# Supporting Information

## Improving glycoproteomic analysis workflow by systematic evaluation of glycopeptide enrichment, quantification, mass spectrometry approach, and data analysis strategies

Zhenyu Sun[1]; T. Mamie Lih [1]; Jongmin Woo[1]; Liyuan Jiao[1]; Yingwei Hu[1]; Yuefan Wang[1]; Hongyi Liu[1]; Hui Zhang[1*]

[1] Department of Pathology, Oncology, Chemical and Biomolecular Engineering, and Urology, Johns Hopkins University School of Medicine, Baltimore, MD 21231

# Contents

# Experimental section

**SepPak C18 desalting of digested peptides.** After digestion, samples were centrifugated at 14000×g for 15 min, the supernatants containing resulting digested tryptic peptides were desalted with C18 cartridges. Cartridges were conditioned in 1 mL ACN twice and balanced twice each with 1 mL 50% ACN (v/v) 0.1% FA (v/v) and 0.1% TFA (v/v). The samples were loaded and washed in 1 mL 0.1% TFA (v/v) twice, followed by 1 mL 1% FA (v/v). and then elution using 500 µL 50% ACN (v/v) 0.1% FA (v/v) twice. The samples were dried in a vacuum centrifuge and stored at -80 ℃.

**TMTpro labeling of peptides.** TMTpro labeling procedure was performed according to the manufacturer's instruction (ThermoFisher). In brief, a total of 300 µg peptides were dissolved in 60 µL of 100 mM HEPES. TMTpro reagents were dissolved in ACN. A total of 60 µL of each TMTpro reagent was added to the corresponding aliquot of peptides. The TMT labeling reaction was incubated at room temperature for 1h with shaking. The reaction was quenched with 5% hydroxylamine at room temperature for 15 min. Then, peptides were desalted using SepPak C18 and subsequently dried by vacuum centrifugation.

**Enrichment of glycopeptides by ZIC-HILIC enrichment**. Intact glycopeptides were selectively enriched using a ZIC-HILIC micro-column, which was prepared in-house following a previously published protocol literature [1]. Briefly, ZIC-HILIC column was equilibrated in 200 µL 80% ACN (v/v) 1% TFA (v/v) and repeated for three times. The peptides were loaded into the ZIC-HILIC micro-column, and then the column was washed with 200 µL 80% ACN (v/v) 1% TFA (v/v) for three times. Enriched glycopeptides were eluted with 200 µL 0.1% TFA for three times. And subsequently dried by vacuum centrifugation and stored at −80 °C before LC-MS/MS analysis.

**Enrichment of glycopeptides by MAX enrichment.** To enrich glycopeptides using MAX enrichment, MAX cartridge was sequentially conditioned with 1 mL of methanol twice, 1 mL of ACN twice, 1 mL of 100 mM triethylammonium acetate twice, 1 mL of water twice, and 1 mL of 95% ACN (v/v) 1% TFA (v/v) twice. Samples were dissolved with 1 mL of 95% ACN (v/v) 1% TFA (v/v) and loaded onto cartridge twice [16]. The non-glycopeptides were eluted out by 95% ACN (v/v) 1% TFA (v/v) for four times. Finally, intact glycopeptides were eluted by 500 µL of 50% ACN (v/v) 0.1% TFA (v/v), dried and stored at −80 °C before LC-MS/MS analysis.

**Basic reverse phase fractionation.** The TMT-labeled samples were reconstituted and loaded onto a 4.6 mm x 250 mm reverse phase (RP) Zorbax 300 A Extend-C18 column with 3.5 µm size beads. Peptides were separated into 24 fractions by following previously established protocol using off-line HPLC fractionation [2]. The fractions were later combined into 12 fractions for LC-MS/MS analysis.

