The Mouse N-Glycome Atlas – High-resolution N-glycan analysis of 23 mouse tissues.

Johannes Helm[1], Stefan Mereiter[2], Tiago Oliveira[2], Josef M. Penninger[2,3], Friedrich Altmann[1], Johannes Stadlmann[1,4].

Affiliations

1 Institute of Biochemistry, Department of Chemistry, University of Natural Resources and Life Sciences (BOKU), Muthgasse 18, 1190 Vienna, Austria.

2 Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), Vienna BioCenter (VBC), Dr. Bohr-Gasse 3, 1030 Vienna, Austria.

3 Department of Medical Genetics, Life Sciences Institute, University of British Columbia, Vancouver Campus, 2350 Health Sciences Mall, Vancouver, BC Canada V6T 1Z3.

4 BOKU Core Facility Mass Spectrometry, University of Natural Resources and Life Sciences (BOKU), Muthgasse 18, 1190 Vienna, Austria.

Correspondence

Johannes Stadlmann
Institute of Biochemistry, Department of Chemistry,
University of Natural Resources and Life Sciences (BOKU)
Muthgasse 18
A-1190 Vienna
Austria.
email: j.stadlmann@boku.ac.at

Abstract (maximum 170 words)

Over 50% of all human proteins are post-translationally modified by glycans, which alter their functions in fundamental biological processes. Mice (mus musculus) represent the most common mammalian model organism to study fundamental biological processes, including glycosylation. Because of isomeric variance, the comprehensive analysis of protein-linked N-glycans is a challenging task. Coupled with high-resolution mass-spectrometry, porous graphitic carbon (PGC)-LC is a powerful and widely used method for isomer-specific glycan analysis. Here we report on a consistent and complete PGC-LC-MS/MS N-glycome dataset of 23 different mouse tissues, critically complementing existing glycan data-repositories. We used PGC-LC to chromatographically separate even closely related N-glycan isomers and an Orbitrap Exploris 480 mass-spectrometer for data-acquisition. Multivariate data analysis revealed tissue specific N-glycome signatures, highlighted organ-intrinsic regulations of glycobiological pathways and confirmed prior glycobiological knowledge, as exemplified by the brain N-glycome. This dataset can be used for fundamental glycobiological research, glycan analytical benchmarking, the development of new mass-spectrometric data analysis tools, glycobiological pathway modelling and simulation, as well as for integrative systems biology.

Introduction

Protein glycosylation, the covalent attachment of simple or complex sugar structures to amino-acid side chains of polypeptides, affects virtually all aspects of biology. Over 50% of all human proteins are subject to post-translational modifications by glycans[1], which alter their functions in fundamental biological processes, such as cell adhesion, signal transduction, intracellular trafficking, essential immune functions[2] or host-pathogen interactions[3].

In the process of mammalian protein glycosylation, glycan structures are co- or post-translationally linked to nascent polypeptide chains, which then become extensively processed along the secretory pathway, throughout the Golgi-network to the cell surface. In protein N-glycosylation, arguably the currently best understood form of protein glycosylation, a large, highly conserved N-glycan precursor structure (comprising 2 N-acetylglucosamine, 9 mannose and 3 glucose molecules) is *en bloc* covalently linked to the side chains of specific asparagine-residues within the endoplasmic reticulum (ER). Subsequent N-glycan processing is a non-template driven process which critically depends on the coordinated action of dedicated enzymes, a highly specific set of glycosyltransferases and glycosidases[1]. Perturbations in any of these tightly interconnected and subtly tuned cellular processes hold the potential to result in differentially processed glycoproteins exhibiting aberrant glycan structures. The accumulation of such abnormal glycan structures has long been found as hallmark to many human diseases[3–8].

Owing to its close resemblance to the human physiology, its short generation time and its visible phenotypic variants, the mouse (*mus musculus*) represent the most common mammalian model organism to study fundamental biological processes, including glycosylation[9]. Nevertheless, striking differences between human and murine glycosylation exist. Most importantly, humans lack the Galα1,3Gal epitope (due to an inactivation mutation in α1,3GT (Ggta1)[10]) and the biosynthetic capabilities for the generation of N-glycolylneuraminic acid (due to a mutation in the CMP-NeuAc hydroxylase[11] (Cmahp)). Further differences between human and murine glycosylation have recently also been reported for brain[12] and immune cells[13].

The comprehensive analysis of protein-linked N-glycan structures is a highly challenging task. Due to compositional, structural, positional and anomeric isomers[14], the number of possible N-glycan structures vastly exceeds the number of oligomers formed by other (essentially linear) biopolymers, such as amino acids or nucleotides[15]. This analytical challenge has long been tackled using a variety of glycomics techniques, which all profile different aspects of glycosylation and exhibit specific analytical advantages and disadvantages. Importantly however, the use of this wide range of methodologies, including lectin-arrays[16], MALDI-TOF MS analysis of underivatized[17] , permethylated [18] or derivatized glycans[19,20], fluorescence labelling of glycans followed by either reversed- or normal-phase HPLC analysis[21,22], to various LC-MS and LC-MS/MS analysis workflows[23,24], leads to low comparability between studies.

Generation of glycobiological ground-truth was pioneered by the Consortium of Functional Glycomics glycan profiling initiative (CFG, www.functionalglycomics.org) who generated and made publicly available the hitherto most comprehensive, consistent, and coherent glycome datasets. From 2001 to 2011, the CFG Analytical Glycotechnology Core analysed protein N- and O-linked glycans from 11 human and 16 murine tissues, primarily using MALDI-TOF MS analysis of permethylated glycans. While the CFG's analytical activities have come to an end, these seminal glycan profiling datasets still represent a vital and fundamental resource to the field of mammalian glycobiology.

More recently, owing to its exceptional chromatographic selectivity, porous graphitic carbon (PGC) emerged as a new state of the art in glycome analysis. Coupled to MS, the separation power of PGC, which is able to separate even closely related isomers[25], became a new gold standard in isomer-specific analysis of glycans, and is nowadays routinely applied by many groups for the analysis of both, O-glycans[26–30] and N-glycans[14,30–39]. Complementing existing glycan data-repositories with critical isomer-specific analysis, we here report on and make publicly available a consistent and complete PGC-LC-MS/MS N-glycome dataset of 23 different mouse tissues (n=2), which is, to our knowledge, the currently most comprehensive mouse N-glycome data set. We used a shape selective stationary phase (i.e. porous graphitic carbon, PGC) to chromatographically separate even closely related (PNGase F-released, reduced) N-glycan isomers[14,40], and high-resolution tandem mass-spectrometry (Orbitrap

Exploris 480) for data-acquisition. We anticipate that this data will be useful for fundamental glycobiological research, glycan analytical benchmarking, the development of mass-spectrometric data-interpretation tools (e.g. MS/MS spectral libraries), biological pathway modelling and simulation, as well as for other integrative systems biology approaches.

Results and Discussion

To generate a representative and consistent dataset, we sacrificed two female 2C57BL/6J mice, collected their tissues in duplicates, and processed all samples using identical protein extraction, enzymatic N-glycan release, and glycan-clean-up protocols. Similarly, all samples were analysed using identical LC-MS/MS parameters.

To assess LC-MS raw-data quality, we first established mass-precision and retention time-stability for three representative N-glycan structures that are expected to be present in all 23 tissue samples (i.e. (1) Man5; (2) oligo-mannose isomer Man8"A"; (3) biantennary, di-(2,3-N-acetylneuraminic acid) sialylated and core-fucosylated glycan $Na^{2-3}Na^{2-3}F$). Close inspection of extracted ion chromatograms (XICs) and corresponding mass-spectra confirmed detection, highly reproducible retention-times (Man5 mean R_t = 26.88 min (+/- 0. 26 min); Man8"A" mean R_t = 23.5 min (+/- 0.26 min); $Na^{2-3}Na^{2-3}F$ mean R_t = 34.22 min (+/- 0.39 min)) and mass-precision (i.e. < 3ppm) for all three compounds, across 42 individual samples (Fig. 1). Of note, three individual samples were re-analyzed on a different, yet identically equipped and operated, LC-MS/MS system, which resulted in slightly deviating absolute retention times. The raw data files of the respective samples are indicated by the file name tag "_re".

