Synthesis of 5-dihydroxyboryluridine Phosphoramidite and Its Site-Specific Incorporation into Oligonucleotides as Functional Probe for Thymine DNA Glycosylase

Sam Kavoosi,[§] Debasis Dey[§] and Kabirul Islam*

Department of Chemistry, Pittsburgh, Pittsburgh PA 15260

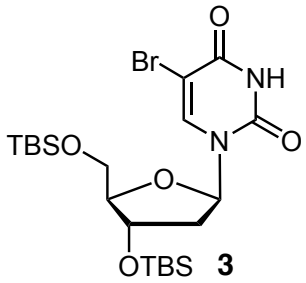
([§]Equal contribution)

1. General materials, methods and equipment

Chemicals: All chemicals were purchased from established vendors (e.g. Sigma-Aldrich, Acros Organics) and used without purification unless otherwise noted. Optima grade acetonitrile was obtained from Fisher Scientific and degassed under vacuum prior to use during HPLC purification. All reactions to prepare uridine analogues were carried out in round bottom flasks and stirred with Teflon®-coated magnetic stir bars under inert atmosphere when needed. Analytical thin layer chromatography (TLC) was performed using EMD 250 micron flexible aluminum backed, UV F_{254} pre-coated silica gel plates and visualized under UV light (254 nm) or by staining with phosphomolybdic acid, ninhydrin or anisaldehyde. Reaction solvents were removed by a Büchi rotary evaporator equipped with a dry ice-acetone condenser. Analytical and preparative HPLC was carried out on an Agilent 1220 Infinity HPLC with diode array detector. Concentration and lyophilization of aqueous samples were performed using Savant Sc210A SpeedVac Concentrator (Thermo), followed by Labconco Freeze-Dryer system.

Proton nuclear magnetic resonance spectra (1H NMR) were recorded on Bruker Ultrashield^{TM} Plus 600/500/400/300 MHz instruments at 24^{o}C. Chemical shifts of ^{1}H, ^{13}C NMR, ^{11}B and ^{31}P spectra are reported as δ in units of parts per million (ppm) relative to tetramethylsilane (δ 0.0) or residual solvent signals: chloroform-d (δ 7.26, singlet), methanol-d_4 (δ 3.30, quintet), and deuterium oxide-d_2 (δ 4.80, singlet). Coupling constants are expressed in Hz. MALDI mass spectra were collected at ultraFlextreme (Bruker) and the data was analyzed using flexAnalysis software. The ESI-MS were recorded on a Q-Exactive^{TM} Thermo Scientific LC-MS with electron spray ionization (ESI) probe.

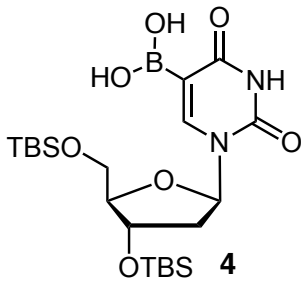
2. Synthesis and characterization of 5-dihydroxyboryldexoyuridine (5boU) phosphoramidite



3

A solution of 5-bromo-2′-deoxyuridine **2** (4.00 g, 13.0 mmol), tert-butyldimethylsilyl chloride (TBDMSCl, 4.40 g, 29.2 mmol) and imidazole (4.00 g, 58.7 mmol) and a pinch of DMAP in DMF (50 mL) was stirred room temperature for 12 h. The reaction mixture was quenched by addition of 50 mL of saturated sodium bicarbonate and then exacted with DCM (3 × 50 mL); evaporated *in vacuo.* The residue was

purified by silica gel column chromatography (20 % ethyl acetate in hexanes) to obtain TBS protected compound **3** (7.0 g, 99%) as white crystalline solid. R_f = 0.35 (ethyl acetate/hexanes, 1:4). ^{1}H-NMR (Chloroform-d, 400 MHz): δ = 8.89 (bs, 1 H), 8.07 (s, 1H), 6.29 (dd, *J*=7.70, 5.78 Hz, 1 H), 4.37-4.43 (m, 1 H), 3.98 (q, *J*=2.15 Hz, 1 H), 3.90 (dd, *J*=11.50, 2.26 Hz, 1 H), 3.76 (dd, *J*=11.52, 2.12 Hz, 1 H), 2.31 (ddd, *J*=13.15, 5.77, 2.49 Hz, 1 H), 1.96-2.05 (m, 1 H), 0.934 (s, 9 H), 0.890 (s, 9 H), 0.144 (s, 3 H), 0.134 (s, 3 H), 0.081 (s, 3 H), 0.072 (s, 3H), ppm; ^{13}C-NMR (Chloroform-d, 100 MHz): δ = 159.03, 149.58, 139.59, 96.98, 88.60, 88.02, 72.59, 63.17, 42.17, 26.21, 25.88, 18.62, 18.14, -4.50, -4.70, -5.15, -5.23 ppm; HRMS (ESI) calcd. for $C_{21}H_{40}O_5N_2{}^{79}BrSi_2$ $[M+H]^+$: 535.1653; found: 535.1632.



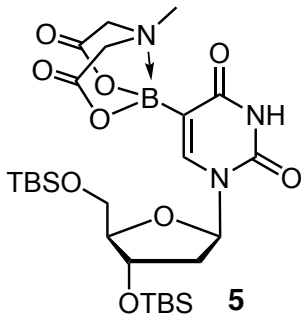
4

To a stirred solution of TBS protected 5-bromo-2′-deoxyuridine **3** (5.00 g, 9.33 mmol) and tributyl borate (6.55 mL, 24.27 mmol)[1] in THF (30 mL) at -78°C was added *n*-BuLi (15.17 mL, 24.27 mmol, 1.6 M) very slowly (0.3 mL/min) through a syringe pump. The reaction was maintained at that temperature for an hour and then stirred at room temperature for 36 hours. The reaction was quenched by addition of saturated aqueous NH_4Cl (50mL) solution and extracted with ethyl acetate (3 x 50mL). Crude product was purified on a silica gel column using ethyl acetate/hexanes (10:1 - 4:1 v/v). Fractions containing the titled product were combined and evaporated to give the required compound **4** (~2 g, 40%). However, majority of the remaining ~2.1 g material eluted as mixture with TBS protected uridine (debrominated starting material) and the unreacted starting material. We proceeded to the next step as further purification proved to be challenging at this stage.

^{1}H NMR (400 MHz, CHLOROFORM-*d*) δ = 0.10 (dd, *J*=4.89, 2.38 Hz, 12 H) 0.91 (d, *J*=1.76 Hz, 18 H) 2.05 (ddd, *J*=13.43, 7.91, 5.77 Hz, 1 H) 2.33 (ddd, *J*=13.18, 5.65, 2.26 Hz, 1 H) 3.77 (d, *J*=3.76 Hz, 2 H) 3.96 - 4.02 (m, 1 H) 4.39 - 4.48 (m, 1 H) 6.21 (dd, *J*=8.16, 5.65 Hz, 1 H) 7.93 (s, 1 H) 8.11 (s, 1 H) ppm.

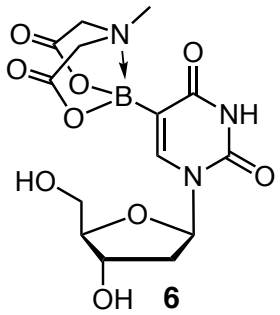
^{13}C NMR (100 MHz, CHLOROFORM-*d*) δ = -5.62, -5.22, -4.86, -4.73, 18.01, 18.37, 24.65, 24.79, 25.74, 25.99, 41.15, 62.92, 72.48, 83.89, 85.90, 88.10, 148.82, 150.08, 163.42 ppm.

ESI-HRMS m/z $[M+H]^+$ calcd for $C_{21}H_{42}O_7N_2BSi_2$: 501.2624, found 501.2624 (1.21ppm).



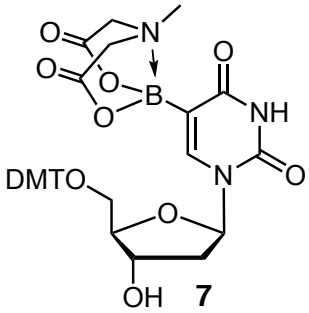
5

For the initial 5-gram reaction: Methyliminodiacetic acid (2.94 g, 19.9 mmol)[2] and the mixture of the compound obtained in the previous step were dissolved in DMSO/Benzene (30:4 mL) and heated to reflux with a Dean-Stark apparatus. Reaction was left on heat for 36 hrs, then brought to room temperature. Water (20 mL) was added and stirred for 20 mins, EtOAc was added and the organic phase was isolated and concentrated, and column chromatographed (2%-5% MeOH in DCM) to yield the MIDA and TBS protected uridine boronate **5** (2.27 g, 40.0% over two steps). ^{1}H-NMR (Chloroform-d, 400 MHz): δ = 9.45 (bs, 1 H), 7.76 (s, 1 H), 6.22 (dd, *J*=8.22, 5.74 Hz, 1 H), 4.37-4.44 (m, 1 H), 4.27 (d, *J*=16.16 Hz, 1 H), 3.86-4.07 (m, 4 H), 3.71-3.83 (m, 2 H), 2.81 (s, 3 H), 2.29 (ddd, *J*=13.18, 5.70, 2.02 Hz, 1 H), 2.00-2.16 (m, 1 H), 0.889 (s, 9 H), 0.886 (s, 9 H), 0.092 (s, 6 H), 0.082 (s, 3 H), 0.078 (s, 3 H), ppm;^{13}C-NMR (CHLOROFRM-d, 100 MHz): δ = 167.67, 167.51, 166.84, 150.59, 145.30, 88.33, 86.10, 72.76, 63.43, 63.35, 46.52, 41.11, 40.92, 26.10, 25.90, 18.53, 18.13, -4.55, -4.69, -5.19, -5.45, ppm; ^{11}B NMR (Chloroform-d, 160 MHz): δ = 12.24 ppm; HRMS (ESI) calcd for $C_{26}H_{47}O_9N_3{}^{11}BSi_2$ $[M+H]^+$: 612.2938; found: 612.2947.

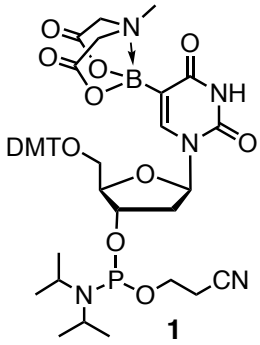


6

To a solution of the MIDA and TBS protected uridine boronate **5** (2.76 g, 4.52 mmol) and pyridine (1.68 mL, 20.8 mmol) in EtOAc (45 mL) in a polypropylene tube was added 70% HF-pyridine (1.79 mL, 19.9 mmol). The solution was stirred for 24 hrs and then quenched with 2 mL of methoxytrimethylsilane. The reaction was concentrated and chromatographed (15-30% MeOH in DCM) to yield the required diol **6** (1.31 g, 75.3%). ^{1}H-NMR (DMSO-d_6/Deuterium Oxide, 400 MHz): δ = 7.68 (s, 1 H), 6.15 (t, *J*=6.78 Hz, 1 H), 4.17-4.37 (m, 3 H), 3.99 (d, *J*=16.85 Hz, 2 H), 3.78 (q, *J*=3.92 Hz, 1 H), 3.49-3.53 (m, 2 H), 2.72 (s, 3 H), 2.07-2.20 (m, 2 H), ppm;^{13}C-NMR (DMSO-d_6/Deuterium Oxide, 100 MHz): δ = 169.16, 169.09, 166.33, 150.83, 144.29. 87.44, 84.55, 70.58, 62.75, 61.41, 46.82, ppm; ^{11}B NMR (DMSO-d_6/Deuterium Oxide, 128 MHz): δ = 11.81 ppm; HRMS (ESI) calcd for $C_{14}H_{19}O_9N_3{}^{11}B$ $[M+H]^+$: 384.1208; found: 384.1222.