**LC-MS/MS analysis.** Intact glycopeptides were analyzed by Orbitrap Exploris™ 480 MS (Thermo Scientific) combined with Easy-nLC™ 1200 system (Thermo Scientific). Samples were reconstituted with 3% ACN in 0.1% formic acid (solvent A) and loaded onto an in-house packed column (75 µm I.D. x 27.5 cm length packed with ReproSil-

Pur 120 C18-AQ, 1.9 μm). For label-free intact glycopeptides, loaded peptides were subjected to the gradient with 300 μL/min of flow rate as follows: 1 to 4% B (90% ACN 0.1% F.A) for 15 min, 4 to 25% B for 90 min, 25 to 45% B for 20 min, 45 to 90% B for 3 min, isocratic 90% B for 8 min, 90 to 50% B for 2 min, and isocratic 50% B for 5 min. The temperature of the column was maintained by a column heater (Phoenix-ST) at 60°C. The ion source of the mass spectrometer was set up with 1.8 kV of electrospray voltage and 300 °C of ion transfer tube temperature. The precursor ion scan was acquired with 120 K resolution at 200 m/z for from 350 to 2,000 m/z range and AGC value was set as $5 \times 10^5$. Precursor ions isolated from 2 m/z width were fragmented by different HCD NCE (**Table S1)** and fragment ions were acquired with 30 K resolution with $1 \times 10^5$ of AGC value for 64 ms of injection time for a total duty cycle (2 s). The peptide charge state screening was enabled to include 2 to 8 ions with a dynamic exclusion time of 45 s to discriminate against previously analyzed ions between +/− 10 ppm.

For TMTpro labeled intact N-glycopeptides, the gradient as follows: 2 to 6% B (90% ACN 0.1% F.A) for 5 min, 6 to 30% B for 95 min, 30 to 60% B for 25 min, 60 to 90% B for 3 min, isocratic 90% B for 5 min, 90 to 50% B for 2 min, and isocratic 50% B for 5 min. The precursor ion scan was acquired with 60 K resolution at 200 m/z for from 350 to 2000 m/z range. Precursor ions isolated from 0.7 m/z. Fragment ions were acquired with turboTMT. Other parameters are the same as label-free methods.

For the 12 fractionated samples, intact glycopeptides were analyzed by Orbitrap Ascend Tribrid MS (Thermo Scientific) combined with Evosep One system. Samples were reconstituted with 0.1% FA and loaded onto Evotips. Fragment ions were acquired with 30 K resolution with $3 \times 10^5$ of AGC value for 64 ms of maximum injection time for a total duty cycle (2 s). Other parameters are same as Orbitrap Exploris™ 480 MS.

**Data analysis.** The raw data files were analyzed using pGlyco3.0, MS- PyCloud, and MSFragger software [26-28]. For label-free samples, the parameters were set as follows: oxidation [M, +15.995 Da] and acetyl [ProteinN-term, +42.011 Da] for variable modifications. Carbamidomethyl [C, +57.022 Da] for fixed modification, maximum 3 missed cleavages, 10 ppm for precursor tolerance, 20 ppm for fragment tolerance, trypsin for enzyme, and a false discovery rate (FDR) of 1%. The human and mouse protein database were downloaded from UniProt/Swiss-Prot (version, 2019-12). The human N-glycan database in pGlyco 3.0 was used. And for MS- PyCloud, and MSFragger-Glyco search, the same N-glycan database was used. For TMTpro labeled samples, set variable modifications, oxidation [M, +15.995 Da], set fix modifications as follow: carbamidomethyl [C, +57.022 Da], Similar settings were used for TMTpro, except TMTpro [any-N term, + 304.207 Da], and TMTpro [K, + 304.207 Da] were selected as fixed modifications in addition to carbamidomethyl [C, +57.022 Da]. pGlycoQuant software was utilized for quantification [3].

# Tables

**Table S1** HCD collision energy for label-free intact glycopeptides

| HCD collision energy | Single/Step |
|---|---|
| 20 | Single |
| 25 | Single |
| 30 | Single |
| 35 | Single |
| 40 | Single |
| 45 | Single |
| 55 | Single |
| 20_30_40 | Step |
| 25_35_45 | Step |

**Table S2** HCD collision energy for TMTpro labeled intact glycopeptides

| HCD collision energy | Single/Step |
|---|---|
| 25 | Single |
| 30 | Single |
| 35 | Single |
| 40 | Single |
| 45 | Single |
| 25_35_45 | Step |
| 35_45_55 | Step |

