Electronic Supplementary Material (ESI) for RSC Advances.
This journal is © The Royal Society of Chemistry 2020

Supporting Information

for

Exploring the Broad Nucleotide Triphosphate and Sugar-1-phosphate Specificities of Thymidylyltransferase Cps23FL from *Streptococcus pneumonia* Serotype 23F

Siqiang Li,[a,b,§] Hong Wang,[a,§] Guoxia Jin,[c] Zonggang Chen,[a] and Guofeng Gu[a,*]

[a]National Glycoengineering Research Center and Shandong Key Laboratory of Carbohydrate Chemistry and Glycobiology, Shandong University, 72 Binhai Road, Qingdao 266237, China

[b]School of Biological and Food Processing Engineering, Huanghuai University, 76 Kaiyuan Road, Zhumadian 463000, China

[c]College of Chemistry, Chemical Engineering and Materials Science, Shandong Normal University, 88 Wenhua Dong Lu, Jinan 250014, China.

*Corresponding Author. +86-532-5863 1408; E-mail: guofenggu@sdu.edu.cn

[§]These authors contributed equally to this paper.

Table of Contents

I. Experimental Section

Materials. *E. coli* harboring *cps23FL*-pET28a was from our lab. BCA protein quantitation kit from Cwbiotech, glactose-1-phosphate (Gal-1-P), glucosamine-1-phosphate ($GlcNH_2$-1-P), galactosamine-1-phosphate ($GalNH_2$-1-P), galacturonic acid-1-phosphate (GalA-1-P) and inorganic pyrophosphatease (YIPP) from Sigma, mannose-1-phosphate (Man-1-P) from Chemme, glucose-1-phosphate (Glc-1-P), 2'-deoxythymidine 5'-triphosphate (dTTP), uridine 5'-triphosphate (UTP), 2'-deoxyuridine 5'-triphosphate (dUTP), 2'-deoxycytidine 5'-triphosphate (dCTP), cytidine 5'-triphosphate (CTP), 2'-deoxyguanosine 5'-triphosphate (dGTP), guanosine 5'-triphosphate (GTP), 2'-deoxyadenosine 5'-triphosphate (dATP), adenosine 5'-triphosphate (ATP) from Sangonbiotech, yeast extract, agar and tryptone from OXOID were purchased, respectively. Other reagents were obtained from Sangonbiotech unless stated otherwise.

Expression and purification of Cps23FL. Expression and purification of Cps23FL were performed as previously reported (Li, S. et al., *Chem. Commun.* **2016**, 52(97): 13995). Specifically, a desired expression strain was directly incubated overnight in LB/Kan medium. After growth overnight (37 ºC, 200 revolutions/minute), this culture was diluted 1:100 with fresh LB/Kan medium typically to a total volume of 1 L. The large scale culture was subsequently grown (37 ºC, 200 revolutions/minute) to mid-log phase (OD600 ≈ 0.6), at which point IPTG was added at the final concentration of 1mM. After growth was continued for an additional 10 hours (25 ºC, 200 revolutions/minute), the cells were collected by centrifugation (10 minutes, 8000 g/minute) and suspended in100 ml of 50 mM Tris-HCl, at pH 7.5, containing 300 mM NaCl on the ice. The desired enzymes were purified with Ni^+ affinity chromatography (GE Healthcare). Enzymes concentrations were calculated by BCA protein quantitation kit, and enzymes were stored at -80 ºC before use.

Substrate specificity analysis of Cps23FL. Enzymatic reactions were performed under standard conditions in duplicates in a total 100 μL solution containing enzymes (6 or 60 μM), $MgCl_2$ (5 mM), YIPP (2 U/mL), sugar-1-P (5 mM), (d)NTP (3 mM) and HEPES (30 mM, pH 7.5). The reactions were performed at 37 ºC for 1, 6, or 48 h and terminated in boiling water bath for 30 s, followed by

centrifugated at 12 300 rpm for 10 minutes. The consumption of (d)NTP and the formation of sugar nucleotide during reactions were analyzed by analytical HPLC (Dionex CarboPac^{TM} PA-100 Analytical Column 4×250mm, 0～1.0 M ammonium acetate). Scale-up reactions were applied and then milligram products were purified by semi-preparative HPLC (Dionex CarboPac^{TM} PA-100, 22×250mm, 0～1.0 M ammonium acetate), then desalted via a size exclusion chromatography using Biogel P-2 column with distilled water as the eluents. Product-containing fractions were combined and lyophilized to give the desired target products, which were fully charaterized with high resolution mass spectrometry (ESI-HRMS, Shimadzu LCMS-IT-TOF mass spectrometer) and nuclear magnetic resonance (NMR, 600M Aglient) analysis of ^{1}H-NMR (600 MHz), ^{13}C-NMR (150 MHz), ^{31}P-NMR (243 MHz), ^{1}H-^{1}H COSY (600 MHz) and ^{1}H-^{13}C HSQC (600/150 MHz).

The optimal reactions were carried out by different concentrations of Cps23FL and of (d)NTP. Different concentrations of (d)NTP (0, 1.5, 3.0, 6.0, 12.0, 24.0 mM final) were added into reaction mixtures containing 30 mM HEPES at pH 7.5, 5 mM Glc-1-P, 5 mM $MgCl_2$, 2U/ml YIPP and 6 μM Cps23FL to optimize concentration of nucleoside triphosphates; whilst different concentrations of Cps23FL (0, 0.06, 0.6, 6, 30 and 60 μM final) was added into the reaction mixtures containing containing 30 mM HEPES at pH 7.5, 5 mM Glc-1-P, 5 mM $MgCl_2$, 2U/ml YIPP and 3 mM ATP to optimize concentration of Cps23FL. The reaction mixture was carried out at 37 °C for 6 h, and then quenched for HPLC analysis.

dUDP-Glc. Yield: 95%. ^{1}H NMR (600 MHz, D_2O): δ 7.80 (d, *J* = 8.4 Hz, 1H, -C*H*=), 6.18 (t, *J* = 6.6 Hz, 1H, H-1), 5.81 (d, *J* = 7.8 Hz, 1H, -C*H*=), 5.46 (dd, *J* = 7.2, 3.6 Hz, 1H, H-1'), 4.47 (m, 2H, H-5a,b), 4.08–4.00 (m, 3H, H-2, H-3, H-4), 3.78–3.73 (m, 1H, H-5'), 3.71 (dd, *J* = 12.6,2.4 Hz, 1H, H-6a'), 3.66–3.60 (m, 2H, H-3', H-6b'), 3.39 (dt, *J* = 10.2, 3.0 Hz, 1H, H-2'), 3.32 (t, *J* = 9.6 Hz, 1H, H-4'), 2.29–2.20 (m, 2H, H-2); ^{13}C NMR (150 MHz, D_2O): δ 166.25, 151.57, 141.81, 128.48, 102.42, 95.48 (d, *J* = 7.5 Hz, C-1'), 83.38 (d, *J* = 9.0 Hz, C-4), 85.32 (C-1), 72.76 (C-3'), 72.72 (C-5'), 71.49 (d, *J* = 9.0 Hz, C-2'), 70.77 (C-3), 69.09 (C-4'), 65.30 (d, *J* = 4.5 Hz, C-

5), 60.22 (C-6'), 38.76 (C-2); ^{31}P NMR (243 MHz, D_2O): δ -11.19 (d, *J* = 20.5 Hz), -12.85 (d, *J* = 20.5 Hz); ESI-TOF HRMS m/z calcd for $C_{15}H_{24}N_2O_{16}P_2$ [M-H] $^{-1}$ 549.0528, found 549.0445.

UDP-Glc. Yield: 94%. ^{1}H NMR (600 MHz, D_2O): δ 7.79 (d, *J* = 8.4 Hz, 1H, -C*H*=), 5.81 (s, 1H, H-1), 5.80 (d, *J* = 8.4 Hz, 1H, -C*H*=), 5.43 (dd, *J* = 7.2, 3.0 Hz, 1H, H-1'), 4.21 (br s, 2H, H-2,3), 4.14–4.00 (m, 3H, H-4, H-5a,b), 3.75–3.70 (m, 1H, H-5'), 3.69 (br d, *J* = 12.6 Hz, 1H, H-6a'), 3.64–3.57 (m, 2H, H-3', H-6b'), 3.49–3.34 (m, 1H, H-2'), 3.30 (t, *J* = 9.6 Hz, 1H, H-4'); ^{13}C NMR (150 MHz, D_2O): δ 166.15, 151.72, 141.50, 102.54, 95.48 (d, *J* = 6.0 Hz, C-1'), 88.27 (C-1), 83.10 (d, *J* = 9.6 Hz), 73.66, 72.74, 72.69, 71.48 (d, *J* = 9.0 Hz), 69.52, 69.04, 64.82 (d, *J* = 6.0 Hz), 60.18; ^{31}P NMR (243 MHz, D_2O): δ -11.33 (d, *J* = 20.5 Hz), -12.96 (d, *J* = 20.5 Hz); ESI-TOF HRMS m/z calcd for $C_{15}H_{23}N_2O_{17}P_2$ [M-H] $^{-1}$ 565.0477, found 565.0439.

dCDP-Glc. Yield: 93%. ^{1}H NMR (600 MHz, D_2O): δ 7.91 (d, *J* = 7.8 Hz, 1H, -C*H*=), 6.13 (t, *J* = 6.6 Hz, 1H, H-1), 6.03 (d, *J* = 7.8 Hz, 1H, -C*H*=), 5.42 (dd, *J* = 7.2, 3.0 Hz, 1H, H-1'), 4.43 (br s, 1H, H-3), 4.09–3.97 (m, 3H, H-4, H-5a,b), 3.75–3.70 (m, 1H, H-5'), 3.68 (d, *J* = 12.6 Hz, 1H, H-6a'), 3.63–3.56 (m, 2H, H-3´, H-6b'), 3.38–3.33 (m, 1H, H-2'), 3.29 (t, *J* = 9.6 Hz, 1H, H-4'), 2.32–2.25 (m, 1H, H-2a), 2.20–2.13 (m, 1H, H-2b); ^{13}C NMR (150 MHz, D_2O): δ 162.81, 142.75, 95.76, 95.46 (d, *J* = 6.0 Hz, C-1'), 86.18 (C-1), 85.64 (d, *J* = 9.0 Hz, C-4), 72.75 (C-3'), 72.67 (C-5'), 71.50 (d, *J* = 8.0 Hz, C-2'), 70.58 (C-3), 69.04 (C-4'), 65.09 (d, *J* = 6.0 Hz, C-5), 60.17 (C-6'), 39.41 (C-2); ^{31}P NMR (243 MHz, D_2O): δ -11.34 (d, *J* = 20.7 Hz), -12.99 (d, *J* = 20.7 Hz); ESI-TOF HRMS m/z calcd for $C_{15}H_{24}N_3O_{15}P_2$ [M-H]$^{-1}$ 548.0688, found 548.0616.

CDP-Glc. Yield: 87%.^{1}H NMR (600 MHz, D_2O): δ 7.83 (d, *J* = 7.2 Hz, 1H, -C*H*=), 5.98 (d, *J* = 7.8 Hz, 1H, -C*H*=), 5.80 (d, *J* = 4.2 Hz, 1H, H-1), 5.42 (dd, *J* = 7.2, 3.0 Hz, 1H, H-1'), 4.19-4.13 (m, 2H, H-2, H-3), 4.13-4.08 (m, 2H, H-4, H-5a), 4.05-4.00 (m, 1H, H-5b), 3.74–3.69 (m, 1H, H-5'), 3.67 (br d, *J* = 12.6 Hz, 1H, H-6a'), 3.62–3.56 (m, 2H, H-3', H-6b'), 3.37–3.32 (m, 1H, H-2'), 3.28 (t, *J* = 9.6 Hz, 1H, H-4'); ^{13}C NMR (150 MHz, D_2O): δ 164.75, 155.92, 141.78, 96.24, 95.49 (d, *J* = 7.5 Hz, C-1'), 89.19 (C-1), 82.63 (d, *J* = 9.0

Hz, C-4), 74.14 (C-2), 72.75 (C-3'), 72.69 (C-5'), 71.49 (d, *J* = 9.0 Hz, C-2'), 69.10 (C-3), 69.04 (C-4′), 64.50 (d, *J* = 4.5 Hz, C-5), 60.18 (C-6); ^{31}P NMR (243 MHz, D_2O): δ -11.30 (d, *J* = 20.5 Hz), -12.94 (d, *J* = 20.5 Hz); ESI-TOF HRMS m/z calcd for $C_{15}H_{24}N_3O_{16}P_2$ [M-H]$^{-1}$ 564.0637, found 564.0632.

dGDP-Glc. Yield: 89%. ^{1}H NMR (600 MHz, D_2O): δ 7.92 (s, 1H, -C*H*=), 6.15 (t, *J* = 7.2 Hz, 1H, H-1), 5.40 (dd, *J* = 6.6, 3.0 Hz, 1H, H-1'), 4.58 (br s, 1H, H-3), 4.08 (br s, 1H, H-4), 4.01–3.96 (m, 2H, H-5a,b), 3.74–3.69 (m, 1H, H-5'), 3.67 (d, *J* = 12.6 Hz, 1H, H-6a'), 3.62–3.54 (m, 2H, H-3', H-6b'), 3.36–3.31 (m, 1H, H-2'), 3.27 (t, *J* = 9.6 Hz, 1H, H-4'), 2.68–2.60 (m, 1H, H-2a), 2.38–2.30 (m, 1H, H-2b); ^{13}C NMR (150 MHz, D_2O): δ 158.89, 153.75, 151.34, 137.52, 116.11, 95.45 (d, *J* = 7.5 Hz, C-1'), 85.46 (d, *J* = 9.0 Hz, C-4), 83.49 (C-1), 72.70 (C-3'), 72.66 (C-5'), 71.49 (d, *J* = 9.0 Hz, C-2'), 71.22 (C-3), 69.07 (C-4'), 65.37 (C-5), 60.19 (C-6'), 38.37 (C-2); ^{31}P NMR (243 MHz, D_2O): δ -11.22 (d, *J* = 20.6 Hz), -12.99 (d, *J* = 20.6 Hz); ESI-TOF HRMS m/z calcd for $C_{16}H_{24}N_5O_{15}P_2$ [M-H]$^{-1}$ 588.0750, found 588.0680.

GDP-Glc. Yield: 91%. ^{1}H NMR (600 MHz, D_2O): δ 7.93 (s, 1H, -C*H*=), 5.75 (d, *J* = 6.0 Hz, 1H, H-1), 5.41 (dd, *J* = 7.2, 3.6 Hz, 1H, H-1'), 4.60 (t, *J* = 6.0 Hz, 1H, H-2), 4.35 (br t, *J* = 3.6 Hz, 1H, H-3), 4.18 (br s, 1H, H-4), 4.06–4.01 (m, 2H, H-5a,b), 3.73–3.68 (m, 1H, H-5'), 3.66 (d, *J* = 12.6 Hz, 1H, H-6a'), 3.64–3.54 (m, 2H, H-3', H-6b'), 3.36–3.32 (m, 1H, H-2'), 3.27 (t, *J* = 9.6 Hz, 1H, H-4'); ^{13}C NMR (150 MHz, D_2O): δ 158.88, 153.83, 151.70, 137.49, 116.15, 95.48 (d, *J* = 7.5 Hz, C-1'), 86.69 (C-1), 83.65 (d, *J* = 9.0 Hz, C-4), 73.50 (C-2), 72.72 (C-3'), 72.69 (C-5'), 71.50 (d, *J* = 9.0 Hz, C-2'), 70.29 (C-3), 69.08 (C-4'), 65.18 (d, *J* = 4.5 Hz, C-5), 60.21 (C-6'); ^{31}P NMR (243 MHz, D_2O): δ -11.25 (d, *J* = 20.5 Hz), -12.94 (d, *J* = 20.5 Hz); ESI-TOF HRMS m/z calcd for $C_{16}H_{24}N_5O_{16}P_2$ [M-H]$^{-1}$ 604.0699, found 604.0606.

dADP-Glc. Yield:73%.^{1}H NMR (600 MHz, D_2O): δ 8.32 (s, 1H, -C*H*=), 8.09 (s, 1H, -C*H*=), 6.36 (t, *J* = 6.6 Hz, 1H, H-1), 5.42–5.39 (dd, *J* = 6.6, 2.4 Hz, 1H, H-1'), 4.59 (br s, 1H, H-3), 4.14 (br s, 1H, H-4), 4.05–3.95 (m, 2H, H-5a,b), 3.73–3.69 (m, 1H, H-5'), 3.66 (d,

J = 12.6 Hz, 1H, H-6a'), 3.60 (t, *J* = 9.6 Hz, 1H, H-3'), 3.56 (dd, *J* = 12.6, 4.8 Hz, 1H, H-6b'), 3.35–3.31 (m, 1H, H-2'), 3.26 (t, *J* = 9.6 Hz, 1H, H-4'), 2.70–2.64 (m, 1H, H-2a), 2.47–2.41 (m, 1H, H-2b); ^{13}C NMR (150 MHz, D_2O): δ 155.11, 152.05, 148.63, 139.97, 118.55, 95.45 (d, *J* = 6.0 Hz, C-1'), 85.65 (d, *J* = 9.0 Hz, C-4), 83.67 (C-1), 72.73 (C-3'), 72.67 (C-5'), 71.52 (d, *J* = 9.0 Hz, C-2'), 71.14 (C-3), 69.10 (C-4'), 65.35 (d, *J* = 6.0 Hz, C-5), 60.22 (C-6'), 38.97 (C-2); ^{31}P NMR (243 MHz, D_2O): δ -11.22 (d, *J* = 20.6 Hz), -12.97 (d, *J* = 20.6 Hz); ESI-TOF HRMS m/z calcd for $C_{16}H_{24}N_5O_{14}P_2$ $[M-H]^{-1}$572.0800, found572.0728.

ADP-Glc. Yield: 78%.^{1}H NMR (600 MHz, D_2O): δ 8.34 (s, 1H, -C*H*=), 8.10 (s, 1H, -C*H*=), 5.98 (d, *J* = 6.0 Hz, 1H, H-1), 5.42 (dd, *J* = 7.2, 3.0 Hz, 1H, H-1'), 4.63–4.57 (m, 1H, H-2), 4.37 (br t, *J* = 4.2 Hz, 1H, H-3), 4.23 (br s, 1H, H-4), 4.06 (br s, 2H, H-5a,b), 3.74–3.69 (m, 1H, H-5'), 3.66 (d, *J* = 12.6 Hz, 1H, H-6a'), 3.60 (t, *J* = 9.6 Hz, 1H, H-3'), 3.57 (dd, *J* = 12.6, 4.2 Hz, 1H, H-6b'), 3.36–3.32 (m, 1H, H-2'), 3.26 (t, *J* = 9.6 Hz, 1H, H-4'); ^{13}C NMR (150 MHz, D_2O): δ 155.45, 152.60, 139.66, 95.48 (d, *J* = 7.5 Hz, C-1'), 86.70 (C-1), 83.80 (d, *J* = 9.0 Hz, C-4), 74.16 (C-2), 72.73 (C-3'), 72.67 (C-5'), 71.52 (d, *J* = 9.0 Hz, H-2'), 70.29 (C-3), 69.07 (C-4'), 65.13 (d, *J* = 6.0 Hz, C-5), 60.20 (C-6'); ^{31}P NMR (243 MHz, D_2O): δ -11.24 (d, *J* = 19.6 Hz), -12.95 (d, *J* = 19.6 Hz); ESI-TOF HRMS m/z calcd for $C_{16}H_{24}N_5O_{15}P_2$ $[M-H]^{-1}$ 588.0750, found 588.0704.

dTDP-Gal. Yield: 91%.^{1}H NMR (600 MHz, D_2O): δ 7.58 (s, 1H, -C*H*=), 6.19 (t, *J* = 7.2 Hz, 1H, H-1), 5.47 (dd, *J* = 7.2, 3.6 Hz, 1H, H-1'), 4.48–4.44 (m, 1H, H-3), 4.11–3.98 (m, 4H, H-4, H-5a,b, H-5'), 3.86 (d, *J* = 2.4 Hz, 1H, H-4'), 3.75 (dd, *J* = 10.2, 3.6 Hz, 1H, H-3'), 3.66–3.52 (m, 3H, H-2', H-6a,b'), 2.26–2.16 (m, 2H, H-2a,b), 1.77 (s, 3H, -C*H*$_3$); ^{13}C NMR (150 MHz, D_2O): δ 166.49, 151.65, 137.23, 111.67, 95.71 (d, *J* = 7.5 Hz, C-1'), 85.27 (d, *J* = 9.0 Hz, C-4), 84.90 (C-1), 71.78 (C-5'), 70.92 (C-3), 69.22 (C-3'), 69.00 (C-4'), 68.33 (d, *J* = 7.5 Hz, C-2'), 65.36 (d, *J* = 4.5 Hz, C-5), 60.91 (C-6'), 38.47 (C-2), 11.55 (-CH_3); ^{31}P NMR (243 MHz, D_2O): δ -11.40 (d, *J* = 21.0 Hz), -12.84 (d, *J* = 21.0 Hz); ESI-TOF HRMS m/z calcd for $C_{16}H_{25}N_2O_{16}P_2$ $[M-H]^{-1}$ 563.0685, found 563.0668.

dTDP-Man. Yield: 88%. ^{1}H NMR (600 MHz, D_2O): δ 7.61 (br d, *J* = 0.6 Hz, 1H, -C*H*=), 6.21 (t, *J* = 7.2 Hz, 1H, H-1), 5.37 (dd, *J* = 7.8, 1.8 Hz, 1H, H-1'), 4.51–4.44 (m, 1H, H-3), 4.07–3.99 (m, 3H, H-4, H-5a,b), 3.90 (dd, *J* = 3.6, 1.8 Hz, H-2'), 3.78 (dd, *J* = 10.2, 3.6 Hz, 1H, H-3'), 3.76–3.69 (m, 2H, H-5', H-6a'), 3.62 (dd, *J* = 12.0, 5.4 Hz, 1H, H-6b'), 3.54 (t, *J* = 9.6 Hz, 1H, H-4'), 2.26–2.18 (m, 2H, H-2a,b), 1.79 (br s, 3H, -C*H*$_3$); ^{13}C NMR (150 MHz, D_2O): δ 166.50, 151.67, 137.27, 111.70, 96.35 (d, *J* = 4.5 Hz, C-1'), 85.30 (d, *J* = 9.0 Hz, C-4), 84.92 (C-1), 73.60 (C-5'), 70.96 (C-3), 70.19 (d, *J* = 9.0 Hz, C-2'), 69.77 (C-3'), 69.00 (C-4'), 66.38, 65.38 (d, *J* = 6.0 Hz, C-5), 60.71 (C-6'), 38.55 (C-2), 11.58 (-*C*H$_3$); ^{31}P NMR (243 MHz, D_2O): δ -11.78 (d, *J* = 20.6 Hz), -13.92 (d, *J* = 20.6 Hz); ESI-TOF HRMS m/z calcd for $C_{16}H_{25}N_2O_{16}P_2$ $[M-H]^{-1}$ 563.0685, found 563.0697.

dTDP-GlcNH$_2$. Yield: 71%. ^{1}H NMR (600 MHz, D_2O): δ 7.55 (s, 1H, -C*H*=), 6.17 (t, *J* = 7.2 Hz, 1H, H-1), 5.66 (dd, *J* = 6.6, 3.0 Hz, 1H, H-1'), 4.44 (br s, 1H, H-3), 4.05–3.98 (m, 3H, H-4, H-5a,b), 3.78–3.72 (m, 2H, H-3', H-5'), 3.70 (d, *J* = 12.6 Hz, 1H, H-6a'), 3.64 (dd, *J* = 12.6, 3.6 Hz, 1H, H-6b'), 3.38 (t, *J* = 9.6 Hz, 1H, H-4'), 3.18 (br d, *J* = 10.2 Hz, 1H, H-2'), 2.25–2.16 (m, 2H, H-2a,b), 1.75 (s, 3H, -C*H*$_3$); ^{13}C NMR (150 MHz, D_2O): δ 166.49, 151.64, 137.21, 111.62, 92.72 (d, *J* = 4.5 Hz, C-1'), 85.13 (d, *J* = 9.0 Hz, C-4), 84.94 (C-1), 73.11 (C-5'), 70.82 (C-3), 69.74 (C-3'), 68.94 (C-4'), 65.51 (d, *J* = 4.5 Hz, C-5), 59.86 (C-6), 54.04 (d, *J* = 9.0 Hz, C-2'), 38.49 (C-2), 11.54 (-*C*H$_3$); ^{31}P NMR (243 MHz, D_2O): δ -11.34 (d, *J* = 21.0 Hz), -13.75 (d, *J* = 21.0 Hz); ESI-TOF HRMS m/z calcd for $C_{16}H_{26}N_3O_{15}P_2$ $[M-H]^{-1}$ 562.0850, found 562.0776.

UDP-Gal. Yield: 86%. ^{1}H NMR (600 MHz, D_2O): δ 7.78 (d, *J* = 7.8 Hz, 1H, -C*H*=), 5.82–5.77 (m, 2H, H-1, -C*H*=), 5.45 (dd, *J* = 7.2, 3.6 Hz, 1H, H-1'), 4.19 (br d, *J* = 3.0 Hz, 2H, H-2,3), 4.12–3.95 (m, 4H, H-4, H-5a,b, H-5'), 3.84 (d, *J* = 3.0 Hz, 1H, H-4'), 3.73 (dd, *J* = 10.2, 3.0 Hz, 1H, H-3'), 3.64–3.59 (m, 1H, H-2'), 3.59–3.51 (m, 2H, H-6a,b'); ^{13}C NMR (150 MHz, D_2O): δ 166.13, 151.72, 141.49, 102.54, 95.72 (d, *J* = 6.0 Hz, C-1'), 88.22 (C-1), 83.14 (d, *J* = 10.5 Hz), 73.66, 71.79, 69.53, 69.19, 68.98, 68.30 (d, *J* = 9.0 Hz), 64.82

(d, *J* = 4.5 Hz), 60.88; ^{31}P NMR (243 MHz, D_2O): δ -11.28 (d, *J* = 20.8 Hz), -12.80 (d, *J* = 20.8 Hz); ESI-TOF HRMS m/z calcd for $C_{15}H_{23}N_2O_{17}P_2$ $[M-H]^{-1}$565.0477, found 565.0460.

dUDP-Gal. Yield: 83%. ^{1}H NMR (600 MHz, D_2O): δ 7.80 (d, *J* = 7.8 Hz, 1H, -C*H*=), 6.18 (t, *J* = 6.6 Hz, 1H, H-1), 5.81 (d, *J* = 7.8 Hz, 1H, -C*H*=), 5.45 (dd, *J* = 7.2, 3.6 Hz, 1H, H-1'), 4.50–4.44 (m, 1H, H-3), 4.08–3.98 (m, 4H, H-4, H-5a,b, H-5'), 3.88 (d, *J* = 3.0 Hz, 1H, H-4'), 3.76 (dd, *J* = 10.2, 3.0 Hz, 1H, H-3'), 3.67–3.63 (m, 1H, H-2'), 3.63–3.55 (m, 2H, H-6a,b'), 2.28–2.19 (m, 2H, H-2a,b); ^{13}C NMR (150 MHz, D_2O): δ 166.23, 151.55, 141.81, 102.40, 95.74 (d, *J* = 7.5 Hz, C-1'), 85.42 (d, *J* = 9.0 Hz, H-4), 85.32 (C-1), 71.81 (C-5'), 70.79 (C-3), 69.23(C-3'), 69.01 (C-4'), 68.33 (d, *J* = 9.0 Hz, C-2'), 65.29 (d, *J* = 6.0 Hz, C-5), 60.91 (C-6'), 38.76 (C-2); ^{31}P NMR (243 MHz, D_2O): δ -11.20 (d, *J* = 19.4 Hz), -12.77 (d, *J* = 19.4 Hz); ESI-TOF HRMS m/z calcd for $C_{15}H_{24}N_2O_{16}P_2$ $[M-H]^{-1}$ 549.0528, found 565.0449.

