JOC The Journal of Organic Chemistry

Article

pubs.acs.org/joc

The N^2-Furfuryl-deoxyguanosine Adduct Does Not Alter the Structure of B-DNA

Pratibha P. Ghodke,[†] Kiran R. Gore,[†,∥] S. Harikrishna,[†] Biswajit Samanta,[†] Jithesh Kottur,[‡,§] Deepak T. Nair,[*,§,⊥] and P. I. Pradeepkumar[*,†]

[†]Department of Chemistry, Indian Institute of Technology Bombay, Mumbai-400076, India

[∥]Department of Chemistry, University of Mumbai, Mumbai-400098, India

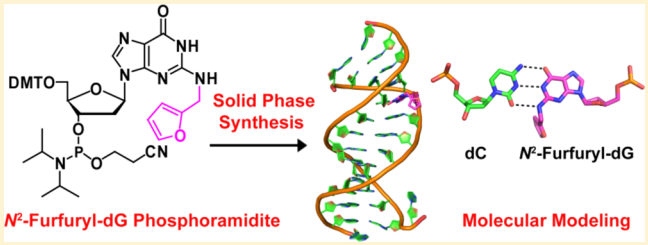
[§]Regional Centre for Biotechnology, NCR Biotech Science Cluster, Third Milestone, Faridabad-Gurgaon Expressway, Faridabad-121001, India

[‡]Manipal University, Manipal-576104, India

[⊥]National Centre for Biological Sciences (NCBS-TIFR), GKVK Campus, Bellary Road, Bangalore-560065, India

Supporting Information

ABSTRACT: N^2-Furfuryl-deoxyguanosine (fdG) is carcinogenic DNA adduct that originates from furfuryl alcohol. It is also a stable structural mimic of the damage induced by the nitrofurazone family of antibiotics. For the structural and functional studies of this model N^2-dG adduct, reliable and rapid access to fdG-modified DNAs are warranted. Toward this end, here we report the synthesis of fdG-modified DNAs using phosphoramidite chemistry involving only three steps. The functional integrity of the modified DNA has been verified by primer extension studies with DNA polymerases I and IV from *E. coli*. Introduction of fdG into a DNA duplex decreases the T_m by ~1.6 °C/modification. Molecular dynamics simulations of a DNA duplex bearing the fdG adduct revealed that though the overall B-DNA structure is maintained, this lesion can disrupt W−C H-bonding, stacking interactions, and minor groove hydrations to some extent at the modified site, and these effects lead to slight variations in the local base pair parameters. Overall, our studies show that fdG is tolerated at the minor groove of the DNA to a better extent compared with other bulky DNA damages, and this property will make it difficult for the DNA repair pathways to detect this adduct.



INTRODUCTION

DNA damage is central to chemical carcinogenesis and occurs due to the continuous exposure to various endogenous and exogenous genotoxic agents.[1] Different carcinogens create a variety of metabolites, which can form significant modifications (adducts) in the DNA strands.[2] Such DNA adducts have the potential to inhibit the DNA replication process.[3] The presence of these adducts leads to recruitment of pathways and molecules that serve to neutralize their deleterious effects on replication and remove them from the genome.[4] One such event is DNA damage response (DDR), which eliminates various DNA adducts by various DNA repair mechanisms, thereby preventing the formation of lethal mutations.[4] Translesion synthesis (TLS) is a damage tolerance pathway in which the damage is bypassed with the assistance of specialized DNA polymerases (Y-family) to rescue DNA replication stalled at these lesions.[5] Failure to neutralize the deleterious effects of the damaged nucleotide on different genomic processes can lead to genetic instability, which generally leads to the death or oncogenic transformation of the cell.[4]

The N^2-position of the deoxyguanosine (dG) is one of the most susceptible sites in DNA, and a large number of adducts are known to form at this site from various genotoxic agents.[2] Many of the N^2-dG adducts such as N^2-benzo[*a*]pyrene, N^2-carboxyethyl, N^2-furfuryl (fdG, Figure 1), etc., are reported to be lethal to the cells.[6−9] The fdG adduct is caused by furfuryl alcohol (FFA, Figure 1), which is used as a flavoring agent.[10] Also, FFA has been detected in several foodstuffs such as coffee, red wine, rice cakes, cooked meat, and milk products.[11] It is mostly produced in heat-treated foods via thermal and acid-catalyzed dehydration of pentoses.[12] Moreover, it can be formed by reduction of hepatocarcinogenic agent furfural.[12] FFA is an established rodent carcinogen, and its carcinogenic effects on humans are being investigated.[13] Recently, FFA-induced fdG adduct has also been detected in human lung specimens.[10,13,14] The metabolic pathway of FFA follows bioactivation of furfuryl alcohol by endogenous sulfotransferases to form a furfuryl sulfate intermediate (Figure 1).[13] This highly unstable intermediate generates an electrophilic carbonium ion, which is prone to react with nucleophilic sites on the DNA and forms furfuryl-modified DNA adducts.[13] fdG (Figure 1) and N^6-furfuryl deoxyadenosine (fdA) are the main exocyclic

Received: October 9, 2015
Published: December 9, 2015

DOI: 10.1021/acs.joc.5b02341
J. Org. Chem. 2016, 81, 502−511

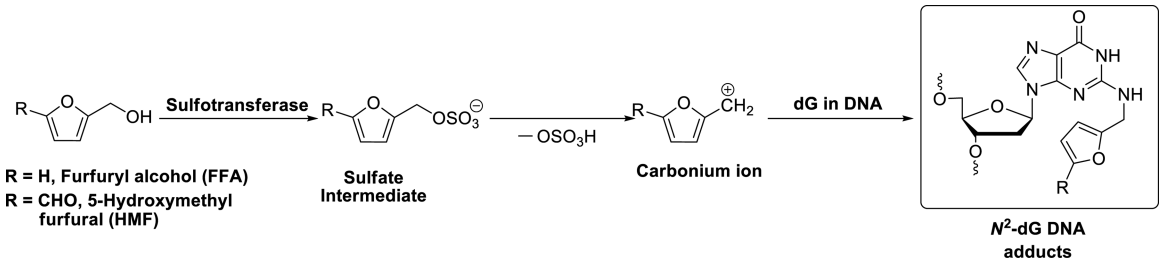


Figure 1. Likely metabolic pathway depicting the formation of the N^2-dG DNA adducts.

adducts, which have been detected so far and associated with mutagenesis and lethality.[10,12] Similar adducts can also be induced by 5-hydroxymethylfurfural (HMF, Figure 1),[15] a carcinogen present in food stuffs, beverages, and cigarette smoke, which is also produced by the Maillard reaction.[15,16]

The fdG DNA adduct is also a structural analogue of the DNA adduct formed from the antimicrobial drug nitrofurazone (NFZ).[6] NFZ can be reduced in the presence of cellular nitroreductases, and the reduced NFZ upon acetylation holds the potential for the formation of N^2-dG adducts; however, the exact structure of NFZ-induced DNA adduct is yet to be reported. The antibacterial action has been attributed to the damage of the DNA caused by NFZ.[6,17] However, bacteria can develop resistance to NFZ by employing polymerase IV (Pol IV), which efficiently and accurately bypasses fdG by TLS.[6,18] Therefore, fdG adduct serves as a model N^2-dG adduct to elucidate the mechanistic insights of TLS by functional and structural studies.[6] To study the effect of this adduct, a rapid and reliable access to the fdG-modified DNAs is warranted. The reported method of synthesis relies on postsynthetic modification, where 2-F-O^6-trimethylsilylethyl (TMSE)-deoxyinosine (2-F-dI)-modified DNA is subjected to fluoro displacement using furfuryl amine to generate final fdG-modified DNAs.[6] However, high quality material in high preparative yield is not guaranteed due to varying yield of the substitution reaction. Moreover, the syntheses of the 2-F-dI-modified building block involve five steps, and it demands the use of highly toxic fluorinating agent HF−pyridine.[19] To address those difficulties, herein we report the synthesis of fdG-modified DNAs employing the corresponding modified phosphoramidite, which is accessed in the least number of steps using less hazardous chemicals.

We have synthesized the fdG nucleoside using two different strategies in single or three steps (Scheme 1). Reductive amination of dG with furfural directly yielded fdG nucleoside in a one-step reaction, while the Buchwald−Hartwig coupling of the furfurylamine with protected 2-chloro-deoxyinosine produced the nucleoside in three steps. The fdG-modified phosphoramidite have been successfully incorporated into DNA strands of different lengths using solid-phase synthesis. The functional integrity of the fdG-modified DNA was validated by primer extension studies using replicative and translesion polymerases from *E. coli*. The effect of fdG modification on the thermal stability of duplex DNAs have been investigated by UV-melting studies, which are further validated by molecular dynamics (MD) simulations.

RESULTS AND DISCUSSION

Synthesis of fdG-Modified Phosphoramidite and Its Incorporation into DNAs. The synthesis of fdG-modified nucleoside **4** and the corresponding phosphoramidite **6** is shown in Scheme 1. In the first route utilizing Buchwald−Hartwig (B−H) coupling, the 2-chloro deoxyinosine **1**[20] was reacted with furfurylamine in the presence of $Pd(OAc)_2$, (*R*)-BINAP, and Cs_2CO_3 in toluene to give modified nucleoside **2** in 57% yield.[20] Deprotection of O^6-*p*-nitrophenylethyl (NPE) by DBU in pyridine produced the diacetate compound **3**, which upon deacetylation using methylamine in ethanol yielded the fdG nucleoside **4**.[21] Protection of 5′-OH group as DMT ether led to the nucleoside **5** (76% yield after three steps from **2**). Phosphitylation at the 3′-OH using 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (CEP-Cl) in DCM furnished fdG phosphoramidite **6** in 74% yield.[22] In an alternate approach, we synthesized fdG nucleoside **4** in 46% yield by reductive amination of dG with furfural in the presence of sodium cyanoborohydride. Though the purification of the product is cumbersome, this method provides access to fdG nucleoside in a single step, and it is, therefore, superior to the chemistry which relies on B−H coupling. Protection of 5′-OH by DMT of **4** yielded the nucleoside **5** in 70% yield. Subsequent phosphitylation produced the fdG phosphoramidite in just three steps with an overall yield of 23% (Scheme 1).

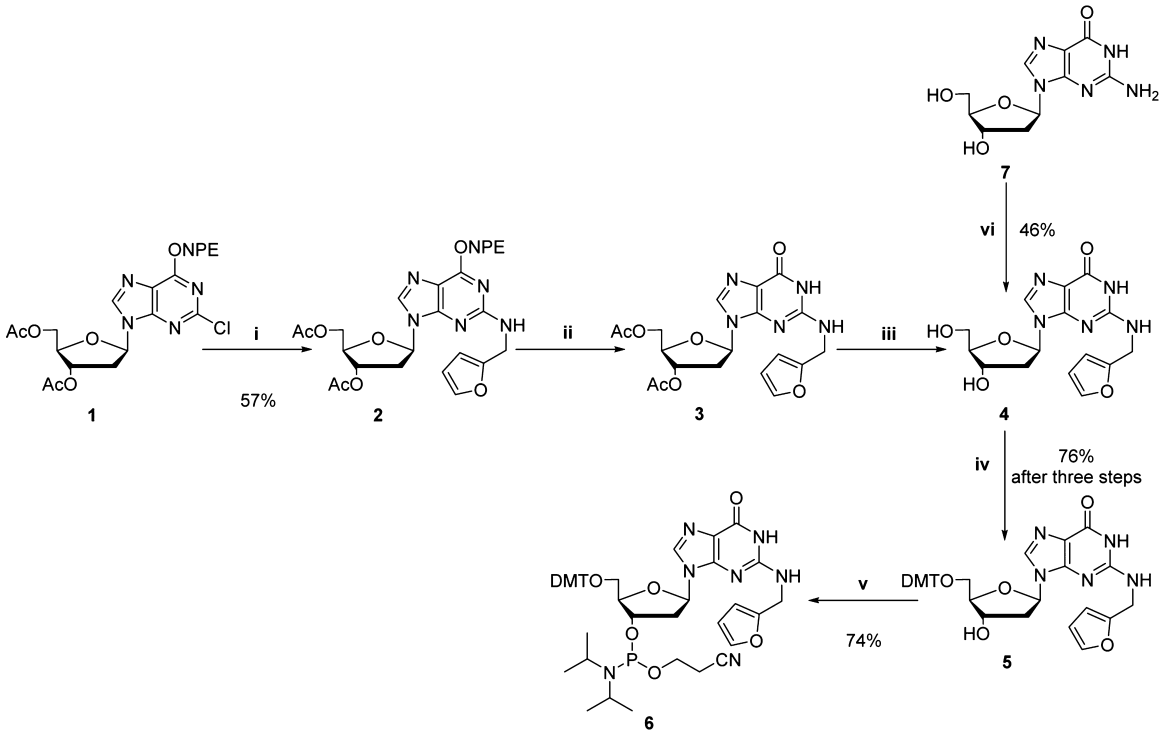
The fdG phosphoramidite obtained was utilized to introduce the fdG modification into the DNA strands of varying lengths (sequences **D1** to **D6** as shown in Table 1) using standard solid-phase synthesis protocols. The 14-mer DNAs, **D1** and **D2**, were used to evaluate the thermal stability of the modified duplexes. The modified DNAs **D3** to **D5** were 18 nt long and had been used in a previous study to crystallize a ternary complex with DNA Pol IV from *E. coli*.[23] The primer extension assays against the fdG site were performed with a 50 nt long DNA, **D6** or 18 nt DNA, **D3**. The chemical integrities of all the modified DNAs were confirmed by mass spectrometric analysis using ESI or MALDI ionization methods (Table 1).

Thermal Melting Studies. UV-melting studies were carried out to examine the effect of fdG modification on DNA duplex stability. The melting temperatures (T_m) are summarized in Table 2. The unmodified duplex (**D7**−**D8**) exhibited T_m of 62.5 while T_m of single (**D1**−**D8** and **D2**−**D7**) and dual (**D1**−**D2**) fdG-modified duplexes were 61.3, 60.5, and 59.6 °C, respectively. These results show that the DNA duplex with single and dual fdG modification show reduction of ∼1.6 °C in T_m per modification. Thermodynamic parameters ($\Delta H°$, $\Delta S°$, and $\Delta G°_{298}$) were also determined from melting curves using a two-state model[24] (Table 2 and Figure S1 of the Supporting Information). No appreciable change in $\Delta H°$ and $\Delta S°$ was observed due to the presence of the fdG in a DNA duplex. The observed slight destabilization may have arisen from the local structural/solvation effects, which were probed using MD simulations (see the respective section for details).

Primer Extension Studies. The functional integrity of the fdG lesion was assessed by probing the effect of this

DOI: 10.1021/acs.joc.5b02341
J. Org. Chem. 2016, 81, 502−511

Scheme 1. Synthesis of N^2-Furfuryl-deoxyguanosine (fdG) Phosphoramidite 6[a]



[a]Reagents and conditions: (i) furfurylamine, $Pd(OAc)_2$, (*R*)-BINAP, Cs_2CO_3, toluene 80 °C, 45 h; (ii) DBU, pyridine, rt, 6 h; (iii) 40% $MeNH_2$ in EtOH (v/v), rt, 16 h; (iv) DMT-Cl, pyridine, rt, 2 h; (v) CEP-Cl, DIPEA, DCM, rt, 40 min; (vi) $NaCNBH_3$, furfural, MeOH, 50 °C, 36 h.

Table 1. FdG-Modified and the Unmodified DNAs and Their Molecular Weights

Code	DNA Sequences (5'-3')	MW (calc.)	MW[a] (found)
D1	GCCGXAATAGCGCA	4377	4379
D2	TGCXCTATTCCGGC	4310	4312
D3	TCTXGGGTCCTAGGACCC	5570	5571
D4	TCTAXGGTCCTAGGACCC	5554	5555
D5	TCTAGGXTCCTAGGACCC	5554	5555
D6	TCCTACCGTGCCTACCTGAACAGCTGGTCA CACTXATGCCTACGAGTACG	15346	15348
D7	GCCGGAATAGCGCA	4297	4299
D8	TGCGCTATTCCGGC	4230	4232

[a]The modified (X = fdG) and unmodified DNAs were characterized using ESI-MS or MALDI-MS.

modification on DNA replication. The ability of the replicative Klenow fragment of DNA polymerase I exo$^+$ (Pol I) and translesion Pol IV to bypass the fdG adduct was tested. The primer extension reactions were performed with 6-carboxyflurescein (6FAM)-labeled primer using either all four dNTPs individually or together. The extension reactions were analyzed by denaturing PAGE (20%), and the products were visualized at 488 nm (Figure 2 and Figure S2 of the Supporting Information). In reactions employing Pol IV, and only dCTP, we observed a prominent extension product of 16 nucleotide length, which corresponds to single dCTP incorporation (Figure 2, and Figure S2 of the Supporting Information). These results validate the fidelity and catalytic efficiency of Pol IV for dCTP incorporation across fdG, which is in agreement with the literature reports.[6,23] Furthermore, the reaction with a mixture of all dNTPs was more efficient with the fdG-containing DNA template compared to that in the unmodified one (Figure 2 and Figure S2 of the Supporting Information).[6,7,23] The better efficacy of Pol IV for TLS across fdG can be attributed to the unique structural features of Pol IV, which helps the enzyme to accommodate the fdG adduct in the active site.[23] As expected, the replicative polymerase Pol I having 3′ to 5′ exonuclease activity (KF exo$^+$) was inefficient to bypass fdG template when compared to unmodified one (Figure S3 of the Supporting Information).[6] As a result, only ~5.5% of the full length product was observed in the experiment using fdG-modified template. Under similar conditions, unmodified template yielded >90% of the full length product (Figure S3 of the Supporting Information).

