JOC The Journal of Organic Chemistry

pubs.acs.org/joc Article

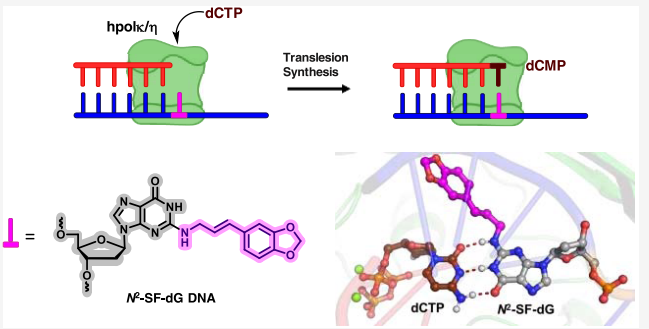
Synthesis of N^2-*trans*-isosafrole-dG-adduct Bearing DNAs and the Bypass Studies with Human TLS Polymerases κ and η

Siddharam Shivappa Bagale, Priyanka U. Deshmukh, Shailesh B. Lad, Akhil Sudarsan, Sruthi Sudhakar, Soumyadeep Mandal, Kiran Kondabagil,* and P. I. Pradeepkumar*

Cite This: J. Org. Chem. 2024, 89, 7680−7691 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Safrole is a natural product present in many plants and plant products, including spices and essential oils. During cellular metabolism, it converts to a highly reactive trans-isosafrole (SF) intermediate that reacts with genomic DNA and forms N^2-SF-dG and N^6-SF-dA DNA adducts, which are detected in the oral tissue of cancer patients with betel quid chewing history. To study the SF-induced carcinogenesis and to probe the role of low fidelity translesion synthesis (TLS) polymerases in bypassing SF adducts, herein, we report the synthesis of N^2-SF-dG modified DNAs using phosphoramidite chemistry. The N^2-SF-dG modification in the duplex DNA does not affect the thermal stability and retains the B-form of helical conformation, indicating that this adduct may escape the radar of common DNA repair mechanisms. Primer extension studies showed that the N^2-SF-dG adduct is bypassed by human TLS polymerases hpolκ and hpolη, which perform error-free replication across this adduct. Furthermore, molecular modeling and dynamics studies revealed that the adduct reorients to pair with the incoming nucleotide, thus allowing the effective bypass. Overall, the results indicate that hpolκ and hpolη do not distinguish the N^2-SF-dG adduct, suggesting that they may not be involved in the safrole-induced carcinogenicity.



INTRODUCTION

Various endogenous and exogenous genotoxic agents constantly attack genomic DNA. The action of genotoxic agents induces the chemical transformation of nucleobases or damages the phosphodiester backbone and creates lesions called DNA adducts or damages.[1] The structurally modified DNA nucleobases are one of the major causes of mutagenesis, which compromises the accuracy of DNA replication.[2] They are formed by different reactive metabolites of chemical carcinogens produced during cellular metabolism.[3−5] Nucleobase of 2′-deoxyguanosine (dG) is more vulnerable to modification by genotoxic agents than the other nucleobases. Along with the O^6 and N^7 position of dG, many genotoxic agents modify the exocyclic amine group N^2 position to form N^2-dG-adducts.[4,6] Cells have evolved various repair mechanisms to repair DNA damages, and despite these, certain DNA damages may escape the repair machinery.[7] The presence of these DNA damages arrests the progression of replicative DNA polymerases (dpols) δ, α, and ε due to the steric bulkiness and stalls the replication fork.[8] The stalled replication fork can lead to cellular apoptosis, which causes aging or age-related diseases.[9,10] Cells also possess a specialized group of DNA polymerases from the Y-family called translesion polymerases η, κ, ι, and Rev1, which bypass the damages and rescue the stalled replication forks.[11] Other than the Y-family polymerases, the replicative polymerase hpolζ, can also bypass the DNA damages.[12] This process is called translesion synthesis (TLS).[13] The Y-family polymerases are characterized as low fidelity and low processivity enzymes due to their, wider active site, and absence of an exonuclease proofreading domain.[14] Thus, the DNA damage tolerance can be error-prone (mutagenic) or error-free.[15] For example, the bypass of cyclobutene pyrimidine TT dimer DNA adduct formed by UV rays by hpolη is error-free, while the bypass of DNA adduct formed by heterocyclic aromatic amine from cooked meat, N^2-IQ-dG, is error-prone.[16,17] Polymerase hpolκ accurately bypasses the N^2-BP-dG DNA adduct formed by polycyclic aromatic hydrocarbon from cigarette smoke, while it performs an error-prone TLS across acetylaminofluorene adduct.[18]

Alkenylbenzenes are secondary plant metabolites, primarily present in many herbs and spices such as tarragon, pimento, anis, clove, cinnamon, and basil. Safrole is one of the

Received: February 9, 2024
Revised: April 17, 2024
Accepted: April 22, 2024
Published: May 13, 2024

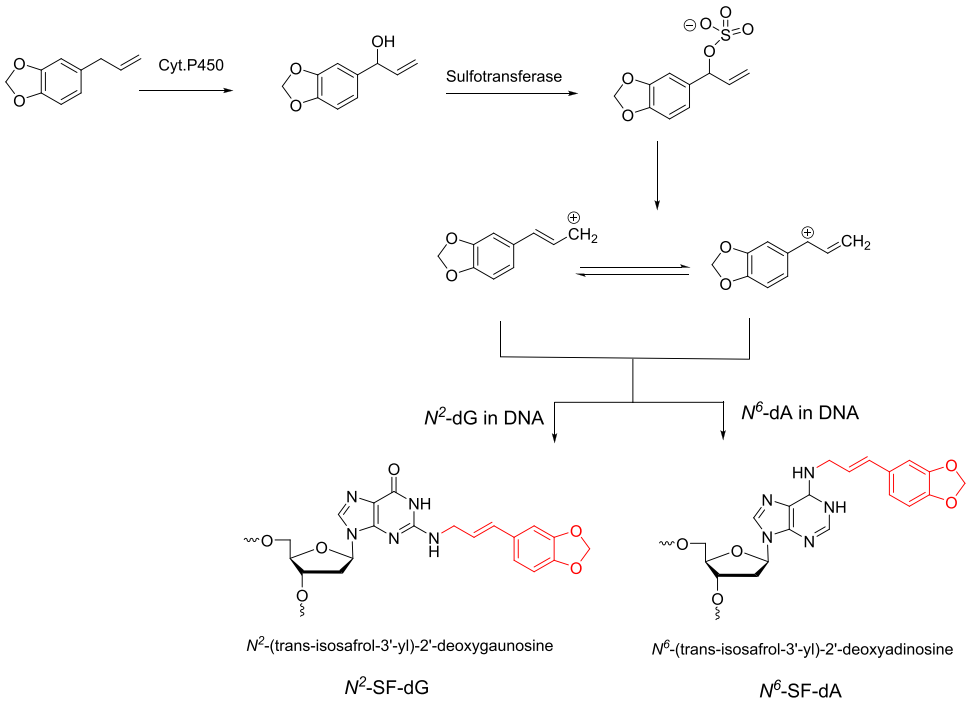
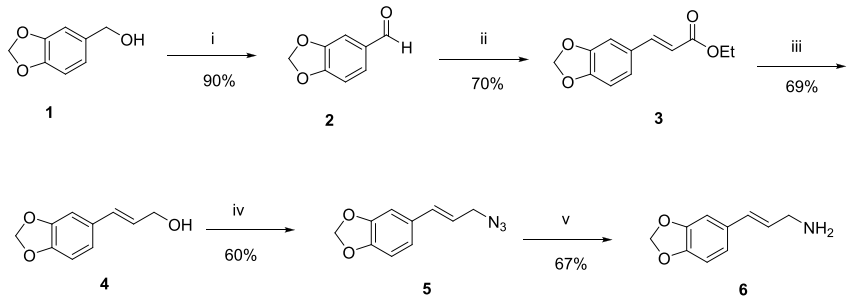


Figure 1. Proposed metabolic pathways for forming the N^2-SF-dG and N^6-SF-dA adducts.

Scheme 1. Synthesis of 3,4-Methylenedioxy-*trans*-cinnamyl Amine[a]



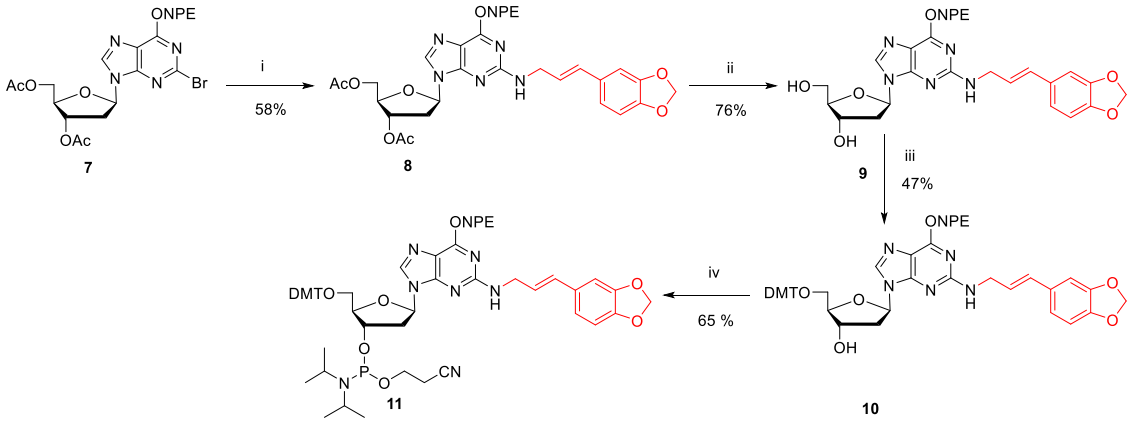
[a]Reagents and conditions: (i) DDQ, dioxane, RT, 3 h; (ii) ethyl bromoacetate, PPh_3, aq. $NaHCO_3$, 110 °C 24 h; (iii) DIBAL-H, Toluene, RT, 12 h; (iv) NaN_3, PPh_3, CCl_4-DMF (1:4), 90 °C, 16 h; (v) PPh_3, H_2O, THF, RT, 18 h.

derivatives of alkylbenzenes present in sassafras oil (85%) which is isolated from root bark *Passaferas albidum*.[19] It is also found in nutmeg (10%) and black pepper (0.01−0.1%).[20] The essential oil of piper betel, a tropical plant species of the *Piperaceae* family contains 28% safrole.[21] Reports suggest that chewing betel quid containing piper betel leaves generates safrole (420 μM) in the saliva,[21−23] which increases the risk for oral squamous cell carcinoma.[24] The N^2-dG (major) and N^6-dA (minor) DNA adducts of safrole were detected in the oral tissue of oral cancer patients, who had betel quid chewing history[21] and were also observed in the HepG2 cell line after exposure to safrole.[21,25] The dietary administration of safrole was found to cause liver cancer (hepatocellular carcinoma) and benign malignant liver tumors and also the exposure to safrole-2′,3′-oxide caused skin cancer in mice.[25,26] The International Agency for Research in Cancer (IARC) has classified safrole as a class 2B carcinogen and is identified as " possibly carcinogenic to humans".[27]

The safrole gets metabolized to an electrophilic intermediate and forms the N^2-*trans*-isosafrole-dG (N^2-SF-dG) and N^6-*trans*-isosafrole-dA (N^6-SF-dA) DNA adducts (Figure 1). Bioactivation of safrole involves a two-step enzymatic reaction where cytochrome P450 oxidizes the benzylic position to form 1′-hydroxy safrole, which is then converted to sulfonic acid ester by sulfotransferase enzyme. The unstable sulfonic acid ester forms a reactive intermediate and forms covalent adducts with DNA (Figure 1).[28−30] Therefore, it is imperative to probe the role of human TLS polymerase in triggering the carcinogenicity induced by safrole DNA adducts.[31] For this, access to N^2-SF-dG and N^6-SF-dA adduct-bearing DNAs is necessary. Herein, we report the synthesis of N^2-SF-dG-modified phosphoramidite and its site-selective incorporation into DNAs by solid-phase synthesis. Primer extension assays with human TLS polymerase hpolκ and hpolη were carried out to study the bypass across the N^2-SF-dG adduct. Thermal melting and CD spectroscopic studies were performed to understand the effect of N^2-SF-dG on the stability and conformation of duplex DNA. Molecular modeling and dynamics studies were carried out to gain insights into the bypass mechanism by hpolκ.

https://doi.org/10.1021/acs.joc.4c00368
J. Org. Chem. 2024, 89, 7680−7691

Scheme 2. Synthesis of N^2-SF-dG Phosphoramidite[a]



[a]Reagents and condition: (i) 3,4-methylenedioxy-*trans*-cinnamyl amine **6**, rac-BINAP, Cs_2CO_3, $Pd(OAc)_2$, toluene 90 °C, 12 h; (ii) 33% $MeNH_2$ in H_2O, RT, 2 h; (iii) DMT-Cl, DMAP, pyridine RT, 28 h; (vi) CEP-Cl, DIPEA, dry DCM, RT, 2 h.

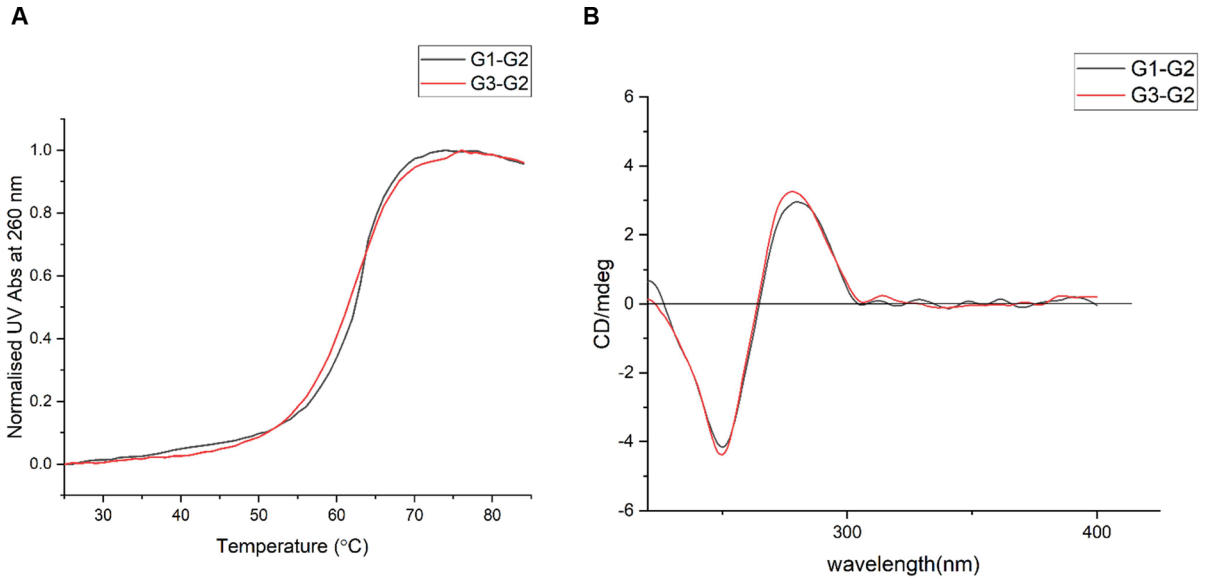


Figure 2. (A) UV thermal melting curve of the unmodified and the N^2-SF-dG modified duplex DNA. Melting curves were obtained using a 5 μM concentration of duplex DNA; (B) CD spectra of the unmodified and the N^2-SF-dG modified duplex DNA. The spectra were recorded using 10 μM concentration of ds DNA. Both UV and CD experiments were performed in phosphate buffer (20 mM phosphate buffer pH 7.4, 100 mM NaCl, 0.1 mM EDTA).

RESULTS AND DISCUSSION

Synthesis of the SF-dG-Modified Phosphoramidite and DNAs. To synthesize N^2-SF-dG-modified DNAs, we first synthesized the N^2-SF-dG-modified phosphoramidite. The synthetic strategy involved Buchwald-Hartwig coupling between amine **6** and bromo-nucleoside **7**. (Schemes 1 and 2) The amine **6** was prepared using a modified strategy starting from piperonyl alcohol **1**.[32] Compound **1** was oxidized to aldehyde using 2,3-dicholro-5,6-dicyno-1,4-benzoquinone (DDQ) to obtain compound **2** with 90% yield.[33] The aldehyde **2** was subjected to Wittig reaction using ethyl bromoacetate in saturated $NaHCO_3$ to produce α, β-unsaturated ester **3** in 70% yield.[34] Then, the ester derivative **3** was reduced to alcohol using DIBAL-H to access alcohol **4** in 69% yield.[35] The alcohol **4** was converted to azide using sodium azide to obtain azide **5** in 60% yield, and the azide **5** was reduced to amine using Staudinger reaction to furnish **6** in 67% yield.[5,36]

The amine **6** (Scheme 1) was subjected to the Buchwald–Hartwig coupling reaction using bromo nucleoside **7** to produce nucleoside **8** in 58% yield[37] (Scheme 2). The acetyl groups of **7** were deprotected by $MeNH_2$ in dioxane produce **9** in 76% yield. Then, the 5′-OH group of **9** was protected by using DMT-Cl to generate **10** in 47% yield. Nucleoside **10** was phosphorylated by using 2-cyanoethyl-*N*,*N*-diisopropylchlor-ophosphoramidite (CEP-Cl) to obtain phosphoramidite **11** in 65% yield.

Using solid-phase DNA synthesis, the N^2-SF-dG-modified phosphoramidite was successfully incorporated into the desired DNA sequences (Table S1 of the Supporting Information).[38] All the modified DNAs were characterized by using MALDI-TOF spectrometry. The 14-mer modified DNA, G3 was used for thermal melting studies and to record CD spectra. The 27-mer modified DNA sequence G4 was used for primer extension studies with human TLS polymerase hpolκ and hpolη.

https://doi.org/10.1021/acs.joc.4c00368
J. Org. Chem. 2024, 89, 7680−7691

UV Thermal Melting and CD Spectroscopy Studies. To examine the effect of N^2-SF-dG modification on DNA duplex stability, UV-melting studies were carried out (Figure 2A). The similar melting temperature (T_m = 62.4 ± 0.6 °C) for both unmodified (G1−G2) and modified (G3−G2) DNA duplexes shows that N^2-SF-dG modification does not appear to affect the duplex stability. CD spectra of the modified duplex showed B-helical confirmation (Figure 2B). Distortion of B-form of double helical conformation due to DNA damage is sensed by repair enzymes, and therefore, the N^2-SF-dG damage may escape the radar of the common repair pathways.[39−41]

Primer Extension Assays. To investigate the fidelity and extension ability of the human TLS polymerases hpolκ and hpolη, we carried out single nucleotide incorporation and full-length extension assays (Figures 3 and 4). Single nucleotide incorporation was carried out using a primer/template duplex DNA substrate (Figures 3A and 4A). The duplex was prepared by annealing a 15-mer FAM labeled primer with a 27-mer unmodified or a 27-mer modified DNA template. The modified DNA template contains the N^2-SF-dG modification at the 16th position. The duplex DNA substrate was incubated with TLS polymerase hpolκ or hpolη and varying concentrations of all four dNTPs at 37 °C for 10 min in individual experiments. Product formation was analyzed on 20% denaturing PAGE (7 M urea) and bands were visualized using fluorescence imaging. Results showed that hpolκ selectively incorporates dCTP across the unmodified (dG) and N^2-SF-dG modified template (Figure 3B,C). Similarly, hpolη also selectively incorporates dCTP across the N^2-SF-dG-modification (Figure 4C). Even at higher concentrations of dCTP, the efficiency to bypass the adduct by hpolη remains lower compared to hpolκ. Thus, the single nucleotide extension assays showed that both human TLS polymerases perform error-free bypass across the N^2-SF-dG adduct. This is in agreement with the literature reports on the bypass of various N^2-dG adducts by hpolκ and hpolη.[42,43]

A full-length extension reaction was performed using the same primer-template duplex used in the single nucleotide incorporation assays (Figure 5). The duplex DNA substrate was incubated in TLS polymerase hpolκ or hpolη and a mixture of all four dNTPs at 37 °C for 2 h. The result showed that hpolκ efficiently bypassed the N^2-SF-dG adduct and extended the primer to form the full-length product (Figure 5B,C). However, though TLS polymerase hpolη extended the primer, it has less efficiency in forming the full-length product (Figure 5D, E), which aligns with the literature reports.[42]

To check the polymerase efficiency in bypassing the N^2-SF-dG adduct during the running start extension reaction, we repeated the full-length extension assay with a 13-mer FAM-labeled primer/template duplex DNA substrate (Figure 6A). The results showed that the primer extended in the unmodified DNA substrate to form the full-length product without any stall (Figure 6B). While in the DNA substrate which contains the N^2-SF-dG modification, hpolκ slightly stalled the extension after the addition of two nucleotides to form the product at the 15th position, due to the presence of modification at the 16th position; furthermore, it extended the primer to form the full-length product (Figure 6C). The primer extended by hpolη in

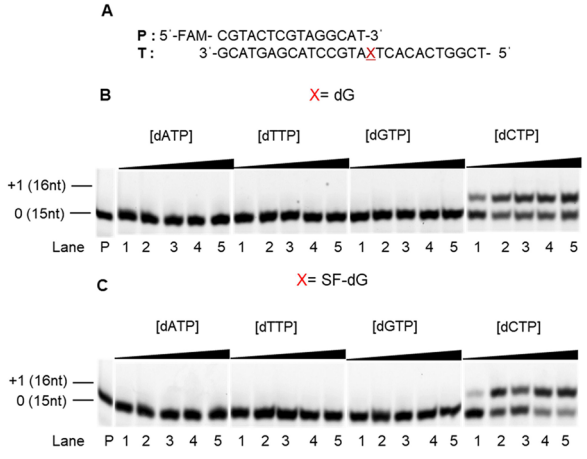


Figure 3. Denaturing PAGE (20%, 7 M urea) of single nucleotide incorporation of hpolκ (5 nM) in the presence of individual dNTPs. (A) Duplex DNA substrate containing 15-mer primer (P) and 27-mer template (T), where (B) X = dG; and (C) X = N^2-SF-dG. All reactions were carried at 37 °C for 0.1, 1, 10, 20, or 40 μM concentrations of individual dNTPs (dATP, dTTP, dGTP, dCTP).

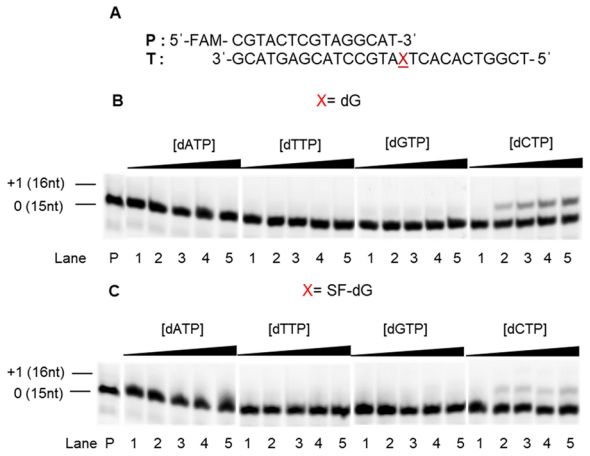


Figure 4. Denaturing PAGE (20%, 7 M urea) of single nucleotide incorporation of hpolη (5 nM) in the presence of individual dNTPs. (A) Duplex DNA substrate containing 15-mer primer (P) and 27-mer template (T), where (B) X = dG; and (C) X = N^2-SF-dG. All reactions were carried at 37 °C for 0.1, 1, 10, 20, or 40 μM concentrations of individual dNTPs (dATP, dTTP, dGTP, dCTP).

