Effect of the Oxidized Guanosine Lesions Spiroiminodihydantoin and Guanidinohydantoin on Proofreading by *Escherichia coli* DNA Polymerase I (Klenow Fragment) in Different Sequence Contexts†

Olga Kornyushyna and Cynthia J. Burrows*

Department of Chemistry, University of Utah, 315 South 1400 East, Salt Lake City, Utah 84112-0850

Received June 23, 2003; Revised Manuscript Received September 11, 2003

ABSTRACT: Oxidative damage to DNA by endogenous and exogenous reactive oxygen species has been directly linked to cancer, aging, and a variety of neurological disorders. The potential mutagenicity of the primary guanine oxidation product 8-oxo-7,8-dihydroguanine (Og) has been studied intensively, and much information is available about its miscoding potential in vitro and in vivo. Recently, a variety of DNA lesions have been identified as oxidation products of both guanine and 8-oxoguanine, among them spiroiminodihydantoin (Sp) and guanidinohydantoin (Gh). To address questions concerning the mutagenic potential of these secondary products of guanine oxidation, the effect of the lesions on proofreading by DNA polymerase was studied in vitro using the Klenow fragment of *Escherichia coli* polymerase I (Kf exo+). For the first time, k_{cat}/K_m values were obtained for proofreading of the X:N mismatches (X = Og, Gh, or Sp; N = A, G, or C). Proofreading studies of the terminal mismatches demonstrated the significance of the sequence context flanking the lesion on the 3′ side. In addition, a sequence dependence was observed for Gh based on the identity of the base on the 5′ side of the lesion providing evidence for a primer slippage mode if N was complementary to the 5′ base. Internal mismatches were recognized by Kf exo+ resulting in the excision of the correct base pairs flanking mismatches from the 5′ side. The absence of a sequence effect for the Gh- and Sp-containing duplexes can be attributed to the severe destabilization of the lesion-containing duplexes that promotes interaction with the exonuclease domain of the Klenow fragment.

An essential feature of all DNA dependent DNA polymerases is the ability to accurately copy DNA in vivo. The high fidelity of this process is determined by the fidelity of each step, nucleotide insertion, base pair extension, proofreading, and the correction of errors by repair mechanisms (*1*−*5*). The efficiency of proofreading reflects a balance between the competing 3′−5′ exonuclease removal of mispairs and polymerization of the next correct nucleotide.

The Klenow fragment of DNA polymerase I (Kf)[1] is one of the best-characterized enzymes of the DNA polymerase family (*6*−*14*). Its 3′−5′ exonuclease mechanism has been extensively studied both structurally and biochemically (*7*, *15*, *16*), and the enzyme has been used in kinetic experiments with undamaged as well as damaged DNA templates (*2*, *17*−*24*). Analyses of the proofreading-proficient DNA polymerases revealed the importance of communication between the polymerase and the exonuclease active sites. Klenow fragment proofreads mismatches processively by shuffling the DNA substrate between the polymerase and the exonuclease active sites without dissociating from the DNA (*8*). The intrinsic affinity of the exonuclease domain for DNA duplexes is related to thermodynamic stability because the 3′−5′ exonuclease of Kf strongly prefers a single-stranded substrate (*6*, *7*, *10*, *15*, *25*, *26*), while a correctly paired template/primer substrate preferentially binds to the polymerase domain.

In general, the exonuclease catalytic rates are enhanced by any factor that destabilizes duplex DNA. Proofreading efficiency is a function of the composition of the mismatch, its location, and the sequence context that surrounds the mismatch as well as the nature and extent of DNA contacts between both the polymerase and the exonuclease domains (*27*−*31*). Fluorescence spectroscopy studies by Carver, et al. (*32*) demonstrated that an AT-rich terminus showed greater partitioning into the exonuclease site of Kf exo+ as compared to a GC-rich terminus, again illustrating that binding to the exonuclease domain is favored by sequences that promote increased melting capacity of the primer terminus. On the basis of these observations, a mechanism for 3′ editing was proposed in which the selectivity for editing is attributed partially to both the melting capacity of the DNA and the weaker binding of the DNA to the polymerase site (*32*).

Even though the increased melting capacity of the primer terminus obviously has an effect on DNA binding to the

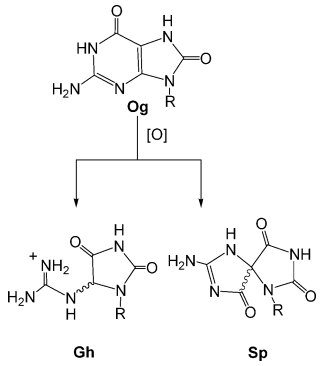
† This work was supported by the NIH (CA 90689). Funds from the NSF (CHE-9807669) and the University of Utah Institutional Funds Committee were provided for purchase of the mass spectrometer used in this study.

* Corresponding author. Phone: (801) 585-7290. Fax: (801) 585-0024. E-mail: burrows@chem.utah.edu.

[1] Abbreviations: Og, 8-oxo-7,8-dihydroguanine; Gh, 5-guanidinohydantoin; Sp, spiroiminodihydantoin (or 2-imino-5,5′-spirodihydantoin); Kf exo+, *Escherichia coli* Klenow fragment DNA polymerase I; ESI-MS, electrospray ionization mass spectrometry.

10.1021/bi0350755 CCC: $25.00
Published on Web 10/17/2003

Scheme 1: Oxidation of Og to Gh and Sp



exonuclease domain, proofreading efficiency does not always correlate with the thermodynamic properties of undamaged DNA duplexes. The free energy difference for partitioning of correct base pairs versus mismatches in undamaged DNA duplexes is found to be −0.6 to −0.7 kcal/mol, while the difference in melting free energies between correct base and mismatches is calculated to be −0.2 kcal/mol for blunt-ended duplexes (*32*, *33*). However, a much better correlation between 3′ editing and thermodynamic stability and the partitioning energy was observed by Morales and Kool (*34*) for oligonucleotides with overhanging ends. They thus concluded that the use of blunt-ended duplexes as a model may be misleading because the overhanging ends are important to overall base-pair stabilities. Studies by the same authors suggested that the stability of the DNA terminus is the major determinant of the editing rate and also pointed to the importance of hydrogen bonding in proofreading (*23*).

Nucleoside studies in our laboratory have recently identified two major pathways of 8-oxo-7,8-dihydroguanine (Og) one-electron oxidation that lead to the formation of 5-guanidinohydantoin (Gh) and 2-imino-5,5′-spirodihydantoin (more commonly referred to as spiroiminodihydantoin, Sp), as shown in Scheme 1 (*35*). Importantly, these lesions also arise from direct oxidation of deoxyguanosine by type I (*35*, *36*) and type II photooxidation (*37*, *38*) and peroxyl radicals (*39*) as well as through the action of a variety of other oxidants on Og (*40*−*43*). Further studies of short oligodeoxynucleotides have shown that the oxidation of Og in duplex DNA led to the formation of Gh as the major product. In single-stranded oligodeoxynucleotides, the reaction was highly temperature dependent leading to Gh at low temperatures and Sp at high temperatures (*44*−*46*). Recently, our laboratory determined the thermal stability of an array of damaged DNA duplexes containing the following products of guanine oxidation: Og, Gh, and Sp. The duplexes were designed to mimic the template/primer DNA systems used in the polymerase kinetic assays and contained lesions in the template strand forming either a terminal mismatch or an internal mismatch placed two bases away from the primer terminus (*46*). These studies showed that the presence of the lesions greatly affected the thermal stability of the duplex, suggesting the present study of the effect of the destabilizing lesions on Klenow fragment proofreading ability.

EXPERIMENTAL PROCEDURES

Materials and Instrumentation. Reagents and enzymes were purchased from the following sources: Na_2IrCl_6 from Alfa Aesar (Ward Hill, MA), Og phosphoramidites from Glen Research (Sterling, VA), [γ-^{32}P] ATP (6000 Ci/mmol) from Amersham (Arlington Heights, IL), and T4 polynucleotide kinase and DNA polymerase I large fragment (Klenow) from New England Biolabs (Beverly, MA). Oligodeoxynucleotides were synthesized with an Applied Biosystems synthesizer (ABI 392 B). ESI mass spectra were recorded on a Micromass Quattro II.

