JOC The Journal of Organic Chemistry

Subscriber access provided by Iowa State University | Library

Article

Synthesis of N2 -deoxyguanosine modified DNAs and the studies on their translesion synthesis by the E. coli DNA polymerase IV

Pratibha P. Ghodke, Praneeth Bommisetti, Deepak T Nair, and Pushpangadan Indira Pradeepkumar

J. Org. Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.joc.8b02082 • Publication Date (Web): 10 Jan 2019

Downloaded from http://pubs.acs.org on January 10, 2019

Just Accepted

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.

ACS Publications

is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Published by American Chemical Society. Copyright © American Chemical Society. However, no copyright claim is made to original U.S. Government works, or works produced by employees of any Commonwealth realm Crown government in the course of their duties.

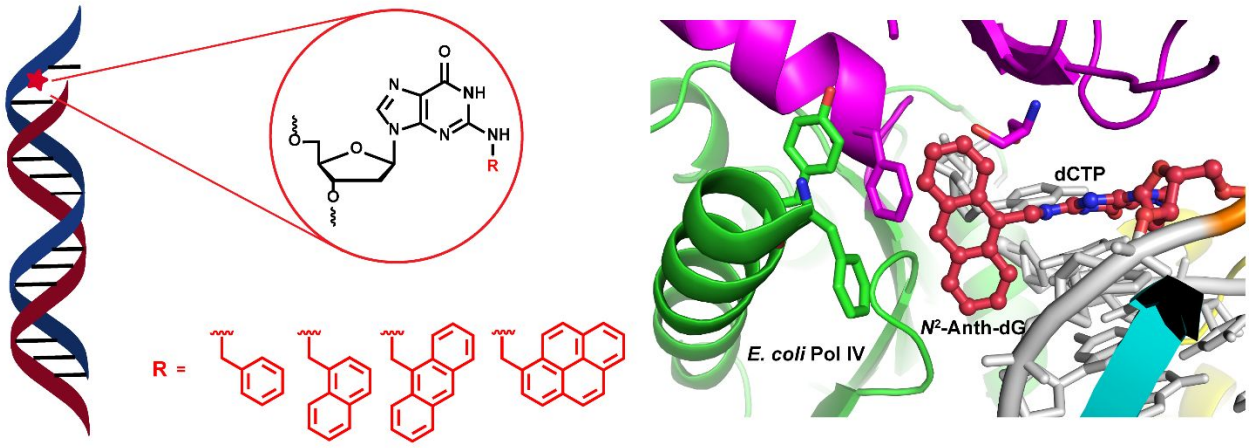
Synthesis of N^2-deoxyguanosine modified DNAs and the studies on their translesion synthesis by the *E. coli* DNA polymerase IV

Pratibha P. Ghodke,[†] Praneeth Bommisetti,[†] Deepak T. Nair,*[,§] P. I. Pradeepkumar*[,†]

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

[§]Regional Centre for Biotechnology, NCR Biotech Science Cluster, 3rd Milestone, Faridabad-Gurgaon Expressway, Faridabad-121001, India

Table of Contents (TOC) Graphic



ABSTRACT

We report the synthesis of N^2-aryl (benzyl, naphthyl, anthracenyl and pyrenyl)-deoxyguanosine (dG) modified phosphoramidite building blocks and the corresponding damaged DNAs. Primer extension studies using *E. coli* Pol IV, a translesion polymerase, demonstrate that translesion synthesis (TLS) across these N^2-dG adducts is error free. But the efficiency of TLS activity decreases with increase in the steric bulkiness of the adducts. Molecular dynamics simulations of damaged DNA-Pol IV complexes reveal the van der Waals interactions between key amino acid residues (Phe13, Ile31, Gly32, Gly33, Ser42, Pro73, Gly74, Phe76, and Tyr79) of the enzyme and adduct that help to accommodate the bulky damages in a hydrophobic pocket to facilitate TLS. Overall, the results presented here provide insights into the TLS across N^2-aryl-dG damaged DNAs by Pol IV.

.

INTRODUCTION

Polycyclic aromatic hydrocarbons (PAHs) are fused benzenoid ring containing compounds that are well known for their carcinogenic and mutagenic properties.[1] The exocyclic amino group of deoxyguanosine (dG) acts as a softer reaction center to form major carcinogenic N^2-dG DNA adducts.[2,3] These adducts can lead to genomic alterations via stalling replicative polymerases. To avoid any detrimental consequence, cells use specialized low-fidelity DNA polymerases belong to Y-family to bypass the damaged sites. This process called translesion synthesis (TLS), an evolutionarily conserved DNA damage tolerance pathway, which can be error-free or error-prone.[4] Among various TLS polymerases, Pol IV from *E. coli* is well known for accurate and efficient bypass of N^2-dG adducts.[5–7] Pol IV is an ortholog of human TLS polymerase, Pol κ.[5] A systematic study on the effects of the steric bulkiness of N^2-dG adducts on TLS by *E. coli* Pol IV is yet to be reported.

Among the N^2-dG DNA adducts, N^2-benzyl-dG (N^2-Bn-dG, Figure 1) is potentially formed from three key carcinogens. *N*-nitroso-*N*-benzylurea (BnNU) and *N*-nitrosobenzylmethylamine (NBnMA) could form N^2-Bn-dG adducts via benzyldiazonium ion [8] and methyl hydroxylation[9] pathway respectively. Whereas, benzyl halides could directly act as a benzylating agent.[10,11] Among these carcinogens, NBnMA is believed to form N^7-benzyl-dG adduct in in *vivo* conditions, since the N^7-position is one of the major site of DNA alkylation.[11] The 9-(Sulfoxymethyl)anthracene is an electrophilic metabolite of 9-hydroxymethyl-anthracene, which was alluded to be responsible for the formation of N^2-(9-anthracenylmethyl)-dG adduct (N^2-Anth-dG, Figure 1).[12] Halomethyl-anthracenes such as bromo and chloromethyl derivatives have as well been suggested to alkylate the N^2-position of dG.[13] 1-Methylpyrene, a common environmental pollutant can enzymatically be activated to 1-(hydroxymethyl)pyrene, which in turn can be converted to 1-(sulfoxymethyl)pyrene (SMP). SMP on reaction with DNA leads to the formation of N^2-(1-pyrenylmethyl)-dG adduct (N^2-

Pyre-dG, Figure 1).[14,15] Though N^2-Bn, N^2-Naph, and N^2-Anth (Figure 1) adducts are yet to be detected in cellular DNA, the N^2-Pyre-dG adduct was detected in DNA from lung, liver, and kidney tissues of mice.[14]

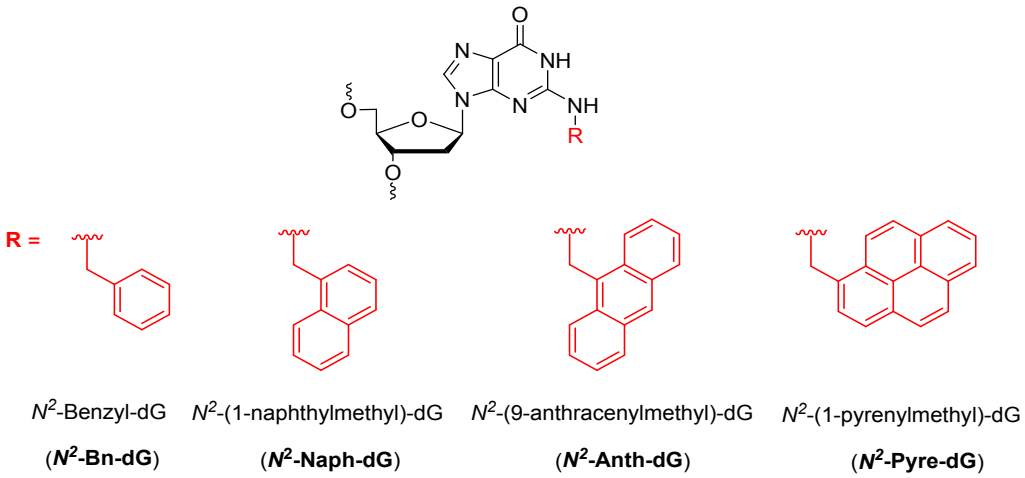


Figure 1. Structures of N^2-dG adducts which originate from polycyclic aromatic hydrocarbons.

To unravel the molecular mechanism of TLS across N^2-aryl-dG adducts formed by non-diol epoxide metabolic pathways by using various biochemical and biophysical methods, the corresponding modified DNA oligonucleotides are required. Herein we report a short and reliable synthesis of N^2-dG adducts of varying steric bulkiness and the corresponding modified DNAs. Using primer extension assays, molecular dynamics studies, we were able to gain insights on the TLS by *E. coli* Pol IV across N^2-(Bn, Naph, Anth, Pyre)-dG damaged templates.

RESULTS AND DISCUSSION

To access various N^2-dG damaged DNAs, either phosphoramidite or post-oligomerization approaches were reported.[16–18] Synthesis of N^2-Bn-dG adduct was achieved by different methods such as substitution of C2-trimethanesulfonate group bearing dG by benzylamine,[19] aromatic nucleophilic substitution of 2-fluoro-2'-deoxyinosine (2-fluoro-dI),[20]and reaction of dG with benzyl bromide,[10,21] reduction of N^2-benzoyl group using reducing agents,[22] reductive amination of dG with benzaldehyde[23] and by Buchwald-Hartwig (B-H) coupling between TBDMS protected 2-bromo-2'-deoxyinosine (2-bromo-dI) with benzylamine.[24] The N^2-Naph-

dG adduct was synthesized using nucleophilic substitution of DMT protected 2-fluoro-dI with the corresponding amine.[47] Synthesis of N^2-Anth-dG adduct was achieved using B-H coupling,[24] or by the hydrolysis of 2'-deoxy-4-desmethylwyosine.[13] Post-oligomerization approach was also utilized for the synthesis of N^2-Anth-dG adduct employing 2-fluoro-dI bearing oligonucleotide.[23] The N^2-Pyre-dG adduct was acessed by fluoro substitution method using 2-fluoro-dI convertible nucleoside and 1-(aminomethyl)pyrene.[25]

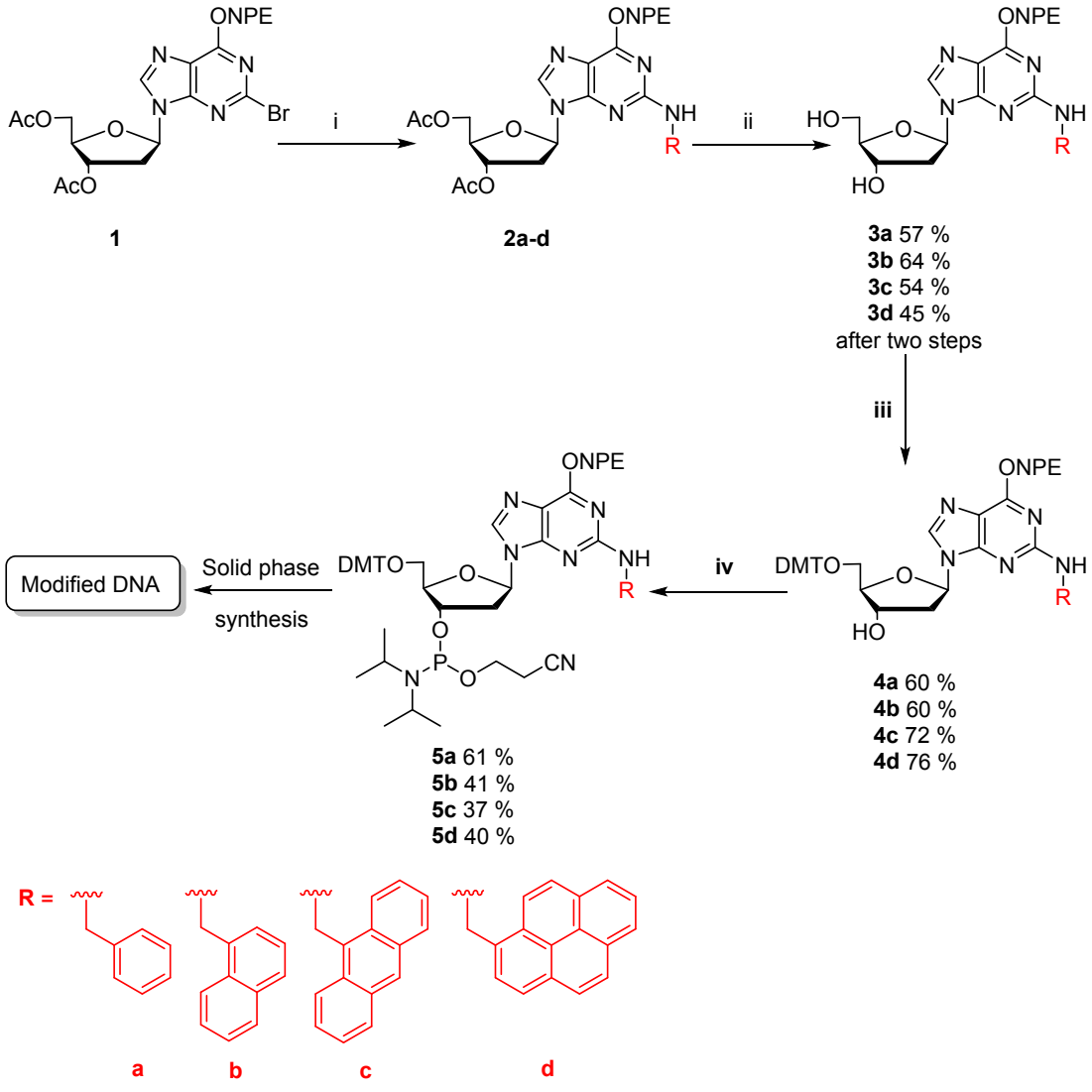
Our objective was to develop a robust protocol to synthesize N^2-aryl-dG modified nucleoside and corresponding oligonucleotides (Figure 1) from a common precursor. To achieve this, first, we attempted various reported methods. However, in our hands, many of those methods were marred by lack of reproducibility, capricious yields, and poor solubility of various intermediates/amidites.[26] In the present work, we report a reliable and efficient method for the synthesis of various N^2-aryl-dG modified phosphoramidite building blocks employing B-H coupling as the key step. These modified building blocks were synthesized in the presence of a versatile *para*-nitrophenylethyl (NPE) protecting group at O^6-position of dG. This route is more convenient since all the bulky modifications can be synthesized from a common precursor with good yields.

