ELECTRONIC SUPPLEMENTARY INFORMATION (ESI)

Enzymatic synthesis of *N*-acetyllactosamine from lactose enabled by recombinant β1,4- galactosyltransferases

Kun Huang,[a] Fabio Parmeggiani,[a] Helene Ledru,[b] Kristian Hollingsworth,[c] Jordi Mas Pons,[b] Andrea Marchesi,[a] Peter Both,[a] Ashley P. Mattey,[a] Edward Pallister,[a] Gregory S. Bulmer,[a] Jolanda M. van Munster,[a] W. Bruce Turnbull,[c] M. Carmen Galan,*[b] Sabine L. Flitsch*[a]

[a] Manchester Institute of Biotechnology (MIB), School of Chemistry, The University of Manchester, 131 Pincess Street, Manchester, M1 7DN, United Kingdom

[b] School of Chemistry, The University of Bristol, Cantock's Close, Bristol, BS8 1TS, United Kingdom

[c] School of Chemistry and Asbury Centre for Structural Molecular Biology, The University of Leeds, Woodhouse Lane, Leeds, LS2 9JT, United Kingdom

Experimental section

Materials and enzymes

UDP, lactose, *p*NP-β-GlcNAc, UDP-Gal, *p*NP-β-Glc, *p*NP-α-GlcNAc, β-N-acetyl-hexosaminidase, *p*NP-β-Lac and LacNAc were purchased either from Sigma-Aldrich or Carbosynth and used without further purification. β1,4-galactosyltransferases (NmLgtB-A, NmLgtB-B, NmLgtH) from *Neisseria meningitides* serogroup A, *Neisseria meningitides* serogroup B and *Neisseria meningitides* respectively were provided by Prozomix Lld, Haltwhistle, UK as purified protein suspended in $(NH_4)_2SO_4$ 3.2 M. The protein was centrifuged (20000*g*, 5 min, 4°C), the supernatant was discarded and an equal volume of Tris buffer (20 mM pH 8.0) was added to resuspend the protein pellet. The protein sample was centrifuged again and the supernatant containing enzyme was kept. A vivaspin centrifugal filter (10 kDa cut-off) was utilised to concentrate the enzyme. Protein quantification was carried out by the Pierce BCA protein assay kit (Thermo Fisher Scientific). The enzyme solution was aliquoted, snap-frozen in liquid nitrogen and stored at –80°C until use. SDS-PAGE was performed to determine protein purity (Figure S1).

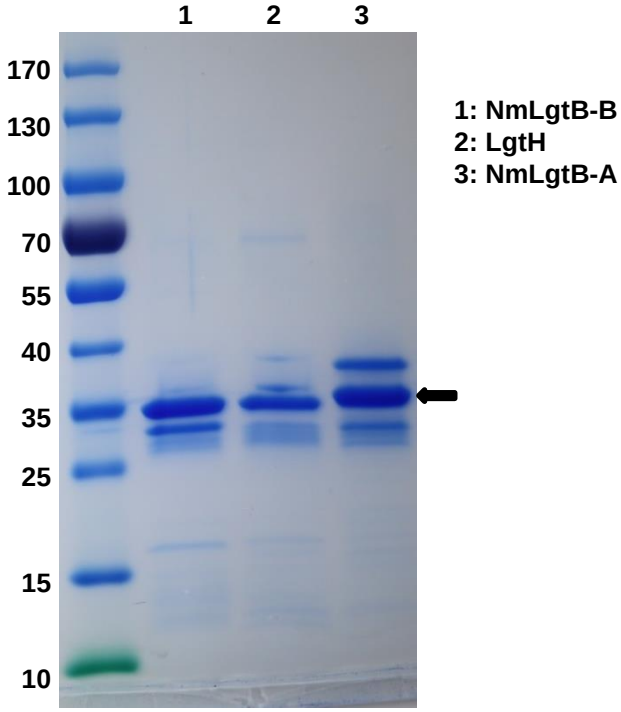


Figure S1. SDS-PAGE analysis of β1,4-galactosyltransferases used in this work. The purified proteins are indicated by a black arrow.

Amino acid sequences of recombinant β4GalTs

Neisseria meningitidis serogroup A strain Z2491 NmLgtB-A Uniprot No: P57033
MW: 34.1 kDa

(MGSSHHHHHHSSGLVPRGSH)MQNHVISLASAAERRAHITDTFGVRGIPFQFFDALMPSERLEQAMAELVPGLS
AHPYLSGVEKACFMSHAVLWKQALDEGLPYIAVFEDDVLLGEGAEKFLAEDAWLQERFDPDSAFIVRLETMFMHV
LTSPSGVADYCGRAFPLLESEHWGTAGYIISRKAMWFFLDRFAALPSEGLHPVDWMMFGNPDDRERMPVCQLNPA
LCAQELHYAKFHDQNSALGSLIEHDRCLNSKQQRRDSPANTFKHRLIRALTKISREREKRRQRREQLIGKIIVPF
Q

Neisseria meningitidis serogroup B strain MC58 NmLgtB-B Uniprot No: Q51116
MW: 33.7 kDa

(MGSSHHHHHHSSGLVPRGSH)MQNHVISLASAAERRAHIADTFGRHGIPFQFFDALMPSERLEQAMAELVPGLS
AHPYLSGVEKACFMSHAVLWKQALDEGLPYITVFEDDVLLGEGAEKFLAEDAWLQERFDPDTAFIVRLETMFMHV
LTSPSGVADYCGRAFPLLESEHWGTAGYIISRKAMRFFLDRFAALPPEGLHPVDLMMFSDFFDREGMPVCQLNPA
LCAQELHYAKFHDQNSALGSLIEHDRLLNRKQQRRDSPANTFKHRLIRALTKISREREKRRQRREQFIVPFQ

Neisseria meningitidis NmLgtH Uniprot No: Q2TIJ3
MW: 32.8 kDa

(MGSSHHHHHHSSGLVPRGSH)MQNHVISLASAAERRAHIADTFGRHGIPFQFFDALMPSERLEQAMAELVPGLS
AHPYLSSVEKACFMSHVVLWKQALDEGVPYVAVFEDDVLLGEGAEKFLDEDAWLQERFDKDSAFIVRLETMFMHV
LTSPSGVADYCGRAFPLLESEHWGMAGYIISRKAMRFFLDRFAVLPSERLKAVDWMLFSSFLDKGGMTVCQLTPA
LCIQSETLPSQLKNGRQESYRNRRSPKVLLKRALGKIGREIERARERKRQKKLEKHLGRHVVPFE

```
LgtB-A : MQNHVISLASAAERRAHITDTFGVRGIPFQFFDALMPSERLEQAMAELVPGLSAHP :  56
LgtB-B : MQNHVISLASAAERRAHIADTFGRHGIPFQFFDALMPSERLEQAMAELVPGLSAHP :  56
LgtH   : MQNHVISLASAAERRAHIADTFGRHGIPFQFFDALMPSERLEQAMAELVPGLSAHP :  56

LgtB-A : YLSGVEKACFMSHAVLWKQALDEGLPYIAVFEDDVLLGEGAEKFLAEDAWLQERFD : 112
LgtB-B : YLSGVEKACFMSHAVLWKQALDEGLPYITVFEDDVLLGEGAEKFLAEDAWLQERFD : 112
LgtH   : YLSSVEKACFMSHVVLWKQALDEGVPYVAVFEDDVLLGEGAEKFLDEDAWLQERFD : 112

LgtB-A : PDSAFIVRLETMFMHVLTSPSGVADYCGRAFPLLESEHWGTAGYIISRKAMWFFLD : 168
LgtB-B : PDTAFIVRLETMFMHVLTSPSGVADYCGRAFPLLESEHWGTAGYIISRKAMRFFLD : 168
LgtH   : KDSAFIVRLETMFMHVLTSPSGVADYCGRAFPLLESEHWGMAGYIISRKAMRFFLD : 168

LgtB-A : RFAALPSEGLHPVDWMMFGNPDDRERMPVCQLNPALCAQELHYAKFHDQNSALGSL : 224
LgtB-B : RFAALPPEGLHPVDLMMFSDFFDREGMPVCQLNPALCAQELHYAKFHDQNSALGSL : 224
LgtH   : RFAVLPSERLKAVDWMLFSSFLDKGGMTVCQLTPALCIQS----------ETLPSQ : 214

LgtB-A : IEHDRCLNSKQQRRDSPANTFKHRLIRALTKISREREKRRQRR-EQLIGKIIVPFQ : 279
LgtB-B : IEHDRLLNRKQQRRDSPANTFKHRLIRALTKISREREKRRQRR-EQFI----VPFQ : 275
LgtH   : LKNGRQESYRNRR--SPKVLLKRALGKIGREIERARERKRQKKLEKHLGRHVVPFE : 268
```

Figure S2. Amino acid sequence alignment of β4GalTs.

Enzymatic activity assay of β4GalTs

The assay of direct β4GalTs activity was carried out in a total volume of 20 µL containing Tris buffer (50 mM, pH 8.0), $MnCl_2$ (10 mM), $MgCl_2$ (10 mM), UDP-Gal (5 mM), GlcNAc-ITag **12** or Glc-ITag **11** (0.5 mM) and β4GalT (~2 µL). The mixtures were incubated at 37°C for 14 h and analysed by MALDI-ToF.

Transgalactosylation activity assay of β4GalTs using lactose as galactose donor

The assay of reverse β4GalTs activity was carried out in a total volume of 20 µL containing Tris buffer (50 mM, pH 8.0), $MnCl_2$ (10 mM), UDP (5 mM), Lac (10 mM), GlcNAc-ITag or Glc-ITag (0.5 mM) and β4GalT (~2 µL). The mixtures were incubated at 37°C for 14 h and analysed by MALDI-ToF.

Transgalactosylation activity assay of NmLgtB-B using *p*NP-β-Lac, LacNAc, *p*NP-α-Gal and *p*NP-β-Gal as galactose donor

In total 20 µl reaction mixture containing Glc-ITag or GlcNAc-ITag (0.25 mM), galactose donor (*p*NP-β-Lac, LacNAc, *p*NP-α-Gal and *p*NP-β-Gal, 2.5 mM), UDP (5 mM), sodium acetate buffer (50 mM, pH5), $MnCl_2$ (10 mM) and NmLgtB-B (18.4 µg) was incubated at 37°C for overnight. Product was analysed by MALDI-ToF.

Representative MALDI traces

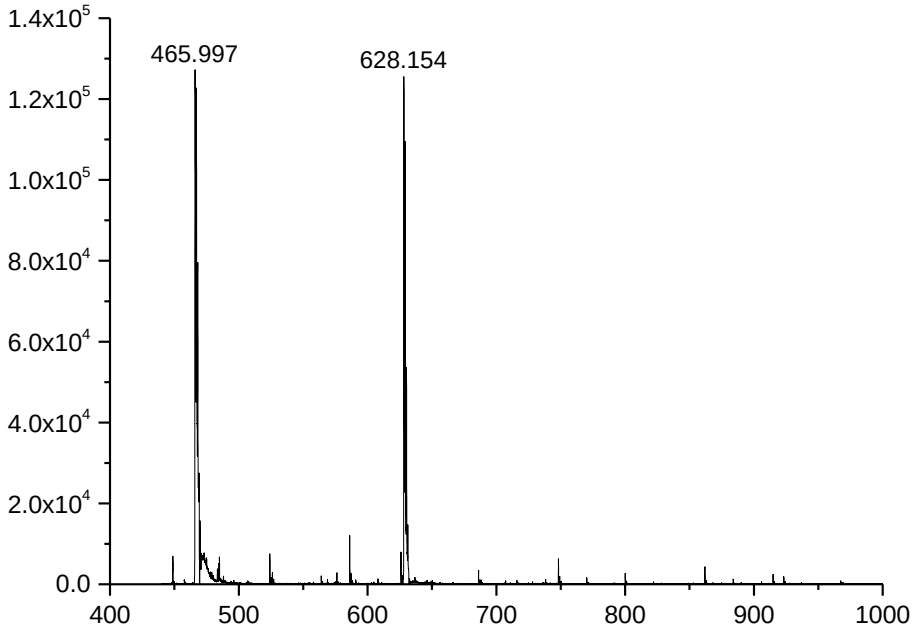


Figure S3. Transgalactosylation activity of NmLgtB-B against Glc-ITag **11**. A peak of m/z 628.154 was observed, corresponding to the expected Lac-ITag **13** $[M]^+$ 628.271.

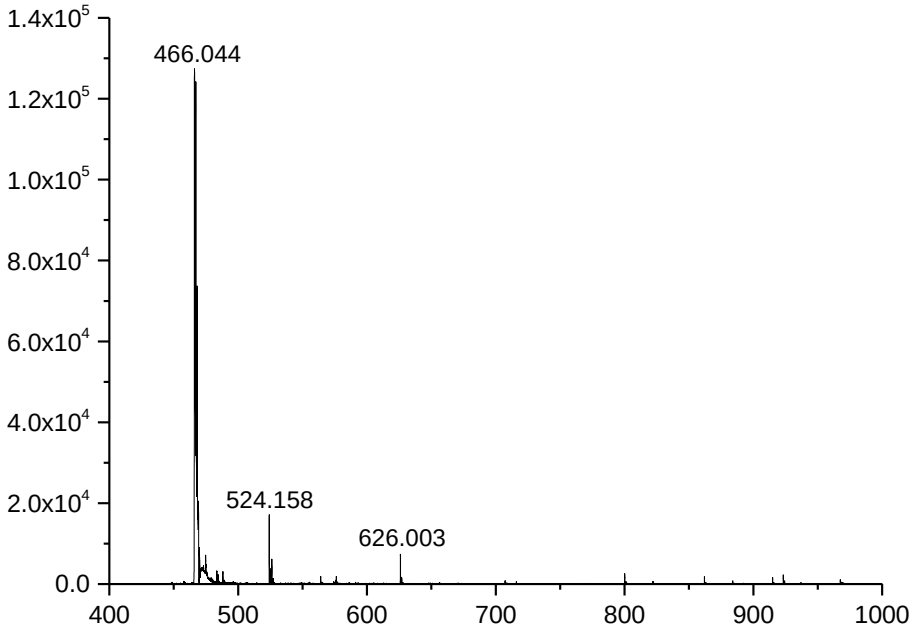


Figure S4. Transgalactosylation activity of NmLgtH against Glc-ITag **11**.

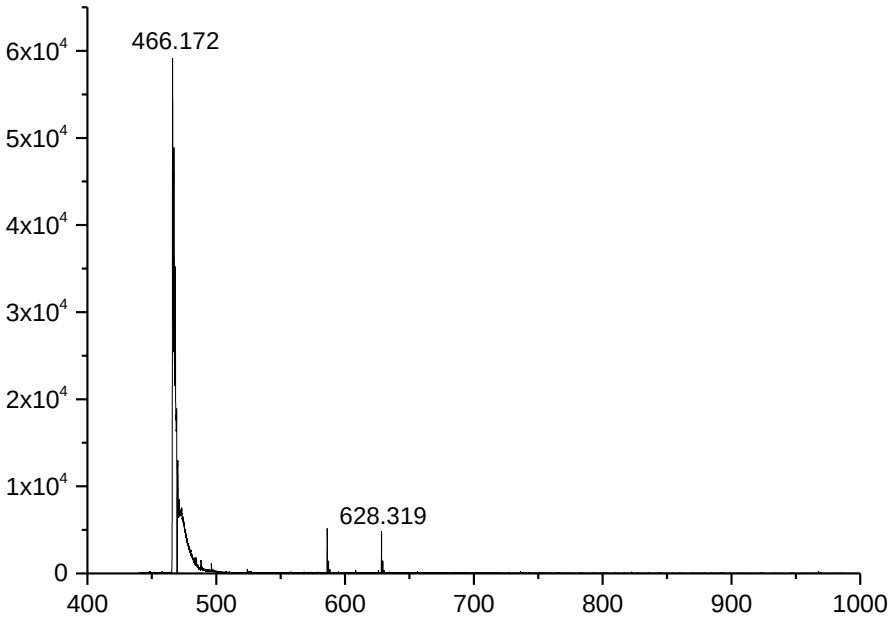


Figure S5. Transgalactosylation activity of NmLgtB-A against Glc-ITag **11**. A peak of m/z 628.319 was observed, corresponding to the expected Lac-ITag **13** $[M]^+$ 628.271.