DOI: 10.1021/acs.joc.5b02341
J. Org. Chem. 2016, 81, 502−511

Table 2. DNA Duplexes Reported in This Study and Their Thermodynamic Parameters

Code	Duplex	T_m[a] (°C)	ΔT_m[b]/ Mod (°C)	$\Delta H°$ (kcal/mol)	$\Delta S°$ (eu)	$T\Delta S°$ (kcal/mol)	$\Delta G°$ (kcal/mol)
D7-D8	5'-GCCGGAATAGCGCA-3' 3'-CGGCCTTATCGCGT-5'	62.5 ±0.8	-	−67.8 ±1.1	−202 ±4.1	−60.2	−7.5 ±0.1
D1-D8	5'-GCCGXAATAGCGCA-3' 3'-CGGCCTTATCGCGT-5'	61.3 ±0.8	−1.2	−67.1 ±0.6	−200 ±2.0	−59.8	−7.2 ±0.1
D2-D7	5'-TGCXCTATTCCGGC-3' 3'-ACGCGATAAGGCCG-5'	60.5 ±0.1	−2.0	−65.9 ±1.2	−196 ±3.7	−58.9	−7.0 ±0.1
D1-D2	5'-GCCGXAATAGCGCA-3' 3'-CGGCCTTATCXCGT-5'	59.6 ±0.4	−1.7	−67.3 ±0.6	−200 ±1.2	−60.3	−7.0 ±0.1

[a]Thermal denaturation was performed at 25 °C using phosphate buffer (100 mM NaCl, 20 mM phosphate buffer, 0.1 mM EDTA, pH 7.4), and monitored at 260 nm with 0.5 °C min^{-1} in the range 20−80 °C using three cycles. The concentration of each DNA strand was 1 μM. The fdG modification is highlighted in red and represented as a X. [b]ΔT_m represents the $[T_m(DNA_{mod}) - T_m(DNA_{unmod})]$. The T_m values reported are average of three independent measurements with the estimated standard deviation. Thermodynamic parameters were calculated from the melting curves using a two state model and error propagation of thermodynamic parameters were calculated using a reported protocol.[24]

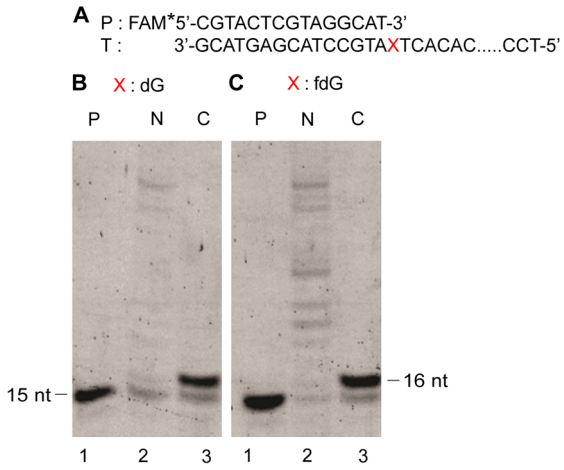


Figure 2. PAGE (20%, 8 M urea) of primer extension reactions with *E. coli* DNA polymerase IV. (A) Partial sequence of primer and template (X = dG or fdG); (B) and (C) lane 1: primer only; lane 2: reaction with mixture of all dNTPs; lane 3: reaction with dCTP.

The translesion polymerases exhibit low processivity, and they are generally not efficient in synthesizing full length products.[5,25−27] To verify this, running start experiments were performed using 11-mer primer and 18 mer template (**D3**, Table 1) in which the modification site is three nucleotides away from the starting site (Figure S4A of the Supporting Information). Results clearly showed that, though the bypass across fdG is facilitated, formation of full length product was only observed in experiments employing a large excess of enzyme (Figure S4B and S4C of the Supporting Information).[6,7,23]

To validate the ability of Pol I, which lacks 3′ to 5′ exonuclease activity (KF exo$^-$), to perform TLS across fdG modification, the running start experiments were also carried out using the same template primer system as described above (Figure S5A of the Supporting Information). The results indicated that the KF exo$^-$ was capable of bypassing fdG modification in a manner similar to that by Pol IV (Figure S5B and S5C of the Supporting Information). Overall, the experiments with TLS as well as replicative polymerases clearly underscore the functional integrity of the chemically synthesized fdG DNA templates.

Molecular Modeling Studies. To evaluate the structural and conformational flexibility of fdG-containing DNA and to validate the effect of fdG on thermal stability of the duplex, MD simulations in explicit solvent were performed for 300 ns using AMBER 14.[28] In brief, the energy-optimized (HF/6-31G(d), Gaussian 09[29]) and RESP charge-fitted fdG nucleotide (Figure S6 of the Supporting Information) was placed in a B-DNA sequence, which was used in the thermal melting studies (**D1−D8**, Table 2). To compare the possible structural changes induced by the fdG in a duplex, the corresponding unmodified DNA (**D7−D8**, Table 2) was also simulated for 300 ns. The average RMSD of heavy atoms with respect to the corresponding equilibrated structure of the unmodified and the modified DNA was ∼2.8 Å and ∼3.5 Å, respectively (Figure S7 of the Supporting Information). The final MD snapshot of modified and unmodified DNA (Figure 3) showed no significant changes in the global structure of the modified DNA in comparison to the unmodified structure. The local architectures of B-DNA around fdG and the corresponding

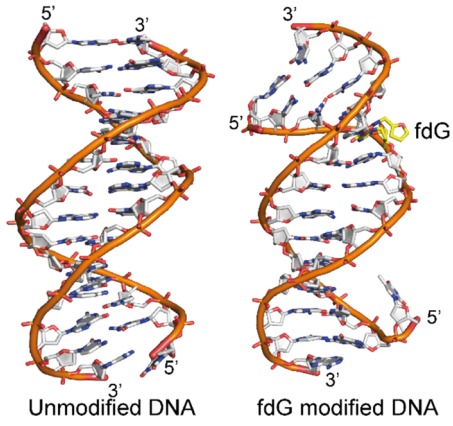


Figure 3. Final MD simulation (300 ns) snapshots of the unmodified and fdG-modified DNA show the global conformation of the B-DNA. The fdG adduct is highlighted in yellow color, the backbone is represented as a cartoon, and the atoms are shown as sticks.

DOI: 10.1021/acs.joc.5b02341
J. Org. Chem. 2016, 81, 502−511

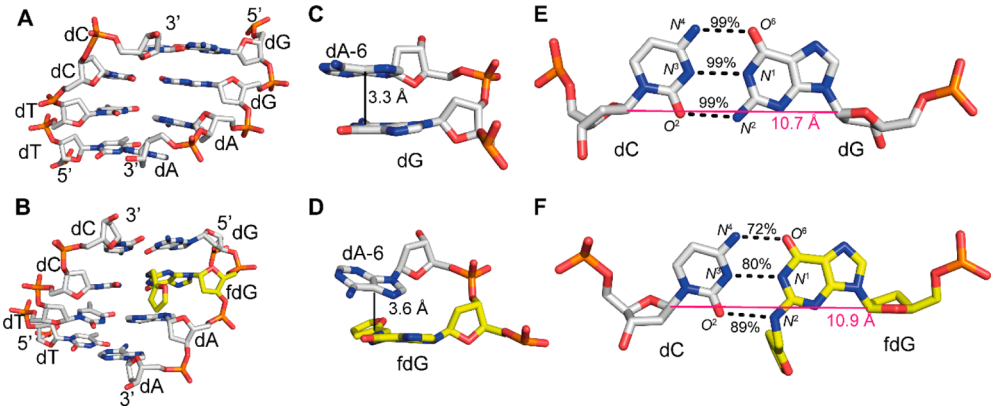


Figure 4. MD snapshot (300 ns) of unmodified and fdG-containing modified DNA depicting the W−C H-bonding and stacking interactions. (A) Architecture of B-DNA in the unmodified region. (B) Architecture of B-DNA around the fdG in the modified DNA. (C) Stacking interaction between the dG and the successive nucleotide dA in the unmodified DNA. (D) Stacking interaction between the fdG and the successive nucleotide dA in the modified DNA. (E) W−C base pairing and H-bonds between dG and dC atoms in the unmodified DNA. (F) W−C base pairing and H-bonds between fdG and dC atoms in the modified DNA. The occupancies of H-bonds are mentioned in percentage of the total simulation time. The black dashed line indicates the W−C H-bonds between the bases, the magenta line indicates the C1′−C1′distance between the bases, and the black line indicates the stacking distance between the bases. The fdG adduct is highlighted in yellow color, and atoms are represented as sticks.

unmodified region are shown in Figure 4A and 4B. There are only slight disturbances in the stacking and base pairing interactions observed at the modified site.

To probe the stacking interactions in detail, the intrastrand phosphate distances between the bases were calculated from the 300 ns MD trajectories (Figure S8 of the Supporting Information). The results indicated that the average phosphate intrastrand phosphate atom (P−P) distance around fdG in the modified DNA (∼6.7 Å ± 0.7) was slightly higher in comparison to that in the unmodified DNA (∼5.7 Å ± 0.5). This high phosphate distance is a sign of relatively weak stacking interactions between the fdG and the neighboring base (Figure 4C and 4D). To verify this inference further, the percentages of stacking interactions around the fdG-modified site were calculated from the MD simulations (Table S1 of the Supporting Information). The stacking interactions of fdG with dA were found to be present during ∼68% of the total simulation time. In the case of the unmodified DNA the stacking interaction of dG with dA was found to be present during 99% of the total simulation time (Table S1 of the Supporting Information). The distance and percentage occupancies of W−C H-bonds between the fdG and dC in the modified DNA and the corresponding parameters for the dG and dC in the unmodified DNA were also probed during the MD simulations (Figure 4E, 4F, and S9 of the Supporting Information). The results indicate that the occupancies of H-bonds were slightly less in the fdG-modified site. The average H-bond distance between the $O^6\cdots N^4$ atoms was found to be fluctuating between 2.92 to 7.66 Å (Figure S9 of the Supporting Information). The average distances between the $N^1\cdots N^3$ and the $N^2\cdots O^2$ atoms were found to be 2.81 ± 0.12 Å and 2.91 ± 0.09 Å, respectively (Figure S9 and Table S2 of the Supporting Information). The distances and percentage occupancies indicate that W−C H-bond in the unmodified DNA (dG:dC) was more or less intact during the course of MD simulations (Figure S9 and Table S2 of the Supporting Information). Furthermore, the C1′−C1′ distance between the fdG:dC was found to be ∼10.9 ± 0.2 Å, which was similar to the ∼10.7 ± 0.2 Å distance observed for the unmodified dG:dC pair (Figure 4E and 4F). In the crystal structures of Pol IV in complex with DNA bearing the fdG adduct, the fdG:dC base pair was found to retain the hydrogen bonding and stacking interactions observed in canonical Watson−Crick base pairs (PDB entries: 4Q44 and 4Q45).[23]

The sugars fdG in the modified and the corresponding dG in the unmodified DNAs predominantly adopt C2′-endo puckering, and the χ was found to be −142° ± 22 (fdG) and −151° ± 26 (dG) during the course of MD simulations. The average of the six backbone dihedral angles (α, β, γ, ε, δ, and ζ) during the MD simulations for the fdG nucleotide in the modified DNA were found to be −71.08° ± 11.8, 110.9° ± 20.9, 55.9° ± 8.7, 120.1° ± 19.5, 88.02° ± 36.8, and −72.76° ± 26.8, respectively, whereas those for the dG in the unmodified were found to be −67.5° ± 30.2, 158.6° ± 44.3, 43.2° ± 38.6, 104.8° ± 28.5, 138.3° ± 42.8, and −75.9° ± 38.8, respectively. Among these, β and δ were found to be more deviated in the fdG in comparison to the dG in the unmodified DNA. The variations in the β and δ backbone dihedral angles seem to affect the stacking interactions and W−C H-bondings.[30]

The six local base pair parameters were also calculated using X3DNA from the MD simulation trajectories of the unmodified and modified DNA (Figure 5). The stretch values in both the unmodified and modified DNA were quite similar along the sequences (Figure 5B), except at the end part of the DNA, which is due to the fraying effect that occurred during the course of simulations. The stagger and buckle parameters were also found to be similar in both the unmodified and the modified DNA (Figure 5C and 5D). These results indicate that there is no global helical bending in the B-DNA geometry. The propeller values were found to be higher in the fdG-modified DNA than those in the unmodified DNA along the sequence (Figure 5E). The shear and opening parameters (Figure 5A and 5F) deviated only around the fdG site of the modified DNA in comparison to the corresponding unmodified site.

The number of water molecules and their percentage occupancies around the modified site were also calculated. On average, ∼3 water molecules were found to be present in the minor groove of the dG:dC base pair throughout the MD simulations whereas only one water molecule was found to be present in the minor groove of the fdG:dC base pair in the modified DNA (Table S3 of the Supporting Information). The hydration was also affected in the consecutive base pair dT:dA,

DOI: 10.1021/acs.joc.5b02341
J. Org. Chem. 2016, 81, 502−511

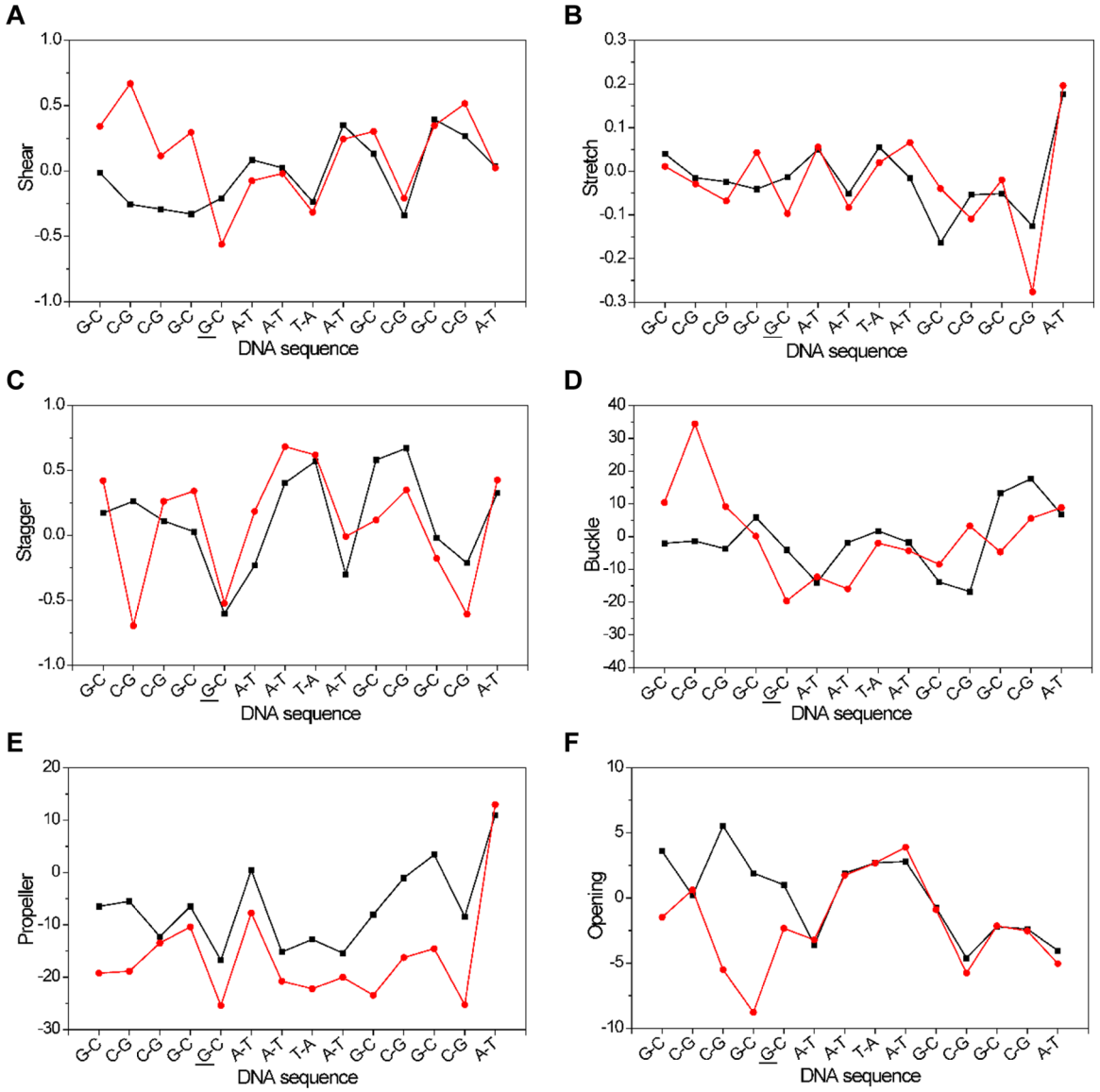


Figure 5. Six local base pair parameters of unmodified and fdG-modified DNAs. Local base pair parameters for each base pair were calculated from the averaged structure obtained from MD simulations of unmodified (black) and modified (red) DNA. (A) Shear, (B) stretch, (C) stagger, (D) buckle, (E) propeller, and (F) opening parameters were plotted for each base pair of the unmodified and modified DNA. The shear, stretch, and stagger values are reported in angstroms (Å), and buckle propeller and opening values are reported in degrees. The modified positions are highlighted by underlines. All the parameters were calculated using X3DNA.

in which two water molecules were occupied for ∼68% of the total simulation time. Such alterations in the hydration pattern were also observed at the minor groove of the N^2-lucidin-dG adduct-bearing DNA.[31]

The furfural moiety of the fdG adduct was predicted to be positioned toward the 3′-end of the minor groove of the B-DNA. The dihedral angle C2−N2−C11*−C12* (Ω, Figure S4) connecting the furfural and guanine was found to be ∼130.3° ± 20.2 in the first 115 ns of simulations, after which the Ω was found to be ∼171.5° ± 7.3 until the end of the simulations time. This denotes the conformational fluctuation of the furfural moiety linked with the dG base, which can also lead to disturbances in the minor groove hydrations. The Ω in the crystal structure of fdG-containing modified DNA in complex with DNA Pol IV was found to be 177.2° and 140° (PDB entry: 4Q45, 4Q44).[23] This shows that the observations from the MD simulations are broadly in agreement with those from the crystal structures.

