Electronic Supplementary Material (ESI) for ChemComm.
This journal is © The Royal Society of Chemistry 2021

Supplementary Information

Linkage-selective derivatization for glycosylation site- and glycoform-specific characterization of sialic acid isomers by mass spectrometry

Ye Peng[a,b], Bing Gu[c,*], Zhenyu Sun[a], Yueyue Li[a], Ying Zhang[a,b,*], Haojie Lu [a,b,*]

[a] Institutes of Biomedical Sciences and Shanghai Cancer Center, Fudan University. Shanghai 200032 (China).

[b] Department of Chemistry and NHC Key Laboratory of Glycoconjugates Research, Fudan University. Shanghai 200032 (China).

[c] Laboratory Medicine of Guangdong Provincial People's Hospital and Guangdong Academy of Medical Sciences. Guangzhou Guangdong 510000 (China).

Corresponding Authors:

* Ying Zhang, Phone: (+) 86-21-54237618. Fax: (+) 86-21-54237961. E- mail: ying@fudan.edu.cn

* Haojie Lu, Phone: (+) 86-21-54237618. Fax: (+) 86-21-54237961. E- mail: luhaojie@fudan.edu.cn

* Bing Gu, Phone: (+) 86-20-83827812. Fax: (+)86-20-83827812. E-mail: gb20031129@163.com

CONTENTS

MATERIALS AND METHODS......3
Materials and reagents. 3
N-glycopeptides purification using crystallized cellulose......3
Two-step sequential methylamidation for sialic acid linkage differentiation. 3
IgG extraction from human serum samples. 4
Enzymatic digestion of standard glycoproteins and extracted IgG. 4
MALDI-MS analysis. 4
LC-ESI-MS/MS Analysis......5
Data analysis. 5
RESULTS AND DISCUSSION 6
Correction the overlap of isotopic clusters 6
Supporting Figures 8
Fig. S1 MALDI-MS spectra of 3-SGP (a) and 6-SGP (b) after weak-activation methylamidation. The black circles denote side products resulting from dehydration.. 8
Fig. S2 MALDI-MS/MS spectra of precursor with m/z=2864.17 (a) and m/z=3019.31 (b), recorded from the same sample of Figure S1 9
Fig. S3 MALDI-MS spectra of 3-SGP (a) and 6-SGP (b) after weak-activation methylamidation performed using DMSO as solvent......10
Fig. S4 MALDI-MS spectra of 3-SGP (a) and 6-SGP (b) after weak-activation methylamidation using d0-MA·HCl performed under optimized derivatization condition. The optimized reaction solution is composed of 750 mM EDC·HCl, 250 mM HOBt and 2 M d0/d3-MA·HCl, with water as solvent. The reaction was performed at 37 °C for 2 h.. 11
Fig. S5 Zoomed MALDI-MS spectra of 3-SGP and 6-SGP mixed at ratio of 10:1 to 1:10, after two-step forward methylamidation (a). The result of linear fitting of expected ratio of 3-SGP/6-SGP vs detected ratio of that during forward (b) and reverse (c) labeling......12
Fig. S6 Representative ESI-MS spectrum of N-glycopeptide EEQYNSTYR-G2FS, obtained during analysis of reverse labelled standard human IgG. The black circle and square denote monoisotopic peak of α-2,3-linked and α-2,6-linked EEQYNSTYR-G2FS, respectively. 12
Fig. S7 ESI-MS/MS spectra of N-glycopeptide TTCHVLDPTPLAN#CSVR-H5N4S2 (a) and SN#GSYLQLVE-H6N5S4 (b) from heavy-light methylamidated bovine fetuin. 13
Fig. S8 SDS-PAGE of standard IgG and IgG extracted from human serum (3 replicates), stained by Coomassie Brilliant Blue 14
Supporting Tabels......15
Table S1 Clinical features of the 10 liver cancer patients and 10 healthy individuals enrolled in this study. 15
Table S2 N-glycopeptides detected from standard human IgG 15
Table S3 N-glycopeptides detected from bovine fetuin 17
REFERENCES......20

MATERIALS AND METHODS

Materials and reagents.

Methylamine hydrochloride (d0-MA·HCl), 1-(3-(dimethylamino)propyl)-3-ethylcarbodiimide hydrochloride (EDC·HCl) and 1-hydroxybenzotriazole monohydrate (HOBt) were purchased from Shanghai Aladdin Bio-Chem Technology Co., LTD (Shanghai, China). Dimethyl sulfoxide (DMSO) is purchased from Sinopharm Chemical Reagent Co., Ltd. (Shanghai, China). Methyl-d3-amine hydrochloride (d3-MA·HCl), (7-azabenzotriazol-1-yloxy)trispyrrolidinophosphonium hexafluorophosphate (PyAOP), N-methylmorpholine (NMM), bovine fetuin (fetuin), asialofetuin from bovine pancreas (ASF), trypsin from porcine pancreas, Glu-C from Staphylococcus aureus V8, protein G agarose, Roche cOmplete™ Protease Inhibitor Cocktail tablets, sodium hydroxide (NaOH), ammonium bicarbonate (ABC), urea, super DHB (mixture of 2,5-dihydroxybenzoic acid and 2-hydroxy-5-methoxybenzoic acid), triethylammonium bicarbonate buffer (TEAB, 1 M aqueous solution), RIPA Lysis Buffer, 10X, phosphate buffer saline (PBS) 1X, dithiothreitol (DTT), iodoacetamide (IAA), crystallized cellulose and trifluoroacetic acid (TFA) were purchased from Sigma-Aldrich (MO, USA). α-2,6-sialylglycopeptide (6-SGP) was purchased from Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan). HPLC-grade methanol and acetonitrile (ACN) were purchased from Merck (Darmstadt, Germany). α-2,3-sialylglycopeptide (3-SGP) was purchased from Fushimi Pharmaceutical Co., Ltd. (Kagawa, Japan). HyperSepTM HypercarbTM porous graphitized carbon (PGC) micro-column and Pierce™ BCA Protein Assay Kit were purchased from Thermo Fisher (CA, USA). Sep-Pak Vac C18 cartridges were purchased from Waters (MA, USA). Centrifugal filters with MWCO of 3 kDa and 10 kDa were purchased from Millipore (MA, USA). Centrifugal spin column with 0.22 μm acetate cellulose filter was purchased from Agilent (CA, USA). Distilled water was purified by a Milli-Q system (MA, USA).
Serum samples of healthy individuals and liver cancer patients were obtained from Guangdong Provincial People's Hospital and The Affiliated Hospital of Xuzhou Medical University. The clinical features of these samples can be referred to Table S1. The research was handled under an institutional review board-approved protocol with ethical and legal standards.

N-glycopeptides purification using crystallized cellulose.

Six mg of crystallized cellulose is added into a 0.6 mL Eppendorf tube and washed by water for twice, 5 min each. Then the cellulose powder is equilibrate using 83% (v/v) ACN with 1% TFA. The samples to be purified is also adjusted to 83% (v/v) ACN with 1% TFA and added into the cellulose suspension, followed by gentle shaking for 30 min. A small piece of cotton wool (~0.5 mg) is packed at the end of a 10 μL pipet tip to support the cellulose powder. Then, the tip was set upon a 2.0 mL Eppendorf tube. The cellulose suspension is pipetted onto the tip. A one-minute interval was set between adding solution and centrifugation at 10,000 g for 1 min to allow efficient soaking of cellulose during every operation. After loading, the tip is washed four times with 83% (v/v) ACN with 1% TFA. Finally, the glycopeptides are eluted three times using 0.1% TFA.

Two-step sequential methylamidation for sialic acid linkage differentiation.

The N-glycopeptide was lyophilized. 10 μL of reaction solution (750 mM EDC·HCl, 250 mM HOBt and 2 M d0/d3-MA·HCl, in water) was added. The mixture was incubated at 37 °C for 2 h with gentle shaking. The reaction was quenched by adding 50 μL of ACN, and 260 μL of 83% (v/v) ACN with 1% TFA. The solution was then put through cellulose purification.

The purified N-glycopeptide from previous reaction was lyophilized. Ten μL of 5 M d0/d3-MA·HCl (in 1 M TEAB) was added. The mixture was incubated at 37 °C for 1 h with gentle shaking, followed by lyophilization. Then 20 μL of 125 mM PyAOP (in 15% (v/v) NMM/DMSO) was added. The mixture was incubated at 37 °C for 1 h. The reaction was quenched by adding 100 μL of ACN and 280 μL of 83% (v/v) ACN with 1% TFA. The solution was then put through cellulose purification.

IgG extraction from human serum samples.

