Supporting Information

Enzymatic Synthesis of UDP-GlcNAc/UDP-GalNAc Analogs Using *N*-acetylglucosamine 1-phosphate uridyltransferase (GlmU)

Wanyi Guan, Li Cai, Junqiang Fang, Baolin Wu, Peng George Wang*

Contents:

A. Materials and Chemical Reagents. *Pfu* DNA polymerase, restriction enzymes and T4 DNA ligase were purchased from Fermentas (Glen Burnie, MD). Plasmid Mini Kit I and Gel Extraction Kit were purchased from Omega Bio-Tek (Norcross, GA). HisTrap HP column (5 mL) and HiTrap Desalting column (5 mL) were purchased from GE healthcare (Piscataway, NJ). Isopropyl-1-thio-β-D-galactopyranoside (IPTG) was purchased from Acros Organics (Geel, Belgium). Uridine 5′-triphosphate disodium salt hydrate (UTP), *p*-Nitrophenyl α-D-galacto-pyranoside, imidazole, ammonium bicarbonate, DEAE-cellulose were purchased from Sigma-Aldrich (St. Louis, MO). Bio-Gel P-2 Gel was purchased from Bio-rad Laboratories (Hercules, CA). Amicon Ultra 10,000 MWCO was purchased from Millipore (Billerica, MA).

B. Bacterial Strains and Plasmids. *Escherichia coli* DH5α [*lacZDM15 hsdR recA*] was purchased from Gibco-BRL (Gaithersburg, MD). *E. coli* BL21 (DE3) [F^- *ompT hsd*S_B(r^-_B m^-_B) *gal dcm* (DE3)] and plasmid pET15b were purchased from Novagen (Carlsbad, CA). *E. coli* K12 (substrain MG1655) was from ATCC (#47076).

C. Construction of Recombinant Plasmid. The *glmU* gene was amplified from genomic DNA of *Escherichia coli* K12 by polymerase chain reaction (PCR) using primers: 5'-ACTG**CATATG**TTGAATAATGCTATGAGCG-3' and 5'-CTGAC**GGATCC**TCACTTTTTCTTTACCGGACG-3'. The PCR fragment was digested with *Bam*HI and *Nde*I (sites in bold) and inserted into the corresponding sitesplasmid pET-15b. The recombinant plasmid was transformed into *E. coli* DH5α cells for DNA sequencing. After sequence confirmation, the plasmid was transformed into *E. coli* BL21 (DE3) for protein expression.

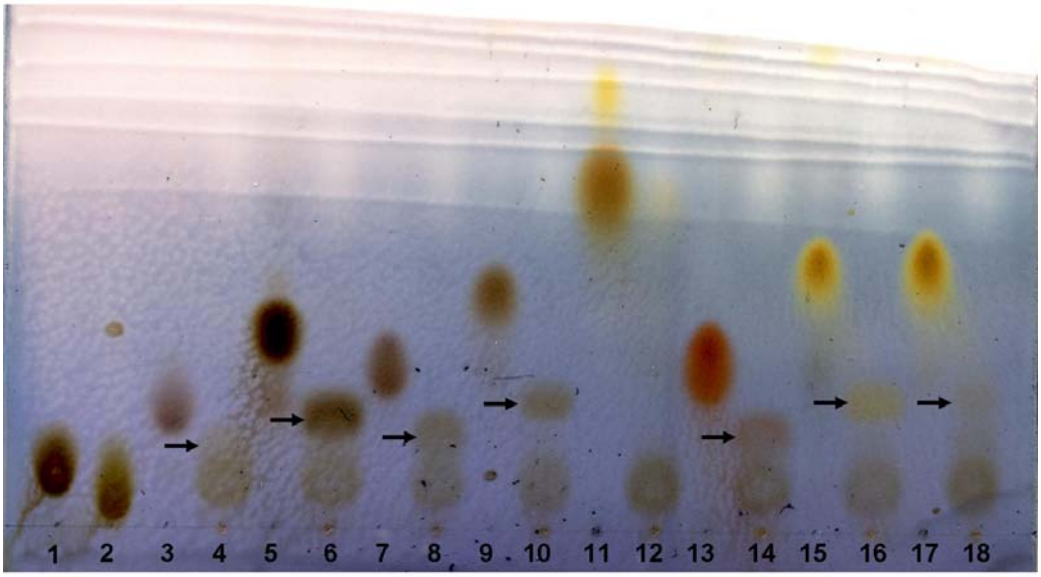
D. Overexpression and Purification of GlmU. *E. coli* BL21 (DE3) transformant containing the pET15b-*glmU* plasmid was grown in LB-ampicillin medium at 37 ^oC until OD_{600} reached 0.7. IPTG was added to a final concentration of 0.1 mM to induce protein to express at 16 ^oC for 20 h. The cells were harvested by centrifugation at 4,000 × *g* for 30 min. The cell pellet was resuspended in 20 mM sodium phosphate (pH 7.4) and sonicated on ice. The lysate was clarified by centrifugation at 20,000 × *g* for 40 min at 4 ^oC and the supernatant was applied to HisTrap^{TM} HP column. The column was then washed with washing buffer (20 mM sodium phosphate, pH 7.4, 100 mM NaCl, 10 mM imidazole) and the enzyme was eluted with elution buffer (20 mM sodium phosphate, pH 7.4, 100 mM NaCl, 500 mM imidazole). The eluant was desalted by HiTrap^{TM} Desalting HP column with 50 mM Tris-HCl buffer, pH 7.5. The fractions containing the purified enzyme were combined and concentrated with Amicon Ultra and the enzyme was stored in 20% glycerol (50 mM Tris-HCl, pH 7.5) at -20 ^oC.

E. Activity Measurement. The activity assay of recombinant GlmU was carried out in a mixture containing 100 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 5 mM GlcNAc-1-P, 5 mM UTP, and various amount of GlmU incubated at 37 ^oC. The reaction was quenched by boiling the mixture for 2 min followed by centrifugation. The supernatant was analyzed by thin-layer chromatography [n-butanol/acetic acid/water = 2:1:1 (v/v/v)].

