Biochemistry

pubs.acs.org/biochemistry Article

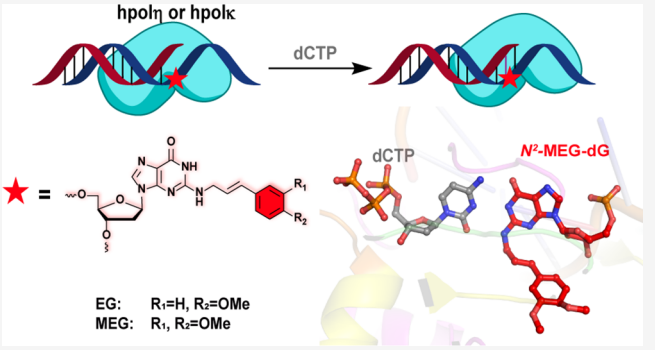
Human Translesion Synthesis Polymerases polκ and polη Perform Error-Free Replication across N^2-dG Methyleugenol and Estragole DNA Adducts

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

Cite This: https://doi.org/10.1021/acs.biochem.2c00663 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The secondary metabolites of polypropanoids, methyleugenol (MEG), and estragole (EG), found in many herbs and spices, are commonly used as food flavoring agents and as ingredients in cosmetics. MEG and EG have been reported to cause hepatocarcinogenicity in rodents, human livers, and lung cells. The formation of N^2-dG and N^6-dA DNA adducts is primarily attributed to the carcinogenicity of these compounds. Therefore, these compounds have been classified as "possible human carcinogens" by the International Agency for Research on Cancer and "reasonably anticipated to be a human carcinogen" by the National Toxicology Program. Herein, we report the synthesis of the N^2-MEG-dG and N^2-EG-dG modified oligonucleotides to study the mutagenicity of these DNA adducts. Our studies show that N^2-MEG-dG and N^2-EG-dG could be bypassed by human translesion synthesis (TLS) polymerases hpolκ and hpolη in an error-free manner. The steady-state kinetics of dCTP incorporation by hpolκ across N^2-MEG-dG and N^2-EG-dG adducts show that the catalytic efficiencies (k_{cat}/K_m) were ~2.5- and ~4.4-fold higher, respectively, compared to the unmodified dG template. A full-length primer extension assay demonstrates that hpolκ exhibits better catalytic efficiency than hpolη. Molecular modeling and dynamics studies capturing pre-insertion, insertion, and post-insertion steps reveal the structural features associated with the efficient bypass of the N^2-MEG-dG adduct by hpolκ and indicate the reorientation of the adduct in the active site allowing the successful insertion of the incoming nucleotide. Together, these results suggest that though hpolκ and hpolη perform error-free TLS across MEG and EG during DNA replication, the observed carcinogenicity of these adducts could be attributed to the involvement of other low fidelity polymerases.



INTRODUCTION

Biological systems have a strong ability to replicate their genomes with high efficiency and accuracy, but sometimes, it gets affected due to DNA damage.[1,2] DNA damage occurs due to continuous exposure to various endogenous and exogenous agents that affect the genetic integrity by introducing mutations or blocking the DNA replication process.[2,3] Most chemical carcinogens generate different kinds of reactive metabolites, which induce nucleobase modifications. Mostly, the N^2-position of dG and N^6-position of dA are prone to be attacked by several reactive electrophiles generated during metabolism. Among these, nucleobase adducts with N^2-dG modification are more prevalent.[4] Depending on the chemical attributes, N^2-dG adducts are mainly located on the minor groove of DNA and can disturb base pairing and the interaction with the replicative polymerases.[5,6] The unrepaired adducts can cause mutations responsible for cancer.[7,8] Though multiple repair mechanisms are employed to tackle these lesions, some can be persistent and lead to the blockage of the replication process. To overcome this, cells use specialized DNA polymerases from the Y family that bypass lesions at the stalled replication forks, and this process is called translesion synthesis (TLS).[9,10] TLS polymerases lack the proofreading exonuclease domain, a distinctive feature of replicative polymerases.[11] DNA damage tolerance in the human cell may be error-prone or error-free, and the fidelity of TLS bypass mainly depends on the type of DNA lesion and the polymerase used. Human TLS polymerases include hpolκ, hpolη, hpolι, Rev1, and hpolζ, with each having the unique potential to bypass different DNA lesions and exhibit varying levels of

Received: November 27, 2022
Revised: June 13, 2023
Published: July 24, 2023

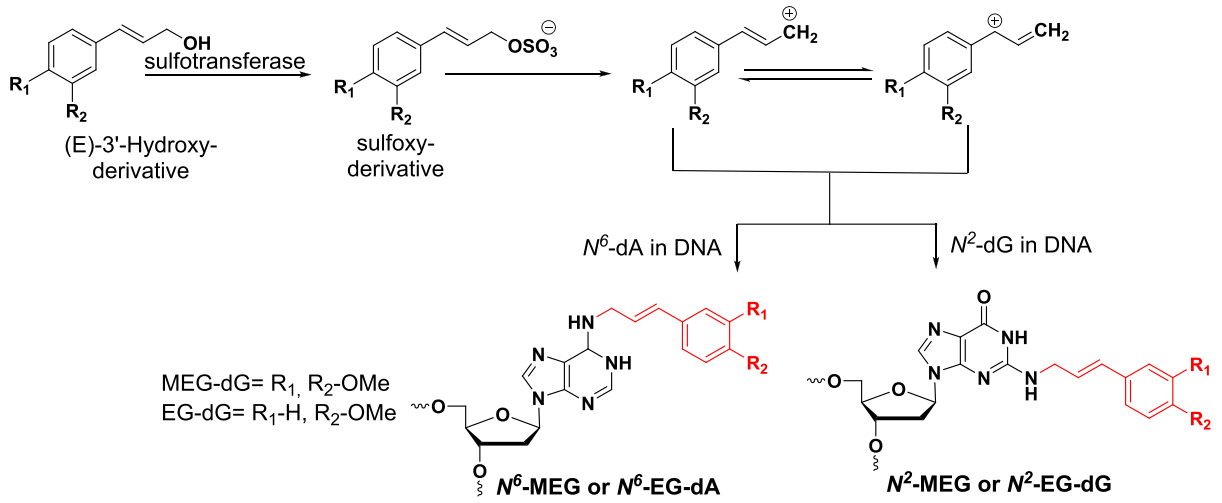


Figure 1. Proposed metabolic pathways for forming the N^2-MEG/EG-dG and N^6-MEG/EG-dA adducts.

fidelity.[12] Several reports are available on DNA adducts that human TLS polymerases can bypass in an error-free or error-prone manner *in vitro* and inside the cells.[13,14] For example, the DNA adducts formed at N^2-dG of benzo[*a*]pyrene-7,8-dihydrodiol-9,10-epoxide are bypassed in an error-free manner by hpolκ.[15,16] Similarly, the N^2-IQ-dG (2-amino-3-methylimidazo[4,5-*f*]quinoline) adduct is bypassed in an error-free manner by hpolκ but in an error-prone way by hpolη.[17]

Alkenylbenzenes are naturally occurring compounds present in about 450 plant species, including herbs such as tarragon, pimento, anis, clove, and basil.[18−20] Alkenylbenzenes like methyleugenol, estragole, and safrole are secondary metabolites of polypropanoids from essential oils produced in the plants and commonly used as flavoring agents in food and cosmetics.[18,20,21] Methyleugenol and estragole form N^2-dG and N^6-dA DNA adducts after sidechain hydroxylation followed by sulfation using the sulfotransferase enzyme (SULTs)[22] (Figure 1). The scientific community on food of the European Commission (EC) categorized both methyleugenol and estragole as genotoxic and carcinogenic.[21] Further, methyleugenol was classified as "possible human carcinogens" by the International Agency for Research on Cancer (IARC) and "reasonably anticipated to be a human carcinogen" by the National Toxicology Program.[22−24] The potent carcinogenic threats of methyleugenol are calculated to be about 25- and 100-fold greater than benzo[*a*]pyrene and PhIP, respectively, which are well-known carcinogens.[25] Due to the hepatocarcinogenicity observed in rodents, human lungs, and liver cells, they have been banned as food additives in certain countries.[26] *In vivo*, adduct formation with MEG and EG has been studied by measuring the adduct levels in the kidney and liver of mice.[27,28] The N^2-MEG-dG DNA adduct levels found in liver cells collected from human subjects (maximal level of 37 per 10^8 nucleosides) were sufficiently close to the levels observed for various heptocarcinogens in rodents.[27] Later, it was also detected in human lung samples,[29] and therefore, continued exposure to MEG imparts carcinogenic risk factors in humans.[27] Similarly, prolonged dietary exposure to EG can also elevate the levels of N^2-dG and N^6-dA adducts in humans.[30] Thus, it is important to study the mutagenicity of N^2-MEG-dG and N^2-EG-dG adducts using human polymerases. Herein, we report the synthesis of modified DNAs using N^2-MEG-dG and N^2-EG-dG phosphoramidite chemistry and their bypass by human TLS polymerases hpolκ and hpolη using primer extension assays and steady-state kinetics. Insights into the mechanism of TLS action of the N^2-MEG-dG adduct have been explored using molecular modeling and dynamics studies.

EXPERIMENTAL SECTION

General Experimental. All the required chemicals and solvents were obtained from commercial sources. DCM (dichloromethane), pyridine, acetonitrile, DIPEA (*N*,*N*-diisopropylethylamine), and toluene were dried using calcium chloride. Thin layer chromatography (TLC) was performed on silica gel plates pre-coated with fluorescent indicators and visualized under UV light (260 nm). Silica gel (100−200 mesh) and basic or neutral alumina (60−325 mesh) were used in column chromatography to purify the compounds. ^{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 (Bruker). Chemical shifts (δ) in parts per million were referenced to the residual signal of TMS (0 ppm) or residual proton signal from the 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 a Q-TOF analyzer (Bruker). The mass spectra of the oligonucleotides were obtained by positive reflection mode in a MALDI-TOF spectrometer (Bruker).

Method A: General Procedure for Wittig Reaction. Aldehyde (1 equiv) and PPh_3 (1.5 equiv) were dissolved in saturated aq. $NaHCO_3$ solution (30 mL) followed by addition of ethyl bromoacetate (3 equiv). The resulting reaction mixture was stirred at 85 °C for 12 h. After completion of the reaction, the reaction mixture was poured into saturated solution of $NaHCO_3$ (20 mL). The product was extracted with ethyl acetate (3 × 30 mL), the combined organic extract was dried over Na_2SO_4, and the solvent was removed under

https://doi.org/10.1021/acs.biochem.2c00663
Biochemistry XXXX, XXX, XXX−XXX

reduced pressure. The residue was purified by column chromatography on silica gel to obtain the olefin product.

Method B: General Procedure for Aldehyde Reduction. *trans*-Cinnamate (1 equiv) was dissolved in toluene (10 mL), and then, DIBAL-H (1.6−3.5 equiv) was added dropwise at 0 °C. Then, the reaction was removed from the ice bath, and the reaction mixture was stirred at room temperature for 12 h. After completion of the reaction, the reaction mixture was poured into H_2O (20 mL). The product was extracted with ethyl acetate (2 × 30 mL), dried over Na_2SO_4, and filtered. The solvent was evaporated under reduced pressure, and the crude compound was purified by silica gel column chromatography to afford alcohol.

Method C: General Procedure for Azidation. *trans*-Cinnamyl alcohol (1 equiv) was dissolved in CCl_4-DMF (13.6 mL, 1:4, v/v) followed by the addition of PPh_3 (2.1 equiv) and sodium azide (1.1 equiv). The combined reaction mixture was stirred at 90 °C for 12 h. After completion of the reaction, the reaction mixture was quenched with water and extracted using ethyl acetate (2 × 35 mL). Then, the organic layer was washed with water (2 × 25 mL), the combined organic layer was dried over Na_2SO_4, and the volatiles were evaporated under reduced pressure. Purification by silica gel column chromatography afforded azide.

Method D: General Procedure for Azide Reduction. *trans*-Cinnamyl azide (1 equiv) was dissolved in THF (3.9 mL) followed by addition of PPh_3 (1.2−1.5 equiv) at 0 °C. The resulting reaction mixture was removed from the ice bath, and water (0.7 mL) was added; further, the reaction mixture was allowed to stir at room temperature for 4 h. After completion of the reaction, the solvent was evaporated on reduced pressure and purified by column chromatography using neutral alumina to obtain amine.

Method E: General Procedure for Bucwald−Hartwig Coupling. In an oven-dried round bottom flask, *rac*-BINAP (0.2 equiv) and $Pd(OAc)_2$ (0.1 equiv) were dissolved in dry toluene (5.8 mL) and allowed to stir at room temperature for 5 min followed by the addition of Cs_2CO_3 (1.4 equiv), *trans*-cinnamyl amine (1 equiv), and bromo nucleoside **6** (1 equiv). The reaction mixture was heated in an oil bath at 85 °C for 10−12 h. After the completion of the reaction, the reaction mixture was passed through a celite pad and washed with ethyl acetate (50 mL). The filtrate was evaporated under reduced pressure and purified by column chromatography to obtain coupled nucleoside.

Method F: General Procedure for Deacylation. Coupled nucleoside (1 equiv) was dissolved in 1,4-dioxane (10.3 mL, 14 mL/mmol). Then, 33% aq. $MeNH_2$ (5.1 mL, 7 mL/mmol) was added and stirred the reaction mixture at room temperature for 2−4 h. After the completion of the reaction, the reaction mixture was concentrated under reduced pressure. Then, the resultant residue was purified by silica gel column chromatography to obtain deacylated nucleoside.

Method G: General Procedure for 5′-O-DMT Protection. Deacylated nucleoside (1 equiv) was added in a microwave tube and dissolved in dry pyridine (1.8 mL, 10 mL/mmol). To this, DMT-Cl (1.1 equiv) and DMAP (0.1 equiv) were added. The reaction mixture was stirred under microwave conditions (150 W) at 85 °C for 30 min. The crude mixture was quenched with water, then extracted with DCM (3 × 15 mL), and washed with saturated $NaHCO_3$ (10 mL). The combined organic extract was dried over Na_2SO_4, and the volatiles were evaporated under reduced pressure. The remaining residue was purified by column chromatography on silica gel to furnish 5′-*O*-DMT protected nucleoside.

Method H: General Procedure for 3′-O-Phosphytilation. DMT protected nucleoside (1 equiv) in DCM (2.3 mL), DIPEA (5 equiv), and CEP-Cl (2 equiv) was added. The reaction mixture was stirred at room temperature for 40 min to 2 h. After completion of the reaction, MeOH (0.5 mL) was added and stirred for 30 min. The reaction mixture was diluted with DCM (30 mL) and washed with $NaHCO_3$ (3 × 10 mL). Then, the organic layer was dried over Na_2SO_4, and the solvent was evaporated under reduced pressure to give a residue, which was purified by silica gel column chromatography to furnish phosphoramidite.

Preparation of 3,4-Dimethoxy-*trans*-cinnamate (2a). Method A was followed using 3,4-dimethoxy benzaldehyde (1 g, 6.0 mmol) and PPh_3 (2.3 g, 9.0 mmol) in saturated aq. $NaHCO_3$ solution (30 mL) followed by addition of ethyl bromoacetate (2.9 mL, 18.0 mmol) to obtain the colorless liquid **2a**.[31] Yield: 88% (1.25 g); R_f = 0.4 (30% ethyl acetate in petroleum ether); ^{1}H NMR: (400 MHz, $CDCl_3$) δ: 7.59 (d, J = 16 Hz, 1H), 7.07 (dd, J = 8.4, 2.0 Hz, 1H), 7.01 (d, J = 2 Hz, 1H), 6.82 (d, J = 8.0 Hz, 1H), 6.24 (d, J = 15.6 Hz, 1H), 4.24 (q, J = 7.2 Hz, 2H), 3.87 (s, 6H), 1.31 (t, J = 7.2 Hz, 3H); ^{13}C NMR δ: (100 MHz, $CDCl_3$): 167.2, 151.1, 149.2, 144.5, 127.4, 122.6, 116.0, 111.0, 109.5, 60.4, 56.0, 55.9, 14.4; HRMS (ESI): Calcd for $C_{13}H_{16}O_4$, $[M + Na]^+$ 259.0941; found, $[M + Na]^+$ 259.0940 (Δm = +0.0001 and error = +0.4 ppm).

Preparation of 3,4-Dimethoxy-*trans*-cinnamyl Alcohol (3a). Method B was followed using compound **2a** (500 mg, 2.1 mmol) in toluene (10 mL) and DIBAL-H (1.3 mL, 3.5 mmol, 1.0 M in THF) to afford yellowish liquid **3a**.[32] Yield: 76% (310 mg); R_f = 0.3 (30% ethyl acetate in petroleum ether); ^{1}H NMR: (400 MHz, $CDCl_3$) δ: 6.89−6.85 (m, 2H), 6.75 (d, J = 8 Hz, 1H), 6.48 (d, J = 15.6 Hz, 1H), 6.18 (dt, J = 11.6, 5.8 Hz, 1H), 4.24 (d, J = 5.8 Hz, 2H), 3.83 (s, 3H), 3.82 (s, 3H); ^{13}C NMR: (100 MHz, $CDCl_3$) δ: 148.9, 148.8, 130.8, 129.8, 126.6, 119.6, 111.1, 108.9, 63.6, 55.9, 55.8; HRMS (ESI): Calcd for $C_{11}H_{14}O_3$, $[M + H]^+$ 195.0942; found, $[M + H]^+$ 195.0940 (Δm = +0.0002 and error = +0.2 ppm).

Synthesis of 3,4-Dimethoxy-*trans*-cinnamyl Azide (4a). Method C was followed using compound **3a** (530 mg, 2.7 mmol), sodium azide (212 mg, 3.2 mmol), and PPh_3 (1.4 g, 5.7 mmol) in CCl_4-DMF (13.6 mL, 1:4, v/v) to afford the light greenish liquid **4a**. Yield: 71% (420 mg); R_f = 0.4 (5% ethyl acetate in petroleum ether); ^{1}H NMR: (400 MHz, $CDCl_3$) δ: 6.94−6.92 (m, 2H), 6.82 (d, J = 8.8 Hz, 1H), 6.57 (d, J = 15.6 Hz, 1H), 6.13 (dt, J = 15.6, 6.8 Hz, 1H), 3.92 (dd, J = 7.8, 1.0 Hz, 2H), 3.90 (s, 3H), 3.88 (s, 3H); ^{13}C NMR: (100 MHz, $CDCl_3$) δ: 149.4, 149.2, 134.5, 129.1, 120.4, 120.1, 111.2, 109.0, 56.0, 55.9, 53.3; HRMS (ESI): Calcd for $C_{11}H_{13}O_2N_3$, $[M + Na]^+$ 242.0900; found, $[M + Na]^+$ 242.0899 (Δm = +0.0001 and error = +0.6 ppm).

Synthesis of 3,4-Dimethoxy-*trans*-cinnamyl Amine (5a). Method D was followed using compound **4a** (350 mg, 1.59 mmol) and PPh_3 (623.5 mg, 2.38 mmol) in THF (3.9 mL) and water (0.7 mL) to obtain light yellow liquid **5a**. Yield: 74% (228 mg); R_f = 0.3 (10% MeOH in DCM); ^{1}H NMR: (400 MHz, $CDCl_3$) δ: 6.90 (d, J = 1.6 Hz, 1H), 6.86 (dd, J = 8.2, 1.7 Hz, 1H), 6.74 (d, J = 8.2 Hz, 1H), 6.44 (d, J = 15.8 Hz, 1H), 6.15 (dt, J = 15.8, 6.3 Hz, 1H), 3.85 (s, 3H), 3.82 (s, 3H), 3.50 (d, J = 6.2 Hz, 2H); ^{13}C NMR: (100 MHz, $CDCl_3$) δ: 148.6, 148.2, 129.9, 128.9, 128.8, 118.9, 110.8, 108.3, 55.4, 55.3, 43.8; HRMS (ESI): Calcd for $C_{11}H_{15}O_2N$, $[M + Na]^+$

https://doi.org/10.1021/acs.biochem.2c00663
Biochemistry XXXX, XXX, XXX−XXX

216.0995; found, $[M + Na]^+$ 216.0994 (Δm = +0.0001 and error = +0.7 ppm).

Synthesis of N^2-Methy-(3,4-dimethoxycinnamyl)-O^6-(2-(4-nitrophenyl)-ethyl)-3′,5′-diacetyl-2′-deoxyguanosine (7a). Method E was followed using *rac*-BINAP (99 mg, 0.15 mmol) and $Pd(OAc)_2$ (23 mg, 0.10 mmol) in dry toluene (5.8 mL) followed by the addition of Cs_2CO_3 (242 mg, 0.74 mmol), 3,4-dimethoxy-*trans*-cinnamyl amine **5a** (113 mg, 0.58 mmol), and bromo nucleoside **6** [synthesized using the previously reported protocol][33] (300 mg, 0.53 mmol). The reaction mixture was stirred at 85 °C for 12 h to obtain compound **7a** as a yellowish solid. Yield: 59% (213 mg); R_f = 0.3 (50% ethyl acetate in petroleum ether); mp 64−68 °C; ^{1}H NMR: (400 MHz, $CDCl_3$) δ: 8.12 (d, J = 8.7 Hz, 2H), 7.70 (s, 1H), 7.44 (d, J = 8.6 Hz, 2H), 6.90 (d, J = 1.8 Hz, 1H), 6.87 (dd, J = 8.2, 1.8 Hz, 1H), 6.79 (d, J = 8.2 Hz, 1H), 6.50 (d, J = 15.8 Hz, 1H), 6.25 (t, J = 6.8 Hz, 1H), 6.16 (dt, J = 15.8, 5.9 Hz,1H), 5.47 (m, 1H), 5.16 (t, J = 5.8 Hz, 1H), 4.71 (t, J = 6.7 Hz, 2H), 4.45 (m, 1H), 4.32 (m, 2H), 4.22 (m, 2H), 3.86 (s, 6H), 3.25 (t, J = 6.7 Hz, 2H), 3.11 (m, 1H), 2.52 (ddd, J = 9.3, 6.3, 3 Hz, 1H), 2.11 (s, 3H), 2.05 (s, 3H); ^{13}C NMR: (100 MHz, $CDCl_3$) δ: 170.6, 170.3, 160.7, 158.8, 153.6, 149.0, 148.8, 146.8, 146.0, 137.6, 131.1, 129.9, 129.8, 124.6, 123.7, 119.4, 115.7, 111.1, 108.7, 84.5, 82.2, 74.5, 66.0, 63.9, 55.9, 55.8, 44.1, 36.5, 35.2, 21.0, 20.8; HRMS (ESI): Calcd for $C_{33}H_{36}O_{10}N_6$, $[M + H]^+$ 677.2566; found, $[M + H]^+$ 677.2565 (Δm = +0.0001 and error = +0.1 ppm).

Synthesis of N^2-Methy-(3,4-dimethoxycinnamyl)-O^6-(2-(4-nitrophenyl)-ethyl)-2′-deoxyguanosine (8a). Method F was followed using compound **7a** (500 mg, 0.73 mmol) and 33% aq. $MeNH_2$ (5.1 mL, 7 mL/mmol) in 1,4-dioxane (10.3 mL, 14 mL/mmol). The reaction mixture was stirred at room temperature for 2 h to obtain compound **8a** as a fluffy yellowish solid. Yield: 81% (352 mg); R_f = 0.3 (15% MeOH in DCM); mp 69−71 °C; ^{1}H NMR: (400 MHz, $CDCl_3$) δ: 8.10 (d, J = 8.4 Hz, 2H), 7.60 (s, 1H), 7.41 (d, J = 8.4 Hz, 2H), 6.88−6.82 (m, 2H), 6.78 (d, J = 8 Hz, 1H), 6.48 (d, J = 15.6 Hz, 1H), 6.22−6.18 (m, 1H), 6.15−6.08 (m, 1H), 5.22 (t, J = 5.5 Hz, 1H), 4.75−4.71 (m, 3H), 4.17 (t, J = 5.2 Hz, 3H), 3.95−3.92 (m, 1H), 3.86 (s, 3H), 3.85 (s, 3H), 3.75 (brs, 1H), 3.25 (t, J = 6.8 Hz, 2H), 3.03−2.98 (m, 1H), 2.24 (dd, J = 13.3, 5.6 Hz, 1H); ^{13}C NMR: (100 MHz, $CDCl_3$) δ: 161.0, 158.3, 152.8, 149.1, 148.9, 146.8, 145.9, 139.1, 131.4, 129.9, 129.7, 124.4, 123.8, 119.5, 116.5, 111.1, 108.7, 89.2, 87.4, 73.4, 66.1, 63.5, 56.0, 55.9, 44.3, 40.2, 35.2; HRMS (ESI): Calcd for $C_{29}H_{32}O_8N_6$, $[M + H]^+$ 593.2354; found, $[M + H]^+$ 593.2353 (Δm = +0.0001 and error = +0.1 ppm).

Synthesis of N^2-Methy-(3,4-dimethoxycinnamyl)-O^6-(2-(4-nitrophenyl) ethyl)-5′-(4,4′-dimethoxytrityl)-2′-deoxyguanosine (9a). Method G was followed using compound **8a** (200 mg, 0.33 mmol), DMT-Cl (125 mg, 0.37 mmol), and DMAP (3.6 mg, 0.03 mmol) in dry pyridine (4.9 mL, 15 mL/mmol) to get a yellowish solid. Yield: 47% (139 mg); R_f = 0.3 (50% ethyl acetate in petroleum ether); mp 103−105 °C; ^{1}H NMR: (400 MHz, $CDCl_3$) δ. 8.11 (d, J = 8.4 Hz, 2H), 7.68 (s, 1H), 7.10 (d, J = 8.4 Hz, 2H), 7.38 (d, J = 6.8 Hz, 2H), 7.29−7.26 (m, 4H), 7.24−7.17 (m, 2H), 6.89−6.84 (m, 2H), 6.80−6.76 (m, 5H), 6.46 (t, J = 15.6 Hz, 1H), 6.31 (t, J = 6.4 Hz, 1H), 6.14−6.07 (m, 1H), 4.97 (t, J = 5.2 Hz, 1H), 4.72−4.65 (m, 3H), 4.13−4.08 (m, 3H), 3.86 (s, 3H), 3.85 (s, 3H), 3.74 (s, 6H), 3.42−3.32 (m, 2H), 3.25 (t, J = 6.8 Hz, 2H), 2.85−2.78 (m, 1H), 2.45−2.39 (m, 1H); ^{13}C NMR: (100 MHz, $CDCl_3$) δ: 160.6, 158.7, 158.6, 153.8, 149.1, 148.9, 146.9, 146.1, 144.6, 137.6, 135.7, 131.1, 130.1, 130.0, 129.9, 128.2, 128.0, 127.0, 124.8, 123.8, 119.5, 115.5, 113.3, 111.2, 108.8, 86.6, 85.8, 83.8, 72.9, 66.0, 64.0, 56.0, 55.9, 55.3, 44.1, 39.6, 35.3; HRMS (ESI): Calcd for $C_{50}H_{50}O_{10}N_6$, $[M + H]^+$ 895.3661; found, $[M + H]^+$ 895.3661.

Synthesis of N^2-Methy-(3,4-dimethoxycinnamyl)-O^6-(2-(4-nitrophenyl) ethyl)-5′-(4,4′-dimethoxytrityl)-2′-deoxyguanosine phosphoramidite (10a). Method H was followed using nucleoside **9a** (200 mg, 0.22 mmol) in DCM (2.2 mL), DIPEA (0.19 mL, 1.12 mmol), and CEP-Cl (0.1 mL, 0.44 mmol). The reaction mixture was stirred at room temperature for 2 h to furnish a yellowish solid. Yield: 65% (156 mg); R_f = 0.4 (50% ethyl acetate in petroleum ether +2% Et_3N); mp 75−81 °C; ^{31}P NMR: (162 MHz, $CDCl_3$) δ: 148.83, 148.71; HRMS (ESI): Calcd for $C_{59}H_{67}O_{11}N_8P$, $[M + H]^+$ 1095.4740; found, $[M + H]^+$ 1095.4735 (Δm = +0.0005 and error = +0.4 ppm).

Preparation of 4-Methoxy-*trans*-cinnamate (2b). Method A was followed using 4-methoxy benzaldehyde (500 mg, 3.67 mmol) and PPh_3 (1.4 g, 5.50 mmol) in saturated aq. $NaHCO_3$ solution (15 mL) followed by the addition of ethyl bromoacetate (1.2 mL, 11.0 mmol) to obtain colorless liquid **2b**.[31] Yield: 87% (660 mg); R_f = 0.4 (30% ethyl acetate in petroleum ether); ^{1}H NMR: (400 MHz, $CDCl_3$) δ: 7.63 (d, J = 15.6 Hz, 1H), 7.45 (d, J = 8.4 Hz, 2H), 6.88 (d, J = 8.8 Hz, 2H), 6.29 (d, J = 15.6 Hz, 1H), 4.25 (q, J = 6.8 Hz, 2H), 3.82 (s, 3H), 1.32 (t, J = 6.8 Hz, 3H); ^{13}C NMR: (100 MHz, $CDCl_3$) δ: 167.4, 161.4, 144.3, 129.8, 127.3, 115.8, 114.4, 60.4, 55.4, 14.4; HRMS (ESI): Calcd for $C_{12}H_{14}O_3$, $[M + Na]^+$ 229.0835; found, $[M + Na]^+$ 229.0835.

Preparation of 4-Methoxy-*trans*-cinnamyl Alcohol (3b). Method B was followed using compound **2b** (1 g, 4.85 mmol), in toluene (20 mL), and DIBAL-H (3 mL, 16.90 mmol of 1.0 M in THF) to obtain a light greenish solid **3b**.[32] Yield: 73% (584 mg); R_f = 0.3 (30% ethyl acetate in petroleum ether); mp 62−65 °C; ^{1}H NMR: (500 MHz, CDCl3) δ: 7.32 (d, J = 8.5 Hz, 2H), 6.85 (d, J = 8.5 Hz, 2H), 6.55 (d, J = 15.6 Hz, 1H), 6.23 (dt, J = 15.8, 5.8 Hz, 1H), 4.30 (t, J = 4.5 Hz, 2H), 3.80 (s, 3H); ^{13}C NMR: (125 MHz, CDCl3) δ: 131.1, 129.5, 127.8, 126.4, 114.2, 64.1, 55.5; HRMS (ESI): m/z calcd for $C_{10}H_{12}O_2$, $[M + H]^+$ 164.0837; found, $[M + H]^+$ 164.0836 (Δm = +0.0001 and error = +0.3 ppm).

Synthesis of 4-Methoxy-*trans*-cinnamyl azide (4b). Method C was followed using compound **3b** (600 mg, 3.65 mmol), sodium azide (285 mg, 4.38 mmol), and PPh_3 (2.01 g, 7.66 mmol) in CCl_4-DMF (18.2 mL, 1:4 v/v) to obtain yellowish liquid **4b**. Yield: 69% (411 mg); R_f = 0.4 (5% ethyl acetate in petroleum ether); ^{1}H NMR: (400 MHz, $CDCl_3$) δ: 7.34 (d, J = 8.6 Hz, 2H), 6.87 (d, J = 8.8 Hz, 2H), 6.59 (d, J = 15.7 Hz, 1H), 6.10 (dt, J = 15.7, 6.7 Hz, 1H), 3.91 (d, J = 6.8 Hz, 2H), 3.81 (s, 3H); ^{13}C NMR: (100 MHz, $CDCl_3$) δ: 159.8, 134.4, 128.9, 128.0, 120.2, 114.2, 55.5 53.4; HRMS (ESI): Calcd for $C_{10}H_{11}N_3O$, $[M + H]^+$ 189.0902; found, $[M + H]^+$ 189.0904 (Δm = −0.0002 and error = −0.3 ppm).

Synthesis of 4-Dimethoxy-*trans*-cinnamyl amine (5b). Method D was followed using compound **4b** (500 mg, 3.06 mmol), PPh_3 (962 mg, 3.67 mmol) in THF (7.6 mL), and water (1.5 mL) to obtain greenish viscous liquid **5b**. Yield: 75% (375 mg); R_f = 0.3 (10% MeOH in DCM); ^{1}H NMR: (400 MHz, $CDCl_3$) δ: 7.29 (d, J = 8.4 Hz, 2H), 6.84 (d, J = 8.8 Hz, 2H), 6.43 (d, J = 15.6 Hz, 1H), 6.17 (td, J = 15.8, 6.0 Hz, 1H), 3.79 (s, 3H), 3.45 (dd, J = 6.1, 1.5 Hz, 2H), 1.38 (s, 2H); ^{13}C NMR: (100 MHz, $CDCl_3$) δ: 159.1, 130.1, 129.3, 129.0,

https://doi.org/10.1021/acs.biochem.2c00663
Biochemistry XXXX, XXX, XXX−XXX

127.5, 114.1, 55.4, 44.6; HRMS (ESI): Calcd for $C_{10}H_{13}NO$, $[M + H]^+$ 163.0997; found, $[M + H]^+$ 163.0995 (Δm = +0.0002 and error = +0.4 ppm).

Synthesis of N^2-Methy-(4-methoxycinnamyl)-O^6-(2-(4-nitrophenyl)-ethyl)-3′,5′-diacetyl-2′-deoxyguanosine (7b). Method E was followed using *rac*-BINAP (197 mg, 0.31 mmol) and $Pd(OAc)_2$ (51 mg, 0.21 mmol) in dry toluene (10 mL) followed by the addition of Cs_2CO_3 (522 mg, 1.48 mmol), 4-methoxy-*trans*-cinnamyl amine **5b** (190 mg, 1.16 mmol), and bromo nucleoside **6** [synthesized using the previously reported protocol][33] (600 mg, 1.06 mmol). The reaction mixture was stirred at 85 °C for 10 h to obtain compound **7b** as a fluffy yellowish solid. Yield: 62% (425 mg); R_f = 0.3 (50% ethyl acetate in petroleum ether); mp 55−59 °C; ^{1}H NMR: (400 MHz, $CDCl_3$) δ: 8.13 (d, J = 8 Hz, 2H), 7.71 (s, 1H), 7.45 (d, J = 8.4 Hz, 2H), 7.27 (d, J = 8.4 Hz, 2H), 6.83 (d, J = 8.4 Hz, 2H), 6.52 (d, J = 15.6 Hz, 1H), 6.27 (t, J = 6.8 Hz, 1H), 6.16 (dt, J = 15.8, 6 Hz, 1H), 5.48 (m, 1H), 5.14 (t, J = 5.6 Hz, 1H), 4.73 (t, J = 6.8 Hz, 2H), 4.45−4.40 (m, 1H), 4.31−4.26 (m, 2H), 4.20 (t, J = 5.6 Hz, 2H), 3.79 (s, 3H), 3.26 (t, J = 6.4 Hz, 2H), 3.11−3.03 (m, 1H), 2.50 (ddd, J = 14.1, 6.3, 2.9 Hz, 1H), 2.11 (s, 3H), 2.05 (s, 3H); ^{13}C NMR: (100 MHz, $CDCl_3$): 170.6, 170.4, 160.8, 159.3, 158.9, 153.7, 146.9, 146.1, 137.7, 131.0, 130.0, 129.6, 127.5, 124.4, 123.8, 115.8, 114.1, 84.6, 82.3, 74.6, 66.0, 64.0, 55.4, 44.3, 36.6, 35.3, 21.1, 20.9; HRMS (ESI): m/z calcd for $C_{32}H_{34}O_9N_6$, $[M + H]^+$ 647.2460; found, $[M + H]^+$ 647.2460.

