THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

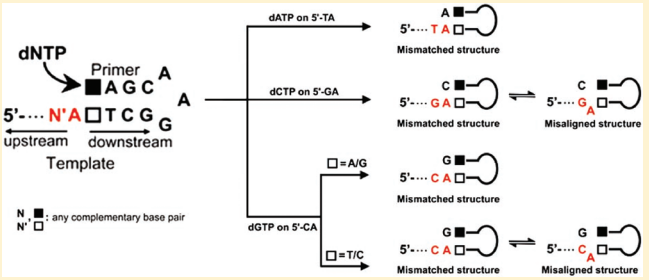
Sequence Context Effect on Strand Slippage in Natural DNA Primer–Templates

Lai Man Chi and Sik Lok Lam*

Department of Chemistry, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong

Supporting Information

ABSTRACT: Strand slippage has been found to occur in primer–templates containing a templating thymine, cytosine, and guanine, leading to the formation of misaligned structures with a single-nucleotide bulge. If remained in the active site of low-fidelity polymerases during DNA replication, these misaligned structures can ultimately bring about deletion mutations. In this study, we performed NMR investigations on primer–template models containing a templating adenine. Similar to our previous results on guanine, adenine templates are also less prone to strand slippage than pyrimidine templates. Misalignment occurs only in primer–templates that form a terminal C·G or G·C base pair. Together with our previous findings on thymine, cytosine, and guanine templates, the present study reveals strand slippage can occur in any kind of natural templating bases during DNA replication, providing insights into the origin of mutation hotspots in natural DNA sequences. In addition to the type of incoming base upon misincorporation, the propensity of strand slippage in primer–templates depends also on the type of templating base, its upstream and downstream bases.



1. INTRODUCTION

High-fidelity DNA replication is essential for maintaining an accurate transfer of genetic information that relies on a variety of factors and processes including selectivity of DNA polymerase, exonucleolytic proofreading function, and post-replication repair activity.[1] Yet low-fidelity DNA replication is also needed to facilitate translesion synthesis of otherwise replication-blocking lesions and allow mutations to fulfill the requirement for evolution of species.[2] Continuous efforts in DNA polymerase research have been improving our understanding of the replication fidelity and recent findings on the proofreading deficient Y-family DNA polymerases have provided new mechanistic pathways for evolution.[3−5] These polymerases have no detectable sequence similarity with other family polymerases.[6,7] Interestingly, they have an open and spacious active site that allows the accommodation of bulky adducts, wobble, or Hoogsteen base pairs,[8−11] suggesting the replicating site structure of primer–templates may also influence the fidelity of DNA replication. In this regard, we performed NMR investigations on primer–template models containing a templating adenine in this study and showed slipped frameshift intermediates can occur in adenine templates with a mismatched cytosine or guanine. Together with our previous findings on thymine,[12,13] cytosine,[14] and guanine[15] templates, our results reveal all four kinds of natural DNA bases possess different levels of strand slippage propensity in primer–templates, providing insights into the origin of hotspots in genetic mutations.

The low-fidelity Y-family DNA polymerases are believed to aid in continuing translesion DNA synthesis by efficiently bypassing damaged templates of bulky DNA adducts and/or replicating undamaged DNA.[5,16] Crystallographic results have shown the presence of undamaged misaligned DNA templates embedded in active sites of these polymerases.[17−19] Such misaligned intermediates have been suggested to form in the absence of a polymerase.[20,21] They can be accepted for extension on polymerase binding,[22,23] supporting that the formation and accommodation of a misaligned primer–template in the loose active site can ultimately lead to deletion mutation.[24−26] Despite different mechanisms proposed for strand slippage,[25,27−31] mutation rate in vivo varies widely as it depends on the DNA sequence, its surrounding context, and orientation.[1,32] Nevertheless, the sequence context effect of natural DNA template remains not well understood, especially under what situations it will lead to strand slippage in primer–templates.

Although various types of DNA polymerases are likely to interact with primer–template differently, it is our primary interest to understand from the DNA substrate point of view how sequence context affects the replicating site structures of primer–templates. Previously, we performed NMR structural investigations on primer–template oligonucleotide models that mimic the situation in which a deoxyribonucleotide triphos-

Received: December 5, 2011
Revised: January 18, 2012
Published: February 3, 2012

dx.doi.org/10.1021/jp211666k | J. Phys. Chem. B 2012, 116, 1999–2007

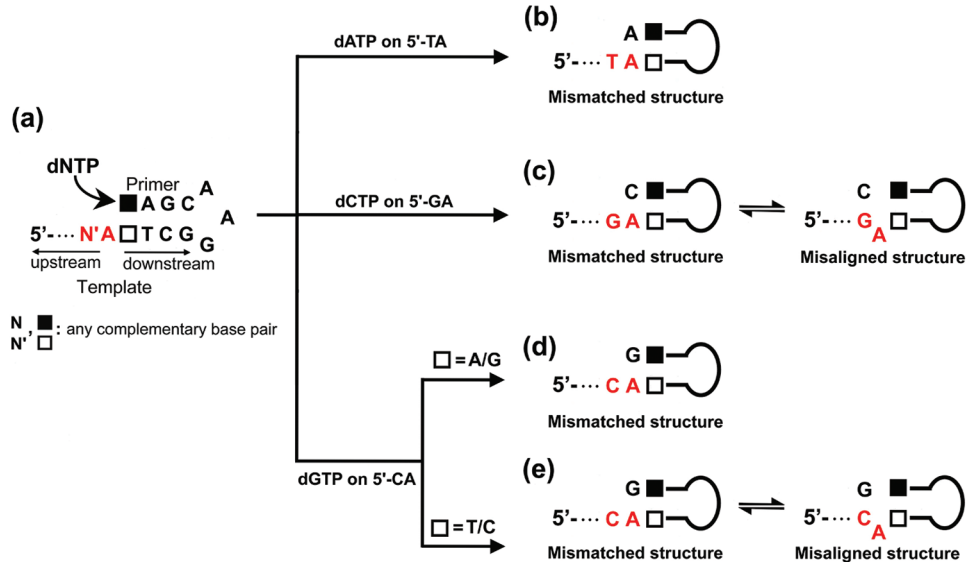


Figure 1. DNA primer–template models used in this study. (a) All models were designed to form a hairpin with a 5′-GAA loop. The top and bottom strands mimic the primer and template, respectively. The upstream and downstream directions were defined from the perspective of the template strand. (b) For 5′-TA templates, only the mismatched conformer with a terminal A·A was observed. (c) For 5′-GA templates, both the mismatched conformer with a terminal C·A mispair and the misaligned conformer with an A-bulge were observed. (d) For 5′-CA templates, only the mismatched conformer with a terminal G·A mispair was observed when the base downstream of the templating adenine was a purine. (e) When the downstream base became a pyrimidine, the misaligned conformer was also present.

phate (dNTP) has just been incorporated opposite a thymine,[12,13] cytosine,[14] and guanine[15] template, revealing that guanine template is less prone to strand slippage than pyrimidine templates. Strand slippage occurs in pyrimidine templates regardless of the type of resulting terminal base pair in the misaligned primer–templates but a terminal G·C Watson–Crick base pair is necessary for guanine templates. The significance of the base upstream and downstream of the templating base has also been demonstrated. In order to obtain a thorough understanding of the sequence context effect on strand slippage, we performed NMR investigations on primer–template models containing 5′-TA, 5′-GA, and 5′-CA templates in this study (Figure 1). Conclusively, the propensity of strand slippage in primer–templates containing natural DNA sequences has been found to depend on the type of (i) the templating base, (ii) the base upstream of the templating base, (iii) the base downstream of the templating base, and (iv) the incoming base.

