Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry. This journal is © The Royal Society of Chemistry 2017

An Enzymatic Strategy to Asymmetrically Branched N-glycans

Angie D. Calderon,^a Jun Zhou,^a Wanyi Guan,^{a,b} Zhigang Wu,^a Yuxi Guo,^a Jing Bai,^b Qing Li,^b Peng George Wang,^a Junqiang Fang^{*c} and Lei Li^{*a}

^aDepartment of Chemistry and Center for Diagnostics & Therapeutics, Georgia State University, Atlanta, GA 30303

^bCollege of Life Science, Hebei Normal University, Shijiazhuang, Hebei 050024, China

[°]National Glycoengineering Research Center, Shandong Provincial Key Laboratory of Glycochemistry and Glycobiology, Shandong University, Jinan, Shandong 250100, China

Supplementary Information

I.	Materials and enzymes	S2
II.	Expression and purification of human MGAT5	S2
III.	Isolation of symmetric N-glycans N001 and N301 from bovine fetuin	S3
IV.	Label N-glycans with 2-AA and AEAB	S4
V.	Detailed substrate specificity of MGAT5	S5
VI.	Substrate specificity study of β -galactosidase from <i>Escherichia coli</i>	S 7
VII.	Generation of asymmetric N-glycan structures using LacZ	S 8
VIII.	Glycosyltransferase catalyzed reactions	S9
IX.	HPLC, MALDI-MS and NMR analysis of purified N-glycans	S11
X.	References	S31

I. Materials and enzymes

Fetuin from fetal bovine serum was purchased from Sigma-Aldrich. Thermosensitive Alkaline Phosphatase (FastAP) was purchased from Thermo Scientific. *Escherichia coli* β-galactosidase (LacZ), PNGase F and neuraminidase from *Clostridium perfringens* were purchased from Sigma-Aldrich. Human FUT8 was expressed and purified as we did previously.¹ Sugar nucleotides uridine 5'-diphospho-GlcNAc (UDP-GlcNAc), guanosine 5'-diphospho-L-fucose (GDP-Fuc) and CMP-N-Glycolylneuraminic acid (CMP-Neu5Gc) were prepared as described previously.^{2,3}

II. Expression and purification of human MGAT5

Recombinant human mannosyl-a1,6-glycoprotein β1,6-N-acetyl-glucosaminyltransferase (MGAT5) were expressed using Sf9 insect cells as previously described.^{1,4} The baculovirus stocks for the expression of each of the hGTs were kindly provided by Dr. Donald Jarvis group from University of Wyoming NIH grant number RR005351. Briefly, Sf9 cells were infected at a cell density of 1×10⁶ C/mL. The baculovirus titer was estimated using insect cell density at time of infection and a conservative cellular baculovirus maximum production rate of 100 pfu/cell for 1×10^{6} C/mL meaning 1×10^{8} pfu/mL. BIIC stock was thawed in a 37 °C water bath and diluted 1:100 into cell-free medium, then 10 mL of infected cell-free medium was added to 1 L of uninfected cells (1:100 dilution). The medium was collected after 72 h of expression followed by purification. The purification of the secreted MGAT5 was performed by diluting the medium 1:1 with equilibration buffer (50 mM Tris-HCl pH 7.5, 0.5 M NaCl) and loading the expression medium to an ÄKTA pure fast protein liquid chromatography system (FPLC) from GE healthcare life sciences with a 1 mL pre-equilibrated His-Trap Excel affinity column (GE Healthcare)⁵. The column was rinsed using wash buffer (50 mM Tris-HCl pH 7.5, 0.5 M NaCl, 50 mM imidazole), and eluted with 10 mL elution buffer (50 mM Tris-HCl pH 7.5, 0.5 M NaCl, 250 mM imidazole). Enzyme was then desalted and concentrated using Amicon YM-30 ultrafiltration membrane (Millipore) against desalting buffer (50 mM Tris-HCl pH 7.5).⁵ Purified MGAT5 was confirmed by western blot (Figure S1). Around 3 mg of MGAT5 can be purified from 1 L cultures as determined using Pierce BCA protein assay kit (Thermo Fisher). The activity of FUT8 was confirmed using N000 (GlcNAcβ1,2-Manα1,6-(GlcNAcβ1,2-Manα1,3-)Manβ1,4-GlcNAcβ1,4-GlcNAc) as substrate.