Next, we assessed the occurrence and frequency of glycan-derived MS/MS spectra in the dataset. For this we converted all MS/MS spectra from all raw files into .mgf files and automatically extracted intensity values of characteristic, glycan-associated oxonium fragment ions (Fig. 2), using custom code. This analysis confirmed a vast fraction of the tandem mass-spectrometric raw data to derive from glycans (i.e. cooccurrence of HexNAc and HexNAc-Hex derived oxonium ions 204.08779 amu and 366.14059 amu, respectively) across all 46 samples. Remarkably, this analysis also highlighted tissue-specific glycome differences (Fig. 2). For example, the HexNAc-Hex-dHex derived oxonium ions of 512.19849 amu, that are indicative (but -due to fucose-migration reactions upon HCD- not diagnostic[43,44]) for distal Lewis-type fucosylation, were observed at exceptionally high frequency in kidney and seminal vesicles (Fig. 2). Moreover, while most tissues present with comparable frequencies of MS/MS spectra that contain NeuNAc and/or NeuNGc derived oxonium ions (i.e. HexNAc-Hex-NeuNAc of 657.23599 amu and HexNAc-Hex-NeuNGc of 673.23089 amu; indicated by purple and

light grey points, respectively, in Fig. 2), brain samples exhibit a clear over-representation of NeuNAc containing spectra (Fig. 2).

To further corroborate technical validity of the dataset, we used custom code to automatically charge-deconvolute and deisotope all individual mass-spectra, before summing their intensity values across the entire chromatographic elution time-range (i.e. 15 – 40 min) within the mass-range of 1000 to 3000 amu (bin width = 1 amu) of each sample, individually. This analysis allowed us to reconstruct semi-quantitative precursor-ion intensity histograms (Fig. 3) that closely resemble MALDI-TOF spectra frequently used in the field of glycomics (e.g. CFG, www.functionalglycomics.org). Of note, in contrast to generic two-dimensional single-stage (e.g. MALDI-TOF) mass-spectra, which merely allow for glycan-composition analysis, our dataset retains crucial, glycan-structure associated, three-dimensional information (i.e. mass over intensity over PGC-chromatographic retention time), as well as tandem mass-spectrometric information for glycan structure identification.

Based on the precursor-ion intensity histograms generated from our LC-MS/MS dataset, we then performed multivariate data analysis[45]. For this, we selected the 500 most intense mass-bins across all histograms and extracted their sample-specific intensity values for principal component analysis (PCA). This analysis (Fig. 4a) confirmed good reproducibility between sample replicates. Most importantly, however, PCA also highlighted unexpected tissue-specific clustering of our N-glycome datasets (Fig. 4b). For example, we discovered convergent N-glycan signatures for white adipose tissue, lymph node and mammary gland, which may be explained by lymph nodes and the mammary glands being embedded in white adipose tissue. Furthermore, we found major organs of the digestive tract (i.e. colon, ileum, jejunum, duodenum) to cluster, with jejunum and duodenum sharing highly similar N-glycomes. Similarly, PCA clusters revealed common N-glycosylation features for exocrine organs, such as seminal vesicle and pancreas, the central organs of the immune system (i.e. spleen, thymus), and cardio-pulmonary tissues (i.e. heart and lung).

These observations prompted us to further investigate tissue-specific N-glycosylation events by targeted N-glycome data-analysis. For this, in a next step, we also automatically extracted the intensity values of all theoretical N-glycan masses, which we derived from a functional *in silico* mouse N-glycosylation

pathway model (Supplemental Data), from previously charge-deconvoluted and deisotoped datasets, at high mass-precision, using custom code. More specifically, we extracted precursor-ion intensity values of 950 masses (mass-tolerance +/- 0.05 Da.), eluting over the entire chromatographic elution time-range (i.e. 15 – 40 min), from each sample. Then, we projected glycan-mass intensity values onto our mouse N-glycobiosynthetic map. Visualizations of these data projections (Fig. 5) not only confirmed substantial differences in N-glycan expression levels between tissues, but also highlighted tissue-dependant regulations of specific glycosylation-pathway branches (Fig. 5).

For example, next to the most abundant N-glycan compositions detected for liver (i.e. complex type, NeuNGc-terminated N-glycans N4H5G2 and N4H5F1G2), our network analyses also highlighted the remarkable overexpression of phosphorylated oligomannosidic glycans (with compositions ranging from N2H3p1 to N2H6p1) at the expense of a concomitant underrepresentation of non-phosphorylated oligomannose-type glycans in this tissue.

By contrast, the kidney N-glycome features the unusually high expression of multi-fucosylated (e.g. N5H4F2, N5H4F3) and bisected (N5H5F3, N6H6F4) N-glycans, which is accompanied by a marked underrepresentation of sialylated N-glycans, when compared to most other tissues. This is in line with previous work[46] that shows high expression levels for α1,3 fucosyltransferase Fut9 and the neuraminidases Neu1 and Neu2, as well as low expression levels of sialyltransferase St3gal4 in kidney. Furthermore, expression profiling[46] also showed that Mgat3, the enzyme responsible for the modification of glycans with a bisecting GlcNAc, is highly expressed in the kidney.

We also compared N-glycosylation profiles of PCA-clustering tissues, such as spleen and thymus, which are both central organs of the immune system. Unexpectedly, these two morphologically very distinct tissues exhibit highly similar glycanbiosynthetic pathways, resulting in a clear under expression of oligomannosidic structures and an over expression of biantennary glycans capped with N-glycolylneuraminic acid, with and without a fucose residue. Similarly, also exocrine tissues (e.g. pancreas, seminal vesicle) were found to show convergent N-glycosylation pathways, featuring multi-fucosylation and alpha-galactosylation (e.g. N4H7F1, N5H8F2).

Our cross-tissue N-glycome analyses show that the brain is exhibiting a rather unique N-glycosylation pattern, which is greatly dominated by the abundance of a single oligomannosidic N-glycan (i.e. Man5). Brain N-glycosylation is known to be crucial for neurodevelopmental processes and for normal adult brain functions[47–50] and aberrant glycosylation often induces neurological disfunctions and is involved in other diseases like brain cancer[51]. The importance of glycosylation and the implications of aberrant glycosylation in this organ has been described by others[31,52–56] and has recently been reviewed by Klaric et al[57]. Given the wealth of glycome information on this organ, we chose the brain to illustrate the advantages of isomer resolved PGC-LC-MS/MS data set and to demonstrate comparability with other studies.

In good agreement with recent work of Williams et al.[53], our dataset corroborates that the murine brain N-glycome shows a distinct prevalence of lower molecular weight structures (< 2000 m/z) when compared to other tissues (Fig. 3). Furthermore, our data confirms[53] a generally high abundance of oligo-mannosidic structures ranging from Man5-Man9, with Man5 being by far the most abundant glycan structure in both samples. Specifically, Man5 was found with approximately three times higher intensity than the second most abundant glycan structure (i.e. N4H3F1) in brain samples. In contrast to previous studies[53], however, which report oligomannosidic glycan contents of approx. 60 %, our data suggest a content of merely 20-25 % (Figs. 3, 6). Even when considering NH_4^+ adduct-ion formation of oligomannosidic glycans their amount does not exceed 35% in our mouse brain dataset (data not shown).

Besides the exceptional abundance of oligomannose structures, we also found the brain N-glycome to show a distinct prevalence of biosynthetic pathways towards the synthesis of smaller, truncated glycans (e.g. N2H3F1, N3H3F1). We hypothesize that this reflects the unusual enzymatic activity of the brain-specific N-acetyl-hexosaminidase B (Hexb), which cleaves the GlcNAc-residue linked to the α1,3 branching mannose and is not part of canonical biosynthetic pathways for N-glycans[58]. Furthermore, our data show increased expression levels for multi-fucosylated and bisected glycans (e.g. N5H5F1, N5H5F2G1) when compared to other tissues, which is, again, in good agreement with brain gene-expression profiling data, that show Mgat3 and Fut9 to be upregulated[46]. Our data show that the content

of fucosylated structures (i.e. containing 1-5 fucose residues) in the brain range from 36-42 %, whereas the content of multi-fucosylated structures (i.e. more than 1 fucose residues) ranges from 12-14 %. Again this is in good agreement with Williams et al, reporting a fucosylated glycan content of 35 %[53]. The high amount of multi-fucosylated N-glycans can be attributed to FuT9, which determines Lewis X expression in the brain[59].

Next, we analysed the relative abundance of sialylated glycans in mouse brain. Intriguingly, almost all sialylated structures of the murine brain are modified with N-acetylneuraminic acid, rather than with N-glycolylneuraminic acid (Fig. 2). This is in agreement with prior results[53] and further highlights sample quality, as murine blood, the major source of contamination in biological specimen has been reported to exhibit an unusually high N-glycolylneuraminic acid content[60]. Furthermore, we found that 18-25 % of all glycan signals are derived from sialylated glycan structures (Fig. 6). While this is well in line with the results of Barboza et al. (i.e. approx. 20 %)[54], it is in stark contrast to Williams et al.[53], who report a sialylated glycan content of only 1-3 %[53].

