Solution Structure of Duplex DNA Containing the Mutagenic Lesion 1,N^2-Etheno-2′-deoxyguanine[†,‡]

Tanya Zaliznyak, Mark Lukin, Francis Johnson, and Carlos de los Santos*

Department of Pharmacological Sciences, Stony Brook University, School of Medicine, Stony Brook, New York 11794-8651

Received November 12, 2007; Revised Manuscript Received January 30, 2008

ABSTRACT: We have used high-resolution NMR spectroscopy and molecular dynamics simulations to determine the solution structure of DNA containing the genotoxic lesion 1,N^2-etheno-2′-deoxyguanosine (εG), paired to dC. The NMR data suggest the presence of a major, minimally perturbed structure at neutral pH. NOESY spectra indicate the presence of a right-handed helix with all nucleotides in *anti*, 2′-deoxyribose conformations within the C2′-endo/C1′-exo range and proper Watson–Crick base pair alignments outside the lesion site. The εG residue remains deeply embedded inside the helix and stacks between the flanking base pairs. The lesion partner dC is extrahelical and is located in the minor groove of the duplex, where it is highly exposed to solvent. Upon acidification of the sample, a second conformation at the lesion site of the duplex emerges, with protonation of the lesion partner dC and possible formation of a Hoogsteen base pair. Restrained molecular dynamics simulations of the neutral-pH structure generated a set of three-dimensional models that show εG inside the helix, where the lesion is stabilized by stacking interactions with flanking bases but without participating in hydrogen bonding. The lesion counterbase dC is displaced in the minor groove of the duplex where it can form a hydrogen bond with the sugar O4′ atom of a residue 2 bp away.

The industrial use of polyvinyl chloride (PVC), the building block in the production of the polymer plastics, continues to grow. Vinyl chloride is a well-established environmental hazard whose exposure is associated with a rare form of human cancer, angiosarcoma of the liver (*1*). It generates various types of DNA lesions that produce base pair substitutions as well as other genotoxic actions, including sister chromatid exchange, gene conversion events, and chromosomal aberrations (*2*). The metabolic oxidation of vinyl chloride by CYP2E1 generates 2-chloroacetaldehyde and 2-chloroethylene oxide, compounds that may react with DNA to form etheno lesions, such as 3,N^4-ethenocytosine (εdC), 1,N^6-ethenoadenine (εdA), N^2,3-ethenoguanine (N^2,3-εdG), and 1,N^2-ethenoguanine (1,N^2-εdG) (*3–6*). The same adducts also arise from the reaction of endogenous lipid peroxidation products with nucleobases (*7, 8*), suggesting that etheno bases could be a generic risk factor in sporadic carcinogenesis. The number of etheno adducts formed in rodent tissues after exposure to vinyl chloride varies significantly, and 1,N^2-εdG has still to be detected in vivo (*9*). In contrast, treatment of calf thymus DNA with ethyl linoleate or 4-hydroxynonenal generates significant levels of 1,N^2-εdG, levels even higher than those of N^2,3-εdG, indicating that the former lesion would play a predominant role in carcinogenesis derived from endogenous oxidative processes (*10*).

All etheno adducts are genotoxic, causing mostly base substitution mutations. 1,N^2-εdG leads to G → T, G → C, and G → A substitutions in bacteria, while it triggers G → A and G → T mutations in mammalian cells (*11, 12*). In addition, it can generate double mutants, as well as one- and two-base deletions and rearrangements near the lesion site (*12*). Repair activities present in prokaryotic and eukaryotic cells remove etheno bases from DNA, primarily by base excision repair (BER). In *Escherichia coli*, the mismatch-specific uracil glycosylase (MUG) incises DNA containing εdC or 1,N^2-εdG lesions and AlkA operates on N^2,3-εdG and εdA (*13–15*). Mammalian alkyl-*N*-purine-DNA glycosylase (ANPG) can remove εdA and 1,N^2-εdG lesions, and mismatch-specific thymine-DNA glycosylase (hTDG) deals with εdC adducts (*13, 16–18*). The structural determinants that mediate the recognition of a given etheno lesion by a specific DNA glycosylase remain poorly understood.

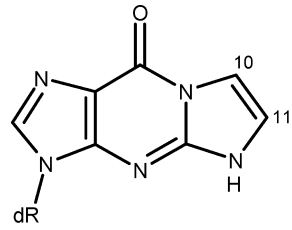
The solution structures of duplexes containing a single εdA or εdC adduct paired to their respective Watson–Crick partners, or in the context of different mutagenic intermediates, show that the lesions induce only a local and relatively minor perturbation of the B-form conformation, without disrupting the Watson–Crick alignment on lesion-flanking base pairs. In all these duplexes, the etheno residue remains inside the helix where it is stabilized by hydrogen bond formation and stacking interactions with lesion-flanking base pairs (*19–24*). Recently, Shanmugam and co-workers reported the conformation of duplex DNA containing the 1,N^2-εdG lesion (*25, 44*). At pH 5.2, εdG adopts a *syn* conformation forming a Hoogsteen base pair with its protonated

[†] NIH Grants CA47995 and CA77094 supported this research.

[‡] This structure was initially presented at the 232nd National Meeting of the American Chemical Society, Division of Chemical Toxicology, San Francisco CA, September 10–14 2006 Atom coordinates of the εdG·dC duplex are deposted in the Protein Data Bank, accession 2kly.

* To whom correspondence should be addressed. Phone: (631) 444-3649. Fax: (631) 444-3218. E-mail: cds@pharm.sunysb.edu.

10.1021/bi7022514 CCC: $40.75
Published on Web 03/29/2008



1,*N*²-etheno-dG (εG)

C1 G2 T3 A4 C5 εG6 C7 A8 T9 G10 C11
G22 C21 A20 T19 G18 C17 G16 T15 A14 C13 G12

FIGURE 1: Chemical structure of 1,N^2-ethenodeoxyguanosine and numbering scheme of the duplex used in this study.

dC^+ partner. At basic pH, 1,N^2-εdG remains *anti* and stacks between flanking bases, while displacing its counterbase in the major groove of the duplex. At neutral pH, NMR[1] signals are broad, presumably due to the exchange between the acidic and basic conformations, a fact that precluded determination of the duplex structure.

We report here the solution structure of a d(C-G-T-A-C-εG-C-A-T-G-C)·d(G-C-A-T-G-C-G-T-A-C-G) oligodeoxynucleotide duplex (throughout this paper called the εdG duplex), as determined by high-resolution NMR spectroscopy and restrained molecular dynamics (MD) simulations. At pH 6.9 and 27 °C, the εdG duplex adopts a major conformation, which has common features and also some important differences with the structure determined by Shanmugam and co-workers at basic pH, with the lesion present in the d(...T-εG-G...) sequence. The chemical structure of the 1,N^2-εdG adduct and the sequence and numbering of the duplex used in our study are shown in Figure 1.

