Egg yolk Sialylglycopeptide: purification isolation and characterization of *N*-glycans from minor glycopeptide species

Javier Mauricio Melo Diaz[a,b] ,Alan B. Moran[a,c] Simon R. Peel[a] ,Jenifer L Hendel[a] ,Daniel I. R. Spencer[a]

a Ludger Ltd, Culham Science Centre, Oxfordshire, Abingdon, UK
b Department of chemistry, Royal College of Surgeons in Ireland, 123 St Stephen's Green, Dublin.
c Leiden University Medical Center, Center for Proteomics and Metabolomics, 2300 RC Leiden, the Netherlands

Supplementary information

Table of Supplementary figures 1
1.0 Materials and reagents 3
2.0 Experimental procedures 3
2.1 SGP extraction 3
2.2 Vacuum assisted activated charcoal-celite chromatography 3
2.3 Preparative glycan release and purification 4
2.4 N-glycan labelling and derivatization 5
2.5 Analytical HPLC methods 6
2.6 MALDI-TOF-MS 6
2.7 Chemical and enzymatic modifications 7
2.8 Software used for image composition. 7
2.9 NMR Spectroscopy 8
3.0 Supplementary Figures 12
3.1 Regioselectivity of EEA 13
References 31

Table of Supplementary figures

S. Figure 1 UHPLC and ^{1}H-NMR Comparison of SGP obtained from egg yolk powder and from fresh eggs. Base peak chromatogram for the HILIC-MS separation of crude glycopeptides. A) Dry egg yolk B) Fresh egg yolk. 500 MHz ^{1}H-NMR C) for SGP from dry yolk, D) SGP from fresh egg yolk. 12

S. Figure 2 UHPLC analysis of procainamide labelled A2G2S2, before and after EEA. A) UHPL trace for procainamide labelled A2G2S2 standard. B) UHPLC trace for EEA-procainamide labelled A2G2S2 standards. C) ESI-MS procainamide labelled A2G2S2 standard. D) ESI-MS for EEA-procainamide label A2G2S2 E) ESI-MS/MS (815.41 m/z) from procainamide labelled A2G2S2 standard. F) ESI-MS/MS (834.12 m/z) from EEA-procainamide labelled A2G2S2(6) standard. 14

S. Figure 3 UHPLC analysis of procainamide labelled FA2G2S2, before and after EEA. A) UHPL trace for procainamide labelled FA2G2S2 standard. B) UHPLC trace for EEA-procainamide labelled FA2G2S2 standards. 1) ESI-MS procainamide labelled FA2G2S2 standard. 1') ESI-MS for EEA-procainamide label FA2G2S2(6). 15

S. Figure 4UHPLC analysis of procainamide labelled A3G3S3, before and after EEA. A) UHPL trace for procainamide labelled A3G3S3 standard. B) UHPLC trace for EEA-procainamide labelled A3G3S3 standards. 1) ESI-MS procainamide labelled A3G3S3 RT 34.0-34.2 min standard. 1') ESI-MS for EEA-procainamide label A3G3S2(3)S(6). 2) ESI-MS procainamide labelled A3G3S3 RT 35.3-36.2 min 2') ESI-MS for EEA-procainamide label A3G3S1(3)S2(6). 16

S. Figure 5 ESI- MS/MS procainamide labelled A3G3S3 A) ESI-MS/MS (1034.20) procainamide labelled A3G3S3 RT 35.6 min, peak 2 S.Figure 4, standard. B) ESI-MS/MS for EEA-procainamide label A3G3S1(3)S2(6). (945.81) RT 29.1 min, peak 2' S.Figure 4. ...17

S. Figure 6 ESI-MS of procainamide labelled N-glycans from Figure 2. Colum A). ESI-MS of procainamide labelled N-glycans. Column B). ESI-MS of EEA-procainamide labelled glycan. 18

S. Figure 6 ESI-MS of procainamide labelled N-glycans from Figure 3A. 19

S. Figure 7 Base peak chromatograms for the UHPLC-HILIC-MS separation of the crude egg yolk glycopeptides. 20

S. Figure 8 ESI-MS of Glycopeptides (A-C) or glycan (D-F) species identified on S. Figure7 21

S.Figure 9 A) UV 214 nm HPLC PGC preparative purification of A2[3]G1S1 and A2G2S2, peaks identified with an asterisk correspond to glycan containing peaks B) UHPLC 2AB labelling of PGC purified N-glycans C) Purification yield estimation, based on the abundance of each glycan species obtained from procainamide labelling. 22

S. Figure 10 ^{1}H-NMR 500 MHz in D_2O of the purified N-glycans A) A2[3]G1S1 B) A2G2S2 23

S. Figure 11 ^{1}H-NMR 500 MHz in D_2O of the anomeric region of the purified N-glycans A) A2[3]G1S1 B) A2G2S2 24

S. Figure 12 MALDI-MS/MS expected fragmentation for A2[6]G1 25

S. Figure 13. MALDI MS and MS/MS analysis for A2[3]G1F and A2G2F2 A) MALDI MS for A2G1F B) MS/MS for A2G1F C) MALDI MS for A2G2F2, D) MALDI MS/MS of A2G2F2.....26

S. Figure 14 Retention time-based confirmation of sialic acid linkage, using a chemoenzymatic produced glycan. A) Initial A2[3]G1S1(6). B) A2[3]G1, obtained after sialic acid removal. C) A2[3]G1S1(3), obtained after sialylation reaction 27

S. Figure 15 Extracted Ion chromatogram at 1033.70 m/z, A) Procainamide labelled N-glycans from bovine fetuin. B) Procainamide labelled N-glycans from S3 fraction C) procainamide labelled N-glycans after antennary fucosylation. Different isomers are detected 4.1,4.0. 4.2 Correspond to a triantennary glycan with two, one and zero 2,3 linked sialic acid, respectively. 28

S. Figure 16 A) procainamide labelled N-glycans after antennary fucosylation B) Extracted ion chromatogram at 1082.70 m/z for the UHPLC trace of procainamide labelled N-glycans after fucosylation reaction C) ESI-MS/MS of the different isomeric species. 5.1, 5.0, 5.2.....29

S. Figure 17 A) procainamide labelled N-glycans after antennary fucosylation. B) Extracted ion chromatogram a 1131.10 m/z corresponding to the doubly fucosylated product. C) ESI-MS for 5.3. 30

S. Figure 18Additional side reaction detected during fucosylation A) Extracted ion chromatogram at 657.2 m/z and 804.3 m/z before fucosylation reaction. B) Extracted ion chromatogram at 657.2 m/z and 804.3 m/z after fucosylation reaction C) ESI-MS of potential fucose acceptor D) ESI-MS of additional fucosylated product. 31

1.0 Materials and reagents

Unless stated below all reagents including Sartorius Vivaspin ultrafiltration unit MWCO 10 kDa, were obtained from Sigma Aldrich. Fresh eggs were acquired at local supermarket, egg yolk powder was acquired from Pulviver (Bastogne, Belgium). PD-10 desalting columns were obtained from Cytiva. LudgerTag 2AB, LudgerTag PROC labelling kits, LudgerClean S cartridges, EC-50 SPE cartridges, LudgerSep N Buffer x40 concentrate, 6M HCl, Ludger BioQuant chitotriose, PNGase-F-QA-Bio E-PNG05, 2AB-labelled A2G1 glycan standards, and unlabelled glycan standards A3, A2F(FA2G2S2), A2(A2G2S2) were obtained from Ludger Ltd (Abingdon, UK). Acetonitrile, methanol, ethyl acetate, and acetone were acquired from Romil Ltd (Cambridge, UK). Toyopearl® GigaCap Q-650M, 75 µm, anion exchange resin was obtained from Tosoh Bioscience (Japan). Thermo Scientific™ FastAP Thermosensitive Alkaline Phosphatase (1 U/µL) was obtained from Fisher-Scientific. *Photobacterium phosphoreum* alpha/beta α-2,3 sialyltransferase (Tt-88) and *Helicobacter pyllori* Fucosyltransferase (Hp α1,3FT/ Tt-460) were obtained as part of the Glycoenzymes for Bioindustries consortium and expressed by Prozomix (Haltwhistle, UK). CMP-Neu5Ac-Sodium salt and GDP-fucose-sodium salt were obtained from Carbosynth (Compton, UK).

2.0 Experimental procedures

2.1 SGP extraction

2.1.1 SGP Extraction from dry yolk powder used for preliminary experiments

The procedure from[1,2] was adapted with minor changes. All the processes were carried out at room temperature. Briefly, to 200 g of dry yolk powder, 450 mL of ethyl acetate (~1:2.2 solid to liquid ratio) were added to remove lipids and allowed to stir for 1h. The solid residue was filtered using a sintered glass funnel and the extraction repeated with 450 mL of ethyl acetate. The solid residue was allowed to dry and then it was washed twice with 200 mL of 70% acetone in water (v/v), allowing for 15 min for each washing. SGP was extracted by washing the solid residue with 200 mL of 40% acetone in water (v/v) twice. After filtering through a sintered glass funnel. Acetone was removed using a rotary evaporator. The concentrated extract was separated into tubes and dried using a ThermoSavant Speedvac for 8 h. To produce a yellow solid, this residue was reconstituted in 5 mL of DI water and purified using vacuum-assisted activated charcoal-celite chromatography as described below.

2.1.2 SGP extraction from fresh egg yolks

The process of extraction of SGP from fresh eggs was adapted from the purification protocols of IgY purification [3]. Briefly, starting from fresh eggs purchased from a local supermarket. The yolks from 48 eggs were manually separated from the white. Yolks were punctured using a pipette tip and their contents drained into a measuring cylinder, yolks sacks and chalazas were discarded. Total volume of yolk was determined, and DI water was added to achieve a 1:8 dilution and mixed to achieve homogeneity. The suspension was transferred to a suitable container and allowed to freeze. The solid solution was thawed at room temperature producing a precipitate and a supernatant. The supernatant was filtered through celite and concentrated using a rotary evaporator. To minimize loss of sialic acid during evaporation the temperature was not allowed to exceed 35 °C. 1-octanol (100 µL/ 100 mL of solution) was added to decrease bubbling. For 48 eggs, the volume of supernatant that was obtained was about 4.5 L, which was concentrated down to a volume of 1.2 L. Pure acetone was added to the concentrated sample to achieve a final 40% ratio (v/v) The solution was allowed to decant at -20 °C and then centrifuged at 2500xg for 15 min to remove the precipitate. The supernatant was concentrated by rotary evaporation to a volume of 250 mL and split into 50 mL portions for subsequent purification by vacuum assisted activated charcoal-celite chromatography.

2.2 Vacuum-assisted activated charcoal-celite chromatography

2.2.1Column packing

Equal amounts of activated charcoal and celite were weighed and mixed with DI water to form a suspension and allowed to decant for 1 h, after which the supernatant containing fine particles was discarded. This process was repeated twice. To a plastic funnel equipped with a bottom frit of 10 µm, ~1 g of dry celite was added, and dispersed using DI water. The funnel was placed on top of a vacuum flask and, the vacuum was applied to compact the celite into a thin layer. The carbon slurry was then added gradually with the vacuum turned on to achieve a column volume (CV) of ~ 35 mL, a small amount of cotton wool was added to the top of the column.

2.2.2 Glycopeptide purification

For the glycopeptide purification either 50 mL fresh egg solutions or 5 mL solution from dry eggs were added on top of a carbon-celite column. The column was washed stepwise with 100 mL (~3CV) portions of 0, 5, 10 and 25 % MeCN (v/v). vacuum was applied to allow for faster separation[4]. Carbohydrate presence in fractions was monitored by H_2SO_4-resorcinol hexose test. Glycopeptides eluted at 25 % MeCN. Acetonitrile was removed by rotary evaporation, and the samples freeze-dried. Dried samples were resuspended in water and centrifuged to remove insoluble material e.g., celite. The supernatant was then subjected to 2xG25 Sephadex purification. Each fraction was analysed using, MALDI-MS SGP containing fractions (SGP-A2[3]G1S1: 2410.6 [M-H]-, SGP: 2885.13 [M+Na-H]- or 2901.7 [M+K-H]- negative mode), were collected and freeze dried, to yield a white fluffy powder, based on the weight of the dry residue weight, about 30 mg of crude SGP can be obtained from ~10 eggs, or ~ 50 mg from 200 g of yolk powder).

2.2.3 resorcinol test for fractions

Samples to be tested were spotted on a silica-TLC plate and allow to dry. Spotted plates were dipped into the resorcinol reagent (0.1 % (w/v) Resorcinol in 95% EtOH, 1% H_2SO_4) and heated on a hot plate 100 °C for five minutes, with a cherry-red colour indicating the sample was positive for carbohydrates.

2.2.4 Gel filtration using 2xG25 HiPrep 26/10

Dry samples were resuspended 5mL of water injected to a 2x High prep 26/10 Sephadex-G25 column. Elution was performed with water at a flow rate of 1mL/mL, with 5 mL fractions collected starting when the elution volume reached about 15 mL. Fractions were analysed by HPAE-PAD and or MALDI-MS, Glycan/glycopeptide containing fractions were pooled and freeze dried.

2.3 Preparative glycan release and purification

2.3.1 SGP release using PNGase-F

30 mg of crude SGP was dissolved into 2 mL of 50 mM NaH_2PO_4-$NaHPO_4$ pH 6.0, 0.05% NaN_3 20 µL of PNGase-F-QA-Bio E-PNG05 were added, and the reaction allowed to incubate for 1 d at 37 °C, after which an additional 10 µL were added. (30 µL total) and allowed to procced for additional 12 h, the reaction was monitored with using HPAE-PAD (1 µL reaction mixture+29 µL DI water) and MALDI-MS (0.5 µL of the 1:30 dilution). Released *N*-glycans were separated from PNGaseF using Vivaspin ultrafiltration unit MWCO 10 kDa, salts were removed using 2xG25 HiPrep 26/10 Sephadex columns as previously indicated.

2.3.2 Anion Exchange Chromatography

After desalting released glycan were dried using a ThermoSavant Speedvac, and resuspended to a volume of 2.2 mL, injection of 2.0 mL was applied to an anion exchange column (GigaCap Q-650M, ,25mmx80mm, 75 µm Toyopearl®) column volume 40 mL, on a Dionex Ultimate 3000 UHPLC with an automatic fraction collection. Elution was performed at 3.0 mL/min, with detection at 214nm. Using a two solvent system, solvent A was 10 % MeCN, and solvent B was 250 mM ammonium formate.

S. Table 1 Chromatographic method used for anion exchange separation.

Time(min)	0	20	65	78	80	110
%B (250 mM ammonium formate)	0	0	100	100	0	0

After anion exchange chromatography three N-glycans pools S1, S2 and S3 were collected (~50 mL) from each pool an aliquot of 35 µL was dried down and procainamide labelled. S1 and S2 were further purified by porous graphitised chromatography (PGC), S3 was desalted using a PD10 column.

2.3.3 PD10 desalting.

A PD10 column was used according to the supplier instructions. The column was equilibrated with 30 mL of DI water. 1.0 mL of glycan sample was introduced and followed by 1.5 mL of DI water. Elution was carried out with 3.5 mL of DI water, collecting 1.0 mL fractions. After MALDI-MS analysis, glycan containing fractions were combined.

