Cleavage of fragments containing DNA mismatches by enzymic and chemical probes

James BROWN*[1], Tom BROWN† and Keith R. FOX*[2]

*Division of Biochemistry and Molecular Biology, School of Biological Sciences, University of Southampton, Bassett Crescent East, Southampton SO16 7PX, U.K., and †Department of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, U.K.

We prepared synthetic 50-mer DNA duplexes, each containing four mismatched base-pairs in similar positions. We examined their cleavage by DNases I and II, micrococcal nuclease (MNase), methidiumpropyl-EDTA-Fe(II) [MPE-Fe(II)] and hydroxyl radicals. We find that single mismatches only produce subtle changes in the DNase I-cleavage pattern, the most common of which is attenuated cleavage at locations 2–3 bases on the 3′-side of the mismatch. Subtle changes are also observed in most of the DNase II-cleavage patterns, although GT and GG inhibit the cleavage over longer regions and generate patterns that resemble footprints. MNase cleaves the heteroduplexes at the mismatches themselves (except for CC), and in some cases cleaves CpG and CpC steps. None of the mismatches causes any change in the cleavage patterns produced by hydroxyl radicals or MPE-Fe(II). We also examined the cleavage patterns of fragments containing tandem GA mismatches in the sequences RGAY/RGAY and YGAR/YGAR (R, purine; Y, pyrimidine). RGAY causes only subtle changes in the cleavage patterns, which are similar to those seen with single mismatches, except that there are no changes in MNase cleavage. However, YGAR inhibits DNases I and II cleavage over 4–6 bases, and attenuates MPE-Fe(II) and hydroxyl radical cleavage at 2 bases. These changes suggest that this mismatch has a more pronounced effect on the local DNA structure. These changes are discussed in terms of the structural and dynamic effects of each mismatch.

Key words: DNA mismatch, DNase I, hydroxyl radical, methidiumpropyl-EDTA-Fe(II).

INTRODUCTION

Correct base-pairing is essential for the maintenance of genetic fidelity. DNA mismatches, which can lead to mutations, arise as a result of misincorporation during replication, from chemical damage or as intermediates in recombination. All cells therefore possess several mechanisms for repairing mismatched DNA base-pairs [1,2]. Several mechanisms have been proposed for the ways in which repair enzymes might recognize mismatched or damaged DNA, including structural recognition of the distorted DNA duplex, direct read-out of substituents on the mismatched base-pair and changes in the local DNA dynamics [3]. In the present study, we use three enzymic and two chemical cleavage agents to probe the changes that occur when mismatches are introduced into synthetic DNA fragments.

There are eight possible base mismatches and, although it was initially considered that all these would be destabilizing [4,5], it is now known that they often form stable, hydrogen-bonded pairs [6,7]. The structures of some of the best characterized mismatch pairs are shown in Figure 1. The GT pair is one of the most stable mismatches [5,8], and NMR and X-ray crystallography have shown that it adopts a hydrogen-bonded 'wobble'-type conformation [4,9,10]. Small localized backbone perturbations are observed around the GT pair, but these only extend as far as the neighbouring Watson–Crick pair and are dependent on sequence context [8,9,11,12]. In addition, a duplex containing tandem GT mismatches shows only local structural perturbations with little thermodynamic destabilization [10]. Similar distortions are induced by AC mismatches [13]. Two pH-dependent structures also show a 'wobble' conformation, with the most stable structure involving a protonated adenine [7,14,15]. The AC mismatch fits into the B-DNA duplex with only slight distortion and with efficient base stacking [15].

The GA mispair can adopt four different structures [4,7,16–21], which are affected by factors such as pH and sequence context. The G(*anti*)A(*syn*) causes little local or global distortion of the B-DNA helix [11,17,20], whereas the A(*anti*)G(*syn*) mismatch has poor base stacking within the helix which perturbs the DNA backbone [22].

The GG mismatch also adopts a configuration which depends on the local sequence [23] and can be one of the most stable of mismatches [24]. Oligomers containing GG mismatches are globally in the B-form, with a G(*anti*)G(*syn*) conformation [23,25]. In contrast, the G(*anti*)G(*anti*) conformation is highly destabilizing and has poor base stacking within the helix [23,26]. Few data are available about the properties of the AA mismatch, although its stability is known to be sequence-dependent [27].

Little is known about pyrimidine:pyrimidine mismatches, possibly because they are less stable [24,25,27]. Thermodynamic studies on TT, CT and CC mismatches demonstrate that all but the TT are destabilizing in all sequence contexts [27,28].

Therefore it appears that, although some mismatches have properties which are different from those of Watson–Crick base-pairs, not all mismatches destabilize DNA and some fit within a B-DNA helix without causing any distortion [7].

Another duplex lesion that has been extensively studied is the tandem GA mismatch (i.e. 5′-GpA-3′ paired with 5′-GpA-3′).

Abbreviations used: MNase, micrococcal nuclease; MPE, methidiumpropyl-EDTA; R, purine; Y, pyrimidine.

[1] Present address: Division of Structural Biology, University of Oxford, Henry Wellcome Building of Genomic Medicine, Roosevelt Drive, Oxford OX3 7BN, U.K.

[2] To whom correspondence should be addressed (e-mail K.R.Fox@soton.ac.uk).

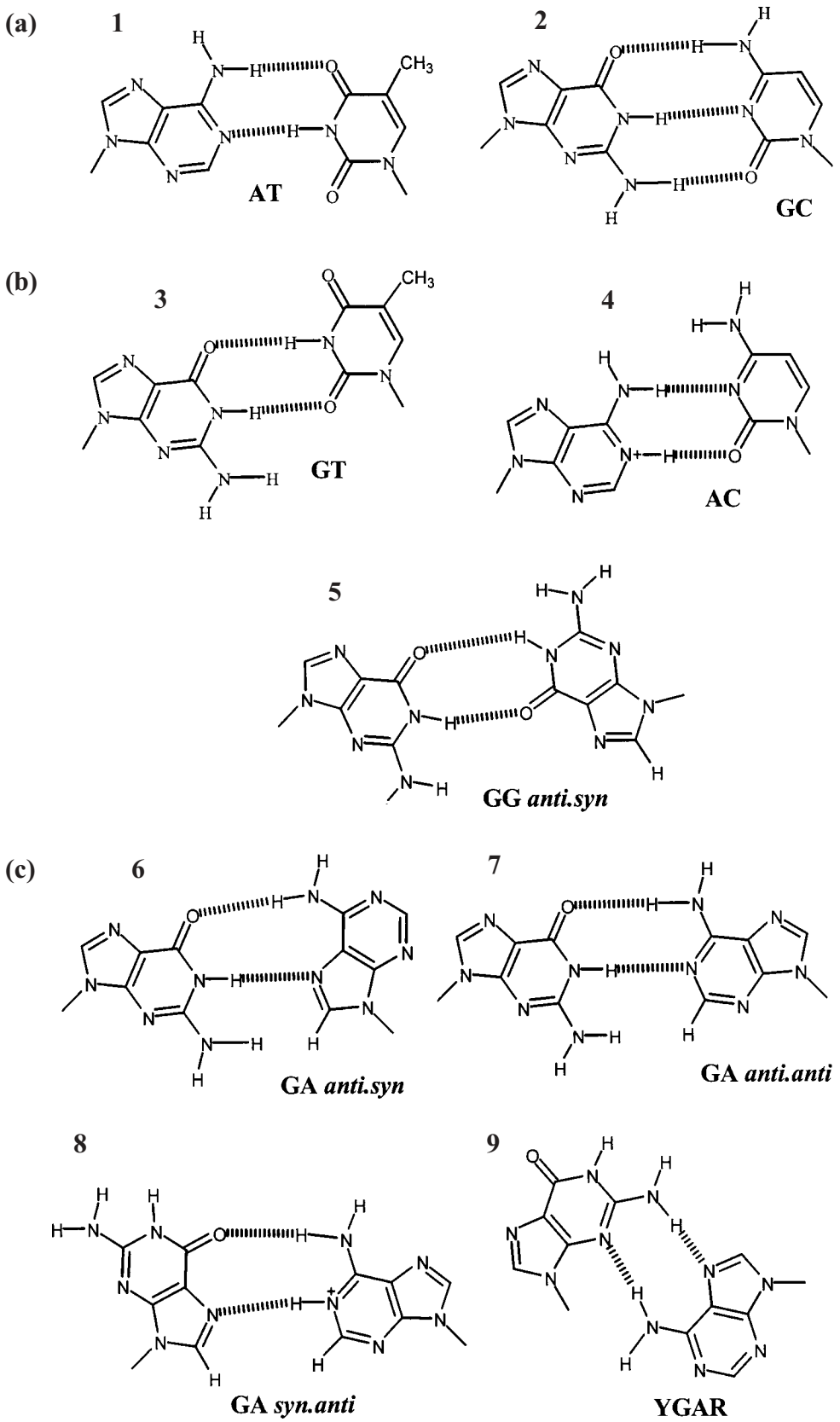

Figure 1 Structures of the best-characterized mismatched base-pairs

(**a**) Structures **1** and **2** show the Watson–Crick base AT and GC base-pairs. (**b**) Structures **3**–**5** show the GT, A + C and GG mismatch base-pairs. (**c**) Structures **6**–**9** show four possible arrangements for the GA mismatch (*anti.syn*, *anti.anti*, *syn.anti* and the structure of tandem GA mismatches in the sequence context YGAR).

