The Exocyclic 1,N^2-Deoxyguanosine Pyrimidopurinone M_1G Is a Chemically Stable DNA Adduct When Placed Opposite a Two-Base Deletion in the $(CpG)_3$ Frameshift Hotspot of the *Salmonella typhimurium hisD3052* Gene†

Nathalie C. Schnetz-Boutaud,‡ Sam Saleh,§ Lawrence J. Marnett,|| and Michael P. Stone*,‡

Departments of Chemistry, Pharmacology, and Biochemistry, Center in Molecular Toxicology, A.B. Hancock, Jr., Memorial Laboratory for Cancer Research, Vanderbilt−Ingram Cancer Center, Vanderbilt University, Nashville, Tennessee 37235

Received June 15, 2001; Revised Manuscript Received September 18, 2001

ABSTRACT: The pyrimidopurinone adduct M_1G [3-(2′-deoxy-β-D-erythro-pentofuranosyl)pyrimido[1,2-*a*]-purin-10(3*H*)-one], formed in DNA upon exposure to malondialdehyde or base propenals, was incorporated into 5′-d(ATCGCMCGGCATG)-3′·5′-d(CATGCCGCGAT)-3′, where M = M_1G. This duplex contained a two-nucleotide bulge in the modified strand, and was named the M_1G-2BD oligodeoxynucleotide. It provided a model for −2 bp strand slippage deletions associated with the $(CpG)_3$-iterated repeat hotspot for frameshift mutations from the *Salmonella typhimurium hisD3052* gene. M_1G was chemically stable in the M_1G-2BD duplex at neutral pH. The two-base bulge in the M_1G-2BD oligodeoxynucleotide was localized and consisted of M_1G and the 3′-neighbor deoxycytosine. The intrahelical orientation of M_1G was established from a combination of NOE and chemical shift data. M_1G was in the anti conformation about the glycosyl bond. The 3′-neighbor deoxycytosine appeared to be extruded toward the major groove. In contrast, when M_1G was placed into the corresponding fully complementary $(CpG)_3$-iterated repeat duplex at neutral pH, spontaneous and quantitative ring-opening to N^2-(3-oxo-1-propenyl)-dG (the OPG adduct) was facilitated [Mao, H., Reddy, G. R., Marnett, L. J., and Stone, M. P. (1999) *Biochemistry 38*, 13491−13501]. The structure of the M_1G-2BD duplex suggested that the bulged sequence lacked a cytosine amino group properly positioned to facilitate opening of M_1G and supports the notion that proper positioning of deoxycytosine complementary to M_1G is necessary to promote ring-opening of the exocyclic adduct in duplex DNA. The structure of the M_1G-2BD duplex was similar to that of the structural analogue 1,N^2-propanodeoxyguanosine (PdG) in the corresponding PdG-2BD duplex [Weisenseel, J. P., Moe, J. G., Reddy, G. R., Marnett, L. J., and Stone, M. P. (1995) *Biochemistry 34*, 50−64]. The fixed position of the bulged bases in both instances suggests that these exocyclic adducts do not facilitate transient bulge migration.

The exocyclic guanine adduct M_1G[1] [3-(2′-deoxy-β-D-erythro-pentofuranosyl)pyrimido[1,2-*a*]-purin-10(3*H*)-one] arises in DNA from multiple sources. Malondialdehyde is a toxic and mutagenic metabolite produced by lipid peroxidation and prostaglandin biosynthesis [for a review, see (*1*)]. MDA exists in solution primarily as β-hydroxyacrolein and reacts with DNA as a bis-electrophile to form M_1G (*2*−*6*). Alternatively, M_1G arises as a consequence of DNA oxidative damage, resulting in the formation of base propenals that can transfer their oxopropenyl group to deoxyguanosine (*7*) (Scheme 1). The base propenals are significantly more potent than MDA in forming M_1G (*8*).

M_1G has been identified in DNA from rodent (*9*) and human (*10, 11*) tissue samples, as have other exocyclic purine lesions (*12*−*14*), suggesting their formation in vivo. M_1G has been quantitated by mass spectroscopic (*15, 16*), postlabeling (*17, 18*), and immunochemical (*19*) techniques. It represents the most abundant exocyclic DNA adduct present endogenously in human DNA (*16*−*18, 20*). Site-specific mutagenesis experiments indicate that it is an efficient premutagenic lesion (*21*). Thus, M_1G is a likely mediator of human genetic disease.

The M_1G adduct is chemically stable in nucleotides and single-stranded DNA at neutral pH. Under basic conditions,

† This work was supported by NIH Grants CA-55678 (M.P.S.) and CA-47479 (L.J.M.). Funding for the NMR spectrometer was supplied by NIH Grant RR-05805, and the Vanderbilt Center in Molecular Toxicology (ES-00267). The National Magnetic Resonance Facility at Madison was funded by the University of Wisconsin, NSF Grants DMB-8415048 and BIR-9214394, NIH Grants RR-02301, RR-02781, and RR08438, and the USDA. Portions of this work were presented at the Sixth International Symposium on Biological Reactive Intermediates held at the Universite Rene Descartes, Paris, France, July 16−20, 2000.

* To whom correspondence should be addressed. Telephone: (615) 322-2589; FAX: (615) 343-1234; E-mail: stone@toxicology.mc.vanderbilt.edu.

‡ Department of Chemistry.

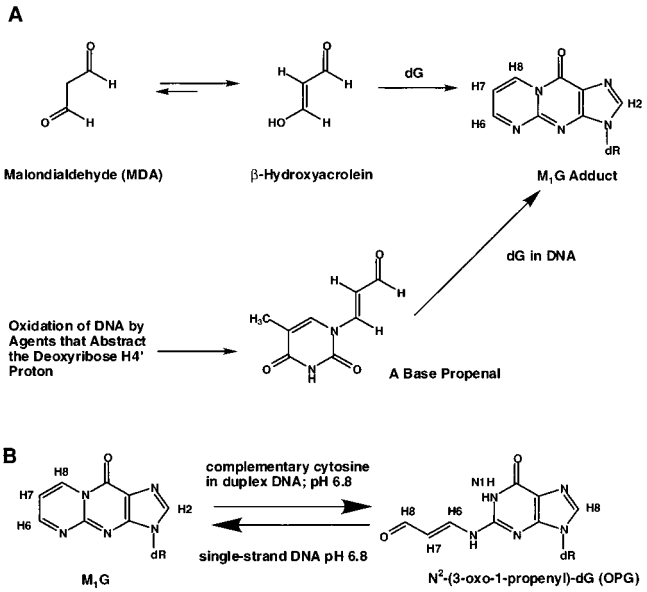
§ Department of Pharmacology.

|| Department of Biochemistry.

[1] Abbreviations: DQF-COSY, double-quantum-filtered correlation spectroscopy; M_1G, 3-(2′-deoxy-β-D-erythro-pentofuranosyl)pyrimido-[1,2-*a*]-purin-10(3*H*)-one; M_1G-2BD, 5′-d(ATCGCMCGGCATG)-3′·5′-d(CATGCCGCGAT)-3′ where M = M_1G; NOESY, two-dimensional NOE spectroscopy; OPG, N^2-(3-oxo-1-propenyl)-dG; PdG, 1,N^2-propanodeoxyguanosine; PEM, potential energy minimization; PdG-2BD, 5′-d(ATCGCPCGGCATG)-3′·5′-d(CATGCCGCGAT)-3′ where P = PdG; rMD, restrained molecular dynamics; rmsd, root-mean-square-deviation; R_1^x, sixth root factor residual; TPPI, time-proportional phase increment.

10.1021/bi011242u CCC: $20.00
Published on Web 12/01/2001

Scheme 1[a]



[a](A) Formation of M_1G from malondialdehyde or from base propenals; thymine propenal is shown as a representative base propenal. Note the numbering scheme for M_1G in which the imidazole proton is H2, corresponding to the H8 proton in purines. The exocyclic ring protons are numbered H6, H7, and H8. (B) M_1G is stable in single-stranded DNA. In duplex DNA when placed opposite deoxycytosine, it is spontaneously and quantitatively converted to N^2-(3-oxo-1-propenyl)-dG, the OPG derivative.

Scheme 2: M_1G-2BD Oligodeoxynucleotide Contains a Two-Base Deletion in the Complementary Strand[a]

```
5'-A-2T-1C1 G2 C3 M4 C5 G6 G7 C8 A9 T10G11-3'
3'-T22A21G20C19G18       C17C16G15T14A13C12-5'
```

[a]The nucleotide numbering scheme is derived to be consistent with previous studies on this iterated repeat sequence from the *hisD3052* gene (*45*).

it spontaneously converts to its N^2-(3-oxo-1-propenyl)-dG (OPG) derivative. However, when M_1G is placed at neutral pH into duplex DNA opposite deoxycytosine, a rapid, spontaneous, and quantitative conversion to the OPG derivative is facilitated. This is reversible, such that upon denaturation of the DNA duplex, M_1G is regenerated (Scheme 2). Ring-opening does not occur at neutral pH in duplex DNA if thymine is placed opposite to M_1G. These observations led to the proposal that cytosine in duplex DNA catalyzes the transformation of M_1G to its ring-opened OPG derivative, probably by a mechanism involving the exocyclic amino group of the complementary cytosine (*22*). Subsequently, the major deoxyguanosine adduct derived from acrolein, γ-hydroxyl-1,N^2-propano-2′-deoxyguanosine, was also reported to exist in duplex DNA opposite deoxycytosine primarily as its ring-opened derivative (*23*).