Synthesis of N^2-Methy-(4-methoxycinnamyl)-O^6-(2-(4-nitrophenyl)-ethyl)-2′-deoxyguanosine (8b). Method F was followed using compound **7b** (500 mg, 0.77 mmol) and 33% aq. $MeNH_2$ (5.4 mL; 7 mL/mmol) in 1,4-dioxane (10.8 mL, 14 mL/mmol). The reaction mixture was stirred at room temperature for 4 h to obtain compound **8b** as a fluffy yellowish solid. Yield: 80% (349 mg). R_f = 0.3 (15% MeOH in DCM); mp 68−72 °C; ^{1}H NMR: (400 MHz, $CDCl_3$) δ: 8.08 (d, J = 8 Hz, 2H), 7.63 (s, 1H), 7.40 (d, J = 8 Hz, 2H), 7.23 (J = 8.4 Hz, 2H), 6.80 (d, J = 8.4 Hz, 2H), 6.47 (d, J = 15.6 Hz, 1H), 6.21 (t, J = 5.6 Hz, 1H), 6.13−6.06 (m, 1H), 5.28 (s, 1H), 4.71−4.70 (m, 3H), 4.15−4.14 (m, 3H), 3.92−3.89 (m, 1H), 3.77 (s, 3H), 3.74−3.71 (m, 1H), 3.23 (t, J = 6.4 Hz, 2H), 2.97 (brs, 1H), 2.26−2.22 (m, 1H); ^{13}C NMR: (100 MHz, $CDCl_3$) δ: 161.0, 159.3, 158.4, 152.8, 146.9, 145.9, 139.1, 131.2, 130.0, 129.5, 127.5, 124.2, 123.8, 116.3, 114.1, 89.1, 87.3, 73.3, 66.2, 63.5, 55.4, 44.4, 40.2, 35.2; HRMS (ESI): Calcd for $C_{28}H_{30}O_7N_6$, $[M + Na]^+$ 585.2068; found, $[M + Na]^+$ 585.2072 (Δm = −0.0004 and error = −0.6 ppm).

Synthesis of N^2-Methy-(4-methoxycinnamyl)-O^6-(2-(4-nitrophenyl)-ethyl)-5′-(4,4′-dimethoxytrityl)-2′-deoxyguanosine (9b). Method G was followed using compound **8b** (100 mg, 0.17 mmol), DMT-Cl (66 mg, 0.19 mmol), and DMAP (2 mg, 0.01 mmol) in dry pyridine (1.8 mL, 10 mL/mmol) to furnish a yellowish solid. Yield: 60% (89 mg); R_f = 0.3 (50% ethyl acetate in petroleum ether); mp 101−106 °C; ^{1}H NMR: (400 MHz, $CDCl_3$) δ: 8.15 (d, J = 8.4 Hz, 2H), 7.73 (s, 1H), 7.47−7.42 (m, 4H), 7.33−7.28 (m, 8H), 7.23−7.19 (m, 1H), 6.86 (d, J = 8.8 Hz, 2H), 6.81 (dd, J = 8.7, 1.7 Hz, 4H), 6.51 (d, J = 15.6 Hz, 1H), 6.35 (t, J = 6.4 Hz, 1H), 6.16−6.09 (m, 1H), 5.01 (t, J = 5.6 Hz, 1H), 4.75 (t, J = 6.8 Hz, 2H), 4.68−4.65 (m, 1H), 4.14−4.12 (m, 3H), 3.83 (s, 3H), 3.78 (s, 6H), 3.44−3.36 (m, 2H), 3.29 (t, J = 6.4 Hz, 2H), 2.48−2.42 (m, 1H); ^{13}C NMR: (100 MHz, $CDCl_3$) δ: 160.5, 159.2, 158.7, 158.5, 153.8, 146.7, 146.0, 144.6, 139.6, 137.7, 135.7, 130.8, 130.0, 129.9, 129.5, 129.2, 128.1, 127.9, 127.4, 126.9, 124.4, 123.7, 115.2, 114.0, 113.2, 86.5, 86.09, 84.0, 72.5, 65.9, 64.0, 55.3, 55.2, 44.1, 39.6, 35.2; HRMS (ESI): Calcd for $C_{49}H_{48}O_9N_6$, $[M + H]^+$ 865.3556; found, $[M + H]^+$ 865.3557 (Δm = −0.0001 and error = −0.2 ppm).

Synthesis of N^2-Methy-(4-methoxycinnamyl)-O^6-(2-(4-nitrophenyl)-ethyl)-5′-(4,4′-dimethoxytrityl)-2′-deoxyguanosine Phosphoramidite (10b). Method H was followed using compound **9b** (200 mg, 0.23 mmol) in DCM (2.3 mL), DIPEA (0.2 mL, 1.15 mmol), and CEP-Cl (0.1 mL, 0.46 mmol). The reaction mixture was stirred at room temperature for 40 min to furnish a yellowish solid. Yield: 66% (162 mg); R_f = 0.4 (50% ethyl acetate in petroleum ether +2% Et_3N); mp 70−74 °C; ^{31}P NMR: (162 MHz, $CDCl_3$) δ: 148.80, 148.69; HRMS (ESI): Calcd for $C_{58}H_{65}O_{10}N_8P$, $[M + H]^+$ 1065.4634; found, $[M + H]^+$ 1065.4632 (Δm = +0.0002 and error = +0.2 ppm).

Oligonucleotide Synthesis. DNA sequences were synthesized on a K&A automated DNA synthesizer. The unmodified and MEG-dG and EG-dG modified DNA sequences were synthesized on a 1 μmol scale using controlled pore glass (CPG) as a solid support. The coupling time used for the unmodified phosphoramidites was 2 min and for the MEG-dG and EG-dG modified phosphoramidites was 10 min. 5-Ethylthio-1*H*-tetrazole (ETT) was used as the coupling agent. The deprotection of MEG-dG and EG-dG modified oligos was carried out in four different steps.[33] The first step involved the selective deprotection of the cyanoethyl group using 10% diethylamine in acetonitrile (ACN) at room temperature for 10 min. The second step involved the removal of the NPE group with 1 M DBU in ACN (1 mL) for 1 h at room temperature. Then, the oligonucleotides were treated with 30% aq. NH_3 for cleavage of the CPG support at room temperature for 3 h followed by treatment with 30% aq. NH_3 at 55 °C for 16 h to remove the base-protecting groups. The supernatant layer was collected, and the CPG beads were washed with water (1 × 300 μL). The combined aqueous layer was evaporated on a vacuum concentrator to get the crude DNA, which was further purified using 20% denaturing PAGE (7 M urea) at 30 W for 3 h with 1× TBE running buffer (89 mM each Tris and boric acid and 2 mM EDTA, pH 8.3). The gel thickness was 1 mm, and the gel dimension was 20 × 30 mm. The gel was visualized under a UV lamp (260 nm), and desired DNA bands were marked. The gel bands were crushed, and 15 mL of TEN (10 mM Tris, 1 mM EDTA, and 300 mM NaCl, pH 8.0) was used. It was further shaken for 16 h at room temperature to recover the DNA. Finally, desalting of oligonucleotides was carried out using a C18 Sep-Pak column. Desalted DNAs were dissolved in water, and the absorbance was measured at 260 nm in a UV−visible spectrometer using molar extension coefficients (ε); for **T1** and **T2**, 214,000 L m^{-1} cm^{-1}. All the oligonucleotides were characterized using MALDI-TOF in positive reflection mode, and the molecular weights are provided in Table S5.

Expression and Purification of hpolκ and hpolη. The constructs pBG101-hpolκ (encoding hpolκ amino acids 19−526; UniProt Accession Q9UBT6) and pET28a-hpolη (encoding hpolη amino acids 1−432; UniProt Accession Q9Y253) were kind gifts from Prof. Guengerich's lab, Department of Biochemistry, Vanderbilt University, USA). Both plasmids were transformed into *Escherichia coli* BL21 (DE3) RIPL cells for expressing and purifying the respective polymerases as described below.

https://doi.org/10.1021/acs.biochem.2c00663
Biochemistry XXXX, XXX, XXX−XXX

Expression and Purification of hpolκ. *E. coli* BL21 (DE3) RIPL cells containing the pBG101-hpolκ were grown at 37 °C in the presence of antibiotics (chloramphenicol 25 μg/mL and kanamycin 50 μg/mL) till OD_{600} reaches 0.6. The protein expression was then induced using 1 mM IPTG, and the cells were further incubated at 16 °C for 16 h at 130 rpm. Cells were harvested by centrifugation and resuspended in a binding buffer containing 50 mM Tris−HCl (pH 7.4), 150 mM NaCl, 1 mM DTT, and 20 mM imidazole.[34] After ultrasonication, the lysate was centrifuged at 12,500*g* for 30 min. The supernatant was then passed through a His-Trap column (GE healthcare) pre-equilibrated with a binding buffer. Protein was eluted using a linear imidazole gradient. Fractions containing the desired protein were identified using SDS-PAGE. Then, protein containing fractions were collected and concentrated to 2 mL using Vivaspin turbo centrifugal concentrators (Merck). The concentrated protein was further treated with PreScission protease (GE Healthcare; 2 U/100 microgram protein) at 4 °C for 15 h to remove the GST tag. The GST tag was then separated by passing the protein through 0.5 mL GST beads. Further, the flow-through was concentrated to 2 mL and subjected to gel filtration chromatography using a 16/600 Superdex 200 gel filtration column connected to an AKTA Purifier system (GE Healthcare). The resulting protein fractions were analyzed on SDS-PAGE to check purity, pooled, and concentrated to about 200 μL. The fractions were then stored at −80 °C. The purity of the proteins was assessed with SDS-PAGE followed by Coomassie Blue staining.

Expression and Purification hpolη. hpolη was expressed similar to hpolκ, except for 0.5 mM IPTG that was used for inducing protein expression.[35] The cell pellets were resuspended in a binding buffer containing 50 mM Tris−HCl (pH 7.5), 500 mM NaCl, 1 mM DTT, and 20 mM imidazole and lysed by ultra-sonication. The lysate was centrifuged at 12,500*g* and loaded onto a His-Trap column. Protein was eluted using a linear imidazole gradient. Fractions containing the protein were collected and concentrated to 2 mL. The concentrated protein was further subjected to PreScission protease cleavage (GE Healthcare; 2 U/100 microgram protein) at 4 °C for 15 h on slow rocking. The protein fractions were collected, concentrated to 2 mL by using Vivaspin turbo centrifugal concentrators (Merck), and further subjected to gel filtration chromatography using a 16/600 Superdex 200 gel filtration column connected to an AKTA Purifier system (GE Healthcare). The resulting protein fractions were pooled, concentrated, and then stored at −80 °C. The protein concentration was measured by the Bradford assay using bovine serum albumin (BSA) as a standard.

Full-Length Extension Assays. A primer-template complex was generated by annealing a 15-mer FAM (6-carboxyfluorescein) labeled DNA to either unmodified 27-mer or MEG-dG or EG-dG modified templates. The primer and template DNA were mixed in a 1:1 molar ratio followed by heating at 95 °C for 5 min and slow cooling to room temperature.[36] Full-length extension reactions were performed at 37 °C by incubating the primer-template complex (500 nM) with 5 nM of hpolκ or hpolη in a 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. The reactions were 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 a 5 μL loading dye containing 1 mg/mL bromophenol blue, 20 mM EDTA, and 80% formamide (v/v). Reactions were heated at 95 °C for 5 min, and the products were separated using 20% PAGE. Gels were visualized using a Typhoon FLA 9500 scanner (GE Healthcare).

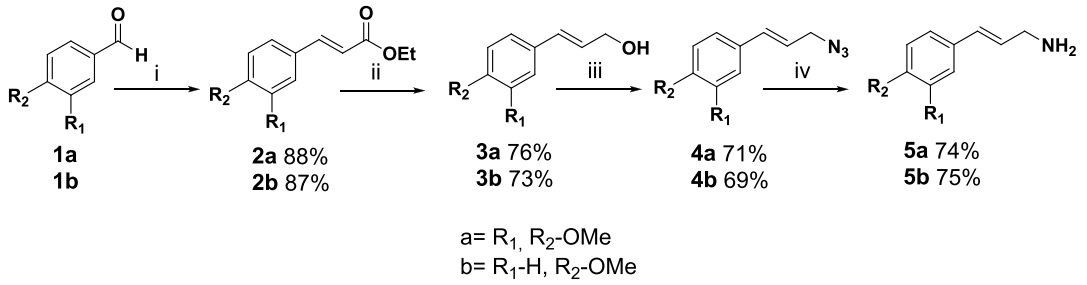
Single-Nucleotide Insertion Assays. A 15-mer FAM labeled primer was annealed to unmodified 27-mer or 27-mer templates containing either the MEG-dG or EG-dG adduct that was used for single-nucleotide insertion assays. The primer-template substrate (500 nM) was incubated with 5 nM of hpolκ or hpolη in the polymerase buffer described above. Varying concentrations (0.1, 1, 10, 20, or 40 μM) of all four individual dNTPs were added to initiate the reaction (total reaction volume of 10 μL). Reactions were incubated at 37 °C for 10 min and stopped using the gel loading dye. Samples were heated at 95 °C for 5 min and then separated using a 20% PAGE. Gels were visualized as described above for full-length extension assays.

Steady-State Kinetics Analysis. Steady-state kinetics were performed using the same primer-template DNA complex used for full-length extension and single-nucleotide insertion assays. However, the initial DNA concentration of the duplex was maintained in excess (2 μM). The enzyme and dNTP concentrations were optimized to maintain the steady-state conditions, i.e., product formation is ≤30% of the total substrate used.[36] For hpolκ, the 4 nM enzyme was incubated with the primer-template complex containing either the undamaged or MEG-dG or EG-dG adducts for 5 min. dCTP concentrations of 0, 0.5, 1, 2, 5, 10, and 20 μM were used for MEG-dG and 0, 0.5, 1, 2, 5, 10, 20, and 50 μM for EG-dG. For hpolη, a 5 nM enzyme and incubation time of 10 min were used along with dCTP concentrations of 0, 2, 3, 4, 5, 10, 15, 20, and 30 μM for MEG-dG and 0, 2, 3, 4, 5, 10, 15, 20, 30, and 40 μM for EG-dG. All reactions were pre-incubated at 37 °C for 5 min before adding dCTP. Reactions were stopped using a gel loading dye, and products were separated and visualized as described previously for full-length extension assays. All reactions were performed in duplicates, and the data points shown indicate mean values along with the standard deviation (±SD). The integrated densities of the products and substrates were calculated using the ImageJ software. The ratio of product to the total DNA substrate was used to estimate the k_{cat} and K_m by fitting the data to a Michaelis−Menten hyperbolic equation using the Origin software (OriginLab, U.S.A.).[37]

General Procedures for MD Simulation. *Starting Structures.* The starting structures for the insertion and post-insertion stages were derived from the X-ray crystal structure of DNA with the lucidin adduct in complex hpolκ (PDB ID: 5W2A and 5W2C, respectively). The missing residues in the protein were added by homology modeling with the help of the Swiss model.[38] The chemically modified dNTPs in the crystal structures were replaced by normal dNTPs, and previously calculated force field parameters were used (http://upjv.q4md-forcefieldtools.org/REDDB/projects/F-90/). The structure of the N^2-MEG-dG adduct was generated using Gauss View 6.0 and loaded into the R.E.D Server[39−41] along with a dimethylphosphate molecule to generate the force field parameters needed to define the adduct. The damaged nucleotide was incorporated into the DNA system using the xleap module of AMBER 18, and the coordinate and topology files for the complexes were generated. The unmodified dG

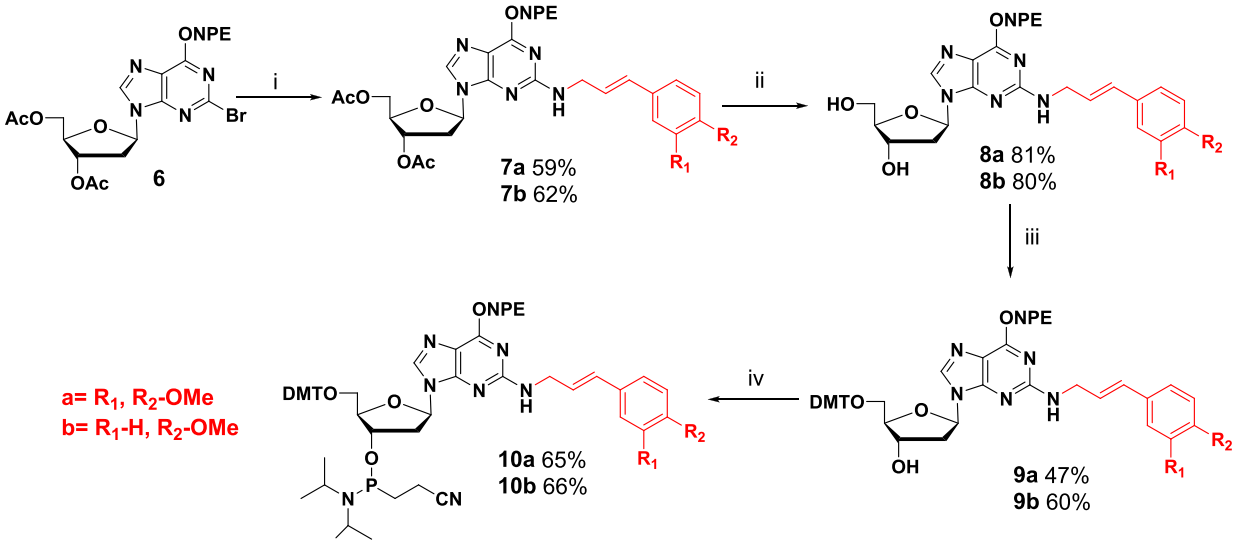
https://doi.org/10.1021/acs.biochem.2c00663
Biochemistry XXXX, XXX, XXX−XXX

Scheme 1. Synthesis of the 3,4-Dimethoxy-*trans*-cinnamyl Amine or 4-Methoxy-*trans*-cinnamyl Amine[a]



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

Scheme 2. Synthesis of the N^2-MEG-dG and N^2-EG-dG Phosphoramidites[a]



[a]Reagents and conditions: (i) compound **6**, $Pd(OAc)_2$, *rac*-BINAP, Cs_2CO_3, toluene, 85 °C, 12 h; (ii) 33% $MeNH_2$ in 1,4-dioaxne, RT, 4 h; (iii) DMT-Cl, pyridine, 85 °C, 30 min, MW (150 W); and (iv) CEP-Cl, DIPEA, DCM, RT, 2 h.

complex was also generated as a control system. The AMBER force fields ff14SB[42] and parmbsc1[43] were used for protein and DNA, respectively.

Simulation Protocol. The complexes generated were subjected to a short energy minimization run of 5000 steps to reduce bad contacts. The system was then neutralized by adding Na^+ ions and solvated using a TIP3P water box extending 10 Å from the atoms. Two-stage energy minimization was done for 10,000 steps each (5000 steps of steepest descent cycles followed by 5000 steps of conjugate gradient cycles). In the first step, the protein-DNA complex was restrained with a force constant of 10 kcal/mol·Å^2, which was later removed during the second minimization step. The system was then heated from 0 to 300 K over 100 ps under constant volume conditions with a restraint of 10 kcal/mol·Å^2. The restraints were slowly removed over eight stages of equilibration (with restraints of 10, 8, 5, 4, 3, 2, 1, and 0.5 kcal/mol·Å^2). Each equilibration stage was followed by energy minimization (5000 steepest descent and 5000 conjugate gradient cycles), and the minimized structure was used as the starting structure for the next equilibration stage. After eight stages, 100 ps of equilibration was run under NPT conditions without any restrictions before proceeding to the production run. Unrestrained dynamics (500 ns) (under NPT conditions) was performed on all the complexes. The production runs were performed using the GPU accelerated version of particle mesh Ewald molecular dynamics (PMEMD)[44,45] implemented in AMBER 18.[46] The trajectories were visualized using PyMOL, and the MD trajectories were analyzed by the CPPTRAJ[47] module of AMBER 18. Berendsen[48] weak coupling barostat was used to maintain the pressure at 1 atm with a pressure relaxation time of 2 ps, and a Langvein[49] thermostat was used to maintain the temperature at 300 K. All the bonds involving hydrogen atoms were constrained using the SHAKE algorithm with a time step of 2 fs.

Trajectory Analysis. RMSD of the backbone atoms, RMSF, interatomic distances, and dihedral angles were calculated using the CPPTRAJ module of AMBER 18. A total of 250,000 frames were used for the analysis. The hydrogen bonding occupancies were determined by using the "hbond" command of the CPPTRAJ module, with a donor−acceptor distance cutoff value of 3.4 Å and a donor−hydrogen−acceptor angle cutoff of 120°. The hydrogen bonding interaction energies and the base stacking energies were determined by using the "lie" command of the CPPTRAJ module. Only nucleobase moieties of the interacting nucleotides were considered for the interaction energy calculations. Snapshots from the trajectory were clustered into five ensembles by employing the hierarchical agglomerative algorithm implemented in CPPTRAJ. Pair-wise interaction energy between the adduct

https://doi.org/10.1021/acs.biochem.2c00663
Biochemistry XXXX, XXX, XXX−XXX

and the surrounding amino acid residues was calculated for the insertion stages by using the molecular mechanics/generalized born surface area (MM/GBSA)[50] method implemented in AMBER 18. The amino acids in close proximity to the DNA lesion and incoming nucleotides were considered for this analysis.

RESULTS AND DISCUSSION

Synthesis and Characterization of N^2-MEG-dG and N^2-EG-dG Modified DNAs. Our initial goal was to synthesize the N^2-MEG-dG and N^2-EG-dG modified phosphoramidites for their incorporation into DNA. Toward this end, we began preparing amines **5a** and **5b** (Scheme 1). Compounds **1a** and **1b** were converted into the corresponding α,β-unsaturated ester derivatives **2a** and **2b** by reacting with ethyl bromoacetate in saturated $NaHCO_3$ to give 88 and 87% yields, respectively.[31]

Then, ester derivatives were reduced to alcohols using DIBAL-H in dry toluene with **3a** in 76% and **3b** in 73% yields.[32] Alcohols were converted into azides by using sodium azide in CCl_4:DMF (1:4) to produce **4a** in 71% and **4b** in 69% yields.[51] Then, **4a** and **4b** were subjected to Staudinger reduction to give respective amines **5a** and **5b** in 74 and 75% yields, respectively.[52]

The key step to achieve the required modified phosphoramidites involved Buchwald−Hartwig coupling reaction. The amines **5a** and **5b** were coupled with 2-bromo-dG **6** to afford **7a** in 59% and **7b** in 62% yields (Scheme 2).[53−55] Then, coupled products were further subjected to deacetylation by using $MeNH_2$ in 1,4-dioxane to get nucleosides **8a** and **8b** in 81 and 80% yields, respectively. DMT protection of 5′-OH of compounds **8a** and **8b** was carried out under microwave conditions (150 W) at 85 °C for 30 min to afford compound **9a** in 47% and **9b** in 60% yields. DMT-protected nucleosides were further subjected to phosphorylation by using 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (CEP-Cl) to furnish respective phosphoramidites **10a** in 65% and **10b** in 66% yields.

After successfully synthesizing N^2-MEG-dG and N^2-EG-dG modified phosphoramidites, they were incorporated into desired DNA sequences (Table S5) using solid-phase DNA synthesis. The chemical integrity of modified DNAs was confirmed by MALDI-TOF spectrometry. The modified DNA sequences of **T1** (N^2-MEG-dG) and **T2** (N^2-EG-dG) were used for primer extension as well as steady-state kinetic studies with TLS polymerases hpolκ and hpolη.

Translesion Synthesis across N^2-MEG-dG and N^2-EG Adducts by hpolκ and hPolη Is Error-Free. Human TLS polymerase encoding constructs were obtained as pBG101-hpolκ (encoding hpolκ amino acids 19−526) and pET28a-hpolη (encoding hpolη amino acids 1−432). Both these plasmids were transformed into *E. coli* BL21 (DE3) RIPL cells for expression. The proteins were purified to near homogeneity using Ni^{2+}-NTA and gel filtration chromatography. The purified proteins hpolκ (∼58 kDa) and hpolη (∼48 kDa) were used for all the assays (Figure S1).

To check the ability of human TLS polymerases to bypass a MEG-dG or EG-dG adducts, we performed full-length extension assays using hpolκ and hpolη with 27-mer templates containing either an unmodified dG or the MEG-dG or EG-dG modifications at the 16th position (Figure 2A). Extension of the 15-mer-FAM labeled primer beyond the modified base by hpolκ (Figure 2B−D) and hpolη (Figure 2E−G) was monitored by stopping the reactions at desired time points

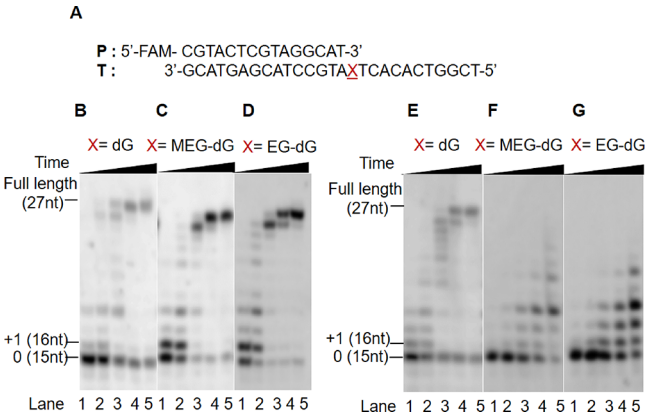


Figure 2. Full-length extension assay with hpolκ and hpolη in the presence of a mixture of dNTPs. (A) 15-mer primer and 27-mer template DNA sequence, where (B) X = dG, (C) X = MEG-dG, and (D) X = EG-dG in the presence of 5 nM hpolκ and (E) X = dG, (F) X = MEG-dG, and (G) X = EG-dG in the presence of 5 nM hpolη. All reactions were performed at 37 °C for 5, 10, 30, 60, and 120 min.

and visualizing the products upon a 20% PAGE as described in the Experimental Section. Results showed that hpolκ efficiently bypasses both MEG and EG-DNA adducts and extends to form the full-length product in high efficiency (Figure 2B−D). However, though hpolη could bypass the adducts, it could not form full-length extension products (Figure 2E−G). Inside the cells, a second polymerase-like polζ could act as an extender to continue the initial incorporation by hpolη.[56] We also performed a running start assay using a 13-mer primer with the adduct at the 16th position (Figure S2A). We observed that while extending the modified primer/template DNA substrates, hpolκ shows a slight pause at the 15th nucleotide position. This pause is absent while extending the unmodified primer/template DNA substrate (Figure S2B−D). Additionally, we also found that nearly all the modified DNA substrate (MEG-dG and EG-dG) is converted to the full-length form by hpolκ. For hpolη, however, the extension of the unmodified DNA is more efficient than MEG-dG or EG-dG modifications (Figure S2E−G). These results are in line with the previous reports on the TLS ability of hpolκ and hpolη across various N^2-dG adducts.[57,58] The bypass efficiency of hpolκ and hpolη for the unmodified and the MEG-dG and EG-dG modified DNA substrates was further explored using steady-state enzyme kinetics.

Further, to determine the fidelity of the bypass, we first performed a single-nucleotide incorporation assay (Figures 3A and 4A). We observed that hpolκ selectively incorporates dCTP across dG, EG-dG, and MEG-dG DNA templates (Figure 3B−D). No other nucleotide was incorporated within the concentration range tested across the unmodified or modified templates. Similar to hpolκ, hpolη also bypassed the modified base only with the complementary dC nucleotide (Figure 4B−D). Thus, overall, these studies reveal that both TLS polymerases could perform error-free bypass of the MEG-dG and EG-dG adducts.[57,58]

Steady-State Kinetics of Bypass across N^2-MEG-dG and N^2-EG Adducts by hpolκ and hpolη. To gain further insights into the bypass of MEG-dG and EG-dG DNA lesions by TLS polymerases, we performed steady-state kinetic studies. We determined the kinetic parameters for incorporating dCTP across the unmodified or modified substrates as both polymerases only incorporate the complementary base

https://doi.org/10.1021/acs.biochem.2c00663
Biochemistry XXXX, XXX, XXX−XXX

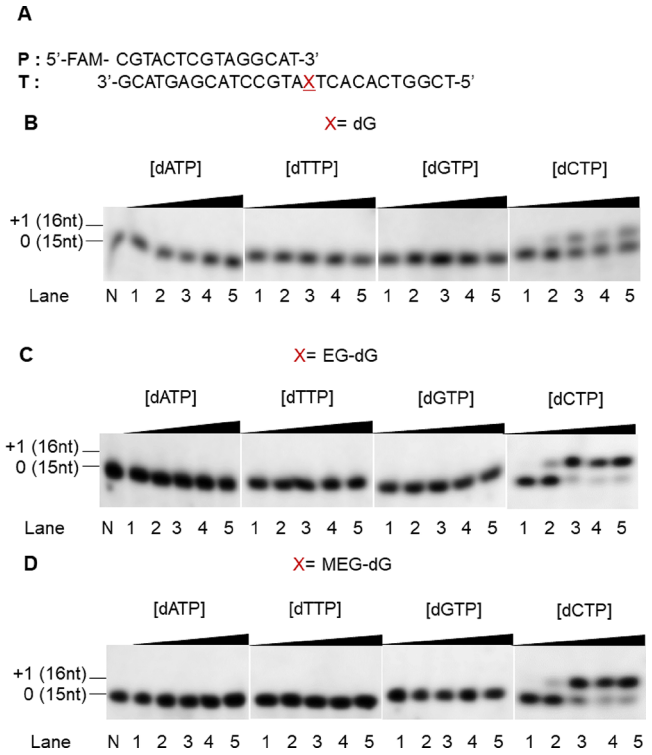


Figure 3. Single-nucleotide extension assay with hpolκ in the presence of individual dNTPs. (A) 15-mer primer and 27-mer template DNA sequence, where (B) X = dG, (C) X = EG-dG, and (D) X = MEG-dG. The reaction was performed in the presence of 5 nM hpolκ. All reactions were performed at 37 °C for 0.1, 1, 10, 20, or 40 μM with individual dNTPs (dATP, dTTP, dGTP, and dCTP).

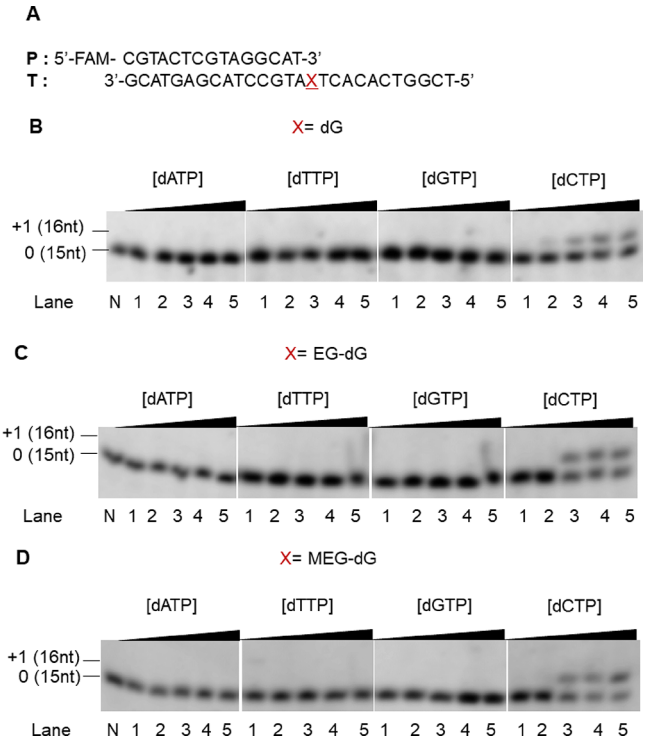


Figure 4. Single-nucleotide extension assay with hpolη in the presence of individual dNTPs. (A) 15-mer primer and 27-mer template DNA sequence, where (B) X = dG, (C) X = EG-dG, and (D) X = MEG-dG. The reaction was performed in the presence of 5 nM hpolη. All reactions were performed at 37 °C for 0.1, 1, 10, 20, or 40 μM with individual dNTPs.

(dCTP) across dG, MEG-dG, and EG-dG substrates. Reactions were performed similarly to that of single-nucleotide incorporation assays except in the presence of excess substrate DNA and varying concentrations of the dCTP. The products formed were separated on a 20% PAGE and quantified, and data points were fit to a hyperbolic curve to obtain the K_m and k_{cat} values (Figure S3 and Figure 5). For hpolκ, the K_m for dCTP incorporation across unmodified dG was found to be 2.4 (±0.8) (Table 1). The K_m values for the two modified DNA substrates MEG-dG and EG-dG were found to be 1.4 (±0.1) and 0.8 (±0.1), which are approximately 1.7- and 3-fold lower when compared with the unmodified dG, respectively (Table 1 and Figure S4). The enzyme turnover number or k_{cat} changes about 1.4 for the modified substrates compared to the unmodified dG, whereas the specificity constants k_{cat}/K_m for hpolκ of MEG-dG and EG-dG substrates are 2.5- and 4.4-fold higher than the unmodified dG substrate, respectively. Thus, the K_m and k_{cat}/K_m values obtained for MEG-dG and EG-dG substrates clearly show that these modifications can be bypassed more efficiently by hpolκ compared to the unmodified substrate. Guengerich and co-workers have reported similar results where the N^2-Bn-dG adduct was bypassed more efficiently by hpolκ compared to the unmodified substrate, and they proposed the formation of the activated enzyme-DNA-dNTP complex as the rate-limiting step.[57]

The K_m of hpolη for dCTP incorporation across the unmodified dG substrate was found to be 3.3 (±0.1) (Table 1). The K_m for the two modified DNA substrates, MEG-dG and EG-dG, were found to be 9.3 (±2.0) and 2.1 (±0.4), respectively (Table 1). Thus, the K_m of hpolη toward the MEG-dG substrate is around 2.8-fold higher and about 1.5-fold lower for the EG-dG substrate as compared to the unmodified dG template (Table 1 and Figure S5). The catalytic efficiency (k_{cat}) of hpolη is highest for the insertion of dCTP across the unmodified dG and follows the order dG > EG-dG > MEG-dG. The specificity constant (k_{cat}/K_m) is nearly the same for unmodified dG and EG-dG substrates. However, the MEG-dG modification decreased k_{cat}/K_m by about 6-fold, thereby exhibiting less substrate efficiency toward hpolη. A comparative analysis of both the polymerases thus showed that hpolκ has high affinity and catalytic efficiency while bypassing the MEG-dG and EG-dG substrates and may be the major polymerase involved in translesion synthesis across these DNA adducts *in vivo*. Although both the polymerases displayed error-free bypass of the DNA lesions, *in vitro*, further studies are required to understand the influence of other factors, such as other TLS polymerases, accessory proteins, on the fidelity and processivity of TLS across these lesions under cellular conditions.