2. EXPERIMENTAL METHODS

2.1. Sample Design. All DNA samples were designed to form a hairpin with the top and bottom strands mimicking the primer and template, respectively (Figure 1a). The 5′-GAA hairpin loop connects the primer and template strands in order to simplify the sample preparative work.[33] These samples mimic the situation in which a dNTP has just been incorporated opposite the templates. The 5′-overhang region of the samples represents the template sequence, whereas the 3′-terminal nucleotide represents the newly incorporated dNTP at the primer terminus.

2.2. Sample Preparation. All DNA samples were synthesized using an Applied Biosystems model 394 DNA synthesizer and purified using denaturing polyacrylamide gel electrophoresis and diethylaminoethyl Sephacel anion exchange column chromatography. NMR samples were prepared by dissolving 0.5 μmol of purified DNA samples into 500 μL of buffer solution containing 150 mM sodium chloride, 10 mM sodium phosphate (pH 7.0), and 0.1 mM 2,2-dimethyl-2-silapentane-5-sulfonic acid.

2.3. NMR Analysis. All NMR experiments were performed using Bruker ARX-500 and AV-500 spectrometers operating at 500.13 and 500.30 MHz, respectively. For studying labile proton resonance signals, the samples were prepared in a 90% H_2O/10% D_2O buffer solution. One-dimensional (1D) imino proton spectra were acquired using the water suppression by gradient-tailored excitation (WATERGATE) pulse sequence.[34,35] Two-dimensional (2D) WATERGATE-nuclear Overhauser effect spectroscopy (NOESY) experiments were performed with a mixing time of 300 ms. For studying nonlabile proton signals, the solvent was exchanged with a 100% D_2O solution. For the D_2O samples, the 2D NOESY experiments were also performed with a mixing time of 300 ms. Data sets of 4096 × 512 were collected for all 2D NOESY experiments. The acquired data were zero-filled to give 4096 × 4096 spectra with a cosine window function applied to both dimensions. In general, these experiments were conducted at 25 °C unless stated otherwise. In order to better observe and resolve the labile and nonlabile signals, some of these experiments were repeated at lower temperatures.

3. RESULTS

In this study, we performed high-resolution NMR spectroscopic investigations focusing on 1D imino proton and 2D nuclear Overhauser effect (NOE) analyses to probe the replicating site structures of 5′-TA, 5′-GA, and 5′-CA templates after the incorporation of dATP, dCTP, and dGTP, respectively (Figure 1). Sequential assignments were made from the fingerprint regions in the 2D NOESY spectra in H_2O and/or D_2O using standard methods[36] (see also Supporting Information, S1–S14). Unlike some guanine templates,[15] all adenine template samples adopted the designed hairpin conformation instead of forming homoduplexes as revealed

dx.doi.org/10.1021/jp211666k | J. Phys. Chem. B 2012, 116, 1999–2007

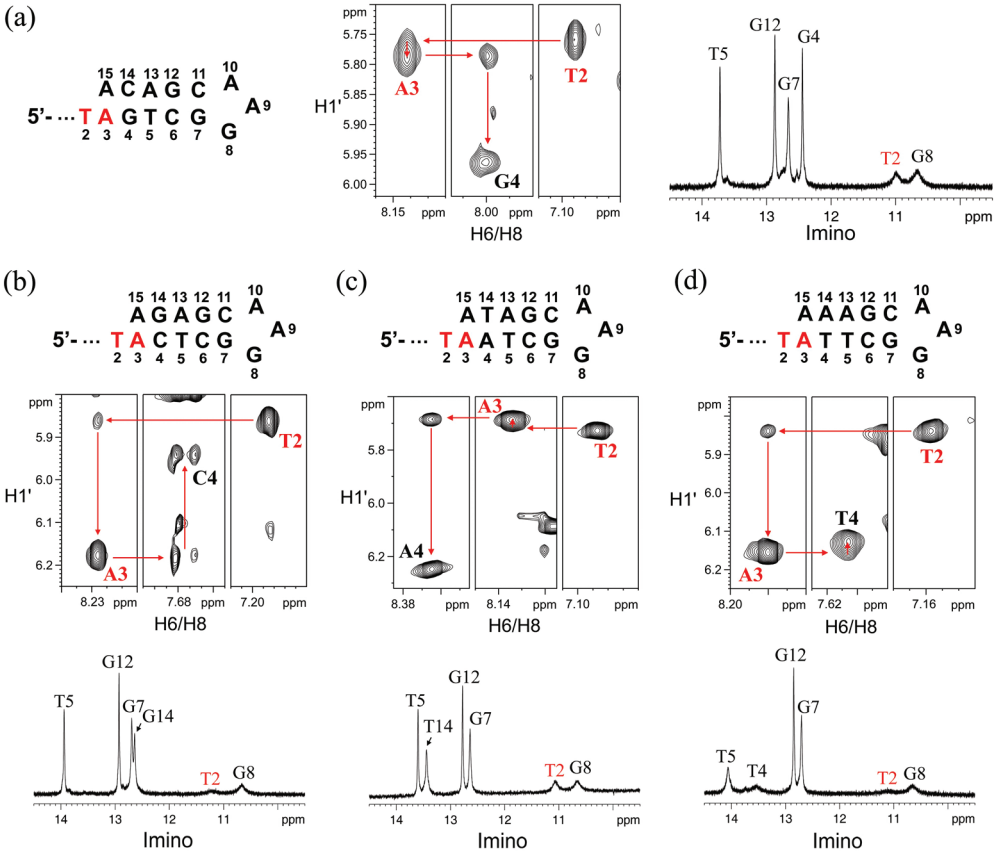


Figure 2. NMR evidence for the mismatched conformers in 5′-TA templates. (a) Mismatched conformer was formed after the incorporation of a dATP opposite a 5′-TA template. At 25 °C, NOESY H6/H8−H1′ fingerprint region shows the sequential T2−A3 and A3−G4 NOEs. At 5 °C, the small imino signal at ∼11.0 ppm was assigned to the overhang T2 that was stacked on the terminal A·A mispair. The small imino signal at ∼10.7 ppm corresponds to the G8 imino signal of sheared G8·A10 mispair in the GAA loop. Similarly, the same NOESY and imino regions support the formation of mismatched conformers in 5′-TA templates upon substituting C14·G4 with (b) G14·C4, (c) T14·A4, and (d) A14·T4 base pairs.

by the characteristic chemical shifts of the GAA triloop[33] and the absence of broadened thymine methyl peaks.

