Supporting Information

For

A One-step Chemoenzymatic Labeling Strategy for Probing Sialylated Thomsen−Friedenreich Antigen

Liuqing Wen,* Ding Liu, Yuan Zheng, Kenneth Huang, Xuefeng Cao, Jing Song, Peng George Wang*

Department of Chemistry, Georgia State University, Atlanta, Georgia, 30303, United States

Table of contents

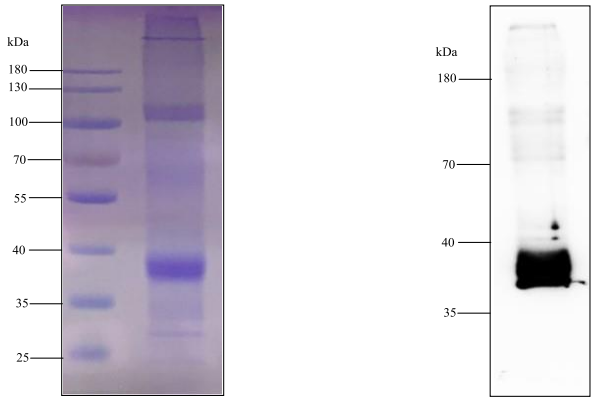


Figure S1. ST6GalNAc-IV was expressed and purified as described in experiment part. The purified ST6GalNAc-IV was resolved by SDS-PAGE and stained with Coomassie brilliant blue (left graph). The purified ST6GalNAc-IV was resolved by SDS-PAGE and the blot was probed using anti-His primary antibody (1:500) and Goat anti-Mouse secondary antibody (1:10000) conjugated with HRP (right graph).

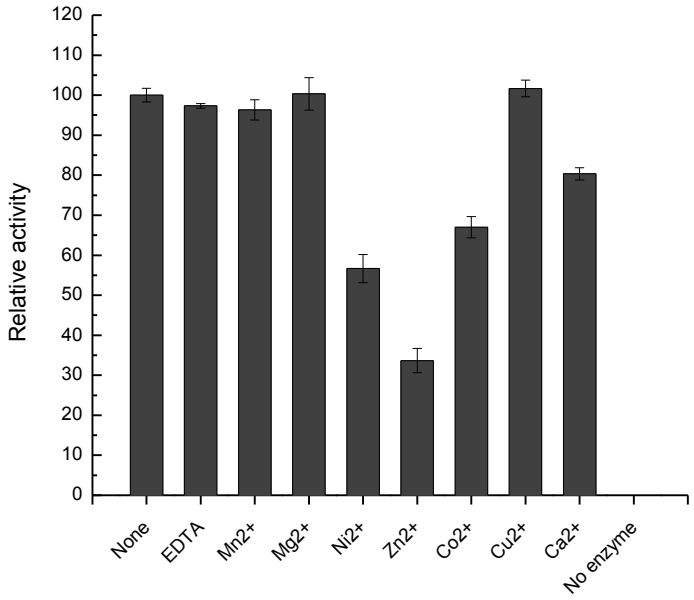


Figure S2. Effects of metal ions and EDTA on the activity of ST6GalNAc-IV. A 20 ul mixture (PBS, pH 7.4) containing 10 mM of Neu5Acα(2-3)Galβ(1-4)GalNAc-α-Bn, 10 mM of CMP-Neu5Ac (**1**. Table 1), 15 ug of ST6GalNAc-IV, with 10 mM of EDTA or 10 mM of metal ions (Mg^{2+}, Mn^{2+}, Ni^{2+}, Zn^{2+}, Co^{+}, Cu^{2+}, or Ca^{2+}) was incubated at 37°C for 1 hour. The reaction without enzyme was performed as a control. The reaction was stopped by diluting the reaction five times with cooled buffer of acetonitrile/100 mM aqueous ammonium acetate pH 4.5 (60% acetonitrile). The reaction was quantified by measuring the newly formed CMP by HPLC.

Figure S3. Chemoenzymatic labeling of cell surface sialyl-T antigen using conventional reporter pair **3** and **12** (long exposure, 40s). Cell surface of sialyl-T (MCF 7 cells) were labeled and biotinylated as described in experiment part. Labeled proteins from cell lysates were detected by Western blotting with streptavidin conjugated with HRP.

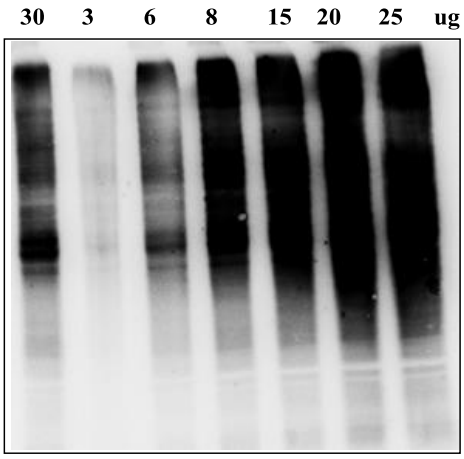


Figure S4. Comparison of reporter pair **4** and **13** (two-step labeling) and probe **10** (one-step labeling) by western blot. Cell surface of sialyl-T (MCF7) were biotinylated as described in experiment part. 30 μg of proteins from two-step labeling group (the most left lane) and 3 to 25 μg of proteins from one-step labeling group were resolved by SDS-PAGE and the blot was probed with Western blotting with streptavidin conjugated with HRP.

Materials and general methods. All cell culture reagents were purchased from Gibco. Streptavidin HRP, goat anti-mouse HRP, anti-his and anti-actin were purchased from Abcam. Alexa Fluor 488-conjugated streptavidin, High Capacity Streptavidin Agarose Resin, BCA Protein Assay Kit, Prestained Protein Ladder, RIPA buffer, and avidin agarose were from Thermo Scientific. BTTES, Biotin-PEG4-alkyne, Biotin-PEG4- picolyl azide were from Click Chemistry Tools (Scottsdale, AZ). Protease inhibitor cocktail was from Millipore. All other chemicals unless otherwise stated were purchased from Sigma without further purification. ^{1}H-NMR and ^{13}C-NMR spectra were recorded on a Bruker 400-MHz NMR spectrometer (D_2O as the solvent). High resolution electrospray ionization (ESI) mass spectra were obtained using LC-MS (Thermo HPLC-Orbitrap Elite) equipping C-18 column. High performance liquid chromatography (HPLC) was performed on a Shimadzu SPD-20A equipped with ultraviolet (UV, 254 nm) detector. The HPLC columns used in this work are ZIC®-cHILIC (Merck, Darmstadt, Germany). The column was eluted at 30°C with acetonitrile/100 mM aqueous ammonium acetate pH 4.5 (65% acetonitrile) at a flow rate of 0.6 ml/min. Gel filtration chromatography was performed using a column (100 cm × 2.5 cm) packed with Bio-Gel P-2 fine resins (Bio-Rad, Hercules, CA). Oligosaccharides that used in this work were synthesized using the enzymes reported previously.[1] T47D and MDA-Mb231 cell lines were kindly provided by Dr. Yuan Liu and Dr. Ritu Aneja from Georgia State University. Other cell lines were from ATCC.

Chemical synthesis. Neu5Ac and D-mannosamine·HCl were purchased from Carbosynth LLC. 9-N_3-Neu5Ac (**S2**) was synthesized from Neu5Ac as described previously.[2] **S3**, **S4**, **S6**, and **S7** were synthesized from D-Mannosamine hydrochloride by typical condensation reaction between carboxylic group and amine group.[3] **S5** and **S8** were synthesized by the method reported previously.[4] **S9** was synthesized from **S5** by the method reported previously.[5]

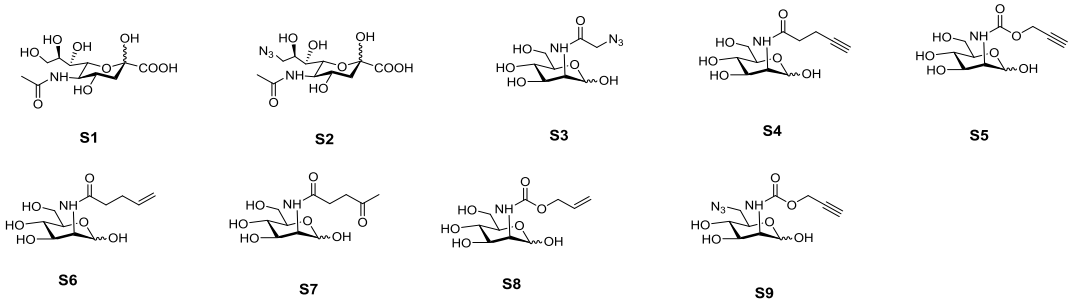
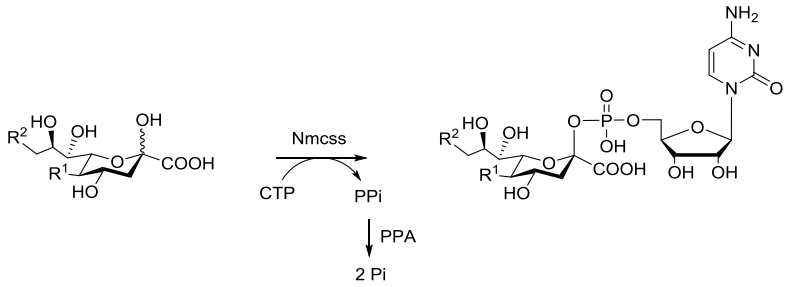


Figure S5. Neu5Ac derivatives (S1 and S2) or ManNAc derivatives (**S3** to **S9**) that used as precursors for the synthesis of CMP-Neu5Ac derivatives



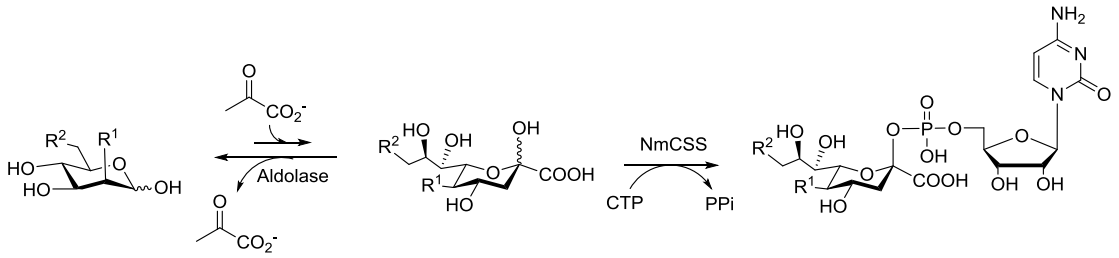
Scheme S1. chemoenzymatic synthesis of **1** and **2** from **a** or **b**

General reaction procedures for the synthesis of CMP-Neu5Ac (**1**) and CMP-9-N_3-Neu5Ac (**2**) from Neu5Ac (**S1**) or 9-N_3-Neu5Ac (**S2**) (Scheme S1).[6] The preparative synthesis of CMP-Neu5Ac analogues was carried in a 40 ml of solution containing 20 mM of Tris-HCl (pH 7.5), 10 mM of **S1** or **S2**, 15 mM of CTP, 5 mM of Mg^{2+} and 5 mg of CMP-sialic acid synthetase from *Neisseria meningitis* (NmCSS). 1 mg of PPA was added to improve the reaction yield by hydrolyzing the newly formed inorganic pyrophosphate. The reaction was carefully shaken at 37°C for 4 hours and monitored by TLC (EtOAc/MeOH/H2O/HOAc=5:2:1.4:0.4). Once reaction finished, product was purified by the method reported previously,[7,8] and was desalted by Bio-Gel P-2 column to afford final product.

Compound 1 (CMP-Neu5Ac). 196 mg; yield, 80%; ^{1}H NMR (400 MHz, D_2O) δ 7.97 (d, *J* = 7.6 Hz, 1H), 6.12 (d, *J* = 7.5 Hz, 1H), 5.99 (d, *J* = 4.4 Hz, 1H), 4.35 (t, *J* = 4.4 Hz, 1H), 4.31 (t, *J* = 4.8 Hz, 1H), 4.24 (d, *J* = 5.0 Hz, 3H), 4.15 (d, *J* = 10.4 Hz, 1H), 4.08 (td, *J* = 10.7, 4.6 Hz, 1H), 4.03 – 3.85 (m, 3H), 3.63 (dd, *J* = 11.7, 6.5 Hz, 1H), 3.46 (d, *J* = 9.6 Hz, 1H), 2.50 (dd, *J* = 13.2, 4.6 Hz, 1H), 2.06 (s, 3H), 1.65 (td, *J* = 12.3, 5.7 Hz, 1H) ; ^{13}C NMR (101 MHz, D_2O) δ 174.70, 174.32, 166.13, 157.76, 141.58, 100.07, 99.99, 96.58, 89.02, 82.94, 82.86, 74.24, 71.74, 69.55, 69.32, 68.80, 66.83, 64.82, 62.91, 51.77,41.02, 22.08. HRMS (ESI): m/z calcd for $C_{20}H_{30}N_4O_{16}P$ $[M-H]^-$ 613.1394; found 613.1403.

Compound 2 (CMP-9-N_3-Neu5AC). 212 mg; yield, 83%; ^{1}H NMR (400 MHz, D_2O) δ 7.95 (d, *J* = 7.5 Hz, 1H), 6.13 (d, *J* = 7.5 Hz, 1H), 6.00 (d, *J* = 4.0 Hz, 1H), 4.40 – 4.30 (m, 2H), 4.26 (d, *J* = 4.0 Hz, 3H), 4.16 (d, *J* = 10.4 Hz, 1H), 4.12 – 4.06 (m, 2H), 3.96 (t, *J* = 10.2 Hz, 1H), 3.66 (d, *J* = 12.1 Hz, 1H), 3.52 (dd, *J* = 15.1, 7.7 Hz, 2H), 2.51 (dd, *J* = 13.1, 4.3 Hz, 1H), 2.08 (s, 3H), 1.71-1.64 (m, 1H); ^{13}C NMR (100 MHz, D_2O) δ 174.73, 174.43, 165.97, 157.56, 141.59, 100.06, 96.58,

89.18, 82.96, 82.89, 74.31, 71.60, 69.43, 69.23, 68.38, 66.81, 65.01, 64.95, 53.15, 51.81, 41.10, 22.15. HRMS (ESI): m/z calcd for $C_{20}H_{29}N7O_{15}P$ $[M-H]^-$ 638.1459; found 638.1474.



Scheme S2. chemoenzymatic synthesis of CMP-Neu5Ac analogues from ManNAc derivatives

General reaction procedures for the synthesis of CMP-Neu5Ac analogues from ManNAc derivatives (Scheme S2).[6] The preparative synthesis of CMP-sialic acid analogues was carried in a 20 ml of solution containing 20 mM of Tris-HCl (pH 7.5), 10 mM of ManNAc derivative, 30 mM of pyruvate, 20 mM of CTP, 5 mM of Mg^{2+} and 5 mg of sialic acid aldolase from *E. coli* K12, 5 mg of CMP-sialic acid synthetase from *Neisseria meningitis* (NmCSS). 1 mg of PPA was added to improve the reaction yield by hydrolyzing the newly formed inorganic pyrophosphate. The reaction was carefully shaken at 37°C for 4 hours and monitored by TLC (EtOAc/MeOH/H2O/HOAc=5:2:1.4:0.4). Once reaction finished, product was purified by the method reported previously,[7,8] and was desalted by Bio-Gel P-2 column to afford final product.

Compound 3 (CMP-Neu5Az). 192 mg (start from 0.4 mmol of **c**); yield, 73%; ^{1}H NMR (400 MHz, D_2O) δ 7.93 (d, *J* = 7.6 Hz, 1H), 6.09 (d, *J* = 7.6 Hz, 1H), 5.96 (d, *J* = 4.4 Hz, 1H), 4.36 – 4.30 (m, 1H), 4.30 – 4.26 (m, 1H), 4.27-4.18 (m, 3H), 4.12 – 4.09 (m, 1H), 4.07 (s, 2H), 4.02 (d, *J* = 10.3 Hz, 1H), 3.95 – 3.89 (m, 1H), 3.88 (d, *J* = 2.4 Hz, 1H), 3.85 (d, *J* = 2.3 Hz, 1H), 3.60 (dd, *J* = 11.8, 6.5 Hz, 1H), 3.43 (d, *J* = 9.6 Hz, 1H), 2.48 (dd, *J* = 13.2, 4.7 Hz, 1H), 1.68-1.60 (m, 1H); ^{13}C NMR (100 MHz, D_2O) δ 174.30, 170.93, 166.02, 157.62, 141.55, 100.03, 96.55, 89.07, 82.89, 74.22, 71.47, 69.59, 69.30, 68.72, 66.57, 64.87, 62.87, 51.90, 41.11. HRMS (ESI): m/z calcd for $C_{20}H_{29}N_7O_{16}P$ $[M-H]^-$ 654.1408; found 654.1416.

Compound 4 (CMP-Neu5AI). 101 mg; yield, 77%; ^{1}H NMR (400 MHz, D_2O) δ 7.96 (d, *J* = 7.5 Hz, 1H), 6.11 (d, *J* = 7.5 Hz, 2H), 5.98 (d, *J* = 4.1 Hz, 1H), 4.32 (dd, *J* = 11.7, 4.3 Hz,24H), 4.24 (d, *J* = 4.5 Hz, 3H), 4.16 (d, *J* = 10.4 Hz, 1H), 4.12-4.06 (m, 1H), 4.00 (d, *J* = 10.2 Hz, 1H), 3.92 (dd, *J* = 23.2, 10.6 Hz, 2H), 3.58 (dd, *J* = 15.9, 8.2 Hz, 2H), 2.55 – 2.45 (m, 5H), 2.40 (s, 1H),

1.69-1.62 (m, 1H); ^{13}C NMR (100 MHz, D_2O) δ 175.29, 174.33, 166.04, 157.64, 141.58, 100.08, 96.57, 89.06, 83.51, 82.83, 74.22, 71.74, 70.40, 69.74, 69.31, 68.91, 66.70, 64.82, 63.05, 51.76, 41.15, 34.74, 14.60. HRMS (ESI): m/z calcd for $C_{23}H_{32}N_4O_{16}P$ $[M-H]^-$ 651.1551; found 651.1568.