A refined structure was obtained for the OPG derivative of M_1G in d(ATCGCXCGGCATG)·d(CATGCCGCGCGAT), X = OPG (*24*). The structure of the OPG lesion embedded into this sequence was of interest because MDA, a small alkylating agent, induced frameshift mutations (*25*). This sequence contains the *Salmonella typhimurium hisD3052* gene $d(CpG)_3$ frameshift mutation hotspot. In this structure, OPG maintained stacking interactions with neighboring bases. It was not Watson−Crick hydrogen-bonded. The cytosine complementary to OPG was pushed toward the major groove but maintained partial stacking with its neighboring bases. The modified guanine remained in the anti conformation, while the OPG 3-oxo-1-propenyl moiety was positioned in the minor groove of the duplex.

The frameshift mutations in the $d(CpG)_3$ hotspot are −2 bp deletions. Replication bypass studies in vitro suggest that DNA polymerase β can induce two-base deletions when replicating past M_1G in this sequence context.[2] Thus, it was of interest to examine an oligodeoxynucleotide containing a two-base (CpG) deletion opposite M_1G. We named this oligodeoxynucleotide M_1G-2BD (Scheme 1). This work reports the solution structure of M_1G-2BD using a simulated annealing protocol in which molecular dynamics calculations were restrained by a combination of ^{1}H NOE data and vicinal coupling constants. One significant finding is that M_1G is chemically stable in the M_1G-2BD oligodeoxynucleotide. It does not undergo spontaneous opening of the deoxyguanosine 1,N^2-exocyclic ring to the OPG adduct under these conditions. These structural studies provide the first glimpse of the intact M_1G adduct within a DNA duplex. The work corroborates the previous observation that spontaneous ring-opening of M_1G to the OPG lesion in duplex DNA requires the presence of deoxycytosine complementary to the M_1G lesion (*22*). It supports the notion that the exocyclic amino group of deoxycytosine mediates formation of the OPG lesion in duplex DNA, and in the absence of an appropriately positioned amino group, M_1G is chemically stable. A second significant finding is that the resulting two-base bulge is localized and involves M_1G and its 3′-neighbor cytosine. M_1G is in the anti conformation about the glycosyl bond, and inserted into the helix. The 3′-neighbor cytosine is pushed toward the major groove. The structure is similar to that previously reported for the structural analogue 1,N^2-propano-dG (PdG) in the corresponding PdG-2BD sequence context (*26*). The results suggest that M_1G and PdG do not promote transient bulge migration (*27*−*29*) in the *hisD3052* $(CG)_3$ iterated repeat sequence.

MATERIALS AND METHODS

Oligodeoxynucleotide Synthesis. The unmodified oligodeoxynucleotides were synthesized by the Midland Certified Reagent Co. (Midland, TX) and purified by anion-exchange chromatography. M_1G was synthesized, purified, and incorporated into the oligodeoxynucleotide as described (*30*, *31*). The extinction coefficients of the oligodeoxynucleotides were calculated on the basis of nearest-neighbor analysis: 1.04×10^5 M^{-1} cm^{-1} for the modified strand 5′-d(ATCGCMCGGCATG)-3′ and 1.01×10^5 M^{-1} cm^{-1} for the complementary strand 5′-d(CATGCCGCGAT)-3′ (*32*). An excess of nonmodified strand was combined with the modified strand, in a buffer consisting of 10 mM NaH_2PO_4, 0.1 M NaCl, 50 μM Na_2EDTA at pH 7.0. The solution was heated at 95 °C for 5 min and then allowed to cool slowly to room temperature. The resulting mixture of single-stranded and duplex DNA was equilibrated in 10 mM NaH_2PO_4, 0.1 M NaCl, 50 μM Na_2EDTA at pH 7.0 on a column containing

[2] Riggins, J. N., and Marnett, L. J. (2001) Malondialdehyde−Deoxyguanosine Adducts M_1G and N^2OPG Block Replication by Human DNA Polymerase β and Induce Frameshift Mutations in Vitro. 222nd American Chemical Society National Meeting, Chicago, IL, Aug 2001, Division of Chemical Toxicology, Collected Abstracts.

DNA Grade Biogel hydroxylapatite (Bio-Rad Laboratories, Richmond, CA). The DNA was eluted off the hydroxylapatite with a gradient from 10 to 200 mM NaH_2PO_4, pH 7.0, to separate single-stranded from double-stranded duplex DNA. The duplex was lyophilized, resuspended in 1 mL of H_2O, and desalted on a Sephadex G-25 column. The sample was lyophilized.

UV Spectroscopy. UV melting assays were performed with 8−9 μM DNA strand concentration in 1 mL of 10 mM NaH_2PO_4, 1.0 M NaCl, 50 μM Na_2EDTA at pH 7.0 in a 1.0 cm quartz cuvette to obtain a total initial absorbance of ~1.0 at 260 nm. Samples were heated in the UV spectrometer at a rate of 1 °C/min while observing the absorbance at 260 nm. The melting temperature, T_m, was determined as the midpoint of the double-strand/single-strand transition.

NMR Spectroscopy. The modified duplex was prepared at a concentration of 2 mM. For observation of nonexchangeable protons, the sample was dissolved in 0.5 mL of 10 mM NaH_2PO_4, 0.1 M NaCl, 50 μM Na_2EDTA at pH 7.0. The sample was exchanged 3 times with 99.9% D_2O and dissolved in 99.96% D_2O. For observation of exchangeable protons, the sample was dissolved in 0.5 mL of 10 mM NaH_2PO_4, 0.1 M NaCl, 50 μM Na_2EDTA at pH 7.0. The sample was lyophilized and suspended in 0.5 mL of 9:1 H_2O/D_2O. ^{1}H NMR spectra were recorded at 600.13 and 500.13 MHz. The temperature was controlled at 20 ± 0.5 °C for the observation of nonexchangeable protons and at 5 ± 0.5 °C for the exchangeable protons. Chemical shifts were referenced to the water resonance at 4.81 ppm at 20 °C and 4.97 ppm at 5 °C.

NOESY and DQF-COSY experiments were performed at a frequency of 600.13 MHz. Phase-sensitive NOESY spectra used for resonance assignment were recorded using TPPI phase cycling with a mixing time of 200 ms. For examination of exchangeable protons, phase-sensitive NOESY were carried out using a field gradient Watergate pulse sequence for water suppression (*33*). The spectra were recorded at 5 °C with a mixing time of 150 ms. These experiments were generally recorded with 1024 real data points in the d1 dimension and 2048 real data points in the d2 dimension. A relaxation delay of 2.0 s was used for these experiments. To derive distance restraints, NOESY spectra were recorded at mixing times of 150, 200, and 250 ms.

The data were processed using FELIX (Accelrys, Inc., San Diego, CA) or X-WIN NMR (Bruker Instruments, Inc., Billerica, MA) on an Octane workstation (Silicon Graphics, Inc., Mountain View, CA). The data in the d1 dimension were zero-filled to give a matrix of 2048 × 2048 real points. A skewed sine-bell square apodization function with a 90° phase shift and a skew factor of 1.0 was used in both dimension.

Restrained Molecular Dynamics Calculations. The rMD calculations followed a protocol which was essentially the same as that developed by Weisenseel et al. for the structure of the corresponding PdG-2BD oligodeoxynucleotide (*26*). Classical B-DNA and A-DNA were used as the reference structures to create starting structures for the refinement (*34*). The 13−11-mer duplex d(ATCGCMCGGCATG)·d(CATGCCGCGAT) was built starting with the model and procedure used by Weisenseel et al. (*26*). M_1G was constructed at the fourth position using the BUILDER module of INSIGHT II (Accelrys, Inc.). The A and B structures were energy-minimized by the conjugate gradients method for 200 iterations without experimental restraints to give starting IniA and IniB used for the subsequent relaxation matrix analysis and molecular dynamics calculations. Footprints were drawn around the NOE cross-peak for the NOESY spectrum measured with a mixing time of 250 ms to define the size and the shape of the individual cross-peak using FELIX. The same set of footprints was applied to spectra measured with the other mixing times. Cross-peak intensities were determined by volume integration of the areas under the footprints. The intensities were combined as necessary with intensities generated from complete relaxation matrix analysis of a starting DNA structure to generate two hybrid intensity matrixes using MARDIGRAS (*35*) for each of the three mixing times. Calculations using DNA starting models generated by INSIGHT II, NOE experiments with three mixing times, and a DNA correlation time of 2, 3, and 4 ns yielded 18 sets of distances. These data were pooled; average values of all minimum and maximum distances were used in setting error bounds to give the experimental NOE restraints used in subsequent molecular dynamics calculations (*36*). For partially overlapped cross-peaks, lower and upper error bonds on the resulting distances were increased. These distance restraints were divided into classes on the basis of the confidence factor obtained from MARDIGRAS. Additional restraints were included. The deoxyribose data were consistent with the C2′-endo sugar ring conformation (*37*). The backbone torsion angles ϵ and ζ were restrained to 165 ± 20° and −120 ± 20°, respectively (*38*). Empirical Watson−Crick hydrogen bonding restraints were used except for M_1G and C^5. These were similar to those used previously in structural determinations of oligodeoxynucleotides (*39*).

Energy minimization and restrained molecular dynamics calculations were carried out with X-PLOR (*40*). All calculations were based on an energy function approach in which the total energy was the sum of the empirical energy of the molecule and the effective energy, comprised of the restraint energy terms. The CHARMM force field contributed the usual terms for the bonds, bond angles, torsion angles, tetrahedral and planar geometry, hydrogen bonding, and nonbonding interactions, including van der Waals and electrostatic forces. The electrostatic term used the Coulombic function based on a reduced charge set of partial charges and a distance-dependent dielectric constant of 4.0, to mimic solvent screening of charge. The van der Waals term was approximated using the Lennard-Jones potential energy function. The nonbonded pair list was updated if any atom moved more than 0.5 Å, and the cutoff radius for the nonbonded interaction was 11 Å. All calculations were run in vacuo without explicit counterions. The effective energy function was comprised of two terms describing distances and dihedral restraints, both of which were in the form of a standard square-well potential (*41*). Bond lengths involving hydrogens were fixed with the SHAKE algorithm (*42*) during molecular dynamics calculations.

