Nitrogen mustards (NMs) are DNA-alkylating compounds that represent the earliest anticancer drugs. However, clinical use of NMs is limited because of their own mutagenic properties. The mechanisms of NM-induced mutagenesis remain uncertain. The major product of DNA alkylation by NMs is a cationic NM-N7-dG adduct that can yield the imidazole ring-fragmented lesion, N5-NM-substituted formamidopyrimidine (NM-Fapy-dG). Characterization of this adduct is complicated because it adopts different conformations, including both a canonical β- and an unnatural α-anomeric configuration. Although formation of NM-Fapy-dG in cellular DNA has been demonstrated, its potential role in NM-induced mutagenesis is unknown. Here, we created site-specifically modified single-stranded vectors for replication in primate (COS7) or Escherichia coli cells. In COS7 cells, NM-Fapy-dG caused targeted mutations, predominantly G → T transversions, with overall frequencies of ~11–12%. These frequencies were ~2-fold higher than that induced by 8-oxo-dG adduct. Replication in E. coli was essentially error-free. To elucidate the mechanisms of bypass of NM-Fapy-dG, we performed replication assays in vitro with a high-fidelity DNA polymerase, Saccharomyces cerevisiae polymerase (pol) δ. It was found that pol δ could catalyze high-fidelity synthesis past NM-Fapy-dG, but only on a template subpopulation, presumably containing the α-anomeric adduct. Consistent with the low mutagenic potential of the β-anomer in vitro, the mutation frequency was significantly reduced when conditions for vector preparation were modified to favor this configuration. Collectively, these data implicate the α-anomer as a major contributor to NM-Fapy-dG-induced mutagenesis in primate cells.

This work was supported by National Institutes of Health, NCI Grants T32 CA106195, P01 CA160032, and P30 CA068485 and NIEHS Grants P30 ES000267 and T32 ES0007060. The authors declare that they have no conflicts of interest with the contents of this article. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This article contains supplemental Figs. S1 and S2.

1 To whom correspondence should be addressed. Tel.: 503-494-9957; Fax: 503-494-6831; E-mail: lloydst@ohsu.edu.
2 The abbreviations used are: NM, nitrogen mustard; Fapy, formamidopyrimidine; NM-Fapy-dG, N5-NM-substituted formamidopyrimidine dG; Me-Fapy-dG, Nβ-methyl substituted formamidopyrimidine dG; 8-oxo-dG, 8-oxo-7,8-dihydro-2′-dG; Endo IV, Endonuclease IV; pol, polymerase.

Nitrogen mustards (NMs) are bifunctional alkylating compounds derived from bis(2-chloroethyl)amine (Fig. 1A). Presenting the earliest of the anticancer drugs, these agents, particularly cyclophosphamide and melphan, are still employed for the treatment of hematologic malignancies and various solid tumors (1, 2) and also as components of conditioning regimens for hematopoietic stem cell transplantation (3). The therapeutic effect of NMs is attributed to their ability to form DNA interstrand cross-links and thus, interfere with critical biological processes that require separation of DNA strands (4). However, the clinical use of NMs is limited because they are themselves highly carcinogenic, with the incidence of secondary malignancies, mostly acute myeloid leukemia, being significant in patients treated with these agents (5, 6). Although the carcinogenic potential of NMs has been recognized for decades, DNA lesions that are primarily responsible for NM-induced mutations remain undetermined.

Reaction of NMs with DNA yields a multitude of products. The most common site of alkylation is N7-dG (7–13) with the cationic monoadduct (N7-NM-dG) being the major initial species (Fig. 1B). Other known sites of modification include N1- and N3-dA (11, 12, 14, 15), N3-dC (15, 16), and O6-dG (17). Furthermore, the alkylated base can react with another DNA base (7, 10–13, 15) or with a cysteine thiol within proteins (18, 19) to form DNA–DNA or DNA–protein cross-links, respectively. Despite the recognized clinical significance, DNA interstrand cross-links represent only a small fraction of the NM-induced DNA lesions (13).

In addition to cross-linking, the cationic N7-NM-dG adduct can undergo a secondary reaction that proceeds via the addition of a hydroxide to the C8 position with a subsequent imidazole ring-opening, yielding the N5-NM-substituted formamidopyrimidine adduct (NM-Fapy-dG) (Fig. 1C). Although the possibility of such ring fragmentation had been reported in the early 1980s (8–10, 20), its biological relevance had remained uncertain until 2015 when Rizzo, Turesky, and colleagues (21) established an ion-trap multistage scan MS method to simultaneously measure various dG adducts of bis(2-chloroethyl)ethylenimine (Fig. 1A). Following treatment of calf thymus DNA in vitro, the NM-Fapy-dG adducts were formed at a level ~20% that of the N7-NM-dG adducts. When these studies were extended to a tumor cell line that was exposed to bis(2-chloro-
ethyl)ethylamine, NM-Fapy-dG was detected at the level of 180 adducts per 10^7 nucleosides, representing ~13% of all adducts measured. DNA cross-links were also detected in the products from both of the experimental systems described above. Interestingly, a fraction of these cross-links contained dGs that underwent ring-opening. These data demonstrated that N7-NM-dG monoadducts or cross-links are routinely converted to the corresponding Fapy-dG species and thus, contribution of these ring-fragmented lesions to NM-induced mutagenesis should be considered.