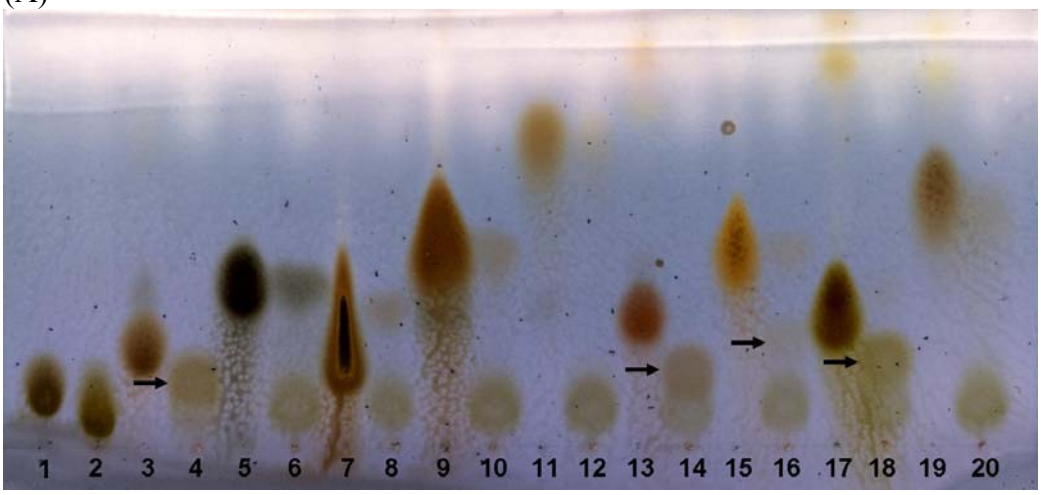
F. Production of UDP-GlcNAc/UDP-GalNAc and their analogs. The reaction mixture (4 mL) contained 10 mM GlcNAc1-P/GalNAc-1-P or their analogs, 15 mM UTP, 10 mM $MgCl_2$, and 1 mg/mL GlmU in 100 mM Tris-HCl buffer (pH 7.5). After incubation at 37 ^oC for 12 h, the reaction mixture was monitored by TLC (**Fig. S1**). The solvent was removed by lyophilization and the residue was diluted with ammonium bicarbonate solution (final concentration 50 mM). The mixture was applied to a column (10×170 mm) of DEAE-cellulose (HCO_3^- form) equilibrated with 50 mM ammonium bicarbonate. The column was eluted with a linear gradient of 50–300 mM NH_4HCO_3. The fractions containing the products were collected and concentrated by lyophilization. The residue was then dissolved in ddH_2O (1 mL) and applied to a Bio-Gel P-2 column. ddH_2O was used as eluent, and fractions containing the products were collected and lyophilized.

G. Enzymatic Synthesis of Trisaccharide and Disaccharide using LgtA. For trisaccharide (6-deoxy-GlcNAcβ1-3Lac-1-OBn) synthesis, 3.5 mg Lac-1-OBn, 5 mg UDP-6-deoxy-GlcNAc, and 10 mM $MnCl_2$ were incubated with 0.5 mg/mL LgtA in 100 mM Tris-HCl (pH 7.5) in a total volume of 1.6 mL at room temperature for 16 h. The synthesis of 6-deoxy-GlcNAcβ1-3Gal*p*NO_2Ph was performed in a 1 mL mixture containing 5 mM UDP-6-deoxy-GlcNAc, 5 mM *p*-nitrophenyl-α-D-galactopyranoside, 10 mM $MnCl_2$, and 0.5 mg/mL LgtA in 100 mM Tris-HCl (pH 7.5). The reaction was incubated at room temperature for 16 h. After completion, the mixtures were briefly boiled for 2 min and then centrifuged to remove protein. The supernatant was concentrated under reduced pressure and the residue was applied to a Bio-Gel P-2 column (1.5 × 90 cm) using ddH_2O as eluant. Fractions containing products were monitored by TLC and MS and pooled. LgtA was overexpressed with vector pMCSG7 in *E. coli* BL21 (DE3).

H. Physical Properties and Spectroscopic Measurements. Proton nuclear magnetic resonance (^1H NMR) spectra and carbon-13 (^{13}C) NMR spectra were recorded on a Bruker DPX 400 spectrometer at 400 MHz and 100 MHz respectively, or on a Bruker DRX 500 spectrometer at 500 MHz and 125 MHz respectively. Chemical shifts and coupling constants are reported in ppm and Hz respectively. The high-resolution mass spectra were recorded on a Bruker MicrOTOF spectrometer. Analytical TLC was carried out on silica gel 60 F254 aluminum-backed plates (E. Merck).



(A)



(B)

Figure S1 (A) *In vitro* GlmU reaction (after 12 hrs) with GlcNAc-1-P derivatives as substrates. Lane 1: UDP-GlcNAc; lane 2: UTP; lanes 3&4: GlcNAc-1-P and reaction (compound **1**); lanes 5&6: GlcNAz-1-P and reaction (compound **2**); lanes 7&8: GlcNPr-1-P and reaction (compound **3**); lanes 9&10: GlcNBu-1-P and reaction (compound **4**); lanes 11&12: GlcNBz-1-P and reaction (compound **5**, very low yield); lanes 13&14: 6-deoxyGlcNAc-1-P and reaction (compound **6**); lanes 15&16: 6-N_3GlcNAc-1-P and reaction (compound **7**); lanes 17&18: AllNAc-1-P and reaction (compound **8**). Arrows point to product spots. n-BuOH:HAc:H_2O 2:1:1. (B) *In vitro* GlmU reaction (after 12 hrs) with GalNAc-1-P derivatives as substrates. Lane 1: UDP-GlcNAc; lane 2: UTP; lanes 3&4: GalNAc-1-P and reaction (compound **9**); lanes 5&6: GalNAz-1-P and reaction (compound **10**); lanes 7&8: GalNPr-1-P and reaction (compound **11**); lanes 9&10: GalNBu-1-P and reaction (compound **12**); lanes 11&12: GalNBz-1-P and reaction (compound **13**); lanes 13&14: 6-deoxyGalNAc-1-P and reaction (compound **14**); lanes 15&16: 6-N_3GalNAc-1-P and reaction (compound **15**); lanes 17&18: 4-deoxyGalNAc-1-P and reaction (compound **16**); lanes 19&20: 4-N3GalNAc-1-P and reaction (compound **17**). Arrows point to product spots. n-BuOH:HAc:H_2O 2:1:1.

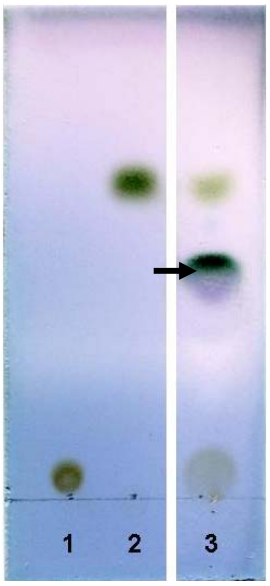


Figure S2 TLC for *in vitro* synthesis of disaccharide. Lane 1. UDP-6DeOGlcNAc (comp. **6**); Lane 2. *p*-nitrophenyl α-D-galacto-pyranoside; Lane 3. reaction mixture of disaccharide synthesis with UDP-6DeOGlcNAc (comp. **6**) as sugar donor. Arrow points to product spot. n-BuOH:HAc:H_2O 2:1:1.