# Results and discussion

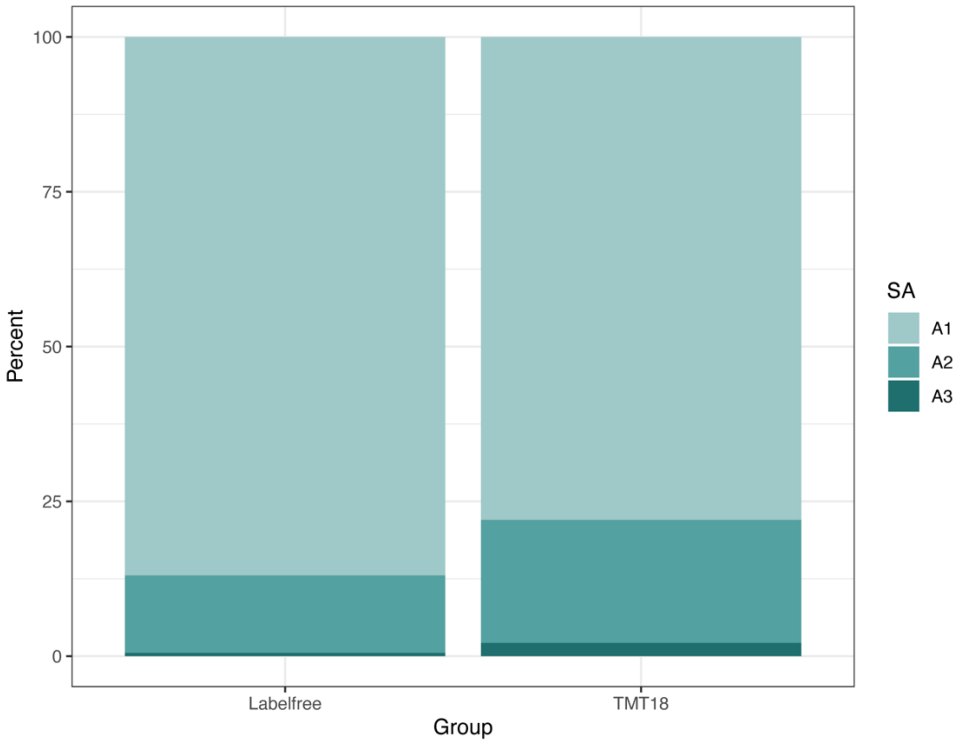


**Figure S1** The identification number of salic acid-containing intact glycopeptides from unlabeled and TMTpro labeled intact glycopeptides.