Oligonucleotide Synthesis, DNA Oxidation with Na_2IrCl_6, and Analysis of the Oxidation Products. Deoxyribonucleotides with or without modification were synthesized on an Applied Biosystems Model 392B synthesizer using the manufacturer's protocols. Oligomers containing Og were manually deprotected with concentrated NH_4OH with 0.25 M β-mercaptoethanol (to prevent unwanted oxidation of Og) at 65 °C for 17 h, after which NH_3 was removed in vacuo. Gel electrophoresis on 15 or 20% polyacrylamide 0.8 mm gels with 7 M urea at 40 W provided the purified oligonucleotides. The purity of synthesized and modified duplexes was confirmed by ESI mass spectral analysis on a Micromass Quattro II instrument. Selective oxidation yielding either Sp or Gh was achieved by treating the corresponding Og-containing templates with Na_2IrCl_6 followed by ESI-MS analysis of the samples as described previously (*46*). HPLC analysis indicated that the TGhA and CGhG templates were 98% pure with a small impurity of Og, although no Og was detectable by mass spectrometry (see Supporting Information). All kinetic experiments were performed on unpurified Og- and Gh-containing templates. The CSpG and TSpA templates were 90−95% pure with the principal impurity being Gh. Further purification of the CXG and TXA (X = Gh, Sp) templates was accomplished by HPLC using a Dionex DNAPac PA-100 column. The isocratic buffer systems consisting of 35% solvent A and 65% solvent B and 44% solvent A and 56% solvent B were used to purify the TSpA and TGhA templates, respectively. Purification of the CSpG and CGhG templates required isocratic buffer systems consisting of 37% solvent A and 63% B and 42% solvent A and 58% solvent B, respectively. Solvent A was 10% acetonitrile and 90% H_2O, while solvent B was 10% acetonitrile and 90% 1.5 M ammonium acetate (pH 7). The flow rate was 1.0 mL min^{-1}, and UV spectra were recorded at 260 nm. Prior to use, all resulting oligodeoxynucleotides were dialyzed using 2000 MWCO dialysis tubing for 24 h against H_2O, followed by dialysis against 1 mM NaCl for an additional 24 h. The purity of the oligodeoxynucleotides was determined by HPLC using a Dionex DNAPac PA-100 column and a linear gradient of 35−100% B in 30 min. The resulting HPLC traces are presented in the Supporting Information. Gh consists of two interconverting diastereoisomers that could not be stably isolated. Sp is formed as two diastereoisomers. Previous work (*46*) indicated that the two structures behave similarly in the steady-state kinetic experiments, so they were used as a mixture for these studies. Since HPLC analysis revealed that Sp-containing templates are the least pure among all studied oligonucleotides, steady-state kinetic experiments were repeated for these templates and used for interpretation of the results (see Supporting Information).

End Labeling of Primers and Template/Primer Annealing. The 20nt primers were 5′ end labeled using T4 polynucleotide kinase and [γ-^{32}P]ATP. Unreacted [γ-^{32}P]ATP was

separated from labeled oligonucleotides with a MicroSpin G-25 column (Amersham Pharmacia Biotech). The primers were annealed to the templates in 10 μL solutions of 1x EcoPol buffer (10 mM Tris-HCl, 5 mM $MgCl_2$, and 7.5 mM DTT, pH 7.5) with a template/primer nanomolar ratio of 2:1, yielding a final concentration of 100 and 50 nM for template and primer, respectively.

Qualitative Analysis of Proofreading by Klenow Fragment (exo+) in Different Sequence Contexts. All DNA proofreading reactions were started by mixing a solution containing the template/primer mixture with a solution of DNA polymerase. The final concentrations were 25 nM for template/primer and 0.5 units (370 nM) for the enzyme. Reaction mixtures were incubated at room temperature for 30, 90, and 180 s. Reactions were quenched with 2 μL of termination solution (95% formamide, 0.1% bromophenol blue, and 0.1% xylene cyanol). The samples were heated at 90 °C for 3 min and loaded onto a 15% polyacrylamide/7 M urea gel and then visualized by storage phosphor autoradiography (Molecular Dynamics Storm 840) using ImageQuant software (version 4.2a).

Steady-State Kinetic Analysis of Proofreading in Different Sequence Context. Reactions run under steady-state conditions were all catalyzed by Kf exo+ thermostated at 25 °C. In all reactions, the final concentrations of the annealed template/primer were varied from 50 to 5 nM. The amounts of polymerase used for each template/primer system were adjusted in the preliminary experiments to ensure steady-state conditions as described (*4*, *22*, *23*). The steady-state proofreading reactions for TXA/AT14 and CXG/AC14 (X = Og, Sp) duplexes utilized 3.68 nM Kf exo+. However, TGhA/AT14 and CGhG/AC14 systems required 7.36 nM enzyme. When G was annealed opposite the lesions in the 14nt duplexes, TOgA/GT14 and COgG/GC14 utilized 2.5 nM, TSpA/GT14 and CSpG/GC14 3.68 nM, TGhA/GT14 2.5 nM, and CGhG/CC14G 2 nM Kf exo+, respectively. In the case of the 16nt duplexes, reaction conditions were adjusted to yield less than 25% excision as follows: TXA/AAT16 (X = Og, Gh, or Sp) required 7.36 nM enzyme, CGhG/GAC16 utilized 29.44 nM, and CSpG/GAC16 utilized 22.1 nM, respectively. When C is placed opposite the lesion, the TXA/ACT16 substrates (X = Gh or Sp) utilized 0.92 nM Kf exo+, while TOgA/ACT16 required 1.84 nM, COgG/GCC16 utilized 3.68 nM, and CXG/GCC16 (X = Gh or Sp) required 1.84 nM enzyme. All CXG/GGC16 and TXA/AGT16 (X = Og, Gh, or Sp) substrates utilized 1.84 nM Kf exo+. Finally, the unmodified TGA/ACT16 control utilized 1.84 nM enzyme. Kinetic parameters for CGG/GCC16 and COgG/GAC16 systems were not obtained since the conversion of the original primers was too slow to be quantitated accurately under the conditions employed. The reactions were quenched with 2 μL of a termination solution and applied on a 15% denaturing polyacrylamide gel. The excision of a 3′ terminus of the radiolabeled primer was quantified by storage phosphor autoradiography. The initial velocity (*V*, nM/min) for the excision was obtained based on the equation

$$V = [S]AI_1/(I_0 + I_1)/3 \quad (1)$$

where [S] is the concentration of template/primer, and I_0 and I_1 are band intensities of the original primer (14nt or 16nt) and the excised product (13nt or 15nt) after Kf exo+ reaction, respectively. k_{cat} and K_m values were obtained from nonlinear regression fitting of the data points using KaleidaGraph version 3.09 (Synergy Software, Reading, PA). Weighted error analysis was performed using KaleidaGraph version 3.09; errors for the k_{cat}/K_m ratios were obtained by standard error propagation methods. All the experiments were performed at least three times.

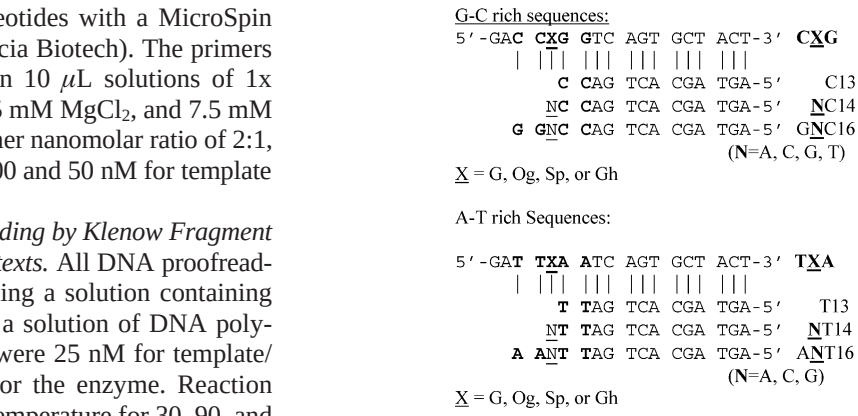


FIGURE 1: Sequences of templates and primers used in proofreading studies.

RESULTS

Substrate Design and Analysis. To study the processing of X:N (X = Og, Gh, or Sp; N = C, A,or G) mismatches by Kf exo+, we designed an array of DNA duplexes that mimic template/primer systems. To investigate the effect of sequence context and the location of the mismatch on proofreading efficiency by Kf exo+, the X:N mismatches were flanked by either an AT-rich or a GC-rich region as shown in Figure 1. In these studies, all the mismatched pairs were formed by annealing a corresponding primer to the template containing a specific lesion site. Particularly, we studied two groups of 18-mer DNA templates that contained either a high G−C content (CXG = 5′-CC**X**GG-3′ within an 18nt) or a high A−T content (TXA = 5′-TT**X**AA-3′ within an 18nt) flanking the lesion site (Figure 1), where X was either G, Og, Sp, or Gh. A group of complementary primers was then annealed to these templates. As controls, two 13-mer primers (C13 = 3′-CC... and T13 = 3′-TT...) were synthesized and annealed immediately upstream of the lesion site. To examine the proofreading of the bases incorporated opposite these lesions, a subset of 14-mer primers (NC14 = 3′-**N**CC... and NT14 = 3′-**N**TT...; where N = A, C, or G; see Figure 1 for the remainder of the sequence) containing each one of the canonical bases directly across from the lesion site was analyzed. Primer extension studies in various sequence contexts showed that Kf exo− possesses an ability to extend past the oxidized lesions, and the amounts of full extension are sequence dependent (*46*). The effects of these lesions within a duplex were also determined using a subset of 16-mer primers with internal mismatches (GNC16 = 3′-GG**N**CC... and ANT14 = 3′-AA**N**TT...; where N = A, C, or G) that contained two more bases located downstream of the lesion site.