The purification of IgG from human serum was performed using protein G agarose. Forty μL of protein G agarose suspension was added into a spin column with 0.22 μm acetate cellulose filter, followed by centrifugation to remove the solvent. Then, 200 μL of RIPA buffer with protease inhibitor cocktail (prepared by adding 1 tablet of Roche cOmplete™ Protease Inhibitor Cocktail tablet for every 10 mL of 1X RIPA buffer) was added, followed by gentle shaking at room temperature for 5 min and centrifugation to remove the liquid. The agarose was washed again using the process above. Ten μL of serum sample to be extracted was mixed with 200 μL of RIPA buffer with protease inhibitor cocktail, and added into the spin column. The spin column was incubated at room temperature for 4 h with gentle shaking. The liquid was removed by centrifugation. Then, the agarose was washed four times using 200 μL of RIPA buffer (without protease cocktail) and twice using 200 μL 1X PBS, each with gentle shaking at room temperature for 5 min. The captured IgG was eluted by 100 μL 0.1% TFA for three times. The eluents were combined, neutralized using 5 μL of 1 M ABC, and put through BCA assay quantification and subsequent digestion.

Enzymatic digestion of standard glycoproteins and extracted IgG.

For tryptic digestion, the standard IgG, fetuin and ASF in powder form were dissolved in 50 mM ABC at final concentration of 1 mg/mL. The eluent containing IgG extracted from human serum was diluted 2 times using 100 mM ABC.

Before digestion, the protein solution was put through reductive alkylation. First, 1 M DTT solution was added to a final concentration of 10 mM, followed by water bath at 56 °C for 30 min. Then, after the solution is cooled to room temperature, 1 M IAA solution was added to a final concentration of 25 mM. The mixture was kept in dark for 30 min. After reductive alkylation, trypsin was added at a mass ratio of 1:20 (trypsin: protein), followed by incubation at 37 °C for 16 h. For tryptic digestion of IgG, the solution containing digested peptides was acidified using 10% TFA and desalted using C18 cartridges. The C18 cartridge was activated using 1 mL ACN, washed by 1 mL 0.1% TFA. The sample was loaded, and the flow-through was collected and re-loaded, for three times. Then, the cartridge was washed by 1 mL of 0.1% TFA for three times. The captured peptides were eluted using 400 μL of 40% (v/v) ACN with 0.1% TFA. For tryptic and Glu-C digestion of fetuin and ASF, Glu-C was added into the solution containing tryptic digested peptides at a mass ratio of 1:20 (Glu-C: protein), followed by incubation at 37 °C for 16 h. After digestion, the solution was acidified using 10% TFA, and desalted using C18 cartridges as described above.

MALDI-MS/MS analysis.

All MALDI-MS analysis were performed on a 5800 TOF/TOF analyzer (AB SCIEX, MA, USA) equipped with a pulsed Nd:YAG laser at an excitation wavelength of 355 nm, repetition rate of 400 Hz, and acceleration voltage of 20 kV. Samples were spotted on a MALDI target plate and allowed to air dry. Then 1 μL of Super-DHB matrix solution (10 mg/mL, in 50% ACN with 0.1 mM NaOH or 0.1% TFA) or CHCA matrix solution (5 mg/mL, in 50% ACN with 0.1% TFA) was deposited on the same position and dried before plate loading. MALDI-MS spectra were analyzed by Data Explorer™ software. Peak list was generated automatically by the same software with the

following peak detection parameter: 50% centroid, S/N threshold=3, noise window width=250. The intensity and S/N values of peaks were all retrieved from the peak list.

For MALDI-MS/MS analysis, ambient air was utilized as the collision gas with medium pressure of 10^{-6} Torr. Energy of 1 kV was used for collision-induced dissociation (CID), and 2000 acquisitions were accumulated for each MS/MS spectrum. The peak detection criteria used were a minimum S/N of 3, local noise window width (m/z) of 200, and minimum full-width half-maximum (bins) of 2.9.

LC-ESI-MS/MS Analysis.

LC-ESI-MS/MS analysis was performed on an Orbitrap Fusion Tribrid (Thermo Scientific) coupled with an EASY-nano-LC system (Thermo Scientific) without trap column. The N-glycopeptides were separated by an Acclaim PepMapTM RSLC C18 column (75 µm x 25 cm, nanoViper) with a flow rate of 300 nL/min. Solvent A was water containing 0.1% formic acid. Solvent B was acetonitrile containing 0.1% formic acid. The gradient was 60 min in total: 0-45 min, 1%-30% B; 45-50 min, 30%-50% B; 50-54 min, 50%-90% B; 54-60 min, 90% B. The MS parameters for N-glycopeptide analysis were set as follow. MS1: scan range (m/z) = 350-2000; resolution = 30 k; AGC target = 400,000; maximum injection time = 50 ms; dynamic exclusion after n times, n = 1; dynamic exclusion duration = 20 s; each selected precursor was subject to one HCD-MS/MS. MS/MS: isolation window = 1.8; detector type = Orbitrap; resolution = 30 k; AGC target =500,000; maximum injection time = 250 ms; HCD collision energy = 30% (NCE).

Data analysis.

XCalibur 4.0 was used for data processing. Raw data was used directly without any further processing for automated N-glycopeptide identification using Byonic. The following parameters were used for IgG search: mass tolerance for precursors and fragment ions were set as ± 10 ppm. and ± 20 ppm., respectively. The enzyme was full-trypsin and maximal missed cleavage was 0. Cysteine carbamidomethylation (+57.021 Da), aspartic acid, glutamic acid (non-N-terminal), and C-Terminal methylamidation (+13.032 Da for d0-methylamine, +16.050 Da for d3-methylamine) were set as fixed modification. Variable modifications contained oxidation on Met (M +15.995 Da, rare 2), pyroglutamate conversion on N-terminal glutamic acid (N-term E, -18.011 Da, common 1) and N-glycosylation (common 1), FDR (False discovery rate) < 1%. The IgG glycan database was modified from previous report[1], with the mass shift of d0- and d3-methylamidation for each sialic acid added. The protein databases contain all four IgG subclasses: IgG1 (UniProt ID, P01857), IgG2 (P01859), IgG3 (P01860) and IgG4 (P01861). For search of fetuin and ASF, the parameters were same except the following modifications: enzyme was set as combination of trypsin and Glu-C, the maximal missed cleavage was 1, the N-glycan database was modified from the built-in database (N-glycan 182 human no multiple fucose, with the mass shift of d0- and d3-methylamidation for each sialic acid added) and the protein database was derived from Uniprot (Uniprot ID, P12763). The N-glycopeptides identified were filtered to 1% FDR and Byonic score ≥150. For relative quantification analysis of IgG N-glycopeptide from human serum, the m/z and retention time window of non-sialylated N-glycopeptide or that containing 2,6-linked sialic acid were obtained from Byonic search result, which were utilized to extract their peak area. The peak area of N-glycopeptide containing 2,3-linked sialic acid was extracted according to the calculated theoretical m/z, with similar retention time window (within 0.15 min) of its 2,6-linked counterpart.

Deconvolution and interpretation of isotopic clusters.

To obtain the relative abundance of each linkage isomer with the same composition, following procedures were performed. Briefly, for a peak cluster annotated with a glycopeptide composition, the chemical composition of the

native glycan was retrieved from Glyco Mass Calculator (National Institute of Standards and Technology). Then, the chemical composition and mass of each linkage isomer after derivatization were calculated by adding the mass of glycan, peptide backbone and derivatization tags. The signal intensity of peaks within the isotope envelop was extracted for further use. Isotopic cluster pattern and abundance distribution of each linkage isomer was obtained from enviPat Web 2.4[2]. The relative abundance of each linkage isomer was calculated according to the abundance distribution and the signal intensity of each peak. A detailed example can be seen in Supplementary Information.

RESULTS AND DISCUSSION

Correction the overlap of isotopic clusters

Due to the existence of naturally-occur isotopes like ^{13}C and ^{15}N, analytes usually appear as a cluster of peaks called isotopic peaks. In this isotope-coded derivatization method, there is 3 Da difference between light- and heavy-labelling reagents we used. So, it can be expected that some of the isotopic peaks of light-labelled analyte will overlap with that of heavy-labelled one. Such overlapping will lead to inaccuracy during relative quantification. Thus, correction of the overlap should be performed. Taking the tri-antenna tri-sialylated N-glycans with the composition of Hex5HexNAc4Sia2 (H5N4Sia2 for short), attached to peptide with sequence of EEQYNSTYR for example, the chemical composition for un-derivatized N-glycopeptide is $C_{140}H_{218}N_{20}O_{85}$. After heavy-light linkage-specific isotopic methylamidation, all 2,3-linked sialic acids (^{3}Sia for short) are labelled with d0-methylamine and 2,6-linked ones (^{6}Sia for short) are labelled with d3-methylamine. For the peptide backbone, the N-terminal glutamic acid turns to pyroglutamic acid. The carboxylic acids at the side chain and the C-terminal are labelled with d3-methylamine. The chemical compositions of three linkage isomers after derivatization are: H5N4^{3}Sia2: $C_{144}H_{216}N_{24}O_{80}D_{6}$; H5N4^{6}Sia1^{3}Sia1: $C_{144}H_{216}N_{24}O_{80}D_{9}$; H5N4^{6}Sia2:$C_{144}H_{216}N_{24}O_{80}D_{12}$. The exact mass of each compositions is: 3579.49 Da, 3852.51 Da and 3585.53 Da.