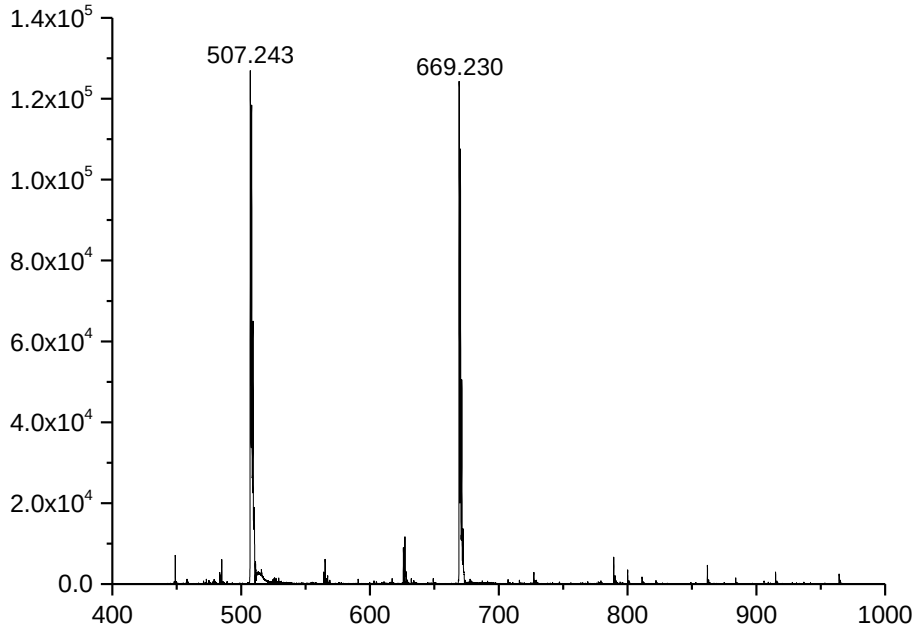


Figure S6. Transgalactosylation activity of NmLgtB-B against GlcNAc-ITag **12**. A peak of m/z 669.230 was observed, corresponding to the expected LacNAc-ITag **14** $[M]^+$ 669.298.

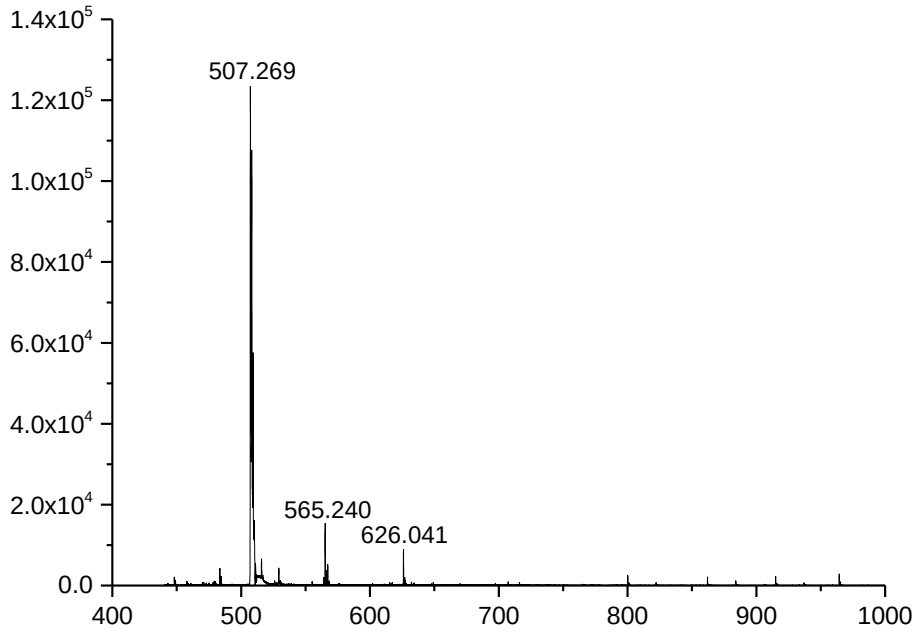


Figure S7. Transgalactosylation activity of NmLgtH against GlcNAc-ITag **12**.

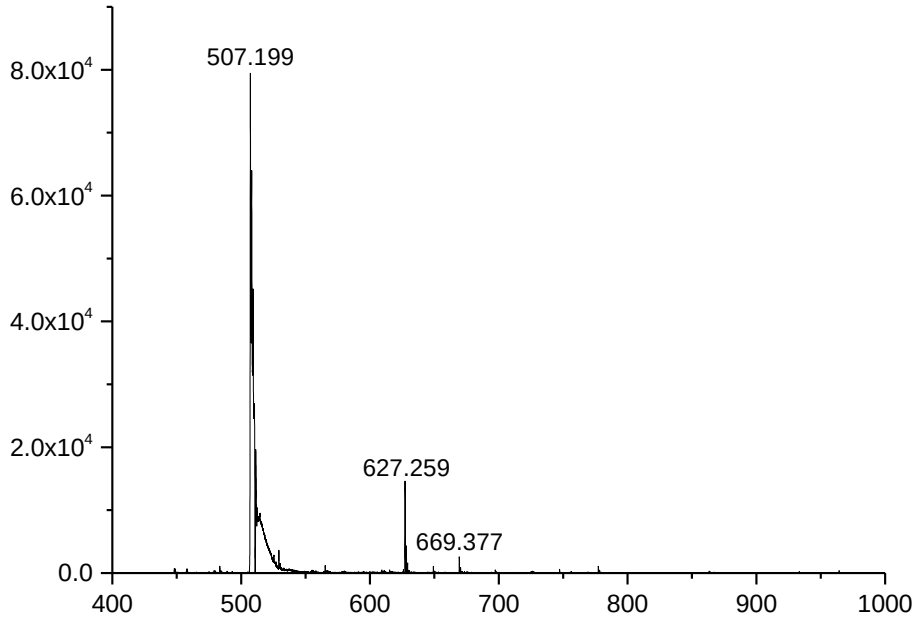


Figure S8. Transgalactosylation activity of NmLgtB-A against GlcNAc-ITag **12**. A peak of m/z 669.377 was observed, corresponding to the expected LacNAc-ITag **14** $[M]^+$ 669.298.

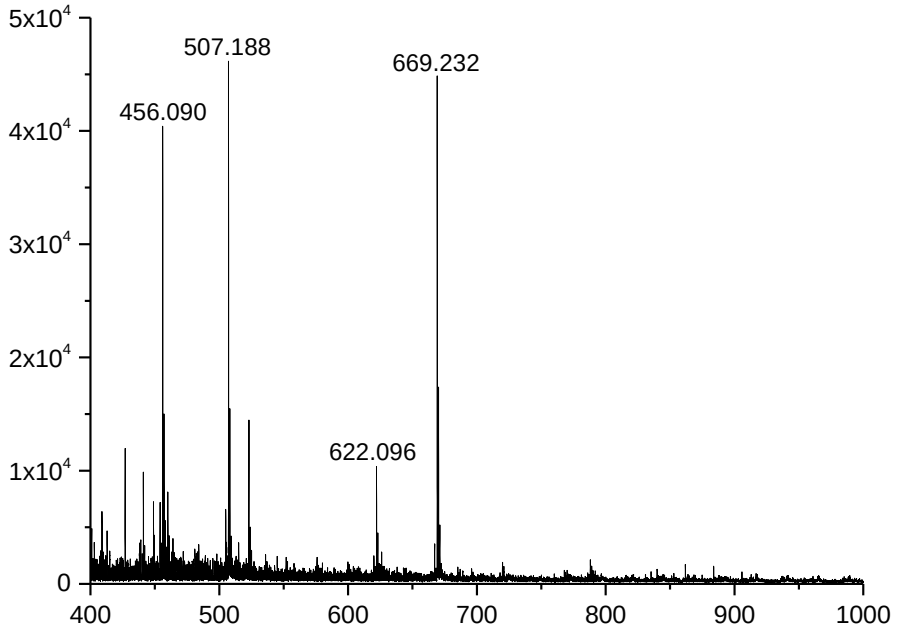


Figure S9. Transgalactosylation activity of NmLgtB-B against GlcNAc-ITag **12** using *p*NP-β-Lac as galactose donor. A peak of m/z 669.232 was observed, corresponding to the expected LacNAc-ITag **14** $[M]^+$ 669.298.

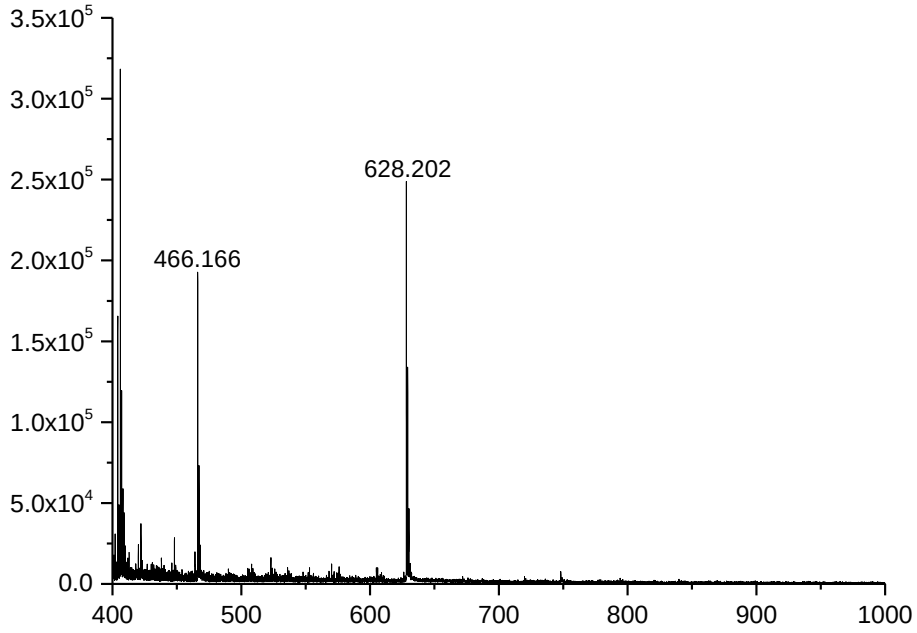


Figure S10. Transgalactosylation activity of NmLgtB-B against Glc-ITag **11** using LacNAc as galactose donor. A peak of m/z 628.202 was observed, corresponding to the expected Lac-ITag **13** $[M]^+$ 628.271.

Transgalactosylation reaction condition optimisation

Influence of pH. The reaction was performed in a total volume of 20 µL containing $MnCl_2$ (10 mM), UDP (5 mM), *p*NP-β-GlcNAc (1 mM), Lac (10 mM), NmLgtB-B (17.6 µg) in different buffers (50 mM, sodium acetate 4.0-5.5, MES 6.0, MOPS 7.0-7.5, Tris 8.0-9.0).
Influence of UDP. The reaction was performed in a total volume of 20 µL containing sodium acetate buffer (50 mM, pH 5.0), $MnCl_2$ (10 mM), *p*NP-β-GlcNAc (1 mM), Lac (10 mM), NmLgtB-B (17.6 µg) and varying concentrations of UDP (0-6 mM).
Influence of Lac. The reaction was performed in a total volume of 20 µL containing sodium acetate buffer (50 mM, pH 5.0), $MnCl_2$ (10 mM), *p*NP-β-GlcNAc (1 mM), UDP (2 mM), NmLgtB-B (17.6 µg) and varying concentrations of Lac (1-30 mM).
Analysis. All assays were performed in triplicate. The reactions were incubated at 37°C for 14 h, then quenched by addition of MeCN (20 µL). Products were analysed by HPLC and the conversion was calculated by integration of the *p*NP-LacNAc and *p*NP-GlcNAc peak areas (UV detection at 300 nm).

Direct detection of UDP-Gal during the transgalactosylation with NmLgtB-B

The reaction was performed in a total volume of 20 µL containing sodium acetate buffer (50 mM, pH 5.0), $MnCl_2$ (10 mM), UDP (5 mM), Lac (10 mM), NmLgtB-B (17.6 µg). The mixture was incubated at 37°C for 14 h, then quenched by addition of MeCN (20 µL). The formation of UDP-Gal was verified by HRMS (Figure S11).

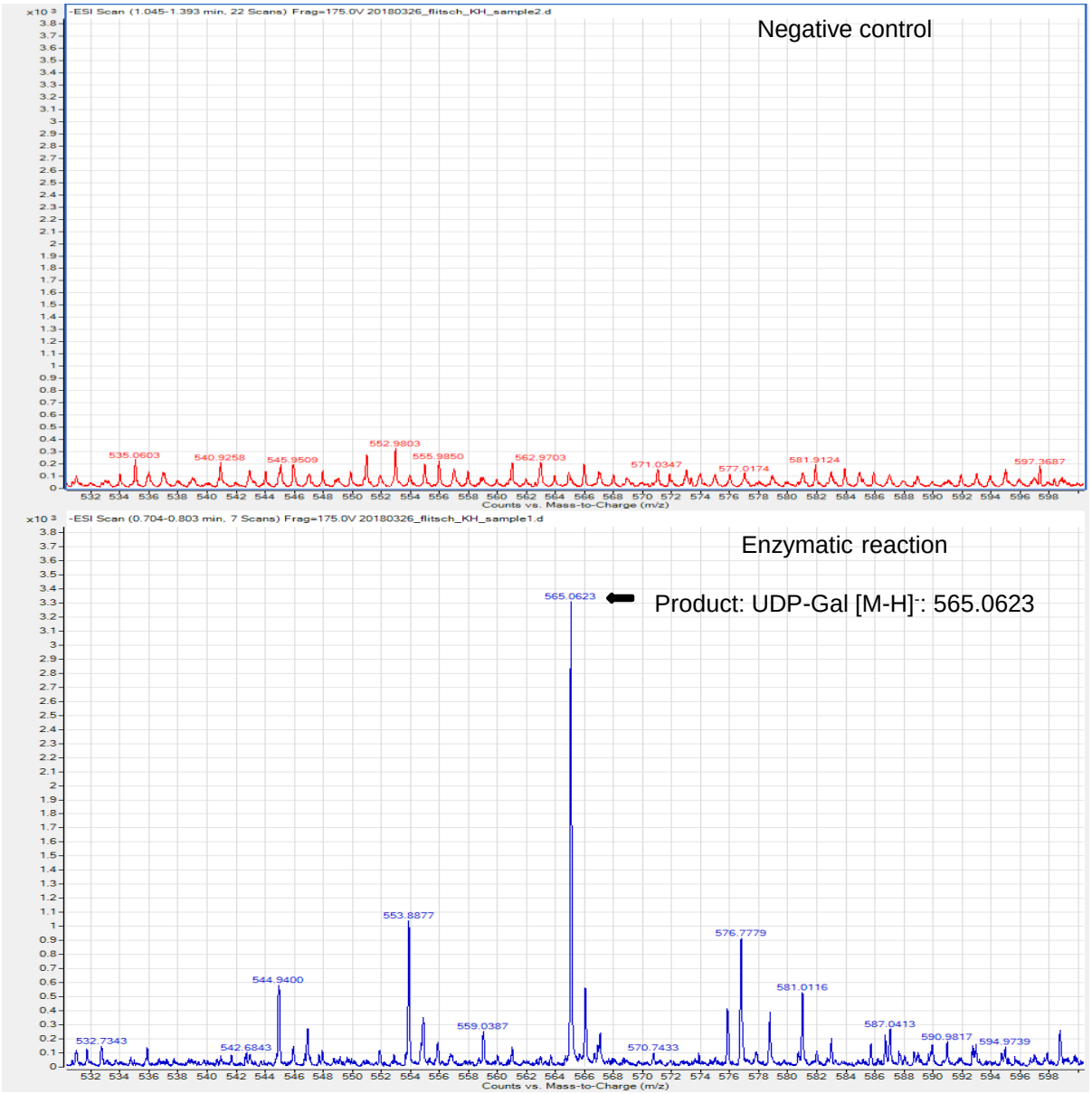


Figure S11. HRMS analysis showing UDP-Gal formation by NmLgtB-B. A peak of m/z 565.0623 was observed in the enzymatic reaction mixture, corresponding to the expected UDP-Gal $[M-H]^-$ 565.0550.