Eventually, to also showcase the advantage of PGC-LC-MS/MS in glycome analysis, we specifically analysed the brain sample dataset for all N-glycan structures with a composition 5 hexoses, 4 N-acetyl-hexosamines and 1 deoxy-hexose (i.e. N4H5F1). The XIC (Fig. 7) shows that at least four isomers are present at the MS1 level. Based on relative retention time recently described[14,31], we could identify two hybrid type glycans (one of them with an elongated bisecting GlcNAc) and a bi-antennary galactosylated structure with a core fucose. Importantly, these biosynthetically distant, yet isobaric glycan structures of identical composition would not be discriminated without chromatographic separation of the sample. This highlights that elution rules and relative elution positions for PGC chromatography, which are extensively described elsewhere[14,25,31,34,61,62], are a valuable tool for the exact identification of isomers and a vital resource for the analysis of this dataset. In addition, the highly reproducible MS/MS data provided here can be used to confirm particular isomers or to identify unknown structures via spectral matching, as already described elswhere[61].

Taken together, we here report the, to our knowledge, the currently most comprehensive mouse N-glycome data set, comprising consistent PGC-LC-MS/MS analysis of 23 different mouse tissues (n=2),

that exceeds and thus crucially complements currently existing primary glycobiological data-repositories, essential for modern integrative system biology efforts.

Methods

Preparation of mouse tissues

Two Male and two female 2C57BL/6J mice were obtained from the Jackson Laboratory (PRID: IMSR_JAX:000664, Bar Harbor, ME). C57BL/6J mice were bred in the licensed breeding facility of the Institute of Molecular Biotechnology (IMBA, Vienna, AT), under a 14hrs/10hrs light/dark cycle. Food and water were available ad libitum. All mice were euthanized with carbon dioxide at the age of 13 weeks. All tissues were collected in triplicates from either female (F) of male (M) mice, and included: Skin from ear (F), whole esophagus (F), whole stomach (F), whole pancreas (F), middle section of duodenum (F), middle section of jejunum (F), middle section of ileum (F), middle section of colon (F), whole spleen (F), left kidney (F), whole liver (F), whole heart (F), whole lung (F), whole brain (F), quadriceps muscle (F), epididymal white adipose tissue (F), interscapular brown adipose tissue (M), inguinal mammary fat pad including mammary glands and excluding inguinal lymph node (F), inguinal lymph node (F), whole thymus (F), whole uterus (F), left testicle (M), whole penis (M), whole bladder (M), serum (submandibular bleeding, M), left seminal vesicle (M). Prior to freezing, stomach, duodenum, jejunum, ileum, colon and bladder were opened and cleaned with PBS from food, fecal matter and/or urine. Serum was collected from blood with Microtainer SST tubes (BD Biosciences, 365968). All samples were immediately after collection snap frozen in liquid nitrogen and stored at -80°C until further processing.

N-glycan extraction

Each mouse tissue was transferred into a Falcon tube and mixed with 100 mM ammonium bicarbonate buffer containing 20 mM ditiothreitol and 2 % sodium dodecyl sulphate in a total volume of 2 mL. After homogenization with an Ultra-Turrax T25 disperser, the homogenate was incubated at 56 °C for 30 min. After cooling down, the solution was brought to 40 mM iodoacetamide and incubated at room temperature in the dark for 30 min. After centrifugal clarification, chloroform-methanol extraction of the proteins was carried out using standard protocols[41]. In brief, the supernatant was mixed with four volumes of methanol, one volume of chloroform and three volumes of water in the given order. After

centrifugation of the mixture for 3 min at 2500 rpm, the upper phase was removed, 4 mL of methanol were added, and the pellet was resuspended. The solution was centrifuged, the supernatant was removed, and the pellet was again resuspended in 4 mL methanol. The last step was repeated two more times. After the last methanol washing step, the pellet was dried at room temperature. The dried pellet was taken up in 50 mM ammonium acetate o (pH 8.4). For N-glycan release 2.5 U N-glycosidase F were added and the resulting mixture was incubated overnight at 37°C. The reaction mixture was acidified with some drops of glacial acetic acid and centrifuged at 4000×g. The supernatant was loaded onto a Hypersep C18 cartridge (1000 mg; Thermo Scientific, Vienna) that had been previously primed with 2 mL of methanol and equilibrated with 10 mL water. The sample was applied, and the column was washed with 4 mL water. The flow-through and the wash solution were collected and subjected to centrifugal evaporation. Reduction of the glycans was carried out in 1 % sodium borohydride in 50 mM NaOH at room temperature overnight. The reaction was quenched by the addition of two drops glacial acetic acid. Desalting was performed using HyperSep Hypercarb solid-phase extraction cartridges (25 mg) (Thermo Scientific, Vienna). The cartridges were primed with 450 µL methanol followed by 450 µL 80% acetonitrile in 10 mM ammonium bicarbonate. Equilibration was carried out by the addition of three times 450 µL water. The sample was applied and washed three times with 450 µL water. N-glycans were eluted by the addition of two times 450 µL 80% acetonitrile in 10 mM ammonium bicarbonate. The eluate was subjected to centrifugal evaporation and the dried N-glycans were taken up in 20 µL HQ-water. 5 µL of each sample were subjected to LC-MS/MS analysis

LC-MS/MS analysis

LC-MS analysis was performed on a Dionex Ultimate 3000 UHPLC system coupled to an Orbitrap Exploris 480 Mass Spectrometer (Thermo Scientific). The purified glycans were loaded on a Hypercarb column (100 mm × 0.32 mm, 5 µm particle size, Thermo Scientific, Waltham, MA, USA) with 10 mM ammonium bicarbonate as the aqueous solvent A and 80 % acetonitrile in 50 mM ammonium bicarbonate as solvent B. The gradient was as follows: 0 – 4.5 min 1 % B, from 4.5 – 5.5 min 1 – 9 % B, from 5.5 – 30 min 9 - 20 % B, from 30 – 41.5 min 20 - 35 % B, from 41.5 – 45 min 35 – 65 % B, followed by an equilibration period at 1 % B from 45 –55 min. The flowrate was 6 µL/min. MS analysis

was performed in data-dependent acquisition (DDA) mode with positive polarity from 500 – 1500 m/z using the following parameters: resolution was set to 60,000 with a normalized AGC target of 300%. The 10 most abundant precursors (charge states 2-6) within an isolation window of 1.4 m/z were selected for fragmentation. Dynamic exclusion was set at 20 s (n=1) with a mass tolerance of ± 10 ppm. Normalized collsion energy (NCE) for HCD was set to 20, 25 and 30% and the precursor intensity threshold was set to 8000. MS/MS spectra were recorded with a resolution of 15,000, using a normalized AGC target of 100% and a maximum accumulation time of 100 ms.

Data Analysis

For the generation of oxo-plots, MS2 data (in the .raw file format) were extracted and converted into .mgf files using the "vendor"-specific peak-picking algorithm implemented in MSConvert (version 3.0.22133-5eed1a6) and removing zero-samples. MS2 spectra were filtered for (N-)glycan specific oxonium-ions (i.e. 204.08779, 366.14059, 512.19849, 657.23599, 407.16719, 673.23089, 803.29389, 819.28879; mass-tolerance +/- 0.05 amu) using a custom-made perl-script ("MS2OxoPlot.pl"), binned according to retention time (bin size = 10s), counted and visualized as scatter-plots using Excel365 (Microsoft).

N-Glycan intensity heat-maps, histograms and extracted ion chromatograms (XICs) were all re-constructed from MS1data, that were extracted and converted into the .ms1 file format using the "vendor"-specific peak-picking algorithm (implemented in MSConvert version 3.0.22133-5eed1a6) and applying an absolute intensity threshold of 1E5. Individual MS1 scans were then converted into the .mgf file format using a custom made perl-script ("MS1toPseduoMGF.pl"), charge-deconvoluted and deisotoped using a custom-made perl-script ("SugarQBits_DD.pl"). Eventually, MS1 intensity values were extracted from the charge-deconvoluted and deisotoped MS1 data using a custom-made perl-script ("SugarQBits_RepX.pl") into .csv files. For non-targeted data analysis, the highest intensity values in mass bins of 1 amu (+/- 0.5 amu around the center-mass) were extracted in the mass range from 500.3 to 4500.3 amu over the full retention time-range. For N-glycan specific MS1 data analyses, the intensity values of 951 N-glycan specific masses (+/- 0.05 amu; "mouseNglycanDB.csv") were extracted ("SugarQBits_RepX.pl").