The chemical compositions are input into enviPat Web 2.4 to obtain the isotopic distribution of each isomer in MS. The parameters are: adducts: positive-M+3H; output: centroid. The relative abundance distribution of each isomers is listed below (peaks with less than 1% intensity of that of the base peak are ignored).

For H5N4^{3}Sia2, the monoisotopic peak is 1194.17 Da. The relative intensity of each peak is 58.76% (+0), 100.00% (+1), 94.26% (+2), 63.84% (+3), 34.46% (+4), 15.67% (+5), 6.18% (+6) and 2.16% (+7).

For H5N4^{6}Sia1^{3}Sia1, the monoisotopic peak is 1195.17 Da. The relative intensity of each peak is 58.74% (+0), 100.00% (+1), 94.27% (+2), 63.86% (+3), 34.48% (+4), 15.67% (+5), 6.19% (+6) and 2.17% (+7).

For H5N4^{6}Sia2, the monoisotopic peak is 1196.18 Da. The relative intensity of each peak is 58.73% (+0), 100.00% (+1), 94.29% (+2), 63.88% (+3), 34.49% (+4), 15.68% (+5), 6.19% (+6) and 2.17% (+7).

Assume that the total actual peak intensity of each isomer to be A_1 (H5N4^{3}Sia2), A_2 (H5N4^{6}Sia1^{3}Sia1), and A_3 (H5N4^{6}Sia2).

To increase the quantification sensitivity and simplify calculation, the signal intensity of the first three isotope peaks (+0, +1, +2) are extracted from peak list as I_{1-1}, I_{1-2}, I_{1-3}, I_{2-1}, I_{2-2}, I_{2-3}, and I_{3-1}, I_{3-2}, I_{3-3}.

For isomer 1 (H5N4^{3}Sia2), as the ions are distributed among the isotopic envelope and the three peaks are not interfered by other isomers, after normalization, there is:

$$A_1 = \frac{58.76 + 100 + 94.26 + 63.84 + 34.46 + 15.67 + 6.18 + 2.16}{58.76 + 100 + 94.26}(I_{1-1} + I_{1-2} + I_{1-3}) = \frac{375.33}{253.02}(I_{1-}$$
$$)$$

For isomer 2 (H5N4[6]Sia1[3]Sia1), considering that the first three isotopic peaks of it are overlapped by the +3, +4 and +5 isotopic peaks of isomer1, there is:

$$A_2 = \frac{58.74 + 100 + 94.27 + 63.86 + 34.48 + 15.67 + 6.19 + 2.17}{58.74 + 100 + 94.27}((I_{2-1} + I_{2-2} + I_{2-3} - \frac{63.84 + 34.4}{58.76 + 10} \; + I_{1-2} + I_{1-3})) = \frac{375.38}{253.01}((I_{2-1} + I_{2-2} + I_{2-3} - \frac{113.97}{253.02}(I_{1-1} + I_{1-2} + I_{1-3}))$$

For isomer 3 (H5N4[6]Sia2), whose first three isotopic peaks are overlapped with the +6 and +7 isotopic peak of isomer 1, and the +3, +4 and +5 isotopic peaks of isomer 2, there is:

$$A_3 = \frac{58.73 + 100 + 94.29 + 63.88 + 34.49 + 15.68 + 6.19 + 2.17}{58.73 + 100 + 94.29}((I_{3-1} + I_{3-2} + I_{3-3} - \frac{63.86 + 34.4}{58.74 + 10} \; - \frac{6.18 + 2.16}{58.76 + 100 + 94.26}(I_{1-1} + I_{1-2} + I_{1-3})) = \frac{375.43}{253.02}((I_{3-1} + I_{3-2} + I_{3-3}) - \frac{114.01}{235.01}(I_{2-1} + I_{2-2} + I_{2-3} - \frac{113.97}{235.02}(I_{1-1} + I_{1-2} + I_{1-3})) - \frac{8.34}{253.02}(I_{1-1} + I_{1-2} + I_{1-3}))$$

Because the three linkage isomers are also isotopic isomers, the ionization efficiency of them can be assumed to be the same during MS-analysis. Thus, the abundance ratio of these isomers equals the ratio of their total actual peak intensity. Through calculating A_1~A_3 from $I_{1\text{-}1}$~$I_{3\text{-}3}$, which can be read from the peak list, using equations listed above, the relative abundance of the linkage isomers can be obtained. And because the coefficient among the linkage isomers with the same composition is quite similar (with variation of around 0.02%), they are omitted during calculation of relative abundance for simplification. For example, the relative abundance of isomer 1 (R_1 for short) would be:

$$R_1 = \frac{A_1}{A_1 + A_2 + A_3} = \frac{\frac{375.33}{253.02}(I_{1-1} + I_{1-2} + I_{1-3})}{\frac{375.33}{253.02}(I_{1-1} + I_{1-2} + I_{1-3}) + \frac{375.38}{253.01}((I_{2-1} + I_{2-2} + I_{2-3} - \frac{113.97}{253.02}(I_{1-1} + I_{1-2} + I_{1-3})) + \frac{375.43}{253.02}((I_{3-1} + I_{3-2} + I_{3-3}) - \frac{114.01}{235.01}(\cdots}$$

$$\approx \frac{(I_{1-1} + I_{1-2} + I_{1-3})}{(I_{1-1} + I_{1-2} + I_{1-3}) + ((I_{2-1} + I_{2-2} + I_{2-3} - \frac{113.97}{253.02}(I_{1-1} + I_{1-2} + I_{1-3})) + ((I_{3-1} + I_{3-2} + I_{3-3}) - \frac{114.01}{235.01}(I_{2-1} + I_{2-2} + I_{2-3} - \frac{113.97}{235.02}(I_{1-1} \cdots}$$