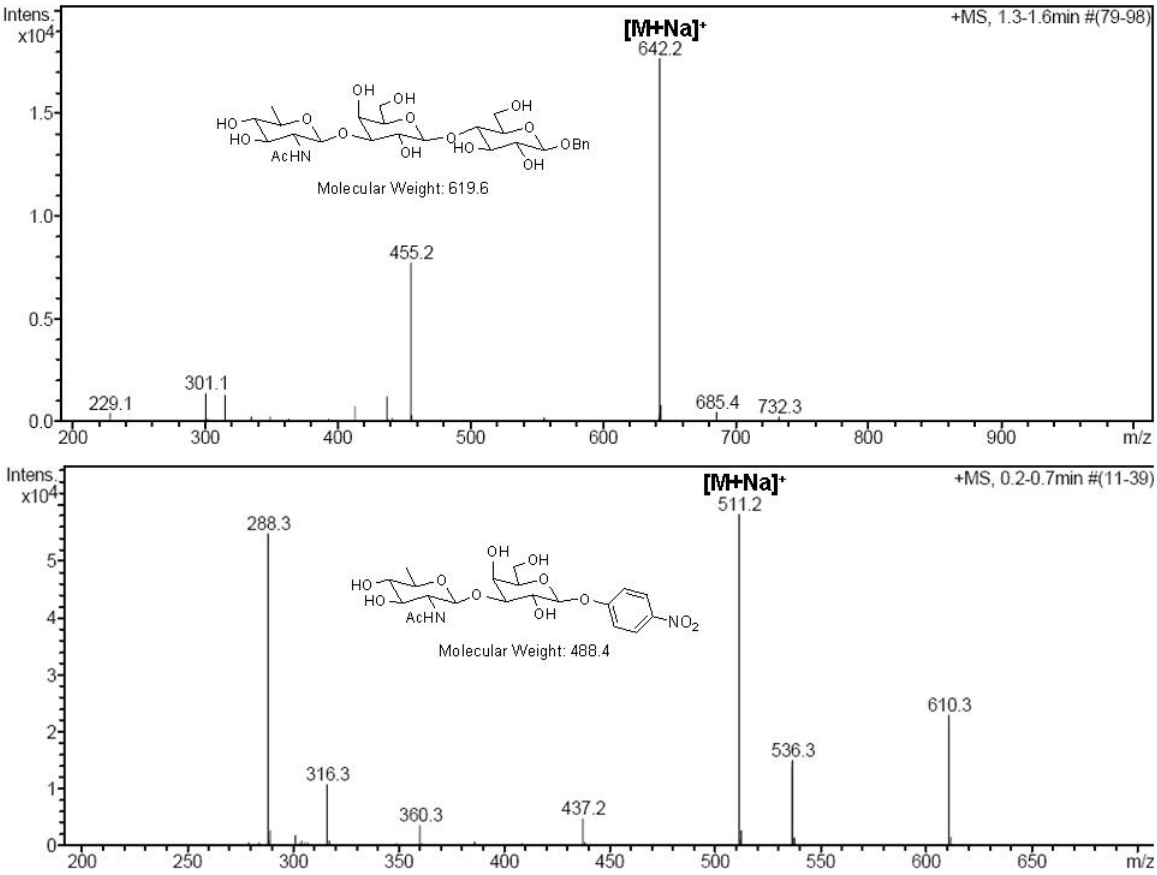
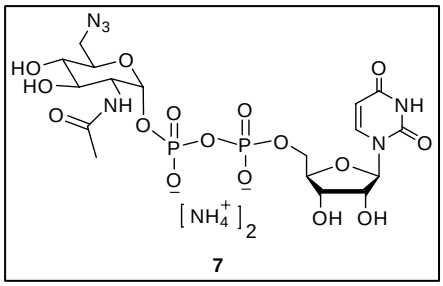


Figure S3 Mass spectra of reaction mixture of enzymatic synthesis of trisaccharide and disaccharide using glycosyltransferase LgtA.

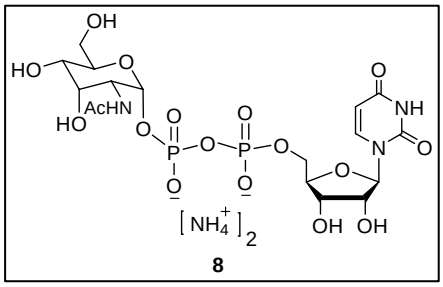
HRMS (trisaccharide upper one) calcd for $C_{27}H_{41}NO_{15}Na$ $(M+Na)^+$ 642.2368, found 642.2343 m/z.

HRMS (disaccharide lower one) calcd for $C_{20}H_{28}N_2O_{12}Na$ $(M+Na)^+$ 511.1534, found 511.1552 m/z.



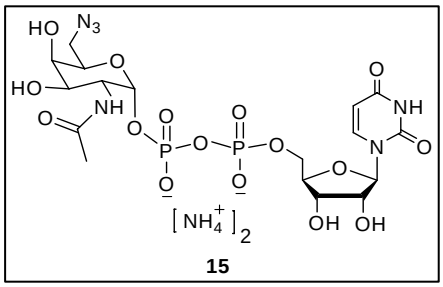
7

Uridine 5′-diphospho-2-acetamido-6-azido-2,6-dideoxy-α-D-glucopyranose diammonium salt: ^{1}H NMR (400 MHz, D_2O): δ = 7.98 (d, *J* = 8.1 Hz, 1H), 5.97-5.99 (m, 2H), 5.51 (dd, *J* = 7.2, 3.3 Hz. 1H), 4.36-4.40 (m, 2H), 4.20-4.30 (m, 3H), 4.00-4.09 (m, 2H), 3.81 (t, *J* = 9.2 Hz, 1H), 3.74 (dd, *J* = 13.5, 2.5 Hz, 1H), 3.57-3.65 (m, 2H), 2.08 (s, 3H); ^{31}P NMR (162 MHz, D_2O): δ = -11.3 (d, *J* = 21.4 Hz), -13.2 (d, *J* = 21.4 Hz); HRMS (ESI) calcd for $C_{17}H_{25}N_6O_{16}P_2$ $(M-H)^-$ 631.0808, found 631.0806 *m*/*z*.



