JOC The Journal of Organic Chemistry

Article

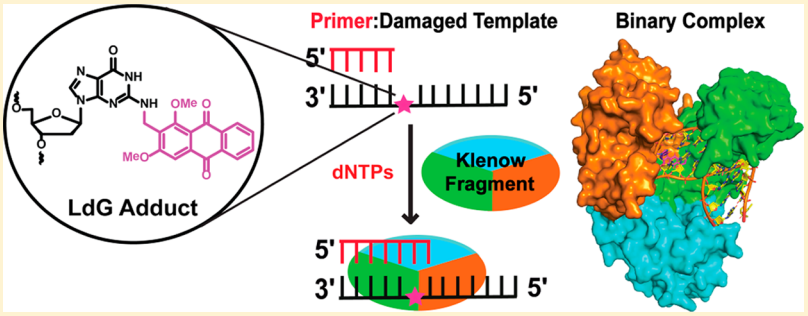
pubs.acs.org/joc

Synthesis and Polymerase-Mediated Bypass Studies of the N^2-Deoxyguanosine DNA Damage Caused by a Lucidin Analogue

Pratibha P. Ghodke, S. Harikrishna, and P. I. Pradeepkumar*

Department of Chemistry, Indian Institute of Technology Bombay, Mumbai 400076, India

Supporting Information



ABSTRACT: Lucidin is a genotoxic and mutagenic hydroxyanthraquinone metabolite, which originates from the roots of *Rubia tinctorum* L. (madder root). It reacts with exocyclic amino groups of DNA nucleobases and forms adducts/lesions leading to carcinogenesis. To study the effect of lucidin-induced DNA damage, herein, we report the first synthesis of a structural analogue of lucidin [N^2-methyl-(1,3-dimethoxyanthraquinone)-deoxyguanosine, LdG] embedded DNAs utilizing phosphoramidite strategy. LdG modification in a DNA duplex imparts destabilization (ΔT_m ~5 °C/modification), which is attributed to the unfavorable contribution from the enthalpy. Primer extension studies using the Klenow fragment (exo$^-$) of *Escherichia coli* DNA polymerase I demonstrate that bypass of LdG modification is error prone as well as slow compared to that across the unmodified sites. Molecular dynamics simulations of the binary complex of *Bacillus* fragment polymerase (homologue of the Klenow fragment) and LdG-DNA duplex elucidate the structural fluctuations imparted by the LdG lesion, as well as the molecular mechanism of bypass at the lesion site. Overall, the results presented here show that the lucidin adduct destabilizes DNA structure and reduces fidelity and processivity of DNA synthesis.

INTRODUCTION

The integrity of the genomic DNA is disturbed by continuous exposure to exogenous and endogenous agents, which leads to DNA damage. This is a major threat, and the first significant step of chemical carcinogenesis.[1] In general, the DNA damages are recognized by DNA polymerases during replication and are repaired by various DNA repair mechanisms.[2] If DNA damage sites are unrepaired, this can result in blockage, slippage, misinsertion, or mutations in daughter cells, leading to cancer.[3] Therefore, it is essential to understand the interaction of polymerases with damaged DNAs at the molecular level. Each DNA polymerase has a unique role and shows varying fidelity toward individual DNA damage.[4] Most of the replicative DNA polymerases are not able to tolerate the damages formed in the genomic DNA. However, DNA polymerases belonging to the Y-family can continue the DNA replication in the presence of damages employing a mechanism known as translesion synthesis (TLS) or bypass mechanism.[5] These Y-family polymerases are devoid of intrinsic proofreading exonuclease activities, which are the characteristics of replicative polymerases.[5,6]

The DNA damages can be caused by chemical carcinogens and mutagens, which are electrophilic in nature, or they get converted to electrophiles through various metabolic pathways.[7] These species react with nucleophilic sites of DNA, thereby altering the structural integrity by forming covalent bonds with the DNA bases known as adducts or lesions.[8] Various reactive metabolites of carcinogens show different reactivity toward the nucleophilic centers of DNA bases.[9] Among the various nucleophilic centers of DNA, the N^2 position of the guanine is vulnerable toward the attack by a wide variety of electrophiles.[9] Importantly, the N^2-dG DNA adducts are located in the minor groove of DNA, and depending upon their chemical and structural attributes, they have potential to disrupt standard Watson−Crick (W−C) base pairing, which lead to the stalling of the replication process.[10,11] For example, polycyclic aromatic hydrocarbons (PAH) such as the benzo[*a*]pyrene (BP) get converted to carcinogenic diol-epoxides (BPDE) and its *o*-quinone derivatives via different metabolic pathways to form N^2-dG and N^6-dA adducts.[12] Similarly, the number of other PAHs like the chrysene, benz[*a*]anthracene, aminofluorene (AF) etc., get activated enzymatically to their reactive metabolites, which can form N^2-dG adducts.[13,14]

Received: November 18, 2014
Published: January 9, 2015

DOI: 10.1021/jo502627b
J. Org. Chem. 2015, 80, 2128−2138

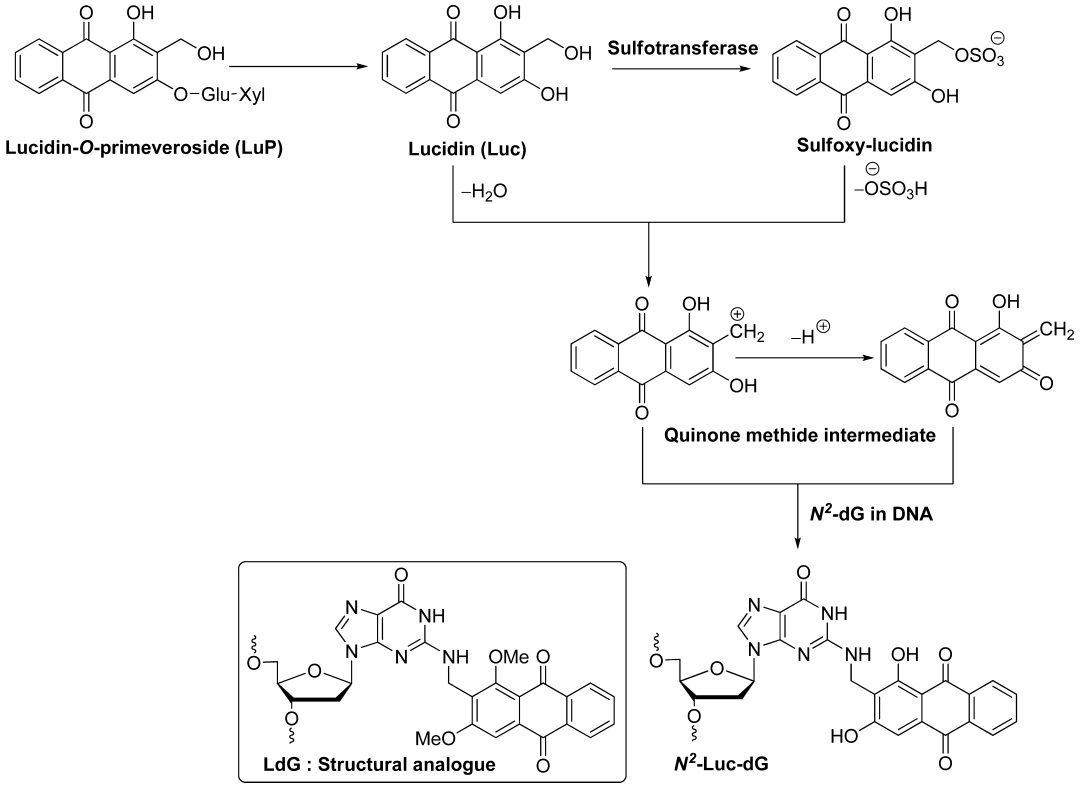
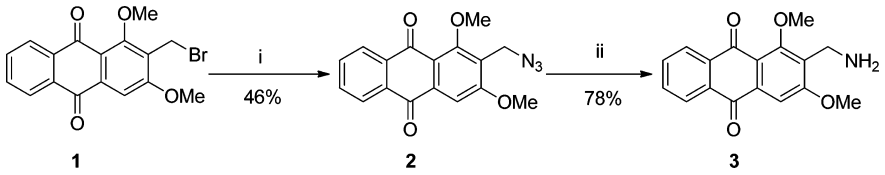


Figure 1. Proposed metabolic pathways for the formation of the N^2-Luc-dG adduct, and the structure of its methyl protected analogue (LdG).

Scheme 1. Synthesis of 2-(Aminomethyl)-1,3-dimethoxyanthraquinone 3[a]



[a]Reagents and conditions: (i) Sodium azide, DMF, rt, 30 h and (ii) PPh_3, THF:H_2O, rt, 2 h.