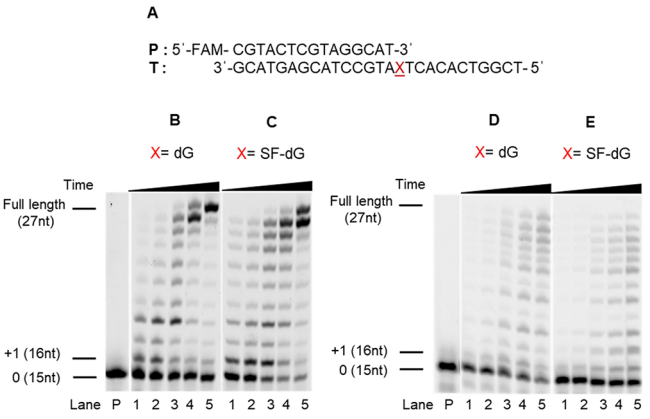


Figure 5. Denaturing PAGE (20%, 7 M urea) of full-length extension assay with hpolκ and hpolη (5 nM) in the presence of a mixture of dNTPs. (A) Duplex DNA substrate containing 15-mer primer and 27-mer templates; (B) X = dG; (C) X = N^2-SF-dG reaction carried out in the presence of hpolκ; (D) X = dG; and (E) X = N^2-SF-dG reaction carried out in the presence of hpolη. All reactions were incubated at 37 °C, and time points were taken at 5, 10, 30, 60, and 120 min.

https://doi.org/10.1021/acs.joc.4c00368
J. Org. Chem. 2024, 89, 7680−7691

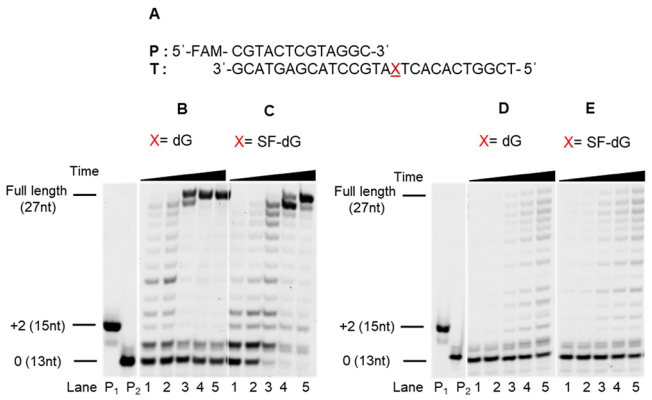


Figure 6. Denaturing PAGE (20%, 7 M urea) of running start full-length extension assay with hpolκ and hpolη (5 nM) in the presence of a mixture of dNTPs. (A) Duplex DNA substrate containing 13-mer primer and 27-mer template; (B) X = dG; (C) X = N^2-SF-dG reaction carried out in the presence of hpolκ.; (D) X = dG; and (E) X = N^2-SF-dG reaction carried out in the presence of hpolη. All reactions were incubated at 37 °C, and time points were taken at 5, 10, 30, 60, and 120 min.

the N^2-SF-dG modified the DNA substrate to produce the full-length product with much lower efficiency than hpolκ (Figure 6E).

Molecular Modeling and Dynamics Studies. To gain atomistic details of the mechanism of TLS action on the N^2-SF-dG adducted DNA, 500 ns molecular dynamic (MD) simulations were carried out on the damaged DNA-hpolκ complex in its preinsertion, insertion, and postinsertion stages (Figure 7A). The preparation of the complexes was carried out by replacing the lucidin adduct by N^2-SF-dG (Figure S1 of the Supporting Information), in the X-ray crystal structures of N^2-dG-lucidin adducted DNA in complex with hpolκ (PDB ID: 5W2A as a template for the preinsertion and insertion stages and PDB ID:5W2C as a template for the postinsertion stage) while keeping the DNA sequence same (Figure 7B). The structural stability of the system was analyzed using root-mean-square deviation (RMSD) calculation of the protein and DNA (Figure S2A,B of the Supporting Information). The complexes were stable in all three stages of replication with an average RMSD value of less than 3 Å for the protein backbone and less than 2.5 Å for the DNA backbone. The RMSD values of the DNA backbone for the insertion stage were slightly higher compared to the preinsertion stage, indicating that the advent of the incoming nucleotide opposite the adduct in the insertion stage may be associated with some structural rearrangements within the DNA structure. The root means square fluctuation (RMSF) calculations revealed that the amino acid residues at the beginning of the N-Clasp domain show the most significant fluctuations, followed by some residues in the palm and thumb

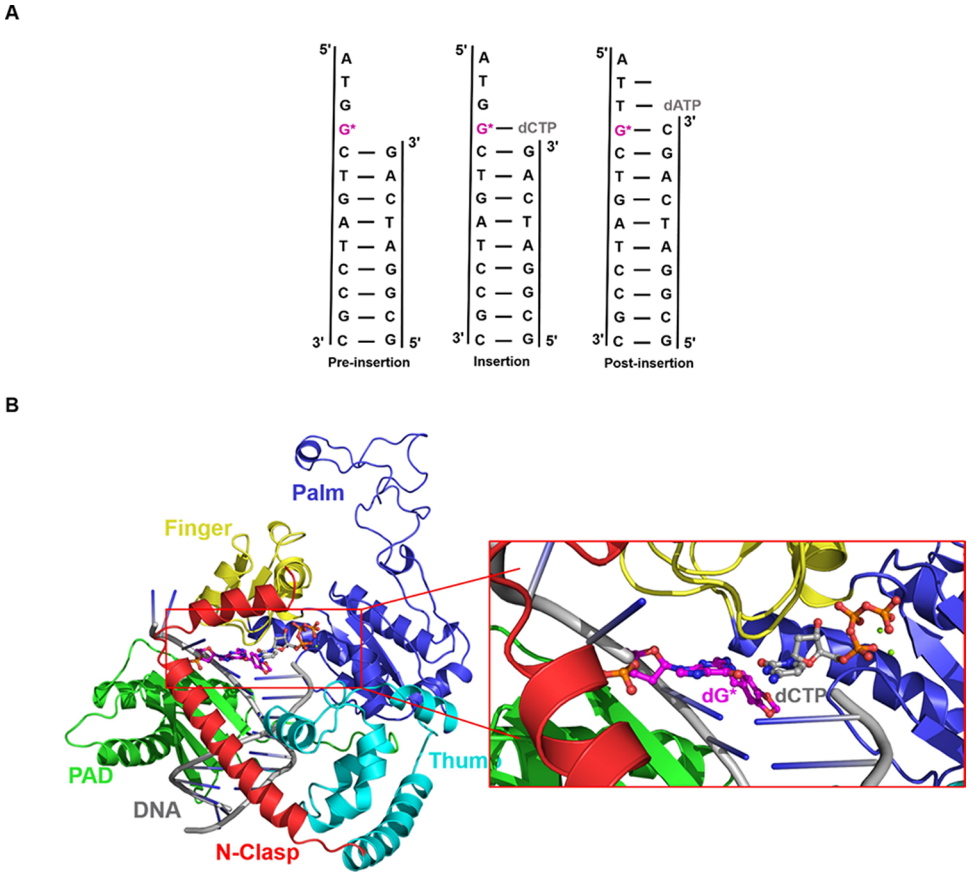


Figure 7. (A) Illustration showing the DNA duplex sequence used for MD simulation in the preinsertion, insertion, and postinsertion stages; (B) representative image of the insertion stage of N^2-SF-dG hpolκ-DNA complex used in the MD simulations. The palm, thumb, N-clasp, finger, and polymerase-associated domain (PAD) of hpolκ are shown in purple, cyan, red, yellow, and green colors, respectively. The red-colored rectangle represents the active site. The zoomed-in view shows the N^2-SF-dG adduct (represented in magenta and labeled as dG*) and the dCTP (represented in gray) in the enzyme's active site.

https://doi.org/10.1021/acs.joc.4c00368
J. Org. Chem. 2024, 89, 7680−7691

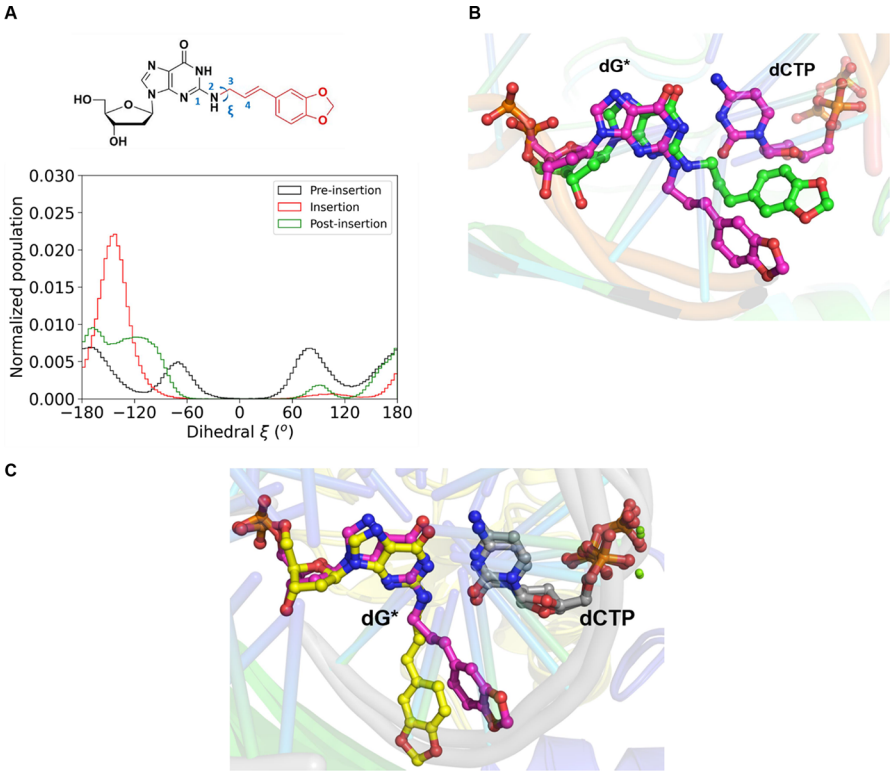


Figure 8. (A) Structure of adduct representing the dihedral angle ξ between the atoms C2, N2, C12, and C14 labeled 1−4, respectively, and the histogram representing the dihedral values in all the three stages of replication; (B) superimposed images of the cluster 1 of preinsertion (green) and insertion (magenta); (C) superimposed images of the cluster 1 (magenta) and cluster 2 (yellow) of the insertion stage. In all the snapshots, oxygen atoms are shown in red, nitrogen atoms in blue, phosphate atoms in orange, and dCTP in gray. dG* represents the adduct.

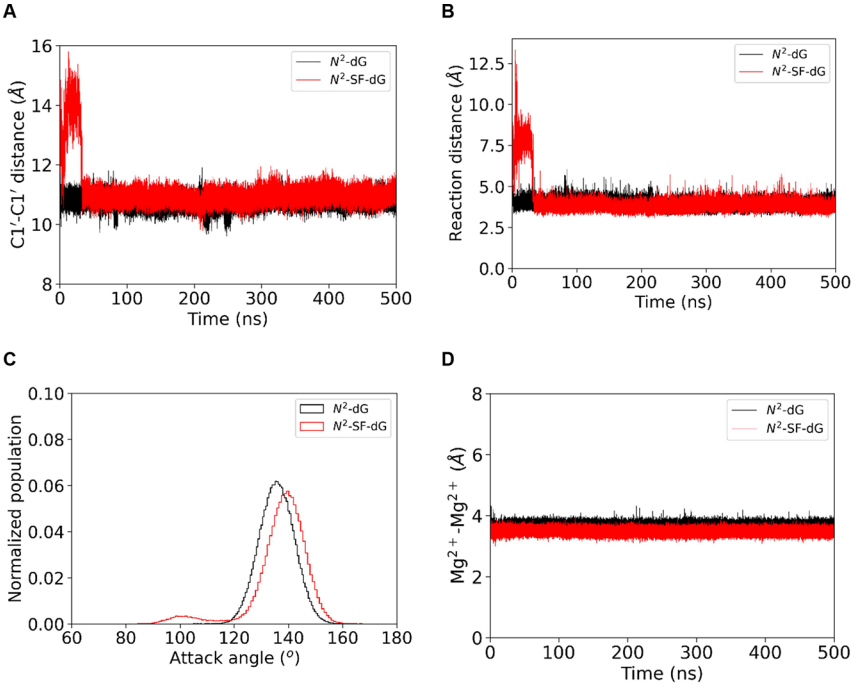


Figure 9. (A) Plot showing the distance between the C1′ atom of N^2-dG/N^2-SF-dG nucleotide and the C1′ atom of the dCTP; (B) plot depicting the reaction distance; (C) histogram showing the probability distribution of the attack angle of the incoming nucleotide; and (D) plot depicting the Mg^{2+}−Mg^{2+} distance with respect to time. All the calculations are carried out from the 500 ns MD trajectories.

domains (Figure S3A of the Supporting Information). The amino acid residues surrounding the adduct in the active site pocket were stable throughout the length of the simulation in all three stages. The free (unpaired) nucleotides present toward the 5′-end of the adduct showed maximum fluctuation in the case of DNA, followed by the ones present at the primer

https://doi.org/10.1021/acs.joc.4c00368
J. Org. Chem. 2024, 89, 7680−7691

strand terminals. The DNA adduct exhibited minor fluctuations through all the stages of replication (Figure S3B of the Supporting Information).

Preinsertion Stage. The preinsertion stage happens before the insertion of dNTP in the active site (Figure 7A). The absence of dNTP allows the adduct to have multiple orientations (Figure S4A of the Supporting Information). To understand these conformations, the average value of the dihedral angle ξ (Figure 8A) was calculated for the major cluster obtained after hierarchical agglomerative clustering of the frames from the trajectory. The value of ξ for the damaged nucleotide exhibited peaks around ∼ −166°, −70°, and 80°, which suggests that the adduct in the preinsertion stage was mainly oriented in a way that obstructed the insertion of dNTP. This was further confirmed by the visualization of the superimposed major clusters of the preinsertion and insertion stage (Figure 8B).

However, during the insertion stage, the adduct reoriented itself to accommodate the dCTP in the active site (Figure 8C), which is discussed further later in the results. The presence of the adduct also did not alter the structural stability of the duplex (Figure S2B of the Supporting Information). The C1′−C1′ distances between opposite base pairs were also constant at around 10.8 Å.

Insertion Stage. The structural parameters associated with the dCTP orientation for successful replication of DNA past the adducts were measured as reported previously.[5,44] The parameters are C1′−C1′ distance (10.8 Å), reaction distance [between Pα of the dNTP and O3′ of the terminal base of primer (3.5 Å)], the attack angle [between O3′ (primer base), Pα (dNTP), and O$\alpha\beta$ (dNTP) (150°−180°)], the distance between the catalytic Mg^{2+} ions (3.5−4.4 Å), and Watson−Crick hydrogen bonds between the template base and the dNTP with sufficient strength (−24 to −27 kcal/mol for the G−C base pair and −10 to −14 kcal/mol for the A−T base pair). Together, the optimal values of these parameters determine the replication efficiency.

The H-bond occupancy analysis revealed that N^2-SF-dG and dCTP form strong interactions with >90% occupancy (Table S2 of the Supporting Information) and a cumulative interaction strength of −23 ± 1.5 kcal/mol, consistent with the reported values. The Watson−Crick pairing between the rest of the duplex base pairs remained intact throughout the simulation, except for the terminal 3′ base pair, which was susceptible to fraying. The C1′−C1′ distance between opposite base pairs was also consistent at around 10.8 Å, indicating that the DNA duplex does not undergo drastic structural changes upon incorporating the modified N^2-SF-dG adduct (Figure 9A). The average value of the reaction distance was ∼3.8 Å, and the attack angle was around 139.2° (Figure 9B, C). The average distance between the catalytic Mg^{2+} and the nucleotide-binding Mg^{2+} was 3.5 ± 0.1 Å (Figure 9D). Although there are slight deviations, the overall orientation of the dCTP is found to be satisfactory for successful replication.

The interaction of the DNA adduct at the active site was explored in detail. N^2-SF-dG engages in hydrogen bonding interaction with various residues present in the active site, such as Met-135, Ser-137, and Thr-421 (Figure 10A). These interactions stabilize the adduct and hold them in site for replication. The DNA adduct also involves other noncovalent (mainly van der Waals) interactions with the surrounding residues (Figure 10A). The MM/GBSA method was employed to quantify these further, and the average interaction energies

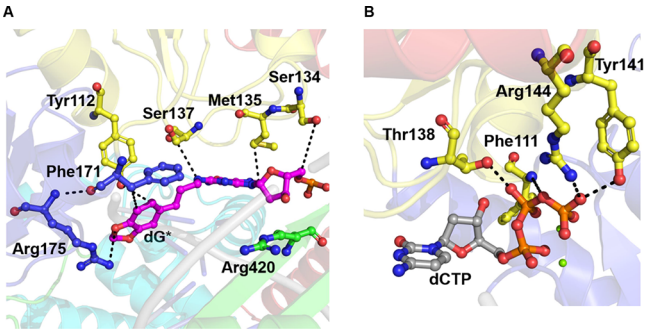


Figure 10. (A) Interaction of the modified N^2-SF-dG adduct (dG*) with amino acid residues surrounding the hpolκ active site. These amino acid residues belong in the finger, PAD, and palm domains; (B) interactions that hold the dCTP in the active site for efficient bypass of the adduct.

between the adduct and the active site residues surrounding it were calculated (Table S3 of the Supporting Information). Major noncovalent interactions of the adduct were observed with the residues from the finger (Tyr-112, Ser-137, Met-135), palm (Phe-171, Arg-175), and PAD (Arg-420) domains (Table S3 of the Supporting Information). Similarly, the incoming nucleotide participates in the hydrogen bonding interactions with the amino acid residues of the finger domain, mainly Phe-111, Thr-138, Arg-144, and Tyr-141 (Figure 10B).

After clustering the insertion stage trajectory, the major ensemble obtained had an occupancy of around 97%, and the adduct showed an average dihedral angle of −144.8° ± 12.3°. Visual inspection of these clusters revealed that this orientation does not disrupt the entry of the incoming nucleotide into the active site, unlike the conformation occupied by the adduct in the preinsertion stage (Figure 8B). This implies that the adduct changes its orientation during the insertion stage to allow the proper pairing with dCTP and engage in stabilizing interactions with the active site amino acid residues.

Postinsertion Stage. The postinsertion (extension) stage occurs after the insertion stage and is characterized by the primer moving to the base 5′ to the N^2-SF-dG for replication (Figure 7A). The incoming nucleotide (dATP) interacts with the 5′-dT with two hydrogen bonds of strengths ∼99 and 97%, respectively (Table S4 of the Supporting Information). The average interaction energy between them is −9.6 ± 1 kcal/mol, slightly less than the theoretical value. The pairing of the dATP with dT did not significantly impact the 3 H-bonds of the base pairing between the already replicated adduct with an interaction energy of −24.8 ± 1.7 kcal/mol and percentage H-bond occupancy >90% (Table S5 of the Supporting Information). The same applies to the rest of the DNA base pairs; the base pairing remains intact, and the C1′-C1′ distance between the opposite base pairs is around 10.8 Å. The value of the reaction distance between the dATP and dT was ∼5 Å, and the attack angle was ∼147° for the postinsertion stage, implying that the incoming nucleotide is favorably oriented in the active site.

A major cluster of ∼55% occupancy and a minor cluster of ∼39% occupancy were obtained after clustering the trajectory into 5 ensembles. Analysis of the dihedral angle ξ of the major cluster reveals that the adduct occupies a conformation in the active site with an average dihedral angle of ∼ −140° and −170° in the central and minor clusters, respectively. The representative snapshots of the major clusters of the post-

https://doi.org/10.1021/acs.joc.4c00368
J. Org. Chem. 2024, 89, 7680−7691

insertion stage were superimposed to visualize the same (Figure S4B). This suggests that the orientation of the adduct in the active site does not interfere with the replication past the adduct.

CONCLUSIONS

To probe the role of translesion DNA synthesis on the reported carcinogenicity of natural product safrole, we have achieved the synthesis of N^2-SF-dG modified DNAs using phosphoramidite chemistry. Primer extension studies indicated human TLS polymerase hpolκ and hpolη bypass the N^2-SF-dG adduct in an error-free manner. The hpolκ bypasses the adduct with high efficiency and processivity compared to hpolη. Thermal melting and CD spectroscopic studies show that the presence of the adduct does not affect the stability of duplex DNA and B helical conformation. Extensive molecular modeling and dynamics studies using the preinsertion, insertion, and postinsertion DNA-hpolκ complexes showed the structural features associated with the effective adduct bypass by hpolκ. Orientation of the adduct alters at the active site, thus allowing efficient nucleotide insertion. Overall, our findings suggest that human TLS polymerase hpolκ and hpolη are not involved in the carcinogenicity caused by safrole. However, other low-fidelity polymerases such as Polζ or Rev1 may be involved in error-prone bypass across this adduct, and we plan to pursue such studies in the future.

EXPERIMENTAL SECTION

General. All the required chemicals were obtained from commercial sources. Acetonitrile (ACN), dichloromethane (DCM), toluene, pyridine, *N*, *N*-Diisopropylethylamine (DIPEA), triethylamine (Et_3N), and dioxane were dried using calcium hydride and stored in oven-dried molecular sieves (4 Å). The progress of the reaction was monitored by thin-layer chromatography (TLC). TLC was performed on silica gel plates precoated with fluorescent indicators and visualized under UV light (260 nm) or by spraying a solution of *p*-anisaldehyde, or $KMnO_4$, or 5% conc. H_2SO_4 in ethanol (v/v). Silica gel (100−200 mesh) was used in column chromatography to purify the compound. ^{1}H NMR (400 and 500 MHz), ^{13}C NMR (100 and 125 MHz), and ^{31}P NMR (162 MHz) were recorded on 400 and 500 MHz NMR instruments. Chemical shifts (δ) in parts per million were referenced to the residual signal of TMS (0 ppm) or residual proton signal from deuterated solvent: $CDCl_3$ (7.26 ppm) for the ^{1}H NMR spectra, and $CDCl_3$ (77.2 ppm) for the ^{13}C NMR spectra. Multiplicities of ^{1}H NMR spin couplings are reported as s for singlet, bs for broad singlet, d for doublet, dt for doublet of triplets, dd for doublet of doublets, ddd for doublet of doublet of doublets, or m for multiplets and overlapping spin systems. Values of coupling constants (J) are reported in Hz. High-resolution mass spectra (HRMS) were obtained in a positive ion electrospray ionization (ESI) mode using the Q-TOF analyzer. Mass spectra of the oligonucleotides were obtained by positive reflection mode in the MALDI-TOF spectrometer.