Figure S1. Western blot analysis of MGAT5 using anti-His antibody as the primary antibody.

III. Isolation of symmetric N-glycans N001 and N3001 from bovine fetuin



Figure S2. Scheme for the generation of symmetric *N*-glycans **N001** and **N301** in one-pot reaction (PNGaseF and neuraminidase) and HPLC separation.



Figure S3. MALDI-MS analysis of N-glycan containing fractions from G25 gel filtration. **N301** with larger molecular weight elute faster than **N001**.

N001 and **N301** were generated from one-pot enzymatic digestion of bovine fetuin with PNGaseF and α -neuraminidase (**Figure S2**). Briefly, the 20 mL reaction system contains, 400 mg of bovine fetuin, 20 mM Tris-HCl, pH 8.0, 2 mM DTT, 10 units of PNGaseF and 20 units of neuraminidase from *Clostridium perfringens*. The reaction was allowed to proceed for 2 days at 37 °C and quenched by the addition of 3 volumes of ice-cold acetone and store at 4 °C for 1 hour. Protein was the removed by brief centrifuge, and the supernatant was concentrated by rotatory evaporation. After HPLC analysis using a Water amide column (130 Å, 5 µm, 4.6 mm × 250 mm) (Solvent A: Water, Solvent B: Acetonitrile, gradient elution with B% from 65% to 50% in 25 min, flow rate 1 mL/mL), the sample was loaded on a G25 gel filtration column (2.5 cm × 100 cm) for purification. One µL from each fraction tube (5 mL each) was spotted on a TLC plate, after dry, visualized by anisaldehyde sugar stain. Tubes with sugars were then analyzed by MALDI-MS and fractions contain *N*-glycans **N001** and/or **N301** (**Figure S3**) were further separated by HPLC using a Waters amide column (130 Å, 5 µm, 10 mm × 250 mm) with similar separation method as described above (flow rate is 4 mL/min). **N001** was characterized as did previously,⁶ and **N301** was characterized by HPLC-ELSD, MALDI-MS, and NMR (see IX).

<u>Calculation of yields</u>: Each fetuin protein contains 3 N-glycans, and around 80% is tri-antennary, 16% is bi-antennary as described.⁷ Calculation of yields are illustrated as below. The relatively low overall yield maybe caused by 1) fetuin purities and/or 2) incomplete digestion and product loss in purification steps.

Yield of Tri-antennary N301 =
$$26 \text{ mg (purified N301)}$$
X 100% = 65% $400 \text{ mg (total Fetuin)}$
 $48400 (MW of Fetuin)$ X 3 X 80% X 2005.7 (MW of N301)X 100% = 65%Yield of Tri-antennary N001 = $4.1 \text{ mg (purified N001)}$
 $\frac{400 \text{ mg (total Fetuin)}}{48400 (MW of Fetuin)}$ X 3 X 16% X 1640.6 (MW of N001)X 100% = 63%

IV. Label *N*-glycans with 2-AA and AEAB

2-AA (2-Aminobenzoic acid) was used to label **N001**, **N000**, **N111** and **N211** via reductive animation for LacZ specificity studies. AEAB (N-(aminoethyl)-2-amino benzamide) was used to label **N301**, **N300**, **N391** and **N382** as well as previously synthesized N-glycans⁶ (label data not shown) for MGAT5 specificity studies. 2-AA or AEAB reactions were performed⁸ using the following conditions: Dissolve 5 mg of 2-AA or AEAB in 100 μ L solution of DMSO/AcOH (7:3 v/v), and add 6 mg of NaCNBH₃ in the previous solution to make final label solution. To every 100 μ g glycan sample 5 μ L of labelling solution was added, and incubated in a water bath for 2 h at 65 °C. Labeled glycans were precipitated with acetonitrile (10 volumes) and incubated at -20 °C for 1 h. After centrifugation (13,000 g for 10 min), the supernatant was removed. The pellet was dried in a vacufuge for 15 min. Labeled glycan was diluted in 100 μ L of water and purified using a Hypercarb Porous Graphitic Carbon LC Column as described.⁸ Sample concentration was determined using the peak area and comparing to 10 nmol of known standard (AEAB labelled LNnT tetrasaccharide).

