

## Electronic Supplementary Information for

### Synthesis and Properties of 2'-O,4'-C-Spirocyclopropylene Bridged Nucleic Acid (scpBNA), an Analogue of 2',4'-BNA/LNA Bearing a Cyclopropane Ring

Takao Yamaguchi, Masahiko Horiba, and Satoshi Obika\*

#### Contents

<b>1. General</b>	<b>S2</b>
<b>2. Synthesis of scpBNA monomer and phosphoramidite</b>	<b>S2–S11</b>
<b>Figure S1 to Figure S30.</b> $^1\text{H}$ , $^{13}\text{C}$ , and $^{31}\text{P}$ -NMR spectra of synthesized compounds.	<b>S12–S41</b>
<b>3. Synthesis, purification and characterization of oligonucleotides</b>	<b>S42</b>
<b>Table S1.</b> Yields and MALDI-TOF MS data for the oligonucleotides	<b>S42</b>
<b>Figure S31.</b> MALDI-TOF MS spectra and HPLC charts of all new oligonucleotides.	<b>S43–S46</b>
<b>4. UV melting experiments</b>	<b>S47</b>
<b>Figure S32.</b> UV melting curves for the duplexes formed between representative oligonucleotides and ssRNA.	<b>S47</b>
<b>Figure S33.</b> UV melting curves for the duplexes formed between representative oligonucleotides and ssDNA.	<b>S47</b>
<b>5. Nuclease resistance study</b>	<b>S48</b>
<b>Figure S34.</b> Nuclease resistance of 5'-(TTTTTTTTX)-3' against <i>Crotalus admanteus</i> venom phosphodiesterase (CAVP).	<b>S48–S49</b>
<b>6. Figure S35.</b> Energy-minimized structures of 2',4'-BNA/LNA and scpBNA.	<b>S50</b>
<b>7. Mismatch discrimination and thermodynamic data of scpBNA</b>	<b>S51</b>
<b>Table S2.</b> $T_m$ values ( $^{\circ}\text{C}$ ) of duplexes formed between ONs and ssRNA with or without one base mismatch	<b>S51</b>
<b>Table S3.</b> $T_m$ values ( $^{\circ}\text{C}$ ) of duplexes formed between ONs and ssDNA with or without one base mismatch	<b>S51</b>
<b>Table S4.</b> Thermodynamic data of duplexes formed between ON2 and ssRNA or ssDNA	<b>S51</b>
<b>Figure S36.</b> Van't Hoff plots of $T_m$ values of duplexes formed between ON2 and ssRNA or ssDNA.	<b>S52</b>
<b>8. Supplementary references</b>	<b>S52</b>

## 1. General

Dry dichloromethane, *N,N*-dimethylformamide, tetrahydrofuran, acetonitrile and pyridine were used as purchased. <sup>1</sup>H NMR (400 MHz), <sup>13</sup>C NMR (100 MHz) and <sup>31</sup>P NMR (162 MHz) spectra were recorded on a JEOL JNM-ECS-400 spectrometer. <sup>1</sup>H NMR (300 MHz) and <sup>13</sup>C NMR (75 MHz) spectra were recorded on a JEOL JNM-AL-300 spectrometer. Chemical shift values are expressed in  $\delta$  values (ppm) relative to internal tetramethylsilane (0.00 ppm), residual CHCl<sub>3</sub> (7.26 ppm) or CHD<sub>2</sub>OD (3.31 ppm) for <sup>1</sup>H NMR, and internal tetramethylsilane (0.00 ppm), chloroform-*d*<sub>1</sub> (77.16 ppm) or methanol-*d*<sub>4</sub> (49.00 ppm) for <sup>13</sup>C NMR, and 85% H<sub>3</sub>PO<sub>4</sub> (0.00 ppm) as external standard for <sup>31</sup>P NMR. IR spectra were recorded on a JASCO FT/IR-4200 spectrometer. Optical rotations were recorded on a JASCO DIP-370 instrument. MALDI-TOF mass spectra of all new compounds were measured on SpiralTOF JMS-S3000. MALDI-TOF mass spectra of oligonucleotides were measured on a Bruker Daltonics Autoflex II TOF/TOF mass spectrometer. For column chromatography, Fuji Silysia PSQ-100B or FL-100D silica gel was used. For flash column chromatography, Fuji Silysia PSQ-60B or FL-60D silica gel was used. For high performance liquid chromatography (HPLC), SHIMADZU LC-10AT<sub>vp</sub>, SPD-10A<sub>vp</sub> and CTO-10<sub>vp</sub> were used.

## 2. Synthesis of scpBNA monomer and phosphoramidite.

### 3,5-Di-*O*-benzyl-4-*C*-formyl-1,2-*O*-isopropylidene- $\alpha$ -D-ribose (2)

The compound **2** was prepared *via* a different procedure from the reported one (Morita, K.; Takagi, M.; Hasegawa, C.; Kaneko, M.; Tsutsumi, S.; Sone, J.; Ishikawa, T.; Imanishi, T.; Koizumi, M. *Bioorg. Med. Chem.* **2003**, *11*, 2211–2226.). To the solution of **1** (7.38 g, 18.5 mmol) in dry dichloromethane (100 mL) was added Dess-Martin periodinane (9.41 g, 22.2 mmol) at 0 °C, and the reaction mixture was stirred at room temperature for 40 min under N<sub>2</sub> atmosphere. After completion of the reaction, saturated aq. Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> and saturated aq. NaHCO<sub>3</sub> were added, and the resulting mixture was further stirred for 10 min. The organic layer was then removed under reduced pressure, and the residual aqueous solution was extracted with Et<sub>2</sub>O. The combined organic layer was washed with water and brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated to afford **2** (7.61 g, quant.) as a colorless

oil.

Compound **2**:  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  1.34 (s, 3H), 1.60 (s, 3H), 3.61, 3.67 (AB,  $J = 11.0$  Hz, 2H), 4.36 (d,  $J = 4.4$  Hz, 1H), 4.46, 4.52 (AB,  $J = 12.2$  Hz, 2H), 4.59, 4.71 (AB,  $J = 11.9$  Hz, 2H), 4.60 (dd,  $J = 3.7, 4.4$  Hz, 1H), 5.84 (d,  $J = 3.7$  Hz, 1H), 7.21–7.37 (m, 10H), 9.91 (s, 1H);  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  26.1, 26.5, 69.1, 72.8, 73.8, 78.3, 79.6, 89.7, 104.8, 114.1, 127.7, 127.8, 128.0, 128.1, 128.4, 128.5, 137.0, 137.5, 200.0; IR (KBr): 2985, 2973, 2866, 1731, 1496, 1213, 1165, 1103, 1020, 739, 699  $\text{cm}^{-1}$ ;  $[\alpha]_{\text{D}}^{29} +27.1$  (c 1.03, MeOH); HRMS (MALDI) Calcd. for  $\text{C}_{23}\text{H}_{26}\text{O}_6\text{Na}$   $[\text{M}+\text{Na}]^+$ : 421.1622, Found 421.1620.

### **3,5-Di-*O*-benzyl-4-*C*-carboxy-1,2-*O*-isopropylidene- $\alpha$ -D-ribose (3)**

To the solution of **2** (7.61 g, 19.1 mmol) in acetonitrile (100 mL) was added aq. sodium dihydrogen orthophosphate (0.2 M, 20 mL, 3.82 mmol) and aq. hydrogen peroxide (30 wt.%, 2.3 mL, 21.0 mmol). After aq. sodium chlorite (0.75 M, 38 mL, 28.6 mmol) was added dropwise to the reaction mixture at 0 °C, the reaction mixture was stirred at room temperature for 1 h. After addition of  $\text{Na}_2\text{S}_2\text{O}_3$ , the resulting mixture was stirred at room temperature for 10 min. The organic layer was then removed under reduced pressure, and the residual aqueous solution was extracted with AcOEt. The combined organic layer was washed with water and brine, dried over  $\text{Na}_2\text{SO}_4$ , and concentrated to afford **3** (7.61 g, 96%) as a white solid.

Compound **3**:  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  1.34 (s, 3H), 1.58 (s, 3H), 3.72, 3.77 (AB,  $J = 10.8$  Hz, 2H), 4.30 (d,  $J = 4.5$  Hz, 1H), 4.49, 4.55 (AB,  $J = 11.9$  Hz, 2H), 4.65 (dd,  $J = 4.3, 4.5$  Hz, 1H), 4.69, 4.80 (AB,  $J = 11.9$  Hz, 2H), 5.83 (d,  $J = 4.3$  Hz, 1H), 7.21–7.40 (m, 10H);  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$  25.3, 26.4, 71.8, 73.4, 74.0, 78.1, 78.5, 104.9, 114.6, 127.8, 128.0, 128.2, 128.4, 128.6, 128.7, 136.5, 137.4, 170.0; IR (KBr): 3171, 2985, 2937, 2870, 1768, 1497, 1163, 1098, 1020, 740, 698  $\text{cm}^{-1}$ ;  $[\alpha]_{\text{D}}^{26} +42.3$  (c 1.01, MeOH); HRMS (MALDI) Calcd. for  $\text{C}_{23}\text{H}_{26}\text{O}_7\text{Na}$   $[\text{M}+\text{Na}]^+$ : 437.1571, Found 437.1570.

### **3,5-Di-*O*-benzyl-1,2-*O*-isopropylidene-4-*C*-methoxycarbonyl- $\alpha$ -D-ribose (4)**

To the solution of **3** (7.61 g, 18.4 mmol) in dry *N,N*-dimethylformamide (30 mL) was added  $\text{NaHCO}_3$  (15.4 g, 184 mmol) and iodomethane (2.86 mL, 45.9 mmol) at 0 °C under  $\text{N}_2$  atmosphere. After stirring at room temperature for 20 h, saturated aq.  $\text{Na}_2\text{S}_2\text{O}_3$  was added and the resulting

mixture was extracted with Et<sub>2</sub>O. The combined organic layer was washed with water and brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated to afford **4** (7.35 g, 93%) as a white solid.

Compound **4**: <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 1.38 (s, 3H), 1.64 (s, 3H), 3.67, 3.82 (AB, *J* = 10.3 Hz, 2H), 3.75 (s, 3H), 4.25 (d, *J* = 5.0 Hz, 1H), 4.49, 4.54 (AB, *J* = 11.9 Hz, 2H), 4.59, 4.77 (AB, *J* = 12.2 Hz, 2H), 4.67 (dd, *J* = 4.2, 5.0 Hz, 1H), 5.89 (d, *J* = 4.2 Hz, 1H), 7.24–7.27 (m, 10H); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 26.3, 27.3, 73.0, 73.7, 73.9, 79.4, 80.5, 89.7, 127.7, 127.7, 127.8, 127.9, 128.4, 128.6, 137.7, 169.4; IR (KBr): 2985, 2949, 2869, 1763, 1733, 1497, 1160, 1106, 1028, 738, 698 cm<sup>-1</sup>; [α]<sub>D</sub><sup>27</sup> +31.5 (c 1.00, MeOH); HRMS (MALDI) Calcd. for C<sub>24</sub>H<sub>28</sub>O<sub>7</sub>Na [M+Na]<sup>+</sup>: 451.1727, Found 451.1732.

### **3,5-Di-*O*-benzyl-4-*C*-(1-hydroxycyclopropyl)-1,2-*O*-isopropylidene- $\alpha$ -D-ribofuranose (**5**)**

To the solution of **4** (12.5 g, 29.0 mmol) in dry tetrahydrofuran (290 mL) was added tetraisopropyl orthotitanate (8.59 mL, 29.0 mmol) and 1M ethylmagnesium bromide in tetrahydrofuran (145 mL, 145 mmol) at 0 °C under N<sub>2</sub> atmosphere. After stirring at room temperature for 6 h, saturated aq. NH<sub>4</sub>Cl was added and the organic layer was then removed under reduced pressure. The residual aqueous solution was filtered through celite and extracted with AcOEt. The combined organic layer was washed with water and brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated. The crude product was purified by column chromatography (SiO<sub>2</sub>, *n*-hexane : AcOEt = 6 : 1) to afford **5** (6.80 g, 55%) as a yellow paste.

Compound **5**: <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 0.56–0.68 (m, 3H), 1.16–1.21 (m, 1H), 1.39 (s, 3H), 1.61 (s, 3H), 3.37 (s, 1H), 3.48, 3.93 (AB, *J* = 9.8 Hz, 2H), 4.33 (d, *J* = 5.7 Hz, 1H), 4.43, 5.00 (AB, *J* = 11.6 Hz, 2H), 4.45, 4.54 (AB, *J* = 12.0 Hz, 2H), 4.84 (dd, *J* = 4.5, 5.7 Hz, 1H), 5.88 (d, *J* = 4.5 Hz, 1H), 7.26–7.39 (m, 10H); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 8.64, 11.2, 27.1, 27.9, 56.3, 73.1, 73.8, 75.5, 80.3, 82.0, 89.1, 106.6, 114.6, 127.5, 127.6, 127.8, 128.0, 128.6, 128.8, 137.9, 138.2; IR (KBr): 2935, 2867, 1496, 1454, 1252, 1099, 1027, 741, 699 cm<sup>-1</sup>; [α]<sub>D</sub><sup>29</sup> +93.5 (c 1.02, MeOH); HRMS (MALDI) Calcd. for C<sub>25</sub>H<sub>30</sub>O<sub>6</sub>Na [M+Na]<sup>+</sup>: 449.1935, Found 449.1939.

### **3,5-Di-*O*-benzyl-4-*C*-[1-(*tert*-butyldimethylsilyloxy)cyclopropyl]-1,2-*O*-isopropylidene- $\alpha$ -D-ribofuranose (**6**)**

To the solution of **5** (2.55 g, 5.99 mmol) in dry dichloromethane (50 mL) was added 2,6-lutidine (2.09 mL, 18.0 mmol) and *tert*-butyldimethylsilyl trifluoromethanesulfonate (2.75 mL, 12.0 mmol) at 0 °C under N<sub>2</sub> atmosphere. After stirring at room temperature for 2 h, saturated aq. NaHCO<sub>3</sub> was added and the resulting mixture was extracted with AcOEt. The combined organic layer was washed with water and brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated. The crude product was purified by column chromatography (SiO<sub>2</sub>, *n*-hexane : AcOEt = 15 : 1 to 5 : 1) to afford **6** (2.92 g, 90%) as a yellow oil.

Compound **6**: <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ -0.06 (s, 3H), -0.02 (s, 3H), 0.57–0.77 (m, 3H), 0.75 (s, 9H), 1.20–1.25 (m, 1H), 1.34 (s, 3H), 1.43 (s, 3H), 3.46, 3.92 (AB, *J* = 9.5 Hz, 2H), 4.00 (d, *J* = 5.7 Hz, 1H), 4.42, 4.61 (AB, *J* = 12.0 Hz, 2H), 4.52, 4.86 (AB, *J* = 11.4 Hz, 2H), 4.95 (dd, *J* = 4.5, 5.7 Hz, 1H), 5.87 (d, *J* = 4.5 Hz, 1H), 7.19–7.43 (m, 10H); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ -3.4, -3.2, 7.7, 10.2, 17.8, 25.7, 27.1, 28.5, 57.5, 73.4, 73.8, 76.2, 80.0, 83.3, 90.3, 106.1, 114.5, 126.8, 126.9, 127.6, 127.8, 127.8, 128.5, 138.0, 139.2; IR (KBr): 2929, 2858, 1497, 1455, 1279, 1254, 1106, 1040, 733, 696 cm<sup>-1</sup>; [α]<sub>D</sub><sup>29</sup> +53.6 (c 1.01, MeOH); HRMS (MALDI) Calcd. for C<sub>31</sub>H<sub>44</sub>O<sub>6</sub>NaSi [M+Na]<sup>+</sup>: 563.2799, Found 563.2809.

### **1-[3,5-Di-*O*-benzyl-4-*C*-[1-(*tert*-butyldimethylsilyloxy)cyclopropyl]-2-*O*-methanesulfonyl-β-D-ribofuranosyl]thymine (10)**

To the solution of **6** (8.04 g, 14.9 mmol) in acetic acid (17.0 mL, 0.30 mol) was added acetic anhydride (28.2 mL, 0.30 mol) and trifluoroacetic acid (3.20 mL, 44.7 mmol) at 0 °C. After stirring at room temperature for 5 h, saturated aq. NaHCO<sub>3</sub> was added and the resulting mixture was extracted with AcOEt. The combined organic layer was washed with water and brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated. The crude product **7** (9.43 g) was used immediately for the next reaction without further purification.

To the solution of **7** (9.43 g) in dry acetonitrile (140 mL) was added thymine (5.63 g, 44.6 mmol), *N,O*-bis-trimethylsilylacetoamide (18.2 mL, 74.3 mmol) and trimethylsilyl trifluoromethanesulfonate (4.03 mL, 22.3 mmol) at room temperature under N<sub>2</sub> atmosphere. After refluxing for 2 h, saturated aq. NaHCO<sub>3</sub> was added and the resulting mixture was extracted with AcOEt. The combined organic layer was washed with water and brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated to afford **8** (8.26 g). The crude product **8** was used for the next reaction without further

purification. For an analysis of compound **8**, a small amount of the crude product was purified by silica gel column chromatography (SiO<sub>2</sub>, *n*-hexane : AcOEt = 3 : 1).

To the solution of **8** (8.26 g) in tetrahydrofuran (150 mL) was added aq. methylamine (40 wt.%, 30.4 mL, 0.73 mol) at 0 °C, and the reaction mixture was stirred at room temperature for 4 h. After completion of reaction, the resulting mixture was concentrated and extracted with AcOEt. The combined organic layer was washed with water and brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated. The crude product **9** (7.50 g) was used for the next reaction without further purification. For an analysis of compound **9**, a small amount of the crude product was purified by silica gel column chromatography (SiO<sub>2</sub>, *n*-hexane : AcOEt = 2 : 1).

To the solution of **9** (7.50 g) in dry pyridine (120 mL) was added methanesulfonyl chloride (1.43 mL, 18.5 mmol) at 0 °C, and the reaction mixture was stirred at room temperature for 4 h under N<sub>2</sub> atmosphere. After addition of water, the resulting mixture was extracted with AcOEt. The combined organic layer was washed with water and brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated. The crude product was purified by column chromatography (SiO<sub>2</sub>, *n*-hexane : AcOEt = 3 : 2) to afford **10** (7.39 g, 72%, 4 steps) as a white solid.

Compound **7**: <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ -0.07 (s, 3/2H), -0.02 (s, 3/2H), 0.00 (s, 3/2H), 0.04 (s, 3/2H), 0.59–0.83 (m, 4H), 0.75 (s, 9/2H), 0.78 (s, 9/2H), 1.92 (s, 3/2H), 1.98 (s, 3/2H), 2.01 (s, 3/2H), 2.09 (s, 3/2H), 3.44 (d, *J* = 9.9 Hz, 1/2H), 3.57 (d, *J* = 9.9 Hz, 1/2H), 3.91 (d, *J* = 9.9 Hz, 1/2H), 3.94 (d, *J* = 10.2 Hz, 1/2H), 4.30 (d, *J* = 5.1 Hz, 1/2H), 4.40–4.59 (m, 7/2H), 4.67 (d, *J* = 2.4 Hz, 1/2H), 4.71 (d, *J* = 2.4 Hz, 1/2H), 4.88 (d, *J* = 3.9 Hz, 1/2H), 4.92 (d, *J* = 3.6 Hz, 1/2H), 5.44 (t, *J* = 5.1 Hz, 1/2H), 5.57 (d, *J* = 5.7 Hz, 1/2H), 6.20 (d, *J* = 5.1 Hz, 1/2H), 6.39 (d, *J* = 5.1 Hz, 1/2H), 7.26–7.39 (m, 10H); HRMS (MALDI) Calcd. for C<sub>32</sub>H<sub>44</sub>O<sub>8</sub>NaSi [M+Na]<sup>+</sup>: 607.2698, Found 607.2701.

Compound **8**: <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ -0.01 (s, 3H), 0.03 (s, 3H), 0.65–0.74 (m, 2H), 0.78 (s, 9H), 0.95–1.03 (m, 2H), 1.56 (d, *J* = 1.2 Hz, 3H), 1.97 (s, 3H), 3.63, 4.03 (AB, *J* = 9.8 Hz, 2H), 4.46 (d, *J* = 5.0 Hz, 1H), 4.50, 4.95 (AB, *J* = 11.3 Hz, 2H), 4.62, 4.72 (AB, *J* = 11.6 Hz, 2H), 5.50 (dd, *J* = 5.0, 8.9 Hz, 1H), 6.23 (d, *J* = 8.9 Hz, 1H), 7.26–7.44 (m, 10H), 7.66 (d, *J* = 1.2 Hz, 1H), 7.86 (s, 1H); <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>) δ -3.3, -3.0, 7.3, 10.7, 12.3, 18.0, 20.8, 25.8, 58.1, 74.1, 74.4, 75.3, 75.3, 80.9, 84.9, 87.8, 111.6, 127.4, 127.6, 128.0, 128.3, 128.4, 129.0, 136.1, 137.1, 138.7, 150.8, 163.6, 170.8; IR (KBr): 3499, 2955, 2929, 1714, 1683, 1470, 1274, 1233, 1127, 1075,

1036, 733, 699  $\text{cm}^{-1}$ ;  $[\alpha]_{\text{D}}^{24}$   $-46.9$  (c 0.99, MeOH); HRMS (MALDI) Calcd. for  $\text{C}_{35}\text{H}_{46}\text{N}_2\text{O}_8\text{NaSi}$   $[\text{M}+\text{Na}]^+$ : 673.2916, Found 673.2917.

Compound **9**:  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  0.03 (s, 3H), 0.06 (s, 3H), 0.68–0.79 (m, 3H), 0.81 (s, 9H), 0.94–0.98 (m, 1H), 1.60 (d,  $J = 1.4$  Hz, 3H), 2.86 (d,  $J = 12.0$  Hz, 1H), 3.60, 4.02 (AB,  $J = 9.6$  Hz, 2H), 4.21 (d,  $J = 5.4$  Hz, 1H), 4.51–4.60 (m, 3H), 4.69 (B part of an AB system,  $J = 11.7$  Hz, 1H), 5.20 (B part of an AB system,  $J = 10.8$  Hz, 1H), 5.83 (d,  $J = 8.1$  Hz, 1H), 7.32–7.42 (m, 10H), 7.59 (d,  $J = 1.4$  Hz, 1H), 8.38 (s, 1H);  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$   $-3.2$ ,  $-2.9$ , 7.4, 10.8, 12.3, 18.1, 25.9, 58.2, 74.1, 74.5, 74.7, 75.8, 82.7, 86.8, 87.8, 111.4, 127.9, 128.0, 128.2, 128.5, 128.7, 129.0, 136.1, 137.1, 137.9, 151.1, 163.5; IR (KBr): 3422, 2955, 2929, 1699, 1470, 1277, 1254, 1129, 1087, 1036, 751, 698  $\text{cm}^{-1}$ ;  $[\alpha]_{\text{D}}^{26}$   $-45.1$  (c 1.00, MeOH); HRMS (MALDI) Calcd. for  $\text{C}_{33}\text{H}_{44}\text{N}_2\text{O}_7\text{NaSi}$   $[\text{M}+\text{Na}]^+$ : 631.2810, Found 631.2814.

Compound **10**:  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$   $-0.01$  (s, 3H), 0.04 (s, 3H), 0.58–1.02 (m, 4H), 0.78 (s, 9H), 1.56 (s, 3H), 2.89 (s, 3H), 3.63, 4.03 (AB,  $J = 9.8$  Hz, 2H), 4.37 (d,  $J = 4.8$  Hz, 1H), 4.63, 4.71 (AB,  $J = 11.6$  Hz, 2H), 4.78, 4.94 (AB,  $J = 11.0$  Hz, 2H), 5.58 (dd,  $J = 4.8, 8.7$  Hz, 1H), 6.23 (d,  $J = 8.7$  Hz, 1H), 7.26–7.41 (m, 10H), 7.60 (s, 1H), 8.13 (s, 1H);  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ )  $\delta$   $-3.4$ ,  $-3.1$ , 7.1, 10.6, 12.2, 17.9, 25.7, 38.2, 57.9, 73.9, 74.0, 75.1, 77.2, 81.0, 84.5, 87.6, 111.9, 127.4, 127.5, 127.9, 128.2, 128.4, 128.9, 135.3, 136.7, 138.3, 150.6, 163.2; IR (KBr): 3414, 2926, 1696, 1454, 1363, 1127, 1072, 1038, 748, 698  $\text{cm}^{-1}$ ;  $[\alpha]_{\text{D}}^{31}$   $-48.2$  (c 0.96, MeOH); HRMS (MALDI) Calcd. for  $\text{C}_{34}\text{H}_{46}\text{N}_2\text{O}_9\text{NaSiS}$   $[\text{M}+\text{Na}]^+$ : 709.2586, Found 709.2582.