To facilitate investigations on the structure and function of the NM-Fapy-dG adducts, the Rizzo group developed chemistry for the site-specific synthesis of oligodeoxynucleotides containing bis(2-chloroethyl)ethylamine-derived Fapy-dG (Fig. 1A) and employed Escherichia coli endonuclease IV (Endo IV), an enzyme that selectively incises double-stranded DNA at β-nucleotides (22, 23) to probe the anomeric configuration of NM-Fapy-dG (24). These data showed that similar to unsubstituted (25) or methyl-substituted Fapy-dG (Me-Fapy-dG) (26), NM-Fapy-dG exists as a mixture of interconverting β- and α-anomeric forms (Fig. 1, C and D), with the former being predominant in duplex DNA.

Although modifications at the N7 position of dG are generally considered benign with regard to mutagenicity, the imidazole ring-fragmented dG adducts are known to induce significant levels of mutations in primate cells. When site-specifically modified vectors containing an unsubstituted Fapy-dG were replicated in either simian kidney COS7 (27, 28) or human embryonic kidney 293T (29) cells, the adduct caused mutations, predominantly G → T transversions and G → A transitions, at frequencies of ~8 to 30%. Me-Fapy-dG also manifested a strong mutator phenotype in COS7 cells, inducing mutations at frequencies of ~9 to 21%, with G → T transversions being most common (30). In contrast to this error-prone bypass observed in primate cells, both Fapy-dG (28, 31) and Me-Fapy-dG (32) were essentially non-mutagenic in E. coli.

The present study addressed the mutagenic potential and mechanisms of replication bypass of NM-Fapy-dG. Single-stranded vectors containing a site-specific NM-Fapy-dG were replicated in COS7 and E. coli cells, and frequencies and spectra of mutations were assessed. Furthermore, the effect of NM-Fapy-dG on DNA synthesis was investigated using Saccharomyces cerevisiae pol δ, a high-fidelity DNA polymerase that represents one of the two major nuclear polymerases in eukaryotic cells. The 8-oxo-dG adduct was tested under identical conditions as a reference lesion. In addition, differential effects on replication were evaluated for the two anomeric forms of NM-Fapy-dG.

**Results**

**The general design of mutagenesis assays**

Prior investigations have demonstrated that in mammalian cells, replication of single-stranded shuttle vectors containing either Fapy-dG (27–29) or Me-Fapy-dG (30) could result in targeted mutations in up to ~30% replication events, with frequencies and spectra being modulated by the local sequence context. In the case of Me-Fapy-dG, a subset of misincorporations was hypothesized to be directed by the 5′-neighboring nucleotide, as the result of primer/template misalignment, fol-
Mutagenesis of nitrogen mustard Fapy-dG

Table 1

Mutations generated following replication of site-specifically modified vectors in COS7 cells

The data in parentheses are the mean percentages of mutations with corresponding standard errors calculated from three or four independent transfections of COS cells. Because of variabilities, the mean percentages are not exactly equal to percentages calculated from combined results. The standard errors for the mean frequencies < 1% are not reported. — indicates no mutations were detected.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>8-Oxo-dG</td>
<td>5’-CXG-3’</td>
<td>344</td>
<td>12’ (3.7 ± 1.3)</td>
<td>—</td>
<td>—</td>
<td>2’ (0.6)</td>
<td>2’ (0.6)</td>
</tr>
<tr>
<td>NM-Fapy-dG</td>
<td>5’-TXG-3’</td>
<td>284</td>
<td>14’ (3.0 ± 0.4)</td>
<td>1 (0.3)</td>
<td>4 (1.2 ± 0.1)</td>
<td>2’ (0.6)</td>
<td>2’ (0.6)</td>
</tr>
<tr>
<td>NM-Fapy-dG</td>
<td>5’-TXG-3’</td>
<td>355</td>
<td>29 (8.3 ± 0.8)</td>
<td>6 (1.8 ± 1.0)</td>
<td>1 (0.3)</td>
<td>5 (1.5 ± 0.7)</td>
<td>—</td>
</tr>
<tr>
<td>NM-Fapy-dG</td>
<td>5’-TXG-3’</td>
<td>319</td>
<td>29 (8.8 ± 1.2)</td>
<td>1 (0.3)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

* Including GG to TT.
* Including deletion CG.
* Including CG to TT.
* Including GCXGG to TTGG.
* Including 5’-C to G.
* Including 3’-G to T.

followed by primer realignment and extension. Considering the structural similarity of various imidazole ring-fragmented dG adducts, the current study investigated the mutagenic properties of NM-Fapy-dG and in particular, addressed the potential effect on mutagenesis of the 5’-nucleotide. The adduct in either a 5’-CXG-3’ or 5’-TXG-3’ sequence context was incorporated into a single-stranded pSBL vector as illustrated in supplemental Fig. S1. Identically designed vectors were also engineered to contain 8-oxo-dG. This adduct was selected to serve as a reference lesion because the NM-Fapy-dG and 8-oxo-dG adducts have a certain structural similarity and in particular, both retain the canonical Watson-Crick hydrogen-bonding face of dG.