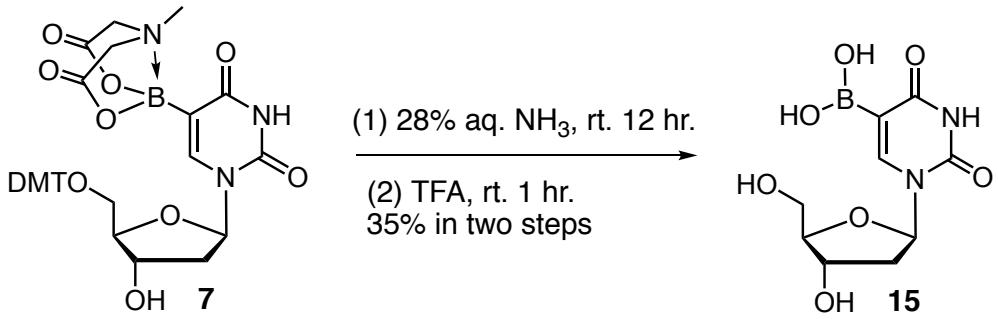


A mixture of the diol **6** (1.30 g, 3.39 mmol) and dimethoxytrityl chloride (DMTrCl) (1.26 g, 3.73 mmol) was dried on vacuum for 1 hour. To the dried material was added pyridine (10 mL) and the reaction was stirred overnight. The reaction was quenched with MeOH (0.5 mL) and concentrated. The residue was chromatographed (5-15% MeOH in DCM w/ 0.1% TEA). Residual starting material was recovered, and the reaction was repeated to yield the DMTr protected uridine boronate **7** (0.682 g, 29.3%). ^{1}H-NMR (DMSO-d_6/Deuterium Oxide, 400 MHz): δ = 7.51 (s, 1 H), 7.38 (d, *J*=7.48 Hz, 2 H), 7.17-7.33 (m, 7 H), 6.85-6.92 (m, 4 H), 6.09 (t, *J*=6.46 Hz, 1 H), 4.24-4.37 (m, 2 H), 4.07 (q, *J*=5.15 Hz, 1 H), 3.99 (dd, *J*=16.85, 7.40 Hz, 1 H), 3.85-3.92 (m, 1 H), 3.73 (s, 3 H), 3.72 (s, 3 H), 2.69 (s, 3 H), 2.20 (t, *J*=5.98 Hz, 2 H), ppm;^{13}C-NMR (DMSO-d_6/Deuterium Oxide, 100 MHz): δ = 169.06, 168.95, 166.35, 158.14, 150.61, 144.91, 144.19, 129.84, 129.79, 127.95, 127.80, 126.76, 113.30, 112.89, 85.77, 85.32, 70.56, 63.99, 62.74, 55.10, 46.63, ppm; ^{11}B NMR (DMSO-d_6/Deuterium Oxide, 128 MHz): δ = 12.91; HRMS (ESI) calcd for $C_{35}H_{37}O_{11}N_3{}^{11}B$ $[M+H]^+$: 686.2515; found: 686.2506.

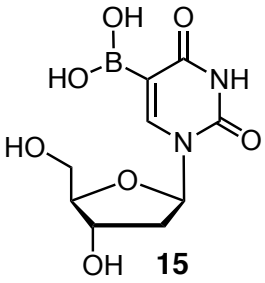


The starting DMTr protected uridine boronate **7** (0.682 g, 0.994 mmol) was dissolved in DCM (20 mL) and DIPEA (1.73 mL, 9.94 mmol), and immediately the solution was degassed and flushed with N_2. Then 2-Cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite (0.555 mL, 2.49 mmol) was added. After 1.5 hours the reaction was concentrated and chromatographed (30-70% ACN in DCM w/ 0.1% TEA) to yield 5boU phosphoramidite **1** (0.631 mg, 71.6%) as a mixture of diastereomers. ^{1}H-NMR (Acetonitrile-d_3, 400 MHz): δ = 9.17 (bs, 1 H), [7.61 (s, 0.5 H), 7.59 (s, 0.5 H)], 7.40-7.49 (m, 2 H), 7.16-7.38 (m, 7 H), 6.80-6.95 (m, 4 H), 6.11 (q, *J*=6.88 Hz, 1 H), 4.31.4.46 (m, 1 H), 3.92-4.13 (m, 5 H), 3.68-3.84 (m, 7 H), 3.46-3.68 (m, 3 H), 3.20-3.38 (m, 2 H), 2.71 (s, 3 H), 2.62 (t, *J*=5.86 Hz, 1 H), 2.52 (t, *J*=5.92 Hz, 1 H), 2.25-2.48 (m, 2 H), 1.09-1.26 (m, 9 H), 1.04 (d, *J*=6.72 Hz, 1 H), ppm; ^{13}C-NMR (Acetonitrile -d_3, 125 MHz): δ = 169.29, 169.26, 169.24, 169.22, 167.45, 159.72, 151.56, 151.55, 146.1, 145.79, 145.73, 145.67, 145.61, 136.92, 136.89, 136.86, 136.85. 131.08, 129.03, 128.91, 127.82, 119.56, 119.41, 114.19, 87.35, 87.31, 86.87, 86.76, 86.07, 86.03, 85.74, 85.68, 74.57, 74.43, 74.28, 74.15, 64.62, 64.54, 63.96, 59.64, 59.55, 59.49, 59.40, 55.94, 55.93, 47.61, 44.15, 44.05, 39.45, 39.42, 39.37, 24.97, 24.94, 24.91, 24.87, 24.84, 21.09,

21.04, 21.01, 20.95, ppm; ^{31}P-NMR (Acetonitrile -d_3, 162 MHz): δ = 148.12, 147.91 ppm. ^{11}B NMR (Acetonitrile -d_3, 160 MHz): δ = 10.92; HRMS (ESI) calcd for $C_{44}H_{54}O_{12}N_5BP$ $[M+H]^+$: 886.3594; found: 886.3615.



Scheme S1: Synthesis of 5-dihydroxyboryl-deoxyuridine (5boU) **15**



The DMTr protected uridine boronate **7** (50.0 mg, 0.073 mmol) was dissolved in methanol (1 mL). Then a solution of 28% aqueous ammonia (2 mL) was added and the reaction was stirred overnight. The ammonia was removed by *rotovap* and additional methanol (4 mL) was added to dissolve the resultant precipitate. TFA (0.5 mL) was then added to the solution and stirred for 1 hr, after which the reaction was concentrated down. To the residue was added water (5 mL) and successively washed with ether (x3). The aqueous layer was syringe filtered through a *Millex-GP* Filter Unit (PES Membrane, 0.22 μm), and purified using HPLC. Column: *X-Bridge* BEH C18 OBD Prep Column 130 Å, 5 μm, 10 mm x 250 mm. Solvent A: Water Solvent B: Acetonitrile, a linear gradient over 15 mins from solvent A to B at 4 mL/min was used. The peak at retention time 6.75 mins (λ 260 nm) was collected to yield 5-dihydroxyboryldeoxyuridine **15** (7.0 mg, 35%). ^{1}H NMR (Deuterium Oxide, 500 MHz): δ = 8.17 (s, 1 H), 6.25 (t, *J*=6.63 Hz, 1 H), 4.47 (dt, *J*=8.50, 3.20 Hz, 1 H), 4.01-4.06 (m, 1 H), 3.84 (dd, *J*=12.50, 3.45 Hz, 1 H), 3.76 (dd, *J*=12.50, 5.00 Hz, 1 H), 2.34-2.45 (m, 2 H), ppm; ^{13}C-NMR (Deuterium Oxide, 125 MHz): δ = 169.40, 151.49, 148.71, 86.72, 85.65, 70.29, 61.02, 38.88, ppm; ^{11}B NMR (Deuterium Oxide, 160 MHz): δ = 19.39 ppm; HRMS (ESI) calcd for $C_9H_{12}O_7N_2{}^{11}B$ $[M-H]^-$: 271.0738; found: 271.0722.

3. Synthesis and purification of oligonucleotides. DNA Oligonucleotides (Table S1) were synthesized using standard DNA phosphoramidite monomers (Glen Research) in an EXPIDITE Nucleic Acid Synthesis System (PerSeptive Biosystems). The phosphoramidites of natural deoxynucleotides (A, T, G and C) were purchased from Glen Research and 6-Fluorescein (FAM) Phosphoramidite from Millipore-Sigma. Standard synthesis protocols were followed. To ensure good coupling elongated (4 min 30 sec) coupling times were applied for the coupling of modified bases and for standard bases (2 min) normal coupling was applied. FAM group was added at the 5'position of the oligonucleotides. Both FAM-labeled and unlabeled oligomers were removed from the resin and deprotected with ammonium hydroxide (28% v/v) at room temperature for overnight. DMTr containing DNA was purified initially through a Poly Pak II purification cartridge (Glen Research). FAM-labeled and Poly Pak purified DNA were then purified through HPLC on a C-18 column with the following gradient: Solvent A: 0.1 M TEAA pH 7, Solvent B: Acetonitrile; 0 min 5% B, 10 min 40% B, 15 min 100% B with a flow rate of 4 mL/min. The appropriate fractions were collected, for FAM labeled DNA only the bands containing absorbances at both 260 and 495 nm were collected. All fractions were concentrated down with a SpeedVac concentrator, and if required the DNA was also gel purified through agarose and extracted with QIAEX II Gel Extraction Kit. All oligomers were re-dissolved in either nuclease free water or duplex buffer (100 mM KOAc and 30 mM HEPES pH 7.5, Integrated DNA Technology) and were confirmed by ESI LC-HRMS.

4. Expression and purification of TDG in *E. coli*. The N-terminal 6xHis-tagged human TDG bacterial expression construct pET28 kanamycin-resistant vector was obtained from Alexander C. Drohat laboratory, University of Maryland.[3,4] The wild type TDG plasmid was transformed into *Escherichia coli* BL21 star (DE3) competent cells (Invitrogen) using pET28 kanamycin-resistant vector. A single colony was picked up and grown overnight at 37°C in 10 mL of Luria-Bertani (LB) broth in the presence of 50 μg/mL kanamycin. The culture was diluted 100-fold and allowed to grow at 37°C to an optical density (OD_{600}) of 0.8, and protein expression was induced overnight at 17°C with 0.25 mM IPTG in an Innova 44® Incubator shaker (New Brunswick Scientific). The cells were harvested and resuspended in buffer containing 50 mM Tris (pH 8.0), 150 mM NaCl, 10 % glycerol, 5 mM β-mercaptoethanol, 25 mM imidazole, DNase, lysozyme and protease inhibitor tablets. The cells were lysed by pulsed sonication (Qsonica-Q700) and centrifuged at 13,000 rpm for 50 mins at 4°C. The soluble extracts were incubated with Ni-NTA agarose resin

(Thermo) according to manufacturer's protocol. The beads were washed with 20 column volumes of wash buffer (50 mM Tris pH 8.0, 150 mM NaCl, 10 % glycerol, 5 mM β-mercaptoethanol and 25 mM imidazole). The protein was eluted with 50 mM Tris pH 8.0, 150 mM NaCl, 5 mM β-mercaptoethanol and 400 mM imidazole. Eluted protein was subjected to further purification by gel filtration chromatography (Superdex-75) using AKTA pure FPLC system (GE healthcare) with buffer containing 50 mM Tris pH 8.0, 150 mM NaCl and 10 % glycerol. The purified protein was concentrated using Amicon-Ultra 3K centrifugal filter device (Merck Millipore Ltd.). The concentration of the protein was determined using Bradford assay kit (BioRad Laboratories) with BSA as standard. The concentrated proteins were aliquoted and stored at -80ºC.