Time course reaction for *p*NP-β-LacNAc formation

The reaction was executed at 37°C in a total volume of 20 µL containing $MnCl_2$ (10 mM), UDP (2 mM), *p*NP-β-GlcNAc (1 mM), Lac (20 mM), NmLgtB-B (3.73 µg) in sodium acetate buffer (50mM, pH 5). The reaction was quenched by adding equal volume of 100% methanol after 1, 2, 3, 4, 5, 6, 7 and 9 hours.

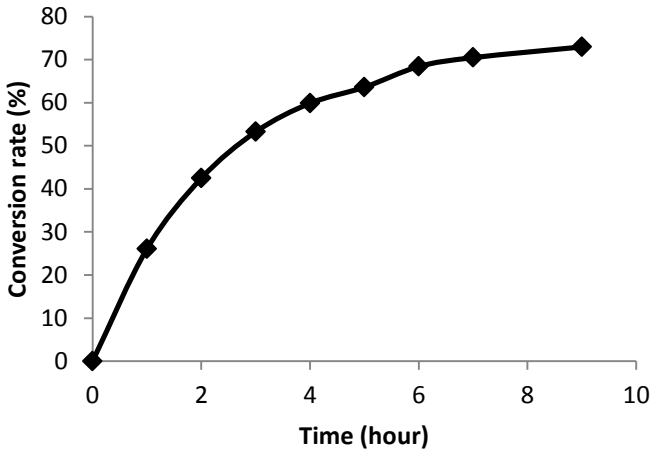


Figure S12. Time course reaction for *p*NP-β-LacNAc formation.

Analytical methods

MALDI-ToF. Crude reaction mixtures (0.5 µL) were spotted on target plates, then mixed with THAP matrix (0.5 µL, 10 mg/mL solution in acetone) and dried in ambient temperature. The product was analysed in positive mode on a Bruker Ultraflex MALDI-ToF instrument.

HRMS. HRMS analysis were performed using an Agilent 1200 series LC system, coupled to an Agilent 6520 QTOF mass spectrometer, ESI positive mode. The sample (2 µL) was flow-injected into 0.3 mL min^{-1} MeCN/H_2O 1:1 + formic acid 0.1% v/v. The data was analyzed using Agilent MassHunter software.

HPLC. Reverse-phase HPLC analysis was performed on an Agilent 1200 series LC system equipped with a Luna C18 250 × 2 mm reverse phase column, according to the following method. Mobile phase A: 50 mM ammonium formate in water, pH 4.5. Mobile phase B: acetonitrile. Flow rate: 0.6 mL/min. Gradient: 0-5 min isocratic 5% B, 5-15 min linear gradient 5-50% B, 15-25 min isocratic 5% B.

Preparative scale synthesis of *p*NP-β-LacNAc 16

The preparative scale synthesis of *p*NP-β-LacNAc was carried out in a total volume of 5 mL containing sodium acetate buffer (50 mM, pH 5.0), $MnCl_2$ (10 mM), *p*NP-β-GlcNAc (17.5 mg), UDP (50 mg), Lac (345 mg), NmLgtB-B (1.5 mg). The mixture was incubated at 37°C for 14 h and monitored by HPLC. When the conversion reached its maximum (Figure S13), β-N-acetylhexosaminidase 2 mg in total was added to the reaction mixture to hydrolyse the unreacted starting material *p*NP-β-GlcNAc at 37°C overnight. The enzymes were removed by ultrafiltration (10 kDa Vivaspin centrifugal filter). The filtrate was lyophilized and re-solubilised in distilled water. The product was purified by preparative HPLC on a Luna C18 (2), 100Å, 250×15 mm column with UV detection at 300 nm. Mobile phase A: 50 mM ammonium formate pH 4.5. Mobile phase B: acetonitrile. Method: flow rate 10 mL· min^{-1}, 0-5 min isocratic 5% B, 5-15 min linear gradient from 5 to 50% B, 15-20 min isocratic 5% B. The product fractions were pooled, lyophilized and the product (12.9 mg, 51% yield) was characterised by NMR and HRMS.

^{1}H NMR (600 MHz, D_2O) δ 8.23 – 8.20 (m, 2H, Ar*H*), 7.18 – 7.15 (m, 2H, Ar*H*), 5.33 (d, $J_{1,2}$ = 8.4 Hz, 1H, H-1), 4.50 (d, $J_{1',2'}$ = 7.8 Hz, 1H, H-1'), 4.10 – 4.05 (m, 1H, H-2), 4.03 – 3.93 (m, 1H, H-6a), 3.93 – 3.91 (m, 1H, H-4'), 3.90 – 3.82 (m, 4H, H-3, H-4, H-5, H-6b), 3.81 – 3.72 (m, 3H, H-5', H-6'ab), 3.67 (dd, $J_{2',3'}$ = 10.0 Hz, $J_{3',4'}$ = 3.3 Hz, 1H, H-3'), 3.55 (dd, $J_{1',2'}$ = 7.8 Hz, $J_{2',3'}$ = 10.0 Hz, 1H, H-2'), 2.01 (s, 3H, C(O)C*H*$_3$); **^{13}C-NMR** (151 MHz, D_2O) δ 174.9 (*C*(O)CH_3), 161.6 (Cq, Ar), 142.6 (Cq, Ar), 126.0 (*C*H, Ar), 116.4 (*C*H, Ar), 102.9 (C-1'), 98.4 (C-1), 77.9 (C-4), 75.3 (C-5'), 75.1 (C-5), 72.5 (C-3'), 72.0 (C-3), 71.0 (C-2'), 68.5 (C-4'), 61.0 (C-6'), 59.7 (C-6), 54.8 (C-2), 22.1 (C(O)*C*H_3); **HRMS (ESI)** m/z for $C_{20}H_{28}N_2O_{13}$: $[M+H]^+$ calcd. 505.1664, found 505.1668.

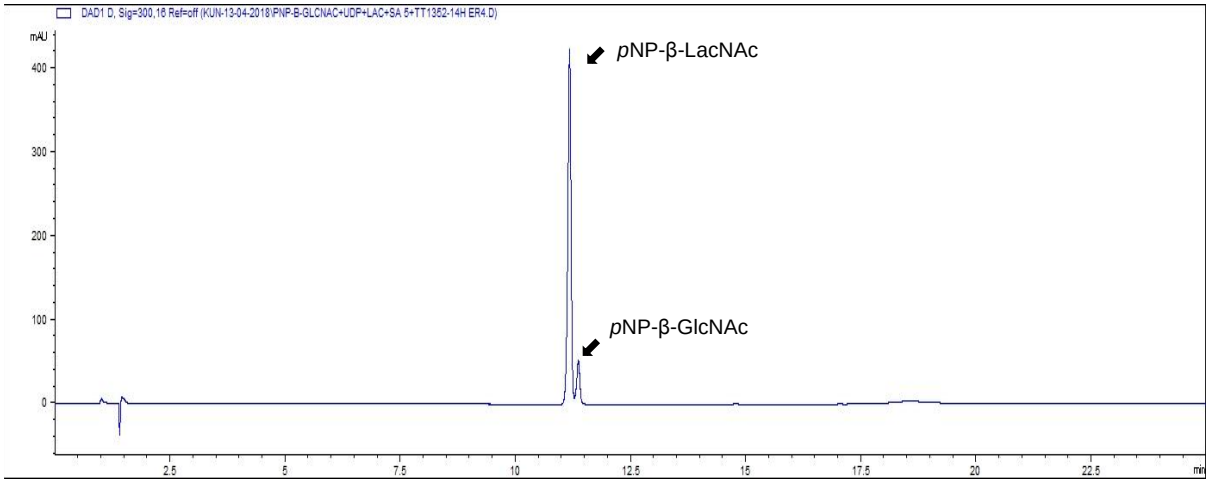
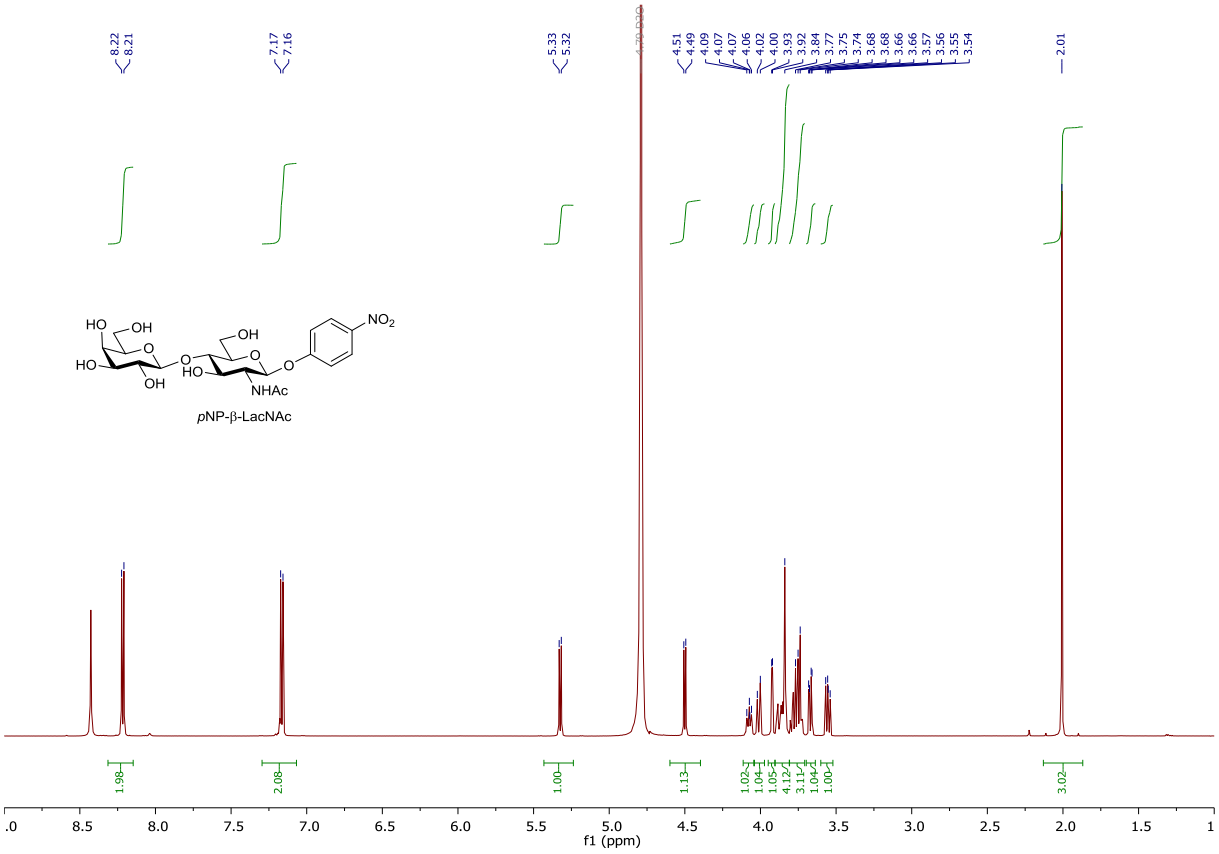
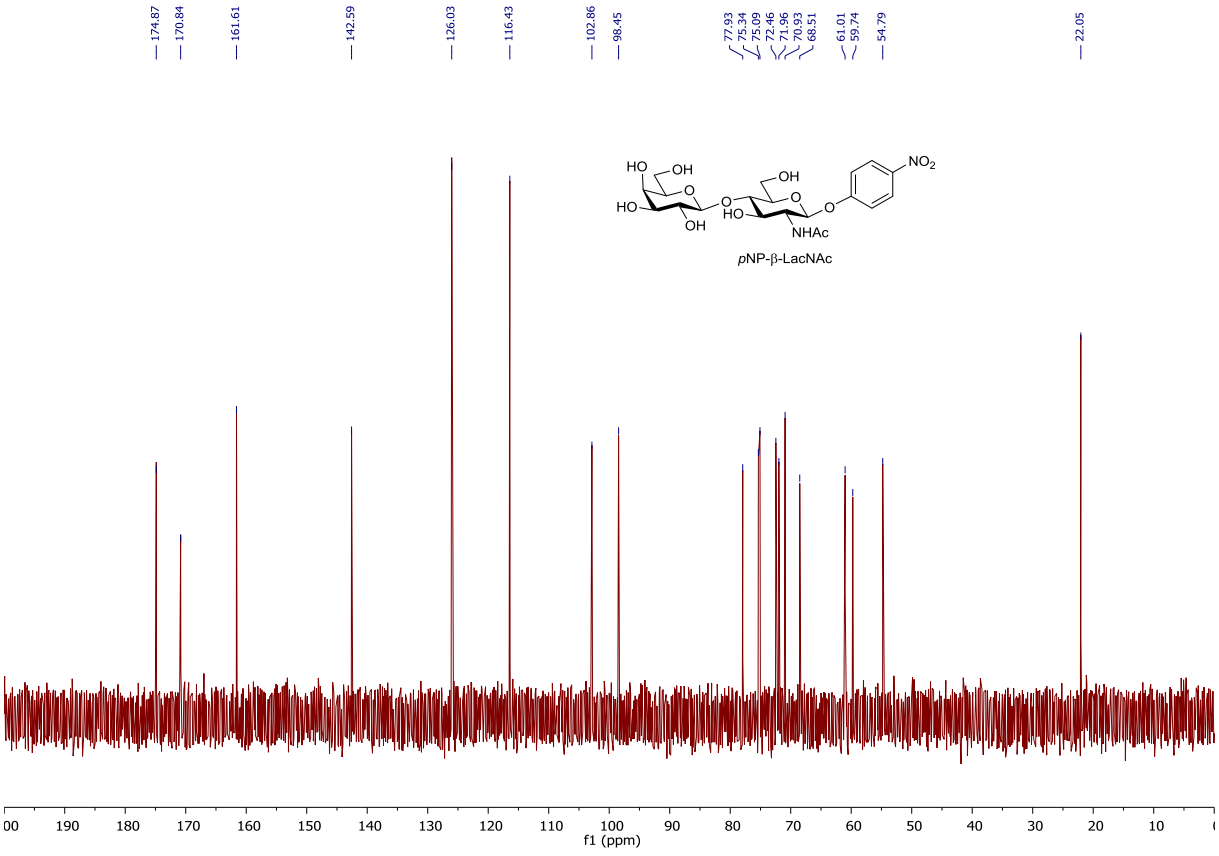
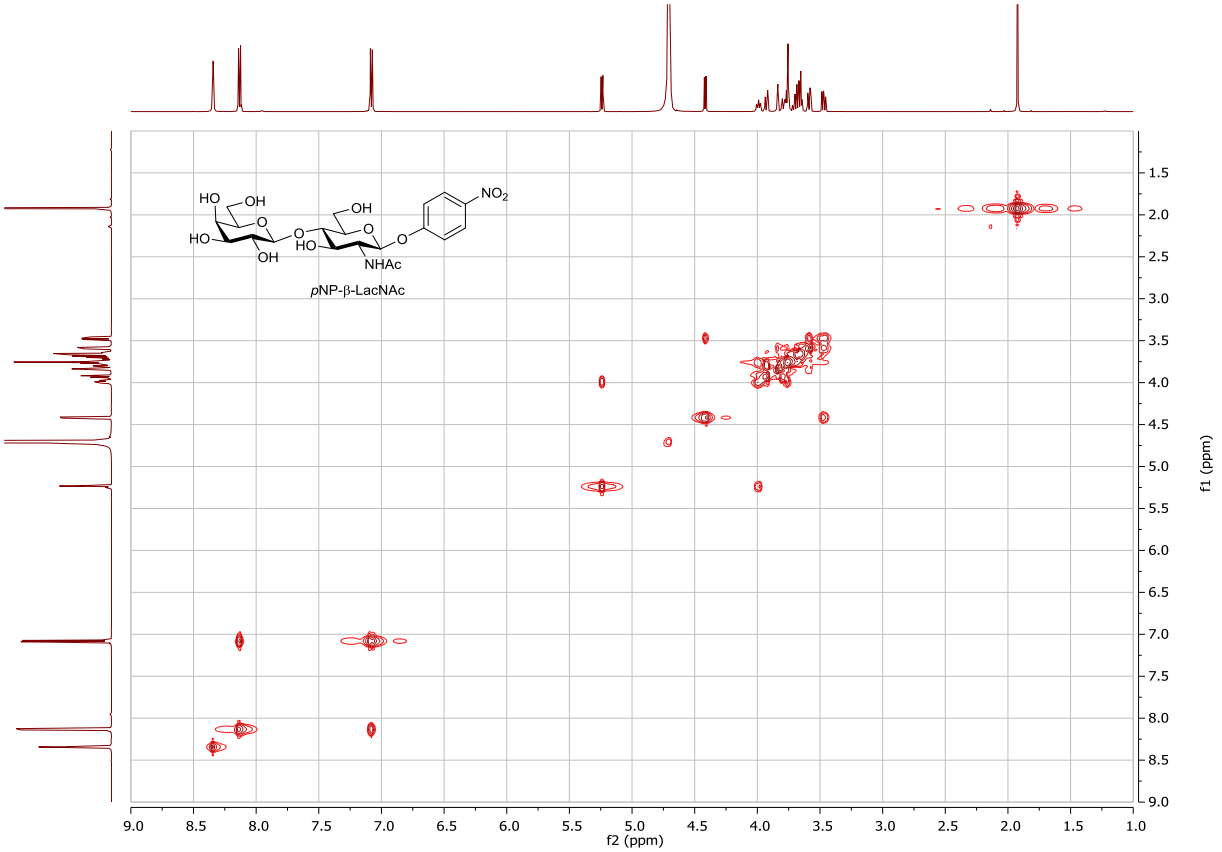