Compound 5 (CMP-SiaNProc). 95 mg; yield, 73%; ^{1}H NMR (400 MHz, D_2O) δ 7.94 (d, *J* = 7.6 Hz, 1H), 6.09 (d, *J* = 7.5 Hz, 1H), 5.96 (d, *J* = 4.3 Hz, 1H), 4.67 (d, *J* = 2.0 Hz, 2H), 4.31 (d, *J* = 4.8 Hz, 1H), 4.30 – 4.25 (m, 1H), 4.21 (d, *J* = 5.0 Hz, 3H), 4.13 (d, *J* = 10.5 Hz, 1H), 4.08 – 3.98 (m, 1H), 3.95–3.84 (m, 2H), 3.72–3.65 (m, 1H), 3.65–3.57 (m, 1H), 3.52 (d, *J* = 9.5 Hz, 1H), 2.88 (t, *J* = 2.3 Hz, 1H), 2.46 (dd, *J* = 13.2, 4.6 Hz, 1H), 1.62 (td, *J* = 12.8, 5.7 Hz, 1H); ^{13}C NMR (100 MHz, D_2O) δ 174.35, 165.96, 157.54, 157.45, 141.56, 99.95, 96.53, 89.02, 82.88, 82.81, 78.48, 75.65, 74.21, 71.82, 69.68, 69.25, 68.73, 66.86, 64.82, 62.95, 53.34, 52.87, 41.09. HRMS (ESI): m/z calcd for $C_{22}H_{30}N_4O_{17}P$ $[M-H]^-$ 653.1344; found 653.1360.

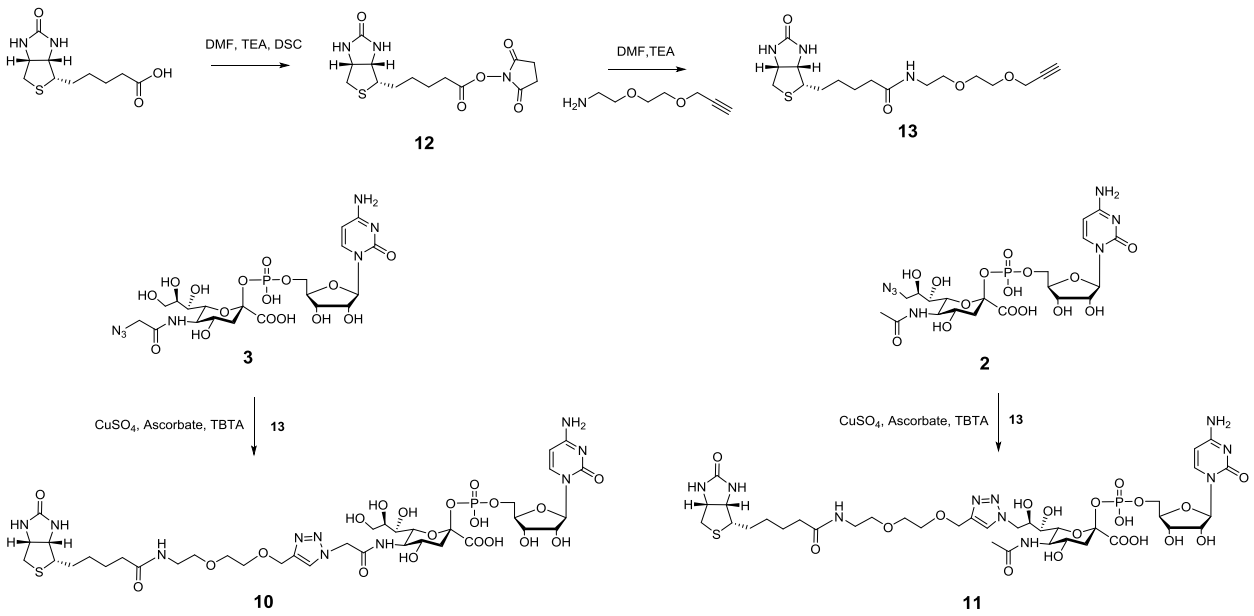
Compound 6 (CMP-SiaNPtl). 99mg; yield, 76%. ^{1}H NMR (400 MHz, D_2O) δ 7.95 (d, *J* = 7.5 Hz, 1H), 6.12 (d, *J* = 7.5 Hz, 1H), 5.99 (d, *J* = 4.1 Hz, 1H), 5.96 – 5.81 (m, 1H), 5.12 (d, *J* = 17.2 Hz, 1H), 5.06 (d, *J* = 10.2 Hz, 1H), 4.34 (dd, *J* = 11.1, 4.4 Hz, 2H), 4.26 (d, *J* = 4.7 Hz, 3H), 4.15 (d, *J* = 10.5 Hz, 1H), 4.09 (td, *J* = 10.9, 4.8 Hz, 1H), 4.01 – 3.87 (m, 3H), 3.61 (dd, *J* = 11.6, 6.8 Hz, 1H), 3.48 (d, *J* = 9.4 Hz, 1H), 2.51 (dd, *J* = 13.2, 4.4 Hz, 1H), 2.46 – 2.36 (m, 4H), 1.67 (td, *J* = 12.5, 5.6 Hz, 1H); ^{13}C NMR (101 MHz, D_2O) δ 176.85, 174.38, 165.97, 157.56, 141.59, 137.14, 115.82, 100.10, 96.59, 89.18, 82.89, 74.23, 71.75, 69.75, 69.34, 68.90, 66.71, 64.90, 63.08, 51.70, 41.21, 35.28, 29.39. HRMS (ESI): m/z calcd for $C_{23}H_{34}N_4O_{16}P$ $[M-H]^-$ 653.1707; found 653.1723.

Compound 7 (CMP-SiaLeV). 113mg; yield, 84%. ^{1}H NMR (400 MHz, D_2O) δ 7.93 (d, *J* = 7.6 Hz, 1H), 6.09 (d, *J* = 7.4 Hz, 1H), 5.96 (d, *J* = 4.1 Hz, 1H), 4.32 (d, *J* = 4.3 Hz, 1H), 4.30 – 4.26 (m, 1H), 4.22 (d, *J* = 4.6 Hz, 3H), 4.13 (d, *J* = 10.5 Hz, 1H), 4.10 – 4.01 (m, 1H), 3.94 – 3.85 (m, 3H), 3.61 (dd, *J* = 11.6, 6.7 Hz, 1H), 3.43 (d, *J* = 9.5 Hz, 1H), 2.86 (dd, *J* = 12.7, 6.3 Hz, 2H), 2.54-2.45 (m, 3H), 2.20 (s, 3H), 1.66-1.54 (m, 1H); ^{13}C NMR (100 MHz, D_2O) δ 213.86, 175.78, 174.37, 165.92, 157.51, 141.55, 99.98, 96.55, 89.10, 82.88, 74.20, 71.76, 69.62, 69.30, 68.77, 66.69, 64.81, 62.97, 51.73, 41.10, 38.17, 29.50, 29.15. HRMS (ESI): m/z calcd for $C_{23}H_{34}N_4O_{17}P$ $[M-H]^-$ 669.1657; found 669.1670.

Compound 8 (CMP-SiaNAloc). 106 mg; yield, 80%. ^{1}H NMR (400 MHz, D_2O) δ 7.92 (d, *J* = 7.4 Hz, 1H), 6.08 (d, *J* = 7.4 Hz, 1H), 6.02 – 5.87 (m, 2H), 5.30 (d, *J* = 17.3 Hz, 1H), 5.23 (d, *J* =

10.5 Hz, 1H), 4.61-4.51 (m, 2H), 4.32 – 4.27 (m, 2H), 4.23 (brs, 3H), 4.14 (d, *J* = 10.4 Hz, 1H), 4.04 (td, *J* = 10.2, 4.5 Hz, 1H), 3.92 – 3.87 (m, 2H), 3.70 – 3.58 (m, 2H), 3.53 (d, *J* = 9.4 Hz, 1H), 2.47 (dd, *J* = 12.8, 3.8 Hz, 1H), 1.64 (td, *J* = 12.2, 5.5 Hz, 1H); ^{13}C NMR (100 MHz, D_2O) δ 174.41, 165.98, 158.27, 157.59, 141.50, 132.75, 117.19, 99.97, 96.55, 89.12, 82.85, 74.20, 71.94, 69.71, 69.28, 68.74, 66.91, 65.86, 64.84, 62.97, 53.22, 41.04. HRMS (ESI): m/z calcd for $C_{22}H_{32}N_4O_{17}P$ $[M-H]^-$ 655.1500; found 655.1515.

Compound 9 (CMP-9-N$_3$-SiaNProc). 87 mg; yield, 64%.^{1}H NMR (400 MHz, D_2O) δ 7.94 (d, *J* = 7.5 Hz, 1H), 6.11 (s, 1H), 5.96 (d, *J* = 4.2 Hz, 1H), 4.71 – 4.64 (m, 2H), 4.31 – 4.22 (m, 2H), 4.22 (s, 3H), 4.13 (d, *J* = 10.5 Hz, 1H), 4.06-4.00 (m, 2H), 3.65 (dd, *J* = 21.9, 11.5 Hz, 2H), 3.55 (d, *J* = 9.5 Hz, 1H), 3.47 (dd, *J* = 13.0, 5.9 Hz, 1H), 2.89 (s, 1H), 2.46 (dd, *J* = 13.1, 4.4 Hz, 1H), 1.62 (td, *J* = 12.4, 5.7 Hz, 1H); ^{13}C NMR (100 MHz, D_2O) δ 174.31, 166.06, 157.42, 141.54, 99.96, 96.14, 89.02, 82.97, 78.48, 75.65, 74.32, 71.68, 69.37, 69.20, 68.44, 66.85, 64.89, 53.34, 53.09, 52.88, 40.98. HRMS (ESI): m/z calcd for $C_{22}H_{29}N_7O_{16}P$ $[M-H]^-$ 678.1408; found 678.1390.



Scheme S3. Chemoenzymatic synthesis of **10** and **11**

Biotin N-hydroxysuccinimide (biotin-NHS); 12. To a solution of 1.96 g of biotin (8 mmol) in 30 ml of DMF was added 5.6 ml of TEA (40 mmol) followed and *N*,*N*'-disuccinimidyl carbonate (DSC, 2.60 g, 9.6 mmol). The reaction mixture was stirred at room temperature for overnight. Reaction solution was concentrated under vacuum, and the residue was washed three times by

diethyl ether to afford rather pure **13** biotin-NHS (2.6 g, 98%). This was used in the next step without further purification.

Biotin-PEG2-alkyne; 13. To a solution of 1023 mg of biotin-NHS (3 mmol) in 20 ml DMF, 500 mg of 2-[2-(Propargyloxy)ethoxy]ethylamine (3.5 mmol), 2.8 ml of TEA (20 mmol) were added. The reaction mixture was stirred at room temperature for three hours and monitor by TLC. Once reaction finished, DMF was removed under vacuum, and the residue was purified by flash chromatography on silica gel, which eluted with MeOH: EA gradually from 1:50 to 1:4 to give the product **2** as a white solid (0.96 g, 90% yield). ^{1}H NMR (400 MHz, $CDCl_3$) δ 6.94 (s, 1H), 6.84 (s, 1H), 6.13 (s, 1H), 4.47 – 4.44(m, 1H), 4.27 – 4.24 (m, 1H), 4.14 (d, *J* = 1.8 Hz, 2H), 3.61 (dd, *J* = 14.4, 5.0 Hz, 4H), 3.52 (t, *J* = 4.9 Hz, 2H), 3.39– 3.37 (m, 2H), 3.08 (dd, *J* = 11.6, 7.0 Hz, 1H), 2.84 (dd, *J* = 12.8, 4.7 Hz, 1H), 2.68 (d, *J* = 12.8 Hz, 1H), 2.46 (s, 1H), 2.17 (t, *J* = 7.4 Hz, 2H), 1.71-1.57 (m, 4H), 1.38 (dd, *J* = 14.8, 7.4 Hz, 2H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 173.65, 164.57, 79.41, 75.02, 69.95, 69.84, 68.94, 61.83, 60.29, 58.35, 55.79, 40.56, 39.09, 36.00, 28.36, 28.08, 25.66.

Compound 10 (CMP-5-biotin-PEG2-triazole-Sialic acid). To a solution of 98 mg of **3** (0.15 mmol) in 10 ml of water, 10 ml of methanol containing 74 mg of **13** (0.2 mmol) was added. Then, 5.3 mg of TBTA, 2.5 mg of Copper(II) sulfate pentahydrate, and 3.5 mg of Ascorbic acid were added. The reaction mixture was stirred at room temperature for one hour and monitor by TLC. Once reaction finished, the solvent was removed under vacuum and the residue was re-dissolved in water. The concentrated solution was desalted by Bio-Gel P-2 column to afford 124 mg of white solid (81% yield). ^{1}H NMR (400 MHz, D_2O) δ 8.00 (s, 1H), 7.85 (d, *J* = 7.5 Hz, 1H), 6.00 (d, *J* = 7.5 Hz, 1H), 5.87 (d, *J* = 4.3 Hz, 1H), 5.25 (s, 2H), 4.60 (s, 2H), 4.47 (dd, *J* = 7.6, 5.0 Hz, 1H), 4.29 – 4.23 (m, 2H), 4.19 (t, *J* = 4.7 Hz, 1H), 4.14 -4.11 (m, 4H), 4.04 (td, *J* = 10.7, 4.8 Hz, 1H), 3.91 (t, *J* = 10.3 Hz, 1H), 3.87 – 3.81 (m, 1H), 3.78 (d, *J* = 11.6 Hz, 1H), 3.60 (dd, *J* = 15.3, 4.9 Hz, 5H), 3.51 – 3.48 (m, 2H), 3.38 (d, *J* = 9.6 Hz, 1H), 3.27 (t, *J* = 5.0 Hz, 2H), 3.20 – 3.15 (m, 1H), 2.85 (dd, *J* = 13.0, 4.8 Hz, 1H), 2.65 (d, *J* = 13.0 Hz, 1H), 2.40 (dd, *J* = 13.2, 4.4 Hz, 1H), 2.14 (t, *J* = 7.1 Hz, 2H), 1.58 – 1.39 (m, 5H), 1.27 (dd, *J* = 14.7, 7.2 Hz, 2H); ^{13}C NMR (100 MHz, D_2O) δ 176.81, 174.27, 168.18, 166.03, 165.21, 157.63, 143.96, 141.54, 126.50, 99.99, 99.91, 96.54, 89.09, 82.88, 82.80, 74.23, 71.52, 69.58, 69.36, 69.30, 69.00, 68.85, 68.75, 66.67, 64.87, 63.01, 62.87, 62.02, 60.20, 55.35, 52.18, 40.99, 39.73, 38.86, 35.43, 27.85, 27.65, 25.13. HRMS (ESI): m/z calcd for $C_{37}H_{56}N_{10}O_{20}PS$ $[M-H]^-$ 1023.3131; found 1023.3158.

Compound 11 (CMP-9-biotin-PEG2- triazole-Sia5NAc). To a solution of 96 mg of **2** (0.15 mmol) in 10 ml of water, 10 ml of methanol containing 74 mg of **13** (0.2 mmol) was added. Then, 5.3 mg of TBTA, 2.5 mg of Copper(II) sulfate pentahydrate, and 3.5 mg of Ascorbic acid were added. The reaction mixture was stirred at room temperature for one hour and monitor by TLC. Once reaction finished, the solvent was removed under vacuum and re-dissolved in water. The product was purified by Bio-Gel P-2 column to afford 119 mg of white solid (79% yield). ^{1}H NMR (400 MHz, D_2O) δ 8.06 (s, 1H), 7.85 (d, *J* = 7.5 Hz, 1H), 6.01 (d, *J* = 6.8 Hz, 1H), 5.95 (d, *J* = 4.5 Hz, 1H), 4.64 (s, 2H), 4.55 (dd, *J* = 7.6, 4.9 Hz, 1H), 4.46 (dd, *J* = 14.3, 8.1 Hz, 1H), 4.35 (dd, *J* = 7.7, 4.5 Hz, 1H), 4.27 – 4.03 (m, 9H), 3.94 (t, *J* = 10.3 Hz, 1H), 3.70-3.65 (m, 4H), 3.57 (t, *J* = 5.1 Hz, 2H), 3.34 (d, *J* = 7.1 Hz, 3H), 3.27 – 3.22 (m, 1H), 2.92 (dd, *J* = 13.0, 4.8 Hz, 1H), 2.72 (d, *J* = 13.0 Hz, 1H), 2.47 (dd, *J* = 13.1, 4.4 Hz, 1H), 2.21 (t, *J* = 7.1 Hz, 2H), 2.01 (s, 3H), 1.68 – 1.46 (m, 5H), 1.35 (dd, *J* = 14.7, 7.2 Hz, 2H); ^{13}C NMR (101 MHz, D_2O) δ 176.78, 174.67, 174.23, 165.86, 165.18, 157.49, 143.59, 141.47, 125.75, 99.99, 96.48, 88.81, 83.04, 74.32, 71.57, 69.82, 69.51, 69.34, 68.88, 68.32, 66.65, 65.05, 63.01, 62.01, 60.19, 55.35, 52.95, 51.75, 41.03, 39.73, 38.86, 35.42, 27.85, 27.66, 25.13, 22.07. HRMS (ESI): m/z calcd for $C_{37}H_{56}N_{10}O_{19}PS$ $[M-H]^-$ 1007.3182; found 1007.3207.