2.3.4 Porous graphitised carbon preparative chromatography.

Sample volume was reduced using a rotary evaporator to a volume of about 5 mL. Injections of 1.5 mL were applied to a semi-preparative porous graphitised carbon (PGC) column (Hypercarb ™,10mm i.d x250 mmm, 7 µm, Thermo Fisher) at 40°C on a Dionex Ultimate 3000 UHPLC with an automatic fraction collection. Elution was performed at a flow rate of 1.5 mL/min, using a two solvent gradient. Solvent A was 1.0 % MeCN + 0.1 % TFA, and solvent B was 50% MeCN + 0.1 % TFA, the column was equilibrated with solvent A. The gradient elution parameters were 0-50% B linear gradient within 80 min. Glycan peaks were collected as they eluted from the column. Collection time was 15 seconds affording fractions of 375 µL. UV detection was at 214 nm, and these fractions were analysed by MALDI-MS and HPAE-PAD. Glycan containing fractions were combined and dried down using a Thermo Scientific Savant SpeedVac vacuum concentrator.

S. Table 2 Gradient parameters used for PGC-preparative chromatography.

Time(min)	0	0	10	90	110	112	121	122
%B (50% MeCN, 0.1% TFA)	0	0	0	50	70	0	0	0

2.4 N-glycan labelling and derivatization

2.4.1Analytical PNGase-F release

PNGase-F release was carried using a LudgerZyme recombinant PNGase F Kit as follows: to 4 µL of glycopeptide (~80 µg), 5 µL of DI water and 1 µL of 10X denaturation solution (5% SDS 400 mM DTT) was added and the samples heated for 5 min at 100 °C. After cooling down to room temperature, 5µL of H_2O ,2 µL of 10% NP-40 solution., and 2 µL of 10x Reaction buffer concentrate (500 mM sodium phosphate) pH 7.7, are added and briefly centrifuged. To this mixture is added 1 µL of LudgerZyme Peptide N-glycosidase F (500,000 units/mL). Prior to labelling the samples are concentrated to dryness on a Thermosavant speedvac, and procainamide labelled.

2.4.2 aminobenzamide (2AB) labelling

2-aminobenzamide (2AB) labelling was used to estimate the amount of free reducing-end N-glycans after PGC purification by comparing against a labelled chitotriose quantified standard (Ludger BioQuant chitotriose). Labelling was carried out using LudgerTag 2AB labelling kit as follows: 10 µl of labelling solution (2AB 0.3 M, $NaBH_3CN$ 1 M, dissolved in acetic acid: DMSO, (3:7)) was added to 5µL of the sample, briefly vortexed and incubated at 65°C for 3h. Excess reagents were removed using LudgerClean S cartridges using the procedure below.

2.4.3 Procainamide-labelling

Free reducing-end N-glycans were labelled with procainamide using the LudgerTag Proc labelling kit, to a dry glycan aliquot, 20 µL of labelling solution was added (procainamide-HCl 0.2 M, $NaBH_3CN$0.5 M dissolved in acetic acid: DMSO: H_2O, (1.5:3.5:5)) briefly vortexed and incubated for 1 h at 65°C. Excess labelling reagents were removed using LudgerClean S cartridges.

2.4.4 Removal of excess labelling reagents

LudgerClean S cartridges were conditioned successively with 1mL of DI water, 5 mL of 30 % acetic acid (v/v), and 1 mL of acetonitrile. 2AB or procainamide labelled samples were then spotted on the cartridge and allowed to adsorb for 15 min. The excess dye was washed first with 1 mL of acetonitrile and then 2x5 mL 96% aqueous acetonitrile. Labelled N-glycans were eluted with 1 mL of DI water.

2.4.5 Ethyl esterification amidation (EEA) of sialic acids

2.4.5 Ethyl esterification amidation (EEA) of procainamide labelled N-glycans

Sialic acid linkage specific derivatization by ethyl esterification-amidation (EEA) was performed as described previously [5]. For free reducing end glycan standards (FA2G2S2, A3 and A2G2S2) procainamide labelling was performed as previously described, followed by the ethyl esterification protocol. In this case, the labelled standards were reconstituted in 100 µL and 3 µL (~100 pmols) was used for the reaction with 60 µL of ethyl esterification reagent (250 mM EDC and 250mM HOBt in EtOH). Incubation occurred for 30 min at 37 °C followed by the addition of 12µL of 28% (w/w) NH_3. Then, another incubation was performed for 30 min at 37 °C. Following this, 600 µL MeCN was added to the samples. Samples were loaded to a pre-equilibrated LudgerClean Procainamide clean-up plate (conditioning was done with 200 µL of 70 % ethanol, 200 µL of water, and 200 µL of MeCN, applying vacuum

between each addition) by the addition of 300 µL and draining for 5 min, followed the further addition of 300 µL of the remaining sample. allowing the loaded EEA-N-glycans to drain by gravity, the plate was washed with 200 µL of acetonitrile (5x times to remove excess reagents), EEA-glycans were eluted from the membrane with 300 µL of water.

Procainamide labelled glycans obtained from SGP were processed as follows. 1µL of procainamide labelled N-glycans (~100 pmoles) were added to 20 µL of ethyl esterification reagent (250 mM EDC and 250mM HOBt in EtOH) and incubated for 30 min at 37 °C followed by the addition of 4µL of 28% (w/w) NH_3. Incubation was performed for an additional 30 min at 37 °C. Following this incubation, 24 µL MeCN was added to the samples. After derivatization. EEA- N-glycans were cleaned using EC-50 SPE carbon resin cartridges. The cartridges were conditioned with successively 1mL~MeOH,1mL NaOH 0.1 M,1 mL 30% acetic acid (v/v) and 1 mL H_2O+ 0.1 % TFA (v/v). 500 µL of DI water was added to the samples after the EEA reaction and loaded onto the cartridges after which they are washed with 1 mL H_2O+0.1 % TFA(v/v), and 5% MeCN + 0.1% TFA (v/v) to elute non-glycan impurities and salts. EEA-N-glycans were eluted with 0.8 mL of 50 % MeCN + 0.1% TFA (v/v) and dried using a Thermosavant speedvac. EEA-N-glycans were reconstituted in DI water and analysed using the same UHPLC-MS gradient used for procainamide labelled N-glycans.

2.5 Analytical HPLC methods

2.5.1. HPAE-PAD analysis

Where indicated in the main text, analyses by HPAE-PAD were conducted on a Thermo Fisher IC 5000 system with electrochemical detection. Glycan containing samples were separated on a Thermo Scientific Dionex CarboPac PA1 column set (4x30 mm and 4x 250 mm,10 µm) at 30 °C. with the following conditions: There were two mobile phases A and B, containing 100 mM sodium hydroxide and 1M sodium acetate-100mM sodium hydroxide, respectively. The separation started with a gradient from 2- 25 % of B within 42 min. 1:10 or 1:30 v/v for injection. The injection volume was 10 µL and the flow rate was 1.0mL / min. Detection was performed using a square waveform and a disposable gold working electrode.

2.5.2 UHPLC- of 2AB labelled samples

20 µL of 2AB labelled N-Glycan solution were diluted with 80µL of MeCN, 10 mL of sample were injected into a Waters ACQUITY UHPLC BEH-glycan column (2.1X150 mm, 1.7 µm) at 60 °C on a Waters Acquity System with florescence detection (excitation 250nm, emission 428 nm). Running conditions were as follows Solvent A was 50 mM ammonium formate (pH 4.4) made from LudgerSep N Buffer x40 Concentrate, and solvent B was MeCN. Gradient elution was performed at 560 µL/min under the following conditions, linear gradient of 78-58 B% within 24.7 min.

2.5.3 UHPLC-MS For procainamide labelled samples.

20 µL of procainamide labelled solution was diluted with 80µL of MeCN (~80-190 pmol of N-glycans per vial). 20 µL of the sample was injected into a Waters Acquity UHPLC BEH-glycan column (2.1X150 mm, 1.7 µm) at 60°C On a Dionex Ultimate 3000 UHPLC with fluorescent detection (excitation 310nm, emission 370 nm) attached to a Bruker Amazon Speed ETD mass spectrometer. Data acquisition was controlled by HyStar version 3.2 (Bruker). Running conditions were as follows; solvent A was 50 mM ammonium formate (pH 4.4) made from Ludgerx40 stock buffer, and solvent B was MeCN. Elution was performed at 400 µL/min using a linear gradient of 76-55% B within 53.4 min. The mass spectrometer was used in enhanced resolution mode with positive ion setting at a mw range of 400–1700 m/z. Nebuliser pressure was 14.5 psi, capillary voltage was 4500 V, nitrogen flow was 10 L/min, ion charge control (ICC) target was 200,000, maximum accumulation time was 50.00 ms, singly charged ions were excluded. The same conditions were used unlabelled glycopeptides (1 mL of crude glycopeptide 0.02 mg/µL), 19 µL DI water, 80 µL MeCN.

2.6 MALDI-TOF-MS

The MALDI-TOF mass spectrometry experiments were performed using a Bruker Auto-flex Speed (Bruker Daltonics, Bremen, Germany). Sample preparation was as follows; 0.5 µL of Super-DHB matrix (50 mg/mL in 50:50 [v/v] acetonitrile: DI water) was spotted on a ground steel target. 0.5 µL of the sample in water was added on top and allowed to dry. The spectrometer was operated in positive ion mode. Spectra were acquired in the mass range 900–3500 m/z for MS (reflector voltage, lens voltage respectively 20.9kV, 8.35 kV) at a laser intensity of 50%. For MS/MS (reflector voltage, lift voltage respectively 27 kV, 19 kV). The MS data were further processed using Flex Analysis 3.5.

2.7 Chemical and enzymatic modifications

2.7.1 Acid catalysed desialylation

Samples were dissolved in 1 mL of water, HCl 6M (8µL) was added to achieve a final concentration of 50 mM, samples were left in a heating block overnight at 60°C, the progress of the reaction was followed with HPAE-PAD and MALDI-MS, after completion samples were neutralised with 2.8 %(v/v) NH_3 then desalted with a PD10-column.

2.7.2 Fucosylation reaction.

Enzyme (Hp α1,3FT-Tt-460) from *Helicobacter pylori* [6] was obtained from Prozomix in the form of an ammonium sulphate precipitate. The protein precipitate was resuspended in 50 mM Tris-HCl pH 7.8 and quantified using a BCA kit (Thermo, UK) against a BSA standard. The enzyme concentration in the reaction was adjusted to 1mg/mL. Stock GDP-L-fucose solution was prepared by dissolving solid reagent in 50 mM Tris-HCl, 10 mM $MnCl_2$, pH 7.8 buffer. The GDP-L-fucose solution was added to dry N-glycans to achieve a concentration of acceptor and donor of 4mM and 8mM, respectively. FastAP 1U/200 µL, was added to drive the reaction forward, the solution was allowed to incubate at 37 °C for 12 h. The reaction was monitored using MS-MALDI. Protein was removed using a Vivaspin ultrafiltration unit MWCO 10 kDa, and the N-glycans desalted using a 2x High prep 26/10 Sephadex-G25 column.

2.7.3 Alpha-2,3 sialyltransferase reaction

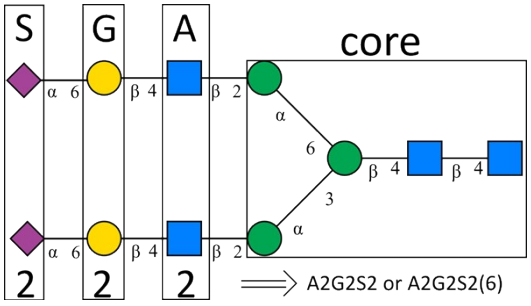
Alpha-2,3 Sialic acid transferase from *Photobacterium phosphoreum* [7] and was obtained from Prozomix in the form of an ammonium sulphate precipitate. The protein precipitate was resuspended in 50 mM Tris-HCl pH 7.8 and quantified using a BCA kit (Thermo, UK) against a BSA standard. The enzyme concentration in the reaction was adjusted to 0.1mg/mL. Stock CMP-Neu5Ac solution was prepared by dissolving solid reagent in 50 mM Tris-HCl,10 mM $MgCl_2$ pH 8.0. CMP-NeuAc solution was added to dry N-glycans to achieve a concentration of acceptor and donor of 4mM and 8mM, respectively. FastAP 1U/200 µL was also added. The solution was incubated at 37 °C for 1-3h and the extent of reaction was monitored using MS-MALDI, and 2AB labelling. Reaction was stopped by addition of methanol to achieve a concentration of 50% and stored at – 20 °C. Precipitated proteins were removed by centrifugation.

2.8 Software used for image composition.

N-glycans were represented according to the symbol nomenclature for N-glycans (SNFG) using glycoworkbwench 2.0. 7.[8] Chemical structures and schemes were drawn using ChemDraw 20.1.1. NMR data was produced using TopSpin 4.1.3. UHPLC-ESI-MS was processed using Bruker Compass DataAnalysis 4.4. MALDI-TOF data was processed using Flex Analysis 3.5. additional graphs were produced using RStudio (2021). For composite image creation Inkscape 0.92 Open-Source Scalable Vector Graphics Editor was used.

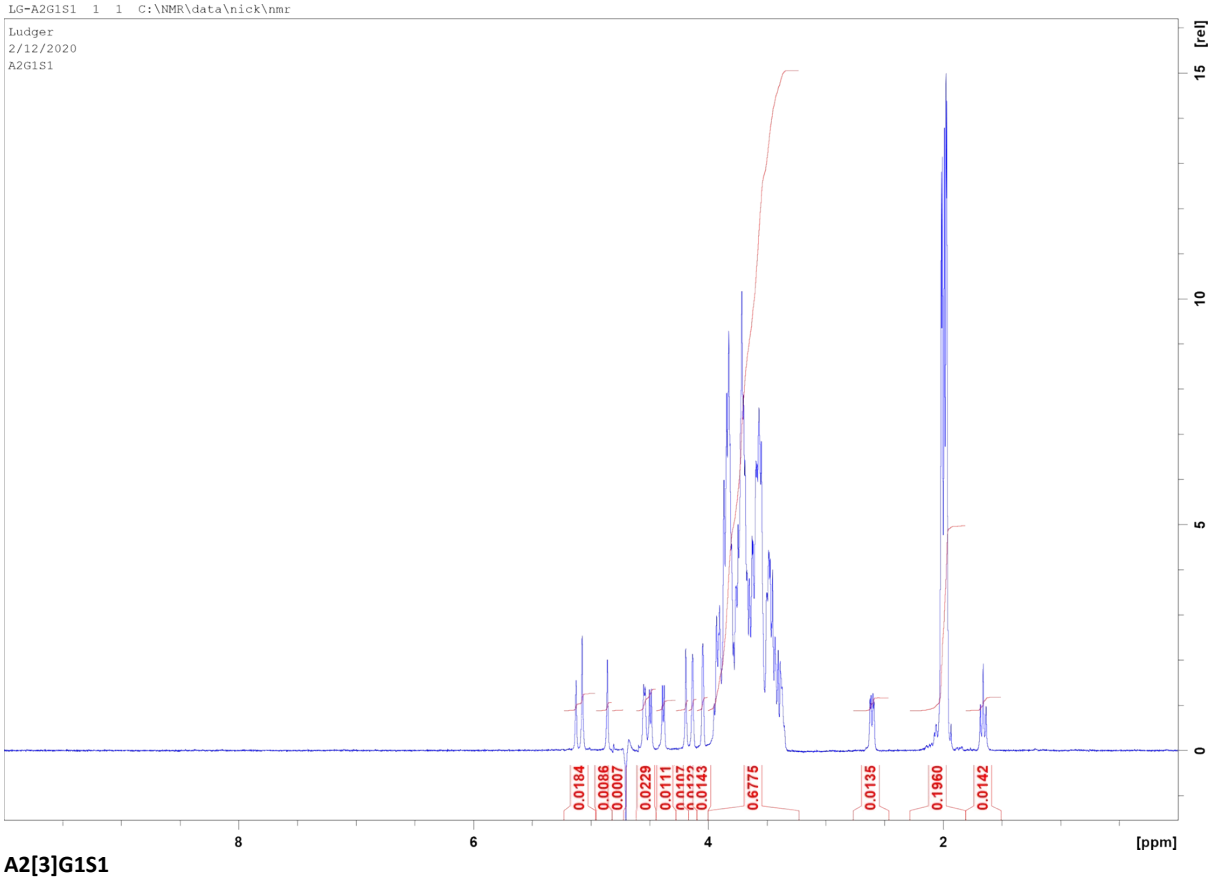
2.9 Oxford Notation for N-Glycan abbreviation

The Oxford notation is based on building up N-glycan structures and it can be used to denote very complex glycans. All N-glycans have two core GlcNAcs; F at the start of the abbreviation indicates a core fucose; Ax, where x - number of the antenna (GlcNAc) on trimannosyl core; Gx, where x - number of linked galactose on antenna; [3]G1 and [6]G1 indicates that the galactose is on the antenna of the α1-3 or α1-6 mannose; Sx, where x - number of sialic acids linked to galactose. Numbers in brackets are used to indicate linkages, where known.