Synthesis of N^2-Bn, N^2-Naph, N^2-Anth, N^2-Pyre-dG modified phosphoramidites 5a-d

The synthetic Scheme 1 was utilized to access N^2-modified-dG (N^2-Bn, N^2-Naph, N^2-Anth, N^2-Pyre-dG) phosphoramidite building blocks. Using the protected 2-Br-dI, **1**, as a common precursor,[27] the C-N bond formation was carried out with different polycyclic aromatic amines in presence of $Pd(OAc)_2$, (*R*)-BINAP, Cs_2CO_3 in toluene.[28] For N^2-Bn and N^2-Naph-dG modifications, B-H coupling were carried out using commercially available amine while for N^2-Anth-dG and N^2-Pyre-dG modifications, the respective amines were prepared by reported methods.[25,29,30] Deprotection of acyl protecting group was accomplished using 33% $MeNH_2$ in EtOH (v/v) to obtain N^2-modified-dG diols (**3a-d**). Final modified phosphoramidites **5a-d** were

synthesized by the protection of 5'-OH group using DMT-Cl, followed by phosphitylation reaction using CEP-Cl, DIPEA in DCM.

Scheme 1. Synthesis of N^2-Bn, N^2-Naph, N^2-Anth and N^2-Pyre-dG modified phosphoramidites 5a-d



Reagent and conditions: (i) $Pd(OAc)_2$, (*R*)-BINAP, respective amines, Cs_2CO_3, toluene, 85 °C; (ii) 33% $MeNH_2$ in EtOH, rt; (iii) DMT-Cl, pyridine, rt; (iv) CEP-Cl, DIPEA, DCM, rt.

Solid phase synthesis of N^2-dG modified oligonucleotides

The N^2-dG modified phosphoramidite building blocks (**5a-5d**) were used to synthesize the desired modified oligonucleotides. The presence of versatile NPE protecting group at O^6-position of dG provided enhanced solubility of respective phosphoramidites during solid phase

DNA synthesis.[39,50,31] We have chosen NPE protecting group because of its chemical inertness and its stability in the presence of mild acids and bases.[32] The N^2-aryl-dG damaged DNA sequences (Table 1) were synthesized using automated DNA synthesizer employing appropriate CPG solid supports. The 18-mer DNA sequences (Table 1) were utilized for crystallization with complementary 14-mer primer, incoming nucleotide and *E. coli* Pol IV.[33] These sequences were designed for easy assembling of the three functional complexes of Pol IV with three different DNA duplexes containing the N^2-dG adduct and incoming nucleotide. The 50-mer N^2-dG DNA template sequences were synthesized for the primer extension studies (Table 1). Deprotection of N^2-dG modified DNAs was carried out in four different steps. The first step involved the selective deprotection of cyanoethyl group using 10% diethylamine in ACN.[34] The deprotection of NPE group was performed using 1 M DBU in ACN.[23] The cleavage of the CPG support and deprotection of the base protecting groups were carried out separately by treatment of aq. NH_3. Modified DNAs were PAGE purified and the integrity of modified DNAs was confirmed by ESI-MS or MALDI-TOF (Table 1).

Table 1. N^2-dG damaged oligonucleotides and their molecular weights

Code	DNA Sequences (5'-3')	Mol wt. (Calcd.)	Mol wt. (Found)	Error, % (found−calcd.)
D1	TCTAG1GGTCCTAGGACCC	5566.7	5566.0	−0.01
D2	TCTG1GGGTCCTAGGACCC	5581.7	5581.9	0.004
D3	TCTAGGG1TCCTAGGACCC	5566.7	5567.3	0.01
D4	TCCTACCGTGCCTACCTGAACAGCTG GTCACACTG1ATGCCTACGAGTACG	15358	15354.8	−0.02
D5	TCTAG2GGTCCTAGGACCC	5616.8	5616.8	0
D6	TCTG2GGGTCCTAGGACCC	5632.8	5635.1	0.04
D7	TCTAGGG2TCCTAGGACCC	5616.8	5617.4	0.01
D8	TCCTACCGTGCCTACCTGAACAGCT GGTCACACTG2ATGCCTACGAGTACG	15408.1	15409.1	0.007

D9	TCTAG3GGTCCTAGGACCC	5666.9	5667.6	0.01
D10	TCTG3GGGTCCTAGGACCC	5682.9	5682.0	−0.02
D11	TCTAGGG3TCCTAGGACCC	5666.9	5666.9	0.001
D12	TCCTACCGTGCCTACCTGAACAGCTG GTCACACTG3ATGCCTACGAGTACG	15458.2	15453.8	−0.03
D13	TCTAG4GGTCCTAGGACCC	5690.9	5690.2	−0.01
D14	TCTG4GGGTCCTAGGACCC	5706.9	5705.8	−0.02
D15	TCTAGGG4TCCTAGGACCC	5690.9	5689.9	−0.02
D16	TCCTACCGTGCCTACCTGAACAGCTG GTCACACTG4ATGCCTACGAGTACG	15482.2	15480.9	−0.01

Structures of modifications G1 (N^2-Bn-dG), G2 (N^2-Naph-dG), G3 (N^2-Anth-dG) and G4 (N^2-Pyre-dG) are shown in Figure 1. The modified DNAs were characterised by MALDI-TOF in the positive reflectron/linear mode or negative ion electrospray ionization (ESI) technique (calculated and found). *Only **D2** was characterized by ESI-MS in negative mode.

Primer extension studies using *E. Coli* DNA PoI IV

Primer extension reactions with *E. coli* Pol IV were carried out using a primer-template system as shown in Figure 2A. Reactions were carried out with individual dNTP or with a mixture of dNTPs employing unmodified and modified DNA templates and were monitored at various time courses from 30 sec to 2 h (Figure 2 and Figures S1 to S5 of the supporting information). In case of single nucleotide incorporation reactions, we observed a prominent extension product of 16 nucleotide length, which corresponds to correct base (dCTP) incorporation across all the N^2-dG modified templates (Figure 2B to 2F, lane C). With the use of 4 nM of Pol IV, ~ 94% incorporation of the correct base was observed opposite to unmodified dG template at 1.5 h time period (Figure 2B, lane C). For the N^2-Bn-dG modification ~ 80% (Figure 2C, lane C), for the N^2-Naph-dG modification ~ 75% (Figure 2D, lane C), for the N^2-Anth-dG modification ~ 62% (Figure 2E, lane C), and for the N^2-Pyre-dG modification ~ 64% (Figure 2F, lane C) incorporation products were observed after 1.5 h. These results reveal that the TLS employing Pol IV is error free irrespective of the steric bulkiness of the N^2-dG adducts.[5,35] However,

incorporation efficiency of Pol IV decreases with increasing steric bulkiness from the N^2-Bn to the N^2-Pyre-dG adducts. It should be noted that in addition to the correct incorporation, a trace amount of dTTP misincorporation was also observed opposite to the N^2-Bn-dG (~ 10%, Figure 2C, lane T), and the N^2-Naph-dG modifications (~ 4%, Figure 2D, lane T).

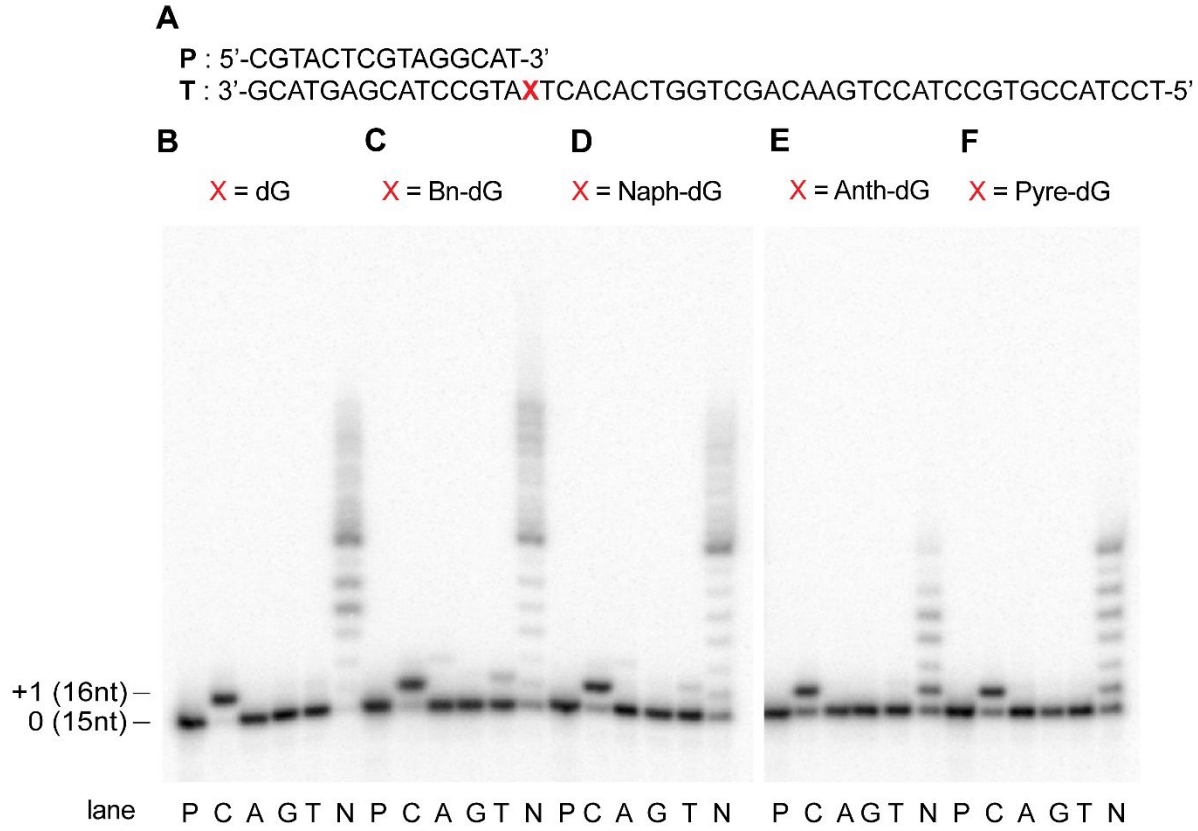


Figure 2. PAGE (20%, 7 M urea) of primer extension reactions with single and with the mixture of dNTPs using unmodified (dG) and modified (N^2-dG adducts) templates employing *E. coli*. Pol IV (A) Complete sequence of the primer, P (15-mer) and the template, T (50-mer); (B) reactions with dG template (X = dG) (lanes: P:primer, C:dCTP, A:dATP, G:dGTP, T:dTTP, N:mixture of dNTPs); (C) reactions with N^2-Bn-dG template; (D) reactions with N^2-Naph-dG template; (E) reactions with N^2-Anth-dG template; and (F) reactions with N^2-Pyre-dG template. All reactions were carried out at 37 °C for 1.5 h time course.

