Analytica Chimica Acta 1179 (2021) 338840

Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

2-Dimensional ultra-high performance liquid chromatography and DMT-MM derivatization paired with tandem mass spectrometry for comprehensive serum N-glycome characterization

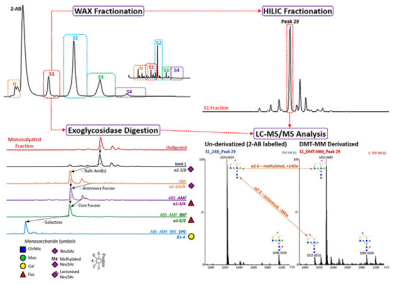
Josh Smith [a, b], Silvia Millán-Martín [a], Stefan Mittermayr [a], Vivian Hilborne [c], Gavin Davey [b], Karol Polom [d, e], Franco Roviello [f], Jonathan Bones [a, g, *]

[a] Characterisation and Comparability Laboratory, The National Institute for Bioprocessing Research and Training, Foster Avenue, Mount Merrion, Co. Dublin, A94 X099, Ireland
[b] School of Biochemistry and Immunology, Trinity Biomedical Sciences Institute, Trinity College Dublin, Dublin 2, D02 R590, Ireland
[c] Northeastern University, 360 Huntington Avenue, Boston, MA, 02115, USA
[d] Department of General Surgery and Surgical Oncology, University of Siena, Siena, Italy
[e] Department of Surgical Oncology, Medical University of Gdansk, Gdansk, Poland
[f] Department of Medicine, Surgery and Neurosciences, University of Siena, Siena, Italy
[g] School of Chemical and Bioprocess Engineering, University College Dublin, Belfield, Dublin 4, D04 V1W8, Ireland

HIGHLIGHTS

- In depth characterization of the serum N-glycome using multidimensional chromatography and high resolution mass spectrometry.
- Identification of low level modified N-glycans bearing acetylation and sulphation.
- Application of label free quantitation for differential glycomics.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 16 April 2021
Received in revised form
2 July 2021
Accepted 4 July 2021
Available online 9 July 2021

Keywords:
Glycosylation
Sialic acid
Glycomics
Mass spectrometry
Liquid chromatography
Characterization

ABSTRACT

Glycosylation is a prominent co- and post-translational modification which contributes to a variety of important biological functions. Protein glycosylation characteristics, particularly N-glycosylation, are influenced by changes in one's pathological state, such as through the presence of disease, and as such, there is great interest in N-glycans as potential disease biomarkers. Human serum is an attractive source for N-glycan based biomarker studies as circulatory proteins are representative of one's physiology, with many serum proteins containing N-glycosylation. The difficulty in comprehensively characterizing the serum N-glycome arises from the absence of a biosynthetic template resulting in great structural heterogeneity and complexity. To help overcome these challenges we developed a 2-dimensional liquid chromatography platform which utilizes offline weak anion exchange (WAX) chromatography in the first dimension and hydrophilic interaction liquid chromatography (HILIC) in the second dimension to separate N-glycans by charge, corresponding to degree of sialylation, and size, respectively. Performing these separations offline enables subsequent derivatization with 4-(4,6-Dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride (DMT-MM) for sialic acid linkage determination and the identification of

* Corresponding author. Characterisation and Comparability Laboratory, The National Institute for Bioprocessing Research and Training, Foster Avenue, Mount Merrion, Co. Dublin, A94 X099, Ireland.
E-mail address: jonathan.bones@nibrt.ie (J. Bones).

https://doi.org/10.1016/j.aca.2021.338840

sialic acid linkage isomers. Subsequent tandem mass spectrometry analysis revealed the identification of 212 complete and partial N-glycan structures including low abundant N-glycans containing acetyl and sulphate modifications. The identifications obtained through this platform were then applied to N-glycans released from a set of stage 3 gastric cancer serum samples obtained from patients before (pre-op) and after (post-op) tumour resection to investigate how the serum N-glycome can facilitate differentiation between the two pathological states.

1. Introduction

Glycosylation is a prominent co- and post-translational protein modification known to play an integral role in a wide variety of biological functions [1]. Changes in protein glycosylation have been shown to occur in many disease states including cancer, various genetic diseases and neurodegenerative diseases such as Alzheimer's, creating significant interest in potentially utilizing glycans and glycoproteins for disease biomarkers and therapeutics [2–8]. Human serum is a common biological source used in biomarker studies and contains a large number of glycosylated proteins [9]. Complete characterization of the human serum glycome is a key objective to improve our understanding of how changes in biological states can impact protein glycosylation and thus protein function. However, the large variety of glycan structures, their associated diverse micro-heterogeneity and the lack of a standard biological template makes glycan characterization a difficult task.

The structural diversity and complexity of N-glycans makes the complete characterization of the serum N-glycome difficult to perform. Multiple studies have been conducted to achieve this aim, but the number of N-glycans identified during these studies has widely varied [10–18]. A variety of reasons can account for this disparity, but the greatest is the technological advancements that have improved analytical separation and analysis capabilities, allowing for greater glycan separation along with better detection of low abundant glycans. Improved technologies, such as sub 2 μm and core shell hydrophilic interaction column stationary phases developed in parallel with ultra-high pressure liquid chromatography (UHPLC) platforms have greatly improved glycan separations capabilities. The development of high-resolution hybrid mass spectrometry (MS) instruments such as quadrupole-time of flight (Q-ToF) and quadrupole-Orbitrap MS has also aided in improving structural identification, as generated tandem MS (MS^2) fragments can help distinguish between positional and linkage glycan isomers. Despite these advancements, separations platforms capable of obtaining one glycan per peak have yet to be realized with complex N-glycan samples such as human serum. As such, multi-dimensional separations platforms are still needed for mining the depth of the serum N-glycome.

Here, a 2-dimensional liquid phase separation platform paired with MS^2 analysis was employed to perform comprehensive profiling of the serum N-glycome. Weak anion exchange (WAX) chromatography was performed in the first dimension to separate N-glycans by their charge, which corresponds to degree of sialylation or the presence of other charged substituents, while hydrophilic interaction liquid chromatography (HILIC), which separates glycans by size and hydrophilicity, was applied in the second dimension. During HILIC separation, each peak was collected for subsequent derivatization with 4-(4,6-Dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride (DMT-MM) and UHPLC-MS^2 analysis. DMT-MM derivatization enables the determination of sialic acid linkage for each glycan and thus the detection of sialic acid linkage isomers present in each individually collected peak. We fully characterized 64 structures and partially characterized 148 structures for a total of 212 structures including sialic acid-based isomers. These identifications were then applied to N-glycans released from a set of stage 3 gastric cancer (GC) serum samples obtained from patients before (pre-op) and after tumour resection (post-op) to profile how the glycan profile may facilitate differentiation between the two pathological states. It was detected that the differentiation between the pre-op and post-op states became more pronounced as the stage of the cancer became more severe, with Stage 3C cancer showing complete differentiation between the pre-op and post-op states.

2. Materials and methods

2.1. Chemicals and solvents

All reagents and solvents used were ACS reagent grade or better. Acetic acid (HAc, glacial), urea, iodoacetamide (IAA), ammonium bicarbonate (ABC), formic acid (FA, 98–100%), anthranilamide (2-AB), sodium cyanoborohydride, dimethyl sulfoxide (DMSO), 4-(4,6-Dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride (DMT-MM), ammonium acetate, acetonitrile (ACN – for HPLC, gradient grade, ≥99.9%), Methanol (for HPLC, ≥99.9%) and trizma hydrochloride were all obtained from Sigma Aldrich (St. Louis, MO, USA). 20 × modified Dulbecco's phosphate buffered saline (PBS), dithiothreitol (DTT), water (Optima® LC-MS grade), ACN (Optima® LC-MS grade) and tris base were obtained from Fisher Scientific (Waltham, MA, USA). Peptide N-Glycosidase F (PNGase F), *Arthrobacter ureafaciens* α(2–3/6/8)-neuraminidase (ABS), *Streptococcus pneumoniae* α(2–3)-neuraminidase (NAN 1), almond meal α(1–3/4)-fucosidase (AMF), bovine kidney α(1–2/6)-fucosidase (BKF), and *S. pneumoniae* β(1–4)-galactosidase (SPG) were purchased from New England Biolabs (Ipswich, MA, USA). Water was obtained from an Arium® pro Ultrapure Water System (Sartorius, Göttingen, Germany). Pooled human serum was obtained from 100 healthy adult male and female blood donors (U.K. Blood Transfusion Service). Cancer patient serum samples were obtained with informed patient consent from individual male and female patients with stage 3 GC (Department of Surgical Oncology, University of Siena, Italy).

2.2. Serum depletion (2P depletion)

Serum was depleted of human albumin and Immunoglobulin G (IgG, IgG 1, 2 and 4 isotypes) on a 4.6 mm i.d. × 50 mm Multiple Affinity Removal System (MARS) Human Albumin column (Agilent – Santa Clara, CA, USA) in series with a 0.7 cm i.d. × 3.5 cm, 1 mL HiTrap Protein A HP column (GE, Fairfield, CT, USA) using an Agilent 1200 series HPLC (Santa Clara, CA, USA) consisting of an autosampler, degasser, quaternary pump, column compartment and VWD detector. Instrument settings were as follows: autosampler temperature: 5 °C, column temperature: ambient, VWD wavelength

(λ): 280 nm. Injections of 200 μL were prepared by diluting 40 μL of serum to 200 μL with 1 × PBS. The depletion was performed using a step gradient elution, solvent A was 1 × PBS and solvent B was 0.5 M HAc. Samples were loaded using 100% A at 0.3 mL.min^{-1} (for 6 min) and then from 0.3 mL.min^{-1} to 1.0 mL.min^{-1} (over 2 min) with a 1 min hold. The bound albumin and IgG fraction was eluted using a step gradient to 100% B at 1.0 mL.min^{-1} (over 0.1 min) and held for 3 min. Following elution the column was re-equilibrated using a return to 100% A at 1.0 mL.min^{-1} (over 0.1 min) followed by a 3.7 min hold before return to initial flow rate of 0.3 mL.min^{-1} (over 0.1 min). The unbound fraction (in PBS) was collected at 0 °C on ice. Samples were pooled and concentrated for 60 min at 4000×*g* at 4 °C using Amicon Ultra-15 3 KDa MWCO centrifugal filter devices (Millipore, Bedford, MA, USA). Protein concentration was determined using a BCA protein assay (Thermo Fisher Scientific, Waltham, MA, USA) and concentrated samples were stored at −30 °C if not immediately used for subsequent N-glycan release.

2.3. N-glycan release, 2-AB labelling and clean-up

Glycan release was performed in 0.5 mg aliquots using a modified version of the FASP N-glycan release method [19]. 2P depleted human serum was buffer exchanged into 8 M urea in 0.1 M tris, pH 8.5 solution (UA) and applied to Amicon Ultra 0.5 mL centrifugal filters with a 10 kDa molecular weight cut off (Sigma Aldrich, St. Louis, MO, USA). After 2 washes with UA, the sample was reduced with DTT in UA at a final concentration 10 mM DTT for 30 min at 65 °C on a thermomixer with shaking at 600 rpm and, following the removal of DTT by centrifugation, 14,000×*g*, alkylated with IAA in UA at a final concentration 50 mM IAA for 30 min at room temperature (RT) in dark. After removal of IAA, buffer exchange in to 50 mM ABC, pH 8.0 was performed with 3 washes of ABC. Overnight (16 h) N-glycan release using 1000 units of PNGase F was performed through incubation of samples on filters at 40 °C on a thermomixer (Eppendorf, Hamburg, Germany) shaking at 600 rpm. Released glycans were collected and samples on filters were washed twice with water to ensure total glycan collection. Collected glycans were evaporated to dryness *via* vacuum centrifugation (Thermo, Waltham, MA, USA), reconstituted in 20 μL of 1% (v.v^{-1}) formic acid and incubated for 40 min at RT to promote conversion of the released glycosylamines to reducing sugars and subsequently dried *via* vacuum centrifugation. Dried samples were stored at −30 °C if not immediately used for 2-AB labelling [20].