Multivariate data analysis was performed using ClustVis [45]. All other data visualization was performed using Excel365 (Microsoft) or Cytoscape version 3.9.1.

Data Records

The raw mouse N-glycome PGC-LC-MS/MS raw-data (".raw") and processed datasets referred to in this manuscript (in the ".csv" file format) are available for review via the GlycoPOST repository[42] (URL below). Tissue and sample replicate numbers are indicated in the respective file names.

https://glycopost.glycosmos.org/preview/20098543276311cb410c7de (PIN CODE: 9378).

Usage Notes

R-scripts and a brief manual for data-visualization of extracted LC-MS data using these scripts, a Skyline document file ("mouseNglycomeAtlas_TransitionList.sky") for the targeted analysis of all N-glycan compositions considered in this study ("mouseNglycanDB.csv"), are provided as Supplementary Information to this manuscript. The N-glycosylation biosynthesis pathway model (NglycoScape.xls) is provided as Supplementary Information to this manuscript.

Code Availability

All custom code and software for MS and MS/MS data pre-processing is freely available at: http://homepage.boku.ac.at/jstadlmann.

Acknowledgements

J.H. was partly funded by the European Commission (Newcotiana 760331). S.M. is funded by the ESPRIT-Programme of the Austrian Science Fund (FWF, Project number: ESP 166). J.M.P. received funding from the T. von Zastrow foundation, the FWF Wittgenstein award (Z 271-B19), the Austrian Academy of Sciences, the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No 101005026, and the Canada 150 Research Chairs Program F18-01336 as well as the Canadian Institutes of Health Research COVID-19 grants F20–02343 and F20-02015. This project was supported by MS equipment kindly provided the EQ-BOKU VIBT GmbH and the BOKU Core Facility Mass Spectrometry.

Abbreviations

PGC – porous graphitic carbon

HCD – higher energy collisional dissociation

AGC – automated gain control

NeuNAc or Na – N-acetyneuraminic acid

NeuNGc – N-glycolylneuraminic acid

HexNAc – N-acetylhexosamine

N5H5F1A1G1 – glycan structure composed of 5 N-acetylhexosamine (N), 5 hexose (H), 1 fucose (F), 1 N-acetylneuraminic acid (A) and 1 N-glycolylneuraminic acid (G) residues; phosphate groups are indicated by “p”.

Man5-9 – oligo mannosidic N-glycan composed of 2 N-acetylglucosamine residues and 5-9 mannose residues

References

1. Varki, A. *et al.* Essentials of Glycobiology. *Essentials Glycobiol.* (2022) doi:10.1101/9781621824213.

2. Ugonotti, J., Chatterjee, S. & Thaysen-Andersen, M. Structural and functional diversity of neutrophil glycosylation in innate immunity and related disorders. *Mol. Aspects Med.* **79**, 100882 (2021).

3. Zhang, X. L. & Qu, H. The Role of Glycosylation in Infectious Diseases. *Adv. Exp. Med. Biol.* **1325**, 219–237 (2021).

4. Rudman, N., Gornik, O. & Lauc, G. Altered N-glycosylation profiles as potential biomarkers and drug targets in diabetes. *FEBS Lett.* **593**, 1598–1615 (2019).

5. Verhelst, X. *et al.* Protein Glycosylation as a Diagnostic and Prognostic Marker of Chronic Inflammatory Gastrointestinal and Liver Diseases. *Gastroenterology* **158**, 95–110 (2020).

6. Reily, C., Stewart, T. J., Renfrow, M. B. & Novak, J. Glycosylation in health and disease. *Nat. Rev. Nephrol.* **15**, 346–366 (2019).

7. Mereiter, S., Balmaña, M., Campos, D., Gomes, J. & Reis, C. A. Glycosylation in the Era of Cancer-Targeted Therapy: Where Are We Heading? *Cancer Cell* **36**, 6–16 (2019).

8. Thomas, D., Rathinavel, A. K. & Radhakrishnan, P. Altered glycosylation in cancer: A promising target for biomarkers and therapeutics. *Biochim. Biophys. Acta - Rev. Cancer* **1875**, 188464 (2021).

9. Orr, S. L. *et al.* A phenotype survey of 36 mutant mouse strains with gene-targeted defects in glycosyltransferases or glycan-binding proteins. *Glycobiology* **23**, 363–380 (2013).

10. Macher, B. A. & Galili, U. The Galalpha1,3Galbeta1,4GlcNAc-R (alpha-Gal) epitope: a carbohydrate of unique evolution and clinical relevance. *Biochim. Biophys. Acta* **1780**, 75–88 (2008).

11. Varki, A. Loss of N-glycolylneuraminic acid in humans: Mechanisms, consequences, and implications for hominid evolution. *Am. J. Phys. Anthropol.* **Suppl 33**, 54 (2001).

12. C, A., RA, K. & B, S. Do rodent and human brains have different N-glycosylation patterns? *Biol. Chem.* **382**, 187–194 (2001).

13. Antonopoulos, A., North, S. J., Haslam, S. M. & Dell, A. Glycosylation of mouse and human immune cells: insights emerging from N-glycomics analyses. *Biochem. Soc. Trans.* **39**, 1334–1340 (2011).

14. Helm, J. *et al.* Bisecting Lewis X in Hybrid-Type N-Glycans of Human Brain Revealed by Deep Structural Glycomics. *Analytical Chemistry* 93 (45), 15175-15182 (2021).

15. Gabius, H. J. & Kayser, K. Introduction to glycopathology: the concept, the tools and the perspectives. *Diagn. Pathol.* **9**, (2014).

16. Nagai-Okatani, C. *et al.* LM-GlycomeAtlas Ver. 2.0: An Integrated Visualization for Lectin Microarray-based Mouse Tissue Glycome Mapping Data with Lectin Histochemistry. *J. Proteome Res.* **20**, 2069–2075 (2021).

17. Lee, J. W. *et al.* Potential of MALDI-TOF-based serum N-glycan analysis for the diagnosis and surveillance of breast cancer. *Sci. Reports 2020 101* **10**, 1–8 (2020).

18. Lin, Z. & Lubman, D. M. Permethylated N-glycan analysis with mass spectrometry. *Methods Mol. Biol.* **1007**, 289–300 (2013).

19. Zhang, Y. *et al.* Sensitive and robust MALDI-TOF-MS glycomics analysis enabled by Girard's reagent T on-target derivatization (GTOD) of reducing glycans. *Anal. Chim. Acta* **1048**, 105 (2019).

20. Reiding, K. R., Blank, D., Kuijper, D. M., Deelder, A. M. & Wuhrer, M. High-Throughput Profiling of Protein N - Glycosylation by MALDI- TOF-MS Employing Linkage-Specific Sialic Acid Esterification. *Analytical Chemistry* 86 (12), 5784-5793 (2014).

21. Kinoshita, M., Saito, A., Yamamoto, S. & Suzuki, S. A practical method for preparing fluorescent-labeled glycans with a 9-fluorenylmethyl derivative to simplify a fluorimetric HPLC-based analysis. *J. Pharm. Biomed. Anal.* **186**, 113267 (2020).

22. Zhang, L., Luo, S. & Zhang, B. Glycan analysis of therapeutic glycoproteins. *MAbs* **8**, 205–215 (2016).

23. Veillon, L. *et al.* Characterization of isomeric glycan structures by LC-MS/MS. *Electrophoresis* **38**, 2100–2114 (2017).

24. Peng, W. *et al.* MS-based glycomics and glycoproteomics methods enabling isomeric characterization. *Mass Spectrom. Rev.* (2021) doi:10.1002/MAS.21713.

25. Pabst, M., Bondili, J. S., Stadlmann, J., Mach, L. & Altmann, F. Mass + retention time = structure: A strategy for the analysis of N-glycans by carbon LC-ESI-MS and its application to fibrin N-glycans. *Anal. Chem.* **79**, 5051–5057 (2007).

26. Saldova, R. & Wilkinson, H. Current Methods for the Characterization of O-Glycans. *J. Proteome Res.* **19**, 3890–3905 (2020).

27. Wilkinson, H. *et al.* The O-Glycome of Human Nigrostriatal Tissue and Its Alteration in Parkinson's Disease. *J. Proteome Res.* **20**, 3913–3924 (2021).

