The minimum information required for a glycomics experiment (MIRAGE): reporting guidelines for capillary electrophoresis

Data Reporting Examples

Supplementary Information, published in *Glycobiology* doi:10.1093/glycob/cwac021

Guinevere S.M. Lageveen-Kammeijer, Erdmann Rapp, Deborah Chang, Pauline M Rudd,

Carsten Kettner,* Joseph Zaia

*To whom correspondence should be addressed:
Tel: +49-(0)69-7167-3221; Fax: +49-(0)69 7167-3219; e-mail: ckettner@beilstein-institut.de

SUPPLEMENTARY INFORMATION INCLUDES:

SUPPLEMENTARY INFORMATION 1 - EXAMPLE OF CE GUIDELINES APPLIED FOR A GLYCOMICS STUDY (LIF)

1. General features	
Date stamp	2014-11-17
Responsible person/role	René Hennig & Erdmann Rapp Max Planck Institute for Dynamics of Complex Technical Systems Sandtorstrasse 1, 39106 Magdeburg, Germany and glyXera GmbH, Leipziger Straße 44, 39120 Magdeburg, Germany r.hennig@glyXera.com; rapp@mpi-magdeburg.mpg.de ORCID: (RH) 0000-0001-9172-7982; (ER) 0000-0001-6618-2626
Experiment type	xCGE-LIF (multiplexed capillary gel electrophoresis with laser induced fluorescence detection)
Experiment aim	Investigate the long-term stability of the human plasma *N*-glycome of individuals over a time period of up to six years to evaluate the potential of *N*-glycans as biomarkers for personalized diagnostics (Hennig, R., Cajic, S., et al. 2016).

2. Sample preparation

Samples were collected from five healthy volunteers (male, nonsmoking, same age: all 28 years old in 2011) within the time course of one year: for three months at weekly intervals, followed by three months at biweekly and six months at monthly intervals. Additional samples were taken within the time span of six years. Several environmental factors were documented like illness, diet, allergy and alcohol consumption (data collected for the period two days before sampling).

N-glycans were released from the plasma proteins as described previously (Ruhaak, L.R., Hennig, R., et al. 2010), with some slight modifications. Briefly, a total of 2 µL of plasma was mixed with 4 µL 2% SDS_{PBS} (*w/v*; 2 % sodium dodecyl sulfate (≥99%, A2572, AppliChem) and 98% PBS_{aq} (phosphate buffered saline; 10× concentrated, BioReagent, P5493-1L, Sigma-Aldrich) and incubated for 10 min at 60 °C in a 96-well polypropylene microplates (651201, Greiner Bio-One. After incubation, the remaining SDS was neutralized with the addition of 4 µL 8% $IGEPAL_{PBS}$ (*v/v;* 8% IGEPAL for molecular biology, I8896, Sigma-Aldrich with 92% PBS_{aq}). Subsequently, the *N*-glycans were released by adding 0.5 units of PNGase F (Peptide N-glycosidase F, BioReagent ≥95%, P7367, Sigma-Aldrich) in 1 µL PBS_{aq}, incubation was performed for 3 h at 37 °C. A total of 2 µL of released *N*-glycans were labeled with 20 mM APTS (aminopyrene-1,3,6-trisulfonic acid for fluorescence, ≥96.0%, 09341, Sigma-Aldrich) in 3.6 M citric $acid_{aq}$ (ACS grade for analysis, 1002440500, Merck-Millipore) and 2 µL 0.2 M 2-picoline-borane (≥95%, 654213, Sigma-Aldrich) in DMSO (dimethylsulfoxide for HPLC, ≥99.7%, 34869, Sigma-Aldrich). The labeling reaction was stopped by adding 100 µL of 80% $MeCN_{aq}$ (*v/v*; 80% acetonitrile, LC-MS grade > 99.5%, 34967, Sigma-Aldrich with 20% deionized water with R > 18.2 MΩ/cm (MQ) was produced by a Gradient A10 system from Merck-Millipore). After the MeCN addition, the samples were mixed carefully.

The free APTS, reducing agent and other contaminants were removed by hydrophilic interaction chromatography based solid phase extraction (HILIC-SPE) as described previously (Hennig, R., Rapp, E., et al. 2015). Briefly, a 100 mg/mL Bio-Gel P10 (150-4144, Bio-Rad, Germany) suspension dissolved in a MQ/ethanol (≥99.8%, T868.3, Carl Roth)/MeCN mixture (70:20:10%; *v/v/v*) was added to an AcroPrep™ 96-well GHP Filter Plates (5030, Pall Corporation). By applying vacuum with a vacuum manifold (Merck-Milipore) the solvents were removed. The wells were prewashed and equilibrated by × 200 µL MQ and 3 × 200 µL 80% $MeCN_{aq}$, respectively. The labelled samples were loaded onto the filterplate, to improve glycan binding the plate was shaken at 500 rpm for 5 min using a Thermomixer. Subsequently, the wells were washed with 5 × 200 µL 80% $MeCN_{aq}$ with 100 mM triethylamine (≥99.5%, 471283, Sigma-Aldrich)

which was adjusted to a pH of 8.5 using acetic acid (for luminescence, 45725, Sigma-Aldrich) and 3 × 200 µL $MeCN_{aq}$. All washing solutions were incubated for 2 min and removed by vacuum. To swell the BioGel, 1 × 100 µL MQ was added and the samples were eluted by applying 2 × 200 µL MQ to each well. Each MQ step was followed by an incubation of 5 min at 500 rpm on the Thermomixer. The eluates were collected by vacuum in 96-well storage plates (0.8 mL, AB-0756, Thermo Scientific). The eluates of the purified samples were pooled and stored at −20 °C (for max. 48 h) or analyzed directly by xCGE-LIF.

3. Equipment

The Genetic Analyzer model 3130xl was used (Applied Biosystems™, ThermoFischer Scientific) was equipped with a 50 cm 16-capillary array (4315930, Life Technologies) and filled with POP-7™ polymer (4363785, Life Technologies). The system is fully automated, including the loading and replacement of the polymer and can hold 96- or 384-well format microtiter plates. It is equipped with an Argon laser with a primary excitation at 488 nm and detection is performed > 520 nm.

4. Type of analysis

Type of cal./norm.	The xCGE-LIF generated data files (.ab1 / .fsa format) were directly loaded and processed by glyXtool™ (glyXera). The same software tool allowed migration time alignment (normalization) using the glyXalign internal standard (glyXera). Doing so, the electropherograms were transformed to "*N*-glycan fingerprints".
Analysis level	Qualitative and relative quantitative glycoprofiling of the total plasma *N*-glycome.