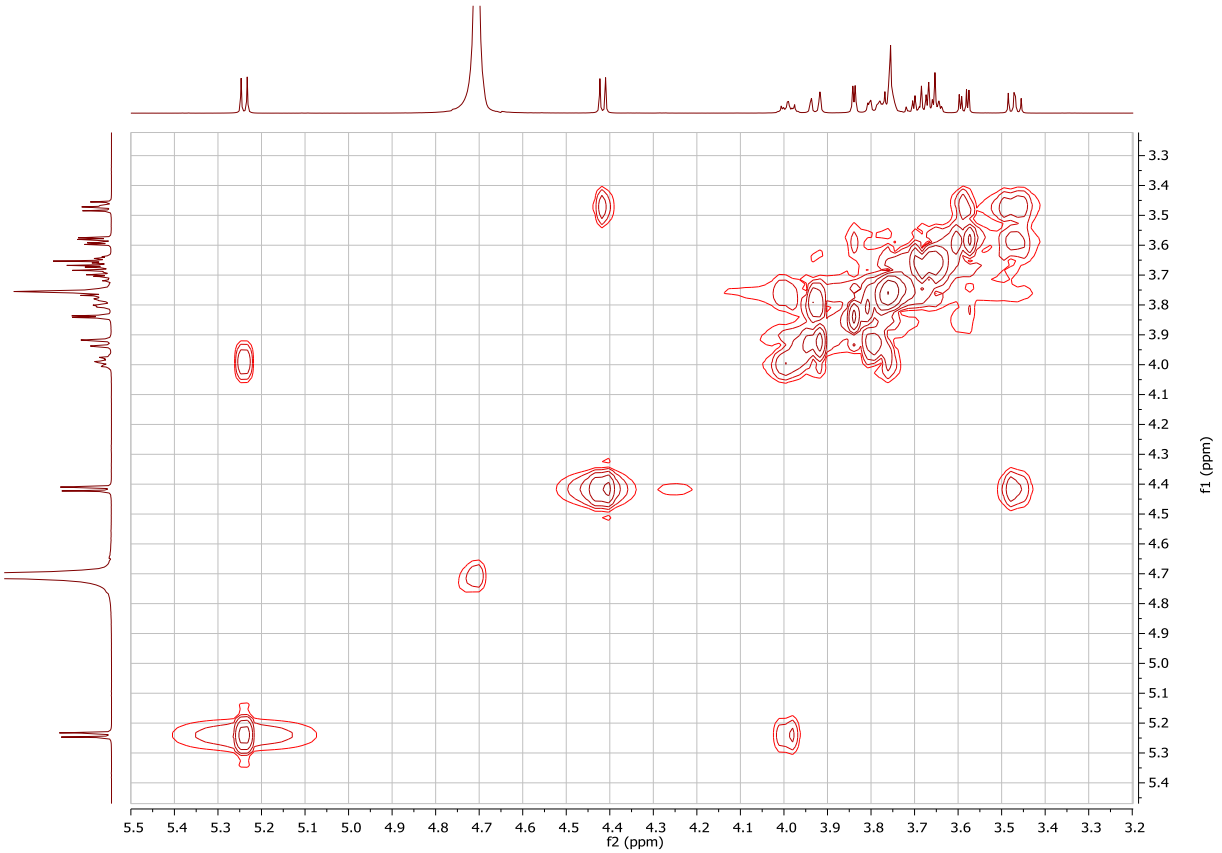
Figure S13. HPLC analysis of preparative synthesis of *p*NP-β-LacNAc **16**.

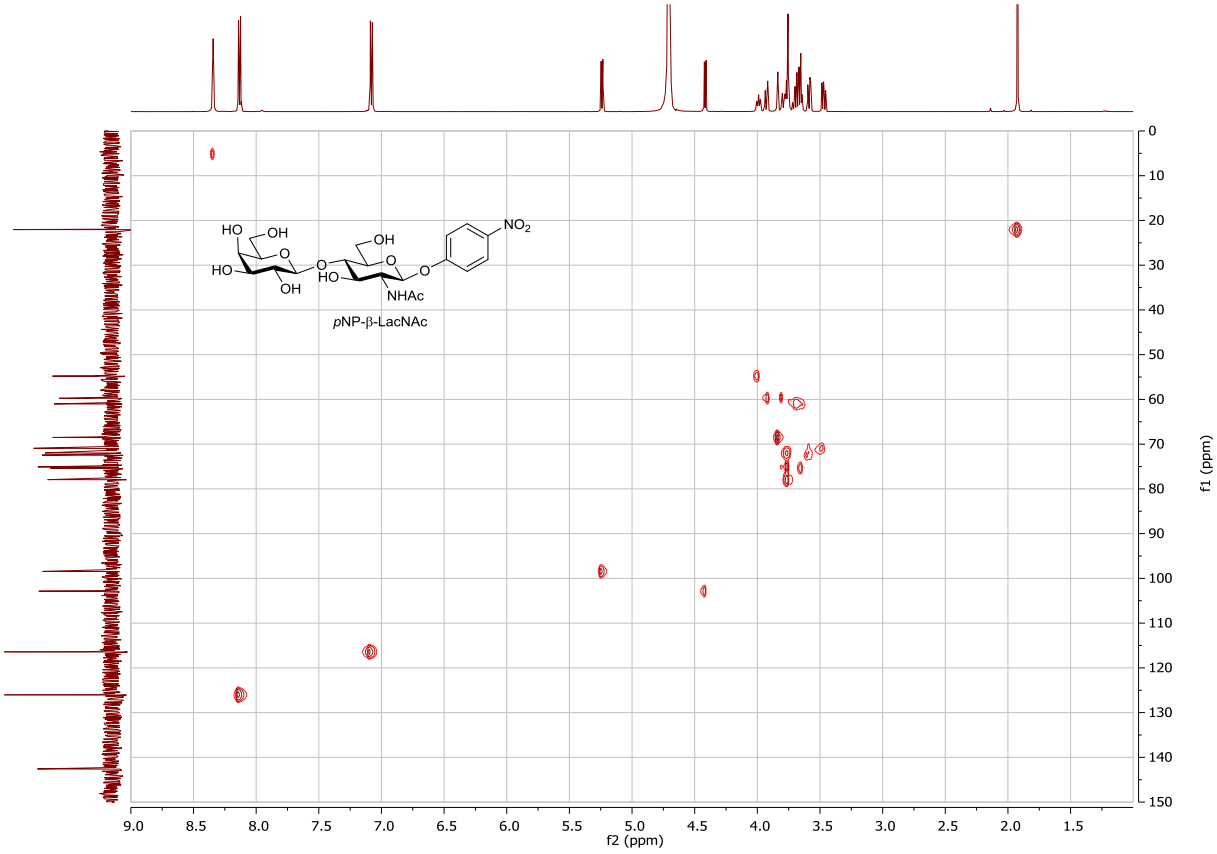
NMR and HRMS spectra for the preparative synthesis of *p*NP-β-LacNAc 16

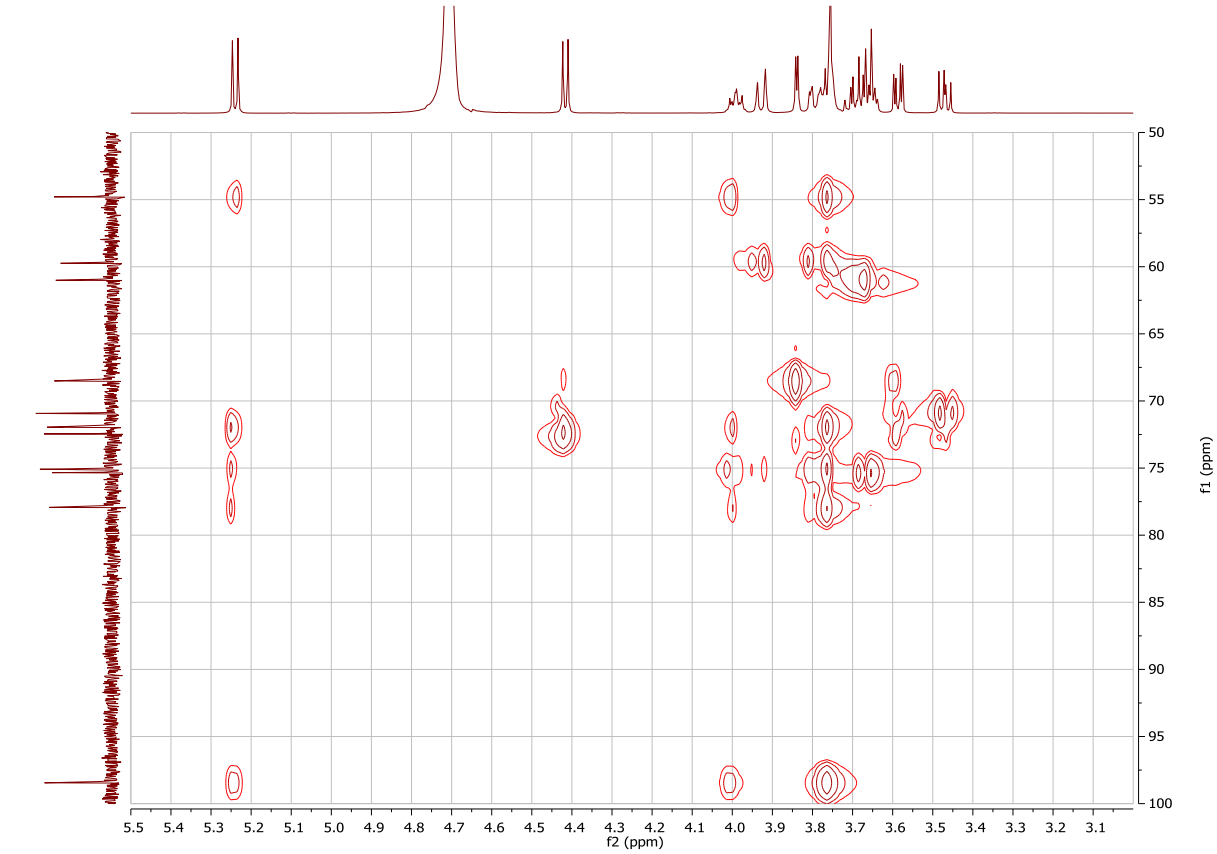





pNP-β-LacNAc
f2 (ppm)
f1 (ppm)


f2 (ppm)
f1 (ppm)


HO
OH
O
HO
OH
O
HO
OH
O
O
NHAc
NO2
pNP-β-LacNAc
f2 (ppm)
f1 (ppm)


f2 (ppm)
f1 (ppm)