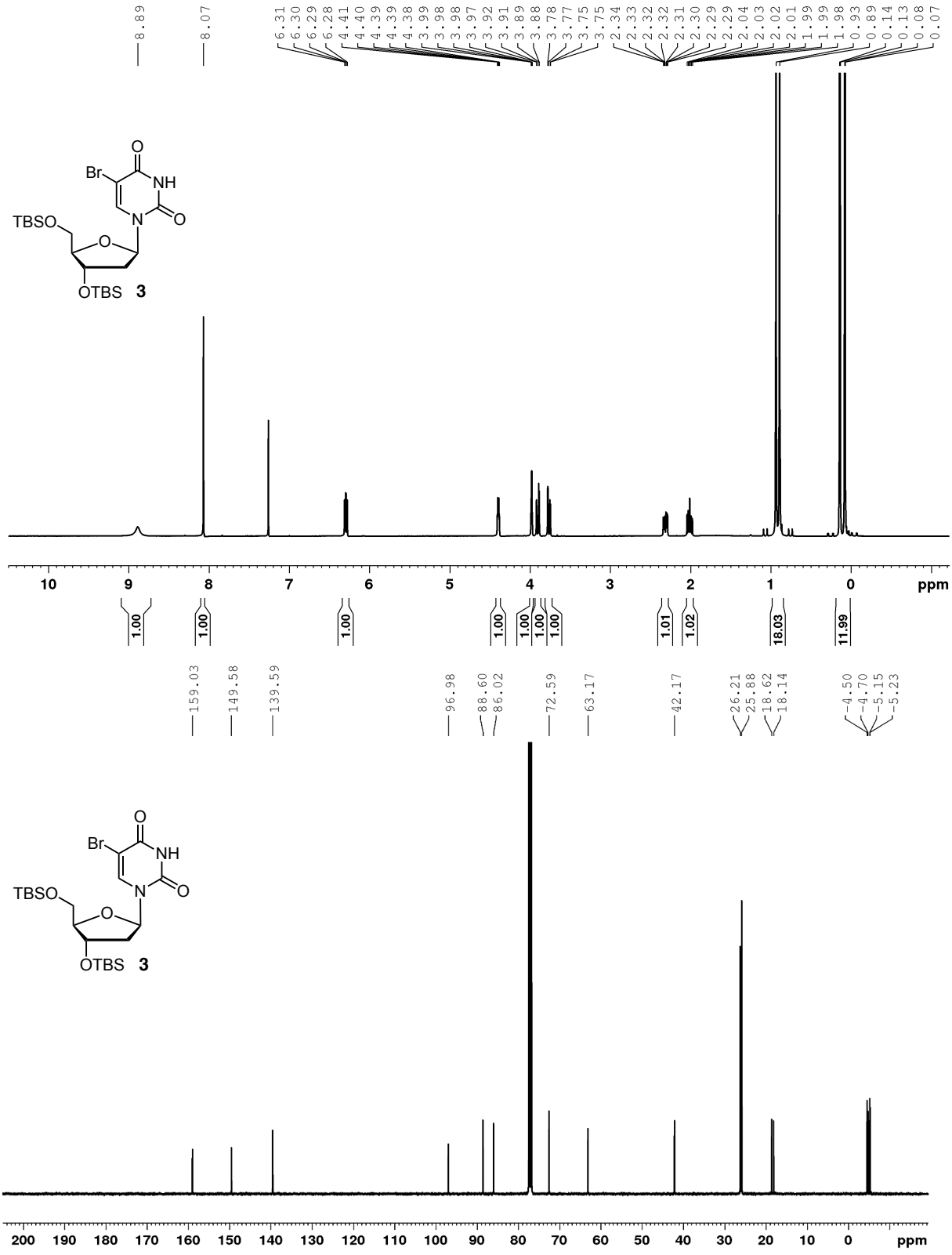
5. Methods for TDG Assay. The substrate DNAs were first duplexed using Eppendorf Thermomixer C with the following parameters in duplex buffer (100 mM KOAc and 30 mM HEPES pH 7.5): (1) 50°C for 3 mins at 1000 rpm, (2) Cool to room temperature over 1 hr. For *in vitro* enzymatic activity assays, 0.125 μM of double-stranded fluorescein labeled 32-nt DNA containing T or 5boU modification, mismatched to G, was incubated with 5 μM hTDG in a 10 μL assay containing 10 mM Tris-HCl pH 8.0, 1 mM EDTA, and 0.05% BSA.[5] The samples were incubated for 1 hr. at 37°C. The reactions were quenched, and the abasic sites cleaved with 2 μL of 1 M NaOH that was heated to 95°C for 10 min. Then the loading buffer was added which contained 10 μL of formamide, 1 μL of 100 mM EDTA, and 2 μL of Novex™ TBE-Urea Sample Buffer (2X) was added and heated to 95°C for 10 mins. 25 μL of the final solution was loaded onto a 15%- acrylamide/24%-formamide urea gel and ran at 200 V in TBE (1X) buffer for 1 hr. Gels were immediately imaged on Bio-Rad ChemiDoc™ XRS+platform. Similar assay protocol was employed for double-stranded fluorescein labeled 22-nt DNA containing 5boU modification mismatched to G.

For *in vitro* enzymatic inhibition assay, 2.5-5.0 μM TDG was incubated with 50 μM inhibitor in a 10 μL assay containing 10 mM Tris-HCl Ph 8.0, 1 mM EDTA, and 0.1% BSA. The samples were incubated on ice for 30 mins, then 0.125 μM of double-stranded fluorescein labeled 32-nt DNA containing T:G mismatch was added and the assay was placed at 37°C for 1 hr. The reactions were quenched, and the abasic sites cleaved with 1 μL of 1 M NaOH that was heated to 95°C for 10 min. Then the loading buffer was added which contained 10 μL of formamide, 1 μL of 100 mM EDTA, and 2 μL of Novex™ TBE-Urea Sample Buffer (2X) was added and heated to 95°C for 10

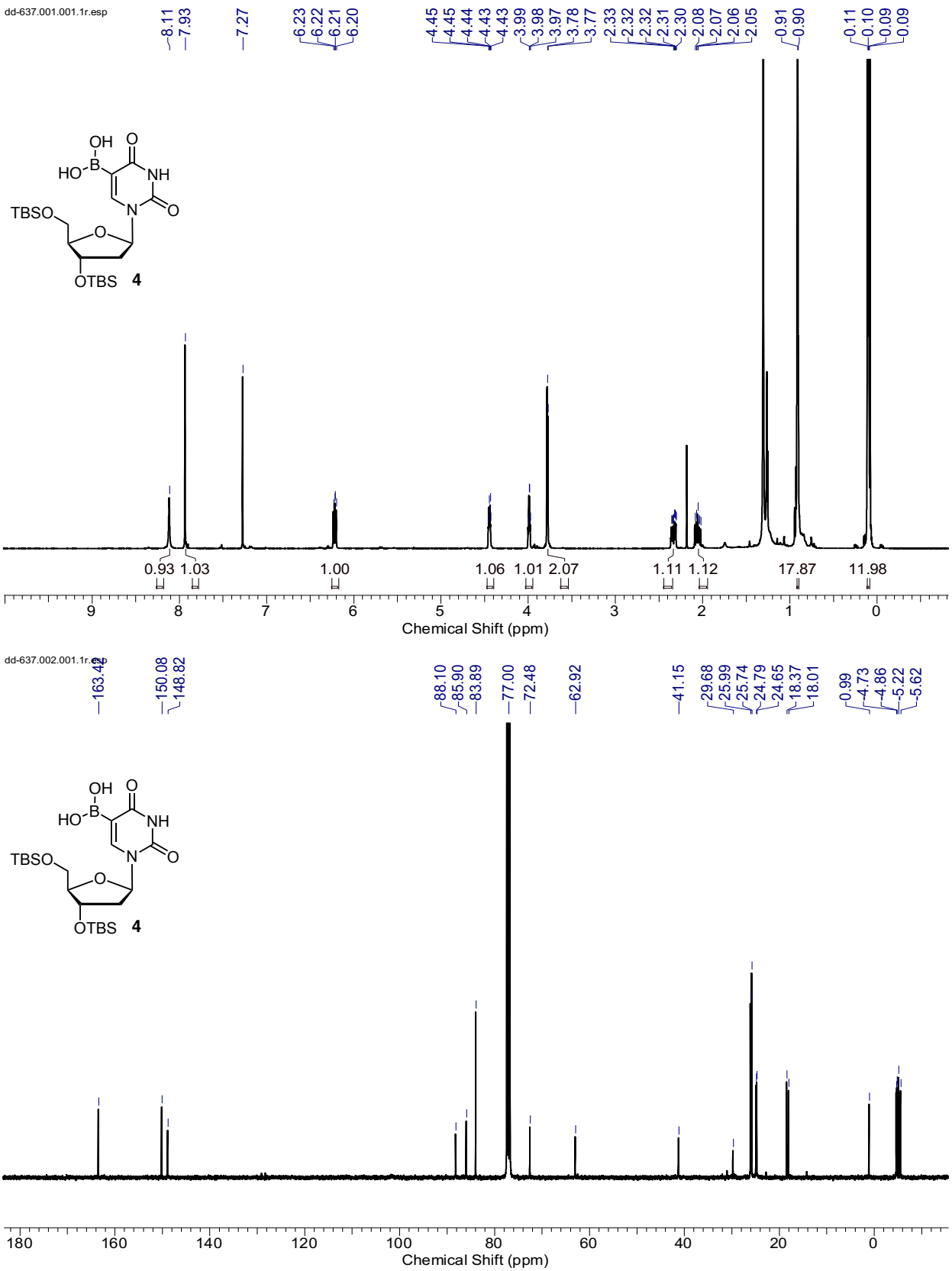
mins. 25 µL of the final solution was loaded onto a 15%- acrylamide/24%-formamide urea gel and ran at 200 V in TBE (1X) buffer for 1 hr. Gels were immediately imaged on Bio-Rad ChemiDoc™ XRS+ platform.

For electrophoretic mobility shift assay (EMSA), in a 10 µL 0.125 µM 32-nt FAM labeled dsDNA containing either T:G or bU:G mismatch was incubated with 10 mM Tris-HCl pH 8, 1 mM EDTA, 5% glycerol, and 5 µM TDG at room temperature. After 30 mins, 2 uL of 6x orange loading dye (Thermo Scientific) was added and the samples were loaded onto a 5% Native-Page containing 0.5X TBE. The samples were run at 4°C at 150V in 0.5X TBE for 30 mins. Gels were immediately imaged on Bio-Rad ChemiDoc™ XRS+ platform.