To probe the possibility of intercalation of the furfural moiety into the B-DNA, the base pair opening angle of fdG:dC was studied using umbrella sampling MD simulations following reported protocols.[32,33] The results showed that the opening angle of fdG base is −6° deviated with respect to the dG base in the unmodified DNA (Figure S10 of the Supporting Information). The opening angle of the unmodified DNA was found to be similar to that reported in the literature.[32] The results also indicate that the opening of the fdG base into the major groove of DNA is not energetically favorable (Figure S10B of the Supporting Information). Overall, the data from the umbrella sampling simulations rule out the intercalation mode of the furfuryl moiety into the DNA helix.

In summary, the results from the MD simulation studies of the unmodified and fdG-modified DNAs suggest that the fdG

DOI: 10.1021/acs.joc.5b02341
J. Org. Chem. 2016, 81, 502−511

adduct slightly disturbs W−C H-bonds, stacking interactions with the successive nucleotide, and minor groove hydrations, that results in only minor local structural disturbances in the B-DNA. These results may explain the slight destabilizing effect imparted by the fdG adduct observed in the UV-melting studies.

SUMMARY AND CONCLUSIONS

We have achieved the cost- and labor-efficient synthesis fdG-modified DNAs using phosphoramidite chemistry. Our strategy employing reductive amination is the shortest route to access large quantities of fdG phosphoramidite.[34] Functional assays showed that replicative polymerase Pol I was inefficient to bypass the fdG adduct, while TLS polymerase Pol IV was able to bypass fdG adduct in a proficient and accurate manner. Thermal melting studies of fdG-modified duplex showed slight destabilization by ∼1.6 °C in T_m. Molecular modeling and atomistic MD simulation (300 ns) studies of the unmodified and the fdG-modified DNAs revealed the conformational flexibility induced by the fdG adduct in the B-DNA. Results suggest that there are slight disturbances in W−C H-bonding, stacking interactions, and hydrations in the minor groove around the fdG adduct site. These effects in combination may account for the slightly reduced destabilization of the fdG-modified DNA observed in the UV-melting studies. Overall, there are only marginal changes in the structure of B-DNA due to the presence of the fdG adduct, and hence the DNA repair machinery will find it difficult to detect and repair this adduct.[35] The inability of DNA repair machinery to remove the adduct and its potent ability to block replication by high fidelity DNA polymerases may be responsible for the proto-oncogenic nature of this adduct.

EXPERIMENTAL SECTION

General. All the necessary solvents and chemicals were obtained from commercial sources. Acetonitrile, DCM, pyridine, and DIPEA were dried over calcium hydride, and toluene was dried using calcium chloride. Thin layer chromatography (TLC) was performed using precoated silica gel plates and visualized by UV light (254 nm) or by spraying a solution of 5% (v/v) concd H_2SO_4 in ethanol on silica plates followed by heating. For column chromatography, 100−200 mesh size silica gel was used. ^{1}H (400 and 500 MHz), ^{13}C (100 MHz), and ^{31}P NMR (162 MHz) were recorded on 400 and 500 MHz NMR instruments. Chemical shifts were reported in parts per million (δ) downfield from TMS (0 ppm) 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) or DMSO-d_6 (2.5 ppm) for ^{1}H NMR spectra, and $CDCl_3$ (77.2 ppm) or CD_3OD (49.1 ppm) or DMSO-d_6 (39.5 ppm) for ^{13}C NMR spectra. Multiplicities of ^{1}H NMR spin couplings are mentioned as s for singlet, br s for broad singlet, d for doublet, q for quartet, quint for quintet, td for triplet of doublets, dd for doublet of doublets, ddd for doublet of doublet of doublets, or m for multiplet and overlapping spin systems. Values for the coupling constants (J) are reported in hertz (Hz). High resolution mass spectra (HRMS) were obtained in positive ion electrospray ionization (ESI) mode using a Q-TOF analyzer. Mass spectra of oligonucleotides were obtained by ESI or MALDI in negative mode using TOF analyzers.

3′,5′-*O*-Diacetyl-*N*^{2}-furfuryl-*O*^{6}-(2-(4-nitrophenyl)ethyl)-2′-deoxyguanosine (2). To a solution of compound **1** (1.8 g, 3.46 mmol) in dry toluene (72 mL) was added furfurylamine (0.50 mL, 5.88 mmol). The mixture was degassed for 8 min by passing nitrogen. $Pd(OAc)_2$ (78 mg, 0.35 mmol), (*R*)-BINAP (301 mg, 0.48 mmol) and Cs_2CO_3 (1.7 g, 4.84 mmol) were added to this, and the whole mixture was purged with nitrogen. Reaction was then stirred at 80 °C for 45 h. The reaction mixture was filtered through a Celite pad and washed with ethyl acetate (250 mL). The filtrate was concentrated to give a dark brown residue, which was purified by column chromatography (60% ethyl acetate in petroleum ether) to yield solid brown compound **2** (1.15 g, 57%); R_f = 0.4, (60% ethyl acetate in petroleum ether); mp 102−103 °C; ^{1}H NMR (400 MHz, $CDCl_3$) δ: 8.11 (d, J = 8 Hz, 2H), 7.72 (s, 1H), 7.43 (d, J = 8.0 Hz, 2H), 7.33−7.32 (m, 1H), 6.29 (dd, J = 3.0, 2.0 Hz, 1H), 6.25 (t, J = 7.0 Hz, 1H), 6.20 (d, J = 2.8 Hz, 1H), 5.46−5.44 (m, 1H), 5.40 (t, J = 5.6 Hz, 1H), 4.69 (t, J = 6.8 Hz, 2H), 4.61 (d, J = 5.8 Hz, 2H), 4.40 (q, J = 6.9 Hz, 1H), 4.31−4.24 (m, 2H), 3.23 (t, J = 6.8 Hz, 2H), 3.06 (quint, 7.0 Hz, 1H), 2.49 (ddd, J = 14.1, 6.3, 2.9 Hz, 1H), 2.11 (s, 3H), 2.04 (s, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ: 170.6, 170.3, 160.7, 158.5, 153.5, 152.7, 146.8, 146.1, 141.9, 137.8, 130.0, 123.7, 115.8, 110.5, 106.8, 84.5, 82.2, 74.6, 66.0, 63.9, 63.7, 39.2, 36.5, 35.2, 21.0, 20.8; HRMS: m/z calcd for $C_{27}H_{29}N_6O_9$ $[M + H]^+$ 581.1996, found $[M + H]^+$ 581.1985 (Δm −0.0011, error −1.8 ppm)

5′-*O*-(4,4′-Dimethoxytrityl)-*N*^{2}-furfuryl-2′-deoxyguanosine (5). Nucleoside **2** (1.05 g, 1.8 mmol) was dissolved in 0.5 M DBU in pyridine (37 mL) and was stirred at room temperature for 6 h. Then 50% acetic acid was added to neutralize the reaction, followed by dilution with DCM (150 mL). The organic layer was washed with saturated $NaHCO_3$ (50 mL) and water (2 × 70 mL). The DCM layer was dried over Na_2SO_4, evaporated, and dried. The crude diacetate compound **3** was stirred with 40% $MeNH_2$ in EtOH (31 mL, v/v) for 16 h. Solvent was removed, and the residue was coevaporated with dry pyridine (2 × 5 mL) and was dissolved in the same solvent (9 mL). To this solution was added DMT-Cl (730 mg, 2.16 mmol), and the reaction was stirred at room temperature for 2 h. The reaction mixture was diluted with DCM (100 mL) and washed with saturated $NaHCO_3$ (60 mL) and water (2 × 100 mL). The DCM layer was dried over Na_2SO_4 and evaporated. Crude compound was purified by column chromatography (1% MeOH in DCM + 2% Et_3N) to furnish yellow solid **5** (900 mg, 76% after three steps); R_f = 0.55 (ethyl acetate + 2% Et_3N); mp 75−76 °C; ^{1}H NMR (400 MHz, CD_3OD) δ: 7.83 (br s, 1H), 7.44−7.41 (m, 1H), 7.41−7.36 (m, 2H), 7.30−7.24 (m, 4H), 7.23−7.13 (m, 3H), 6.81−6.73 (m, 4H), 6.35−6.32 (m, 1H), 6.32−6.29 (m, 1H), 6.28−6.24 (m, 1H), 4.65−4.56 (m, 3H), 4.14−4.08 (m, 1H), 3.75 (d, J = 0.9 Hz, 6H), 3.39 (dd, J = 8.0, 4.0 Hz, 1H), 3.24 (dd, J = 8.0, 4.0 Hz, 1H), 2.95−2.86 (m, 1H), 2.38 (ddd, J = 13.6, 6.8, 4.4 Hz, 1H); ^{13}C NMR (100 MHz, DMSO-d_6) δ: 158.0, 158.0, 156.6, 152.1, 151.7, 150.1, 144.9, 142.3, 136.2, 135.6, 135.5, 129.6, 127.7, 126.6, 117.4, 113.0, 110.5, 107.2, 85.8, 85.4, 83.0, 70.8, 64.4, 54.9, 38.7, 37.4; HRMS: m/z calcd for $C_{36}H_{36}N_5O_7$ $[M + H]^+$ 650.2605, found $[M + H]^+$ 650.2615 (Δm −0.001, error −1.5 ppm)

3′-*O*-[2-Cyanoethoxy-(*N*,*N*′-diisopropylamino)phosphino]-5′-*O*-(4,4′-dimethoxytrityl)-*N*^{2}-furfuryl-2′-deoxyguanosine (6). To a solution of compound **5** (200 mg, 0.30 mmol) in DCM (4 mL) were added DIPEA (0.4 mL, 2.45 mmol) and CEP-Cl (181 mg, 0.77 mmol), and the reaction mixture was stirred at room temperature for 40 min. After completion of the reaction, MeOH (0.05 mL) was added and stirred for 15 min. Reaction mixture was diluted with DCM (100 mL), washed with $NaHCO_3$ (3 × 50 mL), dried over Na_2SO_4, and evaporated under reduced pressure. The crude compound was purified by column chromatography (1% MeOH in DCM + 2% Et_3N) to give brown solid **6** (190 mg, 74%); R_f = 0.6, (ethyl acetate + 2% Et_3N); mp 75−76 °C; ^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.51 (d, J = 9.5 Hz, 1H), 7.41 (dd, J = 7.0, 4.6 Hz, 2H), 7.33−7.27 (m, 4H), 7.25−7.14 (m, 4H), 6.82−6.75 (m, 4H), 6.29−6.22 (m, 1H), 6.21−6.12 (m, 2H), 4.71−4.62 (m, 1H), 4.50−4.40 (m, 2H), 4.29−4.22 (m, 1H), 3.74 (d, J = 4.0 Hz, 6H), 3.71−3.54 (m, 4H), 3.39−3.25 (m, 2H), 2.73−2.63 (m, 1H), 2.58 (t, J = 6.4 Hz, 1H), 2.53−2.46 (m, 1H), 2.44 (t, J = 6.4 Hz, 1H), 1.21−1.14 (m, 9H), 1.10 (d, J = 6.7 Hz, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ: 159.5, 158.7, 158.7, 152.64, 152.1, 152.1, 151.4, 151.3, 144.7, 141.9, 135.9, 135.8, 135.8, 135.7, 130.2, 130.1, 128.3, 128.3, 128.0, 127.1, 127.0, 117.6, 117.5, 113.3, 110.5, 110.4, 107.5, 107.5, 86.6, 86.6, 85.8, 85.8, 85.6, 85.5, 84.2, 84.1, 74.6, 74.4, 74.3, 74.1, 64.0, 63.8, 58.6, 58.5, 58.4, 58.3, 55.4, 43.6, 43.5, 43.5, 43.4, 39.7, 38.3, 24.7, 24.7, 24.6, 20.5, 20.5, 20.4, 20.3; ^{31}P NMR (162 MHz, $CDCl_3$) δ: 149.08, 148.67; HRMS: m/z calcd for $C_{45}H_{53}N_7O_8P$ $[M + H]^+$ 850.3693, found $[M + H]^+$ 850.3676 (Δm −0.0017, error −2.1 ppm)

DOI: 10.1021/acs.joc.5b02341
J. Org. Chem. 2016, 81, 502−511

N^2-Furfuryl-2′-deoxyguanosine (4). To a solution of deoxyguanosine **7** (1 g, 3.50 mmol) in dry methanol (50 mL) were added furfural (4.4 mL, 52.5 mmol) and $NaCNBH_3$ (660 mg, 10.5 mmol), and the reaction mixture was heated at 50 °C for 12 h. Then an additional batch of furfural (4.4 mL, 52.5 mmol) and $NaCNBH_3$ (660 mg, 10.5 mmol) was added and stirred further for 24 h. The reaction mixture was filtered through a Celite pad, washed with methanol (250 mL), and evaporated. The crude compound was purified using column chromatography (20% MeOH in DCM) to yield brown solid **4** (562 mg, 46%); R_f = 0.6 (20% MeOH in DCM); mp 143−144 °C; ^{1}H NMR (400 MHz, CD_3OD) δ: 7.99 (br s, 1H), 7.46−7.44 (m, 1H), 6.38−6.36 (m, 1H), 6.35−6.30 (m, 2H), 4.58 (br s, 2H), 4.56−4.51 (m, 1H), 4.01−3.96 (q, J = 4.0 Hz, 1H), 3.76 (dd, J = 12.0, 4.0 Hz, 1H), 3.70 (dd, J = 12.0, 4.0 Hz, 1H), 2.79−2.69 (m, 1H), 2.39 (ddd, J = 14.5, 6.4, 3.5 Hz, 1H); ^{13}C NMR (100 MHz, CD_3OD and $CDCl_3$) δ: 160.2, 152.3, 152.0, 151.7, 142.8, 137.1, 116.3, 111.0, 108.6, 88.0, 84.5, 71.8, 62.7, 40.7, 38.2; HRMS: m/z calcd for $C_{15}H_{17}N_5O_5Na$ $[M + Na]^+$ 370.1127, found $[M + Na]^+$ 370.1143 (Δm +0.0016, error +4.2 ppm)

Oligonucleotide Synthesis and Purification. The unmodified and the fdG-modified DNA sequences were synthesized on 1 μmol scale employing an automated DNA/RNA synthesizer using a controlled pore glass (CPG) solid support. The coupling time used for the unmodified amidites was 2 min and for the modified one was 10 min. 5-Ethylthio-1*H*-tetrazole (ETT) was used as the coupling agent. The deprotection of DNA was performed employing a 1:1 mixture of 41% methylamine in water and 30% aq NH_3 (1 mL, v/v) at 65 °C for 18 min. The supernatant layer was collected, and the CPG beads were washed with a 1:1 mixture of EtOH and water (300 μL). The obtained crude DNA was further purified by 20% denaturing PAGE (7 M urea) at 30 W for 3.5 h using 1X TBE buffer (89 mM each of Tris-HCl and boric acid and 2 mM EDTA, pH 8.3). The gel dimension was 20 × 30 cm, and gel thickness was 1 mm. The gel was visualized under a UV lamp (260 nm), and the desired DNA bands were excised. DNA was extracted from the gel using 15 mL of TEN buffer (10 mM Tris, 1 mM EDTA, and 300 mM NaCl, pH 8.0) by shaking at room temperature for 12 h. Desalting of DNA samples was carried out using a Sep-Pak (C-18) column using standard protocols. The final concentration was measured at 260 nm in a UV−vis spectrophotometer using molar extinction coefficients (ε): for **D1** and **D7** 139500 L m^{-1} cm^{-1}; for **D2** and **D8** 121300 L m^{-1} cm^{-1}; for **D3**, **D4**, and **D5** 163900, 163900, and 167800 L m^{-1} cm^{-1}, respectively, and for **D6** 464200 L m^{-1} cm^{-1}.

Thermal Melting Experiments. The unmodified and the fdG-modified DNA duplex solutions were prepared using 14-mer complementary DNA strands of equimolar concentration (1 μM) in phosphate buffer (20 mM Na_2PO_4, 100 mM NaCl, 0.1 mM EDTA, pH 7.4). Annealing of the strands was performed by heating at 95 °C for 3 min followed by slow cooling to room temperature. In T_m measurements, the change in absorbance at 260 nm was measured with increase in temperature using a UV spectrophotometer coupled with a Peltier-controller. Samples were equilibrated at 25 °C for 10 min before starting the experiment and then heated to 80 °C in an increment of 0.5 °C per minute. To decipher T_m and $\Delta H°$ values, the data of change in absorbance versus increase in temperature were fit to a two-state model with a sloping baseline and analyzed using a Marquadt algorithm for nonlinear curve fitting using KaleidaGraph (version 3.5).[24] The following equations were used to calculate $\Delta S°$ and $\Delta G°$ values: $\Delta S° = \Delta H°/T_m$ and $\Delta G° = \Delta H°(1 - T/T_m)$. All experiments were triplicated, and the T_m values reported are the average of three independent measurements (Table 2). The error propagations for the thermodynamic parameters were derived using reported protocols.[24]

5′-Radiolabeling of DNAs. The 11-mer primer (5′-GGGTCCTAGGA-3′) and 18-mer standard (3′CCCAG-GATCCTGGGGTCT-5′) were radiolabeled using [γ-^{32}P]ATP. The 25 pmol of each DNA strand was mixed with [γ-^{32}P]ATP (20 μCi) in 1X PNK buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, 5 mM DTT, 0.1 mM spermidine, 0.1 mM EDTA, pH 7.6) and T4 polynucleotide kinase (PNK) enzyme (5 U) in a total reaction volume of 10 μL. Reaction mixture was incubated at 37 °C for 1 h and then kept at 70 °C for 3 min to deactivate the PNK enzyme. The purification of radiolabeled DNAs were carried out using QIAquick nucleotide removal kit.

