

## Supplementary Information

### Gold Nanoparticle-Based Multivalent Carbohydrate Probes: Selective Photoaffinity Labeling of Carbohydrate-Binding Proteins

Kaori Sakurai,\* Yuki Hatai and Ayumi Okada

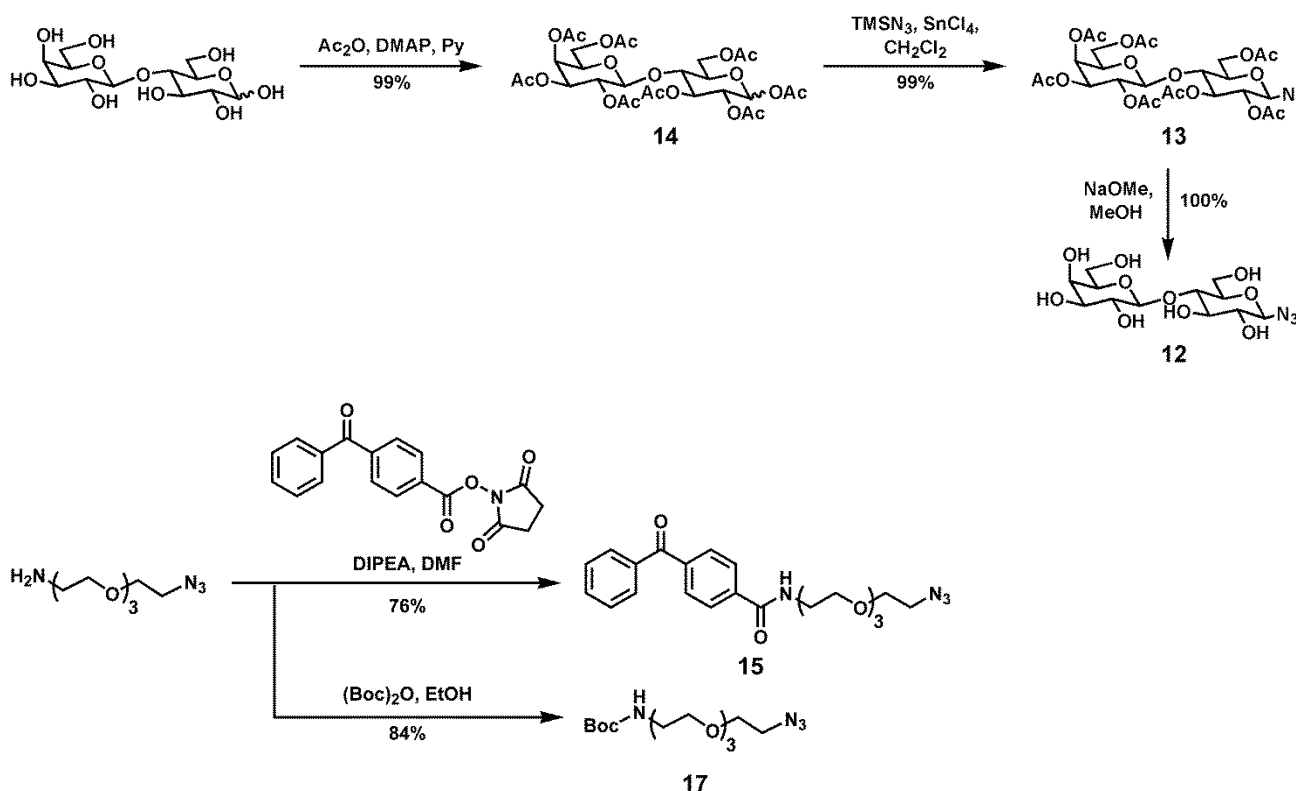
Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Tokyo 184-8588, Japan

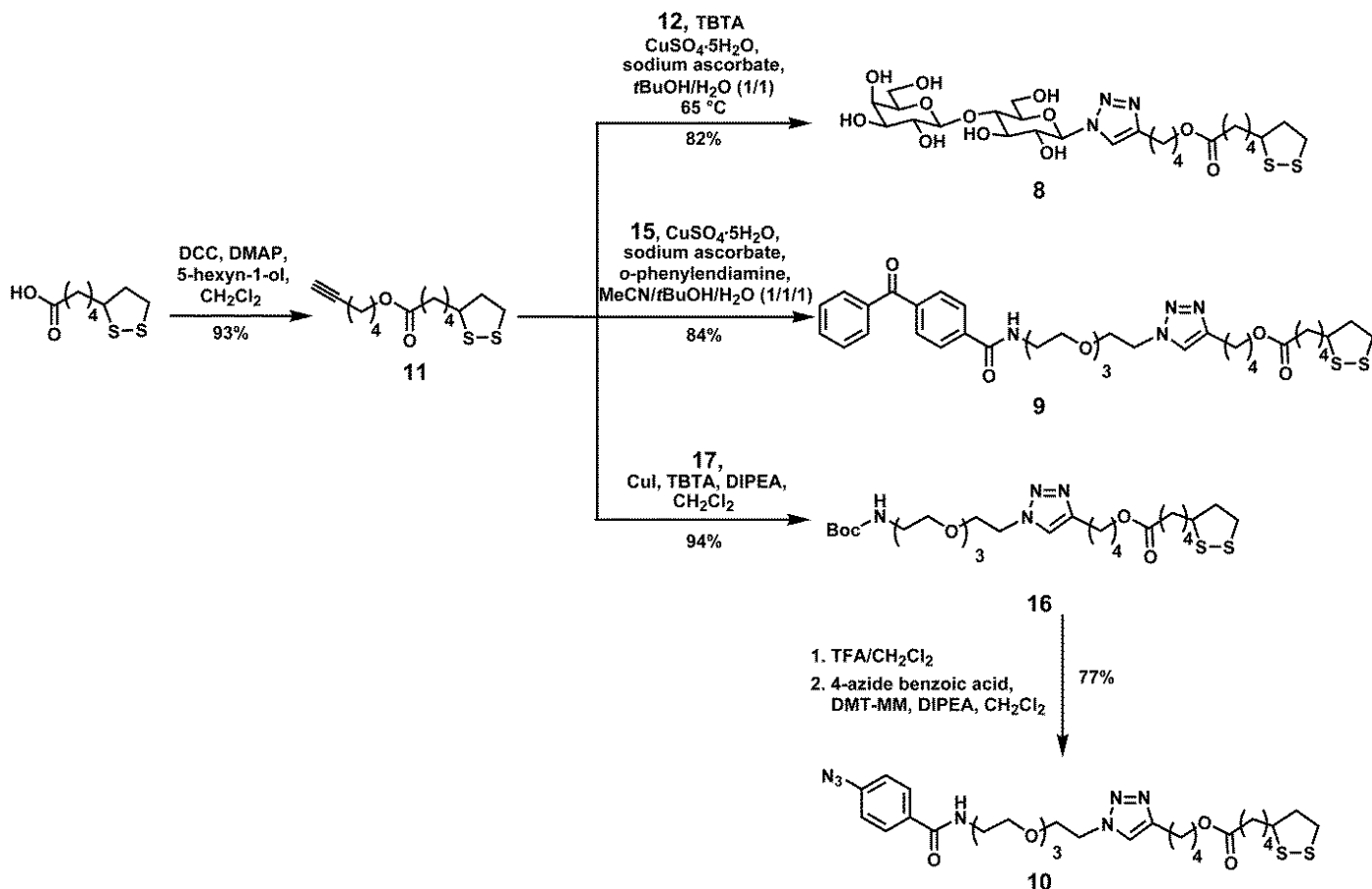
1. General experimental methods .....	2
2. Synthesis of functional groups conjugated to lipoic acid ( <b>8–17</b> ). .....	2
3. Synthesis and characterization of photoaffinity probe <b>1–7</b> .....	7
3-1. Assembly of photoaffinity probe <b>1–7</b> .....	7
3-2. SDS-PAGE analysis .....	7
3-3. UV-VIS spectrometric analysis .....	8
3-4. MALDI-TOF-MS analysis .....	8
3-5. LC-MS analysis of probe <b>5</b> to determine the ligand density for <b>8</b> and <b>10</b> .....	8
4. Affinity pull-down assay.....	9
5. Photoaffinity labeling experiments. ....	11
5.1. PAL reactions with carbohydrate-binding proteins.....	11
5.2. PAL reactions with PNA in HeLa lysate.....	12
5.3. MALDI-TOF-MS analysis of probe crosslinked PNA obtained by PAL reactions. ....	12
6. NMR spectra .....	14

## 1. General experimental methods

$^1\text{H}$  and  $^{13}\text{C}$  NMR spectra were recorded at 293 K on JEOL ECX 300 or 400 using 5 mm z-gradient probes and the data were processed with Delta software. The chemical shifts are denoted in  $\delta$  (ppm) relative to residual solvent peaks as internal references ( $\text{CDCl}_3$ ,  $^1\text{H}$  NMR, 7.26 ppm,  $^{13}\text{C}$  NMR, 77.0 ppm;  $\text{DMSO}-d_6$ ,  $^1\text{H}$  NMR, 2.50 ppm,  $^{13}\text{C}$  NMR, 39.5 ppm; pyridine- $d_5$ ,  $^1\text{H}$  NMR, 8.71 ppm,  $^{13}\text{C}$  NMR, 150 ppm). Positive ion ESI-TOF-MS data were obtained by JEOL AccuTOF mass spectrometer. MALDI-TOF-mass spectra were acquired by AB SCIEX TOF/TOF<sup>TM</sup> 5800 (AB SCIEX). Unless noted otherwise, all chemical reagents were purchased from Wako Chemicals, TCI and Sigma-Aldrich. Peanut agglutinin (PNA) and *Erythrina cristagalli* lectin (ECA) were purchased from Sigma-Aldrich. *Ricinus communis* agglutinin RCA was purchased from Vector Laboratories, INC. Gold nanoparticle was purchased from BBI solutions. Flash column chromatography was performed using Silica gel 60 (spherical, particle size 40-100  $\mu\text{m}$ ; Kanto). Fluorescence imaging was scanned with Typhoon 8600 (GE Healthcare Science). A UV-visible spectrum was recorded on a NanoDrop ND-1000 spectrophotometer and VARIAN CARY 50 conc UV-visible spectrophotometer. LC-MS experiments were performed on Hitachi HPLC 2100 equipped with microTOF-QII (BRUKER).

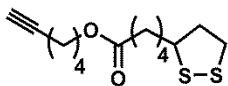
## 2. Synthesis of functional groups conjugated to lipoic acid (8–17).





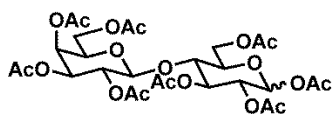
**Scheme S1.** Synthesis of lipoic acid conjugated Lac **8** and the photoreactive group (**9** and **10**).

#### Compound **11**



To a solution of  $\alpha$ -lipoic acid (0.50 g, 2.4 mmol) in dry  $\text{CH}_2\text{Cl}_2$  (22 mL) was added 5-hexyn-1-ol (24  $\mu\text{L}$ , 2.4 mmol), dicyclohexylcarbodiimide ( $9.1 \times 10^2$  mg, 4.4 mmol) and *N*-(4-pyridyl) dimethylamine (19 mg, 0.16 mmol) at  $0^\circ\text{C}$ . The mixture was stirred for 3 hours at room temperature. The reaction mixture was diluted with MilliQ water, extracted twice by EtOAc, and washed with brine, dried over  $\text{Na}_2\text{SO}_4$ , filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography (hexane/EtOAc = 1/19 to 1/9) to give compound **11** (0.59 g, 2.0 mmol, 93%).  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  4.05 (t,  $J = 5.9$  Hz, 2H), 3.57-3.48 (m, 1H), 3.18-3.02 (m, 2H), 2.47-2.37 (m, 1H), 2.27 (dd,  $J = 7.2, 6.9$  Hz, 2H), 2.19 (m, 2H), 1.93 (m, 1H), 1.90-1.81 (m, 1H), 1.76-1.50 (m, 8H), 1.48-1.34 (m, 2H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{CDCl}_3$ ):  $\delta$  173.3, 83.7, 68.6, 63.6, 56.1, 40.0, 38.3, 34.4, 33.9, 28.6, 27.5, 24.8, 24.5, 17.9; HRMS (ESI-TOF) calculated for  $\text{C}_{14}\text{H}_{22}\text{NaO}_2\text{S}_2$  ( $\text{M}+\text{Na}$ ) $^+$ : 309.0959; found: 309.0955.

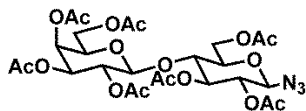
#### Compound **14**



To a solution of D-(+)-lactose (0.50 g, 1.5 mmol) in pyridine (1.9 mL, 23 mmol) was added acetic anhydride (2.2 mL, 23 mmol) and *N*-(4-pyridyl) dimethylamine (3.6 mg, 0.029 mmol) under argon atmosphere at  $0^\circ\text{C}$ . The reaction mixture was stirred for 30 min then was allowed to warm to room temperature. After 67 hours, the reaction mixture was quenched by addition of 2 M HCl (10 mL). The mixture was extracted with  $\text{CH}_2\text{Cl}_2$ , which was washed with saturated aqueous

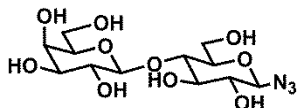
NaHCO<sub>3</sub>, brine then dried over Na<sub>2</sub>SO<sub>4</sub>, and filtered. The residue was concentrated *in vacuo* to give compound **14** (0.98 g, 1.5 mmol, 99%). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ 6.25 (d, *J* = 3.8 Hz, 1H), 5.46 (t, *J* = 9.6 Hz, 1H), 5.36 (d, *J* = 3.1 Hz, 1H), 5.15-5.09 (m, 1H), 5.03-4.94 (m, 2H), 4.49-4.43 (m, 2H), 4.19-3.99 (m, 4H), 3.90-3.78 (m, 2H), 2.18 (s, 3H), 2.16 (s, 3H), 2.12 (s, 3H), 2.06 (s, 9H), 2.01 (s, 3H), 1.97 (s, 3H); <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>): δ 170.3, 170.2, 170.1, 170.0, 169.9, 169.6, 169.1, 168.9, 129.0, 128.2, 125.2, 101.2, 88.9, 75.7, 70.9, 70.6, 70.4, 69.5, 69.3, 69.0, 66.5, 61.4, 60.7, 21.4, 20.9, 20.8, 20.6, 20.5; HRMS (ESI-TOF) calculated for C<sub>28</sub>H<sub>38</sub>NaO<sub>19</sub> (M+Na)<sup>+</sup>: 701.1905; found 701.1905.

#### Compound 13



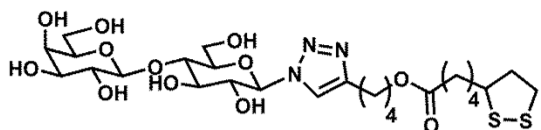
To a solution of compound **14** (0.21 g, 0.31 mmol) and trimethylsilyl azide (0.057 mL, 0.43 mmol) in dry CH<sub>2</sub>Cl<sub>2</sub> (0.62 mL) was added tin (IV) chloride (0.093 mL, 0.093 mmol) under argon atmosphere. The reaction mixture was stirred at room temperature for 36 hours, then the residue was concentrated *in vacuo* and purified by flash column chromatography (hexane/EtOAc = 9/1 to 6/4) to give compound **13** (0.20 g, 0.31 mmol, 99%). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ 5.20 (d, *J* = 3.1 Hz, 1H), 5.08 (dd, *J* = 8.9, 9.3 Hz, 1H), 4.98-4.92 (m, 1H), 4.83 (m, 1H), 4.72 (dd, *J* = 8.9, 9.3 Hz, 1H), 4.55 (d, *J* = 9.3 Hz, 1H), 4.42-4.35 (m, 2H), 4.02-3.96 (m, 3H), 3.80 (dd, *J* = 6.9, 6.5 Hz, 1H), 3.71 (dd, *J* = 9.6, 8.9 Hz, 1H), 3.64-3.55 (m, 1H), 2.01 (s, 3H), 1.99 (s, 3H), 1.93-1.91 (s, 12H), 1.82 (s, 3H); <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>): δ 169.9 × 2, 169.8, 169.6, 169.3, 169.1, 168.7, 100.7, 87.2, 77.2, 75.4, 74.4, 72.1, 70.5, 70.3, 68.7, 66.3, 61.5, 60.5, 20.4, 20.3, 20.2 × 4, 20.1; HRMS (ESI-TOF) calculated for C<sub>26</sub>H<sub>35</sub>N<sub>3</sub>NaO<sub>17</sub> (M+Na)<sup>+</sup>: 684.1864; found 684.1864.