MATERIALS AND METHODS

Synthesis and Purification of 2′-Deoxynucleotide Oligomers. The 11-mer oligonucleotide containing a single 1,N^2-ethenodeoxyguanosine lesion was prepared following a postsynthetic modification approach previously reported (*26*). Briefly, we prepared an N^2-protected deoxyguanosine having a 2,3-diacetoxypropane chain at the N1 position. Following conversion of this nucleoside to the 5′-*O*-DMT-3′-*O*-phosphoramidite derivative, we employed standard solid-phase synthesis to synthesize the deoxynucleotide oligomer containing the modified dG at its center. After cleavage of synthesis products from the solid support by treatment of the resin with a concentrated ammonium solution, sodium periodate was used to oxidize the diol moiety, generating the εdG-containing oligonucleotide. The unmodified oligodeoxynucleotide was synthesized using standard solid-phase chemical methods. Damaged and undamaged oligodeoxynucleotides were purified by HPLC methods following protocols previously described (*27*). Oligomer composition was confirmed by electrospray ionization mass spectroscopy using a Micromass Quattro LC system. Oligodeoxynucleotides were desalted using a Sephadex G-25 column and converted to the sodium salt by percolation through a Dowex 50W ion-exchange resin.

Duplex Formation and Sample Preparation. Proper strand stoichiometry was obtained by following the intensity of NMR proton signals during gradual addition of the unmodified strand to the εdG-containing strand, at 55 °C. The typical NMR sample consisted of approximately 300 OD_{260} units of duplex dissolved in 0.6 mL of 10 mM phosphate buffer (pH 6.9) containing 50 mM NaCl and 1 mM EDTA, for a final duplex concentration of approximately 4.2 mM. The sample pH was measured and adjusted to the desired value before exchange with D_2O. The duplex sample was lyophilized and dissolved in 99.96% D_2O (D_2O buffer) before NMR data collection. In addition, the sample was lyophilized and dissolved in a 90% H_2O/10% D_2O mixture (H_2O buffer) for collection of the exchangeable proton spectra. The pH dependence study started with the sample dissolved at neutral pH followed by smooth acidification caused by the addition of microliter amounts of a 0.5 M HCl solution. Reported pH values are meter readings at room temperatures.

NMR Experiments. One- and two-dimensional NMR spectra were recorded using Varian (Inova) spectrometers operating at a field strengths of 11.75 or 14.1 T. Proton chemical shifts were referenced relative to sodium 3-(trimethylsilyl)propionate-2,2,3,3-d_4 at 0 ppm. Phase-sensitive (*28*) NOESY (50, 150, 200, and 300 ms mixing times), TOCSY (70 and 120 ms mixing times), COSY, and DQF-COSY spectra in D_2O buffer were recorded at 27 °C. The residual water signal was suppressed by presaturation during the repetition delay of 1.5 s. Phase-sensitive proton NOESY spectra (120 and 220 ms mixing times) in H_2O buffer were recorded at 5 °C, using a "jump and return" reading pulse (*29*). NMR data were processed and analyzed using Felix (Accelrys Inc., San Diego, CA) running on Silicon Graphics workstations. Data sets consisted of 2048 and 300 complex data points in the t_2 and t_1 dimensions, respectively. Prior to Fourier transformation, shifted sine-bell window functions were used to smooth the time domain data and prevent discontinuation artifacts at the last point. No baseline correction was applied to the frequency domain data. Residual dipolar couplings (RDCs) were calculated at 800 MHz using a combination of gradient-enhanced $^1H-^{13}C$ HSQC and gradient-selected TROSY experiments (*30*). The liquid crystal phase consisted of 5% of a pentaethylene glycol monododecyl ether/*n*-hexanol mixture.

Refinement of the Duplex Structure. Restrained MD simulations were performed using X-PLOR 3.1 (*31*) running on Silicon Graphics computers. Simulations were run *in vacuo* using an all-atom force field derived from CHARMM (*32*) with the dielectric constant set to 4 (*33*). Partial charges on the etheno bridge atoms were derived by analogy with our previous studies with εdC and εdA (*19, 21, 22*), resulting in a neutral εdG base. Partial atomic charges on phosphate groups were not reduced, resulting in deoxynucleotide residues with a net charge of −1. Interproton distances were calculated by subjecting a canonical B-form duplex structure to 1000 steps of potential energy minimization, using a sole penalty function proportional to the difference between back-calculated and experimental NOE intensities (*34*). Interproton distances were extracted from the atom coordinates of the energy-minimized model. NOE cross-peak intensities were measured from all NOESY (50, 150, 200,

[1] Abbreviations: NMR, nuclear magnetic resonance; EDTA, disodium ethylenediaminetetraacetate; TSP, sodium 3-trimethylsilyl propionate-2,2,3,3-d_4; NOESY, nuclear Overhauser effect spectroscopy; COSY, correlation spectroscopy; DQF-COSY, double-quantum-filtered correlation spectroscopy; TOCSY, total correlation spectroscopy; NOE, nuclear Overhauser effect.

and 300 ms mixing times) spectra recorded in D_2O buffer using Felix and used for distance calculations. A grid search showed that 1.56 ns was the best-fit isotropic correlation time for calculation of the full relaxation matrix, and this value was used for distance computations. Five hundred sixty-one experimental distances were enforced during the molecular dynamics calculations by using square-well potential energy functions with boundaries of ±0.6 Å to the values computed from the NOESY spectra. Similarly, boundaries of ±0.9 Å were used for interproton distances derived from cross-peaks only present in the 300 ms mixing time NOESY spectrum or from overlapping NOE peaks. Our molecular dynamics protocol consisted of a heating step, where the temperature of the system was increased from its initial value to 500 K in 20 ps, followed by a high-temperature step, during which time the temperature of the simulation remained constant at 500 K. After the high-temperature step, the system was cooled to 300 K over a period of 20 ps and equilibrated at this temperature for an additional 40 ps of molecular dynamics. Experimental distance restraints were softly introduced during the heating step of the simulations, using a penalty constant that gradually increased to its maximum value of 300 kcal mol^{-1} $Å^{-2}$ and remained enforced until completion of the runs. Watson–Crick hydrogen bonds were enforced on all canonical base pairs of the duplex, including those flanking the εdG•C base pair, using distance-dependent potential energy functions. Backbone dihedral angles were restrained by an empirical square-well potential function with a width encompassing the range of A- and B-form DNA conformations. Sugar puckers and glycosidic torsion angles were not restrained. Twenty structures were computed using five initial temperatures (140, 145, 150, 155, and 160 K) and four different high-temperature step lengths (35, 40, 45, and 55 ps). The last atomic coordinate frame of the simulations was energy-minimized, generating the ensemble of distance-refined structures. A set of seven refined models converged with a pairwise root-mean-square deviation (rmsd) of atom position of <1.5 Å. They were averaged and energy-minimized, producing the average three-dimensional model presented here. Refined models were analyzed using Chimera (*35*) and structural parameters obtained with Curves (*36*).