Supporting Figures

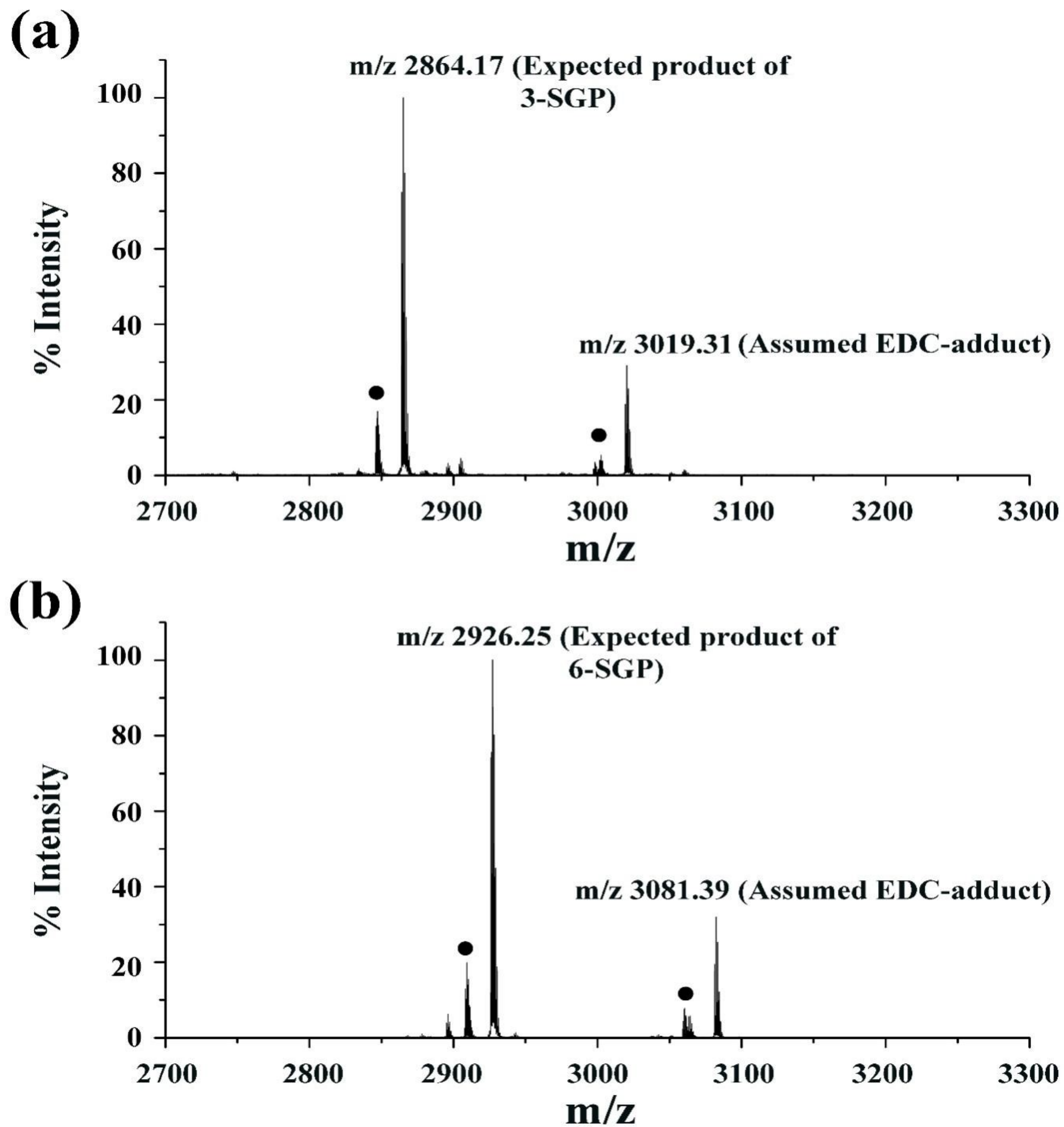


Fig. S1 MALDI-MS spectra of 3-SGP (a) and 6-SGP (b) after weak-activation methylamidation. The black circles denote side products resulting from dehydration.

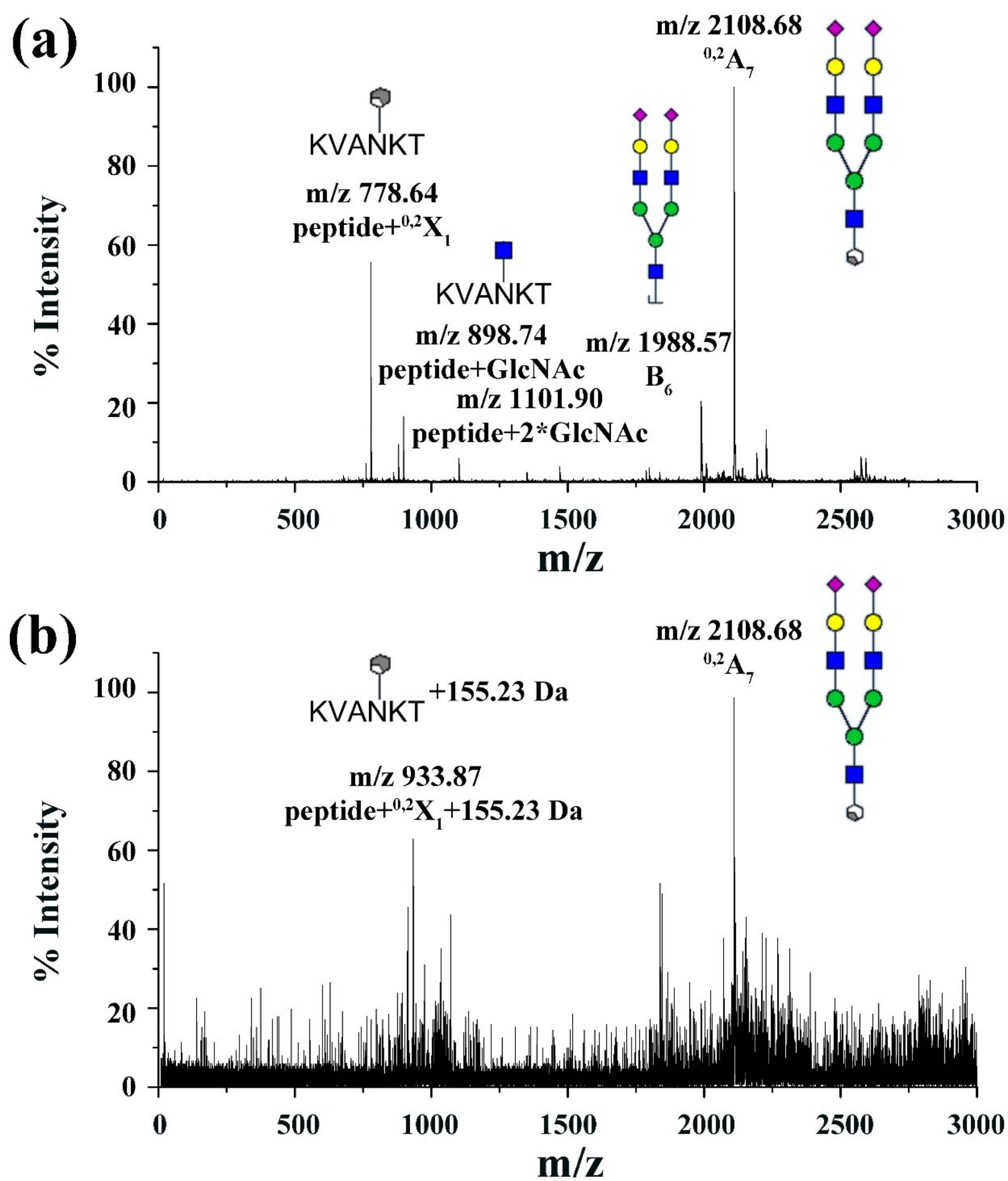


Fig. S2 MALDI-MS/MS spectra of precursor with m/z=2864.17 (a) and m/z=3019.31 (b), recorded from the same sample of Figure S1

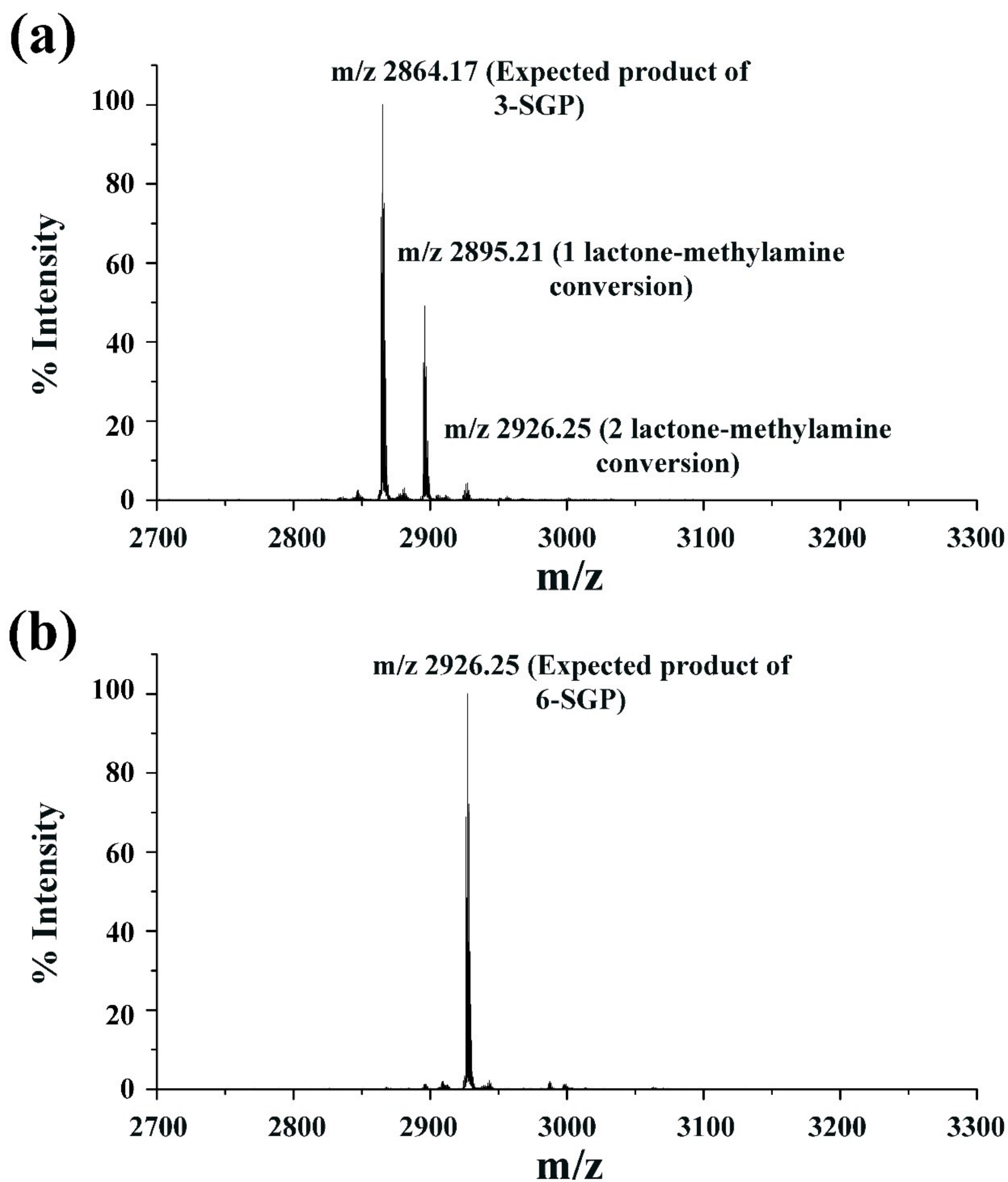


Fig. S3 MALDI-MS spectra of 3-SGP (a) and 6-SGP (b) after weak-activation methylamidation performed using DMSO as solvent.