NM-Fapy-dG-induced mutagenesis in COS7 cells

The pSBL vectors containing 8-oxo-dG or NM-Fapy-dG in either 5’-CXG-3’ or 5’-TXG-3’ sequence context were replicated in COS7 cells. Following isolation of the progeny DNAs, individual clones were obtained by transforming E. coli DH5α cells and analyzed for mutations using a combination of differential hybridization technique and Sanger sequencing. Analyses of control vectors that contained dG at the target site were performed in parallel experiments. These results demonstrated essentially error-free replication of non-damaged vectors as we have recently reported (28). Although our experimental approach was not designed to evaluate the bypass efficiency, no correlation was noticed between the yield of bacterial transformants and the type of DNA replicated. Thus, the NM-Fapy-dG lesion is unlikely to represent a very strong block for the mammalian replication machinery. The data on the mutagenic properties of 8-oxo-dG and NM-Fapy-dG were collected from a minimum of three independent transfections of COS7 cells and are shown in Table 1.

Replication of 8-oxo-dG-containing DNAs resulted in a simple spectrum of mutations, with G → T transversions being most common in both sequence contexts (~3.7% in 5’-CXG-3’ and ~5.0% in 5’-TXG-3’). In addition, low levels (~1.2%) of various deletions were observed only in the 5’-CXG-3’ sequence. The differences in frequencies of both G → T transversions and total targeted mutations were not statistically significant between the two sequence contexts. These results generally agree with data previously reported (27, 29, 33) and are consistent with established mechanisms of replication past 8-oxo-dG (34–43), in which the adducted base remains in the canonical anti conformation opposite an incoming dC (Fig. 1E), but assumes the syn conformation opposite dA (Fig. 1F) to facilitate generation of G → T transversions. It cannot be excluded, however, that a subset of these transversions, occurring in the 5’-TXG-3’ sequence context could be formed via a primer/template misalignment.

Relative to 8-oxo-dG, the NM-Fapy-dG adduct caused targeted mutations at ~2-fold higher frequencies (~11.6 versus ~4.4%, p = 0.019, in 5’-CXG-3’ and ~10.6 versus ~5.3%, p = 0.032, in 5’-TXG-3’). G → T transversions were most common in both sequence contexts (~8.3% in 5’-CXG-3’ and ~8.8% in 5’-TXG-3’), although other base substitutions also occurred. The spectra of base substitutions were minimally affected by the local sequence context. Frequencies of G → C transversions in 5’-CXG-3’ were slightly higher than in 5’-TXG-3’ (~1.8 and ~0.3%, respectively), but the difference was not statistically significant. Deletions were detected exclusively in the 5’-CXG-3’ context and were at near background levels. Thus, whereas the structure of the correctly matched pair can be easily envisioned for NM-Fapy-dG (Fig. 1G), the nature of replication intermediates leading to mutagenesis is less apparent. Conventional primer-template misalignment models can only explain a small fraction of NM-Fapy-dG-induced mutations. Specifically with regard to the major mutation, a G → T transversion, the recognized model of mispair involving the Hoogsteen face in 8-oxo-dG (Fig. 1F) cannot be applied to NM-Fapy-dG, because this face is affected by the NM modification (Fig. 1C).

NM-Fapy-dG-induced mutagenesis in E. coli

The pSBL vectors containing NM-Fapy-dG in the 5’-CXG-3’ sequence were replicated in wild-type E. coli (ZK126, W3110ΔlacU169tna-2) and the isogenic mutant lacking all three DNA damage-inducible ( SOS) DNA polymerases, pol II, pol IV, and pol V (SF2018, ZK126 polB::Spc′ dinB::Kan′ umuDC::Cam′). After obtaining ampicillin-resistant clones, frequencies and spectra of mutations were assessed by a combination of differential hybridization and Sanger sequencing. The data collected from the two independent vector preparations demonstrated that in both E. coli strains, replication bypass of NM-Fapy-dG was significantly more accurate than in COS7 cells. Specifically, no mutations were detected in 152 clones analyzed for the wild-type, and only two of 175 clones
standard deviations were calculated from three independent experiments. Shown. The mean percentages of the bypass products with corresponding all four dNTPs under running start conditions. A32P-labeled utilized (40, 44, 45).

ficient form, with both versions being extensively character- tion, which were A → A transition at the 3'-dG and deletion of the 5'-dC. The latter observation implies that in E. coli, high-fidelity DNA polymerases, pol I and/or pol III, can bypass the NM-Fapy-dG adduct, predominantly producing the correct replication products.

