Supporting Information

An Effective Enrichment Strategy using Boronic Acid-Functionalized Mesoporous Graphene–Silica Composites for Intact N- and O-linked Glycopeptide Analysis in Human Serum

Siyuan Kong[1], Quanqing Zhang[1], Lujie Yang[1], Yuanyu Huang[1], Mingqi Liu[1], Guoquan Yan[1], Huanhuan Zhao[1], Mengxi Wu[1,2], Xiangmin Zhang[1,2,3], Pengyuan Yang[1,2,3,*], Weiqian Cao[1,2,*]

1. The Fifth People's Hospital, Fudan University, and the Shanghai Key Laboratory of Medical Epigenetics, the International Co-laboratory of Medical Epigenetics and Metabolism, Ministry of Science and Technology, Institutes of Biomedical Sciences, Fudan University, Shanghai, 200032, China
2. NHC Key Laboratory of Glycoconjugates Research, Fudan University, Shanghai, 200032, China
3. Department of Chemistry, Fudan University, Shanghai, 200043, China

* To whom correspondence should be addressed:

Pengyuan Yang (TEL:21-54237961, Email: pyyang@fudan.edu.cn)

Weiqian Cao (TEL: 21-54237416, Email: wqcao@fudan.edu.cn)

Table of contents

Experiment details

S1 Materials and chemicals

Hexadecyl trimethyl ammonium bromide (CTAB), ethanol, ethylene glycol (analytical grade), 2,5-dihydroxybenzoic acid (2,5-DHB), ammonium solution (28 wt%), ammonium bicarbonate (NH_4HCO_3, ABC) and bovine serum albumin (BSA) were provided by Shanghai Chemical Reagent (China). Standard glycoproteins including immunoglobin G (IgG) from human serum, fetuin from fetal calf serum, horseradish peroxidase (HRP) from horseradish), acetonitrile (ACN, mass spectrum grade), trifluoroacetic acid (TFA), and tetraethyl orthosilicate (TEOS) were purchased from Merck (Darmstadt, Germany). Cyclohexane was obtained from Sinopharm Chemical Reagent (China). Urea, dithiothreitol (DTT) and iodoacetamide (IAA) were obtained from Bio-Rad (Hercules, CA, USA). PNGase F (500 units/μL) and OpeRATOR were obtained from New England Biolabs (MA, USA) and Genovis Inc. (MA, USA), respectively. The commercial materials of Sepharose 4FF and Bond Elut PBA were purchased from GE Healthcare (Uppsala, Sweden) and Agilent Technologies (Santa Clara, CA) severally. Human serum samples from healthy volunteers were kindly provided by the Fudan University Shanghai Zhongshan Hospital and were stored at -80 °C before analysis. Informed consent was obtained from each volunteer. The research followed the tenets of the Declaration of Helsinki and was approved by the Ethics Committee of the Fudan University Shanghai Zhongshan Hospital. The deionized water used in all

experiments was purified using a Milli-Q water system.

S2 Synthesis of mesoporous silica-coated graphene (GO@mSiO$_2$)

Graphene oxide (100 mg) was added to deionized water (20 mL) and uniformly dispersed by sonication for a total of 20 min with 5 seconds' intervals. A solution of NaOH (0.1 mol/L) was added to the GO solution to adjust the pH to 11.8 to 12.8. The solution was mechanically stirred at 85 °C overnight. Then, 0.5 g of CTAB, aqueous ammonia solution (0.8 mL) (28 wt%) and processed GO (20 mg) were added to the deionized water (50 mL) mixture, and mechanical stirring was continued for 10 min at room temperature. After 10 min, cyclohexane (10 mL) containing TEOS (500 μL) was slowly added dropwise into the mixture with gentle mechanical stirring over a period of 12 h. This operation effectively forms a two-phase liquid system with CTAB micelle expansion. Then, the lower liquid was drawn out through a syringe, resuspended in deionized water and centrifuged (10,000x g, 5 min). After centrifugation, the lower liquid was drawn out again. Next, the separated powder was washed 3 times with deionized water and ethanol. The synthesized composite was refluxed with acetone at 80 °C for 12 h to remove the organic phase and template. Finally, before further use, the obtained sample was dried overnight in a vacuum at 65 °C.