RESULTS

NMR Characterization of the εdG Duplex. The one-dimensional proton spectrum of the εdG duplex recorded at 27 °C in D_2O buffer (pH 6.9) (Figure 1S, Supporting Information) shows both relatively sharp and partially resolved proton signals, suggesting that the sample is amenable for complete structural studies. Assignment of the nonexchangeable proton signals resulted from the analysis of NOESY, COSY, and TOCSY spectra recorded at the same temperature following standard procedures (*37, 38*). Figure 2 shows the fingerprint region of a 500 MHz NOESY (300 ms mixing time) spectrum recorded with the εdG duplex dissolved in D_2O buffer (pH 6.9). Characteristic of a right-handed helix, each base proton (purine H8 or pyrimidine H6) of the duplex exhibits NOE interactions with the H1′ sugar protons of the same and 5′-flanking nucleotides. Similarly, each adenine H2 shows intrastrand NOE peaks with the H1′ protons of the same and 3′-flanking residues as well as interstrand cross-peaks with the H1′ proton of its Watson–Crick partner and that of 3′-flanking nucleotides (Figure 2, peaks A−E and G−M).

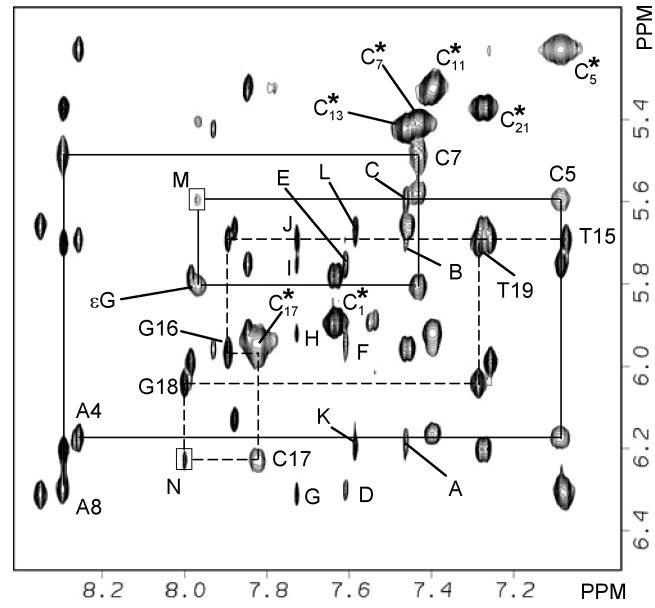


FIGURE 2: Fingerprint region of a 500 MHz NOESY (300 ms mixing time) spectrum recorded at 27 °C, with the sample dissolved in D_2O buffer (pH 6.9), depicting interactions at the central 5 bp segment of the εdG duplex. Each base proton (purine H8 or pyrimidine H6) exhibits NOE cross-peaks to its own and 5′-flanking H1′ sugar protons on the lesion-containing (solid lines) and unmodified (dashed lines) strands. Labels identify intraresidue base−H1′ cross peaks and asterisks intraresidue cytosine H5−H6 interactions, and other labels are assigned as follows: (A) dA4 H2−dA4 H1′ and dA4 H2−dA20 H1′, (B) dA4 H2−T19 H1′, (C) dA4 H2−dC5 H1′, (D) dA8 H2−dA8 H1′, (E) dA8 H2−T9 H1′, (F) dA8 H2−dC17 H5, (G) dA14 H2−dA14 H1′, (H) dA14 H2−dG10 H1′, (I) dA14 H2−T9 H1′, (J) dA14 H2−T15 H1′, (K) dA20 H2−dA20 H1′ and dA20 H2−dA4 H1′, (L) dA20 H2−T3 H1′ and dA20 H2−dC21 H1′, (M) εdG H8−dC5 H1′, and (N) dG18 H8−dC17 H1′.

The presence of the εdG lesion causes only minor perturbations in this region of the NOESY spectrum. The intensity of the dC5 H1′−εdG H8 and dC17 H1′−G18 H8 NOE peaks is rather weak (Figure 2, boxed peaks M and N), a fact that could reflect longer distances between these protons at the 5′-side of the lesion or increased dynamics at the lesion site. A second unusual feature is the chemical shift value of dC17 H5, H6, and H1′ proton signals, which are downfield from the values observed for other cytosines of the duplex (Table 1S, Supporting Information). In addition, the dC17 H5 proton signal displays NOE connectivity to the dA8 H2 proton (Figure 2, peak F), which is located in the minor groove of the duplex. This latter observation suggests that dC17 is extruded into the minor groove of the duplex and oriented toward the 3′-end of the lesion-containing strand. In agreement with the C17 nonstacked position is the observation that the RDC of the C_6-H_6 interaction is 20 Hz, a value that is quite different from those measured for all other bases of the εdG duplex (Table 2S,Supporting Information). The extrahelical location of C17 would increase solvent exposure of this residue, explaining the downfield shift of its H5, H6, and H1′ proton signals.

Assignment of the εdG H10 and εdG H11 protons followed the analysis of TOCSY spectra recorded in D_2O buffer (pH 6.9) at 27 °C. As shown in Figure 2S (Supporting Information), there is only one TOCSY cross-peak between

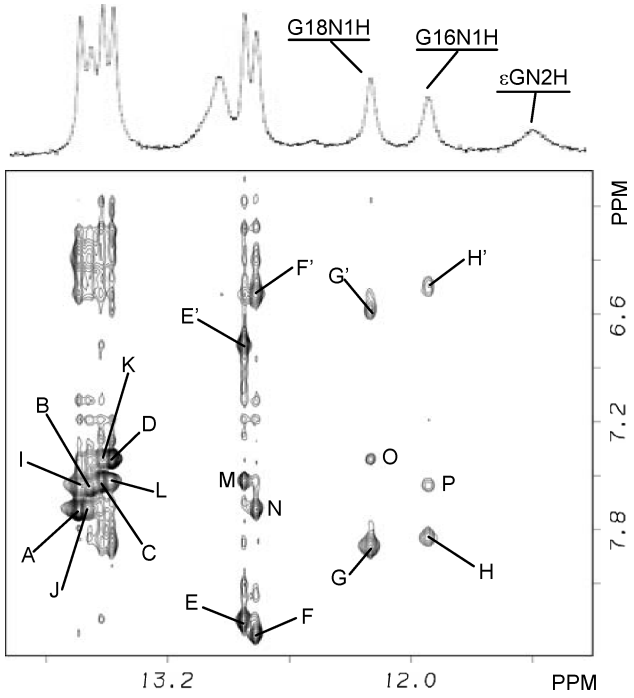


FIGURE 3: Expanded counter plot of a 500 MHz NOESY (120 ms mixing time) spectrum, recorded with the εdG duplex dissolved in H_2O buffer (pH 6.9) at 5 °C, displaying interactions between imino and amino/base protons. Labeled peaks are assigned as follows: (A) T9 N3H−dA14 H2, (B) T15 N3H−dA8 H2, (C) T3 N3H−dA20 H2, (D) T19 N3H−dA4 H2, (E) dG2 N1H−dC21 $N4H_b$, (E′) dG2 N1H−dC21 $N4H_{nb}$, (F) dG10 N1H−dC13 $N4H_b$, (F′) dG10 N1H−dC13 $N4H_{nb}$, (G) dG18 N1H−dC5 $N4H_b$, (G′) dG18 N1H−dC5 $N4H_{nb}$, (H) dG18 N1H−dC5 $N4H_b$, (H′) dG18 N1H−dC5 $N4H_{nb}$, (I) T9 N3H−dA8 H2, (J) T15 N3H−dA14 H2, (K) T3 N3H−dA4 H2, (L) T19 N3H−dA20 H2, (M) dG2 N1H−dA20 H2, (N) dG10 N1H−dA14 H2, (O) dG18 N1H−dA4 H2, and (P) dG16 N1H−dA8 H2.