UDP-Man. Yield: 78%. ^{1}H NMR (600 MHz, D_2O): δ 7.81 (d, *J* = 7.8 Hz, 1H, -C*H*=), 5.86–5.79 (m, 2H, H-1, -C*H*=), 5.36 (br d, *J* = 7.8 z, 1H, H-1'), 4.24–4.19 (m, 2H, H-2,3), 4.14 (br s, 1H, H-4), 4.11–4.01 (m, 2H, H-5a,b), 3.89 (br s, 1H, H-2'), 3.77 (dd, *J* = 10.2, 3.0 Hz, 1H, H-3'), 3.75–3.68 (m, 2H, H-5', H-6a'), 3.62 (dd, *J* = 12.0, 5.4 Hz, H-6b'), 3.53 (t, *J* = 9.6 Hz, 1H, H-4'); ^{13}C NMR (150 MHz, D_2O): δ 166.16, 151.74, 141.52, 102.57, 96.37 (d, *J* = 4.5 Hz, C-1'), 88.27 (C-1), 83.15 (d, *J* = 9.0 Hz, C-4), 73.70 (C-2), 73.60 (C-5'), 70.28 (d, *J* = 9.0 Hz, C-2'), 69.75 (C-3'), 69.56 (C-3), 66.35 (C-4'), 64.80 (d, *J* = 4.5 Hz, C-5), 60.70 (C-6'); ^{31}P NMR (243 MHz, D_2O): δ -11.62 (d, *J* = 20.7 Hz), -13.91 (d, *J* = 20.7 Hz); ESI-TOF HRMS m/z calcd for $C_{15}H_{23}N_2O_{17}P_2$ $[M-H]^{-1}$ 565.0477, found 565.0500.

GDP-Man. Yield: 13%. ESI-TOF HRMS m/z calcd for $C_{16}H_{24}N_5O_{16}P_2$ $[M-H]^{-1}$ 604.0699, found 604.0715.

UDP-GlcNH$_2$. Yield: 12%. ESI-TOF HRMS m/z calcd for $C_{15}H_{24}N_3O_{16}P_2$ $[M-H]^{-1}$564.0637, found 564.0607.

dADP-GlcNH$_2$. Yield: 21%. ESI-TOF HRMS m/z calcd for $C_{16}H_{25}N_6O_{13}P_2$[M-H]$^{-1}$571.0960, found 571.0910.

dTDP-GalNH$_2$. Yield: 5%. ESI-TOF HRMS m/z calcd for $C_{16}H_{26}N_3O_{15}P_2$ $[M-H]^{-1}$ 562.0850, found 562.0781.

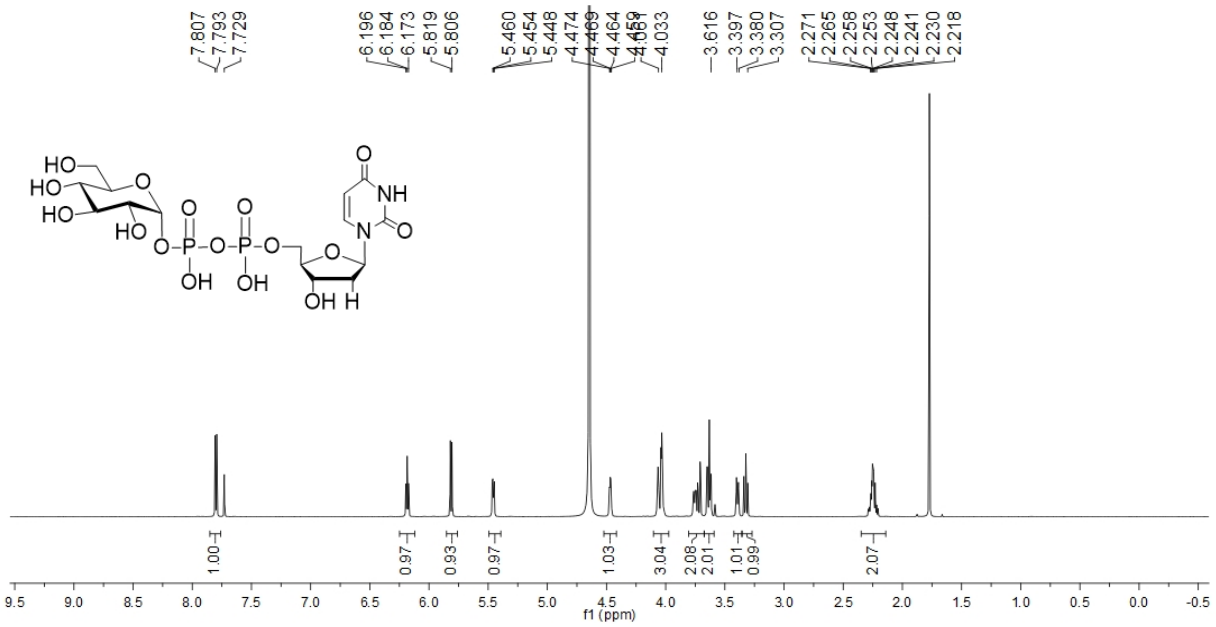
II. HPLC analysis of sugar-nucleotides

Table S1. Retention time of various sugar-nucleotides[a]

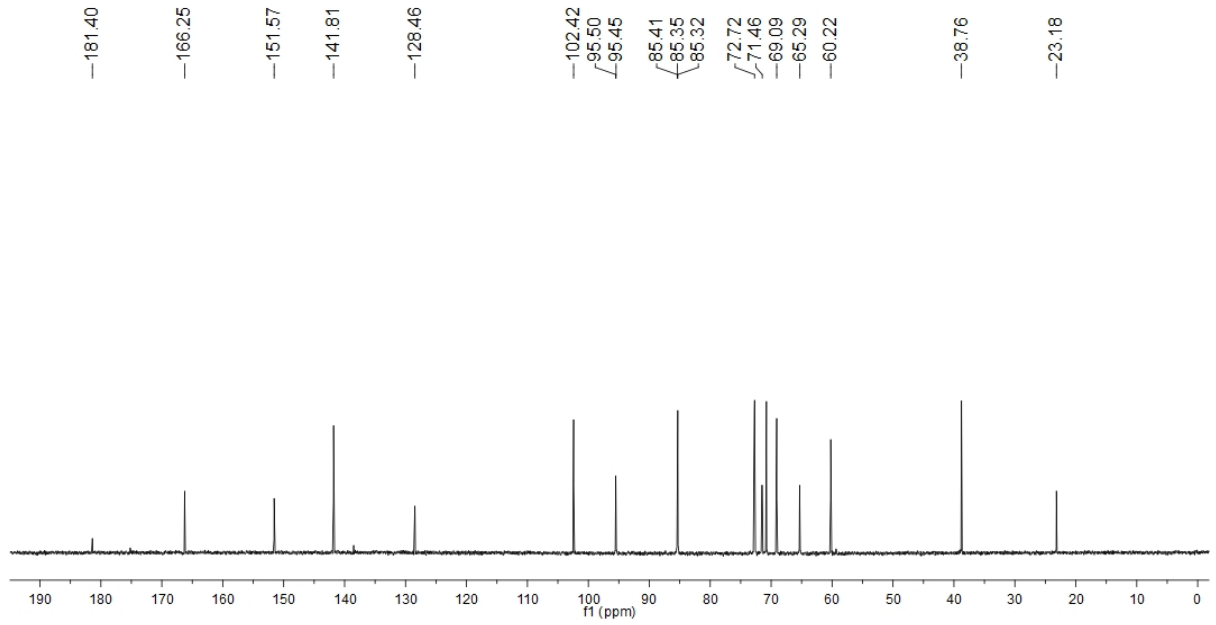
Entry	(d)NDP-Sugar	Retention Time
1	dTDP-Glc	17.35
2	dUDP-Glc	21.88
3	UDP-Glc	17.16
4	dCDP-Glc	16.07
5	CDP-Glc	15.87
6	dGDP-Glc	26.66
7	GDP-Glc	26.70
8	dADP-Glc	22.22
9	ADP-Glc	22.52
10	dTDP-Gal	16.38
11	dTDP-Man	16.24
12	dTDP-$GlcNH_2$	8.94
13	dTDP-$GalNH_2$	8.35
14	UDP-Gal	16.39
15	dUDP-Gal	21.82
16	UDP-Man	16.32
17	GDP-Man	26.27
18	UDP-$GlcNH_2$	8.91
19	dADP-$GlcNH_2$	11.48

[a]HPLC conditions: DionexCarboPac^{TM}PA-100 column;
4 × 250 mm size; 0~1.0 M ammonium acetate gradient eluents.

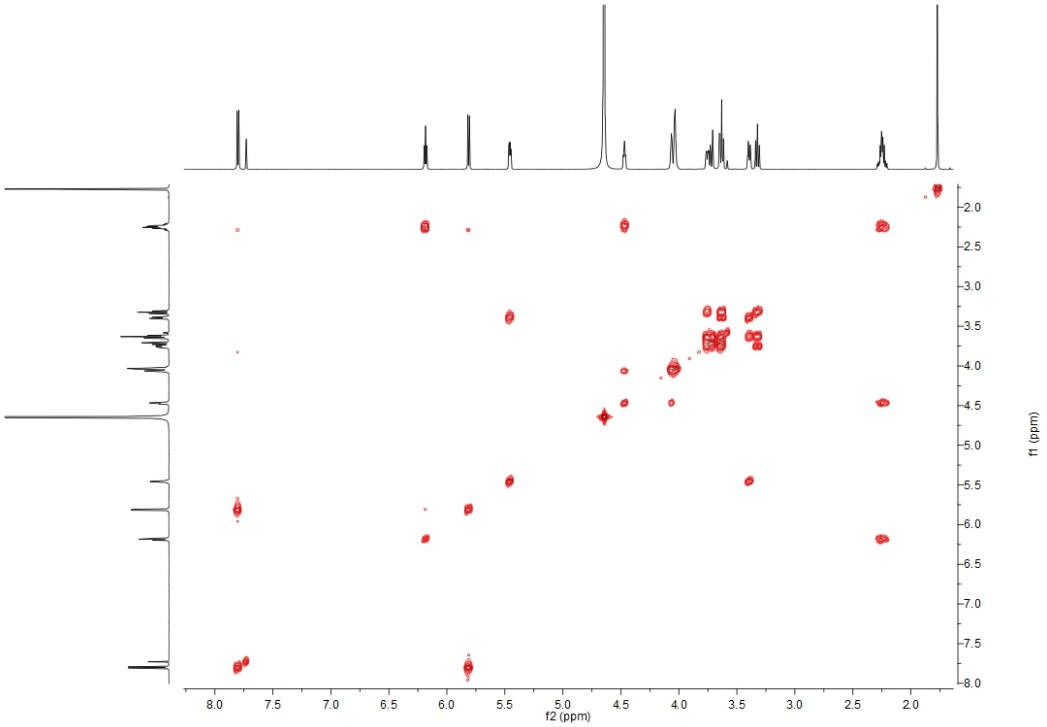
III. MS and NMR spectra



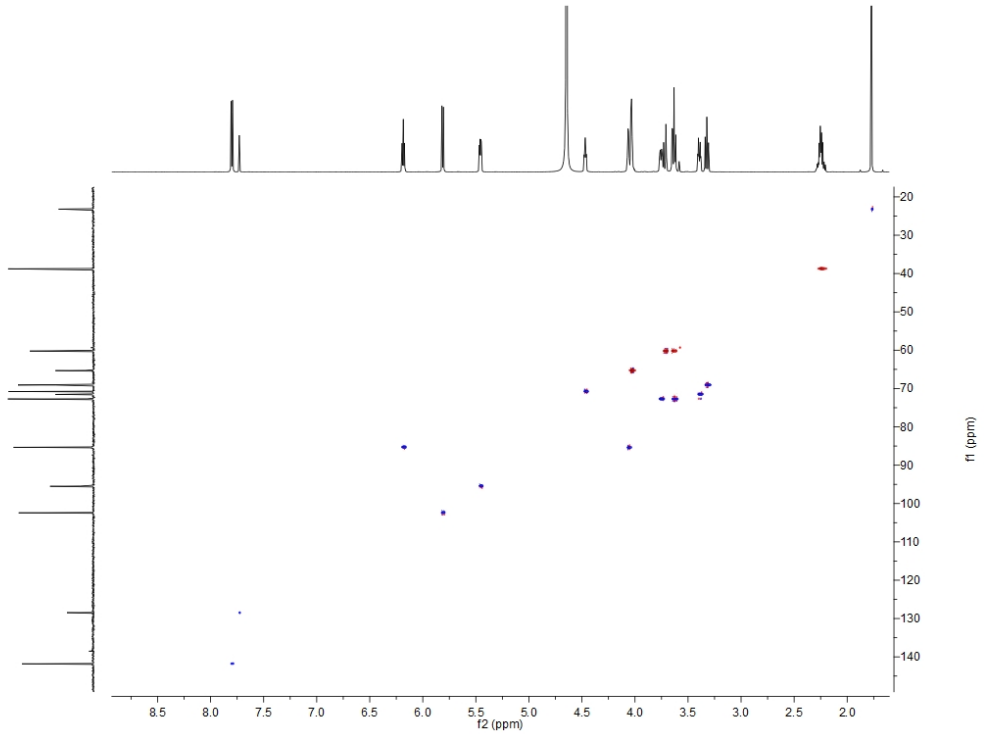
^{1}H NMR spectrum of dUDP-Glc (600 MHz, D_2O)



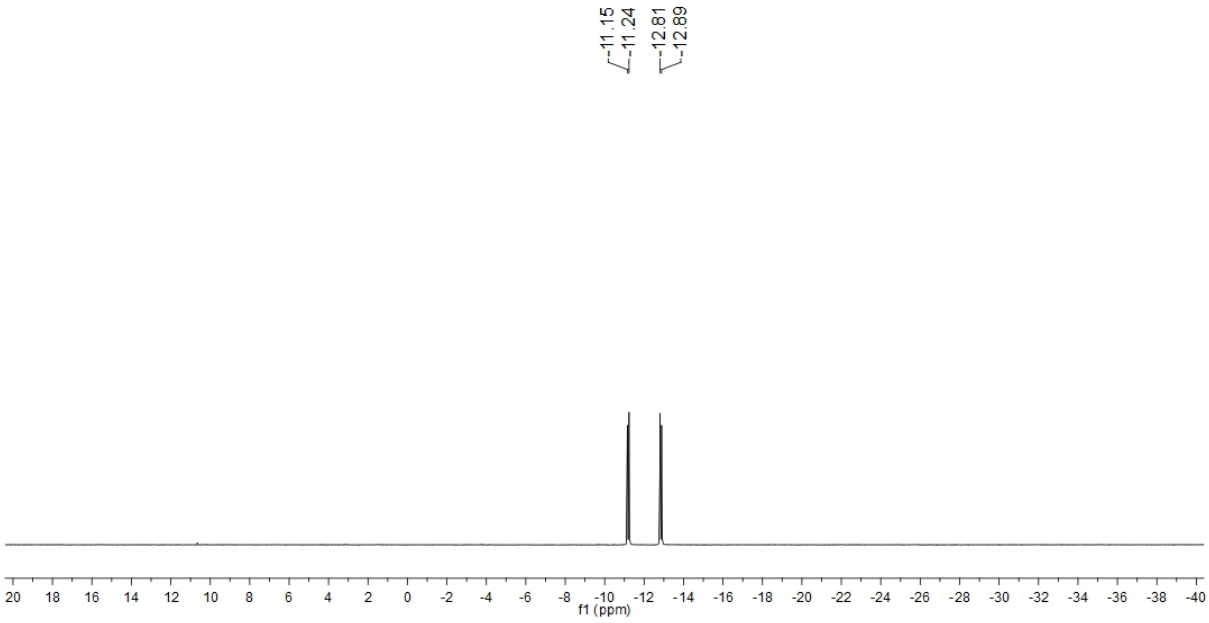
^{13}C NMR spectrum of dUDP-Glc (150 MHz, D_2O)



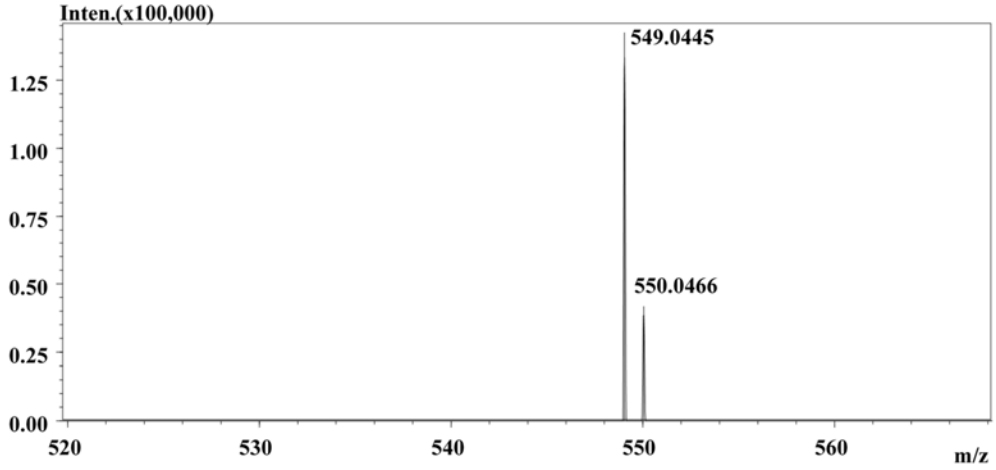
^{1}H-^{1}H COSY spectrum of dUDP-Glc (600 MHz, D_2O)



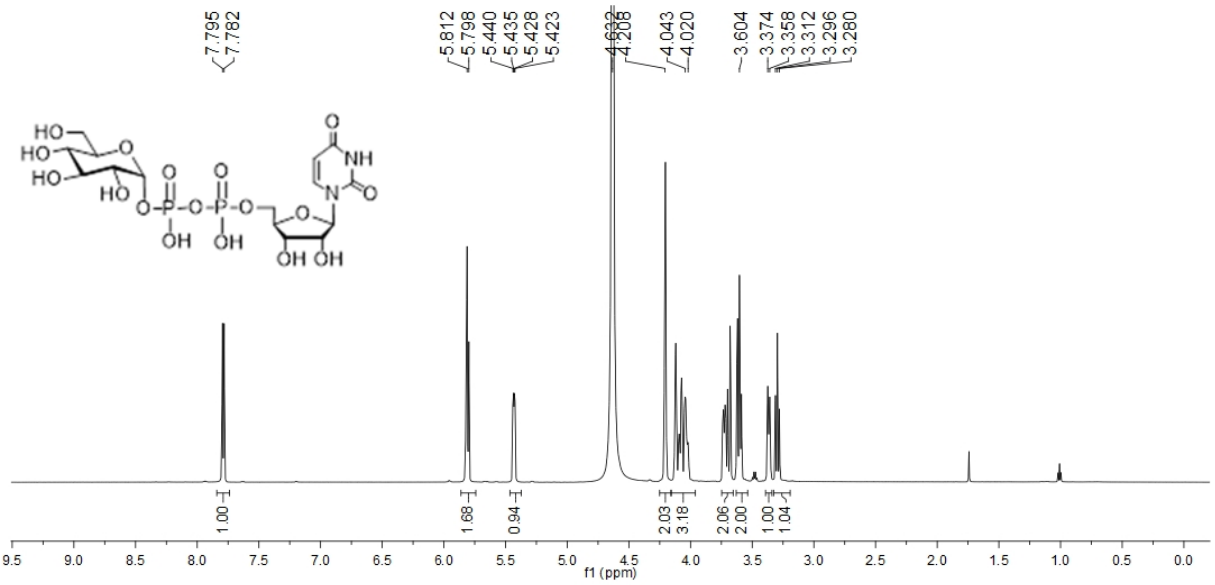
^{1}H-^{13}C HSQC spectrum of dUDP-Glc (600/150 MHz, D_2O)



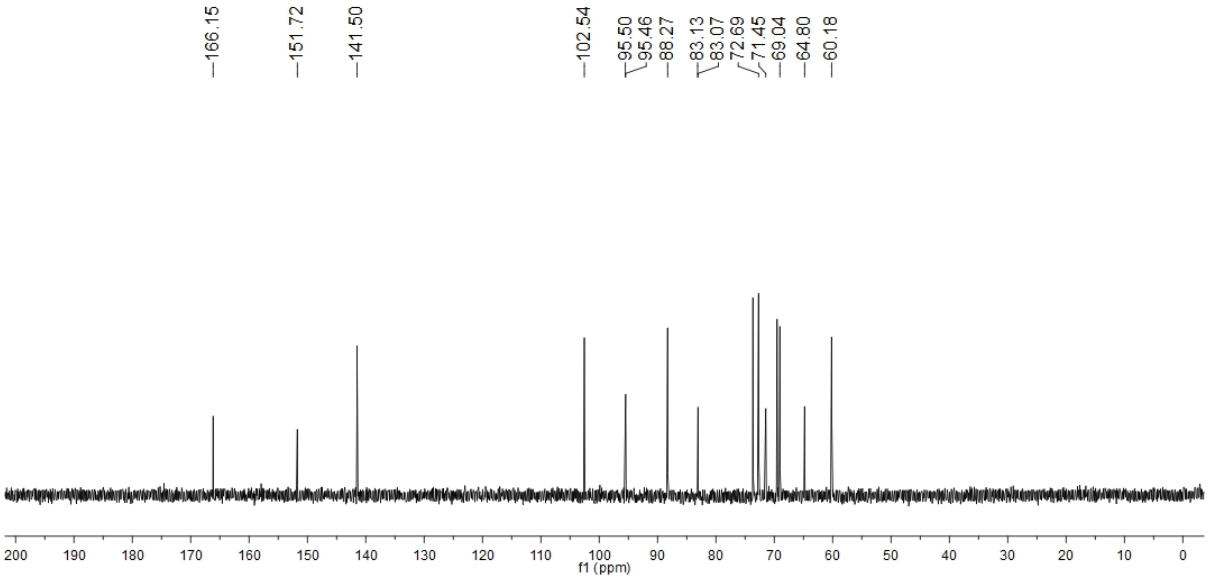
^{31}P NMR spectrum of dUDP-Glc (243 MHz, D_2O)



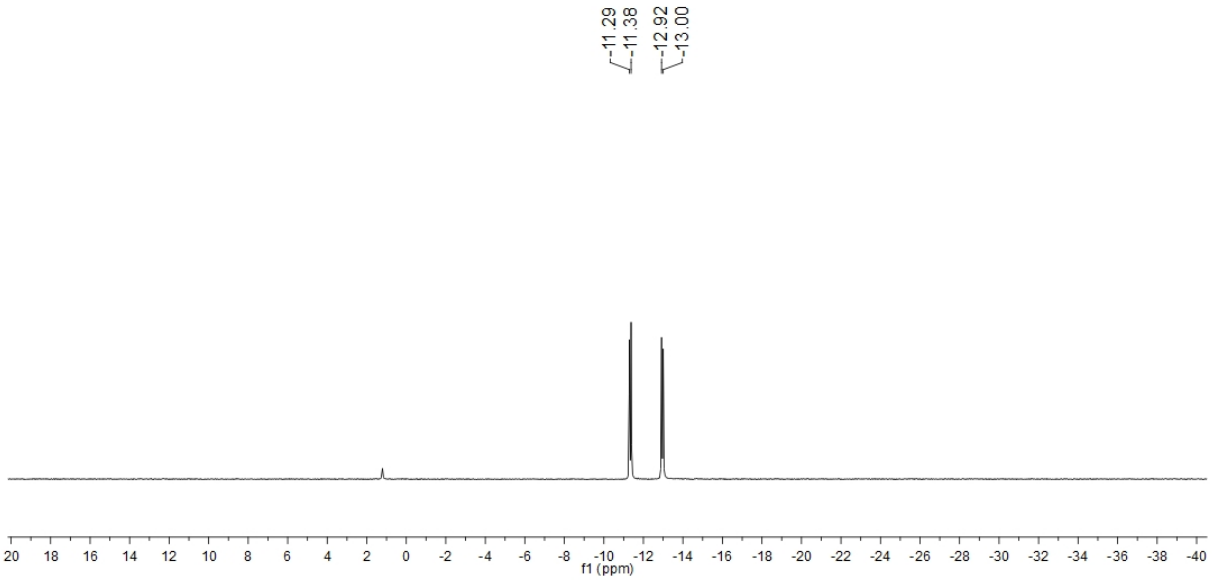
ESI-HRMS spectrum of dUDP-Glc



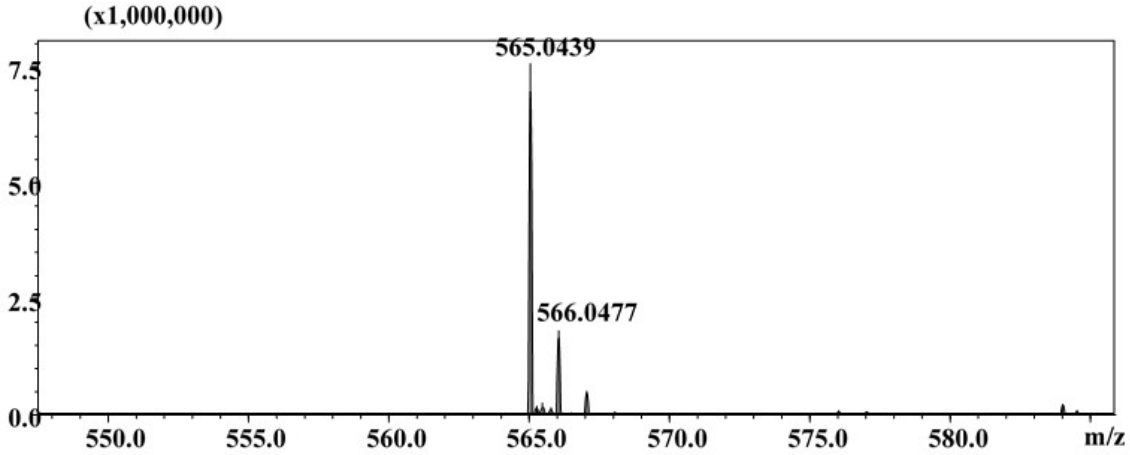
^{1}H NMR spectrum of UDP-Glc (600 MHz, D_2O)



