Published in final edited form as:

Nat Protoc. 2011 March ; 6(3): 253–269. doi:10.1038/nprot.2010.176.

Integrated mass spectrometry–based analysis of plasma glycoproteins and their glycan modifications

Hong Wang^{1,2}, Chee-Hong Wong¹, Alice Chin¹, Ayumu Taguchi¹, Allen Taylor¹, Samir Hanash¹, Sadanori Sekiya³, Hidenori Takahashi³, Masaki Murase³, Shigeki Kajihara⁴, Shinichi Iwamoto³, and Koichi Tanaka³

¹Molecular Diagnostics Program, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA

²Ventana Medical Systems, Oro Valley, Arizona, USA

³Koichi Tanaka Mass Spectrometry Research Laboratory, Shimadzu Corporation, Nakagyo-ku, Kyoto, Japan

⁴Technology Research Laboratory, Shimadzu Corporation, Seika-cho, Soraku-gun, Kyoto, Japan

Abstract

We present a protocol for the identification of glycosylated proteins in plasma followed by elucidation of their individual glycan compositions. The study of glycoproteins by mass spectrometry is usually based on cleavage of glycans followed by separate analysis of glycans and deglycosylated proteins, which limits the ability to derive glycan compositions for individual glycoproteins. The methodology described here consists of 2D HPLC fractionation of intact proteins and liquid chromatography–multistage tandem mass spectrometry (LC-MS/MSⁿ) analysis of digested protein fractions. Protein samples are separated by 1D anion-exchange chromatography (AEX) with an eight-step salt elution. Protein fractions from each of the eight AEX elution steps are transferred onto the 2D reversed-phase column to further separate proteins. A digital ion trap mass spectrometer with a wide mass range is then used for LC-MS/MSⁿ analysis of intact glycopeptides from the 2D HPLC fractions. Both peptide and oligosaccharide compositions are revealed by analysis of the ion fragmentation patterns of glycopeptides with an intact glycopeptide analysis pipeline.

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>.

Correspondence should be addressed to H.W., hong.wang.hw3@ventana.roche.com.

AUTHOR CONTRIBUTIONS: H.W., K.T. and S.H. conceived and supervised the project. H.W. contributed to the development and evaluation of the methods and combined the data and wrote the protocol. C.-H.W. designed and developed the analysis software. A.C. performed the sample fractionation. H.W. and S.S. contributed to the mass spectrometry analysis of the sample. H.T. and S.I. optimized the MALDI-DIT instrument. A. Taguchi, A. Taylor and A.C. prepared the analyzed sample. H.T., M.M. and S.K. designed and developed the MALDI-DIT software.

COMPETING FINANCIAL INTERESTS: The authors declare no competing financial interests.

Note: Supplementary information is available via the HTML version of this article.

INTRODUCTION

Glycosylation is the most widespread and complex form of protein post-translational modification, as more than 50% of proteins in humans are glycosylated^{1,2}. Glycosylation has a key role in protein folding, stability, activity, trafficking, molecular recognition and immunogenicity. Aberrant glycosylation occurs in numerous diseases^{3–5}. In cancer, changes in oligosaccharide structures occur as part of initial oncogenic transformation, as well as part of key events in promoting tumor cell invasion and metastasis⁶. An immune response may occur in cancer resulting from aberrant glycosylation^{6–12}. Therefore, profiling glycoproteins in disease states may have relevance to the development of biomarkers as well as therapeutics that target proteins with aberrant glycosylation.

Most protein glycosylation consists of the attachment of oligosaccharides to either the side chain of asparagine (*N*-linked) in an Asp-Xxx-Y (Y is Ser or Thr or occasionally Cys) motif, where Xxx is any amino acid except proline¹³, or to Ser or Thr (*O*-linked) unrelated to a consensus sequence¹⁴. For *N*-linked glycoproteins, processing of the sugar groups occurs post-translationally in the lumen of the endoplasmic reticulum and continues in the Golgi apparatus. *N*-linked protein glycosylation pathways in eukaryotes can be divided into two major processes: the assembly of the lipid-linked oligosaccharide at the membrane of the endoplasmic reticulum and the transfer of the oligosaccharide from the lipid anchor dolichyl pyrophosphate to selected asparagine residues of nascent polypeptides with the recognition of asparagine residues in the sequence Asp-Xxx-Ser/Thr (where Xxx can be any amino acid except proline)¹³. Attachment of sugars in *O*-linked glycoproteins occurs post-translationally in the Golgi apparatus, most commonly by a *N*-acetyl galactosaminyltransferase that transfers a *N*-acetyl-galactosamine residue to the side chain of a serine or a threonine residue. Subsequently, a stepwise enzymatic elongation by specific transferases yields several core structures, which are further elongated or modified by sialylation, sulfatation, acetylation, fucosylation or polylactosamine extension¹⁴. Whereas the biosynthesis of nucleic acids and proteins is carried out by template mechanisms, glycoproteins are synthesized by a complex nontemplate process involving sugar nucleotide synthases, glycosyltransferases and glycosidases. Protein glycosylation introduces considerable heterogeneity through the generation of different 'core' oligosaccharides and the variable addition of outer arm sugars as well as incomplete enzymatic trimming of the attached sugar side chains. There are currently two major strategies for the analysis of glycoproteins in complex mixtures, such as biological fluids, plasma or serum. One is the affinity capture-based enrichment of glycoproteins followed by removal of their glycans and analysis of nonglycopeptides to determine their protein identity by mass spectrometry. Although this approach allows the determination of *N*-linked glycosylation sites, the information about glycans is lost during sample processing^{11,15–18}. An alternative glycomics-based approach focuses on the analysis (without knowledge of protein identity and glycosylation sites) of pooled glycans removed from their proteins^{8,19–23}. Thus, removal of glycans from the proteins to which they are attached, followed by profiling of either protein backbones or glycans, provides limited information regarding glycan composition of individual glycoprotein of interest. Consequently, there is a need for improved strategies for

glycoprotein analysis that allow the elucidation of glycan composition of glycoproteins that occur in complex mixtures.

We have previously implemented a mass spectrometry-based strategy for in-depth analysis of the plasma proteome^{24–29}. Here, we present a protocol in which analyses of protein sequence, glycosylation sites and glycan compositions are integrated. The protocol has been applied to the analysis of plasma glycoproteins. High-resolution 2D HPLC-based fractionation of intact proteins is coupled with high-precision mass spectrometry of digested fractions for measurement of intact glycopeptides/peptides and their fragments. An automated data analysis pipeline is utilized for identification of glycoproteins and glycan attachment sites (see Fig. 1 for general overview of the procedure). The protocol described here has been applied to the plasma glycoproteome and can also be applied, in principle, to the study of glycoproteomes of other biological fluids. Alternative sample preparation procedures may be appropriate for other sample types^{30,31}.

Experimental design

Sample preparation—Human plasma samples are processed by immunodepletion chromatography followed by an online, intact protein-based 2D HPLC separation procedure according to a protocol published by our group^{24,28}. The dynamic range of plasma protein abundance spans at least 12 orders of magnitude³². Analysis of sufficient plasma sample to detect low-abundance proteins invariably means excessive loading of albumin and several other high-abundance proteins that constitute > 90% of the total protein mass of plasma. Therefore, removal of high-abundance proteins from plasma has become a useful step for plasma sample preparation. We have included in this protocol an immunodepletion step to remove abundant proteins, including albumin, IgG, IgA, transferrin, haptoglobin and α -1-antitrypsin. Other immunodepletion approaches are available for removal of a greater number of high-abundance proteins from plasma^{33,34}.

The flow-through fractions (i.e., after removal of the high-abundance proteins) from immunodepletion are collected and concentrated with a centrifugal filter device (3-kDa M_r cutoff). After reduction (i.e., the cleavage of protein disulfide bonds formed between the thiol groups of cysteine residues via reduction with DTT) and alkylation (i.e., blocking of the thiols by alkylation with acrylamide), the protein sample is subjected to an intact protein-based 2D HPLC separation procedure. Intact protein-based separation before protein digestion for LC-MS/MSⁿ analysis is crucial for the assessment of protein post-translational modifications, notably, glycosylation. Intact protein-based separation is carried out gel-free, using an online orthogonal 2D HPLC system consisting of 1D AEX, and 2D reversed-phase (RP) chromatography²⁸. Usually, we process 0.6–1.0 ml of plasma and load 5–10 mg of immunodepleted plasma onto the 2D HPLC system. Typical chromatograms following immunodepletion and 2D HPLC are presented in Figure 2. The collected 2D HPLC protein fractions are immediately transferred to an -80°C freezer or instantly frozen with liquid nitrogen for subsequent lyophilization. The lyophilized protein fractions are subjected to in-solution digestion with trypsin followed by glycopeptide capture for subsequent LC-MS/MS analysis (see Fig. 1).

Glycopeptide enrichment—Glycopeptides from the digested protein fractions are enriched by hydrophilic interaction chromatography as reported by Wada *et al.*³⁵. We optimized this procedure by increasing washing times to remove nonglycopeptides from Sepharose CL-4B and increase glycopeptide recovery. Using transferrin as a test case, we achieved an average recovery of transferrin glycopeptides of ~80% (see Supplementary Table 1). Some non-glycopeptide binding to Sepharose CL-4B was observed; however, sample complexity was reduced compared with the non-treated sample (see Supplementary Fig. 1). Moreover, when glycopeptide mixtures were spotted onto the matrix-assisted laser desorption/ionization (MALDI) plate with an online reversed-phase liquid chromatography (RPLC) AccuSpot instrument (Shimadzu), the nonglycopeptides were well resolved from the glycopeptides. Thus, glycopeptide ionization efficiency is greatly improved by reducing signal suppression from nonglycopeptides in the same well on the MALDI plate.

