Electronic Supplementary Material (ESI) for Analyst.

Supporting Information

Highly efficient enrichment of N-glycopeptides by two-dimensional Hf-based metal-organic framework nanosheets

Shi-Shu Yang,[†,a] Chen Wang,[†,a] Jing Xiao,[†] Xizhong Yu,[§] Wenbin Shang,[§] David Da Yong Chen[†,‡,*] and Zhi-Yuan Gu[†,*]

[†]*Jiangsu Key Laboratory of Biofunctional Materials, Jiangsu Collaborative Innovation Center of Biomedical Functional Materials, College of Chemistry and Materials Science, College of Life Sciences, Nanjing Normal University, Nanjing 210023, China*

[§]*Key Laboratory for Metabolic Diseases in Chinese Medicine, First College of Clinical Medicine, Nanjing University of Chinese Medicine, Nanjing 210023, China*

[‡]*Department of Chemistry, University of British Columbia, Vancouver, BC, Canada, V6T 1Z1*

[*]*To whom correspondence should be addressed. E-mail: chen@chem.ubc.ca; guzhiyuan@njnu.edu.cn*

Chemicals and Instrumentation

Acetonitrile (ACN) was obtained from Merck (Darmstadt, Germany). Peptide-N-glycosidase (PNGase F), trifluoroacetic acid (TFA), horseradish peroxidase (HRP), hafnium chloride ($HfCl_4$) and trypsin (TPCK treated) were purchased from Sigma-Aldrich (St. Louis, MO). Ammonium bicarbonate (NH_4HCO_3), bovine serum albumin (BSA, from bovine milk), hafnium oxide (HfO_2), formic acid (FA) and ammonia solution (ACS, 28.0-30.0% NH_3 basis) were obtained from Aladdin (Shanghai, China). 2,5-dihydroxybenzoic acid (DHB) was obtained from TCI (Shanghai, China). 1,3,5-(4-carboxyphenyl)-benzene (H_3BTB) was purchased from Henghua Sci. & Tec. Co., Ltd. (Jinan, China). N,N-dimethylformamide (DMF) was obtained from Sinopharm Chemical Reagent Co., Ltd (Shanghai, China). Ultrapure water (18.4 MΩ·cm) was prepared by using an ELGA purification system (Veolia Water Solutions & Technologies, UK).

X-ray diffraction (XRD) patterns were obtained using a Rigaku D/MAX-2500 diffractometer (Tokyo, Japan) with CuKα radiation (1.54056 Å). Transmission electron microscopy (TEM) images were collected on an H7650 TEM (JEOL Ltd., Japan) operated at an accelerating voltage of 120 kV. Atomic force microscopy (AFM) measurements were performed using a PicoPlus AFM in tapping mode (Agilent, US). Scanning electron microscopy (SEM) images were collected on a JSM-7600F SEM (JEOL Ltd., Japan).

Synthesis of Hf-BTB nanosheets[S1]

2-D Hf-BTB nanosheets were synthesized via a solvothermal method. Briefly, $HfCl_4$ (14 mg) and H_3BTB (12.5 mg) were dissolved in a mixture of dimethylformamide (DMF) (5 mL), HCOOH (1.113 g) and water (120 μL) in a Pyrex vial. The vial was then heated at 120 °C for 48 h in an oven. The 2-D Hf-BTB nanosheets were collected by centrifugation, and washed three times with DMF and ethanol, respectively. The 2-D Hf-BTB nanosheets were then dried at 80 °C for 8 h in a vacuum oven.

Sample Preparation

Glycoprotein HRP standard (1 mg) was dissolved in a solution of ammonium bicarbonate (50 mM) to give a final concentration of 1 μg μL^{-1}. The resulting solution was boiled at 100 ºC for 10 min and then cooled to room temperature. Trypsin was added to give a 50:1 ratio for protein:trypsin. The mixture was incubated in a water bath at 37 ºC for 20 h. Finally, the sample was lyophilized at -40 ºC and stored at -20 ºC .

Human serum samples were collected from patients in Jiangsu Province Hospital of TCM (Affiliated Hospital of Nanjing University of Chinese Medicine) according to their standard clinical procedures. All human serum samples were obtained with patient consent, and the handling of human serum samples was performed in observance with all relevant laws and institutional guidelines. In addition, ethical approval was obtained from the Institutional Committee.