### **1-[(1*R*,4*R*,6*R*,7*S*)-7-benzyloxy-4-benzyloxymethyl-2,5-dioxaspiro(bicyclo[2.2.1]heptane-3,1'-cyclopropan)-6-yl]thymine (13)**

To the solution of **10** (3.19 g, 4.64 mmol) in tetrahydrofuran/ ethanol (150 mL, 3 : 2) was added 4M aq. sodium hydroxide (60 mL, 0.23 mol) at 0 °C, and the reaction mixture was stirred at room temperature for 12 h. After addition of aq. HCl, the resulting mixture was concentrated and extracted with AcOEt. The combined organic layer was washed with water and brine, dried over  $\text{Na}_2\text{SO}_4$ , and concentrated. The crude product **11** (2.71 g) was used for the next reaction without further purification. For an analysis of compound **11**, a small amount of the crude product was purified by column chromatography ( $\text{SiO}_2$ , *n*-hexane : AcOEt = 1.7 : 1).

To the solution of **11** (2.71 g) in dry pyridine (50 mL) was added trifluoromethanesulfonic

anhydride (3.65 mL, 22.3 mmol) at 0 °C, and the reaction mixture was stirred at room temperature for 12 h under N<sub>2</sub> atmosphere. After addition of water, the resulting mixture was extracted with AcOEt. The combined organic layer was washed with water and brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated. The crude product **12** (4.12 g) was used immediately for the next reaction without further purification.

To the solution of **12** (4.12 g) in tetrahydrofuran (250 mL) was added 1M tetrabutylammonium fluoride in tetrahydrofuran (13.9 mL, 13.9 mmol) at 0 °C, and the reaction mixture was stirred at room temperature for 2 h. After completion of reaction, the reaction mixture was concentrated. The crude product was purified by column chromatography (SiO<sub>2</sub>, *n*-hexane : AcOEt = 3 : 2) to afford **13** (710 mg, 32%, 3 steps) as a white solid.

Compound **11**: <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ -0.06 (s, 3H), 0.00 (s, 3H), 0.57–0.91 (m, 4H), 0.74 (s, 9H), 1.80 (d, *J* = 0.9 Hz, 3H), 3.81, 4.15 (AB, *J* = 9.9 Hz, 2H), 4.15 (s, 1H), 4.22 (dd, *J* = 3.7, 11.9 Hz, 1H), 4.63, 4.76 (AB, *J* = 11.5 Hz, 2H), 4.67, 4.72 (AB, *J* = 11.9 Hz, 2H), 4.96 (d, *J* = 11.9 Hz, 1H), 6.03 (d, *J* = 3.7 Hz, 1H), 7.26–7.44 (m, 11H), 8.13 (s, 1H); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ -3.4, -3.2, 7.9, 10.9, 12.6, 17.8, 25.6, 58.1, 73.2, 74.3, 74.4, 74.5, 86.9, 87.2, 87.4, 108.8, 127.0, 127.5, 128.2, 128.4, 129.0, 129.0, 135.6, 137.4, 137.9, 149.9, 163.5; IR (KBr): 2954, 1703, 1669, 1472, 1286, 1254, 1097, 1042, 738, 696 cm<sup>-1</sup>; [α]<sub>D</sub><sup>30</sup> +36.3 (c 1.00, MeOH); HRMS (MALDI) Calcd. for C<sub>33</sub>H<sub>44</sub>N<sub>2</sub>O<sub>7</sub>NaSi [M+Na]<sup>+</sup>: 631.2810, Found 631.2813.

Compound **12**: <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 0.04 (s, 3H), 0.07 (d, *J* = 0.9 Hz, 3H), 0.66–0.88 (m, 13H), 1.25 (s, 3H), 3.50, 3.62 (AB, *J* = 7.7 Hz, 2H), 4.45, 4.56 (AB, *J* = 8.7 Hz, 2H), 4.60 (d, *J* = 2.7 Hz, 1H), 4.64, 4.75 (AB, *J* = 8.7 Hz, 2H), 5.58 (dd, *J* = 2.7, 4.2 Hz, 1H), 6.49 (d, *J* = 4.2 Hz, 1H), 7.29–7.41 (m, 10H), 7.46 (d, *J* = 0.6 Hz, 1H), 8.08 (s, 1H).

Compound **13**: <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 0.65–0.76 (m, 2H), 0.91–1.01 (m, 2H), 1.62 (d, *J* = 0.9 Hz, 3 H), 3.50, 3.63 (AB, *J* = 11.0 Hz, 2H), 4.04 (s, 1H), 4.51–4.59 (m, 4H), 4.70 (B part of an AB system, *J* = 12.0 Hz, 1H), 5.73 (s, 1H), 7.26–7.39 (m, 10H), 7.51 (d, *J* = 0.9 Hz, 1H), 8.33 (s, 1H); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 5.3, 9.9, 12.4, 64.1, 68.4, 72.2, 74.0, 77.1, 87.1, 87.6, 110.3, 127.7, 127.8, 127.8, 128.2, 128.6, 128.7, 128.7, 135.1, 137.3, 137.5, 150.0, 164.1; IR (KBr): 3512, 3031, 1693, 1455, 1269, 1108, 1054, 761, 738, 699 cm<sup>-1</sup>; [α]<sub>D</sub><sup>22</sup> +55.3 (c 1.00, MeOH); HRMS (MALDI) Calcd. for C<sub>27</sub>H<sub>28</sub>N<sub>2</sub>O<sub>6</sub>Na [M+Na]<sup>+</sup>: 499.1840, Found 499.1829.



**2,2'-Anhydro-1-[3,5-di-*O*-benzyl-4-*C*-(1-hydroxycyclopropyl)- $\beta$ -D-arabinopentofuranosyl]thymine (14)**

To the solution of **10** (459 mg, 0.67 mmol) in tetrahydrofuran (25 mL) was added 1M tetrabutylammoniumfluoride in tetrahydrofuran (0.67 mL, 0.67 mmol) at 0 °C, and the reaction mixture was stirred at room temperature for 5 h. After addition of water, the resulting mixture was extracted with AcOEt. The combined organic layer was washed with water and brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated. The crude product was purified by column chromatography (SiO<sub>2</sub>, CHCl<sub>3</sub> : CH<sub>3</sub>OH = 50 : 1 to 20 : 1) to afford **14** (290 mg, 91%) as a white solid.

Compound **14**: <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>)  $\delta$  0.65–0.75 (m, 4H), 1.92 (d, *J* = 0.9 Hz, 3H), 3.24 (s, 1H), 3.31, 3.60 (AB, *J* = 10.5 Hz, 2H), 4.24, 4.33 (AB, *J* = 12.2 Hz, 2H), 4.59–4.63 (m, 2H), 4.85 (B part of an AB system, *J* = 11.7 Hz, 1H), 5.36 (dd, *J* = 2.1, 6.2 Hz, 1H), 6.15 (d, *J* = 6.2 Hz, 1H), 7.08–7.12 (m, 3H), 7.26–7.40 (m, 8H); <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>)  $\delta$  10.0, 10.8, 14.2, 56.5, 70.9, 73.3, 73.8, 85.9, 87.0, 89.7, 90.8, 118.9, 128.0, 128.1, 128.1, 128.5, 128.7, 129.0, 130.2, 136.1, 136.9, 159.6, 172.6; IR (KBr): 3330, 3069, 2923, 1665, 1633, 1556, 1487, 1128, 1087, 736, 700 cm<sup>-1</sup>; [ $\alpha$ ]<sub>D</sub><sup>26</sup> –2.24 (c 1.00, MeOH); HRMS (MALDI) Calcd. for C<sub>27</sub>H<sub>29</sub>N<sub>2</sub>O<sub>6</sub> [M+H]<sup>+</sup>: 477.2020, Found 477.2024.

**1-[(1*R*,4*R*,6*R*,7*S*)-7-benzyloxy-4-benzyloxymethyl-2,5-dioxaspiro(bicyclo[2.2.1]heptane-3,1'-cyclopropan)-6-yl]thymine (13)**

To the solution of **14** (1.62 g, 3.40 mmol) in *N,N*-dimethylformamide (35 mL) was added potassium carbonate (1.41 g, 10.2 mmol) at 0 °C, and the reaction mixture was stirred at 90 °C for 20 h. After addition of water, the resulting mixture was extracted with Et<sub>2</sub>O. The combined organic layer was washed with water and brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated. The crude product was purified by column chromatography (SiO<sub>2</sub>, *n*-hexane : AcOEt = 1 : 1) to afford **13** (1.23 g, 77%) as a white solid.

**1-[(1*R*,4*R*,6*R*,7*S*)-2,5-Dioxaspiro(bicyclo[2.2.1]heptane-3,1'-cyclopropan)-7-hydroxy-4-hydroxymethyl-6-yl]thymine (15)**

To the solution of **13** (2.58 g, 5.46 mmol) in AcOEt (50 mL) was added palladium hydroxide 20% on carbon (1.24 g). The reaction flask was degassed a few times with H<sub>2</sub> and the reaction mixture

was stirred at room temperature for 1 h under H<sub>2</sub> atmosphere. After completion of reaction, the reaction mixture was filtered, washed by AcOEt and concentrated. The crude product was purified by column chromatography (SiO<sub>2</sub>, *n*-hexane : AcOEt = 1 : 5) to afford **15** (1.53 g, 95%) and **16** (70 mg, 5%) as white solids.

Compound **15**: <sup>1</sup>H NMR (300 MHz, CD<sub>3</sub>OD) δ 0.70–0.92 (m, 4H), 1.89 (d, *J* = 1.1 Hz, 3H), 3.56, 3.74 (AB, *J* = 12.8 Hz, 2H), 4.19 (s, 1H), 4.32 (s, 1H), 5.63 (s, 1H), 7.79 (d, *J* = 1.1 Hz, 1H); <sup>13</sup>C NMR (75 MHz, CD<sub>3</sub>OD) δ 5.1, 9.9, 12.6, 56.8, 68.7, 71.9, 81.0, 88.1, 89.9, 110.7, 137.0, 151.9, 166.5; IR (KBr): 3479, 3076, 1695, 1472, 1269, 1105, 1041 cm<sup>-1</sup>; [α]<sub>D</sub><sup>20</sup> +25.2 (c 1.01, MeOH); HRMS (MALDI) Calcd. for C<sub>13</sub>H<sub>16</sub>N<sub>2</sub>O<sub>6</sub>Na [M+Na]<sup>+</sup>: 319.0901, Found 319.0882.

Compound **16**: <sup>1</sup>H NMR (300 MHz, CD<sub>3</sub>OD) δ 0.96–1.01 (m, 6H), 1.89 (d, *J* = 1.2 Hz, 3H), 2.29–2.41 (m, 1H), 3.65, 3.73 (AB, *J* = 11.6 Hz, 2H), 4.15 (d, *J* = 5.3 Hz, 1H), 4.53 (dd, *J* = 5.3, 8.3 Hz, 1H), 5.93 (d, *J* = 8.3 Hz, 1H), 7.95 (d, *J* = 1.2 Hz, 1H); <sup>13</sup>C NMR (75 MHz, CD<sub>3</sub>OD) δ 12.5, 17.4, 18.9, 32.3, 63.6, 74.1, 75.6, 88.5, 92.0, 111.8, 138.9, 153.2, 166.4; IR (KBr): 3375, 2968, 1692, 1474, 1279, 1114, 1087 cm<sup>-1</sup>; [α]<sub>D</sub><sup>30</sup> -25.7 (c 1.04, MeOH); HRMS (MALDI) Calcd. for C<sub>13</sub>H<sub>20</sub>N<sub>2</sub>O<sub>6</sub>Na [M+Na]<sup>+</sup>: 323.1214, Found 323.1212.

**1-[(1*R*,4*R*,6*R*,7*S*)-2,5-Dioxaspiro(bicyclo[2.2.1]heptane-3,1'-cyclopropan)-4-[[bis(4-methoxyphenyl)(phenyl)methoxy]methyl]-7-hydroxy-6-yl]thymine (**17**)**

To the solution of **15** (873 mg, 2.95 mmol) in dry pyridine (60 mL) was added 4,4'-dimethoxytrityl chloride (1.50 g, 4.42 mmol) at 0 °C, and the reaction mixture was stirred at room temperature for 9 h under N<sub>2</sub> atmosphere. After addition of water, the resulting mixture was extracted with CH<sub>2</sub>Cl<sub>2</sub>. The organic layer was washed with water and brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated. The crude product was purified by column chromatography (SiO<sub>2</sub>, 0.5% triethylamine in *n*-hexane : AcOEt = 1 : 1 to 1 : 5) to afford **17** (1.72 g, 97%) as a white solid.

Compound **17**: <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 0.48–0.54 (m, 1H), 0.74–0.96 (m, 3H), 1.73 (s, 3H), 2.18 (d, *J* = 9.6 Hz, 1H), 3.16, 3.33 (AB, *J* = 11.0 Hz, 2H), 3.79 (s, 3H), 3.80 (s, 3H), 4.30 (d, *J* = 9.6 Hz, 1H), 4.43 (s, 1H), 5.76 (s, 1H), 6.85 (d, *J* = 8.4 Hz, 4H), 7.22–7.35 (m, 7H), 7.45 (d, *J* = 7.8 Hz, 2H), 7.65 (s, 1H), 8.39 (s, 1H); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 5.3, 9.6, 12.7, 55.3, 58.0, 68.0, 72.6, 79.7, 86.8, 87.0, 88.0, 110.6, 113.4, 127.2, 128.1, 128.2, 130.1, 130.2, 134.8, 135.2, 135.4, 144.4, 150.1, 158.8, 164.3; IR (KBr): 3430, 2933, 1696, 1509, 1254, 1177, 1053, 829, 757 cm<sup>-1</sup>;

$[\alpha]_D^{21}$  -16.2 (c 1.00, MeOH); HRMS (MALDI) Calcd. for  $C_{34}H_{34}N_2O_8Na$   $[M+Na]^+$ : 621.2207, Found 621.2208.

**1-[(1*R*,4*R*,6*R*,7*S*)-7-[(2-Cyanoethoxy)(diisopropylamino)phosphinoxy]-2,5-dioxaspiro(bicyclo[2.2.1]heptane-3,1'-cyclopropan)-4,4'-dimethoxytrityloxymethyl -6-yl]thymine (18)**

To the solution of **17** (192 mg, 0.32 mmol) in dry acetonitrile (4 mL) was added *N,N*-diisopropylethylamine (0.17 mL, 0.96 mmol) and 2-cyanoethyl-*N,N*-diisopropylphosphoramidochloridite (0.11 mL, 0.48 mmol) at 0 °C under  $N_2$  atmosphere. After stirring at room temperature for 5 h, the reaction mixture was concentrated and the crude product was purified by column chromatography ( $SiO_2$ , 0.5% triethylamine in *n*-hexane : AcOEt = 2 : 1) to afford **18** (222 mg, 87%) as a white solid.

Compound **18**:  $^1H$  NMR (300 MHz,  $CDCl_3$ )  $\delta$  0.40–0.44 (m, 1H), 0.71–0.87 (m, 3H), 0.98 (d,  $J$  = 6.9 Hz, 3H), 1.07 (d,  $J$  = 6.9 Hz, 3H), 1.12 (d,  $J$  = 6.9 Hz, 3H), 1.15 (d,  $J$  = 6.6 Hz, 3H), 1.67 (s, 3/2H), 1.68 (s, 3/2H), 2.37–2.41 (m, 1H), 2.53–2.63 (m, 1H), 3.15–3.30 (m, 2H), 3.49–3.57 (m, 3H), 3.63–3.73 (m, 1H), 3.79 (s, 3H), 3.80 (s, 3H), 4.36 (d,  $J$  = 6.9 Hz, 1/2H), 4.41 (d,  $J$  = 8.7 Hz, 1/2H), 4.60 (s, 1/2H), 4.63 (s, 1/2H), 5.77 (s, 1H), 6.82–6.87 (m, 4H), 7.24–7.35 (m, 7H), 7.43–7.45 (m, 2H), 7.70 (s, 1/2H), 7.73 (s, 1/2H), 8.21 (s, 1H);  $^{31}P$  NMR (162 MHz,  $CDCl_3$ )  $\delta$  148.6; LRMS (FAB)  $m/z$  = 799 ( $M+H$ ) $^+$ ; HRMS (FAB) Calcd. for  $C_{43}H_{52}O_9N_4P$  799.3472, Found 799.3475.

Figure S1. Compound 2 (<sup>1</sup>H NMR, CDCl<sub>3</sub>, 400MHz)

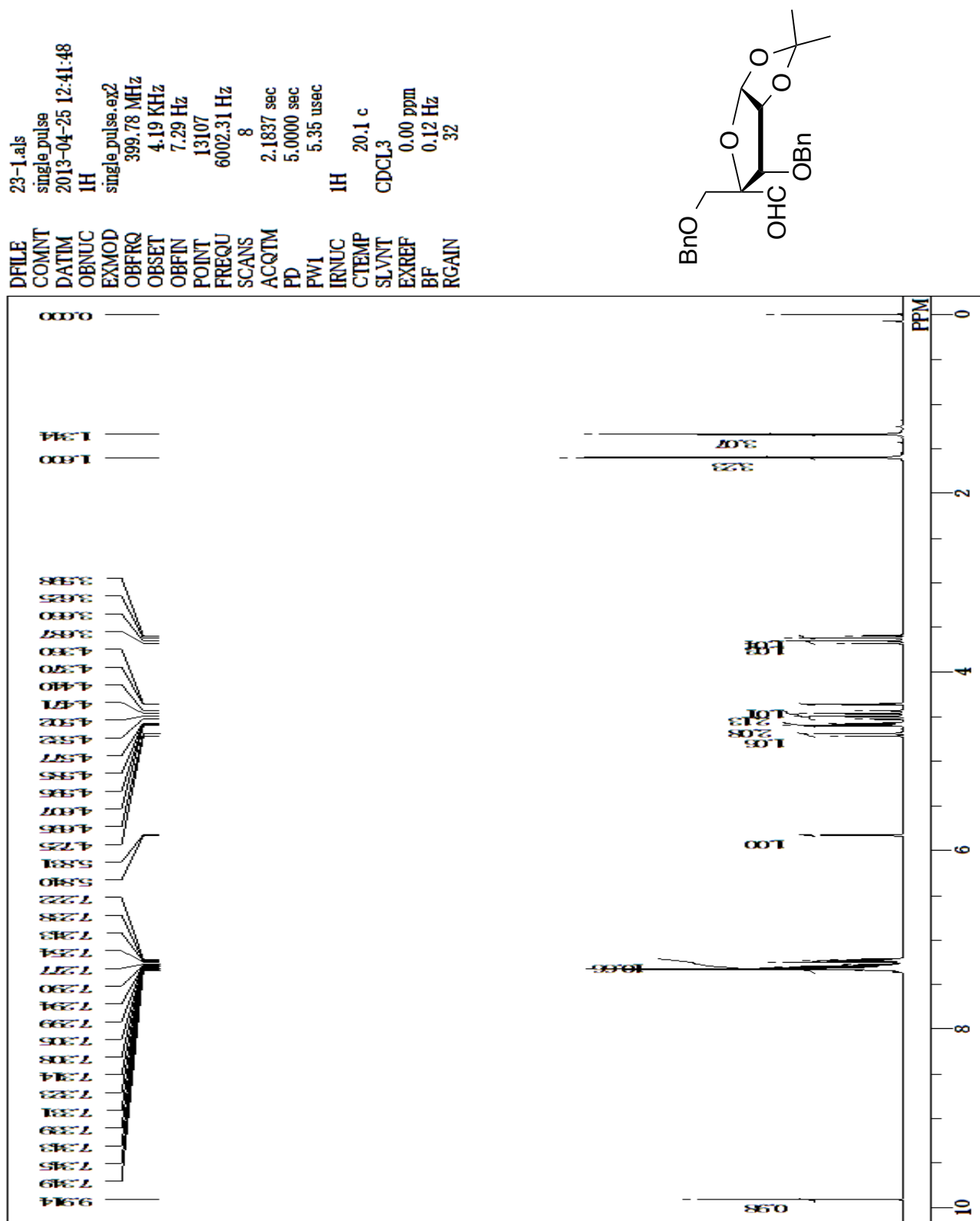


Figure S2. Compound 2 (<sup>13</sup>C NMR, CDCl<sub>3</sub>, 100 MHz)

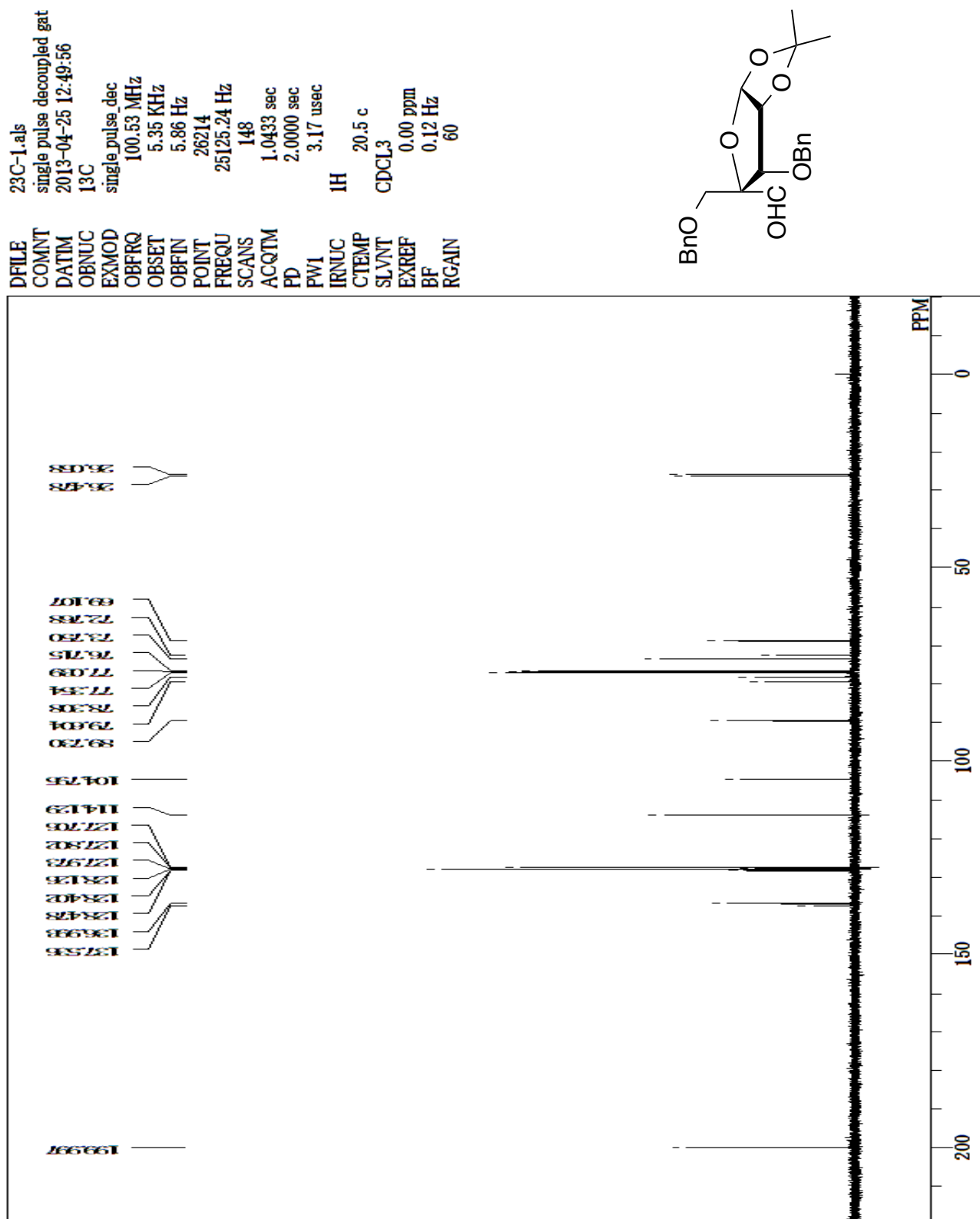


Figure S3. Compound 3 (<sup>1</sup>H NMR, CDCl<sub>3</sub>, 400 MHz)