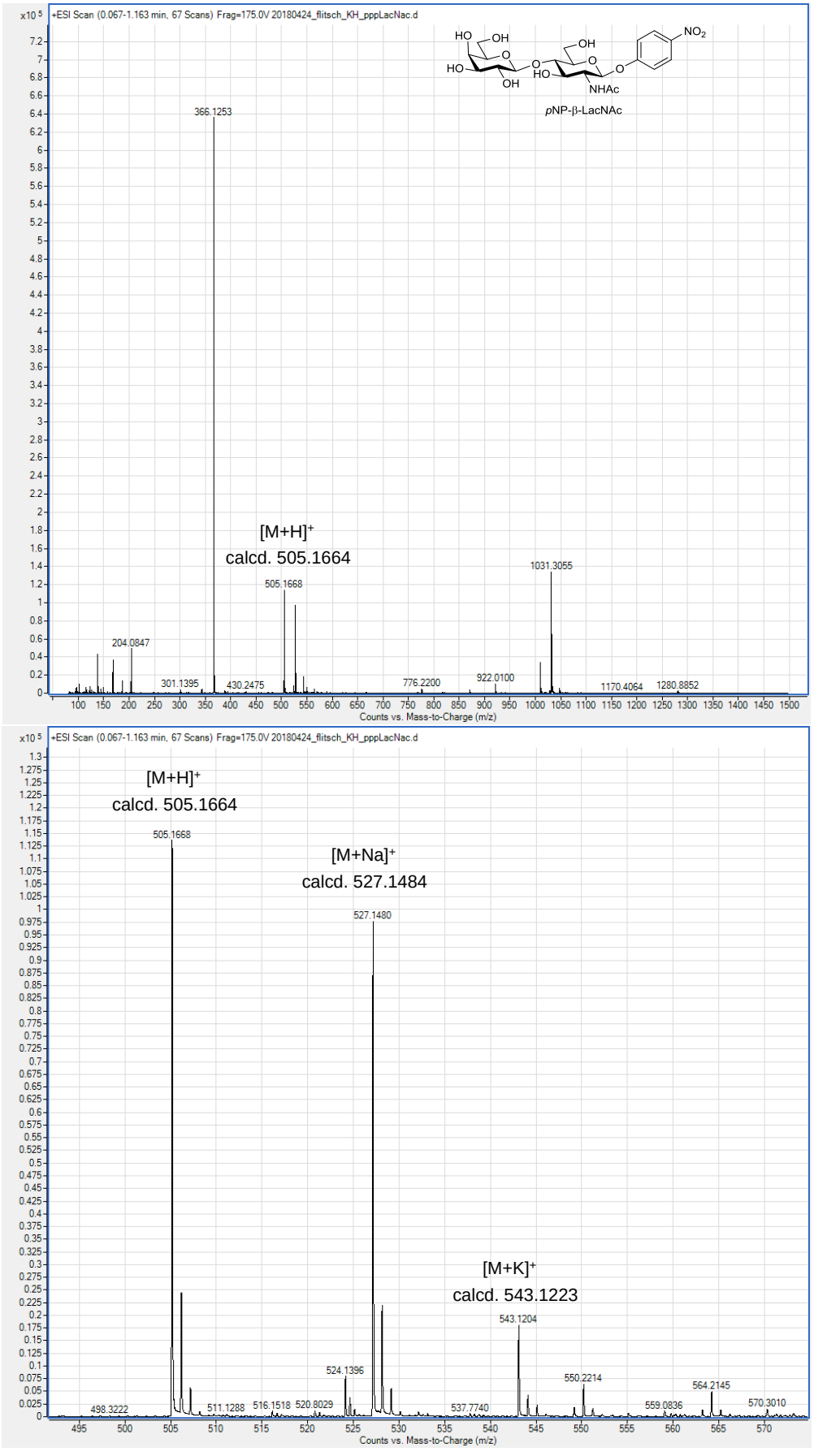
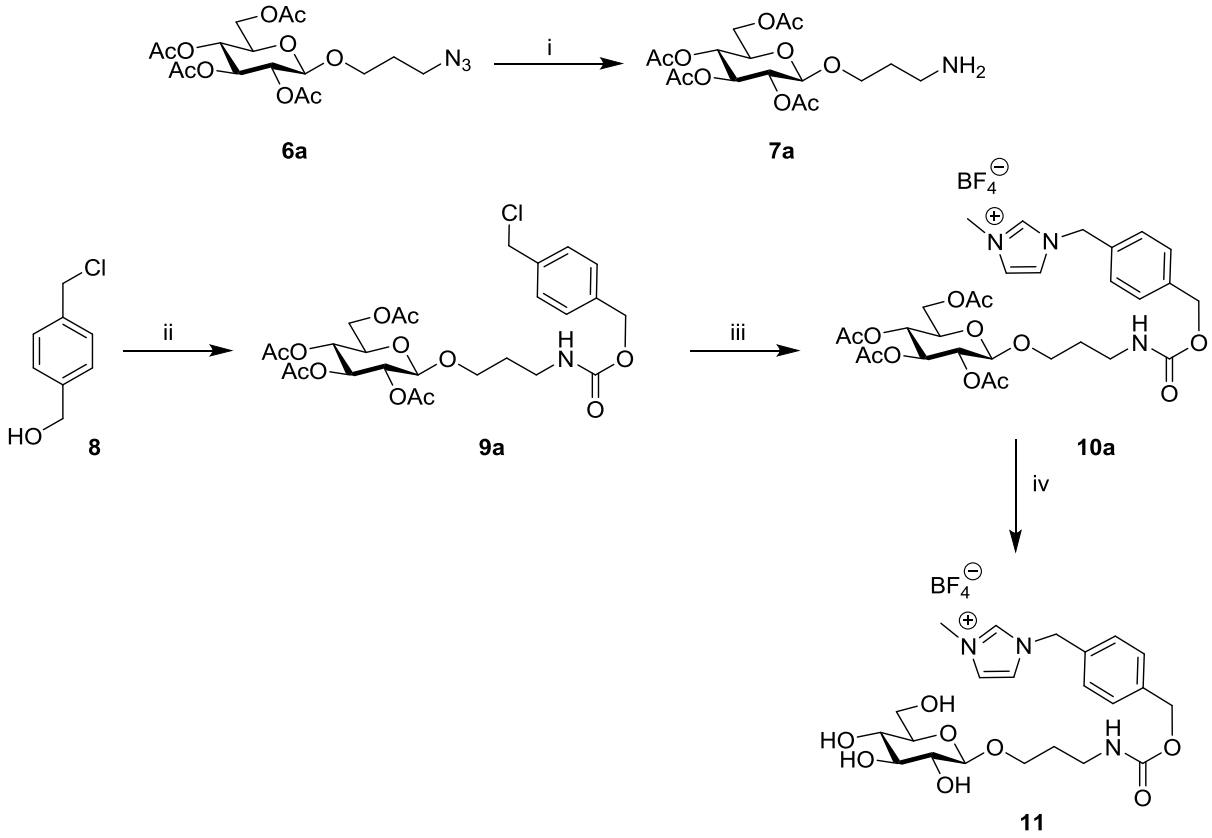


Figure S14. HRMS for purified *p*NP-β-LacNAc

Chemical synthesis of Glc-ITag 11



Scheme S1. Reagents and conditions for the synthesis of **11**: i) H_2, Pd/C, THF, rt, 1 h; ii) CDI, THF, rt, 30 min; then Glc-O$(CH_2)_3$-NH_2 (**7a**), 15 h (44% over 2 steps); iii) 1-methylimidazole, KBF_4, CH_3CN, 80 °C, 16 h (54%); iv) Et_3N, MeOH, 16 h (78%).

3-Azidopropyl 2,3,4,6-tetra-*O*-acetyl-*β*-D-glucopyranoside (6a)

To a stirring solution of 3-bromopropyl 2,3,4,6-tetra-*O*-acetyl-*β*-D-glucopyranoside (5.14 g, 10.95 mmol) in anhydrous DMF (96 mL) at rt, was added NaN_3 (2.14 g, 32.86 mmol) and the mixture stirred for 16 h at 80 ºC. The solvent was then evaporated and the resulting residue diluted with H2O (50 mL). Extraction with DCM (3 x 50 mL) and drying over anhydrous $MgSO_4$. The filtrates were concentrated under reduced pressure and the dried residue was subsequently purified by column chromatography (Toluene:EtOAc 8:2) to give **6a** (4.20 g, 75 %) as a pale yellow solid.

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.21 (dd, *J* = 9.7, 9.2 Hz, 1H, H-3), 5.08 (dd, *J* = 10.0, 9.2 Hz, 1H, H-4), 4.99 (dd, *J* = 9.7, 8.0 Hz, 1H, H-2), 4.51 (d, *J* = 8.0 Hz, 1H, H-1), 4.26 (dd, *J* = 12.3, 4.7 Hz, 1H, H-6a), 4.15 (dd, *J* = 12.3, 2.6 Hz, 1H, H-6b), 3.95 (dt, *J* = 9.8, 5.6 Hz, 1H, OC*H*HCH_2), 3.70 (*app*. ddd, *J* = 10.0, 4.7, 2.6 Hz, 1H, H-5), 3.61 (*app*. ddd, *J* = 9.8, 7.7, 5.0 Hz, 1H, OCH*H*CH_2), 3.37 (td, *J* = 6.2, 2.3 Hz, 2H, $CH_2CH_2N_3$), 2.09 (s, 3H, C*H*3, Ac), 2.05 (s, 3H, C*H*3, Ac), 2.03 (s, 3H, CH_3, Ac), 2.01 (s, 3H, CH_3, Ac), 1.91 – 1.79 (m, 2H, OCH_2CH_2); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 170.8 (*C*O, Ac), 170.4 (*C*O, Ac), 169.5 (*C*O, Ac), 169.4 (*C*O, Ac), 101.0 (C-1), 72.9 (C-3), 72.0 (C-5), 71.4 (C-2), 68.6 (C-4), 66.6 (OCH_2CH_2), 62.1 (C-6), 48.1 ($CH_2CH_2N_3$), 29.1 (OCH_2CH_2), 20.9 (CH_3, Ac), 20.8 (CH_3, Ac), 20.8 (CH_3, Ac), 20.7 (CH_3, Ac). The spectroscopic data is in accordance with the data reported in the literature.[1]

3-(4-(Chloromethyl)benzyl carbamate) propyl 2,3,4,6-tetra-*O*-acetyl-*β*-D-glucopyranoside (9a)

To a stirring solution of **6a** (1.00 g, 2.32 mmol) in THF (23 mL) at rt, Pd/C (120 mg, 1.16 mmol) was added. The mixture was stirred under a H_2 atmosphere for 1 h, upon which time the mixture was filtered through celite and washed with EtOAc. The filtrates were concentrated under reduced pressure to give **7a** (940 mg, quantitative) as a black residue, which was taken to the next step

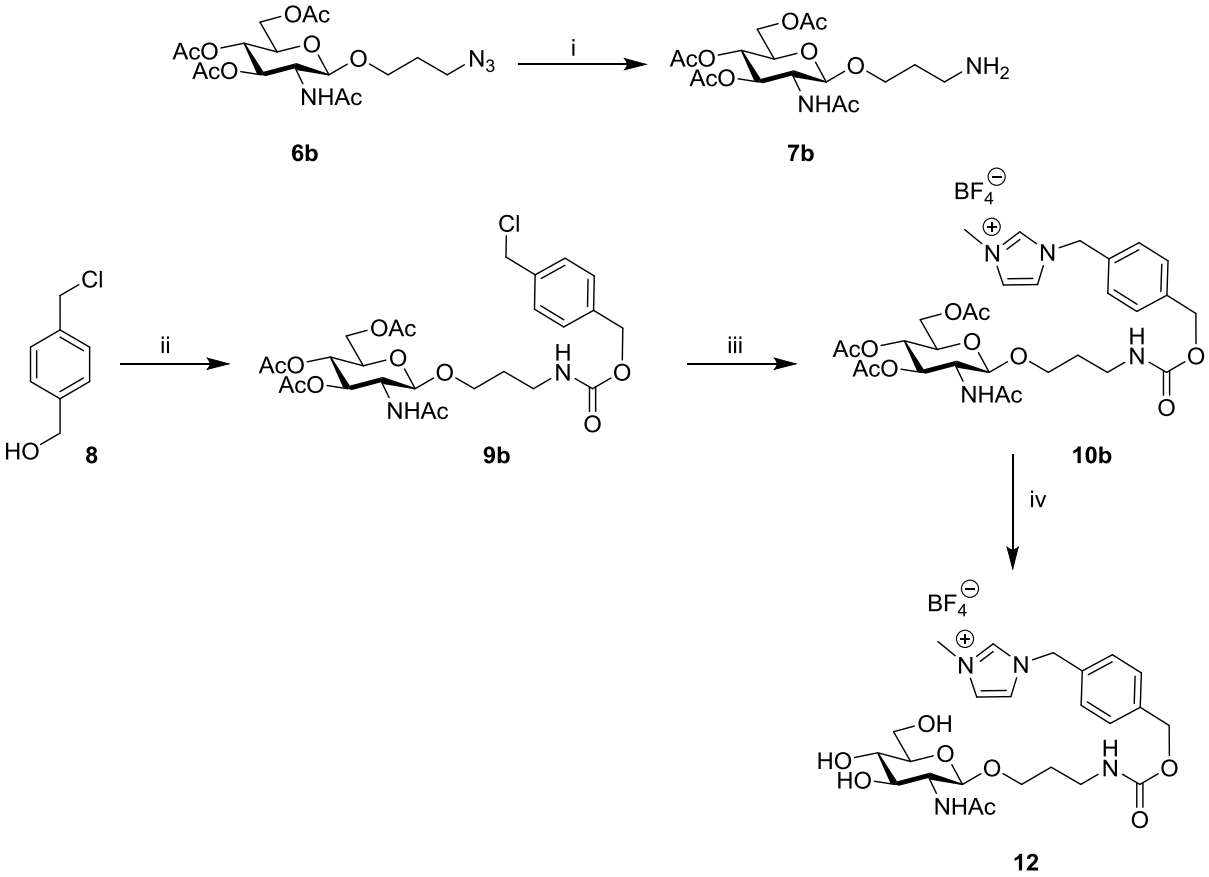
without further purification. To a stirring solution of 4-chloromethylphenyl methanol (440 mg, 2.78 mmol) dissolved in anhydrous THF (20 mL) at rt was added CDI (450 mg, 2.78 mmol) and the mixture stirred for 0.5 h, upon which time, a solution of **7a** (940 mg, 2.32 mmol) in anhydrous THF (3 mL) at rt was added dropwise and stirred for another 16 h. The reaction mixture was then diluted with DCM and the organic layer washed with 1 M HCl (3 x 5 mL) and brine (3 x 5 mL) before drying over anhydrous $MgSO_4$. The dried residue was purified by flash column chromatography (Hexane:EtOAc 6:4) to give **9a** (595 mg, 44% over 2 steps) as a pale orange foam.
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.34 (d, *J* = 8.3 Hz, 2H, C*H*, Ar), 7.29 (d, *J* = 8.3 Hz, 2H, C*H*, Ar), 5.19 (dd, *J* = 9.7, 9.5 Hz, 1H, H-3), 5.15 – 5.02 (m, 4H, H-4 & OC*H*2Ar), 4.97 (dd, *J* = 9.7, 8.0 Hz, 1H, H-2), 4.57 (s, 2H, ArC*H*2Cl), 4.49 (d, *J* = 8.0 Hz, 1H, H-1), 4.22 (dd, *J* = 12.3, 4.7 Hz, 1H, H-6a), 4.14 (dd, *J* = 12.3, 2.4 Hz, 1H, H-6b), 3.95 – 3.76 (m, 1H, OC*H*HCH_2), 3.67 (ddd, *J* = 10.2, 4.7, 2.4 Hz, 1H, H-5), 3.58 (dd, *J* = 7.7, 3.4 Hz, 1H, OCH*H*CH_2), 3.30 (dt, *J* = 12.5, 6.4 Hz, 1H, CH_2C*H*HNH), 3.21 (dt, *J* = 13.5, 6.4 Hz, 1H, CH_2CH*H*NH), 2.04 (s, 3H, CH_3, Ac), 2.01 (s, 6H, 2 x CH_3, Ac), 1.99 (s, 3H, CH_3, Ac), 1.84 – 1.66 (m, 2H, OCH_2CH_2); **^{13}C NMR** (126 MHz, $CDCl_3$) δ 170.8 (*C*O, Ac), 170.3 (*C*O, Ac), 169.5 (*C*O, Ac), 169.5 (*C*O, Ac), 156.4 (N*C*OO), 137.4 (Cq, Ar), 137.1 (Cq, Ar), 128.9 (*C*H, Ar), 128.5 (*C*H, Ar), 100.7 (C-1), 72.8 (C-3), 72.0 (C-5), 71.3 (C-2), 68.5 (C-4), 67.7 (OCH_2CH_2), 66.1 (OCH_2Ar), 62.0 (C-6), 46.0 (ArCH_2Cl), 38.3 (CH_2CH_2NH), 29.6 (OCH_2CH2), 20.8 (CH_3, Ac), 3 x 20.7 (CH_3, Ac); **ESI-HRMS** for $C_{26}H_{34}ClNO_{12}Na^+$, $[M+Na]^+$; calculated 610.1662; found 610.1696; **v max** 2960, 2882, 1755, 1524, 1372, 1226, 1066, 1039; $[\alpha]_D^{24}$ = –26.7 (*c* = 0.3, MeOH).