Standard Samples

Hf-BTB nanosheets were synthesized according to previously reported procedures.[S1] The enrichment procedure included preconditioning, sample loading, washing, and elution. Briefly, Hf-BTB nanosheets (0.5 mg) were first washed three times with loading buffer (100 μL, ACN/H_2O/FA = 85/14.9/0.1, v/v/v). Second, peptide mixture (1 μL) was incubated with Hf-BTB nanosheets for 10 min in loading buffer. Third, unadsorbed peptides were washed away with 3 rinses of loading buffer. Finally, adsorbed peptides were eluted from the nanosheets with eluting buffer (10 μL, ACN/H_2O/FA = 30/69.9/0.1, v/v/v) and analyzed by MALDI-TOF MS.

The procedure for the anti-interference experiment was the same as above. BSA was added to obtain the anti-interference capacity of the Hf-BTB nanosheets. The mixtures (HRP digest and BSA digest at molar ratios of 1:10, 1:100 and 1:1000) were enriched and the eluent was collected for analysis by MALDI-TOF MS.

For the analysis by MALDI-TOF MS, the eluent (1 μL) was first deposited on the MALDI plate. Second, it was covered with 2,5-DHB matrix aqueous solution (1 μL, 20 mg mL^{-1} DHB in 1% H_3PO_4 aqueous solution). Glycopeptides were analyzed by MALDI-TOF MS (Bruker Daltonics, Germany) in positive-ion mode with Smartbeam™-II laser technology. The measurements were made in reflection mode with a shot number of 500 and an acceleration voltage of +25 kV.

Real Biological Samples

Digested human serum samples were enriched by using Hf-BTB nanosheets. The eluent obtained was deglycosylated by a solution of deionized water (17 μL), 10×G7 reaction buffer (2 μL, 0.5 M Na_3PO_4) and PNGase F (1 μL, 500 units/mL), and then incubated at 37 ºC for 18 h. The deglycosylated peptides were analyzed by LC-MS/MS.

For LC MS/MS detection, enriched glycopeptides were desalted with StageTip C18 columns. RPLC-ESI-MS/MS was then used to analyze the samples. LC-MS/MS detection was carried out on a hybrid quadrupole-TOF LC/MS/MS mass spectrometer (TripleTOF 5600+, AB Sciex) equipped

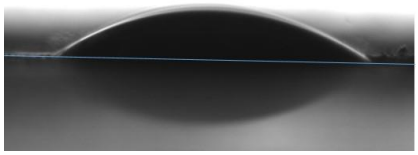
with a nanospray source. Peptides were first loaded onto a C18 trap column (5 μm, 5×0.3 mm, Agilent Technologies) and then eluted into a C18 analytical column (75 μm×150 mm, 3 μm particle size, 100 Å pore size, Eksigent). Mobile phase A (3% DMSO, 97% H_2O, 0.1% formic acid) and mobile phase B (3% DMSO, 97% ACN, 0.1% formic acid) were used to establish a 30 min gradient as follows: 0 min in 5% B, 15 min of 5-35% B, 1 min of 35-80% B, 5 min 80% B, followed by 0.1 min of 80-5% B, and finally 5% B for 8.9 min. A constant flow rate of 300 nL min^{-1} was used. MS scans were conducted from 350 to 1500 amu, with a 250 ms time span. For MS/MS analysis, each scan cycle consisted of one full-scan mass spectrum (m/z 100 to 1500 over 50 ms) followed by 40 MS/MS events. The threshold count was set to 120 cps and charge states were screened from +2 to +5 to activate MS/MS accumulation. Former target ion exclusion was set for 18 s.

In the database search, raw data from TripleTOF 5600+ were analyzed with the Proteome Discoverer software (Thermo Fisher Scientific, version 1.4) with a Sequest HT search engine. Data were searched against the Human UniProtKB/SwissProt database using the following parameters: precursor mass tolerance, 20 ppm; fragment mass tolerance, 0.05 Da. Two missed cleavages were allowed by trypsin digestion. False discovery rates (FDR) were controlled to be lower than 1% by the percolator algorithm at both the peptide and protein level. Carbamidomethyl on cysteine was set as a fixed modification. Oxidation of methionine and deamidation of asparagine were set as variable modifications.