3.1. Incorporation of a dATP Opposite a 5′-TA Template. When dATP was incorporated opposite a 5′-TA template, a terminal A·A mismatch was formed and no misalignment was observed (Figure 1b, see also Supporting Information, S15). This was supported by the sequential NOEs between T2 H1′−A3 H8 and A3 H1′−G4 H8 (Figure 2a), together with the absence of unusual T2 H1′−G4 H8 NOE and T2 imino signal that could come from a potential T2·A15 Watson−Crick base pair. The small imino signal at ∼11.0 ppm was assigned to the overhang T2 which stacked on the terminal A·A mispair, as supported by its chemical shift which is similar to that of a reference sequence containing a terminal T·A Watson−Crick base pair (Supporting Information, S16). An intense and sharp G4 imino signal was observed (Figure 2a), suggesting the terminal A·A mismatch has little effect on the penultimate C14·G4 base pair.

To investigate if the base downstream of 5′-TA template affects the alignment of primer-template, we substituted C14·G4 base pair with G14·C4, T14·A4, and A14·T4 base pairs, respectively. These substitutions were evidenced by the presence of G14 (Figure 2b), T14 (Figure 2c), and T4 (Figure 2d) Watson−Crick imino signals. In all three cases, a mismatched conformer with a terminal A·A mispair was observed as revealed by the sequential NOEs between the second to third and third to fourth nucleotides and the presence of overhang T2 imino signals at ∼11.0 ppm. The absence of unusual NOEs between the second and fourth nucleotides and T2 Watson−Crick imino signals also support the mismatched conformers. In our previous studies on thymine,[12,13] cytosine,[14] and guanine[15] templates, realignment leading to the formation of misaligned structures with a single-nucleotide bulge was observed when the templating base and its downstream base were the same. However, no realignment was observed when there were two successive adenines in the template (Figure 2c). When the downstream base pair was A14·T4 (Figure 2d), which resulted in two successive thymines in the template and three successive adenines in the primer, the template end became more flexible, as reflected by the broadened T4, and weakened T5 and overhang T2 imino signals. No evidence for other conformers has been observed and the higher flexibility might be due to the lower stacking energy of A14·T4 on A13·T5 when compared to that of T14·A4 on A13·T5.[37]

3.2. Incorporation of a dCTP Opposite a 5′-GA Template. When dCTP was incorporated opposite a 5′-GA template, both misaligned conformers containing an A-bulge and mismatched conformers with a terminal C·A mispair were observed (Figure 1c, see also Supporting Information, S17). The presence of a misaligned conformer was evidenced by the unusual G2 H1′−G4 H8 NOE while the sequential G2 H1′−A3 H8 and A3 H1′−G4 H8 NOEs support the presence of a mismatched conformer (Figure 3a). Owing to the chemical

dx.doi.org/10.1021/jp211666k | J. Phys. Chem. B 2012, 116, 1999−2007

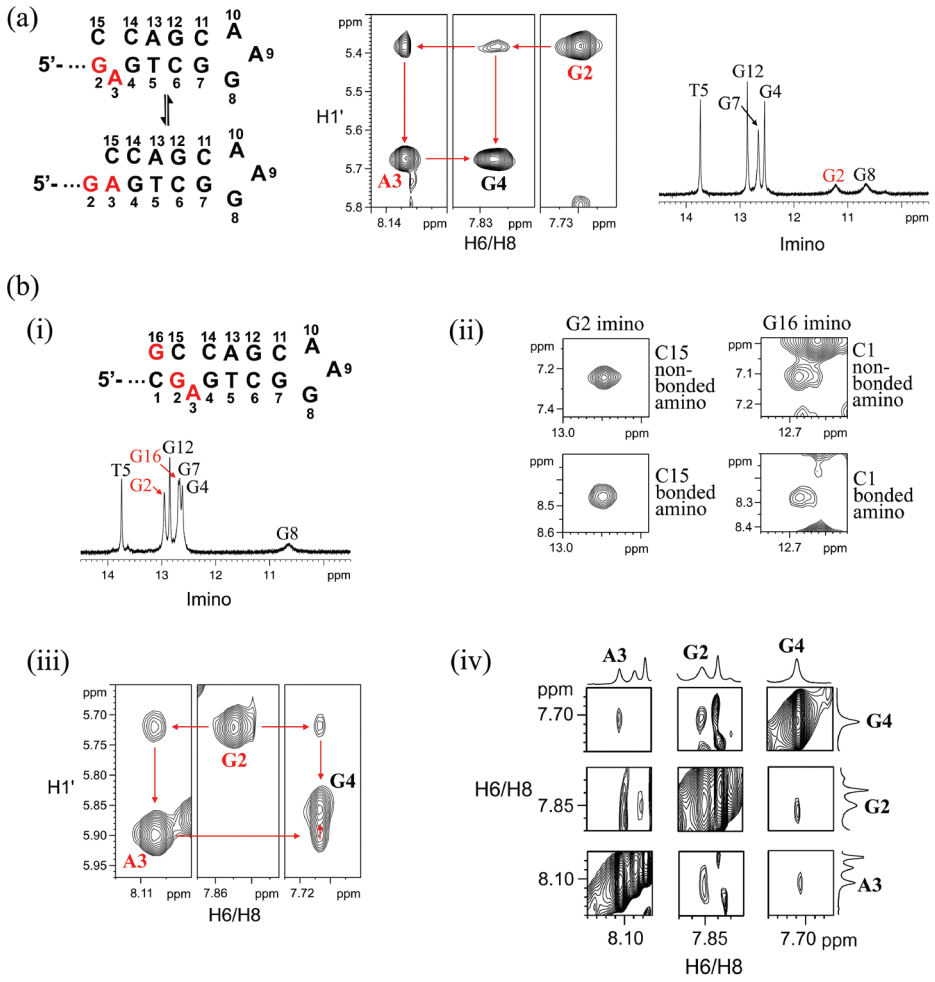


Figure 3. (a) Both the misaligned and mismatched conformers were present in the 5′-GA template with an opposite C. At 25 °C, the unusual G2−G4 NOE supports the misaligned conformer whereas the sequential G2−A3 and A3−G4 NOEs support the mismatched conformer. At 5 °C, the imino signal at ∼11.2 ppm corresponds to the averaged G2 of the two conformers. (b) After the addition of a terminal G·C base pair, a stabilized misaligned conformer containing an intrahelical A-bulge was observed. The presence of (i) G2 and G16 imino signals and (ii) NOEs between G2 imino−C15 amino and G16 imino−C1 amino protons at 5 °C indicates the formation of G2·C15 and G16·C1 Watson−Crick base pairs. (iii) At 25 °C, the G2−G4 NOE supports the A-bulged misaligned conformer, and the sequential G2−A3 and A3−G4 NOEs, together with the chemical shift of A3 H8, and (iv) the base−base NOEs between G2−A3 and A3−G4 support the A-bulge being intrahelical.