Primer Extension Studies. For these studies, Pol I and Pol IV from *E. coli* have been used. Pol IV was purified using published procedure.[36] The primer extension reactions were carried out at 37 °C for 2 h using a 5′-FAM-labeled primer. The reaction mixture contained 40 nM of DNAs, 50 μg/mL of BSA, 0.1 mM of ammonium sulfate, 2.5 mM $MgCl_2$, 5 μM of dCTP, 4 μL of 5X assay buffer (125 mM Tris-Cl, 5 mM DTT, pH = 8.0), and 4 nM of Pol I or Pol IV in total 20 μL of volume. The reactions were terminated using stop solution (80% formamide, 1 mg/mL xylene cyanol, 1 mg/mL bromophenol blue, and 20 mM EDTA). Then the reaction mixtures were incubated at 95 °C for 2 min and transferred to ice for 10 min. Finally, the reactions products (15 μL) were loaded on to 20% polyacrylamide gel (8 M urea). The resolved reaction products were observed by excitation at 488 nm, and the bands were recorded using a fluorescence image scanner.

The running start experiments using Pol IV were carried out using 10 nM of DNAs (1:1 ratio of primer and template) containing traces of 5′-radiolabeled primer. The mixture was annealed, and the reactions were performed using same reaction conditions used for primer extension assay employing a mixture of dNTPs (250 μM). In reactions using KF exo$^-$, the reaction mixture contained 10 nM DNAs, 250 μM concentration of dNTPs, and 1X polymerase buffer (50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9) in a total reaction volume of 20 μL. The reactions were terminated as mentioned previously. The desired products were loaded on a 20% denaturing PAGE (7 M urea). The autoradiograms were generated using a phosphorimager, and the bands were quantified by using ImageQuantTL software.

Molecular Modeling Studies. The force field parameters and RESP charges were developed for the fdG adduct from the energy-optimized structure using Gaussian 09.[29,37] The B-form structure of the unmodified DNA sequence (**D7−D8**, Table 2) was built using nucleic acid builder module in AMBER 14.[28] The fdG-modified ds DNA sequence (**D1−D8**, Table 2) was used for the MD simulations studies. Using xleap module in AMBER 14, dG in the unmodified DNA was replaced by the fdG nucleotide to generate the modified DNA, and AMBER parmbsc0 force field was used for the DNA.[38] The unmodified and modified structures were initially minimized using 5000 steps of steepest descent minimization. Sodium ions were used to neutralize the backbone of the DNA, and an octahedral TIP3P water box was extended up to 8 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 100 ps with the restraint force for 10 kcal/mol Å^2. The equilibration was carried for 750 ps in 1 atm pressure at a constant temperature of 300 K. Minimization and equilibration was carried using SANDER in AMBER 14. After equilibration, the MD simulations of the unmodified and fdG-modified duplex DNA was carried for 300 ns each using PMEMD in the GPU accelerated version of AMBER 14. MD trajectories were saved for every 2 ps. The RMSD, hydration in the major and the minor groove, stacking interactions, W−C H-bond distances, and occupancies were calculated using CPPTRAJ module in AMBER 14. Six local base pair parameters, torsion angles of the DNA backbone, and the intrastrand phosphate distances were calculated using X3DNA.[39]

The final structures of the unmodified and fdG-modified DNA obtained from the 300 ns of unrestrained MD simulations were taken as the initial structure to calculate the base pair opening parameters in the umbrella sampling simulations. The opening of the dG:dC and the fdG:dC base pairs was probed. The protocols for generating the angle constraint and the umbrella sampling simulation protocol were adapted from the method developed by Lavery and co-workers.[32,33] Simulations were carried out with increasing and decreasing values of opening angle (θ). The θ was increased by 4°, and a force constant of 0.05 kcal mol^{-1} deg^{-2} was added to the structure. The conformation developed for the each opening angle was solvated, minimized, and

DOI: 10.1021/acs.joc.5b02341
J. Org. Chem. 2016, 81, 502−511

equilibrated for 300 ps. A further 1 ns of MD simulations was carried out to sample the fluctuation of θ. The conformation obtained from the end of the simulations for a specific θ was used as the starting structure to generate the next conformation of θ. In total, 50 conformations of θ were generated to study the opening angle ($\pm 100°$), which results in 50 ns of simulations. The sampling of all the θ values was verified using the adjacent overlapping in the histogram, and then the potential of mean force or the free energy associated with the base opening was constructed using the WHAM algorithm using the histogram.[40] MD trajectories were visualized using UCSF Chimera. Figures were rendered using PyMOL.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.joc.5b02341.

UV melting curves of the unmodified and the fdG-modified DNAs. Nucleotide incorporation using Pol IV DNA polymerase. Nucleotide incorporation using Pol I DNA polymerase exo$^+$ (KF exo$^+$). Running start experiments using Pol IV DNA polymerase. Running start experiments using the Pol I DNA polymerase exo$^-$ (KF exo$^-$). Energy-optimized structure of the fdG nucleotide. Time-dependent RMSD graphs of the unmodified and the fdG-modified DNAs. Intrastrand phosphate distances of the unmodified and the fdG-modified DNAs. W−C H-bond distances in the unmodified and in the fdG-modified DNAs. Umbrella sampling MD simulations of the base pair opening. Percentage occupancies of stacking interactions in the unmodified and fdG-modified nucleotide steps. Watson−Crick H-bond occupancies of the unmodified and the fdG-modified DNAs. Major and minor groove hydration of the unmodified and the fdG-modified DNAs. ^{1}H, ^{13}C, and ^{31}P NMR spectra for all new compounds. MALDI or ESI-MS spectra for fdG-modified DNAs (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

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

Author Contributions

#P.P.G and K.R.G contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work is financially supported by grants from Department of Biotechnology (DBT)-Government of India (grant no.: BT/PR8265/BRB/10/1228/2013). We are thankful to Dr. Claudia Höbartner (MPIbpc-Göttingen), Prof. K.V.R Chary, and Ms. Gitanjali A. Dhotre (TIFR-Mumbai) and the central facility supported by IRCC-IIT Bombay for providing ESI/MALDI spectra of the modified DNAs. Computer Centre, IIT Bombay, and NPSF-CDAC, Pune, are gratefully acknowledged for providing high performance computing facilities. P.P.G. and J.K. thank the Council of Scientific and Industrial Research (CSIR), and S.H. thanks the Department of Atomic Energy-Board of Research in Nuclear Sciences (DAE-BRNS), and IRCC-IIT Bombay, for the Ph.D. fellowships.

■ REFERENCES

(1) Garner, C. R. *Mutat. Res., Fundam. Mol. Mech. Mutagen.* **1998**, *402*, 67−75.
(2) Dipple, A. *Carcinogenesis* **1995**, *16*, 437−441.
(3) Washington, M. T.; Carlson, K. D.; Freudenthal, B. D.; Pryor, J. M. *Biochim. Biophys. Acta, Proteins Proteomics* **2010**, *1804*, 1113−1123.
(4) Geacintov, N. E.; Broyde, S. In *The Chemical Biology of DNA Damage*; Wiley-VCH Verlag GmbH & Co. KGaA: New York, 2010; pp 1−20.
(5) Sale, J. E.; Lehmann, A. R.; Woodgate, R. *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 141−152.
(6) Jarosz, D. F.; Godoy, V. G.; Delaney, J. C.; Essigmann, J. M.; Walker, G. C. *Nature* **2006**, *439*, 225−228.
(7) Jarosz, D. F.; Cohen, S. E.; Delaney, J. C.; Essigmann, J. M.; Walker, G. C. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 21137−21142.
(8) Cheng, S. C.; Hilton, B. D.; Roman, J. M.; Dipple, A. *Chem. Res. Toxicol.* **1989**, *2*, 334−340.
(9) Pischetsrieder, M.; Seidel, W.; Münch, G.; Schinzel, R. *Biochem. Biophys. Res. Commun.* **1999**, *264*, 544−549.
(10) Larsen, J. C. *EFSA J.* **2011**, *9*, 1−44.
(11) Maga, J. A.; Katz, I. *Crit. Rev. Food Sci. Nutr.* **1979**, *11*, 355−400.
(12) Monien, B. H.; Herrmann, K.; Florian, S.; Glatt, H. *Carcinogenesis* **2011**, *32*, 1533−1539.
(13) Sachse, B.; Meinl, W.; Glatt, H.; Monien, B. H. *Carcinogenesis* **2014**, *35*, 2339−2345.
(14) Høie, A. H.; Monien, B. H.; Sakhi, A. K.; Glatt, H.; Hjertholm, H.; Husøy, T. *Mutagenesis* **2015**, *30*, 643−649.
(15) Monien, B. H.; Schumacher, F.; Herrmann, K.; Glatt, H.; Turesky, R. J.; Chesné, C. *Anal. Chem.* **2015**, *87*, 641−648.
(16) Monien, B. H.; Engst, W.; Barknowitz, G.; Seidel, A.; Glatt, H. *Chem. Res. Toxicol.* **2012**, *25*, 1484−1492.
(17) Whiteway, J.; Koziarz, P.; Veall, J.; Sandhu, N.; Kumar, P.; Hoecher, B.; Lambert, I. B. J. *J. Bacteriol.* **1998**, *180*, 5529−5539.
(18) Lu, C.; McCalla, D. R.; Bryant, D. W. *Mutat. Res., Genet. Toxicol. Test.* **1979**, *67*, 133−144.
(19) DeCorte, B. L.; Tsarouhtsis, D.; Kuchimanchi, S.; Cooper, M. D.; Horton, P.; Harris, C. M.; Harris, T. M. *Chem. Res. Toxicol.* **1996**, *9*, 630−637.
(20) Dai, Q.; Ran, C.; Harvey, R. G. *Org. Lett.* **2005**, *7*, 999−1002.
(21) Eritja, R.; Robles, J.; Aviñó, A.; Alberico, F.; Pedroso, E. *Tetrahedron* **1992**, *48*, 4171−4182.
(22) Honcharenko, D.; Varghese, O. P.; Plashkevych, O.; Barman, J.; Chattopadhyaya, J. *J. Org. Chem.* **2006**, *71*, 299−314.
(23) Kottur, J.; Sharma, A.; Gore; Kiran, R.; Narayanan, N.; Samanta, B.; Pradeepkumar, P. I.; Nair; Deepak, T. *Structure* **2015**, *23*, 56−67.
(24) Siegfried, N. A.; Metzger, S. L.; Bevilacqua, P. C. *Biochemistry* **2007**, *46*, 172−181.
(25) Johnson, R. E.; Washington, M. T.; Haracska, L.; Prakash, S.; Prakash, L. *Nature* **2000**, *406*, 1015−1019.
(26) Woodgate, R. *Mutat. Res., DNA Repair* **2001**, *485*, 83−92.
(27) Shachar, S.; Ziv, O.; Avkin, S.; Adar, S.; Wittschieben, J.; Reißner, T.; Chaney, S.; Friedberg, E. C.; Wang, Z.; Carell, T.; Geacintov, N.; Livneh, Z. *EMBO J.* **2009**, *28*, 383−393.
(28) Case, D. *AMERR 14*; University of California, San Francisco, 2014.
(29) Frisch, M. *Guassian 09, revision A.02*; Guassian Inc., Wallingford, CT, 2009.
(30) Saenger, W. In *Principles of Nucleic Acid Structure*; Springer: New York, 1984; pp 51−104.
(31) Ghodke, P. P.; Harikrishna, S.; Pradeepkumar, P. I. *J. Org. Chem.* **2015**, *80*, 2128−2138.
(32) Giudice, E.; Várnai, P.; Lavery, R. *Nucleic Acids Res.* **2003**, *31*, 1434−1443.
(33) Várnai, P.; Lavery, R. *J. Am. Chem. Soc.* **2002**, *124*, 7272−7273.
(34) Gore, K. R.; Nair, D. T.; Pradeepkumar, P. I. Synthesis of N^2-furfuryl deoxyguanosine phosphoramidite and modified oligonucleotides. Indian Patent Application 0454/MUM/2013; Feb 18, 2013.
(35) Marteijn, J. A.; Lans, H.; Vermeulen, W.; Hoeijmakers, J. H. J. *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 465−481.

DOI: 10.1021/acs.joc.5b02341
J. Org. Chem. 2016, 81, 502−511

(36) Sharma, A.; Kottur, J.; Narayanan, N.; Nair, D. T. *Nucleic Acids Res.* **2013**, *41*, 5104−5114.
(37) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269−10280.
(38) Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E., III; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817−3829.
(39) Lu, X.-J.; Olson, W. K. *Nat. Protoc.* **2008**, *3*, 1213−1227.
(40) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 1011−1021.

DOI: 10.1021/acs.joc.5b02341
J. Org. Chem. 2016, 81, 502−511

Supporting Information

The N^2-Furfuryl-deoxyguanosine (fdG) Adduct Does Not Alter the Structure of B-DNA

Pratibha P. Ghodke,[†] Kiran R. Gore,[†,‖] S. Harikrishna,[†] Biswajit Samanta,[†] Jithesh Kottur,[§,‡] Deepak T. Nair*[,§,⊥] and P. I. Pradeepkumar*[,†]

[†]*Department of Chemistry, Indian Institute of Technology Bombay, Mumbai-400076, India*

[‖]*Department of Chemistry, University of Mumbai, Mumbai-400098, India*

[§]*Regional Centre for Biotechnology, NCR Biotech Science Cluster, Third Milestone, Faridabad-Gurgaon Expressway, Faridabad-121001, India*

[‡]*Manipal University, Manipal-576104, India*

[⊥]*National Centre for Biological Sciences (NCBS-TIFR), GKVK Campus, Bellary Road, Bangalore-560065, India*

Email: deepak@rcb.res.in or pradeep@chem.iitb.ac.in

TABLE OF CONTENTS

UV melting curves of the unmodified and the fdG modified DNAs

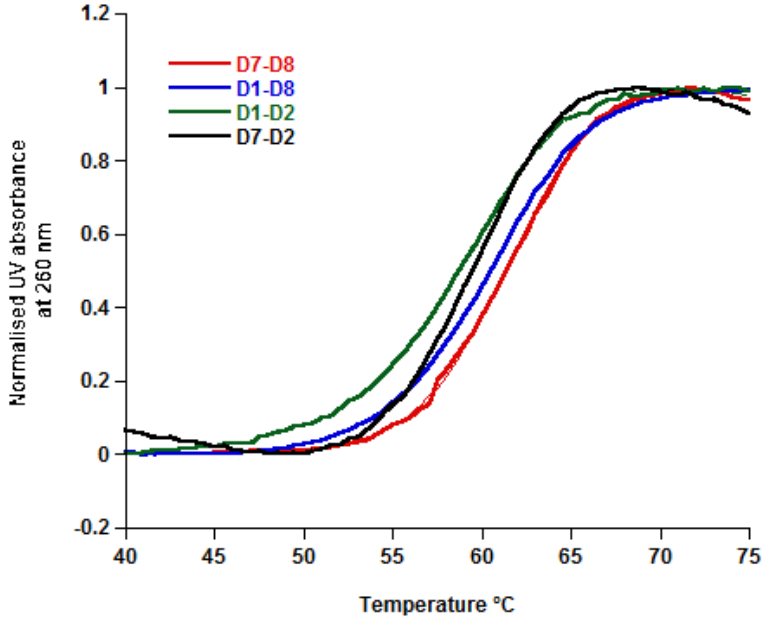


Figure S1. UV monitored thermal dissociation of the unmodified and the fdG modified DNA duplexes (1 µM) were obtained using buffer 100 mM NaCl, 20 mM Na_2PO_4, 0.1 mM EDTA, pH 7.4. Denaturation curves were obtained at 260 nm at a heating rate of 0.5 °C/min. The T_m values reported are average of three measurements and shown in Table 2 (Main Text)

Nucleotide incorporation using Pol IV DNA polymerase

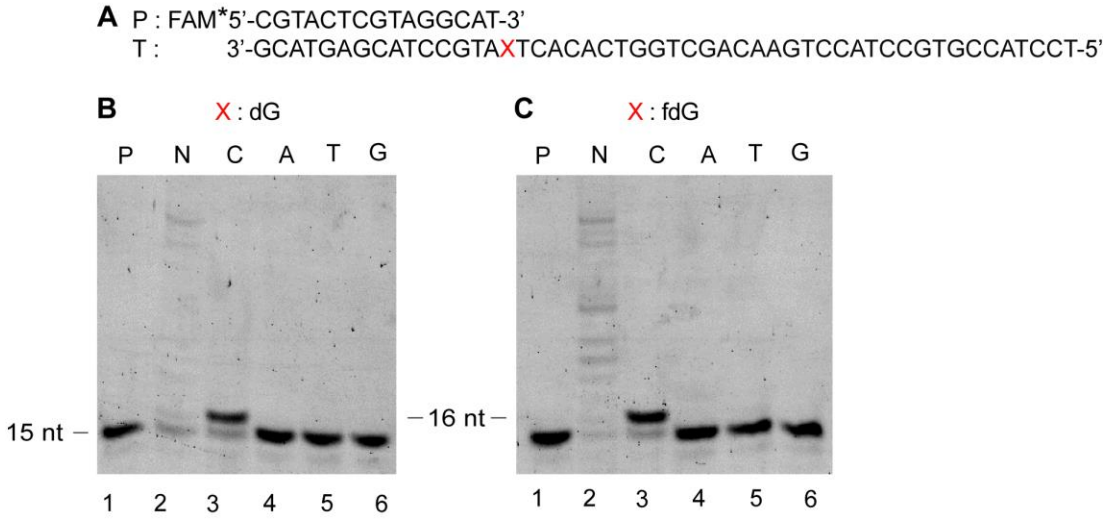


Figure S2. PAGE (20%, 8 M urea) of primer extension reactions with *E. coli* DNA Polymerase IV (A) Complete sequence of template and primer; (B) X = dG and (C) X = fdG; Lane 1: primer only; lane 2: reaction with mixture of all dNTPs; lanes 3 to 6: reactions with dCTP or dATP or dTTP or dGTP.