Recovery Experiment[S2, S3]

Stable isotope dimethyl labeling was chosen to determine the recovery yield of the Hf-BTB nanosheets. The same amount digests of HRP were labeled by light isotopes (CH_2O) and heavy isotopes (CD_2O), respectively. The heavy labeled peptides were then enriched by using Hf-BTB nanosheets and the eluent was collected. Next, the light labeled HRP peptides were mixed with the eluent. This mixture was also enriched by Hf-BTB nanosheets. This final eluent was analyzed by MALDI-TOF MS. The recovery (D/H) yield was calculated according to the peak intensities, using the ratio of heavy peptides to light ones.

Table S1. The contact angle of 2-D Hf-BTB nanosheets.

Material	Image	Contact angle
Hf-BTB nanosheets		34.0°

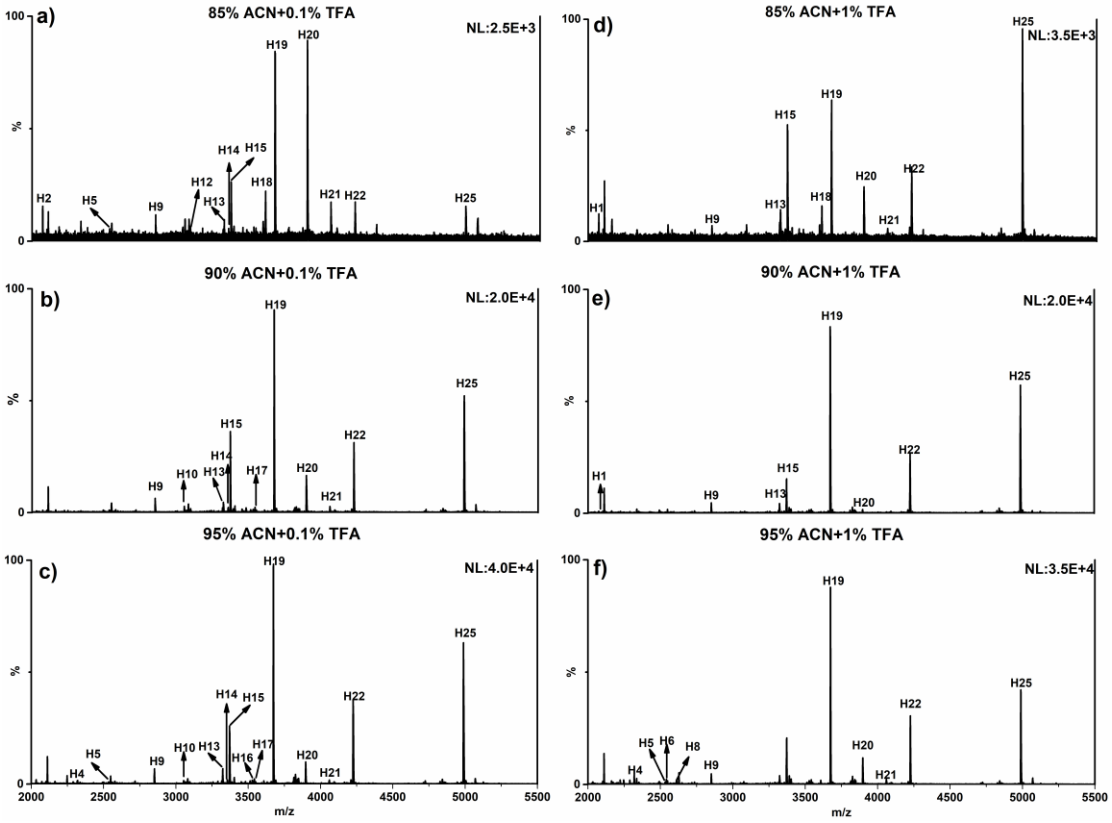


Fig. S1 MALDI-TOF mass spectra of glycopeptides from tryptic digest of HRP (2.5×10^{-6} M) enriched by Hf-BTB nanosheets with different ACN/TFA loading buffer solutions: (a) 85% ACN/0.1% TFA, (b) 90% ACN/0.1% TFA, (c) 95% ACN/0.1% TFA, (d) 85% ACN/1% TFA, (e) 90% ACN/1% TFA, (f) 95% ACN/1% TFA (all unmarked peaks are from non-glycopeptides; H1: No. 1 glycopeptide from HRP; NL: Normalized Level). Glycopeptide peaks identified are marked with the symbol "HX" (X = Arabic numerals). Detected glycopeptides with m/z values and amino acid sequences are listed in Table S2.