V. Detailed substrate specificity of MGAT5

A detailed substrate specificity study of MGAT5 was performed using 71 chemically or chemoenzymatically synthesized *N*-glycans and 4 tri-antennary *N*-glycans obtained in the study as acceptors (**Figure S4**). Reactions were performed in a 96 well PCR plate, each well (10 μ L total volume) contains one AEAB⁸ labelled *N*-glycan substrate (50 μ M), UDP-GlcNAc (10 mM), MES buffer (100 mM, pH 7.0), MgCl₂ (15 mM), 0.2% Triton X-100, BSA (1 mg/ml), and MGAT5 (0.5 mg/mL). Reactions were incubated at 37 °C for 1 h in an Eppendorf thermocycler (Mastercycler Pro), followed by enzyme inactivation at 95 °C for 5 min. A negative control was also set up using **N000** as substrate without enzyme. After centrifugation (13,000 g for 10 min), the supernatant was analyzed by HPLC using an analytical Waters XBridge BEH amide column (130 Å, 5 μ m, 4.6 × 250 mm), detected by UV detector SPD-20A (330 nm) and a fluorescence detector RF-10Axl (330 nm excitation and 420 nm emission). Gradient elution (%B: 65-50% within 30 min) was performed with Solvent A (100 mM ammonium formate, pH 3.2) and Solvent B (Acetonitrile). The conversion rates (average of three replicated reactions, see **Table S1**) were calculated as % = Product peak area/ (Product peak area + Substrate peak area) × 100.



Figure S4. MGAT5 catalyzed reaction and substrate specificity of human MGAT5 using 71 *N*-glycans as potential acceptors.

Table S1. Summary of specificity studies for MGAT5.

Glycan Substrate	Conversion (%)
N01X (9 glycans)	ND
N02X (6 glycans)	ND

N03X (6 glycans)	ND
N040	24.0 ±2.5
N04X (5 glycans)	ND
N050	45.6 ±4.1
N05X (5 glycans)	ND
N000	72.3 ±7.0
N110	ND
N210	45.5 ±6.5
N111	26.4 ±2.2
N112	10.6 ±2.1
N113	37.5 ±5.2
N114	29.7 ±5.7
N115	ND (After Overnight, 4%)
NXXX (30 glycans)	ND
N6000	61.2 ±7.5
N301	ND
N300	19.6 ± 1.0
N391	13.8 ±1.4
N382	18.4 ±3.4

The results showed that MGAT5 recognizes N-glycans with an agalactosylated and unprotected GlcNAc residue on the α 1,6-mannose branch (N000, N040, N050, and N6000), but not those either without a GlcNAc residue (N020), galactosylated (N211) or with protected GlcNAc residue (N110) on the branch. Among these N-glycans, MGAT5 showed the highest activity towards bi-antennary complex N-glycan structures terminated with GlcNAc at the non-reducing end of both α 1,3- and α 1,6-mannose branches N000 (72.3% conversion) and slightly decreased activities towards N6000 (61.2% conversion), N210 (45.5% conversion), and N050 (45.6% conversion) indicated that α 1,6-fucosylation, protection or removal of GlcNAc residue at the α 1,3-mannose branch can be well tolerated. Removal of α 1,3-mannose branch N040 (24.0% conversion), galactosylation N111 (26.4% conversion), α 1,6 sialyation N113 (37.5% conversion) and α 1,3-fucosylation N114 (29.7% conversion) will further decrease the efficiency of MGAT5. The presence of $\alpha 2,3$ sialic acid N112 (10.6% conversion) and N115 (ND), respectively is not preferred by MGAT5. MGAT5 product from Sialyl Lewis x structure N115 was only detected after overnight reaction. MGAT5 also recognized tri-antennary N-glycans (N300, N391, and N382). The enzyme did not recognize N301. Overall, the substrate specificity study using a library of structurally defined N-glycans confirmed that MGAT5 has a strict requirement towards the α 1,6-mannose branch, and a fairly relaxed preference towards the α 1,3-mannose branch.