On-plate enrichment of glycopeptides—The use of MALDI plates coated with boronate for on-plate glycopeptide enrichment before MALDI-MS analysis requires repeat wash-dry cycles for desalting, which reduces sample recovery and analysis throughput³⁶. We developed a protocol for online RPLC-MALDI plate spotting without relying on wash-dry cycles. The enriched glycopeptide mixture is separated using a capillary RPLC column and the resolved fractions are coaxially mixed online with matrix (2,5-dihydroxybenzoic acid, DHB) and spotted onto the MALDI plate with an AccuSpot. The efficiency of glycopeptide enrichment with two types of MALDI plates (μ Focus plate and metal plate) was evaluated based on glycopeptide peak intensity measured by MALDI-DIT (digital ion trap, DIT) (**GP-1**: CGLVPV LAENY $\underline{\text{N}}$ *(5Hex4HexNAc2Sia)K and **GP-2**: QQQH $\underline{\text{L}}$ FGSN*(5Hex 4HexNAc2Sia)VTDCSG $\underline{\text{N}}$ FCLFR) as shown in Figure 3. We found that MS signals obtained using a μ Focus plate are much higher than those obtained with metal plate. The efficiency of glycopeptide enrichment with the μ Focus plate is 1.5–2.5 times greater than the metal plate. We also found that DHB and glycopeptides result in heterogeneous crystallization on the metal plate, and the crystals are randomly distributed on the plate, which complicates automated measurements. However, the co-crystallized glycopeptides and DHB form a crystal ring with an ~200- μ m section around the center of the μ Focus plate. DHB usually results in heterogeneous crystallization with air-drying, complicating automated analysis due to spot-to-spot variation. As μ Focus plates have a hydrophobic coating on the surface, DHB and glycopeptides are focused against the hydrophobic layer toward the hydrophilic center of the plate to form a ring of homogenous crystals. Glycopeptides are distributed across the 200- μ m thickness of the ring, with reduced crowding of crystals at the center of the plate (Fig. 3a). The peak intensities of glycopeptides **GP-1** and **GP-2** were highest at the outer edge of the ring ('Outside' in Fig. 4) and decreased to the lowest at the center of the plate ('Center' in Fig. 4). The peak intensity distribution of the two glycopeptides at four different locations in the plate is shown in Figure 4, and their corresponding mass spectra are shown in Figure 5. The signal-to-noise observed was fairly stable across the ring with a gradual decrease in peak intensity from the outside to the inside, which became worse at the center of the plate. Thus, this *in situ* enrichment approach using μ Focus plates is effective and sensitive for analysis of glycopeptides, and provides a high-throughput format for online RPLC spotting for automated measurement of glycopeptides using MALDI-MS. An alternative matrix such as

super-DHB or another quick recrystallization procedure may also have utility for crystallization^{37,38}.

Mass spectrometry analysis—In principle, the protocol described here can be implemented on mass spectrometry instruments that have the capability for MS/MS measurement. In general, a high-resolution instrument, such as a time of flight, Orbitrap, or Fourier transform (FT), should be used for detection of the precursor mass-to-charge ratio (m/z) with high mass accuracy (< 20 p.p.m.). An instrument with a faster scan rate and higher sensitivity, such as an ion trap–quadrupole instrument, is usually selected for MS/MS analysis, although the mass accuracy for measured fragment ions is lower compared with the measurement of the precursor ions. As illustrated in Figure 1, we use LC–electrospray ionization (ESI)–linear trap quadrupole/ Fourier transform (LTQ/FT) for glycoprotein identification based on non-glycopeptide constituents, FT is used to detect precursors and LTQ is used to detect the fragment ions. Both FT and LTQ are used in positive ion mode. We used LC–MALDI–DIT for glycoprotein characterization based on glycopeptides analysis. DIT provides a wide m/z scan range (m/z 1,600–9,100) and low m/z cutoff ($\sim 10\%$ of the precursor m/z)³⁹. Glycopeptides and their fragment ions are measured with higher resolution positive ion mode by MALDI–DIT MS/MS. For glycopeptide analysis, MALDI–quadrupole/ion trap (Q/IT) or MALDI–quadrupole–time of flight (Q/TOF) may be used and such instruments are available from most of the major vendors.

Data analysis—Acquired spectra are automatically processed through an in-house built intact glycopeptide analysis pipeline (IGAP) to obtain integrated information pertaining to glycoprotein amino acid sequences, glycosylation sites and glycan compositions (See Fig. 6 for an overview). In brief, LC–ESI–MS/MS data are first converted to mzXML format using ReAdW software to generate the peak list for protein database searches. X!Tandem⁴⁰ was used as the search engine and the International Protein Index⁴¹ human protein knowledge base was used as the database. Other search engines such as Mascot⁴² or SEQUEST⁴³ may be used. The identified proteins constituted the ‘refinement protein database’ for glycopeptide analysis, which is essentially the list of confidently identified proteins in the sample using LC–ESI–MS/MS. The use of the refinement protein database for glycopeptide analysis is based on the assumption that for each *N*-glycosylated peptide, an unmodified peptide from the glycosylated protein is also present.

The acquired glycopeptide tandem spectra (MS/MS) using MALDI–DIT–MS/MS are processed through IGAP to compute the theoretical tryptic peptides from the generated Refinement Protein Database with 0, 1 and 2 missed cleavages. Only peptides containing the consensus site Asn–Xaa–Ser/Thr/Cys (where Xaa is not Pro), are considered as the potential peptide moiety of the *N*-glycopeptide, referred to as *N*-peptide. For every *N*-glycopeptide (combination of *N*-peptide and *N*-glycan) with a 0.5-Da tolerance of the precursor $[M + H]^+$, its theoretical spectrum is generated and matched against the measured spectrum. To compensate for the selection of the most intense peak for the precursor, the precursor tolerance is expanded by considering the same tolerance window centered on the first, second and third ¹³C isotope peaks for *N*-glycopeptide. Acrylamidation of cysteine and deamination of N-terminal glutamine are considered for *N*-peptide when generating a

potential *N*-glycopeptide. Both b-ion and y-ion series of *N*-peptide are calculated along with the y-ion series and $X^{0,2}$ of *N*-glycan with intact peptide backbones for the theoretical spectrum. During MS/MS fragmentation, the inner core *N*-acetyl-glucosamine undergoes the $X^{0,2}$ cross-ring ring cleavage, resulting in mass differences of 120 and 83 Da. The ion fragment match tolerance is 0.5 Da. The measured spectrum is cumulative intensity normalized⁴⁴. The measured m/z range is divided into 100-Th regions and the eight most intensive peaks in each such region are extracted to form the intensity spectrum, which the theoretical spectrum is matched against. Each match produces a paired metric (score, probability of random match) for both the glycan moiety and the glycopeptide. The glycopeptide MS/MS ion search process is further illustrated in Figure 7. These scorings are adapted from A-Score, as described by Beausoleil *et al.*⁴⁵. The collected matches are ranked by glycan moiety score, glycopeptide probability of random match and glycopeptide score. The top 20 glycan composition results are retained and further narrowed down based on other available information, such as cysteine acrylamidation and composition plausibility of swapping between one *N*-acetylneuraminic acid (291) and two deoxyhexoses ($292 = 146 + 146$).

MATERIALS

REAGENTS

- Human whole blood ! **CAUTION** Human blood may transmit infectious diseases such as hepatitis B or HIV. Use protective gloves, dispose of contaminated sharp waste in safety containers at the point of use, dispose of contaminated waste after autoclaving.
- ! **CAUTION** Experiments with human subjects must comply with institutional and national guidelines.
- Trypsin (mass spectrometry grade; Promega, cat. no. V511A; store at 4 °C)
- DL-DTT (Promega, cat. no. V3155) (store at -20 °C)
- Acrylamide (Fluka, cat. no. 01696) (store at 4 °C, in dark)
- Octyl β -D-glucopyranoside (Sigma-Aldrich, cat. no. 03757) (store at -20 °C)
- Ammonium bicarbonate (Thermo Fisher, cat. no. A643-500)
- Trizma base (Sigma-Aldrich, cat. no. T6066)
- Water, HPLC grade (Thermo Fisher, cat. no. W5-4)
- Acetonitrile (Thermo Fisher, cat. no. A988-4) ! **CAUTION** Highly flammable and toxic. Use appropriate protective attire (lab coats, safety glasses and latex gloves) and pipette under fume hood.
- 2-Propanol (Thermo Fisher, cat. no. A451-4) ! **CAUTION** Highly flammable and toxic. Use appropriate protective attire (lab coats, safety glasses and latex gloves) and pipette under fume hood.

