Supporting Information

Dual-Zwitterion Functionalized Ultra-Hydrophilic Metal-Organic Framework with Ingenious Synergy for Enhanced Enrichment of Glycopeptides

Dapeng Li, Jinghui Zhang, Guangshan Xie, Fenfen Ji, Xiaojian Shao, Lin Zhu, Zongwei Cai*

State Key Laboratory of Environmental and Biological Analysis, Department of Chemistry, Hong Kong Baptist University, Hong Kong SAR, P. R. China

*Corresponding author:

Prof. Zongwei Cai

E-mail: zwcai@hkbu. edu.hk (Z.C.).

Phone: +852-34117070. Fax: 34117348.

ORCID: 0000-0002-7013-5547

Hong Kong Baptist University, 224 Waterloo Road, Kowloon Tong, Hong Kong, SAR, P. R. China

EXPERIMENTAL SECTION:

Materials:

Zinc nitrate hexahydrate, 2-methylimidazole (2-MIM), chloroauric acid ($HAuCl_4$), L-Cysteine (Cys), reduced glutathione (GSH), 2, 5-Dihydroxybenzoic acid (DHB), ammonium bicarbonate (NH_4HCO_3), urea (UA), phosphoric acid (H_3PO_4), formic acid (FA), horseradish peroxidase (HRP), human serum immunoglobulin G (human IgG), bovine serum albumin (BSA), trypsin, dithiothreitol (DTT), iodoacetamide (IAA) were purchased from Sigma-Aldrich (St. Louis, MO, USA) without further purification. Peptide-N-glycosidase (PNGase F) was from New England Biolabs (Ipswich, MA, USA). Deionized water was purified with a Milli-Q system (Millipore, Bedford, MA, USA). HPLC-grade acetonitrile (ACN) and methanol (MeOH) were from VWR Chemicals BDH (Visalia, CA, USA). MEM medium, fetal bovine serum (FBS), and sodium pyruvate used for HepG2 cell were from Gibco (Thermo Fisher Scientific, MA, USA).

Characterization:

Transmission electron microscope (TEM) images were obtained by Tecnai G2 20 S-TWIN. FE-scanning electron microscope (FE-SEM) images were carried out by LEO 1530, operating at 20 kV. Powder X-ray diffraction (XRD) patterns were recorded by Bruker D8 Advance diffractometer nickel-filtered Cu Kα radiation (λ = 1.5406 Å) at 40 kV and 40 mA. X-ray photoelectron spectroscopy (XPS) spectra were obtained from an SKL-12 X. Fourier transforms infrared spectroscopy (FT-IR) spectra were recorded on a Perkin Elmer Spectrometer. Thermogravimetric analysis (TGA) was performed on a PE thermal analyzer with a ramp rate of 10 °C min^{-1} from 30 up to 900 °C. Fluorescence spectra were characterized by Horiba FluoroMax-4 SpectroFluorometer.

Synthesis of AuGC/ZIF-8 nanocomposite

The AuGC nanoclusters (NCs) were prepared by a method similar to previous report[1] with modification. Briefly, freshly prepared aqueous solutions of $HAuCl_4$ (20 mM, 0.50

mL), GSH (100 mM, 0.15 mL) and Cys (50 mM, 0.15 mL) were mixed with 4.20 mL of deionized water at room temperature and then heated to 70 °C to react for 24 h under gentle stirring. The AuGC solution could be stored at 4 °C for further use. ZIF-8 was generally synthesized according to the previous report[2] with slight modification. Briefly, the zinc nitrate aqueous solution (0.29 g, 2 mL) was poured into the 2-MIM solution (5.67 g, 20 mL), and then the mixed solution was stirred for 30 min under room temperature. The product was collected by centrifuging (16000 g, 10 min), washed with deionized water three times and then vacuum dried in a lyophilizer. Finally, the AuGC/ZIF-8 nanocomposites could be easily obtained by hybridization, where ZIF-8 powder (20 mg) dispersed in AuGC solution (5 mL) and stirred for 3 h. The product was collected, washed with deionized water, and dried in vacuum for further use.

Tryptic Digest of Model Glycoproteins and Proteins of Biological samples

1 mg of HRP or IgG was dissolved in 0.5 mL 25 mM NH_4HCO_3 respectively and denatured at 95 °C for 10 min. After cooling down to room temperature, trypsin was added with the ratio of trypsin: protein at 1:40 (w/w). The mixture was then incubated at 37 °C for 16 h followed with lyophilization and stored at -20 °C for further use.

Cells lysates of HepG2 were prepared in a lysis buffer (8 M urea, 0.1 M NH_4HCO_3, 0.1 % SDS, protease inhibitor cocktail 1 tablet/10 mL) by ultrasonic cell disruptor[3]. After acetone precipitation, the proteins were redissolved and quantified by BCA method. To digest the proteins mixture, DTT (final concentration of 5 mM) was added, and the mixture was incubated at 56 °C for 30 min and alkylated by IAA (final concentration of 15 mM) at 37 °C in the dark for 45 min. Then, the mixture was 10-fold diluted by 50 mM NH_4HCO_3, and trypsin was added with enzyme-protein ratio 1:40 (w/w) to digest proteins into peptides at 37 °C for 24 h. The tryptic digest was desalted by an Oasis HLB cartridge, lyophilized and stored at -80 °C.

The mouse liver proteins were extracted following a procedure previously reported[4]. Briefly, the tissues were homogenized in a lysis buffer (4 % SDS, 0.1 M NH_4HCO_3, 8

M urea) by using a high throughput tissue grinding machine. The homogenates were centrifuged, and then the 10-fold volume of acetone was added to the supernatant to precipitate proteins at -20 °C overnight. The following procedure of proteins digest was similar to HepG2 cell lysates mentioned above.