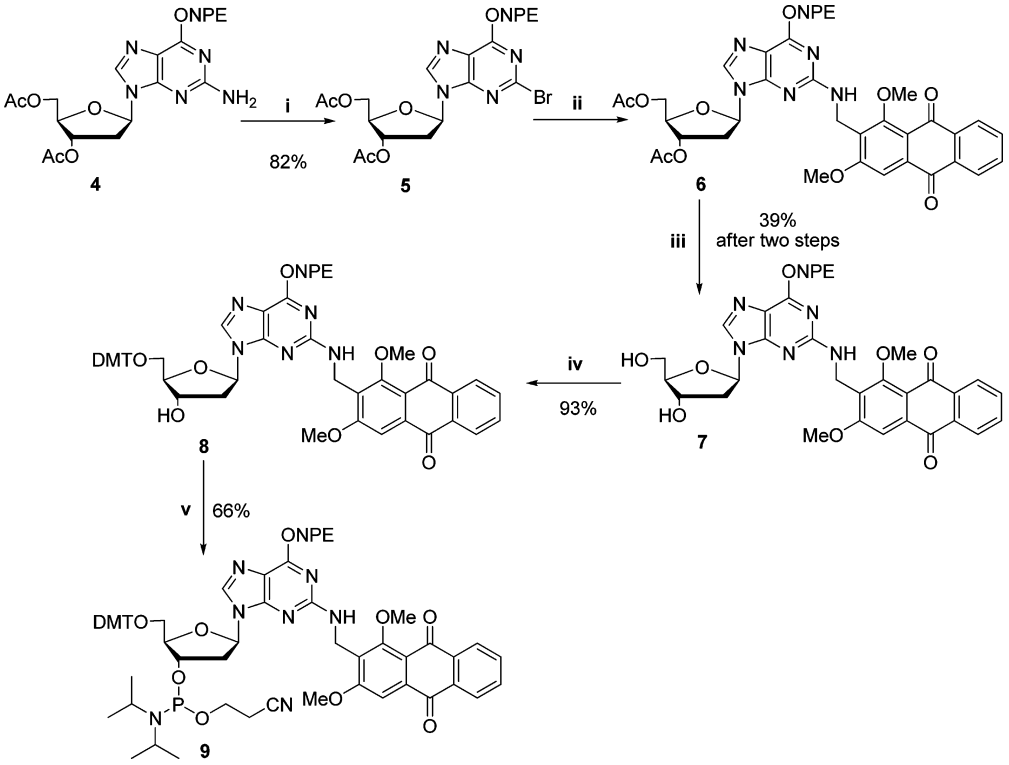
The red colored dye extract from roots of the *Rubia tinctorum* L. (madder root) has been widely used as a dying agent, as well as an additive in a variety of food and drinks.[15] The madder root contains various carcinogenic hydroxyantharaquinones and their glycoside conjugates like lucidin-*O*-primeveroside (LuP, Figure 1) and alizarinprimeveroside (AIP).[16] The LuP and AIP glycosides are hydrolyzed into different derivatives of hydroxyanthraquinones named lucidin and alizarin, respectively.[17] Lucidin is a potent carcinogen, and it is capable of forming N^2-dG and N^6-dA DNA adducts utilizing a possible sulfotransferase metabolic pathway or by a dehydration process (Figure 1).[18,19] Lucidin is found to be mutagenic and genotoxic to bacteria and mammalian cells and forms DNA adducts in mice.[20] Specific DNA adduct formation by lucidin under *in vivo* conditions has been studied by measuring the adduct levels in kidneys and livers of rats.[21] To study the lucidin-induced chemical carcinogenesis and the effect of this adduct on DNA replication, it is necessary to develop robust chemical methods to synthesize the lucidin-modified DNA. To this end, herein, we report the synthesis of a structural analogue of N^2-Luc-dG-modified DNA using phosphoramidite chemistry (LdG, Figure 1). Primer extension studies using Klenow fragment exo$^-$ (KF$^-$) of *Escherichia coli* DNA polymerase I have been utilized to study the DNA synthesis across the LdG adduct. The structural fluctuations imparted by the LdG modification at the primer−template interface during the polymerase action have been probed utilizing molecular modeling and dynamics simulations. The effects of the LdG adduct on the stability and thermodynamics of the DNA duplex have also been explored utilizing thermal melting experiments as well as molecular dynamics (MD) simulations.

RESULTS AND DISCUSSION

Synthesis of the LdG-Modified Phosphoramidite and DNAs. Our initial goal was to synthesize the N^2-Luc-dG (Figure 1)-modified phosphoramidite for its incorporation into DNA. Toward this end, it was planned to achieve the synthesis of 2-(aminomethyl)-1,3-diacetoxyanthraquinone, which is an acetyl-protected lucidin amine (Figure S1 of the Supporting Information) required to assemble the N^2-dG nucleoside via Buchwald-Hartwig coupling. This strategy was envisaged because the acetate group can easily be deprotected at the oligonucleotide level to get the lucidin-modified DNAs. But, all our efforts to synthesize the respective amine were not successful, perhaps due to steric hindrance encountered by one of the hydroxy group of the lucidin during acylation (Figure S1 of the Supporting Information).[22] To circumvent this, a methyl-protected structural analogue of lucidin (LdG, Figure 1) was

DOI: 10.1021/jo502627b
J. Org. Chem. 2015, 80, 2128−2138

Scheme 2. Synthesis of the LdG Phosphoramidite 9[a]



[a]Reagents and conditions: (i) $SbBr_3$, TBN, DBM, −10 to 0 °C, 4 h; (ii) compound **3**, $Pd(OAc)_2$, (*R*)-BINAP, Cs_2CO_3, toluene, 85 °C, 10 h; (iii) 33% $MeNH_2$ in EtOH (v/v), rt, 2 h; (iv) DMT-Cl, pyridine, rt, 10 h; and (v) CEP-Cl, DIPEA, DCM, rt, 1.5 h.

Table 1. T_m and the Thermodynamic Parameters of the Unmodified and LdG Modified DNA Duplexes[a]

code	duplex	T_m[a] (°C)	ΔT_m[b]/mod (°C)	$\Delta H°$ ($\Delta\Delta H°$)[c] (kcal/mol)	$\Delta S°$ (eu)	$T\Delta S°$ ($\Delta T\Delta S°$)[d] (kcal/mol)	$\Delta G°_{298}$ ($\Delta\Delta G°$)[e] (kcal/mol)
D1-D2	5′-GCCGGAATAGCGCA-3′	64.5±0.50	–	−66.0±0.59	−195±1.4	−58.2	−7.7±0.06
	3′-CGGCCTTATCGCGT-5′						
D3-D2	5′-GCCG**G***AATAGCGCA-3′	59.1±0.60	−5.4	−61.7±0.45	−185±0.7	−55.4	−6.3±0.04
	3′-CGGCCTTATCGCGT-5′			(+4.3)		(+2.8)	(+1.4)
D3-D4	5′-GCCG**G***AATAGCGCA-3′	54.4±0.47	−4.7	−44.3±0.63	−135±1.8	−40.3	−3.9±0.05
	3′-CGGCCTTATC**G***CGT-5′			(+21.7)		(+17.9)	(+3.8)

[a]The thermal denaturation was performed at 25 °C (298 K) using DNA duplex (5 μM) in phosphate buffer (100 mM NaCl, 20 mM sodium phosphate, 0.1 mM EDTA, pH 7.4). The LdG (**G***) modified nucleotide is highlighted as starred and shown in bold. [b]ΔT_m represents the [T_m (DNA_{mod})− T_m (DNA_{unmod})]. The T_m values reported are the average of 3 independent measurements with the estimated standard deviation. $\Delta H°$, $\Delta S°$, and $\Delta G°$ are average values emerged from 3 independent melting curve analysis using a two state model. The error propagation for $\Delta H°$, $\Delta S°$, and $\Delta G°_{298}$ were calculated using a reported protocol.[27] [c]$\Delta\Delta H°$, [d]$\Delta T\Delta S°$, and [e]$\Delta\Delta G°$ represent the differences in the thermodynamic parameters of modified duplex with respect to those of the unmodified duplex.

synthesized utilizing 2-(aminomethyl)-1,3-dimethoxyanthraquinone **3**, which was accessed from a bromo precursor **1** as outlined in Scheme 1.[22] A bromo group of **1** was displaced to form the azide **2** using sodium azide in DMF in a 46% yield. Staudinger reduction of azide compound **2** to the corresponding amine **3** was accomplished in a 78% yield using PPh_3 in the THF−water solvent system.

The protected bromo nucleoside **5** required for the Buchwald-Hartwig coupling was synthesized utilizing O^6-*para*-nitrophenylethyl (NPE)-protected deoxyguanosine **4** (Scheme 2).[23] Nucleoside **4** was subjected to diazotization followed by bromination using $SbBr_3$ and *t*-butyl nitrite (TBN) in dibromomethane (DBM) to furnish -bromo nucleoside **5** in 82% yield.[24] Coupling between **5** and dimethoxyanthraquinone **3**, in the presence of $Pd(OAc)_2$, (*R*)-BINAP, and Cs_2CO_3 in toluene afforded compound **6**.[25] The deacetylation of crude compound **6** was achieved using $MeNH_2$ in ethanol to obtain the diol nucleoside **7** in 39% yield (after two steps from **5**). The 5′-OH of **7** was protected using DMT-Cl in pyridine to afford tritylated nucleoside **8** in 93% yield. Nucleoside **8** was phosphitylated using 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (CEP-Cl) and *N*,*N*-diisopropylethylamine (DIPEA) in DCM to obtain the LdG phosphoramidite **9** in 66% yield.[26] The O^6-NPE group is retained in the phosphoramidite to enhance the solubility during the solid phase synthesis.

The LdG-modified phosphoramidite building block **9** was successfully incorporated into the DNA sequences (Table S6 of the Supporting Information) employing an automated DNA synthesizer. The sequences **D3** and **D4** are the LdG-modified 14-mer DNAs, which were synthesized to evaluate the thermal stability of the modified duplexes. The sequence **T** is a LdG-

DOI: 10.1021/jo502627b
J. Org. Chem. 2015, 80, 2128−2138

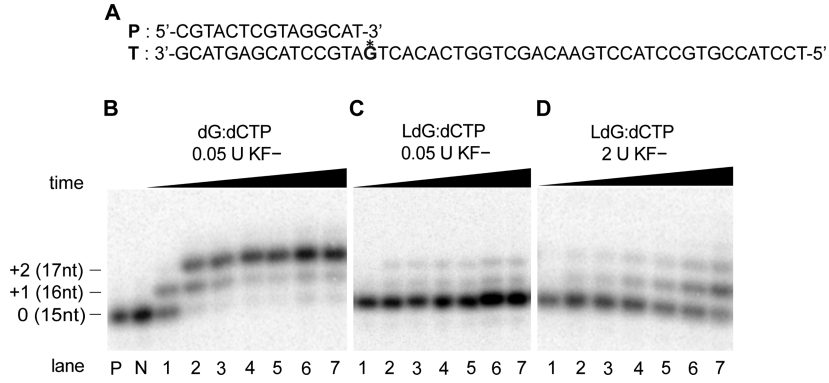


Figure 2. PAGE (20%, 7 M urea) of single nucleotide incorporation assay using KF⁻. (A) Complete sequence of the template, T (**G*** = LdG) and the primer, P; (B) unmodified template (dG:dCTP) with 0.05 U of KF⁻; (C) modified template (LdG:dCTP) with 0.05 U of KF⁻; (D) modified template (LdG:dCTP) with 2 U of KF⁻. Lane P, primer; lane N, negative reaction control (no dCTP); lanes 1 to 7, primer extension reactions with 100 μM dCTP in a different time course from 0.5, 5, 10, 30, 60, 120, and 240 min. All reactions were carried out at 37 °C in the polymerase buffer (50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9).

modified 50-mer DNA template utilized for the primer extension studies.