- 1-Butanol (Thermo Fisher, cat. no. A383-1) ! **CAUTION** Highly flammable and toxic. Use appropriate protective attire (lab coats, safety glasses and latex gloves) and pipette under fume hood.
- Ethanol (Thermo Fisher, cat. no. AC61509-5000) ! **CAUTION** Highly flammable and toxic. Use appropriate protective attire (lab coats, safety glasses and latex gloves) and pipette under fume hood.
- Immunodepletion Buffer A (Agilent Technologies, cat. no.5185-5987)
- Immunodepletion Buffer B (Agilent Technologies, cat. no.5185-5988)
- Transferrin (Sigma Aldrich, cat. no. 90190) (store at 4 °C)
- 2,5-dihydroxybenzoic acid (DHB; LaserBio Labs, cat. no. M003) (store at 4 °C, in dark) ! **CAUTION** Harmful if inhaled, if in contact with skin or if swallowed.
- Sepharose CL-4B (GE Healthcare, cat. no.17-0150-01) (store at 4 °C)
- Hydrochloric acid (Thermo Fisher, cat. no. A466-250) ! **CAUTION** Highly corrosive. Use appropriate protective attire (lab coats, safety glasses and latex gloves) and pipette under fume hood.
- Formic acid (Pierce, cat. no. 28905) ! **CAUTION** Highly corrosive. Use appropriate protective attire (lab coats, safety glasses and latex gloves) and pipette under fume hood.
- Trifluoroacetic acid (TFA; Supelco, cat. no. 3-3077) ! **CAUTION** Highly corrosive. Use appropriate protective attire (lab coats, safety glasses and latex gloves) and pipette under fume hood.
- Urea (Thermo Fisher, cat. no. U15-3)
- Sodium chloride (Thermo Fisher, cat. no. S271-1)
- Magic C18 packing materials 100A, 5 µm (Michrom Bioresources, cat. no. PM5/66100/00)
- Tris-HCl
- Hydrochloric acid

EQUIPMENT

- BD Vacutainer plastic EDTA tubes (10 ml), 16×100 mm (BD, cat. no. 366643)
- Online 2D HPLC system (Shimadzu) consists of (i) one injector with a 10-µl sample loop, (ii) one UV detector (SPD-20), (iii) two binary pumps (LC-20AD), (iv) one quaternary pump (LC-10Ai), (v) one degasser (DGU-20A5), (vi) one column oven (CTO-20AC), (vii) four fraction collectors (FRC-10A), (viii) one controller (SCL-10A). The whole system is controlled by EZ Start software (version 7.4)
- AccuSpot (Shimadzu), a spotting device for MALDI plate, controlled by MainDlg (version 2.03)

- MALDI-DIT (Shimadzu), controlled by DIT Console (version 1.0)
- LTQ-FT (Thermo Scientific), controlled by Xcalibur (version 2.0.7)
- Nanoacquity UPLC (Waters Corporation), controlled by MassLynx (version 4.1)
- NanoLC 1D (Eksigent Technologies), controlled by EksigentStart (version 2.08)
- 384 MALDI target plate, stainless steel (Shimadzu)
- 384 MALDI target plate, μ Focus (Hudson Surface Technology)
- 384 MALDI target plate, glyco-affinity (Hudson Surface Technology)
- Immunodepletion column, 10×100 mm (Agilent Technologies, cat. no. HU-6HC)
- Anion exchange column, 7.5×150 mm (Column Technology, cat. no. NA75150WP)
- Reversed-phase column, 4.6×150 mm (Column Technology, cat. no. RP5D46150)
- Onyx Monolithic C18, 0.1×150 mm (Phenomenex, cat. no. CH0-7646)
- Magic C18 resin (packed in house)
- Fritted capillary tubing, 0.075×500 mm (New Objective, cat. no. PF360-75-15-N-5)
- Trap-column, 0.18×20 mm (Waters, cat. no. 186002841)
- Software (DITViewer, version 1.1 (Shimadzu); ReAdW version 1.2; X!Tandem version 2005.12.01; Microsoft Excel 2003 or later; Intact Glycopeptide Analysis Pipeline-IGAP (see EQUIPMENT SETUP))
- $0.22\text{-}\mu\text{m}$ RC syringe filter (MicroSolv, cat. no. 58022-R13-C)
- Amicon Ultra YM-3 centrifugal filters (M_r cutoff 3 kDa) (Millipore, cat. no. UFC900324)
- Allegra 64R High-Speed Refrigerated Benchtop Centrifuge (Beckman Coulter)
- FreeZone Plus 12-Liter Cascade Console Freeze-Dry System (Labconco)
- SpeedVac System (GMI, SC 210A)
- Sample loops, polyethyl ethyl ketone (PEEK; 2 ml, 10 ml)

▲ CRITICAL Although we have used the MALDI-DIT from Shimadzu Corporation for protein glycosylation analysis, the protocol for protein glycosylation can also be implemented using other MALDI instruments (e.g., MALDI-QIT from Shimadzu Corporation). We use ESI-LTQFT from Thermo Scientific for protein identification on the basis of the analysis of nonglycopeptides. However, the method for protein identification can also be carried out with an ESI-LTQ Orbitrap from Thermo Scientific, or with ESI-QTOF instruments from Bruker Daltonics, Waters, Agilent or Applied Biosystems.

REAGENT SETUP

Human plasma—Human whole blood was drawn from adult subjects at a Food and Drug Administration–approved facility at the Fred Hutchinson Cancer Research Center. The blood is collected into a BD Vacutainer tube (10 ml) containing ~1.8 mg K₂ EDTA per ml blood. Invert the tube gently 10 times to mix blood and anticoagulant. The blood is then centrifuged for 10 min at 1300g at room temperature (22 °C). Carefully aspirate the supernatant (plasma) and transfer the plasma into cryovials. Store at –80 °C until use. !

CAUTION Human blood may transmit infectious diseases, such as hepatitis B or HIV. Use protective gloves and dispose of contaminated sharps in appropriate safety containers at the point of use. Dispose of contaminated waste after autoclaving it. ! **CAUTION** Experiments using human blood must comply with the legal requirements of the country and guidelines of the institution.

Protein denaturation solution—8 M urea, 100 mM Tris-HCl, and 0.5% (wt/vol) octyl β -D-glucopyranoside, pH 8.5 (pH is adjusted by adding 1 M HCl).

▲ **CRITICAL** Prepare fresh immediately after standard protein (transferrin) is weighed or immunodepleted plasma is concentrated.

Protein reduction solution—1 M DTT in water (freshly made).

Protein alkylation reagent—Dry acrylamide (freshly weighed, stored in the dark at 22 °C for up to 24 h).

Protein digestion buffer—50 mM ammonium bicarbonate in 2% acetonitrile/98% H₂O (vol/vol) (freshly made).

Binding solution (for glycopeptide enrichment)—1-Butanol/ethanol/H₂O, 4:1:1 (vol/vol/vol). Store at 22 °C for up to 1 year.

Eluting solution (for glycopeptide enrichment)—Ethanol/H₂O, 1:1 (vol/vol). Store at 22 °C for up to 1 year.

Buffers for immunodepletion chromatography

Buffer A: Equilibration/ loading/washing (Agilent Technologies, cat. no.5185-5987).

Buffer B: Elution (Agilent Technologies, cat. no.5185-5988).

Mobile phase for 2D HPLC protein separation

AEX mobile phase A: 20 mM Tris-HCl, 6% (vol/vol) isopropanol, and 4 M urea, pH 8.5 (pH is adjusted by adding 1 M HCl) (must freshly made before use).

AEX mobile phase B: 20 mM Tris-HCl, 6% (vol/vol) isopropanol, 4 M urea, and 1 M NaCl, pH 8.5 (pH is adjusted by adding 1 M HCl) (must be freshly made before use). ▲

CRITICAL The urea-containing solution used in AEX separation should be prepared

freshly and immediately used at room temperature to reduce nonspecific protein carbamylation.

RP mobile phase A: HPLC grade water/acetonitrile/TFA, 95:5:0.1 (vol/vol/vol).

RP mobile phase B: HPLC grade water/acetonitrile/TFA, 10:90:0.1 (vol/vol/vol).

Mobile phase for RPLC-MS/MS

RP mobile phase A: 0.1% formic acid in HPLC grade water (vol/vol).

RP mobile phase B: 0.1% formic acid in acetonitrile (vol/vol).

DHB matrix solution: 5 mg DHB in 1 ml of 30% acetonitrile/70% H₂O (vol/vol); prepare fresh before use.

EQUIPMENT SETUP

LC-ESI-MS/MS analysis—Couple NanoLC 1D with ESI-LTQFT for non-glycopeptide analysis with online LC-ESI-LTQFT MS/MS. A RPLC column (0.075 × 250 mm) packed with Magic C18 in-house is used with a flow rate of 300 nl min⁻¹.

LC-MALDI-MS/MS analysis—Couple Nanoacquity UPLC with AccuSpot for glycopeptide analysis with offline LC-MALDI-DIT MS/MS. A monolithic C18 column (0.1 × 150 mm) is used with a flow rate of 1 µl min⁻¹.

IGAP software installation—Download the three supplementary zip folders: Supplementary Program, Supplementary Data 1 and Data 2. Unzip these files (e.g., with WinZip) and save them into a folder named 'IGAP' on the computer to be used. No installation script needs to be executed.

PROCEDURE

Sample preparation: immunodepletion chromatography ● TIMING 50 min

- 1| Equilibrate the immunodepletion column with Buffer A for 10 min at 3 ml min⁻¹.
- 2| Wash the 2-ml sample loop (PEEK) with 5 ml of Buffer A.
- 3| Mix 300 µl of plasma with 1.5 ml of Buffer A and filter the diluted plasma through a 0.22-µm RC syringe filter.

▲ **CRITICAL STEP** Mixing the plasma sample with Buffer A (1:5, vol/vol) to dissociate low-abundance proteins that are bound to high-abundance proteins such as albumin is crucial for better recovery of low-abundance proteins after immunodepletion approach.

? TROUBLESHOOTING

- 4| Load the diluted plasma onto the 2-ml sample loop and start the step elution (see Box 1 for a recommended step elution program). The low-abundance proteins

are in the flow-through fraction (see Fig. 2a for a typical immunodepletion chromatogram).

■ **PAUSE POINT** Low-abundance proteins can be stored at -80°C for months until use or may immediately be concentrated (see Step 5).

- 5| Concentrate the flow-through fraction (i.e., the low-abundance protein fraction) with the Amicon YM-3 filter at 3,000g and 4°C . It usually takes 60–75 min to concentrate the sample from 15 ml down to 0.5 ml.