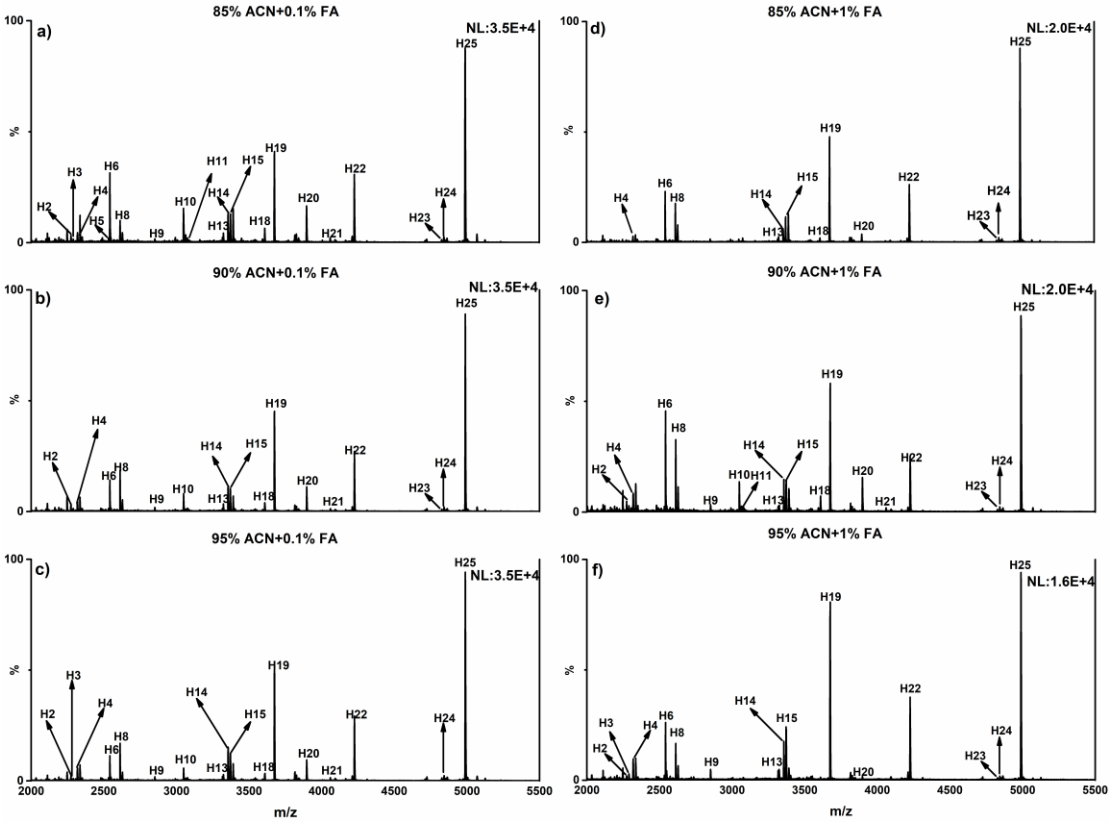


Fig. S2 MALDI-TOF mass spectra of glycopeptides from tryptic digest of HRP (2.5×10^{-6} M) enriched by Hf-BTB nanosheets with different ACN/FA loading buffer solutions: (a) 85% ACN/0.1% FA, (b) 90% ACN/0.1% FA, (c) 95% ACN/0.1% FA, (d) 85% ACN/1% FA, (e) 90% ACN/1% FA, (f) 95% ACN/1% FA (all unmarked peaks are from non-glycopeptides; H1: No. 1 glycopeptide from HRP; NL: Normalized Level). Glycopeptide peaks identified are marked with the symbol "HX" (X = Arabic numerals). Detected glycopeptides with m/z values and amino acid sequences are listed in Table S2.