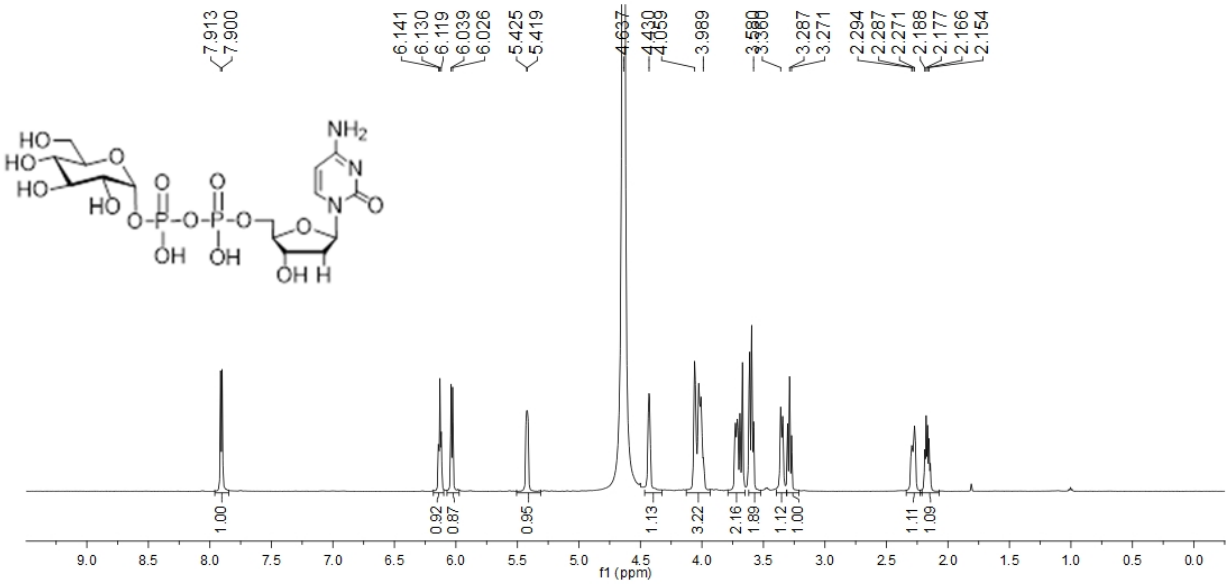
^{13}C NMR spectrum of UDP-Glc (150 MHz, D_2O)



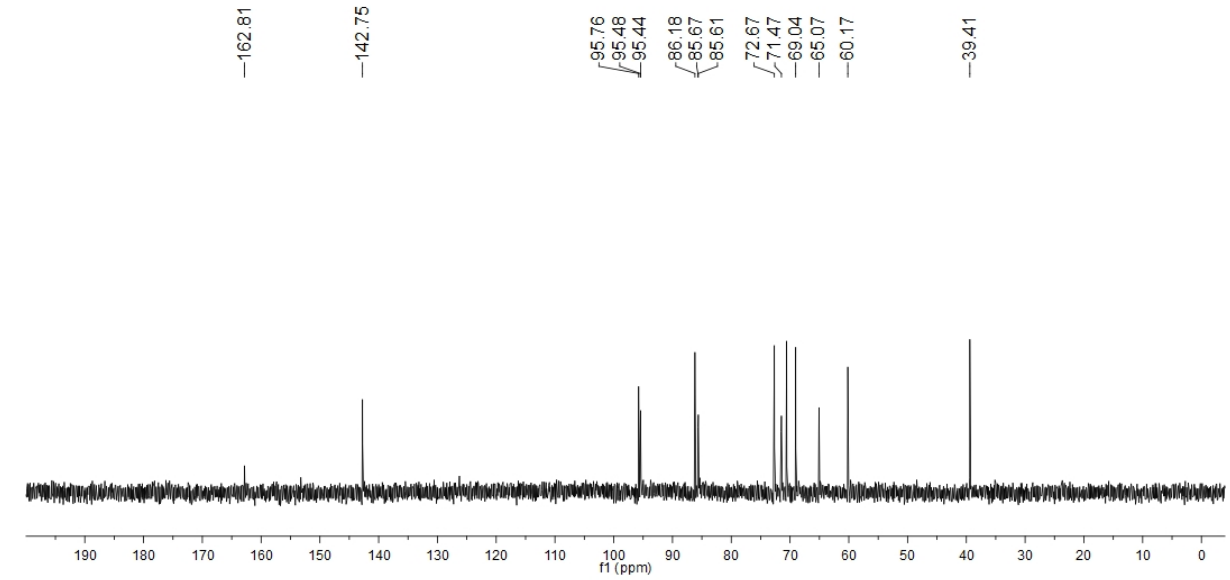
^{31}P NMR spectrum of UDP-Glc (243 MHz, D_2O)



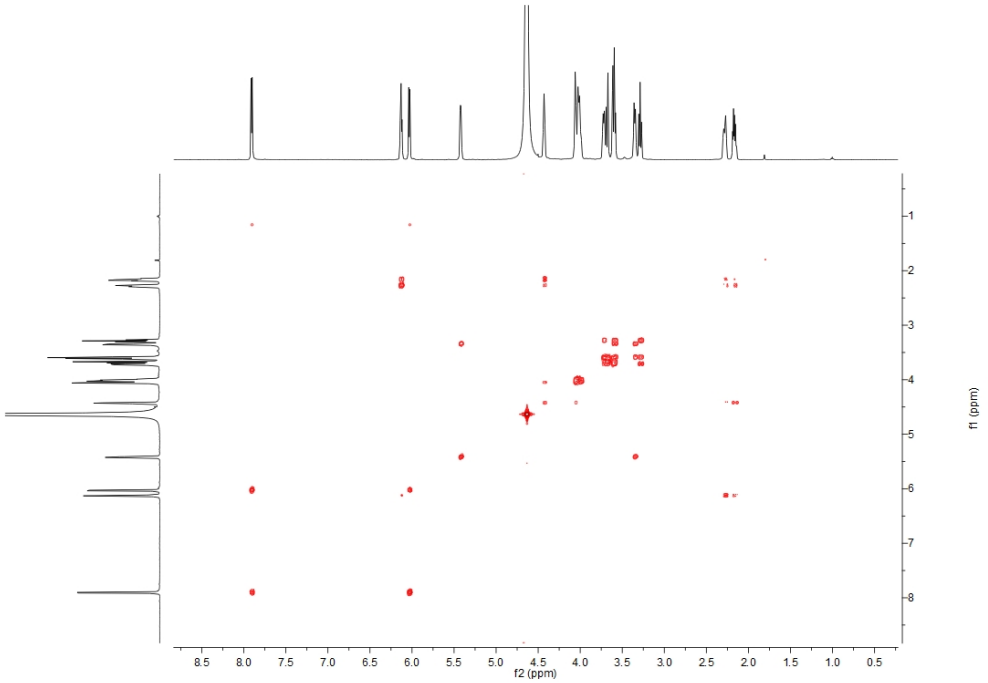
ESI-HRMS spectrum of UDP-Glc



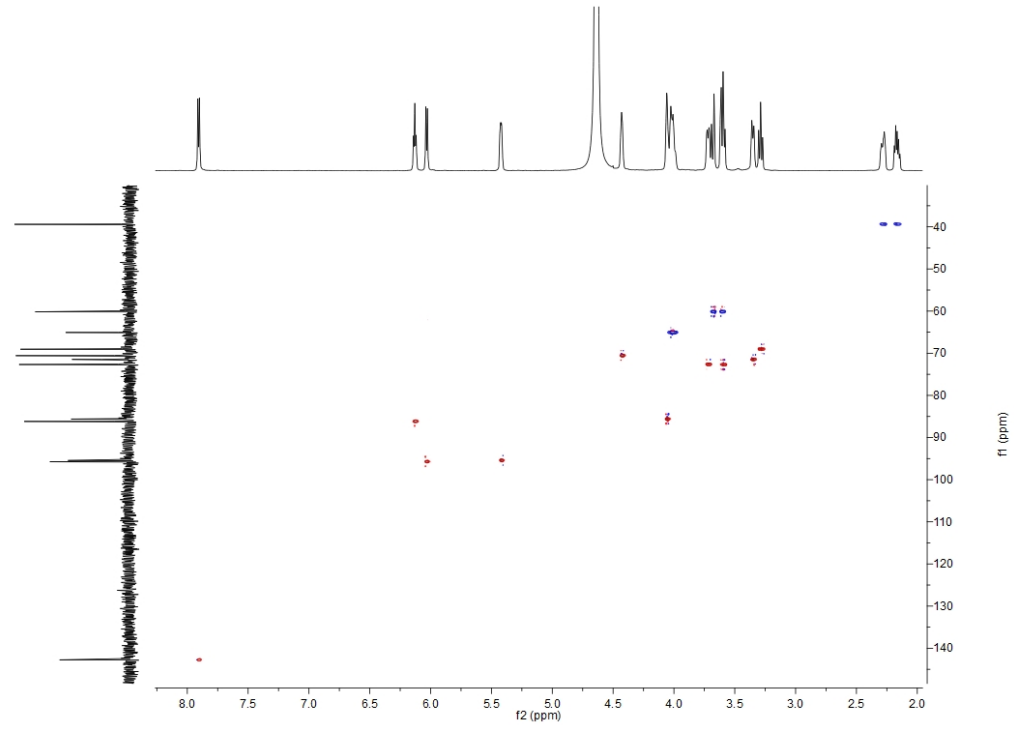
^{1}H NMR spectrum of dCDP-Glc (600 MHz, D_2O)



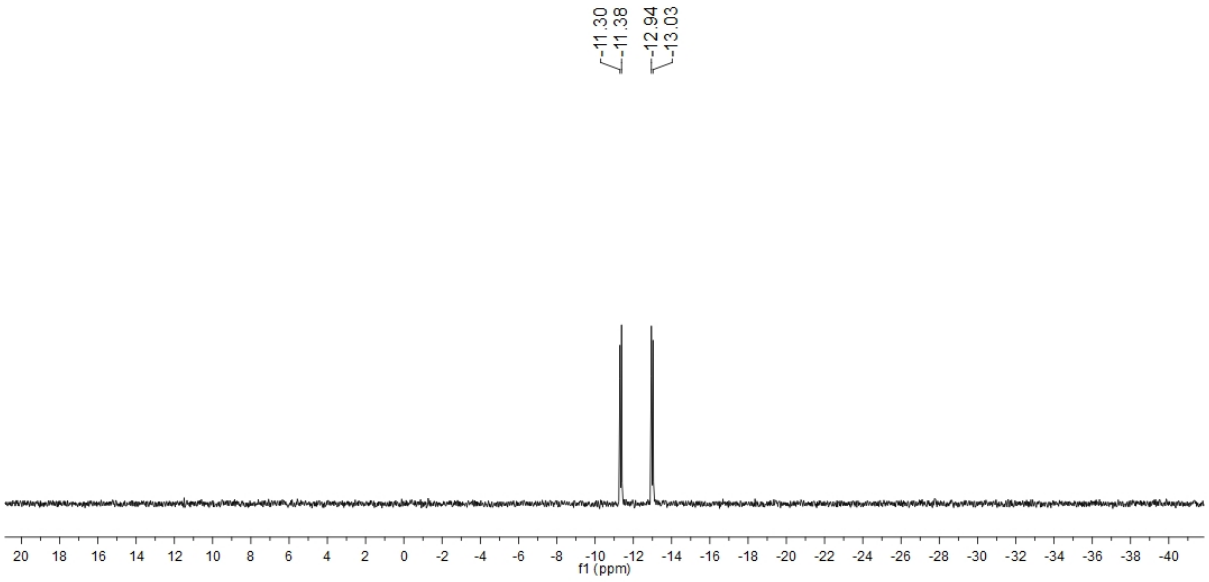
^{13}C NMR spectrum of dCDP-Glc (150 MHz, D_2O)



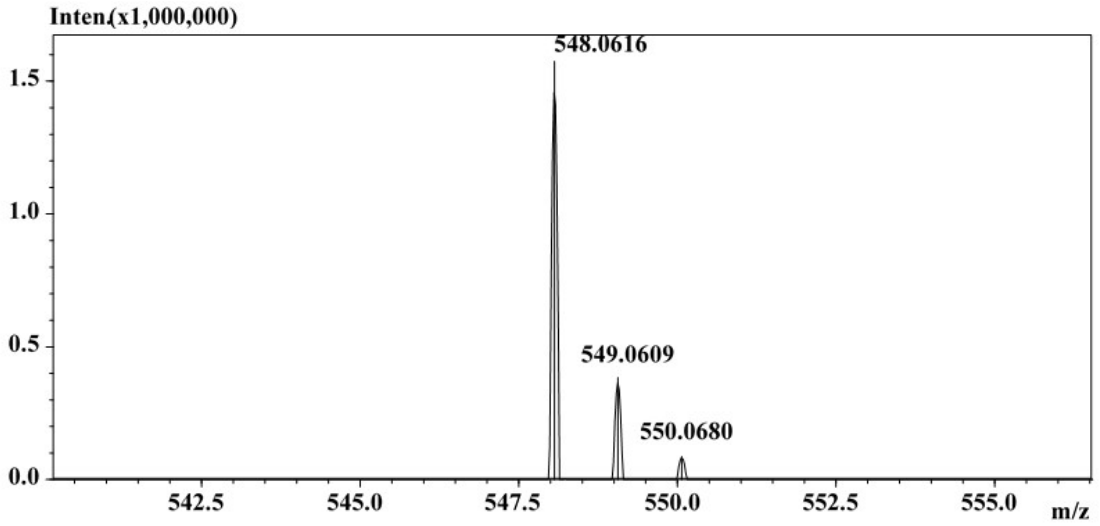
^{1}H-^{1}H COSY spectrum of dCDP-Glc (600 MHz, D_2O)



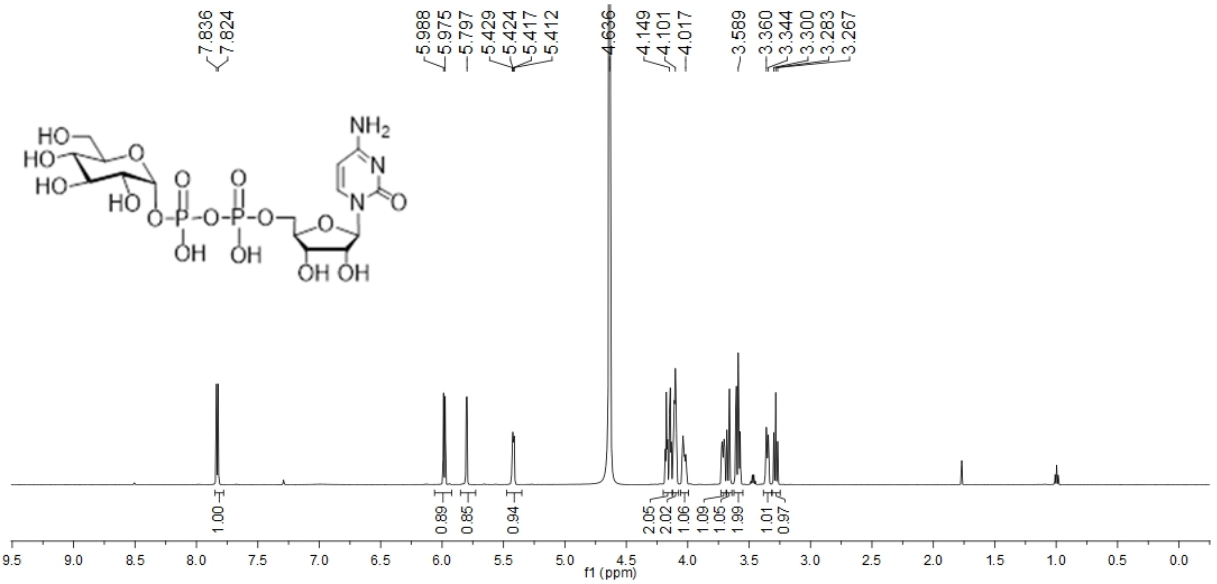
^{1}H-^{13}C HSQC spectrum of dCDP-Glc (600/150 MHz, D_2O)



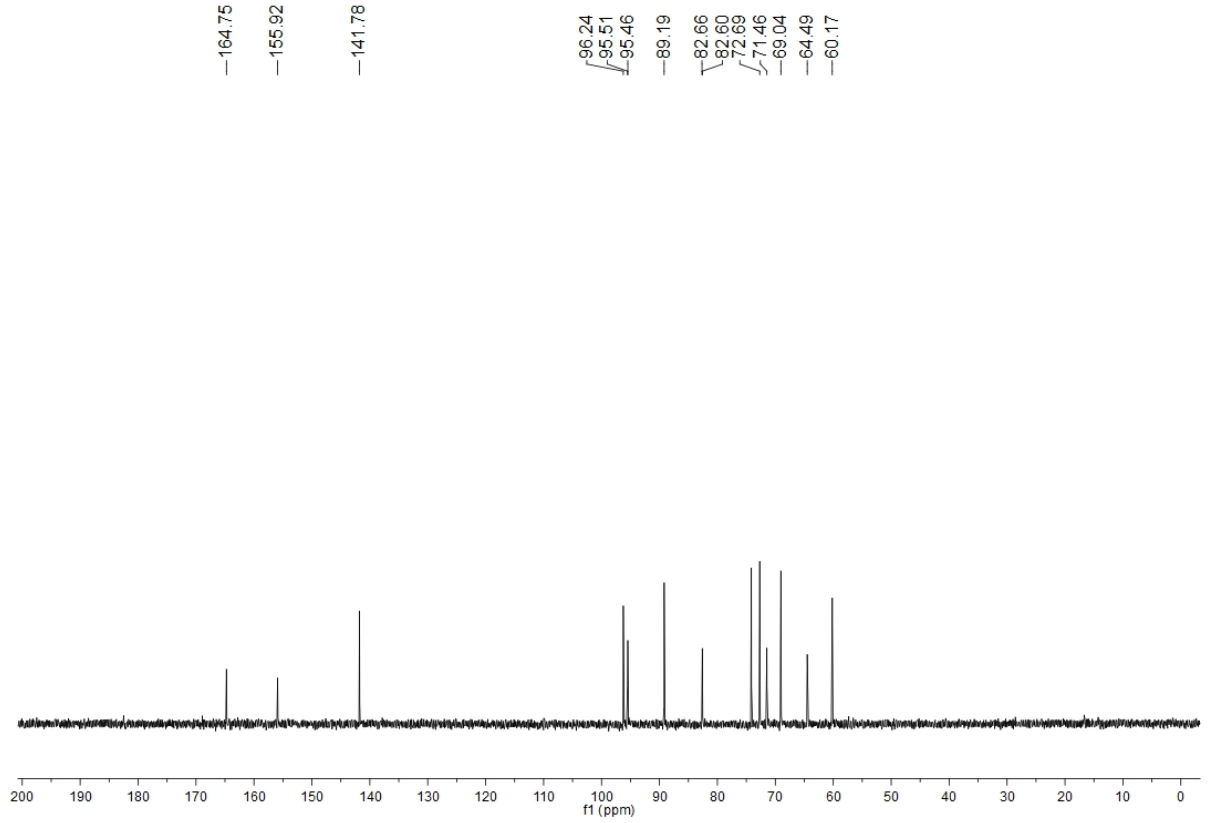
^{31}P NMR spectrum of dCDP-Glc (243 MHz, D_2O)



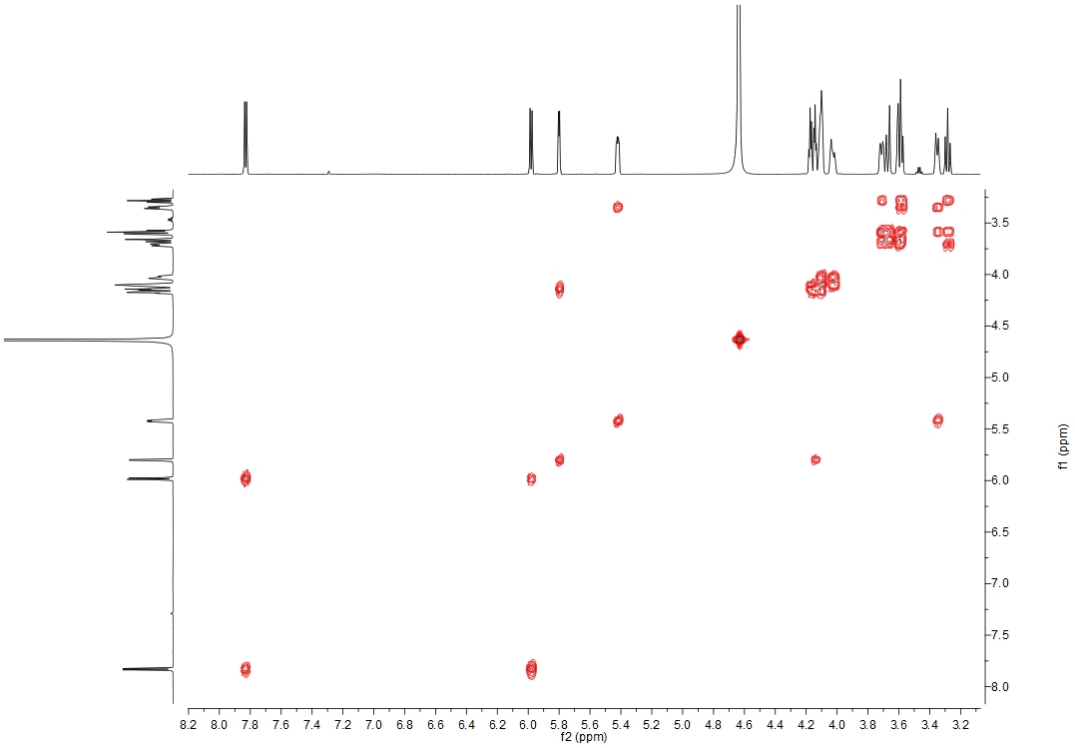
ESI-HRMS spectrum of dCDP-Glc



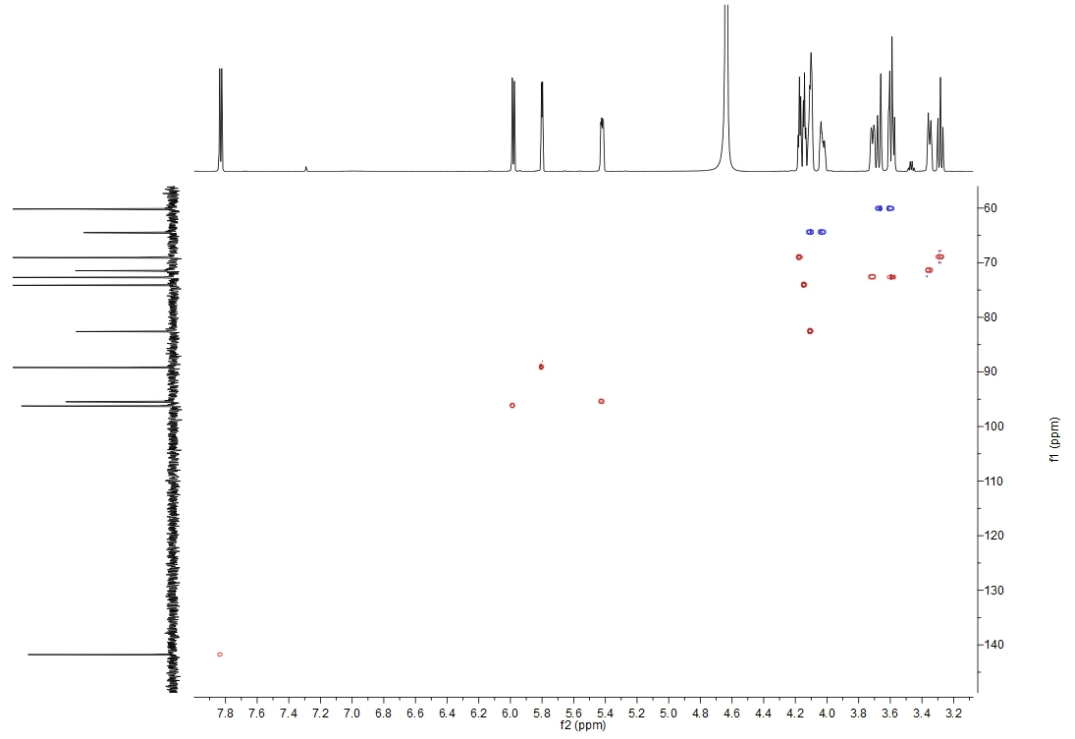
^{1}H NMR spectrum of CDP-Glc (600 MHz, D_2O)



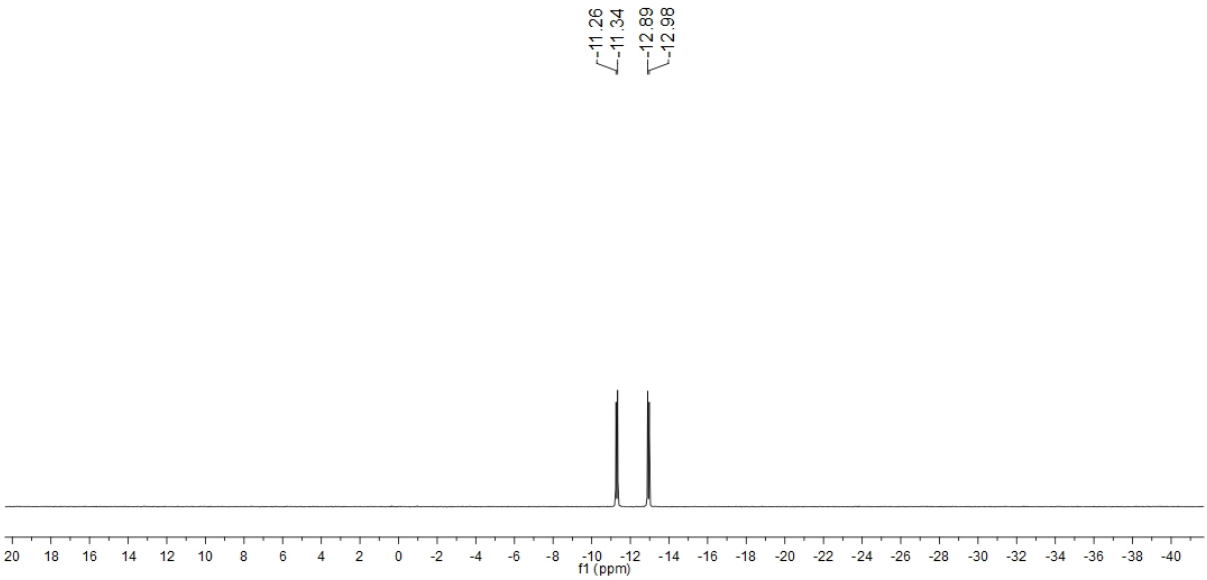
^{13}C NMR spectrum of CDP-Glc (150 MHz, D_2O)



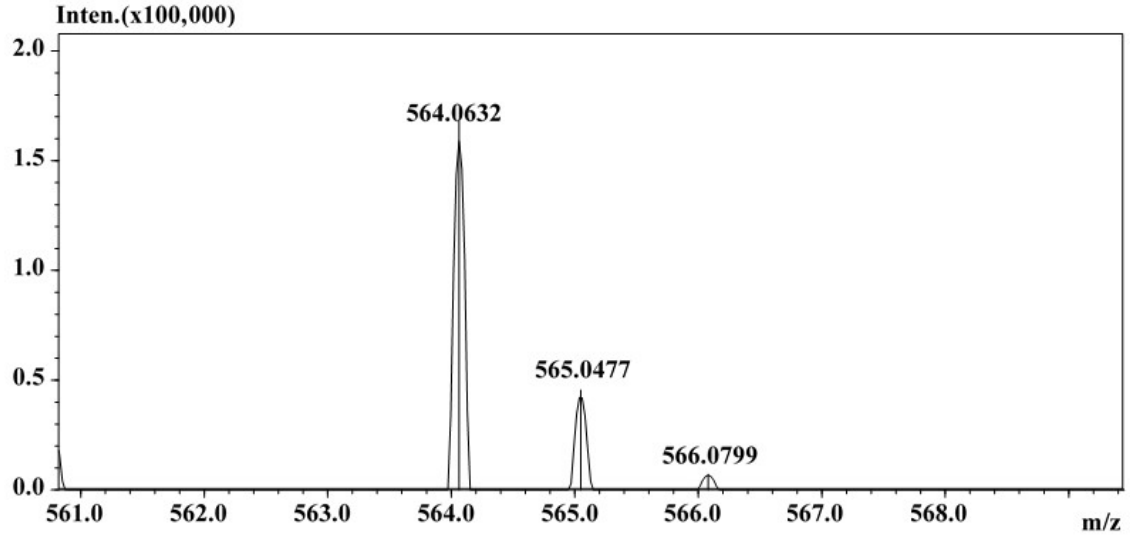
^{1}H-^{1}H COSY spectrum of CDP-Glc (600 MHz, D_2O)