The Protocol of the Enrichment Process

The workflow of enrichment of glycopeptides was shown in Scheme 1b. Briefly, 100 μg of AuGC/ZIF-8 was dispersed in 100 μL loading buffer (ACN/H_2O/TFA= 95/4.9/0.1, v/v/v) containing 30 ng/μL HRP or IgG tryptic digests. The mixture was incubated at 37 °C for 30 min with gentle shaking. The glycopeptide-captured AuGC/ZIF-8 was washed with washing buffer (ACN/H_2O/FA, 85:15:0.5, v/v/v) three times. After that, enriched glycopeptides were eluted by 15 μL eluent buffer (ACN/H_2O/FA, 30:70:0.5, v/v/v) for 10 min at 25 °C and subsequently analyzed by MALDI-TOF MS.

Glycopeptides enrichment from HepG2 cell or mouse liver was conducted as follows: 500 μg of AuGC/ZIF-8 was added to 200 μg of peptide mixture. The enrichment was gently carried out for 60 min at 37 °C and then washed with 100 μL washing buffer for three times. Glycopeptides were eluted with 3 × 100 μL of eluent and lyophilized. Then, the lyophilized glycopeptides were redissolved in 25 μL 10 mM NH_4HCO_3 solution, followed by addition of 1 μL of PNGase F to deglycosylate at 37 °C for 16 h. The resulting solution was lyophilized and further analyzed by nano-LC-MS/MS.

Recovery Evaluation

The enrichment recovery was investigated by using dimethyl labeling method[5]. Briefly, 50 μg of IgG digests were labeled by light and heavy isotopes, respectively. The heavy-tagged peptides were enriched byAuGC/ZIF-8, and the eluent was mixed with the light-tagged IgG digest. The mixture was further enriched by AuGC/ZIF-8, and the eluted glycopeptides were analyzed by MALDI-TOF MS. Five typical

glycopeptides were selected to calculate the recovery by the intensity ratio of heavy and light tagged glycopeptides.

Mass Spectrometry Analysis

All MALDI-TOF MS analysis was performed on a Bruker Rapiflex MALDI-TOF MS (Bruker, Daltonics, Germany). 1 μL of eluent was mixed with 1 μL of DHB matrix (30 mg/mL, ACN/H_2O/TFA, 70:29:1, v/v/v/) on the steel plate and dried for MS analysis. MALDI-TOF MS spectra were obtained in the positive ion reflector mode with a 337 nm nitrogen laser at 20 kV of accelerating voltage. The mass ranges from m/z 1000 to 5500 were recorded under the constant laser intensity. The results were analyzed by the FlexAnalysis software (version 4.0) at the signal-to-noise (S/N) ratio of 3:1 as the lowest limit of detection.

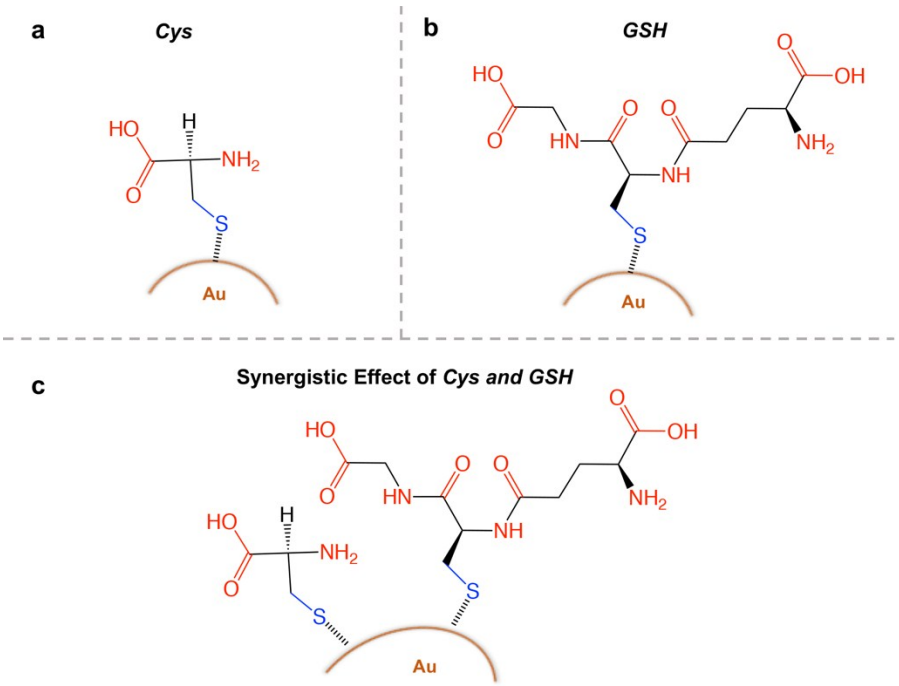
All LC-MS/MS was performed on an Ultimate 3000 Nano-LC system equipped with Orbitrap Fusion™ Tribrid™ Mass Spectrometer (Thermo, San Jose, CA) using a nano-ESI ion source. The samples were redissolved in 0.1% formic acid, loaded onto a precolumn for 10 min, and separated on a C18 capillary column. The mobile phase A (0.1% FA in H_2O) and B (ACN) were used for gradient separation, 0 min, 2% B; 10 min, 5 % B; 90 min, 23 % B; 97 min, 35 % B; 102 min, 80 % B; 107 min, 80 % B; 112 min, 2 % B; 120 min, 2 % B. The flow rate was 300 nL/min. The spray voltage was set at 2.2 kV. All MS and MS/MS spectra were acquired in the data-dependent acquisition mode, and the full mass scan was acquired from m/z 350 to 1550 with a resolution of 60,000.

