Published in final edited form as:

Chem Res Toxicol. 2018 September 17; 31(9): 924–935. doi:10.1021/acs.chemrestox.8b00135.

Configurational and Conformational Equilibria of *N*⁶-(2-Deoxy-*D*-erythro-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-5-*N*-methylformamidopyrimidine (MeFapy-dG) Lesion in DNA

Stephanie N. Bamberger, Chanchal K. Malik, Markus W. Voehler, Summer K. Brown, Hope Pan, Tracy L. Johnson-Salyard, Carmelo J. Rizzo, and Michael P. Stone*

Department of Chemistry, Vanderbilt University Center for Structural Biology, Vanderbilt Center in Molecular Toxicology, and the Vanderbilt-Ingram Cancer Center, Vanderbilt University, Nashville, Tennessee 37235, United States

Abstract

The most common lesion in DNA occurring due to clinical treatment with Temozolomide or cellular exposures to other methylating agents is 7-methylguanine (N7-Me-dG). It can undergo a secondary reaction to form *N*⁶-(2-deoxy-*D*-erythro-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-5-*N*-methylformamidopyrimidine (MeFapy-dG). MeFapy-dG undergoes epimerization in DNA to produce either α or β deoxyribose anomers. Additionally, conformational rotation around the formyl bond, C5–*N*⁵ bond, and glycosidic bond may occur. To characterize and quantitate the mixture of these isomers in DNA, a ¹³C-MeFapy-dG lesion, in which the CH₃ group of the MeFapy-dG was isotopically labeled, was incorporated into the trimer 5'-TXT-3' and the dodecamer 5'-CATXATGACGCT-3' (X = ¹³C-MeFapy-dG). NMR spectroscopy of both the trimer and dodecamer revealed that the MeFapy-dG lesion exists in single strand DNA as ten configurationally and conformationally discrete species, eight of which may be unequivocally assigned. In the duplex dodecamer, the MeFapy-dG lesion exists as six configurationally and conformationally discrete species. Analyses of NMR data in the single strand trimer confirm that for each deoxyribose anomer, atropisomerism occurs around the C5–*N*⁵ bond to produce *R*_a and *S*_a atropisomers. Each atropisomer exhibits geometrical isomerism about the formyl bond yielding *E* and *Z* conformations. ¹H NMR experiments allow the relative abundances of the species to be determined. For the single strand trimer, the α and β anomers exist in a 3:7 ratio, favoring the β anomer. For the β anomer, with respect to the C5–*N*⁵ bond, the *R*_a and *S*_a atropisomers are equally populated. However, the *Z* geometrical isomer of the formyl moiety is preferred. For the α anomer, the *E*–*S*_a isomer is present at 12%, whereas all other isomers are present at 5–7%. DNA processing enzymes may differentially recognize different isomers of the MeFapy-dG lesion. Moreover, DNA sequence-specific differences in the populations of configurational and conformational species may modulate biological responses to the MeFapy-dG lesion.

Graphical Abstract

*Corresponding Author michael.p.stone@vanderbilt.edu. Phone 615-322-2589.

Notes

The authors declare no competing financial interest.

INTRODUCTION

The development of temozolomide (TMZ)¹ treatment adjuvant with radiotherapy has played a significant role in the clinical management of malignant gliomas.^{2–9} TMZ is a S_N1-type DNA methylating agent that can pass through the blood–brain barrier.^{1,10} Upon hydrolysis, it forms a methyldiazonium ion capable of alkylating DNA at the N7-dG, N3-dA, and O⁶-dG positions, and also the N1-dA and N3-dC positions, particularly in single strand DNA.¹¹ The N7-Me-dG lesion accounts for 70% of the DNA adducts,^{12–18} with N3-dA and O⁶-dG adducts accounting for most of the remaining DNA damage.¹⁰ DNA is also exposed to methylating agents from other sources^{19–21} including cigarette smoke,^{22–25} food,^{26–28} environmental exposures,^{29,30} workplace chemicals,^{31,32} and endogenous alkylating agents.³³ DNA methylation is cytotoxic and it is also genotoxic. TMZ treatment has been correlated with the occurrence of secondary hematopoietic malignancies.^{34–36}

The major N7-Me-dG lesion is not bulky and does not interfere with Watson–Crick base pairing. However, it carries a positive charge on the guanine imidazole ring and may undergo depurination to an apurinic (AP) site, which can be repaired by AP endonucleases,^{37–39} or undergo base-catalyzed opening to produce N⁶-(2-deoxy-D-*erythro*-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-5-*N*-methylformamidopyrimidine (MeFapy-dG) (Scheme 1).^{18,40,41} The MeFapy-dG lesion was identified in hepatic cells of rats treated with *N,N*-dimethylnitrosamine, 1,2 dimethylhydrazine, and *N*-methylnitrosourea,⁴² and in bladder epithelial cells of rats treated with *N*-methylnitrosourea,⁴³ although in a different study it was not detected.⁴⁴ Kadlubar et al.⁴³ treated rats with MNU, and sacrificed animals at various time points following exposure. In the initial 2 h after dosing, bladder epithelial cells exhibited levels of N7-Me-dG almost 16-fold higher than levels of MeFapy-dG. The latter was formed at about half the rate as the formation of the O⁶-Me-dG adduct.⁴⁵ However, at 21 d MeFapy-dG was the major lesion detected in cellular DNA, over seven-fold more prominent than the O⁶-Me-dG adduct, while levels of N7-Me-dG dropped to only 0.65% of the initial value. These data suggested that while MeFapy-dG forms slowly in cellular DNA as compared to the O⁶-Me-dG adduct,⁴⁵ it is persistent. Unlike the N7-Me-dG lesion, it inhibits DNA synthesis.^{46–48} MeFapy-dG is mutagenic in mammalian cells⁴⁹ and is bypassed by Y-family polymerases in an error-prone manner.^{50,51} The chemistry and biology of the MeFapy-dG lesion is complex. Anomerization, the acid catalyzed ring opening of the deoxyribose ring followed by closing to form either the natural β anomer or unnatural α anomer, occurs.^{18,52–58} Conformational isomerization also arises from rotation around the formyl, C⁵–N⁵, and glycosidic bonds.^{40,42,43,52,58–60} These different configurational and conformational isomers are consistent with the observation that the

MeFapy-dG lesion exhibits poor chromatographic properties,⁶¹ making quantitative mass spectrometry approaches toward its quantitation under physiological conditions in cellular DNA challenging.⁶² NMR based approaches have potential to facilitate the quantitation of the types and amounts of specific configurational and conformational isomers of the MeFapy-dG lesion in DNA, as they allow sequence-specific elucidation of the identities of the MeFapy-dG isomers as well as their relative amounts.^{58,60} NMR studies by Guengerich and Humphreys⁶⁰ utilized ¹⁵N isotope labeling, which established that MeFapy-dG existed as the *N*⁵-formyl as opposed to the *N*⁶-formyl regioisomer.

Herein, a ¹³C-MeFapy-dG lesion, in which the CH₃ group of the MeFapy-dG was isotopically labeled, was incorporated into the trimer 5'-TXT-3', and the dodecamer 5'-CATXA-TGACGCT-3' (X = ¹³C-MeFapy-dG).⁶¹ NMR reveals that in the 5'-TXT-3' trimer the MeFapy-dG lesion exists as ten spectroscopically identifiable species, eight of which can be unambiguously assigned as configurationally and conformationally unique species. NOESY data confirm that for each deoxyribose anomer, atropisomerism occurs around the C⁵-*N*⁵ bond to produce *R*_a and *S*_a isomers. Each atropisomer also exhibits geometrical isomerism about the formyl bond yielding *E* and *Z* conformations. The relative abundances of the various configurational and conformational species have been determined in the 5'-TXT-3' trimer.

MATERIALS AND METHODS

Oligodeoxynucleotide Synthesis.

The ¹³C-labeled MeFapy-dG phosphoramidite was synthesized using ¹³C-labeled methyl iodide using methodology described for synthesis of the unlabeled MeFapy-dG.⁶¹ The incorporation of the ¹³C-labeled MeFapy-dG phosphoramidite into 5'-TXT-3' and 5'-CATXATGACGCT-3' (X = ¹³C-MeFapy-dG) was performed using a Millipore Expedite 8909 Nucleic Acid Synthesis System on a 1 μmol scale using Ultramild Expedite reagents (Glen Research, Sterling, VA) with the standard synthetic protocols for the couplings of the unmodified bases. The couplings of the MeFapy-dG and ¹³C-MeFapy-dG phosphoramidites were performed offline for 30 min as was described by Elmquist et al.⁶³ Fapy-dG nucleotides can isomerize to the pyranose form when the 5'-hydroxyl group is deprotected during solid phase synthesis. The dimethoxytrityl (DMTr) group of the MeFapy-dG or ¹³C-MeFapy-dG was removed using a "short" deprotection cycle (160 μL of Cl₃CCO₂) to minimize rearrangement to the undesired ribopyranose forms. Subsequent to the addition of the MeFapy-dG or ¹³C-MeFapy-dG nucleotides the remainder of the nucleotide additions were performed online using the standard protocols. The completed oligodeoxynucleotides were cleaved from the solid support and deprotected by dissolving in 0.1 M NaOH (1.2 mL). The solution was stirred at room temperature for 24 h. It was then adjusted to pH 7 with 0.1 M acetic acid. Absorbance spectroscopy measurements at 260 nm were used to establish oligodeoxynucleotide concentrations. The calculated extinction coefficients of the modified trimer (5'-TXT-3'), modified dodecamer (dodecamer 5'-CATXAT-GACGCT-3'), and dodecamer complementary strand were $2.93 \times 10^4 \text{ L mol}^{-1} \text{ cm}^{-1}$, $1.14 \times 10^5 \text{ L mol}^{-1} \text{ cm}^{-1}$, and $1.18 \times 10^5 \text{ L mol}^{-1} \text{ cm}^{-1}$, respectively;⁶⁴ the extinction coefficients were not corrected for the presence of the MeFapy-dG lesion.

HPLC.