Thermal Melting Studies. To deduce the effect of the LdG modification on the DNA duplex stability, UV melting studies were carried out (Table 1 and Figure S2 of the Supporting Information). In comparison to the unmodified duplex (**D1**-**D2**), there was a significant decrease in T_m (ΔT_m ~5 °C/LdG unit) observed for LdG-modified duplexes (**D3**-**D2** and **D3**-**D4**). To understand the factors responsible for these effects, all the thermodynamic parameters ($\Delta H°$, $\Delta S°$, and $\Delta G°_{298}$) were calculated from the absorbance versus temperature melting curves, by fitting them into a two-state model (Figure S2 of the Supporting Information).[27] The decrease in T_m due to the presence of LdG modification was reflected in the +1.4 to 3.8 kcal/mol drop in $\Delta G°_{298}$ values. Thermodynamic parameters show that the LdG provides an entropic advantage ($\Delta T\Delta S°$ = +2.8 to 17.9 kcal/mol) to the duplex, while it had an adverse effect on enthalpy ($\Delta\Delta H°$ = +4.3 to 21.7 kcal/mol). Therefore, the net destabilization of the duplex can be attributed to the compensation of favorable contribution from entropy by the unfavorable contribution of enthalpy. Our molecular modeling and dynamics studies suggest that the LdG adduct imparts conformational flexibility due to which the base stacking, base pairing, and hydration were disturbed at the modified site (see the respective section for details). The observed thermodynamic parameters validate those findings.

Primer Extension Studies. To evaluate the bypass and the extension ability of the LdG-modified DNA by the KF⁻, single nucleotide incorporation as well as full length extension experiments were carried out. It has been reported that the KF⁻ is able to bypass various N^2-dG adducts, and it lacks proofreading activity, which results in the error prone replication.[10,28,29] Single nucleotide incorporation experiments were performed using the dCTP and LdG-modified DNA template by varying concentration of enzyme to study the correct base incorporation opposite to the LdG adduct (Figure 2, panels B–D). The primer–template system used in these experiments is shown in Figure 2A. In the reactions employing 0.05 units of KF⁻ with the unmodified template, both the incorporated (+1, 16nt) and extended (+2, 17nt) products were observed (Figure 2B). After 4 h, 10% of incorporation and 89% of extension were visible (lane 7, Figure 2B). These results validate the highly efficient incorporation as well as extension across the unmodified dG, even at a relatively low concentration of KF⁻. However, since the +2 site of the template contains dT, the extension of primer in the presence of dCTP suggests low fidelity of the KF⁻ during the extension process. However, under similar reaction conditions and time, the LdG-modified template showed only 9% incorporation and 5% extension of the primer (lane 7, Figure 2C). This indicates that neither significant incorporation nor extension was achieved by the polymerase across LdG when compared to dG. It should be noted that by increasing the enzyme concentration to 2 units, 52% of incorporation and 14% of extension was observed across LdG after 4 h (lane 7, Figure 2D). These results clearly show that compared to correct base incorporation across LdG, furthering the extension is significantly slower. This may be attributed to the structural distortions at the modified site as well as the steric clashes caused by the bulky adduct with KF⁻.[30] These findings are supported by our molecular modeling studies (see the respective section for details).

To study the miscoding potential opposite LdG, single nucleotide misincorporation experiments were performed using 2 units of KF⁻ in the presence of dATP, dGTP, and dTTP (Figure 3, panels A–C). In the case of dATP, 84% of +2 and 16% of +3 extended products were observed after 4 h (lane 7, Figure 3A). This shows that dATP incorporation was quite rapid, and the incorporated product immediately gets extended due to the presence of dT at a +2 site. The dGTP was incorporated up to 60% along with the formation of 25% of +2 and 7% of +3 extended products (lane 7, Figure 3B). Similarly, dTTP

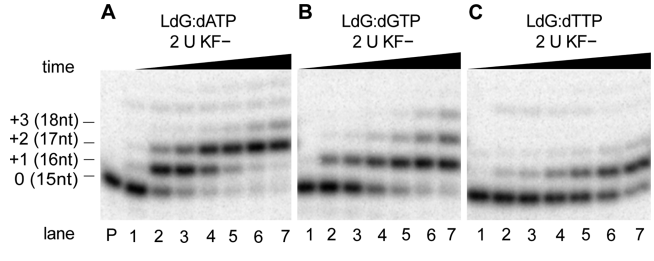


Figure 3. PAGE (20%, 7 M urea) of single nucleotide misincorporation using KF⁻. The sequences of the primer and the LdG-modified template are same as in Figure 2A. (A) dATP; (B) dGTP; (C) dTTP. Lane P, primer; lanes 1 to 7, reactions for different time course from 0.5, 5, 10, 30, 60, 120, and 240 min. All the reactions were carried out at 37 °C using 2 U of KF⁻ in the polymerase buffer (50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9).

DOI: 10.1021/jo502627b
J. Org. Chem. 2015, 80, 2128–2138

incorporation across LdG was found to be 67% along with the formation of 8% of a +2 extended product (lane 7, Figure 3C). Overall, these results show that the efficiency of KF$^-$ for dATP misincorporation is higher than other dNTPs, and the rate of incorporation was found to be in the following order: dATP > dGTP > dTTP > dCTP. These results are in line with the reported findings for the widely studied BP adduct and others, where the KF$^-$ showed preference for dATP incorporation over other nucleotides.[31−35] The reason for the better tolerance of dA across the adduct site has been attributed to the efficient stacking of purine bases with the polycyclic aromatic DNA adducts. This facilitates the formation of right conformation required for the base incorporation.[32]

To find the ability of KF$^-$ to polymerize the full length product, primer extensions of the LdG-modified and the unmodified DNA templates were carried out in the presence of a mixture of all four dNTPs (Figure 4, panels A−C). With the use

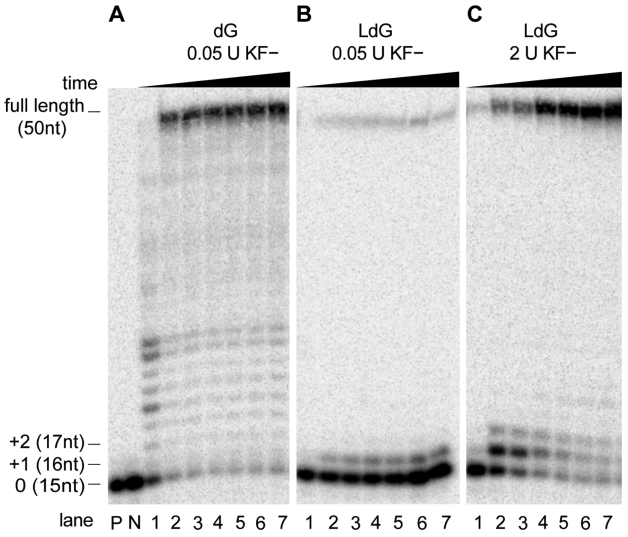


Figure 4. PAGE (20%, 7 M urea) of the full length extension reactions using KF$^-$ with all the dNTPs. The sequences of the primer and the template are the same as in Figure 2A. (A) Unmodified template (dG) with 0.05 U of KF$^-$; (B) modified template (LdG) with 0.05 U of KF$^-$; (C) modified template (LdG) with 2 U of KF$^-$. Lane P, primer; lane N, negative reaction control (no dNTPs); lanes 1 to 7, primer extension reactions with 100 μM of each dNTPs in different time course from 0.5, 5, 10, 30, 60, 120, and 240 min. All the reactions were carried out at 37 °C in the polymerase buffer (50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9).

of 0.05 units of the KF$^-$, the unmodified template yielded 86% of the full length product after 4 h (lane 7, Figure 4A). However, under the identical conditions and time, the LdG-modified template showed only 2% of the full length product (lane 7, Figure 4B). Nevertheless, as expected, at a higher unit of enzyme (2 U), 84% of the fully extended primer was observed in the presence of the LdG-modified template (lane 7, Figure 4C). This observation is in contrast with the behavior of N^2-dG adducts formed by BP and AF toward KF$^-$. In the case of BP, the enzyme is unable to extend the primer after the adduct site, while for AF the extension beyond the +3 site is strongly blocked.[35,36] To investigate the efficiency of full length extension reactions, additional experiments were carried out using the 11-mer primer and the LdG-modified template, in which the modification is five nucleotides away from the initiating point (Figure S3A of the Supporting Information). When 0.05 units of the KF$^-$ was used, after 4 h, the unmodified template yielded 96% of the full length product (lane 7, Figure S3B of the Supporting Information). While, under similar conditions, there was no significant product formation with the LdG-modified template, which clearly shows that DNA synthesis has been halted at the modification site (lane 7, Figure S3C of the Supporting Information). However, using a higher unit of enzyme (2 U KF$^-$), the 92% of the fully extended product was observed (lane 7, Figure S3D, of the Supporting Information). These results clearly demonstrate that the incorporation and extension efficiency of the LdG-modified templates by the KF$^-$ is independent of the position of the lesion site. Overall, the primer extension studies reveal that KF$^-$ is able to perform TLS across LdG DNA adduct albeit at low fidelity and processivity.