3-(4-(3-Methylimidazolium)methylbenzyloxycarbonylamino)propyl 2,3,4,6-tetra-*O*-acetyl-*β*-D-glucopyranoside (10a)
To a stirring solution of **9a** (330 mg, 0.56 mmol) in MeCN (6 mL) at rt was added KBF_4 (280 mg, 2.24 mmol) and 1-methylimidazole (180 mg, 0.18 mmol) and the mixture stirred at 80 ºC for 16 h. On cooling to rt, KBF_4 was removed by filtration before the solution was concentrated under reduced pressure. The dried residue was subsequently purified via preparative HPLC with a DCM/MeOH to give **10a** (220 mg, 54 %) as a white foam.
^{1}H NMR (500 MHz, CD_3OD) δ 9.03 (s, 1H, NC*H*N), 7.66 – 7.57 (m, 2H, N(C*H*)$_2$N), 7.46 (s, 4H, 4 x C*H*, Ar), 5.45 (s, 2H, ArCH_2N), 5.27 (dd, *J* = 10.2, 9.5 Hz, 1H, H-3), 5.11 (s, 2H, OCH_2Ar), 5.03 (t, *J* = 9.9 Hz, 1H, H-4), 4.91 (dd, *J* = 9.5, 8.1 Hz, 1H, H-2), 4.69 (d, *J* = 8.0 Hz, 1H, H-1), 4.29 (dd, *J* = 12.3, 4.6 Hz, 1H, H-6a), 4.15 (dd, *J* = 12.4, 2.4 Hz, 1H, H-6b), 3.96 (s, 3H, NCH_3), 3.93 – 3.81 (m, 2H, H-5 & OC*H*HCH_2), 3.63 (dt, *J* = 10.1, 6.2 Hz, 1H, OCH*H*CH_2), 3.20 (*app*. h, *J* = 6.7 Hz, 2H, CH_2CH_2NH), 2.06 (s, 3H, CH_3, Ac), 2.04 (d, *J* = 2.2 Hz, 6H, 2 x C*H*3, Ac), 1.99 (s, 3H, CH_3, Ac), 1.77 (*app*. p, *J* = 5.4 Hz, 2H, OCH_2CH_2); **^{13}C NMR** (126 MHz, CD_3OD) δ 172.3 (*C*O, Ac), 171.6 (*C*O, Ac), 171.3 (*C*O, Ac), 171.2 (*C*O, Ac), 158.6 (N*C*OO), 138.0 (N*C*HN), 136.6 (Cq, Ar), 134.8 (Cq, Ar), 129.8 (*C*H, Ar), 129.5 (*C*H, Ar), 125.3 (N*C*HCHN), 123.7 (NCH*C*HN), 101.7 (C-1), 74.3 (C-3), 72.9 (C-2), 72.8 (C-5), 69.9 (C-4), 68.5 (OCH_2CH_2), 66.6 (OCH_2Ar), 63.1 (C-6), 53.8 (ArCH_2N), 38.9 (CH_2CH_2NH), 36.6 (NCH_3), 30.8 (OCH_2CH_2), 3 x 20.6 (CH_3, Ac), 20.5 (CH_3, Ac); **ESI-HRMS** for $C_30H_{40}N_3O_{12}^+$, $[M^+]$; calculated: 634.2607; found: 634.2610; **v max** 3153, 3108, 2926, 2879, 1699, 1534, 1292, 1077, 1038; $[\alpha]_D^{24}$ = –9.0 (*c* = 6.7, MeOH).

3-(4-(3-Methylimidazolium) methylbenzyloxycarbonylamino) propyl-*β*-D-glucopyranoside (11)
To a stirring solution of **10a** (220 mg, 0.30 mmol) in methanol (7 mL) at rt was added Et3N (1.1 mL, 8.10 mmol) and the reaction stirred for 16 h. The mixture was concentrated under reduced pressure and the dried residue purified via preparative RP-HPLC with a H_2O/MeOH gradient to give the **11** (130 mg, 78%) as a colourless oil.

^{1}H NMR (400 MHz, D_2O) δ 8.72 (s, 1H, NC*H*N), 7.50 – 7.31 (m, 6H, 4 x C*H*, Ar & N(C*H*)$_2$N), 5.38 (s, 2H, ArCH_2N), 5.10 (s, 2H, OCH_2Ar), 4.42 (d, *J* = 7.9 Hz, 1H, H-1), 4.03 – 3.79 (m, 5H, H-6a & OC*H*HCH_2 & NCH_3), 3.79 – 3.63 (m, 2H, H-6b & OCH*H*CH_2), 3.52 – 3.33 (m, 3H, H-3 & H-4 & H-5), 3.30 – 3.15 (m, 3H, H-2 & CH_2CH_2NH), 1.80 (*app*. p, *J* = 6.5 Hz, 2H, OCH2C*H*2); **^{13}C NMR** (101 MHz, D2O) δ 158.3 (HN*C*OO), 137.7 (N*C*HN), 136.1 (Cq, Ar), 133.4 (Cq, Ar), 128.8 (2 x *C*H, Ar), 128.2 (2 x *C*H, Ar), 123.8 (N*C*HCHN), 122.3 (NCH*C*HN), 102.2 (C-1), 75.9 (C-3), 75.7 (C-5), 73.1 (C- 2), 69.7 (C-4), 67.7, (OCH_2CH_2), 66.2 (OCH_2Ar), 60.8 (C-6), 52.5 (ArCH_2N), 37.4 (CH_2CH_2NH), 35.7 (NCH_3), 28.9 (OCH2*C*H2); **MALDI-HRMS** for $C_{22}H_{32}N_3O_8^+$, $[M^+]$; calculated: 466.2184; found: 466.2192; **v max** 3388, 3159, 2935, 2890, 1698, 1532; $[\alpha]_D^{23}$ = – 8.9 (*c* = 0.9, MeOH).

Chemical synthesis of GlcNAc-ITag 12



Scheme S2. Reagents and conditions for the synthesis of **12**: i) H_2, Pd/C, THF, rt, 1 h (quant.); ii) CDI, THF, rt, 30 min; then GlcNAc-O$(CH_2)_3$-NH_2 (**7b**), 15 h (54%); iii) 1-methylimidazole, KBF_4, CH_3CN, 80 °C, 15 h (61%); iv) Et_3N, MeOH, 16 h (63%).

3-Azidopropyl 2-acetamido-2-deoxy-3,4,6-tri-*O*-acetyl-*β*-D-glucopyranoside (6b)
To a stirring solution of 3-bromopropyl 2-acetamido-2-deoxy-3,4,6-tri-*O*-acetyl-*β*-D-glucopyranoside (5.00 g, 10.7 mmol) in DMF (110 mL) at rt, was added NaN_3 (2.08 g, 32.0 mmol). The resulting mixture was warmed to 80°C and left stirring for 16 h. The mixture was then cooled to rt, filtered through celite and then concentrated under reduced pressure. The dried residue was then purified by flash column chromatography (98:2 DCM:MeOH) to give **6b** (4.96 g, 85%) as a pale orange solid.
^{1}H NMR (400 MHz, $CDCl_3$) δ 5.48 (d, *J* = 8.9 Hz, 1H, N*H*Ac), 5.23 (dd, *J* = 10.7, 9.3 Hz, 1H, H-3), 5.07 (dd, *J* = 9.9, 9.3 Hz, 1H, H-4), 4.61 (d, *J* = 8.3 Hz, 1H, H-1), 4.25 (dd, *J* = 12.3, 4.8 Hz, 1H, H-6a), 4.14 (dd, *J* = 12.3, 2.5 Hz, 1H, H-6b), 4.00 – 3.84 (m, 2H, H-2 & OC*H*HCH_2), 3.69 (ddd, *J* = 9.9, 4.8, 2.5 Hz, 1H, H-5), 3.59 (*app*. ddd, *J* = 9.8, 8.2, 4.8 Hz, 1H, OCH*H*CH_2), 3.45 – 3.27 (m, 2H, $CH_2CH_2N_3$), 2.09 (s, 3H, CH_3, Ac), 2.03 (s, 3H, CH_3, Ac), 2.03 (s, 3H, CH_3, Ac), 1.96 (s, 3H, CH_3, NHAc), 1.94 – 1.70 (m, 2H, OCH_2CH_2); **^{13}C NMR** (101 MHz, $CDCl_3$) δ 171.1 (*C*O, NHAc), 170.8 (*C*O, Ac), 170.2 (*C*O, Ac), 169.5 (*C*O, Ac), 101.2 (C-1), 72.5 (C-3), 72.1 (C-5), 68.7 (C-4), 66.4 (OCH2CH2), 62.2 (C-6), 54.7 (C-2), 48.2 ($CH_2CH_2N_3$), 29.1 (OCH_2CH_2), 23.5 (CH_3, NHAc), 20.9 (CH_3, Ac), 20.8 (CH_3, Ac), 20.8 (CH_3, Ac). The spectroscopic data is in accordance with the data reported in the literature.[2]

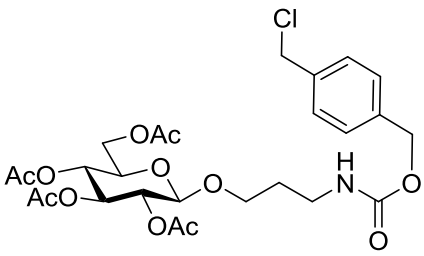
3-(4-(Chloromethyl)benzyl carbamate) propyl 2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy- *β*-D-glucopyranoside (9b)
To a stirring solution of **6b** (300 mg, 0.70 mmol) in THF (7 mL) at rt was added Pd/C (37 mg, 0.35 mmol). The mixture was stirred under a H_2 atmosphere for 1 h after which the mixture was diluted in EtOAc (20 mL) and filtered through celite. The filtrates were then concentrated under reduced

pressure to give **7b** (270 mg, *quant.*) as a greyish residue. The compound was used directly in the next step without further purification. To a stirring solution of 4-chloromethylphenyl methanol (131 mg, 0.84 mmol) in anhydrous THF (50 mL) was added CDI (99 mg, 6.13 mmol) and the mixture was left stirring for 0.5 h, upon which time, a solution of **7b** (281 mg, 0.70 mmol) in anhydrous THF (3 mL) at rt was added dropwise and the mixture left stirring for another 16 h. The reaction mixture was then diluted in DCM (20 mL) and the organic layer washed with 1 M HCl (5 mL), followed by brine (5 mL) before drying over anhydrous $MgSO_4$. The solution was then concentrated under reduced pressure and the dried residue purified by flash column chromatography (98:2 DCM:MeOH) to give **9b** (222 mg, 54%) as a white solid.
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.52 – 7.30 (m, 4H, C*H*, Ar), 5.15 (t, *J* = 10.0 Hz, 1H, H-3), 5.12 – 5.02 (m, 2H, OCH_2Ar), 4.96 (t, *J* = 9.7 Hz, 1H, H-4), 4.63 (s, 2H, ArCH_2Cl), 4.52 (d, *J* = 8.4 Hz, 1H, H-1), 4.25 (dd, *J* = 12.3, 4.6 Hz, 1H, H-6a), 4.10 (dd, *J* = 12.3, 2.4 Hz, 1H, H-6b), 3.90 – 3.79 (m, 2H, OC*H*HCH$_2$), 3.72 (ddd, *J* = 10.3, 4.0, 2.3 Hz, 1H, H-5), 3.52 (ddd, *J* = 9.8, 7.4, 6.1 Hz, 1H, OCH*H*CH$_2$), 3.28-3.20 (m, 1H, CH$_2$C*H*HNH), 3.12 (dt, *J* = 13.3, 6.4 Hz, 1H, CH$_2$CH*H*NH), 2.03 (s, 3H, CH_3-CO), 2.00 (s, 3H, CH$_3$-CO), 1.97 (s, 3H, CH$_3$-CO), 1.90 (s, 3H, CH_3-CONH), 1.81 – 1.71 (m, 2H, -OCH$_2$CH_2). **^{13}C NMR (126 MHz, $CDCl_3$)** δ 172.1 (CH$_3$CONH), 171.0 (CH$_3$CO), 170.5 (CH$_3$CO), 169.8 (CH$_3$CO), 157.4 (OCONH), 137.7 (C-Ar), 137.4 (C-Ar) , 128.5 (C-Ar), 127.7 (C-Ar), 100.8 (C-1), 72.9 (C-3),71.5 (C-5), 68.7 (C-4), 66.8 (C-1'), 65.5 (CH$_2$-O), 61.8 (C-6), 54.0 (C-2), 45.1 (CH$_2$-Cl), 37.2 (C-3'), 29.4 (C-2'), 21.4 (CH$_3$-CONH), 19.2 (CH$_3$-CO), 19.2 (CH$_3$-CO), 19.2 (CH$_3$-CO). **ESI-HRMS** for $C_{26}H_{36}ClN_2O_{11}^+$, $[M+H]^+$; calculated: 587.2002; found: 587.2057; **v max** 3325, 3303, 2954, 1735, 1687, 1659, 1539, 1372, 1231, 1035; $[\alpha]_D^{23}$ = – 8.0 (*c* = 0.5, MeOH); **m.p.** 164-165 ºC.

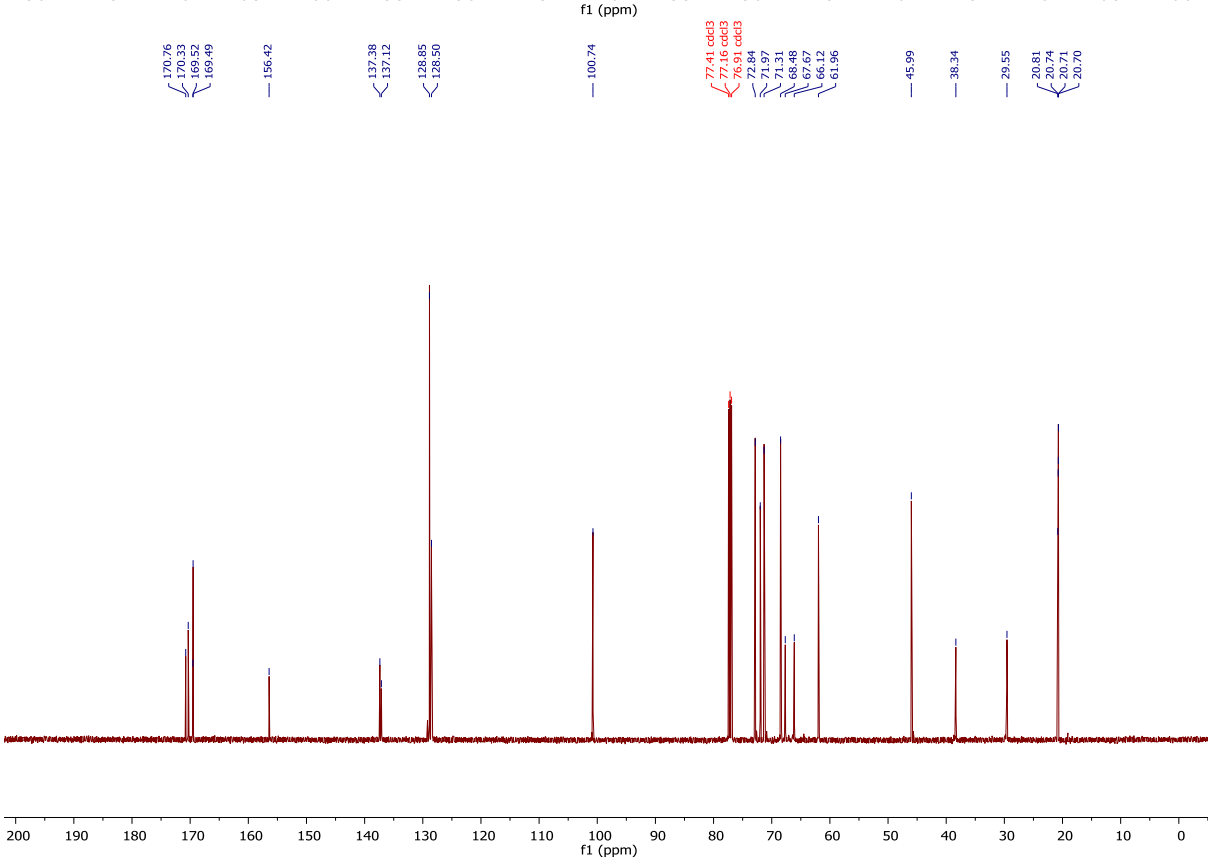
3-(4-(3-Methylimidazolium)methylbenzyl carbamate) propyl 2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-*β*-D-glucopyranoside (10b)
To a stirring solution of **9b** (200 mg, 0.34 mmol) in MeCN (13.5 mL) was added KBF_4 (129 mg, 1.02 mmol) and 1-methylimidazole (0.08 mL, 1.02 mmol) and the mixture left stirring at 80 ºC for 16 h. On cooling to rt, KBF_4 was removed by filtration before the solution was concentrated under reduced pressure. The dried residue was subsequently purified by preparative HPLC with a DCM/MeOH gradient to give **10b** (149 mg, 61 %) as a pale yellow solid.
^{1}H NMR (400 MHz, MeOD) δ 8.92 (s, 1H, NCHN), 7.64-7.50 (m, 2H, NCHCHN), 7.42 (s, 4H, H-Ar), 5.39 (s, 2H, ArCH_2N), 5.17 (dd, *J* = 10.5, 9.3 Hz, 1H, H-3), 5.08 (s, 2H,OCH_2Ar), 4.96 (t, *J* = 9.7 Hz, 1H, H-4), 4.58 (d, *J* = 8.5 Hz, 1H, H-1), 4.25 (dd, *J* = 12.3, 4.7 Hz, 1H, H-6a), 4.11 (dd, *J* = 12.3, 2.4 Hz, 1H, H-6b), 3.91 (s, 3H, CH$_3$-N), 3.91 – 3.80 (m, 2H, H-2, OCH*H*CH$_2$), 3.76 (ddd, *J* = 10.1, 4.7, 2.4 Hz, 1H, H-5), 3.56 (dt, *J* = 10.0, 6.1 Hz, 3H, H-2, H-5, OC*H*HCH$_2$), 3.26-3.08 (m, 2H, OCH$_2$CH_2NH), 2.02 (s, 3H, CH$_3$CO), 2.00 (s, 3H, CH$_3$CO), 1.97 (s, 3H, CH$_3$CO), 1.89 (s, 3H, CH$_3$CONH), 1.73 (p, *J* = 6.6 Hz, 2H, OCH$_2$CH_2). **^{13}C NMR (126 MHz, MeOD)** δ 172.2 (NHCOCH$_3$), 171.0 (COCH$_3$), 170.5 (COCH$_3$), 170.0 (COCH$_3$), 157.3 (OCONH), 138.6 (C-Ar), 133.4 (C-Ar), 128.5 (C-Ar), 128.1 (C-Ar), 123.8 (N*CH*CHN), 122.2 (NCH*CH*N), 100.7 (C-1), 72.9 (C-3), 71.5 (C-2), 68.8 (C-4), 67.0 (C-1'), 65.3 (CH$_2$-O), 61.9 (C-6), 54.0 (C-5), 52.3 (CH$_2$-Im), 37.4 (C-3'), 35.1 (CH$_3$-N), 29.3 (C-2'), 21.4 (CH$_3$CONH), 19.3 (CH$_3$CO), 19.2 (CH$_3$CO), 19.2 (CH$_3$CO). **ESI-HRMS** for $C_{30}H_{41}N_4O_{11}^+$, $[M^+]$; calculated: 633.2766; found: 633.2770; **vmax** 3435, 3290, 2947, 1745, 1727, 1655, 1547, 1504, 1376, 1239, 1032; $[\alpha]_D^{24}$ = –12.0 (*c* =1.7, MeOH); **m.p.** 86-88 ºC.