Replication bypass of the NM-Fapy-dG adduct by DNA polymerase δ

To further elucidate the mechanisms of replication past the NM-Fapy-dG adduct, the ability of pol δ to catalyze either error-free or error-prone bypass of this adduct was assessed. The reason for testing pol δ was that this polymerase represents one of the two major nuclear DNA polymerases in eukaryotic cells and therefore, it encounters a DNA lesion prior to recruitment of any specialized polymerase. S. cerevisiae pol δ was utilized because this enzyme is available as an exonuclease-deficient mutant, in addition to its exonuclease-proficient form, with both versions being extensively characterized (40, 44, 45).

The initial set of reactions was conducted in the presence of all four dNTPs under running start conditions. A 32P-labeled 20-mer oligodeoxynucleotide primer was hybridized with control dG, NM-Fapy-dG, and 8-oxo-dG templates with the lesion in the 5’-CXG-3’ sequence context to initiate DNA synthesis six nucleotides upstream of the modified site (-6 primer) (Fig. 2A). These data demonstrated that both lesions inhibited DNA polymerization, one nucleotide before the adducted site (-1 site) and opposite to it (0 site) (Fig. 2B). NM-Fapy-dG appeared to be a stronger block to replication than 8-oxo-dG. In partic- ular, at 2 nM pol δ concentration, the amount of extension products beyond 8-oxo-dG was approaching the corresponding value calculated for the control dG (~77 versus ~93%). Under identical conditions, a significantly lower fraction of primers (~36%) was extended beyond the NM-Fapy-dG site. Surprisingly, accumulation of the bypass products on NM-Fapy-dG-containing template was minimally changed when the final enzyme concentration was increased 7-fold to 14 nM (~40%). The lack of linear relationships between the enzyme concentration and the product formation was also observed for NM-Fapy-dG that was positioned in the 5’-TXG-3’ sequence context (supplemental Fig. S2). An exonuclease-deficient pol δ was similarly limited in its ability to replicate past NM-Fapy-dG (data not shown). We reasoned that under these experimental conditions, NM-Fapy-dG existed as a mixture of at least two species that differentially affect the ability of pol δ to carry out DNA synthesis. Relevant to this observation, prior analyses demonstrated that NM-Fapy-dG exists as a mixture of unnatural α- and natural β-anomeric configurations (24). Thus, we hypothesized that although NM-Fapy-dG in its β-anomeric form could be engaged in extendable base pairs (Fig. 1, C and G), the α-anomer (Fig. 1D) represented a severe block to pol δ-catalyzed primer extension.

To evaluate the accuracy of nucleotide incorporation by pol δ opposite NM-Fapy-dG, reactions were carried out in the presence of individual dNTPs using a 20-mer −1 primer (Fig. 3A). In the presence of dCTP, primer extension dominated over degradiation on all five DNA templates tested, including control dC and NM-Fapy-dG in the 5’-TXG-3’ and 5’-CXG-3’ sequences, and 8-oxo-dG in the 5’-CXG-3’ sequence (Fig. 3, B and C). In the presence of other dNTPs, the formation of the extension products was minimal or not observed. Consistent with the analyses on pol δ-catalyzed primer extensions in the presence of all four dNTPs (Fig. 2), incorporations of dC were inhibited opposite 8-oxo-dG (~64 versus ~85% control, p = 0.031) and even more notably opposite NM-Fapy-dG (~51 versus ~90% control, p < 0.0001, and ~49 versus ~85% control, p = 0.005, in the 5’-TXG-3’ and 5’-CXG-3’ context, respectively) (Fig. 3C).

Single-nucleotide incorporation experiments were also conducted using exonuclease-deficient pol δ (Fig. 3D). In agreement with published data (40), this enzyme demonstrated a low fidelity of nucleotide incorporation opposite 8-oxo-dG. The primers were extended better in the presence of dATP than dCTP (~78 and ~64%, p = 0.030). This result was in contrast to more selective nucleotide incorporations observed on NM-Fapy-dG-containing templates; in both sequence contexts, higher levels of the extension products were formed in reactions supplemented with dCTP (~49 and ~51% in the 5’-TXG-3’ and 5’-CXG-3’ context, respectively) than any other dNTP (Fig. 3D). The difference between incorporation of dC and the next preferred nucleotide (dA in 5’-TXG-3’ and dG in 5’-CXG-3’) was statistically significant (p = 0.043 and 0.002, respectively). Thus, even in the absence of proofreading activity, incorporation opposite NM-Fapy-dG by pol δ appeared to be accurate.