#### Compound 12



Compound **13** (1.9 mg, 0.29 mmol) dissolved in MeOH (2.9 mL) was added sodium methoxide (5.0 M in MeOH, 0.21 mL, 1.0 mmol) dropwise. The reaction mixture was stirred at room temperature for 1 hour. Amberlite IR-120 (PLUS) ion-exchange resin was added and the reaction mixture was stirred for 1 hour. The resin was then removed by filtration and the resulting solution was concentrated *in vacuo* to give compound **12** (97 mg, 0.26 mmol, 92%). <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>): δ 5.66 (d, *J* = 5.9 Hz, 1H), 5.09 (d, *J* = 4.5 Hz, 1H), 4.80-4.77 (m, 2H), 4.70-4.64 (m, 2H), 4.57 (d, *J* = 8.6 Hz, 1H), 4.52 (d, *J* = 4.8 Hz, 1H), 4.20 (d, *J* = 7.6 Hz, 1H), 3.79-3.73 (m, 1H), 3.64-3.57 (m, 2H), 3.54-3.42 (m, 4H), 3.31-3.30 (m, 4H), 3.08-3.01 (m, 1H); <sup>13</sup>C NMR (75 MHz, DMSO-*d*<sub>6</sub>): δ 103.8, 89.7, 79.9, 77.0, 75.6, 74.8, 73.2, 73.1, 70.6, 68.2, 60.5, 60.1; HRMS (ESI-TOF) calculated for C<sub>12</sub>H<sub>21</sub>N<sub>3</sub>NaO<sub>10</sub> (M+Na)<sup>+</sup>: 390.1124; found 390.1120.

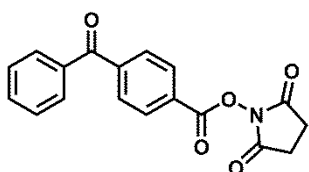
#### Compound 8



Compound **12** (9.6 mg, 26 μmol) and compound **11** (9.5 mg, 33 μmol) were dissolved in 2.7 mL of *t*BuOH/MilliQ water mixture (1/1) and tris [(1-benzyl-1*H*-1, 2, 3-triazol-4-yl) methyl] amine (1.0 mg, 2.7 μmol), sodium ascorbate (1.1 mg, 5.4 μmol), and copper sulfate pentahydrate (0.70 mg, 2.7 μmol) were added at room temperature and the resultant solution was stirred for 8 hours at 65 °C. The Quadra Pure-IDA<sup>®</sup> resin was added to the reaction mixture, which was further stirred

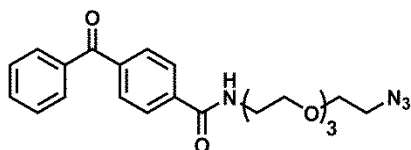
for 2 hours under an atmospheric condition. The reaction mixture was filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography (CHCl<sub>3</sub> after that CH<sub>2</sub>Cl<sub>2</sub>/MeOH/H<sub>2</sub>O = 60/10/1 to 65/25/4) to give compound **8** (14 mg, 0.021 mmol, 82%). <sup>1</sup>H NMR (300 MHz, pyridine-*d*<sub>5</sub>): δ 8.13 (s, 1H), 6.30 (d, *J* = 9.3 Hz, 1H), 5.13 (d, *J* = 7.9 Hz, 1H), 4.77-4.71 (m, 1H), 4.58-4.37 (m, 8H), 4.19-4.09 (m, 5H), 3.52-3.43 (m, 1H), 3.10-2.95 (m, 2H), 2.76 (t, *J* = 7.2 Hz, 2H), 2.32-2.18 (m, 3H), 1.77-1.24 (m, 11H); <sup>13</sup>C NMR (75 MHz, pyridine-*d*<sub>5</sub>): δ 173.4, 147.8, 124.2, 121.0, 105.9, 88.9, 81.0, 79.6, 77.4, 77.2, 75.2, 73.5, 72.4, 70.1, 64.1, 62.1, 61.5, 56.8, 40.4, 38.7, 34.8, 34.1, 29.0, 28.6, 26.1, 25.0; HRMS (ESI-TOF) calculated for C<sub>26</sub>H<sub>43</sub>N<sub>3</sub>NaO<sub>12</sub>S<sub>2</sub> (M+Na)<sup>+</sup>: 676.2186; found 676.2181.

#### 4-benzoyl benzoic NHS ester



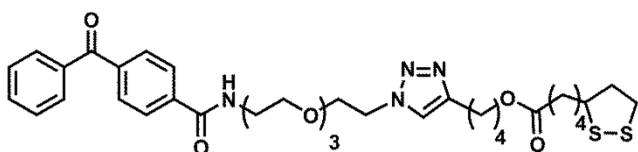
The title compound was prepared according to the published protocol.<sup>1</sup> To a solution of 4-benzoylbenzoic acid (1.0 g, 4.4 mmol) and *N*-hydroxysuccinimide (1.0 g, 8.8 mmol) in dioxane (14 mL) was added *N,N'*-dicyclohexyl carbodiimide (1.1 g, 5.3 mmol) at 0 °C and the reaction mixture was stirred for 21 hours at room temperature. The reaction was quenched by adding dioxane (14 mL) and AcOH (0.30 mL) and subsequent stirring for 30 minutes. The resulting mixture was filtered and concentrated *in vacuo*.

#### Compound **15**



To a solution of 4-benzoyl benzoic NHS ester (51 mg, 0.16 mmol) in DMF (0.52 mL) was added to DIPEA (30 μL, 0.17 mmol) and 1-amino-11-azide-3, 6, 9-trioxaundecane (31 μL, 0.16 mmol) at 0 °C. The solution was stirred for 21 hours at room temperature. The reaction mixture was diluted with EtOAc and washed with MilliQ water and brine, dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH = 100/0, 49/1, 19/1 to 93/7) to give compound **15** (51 mg, 0.12 mmol, 76%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ 7.86 (d, *J* = 8.2 Hz, 2H), 7.74 (d, *J* = 8.2 Hz, 2H), 7.70 (d, *J* = 7.3 Hz, 2H), 7.53 (t, *J* = 7.3 Hz, 1H), 7.41 (t, *J* = 7.8 Hz, 2H), 7.15 (s, 1H), 3.64-3.52 (m, 14H), 3.25 (t, *J* = 5.0 Hz, 2H); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>): δ 195.7, 166.4, 139.6, 137.6, 136.8, 132.6, 129.7 × 2, 129.7 × 2, 128.2 × 2, 126.8 × 2, 70.3 × 2, 70.2, 70.0, 69.7, 69.3, 50.3, 39.7; HRMS (ESI-TOF) calculated for C<sub>22</sub>H<sub>26</sub>N<sub>4</sub>NaO<sub>5</sub> (M+Na)<sup>+</sup>: 449.1801; found: 449.1801.

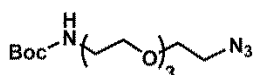
#### Compound **9**



To a solution of **11** (25 mg, 89 μmol) in MeCN (95 μL), *t*BuOH (95 μL) and MilliQ water (95 μL), were added **15** (37 mg, 86 μmol), *o*-phenylenediamine (1.4 mg, 13 μmol), sodium ascorbate (1.8 mg, 8.9 μmol) and copper sulfate pentahydrate (1.1 mg, 4.4 μmol). The reaction mixture was stirred at room temperature under argon atmosphere for 19 hours. The

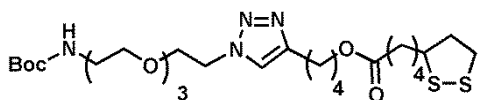
Quadra Pure-IDA<sup>®</sup> resin was added followed by 2 hours stirring under an air atmosphere. The reaction mixture was then filtered with MeOH/CHCl<sub>3</sub> (1/9). The residue was purified by flash column chromatography (EtOAc/MeOH = 100/0, 19/1 to 9/1) to give compound **9** (42 mg, 0.059 mmol, 67%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ 7.91 (d, *J* = 8.7 Hz, 2H), 7.81 (d, *J* = 7.8 Hz, 2H), 7.76 (m, 2H), 7.61-7.57 (m, 1H), 7.47 (dd, *J* = 7.3, 7.8 Hz, 2H), 7.40 (s, 1H), 7.11 (br, 1H), 4.43 (t, *J* = 6.0 Hz, 2H), 4.05 (t, *J* = 6.0 Hz, 2H), 3.81 (t, *J* = 5.0 Hz, 2H), 3.67-3.49 (m, 12H), 3.17-3.03 (m, 2H), 2.76-2.68 (m, 2H), 2.46-2.38 (m, 1H), 2.27 (t, *J* = 7.3 Hz, 2H), 2.19 (s, 1H), 1.91-1.83 (m, 1H), 1.74-1.56 (m, 8H), 1.51-1.35 (m, 2H); <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>): δ 195.9, 173.5, 166.5, 147.4, 139.8, 137.8, 136.9, 132.8, 130.0 × 2, 129.9 × 2, 128.4 × 2, 127.0 × 2, 121.7, 70.4 × 2, 70.3, 70.1, 69.6, 69.4, 64.0, 56.2, 49.9, 40.1, 39.8, 38.4, 34.5, 34.0, 28.7, 28.1, 25.8, 25.1, 24.6; HRMS (ESI-TOF) calculated for C<sub>36</sub>H<sub>48</sub>N<sub>4</sub>NaO<sub>7</sub>S<sub>2</sub> (M+Na)<sup>+</sup>: 735.2862; found: 735.2863.

#### Compound 17



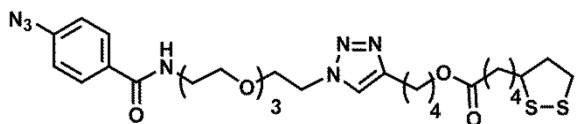
The title compound was prepared according to a published protocol.<sup>2</sup> To a solution of 1-amino-11-azide-3, 6, 9-trioxaundecane (0.10 g, 0.46 mmol) in EtOH (0.92 mL) was added di-tert-butyl pyrocarbonate (0.11 g, 0.50 mmol) at 0 °C. The reaction mixture was stirred at room temperature for 22 hours, evaporated, and diluted with CH<sub>2</sub>Cl<sub>2</sub> (20 mL) and MilliQ water (20 mL). The mixture was extracted twice with CH<sub>2</sub>Cl<sub>2</sub>, and the combined extracts were washed with brine, dried over sodium sulfate, filtered, and concentrated *in vacuo*. The residue was purified by flash column chromatography (hexane/EtOAc = 1/1) to give compound **17** (0.12 g, 0.38 mmol, 84%). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ 5.03 (br, 1H), 3.66-3.54 (m, 10H), 3.48 (t, *J* = 5.2 Hz, 2H), 3.34 (t, *J* = 4.8 Hz, 2H), 3.25 (t, *J* = 4.8 Hz, 2H), 1.39 (s, 9H); <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>): δ 155.8, 130.7, 128.6, 78.9, 70.5, 70.4, 70.0, 69.9, 50.5, 40.2, 28.2 × 3; HRMS (ESI-TOF) calculated for C<sub>13</sub>H<sub>26</sub>N<sub>4</sub>NaO<sub>5</sub> (M+Na)<sup>+</sup>: 341.1801; found: 341.1805.

#### Compound 16



To a solution of compound **17** (86 mg, 0.27 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (2.7 mL) was added compound **5** (78 mg, 0.27 mmol), copper (I) iodide (0.5 mg, 2.7 μmol), tris [(1-benzyl-1*H*-1, 2, 3-triazol-4-yl) methyl] amine (1.4 mg, 2.7 μmol) and *N,N*-diisopropylethylamine (4.7 μL, 27 μmol) and the reaction mixture was stirred at room temperature for 21 hours. The Quadra Pure-IDA<sup>®</sup> resin was added to stir for 2 hours. The reaction mixture was then filtered with CH<sub>2</sub>Cl<sub>2</sub>. The residue was concentrated *in vacuo* and was purified by flash column chromatography (CHCl<sub>3</sub>/MeOH/AcOH = 99/1/0.1) to give compound **16** (0.15 g, 0.26 mmol, 94%). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>): δ 7.44 (br, 1H), 5.09 (br, 1H), 4.46 (t, *J* = 5.0 Hz, 2H), 4.03 (t, *J* = 6.0 Hz, 2H), 3.81 (t, *J* = 5.0 Hz, 2H), 3.55 (d, *J* = 3.8 Hz, 8H), 3.53-3.46 (m, 3H), 3.25-3.24 (m, 2H), 3.16-3.00 (m, 2H), 2.69 (m, 2H), 2.45-2.35 (m, 1H), 2.25 (t, *J* = 7.4 Hz, 2H), 1.90-1.79 (m, 1H), 1.74-1.52 (m, 8H), 1.50-1.41 (m, 2H), 1.37 (s, 9H); <sup>13</sup>C NMR (75 MHz, CDCl<sub>3</sub>): δ 175.0, 173.5, 155.9, 147.2, 121.9, 79.1, 70.3 × 2, 70.0, 69.4, 63.9, 56.2, 50.0, 40.1, 40.0, 38.3, 34.4, 33.9, 28.6, 28.2 × 3, 28.0, 25.7, 24.9, 24.5, 20.7; HRMS (ESI-TOF) calculated for C<sub>27</sub>H<sub>48</sub>N<sub>4</sub>NaO<sub>7</sub>S<sub>2</sub> (M+Na)<sup>+</sup>: 627.2862; found: 627.2867.

#### Compound 10



To a solution of compound **16** (33 mg, 55  $\mu$ mol) in  $\text{CH}_2\text{Cl}_2$  (2.0 mL) was added trifluoroacetic acid (0.50 mL) at 0  $^\circ\text{C}$ . After stirring at room temperature for 1 hour, the reaction mixture was coevaporated with toluene and was used in next step without further purification. The mixture was dissolved with  $\text{CH}_2\text{Cl}_2$  (0.17 mL) and was added DIPEA (18  $\mu$ L, 0.11 mmol) and 4-azide benzoic acid (8.5 mg, 52  $\mu$ mol). DMT-MM (14 mg, 52  $\mu$ mol) were added at 0  $^\circ\text{C}$  and the reaction mixture was stirred at room temperature for 18 hours. The reaction mixture was washed with MilliQ water three times, brine, dried over with  $\text{Na}_2\text{SO}_4$ , filtered and concentrated *in vacuo*. The residue was purified by flash column chromatography ( $\text{CH}_2\text{Cl}_2/\text{MeOH} = 97/3, 95/5$  to  $90/10$ ) to give compound **10** (26 mg, 39  $\mu$ mol, 76%).  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  7.82 (d,  $J = 8.6$  Hz, 2H), 7.41 (s, 1H), 7.05 (d,  $J = 8.6$  Hz, 2H), 6.87 (br, 1H), 4.46 (t,  $J = 5.2$  Hz, 2H), 4.07 (t,  $J = 6.0$  Hz, 2H), 3.83 (t,  $J = 5.2$  Hz, 2H), 3.65-3.51 (m, 13H), 3.22-3.06 (m, 2H), 2.72 (t,  $J = 7.2$  Hz, 2H), 2.51-2.40 (m, 1H), 2.30 (t,  $J = 7.4$  Hz, 2H), 1.95-1.84 (m, 1H), 1.75-1.59 (m, 8H), 1.57-1.39 (m, 2H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{CDCl}_3$ ):  $\delta$  173.6, 166.4, 147.5, 143.2, 131.0, 128.9  $\times$  2, 121.8, 118.9  $\times$  2, 70.5, 70.4, 70.1, 69.8, 69.5, 64.0, 56.3, 50.0, 40.2, 39.8, 38.5, 34.6, 34.0, 31.1, 28.7, 28.2, 25.9, 25.2, 24.6; HRMS (ESI-TOF) calculated for  $\text{C}_{29}\text{H}_{43}\text{N}_7\text{NaO}_6\text{S}_2$  ( $\text{M}+\text{Na}$ ) $^+$ : 672.2613; found: 672.2614.