shift of G imino in G·C or C·G Watson−Crick base pair being ∼12.5 to 13.0 ppm[38] and that as an overhang being ∼10.3 ppm, the small imino signal at ∼11.2 ppm was assigned as the averaged G2 signal of the two conformers at 5 °C (Supporting Information, S18). In our previous studies on primer−templates containing a templating guanine, we had found both the unusual and sequential NOEs in a stable G-bulged misaligned conformer.[15] In order to verify the above assignment, we stabilized the A-bulged misaligned conformer by adding a terminal G·C base pair. The presence of G2 and G16 imino signals (Figure 3b, i), and G2 imino−C15 amino and G16 imino−C1 amino NOEs (Figure 3b, ii) indicates the formation of G2·C15 and G16·C1 base pairs, which hampered the formation of mismatched conformer and caused the G2 imino signal to shift from ∼11.2 ppm to the downfield C·G Watson−Crick imino region. This downfield shift of G2 imino signal supports the assignment and the presence of conformational exchange between the mismatched and misaligned conformers in the 5′-GA template. In this stable misaligned A-bulged conformer, apart from the characteristic G2−G4 NOE (Figure 3b, iii), the sequential NOEs between G2−A3 and A3−G4 were also observed, suggesting A3 remains partially stacked between the G2·C15 and G4·C14 base pairs. This partially stacked intrahelical A-bulge was also supported by the presence of base−base NOEs between G2 H8−A3 H8, A3 H8−G4 H8, and G2 H8−G4 H8 (Figure 3b, iv) and the chemical shift of A3 H8. In the absence of ring-current effect, the chemical shift of this proton has been found to be ∼8.6 ppm.[39] The current value of ∼8.1 ppm indicates the presence of some ring-current effect which comes from the partial stacking of A3.

Similarly, we substituted C14·G4 base pair with A14·T4, T14·A4, and G14·C4 base pairs, respectively, to investigate the effect of base downstream of 5′-GA template. These substitutions were evidenced by the T4 (Figure 4a), T14 (Figure 4b), and G14 (Figure 4c) Watson−Crick imino signals. In all three cases, both the misaligned and mismatched conformers were observed. The presence of misaligned conformers was evidenced by the unusual NOEs between the second and fourth nucleotides. While the sequential NOEs from the second to third and third to fourth nucleotides alone may not suffice for the presence of mismatched conformers, the

dx.doi.org/10.1021/jp211666k | J. Phys. Chem. B 2012, 116, 1999−2007

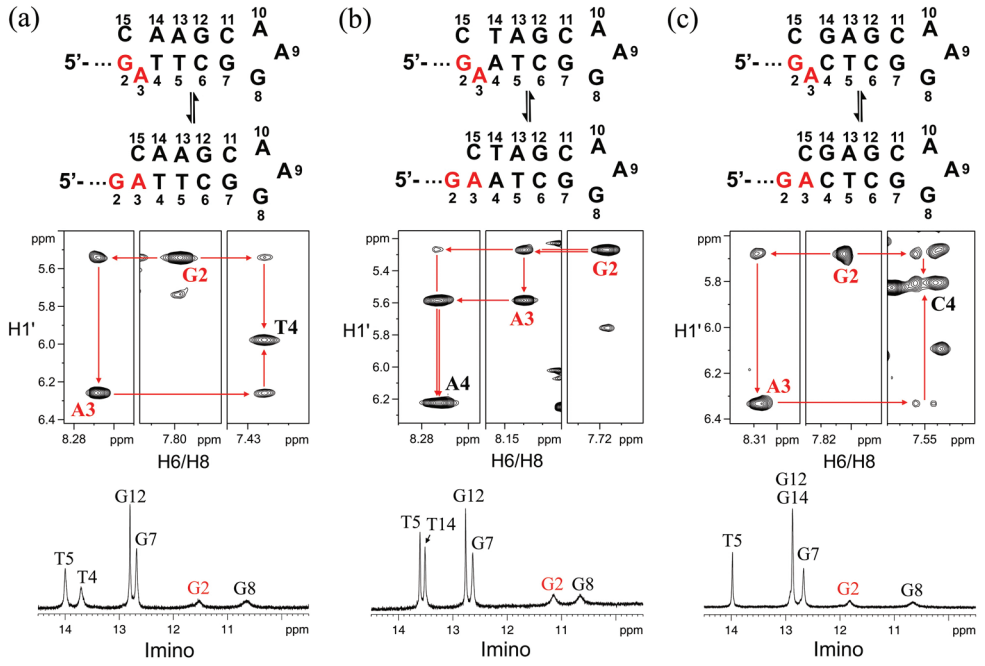


Figure 4. As evidenced by the NOE and imino signals, both the misaligned and mismatched conformers were present in 5′-GA templates upon substituting C14·G4 with (a) A14·T4, (b) T14·A4, and (c) G14·C4 base pairs.

small broad G2 imino signals in the range of 11−12 ppm suggest they came from the averaged signals between the exchanging misaligned and mismatched conformers. According to the G2 imino proton chemical shift, the relative abundance of misaligned conformer was highest when C14·G4 was substituted with G14·C4 (Figure 4c), followed by A14·T4 (Figure 4a), and then T14·A4 (Figure 4b). This order may be because the relative stabilization is highest when the terminal C15·G2 formed in the misaligned conformer stacks on G14·C4, followed by A14·T4 and then T14·A4.[37]