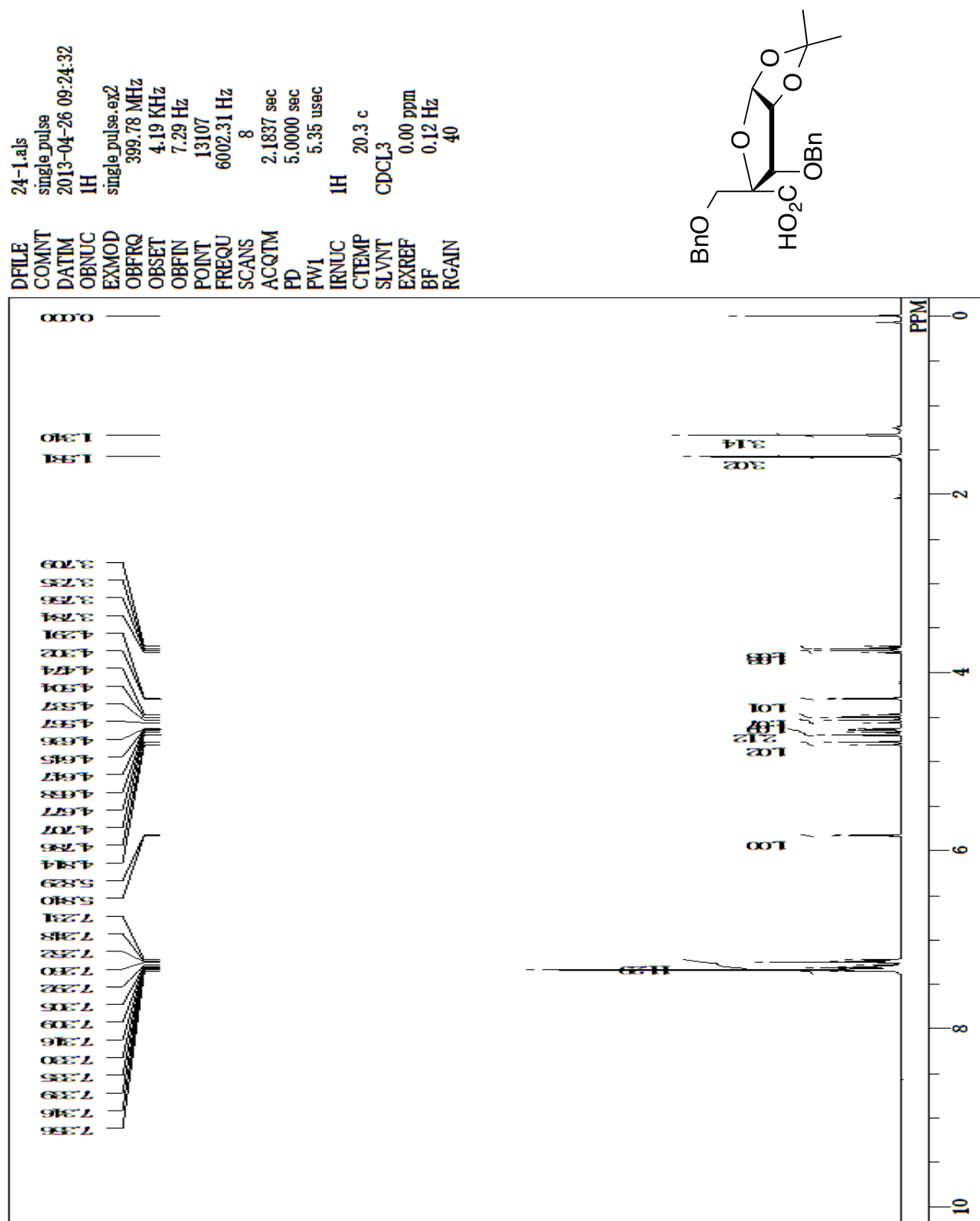
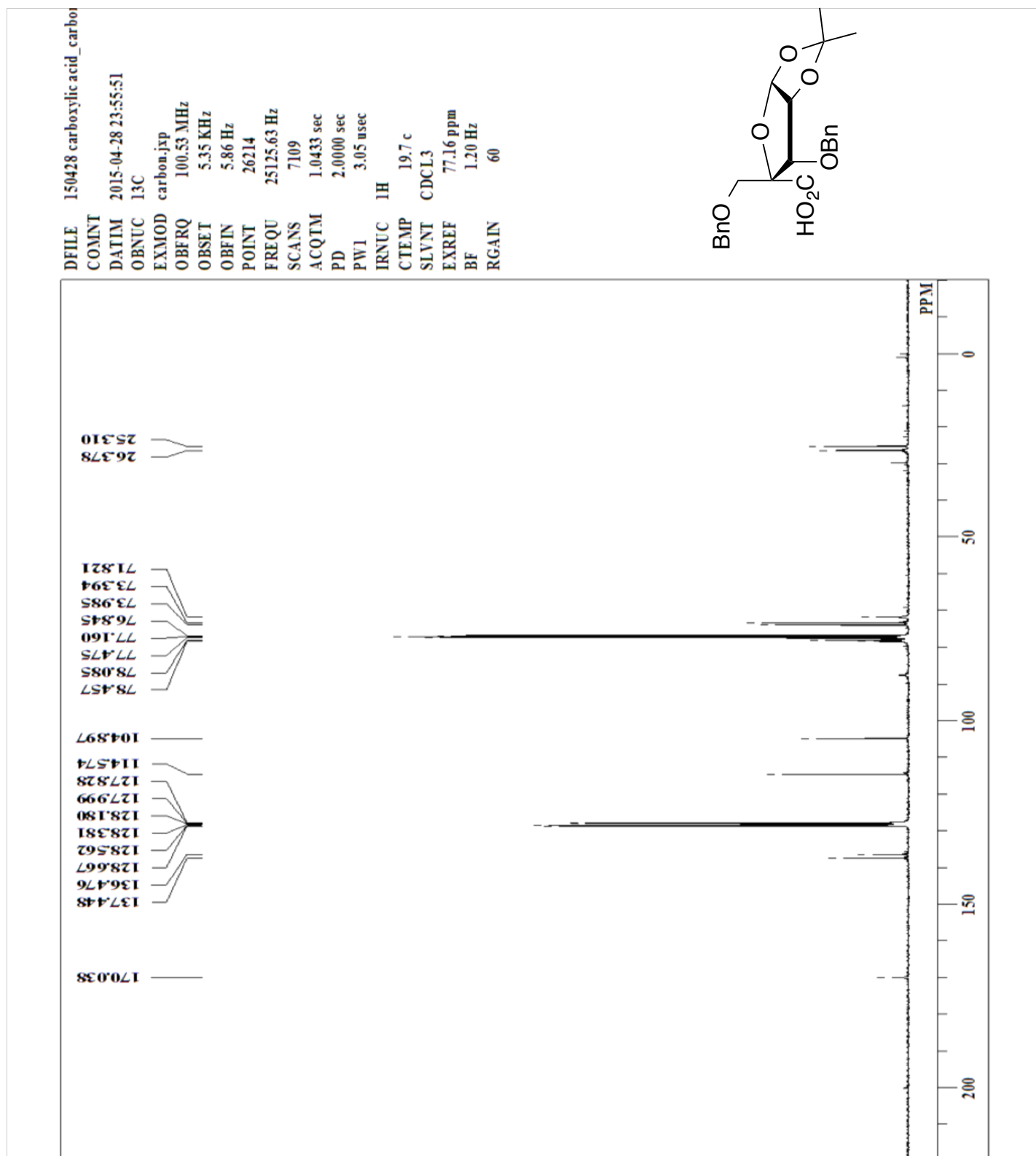


Figure S4. Compound 3 (<sup>13</sup>C NMR, CDCl<sub>3</sub>, 100 MHz)



**Figure S5.** Compound **4** (<sup>1</sup>H NMR, CDCl<sub>3</sub>, 400 MHz)

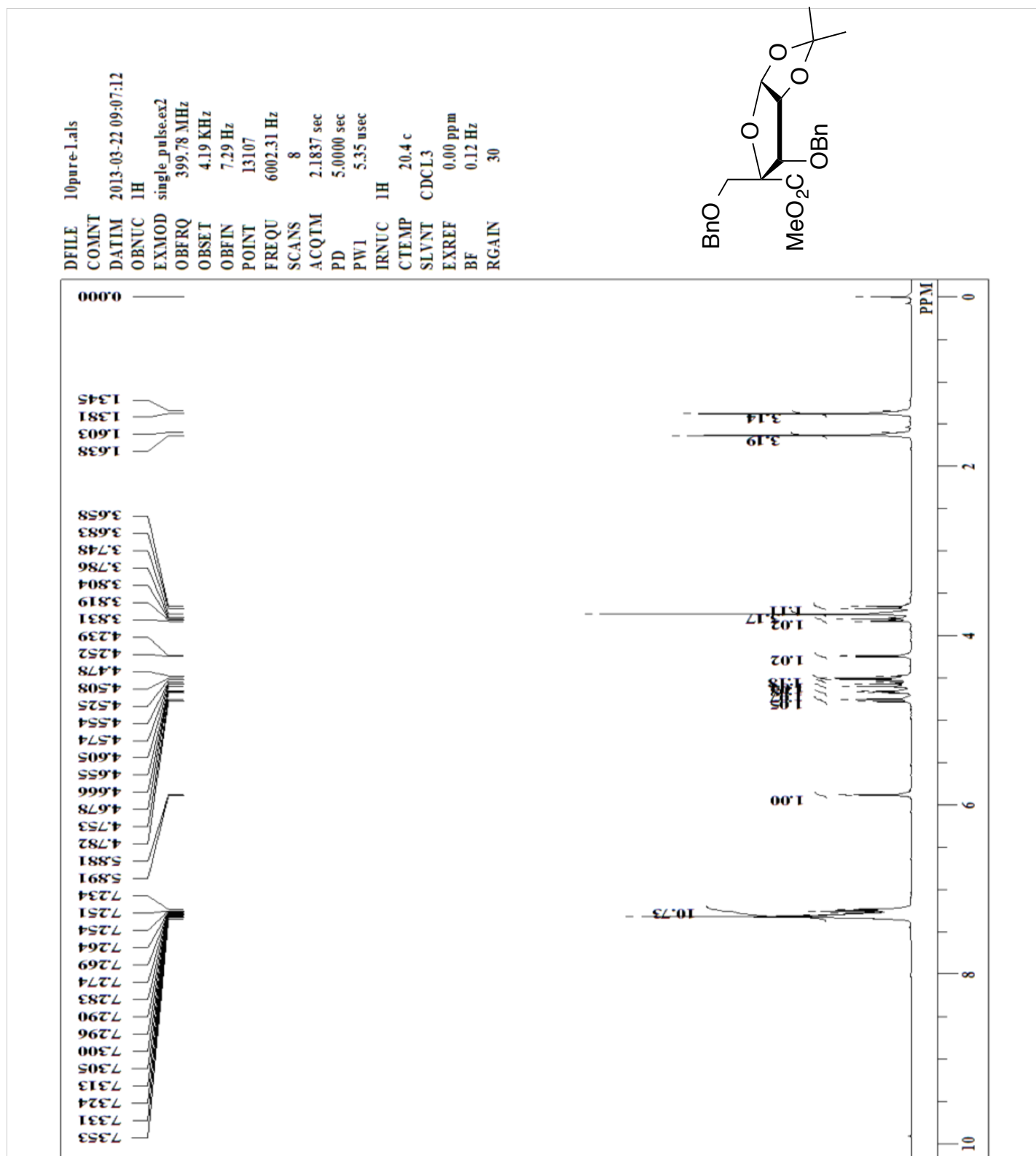




Figure S6. Compound 4 (<sup>13</sup>C NMR, CDCl<sub>3</sub>, 100 MHz)

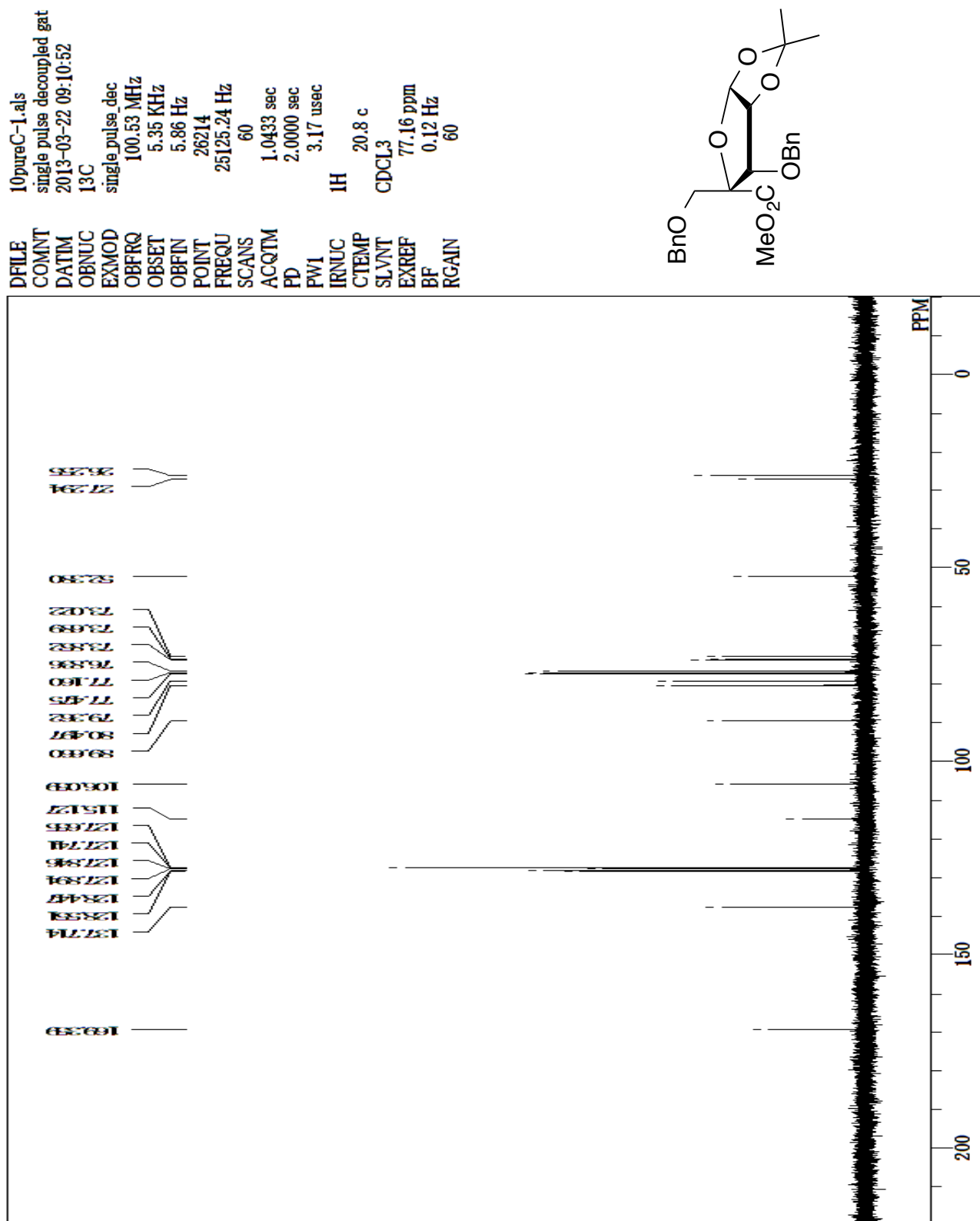


Figure S7. Compound 5 (<sup>1</sup>H NMR, CDCl<sub>3</sub>, 300 MHz)

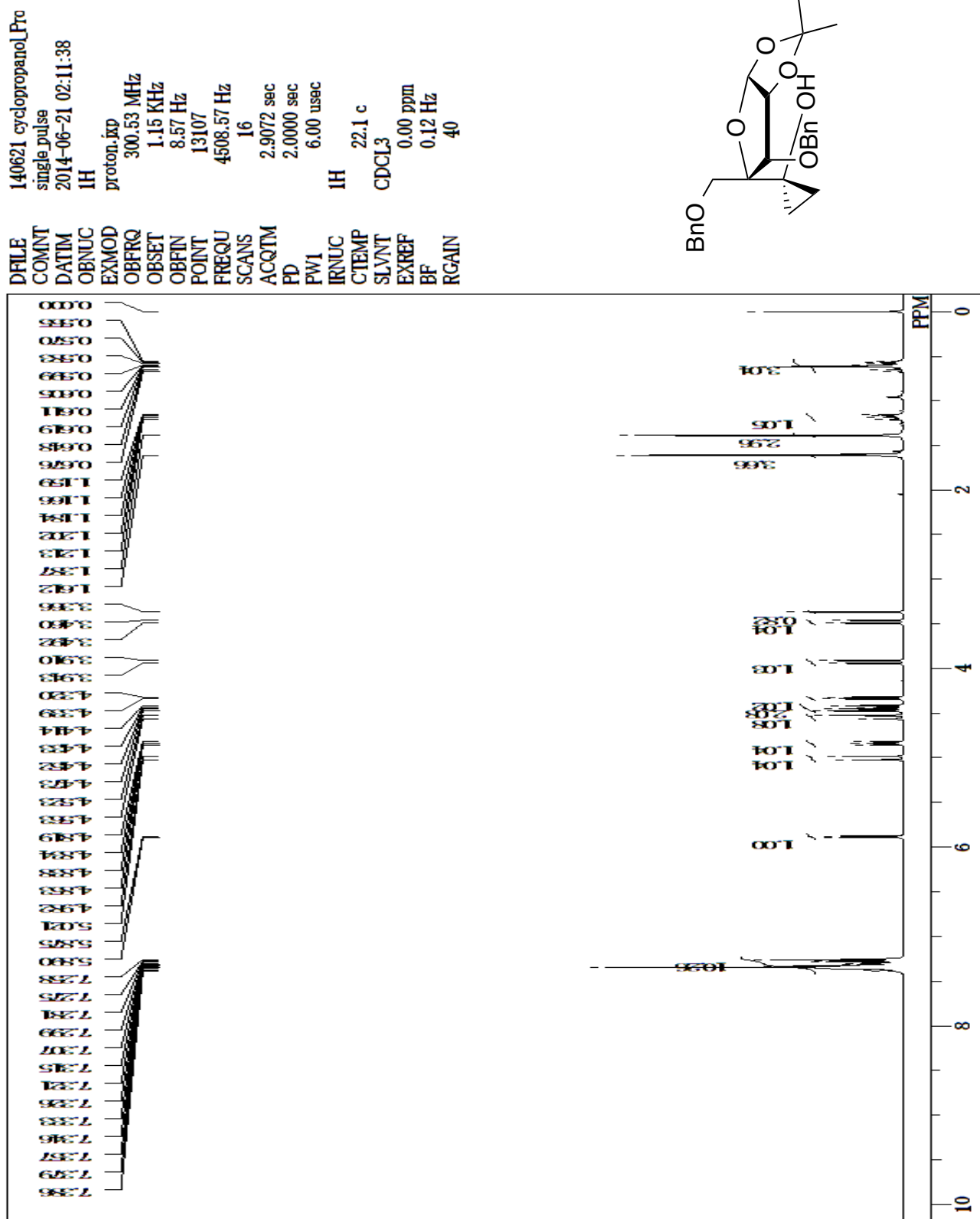


Figure S8. Compound 5 (<sup>13</sup>C NMR, CDCl<sub>3</sub>, 100 MHz)

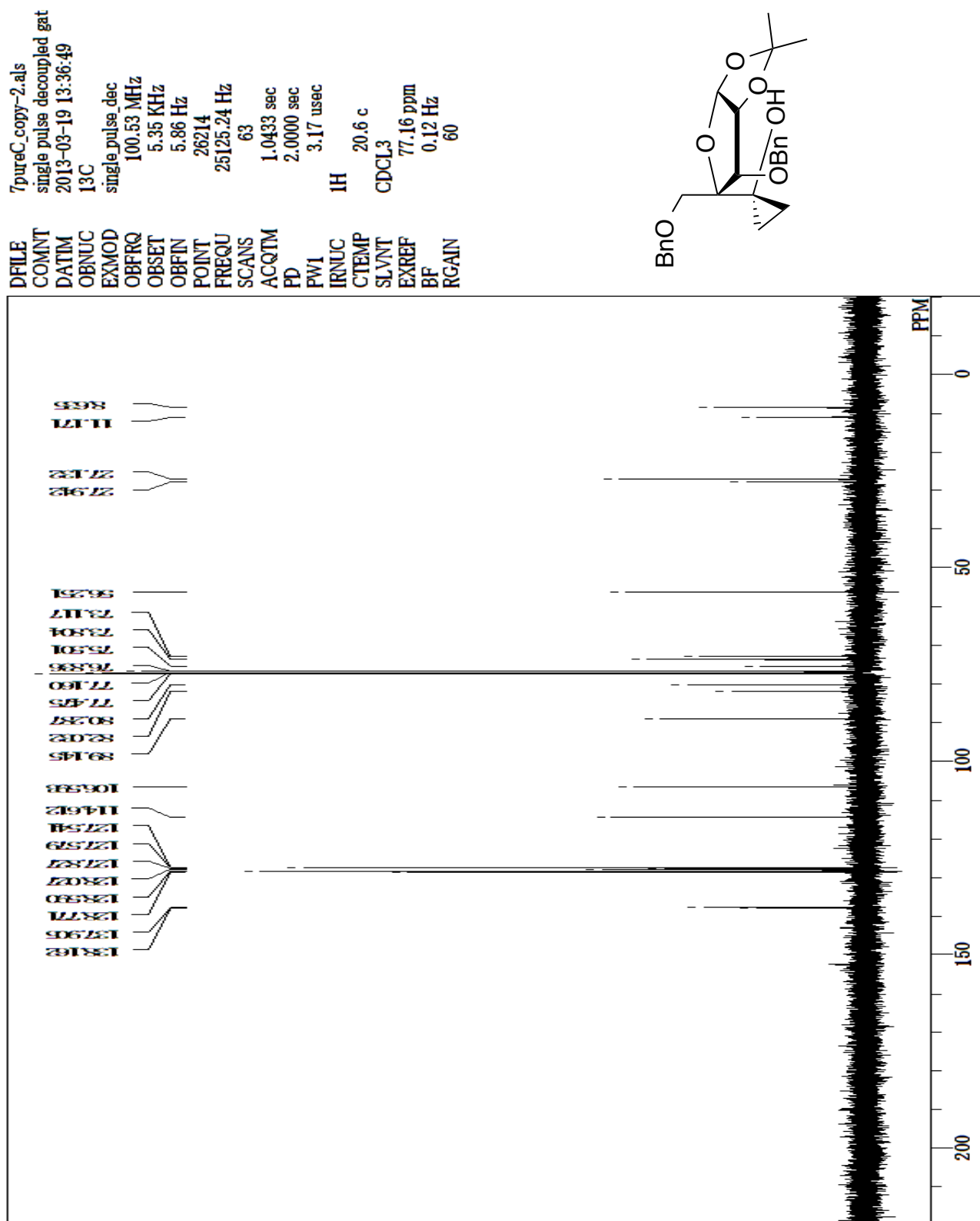


Figure S9. Compound 6 (<sup>1</sup>H NMR, CDCl<sub>3</sub>, 300 MHz)

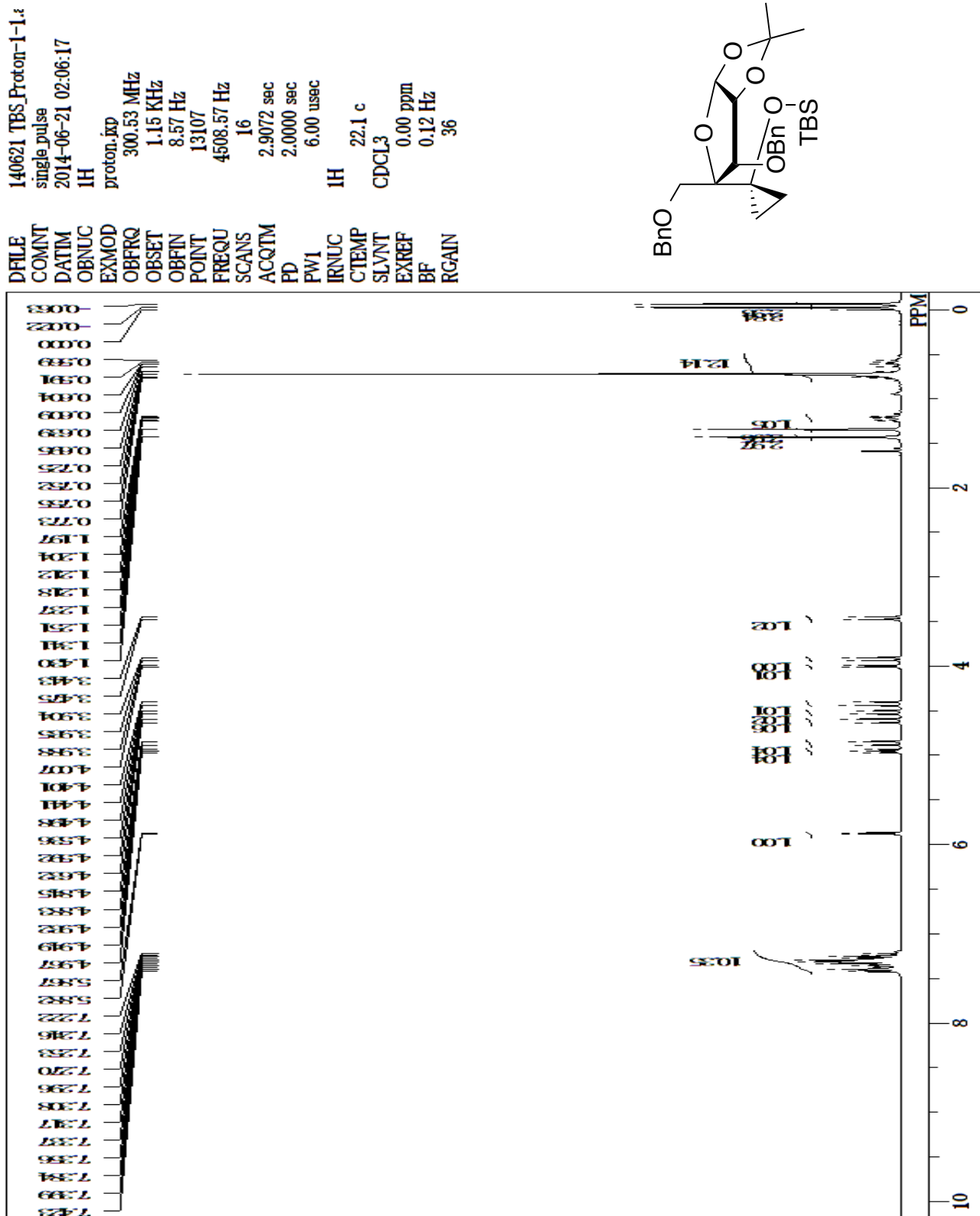


Figure S10. Compound 6 (<sup>13</sup>C NMR, CDCl<sub>3</sub>, 100 MHz)