5. Run Processes

From the purified samples, a total of 1 µL was mixed with 9 µL Hi-Di™ Formamide (HiDi) (4311320, ThermoFischer Scientific). The complete mixture was transferred to a MicroAmp® Optical 384-well Reaction Plate (4309849, ThermoFischer Scientific) and sealed with a 384-well plate septa (4315934, ThermoFischer Scientific). Air bubbles at the bottom of the wells were removed by centrifugation (200 × g for 1 min) and the plate was transferred in to the xCGE-LIF platform. The samples were kept at room temperature, prior to being electrokinetically injected and analyzed with 1 × running buffer (10-times diluted Running Buffer, 402824, ThermoFischer Scientific), a voltage of +15 kV at a capillary temperature of 30-60 °C. Data was collected for 40 min. It is equipped with an Argon laser with a primary excitation at 488 nm and detection is performed > 520 nm. Between each analysis the polymer was automatically replaced.

6. Detection

Laser induced fluorescence (LIF) detection was performed with an Argon laser. The excitation wavelength was 488 nm and wavelength resolved detection was performed > 520 nm. Detector calibration has been carried out using the glyXcal-λ standard (glyXera).

7. Electropherogram/-chromatogram, data processing.

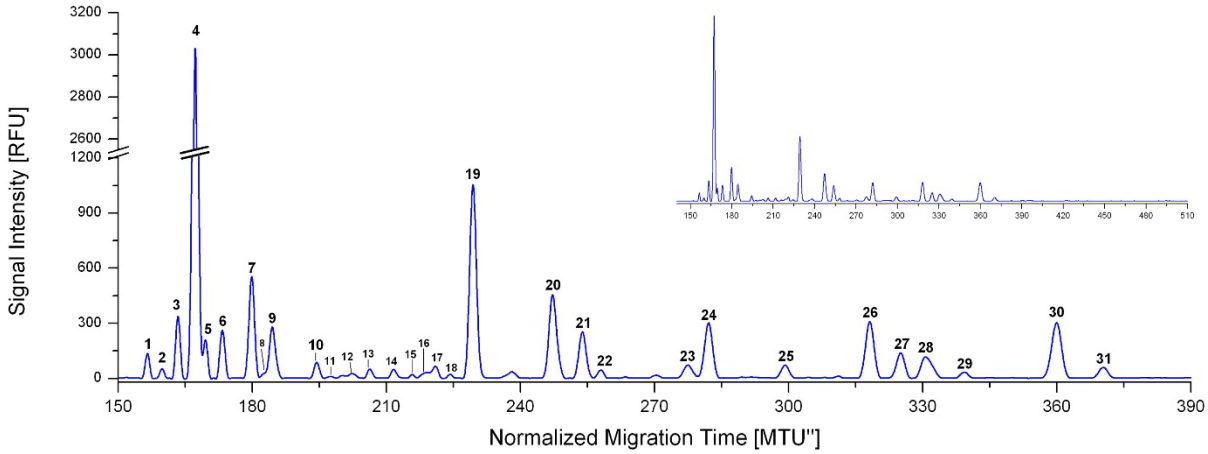
The xCGE-LIF raw data files were directly loaded and processed by glyXtool™ (glyXera) (Hennig, R., Reichl, U., et al. 2011). The same software tool aligned (normalized) the data using the glyXalign™ standard (glyXera), transforming the electropherograms into "*N*-glycan fingerprints", and performed automated data smoothing, background adjustment as well as peak picking (s/n > 3). Using glyXtool, automated

structural assignment of *N*-glycan peaks was performed (**Figure S-1** and **Table S-2**), based on aligned (normalized) migration time matching to the integrated *N*-glycan database glyXbase™ (glyXera). Eventually, automated relative quantification (s/n > 10) was performed automated by glyXtool.

Extended glycoprofiling was performed via exoglycosidase sequencing:

Materials	Exoglycosidases α(2-3) sialidase (Sialidase S, GK80020), α(2-3,6,8) sialidase (Sialidase A, GK80040), α(1-2,3,4,6) fucosidase (GKX-5006) and α(1-2,3,6) mannosidase (GKX-5010) were purchased from Prozyme. Exoglycosidases β(1-4) galactosidase (P0730L) and β(1-2,3,4,6)-*N*-Acetylglucosaminidase (P0732L) were purchased from New England Biolabs GmbH. α(1-3,4) fucosidase (E-F134) was purchased from QA-Bio.
Exoglycosidase Preparation	All exoglycosidase digestions were performed under the recommended reaction conditions from the suppliers. For all experiments, 7-8 µL of the purified APTS-labeled *N*-glycans derived from plasma was used.
Reaction Time	Reaction times for the different exoglycosidase digestions were performed according to the times recommended by the respective suppliers.
Control	Activities and specificities of all exoglycosidases were carefully tested by incubation with APTS labelled N-glycans derived from bovine fetuin (for SiaA, SiaS), bovine IgG (for aFUCase, b14GALase and bNAcGLUase), bovine ribonuclease B (for aMANase) and human lactotransferrin (for a34FUCase). Samples were accordingly treated as the controls.
Protocol	See the protocols available from the suppliers.

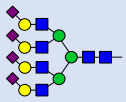
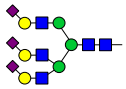
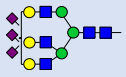
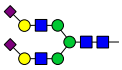
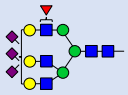
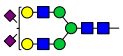
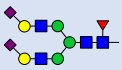
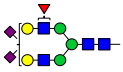
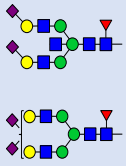
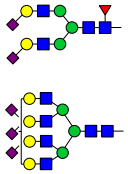
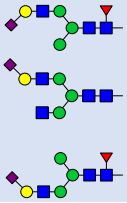
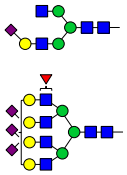
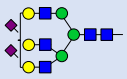
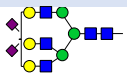
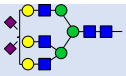
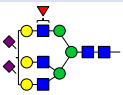
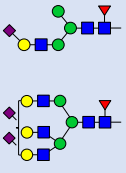
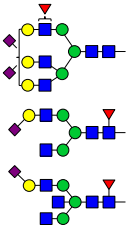
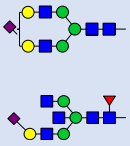
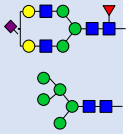
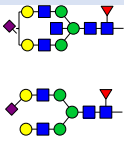
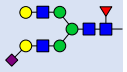
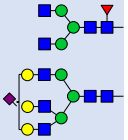
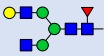
Supplementary Figure S-1

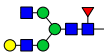
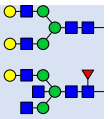
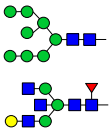
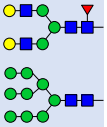
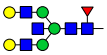


Supplementary Figure S-1. An example of a human plasma *N*-glycan fingerprint of total plasma *N*-glycome generated by xCGE-LIF, derived from a native (sialylated) frozen normal control plasma. Released *N*-glycans were labelled with APTS. The 31 most abundant *N*-glycan peaks are marked with numbers , for more details of the numbered peaks see **Supplementary Table S-1**. The inset illustrates the entire migration time range of the xCGE-LIF measurement. Figure adapted from (Hennig, R., Cajic, S., et al. 2016).