28. Stavenhagen, K., Kolarich, D. & Wuhrer, M. Clinical Glycomics Employing Graphitized Carbon Liquid Chromatography–Mass Spectrometry. *Chromatographia* **78**, 307–320 (2014).

29. Jensen, P. H., Karlsson, N. G., Kolarich, D. & Packer, N. H. Structural analysis of N- and O-glycans released from glycoproteins. *Nat. Protoc. 2012 77* **7**, 1299–1310 (2012).

30. Mereiter, S. *et al.* Glycomic analysis of gastric carcinoma cells discloses glycans as modulators of RON receptor tyrosine kinase activation in cancer. *Biochim. Biophys. Acta - Gen. Subj.* **1860**, 1795–1808 (2016).

31. Helm, J. & Hirtler, L. Towards Mapping of the Human Brain N-Glycome with Standardized

Graphitic Carbon Chromatography. *Biomolecules* 12(1), 85 (2022).

32. Pabst, M. & Altmann, F. Influence of Electrosorption, Solvent, Temperature, and Ion Polarity on the Performance of LC-ESI-MS Using Graphitic Carbon for Acidic Oligosaccharides. *Analytical Chemistry* 80 (19), 7534-7542 (2008).

33. Everest-Dass, A. V., Abrahams, J. L., Kolarich, D., Packer, N. H. & Campbell, M. P. Structural feature ions for distinguishing N- and O-linked glycan isomers by LC-ESI-IT MS/MS. *J. Am. Soc. Mass Spectrom.* **24**, 895–906 (2013).

34. Ashwood, C., Pratt, B., Maclean, B. X., Gundry, R. L. & Packer, N. H. Standardization of PGC-LC-MS-based glycomics for sample specific glycotyping. *Analyst* **144**, 3601–3612 (2019).

35. Young, C. *et al.* In-House Packed Porous Graphitic Carbon Columns for Liquid Chromatography-Mass Spectrometry Analysis of N-Glycans. *Front. Chem.* **9**, 388 (2021).

36. Adamczyk, B., Stöckmann, H., O'Flaherty, R., Karlsson, N. G. & Rudd, P. M. High-throughput analysis of the plasma N-glycome by UHPLC. *Methods Mol. Biol.* **1503**, 97–108 (2017).

37. Chen, C. H. *et al.* Enhancement of fucosylated N-glycan isomer separation with an ultrahigh column temperature in porous graphitic carbon liquid chromatography-mass spectrometry. *J. Chromatogr. A* **1632**, 461610 (2020).

38. Abrahams, J. L., Packer, N. H. & Campbell, M. P. Relative quantitation of multi-antennary N-glycan classes: combining PGC-LC-ESI-MS with exoglycosidase digestion. *Analyst* **140**, 5444–5449 (2015).

39. Zhou, S. *et al.* Isomeric Separation of Permethylated Glycans by Porous Graphitic Carbon (PGC)-LC-MS/MS at High Temperatures. *Anal. Chem.* **89**, 6590–6597 (2017).

40. Pabst, M., Bondili, J. S., Stadlmann, J., Mach, L. & Altmann, F. Mass + retention time = structure: A strategy for the analysis of N-glycans by carbon LC-ESI-MS and its application to

fibrin N-glycans. *Anal. Chem.* **79**, 5051–5057 (2007).

41. Jiang, L., He, L. & Fountoulakis, M. Comparison of protein precipitation methods for sample preparation prior to proteomic analysis. *J. Chromatogr. A* **1023**, 317–320 (2004).

42. Watanabe, Y., Aoki-Kinoshita, K. F., Ishihama, Y. & Okuda, S. GlycoPOST realizes FAIR principles for glycomics mass spectrometry data. *Nucleic Acids Res.* **49**, D1523–D1528 (2021).

43. Lettow, M. *et al.* The role of the mobile proton in fucose migration. *Anal. Bioanal. Chem.* **411**, 4637–4645 (2019).

44. Mucha, E. *et al.* Fucose Migration in Intact Protonated Glycan Ions: A Universal Phenomenon in Mass Spectrometry. *Angew. Chem. Int. Ed. Engl.* **57**, 7440–7443 (2018).

45. Metsalu, T. & Vilo, J. ClustVis: A web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res.* **43**, W566–W570 (2015).

46. Nairn, A. V. *et al.* Regulation of Glycan Structures in Animal Tissues: TRANSCRIPT PROFILING OF GLYCAN-RELATED GENES. *J. Biol. Chem.* **283**, 17298–17313 (2008).

47. Schnaar, R. L., Gerardy-Schahn, R. & Hildebrandt, H. Sialic acids in the brain: Gangliosides and polysialic acid in nervous system development, stability, disease, and regeneration. *Physiol. Rev.* **94**, 461–518 (2014).

48. Klaus, C. *et al.* Reduced sialylation triggers homeostatic synapse and neuronal loss in middle-aged mice. *Neurobiol. Aging* **88**, 91–107 (2020).

49. Hildebrandt, H. *et al.* Imbalance of neural cell adhesion molecule and polysialyltransferase alleles causes defective brain connectivity. *Brain* **132**, 2831–2838 (2009).

50. Weinhold, B. *et al.* Genetic ablation of polysialic acid causes severe neurodevelopmental defects rescued by deletion of the neural cell adhesion molecule. *J. Biol. Chem.* **280**, 42971–

42977 (2005).

51. Veillon, L., Fakih, C., Abou-El-Hassan, H., Kobeissy, F. & Mechref, Y. Glycosylation Changes in Brain Cancer. *ACS Chem. Neurosci.* **9**, 51–72 (2018).

52. Lee, J. *et al.* Spatial and temporal diversity of glycome expression in mammalian brain. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 28743–28753 (2020).

53. Williams, S. E. *et al.* Mammalian brain glycoproteins exhibit diminished glycan complexity compared to other tissues. *Nat. Commun.* **13**, (2022).

54. Barboza, M. *et al.* Region-Specific Cell Membrane N-Glycome of Functional Mouse Brain Areas Revealed by nanoLC-MS Analysis. *Mol. Cell. Proteomics* 100130 (2021) doi:10.1016/j.mcpro.2021.100130.

55. Rebelo, A. L. *et al.* Complete spatial characterisation of N-glycosylation upon striatal neuroinflammation in the rodent brain. *J. Neuroinflammation* **18**, 1–19 (2021).

56. Samal, J., Saldova, R., Rudd, P. M., Pandit, A. & O'Flaherty, R. Region-Specific Characterization of N-Glycans in the Striatum and Substantia Nigra of an Adult Rodent Brain. *Anal. Chem.* **92**, 12842–12851 (2020).

57. Klarić, T. S. & Lauc, G. The dynamic brain N-glycome. *Glycoconj. J.* **39**, 443–471 (2022).

58. Okamoto, Y., Omichi, K., Yamanaka, S., Ikenaka, K. & Hase, S. Conversion of brain-specific complex type sugar chains by N-acetyl-β-D- hexosaminidase B1. *J. Biochem.* **125**, 537–540 (1999).

59. Nishihara, S. *et al.* ,3-Fucosyltransferase IX (Fut9) determines Lewis X expression in brain. *Glycobiology* **13**, 445–455 (2003).

60. Reiding, K. R., Hipgrave Ederveen, A. L., Rombouts, Y. & Wuhrer, M. Murine Plasma N-Glycosylation Traits Associated with Sex and Strain. *J. Proteome Res.* **15**, 3489–3499 (2016).

61. Alagesan, K. *et al.* Better together – relative retention time plus spectral matching improves

automated glycan characterization using PGC-nLC-IT-ESI-MS / MS We thank Max Planck Society for their generous financial support . KA was supported by PhD scholarship from Beilstein-. 1–22 (2021).

62. Abrahams, J. L., Campbell, M. P. & Packer, N. H. Building a PGC-LC-MS N-glycan retention library and elution mapping resource. *Glycoconj. J.* **35**, 15–29 (2018).

Figures

Figure 1.