Structural Stability of the DNA-hpolκ Complexes. Molecular dynamics simulations were performed to shed further light on the molecular mechanisms of DNA replication past the adduct and its accommodation in the active site of hpolκ. The complexes for the simulation (Figure 6A) were prepared from the X-ray crystal structures of N^2-dG-lucidin adducted DNA in complex with hpolκ (PDB ID: 5W2A for the pre-insertion and insertion stages and PDB ID:5W2C for the post-insertion stage)[59] by replacing the adduct with N^2-MEG-dG (Figure S6). The DNA sequences used in the simulation were identical to those in the X-ray crystal structures. The sequences of DNA duplexes used in all the three stages of

https://doi.org/10.1021/acs.biochem.2c00663
Biochemistry XXXX, XXX, XXX−XXX

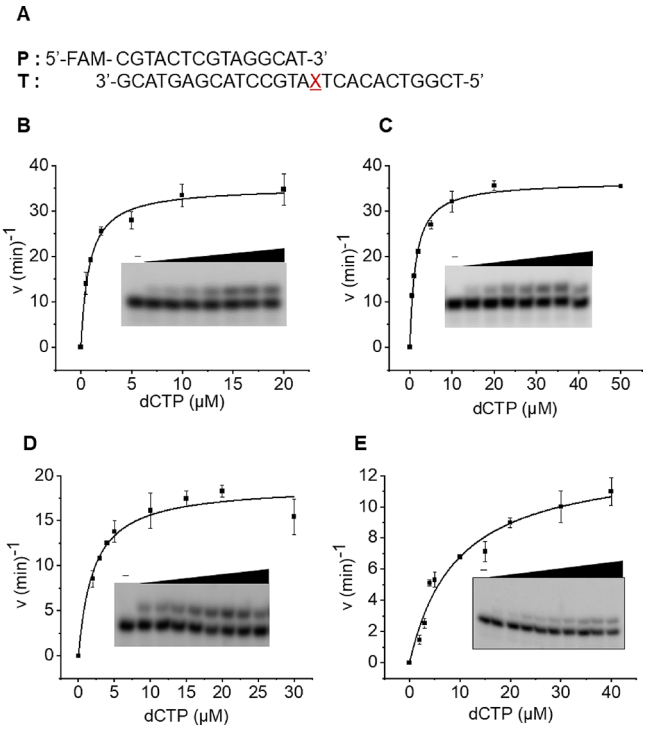


Figure 5. Steady-state kinetic analysis of dCTP insertion by hpolκ and hpolη. (A) 15-mer primer and 27-mer template DNA sequence, where X is dG, MEG-dG, or EG-dG. Reactions were performed in the presence of (B) 4 nM hpolκ with EG-dG, (C) 4 nM hpolκ with MEG-dG, (D) 5 nM hpolη with EG-dG, and (E) 5 nM hpolη with MEG-dG. All reactions were performed in duplicates at 37 °C with varying concentrations of dCTP (see the Experimental Section). The data points indicate mean values with standard deviation (±SD).

Table 1. Steady-State Kinetic Analysis of Insertion Opposite to N^2-MEG-dG and N^2-EG-dG Modified DNA

polymerase	DNA substrate	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m (μM^{-1} min^{-1})
hpolκ	unmodified	25.4 (±2.9)	2.4 (±0.8)	10.7 (±0.8)
	MEG	36.5 (±0.8)	1.4 (±0.1)	27.0 (±0.2)
	EG	35.3 (±1.0)	0.8 (±0.1)	46.9 (±11.2)
hpolη	unmodified	26.7 (±2.1)	3.3 (±0.1)	8.2 (±0.5)
	MEG	13.1 (±1.0)	9.3 (±2.0)	1.4 (±0.1)
	EG	18.9 (±0.9)	2.1 (±0.4)	9.3 (±2.1)

replication (pre-insertion, insertion, and post-insertion) are shown in Figure 6B. MD simulations (500 ns) were performed for each of the stages of the replication process. For comparison, MD simulation was also performed on the insertion stage of the unmodified dG complex. The resultant output trajectories were analyzed to gain insights into the process of translesion synthesis.

The stability of the system was assessed by calculating the average root mean square deviation (RMSD) of the protein and DNA backbone atoms of the complex with respect to the first frame. The residues that constitute the protein's modeled loop were not considered while calculating the RMSD. For all three stages (pre-insertion, insertion, and post-insertion stages), the average RMSD values were less than 3 Å, indicating that the complex was sufficiently stable throughout the simulation (Figure S7). The root mean square fluctuation (RMSF) of the amino acid residues in the protein follows a similar trend in all the three stages. RMSF showed major

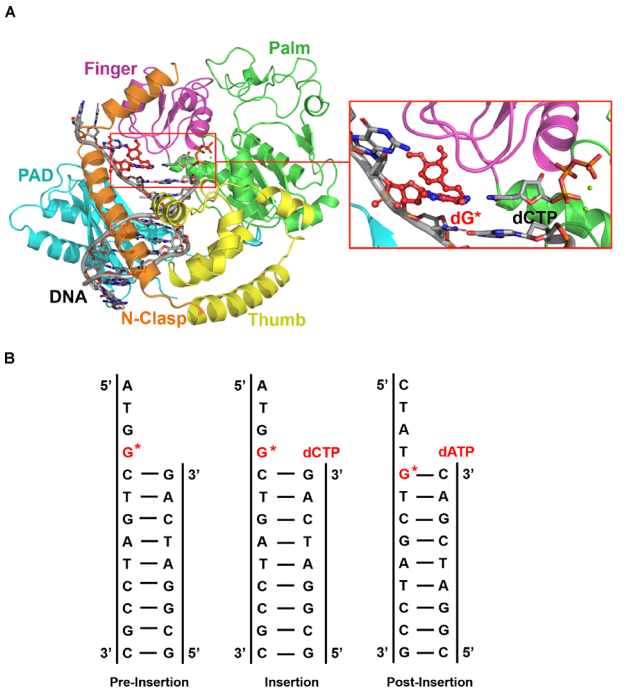


Figure 6. (A) Structure of the insertion stage hpolκ-DNA complex with the N^2-MEG-dG adduct used for MD simulations. The N-clasp, palm, finger, thumb, and polymerase-associated domain (PAD) of hpolκ are highlighted in orange, green, magenta, yellow, and cyan colors, respectively. The red-colored rectangle highlights the active site. The zoomed-in view shows the N^2-MEG-dG adduct (represented in red) and the incoming nucleotide (represented in gray) in the enzyme's active site; (B) DNA duplex systems used for MD simulation in the pre-insertion, insertion, and post-insertion stages.

fluctuations of the residues that constitute the modeled loop of the protein followed by a few residues in the N-clasp, palm, and thumb domains (Figure S8). However, these fluctuating residues did not interact directly with the DNA adduct. The residues constituting the enzyme's active site did not exhibit significant fluctuations, implying that the active site did not undergo substantial conformational changes throughout the simulation. In the case of DNA, the nucleotides present 5′ to the DNA lesion and the terminal base pairs at the 3′ end of the template strand showed significant fluctuation due to the terminal fraying. The DNA adduct showed comparatively larger fluctuation in the insertion stage than in the post-insertion stage (Figure S8).

***N^2*-MEG-dG Adduct Adapts Multiple Orientations in the Pre-Insertion Stage.** The main objective of simulating the pre-insertion stage was to gain insights into the structural preferences of the DNA lesion ahead of dNTP insertion. The analysis of the dihedral angle ξ (Figure 7A) of the adduct in the pre-insertion stage complex reveals that the adduct adapts multiple orientations in the active site, some of which allow dNTP to pair with the lesion in the insertion stage and some orientations, which obstruct the pairing. However, observations made in the simulation of the insertion stage revealed that although the adduct in its pre-insertion stage adapts orientations that blocked the incorporation of the incoming nucleotide, the adduct reoriented itself during the insertion stage to allow dNTP to pair with the damaged base (Figure 7C,D). Hydrogen bonds and non-covalent interactions between the adduct and the amino acid residues of the

https://doi.org/10.1021/acs.biochem.2c00663
Biochemistry XXXX, XXX, XXX−XXX

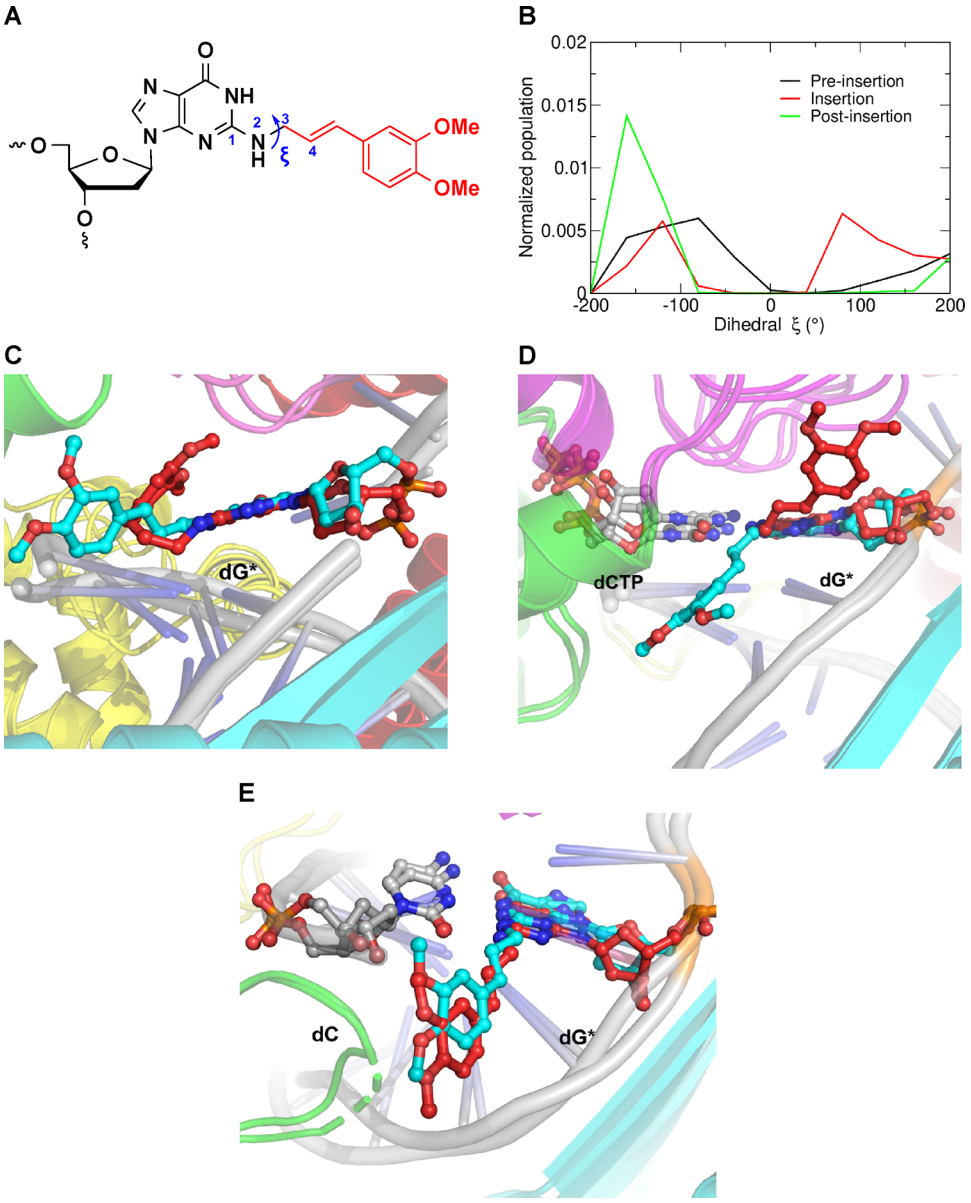


Figure 7. (A) Representative structure showing the dihedral angle ξ described by the atoms C2, N2, C12, and C14 labeled 1−4, respectively, in the N^2-MEG-dG lesion. (B) The plot showing the population of dihedral angle ξ representing the orientation of the lesion relative to the adduct. (C) Superimposed images of the ensembles dictating the orientation of the N^2-MEG-dG adduct in the active site of the pre-insertion complex, (D) insertion complex, and (E) post-insertion complex. The carbon atoms of cluster 1 are represented in red, and those of cluster 2 are represented in cyan, oxygen atoms are in red, nitrogen atoms are in blue, and phosphate atoms are in orange. dG* represents the adduct.

polymerase influence the configuration of the adduct in the active site.

Snapshots of the MD trajectory of the pre-insertion stage complex were clustered into five ensembles using the hierarchical agglomerative clustering method. The representative structures of major ensembles were superimposed to observe the different orientations occupied by the N^2-MEG-dG adduct in the active site (Figure 7B). The superimposed structures of the N^2-MEG-dG adduct in the pre-insertion stage revealed that the adduct in the major ensemble orients itself in a way that it might restrict the entry of the incoming nucleotide, stalling the insertion stage (Figure 7C,D).

Structural and Energetic Parameters Support the Successful Incorporation of Incoming Nucleotide across the N^2-MEG-dG Adduct. Specific structural parameters must be satisfied for successful replication past the DNA lesions.[60] These factors include an optimum C1′−C1′ distance (10.8 Å), Watson−Crick hydrogen bonds between the damaged base and the incoming nucleotide, proper orientation of the incoming nucleotide in the active site, and optimum distance between the catalytic Mg^{2+} ions (3.5−4.4 Å). The appropriate orientation of the incoming nucleotide with respect to the 3′ terminus of the primer strand is defined by the values of the reaction distance [distance between Pα of the incoming nucleotide and O3′ of the terminal primer base (3.5 Å)] and the attack angle [angle between O3′ (primer base), Pα (dNTP), and O$\alpha\beta$ (dNTP) (150−180°)]. The Watson−Crick hydrogen bonds should be consistent throughout the simulation 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).[61,62] All these criteria collectively determine the efficiency of replication.

https://doi.org/10.1021/acs.biochem.2c00663
Biochemistry XXXX, XXX, XXX−XXX

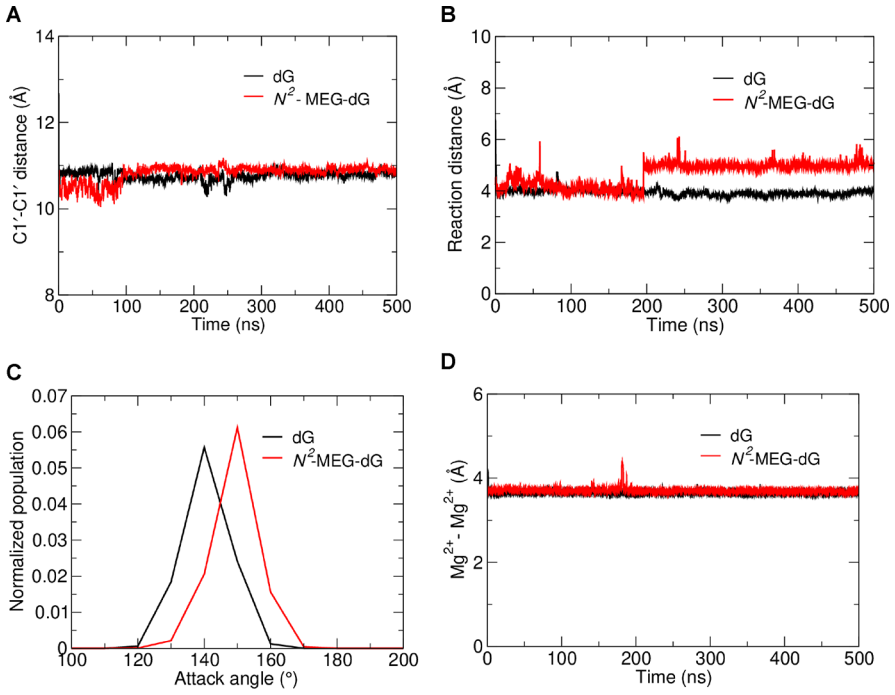


Figure 8. (A) Graph depicting the variation of the distance between the C1′ atom of N^2-dG/N^2-MEG-dG nucleotide and the C1′ atom of the incoming nucleotide over the total simulation time (500 ns); (B) graph depicting the variation of the reaction distance with respect to the simulation time; (C) histogram showing the probability distribution of the attack angle of the incoming nucleotide; (D) graph depicting the variation of the Mg^{2+}−Mg^{2+} distance. The running average of the distances is represented here.

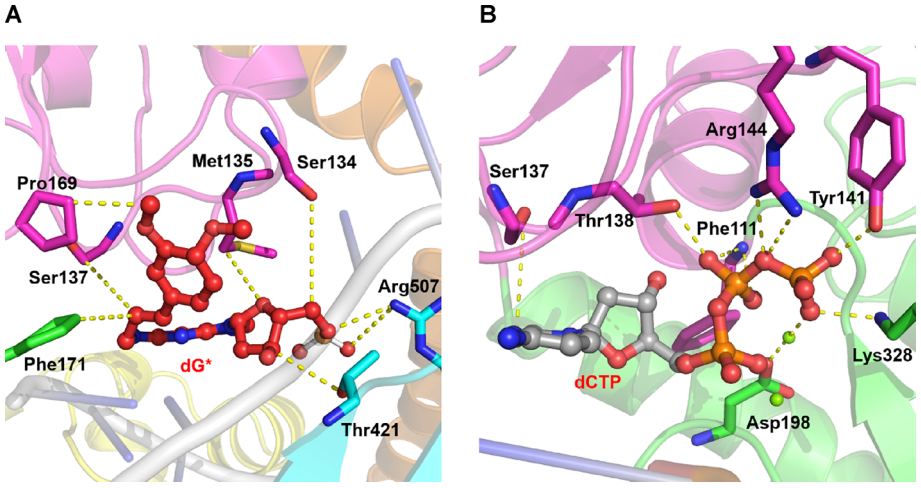


Figure 9. (A) Interaction of the N^2-MEG-dG adduct (dG*) with various residues surrounding in the enzyme's active site. The interactions are mainly with residues in the finger, PAD, and palm domains; (B) various interactions that position the incoming nucleotide (dCTP) in the active site for efficient bypass of the lesion. The incoming nucleotide shows multiple interactions with residues in palm and finger domains.

The insertion stage is characterized by the presence of the incoming nucleotide (dCTP) in the active site. The incoming nucleotide strongly interacts with the DNA lesion via hydrogen bonds (Table S1). The three hydrogen bonds combined had a total interaction strength of −22.6 ± 1.7 kcal/mol, which is slightly less than the reported value (−24 to −27 kcal/mol). The C1′−C1′ distance between the adduct and the incoming nucleotide is around 10.8 Å (Figure 8A). The average value of the reaction distance was 4.66 ± 0.5 Å (Figure 8B), which is slightly higher than the reported value of 3.5 Å, and that of the attack angle was 144.2 ± 6.2° (Figure 8C), which lies close to the acceptable range (150−180°). The average distance between the catalytic Mg^{2+} and the nucleotide-binding Mg^{2+} was 3.7 ± 0.1 Å (Figure 8D), within the optimum range (3.5−4.4 Å). The values of the structural parameters observed for the N^2-dG unmodified complex were close to those observed for the N^2-MEG-dG complex (Figure 8). All these factors point toward a favorable replication where the DNA polymerase enzyme successfully pairs a nucleotide to the DNA lesion. The RMSD, RMSF (calculated using CPPTRAJ), and local base pair parameter (calculated by X3DNA)[63] values of the unmodified and N^2-MEG-dG modified duplexes showed similar values, specifically at the site of modification (Figure S9A−D). These results are similar to the report by Yang and co-workers, where they found minimal distortion in the EG-modified DNA duplex.[22]

https://doi.org/10.1021/acs.biochem.2c00663
Biochemistry XXXX, XXX, XXX−XXX

Similar to the pre-insertion stage, the N^2-MEG-dG adduct in the insertion stage shows hydrogen bonding and other non-covalent interactions with the amino acid residues of hpolκ (Figure 9A). The backbone oxygen atoms of the damaged nucleotide were observed to show electrostatic interactions with the polar and charged amino acids present in the PAD domain. The oxygen atoms present in the methoxy groups of the lesion were found to exhibit weak hydrogen bonding interactions with SER-132 and ARG-175 residues. The incoming nucleotide was held opposite to the DNA lesion by hydrogen bonding and non-covalent interactions with the amino acid residues PHE-111, SER-137, THR-138, TYR-138, TYR-141, ARG-144, ASP-198, and LYS-328 in the binding pocket (Figure 9B). These interactions stabilize the lesion and incoming nucleotide in the active site and facilitate favorable replication. The average non-covalent interaction energies between the N^2-MEG-dG adduct and various amino acid residues enveloping it in the active site were determined by employing the MM/GBSA method (Table S2). MET-135, SER-137, and PRO-169 showed van der Waals interactions, while ASP-172, ARG-175, and ARG-420 showed electrostatic interactions with the DNA lesion. The unmodified dG complex however exhibited strong interactions only with MET-135 amino acid. The major cluster of the insertion complex revealed that the adduct adapts orientations that facilitate pairing dCTP with the damaged base (Figure 7D).

The visual inspection of the major cluster of the N^2-dG and the N^2-MEG-dG complexes was carried out to understand the interactions responsible for the higher efficient bypass of the adduct compared with the unmodified complex. The N^2-MEG-dG adduct is present in a cavity surrounded by amino acids TYR-112, VAL-130, SER-132, SER-134, MET-135, PRO-169, and PHE-171, which creates a predominantly hydrophobic pocket for the adduct occupancy (Figure S10 and Figure 9A). The comparison of the MM/GBSA energy revealed that additional non-covalent interactions are present between the adduct and the amino acids in the hydrophobic cavity. Though a similar cavity is present in the unmodified complex, there is no interaction with the pocket due to the absence of the lesion (Figure S10). This prevents the additional stabilization due to non-covalent interactions seen in the N^2-MEG-dG complex. This additional stabilization can be attributed for the efficient bypass of the N^2-MEG-dG adduct compared to the unmodified dG by hpolκ.

hpolκ Can Successfully Extend the Replication Process Past the DNA Lesion. In the post-insertion (extension) stage, the base 5′ to the lesion (dT) was observed to pair with the incoming nucleotide (dATP) by two hydrogen bonds with percentage occupancies of 94 and 99% (Tables S3 and S4) and an average interaction energy of −9.2 ± 1.2 kcal/mol. The N^2-MEG-dG:dC base pair formed during the insertion stage maintained Watson−Crick pairing in the post-insertion stage with an interaction energy of −23.7 ± 2 kcal/mol. The C1′−C1′ distance between the incoming nucleotide and dT was around 11.1 Å. The average reaction distance between the Pα of the incoming nucleotide and O3′ of the terminal primer base was 4.89 ± 1 Å, and the average value of the attack angle was 139.7 ± 6.7°. Even though the structural parameters that dictate the orientation of the incoming nucleotide deviate slightly from the values reported in the literature,[60] the incoming nucleotide was still found to form Watson−Crick base pairing with the template base, indicating that hpolκ is able to bypass the N^2-MEG-dG adduct and extend the replication process upstream. Similar to the insertion stage, the cluster of the post-insertion stage revealed that the adduct orients itself toward the minor groove, allowing hpolκ to extend the replication process past the DNA lesion (Figure 7E). From the dihedral analysis (Figure 7B), it is clear that the adduct exhibits the same orientation in the pre-insertion and post-insertion stages, unlike that in the incorporation stage.

CONCLUSIONS

Plant metabolites MEG and EG form carcinogenic N^2-dG DNA adducts. We have developed robust protocols to synthesize lesion-bearing DNAs to study their mutagenicity during replication. Primer extension studies using human TLS polymerases hpolκ and hpolη showed that TLS across these adducts is significantly error-free, and hpolκ is more efficient in bypassing the adducts. Molecular modeling and dynamics studies revealed the atomistic features, which satisfy the correct incorporation of dCTP across the adducted base. These findings show that hpolκ and hpolη bypass the MEG and EG adducts with high fidelity in humans. However, the carcinogenicity observed inside the cells could be due to the involvement of other low-fidelity polymerases like Rev1 or Polζ in bypassing these adducts.[17,64] Further, studies are required to substantiate these aspects.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.biochem.2c00663.

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

Accession Codes

hpolκ: UniProt accession code: Q9UBT6; hpolη: UniProt accession code: Q9Y253.

AUTHOR INFORMATION

Corresponding Authors

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

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*

https://doi.org/10.1021/acs.biochem.2c00663
Biochemistry XXXX, XXX, XXX−XXX

Tanvi Aggarwal − *Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Mumbai 400076, India*
Soumyadeep Mandal − *Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Mumbai 400076, India*
Siddharam Shivappa Bagale − *Department of Chemistry, Indian Institute of Technology Bombay, Mumbai 400076, India*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.biochem.2c00663

Author Contributions

The manuscript was completed through the contributions of all authors. All authors have approved the final version of the manuscript.

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/2021/001992), the Government of India to P.I.P. Research in the KK lab is supported by grants from the Department of Biotechnology, DBT [BT/PR35928/BRB/10/1841/2019] and the Board of Research in Nuclear Sciences, BRNS [58/14/11/2020-BRNS/37188]. We are thankful to Prof. F. Peter Guengrich (Vanderbilt University) for his generous gift of plasmid constructs pBG101-hpolκ and pET28a-hpolη. We also thank Dr. Pratibha P. Ghodke for the initial association with this project and discussions. We are thankful to IRCC-IIT Bombay for providing access to the central MALDI facility and the IIT Bombay-Spacetime HPC facility. We are also grateful to the Institute of Eminence (IOE)/RIFC-supported central facility (Department of Chemistry) for providing access to the DNA synthesizer and DST-FIST (SR/FST/CS-II/2017/37) support for NMR and HRMS facilities. P.U.D. and S.M. thank the Council of Scientific and Industrial Research (CSIR), S.B.L. thanks IRCC-IIT Bombay, S.S. thanks Prime Minister's Research Fellowship (PMRF), and T.A. and S.S.B. thank IIT Bombay for the Ph.D. fellowships.

■ REFERENCES

(1) Andersson, D. I.; Koskiniemi, S.; Hughes, D. Biological roles of translesion synthesis DNA polymerases in eubacteria. *Mol. Biol.* **2010**, *77*, 540−548.
(2) Lord, C. J.; Ashworth, A. The DNA damage response and cancer therapy. *Nature* **2012**, *481*, 287−294.
(3) Barnes, J. L.; Zubair, M.; John, K.; Poirier, M. C.; Martin, F. L. Carcinogens and DNA damage. *Biochem. Soc. Trans.* **2018**, *46*, 1213−1224.
(4) Dipple, A. DNA adducts of chemical carcinogens. *Carcinogenesis* **1995**, *16*, 437−441.
(5) Broyde, S.; Wang, L.; Zhang, L.; Rechkoblit, O.; Geacintov, N. E.; Patel, D. J. DNA adduct structure−function relationships: comparing solution with polymerase structures. *Chem. Res. Toxicol.* **2008**, *21*, 45−52.
(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*, 55−82.
(7) Chatterjee, N.; Walker, G. C. Mechanisms of DNA damage, repair, and mutagenesis. *Environ. Mol. Mutagen.* **2017**, *58*, 235−263.
(8) Totsuka, Y.; Watanabe, M.; Lin, Y. New horizons of DNA adductome for exploring environmental causes of cancer. *Cancer Sci.* **2021**, *112*, 7−15.
(9) Goodman, M. F.; Woodgate, R. Translesion DNA polymerases. *Cold Spring Harbor Perspect. Biol.* **2013**, *5*, a010363.
(10) 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.
(11) Vaisman, A.; Woodgate, R. Translesion DNA polymerases in eukaryotes: what makes them tick? *Crit. Rev. Biochem. Mol. Biol.* **2017**, *52*, 274−303.
(12) Prakash, S.; Johnson, R. E.; Prakash, L. Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. *Annu. Rev. Biochem.* **2005**, *74*, 317−353.
(13) Stern, H. R.; Sefcikova, J.; Chaparro, V. E.; Beuning, P. J. Mammalian DNA polymerase kappa activity and specificity. *Molecules* **2019**, *24*, 2805.
(14) Yagi, T.; Fujikawa, Y.; Sawai, T.; Takamura-Enya, T.; Ito-Harashima, S.; Kawanishi, M. Error-prone and error-free translesion DNA synthesis over site-specifically created DNA adducts of aryl hydrocarbons (3-nitrobenzanthrone and 4-aminobiphenyl). *Toxicol. Res.* **2017**, *33*, 265−272.
(15) Zhang, Y.; Yuan, F.; Wu, X.; Wang, M.; Rechkoblit, O.; Taylor, J.-S.; Geacintov, N. E.; Wang, Z. Error-free and error-prone lesion bypass by human DNA polymerase κ in vitro. *Nucleic Acid Res.* **2000**, *28*, 4138−4146.
(16) Avkin, S.; Goldsmith, M.; Velasco-Miguel, S.; Geacintov, N.; Friedberg, E. C.; Livneh, Z. Quantitative analysis of translesion DNA synthesis across a benzo [a] pyrene-guanine adduct in mammalian cells: the role of DNA polymerase κ. *J. Biol. Chem.* **2004**, *279*, 53298−53305.
(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*, 1549−1559.
(18) Herrmann, K.; Engst, W.; Meinl, W.; Florian, S.; Cartus, A. T.; Schrenk, D.; Appel, K. E.; Nolden, T.; Himmelbauer, H.; Glatt, H. Formation of hepatic DNA adducts by methyleugenol in mouse models: drastic decrease by Sult1a1 knockout and strong increase by transgenic human SULT1A1/2. *Carcinogenesis* **2014**, *35*, 935−941.
(19) Dang, H. N. P.; Quirino, J. P. Analytical separation of carcinogenic and genotoxic alkenylbenzenes in foods and related products. *Toxins* **2021**, *13*, 387.
(20) Götz, M. E.; Sachse, B.; Schäfer, B.; Eisenreich, A. Myristicin and Elemicin: Potentially Toxic Alkenylbenzenes in Food. *Foods* **2022**, *11*, 1988.
(21) Eisenreich, A.; Götz, M. E.; Sachse, B.; Monien, B. H.; Herrmann, K.; Schäfer, B. Alkenylbenzenes in Foods: Aspects Impeding the Evaluation of Adverse Health Effects. *Foods* **2021**, *10*, 2139.
(22) Yang, S.; Diem, M.; Liu, J. D. H.; Wesseling, S.; Vervoort, J.; Oostenbrink, C.; Rietjens, I. M. C. M. Cellular levels and molecular dynamics simulations of estragole DNA adducts point at inefficient repair resulting from limited distortion of the double-stranded DNA helix. *Arch. Toxicol.* **2020**, *94*, 1349−1365.
(23) International Agency for Research on Cancer Overall Evaluations of Carcinogenicity: An Updating of IARC Monographs Volumes 1 to 42. In *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*; 1987, *7*, 1−440.
(24) National Toxicology Program. NTP Toxicology and Carcinogenesis Studies of Methyleugenol (CAS NO. 93-15-2) in F344/N Rats and B6C3F1 Mice (Gavage Studies). *Natl. Toxicol. Program Tech. Rep. Ser.* **2000**, *491*, 1−412.
(25) Benford, D.; Bolger, P. M.; Carthew, P.; Coulet, M.; DiNovi, M.; Leblanc, J.-C.; Renwick, A. G.; Setzer, W.; Schlatter, J.; Smith, B.; Slob, W.; Williams, G.; Wildemann, T. Application of the Margin of Exposure (MOE) approach to substances in food that are genotoxic and carcinogenic. *Food Chem. Toxicol.* **2010**, *48*, S2−S24.

