Published in final edited form as:

Mol Biosyst. 2012 November 2; 8(11): 2850–2856. doi:10.1039/c2mb25268f.

Large-scale quantitative glycoproteomics analysis of site-specific glycosylation occupancy†

Sheng Pan^a, Yasuko Tamura^a, Ru Chen^a, Damon May^b, Martin W. McIntosh^b, and Teresa A. Brentnall^a

Sheng Pan: shengp@medicine.washington.edu

^aDepartment of Medicine, University of Washington, Seattle, WA 98195, USA

^bFred Hutchinson Cancer Research Center, Molecular Diagnostics Program, Seattle, WA 98109, USA

Abstract

Disease-associated aberrant glycosylation may be protein specific and glycosylation site specific. Quantitative assessment of glycosylation changes at a site-specific molecular level may represent one of the initial steps for systematically revealing the glycosylation abnormalities associated with a disease or biological state. Comparative quantitative profiling of glycoproteome to provide accurate quantification of site-specific glycosylation occupancy has been a challenging task, requiring a concerted approach drawing from a variety of techniques. In this report, we present a quantitative glycoproteomics method that allows global scale identification and comparative quantification of glycosylation site occupancy using mass spectrometry. We further demonstrated this approach by quantitatively characterizing the N-glycoproteome of human pancreas.

Introduction

Protein glycosylation is one of the most common forms of post-translational modification and plays an important role in many biological processes and pathways. Diseases such as cancer can induce changes of protein glycosylation resulting in altered glycan structures and abnormal glycosylation occupancy. The importance of identifying these aberrant changes is underscored by the fact that many cancer biomarkers are glycoproteins.^{1–3} In addition, non-diseased states such as the regulation of proteins that govern T cell immunologic function and embryonic neurologic development are regulated through glycosylation.^{4,5} While a mechanistic understanding of the implications of glycosylation pathways in many biological systems remains to be resolved, glycosylation has been recognized as an important molecular feature in diseases such as cancer.

Disease-associated aberrant glycosylation may be protein-specific and glycosylation site-specific. Thus, quantitative assessment of glycosylation changes at a site-specific molecular level from either a protein or a glycan perspective may represent one of the initial steps for systematically revealing the glycosylation changes and abnormalities associated with a disease. One focus in current glycoproteomics has been to develop a robust and sensitive technique that affords large-scale quantitative profiling of site-specific glycosylation occupancy in a complex system, *i.e.*, a quantitative assessment of a large number of glycosylation sites to determine whether a specific protein glycosylation site is hypo-, hyper-, or neo-glycosylated. Different strategies have been applied for the mass

†Electronic supplementary information (ESI) available. See DOI: 10.1039/c2mb25268f

Correspondence to: Sheng Pan, shengp@medicine.washington.edu.

spectrometry-based quantitative profiling of the glycoproteome of tissues,^{6,7} serum^{8–14} or plasma,^{15–18} and other bodily fluids.¹⁹ Some of the current strategies approach glycoprotein quantification by enriching the glycoproteome at the protein level followed by quantitative analysis. Such a strategy provides overall glycoprotein quantification but lacks the specificity to quantitatively address glycosylation occupancy density at the individual glycosylation site within a glycoprotein. Other approaches target glycopeptide for quantification, but enrich glycopeptides or glycoproteins prior to stable isotopic labeling; thus can lead to variations in quantification due to the glyco-enrichment steps. In this report, we describe a mass spectrometry-based glycoproteomics method for global quantitative profiling of glycosylation occupancy at a site-specific level. This method has been proven by our lab and others²⁰ to be robust, sensitive and cost effective, and is compatible with a variety of mass spectrometers and applicable to an assortment of sample types, including tissue, cell lysate, plasma/serum, and other bodily fluids. We characterized this method by quantitatively profiling the N-glycoproteome of human pancreas.