Expression of ST6GalNAc-IV in Sf9 cells. His-tagged soluble human ST6GalNAc-IV was produced in insect cells.[9] The 100 μL of P1PP1 baculovirus stock and Sf9 cells for the expression of the ST6GalNAc-IV were kindly provided by Dr. Donald Jarvis from University of Wyoming. Sf9 cells were cultured in serum-free medium (SF900 II, Invitrogen) at 27°C, 125 rpm. P2 virus stock was obtained by infecting 25 mL of Sf9 cells at a concentration of 2×10^6 C/mL with 25 μL of P1PP1 baculovirus stock. After about 72 hours infection (10%~15% death rate), the cells were removed by centrifugation. The supernatant was collected as P2 virus stock with an estimate of virus titer 5×10^7. P3 virus stock was obtained by infecting 100 mL of Sf9 cells at a concentration of 2×10^6 C/mL with 12 ml of P2 virus stock with an estimated MOI of approximately 3 pfu/cell. After about 72 hours infection (10%~20% death rate), the cells were removed by centrifugation. The supernatant was collected as P3 virus stock with an estimate of virus titer 5×10^8. P3 virus was directly used for large scale protein expression. Sf9 cells were scaled up to 2 liters with a concentration of 3×10^6 C/mL. Then, 36 ml of P3 virus was added with an estimated MOI=3. The infected cells were cultured in another 72 hours to allow the expression of ST6GalNAc-IV. ST6GalNAc-IV was expressed in both inside of cells and medium supernatant. The protein was

purified by Ni-NTA agarose column. Then, the purified proteins were concentrated and desalted with 10 kDa molecular weight cut-off (Millipore, MWCO) spin filters for further use. Protein concentration was determined by Pierce BCA Protein Assay Kit (Invitrogen).

Donor specificity study of ST6GalNAc-IV with CMP-Neu5Ac analogues. A 20 ul mixture (PBS, pH 7.4) containing 10 mM of Neu5Acα(2-3)Galβ(1-4)GalNAc-α-Bn, 10 mM of CMP-Neu5Ac analogues (Entry 1 to 12. Table 1), 15 ug of ST6GalNAc-IV, was incubated at 37°C for 1 hour. The reaction was stopped by diluting the reaction five times with cooled buffer of acetonitrile/100 mM aqueous ammonium acetate pH 4.5 (60% acetonitrile). The diluted solution was analyzed by HPLC equipped with UV detector at 254 nm using ZIC®-cHILIC column. The column was eluted at 30°C with acetonitrile/100 mM aqueous ammonium acetate pH 4.5 (65% acetonitrile) at a flow rate of 0.6 ml/min. The enzyme activity was valued by the newly formed CMP. The relative activity data was shown in Figure 1.

Chemoenzymatic synthesis of sialyl-T-related oligosaccharides for acceptor study of ST6GalNAc-IV. Entries 1 to 4, entry 6 and entry 7 were synthesized from GalNAc-α-Bn (entry 5) by stepwise assembly strategy employing bacterial α-2,3-sialyltransferase (PmsT1),[10] α-2,6-sialyltransferase(Pd26T),[11] and D-galactosyl-β1–3-*N*-acetyl-D-hexosamine phosphorylase (BiGalHexNAcP).[12] Entry 8, 9 10, and 11 were prepared as reported previously.[1]

Acceptor specificity study of ST6GalNAc-IV with compound 10 and representative structures. A 20 ul mixture (PBS, pH 7.4) containing 10 mM of acceptor (Entries 1 to 12, Table 1), 10 mM of **10**, and 15 ug of ST6GalNAc-IV was incubated at 37°C for 30 min. The reaction was stopped by diluting the reaction five times with cooled buffer of acetonitrile/100 mM aqueous ammonium acetate pH 4.5 (60% acetonitrile). The reaction was analyzed by HPLC equipped with UV detector at 254 nm using ZIC®-cHILIC column. The column was eluted at 30°C with acetonitrile/100 mM aqueous ammonium acetate pH 4.5 (65% acetonitrile) at a flow rate of 0.6 ml/min. The relative activity data was shown in Table 1.

Cell Culture. HT29, MDA-mb-231, Raw 264.7, T47D cells grown in DMEM medium supplemented with 10% fetal bovine serum (FBS), 100 units/mL penicillin, and 0.1 mg/mL streptomycin (Gibco). MCF 7 and HEK293 grown in MEM medium supplemented with 10% fetal bovine serum (FBS), 100 units/mL penicillin, and 0.1 mg/mL streptomycin (Gibco). In all cases, cells were incubated in a 5% CO2 humidified chamber at 37 °C.

Detection cell surface sialyl-T antigen by western blot. Cells were cultured as described above and harvested by centrifugation (1000 g, 5 min). After wash three times using PBS, Cells were labeled in suspension at 37°C for 1 hour in labeling buffer containing ST6GalNAc-IV (100 μg/ml) and probe (**3**, **4** or **10**, 200 μM), while the reactions with the absence of ST6GalNAc-IV were performed parallel as controls. The labeling reaction was stopped by two washes with PBS (centrifugation, 1000g, 5 min). The cells were further disrupted by RIPA buffer supplemented with protease inhibitor cocktail. Cell debris was removed by centrifugation (12000 g, 15 min), and the supernatant solution was diluted to 1 mg/ml using water. To perform biotinylation (probe **10** without this step), 500 μM of $CuSO_4$, 2 mM of BTTES, 2 mM of sodium ascorbate and 100 uM of probe **12** or **13** were added to biotinylate probe **3** or **4** group, respectively . After the reaction was carried at room temperature for 1 hour, the proteins were precipitated using methanol/chloroform/water as described above and washed twice using MeOH. The proteins were redissolved in 2% SDS. About 30 ug of protein was loaded for SDS-PAGE separation and western blot analysis.

Western Blotting. The purified, labeled samples from above was separated by SDS-PAGE and transferred to a polyvinylidene difluoride (PVDF) membrane (Millipore). The membrane was blocked in 3% BSA in TBST (50 mM Tris–HCl, 150 mM NaCl, 0.05% Tween 20, pH 7.4) at room temperature for 2 hours. The membrane was incubated streptavidin-linked horseradish peroxidase (1:20000) at 4°C overnight or at room temperature for 2 hours. The membrane was washed three times with TBST for 10 min, and the blots were developed using ECL reagents and ImageQuant LAS 4000 mini imager (GE Healthcare). For Anti-Actin or Anti-His western blot analysis, the membrane was incubated with Anti-Actin or Anti-His in TBST containing 3% BSA at 4°C overnight. The membrane was washed three times and incubated with second antibody (Goat anti-mouse, 1: 5000) as room temperature for 1hour. The membrane was washed three times with TBST for 10 min, and the blots were developed using ECL reagents and ImageQuant LAS 4000 mini imager (GE Healthcare).

Chemoenzymatic imaging cell surface sialyl-T antigen by Fluorescence microscopy. MCF7 cells were cultured 24 hours to allow adhesion. Monolayers were washed three times using PBS containing 3% FBS. Then, cells were enzymatically labeled at 37°C for 1 hour in labeling buffer containing ST6GalNAc-IV (100 μg/ml) and probe (**3**, **4** or **10**, 200 μM). Group that use probe **10** was washed three time for next step. After three washes with cooled PBS, groups that used probe

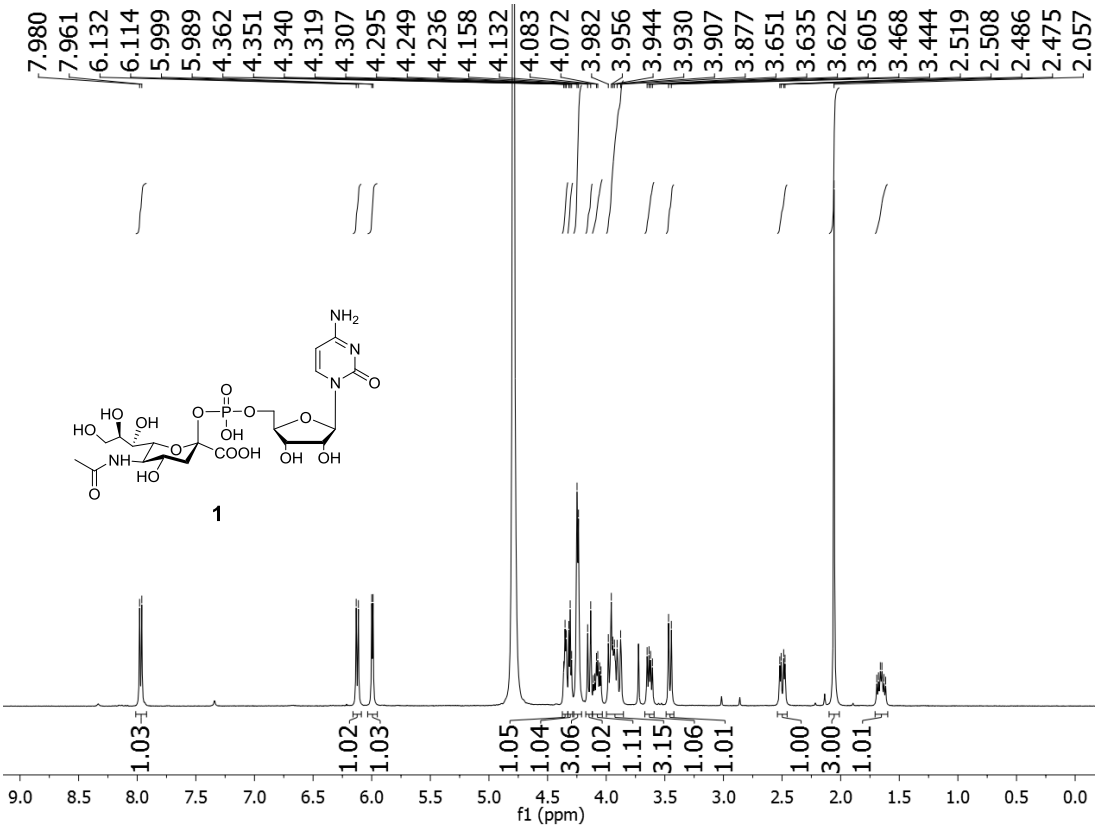
3 or **4** were biotinylated by CuAAc (50 μM **12** or **13**, 100 μM $CuSO_4$, 1 mM of BTTES, and 1 mM of sodium ascorbate) with probe **12** or **13**, respectively at RT for 10 min. The reaction was stopped by three washes with PBS. Next, the cells incubated with streptavidin–Alexa Fluor 488 (10 ug/ml) in cooled PBS at 4 °C for 30 min in dark. Cells were washed twice with PBS and fixed with formaldehyde (4 % in PBS) at RT for 15 min. The nucleus was labeled with DAPI before imaging by fluorescence microscope.

Detection of cell surface sialyl-T antigen by Flow Cytometry. Cells were labeled in suspension at 37°C for 1 hour in labeling buffer containing ST6GalNAc-IV (100 μg/ml) and probe (**3**, **4** or **10**, 200 μM). The labeling reaction was stopped by three washes with PBS (centrifugation, 1000g, 5 min). Labeled cells that used of probe **10** was directly used for next staining step. Labeled cells that used probe **3** and **4** were biotinylated by CuAAc (50 μM **12** or **13**, 100 μM $CuSO_4$, 1 mM of BTTES, and 1 mM of sodium ascorbate) with probe **12** or **13**, respectively at RT for 10 min. The reaction was stopped by three washes with PBS (centrifugation, 1000g, 5 min). Next, the cells were incubated with streptavidin–Alexa Fluor 488 (10 ug/ml) in cooled PBS at 4 °C for 30 min in dark. After three washes with PBS (centrifugation, 1000g, 5 min), the cell was resuspended in PBS for flow cytometry analysis. For each experiment, 10,000 live cells were analyzed, and data analysis was performed on FlowJo.

LC-MS/MS Proteomic Analysis.

Cell surface proteins from MCF7 and HT29 were labeled as described above. the cells were disrupted by RIPA buffer supplemented with protease inhibitor cocktail. The proteins were extracted by MeOH: $CHCl_3$: H_2O (4:6:4.5) and washed three times by MeOH. This process can remove lipids and other impurities. To capture biotinylated proteins by avidin beads, the protein pellet was resuspended in 2 mL of resuspend buffer (4%SDS + 2M urea in 50mM ABC pH=7.5). 250 μl of avidin beads (Thermo Sciecntific, USA) were washed three times with PBS (1 mL). Beads were first pre-blocked with 5 mg/ml BSA in the resuspend buffer at rt for 3 h, washed three times with PBS (1 mL), and incubated with the labeled proteins at rt for 3 h. The beads were washed successively with 2% SDS in PBS, 8 M urea with 250 mM ammonium bicarbonate (ABC), PBS, and 0.05 M ABC. Samples were next added with 40 μL of 4× loading buffer, heated for 10 min at 95 °C, and resolved in 10% SDS-PAGE. Each lane of the SDS-PAGE gel was sliced into $1\times1mm^2$ fractions, and each excised gel slice was placed in a microcentrifuge tube. The gel slices

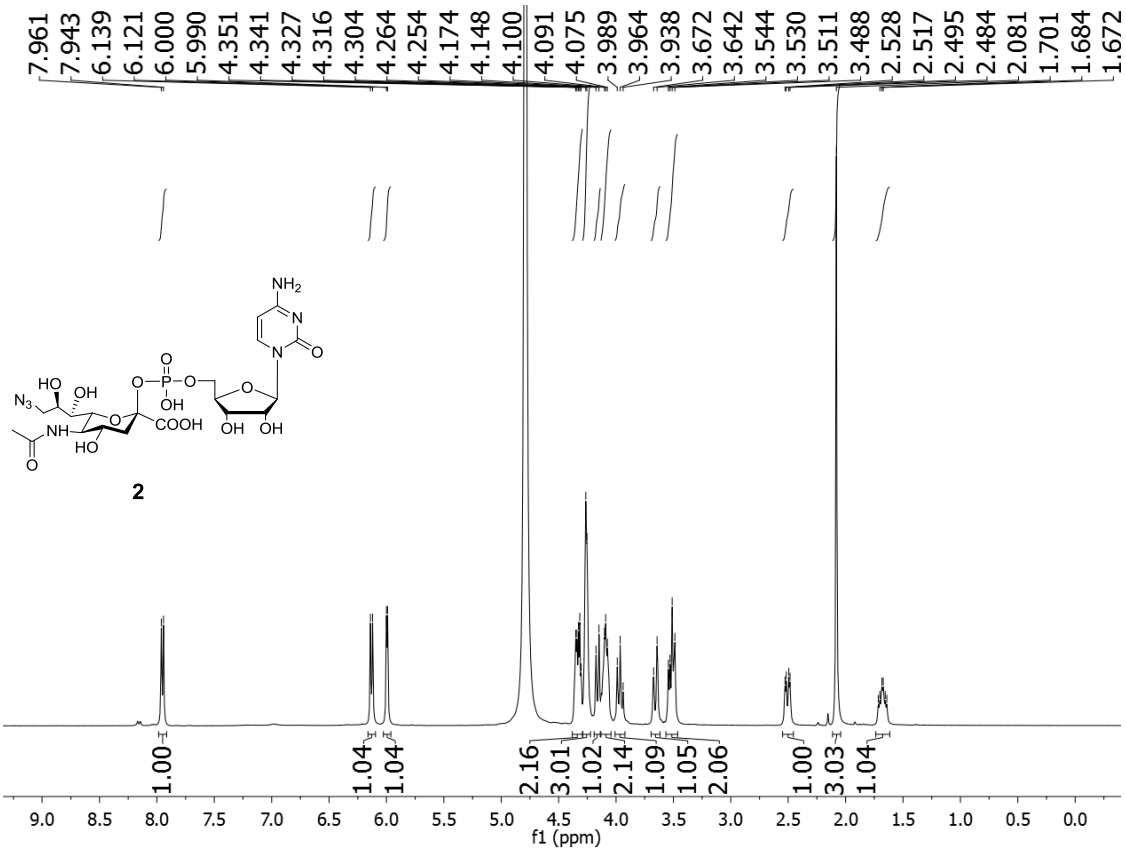
were destained twice with a 1:1 solution of 50 mM of ABC/acetonitrile for 30 min, and then dehydrated in 100% acetonitrile. The gel slices were rehydrated with 20 mM DTT in 50 mM ABC, and incubated for 1 hour at 50°C to reduce all the thiols in the gel. The slices were next incubated with 20 mM iodoacetamide in 50 mM ABC for 1 hour at rt in the dark. The gel slices were dehydrated in 100% acetonitrile. Gel pieces were rehydrated in a trypsin solution (1 ng/µl) and incubated at 37 °C in Oven for 16 h. The peptides were eluted in 67% acetonitrile in H_2O with 1% TFA (200 µL, twice), and dried by SpeedVac. Samples were then subjected to an Nano RP HPLC coupled to an LTQ-Orbitrap Elite mass spectrometer (Thermo Fisher). Peptides were separated on an EASY-Spray PepMap C18 Column (75 µm × 15 cm, 3 µm, Thermo Fisher, US) with a gradient running from 80% buffer A [LC-MS grade water with 0.1% (vol/vol) formic acid] and 20% buffer B [80%LC-MS grade acetonitrile with 0.1% (vol/ vol) formic acid] to 40% B over 35 min at 300 nl/min, next ramping to 99% B over 20 min and holding at 99% B for 10 min. LTQ-Orbitrap Elite mass spectrometer was operated in the data-dependent mode. A full-scan survey MS experiment (m/z range from 375 to 1800; automatic gain control target, 1,000,000 ions; resolution at 400 m/z, 60,000; maximum ion accumulation time, 50 ms) was acquired by the Orbitrap mass spectrometer, and 20 most intense ions were fragmented by collision-induced dissociation (CID). The other conditions were as follows: capillary temperature (200 °C), collision energy (35 eV). Peptides were identified using Proteome Discoverer™ Software and were searched against the Uniprot-human fasta file (2017_08 Release, 20,214 reviewed entries) was used for MS/MS spectra matching. The confident proteins were obtained from three repeats. Proteins that were commonly found in control sample, and any cytosolic or nuclear proteins were omitted.