▲ **CRITICAL STEP** Set the centrifuge temperature to 4°C and balance the Amicon YM-3 filters before starting the centrifuge. For biological fluids other than plasma and serum (e.g. urine, amniotic fluid, saline) that have special characteristics, (low total protein concentration, low albumin, low IgG concentration, high salt concentration, etc.), the sample (1 ml or more) is processed with a centrifugal filter device (the Amicon YM-3 filter at 3,000g, 4°C for 60 min) to 0.1 ml to remove salt and concentrate proteins. The sample is subsequently fractionated by 2D HPLC.

- 6| After immunodepletion, the total protein mass is expected to be reduced by ~90%. This is useful to keep in mind when evaluating the efficiency of the immunodepletion column. The lifetime of the immunodepletion column is ~200 injections, and efficiency is between 90 and 95%.

? TROUBLESHOOTING

Sample preparation: 2D HPLC protein separation ● TIMING 9.5 h

- 7| Mix 0.5 ml of concentrated sample with 0.5 ml of protein denaturation solution and measure protein concentration using the Bradford method⁴⁶.
- 8| Reduce protein disulfide bonds for 2 h at room temperature by adding 1M DTT to the sample (i.e., add 0.66 mg DTT per mg protein).

▲ **CRITICAL STEP** As samples are in the protein denaturation solution that contains 8 M urea, protein reduction should be performed at 22°C to avoid protein carbamylation due to isocyanic acid degradation from urea at higher temperatures.

- 9| Alkylate cysteines with acrylamide immediately after protein reduction by adding dry acrylamide to the sample (i.e., add 7.1 mg acrylamide per mg protein). Protein alkylation takes 1 h at room temperature in the dark.

▲ **CRITICAL STEP** As for protein reduction, protein alkylation should be done at room temperature to avoid protein carbamylation. Acrylamide is highly light sensitive and unstable. Protein alkylation needs to be performed in the dark immediately after protein reduction. As an alternative, iodoacetamide can be used as the alkylation reagent.

- 10| Dilute sample with AEX mobile phase A to 9 ml immediately after protein alkylation. Filter the diluted sample through a $0.22\text{-}\mu\text{m}$ RC syringe filter and

load the filtered sample onto a 10-ml sample loop (PEEK). Start online 2D HPLC protein separation at room temperature. We recommend an eight-step elution for the first-dimensional AEX chromatography (see Box 2 for a recommended eight-step elution). The discrete fraction of absorbed proteins displaced from the AEX column during each of the eight steps is alternatively online-loaded onto one of the two RP columns for sequential 2D RP chromatography separation (see Box 3 for a recommended RP gradient elution). See Figure 2b for a typical 2D HPLC chromatogram.

▲ CRITICAL STEP Protein fractions eluted from 1D AEX chromatography are trapped onto the RP column for 46 min. To achieve higher resolution separation and prevent salt precipitation within the RP column in the presence of acetonitrile, it is critical to perform desalting with 5% RP mobile phase B for 5 min before initiating the 2D RP chromatography (see Box 3 for a recommended RP gradient program).

? TROUBLESHOOTING

- 11| Collect protein fractions from the RP columns and store them at -80°C or freeze the protein fractions in liquid nitrogen.
- 12| Lyophilize protein fractions with the freeze-dry system. It usually takes 30 h to dry the protein fractions completely.

▲ CRITICAL STEP The collected 2D HPLC protein fractions should be totally frozen before being subjected to lyophilization. We recommend that 2D HPLC protein fractions should not be dried by SpeedVac because it requires longer processing time at room temperature.

Box 2

RECOMMENDED 2D HPLC PROGRAM—1D AEX STEP GRADIENT

Dilute the alkylated plasma (~0.5 ml) to 9 ml with AEX mobile phase A. Filter the sample solution with an 0.22- μm syringe filter and load the sample to a 10-ml sample loop. We recommend performing 2D HPLC separation with the following AEX step elution, using AEX mobile phases A and B for 2D HPLC protein separation (described in REAGENT SETUP):

Column equilibration

- 15 min of 100% AEX mobile phase A, 2 ml min^{-1}

Step elution

• Step 1 (46 min)	Inject sample: 100% buffer A, 0.8 ml min^{-1}
• Step 2 (46 min)	5% Mobile phase B, 0.8 ml min^{-1}
• Step 3 (46 min)	10% Mobile phase B, 0.8 ml min^{-1}
• Step 4 (46 min)	15% Mobile phase B, 0.8 ml min^{-1}
• Step 5 (46 min)	20% Mobile phase B, 0.8 ml min^{-1}
• Step 6 (46 min)	30% Mobile phase B, 0.8 ml min^{-1}

- Step 7 (46 min) 50% Mobile phase B, 0.8 ml min⁻¹
- Step 8 (46 min) 100% Mobile phase B, 0.8 ml min⁻¹

In each salt step, a discrete fraction of the absorbed proteins is displaced from the AEX column onto one of the reversed-phase (RP) columns (there are two RP columns in the 2D HPLC system). An acetonitrile gradient (see Box 3) further separates these proteins during 2D RP chromatography and the eluted fractions are online-collected by a fraction collector.

Box 3

RECOMMENDED 2D HPLC PROGRAM— 2D RP GRADIENT (UV = 280 NM)

A discrete fraction of the absorbed proteins in AEX column is eluted and trapped onto one of the RP columns. We recommend performing 2D HPLC separation with the following RP gradient elution, using RP mobile phases A and B for 2D HPLC protein separation (described in REAGENT SETUP):

RP gradient elution

- 0.00 min 5% Buffer B, 2.1 ml min⁻¹
- 5.00 min 5% Buffer B, 2.1 ml min⁻¹ (for desalting)
- 10.0 min Start to collect fractions: 5% buffer B, 2.1 ml min⁻¹, 3 fractions per min, 0.7 ml per fraction
- 12.0 min 20% Buffer B, 2.1 ml min⁻¹
- 30.0 min 65% Buffer B, 2.1 ml min⁻¹
- 32.0 min 65% Buffer B, 2.1 ml min⁻¹
- 33.0 min 90% Buffer B, 2.1 ml min⁻¹
- 35.0 min 90% Buffer B, 2.1 ml min⁻¹
- 35.0 min 100% Buffer B, 3.0 ml min⁻¹
- 37.0 min 5% Buffer B, 2.1 ml min⁻¹
- 40.0 min Stop to collect fractions: 5% buffer B, 2.1 ml min⁻¹
- 46.0 min 5% Buffer B, 2.1 ml min⁻¹

Collect protein fractions from the RP columns and store at -80 °C or freeze the protein fractions with liquid nitrogen immediately after 2D HPLC separation for subsequent lyophilization.

Sample preparation: protein in-solution digestion ● TIMING 5.5 h

- 13| Add 20 µg of trypsin to 2.5 ml of protein digestion buffer to make a trypsin digestion solution with a trypsin concentration of 8 ng µl⁻¹.
- 14| Suspend individual lyophilized protein fractions in 50 µl of trypsin digestion solution and perform protein in-solution digestion at 37 °C for 5 h.

▲ **CRITICAL STEP** Addition of 2% (vol/vol) acetonitrile to the trypsin digestion solution (ammonium bicarbonate) is useful to maintain trypsin activity. Digestion usually takes 5 h at 37 °C, after which fractions are maintained at 4 °C for subsequent glycopeptide enrichment.

Sample preparation: glycopeptide enrichment ● TIMING 2 h (room temperature)

- 15| *Activation:* Mix 30 μ l of Sepharose CL-4B with 0.5 ml of eluting solution (ethanol/H₂O, 1:1, vol/vol) in a microcentrifuge tube, gently mixing for 30 s. Centrifuge at 3,000g for 30 s and remove the supernatant.
- ▲ **CRITICAL STEP** This step is needed to remove nonspecifically bound compounds from Sepharose CL-4B to reduce contamination and increase binding capacity.
- 16| *Equilibration:* Add 0.5 ml binding solution to activated Sepharose CL-4B, gently mixing for 30 s. Centrifuge at 3,000g for 30 s and remove the supernatant. Repeat this step three times.
- ▲ **CRITICAL STEP** This step is performed to condition Sepharose CL-4B; eluting solution should be removed completely.
- 17| *Adsorption:* Add 1 ml of binding solution to the conditioned Sepharose CL-4B, then add 50 μ l of digested protein solution, shaking and incubating for 45 min at room temperature. Then, centrifuge at 3,000g for 30 s and transfer the supernatants (nonglycopeptides) to a new microcentrifuge tube. Store at -20°C until use.
- 18| *Wash:* Add 0.5 ml of binding solution to Sepharose CL-4B, gently shaking for 30 s. Centrifuge at 3,000g for 30 s and discard the supernatant. Repeat this step three times.
- 19| *Elution:* Add 0.2 ml of eluting solution, shaking and incubating for 25 min at room temperature. Then, centrifuge at 3,000g for 30 s and transfer the supernatant (glycopeptides) to a new microcentrifuge tube. Repeat this step three times and combine the supernatants; store at -20°C until use. !
- CAUTION** 1-Butanol and ethanol are highly flammable and toxic. Use appropriate protective attire (lab coats, safety glasses and latex gloves) and perform all steps under a fume hood.