To verify the efficiency of *E. coli* Pol IV towards full length extension of the primer, running start experiments were carried out using 15-mer and 11-mer primers with the modified DNA templates. Running start experiments using 15-mer primer, showed that the DNA Pol IV is unable to synthesize full length products irrespective of the chemical nature of the damage (Figures 3B to F). Similar results were obtained in running start experiments with 11-mer primer (Figures S6 to S7 of the supporting information). As expected, in the case of unmodified

template, we observed traces of the full-length product. These results are in line with the low processivity reported for *E. coli* Pol IV, which is a hallmark of this family of DNA polymerases.[5,36] It should be noted that Pol IV belongs to the Y-family of DNA polymerases and members of this family are known to bypass DNA lesions and rescue replication stalled at damaged nucleotides. The majority of Y-family DNA polymerases serve only to extend the replication fork past the DNA lesion after which they are replaced by replicative DNA polymerases. [4]

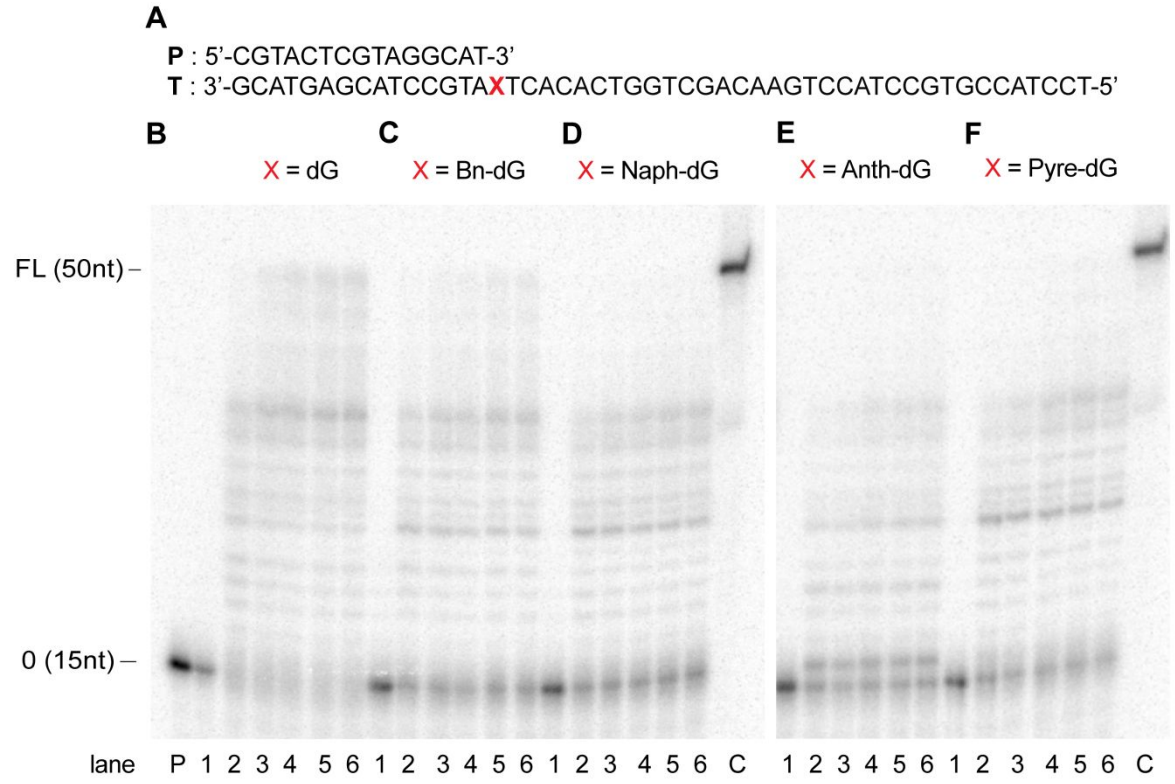


Figure 3. PAGE (20%, 7 M urea) of full-length extension reactions employing Pol IV from *E. coli* with all dNTPs. (A) Complete sequence of template, T and primer, P (15-mer); (B) reactions with dG template (X = dG); (C) reactions with N^2-Bn-dG template; (D) reactions with N^2-Naph-dG template; (E) reactions with N^2-Anth-dG template; (F) reactions with N^2-Pyre-dG template. Lane P, primer; lanes 1 to 6: primer extension reactions with 250 μM of mixture of dNTPs in different time course from 30 sec., 30 min., 1, 2, 4 and 8 h; lane C, 50-mer size standard. All the reactions were carried out at 37 °C.

Molecular modelling and dynamics studies

Molecular dynamics (MD) simulations have been carried out to elucidate the factors that facilitate TLS across N^2-dG DNA adducts by *E. coli* Pol IV. The reported X-ray crystal structure of furfuryl modified N^2-dG DNA adduct in complex with the *E.coli* Pol IV (PDB ID: 4Q43[6]) was used to prepare the starting structures. The N^2-dG adducts, which were used in the experimental studies (Figure 1) were incorporated into the DNA (Figure S8, Supporting information). A total of four DNA-Pol IV ternary complexes (enzyme, DNA and incoming

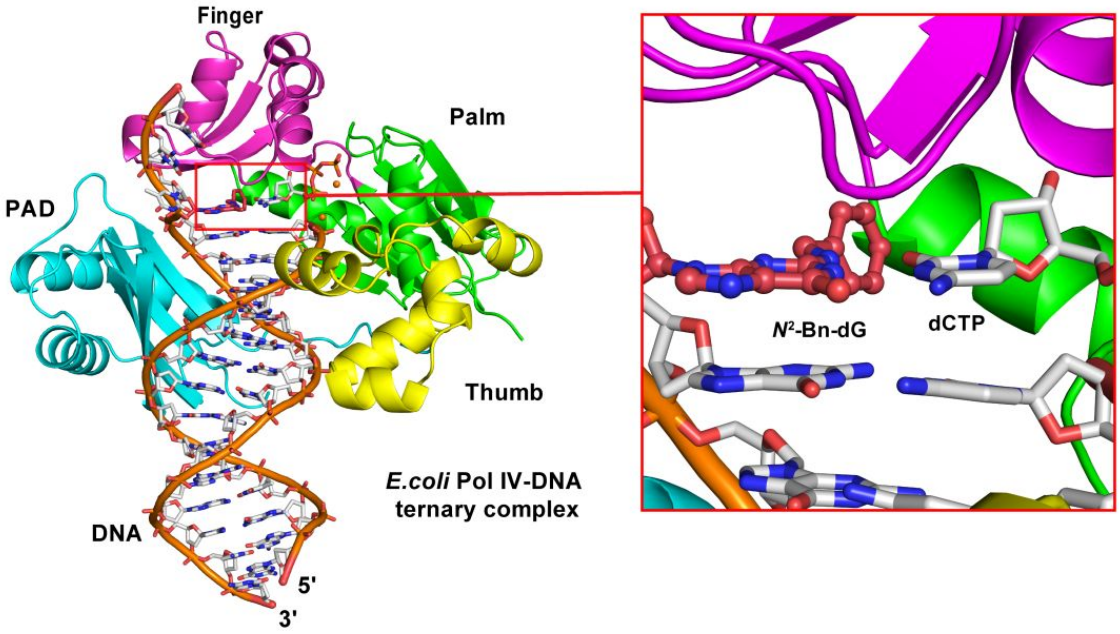


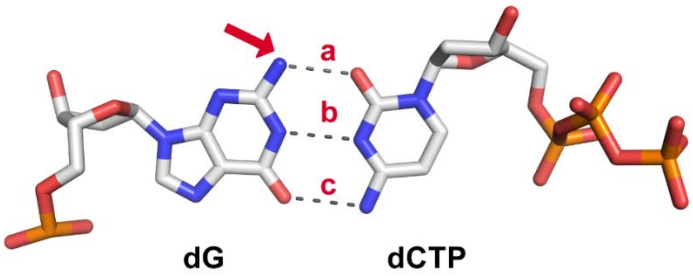
Figure 4. DNA-Pol IV ternary complex used for MD simulations. The initial structure of N^2-Bn complex prepared from fdG containing crystal structure (PDB ID: 4Q43). The red coloured rectangle is used to highlight the active site and zoomed view of the active site. Other complexes are shown in Figure S9, Supporting Information. Cyan, magenta, green and yellow are used for polymerase associated domain (PAD), finger, palm and thumb domains respectively. DNA backbone is represented in orange-red colour. The N^2-modified residue is highlighted in desire red colour. Carbons atoms are represented in white (desire red for N^2-Bn-dG), nitrogen atoms in blue, oxygen atoms in red, and magnesium ions in orange colour.

nucleotide components) were prepared and were subjected to 100 ns of unrestrained MD simulations (Figure 4 and Figure S9, Supporting information). Complexes were generated only for the insertion stage of dCTP. Trajectory analysis of each of these complexes illuminated the key interactions between the DNA and Pol IV at the active site.

Stability of the DNA – Pol IV ternary complexes

To confirm the adequate sampling and the stability of DNA-Pol IV ternary complexes, root mean square deviations (RMSD) of backbone atoms of DNA and protein were calculated separately with respect to the first frame. Protein backbone showed marginal deviations with an average RMSD value < 1.8 Å for all the complexes (Figure S10, Supporting information). The DNA backbone showed considerable fluctuations with an average RMSD value > 2.0 Å. In order to further investigate such large fluctuations in the DNA backbone, root mean square fluctuation (RMSF) values were calculated for DNA with respect to the first frame. Only a few nucleotides which don't interact with protein showed larger fluctuations, whereas the nucleotides around the active site showed fewer fluctuations (Figure S11, Supporting information). RMSF values of protein residues showed only nominal fluctuations in all the complexes (Figure S12, Supporting information). Also, the distance between the α phosphate of dCTP and the O atom of 3′-OH of the primer terminal dG is < 4.0 Å in all the complexes, indicative of a productive conformation for the nucleotide incorporation at the insertion stage. The incoming nucleotide dCTP forms Watson-Crick (W-C) H-bonding with dG at the active site. These W-C H-bonds were monitored through the course of dynamics and the percentage occupancy of each of H-bond was calculated (Table 2). The H-bond (a) formed by dG:N^2-H ⋯ O2:dCTP is affected to some extent in all the complexes. The interaction between the surrounding amino acids (AAs) with the N^2-modifications, made the orientation of the N^2-modified dG unfavourable for W-C H-bonding at certain time frames during the dynamics. The H-bond labelled (a) in Table 2 is most affected in Anth complex in which the H-bond (b) was also disturbed (Table 2).

Table 2: **Percentage occupancies of each H-bond at the active site of *E. coli* Pol IV**



Complex with adduct	a	b	c
N^2-Bn-dG	92.08 %	95.26 %	93.01 %
N^2-Naph-dG	89.88 %	95.59 %	91.77 %
N^2-Anth-dG	84.58 %	88.16 %	93.04 %
N^2-Pyre-dG	88.18 %	95.94 %	92.27 %

The dotted lines indicate the H-bond present between the dG and incoming nucleotide dCTP. Red arrow indicates the position of modification

Clustering analysis

MD snapshots (100 ns) of each of the complex were clustered into 5 ensembles based on hierarchical agglomerative approach.[37] The percentage of each ensemble out of 100 ns of dynamics is tabulated in Table 3. To identify the differences in the major ensembles, representative structures of ensemble 1 and ensemble 2 were superimposed for all the complexes. These representative structures differed primarily in the orientations of terminal nucleotides of DNA. The similarities in the orientations of the AAs surrounding the modifications are highlighted in Figure 5. The residues surrounding N^2-Anth-dG adduct showed least differences compared to other adducts. Minor differences were observed in the orientations of Ser42 and Phe76 residues in the other complexes.[5,6]

Table 3. Percentage occupancies of each ensemble out of 100 ns of dynamics.