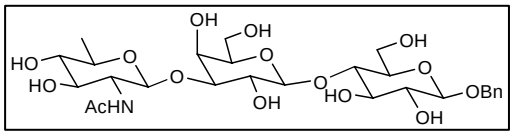
8

Uridine 5′-diphospho-2-acetamido-2-deoxy-α-D-allopyranose diammonium salt: ^{1}H NMR (400 MHz, D_2O): δ = 7.96 (d, *J* = 8.1 Hz, 1H), 5.96-6.00 (m, 2H), 5.53 (dd, *J* = 7.1, 3.8 Hz. 1H), 4.36-4.40 (m, 2H), 4.21-4.30 (m, 3H), 4.08-4.15 (m, 3H), 3.92 (dd, *J* = 12.6, 2.3 Hz, 1H), 3.82 (dd, *J* = 12.6, 4.5 Hz, 1H), 3.77 (dd, *J* = 10.5, 3.2 Hz, 1H), 2.10 (s, 3H); ^{31}P NMR (162 MHz, D_2O): δ = -11.0 (d, *J* = 19.9 Hz), -13.0 (d, *J* = 21.7 Hz); HRMS (ESI) calcd for $C_{17}H_{26}N_3O_{17}P_2$ $(M-H)^-$ 606.0743, found 606.0730 *m*/*z*.