Results and discussion

Analytical flow

The challenge of an in-depth, quantitative analysis of a glycosylation profile in a complex biological sample, such as plasma/serum or tissue lysate, arises from not only the enormous complexity in glycoprotein constituent, but also the intricacy within the molecule of a glycoprotein. Quantitative profiling of glycoproteomes typically requires a concerted approach drawing from different techniques. As illustrated in Fig. 1, the integrated quantitative glycoproteomic approach consists of five major steps: (1) sample preparation, (2) dimethylation stable isotopic labeling for mass spectrometric quantification, (3) glycopeptide capturing, (4) high-resolution tandem mass spectrometric analysis and (5) bioinformatic data analysis for glycopeptide identification and quantification. Equal amounts of the compared samples (*e.g.* disease *versus* normal) are first subjected to trypsin digestion followed by formaldehyde-based differential dimethyl labeling,²¹ in which the compared samples are labeled with either light (H) or heavy (D) versions of formaldehyde, individually, to demark the different sample origins. The labeled samples are then combined, and the glyco-peptides are enriched using either hydrazide chemistry-based solid phase extraction^{10,22–26} or lectin affinity column^{27–33} followed by the enzymatic or chemical removal of glycans. The analysis of the glycopeptides using a high-resolution mass spectrometer allows precise mapping of individual glycosylation sites on the glycopeptides identified. For instance, N-glycosylation sites cleaved by PNGase F can be precisely mapped using the consensus sequence of Asn-X-Ser/Thr (X = any amino acid except proline) in which asparagine is converted to aspartic acid following enzyme cleavage, introducing a mass difference of 0.9840 Daltons. The identified N-glycosylation sites can be further confirmed by database annotation. The quantification of a glycopeptide carrying a specific glycosylation site can be achieved by using the intensity ratio between the heavy and light isotopic forms of the glycopeptide.

Mapping N-linked glycosylation sites in the pancreas glycoproteome

We characterized this method by analyzing the N-glycoproteome of human pancreas tissue using hydrazide chemistry-based solid phase extraction for glycopeptide capturing. Since the dimethylation occurs on the N-terminal and lysine residues of a peptide, it is expected that most of the tryptic peptides will be labeled with either the heavy or light form of the dimethyl groups for quantification, conceptually covering the whole proteome. The subsequent glycopeptide capturing process eliminates most of the non-glycopeptides, including deamidated peptides, for mass spectrometric analysis, thus significantly reduces the possibility of false identification of N-glycosylation sites due to deamidation of

asparagine. In fact, using a high-resolution Orbitrap mass spectrometer coupled with nano-liquid chromatography for analysis, more than 94% of the peptides identified with Asn-X-Ser/Thr motif/s and with a PeptideProphet³⁴ probability score > 0.95 are annotated glycopeptides. In this study, 656 unique annotated N-linked glycopeptides derived from 383 non-redundant glycoproteins were identified with a PeptideProphet probability > 0.95 (~1% false discovery rate), and all of these glycopeptides were quantifiable with a heavy and light area. Forty-five percent and 54% of the glycopeptides identified are Lys and Arg terminated, respectively, suggesting that the dimethyl labeling did not significantly impact the ionization efficiency of Lys terminated peptides. Fig. 2a displays the distribution of glycopeptides based on the deviation of their precursor mass from the theoretical value. The majority of the glycopeptides identified show a mass deviation less than 5 ppm. Fig. 2b illustrates a representative MS/MS identification of a de-glycosylated glycopeptide. The y and b ions of the de-glycosylated peptide ED*GTDTVQEEEESPAEGSK (derived from hypoxia up-regulated protein 1, HYOU1, a protein expressed in pancreas, * indicates where the glycan was attached) generated by collision-induced-dissociation (CID) were well identified by the tandem mass spectrometry. It is notable that the N (asparagine) of the Asn-X-Ser/Thr motif was converted to D (aspartic acid) due to the PNGase F enzymatic cleavage of the N-glycan from the peptide.

Our results revealed that the glycoproteins identified with one or two N-linked glycosylation sites were prevalent in the pancreas glycoproteome (Fig. 3). The Gene Ontology annotation of cellular locations and biological processes for these human pancreas-associated glycoproteins was clustered by The Database for Annotation, Visualization and Integrated Discovery (DAVID) program³⁵ (Fig. 4). Most of the N-glycoproteins identified in human pancreas are membrane or extracellular proteins involving in a variety of biological processes, including cellular organization, development, immune system response, adhesion and transport.