Synthetic oligonucleotides containing adjacent GA mismatches form very stable duplexes and have thermodynamic parameters which are similar to those of Watson–Crick base-paired duplexes [29–31]. However, the neighbouring sequence affects the stability, and 5′-YGAR-3′ (R, purine; Y, pyrimidine) is considerably more stable than 5′-RGAY-3′ [30,32–36]. The YGAR sequence is easily accommodated in B-DNA structures [29,30,37]. Strong inter-strand base stacking has also been observed in the non-symmetrical sequence RGAR:YGAY [38]. Further evidence for the stability of tandem GA mismatches is seen with a decamer containing three pairs of adjacent GA mismatches, i.e. a total of 6 out of 10 mismatched bases [39]. Although this was less stable than a duplex in which the GA mismatches had been replaced by GC base-pairs, it was more stable than the one in which the GA mismatches had been replaced by AT [39]. Again, the duplex had an overall B-DNA conformation with strong base stacking [39].

In the present study, we use the known cleavage preferences of three enzymic and two chemical agents to obtain information about DNA structure around different mismatches.

DNase I is widely used as a footprinting agent and cuts the O3′-P bond in double-stranded DNA, nicking the DNA backbone. DNase I-cleavage patterns are uneven, reflecting sequence-dependent variations in local DNA structure [40–42]. The enzyme binds by inserting an exposed loop into the DNA minor groove, bending the DNA towards the major groove [43,44]. $A_n.T_n$ tracts are poor substrates since they possess a narrow minor groove. GC-rich regions are also poor substrates, as they have a more rigid structure, which impedes the bending that is considered to be an essential feature of the enzymic reaction. We would anticipate that changes in DNA global structure, which affect the binding of this enzyme, will produce attenuated cleavage over a number of base-pairs, whereas local changes in backbone geometry will only affect the cleavage of individual bonds.

DNase II is a double-stranded specific nuclease which produces single-strand breaks in DNA, cutting the O5′–P bond. It produces a very uneven cutting pattern [40,41] and regions of good cleavage do not correlate across the two DNA strands. The enzyme cuts best in oligopurine tracts which contain both A and G and is supposed to recognize a structural feature of this oligopurine strand.

Micrococcal nuclease (MNase) cleaves the O5′–P bond with optimum cleavage at XpA and XpT sites [40,45,46]. The cleavage of AT-rich regions is explained by the structure of the enzyme [47–50], which reveals a narrow cleft at the active site that can accommodate only a single DNA strand. AT base-pairs are preferentially cleaved as a result of their lower stability when compared with GC pairs. Since this enzyme is sensitive to local DNA dynamics, we would expect that unstable mismatches would cause enhanced cleavage.

Hydroxyl radicals, produced by the Fenton reaction, cleave DNA in a sequence-independent manner [51,52]. It is supposed that free radicals react with C4′, leading to DNA strand scission. Hydroxyl radical cleavage is attenuated in regions possessing a narrow minor groove, such as $A_n.T_n$ tracts, and at regions of bent DNA [53,54].

Methidiumpropyl–EDTA (MPE)-Fe(II) consists of an intercalator (methidium) to which an EDTA moiety is tethered, generating free radicals at its DNA-binding site [55,56]. Since the reaction depends on intercalation of the phenathridine chromophore, cleavage is attenuated at poor intercalation sites, such as $A_n.T_n$ tracts.

MATERIALS AND METHODS

Chemicals and enzymes

Acrylamide stock solutions (Sequagel and Accugel) were purchased from National Diagnostics (Unit 4, Fleet Business Park, Itlings Lane, Hessle, Hull, U.K.). [α-^{32}P]dATP (specific radioactivity of 3000 Ci/mmol) was purchased from Amersham International (Arlington, IL, U.S.A.). Reverse transcriptase, DNases I and II, MNase and MPE were obtained from Sigma.

Oligonucleotides

Oligonucleotides were purchased from Oswel DNA Services (Southampton, U.K.) and they were synthesized on a 0.2 μM scale and purified by HPLC. Stock solutions were stored at − 20 °C.

Duplex DNA fragments were prepared by mixing equimolar amounts of complementary oligonucleotides in 10 mM Tris/HCl (pH 7.5) containing 0.1 mM EDTA. The mixture was heated to 100 °C and slowly cooled to room temperature (20 °C). The resultant duplexes have 5′-overhangs and were labelled at the 3′-end of the top strand (Figure 2) by filling in with [α-^{32}P]dATP, using reverse transcriptase. The radiolabelled DNA fragments were purified on non-denaturing 10 % (w/v) polyacrylamide gels, eluted and resuspended in 10 mM Tris/HCl (pH 7.5) containing 0.1 mM EDTA so as to give approx. 20 c.p.s./μl as determined on a hand-held Geiger counter.

The sequences of these duplexes (Figure 2) were chosen so as to generate four mismatches in similar positions on each fragment, such that most duplexes share the same labelled strand. This kept the sequence environment as constant as possible while reducing the number of oligonucleotides required. The labelled strands of the Watson–Crick, GT, AC, AG GG, AA and YGAR duplexes are identical.

Cleavage reactions

DNase I

Radiolabelled DNA (2 μl) was mixed with 2 μl of 10 mM Tris/HCl (pH 7.5) containing 10 mM NaCl. Digestion was initiated by adding 2 μl of DNase I diluted in 20 mM NaCl, 2 mM $MgCl_2$ and 2 mM $MnCl_2$. The enzyme concentration was adjusted to achieve single-hit kinetics and it was typically approx. 0.01 unit/ml. The reaction was stopped after 1 min by the addition of 4 μl formamide containing 10 mM EDTA and 0.1 % (w/v) Bromophenol Blue.

MNase

Radiolabelled DNA (2 μl) was mixed with 2 μl of 50 mM Tris/HCl (pH 7.6) containing 2 mM $CaCl_2$. Digestion was initiated by adding 2 μl of MNase diluted in the same buffer. The reaction was stopped after 1 min by the addition of 4 μl of formamide containing 10 mM EDTA and 0.1 % (w/v) Bromophenol Blue.

DNase II

Radiolabelled DNA (2 μl) was mixed with 2 μl of 20 mM ammonium acetate (pH 5.4) containing 1 mM EDTA. Digestion was initiated by adding 2 μl of DNase II diluted in the same buffer. The reaction was stopped after 1 min by adding 4 μl of formamide containing 10 mM EDTA and 0.1 % (w/v) Bromophenol Blue and cooling on solid CO_2.

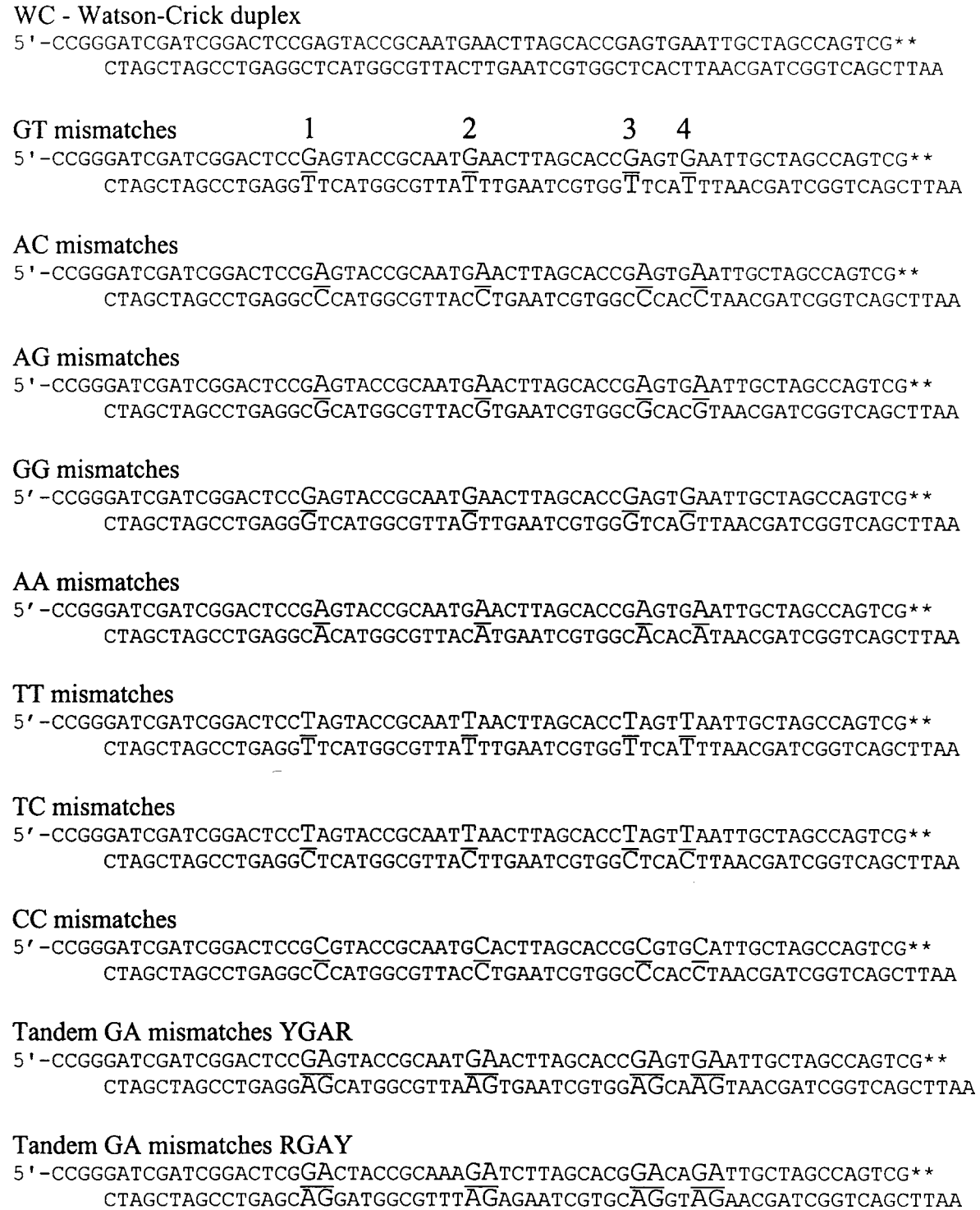

Figure 2 Sequences of the mismatch-containing duplexes used in the present study

The four mismatches in each fragment are numbered 1–4 and are indicated by the larger, underlined text. The fragments were each labelled at the 3′-end of the upper strand with [α-^{32}P]dATP, filling the sticky ends, as indicated by the double asterisks.