The simulated annealing procedure consisted of a total of 50 ps of rMD. The protocol consisted of 30 ps of high-temperature rMD, 5 ps of cooling, and 15 ps of low-temperature rMD. The high temperature was 2500 K, and the low temperature was 300 K. The temperature was controlled by coupling the molecules to a temperature bath with a coupling constant of 0.05 ps (*43*). Ten structures were

Table 1: Optical Melting Temperatures and Sequences (°C; Strand Concentration ~8 μM) for Oligodeoxynucleotides Derived from the *hisD3052* Gene of *Salmonella typhimurium*

abbreviation	sequence	T_m
wt	5′-ATCGCGCGGCATG-3′ 3′-TAGCGCGCCGTAC-5′	73
N^2-OPG	5′-ATCGCOCGGCATG-3′ 3′-TAGCGCGCCGTAC-5′	59
wt-2BD	5′-ATCGCGCGGCATG-3′ 3′-TAGCG- -CCGTAC-5′	58
PdG-2BD	5′-ATCGCPCGGCATG-3′ 3′-TAGCG- -CCGTAC-5′	45
M_1G-2BD	5′-ATCGCMCGGCATG-3′ 3′-TAGCG- -CCGTAC-5′	43

calculated from each starting structure, each being assigned a different random seed for the generation of the initial atomic velocity vectors. The rMD calculations were initialized by assigning a random set of velocities to all the atoms that fit a Maxwell−Boltzmann distribution at 2500 K. The rMD calculations were carried out for 30 000 steps at 2500 K, then cooled to 300 K over 5000 steps, and continued at 300 K for an additional 15 000 steps. An initial force constant of 50.0 kcal mol^{-1} $Å^{-1}$ was used for class 1 distance restraints. Throughout the rMD calculation, the force constants for classes 2, 3, 4, and 5 were set to 90%, 80%, 70%, and 60%, respectively, of the value for class 1. The initial value of the force constant for the base pairing restraints was set at 50.0 kcal mol^{-1} $Å^{-1}$. The force constants were maintained at the initial value for the first 10 000 steps of the rMD calculations; then class 1 force constants were increased to 200 kcal mol^{-1} $Å^{-1}$, and base pairing force constants were increased to 150 kcal mol^{-1} $Å^{-1}$ over the next 10 000 steps. The force constants were maintained for 17 000 steps, scaled down to 70 and 50 kcal mol^{-1} $Å^{-1}$ for class 1 and base pairing restraints, respectively, over 3000 steps and remained at these values for the final 10 000 steps of the rMD calculations. Structure coordinates were archived every 0.1 ps over the final 10 ps of the rMD simulation. Structure coordinates extracted from the final 4 ps of each rMD calculation were averaged and energy-minimized for 200 iterations using a conjugate gradients algorithm. The final refined structure was obtained by minimizing the averaged structure generated from 10 structures from two rMD calculations, 5 from an A-form starting structure, iniA, and 5 from a B-form starting structure, iniB. Back-calculation of NMR data was performed using CORMA (*44*) on the final refined structure excluding two terminal base pairs.

RESULTS

Thermodynamic Stability of M_1G Opposite a Two-Base Deletion. Table 1 compiles results of thermal denaturation assays for M_1G-2BD and related oligodeoxynucleotides. The melting temperature of the native *hisD3052* oligodeoxynucleotide was measured to be 73 °C. The deletion of a CpG dinucleotide repeat unit from the iterated $(CpG)_3$ frameshift-prone sequence reduced the T_m to 58 °C. The unmodified *hisD3052*-2BD oligodeoxynucleotide undergoes the phenomenon of "bulge migration", in which the position of the bulged bases is in rapid equilibrium within the $(CpG)_3$ sequence context (*28*). The presence of bulge migration was evidenced by the NMR spectrum of the unmodified bulge sequence, which showed evidence of disorder in the iterated repeat. Introduction of the M_1G modification into the oligodeoxynucleotide resulted in a further drop in T_m from 58 to 43 °C. This value of T_m for the M_1G-2BD duplex was nearly identical to the value observed for the PdG-2BD duplex (*45*).

Assignments of Nonexchangeable DNA Protons. The sequential assignment for the M_1G-2BD duplex was accomplished using standard protocols (*46*, *47*) (Figure 1). The sequential NOEs were interrupted in both the modified and unmodified strands of the duplex. In the modified strand, the sequential NOE between C^3 H1′ and M_1G H2 (the imidazole proton of M_1G; corresponding to G H8 in the unmodified nucleotide) was missing. In the complementary strand, the C^{17} H1′ to G^{18} H8 NOE was missing. Spectral line broadening was localized adjacent to the M_1G lesion. In Figure 1, line broadening is observed for the cytosine H5 and cytosine H6 resonances of C^3 and C^5 in the modified strand. These two cytosines are the 5′- and 3′-neighbors of the M_1G lesion, respectively. In the complementary strand, line broadening was observed for C^{19} H5 and H6. The complete assignment of the deoxyribose H2′, H2″, and H3′ protons was achieved. Of the H4′ protons, 22 resonances were assigned. Partial assignments were made for the deoxyribose H5′ and H5″ protons; these were in many instances severely overlapped, precluding unequivocal assignments. The assignments are tabulated in Table S1 of the Supporting Information.

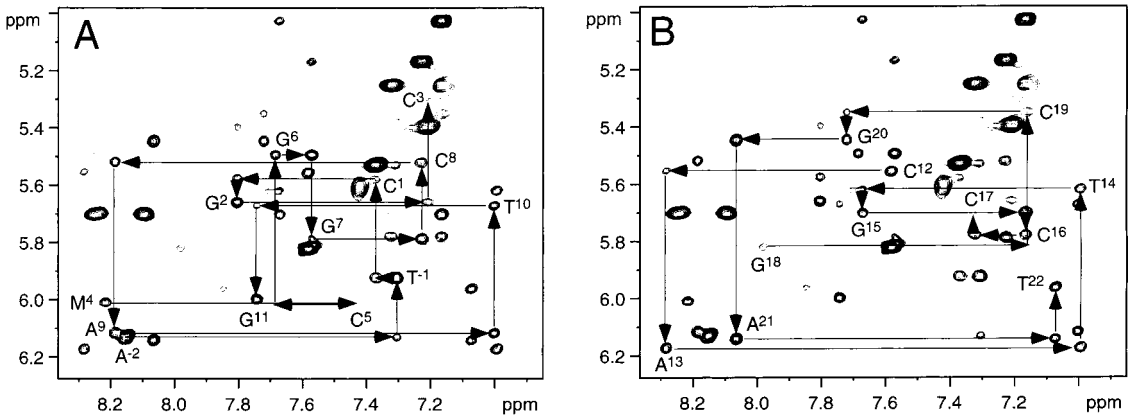


FIGURE 1: Expanded plot showing ^1H NOE sequential connectivities for the M_1G-2BD oligodeoxynucleotide. (A) The modified strand. The sequential connectivities are interrupted at the C^3 H1′ → M^4 H2 step. (B) The complementary strand. The sequential connectivities are interrupted at the C^{17} H1′ → G^{18} H8 step.

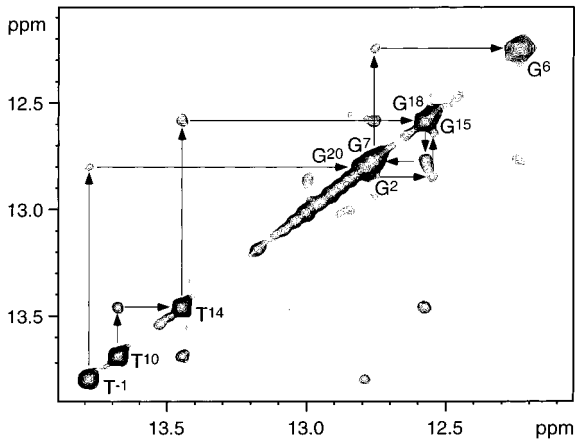


FIGURE 2: Expanded plot showing the sequential NOE connectivities for the imino protons of the M_1G-2BD oligodeoxynucleotide. No connectivity is observed between G^{18} N1H and G^6 N1H.

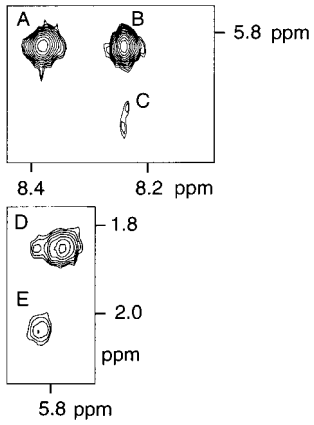


FIGURE 3: Expanded plot showing NOE cross-peaks between M_1G and DNA. (A) The M^4 H8 → M^4 H7 cross-peak. (B) The M^4 H6 → M^4 H7 cross-peak. (C) The M^4 H6 → G^{18} H1′ cross-peak. (D) The M^4 H7 → C^{17} H2′ cross-peak. (E) The M^4 H7 → C^{17} H2″ cross-peak.