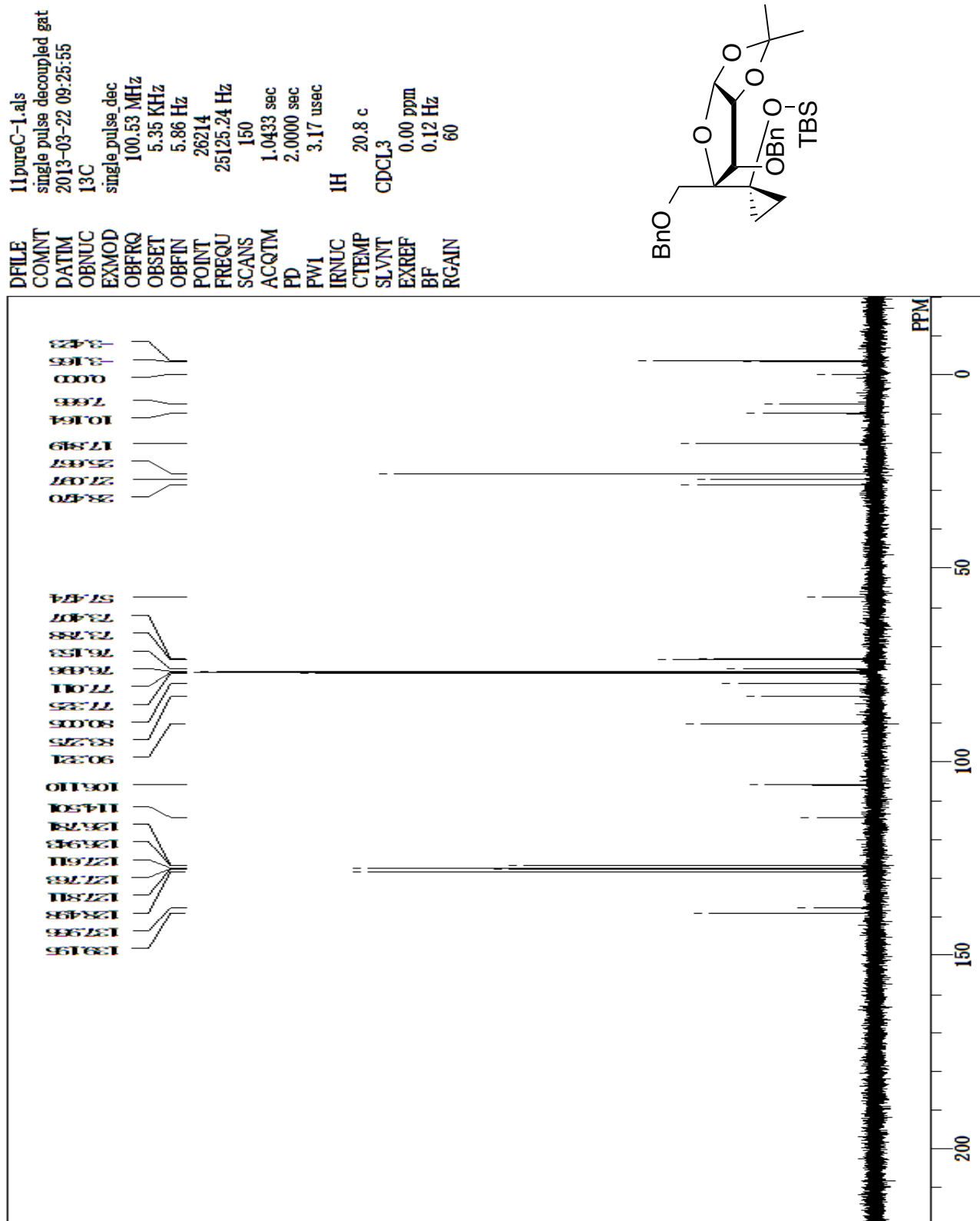




Figure S12. Compound 8 ( $^{13}\text{C}$  NMR,  $\text{CDCl}_3$ , 75 MHz)

DFILE 140702 T column\_Carbon-  
 COMNT single pulse decoupled gai  
 DATIM 03-07-2014 00:07:30  
 OBNUC  $^{13}\text{C}$   
 EXMOD carbon.kxp  
 OBFREQ 75.57 MHz  
 OBSET 5.79 KHz  
 OBFIN 1.08 Hz  
 POINT 26214  
 FREQU 18939.39 Hz  
 SCANS 267  
 ACQTM 1.3841 sec  
 PD 2.0000 sec  
 PW1 3.60 usec  
 IRNUC 1H  
 CTEMP 23.1 c  
 SLVNT  $\text{CDCl}_3$   
 EXREF 77.16 ppm  
 BF 0.12 Hz  
 RGAIN 60

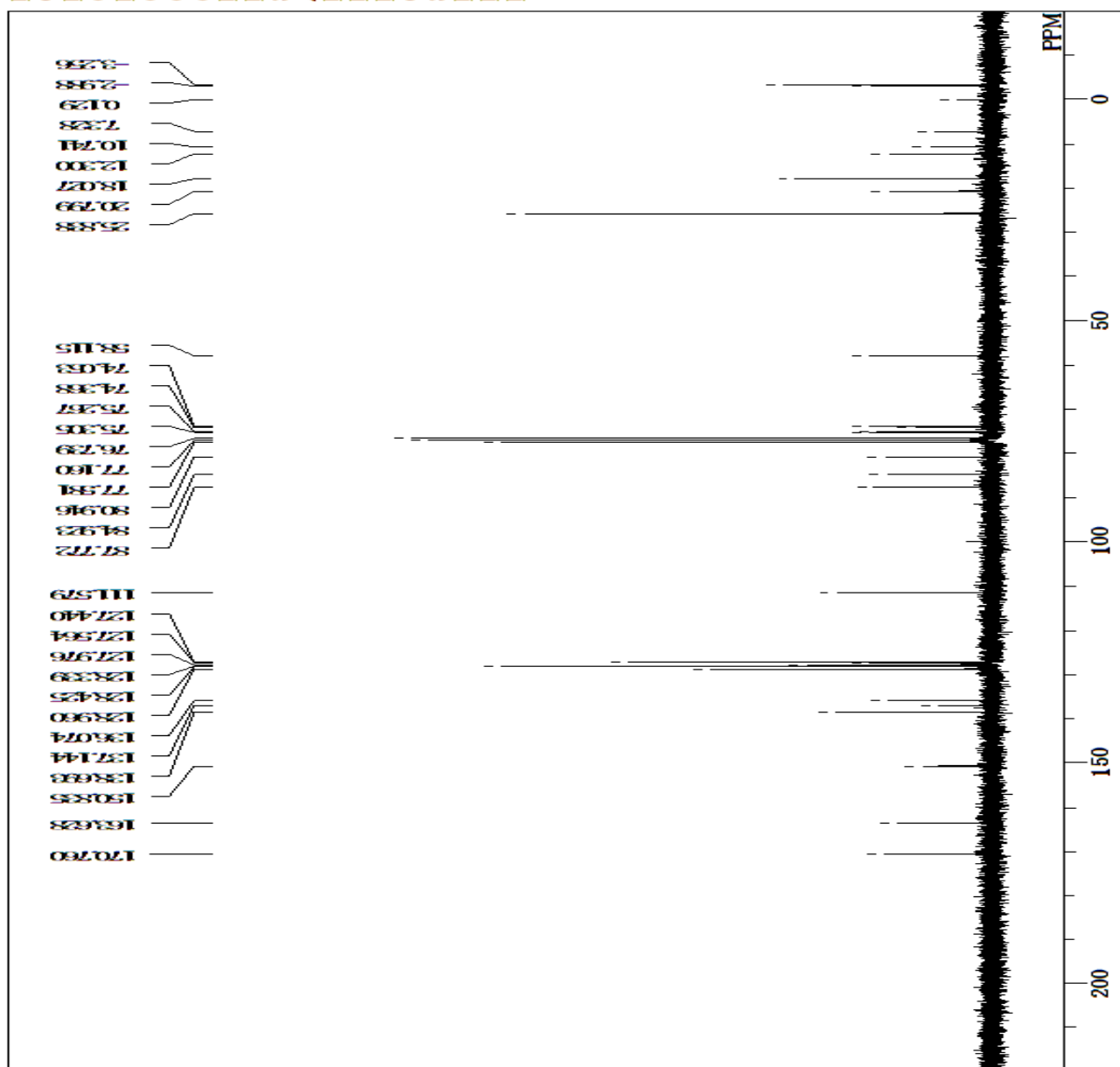
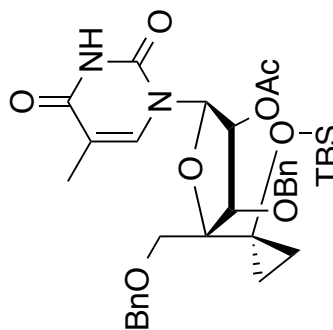


Figure S13. Compound 9 (<sup>1</sup>H NMR, CDCl<sub>3</sub>, 300 MHz)

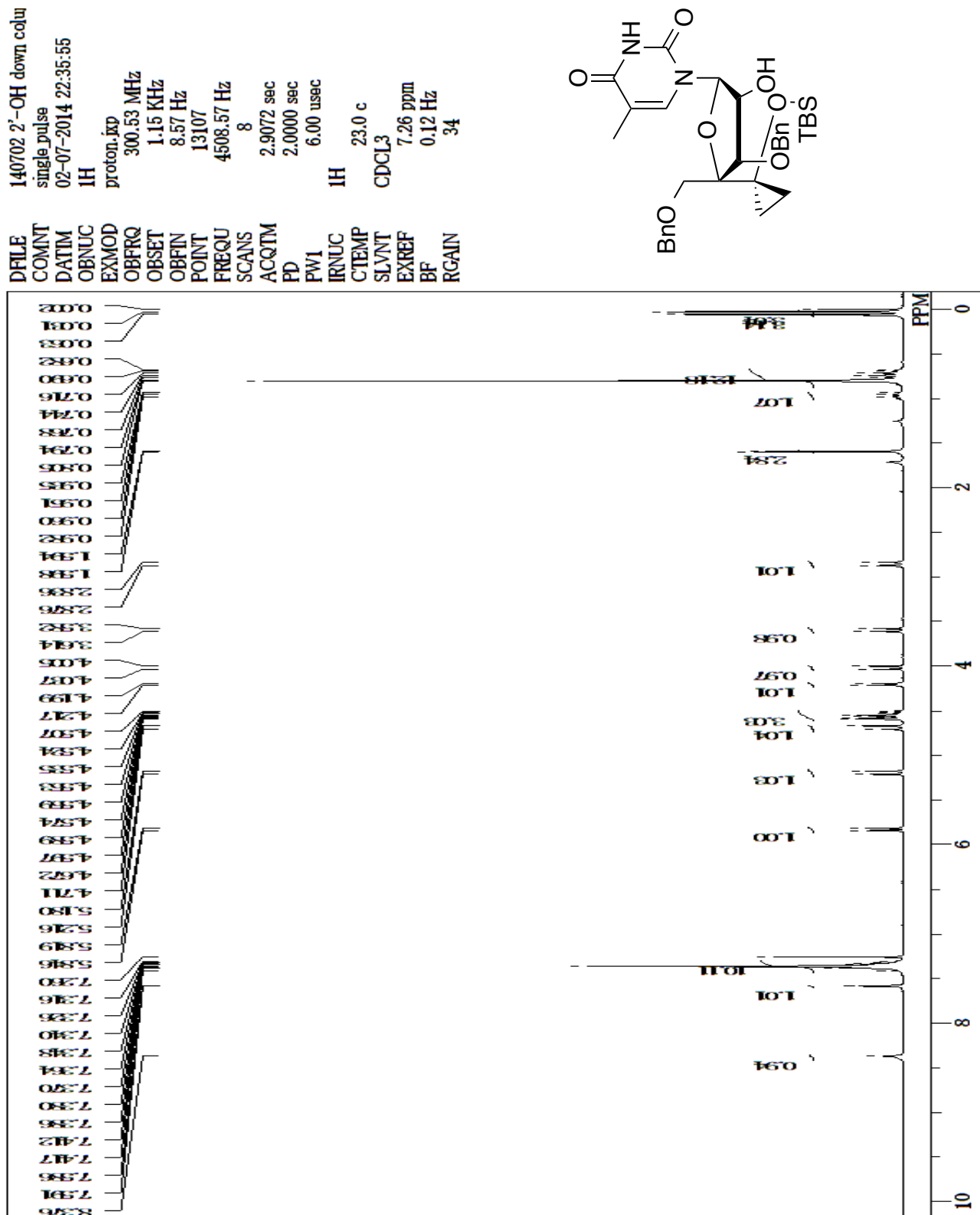




Figure S14. Compound 9 ( $^{13}\text{C}$  NMR,  $\text{CDCl}_3$ , 100 MHz)

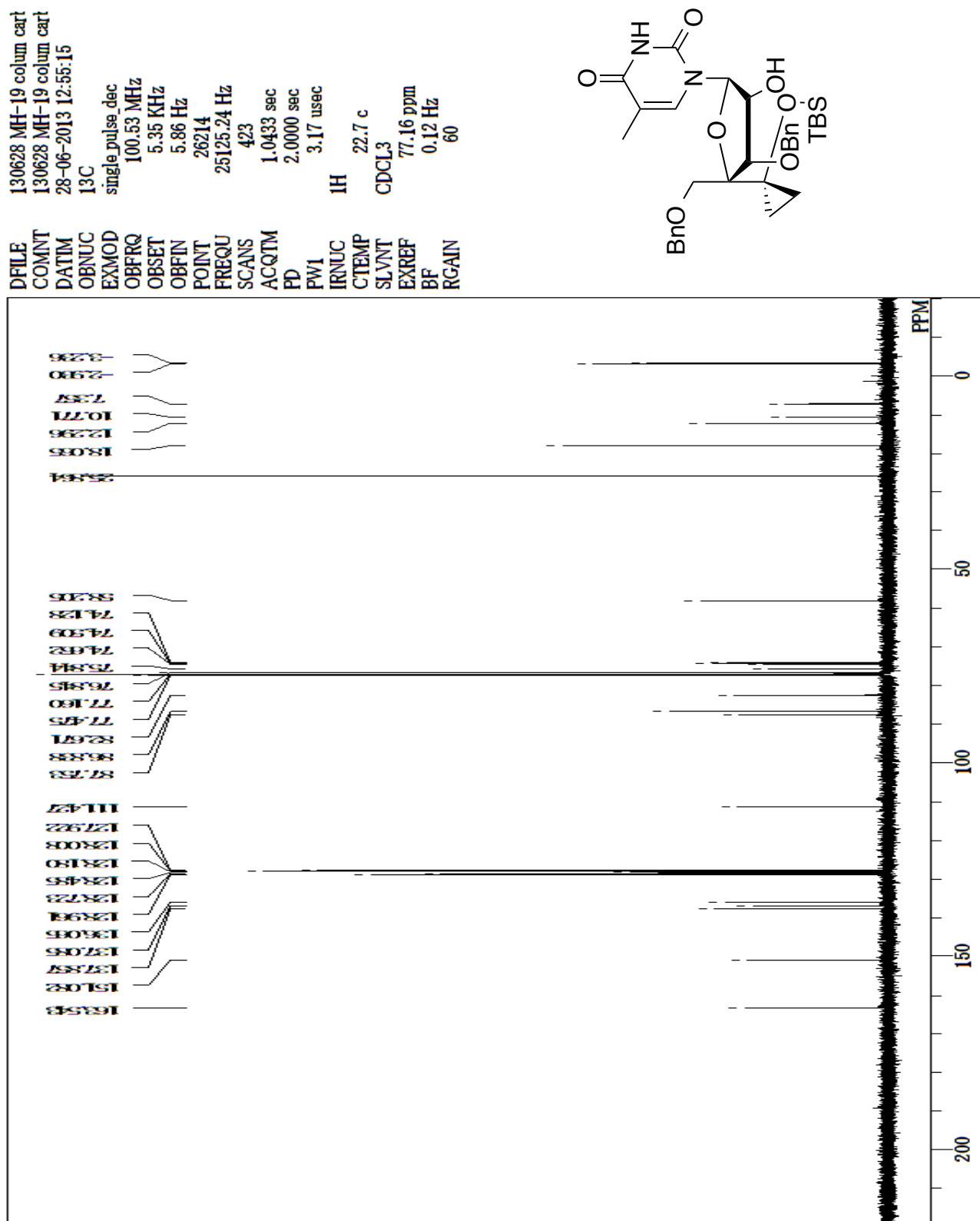


Figure S15. Compound **10** (<sup>1</sup>H NMR, CDCl<sub>3</sub>, 300 MHz)

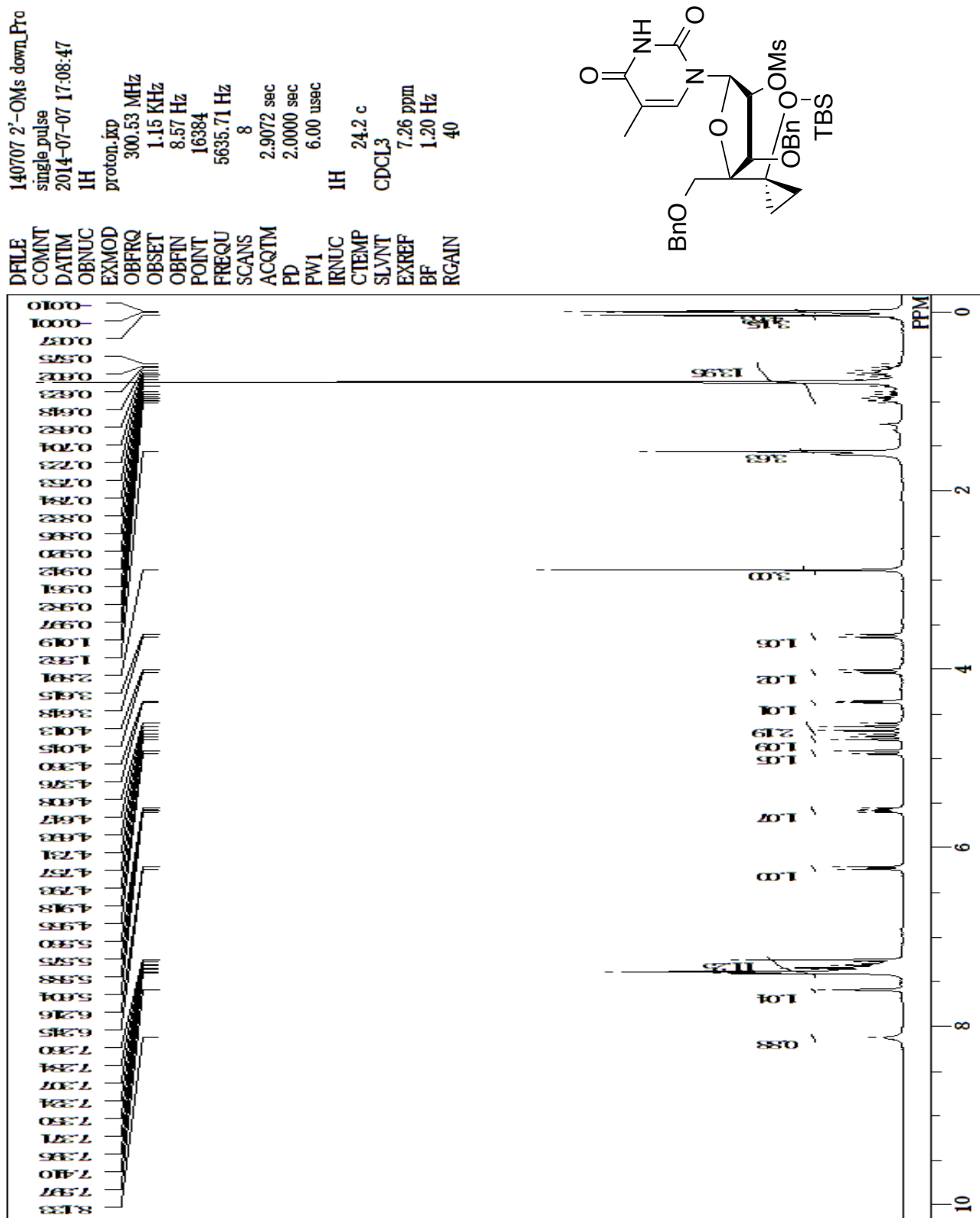


Figure S16. Compound 10 ( $^{13}\text{C}$  NMR,  $\text{CDCl}_3$ , 100 MHz)

DFILE 15pureC\_copy-6.als  
 COMNT single pulse decoupled gat  
 DATIM 2013-03-25 21:30:11  
 OBNUC  $^{13}\text{C}$   
 EXMOD single\_pulse\_dec  
 OBFRQ 100.53 MHz  
 OBSET 5.35 KHz  
 OBFIN 5.86 Hz  
 POINT 26214  
 FREQU 25125.24 Hz  
 SCANS 404  
 ACQTM 1.0433 sec  
 PD 2.0000 sec  
 PW1 3.17 usec  
 IRNUC  $^1\text{H}$   
 CTEMP 20.7 c  
 SLVNT  $\text{CDCl}_3$   
 EXREF 0.00 ppm  
 BF 0.12 Hz  
 RGAIN 60

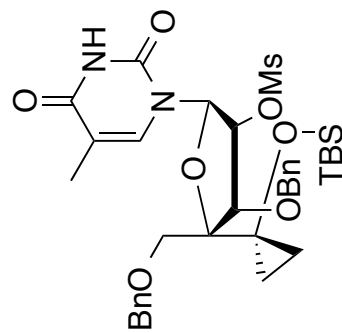
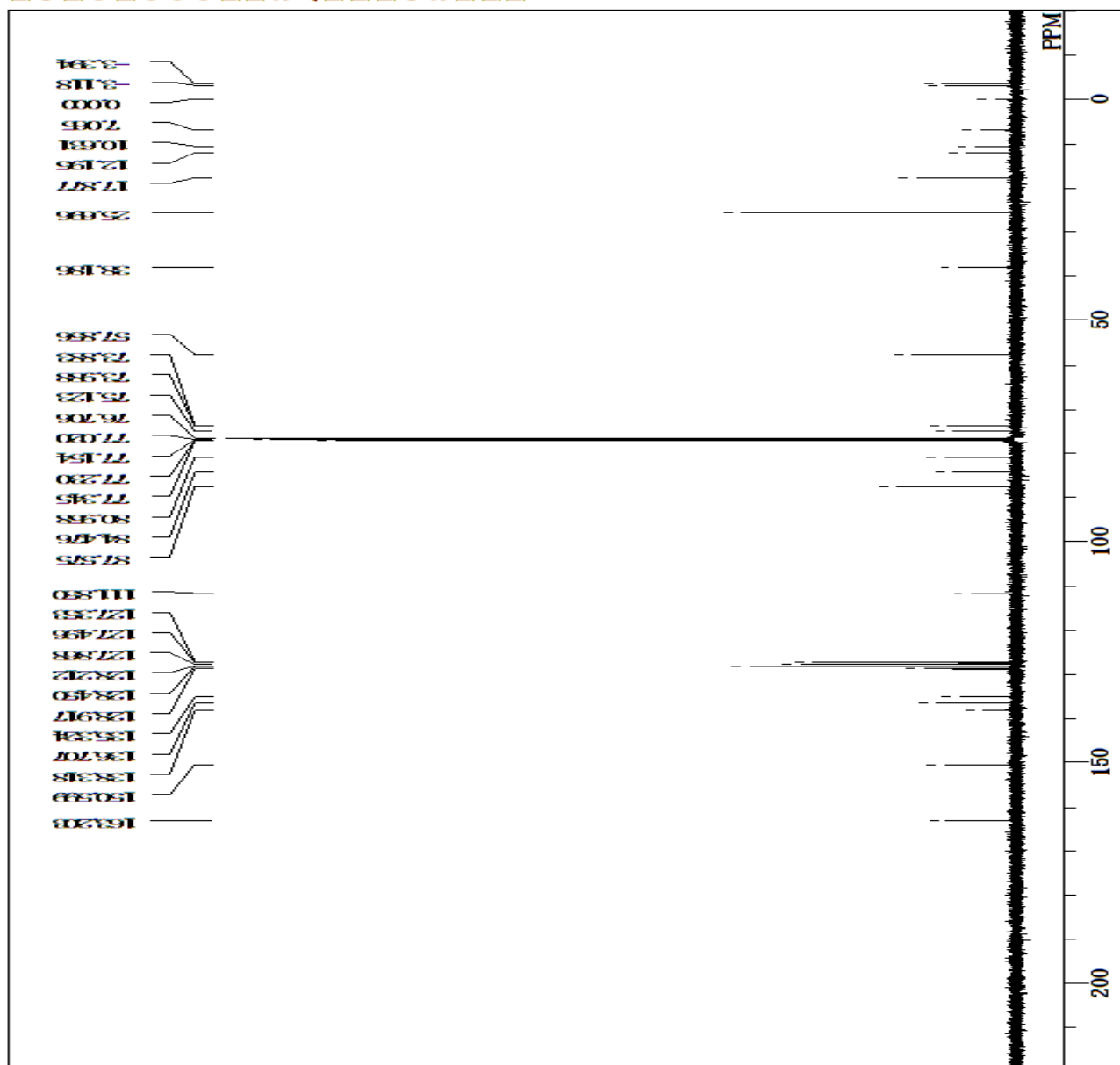