VI. Substrate specificity studies of β-galactosidase from *E. coli* (LacZ)

Substrate specificity assay of LacZ was performed against chemo-enzymatically synthesized *N*-glycans isomers **N111** and **N211** (with 2-AA label at reducing end) as acceptors. Reactions were performed in 10 μ L total volume, containing *N*-glycan substrate (50 μ M), Tris-HCl buffer (100 mM, pH 7.5), and LacZ (1 mg/mL). Reactions were incubated at 37 °C overnight in a water bath. Samples were collected at 0, 5, 10, 15, 20, 30, 45 min, every hour for 8 h, and at 20 h. Reaction was quenched by freezing at -80 °C for 30 min. After centrifugation (13,000 g for 10 min), the supernatant was analyzed by HPLC as above described (fluorescence at 360 nm excitation and 420 nm emission). The conversion rates (average of three replicated reactions, see **Table S2** and **Figure S5**) were calculated as % = Product peak area/ (Product peak area + Substrate peak area) × 100.

	N111-AA	N211-AA
Time	Conversion (%)	Conversion (%)
0 min	ND	ND
5 min	4.1	ND
10 min	8.0	ND
15 min	9.9	ND
20 min	10.9	ND
30 min	16.7	ND
45 min	23.8	2.4
1 h	30.9	3.8
2 h	51.1	5.2
3 h	65.3	8.2
4 h	74.9	10.2
5 h	83.1	11.9
6 h	87.3	15.3
7 h	90.5	16.7
8 h	91.2	18.5
20 h	100.0	28.6

Table S2. Substrate specificity of LacZ. ND, Not Detected. Conversions were monitored by HPLC, average of three replicates were shown.



Figure S5 Time course of LacZ digestion reactions.

VII. Generation of asymmetric *N*-glycan structures using LacZ

The generation of asymmetric *N*-glycan core structures was accomplished by selective digestion of **N001 (Figure S6)** and **N301 (Figure S7)** by LacZ. Digestion reactions by LacZ were performed in 100 mM Tris-HCl buffer (pH 7.5), containing 2 mM glycan acceptor and 1 mg/mL LacZ. A reaction time course was performed for **N001** to determine the best time to collect the intermediate with 1 galactose (**Figure S6**). The percent conversion for the time course was monitor by HPLC-ELSD as previously described. ¹ Larger scale reaction reactions were allowed to proceed for 3 h for **N001** (4 mg, total reaction volume of 1.2 mL) and 3 days for **N301** (20 mg, total reaction volume of 5 mL). The reactions were then quenched by freezing in -80 °C for 30 min, followed with concentration by lyophilization. HPLC was then used to purify target glycans using a semi-preparative amide column (130 Å, 5 μ m, 10 mm × 250 mm). The running conditions are solvent A: 100 mM ammonium formate, pH 3.2; solvent B: acetonitrile; flow rate: 4 mL/min; B%: 60-45% within 25 min. Product containing fraction were then concentrated and lyophilized for characterization. The purity of each glycan was confirmed by HPLC-ELSD using an analytical Waters amide column (130 Å, 5 μ m, 4.6 mm × 250 mm).⁶ The running conditions are solvent A: 100 mM ammonium formate, pH 3.2, solvent B: acetonitrile; flow rate: 1 mL/min; B%: 65-50% within 25 min or 60-45 within 25 min.



Figure S6. LacZ digestion of **N001** over different times, reaction monitored by HPLC-ELSD. A) **N001** control. B) **N211/N111** mixture control. C) **N000** control. 0) Negative control no enzyme. 1)Reaction 0 min. 2) Reaction 1 h. 3) Reaction 2 h. 4) Reaction 3 h, 85% conversion to intermediate product. 5) Reaction 4 h. 6) Reaction 5 h. 7) Reaction 6 h. 8) Reaction 7 h. 9) Reaction 8 h. 10) Reaction overnight. The two adjacent bumps of each peak corresponding to α - and β -anomers of free reducing end *N*-glycans.



Figure S7. LacZ digestion of **N301** over 3 days, each intermediate product was purified and confirmed by MALDI-MS. I) starting material **N301**; II) Intermediate product released 1 galactose residue (**N391**); III) Intermediate product released 2 galactose residues (**N382**); IV) Completely digested product (**N300**).

VIII. Glycosyltransferase catalyzed reactions and synthesis of N-glycans

The purified N-glycan core structures were used as substrates in glycosyltransferase catalyzed reactions (Figure S8 and Figure 5).