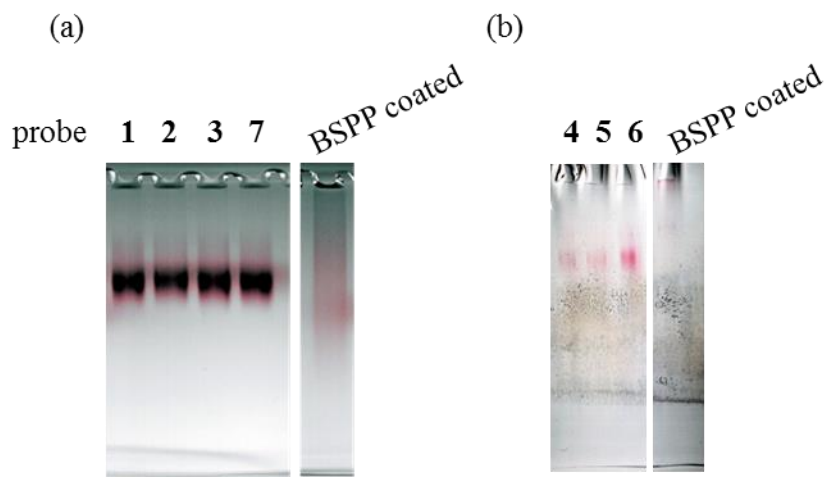
### 3. Synthesis and characterization of photoaffinity probe 1-7

#### 3-1. Assembly of photoaffinity probe 1-7<sup>3</sup>

BSPP was added to the solution of colloidal gold nanoparticles (AuNPs, 5 nm) and the mixture was stirred for 1 hour at 50  $^\circ\text{C}$ . The resulting mixture was centrifuged at  $20,000\times g$  for 90 min and the supernatant was carefully removed. The BSPP-stabilized AuNPs were washed with  $\text{H}_2\text{O}$  three times. A stock solution of compound **8-10** in MeOH (10 mM) was mixed and added to a solution of the BSPP-stabilized gold nanoparticles and the mixture was stirred at room temperature for 24 hours. The functionalized gold nanoparticles were washed with MilliQ water  $\times$  2, MeOH/ MilliQ water (1/1)  $\times$  1 then MilliQ water  $\times$  1. The concentration of the functionalized gold nanoparticles was determined by the visible absorbance at 520 nm.

#### 3-2. SDS-PAGE analysis

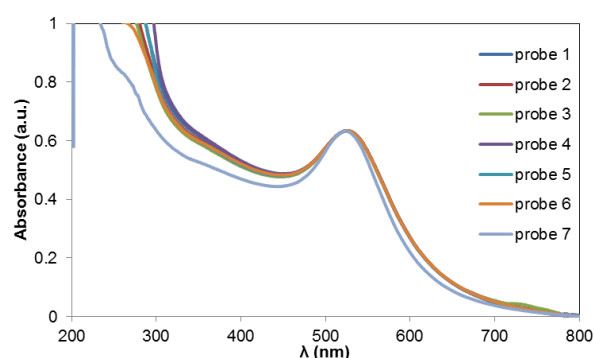
Each of photoaffinity probes **1-7** (4 pmol) was suspended in  $6\times$ LDS sample buffer (0.3 M Tris-HCl, pH 6.8, 10% LDS, 30% glycerol and 0.6 M DTT) and was resolved by 7.5% SDS-PAGE gel.



**Figure S1.** SDS-PAGE analysis of probes **1-7**; (a) probe **1**, probe **2**, probe **3**, probe **7**, BSPP coated AuNPs, (b) probe **4**, probe **5**, probe **6**, BSPP coated AuNPs.

### 3-3. UV-VIS spectrometric analysis

The stock solution photoaffinity probes **1–7** (4 pmol) was diluted in 100  $\mu\text{L}$  of HEPES buffer (10 mM HEPES, pH7.4, 150 mM NaCl, 1 mM  $\text{MnCl}_2$ , 1 mM  $\text{CaCl}_2$ ) and was recorded on VARIAN CARY 50 conc UV-visible spectrophotometer.



**Figure S2.** UV-VIS spectra of probes **1–7**.

### 3-4. MALDI-TOF-MS analysis

The stock solution for photoaffinity probes **1–7** (0.1 pmol) was suspended with 6  $\mu\text{L}$  of a matrix solution (10 mg/mL of 2, 5-dihydroxybenzoic acid in  $\text{MeCN}/\text{H}_2\text{O} = 1/1$ ). This mixture was spotted onto the sample plate for direct MALDI-TOF MS analysis. MALDI-TOF-mass spectra were acquired by AB SCIEX TOF/TOF<sup>TM</sup> 5800 (AB SCIEX). As references, the samples containing free ligands **8**, **9** or **10** were also analyzed and the corresponding peaks were observed at  $m/z = 676.2$  ( $\text{M}+\text{Na}$ )<sup>+</sup> for **8**, at  $m/z = 735.3$  ( $\text{M}+\text{Na}$ )<sup>+</sup> for **9** and at  $m/z = 644.3$  ( $\text{M}+\text{Na}-\text{N}_2$ )<sup>+</sup> for **10**.

Probe number	Compound	Chemical formula	Calculated mass	Observed mass
<b>1</b>	<b>8</b>	$\text{C}_{26}\text{H}_{43}\text{N}_3\text{NaO}_{12}\text{S}_2$ ( $\text{M}+\text{Na}$ ) <sup>+</sup>	676.2	676.6
	<b>9</b>	$\text{C}_{36}\text{H}_{48}\text{N}_4\text{NaO}_7\text{S}_2$ ( $\text{M}+\text{Na}$ ) <sup>+</sup>	735.3	735.7
<b>2</b>	<b>8</b>	$\text{C}_{26}\text{H}_{43}\text{N}_3\text{NaO}_{12}\text{S}_2$ ( $\text{M}+\text{Na}$ ) <sup>+</sup>	676.2	676.4
	<b>9</b>	$\text{C}_{36}\text{H}_{48}\text{N}_4\text{NaO}_7\text{S}_2$ ( $\text{M}+\text{Na}$ ) <sup>+</sup>	735.3	735.7
<b>3</b>	<b>8</b>	$\text{C}_{26}\text{H}_{43}\text{N}_3\text{NaO}_{12}\text{S}_2$ ( $\text{M}+\text{Na}$ ) <sup>+</sup>	676.2	676.4
	<b>9</b>	$\text{C}_{36}\text{H}_{48}\text{N}_4\text{NaO}_7\text{S}_2$ ( $\text{M}+\text{Na}$ ) <sup>+</sup>	735.3	735.6
<b>4</b>	<b>8</b>	$\text{C}_{26}\text{H}_{43}\text{N}_3\text{NaO}_{12}\text{S}_2$ ( $\text{M}+\text{Na}$ ) <sup>+</sup>	676.2	676.2
	<b>10</b>	$\text{C}_{29}\text{H}_{43}\text{N}_5\text{NaO}_6\text{S}_2$ ( $\text{M}+\text{Na}-\text{N}_2$ ) <sup>+</sup>	644.3	644.2
<b>5</b>	<b>8</b>	$\text{C}_{26}\text{H}_{43}\text{N}_3\text{NaO}_{12}\text{S}_2$ ( $\text{M}+\text{Na}$ ) <sup>+</sup>	676.2	676.7
	<b>10</b>	$\text{C}_{29}\text{H}_{43}\text{N}_5\text{NaO}_6\text{S}_2$ ( $\text{M}+\text{Na}-\text{N}_2$ ) <sup>+</sup>	644.3	644.7
<b>6</b>	<b>8</b>	$\text{C}_{26}\text{H}_{43}\text{N}_3\text{NaO}_{12}\text{S}_2$ ( $\text{M}+\text{Na}$ ) <sup>+</sup>	676.2	676.3
	<b>10</b>	$\text{C}_{29}\text{H}_{43}\text{N}_5\text{NaO}_6\text{S}_2$ ( $\text{M}+\text{Na}-\text{N}_2$ ) <sup>+</sup>	644.3	644.1
<b>7</b>	<b>8</b>	$\text{C}_{26}\text{H}_{43}\text{N}_3\text{NaO}_{12}\text{S}_2$ ( $\text{M}+\text{Na}$ ) <sup>+</sup>	676.2	676.3

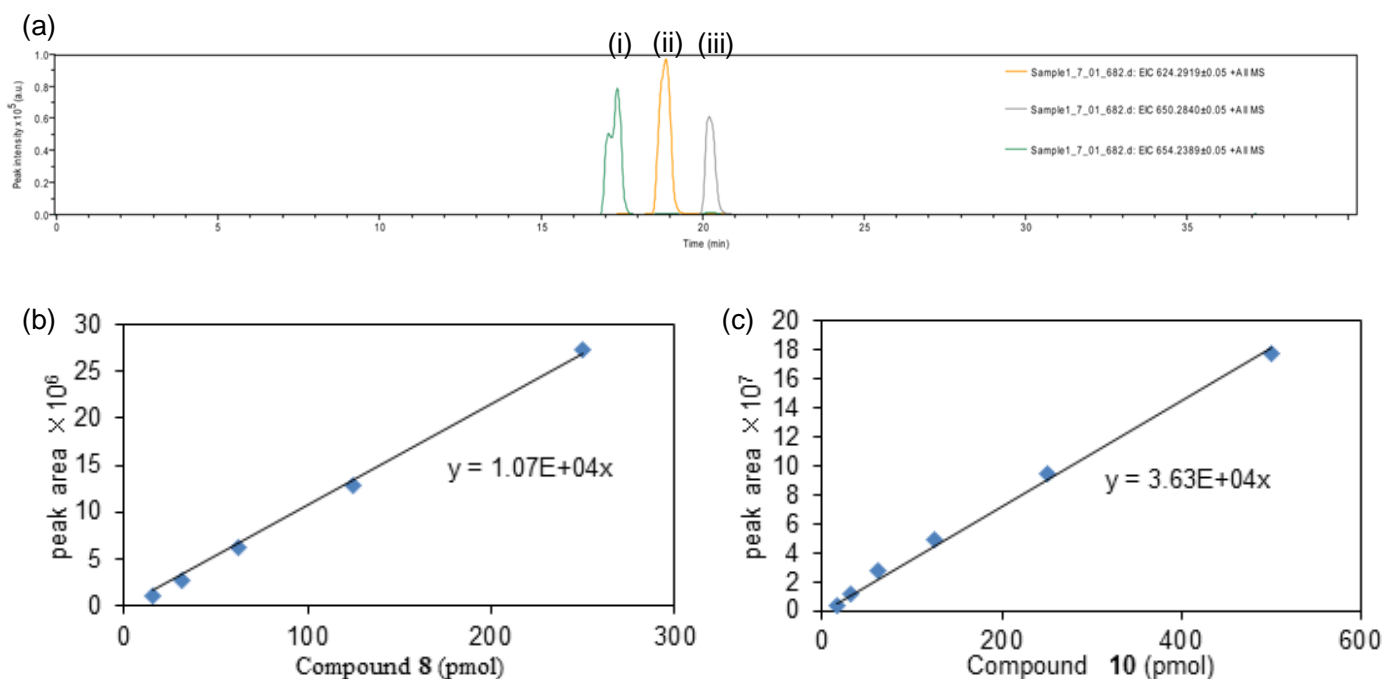
**Table S1.** MALDI-TOF mass peaks corresponding to ligand **8–10** derived from probes **1-7**.

### 3-5. LC-MS analysis of probe **5** to determine the ligand density for **8** and **10**.

Probe **5** (20 pmol) were treated with 2-mercaptoethanol (1 M, 50  $\mu\text{L}$ , in MilliQ water) at 40  $^\circ\text{C}$  for 2 hours. The resulting mixture was centrifuged at 20,000 $\times g$  for 80 minutes to remove AuNPs. The supernatant was collected and the residual 2-mercaptoethanol was removed *in vacuo*. The residue was dissolved in DMSO (50  $\mu\text{L}$ ) was analyzed to determine the amounts of compound **8** and **10** cleaved from the AuNP surface by HPLC using a 2.0 x 30.0 mm (5  $\mu\text{m}$ ) Cadenza 5CD-C18



(Imtakt) at a flow rate of 0.2 mL/min (5–95% MeCN/H<sub>2</sub>O gradient) equipped with micro TOF-QZ-TND1 (BRUKER). The peak area for the total ion count peaks corresponding to compound **8** and **10** was quantified using ESI Compass 1.3 for microTOF Data Analysis Version 4.0 SP1 and by using the standard curve obtained for the total ion count peak intensity of the serially diluted samples of compound **8** and **10** respectively (Figure S3a and b).

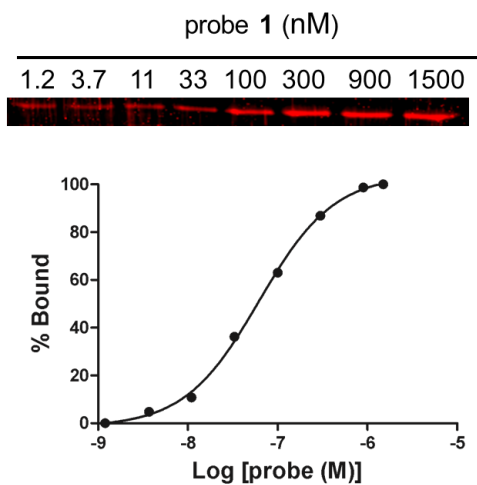


**Figure S3.** (a) Extracted ion chromatogram for LC-MS analysis of the sample containing a mixture of Lac ligand **8** and aryl azide conjugate **10** obtained by cleaving from probe **5**; (i) a peak corresponding to  $(\mathbf{10}+2\text{H}-\text{N}_2)^+$ , (ii) a peak corresponding to  $(\mathbf{10}+\text{H})^+$ , (iii) a peak corresponding to  $(\mathbf{8}+\text{H})^+$ . Standard curves plotted for (b) the peak (iii) corresponding to Lac ligand **8**, (c) the peak (i) and (ii) corresponding to aryl azide conjugate **10**.

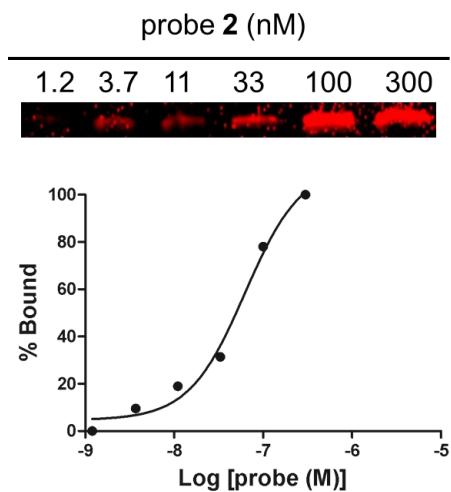
#### 4. Affinity pull-down assay.

PNA (0.5  $\mu\text{g}$ ) were mixed with varied amounts of probe **1–7** (14, 4.7, 1.6, 0.52, 0.17, 0.058 pmol) in 100  $\mu\text{L}$  of HEPES buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 1 mM  $\text{MnCl}_2$ , 1 mM  $\text{CaCl}_2$ ) and the mixture was incubated on rotary mixer at 4  $^\circ\text{C}$  for 2 hours. The mixture was centrifuged at 20,000 $\times g$  for 60 minutes. The supernatant was removed and the precipitated probes were washed with PBS (500  $\mu\text{L}$ ) followed by centrifugation at 20,000 $\times g$  for 1 hour. The washing step was repeated three times to isolate the probe–protein complex in the pellet. The pellet was suspended in LDS sample buffer containing 10% 2-mercaptoethanol and incubated 30 minutes at 95  $^\circ\text{C}$ . The sample solution was resolved by SDS-PAGE using 10% gel. The resultant gel was stained with Flamingo Fluorescent Gel stain and analyzed by fluorescence imaging on Typhoon 8600 (GE Healthcare Science). The  $K_d$  values were obtained by the non-linear curve fitting analysis of the plots of the fluorescence intensity of the PNA bands using Graphpad Prism (Graphpad Software, Inc.).