In addition to the known N-glycosylation sites, six potential N-linked glycosylation sites (derived from five glycoproteins) which are not currently referenced in Uniprot Knowledgebase (<http://www.uniprot.org/>) were identified in human pancreas tissue (Table S1, ESI[†]). All of these newly discovered potential N-linked glycopeptides were consistently identified in at least four out of the five independent experiments with PeptideProphet probability greater than 0.99 (<0.5% false discovery rate). The MS/MS spectra of these glycopeptides (in their de-glycosylated form) are available in the Fig. S1 (ESI[†]). These peptides with the Asn-X-Ser/Thr motif were detected and sequenced multiple times in each experiment. With the consideration of the mass change (0.9840 Da) on asparagines to aspartic acid due to PNGase F enzymatic cleavage, all the peptides listed in Table S1 (ESI[†]) have $a < 5$ ppm mass deviation from the theoretical values. To the best of our knowledge, there is no known potential *in vivo* deamidation of asparagine to aspartic acid on these corresponding asparagine sites. Moreover, these potential N-linked glycopeptides were identified *via* a hydra-zide chemistry-based solid phase extraction process followed by specific cleavage of PNGase F enzyme from the solid phase. In such a process, the majority of *non*-glycosylated peptides were washed off prior to the enzymatic cleavage and thus eliminated from MS analysis. Together, these orthogonal types of evidence strongly suggest that these Asn-X-Ser/Thr motifs are potentially novel N-linked glycosylation sites. The glycan linkages associated with these glycosylation sites warrant further investigation. Whether any of these potential N-linked glycosylation sites are organ- (pancreas-) specific remains unknown at this point. For example, pancreatic secretory granule membrane major glycoprotein GP2, a protein secreted by the human pancreas, was identified with two new

[†]Electronic supplementary information (ESI) available. See DOI: 10.1039/c2mb25268f

N-linked glycosylation motifs in this study, adding to its three known N-linked glycosylation sites.

Global-scale comparative quantification of N-glycosylation occupancy

For quantitative evaluation, two identical pancreas tissue samples with a 1 : 1 weight ratio were differentially labeled with different stable isotopes and compared. Hydrazide chemistry beads were used for glycopeptide capture. Since the glycoproteins were captured at the peptide level after protease digestion, the relative abundance ratio of a glycopeptide quantitatively reflects the absolute difference of glycosylation occupancy on the corresponding glycosylation sites between the two samples compared. Fig. 5a shows an example of mass spectrometric quantification of a formerly glycosylated peptide (ED*GTDTVQEEEESPAEGSK, * indicates where the glycan was attached) derived from hypoxia up-regulated protein 1. The isotopic light and heavy versions of the de-glycosylated peptide eluted concurrently, and their intensities at the MS1 level clearly show that this peptide's quantitative ratio between the two samples is close to 1 (heavy/light = 1.01). Fig. 5b demonstrates the distribution of the intensity ratios (in natural log scale) of the glycopeptides identified in the pancreas tissue sample, which was a combination of two differentially labeled samples with a heavy and light theoretical ratio of 1 : 1. As expected, the majority of the glycopeptides have a near 1:1 heavy-to-light quantitative ratio, showing a narrow natural log ratio distribution centered on 0. For an internal control, we used equal amounts of a non-human glycoprotein standard (yeast invertase 2) spiked into each of the two samples that were compared, to monitor the quantitative variations that could potentially be introduced during the sample preparation process. The mean heavy/light ratio of the N-glycopeptides derived from this standard glycoprotein is 1.1, reasonably close to the theoretical ratio of 1.0. Since the glycopeptides from different sample origins (differentially labeled as light and heavy) were combined and had undergone the same glycopeptide capturing process together, it significantly reduces the variation that may be introduced when individual samples are processed separately for glycopeptide capturing.

Materials and methods

Lysate preparation

The pancreas tissues were obtained from surgical specimens at the University of Washington Medical Center with the approval of the Institutional Review Board. Snap frozen tissue was homogenized in T-per (Thermo Scientific, Rockford, IL) with protease inhibitors and incubated on ice for 15 minutes. To pellet any cell debris, the lysate was spun down at 4 °C and the supernatant was transferred into a new tube. The protein concentration was measured by BCA assay (Pierce, Rockford, IL).