Assignments of Exchangeable DNA Protons. The assignments for the base imino protons of M_1G-2BD were made at 5 °C (Figure 2). Three resonances from thymine imino protons, T^{-1} N3H, T^{10} N3H, and T^{14} N3H, were observed between 13.4 and 13.7 ppm, which indicated formation of a stable duplex on both sides of the unpaired bases. The thymine imino proton T^{22} N3H was broadened, probably due to the exchange with solvent. Seven other resonances were also assigned between 12 and 12.8 ppm corresponding to the guanine imino protons. There was considerable spectral overlap, such that G^{15} N1H and G^{18} N1H were overlapped, as were G^2 N1H, G^7 N1H, and G^{20} N1H. G^{11} N1H, at the 3′ terminus of the oligodeoxynucleotide duplex, was observed, but also considerably broadened due to solvent exchange. An interruption in the sequential connectivities between imino protons was observed at G^{18}, the base pair 5′ to the M_1G lesion. No NOE connectivity was observed between G^{18} N1H and G^6 N1H, which served to localize the adduct between base pairs C^3•G^{18} and G^6•C^{17}. The assignments are tabulated in Table S2 of the Supporting Information.

M_1G Protons. The exocyclic protons H6, H7, and H8 of M_1G were identified as the characteristic aromatic spin system (Figure 3). M_1G H8 was observed at 8.4 ppm, M_1G H7 was observed at 5.8 ppm, and M_1G H6 was observed at 8.2 ppm. The *J* coupling constants for the M_1G protons were measured as $J_{H6,H7} \sim 4$ Hz and $J_{H7,H8} \sim 7$ Hz from a DQF-COSY experiment. Increases in shielding were observed for each of the M_1G protons, as compared to the chemical shifts in the single-stranded oligodeoxynucleotide. M_1G H8 shifted upfield by ~0.9 ppm, H7 shifted upfield by ~1.75 ppm, and H6 shifted upfield by ~0.6 ppm. The upfield chemical shifts were attributed to ring-current effects and led to the conclusion that the M_1G moiety must be inserted into the DNA duplex.

NOE Connectivities between M_1G and DNA. The observation of cross-strand NOEs between M_1G protons and the DNA confirmed that M_1G was inserted into the duplex (Figure 3). These were observed from M_1G H7 to C^{17} H2′ (1.85 ppm) and H2″ (2.03 ppm). A weak cross-strand NOE was observed from M_1G H6 to G^{20} H1′ (5.93 ppm). The NOEs between M_1G and DNA also helped to establish the orientation of M_1G at the location of the two-base bulge in M_1G-2BD.

Torsion Angle Restraints Generated from J-Coupling Data. The pseudorotational equilibria of the deoxyribose sugar moieties were examined by a ^{1}H DQF-COSY experiment. The spectrum, included as Figure S2 of the Supporting Information, showed that all of the deoxyribose rings in the M_1G-2BD oligodeoxynucleotide were predominantly in the C2′-endo conformation, characteristic of a B-type helix (*37*). A ^{1}H-^{31}P HMBC experiment using an IBURP-shaped pulse in the ^{1}H dimension was used to examine the ^{31}P chemical shifts and the $^3J_{31P-1H}$ couplings between H3′ and ^{31}P. The spectrum showed no unusual ^{31}P shifts, and the ^3J couplings indicated that the backbone torsion angles were in the normal range for B-type duplexes (*38*). The glycosyl torsion angles, χ, of all nucleotides were determined to be in the anti conformational range, also consistent with a B-form helix.

Chemical Shift Effects. Figure 4 shows ^{1}H chemical shift comparisons between the M_1G-2BD oligodeoxynucleotide, the PdG-2BD oligodeoxynucleotide (*45*), and the unmodified 13−11-mer. Chemical shift perturbations greater than 0.2 ppm were localized at positions M^4 and C^5 in the modified strand. The greatest shift was observed for the adducted site M^4. For the M^4 H2 (imidazole) proton, a downfield shift of 0.45 ppm was observed when the M_1G-2BD oligodeoxynucleotide was compared with the unmodified 13−11-mer oligodeoxynucleotide. A downfield shift of 0.66 ppm was observed when the M_1G-2BD oligodeoxynucleotide was compared with the PdG-2BD oligodeoxynucleotide. A chemical shift effect was observed for C^5 H6, on the 3′-side of the adduct. An upfield shift of 0.31 ppm was observed when the M_1G-2BD oligodeoxynucleotide was compared with the unmodified 13−11-mer, and a downfield shift of 0.15 ppm was observed when the M_1G-2BD oligodeoxynucleotide was compared with the PdG-2BD oligodeoxynucleotide. The deoxyribose protons at the adduct site exhibited small chemical shift changes as well.

Restrained Molecular Dynamics. The distribution of NOE restraints used in the rMD calculations is shown in Table 2, and the NOEs are listed in Table S3 of the Supporting Information. There was an average of approximately 12 distance restraints per nucleotide. The relatively low number of distance restraints at nucleotide M^4 was offset by the observation of three cross-strand NOEs that allowed the position of M^4 to be determined with reasonable confidence. The NOE restraints detailed in Table 2 were augmented with

Table 2: Distribution of NOE Restraints among Nucleotide Units of the M_1G-2BD Oligodeoxynucleotide[a]

restraints	A^{-2}	T^{-1}	C^{1}	G^{2}	C^{3}	M^{4}	C^{5}	G^{6}	G^{7}	C^{8}	A^{9}	T^{10}	G^{11}	C^{12}	A^{13}	T^{14}	G^{15}	C^{16}	C^{17}	G^{18}	C^{19}	G^{20}	A^{21}	T^{22}
intranucleotide	8	15	6	8	8	4	7	7	10	9	7	10	10	7	10	7	10	12	7	10	9	10	8	11
internucleotide[b]	0	3	6	3	6	0	1	3	3	6	3	4	4	0	3	4	3	6	4	3	0	3	3	4
cross-strand	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	0	0	0
total experimental	8	18	12	11	14	7	8	10	13	15	10	14	14	7	13	11	13	18	13	14	9	13	11	15

[a] Nucleotides $A^{-2}\rightarrow G^{11}$ are the modified strand, and nucleotides $C^{12}\rightarrow G^{22}$ are in the complementary strand. [b] The internucleotide NOEs are listed in the direction $n\rightarrow n-1$.

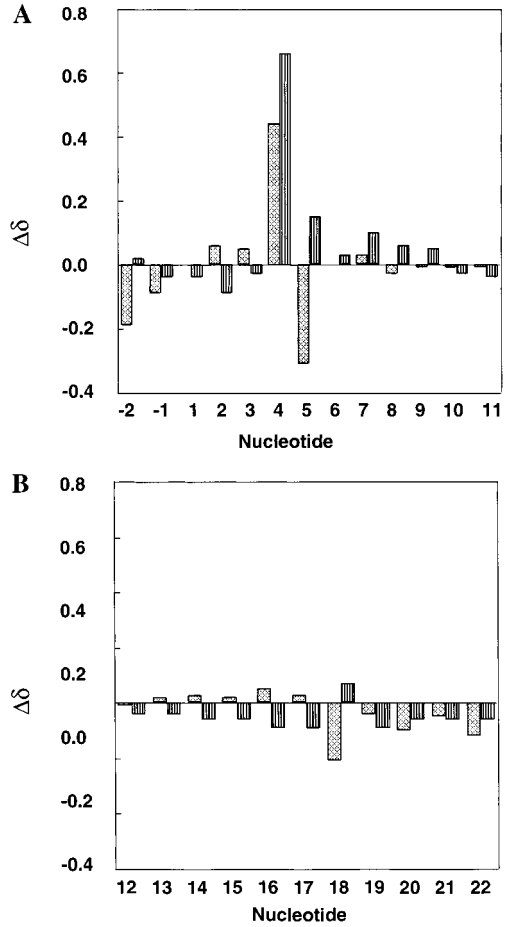


FIGURE 4: Chemical shift changes of nucleotide base protons relative to the unmodified (cross-hatched) and PdG-2BD (striped) oligodeoxynucleotide duplexes. $\Delta\delta = [\delta_{\text{unmodified oligodeoxynucleotide}} - \delta_{\text{modified oligodeoxynucleotide}}]$ (ppm).

torsion angle restraints obtained from the NMR spectra as described above. Watson−Crick hydrogen bonding restraints were also included. Randomly seeded rMD calculations were initiated from both A-form and B-form DNA starting structures. The simulated annealing procedure was essentially similar to that reported by Weisenseel et al. for the corresponding PdG-2BD oligodeoxynucleotide (*26*). It consisted of a total of 50 ps of rMD including 30 ps of high-temperature rMD, 5 ps of cooling, and 15 ps of low-temperature rMD. The high temperature was 2500 K, and the low temperature was 300 K. The temperature was controlled by coupling the molecules to a temperature bath with a coupling constant of 0.05 ps (*43*). Ten structures were calculated from each starting structure, each being assigned a different random seed for the generation of the initial atomic velocity vectors.