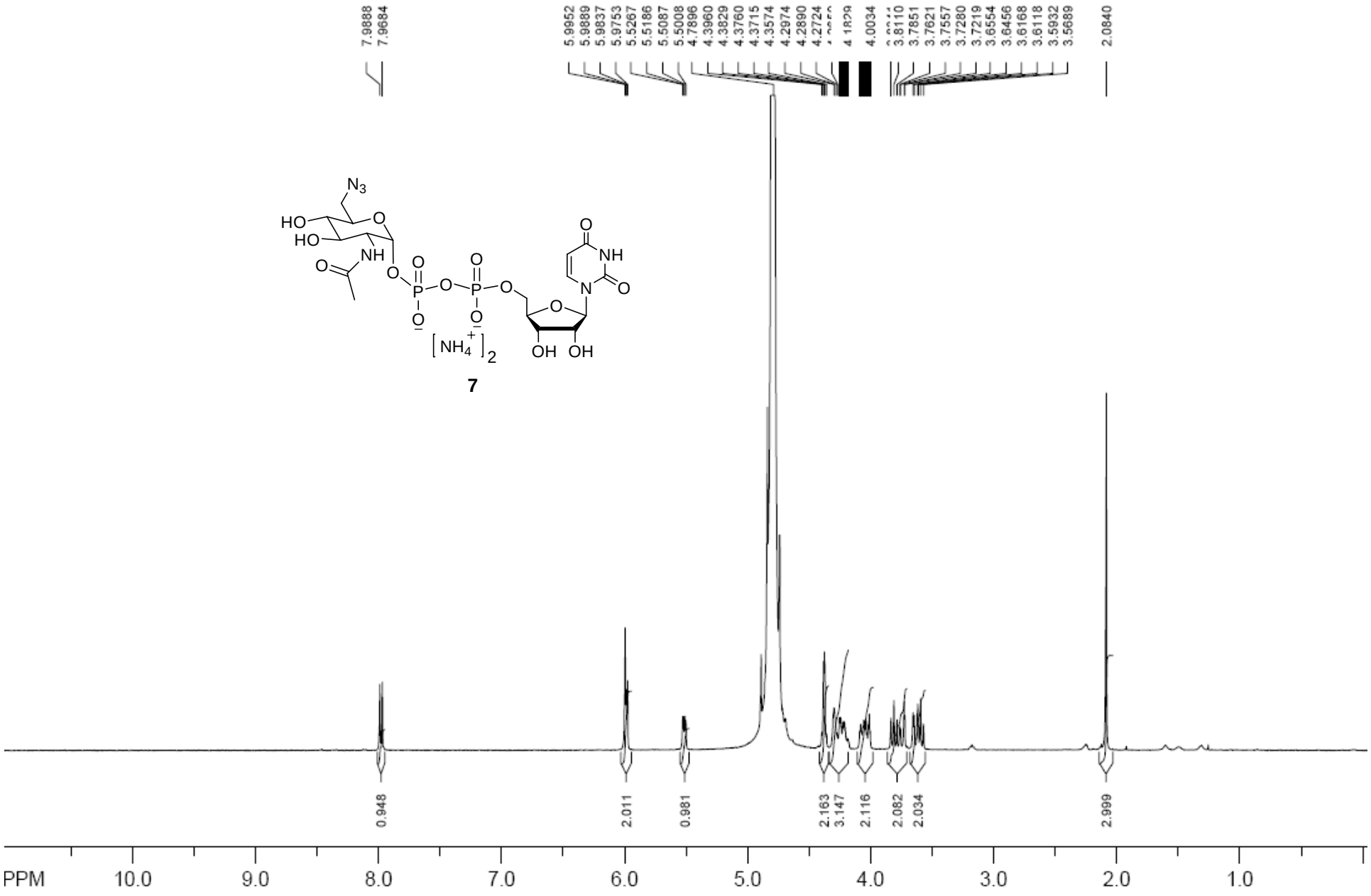


15

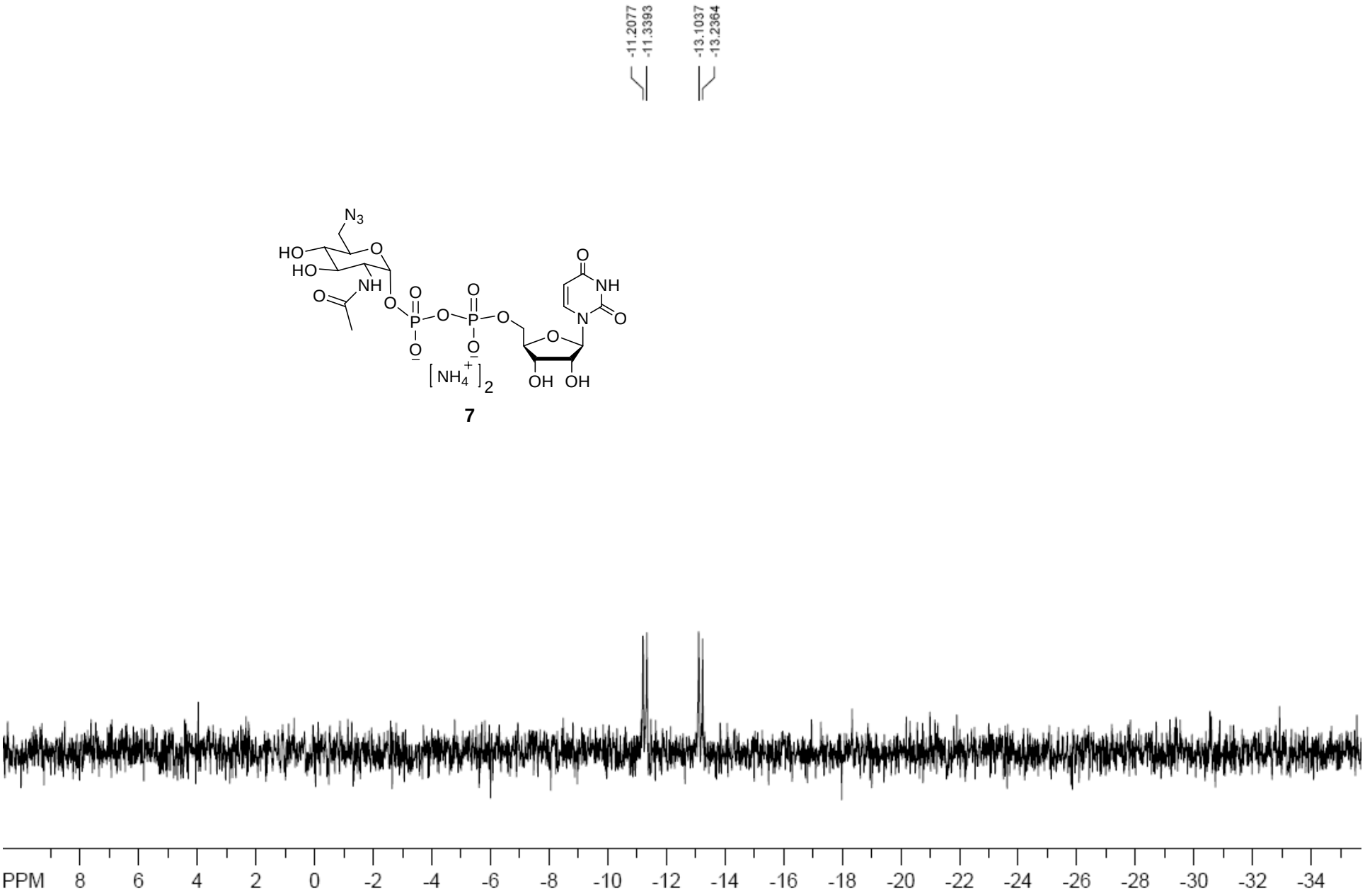
Uridine 5′-diphospho-2-acetamido-6-azido-2,6-dideoxy-α-D-galactopyranose diammonium salt: ^{1}H NMR (400 MHz, D_2O): δ = 7.98 (d, *J* = 8.2 Hz, 1H), 5.97-5.99 (m, 2H), 5.54 (dd, *J* = 7.2, 3.4 Hz. 1H), 4.35-4.39 (m, 2H), 4.19-4.30 (m, 5H), 4.02 (m, 1H), 3.98 (dd, *J* = 10.8, 3.2 Hz, 1H), 3.60 (dd, *J* = 12.8, 7.2 Hz, 1H), 3.50 (dd, *J* = 12.8, 6.0 Hz, 1H), 2.09 (s, 3H); ^{31}P NMR (162 MHz, D_2O): δ = -11.2 (d, *J* = 21.5 Hz), -13.0 (d, *J* = 21.5 Hz); HRMS (ESI) calcd for $C_{17}H_{25}N_6O_{16}P_2$ $(M-H)^-$ 631.0808, found 631.0801 *m*/*z*.



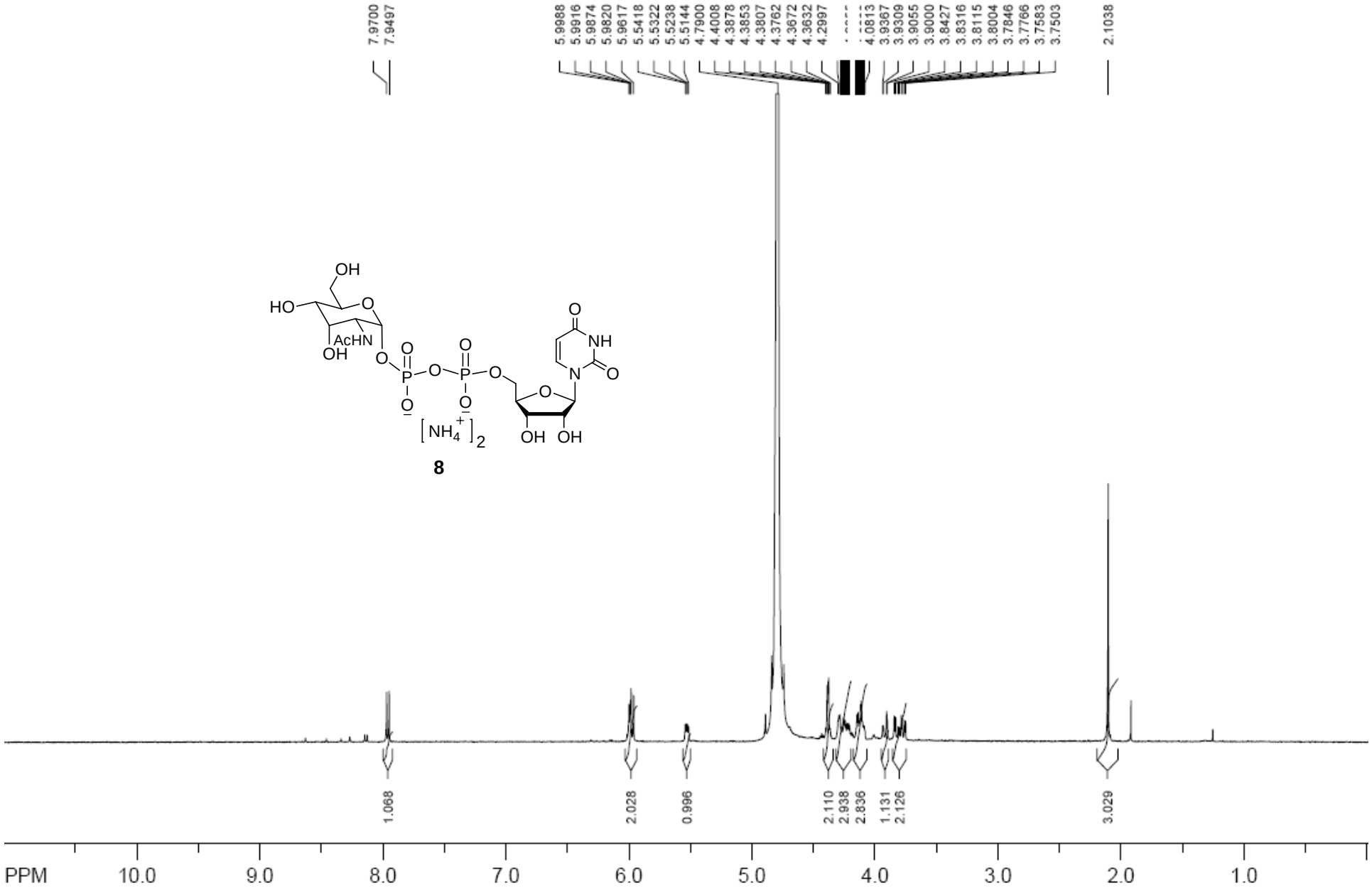
Benzyl-4-O-(3-O-(2-acetamido-2,6-dideoxy-β-D-glucopyranosyl)-β-D-galactopyranosyl)-β-D-glucopyranoside: ^{1}H NMR (500 MHz, D_2O): δ = 7.46-7.54 (m, 5H, arom.), 4.83 (s, 2H), 4.80 (1H, in solvent), 4.73 (d, *J* = 8.6 Hz, 1H), 4.48 (dd, *J* = 7.9, 2.1 Hz, 1H), 4.15 (app t, *J* = 2.3 Hz, 1H), 3.62-3.88 (m ,11H), 3.51-3.58 (m, 2H), 3.27 (t, *J* = 9.3 Hz, 1H), 2.09 (s, 3H, CH_3CO), 1.36 (d, *J* = 6.2 Hz, 3H, CH_3); HRMS (ESI) calcd for $C_{27}H_{41}NO_{15}Na$ $(M+Na)^+$ 642.2368, found 642.2343 m/z.