Molecular Modeling Studies. Our first goal was to understand the possible structural effects of the LdG modification inside a DNA duplex. Toward this end, the energy optimized and RESP charge-fitted LdG-modified nucleotide using Gaussian 09[37] (Figure S4 of the Supporting Information) was placed in a DNA duplex which is used for the thermal melting studies (Table 1, **D3-D2**).[38,39] MD simulations of the LdG-modified DNA duplex was carried out for 50 ns using AMBER 12.[40,41] The system reached the equilibrium after 30 ns as deduced from the RMSD graph (Figure S5 of the Supporting Information); thus, all the analyses were carried out from the 30−50 ns of MD trajectories. The averaged structure from the last 20 ns of the MD simulations is shown in Figure 5A, where the LdG adduct is positioned toward the 3′-end in the minor groove of the duplex DNA. The MD structure shows that the B-form of the LdG-modified DNA is maintained throughout the simulation. This is supported by CD spectra of the LdG-modified DNA duplex, where a positive peak is observed at 280 nm and a negative peak at 250 nm (Figure S6 of the Supporting Information).[42]

The distance between the 3′-dA nucleotide and the LdG adduct was found to be ∼4.2 ± 0.5 Å (Figure 5B). It should be noted that during the course of the simulations, the two consequent dA nucleotides at the 3′-end (Figure 5, panels B and D) moved away from the original position to avoid the steric clashes with the LdG adduct. This movement could disturb the W−C hydrogen bonds between the pairing bases. This is reflected in the H-bond occupancy, where only one W−C H-bond between the LdG and dC was found to be stable and the other two H-bonds were absent (Figure 5C and Table S1 of the Supporting Information). Moreover, the base pairing of the neighboring nucleotides around the LdG adduct was also disturbed (Table S1 of the Supporting Information). The percentage of stacking interactions between the LdG and dC with the neighboring dA and dT were found to be ∼13% of the total simulation time (Figure 5D). Moreover, the hydration around the LdG adduct was severely affected both in the minor and major grooves of the DNA duplex (Table S2 of the Supporting Information). The average of the six sugar−phosphate backbone dihedral angles (α, β, γ, ε, δ, and ζ) were found to be −98.3° ± 12, −178.5° ± 14, 61.1° ± 22, 71.1° ± 18, −163.2° ± 31, and −73.1° ± 36, respectively, with maximum deviation of up to 32° with respect to those of the canonical B-form DNA.[43] In particular, higher deviations were observed for α, δ, and ζ torsions. The χ angle of the glycosidic bond of LdG base was found to be −164.2° ± 16, which is similar to the canonical B-DNA value (−150 to −180°).[43] Overall, the MD simulations of the LdG-modified duplex suggest that due to the position of the bulky adduct in the minor groove, the stacking interactions,

DOI: 10.1021/jo502627b
J. Org. Chem. 2015, 80, 2128−2138

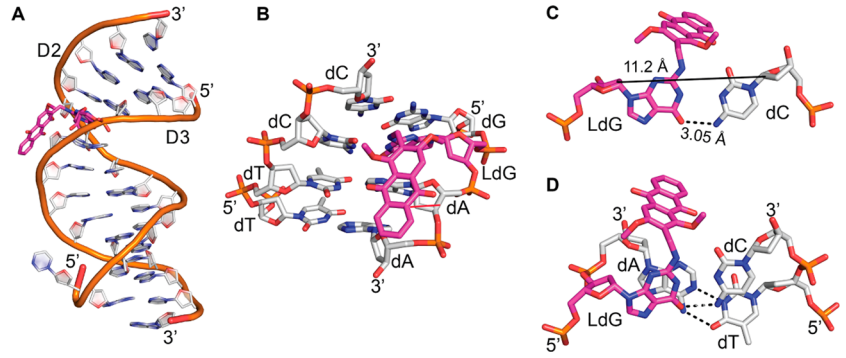


Figure 5. Averaged MD snapshot of the LdG (highlighted in magenta) modified DNA duplex from the last 20 ns of 50 ns simulations. (A) The 14-mer DNA duplex structure is shown in cartoon, and the LdG adduct is represented in sticks. (B) The architecture of DNA duplex around the LdG adduct along with the neighboring nucleotides. (C) W−C H-bonds between LdG and dC. (D) The stacking interactions between the two successive nucleotide pairs in the duplex. The red line in B represents the distance (4.2 Å) between the adduct and the 3′-dA. Black-dotted lines in C and D represent the H-bond between the bases. The C1′−C1′ distance between the pairing nucleotide is represented as a black line in C.

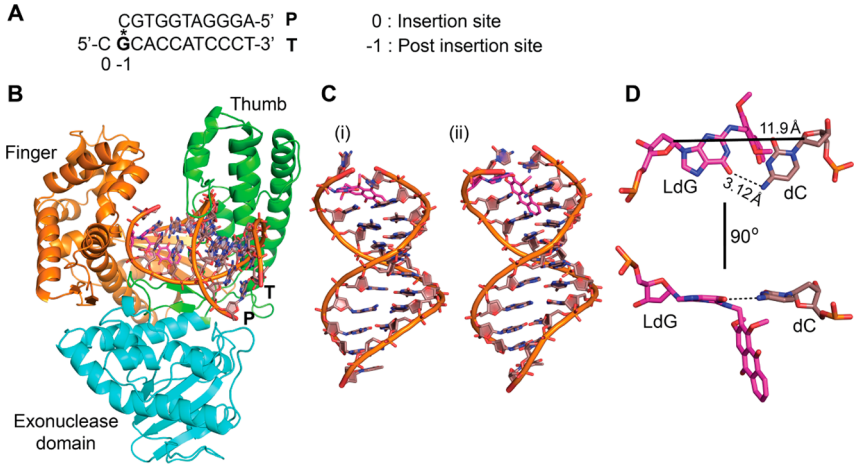


Figure 6. MD snapshots of the template containing LdG adduct-primer complexed with *Bacillus* fragment (BF). (A) The complete sequence of the template, T (**G*** = LdG) and the primer, P. The LdG adduct is present at the post insertion site (−1). (B) The final MD snapshot of the binary complex of the BF and the LdG (highlighted in magenta)-modified DNA. The polymerase and the DNA are represented in cartoon and sticks, respectively. (C) MD simulation snapshots of the two major conformations of T:P duplex DNA obtained from the ensemble analysis. (D) Averaged conformations from 6C of the LdG adduct in T-DNA base paired with the dC of P-DNA. The C1′−C1′ distance between the pairing nucleotide is represented by a black line, and the H-bonding is shown with black dotted lines.

W−C H-bond between bases, and the groove hydration were affected in and around the modified site.

To delineate the possible molecular mechanism of bypass of the LdG adduct by the KF$^-$, a binary complex composed of a DNA duplex containing the LdG in the template and *Bacillus* fragment (BF; PDB entry: 1XC9)[44] was investigated using MD simulations. The rationale to use BF structure is that, it is a homologue of KF$^-$; additionally, the structure contains the BPdG adduct in the template DNA. Initially, the optimized and RESP charge-fitted LdG-modified nucleotide (Figure S4 of the Supporting Information) was placed at the post insertion site (Figure 6A), where the BPdG adduct is present in the BF complex. The modified binary complex was solvated, minimized, and then the unrestrained MD simulations were carried out for 40 ns using AMBER 12.[40]

The final MD snapshot of the binary complex is shown in Figure 6B. The structure shows that the B-form of the template-primer duplex was maintained throughout the simulations. The LdG adduct was positioned in the minor groove of the DNA, in which the adduct makes van der Waals contact with the bases at the 5′-end of the template DNA. The orientation of the LdG adduct in the complex was found to be similar to that of the BPdG adduct.[44] Here, it should be noted that the adducts placed in the major groove are well-tolerated by the polymerases in comparison to the adduct placed in the minor groove of the DNA, which makes key contacts with the polymerase active site.[45]

In the binary complex, the LdG adduct was found to adopt a number of conformations (Figure 6C). This is reflected in the average RMSD (2.5 ± 1.4 Å) of DNA during the course of the simulations (Figure S7 of the Supporting Information). Moreover, the LdG adopts anti ($\chi = -124°$) conformation with higher deviation of up to 20° from the canonical conformation; and the C2′-*endo* conformation was found to be the most prominent sugar pucker. The RMSD analysis revealed that there are two major conformational ensembles for the LdG−DNA duplex in the binary complex: (i) the perpendicularly positioned adduct in the DNA minor groove having one W−C base pair along with dC as seen in Figure 6 (panels C and D) and (ii) the conformer having widened groove size for the duplex and having weak T-shaped van der Waals contact between LdG with the dC and dA bases at the 5′-end of the template strand (Figure 6C and Figure

DOI: 10.1021/jo502627b
J. Org. Chem. 2015, 80, 2128−2138

S8 of the Supporting Information). The representative structures of the two conformational ensembles are shown in Figure 6C. These two major conformational ensembles were preferred for >82% of the total simulation time. The averaged structure of these two conformers having one W−C H-bond between the O^6 of LdG and the hydrogen of N^4 of dC is shown in Figure 6D. The other two W−C H-bonding sites of dC were stabilized by Tyr 714 of BF polymerase (Figure 7).

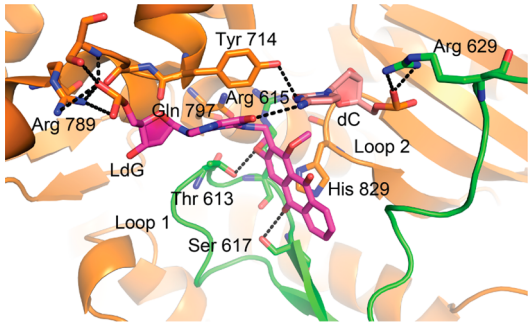


Figure 7. Active site orientation of the LdG adduct and pairing base dC in complex with the *Bacillus* fragment (BF). Polymerase is colored by domains (thumb in green, and finger in orange). The LdG adduct is shown in magenta, the hydrogen bonds between the atoms are represented in black dotted lines. The protein is rendered in cartoon, the atoms shown as sticks, and DNA as sticks along with filled rings. Key residues are labeled with numbers. Tyr 714 makes an H-bond with dC at a distance of 2.75 Å. Arg 629 and Arg 789 form an H-bond with LdG and dC, respectively, at a distance of ∼2.73 Å. Ser 671 forms an H-bond with the carbonyl oxygen atom in the LdG at a distance of 2.63 Å. His 829 stacks on the polycyclic aromatic ring of LdG with a 3.0 Å distance. The W−C H-bond distance between LdG and dC is 3.1 Å. The conformer shown here is the averaged structure depicted in Figure 6C.