Og-containing templates were synthesized and characterized using ESI-MS as described in the Experimental Procedures, and the templates were further oxidized and purified using a published procedure (*47*, *48*). The purity of the Gh-

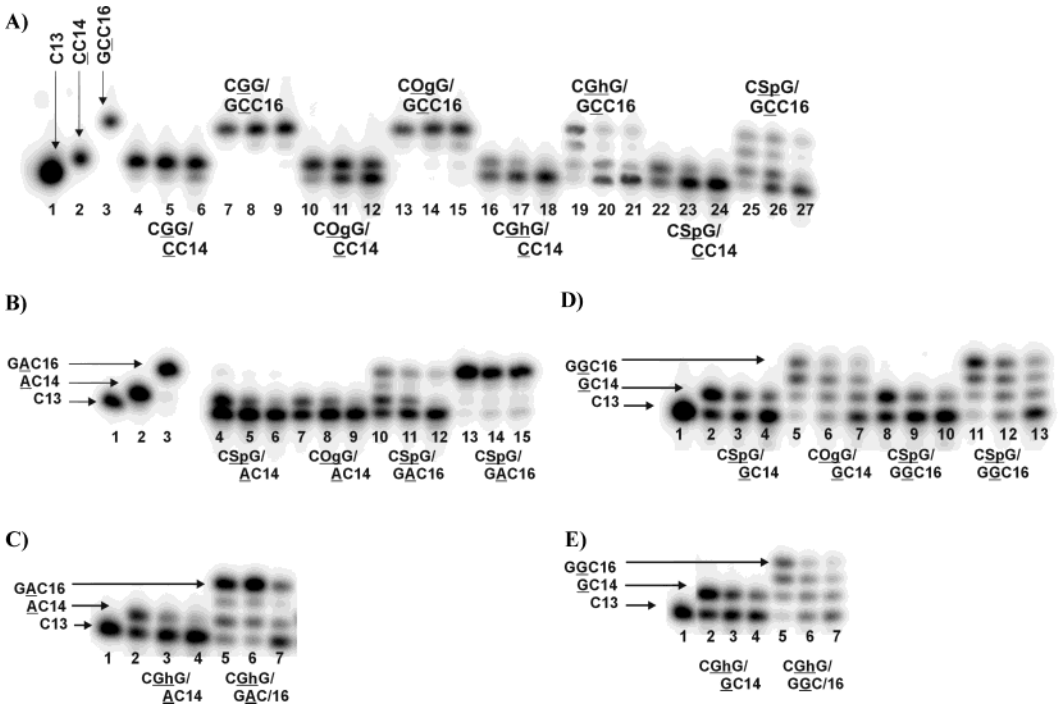


FIGURE 2: Typical gel data for proofreading of the terminal (CXG/NC14) and internal (CXG/GNC16) mismatches by KF exo+. (A) N = C; X = G (lanes 4−9), Og (lanes 10−15), Gh (lanes 16−21), or Sp (lanes 22−27). Template/primer systems were incubated with Kf exo+ for 30 s (lanes 4, 7, 10, 13, 16, 19, 22, and 25), 90 s (lanes 5, 8, 11, 14, 17, 20, 23, 26), and 180 s (lanes 6, 9, 12, 15, 18, 21, 24, 27). (B) N = A; X = Sp (lanes 4−6 and 10−12) or Og (lanes 7−9 and 13−15). Template/primer systems were incubated with Kf exo+ for 30 s (lanes 4, 7, 10, 13), 90 s (lanes 5, 8, 11, 14), and 180 s (lanes 6, 9, 12, 15). (C) N = A; Gh (lanes 1−7). Template/primer systems were incubated with Kf exo+ for 30 s (lanes 2 and 5), 90 s (lanes 3 and 6), and 180 s (lanes 4 and 7). (D) N = G; X = Sp (lanes 4−6 and 10−12), Og (lanes 7−9 and 13−15). Template/primer systems were incubated with Kf exo+ for 30 s (lanes 4, 7, 10, 13), 90 s (lanes 5, 8, 11, 14), and 180 s (lanes 6, 9, 12, 15). (E) N= G; Gh (lanes 1−7). Template/primer systems were incubated with Kf exo+ for 30 s (lanes 2 and 5), 90 s (lanes 3 and 6), and 180 s (lanes 4 and 7). See Experimental Procedures for details.

and Sp-containing samples was once again confirmed by ESI-MS and HPLC as described previously (see Experimental Procedures and Supporting Information).

Qualitative Analysis of Recognition of the Terminal and Internal X:N Mismatches (X = Og, Gh, or Sp; N = C, A, or G) by Kf exo+ in Different Sequence Context. To gain understanding of the basic trends involved in recognition of the terminal and internal X:N mismatches in different sequence contexts by Kf exo+, we examined all the template/primer systems shown in Figure 1 with an excess of the polymerase. Each template/primer system was incubated with an excess of Kf exo+ (see Experimental Procedures) to ensure proper binding of the substrate by polymerase (*22*, *23*, *49*). Templates with X = G (CGG and TGA) were used as important controls in these studies. Additional controls included templates annealed to the 13nt primers (C13 and T13), to the corresponding 14-mers, and to 16-mers in the absence of Kf exo+.

Qualitative analysis of the terminal and internal X:C mismatches is presented in Figures 2A and 3. As expected, normal terminal and internal C:G base pairs in CGG/CC14 and CGG/GCC16 were not excised by the polymerase (Figure 2, lanes 4−9). The terminal Og:C mismatch was visibly excised by Kf exo+ in 180 s (Figure 2A, lanes 10−12 and Figure 3A); however, the addition of two GC base pairs on the 5′ side of the Og:C mismatch stabilized it significantly (Figure 3b). At the same time, terminal and internal Gh:C and Sp:C mismatches were recognized by Kf exo+ in 180 s with similar efficiency (Figure 2A, lanes 16−27). The absence of the mismatch position effect in this case correlates with the low overall stability of these duplexes (*46*).

Proofreading of the terminal and internal X:A and X:G (X = Og, Gh, or Sp) mismatches with an excess of Kf exo+ was also examined. The results of these experiments are presented in Figures 2 and 3 for the GC-rich sequences and in the Supporting Information. These studies indicated the importance of the mismatch identity and position, and the sequence context (GC-rich vs AT-rich) surrounding the mismatch for its recognition by Kf exo+. While the terminal CG base pair in the CGG/GCC16 control duplex was extremely stable to excision by Kf exo+, there was noticeable excision in the TGA/ACT16 control. All the duplexes were formed with the excess of the template strand; thus, high background excision in the TGA/ACT16 control most likely is not due to the presence of the free primer. The lesser overall stability of the TA versus CG base pairs located 5′ to the lesion site is more likely to be accountable for the difference observed for the CGG/GCC16 and TGA/ACT16 duplexes. The same effect was also apparent for the duplexes containing internal Og:C mismatches (Figure 2A and Supporting Information), and editing the TOgA/ACT16 duplex correlates with the background excision in the TGA/ACT16 control. In general, Og:N (N = A or C) mismatches were stable to proofreading especially when placed at the internal position within the DNA duplex, and GC-rich sequences containing oxidized lesions tended to be more stable than their AT-rich counterparts. The Sp:C and Gh:C mismatches

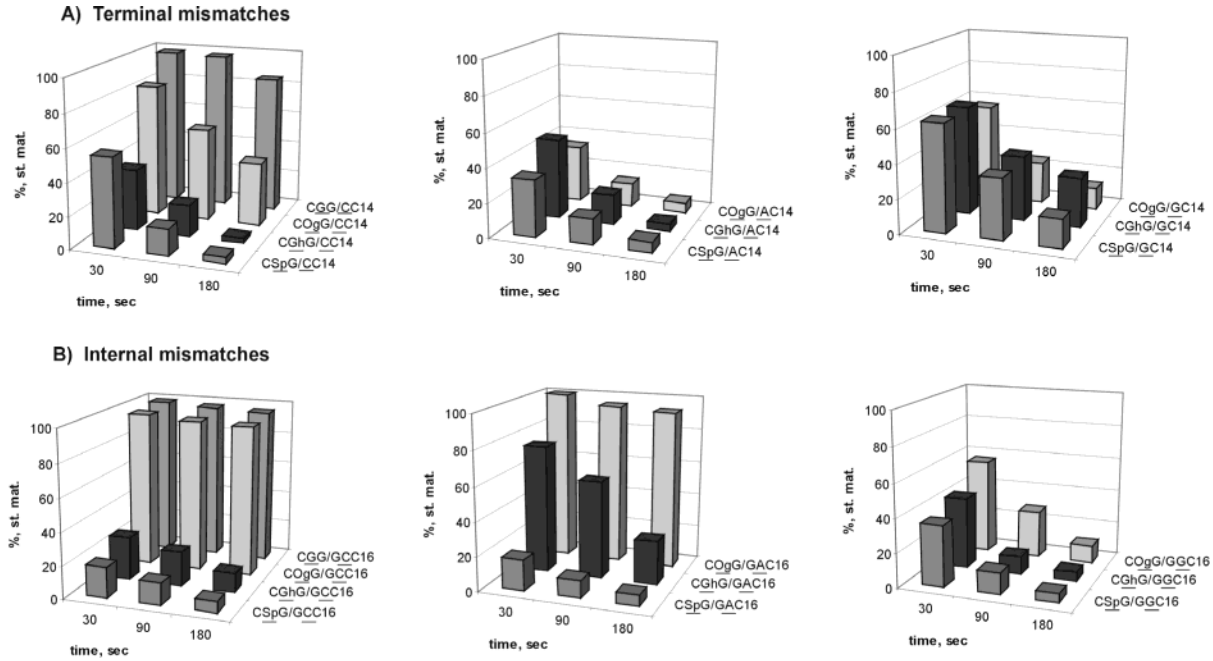


FIGURE 3: Quantitative representation of the proofreading reactions presented in Figure 2. (A) Proofreading of the terminal CXG/NC14 (X = G, Og, Gh, Sp; N = C, A, G) mismatches. (B) Proofreading of the duplexes containing internal CXG/GNC16 (X = G, Og, Gh, Sp; N = C, A, G) mismatches. Bars represent a relative amount of starting material (%) left after incubation of the corresponding duplexes with Kf exo+ for 30, 90, or 180 s as described in Experimental Procedures.