(26) Tremmel, R.; Herrmann, K.; Engst, W.; Meinl, W.; Klein, K.; Glatt, H.; Zanger, U. M. Methyleugenol DNA adducts in human liver are associated with SULT1A1 copy number variations and expression levels. *Arch. Toxicol.* **2017**, *91*, 3329−3339.
(27) Herrmann, K.; Schumacher, F.; Engst, W.; Appel, K. E.; Klein, K.; Zanger, U. M.; Glatt, H. Abundance of DNA adducts of methyleugenol, a rodent hepatocarcinogen, in human liver samples. *Carcinogenesis* **2013**, *34*, 1025−1030.
(28) Paini, A.; Punt, A.; Scholz, G.; Gremaud, E.; Spenkelink, B.; Alink, G.; Schilter, B.; van Bladeren, P. J.; Rietjens, I. M. C. M. In vivo validation of DNA adduct formation by estragole in rats predicted by physiologically based biodynamic modelling. *Mutagenesis* **2012**, *27*, 653−663.
(29) Monien, B. H.; Schumacher, F.; Herrmann, K.; Glatt, H.; Turesky, R. J.; Chesné, C. Simultaneous Detection of Multiple DNA Adducts in Human Lung Samples by Isotope-Dilution UPLC-MS/MS. *Anal. Chem.* **2015**, *87*, 641−648.
(30) Yang, S.; Wesseling, S.; Rietjens, I. M. C. M. Estragole DNA adduct accumulation in human liver HepaRG cells upon repeated in vitro exposure. *Toxicol. Lett.* **2021**, *337*, 1−6.
(31) 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 $_{NaHCO3}$. *J. Org. Chem.* **2007**, *72*, 5244−5259.
(32) Li, H.; Chen, H.; Zhou, Y.; Huang, J.; Yi, J.; Zhao, H.; Wang, W.; Jing, L. Selective Synthesis of Z-Cinnamyl Ethers and Cinnamyl Alcohols through Visible Light-Promoted Photocatalytic E to Z Isomerization. *Chem. − Asian J.* **2020**, *15*, 555−559.
(33) 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*, 2128−2138.
(34) Zhao, L.; Pence, M. G.; Eoff, R. L.; Yuan, S.; Fercu, C. A.; Guengerich, F. P. Elucidation of kinetic mechanisms of human translesion DNA polymerase κ using tryptophan mutants. *FEBS J.* **2014**, *281*, 4394−4410.
(35) Biertümpfel, C.; Zhao, Y.; Kondo, Y.; Ramón-Maiques, S.; Gregory, M.; Lee, J. Y.; Masutani, C.; Lehmann, A. R.; Hanaoka, F.; Yang, W. Structure and mechanism of human DNA polymerase η. *Nature* **2010**, *465*, 1044−1048.
(36) Ghodke, P. P.; Guengerich, F. P. DNA polymerases η and κ bypass N2-guanine-O6-alkylguanine DNA alkyltransferase cross-linked DNA-peptides. *J. Biol. Chem.* **2021**, *297*, No. 101124.
(37) O'Flaherty, D. K.; Guengerich, F. P. Steady-State Kinetic Analysis of DNA Polymerase Single-Nucleotide Incorporation Products. *Curr. Protoc. Nucleic Acid Chem.* **2014**, *59*, 7.21.1−7.2113.
(38) Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C. SWISS-MODEL: An Automated Protein Homology-Modeling Server. *Nucleic Acids Res.* **2003**, *31*, 3381−3385.
(39) 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.
(40) 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*, 7821−7839.
(41) 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*, 10269−10280.
(42) 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*, 3696−3713.
(43) 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*, 55−58.
(44) 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.
(45) 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.
(46) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E., III; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K. *AMBER 2018*; *2018*. Univ. California, San Fr, 2018.
(47) Roe, D. R.; Cheatham, T. E., III PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084−3095.
(48) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(49) Turq, P.; Lantelme, F.; Friedman, H. L. Brownian Dynamics: Its Application to Ionic Solutions. *J. Chem. Phys.* **1977**, *66*, 3039−3044.
(50) 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*, 889−897.
(51) Sagar Reddy, G. V.; Rao, G. V.; 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*, 2233−2237.
(52) Roy, S.; Basak, A. Exploring the scope of Bergman Cyclization mediated cascade reaction of alkenyl enediynes: synthesis of [5] helicene and amino acid appended [4] helicenes. *Tetrahedron* **2013**, *69*, 2184−2192.
(53) Bonala, R. R.; Shishkina, I. G.; Johnson, F. Synthesis of biologically active N2-amine adducts of 2′-deoxyguanosine. *Tetrahedron Lett.* **2000**, *41*, 7281−7284.
(54) Lakshman, M. K.; Hilmer, J. H.; Martin, J. Q.; Keeler, J. C.; Dinh, Y. Q. V.; Ngassa, F. N.; Russon, L. M. Palladium Catalysis for the Synthesis of Hydrophobic C-6 and C-2 Aryl 2′-Deoxynucleosides. Comparison of C− C versus C− N Bond Formation as well as C-6 versus C-2 Reactivity. *J. Am. Chem. Soc.* **2001**, *123*, 7779−7787.
(55) Ghodke, P. P.; Pradeepkumar, P. I. Site-Specific N^2-dG DNA Adducts: Formation, Synthesis, and TLS Polymerase-Mediated Bypass. *Eur. J. Org. Chem.* **2020**, *2020*, 6831−6844.
(56) Lee, Y. S.; Gregory, M. T.; Yang, W. Human Pol ζ purified with accessory subunits is active in translesion DNA synthesis and complements Pol η in cisplatin bypass. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 2954−2959.
(57) 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*, 21062−21072.
(58) Choi, J. Y.; Guengerich, F. P. Adduct size limits efficient and error-free bypass across bulky N2-guanine DNA lesions by human DNA polymerase η. *J. Mol. Biol.* **2005**, *352*, 72−90.
(59) 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*, 2023−2032.
(60) 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*, 2573−2583.

https://doi.org/10.1021/acs.biochem.2c00663
Biochemistry XXXX, XXX, XXX−XXX

(61) Šponer, J.; Leszczynski, J.; Hobza, P. Hydrogen Bonding and Stacking of DNA Bases: A Review of Quantum-Chemical Ab Initio Studies. *J. Biomol. Struct. Dyn.* **1996**, *14*, 117−135.
(62) Šponer, J.; Leszczynski, J.; Hobza, P. Electronic Properties, Hydrogen Bonding, Stacking, and Cation Binding of DNA and RNA Bases. *Biopolymers* **2001**, *61*, 3−31.
(63) Lu, X. J.; Olson, W. K. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* **2003**, *31*, 5108−5121.
(64) Bose, A.; Surugihalli, C.; Pande, P.; Champeil, E.; Basu, A. K. Comparative Error-Free and Error-Prone Translesion Synthesis of N^2-2′-Deoxyguanosine Adducts Formed by Mitomycin C and Its Metabolite, 2,7-Diaminomitosene, in Human Cells. *Chem. Res. Toxicol.* **2016**, *29*, 933−939.

NOTE ADDED AFTER ASAP PUBLICATION

This article published ASAP on July 24, 2023. The Table of Contents/Abstract graphic has been updated and the corrected version was reposted on July 25, 2023.

https://doi.org/10.1021/acs.biochem.2c00663
Biochemistry XXXX, XXX, XXX−XXX

Human TLS polymerases polκ and polη perform error-free replication across N^2-dG methyleugenol and estragole DNA adducts

Priyanka U. Deshmukh, [†] Shailesh B. Lad, [§] Akhil Sudarsan, [†] Sruthi Sudhakar, [†] Tanvi Aggarwal, [§] Soumyadeep Mandal, [§] Siddharam Shivappa Bagale,[†] 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 kirankondabagil@iitb.ac.in or pradeep@chem.iitb.ac.in

Figure S1	SDS-PAGE of purified human polymerases η and κ.....	Page S1
Figure S2	Full length extension assay using hpolκ and hpolη……	Page S2
Figure S3	Steady-state kinetic analysis of dCTP insertion by hpolκ and hpolη with unmodified template (dG).....	Page S3
Figure S4	Comparison of kinetic parameters of hpolκ for bypass of N^2-EG-dG and N^2-MEG-dG DNA.....	Page S4
Figure S5	Comparison of kinetic parameters of hpolη for bypass of N^2-EG-dG and N^2-MEG-dG DNA.....	Page S5
Figure S6	Cartesian coordinates and RESP charges calculated for N^2-MEG-dG	Page S6-S10
Figure S7	Root Mean Square Deviations (RMSD) of the DNA and hpolκ residues.....	Page S11
Figure S8	Root Mean Square Fluctuations (RMSF) of the DNA and hpolκ residues.....	Page S12-S13
Figure S9	Comparison of MD parameters for MEG-modified and unmodified duplex DNAs……	Page S14
Figure S10	Orientation of N^2-dG and N^2-MEG-dG in the predominantly hydrophobic pocket of hpolκ.....	Page S15
Table S1	Percentage occupancies of H-bonds between the incoming nucleotide and the lesion in the insertion stage ……..	Page S16
Table S2	Pairwise interaction energies of the N^2-dG and N^2-MEG-dG with the surrounding amino acids in the active site of the enzyme.....	Page S17
Table S3	Percentage occupancies of H-bonds between the incoming nucleotide and the template base in the post-insertion stage complex.....	Page S17

Table S4 Percentage occupancies of H-bonds between the previously replicated N^2-MEG-dG and dC base pair in the post-insertion stage complex..... Page S18

Table S5 MALDI-TOF mass data of the DNA sequences..... Page S18

^{1}H NMR &^{13}C NMR spectra of compound **2a**…… Page S19-S20

^{1}H NMR &^{13}C NMR spectra of compound **3a**…… Page S21-S22

^{1}H NMR &^{13}C NMR spectra of compound **4a**……. Page S23-S24

^{1}H NMR &^{13}C NMR spectra of compound **5a**……... Page S25-S26

^{1}H NMR &^{13}C NMR spectra of compound **7a**…….. Page S27-S28

^{1}H NMR &^{13}C NMR spectra of compound **8a**……... Page S29-S30

^{1}H NMR &^{13}C NMR spectra of compound **9a**……... Page S31-S32

^{31}P NMR spectra of compound **10a**……... Page S33

^{1}H NMR &^{13}C NMR spectra of compound **2b**……... Page S34-S35

^{1}H NMR &^{13}C NMR spectra of compound **3b**……... Page S36-S37

^{1}H NMR &^{13}C NMR spectra of compound **4b**……... Page S38-S39

^{1}H NMR &^{13}C NMR spectra of compound **5b**……... Page S40-S41

^{1}H NMR &^{13}C NMR spectra of compound **7b**……... Page S42-S43

^{1}H NMR &^{13}C NMR spectra of compound **8b**……... Page S44-S45

^{1}H NMR &^{13}C NMR spectra of compound **9b**……... Page S46-S47

^{31}P NMR spectra of compound **10b**……... Page S48

MALDI Spectrum of **T1**…… Page S49

MALDI Spectrum of **T2**…… Page S50

SDS-PAGE of Purified human polymerases η and κ

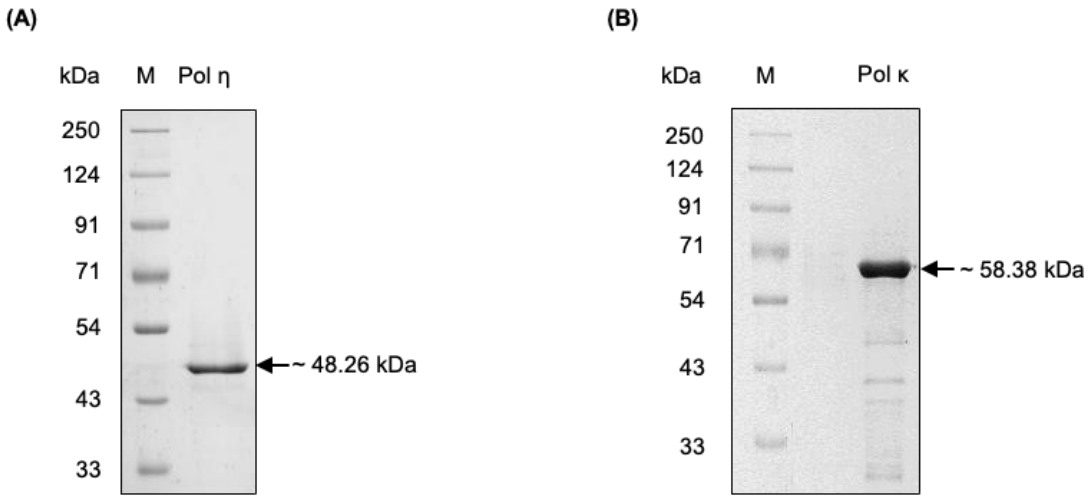


Figure. S1. (A) SDS-PAGE showing purified recombinant human polymerase η. The protein was expressed in *E. coli* BL21(DE3) RIPL cells and purified using Ni-NTA affinity chromatography followed by removal of histidine tag and gel filtration chromatography using 16/600 Superdex 200 pg column. (B) SDS-PAGE showing purified human polymerase κ. The protein was expressed similar to polη and purified using Ni-NTA chromatography and gel filtration chromatography using 16/600 Superdex 200 pg column. Tags on both proteins were cleaved using PreScission protease (Cytiva).

Full length extension assay using hpolκ and hpolη

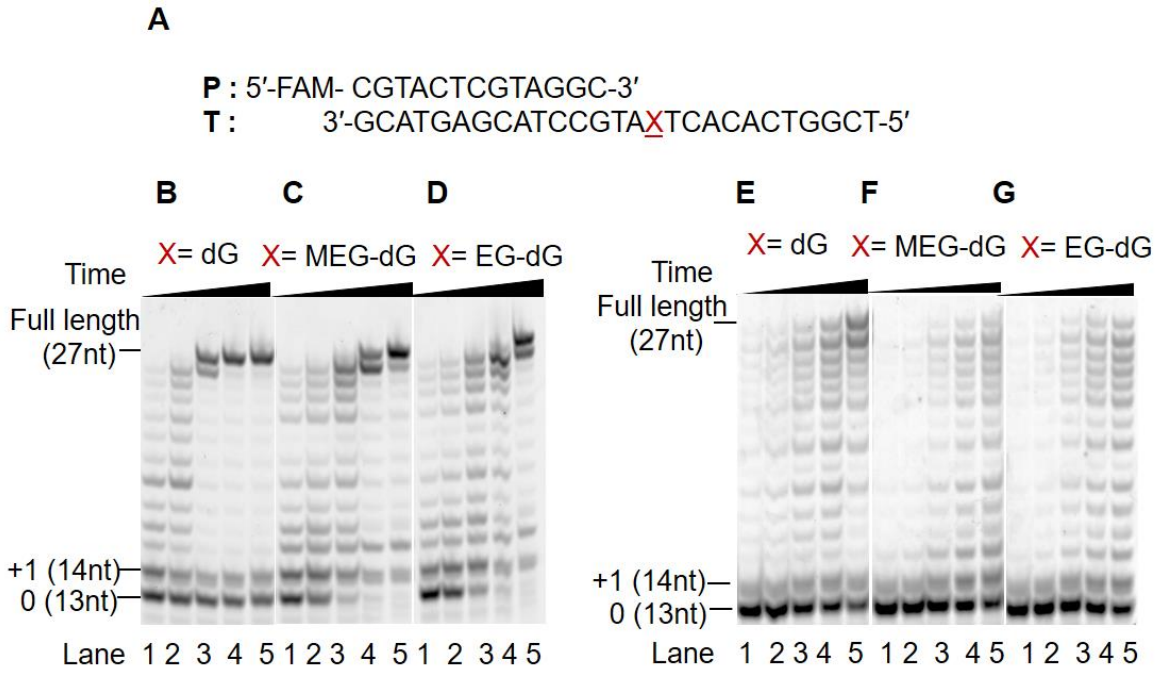


Figure S2. Full-length extension assay with hpolκ and hpolη in the presence of a mixture of dNTPs. (A) 13mer primer and 27mer template DNA sequence, where (B) X= dG; (C) X= MEG-dG; (D) X= EG-dG in the presence of 5 nM hpolκ and (E) X= dG; (F) X= MEG-dG; and (G) X= EG-dG in the presence of 5 nM hpolη. All reactions were performed at 37 °C for 5, 10, 30, 60 and 120 min.

Steady state kinetic analysis of dCTP insertion by hpolκ and hpolη with unmodified template

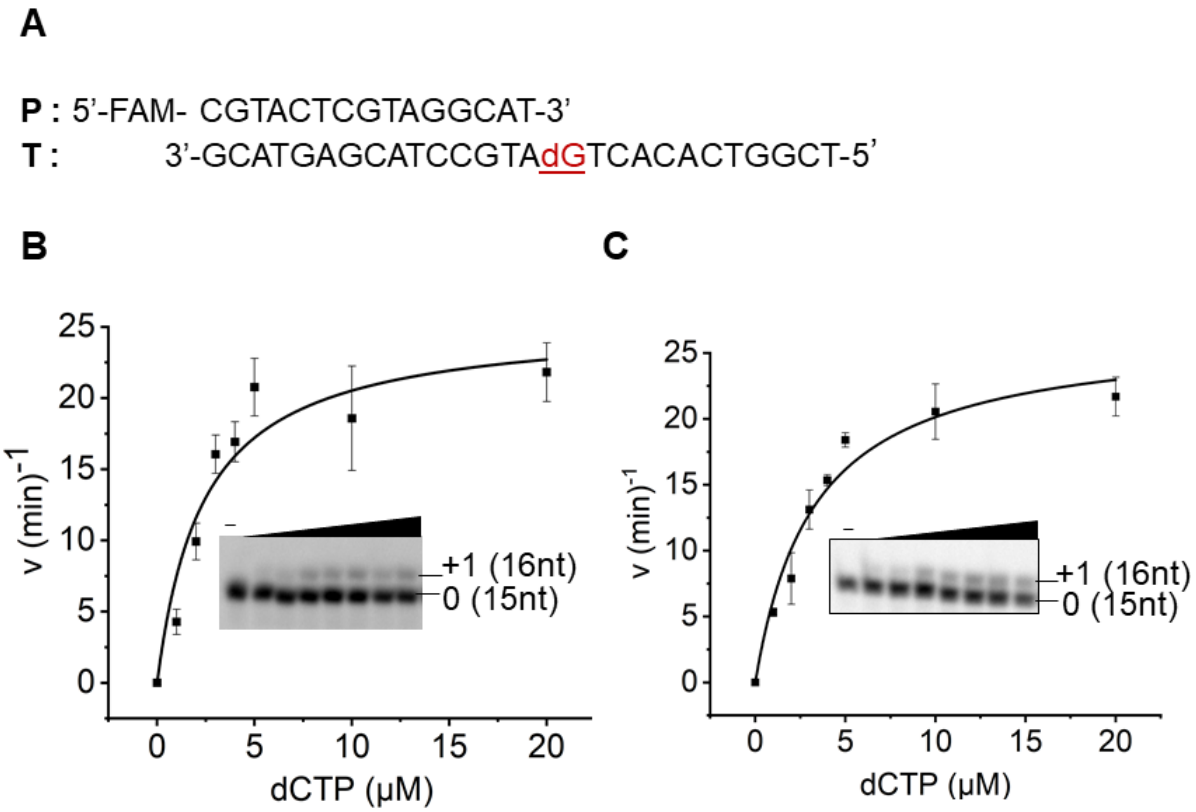


Figure S3. (A) Sequences of 15mer primer and 27mer template DNAs used for the reactions. Reaction was performed in presence of: (B) 5 nM hpolκ; and (C) 4 nM hpolη. All reaction was performed at 37 °C using 0, 2, 3, 4, 5, 10, 15, 20, 30, and 40 μM of dCTP.

Comparison of kinetic parameters of hpolκ for bypass of N^2-EG-dG and N^2-MEG-dG DNA

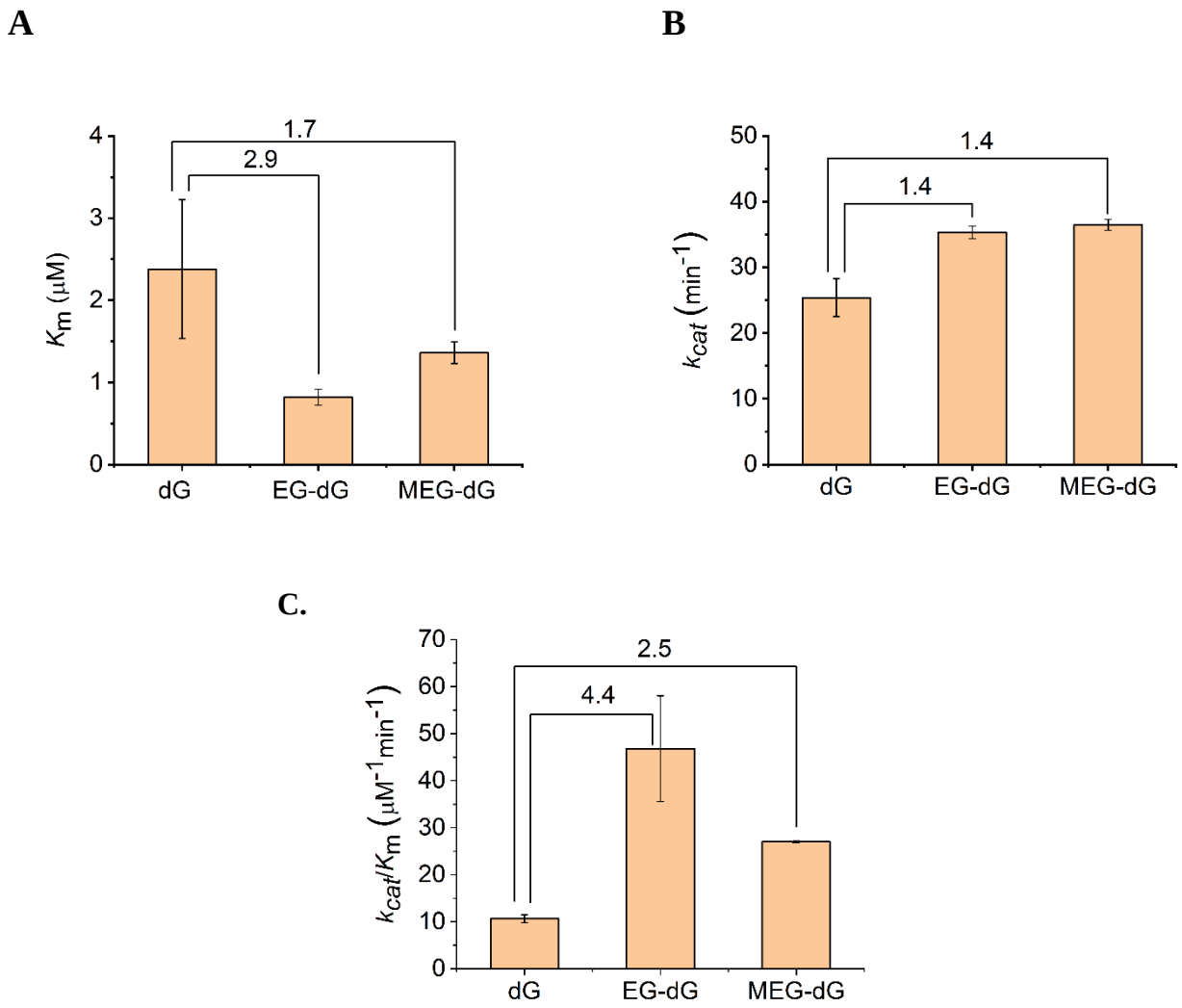


Figure S4. A comparison of (A) K_m; (B) k_{cat}; (C) k_{cat}/K_m values for hPolκ bypass of EG-dG and MEG- dG modified DNA templates with the unmodified template (dG). The numbers on top show the fold change in individual parameters compared to synthesis across unmodified DNA (dG). Error bars represents the standard deviation derived from two independent experiments.

Comparison of kinetic parameters of hPolη for bypass of N^2-EG-dG and N^2-MEG-dG DNA

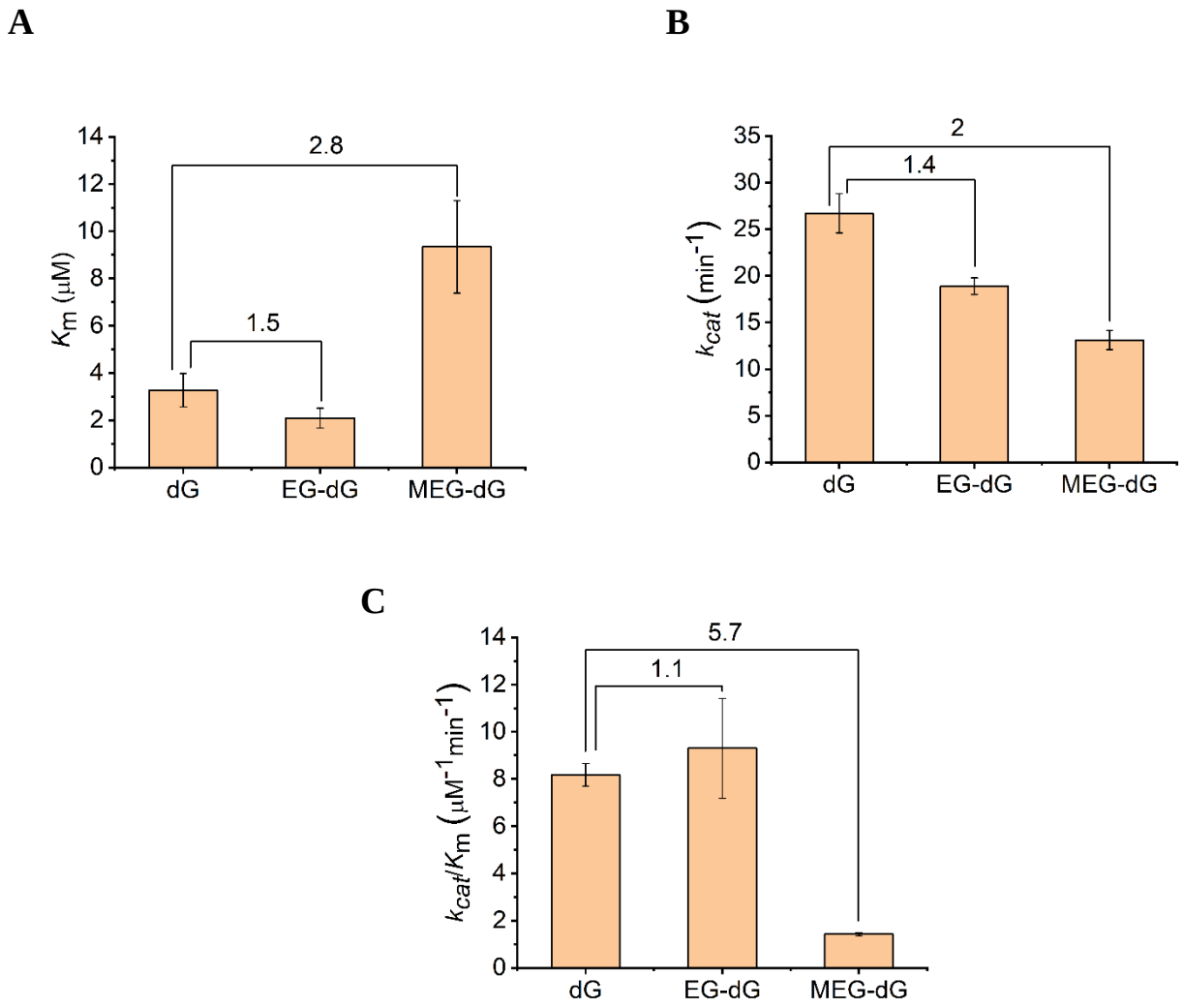
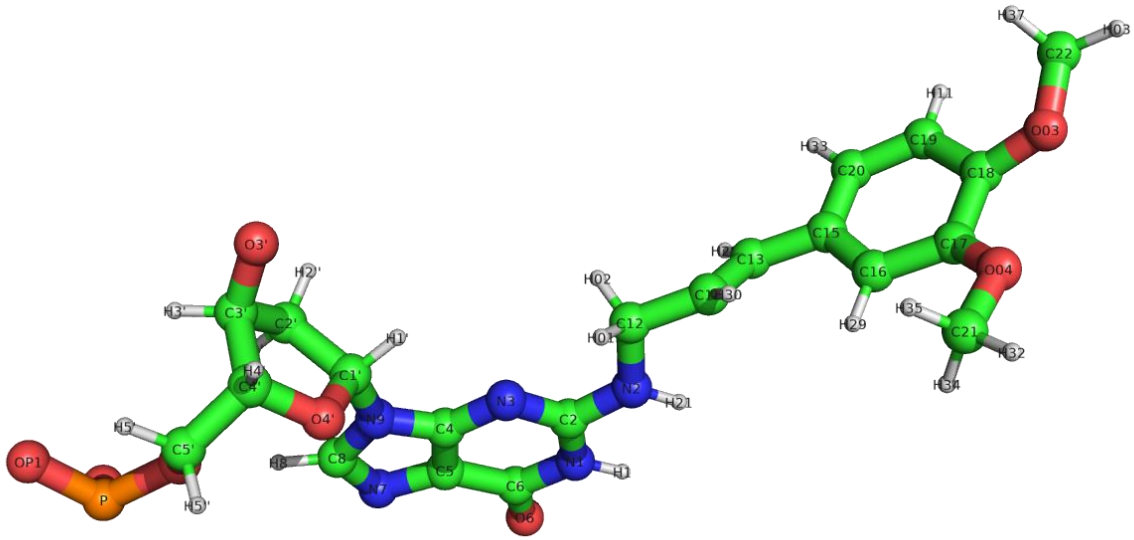


Figure S5. A comparison of (A) K_m, (B) k_{cat} and (C) k_{cat}/K_m values for hPolη bypass of EG-dG and MEG-dG modified DNA templates with the unmodified template (dG). The numbers on top show the fold change in individual parameters compared to synthesis across unmodified DNA (dG). Error bars represent the standard deviation derived from two independent experiments.

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



@<TRIPOS>MOLECULE

U0

58 61 1 0 1

SMALL

USER_CHARGES

@<TRIPOS>ATOM

1	O5'	-2.60838	-1.4277	-0.23197	OS	1	U0	-0.5008	0	****
2	P	-3.73702	-2.55653	0.152949	P	1	U0	1.2138	0	****
3	OP1	-3.90276	-3.46358	-0.99231	O2	1	U0	-0.7915	0	****
4	OP2	-4.83145	-1.8348	0.818645	O2	1	U0	-0.7915	0	****
5	O3'	1.403566	0	0	OS	1	U0	-0.5298	0	****
6	C6	-3.36735	5.336681	-3.46454	C	1	U0	0.5549	0	****
7	O6	-4.34688	5.997367	-3.63544	O	1	U0	-0.5468	0	****
8	N1	-2.18731	5.654167	-4.17618	NA	1	U0	-0.5552	0	****
9	H1	-2.29787	6.404631	-4.82419	H	1	U0	0.3386	0	****
10	C2	-0.99648	4.996723	-4.10349	CA	1	U0	0.6261	0	****
11	N2	-0.01433	5.460182	-4.91136	N2	1	U0	-0.5112	0	****
12	H21	-0.1042	6.397659	-5.23516	H	1	U0	0.3256	0	****

13	O4'	-1.32634	0.364034	-1.94034	OS	1	U0	-0.3342	0	****
14	C4	-1.89589	3.617027	-2.63233	CB	1	U0	0.1874	0	****
15	N3	-0.79503	3.97706	-3.33554	NC	1	U0	-0.5647	0	****
16	O03	6.437632	8.170598	-10.0753	OS	1	U0	-0.3414	0	****
17	C22	7.395882	9.16879	-9.87435	CT	1	U0	0.0078	0	****
18	H03	7.850977	9.345605	-10.837	H1	1	U0	0.0676	0	****
19	H36	6.942334	10.08861	-9.51965	H1	1	U0	0.0676	0	****
20	H37	8.159097	8.850358	-9.17173	H1	1	U0	0.0676	0	****
21	C5	-3.13291	4.196853	-2.62331	CB	1	U0	0.1522	0	****
22	N7	-3.97445	3.52741	-1.76215	NB	1	U0	-0.5484	0	****
23	O04	4.735629	6.316096	-10.6093	OS	1	U0	-0.3262	0	****
24	C21	3.843018	5.313206	-10.9969	CT	1	U0	0.0109	0	****
25	H32	4.000055	5.164664	-12.0545	H1	1	U0	0.0693	0	****
26	H34	2.812943	5.610664	-10.8282	H1	1	U0	0.0693	0	****
27	H35	4.038434	4.382662	-10.4737	H1	1	U0	0.0693	0	****
28	C8	-3.25965	2.589216	-1.27512	CK	1	U0	0.1023	0	****
29	H8	-3.59584	1.861305	-0.56706	H5	1	U0	0.1818	0	****
30	N9	-1.96439	2.582932	-1.75173	N*	1	U0	-0.008	0	****
31	C1'	-0.93184	1.625999	-1.47642	CT	1	U0	0.1063	0	****
32	H1'	-0.06079	1.968322	-2.01987	H2	1	U0	0.1229	0	****
33	C2'	-0.61727	1.399171	0	CT	1	U0	-0.0631	0	****
34	H2'	-1.53556	1.395094	0.573751	HC	1	U0	0.0476	0	****
35	H2"	0.046653	2.1455	0.416492	HC	1	U0	0.0476	0	****
36	C3'	0	0	0	CT	1	U0	0.0557	0	****
37	H3'	-0.28025	-0.55416	0.881564	H1	1	U0	0.2005	0	****
38	C4'	-0.58861	-0.64633	-1.27777	CT	1	U0	0.0296	0	****
39	H4'	0.229023	-0.97579	-1.91398	H1	1	U0	0.0971	0	****
40	C5'	-1.52713	-1.81055	-1.0227	CT	1	U0	-0.0828	0	****
41	H5'	-1.00064	-2.59142	-0.48906	H1	1	U0	0.1312	0	****

42	H5"	-1.86422	-2.219	-1.97127	H1	1	U0	0.1312	0	****
43	C12	1.349741	4.954356	-4.82718	CT	1	U0	-0.004	0	****
44	H01	1.293287	3.874735	-4.81786	H1	1	U0	0.082	0	****
45	H02	1.823135	5.262609	-3.89979	H1	1	U0	0.082	0	****
46	C13	3.217191	6.193231	-5.9134	CD	1	U0	-0.123	0	****
47	H28	3.530791	6.499684	-4.92774	HA	1	U0	0.1489	0	****
48	C14	2.13704	5.433828	-6.01234	CM	1	U0	-0.1482	0	****
49	H30	1.763637	5.118656	-6.97232	HA	1	U0	0.121	0	****
50	C15	4.068125	6.70166	-7.00805	CA	1	U0	-0.0796	0	****
51	C16	3.972118	6.221684	-8.32657	CA	1	U0	-0.1554	0	****
52	H29	3.275904	5.438843	-8.54967	HA	1	U0	0.1359	0	****
53	C17	4.765105	6.723468	-9.3278	C	1	U0	0.1773	0	****
54	C18	5.703556	7.739827	-9.03659	C	1	U0	0.2958	0	****
55	C19	5.806584	8.203458	-7.74695	CA	1	U0	-0.2558	0	****
56	H11	6.514131	8.970634	-7.50026	HA	1	U0	0.159	0	****
57	C20	4.993944	7.682974	-6.73798	CA	1	U0	-0.1787	0	****
58	H33	5.097807	8.066848	-5.73869	HA	1	U0	0.1566	0	****

@<TRIPOS>BOND

1	1	2	1
2	1	40	1
3	2	3	1
4	2	4	1
5	5	36	1
6	6	7	1
7	6	8	1
8	6	21	1
9	8	9	1
10	8	10	1
11	10	11	1

12	10	15	1
13	11	12	1
14	11	43	1
15	13	31	1
16	13	38	1
17	14	15	1
18	14	21	1
19	14	30	1
20	16	17	1
21	16	54	1
22	17	18	1
23	17	19	1
24	17	20	1
25	21	22	1
26	22	28	1
27	23	24	1
28	23	53	1
29	24	25	1
30	24	26	1
31	24	27	1
32	28	29	1
33	28	30	1
34	30	31	1
35	31	32	1
36	31	33	1
37	33	34	1
38	33	35	1
39	33	36	1
40	36	37	1

41 36 38 1

42 38 39 1

43 38 40 1

44 40 41 1

45 40 42 1

46 43 44 1

47 43 45 1

48 43 48 1

49 46 47 1

50 46 48 1

51 46 50 1

52 48 49 1

53 50 51 1

54 50 57 1

55 51 52 1

56 51 53 1

57 53 54 1

58 54 55 1

59 55 56 1

60 55 57 1

61 57 58 1

@<TRIPOS>SUBSTRUCTURE

1 U0 1 **** 0 **** ***

@<TRIPOS>HEADTAIL

P 1

O3' 1

@<TRIPOS>RESIDUECONNECT

1 P O3' 0 0 0 0

Figure S6. Cartesian coordinates and RESP charges calculated for the N^2-MEG-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 DNA and hpolκ residues