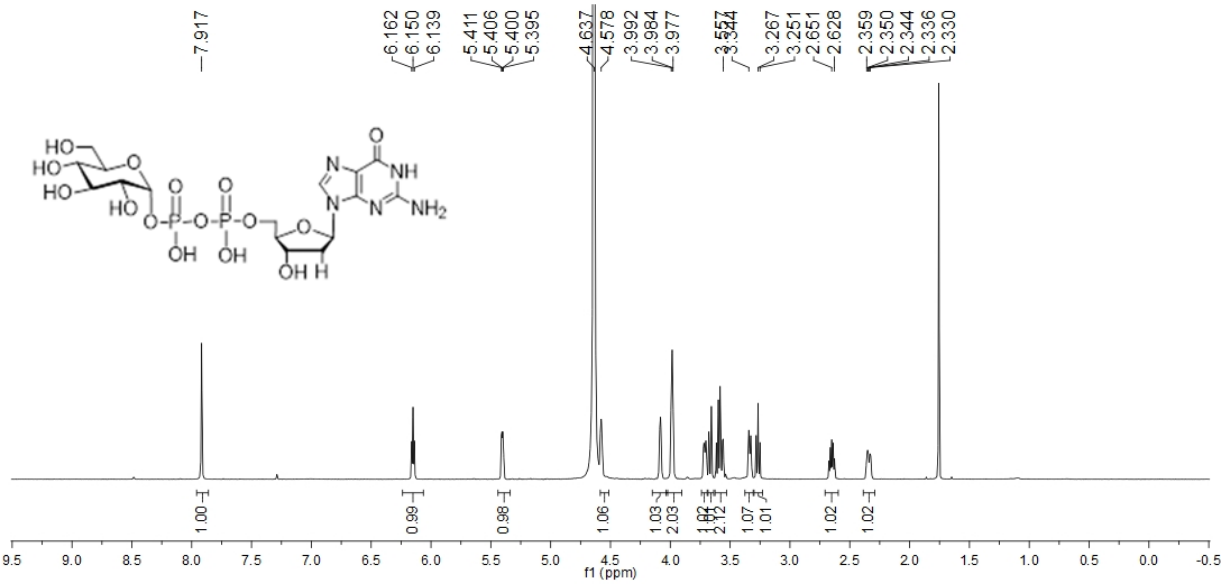
^{1}H-^{13}C HSQC spectrum of CDP-Glc (600/150 MHz, D_2O)



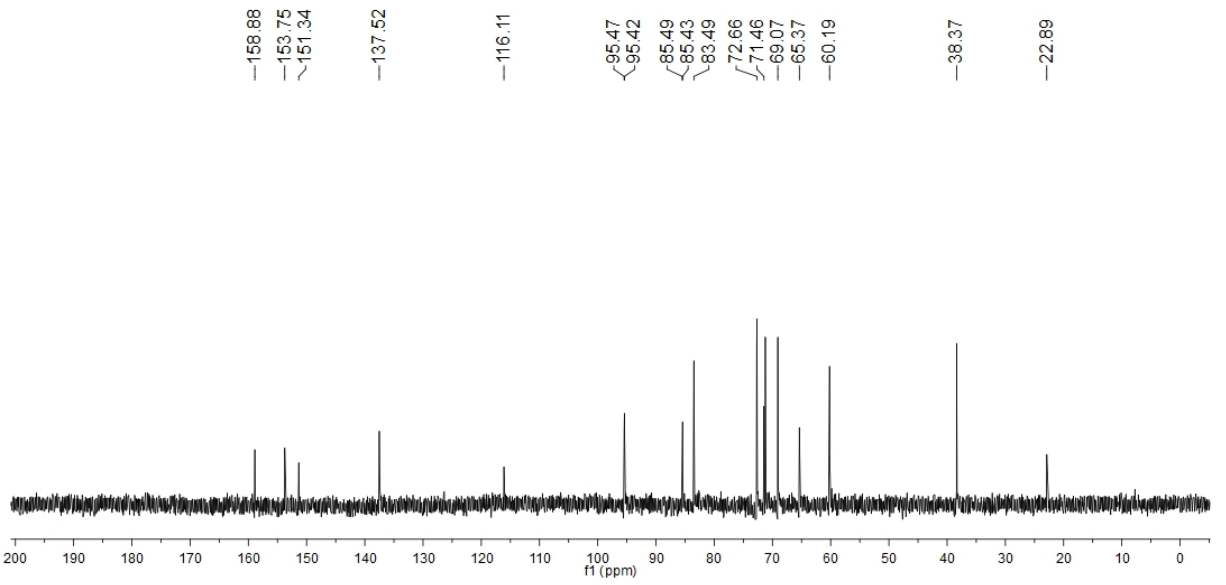
^{31}P NMR spectrum of CDP-Glc (243 MHz, D_2O)



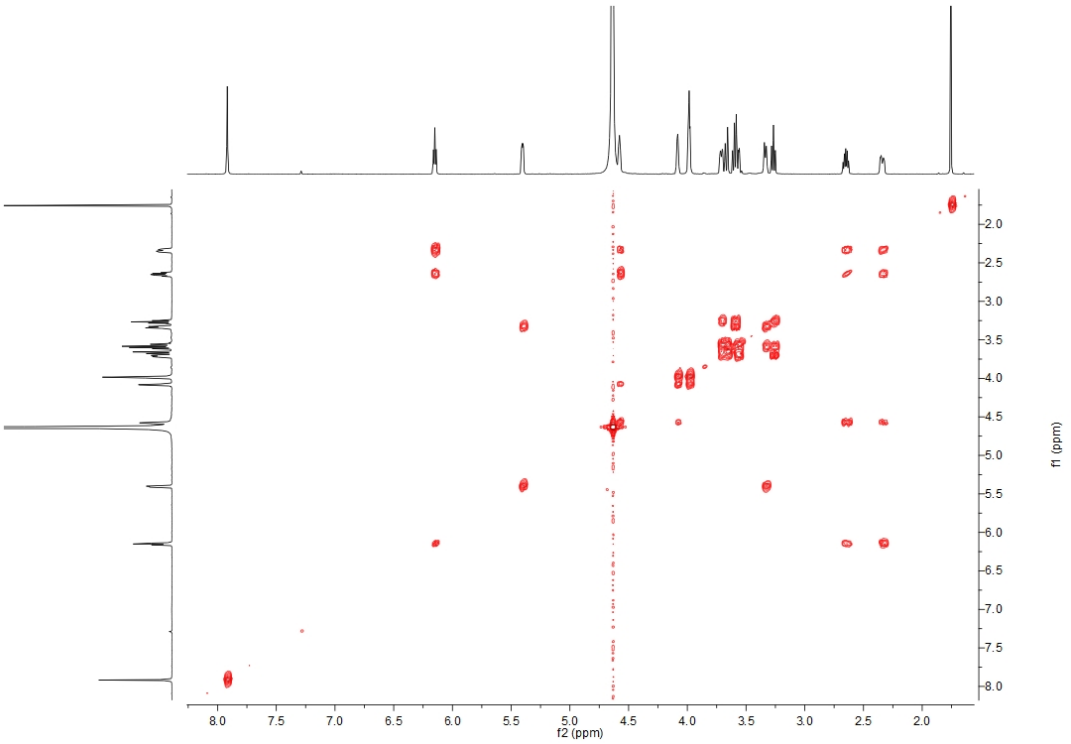
ESI-HRMS spectrum of CDP-Glc



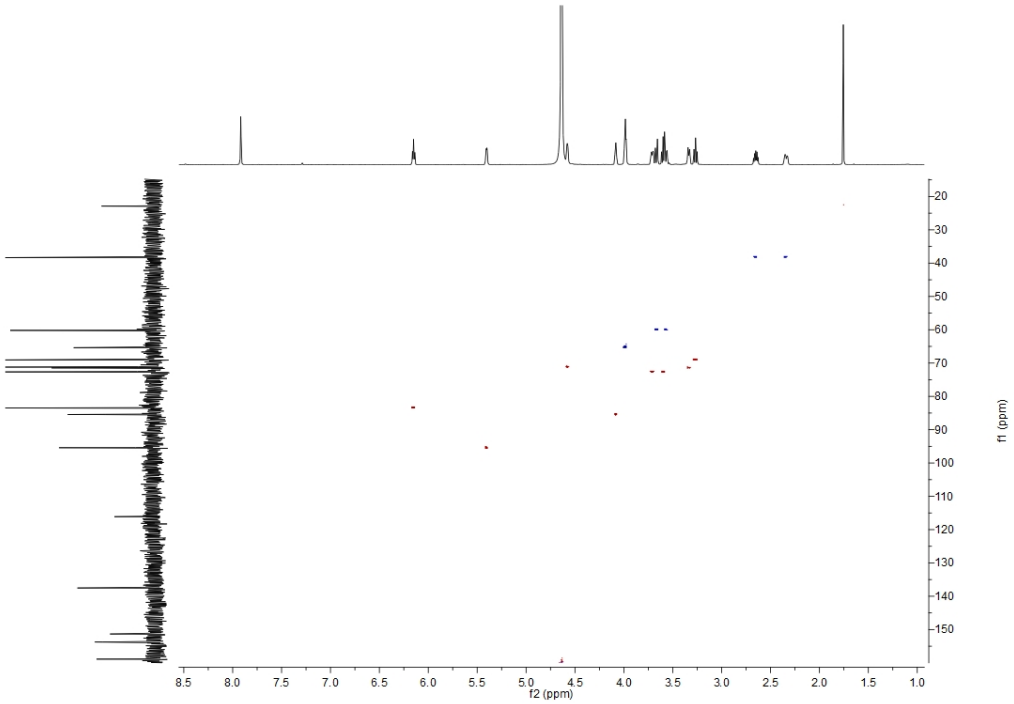
^{1}H NMR spectrum of dGDP-Glc (600 MHz, D_2O)



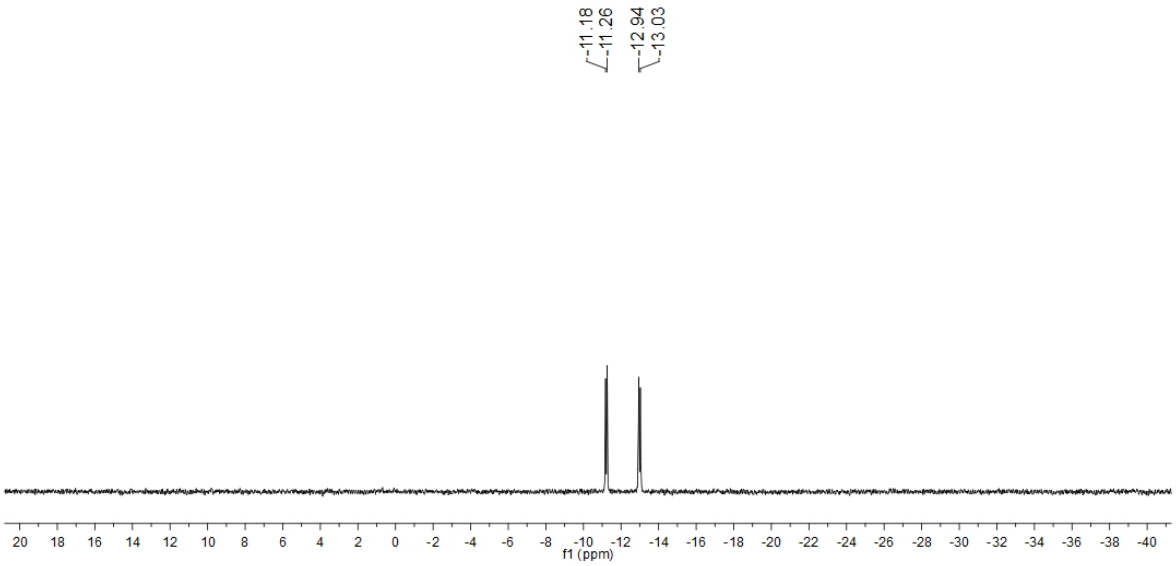
^{13}C NMR spectrum of dGDP-Glc (150 MHz, D_2O)



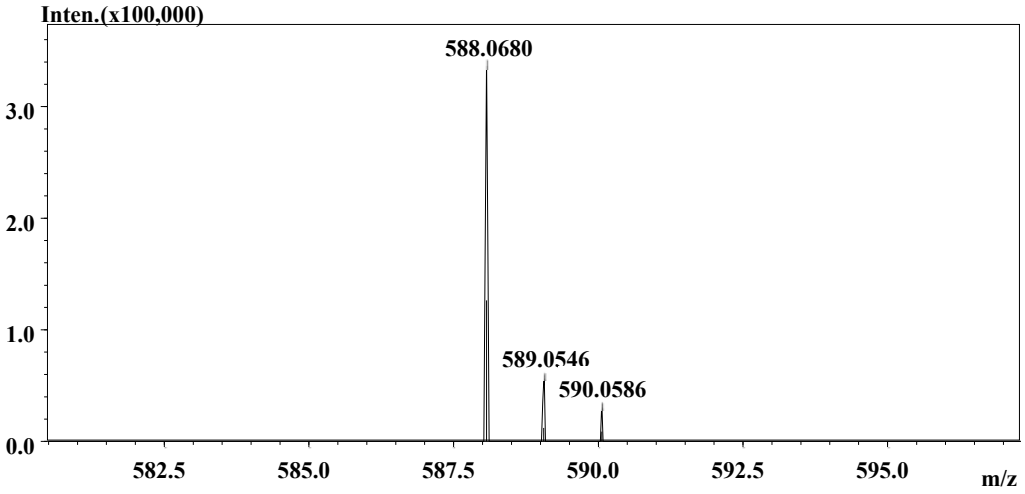
^{1}H-^{1}H COSY spectrum of dGDP-Glc (600 MHz, D_2O)



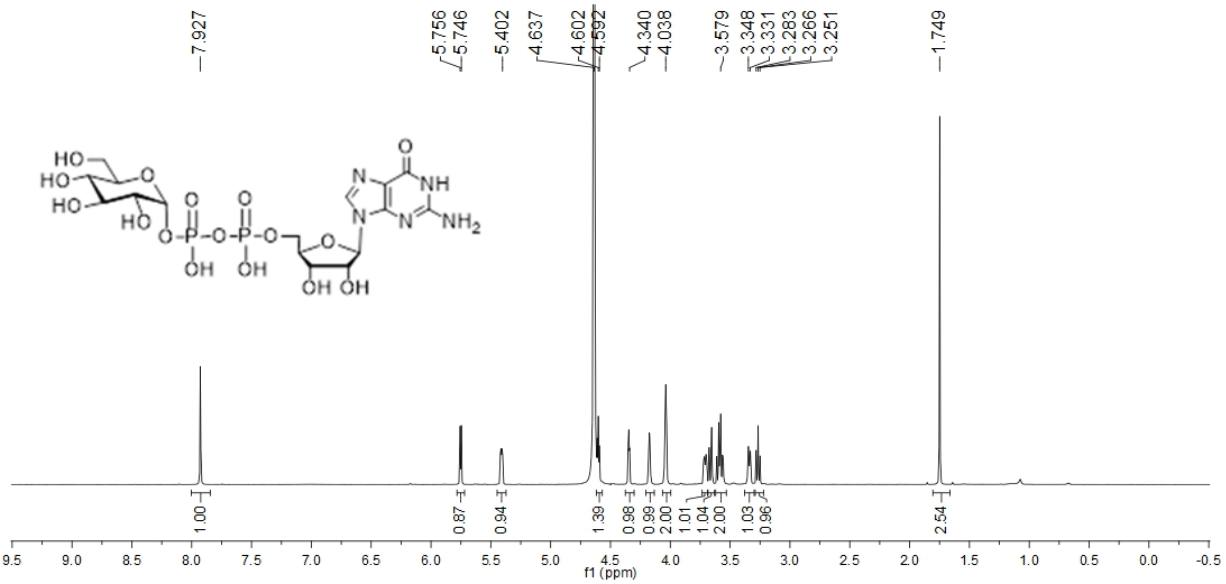
^{1}H-^{13}C HSQC spectrum of dGDP-Glc (600/150 MHz, D_2O)



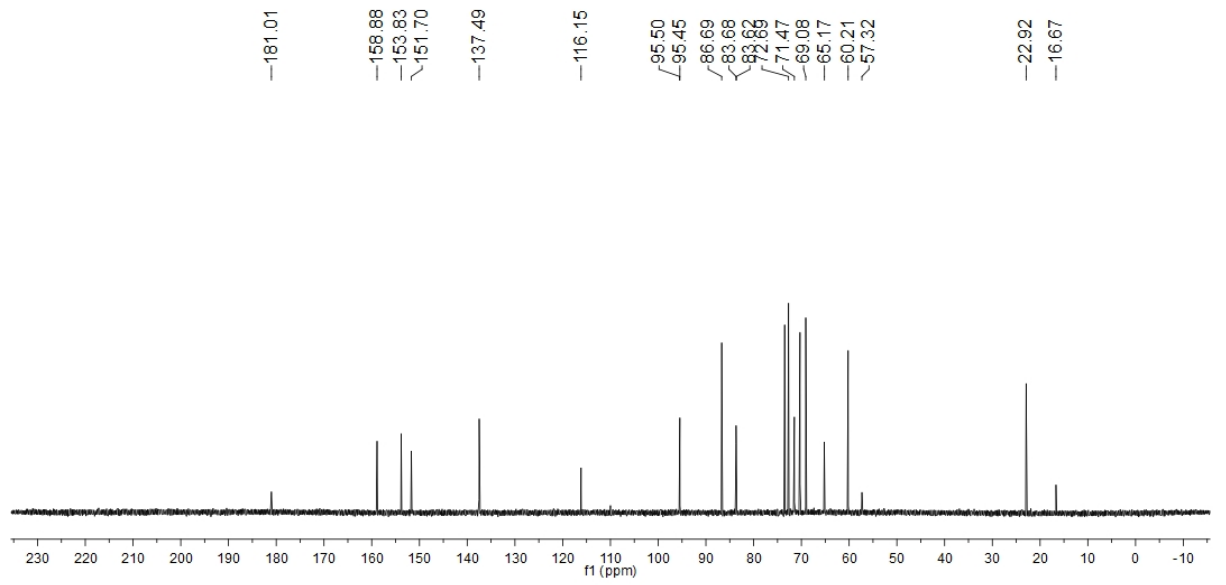
^{31}P NMR spectrum of dGDP-Glc (243 MHz, D_2O)



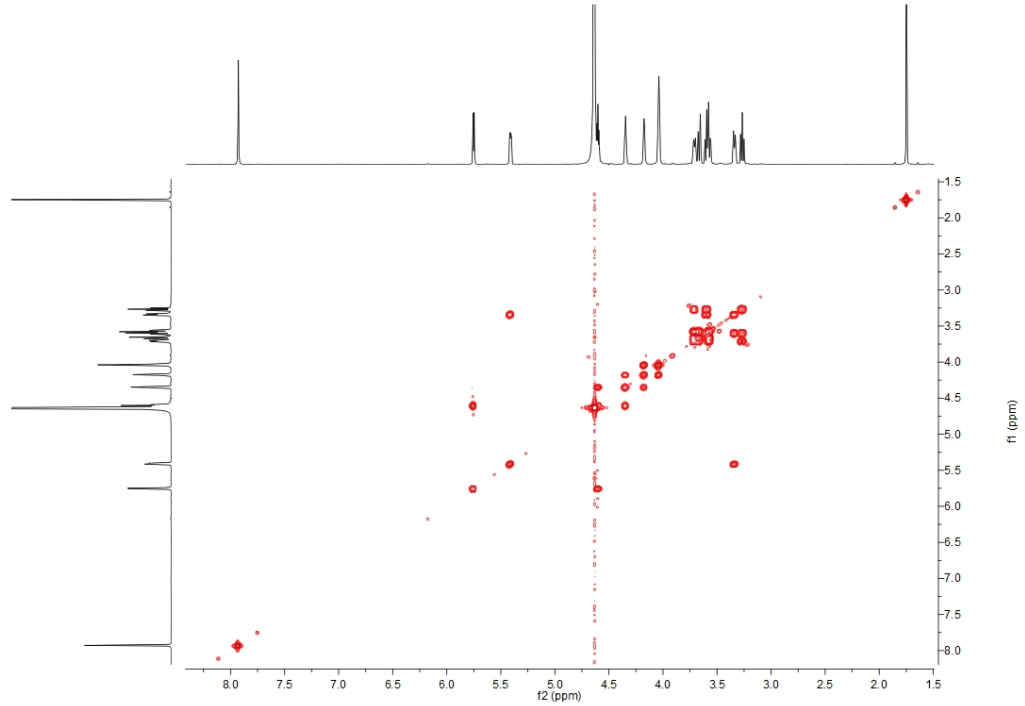
ESI-HRMS spectrum of dGDP-Glc



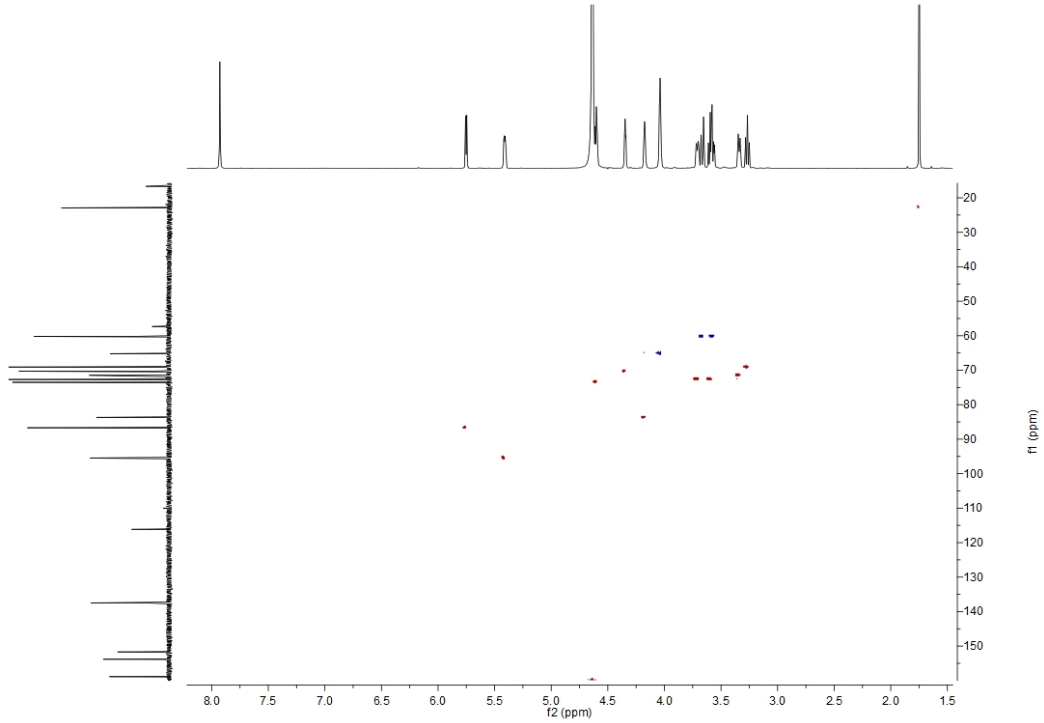
^{1}H NMR spectrum of GDP-Glc (600 MHz, D_2O)



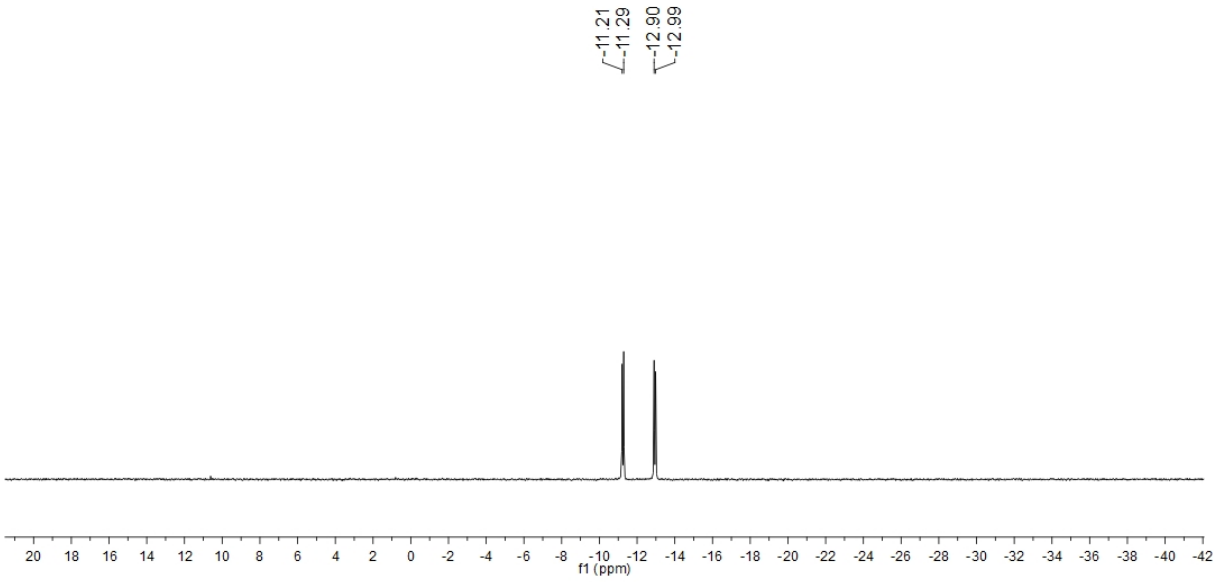
^{13}C NMR spectrum of GDP-Glc (150 MHz, D_2O)



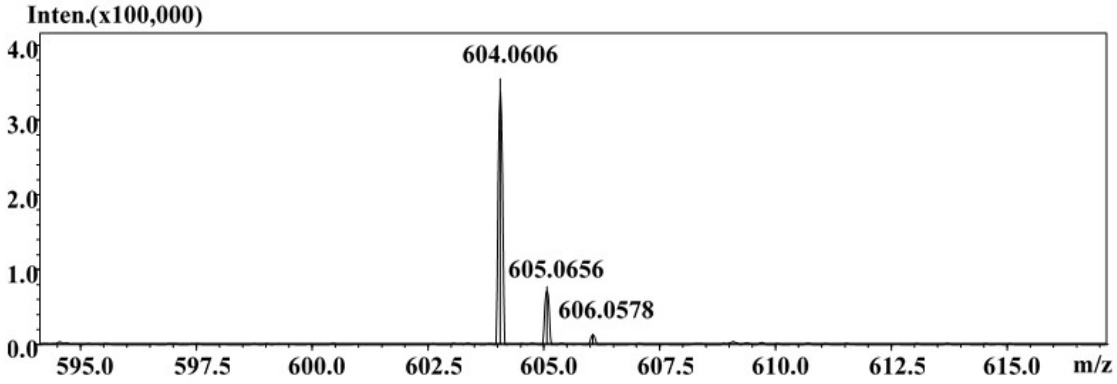
^{1}H-^{1}H COSY spectrum of GDP-Glc (600 MHz, D_2O)