Further experiments were designed to address the extension step of pol δ-catalyzed bypass. A series of 20-mer 0 primers
were hybridized with the 5′-CXG-3′ templates to place either dC or one of the three mismatched nucleotides opposite the adducted site (Fig. 3E), and primer extensions were performed in the presence of dGTP and dCTP. In the given sequence context, sequential incorporations of these nucleotides would generate 22-mer products, with the initially existing mismatches being potentially corrected by exonucleolytic proofreading with subsequent dC insertion. These data demonstrated that on both control and NM-Fapy-dG templates, the reactions likely followed the above scenario (Fig. 3F). Although 20-mer primers had distinct electrophoretic mobilities depending on the identity of the 3′-terminal nucleotide, the 21- and 22-mer extension products appeared indistinguishable. In contrast, the difference in the product mobilities was obvious on 8-oxo-dG-containing template, specifically when dA was placed opposite the adduct.

The in vitro replication analyses suggested that only the β-anomer of NM-Fapy-dG could be bypassed by pol δ and that the outcome of this reaction was likely to be generally accurate.

---

### Table: Primer Extension Products

<table>
<thead>
<tr>
<th>dNTP</th>
<th>TXG</th>
<th>CXG</th>
</tr>
</thead>
<tbody>
<tr>
<td>dG</td>
<td>A C G T</td>
<td>A C G T</td>
</tr>
<tr>
<td>dG</td>
<td>A C G T</td>
<td>A C G T</td>
</tr>
<tr>
<td>dG</td>
<td>A C G T</td>
<td>A C G T</td>
</tr>
<tr>
<td>dG</td>
<td>A C G T</td>
<td>A C G T</td>
</tr>
<tr>
<td>dG</td>
<td>A C G T</td>
<td>A C G T</td>
</tr>
</tbody>
</table>

Figure 3. Fidelity of replication bypass of NM-Fapy-dG by pol δ. A, structure of the −1 primer/template substrates. B, pol δ-catalyzed single nucleotide incorporations. Reactions were carried out for 30 min at 37 °C in the presence of 2 nM pol δ, 5 nM DNA substrates, and 100 μM of an individual dNTP. The control reaction did not contain dNTPs or pol δ. A representative gel image is shown. C, pol δ-catalyzed nucleotide incorporation and exonucleolytic primer degradation in the presence of dCTP. The data represent analyses of images as exemplified in panel B. The mean percentages of the extension and degradation products with corresponding standard deviations were calculated from three or four independent experiments. D, single nucleotide incorporations catalyzed by exonuclease-deficient pol δ. Reactions were carried out for 30 min at 37 °C in the presence of 2 nM D520V pol δ, 5 nM DNA substrates, and 100 μM of an individual dNTP. The mean percentages of the primer extension products with corresponding standard deviations were calculated from three or four independent experiments. E, structures of the 0 primer/template substrates and potential products of primer extensions in the presence of dCTP and dGTP. F, pol δ-catalyzed primer extensions. Reactions were carried out for 20 min at 37 °C in the presence of 2 nM pol δ, 5 nM DNA substrates, 100 μM dCTP, and 100 μM dGTP. A representative gel image is shown.
In contrast to 8-oxo-dG, NM-Fapy-dG did not show the propensity to form an extendable mispair with dA in pol δ-catalyzed reactions. Considering the presence of a significant fraction of the NM-Fapy-dG species that posed a severe, if not complete block to pol δ-catalyzed replication, we did not find it feasible to assess the bypass efficiency by Michaelis-Menten or other conventional kinetic analyses. However, it is worth emphasizing that under non-saturating conditions, the amounts of products beyond the target site were comparable on all three templates tested (Fig. 2B). In the presence of 0.08 nM pol δ, these numbers on dC-, NM-Fapy-dG-, and 8-oxo-dG-containing templates were, respectively ~10, 6, and 7.5%, and in the presence of 0.4 nM pol δ, respectively, ~38, 11, and 19%. These differences are not dramatic, especially given the consideration that on the NM-Fapy-dG-containing template, a fraction of pol δ was likely to be involved in futile attempts to bypass the blocking form of the adduct.