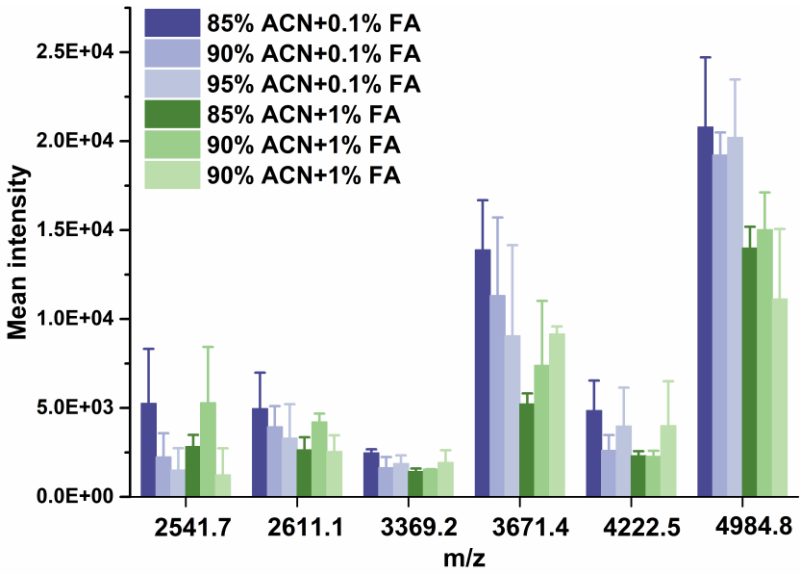


Fig. S3 Effect of different ACN/FA loading buffer solutions on intensities of six selected glycopeptides captured by Hf-BTB nanosheets from 2×10^{-6} M HRP tryptic digest through three parallel tests.

Table S2. Detailed information for detected glycopeptides enriched from tryptic digest of HRP.[S4]

Peak No.	m/z	Glycan composition	Peptide sequences[a]
H1	2074.2	XylMan3GlcNAc2	PN#VSNIVR
H2	2276.5	XylMan2FucGlcNAc2	SILLDN#TTSFR
H3	2290.7	XylMan2GlcNAc2	SILLDN#TTSFR
H4	2321.3	Man2GlcNAc2	MGN#ITPLTGTQGQIR
H5	2531.1	FucGlcNAc	SFAN#STQTFFNAFVEAMDR
H6	2541.7	XylMan3FucGlcNAc2	SSPN#ATDTIPLVR
H7	2591.5	XylMan3FucGlcNAc2	PTLN#TTYLQTLR
H8	2611.1	XylMan3GlcNAc2	MGN#ITPLTGTQGQIR
H9	2850.5	FucGlcNAc	GLIQSDQELFSSPN#ATDTIPLVR
H10	3048.1	XylMan2GlcNAc2	SFAN#STQTFFNAFVEAMDR
H11	3074.3	FucGlcNAc	LHFHDCFVNGCDASILLDN#TTSFR
H12	3087.1	XylMan3FucGlcNAc2	GLCPLNGN#LSALVDFDLR
H13	3321.6	XylMan3GlcNAc2Fuc	QLTPTFYDNSCPN#VSNIVR
H14	3353.3	XylMan3FucGlcNAc2	SFAN#STQTFFNAFVEAMDR
H15	3369.2	XylMan3FucGlcNAc2	SFAN#STQTFFNAFVEAM*DR
H16	3525.0	XylMan3GlcNAc2	GLIQSDQELFSSPN#ATDTIPLVR
H17	3539.7	Man3FucGlcNAc2	GLIQSDQELFSSPN#ATDTIPLVR
H18	3605.5	XylMan3FucGlcNAc2	NQCRGLCPLNGN#LSALVDFDLR
H19	3671.4	XylMan3FucGlcNAc2	GLIQSDQELFSSPN#ATDTIPLVR
H20	3894.4	XylMan3FucGlcNAc2	LHFHDCFVNGCDASILLDN#TTSFR
H21	4056.6	XylMan3GlcNAc2	QLTPTFYDNSC(AAVESACPR)PN#VSNIVR-H_2O
H22	4222.5	XylMan3FucGlcNAc2	QLTPTFYDNSC(AAVESACPR)PN#VSNIVR
H23	4821.5	Man3FucGlcNAc2	LYN#FSNTGLPDPTLN#TTYLQTLR
H24	4837.4	XylMan3FucGlcNAc2, XylMan3GlcNAc2	LYN#FSNTGLPDPTLN#TTYLQTLR
H25	4984.8	XylMan3FucGlcNAc2, XylMan3FucGlcNAc2	LYN#FSNTGLPDPTLN#TTYLQTLR

[a]The N-glycosylation sites are marked with N#. GlcNAc=N-acetylglucosamine, Fuc=fucose, Man=mannose, Xyl=xylose.