Preparation of 3,4-Methylenedioxy Benzaldehyde (2).[33] In a round-bottom flask, commercially available piperanol **1** (2 g, 13.14 mmol) was dissolved in 1,4-dioxane (60 mL), followed by the addition of DDQ (3.04 g, 13.40 mmol). The reaction mixture was stirred at room temperature for 3 h. After completion of the reaction, the solvent was evaporated under reduced pressure, and the crude was purified by silica gel column chromatography (10% ethyl acetate in pet ether) to obtain compound **2** as a white solid with a sweet smell (1.78 g, 90%) R_f = 0.41 (10% ethyl acetate in pet ether); mp: 32−35 °C; ^{1}H NMR (400 MHz, $CDCl_3$) δ 9.80 (s, 1H), 7.41 (dd, J = 7.9, 1.6 Hz, 1H), 7.33 (d, J = 1.6 Hz, 1H), 6.93 (d, J = 7.9 Hz, 1H), 6.07 (s, 2H); ^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$) δ 190.3, 153.1, 148.7, 131.8, 128.6, 108.3, 106.9, 102.1; HRMS (ESI) m/z: $[M + H]^+$ Calcd for $C_8H_6O_3$ 151.0391, found 151.0390.

Preparation of Ethyl 3,4-Methylenedioxy-*trans*-cinnamate (3).[34] In a round-bottom flask, the mixture of aldehyde **2** (1.2 g, 7.99 mmol) and PPh_3 (3.1 g, 11.98 mmol) were dissolved in saturated aq. $NaHCO_3$ (11 mL) followed by the addition of ethyl-bromoacetate (2.6 mL, 23.97 mmol). The reaction was stirred at 110 °C in a heating block for 24 h. The reaction mixture was quenched with aq. $NaHCO_3$ (30 mL). The resulting mixture was extracted with ethyl acetate (3 × 30 mL), and the combined organic extract was dried over Na_2SO_4. The solvent was evaporated under reduced pressure, and the crude was purified by silica gel column chromatography (5% ethyl acetate in pet ether), to obtain compound **3** as a white solid with a sweet smell (1.28 g, 70%); R_f = 0.51 (10% ethyl acetate in pet ether); mp: 53−55 °C; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.57 (d, J = 15.9 Hz, 1H), 7.00 (d, J = 1.6 Hz, 1H), 6.97 (dd, J = 7.9, 1.6 Hz, 1H), 6.78 (d, J = 7.9 Hz, 1H), 6.24 (d, J = 16.0 Hz, 1H) 5.9 (s, 2H), 4.23 (q, J = 7.1 Hz, 2H), 1.30 (t, J = 7.1 Hz, 3H); ^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$) δ 167.1, 149.5, 148.3, 144.2, 128.8, 124.3, 116.1, 108.5, 106.4, 101.5, 60.3, 14.3; HRMS (ESI) m/z: $[M + H]^+$ Calcd for $C_{12}H_{12}O_4$ 221.0822, found 221.0821.

Preparation of 3,4-Methylenedioxy-*trans*-cinnamyl Alcohol (4).[35] In an oven-dried round-bottom flask, trans-cinnamate **3** (1 g, 4.54 mmol) was dissolved in dry toluene (24 mL) followed by dropwise addition of DIBAL-H (2.7 mL, 15.90 mmol, of 1.0 M in toluene) at 0 °C. After being stirred for 5 min, the reaction mixture was removed from the ice bath and further stirred at room temperature for 12 h. The reaction mixture was quenched with H_2O (40 mL). The resulting mixture was extracted with ethyl acetate (3 × 30 mL), and the combined organic extract was dried over Na_2SO_4. The solvent evaporated under reduced pressure, and crude was purified by silica gel column chromatography (20% ethyl acetate in pet ether) to obtain compound **4** as an off-yellow solid (560 mg, 69%); R_f = 0.3 (20% ethyl acetate in pet ether); mp: 77−79 °C; ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.94 (d, J = 1.5 Hz, 1H), 6.83 (dd, J = 8.0, 1.6 Hz, 1H), 6.77 (d, J = 8.0 Hz, 1H), 6.53 (d, J = 15.9 Hz, 1H), 6.21 (dt, J = 15.9, 6.9 Hz, 1H), 5.96 (s, 2H), 4.3 (d, J = 5.3 Hz, 2H), 1.74 (bs, s, 1H); ^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$) δ 148.0, 147.3, 131.1, 130.9, 126.7, 121.1, 108.3, 105.7, 101.0, 63.7; HRMS (ESI) m/z: $[M + H]^+$ Calcd for $C_{10}H_{10}O_3$ 179.0692; found 179.0691.

Preparation of 3,4-Methylenedioxy-*trans*-cinnamyl Azide (5).[36] In an oven-dried round-bottom flask, a mixture of trans-cinnamyl alcohol **4** (320 g, 1.79 mmol) and PPh_3 (987 mg, 3.75 mmol) were dissolved in CCl_4-DMF (67 mL, 1:4 v/v, 5 mL/mmol) followed by portion-wise the addition of sodium azide (139 mg, 2.14 mmol) at RT. The reaction mixture was stirred at 90 °C in a heating block for 16 h. The reaction mixture was allowed to cool at room temperature and then quenched with H_2O (40 mL). The resulting mixture was extracted with ethyl acetate (3 × 30 mL), and the combined organic extract was dried over Na_2SO_4. The solvent evaporated under reduced pressure, and the crude product was purified by silica gel column chromatography (5% ethyl acetate in pet ether), to obtain compound **5** as a yellow liquid (220 mg, 60%) R_f = 0.6 (5% ethyl acetate in pet ether); ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.96 (d, J = 1.7 Hz, 1H), 6.86 (dd, J = 8.0, 1.7 Hz, 1H), 6.79 (d, J = 8.0 Hz, 1H), 6.57 (d, J = 15.6, Hz, 1H), 6.09 (dt, J = 15.7, 6.7 Hz, 1H), 5.98 (s, 2H), 3.93 (d, J = 6.8 Hz, 2H); ^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$) δ 148.1, 147.7, 134.2, 130.4, 121.5, 120.5, 108.3, 105.8, 101.2, 53.0; HRMS (ESI) m/z: $[M + H]^+$ Calcd for $C_{10}H_9N_3O_2$ 204.0768; found 204.0767.

Preparation of 3,4-Methylenedioxy-*trans*-cinnamyl Amine (6).[36] Trans-cinnamyl azide **5** (220 mg, 1.02 mmol) was dissolved in THF (2.7 mL), followed by the addition of PPh_3 (425 mg, 1.62 mmol) at 0 °C in an ice bath. After being stirred for 5 min, the reaction mixture was removed from the ice bath, and water was added (0.6 mL). The reaction was stirred at room temperature for 18 h. After the consumption of the starting material, the solvent was evaporated under reduced pressure and water was removed by azeotropic distillation with toluene. The brown-colored crude was purified by silica gel column chromatography (10% methanol in

https://doi.org/10.1021/acs.joc.4c00368
J. Org. Chem. 2024, 89, 7680−7691

DCM) to obtain compound **6** as a greenish viscous solid (120 mg, 67%); R_f = 0.3 (10% methanol in DCM); ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.90 (d, J = 1.6 Hz, 1H), 6.78 (dd, J = 8.1, 1.6 Hz, 1H), 6.73 (d, J = 7.9 Hz, 1H), 6.40 (d, J = 15.7 Hz, 1H), 6.13 (dt, J = 15.7, 5.9 Hz, 1H), 5.92 (s, 2H) 3.43 (dd, J = 5.9, 1.4 Hz, 2H), 1.86 (s, 2H); ^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$) δ 147.9, 146.9, 131.6, 129.4, 129.1, 120.7, 108.2, 105.5, 101.0, 44.2; HRMS (ESI) m/z: $[M + H]^+$ Calcd for $C_{10}H_{11}NO_2$ 178.0867; found $[M + H]^+$ = 178.0867.

Synthesis of N^2-Methyl-(3,4-methylenedioxy)styrenyl)-O^6-(2-(4-nitrophenyl)ethyl)-3′,5′-diacetyl-2′-deoxyguanosine (8). In an oven-dried round-bottom flask, a mixture of trans-cinnamyl amine **6** (69.1 mg, 0.39 mmol), rac-BINAP (66 mg, 0.10 mmol), $CsCO_3$ (160 mg, 0.49 mmol) and $Pd(OAc)_2$ (15.9 mg, 0.07 mmol) were dissolved in dry toluene (5 mL), and allowed to stir for 5 min. The bromo nucleoside **7** (200 mg, 0.35 mmol; nucleoside **7** was prepared using a protocol reported by us earlier)[37] was added under an inert atmosphere and stirred at 90 °C a heating block for 12 h. The reaction mixture was passed through a Celite pad and washed with ethyl acetate (12 mL). The filtrate was concentrated under reduced pressure and the crude nucleoside was purified by silica gel column chromatography (55% ethyl acetate in pet ether, 1% NEt_3), to obtain coupled nucleoside **8** as a yellow solid (136 mg, 58%); R_f = 0.3 (55% ethyl acetate in pet ether, 1% NEt_3); mp: 82−84 °C; ^{1}H NMR (500 MHz, $CDCl_3$) δ: 8.09 (d, J = 8.6 Hz, 2H), 7.72 (s, 1H), 7.42 (d, J = 8.7 Hz, 2H), 6.85 (s, 1H), 6.71 (dd, J = 17.7, 8.0 Hz, 2H), 6.46 (d, J = 15.8 Hz, 1H), 6.26 (t, J = 6.7 Hz, 1H), 6.11 (dt, J = 15.8, 5.8 Hz, 1H), 5.90 (s, 2H), 5.46−4.45 (m, 1H), 5.29 (t, J = 5.3 Hz, 1H), 4.70 (t, J = 6.5 Hz, 2H), 4.41 (q, J = 6.6 Hz, 1H), 4.30−4.27 (m, 2H), 4.17 (t, J = 5.6 Hz, 2H), 3.23 (t, J = 6.9 Hz, 2H), 3.07 (quintet, J = 2.9 Hz, 1H), 2.4 (ddd, J = 11.4, 3.4, 2.9 Hz, 1H), 2.10 (s, 3H), 2.04 (s, 3H); ^{13}C{^{1}H} NMR (125 MHz, $CDCl_3$) δ: 170.5, 170.3, 160.6, 158.7, 153.5, 148.0, 147.1, 146.7, 146.0, 137.6, 131.2, 130.8, 129.8, 124.8, 123.6, 120.8, 115.5, 108.2, 105.5, 101.0, 84.4, 82.2, 74.5, 65.9, 63.8, 43.9, 36.4, 35.1, 20.9, 20.7; HRMS (ESI) m/z: $[M + H]^+$ Calcd for $C_{32}H_{33}N_6O_{10}$ 661.2253; found 661.2252.

Synthesis of N^2-CH_2−((3,4-Methylenedioxy)styrenyl)-O^6-(2-(4-nitrophenyl)ethyl)-2′-deoxyguanosine (9). The coupled nucleoside **8** (300 mg, 0.45 mmol) was dissolved in 1,4-dioxane 6.0 mL, followed by the addition of 33% aq. $MeNH_2$ (5.1 mL, v/v), and stirred the reaction mixture at room temperature for 2 h. The solvent was evaporated under reduced pressure. The crude nucleoside was purified by silica gel column chromatography (10% methanol in DCM) to obtain nucleoside **9** as a yellow solid (200 mg, 76%); R_f = 0.3 (10% methanol in DCM); mp: 78−80 °C; ^{1}H NMR (400 MHz, $CDCl_3$) δ: 8.13 (d, J = 8.9 Hz, 2H), 7.63 (s, 1H), 7.44 (d, J = 8.6 Hz, 2H), 6.85 (s, 1H), 6.75−6.72 (m, 2H) 6.46 (d, J = 15.7 Hz, 1H), 6.22 (q, J = 6.1 Hz, 1H), 6.09 (dt, J = 15.8, 6.1 Hz, 1H), 5.95 (s, 1H), 5.29 (t, J = 6.3 Hz, 1H), 4.75 (t, J = 6.4 Hz, 3H), 4.17 (m, 3H), 3.95 (dd, J = 12.4, 1.7 Hz, 1H), 3.75 (d, J = 12.4 Hz, 1H), 3.27 (t, J = 6.7 Hz, 2H), 3.02 (m, 1H), 2.26 (ddd, J = 13.6, 8.5, 6.0 Hz, 1H); ^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$) δ: 160.9, 158.2, 152.7, 148.0, 147.2, 146.8, 145.8, 139.0, 131.2, 131.1, 129.8, 124.5, 123.7, 120.8, 116.3, 108.2, 105.5, 101.1, 89.0, 87.2, 73.2, 66.1, 63.4, 44.1, 40.1, 35.1; HRMS (ESI) m/z: $[M + H]^+$ Calcd for $C_{28}H_{29}N_6O_8$ 577.2041; found 577.2042.

Synthesis of N^2-Methyl-(3,4-methylenedioxy)styrenyl)-O^6-(2-(4-nitrophenyl)ethyl)-5′-(4,4′-dimethoxytrityl)-2′-deoxyguanosine (10). The nucleoside **9** (400 mg, 0.69 mmol) was coevaporated with dry pyridine (5 mL) and dissolved in the same solvent (8 mL). To this, DMT-Cl (70 mg, 0.83 mmol) and DMAP (17 mg, 0.27 mmol) were added and stirred at room temperature for 28 h. After completion of the reaction, the solvent was evaporated under radiused pressure. The residue was dissolved in ethyl acetate (30 mL) and washed with aq. $NaHCO_3$ (30 mL). The combined organic extract was dried over Na_2SO_4 and evaporated under reduced pressure. The crude compound was purified by silica gel column chromatography (70% ethyl acetate in pet ether, 1% NEt_3) to obtain nucleoside **10** as a yellow solid (255 mg, 47%); R_f = 0.3 (70% ethyl acetate in pet ether, 1% NEt_3); mp: 89−91 °C; ^{1}H NMR (400 MHz, $CDCl_3$) δ: 8.11 (d, J = 8.7 Hz, 2H), 7.68 (s, 1H), 7.44−7.37 (m, 4H), 7.27 (d, J = 8.5 Hz, 4H), 7.27−7.15 (m, 4H), 6.85 (s, 1H), 6.77 (dd, J = 9.0, 1.5 Hz, 4H), 6.72 (s, 2H), 6.42 (d, J = 15.7 Hz, 1H), 6.29 (t, J = 6.6 Hz, 1H), 6.04 (dt, J = 15.8, 5.9 Hz, 1H), 5.93 (s, 2H), 4.95 (t, J = 5.1 Hz, 1H), 4.71−4.65 (m, 3H), 4.01 (q, J = 4.8 Hz, 3H), 3.75 (s, 6H), 3.41−3.32 (m, 2H), 3.24 (t, J = 6.5 Hz, 2H), 2.86−2.80 (m, 1H), 2.45−2.39 (m, 1H); ^{13}C{^{1}H} NMR (100 MHz, $CDCl_3$) δ: 160.5, 158.6, 158.5, 153.7, 148.0, 147.1, 146.7., 146.0, 144.5, 137.6, 135.6, 131.2, 130.9, 130.0, 129.9, 128.1, 127.9, 126.9, 124.9, 123.7, 120.8, 115.3, 113.1, 108.2, 105.5, 101.0, 86.5, 85.7, 83.8, 77.2, 72.7, 65.9, 63.9, 55.2, 44.0, 39.5, 35.2; HRMS (ESI) m/z: $[M + H]^+$ Calcd for $C_{49}H_{47}N_6O_{10}$ 879.3348; Found 879.3348.

Synthesis of N^2-Methyl-(3,4-methylenedioxy)styrenyl)-O^6-(2-(4-nitrophenyl)ethyl)-5′-(4,4′-dimethoxytrityl)-2′-deoxyguanosine Phosphoramidite (11). The tritylated nucleoside **10** (100 mg, 0.11 mmol) was dissolved in dry DCM (1 mL), followed by the addition of DIPEA (1 mL, 0.57 mmol) and CEP-Cl (0.03 mL, 0.17 mmol). The reaction was stirred at room temperature for 2 h. After completion of the reaction, MeOH (0.5 mL) was added and stirred for the next 15 min. The reaction mixture was diluted with DCM (30 mL), and washed with $NaHCO_3$ (10 mL), a combined organic extract was dried over Na_2SO_4, and evaporated under reduced pressure. The crude amidite was purified by silica gel column chromatography (45% ethyl acetate in pet ether, 1% NEt_3), to obtain amidite **11** as a yellow solid (80 mg, 65%); R_f = 0.3 (50% ethyl acetate in pet ether, 1% NEt_3); mp: 80−82 °C; ^{31}P NMR (162 MHz, $CDCl_3$) δ: 148.84, 148.71; HRMS (ESI) m/z: $[M + H]^+$ Calcd for $C_{58}H_{63}N_8O_{11}P$ 1079.4427; Found 1079.4425.

Oligonucleotide Synthesis. The unmodified and N^2-SF-dG-modified DNA sequences (Table S1) were synthesized on an automatic DNA/RNA synthesizer. Synthesis was performed on a 1 μmol scale using controlled pore glass (CPG) solid support. To couple the phosphoramidites, 5-ethylthio-1H-trtrazole (ETT) was employed as the coupling agent. The coupling time used for the unmodified phosphoramidite was 90 s, and for the N^2-SF-dG phosphoramidite was 10 min. The N^2-SF-dG-modified oligos were deprotected in four different steps. In the first step, the cyanoethyl group was selectively deprotected using 10% diethylamine in acetonitrile (ACN) at room temperature for 10 min. The supernatant solution is removed, followed by washing with ACN (2 × 100 μL). Then, in the second step, the CPG was treated with 1 M DBU in ACN (1 mL) at room temperature for 1 h to remove the NPE group. The supernatant solution is removed, followed by washing with MeOH and ACN (2 × 100 μL). In the third step, the CPG solid support was cleaved in the presence of 30% aq. NH_3 at room temperature for 3 h. The supernatant layer was collected separately followed by washing of CPG beads with water (3 × 100 μL) and evaporated on a vacuum concentrator. Finally, to remove the base protecting groups, the DNAs were treated with 30% aq. NH_3 at 55 °C for 16 h. The ammonia solution is removed on a vacuum concentrator. The crude DNA was purified using 20% denaturing PAGE (7 M urea) with 1X TBE as a running buffer (89 mM each Tris and boric acid and 2 mM EDTA, pH 8.3) at 30 W for 2.5 h. The gel dimension was 20 × 30 mm with 1 mm thickness. The separation of DNA was visualized under a UV lamp (260 nm) by placing the gel on a silica TLC plate, and the desired DNA band was marked. The desired gel bands were crushed, and 15 mL of TEN buffer (10 mM Tris pH 8.0, 1 mM EDTA, 300 nM NaCl) was added and shaken for 16 h at room temperature to recover the DNA from the gel. Finally, the desalting of oligonucleotides was carried out using a C18-SepPack column. Desalted DNAs were dissolved in autoclaved water, the absorbance was measured at 260 nm, and concentrations of DNA were calculated using appropriate molar extinction coefficients. The chemical integrity of all the oligonucleotides was confirmed by MALDI-TOF in positive reflection mode.

Circular Dichroism Studies. The experiment was performed using a CD instrument in a 1 mm path-length quartz cuvette and spectra were recorded in the range of 220 to 500 nm wavelength. Duplexes were prepared using equal concentrations (10 μM) of unmodified and N^2-SF-dG modified complementary DNAs in a phosphate buffer (100 mM NaCl, 20 mM Na_2PO_4, 0.1 mM EDTA,

https://doi.org/10.1021/acs.joc.4c00368
J. Org. Chem. 2024, 89, 7680−7691

pH 7.5). The samples of the DNA duplexes were annealed at 95 °C for 3 min, cooled to room temperature for 3 h, and stored at 4 °C. The experiments were carried out at 25 °C and scanned at 100 nm/min using a response time of 2s. Each spectrum was an average of 3 measurements, baseline-corrected, and analyzed using Origin (version 9.8).

Thermal Melting Studies. UV-melting studies were performed using the UV−vis instrument coupled with a Peltier module. The experiment was monitored at 260 nm in the temperature range of 15−90 °C with a ramp of 0.5°/ min using three heating−cooling cycles. Duplexes were prepared using equal concentrations (5 μM) of unmodified and N^2-SF-dG modified complementary DNAs in a phosphate buffer (100 mM NaCl, 20 mM Na_2PO_4, 0.1 mM EDTA, pH 7.5). The samples of the DNA duplexes were annealed at 95 °C for 3 min, cooled to room temperature for 3 h, and stored at 4 °C. Samples were equilibrated to room temperature for 10 min before starting the experiments. All the experiments were triplicated, and the T*m* values reported are the average of 3 independent measurements with estimated standard deviation. The absorbance versus temperature data were analyzed using a Marquardt nonlinear curve fitting algorithm.

Expression and Purification of hpol*κ* and hpol*η*. Plasmids coding for human pol*κ* (pB101-hpol*κ* encoding 119−526 amino acids; Uniprot accession Q9UBT6) and pol*η* (pET28a-hpol*η* encoding 1−432 amino acid; Uniprot accession: Q9Y253) were kind gifts from the Guengerich Lab (Dept. of Biochemistry, Vanderbilt University, USA). Both the constructs were transformed into *E. coli* BL21 DE3 RIPL cells and expressed and purified as described below.

E. coli BL21 cells harboring pB101-hpoll were grown at 37 °C under double antibiotic selection (chloramphenicol 25 μg/mL and kanamycin 50 mg/mL) until OD600 reached 0.6. Induction for protein expression was carried out with 1 mM IPTG for 16 h at 16 °C with shaking at 130 rpm. Cells were harvested by centrifugation and resuspended in binding buffer (50 mM Tris-HCl; pH 7.4, 150 mM NaCl, I mM DTT, 20 mM Imidazole). Cells were lysed by sonication, and the lysate was centrifuged at 12500g for 30 min. The supernatant was filtered through a 0.45-μm filter and passed through the His-Trap column pre-equilibrated with a binding buffer. Elution of the protein was carried out using a linear gradient of imidazole. Elute fractions were visualized on SDS-PAGE and the pure fractions were concentrated to 2 mL using Vivaspin Turbo centrifugal devices. The concentrate was treated with 6X-His-HRV3C protease (2*U*/100 ug protein) for 3 h at 4 °C and a similar affinity purification using His-Trap was carried out to remove the 6X His Tag from the protein and the His-tagged HRV3C protease. The flowthrough containing the desired protein was collected and concentrated to 2 mL similarly and loaded onto an AKTA purifier with a 16/600 Superdex column for size exclusion chromatography. The resulting elute fractions were visualized on SDS PAGE, and the purest fractions were pooled to 200 μL and stored as aliquots at −80 °C. The purity of the protein was visualized using SDS-PAGE (Figure S4 of Supporting Information). Expression and purification of hpol*η* were carried out using an identical protocol.[5]

Single-Nucleotide Primer Extension Studies. The primer extension reaction was performed using a FAM-labeled primer/template DNA substrate. The duplex was prepared using equal concentration (1:1) of 15-mer FAM labeled primer and 27-mer unmodified or N^2-SF-dG modified DNA template. The samples of the DNA duplexes were annealed at 95 °C for 3 min and cooled to room temperature for 3 h. The primer/template DNA substrate (0.5 μM) was incubated with hpol*κ* or hpol*η* (5 nM) in polymerase buffer containing 50 mM Tris-Cl (pH 7.5), 50 mM NaCl, 5 mM $MgCl_2$, 5% glycerol in water (v/v), 5 mM DTT, and 0.1 mg/mL, BSA. Then, the primer extension reaction is initiated by adding varying concentrations (0.1, 1, 10, 20, 40 μM) of all four individual dNTPs (total volume of reaction 10 μL). The reaction was incubated at 37 °C for 10 min and stopped using gel loading dye containing 1 mg/mL bromophenol blue, 20 mM EDTA, and 80% formamide (v/v). Samples were heated at 95 °C and separated on 20% PAGE (7 M urea). The product formation was visualized using a fluorescence scanner.