7.9888
7.9684
5.9952
5.9889
5.9837
5.9753
5.5267
5.5186
5.5087
5.5008
4.7896
4.3960
4.3829
4.3760
4.3715
4.3574
4.2974
4.2890
4.2724
4.1829
4.0034
3.8110
3.7851
3.7621
3.7557
3.7280
3.7219
3.6554
3.6456
3.6168
3.6118
3.5932
3.5689
2.0840
0.948
2.011
0.981
2.163
3.147
2.116
2.082
2.034
2.999
PPM
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0

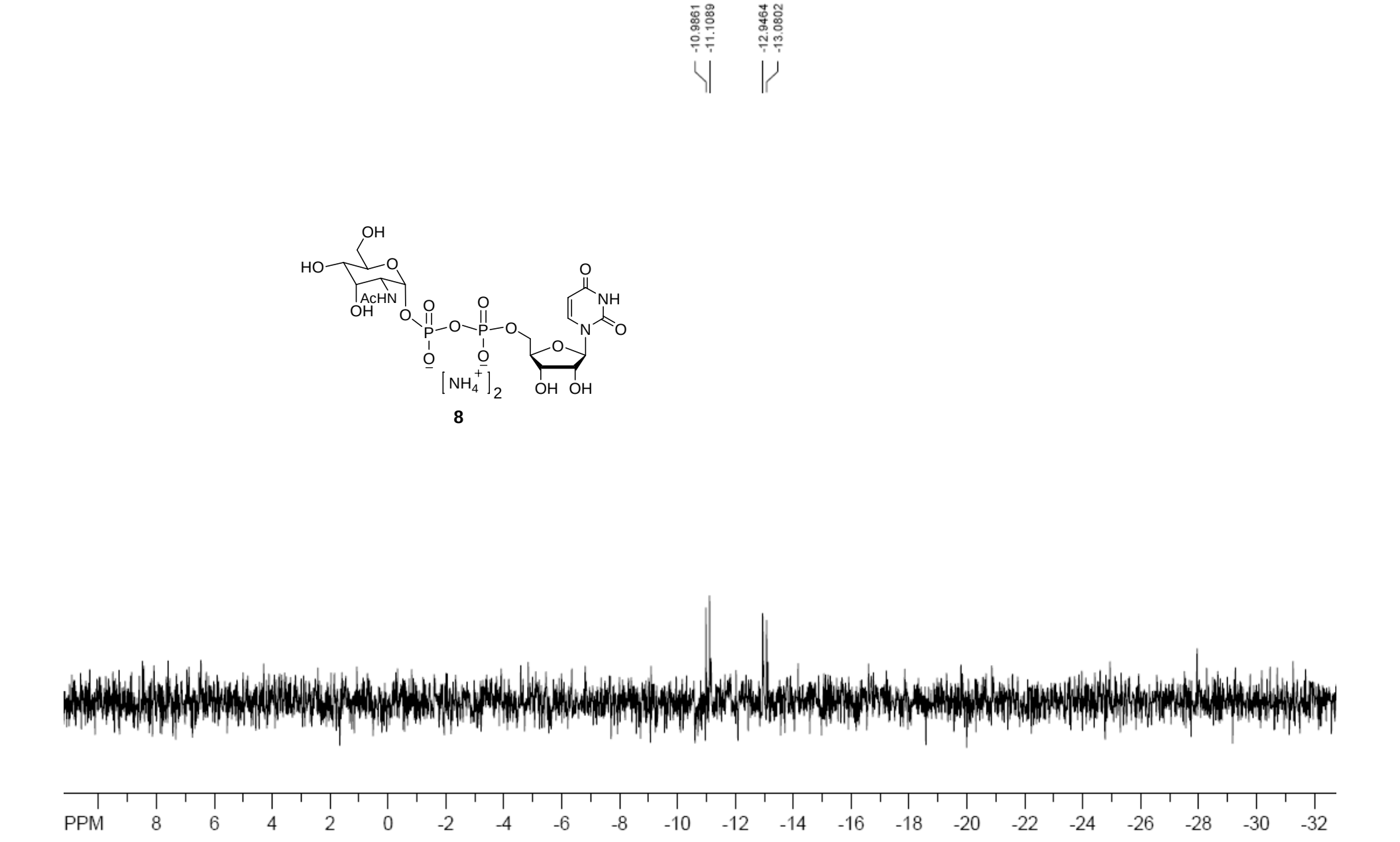
7


-11.2077
-11.3393
-13.1037
-13.2364
N3
HO
HO
O
NH
O
O
P
O
O
P
O
O
O
O
O
NH
N
O
O
OH
OH
[NH4+]2
PPM
8
6
4
2
0
-2
-4
-6
-8
-10
-12
-14
-16
-18
-20
-22
-24
-26
-28
-30
-32
-34