Figure S17. Compound 11 ( $^1\text{H}$  NMR,  $\text{CDCl}_3$ , 400 MHz)

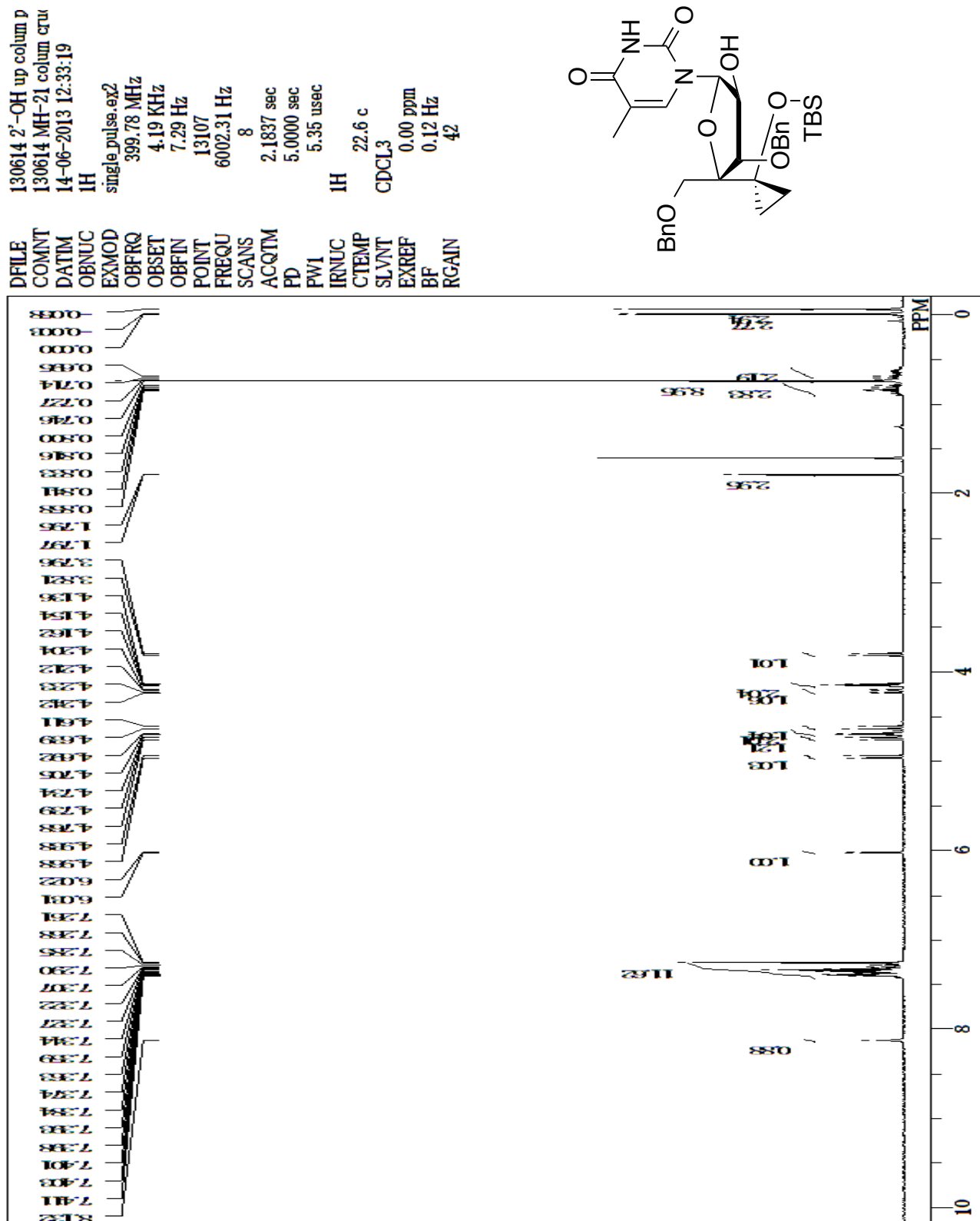


Figure S18. Compound 11 ( $^{13}\text{C}$  NMR,  $\text{CDCl}_3$ , 100 MHz)

DFILE 130614 2'-OH up column p  
 COMNT 130614 MH-21 column pur  
 DATIM 14-06-2013 14:04:32  
 OBNUC  $^{13}\text{C}$   
 EXMOD single\_pulse dec  
 OBFRQ 100.53 MHz  
 OBSET 5.35 KHz  
 OBFIN 5.86 Hz  
 POINT 26214  
 FREQU 25125.24 Hz  
 SCANS 174  
 ACQTM 1.0433 sec  
 PD 2.0000 sec  
 PW1 3.17 usec  
 IRNUC  $^1\text{H}$   
 CTEMP 23.2 c  
 SLVNT  $\text{CDCl}_3$   
 EXREF 0.00 ppm  
 BF 0.12 Hz  
 RGAIN 60

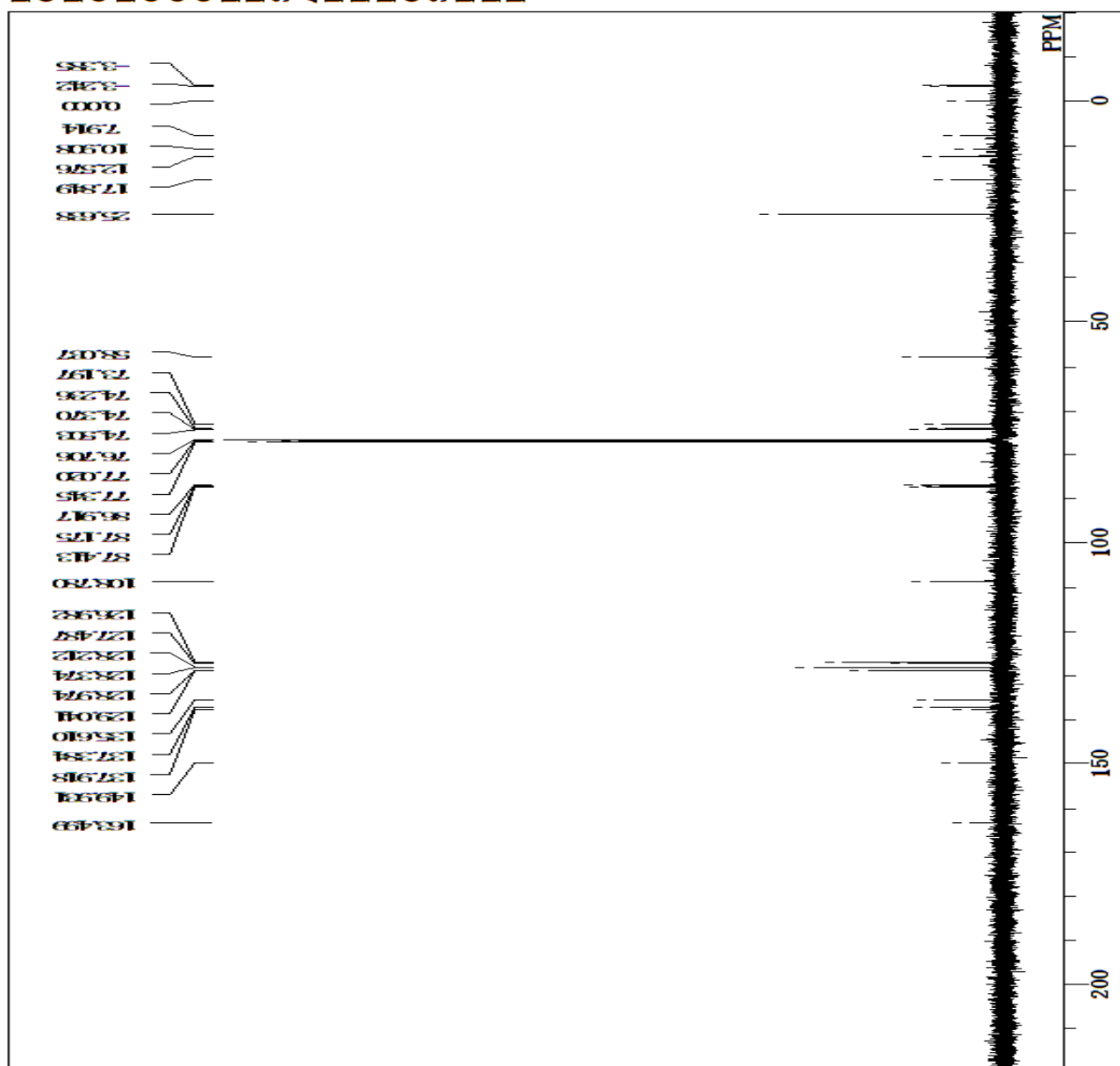
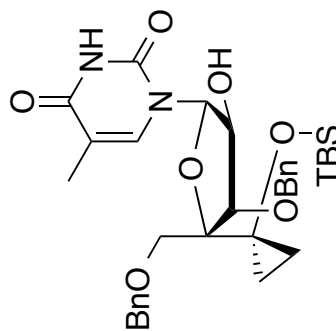


Figure S19. Compound 13 (<sup>1</sup>H NMR, CDCl<sub>3</sub>, 300MHz)

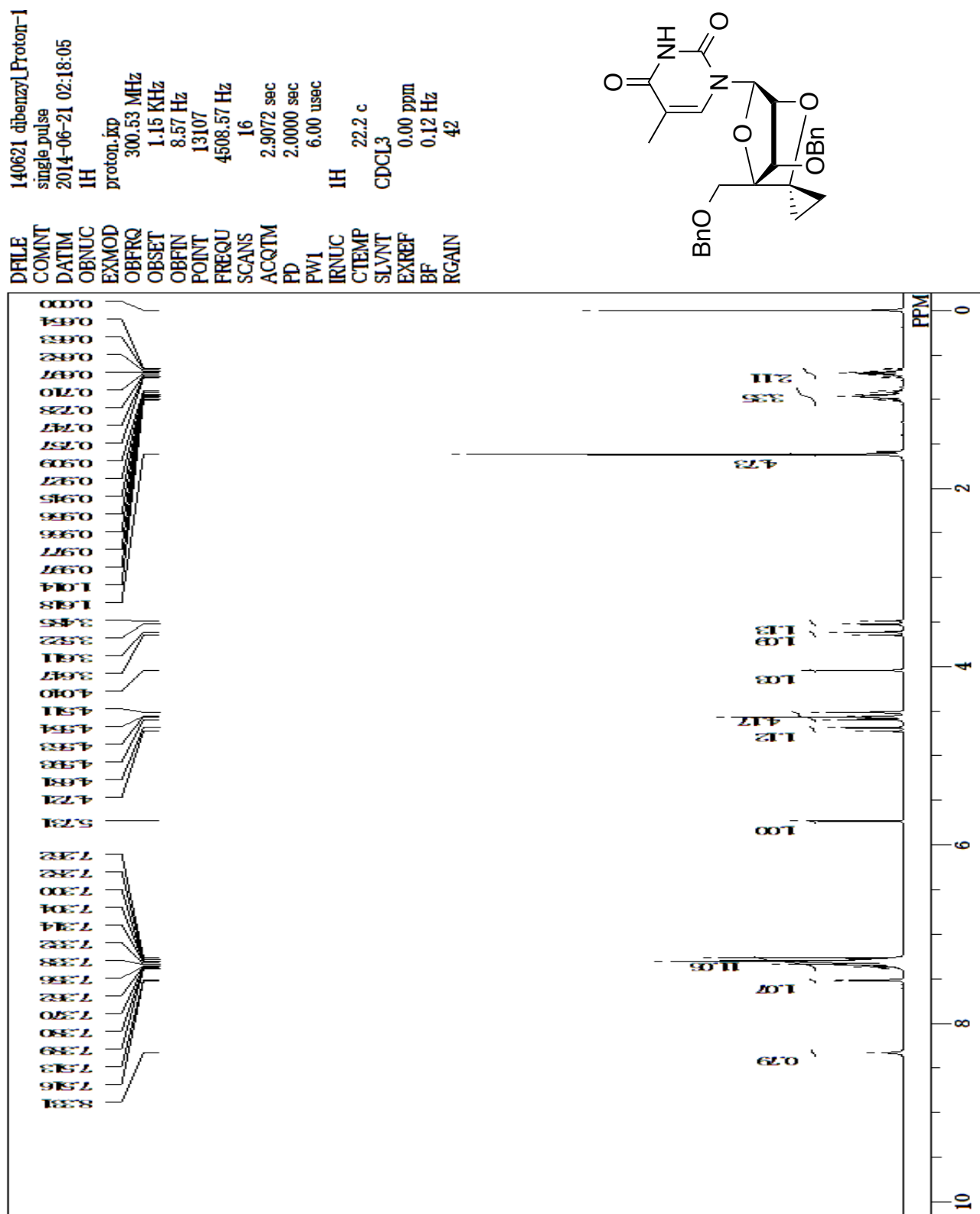


Figure S20. Compound 13 (<sup>13</sup>C NMR, CDCl<sub>3</sub>, 100MHz)

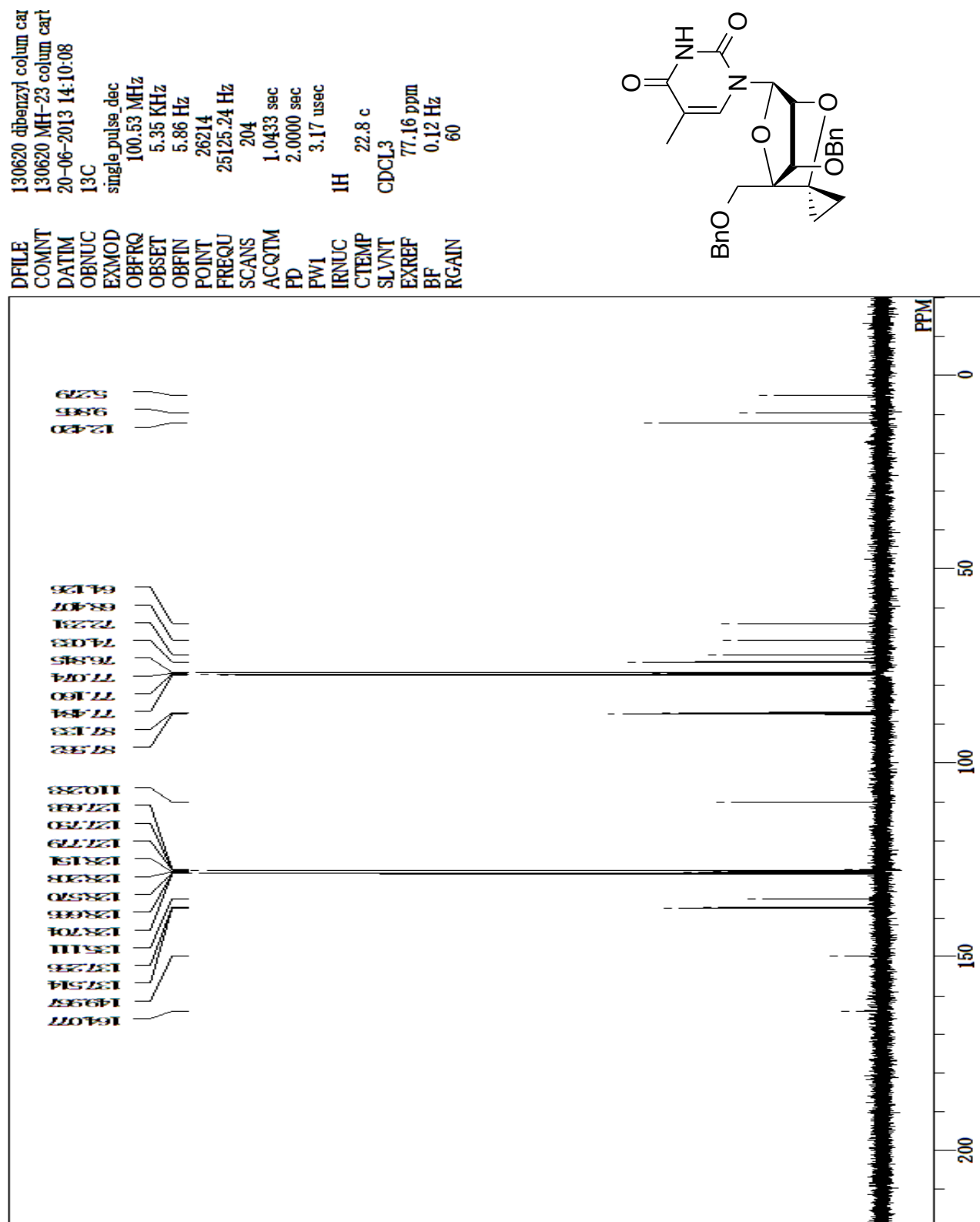


Figure S21. Compound 14 (<sup>1</sup>H NMR, CDCl<sub>3</sub>, 300 MHz)

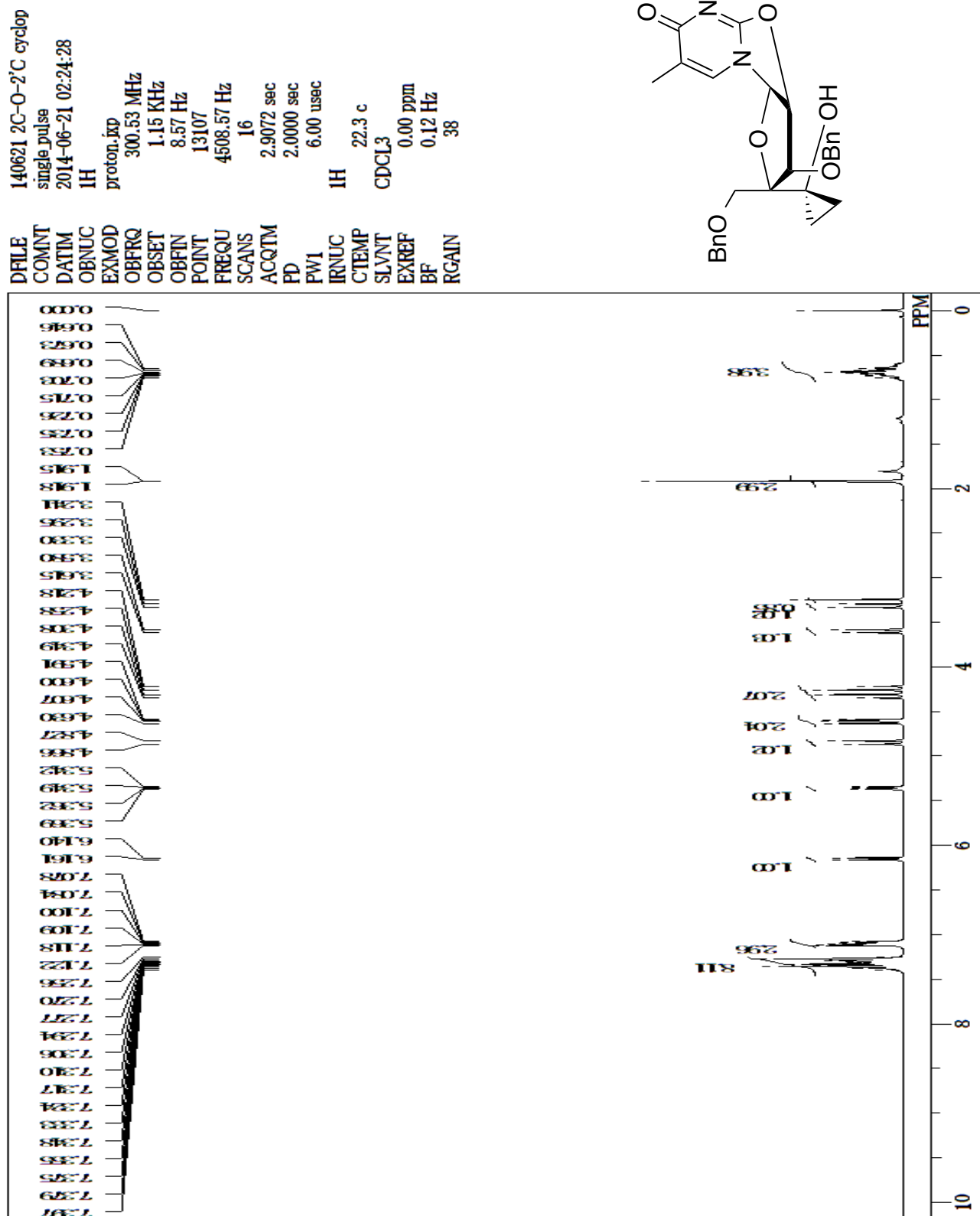




Figure S22. Compound 14 (<sup>13</sup>C NMR, CDCl<sub>3</sub>, 75 MHz)

DFILE 140702 2C-O-2'C cyclop  
 COMNT single pulse decoupled gai  
 DATIM 03-07-2014 00:28:58  
 OBNUC 13C  
 EXMOD carbon.kxp  
 OBFRQ 75.57 MHz  
 OBSET 5.79 KHz  
 OBFIN 1.08 Hz  
 POINT 26214  
 FREQU 18939.39 Hz  
 SCANS 273  
 ACQTM 1.3841 sec  
 PD 2.0000 sec  
 PW1 3.60 usec  
 IRNUC 1H  
 CTEMP 23.1 c  
 SLVNT CDCl3  
 EXREF 77.16 ppm  
 BF 0.12 Hz  
 RGAIN 60

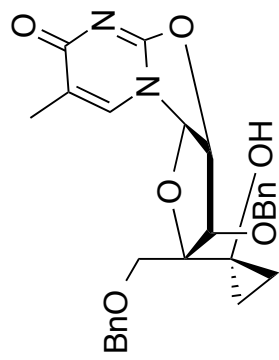
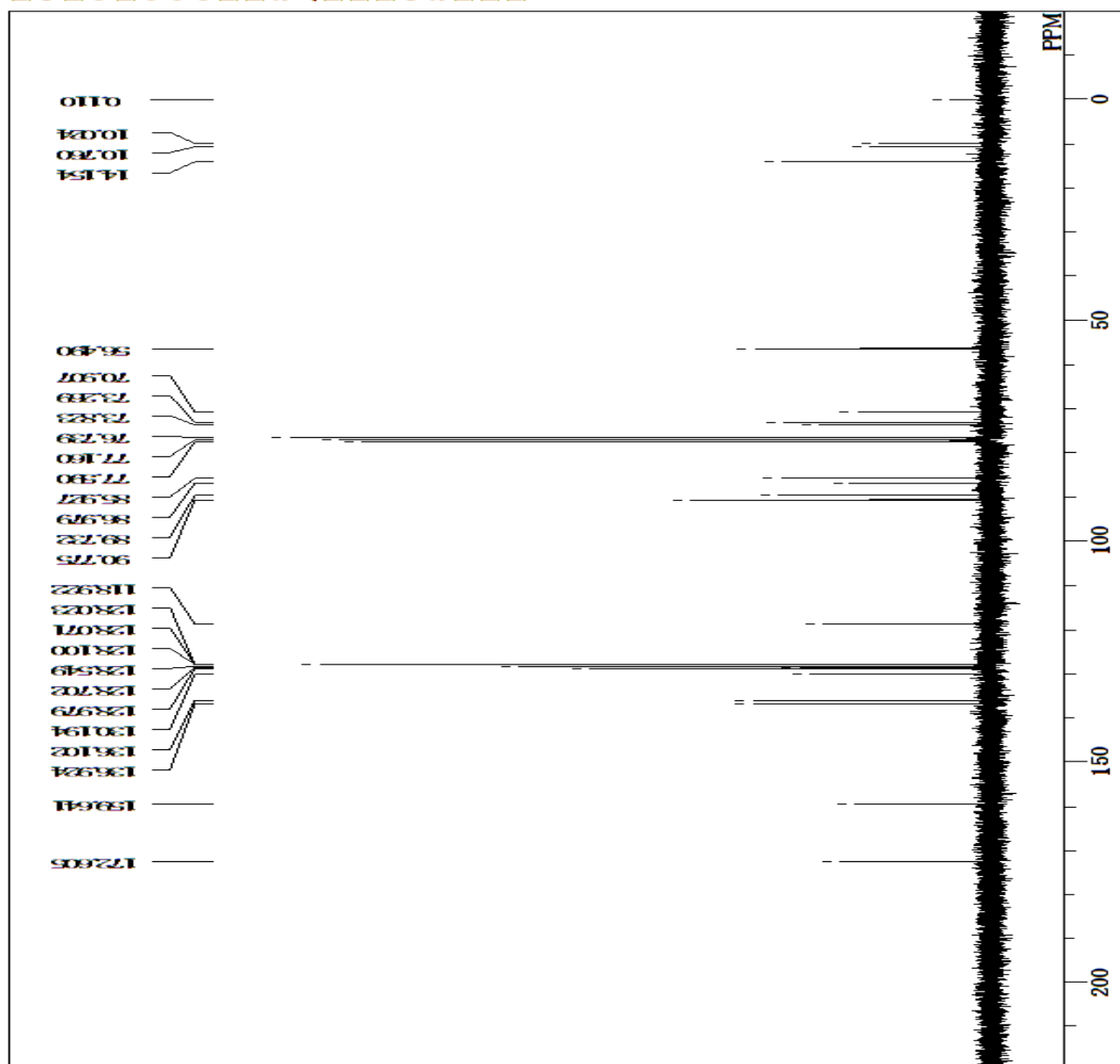