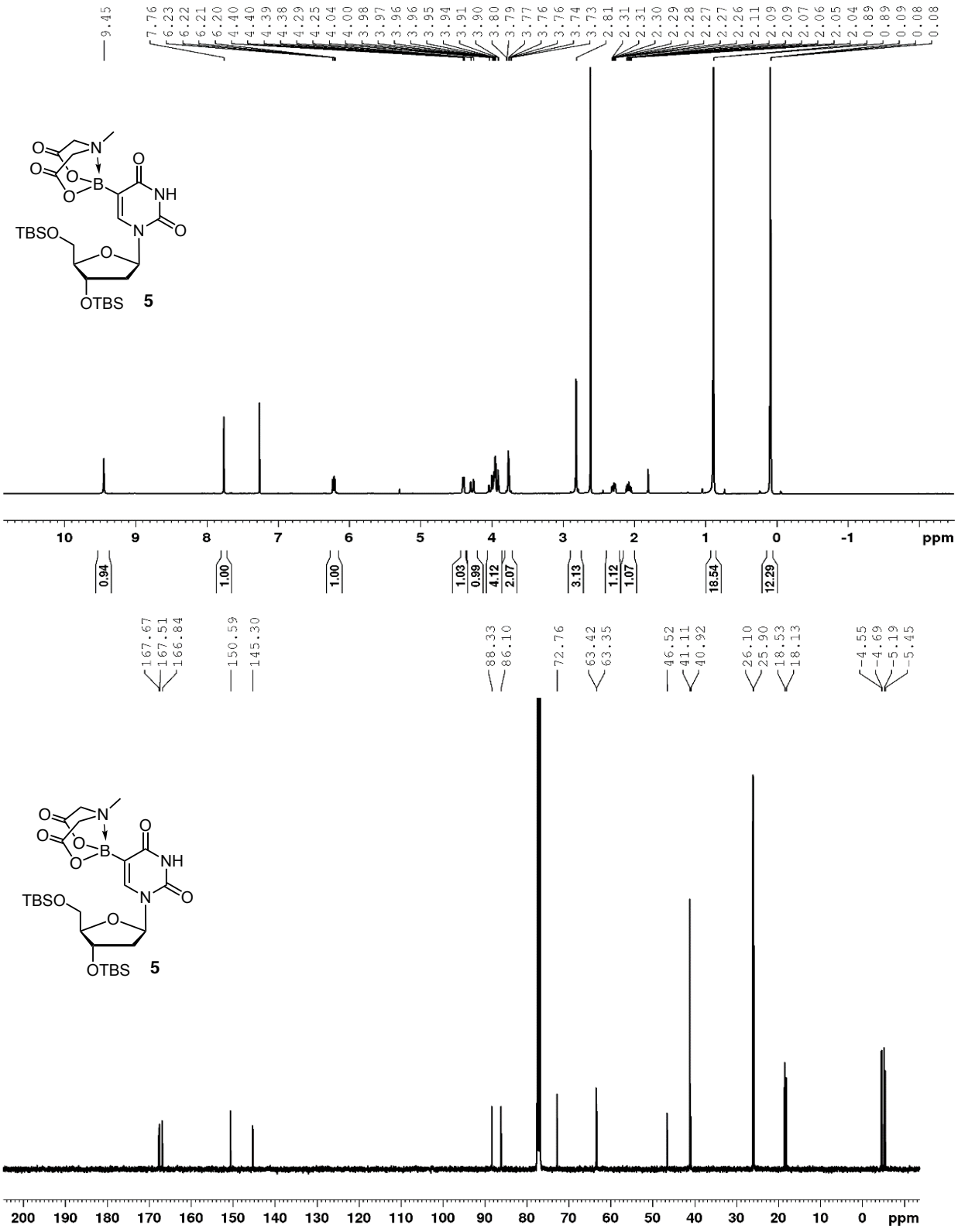
6. Supplementary Figures



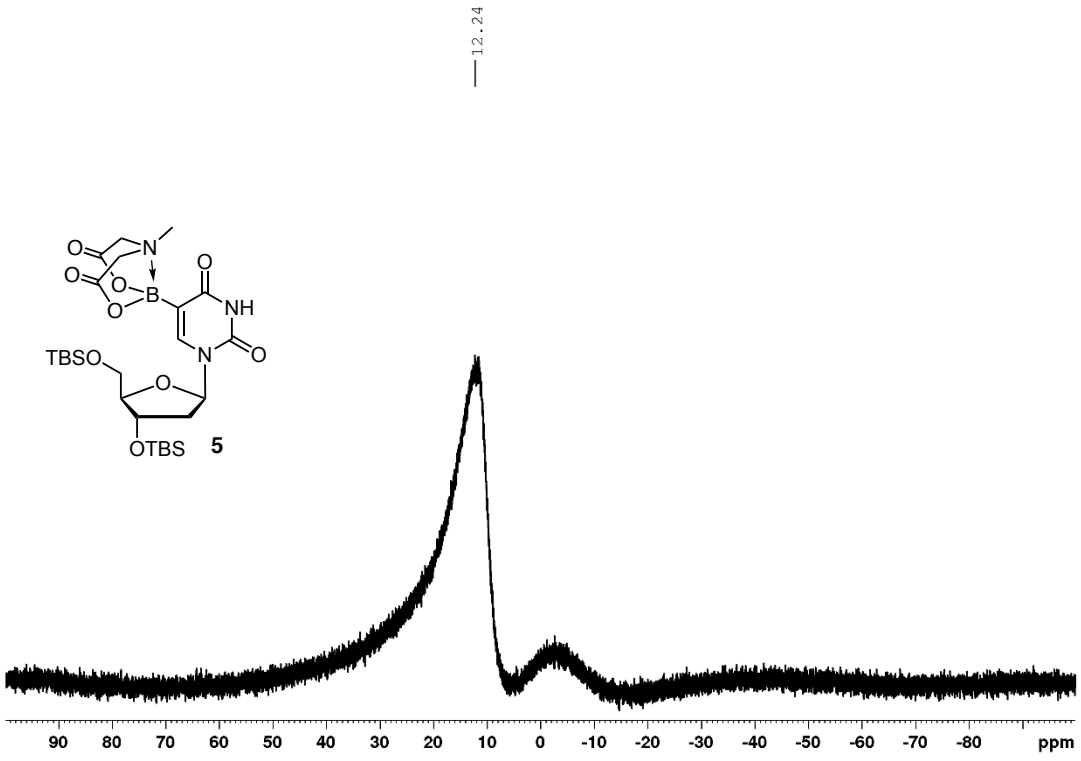
Supplementary Figure S1. ^{1}H and ^{13}C NMR spectra of compound **3** in $CDCl_3$.



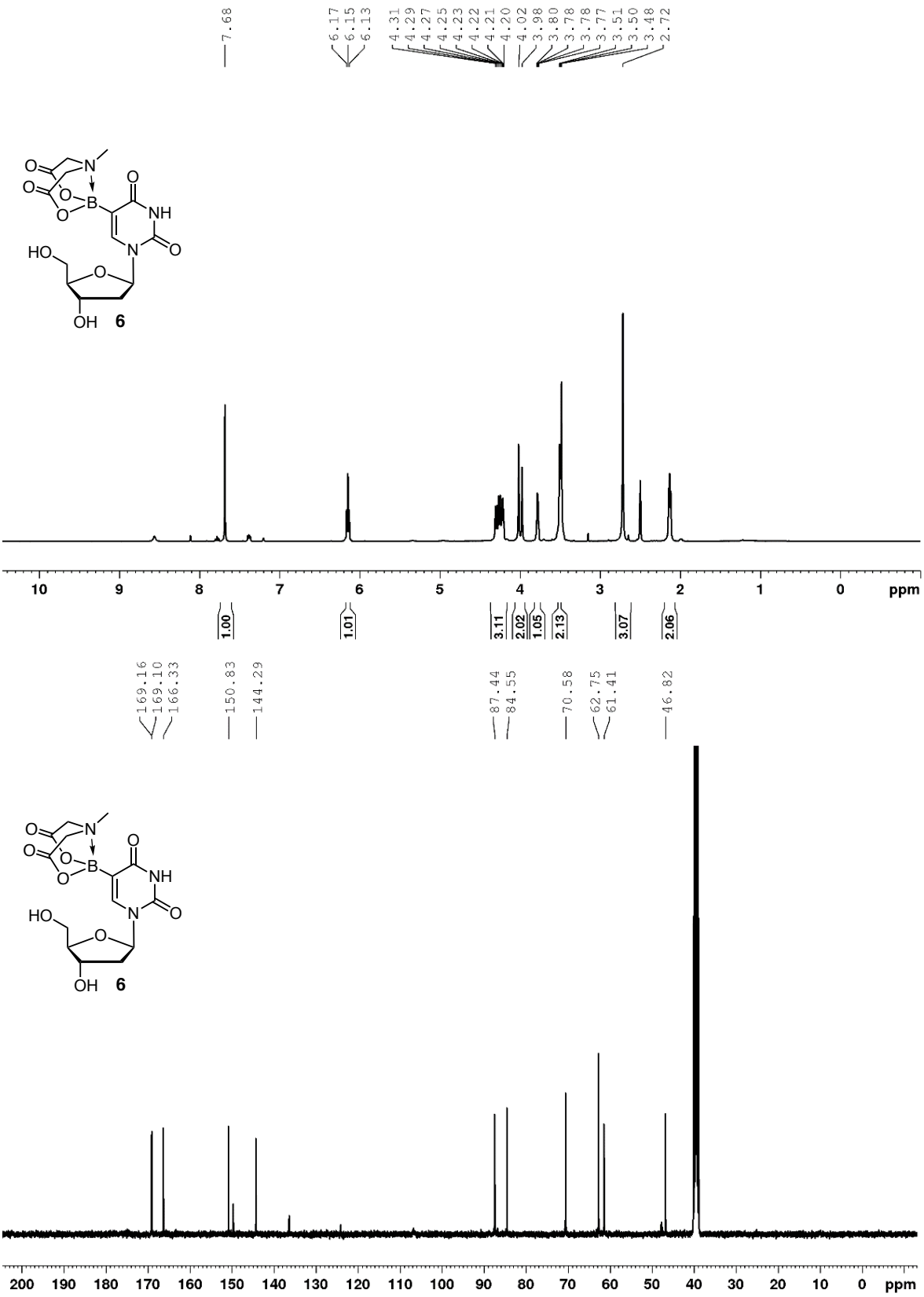
Supplementary Figure S2. ^{1}H and ^{13}C NMR spectra of compound **4** in $CDCl_3$.



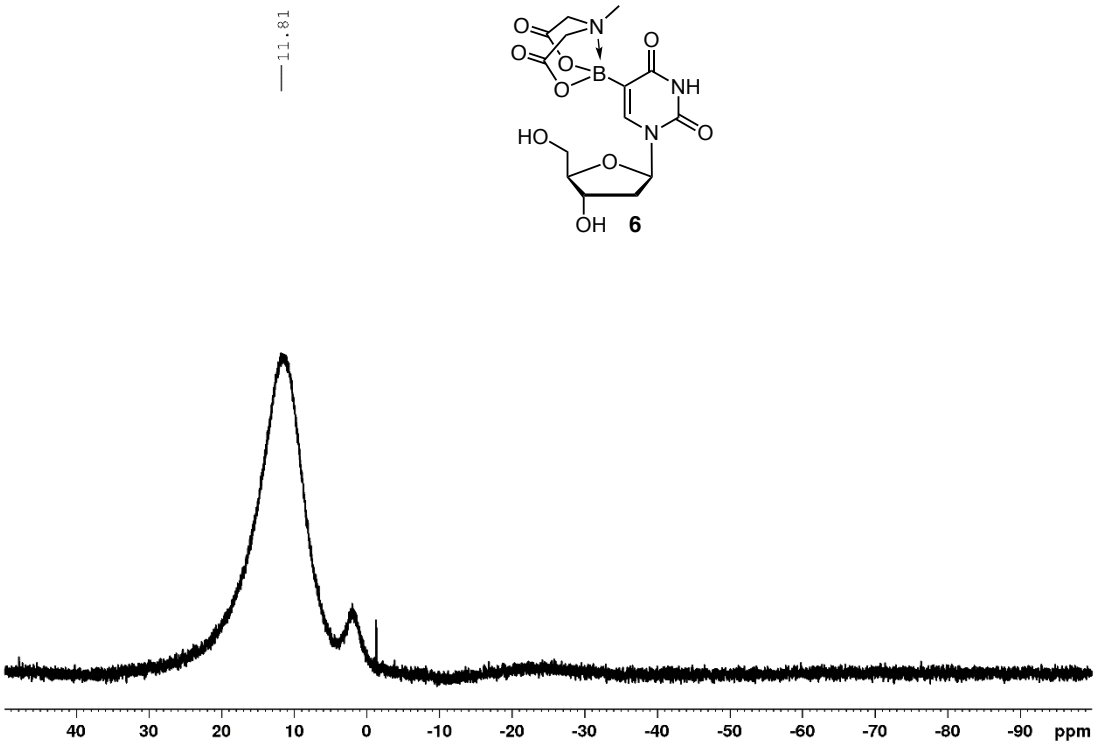
Supplementary Figure S3. ^{1}H and ^{13}C NMR spectra of compound **5** in $CDCl_3$.