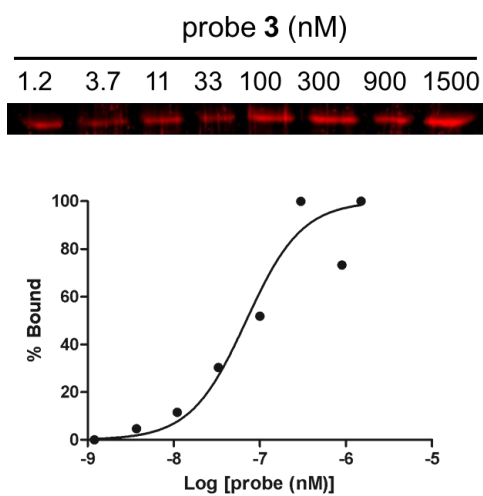
(a)



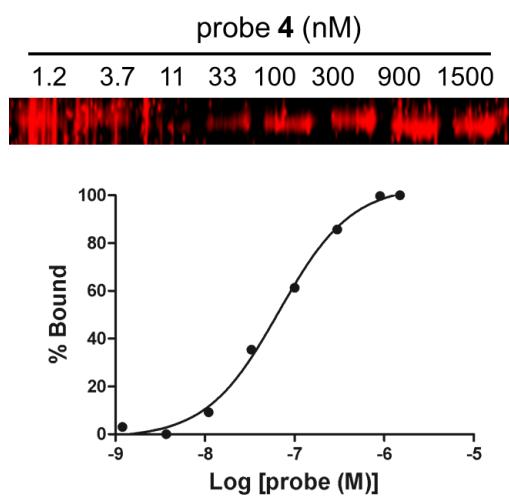
(b)



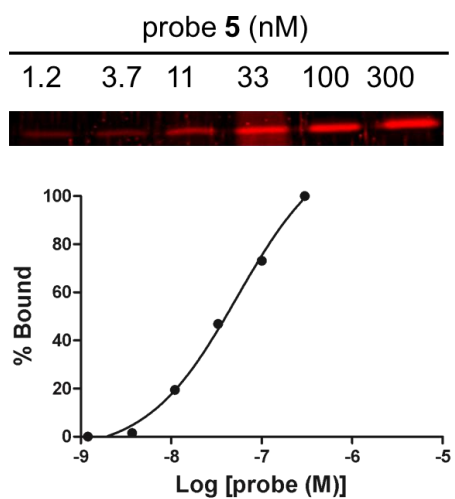
(c)



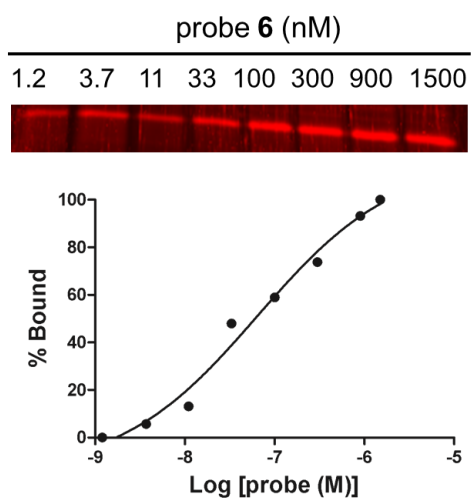
(d)



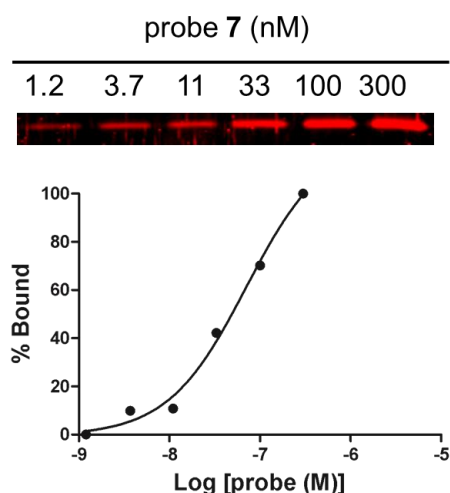
(e)



(f)



(g)

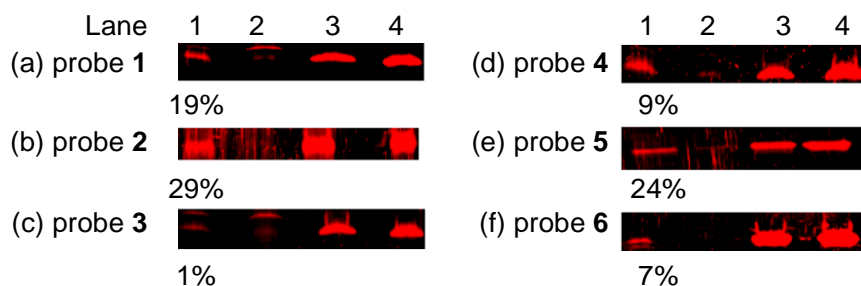


**Figure S4.** A plot showing the binding of probe 1–7 (a-g) to PNA (91 nM) in PBS evaluated by an affinity pull-down assay and gel-based fluorescence imaging.

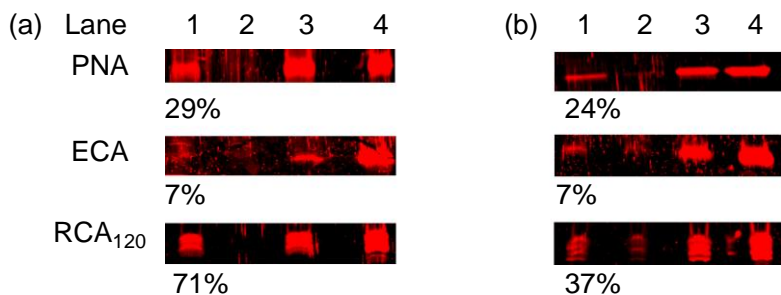
## 5. Photoaffinity labeling experiments.

### 5.1. PAL reactions with carbohydrate binding proteins.

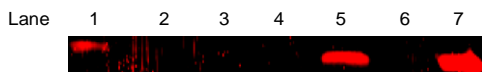
A lactose binding lectin (1.2  $\mu$ g) was mixed with photoaffinity probe 1–7 (10 pmol, 100 nM) in 100  $\mu$ L of HEPES buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 1 mM CaCl<sub>2</sub>, 1 mM MnCl<sub>2</sub>) and the mixture was incubated on a rotary mixer at 4 °C for 2 hours. The mixture was irradiated at 365 nm at a distance of 5 cm on ice for 1 hour using a hand held UV lamp (15W, UVP XX Series). The mixture was centrifuged at 20,000 $\times$ g for 1 hour and the supernatant was removed. 3 M guanidine hydrochloride (Gdn-HCl) in PBS was added to the pellet and the mixture was agitated on a vortex mixer for 1 hour. The mixture was then centrifuged at 20,000 $\times$ g for 1 hour and the supernatant was removed. This washing step was repeated three times. The resultant pellet was further washed with PBS and the mixture was centrifuged at 20,000 $\times$ g for 1 hour and the supernatant was removed. This washing step was repeated one more time. The pellet was suspended in LDS sample buffer containing 10% 2-mercaptoethanol and incubated 30 minutes at 95 °C. The sample was resolved by SDS-PAGE using 10% gel. The resultant gel was stained by Flamingo stain and analyzed by fluorescence imaging on Typhoon 8600 (GE Healthcare Science).



**Figure S5.** Photoaffinity labeling of PNA (100 nM) and each probes (100 nM); (a) probe 1, (b) probe 2, (c) probe 3, (d) probe 4, (e) probe 5, (f) probe 6. Lane 1: crosslinked-PNA enriched by Gdn-HCl wash, lane 2: negative control experiment with no UV irradiation, lane 3: negative control experiment with no UV irradiation and no Gdn-HCl wash, lane 4: 1  $\mu$ g PNA as a reference. The % yields are calculated based on the amounts of the protein used for a given reaction.



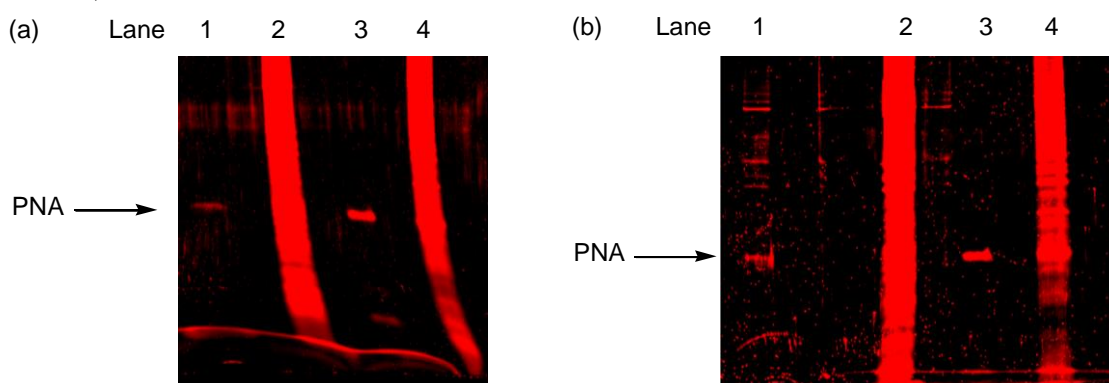
**Figure S6.** Photoaffinity labeling of PNA, ECA, RCA (100 nM) and (a) probe **2** (100 nM), (b) probe **5** (100 nM). Lane 1: crosslinked-PNA enriched by Gdn-HCl wash, lane 2: negative control experiment with no UV irradiation, lane 3: negative control experiment with no UV irradiation and no Gdn-HCl wash, lane 4: 1 µg proteins as a reference. The % yields are calculated based on the amounts of the protein used for a given reaction.



**Figure S7.** Photoaffinity labeling of PNA (100 nM) and probe **5** (100 nM). Lane 1: photocrosslinked-PNA enriched by Gdn-HCl wash, lane 2: PNA crosslinked in the presence of lactose (100 mM) and enriched by Gdn-HCl wash, lane 3: negative control experiment with no UV irradiation and no Gdn-HCl wash, lane 4: same as lane 4 with lactose (100 mM), lane 5: same as lane 3 without stringent washing step, lane 6: same as lane 5 with lactose (100 mM), lane 7: proteins as a reference. The % yields are calculated based on the amounts of the protein used for a given reaction.

### 5.2. PAL reactions with PNA in HeLa lysate.

PNA (1.0 µg) and HeLa cell lysate (100 µg) were mixed with probe **2** and **5** (10 pmol, 100 nM) in 200 µL of the HEPES buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 5 mM CaCl<sub>2</sub>, 5 mM MnCl<sub>2</sub>). The mixture was incubated on a rotary mixer at 4 °C for 2 hours, which was then irradiated at 365 nm at a distance of 5 cm on ice for 1 hour using a hand held UV lamp (15W, UVP XX Series). The mixture was centrifuged at 20,000×g for 1 hour and supernatant was separated. 3 M Gdn-HCl in PBS was added to the pellet and the mixture was vortex for 1 hour. The mixture was then centrifuged at 20,000×g for 1 hour and the supernatant was removed. This washing step was repeated three times. The resultant pellet was further washed with PBS and the mixture was centrifuged at 20,000×g for 1 hour and the supernatant was removed. This washing step was repeated one more time. The pellet was suspended in LDS sample buffer containing 10% 2-mercaptoethanol and incubated 30 minutes at 95 °C. The sample was resolved by SDS-PAGE using 10% gel. The resultant gel was stained by Flamingo stain and analyzed by fluorescence imaging on Typhoon 8600 (GE Healthcare Science).

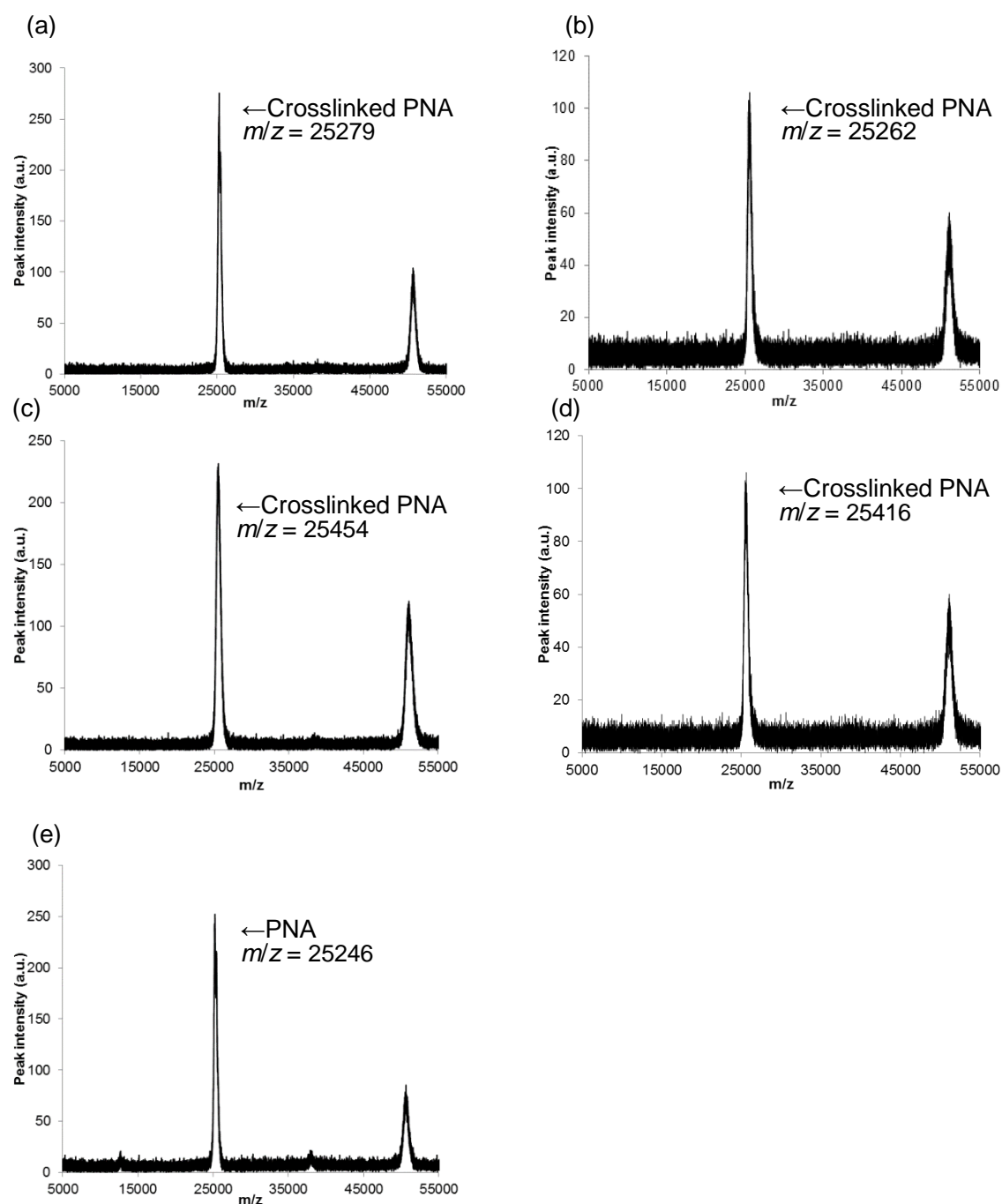


**Figure S8.** Photoaffinity labeling of PNA (1 µg) in the presence of HeLa lysate (100 µg) (a) using probe **2** (50 nM), (b) using probe **5** (50 nM); Lane 1: crosslinked-PNA enriched by Gdn-HCl wash, lane 2: negative control experiment with no UV irradiation and no Gdn-HCl wash, lane 3: 1 µg PNA input as a reference, lane 4: 1 µg PNA mixed with 100 µg HeLa cell lysate as a reference.

### 5.3. MALDI-TOF-MS analysis of probe crosslinked PNA obtained by PAL reaction.

The pellet containing PAL products obtained as described in the section 5.1 or 5.2 were suspended with 6 µL of MALDI

matrix solution composed of sinapinic acid (10 mg/mL) in solution of MeCN: MilliQ water = 1:1. This mixture was spotted onto the sample plate for direct MALDI-TOF MS analysis. MALDI-TOF-mass spectra were acquired by AB SCIEX TOF/TOF™ 5800 (AB SCIEX).