2-AB labelling was carried out using a reaction mixture of 0.37 M 2-AB and 0.95 M sodium cyanoborohydride in DMSO containing 30% (v.v^{-1}) HAc. Dried glycans were mixed with 5 μL of 2-AB labelling solution and incubated at 65 °C for 2 h on a thermomixer shaking at 600 rpm. Samples were cooled to room temperature and cleaned of excess 2-AB using a HILIC clean-up procedure developed for UHPLC using frontal chromatography on a 2.1 mm i.d. × 50 mm, 1.7 μm BEH Glycan column (Waters Corporation, Milford, MA, USA) fitted with 2.1 mm i.d. × 5 mm XBridge® BEH HILIC XP VanGuard® Cartridge packed with 2.5 μm particles. Clean-up was performed using a Dionex Ultimate 3000 Rapid Separation (RS) UHPLC (Thermo Scientific, Germering, Germany) with a 3000 RS dual system solvent pump, RS autosampler, RS column compartment and RS fluorescence (FLR) detector with a Waters Fraction Collector III for cleaned sample collection (Milford, MA, USA). Instrument settings were as follows: Autosampler temperature: 8 °C, column temperature: 40 °C, λ_{ex} was 330 nm, λ_{em} was 420 nm. Injections of 100 μL were prepared by diluting 2-AB labelled samples to a final volume of 20 μL with water. A user defined injection program was used so that 80 μL of 100% ACN was added into the glycan samples by the autosampler followed by mixing. Clean-up was performed at a constant flow rate of 0.5 mL.min^{-1}. Eluting solvents were A: 50 mM ammonium formate (NH_4HCO_2), pH 4.4 B: 100% ACN and C: water. Loading conditions used: 80% solvent B and 20% solvent A (for 4.5 min). Eluting conditions used: 80% B and 20% A to 20% B and 80% solvent C (over 0.01 min) and held for 2.5 min. Column equilibration conditions: 20% B and 80% C to 80% B and 20% A (over 0.01 min) and held for 3.0 min. Cleaned glycans were collected during elution, dried *via* vacuum centrifugation and stored at −30 °C until analysis.

2.4. WAX chromatography

First dimensional N-Glycan separation was performed using WAX chromatography on a 7.5 mm i.d. × 75 mm, 10 μm, 1000 Å BioSuite DEAE Anion-exchange column (Waters Corporation, Milford, MA, USA) using an Acquity H-class ultra-pressure liquid chromatography (UPLC) platform with a BioQuaternary Solvent manager, BioSample manager-FTN, column compartment, FLR detector and Fraction manager under control of Empower 3 chromatography workstation, build 3471 (Waters Corporation, Milford, MA, USA). Instrument settings were as follows: Sample temperature: 5 °C, column temperature: RT, λ_{ex}: 330 nm, λ_{em}: 420 nm, Fraction manager temperature: 5 °C. Injections of 50 μL were prepared by reconstituting N-glycans released from 1 mg of 2P depleted serum in 50 μL of water. Separations were performed using linear gradient elution at 0.75 mL.min^{-1}, solvent A was 20% (v v^{-1}) ACN and solvent B was 0.1 M ammonium acetate, pH 7.0 in 20% (v v^{-1}) ACN. Eluting conditions used: 100% solvent A isocratic hold for 6 min followed by a linear increase in solvent B from 0% to 100% (over 35 min), followed by a hold for 2.5 min at 100% B. The column was re-equilibrated using a 0.5 min switch back to 100% A followed by a hold for 7 min. Glycans were collected at 5 °C in five fractions: neutral (S0), mono-sialylated (S1), di-sialylated (S2), tri-sialylated (S3), and tetra-sialylated (S4). The fractions collected from all runs collected were pooled, dried *via* vacuum centrifugation and stored at −30 °C for subsequent analysis.

2.5. Exoglycosidase sequencing

A panel of recombinant exoglycosidases were used to sequentially digest N-glycans from each WAX fraction to identify the structural composition of the oligosaccharides contained within. This panel consisted of digests with NAN 1, ABS, ABS + AMF, ABS + AMF + BKF and ABS + AMF + BKF + SPG along with an undigested sample. Digests were performed by adding the exoglycosidases as outlined in the combinations above to 2-AB labelled N-glycans corresponding to 400 μg of 2P depleted starting material and incubating in 50 mM sodium acetate buffer, pH 5.5 overnight at 37 °C. 8 mU of each exoglycosidase were used per digestion with the exception of ABS where 20 mU was used. After digestion, enzymes were removed using the HILIC clean-up procedure described in Section 2.3. Collected glycan eluates were dried *via* vacuum centrifugation and stored at −30 °C for subsequent UHPLC-FLR-MS^2 analysis.

2.6. HILIC separation and DMT-MM derivatization

Second dimensional N-glycan separation was performed using HILIC on a 2.1 mm i.d. × 150 mm, 1.7 μm BEH Glycan column (Waters Corporation, Milford, MA, USA) using a Waters Acquity H-class UPLC as described in Section 2.2.4. Nine 50 μL injections prepared in 75% (v.v^{-1}) acetonitrile in water were prepared from each WAX fraction for HILIC separation. Separations were performed using a linear gradient elution at 0.56 mL.min^{-1} at 40 °C, solvent A was 50 mM NH_4HCO_2, pH 4.4 and solvent B was 100% ACN. The gradient used consisted of an initial isocratic hold at 70%

solvent B for 1.47 min, followed by a linear decrease in solvent B from 70% to 53% over 23.34 min. The column was washed using 30% B for 0.75 min with a simultaneous reduction in flow rate from 0.56 mL.min^{-1} to 0.40 mL.min^{-1}, increased back to initial conditions of 70% B over 0.30 min at 0.40 mL.min^{-1}, followed by a return to initial conditions with a subsequent 1.5 min hold. Each individual peak was collected at 5 °C, with the peaks collected from each injection pooled with their respective peak. Collected fractions were dried *via* vacuum centrifugation and stored at −30 °C for DMT-MM derivatization and UPLC-FLR-MS^2 analysis.

For DMT-MM derivatization, each collected HILIC peak fraction was reconstituted in water, with 50% of the sample aliquoted for DMT-MM derivatization. The DMT-MM aliquoted portion along with the remaining 2-AB labelled sample was dried *via* vacuum centrifugation with the 2-AB labelled sample stored at −30 °C for UPLC-FLR-MS^2 analysis. The DMT-MM aliquoted portion was reconstituted in 0.1 M DMT-MM in methanol solution and incubated at 80 °C for 1 h as described by Tousi, F. et al. [20]. After cooling to RT, samples were dried *via* vacuum centrifugation. Dried samples were then immediately analysed using HILIC-UPLC-FLR-MS^2.

2.7. HILIC-UPLC-FLR-Elevated collision energy MS (MS^E) analysis

All HILIC-UPLC-FLR-MS^E analysis was performed using a Waters Acquity I-Class UPLC (Milford, MA, USA) consisting of a Binary Solvent manager, sample manager-FTN, column compartment and FLR detector coupled to a Waters Xevo G2 QToF mass spectrometer (Waters Corporation, Manchester, UK) fitted with a standard electrospray ionization source. Chromatographic separations were performed on a 1.0 mm i.d. × 150 mm, 1.7 μm BEH Amide column (Milford, MA, USA). UPLC instrument settings were as follows: Sample temperature: 5 °C, column temperature: 60 °C, λ_{ex} was 330 nm, λ_{em} was 420 nm. All analysis of WAX digests and GC patient samples were performed using a flow rate of 0.15 mL .min^{-1}, solvent A was 50 mM NH_4HCO_2, pH 4.4 and solvent B 100% ACN. Elution was performed using a 40 min linear gradient method as follows: initial isocratic hold at 72% B for 1.00 min, followed by a decrease from 72 to 57% B over 30 min, then decreased to 30% B over 1.00 min and back to initial conditions over 4.00 min with subsequent 4.00 min hold. Samples were prepared as 10 μL injections in 75% (v.v^{-1}) acetonitrile. For analysis of the individual peak fractions, 10 μL injections for both the 2-AB labelled and DMT-MM derivatized samples were prepared. 2-AB labelled samples were reconstituted in 75% (v.v^{-1}) acetonitrile while the DMT-MM derivatized samples were reconstituted in 75% (v.v^{-1}) acetonitrile in 0.1% FA. Both the DMT-MM derivatized samples and their 2-AB labelled counterparts were analysed using a simple 6.5 min step gradient. Initial conditions were 72% B for 1.00 min, before a step from 72 to 30% B over 0.10 min with subsequent hold for 1.40 min. Starting conditions were restored over 0.10 min and the column was re-equilibrated over 3.9 min.

MS data was acquired in negative ion mode with a source capillary voltage of 1.8 kV. The nebuliser and desolvation gas temperatures were 120 °C and 400 °C, respectively. Cone and desolvation gas flow were 40 L h^{-1} and 600 L h^{-1}, respectively. The cone voltage was set at 50 V in MS mode and ramped from 20 to 90 V in MS^E mode. Data was collected in continuum using 1 second scans between *m/z* 100–2500 Da. Data acquisition and processing was performed using Waters Mass Lynx software V.4.1.

2.8. HILIC-UHPLC-FLR-MS^2 analysis

Glycans from an additional 1 mg of serum were released, labelled and cleaned as described in Section 2.3 before subsequent WAX fractionation was performed as described in Section 2.4. The glycans collected from each WAX fraction were then analysed with a Thermo Q Exactive hybrid quadrupole-Orbitrap (QE) using data-dependent acquisition (DDA) MS^2 analysis coupled to a Thermo Vanquish UHPLC system consisting of a quaternary pump, auto sampler, column pre-heater, column compartment and fluorescence (FLR) detector. N-glycans separation was performed on-line using HILIC on a 1.0 mm i.d. × 150 mm, 1.7 μm BEH Glycan column (Waters Corporation, Milford, MA, USA). Separations were performed using a linear gradient elution at 0.15 mL min^{-1} at 60 °C. Solvent A was 50 mM NH_4HCO_2, pH 4.4 and solvent B was 100% ACN. The gradient used consisted of an initial isocratic hold at 72% solvent B for 1.0 min, followed by a linear decrease in solvent B from 72% to 57% over 30 min. A decrease in solvent B from 57% to 30% was performed over 1 min followed by a return to initial conditions over 4 min to wash the column. The column was then re-equilibrated at initial conditions for 14 min. 10 μL injections were prepared by reconstituting the samples in 75% ACN. Samples were kept at 10 °C in the auto sampler while queueing to run. The column compartment and column pre-heater were both set at 60 °C while the FLR λ_{ex} was 330 nm and FLR λ_{em} was 420 nm.

MS analysis was performed in negative ion mode using a standard ESI source with a source capillary voltage of 3.00 kV. The sheath gas flow, auxiliary gas flow and sweep gas flow rates were 40, 10 and 0 arbitrary units respectively while the capillary temperature and auxiliary gas heater temperature were set at 320 °C and 400 °C respectively. MS^1 data was collected from *m/z* 200–3000 Da using an in-source CID voltage of 15 eV. Top 3 DDA MS^2 analysis was performed with a stepped normalized collision energy from 30 to 45 eV using a 2.0 *m/z* isolation window with no isolation offset, a fixed first mass of *m/z* 150 and a dynamic exclusion of 5 s (See Supplementary Table 1 for full parameters).

2.9. Structural identification

Structural identification of all glycans was performed using a combination of *de novo* annotation of the exoglycosidase LC-MS data and by searching the CFG, Carbbank, GlycomeDB and Glycosciences databases using GlycoWorkBench version 2.1 [21]. Parameters used included the following: derivatization, und; reducing end, 2AB; negative mode; max #H ion, 3 and max # charges, 3. For data generated with the Waters Xevo G2 QToF, only N-glycan matches within a mass accuracy of 20 ppm were considered, while for data generated with the Thermo QE, only N-glycan matches within a mass accuracy of 5 ppm were considered. Whenever possible, MS^2 data was used to confirm structural identification. Fragment options for MS/MS data interpretation included B, Z, Y and C fragments and cross ring fragments (maximum 1) allowing up to 2 maximum number of cleavages (3 for bisecting glycans to investigate D-GlcNAc ion).