Reversed-phase HPLC was performed using a Beckman Coulter system composed of a System Gold 126 solvent module and a 168 diode array detector; 32 Karat software (v. 7.0, Beckmann Coulter, Fullerton, CA) was used for the acquisition, processing, and analysis of chromatographic data. The diode array detector was configured to monitor both 254 and 360 nm wavelengths. Prior to injection, samples were centrifuged using Millipore Ultrafree-MC-HV centrifugal filters (Durapore-PVDF 0.45 μm). The modified oligodeoxynucleotides were purified by reverse-phase gradient HPLC utilizing a Gemini-NX 5 μm C18 110 Å (250 mm $\times$ 10 mm) column. The gradient was established using acetonitrile in a mobile phase of 0.1 M ammonium formate (pH = 7). The injection condition was 3% acetonitrile. After 2 min, a linear gradient to 7% acetonitrile was performed over 28 min. Next, a linear gradient to 80% acetonitrile was performed over 2 min. The final step was a linear gradient (2 min) to restore the injection condition of 3% acetonitrile. The flow rate was 1.5 mL/min. The oligodeoxynucleotides were desalted using Sephadex C18 Light Sep-Pak columns (Waters Co., Milford, MA).

Capillary Gel Electrophoresis.

A Beckman Coulter P/ACE MDQ CE instrument was utilized; data was collected at a voltage of 9.0 kV for 55 min with the UV detector set to 254 nm. The cartridge and sample storage temperatures were 30.0 and 25 °C, respectively.

Mass Spectrometry.

Mass spectrometric characterization was performed using a Voyager MALDI-TOF spectrometer. The ^{13}C -MeFapy trimer (5'-TXT-3' X = ^{13}C -MeFapy-dG), MeFapy trimer (5'-TXT-3'; X = MeFapy-dG), and MeFapy dodecamer (5'-CATXATGACGCT-3'; X = ^{13}C -MeFapy-dG) were prepared in a 3-hydroxypicolinic acid and ammonium citrate matrix; spectra were recorded in negative ion mode. All mass spectra were recorded to ± 1 amu. The theoretical masses of the ^{13}C -labeled trimer, unlabeled trimer, and ^{13}C -labeled dodecamer were 908.6 g mol $^{-1}$, 907.6 g mol $^{-1}$, and 3678.5 g mol $^{-1}$, respectively. The masses recorded for the ^{13}C -labeled trimer, unlabeled trimer, and ^{13}C -labeled dodecamer were 908.3 g mol $^{-1}$, 907.3 g mol $^{-1}$, and 3678.1 g mol $^{-1}$, respectively.

NMR.

Sample Preparation.—The 5'-TXT-3' trimer was prepared in 100 mM NaCl and 10 mM NaH₂PO₄ (pH 8); single strand and duplex dodecamer samples were prepared in 100 mM NaCl, 10 mM NaH₂PO₄, and 50 μM Na₂EDTA (pH 8). The MeFapy-dG trimer sample was prepared at a concentration of 681 μM . The ^{13}C -MeFapy-dG trimer spectra were recorded at a concentration of 511 μM . In each instance, the volume was 600 μL in a 5 mm sample tube. Spectra designed to observe nonlabile protons were collected in buffer containing 99.996% D₂O. Spectra designed to observe labile protons were conducted in buffer containing 9:1 H₂O/D₂O; the Watergate pulse program was used to achieve water suppression.⁶⁵ All spectra were collected using Avance III NMR spectrometers equipped with cryogenic probes and precise temperature control (Bruker BioSpin, Inc., Billerica, MA)

Data Collection.—Typical acquisition parameters for ^1H homo-nuclear experiments were as follows: 2000 complex data points with 512 increments, 32 scans per FID, a sweep width of 5250 Hz in both dimensions, relaxation delay of 2.0 s, and States-TPPI acquisition mode. ⁶⁶ Water suppression was achieved by presaturation. Nuclear Overhauser effect spectroscopy (NOESY)^{67,68} experiments were performed at a ^1H frequency of 900 MHz with 2048 real data points in the t_2 dimension and 512 real data points in the t_1 dimension. These spectra were obtained at mixing times of 300 ms, 500 ms, 800 ms, 1 s, and 1.3 s. The data were zero filled during processing to obtain matrices of 2048×1024 points. COSY⁶⁹ and ^1H – ^1H total correlated spectroscopy (TOCSY)⁷⁰ spectra in 9:1 $\text{H}_2\text{O}/\text{D}_2\text{O}$ were recorded at a ^1H frequency of 600 MHz. The mixing time for the TOCSY experiment was 70 ms. Spectra were obtained with 2048 real data points in the t_2 dimension and 512 real data points in the t_1 dimension; the data were zero filled during processing to obtain matrices of 2048×1024 data points. COSY and TOCSY spectra in 99.996% D_2O were recorded at a ^1H frequency of 800 MHz; the mixing time for the TOCSY experiment was 70 ms. Spectra were obtained with 2048 real data points in the t_2 dimension and 512 real data points in the t_1 dimension; the data were zero filled during processing to obtain matrices of 2048×1024 data points. ^1H – ^{13}C HSQC experiments^{71,72} were performed at a ^1H frequency of 600 MHz (150.9 MHz ^{13}C). The data in the t_2 dimension were collected with 2048 complex data points, t_2 dimension sweep width 7500 Hz, 1024 t_1 increments, t_1 dimension sweep width 7545 Hz, 50 scans per FID, GARP decoupling,⁷³ relaxation delay of 1.4 s, echo-and-antiecho acquisition mode,^{74,75} and pulses were optimized for a 140 Hz $^1J_{\text{ch}}$ -coupling constants. The data were processed to obtain matrices of 2048 real data points in the t_2 dimension and 1024 real data points in the t_1 dimension. Distortionless enhancement by polarization transfer (DEPT)⁷⁶ experiments were performed at a ^{13}C frequency of 150.9 MHz with 6144 and 18 432 scans for the decoupled DEPT-135 and nondecoupled DEPT-135 experiments, respectively. NMR data for the single strand trimer and the duplex dodecamer were collected at 278 K. NMR data for the single strand dodecamer was collected at 298 K.

Data Processing.—NMR spectra were processed with the TOPSPIN software (Bruker BioSpin, Inc., Billerica, MA) and analyzed with SPARKY⁷⁷ and TOPSPIN. NMR spectra were integrated using TOPSPIN. ^1H chemical shifts were referenced to the chemical shift of water with respect to trimethylsilyl propanoic acid (TSP).

RESULTS

^{13}C HSQC NMR.

Figure 1 shows the results of a ^{13}C HSQC experiment conducted on the 5'-TXT-3' (X = ^{13}C -MeFapy-dG) trimer. The ^{13}C -CH₃ signals of the MeFapy-dG adduct were observed as ten spectroscopically unique species, existing in equilibrium and in slow exchange on the NMR time scale. The higher field group of five ^{13}C labeled methyl carbons resonated in the 31–32 ppm chemical shift range, whereas the lower field group of five ^{13}C labeled methyl carbons resonated in the 35–36 ppm chemical shift range. Within each group, there were four strong ^{13}C -HSQC signals and one weak ^{13}C -HSQC signal. Likewise, the protons associated with the ^{13}C -labeled methyl groups appeared in the spectrum as five sets of methyl protons resonating in the 2.5–2.6 ppm range and five sets of methyl protons

resonating in the 2.6–2.8 ppm range. The single stranded dodecamer 5'-CATXATGACGCT-3' (X = ^{13}C -MeFapy-dG) was also examined using ^{13}C HSQC (Figure 1). The ^{13}C -HSQC spectrum of the dodecamer was nearly identical to the trimer sequence. However, greater chemical shift dispersion was observed for the methyl proton peaks in the dodecamer, in which the higher field set resonating between 2.6 and 2.9 ppm and the lower field set resonating between 2.8 and 3.1 ppm. The ^{13}C HSQC spectrum was simplified in duplex DNA. When the dodecamer 5'-CATXATGACGCT-3' (X = ^{13}C -MeFapy-dG) was annealed with the complementary strand 5'-AGCGTCATCATG-3', the number of spectroscopically unique ^{13}C -HSQC peaks was reduced to six strong resonances with two additional extremely weak resonances (Figure 1).

5'-TXT-3' (X = ^{13}C -MeFapy-dG) Trimer.

Geometrical Isomerization of Formyl Group.—The trimer exhibited the least complex NMR spectra and served as a model to characterize the equilibrium distribution of MeFapy-dG isomers in DNA, as it had been employed to examine the Fapy-dG lesion.⁵⁸ The ten unique signals arising from the trimer in the ^{13}C HSQC spectrum were examined utilizing a ^1H -coupled DEPT-135 ^{13}C NMR experiment (Figure 2). A non- ^1H -decoupled spectrum collected under the same conditions showed that each of the ten ^{13}C signals was resolved into a quartet of doublets, arising from a combination of the larger $^2J_{\text{CH}}$ methyl coupling (q) and a smaller $^3J_{\text{CH}}$ between the ^{13}C -labeled methyl carbon (d) and the formyl proton. The $^3J_{\text{CH}}$ values of the downfield group of ^{13}C signals near 35 ppm differed from those of the upfield group near 31.5 ppm, measured as 1.7 and 4.0 Hz, respectively (Figure 2). This allowed assignment of the upfield group of ^{13}C MeFapy-dG resonances (31–32 ppm) as arising from *E* geometrical isomers in which the ^{13}C nucleus and formyl proton exist in a transoid relationship; the downfield group of ^{13}C MeFapy-dG resonances (35–36 ppm) was assigned as *Z* geometrical isomers in which the two nuclei exist in a cisoid relationship (Scheme 2).