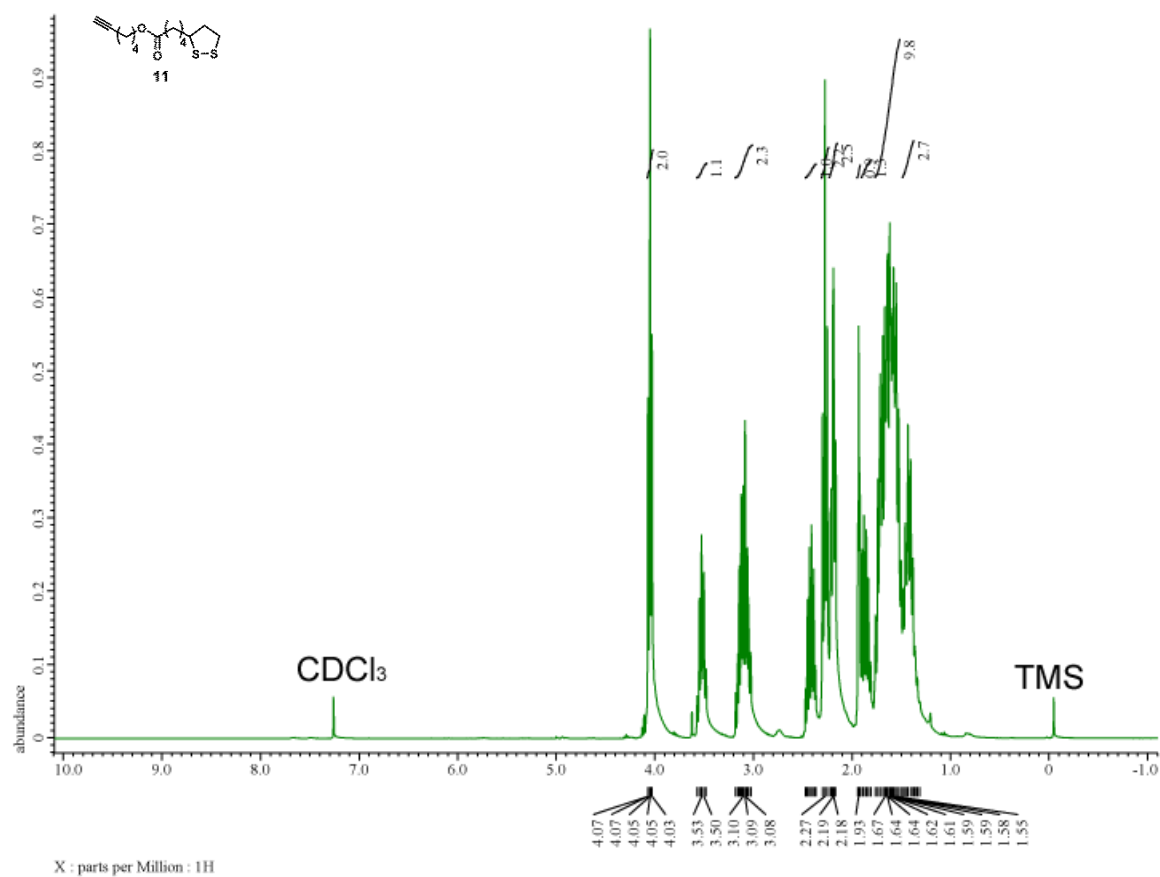


**Figure S9.** MALDI-TOF mass spectra of (a) PNA crosslinked by probe 2, (b) PNA crosslinked by probe 2 in the presence of HeLa cell lysate, (c) PNA crosslinked by probe 5, (d) PNA crosslinked by probe 5 in the presence of HeLa lysate, (e) PNA as a reference (calculated  $M_w = 25189$ ).

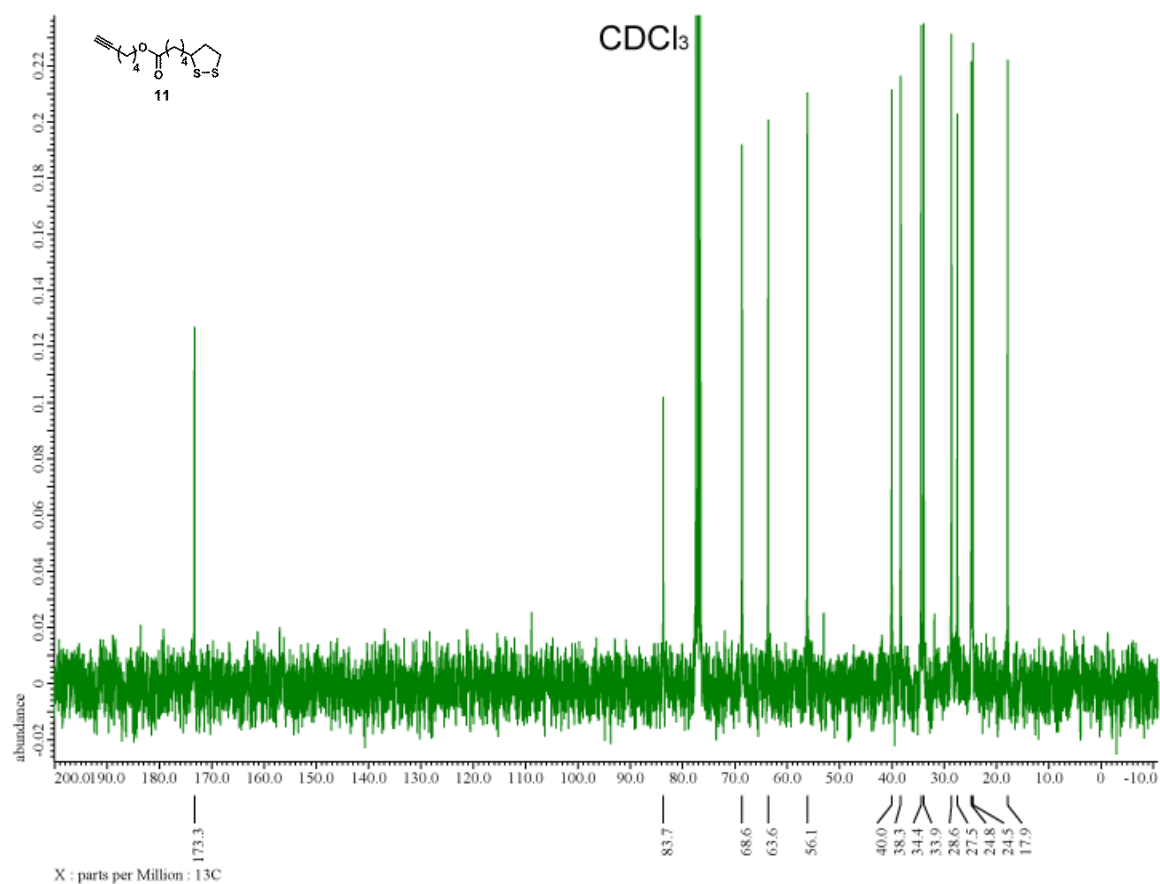
### References

1. K. Sakurai, M. Tawa, A. Okada, R. Yamada, N. Sato, M. Inahara and M. Inoue, *Chem. Asian J.*, 2012, **7**, 1567.
2. K. N. Taavi, S. T. Hyun, R. S. Ashley, J. S. Michael, W. C. Timothy, B. S. Thomas, R. Kanak, A. H. Scott and M. C. Craig, *Nat. Chem. Biol.*, 2012, **7**, 538.
3. K. Suzuki, K. Hosokawa and M. Maeda, *J. Am. Chem. Soc.*, 1997, **119**, 12384.

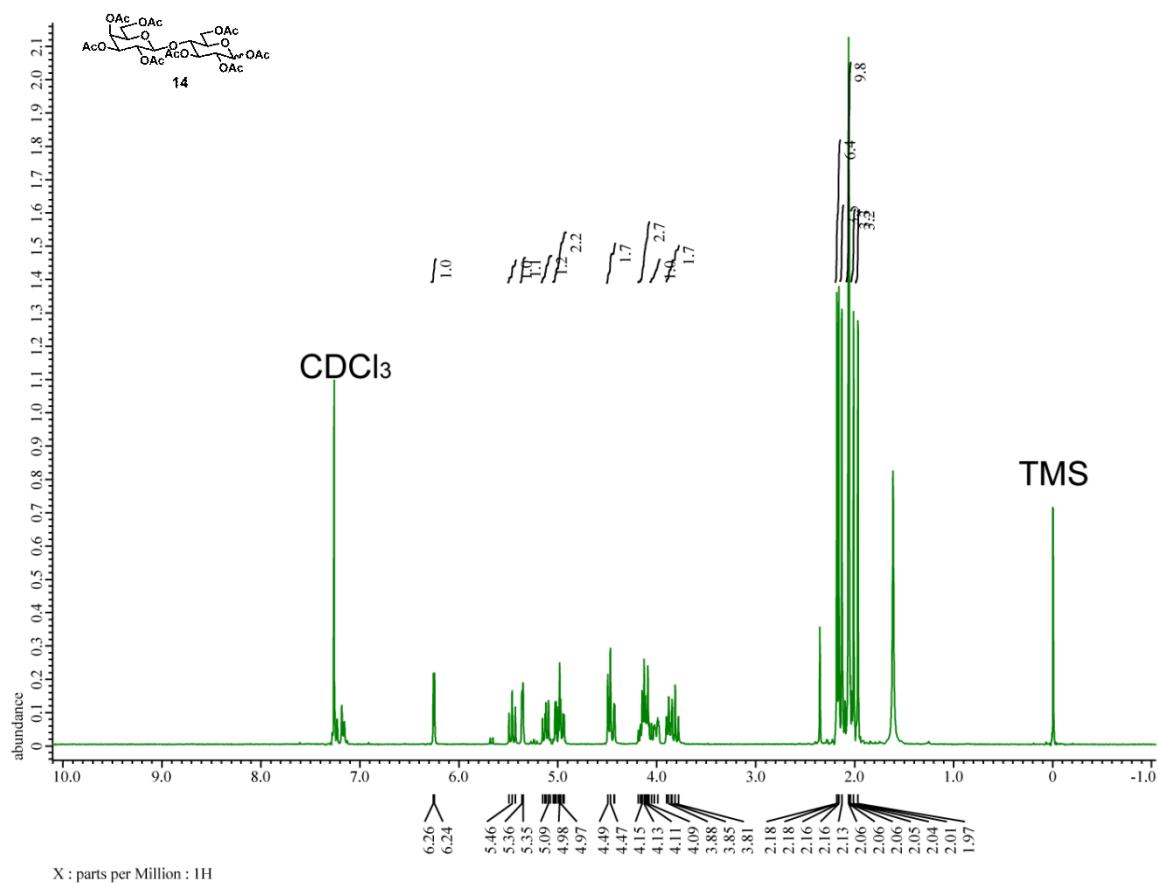
$^1\text{H}$  NMR spectra of compound **11** ( $\text{CDCl}_3$ , 300 MHz).



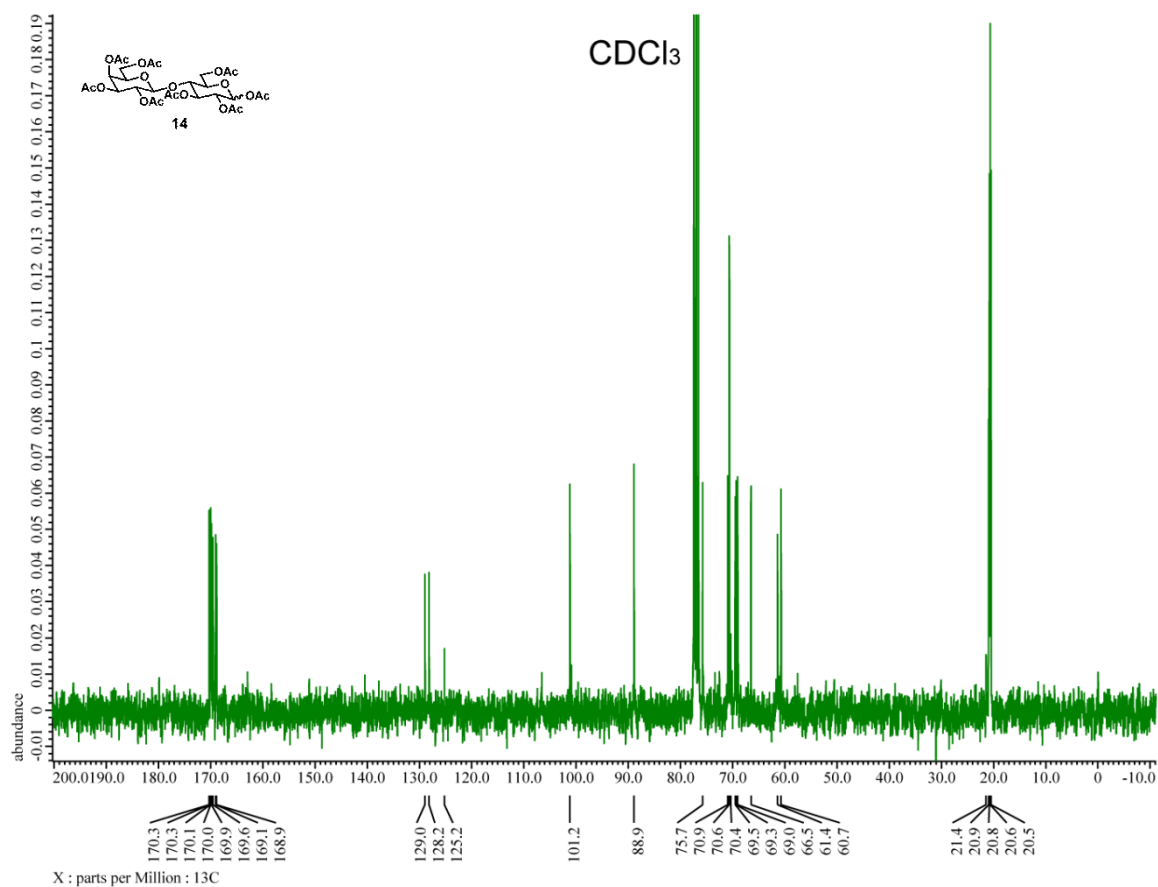
$^{13}\text{C}$  NMR spectra of compound **11** ( $\text{CDCl}_3$ , 75 MHz).



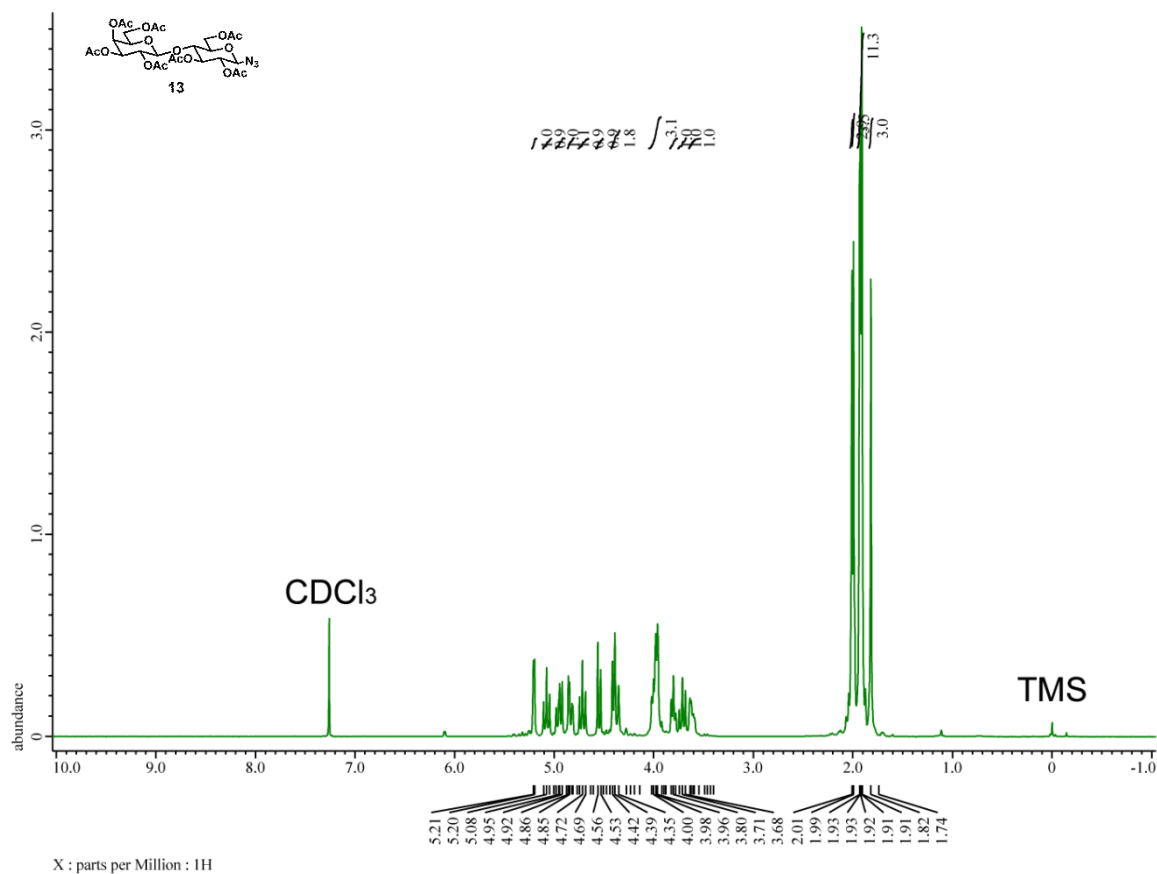
<sup>1</sup>H NMR spectra of compound **14** (CDCl<sub>3</sub>, 300 MHz).