Nucleotide incorporation using Pol I DNA polymerase exo$^+$ (KF exo$^+$)

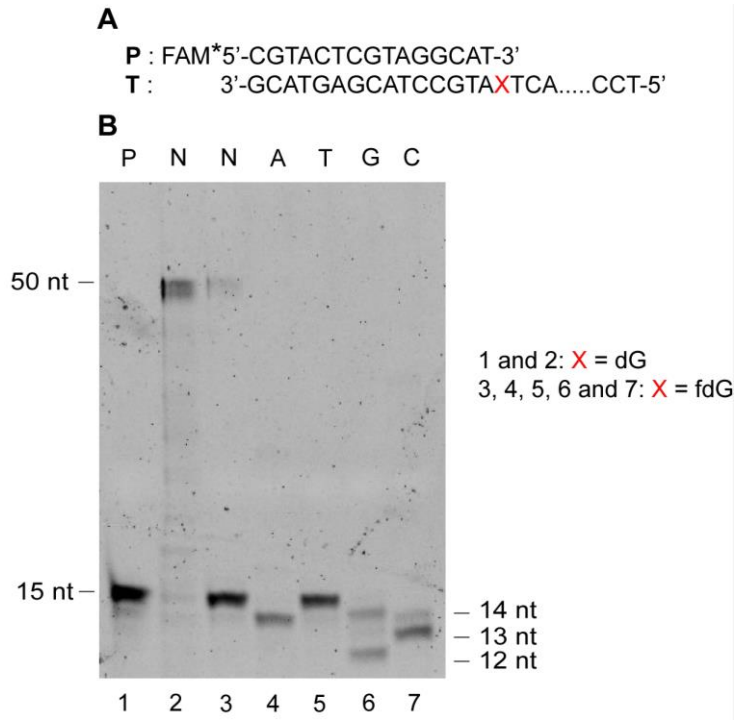


Figure S3. PAGE (20%, 8 M urea) of primer extension reactions with Klenow fragment of *E. coli* DNA Polymerase I (KF exo$^+$) (A) Partial sequence of the template and the primer; and (B) Lanes 2 and 3 correspond to reactions carried out with undamaged dG and fdG templates respectively. Lane 1: primer only; lane 2: reaction with mixture of all dNTPs (positive control), lane 3: reaction with mixture of all dNTPs and lanes 4 to 7: reactions with dATP or dTTP or dGTP or dCTP using fdG modified template. There is ~ 5.5% full length product formation for fdG template (Lane 3). The gel shows that the fdG adduct blocks DNA synthesis by Klenow fragment resulting in adventitious trimming of the primer by the proofreading activity (3' to 5' exonuclease activity) of the enzyme.

Running start experiments using Pol IV DNA polymerase

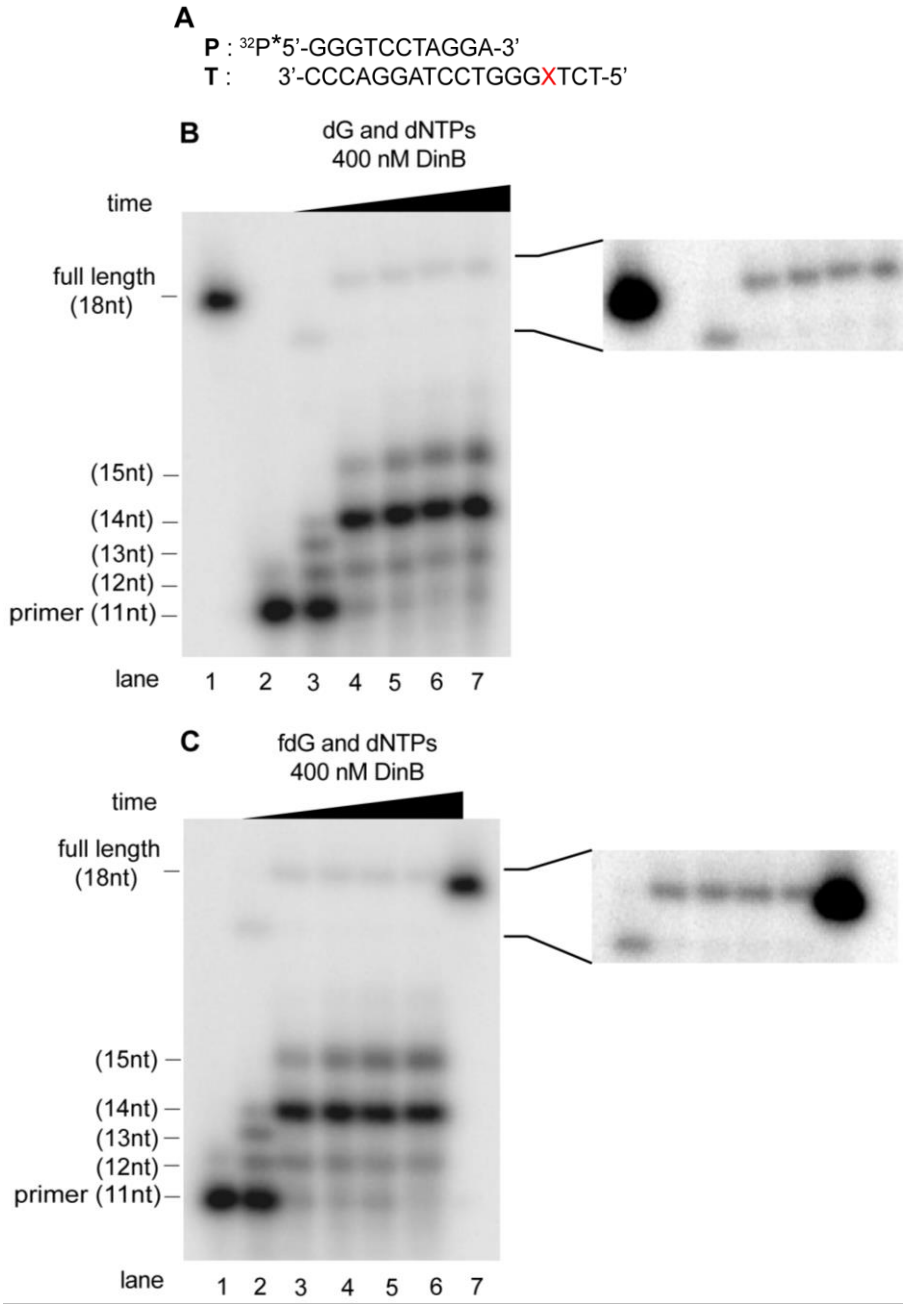


Figure S4. PAGE (20%, 7 M urea) of the ruining start experiments with *E. coli* DNA Polymerase IV (A) Complete sequence of the template and the primer; (B) X = dG, Lane 1: 18 mer size standard, lane 2: primer only, and lanes 3 to 7: reactions with undamaged dG template; and (C) X = fdG, Lane 1: primer only and lanes 3 to 6: reactions with fdG template, lane 7: 18 mer size standard. All the reactions were carried out at 37 °C using mixture of all dNTPs (250 μM) at different time course from 0, 15, 30, 60 and 120 min. Zoom regions (400%) are shown for figure B and C at the full length product sites. After 2 h, ~ 0.8 % of full length product formation was observed with undamaged template while the same for fdG modified template was ~ 2%.

Running start experiments using the Pol I DNA polymerase exo⁻ (KF exo⁻)

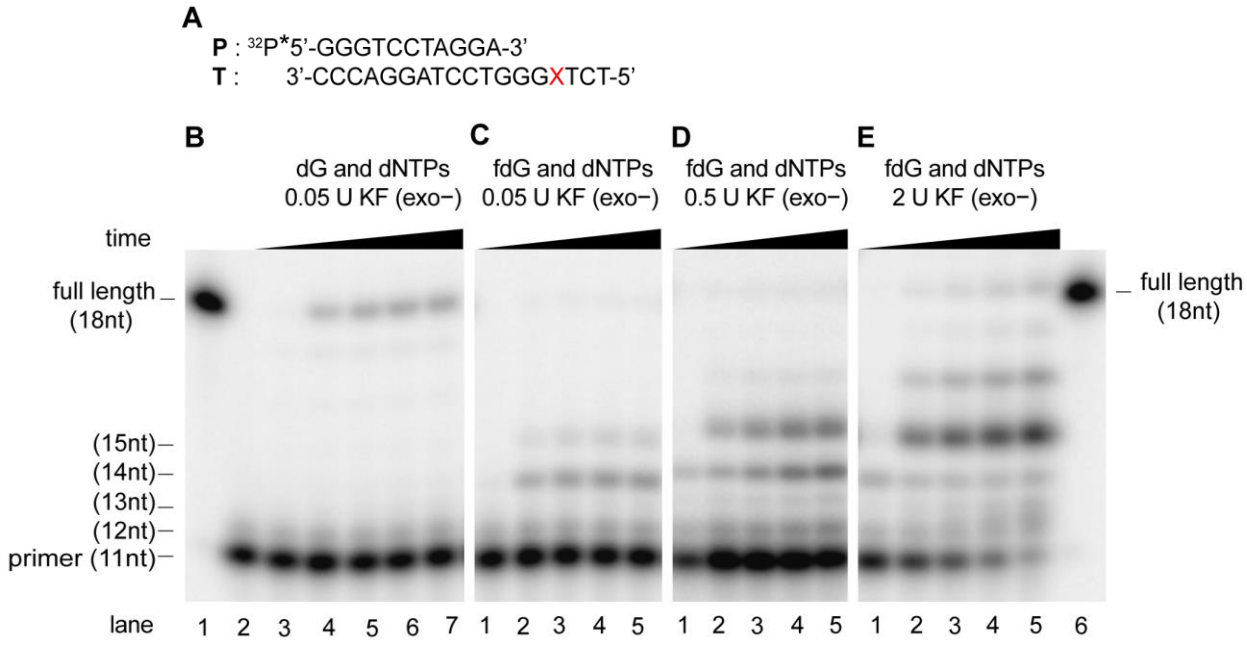
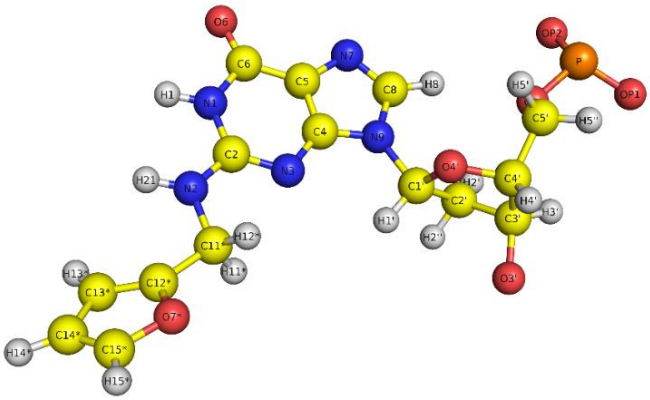


Figure S5. PAGE (20%, 7 M urea) of the full length extension reactions with Klenow fragment of *E. coli* DNA Polymerase I (KF exo⁻). (A) Complete sequence of the template and the primer; (B) Lane 1: 18 mer size standard; lane 2: primer only; lanes 3 to 7: reactions with undamaged dG template using 0.05 U of KF exo⁻; (C) Lanes 1 to 5: reactions with fdG template using 0.05 U of KF exo⁻; (D) Lanes 1 to 5 correspond to reactions carried out with fdG template using 0.5 U of KF exo⁻; and (E) Lanes 1 to 5 correspond to reactions carried out with fdG template using 2 U of KF exo⁻; Lane 6: 18 mer size standard. All the reactions were carried out at 37 °C with the mixture of dNTPs (250 μM) at different time course from 0, 15, 30, 60, 120 min. With the use of 0.05 U after 2 h, undamaged template yielded ~ 6 % of full length product while there was no significant full length product formation (≤ 0.2 %) for the fdG modified template. Use of 0.5 and 2 U of KF exo⁻ with the fdG modified template yielded ~ 0.4 and ~ 3.6% of full length products respectively.

Energy optimized structure of the fdG nucleotide



Atom name	RESP charge	Atom name	RESP charge
OP1	−0.56847	N7	−0.61929
P	1.12900	C5	−0.01664
OP2	−0.56847	C6	0.80052
O5'	−0.21432	O6	−0.64330
C5'	−0.13269	N1	−0.94906
H5'	0.16151	H1	0.44694
H5"	0.16151	C2	0.95026
C4'	−0.15541	N2	−0.60560
H4'	0.15364	H21	0.32470
C3'	0.63856	C11*	−0.16850
H3'	−0.12296	H11*	0.14256
C2'	−0.44110	H12*	0.14256
H2'	0.08308	C12*	0.35183
H2"	0.08308	C13*	−0.29941
O3'	−0.97220	H13*	0.18449
O4'	−0.46502	C14*	−0.24609
C1'	0.47929	H14*	0.19367
H1'	0.11705	C15*	−0.10194
N9	−0.34933	H15*	0.20439
C4	0.39170	O7*	−0.23982
N3	−0.70582		
C8	0.37634		
H8	0.06872		

Figure S6. Energy optimized geometry and calculated RESP charges for the fdG using HF/6-31G(d) basis set in Gaussian 09 program.[1-3]

Time dependent RMSD graphs of the unmodified and the fdG modified DNAs

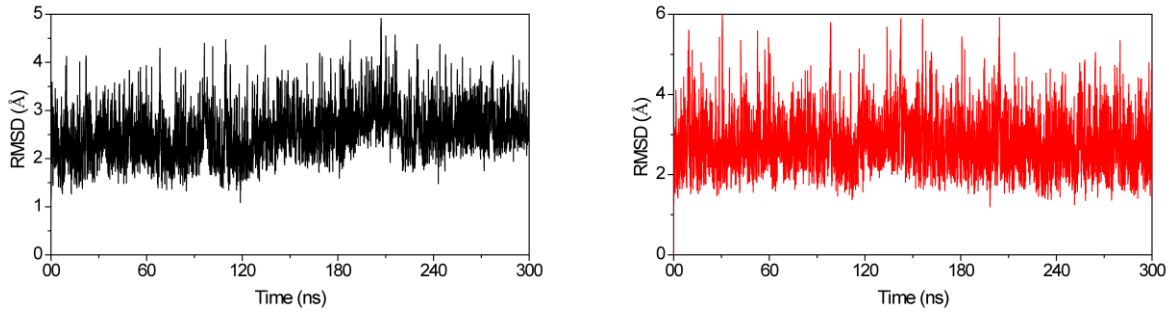


Figure S7. Root mean square deviation graphs of the trajectories obtained from the MD simulations of the unmodified (black) and fdG modified (red) DNA. RMSD values from the MD trajectories were calculated with respect to the corresponding initial structure using CPPTRAJ module in AMBER 14.

Intrastrand phosphate distances of the unmodified and the fdG modified DNAs

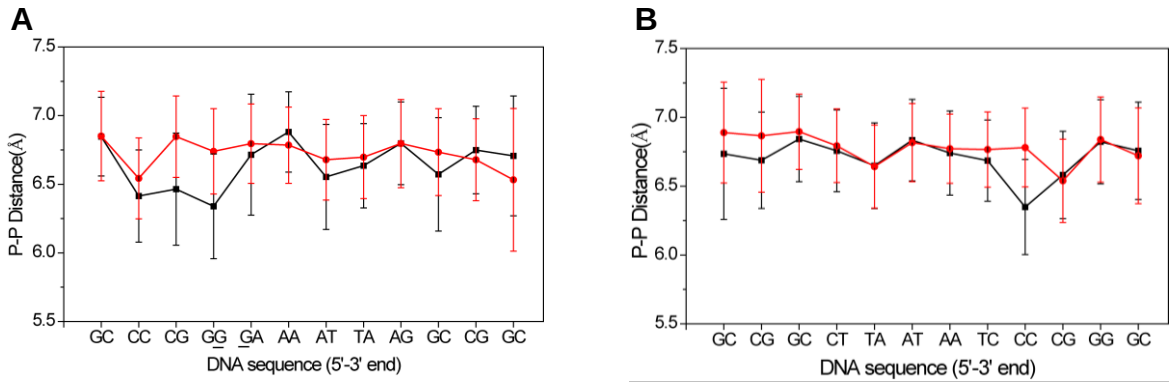


Figure S8. Averaged intrastrand phosphate distance between the successive nucleotides in the modified and unmodified DNA during 300 ns of MD simulations. (A) Intrastrand phosphate distance between the successive nucleotides (5'-3') of the unmodified (**D8**, black) and the fdG modified strand (**D1**, red). The modified site containing fdG is highlighted by underline; (B) Intrastrand phosphate distance between the successive nucleotides in the strand (5'-3') of the unmodified and the fdG modified DNA. The P-P atomic distances were calculated using CPPTRAJ module in AMBER 14.