Epimerization of Deoxyribose α and β Anomers.—A series of ^1H NOESY experiments established the presence of both α and β anomers of the MeFapy-dG deoxyribose ring (Scheme 1). The resonances of the β anomers were identified due to their strong deoxyribose H1'-H4' NOEs and weaker deoxyribose H1'-H3' NOEs (Figure 3).⁵⁸ The 3J coupling between the MeFapy-dG N^6H protons and the deoxyribose H1' protons allowed the assignment of the former from COSY and TOCSY spectra. The β anomers exhibited strong N^6H -H3' NOEs and weak N^6H -H4' NOEs. The β anomers exhibited stronger H1'-H4' NOEs than H1'-H3' NOEs. In contrast, the α anomers had stronger H1'-H3' than H1'-H4' NOEs (Figure 3).⁵⁸ The α anomers were also identified by NOEs between N^6H and deoxyribose H4'. No NOE cross peaks were observed between the N^6H and deoxyribose H3' protons for the α anomers.

Atropisomerism about the C5-N⁵ Bond.—The assignments of R_a and S_a atropisomers around the C⁵-N⁵ bond (Scheme 2) were established by comparing relative intensities of NOEs between the N^6H and CH₃ protons and the formyl protons (Table 1). An inverse relationship was observed between N^6H -CH₃ and N^6H -formyl NOE intensities between atropomers in the same anomeric configuration and conformation around the formyl bond;

the atropomer with a stronger N^6H-CH_3 NOE exhibited a weaker N^6H -formyl NOE. When the NOE intensities were adjusted for the population of each anomeric species, it was possible to establish the assignments of the atropisomers.

Abundances of Configurational and Conformational Isomers.—The isomeric assignment of the MeFapy-dG lesion is summarized in the 2D 1H NOESY data shown in Figure 4. This enabled calculation of the relative populations of each isomeric species (Scheme 2 and Table 2). The α and β anomers existed in a 3:7 ratio, favoring the β anomer. With respect to the C5- N^5 bond, the R_a and S_a atropisomers were equally populated for the β anomer. The Z geometrical isomer of the formyl moiety was preferred. For the α anomer, the $E-S_a$ isomer was present at 12%, whereas all other isomers were present at 5–7%.

5'-CATXATGACGCT-3' Dodecamer.

The ^{13}C -isotope labeling experiments were repeated in the dodecamer 5'-CATXATGACGCT-3' ($X = ^{13}C$ -MeFapy-dG) (Figure 1). A similar result was obtained, resulting in ten unique ^{13}C signals, divided into two groups. The results in the longer oligodeoxynucleotide corroborated the data in the trimer sequence. This showed that the hindered rotation about the C5- N^5 formyl bond and the interconversion between α and β deoxyribose anomers was a characteristic of the MeFapy-dG lesion in longer stretches of DNA, and not due to a greater degree of conformational freedom in the trimer⁵⁸ as compared to the dodecamer.

Duplex Dodecamer.

The dodecamer was annealed with the complementary strand 5'-AGCGTCATCATG-3'. In the duplex DNA, the ^{13}C HSQC NMR experiment revealed six strong resonances rather than ten unique ^{13}C resonances (Figure 1). The six strong ^{13}C resonances were divided into two groups. The results suggested that rotation about the formyl bond, the C5- N^5 bond, and the interconversion between β and β deoxyribose anomers differed in duplex DNA as opposed to single stranded DNA. This was interpreted in terms of the complementary strand of DNA removing specific conformational degrees of freedom with respect to single strand DNA. Two weak HSQC resonances could not be identified unequivocally but likely represent small populations of disfavored conformers of the formyl and the C5- N^5 bonds in duplex DNA.

DISCUSSION

The present results have allowed the first sequence-specific characterization and quantitation of the interconverting species (Schemes 1 and 2) of MeFapy-dG adducts in DNA.^{18,40,42,43,52–60} This characterization is significant because different isomeric species may interact differently with DNA processing enzymes and lead to different biological outcomes.^{78,79} Sequence-specific differences in the populations of these species may also modulate biological responses to the lesion.

Isomeric Characterization of MeFapy-dG Lesion.

In the 5'-TXT-3' trimer, MeFapy-dG exists as ten spectroscopically identifiable and distinct isomeric species (Table 1). Eight of these may be identified unequivocally. Four of these

isomeric species are β deoxyribose anomers, while another four are α deoxyribose anomers (Scheme 1). At equilibrium, the β -MeFapy-dG anomer is favored, and the two anomers exist in a 3:7 α : β ratio. Christov et al.⁵⁷ utilized an *Escherichia coli* endonuclease IV (Endo IV)-based assay⁸⁰ and showed that at equilibrium in duplex DNA the α and β -MeFapy-dG anomers exist in a 20:80 α : β ratio. The unsubstituted Fapy-dG adduct also favors the β anomer.^{53,80} Brown et al.⁵⁵ showed that the β anomer of AFB₁-Fapy-dG is also favored in duplex DNA.^{55,81} The preference for the β anomer is likely due to increased distortion necessary to accommodate the α anomer in DNA. Although there are no structural data available for the α anomer of the MeFapy-dG adduct in DNA, an NMR structure confirmed that the α anomer AFB₁-Fapy-dG adduct distorts DNA.⁸² This was consistent with perturbations of the base stacking register of the DNA. The ϵ and ζ DNA backbone torsion angles were also perturbed.⁵⁵ Other structural studies of α -anomeric nucleotides in DNA have indicated distortion of the duplex.^{83,84} Burgdorf and Carel⁵⁴ examined a chemically protected nucleoside analog of the Fapy-dG lesion and reported a half-life on the order of hours for interconversion between α and β anomers. Christov et al.⁵⁷ showed that in DNA, where the MeFapy-dG adduct is presumably less susceptible to reaction, equilibration between the α and β anomers occurred over a period of 3 days.⁵⁷

In the single strand trimer, each of the α and β anomers of the MeFapy-dG lesion exists as two pairs of rotamers, one set involving atropisomerism about the C5- N^5 bond and the other involving geometrical isomerization of the formyl bond (Scheme 2). Tomasz et al.⁵² identified C5- N^5 atropisomers with the Fapy-dG nucleoside derived from 7-ethyl-dG as well as with the mitomycin C-Fapy-dG adduct. Brown et al.⁵⁵ showed that the R_a atropisomer is favored for the AFB₁-Fapy adduct. NMR studies of the AFB₁-Fapy-dG adduct in DNA showed that intercalation of the AFB₁ moiety on the 5' face of the pyrimidine ring maintains the R_a absolute configuration for the C5- N^5 bond.⁸¹

The second set of MeFapy-dG rotamers involves geometrical isomerism of the formyl bond (Scheme 2). The $^3J_{CH}$ values of the *E* and *Z* geometrical isomers were measured as 4.0 and 1.7 Hz (Figure 2). The potential of the MeFapy-dG lesion to interconvert between *E* and *Z* geometrical isomers of the formamide⁸⁵ is recognized.^{43,52} Raoul et al.⁸⁶ demonstrated the *E* and *Z* geometrical isomers of the Fapy-dA lesion resulting from rotation around the formyl bond existed in a 1:5 trans (*E*)/*cis* (*Z*) ratio, with an energy barrier for interconversion calculated as 75.5 kJ/mol at 370 K. The preference for the *Z* isomer has been attributed to hydrogen bonding of the carbonyl oxygen proximal to the pyrimidine ring with the proton on N^4 .^{52,79} The equilibrium between *E* and *Z* isomers of the AFB₁-Fapy-dG adduct is controlled by major groove hydrogen bonding interactions.⁸⁷

Two minor species seen in the ¹³C-HSQC (Figure 1) and DEPT-135 (Figure 2) spectra may arise from rotation about the glycosidic bond of the MeFapy-dG adduct. The formyl proton resonances and exchangeable N^6 H proton resonances for these two species are not observed. Consequently, no diagnostic NOEs between the Fapy CH₃ protons, the formyl protons, and the N^6 H protons are observed. This is attributed, in part, to the low populations of these species. It is also possible that these species have lower thermal stability and exhibit a greater degree of structural disorder. Rotation of the glycosidic bond into the *syn* conformation would relocate the nonplanar MeFapy-dG formamido moiety to the base

stacking region of the DNA helix. This would be anticipated to constrain rotation about the C5–N⁵ and formyl bonds, which is consistent with the observation that these two species do not exhibit rotational isomerism about the C5–N⁵ and formyl bonds (Figure 1). The NMR spectra suggest that for these minor species the C5–N⁵ and formyl bonds favor a single (lowest energy) conformation. Rotation about the C5–N⁵ bond may also be constrained in duplex DNA, which is consistent with the ¹³C-HSQC NMR data in Figure 1.

The biological consequences of these isomeric states of the MeFapy-dG adduct and their potential for differing interactions with DNA processing enzymes remain to be determined. The observation that these isomers of the MeFapy-dG adduct are simultaneously populated in DNA indicates that they possess similar energies. Small energy differences between these isomers may be negligible compared to the energetics of enzyme–substrate binding. However, the observation of discrete spectroscopic signals for these isomers indicates that energy barriers for interconversions are sufficiently high that rates of interconversion are slow on the millisecond time scale of the NMR experiment. Rates of interconversion are slow at longer time scales as well because some may be transiently separated using HPLC. A complete understanding of the processing of different configurational and conformational isomers of the MeFapy-dG lesion will require further characterization of both thermodynamic and structural differences of MeFapy-dG adducts in complex with specific DNA processing enzymes.

No NMR or crystallographic structural data for the MeFapy-dG adduct are available in DNA, which probably is a reflection of the complex equilibrium mixture of configurational and conformational isomers existing for this adduct. The carbocyclic analog of the β -anomer of the Fapy-dG lesion (cFapy-dG),^{56,88} which cannot undergo anomerization, has been examined in DNA paired to dC.⁸⁹ In this case, the DNA duplex adopted a right-handed helical structure with all deoxyribose sugars in the anti conformation, forming Watson–Crick base pair alignments. The formamido group of cFapy-dG rotated out of the pyrimidine plane and was present in both the *Z* and *E* geometric isomers with an approximate 2:1 ratio. Rotation about the C5–N⁵ bond was not reported, consistent with the notion that rotation about this bond is constrained in DNA.

Sequence Dependence.