Table 1: Steady-State Kinetic Parameters of Proofreading of Terminal Mismatches at 25 °C Placed in the Context of T*X*A/*N*T14 and CXG/NC14 Duplexes (Figure 1), Where X = Og, Gh, or Sp and N = C, A, or G

TXA/NT14	k_{cat}, s^{-1}	K_m, μM	k_{cat}/K_m	CXG/NC14	k_{cat}, s^{-1}	K_m, μM	k_{cat}/K_m
TOgA/CT14	0.0009 ±0.0001	0.077 ±0.004	0.01 ±0.002	COgG/CC14	0.0014 ±0.0001	0.077 ±0.011	0.02 ±0.003
TGhA/CT14	0.0064 ±0.0009	0.028 ±0.004	0.23 ±0.005	CGhG/CC14	0.0055 ±0.0003	0.062 ±0.004	0.10 ±0.008
TSpA/CT14	0.0022 ±0.0006	0.007 ±0.001	0.29 ±0.09	CSpG/CC14	0.0020 ±0.0002	0.013 ±0.003	0.17 ±0.04
TOgA/AT14	0.0034 ± 0.0005	0.016 ± 0.005	0.22 ± 0.08	COgG/AC14	0.0025 ± 0.0009	0.012 ± 0.006	0.25 ± 0.17
TGhA/AT14	0.0021 ± 0.0001	0.004 ± 0.0002	0.6 ±0.04	CGhG/AC14	0.0049 ± 0.0003	0.036 ± 0.007	0.14 ± 0.03
TSpA/AT14	0.0016 ± 0.0001	0.011 ± 0.003	0.15 ± 0.04	CSpG/AC14	0.0028 ± 0.0002	0.023 ± 0.004	0.12 ± 0.02
TOgA/GT14	0.0231 ± 0.0015	0.035 ± 0.011	0.69 ± 0.22	COgG/GC14	0.0088 ± 0.0016	0.013 ± 0.007	0.65 ± 0.19
TGhA/GT14	0.0088 ± 0.0005	0.051 ± 0.018	0.18 ± 0.06	CGhG/GC14	0.0078 ± 0.0013	0.011 ± 0.005	0.61 ± 0.29
TSpA/GT14	0.0030 ± 0.0004	0.031 ±0.008	0.10 ± 0.02	CSpG/GC14	0.0041 ± 0.0008	0.028 ± 0.003	0.06 ± 0.01

were particularly susceptible to proofreading in both sequence contexts.

Steady-State Kinetics of Proofreading of Terminal Mismatches by Kf exo+. To investigate trends involving sequence context and the influence of mismatch identity on the exonucleolytic excision further, we obtained steady-state kinetic parameters for the TXA/NT14 and CXG/NC14 template/primer systems (X = Og, Gh, or Sp; N = A, G, or C; Figure 1) at 25 °C. Terminal mismatches were assembled by annealing 14nt primers to the appropriate 18nt templates. The duplexes were then incubated with the corresponding amounts of Klenow fragment exo+ for 3 min as described in the Experimental Procedures. The initial velocities of proofreading with various concentrations of the substrates were determined using eq 1 (see Experimental Procedures). A typical S−V (substrate concentration vs relative velocity) plot for the TOgA/AT14 duplex is shown in Figure 4A and is characteristic of Michaelis−Menten kinetics. Steady-state parameters were then obtained from the Michaelis−Menten plots by nonlinear regression fitting of the data points. The results of these experiments are summarized in Table 1, and the relative catalytic efficiencies are plotted in Figure 4B for clarity.

When an Og:C pair was placed at the terminal position, the resulting COgG/CC14 and TOgA/CT14 duplexes were extremely resistant to proofreading by Kf exo+ (k_{cat}/K_m = 0.02 ± 0.003 and 0.01 ± 0.002 s^{-1} μM^{-1}, respectively). An Og:A mismatch placed at the terminal end of the duplex yielded similar relative catalytic efficiencies (k_{cat}/K_m) in both AT- and GC-rich sequence contexts (0.22 ± 0.08 and 0.25 ± 0.17 s^{-1} μM^{-1}, respectively), consistent with the qualitative studies described previously. Also in accord with the qualitative studies, excision of the terminal Og:G was three times more efficient than that of the Og:A pair in both

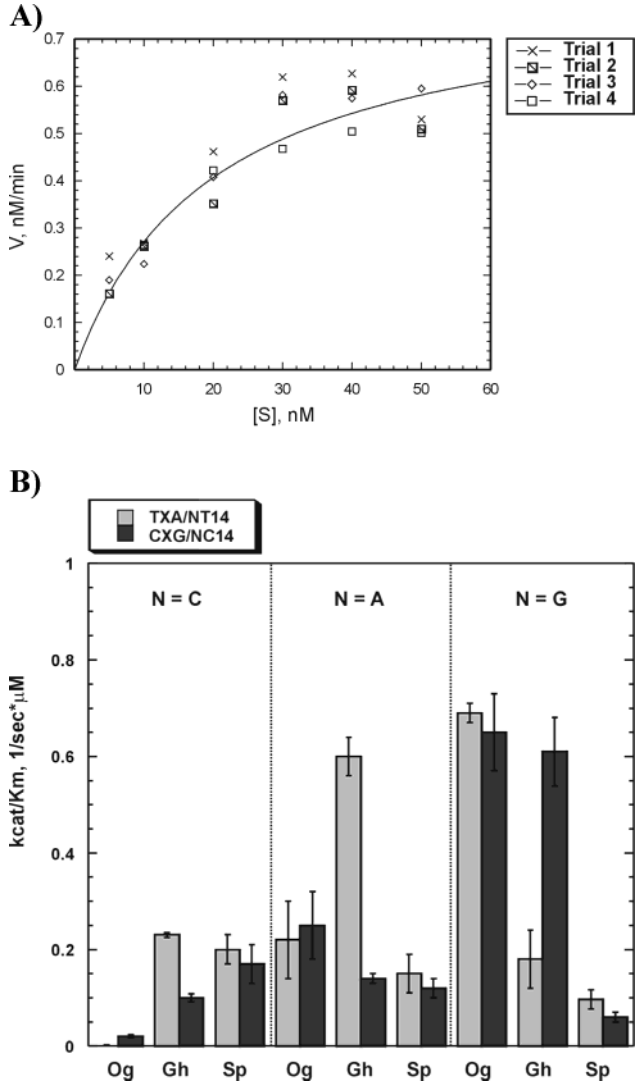


FIGURE 4: Quantitative analysis of proofreading of terminal mismatches. (A) Typical S−V plot obtained for the TOgA/AT14 duplex. (B) Relative catalytic efficiencies of proofreading of terminal mismatches TXA/NT14 and CXG/NC14 (X = Og, Gh, Sp; N = C, A, G).

sequence contexts (0.69 ± 0.22 and 0.65 ± 0.19 s^{-1} μM^{-1} vs 0.22 ± 0.08 and 0.25 ± 0.17 s^{-1} μM^{-1}, respectively).

The terminal Gh:C mismatch was excised two times more efficiently when flanked by two AT base pairs on the 3′ side than the mismatch flanked by two GC base pairs (0.23 ± 0.005 vs 0.10 ± 0.008 s^{-1} μM^{-1}, respectively). The same trend was observed in the case of a terminal Sp:C mismatch where the AT-rich duplex was excised 1.2 times more efficiently than the GC-rich duplex (0.29 ± 0.09 vs 0.17 ± 0.04 s^{-1} μM^{-1}, respectively).

The K_m value for the CGhG/AC14 system is 10 times higher than that for the TGhA/AT14 duplex. The catalytic efficiency of the proofreading of the GC-rich CGhG/AC14 template/primer system is about four times lower than k_{cat}/K_m for TGhA/AT14 (0.14 ± 0.03 and 0.60 ± 0.04 s^{-1} μM^{-1}, respectively). It is noteworthy that the Gh:G terminal mismatch in the GC-rich duplex was excised about four times more efficiently than the same mismatch placed in the AT-rich context (k_{cat}/K_m ratio for the Gh:G placed in the GC-rich duplex is 0.61 ± 0.29 s^{-1} μM^{-1}, and the k_{cat}/K_m value for its AT-rich counterpart is only 0.18 ± 0.06 s^{-1} μM^{-1}).