LC-MS/MS analysis: LC-ESI-LTQFT MS/MS of nonglycopeptides ● TIMING 2.5 h

- 20| Dry the nonglycopeptide sample (see Step 17) by removing the binding solution with the SpeedVac System. Resuspend the nonglycopeptides in 50 μ l of RP mobile phase A (0.1% formic acid in HPLC grade water (vol/vol)).
- 21| Load 20 μ l of nonglycopeptide sample onto a 20- μ l sample loop. Sample is trapped onto the trap column. Desalt with 2% RP mobile phase B (0.1% formic acid in acetonitrile (vol/vol)) for 10 min at 5 μ l min⁻¹.
- 22| Start the nano-RPLC gradient separation with a Magic C18 column at 300 nl min⁻¹ (see Box 4 for a recommended RP gradient) and acquire spectra in data-dependent mode with m/z range of 400 to 1800. Given the LTQ-FT data acquisition duty cycle, the top five most abundant + 2 or + 3 ions in each MS spectrum are selected for MS/MS analysis in automatic data-dependent acquisition mode. Set LTQ-FT instrument parameters as: capillary voltage of 2.0

kV; capillary temperature of 200 °C; FT resolution of 100,000; FT target value of 850,000; and LTQ target value of 10,000. Perform one micro-scan for MS and MS/MS.

? TROUBLESHOOTING

Box 4

RECOMMENDED RPLC GRADIENT FOR LC-MS/MS ANALYSIS

For online RPLC-ESI-LTQFT analysis, a Magic C18 column (0.075 × 250 mm) packed in house is used with a recommended flow rate of 300 nl min⁻¹. For offline RPLC-MALDI-DIT analysis, a Monolithic C18 column (0.1 × 250 mm) is used with a recommended flow rate of 1.0 µl min⁻¹. We recommend performing RPLC separation followed by RP gradient elution for both online RPLC-ESI-LTQFT and offline RPLC-MALDI-DIT, using RP mobile phases A and B for RPLC-MS/MS (described in REAGENT SETUP):

RP gradient elution

- 0.00 min 2% Buffer B
- 1.00 min 7% Buffer B
- 91.0 min 35% Buffer B
- 92.0 min 50% Buffer B
- 102 min 50% Buffer B
- 103 min 95% Buffer B
- 108 min 95% Buffer B
- 109 min 2% Buffer B
- 130 min 2% Buffer B

Load sample onto the C18 trap column and desalt for 10 min before initializing the above RP gradient elution. For online ESI-LTQFT analysis, separated nonglycopeptides are directly analyzed by LTQFT. For offline MALDI-DIT analysis, separated glycopeptide fractions are coaxially mixed online with DHB matrix solution and spotted onto a µFocus MALDI plate at 30 s intervals (see Step 28 of PROCEDURE).

Data analysis of nonglycopeptides ● TIMING Variable

- 23| Search the MS/MS spectra using the X!Tandem software or equivalent database-searching programs against a nonredundant protein database (such as International Protein Index) to identify the peptides and obtain the list of putative proteins in the sample. Set search parameters to include cysteine acrylamidation and methionine oxidation as modifications.
- 24| (Optional) For Mascot, SEQUEST and X!Tandem database search results, statistically analyze the database search results with PeptideProphet and ProteinProphet⁴⁷. Peptides with a PeptideProphet score > 0.2 are grouped by protein sequences using the ProteinProphet. Obtain the list of proteins in which protein probability is not lower than the threshold at an error rate of 5%. See

Supplementary Table 2 for an example of ESI-LTQFT MS/MS analysis for one of the plasma 2D HPLC fractions.

- 25| Prepare a text file of the list of putative proteins as the refinement protein database. Each protein ID must start on a new line. This file will be used for glycopeptide analysis from the data acquired by offline RPLC-MALDI-DIT MS/MS. (See 'RefinementDB.txt' and 'SimplestDB.txt' in the IGAP folder (formed by unzipping Supplementary Program, Supplementary Data 1 and Data 2) for examples.)

LC-MS/MS analysis: LC-MALDI-DIT MS/MS analysis of glycopeptides ● TIMING 5–24 h

- 26| Dry the glycopeptide sample (see Step 19) by removing the eluting solution with the SpeedVac System. Resuspend the glycopeptides in 50 μl of RP mobile phase A (0.1% formic acid in HPLC grade water (vol/vol)).
- 27| Add 20 μl of the glycopeptide sample onto a 20 μl of sample loop. The sample is trapped onto the trap column and desalted with 2% of RP mobile phase B (0.1% formic acid in acetonitrile (vol/vol)) for 10 min at 5 $\mu\text{l min}^{-1}$.
- 28| Start the RPLC gradient and separate the sample with a monolithic column at 1.0 $\mu\text{l min}^{-1}$ (see Box 4 for a recommended RP gradient program). The separated glycopeptides are coaxially mixed online with DHB matrix solution (flow rate, 1.0 $\mu\text{l min}^{-1}$) and automatically spotted onto a μFocus MALDI plate at 30 s intervals.
- 29| Apply offline LC-MALDI-DIT-MS/MS for glycopeptide analysis. The MALDI-DIT mass spectrometer is equipped with a nitrogen laser (337 nm), and argon is used as the collision gas. In MS/MS, precursors with the target mass are isolated by digital asymmetric wave isolation⁴⁸ and fragmented by collision-induced dissociation (CID) in the ion trap. Set MALDI-DIT instrument parameters as: laser power of 60 (arbitrary units; this value is a measure of the change of the transmittance of the laser) and excitation voltage of 0.25 V for CID MS/MS. See Figure 8 for the determination of the excitation voltage for CID MS/MS of transferrin glycopeptides (i.e., **Gp-1** and **Gp-2**). In general, glycan composition information is retrieved from the glycopeptide fragment ions located between [pep-HexNAc]⁺ and [M-2Sia + H]⁺, and peptide sequence information is deduced from peptide fragment ions located below [pep-HexNAc]⁺ ion (Fig. 8). To acquire a rich data set for both glycan composition and peptide sequence, we recommend the excitation voltage of 0.25 V for CID MS/MS.
- 30| Consider deamination of N-terminal glutamine for *N*-peptide when generating a potential *N*-glycopeptide. Protonated peptides with glutamine as the N-terminal amino acid are known to convert the N-terminal glutamine residue into the cyclic pyroglutamyl residue (2-pyrrolidinone-5-carboxylic acid) during CID MS/MS⁴⁹. To determine if glycopeptides with glutamine as the N-terminal amino acid encounter neutral loss by either deamination (–17 Da) or dehydration (–18 Da) during CID MS/MS, we analyzed transferrin glycopeptide (m/z 4165.95, Q*QQHLFGSN(5Hex4HexNAc)VTDCSGNFCLFR) by MALDI-

DIT under different excitation voltages for CID MS/MS, and compared the peak intensities between deamination ions and nondeamination ions for both the intact glycopeptide ($Q^*QQHLFGSN(5Hex4HexNAc)VTDCSGNFCLFR$) and the inner core *N*-acetylglucosamine glycopeptide ($Q^*QQHLFGSN(HexNAc)VTDCSGNFCLFR$) to determine the extent of neutral losses of ammonia (NH_3) from N-terminal glutamine (Q). Partial deamination occurs with lower excitation voltage (< 0.2 V) (Table 1). However, deamination becomes predominant with higher excitation voltage (> 0.2 V). We developed an in-house-built IGAP to analyze MALDI-DIT MS/MS data, and found it necessary to consider the deamination caused by neutral loss of NH_3 (Fig. 9) to elucidate glycopeptide sequences. (Software available in Supplementary Program, Supplementary Data 1 and Data 2.) A high-confidence identification is obtained when deamination (-17 Da) is assigned to all the experimental fragment ions located within the glycan information-enriched region (i.e., m/z 2,400–4,200). This indicates that glycopeptides with glutamine as the N-terminal amino acid encounter significant deamination (-17 Da) during CID MS/MS, which should be taken into account in data analysis. We also found some minor neutral loss by dehydration (-18 Da), likely due to the fact that MALDI generates mostly singly charged ions (positive mode) and the proton is bound predominantly to the basic lysine or arginine residue at the C terminus and has very small ‘residence time’ at peptide bonds and N terminus⁵⁰. The N-terminal glutamine undergoes a NH_3 elimination reaction to form the stable five-member cyclic pyroglutamyl ring (Fig. 9). For a singly charged tryptic glycopeptide, there is less chance to have the N-terminal glutamine protonated. Therefore, dehydration of the N-terminal glutamine is unlikely because of the requirement for charge delocalization between the two nitrogens of N-terminal glutamine to protonate the side-chain amide group and maintain the stability of the formed five-membered cyclic ring.

Data analysis of glycopeptides ● TIMING Variable

- 31| Prepare the *N*-glycan database. The *N*-glycan database is retrieved from the Consortium for Functional Glycomics Glycan Structures Database (<http://www.functionalglycomics.org/glycomics/molecule/jsp/carbohydrate/structure/searchThisStructure.jsp?lincode=Ma3%28Ma6%29Mb4GNb4GN>). As of August 2010, 3,406 *N*-linked glycan records are returned by searching the glycan structures database with the *N*-linked ‘core’ substructure containing three mannose and two *N*-acetylglucosamine residues. The database is cleaned up by removing entries containing xylose (Xyl; e.g., carbNlink_20046_D000), glycolylneuraminic acid (NeuGc; e.g., carbNlink_20147_D000), any modifications unrecognized by IGAP (e.g., [ME] in carbNlink_20011_D000), or structures with sugar or branches without attachment (e.g., carbNlink_21373_P). The remaining entries with the same structure (as indicated by the same IUPAC representation) are collapsed into a single entry (e.g., carbNlink_40082_D000’s structure is represented by seven other IDs), resulting in 1,563 entries. Finally, these entries are further reduced by considering their subfamily classifications

(e.g., high mannose, complex and hybrid) and composition, giving 644 unique records. A copy of this database is bundled with the IGAP download (see EQUIPMENT SETUP). A program is provided to download the latest entries. At the command prompt, run the following command:

```
IGAPUtilities.exe --download --parse -fragmentdb
```

- 32| DIT Console saves acquired MALDI-DIT MS/MS spectra in both its native format (.dit) and the open source format (.mzXML). Use the .mzXML files for searching with IGAP.
- 33| Create a file named 'Input.txt' in the IGAP directory. This file specifies the spectra to be searched, the name of the refinement protein database and the name of the protein database. Below is a sample input file with key configuration settings and notes on each line. (See 'Input.txt' in IGAP download for other possible settings.)

spectra (full pathname to spectrum file or folder of spectrum files)	Specify the full path to a single mzXML/Mascot generic format (MGF) file or the full path to the folder containing mzXML and/or MGF file(s) to be searched
protease (trypsin or other protease used)	Specify the protease used to digest the sample (e.g., trypsin; see 'Input.txt' in IGAP download for list of supported options)
db (protein sequence database)	Specify the name of the protein sequence database in FASTA format
refinedb (refinement protein database)	Specify the name of the refinement protein database created as described in Step 25.
organism (sample's organism type)	Specify the sample's organism (valid options are human, mouse, all)

▲ CRITICAL STEP The protein database specified must be the exact same protein sequence database used to establish the refinement protein database.