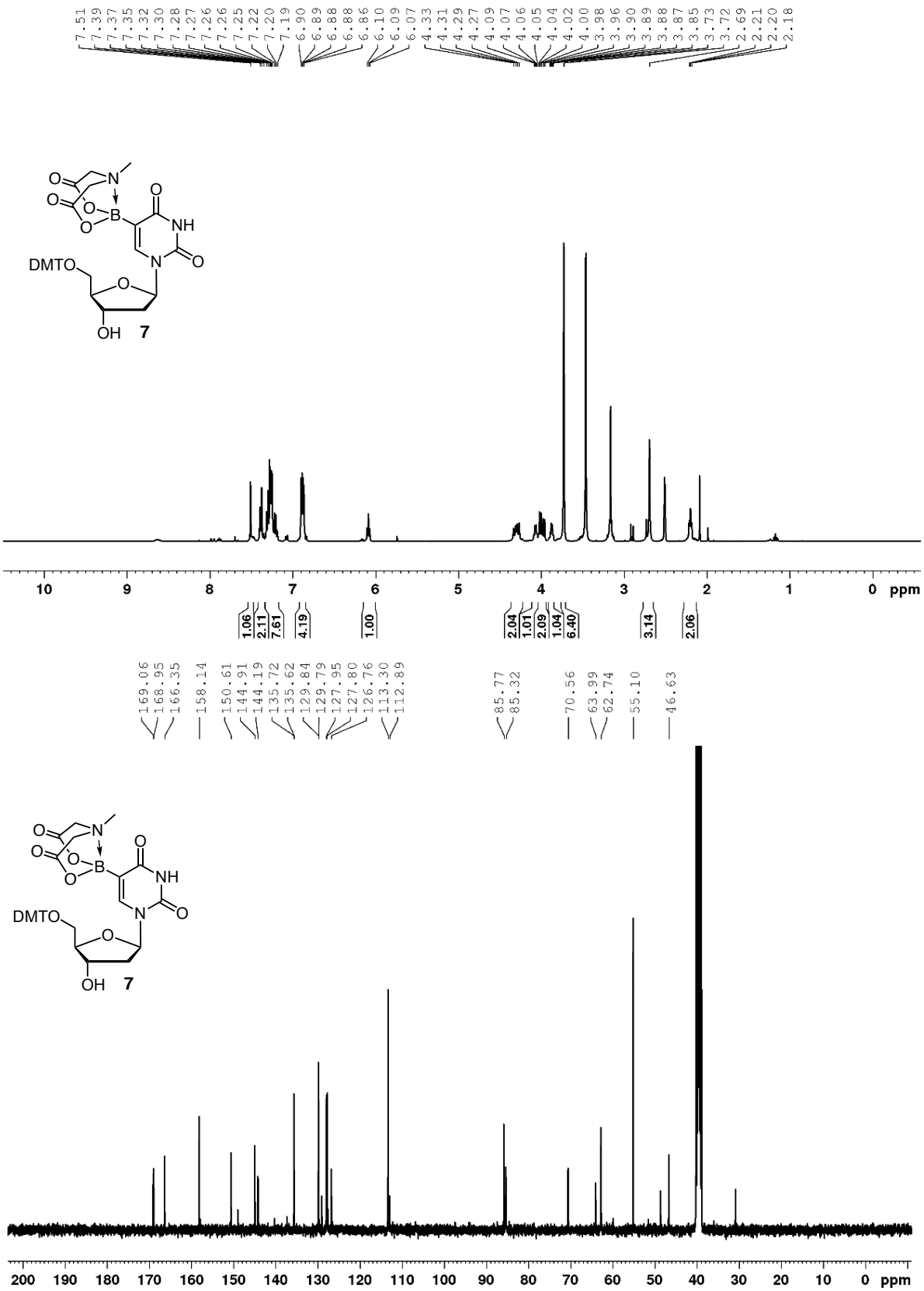
Supplementary Figure S3 Continued. ^{11}B spectrum of compound **5** in $CDCl_3$.



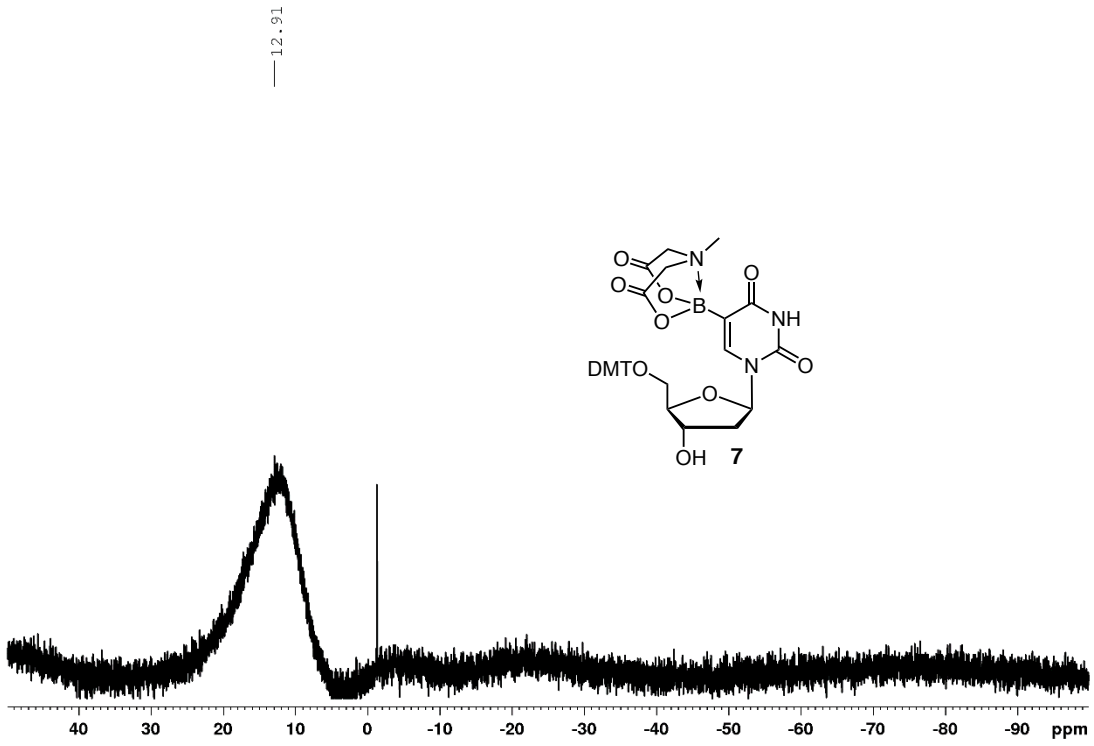
Supplementary Figure S4. ^{1}H and ^{13}C NMR spectra of compound **6** in DMSO-d_6/D_2O.



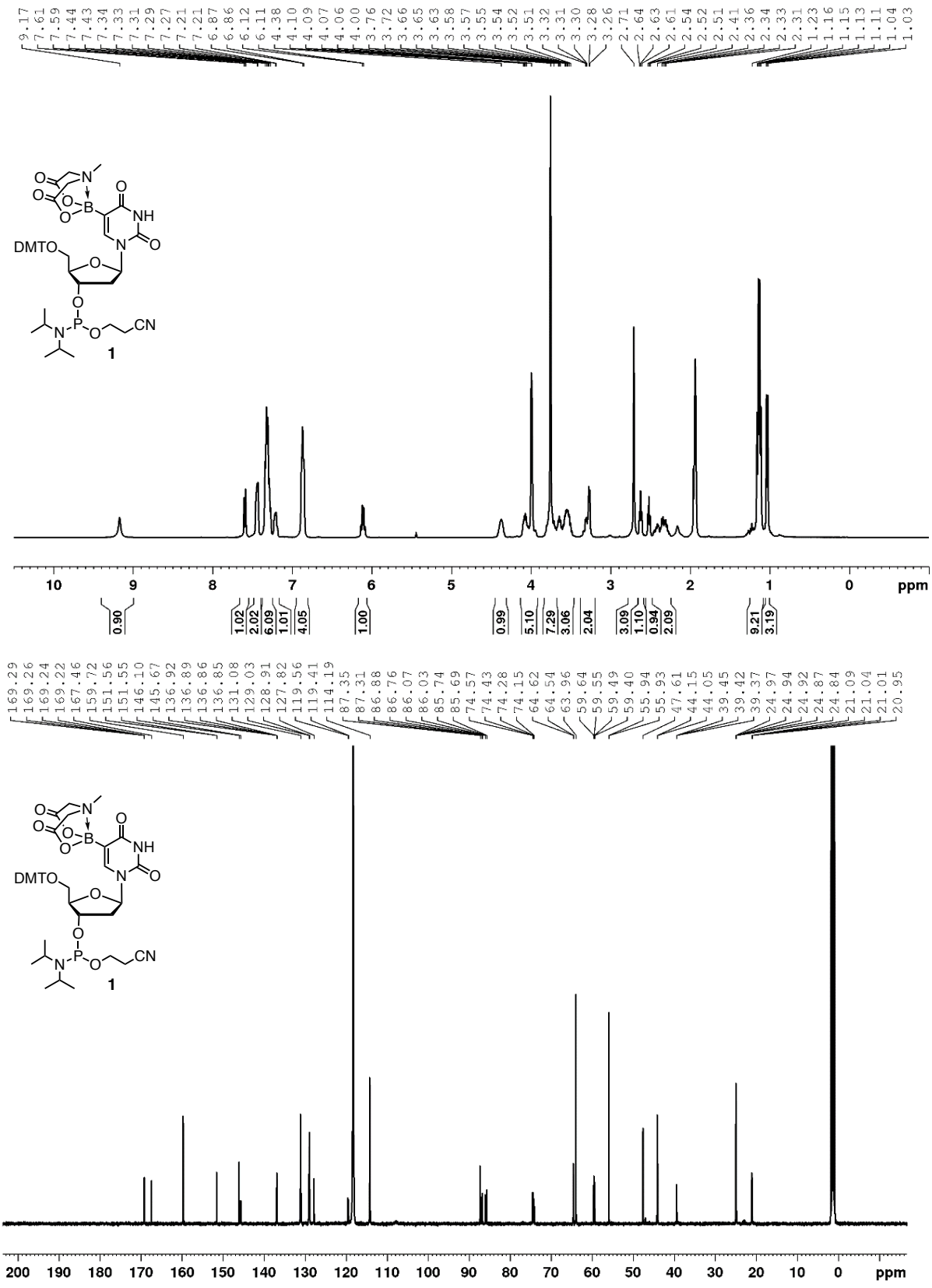
Supplementary Figure S4 continued. ^{11}B spectrum of compound **6** in DMSO-d_6/D_2O.