3.3. Incorporation of a dGTP Opposite a 5′-CA Template. When dGTP was incorporated opposite a 5′-CA template, no misalignment was observed when the base downstream of the templating A was a purine (Figure 1d). When the downstream base was a guanine, only a mismatched conformer containing a terminal G·A mispair was observed as evidenced by the sequential C2−A3 and A3−G4 NOEs (Figure 5a). No unusual NOE was observed between C2 and G4, suggesting no misalignment and further supporting the mismatched conformer. The G15 imino proton of the terminal G·A mispair was found to appear at ∼11.2 ppm at 5 °C. Similar NOE and imino results were also obtained when the downstream base became an adenine (Figure 5b). The small difference in the G15 imino chemical shift is probably due to its involvement in the hydrogen bonds of two different G·A pairing modes that are sequence context dependent.[40−42]

When the base downstream of the templating A was a thymine (Figure 5c) or cytosine (Figure 5d), both the misaligned and mismatched conformers were present. While the sequential second to third and third to fourth NOEs support the mismatched conformer, the unusual NOEs between the second and fourth nucleotides support the misaligned conformer. In addition, the chemical shifts of G15 imino protons were found to be a bit more downfield than those in G·A mispairs but more upfield than those in Watson−Crick base pairs, suggesting the presence of conformational exchange between the misaligned and mismatched conformers. Such exchange process weakened the intensities of T4 (Figure 5c) and G14 (Figure 5d) imino protons, as favorable stacking interactions between the terminal mispair and the penultimate base pair were interrupted due to the formation of intrahelical A-bulge in the misaligned conformers.

4. DISCUSSION

4.1. Misalignment Occurs in Adenine Templates That Form a Terminal C·G or G·C Base Pair. In our previous studies on guanine templates, misalignment was observed only when the primer−templates resulted in a terminal G·C base pair.[15] In this work, our findings on adenine templates were similar to those of guanine templates. Misalignment was only observed when the primer−templates resulted in a terminal C·G or G·C base pair but not an A·T or T·A base pair. Specifically, misalignment occurs in (i) 5′-GA template with an opposite C (Figure 1c) and (ii) 5′-CA template with an opposite G and a downstream pyrimidine (Figure 1e). In both cases, the misaligned conformers were found in exchange with the mismatched conformers. For the former case, the formation of the misaligned conformer may be due to the instability of the terminal C·A mispair in the mismatched conformer,[43] leading to a decrease in the population of the mismatched conformer and an increase in the population of the misaligned conformer. For the latter case, the misaligned conformer may result from the better stacking interactions of the two successive purines in the primer strand after slippage,[44] thus promoting the formation of the misaligned conformer.

4.2. Strand Slippage Is Less Common in Purine Templates. In pyrimidine templates, strand slippage leading to the formation of misaligned conformers has been found to occur regardless of the type of base pair in the resulting primer−templates.[12−14] For guanine[15] and adenine templates, strand slippage is less common because misalignment can only occur when the primer−templates result in a G·C or C·G terminal base pair. Owing to the fact that the bulged purine is more destabilizing than bulged pyrimidine,[45] an A·T or T·A

dx.doi.org/10.1021/jp211666k | J. Phys. Chem. B 2012, 116, 1999−2007

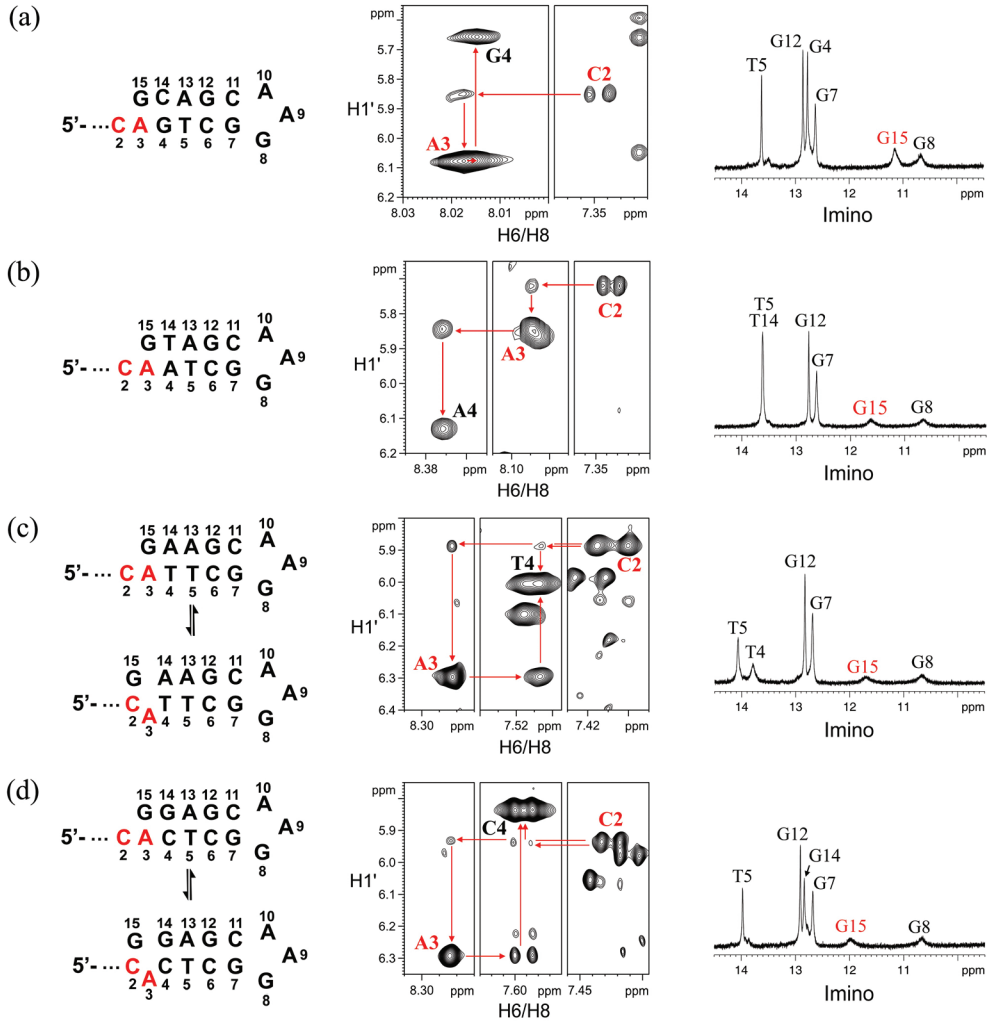


Figure 5. For 5′-CA templates with a downstream (a) G, or (b) A, only the mismatched conformer was formed after the incorporation of a dGTP. At 25 °C, the sequential second to third and third to fourth NOEs support the mismatched conformer. At 5 °C, the small imino signal in ∼11−12 ppm was assigned to G15. For 5′-CA templates with a downstream (c) T or (d) C, both the mismatched and misaligned conformers were formed as revealed by the unusual NOEs between the second and fourth nucleotides. The G15 imino signals were more downfield due to signal averaging of G15·A3 mispair and G15·C2 Watson−Crick base pair in the mismatched and misaligned conformers, respectively.

terminal base pair does not provide sufficient stabilizing interactions to maintain the misaligned conformers for purine templates. As a result, in addition to the type of terminal base pair, the type of templating base also determines the propensity of strand slippage.