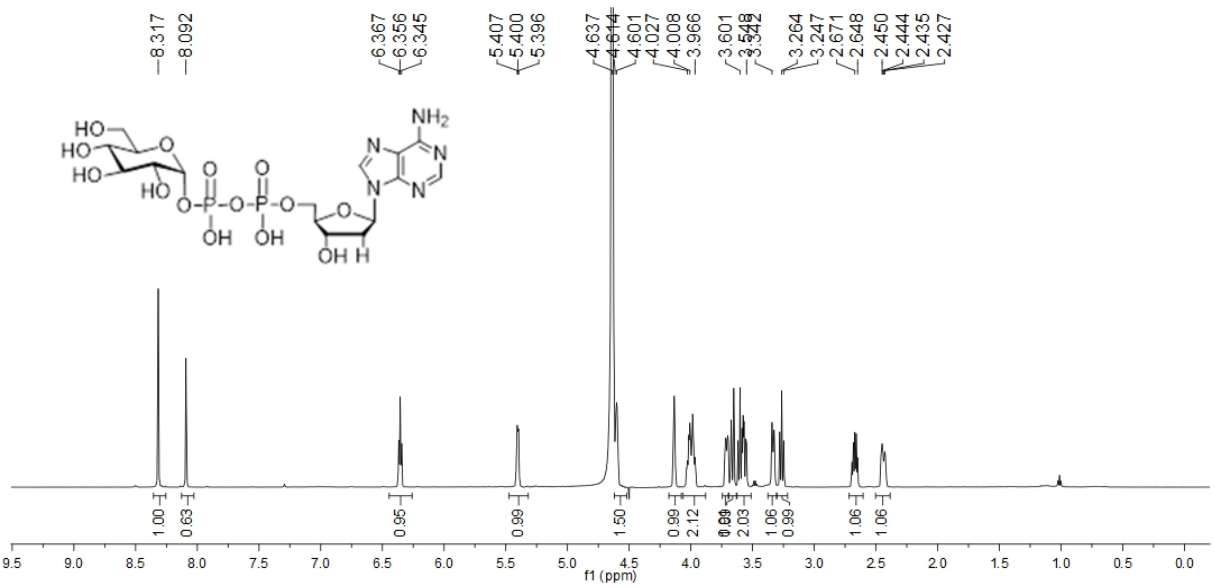
^{1}H-^{13}C HSQC spectrum of GDP-Glc (600/150 MHz, D_2O)



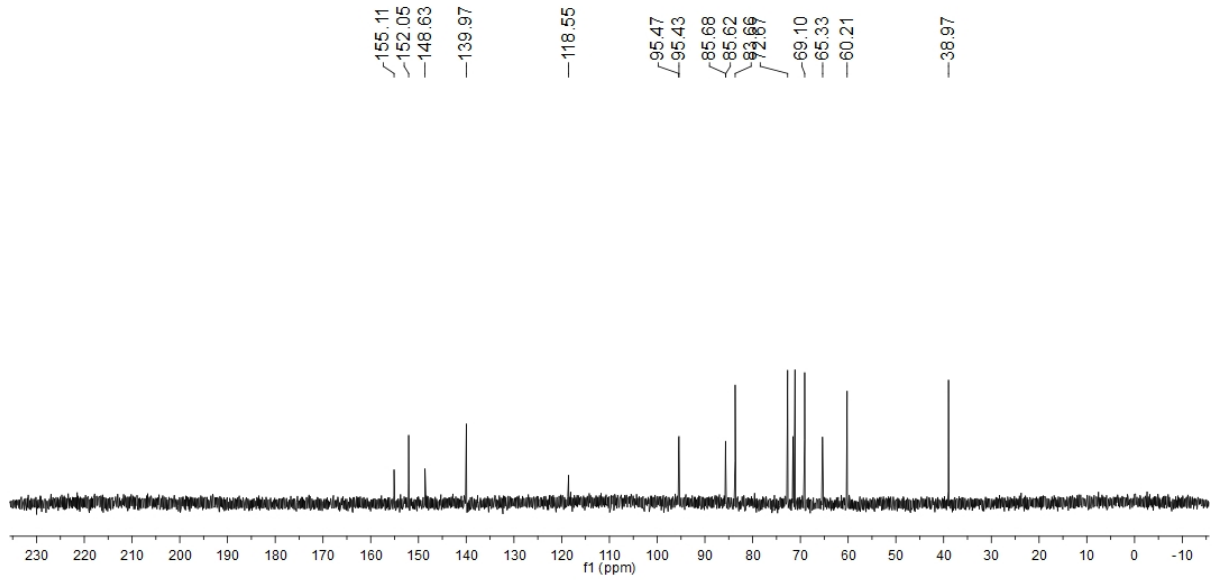
^{31}P NMR spectrum of GDP-Glc (243 MHz, D_2O)



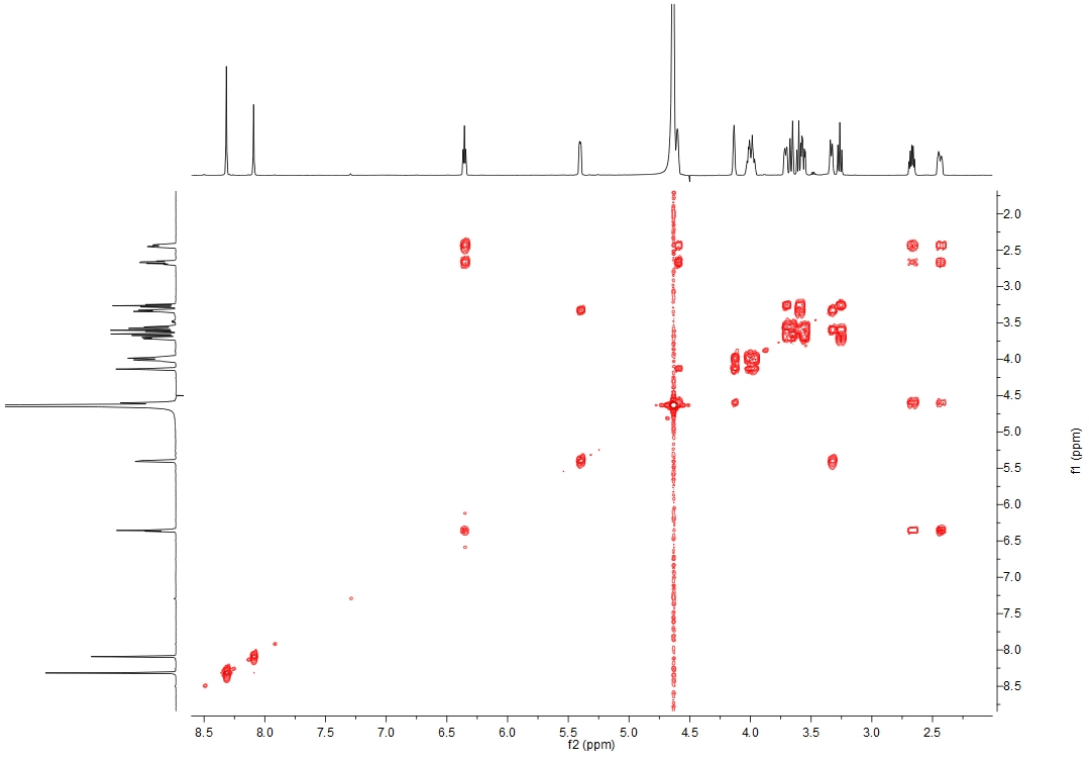
ESI-HRMS spectrum of GDP-Glc



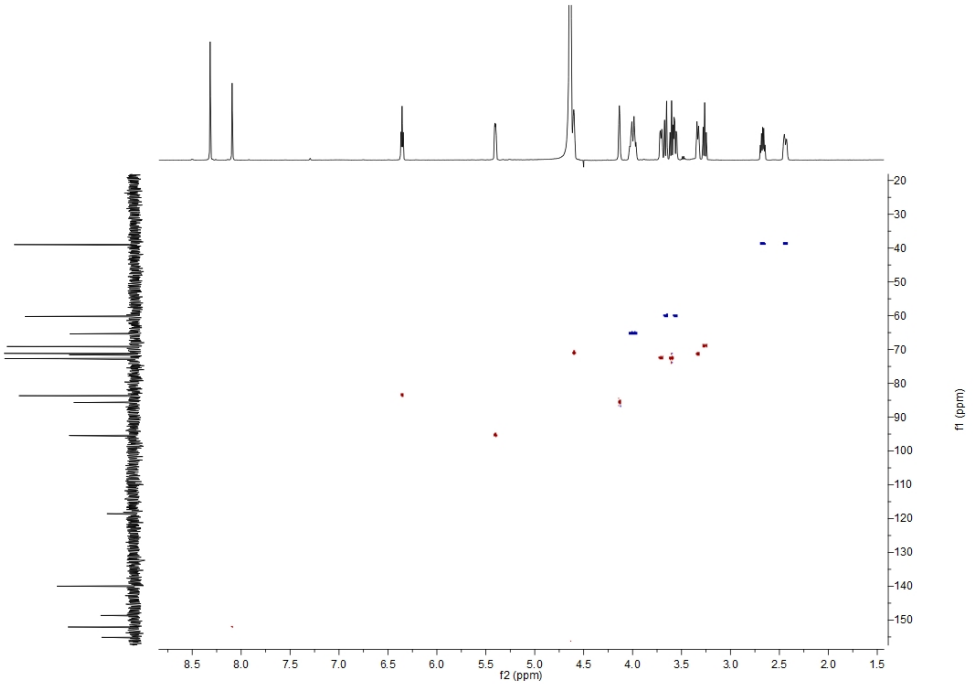
^{1}H NMR spectrum of dADP-Glc (600 MHz, D_2O)



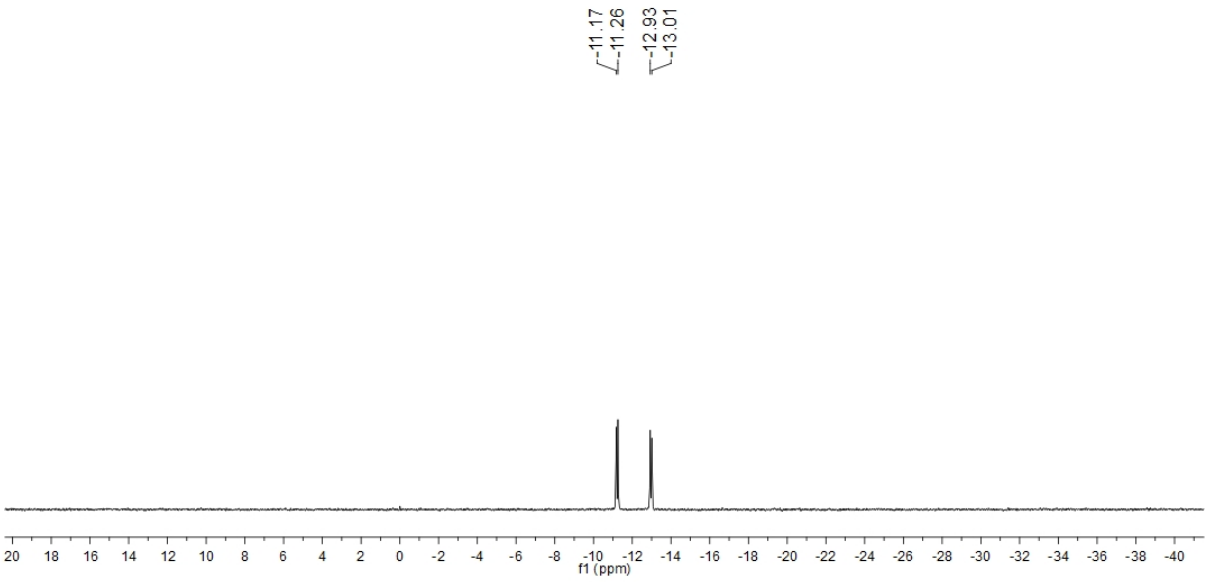
^{13}C NMR spectrum of dADP-Glc (150 MHz, D_2O)



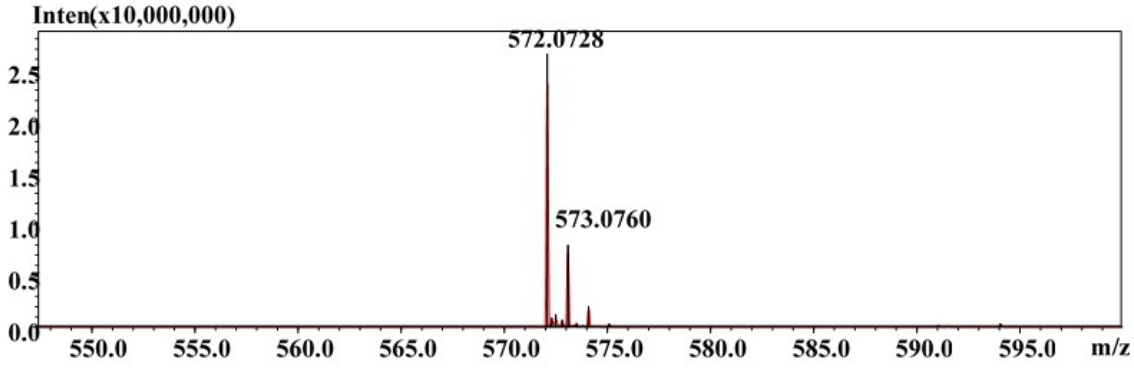
^{1}H-^{1}H COSY spectrum of dADP-Glc (600 MHz, D_2O)



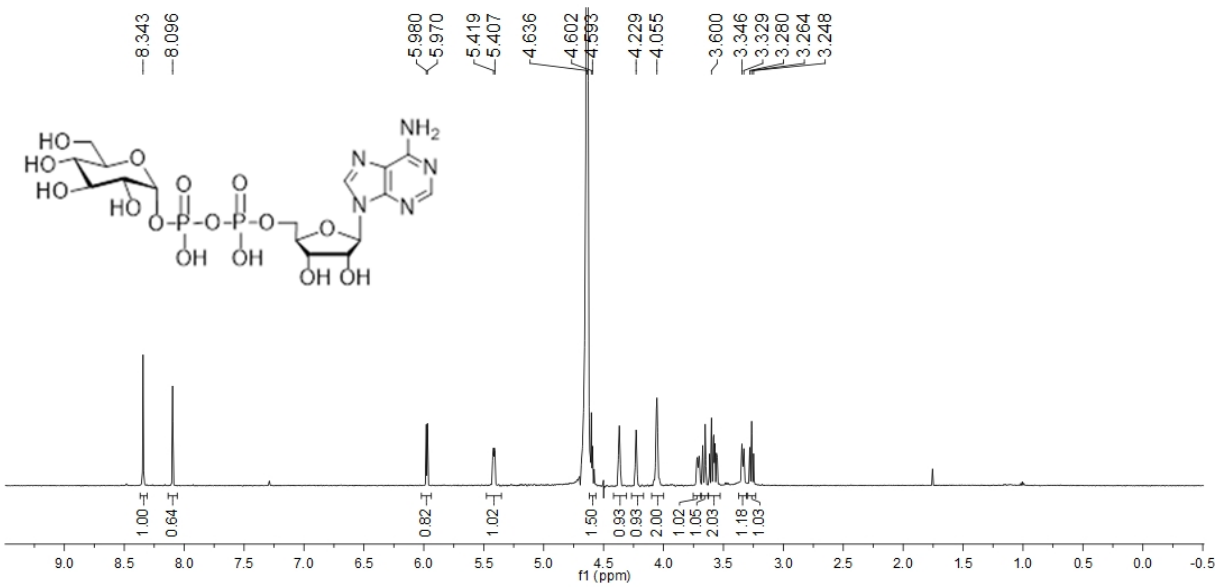
^{1}H-^{13}C HSQC spectrum of dADP-Glc (600/150 MHz, D_2O)



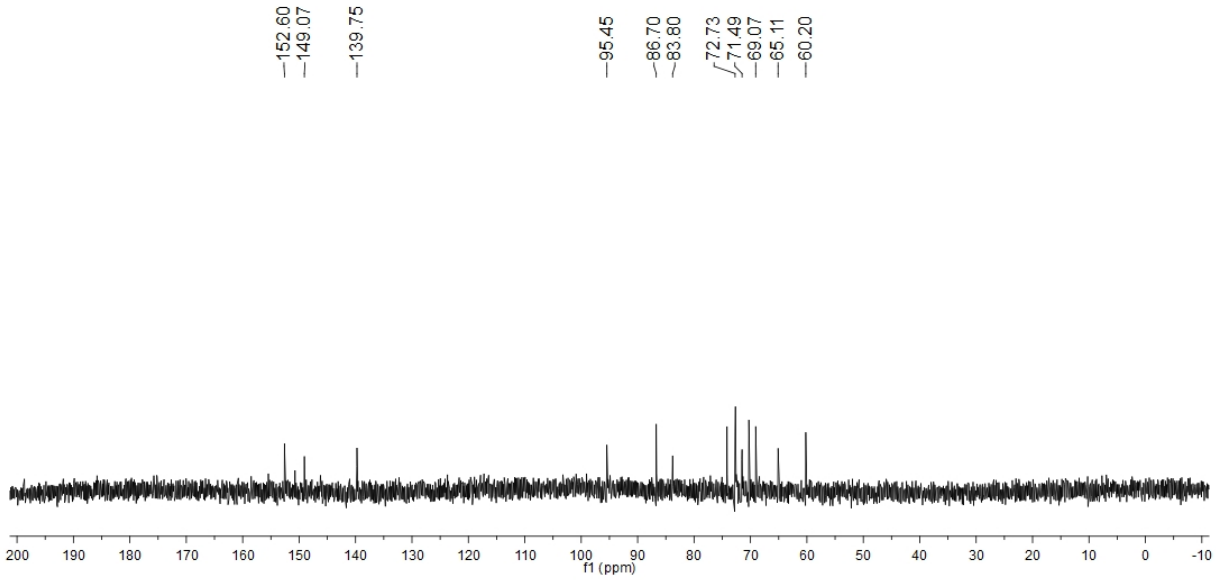
^{31}P NMR spectrum of dADP-Glc (243 MHz, D_2O)



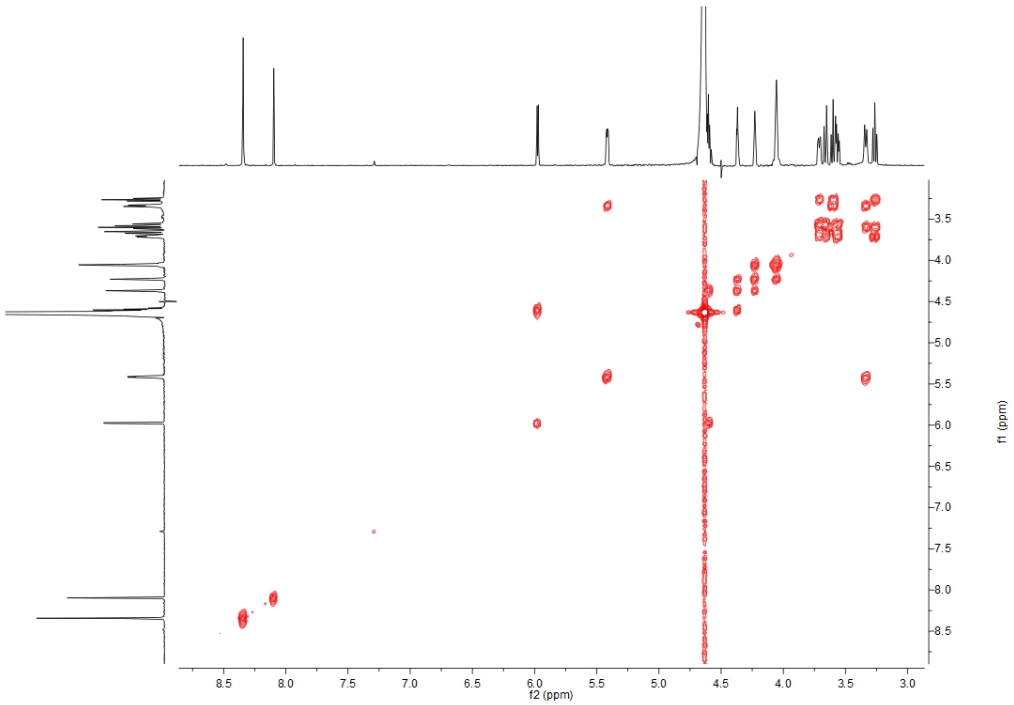
ESI-HRMS spectrum of dADP-Glc



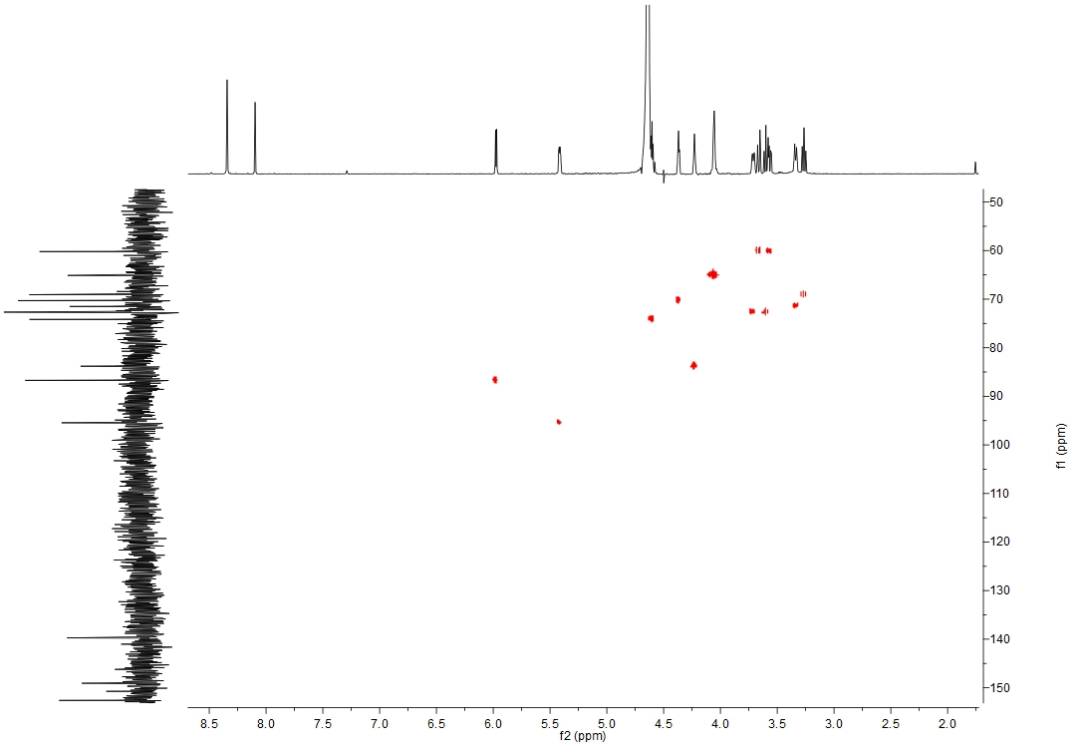
^{1}H NMR spectrum of ADP-Glc (600 MHz, D_2O)



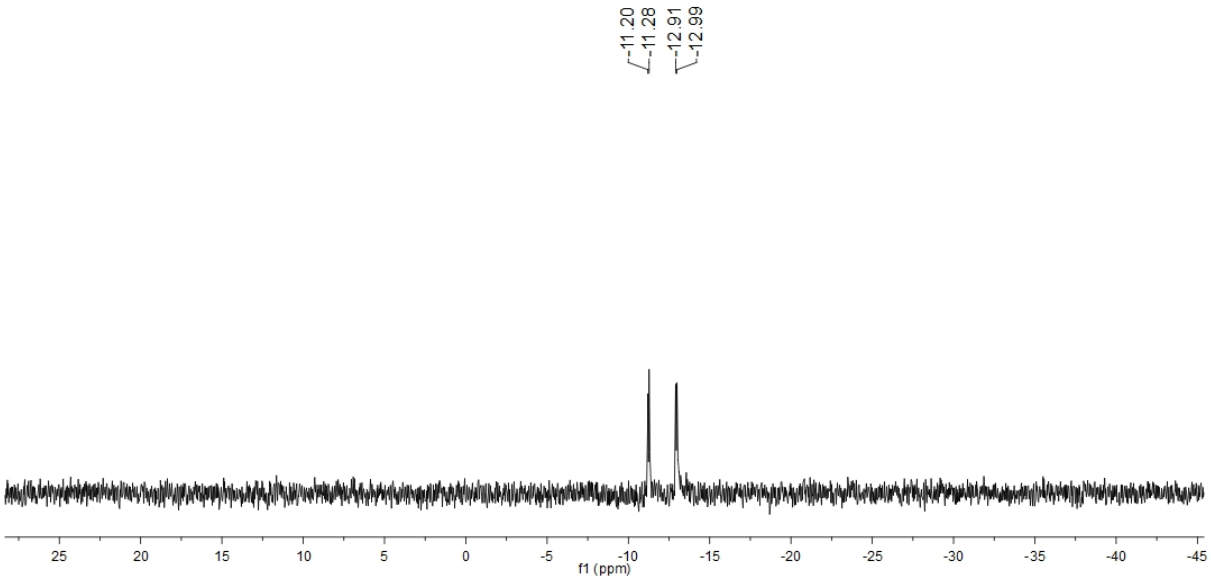
^{13}C NMR spectrum of ADP-Glc (150 MHz, D_2O)



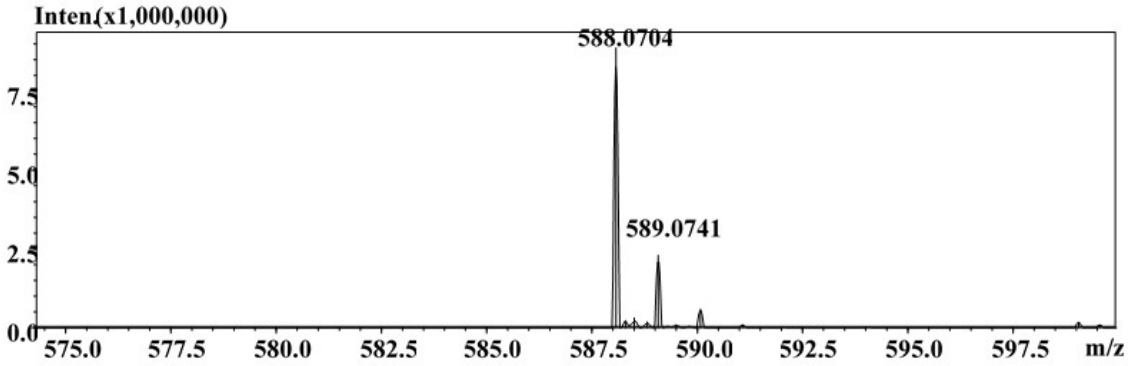
^{1}H-^{1}H COSY spectrum of ADP-Glc (600 MHz, D_2O)