Standing Start and Running Start Full-Length Primer Extension Studies. A standing start full-length extension reaction was carried out by preparing primer/template DNA substrate in the same way as described in the above reaction. For the running start studies, a primer/template DNA substrate was prepared by annealing a 13-mer FAM labeled DNA with 27-mer unmodified or 27-mer N^2-SF-dG modified DNA template. The primer templet DNA substrate (0.5 μM) was incubated at 37 °C with hpol*κ* or hpol*η* (5 nM) in polymerase buffer. The primer extension reaction was initiated by adding 3 μL of an equimolar mixture of all four dNTPs (final concentration of 80 μM) to a total reaction volume of 30 μL. Aliquots of 5 μL were drawn at 5, 10, 30, 60, and 120 min and quenched using 5 μL of gel loading dye. Samples were heated at 95 °C and separated on 20% PAGE (7 M urea). The product formation was visualized using a fluorescence scanner.

Molecular Modeling and Dynamics Studies. The starting structures from the X-ray crystal structure of lucidin adducted DNA in complex with human DNA polymerase *κ* (PDB ID: 5W2A and 5W2C, respectively, for insertion and postinsertion) were used for MD simulations.[45] The complex preparation for the preinsertion, insertion, and postinsertion was carried out using the previously reported protocol.[5] The structure of the safrole adduct was generated using Gauss View 6.0 (Figure S1) and loaded into the R.E.D Server along with a dimethylphosphate molecule to generate the force field parameters needed to define the adduct.[46−48] The modified nucleotide was incorporated into the DNA system using the xleap module of AMBER 1849, and the coordinate and topology files were generated. The AMBER force fields ff14SB[49] and parmbsc1[50] were used for protein and DNA, respectively. The system was then neutralized by adding Na^+ ions and solvated using a TIP3P water box extending 10 Å from the atoms.

The MD simulation for all three complexes was performed using the previously utilized procedure.[5] Briefly, multiple stages of restraint minimization and 8 stages of equilibration were performed. Furthermore, 100 ps of equilibration was run under NPT conditions without restraints. Finally, 500 ns of the production run was carried out using the GPU-accelerated version of PMEMD[51,52] in AMBER 18. All the bonds involving hydrogen atoms were constrained using the SHAKE algorithm with a time step of 2 fs. Analysis of the MD trajectory was performed by the cpptraj[53] module of AMBER 18. The RMSD calculations were performed excluding the modeled loop. The pairwise energy calculations were performed using the MM/GBSA[54] module in the AmberTools 19. The trajectories were visualized using PyMOL (Schrodinger LLC).

■ ASSOCIATED CONTENT

Data Availability Statement

The data underlying this study are available in the published article and its Supporting Information.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.joc.4c00368.

Supporting Figures and Tables from molecular modeling studies, oligonucleotide sequences used for studies, ^{1}H, ^{13}C, and ^{31}P NMR spectra of all compounds, and MALDI spectra of modified oligonucleotides (PDF)

Accession Codes

hpol*κ*: UniProt Accession code- Q9UBT6. hpol*η*: UniProt Accession code- Q9Y253.

■ AUTHOR INFORMATION

Corresponding Authors

Kiran Kondabagil − *Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay,*

https://doi.org/10.1021/acs.joc.4c00368
J. Org. Chem. 2024, 89, 7680−7691

Mumbai 400076, India; orcid.org/0000-0002-7942-023X; Email: kirankondabagil@iitb.ac.in

P. I. Pradeepkumar − *Department of Chemistry, Indian Institute of Technology Bombay, Mumbai 400076, India;* orcid.org/0000-0001-9104-3708; Email: pradeep@chem.iitb.ac.in

Authors

Siddharam Shivappa Bagale − *Department of Chemistry, Indian Institute of Technology Bombay, Mumbai 400076, India*

Priyanka U. Deshmukh − *Department of Chemistry, Indian Institute of Technology Bombay, Mumbai 400076, India*

Shailesh B. Lad − *Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Mumbai 400076, India*

Akhil Sudarsan − *Department of Chemistry, Indian Institute of Technology Bombay, Mumbai 400076, India*

Sruthi Sudhakar − *Department of Chemistry, Indian Institute of Technology Bombay, Mumbai 400076, India*

Soumyadeep Mandal − *Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Mumbai 400076, India*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.joc.4c00368

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work is financially supported by grants from the Science and Engineering Research Board (SERB, CRG/2021/001992CRG/2021/001992), the Government of India to P.I.P. We are thankful to Prof. F. Peter Guengrich (Vanderbilt University), for his generous gift of plasmid constructs pBG101-hpolκ and pET28a-hpolη. We thank IRCC- IIT Bombay for providing access to the central MALDI facility and the Spacetime IIT Bombay for the HPC facilities. We are also grateful to the Institute of Eminence (IOE) /C-supported central facility (Department of Chemistry) for providing access to DNA synthesizer and DST-FIST (SR/FST/CS-II/2017/37) support for NMR and HRMS facilities. We thank Tanvi Aggarwal for her help with protein purification. S.S.B. thanks IIT Bombay, P.U.D. and S.M. thanks the Council of Scientific and Industrial Research (CSIR), S.S. thanks Prime Minister's Research Fellowship (PMRF) and S.B.L. thanks IRCC-IIT Bombay for the Ph.D. fellowships.

■ DEDICATION

Dedicated to Professor Jyoti Chattopadhyaya on the occasion of his 75th birthday.

■ REFERENCES

(1) Gates, K. S. An Overview of Chemical Processes That Damage Cellular DNA: Spontaneous Hydrolysis, Alkylation, and Reactions with Radicals. *Chem. Res. Toxicol.* **2009**, *22* (11), 1747−1760.

(2) Barnes, J. L.; Zubair, M.; John, K.; Poirier, M. C.; Martin, F. L. Carcinogens and DNA Damage. *Biochem. Soc. Trans.* **2018**, *46* (5), 1213−1224.

(3) Bist, I.; Bhakta, S.; Jiang, D.; Keyes, T. E.; Martin, A.; Forster, R. J.; Rusling, J. F. Evaluating Metabolite-Related DNA Oxidation and Adduct Damage from Aryl Amines Using a Microfluidic ECL Array. *Anal. Chem.* **2017**, *89* (22), 12441−12449.

(4) Turesky, R. J.; Le Marchand, L. Metabolism and Biomarkers of Heterocyclic Aromatic Amines in Molecular Epidemiology Studies: Lessons Learned from Aromatic Amines. *Chem. Res. Toxicol.* **2011**, *24* (8), 1169−1214.

(5) Deshmukh, P. U.; Lad, S. B.; Sudarsan, A.; Sudhakar, S.; Aggarwal, T.; Mandal, S.; Bagale, S. S.; Kondabagil, K.; Pradeepkumar, P. I. Human Translesion Synthesis Polymerases polκ and polη Perform Error-Free Replication across N^2-dG Methyleugenol and Estragole DNA Adducts. *Biochemistry* **2023**, *62* (16), 2391−2406.

(6) Hwa Yun, B.; Guo, J.; Bellamri, M.; Turesky, R. J. DNA Adducts: Formation, Biological Effects, and New Biospecimens for Mass Spectrometric Measurements in Humans. *Mass Spectrom. Rev.* **2020**, *39* (1−2), 55−82.

(7) Chatterjee, N.; Walker, G. C. Mechanisms of DNA Damage, Repair, and Mutagenesis. *Environ. Mol. Mutagen.* **2017**, *58* (5), 235−263.

(8) Jackson, S. P.; Bartek, J. The DNA-Damage Response in Human Biology and Disease. *Nature* **2009**, *461* (7267), 1071−1078.

(9) Wang, J. Y. J. Cell Death Response to DNA Damage. *Yale J. Biol. Med.* **2019**, *92* (4), 771−779.

(10) Muradian, K.; Schachtschabel, D. O. The Role of Apoptosis in Aging and Age-Related Disease: Update. *Z. Gerontol. Geriatr.* **2001**, *34* (6), 441−446.

(11) Vaisman, A.; Woodgate, R. Translesion DNA Polymerases in Eukaryotes: What Makes Them Tick? *Crit. Rev. Biochem. Mol. Biol.* **2017**, *52* (3), 274−303.

(12) Nelson, J. R.; Lawrence, C. W.; Hinkle, D. C. Thymine-Thymine Dimer Bypass by Yeast DNA Polymerase ζ. *Science (80-.).* **1996**, *272* (5268), 1646−1649.

(13) Yang, W.; Gao, Y. Translesion and Repair DNA Polymerases: Diverse Structure and Mechanism. *Annu. Rev. Biochem.* **2018**, *87*, 239−261.

(14) Pata, J. D. Structural Diversity of the Y-Family DNA Polymerases. *Biochim. Biophys. Acta* **2010**, *1804* (5), 1124−1135.

(15) 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. Two-Polymerase Mechanisms Dictate Error-Free and Error-Prone Translesion DNA Synthesis in Mammals. *EMBO J.* **2009**, *28* (4), 383−393.

(16) Yoon, J. H.; Prakash, L.; Prakash, S. Highly Error-Free Role of DNA Polymerase η in the Replicative Bypass of UV-Induced Pyrimidine Dimers in Mouse and Human Cells. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (43), 18219−18224.

(17) Bose, A.; Millsap, A. D.; Deleon, A.; Rizzo, C. J.; Basu, A. K. Translesion Synthesis of the N2−2′-Deoxyguanosine Adduct of the Dietary Mutagen IQ in Human Cells: Error-Free Replication by DNA Polymerase κ and Mutagenic Bypass by DNA Polymerases η, ζ, and Rev1. *Chem. Res. Toxicol.* **2016**, *29* (9), 1549−1559.

(18) Zhang, Y.; Yuan, F.; Wu, X.; Wang, M.; Rechkoblit, O.; Taylor, J. S.; Geacintov, N. E.; Wanga, Z. Error-Free and Error-Prone Lesion Bypass by Human DNA Polymerase κ in Vitro. *Nucleic Acids Res.* **2000**, *28* (21), 4138−4146.

(19) Kamdem, D. P.; Gage, D. A. Chemical Composition of Essential Oil from the Root Bark of Sassafras Albidum. *Planta Med.* **1995**, *61* (6), 574−575.

(20) Saputri, F. A.; Mutakin, M.; Lestari, K.; Levita, J. Determination of Safrole in Ethanol Extract of Nutmeg (Myristica Fragrans Houtt) Using Reversed-Phase High Performance Liquid Chromatography. *Int. J. Chem.* **2014**, *6* (3), 14−20.

(21) Chung, Y. T.; Chen, C. L.; Wu, C. C.; Chan, S. A.; Chi, C. W.; Liu, T. Y. Safrole-DNA Adduct in Hepatocellular Carcinoma Associated with Betel Quid Chewing. *Toxicol. Lett.* **2008**, *183* (1−3), 21−27.

(22) Bolton, J. L.; Acay, N. M.; Vukomanovic, V. Evidence that 4-allyl-o-quinones Spontaneously Rearrange to their More Electrophilic Quinone Methides: Potential BioactivationMechanism for the Hepatocarcinogen Safrole. *Chem. Res. Toxicol.* **1994**, *7* (3), 443−450.

(23) Chen, C. L.; Chi, C. W.; Chang, K. W.; Liu, T. Y. Safrole-like DNA Adducts in Oral Tissue from Oral Cancer Patients with a Betel Quid Chewing History. *Carcinogenesis* **1999**, *20* (12), 2331−2334.
(24) Nokovitch, L.; Maquet, C.; Crampon, F.; Taihi, I.; Roussel, L. M.; Obongo, R.; Virard, F.; Fervers, B.; Deneuve, S. Oral Cavity Squamous Cell Carcinoma Risk Factors: State of the Art. *J. Clin. Med.* **2023**, *12* (9), 3264.
(25) Kobets, T.; Cartus, A. T.; Fuhlbrueck, J. A.; Brengel, A.; Stegmüller, S.; Duan, J. D.; Brunnemann, K. D.; Williams, G. M. Assessment and Characterization of DNA Adducts Produced by Alkenylbenzenes in Fetal Turkey and Chicken Livers. *Food Chem. Toxicol.* **2019**, *129* (March), 424−433.
(26) Takeshita, T.; Tao, F.; Kojima, N.; Kanaly, R. A. Triple Quadrupole Mass Spectrometry Comparative DNA Adductomics of Hep G2 Cells Following Exposure to Safrole. *Toxicol. Lett.* **2019**, *300*, 92−104.
(27) IARC. *Agents Classified by the IARC Monographs*; IARC, 2014; Vol. *1*−*132*, pp 1−5.
(28) Jeurissen, S. M. F.; Bogaards, J. J. P.; Awad, H. M.; Boersma, M. G.; Brand, W.; Fiamegos, Y. C.; Van Beek, T. A.; Alink, G. M.; Sudhölter, E. J. R.; Cnubben, N. H. P.; Rietjens, I. M. C. M. Human Cytochrome P450 Enzyme Specificity for Bioactivation of Safrole to the Proximate Carcinogen 1′-Hydroxysafrole. *Chem. Res. Toxicol.* **2004**, *17* (9), 1245−1250.
(29) Ueng, Y. F.; Hsieh, C. H.; Don, M. J.; Chi, C. W.; Ho, L. K. Identification of the Main Human Cytochrome P450 Enzymes Involved in Safrole 1′-Hydroxylation. *Chem. Res. Toxicol.* **2004**, *17* (8), 1151−1156.
(30) Gamage, N.; Barnett, A.; Hempel, N.; Duggleby, R. G.; Windmill, K. F.; Martin, J. L.; McManus, M. E. Human Sulfotransferases and Their Role in Chemical Metabolism. *Toxicol. Sci.* **2006**, *90* (1), 5−22.
(31) Hemeryck, L. Y.; Vanhaecke, L. Diet-Related DNA Adduct Formation in Relation to Carcinogenesis. *Nutr. Rev.* **2016**, *74* (8), 475−489.
(32) Landge, V. G.; Bonds, A. L.; Mncwango, T. A.; Mather, C. B.; Saleh, Y.; Fields, H. L.; Lee, F.; Young, M. C. Amine-Directed Mizoroki-Heck Arylation of Free Allylamines. *Org. Chem. Front.* **2022**, *9* (7), 1967−1974.
(33) Alsharif, M. A.; Raja, Q. A.; Majeed, N. A.; Jassas, R. S.; Alsimaree, A. A.; Sadiq, A.; Naeem, N.; Mughal, E. U.; Alsantali, R. I.; Moussa, Z.; Ahmed, S. A. DDQ as a Versatile and Easily Recyclable Oxidant: A Systematic Review. *RSC Adv.* **2021**, *11* (47), 29826−29858.
(34) El-Batta, A.; Jiang, C.; Zhao, W.; Anness, R.; Cooksy, A. L.; Bergdahl, M. Wittig Reactions in Water Media Employing Stabilized Ylides with Aldehydes. Synthesis of α,β-Unsaturated Esters from Mixing Aldehydes, α-Bromoesters, and Ph_3P in Aqueous $NaHCO_3$. *J. Org. Chem.* **2007**, *72* (14), 5244−5259.
(35) Angle, S. R.; Choi, I.; Tham, F. S. Stereoselective Synthesis of 3-Alkyl-2-Aryltetrahydrofuran-4-Ols: Total Synthesis of (±)-Paulownin. *J. Org. Chem.* **2008**, *73* (16), 6268−6278.
(36) Vidya Sagar Reddy, G.; Venkat Rao, G.; Subramanyam, R. V. K.; Iyengar, D. S. A New Novel and Practical One Pot Methodology for Conversion of Alcohols to Amines. *Synth. Commun.* **2000**, *30* (12), 2233−2237.
(37) Ghodke, P. P.; Harikrishna, S.; Pradeepkumar, P. I. Synthesis and Polymerase-Mediated Bypass Studies of the N^2-Deoxyguanosine DNA Damage Caused by a Lucidin Analogue. *J. Org. Chem.* **2015**, *80* (4), 2128−2138.
(38) Ghodke, P. P.; Pradeepkumar, P. I. Synthesis of N^2-Aryl-2′-Deoxyguanosine Modified Phosphoramidites and Oligonucleotides. *Curr. Protoc. Nucleic Acid Chem.* **2019**, *78* (1), No. e93.
(39) Lee, T. H.; Kang, T. H. DNA Oxidation and Excision Repair Pathways. *Int. J. Mol. Sci.* **2019**, *20* (23), 6092.
(40) Dip, R.; Camenisch, U.; Naegeli, H. Mechanisms of DNA Damage Recognition and Strand Discrimination in Human Nucleotide Excision Repair. *DNA Repair (Amst).* **2004**, *3* (11), 1409−1423.
(41) Min, J. H.; Pavletich, N. P. Recognition of DNA Damage by the Rad4 Nucleotide Excision Repair Protein. *Nature* **2007**, *449* (7162), 570−575.
(42) Choi, J. Y.; Angel, K. C.; Guengerich, F. P. Translesion Synthesis across Bulky N2-Alkyl Guanine DNA Adducts by Human DNA Polymerase κ. *J. Biol. Chem.* **2006**, *281* (30), 21062−21072.
(43) Choi, J. Y.; Guengerich, F. P. Adduct Size Limits Efficient and Error-Free Bypass across Bulky N 2-Guanine DNA Lesions by Human DNA Polymerase η. *J. Mol. Biol.* **2005**, *352* (1), 72−90.
(44) Kathuria, P.; Singh, P.; Sharma, P.; Wetmore, S. D. Replication of the Aristolochic Acid I Adenine Adduct (ALI-N6-A) by a Model Translesion Synthesis DNA Polymerase: Structural Insights on the Induction of Transversion Mutations from Molecular Dynamics Simulations. *Chem. Res. Toxicol.* **2020**, *33* (10), 2573−2583.
(45) Yockey, O. P.; Jha, V.; Ghodke, P. P.; Xu, T.; Xu, W.; Ling, H.; Pradeepkumar, P. I.; Zhao, L. Mechanism of Error-Free DNA Replication Past Lucidin-Derived DNA Damage by Human DNA Polymerase κ. *Chem. Res. Toxicol.* **2017**, *30* (11), 2023−2032.
(46) 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−W517.
(47) 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* (28), 7821−7839.
(48) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97* (40), 10269−10280.
(49) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696−3713.
(50) Ivani, I.; Dans, P. D.; Noy, A.; Pérez, A.; Faustino, I.; Hospital, A.; Walther, J.; Andrio, P.; Goñi, R.; Balaceanu, A.; Portella, G.; Battistini, F.; Gelpí, J. L.; González, C.; Vendruscolo, M.; Laughton, C. A.; Harris, S. A.; Case, D. A.; Orozco, M. Parmbsc1: A Refined Force Field for DNA Simulations. *Nat. Methods* **2016**, *13* (1), 55−58.
(51) 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* (9), 3878−3888.
(52) 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* (5), 1542−1555.
(53) Roe, D. R.; Cheatham, T. E. I. I. I. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9* (7), 3084−3095.
(54) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* **2000**, *33* (12), 889−897.

https://doi.org/10.1021/acs.joc.4c00368
J. Org. Chem. 2024, 89, 7680−7691

Synthesis of N^2 -trans-isosafrole-dG -adduct bearing DNAs and the bypass studies with human TLS polymerases κ and η

Siddharam Shivappa Bagale, [†] Priyanka U. Deshmukh, [†]Shailesh B. Lad, [§] Akhil Sudarsan, [†] Sruthi Sudhakar, [†] Soumyadeep Mandal, [§] Kiran Kondabagil,*, [§] P. I. Pradeepkumar*, [†]

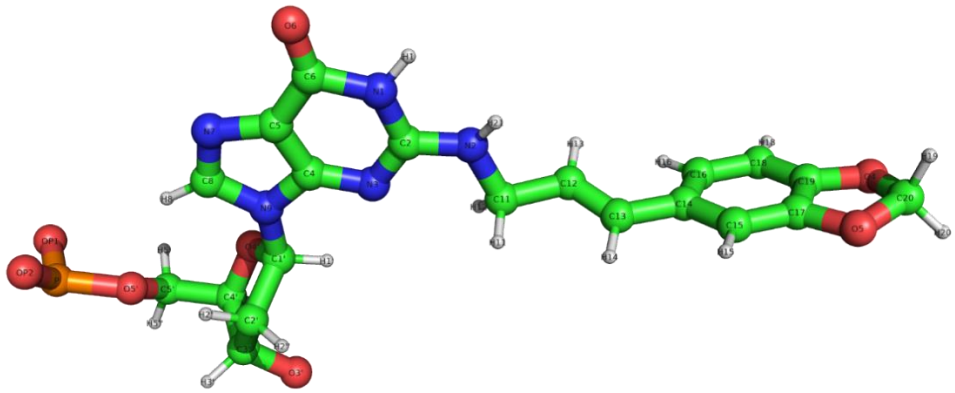
[†] *Department of Chemistry, Indian Institute of Technology Bombay, Powai, Mumbai-400076, India*

[§] *Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Powai, Mumbai-400076, India*

Email: pradeep@chem.iitb.ac.in or kirankondabagil@iitb.ac.in

Table of Content

Cartesian coordinates and RESP charges calculated for N^2-SF-dG



@<TRIPOS>MOLECULE
U0
53 57 1 0 1
SMALL
USER_CHARGES
@<TRIPOS>ATOM

1	O5'	-2.68123	0.862979	-0.51024	OS	1	U0	-0.4899	0	****
2	P	-4.11905	0.976273	-1.29514	P	1	U0	1.2125	0	****
3	OP1	-4.40108	2.395731	-1.55481	O2	1	U0	-0.7918	0	****
4	OP2	-5.03434	0.048399	-0.61461	O2	1	U0	-0.7918	0	****
5	O3'	1.407474	0	0	OS	1	U0	-0.5226	0	****
6	C5'	-1.66846	1.759507	-0.84413	CT	1	U0	-0.0171	0	****
7	H5'	-1.97959	2.782952	-0.65964	H1	1	U0	0.0803	0	****
8	H5"	-1.40339	1.665498	-1.89535	H1	1	U0	0.0803	0	****
9	C4'	-0.44417	1.469488	0	CT	1	U0	0.118	0	****
10	H4'	0.373397	2.083422	-0.35732	H1	1	U0	0.1248	0	****
11	O4'	-0.71052	1.803494	1.345369	OS	1	U0	-0.3704	0	****
12	C1'	-0.49509	0.721274	2.205245	CT	1	U0	0.137	0	****
13	H1'	0.469832	0.806713	2.68096	H2	1	U0	0.1377	0	****
14	N9	-1.47631	0.761404	3.259563	N*	1	U0	0.0023	0	****
15	C8	-2.84191	0.882866	3.117289	CK	1	U0	0.1172	0	****
16	H8	-3.28494	0.951318	2.146648	H5	1	U0	0.1839	0	****
17	N7	-3.46488	0.917107	4.233036	NB	1	U0	-0.5621	0	****
18	C5	-2.47642	0.818338	5.187688	CB	1	U0	0.0841	0	****
19	C6	-2.55499	0.813002	6.620273	C	1	U0	0.6469	0	****
20	O6	-3.49247	0.888574	7.357316	O	1	U0	-0.5705	0	****