Database Searching

The raw LC-MS/MS data files were analyzed by Thermo Scientific Proteome Discoverer software version 1.4.1 with the Swiss-Prot database. The searching parameters were set as follows: precursor mass tolerance: 20 ppm; fragment mass tolerance: 0.02 Da; enzyme: trypsin; missed cleavages: two; fixed modification: carbamidomethyl on cysteine (C, +57.021 Da); variable modifications: oxidation on

methionine (M, +15.995 Da) and deamidation on asparagine (N, +0.984 Da). The percolator algorithm was used to control the false discovery rate (FDR) of peptides below 1%. Only peptides with N-X-S/T (X ≠ Proline) conserved sequence were defined as highly reliable glycopeptides.

Supplementary Figures:



Scheme S1. Single Cys modification on the surface of gold nanocluster (AuNC) with relatively low ratio of hydrophilic groups (a); single GSH modification on the suface of AuNC with relatively large steric hindrance caused by the long space arms (b); Cys and GSH dual functionalized AuNC with synergistic effect of ultra-hydrophilic affinity and low steric hindrance for highly efficient glycopeptides capturing (c).

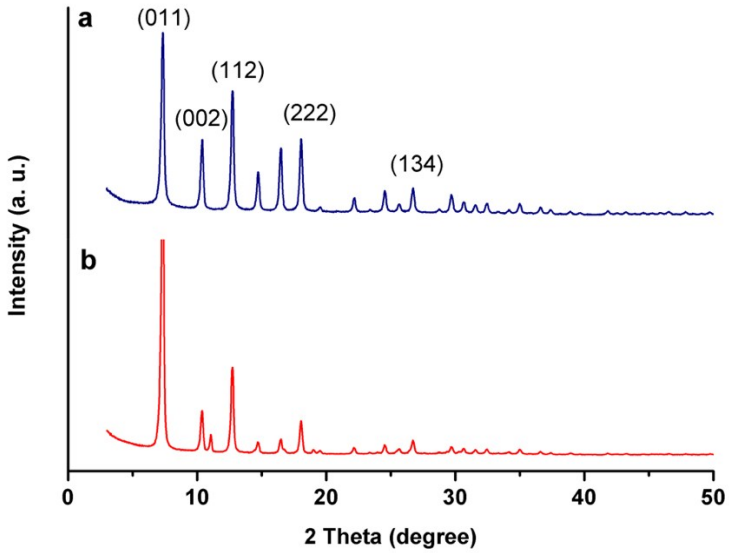


Figure S1. Wide angle XRD pattern of pure ZIF-8 (a) and AuGC/ZIF-8 nanocomposite (b).

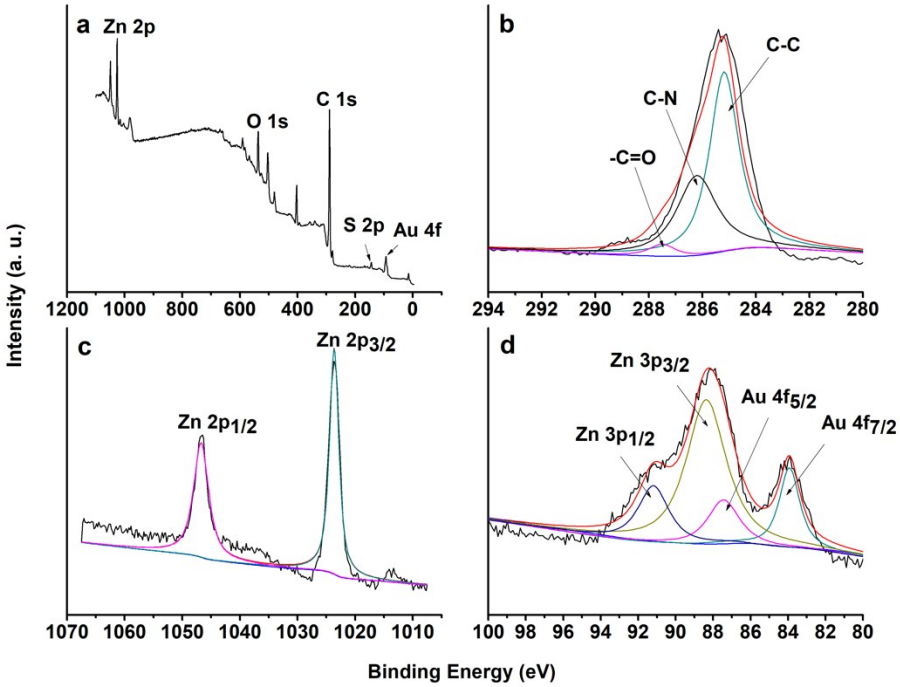


Figure S2. XPS patterns of AuGC/ZIF-8: wide scan XPS spectrum (a) and high-resolution XPS spectrum of C 1s (b), Zn 2p (c), and Au 4f (d).