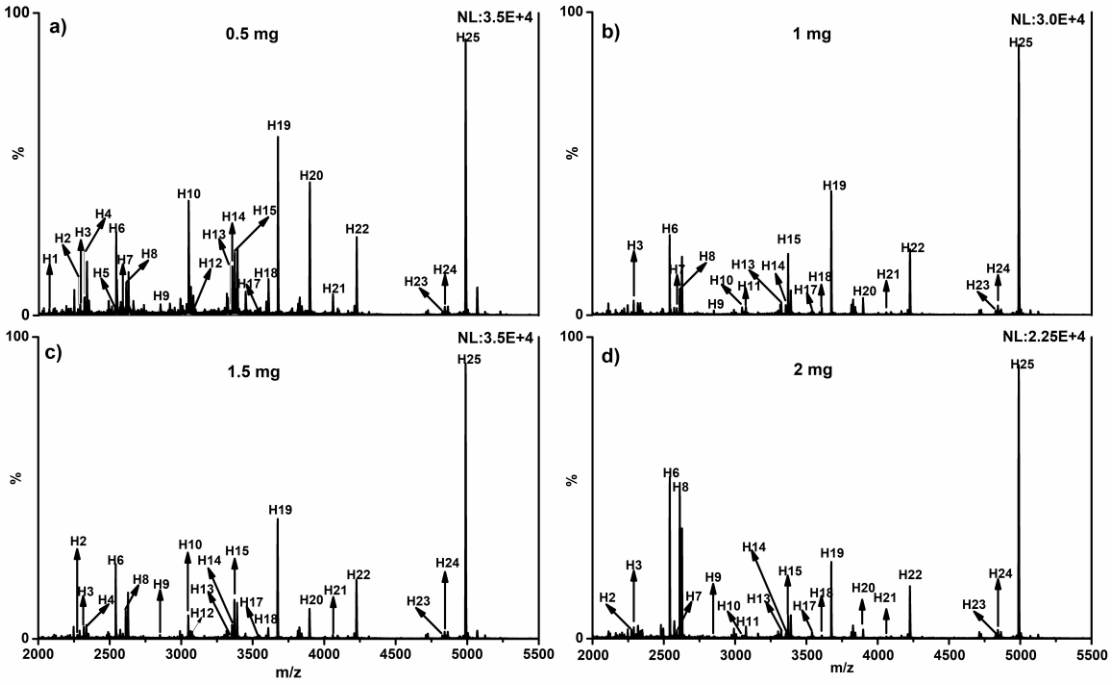


Fig. S4 MALDI-TOF mass spectra of glycopeptides from tryptic digest of HRP (2.5×10^{-6} M) enriched by different amounts of Hf-BTB nanosheets: (a) 0.5 mg, (b) 1 mg, (c) 1.5 mg, (d) 2.0 mg (all unmarked peaks are from non-glycopeptides; H1: No. 1 glycopeptide from HRP; NL: Normalized Level). Glycopeptide peaks identified are marked with the symbol "HX" (X = Arabic numerals). Detected glycopeptides with m/z values and amino acid sequences are listed in Table S2.