Hydroxyl radicals

Radiolabelled DNA (3 μl) was mixed with 3 μl of 10 mM Tris/HCl containing 10 mM NaCl. Hydroxyl radical cleavage was performed by adding 6 μl of a freshly prepared mixture containing 30 μM ferrous ammonium sulphate, 50 μM EDTA, 2 mM ascorbic acid and 0.02 % H_2O_2. The reaction was stopped after 10 min by ethanol precipitation. The DNA was resuspended in 6 μl of formamide containing10 mM EDTA and 0.1 % (w/v) Bromophenol Blue.

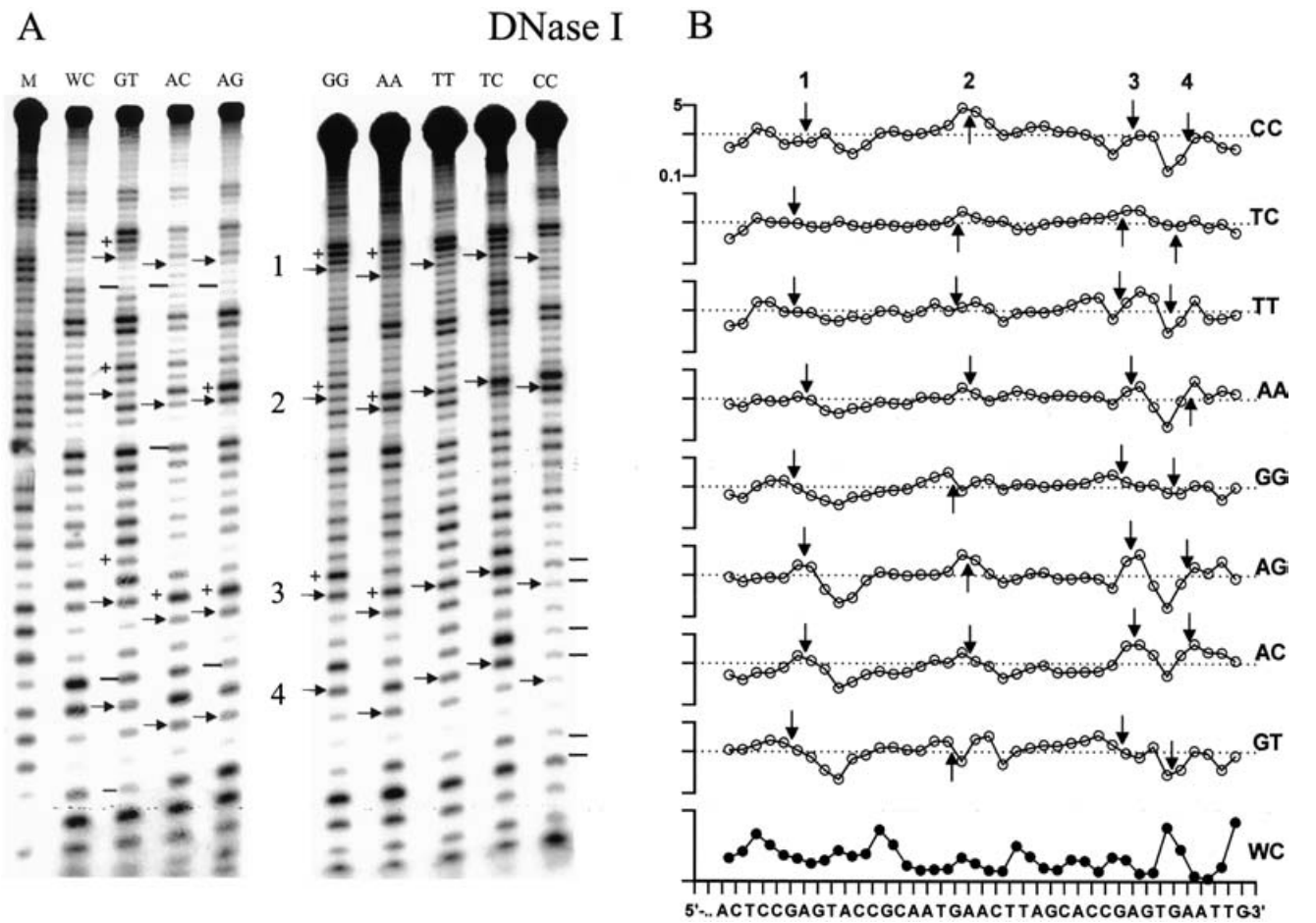

Figure 3 DNase I cleavage of the mismatch-containing fragments

(**A**) DNase I-cleavage patterns of the different fragments. The identity of the mismatch is shown at the top of each gel lane. Arrows indicate the positions of the mismatches, which are numbered 1–4 as shown in Figure 2. Tracks labelled 'M' are Maxam–Gilbert markers specific for purines. The positions of enhanced and attenuated cleavage relative to the Watson–Crick control are indicated by + and −. The fragments were each labelled at the 3′-end of the upper strand shown in Figure 2; the sequences therefore run from 3′ to 5′ from the bottom to the top of the gel. (**B**) Quantitative analysis of the cleavage data, derived from phosphorimage analysis of the digestion patterns. The lower plot shows the relative cleavage of each bond in the Watson–Crick duplex, plotted on a linear scale (arbitrary units). All the other plots are differential cleavage plots showing the intensity of each band in the heteroduplex divided by the intensity of the corresponding band in the Watson–Crick control, plotted on a logarithmic scale. All the vertical scales are the same, and the dotted lines indicate the positions at which cleavage of the heteroduplex is equal to that of the control. Arrows indicate the positions of the mismatches.

MPE-Fe(II)

Radiolabelled DNA (5 μl) was mixed with 5 μl of 20 μM MPE and 20 μM ferrous ammonium sulphate. The mixture was incubated at room temperature for 5 min, before starting the reaction by adding 3 μl of 10 mM dithiothreitol. The reaction was stopped after 45 min by ethanol precipitation.

Gel electrophoresis

Samples were heated at 100 °C for 3 min before cooling on ice and loading on to 40 cm polyacrylamide gel [15 % (w/v)] containing 8 M urea. Gels were run at 1500 V for approx. 2 h before fixing in 10 % (v/v) acetic acid, transferring to Whatmann 3 MM paper and drying under vacuum at 80 °C. Dried gels were exposed to a Kodak storage phosphor screen, which was scanned using a Molecular Dynamics Storm 860 PhosphorImager and analysed using ImageQuant software. Bands in the digest were assigned by comparison with Maxam–Gilbert marker lanes specific for purines. It should be noted that DNase I, hydroxyl radicals, MPE-Fe(II) and Maxam–Gilbert sequencing reactions produce labelled fragments with a phosphate at the 5′-end, whereas DNase II and MNase generate fragments with hydroxyl at the 5′-end. As a result, the products of DNase II and MNase digestion run slower than the marker lanes; this difference is greatest for shorter fragments.

Quantitative analysis

The digestion patterns were quantified by measuring the intensity of each band in the phosphorimaging data using ImageQuant software. The intensity of each band in the heteroduplex cleavage ladder was divided by that of the equivalent band in the control (Watson–Crick) lane, and normalized according to the total radioactivity in each lane, generating differential cleavage plots. These plots show the relative cleavage of each bond in the heteroduplex relative to that in the control. Values < 1 correspond to reduced cleavage in the heteroduplex DNA, whereas values > 1 correspond to enhanced cleavage. The relative cleavage pattern of the control (Watson–Crick) duplexes are also shown in the Figures.