7.980
7.961
6.132
6.114
5.999
5.989
4.362
4.351
4.340
4.319
4.307
4.295
4.249
4.236
4.158
4.132
4.083
4.072
3.982
3.956
3.944
3.930
3.907
3.877
3.651
3.635
3.622
3.605
3.468
3.444
2.519
2.508
2.486
2.475
2.057
1.03
1.02
1.03
1.05
1.04
3.06
1.02
1.11
3.15
1.06
1.01
1.00
3.00
1.01
f1 (ppm)

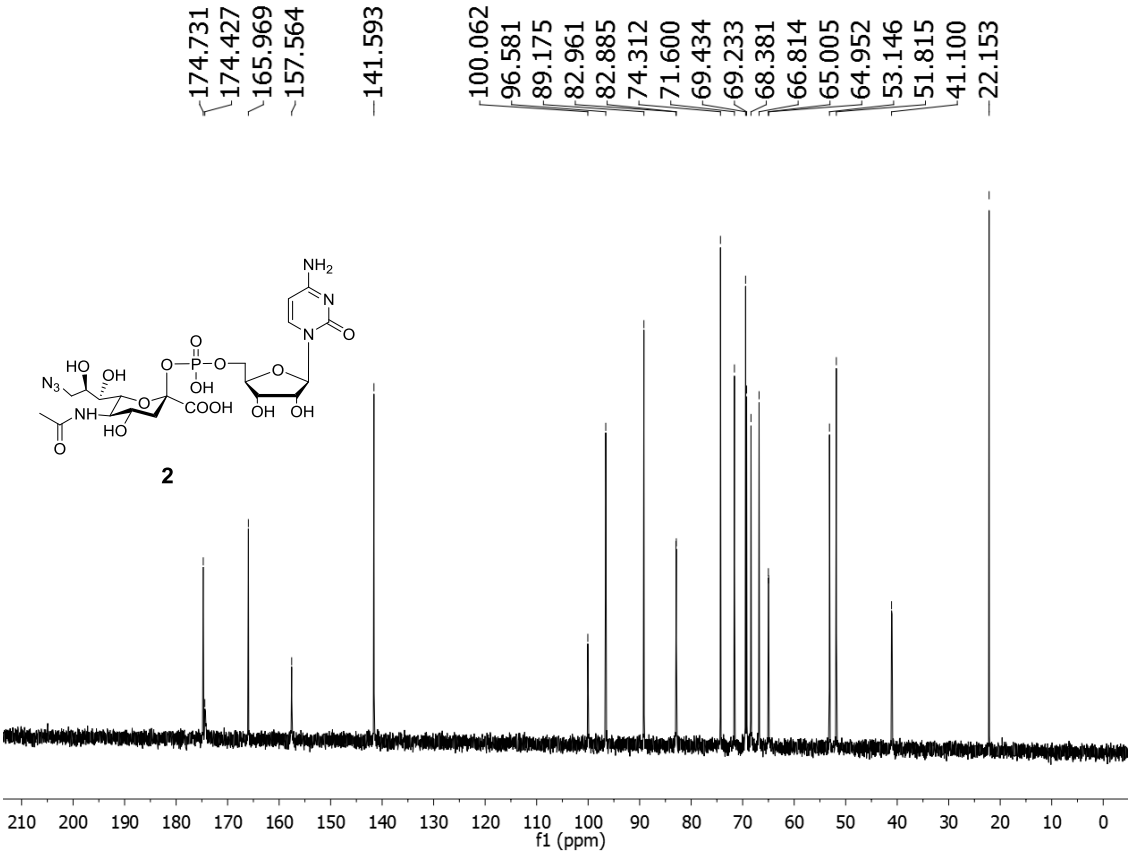
1


174.696
174.321
166.132
157.755
141.576
100.066
99.994
96.578
89.016
82.939
82.859
74.243
71.742
69.554
69.318
68.804
66.832
64.821
62.909
51.767
41.020
22.077
f1 (ppm)

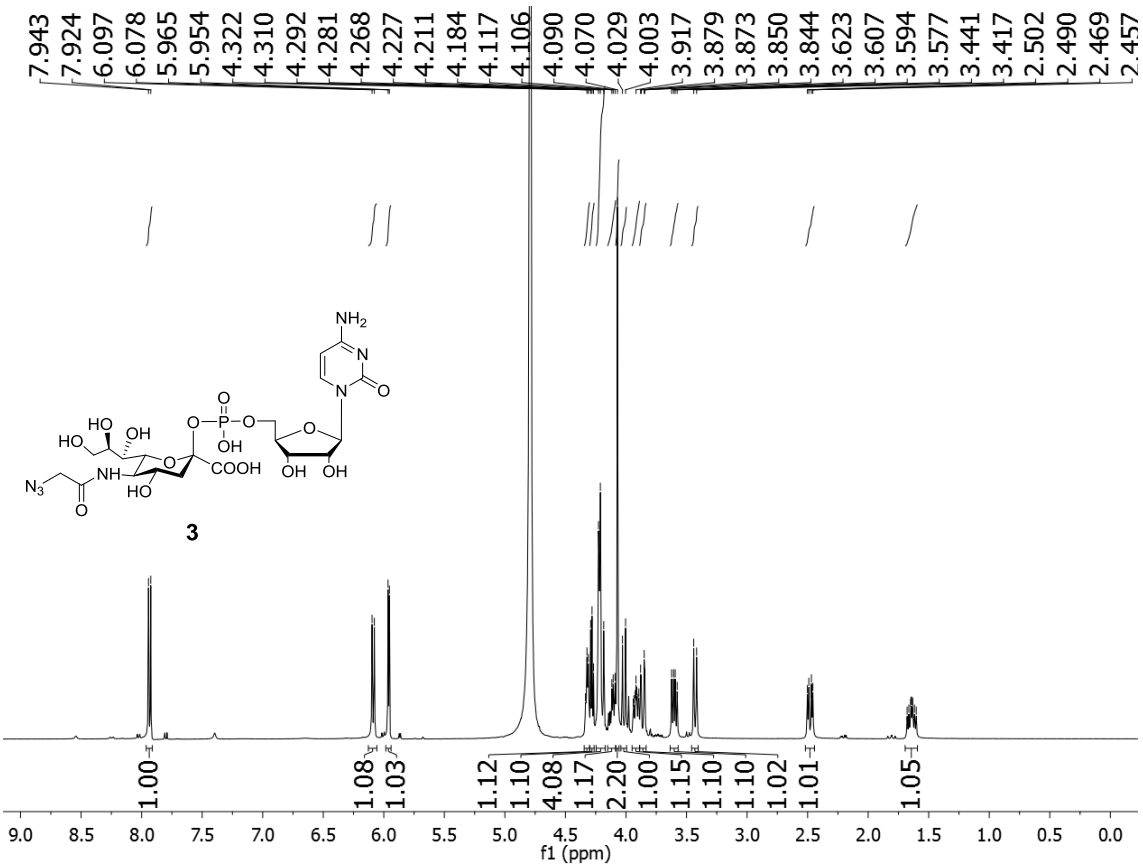
1


7.961
7.943
6.139
6.121
6.000
5.990
4.351
4.341
4.327
4.316
4.304
4.264
4.254
4.174
4.148
4.100
4.091
4.075
3.989
3.964
3.938
3.672
3.642
3.544
3.530
3.511
3.488
2.528
2.517
2.495
2.484
2.081
1.701
1.684
1.672
1.00
1.04
1.04
2.16
3.01
1.02
2.14
1.09
1.05
2.06
1.00
3.03
1.04
f1 (ppm)

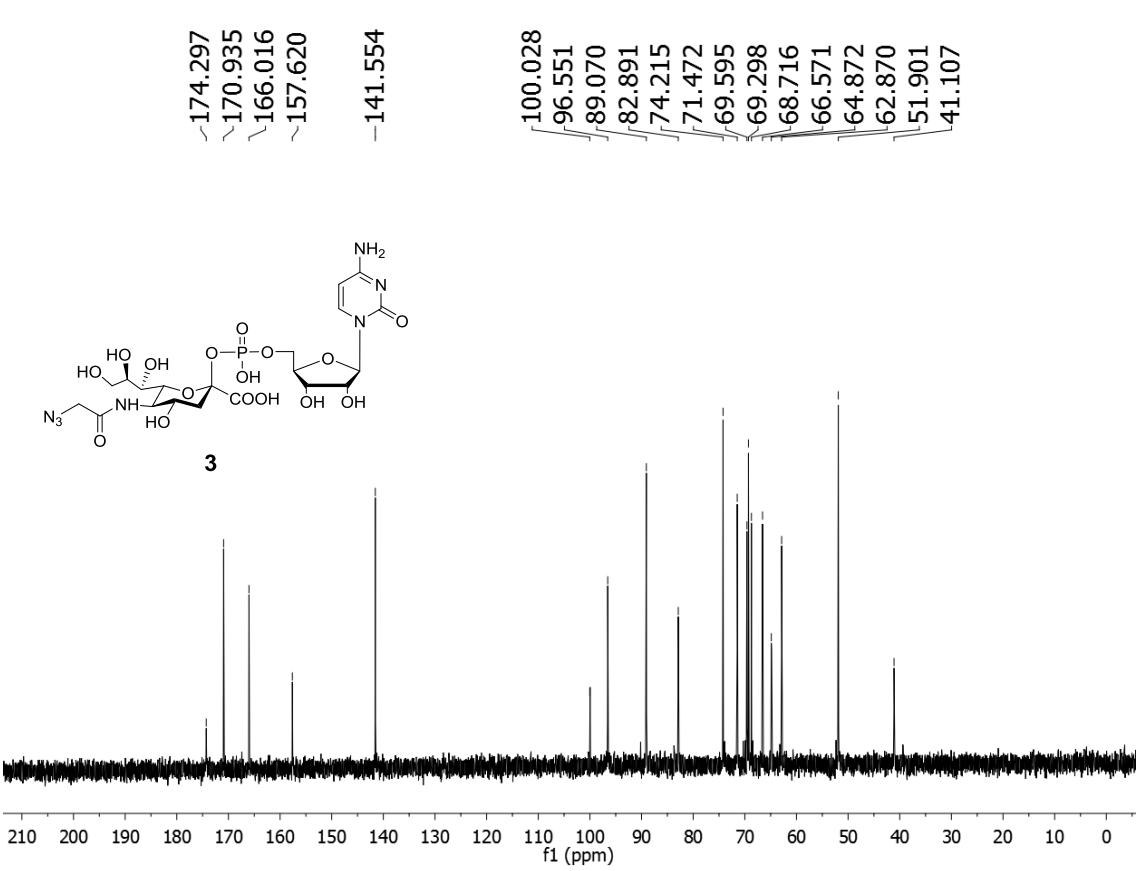
2


174.731
174.427
165.969
157.564
141.593
100.062
96.581
89.175
82.961
82.885
74.312
71.600
69.434
69.233
68.381
66.814
65.005
64.952
53.146
51.815
41.100
22.153
f1 (ppm)

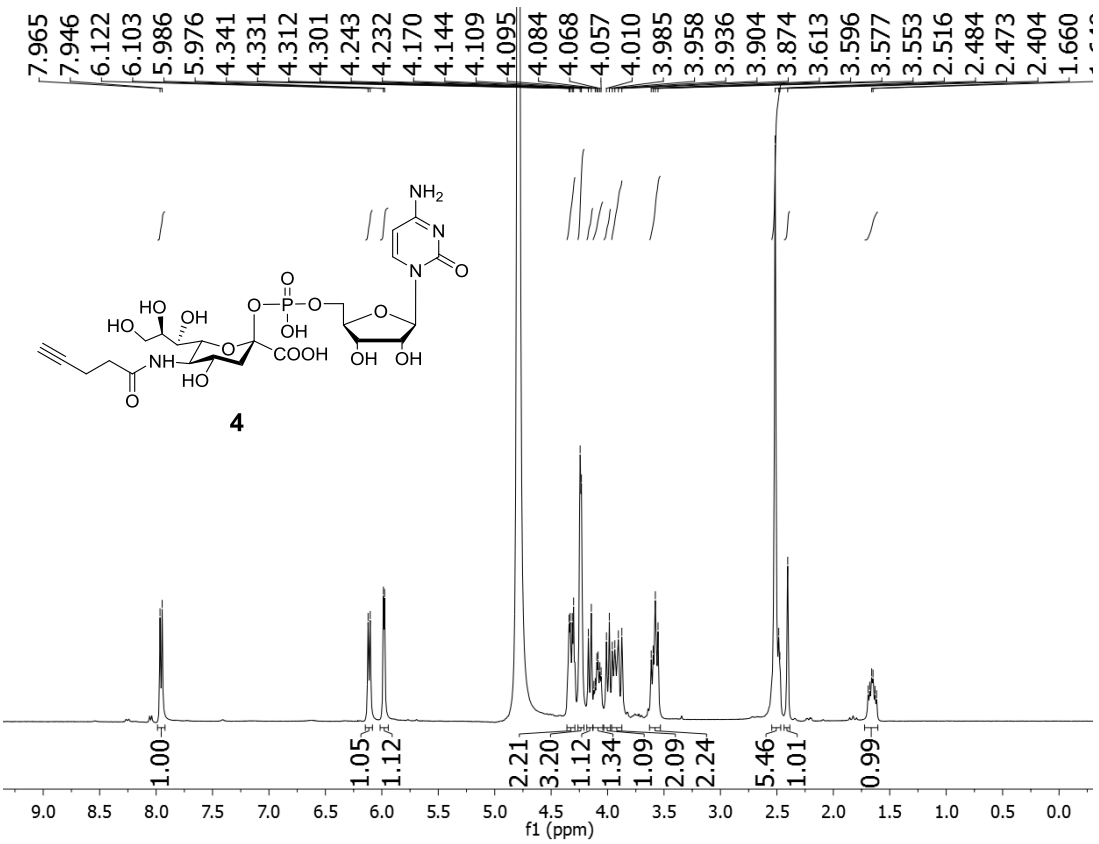
2


7.943
7.924
6.097
6.078
5.965
5.954
4.322
4.310
4.292
4.281
4.268
4.227
4.211
4.184
4.117
4.106
4.090
4.070
4.029
4.003
3.917
3.879
3.873
3.850
3.844
3.623
3.607
3.594
3.577
3.441
3.417
2.502
2.490
2.469
2.457
1.00
1.08
1.03
1.12
1.10
4.08
1.17
2.20
1.00
1.15
1.10
1.10
1.02
1.01
1.05
f1 (ppm)

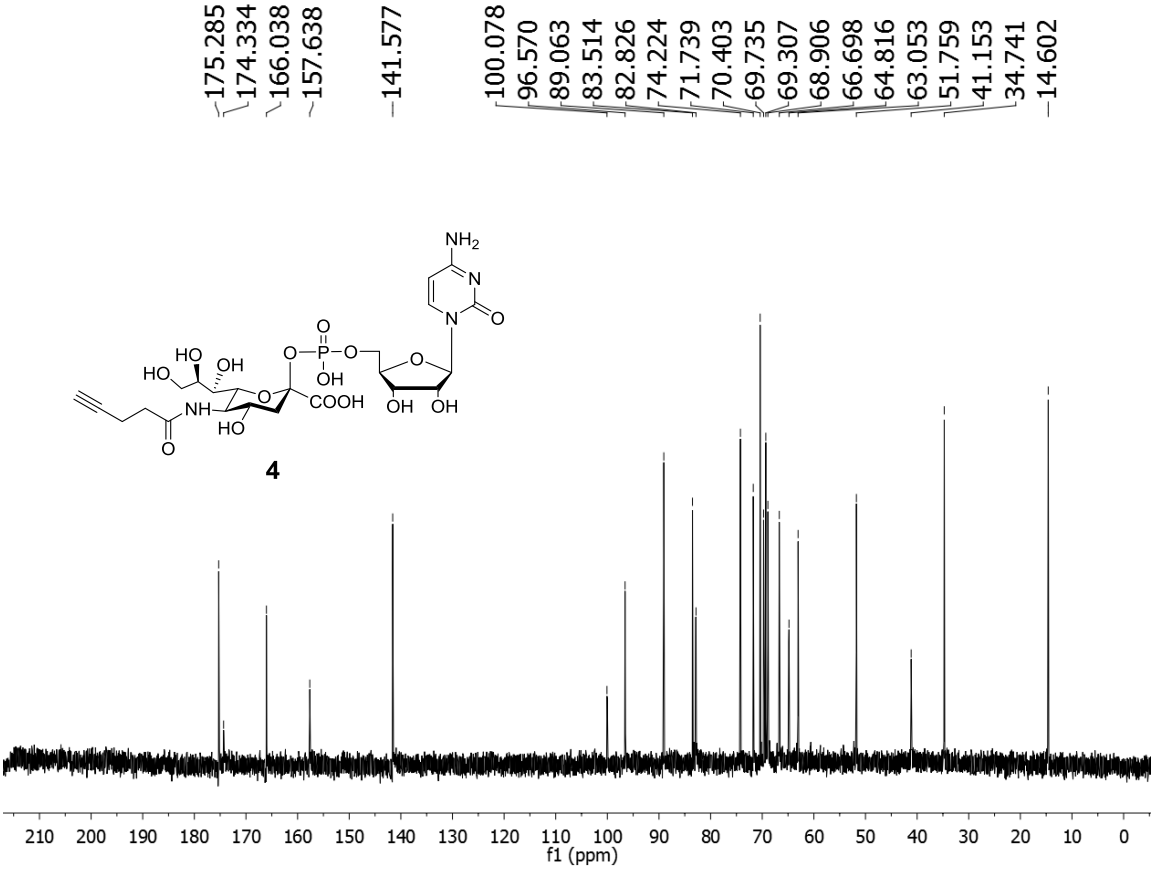
3


174.297
170.935
166.016
157.620
141.554
100.028
96.551
89.070
82.891
74.215
71.472
69.595
69.298
68.716
66.571
64.872
62.870
51.901
41.107
f1 (ppm)