W-C H-bond distances in the unmodified and in the fdG modified DNAs

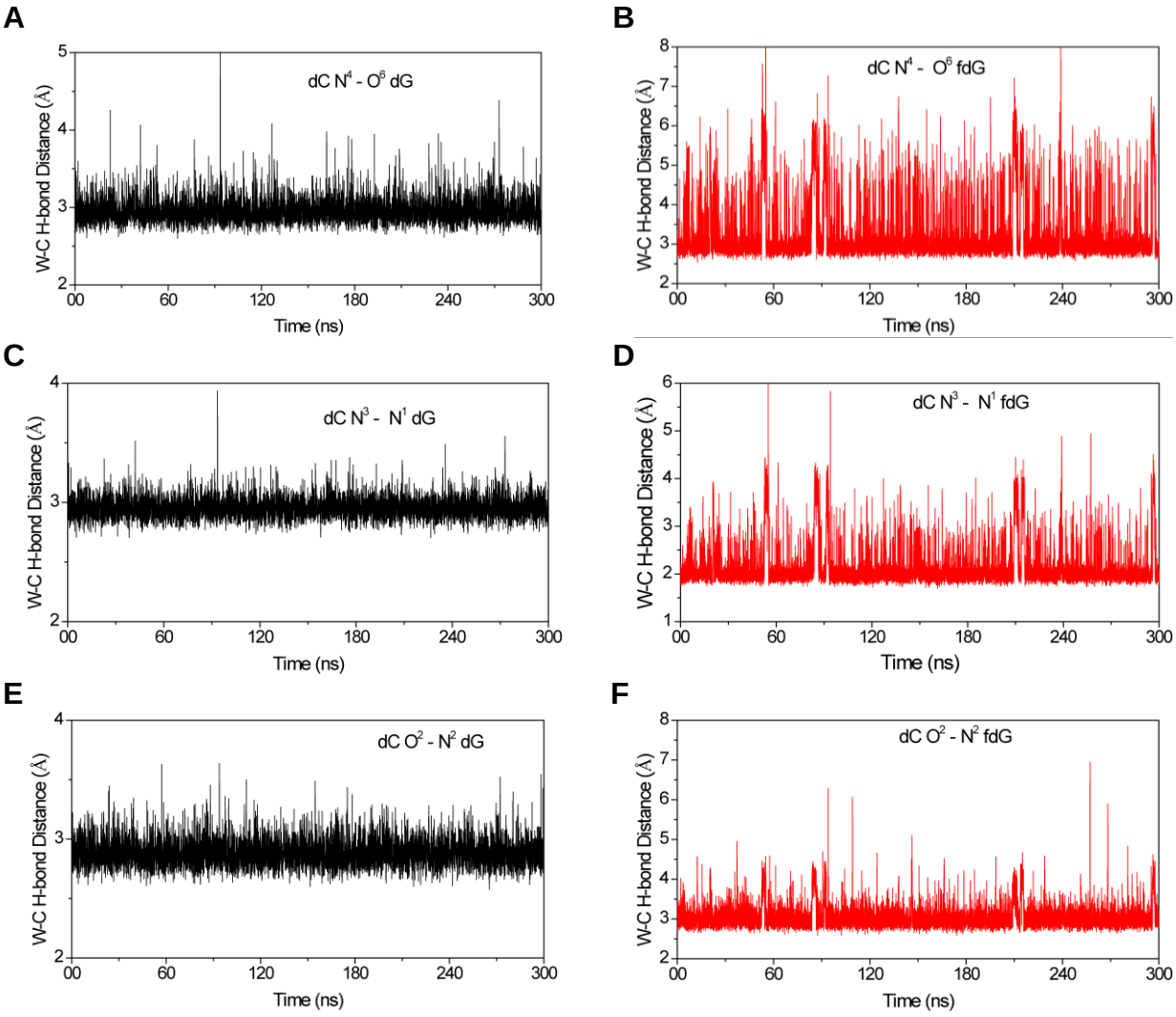


Figure S9. W-C H-bond distances between dG:dC in the unmodified and fdG:dC in the modified DNAs from the 300 ns MD simulations trajectories. (A) and (B) show the distance between $N^4 \cdots O^6$ of dC:dG/fdG base pair in the unmodified and in the fdG modified DNAs; (C) and (D) show the distance between $N^1 \cdots N^3$ of dC:dG/fdG base pair in the unmodified and in the fdG modified DNAs; (E) and (F) show the distance between $O^2 \cdots N^2$ of dC:dG/fdG of base pair in the unmodified and in the fdG modified DNAs. Black lines indicate the distance for the unmodified DNA and the red lines indicate the distance for the fdG modified DNA. All the distances were calculated using CPPTRAJ module in AMBER 14.

Umbrella sampling MD simulations of the base pair opening

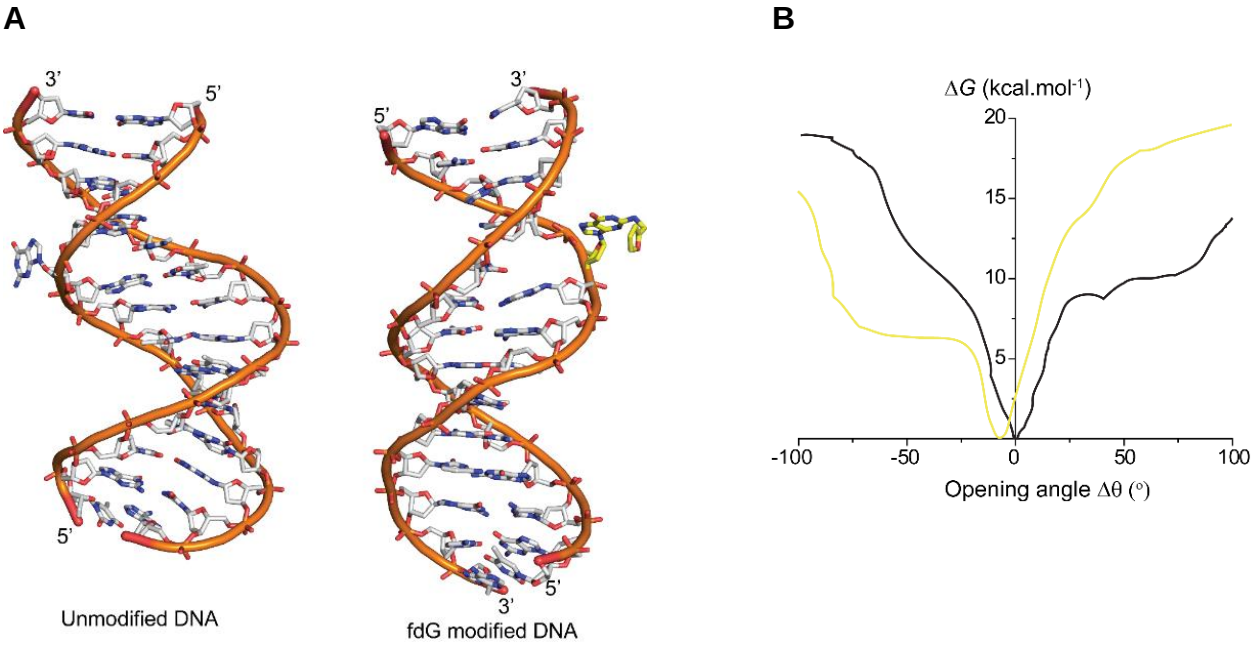


Figure S10. Umbrella sampling MD simulations snapshot of the unmodified and fdG modified DNAs. (A) The snapshots show the orientation of the dG and fdG bases in the unmodified and modified DNA, respectively. The fdG modification is highlighted in yellow color; (B) Free energy (ΔG) curve as a function of relative opening angle ($\Delta\theta$) for the dG:dC (black) and fdG:dC (yellow) bases in the unmodified and modified DNA respectively. Negative and positive openings correspond to the rotation of the nucleotides into the minor and major groove of the DNA, respectively.

Percentage occupancies of stacking interactions in the unmodified and the fdG modified nucleotide steps

Sequence	Unmodified DNA	Modified DNA
dG-**dG**	99%	84%
dG-dA	99%	68%
dA-dA	99%	79%

Table S1. Percentage occupancies of stacking interactions for the unmodified and the fdG modified nucleotide steps in the duplex DNAs during 300 ns of the MD simulations. The modified site is highlighted in red color and bold. The percentage occupancies of the stacking interactions were calculated using CPPTRAJ module in AMBER 14.

Watson-Crick H-bond occupancies of the unmodified and the fdG modified DNAs

Sequence	Unmodified DNA	Modified DNA
dC:dG	99%	96%
dC:**dG**	99%	80%
dT:dA	99%	90%

Table S2. The Watson-Crick H-bond occupancies of the unmodified and the fdG modified DNAs during 300 ns of the MD simulations. The occupancies were calculated in and around the fdG modified site in the duplex DNA. The modified site is highlighted in red color and bold. The percentages of H-bond occupancies were calculated using CPPTRAJ module in AMBER 14.

Major and minor groove hydration of the unmodified and the fdG modified DNAs

Sequence	Unmodified DNA		Modified DNA	
	Major groove	Minor groove	Major groove	Minor groove
dC:dG	3 (96%)	3 (94%)	3 (89%)	2 (90%)
dC:**dG**	3 (95%)	2 (97%)	2 (69%)	1 (85%)
dT:dA	2 (96%)	3 (96%)	2 (81%)	2 (60%)

Table S3. The hydration of the unmodified and the fdG modified duplex DNA during 300 ns of the MD simulations. The number water molecules in the major and minor grooves of the DNAs are shown, and the percentage occupancies of the water molecules are shown in brackets. The modified site is highlighted in red color and bold. Calculations were carried out using CPPTRAJ module in AMBER 14.

References

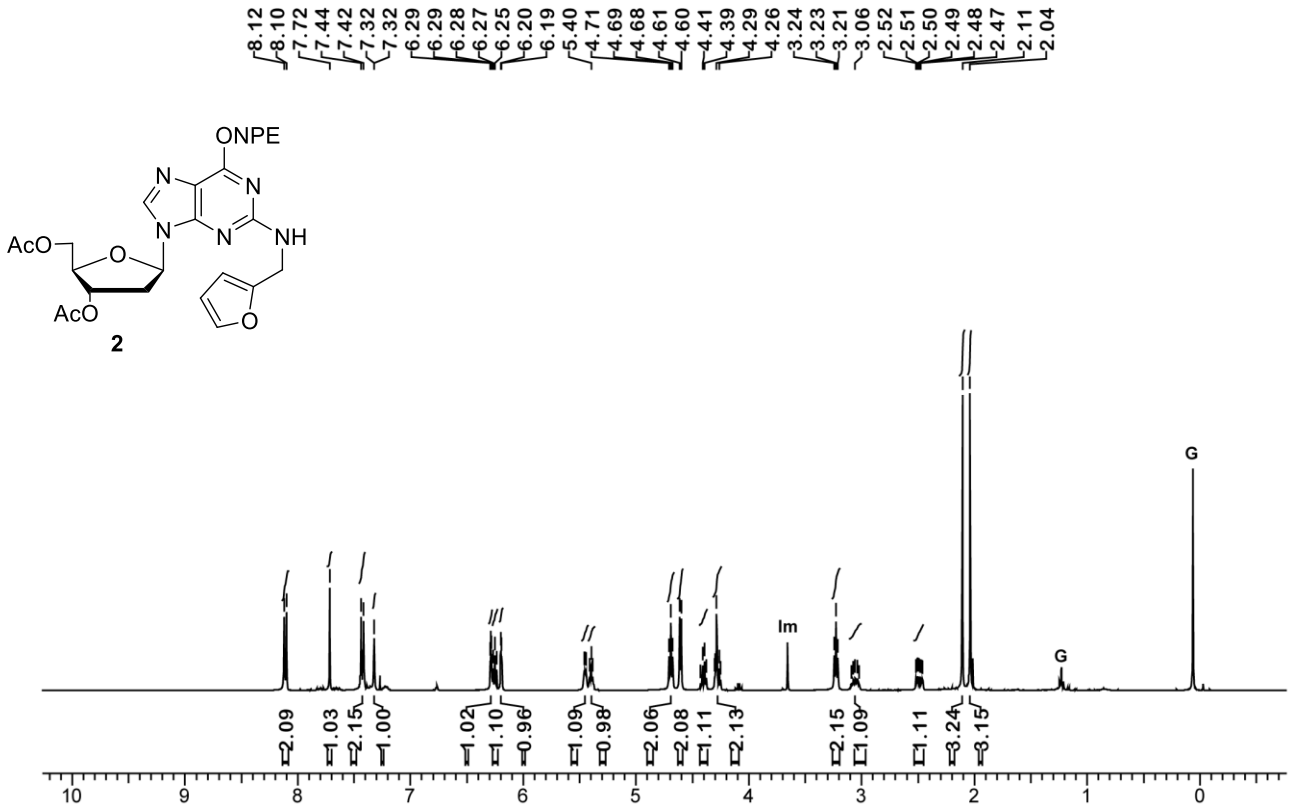
(1) Frisch, M. J.; Gaussian 09, Revision A.02: Gaussian, Inc.: Wallingford, CT, USA, **2009**.

(2) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, 25, 1157–1174.

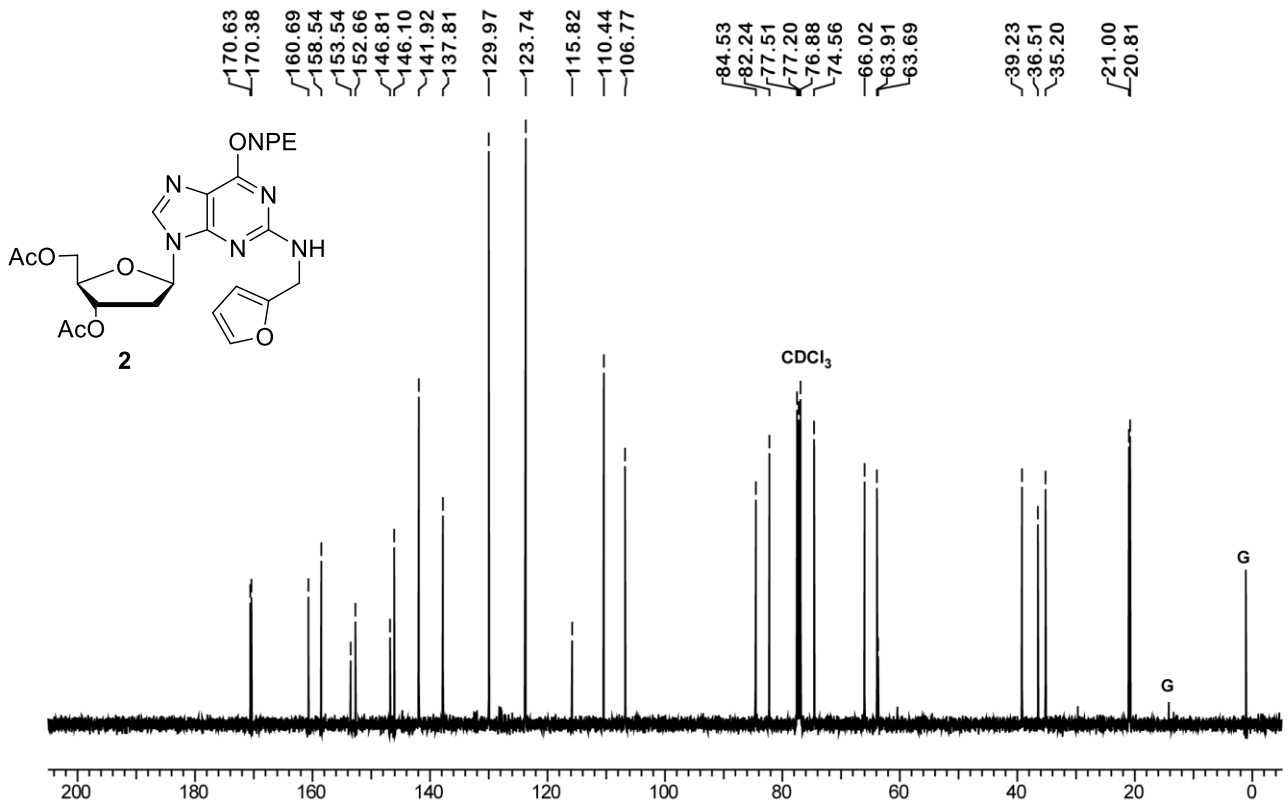
(3) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc*. **1995**, *117*, 5179–5197.