Peptide preparation

Each lysate sample (1000 µg) was mixed with 50 µg of glycoprotein standard (yeast invertase 2, Sigma-Aldrich, St. Louis, MO), diluted in 50 mM ammonium bicarbonate solution and reduced with dithiothreitol (DTT) then alkylated with iodoacetamide. To purify the sample, TCA precipitation was performed by adding 1/4 volume of 100% w/v trichloroacetic acid. The samples were incubated on ice for 10 min and spun down. Pellets were washed twice with ice-cold acetone and air dried before re-suspending them in 300 µl of 50 mM sodium bicarbonate solution. Each lysate was digested with sequencing grade trypsin (Promega, Madison, WI) with a 1 : 50 trypsin-to-sample ratio at 37 °C for 18 hours. The lysate digest was buffer exchanged in 100 µl of 100 mM sodium acetate, pH 5.5.

Stable isotopic labeling

Equal amounts of control and compared sample were separately labeled with formaldehyde- H_2 (light) and formaldehyde- D_2 (heavy) (Sigma-Aldrich, St. Louis, MO), respectively. To label each sample, 5 μ l of 20% labeling agent was added to 100 μ l sample, immediately followed by addition of 5 μ l of freshly prepared 3 M sodium cyanoborohydride solution. The samples were vigorously vortexed and incubated for 15 minutes at the room temperature, with vortexing every few minutes. The light-and heavy-labeled samples were combined into one and purified through C18 purification columns (the Nest Group, Inc., Southborough, MA).

Glycopeptide capturing with hydrazide beads

Peptides were re-suspended in the coupling buffer (0.1 M sodium acetate, 150 mM sodium chloride, pH 5.6) and oxidized with sodium meta-periodate at the final concentration of 10 mM for 1 hour at room temperature, in the dark, with gentle rotation. The excess sodium meta-periodate was quenched by addition of sodium sulfite at the final concentration of 20 mM for 10 min at room temperature with gentle rotation. The sample was then combined with rinsed hydrazide resin beads (Thermo Scientific, Rockford, IL) and coupled at room temperature overnight (>10 hours) with gentle rotation. After coupling, beads were washed with 80% acetonitrile/ 0.1% trifluoroacetic acid solution once, followed by 5 washes with phosphate buffered saline (PBS). Beads were re-suspended in PBS and incubated with PNGase F (New England BioLabs, Ipswich, MA) at 37 °C for 6 hours with vortexing every 30 minutes. Cleaved glycopeptides are collected by centrifuging the sample at $1000 \times g$ for 2 min and taking the supernatant. Beads were washed once with fresh 250 μ l PBS to collect any remaining glycopeptides.

Mass spectrometry analysis

An LTQ-Orbitrap hybrid mass spectrometer (Thermo Fisher Scientific, Waltham, MA) coupled with a nano-flow HPLC (Eksigent Technologies, Dublin, CA) was used in this study. A 2 μ g of sample was injected for the mass spectrometric analysis. The samples were first loaded onto a 1.5 cm trap column (IntegraFrit 100 μ m, New Objective, Woburn, MA) packed with Magic C18AQ resin (5 μ m, 200 Å particles; Michrom Bioresources, Auburn, CA) with Buffer A (water with 0.1% formic acid) at a flow rate of 3 μ L per minute. The peptide samples were then separated by a 27 cm analytical column (PicoFrit 75 μ m, New Objective) packed with Magic C18AQ resin (5 μ m, 100 Å particles; Michrom Bioresources) followed by mass spectrometric analysis. A 90 minute nonlinear LC gradient was used as follows: 5% to 7% Buffer B (acetonitrile with 0.1% formic acid) *versus* Buffer A over 2 minutes, then to 35% over 90 minutes, then to 50% over 1 minute, hold at 50% for 9 minutes, change to 95% over 1 minute, hold at 95% for 5 minutes, drop to 5% over 1 minute and recondition at 5%. The flow rate for the peptide separation was 300 nL per minute. For MS analysis, a spray voltage of 2.25 kV was applied to the nanospray tip. The mass spectrometric analysis was performed using data-dependent acquisition with a m/z range of 400–1800, consisting of a full MS scan in the Orbitrap followed by up to 5 MS/MS spectra acquisitions in the linear ion trap using collision induced dissociation (CID). Other mass spectrometer parameters include: isolation width 2 m/z , target value $1e4$, collision energy 35%, max injection time 100 ms. Lower abundance peptide ions were interrogated using dynamic exclusion (exclusion time 45 seconds, exclusion mass width $-0.55 m/z$ low to $1.55 m/z$ high). Charge state screen was used, allowing for MS/MS of any ions with identifiable charge states +2, +3, and +4 and higher.