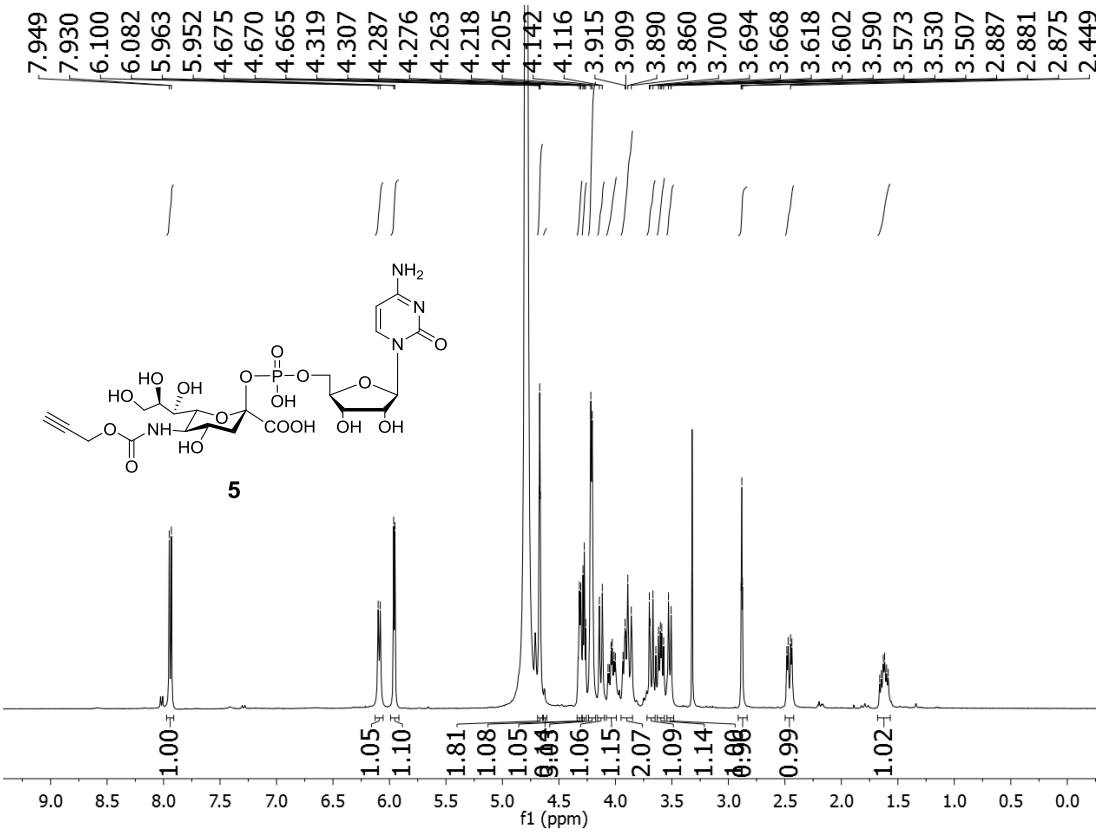
3


7.965
7.946
6.122
6.103
5.986
5.976
4.341
4.331
4.312
4.301
4.243
4.232
4.170
4.144
4.109
4.095
4.084
4.068
4.057
4.010
3.985
3.958
3.936
3.904
3.874
3.613
3.596
3.577
3.553
2.516
2.484
2.473
2.404
1.660
1.00
1.05
1.12
2.21
3.20
1.12
1.34
1.09
2.09
2.24
5.46
1.01
0.99
f1 (ppm)

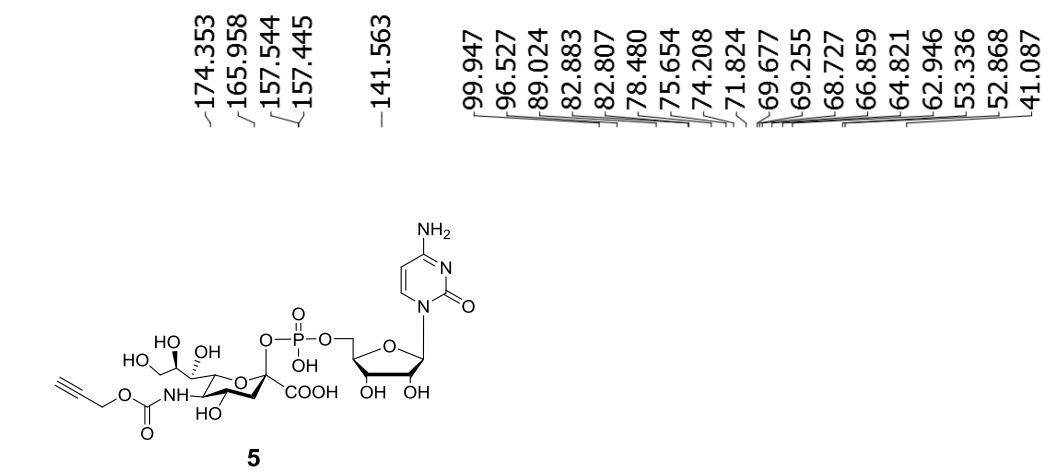
4


175.285
174.334
166.038
157.638
141.577
100.078
96.570
89.063
83.514
82.826
74.224
71.739
70.403
69.735
69.307
68.906
66.698
64.816
63.053
51.759
41.153
34.741
14.602
f1 (ppm)

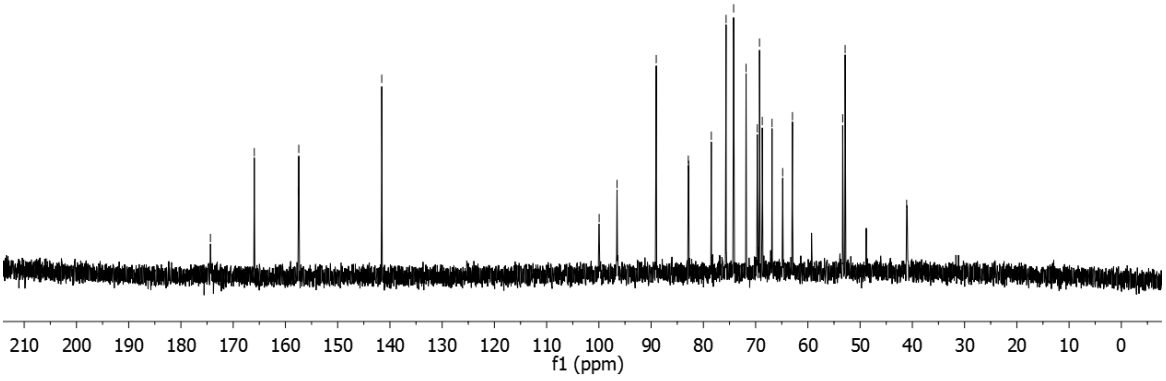
4

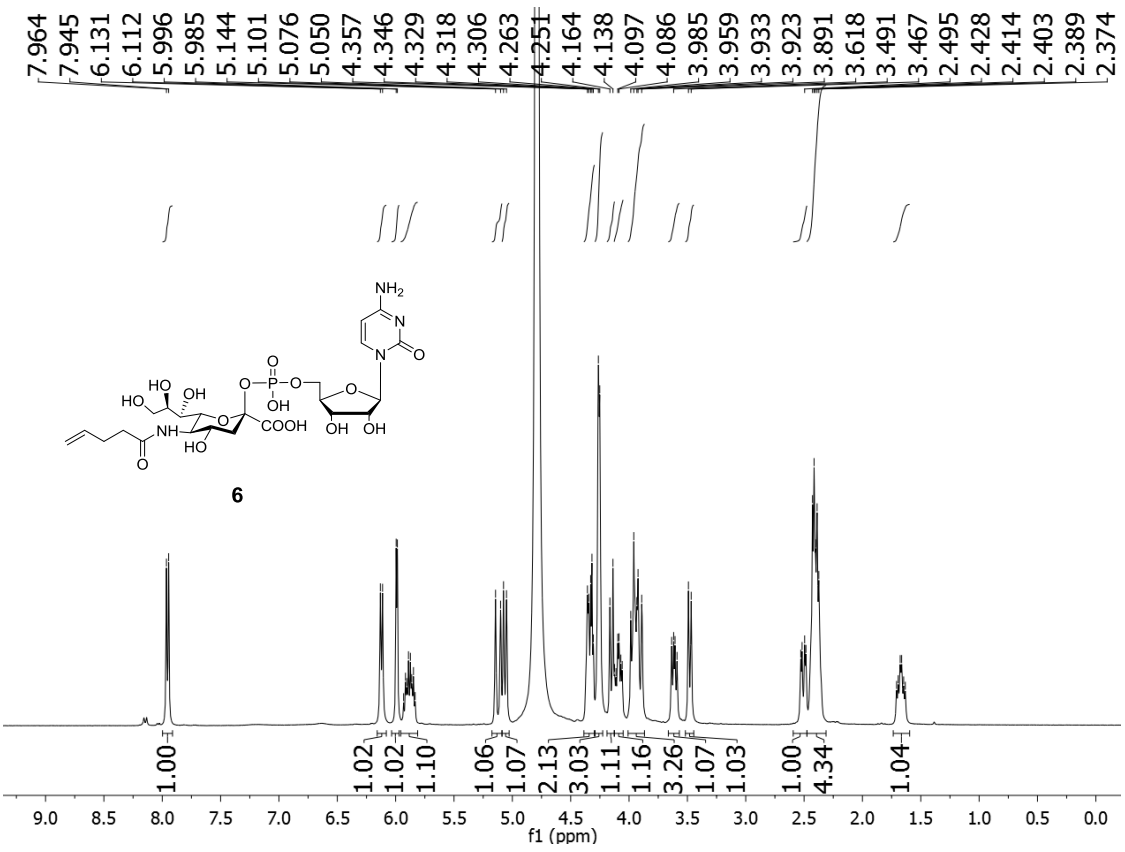

7.949
7.930
6.100
6.082
5.963
5.952
4.675
4.670
4.665
4.319
4.307
4.287
4.276
4.263
4.218
4.205
4.142
4.116
3.915
3.909
3.890
3.860
3.700
3.694
3.668
3.618
3.602
3.590
3.573
3.530
3.507
2.887
2.881
2.875
2.449
1.00
1.05
1.10
1.81
1.08
1.05
1.06
1.15
2.07
1.09
1.14
0.99
1.02
f1 (ppm)

5

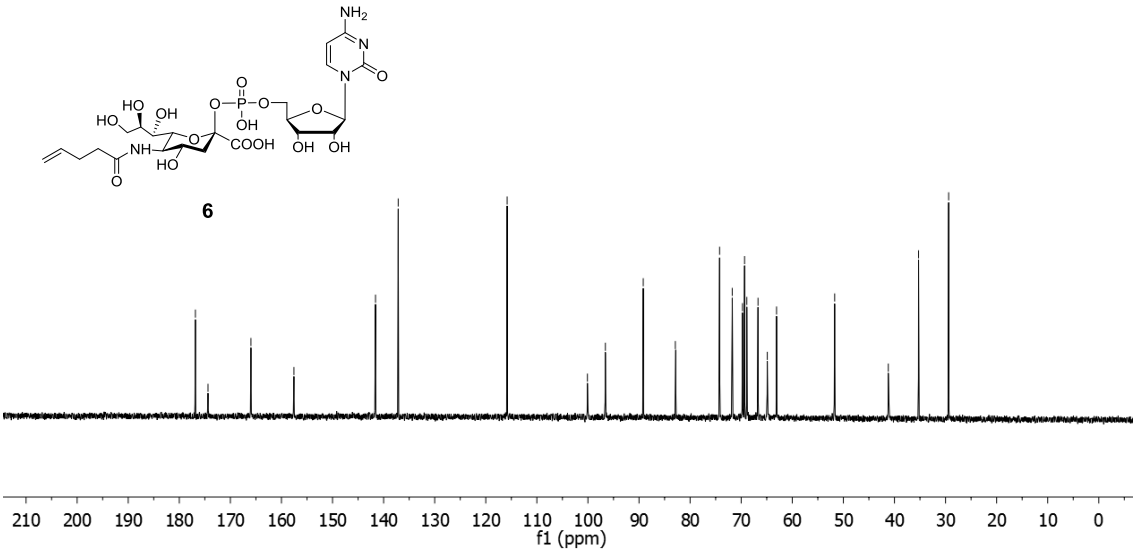

174.353
165.958
157.544
157.445
141.563
99.947
96.527
89.024
82.883
82.807
78.480
75.654
74.208
71.824
69.677
69.255
68.727
66.859
64.821
62.946
53.336
52.868
41.087

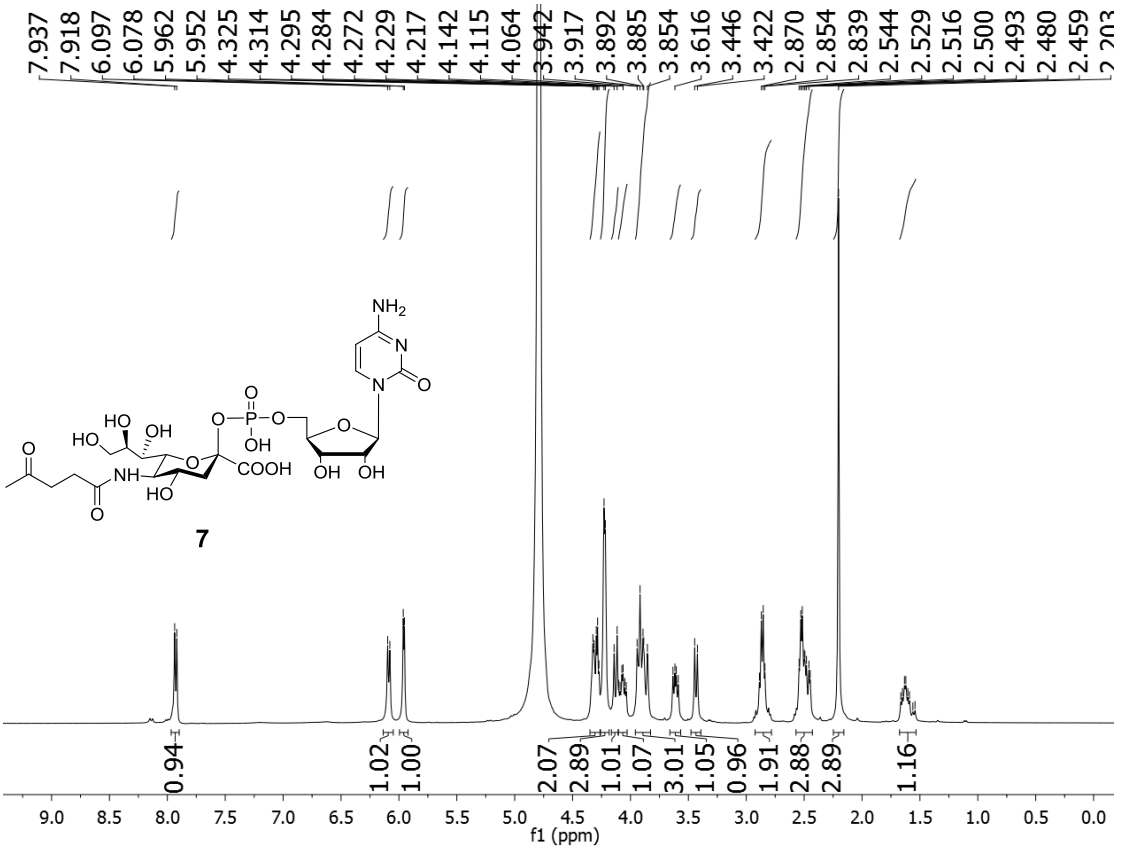
5


f1 (ppm)

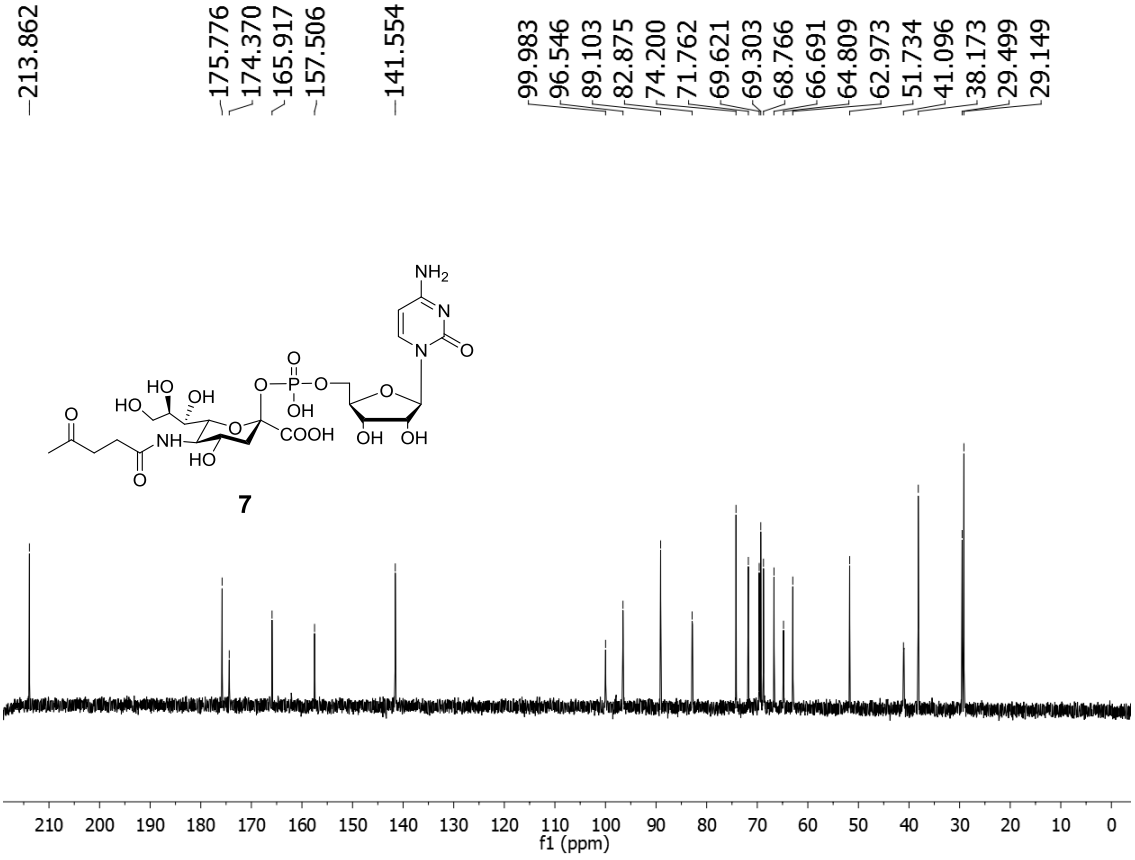

7.964
7.945
6.131
6.112
5.996
5.985
5.144
5.101
5.076
5.050
4.357
4.346
4.329
4.318
4.306
4.263
4.251
4.164
4.138
4.097
4.086
3.985
3.959
3.933
3.923
3.891
3.618
3.491
3.467
2.495
2.428
2.414
2.403
2.389
2.374
6
1.00
1.02
1.02
1.10
1.06
1.07
2.13
3.03
1.11
1.16
3.26
1.07
1.03
1.00
4.34
1.04
f1 (ppm)