2.10 NMR Spectroscopy

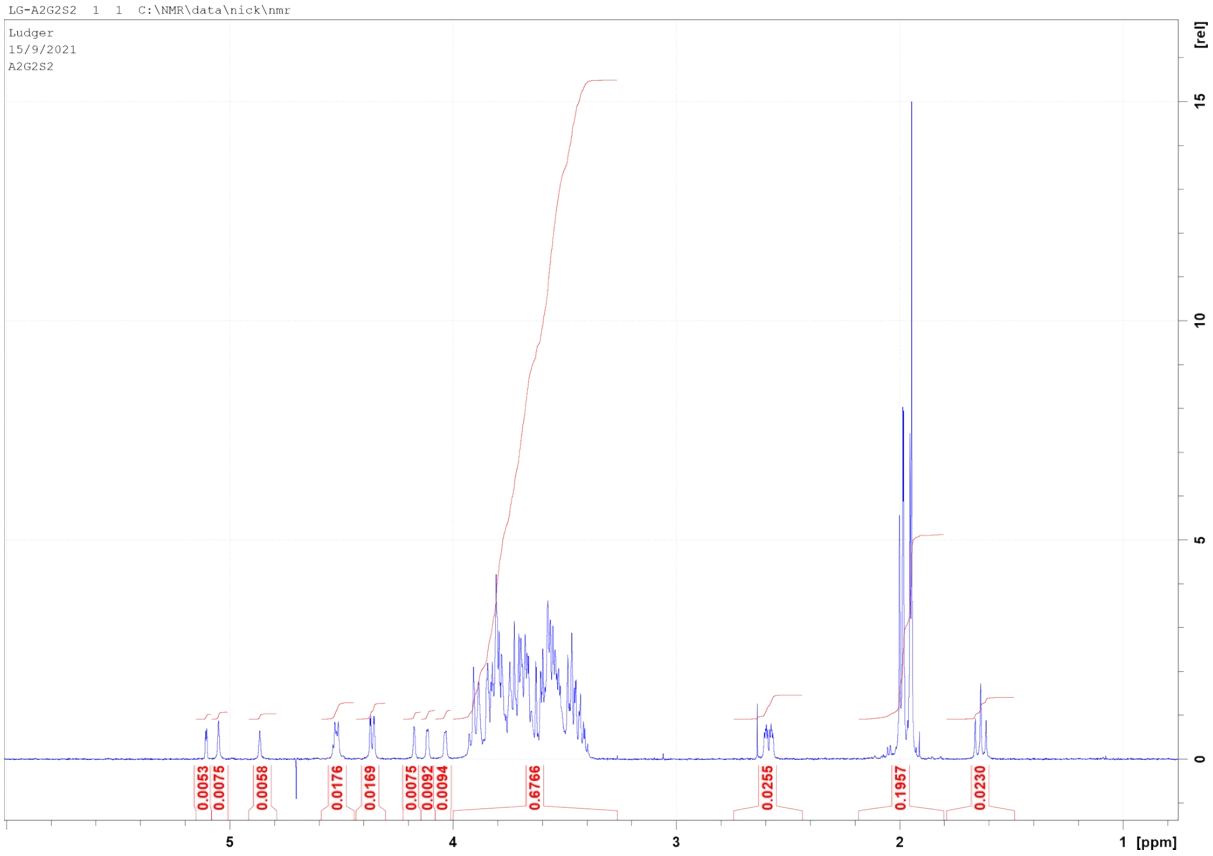
For the acquisition of ^{1}H-NMR, samples were dissolved in D_2O, and spectra were measured using a Brucker AVII 500 spectrometer at 25°C, with an internal standard of HDO at 4.70 ppm



A2[3]G1S1

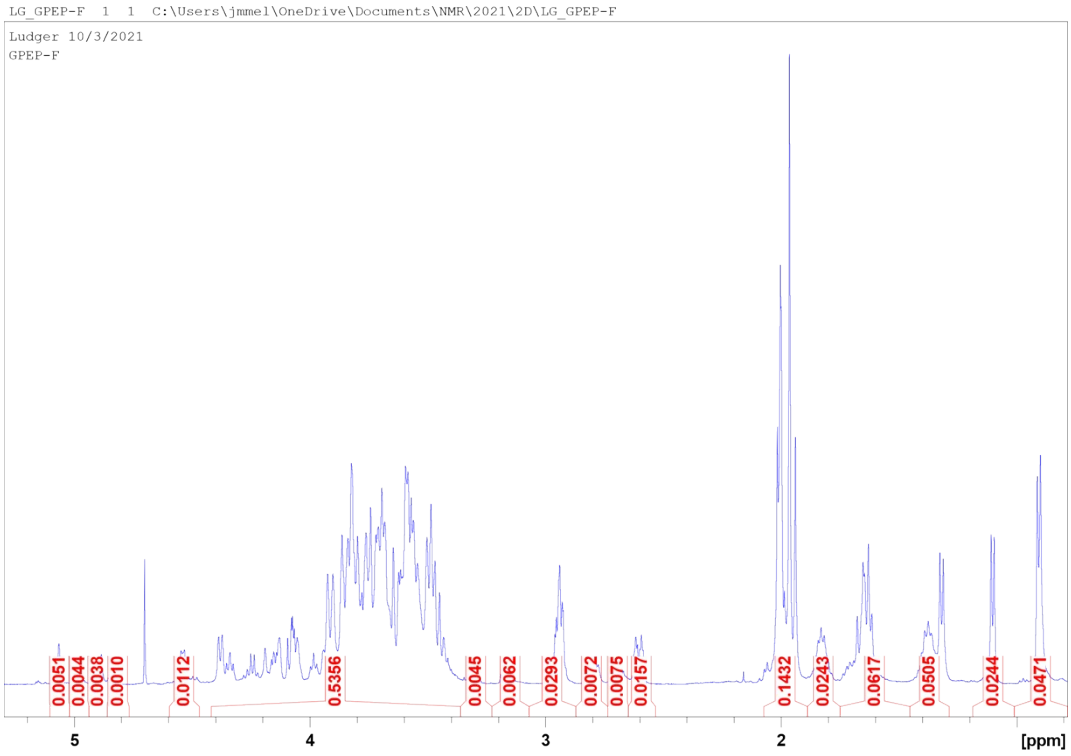
A2[3]G1S1 (93% purity 2AB) :^{1}H NMR (500 MHz, 25°C. internal standard HDO, 4.70 ppm): 5.13 (s, 1H, H-1 , GlcNAc 1), 5.07 (s, 1H, H-1, Man-4), 4.86(s, 1H, H-1 , Man-4'), 4.57(s, 1H, H-1 , GlcNAc-2), 4.46 and 4.50(s, 2H, H-1 , GlcNAc 5 and 5') 4.38(d,1H,H-Gal-6),4.22 (s, 1H, H-2 , Man-3), 4.14(s, 1H, H-2 , Man-4), 4.05(s, 1H, H-2, Man-4'),3.30-4.00-(br m, 55 H H-2,3,4,5,6 of Man,GlcNAc, Gal, and H 4,5,6,7,8,9 of Neu5Ac) 2.61(q, 1H, H-3_{eq} Neu5Ac),1.95-2.05 (br m, 14H, CH_3, GlcNAc, and CH_3 Neu5Ac), 1.66 (t, 1H, H-3_{ax}, Neu5Ac)

A2G2S2



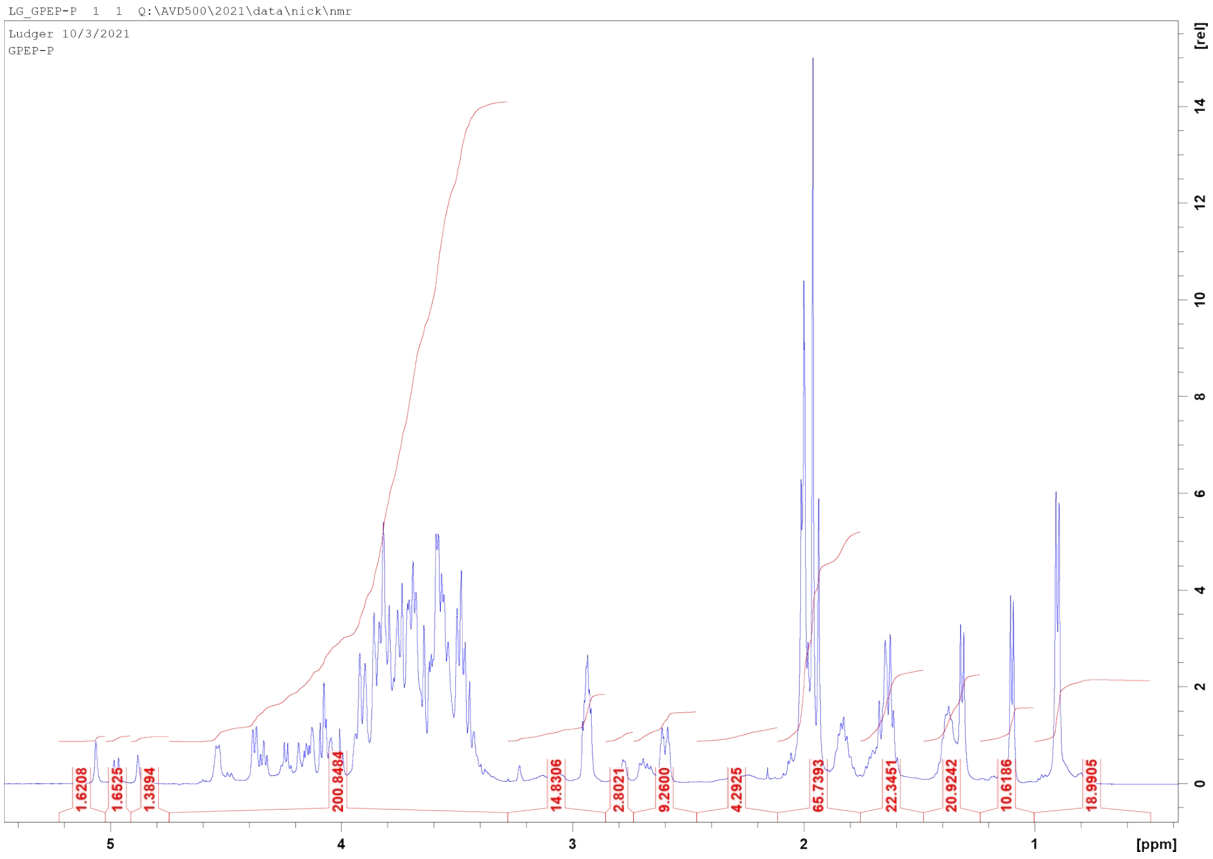
A2G2S2 (94% purity 2AB) ^{1}H NMR (500 MHz, 25°C. internal standard HDO, 4.70 ppm): 5.12 (s, 1H, H-1 , GlcNAc 1), 5.04 (s, 1H, H-1, Man-4), 4.85(s, 1H, H-1 , Man-4') 4.50-4.54(br, m, 3H, H-1 GlcNAc2, GlcNAc 5 and 5') 4.36(dd,2H,1H-Gal-6 and Gal 6'),4.18 (s, 1H, H-2 , Man-3), 4.11(s, 1H, H-2 , Man-4), 4.05(s, 1H, H-2, Man-4'),3.30-4.00-(br m, 68 H H-2,3,4,5,6 of Man ,GlcNAc, Gal, and H-4,5,6,7,8,9 of Neu5Ac) 2.59(q, 2H, H-3_{eq} Neu5Ac),1.90-2.00(br m, 18H, CH_3, GlcNAc, and CH_3 Neu5Ac), 1.64 (t, 2H, H-3_{ax}, Neu5Ac):

SGP from fresh egg yolk



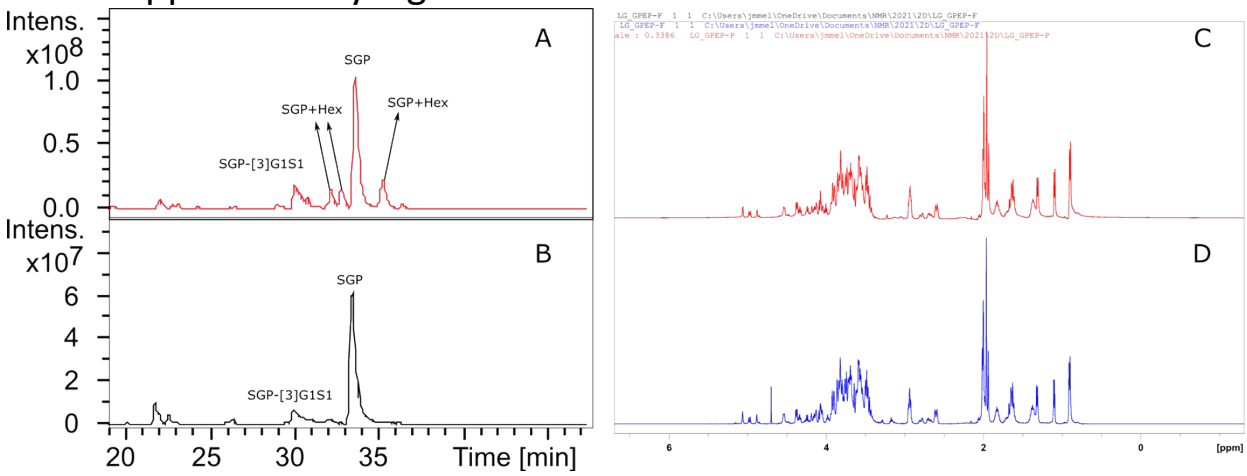
SGP from fresh eggs :^{1}H NMR (500 MHz, 25°C. internal standard HDO, 4.70 ppm): 5.06 (s, 1H, H-1 , Man-4), 5. 4.98 (d, H-1, GlcNAc-1), 4.90 (s, H-1 , Man-4'), 4.47-4.60 (br,m, H-1 , GlcNAc-2, GlcNAc 5 and 5'), 4.39 (d, H-1 , H-Gal-6', H-Gal-6), (3.35-4.35 , br m, H-2,3,4,5,6 of Man,GlcNAc, Gal, and H 4,5,6,7,8,9 of Neu5Ac), 2.97 (m, e CH_2,Lys), 2.59 (q, $H3_{eq}$ Neu5Ac),1.91-2.03 (m,-CH_3 GlcNAc , CH3 Neu5Ac), 1.82 (t, $H3_{ax}$, Neu5Ac), 1.60 (m, b m, dCH_2, Lys , b CH_2-Lys), 1.37 (m, , g-CH_2 Lys), 1.31 (d,-CH_3 Ala), 1.10 (d,-CH_3 Thr), 0.88 (d,-CH_3 Val).