Complex with adduct	Ensemble 1	Ensemble 2	Ensemble 3
N^2-Bn-dG	43.3%	35.4%	14.1%
N^2-Naph-dG	59.9%	23.0%	7.2%
N^2-Anth-dG	56.4%	27.9%	8.2%
N^2-Pyre-dG	47.8%	31.8%	17.8%

Every second frame was considered, and a total of 25000 frames were considered for the calculation. Clustering was performed using hierarchical agglomerative approach, which was implemented by CPPTRAJ[38] module of AMBER14[39]

To find out the similarities between the four ternary complexes at the active site, the representative structures of ensemble 1 of all the complexes were superimposed onto each other (Figure S14, Supporting information). Orientations of the N^2-modified moieties were found to be strikingly similar to that of the furfuryl adduct in the active site of DNA-Pol IV complex as revealed by its crystal structure (PDBID: 4Q43). This implies that the Pol IV orients N^2-modified moieties into a particular conformation to facilitate TLS. To validate this, the dihedral angel labelled ξ (Figure 6) was monitored through the course of dynamics. Bn, Naph and Pyre complexes appeared to follow a pattern but Anth complex showed slight deviations. This difference might be a result of the unique position of linkage of the anthracene moiety to the nucleobase (Figure 1). The average dihedral angle values of 103.18°, 107.58°, 148.13°, 96.31° were observed for Bn, Naph, Anth and Pyre complexes

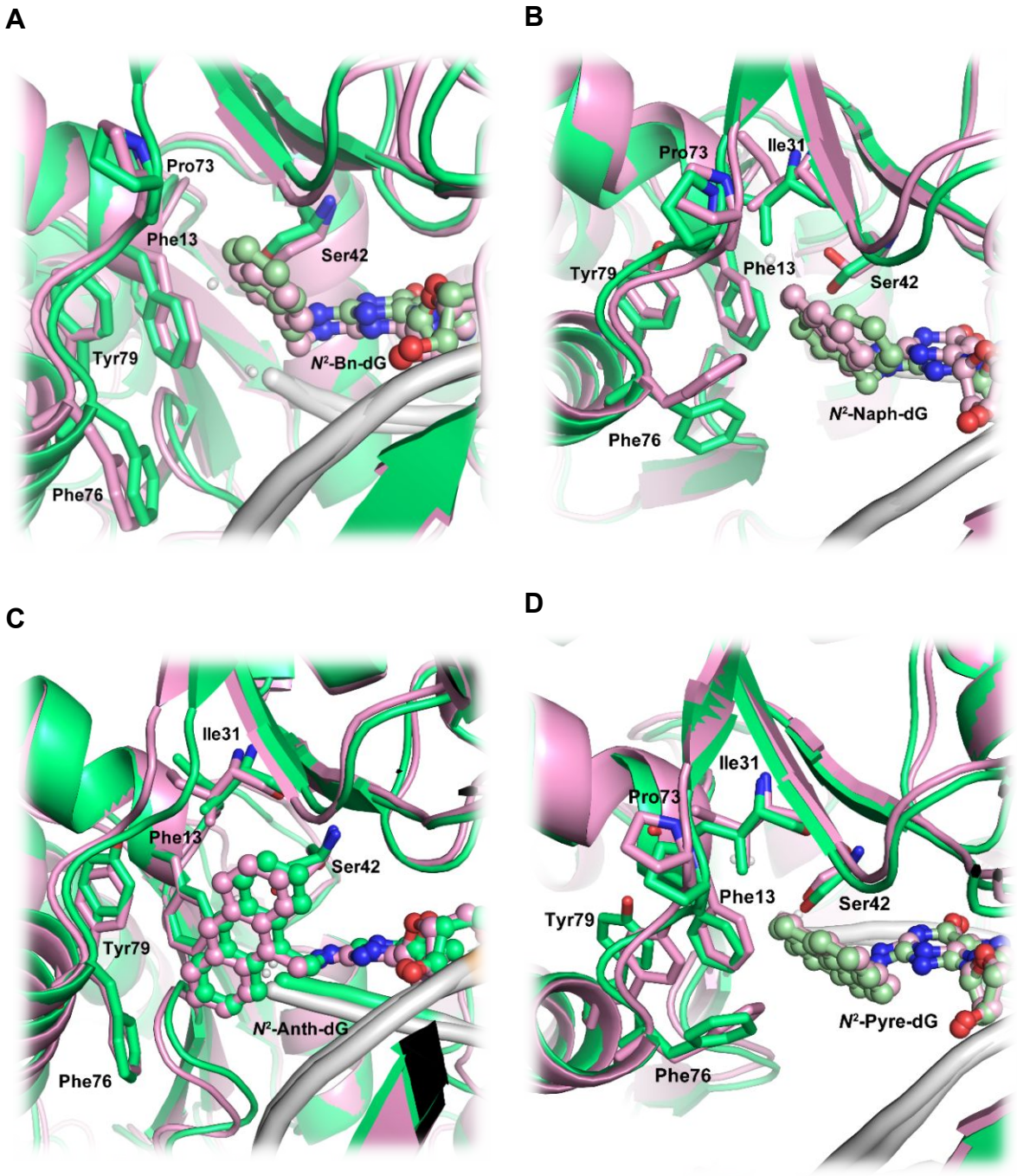


Figure 5. Superposition of representative structures of ensemble 1 and 2 of N^2-dG-modified DNA-Pol IV ternary complexes. Only residues of interest are highlighted. (A) N^2-Bn complex. (B) N^2-Naph complex. (C) N^2-Anth complex. (D) N^2-Pyre complex. Lime green coloured cartoon representation is used for the ensemble 1 and light pink coloured cartoon representation is used for the ensemble 2. DNA backbone is represented in white for both the representative structures. Pale green coloured carbon framework is used for N^2-adducts for ensemble 1 and pale pink coloured carbon framework for ensemble 2. Nitrogen atoms are represented in blue and oxygen atoms in red. All the hydrogens are removed for clarity.

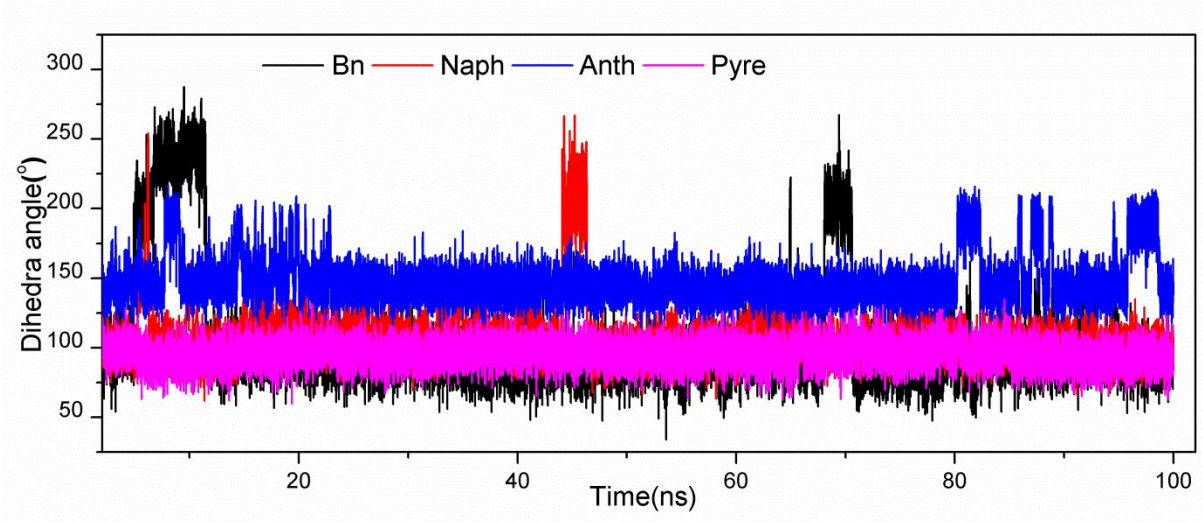


Figure 6: Plot of dihedral Xi (ξ, shown above using red arrow) vs time for all the complexes. Black, red, blue and pink colours were used to represent N^2-Bn, -Naph, -Anth and -Pyre complexes respectively.

respectively. Based on the above results, few important AA residues were identified and the trajectory was further analysed to understand the interaction patterns between these adducts and AAs, and also to probe the nature of the forces responsible for the accommodation of adducts in the active site.

Key non covalent interactions and interaction energies

Average interaction energies of the N^2-dG adducts with the surrounding AA residues were calculated to identify the nature of interactions in the active site. Since our focus is the modified nucleobases, only residues in close proximity to these modifications (within 6 Å) were chosen for the calculations. Interaction energies were calculated using MM/GBSA method implemented in AmberTools package. The total average energy was calculated by summation of van der Waals, electrostatic, polar solvation and nonpolar solvation values (Internal = 0 for reference). The average energy values presented in Table S1, Supporting information helped

to qualitatively identify the chemical nature of these interactions. The results suggest that Ser42 interacts with the dG adducts primarily through van der Waals interactions.[33] Upon visual inspection of trajectories, Ser42 was found to interact with π cloud of aromatic moiety through C-H ⋯π interactions.[40] To verify this, the distance between C^{β} of Ser42 and the centroid of the N^2-modified aromatic moiety was plotted for all the complexes (Figure S14, Supporting information). The distances were in the appreciable range i.e. < 4 Å in all the complexes, favourable for van der Waals interactions. In fact, there is a correlation between this distance and the dihedral Xi (ξ) responsible for the particular orientation of the adducts (Figure S13). When the distance is between 3-4 Å, the orientation of aromatic moiety favours strong π-cloud interactions with the Ser42 residue.

Apart from Ser42, the other major interacting residues displayed similar non-covalent contacts. Phe76 played a crucial role in the complexes with high steric bulkiness, primarily by forming T-shaped C-H⋯ π interactions with the adducts (Figure 5 and 7). The Phe13 also exhibited C-H ⋯π interactions. This residue exhibited significant interactions in Anth complex compared to moderate interactions found in Bn, Anth, and Pyre complexes (Figure 5 and 7). Similarly, the C-H bond of alpha carbon (C^{α}) from Gly32 showed C-H⋯ π interactions in Naph and Pyre (Figure 7B and 7D) complexes.[40] The role of Tyr79 was found to be strategic as it interacts with Phe13 through parallel displaced stacking and with Phe76 through T-shaped interactions, which add strong hydrophobicity to the surroundings (Figure 5 and 7). Ile 31 and Pro73 were also found to form moderate to weak interactions in addition to weak interactions from Gly33 and Gly74 in all the complexes. The RMSF values for these residues in each of the complex are reported in Table S2. The RMSF values were mostly < 1 Å, indicating the importance of these residues in accommodating the adducts.

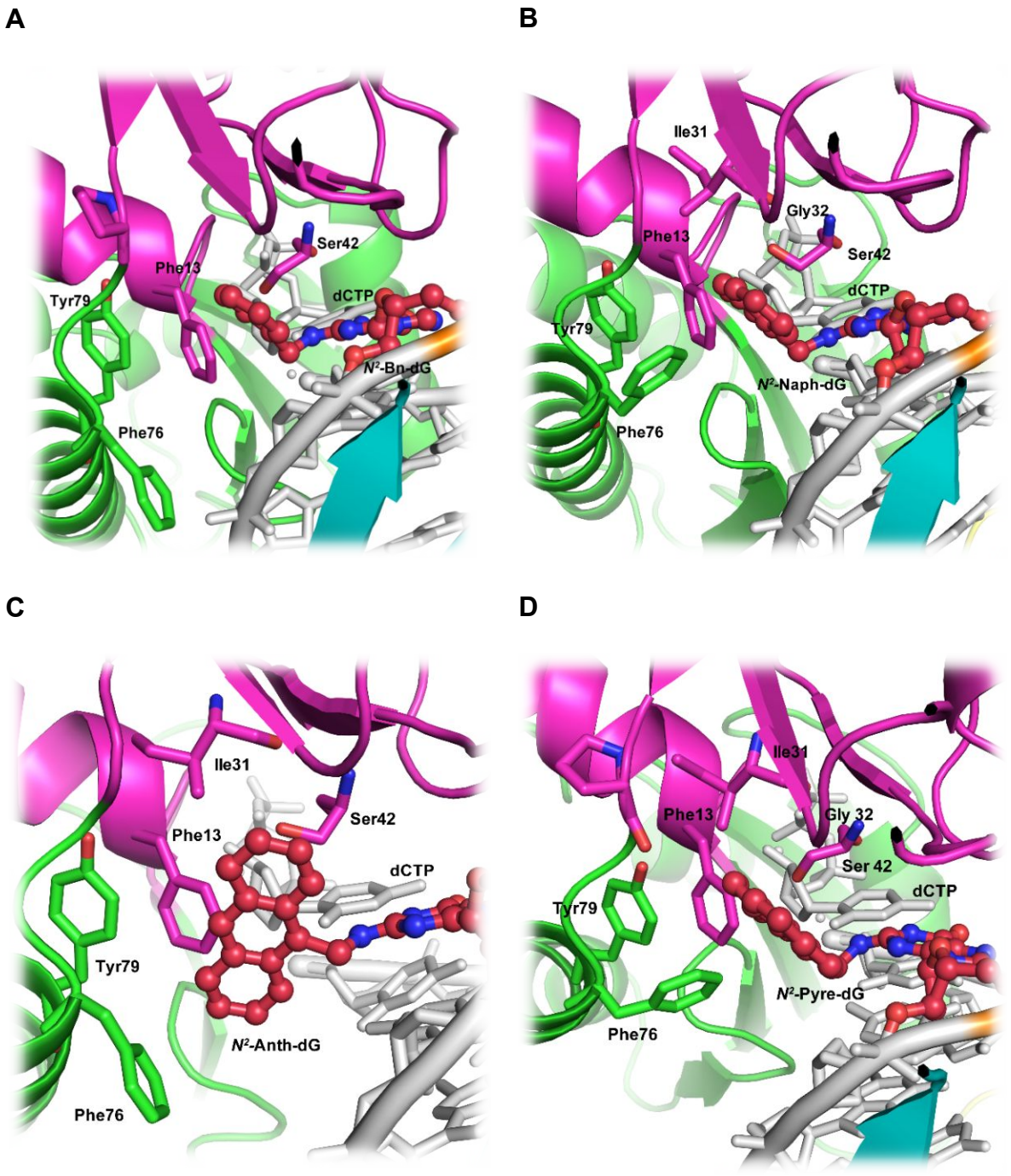


Figure 7: van der Waals interactions between AAs of Pol IV and N^2-adducts found in DNA-Pol IV ternary complexes. (A) N^2-Bn complex. (B) N^2-Naph complex. (C) N^2-Anth complex. (D) N^2-Pyre complex. Phe76 interacts with the N^2-adduct through T-shaped C-H ⋯ π interactions in the Bn, Naph, Anth and Pyre complexes. Carbon atom linked to O-H of Ser42 interacts with π cloud of Naph, Anth and Pyre moieties. The backbone atom of Gly32 as well interacts with adducts through C-H ⋯ π interactions in Naph and Pyre complexes. Phe13 as well forms T-shaped interactions with Anth moiety.