2.10. Relative quantitative analysis of released stage 3 GC N-glycans

N-glycans released from matched serum samples of nine stage 3 cancer patients (three stage 3A, three stage 3B and three stage 3C) taken at the time of surgery (pre-op) and six weeks after tumour resection (post-op) were analysed *via* UPLC-FLR-MS^1 as described in Section 2.7. Each sample was analysed in triplicate with each injection containing N-glycans released from 100 μg of protein. Progenesis QI was used to perform relative quantitative analysis of the pre- and post-op samples. Experiments performed included the analysis of all stage 3 samples and separate analysis of each stage 3 subgroup (stage 3A, 3B and 3C). Data was imported and automatically aligned by the Progenesis QI software, which analyses all runs and selects the optimum run for alignment as a reference and

subsequently aligns all other runs to that reference. Alignment was then checked to ensure that all samples were more than 90% aligned. For the few samples that had less than 90% alignment, manual vectors were added to improve alignment to over 90%. A within-subject experiment design was selected to perform a repeated measures ANOVA experiment between the pre- and post-op samples. Peak picking was set to detect singly, doubly and triply charged ions eluting between 2 and 31 min to cover the range of the method gradient. No minimum intensity or peak width was selected to ensure all spectral features were detected. Compounds were then manually identified by checking the monoisotopic masses and retention times against the glycans identified within this study. All ions not corresponding to such compounds were discarded. Deconvolution of identified glycans was re-examined to ensure all identifications from multiple adducts were considered as a single identification. PCA and ANOVA analysis was then performed based on the identified compounds. For ANOVA analysis glycan significance levels were corrected for multiple comparisons ($P = 0.05/n$, where for all stage 3 analysis $n = 80$, stage 3A analysis $n = 76$, stage 3B analysis $n = 65$ and stage 3C analysis $n = 68$. Here n is the number of N-glycans identified for each stage 3 analysis). Only N-glycans with ANOVA p-values $\leq 6.25 \times 10^{-4}$, 6.58×10^{-4}, 7.69×10^{-4}, and 7.35×10^{-4} were considered to be statistically significant for experiments analysing all stage 3, only stage 3A, only stage 3B and only stage 3C samples, respectively. The normalized abundance values generated from Progenesis QI were then exported and further analysed using non-parametric univariate statistics in SPSS to generate box plots for all glycans exhibiting statistically significant changes in expression level.

3. Results and discussion

3.1. Serum N-glycan fractionation

The complexity of the serum N-glycome requires the utilization of an analytical platform capable of mining for and characterizing low abundant N-glycans. Thus, serum profiling studies have routinely used 2-dimensional separations platforms for N-glycan identification [11,12]. Our platform, as outlined in Fig. 1, follows this same approach but in combination with high resolution mass spectrometry, exoglycosidase digestion sequencing and sialic acid linkage specific chemical derivatization. N-glycans were first released from serum, derivatized at their reducing end with 2-AB and then separated utilizing a 2-dimensional UPLC-FLR-MS platform which uses WAX fractionation in the first dimension followed by HILIC in the second dimension. WAX chromatography allows for the separation of N-glycans based on their degree of sialylation or the presence of other charged substituents and resulted in the generation of five well defined fractions for the N-glycome based on degree of sialyation corresponding to neutral glycans (N or S0) followed by the elution of mono- (S1), di- (S2), tri- (S3) and tetra- (S4) sialylated species. The exception to this, as discussed later, is the presence of glycans containing heavily negatively charged moieties such as sulphate which can then lead to the presence of lesser sialylated N-glycan species in fractions generally containing more heavily sialylated structures.

The application of HILIC is common in N-glycan biomarker studies because of its ability to separate N-glycan structural isomers [4,20,22–25]. However, HILIC analysis of complex samples such as the serum N-glycome, still results in many peaks containing multiple N-glycan structures often with varying degrees of sialyation (Fig. 2). Applying WAX separated N-glycans to HILIC in the second dimension provides confidence in structural characterization as glycans present in WAX peaks contain the same degree of sialyation while the reduced sample complexity allows for the identification of low abundant glycans contained within each WAX fraction. Here, N-glycans released from approximately 7 mg of serum and first fractionated by WAX, were separated by HILIC resulting in the collection of 50–65 HILIC peaks per WAX fraction. The oversaturation of N-glycans was deliberately applied to ensure that low abundance glycans, such as those in the S4 wax fraction, could be detected during subsequent MS analysis. Collection of all individual HILIC peaks allowed for DMT-MM derivatization before UPLC-FLR-MS^E analysis, enabling the determination of all structures contained within a single HILIC peak including sialic acid linkage isomers.

3.2. Serum N-glycan structural determination

HILIC-UPLC-FLR-MS^E analysis of the WAX fractions and associated exoglycosidase digestion panels provides structural composition and isomeric information concerning the glycans present within the WAX fractions (Fig. 3). The presence of each additional enzyme generates a mass shift corresponding to the loss of the specific monosaccharide cleaved by said enzyme generating N-glycan base structures that can be followed through each digest using a bottom to confidently annotate the N-glycan composition. Structural isomers are differentiated according to retention time and analysis of MS^2 digest spectra (Fig. 4). Bisecting and tri-antennary branched glycans can be distinguished by the presence of D, D-18 and E ions in MS^2 spectra, which are product fragment ions generated during negative ion mode ESI analysis that are readily used to determine the composition of a N-glycans 6- or 3-antenna, respectively [26,27]. This is illustrated in Fig. 3A and B using the ABS + AMF + BKF + SPG digest from the S2 WAX fraction to identify the structures in the FLR peaks at 6.23 and 6.43 min who both have an MS^1 *m/z* of 818.8 corresponding to an A2B or A3 structure. For bisecting glycans, the bisecting GlcNAc is part of the D ion (Fig. 4). However, it is readily lost during fragmentation resulting in a prominent $[D - (GlcNAc + H_2O)]^-$ ion and minimal D ion. The MS^2 spectrum of the peak at 6.23 min contains a very prominent ion at *m/z* 508.17 indicative of a $[D - (GlcNAc + H_2O)]^-$ ion due to the presence of a bisecting GlcNAc residue. The lack of the D and D- H_2O ions at *m/z* 729.26 and 711.28 respectively, confirms the structure as A2B. Tri-antennary glycans predominately contain 2 branches extending from the 3-antennary mannose residue and a single branch from the 6-arm mannose. The MS^2 spectrum of the peak at 6.43 min contains a prominent D ion fragment at *m/z* 526.17 representative of a single GlcNAc residue branched from the 6-arm mannose. Here, the fragment at *m/z* 508.17 is a product of a $[D - 18]^-$ ion due to a loss of water and is present at an approximate 1:1 ratio with the D ion. The fragment at *m/z* 507.19 is the E-type fragment indicating the presence of 2 GlcNAc residues on the 3-antennary branching mannose residue, whose presence confirms the structure as A3. These identifications demonstrate that bi-antennary glycans containing a bisecting GlcNAc residue elute marginally before their tri-antennary isomeric counterparts when separated on HILIC. Analysis of the ABS + AMF + BKF + SPG digest for each WAX fraction shows that the presence of bisecting glycan species is inversely related to degree of sialyation with their abundance greatest in the neutral fraction and steadily decreasing throughout until their presence is non-existent in the S3 and S4 WAX fractions. No tri- or tetra-antennary N-glycans found within the neutral, S1 and S2 WAX fractions contained bisecting GlcNAc residues. This is an observation consistent with the common understanding that further N-glycan branching is inhibited by the presence of a bisecting GlcNAc residue, due to conformational changes that hinder the flexibility of the α1-3 arm and generate back-folding of the α1-6 arm, and that extensive branching prevents the N-acetylglucosaminyltransferase responsible for adding

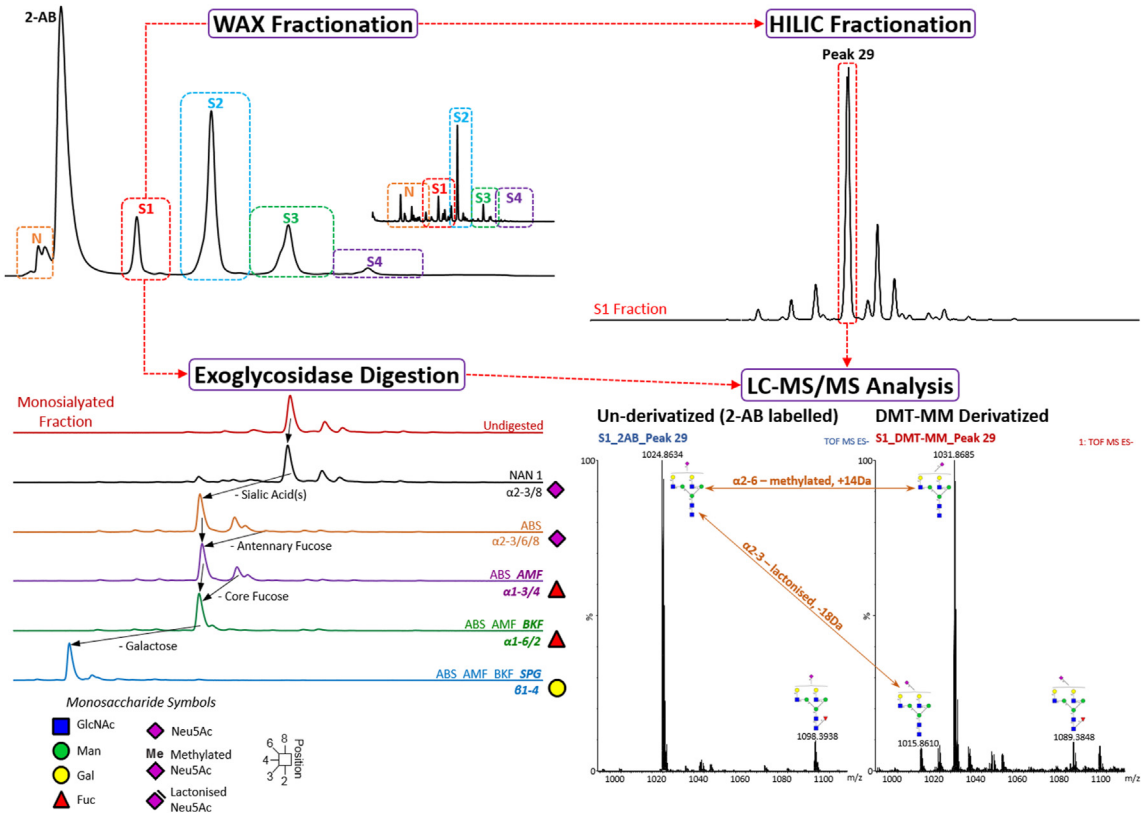


Fig. 1. General schematic of serum N-glycome profiling. Released and 2-AB labelled serum N-glycans are first fractionated offline using WAX to separate glycans by charge, which correlates to the number of sialic acids present on a glycan. Glycans collected from each WAX fraction are either digested with exoglycosidases to obtain structural composition information or are applied to HILIC for second dimensional separation. Each HILIC peak is collected offline and a portion is aliquoted for DMT-MM derivatization. After DMT-MM derivatization, the native aliquot and DMT-MM derivatized aliquot for each peak are run sequentially to obtain sialic acid linkage information.

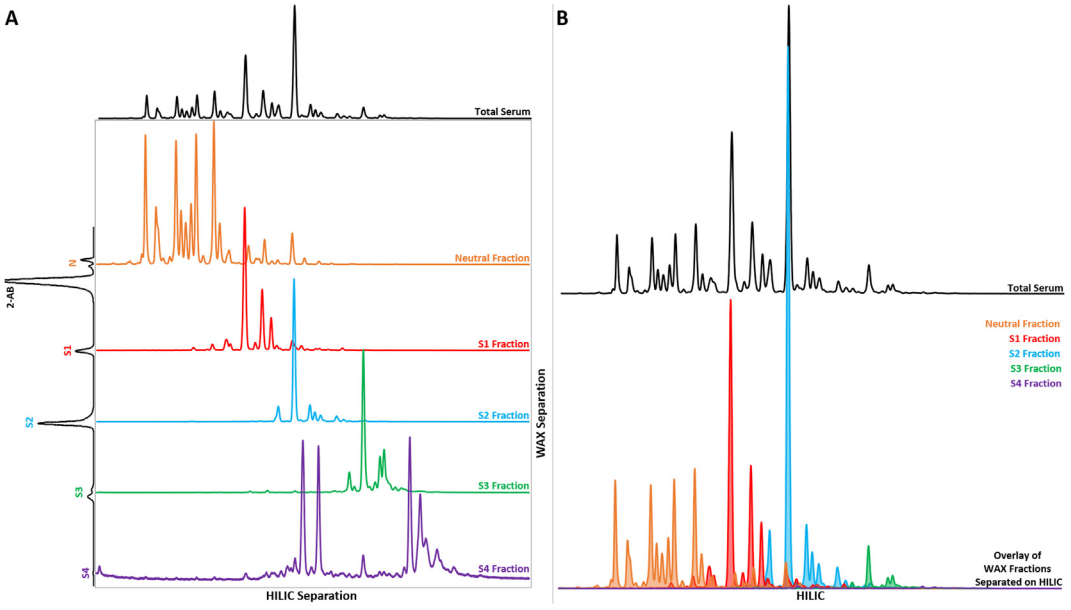


Fig. 2. The complexity of serum results in different sialylated N-glycan species eluting within a single HILIC peak. Separating N-glycans first by WAX allows for the separation of N-glycans by degree of sialyation. **(A)** Illustrates N-glycans first separated by WAX, subsequently applied to HILIC and their contribution to the total human serum N-glycome. Profiles are listed in order of degree of sialyation (From Top: Total serum, N, S1, S2, S3 and S4). **(B)** Illustration of how different sialylated N-glycan species contribute to a single HILIC during total serum N-glycan profiling.

bisecting structures from accessing the core mannose residue [28–33].