SGP from dry egg yolk



SGP from egg yolk powder: ^{1}H NMR (500 MHz, 25°C. internal standard HDO, 4.70 ppm): 5.06 (s, 1H, H-1 , Man-4), 5. 4.97 (d, H-1, GlcNAc-1), 4.81 (s, H-1 , Man-4'), 4.51-4.56 (br,m, H-1 , GlcNAc-2, GlcNAc 5 and 5'), 4.38 (d, H-1 , H-Gal-6', H-Gal-6), (3.28-4.35 , br m, H-2,3,4,5,6 of Man,GlcNAc, Gal, and H 4,5,6,7,8,9 of Neu5Ac), 2.93 (m, e CH_2,Lys), 2.60 (q, $H3_{eq}$ Neu5Ac),1.92-2.02 (m,-CH_3 GlcNAc , CH_3 Neu5Ac), 1.82 (b, t, $H3_{ax}$, Neu5Ac), 1.63 (m, dCH_2, Lys , b-CH2-Lys), 1.37 (m, g-CH_2 Lys), 1.31 (d,-CH_3 Ala), 1.06 (d,- CH_3 Thr), 0.85 (d,-CH_3 Val).

3.0 Supplementary Figures



S. Figure 1 UHPLC and ^{1}H-NMR Comparison of SGP obtained from egg yolk powder and from fresh eggs. Base peak chromatogram for the HILIC-MS separation of crude glycopeptides. A) Dry egg yolk B) Fresh egg yolk. 500 MHz ^{1}H-NMR C) for SGP from dry yolk, D) SGP from fresh egg yolk.

---

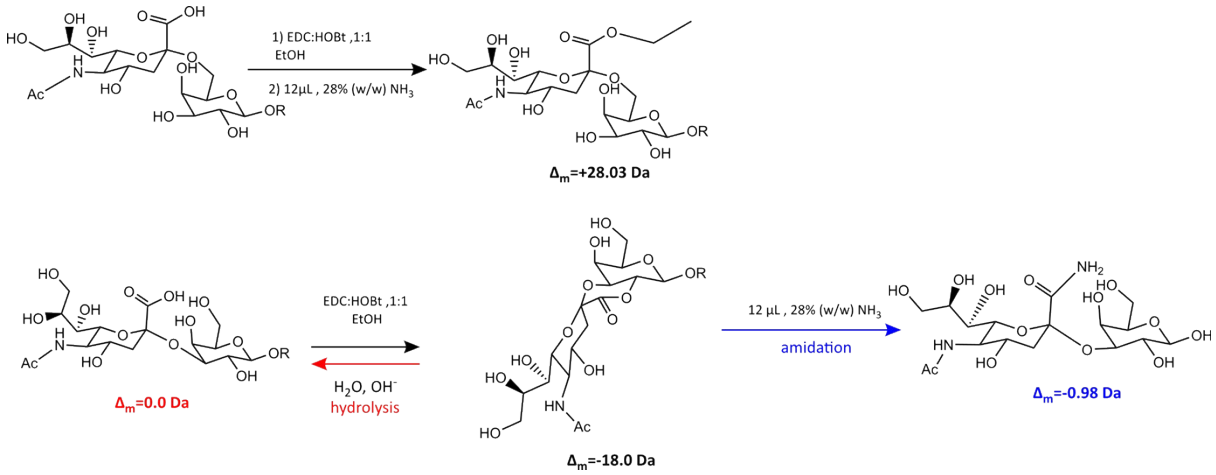
^{1}H-NMR is consistent with that has been reported for pure SGP, however from chromatographic allows to detect other structures are present which cannot be resolved by ^{1}H-NMR due to their structural similarity.

3.1 Regioselectivity of EEA

According to their linkage sialic acids will behave differently towards the Ethyl esterification-amidation conditions in Scheme 1[9]. To check the selectivity of the EEA derivatization strategy three different procainamide labelled standards were processed, representing different types of sialic acid linkage.

S. Table 3 Glycan Standards used for specificity determination of EEA

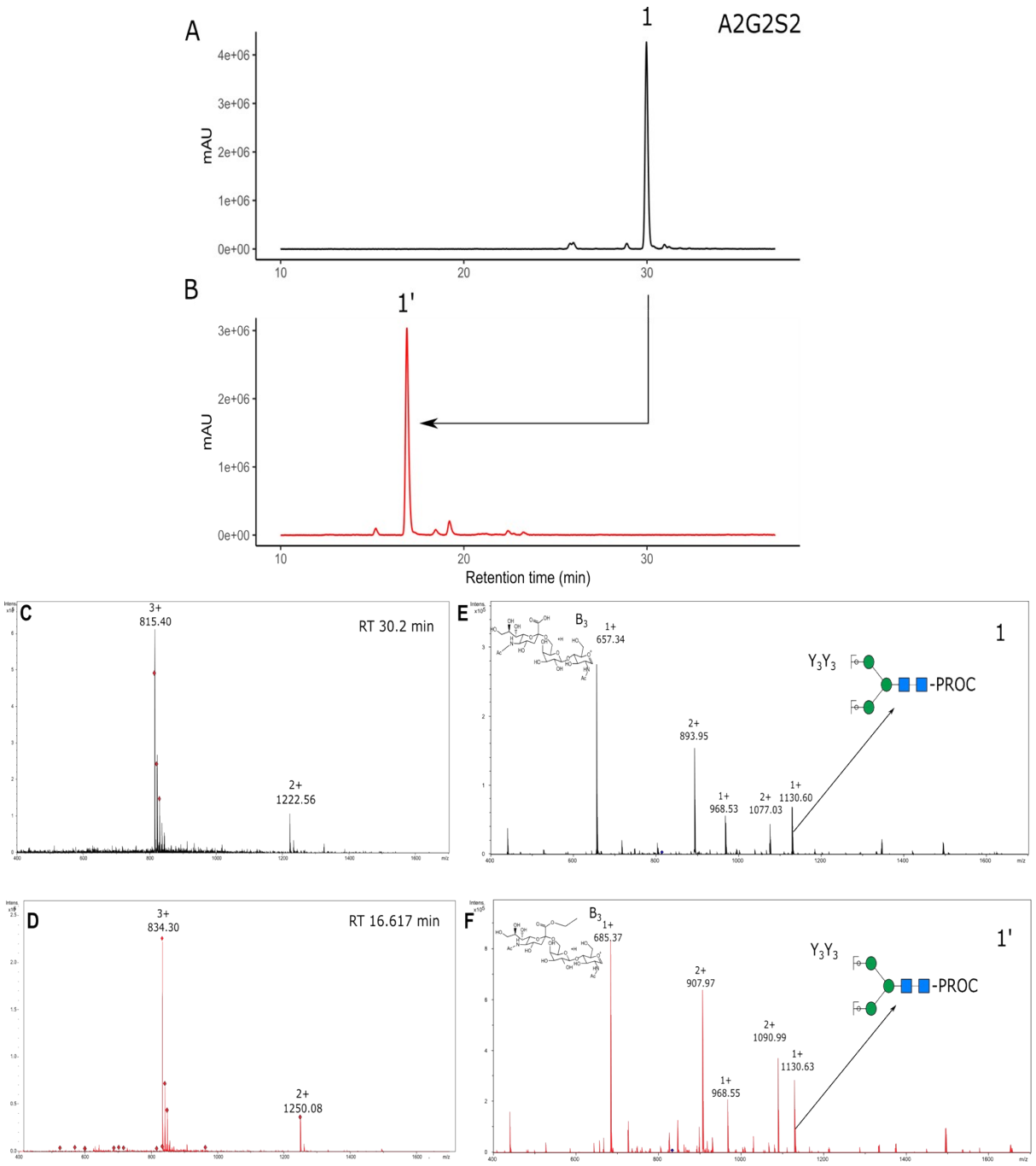
Glycan	Other names	Linkage
A2G2S2	A2	2,6
FA2G2S2	A2F	2,6
A3-1-A3G3S2(3)S1(6)	A3	2,6 and 2,3
A3-2-A3G3S1(3)S2(6)	A3	2,6 and 2,3



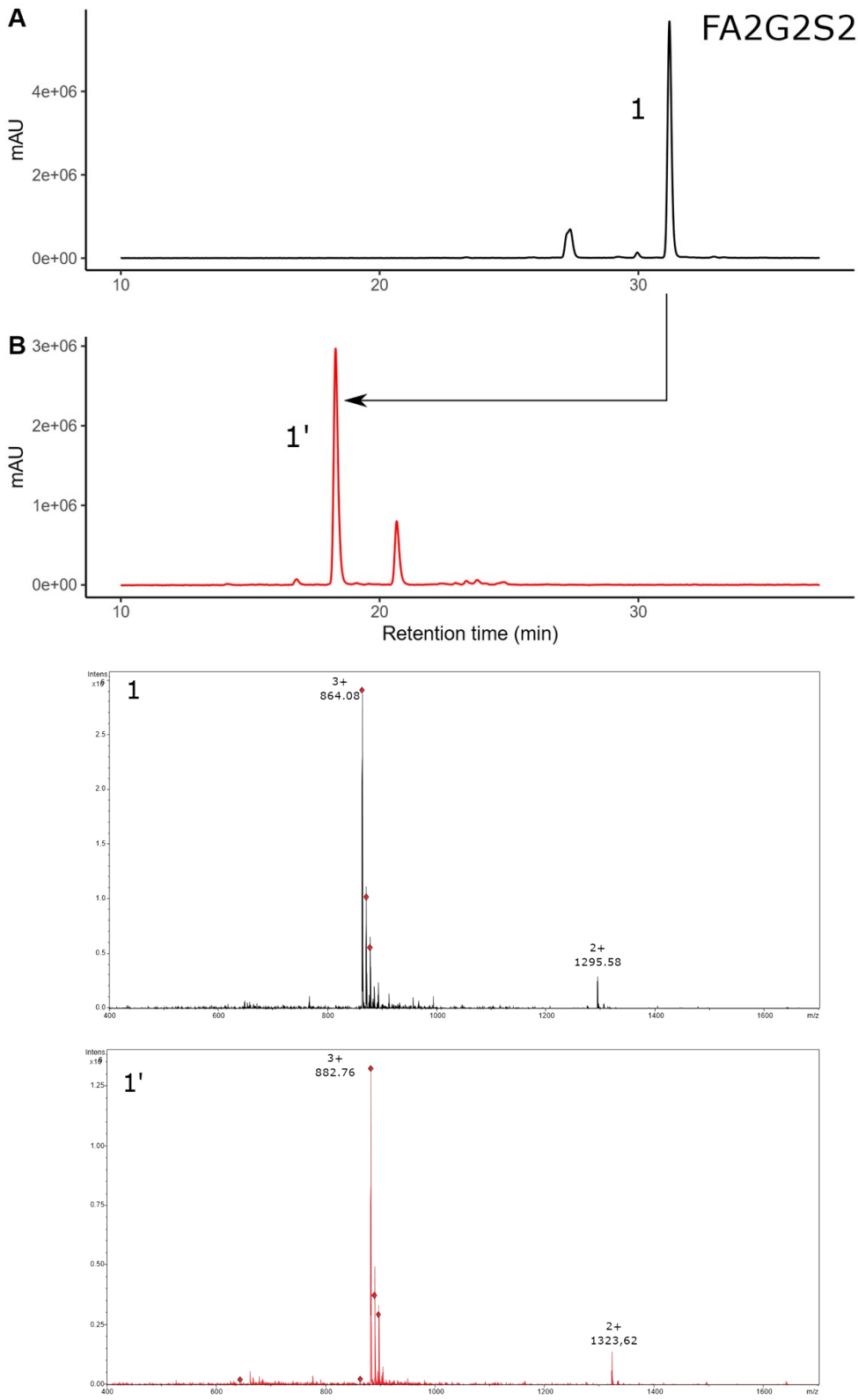
Scheme 1 Reaction outcomes of EEA reaction. A) 2,6 linked sialic acid. B) 2,3 linked sialic acid.