176.846
174.377
165.973
157.557
141.586
137.142
115.819
100.102
96.589
89.177
82.895
74.230
71.752
69.748
69.340
68.901
66.706
64.895
63.080
51.703
41.206
35.276
29.388

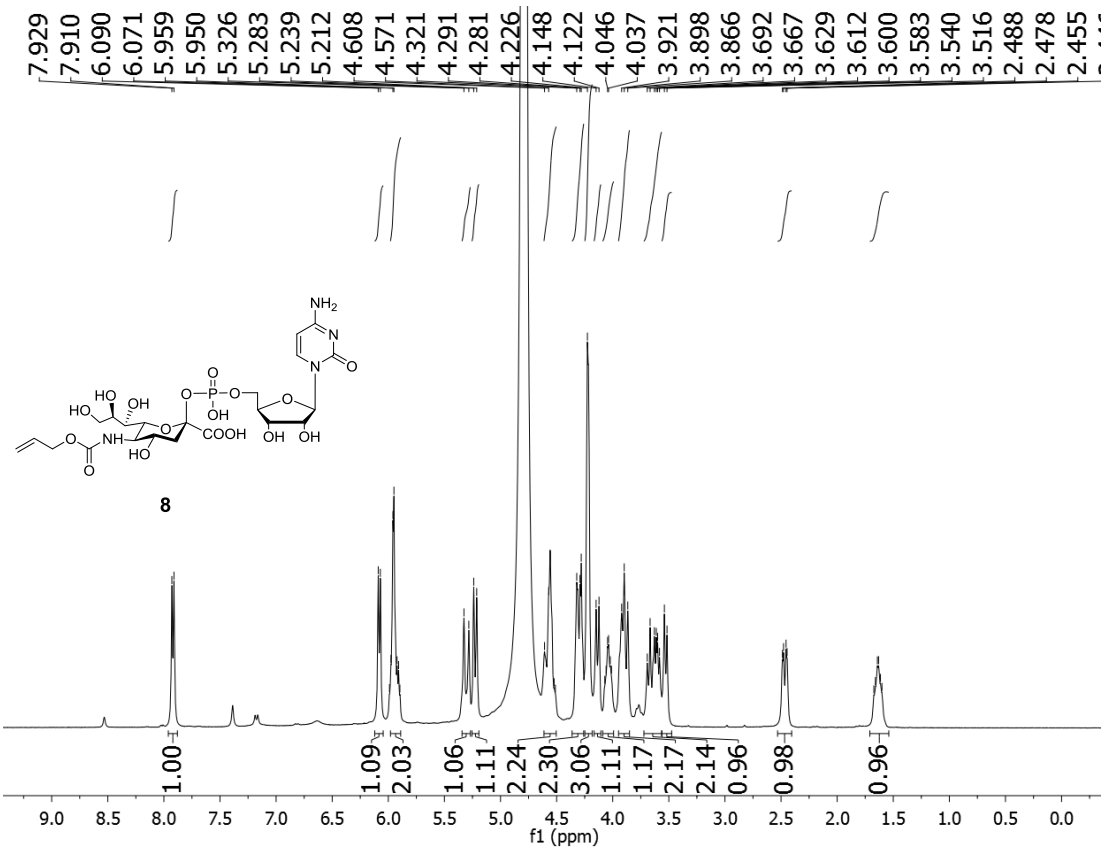

6
f1 (ppm)


7.937
7.918
6.097
6.078
5.962
5.952
4.325
4.314
4.295
4.284
4.272
4.229
4.217
4.142
4.115
4.064
3.942
3.917
3.892
3.885
3.854
3.616
3.446
3.422
2.870
2.854
2.839
2.544
2.529
2.516
2.500
2.493
2.480
2.459
0.94
1.02
1.00
2.07
2.89
1.01
1.07
3.01
1.05
0.96
1.91
2.88
2.89
1.16
f1 (ppm)

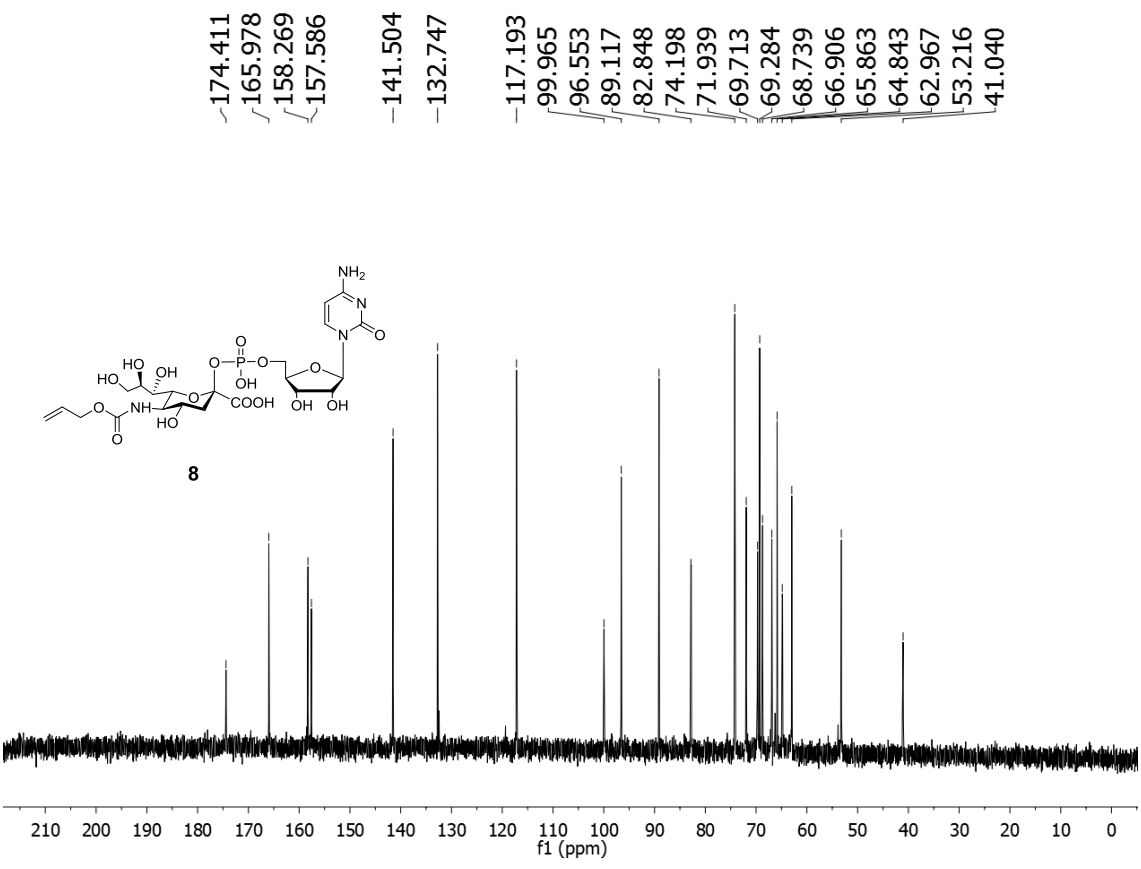
7


213.862
175.776
174.370
165.917
157.506
141.554
99.983
96.546
89.103
82.875
74.200
71.762
69.621
69.303
68.766
66.691
64.809
62.973
51.734
41.096
38.173
29.499
29.149
f1 (ppm)

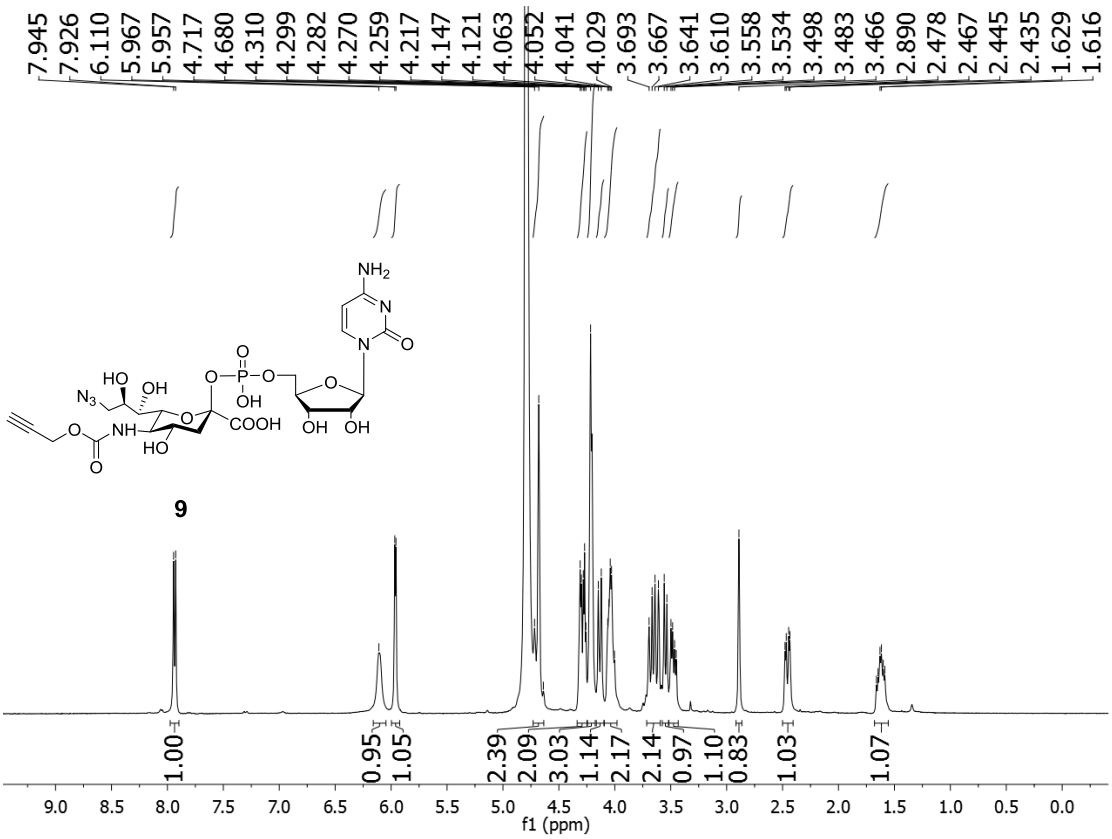
7

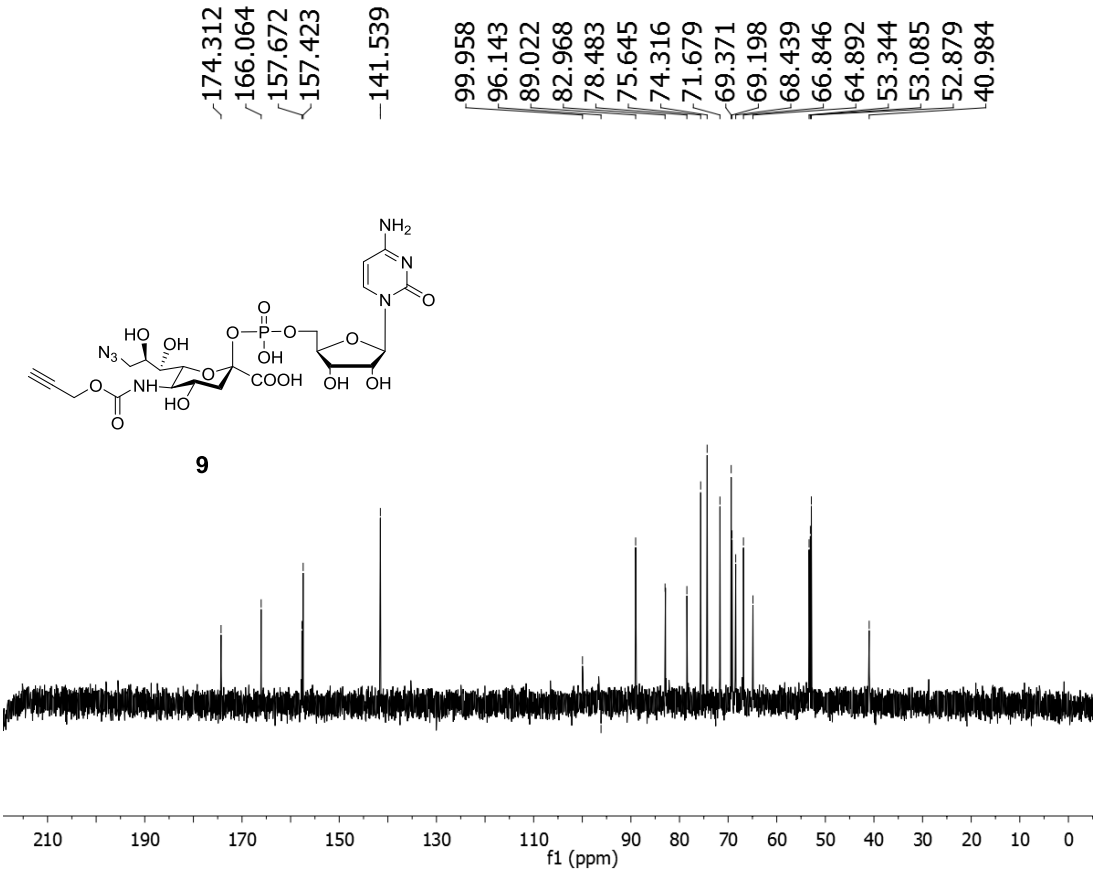

7.929
7.910
6.090
6.071
5.959
5.950
5.326
5.283
5.239
5.212
4.608
4.571
4.321
4.291
4.281
4.226
4.148
4.122
4.046
4.037
3.921
3.898
3.866
3.692
3.667
3.629
3.612
3.600
3.583
3.540
3.516
2.488
2.478
2.455
1.00
1.09
2.03
1.06
1.11
2.24
2.30
3.06
1.11
1.17
2.17
2.14
0.96
0.98
0.96
f1 (ppm)

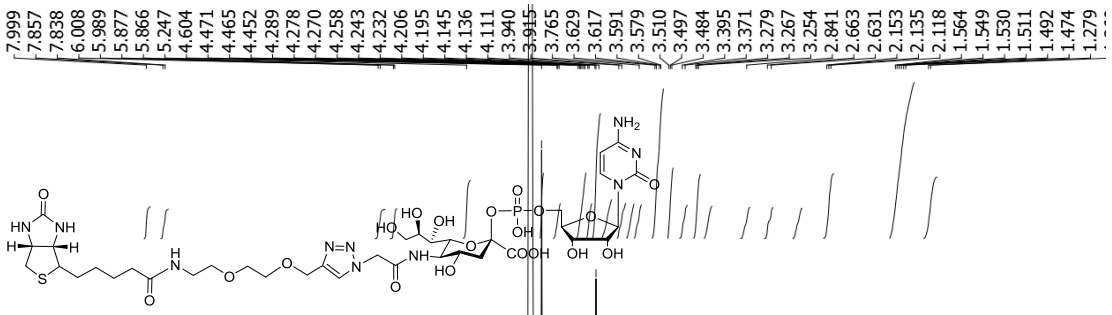
8


174.411
165.978
158.269
157.586
141.504
132.747
117.193
99.965
96.553
89.117
82.848
74.198
71.939
69.713
69.284
68.739
66.906
65.863
64.843
62.967
53.216
41.040
f1 (ppm)

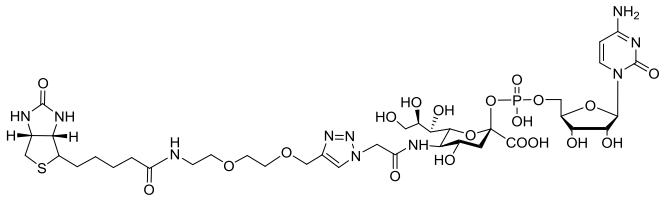
8


7.945
7.926
6.110
5.967
5.957
4.717
4.680
4.310
4.299
4.282
4.270
4.259
4.217
4.147
4.121
4.063
4.052
4.041
4.029
3.693
3.667
3.641
3.610
3.558
3.534
3.498
3.483
3.466
2.890
2.478
2.467
2.445
2.435
1.629
1.616
9
1.00
0.95
1.05
2.39
2.09
3.03
1.14
2.17
2.14
0.97
1.10
0.83
1.03
1.07
f1 (ppm)