4.3. Purine Bulge Partially Stacked within Helix. In pyrimidine templates, the misaligned conformers showed unusual NOEs between the second and fourth nucleotides, that is, the nucleotides flanking the single-nucleotide pyrimidine bulge.[12−14] No sequential second to third or third to fourth NOE was observed in these misaligned conformers. However, in both guanine[15] and adenine templates, owing to the single-nucleotide purine-bulge remains partially stacked within the double helices, these sequential NOEs were present together with the unusual NOEs between the second and fourth nucleotides in the misaligned conformer. These partially stacked purine-bulge structures had also been observed previously by NMR[46] and X-ray crystallography.[17] The unpaired purine-bulge remains intrahelical as the stacking stabilization of the planar bicyclic ring of purine with adjacent base pairs is better than that of the monocyclic ring of pyrimidine.[47] Compared to the extrahelical pyrimidine-bulge,[12−14] the intrahelical purine-bulge affects the stacking stabilization between its flanking Watson−Crick base pairs in the misaligned primer−templates, providing a reason for the lower strand slippage propensity in purine templates.

4.4. Misalignment Occurs in All Natural DNA Sequences. The results of the present work, together with the results on thymine,[12,13] cytosine,[14] and guanine[15] templates provide a clear picture that strand slippage can occur in primer−templates containing any type of natural templating bases. Figure 6 shows a summary of the situations that bring about misalignment of primer−templates in these studies. In guanine templates, misalignment only occurs when there is an upstream cytosine and a downstream purine (Figure 6a). For a templating adenine, misalignment can occur if there is an upstream guanine, or cytosine together with a downstream pyrimidine (Figure 6b). In these two templating purine cases, the misaligned primer−templates are in exchange equilibrium with the mismatched primer−templates. For a templating thymine (Figure 6c) or cytosine (Figure 6d), misalignment can occur whenever the upstream base is different from the templating base. Strand realignment leading to an exchange equilibrium between the misaligned and mismatched primer−

dx.doi.org/10.1021/jp211666k | J. Phys. Chem. B 2012, 116, 1999−2007

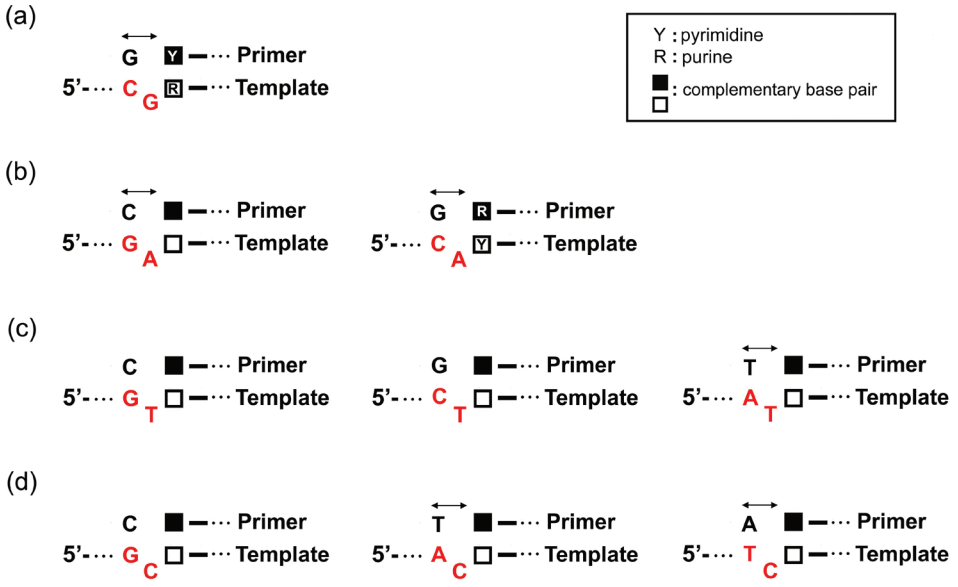


Figure 6. A summary of situations that misalignment can occur in primer–templates containing a templating (a) guanine, (b) adenine, (c) thymine, and (d) cytosine. Black arrows indicate realignment can occur, leading to the formation of mismatched conformers.

templates only occurs if misalignment results in a terminal A·T or T·A base pair. Although it is possible that misalignment of primer–templates may be induced by other molecules that are in contact with the primer–templates during DNA replication, for example, Y-family polymerases,[4,48−51] our overall findings reveal that different natural DNA sequences possess different levels of strand slippage propensity, providing an intrinsic factor for the presence of mutational hotspots in certain genes and the occurrence of natural mutations in evolution, fitness, and immunological diversity.[2]

When low-fidelity polymerases are involved in further processing the misaligned primer–templates during DNA replication, this may bring about −1 frameshifts and/or base substitutions upon realignment of the primer–templates. Yet the major replicative polymerases are high-fidelity in nature and they almost always initially insert the correct dNTP onto properly aligned primer–templates.[52] In addition to the hydrogen bonding between templating bases and incoming dNTPs[53] and the enthalpy–entropy compensation through stripping away the hydrogen-bonded water molecules of the incoming dNTP,[54,55] the exquisite shape complementary of the nascent base pair binding pockets of high-fidelity polymerases[1,53,54,56] also contribute to the high nucleotide selectivity. Several replicative polymerases also contain an associated 3′→5′ exonuclease activity that can excise incorrect bases from the growing DNA chain and allow another attempt at correct synthesis, avoiding errors to occur during DNA replication.[52]

4.5. Factors Governing Strand Slippage in Natural DNA Sequences. On the basis of the summary in Figure 6, the following factors that affect the propensity of strand slippage in primer–templates containing natural DNA sequences have been identified: (i) the type of templating base, (ii) the type of base upstream of the templating base, (iii) the type of base downstream of the templating base, and (iv) the type of incoming base. For the type of templating base, strand slippage has been found to be more common for a templating pyrimidine than a templating purine. As the single-nucleotide purine-bulge formed in the misaligned conformer of primer–templates remains intrahelical, this affects the stacking stabilization between its flanking Watson–Crick base pairs and thereby strand slippage is less common in purine templates. In a previous sequence context study using the data on mutations in coding regions of 19 human loci that cause mendelian diseases, 7 −1 frameshift hotspots were identified.[57] Among these, six involved pyrimidine-nucleotide deletions and one involved purine, agreeing with the propensity of strand slippage in primer–templates concluded in this work.