<u>NMR spectra (^{1}H, ^{13}C, & ^{31}P)</u> (G- Grease, Im- Impurity)

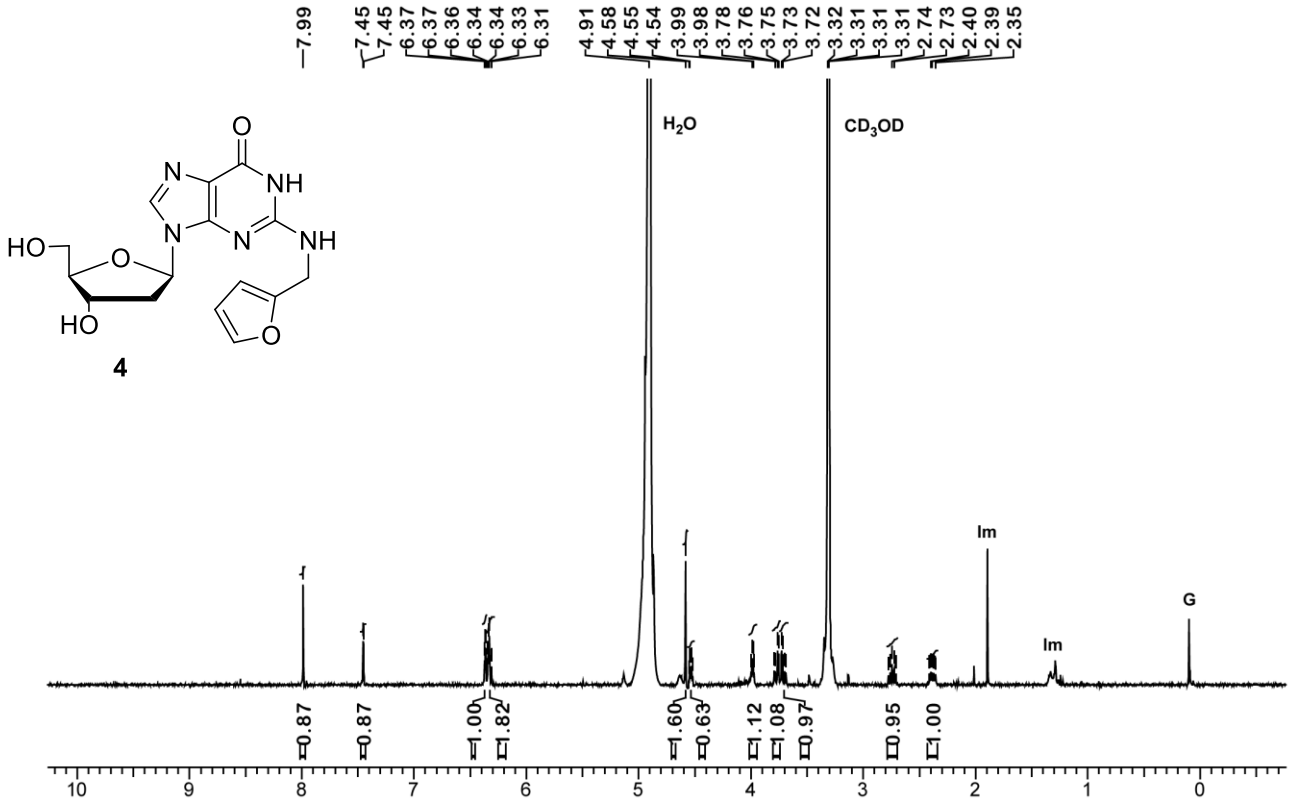
^{1}H NMR for Compound **2**



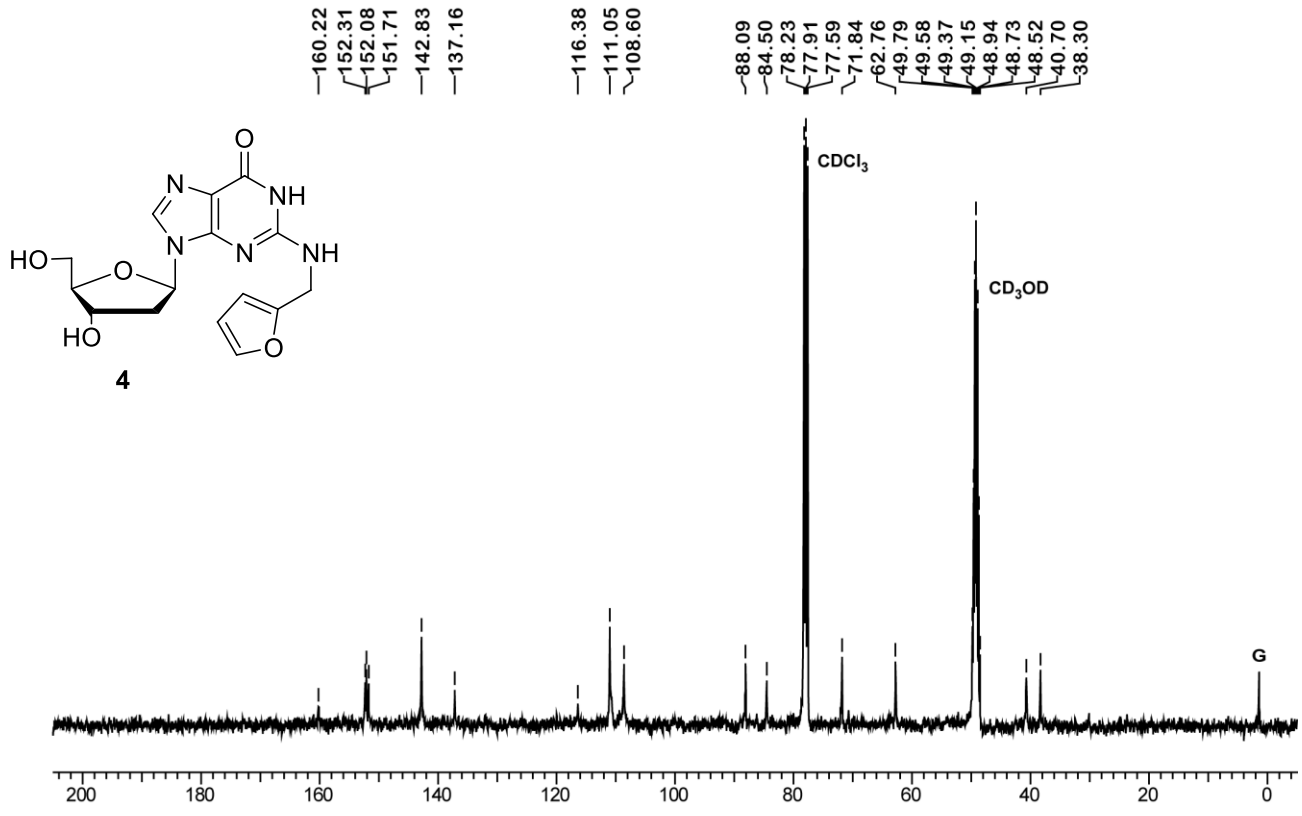
^{13}C NMR for Compound **2**



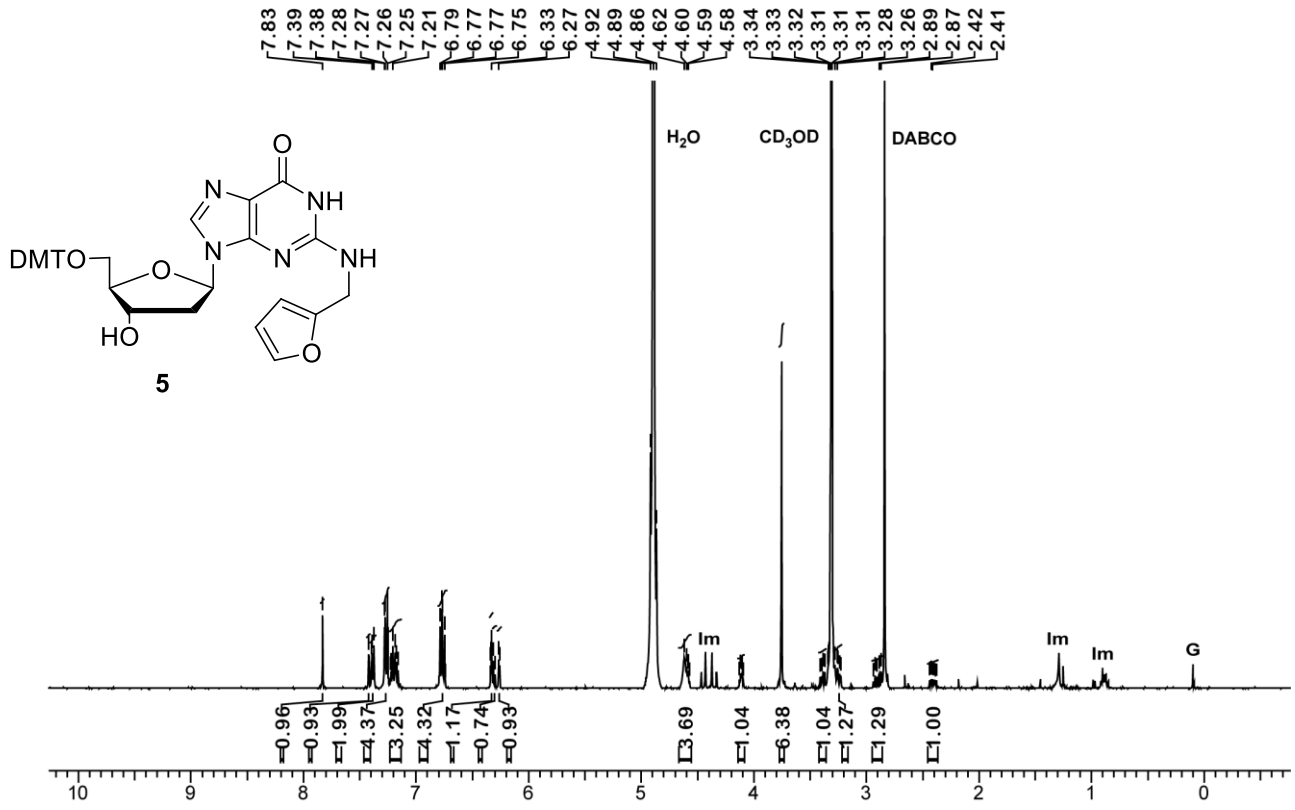
^{1}H NMR for Compound **4**



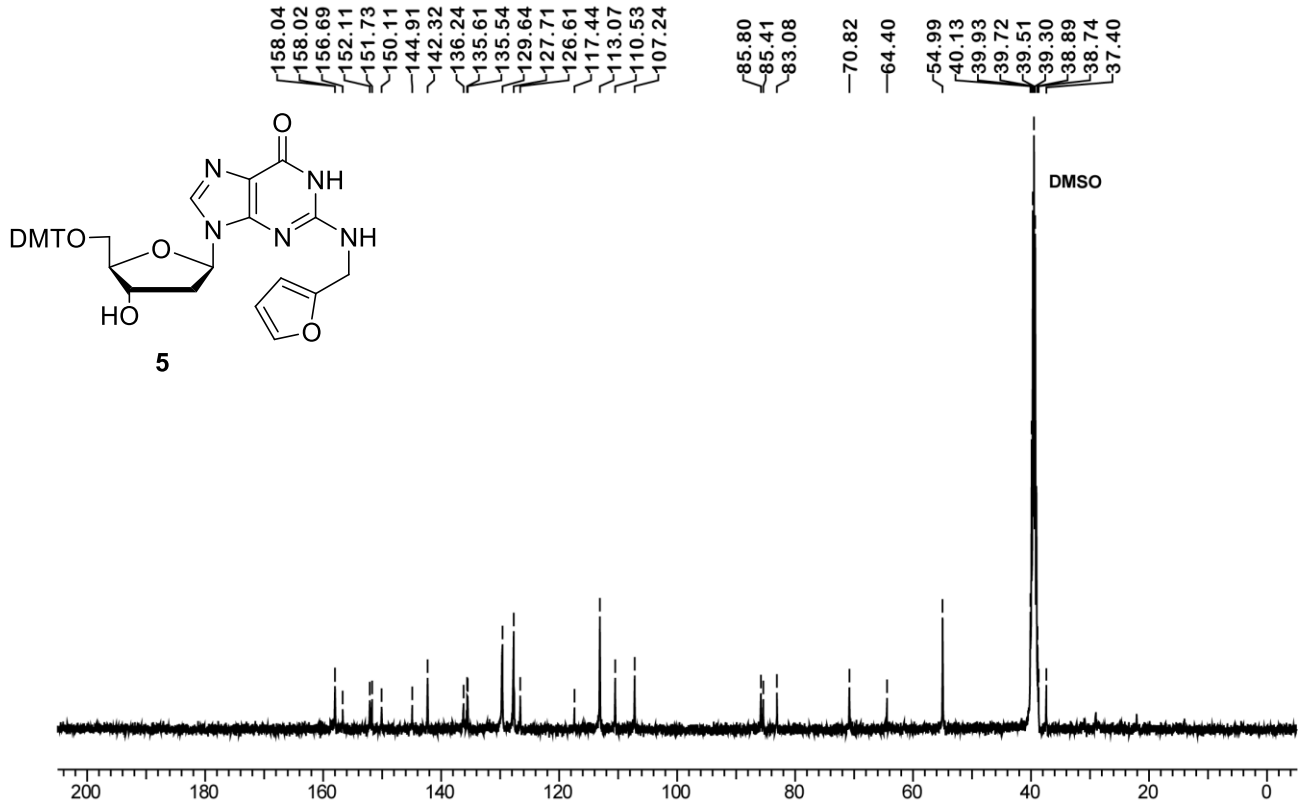
^{13}C NMR for Compound **4**



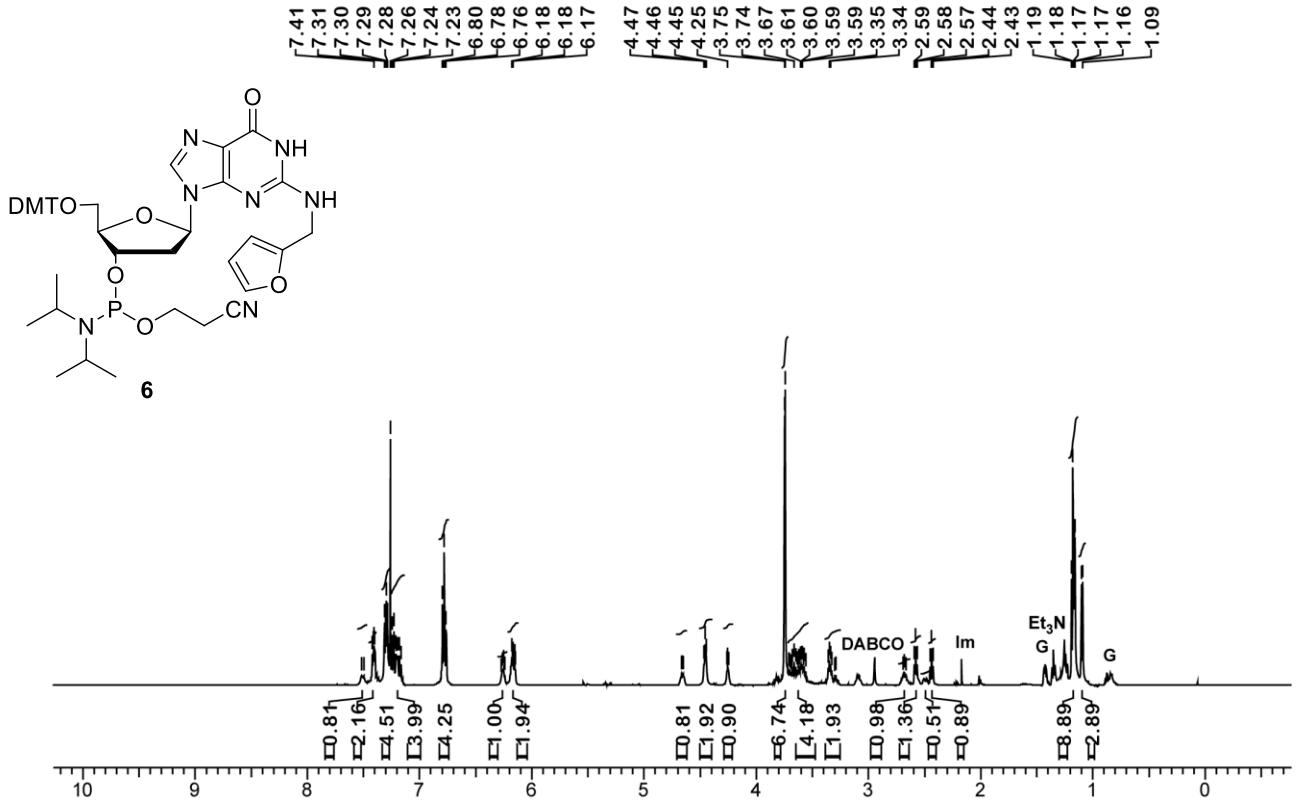
^{1}H NMR for Compound **5**



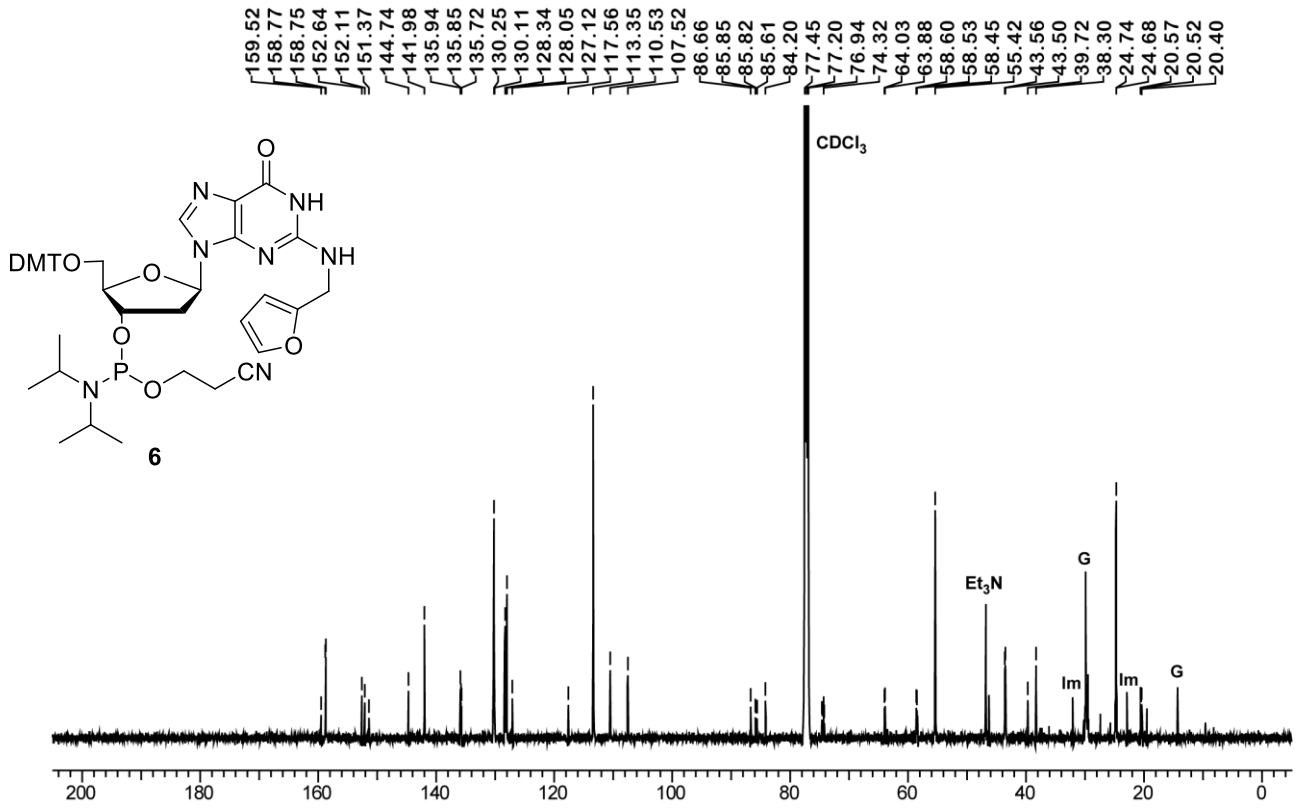
^{13}C NMR for Compound **5**



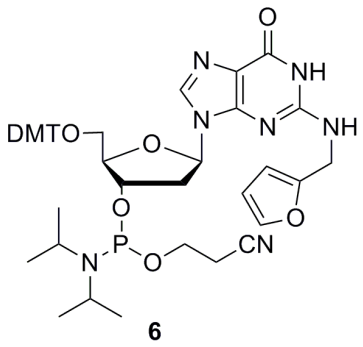
^{1}H NMR for Compound **6**



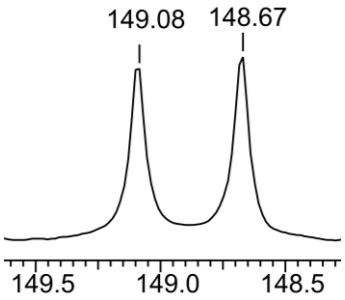
^{13}C NMR for Compound **6**

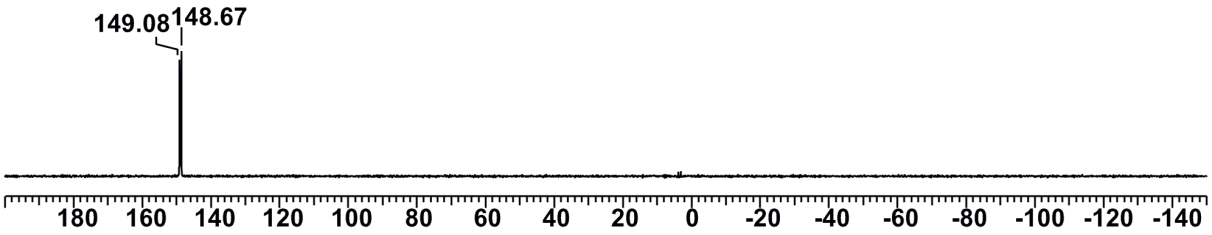


^{31}P NMR for Compound **6**

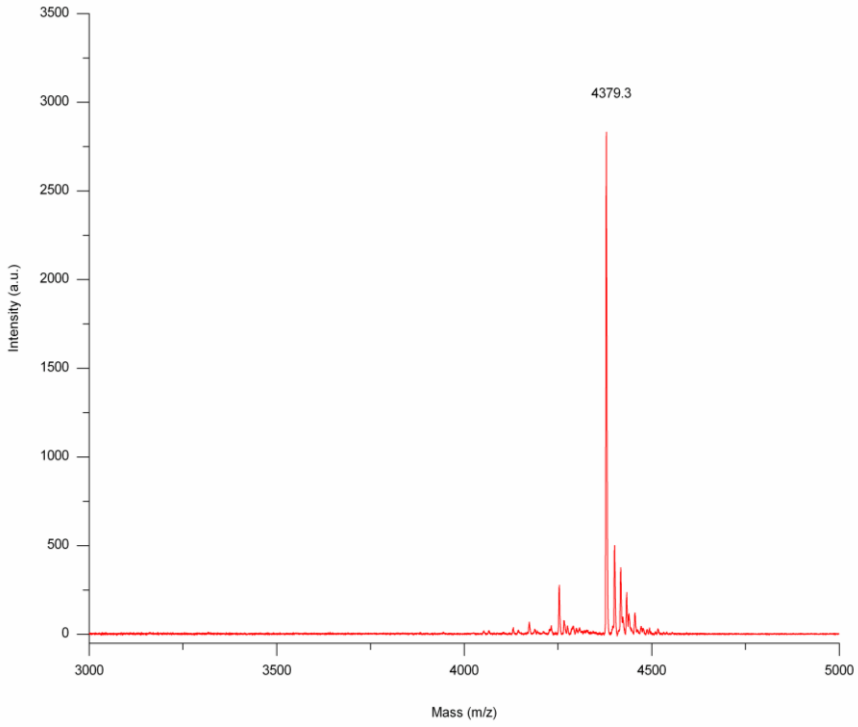


6