Mutagenic properties of anomer forms of NM-Fapy-dG in COS7 cells

As discussed earlier, the origin of the major NM-Fapy-dG-induced mutation, the G → T transversion, could not be universally explained by primer/template misalignments. The patterns of nucleotide incorporations by pol δ were also inconsistent with this initially proposed mechanism. Another possibility to consider would be the non-instructive incorporation of dA opposite the lesion, as has been well-described for an abasic site (46 – 49). We hypothesized that NM-Fapy-dG in its unnatural α configuration (Fig. 1D) could mimic the abasic site structure in the polymerase active site, thus facilitating the formation of the G → T transversions. To address this possibility, we employed a strategy recently developed by us (28) for increasing the fraction of the β-anomeric species in a mixed population of DNA vectors. Following ligation of NM-Fapy-dG-containing vectors, aliquots of these were stored for several days at 4 °C with scaffold oligodeoxynucleotides being intact, thus preserving the double-stranded environment. Based on data from a prior study (24), we anticipated that the equilibrium would shift toward the β-form under these conditions. Further, the vectors were incubated with Endo IV to eliminate the residual α-anomers and following degradation of scaffold oligodeoxynucleotides, immediately transfected into COS7 cells to minimize the possibility of re-equilibration to the initial α/β-anomeric ratio. Because no significant sequence-dependent differences were found with regard to mutagenic outcome, this analysis was limited to the 5′-CGX-3′ sequence context. The data collected from three independent transfections revealed a significant effect of the above manipulations (Fig. 4), with both the overall frequency of targeted mutations and the frequency of G → T transversions decreased more than 2-fold (~4.9 versus ~11.6%, p = 0.014, and ~3.8 versus 8.3%, p = 0.005, respectively). These results strongly suggest that the α-anomer significantly contributed to NM-Fapy-dG-induced mutagenesis in COS7 cells.

Discussion

The present study was designed to evaluate the contribution of the NM-Fapy-dG adduct to mutagenesis associated with the clinical use of NMs. The data revealed that in primate cells, NM-Fapy-dG induced mutations at frequencies of ~11–12%. These frequencies were higher than that caused by 8-oxo-dG in the identical sequence contexts and were comparable with frequencies of mutations reported for the related Fapy-dG (27–29) and Me-Fapy-dG (30) adducts. Relative to 8-oxo-dG, NM-Fapy-dG manifested more complex spectra of mutations, although G → T transversions were most common in both cases. Thus, NM-Fapy-dG has a potential to contribute to NM-induced genomic instability.

The mechanism of mutagenic bypass of NM-Fapy-dG was addressed by altering the identity of the 5′-neighboring nucleotide, with the adduct being placed in either 5′-CGX-3′ or 5′-TXG-3′ sequence. The effect of sequence context on spectra of NM-Fapy-dG-induced base substitutions was not apparent, and deletions were rare. Thus, we conclude that the primer/template-misaligned replication intermediates were not commonly formed in these specific sequences.

Following enrichment of NM-Fapy-dG with the β-anomeric form, frequencies of targeted mutations, particularly G → T transversions, were significantly decreased. This implicates the α-anomer as a major contributor to NM-Fapy-dG-induced mutagenesis. Similar results have been recently reported by us for unsubstituted Fapy-dG, with the major effects being observed on frequencies of both G → A transitions and G → T transversions (28). It is important for understanding the biology of various Fapy-dG adducts that they predominantly exist as the β-anomers in double-stranded DNA (24–26). However, all currently available data on mutagenesis of Fapy-dG (27–29), Me-Fapy-dG (30), and NM-Fapy-dG were generated using single-stranded vectors that are likely to contain a significant portion of the α-anomeric species. Thus, the mutagenic potential of these adducts could have been overestimated in these studies, and consideration

![Figure 4. Modulation of mutagenic properties of NM-Fapy-dG in COS7 cells following enrichment with the β-anomer. The lighted vectors containing NM-Fapy-dG in the 5′-CGX-3′ sequence context were incubated in the presence of the scaffolding oligodeoxynucleotides and treated with Endo IV. Immediately following treatment, scaffolding oligodeoxynucleotides were degraded, and vectors transfected into COS7 cells. The total numbers of clones analyzed were 355 for untreated vectors and 262 for Endo IV-treated vectors. The mean frequencies of mutations with corresponding standard errors were calculated from a minimum of three independent experiments using KaleidaGraph 4.1 software (Synergy Software). The p values were calculated using Student’s t test. The asterisk indicates significant difference (p < 0.05) between untreated versus treated conditions.](image-url)
Mutagenesis of nitrogen mustard Fapy-dG

should be given to an alternative experimental system that allows the adduct to be examined in a double-stranded DNA environment.