S. Table 4 Mass changes due to EEA observed for each of the different standards S. Fig2-4.

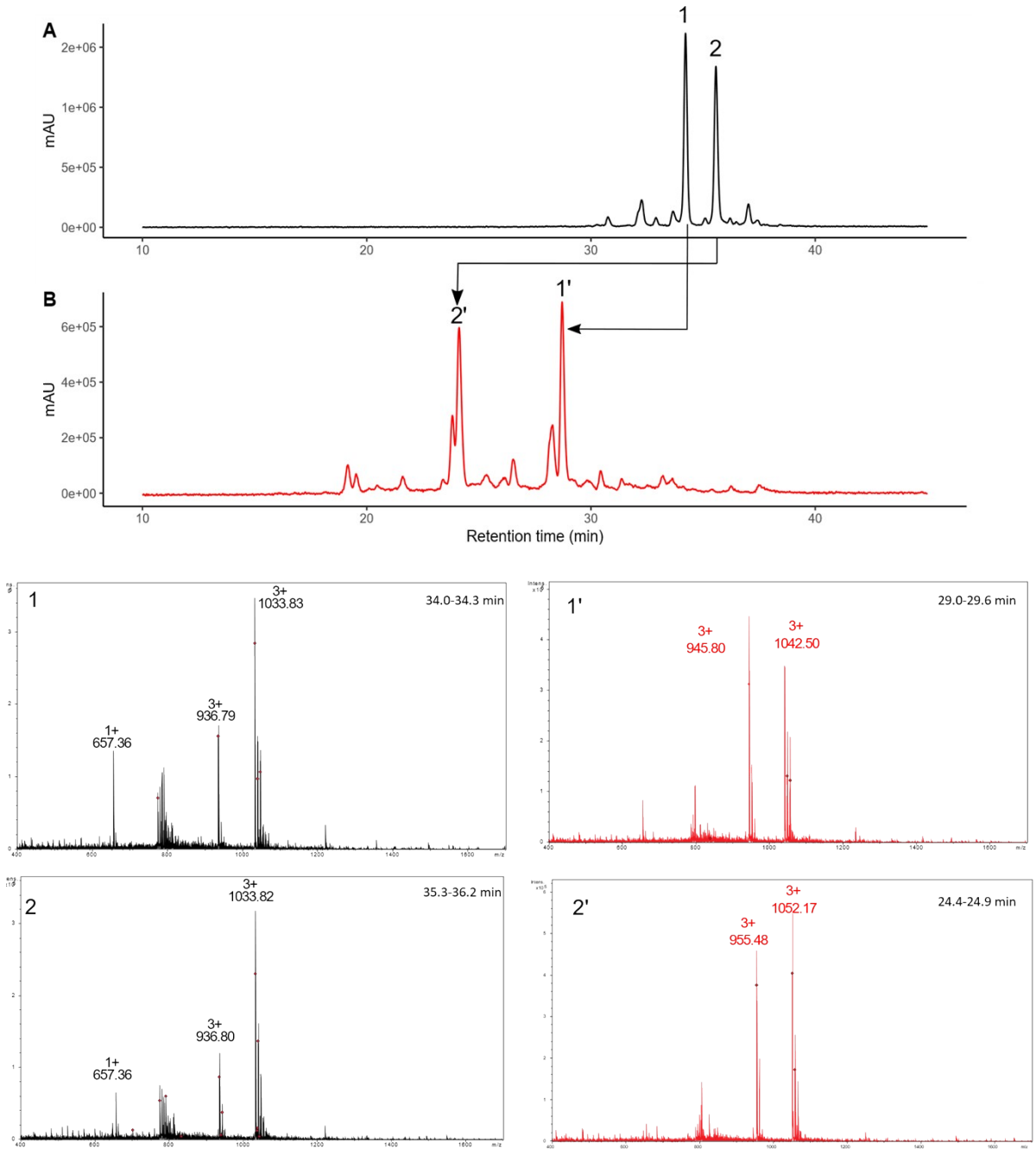
Glycan	Expected monoisotopic m/z before EEA $[M+3H]^{3+}$	Observed monoisotopic m/z before EEA $[M+3H]^{3+}$	ppm	Expected monoisotopic m/z after EEA $[M+3H]^{3+}$	Observed monoisotopic m/z after EEA $[M+3H]^{3+}$	ppm
A2G2S2	814.99	815.4	-501	833.68	834.3	-744
FA2G2S2	863.68	864.08	-465	882.37	882.76	-447
A3-1-A3G3S2(3)S(6)	1033.73	1033.83	-92	1042.42	1042.5	-74
A3-2-A3G3S (3)S2(6)	1033.73	1033.82	-83	1052.09	1052.17	-72



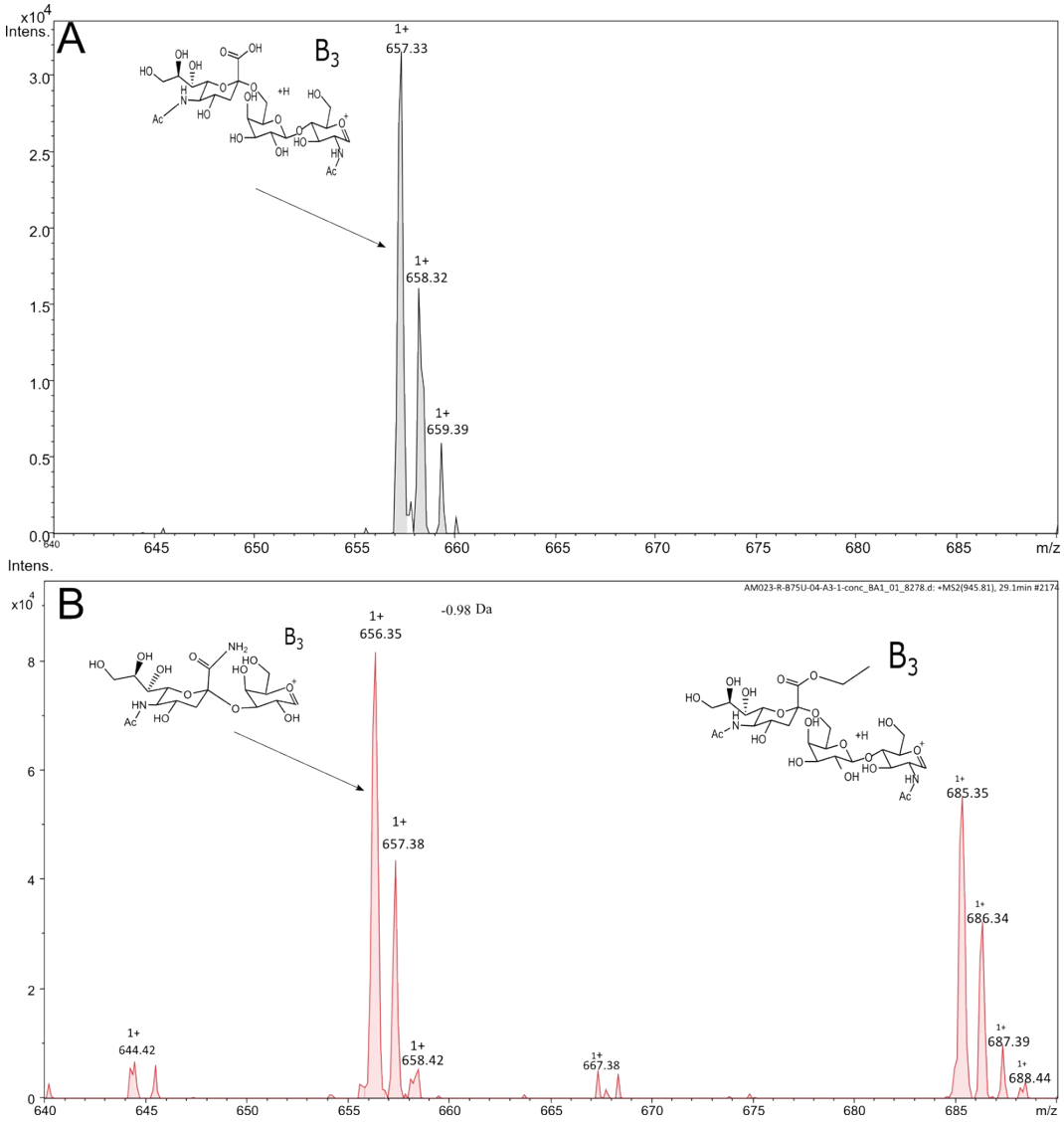
S. Figure 2 UHPLC analysis of procainamide labelled A2G2S2, before and after EEA. A) UHPL trace for procainamide labelled A2G2S2 standard. B) UHPLC trace for EEA-procainamide labelled A2G2S2 standards. C) ESI-MS procainamide labelled A2G2S2 standard. D) ESI-MS for EEA-procainamide label A2G2S2 E) ESI-MS/MS (815.41 m/z) from procainamide labelled A2G2S2 standard. F) ESI-MS/MS (834.12 m/z) from EEA-procainamide labelled A2G2S2(6) standard.



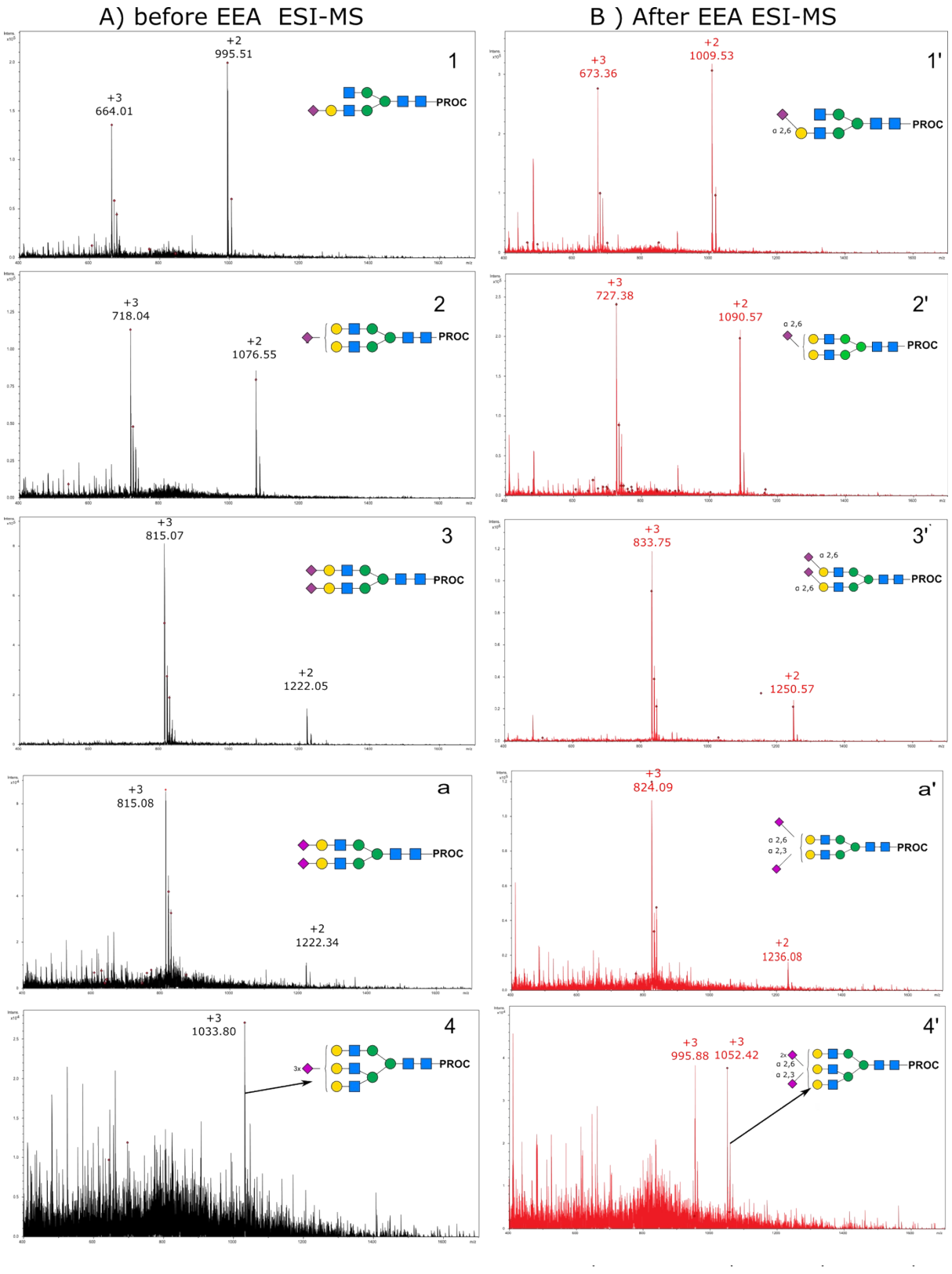
S. Figure 3 UHPLC analysis of procainamide labelled FA2G2S2, before and after EEA. A) UHPL trace for procainamide labelled FA2G2S2 standard. B) UHPLC trace for EEA-procainamide labelled FA2G2S2 standards. 1) ESI-MS procainamide labelled FA2G2S2 standard. 1') ESI-MS for EEA-procainamide label FA2G2S2(6)



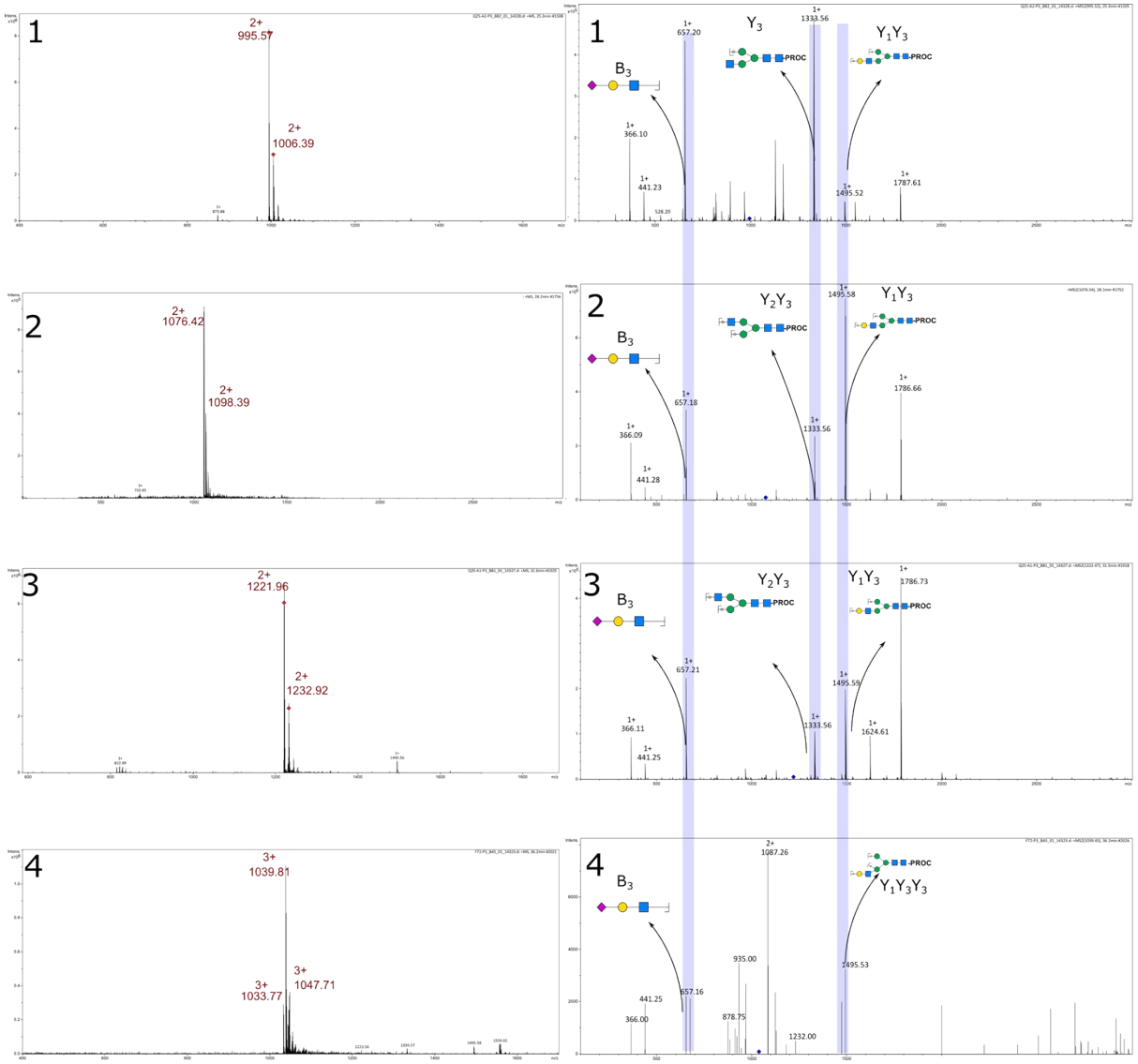
S. Figure 4UHPLC analysis of procainamide labelled A3G3S3, before and after EEA. A) UHPL trace for procainamide labelled A3G3S3 standard. B) UHPLC trace for EEA-procainamide labelled A3G3S3 standards. 1) ESI-MS procainamide labelled A3G3S3 RT 34.0-34.2 min standard. 1') ESI-MS for EEA-procainamide label A3G3S2(3)S(6). 2) ESI-MS procainamide labelled A3G3S3 RT 35.3-36.2 min 2') ESI-MS for EEA-procainamide label A3G3S1(3)S2(6).