Supplementary Table S-1

The 31 most abundant peaks of the *N*-glycan fingerprint of total plasma *N*-glycome (illustrated in **Supplementary Figure 1)** were assigned to their corresponding *N*-glycan structures. Blue square: *N*-acetyl glucosamine, yellow circle: galactose "G", green circle: mannose, red triangle: fucose "F" and purple diamond: *N*-acetyl neuraminic acid "S" ([left pointing: α2,3-linkage and right pointing α2,6-linkage). The number after "A" indicates the number of antennae (occupied with at least one *N*-acetyl glucosamine), "B" indicates bisection, and "F" in front of the "A" indicates core-fucosylation. Numbers in rectangular brackets after "A" indicate that either the 3- or 6-arm is being occupied ([3] or [6], respectively), after "F" indicate a 1,3-linkage ([3] = antennae fucosylation), and behind "S" indicate the sialic acid linkage ([3] = α2,3-linkage and [6] = α2,6-linkage). Curly brackets indicate that it is not defined to which arm the sialic acid(s), respectively, the antenna fucose is bound. Table adapted from (Hennig, R., Cajic, S., et al. 2016).

#	*N*-glycan	Composition
1	A4G4S4[6,6,6,6]	
2	A3G3S3[6,6,6]	
3	A3G3S3[3,6,6]	
4	A2G2S2[6,6]	
5	A3F[3]G3S3[3,6,6]	
6	A2G2S2[3,6]	
7	FA2G2S2[6,6]	
8	A2F[3]G2S2[3,6]	
9	FA2BG2S2[6,6] FA2G2S2[3,6]	
10	FA2G2S2[3,3] A4G4S3[3,6,6]	
11	FA1[3]G1S1[6] FA1[6]G1S1[3] A2[6]G1S1[6]	
12	A2[3]G1S1[6] A4F[3]G4S3[3,6,6]	
13	A3G3S2[6,6]	
14	A3G3S2[3,6]	
15	A3G3S2[3,6]	
16	A3F[3]G3S2[6,6]	
17	FA2[3]G1S1[6] FA3G3S2[6,6]	
18	A3F[3]G3S2[3,6] FA2[6]G1S1[3] FA2[6]BG1S1[6]	
19	A2G2S1[6] FA2[3]BG1S1[6]	
20	FA2G2S1[6] Man5	
21	FA2BG2S1[6] FA2[6]G2S1[3]	
22	FA2[3]G2S1[3]	
23	Man6	
24	FA2G0 A3G3S1[6]	
25	FA2BG0	
26	FA2[6]G1	

27	FA2[3]G1	
28	A2G2 FA2[6]BG1	
29	Man8 FA2[3]BG1	
30	FA2G2 Man9	
31	FA2BG2	

SUPPLEMENTARY INFORMATION 2 - EXAMPLE OF CE GUIDELINES APPLIED FOR A GLYCOMICS STUDY (MS)

1. General features	
Date stamp	2017-08-28
Responsible person/role	Guinevere Lageveen-Kammeijer Leiden University Medical Center, Center for Proteomics and Metabolomics, Postbus 9600, 2300 RC, Leiden, The Netherlands g.s.m.kammeijer@lumc.nl ORCID: 0000-0001-7670-1151
Experiment type	Online t-ITP-CZE-ESI-MS (transient-isotachophoresis - capillary zone electrophoresis -electrospray ionization - mass spectrometry)
Experiment aim	Define the repeatability (intraday variability) of PNGaseF released *N*-glycans from pooled human plasma, followed by a linkage-specific derivatization of sialic acids and labelling procedure with Girard's reagent P by sheathless CE-ESI-MS(Lageveen-Kammeijer, G.S.M., de Haan, N., et al. 2019).

2. Sample preparation

Sample preparation was performed as described previously(Lageveen-Kammeijer, G.S.M., de Haan, N., et al. 2019). Briefly, *N*-glycans were released from all plasma proteins (TPNG) by mixing 100 µL of pooled human plasma (Visucon-F; Affinity Biologicals, Ancaster, ON, Canada) with 200 µL of 2% sodium dodecyl sulfate (Sigma Aldrich; cat. nr. L3771) followed by a 10 min incubation at 60°C. Subsequently, a 200 µL mixture of 2.5×PBS, 2% Nonidet P-40 substitute (Sigma Aldrich; cat. nr. M158 and 10 U Recombinant peptide-N-glycosidase F (Roche Diagnostics, Mannheim, Germany; cat. Nr. 11365177001) was added and the sample was incubated at 37 °C for 17 h. To obtain sialic acid linkage information, the samples were derivatized by ethyl esterification amidation. Only 1 µL of released plasma *N*-glycans (containing the released glycans from 0.2 µL of plasma) was added to 20 µL ethyl esterification reagent (250 mM 1-Ethyl-3-(3-(dimethylamino)propyl) carbodiimide (Fluorochem; cat. nr. 24810) and 250 mM 1-hydroxybenzotriazole (Sigma Aldrich; cat. nr. 54802) in ethanol (Merck; Darmstadt, Germany; cat. nr. 1.00983.1000) and incubated for 1 h at 37 °C. Additionally, 4 µL of 28% ammonium hydroxide (Sigma Aldrich; cat. nr. 221228) was added to the reaction mixture and incubated for 2 h at 37 °C. After incubation, 24 µL HPLC SupraGradient acetonitrile (MeCN; Biosolve, Valkenswaard, Netherlands; cat. nr. 1203502) was added to the mixture and the *N*-glycans were purified by cotton HILIC SPE. The cotton was first washed 5 times with 10 µL ultrapure deionized water from a Q-Gard 2 system (Millipore, Amsterdam, Netherlands; maintained at ≥18 MΩ), followed by 3 times of 10 µL 90% MeCN to condition, the solutions were aspired and dispended. Afterwards the released glycans were adsorbed on to the cotton by pipetting up and down for 20 times. Followed by a wash by applying 3x 90% MeCN (10 µL). Elution was performed in 10 µL MQ. After sialic acid derivatization, released *N*-glycans from plasma were labelled at the reducing end by 1-(Hydrazinocarbonylmethyl)pyridinium chloride (GirP; TCI Development Co. Ltd, Tokyo, Japan; cat. nr. G0030). Five µL of the sialic acid-derivatized, HILIC-purified *N*-glycans were mixed with 20 µL GirP reagent (50 mM GirP in 90% EtOH, 10% glacial acetic acid (HAc; Merck; cat. nr. 1000631000) and incubated for 2 h at 60 °C. After incubation, the samples were dried by vacuum concentration at 60 °C and dissolved in 10 µL MQ for CE-ESI-MS analysis.