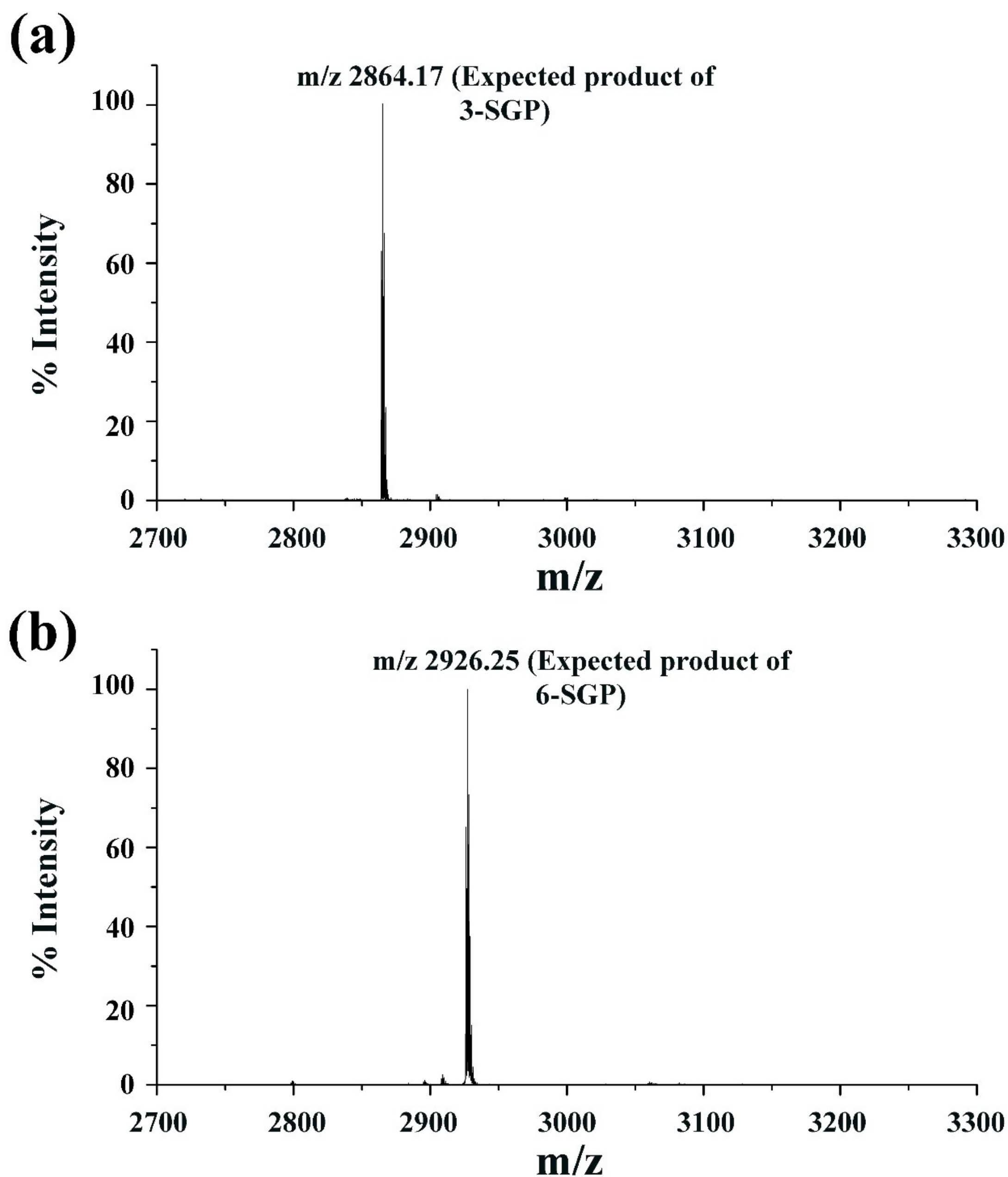


Fig. S4 MALDI-MS spectra of 3-SGP (a) and 6-SGP (b) after weak-activation methylamidation using d0-MA·HCl performed under optimized derivatization condition. The optimized reaction solution is composed of 750 mM EDC·HCl, 250 mM HOBt and 2 M d0/d3-MA·HCl, with water as solvent. The reaction was performed at 37 °C for 2 h.

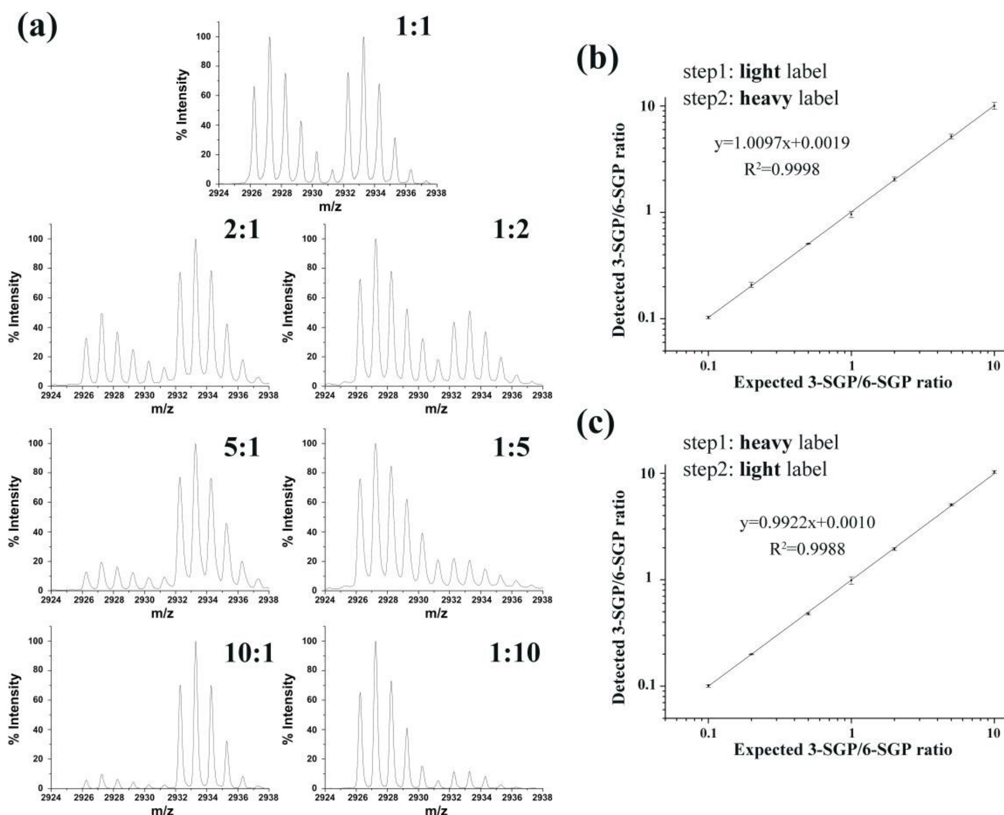


Fig. S5 Zoomed MALDI-MS spectra of 3-SGP and 6-SGP mixed at ratio of 10:1 to 1:10, after two-step forward methylamidation (a). The result of linear fitting of expected ratio of 3-SGP/6-SGP vs detected ratio of that during forward (b) and reverse (c) labeling.