- 34| Launch the search. From the command prompt, run the following command:
- ```
IGAP.exe -i Input.txt -o Output.txt
```

As it runs, IGAP prints out the number of spectra processed and the time taken. The search results are stored in the tab-delimited file 'Output.txt' and the matched peaks annotation in 'Output.txt.annotation'.

### Manual data validation for glycopeptide analysis ● TIMING Variable

- 35| View the results ('Output.txt') in Excel. The results are ranked by glycan moiety score, glycopeptide probability of random match and glycopeptide score. The top 20 glycan composition results are reported. Filter with glycopeptide probability 0.05. (A lower value increases the probability of eliminating a true positive.)

### ? TROUBLESHOOTING

- 36|** Validate the filtered results manually by isotopic peak checking. Manually view the precursor ion in the (MS) mass spectrum using DITViewer to see whether the theoretical isotopic peak agrees with the experimentally identified one. For example, if the measured mass of the glycopeptide is 2,711, and its calculated mass (M) is 2,710, check the peak in the corresponding MS spectrum using DITViewer to confirm that the measured peak is indeed the M + 1 peak. The M + 1 peak is generated by a glycopeptide containing one  $^{13}\text{C}$  (see Supplementary Fig. 2). For glycopeptides with higher mass, the measured mass may be the M + 2 peak. Remove these glycopeptides as false positives if the measured peaks cannot be confirmed as M + 1 peak (or M + 2 peak for higher mass).
- 37|** Check the ion fragmentation (MS/MS) spectrum for *N*-glycopeptides. With DITViewer, visually check that the triplet peaks of the  $\text{X}^{0,2}$  (Cross-ring cleavage of core GlcNAc) of *N*-glycan with the intact peptide backbone are among the tallest peaks. Further visually check that the peaks to the right of the triplet peaks are predominantly sugar mass differences.

### ? TROUBLESHOOTING

- 38|** Check for the presence of glycan forms (with additional monosaccharides) related to this glycopeptide in the MS spectrum. Open the corresponding MS spectrum in DITViewer and visually check for higher  $m/z$  peaks in the ladder next to the precursor with a mass difference of a monosaccharide (such as sialic acid (291), hexose (162) and *N*-acetylhexosamine (hexNAc) (203)) or disaccharide (such as hexose + hexNAc (365)). Include these glycans as part of the glycan moiety.

**▲ CRITICAL STEP** The sialic acid-terminated glycopeptides readily lose sialic acid (see Fig. 10). For this reason, when the sialylated glycopeptides were selected as the precursors for MS/MS analysis, the predominant fragment ions were those ions losing sialic acids. The intensities of other partial glycopeptide ions used for deducing the glycan composition and of partial peptide ions used for deducing the amino acid sequence are usually relatively low (Fig. 10d–f). Therefore, the glycopeptide precursors with loss of sialic acid were selected as the precursors for MS/MS analysis to enhance the efficiency of glycosidic and peptide fragmentations (Fig. 10b,c,g). For sialylated glycopeptides, the number of sialic acids is retrieved from MS measurement and is added to the final glycan composition based on glycopeptide analysis from MS/MS.

### ? TROUBLESHOOTING

Troubleshooting advice can be found in Table 2.

#### ● TIMING

Steps 1–6, Immunodepletion chromatography: 50 min

Steps 7–12, 2D HPLC protein separation: 9.5 h

Steps 13 and 14, Protein in-solution digestion: 5.5 h

Steps 15–19, Glycopeptide enrichment: 2 h (at room temperature)

Steps 20–22, LC-ESI-LTQFT MS/MS of nonglycopeptide: 2.5 h

Steps 23–25, Data analysis of nonglycopeptides: variable, depending on the number of spectra, database size and the search engine used (usually not more than 2 h)

Steps 26–30, LC-MALDI-DIT MS/MS analysis of glycopeptides: 5–24 h

Steps 31–34, Data analysis of glycopeptides: variable, depending on the number of spectra, refinement database size and the number of results passing the filtering (usually a few minutes per spectrum)

Steps 35–38, Manual data validation for glycopeptide analysis: variable, depending on the number of results passing the filtering (usually a few minutes per spectrum)

## ANTICIPATED RESULTS

See Supplementary Table 3 for an example of MALDI-DIT MS/MS analysis of the same plasma 2D HPLC fraction analyzed by ESI-LTQFT MS/MS (See Supplementary Table 2 in Step 24). Our findings indicate that some peptides exhibit heterogeneity in their glycan attachments (see Supplementary Table 3). For example, both 3Hex2GlcNAc1Fuc and 5Hex4GlcNAc are attached to the same complement factor I (CFI) peptide FLNN\*GTCTAEGK, whereas both 3Hex2GlcNAc1Fuc and 5Hex4GlcNAc are attached to a C4B-binding protein beta chain (C4BPB) peptide (EWDN\*TTTECR). Glycosylation microheterogeneity at single glycosylation sites is further demonstrated with six different glycans attached to the same  $\alpha$ -1-acid glycoprotein (ORM2) peptide (NEEYN\*K) (see Fig. 11).

The protocol above provides for an integrated mass spectrometry–based identification of glycopeptides in complex mixtures with simultaneous elucidation of peptide sequences, glycosylation sites, and oligosaccharide compositions. Glycopeptide and glycoprotein analysis has relevance to our understanding of the role and extent of heterogeneity of proteins (on the basis of their glycan modifications) in health and in disease. The search for potential biomarkers representing modified glycoproteins in the blood would particularly benefit from this approach. To our knowledge, the protocol presented here breaks new ground in developing methodology to analyze intact glycoproteins by integrating LC-ESI-MS/MS and LC-MALDI-MS/MS to yield peptide sequence identities, sites of glycan modifications and the composition of glycans. The methodology presented here can be applied to any complex biological samples that contain glycoproteins, including tissue and cellular extracts. However, the immunodepletion step in the sample preparation procedure is only relevant to plasma and some other biological fluids.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.



## References

1. Apweiler R, Hermjakob H, Sharon N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim Biophys Acta*. 1999; 1473:4–8. [PubMed: 10580125]
2. Wong CH. Protein glycosylation: new challenges and opportunities. *J Org Chem*. 2005; 70:4219–4225. [PubMed: 15903293]
3. Cantagrel V, Lefeber DJ, Ng BG, Guan Z, Silhavy JL, Bielas SL, et al. SRD5A3 is required for converting polyprenol to dolichol and is mutated in a congenital glycosylation disorder. *Cell*. 2010; 142:203–217. [PubMed: 20637498]
4. Montpetit ML, Stocker PJ, Schwetz TA, Harper JM, Norring SA, Schaffer L, et al. Regulated and aberrant glycosylation modulate cardiac electrical signaling. *Proc Natl Acad Sci USA*. 2009; 106:16517–16522. [PubMed: 19666501]
5. Ohtsubo K, Marth JD. Glycosylation in cellular mechanisms of health and disease. *Cell*. 2006; 126:855–867. [PubMed: 16959566]
6. Taylor AD, Hancock WS, Hincapie M, Taniguchi N, Hanash SM. Towards an integrated proteomic and glycomic approach to finding cancer biomarkers. *Genome Med*. 2009; 1:57. [PubMed: 19519948]
7. Peracaula R, et al. Altered glycosylation pattern allows the distinction between prostate-specific antigen (PSA) from normal and tumor origins. *Glycobiology*. 2003; 13:457–470. [PubMed: 12626390]
8. Leiserowitz GS, et al. Glycomics analysis of serum: a potential new biomarker for ovarian cancer? *Int J Gynecol Cancer*. 2008; 18:470–475. [PubMed: 17655680]
9. Hakomori S. Glycosylation defining cancer malignancy: new wine in an old bottle. *Proc Natl Acad Sci USA*. 2002; 99:10231–10233. [PubMed: 12149519]
10. Kyselova Z, et al. Breast cancer diagnosis and prognosis through quantitative measurements of serum glycan profiles. *Clin Chem*. 2008; 54:1166–1175. [PubMed: 18487288]
11. Soltermann A, et al. N-glycoprotein profiling of lung adenocarcinoma pleural effusions by shotgun proteomics. *Cancer*. 2008; 114:124–133. [PubMed: 18327805]
12. Peracaula R, Barrabés S, Sarrats A, Rudd PM, de Llorens R. Altered glycosylation in tumours focused to cancer diagnosis. *Dis Markers*. 2008; 25:207–218. [PubMed: 19126965]
13. Gavel Y, Von Heijne G. Sequence difference between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering. *Protein Eng*. 1990; 3:433–442. [PubMed: 2349213]
14. Hansen JE, et al. Prediction of O-glycosylation of mammalian proteins: specificity patterns of UDP-GalNAc: polypeptide N-acetylgalactosaminyl-transferase. *Biochem J*. 1995; 308:801–81. [PubMed: 8948436]
15. Wührer M, Catalina MI, Deelder AM, Hokke CH. Glycoproteomics based on tandem mass spectrometry of glycopeptides. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2007; 849:115–128.
16. Geyer H, Geyer R. Strategies for analysis of glycoprotein glycosylation. *Biochim Biophys Acta*. 2006; 1764:1853–1869. [PubMed: 17134948]
17. Zhang H, Li XJ, Martin DB, Aebersold R. Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat Biotechnol*. 2003; 21:660–666. [PubMed: 12754519]
18. Kaji H, et al. Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify N-linked glycoproteins. *Nat Biotechnol*. 2003; 21:667–672. [PubMed: 12754521]
19. Wada Y, et al. Comparison of the methods for profiling glycoprotein glycans—HUPO Human Disease Glycomics/Proteome Initiative multi-institutional study. *Glycobiology*. 2007; 17:411–422. [PubMed: 17223647]
20. Wührer M, de Boer AR, Deelder AM. Structural glycomics using hydrophilic interaction chromatography (HILIC) with mass spectrometry. *Mass Spectrom Rev*. 2009; 28:192–206. [PubMed: 18979527]