3. Equipment
An in-house Ultratrol coated Silica Surface OptiMS cartridge (SCIEX, Framingham, MA; cat. nr. B07367, sn70171916) was placed in a CESI 8000 system (SCIEX), allowing for a single capillary setup. With an effective and total length of 91 cm and an internal and outer diameter of 30 µm and 150 µm, respectively. The CESI 8000 system was operated by 32 Karat Software (v10.1, build 108, Beckman Coulter, Brea, CA). Prior to usage, the capillary was conditioned by emerging the spray tip in methanol (MeOH) while consecutively the separation and conductive lines were rinsed with MeOH at 100 psi for 10 min and 3 min, respectively. Subsequently, the tip was immersed in LC-MS Ultra water (H_2O, Honeywell, Morris Plains, NJ; cat. nr. 00232141B1BS) and the separation line was rinsed at 100 psi for 10 min consecutively with H_2O, 0.1 M sodium hydroxide (NaOH; Sigma Aldrich; cat. nr. 71686), 0.1 M hydrochloride (HCl), H_2O and the background electrolyte (BGE; 10% HAc), followed by a final rinsing step of the conductive line for 3 min with the BGE at 100 psi. After conditioning the capillary was in-house coated with Ultratrol™ dynamic pre-coat LN (UT; Target Discovery, Palo Alto, CA) as described by Kohler et al.(Kohler, I., Augsburger, M., et al. 2014) with some small adjustments. Briefly, using 29 psi throughout, the separation line was coated by rinsing consecutively for 10 min with MeOH, H_2O, 0.1 M HCl and H_2O, 15 min rinsing with 1 M NaOH followed by 15 min rinsing with H_2O, finally a 30 min rinsing was performed with UT. After the coating procedure, the capillary was rinsed for 30 min with the BGE at 29 psi. All steps were performed with the capillary coolant set at 20 °C.

4. Type of analysis	
Type of cal./norm.	Raw CE-ESI-MS data were calibrated prior to data analysis using a minimum of five signals of the identified *N*-glycan compositions (**Supplementary Table 2**) with Data Analysis 4.2 (Build 395, Bruker Daltonics). After converting the raw files into .mzXml files, targeted data analysis was performed using an adapted version of LaCyTools (v1.0.1, build 8)(Jansen, B.C., Falck, D., et al. 2016). Prior to automated peak integration, all electropherograms were aligned based on 11 glycan peaks that were confirmed by tandem MS and that covered the complete migration range (34.5–39.1 min; **Supplementary Table 2**). An alignment time window of 50 s and a *m/z* window of 0.02 Th were used.
Analysis level	Glycopeptide profiling. The data were manually screened for *N*-glycan compositions based on their exact mass, migration order and previous described structures in literature. *N*-glycan compositions were included for further data analysis, when, in at least two repeated experiments per condition, their mass accuracy was between ±10 ppm, their isotopic pattern did not deviate more than 20% from the theoretical isotopic pattern and their S/N was above 9.

5. Run Processes
All samples (TPNG-EEA-1; TPNG-EEA-2; TPNG-EEA-3) were reconstituted in 10 uL H_2O. In total, 1 µL of 400 mM ammonium acetate (Sigma Aldrich, Steinheim, Germany; cat. nr. A2706) at pH 4.0 and 7.5 µL of H_2O was added to 1.5 uL sample, the mixture was then transferred to a CESI NanoVial (SCIEX, 5043467) and placed in a temperature controlled sample tray (10 °C). Prior to each analysis the capillary was washed with 1 M NaOH (3 min), H2O (4 min), UT (4 min) and with BGE (3 min) all at 29 psi. To ensure the removal of the UT from the outlet of the capillary, an additional 6 min wash with BGE was performed

at 100 psi. Followed by washing the conductive line with BGE for 3 min at 29 psi. Samples were hydrodynamically injected at the long end using 5 psi for 60 sec (~6.8% of the total capillary volume (~43 nL)). After each sample injection, a BGE post plug was injected by applying 0.5 psi for 25 s (~0.3% of the total capillary volume). For each analysis a constant flow was established by applying 0.5 psi and 20 kV over the capillary with 1 min ramping. After 70 min the voltage was ramped down to 1 kV in a time span of 5 min. All steps were performed with the capillary coolant set at 20 °C. For system suitability the following parameters were being monitored and collected: current (~2.7 uA), voltage (+30 kV) and pressure (0.5 psi). Data collection from the detector (MS) was collected for 60 min.