The present analyses of configurational and conformational isomers of the MeFapy-dG lesion were conducted in the trimer 5'-TXT-3', and the dodecamer 5'-CATXATGACGCT-3', in which the lesion is embedded in the 5'-TXA-3' sequence. One may envision that DNA sequence could modulate the relative proportions of specific isomeric species of the MeFapy-dG lesion in DNA. Differing identities of the MeFapy-dG isomers can, theoretically, lead to different interactions with proteins. We hypothesize that stabilization by hydrogen bonding and destabilization due to steric effects will influence the distribution of rotameric isomers of MeFapy-dG. Sequence-specific effects have been observed for the AFB₁-Fapy-dG adduct, where the equilibrium between *E* and *Z* isomers of the formyl group is controlled by major groove hydrogen bonding interactions.⁸⁷ In the 5'-XA-3' sequence (X = AFB₁-Fapy-dG) the *E* isomer of the formamide group is favored, attributed to formation of a hydrogen bond between the formyl oxygen and the N⁶ exocyclic

amino group of the 3'-neighbor adenine. The 5'-XC-3' sequence exhibits a 7:3 *E/Z* ratio. Here, the *E* isomer is favored by a hydrogen bond between the formyl oxygen and the *N*⁴-dC exocyclic amino group of the 3'-neighbor cytosine. In contrast, the 5'-XT-3' and 5'-XY-3' sequences (Y = 7-deaza-dG) cannot form a hydrogen bond between the formyl oxygen and the 3'-neighbor T or Y, respectively, and in these sequences the *Z* isomer is favored. With regard to steric effects one can speculate that if the 3' neighbor of MeFapy-dG is a purine as opposed to a pyrimidine, greater steric hindrance may disfavor the α anomer. It is also possible that sequence-specific hydrogen bonding interactions between the α anomer MeFapy-dG base and its 3' neighbor may differentially stabilize the α anomer.

Repair.

Kadlubar and co-workers showed that the MeFapy-dG adduct is both present in DNA *in vivo* after exposure to methylating agents⁴² and is more persistent than is the initially formed N7-Me-dG lesion.⁴³ Their results suggest that MeFapy-dG may be less efficiently repaired. Asagoshi et al.^{90,91} showed that repair of MeFapy-dG by DNA glycosylases, which flip the damaged Fapy-dG base out of the DNA and into an enzyme binding pocket, is dependent both upon the glycosylase and the identity of the complementary base. In one study, they treated substrates containing MeFapy-dG:N pairs (N = A, G, C, T) as well as a Tg:A pair, with the repair enzymes Endo III, Endo VIII, and mNth1 (murine analog of EndoIII).⁹⁰ The activity of Endo III and Endo VIII for MeFapy-dG was highest with G and the lowest with C, but mNth1 recognized all MeFapy-dG pairs equally well and the activity was comparable to that with the thymine glycol (Tg) lesion. In another study,⁹¹ they examined excision of MeFapy-dG by human 7,8-dihydro-8-oxoguanine DNA glycosylase (hOGG1) and formamidopyrimidine DNA glycosylase (Fpg). They showed that preferences for the base opposite the MeFapy-dG lesion were muted when compared to 8-oxo-dG. Wiederholt et al.⁹² described interactions of DNA containing the Fapy-dG lesion with the Fpg and MutY glycosylases. They concluded that the Fpg glycosylase excises Fapy-dG efficiently. Structural studies by Coste et al.⁹³ showed differences between 8-oxodG and the carbocyclic analog of the β -anomer of the Fapy-dG lesion (cFapy-dG)^{56,88} within the active site of the Fpg glycosylase.⁹³ Although 8-oxodG is accommodated in the active site by rotation about the glycosidic bond into the *syn* conformation, cFapy-dG is accommodated in the *anti* conformation about the glycosidic bond.⁹³ Coste et al.⁹³ further showed that recognition of cFapy-dG by Fpg involves both Watson-Crick face recognition of the damaged base and also recognition of specific rotamers of the Fapy moiety by specific Fpg residues, mediated by structural waters. Burgdorf and Carel⁵⁴ suggested that rotation about the formyl bond facilitates hydrogen bonding between the formamide carbonyl oxygen and the NH of the N-glycosidic bond, providing a potential mechanistic basis for hindering recognition of Fapy-dG by the glycosylase. The formation of productive BER complexes also depends upon the size of the N7-dG adduct, which may alter conformational equilibria in Fapy-dG lesions.⁷⁹

Error-Prone Replication.

The α -anomer of MeFapy-dG inhibits DNA synthesis.⁴⁶⁻⁴⁸ This has also been noted for the Fapy-dG lesion.⁵⁶ Nonetheless, MeFapy-dG is mutagenic, which has been attributed to the β -anomeric configuration of this lesion.⁹⁴ Both the rates of mutation and the spectrum of mutations induced by MeFapy-dG vary depending on the identities of the 5' and 3'

neighboring bases in DNA.^{49,51} Similarly, Kalam et al.⁹⁵ reported differences in the mutational frequency of the Fapy-dG lesion in the 5'-TXT-3' sequence versus the 5'-TGA-3' sequence. Working in primate cells, Earley et al.⁴⁹ determined that both the mutational rate and mutation spectrum MeFapy-dG is sequence dependent; alteration of the 3' and 5' neighbors of the lesion resulted in varying and diverse mutational spectra. They examined nine site-specifically modified oligodeoxynucleotides containing MeFapy-dG. In each instance, MeFapy-dG was mutagenic, predominantly causing G to T transversions, although single and dinucleotide deletions were also found in specific sequence contexts. Single-nucleotide deletions occurred at the adducted site and one nucleotide downstream of the MeFapy-dG adducts.

The mechanistic basis for these mutations remains incompletely understood. Changes in local sequence of the MeFapy-dG caused differences in the mutations and their frequency.⁴⁹ Different local sequences are likely to place the MeFapy-dG adduct in slightly different chemical environments, which will influence the distribution of rotational isomers of both the formamide unit and the glycosidic bond. Structural studies of modified template/primers complexed with specific DNA polymerases show that 8-oxodG may adopt either *syn* or *anti* conformations around the glycosidic bond during replication, with the *syn* conformation being associated with G to T transversions due to Hoogsteen-like pairing with incoming dATP.⁹⁶⁻¹⁰¹ Wiederholt and Greenberg¹⁰² examined the Fapy-dG lesion and suggested that in the *syn* conformation, it may also miscode for dATP insertion during error prone replication, leading to G → T transversions.^{95,103} Likewise, a *syn-anti* interconversion about the glycosidic bond could account for the ability of the MeFapy-dG lesion to cause G to T transversions. In the *syn* conformation, the MeFapy-dG moiety faces into the DNA. Wiederholt and Greenberg¹⁰² propose that geometrical isomerization about the formyl bond and atropisomerism about the C5-N⁵ bond create "T-like" recognition of the Fapy-dG lesion, providing a potential explanation for G to T transversions.¹⁰³ With MeFapy-dG, the CH₃ group must also be accommodated within the stacking of the DNA helix. This is likely to be disfavored thermodynamically so that the *syn* conformation may be present only at low levels in DNA. The observation of the minor peaks in the ¹³C-HSQC (Figure 1) and DEPT-135 (Figure 2) spectra, possibly arising from *syn* conformers, is consistent with this notion.

Gehrke et al.¹⁰⁴ examined the carbocyclic analog of the β -anomer of the Fapy-dG lesion (cFapy-dG)^{56,88} complexed with the *G. stearothermophilus* Pol I polymerase. They characterized both error-free and error-prone replication complexes, the latter involving misinsertion of dATP. In both instances, cFapy-dG maintained the *anti* conformation about the glycosidic bond. The *anti* conformation of the glycosidic bond places the cFapy-dG moiety into the major groove. In this orientation, the cFapy-dG moiety is not involved with Watson-Crick base pairing with incoming dNTPs, consistent with an error-free bypass mechanism. The observation of the *anti* conformation of cFapy-dG in the error-prone bypass structure, with incoming dATP, was unexpected. Gehrke et al.¹⁰⁴ attributed this to the ability of the cFapy-dG lesion to rotate the formamide moiety out of plane from the pyrimidine moiety, as had also been observed in NMR studies,⁸⁹ giving it structural flexibility, ostensibly reducing the need to rotate the cFapy-dG lesion into the *syn* conformation during misinsertion of dATP. Consequently, Gehrke et al.¹⁰⁴ concluded that the G to T transversions associated

with Fapy-dG result from a base shifting and tautomerization type mechanism, as opposed to rearrangement of the lesion into the *syn* conformation about the glycosidic bond.

MeFapy-dG itself has been examined in complex with the error-prone DNA polymerases Dpo4 and pol η .¹⁰⁵ These structures were of the insertion and extension stages, respectively, and both were representative of error-free replication, with MeFapy-dG in the *anti* conformation about the glycosidic bond and forming Watson–Crick pairs with dCTP or dC. These structures confirmed that the formamide group of MeFapy-dG was approximately perpendicular to the pyrimidine ring of the Fapy moiety, as had been noted by Gehrke et al.,¹⁰⁴ and in NMR studies,⁸⁹ rather than in the plane of the ring. Interestingly, the differences in distance between the N^6 H-methyl and N^6 H-formyl protons between *E* and *Z* isomers are predicted to be smaller when the formamide moiety is perpendicular to the pyrimidine ring compared to if the formamide is in the plane of the pyrimidine ring. This is in fact observed in the relative NOE peak strengths between isomers in Table 1. Thus, while a detailed structural analysis of the MeFapy-dG lesion has not been conducted, it seems likely that the MeFapy-dG formamide moiety is also out of plane with the pyrimidine ring in both the trimer and dodecamer DNAs examined herein.

Minko et al.¹⁰⁶ recently demonstrated that the Fapy-dG lesion derived from a nitrogen mustard is mutagenic in primate kidney cells (COS7) *in vitro*, causing targeted G to T transversions at a frequency of ~11–12%. The mutation frequency decreased when the NM-Fapy-dG containing vector was treated with *E. coli* endonuclease IV, which specifically incises the α -anomer. This result suggests that the non-canonical α -anomer is a major contributor to mutagenesis and underscores the importance of understanding the structure dynamics of Fapy-dG lesions.