21. Kirmiz C, et al. A serum glycomics approach to breast cancer biomarkers. *Mol Cell Proteomics*. 2007; 6:43–55. [PubMed: 16847285]
22. Li B, et al. Glycoproteomic analyses of ovarian cancer cell lines and sera from ovarian cancer patients show distinct glycosylation changes in individual proteins. *J Proteome Res*. 2008; 7:3776–3788. [PubMed: 18642944]
23. Abd Hamid UM, et al. A strategy to reveal potential glycan markers from serum glycoproteins associated with breast cancer progression. *Glycobiology*. 2008; 18:1105–1118. [PubMed: 18818422]
24. Wang H, Hanash SM. Increased throughput and reduced carryover of mass spectrometry-based proteomics using a high-efficiency nonsplit nanoflow parallel dual-column capillary HPLC system. *J Proteome Res*. 2008; 7:2743–2755. [PubMed: 18512973]
25. Faca VM, et al. A mouse to human search for plasma proteome changes associated with pancreatic tumor development. *PLoS Med*. 2008; 5:e123. [PubMed: 18547137]
26. Paczesny S, et al. Elafin is a biomarker of graft-versus-host disease of the skin. *Sci Transl Med*. 2010; 2:13ra2.
27. Fukuyama Y, Nakaya S, Yamazaki Y, Tanaka K. Ionic liquid matrixes optimized for MALDI-MS of sulfated/sialylated/neutral oligosaccharides and glycopeptides. *Anal Chem*. 2008; 80:2171–2179. [PubMed: 18275166]
28. Wang H, et al. Intact-protein-based high-resolution three-dimensional quantitative analysis system for proteome profiling of biological fluids. *Mol Cell Proteomics*. 2005; 4:618–625. [PubMed: 15703445]
29. Wang H, Hanash S. Electrospray mass spectrometry for quantitative plasma proteome analysis. *Methods Mol Biol*. 2009; 564:227–242. [PubMed: 19544026]
30. Freiwald A, Sauer S. Phylogenetic classification and identification of bacteria by mass spectrometry. *Nat Protoc*. 2009; 4:732–742. [PubMed: 19390529]
31. Xu T, Wong CC, Kashina A, Yates JR III. Identification of N-terminally arginylated proteins and peptides by mass spectrometry. *Nat Protoc*. 2009; 4:325–332. [PubMed: 19229197]
32. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics*. 2002; 1:845–867. [PubMed: 12488461]
33. Liu T, et al. Evaluation of multi-protein immunoaffinity subtraction for plasma proteomics and candidate biomarker discovery using mass spectrometry. *Mol Cell Proteomics*. 2006; 5:2167–2174. [PubMed: 16854842]
34. Polaskova V, Kapur A, Khan A, Molloy MP, Baker MS. High-abundance protein depletion: comparison of methods for human plasma biomarker discovery. *Electrophoresis*. 2010; 31:471–482. [PubMed: 20119956]
35. Wada Y, Tajiri M, Yoshida S. Hydrophilic affinity isolation and MALDI multiple-stage tandem mass spectrometry of glycopeptides for glycoproteomics. *Anal Chem*. 2004; 76:6560–6565. [PubMed: 15538777]
36. Sparbier K, Koch S, Kessler I, Wenzel T, Kostrzewa M. Selective isolation of glycoproteins and glycopeptides for MALDI-TOF MS detection supported by magnetic particles. *J Biomol Tech*. 2005; 16:407–413. [PubMed: 16522863]
37. Pfenninger A, Karas M, Finke B, Stahl B, Sawatzki G. Matrix optimization for matrix-assisted laser desorption/ionization mass spectrometry of oligosaccharides from human milk. *J Mass Spectrom*. 1999; 34:98–104. [PubMed: 10093211]
38. Steven L, Cohen SL, Chait BT. Influence of matrix solution conditions on the MALDI-MS analysis of peptides and proteins. *Anal Chem*. 1996; 68:31–37. [PubMed: 8779435]
39. Ding L, Sudakov M, Brancia FL, Giles R, Kumashiro S. A digital ion trap mass spectrometer coupled with atmospheric pressure ion sources. *J Mass Spectrom*. 2004; 39:471–484. [PubMed: 15170743]
40. Craig R, Beavis RC. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics*. 2004; 20:1466–1467. [PubMed: 14976030]
41. Kersey PJ, et al. The International Protein Index: an integrated database for proteomics experiments. *Proteomics*. 2004; 4:1985–1988. [PubMed: 15221759]

42. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*. 1999; 20:3551–3567. [PubMed: 10612281]
43. Eng JK, McCormack AL, Yates JR III. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom*. 1994; 5:976–989. [PubMed: 24226387]
44. Na S, Paek E. Quality assessment of tandem mass spectra based on cumulative intensity normalization. *J Proteome Res*. 2006; 5:3241–3248. [PubMed: 17137325]
45. Beausoleil SA, Villén J, Gerber SA, Rush J, Gygi SP. A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat Biotechnol*. 2006; 24:1285–1292. [PubMed: 16964243]
46. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*. 1976; 748:248–254. [PubMed: 942051]
47. Nesvizhskii AI, Keller A, Kolker E, Aebersold RA. Statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem*. 2003; 75:4646–4658. [PubMed: 14632076]
48. Ding, L.; Brancia, FL.; Giles, R.; Smirnov, S.; Nikolaev, E. Advances of Tandem MS Functions with a Digital Ion Trap. *Proc. 53rd ASMS Conf*; 5–9 June 2005; San Antonio, Texas.
49. Baldwin MA, et al. Tandem mass spectrometry of peptides with N-terminal glutamine: Studies on a prion protein peptide. *J Am Soc Mass Spectrom*. 1990; 1:258–264.
50. Neta P, Pu QL, Kilpatrick L, Yang X, Stein SE. Dehydration versus deamination of N-terminal glutamine in collision-induced dissociation of protonated peptides. *J Am Soc Mass Spectrom*. 2007; 18:27–36. [PubMed: 17005415]

**Box 1****RECOMMENDED IMMUNODEPLETION CHROMATOGRAPHY PROGRAM  
(UV = 280 NM)**

Add 1.5 ml of buffer A to 0.3 ml of plasma, vortex the sample thoroughly and filter the sample with a 0.22- $\mu$ m RC syringe filter. Load the sample to a 2-ml sample loop and perform immunodepletion chromatography using the recommended conditions below:

**Column equilibration**

- 10 min of 100% buffer A, 3 ml min<sup>-1</sup>

**Step elution**

- 0.00 min Inject sample, 100% buffer A, 0.5 ml min<sup>-1</sup>
- 6.30 min Start to collect flow-through fraction (low-abundance proteins): 100% Buffer A, 0.5 ml min<sup>-1</sup>
- 25.0 min 100% Buffer A, 0.5 ml min<sup>-1</sup>
- 25.1 min 100% Buffer B, 0.5 ml min<sup>-1</sup>
- 26.3 min Stop to collect flow-through fraction: 100% buffer B, 0.5 ml min<sup>-1</sup>
- 27.0 min 100% Buffer B, 3.0 ml min<sup>-1</sup>
- 29.0 min Start to collect bound fraction (high-abundance proteins): 100% Buffer B, 3.0 ml min<sup>-1</sup>
- 33.0 min Stop to collect bound fraction: 100% buffer B, 3.0 ml min<sup>-1</sup>
- 35.0 min 100% Buffer B, 3.0 ml min<sup>-1</sup>
- 35.1 min 100% Buffer A, 3.0 ml min<sup>-1</sup>
- 45.0 min 100% Buffer A, 3.0 ml min<sup>-1</sup>

Leave the collected flow-through fractions and bound fractions on dry ice or keep them at -80 °C until use.

**Figure 1.**

A general overview of the procedure for profiling human plasma glycoproteomics. A plasma sample is subjected to an immunodepletion chromatography followed by 2D HPLC fractionation. The digested 2D HPLC fractions are then analyzed by LC-MS/MS. Pattern-matching analysis software of mass spectra allows the identification of glycoproteins and their glycan attachments.

**Figure 2.**

Chromatograms for plasma immunodepletion and 2D HPLC chromatography. **(a)** Immunodepletion chromatogram of 300  $\mu$ l of human plasma (see Box 1). **(b)** 2D HPLC chromatograms (2D RP chromatography). There is an eight-step elution for 1D AEX (see Box 2); protein fractions from each of the eight steps are further separated by 2D RP chromatography with a gradient elution (see Box 2).