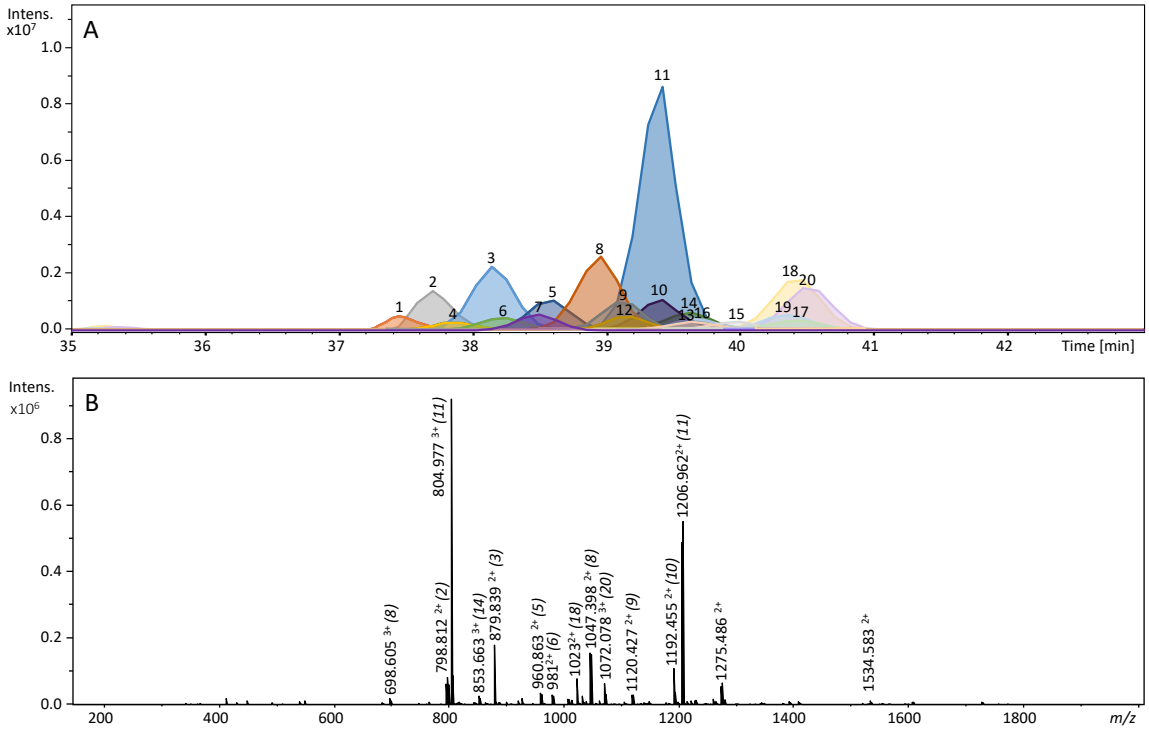
6. Detection

After separation analytes were detected with an Impact HD UHR-QqTOF-MS (Bruker Daltonics, Bremen, Germany) equipped with a nanospray shield (Bruker Daltonics) via a CESI OptiMS Bruker MS adapter kit (SCIEX). Experiments were performed in positive-ionization mode and a stable electrospray was obtained by generating an electrical field between the CE (ground potential) and a negatively charged spray shield (between −1100 and −1300 V). For all analysis, the temperature and flow rate of the drying gas was set at 100°C and 1.2 L/min, respectively. To minimize the in-source decay, the collision cell energy as well as the quadrupole ion energy were set at 3.0 eV and the pre-pulse storage was set at 15.0 µs. To enhance the detection of the glycans species the NanoBooster (Bruker Daltonics) was used, using MeCN as dopant (ca. 4% mole percent) at 0.2 bar. An inhouse made polymer cone enabled the usage of the NanoBooster in combination with the CESI OptiMS Bruker MS adapter(Kammeijer, G.S., Kohler, I., et al. 2016). Fragmentation was performed at 1.00 Hz on the three most abundant precursor ions in a range of *m/z* 150–2000 with a minimum intensity of 4548. Depending on the *m/z* values, the precursor ions were isolated with a width of 8–10 Th. The collision energies were set as a linear curve in a m/z dependent manner, ranging from 55 eV at *m/z* 700 to 124 eV at *m/z* 1800 for all charge states (1–5), applying a basic stepping mode with collision energies of 100% (80% of the time) or 50% (20% of the time).

7. Electropherogram/-chromatogram, data processing.

Raw CE-ESI-MS data were calibrated prior to data analysis using a minimum of five signals of the identified *N*-glycan compositions (**Supplementary Table 2**) with Data Analysis 4.2 (Build 395, Bruker Daltonics). The data were manually screened for *N*-glycan compositions based on their exact mass, migration order and previous described structures in literature. After converting the raw files into .mzXml files, targeted data analysis was performed using an adapted version of LaCyTools (v1.0.1, build 8)(Jansen, B.C., Falck, D., et al. 2016). Prior to automated peak integration, all electropherograms were aligned based on 11 glycan peaks that were confirmed by tandem MS and that covered the complete migration range (34.5–39.1 min; **Supplementary Table 2**). An alignment time window of 50 s and a *m/z* window of 0.02 Th was used. For each glycan composition, the area was integrated of isotopic peaks covering at least 95% of the theoretical isotopic pattern and the background was subtracted based on the local background calculations. *N*-glycan compositions were included for further data analysis, when, in at least two repeated experiments per condition, their mass accuracy was between ±10 ppm, their isotopic pattern did not deviate more than 20% from the theoretical isotopic pattern and their S/N was above 9.

Supplementary Figure S-1



Supplementary Figure S-2. Analysis of derivatized and labelled *N*-glycans released from a pooled human plasma sample. **A**: Extracted ion electropherograms (EIEs) showing the separation by CE-ESI-MS of the 20 most abundant *N*-glycans. **B**: Summed MS spectrum (37.2-40.8 min) of the detected. Peak numbers above EIEs and between brackets are annotated and described in **Supplementary Information, Table S-2**. Separation was achieved using a bare-fused silica capillary after a dynamic neutral coating.

Supplementary Table S-2

Analysis of the 20 major *N*-glycans released from pooled human plasma using a CE-ESI-MS/MS workflow. For the confirmation and quantitation of each glycan the migration times, mass accuracy as well as MS2 spectra are considered. The peak numbers correspond to the assignments shown in **Supplementary Figure S-2**. GlyTouCan accession identifiers are provided for each registered glycan entry to improve database searching. The raw mass spectrometric data files that support the findings of this study are available in MassIVE in .mzXML and .xy format, with the identifier MSV000083478 [https://doi.org/10.25345/C5061Z](Lageveen-Kammeijer, G.S.M., Haan, d.N., et al. 2019). Blue square: *N*-acetyl glucosamine, yellow circle: galactose, green circle: mannose, red triangle: fucose and purple diamond: *N*-acetyl neuraminic acid (left pointing: α2,3-linkage and right pointing α2,6-linkage).

	m/z value used for calibration	MS2 confirmed (either concentrated sample or normal concentration)		$[M]^+$				$[M+H]^{2+}$				$[M+2H]^{3+}$				Relative % area (n=3)		
	Composition	GlyTouCanID	Migration time (min)	*m/z* (calc)	*m/z* (obs)	*m/z* error	ppm error	*m/z* (calc)	*m/z* (obs)	*m/z* error	ppm error	*m/z* (calc)	*m/z* (obs)	*m/z* error	ppm error	Mean	SD	MS2
1		G10609HO	37.5	**1530.558**	1530.565	0.007	4.645	**765.783**	765.785	0.003	3.581	**510.858**				2.48%	0.28%	x
2		G83355KE	37.7	**1596.616**	1596.619	0.002	1.495	**798.812**	798.816	0.004	5.542	**532.877**				5.34%	1.03%	x
3		G17389UM	38.1	**1758.669**	1758.672	0.003	1.653	**879.838**	879.842	0.004	4.672	**586.895**				7.98%	1.11%	x
4		G75698JE	37.9	**1799.696**	1799.690	-0.005	-2.868	**900.351**	900.358	0.006	6.713	**600.570**				1.12%	0.19%	x
5		G36836GD	38.5	**1920.722**	1920.723	0.001	0.401	**960.865**	960.868	0.004	4.044	**640.912**				3.57%	0.23%	x
6		G23280RB	38.3	**1961.748**	1961.751	0.003	1.360	**981.378**	981.380	0.002	2.332	**654.588**				1.66%	0.09%	x
7		G39617JJ	38.5	**2016.717**				**1008.862**	1008.864	0.002	2.060	**672.910**				1.91%	0.18%	x
8		G18609BT	38.7	**2093.791**				**1047.399**	1047.402	0.003	3.025	**698.602**	698.602	0.000	0.127	8.83%	0.88%	x
9		G81413UE	39.1	**2239.849**				**1120.428**	1120.434	0.006	5.445	**747.288**	747.286	-0.001	-1.998	3.60%	0.23%	x
10		G80123ZU	39.4	**2383.902**				**1192.455**	1192.460	0.005	4.576	**795.306**	795.306	0.000	0.312	4.90%	0.24%	x
11		G58235TZ	39.4	**2412.917**				**1206.962**	1206.967	0.005	3.954	**804.977**	804.981	0.004	4.448	29.81%	1.34%	x
12		G86632ZU	39.1	**2442.928**				**1221.968**	1221.975	0.008	6.270	**814.981**	814.989	0.008	9.987	1.95%	0.14%	x
13		G98557ZA	39.6	**2529.960**				**1265.484**	1265.482	-0.001	-0.927	**843.992**	843.990	-0.002	-2.399	1.45%	0.12%	
14		G56749GV	39.6	**2558.975**				**1279.991**	1279.999	0.007	5.663	**853.663**	853.661	-0.002	-2.586	3.30%	0.32%	x
15		G61151LQ	39.6	**2749.034**				**1375.021**	1375.027	0.006	4.447	**917.016**	917.014	-0.002	-2.518	1.31%	0.04%	x
16		G14127XU	39.7	**2762.055**				**1381.531**	1381.534	0.003	2.058	**921.356**	921.358	0.001	1.193	1.75%	0.15%	x
17		G03706EO	40.4	**3039.146**				**1520.077**	1520.082	0.006	3.872	**1013.720**	1013.722	0.002	1.609	1.71%	0.19%	x
18		G77725ZE	40.4	**3068.161**				**1534.584**	1534.582	-0.002	-1.259	**1023.392**	1023.395	0.004	3.545	8.32%	0.83%	x
19		G18188WJ	40.3	**3097.176**				**1549.092**	1549.098	0.006	3.809	**1033.064**	1033.065	0.001	0.920	2.65%	0.19%	x
20		G95997LR	40.5	**3214.219**				**1607.613**	1607.618	0.005	3.057	**1072.078**	1072.082	0.004	4.143	6.38%	0.88%	x