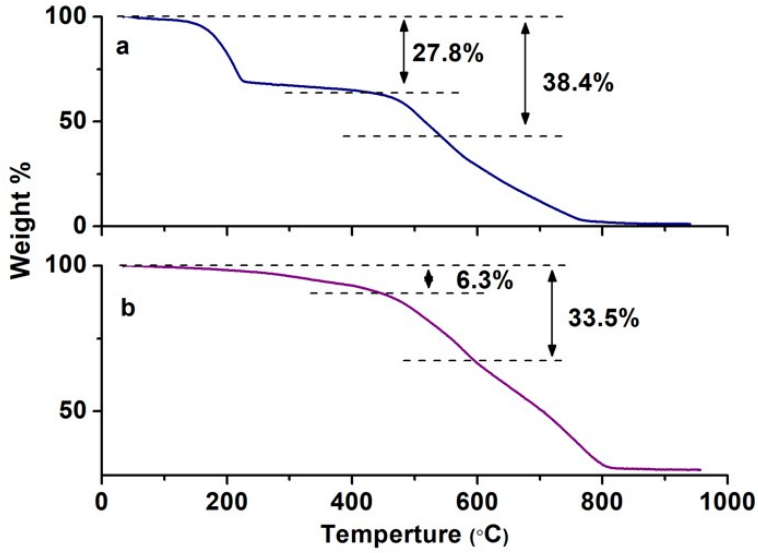


Figure S3. TGA curves of ZIF-8 (a) and AuGC/ZIF-8 (b). The weight loss of the ZIF-8 framework collapse occurred at ≈550 °C, while the AuGC/ZIF-8 nanocomposites decompose at ≈ 600 °C, suggesting AuGC/ZIF-8 nanocomposites was of high thermal stability. the weight loss of ZIF-8 at ≈ 200 °C is 27.8%, while only part of weight loss was of ≈ 6.3% in AuGC/ZIF-8, indicating the loss of molecules from the cavities of ZIF-8 is hampered by the assembly of AuGC.

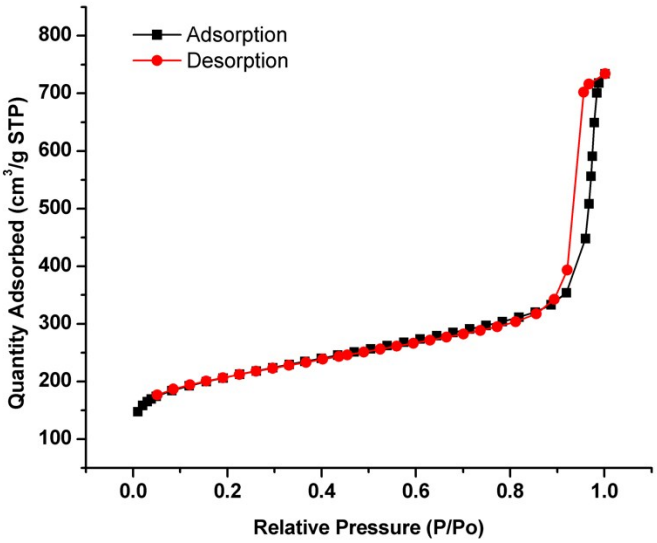


Figure S4. N_2 adsorption–desorption isotherm of AuGC/ZIF-8 nanomaterials.

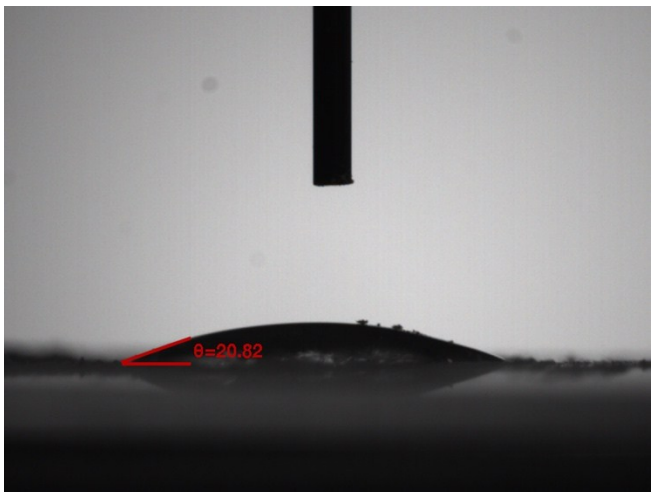


Figure S5. Water contact angle of AuGC/ZIF-8 nanomaterials.

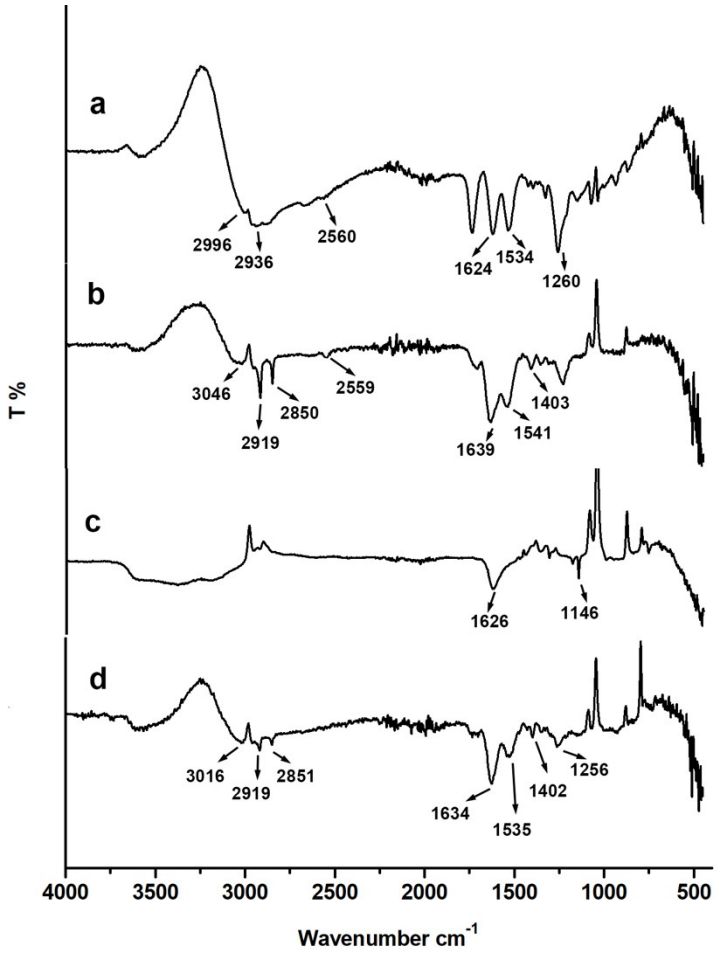


Figure S6. FT-IR fingerprint spectra of Cys (a), GSH (b), ZIF-8 (c), and AuGC/ZIF-8 (d), respectively. 1500-1660 and 1260 cm^{-1} corresponding to COO-, 1403 cm^{-1} corresponding to C-N, 2550-2750 cm^{-1} corresponding to -SH group, 2900-3420 cm^{-1} corresponding to $-NH_2$. The disappearance of 2550 cm^{-1} in AuGC/ZIF-8 attributed to the formation of covalent bonds between Cys/GSH and golden nanoclusters.