Translational and Clinical Aspects.

The potential for cellular repair and replication of DNA methylation damage is of importance in characterizing the efficacy and developing patient-specific approaches of TMZ-based adjuvant therapies. Some tumor cells are better able to repair TMZ-induced DNA alkylation, which diminishes the efficacy of this antitumor agent. Significant effort has been expended upon understanding the chemistry and biology of the less abundant N3-dA and O⁶-dG methylation products in DNA.^{6,10} Cells expressing O⁶-alkylguanine transferase (AGT) encoded by the O⁶-methylguanine-DNA methyl transferase gene (*MGMT*)¹⁰⁷ exhibit refractory behavior to TMZ treatment, whereas epigenetic silencing of this gene increases cytotoxic response.^{3,108,109} Likewise, the repair of 3-methyl-dA and AP sites by the BER pathway is a resistance mechanism to TMZ treatment in glioblastoma.⁶ Consequently, understanding mechanisms of recognition, repair, and the error-prone replication of the various types of DNA methylation damage induced by TMZ treatment will be of importance since it could impact outcomes in chemotherapy.^{6,10}

Summary.

Isotope-edited NMR provides a powerful approach toward evaluating and quantitating complex mixtures of isomeric species, such as arising from MeFapy-dG lesions in DNA, *in situ*.^{58,60,82} This is critical to understanding the chemistry and biology of these lesions.^{78,79}

Here, the CH₃ carbon in the MeFapy-dG moiety was labeled with ¹³C, using chemistry developed by our laboratories,⁶¹ enabling quantitative assessment of configurational and conformational isomers of the lesion. This approach can be applied to additional Fapy-dG-type lesions, for example, as occur from exposures to ^{86,110,111} nitrogen mustards.

ACKNOWLEDGMENTS

We thank Dr. Wade Calcutt for assistance with mass spectrometry.

Funding

We acknowledge funding by NIH Grant P01 CA-160032 (M.P.S. and C.J.R.) and the NIH Training Program in Environmental Toxicology, T32 ES-007028 (S.N.B.). The Vanderbilt-Ingram Cancer Center is funded by NIH Grant No. P30 CA-068485. The Vanderbilt Center in Molecular Toxicology was funded by NIH Grant No. P30 ES-000267. Partial support was provided by instrumentation Grant No. S10 RR05805, and S10 RR025677 and National Science Foundation Instrumentation Grant No. DBI 0922862, the latter funded by the American Recovery and Reinvestment Act of 2009 (Public Law 111-5). Vanderbilt University assisted with the purchase of NMR instrumentation. Funding for Open Access Charge: National Institutes of Health.

REFERENCES

- (1). Newlands ES, Stevens MF, Wedge SR, Wheelhouse RT, and Brock C (1997) Temozolomide: A review of its discovery, chemical properties, pre-clinical development and clinical trials. *Cancer Treat. Rev* 23, 35–61. [PubMed: 9189180]
- (2). Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJ, Belanger K, Brandes AA, Marosi C, Bogdahn U, Curschmann J, Janzer RC, Ludwin SK, Gorlia T, Allgeier A, Lacombe D, Cairncross JG, Eisenhauer E, and Mirimanoff RO (2005) Radiotherapy plus concomitant and adjuvant Temozolomide for glioblastoma. *N. Engl. J. Med* 352, 987–996. [PubMed: 15758009]
- (3). Stupp R, Hegi ME, Mason WP, van den Bent MJ, Taphoorn MJ, Janzer RC, Ludwin SK, Allgeier A, Fisher B, Belanger K, Hau P, Brandes AA, Gijtenbeek J, Marosi C, Vecht BJ, Mokhtari K, Wesseling P, Villa S, Eisenhauer E, Gorlia T, Weller M, Lacombe D, Cairncross JG, and Mirimanoff RO (2009) Effects of radiotherapy with concomitant and adjuvant Temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *Lancet Oncol.* 10, 459–466. [PubMed: 19269895]
- (4). Stupp R, Mayer M, Kann R, Weder W, Zouhair A, Betticher DC, Roth AD, Stahel RA, Majno SB, Peters S, Jost L, Furrer M, Thierstein S, Schmid RA, Hsu-Schmitz SF, Mirimanoff RO, Ris HB, and Pless M (2009) Neoadjuvant chemotherapy and radiotherapy followed by surgery in selected patients with stage IIIB non-small-cell lung cancer: A multicentre phase II trial. *Lancet Oncol.* 10, 785–793. [PubMed: 19604722]
- (5). Chamberlain M (2011) Evolving strategies: Future treatment of glioblastoma. *Expert Rev. Neurother.* 11, 519–532. [PubMed: 21469925]
- (6). Bobola MS, Kolstoe DD, Blank A, Chamberlain MC, and Silber JR (2012) Repair of 3-methyladenine and abasic sites by base excision repair mediates glioblastoma resistance to Temozolomide. *Front. Oncol.* 2, 176. [PubMed: 23230562]
- (7). Stupp R, Taillibert S, Kanner AA, Kesari S, Steinberg DM, Toms SA, Taylor LP, Lieberman F, Silvani A, Fink KL, Barnett GH, Zhu JJ, Henson JW, Engelhard HH, Chen TC, Tran DD, Sroubek J, Tran ND, Hottinger AF, Landolfi J, Desai R, Caroli M, Kew Y, Honnorat J, Idhah A, Kirson ED, Weinberg U, Palti Y, Hegi ME, and Ram Z (2015) Maintenance therapy with tumor-treating fields plus Temozolomide vs. Temozolomide alone for glioblastoma: A randomized clinical trial. *JAMA* 314, 2535–2543. [PubMed: 26670971]
- (8). Seystahl K, Wick W, and Weller M (2016) Therapeutic options in recurrent glioblastoma—An update. *Crit. Rev. Oncol. Hematol.* 99, 389–408. [PubMed: 26830009]
- (9). Lieberman F (2017) Glioblastoma update: Molecular biology, diagnosis, treatment, response assessment, and translational clinical trials. *F1000Research* 6, 1892. [PubMed: 29263783]

- (10). Fu D, Calvo JA, and Samson LD (2012) Balancing repair and tolerance of DNA damage caused by alkylating agents. *Nat. Rev. Cancer* 12, 104–120. [PubMed: 22237395]
- (11). Sedgwick B, Bates PA, Paik J, Jacobs SC, and Lindahl T (2007) Repair of alkylated DNA: Recent advances. *DNA Repair* 6, 429–442. [PubMed: 17112791]
- (12). Reiner B, and Zamenhof S (1957) Studies on the chemically reactive groups of deoxyribonucleic acids. *J. Biol. Chem* 228, 475–486. [PubMed: 13475333]
- (13). Lawley PD, and Thatcher CJ (1970) Methylation of deoxyribonucleic acid in cultured mammalian cells by N-methyl-N'-nitro-N-nitrosoguanidine. The influence of cellular thiol concentrations on the extent of methylation and the 6-oxygen atom of guanine as a site of methylation. *Biochem. J* 116, 693–707. [PubMed: 5435496]
- (14). Beranek DT, Weis CC, and Swenson DH (1980) A comprehensive quantitative analysis of methylated and ethylated DNA using high pressure liquid chromatography. *Carcinogenesis* 1, 595–606. [PubMed: 11219835]
- (15). Beranek DT (1990) Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. *Mutat. Res., Fundam. Mol. Mech. Mutagen* 231, 11–30.
- (16). Smith SA, and Engelward BP (2000) In vivo repair of methylation damage in Aag 3-methyladenine DNA glycosylase null mouse cells. *Nucleic Acids Res* 28, 3294–3300. [PubMed: 10954597]
- (17). Boysen G, Pachkowski BF, Nakamura J, and Swenberg JA (2009) The formation and biological significance of N7-guanine adducts. *Mutat. Res., Genet. Toxicol. Environ. Mutagen* 678, 76–94.
- (18). Gates KS, Nooner T, and Dutta S (2004) Biologically relevant chemical reactions of N7-alkylguanine residues in DNA. *Chem. Res. Toxicol* 17, 839–856. [PubMed: 15257608]
- (19). Kyrtopoulos SA (1998) DNA adducts in humans after exposure to methylating agents. *Mutat. Res., Fundam. Mol. Mech. Mutagen* 405, 135–143.
- (20). Dahlmann HA, Vaidyanathan VG, and Sturla SJ (2009) Investigating the biochemical impact of DNA damage with structure-based probes: Abasic sites, photodimers, alkylation adducts, and oxidative lesions. *Biochemistry* 48, 9347–9359. [PubMed: 19757831]
- (21). Shrivastav N, Li D, and Essigmann JM (2010) Chemical biology of mutagenesis and DNA repair: cellular responses to DNA alkylation. *Carcinogenesis* 31, 59–70. [PubMed: 19875697]
- (22). Hecht SS, and Hoffmann D (1988) Tobacco-specific nitrosamines, an important group of carcinogens in tobacco and tobacco smoke. *Carcinogenesis* 9, 875–884. [PubMed: 3286030]
- (23). Pfeifer GP, Denissenko MF, Olivier M, Tretyakova N, Hecht SS, and Hainaut P (2002) Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. *Oncogene* 21, 7435–7451. [PubMed: 12379884]
- (24). Hecht SS (2003) Tobacco carcinogens, their biomarkers and tobacco-induced cancer. *Nat. Rev. Cancer* 3, 733–744. [PubMed: 14570033]
- (25). Hecht SS (2008) Progress and challenges in selected areas of tobacco carcinogenesis. *Chem. Res. Toxicol* 21, 160–171. [PubMed: 18052103]
- (26). Sen NP, Miles WF, Donaldson B, Panalaks T, and Iyengar JR (1973) Letter: formation of nitrosamines in a meat curing mixture. *Nature* 245, 104–105. [PubMed: 4582759]
- (27). Bartsch H, and Montesano R (1984) Relevance of nitrosamines to human cancer. *Carcinogenesis* 5, 1381–1393. [PubMed: 6386215]
- (28). Lindamood C, 3rd, Bedell MA, Billings KC, Dyroff MC, and Swenberg JA (1984) Dose response for DNA alkylation, [3H]thymidine uptake into DNA, and O₆-methylguanine-DNA methyltransferase activity in hepatocytes of rats and mice continuously exposed to dimethylnitrosamine. *Cancer Res.* 44, 196–200. [PubMed: 6690035]
- (29). Sedgwick B, and Vaughan P (1991) Widespread adaptive response against environmental methylating agents in microorganisms. *Mutat. Res., Fundam. Mol. Mech. Mutagen* 250, 211–221.
- (30). Bianchini F, and Wild CP (1994) 7-Methyldeoxyguanosine as a marker of exposure to environmental methylating agents. *Toxicol. Lett* 72, 175–184. [PubMed: 8202930]
- (31). Swenberg JA, Boysen G, Georgieva N, Bird MG, and Lewis RJ (2007) Future directions in butadiene risk assessment and the role of cross-species internal dosimetry. *Chem.-Biol Interact* 166, 78–83. [PubMed: 17343837]