Calculations initiated from either starting structure converged to similar families of structures. These were examined

Table 3: Root Mean Square (rms) Deviations, Excluding the End Base Pairs, between Various Initial Structures, Intermediate Structures, and Final Average Structures of the M_1G-2BD Oligodeoxynucleotide

	atomic rms difference (Å)
initial structures	
IniA vs IniB	5.60
rms shifts	
IniA vs ⟨rMDA⟩[a]	4.39 ± 0.15
IniB vs ⟨rMDB⟩[b]	2.65 ± 0.04
rms distributions	
⟨rMDA⟩ vs ⟨rMDA⟩	0.22 ± 0.10
⟨rMDB⟩ vs ⟨rMDB⟩	0.10 ± 0.05
⟨rMDA⟩ vs ⟨rMDB⟩	1.32 ± 0.10
⟨rMDA⟩ vs rMD[c]	0.70 ± 0.08
⟨rMDB⟩ vs rMD	0.64 ± 0.04

[a] ⟨rMDA⟩ represents the set of 5 structures that emerged from rMD calculations starting with IniA. [b] ⟨rMDB⟩ represents the set of 5 structures that emerged from rMD calculations starting from IniB. [c] rMD represents the average minimized structure from all 10 rMD calculations.

by pairwise rmsd comparisons tabulated in Table 3. A rmsd of 5.60 Å was measured between the A-form and B-form starting structures. The rmsd between the set of structures emergent from the A-form vs B-form calculations was 1.32 Å. Comparison of the individual sets of structures derived from A-form DNA and B-form DNA between themselves resulted in rmsd values of 0.22 Å for A-form DNA and 0.10 Å for B-form DNA. Comparison of the starting structures with the average structures, which emerged from the rMD calculations starting from either, revealed that the rMD protocol resulted in a greater conformational change for the A-form starting structure than for the B-form starting structure. Hence, the final structures were closer to B-form geometry than to A-form geometry. The B-form starting structure differed from the structures emergent from the rMD calculations, evidenced by an rmsd of 2.65 Å. This was attributed to the presence of the two-base bulge. Figure 5 shows a stereoview of the overall averaged structure for M_1G-2BD as determined from the rMD calculations.

A final structure was obtained by averaging the coordinates from 10 rMD calculations, followed by 200 iterations of potential energy minimization using conjugate gradients. This is shown in CPK representation in Figure 6, in which the view is from the major groove of the duplex. The structural data suggest that the two-base bulge introduced a localized perturbation to the M_1G-2BD oligodeoxynucleotide. Base pairs $A^{-2}\cdot T^{22}$, $T^{-1}\cdot A^{21}$, $C^{1}\cdot G^{20}$, $G^{2}\cdot C^{19}$, and $C^{3}\cdot G^{18}$, located in the 5′-direction from M_1G, form a B-like Watson−Crick hydrogen-bonded duplex. Likewise, base pair $G^{6}\cdot C^{17}$, $G^{7}\cdot C^{16}$, $C^{8}\cdot G^{15}$, $A^{9}\cdot T^{14}$, $T^{10}\cdot A^{13}$, and $G^{11}\cdot C^{12}$, located in the 3′-direction, also form a B-like hydrogen bonded-duplex. At the lesion site, nucleotides M^4 and C^5 are contained within

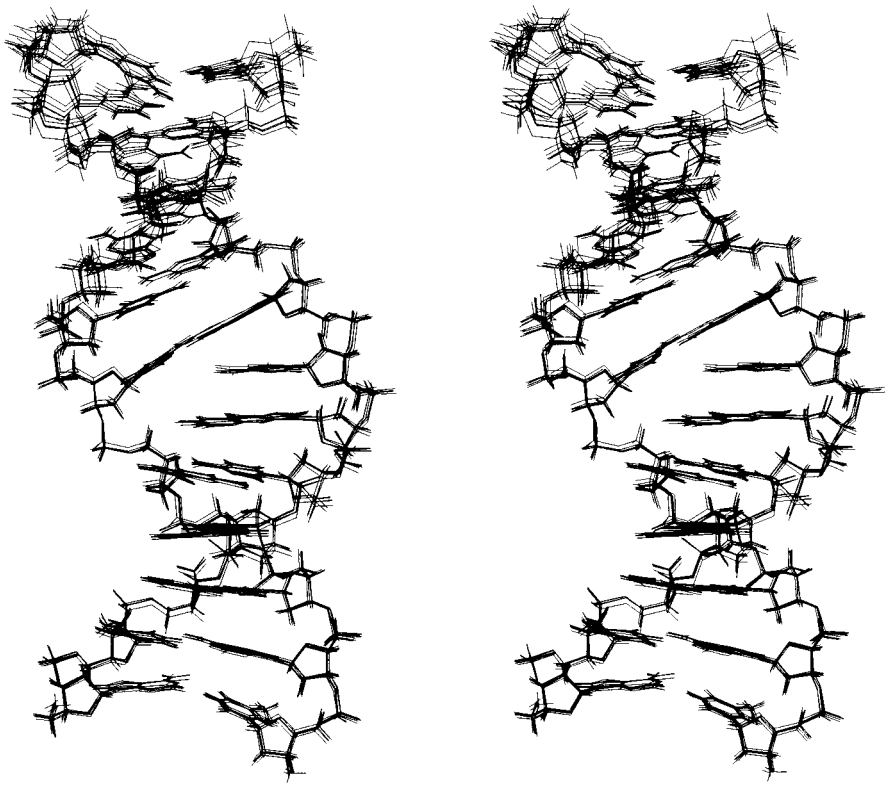


FIGURE 5: Stereoview of the M_1G-2BD duplex showing six superimposed structures obtained from rMD calculations.

a two-base bulge. At the bulged site, M_1G is inserted into the duplex, whereas C^5 appears to be extruded toward the major groove.

Complete Relaxation Matrix Calculations. The results of back-calculation of the intensity matrix from the converged rMD structures are tabulated in Table 4. A plot of the distribution of the sixth-root *R* factor, R_1^x, measured by the difference between NOE intensities calculated by CORMA from the model structure and the intensities measured from the NOESY spectra (*44*), is shown in Figure 7. The inverse sixth-root deviations were calculated with respect to the NOESY experiment measured with a mixing time of 250 ms, using an isotropic molecular correlation time of 4 ns. In general, they indicated reasonable agreement between the calculated intensities and the NOE data. In comparing R_1^x for the two starting structures, the value for the B-DNA starting structure IniB was lower than the R_1^x factor for the A-DNA starting structure IniA. Both were on the order of 0.1 or greater. The values of the total R_1^x decreased to 0.094 and 0.098, respectively, for the rMD structures emergent from either the B-DNA or the A-DNA starting structures. The sixth-root index indicated better agreement with the intraresidue NOE intensities than for the interresidue intensities.

DISCUSSION

The covalent modification of the DNA duplex by reactive electrophiles such as malondialdehyde is presumed to facilitate errors in DNA replication and repair, thus generating a variety of base substitution and frameshift mutations in the genome. When the malondialdehyde adduct M_1G is placed opposite deoxycytosine in DNA, the DNA duplex facilitates its rearrangement into the OPG adduct (*22*). Thus, the DNA duplex can chemically convert one type of damage into another type of damage. Recently, it was reported that the exocyclic γ-OH-1,N^2-propano-dG adduct, which represents the major adduct formed in DNA by acrolein, also undergoes spontaneous ring-opening in duplex DNA (*23*). That observation suggests that many exocyclic ethano and propano lesions in duplex DNA may exist as the corresponding ring-opened structures.

The M_1G and OPG lesions arising from malondialdehyde are chemically distinct and also introduce differing structural perturbations into DNA. M_1G is a bulky lesion that prevents normal Watson−Crick base pairing. In contrast, OPG introduces a relatively minor structural perturbation into the minor groove of the DNA duplex and does not interfere with Watson−Crick base pairing (*24*). Consequently, one might predict that the biological processing of malondialdehyde or base-propenal-induced damage sites in DNA depends on whether the damage exists as the exocyclic lesion M_1G or as the corresponding OPG derivative. Data from site-specific mutagenesis experiments suggest that this is the case. In comparing numbers of mutations from site-specific mutagenesis experiments involving M_1G which can open to the OPG derivative to PdG which cannot, it was found that the PdG adduct gave greater numbers of point mutations (*21, 48*). Other studies with the acrolein γ-hydroxyl-1,N^2-propano-2′-deoxyguanosine adduct, which was also reported to undergo ring-opening (*23*), were found to not be miscoding in vivo (*49, 50*).

The interest in the structure of M_1G embedded into the frameshift-prone *hisD3052*-iterated $(CG)_3$ repeat sequence arose from the observation that MDA, a small alkylating agent, induced frameshift mutations (*25*) despite the fact that these are more commonly associated with DNA intercalators. The iterated repeat contained in the *hisD3052* oligodeoxynucleotide is thought to be prone to frameshifts by slippage

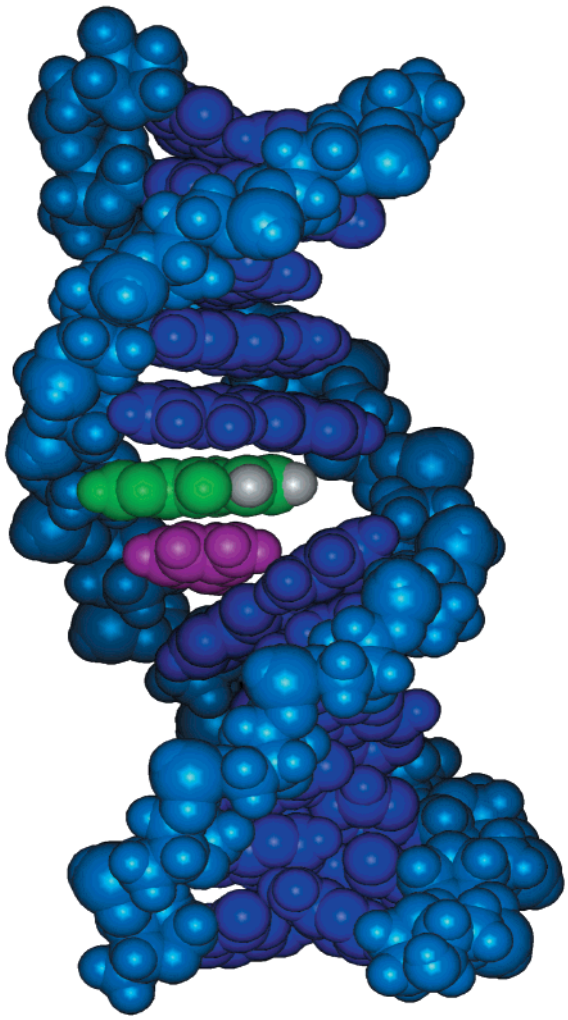


FIGURE 6: Space-filling model of the average structure of the M_1G-2BD oligodeoxynucleotide generated from restrained molecular dynamics calculations. M_1G is shown in green, with the exocyclic ring protons H6, H7, and H8 shown in white. The 3′-neighbor nucleotide C^5 is shown in red.