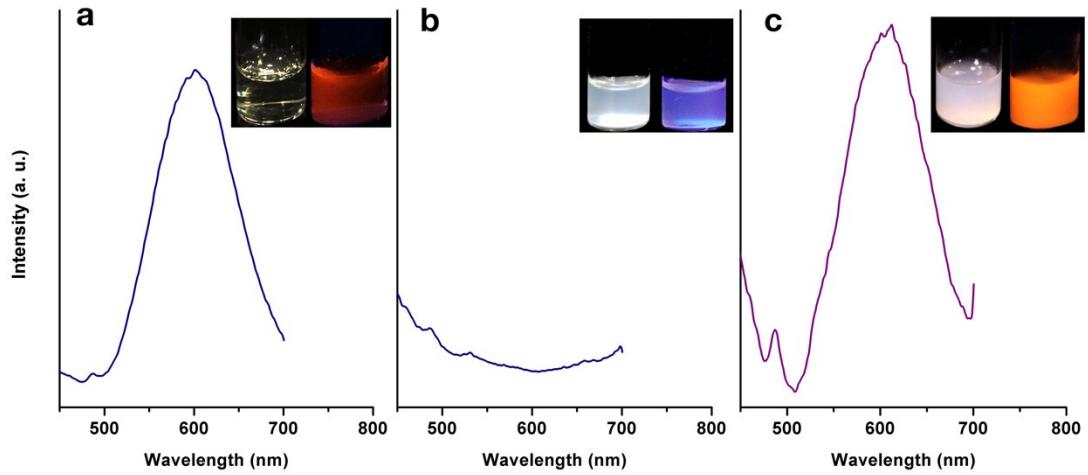


Figure S7. Fluorescent spectra of AuGC (a), ZIF-8 (b), and AuGC/ZIF-8 (c) with excitation wavelength at 365 nm (insets: photographs of the corresponding samples suspended in ethanol aqueous solution illuminated under ambient light (left) and under UV light (365 nm) (right)).

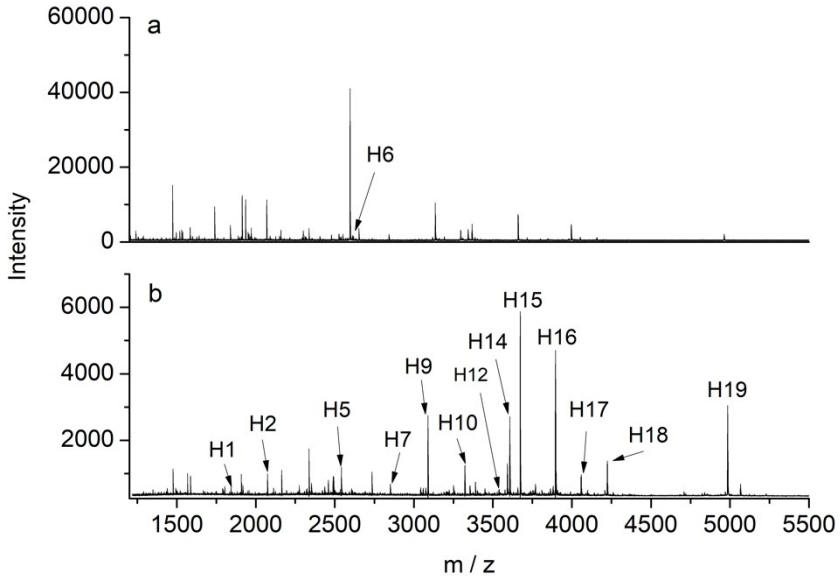


Figure S8. MALDI-TOF MS spectra of HRP tryptic digest (3 μg): (a) direct analysis, (b) after enrichment by AuGC/ZIF-8.

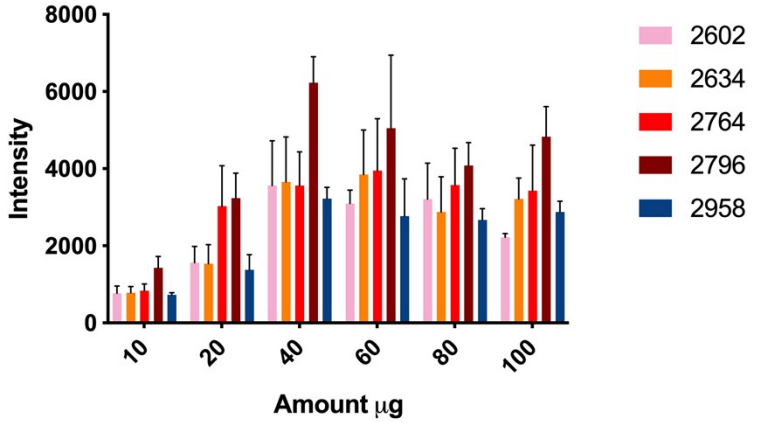


Figure S9. MS intensities of five glycopeptides from 10 μg of IgG tryptic digest after enrichment by various amounts of AuGC/ZIF-8 nanocomposite.