Figure S8. Generation of more *N*-glycan cores using MGAT5 and FUT8 catalyzed reactions.

<u>FUT8 catalyzed reaction</u>: *N*-glycan substrate (5 mM), GDP-fucose (10 mM), MES buffer (100 mM, pH 7.0), 2 U FastAP, and excess FUT8 (0.5 mg/mL). Reactions were incubated at 37 °C overnight or until substrates were totally converted to product. The reactions were then quenched by boiling for 5 min, and subject to HPLC analysis and purification.

<u>MGAT5 catalyzed reaction</u>: *N*-glycan substrate (5 mM), UDP-GlcNAc (10 mM), MES buffer (100 mM, pH 7.0), MgCl₂ (15 mM), 0.2% Triton X-100, BSA (1 mg/ml), and excess MGAT5 (0.5 mg/mL). Reactions were incubated at 37 °C overnight or until substrates were totally converted to product. The reactions were then quenched by boiling for 5 min, and subject to HPLC analysis and purification.

Synthesis of N396: 4 mg of N391 was added to a 0.5 mL reaction system, containing 50 mM Tris-HCl (pH 8.0), 12 mM of GDP-Fuc, 5 mM of MnCl₂, and excess α 1,3-fucosyltransferase from *Helicobacter pylori* Hp3FT (0.2 mg/mL). FastAP (5 U/1 mL) was also added to digest the reaction byproduct GDP to drive reaction forward. Reactions incubated at 37 °C for 4-24 h, and monitored by HILIC-ELSD (amide column, 4.6 mm × 250 mm under a gradient running condition (solvent A: water, pH 3.4; solvent B: acetonitrile; flow rate: 1 mL/min; B%: 65-50% within 25 min). After over 90% acceptor converted, the reaction was quenched and N394 was purified fir next step. Next step reaction contains 50 mM Tris-HCl (pH 8.0), 4 mM of acceptor glycans (3 mg, total reaction volume of 350 µL), 8 mM of UDP-Gal, 5 mM of MnCl₂, and 20 U of bovine β 1,4-galactosyltransferase (bGalT). The reaction was allowed to proceed overnight and subject to HPLC purification to afford N-glycan N395 (3 mg, 93% yield). Finally, N395 was used as a substrate for the synthesis of N396, catalyzed by an α 2,6-sialyltransferase from *Photobacterium damselae* (Pd26ST). Reactions contain 100 mM Tris-HCl (pH 8.0), 4 mM of acceptor glycans (2.6 mg, total reaction volume of 280 µL), 8 mM of CMP-Neu5Ac, and 0.05 mg of Pd26ST. Reactions incubated at 37 °C for overnight and subject to HPLC purification (similar as above with the use of 100 mM Ammonium formate as Solvent A instead of water).

The molecular weight of each *N*-glycans was confirmed by MALDI-MS performed on UltrafleXtreme MALDI TOF/TOF Mass Spectrometer (Bruker). Scan range of MS¹ was according to the molecular weight of *N*-glycans, and reflector mode was used for *N*-glycan analysis. Mass spectra were obtained in positive mode with the following voltage settings: ion source 1 (19.0 kV), ion source 2 (15.9 kV), and lens (9.3 kV). The reflector voltage was set to 20 kV. The laser was pulsed at 7 Hz and the pulsed ion extraction time was set to 400 ns. The laser power was kept in the 25–40% range. Structures of key intermediates and asymmetric triantennary *N*-glycans were confirmed by NMR.

IX. HPLC, MALDI-MS and NMR analysis of N-glycans

All HPLC analyses were performed using the following condition with different gradient elution conditions: Waters amide column: 130 Å, 5 μ m, 4.6 mm × 250 mm; Solvent A : 100 mM Ammonium formate pH 3.2; Solvent B: Acetonitrile; Flow rate: 1 mL/min). Please note that, due to existence of interconverting α - and β -anomers of glycans with free reducing end in solution,⁹ two adjacent bumps or a large shoulder can usually be observed on the glycan peak in HPLC analysis using a HPLIC column. Such a phenomenon was found and reported repeatedly in HPLC analysis of *N*-glycans.^{1, 6, 10, 11} NMR analysis was only performed for major compounds. The structures of other N-glycans were determined by the activity of FUT8 and MGAT5 towards them.