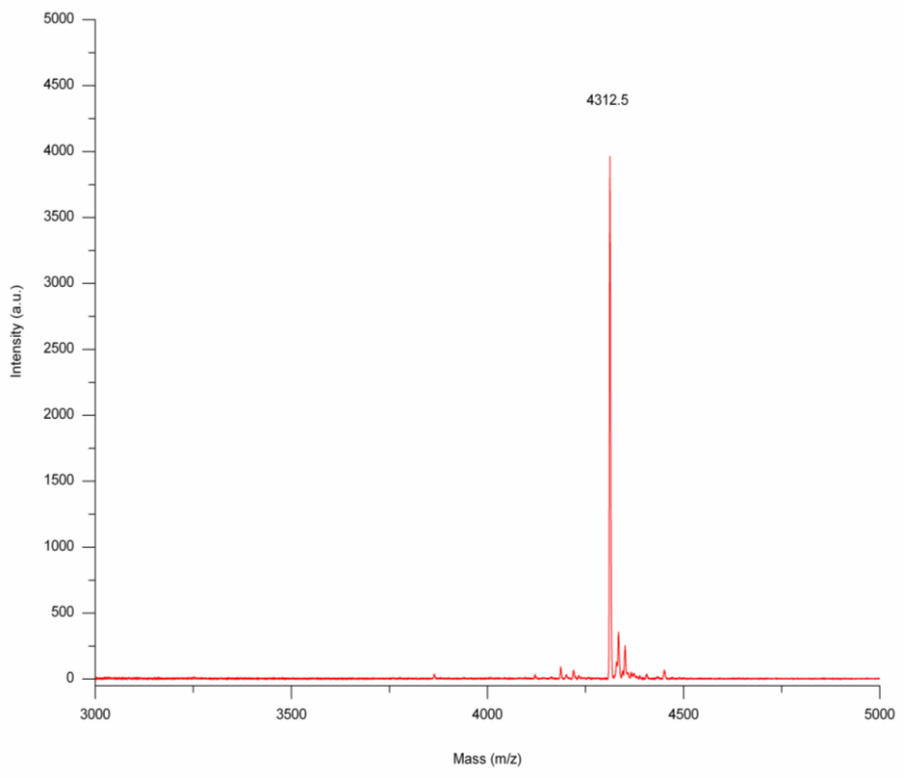




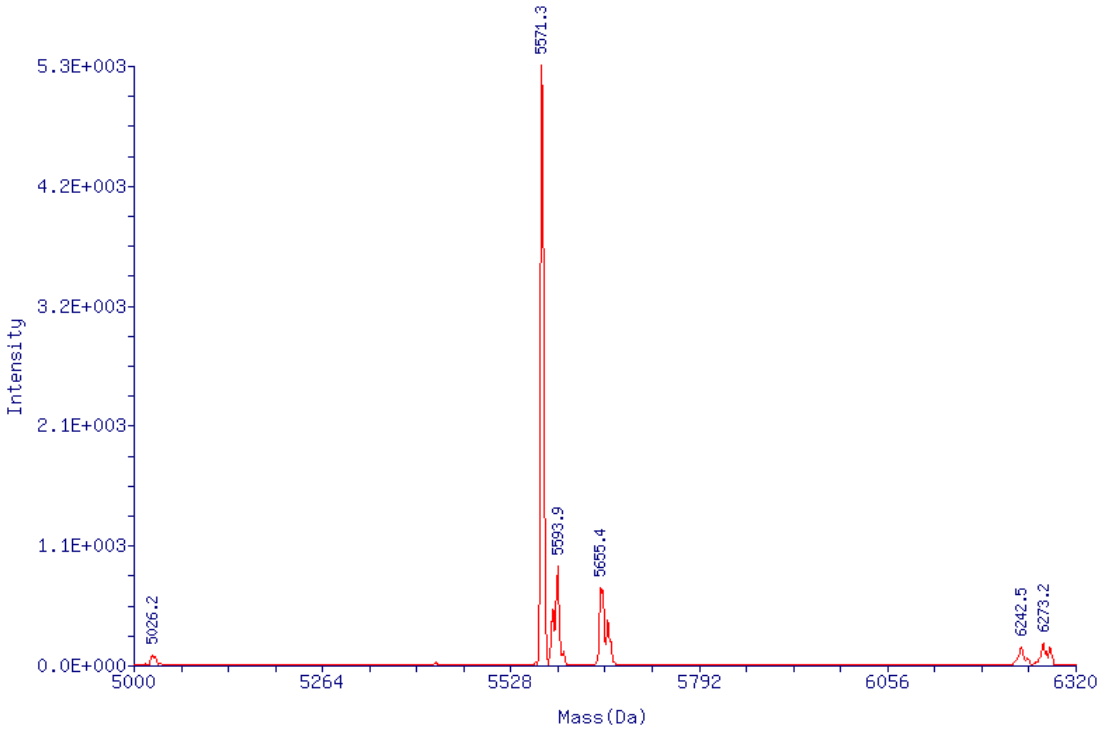
MALDI-MS of **D1**, 5’-GCCGXAATAGCGCA-3’: calc. mass, $[M–H]^-$ 4377; obs. mass, $[M–H]^-$ 4379.3



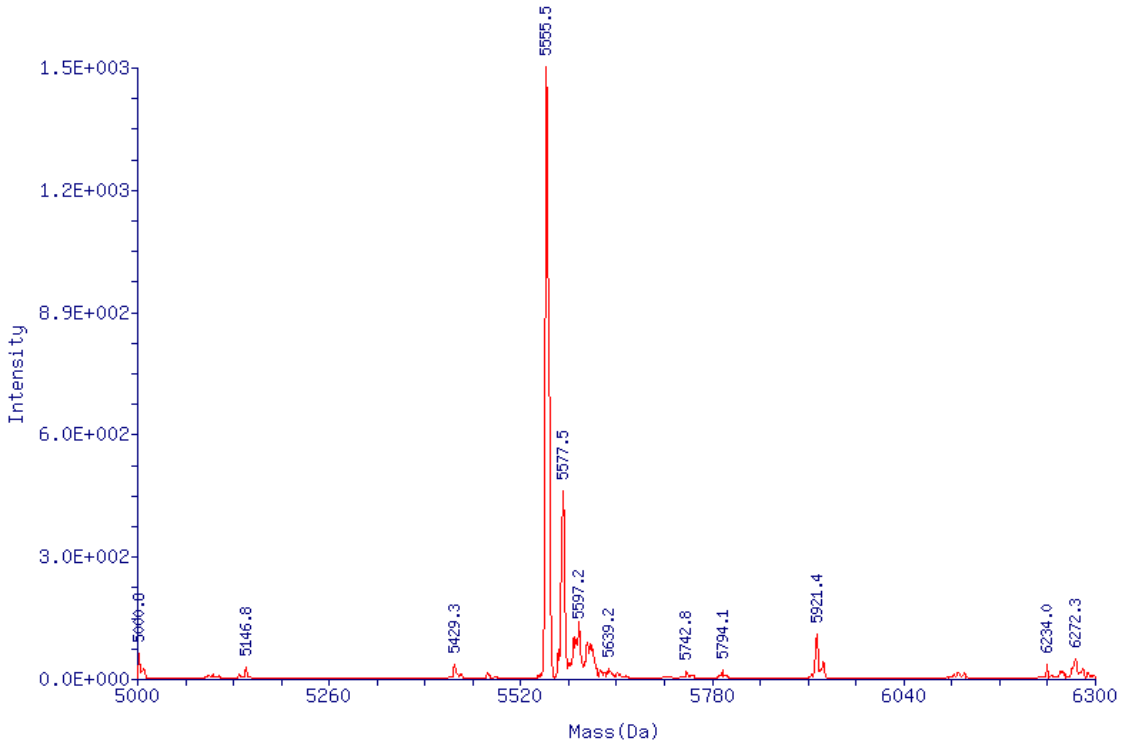
MALDI-MS of **D2**, 5’-TGCXCTATTCCGGC-3’: calc. mass, $[M–H]^-$ 4310; obs. mass, $[M–H]^-$ 4312.5



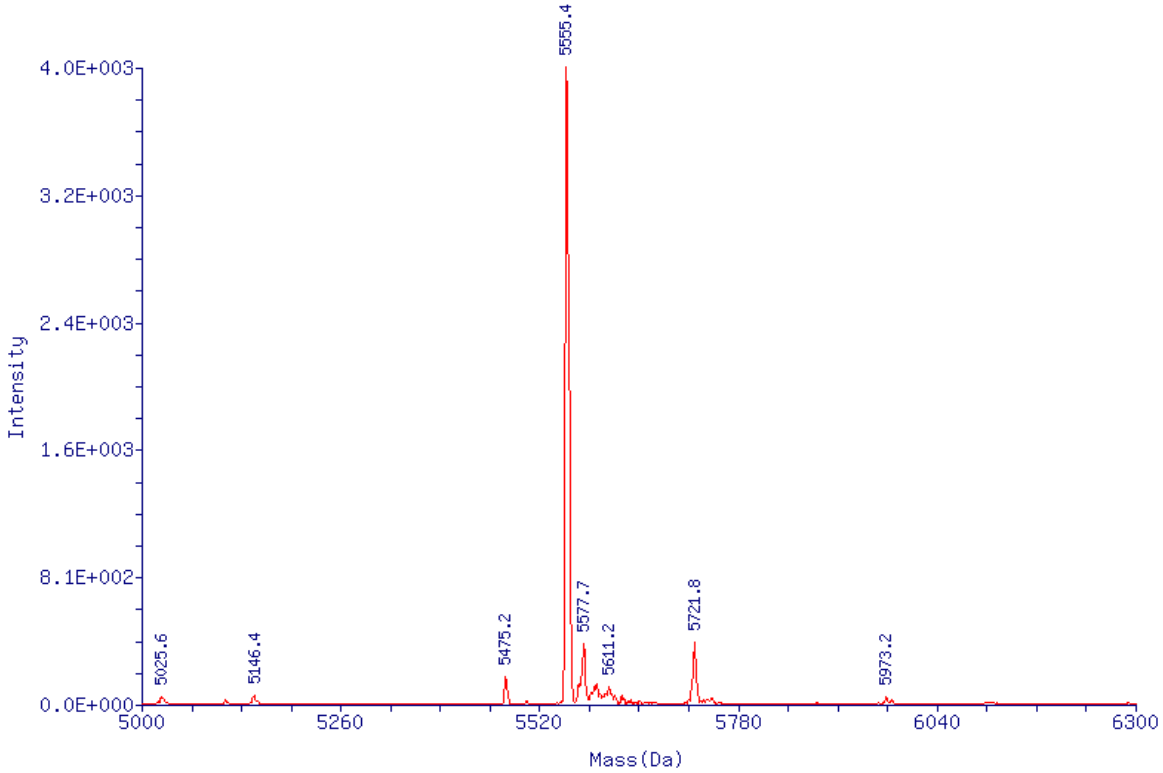
ESI-MS of **D3**, 5’-TCTXGGGTCCTAGGACCC-3’, calc. mass, $[M–H]^-$ 5570; obs. mass, $[M–H]^-$ 5571.3



ESI-MS of **D4**, 5’-TCTAXGGTCCTAGGACCC-3’: calc. mass, [M–H]⁻ 5554; obs. mass, [M–H]⁻ 5555.5



ESI-MS of **D5**, 5’-TCTAGGXTCCTAGGACCC-3’: calc. mass, [M–H]⁻ 5554; obs. mass, [M–H]⁻ 5555.4



ESI-MS of **D6**, 5’-TCCTACCGTGCCTACCTGAACAGCTGGTCACACTXATGCC TACGAGTACG-3’: calc. mass, $[M–H]^-$ 15346; obs. mass, $[M–H]^-$ 15347.9