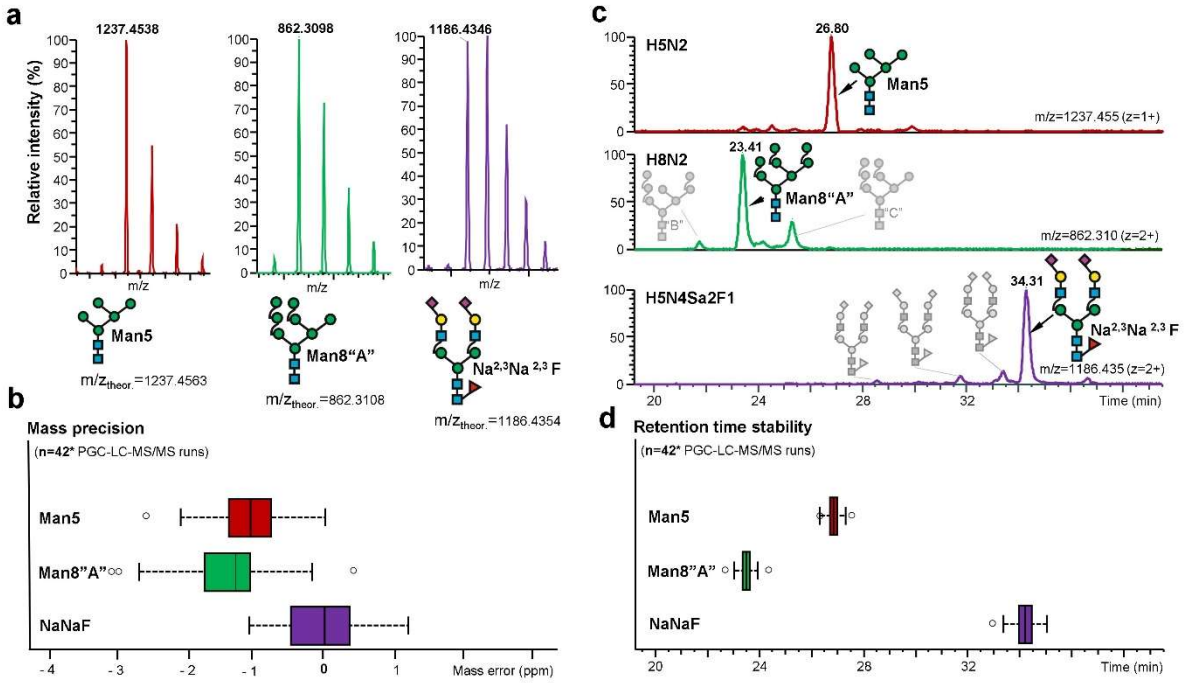

Figure 1 - Technical Validation of the PGC-LC-ESI-MS/MS measurements. a) Representative mass spectra and **b)** box-and-whiskers plots of mass precision for three representative structures (Man5, Man8 "A" and $Na^{2,3}Na^{2,3}F$) across 42individual samples. **c)** Representative XICs and **d)** box-and-whiskers plots of retention time fluctuations for three representative structures (Man5, Man8 "A and $Na^{2,3}Na^{2,3}F$) across 42individual samples. Of note, three samples were reanalyzed on a different LC-MS/MS setup, resulting in deviating absolute retention times. The raw data files of the respective samples are indicated by the file name tag "_re". N-glycan cartoons were drawn according to the symbol nomenclature SNFG (https://www.ncbi.nlm.nih.gov/glycans/snfg.html).

Figure 2.

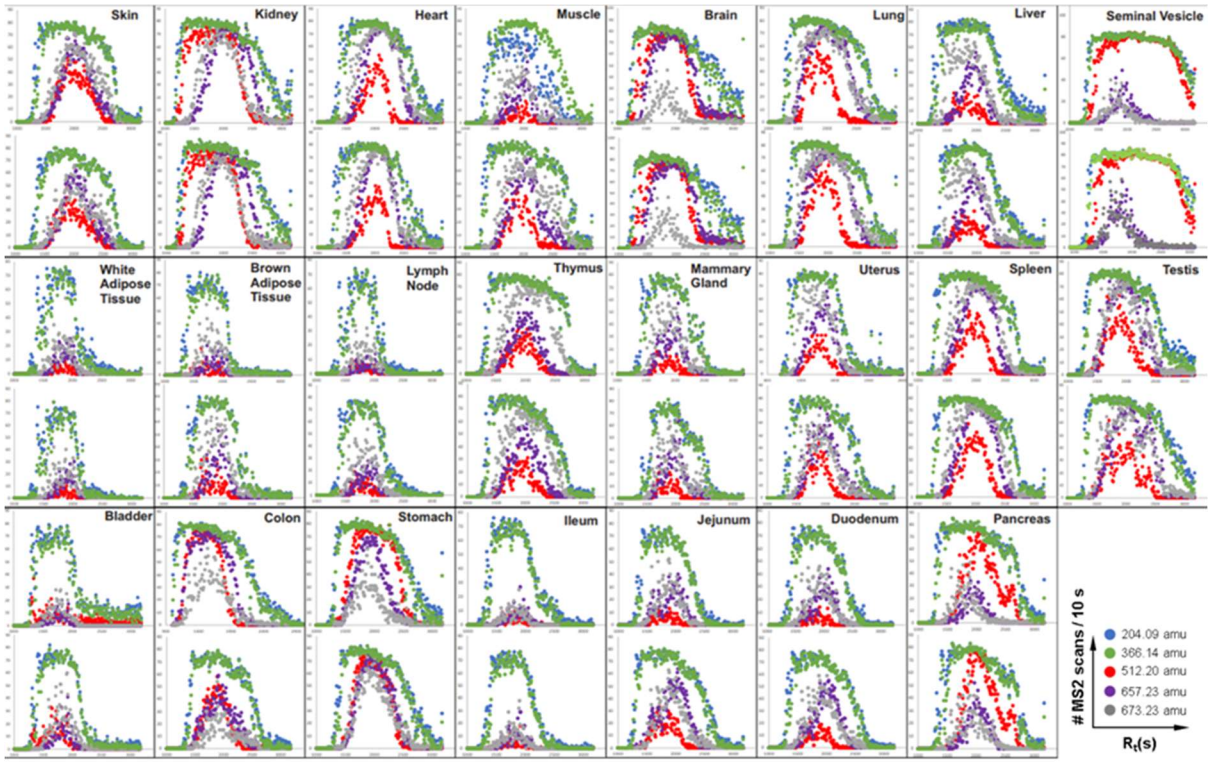

Figure 2 - "Oxoplots" of the N-glycomes across all mouse tissues analysed. Oxoplots depict the frequency of glycan-associated oxonium ions in MS/MS spectra over chromatographic retention time. Green dots represent HexNAc-Hex oxonium ion (366.14 m/z) diagnostic for all N-glycan structures. Red dots represent the HexNAc-Hex-Fuc oxonium ion (512.20 m/z), indicative (yet not diagnostic) for antenna fucosylation. Purple dots represent the HexNAc-Hex-Neu5Ac oxonium ion (657.23 m/z) which is diagnostic for Neu5Ac bearing N-glycan structures. Grey dots represent the HexNAc-Hex-Neu5Gc oxonium ion (673.23 m/z), which is diagnostic for Neu5Gc containing N-glycan structures. For quality control, all N-glycan spectra to be included in the oxoplot were required to contain the HexNAc ion (204.09 m/z; indicated by blue dots).

Figure 3.