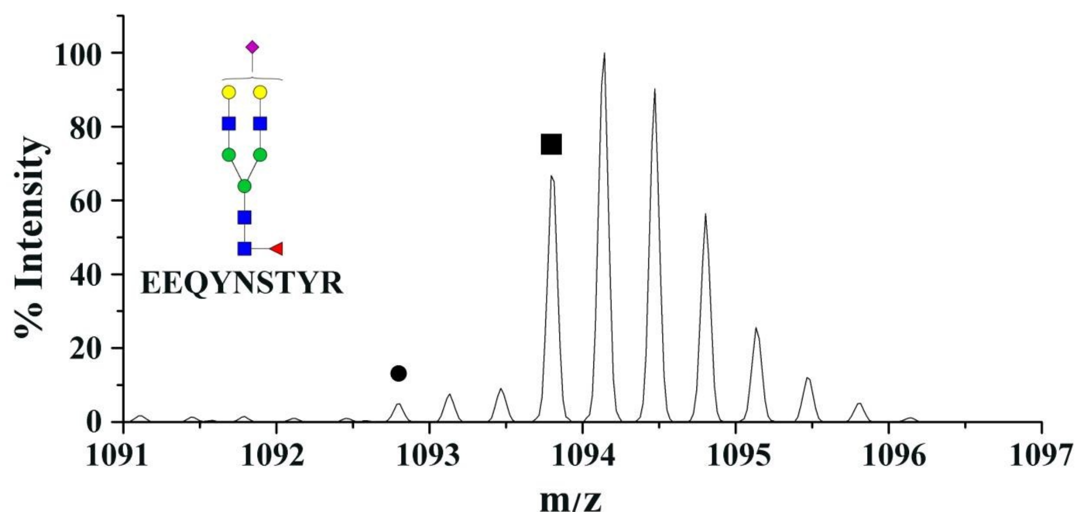
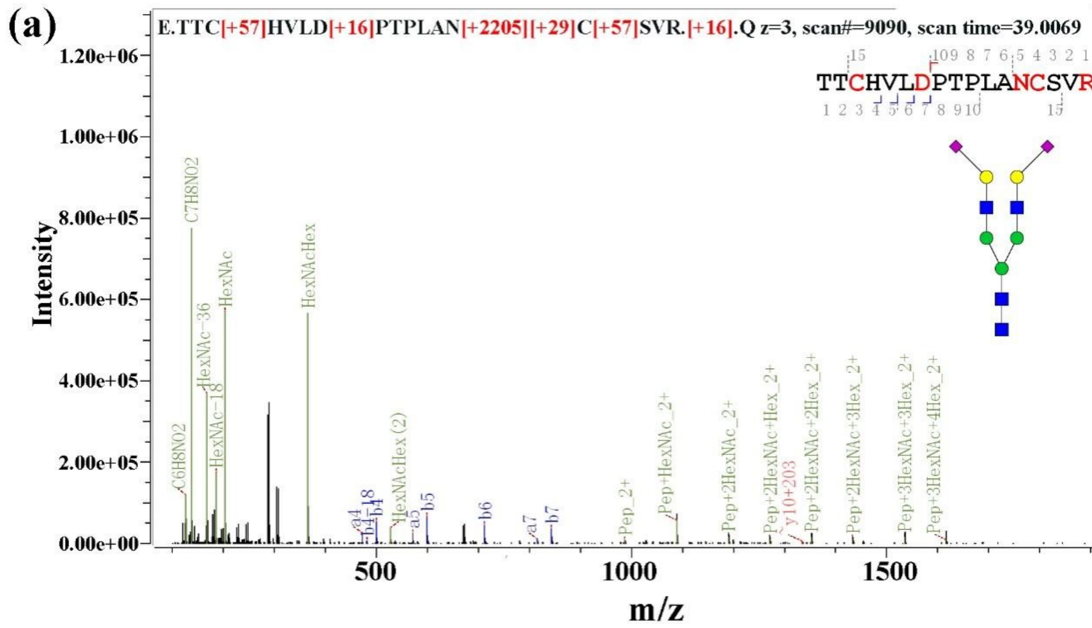


Fig. S6 Representative ESI-MS spectrum of N-glycopeptide EEQYNSTYR-G2FS, obtained during analysis of reverse labelled standard human IgG. The black circle and square denote monoisotopic peak of α-2,3-linked and α-2,6-linked EEQYNSTYR-G2FS, respectively.



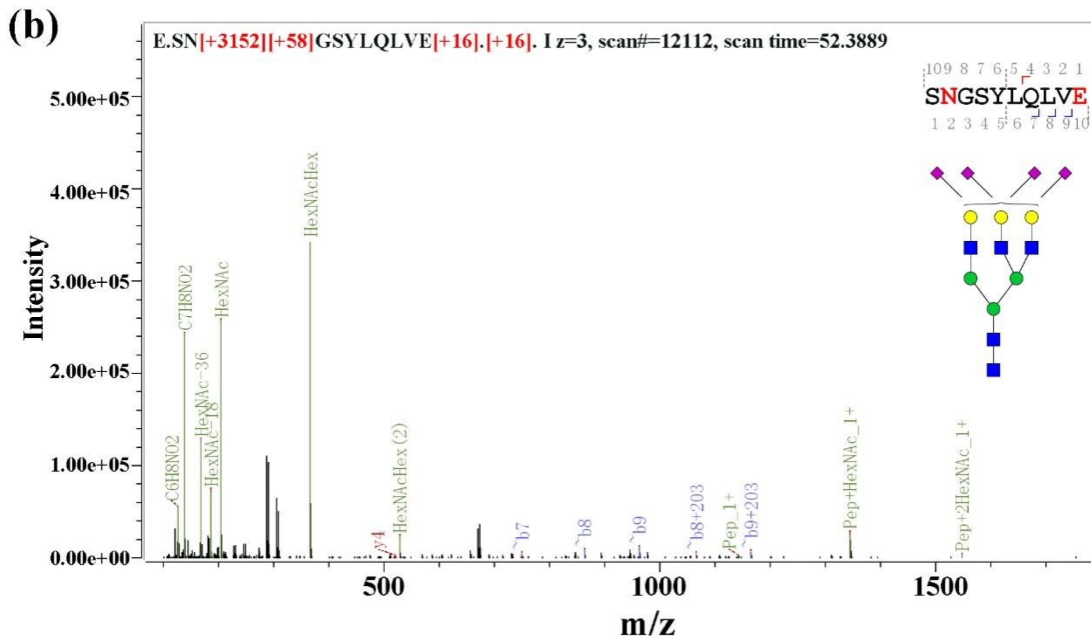


Fig. S7 ESI-MS/MS spectra of N-glycopeptide TTCHVLDPTPLAN#CSVR-H5N4S2 (a) and SN#GSYLQLVE-H6N5S4 (b) from heavy-light methylamidated bovine fetuin.

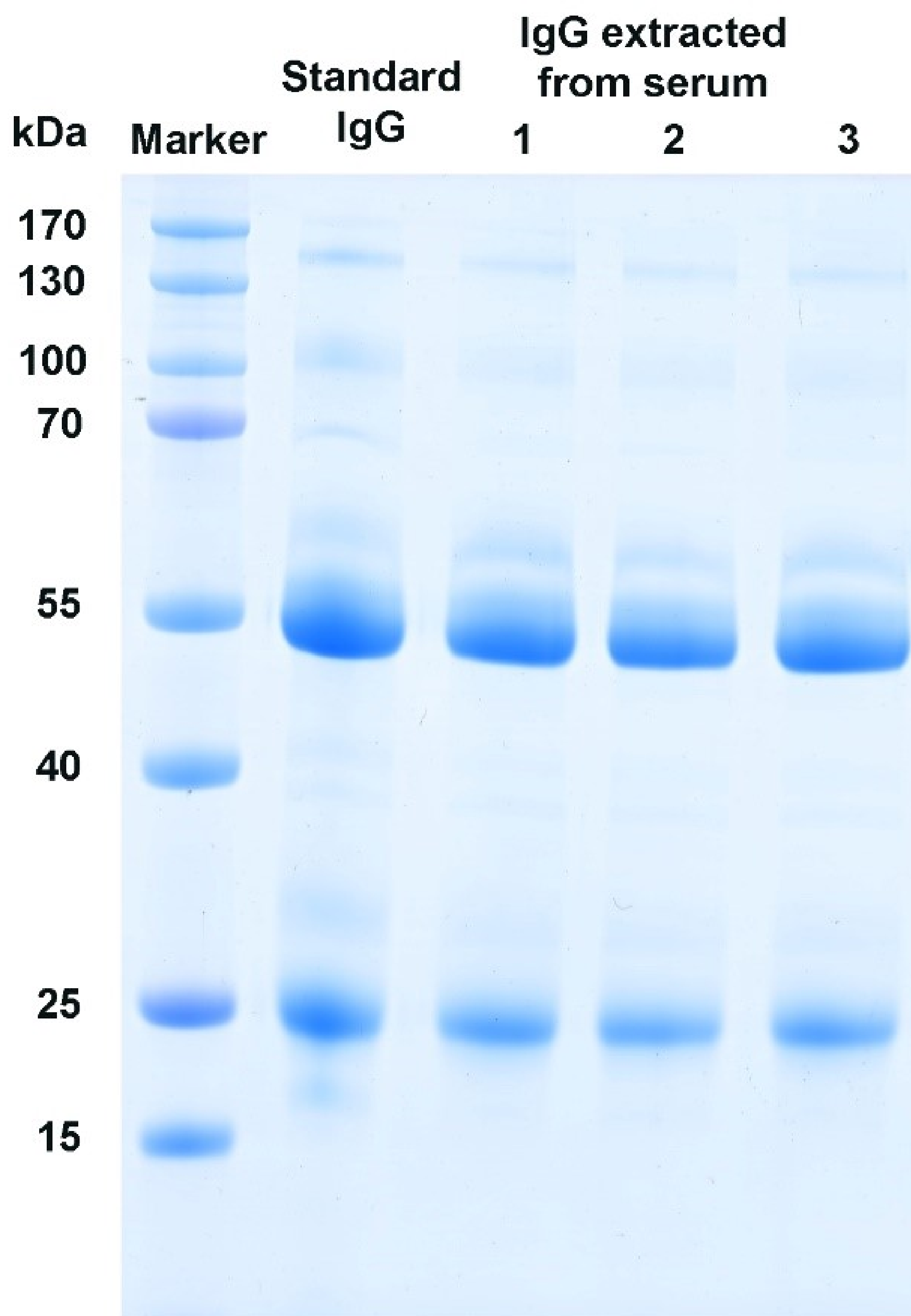


Fig. S8 SDS-PAGE of standard IgG and IgG extracted from human serum (3 replicates), stained by Coomassie Brilliant Blue

Supporting Tabels

Table S1 Clinical features of the 10 liver cancer patients and 10 healthy individuals enrolled in this study.