Table 4: Comparison of Sixth-Root Residual Indices, R_1^x, for Starting Models and Resulting rMD Structures[a]

structure	intraresidue R_1^x	interresidue R_1^x	total R_1^x [b]
IniA	0.120	0.171	0.132
IniB	0.088	0.154	0.104
rMDA	0.094	0.112	0.098
rMDB	0.088	0.113	0.094
rMD	0.089	0.112	0.095

[a] Only the inner 11 base pairs were used in the calculations, to exclude end effects. The mixing time was 250 ms. [b] $R_1^x = \sum|(a_o)_i^{1/6} - (a_c)_i^{1/6}|/\sum|(a_o)_i^{1/6}|$, where (a_o) and (a_c) are the intensities of observed (nonzero) and calculated NOE cross-peaks, respectively.

of either the template or the primer strand during DNA replication (*51*). The *hisD3052* mutation arose from the histidinol dehydrogenase gene of *S. typhimurium* by deletion of a cytosine induced by ICR-191 (*52, 53*). It is reversed by additions and deletions that restore the reading frame but do not necessarily reverse the forward mutation (*54*). The most common reversion is a CG deletion in the reiterated sequence $(CG)_4$ (*55*−*58*). M_1G-2BD models the intermediate structure leading to a two-base deletion, and in which the structural analogue PdG was examined (*26, 45*).

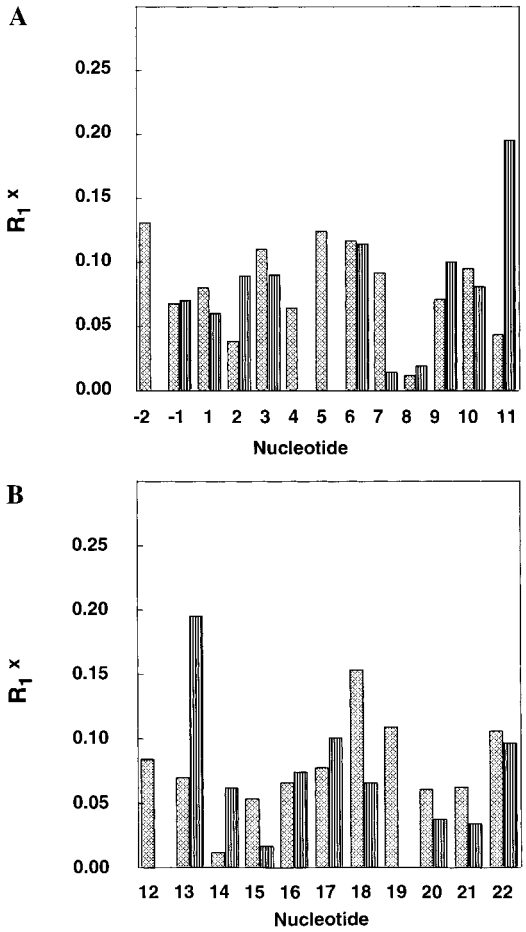


FIGURE 7: Intra- (cross-hatched) and inter-nucleotide (striped) sixth-root R_1^x factors as a function of nucleotide position in the M_1G-2BD oligodeoxynucleotide. (A) The modified strand. (B) The complementary strand.

Structure of the M_1G-2BD Oligodeoxynucleotide. (a) A Localized Bulge at M^4C^5. The observation of resonances from thymine imino protons T^{-1} N3H, T^{10} N3H, and T^{14} N3H indicated formation of a stable duplex on both sides of the unpaired bases M^4C^5 (Figure 2). Theoretically, there were two possible base pairing arrangements for the M_1G-2BD oligodeoxynucleotide resulting in either a C^3M^4 or a M^4C^5 bulge. Only one set of resonances was observed in the ^1H NMR spectra at 20 °C. This suggested a single conformation for the unpaired bases. An alternative explanation, fast exchange between a C^3M^4 and a M^4C^5 bulge, predicted the observation of transferred NOEs, with intensity weighted by the position of the equilibrium and the inverse sixth power of the respective internuclear distances for each of the potential conformations. These were not observed. The observed NOEs were consistent with M^4C^5 being unpaired in a single conformation.

The chemical shifts of the aromatic H8, H6, and H5 and sugar H1′, H2′, and H2″ protons of the unmodified oligodeoxynucleotide were determined at 20 °C from a NOESY experiment and were used to calculate the chemical shift differences ($\Delta\delta_{ppm} = \delta_{ppm\ M_1G\ dinucleotide\ deletion\ duplex} - \delta_{ppm\ unmodified\ duplex}$). The chemical shifts of DNA protons in the major groove (pyrimidine H6 and purine H8) are perturbed at M^4C^5 and on the complementary strand at C^{17} and C^{19}. The chemical shifts of the H2′, and H2″ sugar protons showed particularly large perturbations at C^{17} indicative of

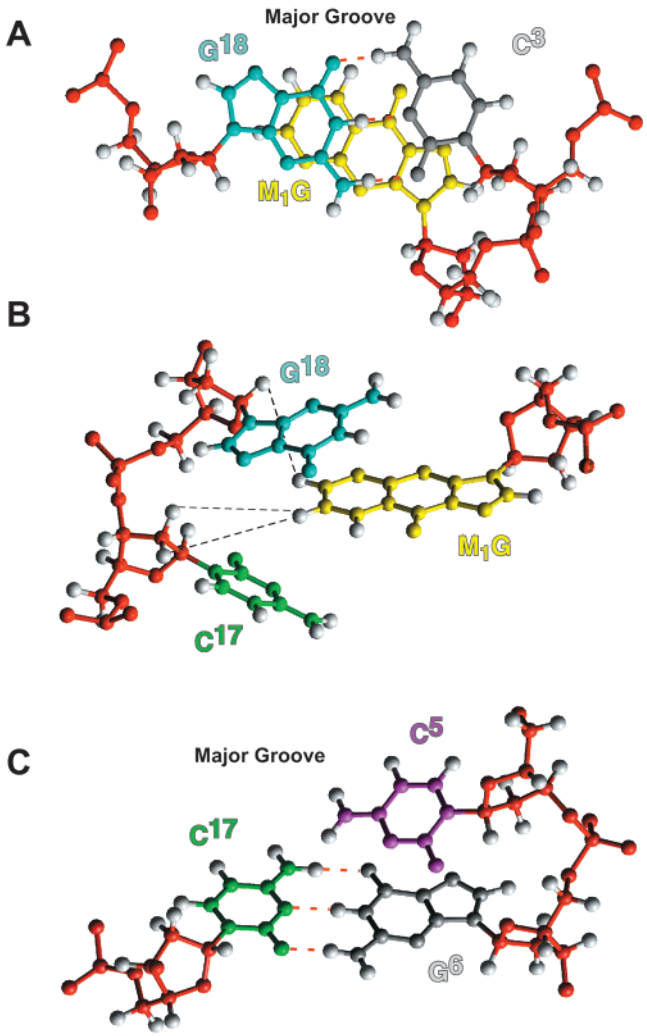


FIGURE 8: Stacking patterns of the M_1G lesion relative to DNA base pairs in the M_1G-2BD oligodeoxynucleotide. (A) M_1G (yellow) is intercalated below the C^3•G^{18} base pair. Note the position of the exocyclic ring of M_1G beneath nucleotide G^{18} (blue). (B) M_1G (yellow) is inserted into the duplex between nucleotides C^{17} (green) and G^{18} (blue) in the complementary strand. The dashed lines show observed NOEs between M_1G and the complementary strand. (C) C^5 (magenta) is not base-paired. It is located above the G^6•C^{17} base pair, and shifted toward the major groove.

a structural perturbation at this position, consistent with a M^4C^5 bulge.

(b) An Intrahelical Orientation for M_1G. Figure 8 shows a detailed representation of the M^4C^5 bulge structure. A stereo depiction of Figure 8 is included in the Supporting Information. The orientation of M_1G was established from a combination of NOE and chemical shift data. Cross-strand NOEs from M_1G H7 to C^{17} H2′, from M_1G H7 to C^{17} H2″, located in the 3′ direction from M_1G, and from M_1G H6 to G^{18} H1′, located in the 5′ direction from M_1G, were consistent with an intrahelical conformation for M_1G (Figure 4). The methylene protons of M_1G were shifted significantly upfield relative to their resonance frequencies in a single-stranded oligodeoxynucleotide containing M_1G; these upfield shifts were as large as 1.7 ppm for M_1G H7. This large upfield shift was attributed to ring current effects from the neighboring base pair. M_1G was concluded to be in the anti conformation about the glycosyl torsion angle. In the syn conformation, M_1G H2 would be proximate to the deoxyribose H1′ proton, and a strong NOE would be detectable. The intensity of the NOESY cross-peak between M_1G H2 and M_1G H1′ was similar to the corresponding purine H8 to purine H1′ cross-peaks from the other nucleotides, consistent with this conclusion. In this arrangement, the exocyclic ring of M_1G is located beneath the aromatic ring of G^{18}, which accounts for the upfield chemical shifts observed for the exocyclic protons of M_1G.

(c) Extrusion of the Unpaired Cytosine toward the Major Groove. The position of the bulged cytosine C^5 is not well-defined at the present level of refinement. The available data suggest this base extrudes toward the major groove (Figure 8). The observation of two weak NOEs from the M_1G H7 and H8 protons to the G^6 N1H imino proton, two bases removed from M_1G in the 3′ direction, provided indirect evidence for an extrahelical conformation for C^5. There was an NOE cross-peak from C^5 H1′ to G^6 H8, but it was weak, indicating a greater-than-normal separation between these two protons.