Regarding the type of base upstream of the templating base, it governs the type of terminal Watson–Crick base pair formed upon misalignment. If the upstream base is a cytosine or guanine, strand slippage is more likely because terminal G·C or C·G base pair will be resulted upon misalignment. These base pairs provide better stabilization for the misaligned conformer than A·T and T·A base pairs. Concerning the type of base downstream of the templating base, it determines the type of Watson–Crick base pair formed next to the terminal base pair or mispair formed in the misaligned or mismatched conformers. This affects the base stacking interactions and thus the relative stabilities of the misaligned and the mismatched conformers. For pyrimidine templates, the bulge formed in misaligned conformers is extrahelical and the effect from the downstream base is negligible. For purine templates, the bulge formed in misaligned conformers remains partially stacked within the helix[15] and thus the downstream base also plays a role in determining slippage. In a mutagenesis study within the lacZα gene present in M13mp2 DNA, −1 frameshifts, and base substitutions were observed at two mutational hotspots in the mutation spectrum produced by pol β.[25] These errors were suggested to result from transient misalignment of primer–templates. Our results showed that these proposed misalignments were possible as they were all observed in our models with the same type of bases downstream and upstream of the templating bases, revealing the significance of local sequence context in DNA replication.

The above three factors are intrinsic in nature as they depend on the sequence context of templates. In contrast, the type of incoming base is an occasional event during DNA replication.

dx.doi.org/10.1021/jp211666k | J. Phys. Chem. B 2012, 116, 1999–2007

The type of incoming base determines the type of terminal mispair formed upon misincorporation. This affects the relative stabilities of the mismatched and misaligned conformers. Therefore, if the stability of the terminal mispair in the mismatched conformer is relatively low,[43] for example, C·C in cytosine templates,[14] it is more likely that strand slippage will occur. However, if the stability of the mispair is relatively high, for example, A·G in guanine templates,[15] it is less likely that strand slippage will occur.

4.6. Insights into the Origin of Deletion Mutation. The present work complements with the results of our previous studies,[12−14] providing a thorough picture of the intrinsic structural features possessed by different primer−templates and revealing that different natural DNA sequences possess different levels of strand slippage propensity. From the thermodynamic point of view, we conclude that the template sequence intrinsically predetermines the likeness of strand slippage. As the spacious active site of low-fidelity polymerase is capable of accommodating a misaligned DNA template[10,17,18,23] and primer extension can occur at the template terminus with a single-nucleotide bulge by low-fidelity polymerase,[23] it is possible that errors resulting from low-fidelity DNA replication are related to the formation propensity of misaligned structures in different templates. Alternatively, it is also possible that the origin of deletion mutation may result from the relative contribution from the kinetic barrier toward the transition state of intermediate upon misincorporation. As individual DNA polymerases are likely to interact with primer−templates differently, the extent to which these misaligned structures occur in DNA polymerases may also depend upon the specific polymerase in question. Further investigation is needed to determine whether misalignment occurs at time scale faster than the rate at which a second nucleotide is incorporated so as to understand the mechanistic pathways for the occurrence of deletion mutation during low-fidelity DNA replication.

5. CONCLUSIONS

Similar to guanine templates,[15] adenine templates are also less prone to strand slippage in primer−templates. Misalignment only occurs when a terminal C·G or G·C base pair is formed in the primer−templates. Together with our previous results on pyrimidine templates,[12−14] we conclude that strand slippage can occur in any kind of natural templating base. Strand slippage occurs more easily in pyrimidine templates than purine. In addition to the type of templating base, the type of base both upstream and downstream of the templating base also plays a significant role in governing strand slippage. Our results in strand slippage propensity of primer−templates coincide with the mutations observed in 19 human loci that cause mendelian diseases[57] and in the lacZα gene present in M13mp2 DNA.[25] The template sequence intrinsically predetermines the likeness of strand slippage, providing insights into the origin of deletion mutation hotspots in natural DNA sequences.

■ ASSOCIATED CONTENT

Supporting Information

Figures showing the sequential assignment, NOESY, imino proton spectra and structural models of primer−templates. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: (852) 3943-8126. Fax: (852) 2603-5057. E-mail: lams@cuhk.edu.hk.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We would like to thank Professors H. K. Lee, H. N. C. Wong, and J. C. Yu for their continuing support on the research activities of our group. The work described in this paper was supported by a General Research Fund (Project No. CUHK401107) and Special Equipment Grant (Project No. SEG/CUHK09) from the Research Grants Council of the Hong Kong Special Administrative Region, and a direct grant (Project No. 2060377) from Research Committee of The Chinese University of Hong Kong.

■ REFERENCES

(1) Kunkel, T. A.; Bebenek, K. *Annu. Rev. Biochem.* **2000**, *69*, 497−529.
(2) Tippin, B.; Pham, P.; Goodman, M. F. *Trends Microbiol.* **2004**, *12*, 288−295.
(3) Ohmori, H.; Friedberg, E. C.; Fuchs, R. P.; Goodman, M. F.; Hanaoka, F.; Hinkle, D.; Kunkel, T. A.; Lawrence, C. W.; Livneh, Z.; Nohmi, T.; Prakash, L.; Prakash, S.; Todo, T.; Walker, G. C.; Wang, Z.; Woodgate, R. *Mol. Cell* **2001**, *8*, 7−8.
(4) Goodman, M. F. *Annu. Rev. Biochem.* **2002**, *71*, 17−50.
(5) Prakash, S.; Johnson, R. E.; Prakash, L. *Annu. Rev. Biochem.* **2005**, *74*, 317−353.
(6) Goodman, M. F.; Tippin, B. *Nat. Rev. Mol. Cell Biol.* **2000**, *1*, 101−109.
(7) Yang, W. *Curr. Opin. Struct. Biol.* **2003**, *13*, 23−30.
(8) Perlow-Poehnelt, R. A.; Likhterov, I.; Scicchitano, D. A.; Geacintov, N. E.; Broyde, S. *J. Biol. Chem.* **2004**, *279*, 36951−36961.
(9) Eoff, R. L.; Irimia, A.; Egli, M.; Guengerich, F. P. *J. Biol. Chem.* **2007**, *282*, 1456−1467.
(10) Nair, D. T.; Johnson, R. E.; Prakash, L.; Prakash, S.; Aggarwal, A. K. *Structure* **2005**, *13*, 1569−1577.
(11) Yang, W. *FEBS Lett.* **2005**, *579*, 868−872.
(12) Chi, L. M.; Lam, S. L. *FEBS Lett.* **2006**, *580*, 6496−6500.
(13) Chi, L. M.; Lam, S. L. *Biochemistry* **2007**, *46*, 9292−9300.
(14) Chi, L. M.; Lam, S. L. *Biochemistry* **2008**, *47*, 4469−4476.
(15) Chi, L. M.; Lam, S. L. *Biochemistry* **2009**, *48*, 11478−11486.
(16) Kunkel, T. A.; Pavlov, Y. I.; Bebenek, K. *DNA Repair* **2003**, *2*, 135−149.
(17) Ling, H.; Boudsocq, F.; Woodgate, R.; Yang, W. *Cell* **2001**, *107*, 91−102.
(18) Garcia-Diaz, M.; Bebenek, K.; Krahn, J. M.; Pedersen, L. C.; Kunkel, T. A. *Cell* **2006**, *124*, 331−342.
(19) Wilson, R. C.; Pata, J. D. *Mol. Cell* **2008**, *29*, 767−779.
(20) Bebenek, K.; Kunkel, T. A. *Cold Spring Harbor Symp. Quant. Biol.* **2000**, *65*, 81−91.
(21) Garcia-Diaz, M.; Kunkel, T. A. *Trends Biochem. Sci.* **2006**, *31*, 206−214.
(22) Wolfle, W. T.; Washington, M. T.; Prakash, L.; Prakash, S. *Genes Dev.* **2003**, *17*, 2191−2199.
(23) Cannistraro, V. J.; Taylor, J. S. *J. Biol. Chem.* **2007**, *282*, 11188−11196.
(24) Kunkel, T. A. *Biochemistry* **1990**, *29*, 8003−8011.
(25) Kunkel, T. A.; Soni, A. *J. Biol. Chem.* **1988**, *263*, 14784−14789.
(26) Fowler, R. G.; Degnen, G. E.; Cox, E. C. *Mol. Gen. Genet.* **1974**, *133*, 179−191.
(27) Streisinger, G.; Okada, Y.; Emrich, J.; Newton, J.; Tsugita, A.; Terzaghi, E.; Inouye, M. *Cold Spring Harbor Symp. Quant. Biol.* **1966**, *31*, 77−84.