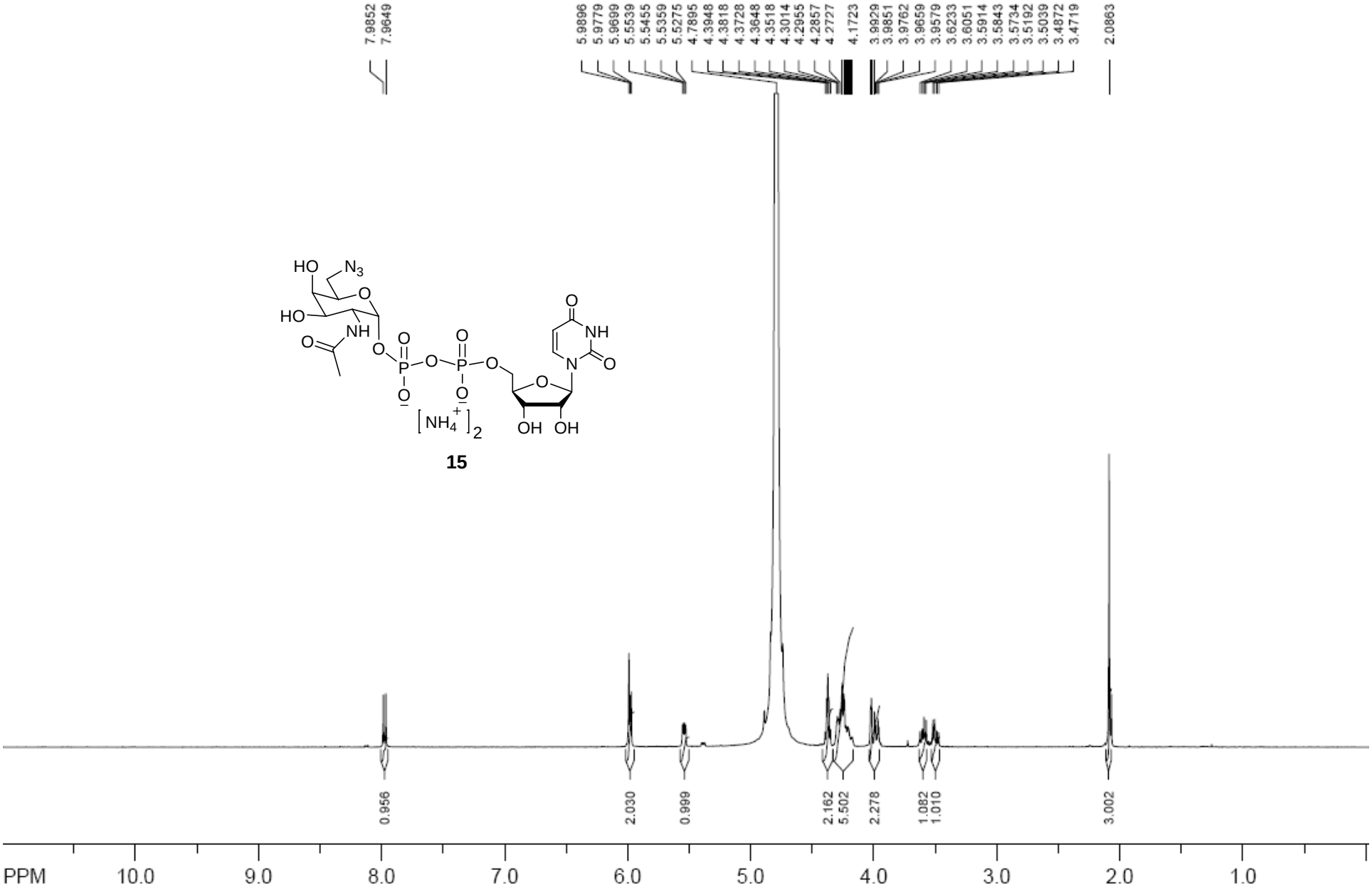
7


7.9700
7.9497
5.9988
5.9916
5.9874
5.9820
5.9617
5.5418
5.5322
5.5238
5.5144
4.7900
4.4008
4.3878
4.3853
4.3807
4.3762
4.3672
4.3632
4.2997
4.0813
3.9367
3.9309
3.9055
3.9000
3.8427
3.8316
3.8115
3.8004
3.7846
3.7766
3.7583
3.7503
2.1038
1.068
2.028
0.996
2.110
2.938
2.836
1.131
2.126
3.029
PPM
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0