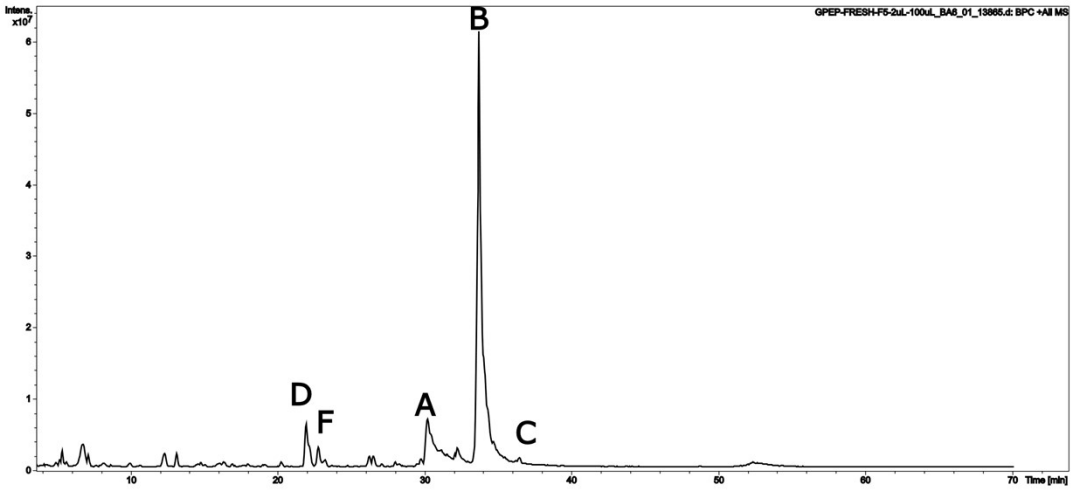
S. Figure 5 ESI- MS/MS procainamide labelled A3G3S3 A) ESI-MS/MS (1034.20) procainamide labelled A3G3S3 RT 35.6 min, peak 2 S.Figure 4, standard. B) ESI-MS/MS for EEA-procainamide label A3G3S1(3)S2(6). (945.81) RT 29.1 min, peak 2' S.Figure 4.



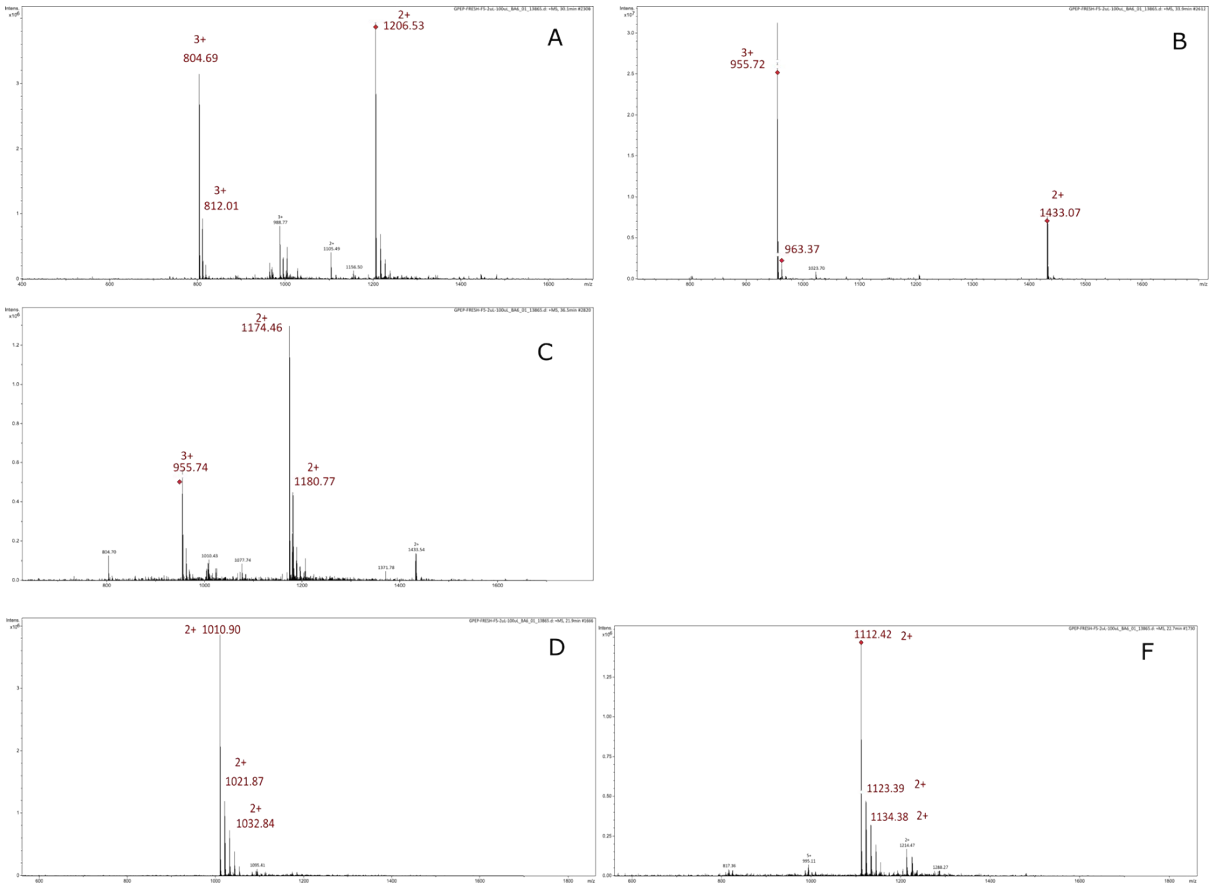
S. Figure 6 ESI-MS of procainamide labelled N-glycans from Figure 2. Colum A). ESI-MS of procainamide labelled N-glycans. Column B). ESI-MS of EEA-procainamide labelled glycan.



S. Figure 7 ESI-MS of procainamide labelled N-glycans from Figure 3A.

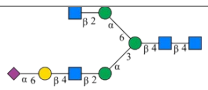
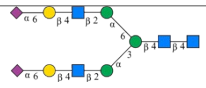
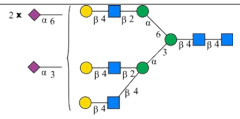
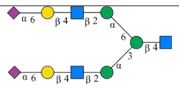
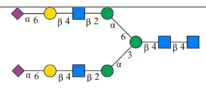


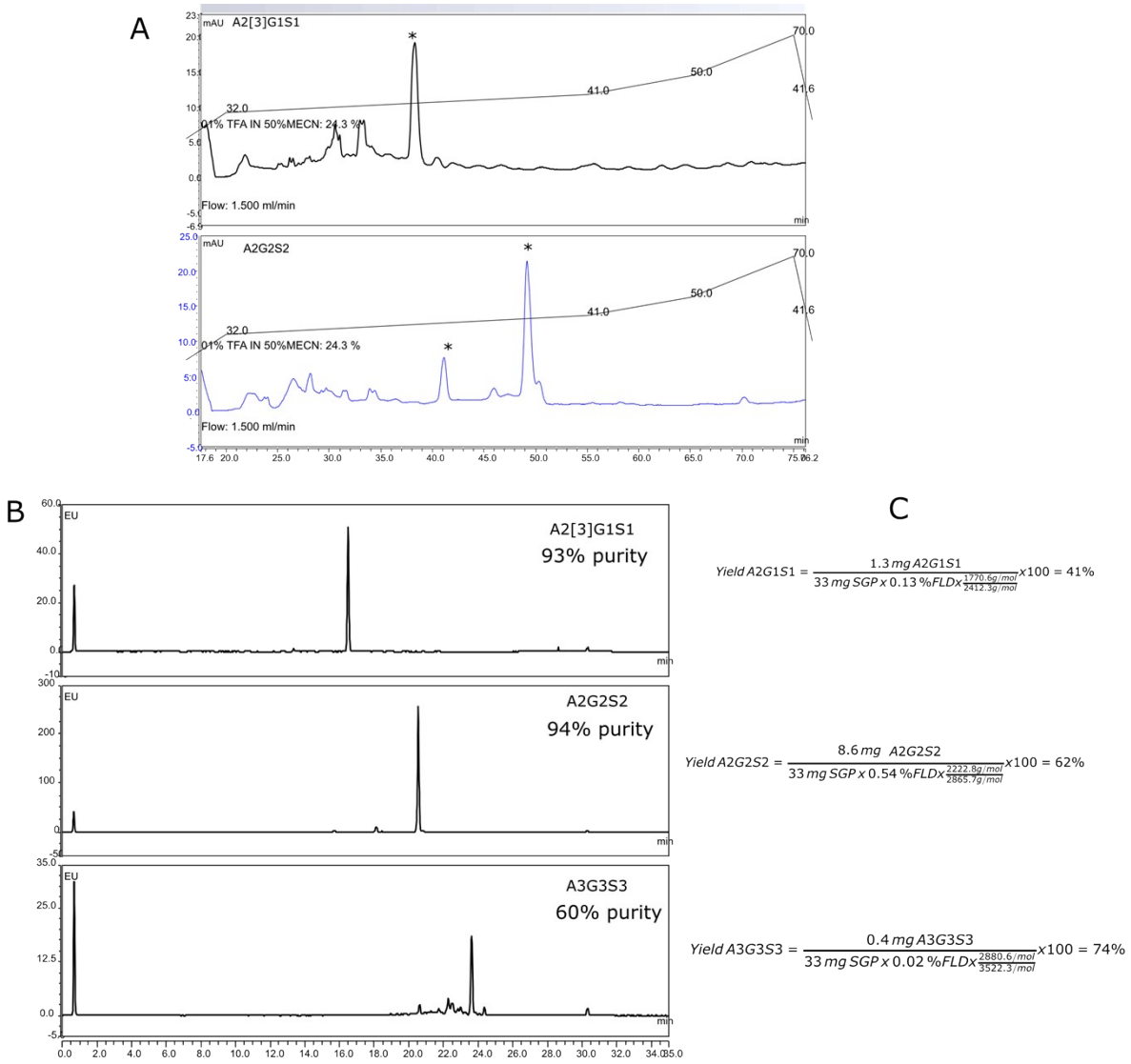
S. Figure 8 Base peak chromatograms for the UHPLC-HILIC-MS separation of the crude egg yolk glycopeptides.



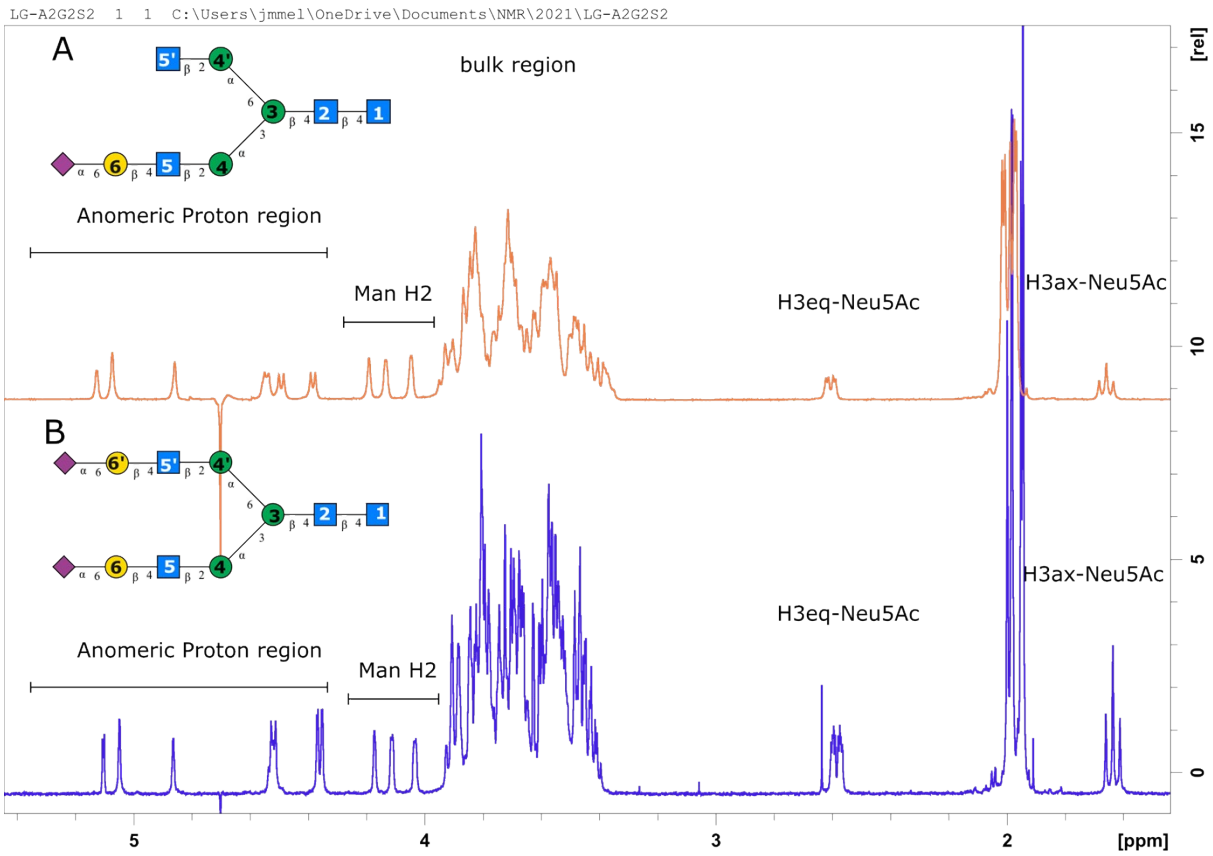
S. Figure 9 ESI-MS of Glycopeptides (A-C) or glycan (D-F) species identified on S. Figure7

S.Table 5 Identification of the different glycopeptide species detected by UHPLC HILIC-MS analysis of crude SGP from S.Figure 8