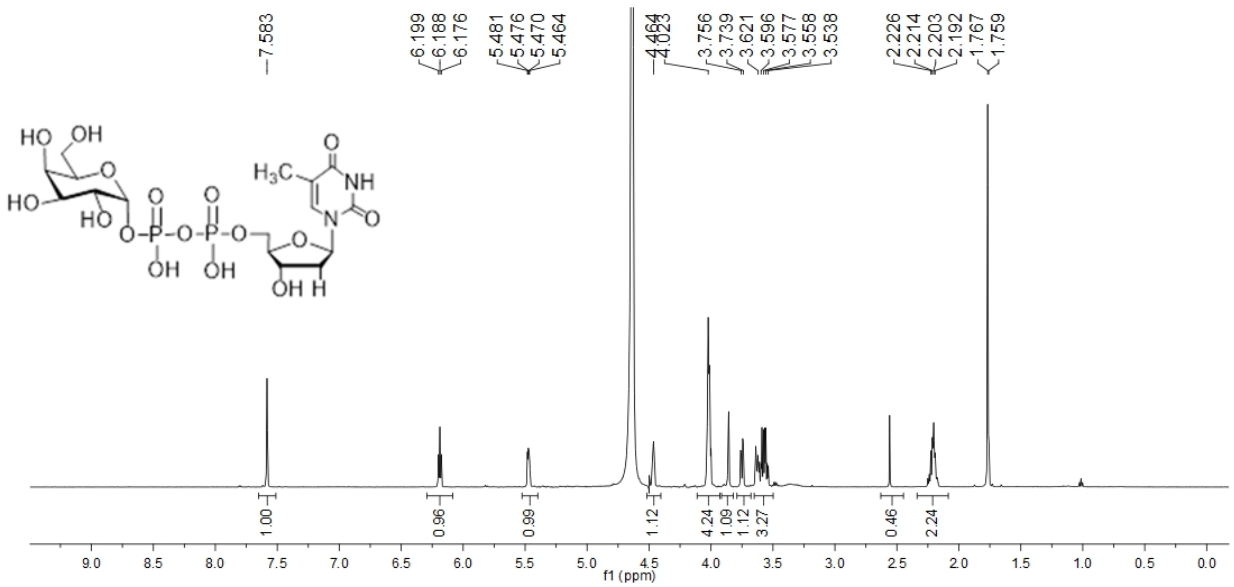
^{1}H-^{13}C HSQC spectrum of ADP-Glc (600/150 MHz, D_2O)



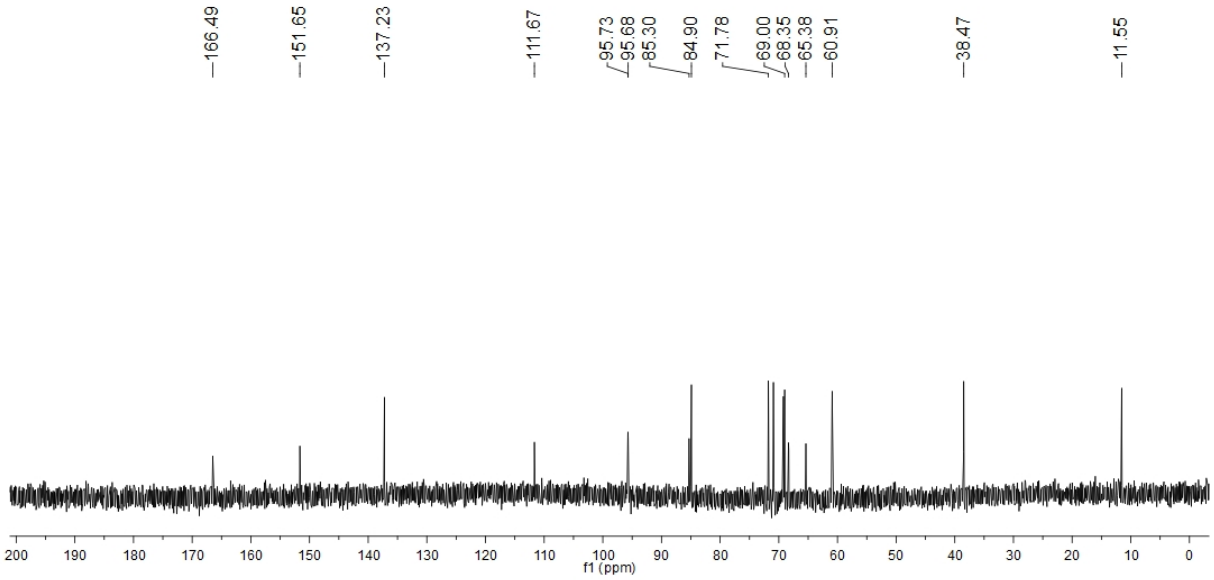
^{31}P NMR spectrum of ADP-Glc (243 MHz, D_2O)



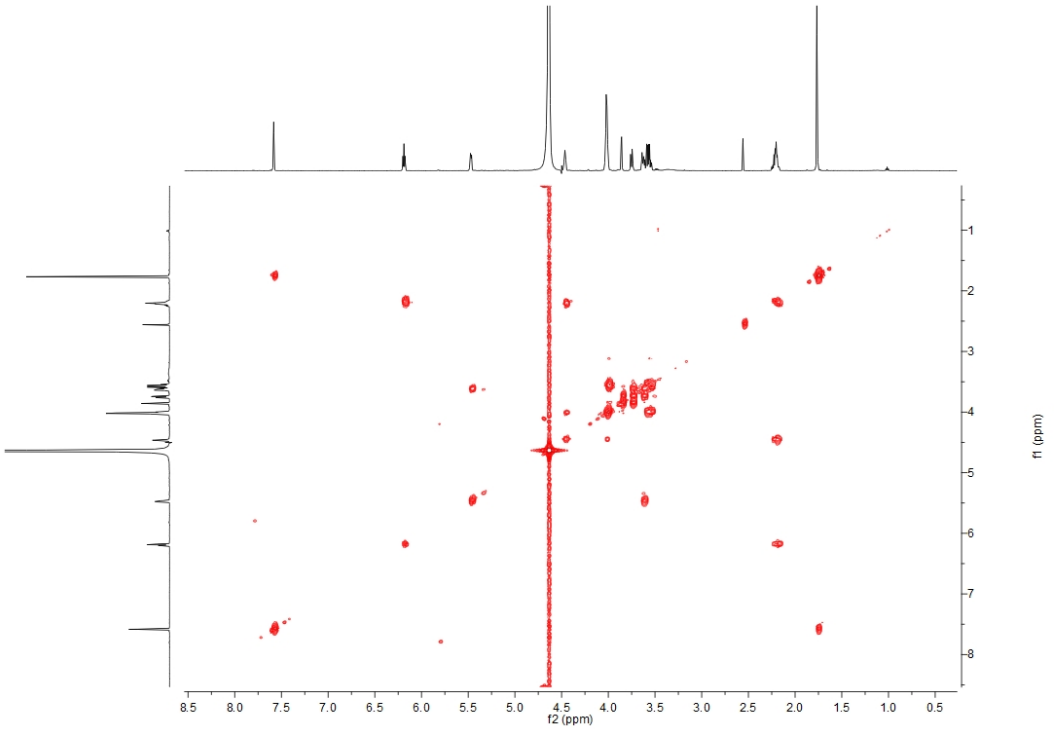
ESI-HRMS spectrum of ADP-Glc



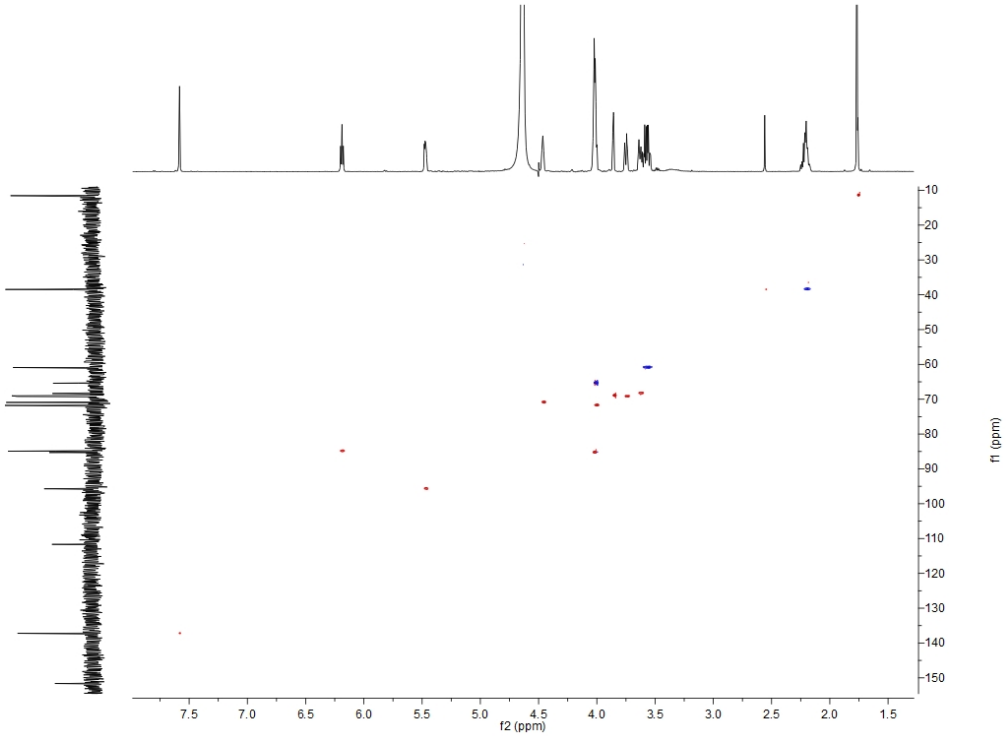
^{1}H NMR spectrum of dTDP-Gal (600 MHz, D_2O)



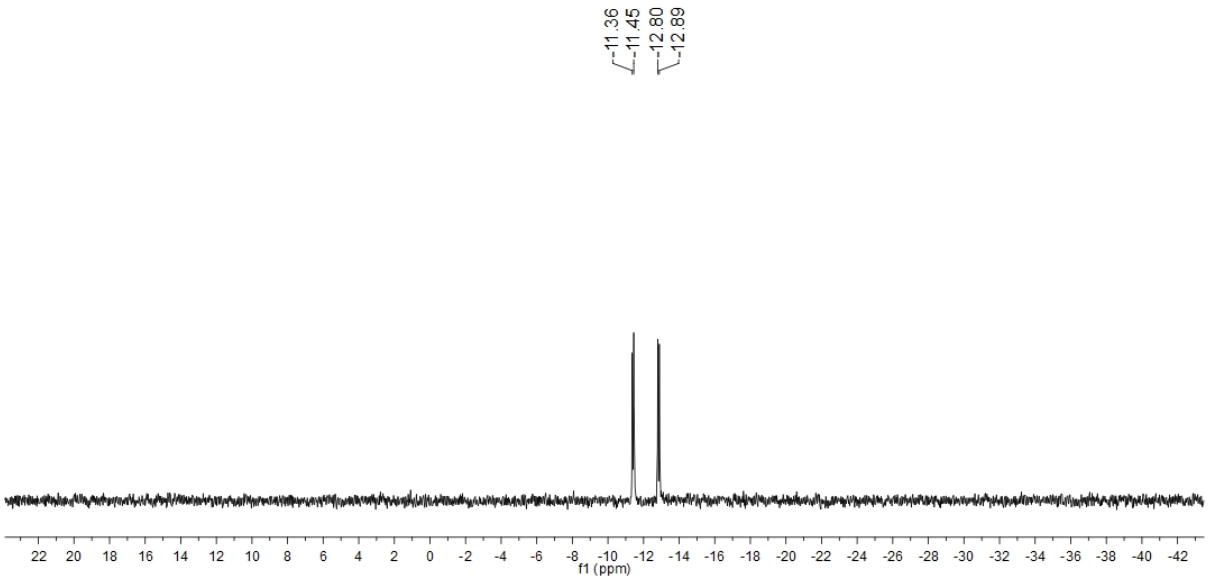
^{13}C NMR spectrum of dTDP-Gal (150 MHz, D_2O)



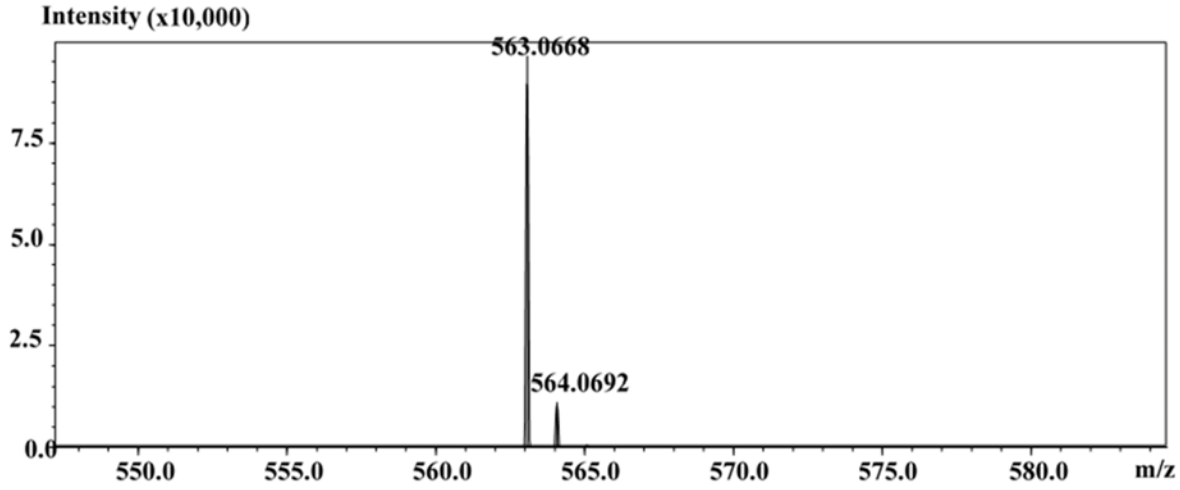
^{1}H-^{1}H COSY spectrum of dTDP-Gal (600 MHz, D_2O)



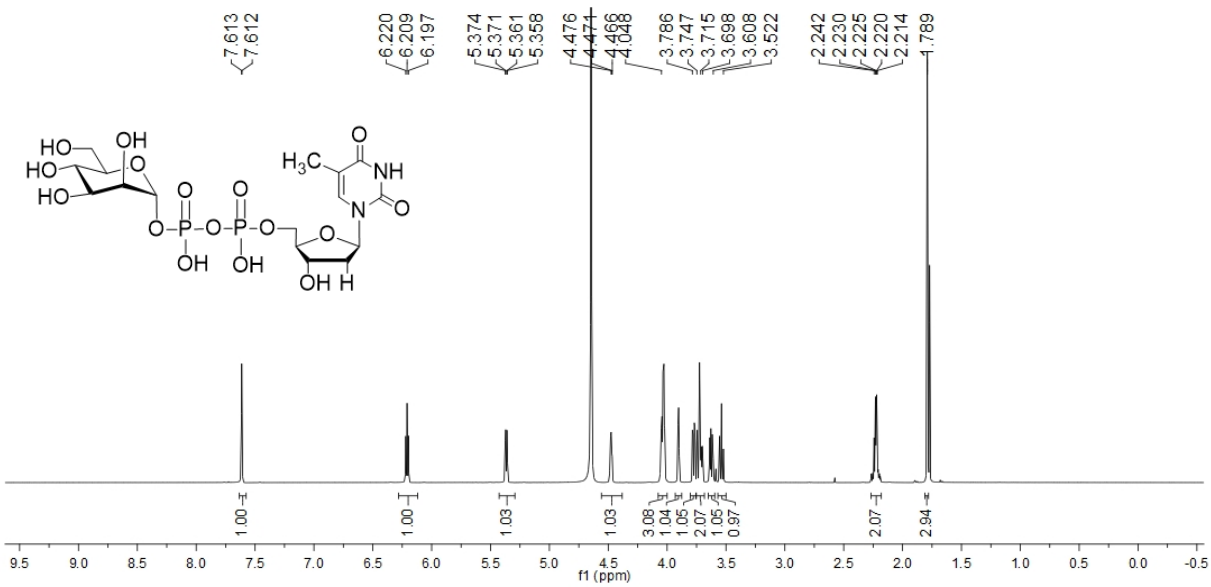
^{1}H-^{13}C HSQC spectrum of dTDP-Gal (600/150 MHz, D_2O)



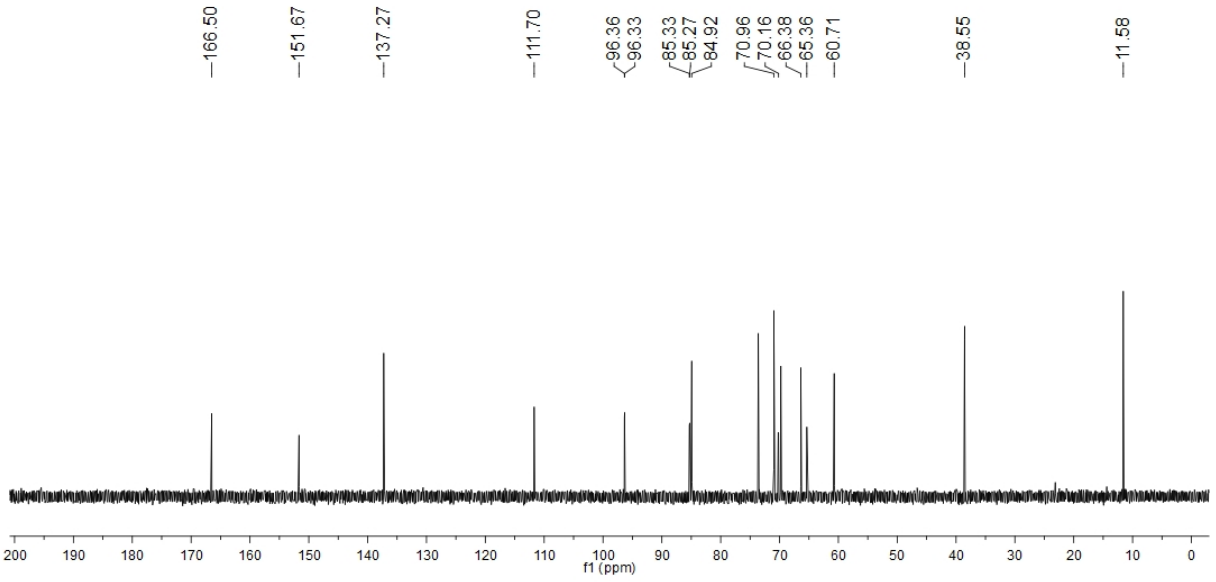
^{31}P NMR spectrum of dTDP-Gal (243 MHz, D_2O)



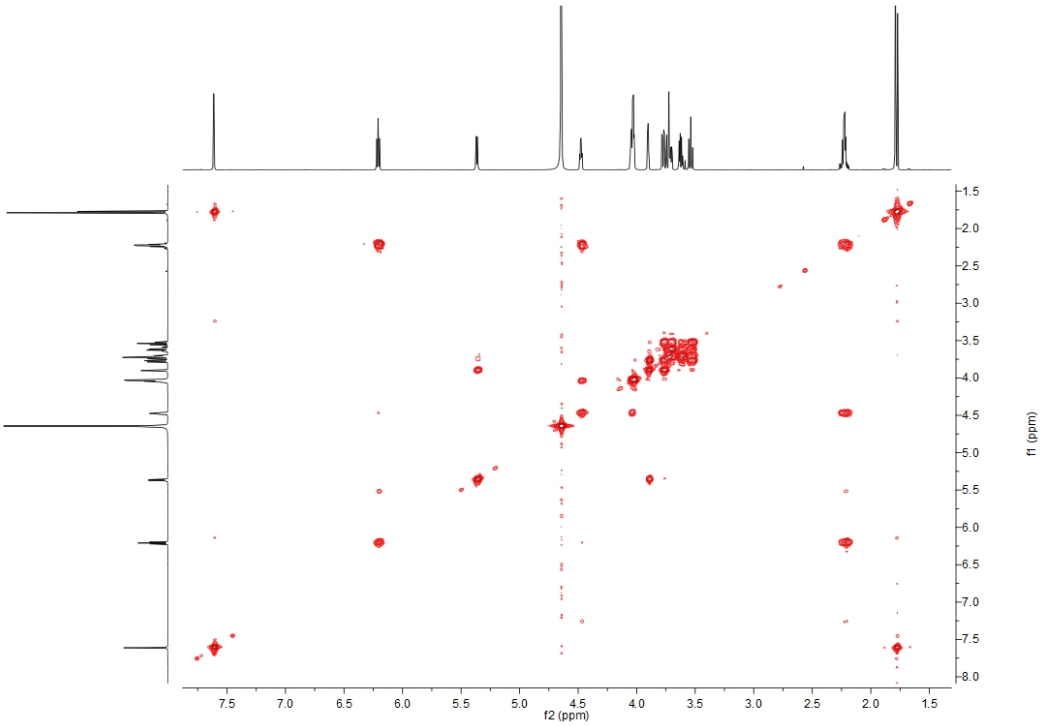
ESI-HRMS spectrum of dTDP-Gal



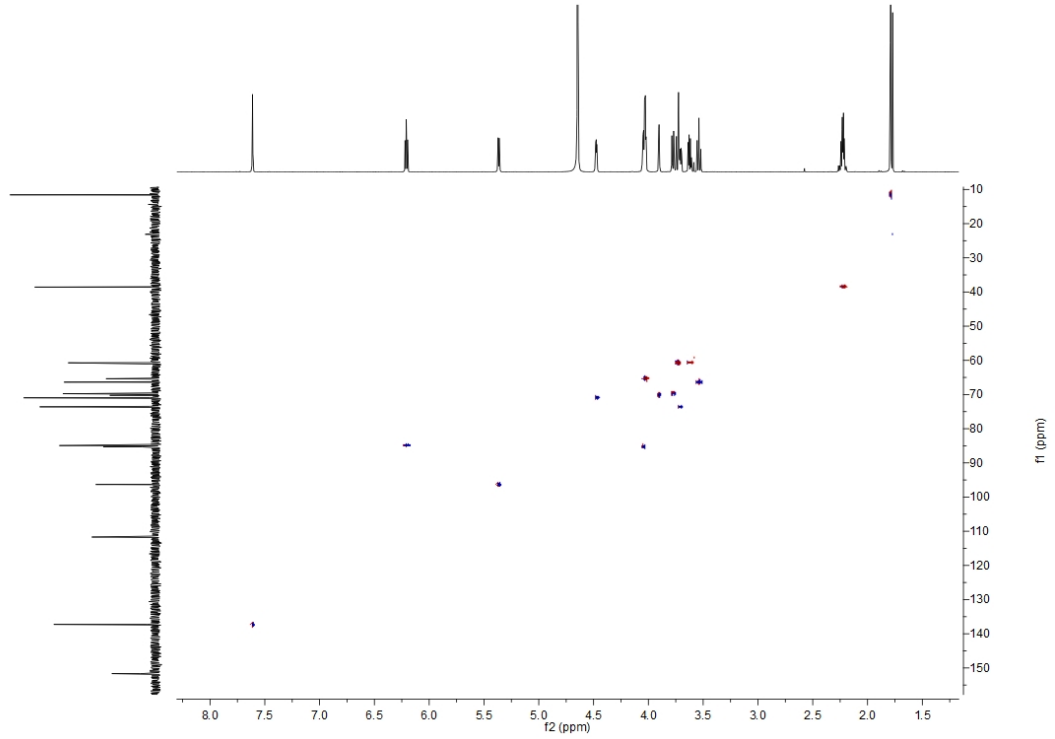
^{1}H NMR spectrum of dTDP-Man (600 MHz, D_2O)



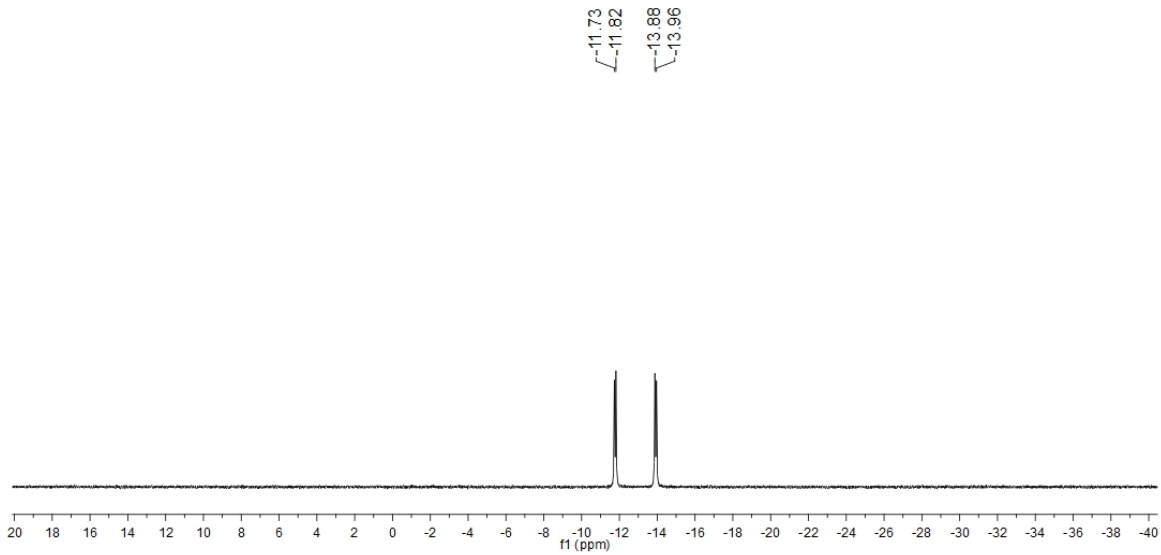
^{13}C NMR spectrum of dTDP-Man (150 MHz, D_2O)



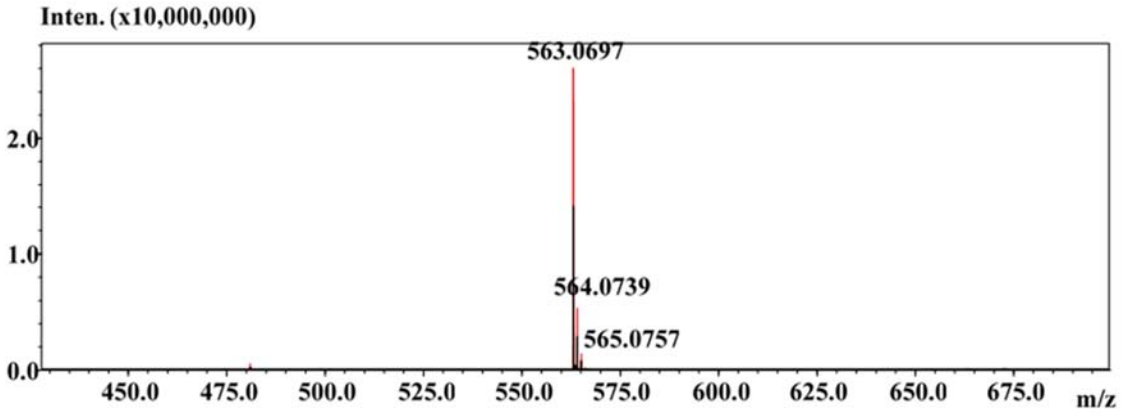
^{1}H-^{1}H COSY spectrum of dTDP-Man (600 MHz, D_2O)