21	N1	-1.25617	0.701639	7.169666	NA	1	U0	-0.6773	0	****
22	H1	-1.23423	0.756312	8.165454	H	1	U0	0.3636	0	****
23	C2	-0.08869	0.61658	6.473386	CA	1	U0	0.7225	0	****
24	N2	1.042371	0.55893	7.221253	N2	1	U0	-0.5491	0	****
25	H21	0.939631	0.230846	8.156098	H	1	U0	0.3361	0	****
26	N3	-0.02948	0.610967	5.184789	NC	1	U0	-0.6109	0	****
27	C4	-1.2465	0.729669	4.596288	CB	1	U0	0.2353	0	****
28	C3'	0	0	0	CT	1	U0	0.08	0	****
29	H3'	-0.37976	-0.53762	-0.86114	H1	1	U0	0.1155	0	****
30	C2'	-0.57067	-0.51768	1.313802	CT	1	U0	-0.1206	0	****
31	H2'	-1.60119	-0.81688	1.179433	HC	1	U0	0.0611	0	****
32	H2"	-0.0126	-1.35252	1.722065	HC	1	U0	0.0611	0	****
33	C11	2.340355	0.306813	6.611548	CT	1	U0	0.0224	0	****
34	H11	2.409453	-0.71137	6.239812	H1	1	U0	0.0776	0	****
35	H12	2.431195	0.966752	5.760194	H1	1	U0	0.0776	0	****
36	C12	3.42587	0.583825	7.611211	CM	1	U0	-0.1371	0	****
37	H13	3.450933	1.586002	8.006384	HA	1	U0	0.1208	0	****
38	C13	4.308383	-0.31955	8.006456	CD	1	U0	-0.1561	0	****
39	H14	4.223736	-1.3228	7.618605	HA	1	U0	0.158	0	****
40	C14	5.419478	-0.10789	8.961417	CD	1	U0	-0.0259	0	****
41	C15	5.891118	-1.21663	9.689377	CA	1	U0	-0.2637	0	****
42	H15	5.46831	-2.19496	9.554973	HA	1	U0	0.1792	0	****
43	C16	6.005306	1.135285	9.151797	CA	1	U0	-0.151	0	****
44	H16	5.671399	1.97247	8.568319	HA	1	U0	0.1674	0	****
45	C17	6.901346	-1.00605	10.58127	CB	1	U0	0.2773	0	****
46	C18	7.041659	1.332429	10.06928	CA	1	U0	-0.3465	0	****
47	H18	7.494126	2.296166	10.20662	HA	1	U0	0.2064	0	****
48	C19	7.464658	0.239966	10.77175	CB	1	U0	0.3387	0	****
49	C20	8.386656	-1.16123	12.21753	CT	1	U0	0.2866	0	****
50	H19	7.983762	-1.13105	13.22222	H2	1	U0	0.0803	0	****
51	H20	9.369509	-1.60464	12.20937	H2	1	U0	0.0803	0	****
52	O5	7.526171	-1.90332	11.38515	OS	1	U0	-0.4088	0	****
53	O4	8.453246	0.145295	11.69552	OS	1	U0	-0.4096	0	****

@<TRIPOS>BOND

1	1	2	1
2	1	6	1
3	2	3	1
4	2	4	1
5	5	28	1

6	6	7	1
7	6	8	1
8	6	9	1
9	9	10	1
10	9	11	1
11	9	28	1
12	11	12	1
13	12	13	1
14	12	14	1
15	12	30	1
16	14	15	1
17	14	27	1
18	15	16	1
19	15	17	1
20	17	18	1
21	18	19	1
22	18	27	1
23	19	20	1
24	19	21	1
25	21	22	1
26	21	23	1
27	23	24	1
28	23	26	1
29	24	25	1
30	24	33	1
31	26	27	1
32	28	29	1
33	28	30	1
34	30	31	1
35	30	32	1
36	33	34	1
37	33	35	1
38	33	36	1
39	36	37	1
40	36	38	1
41	38	39	1
42	38	40	1
43	40	41	1
44	40	43	1

```
45          41      42          1
46          41      45          1
47          43      44          1
48          43      46          1
49          45      48          1
50          45      52          1
51          46      47          1
52          46      48          1
53          48      53          1
54          49      50          1
55          49      51          1
56          49      52          1
57          49      53          1
@<TRIPOS>SUBSTRUCTUR
E
1           U0      1           ****        0           ****    ****
@<TRIPOS>HEADTAIL
P           1
O3'         1
@<TRIPOS>RESIDUECONNECT
1           P       O3'         0           0           0       0
```

Figure S1. Cartesian coordinates and RESP charges calculated for the N^2-SF-dG adduct generated using R.E.D Server. Carbons atoms are represented by green, nitrogen atoms by blue, oxygen atoms by red, phosphorus atoms by orange and hydrogen atoms by white respectively.

Root Mean Square Deviations (RMSD) of the protein and DNA residues

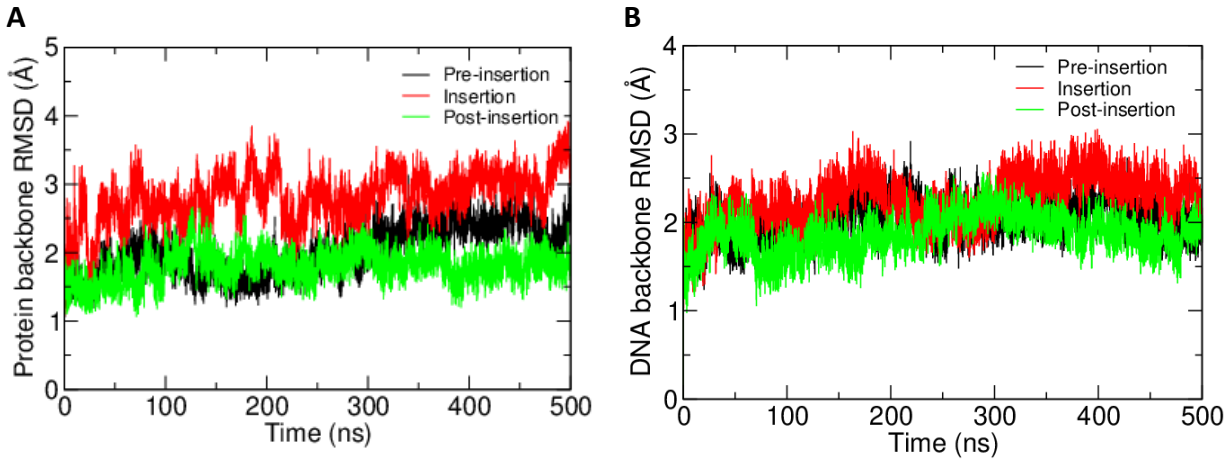
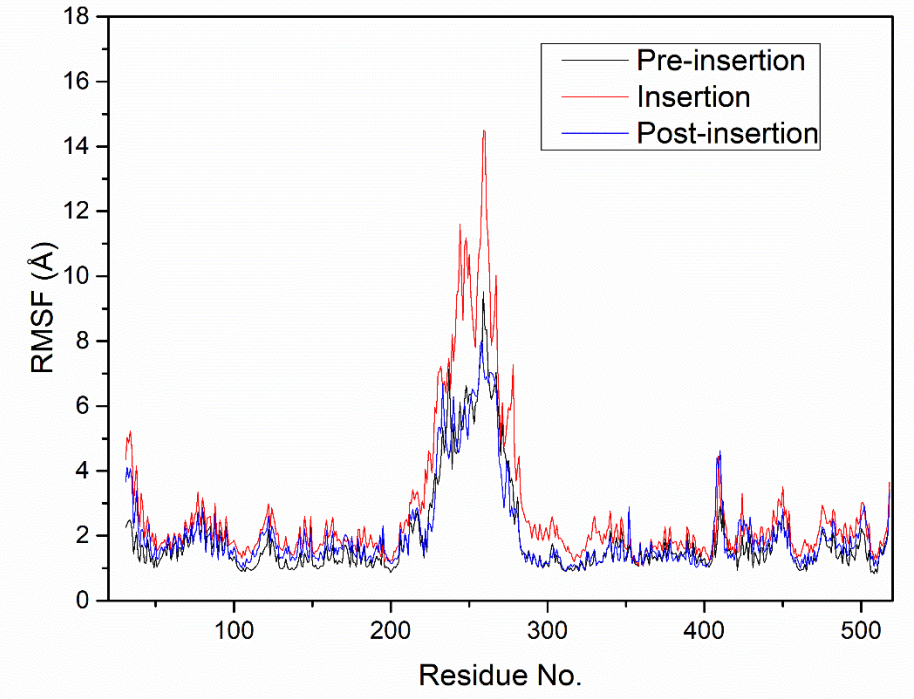


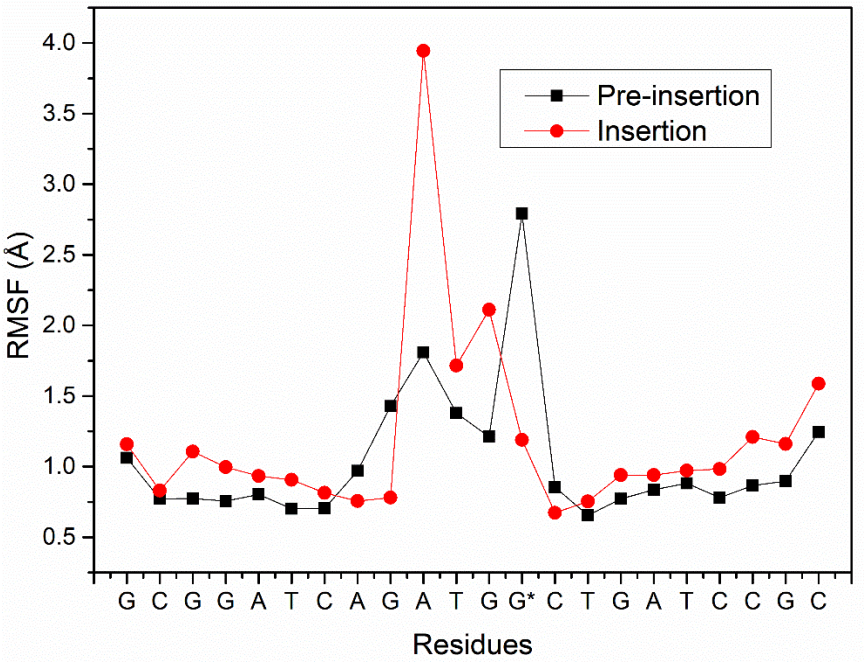
Figure S2. (A) Plot of the RMSD values of protein backbone atoms. and (B) Plot of RMSD values of DNA backbone atoms. The whole trajectory of 500 ns was used for RMSD calculations. The modelled loop is excluded in the protein RMSD plot.

Root Mean Square Fluctuations plots (RMSF) of the protein and DNA residues

A



B



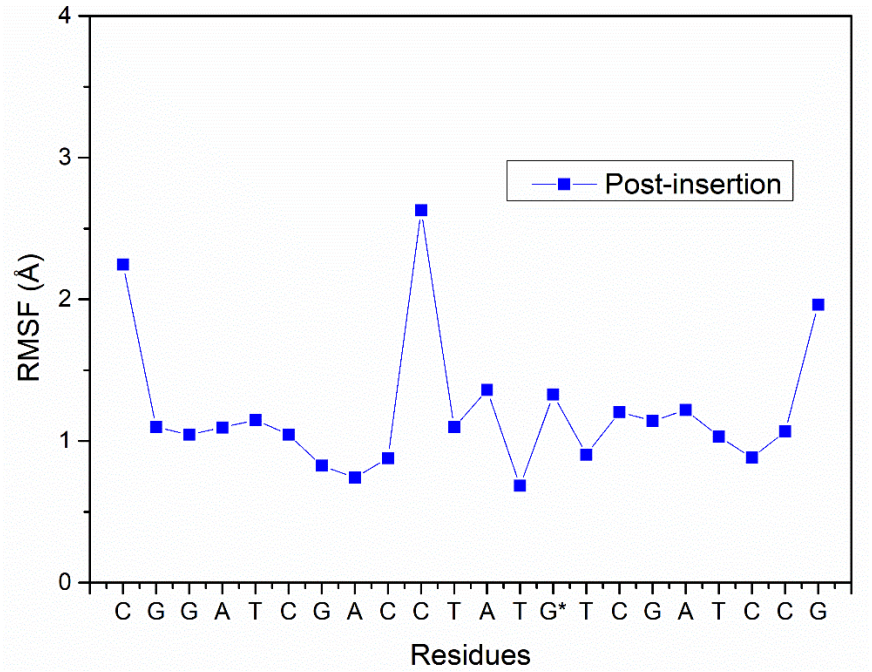


Figure S3. (A) RMSF plot of amino acid residues in protein, and (B) RMSF plots of nucleotides in DNA, the N^2-SF-dG is represented by G*. The black, red, and blue lines represent the pre-insertion, insertion, and post-insertion stages respectively. The whole trajectory of 500 ns was used for RMSF calculations.

Superimposed images of the major clusters of pre-insertion and post-insertion complexes

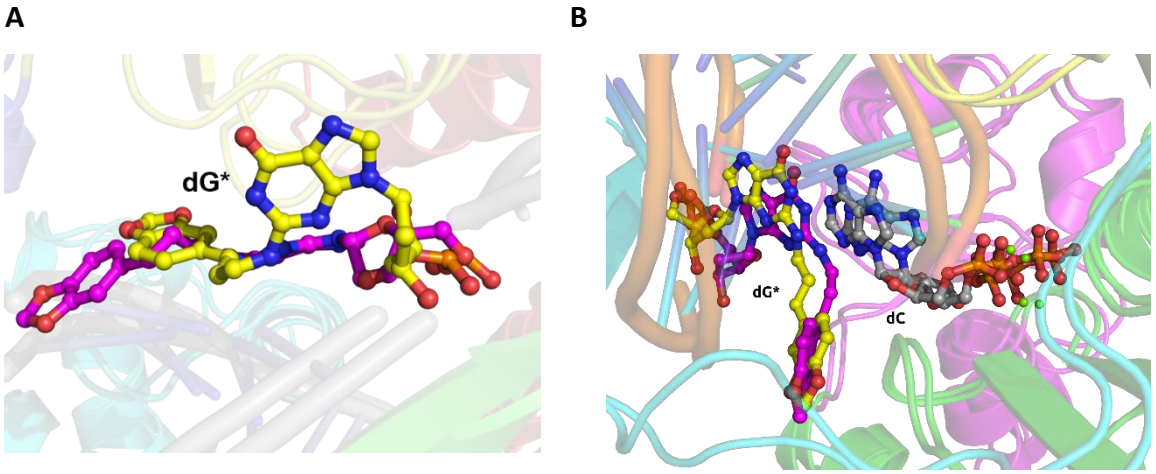


Figure S4. Superimposed images of the representative structures of the major clusters of (A) pre-insertion, and (B) post-insertion complexes. The carbon atoms of cluster 1 are represented in magenta, and those of cluster 2 are represented in yellow, oxygen atoms are in red, nitrogen atoms are in blue, and phosphate atoms are in orange. dG* represents the N^2-SF-dG adduct.

SDS PAGE of purified hpolκ and hpolη

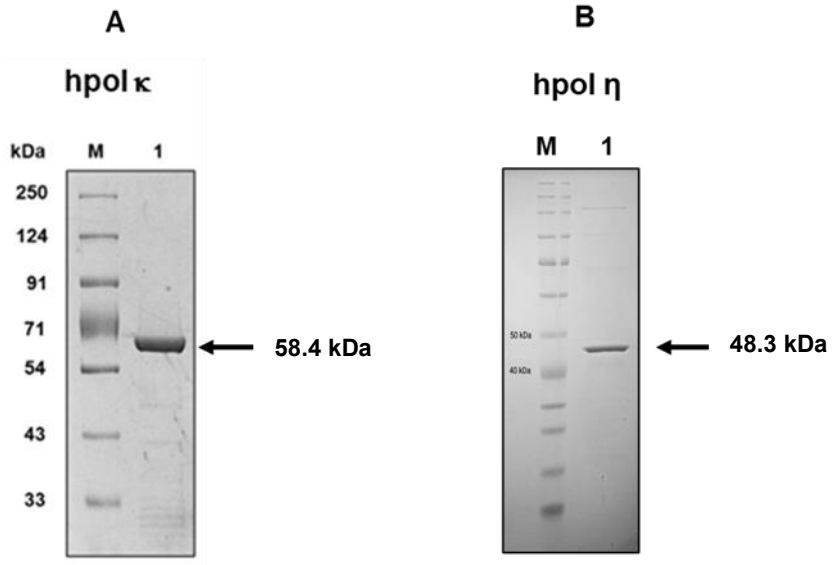


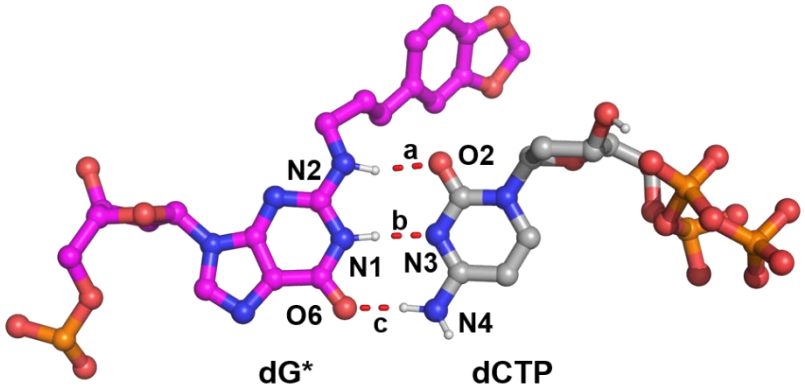
Figure S5. (A) SDS PAGE of purified protein hpol κ in lane 1 at 58.4 kDa, and (B) SDS PAGE of purified protein hpol η in lane 1 at 48.3 kDa

Unmodified and N^2-SF-dG modified DNAs and their molecular weights

Code	Sequence	MW (calc.)	MW (found)
G1	5'-GCCGGAATAGCGCA-3'	4299	4304.3
G2	5'-TGCGCTATTCCGGC-3'	4233	4236.8
G3	5'-GCCGGAATAGCGCA-3'	4458	4460.3
G4	5'-CTGGTCACACTGATGCCTACGAGTACG-3'	8420	8420.0

Table S1. Unmodified and N^2-SF-dG-modified DNAs were characterized using MALDI-TOF in linear positive reflection mode. The SF-dG nucleotide is shown in red colour and underlined.

Percentage occupancies of H-bonds between N^2-SF-dG and dCTP in the insertion stage complex



Bond	Donor Atom	Acceptor Atom	Percentage Occupancy	Average Distance	Average Angle
a	dG*@N2	dCTP@O2	92	3 Å	163°
b	dG*@N1	dCTP@N3	93	2.9 Å	162°
c	dCTP@N4	dG*@O6	94	2.9 Å	163°

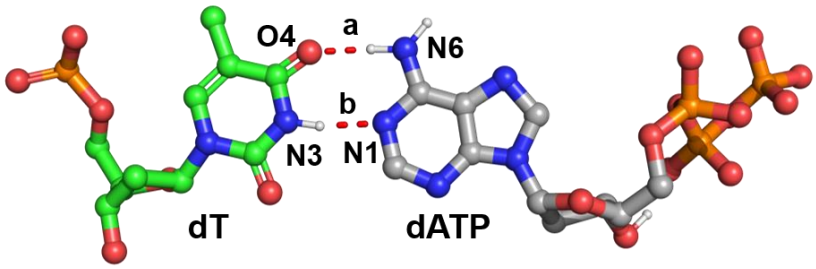
Table S2. The figure represents the Watson-Crick hydrogen bonds between the N^2-SF-dG (dG*) and the incoming nucleotide (dCTP) in the insertion stage. The hydrogen bonds are illustrated as dotted lines. Hydrogen bonds are determined using distance and angle cut-off of 3.4 Å and 120° respectively.

Pairwise interaction energies of the N^2-dG and N^2-SF-dG with the surrounding amino acids in the active site of hpolκ

Residue 1	Residue 2	N^2-SF-dG (kcal/mol)
G* 13	TYR 112	**-2.1 ± 0.7**
G* 13	SER 134	-0.7 ± 0.4
G* 13	MET 135	**-4.2 ± 0.5**
G* 13	SER 137	**-1.5 ± 0.4**
G* 13	PHE 171	**-2.8 ± 0.7**
G* 13	ARG 175	-1.3 ± 0.8
G* 13	ARG 420	**-2.0 ± 0.6**

Table S3. Pairwise interaction energies between the unmodified-dG and the N^2-SF-dG adduct (Residue 1) with the amino acids (Residue 2) present in the active site of hpolκ obtained from MM/GBSA analysis of the insertion stage. The important interactions are highlighted in bold.

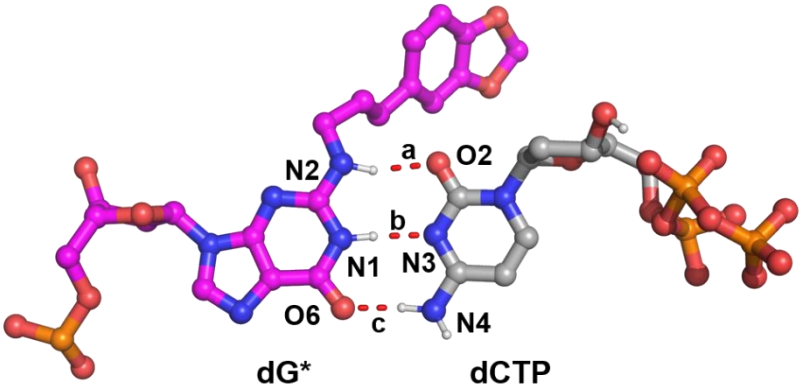
Percentage occupancies of H-bonds between dT and dATP in the post-insertion stage complex



Bond	Donor Atom	Acceptor Atom	Percentage Occupancy	Average Distance	Average Angle
a	dATP@N6	dT@O4	97	3 Å	164°
b	dT@N3	dATP@N1	99	2.9 Å	164°

Table S4. The figure represents the Watson-Crick hydrogen bonds between the template base (dT) and the incoming nucleotide (dATP) in the post-insertion stage. The hydrogen bonds are illustrated as dotted lines. Hydrogen bonds are determined using distance and angle cut-off of 3.4 Å and 120° respectively.