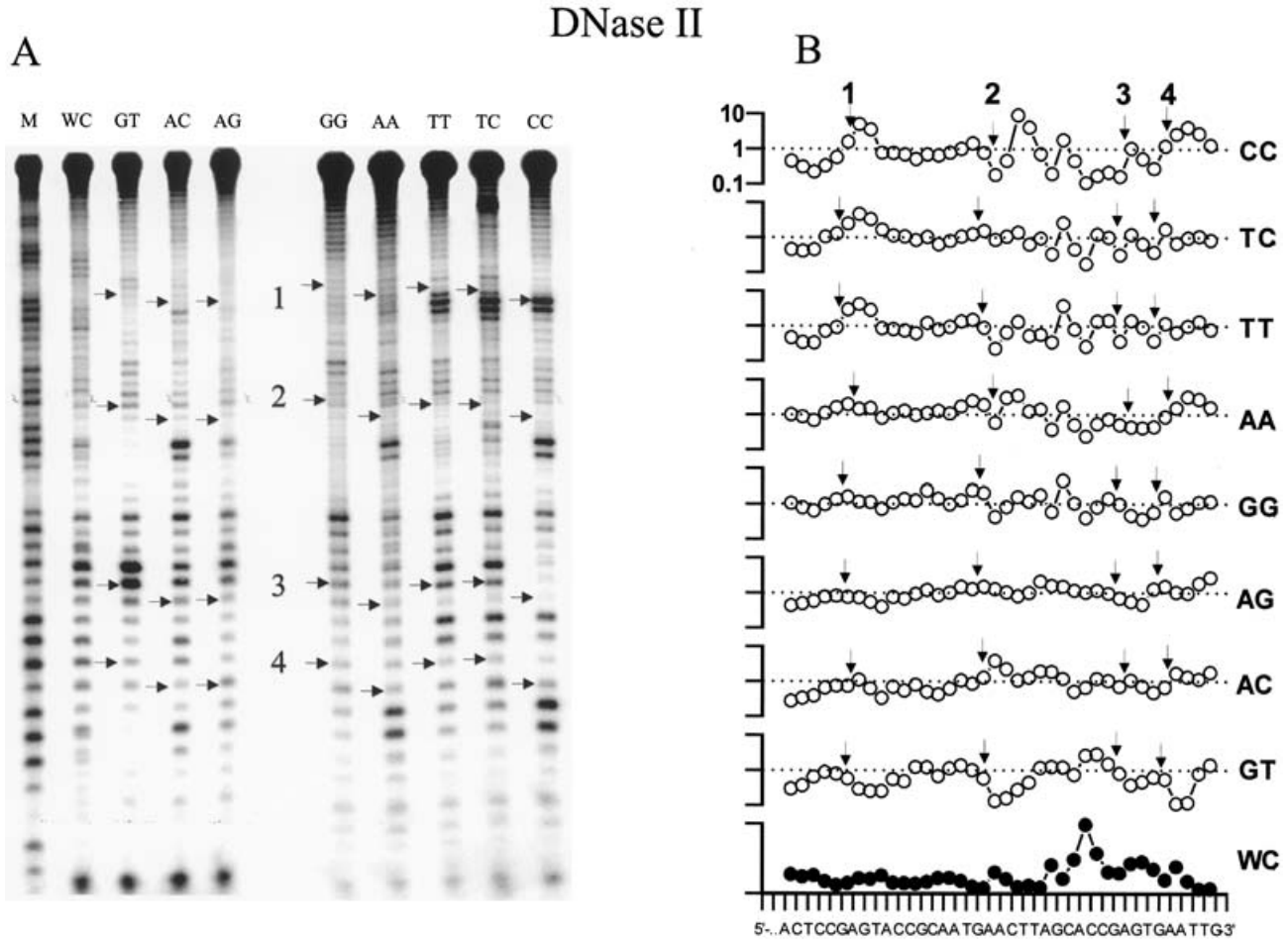

Figure 4 DNase II cleavage of the mismatch-containing fragments

(**A**) DNase II-cleavage patterns of the different fragments. The identity of the mismatch is shown at the top of each gel lane. Arrows indicate the positions of the mismatches, which are numbered 1–4 as shown in Figure 2. Tracks labelled 'M' are Maxam–Gilbert markers specific for purines. The fragments were each labelled at the 3′-end of the upper strand shown in Figure 2; the sequences therefore run from 3′ to 5′ from the bottom to the top of the gel. (**B**) Quantitative analysis of the cleavage data, derived from phosphorimage analysis of the digestion patterns. The lower plot shows the relative cleavage of each bond in the Watson–Crick duplex, plotted on a linear scale (arbitrary units). All the other plots are differential cleavage plots showing the intensity of each band in the heteroduplex divided by the intensity of the corresponding band in the Watson–Crick control, plotted on a logarithmic scale. All the vertical scales are the same, and the dotted lines indicate the positions at which cleavage of the heteroduplex is equal to that of the control. Arrows indicate the positions of the mismatches.

RESULTS

Single mismatches

DNase I

Figure 3 shows the results of DNase I digestion of the synthetic DNA fragments containing each of the eight mismatches, together with the related Watson–Crick homoduplex. The cleavage patterns are shown in Figure 3(A), and the differential cleavage plots derived from these data are shown in Figure 3(B).The locations of the mismatches are indicated by arrows; these are numbered 1–4 from the 5′-end (at the top of the gel). It can be seen that all the fragments are good substrates for DNase I and that, in general, the mismatches do not cause any global changes in the cleavage patterns.

Looking first at the Watson–Crick duplex, the enzyme produces an uneven ladder of cleavage products, in which some bonds are cut better than others. This is most easily seen in the cleavage plot shown in the lower panel of Figure 3(B) and reveals that the best cleavage sites are located at TpC, CpG, TpT, TpG and TpG. These differences are presumed to reflect variations in local DNA structure, in particular the width of the minor groove and DNA flexibility. For the mismatch-containing duplexes, we predict that alterations which affect global DNA structure will prevent DNase I binding (and hence cleavage) over a wide region, whereas those affecting the local conformation should only affect cleavage at the mismatch itself. The observation that the mismatches only produce local changes in the cleavage pattern suggests that they do not drastically affect the DNA structure. It should be noted that the Watson–Crick duplex and the GT, AC, AG, GG and AA heteroduplexes contain the same labelled strand, so that differences in their cleavage patterns can only be due to the presence of a mismatched base on the opposing strand.

DNase I cleavage of the GT-containing heteroduplex is very similar to that of the Watson–Crick control. The cleavage is not affected at any of the GT mismatches (indicated by arrows in Figure 3B), except at site 4 towards the bottom of the gel where two mismatches are separated by only 4 base-pairs. This suggests that the conformation of the scissile bond has not been altered. However, examination of the patterns reveals a reduced cleavage

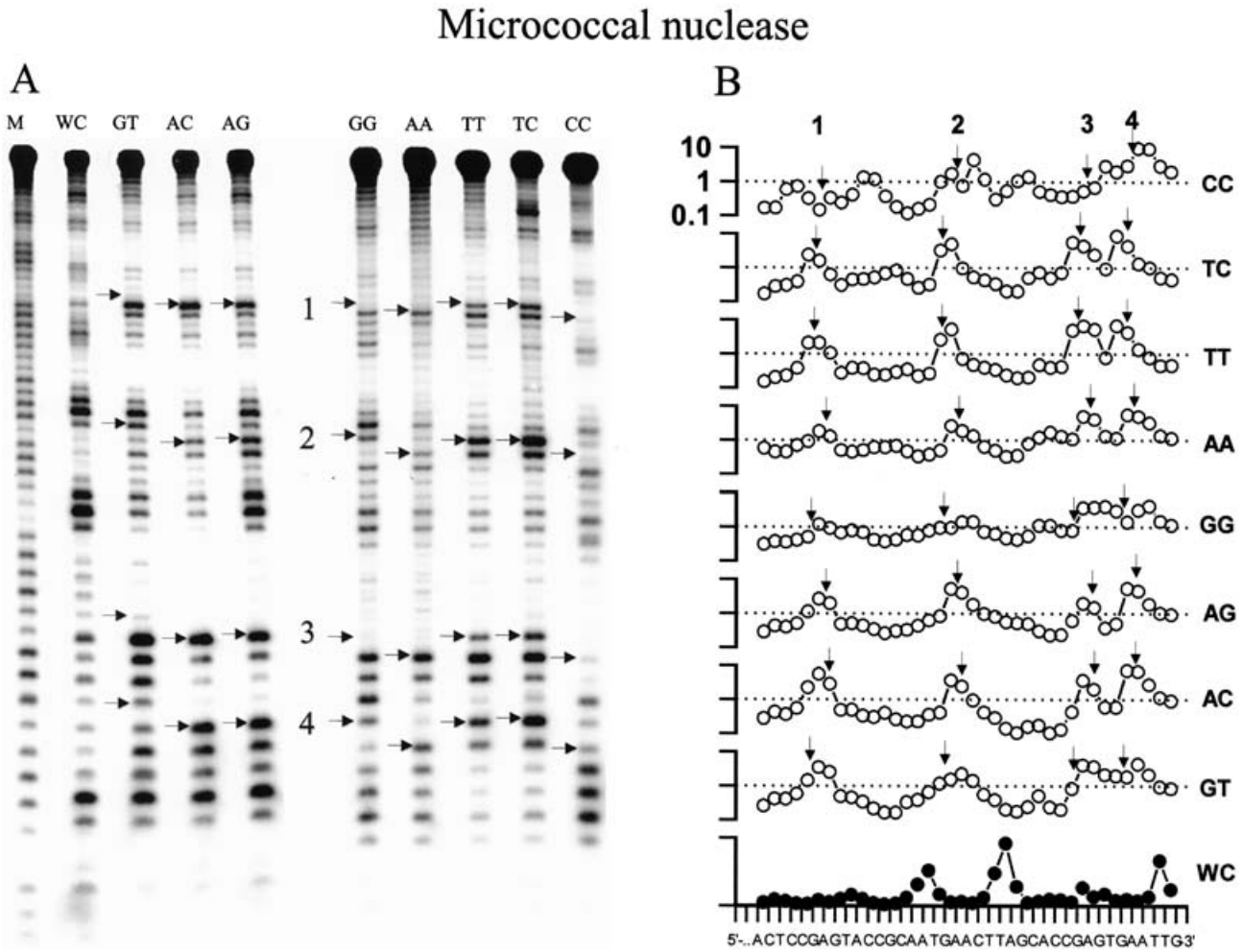

Figure 5 MNase cleavage of the mismatch-containing fragments

(**A**) MNase-cleavage patterns of the different fragments. The identity of the mismatch is shown at the top of each gel lane. Arrows indicate the positions of the mismatches, which are numbered 1–4 as shown in Figure 2. Tracks labelled 'M' are Maxam–Gilbert markers specific for purines. The fragments were each labelled at the 3′-end of the upper strand as shown in Figure 2; the sequences therefore run from 3′ to 5′ from the bottom to the top of the gel. (**B**) Quantitative analysis of the cleavage data, derived from phosphorimage analysis of the digestion patterns. The lower plot shows the relative cleavage of each bond in the Watson–Crick duplex, plotted on a linear scale (arbitrary units). All the other plots are differential cleavage plots showing the intensity of each band in the heteroduplex divided by the intensity of the corresponding band in the Watson–Crick control, plotted on a logarithmic scale. All the vertical scales are the same, and the dotted lines indicate the positions at which cleavage of the heteroduplex is equal to that of the control. Arrows indicate the positions of the mismatches.

at 3 bonds on the 3′-side of mismatches 1, 3 and 4 (the equivalent position at site 2 shows very poor cleavage in the Watson–Crick control). Cleavage is also increased at 1–3 bonds on the 5′-side of mismatches 1, 2 and 3.