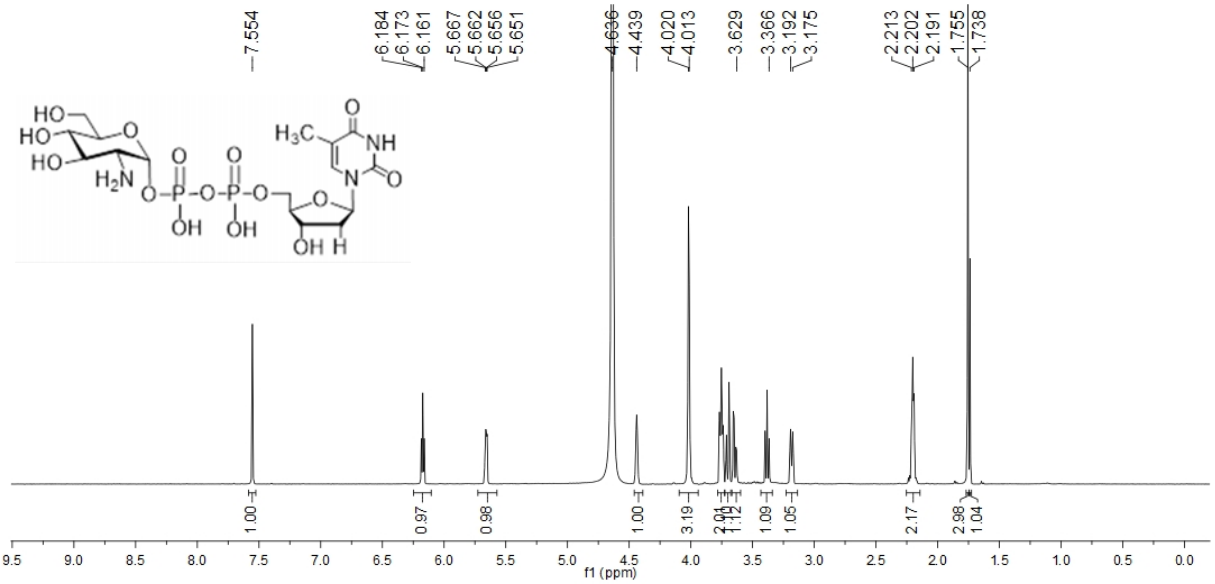
^{1}H-^{13}C HSQC spectrum of dTDP-Man (600/150 MHz, D_2O)



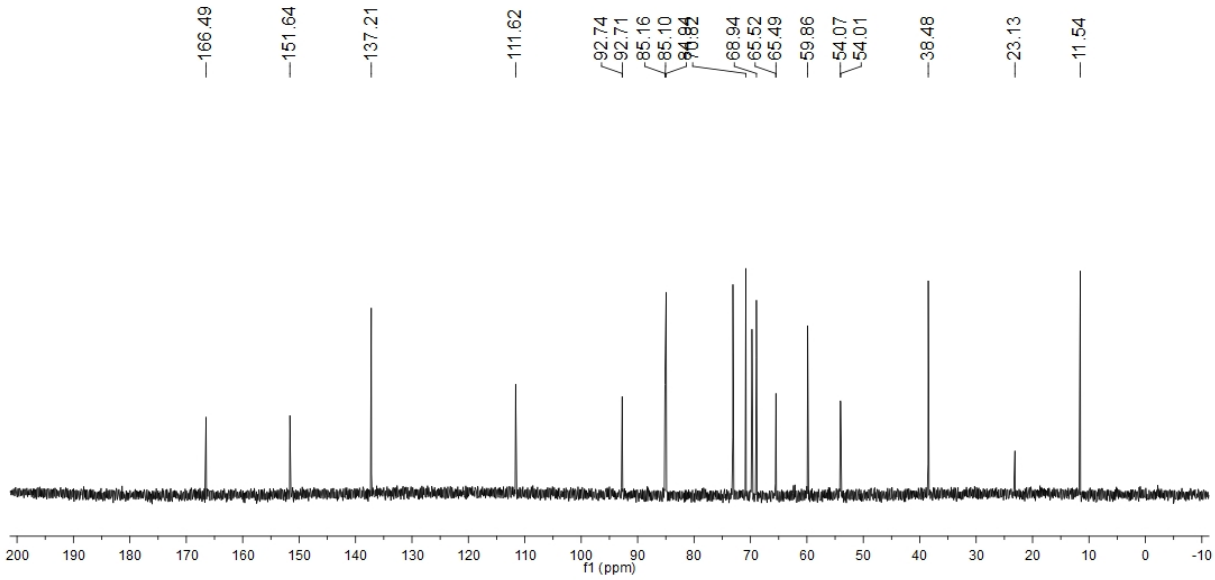
^{31}P NMR spectrum of dTDP-Man (243 MHz, D_2O)



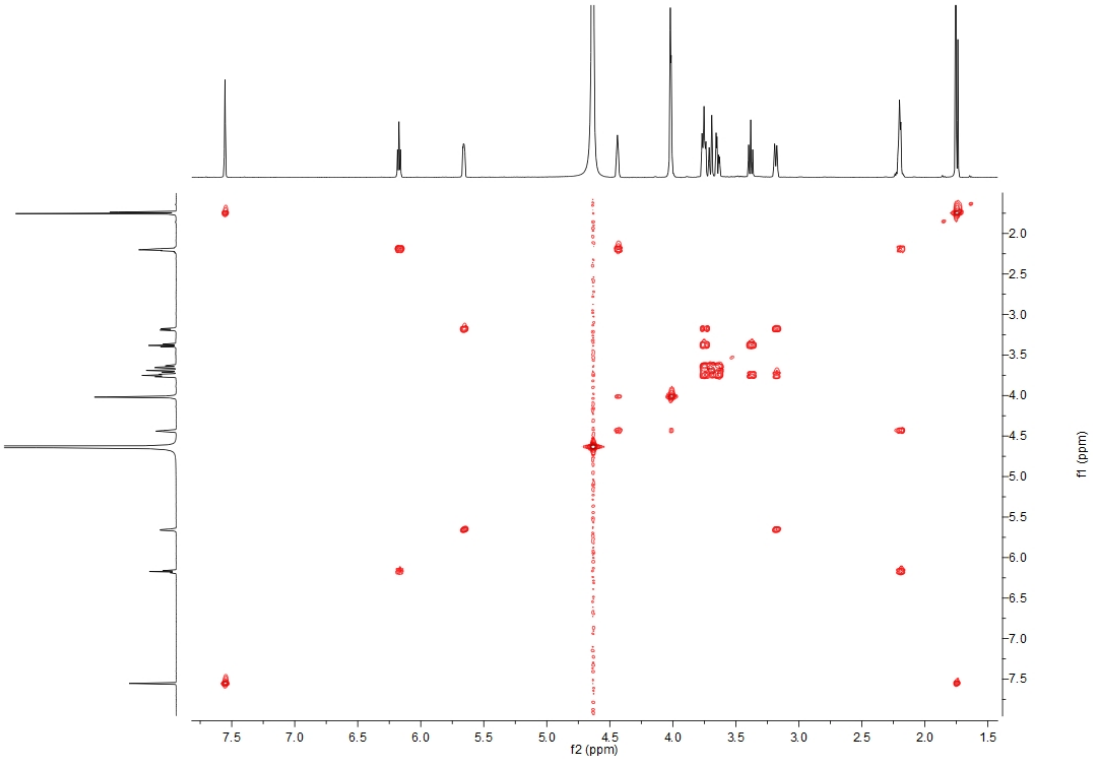
ESI-HRMS spectrum of dTDP-Man



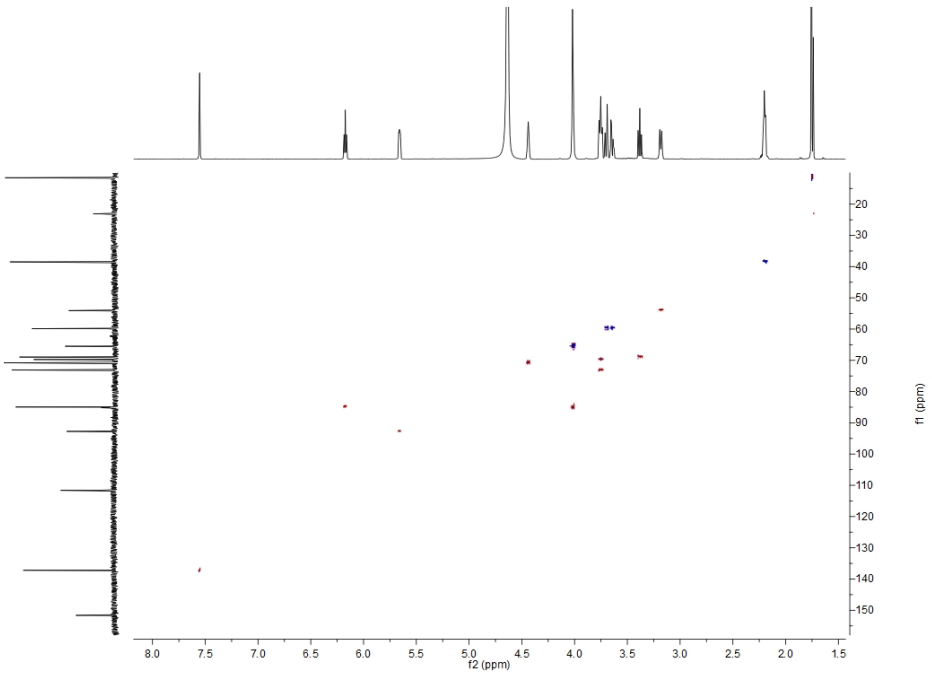
^{1}H NMR spectrum of dTDP-GlcNH$_2$ (600 MHz, D$_2$O)



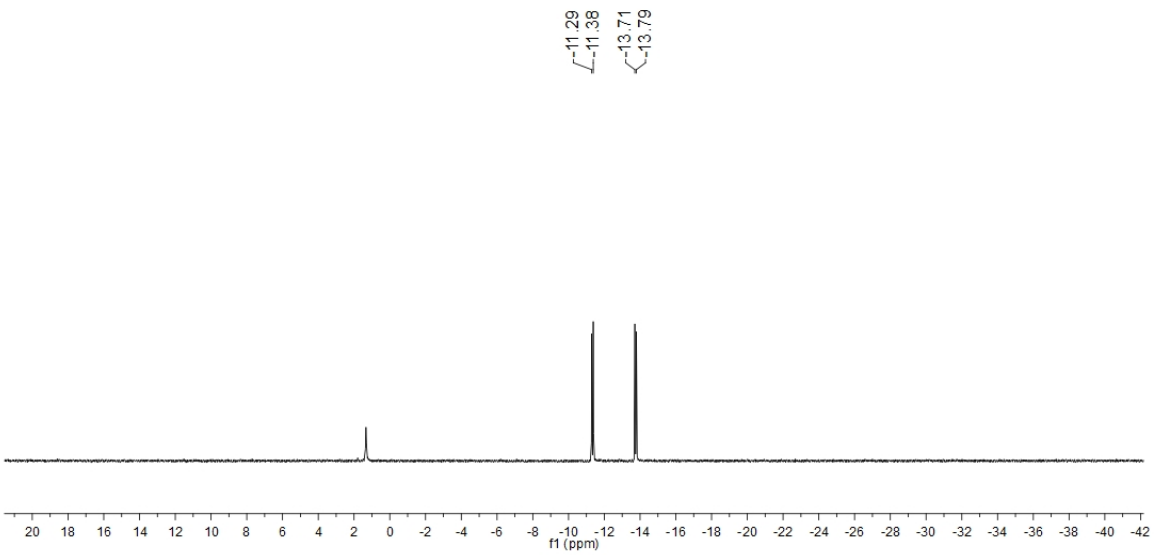
^{13}C NMR spectrum of dTDP-GlcNH$_2$ (150 MHz, D$_2$O)



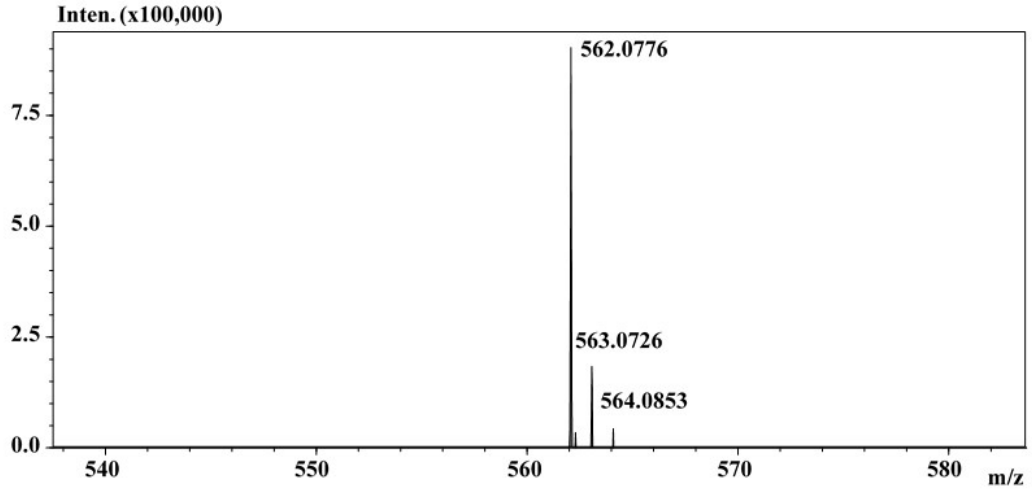
^{1}H-^{1}H COSY spectrum of dTDP-GlcNH$_2$ (600 MHz, D$_2$O)



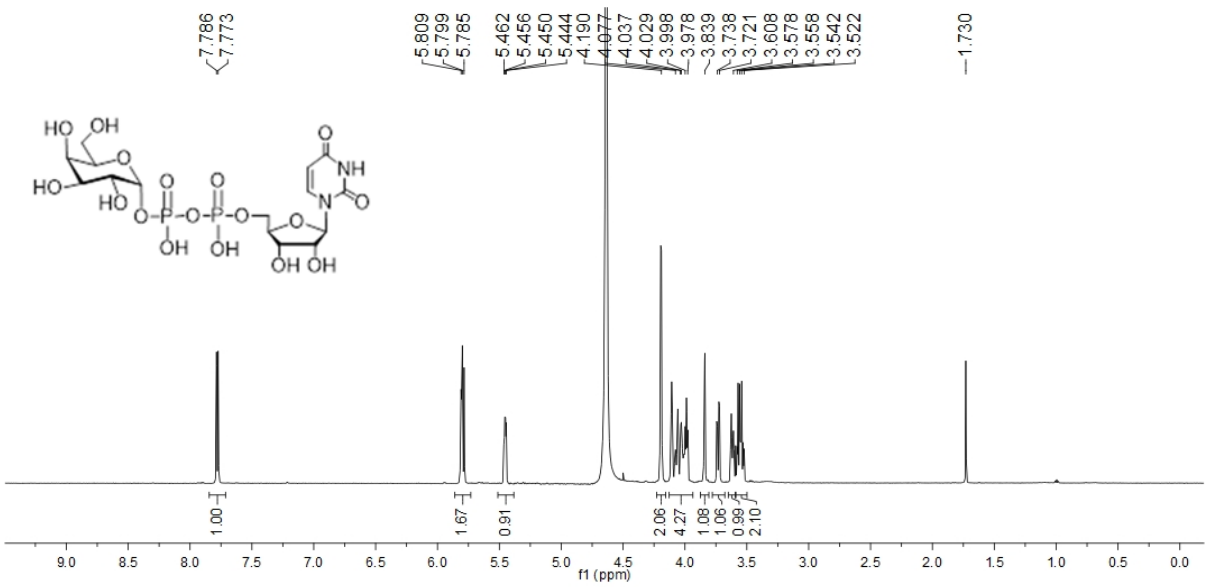
^{1}H-^{13}C HSQC spectrum of dTDP-GlcNH$_2$ (600/150 MHz, D$_2$O)



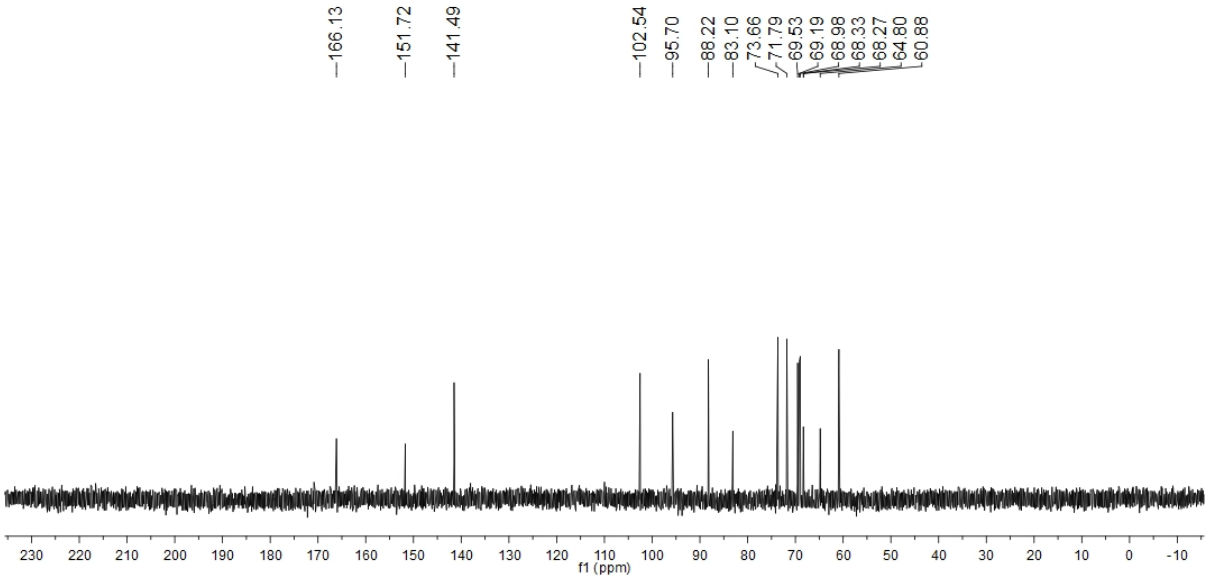
^{31}P NMR spectrum of dTDP-GlcNH$_2$ (243 MHz, D$_2$O)



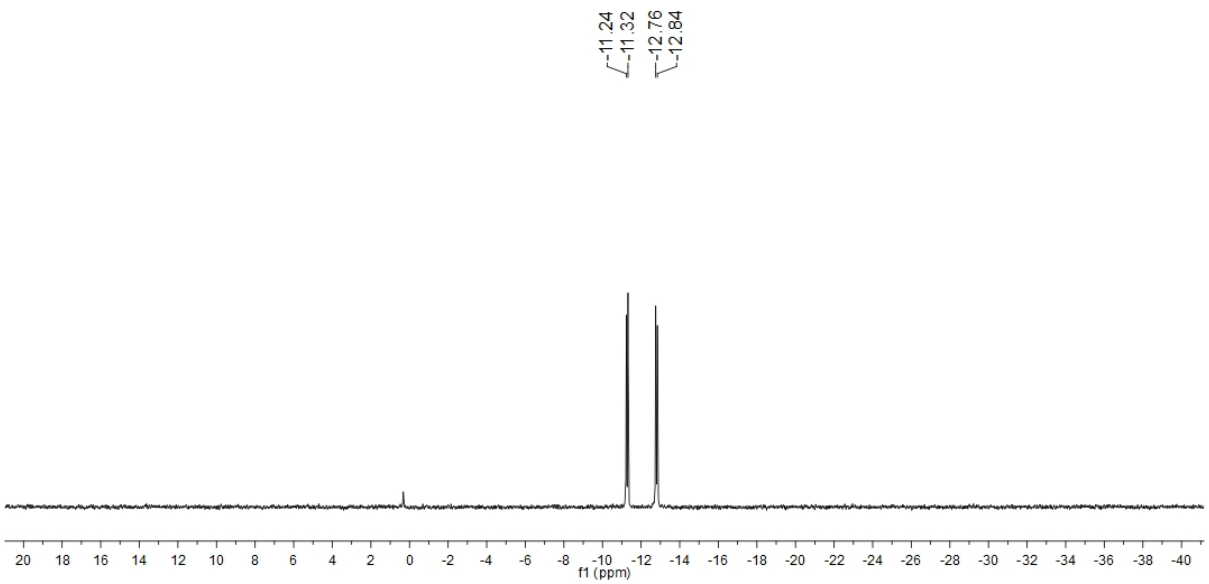
ESI-HRMS spectrum of dTDP-GlcNH$_2$



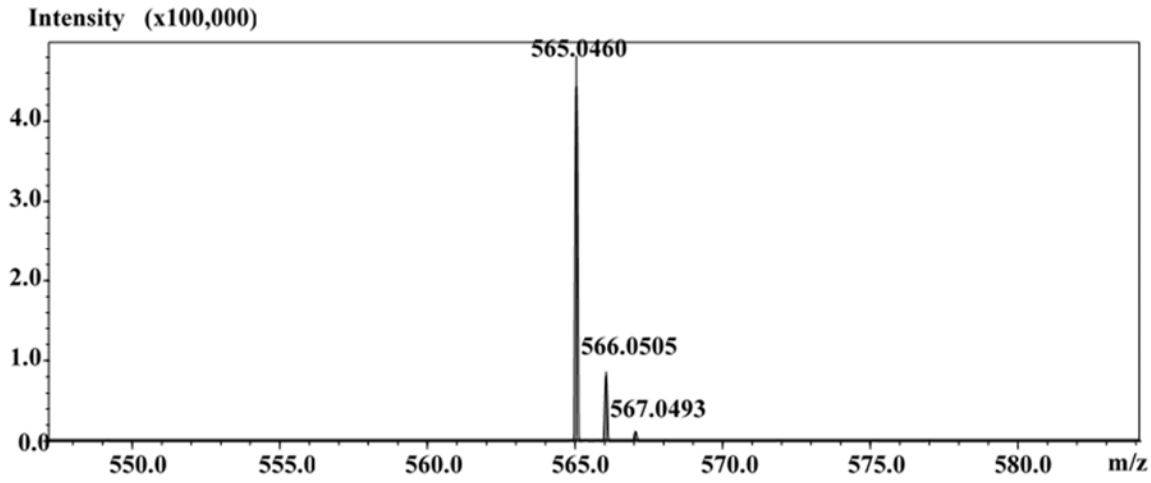
^{1}H NMR spectrum of UDP-Gal (600 MHz, D_2O)



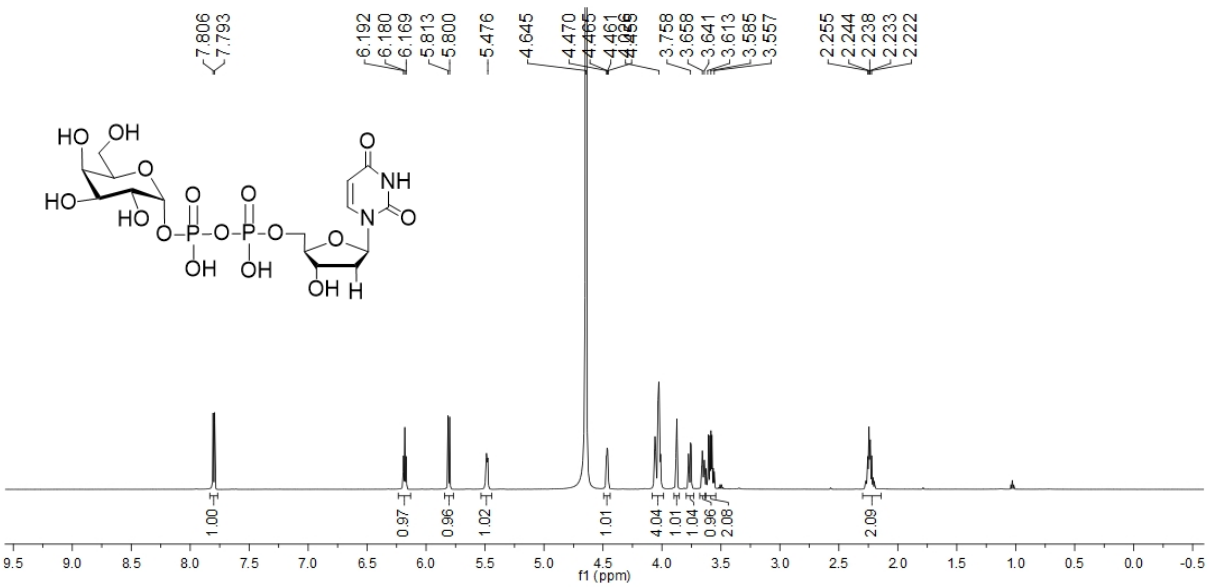
^{13}C NMR spectrum of UDP-Gal (150 MHz, D_2O)



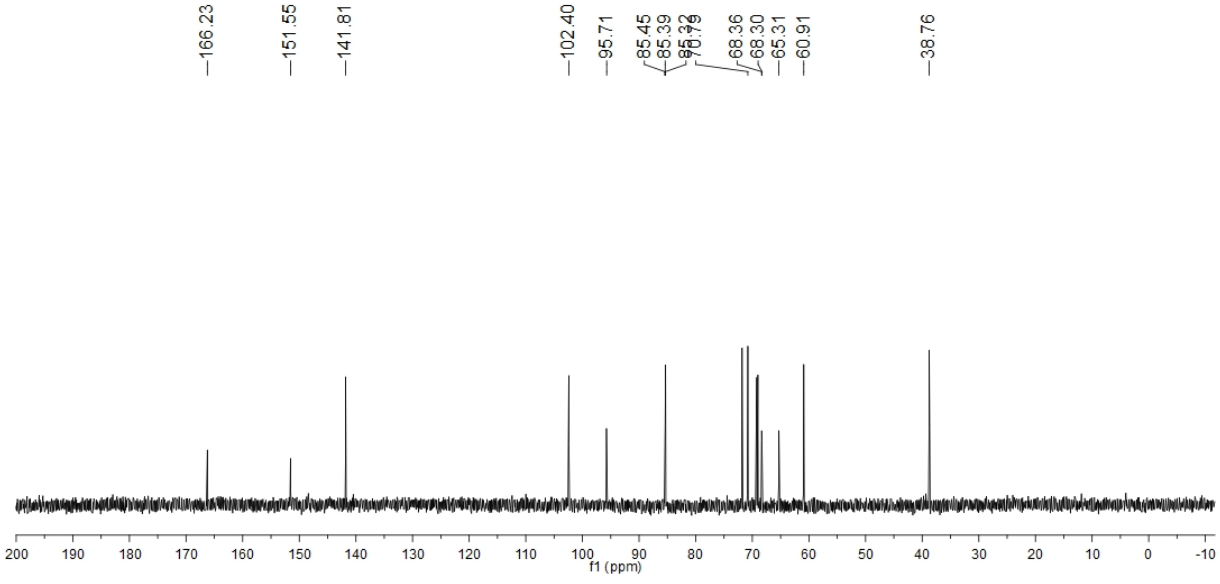
^{31}P NMR spectrum of UDP-Gal (243 MHz, D_2O)