Digest panels also help differentiate between core and antennary fucosylation as illustrated with the S2 WAX fraction in Fig. 3A

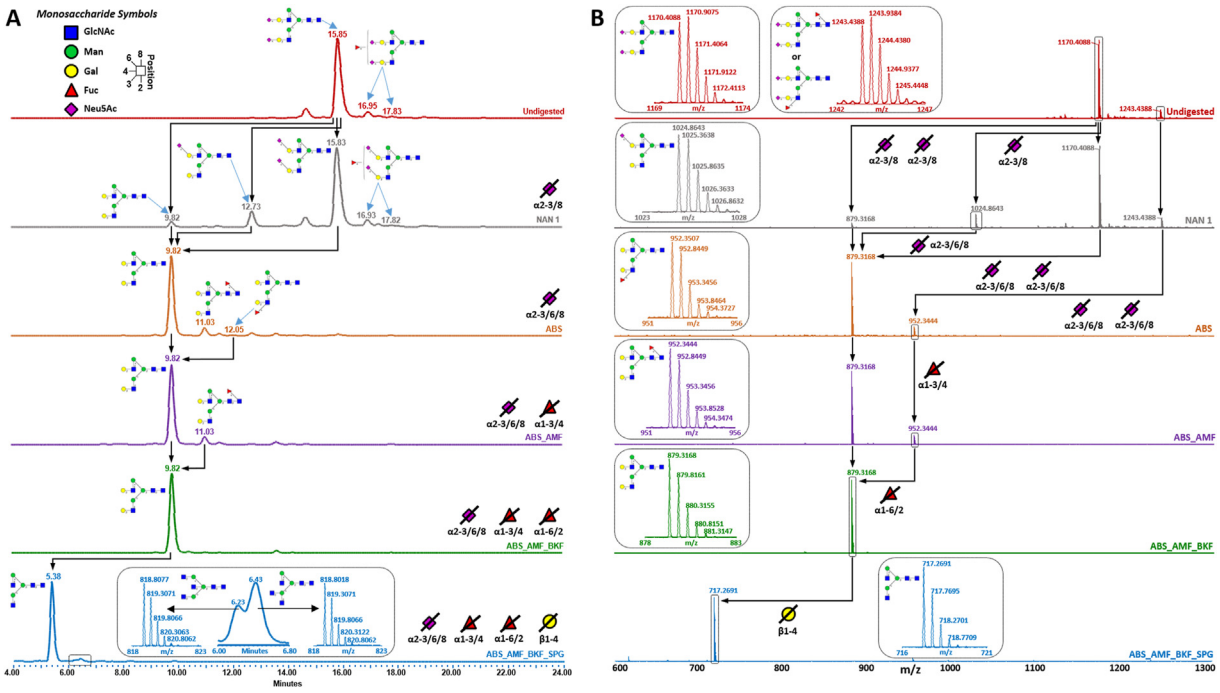


Fig. 3. **(A)** Exoglycosidase digests of each WAX fraction are utilized to determine glycan structure composition. As monosaccharides are cleaved, HILIC retention of N-glycans shifts leftward corresponding to the type of monosaccharide lost. **(B)** MS analysis of the WAX exoglycosidase digests is utilized to confirm N-glycan structures.

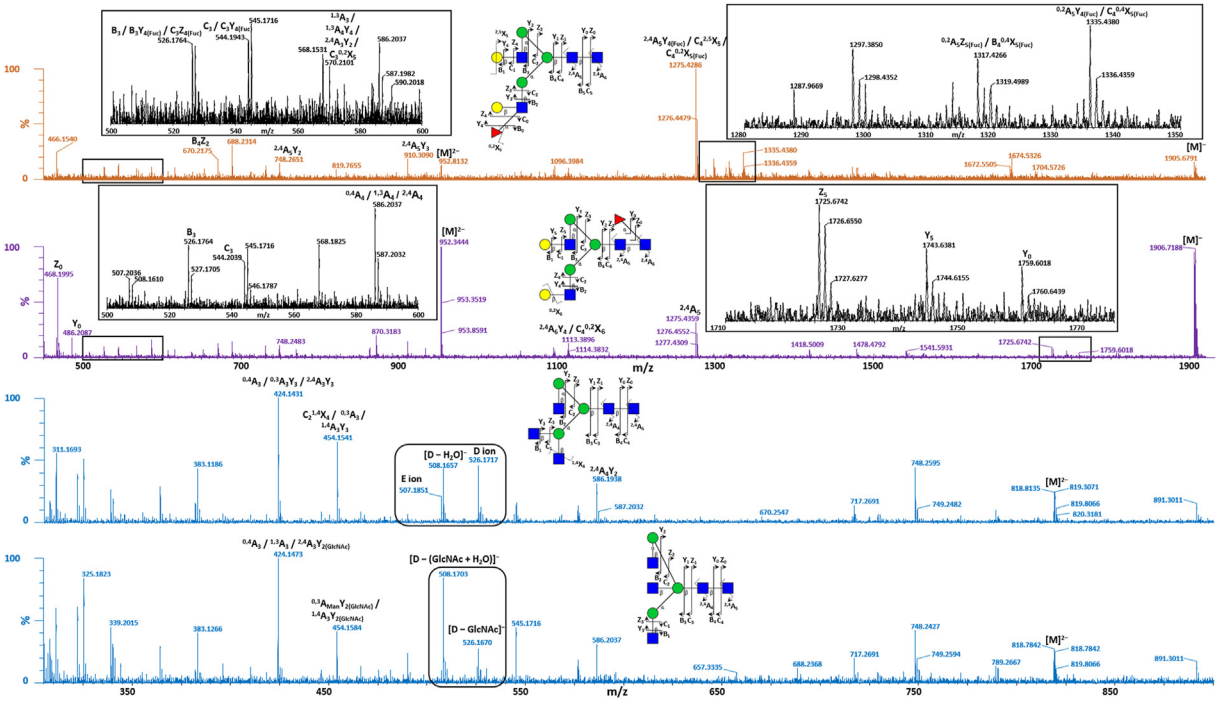


Fig. 4. **(C)** MS^2 analysis is used to identify structural isomers. The presence or lack thereof of specific D and E ions can be used to differentiate between bisecting and branched GlcNAc residues as well as core and antennary fucosylation.

and B. BKF cleaves α1-2/6 linked fucose residues and is commonly used to detect the presence of α1-6 core fucosylation. FLR peaks in ABS + AMF digests, such as the one at 11.03 min in the S2 WAX digest (Fig. 3A), which are not in ABS + AMF + BKF reveal core fucosylation. Here doubly charged ions at *m/z* 952.3 in the MS^1 spectra reveal the structure to be a fucosylated A2G2 glycan. The presence of ions at *m/z* 468.2 and *m/z* 486.2 in the MS^2 spectra representative of *z* and *y* ions produced by Fuc-α1,6-GlcNAc-2AB fragmenting from the glycan structure confirm the presence of core fucosylation as such ions would not be expected to be present in the case of antennary fucosylation (Fig. 4). AMF cleaves α1-3/4 linked fucose and is used to help identify the presence of antennary

fucosylation as illustrated through the comparison of FLR traces for the ABS and ABS + AMF digests of the S2 WAX fraction (Fig. 3A). The peaks at 11.03 min and 12.05 min in the ABS digest are both due to the presence of a fucosylated A2G2 glycan based on MS^1 analysis (Fig. 3B). The glycan at 11.03 min is core fucosylated as its presence was noted in the ABS + AMF digest previously discussed. The digest suggests that the glycan at 12.05 min contains α1-3/4 antennary fucosylation as such a peak does not appear in the ABS + AMF digest. If the glycan contained α1-2 antennary fucosylation, its presence would have been detected in the ABS + AMF digest. The lack of prominent ions at *m/z* 468.2 and *m/z* 486.2 in the MS^2 spectra indicate antennary fucosylation. The presence of an ion at *m/z* 570.2, which is not seen in the MS^2 spectra of the core fucosylated structure (Fig. 4), is representative of the $^{1,3}A_{Man}$ cleavage of a Gal-GlcNAc branch containing an antennary fucose residue. However, due to the low abundance of this particular glycan, MS^2 analysis could not be used to determine which branching arm predominately contains the antennary fucose residue.

The implemented analytical platform also allows for the differentiation of N-glycan isomers with A3 and A3′ tri-antennary base structures known to be present in human serum [34]. The prominent standard A3 structures have the third GlcNAc residue β1-4 linked to the Man residue on the α1-3 branch while the low abundant A3′ structures have the third GlcNAc residue β1-6 linked to the Man residue on the α1-6 branch. As illustrated in the S3 WAX fraction, these structural isomers were separated by about 0.6 min on the HILIC phase with the A3 structures eluting before the A3′ structures (Fig. 2 and Supplementary Table 2). While such structures most commonly occur as tri-sialylated species, the digests of the S1 and S2 WAX fractions show that A3' structures are also present as mono- and di-sialylated structures.

3.3. Sialic acid linkage determination

Different strategies have been described in the recent years to assist in sialic acid linkage determination and recently de Haan, N. et al. published a review detailing all the different derivatization approaches reported in the past decade, which have been implemented in various mass-spectrometry- glycomics workflows and have found clinical glycomics applications [35]. Sialic acid derivatization with DMT-MM enables differentiation between α3-and α6-linked sialic acids [20,36]. When performed in methanol, α2-3 linked sialic acids form a cyclic lactone resulting in a loss of 18 Da from the precursor a mass due to the loss of water during formation, while α2-6 linked sialic acids become methylated resulting in the addition of 14 Da to the precursor mass. Here, a portion of each collected HILIC peak underwent DMT-MM derivatization. Following derivatization, the un-derivatized portion (2-AB labelled only) and derivatized portion (DMT-MM derivatized) of the HILIC peak were analysed sequentially using UPLC-FLR-MS^E. As illustrated in Fig. 5, comparing the mass shifts between the MS^1 spectra of the 2-AB only labelled sample and the DMT-MM derivatized sample enables the determination of sialic acid linkage and can be used to identify N-glycan sialic acid linkage isomers that are present in a single peak. Whenever possible, MS^2 spectra is used for structural composition verification.

DMT-MM derivatization reveals that sialic acid linkage impacts HILIC retention with isomers containing 3-linked sialic acids eluting earlier than their 6-linked counterparts. This observation is consistent with findings in the literature [37,38]. This is best demonstrated within the S1 fraction where the structures A2G2S(3)1 and A2G2S(6)1 are found within peaks 26 and 29 respectively, and the S2 fraction where the structures A2G2S(3,3)2, A2G2S(3,6)2 and A2G2S(6,6)2 are found within peak 22, peak 26 and peak 29 respectively (Fig. 6). This trend is consistent in all fractions. However, for the larger oligosaccharides that elute late in the gradient, there is significant co-elution of linkage isomers due to decreased column efficiency resulting broader peak shape. The ability to still distinguish between such isomers illustrates the benefits of using DMT-MM derivatization for sialic acid linkage determination.