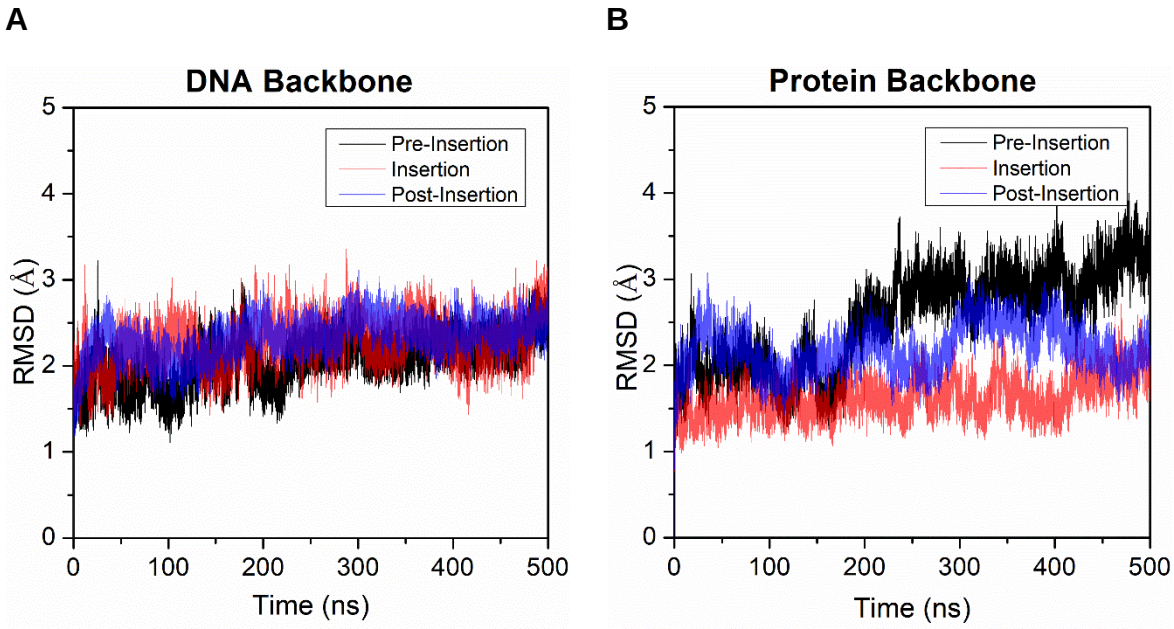
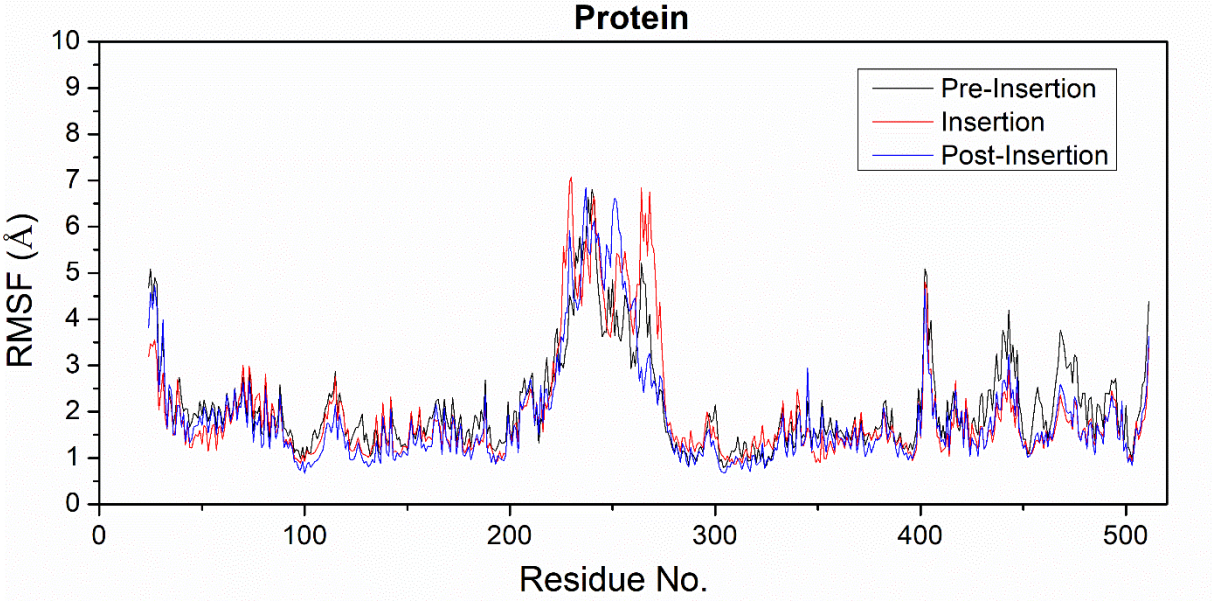


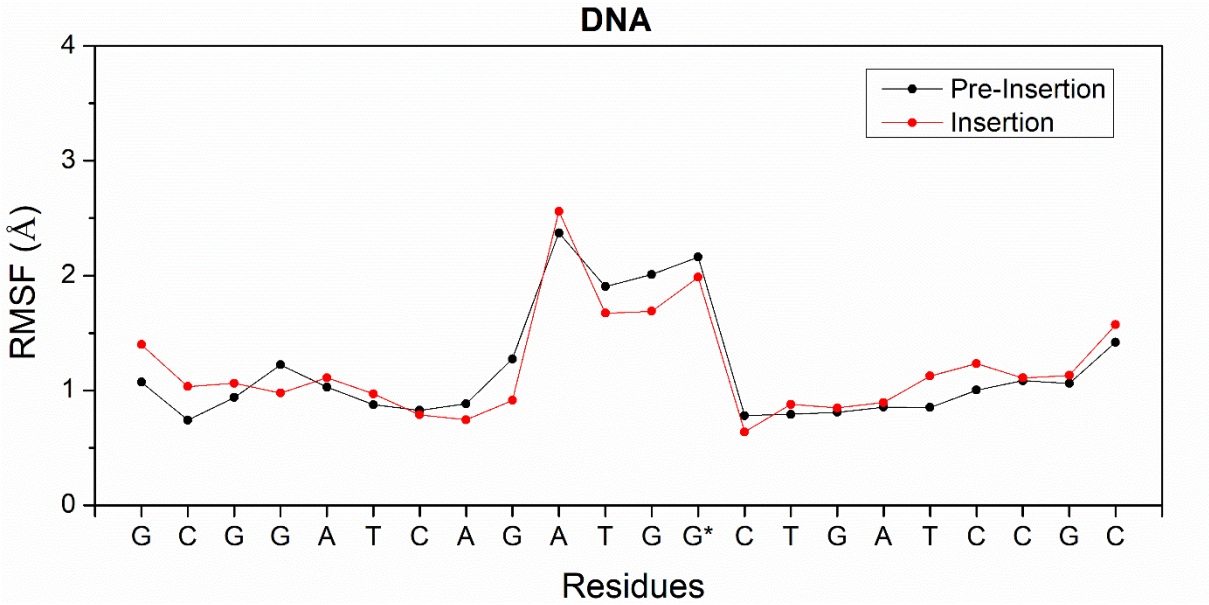
Figure S7. (A) Variation of RMSD values of DNA backbone atoms; (B) Variation of RMSD values of protein backbone atoms. The whole trajectory of 500 ns was used for RMSD calculations.

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

A



B



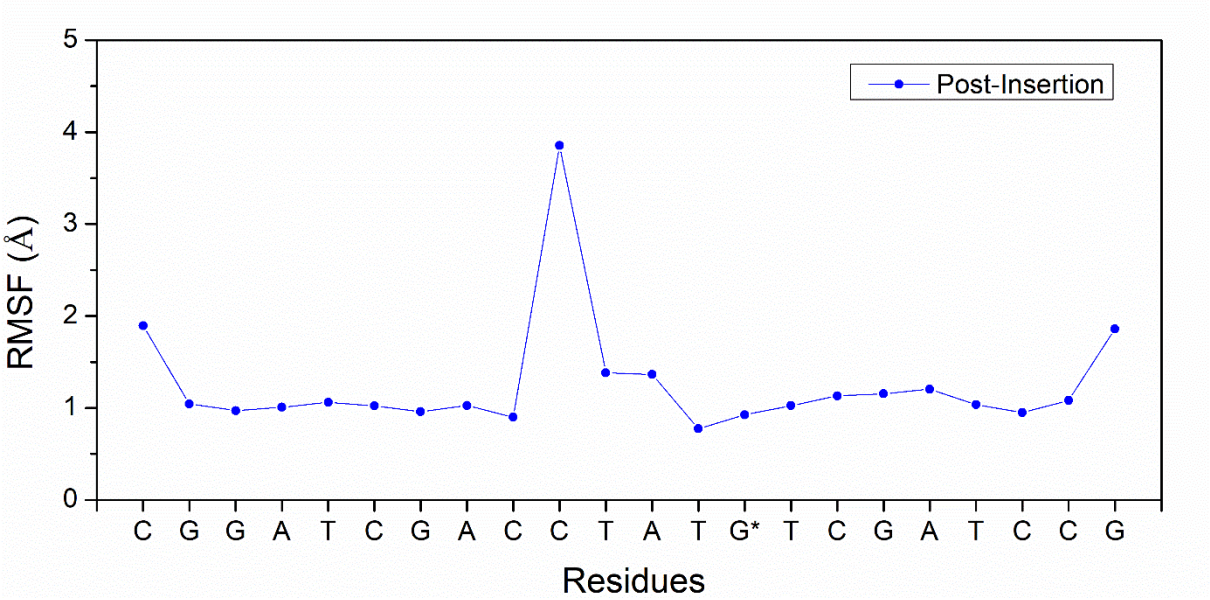
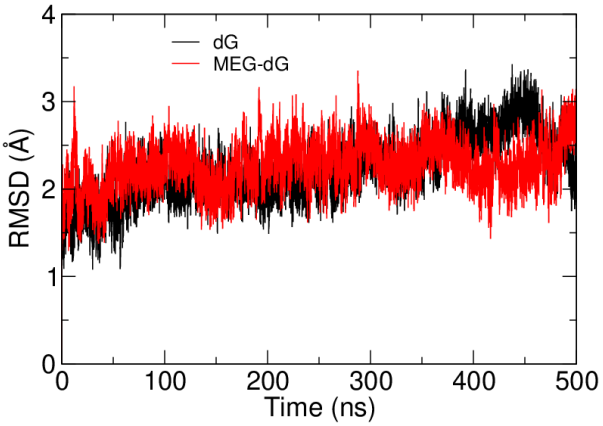


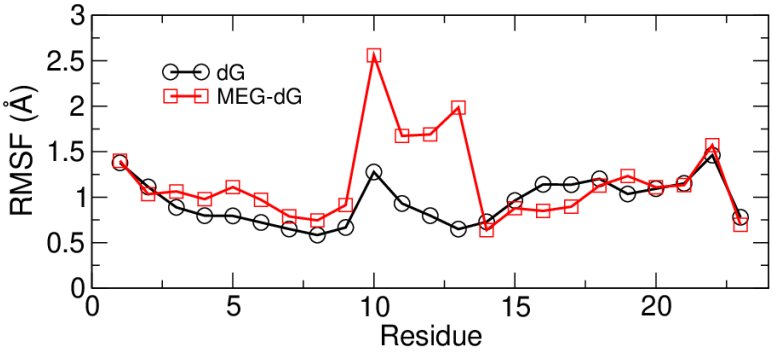
Figure S8. (A) Variation of RMSF values of amino acid residues in protein. The black, red and blue lines represent the pre-insertion, insertion and post-insertion stages respectively. (B) Variation of RMSF values of nucleotides in DNA, the damaged base is represented by G*. The whole trajectory of 500ns was used for RMSF calculations.

Comparison of MD parameters for MEG-modified and unmodified duplex DNAs

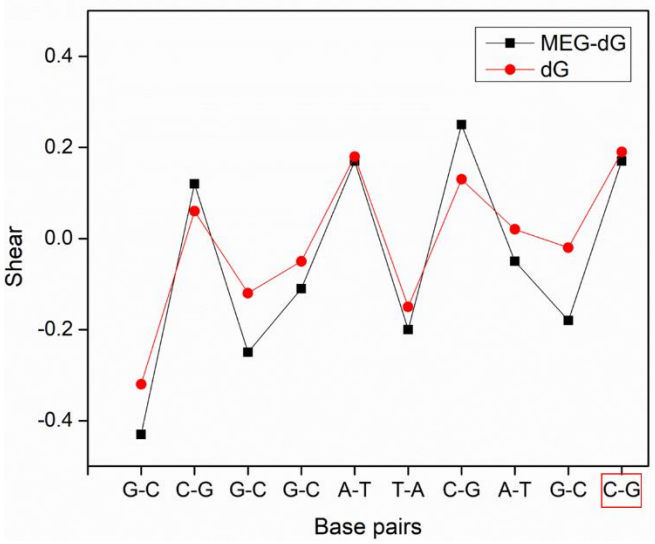
A



B



C



D

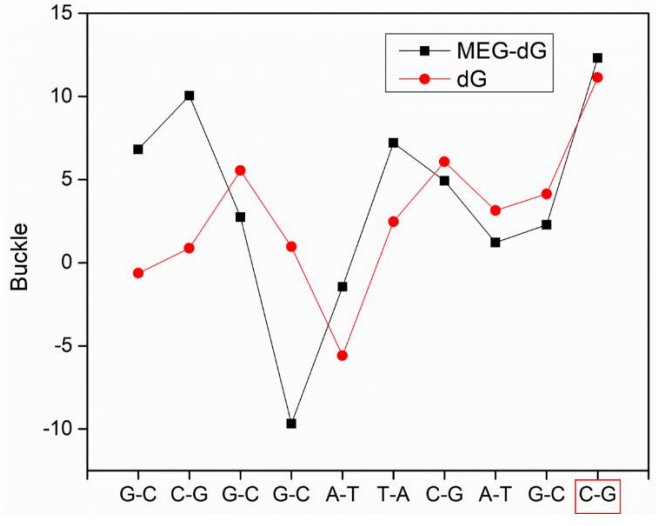


Figure S9. The Variation of (A) RMSD values of DNA backbone atoms; (B) RMSF values of nucleotides; (C) local base pair parameter shear; (D) local base pair parameter buckle of dG and MEG-dG complexes.

Orientation of N^2-dG and N^2-MEG-dG in the predominantly hydrophobic pocket of hpolκ

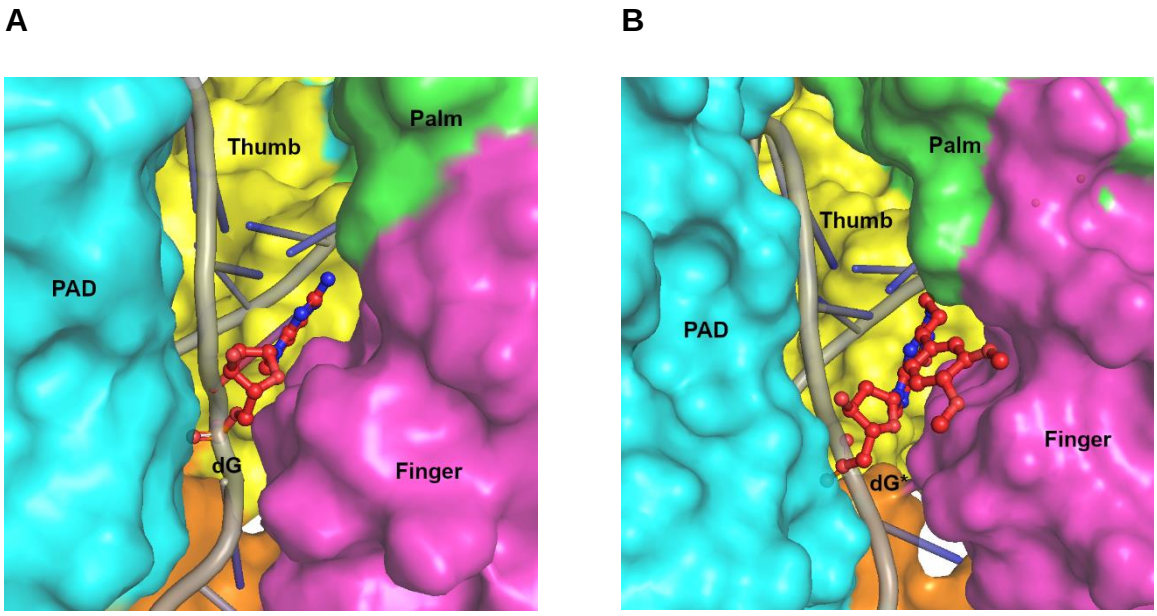
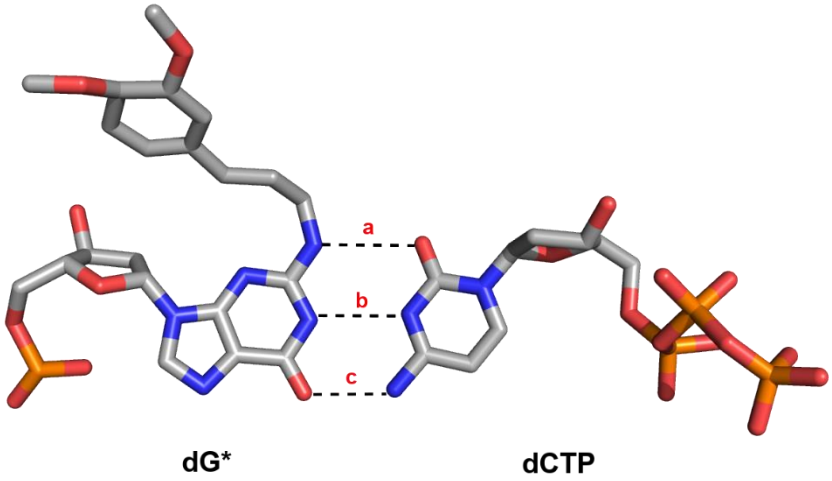


Figure S10. The orientation of (A) N^2-dG unmodified and (B) N^2-MEG-dG in the predominantly hydrophobic pocket of hpolκ. The various domains are marked in the figure. The N^2-MEG-dG (dG*) is represented in red color.

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



Bond	Donor Atom	Acceptor Atom	Percentage Occupancy	Average Distance	Average Angle
a	dG*@N2	dCTP@O2	99%	2.9 Å	160°
b	dG*@N1	dCTP@N3	99%	3 Å	162°
c	dCTP@N4	dG*@O6	98%	3 Å	163°

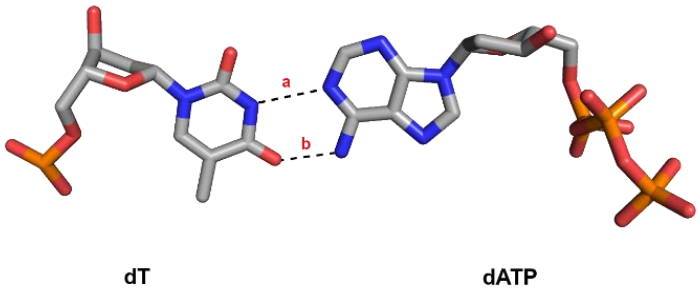
Table S1. The figure represents the Watson-Crick hydrogen bonds between the N^2-MEG-dG (dG*) and the incoming nucleotide (dCTP) in the insertion stage. The hydrogen bonds are illustrated as dotted lines. Hydrogen bonds are determined using simple geometric criteria with donor to acceptor heavy atom distance (less than 3.4 Å) and donor-hydrogen-acceptor angle (greater than 120°).

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

Residue 1	Residue 2	Unmodified (kcal/mol)	N^2-MEG-dG (kcal/mol)
G* 13	**SER 134**	-0.42 ± 0.3	**-1.2 ± 0.5**
G* 13	**MET 135**	**-4.6 ± 0.5**	**-4.8 ± 0.7**
G* 13	**LEU 136**	-0.7 ± 0.3	-0.85 ± 0.3
G* 13	**SER 137**	-0.9 ± 0.3	**-1.62 ± 0.36**
G* 13	**PRO 169**	-0.03 ± 0.01	**-2.1 ± 0.6**
G* 13	**PHE 171**	-0.12 ± 0.06	**-1.4 ± 0.4**
G* 13	**ARG 420**	-0.95 ± 0.2	**-1.25 ± 0.25**

Table S2. Pairwise interaction energies between the unmodified-dG and the N^2-MEG-dG adduct (Residue 1) with the amino acids (Residue 2) present in the active site of hpolκ obtained from MM/GBSA analysis. The important interactions are highlighted in bold.

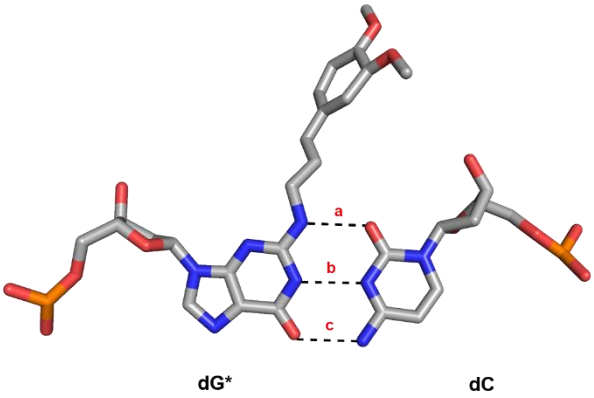
Percentage occupancies of H-bonds between dT and dATP in the post-insertion stage complex



Bond	Donor Atom	Acceptor Atom	Percentage Occupancy	Average Distance	Average Angle
a	dT@N3	dATP@N1	94%	3.1 Å	164°
b	dATP@N6	dT@O4	99%	2.9 Å	163°

Table S3. 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 simple geometric criteria with donor to acceptor heavy atom distance (less than 3.4 Å) and donor-hydrogen-acceptor angle (greater than 120°).

Percentage occupancies of H-bonds between N^2-MEG-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	164°
b	dG*@N1	dCTP@N3	99%	2.9 Å	164°
c	dCTP@N4	dG*@O6	92%	3 Å	160°

Table S4. The figure represents the Watson-Crick hydrogen bonds between the N^2-MEG-dG adduct (dG*) and dC. The hydrogen bonds are illustrated as dotted lines. Hydrogen bonds are determined using simple geometric criteria with donor to acceptor heavy atom distance (less than 3.4 Å) and donor-hydrogen-acceptor angle (greater than 120°).

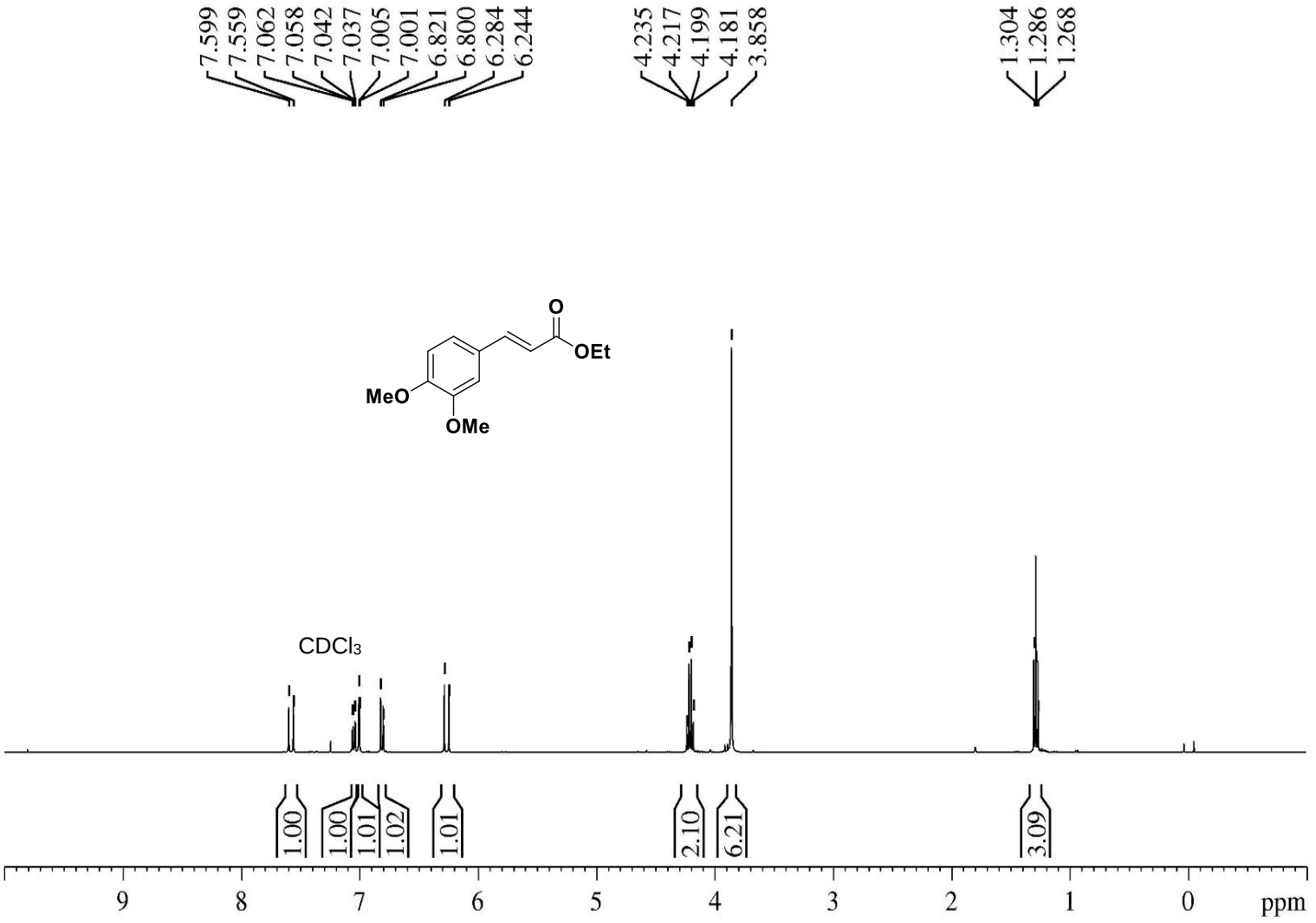
MALDI-TOF mass data of the DNA sequences

Code	DNA sequences (5'-3')	MW (calc.)	MW (Found)
T1	5'-CTGGTCACACTXATGCC TACGAGTACG-3'	8436.4	8434
T2	5'-CTGGTCACACTXATGCC TACGAGTACG-3'	8406.4	8407

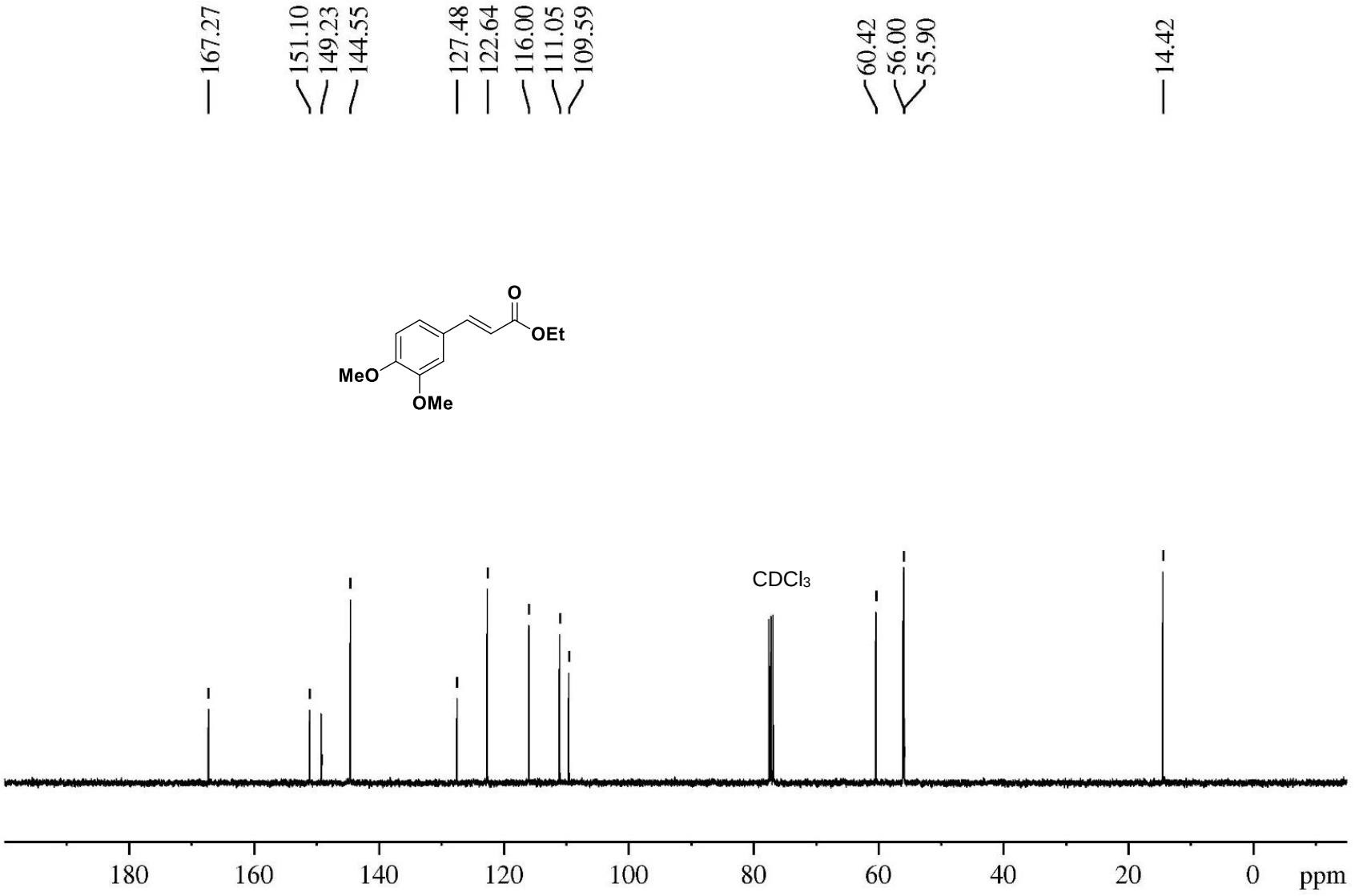
Table S5. T1: X = N^2-MEG-dG and T2: X= N^2-EG-dG modified DNAs were characterized using MALDI-TOF in the positive reflection mode.