protons resonating at 6.46 and 6.12 ppm, in the symmetrical H1′ sugar proton region. In addition, there are no cross-peaks in the aromatic region of the TOCSY spectrum, indicating that the etheno protons are outside of this region. The signals at 6.46 and 6.12 ppm are somewhat broad, generating a relatively weak but evident NOE cross-peak (Figure 2S, Supporting Information), but they lack additional interactions with other protons of the εdG duplex. Table 1S (Supporting Information) lists chemical shift values of the nonexchangeable protons of the εdG duplex at 27 °C and pH 6.9.

Nine partially resolved signals resonate in the 11.90–13.70 ppm region of a one-dimensional spectrum recorded in H_2O buffer (pH 6.9) at 5 °C (Figure 3, top), accounting for all the imino protons of the εdG duplex, with those belonging to the terminal base pairs overlapping at 12.97 ppm. Assignment of the exchangeable protons follows the analysis of the NOESY (120 ms mixing time) spectrum recorded in H_2O buffer, at 5 °C (Figure 3, bottom). Each thymine imino proton exhibits a strong NOE cross-peak to the H2 proton of its adenine partner, within dA·T base pairs of the duplex (Figure 3, peaks A−D). Similarly, guanine imino protons display strong NOE cross-peaks to the hydrogen-bonded and exposed cytosine amino protons, within the dG·C base pairs (Figure 3, peaks E/E′−H/H′). In addition to these intra-base pair interactions, cross-peaks between adenine H2 and imino protons of flanking base pairs (Figure 3, peaks I−P) and among sequential imino protons (data not shown) suggest proper base stacking throughout the εdG duplex. Chemical

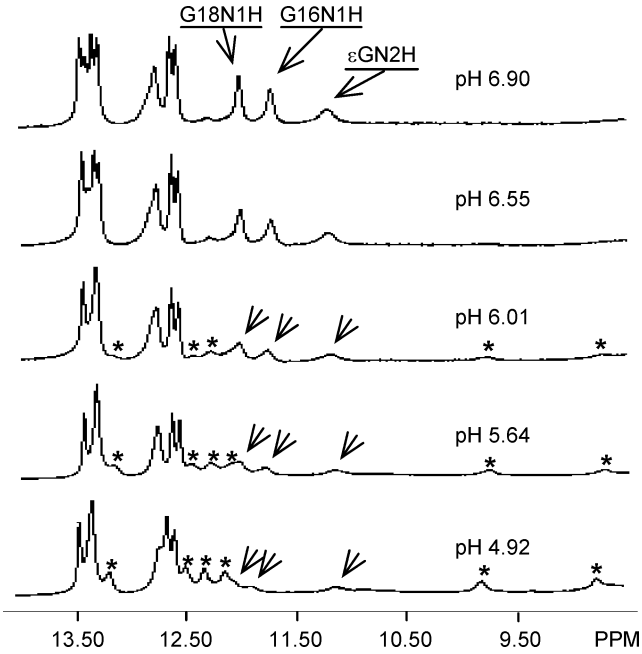


FIGURE 4: pH dependence of the exchangeable protons recorded at 5 °C. Labels identify the lesion site exchangeable proton signals, and asterisks denote resonances observed under acidic conditions.

shift values of the exchangeable protons, recorded at 5 °C and pH 6.9, are also listed in Table 1S (Supporting Information).

Figure 4 shows changes on the exchangeable proton spectrum that occur upon acidification of the sample. At pH 6.01, εdG N2H and the imino proton signals of residue 18 and dG16 become broader than at pH 6.90, suggesting the existence of chemical exchange at the lesion site. At this pH, three new signals appear in the imino proton region of the spectrum and two additional resonances are evident at 9.96 and 8.91 ppm (Figure 4, center spectrum, peaks marked with asterisks). At pH 4.92, resonances of the acid form become stronger and more evident while the εdG N2H, dG18 N1H, and dG16 N1H signals detected at neutral pH are very weak and broad (Figure 4, bottom spectrum). Since the signals originating from the acidic form are rather broad, NOESY spectra recorded in H_2O buffer (pH 4.92) failed to produce useful information and, as a result, proton assignment for the εdG duplex under acidic conditions was not possible.

Solution Structure of the εdG Duplex. The experimental set of NMR restraints drove our computer simulations to a family of converging models, within the B-form DNA conformation, without severe deviations from the idealized covalent geometry (Table 1). Figure 3S (Supporting Information) shows a superposition of the seven models, which have a pairwise rmsd of <1.5 Å (<0.9 Å with respect to the average), whereas Figure 5 displays the energy-minimized average structure seen with the minor group prominent. The εdG duplex is a regular right-handed helix, with non-lesion site residues in the *anti* conformation around their glycosidic torsion angles, sugar conformations in the C2′-endo/C1′-exo range, and Watson–Crick alignments present on all unmodified base pairs, including the lesion-flanking dC5·G18 and dC7·G16 base pairs.

Figures 6 and 7 show close-up views of the lesion site structure, and Table 1 lists some relevant structural param-

Table 1: εdG Duplex Structure

Deviations from Experimental Restraints[c]	
rmsd for NOE distances (Å)	0.10
NOE energy (kcal mol^{-1} Å^{-2})	33
rmsd for bond distances (Å)	0.007
rmsd for bond angles (deg)	2.57
rmsd for dihedral angles (deg)	32
van der Waals energy (kcal/mol)	−347
Relevant Structural Parameters[a]	
εdG *Y* displacement (Å)[b]	−1.7 to −1.9
εdG glycosidic torsion angle (deg)	−63 to −68
dC17 glycosidic torsion angle (deg)	−13 to −73
Lesion Site HB Energy (kcal/mol)[c]	
dC5•dG18	5.3
dC7•dG16	6.7
dC17 N4H−T9 O4′	3.9

[a] Ranges computed from converging structures. [b] The minus sign indicates movement toward the opposite strand. Canonical B-form values are as follows: *Y* displacement = 0.0 Å; glycosidic torsion angle = −98°. [c] Measured on the averaged minimized structure.