Figure S23. Compound 15 (<sup>1</sup>H NMR, CD<sub>3</sub>OD, 300 MHz)

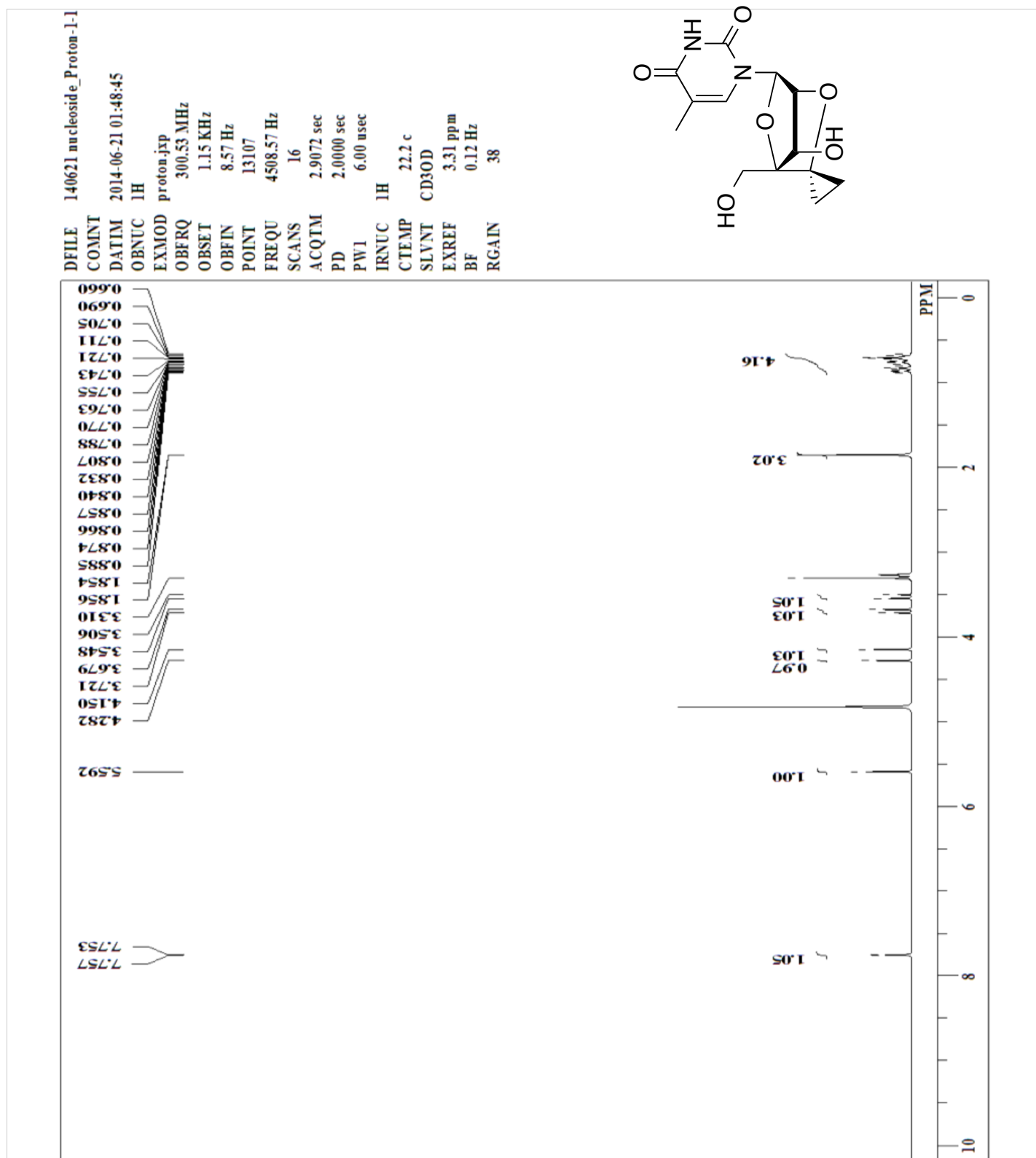


Figure S24. Compound 15 (<sup>13</sup>C NMR, CD<sub>3</sub>OD, 75 MHz)

DFILE 131218 nucleoside pure ce  
 COMNT 131218 MH-24 pure carb  
 DATM 18-12-2013 00:19:07  
 OBNUC <sup>13</sup>C  
 EXMOD single\_pulse\_dec  
 OBFRQ 75.57 MHz  
 OBSET 5.79 KHz  
 OBFIN 1.08 Hz  
 POINT 26214  
 FREQU 18939.11 Hz  
 SCANS 185  
 ACQTM 1.3841 sec  
 PD 2.0000 sec  
 PW1 3.13 usec  
 IRNUC <sup>1</sup>H  
 CTEMP 19.5 c  
 SLVNT CD3OD  
 EXREF 49.00 ppm  
 BF 0.12 Hz  
 RGAIN 60

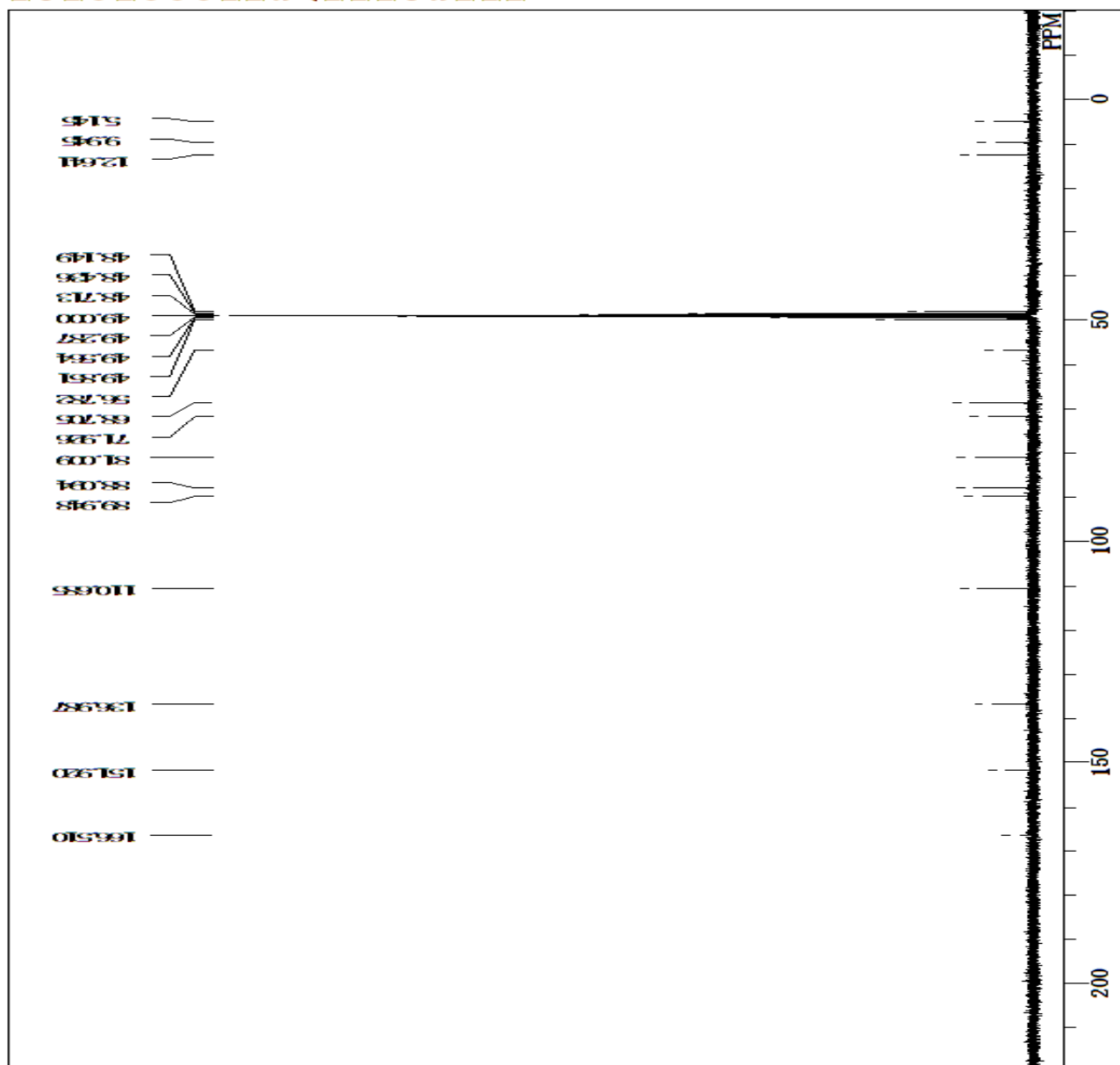
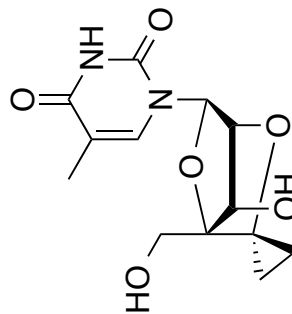


Figure S25. Compound 16 ( $^1\text{H}$  NMR,  $\text{CD}_3\text{OD}$ , 300 MHz)

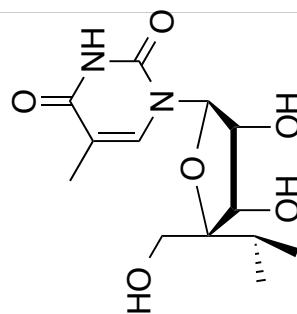
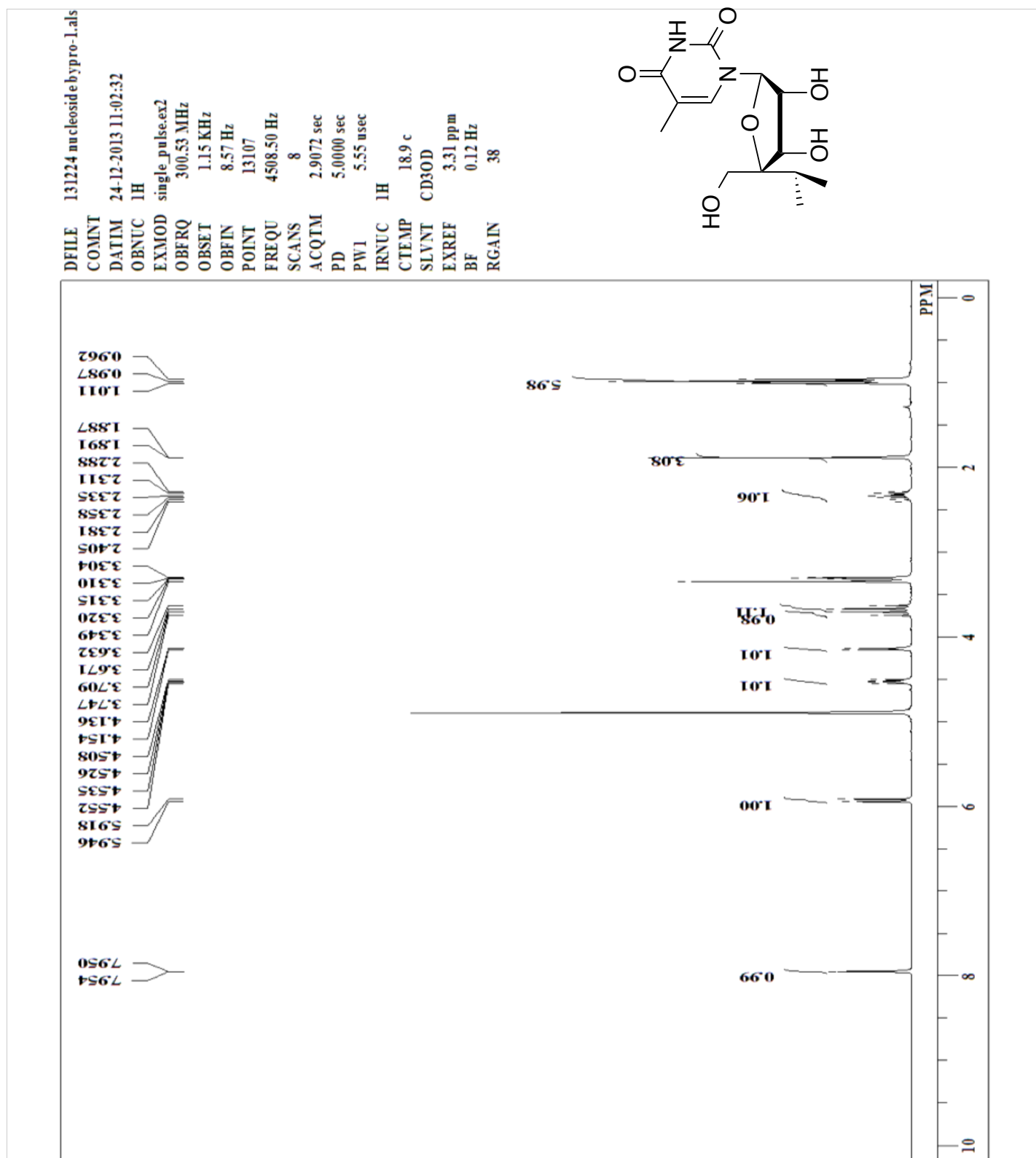


Figure S26. Compound 16 ( $^{13}\text{C}$  NMR,  $\text{CD}_3\text{OD}$ , 75 MHz)

DFILE 140313 MH-24 bypro\_Cat  
 COMINT single pulse decoupled gai  
 DATIM 2014-03-13 23:11:05  
 OBNUC  $^{13}\text{C}$   
 EXMOD carbon.kxp  
 OBFRQ 75.57 MHz  
 OBSET 5.79 KHz  
 OBFIN 1.08 Hz  
 POINT 26214  
 FREQU 18939.39 Hz  
 SCANS 96  
 ACQTM 0.0000 sec  
 PD 2.0000 sec  
 PW1 3.60 usec  
 IRNUC  $^1\text{H}$   
 CTEMP 18.1 c  
 SLVNT  $\text{CD}_3\text{OD}$   
 EXREF 49.00 ppm  
 BF 0.25 Hz  
 RGAIN 60

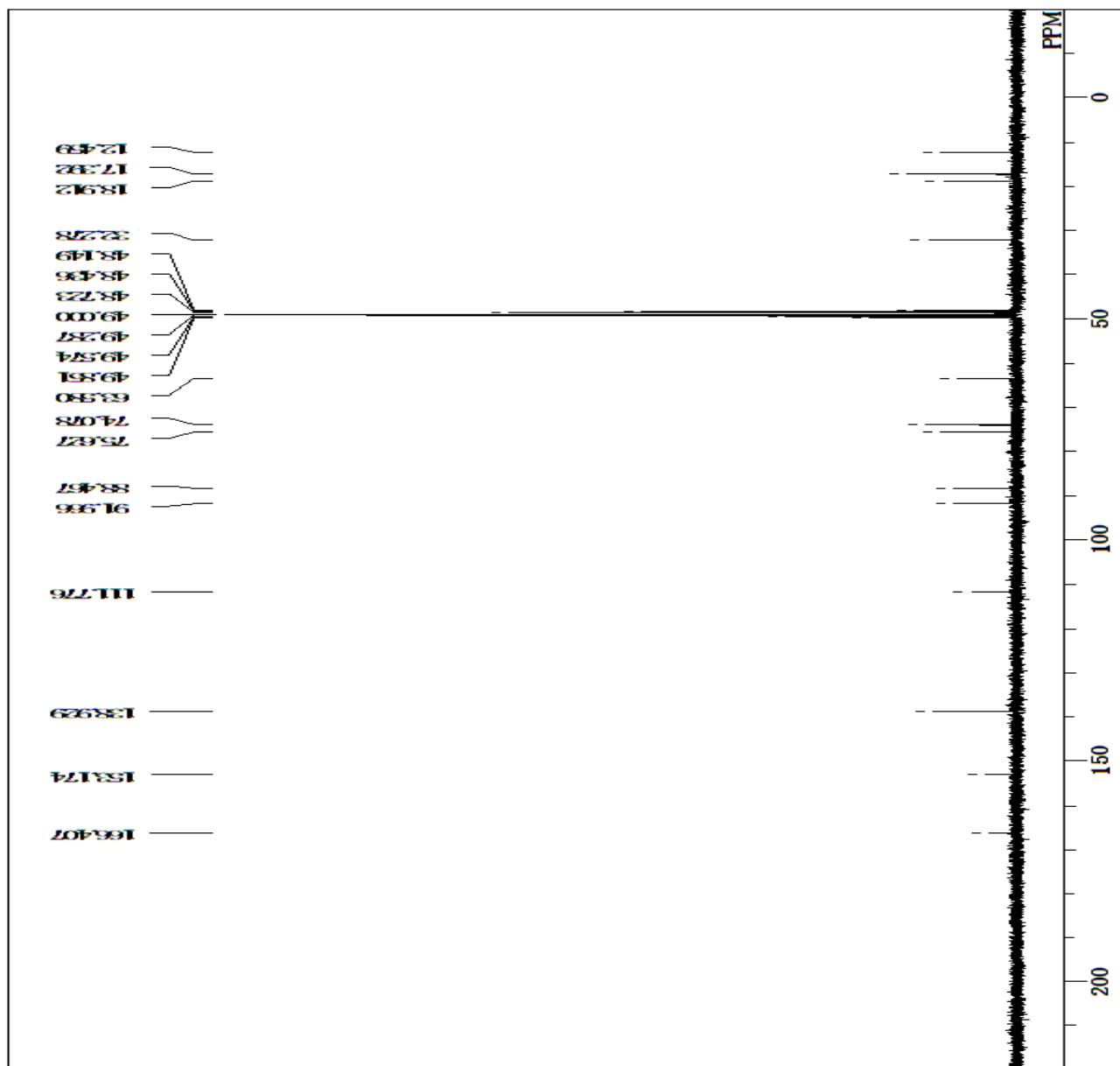
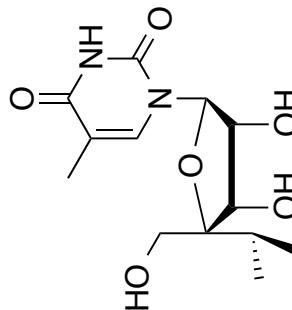


Figure S27. Compound 17 (<sup>1</sup>H NMR, CDCl<sub>3</sub>, 300 MHz)

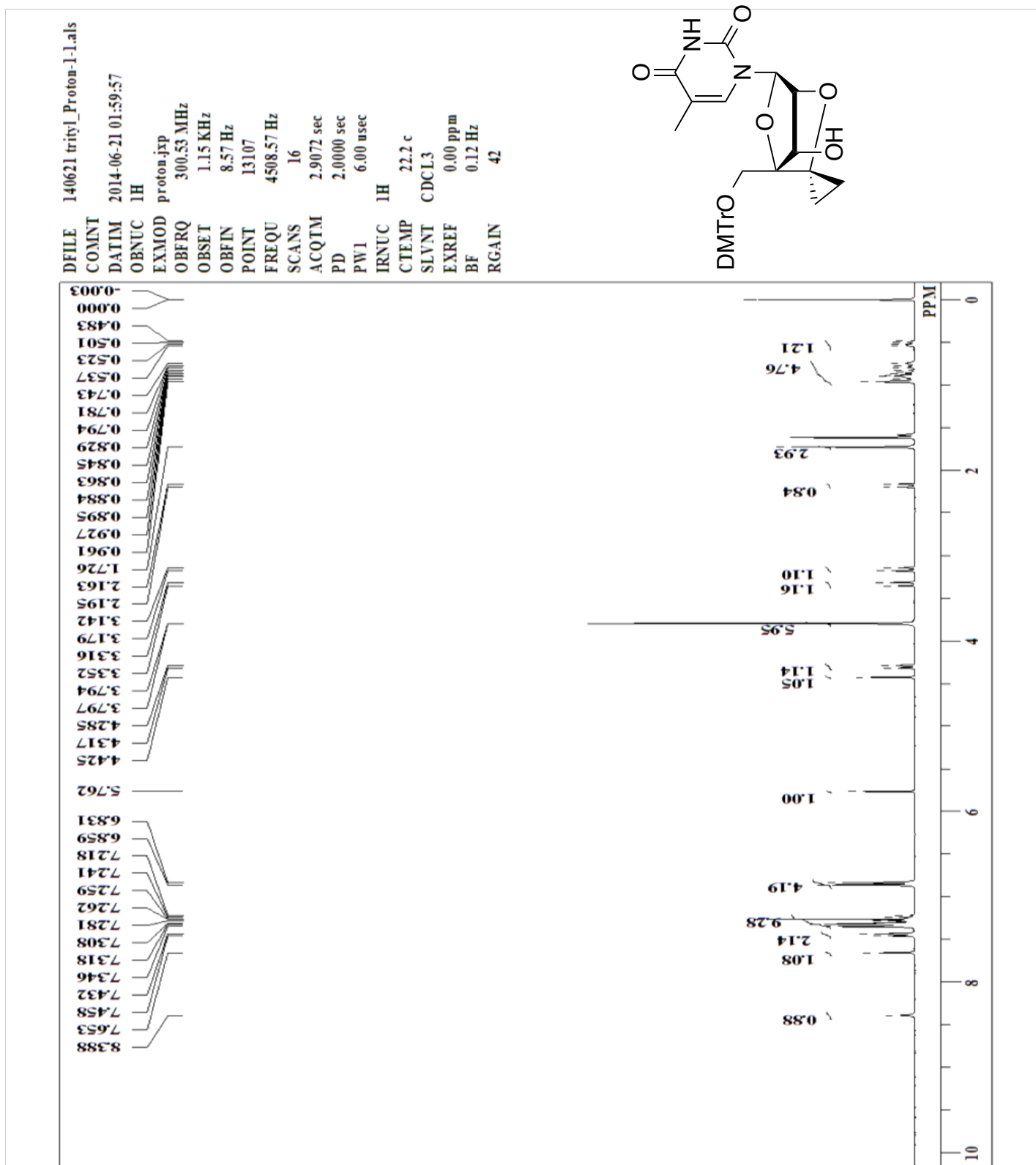


Figure S28. Compound 17 (<sup>13</sup>C NMR, CDCl<sub>3</sub>, 100 MHz)

DFILE 130926 trityl column pure  
 COMNT 130926 MH-25 column pu  
 DATIM 26-09-2013 19:36:23  
 OBNUC <sup>13</sup>C  
 EXMOD single\_pulse dec  
 OBFRQ 100.53 MHz  
 OBSET 5.35 KHz  
 OBFIN 5.86 Hz  
 POINT 26214  
 FREQU 25125.24 Hz  
 SCANS 200  
 ACQTM 1.0433 sec  
 PD 2.0000 sec  
 PW1 3.17 usec  
 IRNUC <sup>1</sup>H  
 CTEMP 22.5 c  
 SLVNT CDCl<sub>3</sub>  
 EXREF 77.16 ppm  
 BF 0.12 Hz  
 RGAIN 60