- (32). Swenberg JA, Bordeerat NK, Boysen G, Carro S, Georgieva NI, Nakamura J, Troutman JM, Upton PB, Albertini RJ, Vacek PM, Walker VE, Sram RJ, Goggins M, and Tretyakova N (2011) 1,3-Butadiene: Biomarkers and application to risk assessment. *Chem.-Biol. Interact.* 192, 150–154.
- (33). Rydberg B, and Lindahl T (1982) Nonenzymatic methylation of DNA by the intracellular methyl group donor S-adenosyl-L-methionine is a potentially mutagenic reaction. *EMBO J.* 1, 211–216. [PubMed: 7188181]
- (34). De Vita S, De Matteis S, Laurenti L, Chiusolo P, Reddiconto G, Fiorini A, Leone G, and Sica S (2005) Secondary Ph+ acute lymphoblastic leukemia after Temozolomide. *Ann. Hematol* 84, 760–762. [PubMed: 16044311]
- (35). Chamberlain MC, and Raizer J (2009) Extended exposure to alkylator chemotherapy: Delayed appearance of myelodysplasia. *J. Neuro-Oncol* 93, 229–232.
- (36). Momota H, Nariata Y, Miyakita Y, Hosono A, Makimoto A, and Shibui S (2010) Acute lymphoblastic leukemia after Temozolomide treatment for anaplastic astrocytoma in a child with a germline TP53 mutation. *Pediatr. Blood Cancer* 55, 577–579. [PubMed: 20658636]
- (37). Tell G, Fantini D, and Quadrifoglio F (2010) Understanding different functions of mammalian AP endonuclease (APE1) as a promising tool for cancer treatment. *Cell. Mol. Life Sci* 67, 3589–3608. [PubMed: 20706766]
- (38). Park JS, Kim HL, Kim YJ, Weon JI, Sung MK, Chung HW, and Seo YR (2014) Human AP endonuclease 1: A potential marker for the prediction of environmental carcinogenesis risk. *Oxid. Med. Cell. Longevity* 2014, 730301.
- (39). Wallace SS (2014) Base excision repair: A critical player in many games. *DNA Repair* 19, 14–26. [PubMed: 24780558]
- (40). Tudek B (2003) Imidazole ring-opened DNA purines and their biological significance. *J. Biochem. Mol. Biol.* 36, 12–19. [PubMed: 12542970]
- (41). Pujari SS, and Tretyakova N (2017) Chemical biology of N5-substituted formamidopyrimidine DNA adducts. *Chem. Res. Toxicol.* 30, 434–452. [PubMed: 27959490]
- (42). Beranek DT, Weis CC, Evans FE, Chetsanga CJ, and Kadlubar FF (1983) Identification of N-5-methyl-N-5-formyl-2,5,6-triamino-4-hydroxypyrimidine as a major adduct in rat-liver DNA after treatment with the carcinogens, N,N-dimethylnitrosamine or 1,2-dimethylhydrazine. *Biochem. Biophys. Res. Commun* 110, 625–631. [PubMed: 6838542]
- (43). Kadlubar FF, Beranek DT, Weis CC, Evans FE, Cox R, and Irving CC (1984) Characterization of the purine ring-opened 7-methylguanine and its persistence in rat bladder epithelial DNA after treatment with the carcinogen N-methylnitrosourea. *Carcinogenesis* 5, 587–592. [PubMed: 6722978]
- (44). Den Engelse L, Menkveld GJ, Debrij RJ, and Tate AD (1986) Formation and stability of alkylated pyrimidines and purines (including imidazole ring-opened 7-alkylguanine) and alkylphosphotriesters in liver DNA of adult-rats treated with ethylnitrosourea or dimethylnitrosamine. *Carcinogenesis* 7, 393–403. [PubMed: 3948325]
- (45). Newbold RF, Warren W, Medcalf ASC, and Amos J (1980) Mutagenicity of carcinogenic methylating agents is associated with a specific DNA modification. *Nature* 283, 596–599. [PubMed: 7354843]
- (46). Boiteux S, and Laval J (1983) Imidazole open ring 7-methylguanine - an inhibitor of DNA-synthesis. *Biochem. Biophys. Res. Commun* 110, 552–558. [PubMed: 6340667]
- (47). O'Connor TR, Boiteux S, and Laval J (1988) Ring-opened 7-methylguanine residues in DNA are a block to *in vitro* DNA synthesis. *Nucleic Acids Res* 16, 5879–5894. [PubMed: 3399381]
- (48). Asagoshi K, Terato H, Ohyama Y, and Ide H (2002) Effects of a guanine-derived formamidopyrimidine lesion on DNA replication: Translesion DNA synthesis, nucleotide insertion, and extension kinetics. *J. Biol. Chem* 277, 14589–14597. [PubMed: 11839760]
- (49). Earley LF, Minko IG, Christov PP, Rizzo CJ, and Lloyd RS (2013) Mutagenic spectra arising from replication bypass of the 2,6-diamino-4-hydroxy-*N*⁵-methyl formamidopyrimidine adduct in primate cells. *Chem. Res. Toxicol* 26, 1108–1114. [PubMed: 23763662]
- (50). Christov PP, Angel KC, Guengerich FP, and Rizzo CJ (2009) Replication past the *N*⁵-methyl-formamidopyrimidine lesion of deoxyguanosine by DNA polymerases and an improved

procedure for sequence analysis of *in vitro* bypass products by mass spectrometry. Chem. Res. Toxicol 22, 1086–1095. [PubMed: 19397282]

- (51). Christov PP, Yamanaka K, Choi JY, Takata K, Wood RB, Guengerich FP, Lloyd RS, and Rizzo CJ (2012) Replication of the 2,6-diamino-4-hydroxy-*N*^δ-(methyl)-formamidopyrimidine (MeFapy-dGuo) adduct by eukaryotic DNA polymerases. Chem. Res. Toxicol 25, 1652–1661. [PubMed: 22721435]
- (52). Tomasz M, Lipman R, Lee MS, Verdine GL, and Nakanishi K (1987) Reaction of acid-activated mitomycin C with calf thymus DNA and model guanines: Elucidation of the base-catalyzed degradation of N7-alkylguanine nucleosides. Biochemistry 26, 2010–2027. [PubMed: 3109476]
- (53). Haraguchi K, and Greenberg MM (2001) Synthesis of oligonucleotides containing Fapy.dG (*N*⁶-(2-deoxy- α , β -D-*erythro*-pentofuranosyl)-2,6-diamino-4-hydroxy-5-formamidopyrimidine). J. Am. Chem. Soc 123, 8636–8637. [PubMed: 11525688]
- (54). Burgdorf LT, and Carell T (2002) Synthesis, stability, and conformation of the formamidopyrimidine G DNA lesion. Chem. - Eur. J 8, 293–301. [PubMed: 11822460]
- (55). Brown KL, Deng JZ, Iyer RS, Iyer LG, Voehler MW, Stone MP, Harris CM, and Harris TM (2006) Unraveling the aflatoxin-FAPY conundrum: Structural basis for differential replicative processing of isomeric forms of the formamidopyrimidine-type DNA adduct of aflatoxin B1. J. Am. Chem. Soc 128, 15188–15199. [PubMed: 17117870]
- (56). Busch F, Pieck JC, Ober M, Gierlich J, Hsu GW, Beese LS, and Carell T (2008) Dissecting the differences between the α and β anomers of the oxidative DNA lesion FaPydG. Chem. - Eur. J 14, 2125–2132. [PubMed: 18196510]
- (57). Christov PP, Banerjee S, Stone MP, and Rizzo CJ (2010) Selective incision of the α -N-methyl-formamidopyrimidine anomer by *Escherichia coli* endonuclease IV. J. Nucleic Acids 2010, 850234. [PubMed: 20798848]
- (58). Lukin M, Minetti CA, Remeta DP, Attaluri S, Johnson F, Breslauer KJ, and de Los Santos C (2011) Novel post-synthetic generation, isomeric resolution, and characterization of Fapy-dG within oligodeoxynucleotides: Differential anomeric impacts on DNA duplex properties. Nucleic Acids Res. 39, 5776–5789. [PubMed: 21415012]
- (59). Boiteux S, Belleney J, Roques BP, and Laval J (1984) Two rotameric forms of open ring 7-methylguanine are present in alkylated polynucleotides. Nucleic Acids Res. 12, 5429–5439. [PubMed: 6462910]
- (60). Humphreys WG, and Guengerich FP (1991) Structure of formamidopyrimidine adducts as determined by NMR using specifically ¹⁵N-labeled guanosine. Chem. Res. Toxicol 4, 632–636. [PubMed: 1807445]
- (61). Christov PP, Brown KL, Kozekov ID, Stone MP, Harris TM, and Rizzo CJ (2008) Site-specific synthesis and characterization of oligonucleotides containing an *N*⁶-(2-deoxy-D-*erythro*-pentofuranosyl)-2,6-diamino-3,4-dihydro-4-oxo-5-N-methyl-formamidopyrimidine lesion, the ring-opened product from N7-methylation of deoxyguanosine. Chem. Res. Toxicol 21, 2324–2333. [PubMed: 19053322]
- (62). Gruppi F, Hejazi L, Christov PP, Krishnamachari S, Turesky RJ, and Rizzo CJ (2015) Characterization of nitrogen mustard formamidopyrimidine adduct formation of bis(2-chloroethyl)ethylamine with calf thymus DNA and a human mammary cancer cell line. Chem. Res. Toxicol. 28, 1850–1860. [PubMed: 26285869]
- (63). Elmquist CE, Stover JS, Wang Z, and Rizzo CJ (2004) Site-specific synthesis and properties of oligonucleotides containing C8-deoxyguanosine adducts of the dietary mutagen IQ. J. Am. Chem. Soc 126, 11189–11201. [PubMed: 15355100]
- (64). Cavaluzzi MJ, and Borer PN (2004) Revised UV extinction coefficients for nucleoside-5'-monophosphates and unpaired DNA and RNA. Nucleic Acids Res. 32, 13e.
- (65). Piotto M, Saudek V, and Sklenar V (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. J. Biomol. NMR 2, 661–665. [PubMed: 1490109]
- (66). Marion D, Ikura M, Tschudin R, and Bax A (1989) Rapid recording of 2D NMR-spectra without phase cycling - Application to the study of hydrogen-exchange in proteins. J. Magn. Reson. 85, 393–399.