3.4. Human serum N-glycome characterization

Utilizing our 2-dimensional separations platform we were able to identify 212 distinct compounds with 64 fully and 148 partially characterized N-glycans including sialic acid linkage isomers (Supplementary Table 2). Partially characterized structures are labelled as such due to a combination of their low abundance and the inability to determine precise location of branching when multiple locations were possible. For example one peak contained an isomer of an A3G3S(3)1 N-glycan. However, due to the peak's low abundance and the presence of multiple other N-glycans within the peak, we were unable to determine the precise branch the sialic acid was linked to. A combination of WAX digests and MS^2 analysis were used to help identify structural isomers within each WAX fraction applied to HILIC.

The ability to detect O-acetylated (OAc) and O-sulphated (SO_3) sialo-N-glycans underscores the ability of this platform to identify low abundant glycans with such glycans previously reported in mammalian and human glycoprotein studies [39–42]. Additionally, Yamada et al. [43] recently published the presence of sulphated and phosphorylated N-glycans in human serum after using a serotonin-immobilized column to enrich minor acidic glycans, although to the best of our knowledge no in-depth human serum characterization study has yet identified acetylated sialo-N-glycan structures. Sialic acid O-acetylation is a common modification that can occur on the hydroxyl groups attached at the C4, C7, C8 and C9 positions and is known to play important roles in various biological processes including disease pathogenesis, viral infection and cancer progression as well as increase potency and half-life of certain biopharmaceuticals [41,44]. In mammals, O-acetylation is most prominently found at the C9 position either through direct acetylation or more commonly through non-enzymatic migration under physiological conditions to the C9 position from the C7 position where it is initially added [41,42,45]. We identified N-glycans containing *mono*-O-acetylation in the S1, S2, and S3 WAX fractions, primarily on the most abundant structures contained within each fraction, *i.e.* in the S2 WAX fraction the most abundant glycan was A2G2S(6,6)2 with its acetylated counterpart OAc-A2G2S(6,6)2 identified as well (Fig. 6). Identified acetylated glycans were found to elute approximately 2 min earlier on HILIC than their unmodified counterparts, an expected occurrence as acetylation can reduce sialic acid hydrophilicity [46]. Due to their low abundance, it was not possible to identify the sialic acid residue containing the O-acetylation modification when more than one sialic acid was present. However, the presence of these acetylated sialo-N-glycans coupled with the common occurrence of this modification could indicate that other acetylated sialo-N-glycan species are present, but beyond the detection capabilities of this platform due to the broad dynamic range of the serum N-glycome.

Sulphated S2 glycans were identified within the S4 WAX fraction. Their presence is a product of the highly negatively charged sulphate group greatly increasing the overall charge of the glycan structures resulting in later elution during WAX chromatography than their non-sulphated counterparts. Comparison of the S2 and S4 HILIC traces reveals that sulphation increases glycan hydrophilicity with the sulphated glycans eluting later that their unmodified glycoforms (Fig. 7). The most prominent identified structures, SO_3-A2G2S(6,6)2 and SO_3-FA2G2S(6,6)2, are the sulphated glycoforms

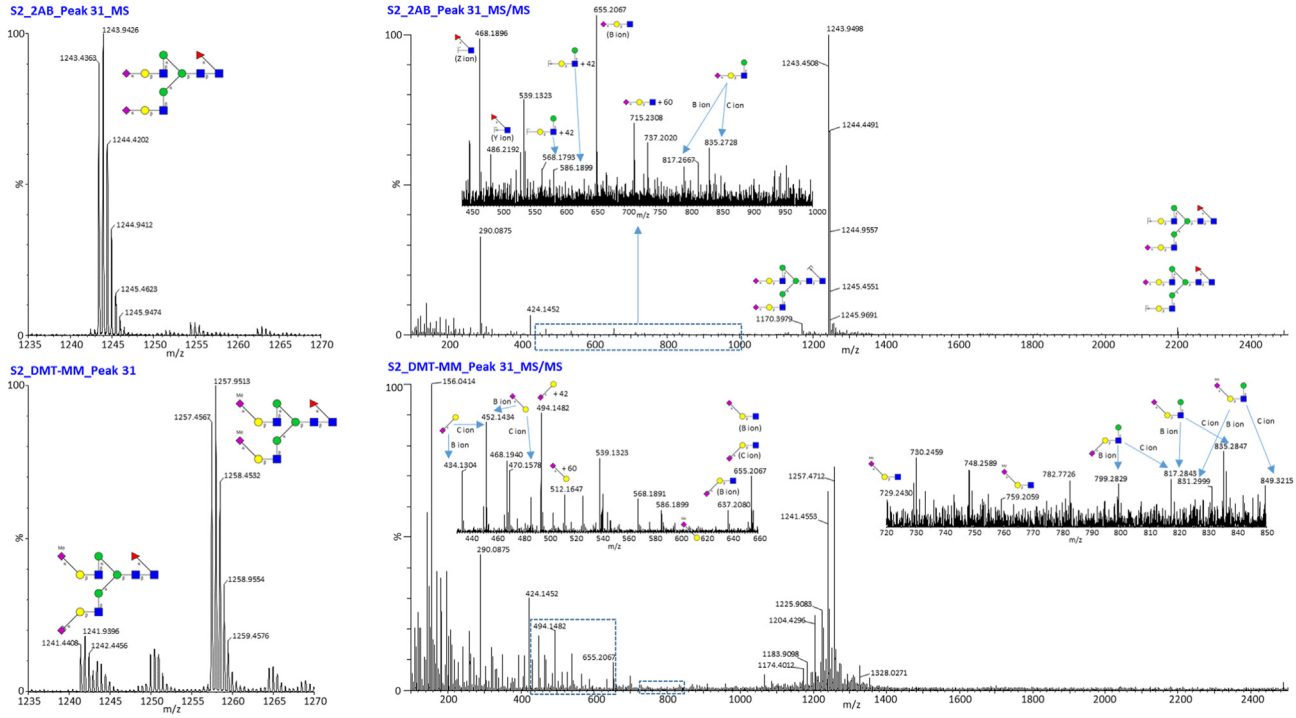


Fig. 5. DMT-MM derivatization allows for the determination of sialic acid linkage position. α2-3 linked sialic acids form a cyclic lactone resulting in a loss of 18 Da from the precursor a mass due to the loss of water during formation, while α2-6 linked sialic acids become methylated resulting in the addition of 14 Da to the precursor mass. MS^2 analysis can be utilized to determine which branched arm a sialic acid is located on.

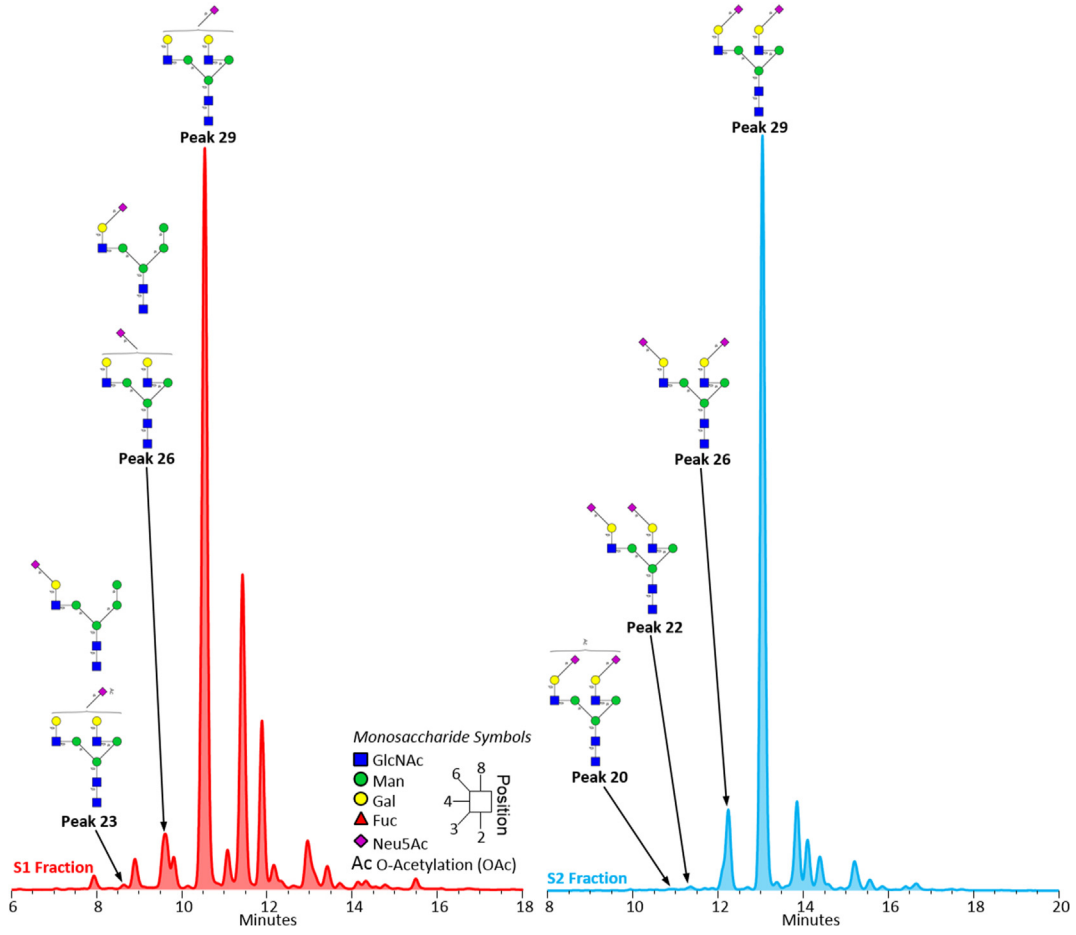


Fig. 6. As illustrated in both the S1 (Left) and S2 (Right) WAX fractions applied to HILIC, N-glycans containing α2-3 linked sialic acids elute earlier than their counterparts containing α2-6 linked sialic acids. *Note: Intensities of the S1 and S2 HILIC profiles are not to scale.*

of some of the most abundant glycan structures present in the serum glycome. Such structures have previously been found on myelin glycoproteins, where they help in the crucial role of forming myelin sheaths to insulate nerve cells within the peripheral and central nervous system, and recently on human serum IgGs [39]. Additionally, sulphation of N-glycans not containing Lewis X, sialyl Lewis X or terminal lacdiNAc features have been identified in multiple cell lines despite not being the target of their respective studies, suggesting that this type of N-glycan sulphation may be more prominent than currently realized [47,48].

The S4 WAX exoglycosidase digests exemplify how the labile nature of sulphate causes difficulties in the identification and structural characterization of sulphated sialo-N-glycans (Fig. 8). Besides the known issue of sulphate migration occurring during MS analysis [49], unsulphated glycoforms of the sulphated N-glycans are present throughout the S4 WAX digest panel suggesting possible loss during sample preparation. For all digests except the undigested run, the sulphated species elutes earlier than the un-sulphated species. This could be because the sulphate occupies a previously free hydroxyl group, reducing interaction with the HILIC stationary phase. For the undigested fraction, the sulphated sialo-N-glycan species elutes after its un-sulphated counterpart as also demonstrated in Fig. 7. While it is not understood why this occurs, it is believed that interactions between the sialic acid and the sulphate ion led to increased retention to the HILIC stationary phase.

The presence of sulphate also appears to impact exoglycosidase efficiency as illustrated by the identification of an SO_3-A2G1 base structure in the ABS + AMF + BKF + SPG digest which suggests incomplete digestion of the Gal residues by SPG. It is suspected that the presence of the sulphate adduct prevents complete digestion from occurring on the branch containing the sulphate modified residue. In this case, it is the Gal residue on the sulphated branch that is not cleaved during digestion. Due to the low abundance of sulphated N-glycan species, the location of the sulphate could not be determined. However, recent studies have identified sulphated N-glycans without a terminal lacdiNAc or sialyl Lewis X motif containing sulphate at the C6 position of GlcNAc (GlcNAc-6-sulphate) and the C3 position of Gal (Gal-3-sulphate) [39,50]. Additionally, the presence of core fucose on the sulphated A2G2S2 glycans appears to stabilize the sulphate substituent. As observed in both the ABS + AMF digest and the ABS only digest, the presence of the un-sulphated FA2G2 structure is low compared to its sulphated counterpart, while the presence of the un-sulphated form of the A2G2 structure is greater than its sulphated glycoform.