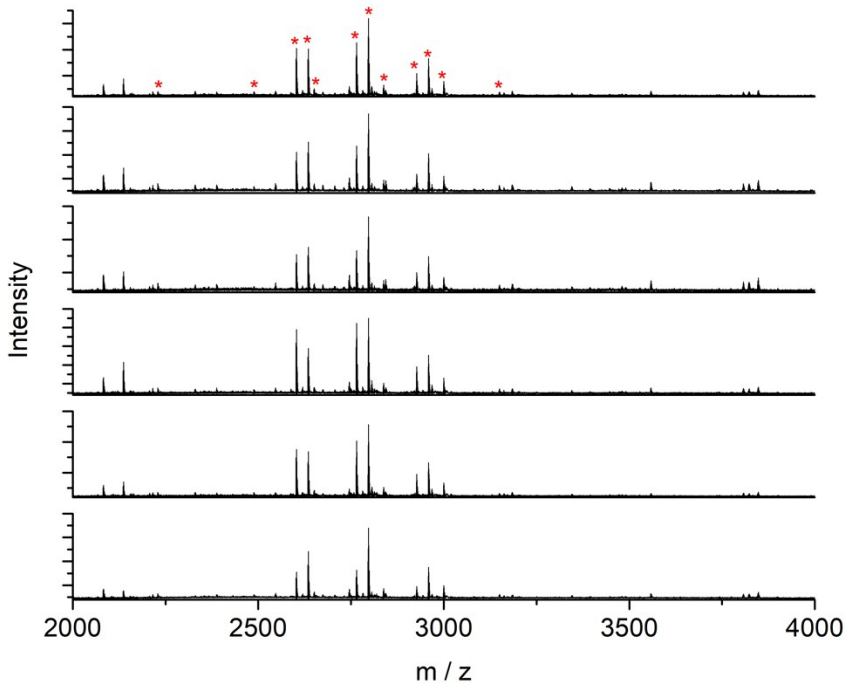


Figure S10. MALDI-TOF MS spectra of tryptic digest mixture of IgG: BSA (1:50, w/w) after enriched by AuGC/ZIF-8 nanocomposite in six different batches.

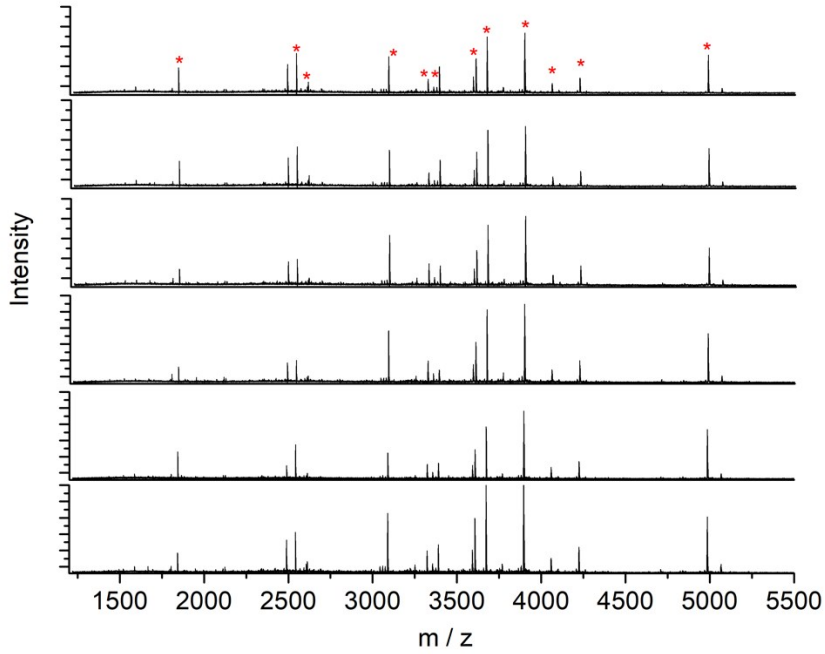


Figure S11. MALDI-TOF MS spectra of tryptic digest mixture of HRP: BSA (1:50, w/w) after enriched by AuGC/ZIF-8 nanocomposite in six different batches.

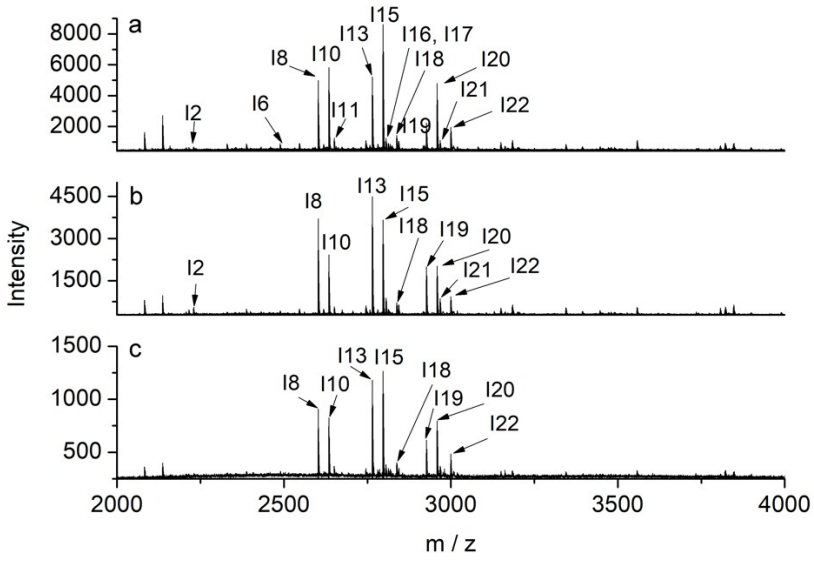


Figure S12. MALDI-TOF MS spectra of the mixture of IgG and BSA tryptic digest with a mass ratio of 1:50 after enriched by AuGC/ZIF-8 (a), AuG/ZIF-8 (b), and AuC/ZIF-8 (c), respectively.