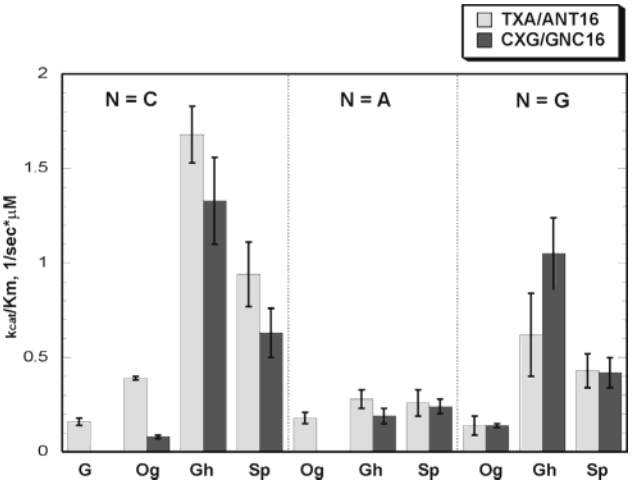


FIGURE 5: Quantitative analysis of proofreading of internal mismatches. Relative catalytic efficiencies of proofreading of internal mismatches TXA/ANT16 and CXG/GNC16 (X = Og, Gh, Sp; N = C, A, G).

Thus, the effect of the sequence context on the efficiency of proofreading for the Gh:A mismatch is exactly the opposite of that for the Gh:G terminal mismatch (see Discussion).

In contrast to the Gh:A mismatch, the sequence context does not seem to play a role in the excision of the Sp:A mismatch (k_{cat}/K_m for the GC-rich system is 0.15 ± 0.04 s^{-1} μM^{-1}, and k_{cat}/K_m for the AT-rich duplex is 0.12 ± 0.02 s^{-1} μM^{-1}). The Sp:G mismatch is slightly more resistant to proofreading by Kf exo+ than the Sp:A mismatch in both GC- and AT-rich sequences (k_{cat}/K_m = 0.06 ± 0.01 and 0.09 ± 0.02 s^{-1} μM^{-1}, respectively).

Effect of Internal Mismatches on Proofreading. Steady-State Kinetics of Proofreading of the TXA/ANT16 and CXG/GNC16 (X = Og, Gh, or Sp; N = A, G, or C) Substrates. To examine the effect of a lesion X on the excision of the correct G or A from the matched 3′ end of the primer, 16nt primers GNC16 and ANT16 were annealed to the corresponding 18nt templates (as shown in Figure 1). The resulting duplexes were then incubated with appropriate amounts of Kf exo+ polymerase as described in the Experimental Procedures. The removal of the single nucleotide G from the correct terminal CG base pair (CXG/GNC16 systems) and A from the correct terminal TA base pair (TXA/ANT16) was quantified. Data were treated as described previously, and the results of these experiments are summarized in Table 2, while relative catalytic efficiencies are plotted in Figure 5 for simplicity.

Kinetic parameters could not be obtained for the undamaged matched CGG/GCC16 duplex that was used as a control in our qualitative studies described previously. The conversion of the GCC16 primer for this extremely stable GC-rich substrate was too slow to be quantified under our conditions. However, we were able to obtain kinetic data for the AT-rich TGA/ACT16 duplex. The catalytic efficiency of the removal of the terminal TA base pair by Kf exo+ (k_{cat}/K_m = 0.16 ± 0.03 s^{-1} μM^{-1}) was used as a reference in the studies of the internal mismatches. When C was placed opposite Og in the TOgA/ACT16 duplex, removal of the correct TA base pair from the primer terminus was 2.5 times more efficient than for the TGA/ACT16 control. At the same time, in the case of the COgG/GCC16 duplex, the ability of

Table 2: Steady-State Kinetic Parameters of the Proofreading of the Internal Mismatches at 25 °C Placed in the Context of T*X*A/A*N*T16 and C*X*G/G*N*C16 Duplexes (Figure 1), where X = G, Og, Gh, or Sp and N = C, A, or G

TXA/ ANT16	k_{cat}, s^{-1}	K_m, μM	k_{cat}/K_m	CXG/ GNC16	k_{cat}, s^{-1}	K_m, μM	k_{cat}/K_m
TGA/ ACT16	0.0058 ±0.0007	0.036 ±0.008	0.16 ±0.03	CGG/ GCC16	N.D.[a]	N.D.	N.D.
TOgA/ ACT16	0.0059 ±0.0002	0.015 ±0.001	0.39 ±0.03	COgG/ GCC16	0.0013 ±0.0007	0.017 ±0.005	0.08 ±0.05
TGhA/ ACT16	0.0084 ±0.0008	0.005 ±0.002	1.68 ±0.15	CGhG/ GCC16	0.0159 ±0.0027	0.012 ±0.003	1.33 ±0.40
TSpA/ ACT16	0.0255 ±0.0045	0.027 ±0.004	0.94 ±0.22	CSpG/ GCC16	0.0372 ±0.0078	0.059 ±0.011	0.63 ±0.18
TOgA/ AAT16	0.0086 ± 0.0019	0.051 ± 0.014	0.18 ±0.05	COgG/ GAC16	N.D.	N.D.	N.D.
TGhA/ AAT16	0.0100 ± 0.0010	0.037 ± 0.006	0.28 ±0.05	CGhG/ GAC16	0.0052 ± 0.0006	0.029 ± 0.009	0.19 ±0.06
TSpA/ AAT16	0.0021 ± 0.0003	0.008 ± 0.003	0.26 ±0.01	CSpG/ GAC16	0.0084 ± 0.0020	0.036 ± 0.012	0.24 ±0.08
TOgA/ AGT16	0.0056 ±0.0007	0.041 ±0.008	0.14 ±0.03	COgG/ GGC16	0.0047 ±0.0003	0.034 ±0.004	0.14 ±0.02
TGhA/ AGT16	0.0076 ±0.0006	0.013 ±0.005	0.62 ±0.24	CGhG/ GGC16	0.0045 ±0.0001	0.005 ±0.001	1.05 ±0.21
TSpA/ AGT16	0.0092 ±0.0017	0.023 ±0.009	0.43 ±0.18	CSpG/ GGC16	0.031 ±0.007	0.078 ±0.028	0.42 ±0.18

[a] N.D.—not determined.

Kf exo+ to recognize this internal mismatch diminished considerably ($k_{cat}/K_m = 0.08 \pm 0.05$ vs 0.39 ± 0.03 s^{-1} μM^{-1} for TOgA/ACT16 duplex, Table 2).

Among the duplexes studied, the proofreading efficiencies are the highest for the Gh:C- and Sp:C-containing substrates. Interestingly, sequence context does not play a large role in this case (Figure 5 and Table 2). However, the recognition of the Gh:C mismatch is 1.4 times more efficient than the recognition of the Sp:C mismatch in both sequence contexts (1.68 ± 0.15 and 1.33 ± 0.40 s^{-1} μM^{-1} vs 0.94 ± 0.22 and 0.63 ± 0.18 s^{-1} μM^{-1}, respectively; Table 2).

The Og:A mismatch was placed within the duplex by annealing the AAT16 and GAC16 16nt primers to the corresponding TOgA and COgG 18nt duplexes as shown in Figure 1. When the Og:A base pair is flanked by two upstream C:G base pairs, the duplex is very stable, and no exonucleolytic activity was observed. The S−V plot could not be obtained in this case since the catalytic efficiency for this substrate was essentially zero. When an Og:A base pair was placed in the high AT context, however, the terminal A from the 3′ end of the AAT16 primer was excised by Kf exo+ after incubation for 3 min allowing the determination of the steady-state kinetic parameters for this case. However, the obtained k_{cat}/K_m value is essentially the same as for the control TGA/ACT16 duplex.

In the case of CGhG/GAC16, the k_{cat} value is half that of the TGhA/AAT16 substrate (0.0052 vs 0.0100 s^{-1}); however, relative catalytic efficiencies for the removal of the terminal base pair from the primer ends coincide within a standard deviation for these two substrates (Figure 5). As seen from Figure 5, the k_{cat}/K_m ratios were the same in the case of TSpA/AAT16 and CSpG/GAC16 duplexes. It should be noted that the k_{cat} value for the CSpG/GAC16 substrate was 4 times higher than for the TSpA/AAT16 duplex (Table 2).

Recognition of the internal X:G (X = Og, Gh, or Sp) mismatch appeared to be independent of the sequence context for all three lesions (Table 2). In the AT-rich sequences, Gh:G and Sp:G induce a more efficient removal of the terminal A from the AGT16 primers when compared to the Og-containing duplex ($k_{cat}/K_m = 0.62 \pm 0.24$ and 0.43 ± 0.18 s^{-1} μM^{-1} vs 0.14 ± 0.03 s^{-1} μM^{-1}, respectively). At the same time, these duplexes do not differ in thermodynamic stabilities (see Supporting Information). The same is true for the CGhG/GGC16, CSpG/GGC16, and COgG/GGC16 substrates ($k_{cat}/K_m = 1.05 \pm 0.21$ and 0.42 ± 0.18 s^{-1} μM^{-1} vs 0.14 ± 0.02 s^{-1} μM^{-1}, respectively).