S3 Preparation of GA (GLYMO-APB)

Boronic acid-bonded GLYMO was prepared by the reaction of GLYMO and APB (denoted as GA). Generally, 50 mg of APB was dissolved in 20 mL of deionized water and the pH was adjusted to 9.18 with 0.1 mol/L NaOH. The APB solution had to be kept in an ice bath for 5 min to avoid hydrolysis of GLYMO. Then, 40 μL of GLYMO was slowly added dropwise into the mixture. When the dropwise addition was complete, the mixture was stirred and allowed to react at 40°C for 6 h. Before adding another 40 μL of GLYMO mixture solution, the temperature of the mixture was decreased to 0 °C in an ice bath. Finally, the reaction was continued at 65°C for another 6 h. The GA solution was stored at 4°C for further use.

S4 Material characterization

The pore structure and morphology were observed by transmission electron microscopy (TEM, JEOL, Tokyo, Japan) with the accelerating voltage set at 120 kV. Scanning electron microscopy (SEM) images were obtained using a GeminiSEM 500 scanning electronic microscope (ZEISS, Germany). Measurements of nitrogen adsorption-desorption for GO@mSiO_2-GLYMO-APB composite was conducted using an ASAP 2010 system (Micromeritics, USA) employing a static-volumetric method at -196 °C (liquid nitrogen temperature), and the Brunauer-Emmett-Teller (BET) surface area was determined using a relative pressure (P/P0) range from 0.01 to 0.98. The Barrett-Joyner-Halenda (BJH) method was used to calculate the distribution curves and pore diameters from the adsorption branch. Fourier transform infrared (FT-IR) spectroscopy was performed using a Thermo Nicolet 380 spectrometer on samples in KBr pellet form (Nicolet 6700, Wisconsin, U.S.A.).

S5 MALDI-TOF MS analysis

Sample solutions were deposited onto MALDI target plates using the dried droplet method. Sample solution, 0.5 μL, was spotted onto the MALDI plate, and then, another 0.5 μL of DHB matrix solution (12.5 mg/mL, 0.1% TFA in 20% ACN/H_2O solution) was added, and the sample was air-dried. MALDI-TOF-MS analysis was conducted using a 5800 MALDI-TOF analyzer (AB Sciex, USA). All mass spectra (1500 laser shots for every spectrum) were acquired in a positive reflection mode with the m/z scan range of 1500–3500 and analyzed using Data Explorer (Version 4.5).

S6 LC-MS/MS analysis

Intact glycopeptides were detected using a nanospray LC-MS/MS on an Orbitrap Fusion Tribrid system (Thermo Fisher Scientific, Waltham, MA, USA) fitted with an EASY-nLC TM1100 system (Thermo Fisher Scientific, Waltham, MA, USA) that included a reverse-phase analytical column without the trap column. For one LC-MS run, samples were loaded onto a C18 column (50 cm × 75 μm i.d. column) and were separated at a flow rate of 300 nL/min. Solvent A was a 0.1% formic acid aqueous solution. Solvent B was acetonitrile containing 0.1% formic acid. The gradient lasted 3 h in total for complex

samples: 1% to 30% in 165 min, an increase to 45% B in 7 min, hold 1 min with 90% B, hold for another 3 min, and then return to 1% B for 4 min. The samples were analyzed by SCE-HCD-MS/MS. The parameters for intact glycopeptide analysis were: (1) MS: scan range (m/z) = 350–2000; resolution = 120 000; AGC target = 200 000; maximum injection time = 100 ms; included charge state = 2–6; dynamic exclusion after n times, n = 1; dynamic exclusion duration = 15 s; the precursors are selected under the “top speed” mode in Thermo Fusion mass spectrometer and each selected precursor was subjected to one HCD-MS/MS; (2) HCD-MS/MS: isolation window = 2; detector type = Orbitrap; resolution = 15 000; AGC target = 500 000; maximum injection time = 250 ms; collision energy = 30%; and stepped collision mode on with an energy difference of ±10% (10% as absolute value in the Orbitrap Fusion).

S7 Optimization of the enrichment experimental conditions

To obtain higher efficiency of enrichment, the experimental parameters of binding buffer ratios and wash times were systematically evaluated. The IgG standard digest (1 μg/uL, 6μL) was add to the GO@mSiO_2-GLYMO-APB (300 μg, 200 μL) and vortexed for 20 s. Using different proportions of ACN and TFA as binding buffer, the effect of six binding buffers on the enrichment efficiency was investigated. Three concentrations of ACN (85%, 90% and 95%) were prepared with 0.5% and 1% TFA. In accordance with the enrichment method described above, the supernatants were collected, lyophilized and stored at -80 °C for subsequent MS analysis. In addition, the number of washing times after the incubation was also a factor that needed to be optimized. Thus, the combinations of different binding buffers and washing times (90% ACN/1% TFA (v/v), 3 washes; 90% ACN/1% TFA (v/v), 6 washes; 95% ACN/1% TFA (v/v), 3 washes; 95% ACN/1% TFA (v/v), 6 washes) were applied to the improvement of enrichment efficiency. Finally, samples were centrifuged to gather the supernatant and freeze-dries for follow-up experimental analysis.