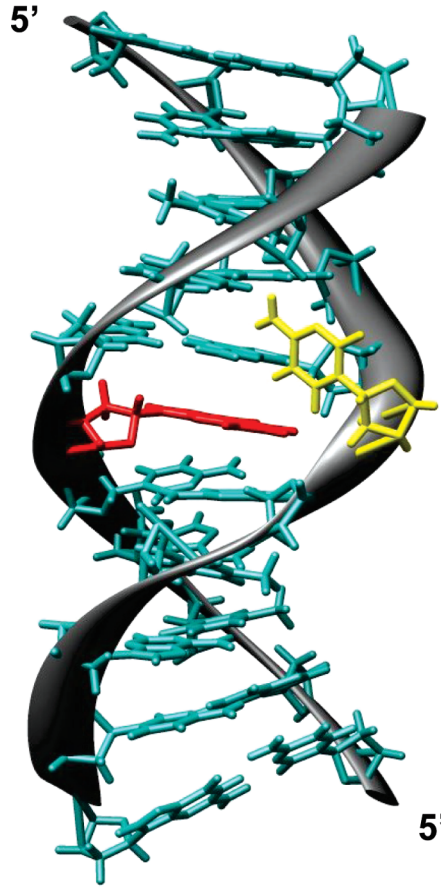


FIGURE 5: Three-dimensional model of the 11-mer εdG duplex, seen with the minor groove prominent. The figure shows the average structure, computed from the converging refined structures, with εdG colored red and the counterbase dC yellow.

eters. The εdG residue is fully intrahelical and appears displaced toward the opposing strand of the duplex, favoring stacking mainly with the dG16 and dG18 residues of flanking base pairs, but without participation of hydrogen bonding (Figures 6 and 7). The lesion partner dC17 is located in the minor groove of the duplex, where it points toward the 3′-end of the damaged strand and forms a strong dC17 N4H−T9 O4′ hydrogen bond (Figure 6). Bases of the εdG•C

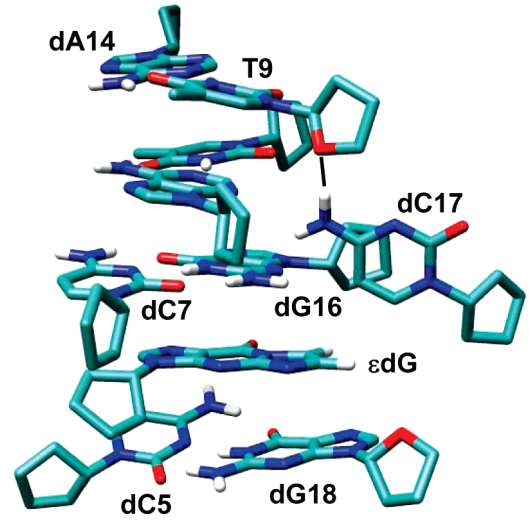


FIGURE 6: Hydrogen bonding interactions at the lesion site of the εdG duplex. Base heavy atoms are colored by type and exchangeable protons white. The black line depicts the dC17 NH4−T9 O4′ hydrogen bond observed in the refined models.

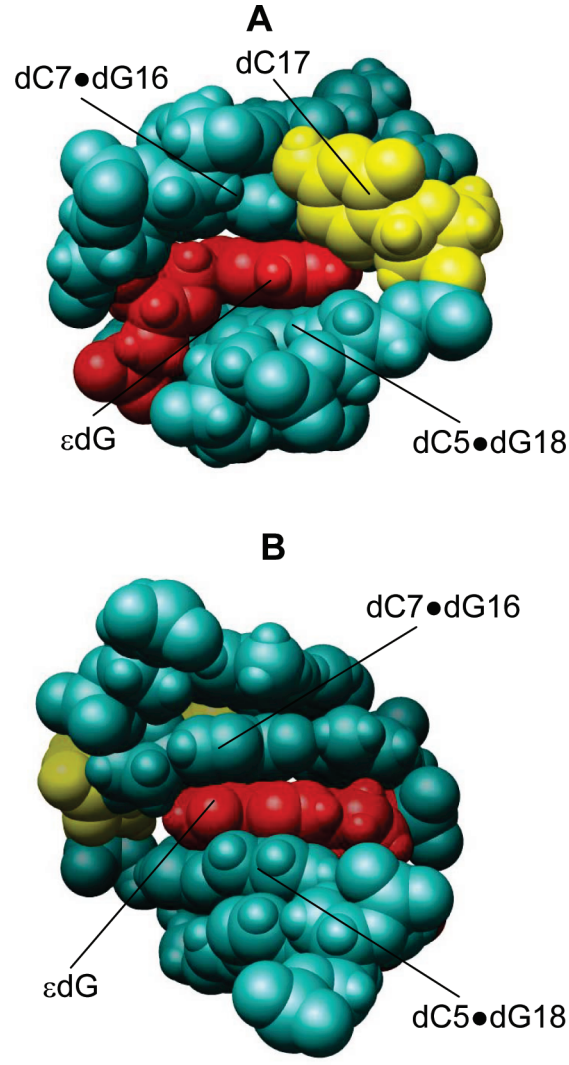


FIGURE 7: Space-filling representations of the central 5 bp segment of the εdG duplex as seen looking down the minor groove (A) or major groove (B) of the helix. εdG is colored red and dC17 yellow.

pair form a 60° angle, with εdG mostly sheltered from the solvent and dC17 largely exposed to water (Figures 6 and

7). The glycosidic torsion (χ) angle of the lesion site residues appears outside the *anti* limit, in the −*syn* range. In the case of εdG, χ adopts a narrow range of values, between −63° to −68°, which is close to the high *anti* range characteristic of canonical B-form DNA. In contrast, the χ angle of dC17 displays a broader range (Table 1). The conformation of lesion-flanking base pairs is quite regular with structural parameters within the B-form family and appropriate WC hydrogen bonding (Figure 6 and Table 1).

DISCUSSION

NMR Spectra of the εdG Duplex. During our analysis of the NMR spectra, it became evident that the εdG lesion induced only small perturbations at the center of the duplex structure. Proton chemical shift differences between the εdG-containing duplex and the unmodified control are significant only at the lesion site and the flanking base pairs (data not shown). The chemical shifts of protons that belong to the four terminal base pairs of the duplex, the intensity of their NOE cross-peaks, and the interproton distances derived from them all fall within the ranges observed for unmodified duplexes (Table 1S, Supporting Information, and Figure 2). Therefore, the subsequent discussion focuses on the structural characteristics of the central 3 bp segment of the duplex.