The hypothesis that the \( \beta \) -anomeric NM-Fapy-dG is not a highly cytotoxic or miscoding lesion agrees with the data of the in vitro replication assays using pol \( \delta \). This polymerase could replicate past the \( \beta \) -anomer and based on the results of single nucleotide incorporations and primer extensions beyond the adducted site, the outcome of overall bypass reaction was accurate. This was in contrast to the propensity of pol \( \delta \) to incorporate a dA opposite the 8-oxo-dG adduct and efficiently extend from this mismatched primer terminus. Thus, even though 8-oxo-dG and NM-Fapy-dG both predominantly induce G \( \rightarrow \) T transversions in primate cells, the underlying molecular mechanisms must be different. It is generally accepted that high-fidelity DNA polymerases, such as pol \( \delta \), are involved in mutagenic bypass of 8-oxo-dG because of their abilities to form and extend a syn-8-oxo-dG:dA mispair (Fig. 1F) (34–38, 40, 41). Based on results of our study, we hypothesize that NM-Fapy-dG in its \( \beta \) -anomeric configuration will be bypassed by these polymerases in a mostly error-free manner. The \( \alpha \) -anomer, on the other hand, is expected to be a much stronger block to replication, thereby necessitating the recruitment of low-fidelity DNA polymerases specialized in translesion synthesis (50). A subset of these must be able to carry out bypass of \( \alpha \) -anomeric NM-Fapy-dG in primate cells, but at the expense of mutations. We also hypothesize that in E. coli, where the arsenal of DNA polymerases is more limited than in mammalian cells (50), bypass of the \( \alpha \) -anomer is extremely poor or impossible. This would be consistent with the known strong potential of the \( \alpha \) -dG lesion to block replication in E. coli (51) and could explain a near background level of mutations in products of replication of NM-Fapy-dG-containing vectors in both strains tested in the present study.

Although the identities of DNA polymerases involved in replication past \( \alpha \) -anomeric NM-Fapy-dG are currently unknown, several candidates can be considered based on the data of investigations on related adducts. Depletion of pol \( \lambda \) from human embryonic kidney 293T cells resulted in a significant reduction of targeted G \( \rightarrow \) T transversions in progenies of Fapy-dG-containing vectors (29). In addition, insertions of dA opposite Me-Fapy-dG were observed in bypass reactions catalyzed by human pol \( \eta \) and pol \( \kappa \) (52). However, both the above studies utilized DNAs with the adducts being placed in the sequence contexts that could facilitate insertions of dA via a primer/template misalignment (5’-TXN-3’ (29) and 5’-THT-3’ (52)). Furthermore, if the mechanism of bypass of the \( \alpha \) -anomer is indeed similar to that of abasic sites, Rev1 and pol \( \zeta \) should be considered, because these polymerases were vital for replication of abasic site-containing vectors in 293T cells (49). Regarding specifically the \( \alpha \) -anomeric lesions, recent analyses revealed that human pol \( \eta \), pol \( \kappa \), and pol \( \iota \) could incorporate nucleotides opposite \( \alpha \)-dG, with pol \( \eta \) manifesting the most accurate and efficient bypass capability (53). The two-subunit yeast pol \( \zeta \) (Rev3-Rev7) appeared to be completely incompetent on \( \alpha \)-dG-containing templates at both incorporation and extension steps (53). Interestingly, our preliminary data demonstrated the ability of four-subunit yeast pol \( \zeta \) (Rev3-Rev7-Pol31-Pol32) (54) to synthesize DNA past NM-Fapy-dG, with the amount of bypass products beyond the lesion site significantly exceeding that observed for pol \( \delta \). The latter result suggested that this enzyme could either bypass the \( \alpha \) form of the adduct or upon binding, modulate the anomeric ratio. Future structural and biochemical analyses will provide insights about mechanisms of NM-Fapy-dG-induced mutagenesis and in particular, address possible roles of individual DNA polymerases in mutagenic bypass of the \( \alpha \) -anomer.

Experimental procedures

Oligodeoxynucleotides

Site-specifically modified oligodeoxynucleotides 5’-GCT-AGCXTGTCCT-3’ and 5’-GCTAGTXGTCCT-3’, where X is NM-Fapy-dG, were synthesized and purified by HPLC as reported previously (24). The corresponding oligodeoxynucleotides containing either dG or 8-oxo-dG at the X site and all other oligodeoxynucleotides were synthesized and purified by Integrated DNA Technologies.

Construction of the pSBL vector

The pSBL vector was created based on the pSB shuttle vector (55) by insertion of a stem-loop structure into the unique BamHI site (supplemental Fig. S1). The details of the construct preparation are described in the legend to supplemental Fig. S1. The f1 origin-dependent production of single-stranded pSBL was performed using DH12S E. coli (ElectroMAX, Invitrogen) following a standard procedure.

Construction of site-specifically modified vectors and mutagenesis assay

Insertion of site-specifically modified 12-mer oligodeoxynucleotides into EcoRV site in a stem region of single-stranded pSBL using uracil-containing scaffold oligodeoxynucleotides (supplemental Fig. S1), replication of vectors in host cells, and analyses of the progeny vectors for mutations by differential hybridization were performed as in our previous studies (28, 30, 56, 57). To enrich the NM-Fapy-dG-containing vectors with the \( \beta \) -anomeric adduct, the ligated vectors were purified by three sequential washes with TE buffer (10 mM Tris-HCl (pH 7.4) and 1 mM EDTA) using Amicon Ultra 100K centrifugal filters (Merck Millipore, Ltd.), and stored for a minimum of 4 days at 4 °C with scaffold oligodeoxynucleotides being intact. The vectors (1 pmol) were incubated with 20 units of E. coli endonuclease IV (New England Biolabs) at 37 °C for 1 h in a buffer provided by supplier (50 mM Tris-HCl (pH 7.9), 100 mM NaCl, 10 mM MgCl₂, 1 mM DTT), and subsequently, with 3 units of T4 DNA polymerase and 5 units of uracil DNA glycosylase (both from New England Biolabs) at 37 °C for 30 min to degrade scaffold oligodeoxynucleotides. Vectors were purified by Bio-Spin P-6 columns (Bio-Rad), and immediately used for transfection of COS7 cells.