S8 Estimation of recovery rate

In the experiment of recovery rate estimation, a total of 24 μg IgG digests were divided into two portions with equal quantities. One portion was directly

treated by PNGase F to obtain the de-glycopeptides and peptide mixtures in $H_2{}^{16}O$. The other portion was used to enrich glycopeptides by GO@mSiO_2-GLYMO-APB. The enriched glycopeptides were further treated by PNGase F in $H_2{}^{18}O$ to obtain the de-glycopeptides isotopically labeled with +2 Da from ^{18}O. Then, the final products of the two portions were mixed and analyzed by MALDI-TOF. The peak intensity ratios of the de-glycopeptides from the two portions were observed and calculated to evaluate the recovery rate. The de-glycopeptides EEQFNSTFR (1157.2 Da) from IgG2/3, EEQYNSTYR (1189.2 Da) from IgG1, and their isotopically labeled de-glycopeptides EEQFNSTFR+2 Da (1159.2 Da) and EEQYNSTYR+2 Da (1191.2 Da) were observed. Comparing the corresponding peak intensities, we could calculate the recovery rate (Figure S6).

S9 Estimation of enrichment specificity for standard glycopeptides

In the experiment of enrichment specificity estimation, a mixture containing two standard glycopeptides (ADKDESTMIN[Hex5HexNAc4NeuAc2]EITLAK and QN[Hex5HexNAc4NeuAc2]CELFEQLGEYK, 1μg/μL, 1μL), BSA tryptic digests (1μg/μL, 6μL) and BSA proteins (1μg/μL, 6μL) was enriched byGO@mSiO_2 or GO@mSiO_2-GLYMO-APB with the same enrichment procedure and analyzed by LC-MS/MS. The enrichment specificity was defined and calculated as the percentage of glycopeptide mass spectra among total mass spectra. [HexNAc+H]+ (m/z=204.08) was set as the diagnostic oxonium ion to determine whether a peptide-spectrum match (PSM) belongs to a glycan/glycopeptide or not.

S10 Glycopeptide enrichment by commercial materials

In Sepharose 4FF-based enrichment, enrichment was performed according to the manufacturer's instruction. Briefly, appropriate amount of Sepharose 4FF was washed by 50% ethanol in water three times at first. Then, peptides were resuspended by solution A (1-butanol:ethanol:water = 41:1, v/v/v), mixed with Sepharose 4FF (volume ratio of material to solution is 1:5)and shook gently for 1 h. Samples were washed for three times by solution A, and eluted by 50% ethanol in water at room temperature for 15 mins twice. The obtained

glycopeptides were freeze-dried and reserved for analysis.

In Bond Elut PBA-based enrichment, the cartridge was washed with 2 mL ACN, 1mL 0.5% TFA in water and 2 mL 0.2 mol/L ammonium acetate (PH 8.5) respectively. Peptides were resuspended by 0.2 mol/L ammonium acetate and loaded into the column using a slow flow rate, followed by washing with 1 mL 0.2 mol/L ammonium acetate and 2mL water. Glycopeptides were eluted using 0.5% FA in water once and 1.0% FA in water twice. The eluents were freeze-dried and reserved for analysis.

Supporting figures

Figure S1 MALDI-TOF MS spectra of IgG tryptic digests after enrichment by GO@mSiO_2-GLYMO-APB with different binding buffers: (A) 85% ACN/0.5% TFA (v/v), (B) 85% ACN/1% TFA (v/v), (C) 90% ACN/0.5% TFA (v/v), (D) 90% ACN/1% TFA (v/v), (E) 95% ACN/0.5% TFA (v/v), (F) 95% ACN/1% TFA (v/v). The N-glycopeptides are marked with red*.