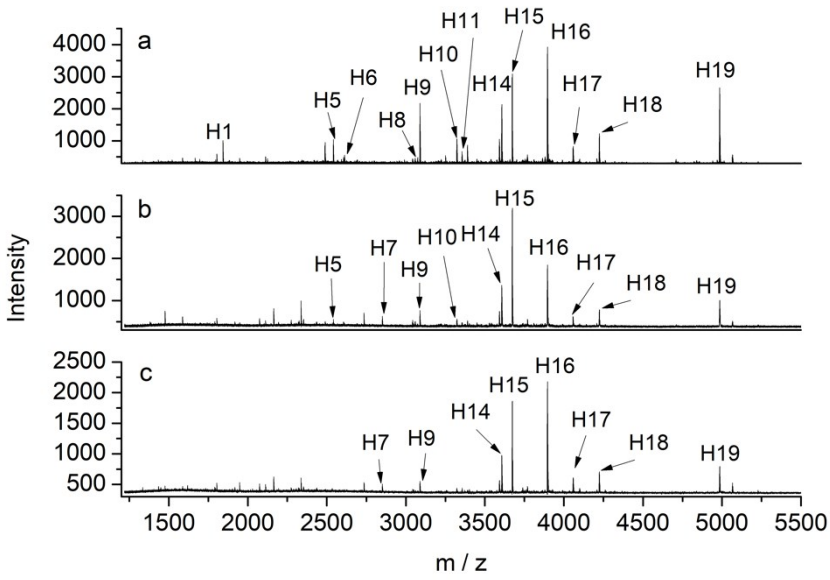


Figure S13. MALDI-TOF MS spectra of the mixture of HRP and BSA tryptic digest with a mass ratio of 1:50 after enriched by AuGC/ZIF-8 (a), AuG/ZIF-8 (b), and AuC/ZIF-8 (c), respectively.

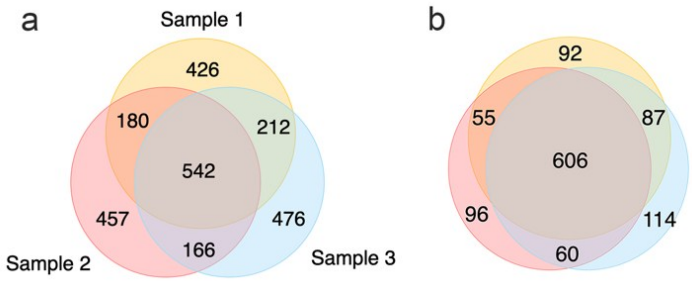


Figure S14. Overlapping of identified glycopeptides (a) and glycoproteins (b) in HepG2 cell proteins after enriched by AuGC/ZIF-8 in three replicates.

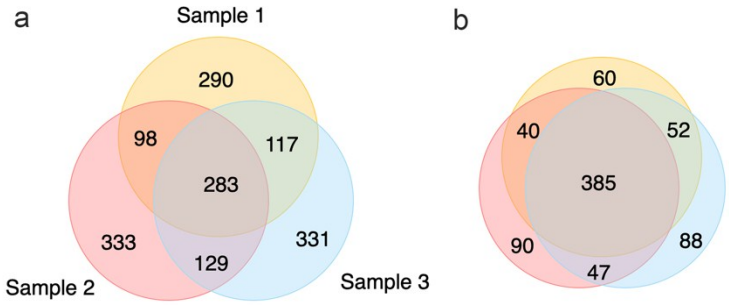


Figure S15. Overlapping of identified glycopeptides (a) and glycoproteins (b) in mouse liver proteins after enriched by AuGC/ZIF-8 in three replicates.

Supplementary Tables:

Table S1 Detailed information of the glycopeptides enriched by AuGC/ZIF-8 from an IgG tryptic digest. N# denotes the N-linked glycosylation site.

NO.	m/z	Glycan composition	Peptides sequence
I1	2209.931	[Hex]4[HexNAc]2	EEQFN#STFR
I2	2228.965	[Hex]3[HexNAc]2[Fuc]1	EEQYN#STYR
I3	2268.911	[Hex]2[HexNAc]3[Fuc]1	EEQYN#STYR
I4	2430.796	[Hex]3[HexNAc]3[Fuc]1	EEQYN#STYR
I5	2455.979	[Hex]3[HexNAc]4	EEQFN#STFR
I6	2487.817	[Hex]3[HexNAc]4	EEQYN#STYR
I7	2594.064	[Hex]4[HexNAc]3[Fuc]1	EEQYN#STYR
I8	2601.873	[Hex]3[HexNAc]4[Fuc]1	EEQFN#STFR
I9	2617.852	[Hex]3[HexNAc]4[Fuc]1, or [Hex]4[HexNAc]4	EEQFN#STYR, or EEQFN#STFR
I10	2633.856	[Hex]3[HexNAc]4[Fuc]1, or [Hex]4[HexNAc]4	EEQYN#STYR, or EEQFN#STYR
I11	2649.849	[Hex]4[HexNAc]4	EEQYN#STYR
I12	2755.100	[Hex]5[HexNAc]3[Fuc]1	EEQYN#STYR
I13	2763.899	[Hex]4[HexNAc]4[Fuc]1	EEQFN#STFR
I14	2779.882	[Hex]4[HexNAc]4[Fuc]1, or [Hex]5[HexNAc]4	EEQFN#STYR, or EEQFN#STFR
I15	2795.885	[Hex]4[HexNAc]4[Fuc]1, or [Hex]5[HexNAc]4	EEQYN#STYR, or EEQFN#STYR
I16	2804.912	[Hex]3[HexNAc]5[Fuc]1	EEQFN#STFR
I17	2811.855	[Hex]5[HexNAc]4	EEQYN#STYR
I18	2836.900	[Hex]3[HexNAc]5[Fuc]1, or [Hex]4[HexNAc]5	EEQYN#STYR or EEQFN#STYR
I19	2925.923	[Hex]5[HexNAc]4[Fuc]1	EEQFN#STFR
I20	2957.912	[Hex]5[HexNAc]4[Fuc]1	EEQYN#STYR
I21	2966.921	[Hex]4[HexNAc]5[Fuc]1	EEQFN#STFR
I22	2998.93	[Hex]4[HexNAc]5[Fuc]1, or [Hex]5[HexNAc]5	EEQYN#STYR, or EEQFN#STYR
I23	3056.245	[Hex]4[HexNAc]4[Fuc]1[NeuAc]1	EEQFN#STFR
I24	3086.367	[Hex]4[HexNAc]4[Fuc]1[NeuAc]1	EEQYN#STYR
I25	3160.941	[Hex]5[HexNAc]5[Fuc]1	EEQYN#STYR
I26	3451.518	[Hex]5[HexNAc]5[Fuc]1[NeuAc]1	EEQYN#STYR