In the binary complex, the LdG appeared to be intact and its motions were stabilized by two different loops in the active site of the BF. Loops from the amino acids 607−618 and 828−831 constitute 16 residues, which form various noncovalent interactions with the adduct and they are highly flexible during the simulations as evident from their RMSD values (∼5.2 ± 2.3 Å). Notably, the amino acid His-829 from the loop 2 stacks on the LdG adduct, which partly restricts the movement of the polycyclic aromatic ring during the course of simulation (Figure 7). In addition, Arg residues at 629 and 789 form hydrogen bonds with the phosphate backbone of both template and primer at the adducted site. Also, the oxygen atom in the carbonyl group of the LdG adduct forms a hydrogen bond with the Ser-617 hydroxyl group. The Arg-615 and Gln-797 in BF polymerase are known to make H-bond contact with the benzo[*a*]pyrene adduct at the post insertion site.[34] To probe the interaction of these two amino acids with the LdG adduct, the distance between them has been computed. Those distances were found to be 3.5 to 4.0 Å (Figure 7), which corresponds to only weak van der Waal contacts. It has been reported that the efficiency of dNTP incorporation by KF^- is significantly lowered when the C1(dG)′-C1′(dC) distance is increased >12.5 Å.[46,47] In the case of the LdG, the average C1′−C1′ distance was found to be 11.9 ± 2.9 Å. Due to this, the KF^- is not able to well-accommodate the LdG adduct, which is indeed reflected in the primer extension experiments.

To find the structural flexibility in BF induced by the LdG adduct, the BF crystal structure in complex with unmodified DNA (PDB entry: 1L3S) was superimposed on the averaged MD structure of BF in complex with LdG-modified DNA (Figure 8, and Figure S9 of the Supporting Information). It is clear from this

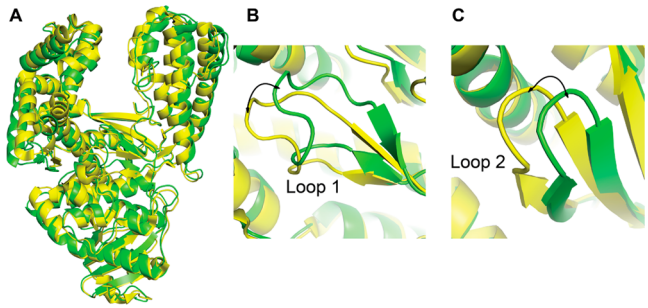


Figure 8. Superimposed structure of *Bacillus* fragment (BF) in complex with the LdG-modified DNA (yellow) and unmodified DNA (green). (A) Superimposed structure of the BF is shown in cartoon. The overall RMSD between the two structures is 2.6 Å. (B) Loop 1 (607−618 aa) of the BF in complex with the modified and unmodified T-DNA. Loop 1 moves away from its original site (5.2 Å) due to the presence of the adduct, as indicated by a black double-headed arrow. (C) Loop 2 (828−831 aa) of BF complexed with modified and unmodified template T-DNA. Loop 2 also moves away from its original position (4.1 Å).

analysis that the loops 1 and 2 of BF move away from its original position by ∼4.5 Å when complexed with the LdG adduct. Thus, the polymerase requires the larger conformational rearrangements to accommodate the LdG. It results in the structural distortions, which create a high-energy barrier to bypass the adduct site. Additionally, to find the distortions in the DNA structure during the course of MD simulations, the backbone torsions and helical parameters of the LdG-modified duplex were calculated and the values were compared with those from the unmodified duplex (PDB entry: 1L3S). The results show that the base pair helical parameters like buckle, propeller, twist, and tilt are significantly deviated around the LdG adduct site (Tables S3−S4 of the Supporting Information). Backbone torsion angles such as α, β, γ, and δ were similar to those of the canonical values; however, ε and ζ were highly deviated (Table S5 of the Supporting Information).

Overall, the results from the molecular modeling and dynamics studies suggest that the LdG occupies the minor groove of the DNA duplex and imparts conformational flexibility due to which lengthening of the C1′(LdG)-C1′(dC) distance was observed at the template−primer interface. Therefore, to facilitate the primer extension, the polymerase undergoes structural rearrangements; also the H-bonding and van der Waals interactions with the amino acid residues at the active site stabilize the LdG:dC base pair.

CONCLUSIONS

In summary, we have achieved the synthesis of the LdG-modified DNAs. Our synthetic strategy involves Buchwald-Hartwig coupling as a key step to access the LdG phosphoramidite building block, which is utilized for the solid phase DNA synthesis. Thermal melting studies revealed that the LdG modification decreases the stability of the DNA duplex by ∼5 °C. The disturbances in W−C H-bonding and base stacking lead to the destabilization of LdG-modified DNA duplexes, which were reflected in the unfavorable contribution of enthalpy. Primer extension studies showed that the KF^- was able to bypass the LdG in an error prone fashion. Adduct bypass by the KF^- has shown more preference for the incorporation of purines in

DOI: 10.1021/jo502627b
J. Org. Chem. 2015, 80, 2128−2138

comparison to pyrimidines. The molecular modeling and MD simulation studies revealed the atomistic interactions of the enzyme and LdG, which helped to unravel the possible molecular mechanism for the accommodation and nucleotide bypass across the adduct. It was evident from these studies that the increase in the C1′−C1′ distance and loss of two W−C H-bonding between LdG and dC, disturbances in the enzyme active site orientations, together may be responsible for the inefficient incorporation of the incoming nucleotide by KF^-. Overall, tools and insights emerging from these studies would pave the way to investigate the structural and functional requirements of possible TLS pathways across lucidin-modified DNAs by both the eukaryotic and prokaryotic Y-family polymerases.

EXPERIMENTAL SECTION

General. All the required chemicals and solvents were obtained from the commercial sources. Dioxane, pyridine, acetonitrile, DIPEA, DCM, and Et_3N were dried using calcium hydride; and the toluene was dried using $CaCl_2$. Thin layer chromatography (TLC) was carried out on the silica gel plates, which were precoated with the fluorescent indicator, visualized by UV light or by dipping into a solution of 5% (v/v) conc. H_2SO_4 in ethanol, and heating. Silica gel of 100−200 mesh size was used for column chromatography. ^{1}H NMR (400 and 500 MHz), ^{13}C NMR (100 MHz), and ^{31}P NMR (162 MHz) were recorded on a 400 or 500 MHz NMR instrument. The chemical shifts are reported in parts per million (ppm) downfield from TMS and referenced to the TMS signal or residual proton signal of the deuterated solvent as follows: TMS (0 ppm) or CD_3OD (3.31 ppm) for the ^{1}H NMR spectra, and $CDCl_3$ (77.2 ppm) or CD_3OD (49.1 ppm) for the ^{13}C NMR spectra. Multiplicities of ^{1}H NMR spin couplings are reported as s for singlet, bs for broad singlet, d for doublet, dt for doublet of triplets, dd for doublet of doublets, ddd for doublet of doublet of doublets, or m for multiplet and overlapping spin systems. Values for the apparent coupling constants (J) are reported in Hz. High-resolution mass spectra (HRMS) were obtained in a positive ion electrospray ionization (ESI) mode using the Q-TOF analyzer. DNA sequences were synthesized using an automated DNA synthesizer. Mass spectra of the oligonucleotides were obtained by positive reflectron mode in a MALDI-TOF spectrometer.

*2-(Azidomethyl)-1,3-dimethoxyanthraquinone (**2**).* 2-(Bromomethyl)-1,3-dimethoxyantharquinone (987 mg, 2.73 mmol) was mixed with sodium azide (284 mg, 4.37 mmol) in 13 mL of dry DMF and stirred for 30 h at room temperature. After completion of the reaction, product was extracted using ethyl acetate (2 × 50 mL), and the organic layer was washed with water (2 × 30 mL), dried over Na_2SO_4, and filtered. Solvent was removed under reduced pressure to afford compound **2** as a yellow solid (413 mg, 46%); R_f = 0.57 (5% ethyl acetate in petroleum ether); mp 140−142 °C; ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.27 (dd, J = 0.9, 7.7 Hz, 1H), 8.23 (dd, J = 1.3, 7.5 Hz, 1H), 7.82−7.71 (m, 2H), 7.70 (s, 1H), 4.52 (s, 2H), 4.07 (s, 3H), 3.98 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 183, 181.1, 163, 161.5, 137.2, 134.9, 134.6, 133.5, 132.5, 127.4, 126.9, 125.2, 120.1, 105.4, 63.3, 56.6, 43.2; HRMS (ESI): calcd for $C_{17}H_{13}N_3NaO_4$, $[M + Na]^+$ 346.0798; found, $[M + Na]^+$ 346.0801 (Δm = +0.0003, error = +0.8 ppm).