Proteomics data analysis

Raw machine output files of MS/MS spectra were converted to mzXML files and searched with X!Tandem³⁶ configured with the *k*-score scoring algorithm,³⁷ against version 3.69 of the human International Protein Index (IPI) database with the addition of yeast invertase 2. The search parameters were as follows: enzyme: trypsin; maximum missed cleavage: 1; fixed modifications: carboxamidomethylation on cysteine, light dimethylation on N-terminus and lysine; variable modifications: oxidation on methionine, difference between light and heavy dimethylation on N-terminus and lysine, enzymatic conversion of asparagine to aspartic acid; parent mono-isotopic mass tolerance: 2.5 Da. Peptide identifications were assigned probability by PeptideProphet.³⁴ Relative quantitation of light and heavy peptide abundance was performed with Xpress³⁸ version 2.1. Proteins present in the sample were inferred using ProteinProphet.³⁹

Conclusion

In this report, we describe an integrated glycoproteomics method that provides quantitative comparison of site-specific glycosylation occupancy on a global scale. This method provides several technical advantages, including significant improvement of quantitative accuracy by minimizing the variations introduced during the glyco-capturing process, simplicity and robustness for sample preparation, cost-effectiveness and compatibility for a variety of mass spectrometers. An assortment of sample types, including tissue, cell lysate/secretome, plasma/serum and other bodily fluids, can be analyzed using this approach for glycoproteomics profiling. We have demonstrated this method through the quantitative characterization of the N-glycoproteome of the human pancreas. For glyco-peptide enrichment, in addition to the widely used approaches based on hydrazide chemistry and lectin affinity, other mechanisms, such as boronic acid,^{40,41} size-exclusion chromatography,⁴² hydrophilic interaction⁴³ and a graphite powder micro-column,⁴⁴ can also be integrated into the pipeline. For the hydrazide chemistry method, glycans with different glycoform variations are oxidized for capturing, thus, this method provides a non-glycoform-specific glycosylation occupancy assessment on a glycosylation site. However, a lectin affinity column with specific affinity can be used to capture glycopeptides with certain glycoforms. In such a setting, the glycopeptide abundance may reflect the occupancy density of specific glycoforms on a glycosylation site. Multiple proteases may be needed in order to quantitatively characterize some specific glycosylation sites, as a tryptic peptide may carry more than one N-linked glycosylation sites or have an *m/z* value beyond the analytical mass range. This pipeline has been demonstrated to be robust for quantitative profiling of N-glycoproteome, and it may be applicable to O-linked and other glycosylations if suitable enzymatic or chemical cleavage methods are used to remove glycans, or intact glycopeptides can be analyzed directly. The development of this low-cost and efficient method provides an effective means for quantitative glycoproteomics profiling with improved accuracy and is particularly useful for quantitative identification of aberrant glycosylation states associated with diseases or other biological settings in a complex sample.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported in part with federal funds from the National Cancer Institute/NIH under grants K25CA137222, R21CA149772, R01CA107209, K07CA116296, and R21CA161575.