The AC-containing heteroduplex shows similar changes, with attenuated cleavage of the bonds located 3 bases below sites 1–3. A similar effect is seen with the AG-containing heteroduplex, which also shows enhanced cleavage above sites 2 and 3. The GG- and AA-containing heteroduplexes show increased cleavage of the bond on the 5′-side of mismatches 1–3. The cleavage patterns with the TT-, TC- and CC-containing heteroduplexes are very similar to that of the Watson–Crick control, although the CC heteroduplex shows very little enzyme cleavage towards the bottom of the gel where two mismatches are close together.

In summary, these single mismatches do not produce any large changes in the DNase I-cleavage patterns, except when two CC mismatches are close together. Instead, the mismatches cause subtle changes in the cleavage patterns, producing attenuated cleavage at bonds located 2–3 bases on their 3′-side and enhanced cutting on the 5′-side.

DNase II

The same heteroduplexes were digested with DNase II and their cleavage patterns are shown in Figure 4(A). Differential cleavage plots derived from these data are shown in Figure 4(B). DNase II produces a less even pattern than DNase I, with few positions of good cleavage. In contrast with DNase I, the mismatches alter the DNase II-cleavage patterns, which show distinct differences from the Watson–Crick control. The Watson–Crick duplex shows good DNase II cleavage in only one location, around the sequence CACC, with maximum cleavage at the ApC bond. This is most easily seen in the cleavage plot shown in the lower panel of Figure 4(B).

The GT-containing heteroduplex shows attenuated cleavage of several bonds on the 3-side of mismatches 1, 2 and 4, producing

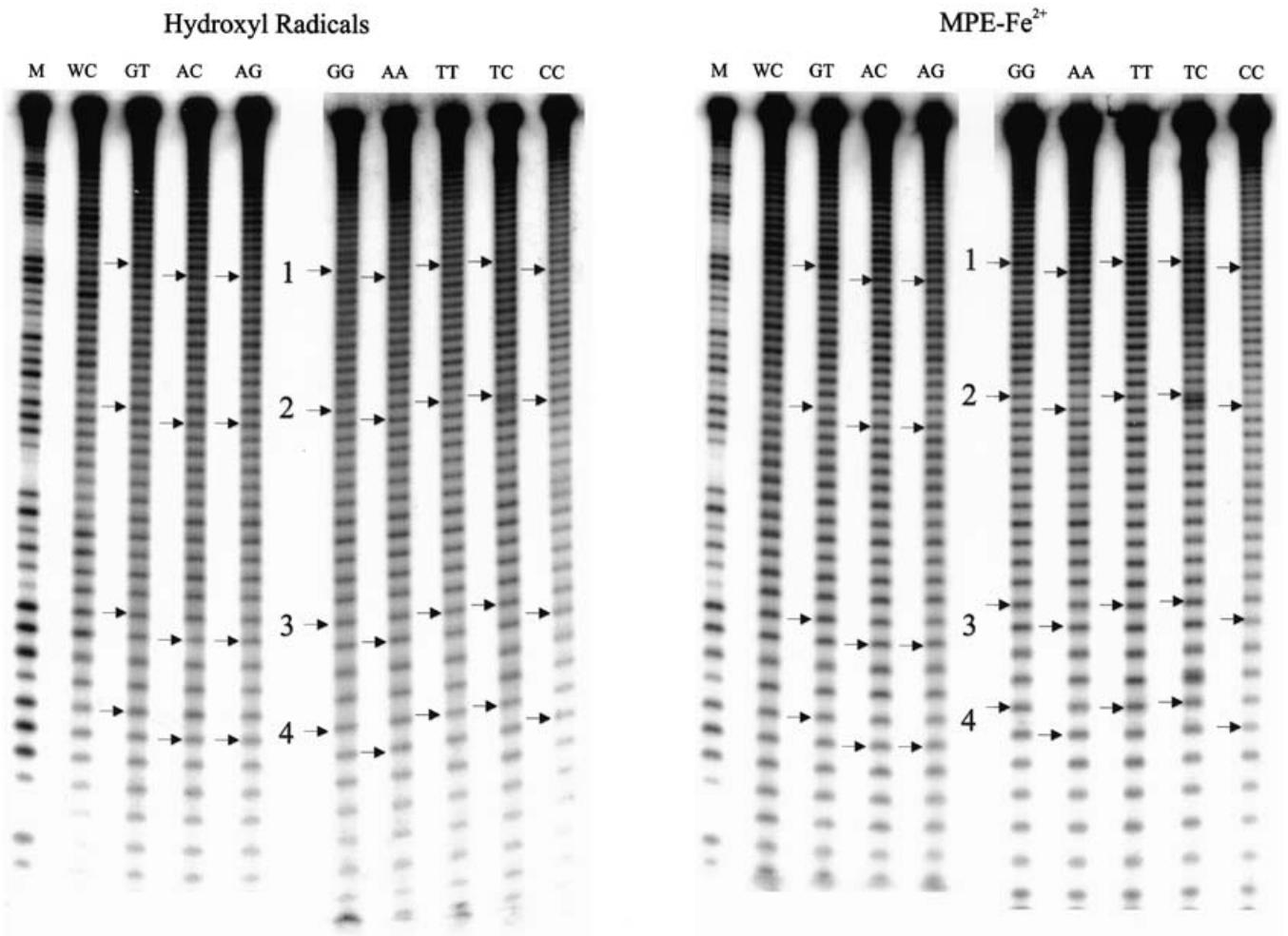

Figure 6 Hydroxyl radical and MPE-cleavage patterns of the mismatch-containing fragments

The identity of the mismatch is shown at the top of each gel lane. Arrows indicate the positions of the mismatches, which are numbered 1–4 as shown in Figure 2. Tracks labelled 'M' are Maxam–Gilbert markers specific for purines. The fragments were each labelled at the 3′-end of the upper strand as shown in Figure 2; the sequences therefore run from 3′ to 5′ from the bottom to the top of the gel.

regions in which there is almost no cleavage, very much like a footprint. In contrast, DNase II cleavage of the AC and AA heteroduplexes is similar to that in the control, although there is enhanced cleavage at 1 or 2 bases on the 3′-side of mismatches 1, 2 and 4. DNase II cleavage of the AG-containing heteroduplex is also very similar to that of the homoduplex, except for a slight attenuation in cleavage on the 3′-side of mismatches 1 and 3. The GG-containing heteroduplex is a poor substrate for this enzyme with reduced cleavage on both sides of each mismatch, producing a pattern similar to that seen with the GT mismatch, except around site 3. Two of the four TT mismatches (sites 1 and 3) show increased cleavage 2–3 bonds on the 3′-side of the mismatch, and a similar pattern is seen with the TC heteroduplex. Finally, the CC heteroduplex produces a pattern that is almost similar to that seen with the AA and AC mismatches, except around site 1.

MNase

Figure 5(A) shows MNase digestion of the same heteroduplexes. Differential cleavage plots derived from these data are shown in Figure 5(B). The different mismatches have clear effects on the cleavage patterns, and in several instances, novel cleavage products are evident in the vicinity of the mispaired bases. Looking first at the control homoduplex, there are three regions of good cleavage (above and below site 2 and below site 4), which are located in the sequences AAT, TTA and AATT, consistent with known preferences of MNase for AT-rich regions. These are most easily seen in the bottom panel of Figure 5(B), which shows the cleavage pattern of the homoduplex.

The heteroduplexes all show increased MNase cleavage at the location of the mismatches. The only exception is the CC heteroduplex for which there is a general reduction in cleavage. In many of these cases MNase cleaves CpG and CpC steps in the vicinity of the mismatch. The GT and GG heteroduplexes show enhanced cleavage at 1–2 bonds on the 3′-side of each mismatch. In contrast, the AC, AG, AA, TT and TC heteroduplexes show enhanced cleavage at the mismatch site itself. The CC heteroduplex shows a different pattern with a reduction in cleavage at all the mismatches, with a slight augmentation of cleavage at 1–2 bonds on their 3′-side.