Several observations indicate the existence of increased dynamics at the lesion site. At pH 6.90, the εdG H10 and εdG H11 signals are relatively broad (Figures 1S and 2S, Supporting Information) and those of εdG H8 and the base (H5 and H6) protons of dC5 and dC7 are broader than analogous signals of non-lesion site residues (Figures 2 and 1S, Supporting Information). Finally, the pH dependence study of the exchangeable protons (Figure 4) produces conclusive evidence that a different conformation of the εdG duplex appears at low pH. Notwithstanding the lack of spectral assignments, the appearance of up to four imino proton signals resonating in the 12.0–13.5 ppm range indicates prevalent WC hydrogen bonding of the low-pH duplex conformation. Furthermore, the pair of exchangeable proton signals at 9.96 and 8.91 ppm strongly suggests protonation of a cytosine residue at acidic pH, most likely the lesion partner dC17, with involvement of its amino group in hydrogen bonding. Several research groups, including our own, have described such a pattern during the formation of d(GCC$^+$) triplets (*39, 40*). Our observation of a low-pH conformation is entirely consistent with the NMR characterization of a εdG duplex at acidic pH communicated by Shanmugam and co-workers (*25*). In their poster, they showed that the duplex adopts a main conformation with εdG in *syn*, forming a Hoogsteen base pair with its protonated dC$^+$ partner, and the lesion-flanking A•T and C•G base pairs adopting WC alignments. The lesion sequence context used for our studies, d(...C-εG-C..), was different from that used by Shanmugam and co-workers, d(...T-εG-G...), suggesting that the acidic form structure of εdG-containing DNA is independent of lesion sequence context.

The NOESY spectrum recorded at pH 6.9 displays the characteristic sequential interactions between the base (purine H8−pyrimidine H6) and the sugar H1′ (Figure 2), H2′, H2″, and H3′ protons (data not shown) of the 5′-flanking residue, indicating that the εdG duplex adopts a regular right-handed helical conformation. The distinctly downfield values of dC17 H5, H6, and H1′ proton chemical shifts (Table 1S, Supporting Information) suggests that dC17 does not stack properly between flanking bases. The high positive dC17-(C_6−H_6) RDC value (20 Hz) indicates that the position of the pyrimidinic ring with respect to the helical axis is very different from that of other bases, arguing also for its nonstacked position (Table 2S, Supporting Information). Furthermore, the presence of a dC17 H5−A8 H2 NOE peak (Figure 2, peak F) clearly establishes the extrahelical position of the lesion partner dC17, more precisely, in the minor groove of the duplex and pointing toward the 3′-end of the damage-containing strand.

Solution Structure of the εdG Duplex. The refined three-dimensional model of the εdG duplex shows excellent agreement with the main NMR observations without significant violations of the idealized covalent geometry (Table 1). The duplex structure is a regular right-handed helix, with the non-lesion-containing residues in the *anti* range, all deoxy sugars in the C2′-endo/C1′-exo conformation and regular WC alignments on all unmodified base pairs (Figure 5). The damaged base, εdG, is located deep inside the helix, stacking between flanking dG16 and dG18 residues and extensively sheltered from solvent exposure (Figures 7 and 4S, Supporting Information). The chemical shift values of the εdG H10 and εdG H11 protons, which are significantly upfield from the base proton region of the spectrum, support the intrahelical orientation of εdG in the refined structure. Furthermore, we have previously reported upfield chemical shift values for the etheno protons of εdC in the εdC•dA duplex (*21*), where the lesion-flanking base pairs were identical to those used in this study and the damaged residue adopted an intrahelical orientation analogous to the one observed for εdG in this study.

The extrahelical location of dC17 readily conforms to experimental restraints. While in the refined MD models dC17 enjoys an increased degree of conformational freedom, with a pairwise rmsd for the C17 residue in the 1.95–0.37 Å range among the converging models (Figure 3S, Supporting Information), the dC17 H5−A8 H2 distance is within the experimentally determined limits in all the structures. In addition, notwithstanding the many factors that affect proton chemical shifts, the downfield values observed for dC17 H5, H6, and H1′ are in qualitative agreement with the solvent-exposed location of this residue. Downfield chemical shift values of dC protons have been observed previously in the case of a *cis-anti*-B[a]P-dG•dC duplex, where the lesion partner residue is positioned in the major groove of the helix and exposed to solvent (*41*), and very recently in the NMR characterization of a εdG•dC duplex at basic pH (*44*). The refined duplex structure shows a strong hydrogen bond between dC17 N4H and T9 O4′, which stabilizes the extrahelical minor groove conformation of dC17 (Figure 6 and Table 1).

Effects of Sequence on the εdG•dC Duplex Structure. There are similarities and differences between our neutral-pH duplex structure, where the damaged residue is in the d(...C-εG-C...) sequence, and the one reported by Shanmugam and co-workers at pH 8.6 (*44*) with the lesion present within the d(...T-εG-G...) context. In both cases, the perturbations from a regular right-handed helix were mostly limited to the lesion-containing base pair, which has εdG properly stacking between flanking bases while its counterbase is

extrahelical. Acidification of the buffer solution causes the appearance of a different lesion site conformation, which has the lesion partner protonated and hydrogen bonded to, most likely, εdG N7 on the Hoogsten edge of the lesion. However, the pK_a of this conformational transition depends strongly on the lesion sequence context, being at approximately neutral pH in Shanmugam's duplex and well below pH 5 in our study. Since G·C base pairs are more stable than A·T base pairs, it is conceivable that the latter breaks more easily, allowing the *anti*−*syn* transition of εdG and the protonation of its counterresidue to occur at higher pH values.

In addition to the pK_a difference, there is a discrepancy in the position of the lesion counterbase, located in the minor groove in our study and in the major groove in Shanmugam's report, which cannot be easily explained by sequence context. We observe the formation of a hydrogen bond that stabilizes the lesion site dC residue extruded in the minor groove of the duplex. Since the heavy atom hydrogen bond acceptor is a sugar O4′ atom, this interaction should, in principle, be independent of duplex sequence. When the dC residue is located in the major groove of the duplex, such interaction is absent and most likely replaced by hydrogen bonding with solvent. Therefore, the mechanisms responsible for this structural difference remain unclear. We are planning to run extensive free MD simulations in the presence of explicit solvent molecules to try to explain this intriguing discrepancy.

Biological Implications. MUG, a bacterial DNA glycosylase that removes thymine or uracil residues from dG·T mismatches, also incises duplex DNA-containing εdC or 1,N^2-εdG lesions. The major groove displacement of εdC or thymine residues, observed in the structures of lesion-containing and dG·T mismatch duplexes, led to the proposal that a positive change of the *X* displacement parameter would be a structural determinant during lesion recognition by MUG (*24*). In the refined models, εdG *X* displacement is normal (−0.72 Å) or slightly changed toward positive values, indicating movement toward the major groove and suggesting that a similar mechanism may take place. However, *X* displacement variations are too small (maximum change of 0.37 Å) and variable to permit an authoritative conclusion.

Y-Family DNA polymerases bypass εdG lesions. In bacteria, Dpo4 incorporates dA or dG opposite εdG, dependent upon the lesion sequence context (*42*). Human DNA polymerase η directs dG incorporation opposite εdG, while DNA polymerases η and κ incorporate dC and T to a similar extent (*43*). The observation that εdG can adopt an *anti* or a *syn* conformation depending upon pH condition and duplex sequence (this work and refs *25* and *44*) would suggest that Y-family DNA polymerases are sensitive to the lesion glycosylic bond, incorporating different bases opposed to each conformation. It would be interesting to determine whether a *syn*−*anti* conformational exchange also occurs on template−primer models containing εdG residues.