DISCUSSION

In the present study, we examined the effect of two novel guanine oxidation products, guanidinohydantoin and spiroiminodihydantoin, on proofreading ability of the *Escherichia coli* Klenow fragment of polymerase I and its dependence on the mismatch position and sequence context surrounding the mismatch. Thermal stability studies previously carried out in our laboratory demonstrated that the presence of the oxidized lesions (especially Gh and Sp) in the duplex affected the overall duplex stability when the lesion was placed either opposite the 3′ end of the primer strand or was flanked by several base pairs on the 5′ side (*46*). It is a well-established fact that the melted ends of the undamaged DNA template/primer duplexes are more susceptible to exonucleolytic proofreading. In an attempt to determine the correlation between duplex stability and proofreading of the mismatches containing oxidized DNA lesions, an array of DNA duplexes was studied that mimicked template/primer systems previously used to determine thermodynamic stabilities of the lesion-containing duplexes. To assess the effect of the mismatch position on the proofreading efficiency by Kf exo+, we designed the DNA duplexes that contained either terminal mismatches or mismatches within the duplex as shown in Figure 1.

Previous in vitro studies in our laboratory established that dATP and dGTP rather than dCTP are inserted opposite Gh and Sp lesions in the DNA template strand by Kf exo−, thus suggesting that these lesions could be highly mutagenic (*46*, *47*, *50*). To gain insight into the recognition of the mismatched bases located at the terminal or internal positions

of the template/primer duplexes by Kf exo+, we examined the X:N pairs, where N = A, G, or C. The insertion of dATP opposite Og makes this lesion mutagenic causing the A → T transversion mutation; however, the correct dCTP is also readily inserted opposite Og by DNA polymerase I, resulting in a nonmutagenic event (*47*, *51*−*53*). At the same time, when Sp or Gh lesions are present in the template strand, the insertion of dCTP by Kf exo− was never observed (*42*, *46*, *47*). However, in situ oxidation of Og placed opposite C in the duplex makes the study of the X:C (X = Sp and Gh) mismatches biologically relevant. Templates containing unoxidized G at position X (templates CGG and TGA, Figure 1) were used as controls in these studies.

Petruska et al. (*27*) have proposed that the efficiency of exonucleolytic proofreading is governed by the relative proportions of A:T and G:C base pairs in the adjacent duplex region. This model is based on the observation that the intrinsic affinity of the exonuclease site for a duplex DNA substrate is correlated with the thermodynamics of terminus melting and that the 3′−5′ exonuclease domain strongly favors a single-stranded substrate (*6*, *7*, *25*, *54*). Indeed, qualitative studies using an excess of the polymerase implied that the sequence context (AT vs GC) surrounding the mismatch and its position relative to the template/primer end play significant roles in the proofreading efficiency by Kf exo+. To investigate these trends further, we obtained steady-state kinetic parameters for the terminal TXA/NT14 and CXG/NC14 template/primer systems as well as for the TXA/ACT16 and CXG/GCC16 duplexes containing internal mismatches (X = Og, Gh, or Sp; Figure 1). The kinetic parameters for proofreading of the terminal mismatches are detailed in Table 1.

Proofreading of Terminal Mismatches. Consistent with qualitative studies, COgG/CC14 and TOgA/CT14 duplexes were extremely resistant to proofreading (k_{cat}/K_m = 0.02 ± 0.003 and 0.01 ± 0.002 s^{-1} μM^{-1}, respectively). Resistance of the terminal Og:C mismatch to proofreading by Kf exo+ was previously reported by Shibutani et al. (*52*); however, kinetic parameters were not determined at the time. Recent thermodynamic stability studies (*46*) demonstrated that the COgG/CC14 duplex is about 1.6 ± 0.7 kcal/mol more stable than the TOgA/CT14 duplex (Table S1, Supporting Information). Interestingly, the Og:C pair flanked by two AT base pairs on the 3′ side was more stable to proofreading than the Og:C pair in the GC-rich duplex. However, the overall proofreading resistance of the Og:C pair correlates with the highest stability of the Og:C-containing duplexes among all studied substrates (Table S1, Supporting Information). Fluorescence spectroscopy studies established that the free energy difference for partitioning of correct base pairs versus mismatches to the exonuclease domain of the Klenow fragment for undamaged DNA duplexes was −0.6 to −0.7 kcal/mol (*32*). Our UV melting experiments (*46*) showed that the Og:C pair in the COgG/CC14 system actually stabilized the duplex as compared to the undamaged G:C-containing substrate by 0.5 kcal/mol (Table S1, Supporting Information).

Single nucleotide insertion studies (*46*) also suggested that Gh and Sp do not interact favorably with cytosine indicating that the identity of the specific mismatch plays a role in polymerase activity of the oxidized lesions. Thus, it is not surprising that the Gh:C and Sp:C mismatches are more efficiently excised than the Og:C mismatch in both sequence contexts (Table 1). Local geometric distortions introduced by the oxidized lesions placed opposite C might be responsible for susceptibility of the damaged GC-rich and the AT-rich duplexes to proofreading by Kf exo+.

In contrast to the Og:C base pair, proofreading of the Og:N terminal mismatches (N = A or G) appears to be sequence independent. Apparently, AT and GC base pairs 3′ to the lesion site do not influence the proofreading of the terminal Og:A pair (k_{cat}/K_m = 0.22 ± 0.08 and 0.25 ± 0.17 s^{-1} μM^{-1}, respectively; Table 1). It should be noted that an Og:A mismatch placed in the GC-rich duplex (COgG/AC14) has approximately the same stability ($\Delta\Delta G$ = −0.6 ± 0.9 kcal/mol) as the undamaged G:C-containing duplex (CGG/CC14), while the $\Delta\Delta G$ value for the AT-rich duplex is about −1.4 ± 0.2 kcal/mol (Table S1, Supporting Information). Thus, only the terminal Og:A mismatch in AT-rich sequence contexts is destabilizing enough to satisfy the free energy partitioning requirement calculated by Carver et al. (*32*, *33*). Efficient excision of the Og:A mismatch in the GC-rich duplex cannot be explained by thermodynamic factors alone. Interaction of this mismatch with the side chains of the residues in the polymerase domain might promote proofreading in this case (*14*). The identity of the opposite base does not change the lack of sequence effect; however, the Og:G mismatch is more efficiently removed from the terminal end of the primer by the polymerase as compared to Og:A. These observations are consistent with the fact that Og adopts the favorable syn conformation to form a Og:A base pair that does not destabilize the DNA duplex.

An interesting phenomenon was observed in the case of the Gh:A and Gh:G mismatches. While Gh:A-containing duplexes were not an exception from the general GC versus AT stability rule (k_{cat}/K_m (CGhG/AC14) = 0.60 ± 0.04 s^{-1} μM^{-1}; k_{cat}/K_m (TGhA/AT14) = 0.14 ± 0.03 s^{-1} μM^{-1}; Table 1), the GC versus AT context dependence for the Gh:G mismatch is exactly reversed as compared to the Gh:A terminal mismatch (k_{cat}/K_m (CGhG/GC14) = 0.61 ± 0.01 s^{-1} μM^{-1}; k_{cat}/K_m (TGhA/GT14) = 0.18 ± 0.06 s^{-1} μM^{-1}; Figure 4B and Table 1). Note that the catalytic efficiencies of proofreading are the same for the Gh:A mismatch in the AT-rich duplex and the Gh:G mismatch in the GC-rich duplex. Examination of the sequence context for these duplexes showed that they provide an opportunity for primer slippage (Figure 6). A terminal A on the 3′ prime end of the primer might pair more easily with the next correct T than with the opposite Gh lesion in the AT-rich duplex (Figure 6A), and a terminal G prefers pairing with the next correct C to the Gh lesion in the GC-rich sequence (Figure 6B).

Wagner et al. previously demonstrated that treatment of DNA by methylene blue plus light led to the formation of lesions other than Og, which efficiently promoted frameshift mutagenesis in *E. coli* (*55*). Recent studies in nucleosides also showed that Gh and Sp can be formed during photooxidation from G and Og via several photooxidation pathways (*36*, *38*). Taken together, these results suggest that Gh and/or Sp lesions in DNA might be responsible for frameshift mutations in *E. coli*. In addition, since the occupancy of the Klenow fragment exonuclease domain is substantially increased when the polymerase encounters extrahelical bases (*24*), the observed increase in susceptibility to proofreading the Gh:A and Gh:G mismatches may indeed be attributed to the formation of the extrahelical base as

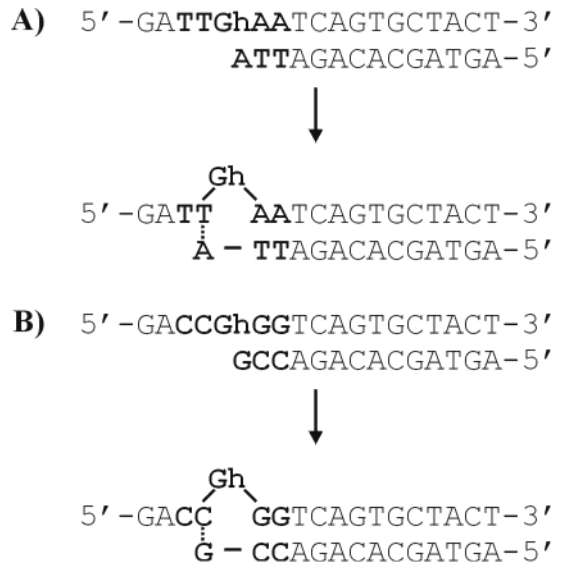


FIGURE 6: Proposed primer slippage model for template/primer systems containing Gh:A and Gh:G terminal mismatches.

proposed in Figure 6. Further substantiation of this idea comes from previous observations (*46*) of full extension past Gh using similar sequences in which sequence-directed primer slippage could occur. A predominant $n - 1$ product was observed in these cases, suggesting that slippage was the principal mechanism of extension.