- (67). Jeener J, Meier BH, Bachmann P, and Ernst RR (1979) Investigation of exchange processes by 2-dimensional NMR spectroscopy. *J. Chem. Phys* 71, 4546–4553.
- (68). Wagner R, and Berger S (1996) Gradient-selected NOESY - A fourfold reduction of the measurement time for the NOESY experiment. *J. Magn. Reson., Ser. A* 123, 119–121. [PubMed: 8980072]
- (69). Aue WP, Bartholdi E, and Ernst RR (1976) Two-dimensional spectroscopy. Application to nuclear magnetic resonance. *J. Chem. Phys* 64, 2229–2246.
- (70). Bax A, Davis DG, and Sarkar SK (1985) An improved method for two-dimensional heteronuclear relayed-coherence-transfer NMR spectroscopy. *J. Magn. Reson* 63, 230–234.
- (71). Bodenhausen G, and Ruben DJ (1980) Natural abundance N-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett* 69, 185–189.
- (72). Willker W, Leibfritz D, Kerssebaum R, and Bermel W (1993) Gradient selection in inverse heteronuclear correlation spectroscopy. *Magn. Reson. Chem* 31, 287–292.
- (73). Shaka AJ, Barker PB, and Freeman R (1985) Computer-optimized decoupling scheme for wideband applications and low-level operation. *J. Magn. Reson.* 64, 547–552.
- (74). Mareci TH, and Freeman R (1982) Echoes and anti-echoes in coherence transfer NMR - Determining the signs of double-quantum frequencies. *J. Magn. Reson.* 48, 158–163.
- (75). Kontaxis G, Stonehouse J, Laue ED, and Keeler J (1994) The sensitivity of experiments which use gradient pulses for coherence-pathway selection. *J. Magn. Reson., Ser. A* 111, 70–76.
- (76). Doddrell DM, Pegg DT, and Bendall MR (1982) Distortionless enhancement of NMR signals by polarization transfer. *J. Magn. Reson* 48, 323–327.
- (77). Goddard TD, and Kneller DG (2006) SPARKY v. 3.113, University of California, San Francisco, CA.
- (78). Ober M, Linne U, Gierlich J, and Carell T (2003) The two main DNA lesions 8-Oxo-7,8-dihydroguanine and 2,6-diamino-5-formamido-4-hydroxypyrimidine exhibit strongly different pairing properties. *Angew. Chem., Int. Ed* 42, 4947–4951.
- (79). Coste F, Ober M, Le Bihan YV, Izquierdo MA, Hervouet N, Mueller H, Carell T, and Castaing B (2008) Bacterial base excision repair enzyme Fpg recognizes bulky N7-substituted-FapydG lesion via unproductive binding mode. *Chem. Biol* 15, 706–717. [PubMed: 18635007]
- (80). Patro JN, Haraguchi K, Delaney MO, and Greenberg MM (2004) Probing the configurations of formamidopyrimidine lesions Fapy dA and Fapy dG in DNA using endonuclease IV. *Biochemistry* 43, 13397–13403. [PubMed: 15491146]
- (81). Mao H, Deng Z, Wang F, Harris TM, and Stone MP (1998) An intercalated and thermally stable FAPY adduct of aflatoxin B1 in a DNA duplex: Structural refinement from ¹H NMR. *Biochemistry* 37, 4374–4387. [PubMed: 9521757]
- (82). Brown KL, Voehler MW, Magee SM, Harris CM, Harris TM, and Stone MP (2009) Structural perturbations induced by the a-anomer of the aflatoxin B₁ formamidopyrimidine adduct in duplex and single-strand DNA. *J. Am. Chem. Soc* 131, 16096–16107. [PubMed: 19831353]
- (83). Aramini JM, Cleaver SH, Pon RT, Cunningham R P., and Germann, M. W. (2004) Solution structure of a DNA duplex containing an a-anomeric adenosine: Insights into substrate recognition by endonuclease IV. *J. Mol. Biol* 338, 77–91. [PubMed: 15050824]
- (84). Johnson CN, Spring AM, Desai S, Cunningham RP, and Germann MW (2012) DNA sequence context conceals alpha-anomeric lesions. *J. Mol. Biol* 416, 425–437. [PubMed: 22227386]
- (85). Oki M (2007) Recent advances in atropisomerism. *Topics Stereochem.* 14, 1–81.
- (86). Raoul S, Bardet M, and Cadet J (1995) Gamma irradiation of 2'-deoxyadenosine in oxygen-free aqueous solutions: Identification and conformational features of formamidopyrimidine nucleoside derivatives. *Chem. Res. Toxicol* 8, 924–933. [PubMed: 8555407]
- (87). Li L, Brown KL, Ma RD, and Stone MP (2015) DNA sequence modulates geometrical isomerism of the trans-8,9-dihydro-8-(2,6-diamino-4-oxo-3,4-dihydropyrimidin-5-yl-formamido)-9-hydroxy aflatoxin B₁ adduct. *Chem. Res. Toxicol* 28, 225–237. [PubMed: 25587868]
- (88). Ober M, Muller H, Pieck C, Gierlich J, and Carell T (2005) Base pairing and replicative processing of the formamidopyrimidine-dG DNA lesion. *J. Am. Chem. Soc* 127, 18143–18149. [PubMed: 16366567]

- (89). Lukin M, Zaliznyak T, Attaluri S, Johnson F, and de Los Santos C (2012) Solution structure of duplex DNA containing a beta-carba-Fapy-dG lesion. *Chem. Res. Toxicol* 25, 2423–2431. [PubMed: 22897814]
- (90). Asagoshi K, Yamada T, Okada Y, Terato H, Ohyama Y, Seki S, and Ide H (2000) Recognition of formamidopyrimidine by *Escherichia coli* and mammalian thymine glycol glycosylases. Distinctive paired base effects and biological and mechanistic implications. *J. Biol. Chem* 275, 24781–24786. [PubMed: 10827172]
- (91). Asagoshi K, Yamada T, Terato H, Ohyama Y, Monden Y, Arai T, Nishimura S, Aburatani H, Lindahl T, and Ide H (2000) Distinct repair activities of human 7,8-dihydro-8-oxoguanine DNA glycosylase and formamidopyrimidine DNA glycosylase for formamidopyrimidine and 7,8-dihydro-8-oxoguanine. *J. Biol. Chem* 275, 4956–4964. [PubMed: 10671534]
- (92). Wiederholt CJ, Delaney MO, Pope MA, David SS, and Greenberg MM (2003) Repair of DNA containing Fapy.dG and its beta-C-nucleoside analogue by formamidopyrimidine DNA glycosylase and MutY. *Biochemistry* 42, 9755–9760. [PubMed: 12911318]
- (93). Coste F, Ober M, Carell T, Boiteux S, Zelwer C, and Castaing B (2004) Structural basis for the recognition of the FapydG lesion (2,6-diamino-4-hydroxy-5-formamidopyrimidine) by formamidopyrimidine-DNA glycosylase. *J. Biol. Chem.* 279, 44074–44083. [PubMed: 15249553]
- (94). Dizdaroglu M, Kirkali G, and Jaruga P (2008) Formamidopyrimidines in DNA: Mechanisms of formation, repair, and biological effects. *Free Radical Biol. Med* 45, 1610–1621. [PubMed: 18692130]
- (95). Kalam MA, Haraguchi K, Chandani S, Loechler EL, Moriya M, Greenberg MM, and Basu AK (2006) Genetic effects of oxidative DNA damages: Comparative mutagenesis of the imidazole ring-opened formamidopyrimidines (Fapy lesions) and 8-oxo-purines in simian kidney cells. *Nucleic Acids Res.* 34, 2305–2315. [PubMed: 16679449]
- (96). Krahn JM, Beard WA, Miller H, Grollman AP, and Wilson SH (2003) Structure of DNA polymerase beta with the mutagenic DNA lesion 8-oxodeoxyguanine reveals structural insights into its coding potential. *Structure* 11, 121–127. [PubMed: 12517346]
- (97). Hsu GW, Ober M, Carell T, and Beese LS (2004) Errorprone replication of oxidatively damaged DNA by a high-fidelity DNA polymerase. *Nature* 431, 217–221. [PubMed: 15322558]
- (98). Briebe LG, Eichman BF, Kokoska RJ, Double S, Kunkel TA, and Ellenberger T (2004) Structural basis for the dual coding potential of 8-oxoguanosine by a high-fidelity DNA polymerase. *EMBO J.* 23, 3452–3461. [PubMed: 15297882]
- (99). Eoff RL, Irimia A, Angel KC, Egli M, and Guengerich FP (2007) Hydrogen bonding of 7,8-dihydro-8-oxodeoxyguanosine with a charged residue in the little finger domain determines miscoding events in *Sulfolobus solfataricus* DNA polymerase Dpo4. *J. Biol. Chem* 282, 19831–19843. [PubMed: 17468100]
- (100). Vasquez-Del Carpio R, Silverstein TD, Lone S, Swan MK, Choudhury JR, Johnson RE, Prakash S, Prakash L, and Aggarwal AK (2009) Structure of human DNA polymerase k inserting dATP opposite an 8-OxoG DNA lesion. *PLoS One* 4, e5766. [PubMed: 19492058]
- (101). Batra VK, Beard WA, Hou EW, Pedersen LC, Prasad R, and Wilson SH (2010) Mutagenic conformation of 8-oxo-7,8-dihydro-2'-dGTP in the confines of a DNA polymerase active site. *Nat. Struct. Mol. Biol* 17, 889–890. [PubMed: 20526335]
- (102). Wiederholt CJ, and Greenberg MM (2002) Fapy.dG instructs Klenow exo- to misincorporate deoxyadenosine. *J. Am. Chem. Soc* 124, 7278–7279. [PubMed: 12071730]
- (103). Patro JN, Wiederholt CJ, Jiang YL, Delaney JC, Essigmann JM, and Greenberg MM (2007) Studies on the replication of the ring opened formamidopyrimidine, Fapy.dG in *Escherichia coli*. *Biochemistry*, Vol. 46, pp 10202–10212. [PubMed: 17691820]
- (104). Gehrke TH, Lischke U, Gasteiger KL, Schneider S, Arnold S, Muller HC, Stephenson DS, Zipse H, and Carell T (2013) Unexpected non-Hoogsteen-based mutagenicity mechanism of FaPy-DNA lesions. *Nat. Chem. Biol* 9, 455–461. [PubMed: 23685671]
- (105). Patra A, Banerjee S, Johnson Salyard TL, Malik CK, Christov PP, Rizzo CJ, Stone MP, and Egli M (2015) Structural basis for error-free bypass of the 5-N-methylformamidopyr-imidine-dG