**Figure 3.**

Comparison of on-plate glycopeptide enrichment with two types of MALDI plates. **Gp-1** and **Gp-2** are transferrin glycopeptides enriched from tryptic digests with Sepharose CL-4B, and subjected to RPLC separation and online spotting onto the MALDI plate using 5 mg ml<sup>-1</sup> DHB as the matrix. (a)  $\mu$ Focus plate. (b) Normal metal plate. **Gp-1**: CGLVPVLAENYn\*(5Hex4HexN Ac2Sia)K; **Gp-2**: QQQHLFGSN\*(5Hex4HexNAc2Sia)VTDCSGNFCLFR.

**Figure 4.**

Distribution of crystallized **Gp-1** and **Gp-2** glycopeptides in a  $\mu$ Focus plate across the spotted area. The locations where the glycopeptides are distributed in the spot on the plate are indicated by Outside, Middle, Inside, and Center. **Gp-1**:

CGLVPVLAENYN\*(5Hex4HexNA c2Sia)K; **Gp-2**: QQQHLFGSN\*(5Hex4HexNAc2Sia)

VTDCSGNFCLFR.

**Figure 5.** MALDI-DIT mass spectra of transferrin glycopeptides **Gp-1** and **Gp-2** at four different positions (outside, middle, inside, center) in the  $\mu$ Focus plate. **Gp-1**: CGLVPVLAENYN\*(5Hex4HexNAc2Sia)K; **Gp-2**: QQQHLFGSN\*(5Hex4HexNAc2Sia)VTDCSGNFCLFR.

**Figure 6.**

Data analysis overview. **(a)** The refinement protein database is created from the protein IDs confidently identified with the LC-ESI-MS/MS data. Each ID is recorded on a separate line. This database is experiment specific. Protein identification may be performed using other search engines. **(b)** The N-glycan database is constructed with the data from Consortium for Functional Glycomics (CFG) Glycan Structures Database. This database only needs to be updated when new entries are found at CFG. **(c)** MALDI-DIT MS/MS spectrum is searched by IGAP, which generates the top 20 glycan composition results in the tab-delimited file 'Output.txt' and the matched peaks annotation in 'Output.txt.annotation'. Manual validation of the results is performed. DITViewer provides visual access to the acquired spectra.

**Figure 7.**

Glycopeptide MS/MS ions search process. **(a)** y-ion series and  $X^{0,2}$  (squares with diagonal lines) of *N*-glycan with intact peptide backbones (peptide sequence NEEYNK) are calculated. **(b)** Both b-ions ( $b_1$ – $b_5$ ) and y-ions ( $y_1$ – $y_5$ ) series of *N*-peptide (peptide NEEYNK) along with an attached GlcNAc (*N*-acetylglucosamine; shown as white squares) at the glycosylation site are calculated. **(c)** The ions are merged to generate the theoretical spectrum. **(d)** Experimental spectrum is cumulative intensity–normalized and matched with the theoretical spectrum. Each match produces a paired metric of score and probability of random match for both glycan moiety (ions in **a**) and glycopeptide (ions from both **a** and **b**).

**Figure 8.**

Effect of collision energy on collision-induced dissociation MS/MS analysis of transferrin glycopeptides **Gp-1** and **Gp-2**. Blue squares represent GlcNAc. Blue squares with “-pep” represent the peptide with its core GlcNAc.  $[M-2Sia + H]^+$  represents the protonated intact glycopeptide with loss of two sialic acids **Gp-1**:

CGLVPVLAENY**N**\*(5Hex4HexNAc2Sia)K; **Gp-2**:

QQQHLFGS**N**\*(5Hex4HexNAc2Sia)VTDCSGNFCLFR. CID, collision-induced dissociation.



**Figure 9.**

Deamination of a glycopeptide with N-terminal glutamine with MALDI-DIT collision-induced dissociation MS/MS. C\*, cysteine with heavy label; D, aspartic acid; P-17, peptide with loss of 17 Da; T, threonine; V, valine. Blue square, GlcNAc (N-acetyl-glucosamine); Green circle, Man (mannose); yellow circle, Gal (galactose).

**Figure 10.**

Precursor selection for MS/MS analysis with MALDI-DIT. **(a)** MALDI-DIT MS (MS-1) spectrum of glycopeptides. **(b–e)** MALDI-DIT MS/MS spectra of selected glycopeptide precursors. Sia, sialic acid.

**Figure 11.** Glycosylation heterogeneity at a single site in  $\alpha$ -1-acid glycoprotein (encoded by ORM2). \*, heavy isotope labeling.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**TABLE 1**  
MALDI-DIT MS/MS deamination of a glycopeptide with an N-terminal glutamine as a function of collision energy.

| Collision energy (V) | Peak intensities for glycopeptide |       |                                  | Peak intensities for GlcNAc |      |                                  |
|----------------------|-----------------------------------|-------|----------------------------------|-----------------------------|------|----------------------------------|
|                      | P                                 | P-17  | Ratio of peak intensity (P-17)/P | P                           | P-17 | Ratio of peak intensity (P-17)/P |
| 0.15                 | 0.35                              | 0.08  | 0.23                             | 0.05                        | 0.07 | 1.4                              |
| 0.17                 | 0.04                              | 0.09  | 2.25                             | 0.01                        | 0.11 | 11                               |
| 0.20                 | 0.001                             | 0.02  | 20                               | 0.003                       | 0.09 | 30                               |
| 0.25                 | 0.0008                            | 0.018 | 22.5                             | 0.002                       | 0.05 | 25                               |
| 0.30                 | 0.0003                            | 0.028 | 93.3                             | 0.002                       | 0.12 | 60                               |

P: QQQLFGSN\*(5Hex4HexNAc2Sia)VTDCSGNFCLFR; P-17: QQQLFGSN\*(5Hex4HexNAc2Sia)VTDCSGNFCLFR – 17.

TABLE 2

Troubleshooting table.

| Step | Problem                                                                                                            | Possible reason                                                                                                                 | Solution                                                                                                                                                                                                                                                                                                                                                                               |
|------|--------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 3    | It is difficult to filter the plasma sample                                                                        | The syringe filter is clogged or the sample concentration is too high                                                           | Mix plasma sample with immunodepletion buffer A (plasma/buffer A, 1:5, vol/vol), then centrifuge the sample at 15,000g for 10 min at 4 °C. Filter the supernatants through the syringe filter                                                                                                                                                                                          |
| 6    | Efficiency of the immunodepletion is low                                                                           | The immunodepletion column is deteriorated                                                                                      | Replace with a new immunodepletion column                                                                                                                                                                                                                                                                                                                                              |
| 10   | It is difficult to filter the alkylated plasma sample                                                              | The sample becomes more viscous after protein alkylation                                                                        | Thoroughly vortex the alkylated sample and change to a new filter                                                                                                                                                                                                                                                                                                                      |
|      | Pressure in the 2D HPLC system is building up                                                                      | The inlet filters and AEX or RP columns may be clogged                                                                          | Replace the affected inlet filter or AEX/RP column if necessary                                                                                                                                                                                                                                                                                                                        |
|      | Pressure in the 2D HPLC system is fluctuating, reduced or completely lost                                          | There may be an air bubble in the pump head or a leak in the 2D HPLC system, detached fittings, or a clogged column             | Purge the pump to expel air bubbles. Check for loose fittings. Retighten all connections. Check for clogging of inlet filters or columns                                                                                                                                                                                                                                               |
|      | Retention times shift                                                                                              | Changing mobile-phase composition                                                                                               | Replace mobile phases                                                                                                                                                                                                                                                                                                                                                                  |
| 22   | High back pressure in the capillary C18 column                                                                     | Fragmentation of packing materials                                                                                              | Repack the column                                                                                                                                                                                                                                                                                                                                                                      |
|      | Unstable spray                                                                                                     | Spray voltage set too low or too high                                                                                           | Optimize the spray voltage                                                                                                                                                                                                                                                                                                                                                             |
| 35   | Few results pass the filtering of glycopeptide probability 0.05                                                    | Low-abundance glycopeptides can result in noisy spectra                                                                         | Confirm that there is a low-information content spectrum (i.e., spectrum with few or no signal peaks) by performing Step 33 on some or all spectra without filtering. Reacquire data from other spots in the same well or acquire data from a fresh sample plate                                                                                                                       |
|      | Expected glycopeptides are not in the results                                                                      | Glycoprotein ID is not present in refinement protein database                                                                   | Indistinguishable proteins identified during the refinement database construction should be included when performing Step 23. Protein that is not confidently identified for inclusion in the refinement protein database should be excluded. Include the protein ID when performing Step 25 and repeat the process if you can demonstrate the presence of the protein by other assays |
|      |                                                                                                                    | Glycan is not present in glycan database                                                                                        | Contact author for glycan database update                                                                                                                                                                                                                                                                                                                                              |
|      | Mass spectrum contains a mixture of ions that appears too complex for one glycopeptide                             | In some cases, co-eluting glycopeptides and peptides may have very similar $m/z$ values and appear in the same isolation window | Check whether at least one of the measured masses in the isolation window matches the calculated mass with the correct isotopic peak with an acceptable delta mass. In this case, the glycopeptide is considered to be a putative true hit                                                                                                                                             |
| 37   | The triplet peaks of the $X^{0.2}$ of $N$ -glycan with the intact peptide backbone are not among the tallest peaks | Insufficient fragmentation                                                                                                      | Check if the peaks in the region closer to the precursor have much higher intensities. Reacquire data if this is the case                                                                                                                                                                                                                                                              |