0

1000

1200

1400



To elucidate the glycosidic linkage profile for triantennary *N*-glycan **N301**, a series of extensive NMR experiments, including ¹H NMR, ¹³C NMR, HSQC, COSY, HSQC-TOCSY, coupled-HSQC, and HMBC, were performed to enable signal assignments and observation of key correlations. In conclusion, the combination of HMBC and HSQC-TOCSY spectra indicated a number of critical three-bond C-H

1800

2000

m/z

1600

correlations, which revealed the presence of 2 antennae at C-3 branch and 1 antenna at C-6 branch with β -1,4-linkage at non-reducing terminal.

Detailed assignments:

1) The assignments of three D-Man residues: β-Man residue was identified from the smaller ${}^{1}J_{C-H}$ value (< 160 Hz) comparing to α-Man (> 170 Hz) calculated by coupled-HSQC spectrum; the resonance of α-Man on upper branch was assigned by the three-bond C-H correlation (${}^{3}J_{C-H}$ in HMBC spectrum) between H-1 Man-4' (4.82 ppm) and C-6 β-Man (65.70 ppm); similarly, the assignment of α-Man on lower branch was achieved by a ${}^{3}J_{C-H}$ correlation between H-1 Man-4 (5.01 ppm) and C-3 β-Man (80.30 ppm). In addition, the individual spin systems of the three mannose residues could be carefully traced by COSY, TOCSY, and HSQC-TOCSY spectra.

2) The elucidation of glycosidic linkages between D-GlcNAc and α -Man: Within the spin system of α -Man residue on the upper branch (Man-4'), only C-2 (76.25 ppm) was found to link with D-GlcNAc residue as evidenced by a single ${}^{3}J_{C-H}$ correlation between H-1 GlcNAc and C-2 Man-4'. However, the attachment of D-GlcNAc can be observed with both C-2 (75.95 ppm) and C-4 (78.05 ppm) on the lower α -Man residue (Man-4) through corresponding ${}^{3}J_{C-H}$ correlations, H-1 GlcNAc/C-2 Man-2 and H-1 GlcNAc/C-4 Man-2.

3) The confirmation of glycosidic linkage between D-Gal and D-GlcNAc: The linkage of D-Gal to the C4 of GlcNAc (β -1,4-linkage) was confirmed by a strong ${}^{3}J_{C-H}$ correlation between H-1 Gal (4.40 – 4.32 ppm) and C-4 GlcNAc (78.45 ppm) in the HMBC spectrum.



¹H NMR (600 MHz, D₂O) δ 5.09 (s, 0.6H, H-1 GlcNAc-1 α), 5.01 (s, 1H, H-1 Man-4), 4.82 (s, 1H, H-1 Man-4'), 4.66 (s, 1H, H-1 Man-3), 4.60 (d, J = 7.6 Hz, 0.4H, H-1 GlcNAc-1 β), 4.53 – 4.43 (m, 4H, H-1 GlcNAc-2, 5, 7, and 7'), 4.40 – 4.32 (m, 3H, H-1 Gal-6, 8, and 8'), 4.16 – 4.07 (m, 2H, H-2 Man-4 and 3), 4.01 (s, 1H, H-2 Man-4'), 3.95 (d, J = 9.3 Hz, 1H), 3.92 – 3.36 (m, 62H), 2.04 – 1.87 (m, 15H, NHAc).

HSQC spectrum for N301









usec

MHz

use use use W MHz

ers

MH: Hz pps

MHz



HSQC-TOCSY spectrum for N301



18168_01 Magodlet

145

599.84

W 169.820 6.792 Hz 37.362

150. 194.34

1024 599.8399994 QSINI (Hz

echo-antiecho 150.8298862 OSINE (Hz

1.00

Hz

58C 58C 58C 58C 58C 58C 58C 58C

MHI USEC USEC USEC USEC USEC

USEC USEC USEC

MH2 H2 DDM

8.0 10 13.996 1.433 150.84



HPLC-ELSD, $T_R = 10.06 \text{ min}$



MALDI-MS, calculated: 1843.6715; found [M+Na]⁺: 1866.6682



¹H NMR and multiple 2D-NMR experiments were performed for the structural elucidation to confirm the glycosidic linkage profile. Essentially, the combination of HMBC and HSQC-TOCSY spectra indicated the structure as the one shown in the scheme (the cleaved Gal situated on upper branch).