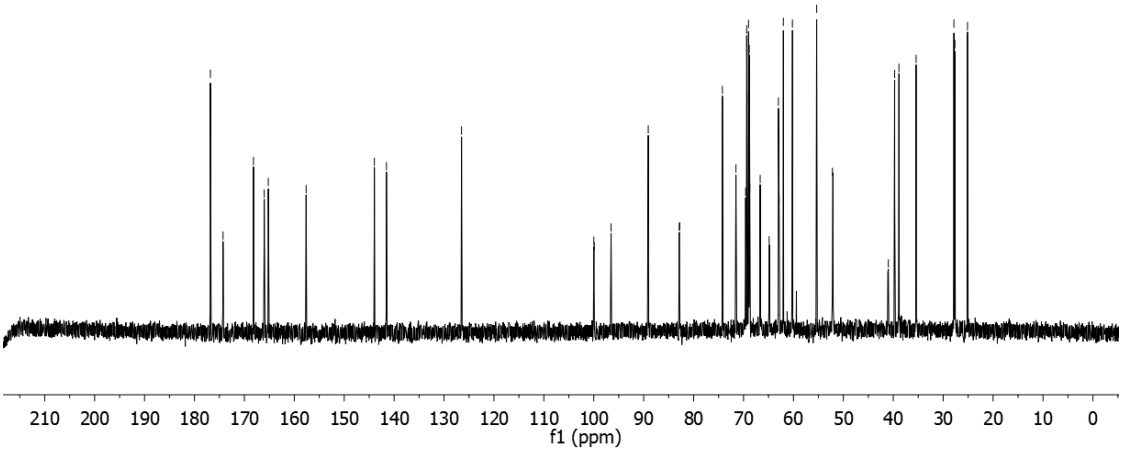

174.312
166.064
157.672
157.423
141.539
99.958
96.143
89.022
82.968
78.483
75.645
74.316
71.679
69.371
69.198
68.439
66.846
64.892
53.344
53.085
52.879
40.984
9
f1 (ppm)

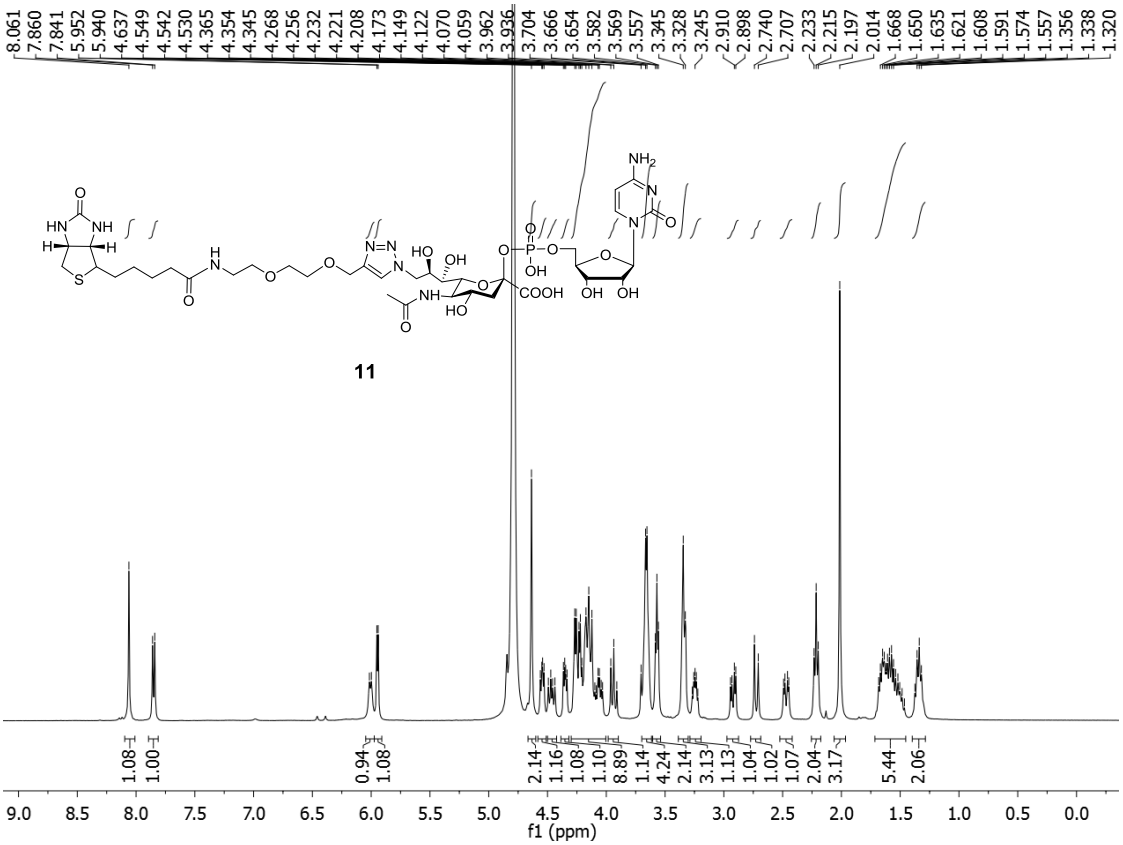


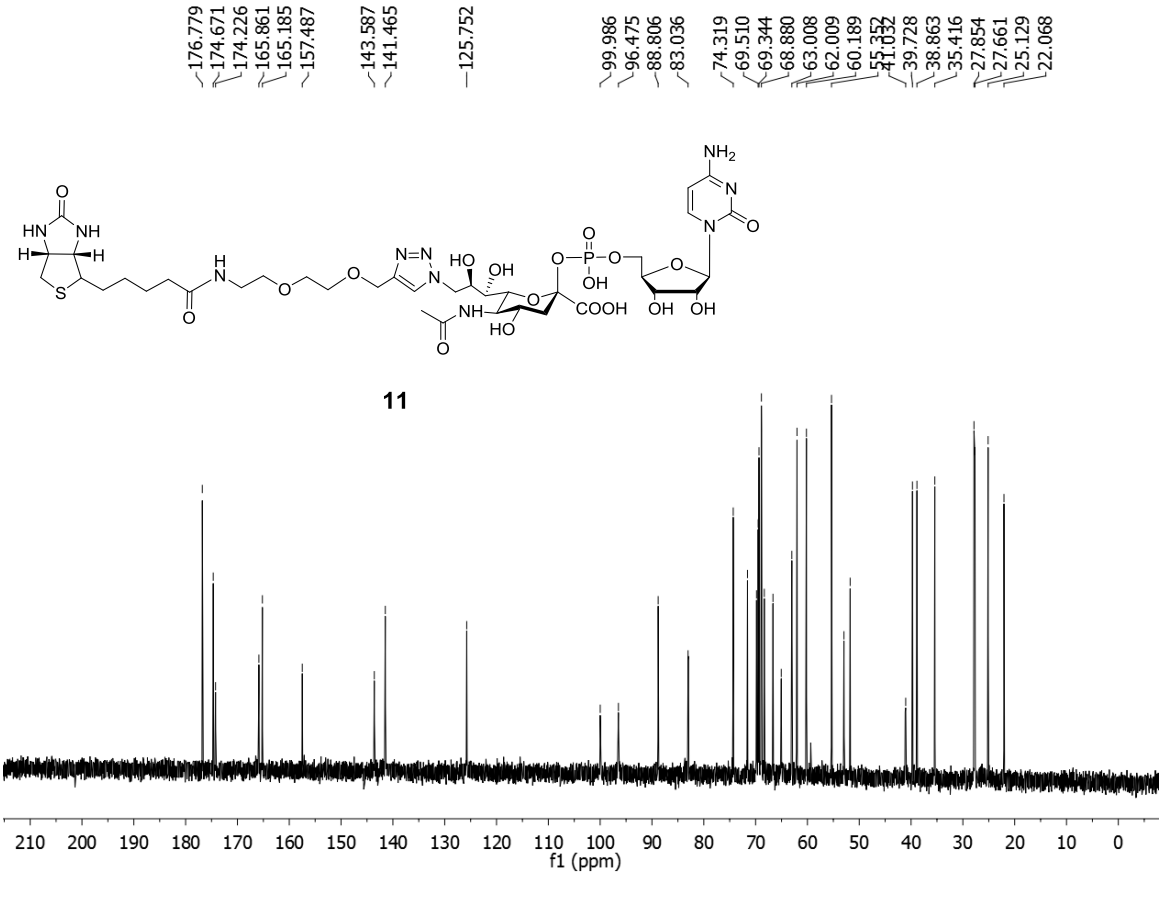
10

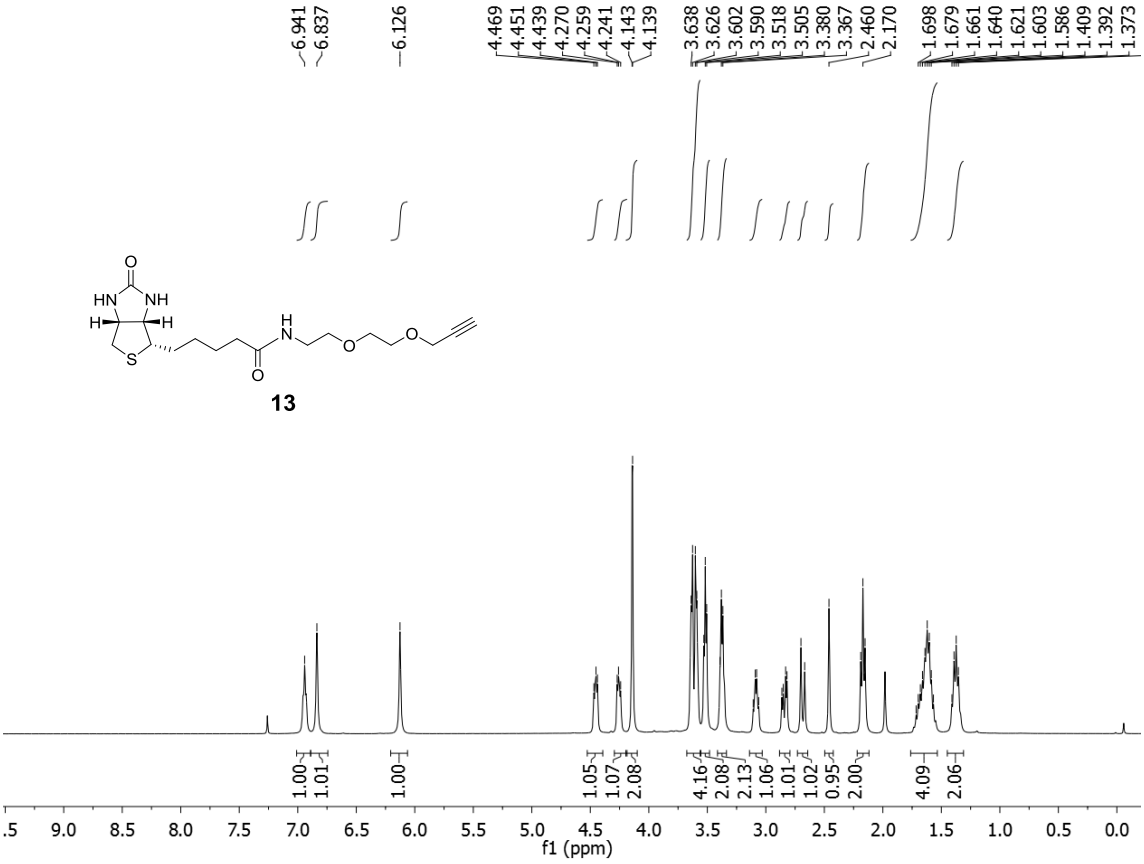


10

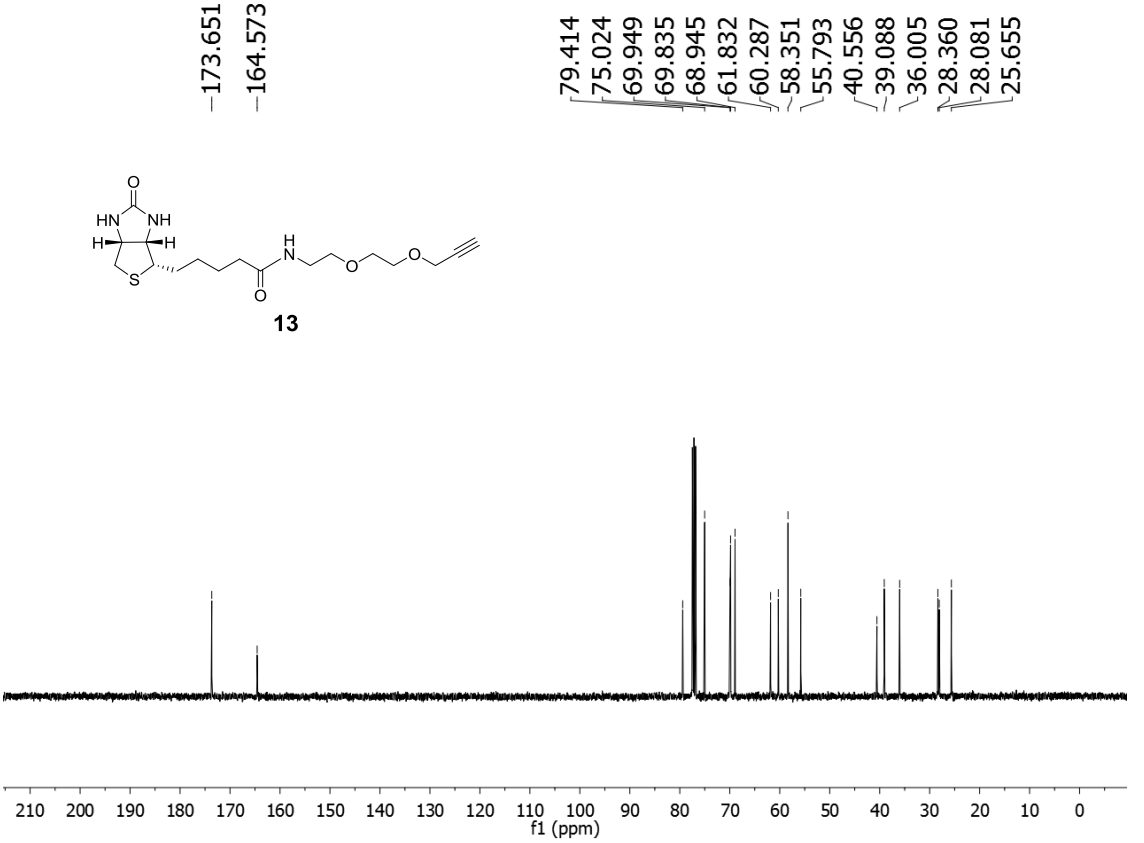



8.061
7.860
7.841
5.952
5.940
4.637
4.549
4.542
4.530
4.365
4.354
4.345
4.268
4.256
4.232
4.221
4.208
4.173
4.149
4.122
4.070
4.059
3.962
3.936
3.704
3.666
3.654
3.582
3.569
3.557
3.345
3.328
3.245
2.910
2.898
2.740
2.707
2.233
2.215
2.197
2.014
1.668
1.650
1.635
1.621
1.608
1.591
1.574
1.557
1.356
1.338
1.320
11
1.08
1.00
0.94
1.08
2.14
1.16
1.08
1.10
8.89
1.14
4.24
2.14
3.13
1.13
1.04
1.02
1.07
2.04
3.17
5.44
2.06
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


176.779
174.671
174.226
165.861
165.185
157.487
143.587
141.465
125.752
99.986
96.475
88.806
83.036
74.319
69.510
69.344
68.880
63.008
62.009
60.189
55.352
41.032
39.728
38.863
35.416
27.854
27.661
25.129
22.068
11
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)


6.941
6.837
6.126
4.469
4.451
4.439
4.270
4.259
4.241
4.143
4.139
3.638
3.626
3.602
3.590
3.518
3.505
3.380
3.367
2.460
2.170
1.698
1.679
1.661
1.640
1.621
1.603
1.586
1.409
1.392
1.373
1.00
1.01
1.00
1.05
1.07
2.08
4.16
2.08
2.13
1.06
1.01
1.02
0.95
2.00
4.09
2.06
f1 (ppm)

13


173.651
164.573
79.414
75.024
69.949
69.835
68.945
61.832
60.287
58.351
55.793
40.556
39.088
36.005
28.360
28.081
25.655
f1 (ppm)