*2-(Aminomethyl)-1,3-dimethoxyanthraquinone (**3**).* 2-(Azidomethyl)-1,3-dimethoxyanthraquinone (413 mg, 1.27 mmol) was dissolved in 3 mL of THF, followed by the addition of triphenylphosphine (502 mg, 1.91 mmol) at 0 °C. The reaction was removed from the ice bath, and water (0.06 mL) was added. The reaction mixture was stirred at room temperature for 3 h. After the completion of the reaction, the solvent was evaporated under reduced pressure. The crude compound was purified by column chromatography in basic alumina (7% methanol in DCM) to obtain compound **3** as a brown colored solid (284 mg, 78%); R_f = 0.44 (10% methanol in DCM); mp: 181−183 °C; ^{1}H NMR (400 MHz, CD_3OD): δ 8.15 (dd, J = 1.1, 7.7 Hz, 1H), 8.10 (dd, J = 1.1, 7.7 Hz, 1H), 7.85−7.72 (m, 2H), 7.57 (s, 1H), 4.02 (s, 3H), 4.01 (s, 2H), 3.91 (s, 3H); ^{13}C NMR (100 MHz, CD_3OD): δ 183.7, 182.4, 164.2, 161.9, 138.1, 136, 135.8, 134.8, 133.7, 128.5, 128.1, 127.7, 120.9, 106.3, 63.2, 57.2, 34.4; HRMS (ESI): calcd for $C_{17}H_{15}NO_4$, $[M + H]^+$ 298.1074; found, $[M + H]^+$ 298.1073 (Δm −0.0001, error −0.4 ppm).

*2-Bromo-O^6-(2-(4-nitrophenyl)ethyl)-3′,5′-diacetyl-2′-deoxyguanosine (**5**).* The amine compound **4** (108 mg, 0.21 mmol) was dried by coevaporation using anhydrous toluene (12 mL) under reduced pressure. Antimony bromide (111 mg, 0.30 mmol) was added to the flask containing dried amine, and the system was purged with N_2 gas. Anhydrous CH_2Br_2 (2.5 mL) was cooled to 0 °C and added to the reaction mixture with continuous stirring. Then the reaction mixture was cooled to −10 °C, and TBN (0.10 mL, 0.73 mmol) was added slowly. The reaction mixture was kept stirring for around 2 h at the same temperature followed by stirring at 0 °C for an additional 2 h. After the completion of the reaction, the reaction mixture was poured into the saturated solution of $NaHCO_3$ (50 mL) in crushed ice and extracted with DCM (3 × 30 mL). The combined organic layers are dried over Na_2SO_4, filtered, and concentrated under reduced pressure. The crude compound was purified by column chromatography (55% ethyl acetate in petroleum ether) to give compound **5** as a pale yellow solid (98 mg, 82%); R_f = 0.55 (70% ethyl acetate in petroleum ether); mp: 134−136 °C; ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.21−8.13 (m, 2H), 8.08 (s, 1H), 7.49 (d, J = 8.5 Hz, 2H), 6.43 (dd, J = 6.0, 8.0 Hz, 1H), 5.39 (td, J = 2.4, 6.0 Hz, 1H), 4.84 (t, J = 6.7 Hz, 2H), 4.40−4.31 (m, 3H), 3.31 (t, J = 6.7 Hz, 2H), 2.81 (ddd, J = 6.4, 8.0, 14.2 Hz, 1H), 2.65 (ddd, J = 2.5, 5.9, 14.2 Hz, 1H), 2.14 (s, 3H), 2.10 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 170.5, 170.4, 160.3, 152.8, 147, 145.4, 143.6, 140.7, 130.1, 123.9, 121.3, 84.8, 82.8, 74.4, 67.7, 63.7, 38.1, 35.1, 21, 20.9; HRMS (ESI): calcd for $C_{22}H_{22}BrN_5O_8$, $[M + H]^+$ 564.0730; found, $[M + H]^+$ 564.0734 (Δm = +0.0004; error = +0.8 ppm).

*N^2-Methyl-(1,3-dimethoxyanthraquinone)-O^6-(2-(4-nitrophenyl)-ethyl)-2′-deoxyguanosine (**7**).* To a screw-capped vial, $Pd(OAc)_2$ (18 mg, 0.07 mmol) and (*R*)-BINAP (142 mg, 0.22 mmol) were added and flushed with N_2 gas. Then dry toluene (8 mL) was added, and the mixture was stirred at room temperature for 5 min. Cs_2CO_3 (375 mg, 1.06 mmol) was added, followed by the addition of 2-(aminomethyl)-1,3-dimethoxyanthraquinone (250 mg, 0.83 mmol) and the bromo nucleoside **5** (430 mg, 0.76 mmol). The vial was purged with N_2 gas, sealed with a Teflon-lined cap, and heated in an oil bath at 85 °C for 10 h. After the completion of the reaction, the reaction mixture was passed through a Celite pad and washed with ethyl acetate (250 mL). The filtrate was concentrated to give a dark brown crude compound **6**. The crude compound **6** (404 mg, 0.51 mmol) was dissolved in 11 mL of methylamine in ethanol (33%, v/v) and was stirred at room temperature for 2 h. After the completion of the reaction, the mixture was concentrated under reduced pressure. The oily residue was purified by silica gel column chromatography (3% methanol in DCM) to obtain compound **7** as a brown colored solid (190 mg, 39% after two steps); R_f = 0.49 (10% methanol in DCM); mp: 127−129 °C; ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.26 (d, J = 7.6 Hz, 1H), 8.22 (d, J = 7.3 Hz, 1H), 8.15 (d, J = 8.2 Hz, 2H), 7.82−7.71 (m, 2H), 7.66 (s, 1H), 7.61 (bs, 1H), 7.51 (d, J = 7.9 Hz, 2H), 6.25−6.17 (m, 1H), 5.55 (bs, 1H), 4.93−4.71 (m, 5H), 4.15 (s, 1H), 4.05 (s, 3H), 3.92 (bs, 1H), 3.89 (s, 3H), 3.75 (d, J = 12.5 Hz, 1H), 3.32−3.27 (m, 2H), 3.04 (bs, 1H), 2.26 (d, J = 8.2 Hz, 1H);^{13}C NMR (100 MHz, $CDCl_3$): δ 183, 181.4, 163.1, 161, 160.9, 158.3, 152.9, 146.9, 146.1, 139, 136.4, 134.8, 134.6, 133.5, 132.4, 130.1, 130, 128.2, 127.3, 126.8, 123.8, 120, 116.4, 105.5, 89, 87.1, 73.2, 66.3, 63.5, 62.8, 56.6, 40.2, 35.3, 35.3; HRMS (ESI): calcd for $C_{35}H_{32}N_6NaO_{10}$, $[M + Na]^+$ 719.2072; found, $[M + Na]^+$ 719.2079 (Δm = +0.0007 and error = +0.9 ppm).

*5′-O-(4,4′-dimethoxytrityl)-N^2-methyl-(1,3-dimethoxyanthraquinone)-O^6-(2-(4-nitrophenyl)ethyl)-2′-deoxyguanosine (**8**).* The compound **7** (140 mg, 0.20 mmol) was coevaporated with dry pyridine and dissolved in the same solvent (2 mL). To this, DMT-Cl (135 mg, 0.40 mmol) was added and stirred for 10 h at room temperature. The reaction mixture was diluted with DCM (100 mL) and washed with saturated $NaHCO_3$ (40 mL) and water (2 × 50 mL). The DCM layer dried over Na_2SO_4 and evaporated. The crude compound **8** was purified by silica gel column chromatography (1% methanol in DCM + 2% Et_3N) to afford compound **8** as a dark brown solid. (187 mg, 93%); R_f = 0.55 (2% methanol in DCM + 2% Et_3N); mp 79−81 °C; ^{1}H NMR (400 MHz, $CDCl_3$) δ: 8.26 (dd, J = 1.0, 7.5 Hz, 1H), 8.22 (dd, J = 1.1, 7.7 Hz,

DOI: 10.1021/jo502627b
J. Org. Chem. 2015, 80, 2128−2138

1H), 8.12 (d, J = 8.5 Hz, 2H), 7.81−7.72 (m, 2H), 7.71 (s, 1H), 7.65 (s, 1H), 7.46 (d, J = 8.5 Hz, 2H), 7.37 (d, J = 7.3 Hz, 2H), 7.29−7.12 (m, 7H), 6.75 (d, J = 8.8 Hz, 4H), 6.39 (t, J = 6.1 Hz, 1H), 5.52 (t, J = 6.3 Hz, 1H), 4.84−4.63 (m, 5H), 4.15 (d, J = 3.5 Hz, 1H), 4.03 (s, 3H), 3.90 (s, 3H), 3.74 (s, 6H), 3.36 (dd, J = 4.5, 10 Hz, 1H), 3.30 (dd, J = 4.6, 10.2 Hz, 1H), 3.25 (t, J = 6.7 Hz, 2H), 2.70 (td, J = 6.6, 13.2 Hz, 1H), 2.54−2.45 (m, 1H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 183, 181.6, 163.3, 161.1, 160.6, 158.6, 158.6, 153.8, 146.9, 146.2, 144.7, 137.3, 136.3, 135.8, 135.8, 134.8, 134.5, 133.5, 132.5, 130.1, 130, 128.7, 128.2, 127.9, 127.3, 126.9, 126.8, 123.8, 119.8, 115.3, 113.2, 105.5, 86.5, 85.9, 83.9, 72.7, 66.1, 64.1, 62.8, 56.6, 55.3, 40.4, 35.3, 35.1; HRMS (ESI): calcd for $C_{56}H_{50}N_6O_{12}$, $[M + H]^+$ 999.3559; found, $[M + H]^+$ 999.3563 (Δm = +0.0004 and error = +0.3 ppm).