Characteristic	Data
Age in years (median, range)	
Liver cancer patient	57.5, 35-85
Healthy individuals	41, 7-64
Gender (female, male)	
Liver cancer patient	1, 9
Healthy individuals	5, 5

Table S2 N-glycopeptides detected from standard human IgG

No.	Protein Name	Glycans NHFAGNa*
1	IgG1	HexNAc(3)Hex(3)Fuc(1)
2	IgG1	HexNAc(3)Hex(3)2,3NeuAc(1)
3	IgG1	HexNAc(3)Hex(4)Fuc(1)
4	IgG1	HexNAc(3)Hex(6)2,6NeuAc(1)
5	IgG1	HexNAc(3)Hex(7)Fuc(1)
6	IgG1	HexNAc(4)Hex(3)
7	IgG1	HexNAc(4)Hex(3)Fuc(1)
8	IgG1	HexNAc(4)Hex(3)2,3NeuAc(1)
9	IgG1	HexNAc(4)Hex(3)2,6NeuAc(1)
10	IgG1	HexNAc(4)Hex(4)
11	IgG1	HexNAc(4)Hex(4)Fuc(1)
12	IgG1	HexNAc(4)Hex(4)Fuc(1)2,6NeuAc(1)
13	IgG1	HexNAc(4)Hex(4)2,6NeuAc(1)
14	IgG1	HexNAc(4)Hex(4)2,6NeuAc(2)
15	IgG1	HexNAc(4)Hex(5)
16	IgG1	HexNAc(4)Hex(5)Fuc(1)
17	IgG1	HexNAc(4)Hex(5)Fuc(1)2,3NeuAc(1)
18	IgG1	HexNAc(4)Hex(5)Fuc(1)2,6NeuAc(1)
19	IgG1	HexNAc(4)Hex(5)Fuc(1)2,6NeuAc(2)
20	IgG1	HexNAc(4)Hex(5)Fuc(2)
21	IgG1	HexNAc(4)Hex(5)2,6NeuAc(1)
22	IgG1	HexNAc(4)Hex(6)
23	IgG1	HexNAc(4)Hex(6)Fuc(1)
24	IgG1	HexNAc(5)Hex(3)
25	IgG1	HexNAc(5)Hex(3)Fuc(1)

26	IgG1	HexNAc(5)Hex(4)
27	IgG1	HexNAc(5)Hex(4)Fuc(1)
28	IgG1	HexNAc(5)Hex(5)
29	IgG1	HexNAc(5)Hex(5)Fuc(1)
30	IgG1	HexNAc(5)Hex(5)Fuc(1)2,6NeuAc(1)
31	IgG2	HexNAc(3)Hex(3)Fuc(1)
32	IgG2	HexNAc(3)Hex(4)Fuc(1)2,3NeuAc(1)
33	IgG2	HexNAc(4)Hex(3)
34	IgG2	HexNAc(4)Hex(3)Fuc(1)
35	IgG2	HexNAc(4)Hex(3)2,6NeuAc(1)
36	IgG2	HexNAc(4)Hex(4)
37	IgG2	HexNAc(4)Hex(4)Fuc(1)
38	IgG2	HexNAc(4)Hex(4)Fuc(1)2,3NeuAc(1)
39	IgG2	HexNAc(4)Hex(4)Fuc(1)2,6NeuAc(1)
40	IgG2	HexNAc(4)Hex(4)2,6NeuAc(1)
41	IgG2	HexNAc(4)Hex(5)Fuc(1)
42	IgG2	HexNAc(4)Hex(5)Fuc(1)2,3NeuAc(1)
43	IgG2	HexNAc(4)Hex(5)Fuc(1)2,6NeuAc(1)
44	IgG2	HexNAc(4)Hex(6)Fuc(1)
45	IgG2	HexNAc(5)Hex(3)
46	IgG2	HexNAc(5)Hex(3)Fuc(1)
47	IgG2	HexNAc(5)Hex(4)
48	IgG2	HexNAc(5)Hex(4)Fuc(1)
49	IgG2	HexNAc(5)Hex(5)
50	IgG2	HexNAc(5)Hex(5)Fuc(1)
51	IgG2	HexNAc(6)Hex(10)Fuc(5)2,3NeuAc(1) 2,6NeuAc(1)
52	IgG3	HexNAc(3)Hex(3)
53	IgG3	HexNAc(3)Hex(3)Fuc(1)
54	IgG3	HexNAc(4)Hex(3)Fuc(1)
55	IgG3	HexNAc(4)Hex(4)Fuc(1)
56	IgG3	HexNAc(4)Hex(5)Fuc(1)
57	IgG3	HexNAc(5)Hex(3)Fuc(1)
58	IgG3	HexNAc(5)Hex(4)Fuc(1)
59	IgG4	HexNAc(3)Hex(3)Fuc(1)
60	IgG4	HexNAc(4)Hex(3)Fuc(1)
61	IgG4	HexNAc(4)Hex(3)2,6NeuAc(1)
62	IgG4	HexNAc(4)Hex(4)Fuc(1)
63	IgG4	HexNAc(4)Hex(4)Fuc(1)2,3NeuAc(1)
64	IgG4	HexNAc(4)Hex(4)Fuc(1)2,6NeuAc(1)
65	IgG4	HexNAc(4)Hex(5)Fuc(1)
66	IgG4	HexNAc(4)Hex(5)Fuc(1)2,3NeuAc(1)
67	IgG4	HexNAc(4)Hex(5)Fuc(1)2,6NeuAc(1)
68	IgG4	HexNAc(5)Hex(3)Fuc(1)

69	IgG4	HexNAc(5)Hex(4)Fuc(1)
70	IgG4	HexNAc(5)Hex(5)Fuc(1)

*：N = HexNAc, H = Hexose, F = Fucose, A = NeuAc, G = NeuGc, Na stands for neutral exchange of H^+ with Na^+.

Table S3 N-glycopeptides detected from bovine fetuin

No	Glycosylation site	Glycans NHFAGNa
1	Asn99	HexNAc(3)Hex(3)2,3NeuAc(1)
2		HexNAc(3)Hex(3)2,6NeuAc(1)
3		HexNAc(3)Hex(8)2,6NeuAc(1)
4		HexNAc(3)Hex(8)2,3NeuAc(1)2,6NeuAc(1)
5		HexNAc(4)Hex(11)2,6NeuAc(1)
6		HexNAc(4)Hex(4)
7		HexNAc(4)Hex(4)2,3NeuAc(1)
8		HexNAc(4)Hex(4)2,6NeuAc(1)
9		HexNAc(4)Hex(4)2,3NeuAc(1)2,6NeuAc(1)
10		HexNAc(4)Hex(5)
11		HexNAc(4)Hex(5)Fuc(1)2,3NeuAc(1)2,6NeuAc(1)
12		HexNAc(4)Hex(5)Fuc(1)2,6NeuAc(2)
13		HexNAc(4)Hex(5)2,6NeuAc(2)
14		HexNAc(4)Hex(5)2,3NeuAc(1)2,6NeuAc(1)
15		HexNAc(4)Hex(5)2,6NeuAc(2)
16		HexNAc(4)Hex(6)Fuc(1)2,3NeuAc(2)
17		HexNAc(4)Hex(6)2,3NeuAc(1)2,6NeuAc(1)
18		HexNAc(4)Hex(7)2,6NeuAc(1)
19		HexNAc(5)Hex(10)2,3NeuAc(1)
20		HexNAc(5)Hex(10)2,6NeuAc(1)
21		HexNAc(5)Hex(10)2,3NeuAc(1)2,6NeuAc(1)
22		HexNAc(5)Hex(3)Fuc(1)
23		HexNAc(5)Hex(4)Fuc(2)2,3NeuAc(2)
24		HexNAc(5)Hex(4)Fuc(2)2,3NeuAc(1)2,6NeuAc(1)
25		HexNAc(5)Hex(4)2,3NeuAc(2)
26		HexNAc(5)Hex(4)2,3NeuAc(1)2,6NeuAc(1)
27		HexNAc(5)Hex(5)Fuc(1)2,3NeuAc(1)
28		HexNAc(5)Hex(5)Fuc(1)2,3NeuAc(2)
29		HexNAc(5)Hex(5)Fuc(1)2,3NeuAc(1)2,6NeuAc(1)
30		HexNAc(5)Hex(5)Fuc(1)2,6NeuAc(2)
31		HexNAc(5)Hex(5)2,3NeuAc(2)
32		HexNAc(5)Hex(5)2,6NeuAc(2)
33		HexNAc(5)Hex(6)Fuc(1)2,6NeuAc(1)
34		HexNAc(5)Hex(6)Fuc(1)2,3NeuAc(1)2,6NeuAc(1)