Hydroxyl radicals and MPE-Fe(II)

The cleavage patterns obtained using hydroxyl radicals and MPE-Fe(II) are shown in Figure 6. Like the homoduplex, all fragments produce fairly even cleavage ladders. Although there are some

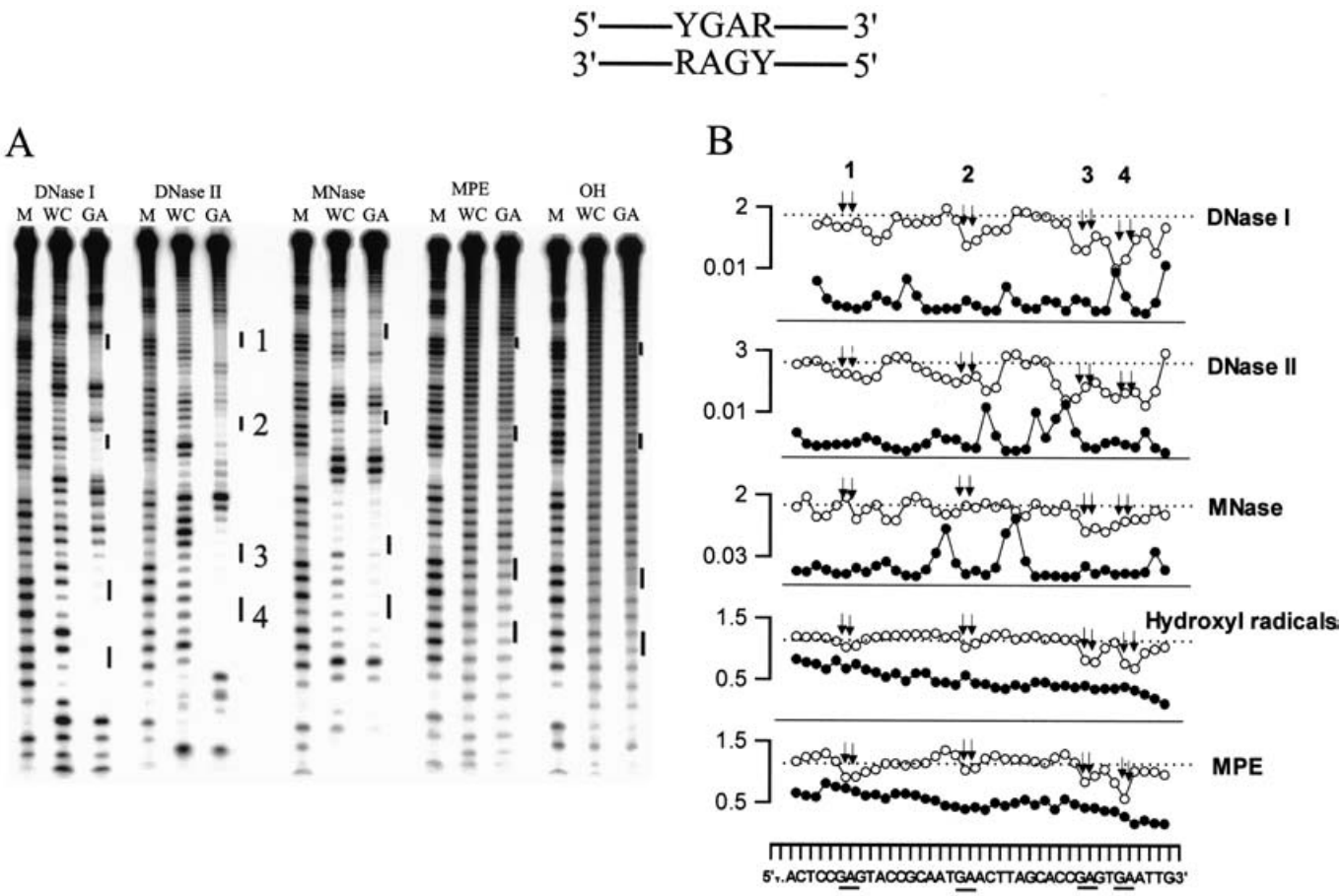

Figure 7 Enzymic and chemical cleavage of the fragments containing tandem GA mismatches in the context YGAR

(**A**) Enzymic and chemical cleavage patterns. The cleavage agents are indicated at the top of each set of three lanes. Tracks labelled 'M' are Maxam–Gilbert markers specific for purines. Tracks labelled 'WC' show results for the Watson–Crick homoduplex, whereas 'GA' indicates the YGAR heteroduplexes. The bars show the positions of the tandem GA mismatches, which are numbered 1–4 as shown in Figure 2. The fragments were each labelled at the 3′-end of the upper strand as shown in Figure 2; the sequences therefore run from 3′ to 5′ from the bottom to the top of the gel. (**B**) Quantitative analysis of the cleavage data, derived from phosphorimage analysis of the digestion patterns. For each cleavage agent, the lower plot (filled circles) shows the relative cleavage of each bond in the Watson–Crick duplex, plotted on a linear scale (arbitrary units). All other plots (open circles) are differential cleavage plots showing the intensity of each band in the heteroduplex divided by the intensity of the corresponding band in the Watson–Crick control, plotted on a logarithmic scale. The dotted lines indicate the positions at which cleavage of the heteroduplex is equal to that of the control. Arrows indicate the positions of the mismatches.

fluctuations in band intensity within each fragment, there are no significant differences between the various heteroduplexes. These results confirm that the mismatches do not produce any major alterations in the dimension of the minor groove and have little effect on intercalation of the methidium chromophore.

Tandem GA mismatches

Several thermodynamic and NMR studies have shown that duplexes which contain two adjacent GA mismatches can be as stable as a Watson–Crick homoduplex. However, the neighbouring bases affect the stability, and 5′-YGAR-3′ is considerably more stable than 5′-RGAY-3′. The results of digestion studies on fragments containing tandem GA mismatches are presented in Figures 7 and 8.

YGAR

Figure 7(A) shows enzymic and chemical cleavage patterns for the YGAR heteroduplex. Differential cleavage plots derived from these data are shown in Figure 7(B). Cleavage by all the probes has been affected by this tandem mismatch. DNase I digestion is clearly inhibited around each mismatch, producing what looks like a footprint, covering 4–6 bases. DNase II digestion also shows a large region of protection around each mismatch. In contrast, cleavage by MNase is hardly affected by the presence of these tandem mismatches, and the cutting is still restricted to the AT-rich regions.

The chemical probes MPE-Fe(II) and hydroxyl radicals produce even ladders of cleavage products in the homoduplex fragment. However, cleavage of the YGAR heteroduplex is attenuated over 2 bonds at each mismatch site.

These results suggest that tandem GA mismatches in the sequence YGAR cause alterations in DNA structure which prevent enzymes from binding over a large region and occlude hydroxyl radicals from entering the minor groove.

RGAY

Figure 8(A) shows the cleavage patterns obtained with the RGAY heteroduplex. Differential cleavage plots derived from these data are shown in Figure 8(B). This mismatch pair has a less

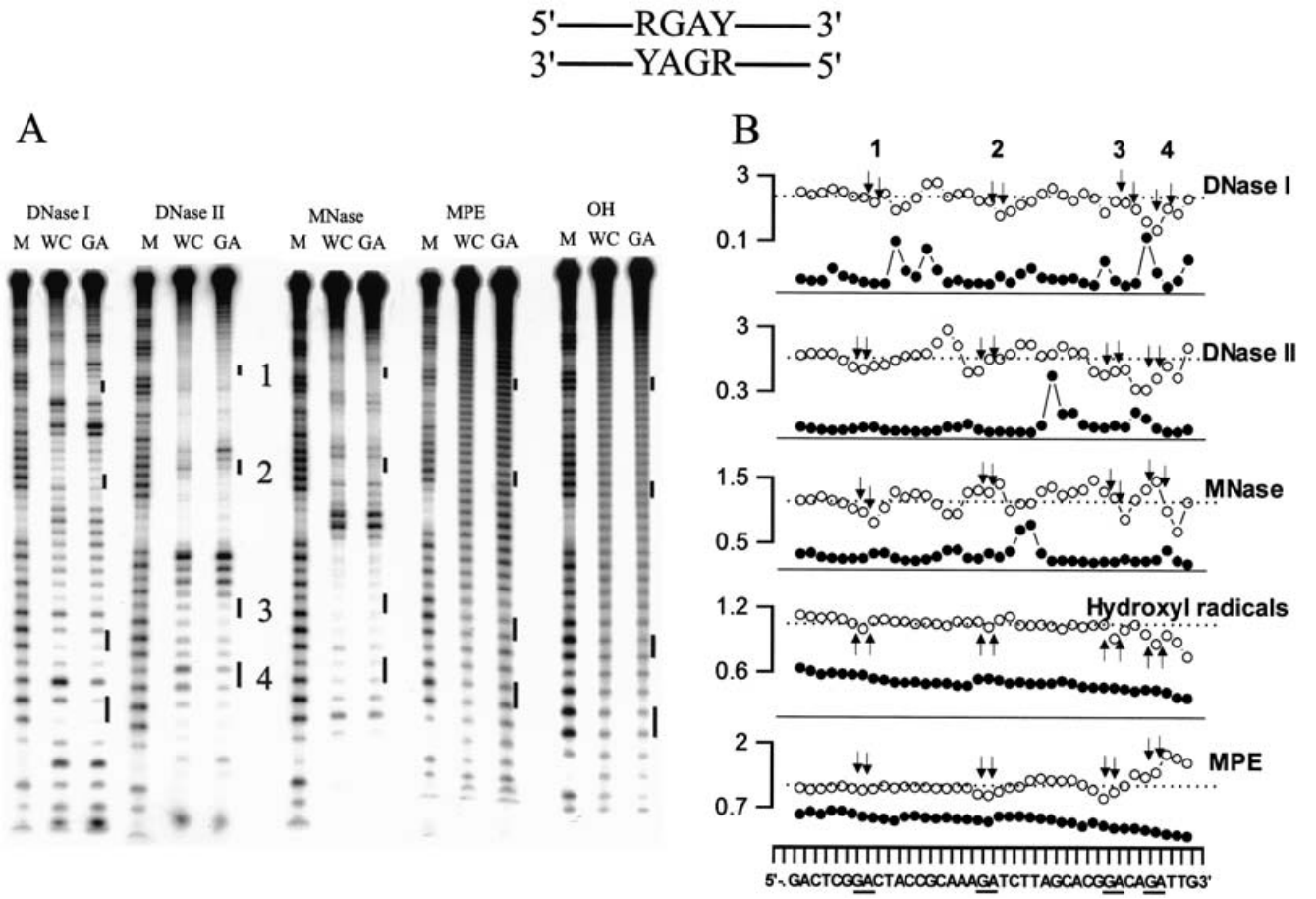

Figure 8 Enzymic and chemical cleavage of the fragments containing tandem GA mismatches in the context RGAY