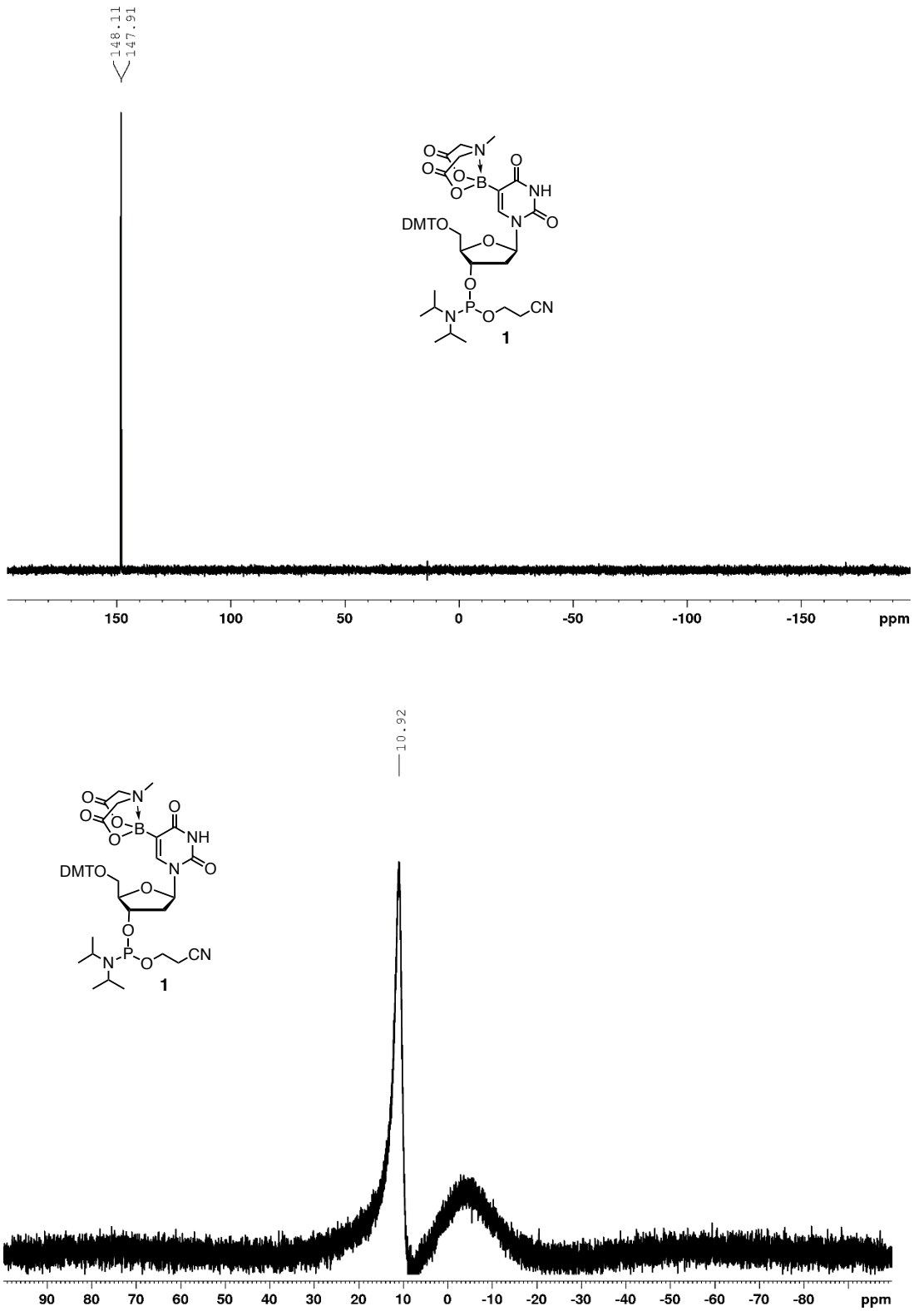
Supplementary Figure S5. ^{1}H and ^{13}C NMR spectra of compound **7** in DMSO-d_6/D_2O.



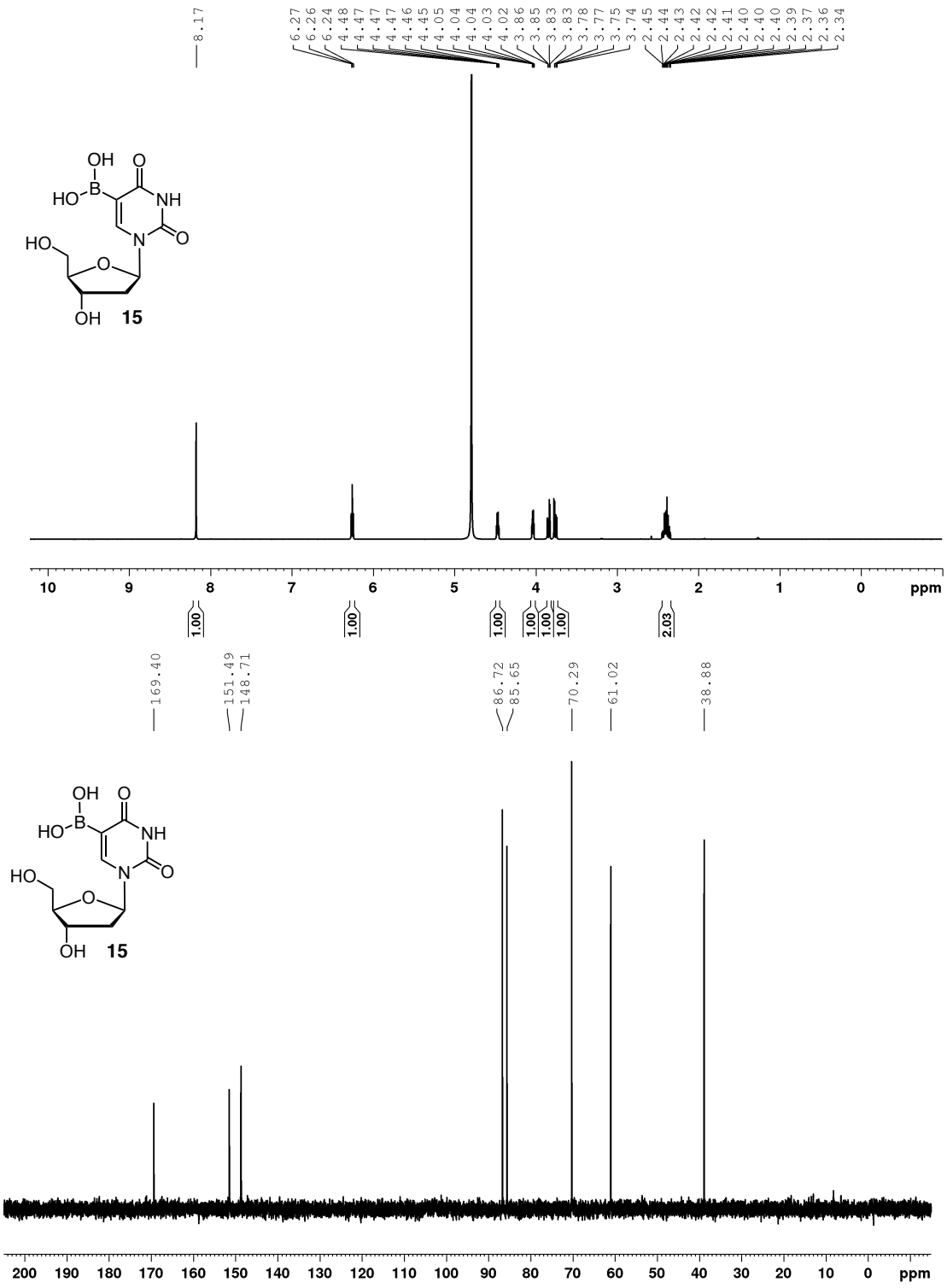
Supplementary Figure S5 continued. ^{11}B NMR spectrum of compound **7** in DMSO-d_6/D_2O.



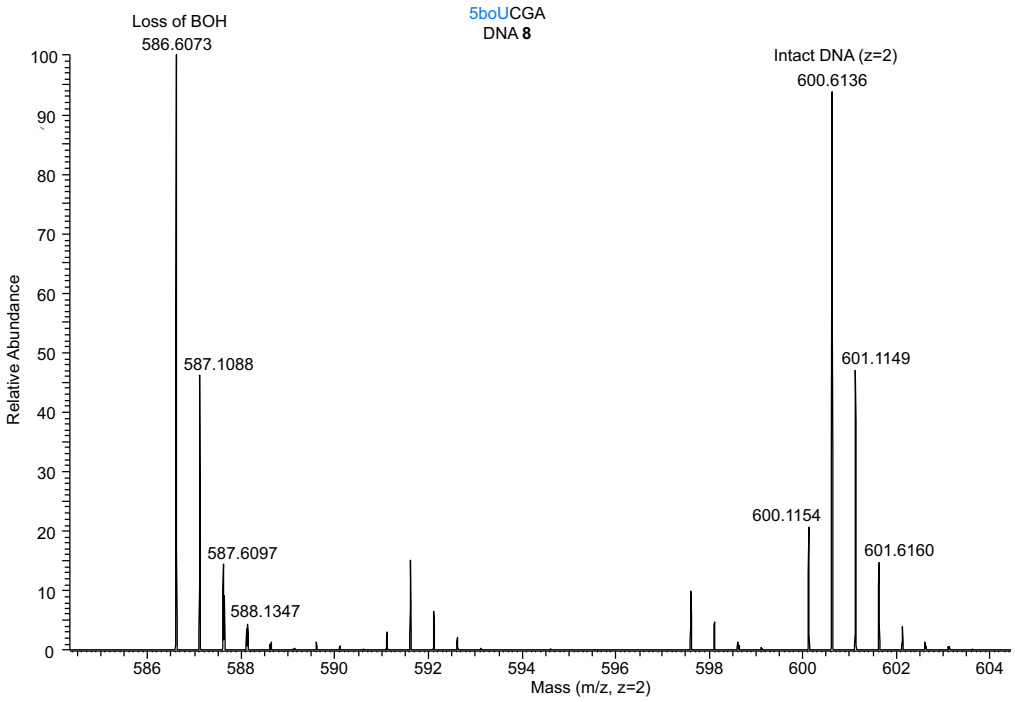
Supplementary Figure S6. ^{1}H and ^{13}C NMR spectra of compound **1** in CD_3CN.



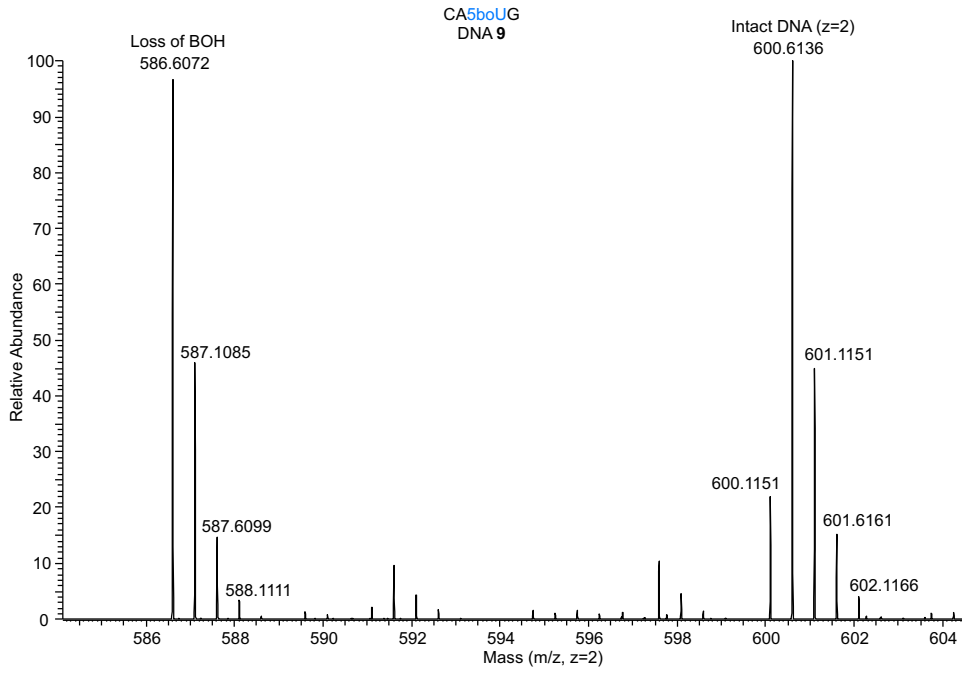
Supplementary Figure S6 continued. ^{31}P and ^{11}B NMR spectra of compound **1** in CD_3CN.