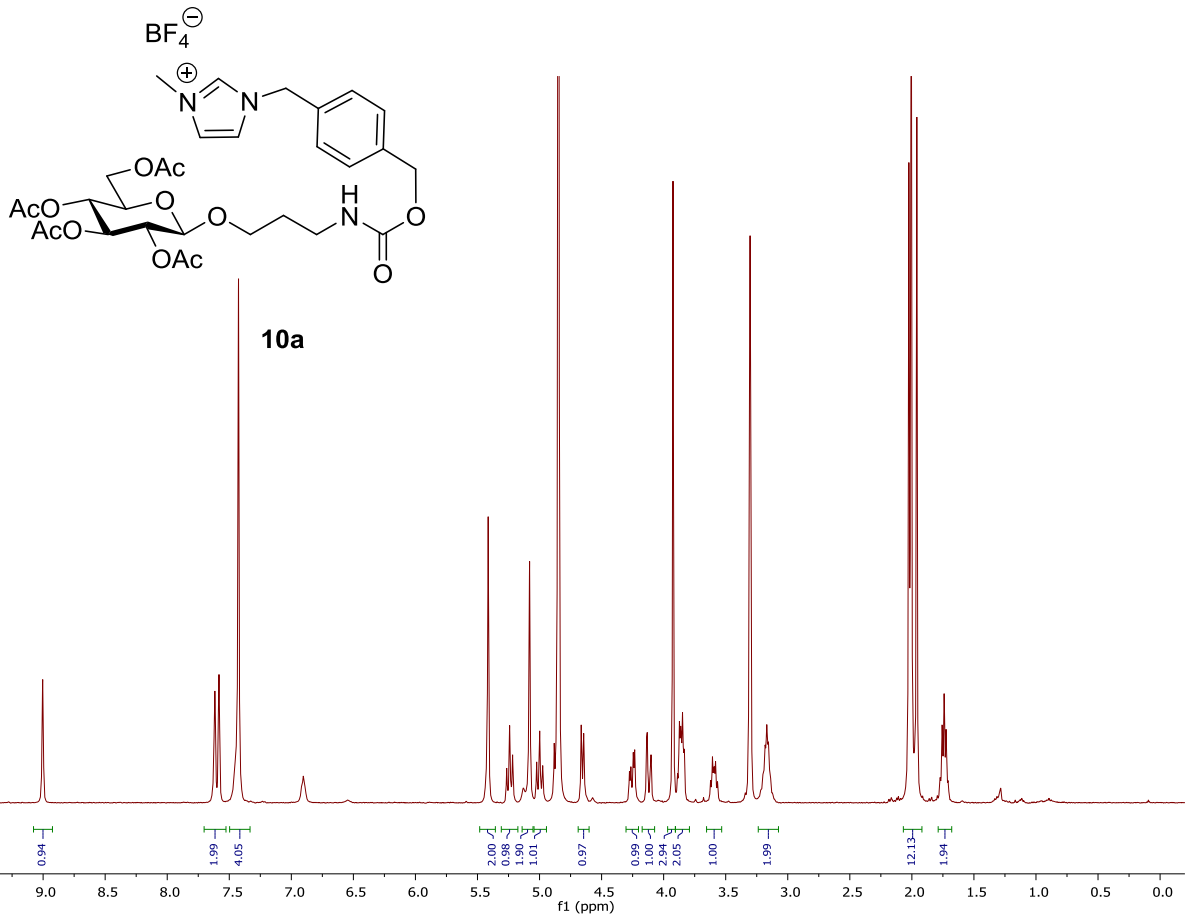
3-(4-(3-Methylimidazolium)methylbenzyl carbamate) propyl 2-acetamido-2-deoxy- *β*-D-glucopyranoside (12)
To a stirring solution of **10b** (280 mg, 0.38 mmol) in methanol (9 mL) at rt was added Et_3N (1.4 mL, 10.6 mmol) and the reaction left stirring for 16 h. The mixture was concentrated under reduced pressure and the dried residue purified by preparative RP-HPLC H_2O/MeOH to give **12** (140 mg, 63%) as a white solid.
^{1}H NMR (500 MHz, D_2O) δ 8.73 (s, 1H, NC*H*N), 7.64 – 7.19 (m, 6H, C*H*, Ar), 5.40 (s, 2H, ArCH_2N), 5.13 (s, 2H, OCH_2Ar), 4.48 (d, *J* = 8.5 Hz, 1H, H-1), 3.97 – 3.89 (m, 2H, OC*H*HCH$_2$ & H-6a), 3.88 (s, 3H, NCH_3), 3.74 (ddd, *J* = 12.8, 4.1, 2.0 Hz, 1H, H-6b), 3.68 (td, *J* = 10.5, 8.5, 2.0 Hz, 1H, H-2), 3.62 (dt, *J* = 10.4, 6.2 Hz, 1H, OCH*H*CH$_2$), 3.53 (ddd, *J* = 10.5, 8.7, 4.7 Hz, 1H, H-3), 3.47 – 3.38 (m, 2H, H-4 & H-5), 3.18 (td, *J* = 12.5, 6.5 Hz, 2H, CH$_2$CH_2NH), 2.01 (s, 3H, CH_3, NHAc), 1.75 (*app*. p, 6.4 Hz, 2H, OCH$_2$CH_2); **^{13}C NMR** (126 MHz, D_2O) δ 175.6 (*C*O, NHAc), 158.2 (HN*C*OO), 137.7 (N*C*HN), 133.3 (Cq, Ar), 131.0 (Cq, Ar), 128.7 (*C*H, Ar), 128.1 (*C*H, Ar), 123.7 (N*C*HCHN), 122.2 (NCH*C*HN), 101.0 (C-1), 75.8 (C-5), 73.7 (C-3), 69.9 (C-4), 67.6 (O*C*H$_2$CH$_2$), 66.1 (O*C*H$_2$Ar), 60.7 (C-6), 55.5 (C-2), 52.4 (Ar*C*H$_2$N), 37.3 (CH$_2$*C*H$_2$NH), 35.6 (N*C*H$_3$), 28.7 (OCH$_2$*C*H$_2$), 22.0 (*C*H3, NHAc); **MALDI-HRMS** for $C_{24}H_{35}N_4O_{8+}$, $[M^+]$; calculated: 507.2449; found: 507.2440; **v max** 3337, 2948, 2837, 1651, 1408, 1014; $[\alpha]_D^{24}$ = –13.3 (*c* = 0.6, MeOH); **m.p.** 126-128 ºC.

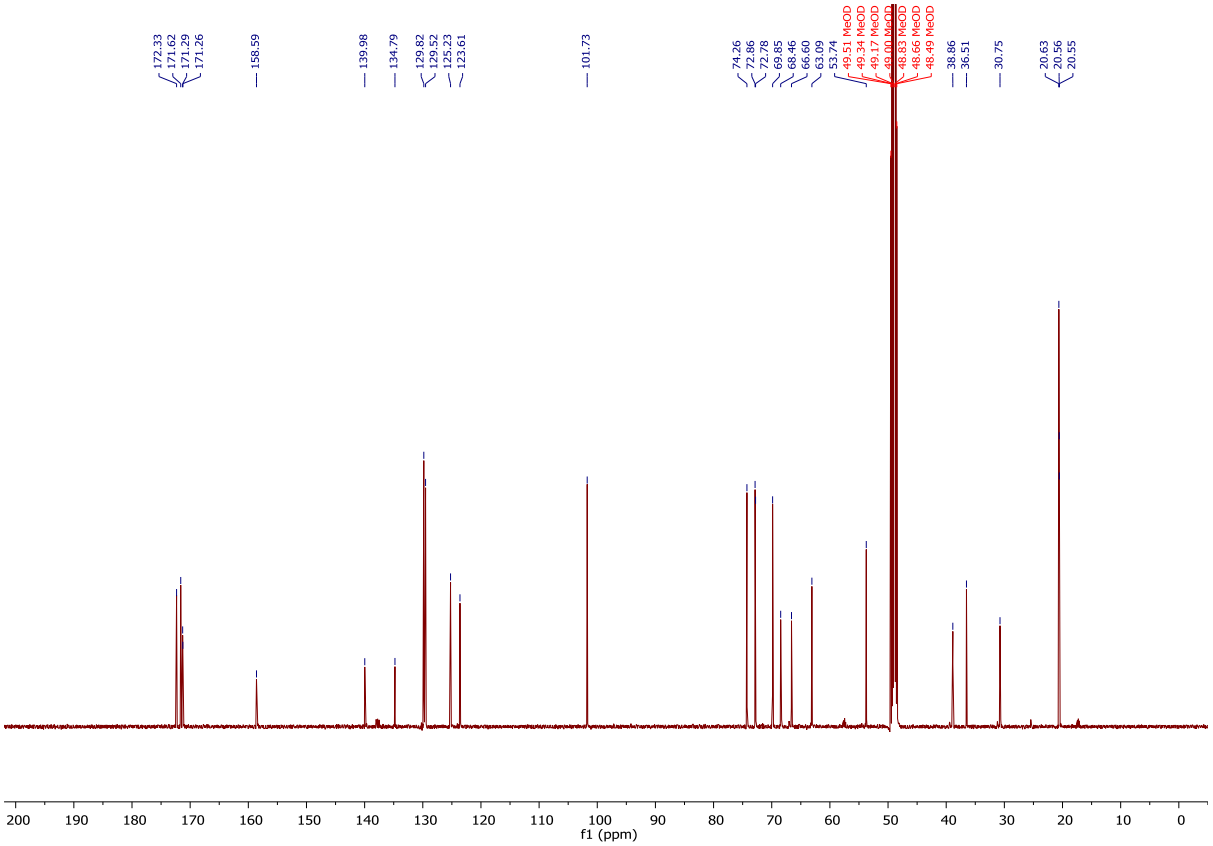
NMR spectra of Glc-ITag 11, GlcNAc-ITag 12 and synthetic intermediates 9a-b and 10a-b

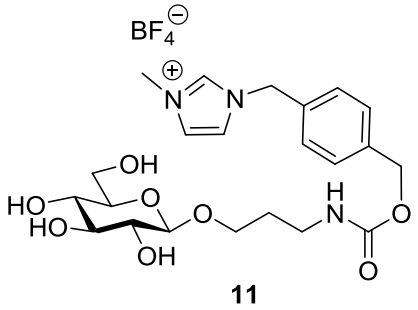


9a



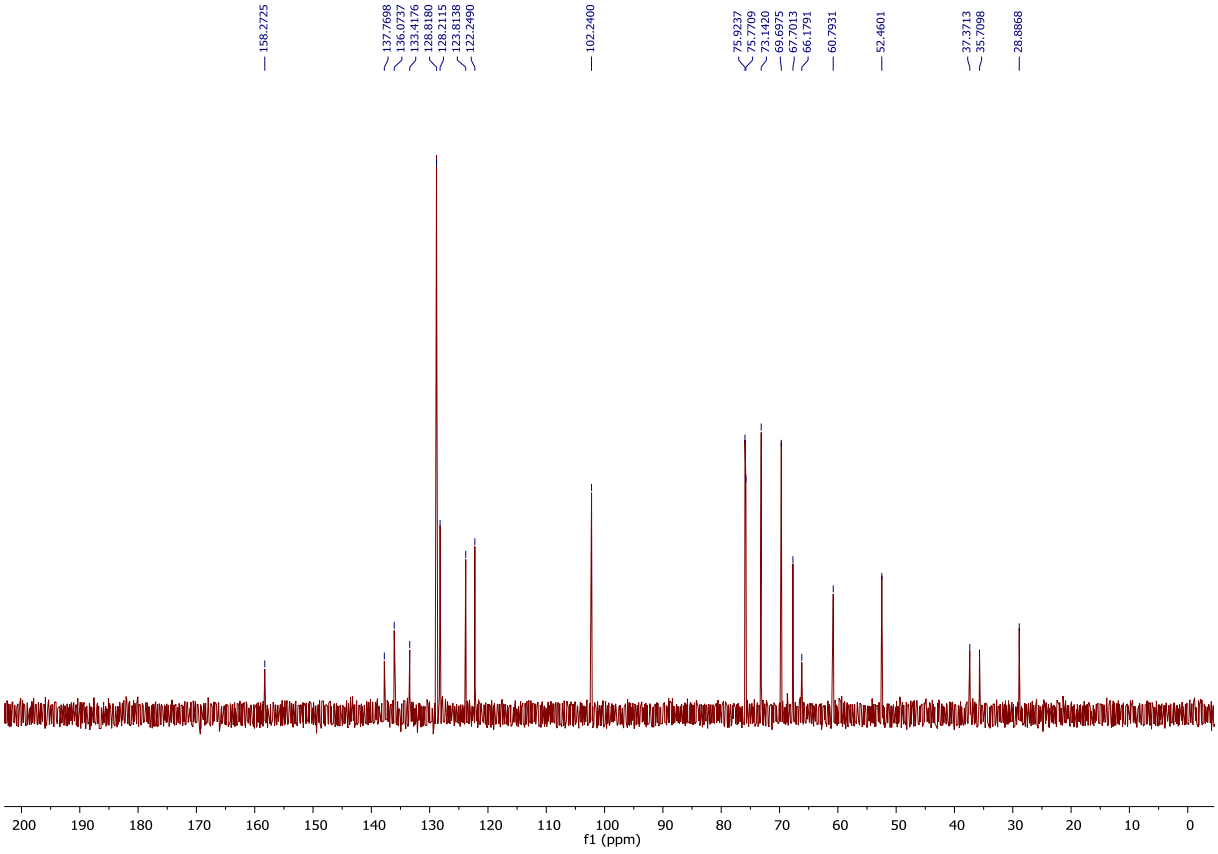

BF4
OAc
AcO
AcO
OAc
O
O
H
N
O
O
N
N
10a
0.94
1.99
4.05
2.00
0.98
1.90
1.01
0.97
0.99
1.00
2.94
2.05
1.00
1.99
12.13
1.94
f1 (ppm)