(**A**) Enzymic and chemical cleavage patterns. The cleavage agents are indicated at the top of each set of three lanes. Tracks labelled 'M' are Maxam–Gilbert markers specific for purines. Tracks labelled 'WC' show results for the Watson–Crick homoduplex, whereas 'GA' indicates the RGAY heteroduplex. The bars show the positions of the tandem GA mismatches, which are numbered 1–4 as shown in Figure 2. The fragments were each labelled at the 3′-end of the upper strand shown in Figure 2; the sequences therefore run from 3′ to 5′ from the bottom to the top of the gel. (**B**) Quantitative analysis of the cleavage data, derived from phosphorimage analysis of the digestion patterns. For each cleavage agent, the lower plot (filled circles) shows the relative cleavage of each bond in the Watson–Crick duplex, plotted on a linear scale (arbitrary units). All the other plots (open circles) are differential cleavage plots showing the intensity of each band in the heteroduplex divided by the intensity of the corresponding band in the Watson–Crick control, plotted on a logarithmic scale. The dotted lines indicate the positions at which cleavage of the heteroduplex is equal to that of the control. Arrows indicate the positions of the mismatches.

pronounced effect on cleavage than the YGAR heteroduplex. There is a slight attenuation of cleavage around the mismatch sites in the DNase I digestions, and at sites 3 and 4 with DNase II. As with the YGAR heteroduplex, MNase cleavage is almost unaffected by the presence of the RGAY mismatch. Similarly, cleavage is not affected by the two chemical probes.

These results suggest that tandem GA mismatches in the RGAY sequence context do not cause substantial changes in the DNA structure; enzymes are still capable of binding and cleaving effectively and the minor groove remains accessible to chemical cleavage agents.

DISCUSSION

Several of the heteroduplexes (most notably GT, AC and AG) show reduced DNase I cleavage at 2–3 bonds on the 3′-side of the mismatched base. This is consistent with the known 3′-staggered cleavage of this enzyme across the two DNA strands and suggests that these mismatches cause small changes in the local DNA structure, which affect binding to the minor groove. If the mismatch only affected the orientation of the cleaved phosphate, this would only be expected to affect DNase I cleavage at the mismatched base itself. The observation that the TC and GG heteroduplexes produce very few changes in DNase I cleavage implies that these mismatches do not significantly affect the local backbone structure.

All the mismatches, except AG, cause changes in digestion by DNase II. In several instances, there is enhanced cleavage of 2–3 bases on the 3′-side of the mismatch. In contrast, the GT and GG heteroduplexes show attenuated cleavage of several bonds on the 3′-side of each mismatch, producing what appears as a small footprint. Interestingly, DNase II cleavage of the AG heteroduplex is very similar to the Watson–Crick duplex.

The mismatches appear to have a much greater effect on MNase cleavage. The three regions of good cleavage in the homoduplex are replaced by four cleavage maxima in the heteroduplexes. Whereas DNase I is sensitive to DNA structure, in particular the minor groove width, MNase is more dependent on local DNA dynamics and cuts only at pA and pT steps. The mode of action of MNase involves separation of the two DNA strands accommodated within a narrow hydrophobic cleft [45,47,48].

Enhanced cleavage of heteroduplexes by MNase may therefore indicate a local decrease in helix stability around the mismatch. These results suggest that mismatches facilitate strand separation and thereby enhance MNase cleavage. However, this is not the case for the CC-containing heteroduplex, for which MNase cleavage is generally poor. These results indicate that all the mismatches (with the exception of CC) alter the local dynamics of the DNA helix.

These mismatches do not affect cleavage by the chemical agents MPE-Fe(II) and hydroxyl radicals. This suggests that they do not affect access to the minor groove or intercalative binding. This is consistent with the relatively small changes in DNase I cleavage, which suggests that there are no major changes in the minor groove width.

The YGAR heteroduplex shows much greater changes in cleavage by all the probes than all the single-base mismatches, and its behaviour is very different from the RGAY duplex. DNase I cleavage produces a footprint-like pattern around the YGAR mismatch and the hydroxyl-radical cleavage is attenuated. These results suggest that the tandem YGAR mismatches cause a significant distortion to the minor groove. Such deformations generate a local DNA structure that does not bind DNase I and this prevents nuclease access over several bases. These structural changes are sufficient to affect local cleavage by hydroxyl radicals and MPE-Fe(II). In contrast, the cleavage by MNase is hardly affected by the presence of this tandem mismatch. This implies that tandem GA mismatches in the YGAR context generate a stable duplex, for which dynamic breathing is similar to a GC-containing Watson–Crick duplex. It therefore appears that YGAR mismatches distort the minor groove and affect the local DNA structure so as to prevent DNase I binding over a wide region. These findings are consistent with NMR structures of tandem GA mismatches in the context of YGAR, which shows that the minor groove is filled by the adenine base [31,33,34,36,37,39]. In contrast, the RGAY duplex shows only minor changes in cleavage compared with the Watson–Crick duplex. Duplexes containing this mismatch are known to adopt a more conventional B-DNA structure, which evidently does not preclude enzymic and chemical cleavage.

This work was supported by a grant from the Biotechnology and Biological Sciences Research Council (Swindon, U.K.).

REFERENCES

1 Friedberg, E. C., Walker, G. C. and Siede, W. (1995) DNA Repair and Mutagenesis, ASM Press, Washington, DC

2 Loeb, L. A. and Kunkel, T. A. (1982) Fidelity of DNA synthesis. Annu. Rev. Biochem. **52**, 429–457

3 Rajski, S. R., Jackson, B. A. and Barton, J. K. (2000) DNA repair: models for damage and mismatch recognition. Mutation Res. **447**, 49–72

4 Patel, D. J., Kozlowski, S. A., Ikuta, S. and Itakura, K. (1984) Dynamics of DNA duplexes containing internal GT, GA, AC, and TC pairs – hydrogen-exchange at and adjacent to mismatch sites. Federation Proc. **43**, 2663–2670

5 Aboulela, F., Koh, D., Tinoco, I. and Martin, F. H. (1985) Base–base mismatches – thermodynamics of double helix formation for $dCA_3XA_3G+dCT_3YT_3G$ (X, Y=A,C,G,T). Nucleic Acids Res. **13**, 4811–4824

6 Hunter, W. N., Brown, T. and Kennard, O. (1986) Structural features and hydration of d(C-G-C-G-A-A-T-T-A-G-C-G) – a double helix containing two GA mispairs. J. Biomol. Struct. Dyn. **4**, 173–191

7 Brown, T. (1995) Oligonucleotide structure by X-ray crystallography and NMR spectroscopy. Aldrichimica Acta **28**, 15–20

8 Allawi, H. T. and SantaLucia, Jr, J. (1998) NMR solution structure of a DNA dodecamer containing single G.T mismatches. Nucleic Acids Res. **26**, 4925–4934

9 Allawi, H. T. and SantaLucia, Jr, J. (1997) Thermodynamics and NMR of internal GT mismatches in DNA. Biochemistry **36**, 10581–10594

10 Kneale, G., Brown, T., Kennard, O. and Rabinovich, D. (1985) G.T base-pairs in a DNA helix – the crystal-structure of d(G-G-G-G-T-C-C-C). J. Mol. Biol. **186**, 805–814

11 Hunter, W. N., Brown, T., Kneale, G., Anand, N. N., Rabinovich, D. and Kennard, O. (1987) The structure of guanosine-thymidine mismatches in B-DNA at 2.5 Å resolution. J. Biol. Chem. **262**, 9962–9970

12 Benevides, J. M., Wang, A. H.-J., Van der Marel, G. A., Van Boom, J. H. and Thomas, G. J. J. (1989) Raman spectral studies of nucleic-acids. 34. Effect of the G.T mismatch on backbone and sugar conformations of Z-DNA and B-DNA – analysis by Raman-spectroscopy of crystal and solution structures of d(CGCGTG) and d(CGCGCG). Biochemistry **28**, 304–310

13 Voigt, J. M. and Topal, M. D. (1990) *O*-6-methylguanine and A.C and G.T mismatches cause asymmetric structural defects in DNA that are affected by DNA sequence. Biochemistry **29**, 5012–5018

14 Allawi, H. T. and SantaLucia, Jr, J. (1998) Nearest-neighbor thermodynamics of internal A.C mismatches in DNA: sequence dependence and pH effects. Biochemistry **37**, 9435–9444

15 Bhaumik, S. R. and Chary, K. V. R. (1998) A^+.C mismatched base-pairing in DNA. Current Sci. **74**, 1008–1012

16 Kan, L. S., Chandrasegaran, S., Pulford, S. M. and Miller, P. S. (1983) Detection of a guanine.adenine base-pair in a decadeoxyribonucleotide by proton magnetic-resonance spectroscopy. Proc. Natl. Acad. Sci. U.S.A. **80**, 4263–4265