SUPPLEMENTARY INFORMATION 3 – EXAMPLE OF CE GUIDELINES APPLIED FOR A GLYCOPROTEOMICS STUDY

Title: CE-MS analysis of glycopeptides	
1. General features	
Date stamp	2018-03-09
Responsible person/role	Joseph Zaia Boston University School of Medicine, Center for Biomedical Mass Spectrometry, 670 Albany St. Fifth floor, Boston, MA 02118 jzaia@bu.edu ORCID: 0000-0001-9497-8701
Experiment type	tITP-CZE-ESI-MS/MS (transient-isotachophoresis - capillary zone electrophoresis - electrospray ionization – tandem mass spectrometry)
Experiment aim	Analysis of glycoprotein proteolytic digests using CE-MS(Khatri, K., Klein, J.A., et al. 2017)
2. Sample	
Sample name(s) and descriptions	Alpha-1-acid glycoprotein (AGP), purified from human plasma (Sigma-Aldrich cat. # G9885)
Sample solution	Sample matrix and background electrolyte (BGE) were from the 908 Devices peptides assay kit (908 Devices, part # 810-00034), used without alterations.
Sample preparation	The methods used for sample preparation were described previously(Chang, D., Hackett, W.E., et al. 2020, Khatri, K., Klein, J.A., et al. 2017). Viral samples were first sonicated in methanol for 20 minutes to disrupt viral membranes, and then dried in a SpeedVac. All glycoprotein samples were denatured and reduced in 100 mM ammonium bicarbonate (J.T. Baker, cat. # JT-3003-01), 2-2-2 trifluoroethanol (TFE; Sigma-Aldrich, cat. # T63002), and 200 mM dithiothreitol (DTT; Sigma-Aldrich, cat # D5545) for one hour at 65°C. Cysteine residues were alkylated in 200 mM iodoacetamide (Bio-Rad, cat. # #1632109) for 1 h at room temperature in the dark; excess iodoacetamide was quenched with the addition of 200 mM DTT, incubating for 1 h at room temperature in the dark. TFE in the solution was diluted to 5% by adding 3:1 mixture of water:100 mM ammonium bicarbonate. Sequencing grade trypsin or chymotrypsin (Promega, cat. # V5111, # V1061) was added at an enzyme:substrate ratio of 1:20, and mixtures were incubated overnight at 37 °C. After digestion, samples were evaporated to dryness. Digested AGP samples were incubated with 200 units of alpha2-3,6,8 neuraminidase (New England Biolabs, cat. # P0720) at 37 °C overnight to remove all terminal sialic acid residues. All samples were cleaned using PepClean C18 spin columns (Pierce Biotechnology, cat. # 89873). Immediately prior to CE-MS analysis, the samples were resuspended to a concentration of 1 µg/µL in ZipChip Sample Diluent.

3. Equipment	
3.1 Instrumentation and other equipment	
Manufacturer, model, catalog number	Thermo-Fisher, Q Exactive 908 Devices, ZipChip CE-ESI interface 908 Devices, ZipChip HR chip (part # 810-00140)
Instrument details	Glass microfluidics interface with etched capillary channels and integrated nano-ESI.
3.2 Control and data collection software	
Manufacturer, name, version	Thermo-Fisher XCalibur version 3.1.66.10 ZipChip standalone software version 1.0
3.3 Capillary	
	908 Devices ZipChip HR (part # 810-00140, lot # HR 002719) is a microfluidic device with a fully integrated capillary and nano-ESI. The capillary length is 22 cm. The ZipChip HR was used without prior conditioning.
4. Type of analysis	
Type of cal./norm.	• No calibration/normalization was used
Analysis level	• Glycopeptide glycoform compositions were assigned based on exact mass and MS/MS spectra
5. Run Processes	
5.1 Run descriptors	
Temperature control of the capillary was not used. No auxiliary data channels were set up. Data collection duration was 15 min for each replicate.	
5.2 Step descriptors (optional)	
Not applicable	
5.3 Sample injection.	
Prior to injection of each sample type, the microfluidics ZipChip was primed with ZipChip Background Electrolyte (BGE) according manufacturer's instructions without modification. Immediately prior to each individual injection, each sample (AGP and Phil82) was resuspended to a concentration of 1 µg/µL in ZipChip Sample Diluent. Sample injection was done manually, one sample at a time, without any temperature control. Three µL of resuspended sample was placed into the small depression at the bottom of the sample well. Using the ZipChip standalone software, a method was run with a load time of 30 s, pressure assist disabled, and analysis run time of 0.1 s. Following this step, 20 µL of sample was placed into the sample well. The final step was to run another method on the ZipChip software with a load time of 0 s, pressure assist enabled, and analysis run time of 15 min. ZipChip's recommended method was to resuspend the desired amount of sample with 20 µl of Sample Diluent and to pipette all 20 µL into the sample well at once, multiple sample	

injections should be able from this solution. However, we found that the sample well would become diluted with BGE, and subsequent injections would result in a loss of peak intensity.