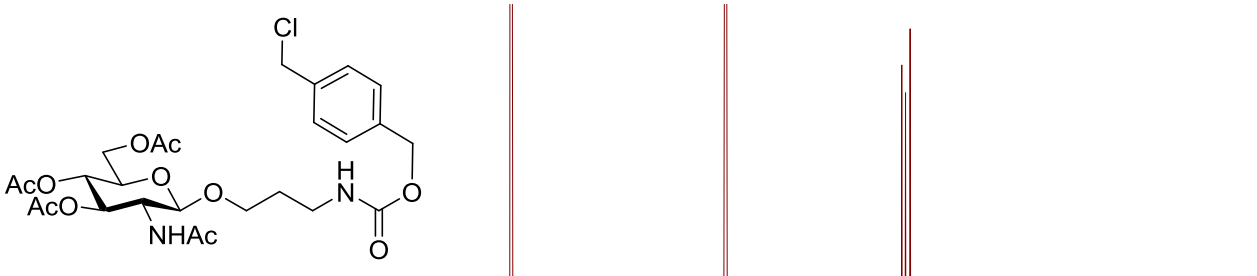

172.33
171.62
171.29
171.26
158.59
139.98
134.79
129.82
129.52
125.23
123.61
101.73
74.26
73.86
72.78
68.85
68.46
66.60
63.09
53.74
49.51 MeOD
49.34 MeOD
49.17 MeOD
49.00 MeOD
48.83 MeOD
48.66 MeOD
48.49 MeOD
38.86
36.51
30.75
20.63
20.56
20.55
f1 (ppm)


BF4
OH
HO
HO
OH
O
O
H
N
O
O

11

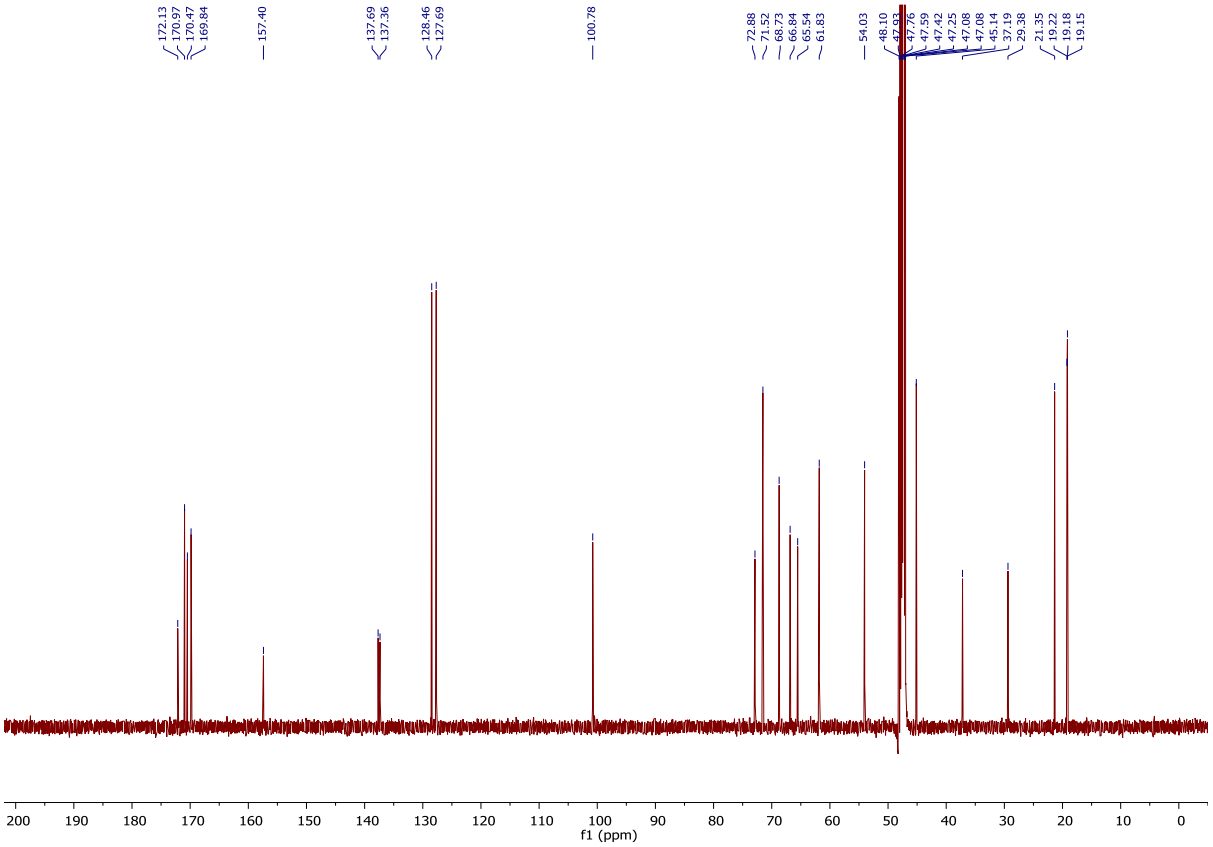
1.02
5.78
1.90
2.07
1.00
4.60
2.00
3.17
2.97
2.06
9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

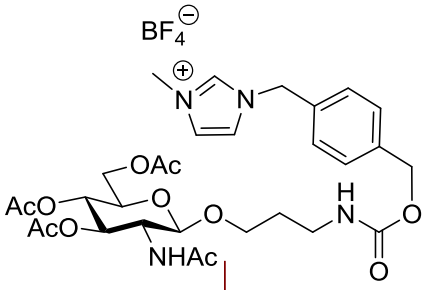

158.2725
137.7698
136.0737
133.4176
128.8180
128.2115
123.8138
122.2490
102.2400
75.9237
75.7709
73.1420
69.6975
67.7013
66.1791
60.7931
52.4601
37.3713
35.7098
28.8868
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)


Cl
OAc
AcO
AcO
O
O
NHAc
H
N
O
O

9b

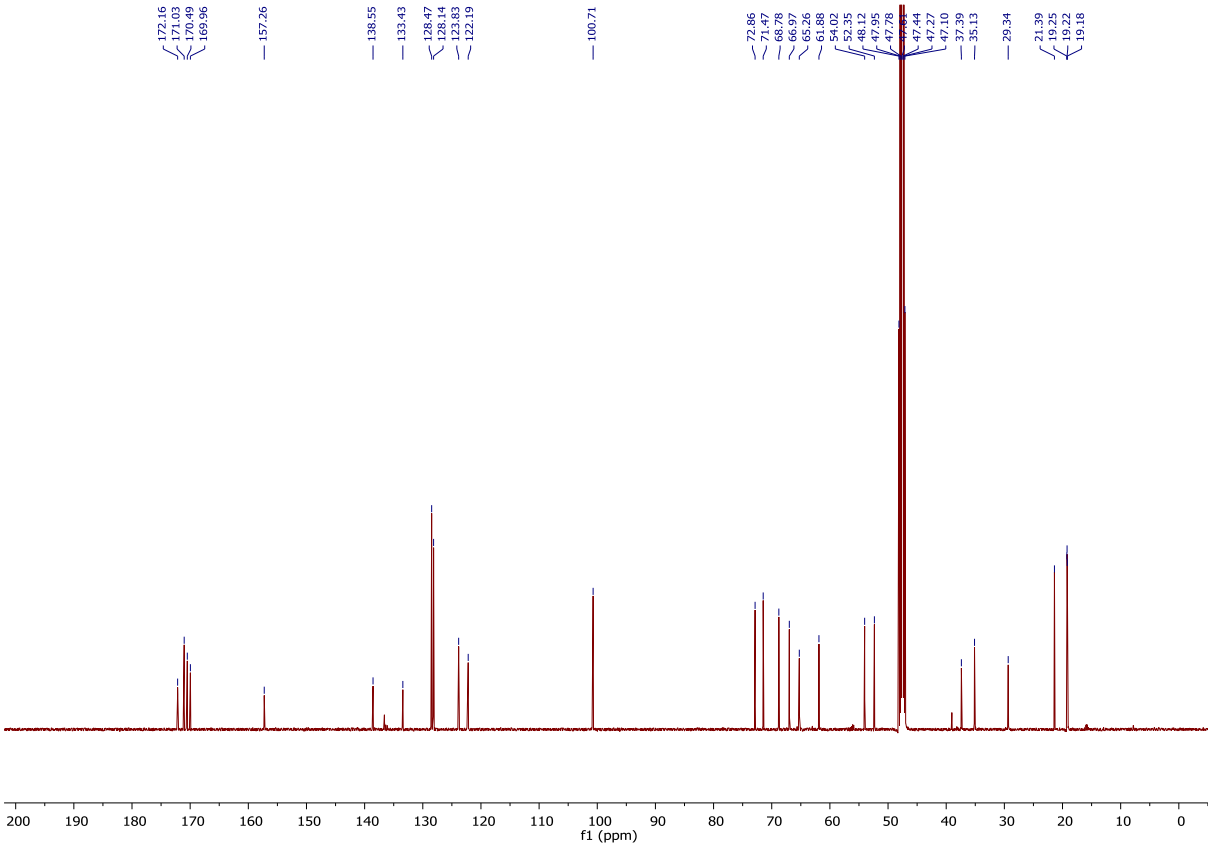
4.49
3.10
1.00
2.20
0.82
1.09
1.06
2.12
1.01
0.88
2.58
0.88
10.13
3.25
2.06
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

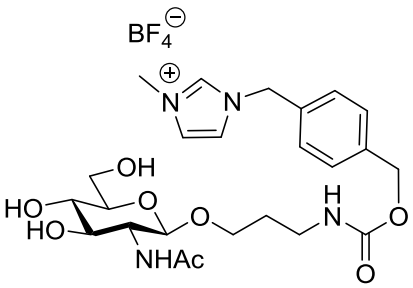

172.13
170.97
170.47
169.84
157.40
137.69
137.36
128.46
127.69
100.78
72.88
71.52
68.73
66.84
65.54
61.83
54.03
48.10
47.93
47.76
47.59
47.42
47.25
47.08
47.08
45.14
37.19
29.38
21.35
19.22
19.18
19.15
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)


BF4
OAc
AcO
AcO
NHAc
H
N
O
O
O
O
N
N

10b

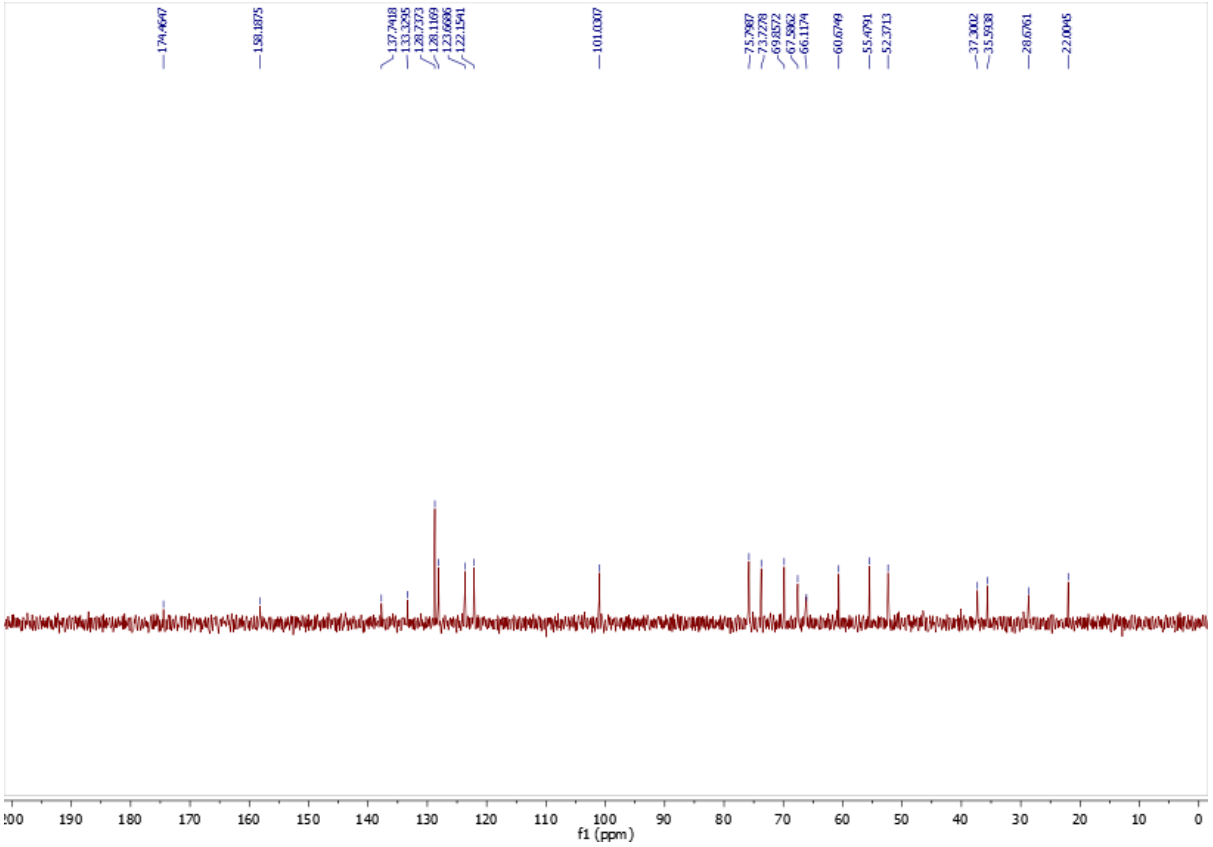
0.88
1.92
3.85
1.93
1.17
1.86
1.10
1.78
0.95
0.95
4.79
1.00
0.98
1.96
11.86
1.94
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


172.16
171.03
170.49
169.96
157.26
138.55
133.43
128.47
128.14
123.83
122.19
100.71
72.86
71.47
68.78
66.97
65.26
61.88
54.02
52.35
48.12
47.95
47.78
47.61
47.44
47.27
47.10
37.39
35.13
29.34
21.39
19.25
19.22
19.18
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)


BF4
N
N
OH
HO
HO
O
O
NHAc
H
N
O
O

12

0.08
6.00
2.02
2.01
0.96
2.00
2.91
0.94
0.97
0.97
0.94
1.96
1.98
2.75
1.98
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)


174.4667
158.1875
137.7418
133.3295
128.7373
128.1169
123.6686
122.1541
101.0307
75.7987
73.7278
69.8572
67.5962
66.1174
60.6749
55.4791
52.3713
37.3002
35.5938
28.6761
22.0045
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

References

1 C. He, S. Wang, M. Liu, C. Zhao, S. Xiang and Y. Zeng, *Org. Biomol. Chem.*, 2016, **14**, 1611–1622.

2 I. Choudhury, N. Minoura and H. Uzawa, *Carbohydr. Res.*, 2003, **338**, 1265–1270.