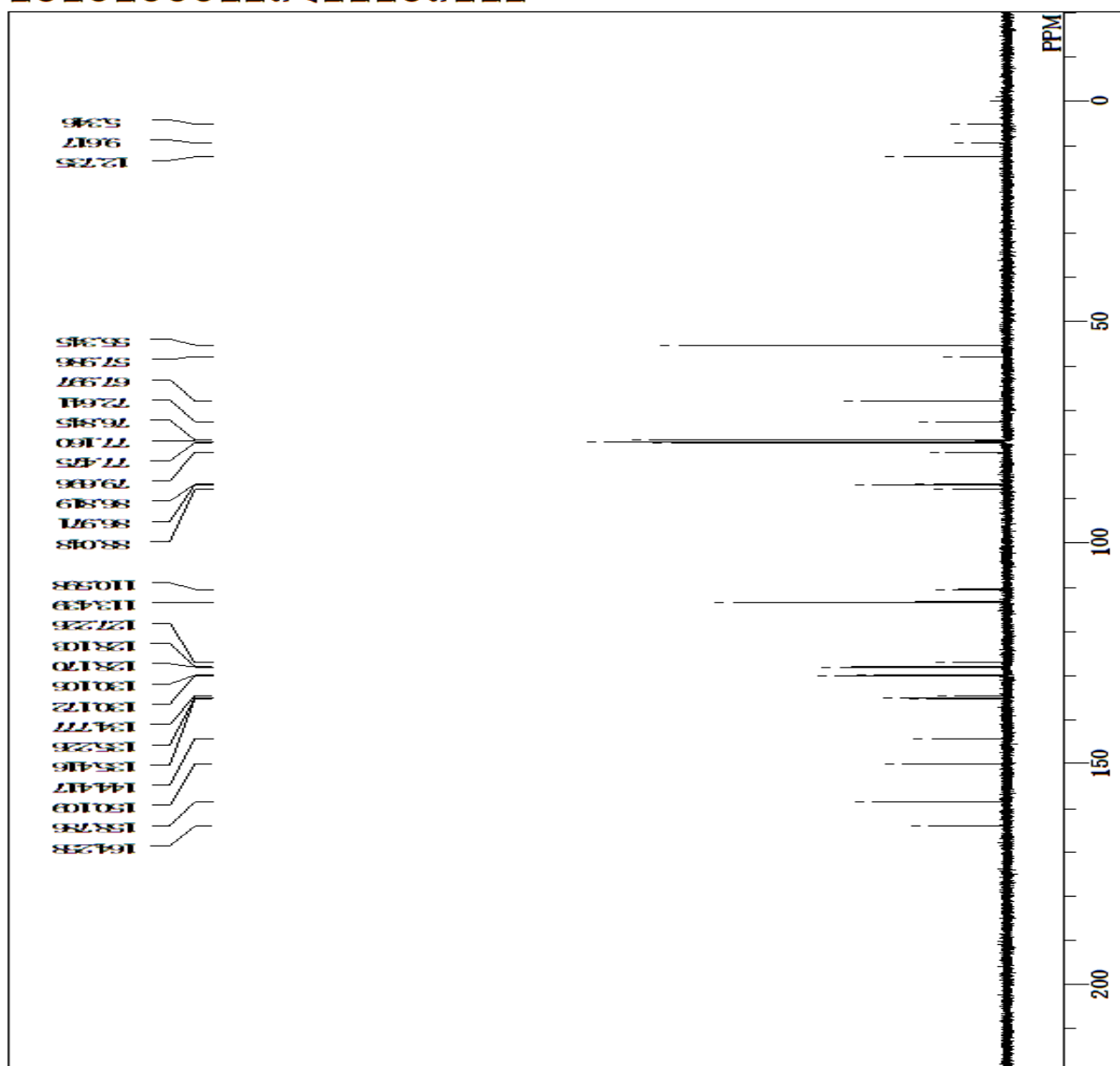
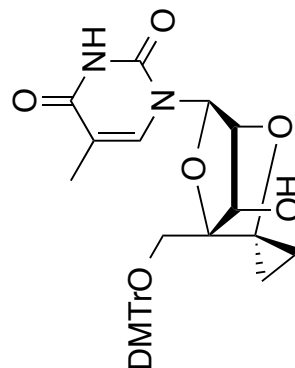


Figure S29. Compound 18 (<sup>1</sup>H NMR, CDCl<sub>3</sub>, 300 MHz)

DFILE 140707 amidite\_Proton-1-  
 COMNT single pulse  
 DATIM 2014-07-07 17:02:25  
 OBNUC 1H  
 EXMOD proton.jpg  
 OBFREQ 300.53 MHz  
 OBSET 1.15 KHz  
 OBFIN 8.57 Hz  
 POINT 16384  
 FREQU 5635.71 Hz  
 SCANS 8  
 ACQTM 2.9072 sec  
 PD 2.0000 sec  
 PW1 6.00 usec  
 IRNUC 1H  
 CTEMP 24.2 c  
 SLVNT CDCL3  
 EXREF 0.00 ppm  
 BF 1.20 Hz  
 RGAIN 40

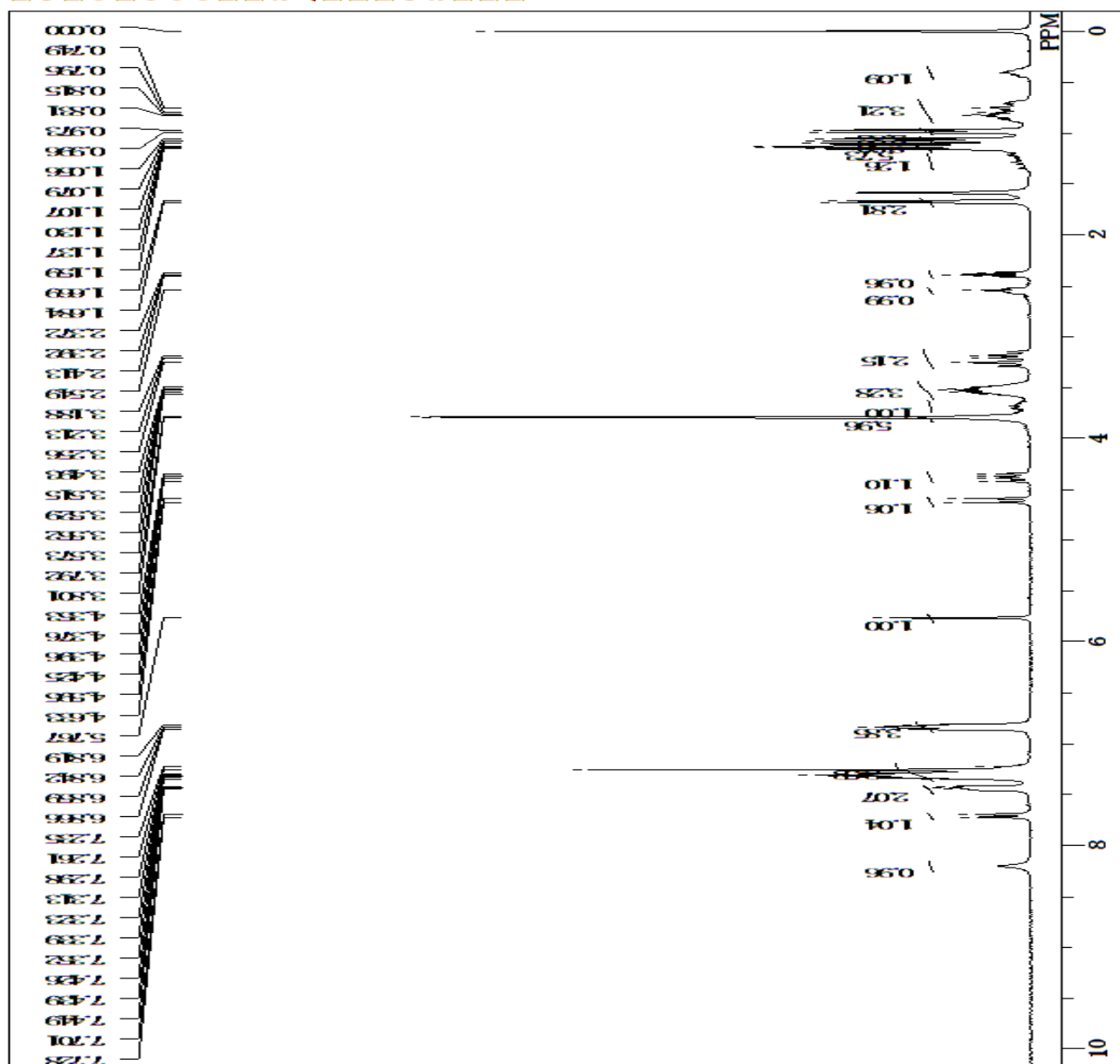
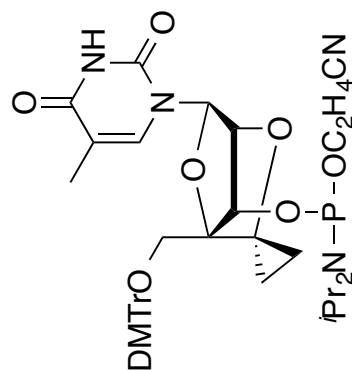
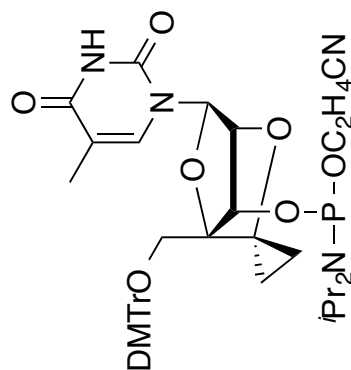
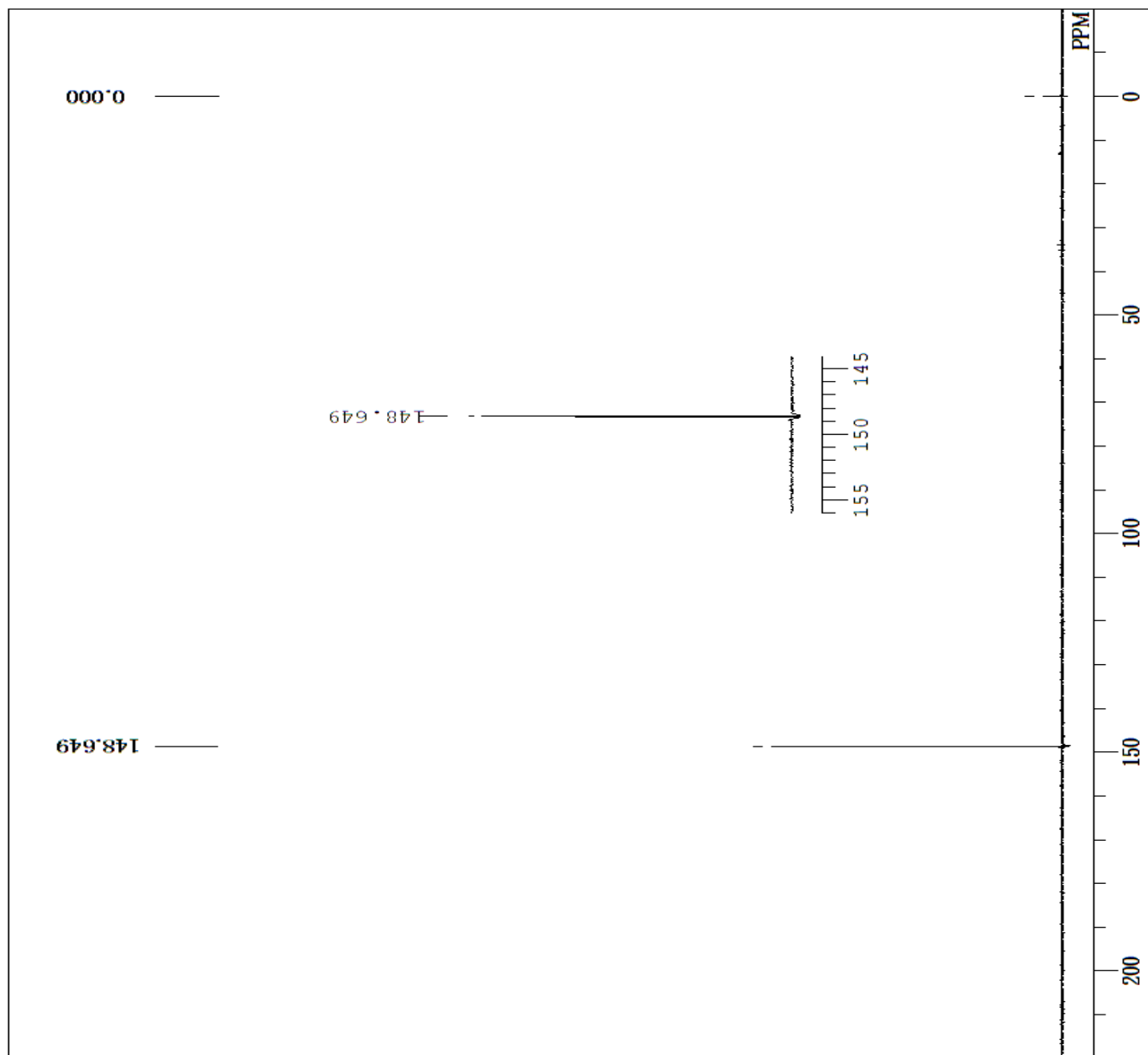




Figure S30. Compound 18 ( $^{31}\text{P}$  NMR,  $\text{CDCl}_3$ , 162 MHz)

131003 amidite phosphor  
DFILE  
COMINT  
DATIM 03-10-2013 17:19:55  
OBNUC 31P  
EXMOD single\_pulse\_dec  
OBFRQ 161.83 MHz  
OBSET 4.69 KHz  
OBFIN 3.09 Hz  
POINT 26214  
FREQU 114283.97 Hz  
SCANS 53  
ACQTM 0.2294 sec  
PD 2.0000 sec  
PW1 4.17 usec  
IRNUC 1H  
CTEMP 22.2 c  
SLVNT  $\text{CDCl}_3$   
EXREF 0.00 ppm  
BF 0.12 Hz  
RGAIN 56



### 3. Synthesis, purification and characterization of oligonucleotides

Synthesis of oligonucleotides modified with scpBNA was performed on an Applied Biosystems Expedite™ 8909 Nucleic Acid Synthesis System on 0.2 μmol scale of **ON2–ON6** and 1.0 μmol scale of **ON14** according to the standard phosphoramidite protocol and 5-[3,5-bis(trifluoromethyl)-phenyl]-1*H*-tetrazole as the activator. In the case of 0.2 μmol scale, the coupling time of phosphoramidite **16** was prolonged from 32 seconds to 8 minutes. In the case of 1.0 μmol scale, that was prolonged from 40 seconds to 10 minutes. The synthesis was carried out in trityl on mode and the solid supported oligonucleotides were treated with concentrated ammonium hydroxide at 55 °C for 12 h. The **ON2–ON6** and the **ON14** were briefly purified with Sep-Pak® Plus C<sub>18</sub> Cartridge and Sep-Pak® Plus C<sub>18</sub> Environmental Cartridge, respectively. The **ON2–ON6** and the **ON14** were further purified by reverse-phase HPLC with Waters XTerra MS C<sub>18</sub> 2.5 μm (10 x 50 mm) columns with a linear gradient of MeCN (6 to 12% over 30 min) in 0.1 M triethylammonium acetate buffer (pH 7.0). The purity of the oligonucleotides were analyzed by reverse-phase HPLC with Waters XTerra MS C<sub>18</sub> 2.5 μm (4.6 x 50 mm) columns and characterized by MALDI-TOF mass spectrometer.

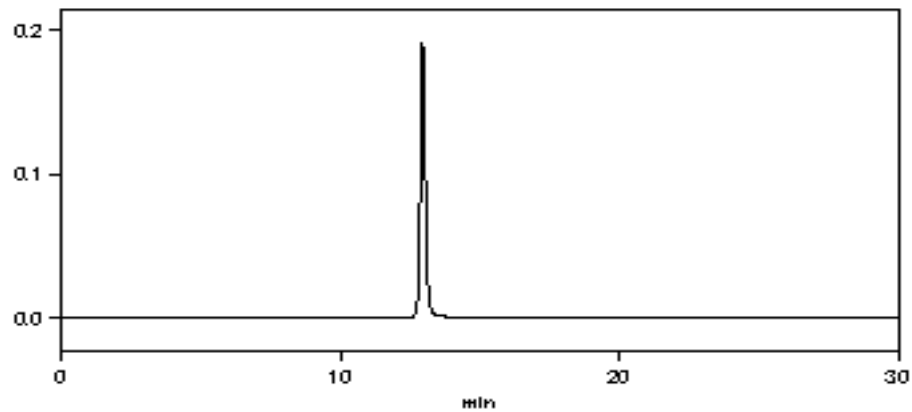
**Table S1.** Yields and MALDI-TOF MS data for the oligonucleotides.

oligonucleotides <sup>a</sup>	Yield (%)	MALDI-TOF MS		
		calcd [M-H] <sup>-</sup>	found [M-H] <sup>-</sup>	
5'-d(GCGTT <b>X</b> TTTGCT)-3'	<b>ON2</b>	26	3686.4	3686.8
5'-d(GCGTT <b>XX</b> TTTGCT)-3'	<b>ON3</b>	12	3740.5	3741.0
5'-d(GCGTT <b>XXX</b> TTTGCT)-3'	<b>ON4</b>	8	3794.5	3794.6
5'-d(GCGTT <b>XTXT</b> GCT)-3'	<b>ON5</b>	36	3740.5	3740.3
5'-d(GCG <b>XTXTXT</b> GCT)-3'	<b>ON6</b>	41	3794.5	3794.6
5'-d(TTTTTTTT <b>X</b> )-3'	<b>ON14</b>	23	2728.8	2728.5

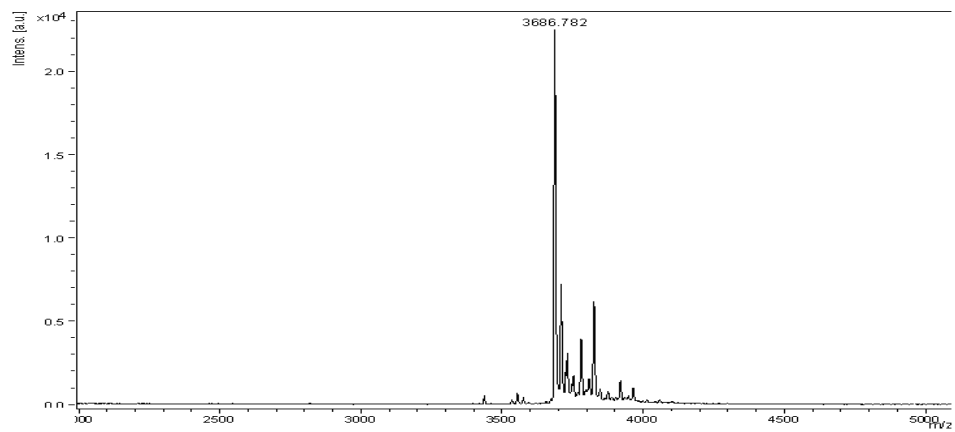
[a] **X** = scpBNA

**Figure S31.** MALDI-TOF MS spectra and HPLC charts of all new oligonucleotides.

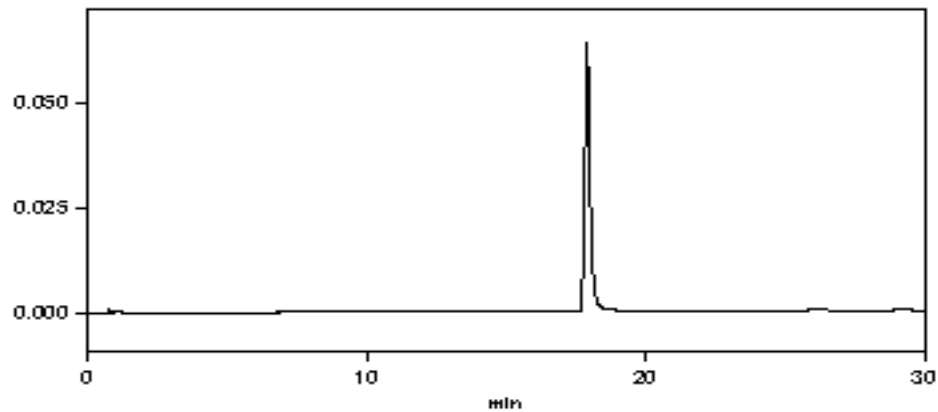
**HPLC (ON2)**



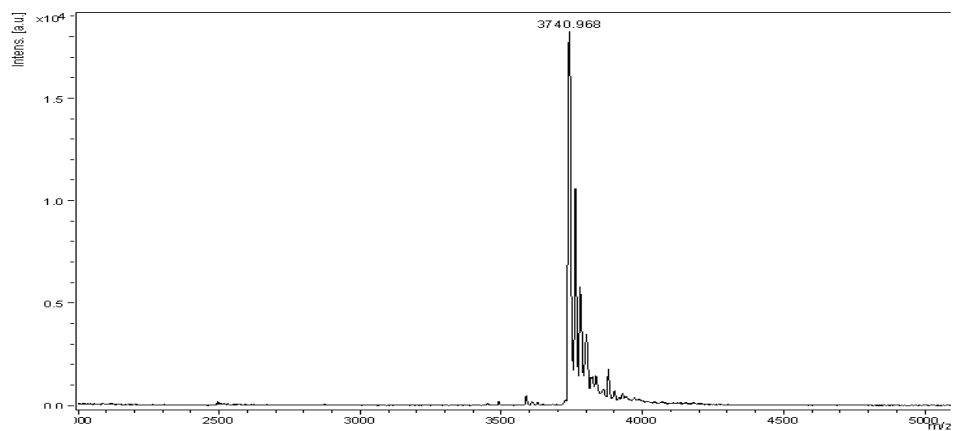
**MALDI-TOF MS (ON2)**



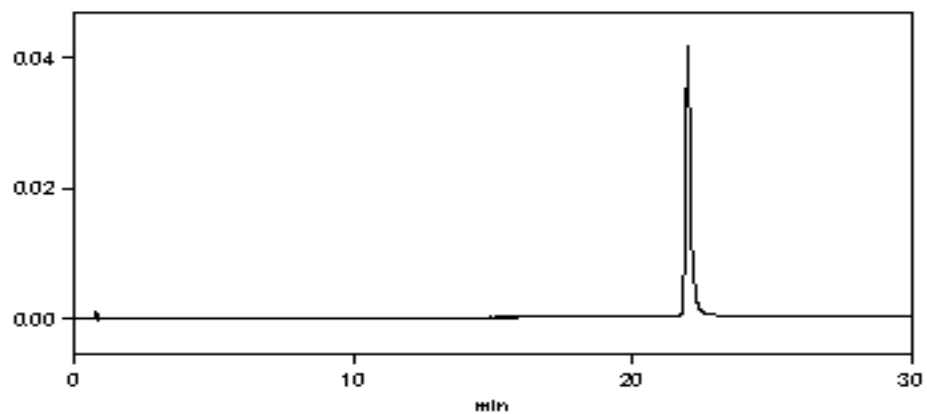
**HPLC (ON3)**



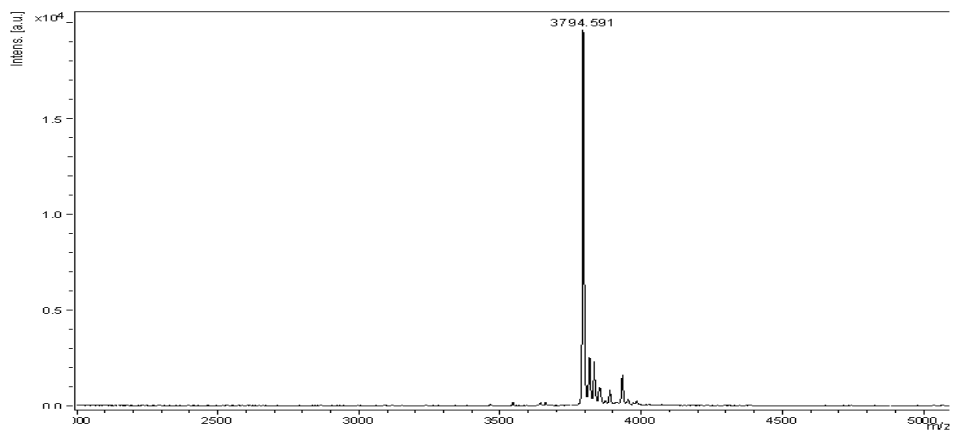
### MALDI-TOF MS (ON3)



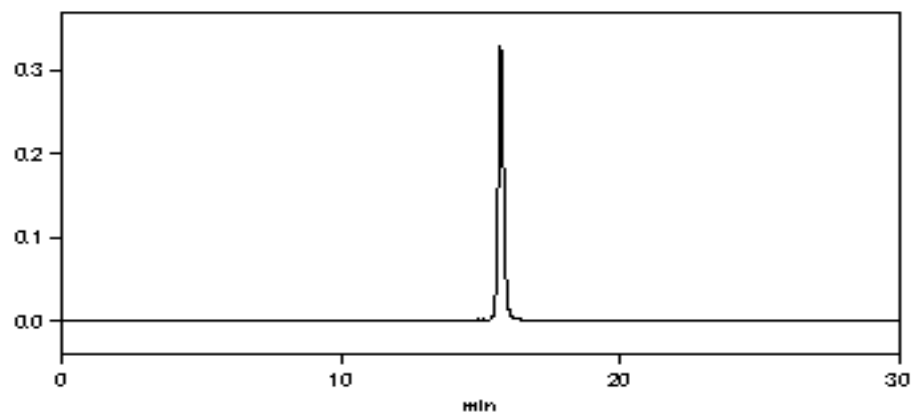
### HPLC (ON4)