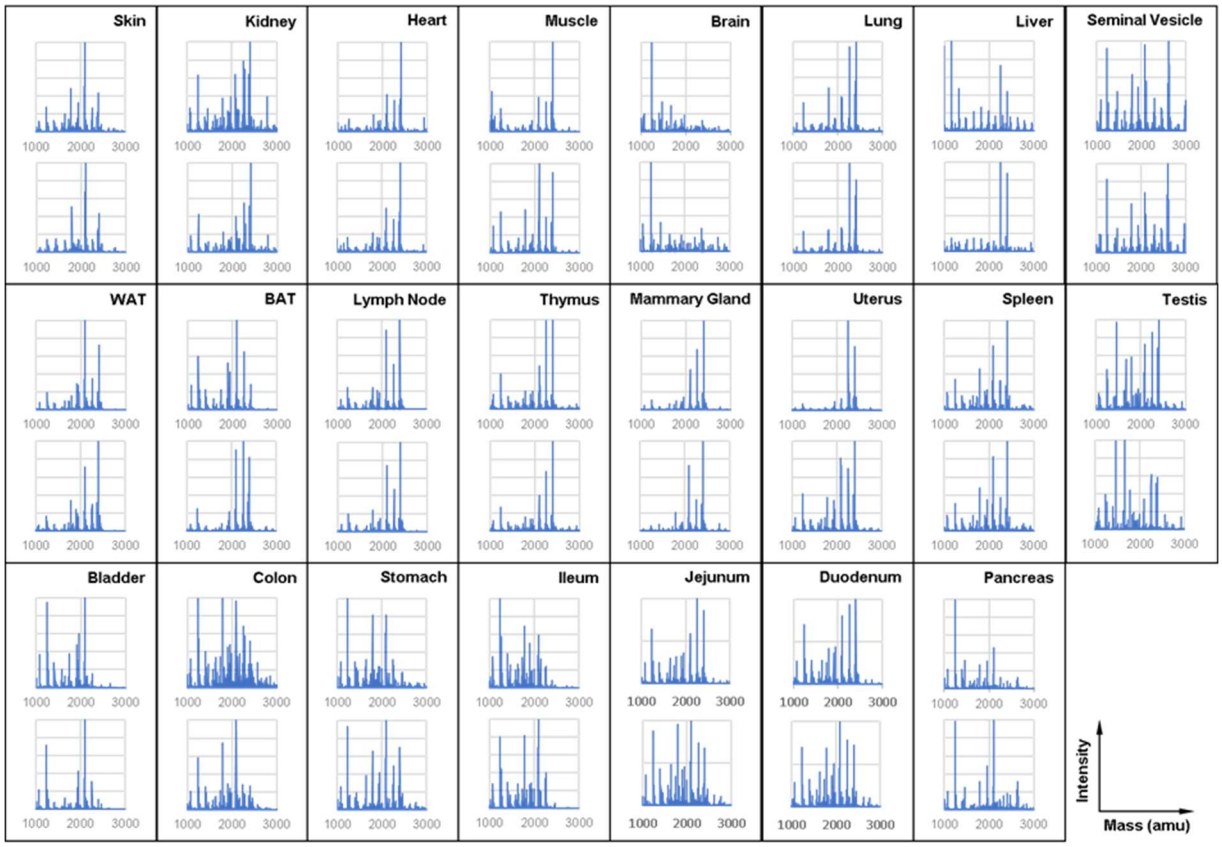

Figure 3 – Reconstructed "pseudo-MALDI"-spectra of mouse N-glycomes of the 23 mouse tissues analysed. Precursor ion intensity histograms (molecular mass range 1000 to 3000 amu; bin width = 1 amu) of all 46 samples were reconstructed from charge-deconvoluted and deisotoped LC-MS1data, summed over the chromatographic retention time-range from 15.0 to 40.0 mins.

Figure 4.

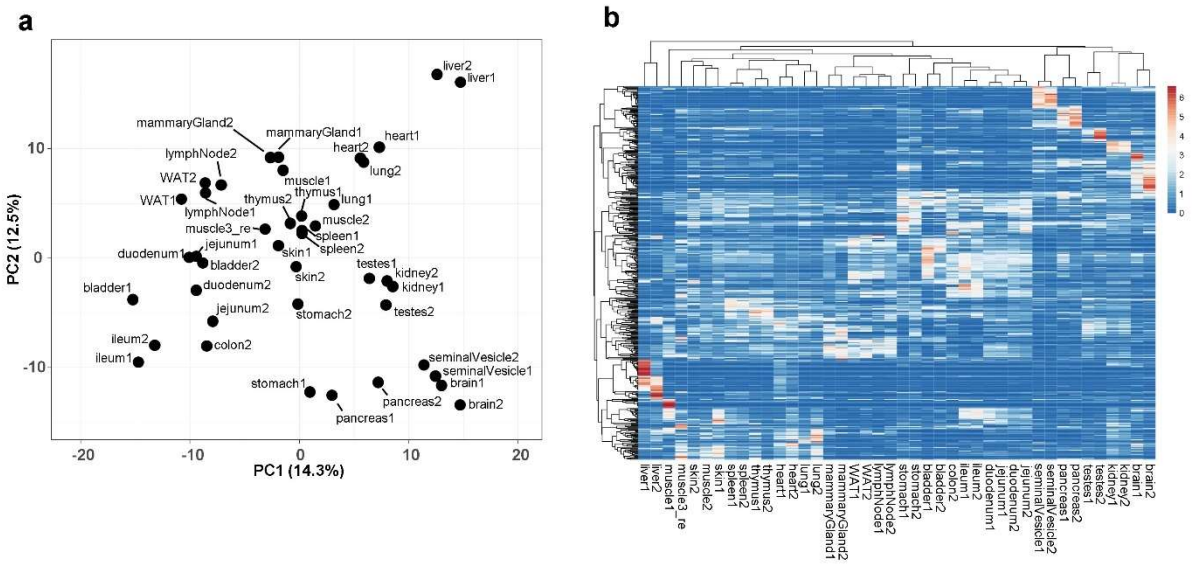

Figure 4 – Principal component analysis of LC-MS data. a) PCA-plot based on the intensity values of the 500 most intense masses extracted across all LC-MS datasets confirms good reproducibility of replicates and highlights clusters of functionally related, yet morphologically distinct mouse tissues. Unit variance scaling was applied to rows; SVD with imputation was used to calculate principal components. X and Y axis show principal component 1 and principal component 2 that explain 14.3% and 12.5% of the total variance, respectively. N = 42 data points. **b)** Heatmap. Unit variance scaling is applied to rows. Both rows and columns are clustered using correlation distance and average linkage. 500 rows, 42 columns.

Figure 5.

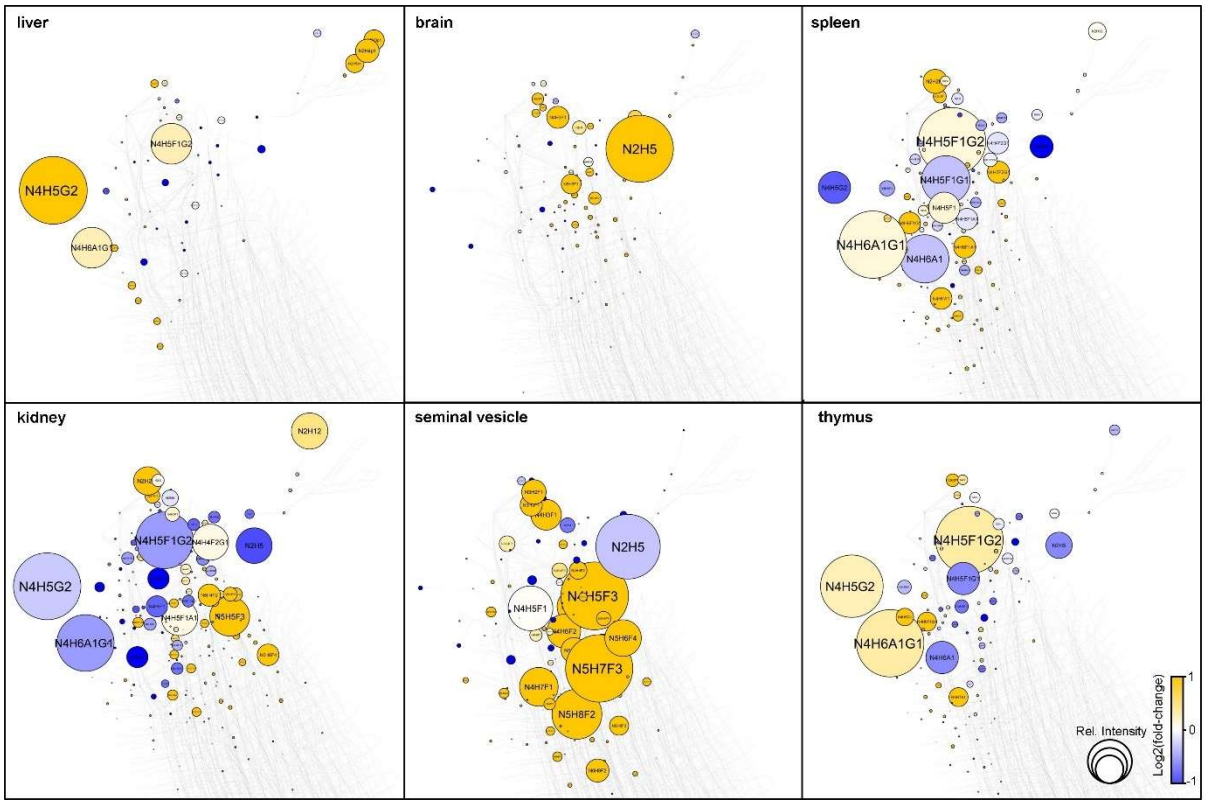

Figure 5 – Projections of glycan-mass intensity values onto a map of the mouse N-glycobiosynthetic pathways. Visualizations highlight substantial differences in N-glycan expression levels and tissue-dependant regulations of specific glycosylation-pathways for 4 tissues (liver, brain, kidney, seminal vesicle) and high similarity between spleen and thymus. Identical sections of the entire network are shown for all figure panels. Circle size indicates sample-specific relative abundance (in %) of each glycan composition. Color scale (orange – up to blue-down) indicates log2 fold-change in glycan composition expression over its mean across all samples. Custom N-glycobiosynthetic pathway maps were visualized using Cytoscape version 3.9.1.