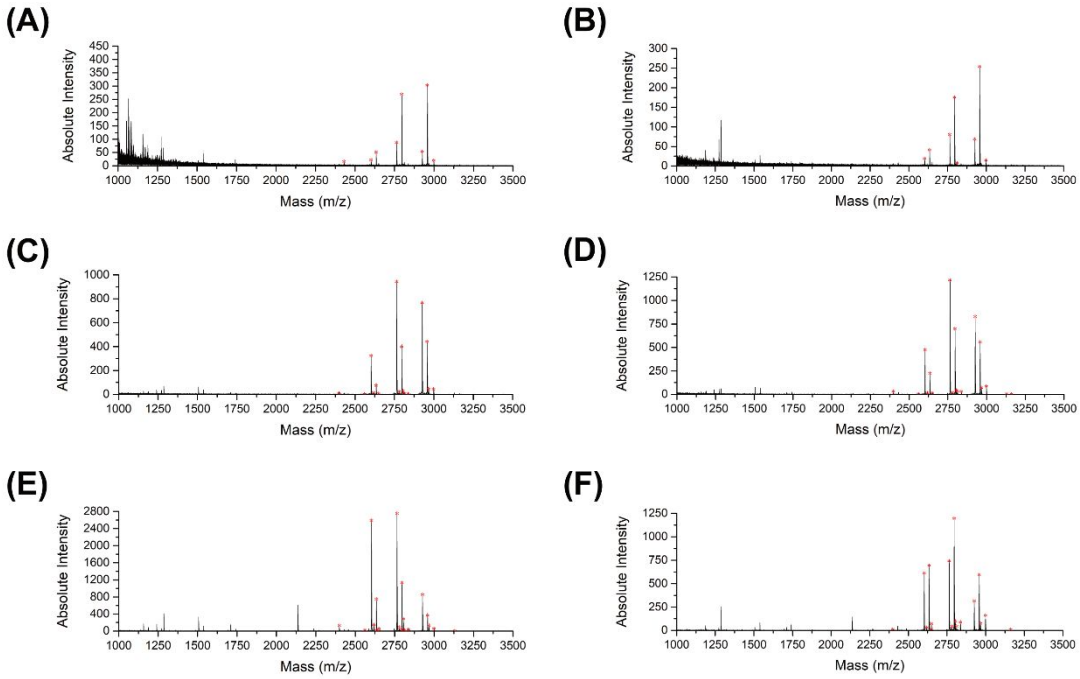


Figure S2 MALDI-TOF MS spectra of IgG tryptic digests after enrichment by GO@mSiO_2-GLYMO-APB with different binding buffers and different washing times. (A) 90% ACN/1% TFA (v/v), 3 washes, (B) 90% ACN/1% TFA (v/v), 6 washes, (C)95% ACN/1% TFA (v/v), 3 washes, (D) 95% ACN/1% TFA (v/v), 6 washes. The N-glycopeptides are marked with red*.

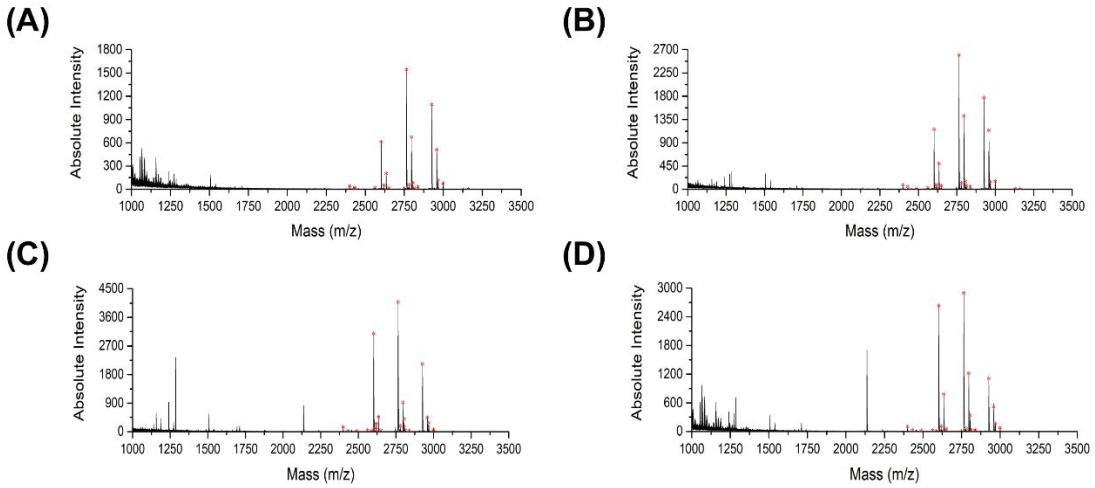


Figure S3 MALDI-TOF MS spectra of 6 μg IgG tryptic digests after enrichment by different amounts of GO@mSiO$_2$-GLYMO-APB composites: (A) 6 μg, (B) 15 μg, (C) 30 μg, (D) 60 μg, (E) 150 μg, (F) 300 μg, (G) 600 μg, and (H) 900 μg. The N-glycopeptides are marked with red*.