Spectral data

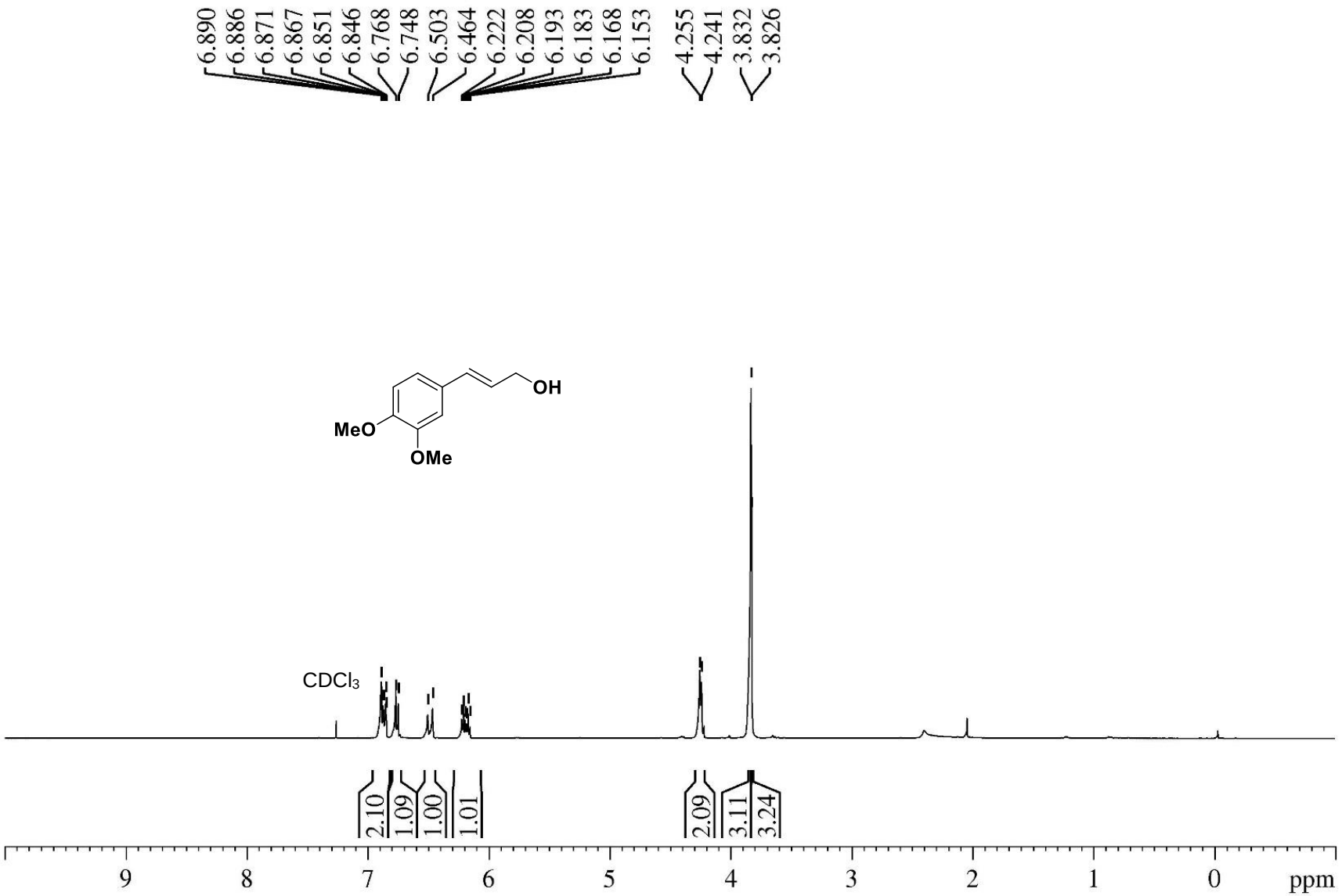
^{1}H spectra of compound 2a



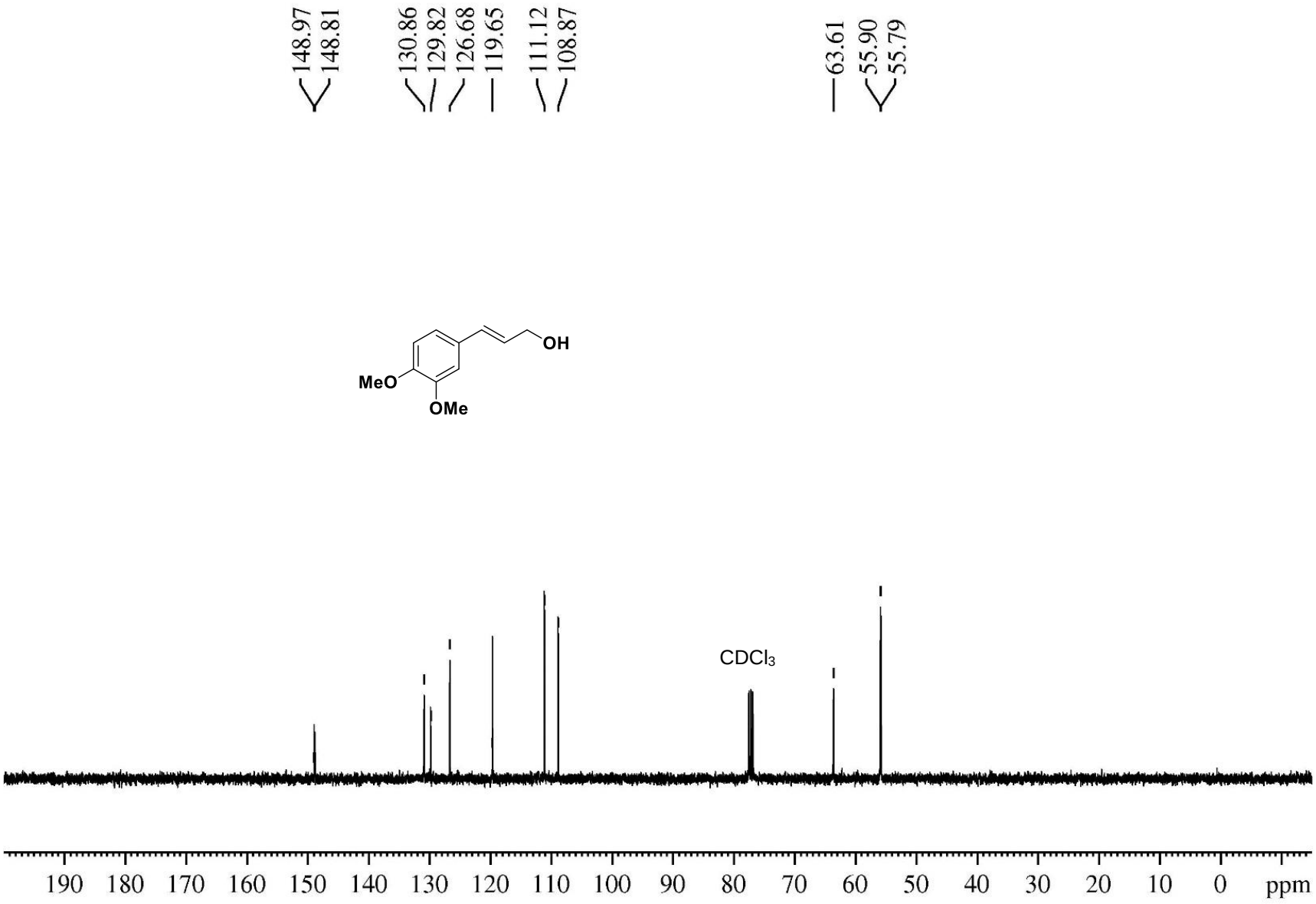
^{13}C spectra of compound 2a



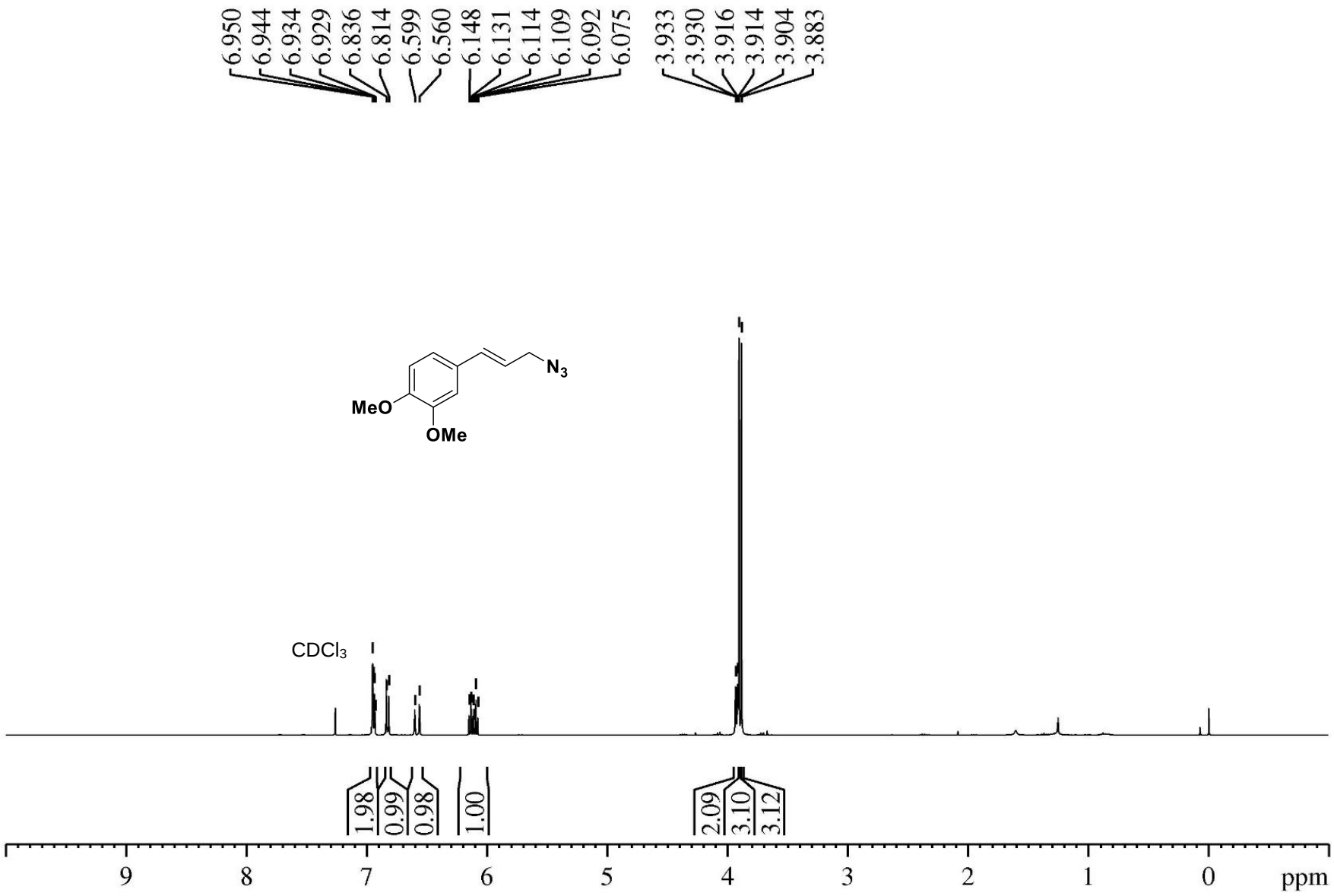
[1]H spectra of compound 3a



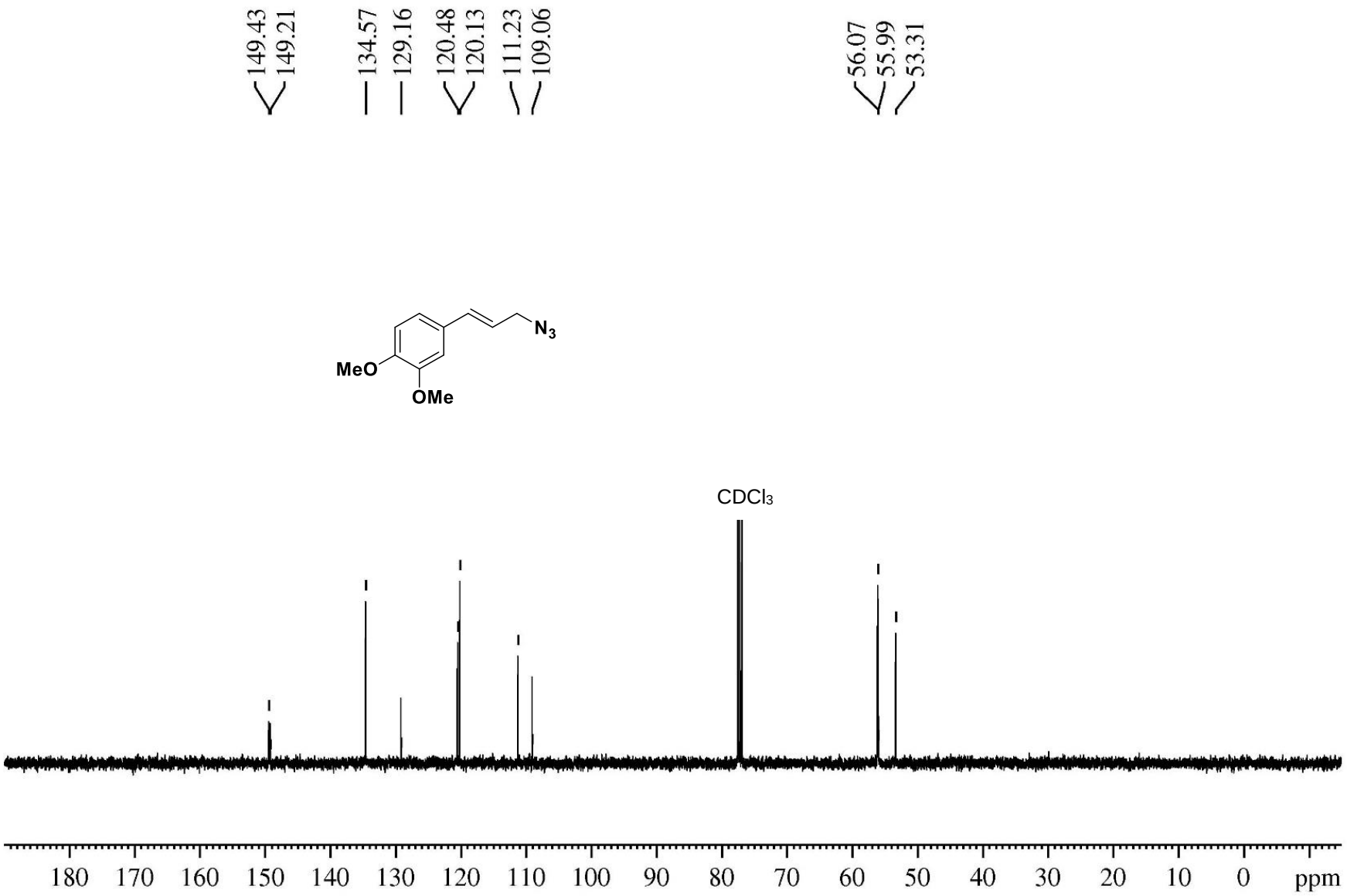
^{13}C spectra of compound 3a



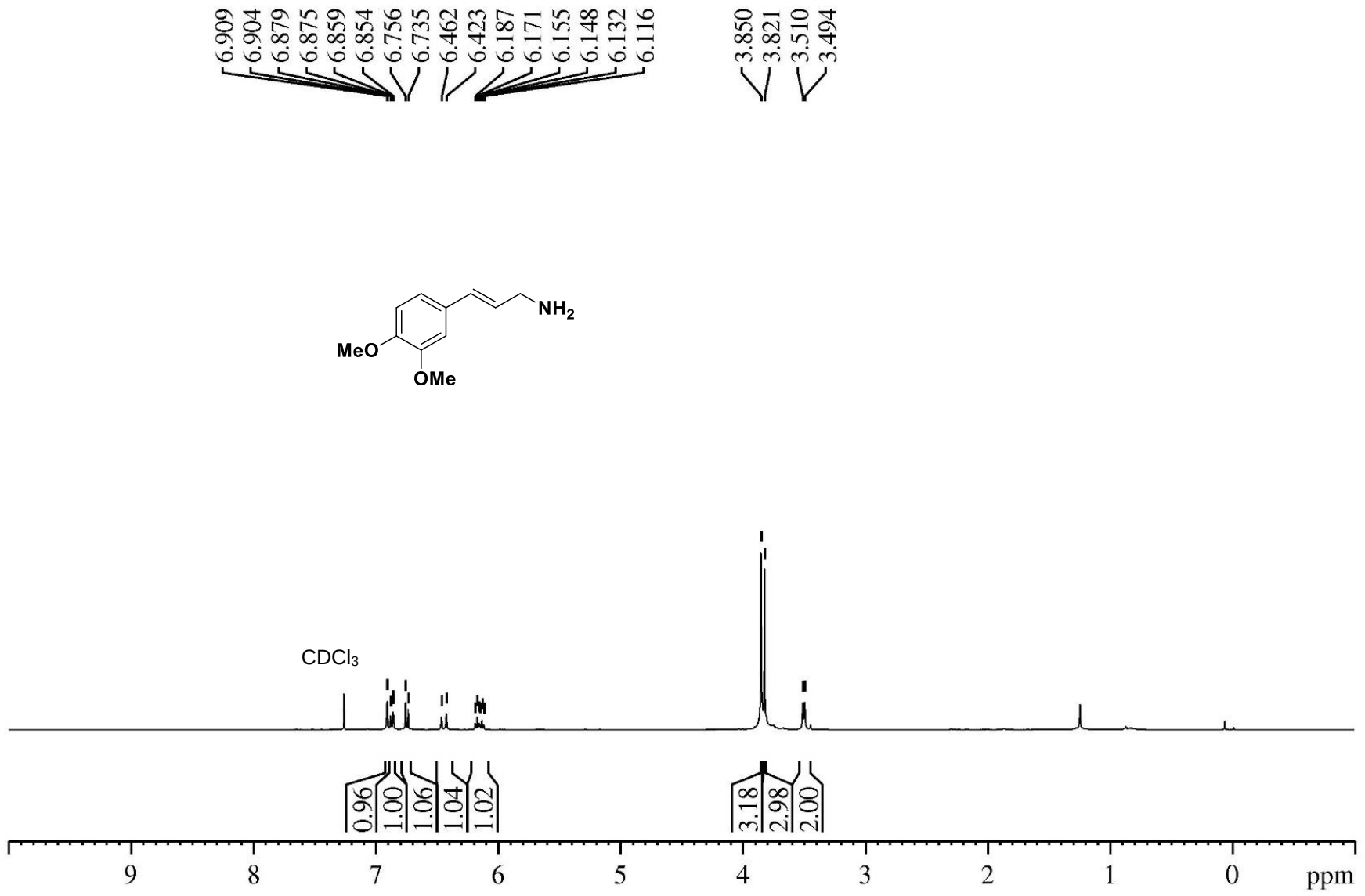
^{1}H spectra of compound 4a



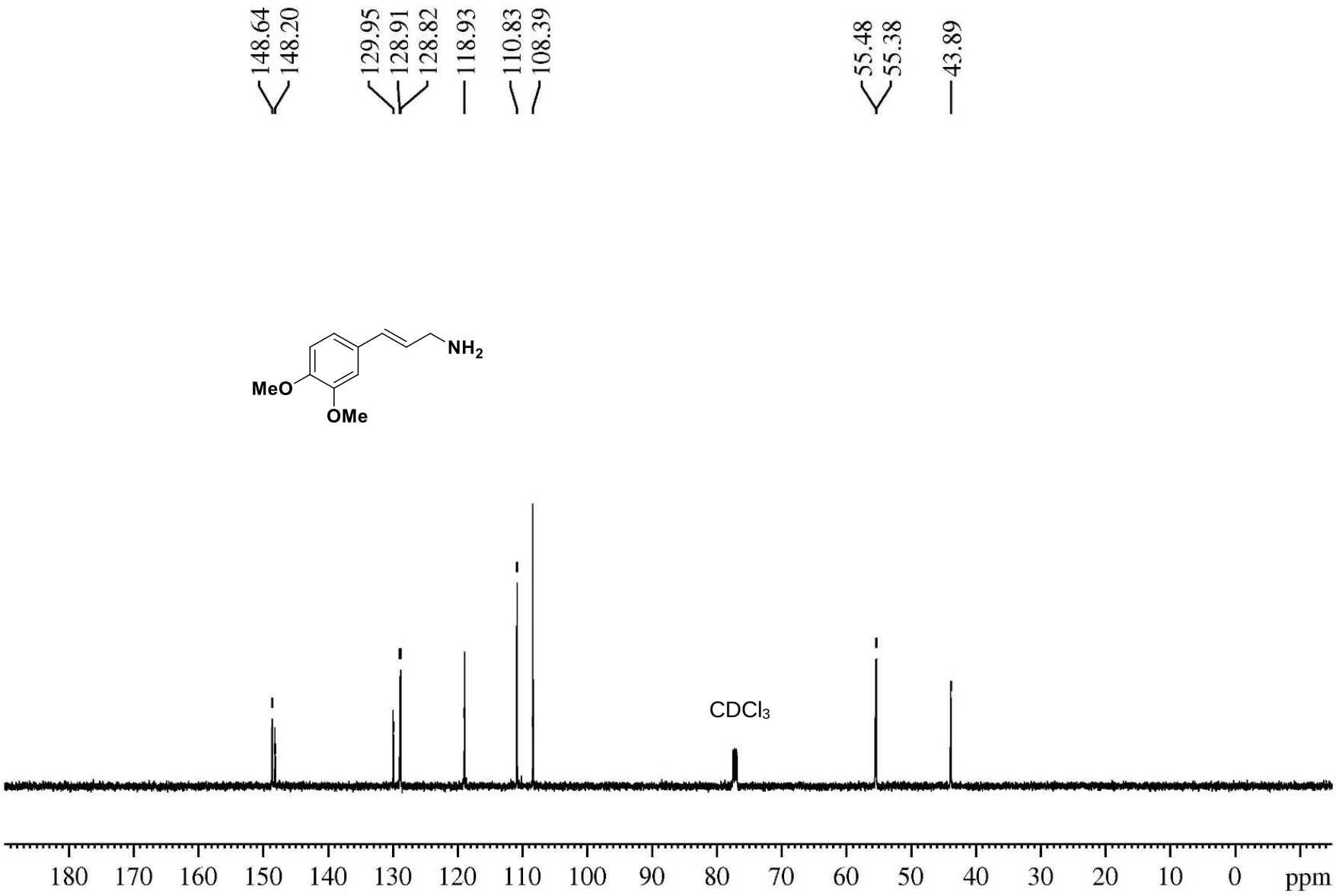
^{13}C spectra of compound 4a



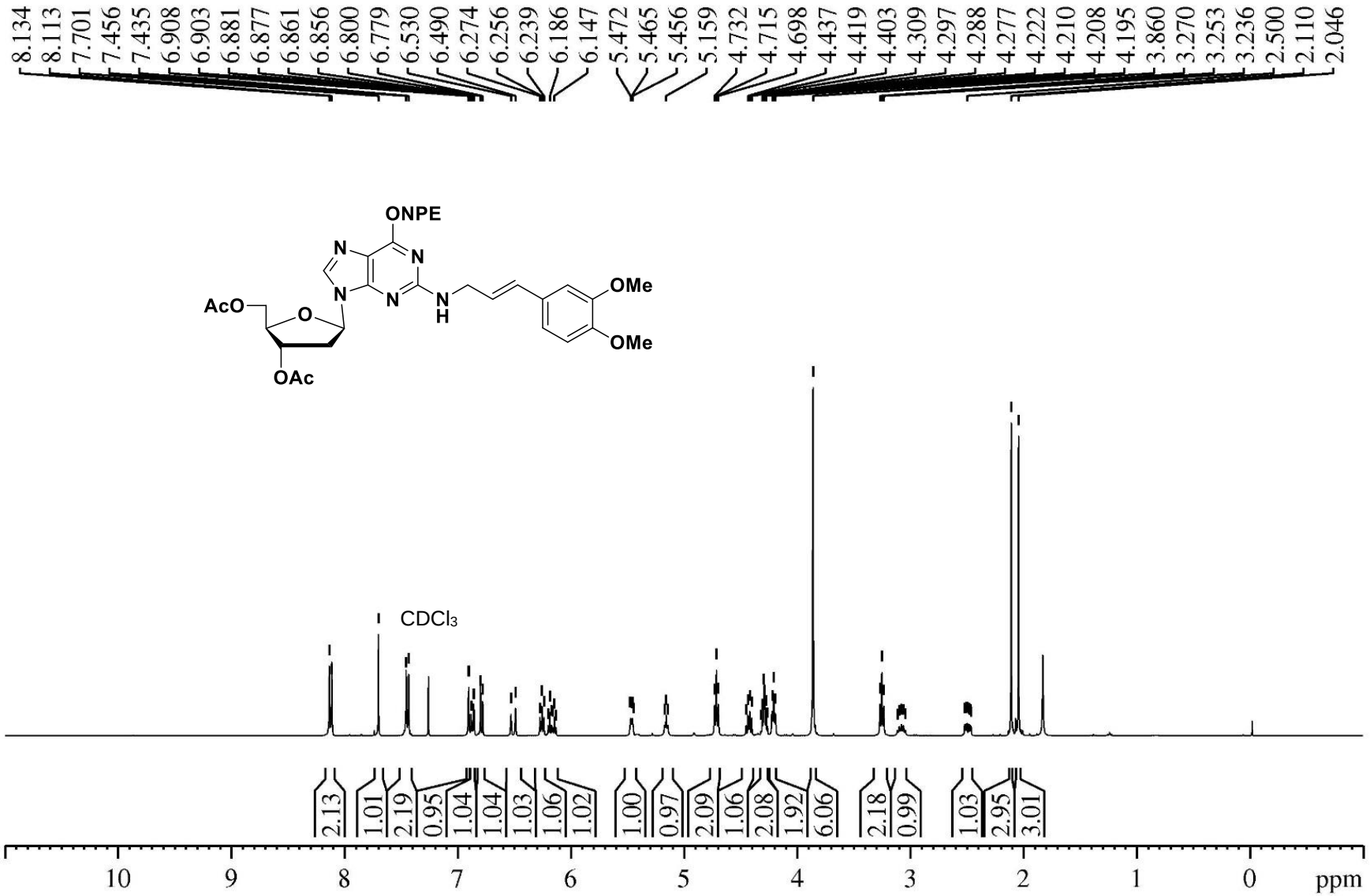
[1]H spectra of compound 5a



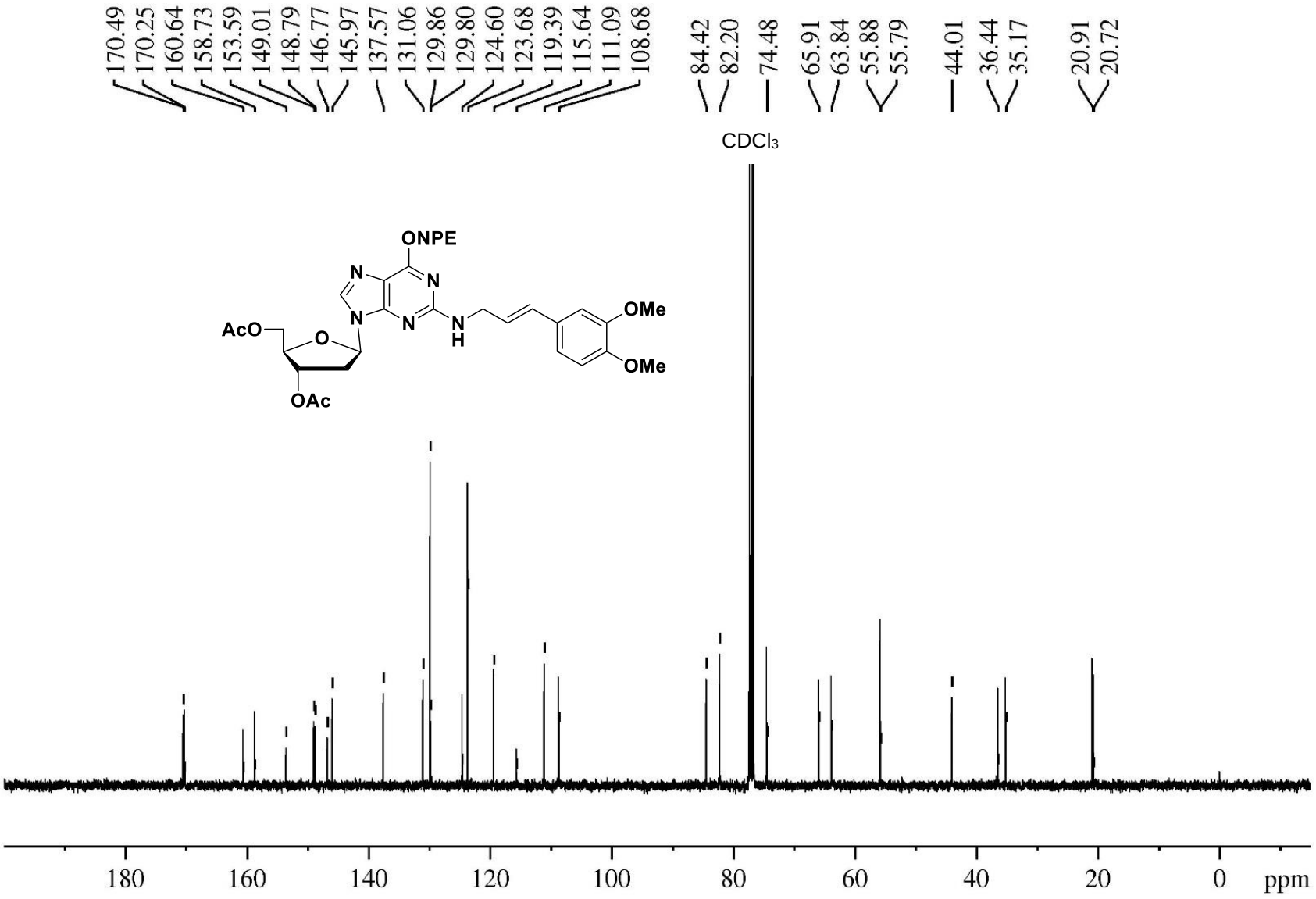
^{13}C spectra of compound 5a



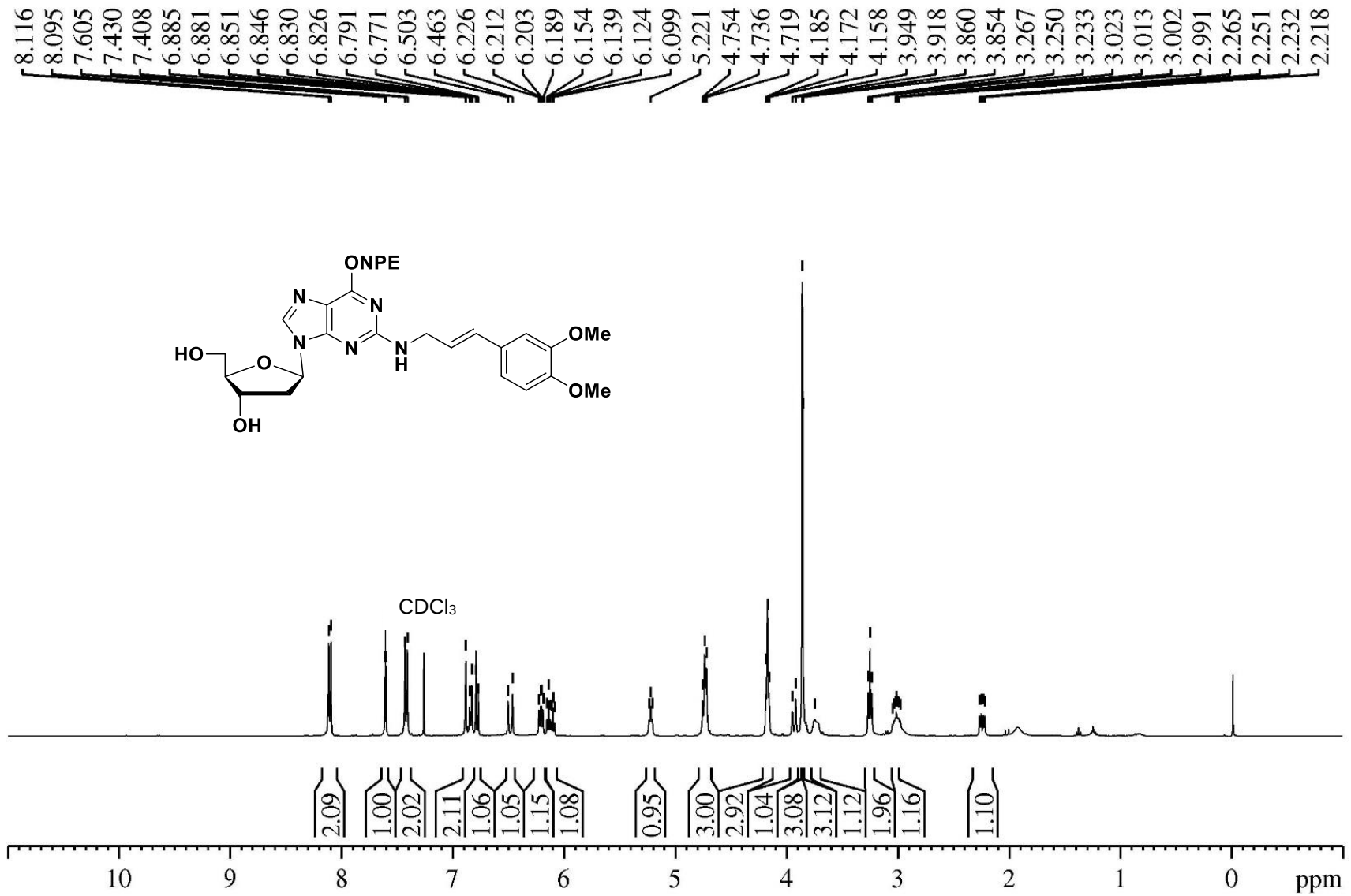
^{1}H spectra of compound 7a



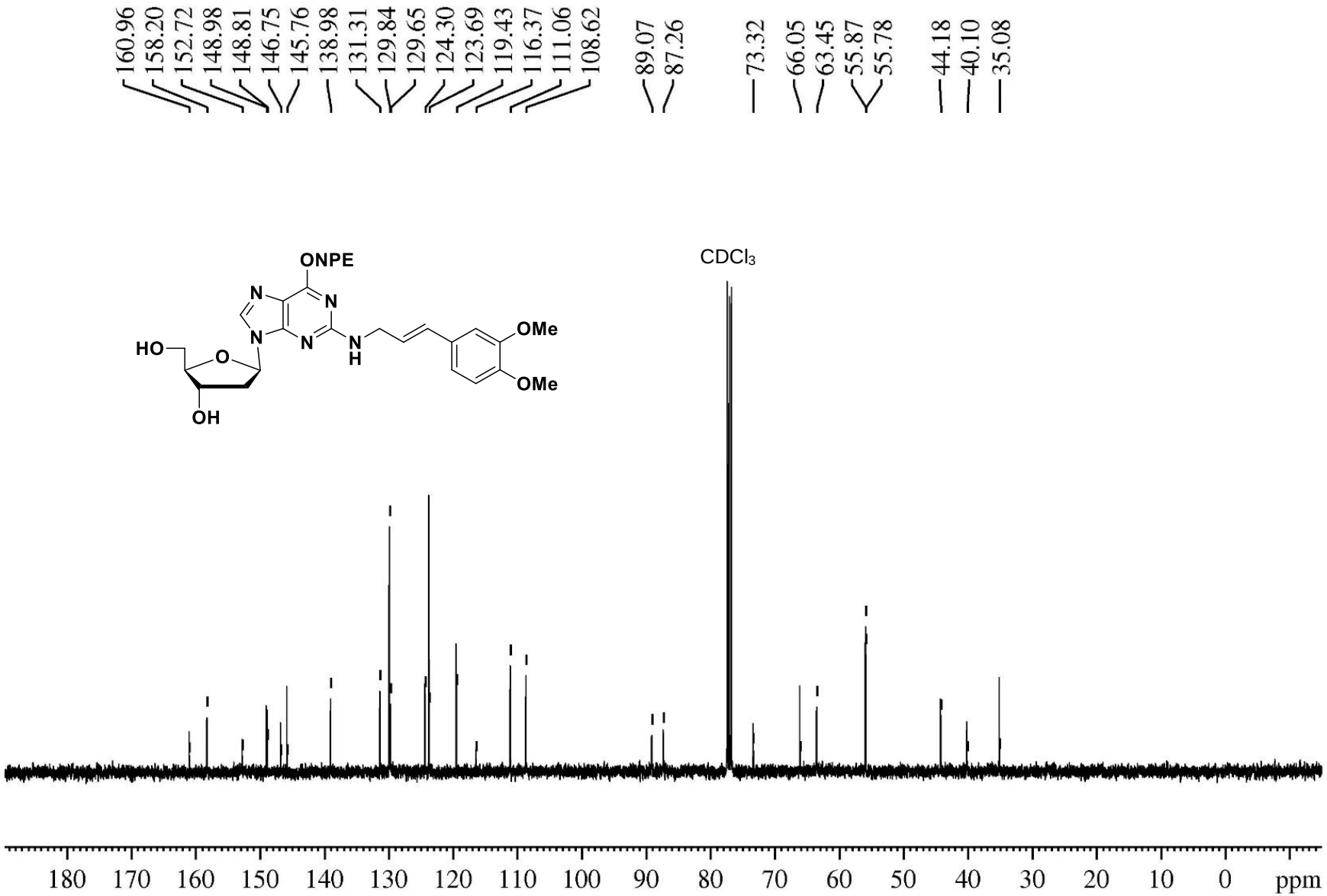
^{13}C spectra of compound 7a



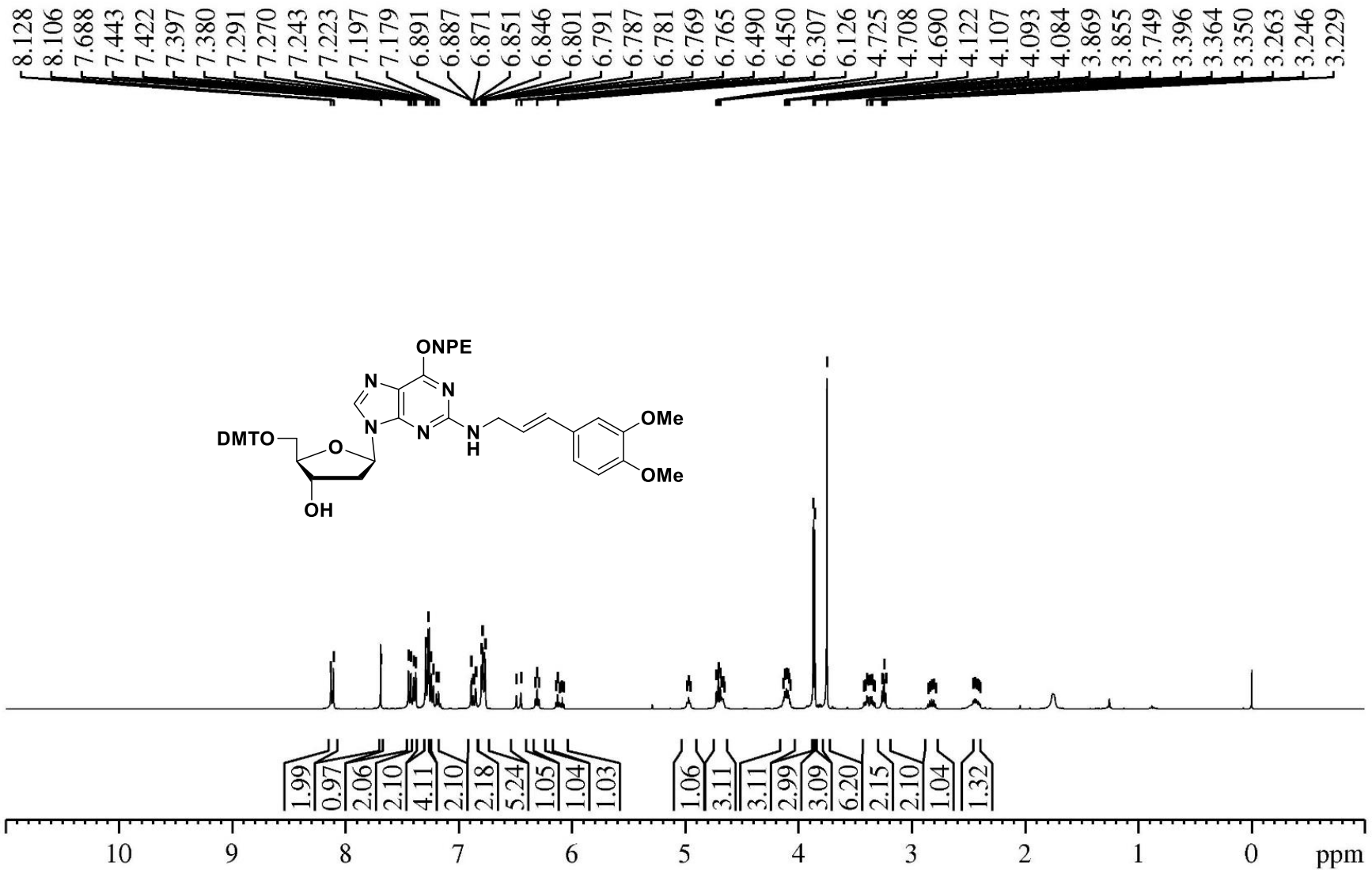
^{1}H spectra of compound 8a