3.5. Relative GC N-glycan analysis

Analysis of the LC-MS data generated for the stage 3 samples in Progenesis QI revealed subgroup specific N-glycan changes. Combined analysis of the stage 3 patient samples found only one glycan, A3G3S(3,6,6)3, to have statistically significant changes in expression levels, with higher expression found in the pre-op samples (Fig. 9: Red box plots). Analysis of the stage 3 sub-groups indicates that this glycan appears to be a product of changes within the stage 3A subgroup as it was significantly expressed in the stage 3A samples, but not in the stage 3B or 3C sub-groups (Fig. 9: Orange box plots). Stage 3B analysis revealed twelve glycans to have significantly increased expression in the cancerous state prior to tumour resection (Fig. 9: Green box plots). A great majority of these glycans contain core fucosylation, a modification which is known to be associated with cancer [51]. In GC, core fucosylation has been found to increase on haptoglobin and in multiple GC cell lines [51–53]. Stage 3C analysis found a total of thirteen glycans including isomers with significantly increased expression after surgery (Fig. 9: Blue box plots). It is thought that these changes are due to the increased inflammatory state following resection.

Two of the sample pairs analysed in this study were from patients with microsatellite instability (MSI). When total stage 3 analysis was performed without these patients, the presence of glycans with significant expression levels (corrected $p < 6.49 \times 10^{-4}$; $n = 77$) dramatically increased, with fourteen glycans identified (Table 1). Many of the glycans were those identified in the analysis of the stage 3B and stage 3C sub-groups to have significant expression. Additionally two glycans, FA2G2S(3)1 and FA3G3S(6)1, were seen to be significantly expressed, but did not show such expression change in the analysis of the stage 3 sub-groups. FA2G2S(3)1 had greater expression in the pre-op samples and thus could be associated with cancer, while A3G3S(6)1 had increased expression after surgery. Interestingly, the glycan A3G3S(3,6,6)3 no longer had significant expression level changes which could signify that increased expression of A3G3S(3,6,6)3

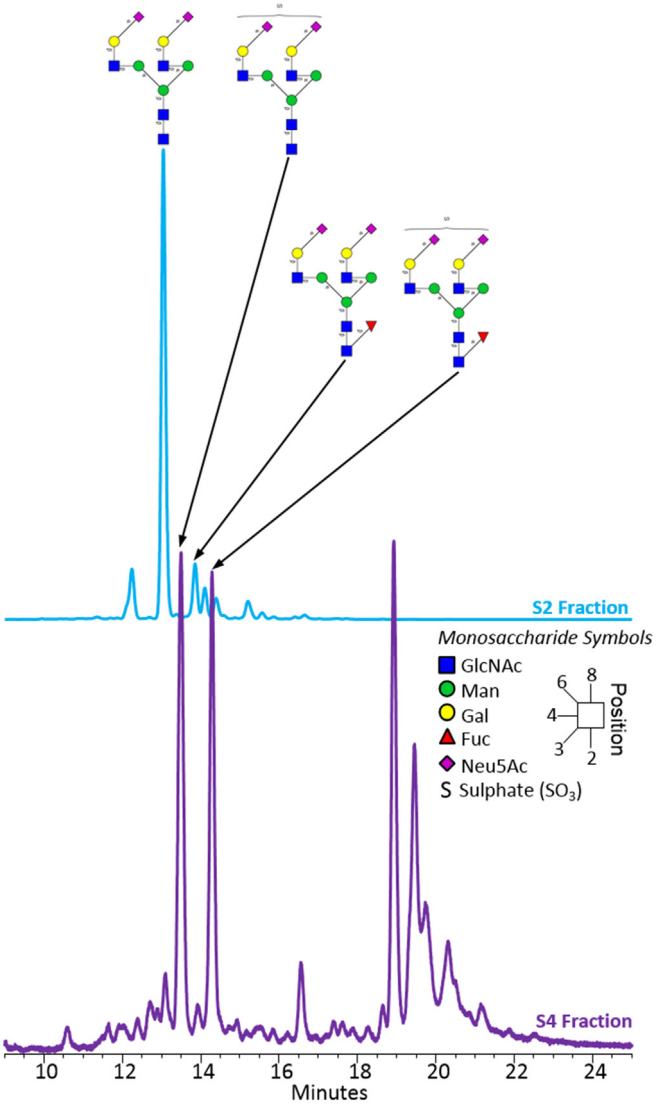


Fig. 7. Comparison of the S2 (Top) and S4 (Bottom) WAX fractions applied to HILIC shows that sulphated sialo-N-glycans elute later than their un-sulphated glycoforms as illustrated with the A2G2S(6,6)2 and FA2G2S(6,6)2 N-glycan structures and their respective sulphated glycoforms. The red circle indicates signifies that sulphation is predicted to occur at the 6-position of one of the antennary GlcNAc residues. The branched arm containing sulphation could not be determined. *Note: Intensities of the S2 and S4 HILIC profiles are not to scale. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)*

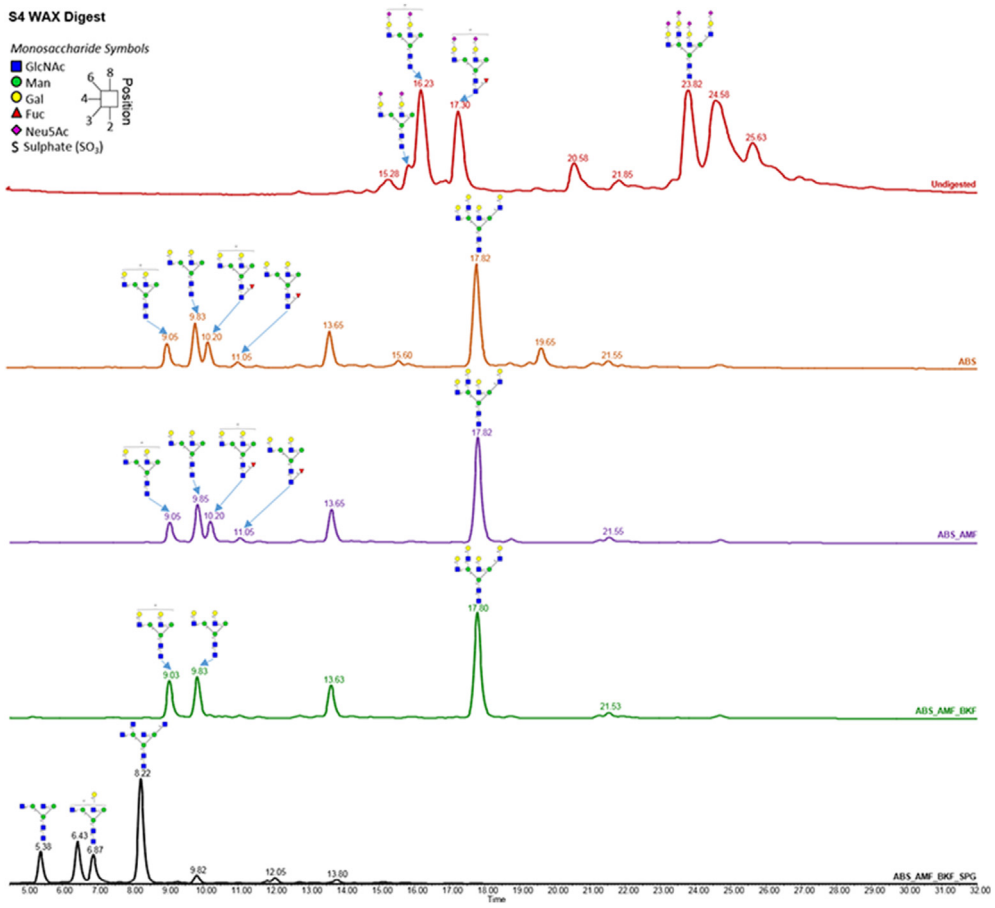


Fig. 8. S4 WAX fraction illustrates the presence of sulphated N-glycan species. The labile nature of sulphate can result in its loss during sample prep resulting in the unexpected presence of the un-sulphated species. The exact location of the sulphate could not be determined. Interestingly, the digests reveal that the un-sialylated sulphated N-glycan species elute before their un-sulphated counterparts, but sulphated sialo-N-glycans elute after their un-sulphated counterparts as previously illustrated in Fig. 7. While the reason for this is not fully understood, it is believed that interactions between the sialic acids and the sulphate increase the retention mechanisms of the sulphated sialio-N-glycans to HILIC causing their later elution. Additionally, the presence of a core fucose appears to increase the stability of the sulphate ion.

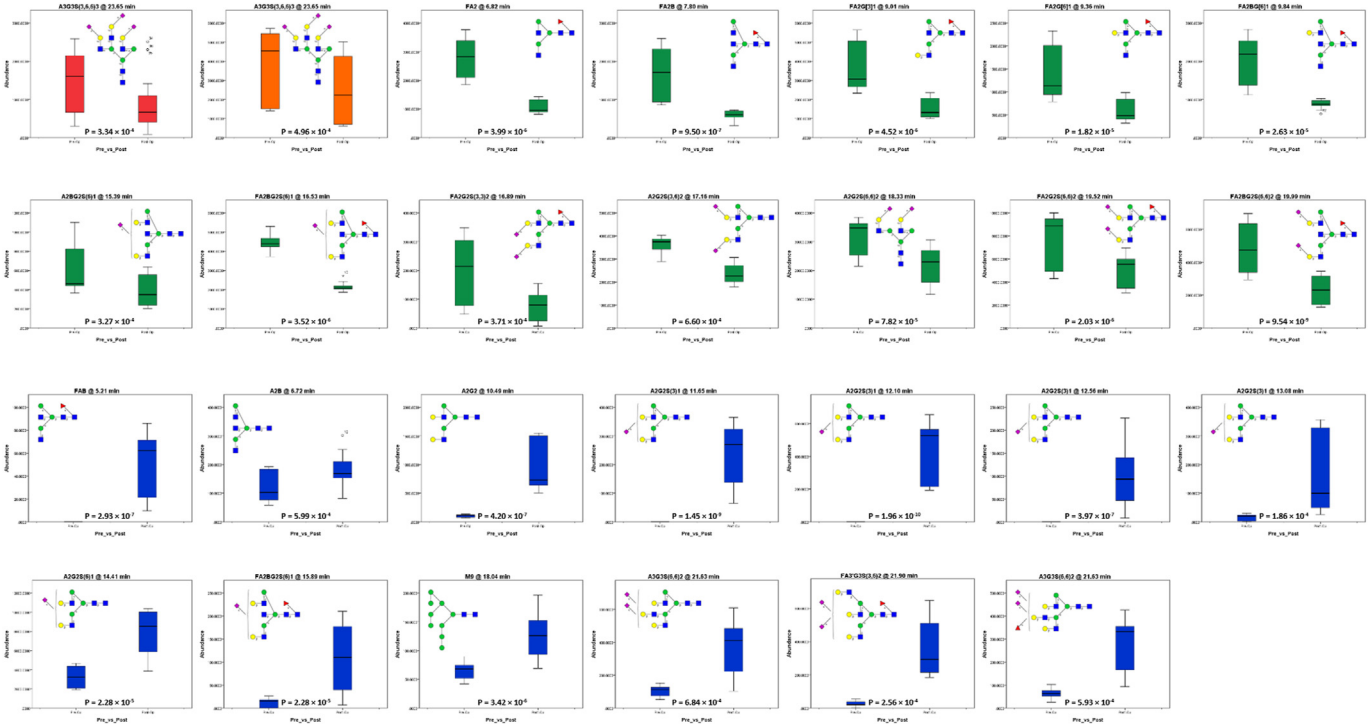
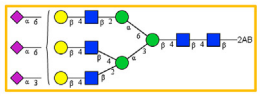
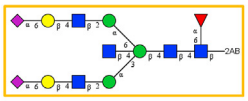
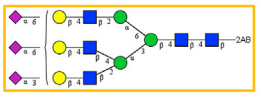
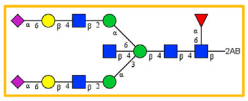
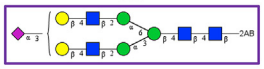
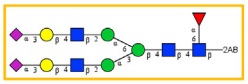
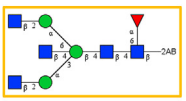
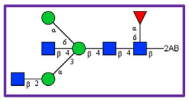
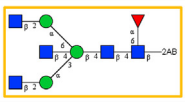
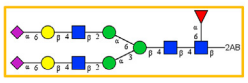
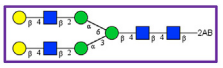
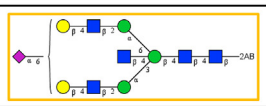
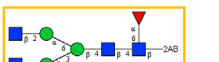
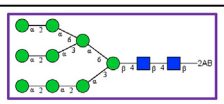
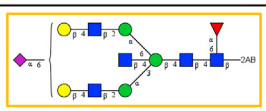
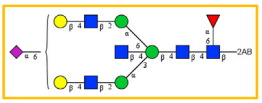
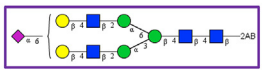
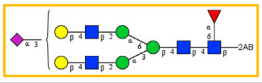
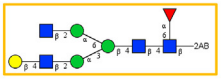
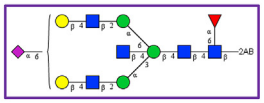
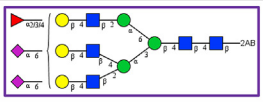
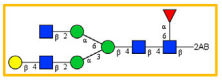
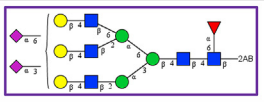
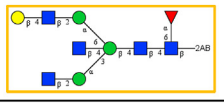
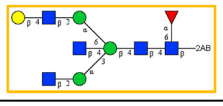
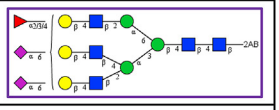
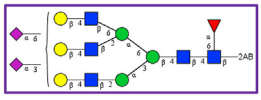
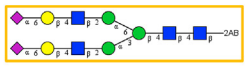
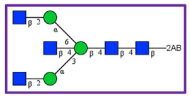
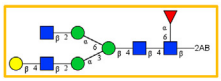
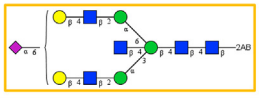
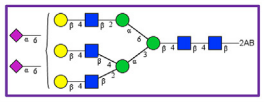
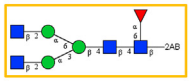
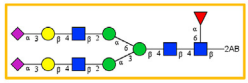
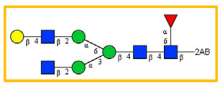
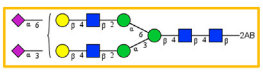
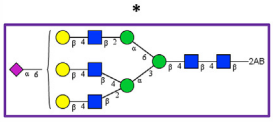
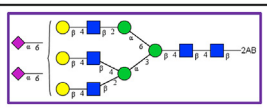