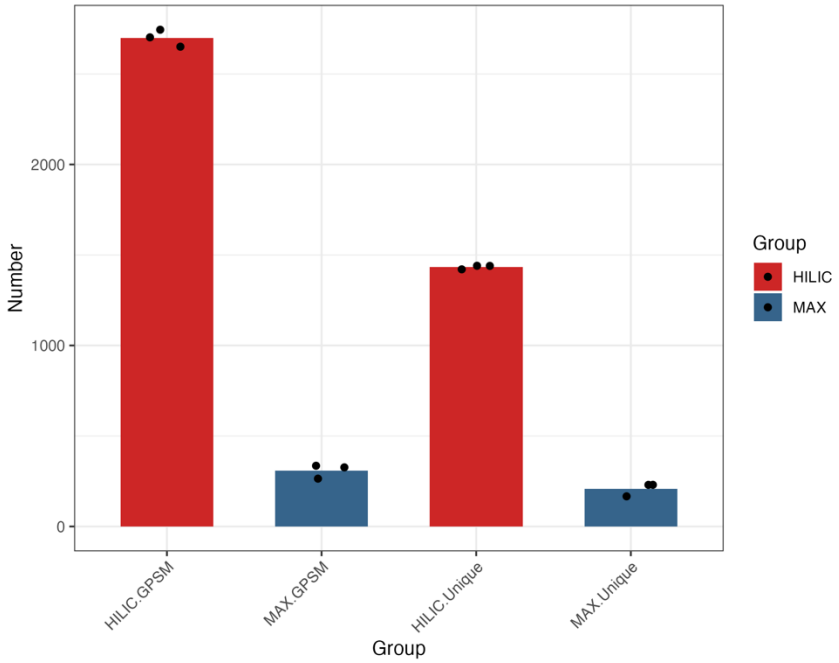


**Figure S2** Comparative performance of ZIC-HILIC and MAX Enrichment strategies for direct glycopeptide enrichment.

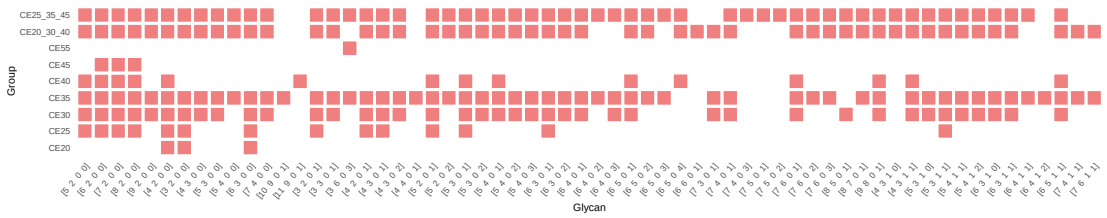


**Figure S3** The N-glycans of LLNINPNK.T (LAPM1) from different collision energies for mass spectrometry analysis. Numbers in the X-axis represent glycans containing the numbers of Hex (H), HexNAc (N), A (NeuAc), F (fucose), respectively.