Structure−Activity Relationship. The observation that M_1G is chemically stable in the M_1G-2BD oligodeoxynucleotide at pH 7 is significant. We previously reported that the cytosine complementary to M_1G in fully duplex DNA plays a critical role in accelerating the rate of ring-opening to OPG (*24*). In duplex DNA at neutral pH, opening of the exocyclic ring to OPG occurred rapidly when cytosine was placed opposite M_1G. At 5 °C, ring-opening to OPG was 90% complete in <5 min (*22*). Additional experiments revealed that OPG did not form in fully duplex DNA when dT was mismatched with M_1G. However, if OPG was placed opposite dT, it did not spontaneously close to M_1G, suggesting that once formed in duplex DNA, the OPG lesion was thermodynamically favored as compared to M_1G (*24*).

The data provide a rationale as to why formation of OPG is not facilitated in the M_1G-2BD oligodeoxynucleotide, whereas in the fully complementary DNA duplex it is (*24*). The structural data suggest that in the M_1G-2BD oligodeoxynucleotide, neither cytosine C^3 or C^5 in the modified strand nor C^{17} in the complementary strand is appropriately positioned to assist in the opening of M_1G. C^3 forms a stable base pair with G^{18} and thus is not available to catalyze ring-opening. Likewise, C^{17} forms a stable base pair with G^6 and also is not available. Under these circumstances, M_1G remains intact. This is consistent with the observation that M_1G is chemically stable in nucleotides and in single-stranded oligodeoxynucleotides at neutral pH. Even under basic conditions (pH 10), formation of OPG occurs slowly, occurring on a time scale of hours.

Comparison with the PdG-2BD Oligodeoxynucleotide. The M_1G-2BD structure shows a number of similarities to the previously examined 1,N^2-propanodeoxyguanosine PdG-2BD duplex structure (*26*, *45*). The earlier studies utilized PdG as a model for M_1G (*59*). Unlike M_1G, PdG cannot undergo ring-opening to form the OPG lesion. In both instances, the two-base bulge was localized at the modified deoxyguanosine and its 3′-neighbor cytosine. The flanking sequences on either side of the two-base bulge were in both instances characterized as normal B-type duplexes. In both instances, the exocyclic M_1G or PdG lesions were inserted into the DNA duplex, and the 3′-neighbor cytosine was unpaired, and extruded toward the major groove. The similarities between

the M_1G-2BD and PdG-2BD duplexes were also reflected in their T_m values (Table 1). In both the M_1G-2BD and PdG-2BD duplexes, the localized bulge at positions X^4C^5 was accompanied by a destabilization of the modified bulge duplexes as compared to the corresponding unmodified 13−11-mer. The NMR spectrum of the unmodified 13−11-mer oligodeoxynucleotide showed evidence of transient bulge migration in which the position of the two-base bulge in the iterated repeat sequence was delocalized (*28*, *29*). The results indicated that site-specific adduction by the exocyclic M_1G adduct at position M^4 hindered transient bulge migration within the $(CpG)_3$-iterated repeat sequence.

Adduct-Induced DNA Bending. An inspection of the structures that emerged from the rMD calculations suggested bending of the duplex axis at the site of the two-base bulge. However, the failure to observe perturbations in the ^{31}P chemical shift or ^3J coupling data associated with the adduct suggests this result should be interpreted cautiously. A ligation ladder electrophoretic mobility assay was not performed for the M_1G-2BD oligodeoxynucleotide. The possibility that the two-base bulge in the M_1G-2BD oligodeoxynucleotide induces a flexible hinge (*60*) into the DNA duplex cannot be ruled out at this juncture. We previously reported the bulged bases in the PdG-2BD oligodeoxynucleotide induced an apparent bend in the DNA, evidenced by altered electrophoretic mobilities in a ligation assay (*45*). Subsequent rMD calculations on the PdG-2BD oligodeoxynucleotide suggested that bending was mediated by a change in the C3′−O3′ torsion angle ϵ at nucleotide C^{17} in the complementary strand (*26*). The resulting structural perturbation to the PdG-2BD duplex was described as two duplex segments of DNA slightly bent with respect to each other, and connected by the X^4C^5 sequence.

Comparison to Other Oligodeoxynucleotides Containing Bulged Nucleotides. Previous studies examined oligodeoxynucleotides containing unpaired bases or bulges of various sequence context and length (*61*−*69*). The present data extend our understanding of how DNA successfully accommodates unpaired bases in both intrahelical and extrahelical environments (*65*), the conformation being dependent both upon the identity of the unpaired base and upon the specific DNA sequence flanking the site (*64*). M_1G models an unpaired purine base, and the observation that it is intrahelical agrees with previous studies showing unpaired purines generally adopt intrahelical conformations in solution (*64*, *70*, *71*). For unpaired pyrimidines, both intra- (*72*, *73*) and extrahelical conformations (*63*, *74*) have been observed. The role of flanking DNA sequences in determining the conformation of bulged pyrimidines was demonstrated by studies demonstrating that bulged pyrimidines embedded in A·T tracts adopt an extrahelical conformation (*62*, *75*).

Biological Implications. The frameshift mutations induced into the iterated CG repeat sequence of the *hisD3052* gene are typically −2 deletion events. Recently, Riggins and Marnett reported that M_1G and its ring-opened derivative OPG block replication by human DNA polymerase β and induce frameshift mutations in vitro.[2] The frameshifts are consistent with mechanisms whereby the modified guanine and adjacent cytosine undergo transient dislocation during replication bypass. The present work demonstrates formation of a stable structure in duplex DNA in which M^4 and C^5 are dislocated opposite a two-base deletion in the complementary strand. The extent to which such transient dislocation might occur during replication is expected to determine the frequency by which M_1G induces −2-base deletions in the *hisD3052* gene. A number of variables can affect strand-misalignment fidelity during replication, including the length of the iterated repeat sequence, the content of the iterated sequence, and the specific DNA polymerase in question. The occurrence of a transient dislocation of M^4 and C^5 during replication could occur either prior to insertion of a nucleotide opposite M_1G or, alternatively, after nucleotide insertion and prior to extension (*76*). It could also involve enzyme dissociation or reassociation with the replication complex.

The similarities between the M_1G-2BD and PdG-2BD duplexes suggest that PdG provides a reasonable *structural* model for M_1G in duplex DNA, even though it provides a poor chemical model in that it cannot undergo ring-opening as does M_1G. It is interesting that PdG is not a strong inducer of −2 frameshifts (*76*). In vitro replication studies with human pol β suggest that M_1G is a strong block to replication, but also can induce low levels of frameshifts.[2] One might speculate that the suppression of transient bulge migration within the M_1G-2BD and PdG-2BD oligodeoxynucleotides, as compared to the corresponding unmodified oligodeoxynucleotide, reduces the potential for frameshift mutagenesis by these adducts. Whether this correlates with suppression of transient strand slippage during replication is not known. On the other hand, malondialdehyde-induced frameshifts might also involve the OPG lesion.[2] Future experiments will examine the behavior of the OPG in the corresponding OPG-2BD oligodeoxynucleotide.

SUMMARY

The M_1G adduct is stable in the M_1G-2BD duplex, in which it is positioned opposite a two-base deletion in the complementary strand. The position of the resulting two-base bulge is fixed at positions M^4C^5. The stability of the M_1G with respect to ring-opening to the OPG derivative corroborates earlier observations that the positioning of deoxycytosine in duplex DNA opposite M_1G plays a critical role in mediating formation of the OPG lesion. The structural and thermodynamic similarities between the M_1G-2BD duplex and the previously examined PdG-2BD duplex are intriguing. In both instances, transient bulge migration is hindered.

ACKNOWLEDGMENT

Mr. Markus Voehler assisted with NMR spectroscopy. Dr. Jason P. Weisenseel assisted with the structural refinement.