ACKNOWLEDGMENT

We thank Ms. Cecilia Torres for the synthesis and purification of oligodeoxynucleotides. Molecular graphics images were produced using the UCSF Chimera from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH Grant P41 RR-01081).

SUPPORTING INFORMATION AVAILABLE

Two tables, listing the proton chemical shifts in the εdG duplex (Table 1S) and RDC values (Table 2S), four figures showing the one-dimensional proton spectrum recorded in 100% D_2O phosphate buffer (pH 6.9) at 27 °C (Figure 1S), the symmetrical H1′ proton region of NOESY (200 ms mixing time) and TOCSY spectra, recorded in D_2O buffer (pH 6.9) at 27 °C (Figure 2S), an overlap view of the converging εdG duplex structures (Figure 3S), and stacking interactions at the lesion site of the εdG duplex (Figure 4S). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Creech, J. L., Jr., and Johnson, M. N. (1974) Angiosarcoma of liver in the manufacture of polyvinyl chloride. *J. Occup. Med. 16*, 150–151.
2. Giri, A. K. (1995) Genetic toxicology of vinyl chloride: A review. *Mutat. Res. 339*, 1–14.
3. Guengerich, F. P., Mason, P. S., Stott, W. T., Fox, T. R., and Watanabe, P. G. (1981) Roles of 2-haloethylene oxides and 2-haloacetaldehydes derived from vinyl bromide and vinyl chloride in irreversible binding to protein and DNA. *Cancer Res. 41*, 4391–4398.
4. Eberle, G., Barbin, A., Laib, R. J., Ciroussel, F., Thomale, J., Bartsch, H., and Rajewsky, M. F. (1989) 1,N^6-Etheno-2′-deoxyadenosine and 3,N^4-etheno-2′-deoxycytidine detected by monoclonal antibodies in lung and liver DNA of rats exposed to vinyl chloride. *Carcinogenesis 10*, 209–212.
5. Fedtke, N., Boucheron, J. A., Walker, V. E., and Swenberg, J. A. (1990) Vinyl chloride-induced DNA adducts. II: Formation and persistence of 7-(2′-oxoethyl)guanine and N^2,3-ethenoguanine in rat tissue DNA. *Carcinogenesis 11*, 1287–1292.
6. Kusmierek, J. T., and Singer, B. (1992) 1,N2-Ethenodeoxyguanosine: Properties and formation in chloroacetaldehyde-treated polynucleotides and DNA. *Chem. Res. Toxicol. 5*, 634–638.
7. Bartsch, H., Barbin, A., Marion, M. J., Nair, J., and Guichard, Y. (1994) Formation, detection, and role in carcinogenesis of ethenobases in DNA. *Drug Metab. Rev. 26*, 349–371.
8. Nair, J., Barbin, A., Guichard, Y., and Bartsch, H. (1995) 1,N^6-Ethenodeoxyadenosine and 3,N^4-ethenodeoxycytine in liver DNA from humans and untreated rodents detected by immunoaffinity/^{32}P-postlabeling. *Carcinogenesis 16*, 613–617.
9. Barbin, A. (2000) Etheno-adduct-forming chemicals: From mutagenicity testing to tumor mutation spectra. *Mutat. Res. 462*, 55–69.
10. Morinello, E. J., Ham, A. J., Ranasinghe, A., Sangaiah, R., and Swenberg, J. A. (2001) Simultaneous quantitation of N^2,3-ethenoguanine and 1,N^2-ethenoguanine with an immunoaffinity/gas chromatography/high-resolution mass spectrometry assay. *Chem. Res. Toxicol. 14*, 327–334.
11. Langouët, S., Mican, A. N., Müller, M., Fink, S. P., Marnett, L. J., Muhle, S. A., and Guengerich, S. P. (1998) Misincorporation of nucleotides opposite five-membered exocyclic ring guanine derivatives by *Escherichia coli* polymerases in vitro and in vivo: 1,N^2-Ethenoguanine, 5,6,7,9-tetrahydro-9-oxoimidazo-[1,2-α] purine, and 5,6,7,9-tetrahydro-7-hydroxy-9-oxoimidazo-[1,2-α]purine. *Biochemistry 37*, 5184–5193.
12. Akasaka, S., and Guengerich, F. P. (1999) Mutagenicity of site-specifically located 1,N^2-ethenoguanine in Chinese hamster ovary cell chromosomal DNA. *Chem. Res. Toxicol. 12*, 501–507.
13. Saparbaev, M., and Laval, J. (1998) 3,N^4-Ethenocytosine, a highly mutagenic adduct, is a primary substrate for *Escherichia coli* double-stranded uracil-DNA glycosylase and human mismatch-specific thymine-DNA glycosylase. *Proc. Natl. Acad. Sci. U.S.A. 95*, 8508–8513.
14. Matijasevic, Z., Sekiguchi, M., and Ludlum, D. B. (1992) Release of N^2,3-ethenoguanine from chloroacetaldehyde-treated DNA by *Escherichia coli* 3 methyladenine DNA glycosylase II. *Proc. Natl. Acad. Sci. U.S.A. 89*, 9331–9334.
15. Saparbaev, S., Kleibl, K., and Laval, J. (1995) *Escherichia coli*, *Saccharomyces cerevisiae*, rat and human 3-methyladenine DNA glycosylases repair 1,N^6-ethenoadenine when present in DNA. *Nucleic Acids Res. 23*, 3750–3755.