Considering the destabilizing effect of the Sp:N (N = A or G) mismatches on the overall duplex stabilities (*46*), it is quite surprising that these mismatches were somewhat resistant to proofreading as compared to their Og:N counterparts (Figure 4B). Resistance to proofreading was especially apparent in case of the Sp:G mismatch, and the k_{cat}/K_m value for this mismatch is about 10-fold lower than that of the Og:G mismatch in both sequence contexts (Table 1). In contrast to Gh:A and Gh:G-containing duplexes, sequence-directed primer slippage did not seem to take place for Sp:A and Sp:G mismatches. Curiously, previous studies of full extension using Kf exo− and Sp-containing templates showed predominant $n - 1$ products similar to the behavior of Gh (*46*). The difference between Gh and Sp must be related to the interaction of the extrahelical base with the protein, possibly due to more severe perturbations resulting from the extrusion of a positively charged base, Gh, from the helix as compared to the neutral spirocycle.

Overall, our studies show that Sp:N and Gh:N terminal mismatches (N = A or G) are resistant to proofreading as compared to Og unless they form an extrahelical base. The base-pairing properties of the Gh:N and Sp:N pairs (N = A or G) and the orientation of the lesions in the context of the DNA duplex are yet to be explored. However, it is clear that the flanking bases have a direct effect on the interactions between the terminal mismatches and the proofreading domain of the Kf exo+. Unusual behavior of the Gh and Sp mismatches might be also due to their interaction with the side chains of the residues important for the polymerase fidelity, namely, Tyr-766, Arg-682, and Lys-635 (*14*). Recently, Johnson et al. demonstrated the importance of proper stacking of the Tyr residue with the template base at the preinsertion site of the polymerase for the prevention of frameshift mutations (*56*). Thus, it is possible that in the appropriate sequence context, extrahelical Gh interacts with Tyr-766 in a way that promotes shuffling of the primer to the exonuclease domain of Kf. Since Gh is positively charged, electrostatic interaction with Arg-682 also might be involved in the shuffling of the Gh-containing mismatches to the exonuclease domain.

Proofreading of Internal Mismatches. Fluorescence spectroscopy studies by Carver et al. showed that in undamaged DNA duplexes, single internal mismatches produced larger effects in the equilibrium partitioning of DNA into the exonuclease site of the Klenow fragment than the same mismatches placed at the primer terminus (*33*). A calculated $\Delta\Delta G$ value relative to the matched sequence was between −1.1 and −1.3 kcal/mol for internal mismatches located two, three, or four bases from the primer terminus. They thus concluded that an internal mismatch might produce a local structural distortion in the double helix that inhibits binding to the polymerase domain of Kf.

Melting studies performed in our laboratory showed that the lesions have a considerable effect on duplex stability when placed in the DNA duplex and flanked by correct base pairs on the 5′ side (*46*). When C<u>X</u>G/G<u>N</u>C16 and T<u>X</u>A/A<u>N</u>T16 (X = Gh or Sp) duplexes were compared to the C<u>X</u>G/C13 and T<u>X</u>A/T13 controls (Figure 1), it became apparent that the addition of two more base pairs does not necessarily increase thermodynamic stability of the duplexes. Even though the duplexes contain additional correct C:G and T:A base pairs, it is likely that the presence of the oxidized lesions causes the duplex ends to melt, thus making them candidates for excision in the presence of the polymerase. These observations, together with preliminary qualitative studies using excess polymerase, set the stage for a more detailed examination of the kinetic behavior of the C<u>X</u>G/G<u>N</u>C16 and T<u>X</u>A/A<u>N</u>T16 duplexes (X = OG, Gh, or Sp; N = A, G, or C; Figures 1 and 5 and Table 2).

Generally, greater rates of exonucleolysis are seen for AT-rich sequences as compared with GC-rich sequences (*57*). While kinetic parameters could not be obtained for the undamaged matched C<u>G</u>G/G<u>C</u>C16 duplex due to extremely slow conversion of the GCC16 primer for this quite stable GC-rich system, we were able to obtain kinetic data for the AT-rich T<u>G</u>A/A<u>C</u>T16 duplex (k_{cat}/K_m = 0.16 ± 0.03 s^{-1} μM^{-1}). This number was used as a reference for background exonucleolysis efficiency by Kf exo+ in the studies of the internal mismatches. The destabilizing effect of the internal Og:C mismatch in an AT-rich sequence (0.9 ± 0.3 kcal/mol, Table S4, Supporting Information) correlates with its recognition by Kf exo+ (k_{cat}/K_m (T<u>G</u>A/A<u>C</u>T16) = 0.16 ± 0.03 s^{-1} μM^{-1}; k_{cat}/K_m (T<u>Og</u>A/A<u>C</u>T16) = 0.39 ± 0.03 s^{-1} μM^{-1}, Table 2). In correlation with relative stabilities of the C<u>Og</u>G/G<u>C</u>C16 and T<u>Og</u>A/A<u>C</u>T16 duplexes ($\Delta\Delta G$ = −3.7 kcal/mol, Table S4), the ability of Kf exo+ to recognize the internal Og:C mismatch in a GC-rich duplex diminished considerably as compared to its AT-rich counterpart (Figure 5). In contrast to the Og:C pair, the internal Og:A mismatch does not perturb the duplex significantly even in the high AT content sequence, and the k_{cat}/K_m ratio for this substrate corresponds to the removal of the TA base pair in the corresponding matched DNA duplex (Table 2 and Figure 5).

Recent thermal stability studies in our laboratory (*46*) showed that the Gh:C and Sp:C mismatches located 2 bp from the primer termini caused a dramatic destabilization

of the corresponding duplexes (up to 3.5 kcal/mol for the AT-rich duplexes and up to 6.7 kcal/mol for the GC-rich duplexes, Table S1). Thus, it is not surprising that the proofreading efficiencies are the highest for the Gh:C and Sp:C-containing substrates.

The effect of the sequence context was not at all apparent in proofreading of the Gh:A mismatch. The GC-rich CGhG/GAC16 duplex is about 2.3±0.1 kcal/mol more stable than the AT-rich TGhA/AAT16 duplex (Table S1). In the case of CGhG/GAC16, relative catalytic efficiencies for the removal of the terminal base pair from the primer ends coincide within the standard deviation for these two substrates (Figure 5). There is also a lack of correlation between the stability of the TSpA/AAT16 and CSpG/GAC16 duplexes and the proofreading efficiency. In general, the Sp:A mismatch had an apparent effect on the duplex stability in both GC- and AT-rich sequences relative to the duplexes containing matched base pairs ($\Delta\Delta G = -6.3 \pm 1.1$ and -3.9 ± 0.3 kcal/mol, respectively; Table S1). However, the CSpG/GAC16 duplex was 1.5 ± 0.3 kcal/mol more stable than the TSpA/AAT16 system (Table S1). As seen from Figure 5B, the k_{cat}/K_m ratios were the same in this case. These discrepancies may not be explained by the primer terminus melting alone, and specific interactions of the internal Gh:A and Sp:A mismatches with the protein side chains may be responsible for the lack of sequence effect in these cases.

Recognition of the internal X:G (X = Og, Gh, or Sp) mismatch appeared to be independent of the sequence context for all three lesions (Table 2). Removal of the terminal TA base pair was just as efficient as the excision of the terminal CG base pair in this case (Figure 5; N = G). While the AT-rich duplexes containing the internal X:G mismatches do not differ in the thermodynamic stabilities (*46*), the Gh:G and Sp:G mismatches promote excision more efficiently than the Og:G mismatch. The same is true for the CGhG/GGC16, CSpG/GGC16, and COgG/GGC16 substrates ($k_{cat}/K_m = 1.05 \pm 0.21$ and 0.42 ± 0.18 s^{-1} μM^{-1} vs 0.14 ± 0.02 s^{-1} μM^{-1}, respectively). Thus, the more efficient recognition of the internal Gh:G and Sp:G mismatches cannot be attributed exclusively to thermodynamic factors. The distortions produced by the oxidized lesions placed opposite G upstream of the duplex terminus may play a role in the apparent affinity of the X:G (X = Gh or Sp)-containing duplexes to the exonuclease domain.

The presented data suggest that it is important to understand the effect of the lesions on duplex structure in different sequences. Proofreading efficiency of the terminal mismatches by Kf exo+ appeared to depend on the thermodynamic stabilities of the corresponding duplexes and/or the mismatch identity. These studies suggested that if the Klenow fragment of DNA polymerase I inserted C after encountering Gh or Sp lesions in vivo, it would be more likely to proofread Gh:C and Sp:C mismatches than to extend past the lesions. Interestingly, studies of the Gh:A and Gh:G mismatches placed in AT- and GC-rich sequences suggested that proofreading might be enhanced by sequence-directed primer slippage leading to the formation of the extrahelical base. Possible formation of the extrahelical base leading to subsequent primer slippage and −1 deletion was also previously reported by our laboratory when the $n - 1$ product was observed during full extension of the primer past both Gh and Sp lesions (*46*). However, only extrusion of Gh from the helix strongly promotes proofreading, possibly due to an unfavorable interaction of its positive charge with the polymerase domain. Kf exo+ is also able to recognize internal mismatches containing Gh and Sp lesions placed at position 3 from the primer terminus resulting in the excision of the correct base pairs flanking the mismatches from the 5′ side. Thermal stability and kinetic studies indicated that the presence of oxidized lesions destabilize the DNA helix and promote recognition of the internal mismatches by the exonuclease domain of the Klenow fragment. Additional studies are underway to fully characterize the structure of the lesions in various sequence contexts in duplex DNA to provide additional insight into the recognition of DNA damage by polymerases.