SUPPORTING INFORMATION AVAILABLE

The Supporting Information material consists of Tables S1 and S2, which detail the ^1H NMR chemical shift assignments; Table S3, the experimental distances and classes of restraints for the M_1G-2BD oligodeoxynucleotide; and Figures S1, which shows force field parametrization values used for M_1G; S2, which shows the expansion of the H1′−H2′H2″ region of the ^1H DQF-COSY spectrum; and S3, which shows a stereoview of Figure 8 (18 pages). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Marnett, L. J. (1999) *IARC Sci. Publ. 150*, 17−27.
2. Basu, A. K., O'Hara, S. M., Valladier, P., Stone, K., Mols, O., and Marnett, L. J. (1988) *Chem. Res. Toxicol. 1*, 53−59.
3. Marnett, L. J., Basu, A. K., O'Hara, S. M., Weller, P. E., Rahman, A. F. M. M., and Oliver, J. P. (1986) *J. Am. Chem. Soc. 108*, 1348−1350.
4. Seto, H., Okuda, T., Takesue, T., and Ikemura, T. (1983) *Bull. Chem. Soc. Jpn. 56*, 1799−1802.
5. Seto, H., Seto, T., Takesue, T., and Ikemura, T. (1986) *Chem. Pharm. Bull. 34*, 5079−5085.
6. Reddy, G. R., and Marnett, L. J. (1996) *Chem. Res. Toxicol. 9*, 12−15.
7. Dedon, P. C., Plastaras, J. P., Rouzer, C. A., and Marnett, L. J. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 11113−11116.
8. Plastaras, J. P., Riggins, J. N., Otteneder, M., and Marnett, L. J. (2000) *Chem. Res. Toxicol. 13*, 1235−1242.
9. Wang, M. Y., and Liehr, J. G. (1995) *Arch. Biochem. Biophys. 316*, 38−46.
10. Chaudhary, A. K., Nokubo, M., Reddy, G. R., Yeola, S. N., Morrow, J. D., Blair, I. A., and Marnett, L. J. (1994) *Science 265*, 1580−1582.
11. Wang, M., Dhingra, K., Hittleman, W. N., Liehr, J. G., de Andrade, M., and Li, D. (1996) *Cancer Epidemiol. 5*, 705−710.
12. Nath, R. G., and Chung, F. L. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 7491−7495.
13. Nath, R. G., and Chung, F. L. (1993) *Proc. Am. Assoc. Cancer Res. 34*, 137.
14. O'Nair, J., Barbin, A., Guichard, Y., and Bartsch, H. (1995) *Carcinogenesis 16*, 613−617.
15. Chaudhary, A. K., Nokubo, M., Marnett, L. J., and Blair, I. A. (1994) *Biol. Mass Spectrom. 23*, 457−464.
16. Rouzer, C. A., Chaudhary, A. K., Nokubo, M., Ferguson, D. M., Reddy, G. R., Blair, I. A., and Marnett, L. J. (1997) *Chem. Res. Toxicol. 10*, 181−188.
17. Vaca, C. E., Fang, J. L., Mutanen, M., and Valsta, L. (1995) *Carcinogenesis 16*, 1847−1851.
18. Fang, J. L., Vaca, C. E., Valsta, L. M., and Mutanen, M. (1996) *Carcinogenesis 17*, 1035−1040.
19. Sevilla, C. L., Mahle, N. H., Eliezer, N., Uzieblo, A., O'Hara, S. M., Nokubo, M., Miller, R., Rouzer, C. A., and Marnett, L. J. (1997) *Chem. Res. Toxicol. 2*, 172−180.
20. Chaudhary, A. K., Reddy, R. G., Blair, I. A., and Marnett, L. J. (1996) *Carcinogenesis 17*, 1167−1170.
21. Fink, S. P., Reddy, G. R., and Marnett, L. J. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 8652−8657.
22. Mao, H., Schnetz-Boutaud, N. C., Weisenseel, J. P., Marnett, L. J., and Stone, M. P. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 6615−6620.
23. de los Santos, C., Zaliznyak, T., and Johnson, F. (2001) *J. Biol. Chem. 276*, 9077−9082.
24. Mao, H., Reddy, G. R., Marnett, L. J., and Stone, M. P. (1999) *Biochemistry 38*, 13491−13501.
25. O'Hara, S. M., and Marnett, L. J. (1991) *Mutat. Res. 247*, 45−56.
26. Weisenseel, J. P., Moe, J. G., Reddy, G. R., Marnett, L. J., and Stone, M. P. (1995) *Biochemistry 34*, 50−64.
27. Woodson, S. A., and Crothers, D. M. (1988) *Biochemistry 27*, 436−445.
28. Woodson, S. A., and Crothers, D. M. (1987) *Biochemistry 26*, 904−912.
29. Woodson, S. A., and Crothers, D. M. (1988) *Biochemistry 27*, 3130−3141.
30. Reddy, G. R., and Marnett, L. J. (1995) *J. Am. Chem. Soc. 117*, 5007−5008.
31. Schnetz-Boutaud, N. C., Mao, H., Stone, M. P., and Marnett, L. J. (2000) *Chem. Res. Toxicol. 13*, 90−95.
32. Borer, P. N. (1975) in *Handbook of biochemistry and molecular biology*, CRC Press, Cleveland, OH.
33. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Mol. Biol. 6*, 661−665.
34. Arnott, S., and Hukins, D. W. L. (1972) *Biochem. Biophys. Res. Commun. 47*, 1504−1509.
35. Borgias, B. A., and James, T. L. (1990) *J. Magn. Reson. 87*, 475−487.
36. Schmitz, U., and James, T. L. (1995) *Methods Enzymol. 261*, 3−44.
37. Kim, S. G., Lin, L. J., and Reid, B. R. (1992) *Biochemistry 31*, 3564−3574.
38. Gorenstein, D. G. (1992) *Methods Enzymol. 211*, 254−286.
39. Tonelli, M., Ragg, E., Bianucci, A. M., Lesiak, K., and James, T. L. (1998) *Biochemistry 37*, 11745−11761.
40. Brunger, A. T. (1992) in *X-Plor. Version 3.1. A system for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.
41. Clore, G. M., Brunger, A. T., Karplus, M., and Gronenborn, A. M. (1986) *J. Mol. Biol. 191*, 523−551.
42. Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977) *J. Comput. Phys. 23*, 327−341.
43. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) *J. Phys. Chem. 81*, 3684−3690.
44. Keepers, J. W., and James, T. L. (1984) *J. Magn. Reson. 57*, 404−426.
45. Moe, J. G., Reddy, G. R., Marnett, L. J., and Stone, M. P. (1994) *Chem. Res. Toxicol. 7*, 319−328.
46. Reid, B. R. (1987) *Q. Rev. Biophys. 20*, 2−28.
47. Patel, D. J., Shapiro, L., and Hare, D. (1987) *Q. Rev. Biophys. 20*, 35−112.
48. Fink, S. P., Reddy, G. R., and Marnett, L. J. (1996) *Chem. Res. Toxicol. 9*, 277−283.
49. VanderVeen, L. A., Hashim, M. F., Nechev, L. V., Harris, T. M., Harris, C. M., and Marnett, L. J. (2001) *J. Biol. Chem. 276*, 9066−9070.
50. Yang, I. Y., Hossain, M., Miller, H., Khullar, S., Johnson, F., Grollman, A., and Moriya, M. (2001) *J. Biol. Chem. 276*, 9071−9076.
51. Streisinger, G., Okada, Y., Enrich, J., Newton, J., Tsugita, A., Terzaghi, E., and Inouye, M. (1966) *Cold Spring Harbor Symp. Quant. Biol. 31*, 77−84.
52. Oeschger, N. S., and Hartman, P. E. (1970) *J. Bacteriol. 101*, 490−504.
53. Hartman, P. E., Ames, B. N., Roth, J. R., Barnes, W. M., and Levin, D. E. (1986) *Environ. Mutagen. 8*, 631−641.
54. McCann, J., Spingarn, N. E., Koburi, J., and Ames, B. N. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 979−983.
55. Isono, K., and Yourno, J. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 1612−1617.
56. Fuscoe, J. C., Wu, R., Shen, N. H., Healy, S. K., and Felton, J. S. (1988) *Mutat. Res. 201*, 241−251.
57. Bell, D. A., Levine, J. G., and DeMarini, D. M. (1991) *Mutat. Res. 252*, 35−44.
58. DeMarini, D. M., Abu-Shakra, A., Gupta, R., Hendee, L. J., and Levine, J. G. (1992) *Environ. Mol. Mutagen. 20*, 12−18.
59. Singh, U. S., Moe, J. G., Reddy, G. R., Weisenseel, J. P., Marnett, L. J., and Stone, M. P. (1993) *Chem. Res. Toxicol. 6*, 825−836.
60. Le, P. T., Harris, C. M., Harris, T. M., and Stone, M. P. (2000) *Chem. Res. Toxicol. 13*, 63−71.
61. Patel, D. J., Pardi, A., and Itakura, K. (1982) *Science 216*, 581−590.
62. Morden, K. M., Gunn, B. M., and Maskos, K. (1990) *Biochemistry 29*, 8835−8845.
63. Morden, K. M., Chu, Y. G., Martin, F. H., and Tinoco, I., Jr. (1983) *Biochemistry 22*, 5557−5563.
64. Morden, K. M., and Maskos, K. (1993) *Biopolymers 33*, 27−36.
65. Rosen, M. A., Live, D., and Patel, D. J. (1992) *Biochemistry 31*, 4004−4014.
66. Rosen, M. A., Shapiro, L., and Patel, D. J. (1992) *Biochemistry 31*, 4015−4026.
67. Plum, G. E., Grollman, A. P., Johnson, F., and Breslauer, K. J. (1992) *Biochemistry 31*, 12096−12102.

68. Aboul-ela, F., Murchie, A. I., Homans, S. W., and Lilley, D. M. (1993) *J. Mol. Biol. 229*, 173−188.
69. Joshua-Tor, L., Frolow, F., Appella, E., Hope, H., Rabinovich, D., and Sussman, J. L. (1992) *J. Mol. Biol. 225*, 397−431.
70. Patel, D. J., Kozlowski, S. A., Marky, L. A., Rice, J. A., Broka, C., Itakura, K., and Breslauer, K. J. (1982) *Biochemistry 21*, 445−451.
71. Nikonowicz, E. P., Meadows, R. P., and Gorenstein, D. G. (1990) *Biochemistry 29*, 4193−4204.
72. van den Hoogen, Y. T., van Beuzekom, A. A., de Vroom, E., Van Der Marel, G. A., Van Boom, J. H., and Altona, C. (1988) *Nucleic Acids Res. 16*, 5013−5030.
73. van den Hoogen, Y. T., van Beuzekom, A. A., Van Den Elst, H., Van Der Marel, G. A., Van Boom, J. H., and Altona, C. (1988) *Nucleic Acids Res. 16*, 2971−2986.
74. Kalnik, M. W., Norman, D. G., Li, B. F., Swann, P. F., and Patel, D. J. (1990) *J. Biol. Chem. 265*, 636−647.
75. Maskos, K., Gunn, B. M., LeBlanc, D. A., and Morden, K. M. (1993) *Biochemistry 32*, 3583−3595.
76. Benamira, M., Singh, U., and Marnett, L. J. (1992) *J. Biol. Chem. 267*, 22392−22400.

BI011242U