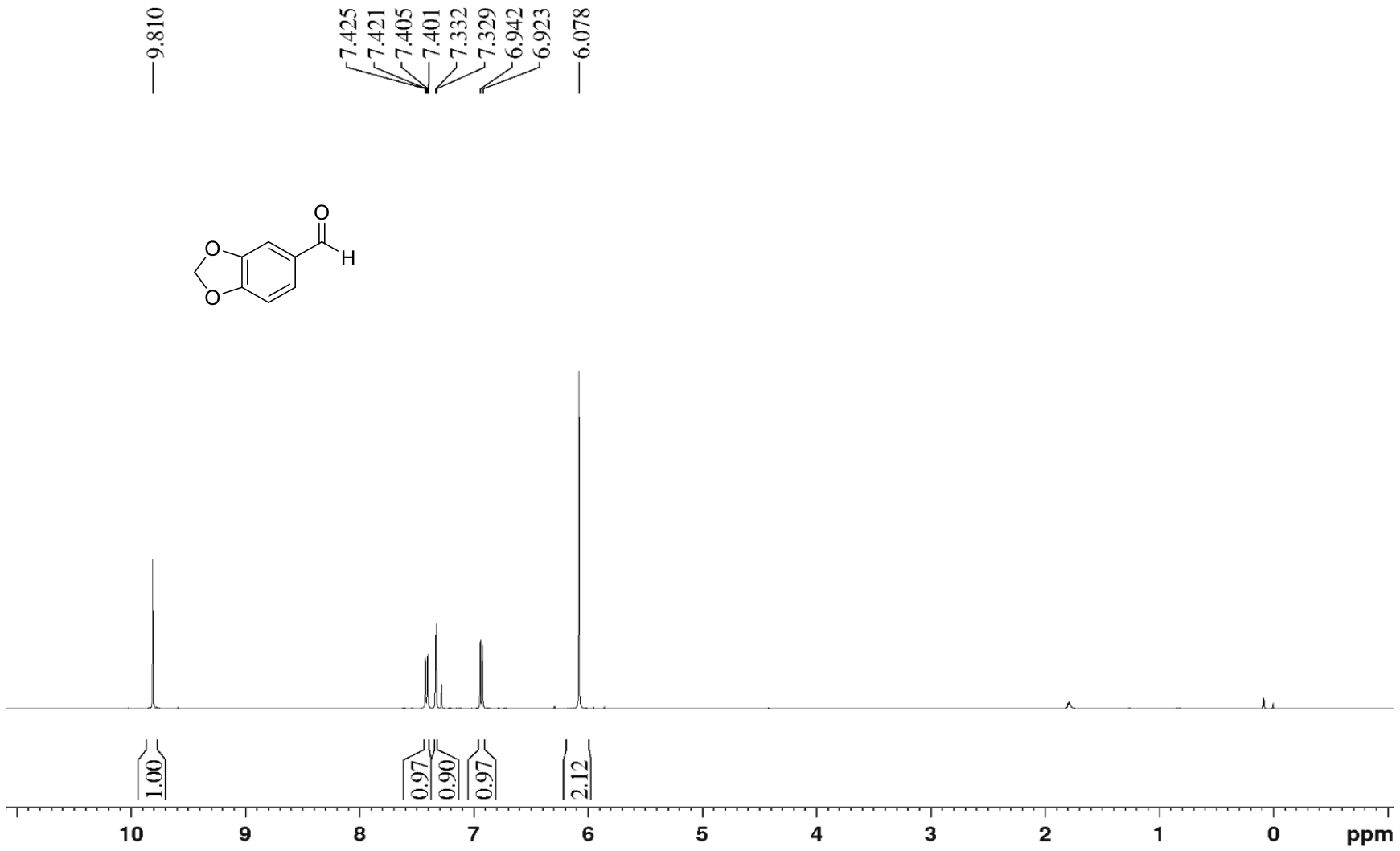
Percentage occupancies of H-bonds between N^2-SF-dG and dC in the post-insertion stage complex



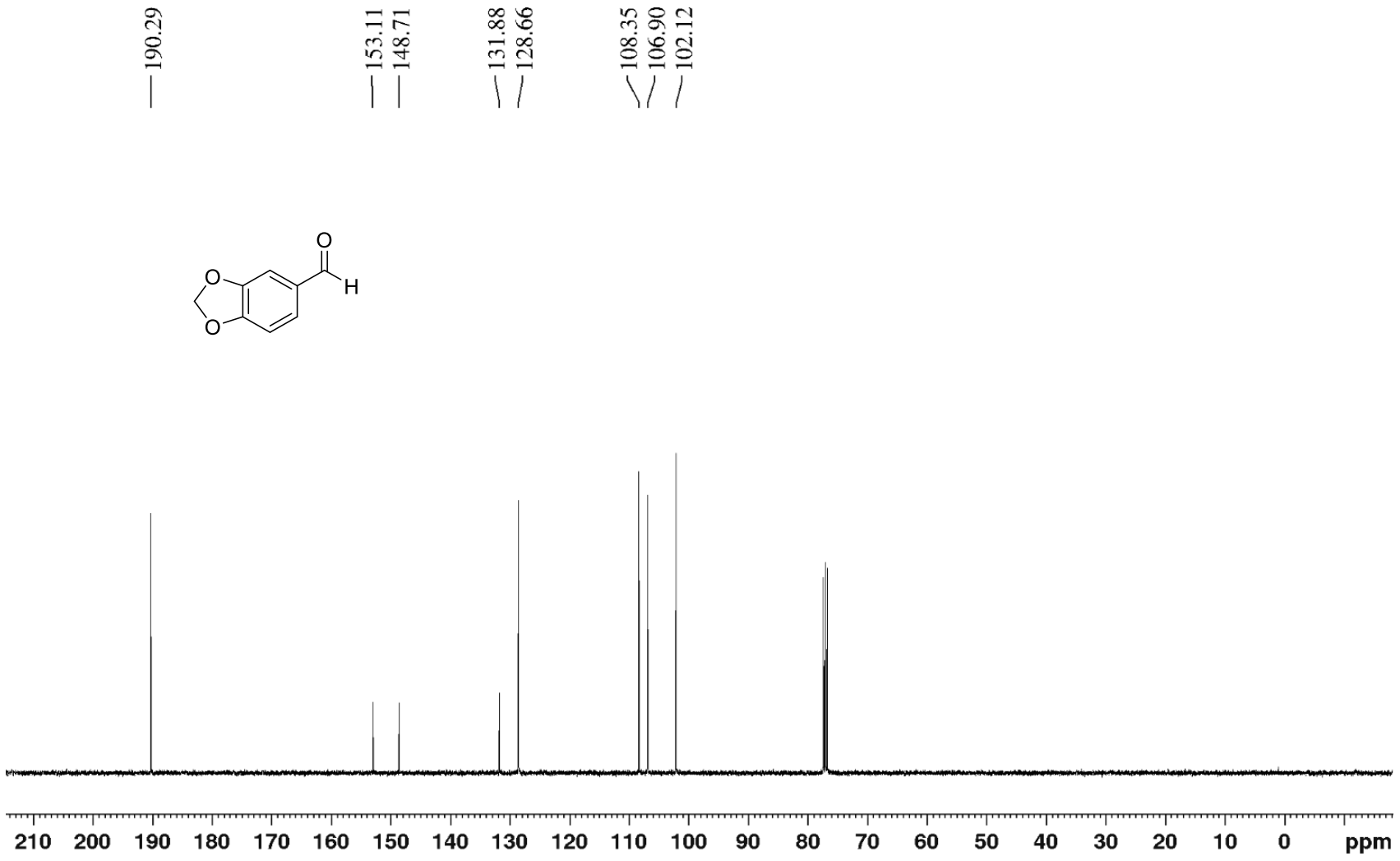
Bond	Donor Atom	Acceptor Atom	Percentage Occupancy	Average Distance	Average Angle
a	dG*@N2	dCTP@O2	99	2.9 Å	163°
b	dG*@N1	dCTP@N3	99	2.9 Å	162°
c	dCTP@N4	dG*@O6	66	3 Å	162°

Table S5. The figure representing the Watson-Crick hydrogen bonds between the N^2-SF-dG adduct (dG*) and dC. The hydrogen bonds are illustrated as dotted lines. The hydrogen bonds are illustrated as dotted lines. Hydrogen bonds are determined using distance and angle cut-off of 3.4 Å and 120° respectively.

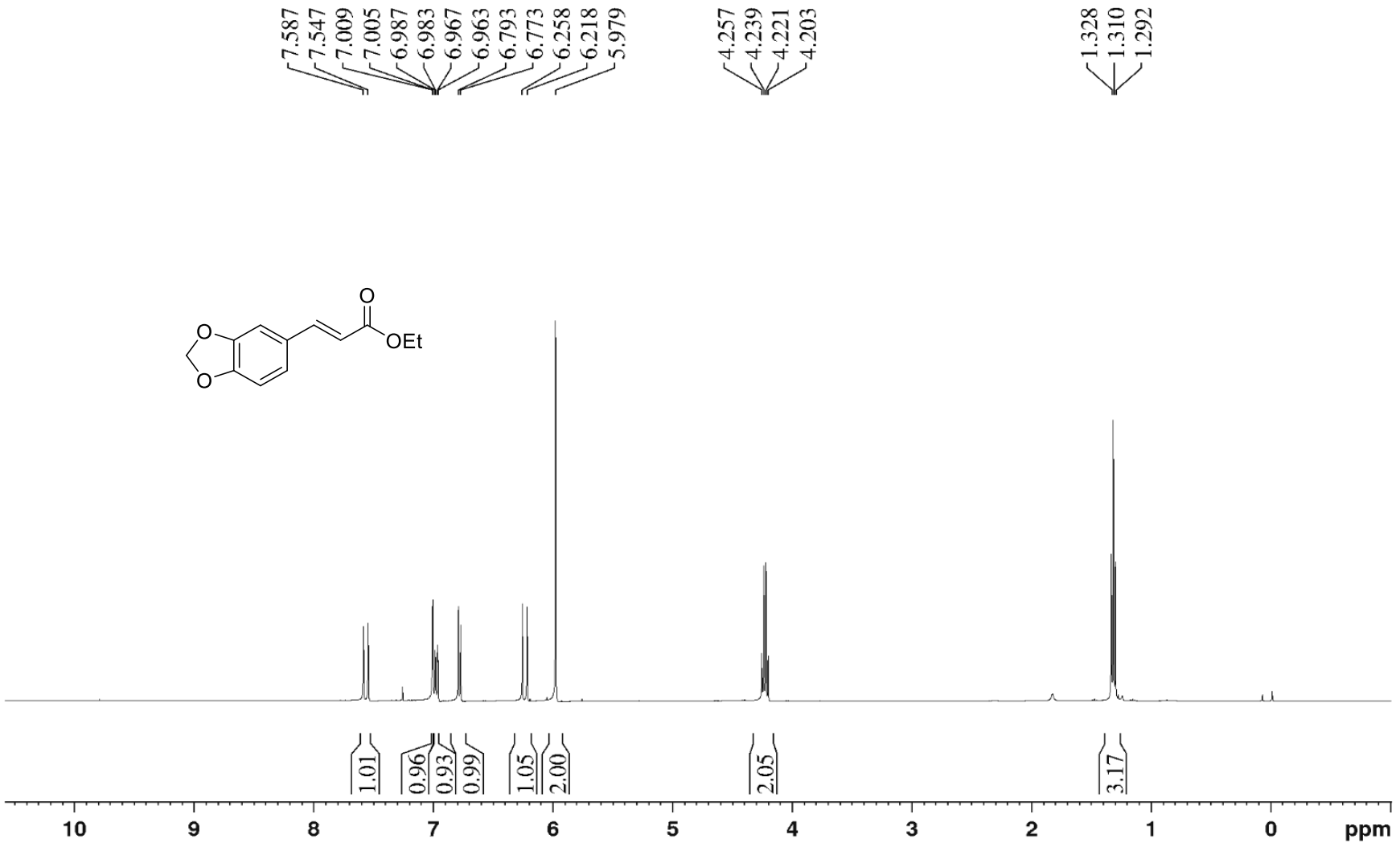
^{1}H NMR (400 MHz, $CDCl_3$) of 3,4-methylenedioxy benzaldehyde (2)



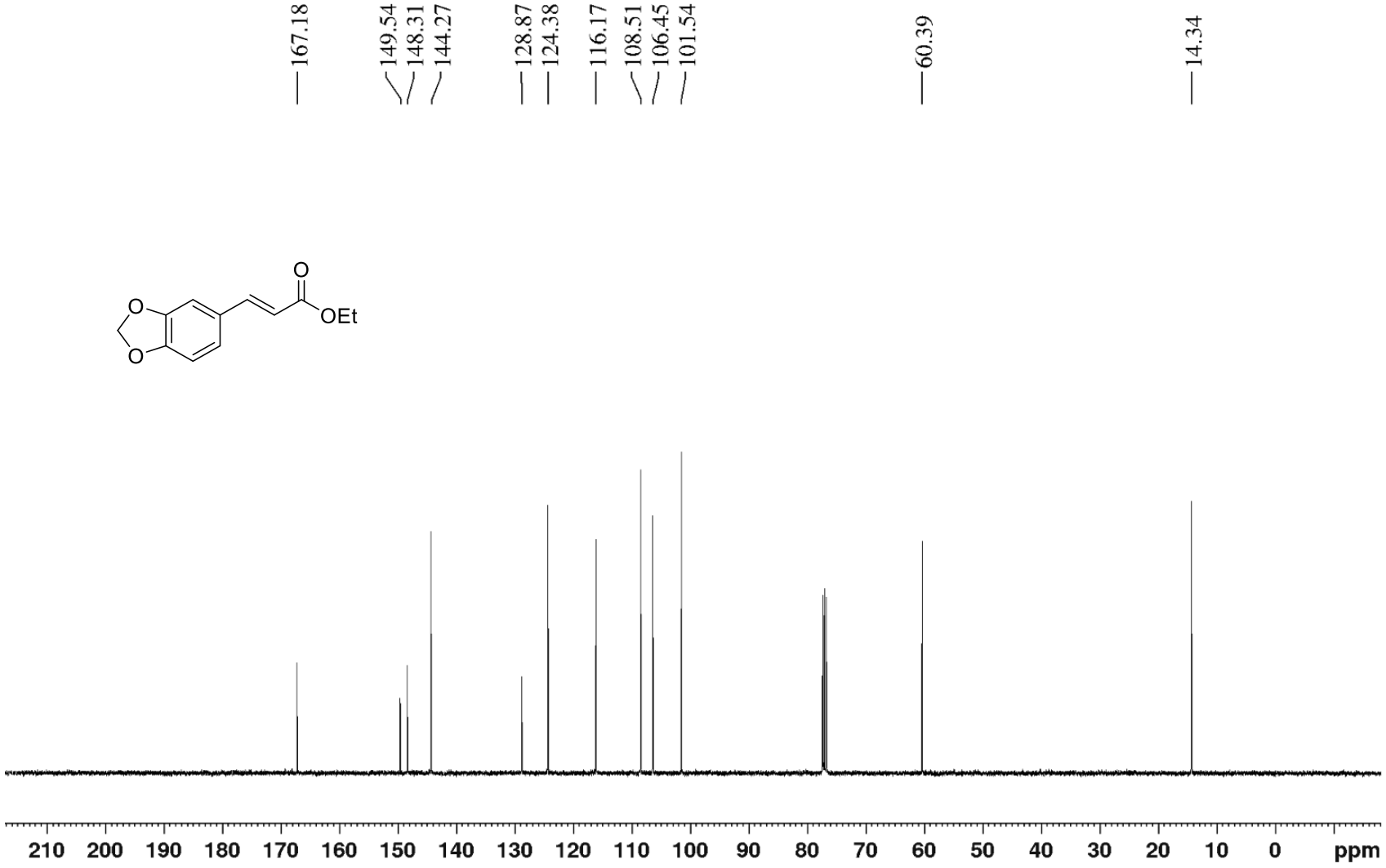
$^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$) of 3,4-methylenedioxy benzaldehyde (2)



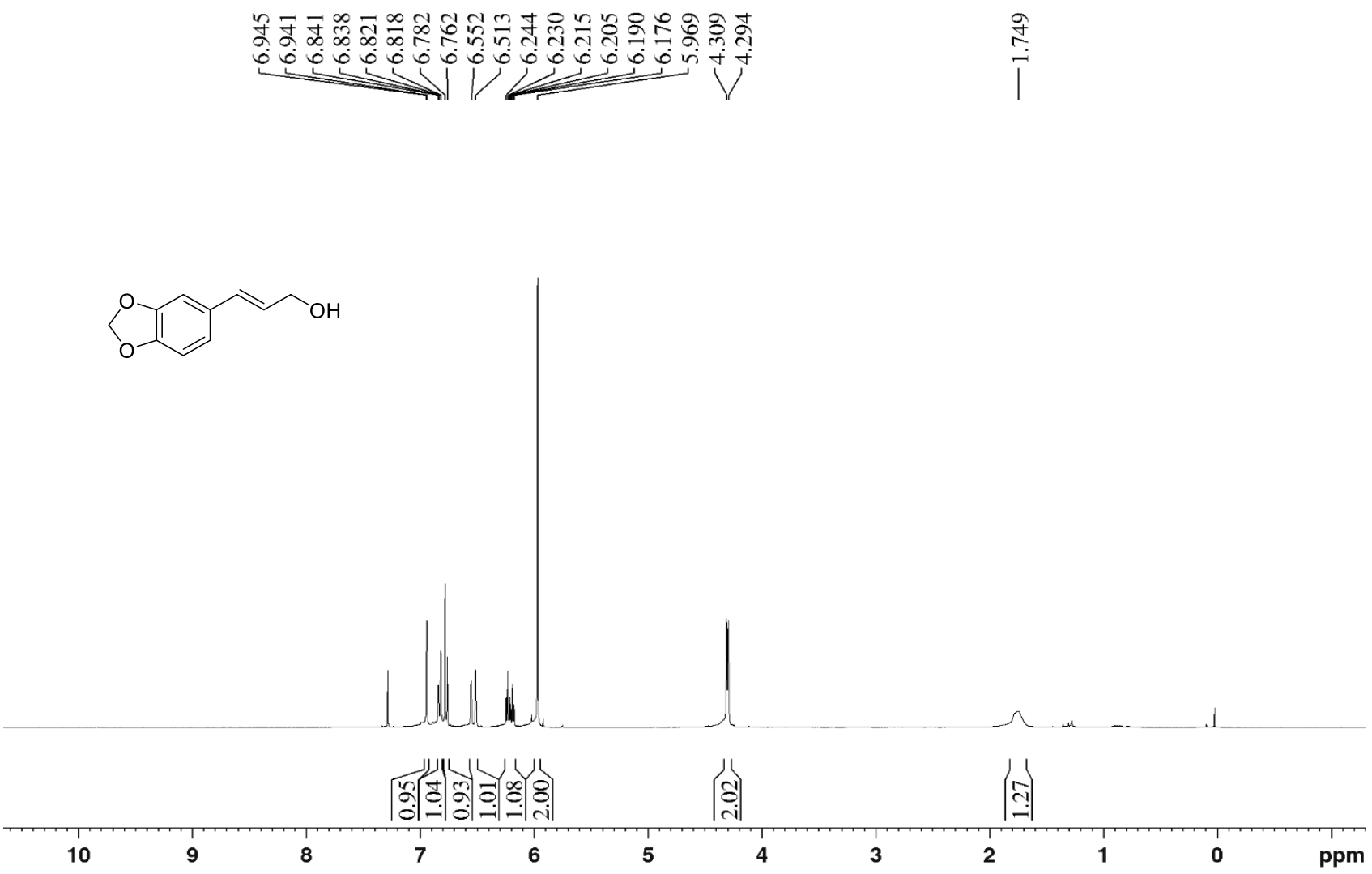
^{1}H NMR (400 MHz, $CDCl_3$) of ethyl 3,4-methylenedioxy-trans-cinnamate (3)



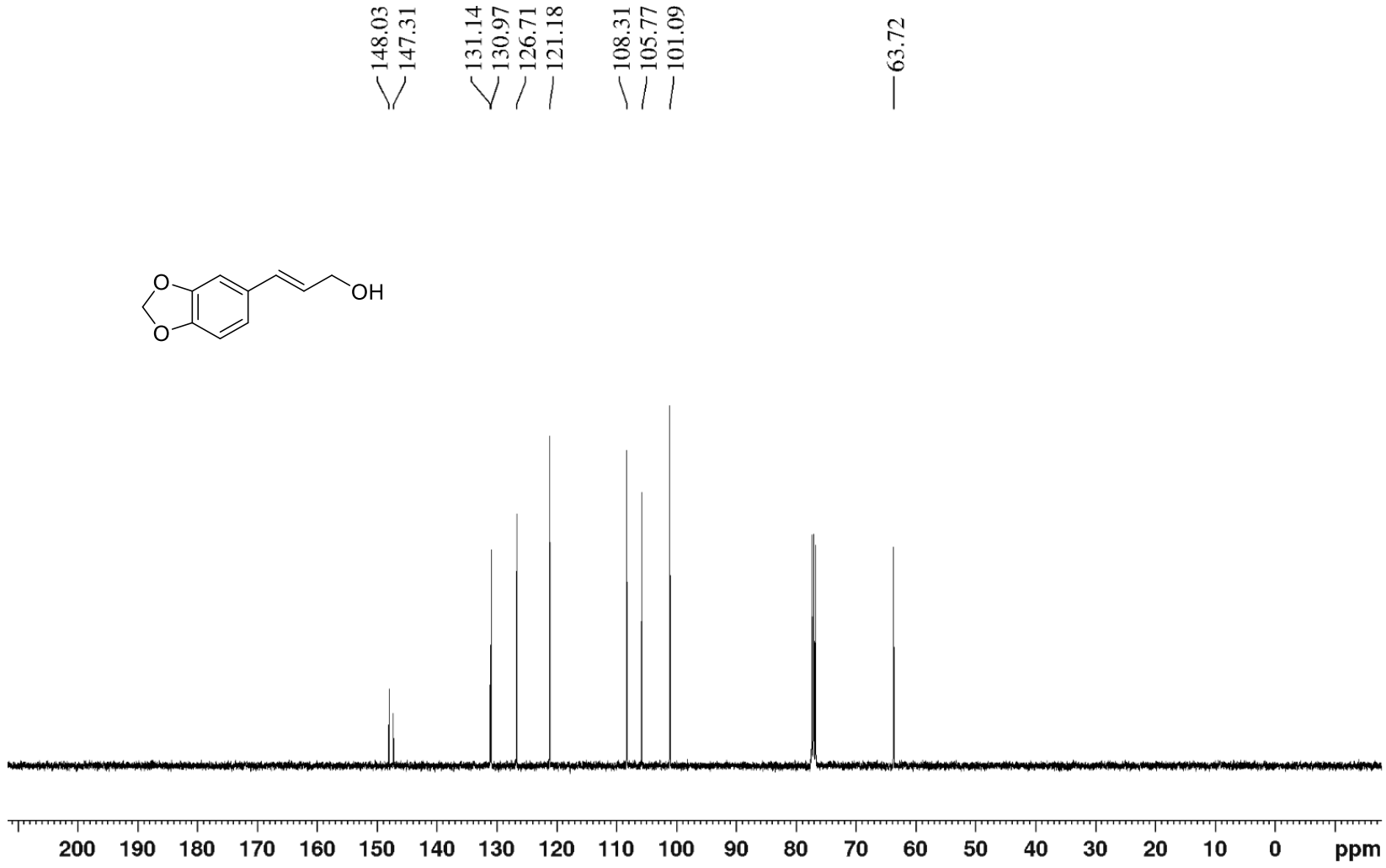
$^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$) of ethyl 3,4-methylenedioxy-trans-cinnamate (3)