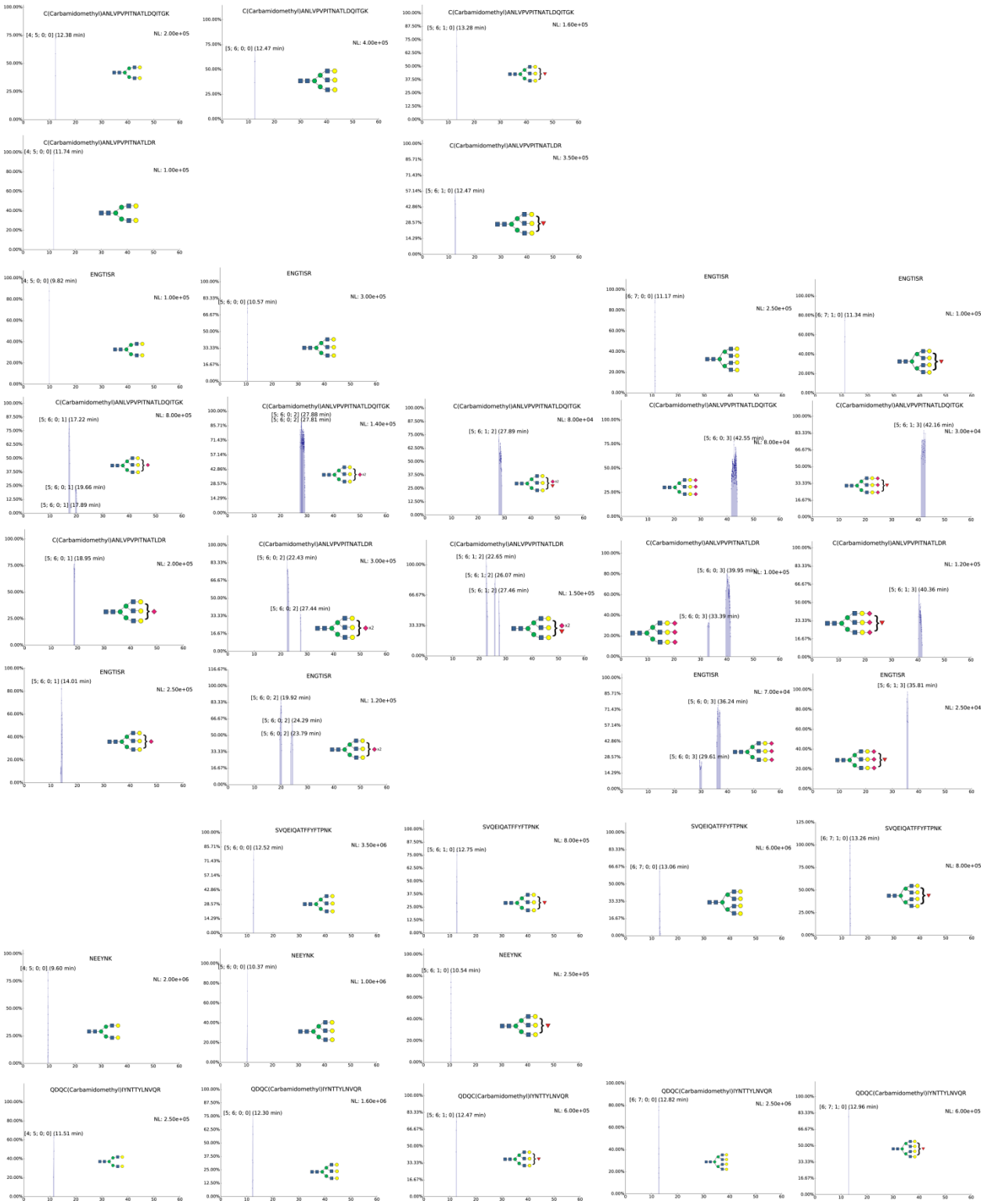
6. Detection

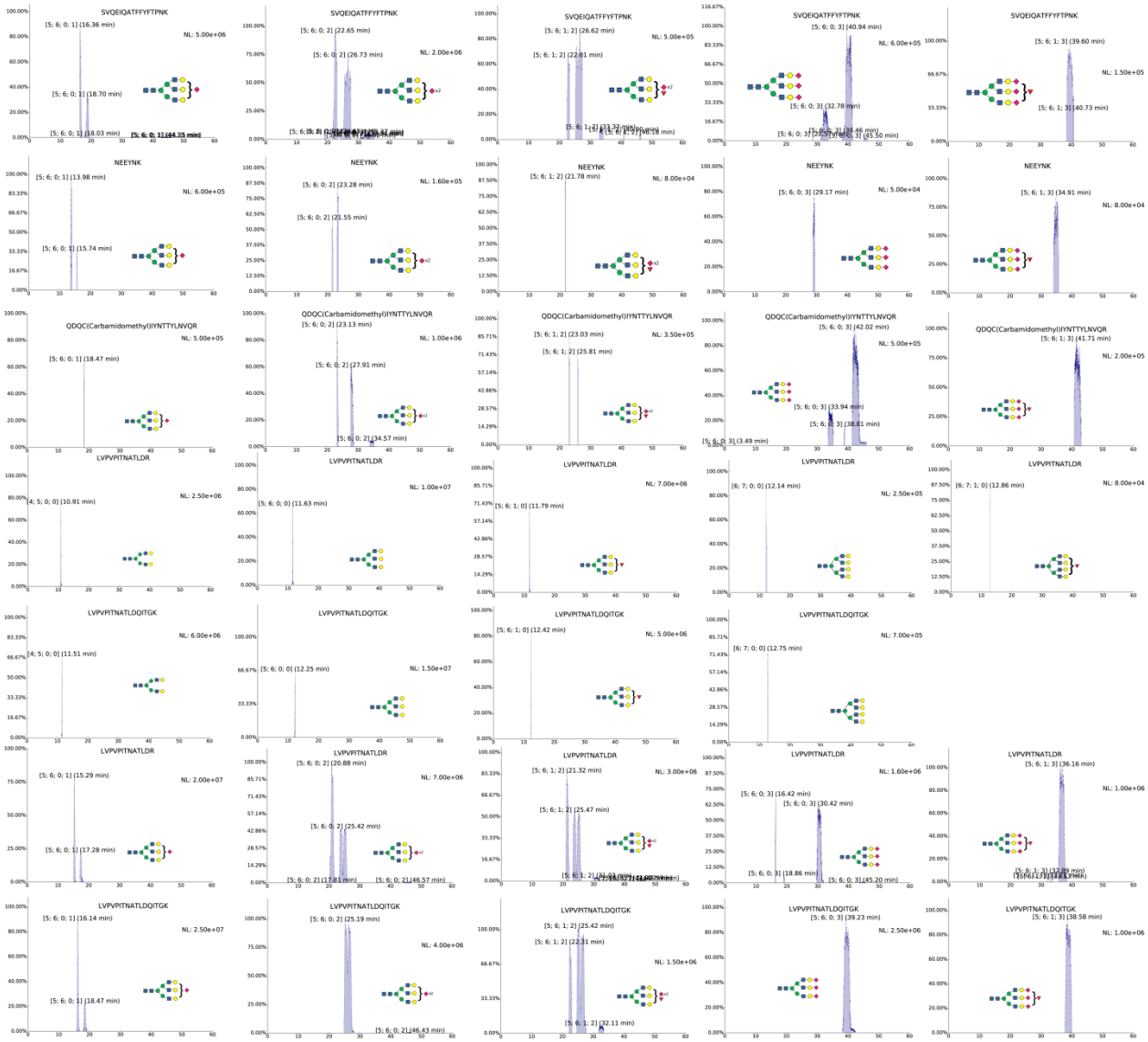
After analyte separation and ionization in the integrated ESI component of the microfluidics chip, detection was performed using a Q Exactive mass spectrometer (Thermo-Fisher, San Jose, CA) in positive ion mode, using data-dependent acquisition MS/MS. The acquisition time was 15 min. MS1 spectra were acquired at 70,000 resolution at *m/z* 400, scan range *m/z* 370-3000, 1 microscan per spectrum, AGC target 3e6, maximum injection time 100 ms. MS2 spectra were acquired at 17,500 resolution at *m/z* 400, 2 microscans per spectrum, AGC target $1e^6$, isolation window *m/z* 4.0, isolation offset *m/z* 0.4, scan range *m/z* 200-2000, fixed first mass of *m/z* 200, exclusion of unassigned charge states and charge state 1, dynamic exclusion 4.0 s, and stepped NCE of 18 and 28. The maximum injection time and intensity thresholds were for some replicates between 100 and 200 ms, and 5e5 and 2e5, respectively. The four most abundant precursor ions per scan were fragmented. Profile data were recorded for both MS1 and MS2 scans.