Fig. 9. Box plots of glycans found to have statistically different expression changes between pre- and post-op serum samples taken from patients with stage 3 gastric cancer. Red box plots are from the analysis of all stage 3 samples, orange box plots are from the analysis of stage 3A samples, green box plots are from the analysis of stage 3B samples and blue box plots are from the analysis of stage 3C samples. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 1
Expression trends of N-glycans with statistical significance across the pathological subgroups of Stage 3 GC.

Stage 3 (All)	Stage 3 (Without MSI)	Stage 3A	Stage 3B	Stage 3C
				Iso 
				
				
				
				
	* 			
				

Stage 3 (All)	Stage 3 (Without MSI)	Stage 3A	Stage 3B	Stage 3C
				
				
				
				
				
	* 			
				

Symbolic representation for N-glycans: GlcNAc – blue square, Fuc – red triangle, Man – green circle, Gal – yellow circle, Sialic acid – purple diamond. Orange border indicates increase of glycan in pre-op samples while purple border depicts increase of glycans in post-op samples. The asterisk (*) marks the glycans found in the analysis of patients without MSI to have significant changes in expression levels that were not found to have significant changes in any of the stage 3 sub-groups. Glycans marked Iso signify that multiple isoforms of the glycan were detected.

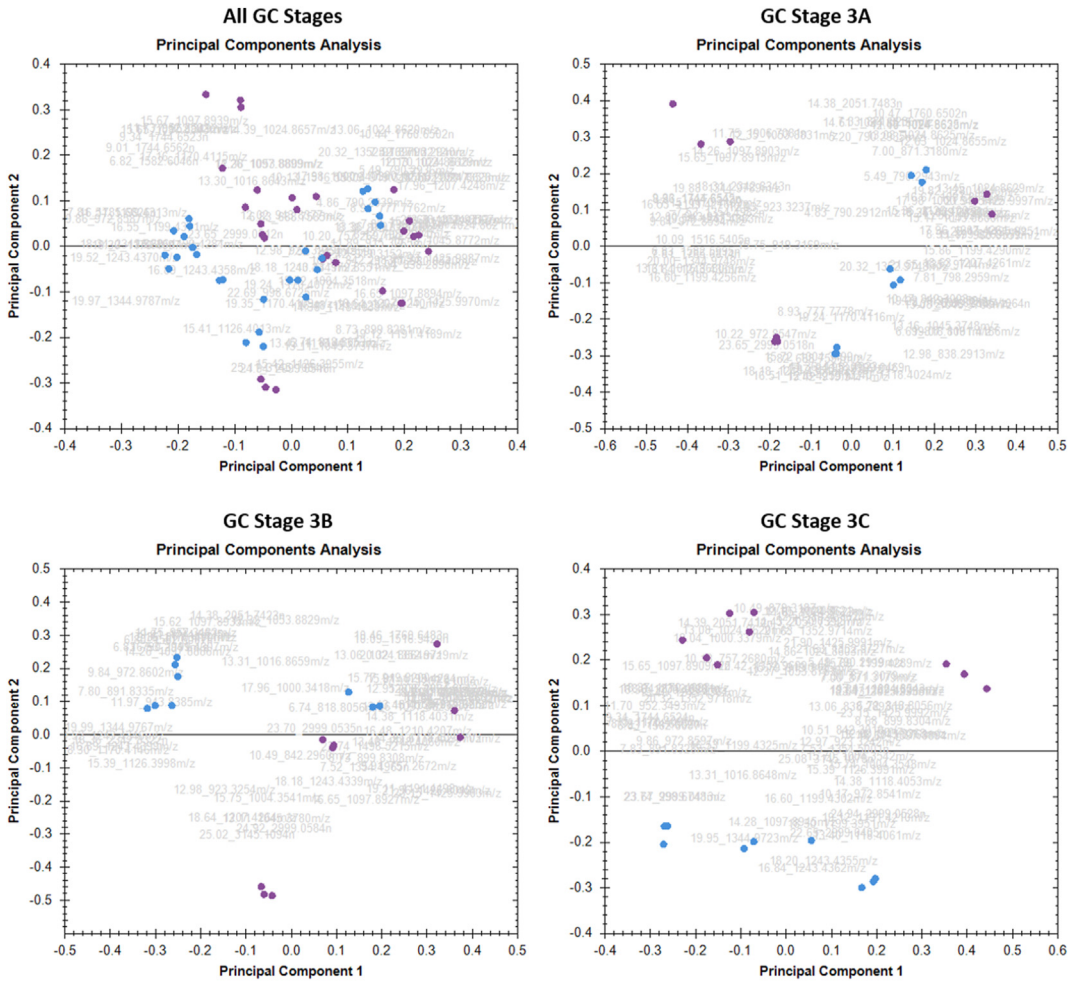


Fig. 10. PCA analysis of differentially expressed N-glycans illustrating how separation between the pre- and post-op samples becomes clearly defined as Stage 3 GC advances.

may be associated with the presence of MSI. However, further investigation into this possibility is necessary as such conclusions cannot be determined with a sample set of this size. Additionally, the N-glycans identified in this study have been previously shown to be cancer associated illustrating the usefulness of this platform for identifying cancer associated N-glycans [54].

PCA analysis was performed to visualize the differential expression behaviour of all the stage 3 samples shows that there is no defining characteristics to differentiate between the pre- and post-op samples (Fig. 10). However, PCA analysis of the stage 3 subgroups illustrates that as the cancer progresses from stage 3A to stage 3C separation between the pre- and post-op samples becomes more clearly defined, with analysis of the stage 3C patients showing distinct grouping of the pre- and post-op samples (Fig. 10). This may be representative of the cancer increasingly infiltrating the surrounding lymph nodes after the tumour has grown through the stomach wall and the potential ability to detect this pathological process in the serum of these patients.

4. Conclusions

An efficient characterization platform consisting of two-dimensional liquid chromatography coupled to time of flight mass spectrometry and coupled with exoglycosidase digestion and sialic acid linkage specific chemical derivatization was developed and applied to the in depth characterization of the serum N-glycome. Highly informative MS^2 product ions enabled the confident structural differentiation of isomeric N-glycans based on the presence of motif specific D- and E-type fragment ions in the negative ion MS^2 spectra. Modified sialic acids, containing additional O-acetylation were identified along with the presence of sulphated bi-antennary sialylated N-glycans. The retention behaviour of these isomeric species and modified structures was also established relative to the parent non-modified N-glycan counterpart. In total, the employed analytical platform facilitated the complete characterization of 64 N-glycans with high confidence and 148 partially annotated structures. The created database of annotated N-glycans was then used in combination with label free LC-MS and Progenesis QI for the determination of differentially expressed N-glycans in the serum of patients with various stages of Stage 3 gastric cancer. Utilization of this label free approach revealed the identification of alterations in the N-glycome that correlated with pathologically significant subgroups within the Stage 3 GC patient group investigated. The range of this platform enables detection of low level glycans not usually seen in total serum characterization platforms. This depth combined with the ability to determine sialic acid linkage, reveals a new level of insight into the serum N-glycome.

CRediT authorship contribution statement

Josh Smith: Methodology, Investigation, Formal analysis,

Writing – original draft. **Silvia Millán-Martín:** Methodology, Investigation, Formal analysis, Writing – original draft, Writing. **Stefan Mittermayr:** Methodology, Investigation. **Vivian Hilborne:** Methodology, Investigation. **Gavin Davey:** Supervision, Writing – original draft, Writing. **Karol Polom:** Methodology, Investigation. **Franco Roviello:** Methodology, Investigation. **Jonathan Bones:** Conceptualization, Funding acquisition, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to gratefully acknowledge the funding provided by the EU Framework Program 7 through the Marie Curie Actions Innovative Training Network grants FP7-PEOPLE-316929 GastricGlycoExplorer.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aca.2021.338840.