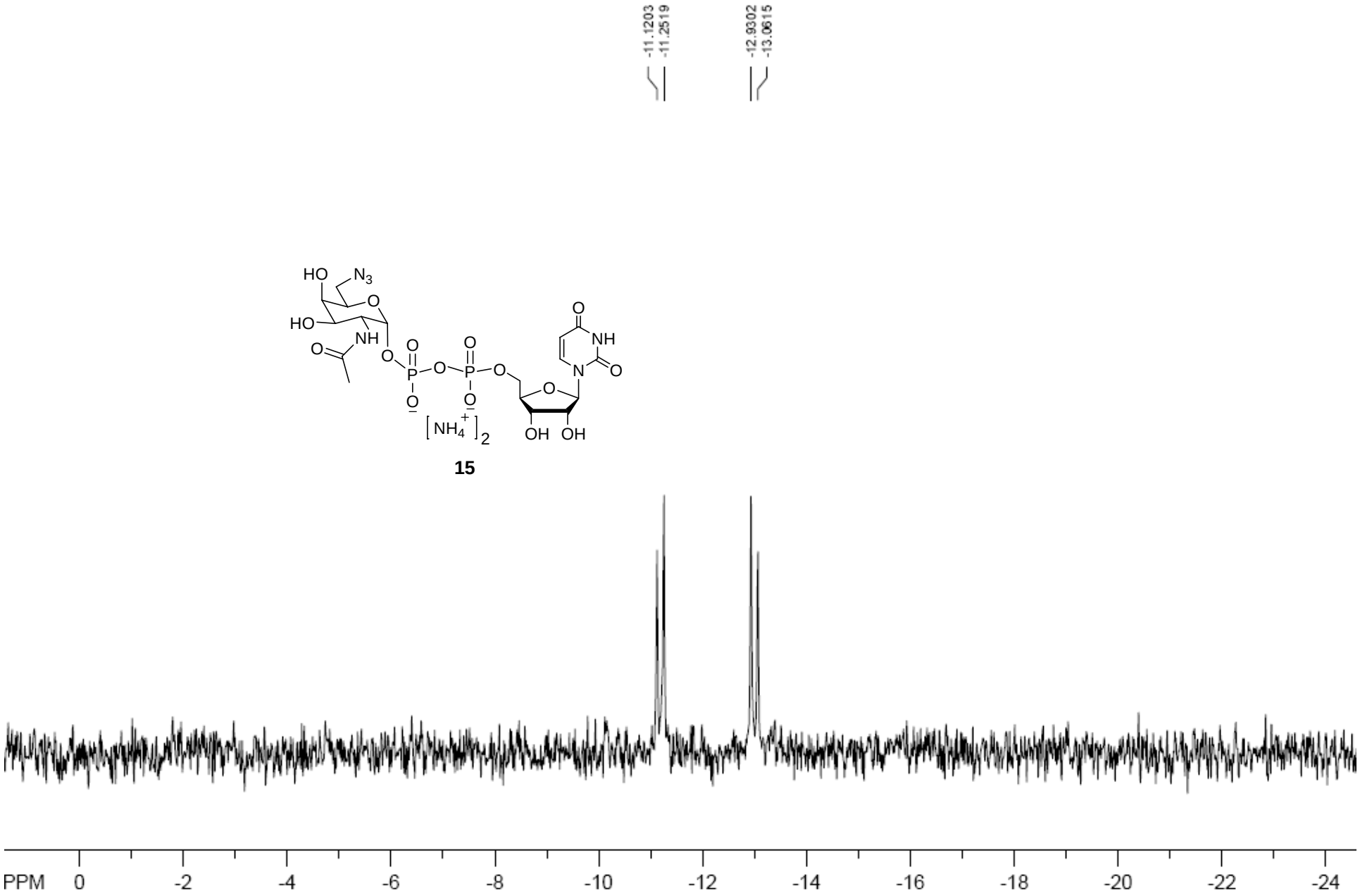
8


-10.9861
-11.1089
-12.9464
-13.0802
OH
HO
O
AcHN
OH
O
O
P
O
O
O
P
O
O
O
O
NH
O
N
O
O
[NH4+]2
OH
OH
PPM
8
6
4
2
0
-2
-4
-6
-8
-10
-12
-14
-16
-18
-20
-22
-24
-26
-28
-30
-32

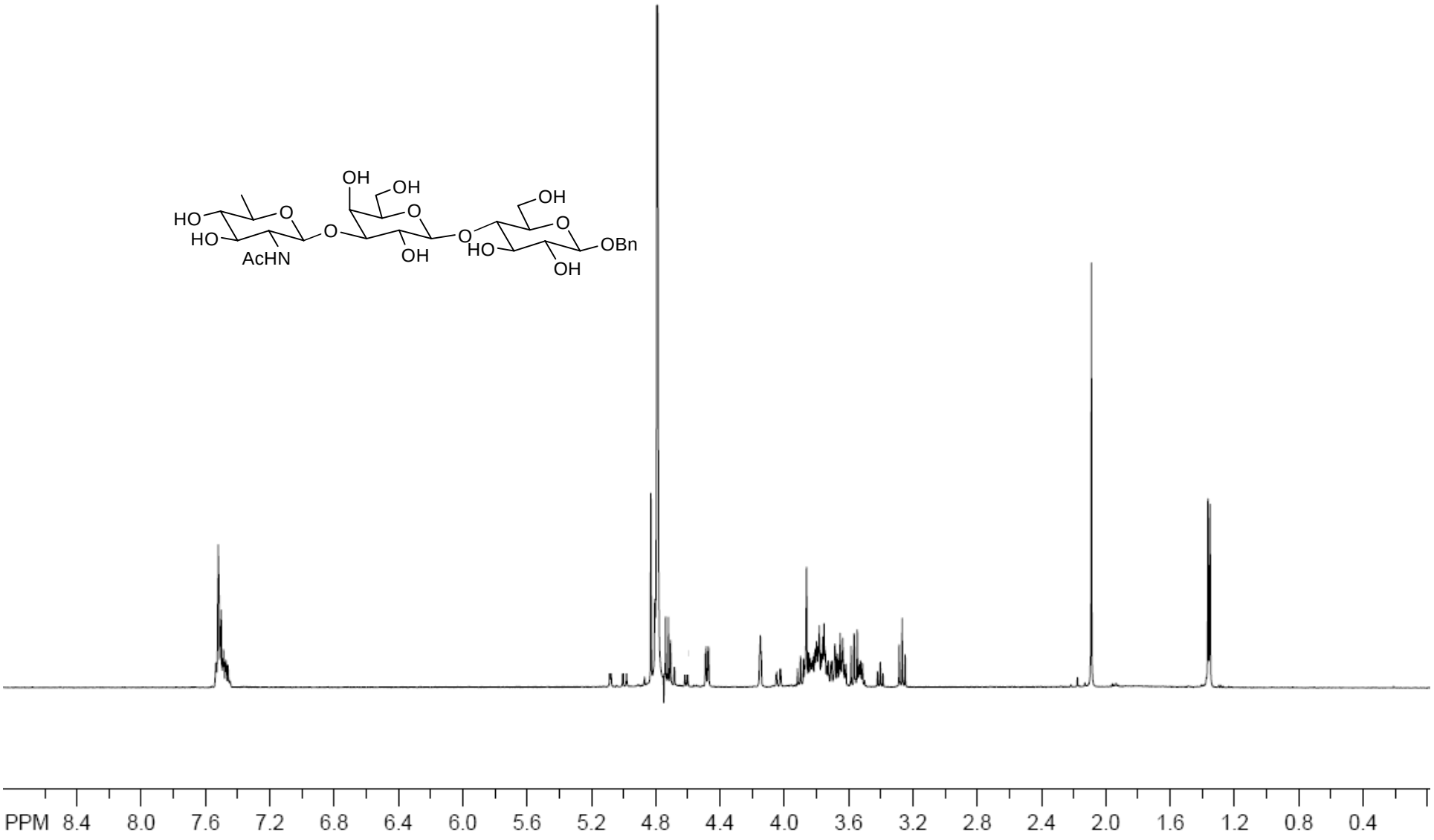
8


7.9852
7.9649
5.9896
5.9779
5.9699
5.5539
5.5455
5.5359
5.5275
4.7895
4.3948
4.3818
4.3728
4.3648
4.3518
4.3014
4.2955
4.2857
4.2727
4.1723
3.9929
3.9851
3.9762
3.9659
3.9579
3.6233
3.6051
3.5914
3.5843
3.5734
3.5192
3.5039
3.4872
3.4719
2.0863
0.956
2.030
0.999
2.162
5.502
2.278
1.082
1.010
3.002
PPM
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0

15


-11.1203
-11.2519
-12.9302
-13.0615
HO
N3
O
HO
NH
O
O
P
O
O
O
P
O
O
O
O
NH
O
N
O
O
OH
OH
[NH4+]2
PPM
0
-2
-4
-6
-8
-10
-12
-14
-16
-18
-20
-22
-24

15


HO
HO
AcHN
O
O
OH
OH
O
OH
O
HO
OH
OH
O
OBn
PPM 8.4 8.0 7.6 7.2 6.8 6.4 6.0 5.6 5.2 4.8 4.4 4.0 3.6 3.2 2.8 2.4 2.0 1.6 1.2 0.8 0.4