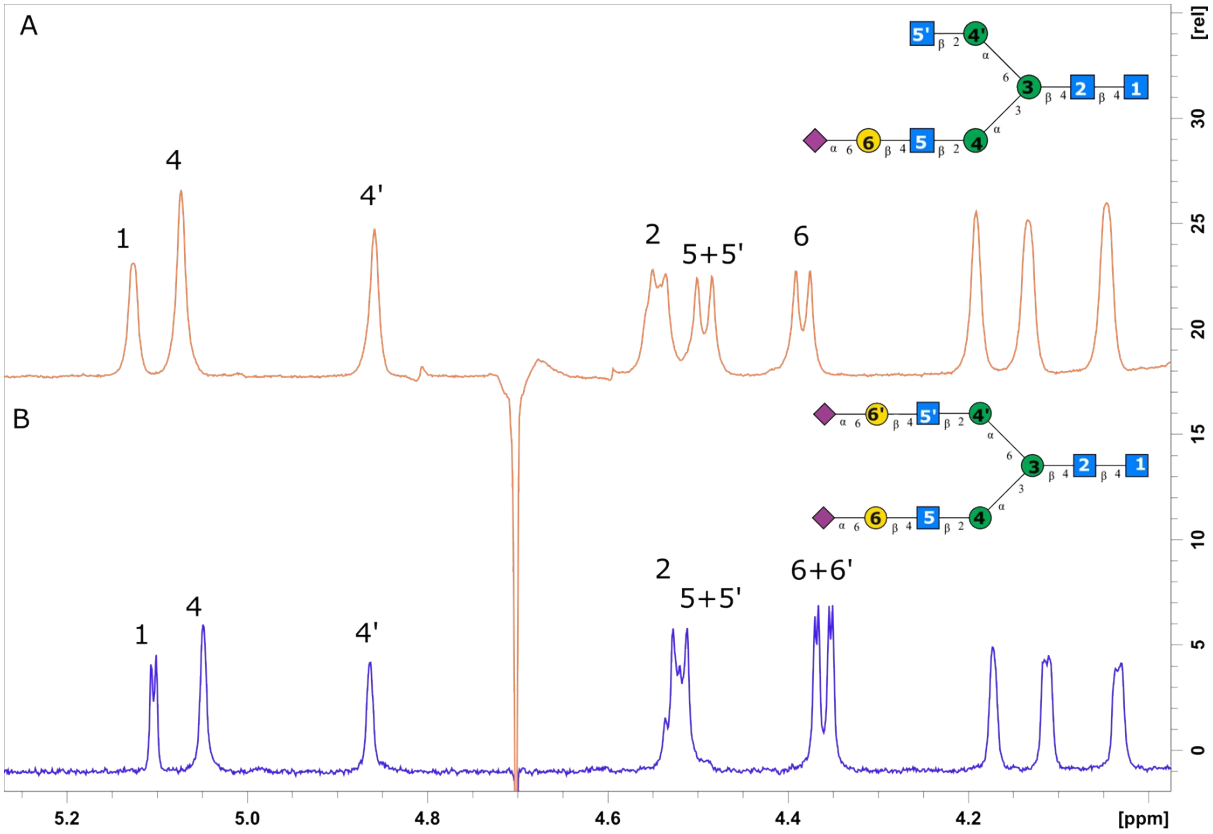
Peak	Composition	N-glycan part	Predicted monoisotopic m/z	Observed monoisotopic m/z
A	KVA**N**KT-A2[3]G1S1		804.69[M+3H]$^{+3}$	804.69[M+3H]$^{+3}$
B	KVA**N**KT-A2G2S2		955.74[M+3H]$^{+3}$	995.72[M+3H]$^{+3}$
C	KVA**N**KT-A3G3S3		1174.46[M+3H]$^{+3}$	1174.46[M+3H]$^{+3}$
D	A2G2S2-(GlcNAc)		1010.86 [M+2H]$^{+2}$	1010.90[M+2H]$^{+2}$
F	A2G2S2		1112.40[M+2H]$^{+2}$	1112.41[M+2H]$^{+2}$



S.Figure 10 A) UV 214 nm HPLC PGC preparative purification of A2[3]G1S1 and A2G2S2, peaks identified with an asterisk correspond to glycan containing peaks B) UHPLC 2AB labelling of PGC purified N-glycans C) Purification yield estimation, based on the abundance of each glycan species obtained from procainamide labelling.

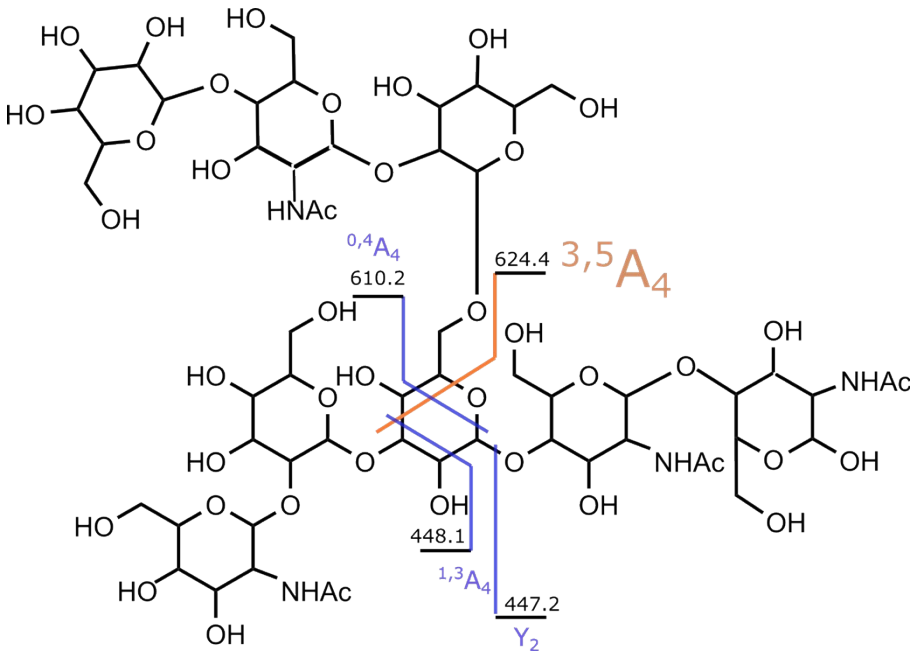


S. Figure 11 ^{1}H-NMR 500 MHz in D_2O of the purified N-glycans A) A2[3]G1S1 B) A2G2S2

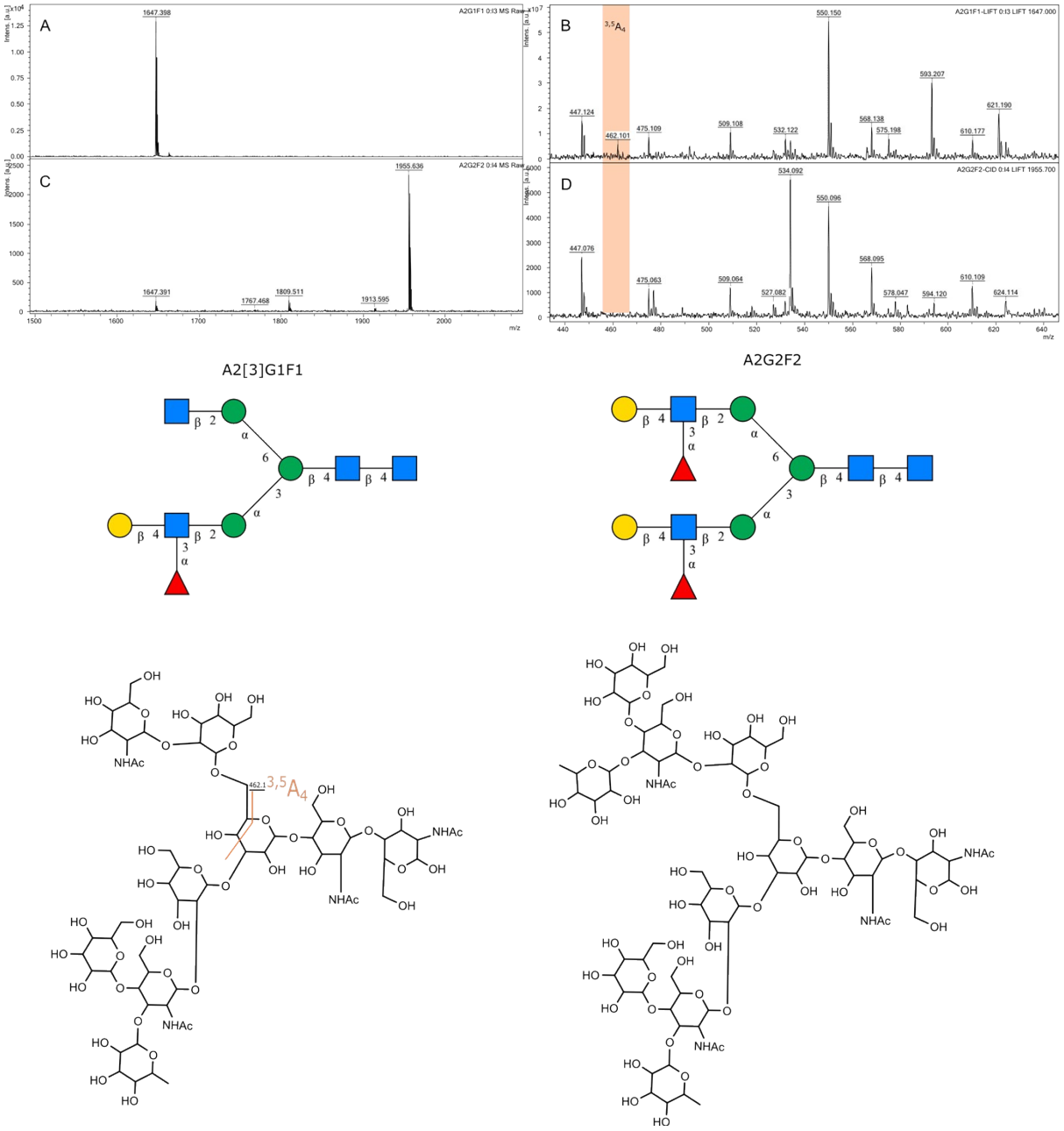


S. Figure 12 ^{1}H-NMR 500 MHz in D_2O of the anomeric region of the purified N-glycans A) A2[3]G1S1 B) A2G2S2

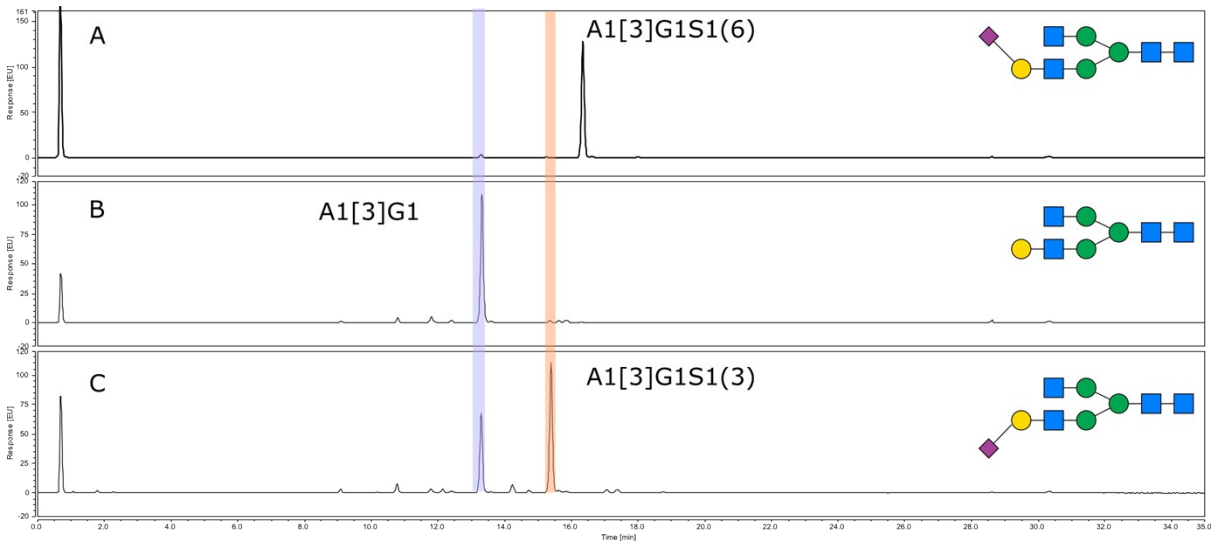
.



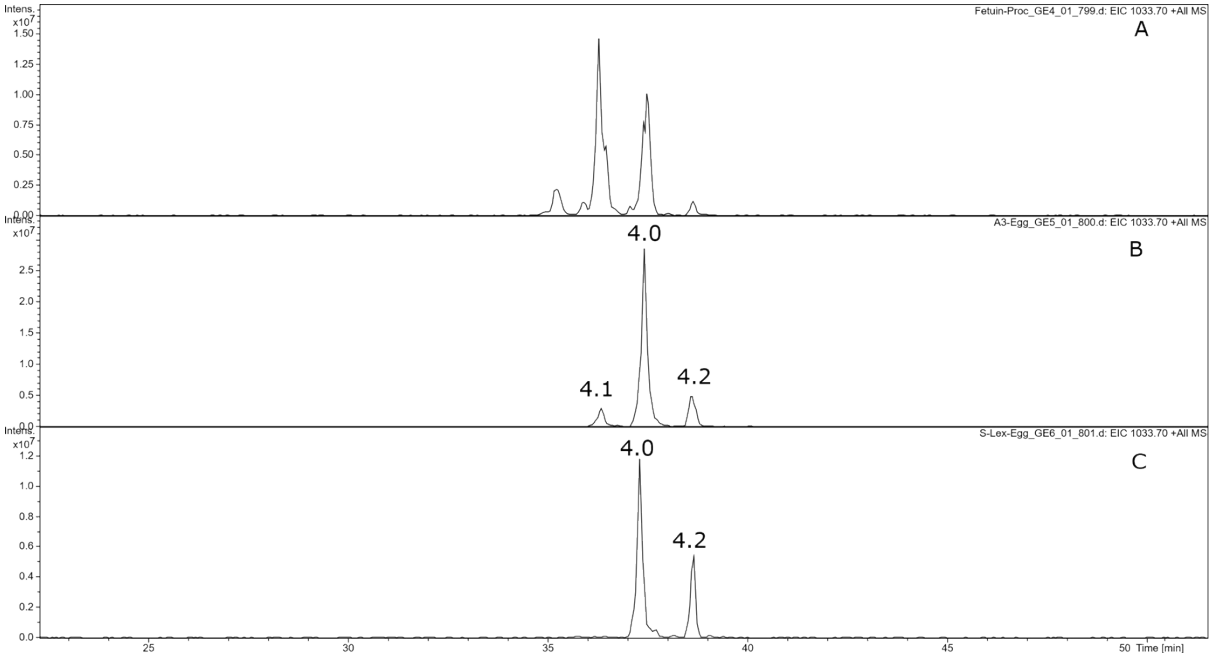
S. Figure 13 MALDI-MS/MS expected fragmentation for A2[6]G1