References

[1] A. Varki, Biological roles of glycans, Glycobiology 27 (2017) 3–49.
[2] R.A. Dwek, Glycobiology: toward understanding the function of sugars, Chem. Rev. 96 (1996) 683–720.
[3] L. Oliveira-Ferrer, K. Legler, K. Milde-Langosch, Role of protein glycosylation in cancer metastasis, Semin. Canc. Biol. 44 (2017) 141–152.
[4] J. Bones, S. Mittermayr, N. O'Donoghue, A. Guttman, P.M. Rudd, Ultra performance liquid chromatographic profiling of serum N-glycans for fast and efficient identification of cancer associated alterations in glycosylation, Anal. Chem. 82 (2010) 10208–10215.
[5] Y. Nakayama, N. Nakamura, D. Tsuji, K. Itoh, A. Kurosaka, Genetic diseases associated with protein glycosylation disorders in mammals, in: M. Puiu (Ed.), Genetic Disorders, InTech, Rijeka (2013) 243–269, pp. Ch. 10.
[6] Y. Kizuka, S. Kitazume, N. Taniguchi, N-glycan and Alzheimer's disease, Biochimica et biophysica acta, 2017.
[7] L. Zhang, S. Luo, B. Zhang, Glycan analysis of therapeutic glycoproteins, mAbs 8 (2016) 205–215.
[8] A.R. Costa, M.E. Rodrigues, M. Henriques, R. Oliveira, J. Azeredo, Glycosylation: impact, control and improvement during therapeutic protein production, Crit. Rev. Biotechnol. 34 (2014) 281–299.
[9] H.J. An, S.R. Kronewitter, M.L. de Leoz, C.B. Lebrilla, Glycomics and disease markers, Curr. Opin. Chem. Biol. 13 (2009) 601–607.
[10] D. Aldredge, H.J. An, N. Tang, K. Waddell, C.B. Lebrilla, Annotation of a serum N-glycan library for rapid identification of structures, J. Proteome Res. 11 (2012) 1958–1968.
[11] T. Song, D. Aldredge, C.B. Lebrilla, A method for in-depth structural annotation of human serum glycans that yields biological variations,, Anal. Chem. 87 (2015) 7754–7762.
[12] R. Saldova, A. Asadi Shehni, V.D. Haakensen, I. Steinfeld, M. Hilliard, I. Kifer, A. Helland, Z. Yakhini, A.L. Borresen-Dale, P.M. Rudd, Association of N-glycosylation with breast carcinoma and systemic features using high-resolution quantitative UPLC, J. Proteome Res. 13 (2014) 2314–2327.
[13] M.S. Bereman, T.I. Williams, D.C. Muddiman, Development of a nanoLC LTQ orbitrap mass spectrometric method for profiling glycans derived from plasma from healthy, benign tumor control, and epithelial ovarian cancer patients, Anal. Chem. 81 (2009) 1130–1136.
[14] S.R. Kronewitter, H.J. An, M.L. de Leoz, C.B. Lebrilla, S. Miyamoto, G.S. Leiserowitz, The development of retrosynthetic glycan libraries to profile and classify the human serum N-linked glycome,, Proteomics 9 (2009) 2986–2994.
[15] C.S. Chu, M.R. Ninonuevo, B.H. Clowers, P.D. Perkins, H.J. An, H. Yin, K. Killeen, S. Miyamoto, R. Grimm, C.B. Lebrilla, Profile of native N-linked glycan structures from human serum using high performance liquid chromatography on a microfluidic chip and time-of-flight mass spectrometry, Proteomics 9 (2009) 1939–1951.
[16] A.G. Gebrehiwot, D.S. Melka, Y.M. Kassaye, I.F. Rehan, S. Rangappa, H. Hinou, T. Kamiyama, S.I. Nishimura, Healthy human serum N-glycan profiling reveals the influence of ethnic variation on the identified cancer-relevant glycan biomarkers, PloS One 13 (2018) e0209515.
[17] B. Meszaros, G. Jarvas, A. Farkas, M. Szigeti, Z. Kovacs, R. Kun, M. Szabo, E. Csanky, A. Guttman, Comparative analysis of the human serum N-glycome in lung cancer, COPD and their comorbidity using capillary electrophoresis, J Chromatogr B Analyt Technol Biomed Life Sci 1137 (2020) 121913.
[18] C. Varadi, V. Hajdu, F. Farkas, I. Gilanyi, C. Olah, B. Viskolcz, The analysis of human serum N-glycosylation in patients with primary and metastatic brain tumors,, Life 11 (2021) 29.
[19] S. Carillo, S. Mittermayr, A. Farrell, S. Albrecht, J. Bones, Glycosylation analysis of therapeutic glycoproteins produced in CHO cells, in: P. Meleady (Ed.), Heterologous Protein Production in CHO Cells: Methods and Protocols, Springer New York, New York, NY, 2017, pp. 227–241.
[20] F. Tousi, J. Bones, W.S. Hancock, M. Hincapie, Differential chemical derivatization integrated with chromatographic separation for analysis of isomeric sialylated N-glycans: a nano-hydrophilic interaction liquid chromatography-MS platform, Anal. Chem. 85 (2013) 8421–8428.
[21] A. Ceroni, K. Maass, H. Geyer, R. Geyer, A. Dell, S.M. Haslam, GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans, J. Proteome Res. 7 (2008) 1650–1659.
[22] S. Albrecht, S. Mittermayr, J. Smith, S.M. Martin, M. Doherty, J. Bones, Twoplex 12/13 C6 aniline stable isotope and linkage-specific sialic acid labeling 2D-LC-MS workflow for quantitative N-glycomics, Proteomics 17 (2017).
[23] F. Tousi, J. Bones, O. Iliopoulos, W.S. Hancock, M. Hincapie, Multidimensional liquid chromatography platform for profiling alterations of clusterin N-glycosylation in the plasma of patients with renal cell carcinoma, J. Chromatogr., A 1256 (2012) 121–128.
[24] J. Bones, J.C. Byrne, N. O'Donoghue, C. McManus, C. Scaife, H. Boissin, A. Nastase, P.M. Rudd, Glycomic and glycoproteomic analysis of serum from patients with stomach cancer reveals potential markers arising from host defense response mechanisms, J. Proteome Res. 10 (2011) 1246–1265.
[25] K. Marino, J. Bones, J.J. Kattla, P.M. Rudd, A systematic approach to protein glycosylation analysis: a path through the maze,, Nat. Chem. Biol. 6 (2010) 713–723.
[26] D.J. Harvey, Fragmentation of negative ions from carbohydrates: part 2. Fragmentation of high-mannose N-linked glycans, J. Am. Soc. Mass Spectrom. 16 (2005) 631–646.
[27] D.J. Harvey, Fragmentation of negative ions from carbohydrates: part 3. Fragmentation of hybrid and complex N-linked glycans, J. Am. Soc. Mass Spectrom. 16 (2005) 647–659.
[28] Y. Kizuka, N. Taniguchi, Enzymes for N-glycan branching and their genetic and nongenetic regulation in cancer, Biomolecules 6 (2016) 25.
[29] M. Nagae, M. Kanagawa, K. Morita-Matsumoto, S. Hanashima, Y. Kizuka, N. Taniguchi, Y. Yamaguchi, Atomic visualization of a flipped-back conformation of bisected glycans bound to specific lectins, Sci. Rep. 6 (2016) 22973.
[30] S. Re, W. Nishima, N. Miyashita, Y. Sugita, Conformational flexibility of N-glycans in solution studied by REMD simulations, Biophys Rev 4 (2012) 179–187.
[31] W. Nishima, N. Miyashita, Y. Yamaguchi, Y. Sugita, S. Re, Effect of bisecting GlcNAc and core fucosylation on conformational properties of biantennary complex-type N-glycans in solution, J. Phys. Chem. B 116 (2012) 8504–8512.
[32] S. André, T. Kožár, S. Kojima, C. Unverzagt, H.J. Gabius, From structural to functional glycomics: core substitutions as molecular switches for shape and lectin affinity of N-glycans, Biol. Chem. 390 (2009) 557.
[33] S. Andre, C. Unverzagt, S. Kojima, M. Frank, J. Seifert, C. Fink, K. Kayser, C.W. von der Lieth, H.J. Gabius, Determination of modulation of ligand properties of synthetic complex-type biantennary N-glycans by introduction of bisecting GlcNAc in silico, in vitro and in vivo, Eur. J. Biochem. 271 (2004) 118–134.
[34] C. Schwedler, M. Kaup, S. Weiz, M. Hoppe, E.I. Braicu, J. Sehouli, B. Hoppe, R. Tauber, M. Berger, V. Blanchard, Identification of 34 N-glycan isomers in human serum by capillary electrophoresis coupled with laser-induced fluorescence allows improving glycan biomarker discovery, Anal. Bioanal. Chem. 406 (2014) 7185–7193.
[35] N. de Haan, S. Yang, J. Cipollo, M. Wuhrer, Glycomics studies using sialic acid derivatization and mass spectrometry, Nature Reviews Chemistry 4 (2020) 229–242.
[36] S.F. Wheeler, P. Domann, D.J. Harvey, Derivatization of sialic acids for stabilization in matrix-assisted laser desorption/ionization mass spectrometry and concomitant differentiation of alpha(2–> 3)- and alpha(2–> 6)-isomers, Rapid communications in mass spectrometry, RCM (Rapid Commun. Mass Spectrom.) 23 (2009) 303–312.
[37] Y. Huang, Y. Nie, B. Boyes, R. Orlando, Resolving isomeric glycopeptide glycoforms with hydrophilic interaction chromatography (HILIC), J. Biomol. Tech. 27 (2016) 98–104.
[38] S. Tao, Y. Huang, B.E. Boyes, R. Orlando, Liquid chromatography-selected reaction monitoring (LC-SRM) approach for the separation and quantitation of sialylated N-glycans linkage isomers, Anal. Chem. 86 (2014) 10584–10590.
[39] T. Yoshimura, A. Hayashi, M. Handa-Narumi, H. Yagi, N. Ohno, T. Koike, Y. Yamaguchi, K. Uchimura, K. Kadomatsu, J. Sedzik, K. Kitamura, K. Kato, B.D. Trapp, H. Baba, K. Ikenaka, GlcNAc6ST-1 regulates sulfation of N-glycans and myelination in the peripheral nervous system, Sci. Rep. 7 (2017) 42257.
[40] S. Hua, H.N. Jeong, L.M. Dimapasoc, I. Kang, C. Han, J.S. Choi, C.B. Lebrilla, H.J. An, Isomer-specific LC/MS and LC/MS/MS profiling of the mouse serum N-glycome revealing a number of novel sialylated N-glycans, Anal. Chem. 85 (2013) 4636–4643.

[41] S.M. Muthana, C.T. Campbell, J.C. Gildersleeve, Modifications of glycans: biological significance and therapeutic opportunities, ACS Chem. Biol. 7 (2012) 31–43.
[42] R. Schauer, Sialic acids: fascinating sugars in higher animals and man, Zoology (Jena) 107 (2004) 49–64.
[43] K. Yamada, K. Suzuki, Y. Hirohata, M. Kinoshita, Analysis of minor acidic N-glycans in human serum, J. Proteome Res. 19 (2020) 3033–3043.
[44] E. Llop, R. Gutierrez-Gallego, J. Segura, J. Mallorqui, J.A. Pascual, Structural analysis of the glycosylation of gene-activated erythropoietin (epoetin delta, Dynepo), Anal. Biochem. 383 (2008) 243–254.
[45] A. Varki, R. Schauer, Sialic Acids, in: A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, M.E. Etzler (Eds.), Essentials of Glycobiology, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY), 2009.
[46] J. Haverkamp, H. van Halbeek, L. Dorland, J.F. Vliegenthart, R. Pfeil, R. Schauer, High-resolution 1H-NMR spectroscopy of free and glycosidically linked O-acetylated sialic acids, Eur. J. Biochem. 122 (1982) 305–311.
[47] S.Y. Yu, L.Y. Chang, C.W. Cheng, C.C. Chou, M.N. Fukuda, K.H. Khoo, Priming mass spectrometry-based sulfoglycomic mapping for identification of terminal sulfated lacdiNAc glycotope, Glycoconj. J. 30 (2013) 183–194.
[48] G. Hernandez Mir, J. Helin, K.P. Skarp, R.D. Cummings, A. Makitie, R. Renkonen, A. Leppanen, Glycoforms of human endothelial CD34 that bind L-selectin carry sulfated sialyl Lewis x capped O- and N-glycans, Blood 114 (2009) 733–741.
[49] D.T. Kenny, S.M. Issa, N.G. Karlsson, Sulfate migration in oligosaccharides induced by negative ion mode ion trap collision-induced dissociation, Rapid communications in mass spectrometry, RCM (Rapid Commun. Mass Spectrom.) 25 (2011) 2611–2618.
[50] J.R. Wang, W.N. Gao, R. Grimm, S. Jiang, Y. Liang, H. Ye, Z.G. Li, L.F. Yau, H. Huang, J. Liu, M. Jiang, Q. Meng, T.T. Tong, H.H. Huang, S. Lee, X. Zeng, L. Liu, Z.H. Jiang, A method to identify trace sulfated IgG N-glycans as biomarkers for rheumatoid arthritis,, Nat. Commun. 8 (2017) 631.
[51] S. Takahashi, T. Sugiyama, M. Shimomura, Y. Kamada, K. Fujita, N. Nonomura, E. Miyoshi, M. Nakano, Site-specific and linkage analyses of fucosylated N-glycans on haptoglobin in sera of patients with various types of cancer: possible implication for the differential diagnosis of cancer, Glycoconj. J. 33 (2016) 471–482.
[52] S.H. Lee, S. Jeong, J. Lee, I.S. Yeo, M.J. Oh, U. Kim, S. Kim, S.H. Kim, S.Y. Park, J.H. Kim, S.H. Park, J.H. Kim, H.J. An, Glycomic profiling of targeted serum haptoglobin for gastric cancer using nano LC/MS and LC/MS/MS, Mol. Biosyst. 12 (2016) 3611–3621.
[53] X. Li, F. Guan, D. Li, Z. Tan, G. Yang, Y. Wu, Z. Huang, Identification of aberrantly expressed glycans in gastric cancer by integrated lectin microarray and mass spectrometric analyses, Oncotarget 7 (2016) 87284–87300.
[54] Y. Lan, C. Hao, X. Zeng, Y.L. He, P.J. Zeng, Z.H. Guo, L.J. Zhang, Serum glycoprotein-derived N- and O-linked glycans as cancer biomarkers, American Journal of Cancer Research 6 (2016) 2390–2415.