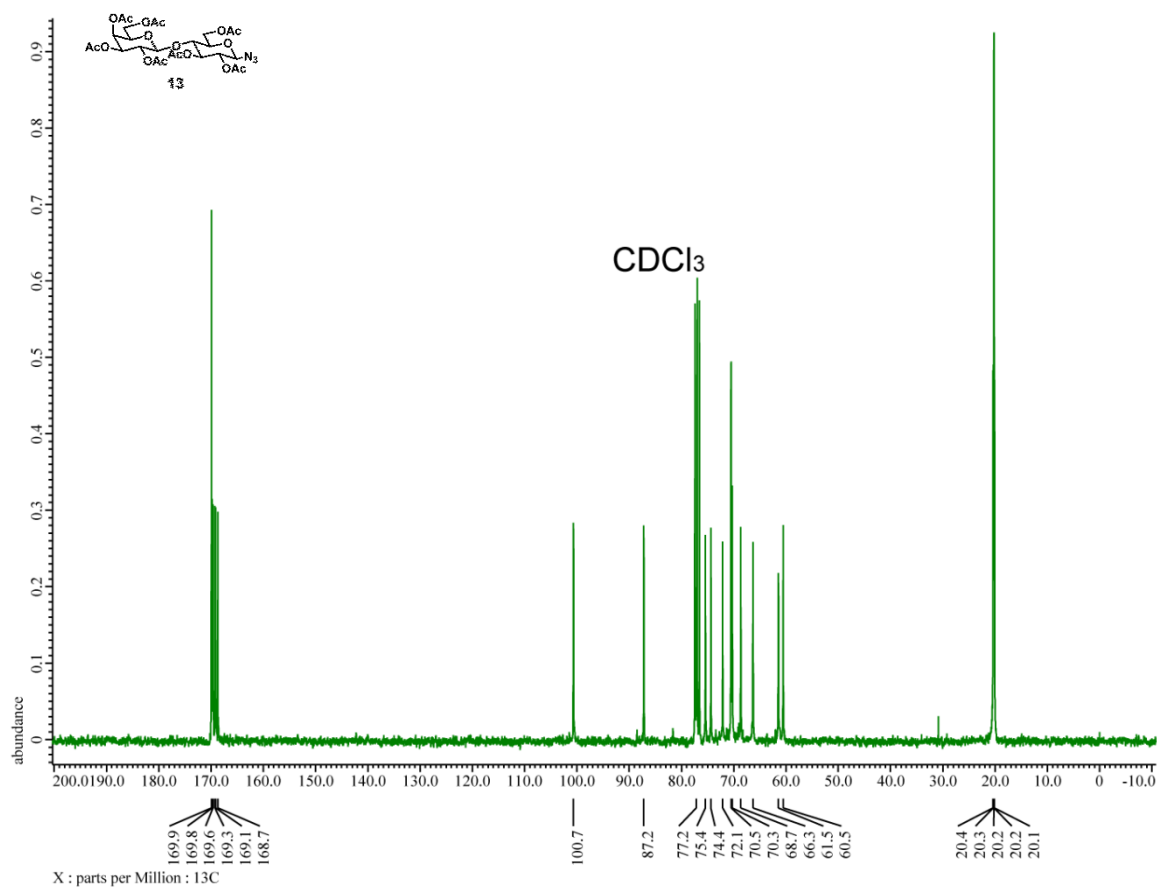
<sup>13</sup>C NMR spectra of compound **14** (CDCl<sub>3</sub>, 75 MHz).



<sup>1</sup>H NMR spectra of compound **13** (CDCl<sub>3</sub>, 300 MHz).

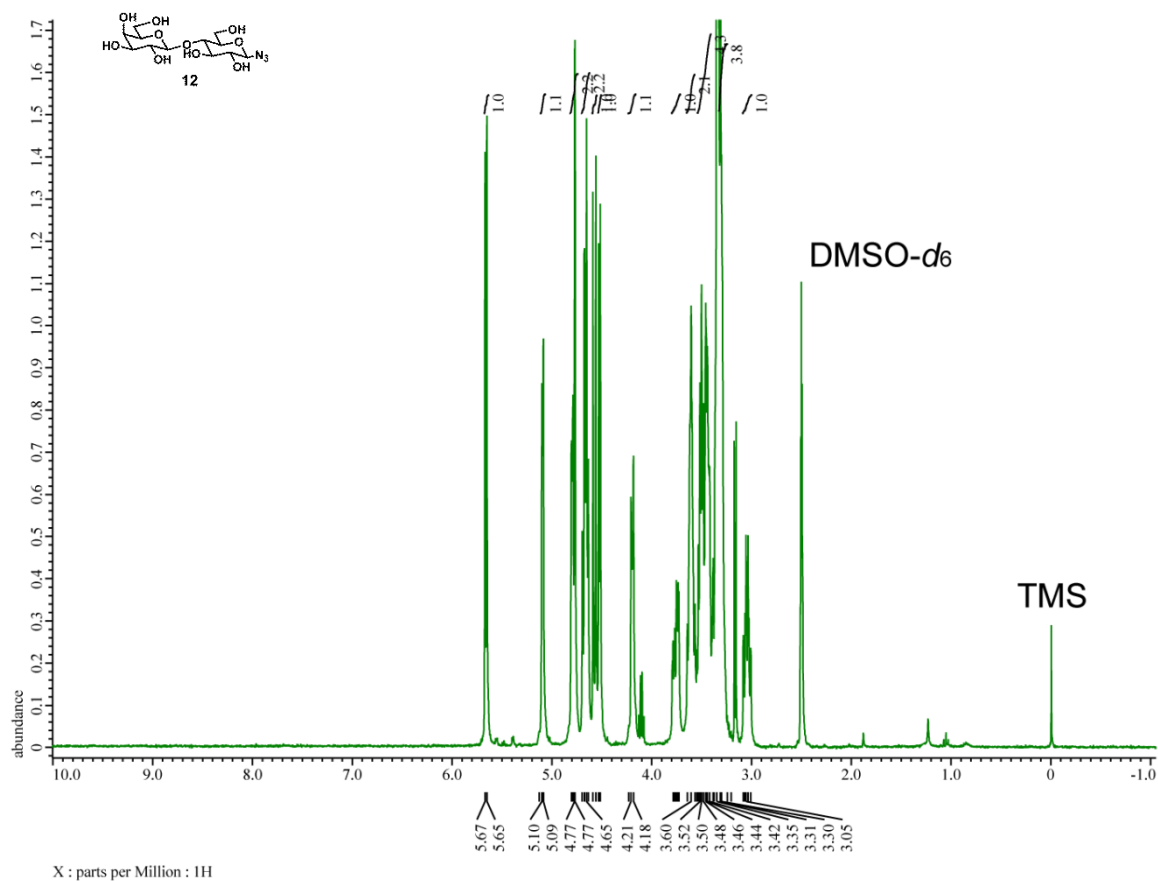


<sup>13</sup>C NMR spectra of compound **13** (CDCl<sub>3</sub>, 75 MHz).

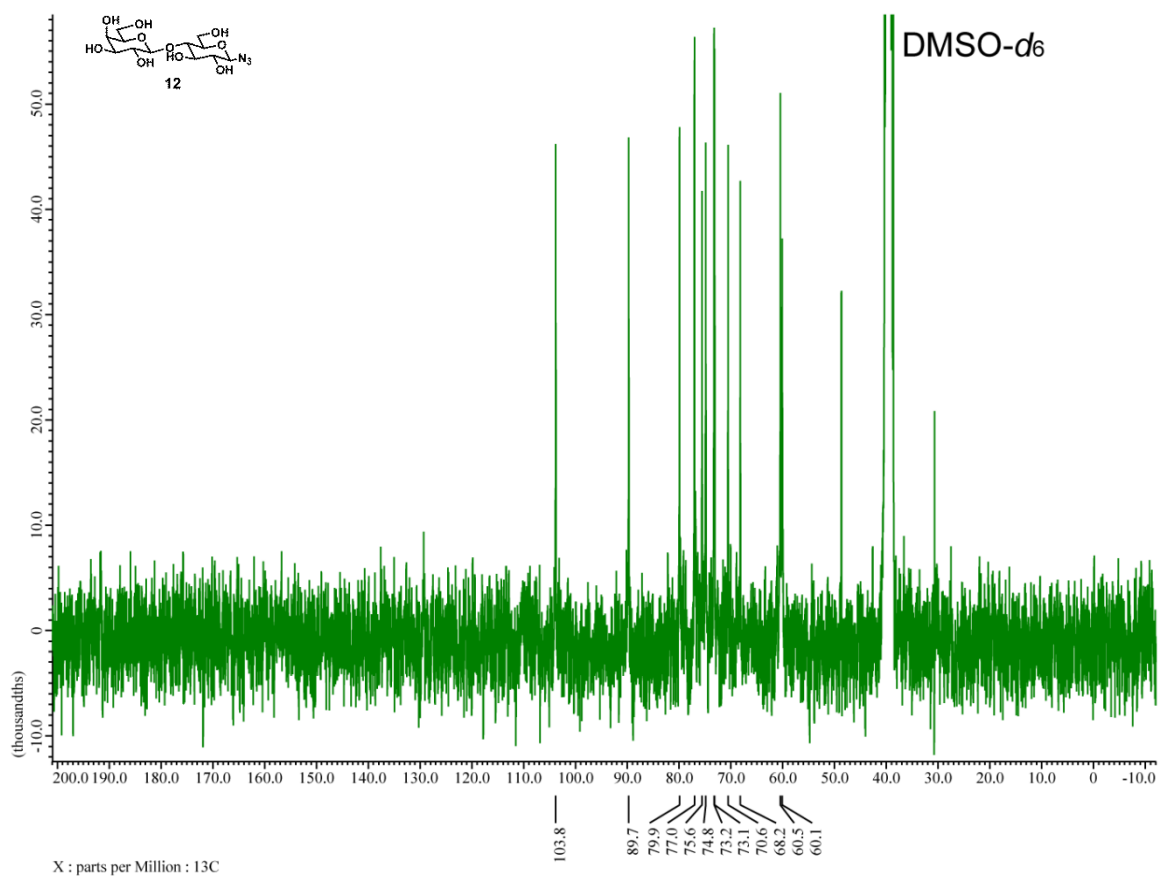




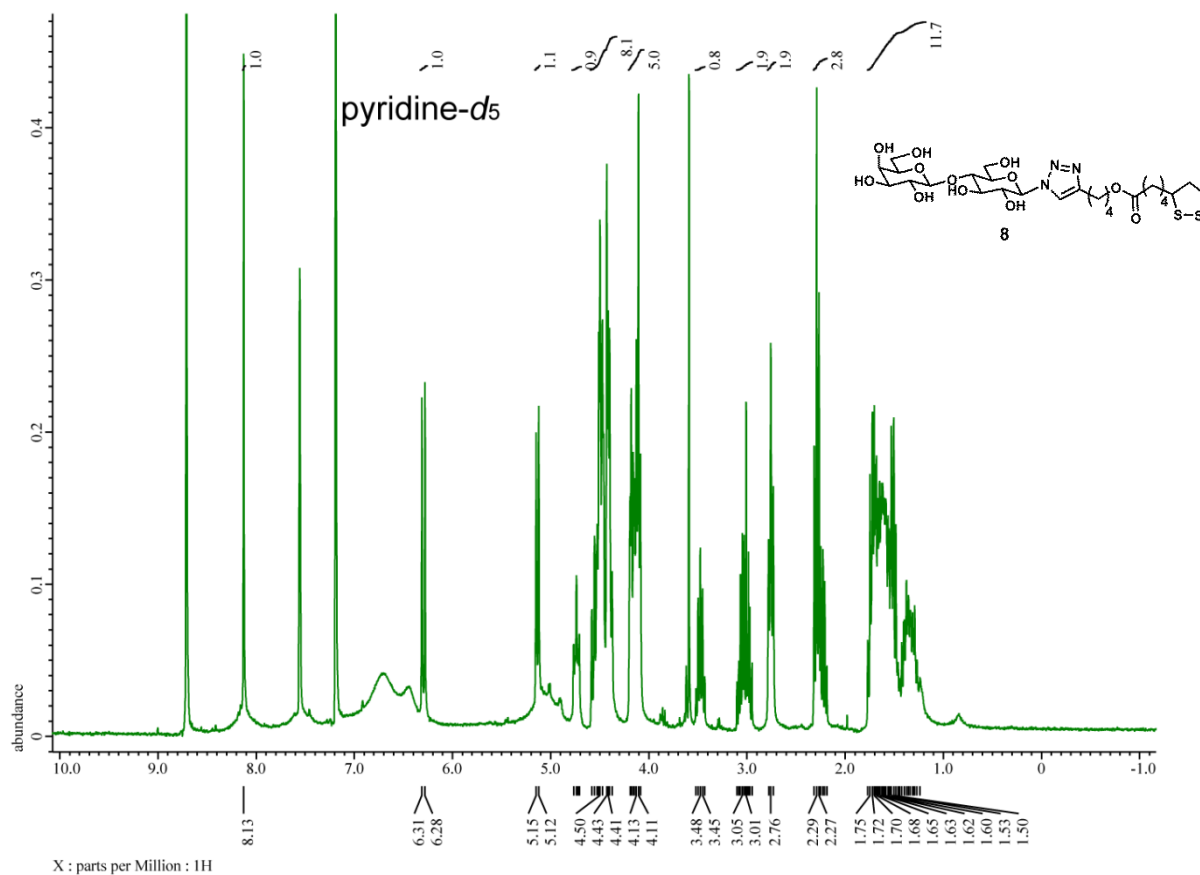
$^1\text{H}$  NMR spectra of compound **12** (DMSO- $d_6$ , 300 MHz).