References

1. Ludwig JA, Weinstein JN. *Nat Rev Cancer*. 2005; 5:845. [PubMed: 16239904]
2. Polanski M, Anderson NL. *Biomarker Insights*. 2007; 1:1. [PubMed: 19690635]
3. Pan S, Chen R, Aebersold R, Brentnall TA. *Mol Cell Proteomics*. 2011; 10:R110. [PubMed: 20736408]
4. Baas S, Sharrow M, Kotu V, Middleton M, Nguyen K, Flanagan-Steet H, Aoki K, Tiemeyer M. *Development (Cambridge, UK)*. 2011; 138:553.
5. Rudd PM, Elliott T, Cresswell P, Wilson IA, Dwek RA. *Science*. 2001; 291:2370. [PubMed: 11269318]
6. Tian Y, Yao Z, Roden RB, Zhang H. *Proteomics*. 2011; 11:4677. [PubMed: 22113853]
7. Chaerkady R, Thuluvath PJ, Kim MS, Nalli A, Vivekanandan P, Simmers J, Torbenson M, Pandey A. *Clin Proteomics*. 2008; 4:137. [PubMed: 20357908]
8. Ueda K, Katagiri T, Shimada T, Irie S, Sato TA, Nakamura Y, Daigo Y. *J Proteome Res*. 2007; 6:3475. [PubMed: 17705522]
9. Wilson CR, Regnier FE, Knapp DW, Raskin RE, Andrews DA, Hooser SB. *Vet Comp Oncol*. 2008; 6:171. [PubMed: 19178677]
10. Zhang H, Li XJ, Martin DB, Aebersold R. *Nat Biotechnol*. 2003; 21:660. [PubMed: 12754519]
11. Liu Z, Cao J, He Y, Qiao L, Xu C, Lu H, Yang P. *J Proteome Res*. 2010; 9:227. [PubMed: 19921957]
12. Qiu R, Regnier FE. *Anal Chem*. 2005; 77:7225. [PubMed: 16285669]
13. Chen R, Tan Y, Wang M, Wang F, Yao Z, Dong L, Ye M, Wang H, Zou H. *Mol Cell Proteomics*. 2011; 10:M110.
14. Ueda K, Takami S, Saichi N, Daigo Y, Ishikawa N, Kohno N, Katsumata M, Yamane A, Ota M, Sato TA, Nakamura Y, Nakagawa H. *Mol Cell Proteomics*. 2010; 9:1819. [PubMed: 20811073]
15. Jung K, Cho W, Regnier FE. *J Proteome Res*. 2009; 8:643. [PubMed: 19099503]
16. Lee HJ, Na K, Choi EY, Kim KS, Kim H, Paik YK. *J Proteome Res*. 2010; 9:308. [PubMed: 19899825]
17. Plavina T, Wakshull E, Hancock WS, Hincapié M. *J Proteome Res*. 2007; 6:662. [PubMed: 17269723]
18. Ishihara T, Fukuda I, Morita A, Takinami Y, Okamoto H, Nishimura S, Numata Y. *J Proteomics*. 2011; 74:2159. [PubMed: 21704746]
19. Lei Z, Beuerman RW, Chew AP, Koh SK, Cafaro TA, Urrets-Zavalía EA, Urrets-Zavalía JA, Li SF, Serra HM. *J Proteome Res*. 2009; 8:1992. [PubMed: 19714880]
20. Parker BL, Palmisano G, Edwards AV, White MY, Engholm-Keller K, Lee A, Scott NE, Kolarich D, Hambly BD, Packer NH, Larsen MR, Cordwell SJ. *Mol Cell Proteomics*. 2011; 10:M110. [PubMed: 21441315]
21. Hsu JL, Huang SY, Chow NH, Chen SH. *Anal Chem*. 2003; 75:6843. [PubMed: 14670044]
22. Liu T, Qian WJ, Gritsenko MA, Camp DG, Monroe ME, Moore RJ, Smith RD. *J Proteome Res*. 2005; 4:2070. [PubMed: 16335952]
23. Pan S, Wang Y, Quinn JF, Peskind ER, Waichunas D, Wimberger JT, Jin J, Li JG, Zhu D, Pan C, Zhang J. *J Proteome Res*. 2006; 5:2769. [PubMed: 17022648]
24. Zhang H, Aebersold R. *Methods Mol Biol (Totowa, N J)*. 2006; 328:177.
25. Sun B, Ranish JA, Utleg AG, White JT, Yan X, Lin B, Hood L. *Mol Cell Proteomics*. 2007; 6:141. [PubMed: 17074749]
26. Zhou Y, Aebersold R, Zhang H. *Anal Chem*. 2007; 79:5826. [PubMed: 17591751]
27. Yang Z, Harris LE, Palmer-Toy DE, Hancock WS. *Clin Chem (Washington, D C)*. 2006; 52:1897.
28. Geng M, Zhang X, Bina M, Regnier F. *J Chromatogr, B: Biomed Sci Appl*. 2001; 752:293. [PubMed: 11270868]
29. Kaji H, Saito H, Yamauchi Y, Shinkawa T, Taoka M, Hirabayashi J, Kasai K, Takahashi N, Isobe T. *Nat Biotechnol*. 2003; 21:667. [PubMed: 12754521]
30. Yang Z, Hancock WS. *J Chromatogr, A*. 2004; 1053:79. [PubMed: 15543974]