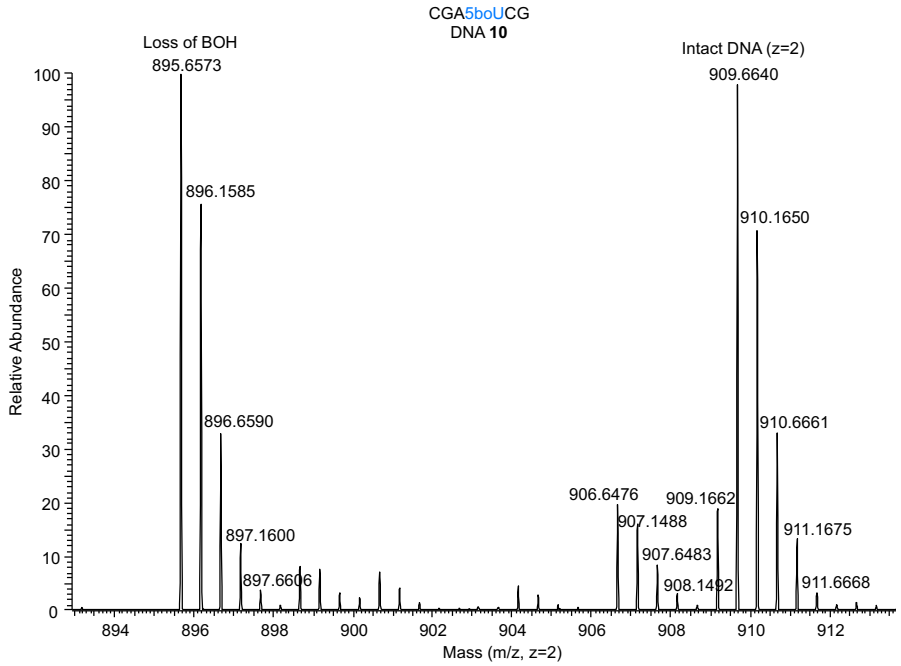
Supplementary Figure S7. ^{1}H and ^{13}C NMR spectra of compound **15** in D_2O.



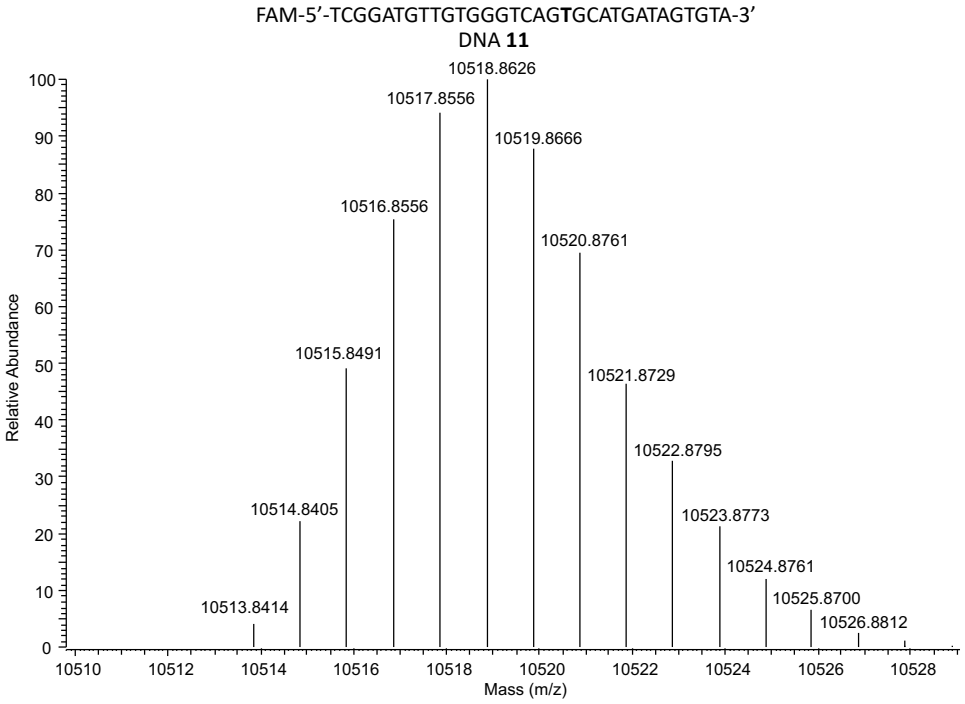
Supplementary Figure S8. ESI HRMS spectrum showing isotopic distribution of DNA **8**.



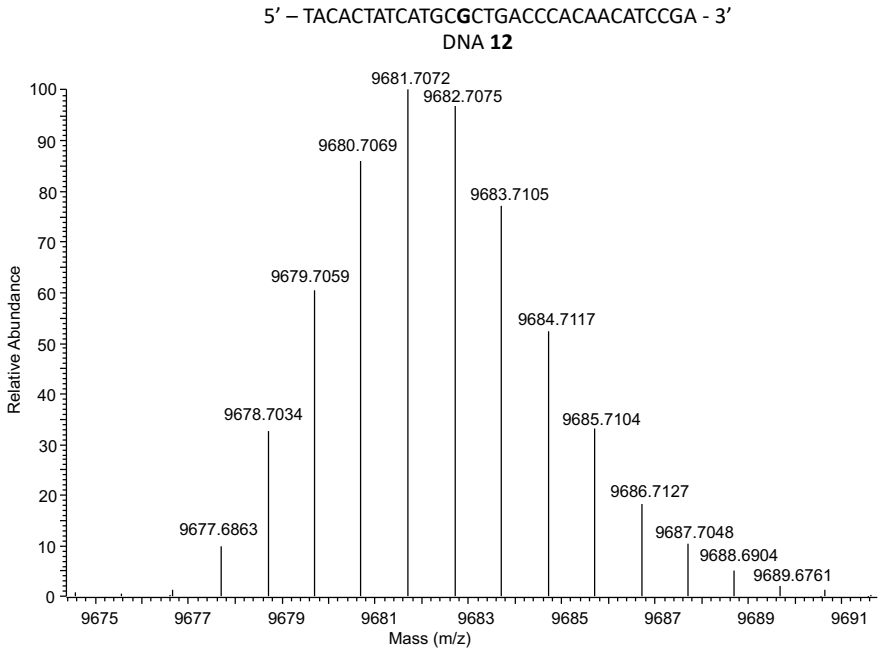
Supplementary Figure S9. ESI HRMS spectrum showing isotopic distribution of DNA **9**.



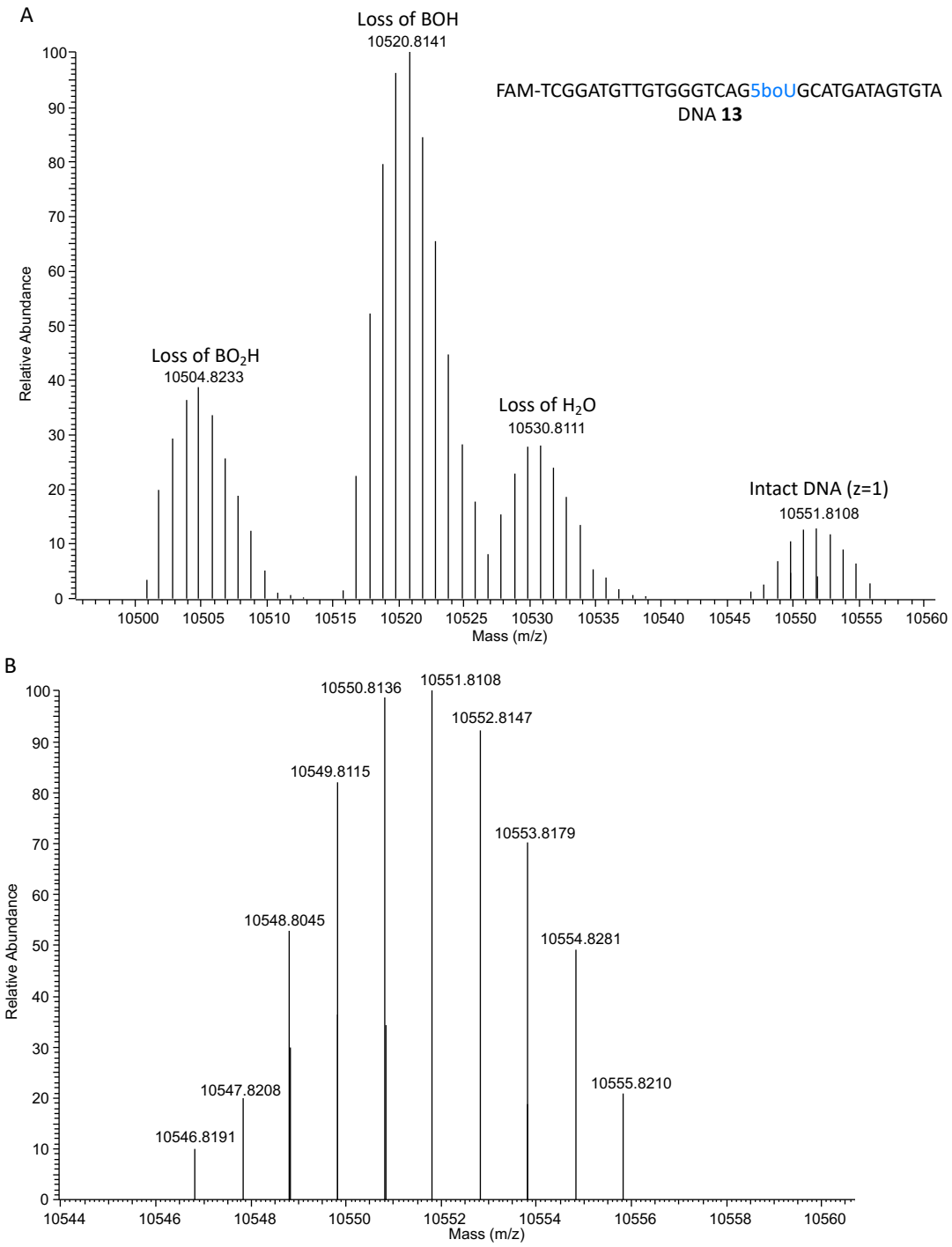
Supplementary Figure S10. ESI HRMS spectrum showing isotopic distribution of DNA **10**.