Table S2 Detailed information of the glycopeptides enriched by AuGC/ZIF-8 from an HRP tryptic digest. N# denotes the N-linked glycosylation site.

NO.	m/z	Glycan composition	Peptides sequence
H1	1842.841	[Xyl]1[Hex]3 [Fuc]1[HexNAc]2	NVGLN#R
H2	2073.113	[Xyl]1[Hex]3 [Fuc]1[HexNAc]2	PN#VSNIVR
H3	2273.156	[Xyl]1[Hex]2 [Fuc]1[HexNAc]2	SILLDN#TTSFR
H4	2322.100	[Hex]2 [HexNAc]2	MGN#ITPLTGTQGQIR
H5	2541.128	[Xyl]1[Hex]3[Fuc]1[HexNAc]2	SSPN#ATDTIPLVR
H6	2611.216	[Xyl]1[Hex]3[[HexNAc]2	MGN#ITPLTGTQGQIR
H7	2850.382	[Fuc]1[HexNAc]1	GLIQSDQELFSSPN#ATDTIPLVR
H8	3073.335	[Fuc]1[HexNAc]1	LHFHDCFVNGCDASILLDN#TTSFR
H9	3088.337	[Xyl]1[Hex]3[Fuc]1[HexNAc] 2	GLCPLNGN#LSALVDFDLR
H10	3321.343	[Xyl]1[Hex]3[Fuc]1[HexNAc]2	QLTPTFYDNSCPN#VSNIVR
H11	3354.298	[Xyl]1[Hex]3[Fuc]1[HexNAc]2	SFAN#STQTFFNAFVEAMDR
H12	3372.248	[Xyl]1[Hex]3[Fuc]1[HexNAc]2	SFAN#STQTFFNAFVEAM*DR
H13	3537.497	[Hex]3[Fuc]1[HexNAc]2	GLIQSDQELFSSPN#ATDTIPLVR
H14	3605.486	[Xyl]1[Hex]3[Fuc]1[HexNAc]2	NQCRGLCPLNGN#LSALVDFDLR
H15	3671.579	[Xyl]1[Hex]3[Fuc]1[HexNAc]2	GLIQSDQELFSSPN#ATDTIPLVR
H16	3894.496	[Xyl]1[Hex]3[Fuc]1[HexNAc]2	LHFHDCFVNGCDASILLDN#TTSFR
H17	4056.530	[Xyl]1[Hex]3[HexNAc]2	QLTPTFYDNSC(AAVESACPR)PN#VSNIVR-H_2O
H18	4221.659	[Xyl]1[Hex]3[Fuc]1[HexNAc]2	QLTPTFYDNSC(AAVESACPR)PN#VSNIVR
H19	4982.884	[Xyl]1[Hex]3[Fuc]1[HexNAc]2, or [Xyl]1[Hex]3[Fuc]1[HexNAc]2	LYN#FSNTGLPDPTLN#TTYLQTLR

Table S3 Recovery of five selected glycopeptides from IgG digest.

m/z	Recovery±S.D. (%, n=3)	Average recovery± S.D (%)
2602	84.94±7.91	
2634	85.84±6.74	
2764	88.20±3.47	87.50±5.39
2796	90.97±3.16	
2958	87.53±5.68	

Separate PDF Files:

Table S4 Identified N-glycoproteins and N-glycopeptides sequence from the tryptic digest of proteins extracted from HepG2 cells after enrichment by AuGC/ZIF-8. (A separate PDF file)

Table S5 Identified N-glycoproteins and N-glycopeptides sequence from the tryptic digest of proteins extracted from mouse liver after enrichment by AuGC/ZIF-8. (A separate PDF file)

REFERENCES

1 Z. Luo, X. Yuan, Y. Yu, Q. Zhang, D. T. Leong, J. Y. Lee and J. Xie, *J. Am. Chem. Soc.*, 2012, **134**, 16662–16670.

2 Y. Luo, S. Fan, W. Yu, Z. Wu, D. A. Cullen, C. Liang, J. Shi and C. Su, *Adv. Mater.*, 2018, **30**, 1704576.

3 Q. Wang, F. Zhao, X. Sun, S. Sun, H. Yu, Y. Zhong, Y. Hu, Z. Chen, P. Zhang and Z. Li, *J. Glycomics Lipidomics*, 2015, **05**, 1–10.

4 Y. Xue, J. Xie, P. Fang, J. Yao, G. Yan, H. Shen and P. Yang, *Analyst*, 2018, **143**, 1870–1880.

5 W. Ma, L. Xu, X. Li, S. Shen, M. Wu, Y. Bai and H. Liu, *ACS Appl. Mater. Interfaces*, 2017, **9**, 19562–19568.