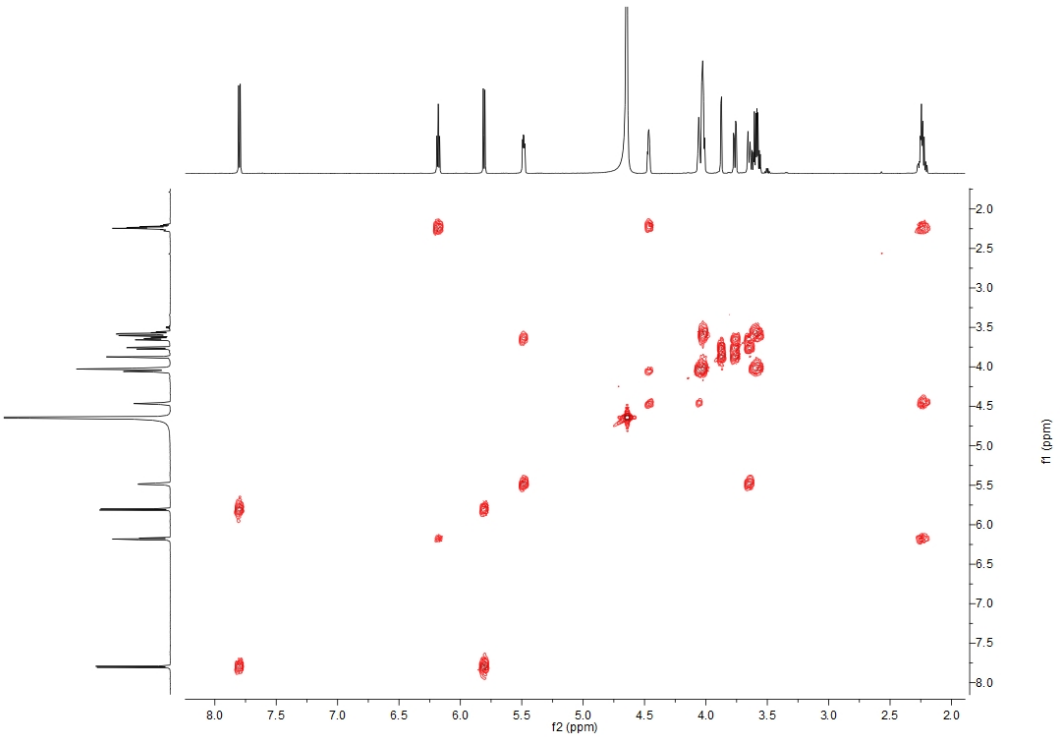
ESI-HRMS spectrum of UDP-Gal



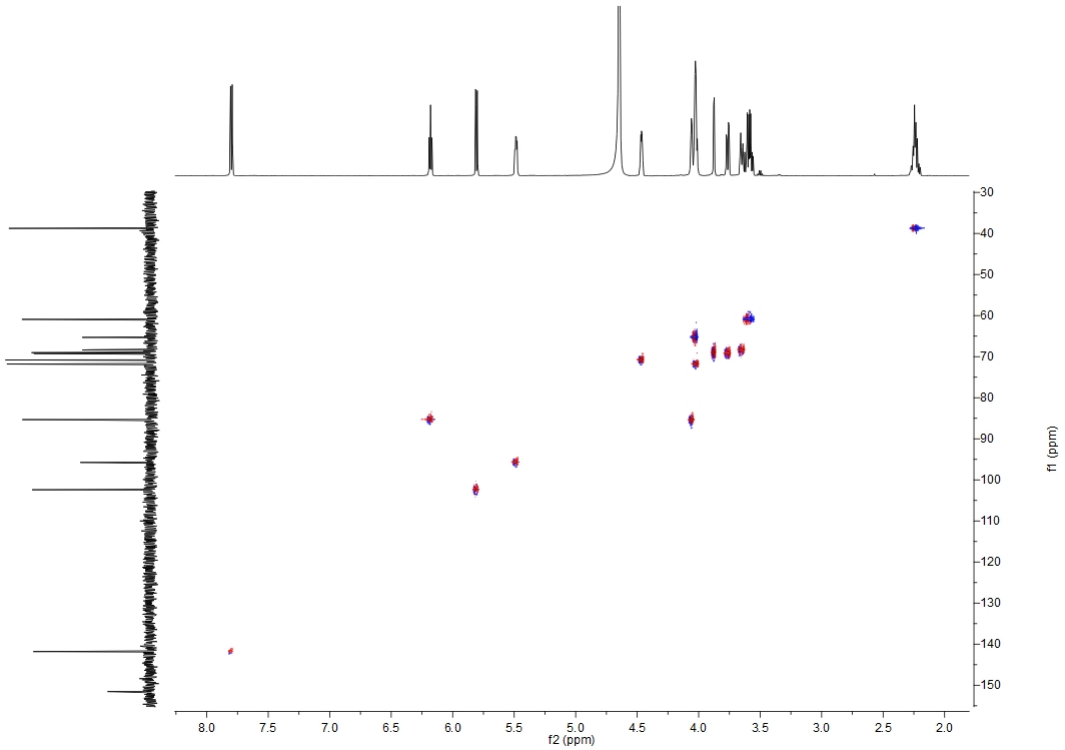
^{1}H NMR spectrum of dUDP-Gal (600 MHz, D_2O)



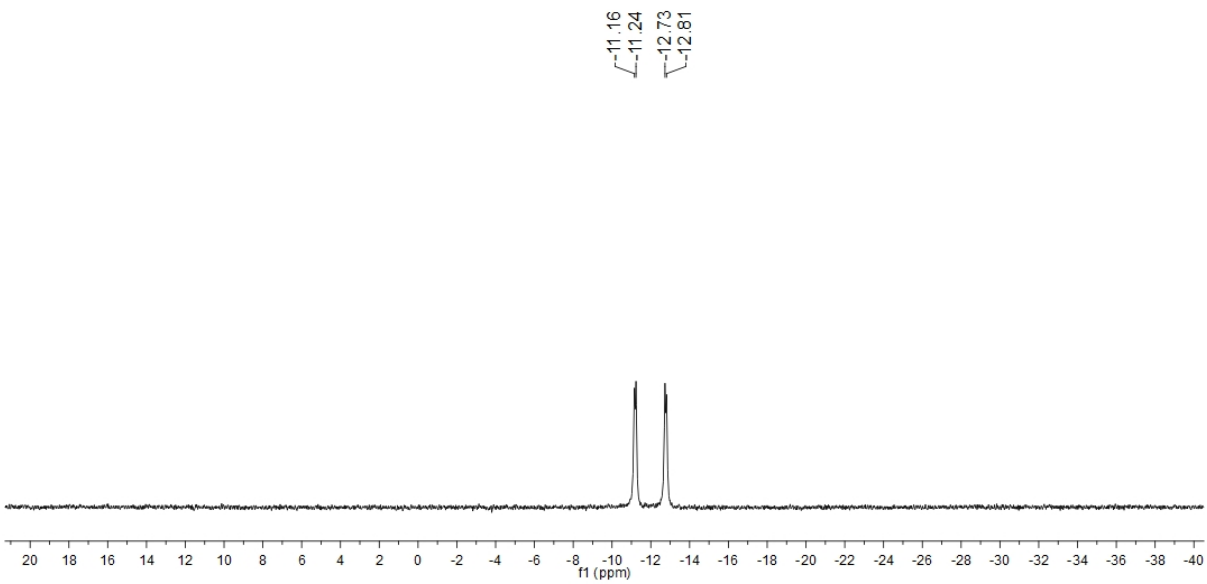
^{13}C NMR spectrum of dUDP-Gal (150 MHz, D_2O)



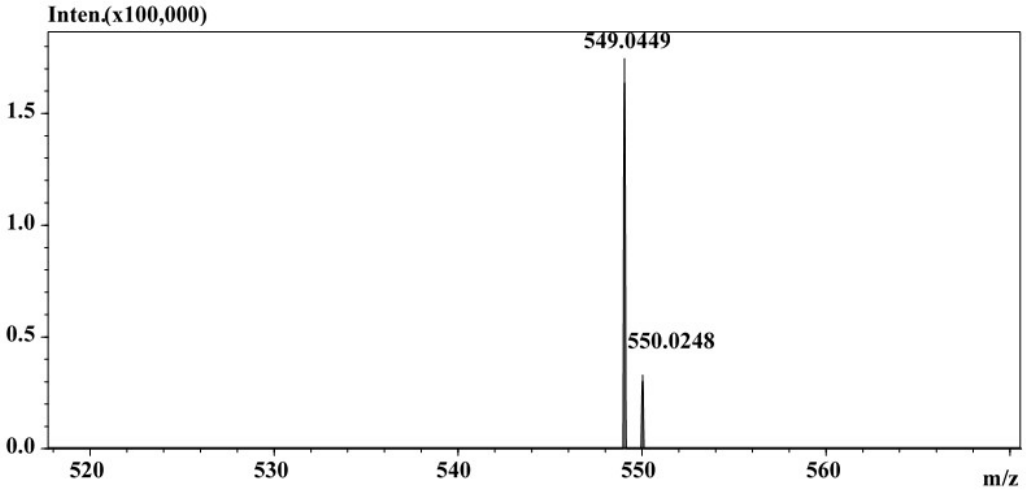
^{1}H-^{1}H COSY spectrum of dUDP-Gal (600 MHz, D_2O)



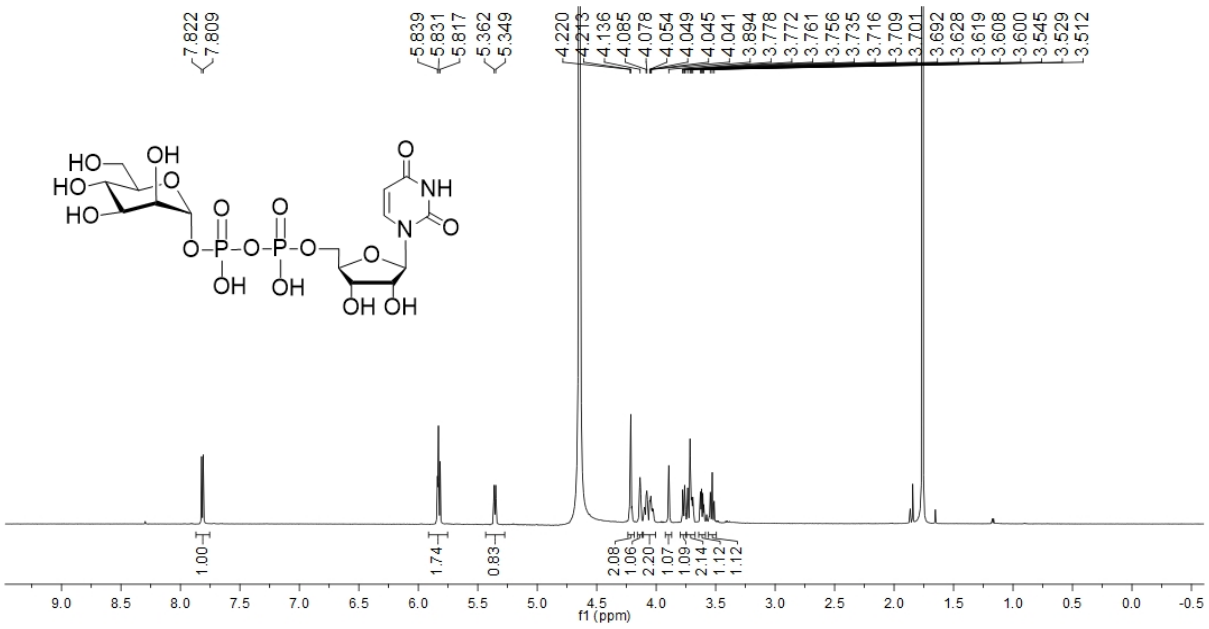
^{1}H-^{13}C HSQC spectrum of dUDP-Gal (600/150 MHz, D_2O)



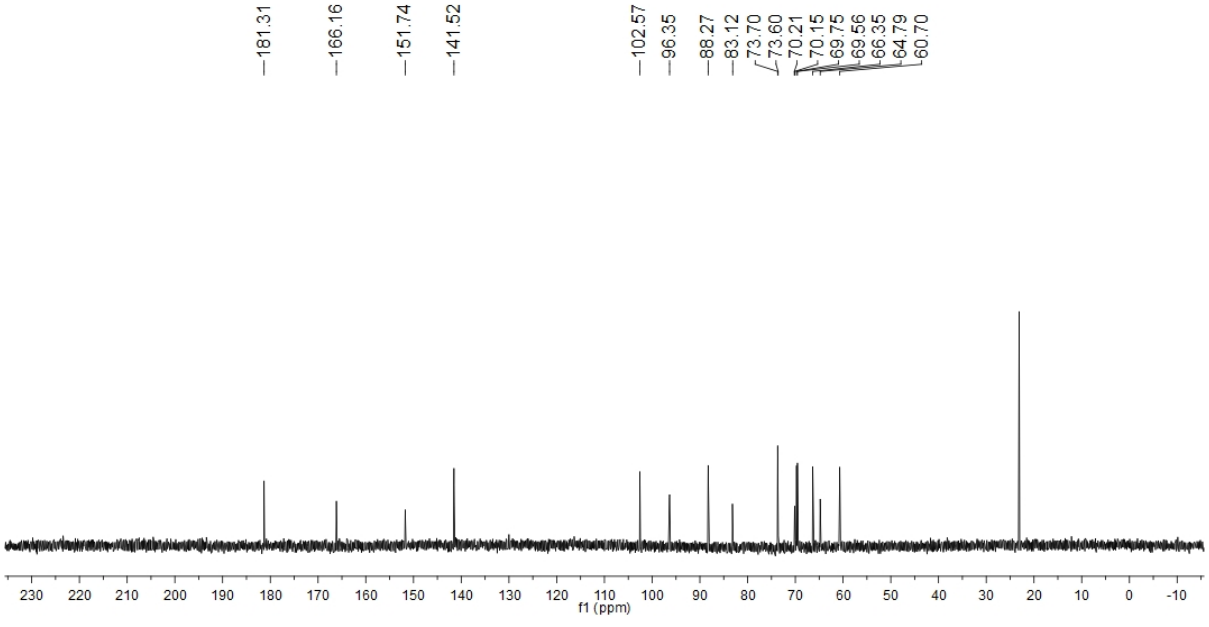
^{31}P NMR spectrum of dUDP-Gal (243 MHz, D_2O)



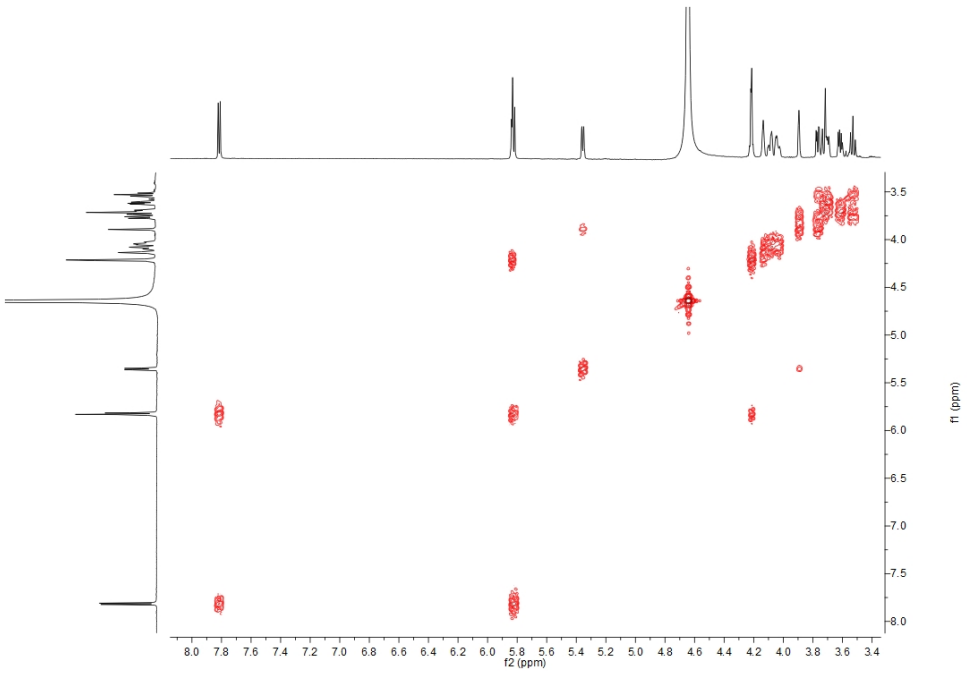
ESI-HRMS spectrum of dUDP-Gal



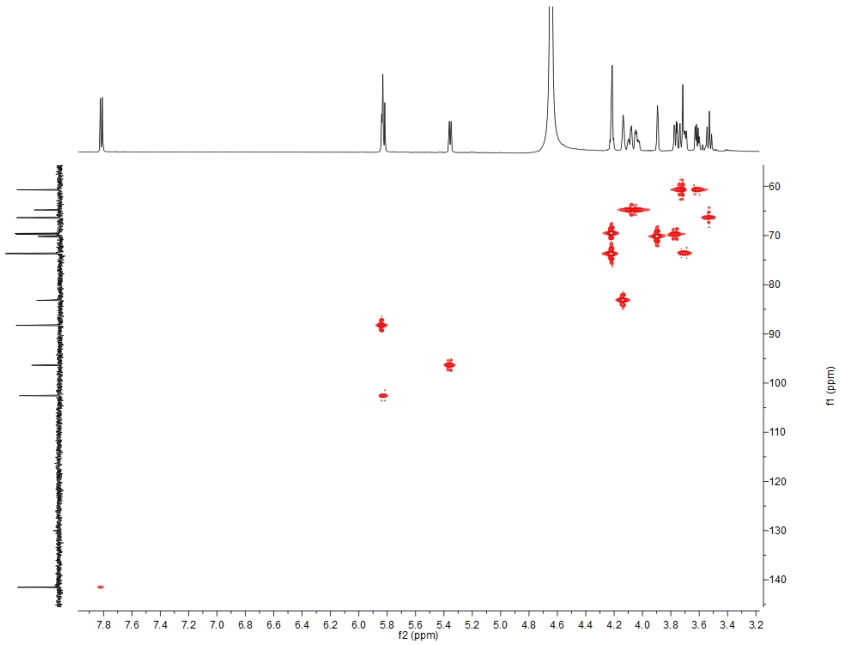
^{1}H NMR spectrum of UDP-Man (600 MHz, D_2O)



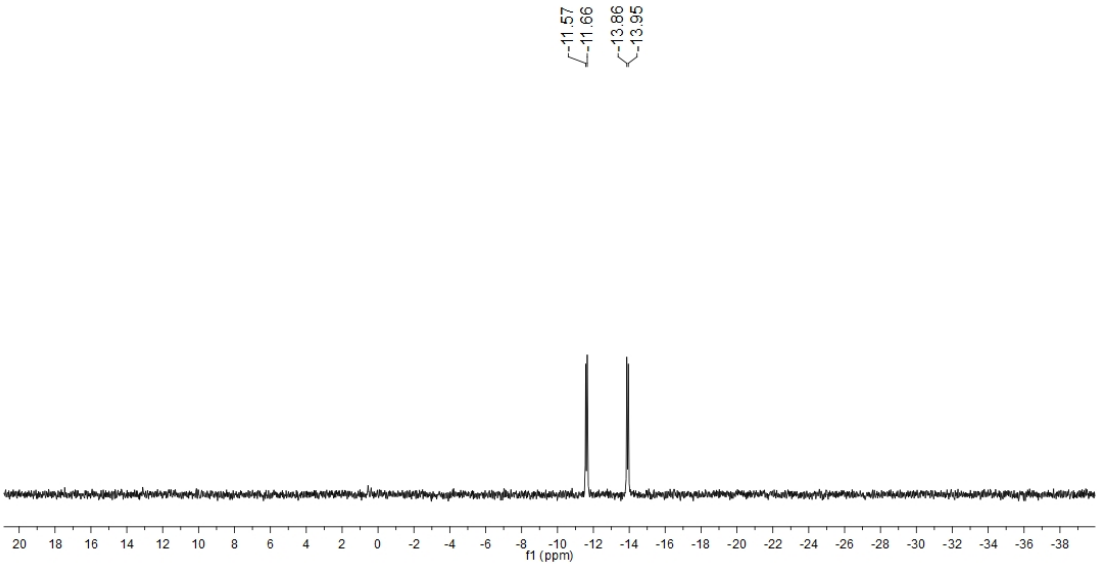
^{13}C NMR spectrum of UDP-Man (150 MHz, D_2O)



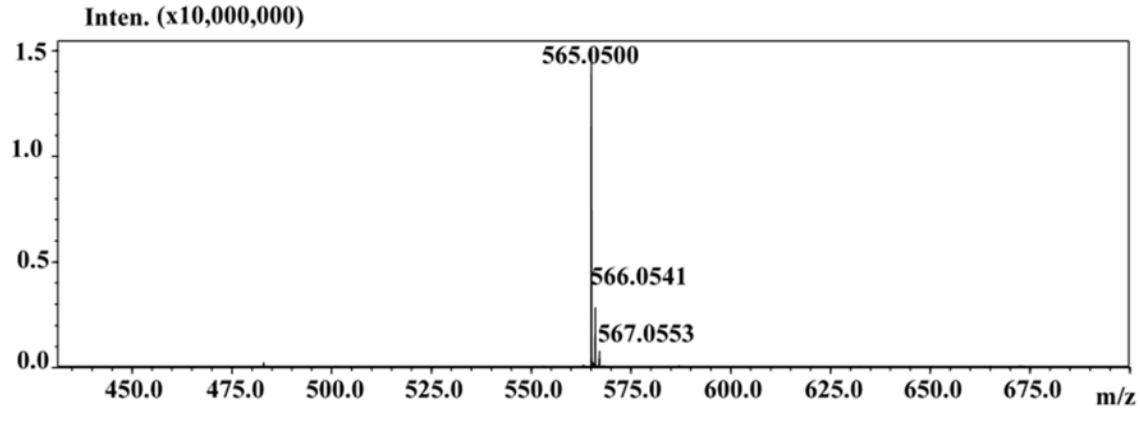
^{1}H-^{1}H COSY spectrum of UDP-Man (600 MHz, D_2O)



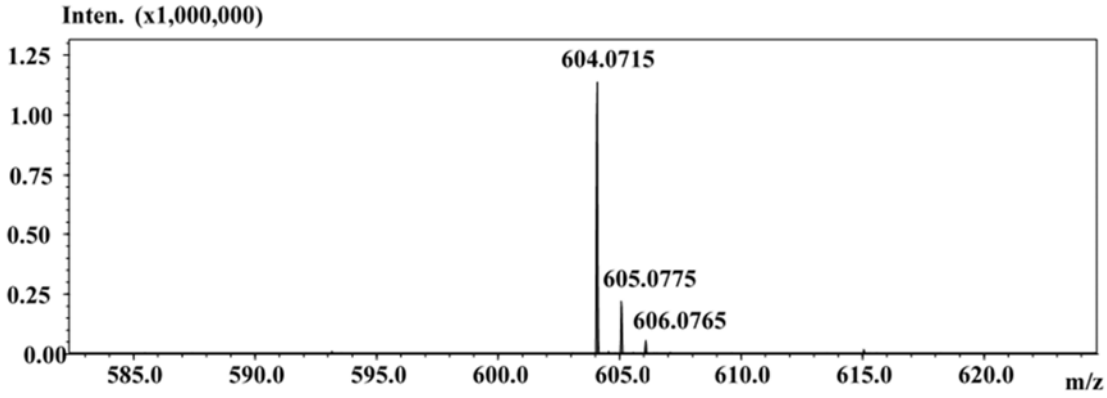
^{1}H-^{13}C HSQC spectrum of UDP-Man (600/150 MHz, D_2O)



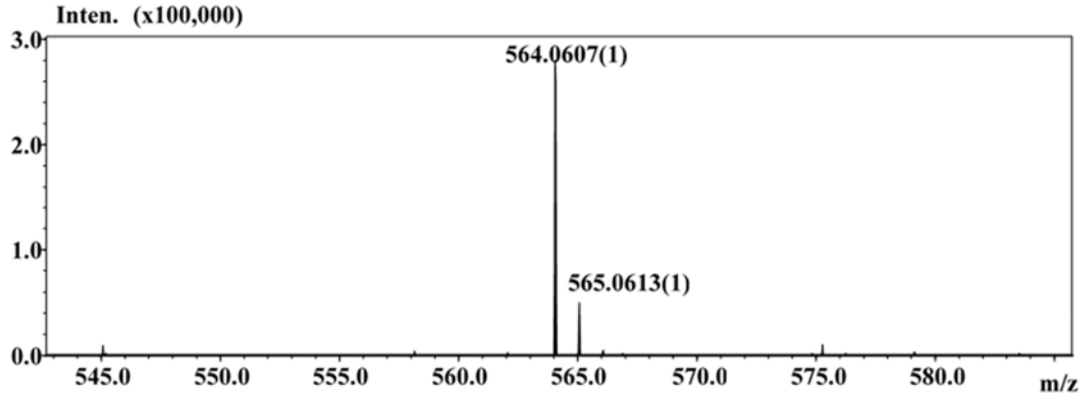
^{31}P NMR spectrum of UDP-Man (243 MHz, D_2O)



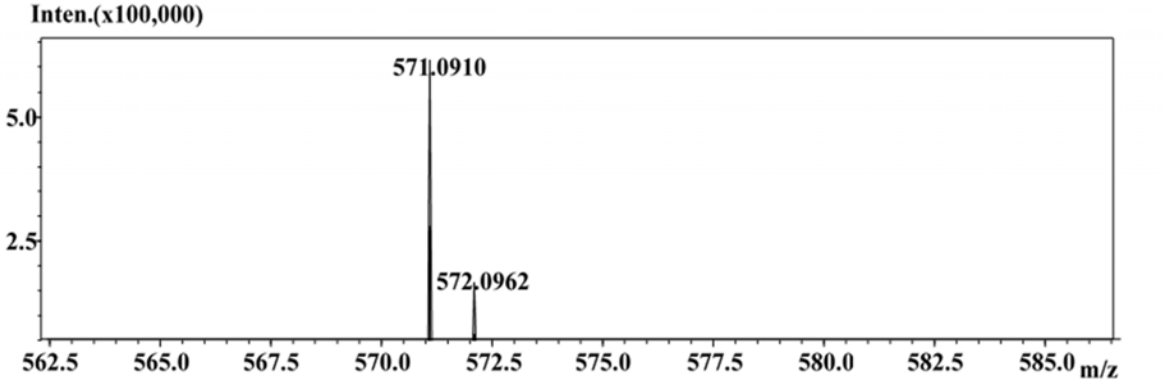
ESI-HRMS spectrum of UDP-Man



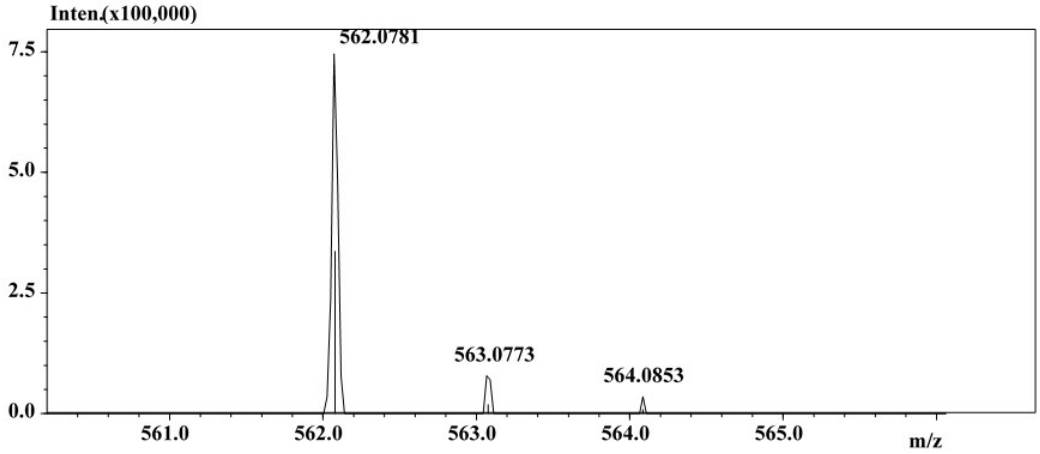
ESI-HRMS spectrum of GDP-Man



ESI-HRMS spectrum of UDP-GlcNH_2



ESI-HRMS spectrum of dADP-GlcNH_2



ESI-HRMS spectrum of dTDP-GalNH_2