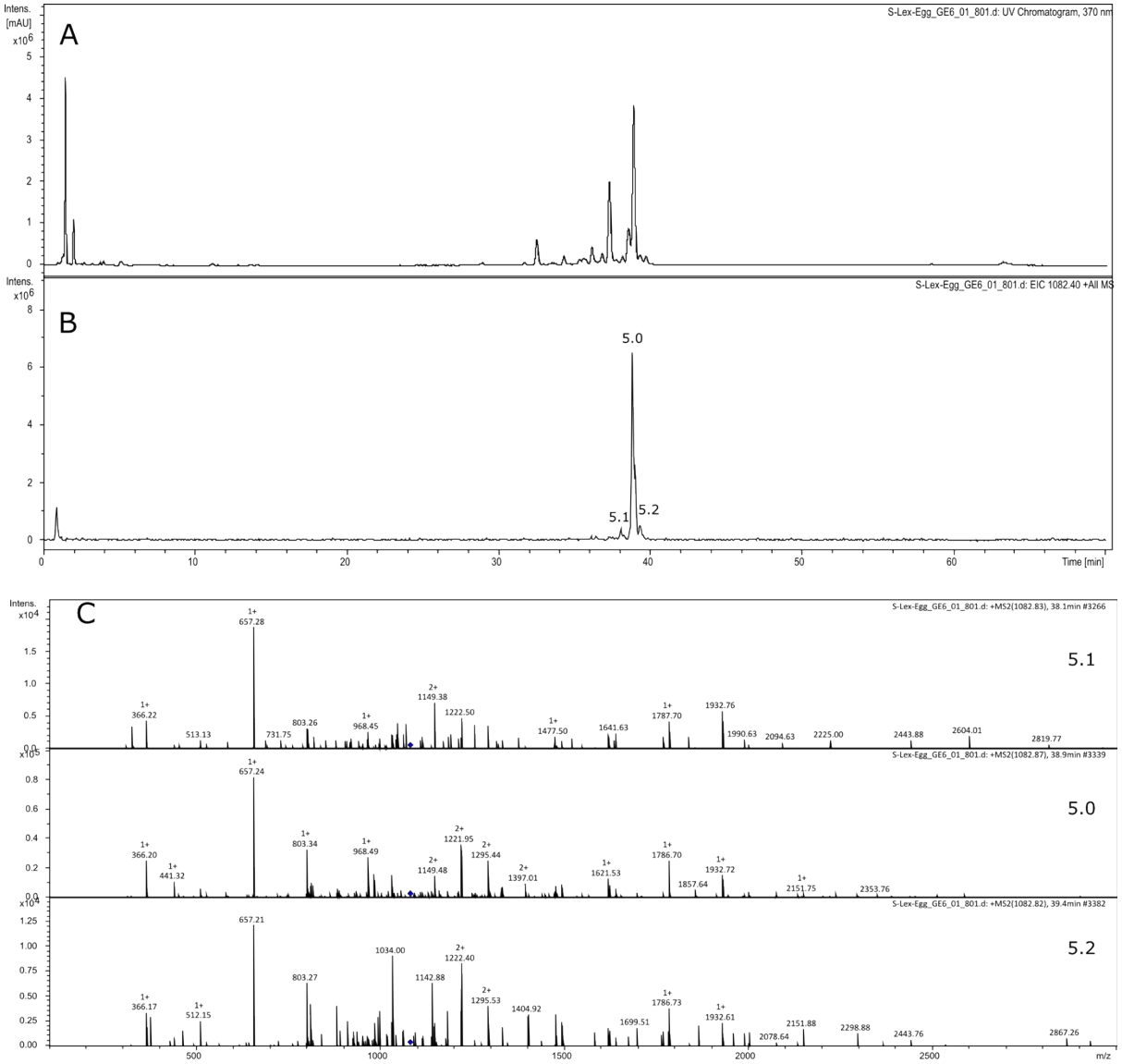
S. Figure 14. MALDI MS and MS/MS analysis for A2[3]G1F and A2G2F2 A) MALDI MS for A2G1F B) MS/MS for A2G1F C) MALDI MS for A2G2F2, D) MALDI MS/MS of A2G2F2



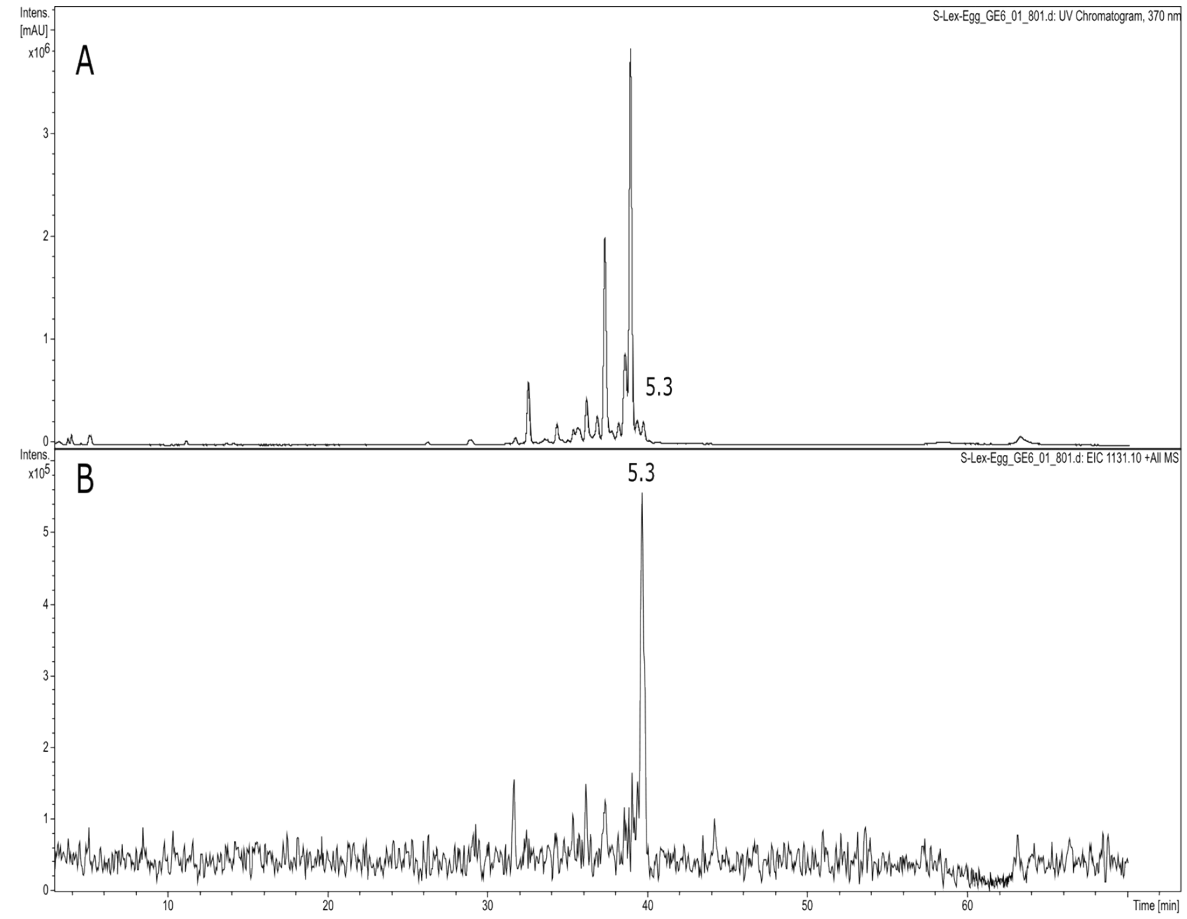
S. Figure 15 Retention time-based confirmation of sialic acid linkage, using a chemoenzymatic produced glycan. A) Initial A2[3]G1S1(6). B) A2[3]G1, obtained after sialic acid removal. C) A2[3]G1S1(3), obtained after sialylation reaction



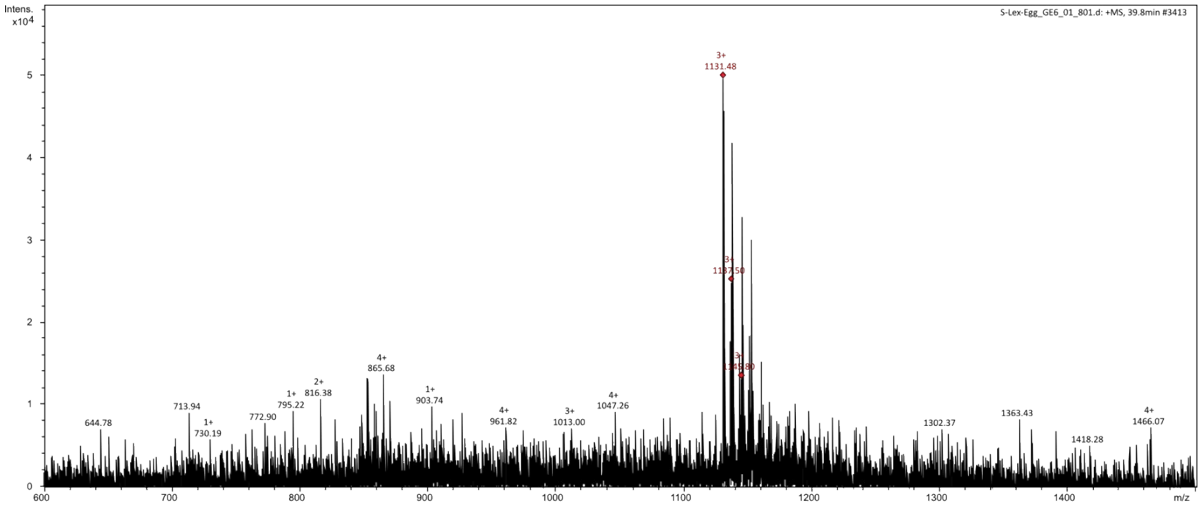
S. Figure 16 Extracted Ion chromatogram at 1033.70 m/z, A) Procainamide labelled N-glycans from bovine fetuin. B) Procainamide labelled N-glycans from S3 fraction C) procainamide labelled N-glycans after antennary fucosylation. Different isomers are detected 4.1,4.0. 4.2 Correspond to a triantennary glycan with two, one and zero 2,3 linked sialic acid, respectively.



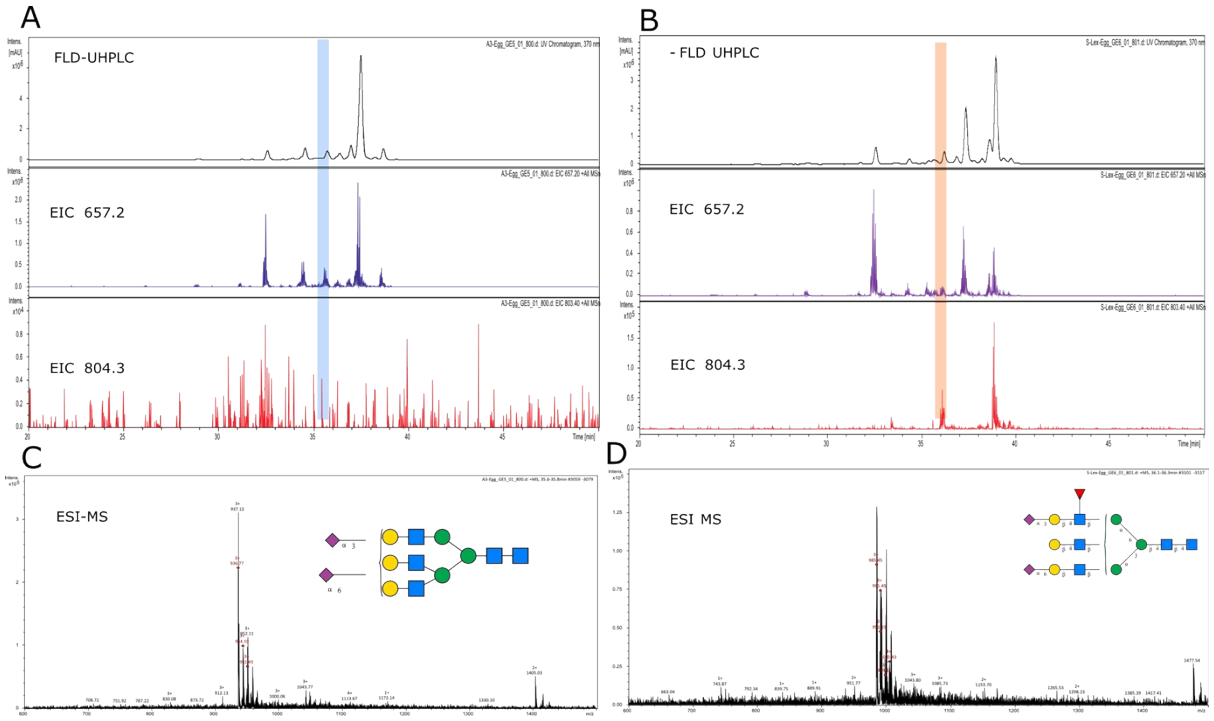
S. Figure 17 A) procainamide labelled N-glycans after antennary fucosylation B) Extracted ion chromatogram at 1082.70 m/z for the UHPLC trace of procainamide labelled N-glycans after fucosylation reaction C) ESI-MS/MS of the different isomeric species. 5.1, 5.0, 5.2.



C



S. Figure 18 A) procainamide labelled N-glycans after antennary fucosylation. B) Extracted ion chromatogram a 1131.10 m/z corresponding to the doubly fucosylated product. C) ESI-MS for 5.3.



S. Figure 19Additional side reaction detected during fucosylation A) Extracted ion chromatogram at 657.2 m/z and 804.3 m/z before fucosylation reaction. B) Extracted ion chromatogram at 657.2 m/z and 804.3 m/z after fucosylation reaction C) ESI-MS of potential fucose acceptor D) ESI-MS of additional fucosylated product.

References

1 L. Liu, A. R. Prudden, C. J. Capicciotti, G. P. Bosman, J. Y. Yang, D. G. Chapla, K. W. Moremen and G. J. Boons, Streamlining the chemoenzymatic synthesis of complex N-glycans by a stop and go strategy, *Nature Chemistry*, 2019, **11**, 161–169.

2 L. Liu, A. R. Prudden, G. P. Bosman and G. J. Boons, Improved isolation and characterization procedure of sialylglycopeptide from egg yolk powder, *Carbohydrate Research*, 2017, **452**, 122–128.

3 R. Schade, E. G. Calzado, R. Sarmiento and P. A. Chacana, Chicken Egg Yolk Antibodies (IgY-technology): A Review of Progress in Production and Use in Research and Human and Veterinary Medicine, *ATLA*, 2005, **33**, 1–25.

4 D. Pedersen and C. Rosenbohm, Dry Column Vacuum Chromatography, *Synthesis (Stuttg)*, 2002, **2001**, 2434–2434.

5 G. S. M. Lageveen-Kammeijer, N. de Haan, P. Mohaupt, S. Wagt, M. Filius, J. Nouta, D. Falck and M. Wuhrer, Highly sensitive CE-ESI-MS analysis of N-glycans from complex biological samples, *Nature Communications*, , DOI:10.1038/s41467-019-09910-7.

6 B. Ma, G. Wang, M. M. Palcic, B. Hazes and D. E. Taylor, C-terminal amino acids of Helicobacter pylori α1,3/4 fucosyltransferases determine Type I and Type II transfer, *Journal of Biological Chemistry*, 2003, **278**, 21893–21900.

7 H. Tsukamoto, Y. Takakura and T. Yamamoto, Purification, cloning, and expression of an α/β-galactoside α-2,3-sialyltransferase from a luminous marine bacterium, Photobacterium phosphoreum, *Journal of Biological Chemistry*, 2007, **282**, 29794–29802.

8 A. Ceroni, K. Maass, H. Geyer, R. Geyer, A. Dell, S. M. Haslam and S. Haslam, GlycoWorkbench: A Tool for the Computer Assisted Annotation of Mass Spectra of Glycans*, *J Proteome Res*, 2008, **7**, 1650–1659.

9 J. I. Furukawa, H. Hanamatsu, T. Nishikaze, H. Manya, N. Miura, H. Yagi, I. Yokota, K. Akasaka-Manya, T. Endo, M. Kanagawa, N. Iwasaki and K. Tanaka, Lactone-Driven Ester-to-Amide Derivatization for Sialic Acid Linkage-Specific Alkylamidation, *Analytical Chemistry*, 2020, **92**, 14383–14392.