The representative structures of other ensembles of each complex had similar interaction patterns as discussed above. Overall, the van der Waals interactions are the major forces that

are responsible for the accommodation of N^2-modified moieties in the active site of Pol IV. The residues Phe13, Ile31, Gly32, Gly33, Ser42, Pro73, Gly74, Phe76 and Tyr79 were found to form a hydrophobic pocket[6] into which the adducts are accommodated and our current MD results support the same. The role of Ser42 was elucidated before[6,41] and our results from the MD simulations underscore the importance of this residue in stabilizing non-polar adducts through van der Waals interactions. General interaction patterns from Phe13, Phe76, Tyr79, and Gly32 were also identified. In summary, our MD studies provide insights on the molecular interactions that are responsible for the accommodation of N^2-dG adducts and their bypass by *E. coli* Pol IV, which is line with the results emerged from the experiments.

SUMMARY AND CONCLUSIONS

In conclusion, we have developed a robust and reliable protocol for the synthesis of N^2-dG modified DNAs. These modified DNAs were utilized to investigate the effect of steric bulkiness of the N^2-dG adducts on the TLS by Pol IV from *E. coli*. The Pol IV was found to be efficient in bypassing these adducts with high fidelity but with low processivity. Molecular modelling and dynamics studies were also carried out incorporating these N^2-dG modified DNAs at the active site of *E. coli.* Pol IV. The results revealed a common mechanism used by the enzyme that involves using its hydrophobic pocket at the active site to accommodate these modifications resulting in the efficient bypass of adducts. Further structural studies to validate these observations are currently being pursed.

EXPERIMENTAL SECTION

All chemicals and dry solvents were obtained from commercial sources and used without any further purification. Toluene, ACN, DCM, Et_3N, DIPEA, dioxane and pyridine were dried using calcium hydride. Thin layer chromatography (TLC) was performed on silica gel plates pre-coated with fluorescent indicator with visualization by UV light (at 260 nm) or by dipping into a solution of 5% conc. H_2SO_4 in EtOH (v/v) and heating. Silica gel (100−200 mesh) was used for column chromatography. ^{1}H NMR (500 or 400 MHz), ^{13}C NMR (125 or 100 MHz), ^{31}P NMR (202 or 162 MHz) were recorded on 500 MHz or 400 MHz instruments. Chemical shifts in parts per million (δ) are reported downfield from TMS (0 ppm) and referenced to the TMS signal or residual proton signal of the deuterated solvent as follows: CD_3OD (3.31 ppm) for ^{1}H NMR spectra, and $CDCl_3$ (77.2 ppm) or CD_3OD (49.1 ppm) for ^{13}C NMR spectra. Multiplicities of ^{1}H NMR spin couplings are reported as s for singlet, br s for broad singlet, d for doublet, t for triplet, br t for broad triplet, dd for doublet of doublet, ddd for doublet of doublet of doublets, quint for quintet or m for multiplet and overlapping spin systems. Values for apparent coupling constants (*J*) are reported in Hz. High resolution mass spectra (HRMS) were obtained from Q-TOF analyzer in positive ion electrospray ionization (ESI) mode. All the modified DNA sequences were synthesized using an automated solid phase synthesizer. Mass spectra of DNA oligonucleotides were obtained by MALDI (in positive linear or reflectron mode) or ESI (in negative mode). *E. coli* Pol IV was expressed and purified in house.[41]

General procedure for the Buchwald-Hartwig coupling and deacetylation

In a screw capped tube, $Pd(OAc)_2$ (0.1 equiv), (*R*)-BINAP (0.3 equiv) were taken and flushed with N_2 gas followed by addition of dry toluene (9.9 mL/mmol). Reaction mixture was stirred at room temperature for 5 min. Cs_2CO_3 (1.4 equiv), respective amines (1.1 equiv) and bromo nucleoside **1** (1 equiv) were added in the sequential manner. Further, the tube was purged with N_2 gas, sealed with a teflon-lined cap and heated in an oil bath at 85 °C. After completion of reaction, the reaction mixture was passed through a celite pad and washed with EtOAc (250 mL). The filtrate was concentrated to give a crude B-H coupled products. Further, the crude compound was subjected directly to deacetylation using 33% $MeNH_2$ in EtOH (v/v, 22 mL/mmol) and stirred at room temperature. After completion of the reaction, the mixture was concentrated under reduced pressure. The oily residue was purified by silica gel column chromatography to obtain deprotected nucleosides.

***N*^{2}-Benzyl-*O*^{6}-(2-(4-nitrophenyl)ethyl)-2'-deoxyguanosine (3a)**

$Pd(OAc)_2$ (5.94 mg, 0.02 mmol), (*R*)-BINAP (50 mg, 0.07 mmol), dry toluene (3 mL), Cs_2CO_3 (131 mg, 0.37 mmol), benzylamine (0.03 mL, 0.29 mmol) and bromo nucleoside **1** (150 mg, 0.26 mmol) were used and heated for 18 h to give a brown crude compound **2a** (156 mg, 0.26 mmol), which was further dissolved in 33% $MeNH_2$ in EtOH (5 mL) and stirred at room temperature for 3 h. Silica gel column chromatography (4% MeOH in DCM) yielded slightly brownish solid **3a** (76 mg, 57% after two steps). R_f = 0.47 (10% MeOH in DCM); mp: 60−63 °C; ^{1}H NMR (400 MHz, $CDCl_3$):δ 8.12–8.07 (m, 2H), 7.61 (s, 1H), 7.33 (d, *J* = 8.6 Hz, 2H), 7.31–7.21 (m, 5H), 6.19 (dd, *J* = 9.4, 5.6 Hz, 1H), 5.58 (br t, 1H), 4.70 (d, *J* = 5 Hz, 1H), 4.64 (t, *J* = 6.9 Hz, 2H), 4.60 (dd, *J* = 5.9, 2.4 Hz, 2H), 4.15 (s, 1H), 3.92 (dd, *J* = 12.6, 1.8 Hz, 1H), 3.71 (d, *J* = 11.3 Hz, 1H), 3.16 (t, *J* = 6.7 Hz, 2H), 2.97 (ddd, *J* = 14.2, 9.1, 5.1 Hz, 1H), 2.23 (dd, *J* = 12.6, 5.6 Hz, 1H); ^{13}C{1H} NMR (100 MHz, $CDCl_3$):δ 160.1, 158.5, 152.8, 146.9, 145.9, 139.3, 139.1, 130, 128.7, 127.3, 127.3, 123.8, 116.5, 89.1, 87.3, 73.3, 66.1, 63.5, 46.1,

40.2, 35.1; HRMS (ESI): Calcd for $C_{25}H_{27}N_6O_6$, $[M + H]^+$ 507.1992; found, $[M + H]^+$ 507.1980 (Δm −0.0012, error −2.4 ppm).

N^2-(1-Naphthylmethyl)-O^6-(2-(4-nitrophenyl)ethyl)-2'-deoxyguanosine (3b)

The $Pd(OAc)_2$ (11 mg, 0.04 mmol), (*R*)-BINAP (88 mg, 0.14 mmol), dry toluene (7 mL), Cs_2CO_3 (214 mg, 0.65 mmol), 1-naphthylmethylamine (0.07 mL, 0.51 mmol) and bromo nucleoside **1** (269 mg, 0.47 mmol) were used and heated for 21 h to give a dark brown crude compound **2b** (260 mg, 0.40 mmol), which was further dissolved in 33% $MeNH_2$ in EtOH (9 mL) and stirred at room temperature for 2.5 h. Silica gel column chromatography (4% MeOH in DCM) yielded yellow solid **3b** (171 mg, 64%, after two steps). R_f = 0.49 (10% MeOH in DCM); mp: 67−70 °C; ^{1}H NMR (400 MHz, $CDCl_3$):δ 8.09–8.02 (m, 1H), 7.97 (d, *J* = 8.3 Hz, 2H), 7.92–7.85 (m, 1H), 7.77 (d, *J* = 8.3 Hz, 1H), 7.60 (s, 1H), 7.57–7.36 (m, 6H), 7.16 (br s, 1H), 6.18 (dd, *J* = 9.2, 5.6 Hz, 1H), 5.51 (br s, 1H), 5.14–4.99 (m, 2H), 4.67 (d, *J* = 5.3 Hz, 1H), 4.60 (t, *J* = 7.0 Hz, 2H), 4.33 (s, 1H), 4.13 (s, 1H), 3.90 (d, *J* = 12 Hz, 1H), 3.68 (dd, *J* = 12.5, 1.8 Hz, 1H),3.07 (t, *J* = 6.7 Hz, 2H), 3.00 (ddd, *J* = 14, 9.5, 5.5 Hz, 1H), 2.22 (dd, *J* = 13.4, 5.6 Hz, 1H); ^{13}C{1H} NMR (100 MHz, $CDCl_3$):δ 161, 158.4, 153, 146.8, 145.6, 139.2, 134.2, 133.9, 131.3, 129.8, 129, 128.1, 126.4, 126, 125.5, 125.1, 123.7, 123.1, 116.4, 89, 87, 73, 66.1, 63.3, 44, 40.1, 35; HRMS (ESI): Calcd for $C_{29}H_{29}N_6O_6$, $[M + H]^+$ 557.2149; found, $[M + H]^+$ 557.2156 (Δm +0.0007, error +1.3 ppm).

N^2-(9-Anthracenylmethyl)-O^6-(2-(4-nitrophenyl)ethyl)-2'-deoxyguanosine (3c)

$Pd(OAc)_2$ (3.9 mg, 0.01 mmol), (*R*)-BINAP (33 mg, 0.05 mmol), dry toluene (2 mL), Cs_2CO_3 (87 mg, 0.24 mmol), 9-(aminomethyl)-anthracene[29] (41 mg, 0.19 mmol), and bromo nucleoside **1** (100 mg, 0.17 mmol) were used and heated for 8 h to give a dark brown crude compound **2c** (127 mg, 0.18 mmol), which was further dissolved in 33% $MeNH_2$ in EtOH (3.7 mL) and stirred at room temperature for 3 h. Column chromatography (3% MeOH in DCM) yielded slightly greenish compound **3c** (59 mg, 54% after two steps). R_f = 0.50 (10% MeOH in DCM);

mp: 83−86 °C; ^{1}H NMR (400 MHz, $CDCl_3$):δ 8.47 (s, 1H), 8.32–8.27 (m, 2H), 8.12 (d, J = 8.5 Hz, 2H), 8.03 (dd, J = 8.5, 1.0 Hz, 2H), 7.64 (s, 1H), 7.54–7.42 (m, 6H), 6.18 (br t, 1H), 5.57–5.47 (m, 2H), 5.18 (t, J = 4.4 Hz, 1H), 4.90–4.71 (m, 2H), 4.63 (br s, 1H), 4.07 (br s, 1H), 3.88–3.77 (m, 1H), 3.63–3.53 (m, 1H), 3.29 (br s, 2H), 3.12–2.92 (m, 1H), 2.25–2.13 (m, 1H); ^{13}C{1H} NMR (100 MHz, $CDCl_3$):δ 161.1, 158.3, 153.1, 146.9, 145.9, 139.2, 131.6, 130.6, 130, 129.3, 128.7, 128.2, 126.6, 125.2, 124.1, 123.8, 116.5, 88.8, 87, 73, 66.3, 63.2, 40.1, 38.9, 35.3; HRMS (ESI): Calcd for $C_{33}H_{31}N_6O_6$, $[M + H]^+$ 607.2300; found, $[M + H]^+$ 607.2297 (Δm −0.0003, error −0.5 ppm).