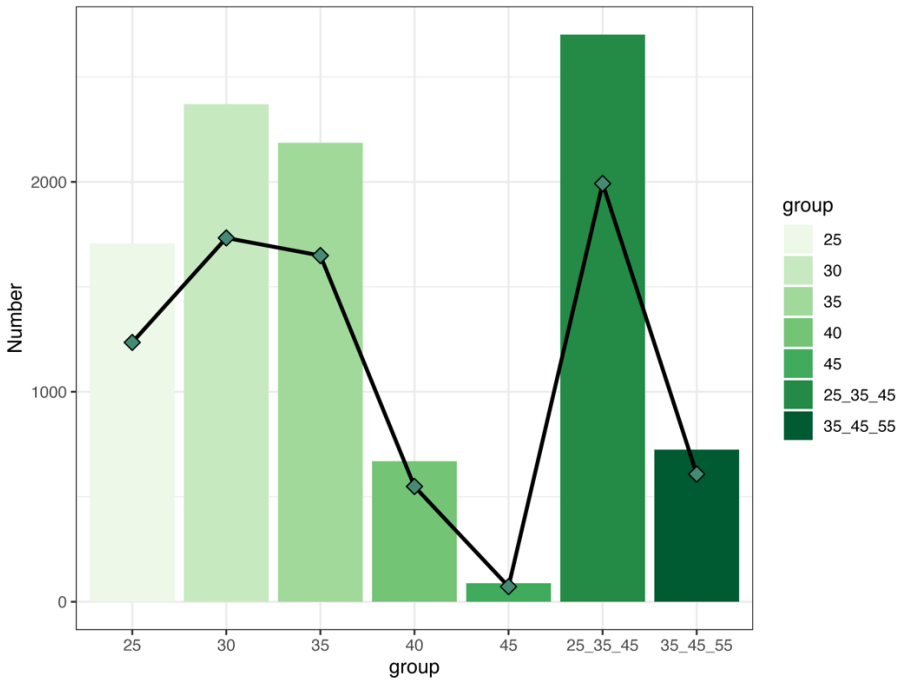


**Figure S4** The characterization of TMT pro labeling intact N- glycopeptides using different collision energies.

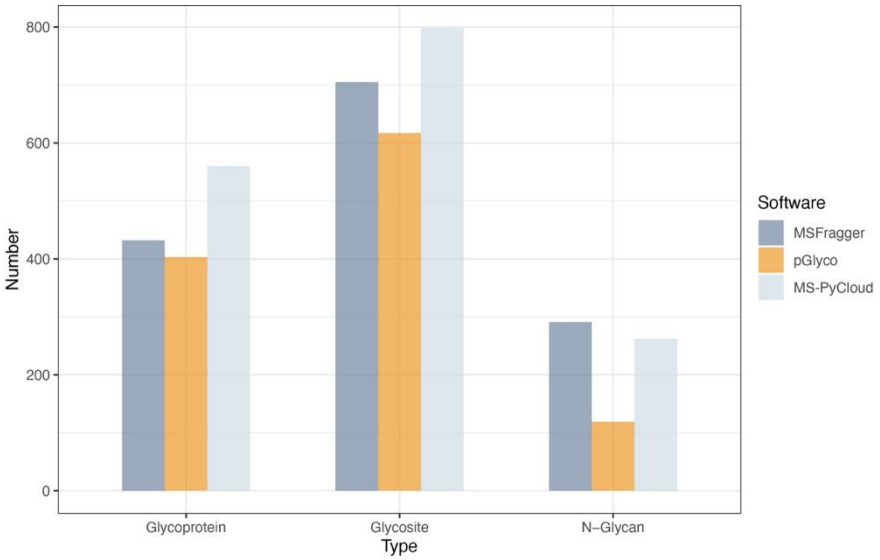


**Figure S5** The number of glycoproteins, glycosites, and N-Glycans identified from different software.

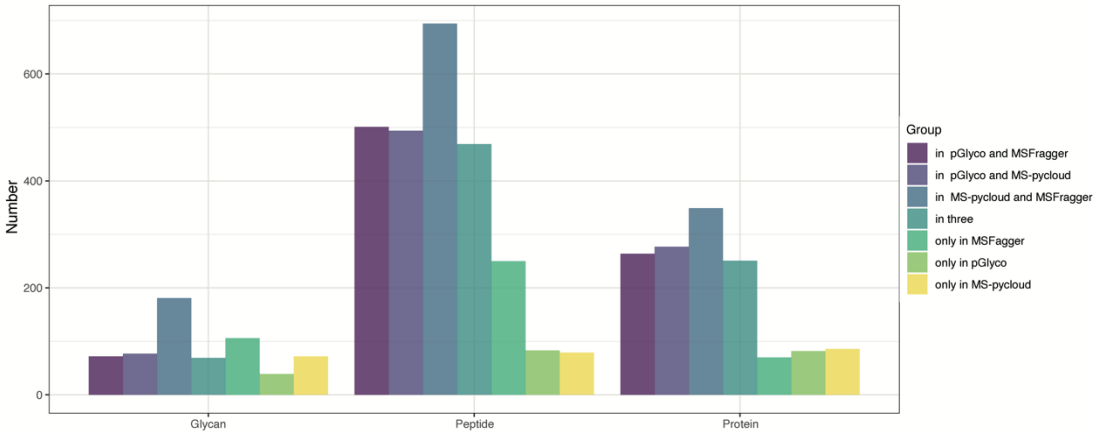


**Figure S6** Distribution of glycans, peptide backbones, and proteins in TMTpro-labeled intact glycopeptides analyzed with different software tools.

# References

(1) Sun, Z. Y.; Ji, G. H.; Wang, G. L.; Wei, L.; Zhang, Y.; Lu, H. J. One step carboxyl group isotopic labeling for quantitative analysis of intact glycopeptides by mass spectrometry. *Chem. Commun.* **2021**, *57*, 4154-4157.
(2) Mertins, P.; Tang, L. C.; Krug, K.; Clark, D. J.; Gritsenko, M. A.; Chen, L. J.; Clauser, K. R.; Clauss, T. R.; Shah, P.; Gillette, M. A.; et al. Reproducible workflow for multiplexed deep-scale proteome and phosphoproteome analysis of tumor tissues by liquid chromatography-mass spectrometry. *Nature Protocols* **2018**, *13*, 1632-1661.
(3) Kong, S. Y.; Gong, P. Y.; Zeng, W. F.; Jiang, B. Y.; Hou, X. H.; Zhang, Y.; Zhao, H. H.; Liu, M. Q.; Yan, G. Q.; Zhou, X. W.; et al. pGlycoQuant with a deep residual network for quantitative glycoproteomics at intact glycopeptide level. *Nat. Commun.* 2022, *13*,7539.