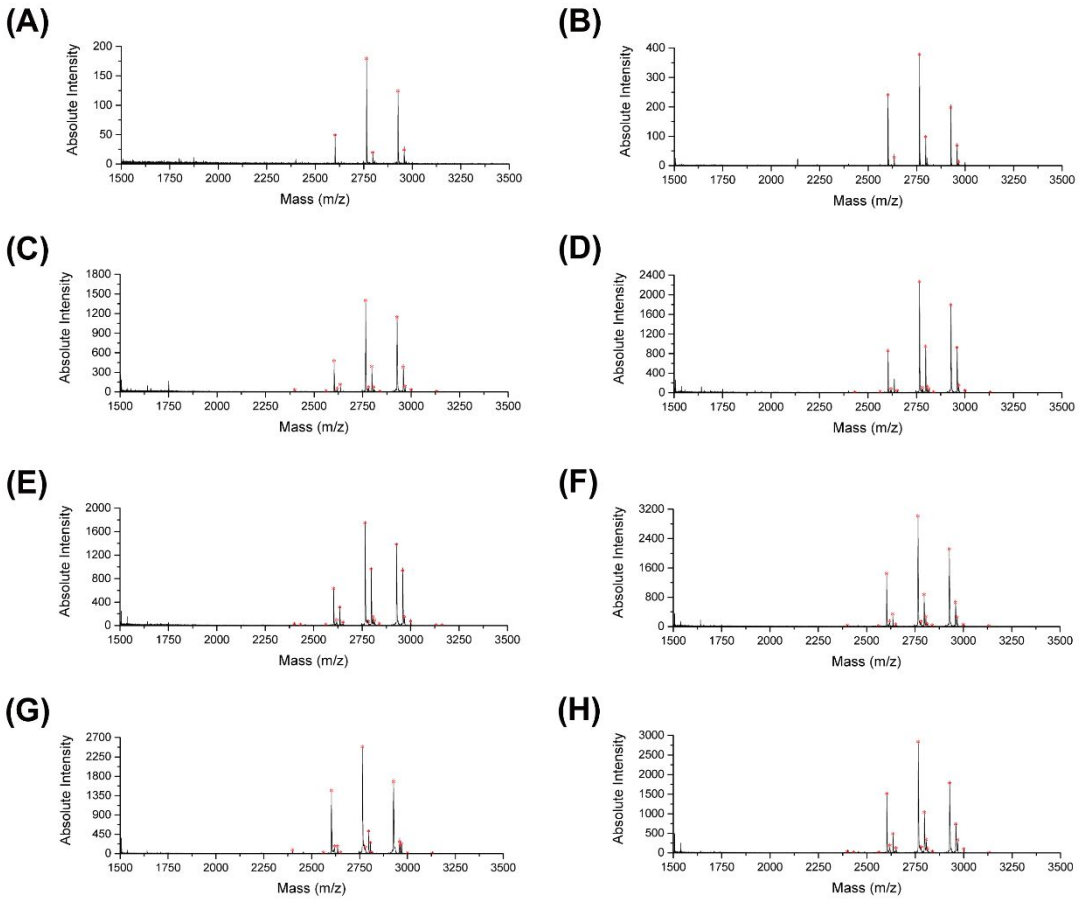


Figure S4 MALDI-TOF MS spectra of different amounts of IgG tryptic digests after enrichment by GO@mSiO$_2$-GLYMO-APB: (A) 20 pmol, (B) 10 pmol, (C) 1 pmol, (D) 100 fmol, (E) 10 fmol and (F) 1 fmol IgG digests. The N-glycopeptides are marked with red*.