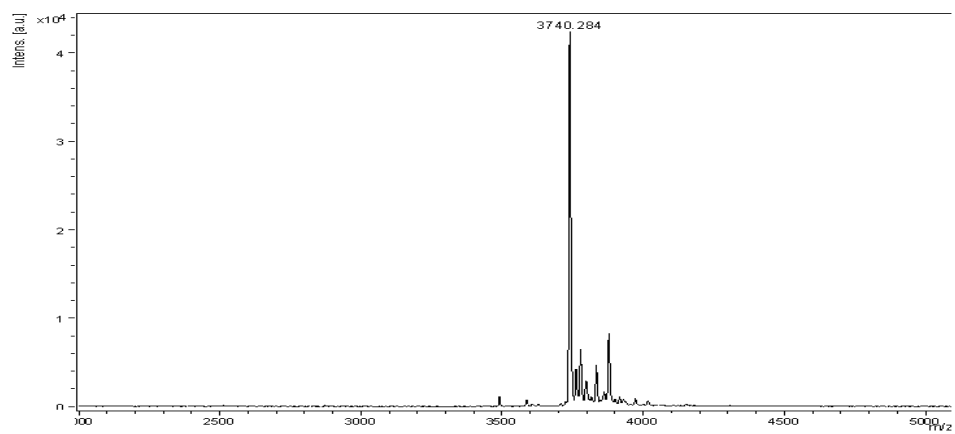
### MALDI-TOF MS (ON4)



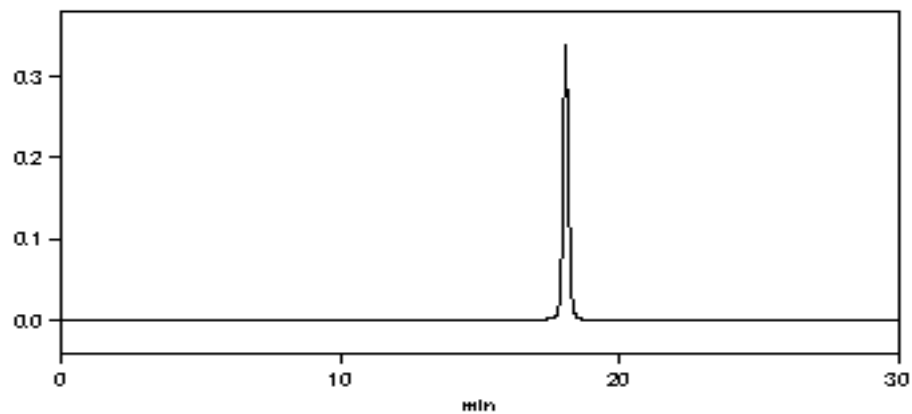
### HPLC (ON5)



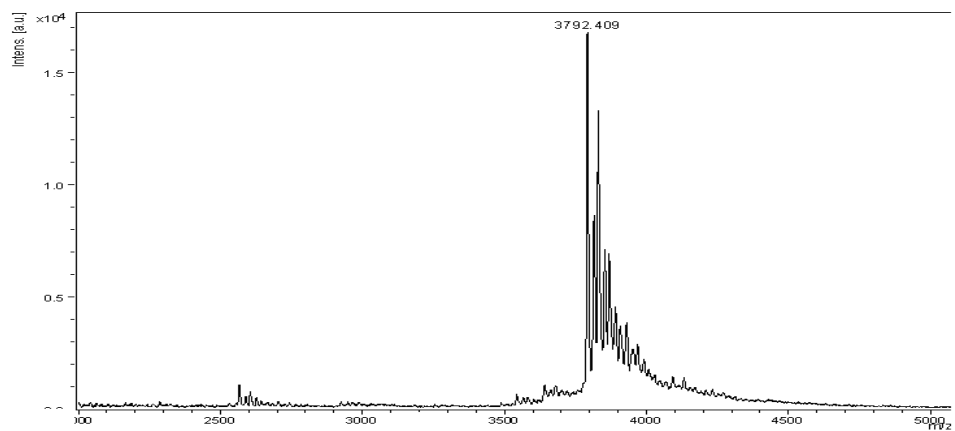
### MALDI-TOF MS (ON5)



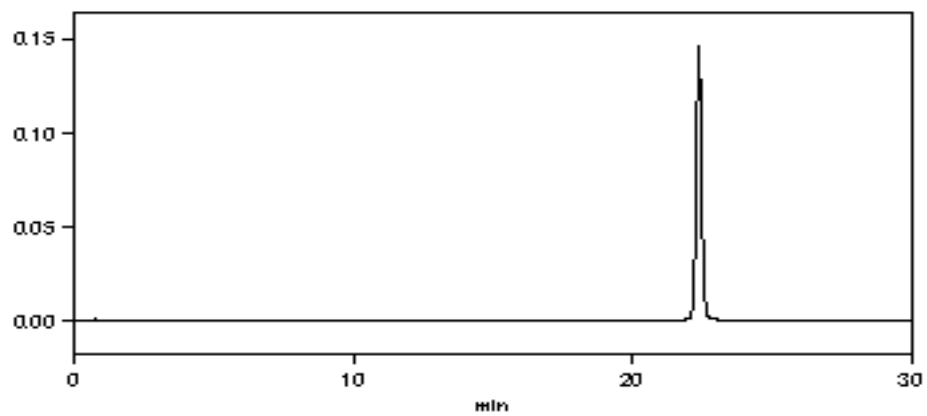
### HPLC (ON6)



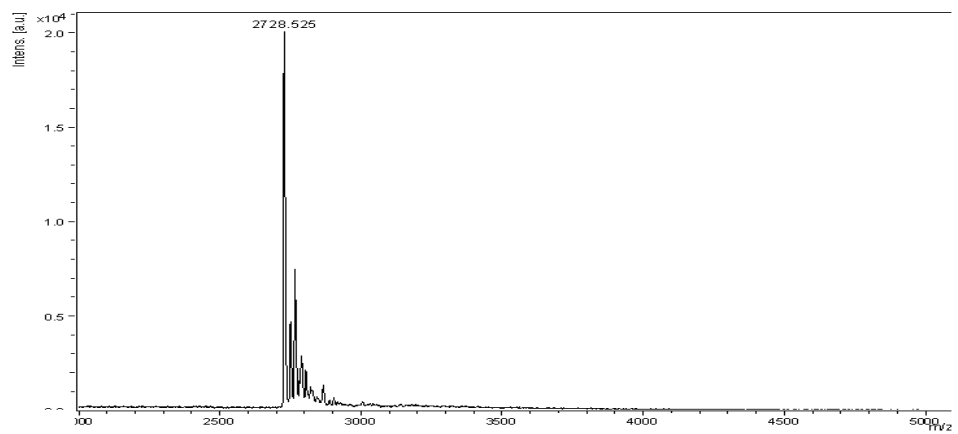
### MALDI-TOF MS (ON6)



### HPLC (ON14)

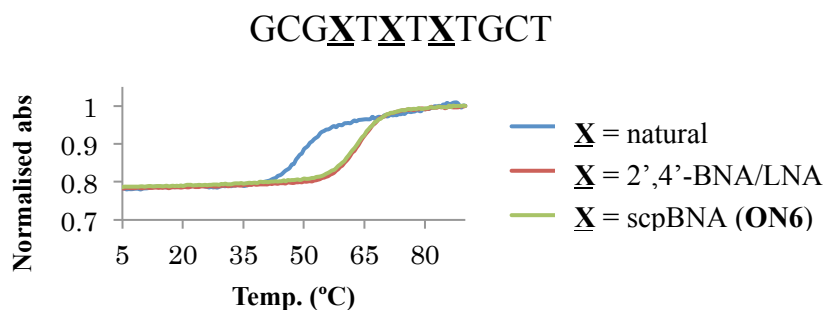


### MALDI-TOF MS (ON14)

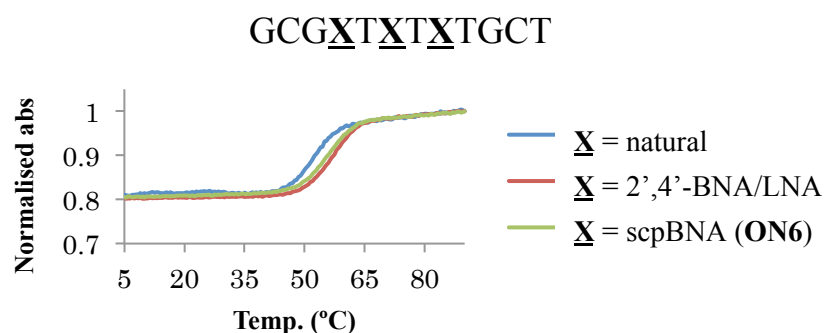


#### 4. UV melting experiments

The UV melting experiments were carried out on SHIMADZU UV-1650PC and SHIMADZU UV-1800 spectrometers equipped with a  $T_m$  analysis accessory. Equimolecular amounts of the target RNA or DNA strand and oligonucleotide were dissolved in buffer (10 mM phosphate buffer at pH 7.2 containing 100 mM NaCl) to give final strand concentration of 4  $\mu$ M. The samples were annealed by heating at 100  $^{\circ}$ C followed by slow cooling to room temperature. The melting profile was recorded at 260 nm from 5 to 90  $^{\circ}$ C at a scan rate of 0.5  $^{\circ}$ C /min. The  $T_m$  value was calculated as the temperature of the half-dissociation of the formed duplexes based on the first derivative of the melting curve.



**Figure S32.** UV melting curves for the duplexes formed between representative oligonucleotides and ssRNA. The sequences of oligonucleotides and ssRNA are 5'-d(GCGXTXTXTGCT)-3' and 5'-d(AGCAAAAACGC)-3', respectively.

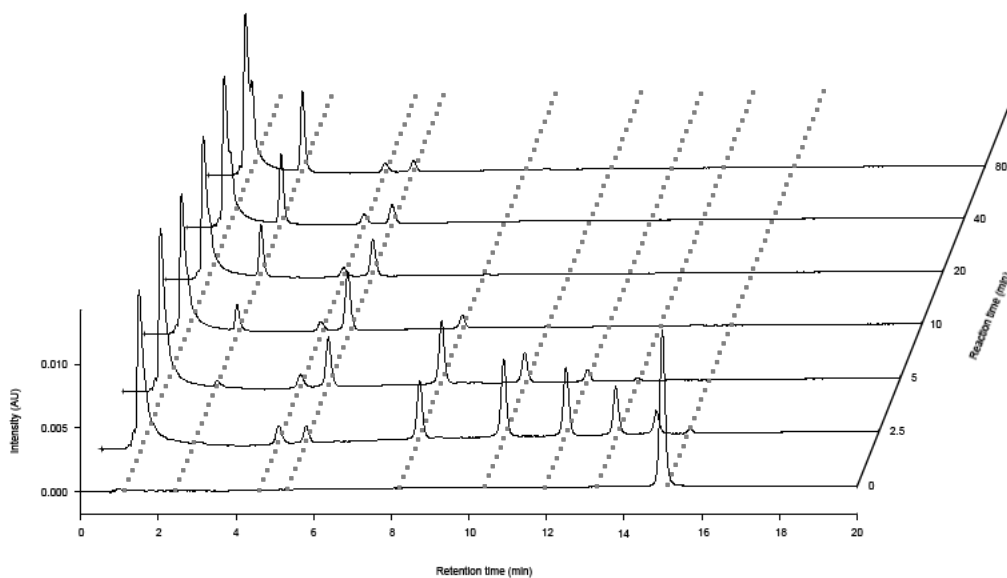


**Figure S33.** UV melting curves for the duplexes formed between representative oligonucleotides and ssDNA. The sequences of oligonucleotides and ssDNA are 5'-d(GCGXTXTXTGCT)-3' and

5'-r(AGCAAAAACGC)-3', respectively.

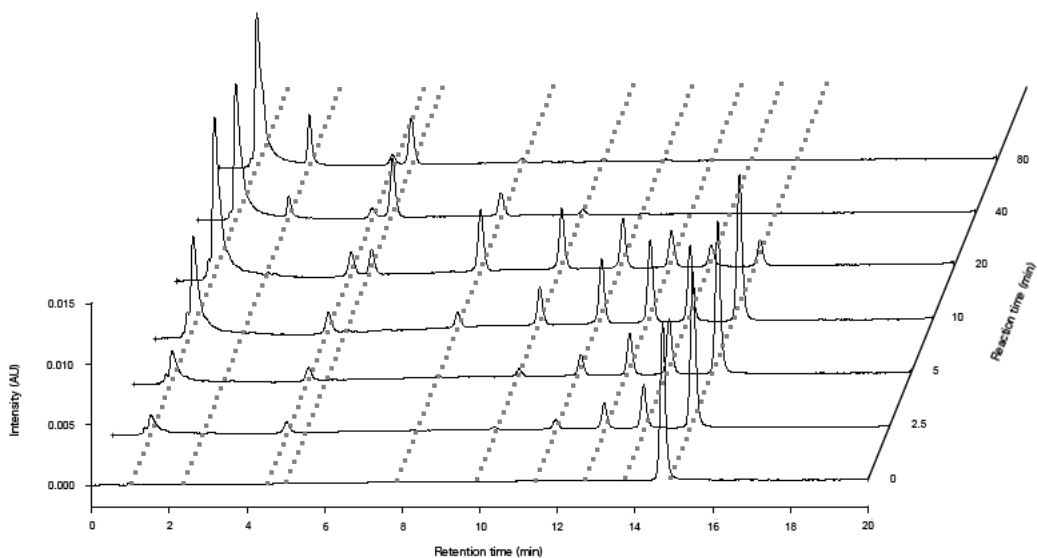
## 5. Nuclease resistance study

The sample solutions were prepared by dissolving 0.56 nmol of oligonucleotides in 50 mM Tris-HCl buffer (pH 8.0) containing 10 mM MgCl<sub>2</sub>. In each sample solutions, 0.14 µg CAVP was added and the cleavage reaction was carried out at 37 °C. A portion of each reaction mixture was removed at time intervals and heated at 90 °C for 2.5 min to deactivate the nuclease. Aliquots of the timed samples were analyzed by reverse-phase HPLC with Waters XBridge™ OST C<sub>18</sub> 2.5 µg (4.6 x 50 mm) columns to evaluate the amount of intact oligonucleotides remaining. The percentage of intact oligonucleotide in each sample was calculated and plotted against the digestion time to obtain a degradation curve.

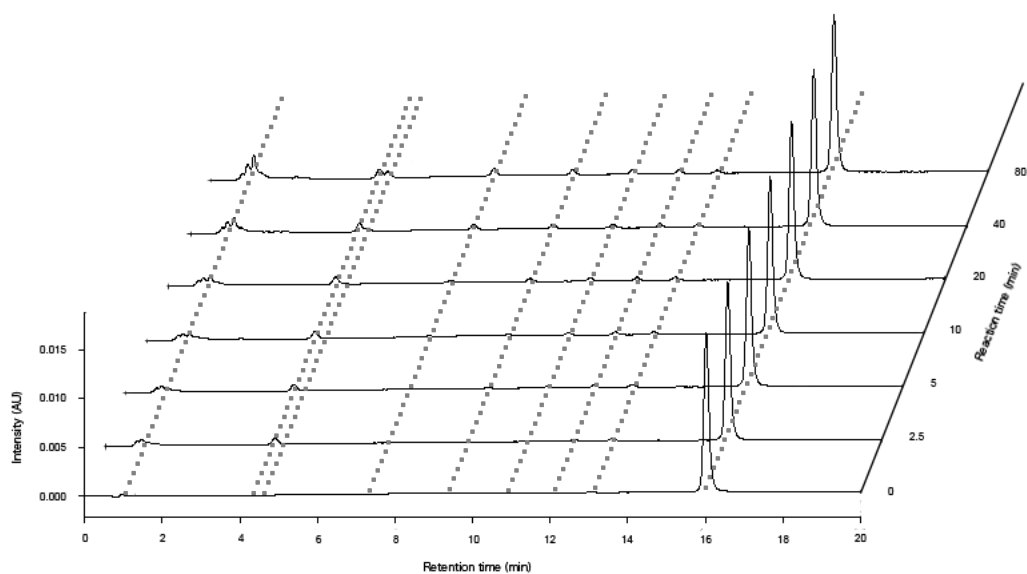


Natural (ON12)



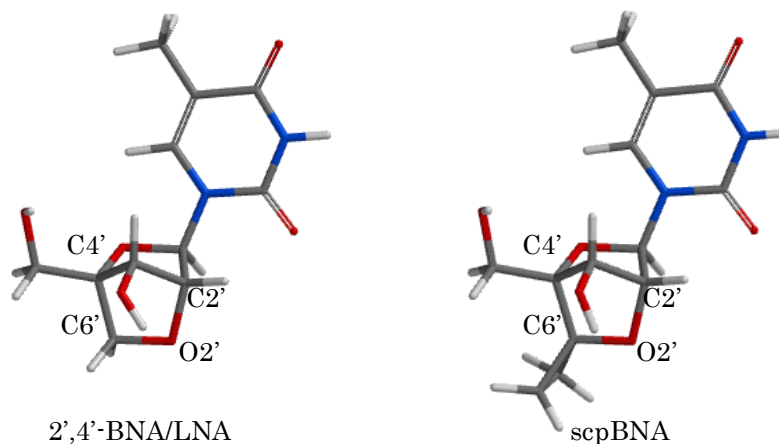


2',4'-BNA/LNA (ON13)



scpBNA (ON14)

**Figure S34.** Nuclease resistance of 5'-(TTTTTTT~~X~~)-3' against *Crotalus adamanteus* venom phosphodiesterase (CAVP). Column: Waters XBridge™ OST C<sub>18</sub> 2.5 μg (4.6 x 50 mm). Mobile Phase: Linear gradient of MeCN (6 to 12 % over 20 min) in 0.1 M triethylammonium acetate (pH 7.0). Flow rate: 1.0 mL/min. Detection: Absorbance at 260 nm.



	2',4'-BNA/LNA	scpBNA
$\nu_0$ :	$-0.71^\circ$	$-1.74^\circ$
$\nu_1$ :	$-35.74^\circ$	$-34.78^\circ$
$\nu_2$ :	$54.45^\circ$	$53.92^\circ$
$\nu_3$ :	$-56.87^\circ$	$-57.17^\circ$
$\nu_4$ :	$37.30^\circ$	$37.72^\circ$
$\nu_{\max}$ :	$57.74^\circ$	$57.54^\circ$
$P$ :	$19.44^\circ$	$20.44^\circ$
O2'-C6'-C4':	$102.96^\circ$	$105.56^\circ$
C2'-O2'-C6':	$104.29^\circ$	$103.31^\circ$
$\delta$ :	$62.13^\circ$	$61.86^\circ$

**Figure S35.** Energy-minimized structures, endocyclic sugar torsion angles  $\nu_0$ - $\nu_4$ , maximum torsion angle  $\nu_{\max}$ , pseudorotation phase angle  $P$ , bond angles of O2'-C6'-C4' and C2'-O2'-C6' and phosphate backbone torsion angle  $\delta$  of 2',4'-BNA/LNA and scpBNA. Theoretical calculation was carried out using HF/6-31G\*\* basis set (Spartan '10, Wavefunction, Inc).  $P$  and  $\nu_{\max}$  are calculated as follows:  $\tan P = (\nu_4 + \nu_1 - \nu_3 - \nu_0) / (2 \cdot \nu_2 \cdot (\sin 36^\circ + \sin 72^\circ))$ ;  $\nu_{\max} = \nu_2 / \cos P$ .

## 7. Mismatch discrimination and thermodynamic data of scpBNA

**Table S2.**  $T_m$  values (°C) of duplexes formed between ONs and ssRNA with or without one base mismatch<sup>a</sup>

oligonucleotides	$T_m$ ( $\Delta T_m = T_m$ [mismatch] – $T_m$ [match]) (°C)			
	$\underline{Z} = A$	G	C	U
<b>ON1</b>	48 <sup>b</sup>	43 (–5) <sup>b</sup>	32 (–16) <sup>b</sup>	33 (–15) <sup>b</sup>
<b>ON2</b>	53	47 (–6)	35 (–18)	38 (–15)
<b>ON7</b>	52	47 (–5)	36 (–16)	39 (–13)

[a] Conditions: 10 mM phosphate buffer (pH 7.2), 100 mM NaCl, and 4  $\mu$ M each oligonucleotide. The  $T_m$  values reflect the average of at least three measurements. The sequences of oligonucleotides are 5'-d(GCGTTXTTTGCT)-3' (X = natural thymidine (**ON1**), scpBNA-T (**ON2**), and 2',4'-BNA/LNA-T (**ON7**)). The sequence of ssRNA is 5'-r(AGCAA $\underline{Z}$ AACGC)-3'. [b] Reference S1.

**Table S3.**  $T_m$  values (°C) of duplexes formed between ONs and ssDNA with or without one base mismatch<sup>a</sup>

oligonucleotides	$T_m$ ( $\Delta T_m = T_m$ [mismatch] – $T_m$ [match]) (°C)			
	$\underline{Z} = A$	G	C	T
<b>ON1</b>	52 <sup>b</sup>	41 (–11) <sup>b</sup>	37 (–15) <sup>b</sup>	38 (–14) <sup>b</sup>
<b>ON2</b>	53	43 (–10)	38 (–15)	40 (–13)
<b>ON7</b>	53	42 (–11)	38 (–15)	41 (–12)

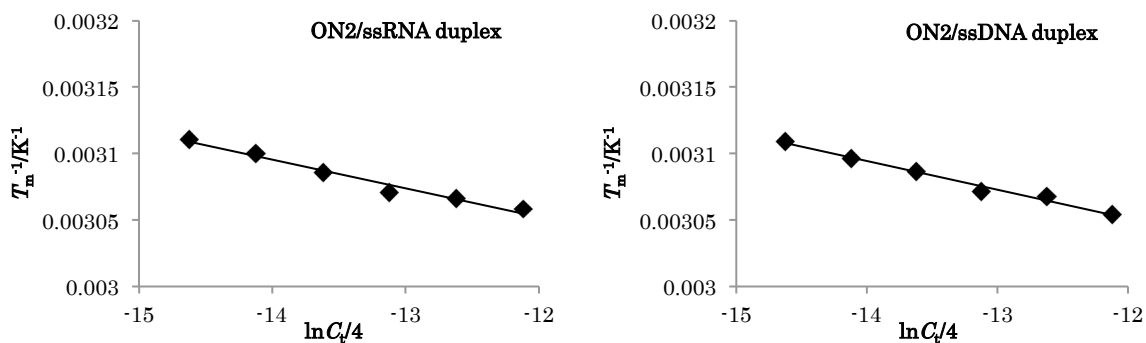
[a] Conditions: 10 mM phosphate buffer (pH 7.2), 100 mM NaCl, and 4  $\mu$ M each oligonucleotide. The  $T_m$  values reflect the average of at least three measurements. The sequences of oligonucleotides are 5'-d(GCGTTXTTTGCT)-3' (X = natural thymidine (**ON1**), scpBNA-T (**ON2**), and 2',4'-BNA/LNA-T (**ON7**)). The sequence of ssDNA is 5'-d(AGCAA $\underline{Z}$ AACGC)-3'. [b] Reference S1.

**Table S4.** Thermodynamic data of duplexes formed between ONs and ssRNA or ssDNA<sup>a,b</sup>

duplexes	$\Delta H^\circ$ (kcal mol <sup>-1</sup> )	$\Delta S^\circ$ (cal K <sup>-1</sup> mol <sup>-1</sup> )	$\Delta G^\circ_{310K}$ (kcal mol <sup>-1</sup> )
<b>ON1/ssRNA</b>	–98.4 <sup>c</sup>	–282 <sup>c</sup>	–10.9 <sup>c</sup>
<b>ON2/ssRNA</b>	–92.0	–257	–12.3
<b>ON7/ssRNA</b>	–87.6 <sup>c</sup>	–244 <sup>c</sup>	–12.0 <sup>c</sup>
<b>ON1/ssDNA</b>	–84.6 <sup>c</sup>	–235 <sup>c</sup>	–11.6 <sup>c</sup>
<b>ON2/ssDNA</b>	–92.0	–257	–12.4
<b>ON7/ssDNA</b>	–83.2 <sup>c</sup>	–230 <sup>c</sup>	–11.9 <sup>c</sup>

[a] Conditions: 10 mM phosphate buffer (pH 7.2), 100 mM NaCl, and 0.89–10.9  $\mu$ M each oligonucleotide (six data points). The  $T_m$  values reflect the average of at least three measurements. The sequences of oligonucleotides are 5'-d(GCGTTXTTTGCT)-3' (X = natural thymidine (**ON1**), scpBNA-T (**ON2**), and 2',4'-BNA/LNA-T (**ON7**)). The sequences of ssRNA and ssDNA are 5'-r(AGCA $\underline{Z}$ AAACGC)-3' and 5'-d(AGCA $\underline{Z}$ AAACGC)-3', respectively. [b] These values are calculated by Van't Hoff plots with six data points. See also Figure S36. [c] Reference S2.

**Figure S36.** Van't Hoff plots of  $T_m$  values of duplexes formed between ON2 and ssRNA or ssDNA.



### 8. Supplementary references

- S1 Y. Mitsuoka, T. Kodama, R. Ohnishi, Y. Hari, T. Imanishi, S. Obika, *Nucleic Acids Res.*, 2009, **37**, 1225–1238.
- S2 K. Morihiro, T. Kodama, Kentefu, Y. Moai, R. N. Veedu, S. Obika, *Angew. Chem. Int. Ed.* 2013, **52**, 5074–5078.