13

Table S1. 78 proteins identified from MCF7

MCF7			Label 1		Label 2		Label 3				
Uniprot Accession	Gene name	Description	Peptides	Spectra count	Peptides	Spectra count	Peptides	Spectra count	Total spectra	Amino acid	MW [kDa]
Q9Y5X1	SNX9	Sorting nexin-9	1	1	1	1	1	1	3	595	66.54962
Q9UNZ2-6	NSFL1C	Isoform 4 of NSFL1 cofactor p47	1	1	1	1	1	1	3	259	28.50472
Q9UJZ1-2	STOML2	Isoform 2 of Stomatin-like protein 2, mitochondrial	5	5	4	4	5	5	14	311	33.31663
Q9UBB4-2	ATXN10	Isoform 2 of Ataxin-10	1	1	1	1	1	1	3	411	46.25717
Q9NQC3-3	RTN4	Isoform 3 of Reticulon-4	1	1	1	1	1	1	3	199	22.38126
Q9H4A4	RNPEP	Aminopeptidase B	1	1	4	4	2	2	7	650	72.54946
Q9H223	EHD4	EH domain-containing protein 4	1	1	2	2	2	2	5	541	61.13653
Q9BPW8	NIPSNAP1	Protein NipSnap homolog 1	1	1	1	1	1	1	3	284	33.28884
Q99829	CPNE1	Copine-1	1	1	1	1	1	1	3	537	59.02155
Q99714-2	HSD17B10	Isoform 2 of 3-hydroxyacyl-CoA dehydrogenase type-2	1	1	2	2	2	2	5	252	25.96758
Q99623-2	PHB2	Isoform 2 of Prohibitin-2	5	6	5	5	4	4	15	261	29.02573
Q99497	PARK7	Protein DJ-1	2	2	2	2	1	1	5	189	19.87849
Q96H20-2	SNF8	Isoform 2 of Vacuolar-sorting protein SNF8	1	1	1	1	1	2	4	257	28.718
Q96CX2	KCTD12	BTB/POZ domain-containing protein KCTD12	1	1	1	1	1	1	3	325	35.67869
Q92597-3	NDRG1	Isoform 3 of Protein NDRG1	1	1	1	1	1	1	3	313	33.62827
Q8WUM4	PDCD6IP	Programmed cell death 6-interacting protein	6	6	6	6	6	6	18	868	95.96312
Q8NC51-4	SERBP1	Isoform 4 of Plasminogen activator inhibitor 1 RNA-binding protein	2	2	2	2	2	2	6	387	42.40128
Q16658	FSCN1	Fascin	5	5	6	6	5	5	16	493	54.49605
Q15046	KARS	Lysine--tRNA ligase	1	1	3	3	4	4	8	597	68.0046
Q15019	SEPT-2	Septin-2	5	7	2	2	4	4	13	361	41.46125
Q07954	LRP1	Prolow-density lipoprotein receptor-related protein 1	1	1	3	3	1	1	5	4544	504.2759
Q07065	CKAP4	Cytoskeleton-associated protein 4	6	6	6	6	4	4	16	602	65.98273
Q07021	C1QBP	Complement component 1 Q subcomponent-binding protein, mitochondrial	2	2	1	1	2	3	6	282	31.34261
Q04917	YWHAH	14-3-3 protein eta	3	5	3	4	5	5	14	246	28.20102
Q02413	DSG1	Desmoglein-1	1	2	7	8	9	10	20	1049	113.6759
Q01518	CAP1	Adenylyl cyclase-associated protein 1	3	3	1	1	2	2	6	475	51.8687
Q01082-3	SPTBN1	Isoform 2 of Spectrin beta chain, non-erythrocytic 1	1	1	1	2	1	2	5	2155	251.2412
P67775-2	PPP2CA	Isoform 2 of Serine/threonine-protein phosphatase 2A catalytic subunit alpha isoform	2	2	2	2	2	2	6	255	29.71567
P63010-3	AP2B1	Isoform 3 of AP-2 complex subunit beta	2	2	3	3	2	2	7	880	98.05567
P61421	ATP6V0D1	V-type proton ATPase subunit d 1	1	1	1	1	1	1	3	351	40.30322
P61026	RAB10	Ras-related protein Rab-10	1	1	1	1	1	1	3	200	22.52659
P61006	RAB8A	Ras-related protein Rab-8A	1	1	1	1	1	1	3	207	23.65319
P54578-2	USP14	Isoform 2 of Ubiquitin carboxyl-terminal hydrolase 14	1	1	1	1	1	1	3	459	52.35233
P53618	COPB1	Coatomer subunit beta	2	2	6	6	6	7	15	953	107.0738
P51572	BCAP31	B-cell receptor-associated protein 31	1	2	1	1	1	1	4	246	27.97401
P51149	RAB7A	Ras-related protein Rab-7a	3	3	3	3	2	2	8	207	23.47484
P50851-2	LRBA	Isoform 2 of Lipopolysaccharide-responsive and beige-like anchor protein	1	1	2	2	1	1	4	2851	317.5013
P49903-2	SEPHS1	Isoform 2 of Selenide, water dikinase 1	1	1	1	1	1	1	3	321	35.456
P49368-2	CCT3	Isoform 2 of T-complex protein 1 subunit gamma	3	3	5	6	1	1	10	507	56.39524
P49327	FASN	Fatty acid synthase	12	14	19	21	14	17	52	2511	273.2543

P48047	ATP5O	ATP synthase subunit O, mitochondrial	2	4	2	2	1	1	7	213	23.26266
P46459-2	NSF	Isoform 2 of Vesicle-fusing ATPase	1	1	1	1	1	1	3	644	71.53859
P35221	CTNNA1	Catenin alpha-1	4	4	6	7	4	4	15	906	100.0085
Q9UQ80-2	PA2G4	Isoform 2 of Proliferation-associated protein 2G4	2	2	1	2	2	2	6	340	38.03539
P27824	CANX	Calnexin	4	4	5	5	2	2	11	592	67.52585
P25705	ATP5A1	ATP synthase subunit alpha, mitochondrial	11	16	12	18	12	14	48	553	59.7136
P20073-2	ANXA7	Isoform 2 of Annexin A7	1	1	2	2	1	1	4	466	50.28367
P18206-2	VCL	Isoform 1 of Vinculin	2	2	4	4	1	1	7	1066	116.6493
P16144-4	ITGB4	Isoform Beta-4D of Integrin beta-4	2	2	3	3	2	2	7	1745	194.3419
P14923	JUP	Junction plakoglobin	1	1	3	3	4	4	8	745	81.69273
P14550	AKR1A1	Alcohol dehydrogenase [NADP(+)]	2	2	2	2	2	3	7	325	36.54986
P13010	XRCC5	X-ray repair cross-complementing protein 5	4	5	6	9	4	5	19	732	82.65228
P12429	ANXA3	Annexin A3	1	1	1	1	1	1	3	323	36.35266
P11413	G6PD	Glucose-6-phosphate 1-dehydrogenase	3	3	1	1	1	1	5	515	59.219
P09429	HMGB1	High mobility group protein B1	4	5	2	3	4	5	13	215	24.87816
P27708	CAD	CAD protein	2	2	4	4	2	2	8	2225	242.8295
P09211	GSTP1	Glutathione S-transferase P	3	10	3	6	2	3	19	210	23.34102
P08727	KRT19	Keratin, type I cytoskeletal 19	12	15	9	13	12	17	45	400	44.07912
P08575-2	PTPRC	Isoform 2 of Receptor-type tyrosine-protein phosphatase C	1	1	1	1	1	1	3	1143	130.815
P08195-2	SLC3A2	Isoform 2 of 4F2 cell-surface antigen heavy chain	1	1	3	3	4	4	8	529	57.90887
P07384	CAPN1	Calpain-1 catalytic subunit	2	2	1	1	1	1	4	714	81.83819
P06744	GPI	Glucose-6-phosphate isomerase	4	7	4	8	4	8	23	558	63.10725
P05787	KRT8	Keratin, type II cytoskeletal 8	22	55	20	47	19	45	147	483	53.67114
P05556-2	ITGB1	Isoform 2 of Integrin beta-1	1	3	1	1	1	1	5	789	87.38855
Q12905	ILF2	Interleukin enhancer-binding factor 2	3	5	2	2	4	4	11	390	43.03519
P55060-3	CSE1L	Isoform 3 of Exportin-2	7	7	9	9	4	4	20	945	107.7089
P04792	HSPB1	Heat shock protein beta-1	3	3	5	6	3	4	13	205	22.76849
P04632	CAPNS1	Calpain small subunit 1	3	3	2	3	3	5	11	268	28.29773
P04083	ANXA1	Annexin A1	12	13	9	10	9	11	34	346	38.68998
P02786	TFRC	Transferrin receptor protein 1	5	5	3	4	5	5	14	760	84.81795
P01857	IGHG1	Immunoglobulin heavy constant gamma 1	1	118	2	54	3	49	221	330	36.08317
P01834	IGKC	Immunoglobulin kappa constant	2	31	2	2	2	2	35	107	11.75777
P00505-2	GOT2	Isoform 2 of Aspartate aminotransferase, mitochondrial	1	1	3	3	2	2	6	387	43.002
P00390-5	GSR	Isoform 4 of Glutathione reductase, mitochondrial	2	2	3	3	2	2	7	440	47.23724
O95793-2	STAU1	Isoform Short of Double-stranded RNA-binding protein Staufen homolog 1	2	3	1	1	2	2	6	496	54.89959
O75534-2	CSDE1	Isoform 2 of Cold shock domain-containing protein E1	1	1	1	1	3	3	5	767	85.69339
O00299	CLIC1	Chloride intracellular channel protein 1	3	3	3	3	3	4	10	241	26.90575
O00161-2	SNAP23	Isoform SNAP-23b of Synaptosomal-associated protein 23	1	1	1	1	1	1	3	158	17.77779

Table S2. 43 proteins identified from HT29

HT29			label 1		label 2		label 3				
Uniprot Accession	Gene name	Description	Peptides	Spectra count	Peptides	Spectra count	Peptides	Spectra count	Total spectra	Amino acid	MW [kDa]
Q9Y446	PKP3	Plakophilin-3	1	1	6	6	3	3	10	797	87.029
Q9UJZ1-2	STOML2	Stomatin-like protein 2, mitochondrial	3	3	1	1	3	3	7	311	33.3166
Q9H223	EHD4	EH domain-containing protein 4	1	1	1	1	2	2	4	541	61.1365
Q99623-2	PHB2	Prohibitin-2	2	3	3	3	3	3	9	261	29.0257
Q92896	GLG1	Golgi apparatus protein 1	2	2	4	4	3	3	9	1179	134.464
Q92673	SORL1	Sortilin-related receptor	2	2	3	3	3	3	8	2214	248.267
Q16643	DBN1	Drebrin	1	1	1	1	1	1	3	649	71.3854
Q16222-2	UAP1	UDP-N-acetylhexosamine pyrophosphorylase	1	1	1	1	1	2	4	505	56.992
Q15149-7	PLEC	Plectin	2	2	4	5	6	6	13	4515	512.292
Q13813-3	SPTAN1	Spectrin alpha chain, non-erythrocytic 1	1	1	5	5	2	2	8	2452	282.108
Q12905	ILF2	Interleukin enhancer-binding factor 2	2	2	1	1	4	5	8	390	43.0352
Q12797-10	ASPH	Aspartyl/asparaginyl beta-hydroxylase	2	2	3	3	2	2	7	729	83.2162
Q02952-3	AKAP12	A-kinase anchor protein 12	3	4	3	3	7	7	14	1677	180.881
Q02413	DSG1	Desmoglein-1	2	3	1	2	1	2	7	1049	113.676
Q01082-3	SPTBN1	Spectrin beta chain, non-erythrocytic 1	1	1	1	1	5	5	7	2155	251.241
P60842-2	EIF4A1	Eukaryotic initiation factor 4A-I	2	2	1	1	1	1	4	347	39.5232
P53396-3	ACLY	ATP-citrate synthase	1	1	4	5	5	5	11	830	91.0408
P50991-2	CCT4	T-complex protein 1 subunit delta	2	2	3	3	5	5	10	509	54.6851
P50990-2	CCT8	T-complex protein 1 subunit theta	1	1	3	3	4	4	8	529	57.6084
P50851-2	LRBA	Lipopolysaccharide-responsive and beige-like anchor protein	1	1	3	3	3	3	7	2851	317.501
P47897-2	IQGAP1	Ras GTPase-activating-like protein IQGAP1	1	1	2	2	2	2	5	764	86.5242
P46940	IQGAP1	Ras GTPase-activating-like protein IQGAP1	6	6	8	8	13	14	28	1657	189.134
P35221	CTNNA1	Catenin alpha-1	2	2	5	5	6	8	15	906	100.009
P30520	ADSS	Adenylosuccinate synthetase isozyme 2	1	1	1	1	1	1	3	456	50.0658
P25705	ATP5A1	ATP synthase subunit alpha, mitochondrial	12	16	7	7	13	18	41	553	59.7136
P23528	CFL1	Cofilin-1	1	1	1	1	2	2	4	166	18.4907
P23396	RPS3	40S ribosomal protein S3	1	1	1	1	2	2	4	243	26.6714
P22234	PAICS	Multifunctional protein ADE2	1	1	1	1	2	2	4	425	47.0491
P18206-2	VCL	Vinculin	2	2	3	3	2	3	8	1066	116.649
P17301	ITGA2	Integrin alpha-2	2	2	2	2	4	5	9	1181	129.214
P16615-2	ATP2A2	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	3	3	8	9	10	12	24	997	109.62
P16144-4	ITGB4	Integrin beta-4	2	2	4	4	7	7	13	1745	194.342
P16070-18	CD44	CD44 antigen	2	2	1	2	1	1	5	340	37.2542
P12814-2	ACTN1	Alpha-actinin-1	4	4	7	7	7	7	18	887	102.644
P11717	IGF2R	Cation-independent mannose-6-phosphate receptor	2	2	2	2	4	4	8	2491	274.199
P11413	G6PD	Glucose-6-phosphate 1-dehydrogenase	1	1	1	1	1	1	3	515	59.219

P11142-2	HSPA8	Heat shock cognate 71 kDa protein	2	2	2	2	11	13	17	493	53.4844
P08758	ANXA5	Annexin A5	1	1	1	1	1	1	3	320	35.9144
P08195-2	SLC3A2	4F2 cell-surface antigen heavy chain	1	1	5	5	2	2	8	529	57.9089
P05556-2	ITGB1	Integrin beta-1	2	2	1	2	1	2	6	789	87.3886
P02748	C9	Complement component C9	1	1	1	1	1	1	3	559	63.1327
P02538	KRT6A	Keratin, type II cytoskeletal 6A	9	12	6	7	5	6	25	564	60.0083
O15031	PLXNB2	Plexin-B2	1	1	2	2	3	5	8	1838	204.997

Table S3. 14 proteins identified in both MCF7 and HT29.

Uniprot Accession	Gene name	Protein Description
Q9UJZ1-2	STOML2	Stomatin-like protein 2, mitochondrial
Q9H223	EHD4	EH domain-containing protein 4
Q99623-2	PHB2	Prohibitin-2
Q12905	ILF2	Interleukin enhancer-binding factor 2
Q02413	DSG1	Desmoglein-1
Q01082-3	SPTBN1	Spectrin beta chain, non-erythrocytic 1
P50851-2	LRBA	Lipopolysaccharide-responsive and beige-like anchor protein
P35221	CTNNA1	Catenin alpha-1
P25705	ATP5A1	ATP synthase subunit alpha, mitochondrial
P18206-2	VCL	Isoform 1 of Vinculin
P16144-4	ITGB4	Integrin beta-4
P11413	G6PD	Glucose-6-phosphate 1-dehydrogenase
P08195-2	SLC3A2	4F2 cell-surface antigen heavy chain
P05556-2	ITGB1	Integrin beta-1

References

(1) Wen, L.; Zheng, Y.; Jiang, K.; Zhang, M.; Kondengaden, S. M.; Li, S.; Huang, K.; Li, J.; Song, J.; Wang, P. G. Two-Step Chemoenzymatic Detection of N-Acetylneuraminic Acid-alpha(2-3)-Galactose Glycans. *J. Am. Chem. Soc.* **2016,** *138*, 11473-11476.

(2) Mbua, N. E.; Li, X.; Flanagan-Steet, H. R.; Meng, L.; Aoki, K.; Moremen, K. W.; Wolfert, M. A.; Steet, R.; Boons, G. J. Selective exo-enzymatic labeling of N-glycans on the surface of living cells by recombinant ST6Gal I. *Angew. Chem. Int. Ed.* **2013,** *52*, 13012-13015.

(3) Laughlin, S. T.; Bertozzi, C. R. Metabolic labeling of glycans with azido sugars and subsequent glycan-profiling and visualization via Staudinger ligation. *Nat. Protoc.* **2007,** *2*, 2930-2944.

(4) Bateman, L. A.; Zaro, B. W.; Chuh, K. N.; Pratt, M. R. N-Propargyloxycarbamate monosaccharides as metabolic chemical reporters of carbohydrate salvage pathways and protein glycosylation. *Chem. Commun.* **2013,** *49*, 4328-4330.

(5) Yu, H.; Chokhawala, H.; Karpel, R.; Yu, H.; Wu, B. Y.; Zhang, J. B.; Zhang, Y. X.; Jia, Q.; Chen, X. A multifunctional Pasteurella multocida sialyltransferase: A powerful tool for the synthesis of sialoside libraries. *J. Am. Chem. Soc.* **2005,** *127*, 17618-17619.

(6) Yu, H.; Yu, H.; Karpel, R.; Chen, X. Chemoenzymatic synthesis of CMP-sialic acid derivatives by a one-pot two-enzyme system: comparison of substrate flexibility of three microbial CMP-sialic acid synthetases. *Bioorg. Med. Chem.* **2004,** *12*, 6427-6435.

(7) Wen, L.; Huang, K.; Wei, M.; Meisner, J.; Liu, Y.; Garner, K.; Zang, L.; Wang, X.; Li, X.; Fang, J.; Zhang, H.; Wang, P. G. Facile Enzymatic Synthesis of Ketoses. *Angew. Chem. Int. Ed.* **2015,** *54*, 12654-12658.

(8) Wen, L. Q.; Huang, K.; Liu, Y. P.; Wang, P. G. Facile Enzymatic Synthesis of Phosphorylated Ketopentoses. *ACS Catal.* **2016,** *6*, 1649-1654.

(9) Jarvis, D. L. Baculovirus-insect cell expression systems. *Methods Enzymol.* **2009,** *463*, 191-222.

(10) Yu, H.; Chokhawala, H.; Karpel, R.; Yu, H.; Wu, B.; Zhang, J.; Zhang, Y.; Jia, Q.; Chen, X. A multifunctional Pasteurella multocida sialyltransferase: a powerful tool for the synthesis of sialoside libraries. *J. Am. Chem. Soc.* **2005,** *127*, 17618-17619.

(11) Yu, H.; Huang, S.; Chokhawala, H.; Sun, M.; Zheng, H.; Chen, X. Highly efficient chemoenzymatic synthesis of naturally occurring and non-natural alpha-2,6-linked sialosides: a P. damsela alpha-2,6-sialyltransferase with extremely flexible donor-substrate specificity. *Angew. Chem. Int. Ed.* **2006,** *45*, 3938-3944.

(12) Yu, H.; Thon, V.; Lau, K.; Cai, L.; Chen, Y.; Mu, S. M.; Li, Y. H.; Wang, P. G.; Chen, X. Highly efficient chemoenzymatic synthesis of beta 1-3-linked galactosides. *Chem. Commun.* **2010,** *46*, 7507-7509.