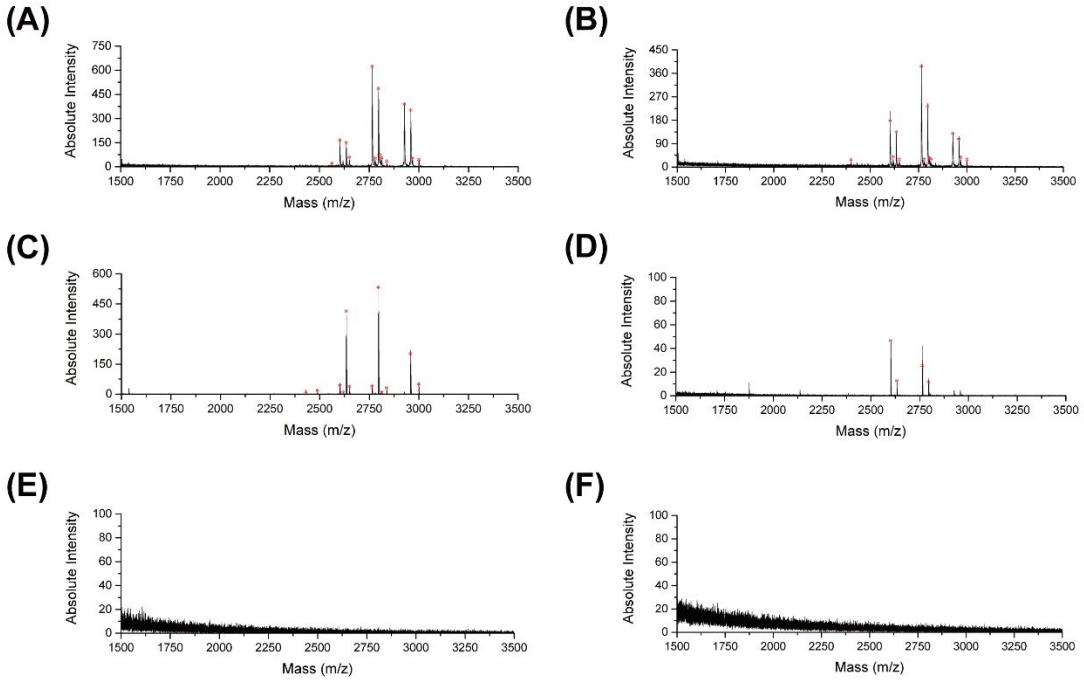


Figure S5 MALDI-TOF MS spectra of a mixture of IgG and BSA tryptic digests with different mass ratios after enrichment by GO@mSiO$_2$-GLYMO-APB: (A) 1:5, (B) 1:50, (C) 1:100, (D) 1:200. The N-glycopeptides are marked with red*.

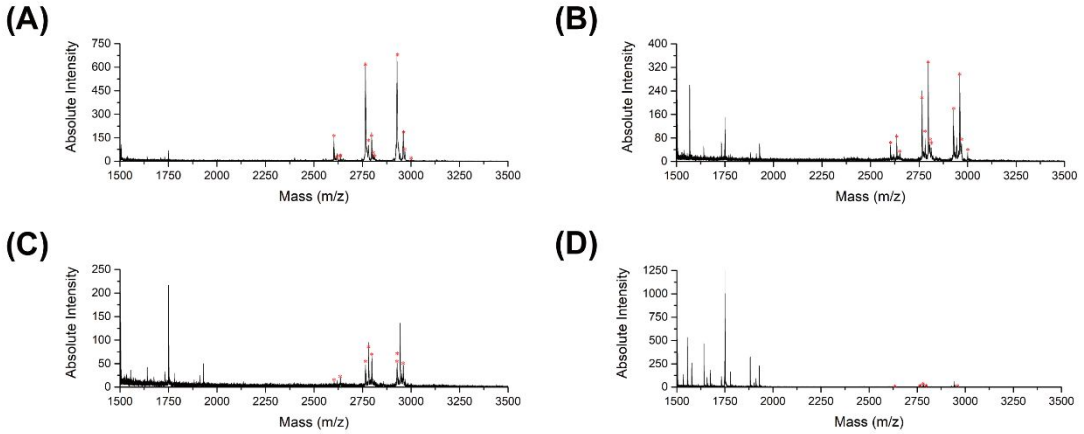
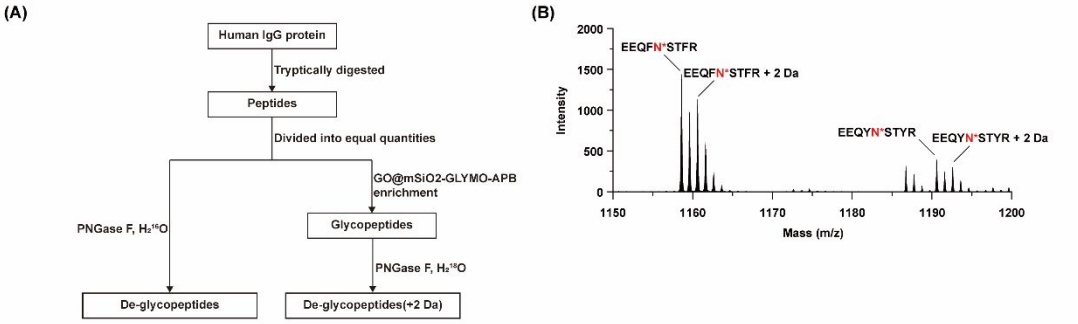


Figure S6 Estimation of GO@mSiO$_2$-GLYMO-APB-based enrichment recovery rate. (A) The workflow of recovery rate experiment. (B) MALDI-TOF mass spectrum of the de-glycopeptide mixtures. (C) The calculation of recovery rate.