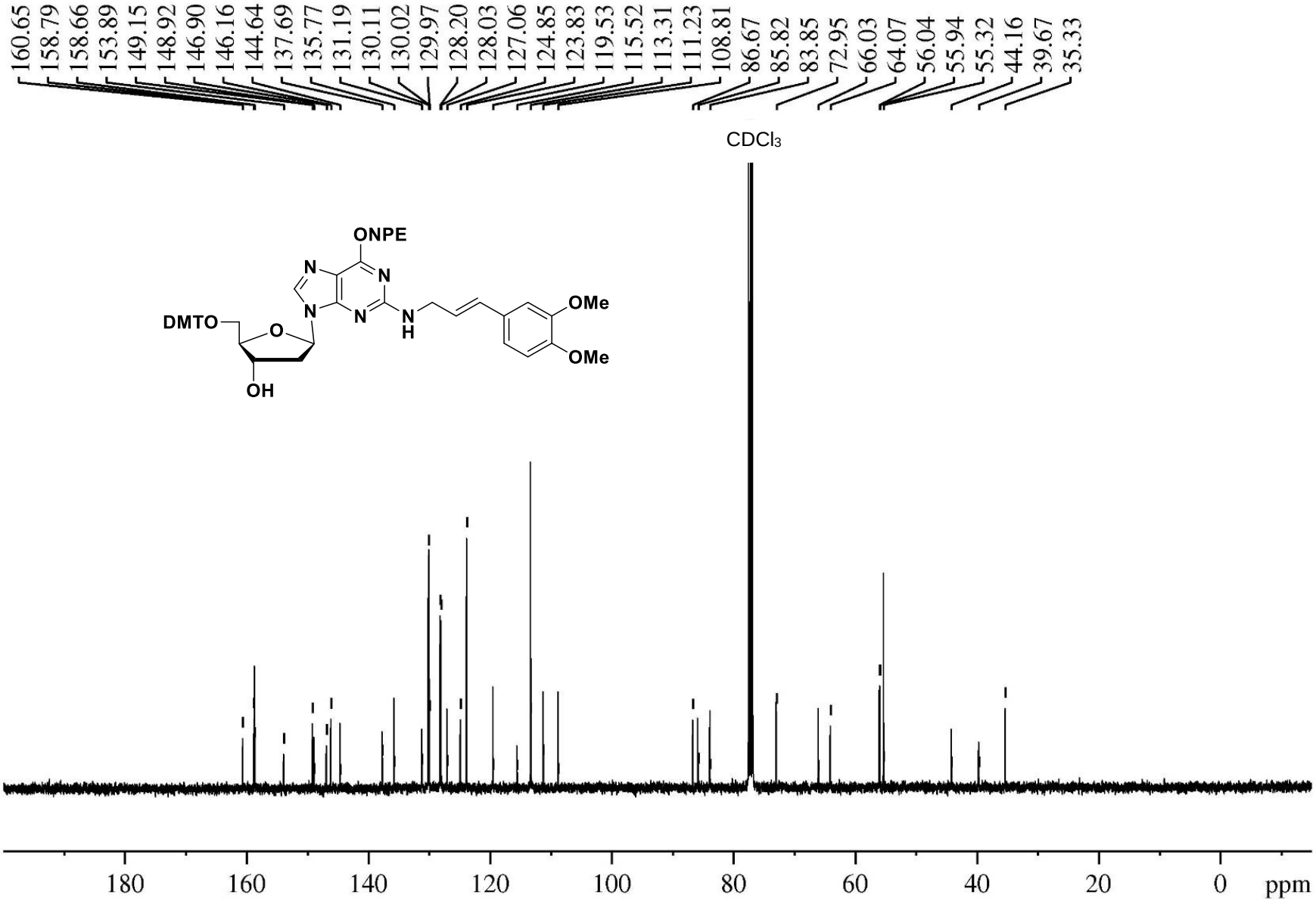
^{13}C spectra of compound 8a



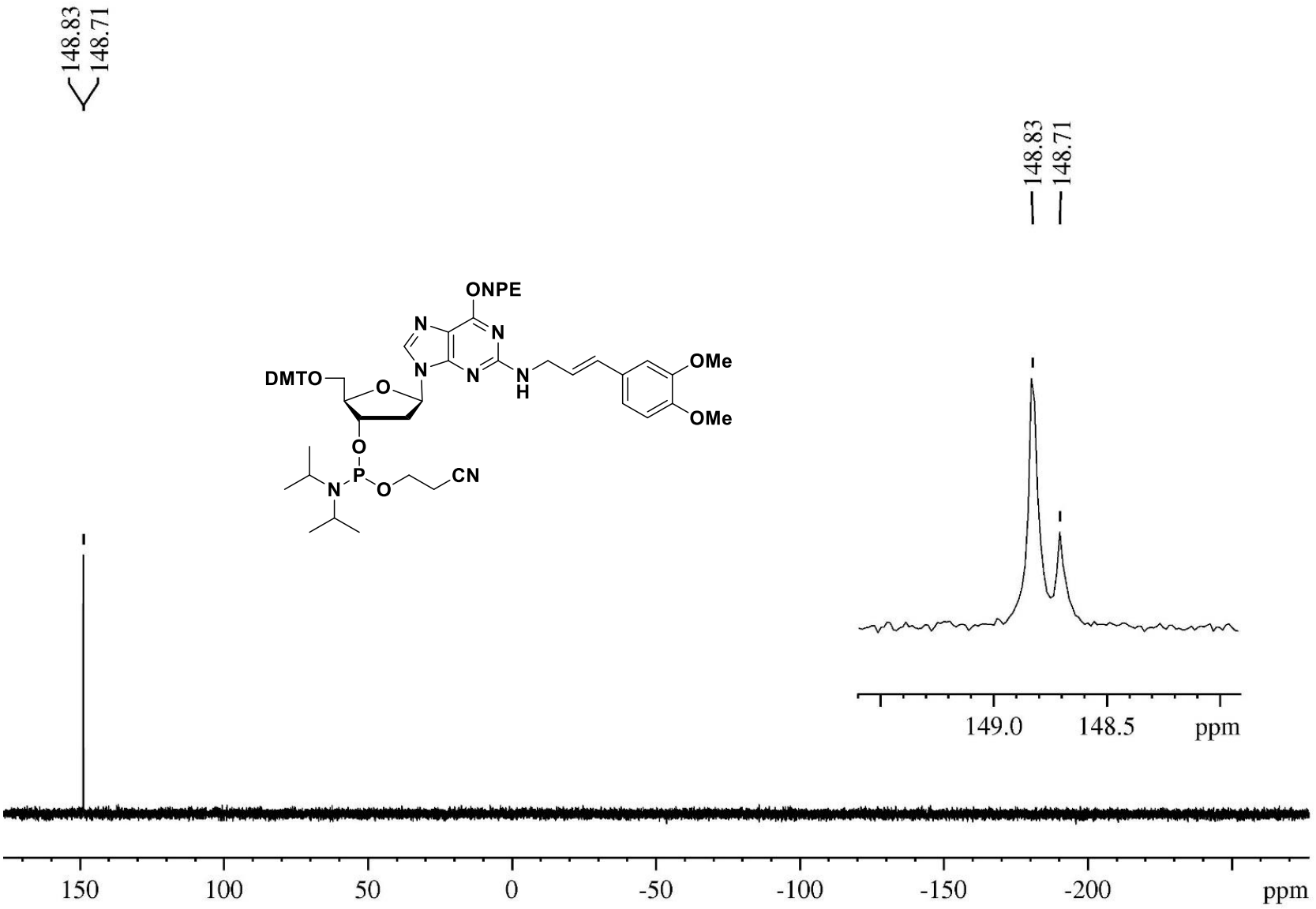
^{1}H spectra of compound 9a



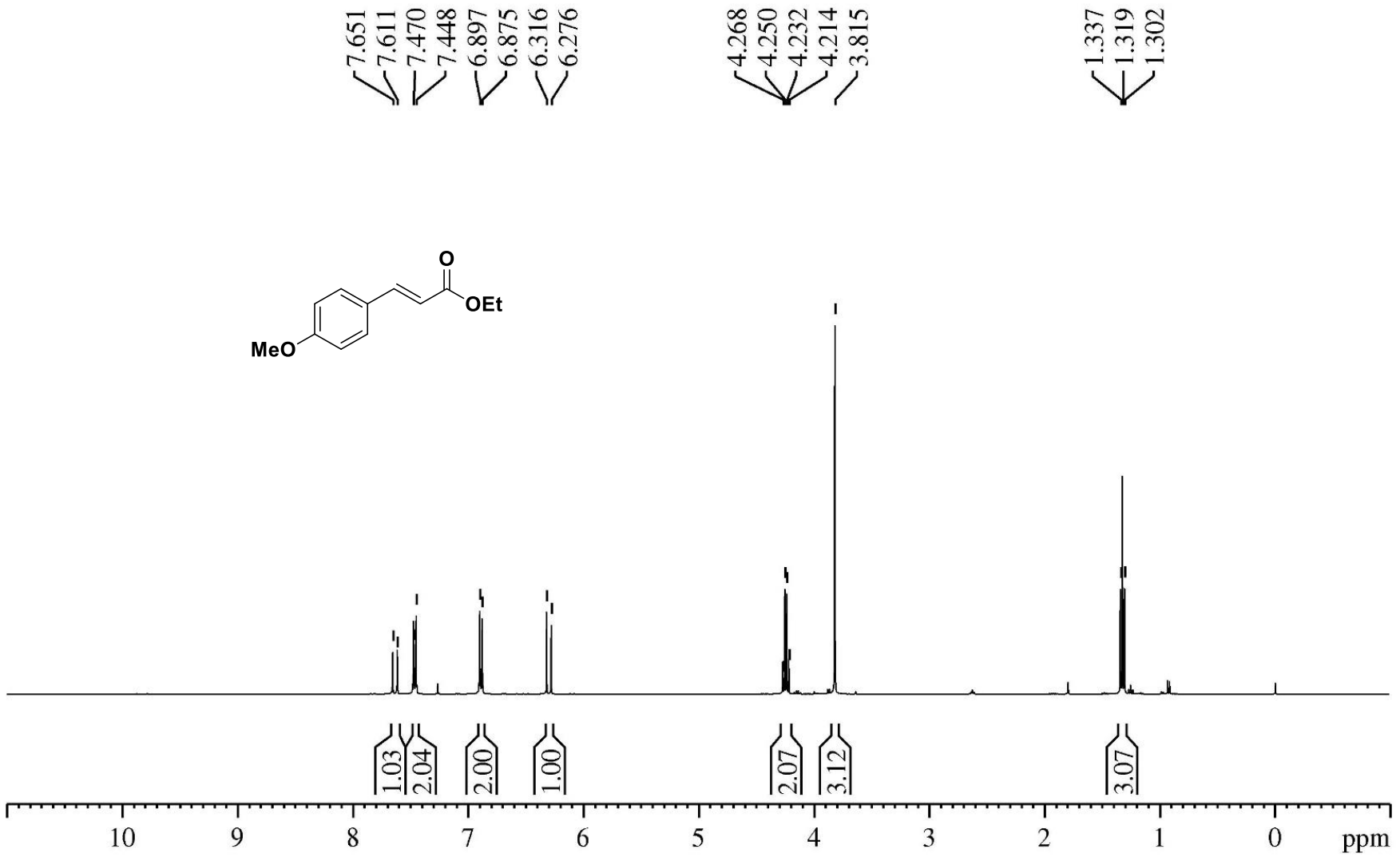
^{13}C spectra of compound 9a



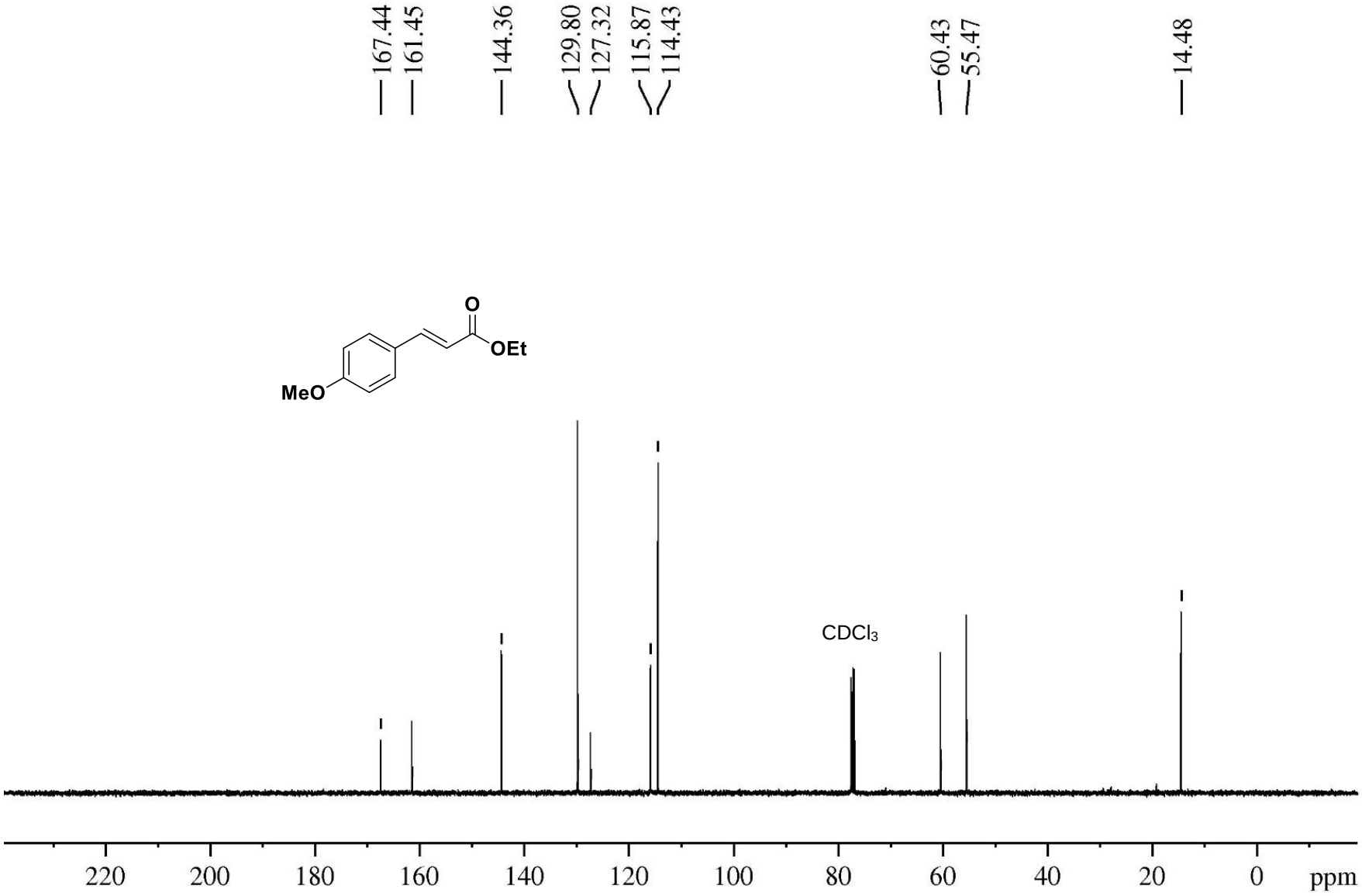
^{31}P spectra of compound 10a



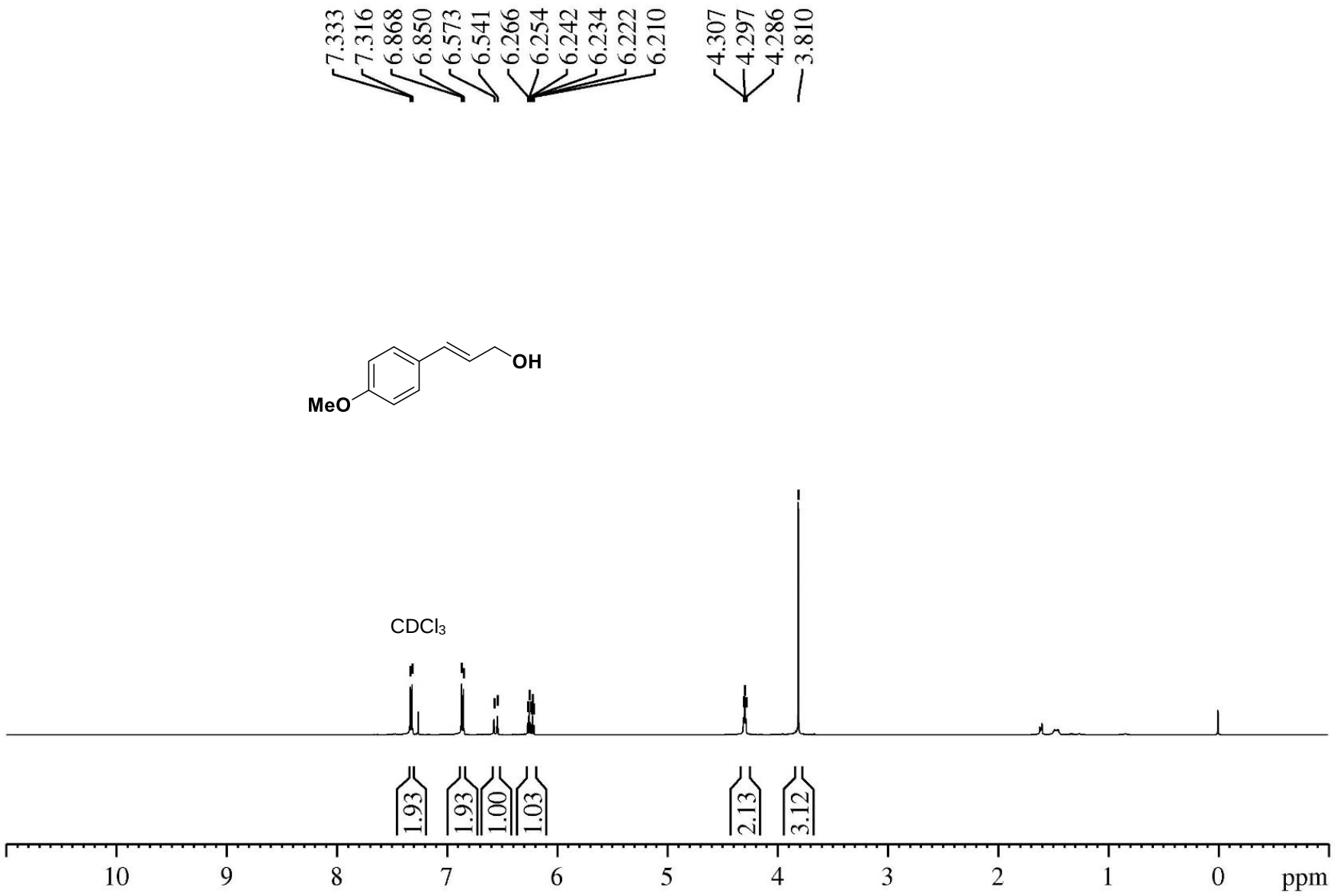
[1]H spectra of compound 2b



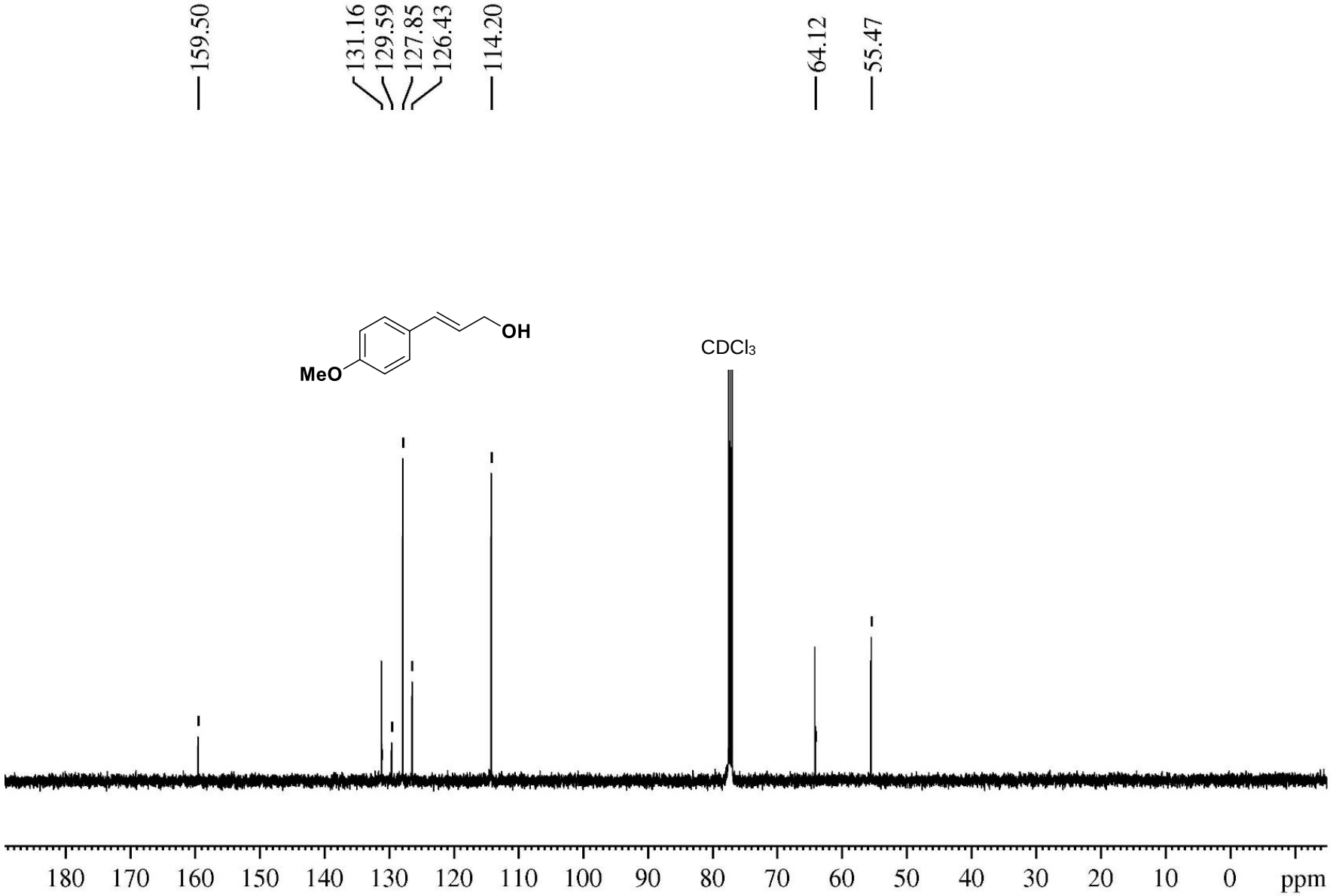
^{13}C spectra of compound 2b



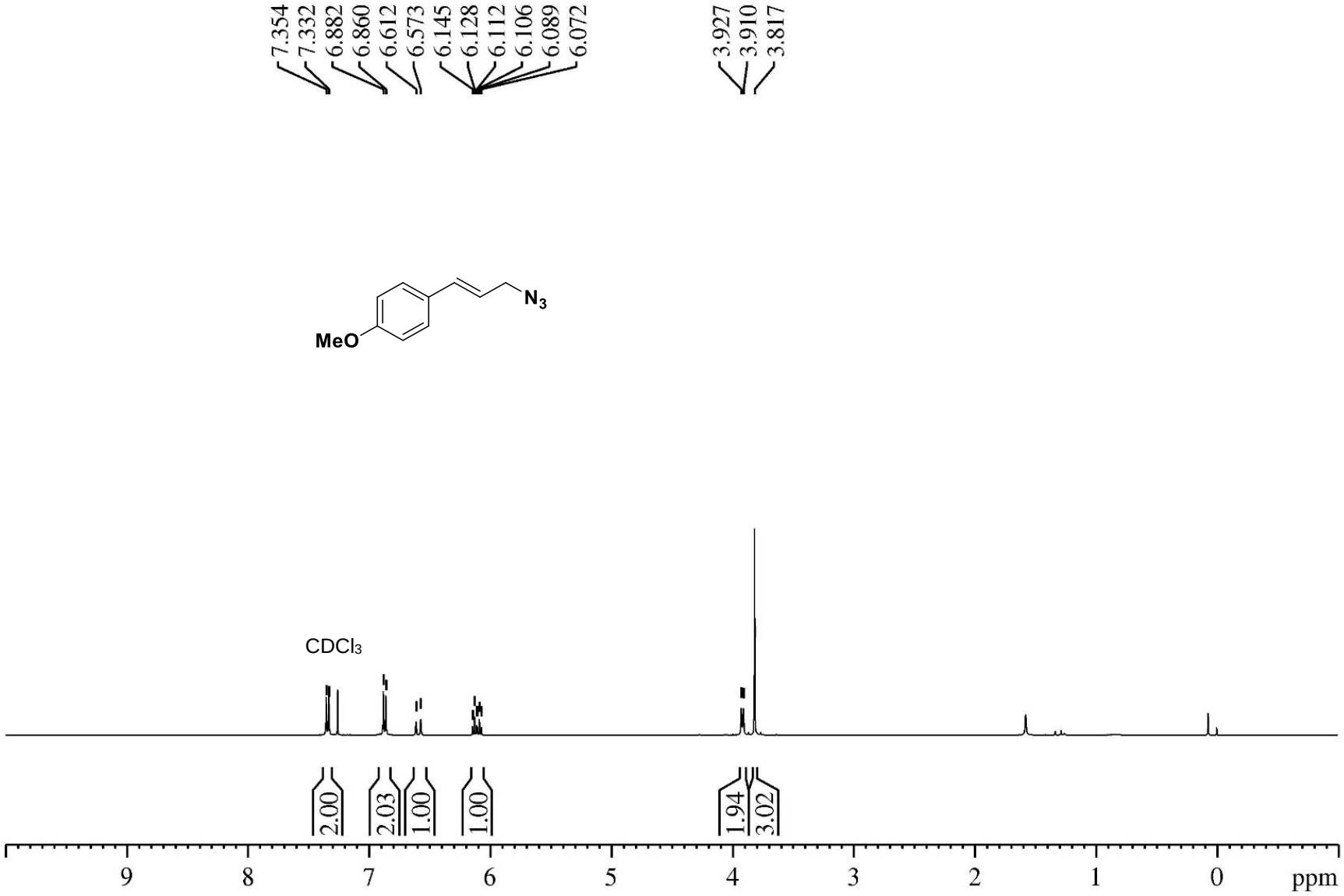
[1]H spectra of compound 3b



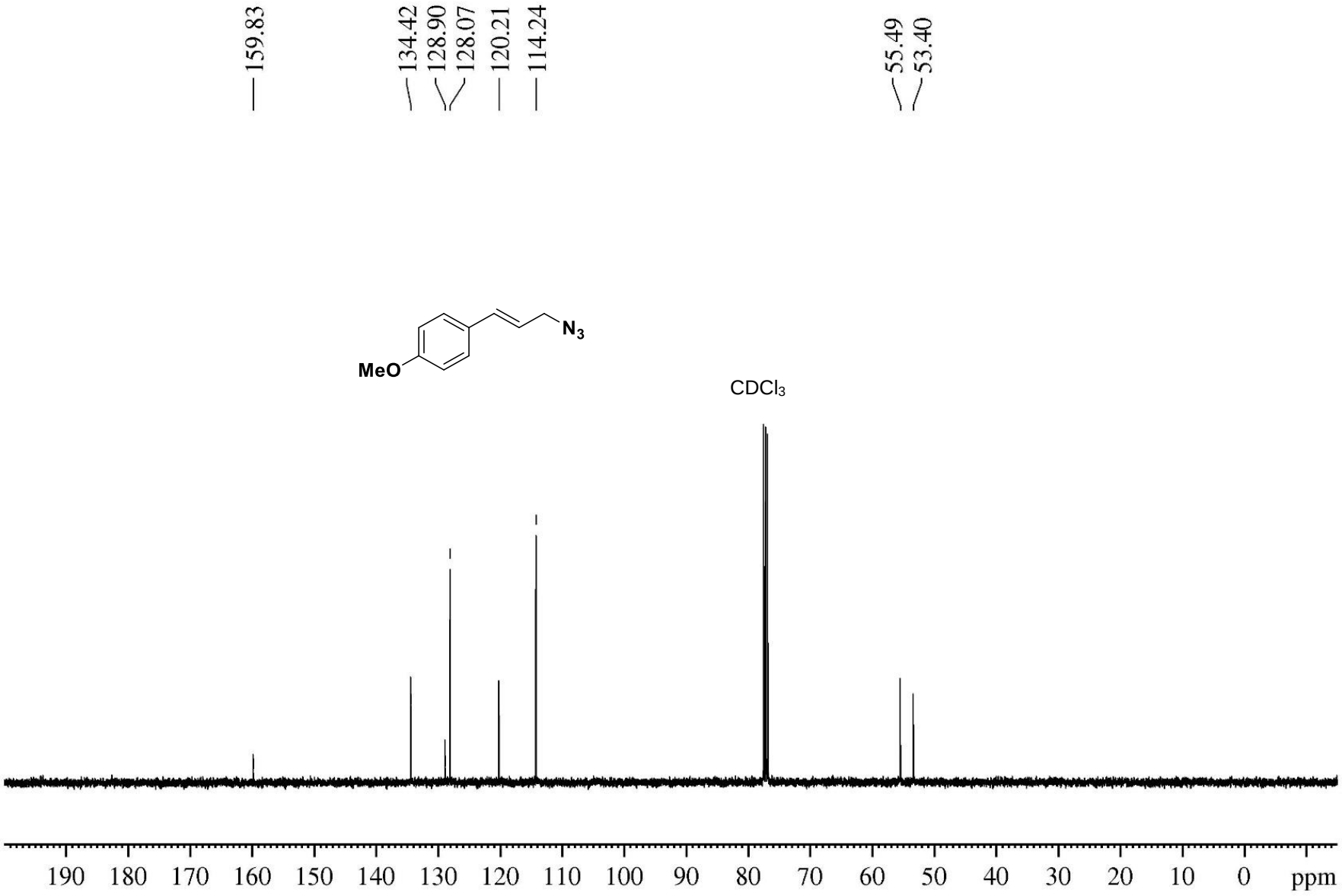
^{13}C spectra of compound 2b



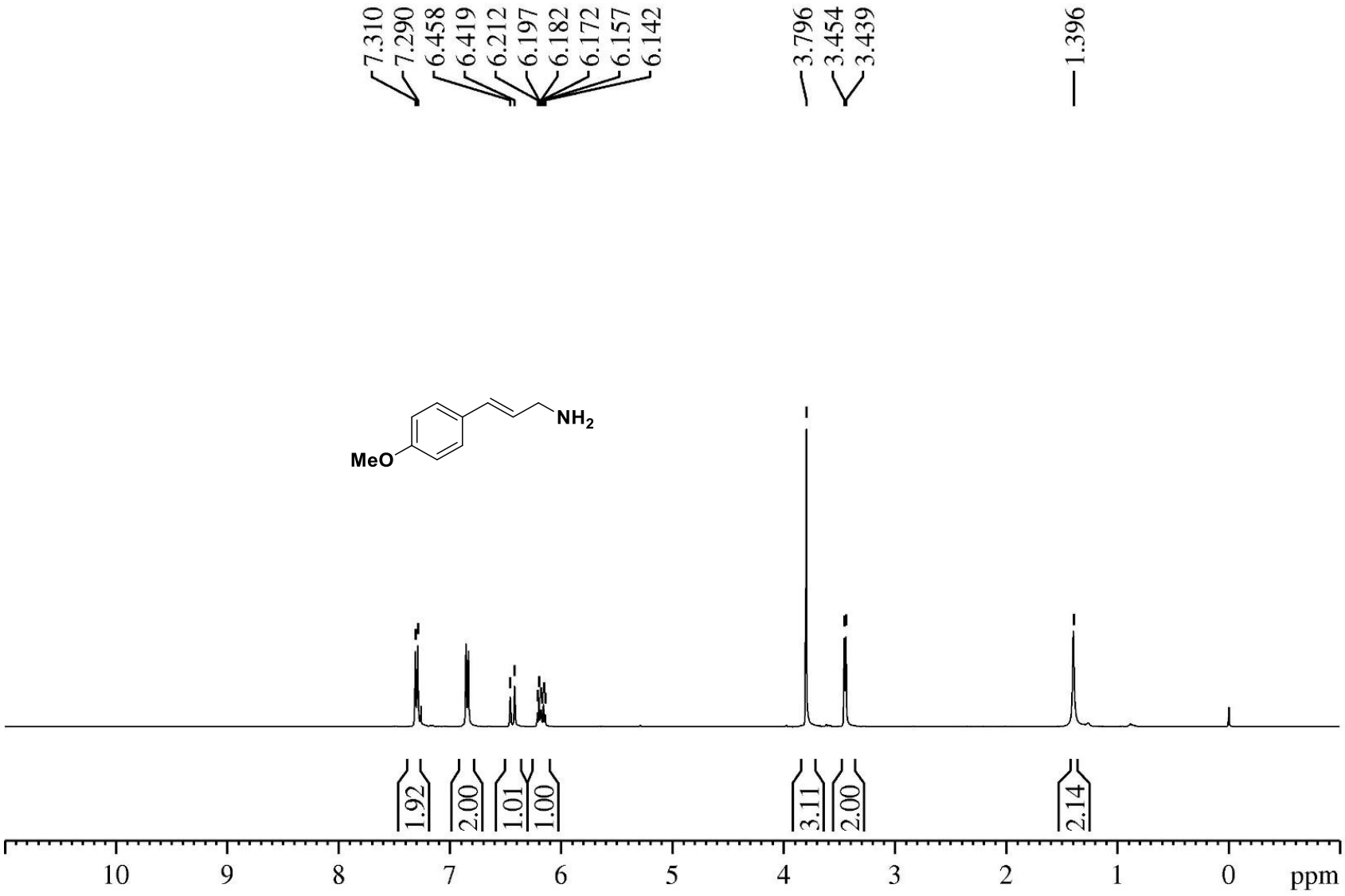
[1]H spectra of compound 4b



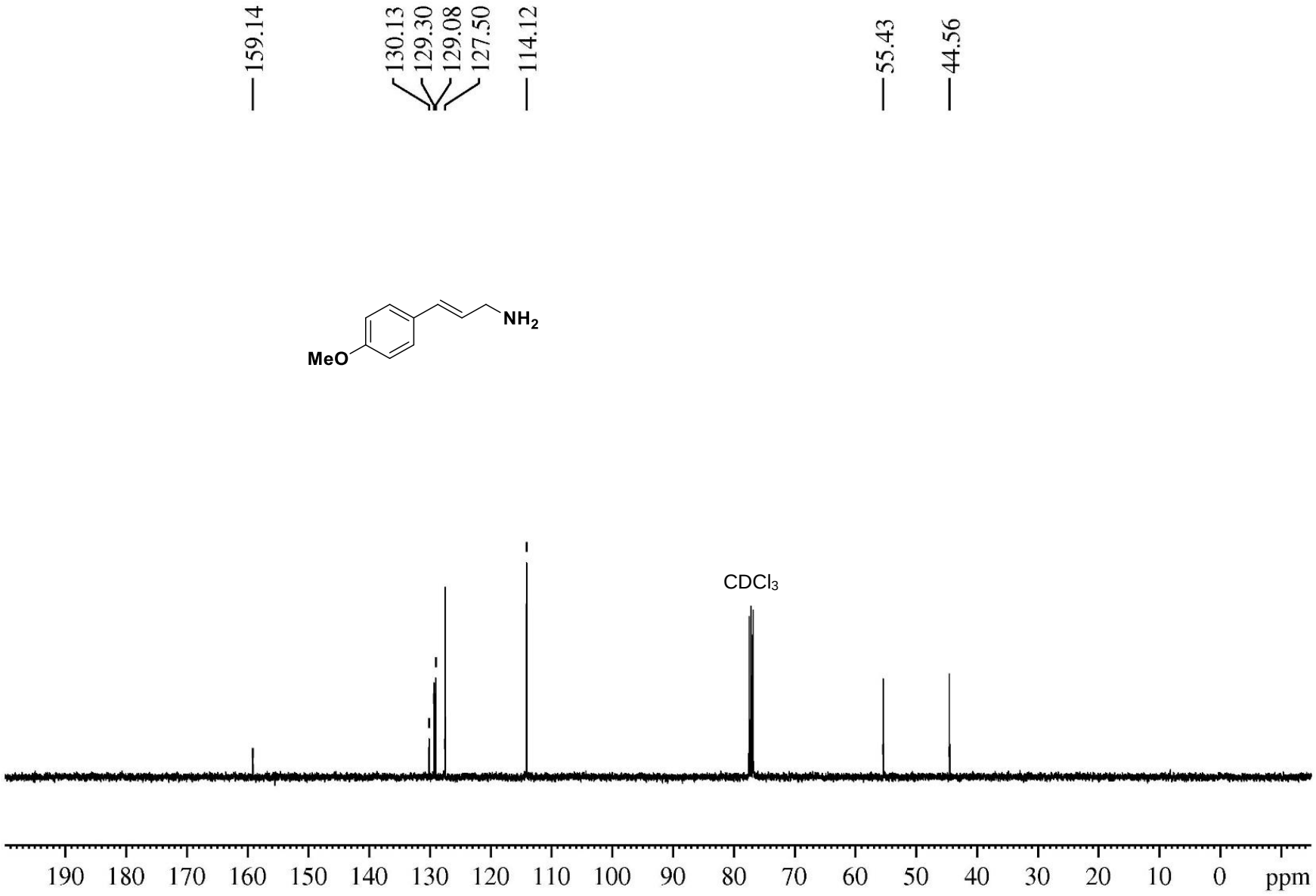
^{13}C spectra of compound 4b



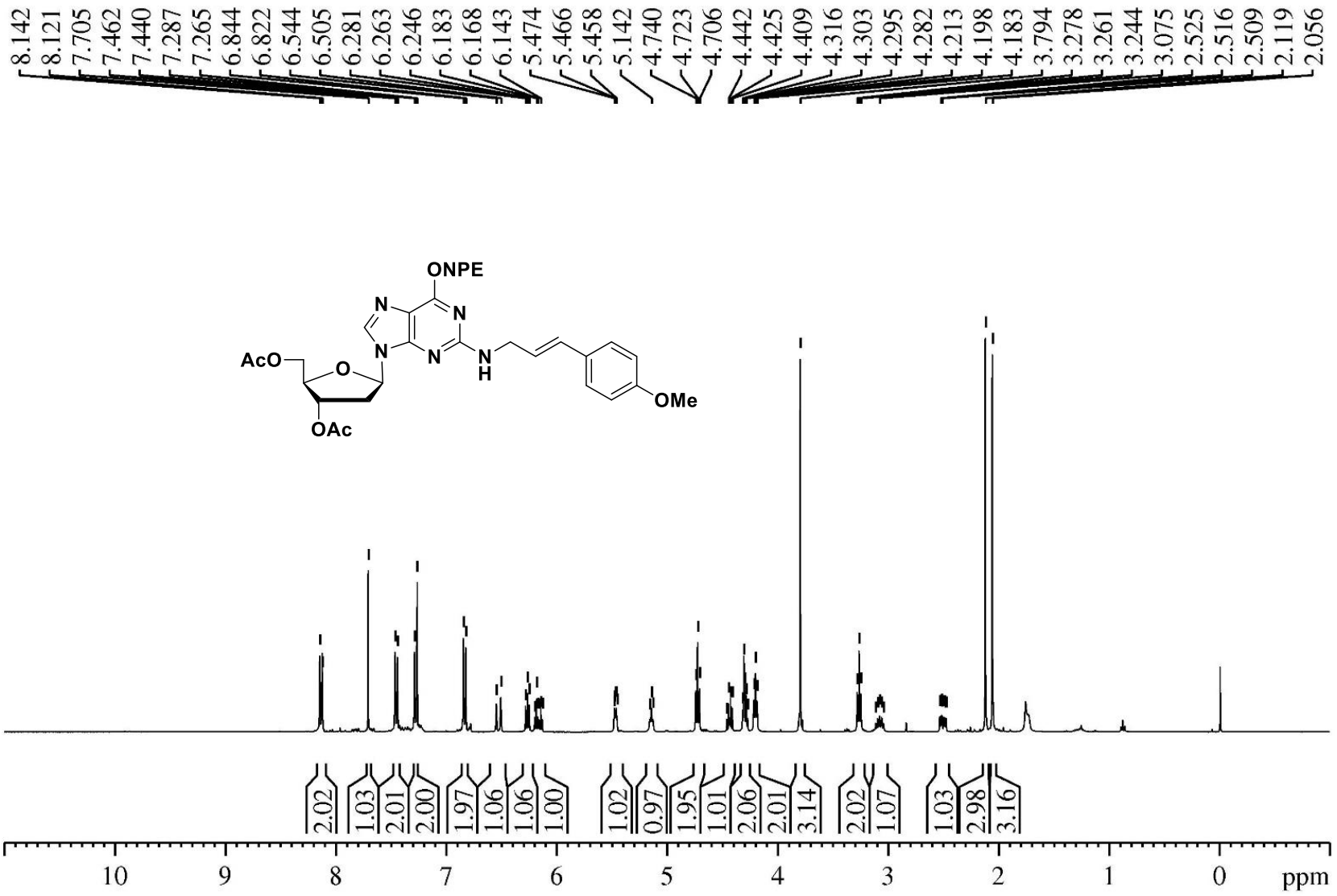
^{1}H spectra of compound 5b