*3′-O-[2-Cyanoethoxy-(N,N′-diisopropylamino)-phosphino]-5′-O-(4,4′-dimethoxytrityl)-N^2-methyl-(1,3-dimethoxyanthraquinone)-O^6-(2-(4-nitrophenyl)ethyl)-2′-deoxyguanosine (**9**).* The tritylated nucleoside **8** (150 mg, 0.15 mmol) was dissolved in DCM (1.5 mL) followed by addition of DIPEA (0.2 mL, 1.20 mmol) and CEP-Cl (71 mg, 0.30 mmol) and stirred at room temperature for 1.5 h. Then, methanol (0.1 mL) was added and stirred for 15 min. The reaction mixture was diluted with DCM (100 mL) and washed with $NaHCO_3$ (3 × 20 mL), dried over Na_2SO_4, and evaporated under reduced pressure. The crude compound was purified by silica gel column chromatography (30% DCM in petroleum ether + 2% Et_3N) to obtain the compound **9** as a brown solid (118 mg, 65%); R_f = 0.56 (75% DCM in petroleum ether + 2% Et_3N); mp: 90−92 °C; ^{31}P NMR (162 MHz, $CDCl_3$): δ 148.73, 148.69; HRMS (ESI): calcd for $C_{65}H_{67}N_8O_{13}P$, $[M + H]^+$ 1199.4638; found, $[M + H]^+$ 1199.4637 (Δm = −0.0001 and error = −0.1 ppm).

Oligonucleotide Synthesis. The unmodified and LdG-modified DNA sequences were synthesized on a 1 μmol scale using controlled pore glass (CPG) solid support. The LdG modification has been introduced into DNA having different lengths (**D3**, **D4**, and **T**, Table S6 of the Supporting Information) using a solid phase synthesizer. The coupling time used for the unmodified phosphoramidites was 2 min and for the LdG-modified phosphoramidite was 6 min. The 5-ethylthio-1H-tetrazole (ETT) was employed as the coupling agent. The deprotection of the LdG-modified oligos was carried out in four different steps. The first step involved the selective deprotection of cyanoethyl group using 10% diethylamine in acetonitrile (ACN) at room temperature for 10 min.[26] The second step involved the removal of the NPE group with 1 M 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) in ACN (1 mL) for 1 h at room temperature.[48] Then, for cleavage of the CPG support, oligonucleotides were treated with 30% aq NH_3 at room temperature for 3 h, followed by treatment with 30% aq NH_3 at 55 °C for 16 h to remove the base-protecting groups.[49] The supernatant layer was collected separately, and the CPG beads were washed with water (1 × 300 μL). The combined aqueous layer was evaporated on speedvac to get the crude DNA, which was further purified using 20% denaturing PAGE (7 M urea) at 30 W for 3 h with 1× TBE running buffer (89 mM each Tris and boric acid and 2 mM EDTA, pH 8.3). The gel thickness was 1 mm, and the gel dimension was 20 × 30 mm. The gel was visualized under a UV lamp (260 nm), and desired DNA bands were marked. The gel bands were crushed, and 15 mL of TEN (10 mM Tris, 1 mM EDTA, 300 mM NaCl, pH 8.0). It was further shaken for 16 h at room temperature for the recovery of DNA from the gel. Finally, desalting of oligonucleotides was carried out using a C18 Sep-Pak column. Desalted DNAs were dissolved in water, the absorbances were measured at 260 nm, and the concentrations were calculated using appropriate molar extinction coefficients. All the oligonucleotides were characterized using MALDI-TOF in positive reflectron mode, and the molecular weights are provided in Table S6 of the Supporting Information.

Thermal Melting Studies. UV-melting studies were carried out at 260 nm in the range of 20−90 °C with a ramp of 0.5 °C min^{-1} using three heating−cooling cycles. Equal concentration (5 μM) of unmodified and the LdG-modified complementary DNAs were dissolved in a phosphate buffer solution: 100 mM NaCl, 20 mM Na_2PO_4, 0.1 mM EDTA, pH 7.4. The samples of the DNA duplexes were annealed at 95 °C for 3 min, cooled to room temperature over a period of 3 h, and stored at 4 °C. Samples were equilibrated at 25 °C for 10 min before starting the experiments. All the experiments were triplicated, and the T_m values reported are the average of 3 independent measurements. The absorbance versus temperature data was analyzed via a Marquadt algorithm for nonlinear curve fitting using KaleidaGraph (version 3.5) to decipher T_m and $\Delta H°$ values.[50] The $\Delta S°$ and $\Delta G°$ values were calculated using following equations: $\Delta S° = \Delta H°/T_m$; $\Delta G° = \Delta H° (1-T/T_m)$. All the thermodynamic parameters were derived from the analysis of three independent melting curves, and the average values are reported in Table 1. The error propagations for $\Delta S°$ and $\Delta G°$ were computed using the methods reported by Bevilacqua and co-workers.[27]

Circular Dichroism Experiments. CD experiments were performed at wavelength in the range of 220−500 nm using a 1 mm path length quartz cuvette. The duplex DNAs (10 μM each strand) were annealed by heating at 95 °C for 5 min in phosphate buffer solution (100 mM NaCl, 20 mM Na_2PO_4, 0.1 mM EDTA, pH 7.4), then cooled to room temperature over a period of 3 h, and stored at 4 °C. The experiments were carried out at 25 °C, scanned at 100 nm/min using a response time of 2 s. Each spectrum was an average of 3 measurements; baseline corrected and analyzed using Origin (version 8.0).

5′-Radiolabeling of Primer DNA. The primer DNA (25 pmol) was mixed with T4 polynucleotide kinase (PNK) enzyme (5 U) and [γ-^{32}P] ATP (20 μCi) in 1× PNK buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, 5 mM DTT, 0.1 mM spermidine, 0.1 mM EDTA, pH 7.6) in a total volume of 10 μL. The reaction mixture was incubated at 37 °C for 1 h followed by deactivation of the PNK enzyme by heating at 70 °C for 3 min. The radiolabeled primer was purified using a QIAquick nucleotide removal kit.

Primer Extension Studies. The primers were mixed with the unmodified and modified template in a molar ratio of 1:1.5 and annealed by heating to 95 °C for 3 min, followed by slow cooling to room temperature for 30 min. After annealing, the primer extension reactions were performed with a primer−template duplex in a total volume of 10 μL, containing 50 nM primer, traces of 5′-labeled primer, and 75 nM template, 100 μM concentration of each dNTPs in 1× polymerase buffer (50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9). The primer extension reactions were initiated by incubation of the reaction mixture at 37 °C for 5 min, followed by the addition of a desired amount of KF^-. The incorporation and misincorporation reactions were carried out using dCTP or dGTP or dATP or dTTP, and the extension was carried out using all dNTPs. The reactions were carried out at 37 °C for the indicated amount of time; aliquots were taken out of the reaction mixture and quenched by the addition of stop solution (80% formamide, 0.025% each bromophenol blue, xylene cyanol, and 50 mM EDTA). The extended products were run on a 20% denaturing PAGE (7 M urea), and an autoradiogram was generated and the bands were quantified by ImageQuantTL software. The percentage of each band was calculated by considering all the observed bands in the PAGE of the reaction using the following formula.

$$\text{percentage incorporation/extension} = \frac{I_{s_1}}{(I_p + I_{s_1} + I_{s_2})} \times 100$$

I_p = intensity of primer band (nonshifted band), I_{s_1} = intensity of shifted band 1, and I_{s_2} = intensity of shifted band 2 and so on. The same formula has been applied for calculating the percentage of each band in the same lane.

Molecular Modeling Studies. The force field parameters and RESP charges were developed for the LdG adduct using previously reported procedures.[38,39] The adduct was placed in the identical ds DNA sequence, which was used in the thermal stability studies (Table 1, **D3-D2**). The B-form structure of the DNA was built using nucleic acid builder module in AMBER 12. The structure was initially minimized using 5000 steps of steepest descent minimization. Sodium ions were used to neutralize the backbone of the DNA; an octahedral TIP3P water box was extended up to 10 Å from any of the solute atoms. Consequently, a steepest descent minimization was carried out for 10000 steps, and then the system was heated from 10 to 300 K in 150 ps with the restraint force for 50 kcal/mol Å^2. The equilibration was carried

DOI: 10.1021/jo502627b
J. Org. Chem. 2015, 80, 2128−2138

out for 700 ps with 1 atm pressure at constant temperature of 300 K. After this, the MD simulation of the duplex system was carried out for 50 ns. The RMSD, hydration, percentage of stacking and W−C H-bond occupancy of the duplex system was calculated using the PTRAJ module of AMBER 12.

The crystal structure of the BF (PDB entry: 1XC9) complexed with the template-containing BPdG adduct, and the primer was utilized as the model for MD simulations. The *ff99SB* with recent force field modifications[41] was used for protein and DNA. The LdG adduct was energy optimized (HF/6-31G*) using Gaussian 09[37], and RESP charges were calculated and fitted using a reported procedure.[38,39] All the minimization and MD simulation were carried out using SANDER and PMEMD modules of the AMBER 12. The Leap module of AMBER 12 have been used to replace the BPdG with the LdG adduct at the template-primer junction. Hydrogen atoms were added to the crystal structure, and the structure was minimized using 500 steps of the steepest descent minimizations. Sodium ions were added to neutralize the net charge of the system using XLeap. A periodic octahedral TIP3P water box was extended up to 10 Å from any of the solute atoms which yielded around 10000 water molecules. Minimization of solvent and counterions for 5000 steps of the steepest descent and 5000 steps of conjugate gradient were carried out. After minimization, 50 ps of initial MD simulations were carried out to relax the water and Na^+ ions, while the protein−DNA complex was detained with a restraint force for 50 kcal/mol Å^2. The system was then heated from 10 to 300 K at constant volume over 100 ps with the restraint force of 15 kcal/mol Å^2 on solute atoms. The equilibrated system was then subjected to the production MD for 40 ns with constant 1 atm pressure and 300 K temperature. Long-range interactions were treated with the particle mesh Ewald method, and a 10 Å cutoff was applied for the nonbonded Lennard-Jones potential. The SHAKE algorithm was used to constrain all the hydrogen atoms. MD trajectories were visualized using UCSF Chimera. RMSD, ensemble, and superimposition analysis were carried out using the ptraj module in AMBER 12. Backbone torsions and base pair helical parameters were calculated using X3DNA.[51] Figures were rendered using PyMOL.