16. Hang, B., Singer, B., Margison, G. P., and Elder, R. H. (1997) Targeted deletion of alkylpurine-DNA-N-glycosylase in mice eliminates repair of 1,N^6-ethenoadenine and hypoxanthine but not of 3,N^4-ethenocytosine or 8-oxoguanine. *Proc. Natl. Acad. Sci. U.S.A. 94*, 12869–12874.
17. Saparbaev, M., Langouet, S., Privezentzev, C. V., Guengerich, F. P., Cai, H., Elder, R. H., and Laval, J. (2002) 1,*N*(2)-Ethenoguanine, a mutagenic DNA adduct, is a primary substrate of *Escherichia coli* mismatch-specific uracil-DNA glycosylase and human alkylpurine-DNA-N-glycosylase. *J. Biol. Chem. 277*, 26987–26993.
18. Hang, B., Medina, M., Fraenkel-Conrat, H., and Singer, B. (1998) A 55-kDa protein isolated from human cells shows DNA glycosylase activity toward 3,N4-ethenocytosine and the G/T Mismatch. *Proc. Natl. Acad. Sci. U.S.A. 95*, 13561–13566.
19. de los Santos, C., Kouchakdjian, M., Yerema, K., Basu, A., Essigman, J., and Patel, D. J. (1991) NMR studies of the exocyclic 1,N^6-ethenodeoxyadenosine adduct (εdA) opposite deoxyguanosine in a DNA duplex. εdA(syn) ·dG(anti) pairing at the lesion site. *Biochemistry 30*, 1828–1835.
20. Kouchakdjian, M., Eisenberg, M., Yarema, K., Basu, A., Essigmann, J., and Patel, D. J. (1991) NMR studies of the exocyclic 1,N^6-ethenodeoxyadenosine adduct (εdA) opposite thymidine in a DNA duplex. Nonplanar alignment of εdA(anti) and dT(anti) at the lesion site. *Biochemistry 30*, 1820–1828.
21. Korobka, A., Cullinan, D., Cosman, M., Grollman, A. P., Patel, D. J., Eisenberg, M., and de los Santos, C. (1996) Solution structure of an oligonucleotide duplex containing the exocyclic lesion 3,N^4-etheno-2′-deoxycytidine opposite 2′-deoxyadenosine determined by NMR spectroscopy and restrained molecular dynamics. *Biochemistry 35*, 13310–13318.
22. Cullinan, D., Korobka, A., Grollman, A. P., Patel, D. J., Eisenberg, M., and de los Santos, C. (1996) NMR solution structure of an oligodeoxynucleotide duplex containing the exocyxlic lesion 3,N^4-etheno-2′-deoxycytidine opposite thymidine: Comparison with the duplex containing deoxyadenosine opposite the adduct. *Biochemistry 35*, 13319–13327.
23. Cullinan, D., Grollman, A. P., Eisenberg, M., and de los Santos, C. (1997) Solution structure of a DNA duplex containing the exocyclic lesion: 3,N^4-Etheno-2′-deoxycytidine opposite 2′-deoxyguanosine. *Biochemistry 36*, 11933–11943.
24. Cullinan, D., Johnson, F., and de los Santos, C. (2000) Solution structure of an 11-mer duplex containing the 3,N^4-ethenocytosine adduct opposite 2′-deoxycytidine: Implications for the recognition of exocyclic lesions by DNA glycosylases. *J. Mol. Biol. 296*, 851–861.
25. Shammugam, G., Kozekov, I. D., Guengerich, F. P., Rizzo, C. J., and Stone, M. P. (2007) Structural characterization of 1,N^2-ethenodeoxyguanosine in the dodecamer 5′-CGCATetheno-GGAATCC-3′. Poster TOXI110, 234th ACS National Meeting, August, 2007, Boston.
26. Huang, Y., Torres, M. C., Iden, C. R., and Johnson, F. (2003) Regioselective synthesis of 1,N^2-etheno-2′-deoxyguanosine and its generation in oligomeric DNA. *Chem. Res. Toxicol. 16*, 708–714.
27. Zaliznyak, T., Bonala, R., Johnson, F., and de los Santos, C. (2006) Structure and Stability of DNA Containing the 3-(Deoxyguanosin-N^2-yl)-2-acetylaminofluorene (dG(N^2)-AAF) Lesion: A Bulky Adduct that Persists in Cellular DNA. *Chem. Res. Toxicol. 19*, 745–752.
28. States, D. J., Habekorn, R. A., and Ruben, D. J. (1982) A Two-Dimensional Nuclear Overhauser Experiment with Pure Absorption Phase in Four Quadrants. *J. Magn. Reson. 48*, 286–292.
29. Plateau, P., and Gueron, M. (1982) Exchangeable Proton NMR without Base-Line Distortion, Using New Strong-Pulse Sequences. *J. Am. Chem. Soc. 104*, 7310–7311.
30. Shawn, M., Douglas, S. M., Chou, J. J., and Shih, W. M. (2007) DNA-nanotube-induced alignment of membrane proteins for NMR structure determination. *Proc. Natl. Acad. Sci. U.S.A. 104*, 6644–6648.
31. Brïger, A. (1993) XPLOR Version 3.1: A system for X-Ray Crystallography and NMR, Yale University Press, New Haven, CT.
32. Brooks, B. R., Buccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem. 4*, 187–217.
33. Friedman, R. A., and Honig, B. (1992) The electrostatic contribution to DNA base-stacking interactions. *Biopolymers 32*, 145–159.
34. Yip, P., and Case, D. A. (1989) A New Method for Refinement of Macromolecular Structures Based on Nuclear Overhauser Effect Spectra. *J. Magn. Reson. 83*, 643–648.
35. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera: A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem. 25*, 1605–1612.
36. Lavery, R., and Sklenar, H. (1988) The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn. 6*, 655–667.
37. de los Santos, C. (1999) Probing DNA Structure by NMR Spectroscopy, in *Comprehensive Natural Products Chemistry* (Kool, E., Ed.) Vol. 7, pp 55–80, Elsevier Science Ltd., Oxford, U.K.
38. van de Ven, F. J. M., and Hilbers, C. W. (1988) Nucleic acids and nuclear magnetic resonance. *Eur. J. Biochem. 178*, 1–38.
39. de los Santos, C., Rosen, M., and Patel, D. (1989) NMR studies of DNA $(R^+)_nC(Y^-)_nC(Y^+)_n$ triple helices in solution: Imino and amino proton markers of T·A·T and C·G·C^+ base triple formation. *Biochemistry 28*, 7282–7289.
40. Sklenář, V., and Feigon, J. (1990) Formation of a stable triplex from a single DNA strand. *Nature 345*, 836–838.
41. Cosman, M., Hingerty, B. E., Luneva, N., Amin, S., Geacintov, N. E., Broyde, S., and Patel, D. J. (1996) Solution conformation of the (−)-cis-anti-benzo[a]pyrenyl-dG adduct opposite dC in a DNA duplex: Intercalation of the covalently attached BP ring into the helix with base displacement of the modified deoxyguanosine into the major groove. *Biochemistry 35*, 9850–9863.
42. Zang, H., Goodenough, A. K., Choi, J. Y., Irimia, A., Loukachevitch, L. V., Kozekov, I. D., Angel, K. C., Rizzo, C. J., Egli, M., and Guengerich, F. P. (2005) DNA adduct bypass polymerization by *Sulfolobus solfataricus* DNA polymerase Dpo4. Analysis and crystal structures of multiple base pair substitution and frameshift products with the adduct 1,N_2-ethenoguanine. *J. Biol. Chem. 280*, 29750–29764.
43. Choi, J. Y., Zang, H., Angel, K. C., Kozekov, I. D., Goodenough, A. K., Rizzo, C. J., and Guengerich, F. P. (2006) Translesion synthesis across 1,N^2-ethenoguanine by human DNA polymerases. *Chem. Res. Toxicol. 19*, 879–886.
44. Shanmugam, G., Goodenough, A. K., Kozekov, I. D., Guengerich, F. P., Rizzo, C. J., and Stone, M. P. (2007) Structure of the 1,N^2-etheno-2′-deoxyguanosine adduct in duplex DNA at pH 8.6. *Chem. Res. Toxicol. 20*, 1601–1611.

BI7022514