N^2-(1-Pyrenylmethyl)-O^6-(2-(4-nitrophenyl)ethyl)-2'-deoxyguanosine (3d)[25]

$Pd(OAc)_2$ (16 mg, 0.07 mmol), (*R*)-BINAP (132 mg, 0.21 mmol), dry toluene (7 mL), Cs_2CO_3 (323 mg, 0.99 mmol), 1-(aminomethyl)pyrene[30] (182 mg, 0.78 mmol), and bromo nucleoside **1** (403 mg, 0.71 mmol) were used and heated in oil bath for 21 h at 85 °C to get crude compound **2d** (547 mg, 0.76 mmol), which was further dissolved in 33% $MeNH_2$ in EtOH (15 mL) and stirred at room temperature for 1 h 15 min. Column chromatography on a silica gel (2.5% MeOH in DCM) yielded yellow compound **3d** (202 mg, 45% after two steps). R_f = 0.49 (10% MeOH in DCM); mp: 75−78 °C; ^{1}H NMR (400 MHz, MeOH-d_4 + $CDCl_3$):δ 8.24 (d, J = 9.1 Hz, 1H), 8.16–8.09 (m, 2H), 8.08–7.89 (m, 7H), 7.72 (s, 1H), 7.66 (br s, 2H), 6.86 (br s, 1H), 6.22 (dd, J = 8.9, 5.9 Hz, 1H), 5.23 (s, 2H), 4.55 (d, J = 5.4 Hz, 1H), 4.38 (t, J = 6.6 Hz, 2H), 4.09 (d, J = 1.3 Hz, 1H), 3.88 (d, J = 12.4 Hz, 1H), 3.73–3.66 (m, 2H), 3.58–3.49 (m, 1H), 2.88 (ddd, J = 14.5, 9.1, 5.6 Hz, 1H), 2.30–2.20 (m, 1H); ^{13}C{1H} NMR (100 MHz, MeOH-d_4 + $CDCl_3$):δ 160.9, 158.5, 152.9, 146.4, 145.5, 139.1, 132.4, 131.3, 130.7, 129.4, 128.4, 127.9, 127.4, 127.2, 126.2, 125.4, 125.2, 124.9, 124.8, 123.4, 122.5, 115.9, 88.8, 87, 72.5, 65.9, 63.2, 44, 40.1, 34.7; HRMS (ESI): Calcd for $C_{35}H_{31}N_6O_6$, $[M + H]^+$ 631.2300; found, $[M + H]^+$ 631.2293 (Δm −0.0007, error −1 ppm).

General procedure for the DMT-protection

The N^2-modified-dG nucleosides (1 equiv) was co-evaporated with dry pyridine (5 mL) and dissolved in the same solvent (10 mL/mmol). To this, DMT-Cl (1.22 equiv) was added and stirred at room temperature. Further, the reaction mixture was diluted with DCM (100 mL) and washed with saturated $NaHCO_3$ (40 mL) and water (2 × 50 mL). DCM layer was dried over Na_2SO_4 and evaporated. The crude DMT-protected compound was purified by silica gel column chromatography (using appropriate solvents along with 2% Et_3N) to afford pure compound.

N^2-Benzyl-O^6-(2-(4-nitrophenyl)ethyl)-5'-(4,4'-dimethoxytrityl)-2'-deoxyguanosine (4a)

Compound **3a** (251 mg, 0.49 mmol) in dry pyridine (5 mL), DMT-Cl (332 mg, 0.99 mmol) were stirred for 24 h. Column chromatography (90% DCM in pet. ether + 2% Et_3N) yielded compound **4a** as a pale yellow solid (237 mg, 60%). R_f = 0.55 (2% MeOH in DCM + 2% Et_3N); mp: 71−74 °C; ^{1}H NMR (400 MHz, $CDCl_3$):δ 8.13–8.09 (m, 2H), 7.68 (s, 1H), 7.39 (d, J = 7.3 Hz, 4H), 7.31–7.26 (m, 8H), 7.25–7.14 (m, 4H), 6.78 (d, J = 8.0 Hz, 4H), 6.27 (t, J = 6.7 Hz, 1H), 5.17 (t, J = 5.5 Hz, 1H), 4.66 (t, J = 6.8 Hz, 2H), 4.63–4.58 (m, 1H), 4.53 (t, J = 5.8 Hz, 2H), 4.09–4.04 (m, 1H), 3.76 (s, 6H), 3.38 (dd, J = 9.8, 4.8 Hz, 1H), 3.34 (dd, J = 10.1, 5.4 Hz, 1H), 3.20 (t, J = 6.8 Hz, 2H), 2.87–2.81 (m, 1H), 2.38 (ddd, J = 13.2, 6.2, 4 Hz, 1H); ^{13}C{1H} NMR (100 MHz, $CDCl_3$):δ 160.6, 158.8, 158.7, 153.8, 146.9, 146.1, 144.7, 139.6, 137.9, 135.8, 130.1, 130, 128.6, 128.2, 128, 127.4, 127.2, 127, 123.8, 115.5, 113.3, 86.6, 85.9, 84, 72.8, 66, 64.1, 55.3, 46.9, 39.5, 35.3. HRMS (ESI): Calcd for $C_{46}H_{45}N_6O_8$, $[M + H]^+$ 809.3299; found, $[M + H]^+$ 809.3329 (Δm +0.003, error +3.7 ppm).

N^2-(1-Naphthylmethyl)-O^6-(2-(4-nitrophenyl)ethyl)-5'-(4,4'-dimethoxytrityl)-2'-deoxyguanosine (4b)

Compound **3b** (171 mg, 0.30 mmol) in dry pyridine (3 mL), DMT-Cl (126 mg, 0.37 mmol) were stirred for 10 h. Column chromatography (1% MeOH in DCM + 2% Et_3N) yielded compound **4b** as a yellow solid (154 mg, 60%). R_f = 0.50 (2% MeOH in DCM + 2% Et_3N);

mp: 85−88 °C; ^{1}H NMR (400 MHz, $CDCl_3$):δ 8.08–7.98 (m, 3H), 7.91–7.86 (m, 1H), 7.77 (d, *J* = 8.0 Hz, 1H), 7.70 (s, 1H), 7.55–7.46 (m, 2H), 7.44–7.35 (m, 4H), 7.30–7.26 (m, 5H), 7.26–7.13 (m, 4H), 6.80–6.72 (m, 4H), 6.29 (t, *J* = 6.6 Hz, 1H), 5.24–5.17 (m, 1H), 5.06–4.88 (m, 2H), 4.63 (t, *J* = 6.9 Hz, 2H), 4.56 (br s, 1H), 4.08–4.01 (m, 1H), 3.74 (d, *J* = 1.3 Hz, 6H), 3.68–3.63 (m, 1H), 3.39 (dd, *J* = 10, 5 Hz, 1H), 3.33 (dd, *J* = 9.8, 4.8 Hz, 1H), 3.14 (br s, 2H), 2.85–2.74 (m, 1H), 2.38 (ddd, *J* = 13.5, 6.5, 4.5 Hz, 1H); ^{13}C{1H} NMR (100 MHz, $CDCl_3$):δ 160.7, 158.7, 158.6, 146.8, 146, 144.7, 137.9, 135.8, 134.9, 134.5, 133.9, 131.4, 130.1, 129.8, 129, 128.2, 128.1, 128, 127, 126.5, 126, 125.6, 123.8, 123.4, 115.7, 113.3, 86.6, 85.8, 84, 72.7, 66, 64.1, 55.3, 44.1, 39.6, 35.2; HRMS (ESI): Calcd for $C_{50}H_{47}N_6O_8$, $[M + H]^+$ 859.3455; found, $[M + H]^+$ 859.3478 (Δ*m* +0.0023, error +2.7 ppm).

N^2-(9-Anthracenylmethyl)-O^6-(2-(4-nitrophenyl)ethyl)-5'-(4,4'-dimethoxytrityl)-2'-deoxyguanosine (4c)

Compound **3c** (162 mg, 0.26 mmol) in dry pyridine (3 mL), DMT-Cl (180 mg, 0.53 mmol) were stirred for 10 h. Column chromatography (1% MeOH in DCM + 2% Et_3N) yielded compound **4c** as a slightly greenish solid (177 mg, 72%). R_f = 0.50 (2% MeOH in DCM + 2% Et_3N); mp: 86−89 °C; ^{1}H NMR (500 MHz, $CDCl_3$):δ 8.46 (s, 1H), 8.29–8.24 (m, 2H), 8.08 (d, *J* = 7.9 Hz, 2H), 8.05–8.00 (m, 2H), 7.75 (s, 1H), 7.50–7.36 (m, 8H), 7.28 (dd, *J* = 5.8, 2.7 Hz, 3H), 7.20–7.15 (m, 3H), 7.12–7.08 (m, 1H), 6.73 (d, *J* = 7.3 Hz, 4H), 6.46–6.28 (m, 1H), 5.51–5.36 (m, 2H), 5.14 (t, *J*= 4.0 Hz, 1H), 4.74–4.59 (m, 2H), 4.09 (br s, 1H), 3.70 (d, *J* = 1.5 Hz, 6H), 3.46–3.40 (m, 1H), 3.38–3.30 (m, 1H), 3.26–3.15 (br s, 2H), 2.91–2.85 (m, 1H), 2.52–2.44 (m, 1H); ^{13}C{1H} NMR (125 MHz, $CDCl_3$):δ 160.7, 158.6, 153.9, 146.8, 146.1, 144.6, 137.7, 135.8, 135.7, 131.6, 130.5, 130.1, 130, 130, 129.4, 129.3, 129.1, 128.3, 128.2, 128, 127, 126.5, 125.2, 124.2, 123.8, 115.5, 113.2, 86.6, 85.9, 83.9, 72.8, 66, 64.1, 55.3, 45.9, 38.7, 35.3; HRMS (ESI): Calcd for $C_{54}H_{49}N_6O_8$, $[M + H]^+$ 909.3606; found, $[M + H]^+$ 909.3608 (Δ*m* +0.0002, error +0.1 ppm).

N^2-(1-Pyrenylmethyl)-O^6-(2-(4-nitrophenyl)ethyl)-5'-(4,4'-dimethoxytrityl)-2'-deoxyguanosine (4d)

Compound **3d** (151 mg, 0.24 mmol) in dry pyridine (3 mL), DMT-Cl (102 mg, 0.30 mmol) were stirred for 12 h. Column chromatography (1% MeOH in DCM + 2% Et_3N) yielded compound **4d** as a pale yellow solid (178 mg, 76%). R_f = 0.55 (2% MeOH in DCM + 2% Et_3N); mp: 77−80 °C; ^{1}H NMR (400 MHz, $CDCl_3$):δ 8.25 (d, *J* = 9.3 Hz, 1H), 8.22–8.16 (m, 2H), 8.10–7.86 (m, 7H), 7.71 (s, 1H), 7.42–7.36 (m, 2H), 7.32–7.26 (m, 5H), 7.25–7.10 (m, 5H), 6.78–6.72 (m, 4H), 6.30 (t, *J* = 6.5 Hz, 1H), 5.37–5.32 (m, 1H), 5.27–5.13 (m, 2H), 4.64–4.50 (m, 3H), 4.09–4.03 (m, 1H), 3.70 (d, *J* = 2.5 Hz, 6H), 3.39 (dd, *J* = 9.9, 4.6 Hz, 1H), 3.35 (dd, *J* = 10, 5.2 Hz, 1H), 3.04 (br s, 2H), 2.87–2.77 (m, 1H), 2.37 (ddd, *J* = 13.4, 6.4, 4.0 Hz, 1H); ^{13}C{1H} NMR (100 MHz, $CDCl_3$):δ 160.7, 158.7, 158.6, 153.9, 146.6, 145.9, 144.7, 137.9, 135.8, 132.4, 131.4, 130.8, 130.1, 129.7, 128.2, 128, 127.5, 127.4, 127, 126.2, 125.5, 125.3, 125, 124.9, 124.8, 123.6, 122.8, 115.7, 113.3, 86.6, 85.9, 84, 72.8, 65.9, 64.1, 55.3, 44.3, 39.6, 35.1; HRMS (ESI): Calcd for $C_{56}H_{48}N_6NaO_8$, $[M + Na]^+$ 955.3426; found, $[M + Na]^+$ 955.3431 (Δ*m* +0.0005, error +0.6 ppm).

General procedure for the phosphitylation

DMT protected nucleoside (1 equiv) was dissolved in DCM (10 mL/mmol) followed by addition of DIPEA (8 equiv) and CEP-Cl (2 equiv). Reaction mixture stirred at room temperature under N_2 atmosphere. After completion, the reaction was quenched by addition of trace amount of MeOH and stirred for 5 to 10 min. Further, 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. Crude compound was purified by silica gel column chromatography (using appropriate solvents along with 2% Et_3N) to obtain phosphitylated compound.

N^2-Benzyl-O^6-(2-(4-nitrophenyl)ethyl)5'-*O*-(4,4'-dimethoxytrityl)-3'-(2-cyanoethyl

diisopropylphosphoramidite)-2'-deoxyguanosine (5a)

Nucleoside **4a** (150 mg, 0.18 mmol), DCM (2 mL), DIPEA (0.25 mL, 1.48 mmol) and CEP-Cl (127 mg, 0.54 mmol) were used and stirred for 2 h. Further, MeOH (0.5 mL) was added and stirred for additional 10 min. Column chromatography (35% DCM in pet. ether + 2% Et_3N) yielded compound **5a** as a pale yellow solid (111 mg, 61%). R_f= 0.65 (90% DCM in pet. ether + 2% Et_3N); mp: 58−61 °C; ^{31}P NMR (162 MHz, $CDCl_3$):δ 148.81, 148.63; HRMS (ESI): Calcd for $C_{55}H_{62}N_8O_9P$, $[M + H]^+$ 1009.4377; found, $[M + H]^+$ 1009.4348 (Δ*m* −0.0029, error −3 ppm).