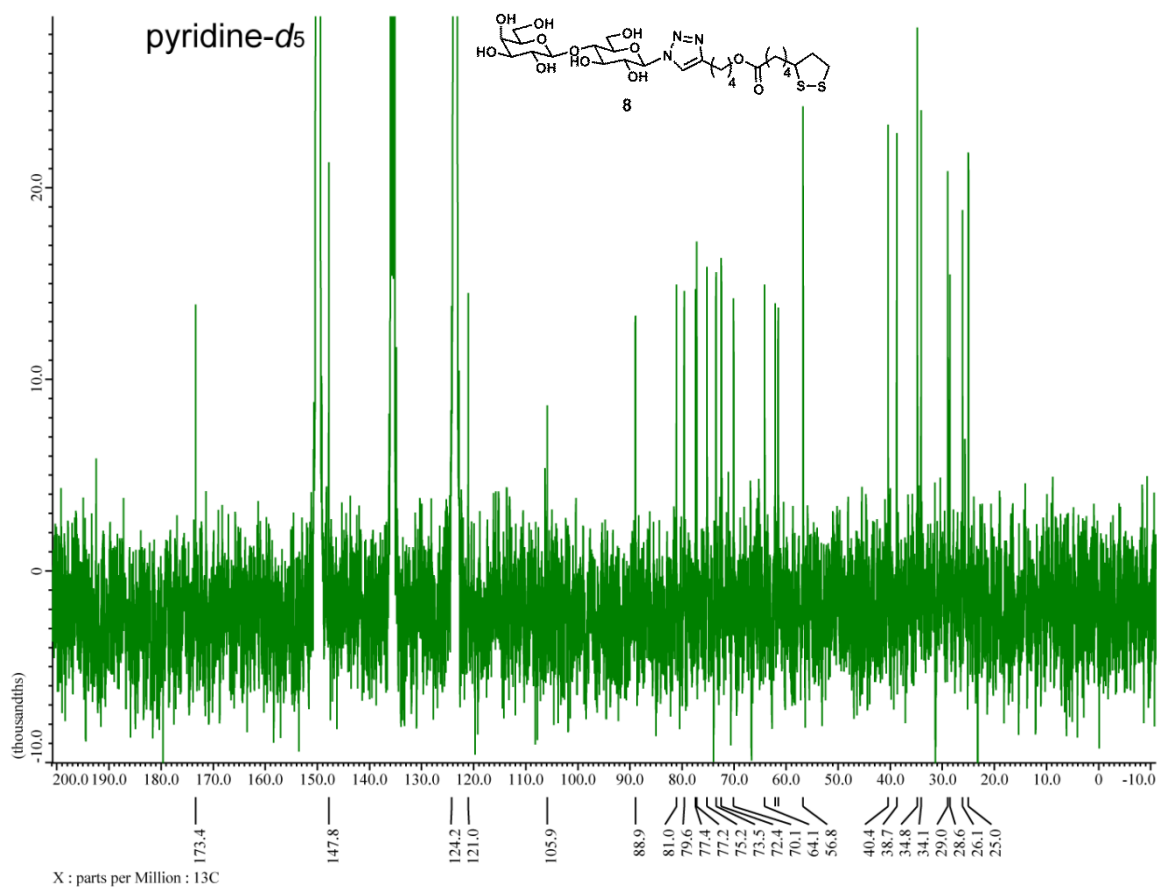
$^{13}\text{C}$  NMR spectra of compound **12** (DMSO- $d_6$ , 75 MHz).



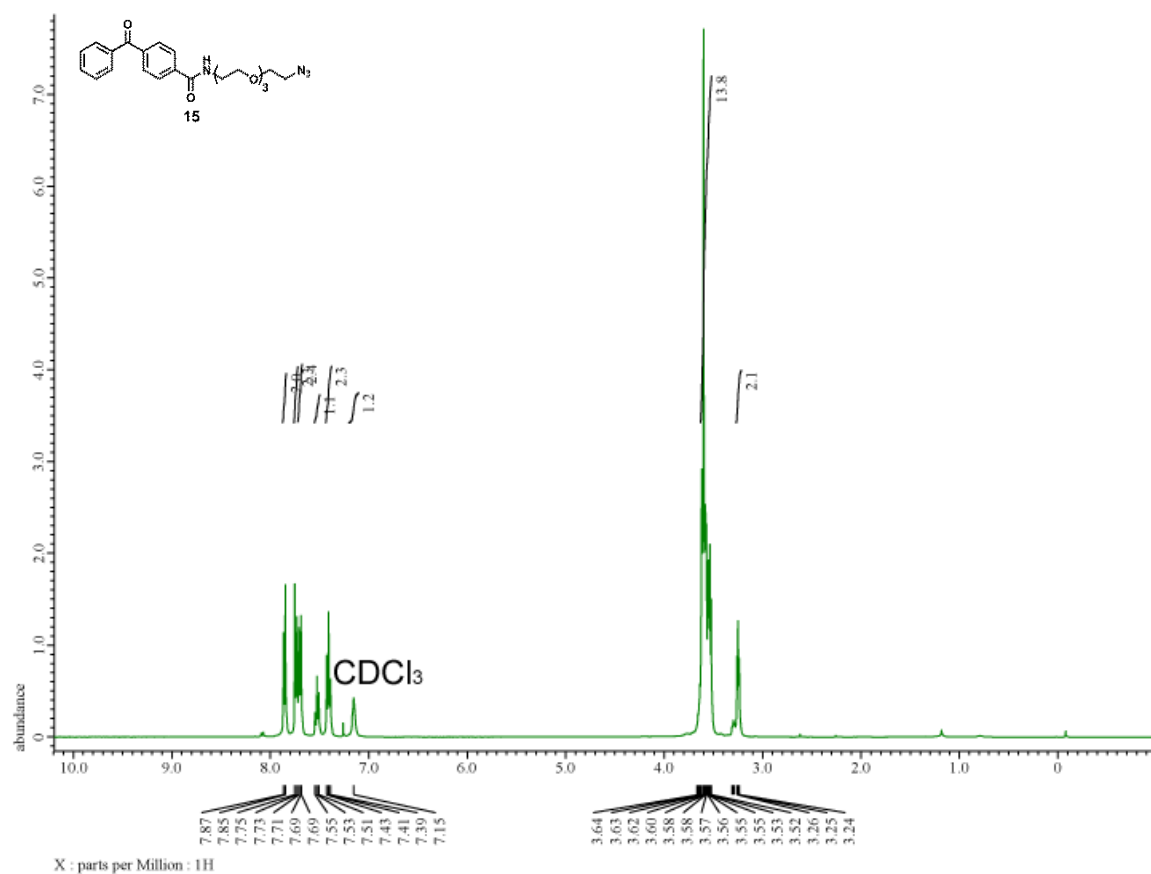
<sup>1</sup>H NMR spectra of compound **8** (pyridine-*d*<sub>5</sub>, 300 MHz).



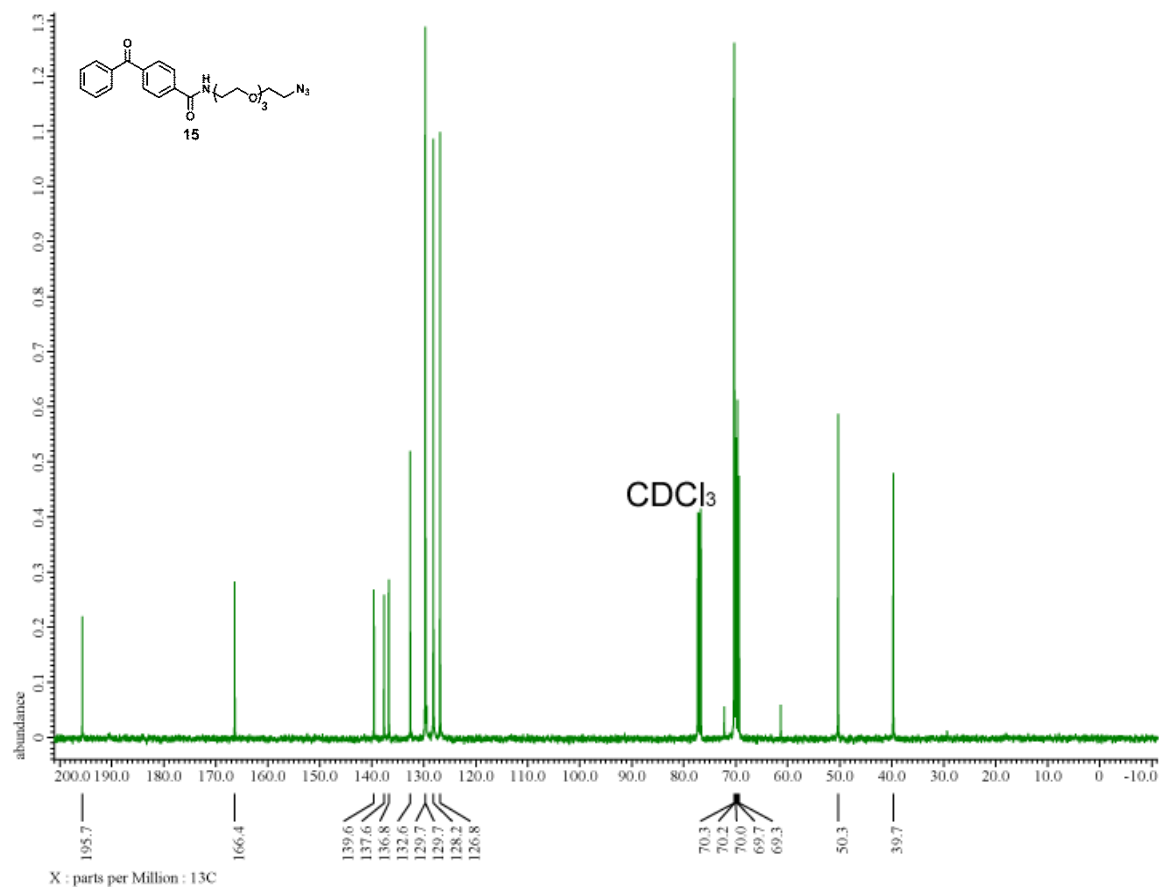
<sup>13</sup>C NMR spectra of compound **8** (pyridine-*d*<sub>5</sub>, 75 MHz).



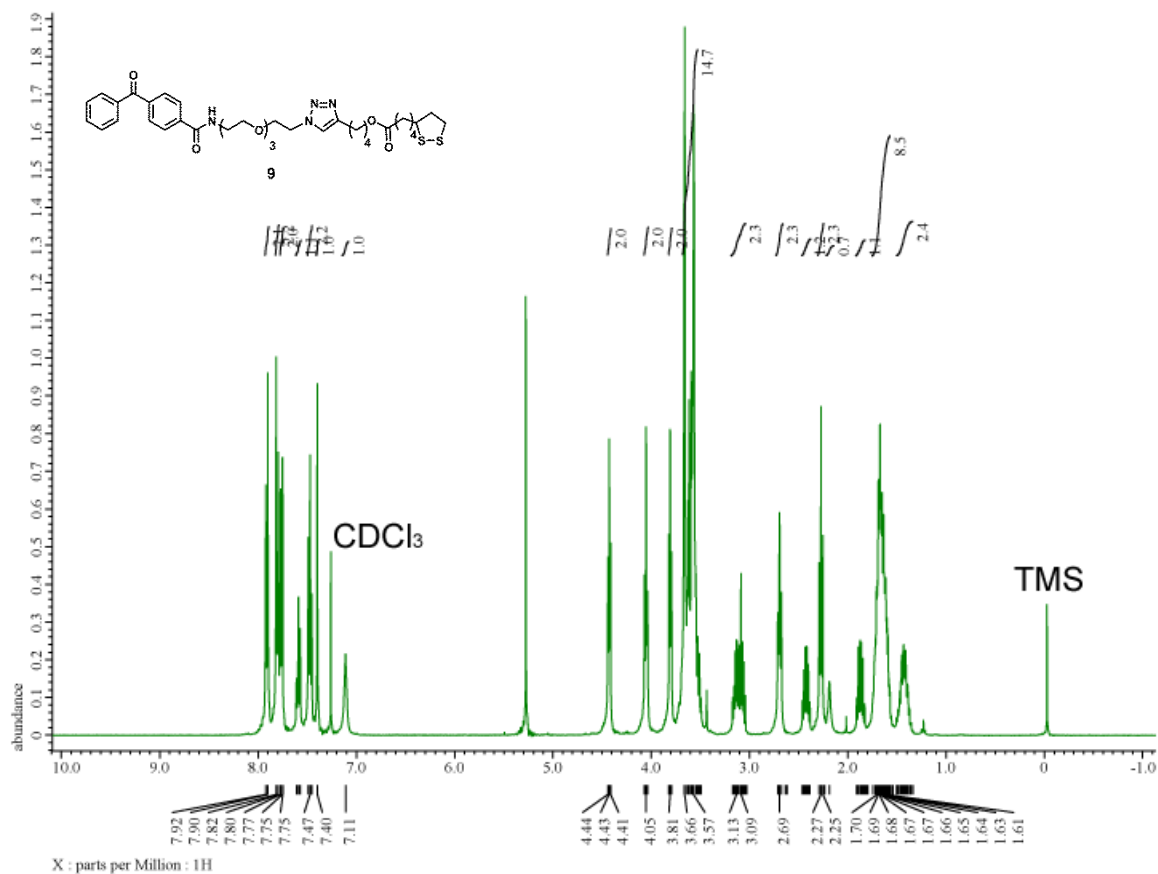
$^1\text{H}$  NMR spectra of compound **15** ( $\text{CDCl}_3$ , 400 MHz).



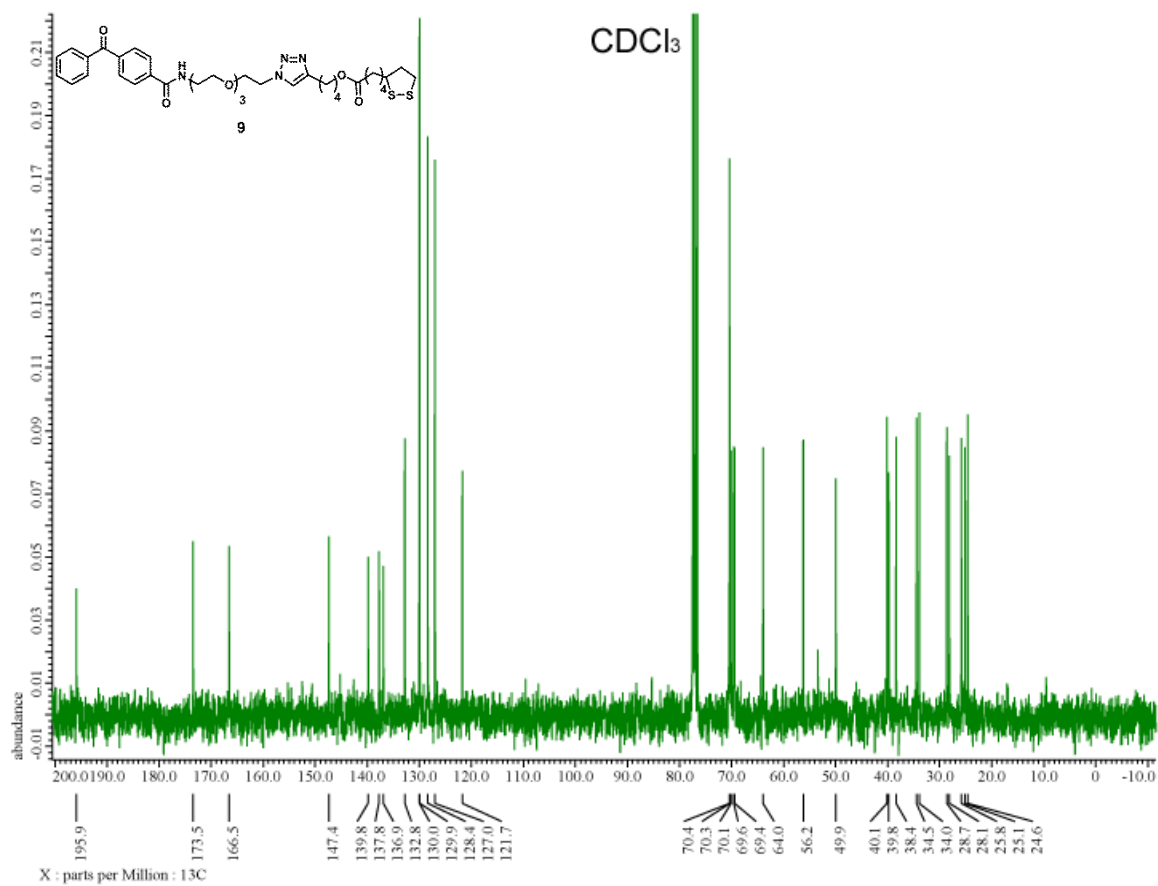
$^{13}\text{C}$  NMR spectra of compound **15** ( $\text{CDCl}_3$ , 100 MHz).