¹H NMR for N391



¹H NMR (600 MHz, H₂O+D₂O) δ 5.13 (s, 0.5H, H-1 GlcNAc-1α), 5.05 (s, 1H, H-1 Man-4), 4.85 (s, 1H, H-1 Man-4'), 4.70 (s, 1H, H-1 Man-3), 4.63 (d, J = 6.6 Hz, 0.5H, H-1 GlcNAc-1β), 4.58 – 4.45 (m, 4H, H-1 GlcNAc-2, 5, 7, and 7'), 4.44 – 4.35 (m, 2H, H-1 Gal-6 and 8), 4.15 (m, 2H, H-2 Man-4 and 3), 4.05 (s, 1H, H-3 Man-4), 3.99 (d, J = 8.3 Hz, 1H, H-2 Man-4'), 3.96 – 3.89 (m, 3H, H-4 Gal-6 and 8, H-2 GlcNAc-1α), 3.89 – 3.80 (m, 8H), 3.80 – 3.29 (m, 45H), 2.08 – 1.91 (m, 15H, NHAc).

HSQC spectrum for N391



HMBC spectrum for N391





ia Parameters 0-ANGIE-1-NMR 16 1 Curre NAME EXPNO PROCN 12.27 spect QXI 1H Zsec sec sec sec sec sec sec NUC1 P1 P2 PL1 PL1W SF01 W MHz NUC2 P3 PL2 PL2W SF02 69.83418 150.9178 GPNA GPNA GPNA GPZ1 GPZ2 GPZ3 F16 30.00 40.10 F1 - Ac TD SF01 FIDRES SW FnMODE F1000L F2 - Pr SI SF WDW SSB LB GB GB PC 2048 600.1300000 SINE 1.00 F1 SI SF WDW SSB LB GB sing parameters 2048 QF 150.9027565 MHz SINE 0 0 Hz

HSQC-TOCSY spectrum for N391









MALDI-MS, calculated: 1681.6187; found [M+Na]⁺: 1704.6020, and [M+K]⁺: 1720.5759









MALDI-MS, calculated: 2046.7509; found [M+Na]⁺: 2069.7421, and [M+K]⁺: 2085.7159





HPLC-ELSD, $T_R = 9.50$ min (The two adjacent peaks corresponding to α - and β -anomer)



MALDI-MS, calculated: 1884.6981; found [M+Na]⁺: 1907.6851



N6382 (HPLC Gradient B%: 65% to 50% within 25 min)



HPLC-ELSD, $T_R = 13.97$ min (The two adjacent peaks corresponding to α - and β -anomer)







HPLC-ELSD, $T_R = 12.40$ min (The two adjacent peaks corresponding to α - and β -anomer)



MALDI-MS, calculated: 1519.5659; found [M+Na]⁺: 1542.5548, and [M+K]⁺: 1558.2585





HPLC-ELSD, $T_R = 12.79$ min (The two adjacent peaks corresponding to α - and β -anomer)



MALDI-MS, calculated: 1722.6453; found [M+Na]⁺: 1745.6322, and [M+K]⁺: 1761.3360





HPLC-ELSD, $T_R = 12.92$ min (The two adjacent peaks corresponding to α - and β -anomer)



MALDI-MS, calculated: 1665.6238; found [M+Na]⁺: 1688.6191





HILIC-ELSD, $T_R = 15.67 \text{ min}$ (The two adjacent peaks corresponding to α - and β -anomer)



MALDI-MS, calculated: 1640.5922; found [M+Na]⁺: 1663.6308



See NMR analysis as previously reported (Chem. Sci., 2015, 6, 5652-5661)



HPLC-ELSD, $T_R = 13.15$ min (The two adjacent peaks corresponding to α - and β -anomer)



MALDI-MS, calculated 1478.5395; found [M+Na]⁺ 1501.5693



¹H NMR and multiple 2D-NMR experiments, including the critical HMBC and HSQC-TOCSY, were performed to elucidate the structure of N211. In addition, the structure of N211 was also confirmed by the comparison with previously reported spectrum.