***N*^{2}-(1-Naphthylmethyl)-*O*^{6}-(2-(4-nitrophenyl)ethyl)5'-*O*-(4,4'-dimethoxytrityl)-3'-(2-cyanoethyldiisopropylphosphoramidite)-2'-deoxyguanosine (5b)**

The tritylated nucleoside **4b** (150 mg, 0.17 mmol), DCM (2 mL), DIPEA (0.25 mL, 1.36 mmol) and CEP-Cl (83 mg, 0.34 mmol) were used and stirred for 2 h 15 min. Further, MeOH (0.5 mL) was added and stirred for additional 10 min. Column chromatography (35% DCM in pet. ether + 2% Et_3N) yielded compound **5b** as a pale yellow solid (74 mg, 41%). R_f= 0.65 (75% DCM in pet. ether + 2% Et_3N); mp: 70−73 °C; ^{31}P NMR (162 MHz, $CDCl_3$):δ 148.72, 148.57; HRMS (ESI): Calcd for $C_{59}H_{64}N_8O_9P$, $[M + H]^+$ 1059.4528; found, $[M + H]^+$ 1059.4529 (Δ*m* +0.00009, error +0.1 ppm).

***N*^{2}-(9-Anthracenylmethyl)-*O*^{6}-(2-(4-nitrophenyl)ethyl)5'-*O*-(4,4'-dimethoxytrityl)-3'-(2-cyanoethyldiisopropylphosphoramidite)-2'-deoxyguanosine (5c)**

Nucleoside **4c** (170 mg, 0.18 mmol), DCM (2 mL), DIPEA (0.26 mL, 1.49 mmol) and CEP-Cl (88 mg, 0.37 mmol) were used and stirred for 1 h. Further, MeOH (0.5 mL) was added and stirred for additional 10 min. Column chromatography (40% DCM in pet. ether + 2% Et_3N) yielded compound **5c** as a slightly greenish solid (75 mg, 37%). R_f = 0.60 (75% DCM in pet. ether + 2% Et_3N); mp: 77−80 °C; ^{31}P NMR (202 MHz, $CDCl_3$):δ 148.85, 148.72; HRMS (ESI):

Calcd for $C_{63}H_{66}N_8O_9P$ $[M + H]^+$1109.4685; found, $[M + H]^+$1109.4686 (Δm +0.00009, error +0.1 ppm).

N^2-(1-Pyrenylmethyl)-O^6-(2-(4-nitrophenyl)ethyl)5'-*O*-(4,4'-dimethoxytrityl)-3'-(2-cyanoethyldiisopropylphosphoramidite)-2'-deoxyguanosine (5d)

Nucleoside **4d** (200 mg, 0.21 mmol), DCM (2.5 mL), DIPEA (0.3 mL, 1.68 mmol) and CEP-Cl (100 mg, 0.42 mmol) were used and stirred for 1 h 15 min. Further, MeOH (0.5 mL) was added and stirred for additional 10 min. Column chromatography (45% DCM in pet. ether + 2% Et_3N) yielded compound **5d** as a pale yellow solid (93 mg, 40%). R_f= 0.66 (75% DCM in pet. ether + 2% Et_3N); mp: 80−83 °C; ^{31}P NMR (202 MHz, $CDCl_3$):δ 148.72; HRMS (ESI): Calcd for $C_{65}H_{66}N_8O_9P$, $[M + H]^+$ 1133.4685; found, $[M + H]^+$1133.4690 (Δm +0.0005, error +0.4 ppm).

Solid phase synthesis, deprotection and purification of N^2-dG modified oligonucleotides

All DNA sequences (Table 1) were synthesized on 1 μmol scale using appropriate controlled pore glass (CPG) solid supports. Unmodified and modified phosphoramidite building blocks were dried under P_2O_5 for 24 h, and then dissolved in anhydrous ACN and stored at 4 °C under molecular sieves (4 Å) for 10 h. Concentration used for the unmodified phosphoramidite was 0.067 M and for modified phosphoramidite was 0.1 M. Coupling time used for unmodified phosphoramidite was 2 min. and for the N^2-dG modified phosphoramidites was 6 min. After solid phase synthesis, deprotection of N^2-Bn, N^2-Naph, N^2-Anth and N^2-Pyre-dG modified oligos were carried out in four different steps. Initially, the deprotection of cyanoethyl group of the modified DNA sequences were carried out using 10% diethylamine in ACN (800 μl, v/v) for 5 min. Supernatant layer was collected and solid support was washed with ACN (2 × 400 μl). The CPG beads was air dried (3 h) and treated with 1 M DBU in ACN (1 mL, v/v) for 1 h at room temperature to remove NPE group.[23] After 1 h, the supernatant was removed and the CPG beads were washed with MeOH (2 × 1 mL) and ACN (3 × 1 mL). Furthermore, the

modified DNAs intact with CPG was treated with 1 mL 30% aq. NH_3 (v/v) for 3 h for the cleavage of solid support. To remove the base protecting groups, the supernatant solution was evaporated and again re-suspended in 1 mL 30% aq. NH_3 (v/v) and heated at 55 °C for 16 h. Further, ammonia layer was evaporated to obtain the pellet of crude DNAs, which were purified by using 20% (7 M urea) denaturing PAGE (30 W, 3 h) with 1 X TBE running buffer (89 mM each Tris and boric acid and 2 mM EDTA, pH ~ 8.3). The gel thickness was 1 mm, and gel dimension was 20 × 30 cm. The desired oligo bands were visualized and marked under the UV lamp (260 nm). Further, the gel bands were cut and crushed into fine particles. The crushed gels were extracted three times with 15 mL of TEN buffer (10 mM Tris, 1 mM EDTA, 300 mM NaCl, pH ~ 8.0) at 37 °C. Finally, desalting of oligos were carried out using Sep-Pak column and the concentrations of N^2-dG modified DNAs were measured at 260 nm in UV-VIS spectrophotometer using the appropriate molar extinction coefficients (ε).

5'-Radiolabeling of primers

The radiolabeling reaction mixture contains the DNA (25 pmol), T4 polynucleotide kinase (PNK) enzyme (5 U) and [γ-^{32}P] ATP (20 μCi) in 1 X PNK buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, 5 mM DTT, 0.1 mM spermidine, 0.1 mM EDTA, pH 7.6). The total reaction volume was 10 μL. The reaction mixture was incubated at 37 °C for 1 h followed by deactivation of PNK enzyme by heating at 70 °C for 3 min. The radiolabeled DNAs (primers and 50-mer DNA standard) were purified using nucleotide removal kit using the protocol provided by the manufacturer.

Primer extension assay

The primer extension reactions were carried out using the mixture of 1:1.2 molar ratios of the primers with the unmodified and N^2-dG modified DNA templates. Initially, the primer-template mixture was annealed by heating at 95 °C for 3 min, followed by slow cooling to room temperature for 30 min. After annealing, the primer extension reactions were performed

with primer-template duplex in a total volume of 20 µl, containing 40 nM primer, traces of 5'-radiolabelled primer, 50 nM template, 50 µg/mL of BSA, 0.1 mM of ammonium sulphate, 2.5 mM $MgCl_2$, 4 µl 5X assay buffer (125 mM Tris-Cl, 5 mM dithiothreitol, pH 8.0), Pol IV (4 nM), and individual dNTP or mixture of dNTPs (10 µM). The primer extension reactions were incubated at 37 °C for 5 min, and the reactions were initiated by the addition of the desired amount of dNTPs. The reactions were conducted for the indicated amount of time, aliquots were withdrawn 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 20% denaturing PAGE (7 M urea), auto radiogram was generated and the bands were quantified by ImageQuantTL software. The running start experiments were carried out using 15-mer and 11-mer primers employing a mixture of dNTPs (250 µM). The reaction conditions were same as described previously. The total reaction volume was contains 40 nM primer, traces of 5'-radiolabelled primer, 50 nM of each template. Finally, the reactions were terminated and analysed as mentioned previously.

Molecular modelling and dynamics studies

The modified nucleotides are constructed using GaussView and submitted to RED Server *Development*[42–45] for the calculation of restraint electrostatic potential (RESP) charges and parameters (bond, angle and dihedral) required for the simulation (Figure S15-S18, Supporting information). Originally reported crystal structure (PDB ID: 4Q43) contains chemically modified dCTP at the active site which was required to prevent catalysis during crystallization. The modified dCTP is replaced with normal dCTP and previously calculated bond, angle and dihedral parameters were used (http://upjv.q4md-forcefieldtools.org/REDDB/projects/F-90/). The amber force fields *ff*99SB[46] and parmbsc0[47] were used for protein and DNA respectively. The parameters for dCTP, and modified nucleotides along with parameters for protein, DNA and ions were fed to Leap program of AmberTools package to prepare input files. The Mg^{2+}

ions reported in the original crystal structure were retained. Additional potassium ions were added to neutralize the system. Each system was then immersed in TIP3P water box with water molecules extending up to 10 Å. Then each complex was minimized in two stages to remove any atomic clashes.

Energy minimization was carried out using SANDER module of Amber 14. In the first step, the N^2-dG adduct was left free and minimized for 10000 steepest descent energy cycles, whereas in the second step the N^2-dG adduct was restricted and minimized for 10000 steepest descent energy minimizations. Later the system was heated from 0.1 K to close to 300K over 100 ps with constant volume. Then the complexes were heated to final temperature (300K) at constant pressure. During the heating, the DNA was restricted with a restraint of 10 kcal/mol-Å^2. This restriction was slowly removed over 8 stages of equilibration (with restraints of 10, 8, 5, 4, 3, 2, 1, 0.5 kcal/mol-Å^2 30 ps each). Each equilibration was carried under constant pressure and temperature (NPT) conditions. The pressure was maintained at 1 atmosphere using Bendersen weak-coupling barostat with a pressure relaxation time of 2 ps. The temperature was maintained at 300K using Langvein dynamics. The equilibration was followed by energy minimization (10000 steepest descent cycles) and the minimized structure was used as starting structure for the subsequent equilibration. After 8 stages, 100 ps of equilibration was run under NPT conditions without any restrictions before proceeding to the production run. 100 ns of unrestrained dynamics (under NPT conditions) was performed on each complex using GPU accelerated version of Particle Mesh Ewald Molecular Dynamics (PMEMD) implemented in Amber 14.[39,48–50] A cut-off value of 10 Å is set for calculations of non-bonded interactions by Particle Mesh Ewald (PME) method. SHAKE algorithm was used to constraint all the bonds involving hydrogen atoms with an integration time step of 2fs. The output coordinate files were saved for every 2 ps and a total of 50 000 snapshots (100 ns) were recorded for each complex.

The trajectories were visualized using UCSF Chimera.[51] CPPTRAJ module of Amber 14 was used for calculations of dihedrals, RMSD values, RMSF values, H-bond occupancies and for clustering of dynamics. Only heavy atoms were considered for the calculation of RMSD values. Distances between Ser and N^2-adducts were calculated using per frame analysis in UCSF Chimera. Distance of 3.0 Å and angle of 150° were used as cut-off values for H-bond calculations. Cut-off distance of 4.0 Å is used for van der Waals interactions. Clustering analysis was carried out employing hierarchical agglomerative approach using mass weighed RMSD. Interaction energies are calculated using MM/GB-SA[52] method implemented as the program MMPBSA.py[53] in AmberTools package. Images are rendered using PyMol (www.pymol.org).

ASSOCIATED CONTENT

Supporting Information

PAGEs of nucleotide incorporation and full length extension assays; additional Figures and Tables from molecular modelling and dynamics studies; ^{1}H, ^{13}C, ^{31}P NMR of all the new compounds; and representative MALDI spectra of N^2-dG modified oligonucleotides. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

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

ACKNOWLEDGEMENTS

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. We also thank Dr. S. Harikrishna for his suggestions on MD studies and Saurja Dasgupta for critically reading the manuscript. Computer Centre, IIT Bombay is gratefully acknowledged for providing high performance computing facilities. P.P.G. thank Ms. Sushree P. Pany for acquiring the mass spectra of phosphoramidites. P.P.G. also thank the Council of Scientific and Industrial Research (CSIR) for the Ph. D. fellowship and P.B. thanks Department of Science and Technology (DST) for the INSPIRE Fellowship.

DEDICATION

Dedicated to Professor Krishna N. Ganesh on the occasion of his 65th birthday.