ACKNOWLEDGMENT

We thank Aym M. Berges for helpful discussions of thermal stability studies and Dr. James G. Muller for assistance with mass spectrometric and HPLC studies.

SUPPORTING INFORMATION AVAILABLE

HPLC data for the purity of lesion-containing synthetic oligodeoxynucleotides, qualitative proofreading studies of the terminal and internal X:N mismatches (X = Og, Gh, Sp; N = C, A, G), and tabular thermodynamic data for lesion-containing template/primer systems. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Echols, H., and Goodman, M. F. (1991) *Annu. Rev. Biochem. 60*, 477−511.
2. Eckert, K. A., and Opresko, P. L. (1999) *Mutat. Res. 424*, 221−236.
3. Kunkel, T. A., and Bebenek, K. (2000) *Annu. Rev. Biochem. 69*, 497−529.
4. Goodman, M. F., Creighton, S., Bloom, L. B., and Petruska, J. (1993) *Crit. Rev. Biochem. Mol. Biol. 28*, 83−126.
5. Goodman, M. F., and Fygenson, D. K. (1998) *Genetics 148*, 1475−1482.
6. Beese, L. S., Derbyshire, V., and Steitz, T. A. (1993) *Science 260*, 352−355.
7. Freemont, P. S., Friedman, J. M., Beese, L. S., Sanderson, M. R., and Steitz, T. A. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8924−8928.
8. Joyce, C. M. (1989) *J. Biol. Chem. 264*, 10858−10866.
9. Dzantiev, L., and Romano, L. J. (2000) *Biochemistry 39*, 356−361.
10. Joyce, C. M., and Steitz, T. A. (1994) *Annu. Rev. Biochem. 63*, 777−822.
11. Clark, J. M., and Beardsley, G. P. (1989) *Biochemistry 28*, 775−779.
12. Minnick, D. T., Astatke, M., Joyce, C. M., and Kunkel, T. A. (1996) *J. Biol. Chem. 271*, 24959−24961.
13. Tuske, S., Singh, K., Kaushik, N., and Modak, M. J. (2000) *J. Biol. Chem. 275*, 23759−23768.
14. Minnick, D. T., Bebenek, K., Osheroff, W. P., Turner, R. M. J., Astatke, M., Liu, L., Kunkel, T. A., and Joyce, C. M. (1999) *J. Biol. Chem. 274*, 3067−3075.
15. Beese, L. S., and Steitz, T. A. (1991) *EMBO J. 10*, 25−33.
16. Derbyshire, V., Grindley, N. D. F., and Joyce, C. M. (1991) *EMBO J. 10*, 17−24.
17. Bloom, L. B., Otto, M. R., Beechem, J. M., and Goodman, M. F. (1993) *Biochemistry 32*, 11247−11258.
18. Miller, H., and Grollman, A. P. (1997) *Biochemistry 36*, 15336−15342.
19. Purmal, A. A., Bond, J. P., Lyons, B. A., Kow, Y. W., and Wallace, S. S. (1998) *Biochemistry 37*, 330−338.
20. Paz-Elizur, T., Takeshita, M., and Livneh, Z. (1997) *Biochemistry 36*, 1766−1773.

21. Shibutani, S., and Grollman, A. P. (1993) *J. Biol. Chem. 268*, 11703−11710.
22. Ide, H., Murayama, H., Sakamoto, S., Makino, K., Honda, K., Nakamuta, H., Sasaki, M., and Sugimoto, N. (1995) *Nucleic Acids Res. 23*, 123−129.
23. Morales, J. C., and Kool, E. T. (2000) *Biochemistry 39*, 2626−2632.
24. Lam, W.-C., Van der Schans, E. J. C., Sowers, L. C., and Millar, D. P. (1999) *Biochemistry 38*, 2661−2668.
25. Brutlag, D., and Kornberg, A. (1972) *J. Biol. Chem. 247*, 241−248.
26. Wang, J., Sattar, A. K. M. A., Wang, C. C., Karam, J. D., Konigsberg, W. H., and Steitz, T. A. (1997) *Cell 89*, 1087−1099.
27. Petruska, J., Goodman, M. F., Boosalis, M. S., Sowers, L. C., Cheong, C., and Tinoco, I., Jr. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6252−6256.
28. Mendelman, L. V., Boosalis, M. S., Petruska, J., and Goodman, M. F. (1989) *J. Biol. Chem. 264*, 14415−14423.
29. Bloom, L. B., Otto, M. R., Eritja, R., Reha-Krantz, L. J., and Goodman, M. F. (1994) *Biochemistry 33*, 7576−7586.
30. Thompson, E. H., Bailey, M. F., van der Schans, E. J., Joyce, C. M., and Millar, D. P. (2002) *Biochemistry 41*, 713−722.
31. Lam, W. C., Thompson, E. H., Potapova, O., Sun, X. C., Joyce, C. M., and Millar, D. P. (2002) *Biochemistry 41*, 3943−3951.
32. Carver, T. E., Hochstrasser, R. A., and Millar, D. P. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 10670−10674.
33. Carver, T. E., Jr., and Millar, D. P. (1998) *Biochemistry 37*, 1898−1904.
34. Morales, J. C., and Kool, E. T. (2000) *Biochemistry 39*, 12979−12988.
35. Luo, W., Muller, J. G., Rachlin, E. M., and Burrows, C. J. (2000) *Org. Lett. 2*, 613−616.
36. Luo, W., Muller, J. G., and Burrows, C. J. (2001) *Org. Lett. 3*, 2801−2804.
37. Adam, W., Arnold, M. A., Grune, M., Nau, W. M., Pischel, U., and Saha-Möller, C. R. (2002) *Org. Lett. 4*, 537−540.
38. Niles, J. C., Wishnok, J. S., and Tannenbaum, S. R. (2001) *Org. Lett. 3*, 963−966.
39. Adam, W., Arnold, M. A., Nau, W. M., Pischel, U., and Saha-Möller, C. R. (2002) *J. Am. Chem. Soc. 124*, 3893−3904.
40. Suzuki, T., Masuda, M., Friesen, M. D., and Ohshima, H. (2001) *Chem. Res. Toxicol. 14*, 1163−1169.
41. Suzuki, T., and Ohshima, H. (2002) *FEBS Lett. 516*, 67−70.
42. Sugden, K. D., Campo, C. K., and Martin, B. D. (2001) *Chem. Res. Toxicol. 14*, 1315−1322.
43. Suzuki, T., Friesen, M. D., and Ohshima, H. (2003) *Chem. Res. Toxicol. 16*, 382−389.
44. Luo, W., Muller, J. G., Rachlin, E. M., and Burrows, C. J. (2001) *Chem. Res. Toxicol. 14*, 927−938.
45. Leipold, M. D., Muller, J. G., Burrows, C. J., and David, S. S. (2000) *Biochemistry 39*, 14984−14992.
46. Kornyushyna, O., Berges, A. M., Muller, J. G., and Burrows, C. J. (2002) *Biochemistry 41*, 15304−15314.
47. Duarte, V., Muller, J. G., and Burrows, C. J. (1999) *Nucleic Acids Res. 27*, 496−502.
48. Muller, J. G., Duarte, V., Hickerson, R. P., and Burrows, C. J. (1998) *Nucleic Acids Res. 26*, 2247−2249.
49. Morales, J. C., and Kool, E. T. (1999) *J. Am. Chem. Soc. 121*, 2323−2324.
50. Henderson, P. T., Delaney, J. C., Muller, J. G., Neeley, W. L., Tannenbaum, S. R., Burrows, C. J., and Essigmann, J. M. (2003) *Biochemistry 42*, 9257−9262.
51. Cheng, K. C., Cahill, D. S., Kasai, H., Nishimura, S., and Loeb, L. A. (1992) *J. Biol. Chem. 267*, 166−172.
52. Shibutani, S., Takeshita, M., and Grollman, A. P. (1991) *Nature 349*, 431−434.
53. Lowe, L. G., and Guengerich, F. P. (1996) *Biochemistry 35*, 9840−9849.
54. Cowart, M., Gibson, K. J., Allen, D. J., and Benkovic, S. J. (1989) *Biochemistry 28*, 1975−1983.
55. Wagner, J., and Fuchs, R. P. P. (1997) *Chem. Res. Toxicol. 10*, 568−574.
56. Johnson, S. J., Taylor, J. S., and Beese, L. S. (2003) *Proc. Natl. Acad. Sci. U.S.A. 100*, 3895−3900.
57. Fersht, A. R., Knill-Jones, J. W., and Twui, W. C. (1982) *J. Mol. Biol. 156*, 37−51.

BI0350755