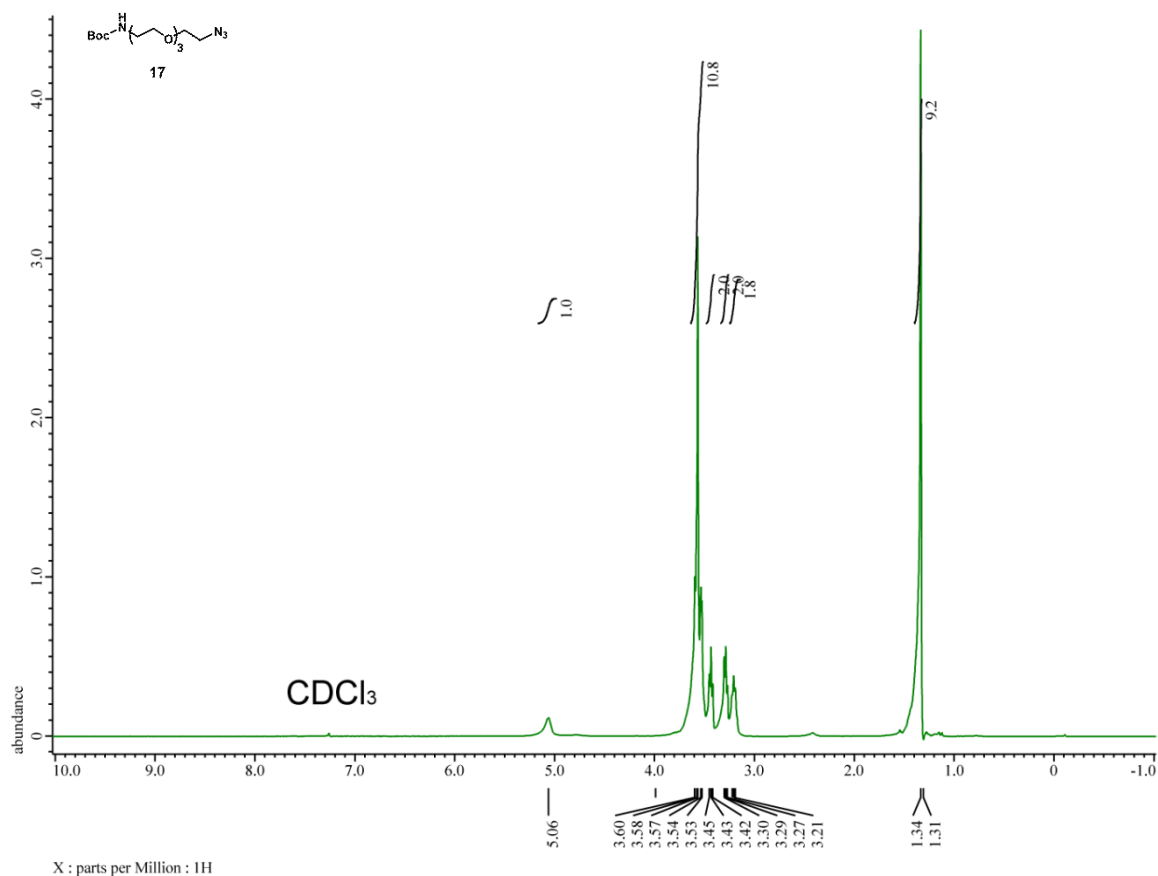
$^1\text{H}$  NMR spectra of compound **9** ( $\text{CDCl}_3$ , 400 MHz).



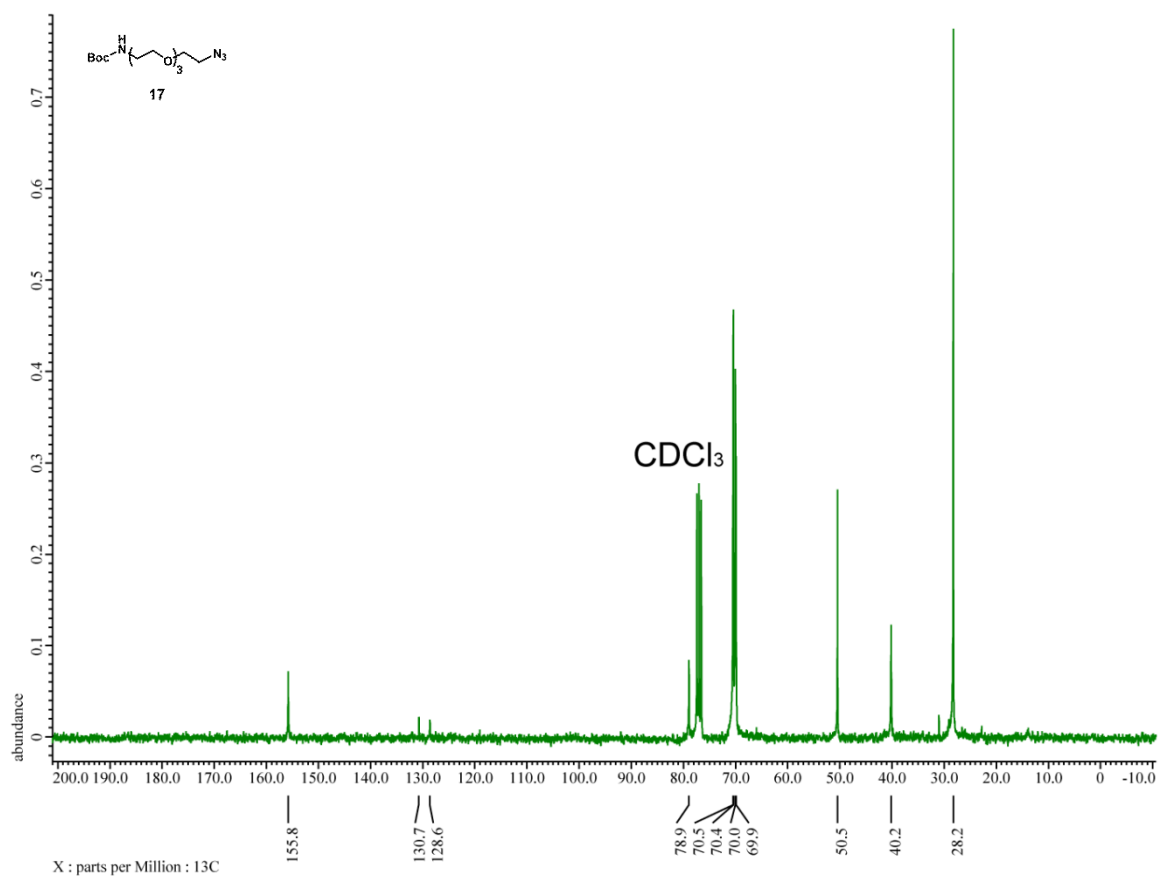
$^{13}\text{C}$  NMR spectra of compound **9** ( $\text{CDCl}_3$ , 100 MHz).



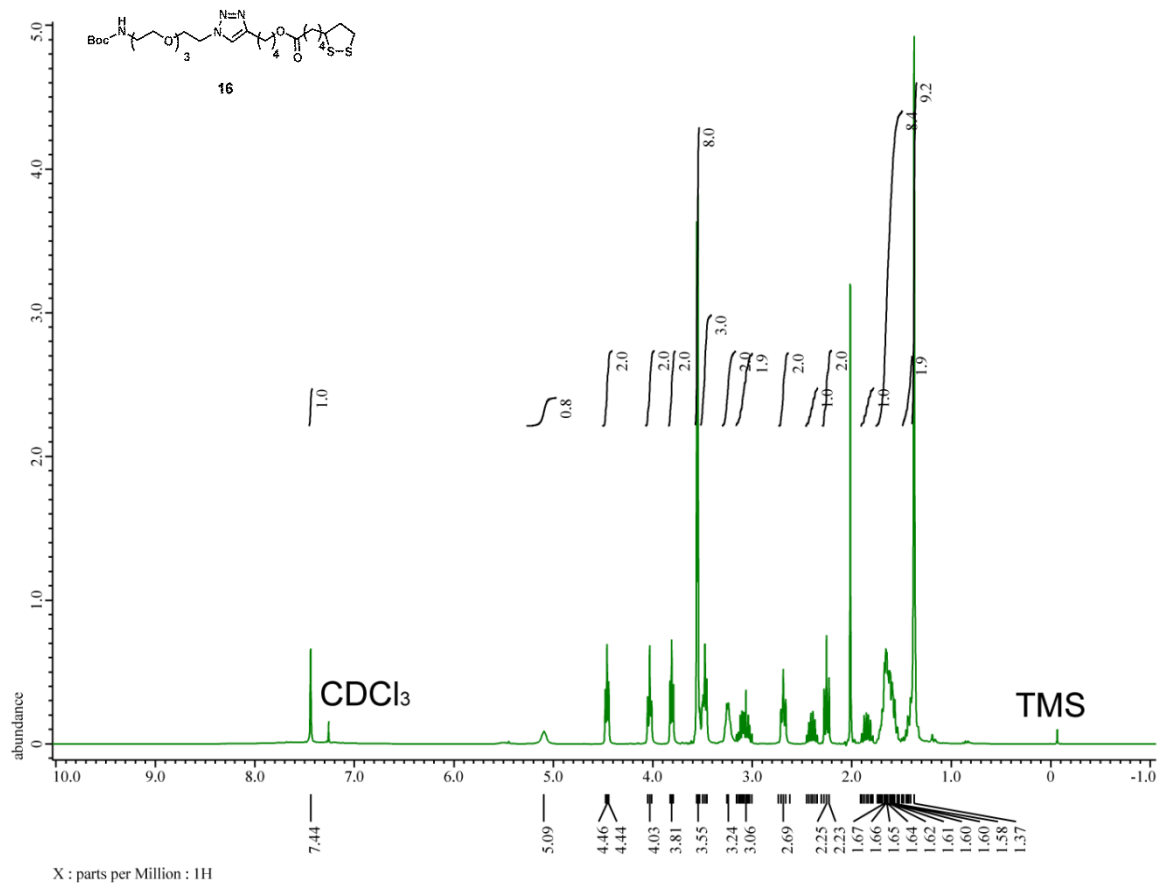
$^1\text{H}$  NMR spectra of compound **17** ( $\text{CDCl}_3$ , 300 MHz).



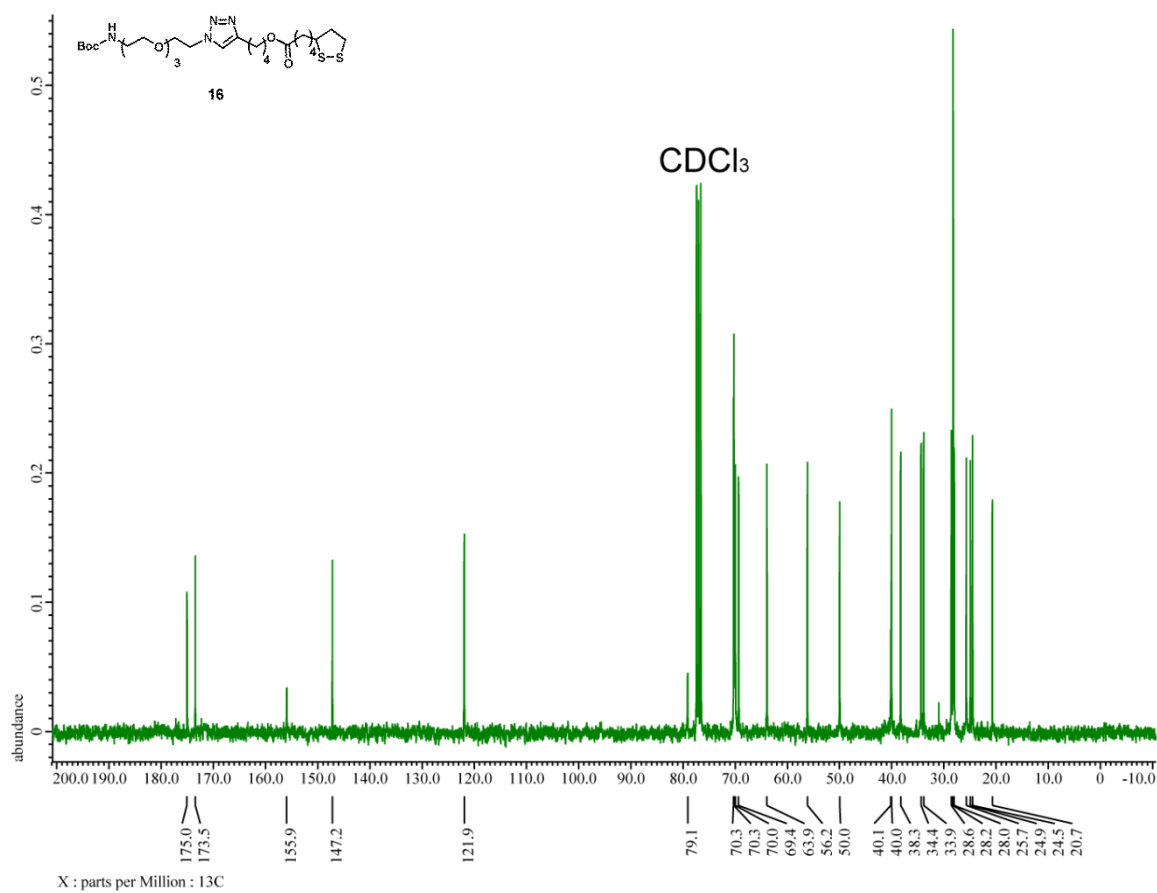
$^{13}\text{C}$  NMR spectra of compound **17** ( $\text{CDCl}_3$ , 75 MHz).



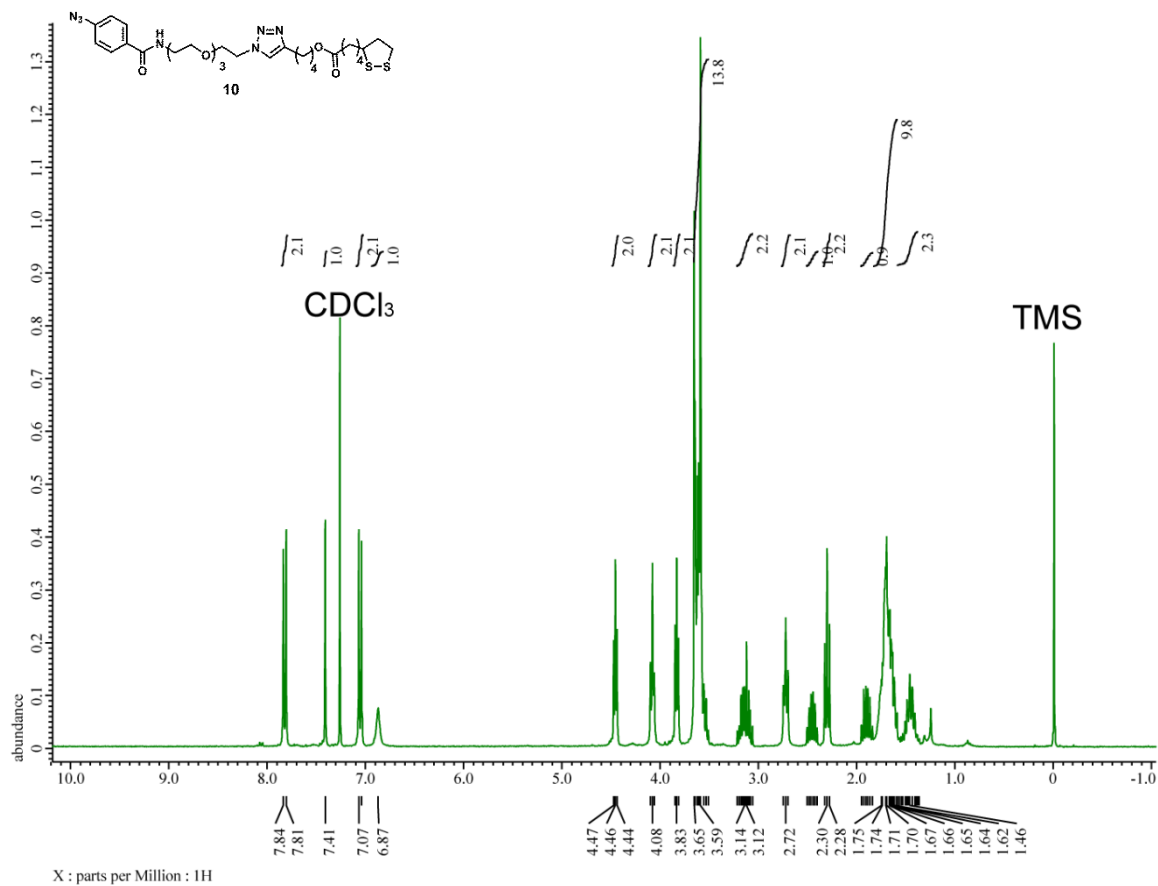
<sup>1</sup>H NMR spectra of compound **16** (CDCl<sub>3</sub>, 300 MHz).



<sup>13</sup>C NMR spectra of compound **16** (CDCl<sub>3</sub>, 75 MHz).



<sup>1</sup>H NMR spectra of compound **10** (CDCl<sub>3</sub>, 300 MHz).



<sup>13</sup>C NMR spectra of compound **10** (CDCl<sub>3</sub>, 75 MHz).