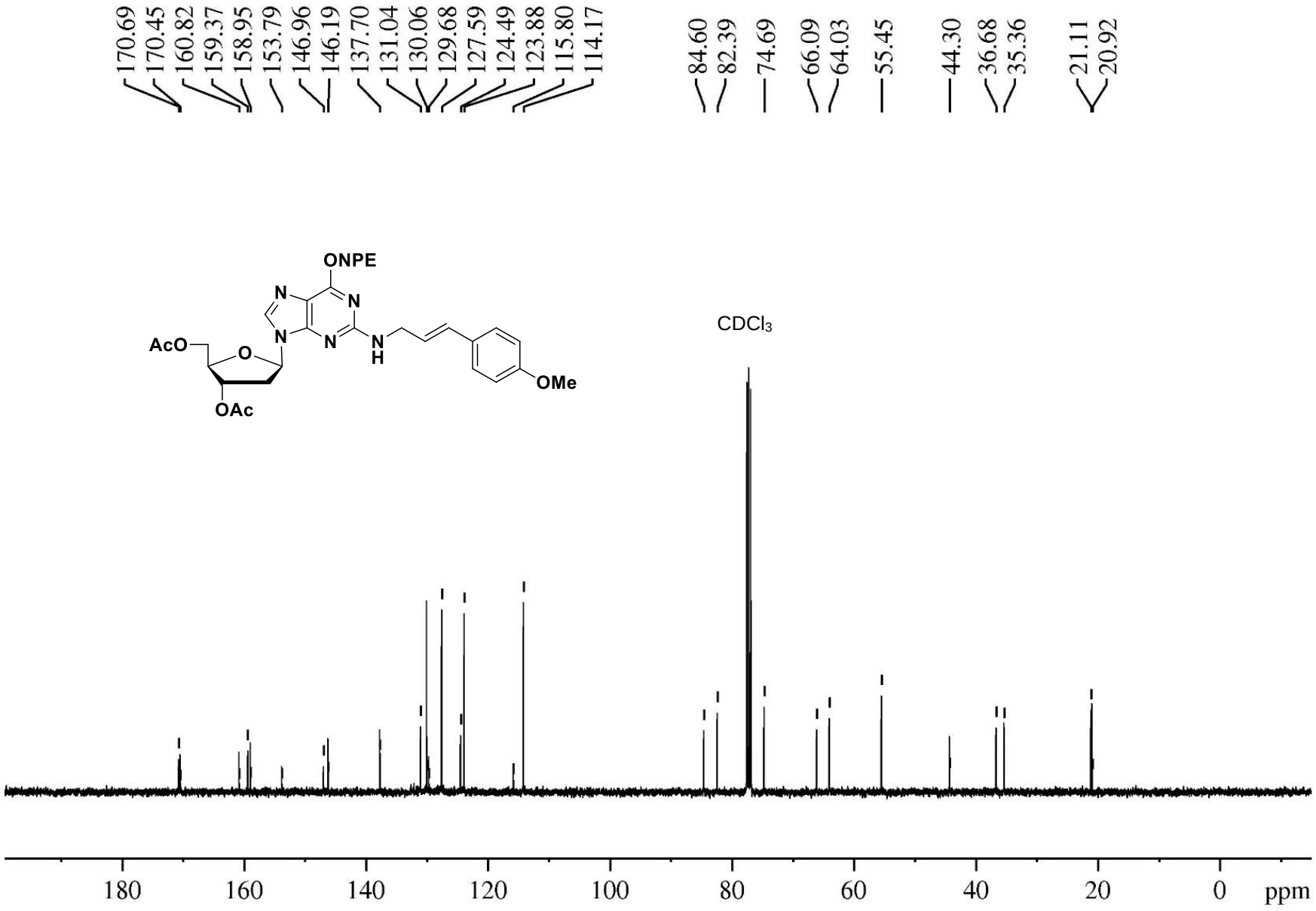
^{13}C spectra of compound 5b



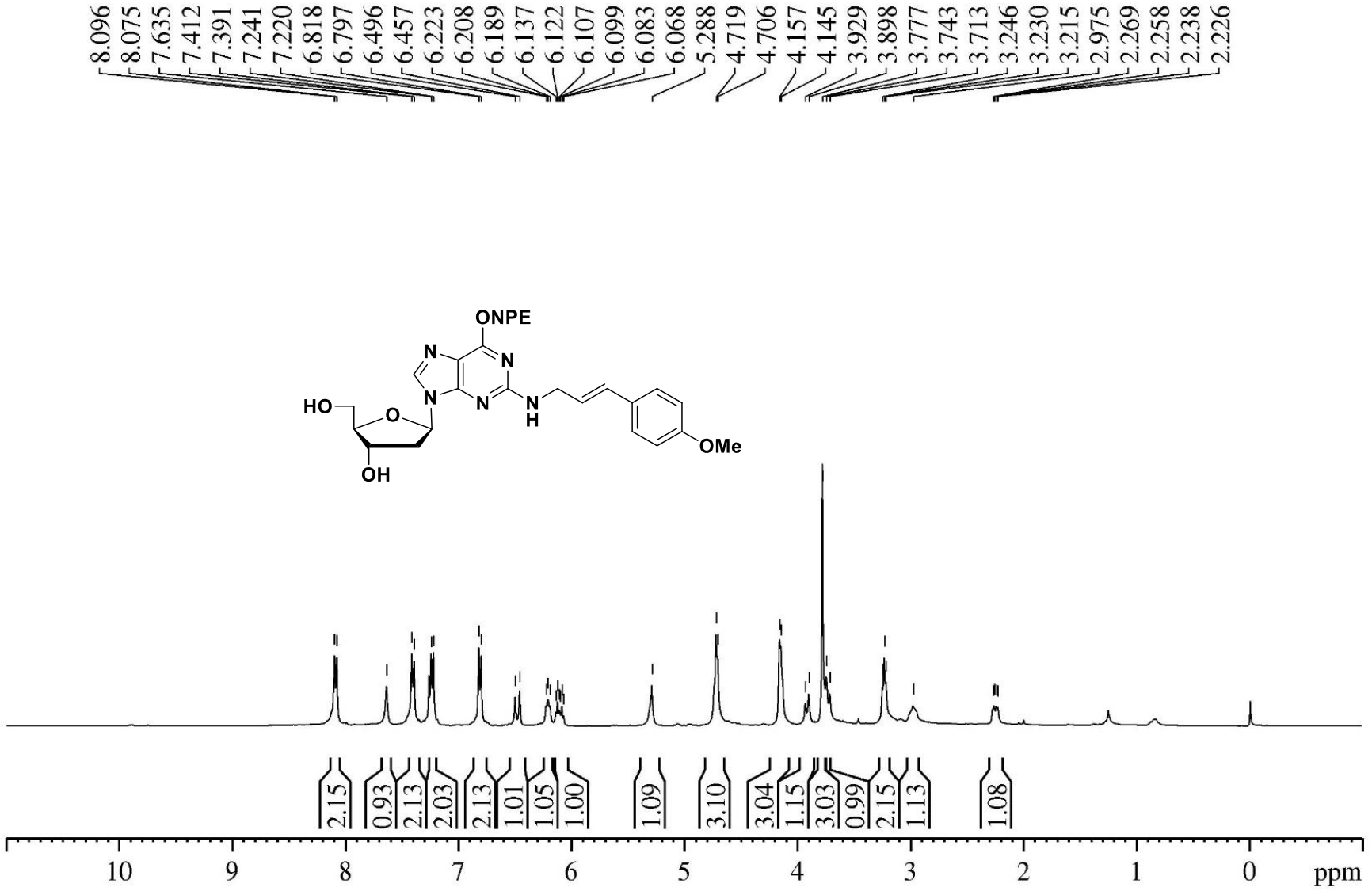
^{1}H spectra of compound 7b



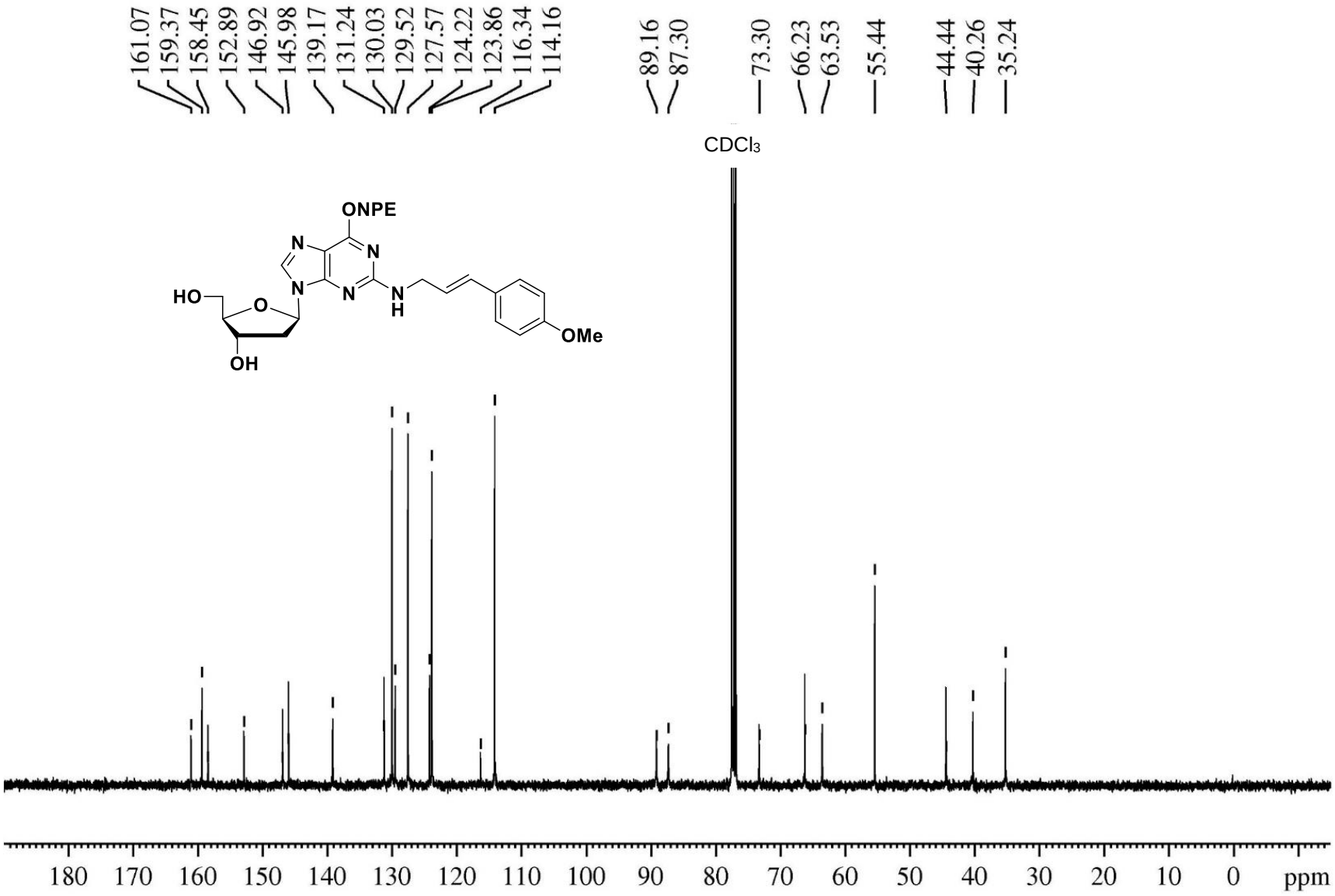
^{13}C spectra of compound 7b



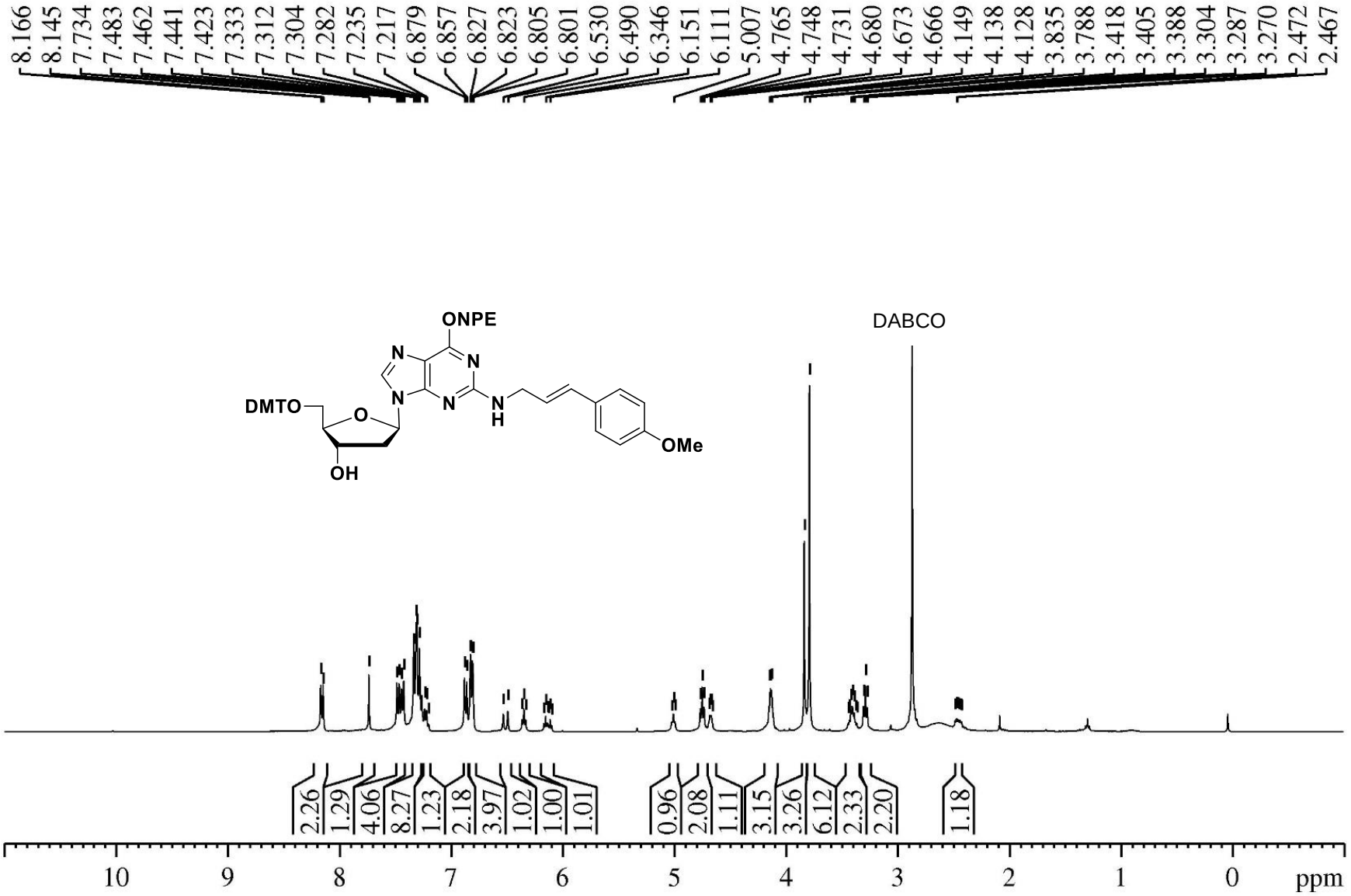
[1]H spectra of compound 8b



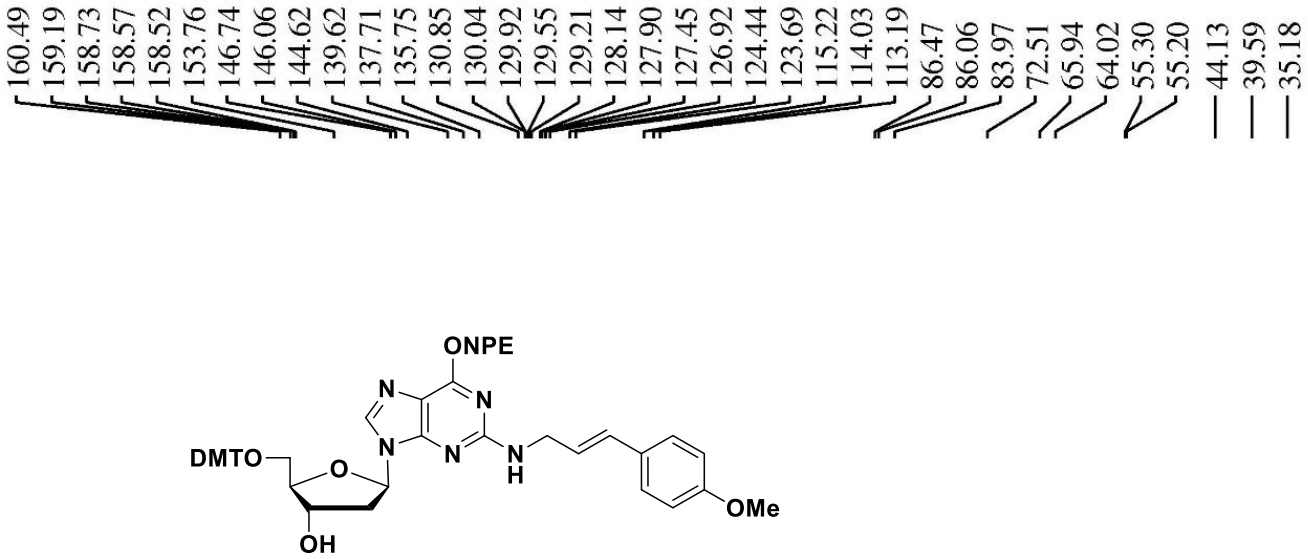
^{13}C spectra of compound 8b

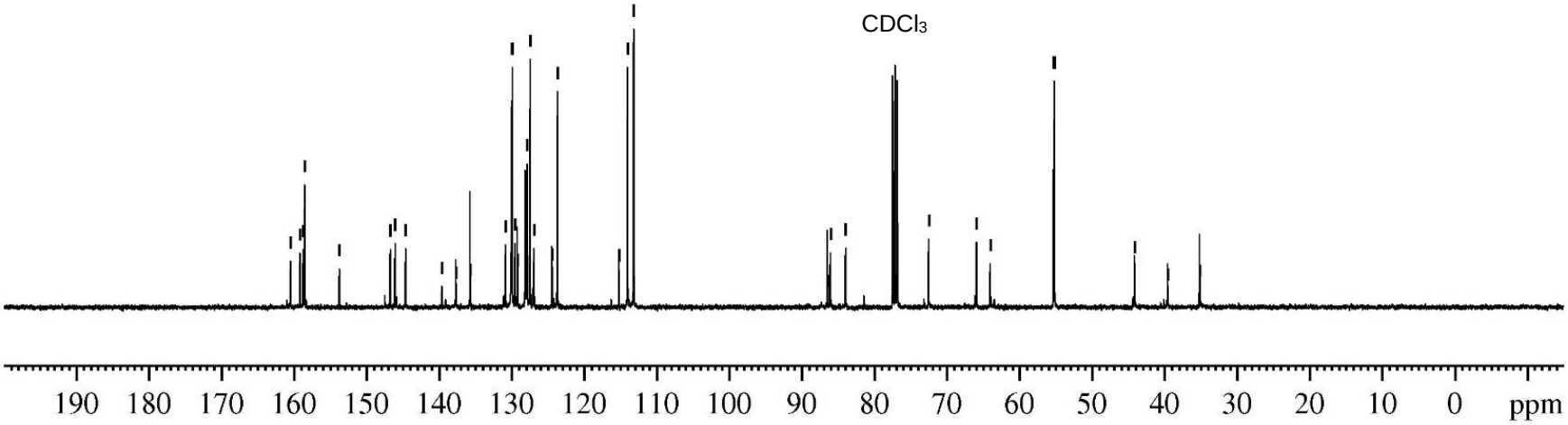


^{1}H spectra of compound 9b

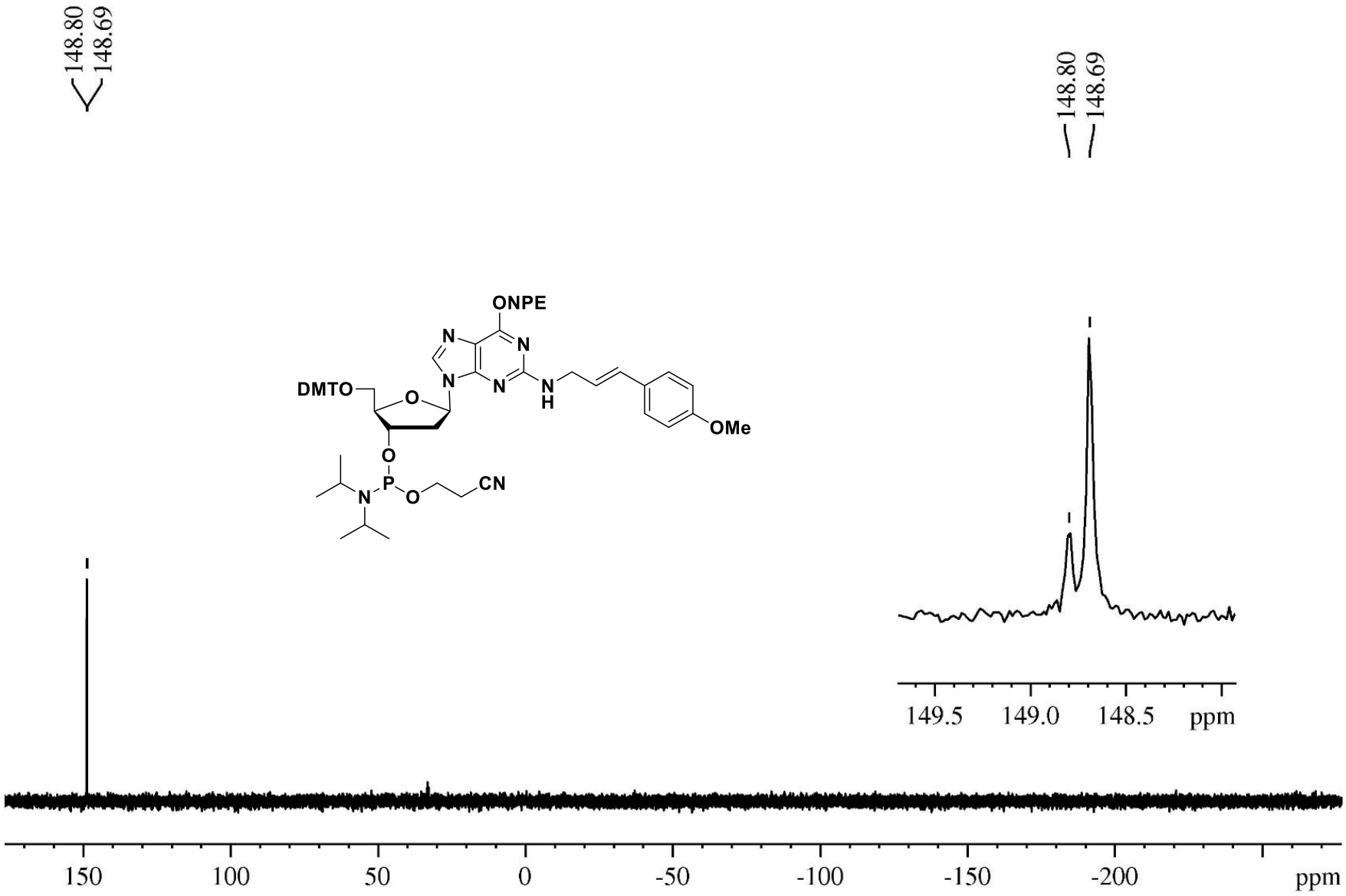


^{13}C spectra of compound 9b

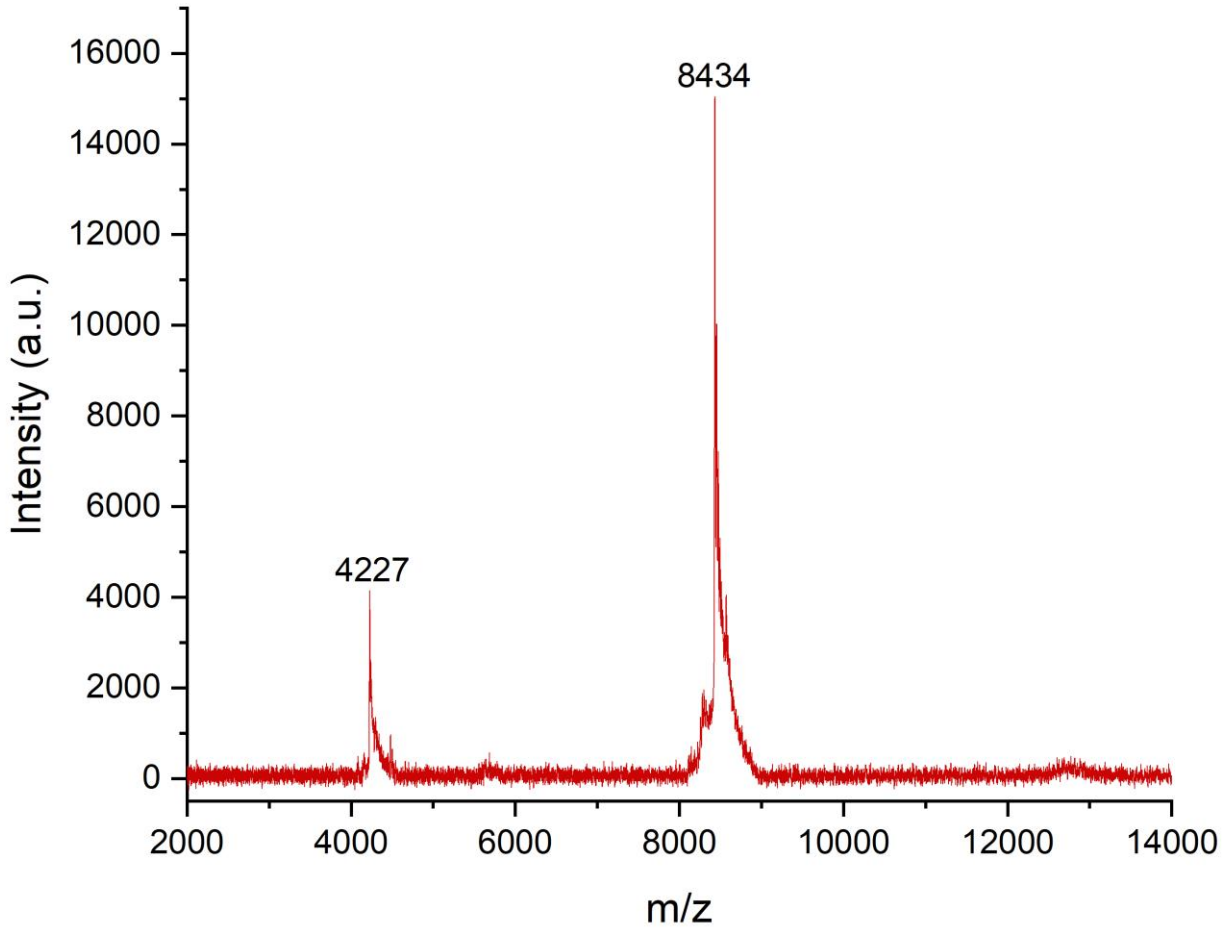




^{31}P spectra of compound 10b



MALDI Spectrum of **T1** (MEG-dG), 5’-CTGGTCACACTXATGCC TACGAGTACG-3’: calc. mass, [M–H]– 8436.4; obs. mass, [M–H]– 8434.0



MALDI Spectrum **T2** (EG-dG), 5’-CTGGTCACACTXATGCC TACGAGTACG-3’: calc. mass, [M–H]– 8406; obs. mass, [M–H]– 8407