lesion by human DNA polymerase ϵ and *Sulfolobus solfataricus* P2 polymerase IV. J. Am. Chem. Soc 137, 7011–7014. [PubMed: 25988947]

- (106). Minko IG, Rizzo CJ, and Lloyd RS (2017) Mutagenic potential of nitrogen mustard-induced formamidopyrimidine DNA adduct: Contribution of the non-canonical α -anomer. J. Biol. Chem 292, 18790–18799. [PubMed: 28972137]
- (107). Tano K, Shiota S, Collier J, Foote RS, and Mitra S (1990) Isolation and structural characterization of a cDNA clone encoding the human DNA repair protein for O^6 -alkylguanine. Proc. Natl. Acad. Sci U. S. A 87, 686–690. [PubMed: 2405387]
- (108). Hegi ME, Diserens AC, Gorlia T, Hamou MF, de Tribolet N, Weller M, Kros JM, Hainfellner JA, Mason W, Mariani L, Bromberg JE, Hau P, Mirimanoff RO, Cairncross JG, Janzer RC, and Stupp R (2005) MGMT gene silencing and benefit from Temozolomide in glioblastoma. N. Engl. J. Med 352, 997–1003. [PubMed: 15758010]
- (109). Aldape K, Zadeh G, Mansouri S, Reifenberger G, and von Deimling A (2015) Glioblastoma: Pathology, molecular mechanisms and markers. Acta Neuropathol 129, 829–848. [PubMed: 25943888]
- (110). Henne T, and Schmahl D (1985) Occurrence of second primary malignancies in man—a second look. Cancer Treat. Rev 12, 77–94.
- (111). Rodriguez MA, Fuller LM, Zimmerman SO, Allen PK, Brown BW, Munsell MF, Hagemeister FB, McLaughlin P, Velasquez WS, Swan F, Jr., et al. (1993) Hodgkin's disease: Study of treatment intensities and incidences of second malignancies. Ann. Oncol 4, 125–131. [PubMed: 8448080]

Figure 1.

^{13}C HSQC spectra of single strand and duplex oligodeoxynucleotides modified site-specifically with ^{13}C -MeFapy-dG, showing unique configurational and conformational species of the MeFapy-dG lesion in DNA. (A) 5'-TXT-3' trimer (X = ^{13}C -MeFapy-dG). (B) 5'-CATXATGACGCT-3' dodecamer (X = ^{13}C -MeFapy-dG). (C) 5'-CATXATGACGCT-3':5'-AGCGTCATCATG-3' dodecamer duplex. The separation of the ^{13}C -labeled methyl peaks into two groups is seen for both the ^{13}C dimension and the ^1H dimension in the HSQC spectra.

Figure 2. Detection of scalar $^3J_{C-H}$ coupling between the ^{13}C -labeled carbon and the formyl proton of the 5'-TXT-3' trimer ($X = ^{13}C$ -MeFapy-dG). (A) 150.9 MHz 1H -decoupled DEPT-135 ^{13}C spectrum. (B) 150.9 MHz non- 1H -decoupled DEPT-135 ^{13}C spectrum. The $^3J_{C-H}$ values of the upfield and downfield groups of methyl peaks were measured as 4 Hz and 1.7 Hz, respectively. The temperature was 278 K.

Figure 3.

Epimerization of the deoxyribose α and β anomers. (A) The α anomers had stronger H1'–H3' (Dipolar coupling IV) than H1'–H4' (Dipolar coupling III) NOEs. The α anomers were also identified by NOEs between N⁶H and deoxyribose H4' (Dipolar coupling II). The ³J coupling between the MeFapy-dG N⁶H protons and the deoxyribose H1' protons (Scalar coupling I) allowed the assignment of the former from COSY and TOCSY spectra (not shown). No NOE cross peaks were observed between the N⁶H and the deoxyribose H3' protons. (B) The β anomers were identified due to their strong deoxyribose H1'–H4' NOEs (Dipolar coupling IV) and weaker deoxyribose H1'–H3' NOEs (Dipolar coupling III). The β anomers exhibited strong N⁶H–H3' NOEs (Dipolar coupling II), and weak N⁶H–H4' NOEs. Again, ³J coupling between the MeFapy-dG N⁶H protons and the deoxyribose H1' protons (Scalar coupling I) allowed the assignment of the former from COSY and TOCSY spectra (not shown). The NOESY spectrum of MeFapy-dG in the 5'-TXT-3' sequence was obtained in buffer containing 9:1 H₂O:D₂O, and 278 K.

Figure 4.

Expanded NOESY spectrum of MeFapy-dG in the 5'-TXT-3' trimer showing NOEs between the CH₃ protons and the formyl protons arising from specific conformational and configurational species of the lesion. The NOESY spectrum was obtained in buffer containing 9:1 H₂O:D₂O, and 278 K.

Scheme 1.

(A) In the Presence of DNA, Temozolomide (TMZ) and Other DNA Methylating Agents Predominantly Form N7-MedG Lesion, Which Carries Positive Charge on Guanine Imidazole Ring. Initially Formed N7-MedG Lesion Can Then Undergo Hydroxide Addition and Ring-Opening into MeFapy-dG Lesion; (B) MeFapy-dG Lesion Can Undergo Hindered Rotation about the C⁵–N⁶ Bond, Formyl Bond (Red), and Glycosidic Bond (Blue); (C) MeFapy-dG Lesion Interconverts between α and β Anomers

Scheme 2.

Geometrical Isomers of MeFapy-dG Formyl Bond and Atropisomers of MeFapy-dG C5–N⁵ bond^a

^a*E* geometrical isomers of the formyl bond exhibit a greater $^3J_{\text{CH}}$ coupling constant than do *Z* geometrical isomers. *R_a* and *S_a* isomers arise when rotation occurs around the C5–N⁵ bond.

Table 1.

Assignment of R_a and S_a Atropisomers about the C5–N⁵ Bond (Scheme 2) Was Made Based on Comparative NOE Intensities between N⁶H Proton and Methyl and Aldehyde Protons of MeFapy-dG Adduct^a

¹³ C isomeric species	N ⁶ H–CH ₃ NOE intensity relative to atropomer population ^b	N ⁶ H–CHO NOE intensity relative to atropomer population ^b	integrated intensity of ¹ H resonance ^c	assignment of MeFapy-dG C5–N ⁵ atropisomer
β Z isomer 1	57%	45%	50%	S_a
β Z isomer 2	43%	55%	50%	R_a
β E isomer 1	51%	56%	54%	S_a
β E isomer 2	49%	44%	46%	R_a
α Z isomer 1	62%	53%	58%	R_a
α Z isomer 2	38%	47%	42%	S_a
α E isomer 1	22%	45%	38%	R_a
α E isomer 2	78%	55%	62%	S_a

^aIt was necessary to correct these NOE intensities for the relative populations of each isomer, as determined from integration of ¹H NMR resonances of the MeFapy-dG N⁶H protons. For the S_a atropisomer, the N⁶H–Me NOE exhibited greater intensity, whereas for the R_a isomer, the N⁶H–CHO NOE exhibited greater intensity.

^bDetermined from a NOESY experiment in buffer containing 9:1 H₂O:D₂O.

^cDetermined from integration of the N⁶ proton in a ¹H spectrum in buffer containing 9:1 H₂O:D₂O.

Table 2.

Calculated Populations of Eight Major Rotational and Configurational Isomers for MeFapy-dG Adduct in 5'-TXT-3' Trimer^a

MeFapy-dG isomer	relative abundance in 5'-TXT-3' sequence ^b
β <i>E S_a</i>	14%
β <i>E R_a</i>	13%
β <i>Z S_a</i>	21%
β <i>Z R_a</i>	21%
α <i>E S_a</i>	12%
α <i>E R_a</i>	7%
α <i>Z S_a</i>	5%
α <i>Z R_a</i>	7%

^aSchemes 1 and 2. The α and β anomers were present in a ratio of 3:7. For the β anomer, the *Z* geometrical isomer of the formyl group was favored, while for the α anomer, the *E* geometrical isomer of the formyl group was favored.

^bDetermined from a ¹H spectrum.