7. Electropherogram/-chromatogram, data processing.

Electropherograms were visualized using Thermo-Fisher Xcalibur Qual Browser version 2.2 SP1.48, a different version than the one used for acquisition above. Raw files were converted to mzML format using MSConvert from ProteoWizard version 3.0.11252 with no filters specified. Glycopeptide glycoform compositions were identified using the GlycReSoft search engine, a complete open-source software package that assigns glycopeptides from tandem mass spectra (https://www.bumc.bu.edu/msr/software/). The input required by GlycReSoft were a peptide hypothesis in mzIdentML format and a glycan hypothesis. The peptide hypothesis was generated from previously acquired LC-MS data. To obtain the mzIdentML files for the peptide hypothesis, deglycosylated aliquots of AGP and Phil82 were run on LC-MS, and then searched using Peaks Studio 8.5. For AGP, the database used was the entire Uniprot Swiss-Prot database with *Homo sapiens* specified for the species taxonomy; For Phil82, the database consisted of the protein sequences for hemagglutinin and neuraminidase from the A/Philippines/2/1982 strain, the internal influenza A protein sequences from A/Puerto Rico/8/1934, all appended to the host proteome, *Gallus gallus*. The deglycosylated sample files were searched trypsin or chymotrypsin as the proteolytic enzyme, with a maximum of three missed cleavages, and one nonspecific cleavage was allowed. Carbamidomethylation of cysteine was set as a fixed modification; oxidation of methionine and deamidation of asparagine were specified as variable modifications. The precursor ion (MS1) mass error tolerance was 10 ppm and the fragment ion (MS/MS) error tolerance was 0.02 Da. Protein identifications required a minimum of two unique peptides. A target-decoy false discovery threshold of 1% at the peptide level was used. For the glycan hypothesis, a combinatorial hypothesis was constructed using the following glycan composition rules: #Hex 3-10, #HexNAc 2-9, #Fuc 0-5, #Neu5Ac 0-4, with #Fuc < #HexNAc and #HexNAc > #Neu5Ac + 1, and the addition of 1 sulfate group as a variable modification. In the GlycReSoft searches, the possibility of 1 ammonium adduct was considered in the identification.

Supplementary Figure S-3





Supplementary Figure S-2. CE-MS extracted ion electropherograms of individual glycopeptides for a tryptic digest of α1-acid glycoprotein (AGP). Supplemental data from a previous publication(Khatri, K., Klein, J.A., et al. 2017) are shown here.

SUPPLEMENTARY INFORMATION - REFERENCES

Chang D, Hackett WE, Zhong L, Wan XF, Zaia J. 2020. Measuring Site-specific Glycosylation Similarity between Influenza a Virus Variants with Statistical Certainty. *Mol Cell Proteomics*, 19:1533-1545.

Hennig R, Cajic S, Borowiak M, Hoffmann M, Kottler R, Reichl U, Rapp E. 2016. Towards personalized diagnostics via longitudinal study of the human plasma N-glycome. *Biochim Biophys Acta*, 1860:1728-1738.

Hennig R, Rapp E, Kottler R, Cajic S, Borowiak M, Reichl U. 2015. N-Glycosylation Fingerprinting of Viral Glycoproteins by xCGE-LIF. *Methods in molecular biology*, 1331:123-143.

Hennig R, Reichl U, Rapp E. 2011. A software tool for automated high-throughput processing of CGE-LIF based glycoanalysis data, generated by a multiplexing capillary DNA sequencer. Poster presented at 5th Glycan Forum, Berlin, Germany.

Jansen BC, Falck D, de Haan N, Ederveen ALH, Razdorov G, Lauc G, Wuhrer M. 2016. LaCyTools: A Targeted Liquid Chromatography-Mass Spectrometry Data Processing Package for Relative Quantitation of Glycopeptides. *J. Proteome Res.*, 15:2198-2210.

Kammeijer GS, Kohler I, Jansen BC, Hensbergen PJ, Mayboroda OA, Falck D, Wuhrer M. 2016. Dopant Enriched Nitrogen Gas Combined with Sheathless Capillary Electrophoresis-Electrospray Ionization-Mass Spectrometry for Improved Sensitivity and Repeatability in Glycopeptide Analysis. *Anal Chem*, 88:5849-5856.

Khatri K, Klein JA, Haserick J, Leon DR, Costello CE, McComb ME, Zaia J. 2017. Microfluidic capillary electrophoresis-mass spectrometry for analysis of monosaccharides, oligosaccharides and glycopeptides. *Anal Chem*, 89:6645-6655.

Kohler I, Augsburger M, Rudaz S, Schappler J. 2014. New insights in carbohydrate-deficient transferrin analysis with capillary electrophoresis-mass spectrometry. *Forensic Sci. Int.*, 243:14-22.

Lageveen-Kammeijer GSM, de Haan N, Mohaupt P, Wagt S, Filius M, Nouta J, Falck D, Wuhrer M. 2019. Highly sensitive CE-ESI-MS analysis of N-glycans from complex biological samples. *Nature communications*, 10:2137.

Lageveen-Kammeijer GSM, Haan dN, Moahupt P, Wagt S, Filius M, Nouta J, Falck D, Wuhrer M. 2019. Highly sensitive CE-ESI-MS analysis of N-glycans from complex biological samples - Data Sets. In: MassIVE editor.

Ruhaak LR, Hennig R, Huhn C, Borowiak M, Dolhain RJ, Deelder AM, Rapp E, Wuhrer M. 2010. Optimized workflow for preparation of APTS-labeled N-glycans allowing high-throughput analysis of human plasma glycomes using 48-channel multiplexed CGE-LIF. *J Proteome Res*, 9:6655-6664.