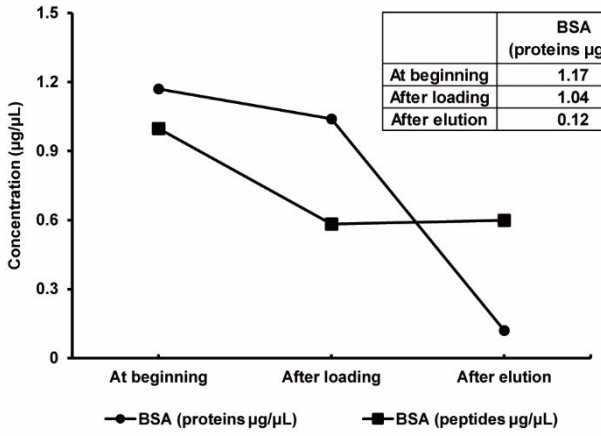
(C)

De-glycopeptides	m/z	Intensity					Recovery rate	RSD	Average recovery rate
		Replicate 1	Replicate 2	Replicate 3	Replicate 4	Average			
EEQFN*STFR	1157.2	722	2709	1408	2664	1876	76.3%	2.8%	76.8%
EEQFN*STFR# (+2 Da)	1159.2	566	1989	1104	1994	1413			
EEQYN*STYR	1189.2	320	1389	383	1613	926	77.4%	0.9%	
EEQYN*STYR# (+2 Da)	1191.2	258	1061	290	1238	712			

N* denotes the N-linked glycosylation sites.
# denotes isotopically labeled de-glycopeptides

$$\text{Recovery rate} = \frac{\text{Intensity of de-glycopeptide (+2 Da)}}{\text{Intensity of de-glycopeptide}} \times 100\%$$

Figure S7 Validation of size exclusion functionality



	BSA (proteins μg/μL)	BSA (peptides μg/μL)
At beginning	1.17	0.998
After loading	1.04	0.583
After elution	0.12	0.599

Figure S8 Annotated tandem mass spectra of 2 standard glycopeptides, ADKDESTMIN[Hex5HexNAc4NeuAc2]EITLAK and QN[Hex5HexNAc4NeuAc2]CELFEQLGEYK enriched by GO@mSiO$_2$ (A, B) and GO@mSiO$_2$-GLYMO-APB (C, D).

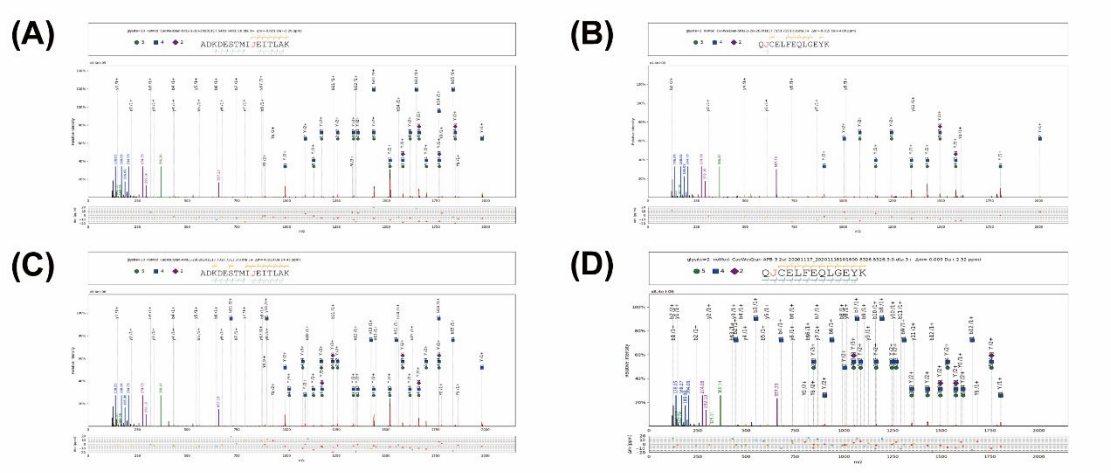


Figure S9 MALDI-TOF MS spectra of IgG tryptic digests after enrichment by GO@mSiO$_2$. The glycopeptides are marked with red number.