17 Brown, T., Hunter, W. N., Kneale, G. and Kennard, O. (1986) Molecular-structure of the G-A base-pair in DNA and its implications for the mechanism of transversion mutations. Proc. Natl. Acad. Sci. U.S.A. **83**, 2402–2406

18 Prive, G. G., Heinemann, U., Chandrasegaran, S., Kan, L. S., Kopka, M. L. and Dickerson, R. E. (1987) Helix geometry, hydration, and G.A mismatch in a B-DNA decamer. Science **238**, 498–504

19 Nikonowicz, E. P., Meadows, R. P., Fagan, P. and Gorenstein, D. G. (1991) NMR structural refinement of a tandem G.A mismatched decamer $d(CCAAGATTGG)_2$ via the hybrid matrix procedure. Biochemistry **30**, 1323–1334

20 Webster, G. D., Sanderson, M. R., Skelly, J. V., Neidle, S., Swann, P. F., Li, B. F. and Tickle, I. J. (1990) Crystal-structure and sequence-dependent conformation of the A.G mispaired oligonucleotide d(CGCAAGCTGGCG). Proc. Natl. Acad. Sci. U.S.A. **87**, 6693–6697

21 Carbonnaux, C., Van der Marel, G. A., Van Boom, J. H., Guschlbauer, W. and Fazakerley, G. V. (1991) Solution structure of an oncogenic DNA duplex containing a G.A mismatch. Biochemistry **30**, 5449–5458

22 Leonard, G. A., Booth, E. D. and Brown, T. (1990) Structural and thermodynamic studies on the adenine guanine mismatch in B-DNA. Nucleic Acids Res. **18**, 5617–5623

23 Lane, A. and Peck, B. (1995) Conformational flexibility in DNA duplexes containing single G.G mismatches. Eur. J. Biochem. **230**, 1073–1087

24 Ke, S. H. and Wartell, R. M. (1993) Influence of nearest-neighbor sequence on the stability of base-pair mismatches in long DNA – determination by temperature-gradient gel-electrophoresis. Nucleic Acids Res. **21**, 5137–5143

25 Brown, T. and Hunter, W. N. (1997) Non-Watson–Crick base associations in DNA and RNA revealed by single crystal X-ray diffraction methods: mismatches, modified bases, and nonduplex DNA. Biopolymers **44**, 91–103

26 Borden, K. L. B., Jenkins, T. C., Skelly, J. V., Brown, T. and Lane, A. N. (1992) Conformational properties of the G.G mismatch in $d(CGCGAATTGGCG)_2$ determined by NMR. Biochemistry **31**, 5411–5422

27 Peyret, N., Seneviratne, P. A., Allawi, H. T. and SantaLucia, Jr, J. (1999) Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A.A, C.C, G.G, and T.T mismatches. Biochemistry **38**, 3468–3477

28 Allawi, H. T. and SantaLucia, Jr, J. (1998) Thermodynamics of internal C.T mismatches in DNA. Nucleic Acids Res. **26**, 2694–2701

29 Li, Y., Zon, G. and Wilson, W. D. (1991) NMR and molecular modeling evidence for a G.A mismatch base-pair in a purine-rich DNA duplex. Proc. Natl. Acad. Sci. U.S.A. **88**, 26–30

30 Li, Y., Zon, G. and Wilson, W. D. (1991) Thermodynamics of DNA duplexes with adjacent G.A mismatches. Biochemistry **30**, 7566–7572

31 Nikonowicz, E. P., Meadows, R. P., Fagan, P. and Gorenstein, D. G. (1991) NMR structural refinement of a tandem G.A mismatched decamer $d(CCAAGATTGG)_2$ via the hybrid matrix procedure. Biochemistry **30**, 1323–1334

32 Cheng, J.-W., Chou, S.-H. and Reid, B. R. (1992) Base-pairing geometry in GA mismatches depends entirely on the neighboring sequence. J. Mol. Biol. **228**, 1037–1041

33 Chou, S.-H., Cheng, J.-W. and Reid, B. R. (1992) Solution structure of $[d(ATGAGCGAATA)]_2$ – adjacent G-A mismatches stabilized by cross-strand base-stacking and B(II) phosphate groups. J. Mol. Biol. **228**, 138–155

34 Ebel, S., Lane, A. and Brown, T. (1992) Very stable mismatch duplexes – structural and thermodynamic studies on tandem G.A mismatches in DNA. Biochemistry **31**, 12083–12086

35 Li, Y. and Agrawal, S. (1995) Oligonucleotides containing G.A pairs – effect of flanking sequences on structure and stability. Biochemistry **34**, 10056–10062

36 Ke, S.-H. and Wartell, R. M. (1996) The thermal stability of DNA fragments with tandem mismatches at a d(CXYG).d(CY′X′G) site. Nucleic Acids Res. **24**, 707–712
37 Lane, A., Martin, S. R., Ebel, S. and Brown, T. (1992) Solution conformation of a deoxynucleotide containing tandem G.A mismatched base-pairs and 3′-overhanging ends in $d(GTGAACTT)_2$. Biochemistry **31**, 12087–12095
38 Chou, S. H., Tseng, Y. Y., Chen, Y. R. and Cheng, J. W. (1999) Structural studies of symmetric DNA undecamers containing non-symmetrical sheared (PuGAPu):(PyGAPy) motifs. J. Biomol. NMR **14**, 157–167
39 Lane, A., Ebel, S. and Brown, T. (1994) Properties of multiple G.A mismatches in stable oligonucleotide duplexes. Eur. J. Biochem. **220**, 717–727
40 Drew, H. R. (1984) Structural specificities of five commonly used DNA nucleases. J. Mol. Biol. **176**, 535–557
41 Drew, H. R. and Travers, A. A. (1984) DNA structural variations in the *Escherichia coli* *Tyr*T-promoter. Cell (Cambridge, Mass.) **37**, 491–502
42 Herrera, J. E. and Chaires, J. B. (1994) Characterization of preferred deoxyribonuclease-I cleavage sites. J. Mol. Biol. **236**, 405–411
43 Suck, D., Lahm, A. and Oefner, C. (1988) Structure refined to 2 Å of a nicked DNA octanucleotide complex with DNase I. Nature (London) **332**, 464–468
44 Suck, D. and Oefner, C. (1986) Structure of DNase-I at 2.0 Å resolution suggests a mechanism for binding to and cutting DNA. Nature (London) **321**, 620–625
45 Flick, J. T., Eissenberg, J. C. and Elgin, S. C. R. (1986) Micrococcal nuclease as a DNA structural probe – its recognition sequences, their genomic distribution and correlation with DNA structure determinants. J. Mol. Biol. **190**, 619–633
46 Fox, K. R. and Waring, M. J. (1987) The use of micrococcal nuclease as a probe for drug-binding sites on DNA. Biochim. Biophys. Acta **909**, 145–155
47 Cotton, F. A., Hazen, E. E. and Legg, M. J. (1979) Staphylococcal nuclease: proposed mechanism of action based on structure of enzyme-thymidine 3′,5′-bisphosphate–calcium ion complex at 1.5 Å. Proc. Natl. Acad. Sci. U.S.A. **76**, 2551–2555
48 Loll, P. J. and Lattman, E. E. (1989) The crystal-structure of the ternary complex of staphylococcal nuclease, Ca^{2+} and the inhibitor pdTp, refined at 1.65 Å. Proteins Struct. Funct. Genet. **5**, 183–201
49 Loll, P. J., Quirk, S., Lattman, E. E. and Garavito, R. M. (1995) X-ray crystal-structures of staphylococcal nuclease complexed with the competitive inhibitor cobalt(II) and nucleotide. Biochemistry **34**, 4316–4324
50 Wang, J., Truckses, D. M., Abildgaard, F., Dzakula, Z., Zolnai, Z. and Markley, J. L. (1997) Solution structures of staphylococcal nuclease from multidimensional, multinuclear NMR: nuclease-H124L and its ternary complex with Ca^{2+} and thymidine-3′,5′-bisphosphate. J. Biomol. NMR **10**, 143–164
51 Tullius, T. D. (1987) Chemical snapshots of DNA – using the hydroxyl radical to study the structure of DNA and DNA–protein complexes. Trends Biochem. Sci. **12**, 297–300
52 Shafer, G. E., Price, M. A. and Tullius, T. D. (1989) Use of the hydroxyl radical and gel-electrophoresis to study DNA structure. Electrophoresis **10**, 397–404
53 Price, M. A. and Tullius, T. D. (1992) Using hydroxyl radical to probe DNA-structure. Methods Enzymol. **212**, 194–219
54 Price, M. A. and Tullius, T. D. (1993) How the structure of an adenine tract depends on sequence context – a new model for the structure of $T_n.A_n$ DNA sequences. Biochemistry **32**, 127–136
55 Hertzberg, R. P. and Dervan, P. B. (1982) Cleavage of DNA with methidiumpropyl-EDTA-iron(II): reaction conditions and product analyses. J. Am. Chem. Soc. **104**, 313–315
56 Hertzberg, R. P. and Dervan, P. B. (1984) Cleavage of DNA with methidiumpropyl EDTA-iron(II) – reaction conditions and product analyses. Biochemistry **23**, 3934–3945

Received 27 November 2002/6 January 2003; accepted 31 January 2003
Published as BJ Immediate Publication 31 January 2003, DOI 10.1042/BJ20021847