REFERENCES

(1) Moorthy, B.; Chu, C.; Carlin, D. J. Polycyclic Aromatic Hydrocarbons: From Metabolism to Lung Cancer. *Toxicol. Sci.* **2015**, *145*, 5–15.
(2) Dipple, A. DNA Adducts of Chemical Carcinogens. *Carcinogenesis* **1995**, *16*, 437–441.
(3) Hemminki, K. Nucleic Acid Adducts of Chemical Carcinogens and Mutagens. *Arch. Toxicol.* **1983**, *52*, 249–285.
(4) Sale, J. E.; Lehmann, A. R.; Woodgate, R. Y-Family DNA Polymerases and their Role in Tolerance of Cellular DNA Damage. *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 141–152.
(5) Jarosz, D. F.; Godoy, V. G.; Delaney, J. C.; Essigmann, J. M.; Walker, G. C. A Single Amino Acid Governs Enhanced Activity of DinB DNA Polymerases on Damaged Templates. *Nature* **2006**, *439*, 225–228.
(6) Kottur, J.; Sharma, A.; Gore, K. R.; Narayanan, N.; Samanta, B.; Pradeepkumar, P. I.; Nair, D. T. Unique Structural Features in DNA Polymerase IV Enable Efficient Bypass of the N2 Adduct Induced by the Nitrofurazone Antibiotic. *Structure* **2015**, *23*, 56–67.
(7) Ikeda, M.; Furukohri, A.; Philippin, G.; Loechler, E.; Akiyama, M. T.; Katayama, T.; Fuchs, R. P.; Maki, H. DNA Polymerase IV Mediates Efficient and Quick Recovery of Replication Forks Stalled at N2-dG Adducts. *Nucleic Acids Res.* **2014**, *42*, 8461–8472.
(8) Moschel, R. C.; Robert Hudgins, W.; Dipple, A. Aralkylation of Guanosine by the Carcinogen N-Nitroso-N-Benzylurea. *J. Org. Chem.* **1980**, *45*, 533–535.
(9) Peterson, L. A. N-Nitrosobenzylmethylamine Is Activated to a DNA Benzylating Agent in Rats. *Chem. Res. Toxicol.* **1997**, *10*, 19–26.
(10) Moon, K.-Y.; Kim, Y. S. Synthesis and Characterization of Oligonucleotides Containing Site-Specific Bulky N2-Aralkylated Guanines and N6-Aralkylated Adenines. *Arch. Pharm. Res.* **2000**, *23*, 139–146.
(11) Kou, Y.; Koag, M.-C.; Lee, S. Structural and Kinetic Studies of the Effect of Guanine-N7 Alkylation and Metal Cofactors on DNA Replication. *Biochemistry* **2018**, *57*, 5105–5116.
(12) Flesher, J. W.; Horn, J.; Lehner, A. F. 9-Sulfooxymethylanthracene Is an Ultimate

Electrophi. *Biochem. Biophys. Res. Commun.* **1998**, *251*, 239–243.
(13) Casale, R.; McLaughlin, L. W. Synthesis and Properties of an Oligodeoxynucleotide Containing a Polycyclic Aromatic Hydrocarbon Site Specifically Bound to the N2-Amino Group of a 2'-Deoxyguanosine Residue. *J. Am. Chem. Soc.* **1990**, *112*, 5264–5271.
(14) Bendadani, C.; Meinl, W.; Monien, B. H.; Dobbernack, G.; Glatt, H. The Carcinogen 1-Methylpyrene Forms Benzylic DNA Adducts in Mouse and Rat Tissues in Vivo via a Reactive Sulphuric Acid Ester. *Arch. Toxicol.* **2014**, *88*, 815–821.
(15) Bendadani, C.; Meinl, W.; Monien, B.; Dobbernack, G.; Florian, S.; Engst, W.; Nolden, T.; Himmelbauer, H.; Glatt, H. Determination of Sulfotransferase Forms Involved in the Metabolic Activation of the Genotoxicant 1-Hydroxymethylpyrene Using Bacterially Expressed Enzymes and Genetically Modified Mouse Models. *Chem. Res. Toxicol.* **2014**, *27*, 1060–1069.
(16) Guengerich, F. P. Interactions of Carcinogen-Bound DNA with Individual DNA Polymerases. *Chem. Rev.* **2006**, *106*, 420–452.
(17) Delaney, J. C.; Essigmann, J. M. Biological Properties of Single Chemical-DNA Adducts: A Twenty Year Perspective. *Chem. Res. Toxicol.* **2008**, *21*, 232–252.
(18) DeCorte, B. L.; Tsarouhtsis, D.; Kuchimanchi, S.; Cooper, M. D.; Horton, P.; Harris, C. M.; Harris, T. M. Improved Strategies for Postoligomerization Synthesis of Oligodeoxynucleotides Bearing Structurally Defined Adducts at the N2 Position of Deoxyguanosine. *Chem. Res. Toxicol.* **1996**, *9*, 630–637.
(19) Steinbrecher, T.; Wameling, C.; Oesch, F.; Seidel, A. Activation of the C2 Position of Purine by the Trifluoromethanesulfonate Group: Synthesis of N2-Alkylated Deoxyguanosines. *Angew. Chem. Int. Ed. English* **1993**, *32*, 404–406.
(20) He, G. X.; Krawczyk, S. H.; Swaminathan, S.; Shea, R. G.; Dougherty, J. P.; Terhorst, T.; Law, V. S.; Griffin, L. C.; Coutré, S.; Bischofberger, N. N2- and C8-Substituted Oligodeoxynucleotides with Enhanced Thrombin Inhibitory Activity in Vitro and in Vivo. *J. Med. Chem.* **1998**, *41*, 2234–2242.
(21) Moon, K. Y.; Moschel, R. C. Effect of Ionic State of 2'-Deoxyguanosine and Solvent on Its Aralkylation by Benzyl Bromide. *Chem. Res. Toxicol.* **1998**, *11*, 696–702.
(22) Sergueeva, Z. A.; Sergueev, D. S.; Shaw, B. R. Rapid and Selective Reduction of Amide Group by Borane-Amine Complexes in Acyl Protected Nucleosides. *Nucleosides Nucleotides Nucleic Acids* **2000**, *19*, 275–282.
(23) Choi, J.; Guengerich, F. P. Analysis of the Effect of Bulk at N2-Alkylguanine DNA Adducts on Catalytic Efficiency and Fidelity of the Processive DNA Polymerases Bacteriophage T7 Exonuclease- and HIV-1 Reverse Transcriptase. *J. Biol. Chem.* **2004**, *279*, 19217–19229.
(24) Bonala, R. R.; Shishkina, I. G.; Johnson, F. Synthesis of Biologically Active N2-Amine Adducts of 2'-Deoxyguanosine. *Tetrahedron Lett.* **2000**, *41*, 7281–7284.
(25) Lee, H.; Luna, E.; Hinz, M.; Stezowski, J. J.; Kiselyov, A. S.; Harvey, R. G. Synthesis of Oligonucleotide Adducts of the Bay Region Diol Epoxide Metabolites of Carcinogenic Polycyclic Aromatic Hydrocarbons. *J. Org. Chem.* **1995**, *60*, 5604–5613.
(26) Kath, J. E.; Jergic, S.; Heltzel, J. M. H.; Jacob, D. T.; Dixon, N. E.; Sutton, M. D.; Walker, G. C.; Loparo, J. J. Polymerase Exchange on Single DNA Molecules Reveals Processivity Clamp Control of Translesion Synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 7647–7652.
(27) Ghodke, P. P.; Harikrishna, S.; Pradeepkumar, P. I. Synthesis and Polymerase-Mediated Bypass Studies of the N2-Deoxyguanosine DNA Damage Caused by a Lucidin Analogue. *J. Org. Chem.* **2015**, *80*, 2128–2138.
(28) Champeil, E.; Pradhan, P.; Lakshman, M. K. Palladium-Catalyzed Synthesis of

Nucleoside Adducts from Bay- and Fjord-Region Diol Epoxides. *J. Org. Chem.* **2007**, *72*, 5035–5045.

(29) Lu, D.; Lei, J.; Tian, Z.; Wang, L.; Zhang, J. $Cu2^{+}$ Fluorescent Sensor Based on Mesoporous Silica Nanosphere. *Dye. Pigment.* **2012**, *94*, 239–246.

(30) Abiraj, K.; Gowda, D. C. Zinc/Ammonium Formate: A New Facile System for the Rapid and Selective Reduction of Oximes to Amines. *J. Chem. Res.* **2003**, *2003*, 332–333.

(31) Dai, Q.; Zheng, G.; Schwartz, M. H.; Clark, W. C.; Pan, T. Selective Enzymatic Demethylation of N2,N2-Dimethylguanosine in RNA and Its Application in High-Throughput TRNA Sequencing. *Angew. Chem. Int. Ed. Engl.* **2017**, *56*, 5017–5020.

(32) Himmelsbach, F.; Schulz, B. S.; Trichtinger, T.; Charubala, R.; Pfleiderer, W. The P-Nitrophenylethyl (NPE) Group. *Tetrahedron* **1984**, *40*, 59–72.

(33) Kottur, J.; Ghodke, P. P.; Nair, D. T.; Pradeepkumar, P. I. Manuscript under Preparation.

(34) Alam, R.; Dixit, V.; Kang, H.; Li, Z. B.; Chen, X.; Trejo, J.; Fisher, M.; Juliano, R. L. Intracellular Delivery of an Anionic Antisense Oligonucleotide via Receptor-Mediated Endocytosis. *Nucleic Acids Res.* **2008**, *36*, 2764–2776.

(35) Sholder, G.; Creech, A.; Loechler, E. L. How Y-Family DNA Polymerase IV Is More Accurate than Dpo4 at DCTP Insertion Opposite an N2-dG Adduct of Benzo[*a*]Pyrene. *DNA Repair (Amst).* **2015**, *35*, 144–153.

(36) Ghodke, P. P.; Gore, K. R.; Harikrishna, S.; Samanta, B.; Kottur, J.; Nair, D. T.; Pradeepkumar, P. I. The N2-Furfuryl-Deoxyguanosine Adduct Does Not Alter the Structure of B-DNA. *J. Org. Chem.* **2016**, *81*, 502–511.

(37) Shao, J.; Tanner, S. W.; Thompson, N.; Cheatham, T. E. Clustering Molecular Dynamics Trajectories: 1. Characterizing the Performance of Different Clustering Algorithms. *J. Chem. Theory Comput.* **2007**, *3*, 2312–2334.

(38) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084–3095.

(39) Case, D. A.; Babin, V.; Berryman, J. T.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham, III, T.; E.; Darden, T. A.; Duke, R. E.; Gohlke, H.; Goetz, A.W.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolossváry, I.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Luchko, T.; Luo, R. . B.; Wu, X; Kollaman, P. A. *"AMBER 14", University of California, San Francisco*; 2014.

(40) Brandl, M.; Weiss, M. S.; Jabs, A.; Sühnel, J.; Hilgenfeld, R. C-H···π-Interactions in Proteins. *J. Mol. Biol.* **2001**, *307*, 357–377.

(41) Sharma, A.; Kottur, J.; Narayanan, N.; Nair, D. T. A Strategically Located Serine Residue Is Critical for the Mutator Activity of DNA Polymerase IV from Escherichia Coli. *Nucleic Acids Res.* **2013**, *41*, 5104–5114.

(42) Vanquelef, E.; Simon, S.; Marquant, G.; Garcia, E.; Klimerak, G.; Delepine, J. C.; Cieplak, P.; Dupradeau, F. Y. R.E.D. Server: A Web Service for Deriving RESP and ESP Charges and Building Force Field Libraries for New Molecules and Molecular Fragments. *Nucleic Acids Res.* **2011**, *39*, W511-7.

(43) Dupradeau, F.-Y.; Pigache, A.; Zaffran, T.; Savineau, C.; Lelong, R.; Grivel, N.; Lelong, D.; Rosanski, W.; Cieplak, P. The R.E.D. Tools: Advances in RESP and ESP Charge Derivation and Force Field Library Building. *Phys. Chem. Chem. Phys.* **2010**, *12*, 7821.

(44) Bayly, C.; Cieplak, P.; Cornell, W.; Kollman, P. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic. *J. Phys. Chem.* **1993**, *97*, 10269–10280.

(45) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. Gaussian 09, Revision D.01. *Gaussian 09, Revision B.01, Gaussian, Inc., Wallingford CT.*

Wallingford CT 2013.

(46) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins Struct. Funct. Genet.* **2006**, *65*, 712–725.

(47) Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers. *Biophys. J.* **2007**, *92*, 3817–3829.

(48) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878–3888.

(49) Götz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput.* **2012**, *8*, 1542–1555.

(50) Le Grand, S.; Götz, A. W.; Walker, R. C. SPFP: Speed without Compromise—A Mixed Precision Model for GPU Accelerated Molecular Dynamics Simulations. *Comput. Phys. Commun.* **2013**, *184*, 374–380.

(51) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera - A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612.

(52) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; et al. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* **2000**, *33*, 889–897.

(53) Miller, B. R.; McGee, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. MMPBSA.Py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321.