■ ASSOCIATED CONTENT

Supporting Information

Synthetic strategy for the 2-(aminomethyl)-1,3-diacetoxyanthraquinone; PAGE of full length extension reactions; energy optimized geometry and calculated RESP charges for the LdG modified nucleotide; time dependent RMSD maps of the protein and DNA complex; MD simulation snapshots from the conformational ensembles; superimposed structure of the BF in complex with unmodified and LdG DNA; Watson-Crick H-bond occupancy of the LdG modified ds DNA; major and minor groove hydration of the LdG modified DNA; local base pair parameters of the template-primer ds DNA; local base pair step parameters of the template-primer ds DNA; backbone torsion angles of nucleotide in the template strand; MALDI-TOF mass data of the DNA sequences; the UV melting curves; and CD spectra for the unmodified and LdG modified DNAs.; ^{1}H, ^{13}C, and ^{31}P NMR spectra of all the new compounds; and MALDI spectra of the oligonucleotides. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: pradeep@chem.iitb.ac.in

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are thankful to Prof. Deepak T. Nair for introducing us to the field of DNA adducts and polymerases and also for his critical reading of the manuscript; Prof. David A. Case for waiving the licensing fee for AMBER12; Prof. Philip C. Bevilacqua for the assistance in error analyses of the thermodynamic parameters; and Prof. I.N.N. Namboothiri for providing access to the melting point apparatus. We are also thankful to the central facility supported by IRCC-IIT Bombay for MALDI spectra and the Computer Centre, IIT Bombay, for use of the HPC facility. This work is financially supported by grants from the Department of Biotechnology (DBT)-Government of India. P.P.G. and S.H. thank the Council of Scientific and Industrial Research (CSIR) and Department of Atomic Energy-Board of Research in Nuclear Sciences (DAE-BRNS), respectively, for their fellowships.

■ REFERENCES

(1) Wogan, G. N.; Hecht, S. S.; Felton, J. S.; Conney, A. H.; Loeb, L. A. *Semin. Cancer Biol.* **2004**, *14*, 473−486.
(2) Geacintov, N. E.; Broyde, S. In *The Chemical Biology of DNA Damage*; Wiley-VCH: Weinheim, Germany, 2010; pp 3−15.
(3) Jackson, S. P.; Bartek, J. *Nature* **2009**, *461*, 1071−1078.
(4) Loeb, L. A.; Monnat, R. J. *Nat. Rev. Genet.* **2008**, *9*, 594−604.
(5) Sale, J. E.; Lehmann, A. R.; Woodgate, R. *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 141−152.
(6) Washington, M. T.; Carlson, K. D.; Freudenthal, B. D.; Pryor, J. M. *Biochim. Biophys. Acta, Proteins Proteomics* **2010**, *1804*, 1113−1123.
(7) Hemminki, K. *Arch. Toxicol.* **1983**, *52*, 249−285.
(8) Sturla, S. J. *Curr. Opin. Chem. Biol.* **2007**, *11*, 293−299.
(9) Dipple, A. *Carcinogenesis* **1995**, *16*, 437−441.
(10) Guengerich, F. P. *Chem. Rev.* **2006**, *106*, 420−452.
(11) Eoff, R. L.; Egli, M.; Guengerich, P. *The Chemical Biology of DNA Damage*; Wiley-VCH: Weinheim, Germany, 2010; pp 299−319.
(12) Kim, H.-Y. H.; Cooper, M.; Nechev, L. V.; Harris, C. M.; Harris, T. M. *Chem. Res. Toxicol.* **2001**, *14*, 1306−1314.
(13) Penning, T. M. In *The Chemical Biology of DNA Damage*; Wiley-VCH: Weinheim, Germany, 2010; pp 131−147.
(14) Turesky, R. J. In *The Chemical Biology of DNA Damage*; Wiley-VCH: Weinheim, Germany, 2010; pp 157−173.
(15) Inoue, K.; Yoshida, M.; Takahashi, M.; Shibutani, M.; Takagi, H.; Hirose, M.; Nishikawa, A. *Food Chem. Toxicol.* **2009**, *47*, 184−191.
(16) Nakanishi, F.; Nagasawa, Y.; Kabaya, Y.; Sekimoto, H.; Shimomura, K. *Plant Physiol. Biochem.* **2005**, *43*, 921−928.
(17) Blömeke, B.; Poginsky, B.; Schmutte, C.; Marquardt, H.; Westendorf, J. *Mutat. Res., Fundam. Mol. Mech. Mutagen.* **1992**, *265*, 263−272.
(18) Ishii, Y.; Okamura, T.; Inoue, T.; Fukuhara, K.; Umemura, T.; Nishikawa, A. *Chem. Res. Toxicol.* **2009**, *23*, 134−141.
(19) Poginsky, B.; Westendorf, J.; Blömeke, B.; Marquardt, H.; Hewer, A.; Grover, P. L.; Phillips, D. H. *Carcinogenesis* **1991**, *12*, 1265−1271.
(20) Westendorf, J.; Pfau, W.; Schulte, A. *Carcinogenesis* **1998**, *19*, 2163−2168.
(21) Ishii, Y.; Inoue, K.; Takasu, S.; Jin, M.; Matsushita, K.; Kuroda, K.; Fukuhara, K.; Nishikawa, A.; Umemura, T. *Chem. Res. Toxicol.* **2012**, *25*, 1112−1118.
(22) Akhtar, M.; Zareen, S.; Yeap, S.; Ho, W.; Lo, K.; Hasan, A.; Alitheen, N. *Molecules* **2013**, *18*, 10042−10055.
(23) Harris, T. M.; Harris, C. M. *Current Protocols in Nucleic Acid Chemistry*; John Wiley & Sons, Inc.: Hoboken, NJ, 2001; pp 1.3.1−1.3.15.
(24) Harwood, E. A.; Sigurdsson, S. T.; Edfeldt, N. B. F.; Reid, B. R.; Hopkins, P. B. *J. Am. Chem. Soc.* **1999**, *121*, 5081−5082.
(25) Champeil, E.; Pradhan, P.; Lakshman, M. K. *J. Org. Chem.* **2007**, *72*, 5035−5045.
(26) Gore, K. R.; Nawale, G. N.; Harikrishna, S.; Chittoor, V. G.; Pandey, S. K.; Höbartner, C.; Patankar, S.; Pradeepkumar, P. I. *J. Org. Chem.* **2012**, *77*, 3233−3245.
(27) Siegfried, N. A.; Metzger, S. L.; Bevilacqua, P. C. *Biochemistry* **2007**, *46*, 172−181.
(28) Terashima, I.; Matsuda, T.; Fang, T.-W.; Suzuki, N.; Kobayashi, J.; Kohda, K.; Shibutani, S. *Biochemistry* **2001**, *40*, 4106−4114.

DOI: 10.1021/jo502627b
J. Org. Chem. 2015, 80, 2128−2138

(29) Zhao, L.; Christov, P. P.; Kozekov, I. D.; Pence, M. G.; Pallan, P. S.; Rizzo, C. J.; Egli, M.; Guengerich, F. P. *Angew. Chem., Int. Ed.* **2012**, *51*, 5466−5469.
(30) Alekseyev, Y. O.; Dzantiev, L.; Romano, L. J. *Biochemistry* **2001**, *40*, 2282−2290.
(31) Arghavani, M. B.; SantaLucia, J.; Romano, L. J. *Biochemistry* **1998**, *37*, 8575−8583.
(32) Alekseyev, Y. O.; Romano, L. J. *Biochemistry* **2000**, *39*, 10431−10438.
(33) Kornyushyna, O.; Berges, A. M.; Muller, J. G.; Burrows, C. J. *Biochemistry* **2002**, *41*, 15304−15314.
(34) Xu, P.; Oum, L.; Beese, L. S.; Geacintov, N. E.; Broyde, S. *Nucleic Acids Res.* **2007**, *35*, 4275−4288.
(35) Doisy, R.; Tang, M.-S. *Biochemistry* **1995**, *34*, 4358−4368.
(36) Alekseyev, Y. O.; Romano, L. J. *Biochemistry* **2002**, *41*, 4467−4479.
(37) Frisch, M. J. *Gaussian 09*, revision A.02; Gaussian, Inc.: Wallingford, CT, 2009.
(38) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(39) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(40) Case, D. A.; *AMBER 12*; University of California: San Francisco, 2012.
(41) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(42) Casale, R.; McLaughlin, L. W. *J. Am. Chem. Soc.* **1990**, *112*, 5264−5211.
(43) Saenger, W. In *Principles of Nucleic Acid Structure*; Springer: New York, 1984; pp 51−104.
(44) Hsu, G. W.; Huang, X.; Luneva, N. P.; Geacintov, N. E.; Beese, L. S. *J. Biol. Chem.* **2005**, *280*, 3764−3770.
(45) Broyde, S.; Wang, L.; Zhang, L.; Rechkoblit, O.; Geacintov, N. E.; Patel, D. J. *Chem. Res. Toxicol.* **2007**, *21*, 45−52.
(46) Zhang, L.; Rechkoblit, O.; Wang, L.; Patel, D. J.; Shapiro, R.; Broyde, S. *Nucleic Acids Res.* **2006**, *34*, 3326−3337.
(47) Barsky, D.; Kool, E. T.; Colvin, M. E. *J. Biomol. Struct. Dyn.* **1999**, *16*, 1119−1134.
(48) Choi, J.-Y.; Guengerich, F. P. *J. Biol. Chem.* **2004**, *279*, 19217−19229.
(49) Blackburn, G. M.; Williams, D. M.; Loakes, D.; Gait, M. J. *Nucleic Acids in Chemistry and Biology*, 3rd ed.; RSC Publishing: Cambridge, 2006; p 143.
(50) Ganguly, M.; Wang, R.-W.; Marky, L. A.; Gold, B. *J. Phys. Chem. B* **2010**, *114*, 7656−7661.
(51) Lu, X.-J.; Olson, W. K. *Nat. Protoc.* **2008**, *3*, 1213−1227.

DOI: 10.1021/jo502627b
J. Org. Chem. 2015, 80, 2128−2138