35		HexNAc(5)Hex(6)Fuc(1)2,6NeuAc(2)
36		HexNAc(5)Hex(6)Fuc(1)2,3NeuAc(3)
37		HexNAc(5)Hex(6)Fuc(1)2,3NeuAc(2) 2,6NeuAc(1)
38		HexNAc(5)Hex(6)Fuc(1)2,6NeuAc(3)
39		HexNAc(5)Hex(6)Fuc(2)2,3NeuAc(2)
40		HexNAc(5)Hex(6)Fuc(2)2,3NeuAc(1)2,6NeuAc(1)
41		HexNAc(5)Hex(6)Fuc(2)2,3NeuAc(3)
42		HexNAc(5)Hex(6)Fuc(2)2,3NeuAc(2) 2,6NeuAc(1)
43		HexNAc(5)Hex(6)Fuc(2)2,6NeuAc(3)
44		HexNAc(5)Hex(6)2,3NeuAc(1)
45		HexNAc(5)Hex(6)2,6NeuAc(1)
46		HexNAc(5)Hex(6)2,3NeuAc(2)
47		HexNAc(5)Hex(6)2,3NeuAc(1)2,6NeuAc(1)
48		HexNAc(5)Hex(6)2,3NeuAc(3)
49		HexNAc(5)Hex(6)2,3NeuAc(2) 2,6NeuAc(1)
50		HexNAc(5)Hex(6)2,3NeuAc(1) 2,6NeuAc(2)
51		HexNAc(5)Hex(6)2,6NeuAc(3)
52		HexNAc(5)Hex(6)2,3NeuAc(3)2,6NeuAc(1)
53		HexNAc(5)Hex(6)2,3NeuAc(2) 2,6NeuAc(2)
54		HexNAc(5)Hex(6)2,6NeuAc(4)
55		HexNAc(5)Hex(7)Fuc(1)
56		HexNAc(5)Hex(7)Fuc(1)2,6NeuAc(1)
57		HexNAc(5)Hex(7)Fuc(1)2,3NeuAc(2)
58		HexNAc(5)Hex(7)Fuc(1)2,3NeuAc(1)2,6NeuAc(1)
59		HexNAc(5)Hex(7)Fuc(1)2,6NeuAc(2)
60		HexNAc(5)Hex(7)Fuc(1)2,3NeuAc(3)
61		HexNAc(5)Hex(7)Fuc(1)2,3NeuAc(1) 2,6NeuAc(2)
62		HexNAc(5)Hex(7)Fuc(1)2,6NeuAc(3)
63		HexNAc(5)Hex(7)Fuc(3)2,3NeuAc(1)2,6NeuAc(1)
64		HexNAc(5)Hex(7)Fuc(3)2,6NeuAc(2)
65		HexNAc(5)Hex(7)2,6NeuAc(1)
66		HexNAc(5)Hex(8)Fuc(2)2,3NeuAc(1)
67		HexNAc(5)Hex(8)2,3NeuAc(1)
68		HexNAc(6)Hex(6)Fuc(1)2,3NeuAc(1)
69		HexNAc(6)Hex(6)Fuc(1)2,6NeuAc(1)
70		HexNAc(6)Hex(6)Fuc(1)2,3NeuAc(2)
71		HexNAc(6)Hex(6)Fuc(1)2,3NeuAc(1)2,6NeuAc(1)
72		HexNAc(6)Hex(6)Fuc(1)2,6NeuAc(2)
73		HexNAc(6)Hex(6)Fuc(1)2,3NeuAc(3)
74		HexNAc(6)Hex(6)Fuc(3)2,3NeuAc(1)
75		HexNAc(6)Hex(6)Fuc(3)2,6NeuAc(1)
76		HexNAc(6)Hex(7)Fuc(1)2,3NeuAc(1)
77		HexNAc(6)Hex(7)Fuc(1)2,3NeuAc(2)
78		HexNAc(6)Hex(7)Fuc(1)2,3NeuAc(1)2,6NeuAc(1)

79		HexNAc(6)Hex(7)Fuc(1)2,6NeuAc(2)
80		HexNAc(6)Hex(7)Fuc(2)2,6NeuAc(1)
81		HexNAc(6)Hex(7)Fuc(2)2,3NeuAc(1)2,6NeuAc(2)
82		HexNAc(6)Hex(7)Fuc(3)2,6NeuAc(1)
83		HexNAc(6)Hex(7)2,3NeuAc(2)
84		HexNAc(6)Hex(7)2,3NeuAc(1)2,6NeuAc(1)
85		HexNAc(6)Hex(7)2,6NeuAc(2)
86		HexNAc(6)Hex(7)2,3NeuAc(1)2,6NeuAc(2)
87		HexNAc(6)Hex(7)2,6NeuAc(3)
88		HexNAc(6)Hex(8)Fuc(1)2,3NeuAc(1)
89		HexNAc(6)Hex(8)Fuc(1)2,6NeuAc(1)
90		HexNAc(6)Hex(8)Fuc(1)2,3NeuAc(3)
91		HexNAc(7)Hex(5)
92		HexNAc(7)Hex(7)2,3NeuAc(2)
93		HexNAc(7)Hex(7)2,6NeuAc(2)
94	Asn156	HexNAc(4)Hex(5)2,3NeuAc(1)2,6NeuAc(1)
95		HexNAc(4)Hex(6)Fuc(1)2,3NeuAc(1)
96		HexNAc(4)Hex(7)2,3NeuAc(1)
97		HexNAc(5)Hex(10)2,6NeuAc(1)
98		HexNAc(5)Hex(10)2,3NeuAc(1)2,6NeuAc(1)
99		HexNAc(5)Hex(7)Fuc(1)2,3NeuAc(2)
100		HexNAc(5)Hex(7)Fuc(1)2,3NeuAc(1)2,6NeuAc(1)
101		HexNAc(5)Hex(8)2,3NeuAc(1)
102		HexNAc(6)Hex(6)Fuc(1)2,3NeuAc(3)
103		HexNAc(6)Hex(7)2,3NeuAc(1)2,6NeuAc(1)
104		HexNAc(3)Hex(3)2,3NeuAc(1)
105		HexNAc(4)Hex(6)Fuc(1)2,3NeuAc(2)
106		HexNAc(5)Hex(6)Fuc(2)2,3NeuAc(1)2,6NeuAc(1)
107		HexNAc(5)Hex(6)2,3NeuAc(2)2,6NeuAc(1)
108		HexNAc(5)Hex(6)2,6NeuAc(3)
109		HexNAc(5)Hex(6)2,3NeuAc(2)2,6NeuAc(2)
110		HexNAc(6)Hex(6)Fuc(1)2,3NeuAc(1)2,6NeuAc(1)
111		HexNAc(6)Hex(6)Fuc(1)2,6NeuAc(2)
112	Asn176	HexNAc(5)Hex(6)Fuc(1)2,6NeuAc(3)
113		HexNAc(5)Hex(6)2,3NeuAc(1)2,6NeuAc(1)
114		HexNAc(5)Hex(6)2,3NeuAc(3)
115		HexNAc(5)Hex(6)2,3NeuAc(2)2,6NeuAc(1)
116		HexNAc(5)Hex(6)2,3NeuAc(3)2,6NeuAc(1)
117		HexNAc(5)Hex(6)2,3NeuAc(2)2,6NeuAc(2)

REFERENCES

1. L. Yang, Z. Sun, L. Zhang, Y. Cai, Y. Peng, T.Cao, Y. Zhang and H. Lu, *Chem. Sci.*, 2019, **10**, 9302-9307.
2. M. Loos, C. Gerber, F. Corona, J. Hollender and H. Singer, *Anal. Chem.*, 2015, **87**, 5738-5744.