¹H NMR (600 MHz, H₂O+D₂O) δ 5.12 (s, 0.6H, H-1 GlcNAc-1α), 5.04 (s, 1H, H-1 Man-4), 4.86 (s, 1H, H-1 Man-4'), 4.70 (s, 1H, H-1 Man-3), 4.63 (d, J = 7.6 Hz, 0.4H, H-1 GlcNAc-1β), 4.56 – 4.46 (m, 3H, H-1 GlcNAc-2, 5, and 5'), 4.41 (d, J = 7.5 Hz, 1H, H-1 Gal), 4.18 (s, 1H), 4.12 (s, 1H), 4.04 (s, 1H), 3.95 – 3.33 (m, 45H), 2.11 – 1.91 (m, 12H, NHAc).





HMBC spectrum for N211



HSQC-TOCSY spectrum for N211





HPLC-ELSD, $T_R = 15.25$ min



MALDI-MS, calculated: 2604.9035; found [M-H]⁻: 2603.8287



¹H NMR for N396



¹H NMR (600 MHz, D₂O) δ 5.11 (s, 0.5H, H-1 GlcNAc-1α), 5.06 - 5.00 (m, 3H, H-1 Man-4, 2*H-1 Fuc), 4.86 (s, 1H, H-1 Man-4'), 4.68 (s, 1H, H-1 Man-3), 4.61 (d, J = 7.5 Hz, 0.5H, H-1 GlcNAc-1β), 4.56 - 4.44 (m, 4H, H-1 GlcNAc-2, 5, 7, and 7'), 4.41 - 4.33 (m, 3H, H-1 Gal-6, 8, and 8'), 4.13 (m, 2H), 4.04 (m, 4H), 3.98 - 3.39 (m, 77H), 2.61 (dd, J = 12.2, 4.3 Hz, 1H, Neu5Gc H-3e), 2.07 - 1.92 (m, 15H, NHAc), 1.66 (t, J = 12.1 Hz, 1H, Neu5Gc H-3a), 1.09 (d, J = 5.7 Hz, 6H, Fuc-CH₃).

HSQC spectrum of N396



VIII. References

- 1. A. D. Calderon, Y. Liu, X. Li, X. Wang, X. Chen, L. Li and P. G. Wang, *Org Biomol Chem*, 2016, **14**, 4027-4031.
- 2. G. Zhao, W. Guan, L. Cai and P. G. Wang, *Nat Protoc*, 2010, **5**, 636-646.
- 3. H. Yu, H. A. Chokhawala, S. Huang and X. Chen, *Nat Protoc*, 2006, **1**, 2485-2492.
- 4. D. J. Wasilko, S. E. Lee, K. J. Stutzman-Engwall, B. A. Reitz, T. L. Emmons, K. J. Mathis, M. J. Bienkowski, A. G. Tomasselli and H. D. Fischer, *Protein Expr Purif*, 2009, **65**, 122-132.
- 5. H. Ihara, Y. Ikeda and N. Taniguchi, *Glycobiology*, 2006, **16**, 333-342.
- L. Li, Y. Liu, C. Ma, J. Qu, A. D. Calderon, B. Wu, N. Wei, X. Wang, Y. Guo, Z. Xiao, J. Song, G. Sugiarto, Y. Li, H. Yu, X. Chen and P. G. Wang, *Chem Sci*, 2015, 6, 5652-5661.
- 7. E. D. Green, G. Adelt, J. U. Baenziger, S. Wilson and H. Van Halbeek, *J Biol Chem*, 1988, **263**, 18253-18268.
- 8. X. Song, B. Xia, S. R. Stowell, Y. Lasanajak, D. F. Smith and R. D. Cummings, *Chem Biol*, 2009, **16**, 36-47.
- 9. C. R. Bertozzi and D. Rabuka, in *Essentials of Glycobiology*, eds. A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart and M. E. Etzler, Cold Spring Harbor (NY), 2nd edn., 2009.
- 10. I. A. Gagarinov, T. Li, J. S. Torano, T. Caval, A. D. Srivastava, J. A. Kruijtzer, A. J. Heck and G. J. Boons, *J Am Chem Soc*, 2017, **139**, 1011-1018.
- 11. T. Li, M. Huang, L. Liu, S. Wang, K. W. Moremen and G. J. Boons, *Chem Eur J*, 2016, **22**, 18742-18746.