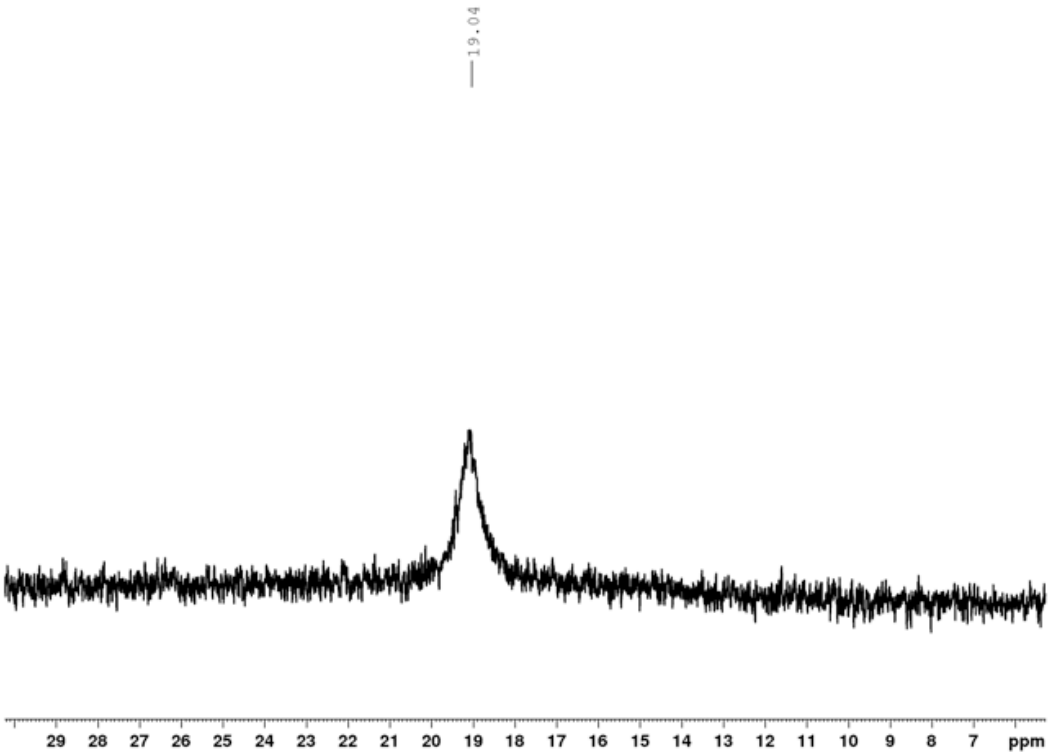
Supplementary Figure S11. ESI HRMS spectrum showing isotopic distribution of DNA **11** (FAM-5'-TCGGATGTTGTGGGTCAG**T**GCATGATAGTGTA-3'). Thymine in bold face generates a T:G mismatch with DNA **12**.



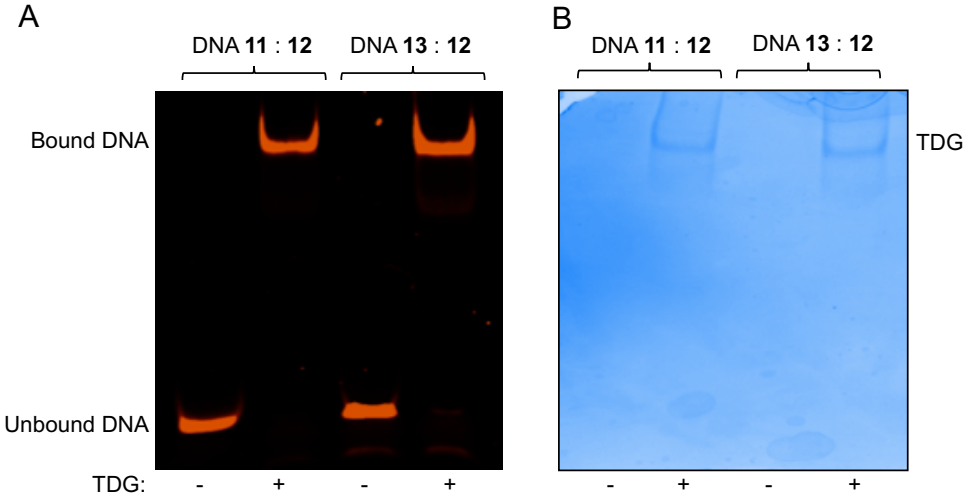
Supplementary Figure S12. ESI HRMS spectrum showing isotopic distribution of DNA **12** (5’–TACACTATCATGC**G**CTGACCCACAACATCCGA-3’). Guanine in bold face generates a T:G mismatch with DNA **11**.



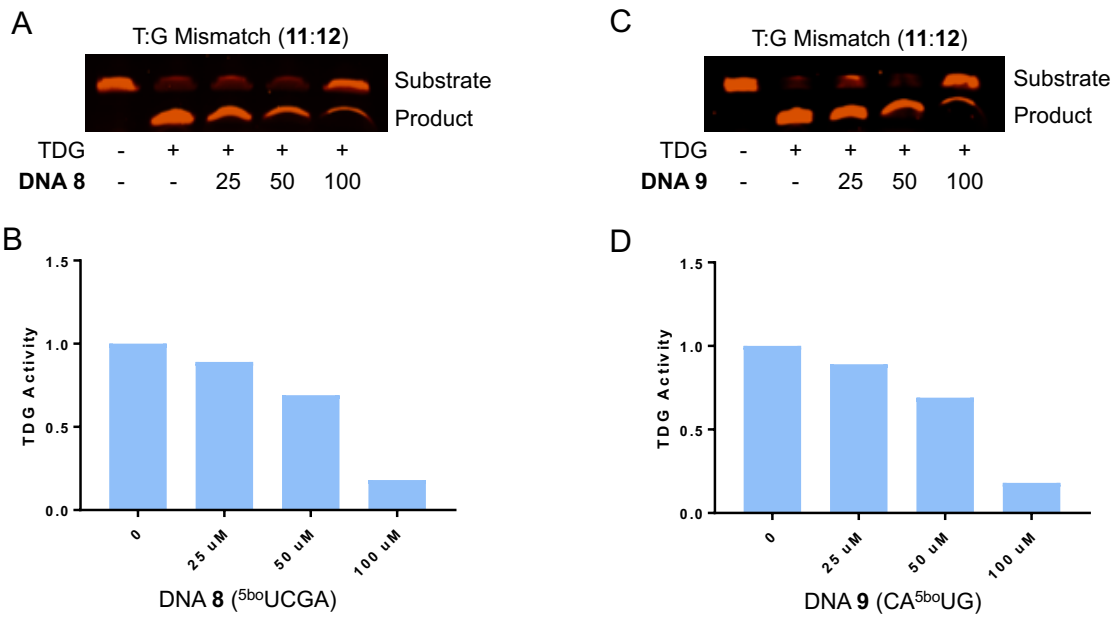
Supplementary Figure S13. (A) ESI HRMS spectrum showing isotopic distribution of DNA **13** (FAM-TCGGATGTTGTGGGTCAG5boUGCATGATAGTGTA). This DNA generates a 5boU:G mismatch with DNA **12**. (B) Expanded view of isotopic distribution of the intact ion peak.



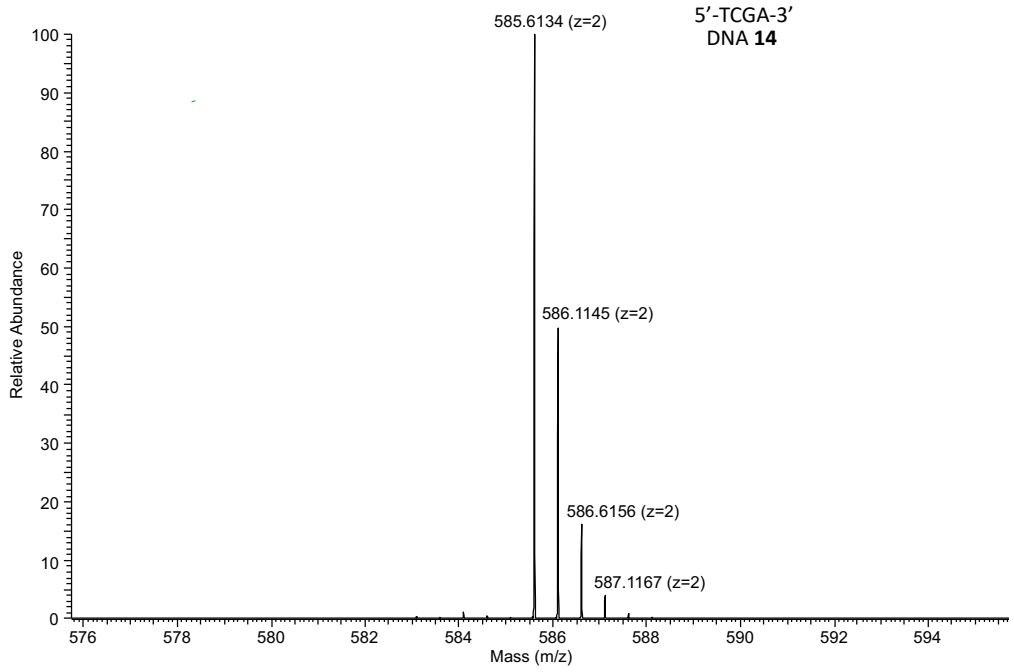
Supplementary Figure S14. ^{11}B NMR spectrum of DNA **13** in D_2O.



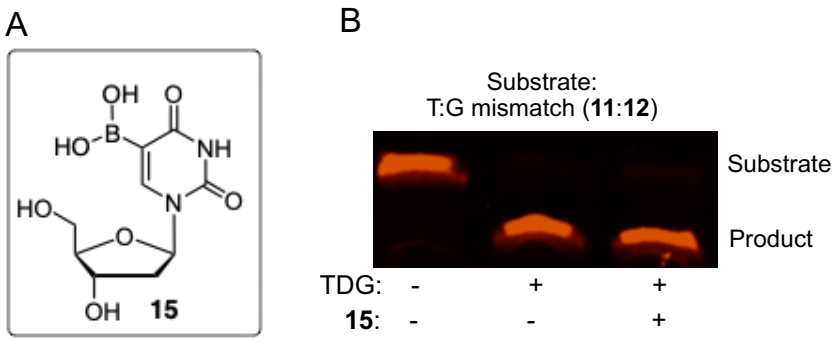
Supplementary Figure S15. (A) Electrophoretic mobility shift assay (EMSA) showing binding of FAM-labeled DNAs **11** and **13** each duplexed with DNA **12** to TDG. DNA **13** contains 5boU unit. In the case, the protein and DNA mixtures were not subjected to alkaline base cleavage prior to gel separation. Unbound and bound DNAs were separated using native-PAGE TBE GEL (5%). Gel was imaged on Bio-Rad ChemiDoc™ XRS+ platform. (B) Coomassie staining of the same gel showing loading of TDG.



Supplementary Figure S16. Dose-dependent inhibition of DNAs **8** and **9** of TDG activity. (A) In-gel fluorescence showing inhibition of TDG by **8**. (B) Bar diagram representation of (A). (C) In-gel fluorescence showing inhibition of TDG by **9**. (D) Bar diagram representation of (C). Quantification of the product formation by TDG was performed using Image-Lab (BioRad) for bar diagrams which were generated using Prism 8.0 (GraphPad).



Supplementary Figure S17. ESI HRMS spectrum showing isotopic distribution of DNA **14** (5'-TCGA-3'). The DNA is self-complementary and palindromic.



Supplementary Figure S18. (A) Chemical structure of 5-dihydroxyboryldeoxyuridine (5boU) **15** synthesized from **7** (Scheme S1, Figure S7). (B) 100 μM of **15** does not inhibit TDG activity on mismatched duplexed DNA of **11** and **12**.

7. References

(1) Schinazi, R. F.; Prusoff, W. H., Synthesis of 5-(dihydroxyboryl)-2'-deoxyuridine and related boron-containing pyrimidines. *J. Org. Chem.* **1985,** *50*, 841-847.

(2) Gillis, E. P.; Burke, M. D., Multistep synthesis of complex boronic acids from simple MIDA boronates. *J. Am. Chem. Soc.* **2008,** *130*, 14084-5.

(3) Pidugu, L. S.; Flowers, J. W.; Coey, C. T.; Pozharski, E.; Greenberg, M. M.; Drohat, A. C., Structural Basis for Excision of 5-Formylcytosine by Thymine DNA Glycosylase. *Biochem.* **2016,** *55*, 6205-6208.

(4) Zhang, L.; Lu, X.; Lu, J.; Liang, H.; Dai, Q.; Xu, G. L.; Luo, C.; Jiang, H.; He, C., Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA. *Nat. Chem. Biol.* **2012,** *8*, 328-30.

(5) Maiti, A.; Drohat, A. C., Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J. Biol. Chem.* **2011,** *286*, 35334-8.