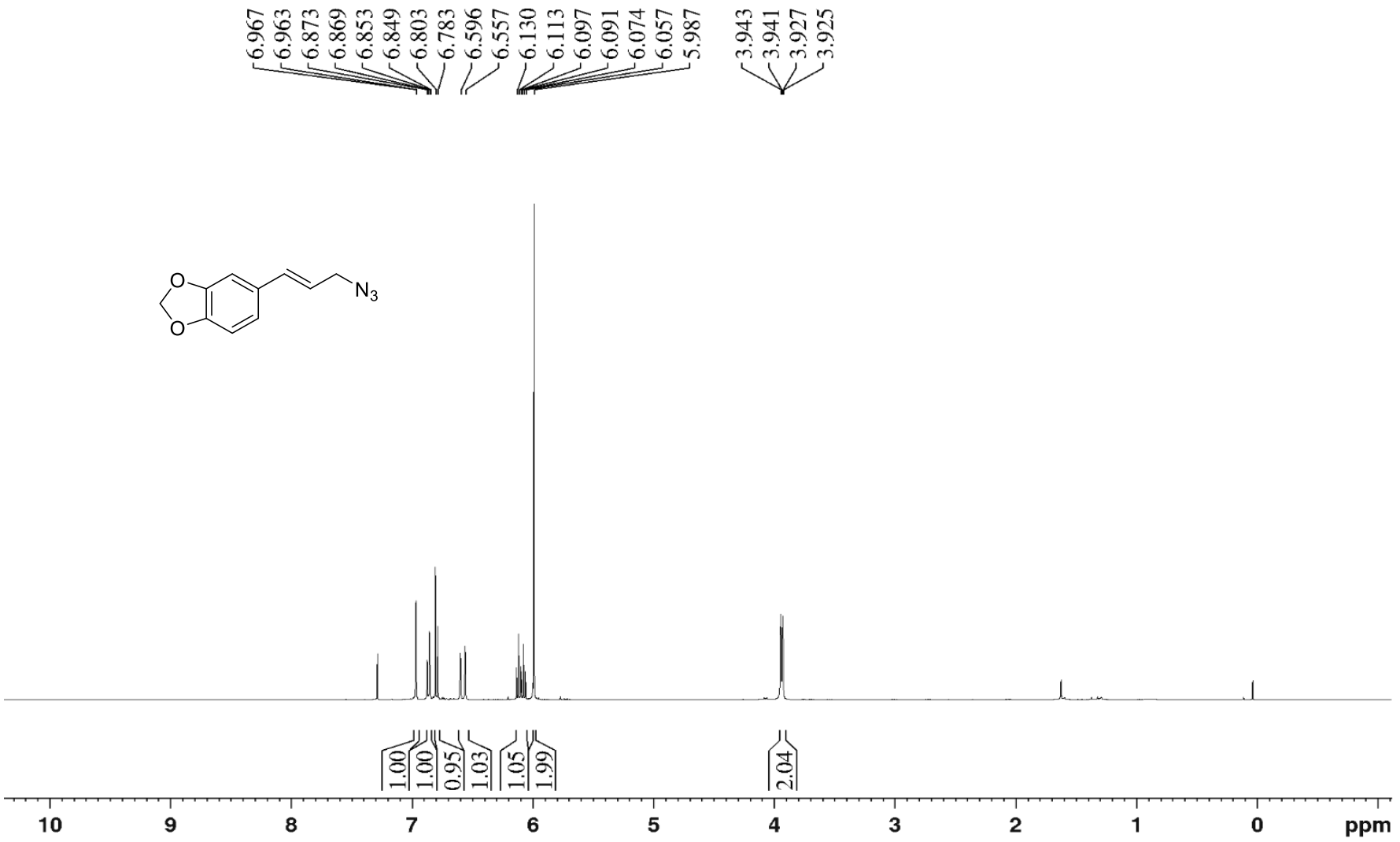
^{1}H NMR (400 MHz, $CDCl_3$) of 3,4-methylenedioxy-trans-cinnamyl alcohol (4)



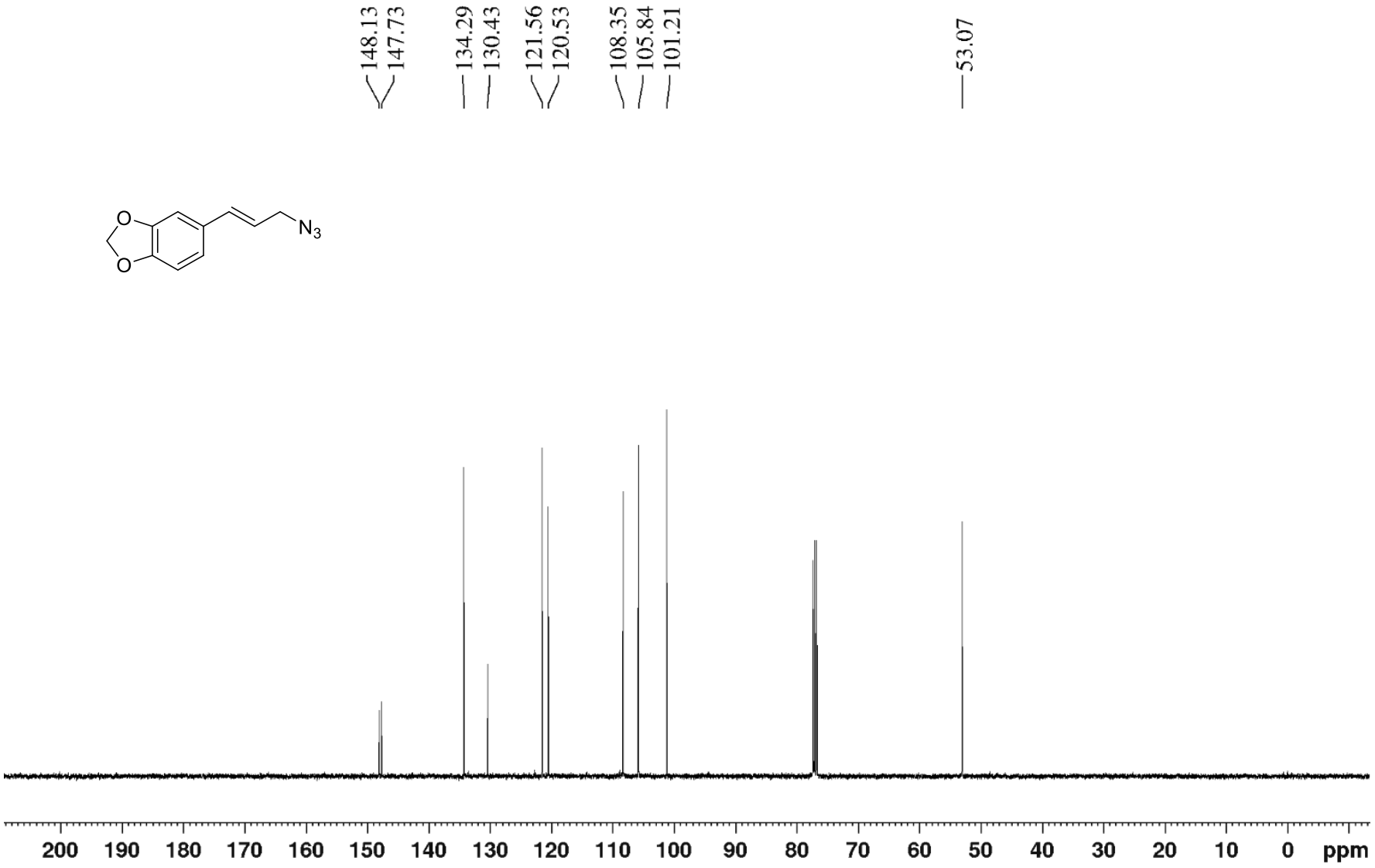
$^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$) of 3,4-methylenedioxy-trans-cinnamyl alcohol (4)



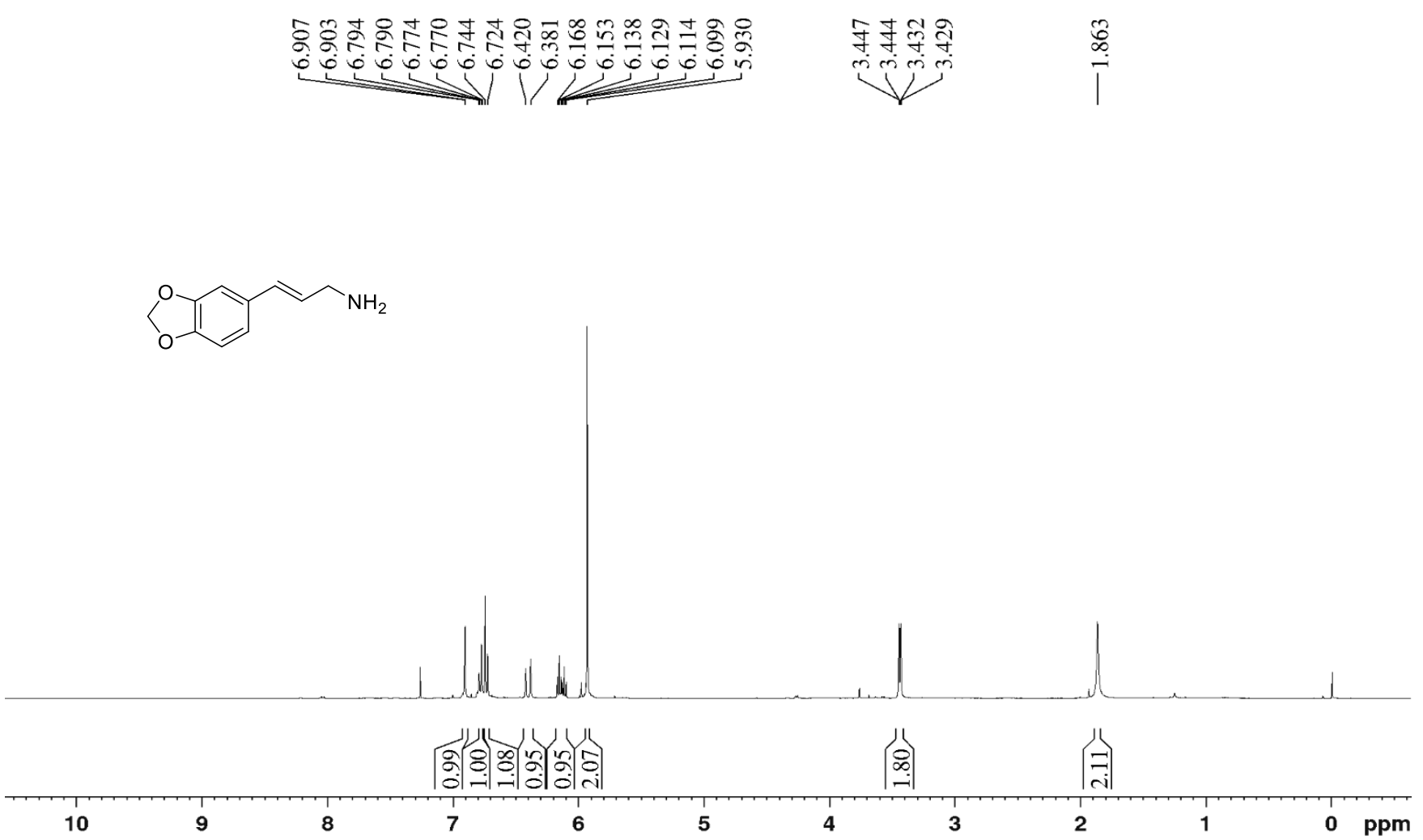
^{1}H NMR (400 MHz, $CDCl_3$) of 3,4-methylenedioxy-trans-cinnamyl azide (5)



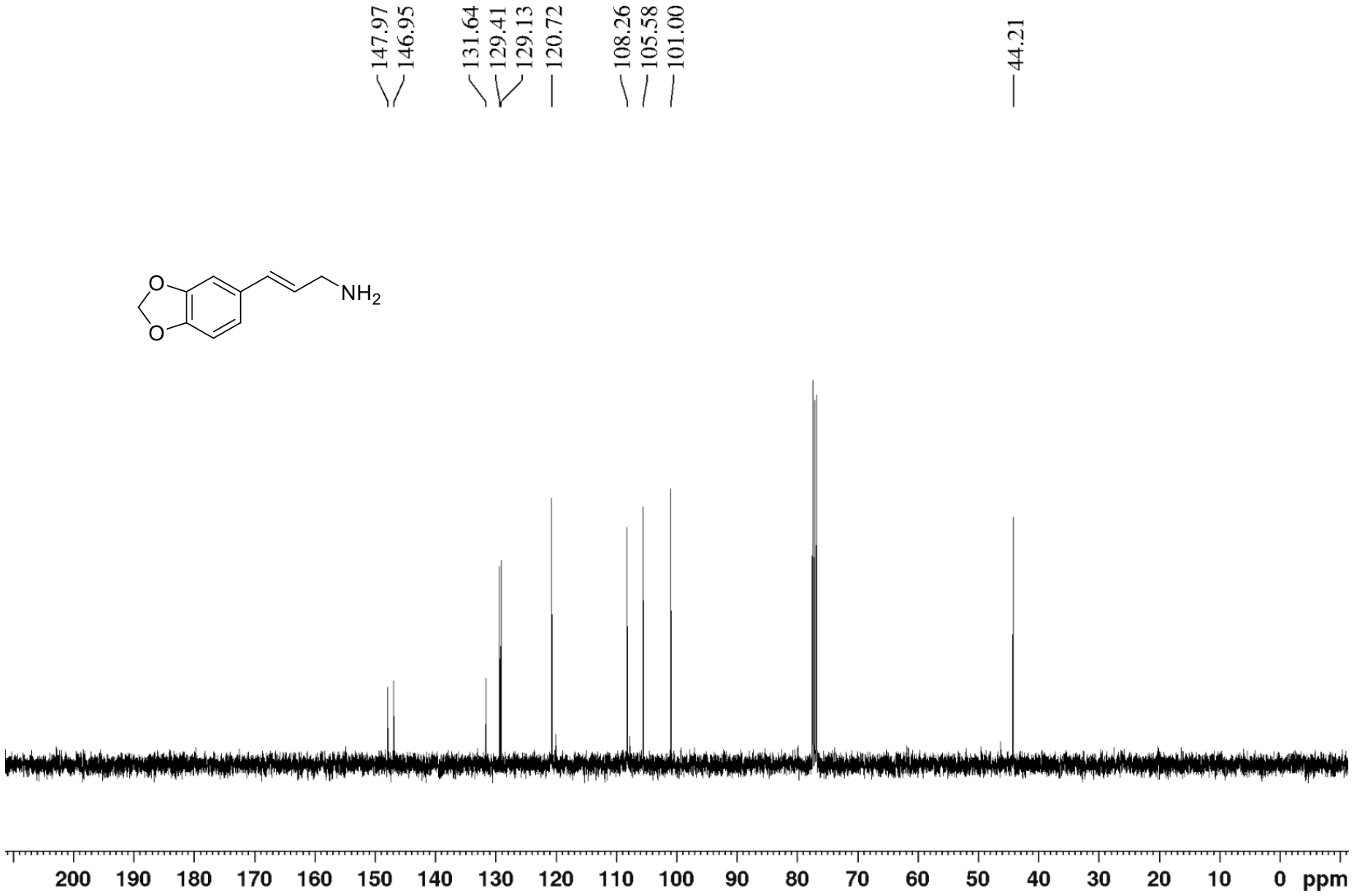
$^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$) of 3,4-methylenedioxy-trans-cinnamyl azide (5)



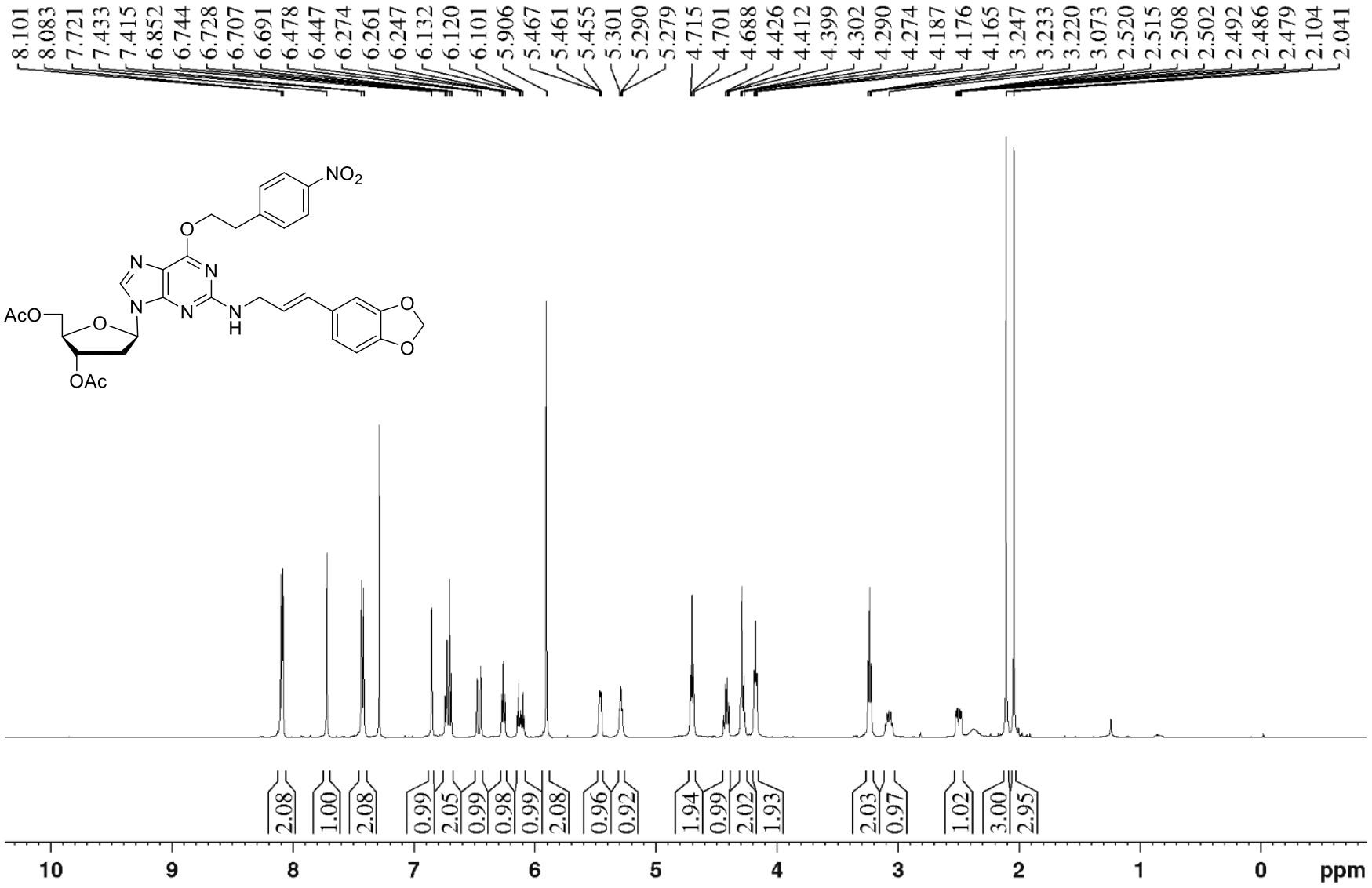
^{1}H NMR (400 MHz, $CDCl_3$) of 3,4-methylenedioxy-trans-cinnamyl amine (6)



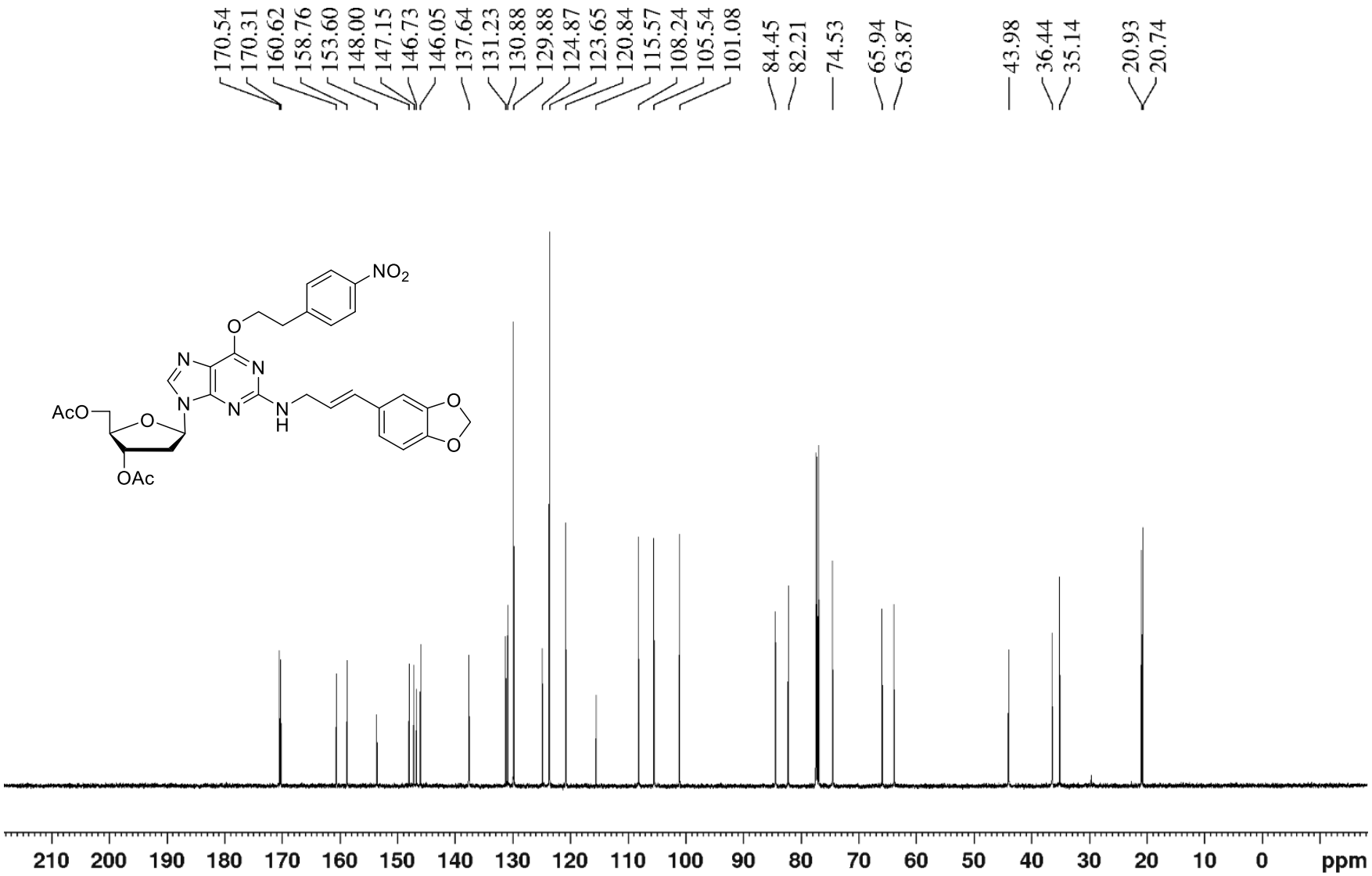
$^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$) of 3,4-methylenedioxy-trans-cinnamyl amine (6)