(28) Fujii, S.; Akiyama, M.; Aoki, K.; Sugaya, Y.; Higuchi, K.; Hiraoka, M.; Miki, Y.; Saitoh, N.; Yoshiyama, K.; Ihara, K.; Seki, M.; Ohtsubo, E.; Maki, H. *J. Mol. Biol.* **1999**, *289*, 835−850.
(29) Kunkel, T. A. *J. Biol. Chem.* **1986**, *261*, 13581−13587.
(30) Kokoska, R. J.; Bebenek, K.; Boudsocq, F.; Woodgate, R.; Kunkel, T. A. *J. Biol. Chem.* **2002**, *277*, 19633−19638.
(31) Efrati, E.; Tocco, G.; Eritja, R.; Wilson, S. H.; Goodman, M. F. *J. Biol. Chem.* **1997**, *272*, 2559−2569.
(32) Harfe, B. D.; Jinks-Robertson, S. *Genetics* **2000**, *156*, 571−578.
(33) Hirao, I.; Kawai, G.; Yoshizawa, S.; Nishimura, Y.; Ishido, Y.; Watanabe, K.; Miura, K. *Nucleic Acids Res.* **1994**, *22*, 576−582.
(34) Piotto, M.; Saudek, V.; Sklenar, V. *J. Biomol. NMR* **1992**, *2*, 661−665.
(35) Sklenar, V.; Piotto, M.; Leppik, R.; Saudek, V. *J. Magn. Reson., Ser. A* **1993**, *102*, 241−245.
(36) Wijmenga, S. S.; van Buuren, B. N. M. *Prog. Nucl. Magn. Reson. Spectrosc.* **1998**, *32*, 287−387.
(37) Ornstein, R. L.; Rein, R.; Breen, D. L.; MacElroy, R. D. *Biopolymers* **1978**, *17*, 2341−2360.
(38) Altona, C.; Faber, D. H.; Westra Hoekzema, A. J. A. *Magn. Reson. Chem.* **2000**, *38*, 95−107.
(39) Wijmenga, S. S.; Kruithof, M.; Hilbers, C. W. *J. Biomol. NMR* **1997**, *10*, 337−350.
(40) Allawi, H. T.; SantaLucia, J. Jr. *Biochemistry* **1998**, *37*, 2170−2179.
(41) Patel, D. J.; Kozlowski, S. A.; Ikuta, S.; Itakura, K. *Biochemistry* **1984**, *23*, 3207−3217.
(42) Webster, G. D.; Sanderson, M. R.; Skelly, J. V.; Neidle, S.; Swann, P. F.; Li, B. F.; Tickle, I. J. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 6693−6697.
(43) Peyret, N.; Seneviratne, P. A.; Allawi, H. T.; SantaLucia, J. Jr. *Biochemistry* **1999**, *38*, 3468−3477.
(44) Rinkel, L. J.; van der Marel, G. A.; van Boom, J. H.; Altona, C. *Eur. J. Biochem.* **1987**, *166*, 87−101.
(45) LeBlanc, D. A.; Morden, K. M. *Biochemistry* **1991**, *30*, 4042−4047.
(46) Woodson, S. A.; Crothers, D. M. *Biochemistry* **1988**, *27*, 3130−3141.
(47) Bommarito, S.; Peyret, N.; SantaLucia, J. Jr. *Nucleic Acids Res.* **2000**, *28*, 1929−1934.
(48) Tang, M.; Pham, P.; Shen, X.; Taylor, J. S.; O'Donnell, M.; Woodgate, R.; Goodman, M. F. *Nature* **2000**, *404*, 1014−1018.
(49) Johnson, R. E.; Prakash, S.; Prakash, L. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 3838−3843.
(50) Johnson, R. E.; Washington, M. T.; Haracska, L.; Prakash, S.; Prakash, L. *Nature* **2000**, *406*, 1015−1019.
(51) Ohashi, E.; Bebenek, K.; Matsuda, T.; Feaver, W. J.; Gerlach, V. L.; Friedberg, E. C.; Ohmori, H.; Kunkel, T. A. *J. Biol. Chem.* **2000**, *275*, 39678−39684.
(52) McCulloch, S. D.; Kunkel, T. A. *Cell Res.* **2008**, *18*, 148−161.
(53) Kool, E. T. *Annu. Rev. Biochem.* **2002**, *71*, 191−219.
(54) Echols, H.; Goodman, M. F. *Annu. Rev. Biochem.* **1991**, *60*, 477−511.
(55) Goodman, M. F.; Creighton, S.; Bloom, L. B.; Petruska, J. *Crit. Rev. Biochem. Mol. Biol.* **1993**, *28*, 83−126.
(56) Goodman, M. F. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 10493−10495.
(57) Kondrashov, A. S.; Rogozin, I. B. *Hum. Mutat.* **2004**, *23*, 177−185.