The hybridization probes were \( ^{32} \)P-labeled oligodeoxynucleotides matching the clones that had no mutations in the inserted sequences (5’-ATGCTAGCGGGTTCCATCG-3’ or 5’-ATGCTAGTGGGTCCATCG-3’ for the 5’-CXC-3’ or 5’-TXG-3’ sequence context, respectively) or acquired G \( \rightarrow \) T.
transversions at the target site (5’-ATGCTAGCTGGTCCATCG-3’ or 5’-ATGCTAGTTGGTCCATCG-3’ for the 5’-CGX-3’ or 5’-TXG-3’ sequence context, respectively). To identify additional clones that contained the inserted sequences, a probe was used that encompassed the junction between the vector and insert sequences. The identities of mutations in the latter clones were determined by Sanger sequencing. The data collected from independent transfections were subjected to the Student’s t test statistical analyses using KaleidaGraph software (version 4.1, Synergy Software).

**Preparation of DNA substrates for in vitro replication assays**

In vitro replication assays used 59-mer oligodeoxynucleotides as templates that were constructed according to the published procedure (56). Briefly, site-specifically modified 12-mer oligodeoxynucleotides (5’-GCTAGCGTCC-3’ or 5’-GCTAGTXGGTCC-3’) were ligated with the flanking 5’ 12-mer (5’-ACGCGCCAGTGGAG-3’) and 3’ 35-mer (5’-GATCCCGCTACTAGTTGCTTCTCGACGTCGTTTTAC-3’) fragments, 500 pmol each, in the presence of equimolar amounts of scaffold oligodeoxynucleotides (5’-GAAGCAGCTAGCCGGATCGGGCATCGCCGGCTAGCTCCTACTGGCCGTGTCTTTTAC-3’ or 5’-GAAGCAACTAGTACGGGATC-3’). These sequences, a probe was used that encompassed the junction between the vector and insert sequences. The identities of mutations in the latter clones were determined by Sanger sequencing. The data collected from independent transfections were subjected to the Student’s t test statistical analyses using KaleidaGraph software (version 4.1, Synergy Software).

**In vitro replication assay**

Three-subunit S. cerevisiae pol δ (Pol3–Pol31–Pol32) and its exonuclease-deficient D520V variant were purified as reported (44, 45). Polymerase bypass reactions were conducted at 37 °C in a buffer composed of 25 mM Tris-HCl (pH 7.5), 10 mM NaCl, 8 mM MgCl₂, 10% glycerol, 100 μg/ml of BSA, and 5 mM DTT. The reactions contained 5 mM DNA substrate and 100 μM dNTPs. Concentrations of pol δ and incubation times varied and are specified in the figures or figure legends. The reaction products were resolved by electrophoresis in a 15% denaturing polyacrylamide gel containing 8 M urea in Tris borate-EDTA buffer and visualized using a Storm PhosphorImager (GE Healthcare) or a Personal Molecular Imager (Bio-Rad). The gel images were analyzed by Image Quant 5.2 software (Molecular Dynamics) or the Personal Molecular Imager built-in software. The mean percentages of bypass and exonucleolytic degradation products with corresponding standard deviations were calculated from a minimum of three independent experiments using KaleidaGraph 4.1 software (Synergy Software). The p values were calculated using Student’s t test.

**References**


**Mutagenesis of nitrogen mustard Fapy-dG**

Author contributions—All authors contributed to design of the study. C. J. R. provided site-specifically modified oligodeoxynucleotides, I. G. M. conducted experiments, analyzed the data, and prepared the manuscript draft with significant intellectual input from C. J. R. and R. S. L.

Acknowledgments—We thank Drs. Louise and Satya Prakash (University of Texas Medical Branch) for a kind gift of the pSB vector, Dr. Peter M. Burgers (Washington University, St. Louis) for providing the yeast pol δ and pol ζ, and Dr. Steven E. Finkel (University of Southern California) for sharing the ZK126 and SF2018 E. coli strains. We also thank Tracy Johnson Salyard and Dr. Plamen P. Christov for synthesis of NM-Fapy-dG-modified oligodeoxynucleotides and Dr. Yan Sha for assistance in performing the mutagenesis assays.
acid (chlorambucil) with 2′-deoxyctydine, 2′-deoxy-5-methylctydine, and thymidine. Chem. Res. Toxicol. 17, 383–391