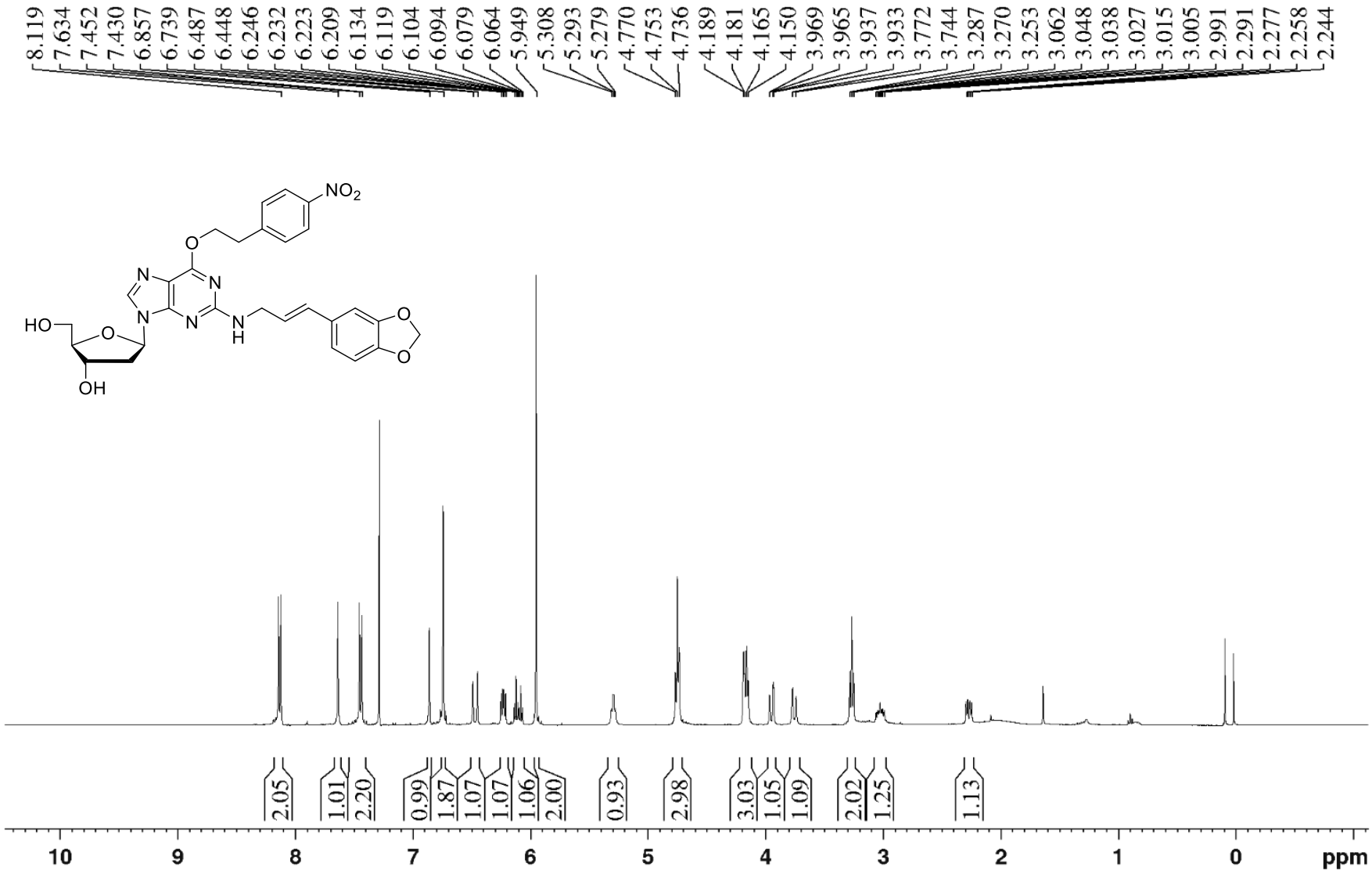
^{1}H NMR (400 MHz, $CDCl_3$) of N^2-methyl-(3,4-methylenedioxy) styrenyl)-O^6-(2-(4-nitrophenyl) ethyl)-3',5'-diacetyl-2′-deoxyguanosine (8)



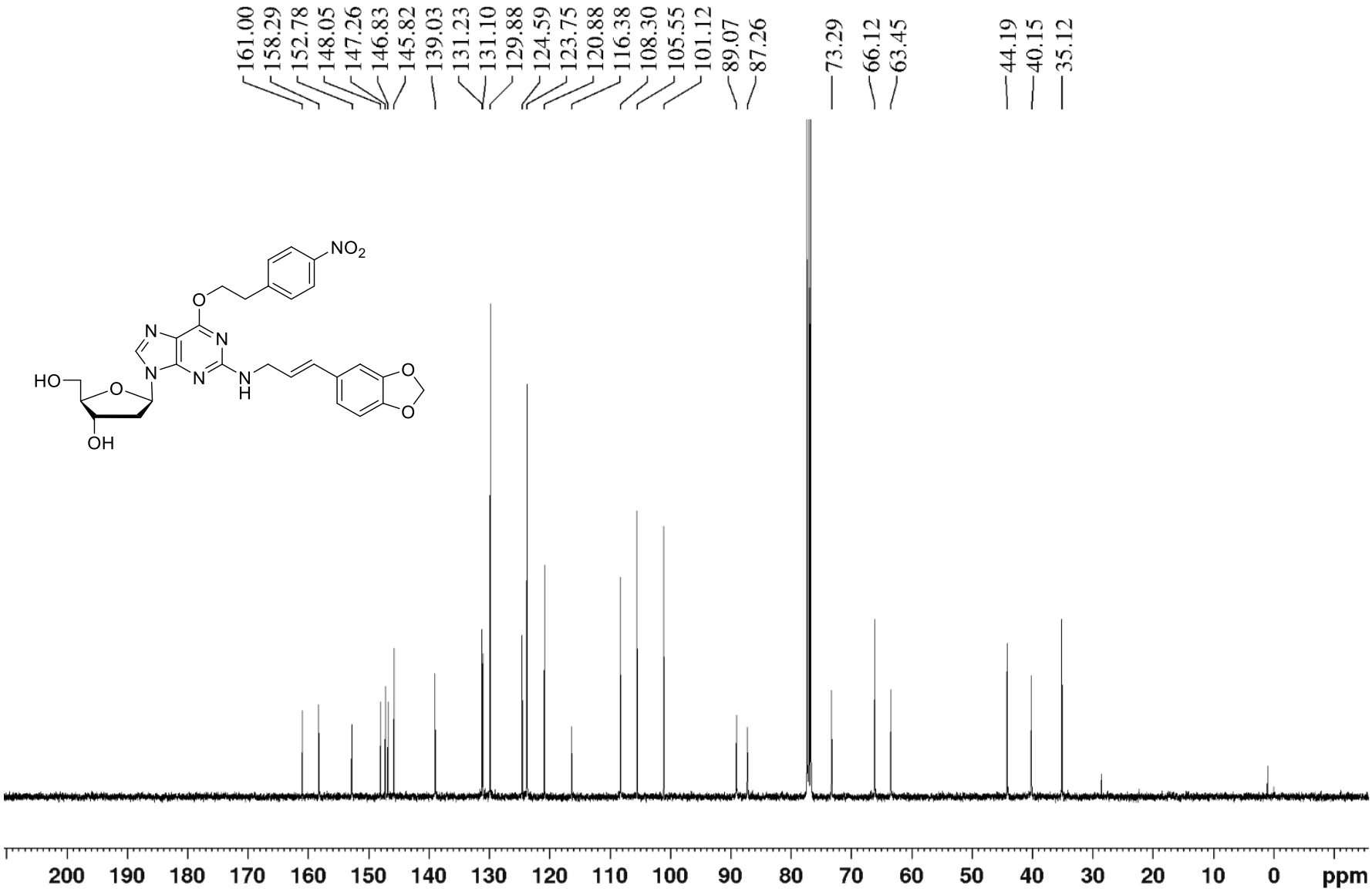
$^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$) of N^2-methyl-(3,4-methylenedioxy) styrenyl)-O^6-(2-(4-nitrophenyl) ethyl)-3',5'-diacetyl-2′-deoxyguanosine (8)



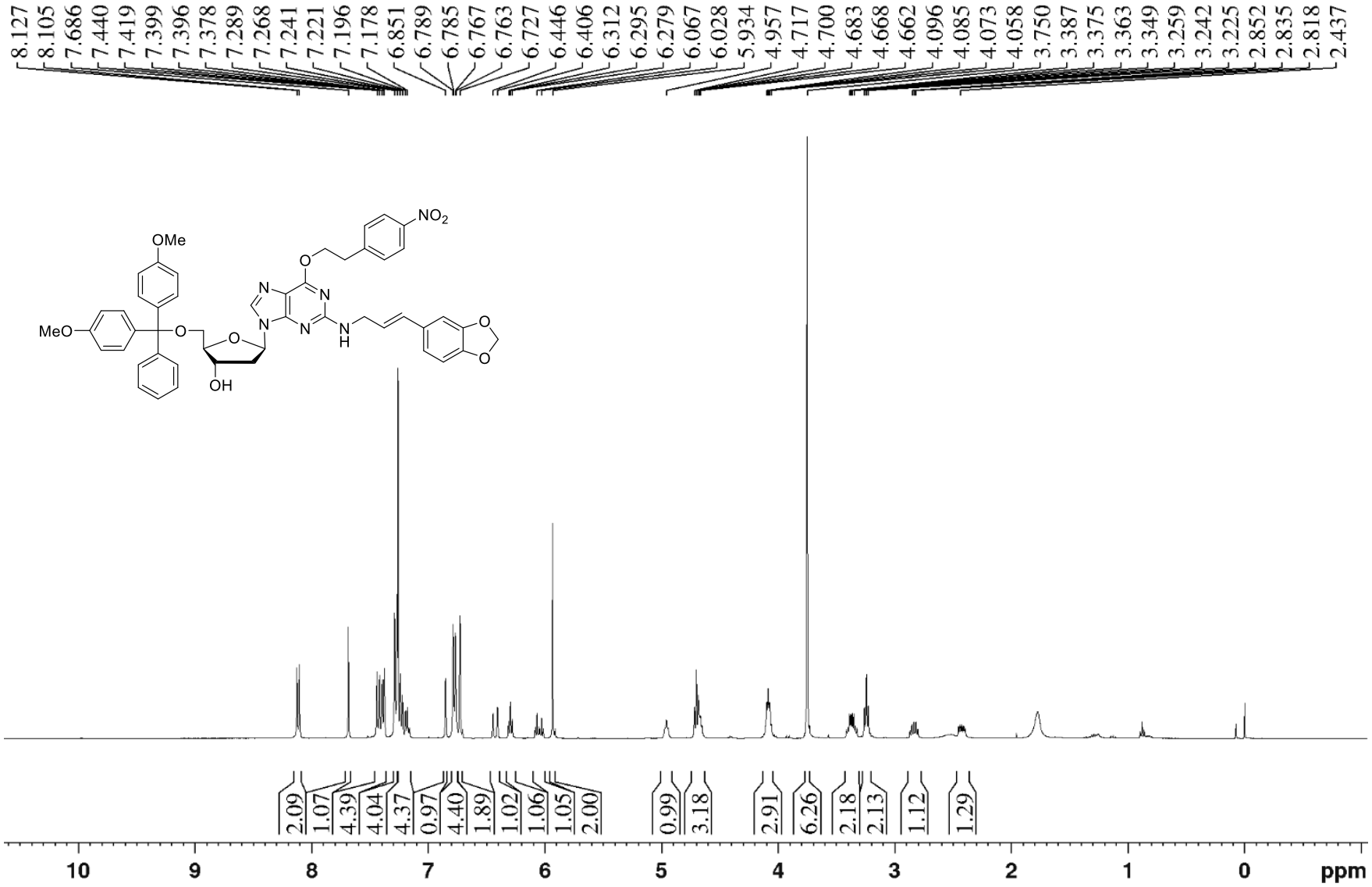
^{1}H NMR (400 MHz, $CDCl_3$) of N^2-CH_2-((3,4-methylenedioxy) styrenyl)-O^6-(2-(4-nitrophenyl) ethyl)-2′-deoxyguanosine (9)



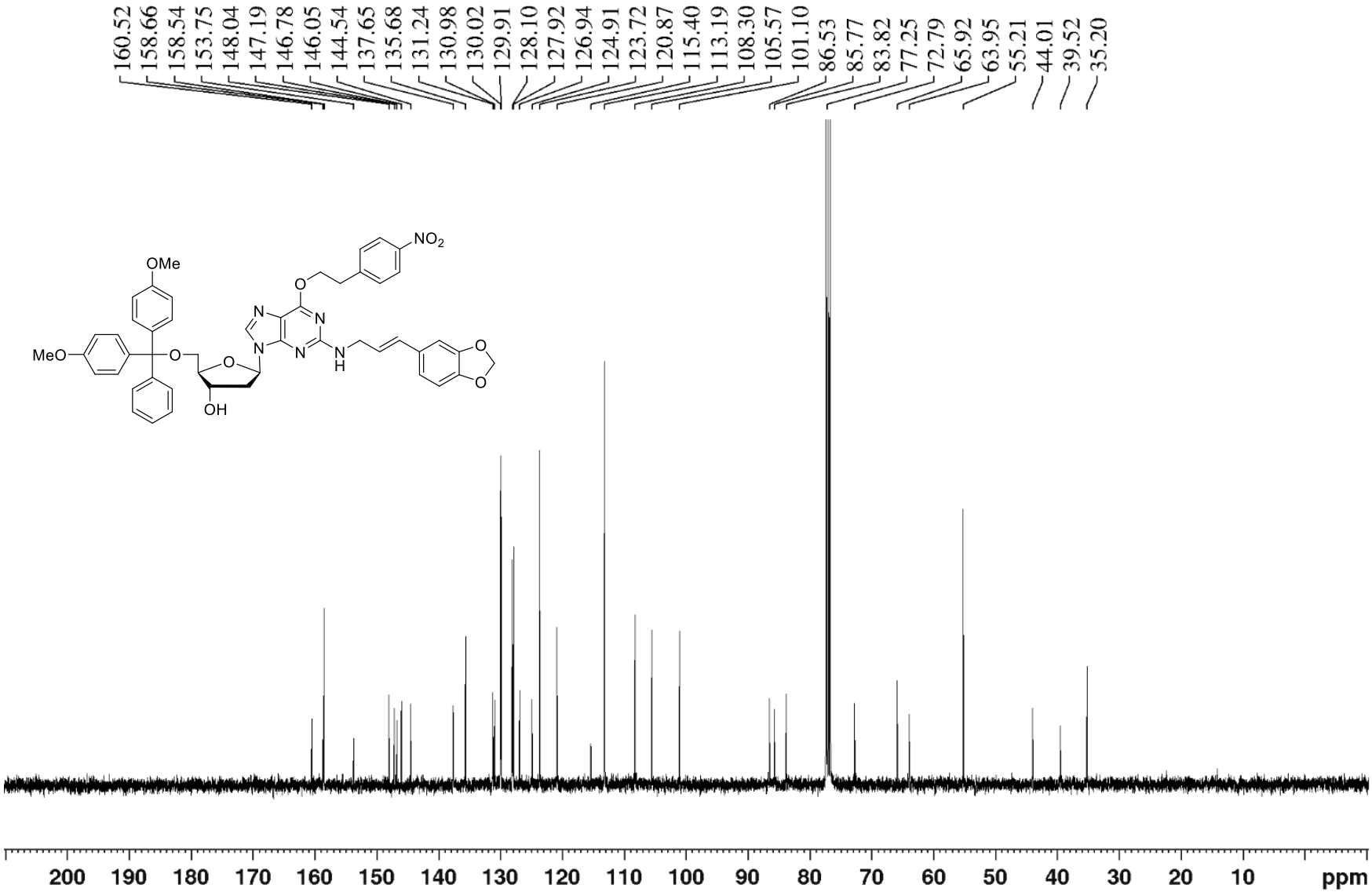
$^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$) of N^2-CH_2-((3,4-methylenedioxy) styrenyl)-O^6-(2-(4-nitrophenyl) ethyl)-2′-deoxyguanosine (9)



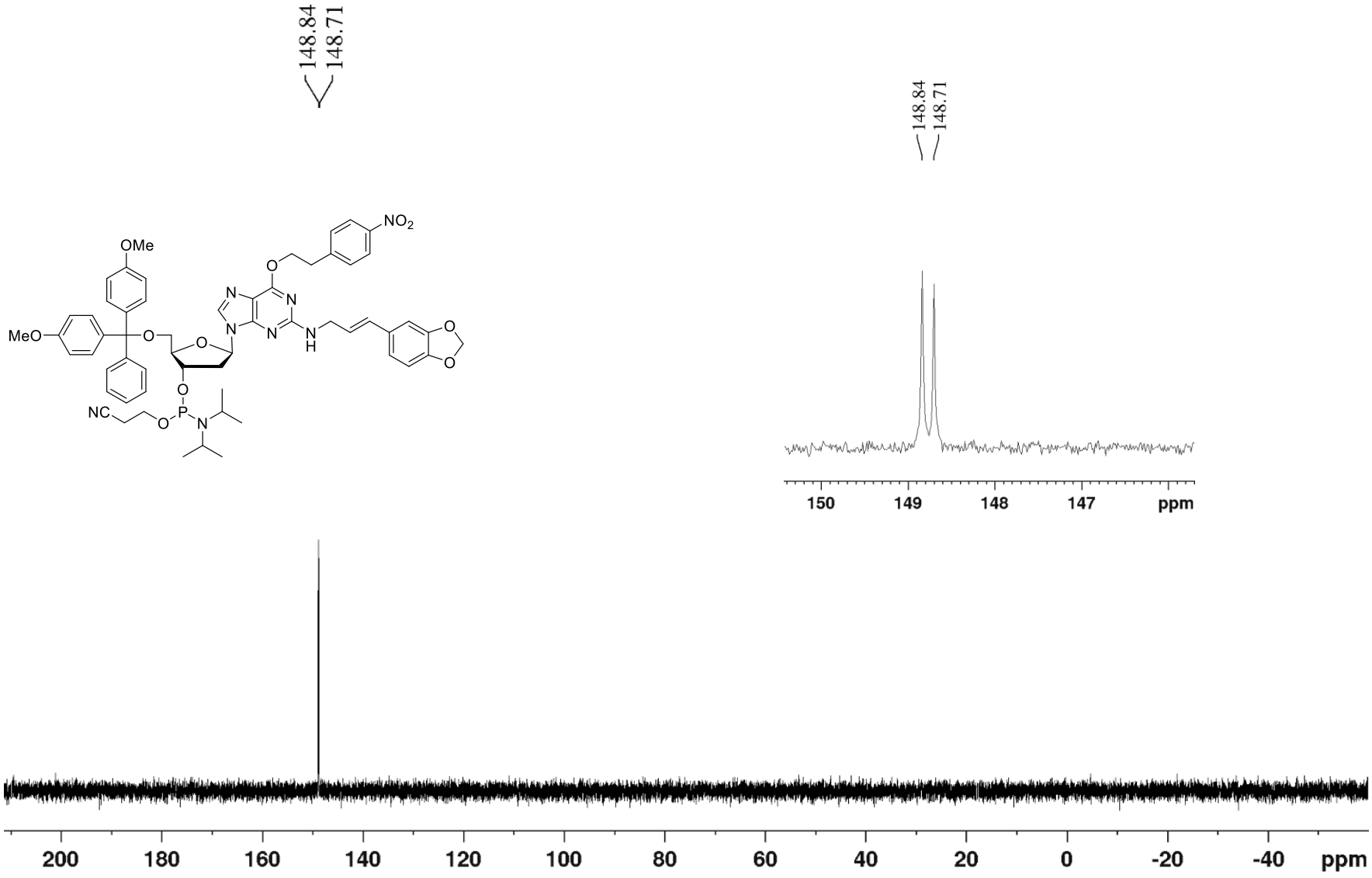
^{1}H NMR (400 MHz, $CDCl_3$) of N^2-methyl-(3,4- methylenedioxy) styrenyl)- O^6-(2-(4-nitrophenyl) ethyl)-5'-(4,4'-dimethoxytrityl)-2′-deoxyguanosine (10)



$^{13}C\{^{1}H\}$ NMR (100 MHz, $CDCl_3$) of of N^2-methyl-(3,4- methylenedioxy) styrenyl)- O^6-(2-(4-nitrophenyl) ethyl)-5'-(4,4'-dimethoxytrityl)-2′-deoxyguanosine (10)

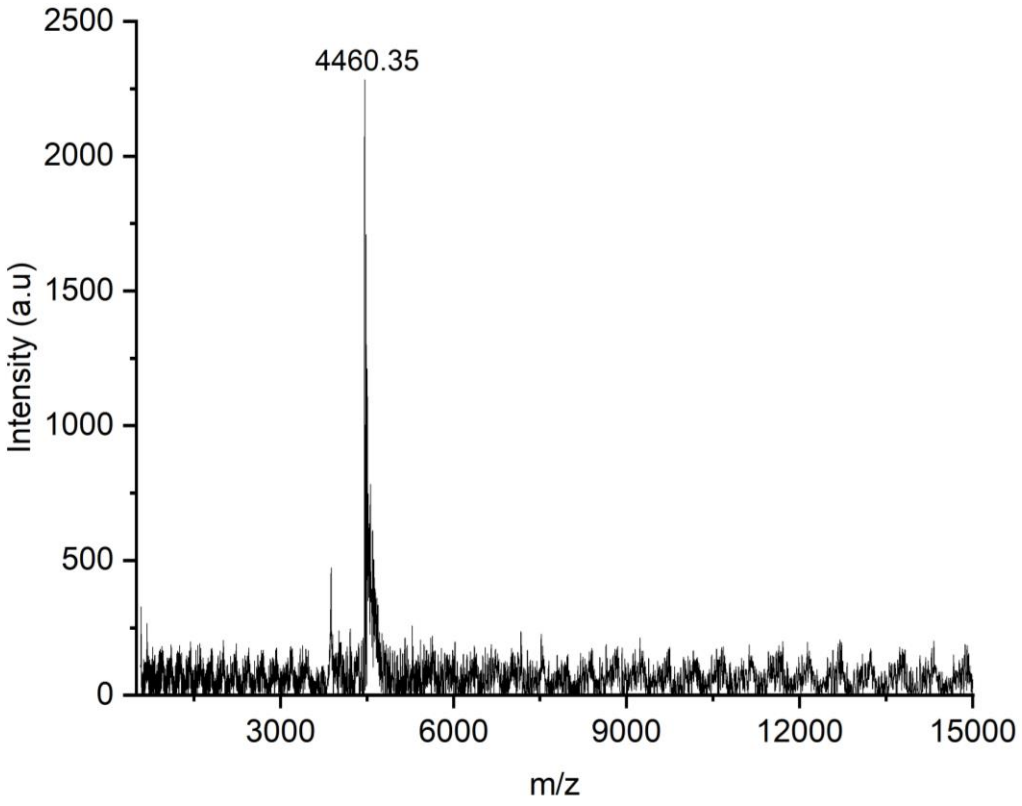


^{31}P NMR (162 MHz, $CDCl_3$) of N^2-methyl-(3,4- methylenedioxy) styrenyl)- O^6-(2-(4-nitrophenyl) ethyl)-5'-(4,4'-dimethoxytrityl)-2′-deoxyguanosine phosphoramidite (11)



MALDI Spectra of (G3)

$[M+H]^+$ Calcd for 5'-GCCGGAATAGCGCA-3' 4408.3 found 4460.3



MALDI Spectra of (G4)

$[M+H]^+$ Calcd for 5'-CTGGTCACACTGATGCCTACGAGTACG-3' 8420.0 found 8420.0