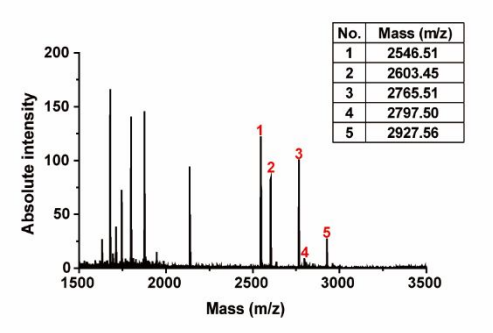
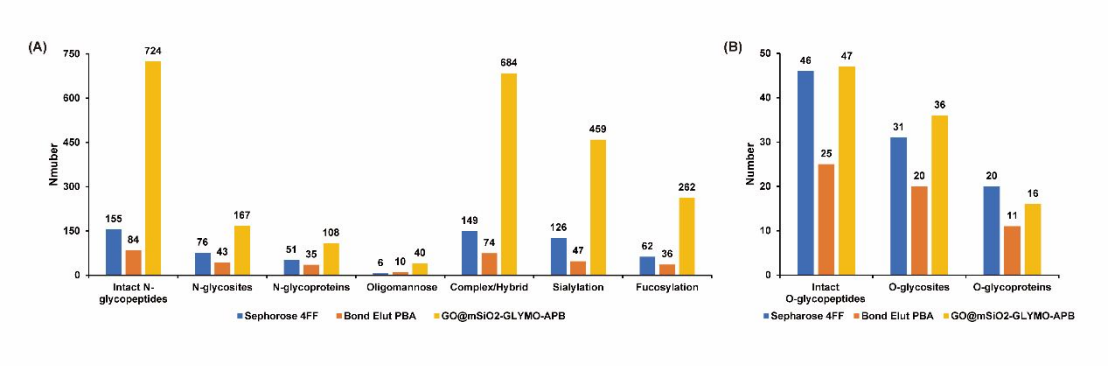


Figure S10 Comparison between GO@mSiO$_2$-GLYMO-APB, Sepharose 4FF and Bond Elut PBA on N- and O- glycopeptides from human serum. (A) The number of identified N-linked intact glycopeptides, glycosites, glycoproteins and the distribution of N-glycopeptides in human serum. N-glycopeptides are sorted into 4 groups (oligomannose, complex/hybrid, sialylation and fucosylation) (B) The number of identified O-linked intact glycopeptides, glycosites, glycoproteins.



Supporting tables

Table S1 Intact N-glycopeptides identified by MALDI-TOF MS without enrichment from human IgG tryptic digests.

Table S2 Intact N-glycopeptides identified by MALDI-TOF MS after enrichment by GO@mSiO$_2$-GLYMO-APB from human IgG tryptic digests.

Table S3 Intact O-glycopeptides identified by LC-MS/MS after enrichment by different amounts of GO@mSiO$_2$-GLYMO-APB from 25 μg fetuin digests.

Table S4 Intact O-glycopeptides identified by LC-MS/MS after enrichment by GO@mSiO$_2$-GLYMO-APB from different amounts of fetuin digests.

Table S5 The enrichment specificity of GO@mSiO$_2$ for standard glycopeptides.

Table S6 The enrichment specificity of GO@mSiO$_2$-GLYMO-APB for standard glycopeptides..

Table S7 Intact N-glycopeptides identified by LC-MS/MS from three standard glycoproteins (fetuin, IgG and HRP).

Table S8 Detailed information of intact N-glycopeptides identified by LC-

MS/MS from three standard glycoproteins (fetuin, IgG and HRP).

Table S9 Intact O-glycopeptides identified by LC-MS/MS from fetuin.

Table S10 Detailed information of intact O-glycopeptides identified by LC-MS/MS from fetuin.

Table S11 Intact N-glycopeptides identified by LC-MS/MS from human serum.

Table S12 Detailed information of intact N-glycopeptides identified by LC-MS/MS from human serum.

Table S13 Intact O-glycopeptides identified by LC-MS/MS from human serum.

Table S14 Detailed information of intact O-glycopeptides identified by LC-MS/MS from human serum.

Table S15 The enrichment specificity of GO@mSiO_2-GLYMO-APB for N-glycopeptides in human serum.

Table S16 The enrichment specificity of GO@mSiO_2-GLYMO-APB for O-glycopeptides in human serum.

Table S17 Intact N-glycopeptides identified by LC-MS/MS after enrichment by Sepharose 4FF from human serum.

Table S18 Intact O-glycopeptides identified by LC-MS/MS after enrichment by Sepharose 4FF from human serum.

Table S19 Intact N-glycopeptides identified by LC-MS/MS after enrichment by Bond Elut PBA from human serum.

Table S20 Intact O-glycopeptides identified by LC-MS/MS after enrichment by Bond Elut PBA from human serum.

Table S21 A list of 15 glycoproteins identified to be both N- and O-glycosylated in human serum.