Figure 6.

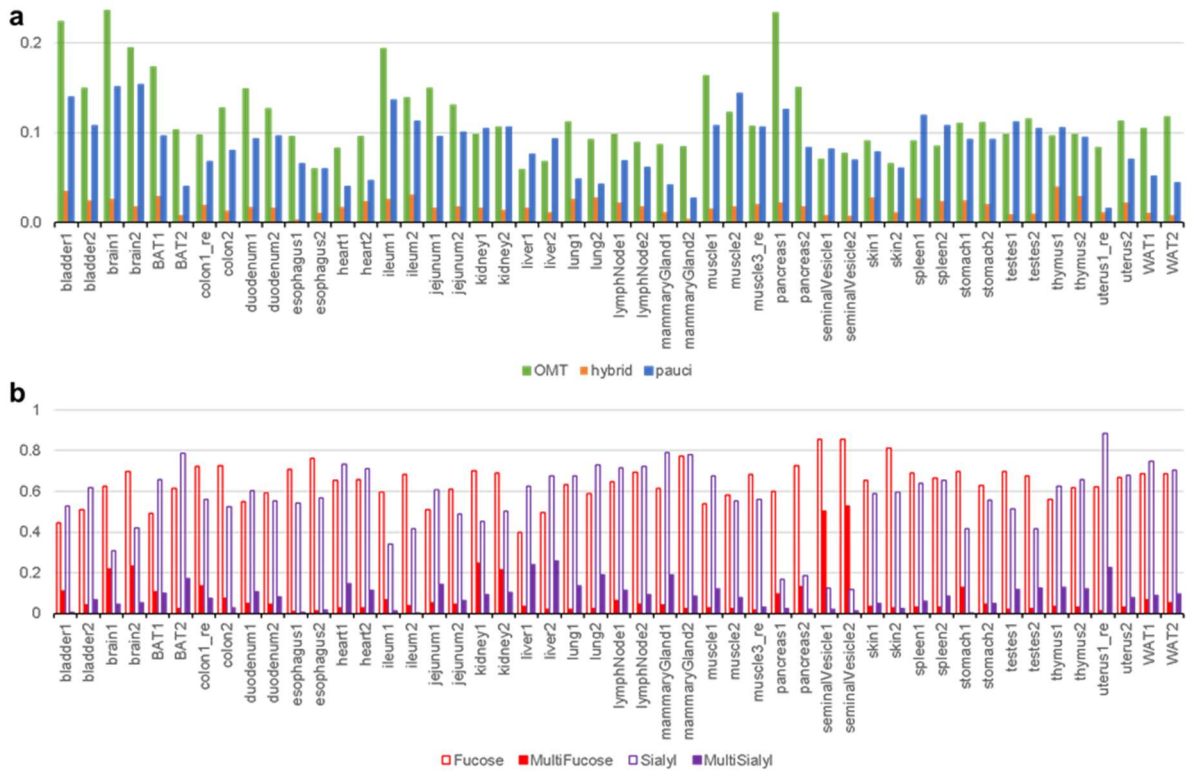

Figure 6 - Relative abundance of N-glycan structural features across all mouse tissues analysed. a) Precursor ion intensity-based, relative abundance of oligomannose type ("OMT") (i.e. Man5 – Man9+Glc3), pauci-mannose type ("pauci") and hybrid-type N-glycan structures. **b)** Precursor ion intensity-based, relative abundance of N-glycan structures with at least one fucose residue ("Fucose"), with two or more fucose-residues ("MultiFucose"), with at least one Neu5Ac or Neu5Gc residue ("Sialyl"), and with two or more Neu5Ac or Neu5Gc residues ("Multisialyl").N-glycan classification is based on N-glycan composition. Ambigous (isobaric) compositions were excluded from N-glycan classification.

Figure 7.

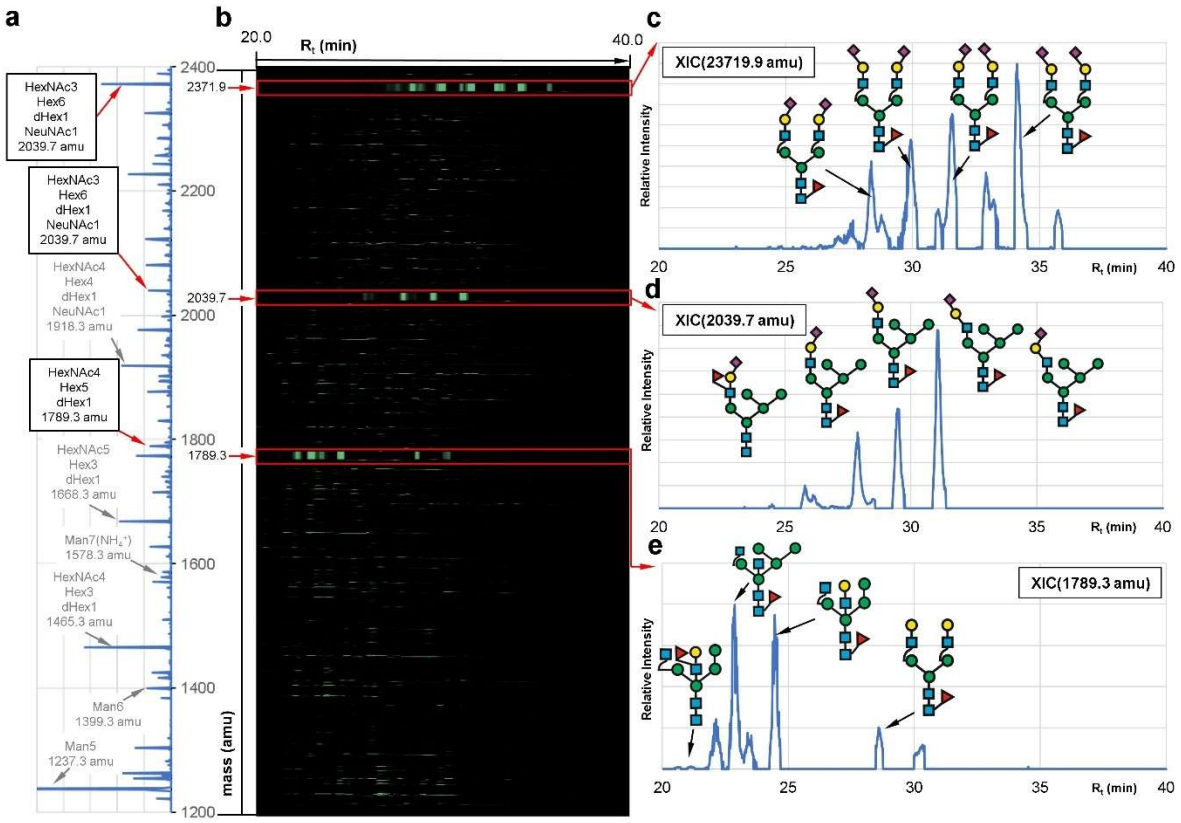

Figure 7 - Examplary graphical representation of the multidimensional PGC-LC-MS data set of the murine brain N-glycome. a) Precursor ion intensity histogram (molecular mass range from 1200 to 2400 amu; bin width = 1 amu), reconstructed from charge-deconvoluted and deisotoped MS1 data, summed over the chromatographic retention time-range from 20.0 to 40.0 mins. The molecular masses of high-intensity N-glycan compositions are indicated by arrows, with three representative N-glycan compositions indicated by red arrows. **b)** Heat-map of precursor ion inensities in the molecular mass range 1200 to 2400 amu (bin width = 1 amu) and the chromatographic retention time-range from 20.0 to 40.0 mins. Magnified heat-maps of the three representative N-glycan compositions (highlighted in panel a) are indicated by red rectangles. **c-e)** XICs of the three representative N-glycan compositions (highlighted in panel a), with arrows indicating the elution time points of the respective structural isomer (depicted as N-glycan cartoons) separted by PGC-LC. N-glycan cartoons were drawn according to the symbol nomenclature SNFG (https://www.ncbi.nlm.nih.gov/glycans/snfg.html).