31. Wang Y, Wu SL, Hancock WS. *Glycobiology*. 2006; 16:514. [PubMed: 16497783]
32. Durham M, Regnier FE. *J Chromatogr, A*. 2006; 1132:165. [PubMed: 16919642]
33. Qiu R, Regnier FE. *Anal Chem*. 2005; 77:2802. [PubMed: 15859596]
34. Keller A, Nesvizhskii AI, Kolker E, Aebersold R. *Anal Chem*. 2002; 74:5383. [PubMed: 12403597]
35. Huang DW, Sherman BT, Lempicki RA. *Nat Protocols*. 2009; 4:44.
36. Craig R, Beavis RC. *Bioinformatics*. 2004; 20:1466. [PubMed: 14976030]
37. Keller A, Eng J, Zhang N, Li XJ, Aebersold R. *Mol Syst Biol*. 2005; 1:2005. [PubMed: 16729052]
38. May D, Law W, Fitzgibbon M, Fang Q, McIntosh M. *J Proteome Res*. 2009; 8:3212. [PubMed: 19309175]
39. Nesvizhskii AI, Vitek O, Aebersold R. *Nat Methods*. 2007; 4:787. [PubMed: 17901868]
40. Sparbier K, Koch S, Kessler I, Wenzel T, Kostrzewa M. *J Biomol Tech*. 2005; 16:407. [PubMed: 16522863]
41. Sparbier K, Wenzel T, Kostrzewa M. *J Chromatogr, B: Anal Technol Biomed Life Sci*. 2006; 840:29.
42. Alvarez-Manilla G, Atwood J III, Guo Y, Warren NL, Orlando R, Pierce M. *J Proteome Res*. 2006; 5:701. [PubMed: 16512686]
43. Hagglund P, Bunkenborg J, Elortza F, Jensen ON, Roepstorff P. *J Proteome Res*. 2004; 3:556. [PubMed: 15253437]
44. Larsen MR, Hojrup P, Roepstorff P. *Mol Cell Proteomics*. 2005; 4:107. [PubMed: 15561728]

Fig. 1. Analytical flow for global quantitative profiling of site-specific glycosylation occupancy. (a) Trypsin digestion, (b) dimethyl labeling, (c) glycopeptide capturing, (d) glycan cleavage, (e) LC MS/MS analysis, (f) database search for peptide/protein identification and quantitative analysis.

Fig. 2. Identification of de-glycosylated N-glycopeptides using high resolution tandem mass spectrometer and stringent identification criteria. (a) Distribution of mass deviation for the de-glycosylated N-glycopeptides identified with a PeptideProphet probability 0.95. (b) Representative MS/MS identification of formerly N-linked glyco-peptide ED*GTDTVQEEEESPAEGSK (HYPOXIA UP-REGULATED PROTEIN 1, * indicates where the glycan was attached. N was converted to D due to PNGase F enzymatic cleavage of N-glycans).

Fig. 3. Distribution of glycoproteins identified in human pancreas tissue with single and multiple annotated N-glycosylation sites.

Fig. 4. Gene ontology annotation of glycoproteins identified in human pancreas tissue. (a) Cellular component, (b) biological process.

Fig. 5.

Quantitative evaluation of the glycoproteomics method by analyzing the N-glycoproteome of human pancreas tissue with a theoretical 1 : 1 heavy-to-light isotopic ratio. (a) An example showing the co-elution of the heavy and light isotopic forms of the formerly N-linked glycopeptide ED*GTDTVQEEEESPAEGSK (HYPOXIA UP-REGULATED PROTEIN 1, * N-linked glycosylation site). The heavy-to-light ratio is 1.01 (theoretical ratio 1.0). (b) The distribution of the heavy-to-light ratios (in natural log scale) of all the N-linked glycopeptides identified in the differentially isotopic labeled pancreas tissue sample. The natural log ratios are narrowly distributed and centered on 0.