y., Uesugi, S., Ikehara, M., Kawase, Y., & Ohtsuka, E. (1991) *Nucleic Acids Res. 19*, 5263−5267.

Patel, D. J., Koslowski, S. A., Nordheim, A., & Rich, A. (1982a) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1413−1417.

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

Peterson, R. D., Bartel, D. P., Szostak, J. W., Horvath, S. J., & Feigon, J. (1994) *Biochemistry 33*, 5357−5366.

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

Privé, G. G., Heinemann, U., Chandrasegaran, S., Kang, L.-S., Kopka, M. L., & Dickerson, R. E. (1987) *Science 238*, 498−504.

Quignard, E., Buu, B., & Fazakerley, G. V. (1986) *J. Magn. Reson. 67*, 342−345.

Roongta, V. A., Jones, C. R., & Gorenstein, D. G. (1990) *Biochemistry 29*, 5245−5258.

Scheek, R. M., Boelens, R., Russo, N., van Boom, J. H., & Kaptein, R. (1984) *Biochemistry 23*, 1371−1376.

Skelly, J. V., Edwards, K. J., Jenkins, T. C., & Neidle, S. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 804−808.

Sowers, L. C., Shaw, B. R., Veigl, M. L., & Sedwick, W. D. (1987) *Mutat. Res. 177*, 201−218.

Tibanyenda, N., de Bruin, S. H., Haasnoot, C. A. G., van der Marel, G. A., van Boom, J. H., & Hilbers, C. W. (1984) *Eur. J. Biochem. 139*, 19−27.

Uesugi, S., Oda, Y., Ikehara, M., Kawase, Y., & Ohtsuka, E. (1987) *J. Biol. Chem. 262*, 6965−6968.

van de Ven, F. J. M., & Hilbers, C. W. (1988) *Nucleic Acids Res. 16*, 5713−5726.

van der Marel, G. A., van Boeckel, C. A. A., Wille, G., & van Boom, J. H. (1981) *Tetrahedron Lett. 22*, 3887−3890.

Weiner, S. J., Kollman, P. A., Case, D. A., Singh, C., Ghio, C., Alagona, G., Profeta, S., & Weiner, P. (1984) *J. Am. Chem. Soc. 106*, 765−784.

Woodson, S. A., & Crothers, D. M. (1988) *Biochemistry 27*, 445−455.

BI960871E