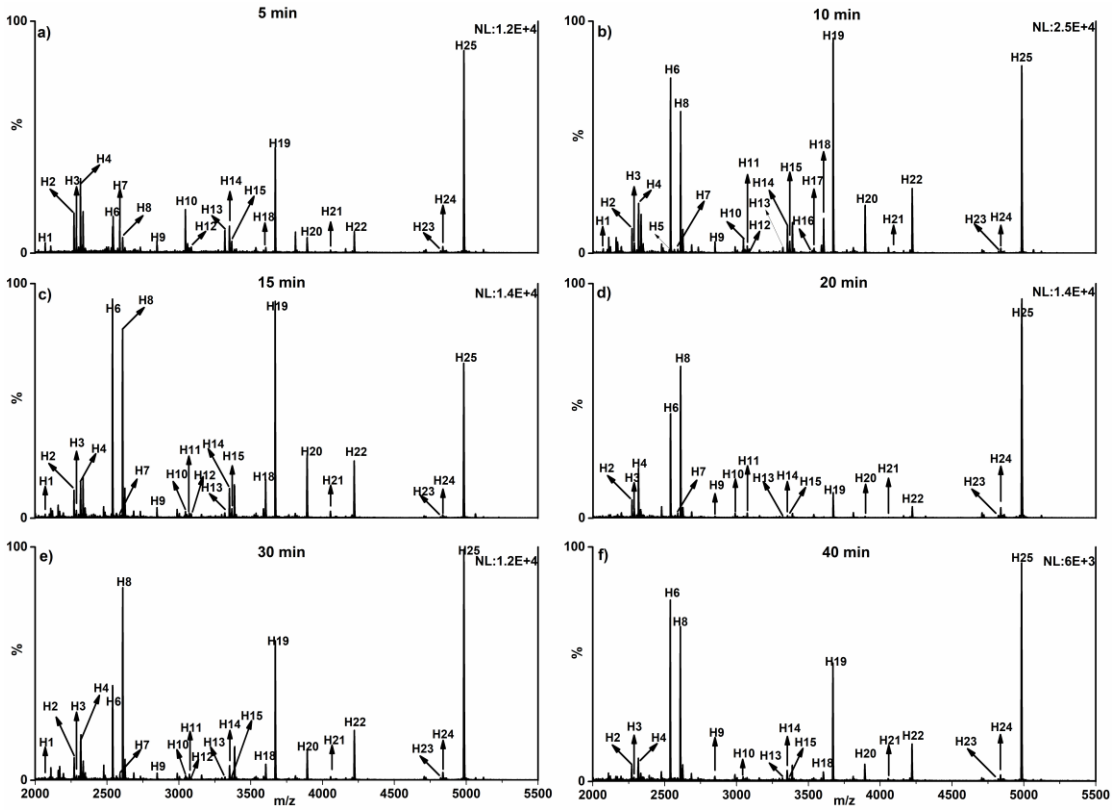


Fig. S5 MALDI-TOF mass spectra of glycopeptides from tryptic digest of HRP (2.5×10^{-6} M) enriched by Hf-BTB nanosheets with different incubation times: (a) 5 min, (b) 10 min, (c) 15 min, (d) 20 min, (e) 30 min, (f) 40 min (all unmarked peaks are from non-glycopeptides; H1: No. 1 glycopeptide from HRP; NL: Normalized Level). Glycopeptide peaks identified are marked with the symbol "HX" (X = Arabic numerals). Detected glycopeptides with m/z values and amino acid sequences are listed in Table S2.

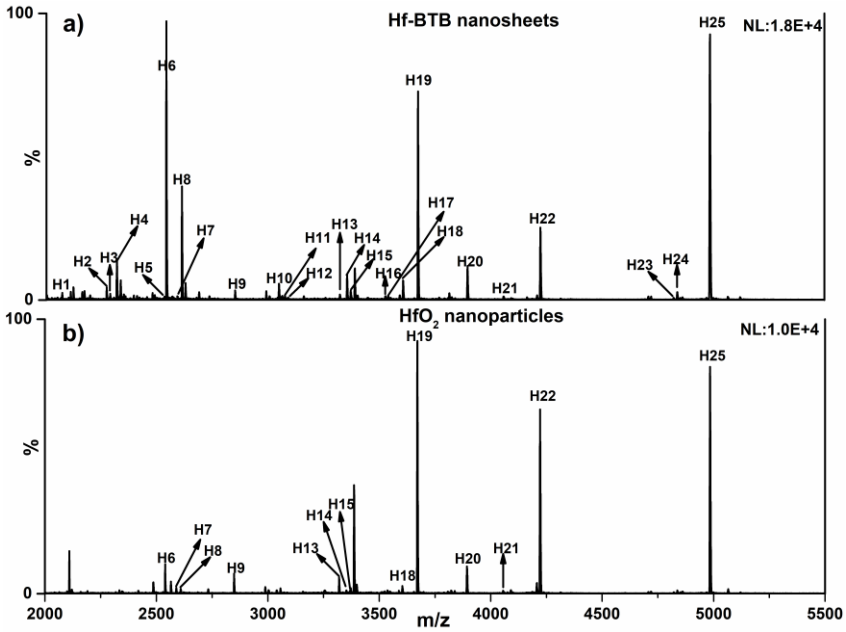


Fig. S6 MALDI-TOF mass spectra of glycopeptides from tryptic digest of HRP (2.5×10^{-6} M) enrichment with different materials: (a) Hf-BTB nanosheets, (b) HfO_2 nanoparticles (all unmarked peaks are from non-glycopeptides; H1: No. 1 glycopeptide from HRP; NL: Normalized Level). Glycopeptide peaks identified are marked with the symbol "HX" (X = Arabic numerals). Detected glycopeptides with m/z values and amino acid sequences are listed in Table S2.

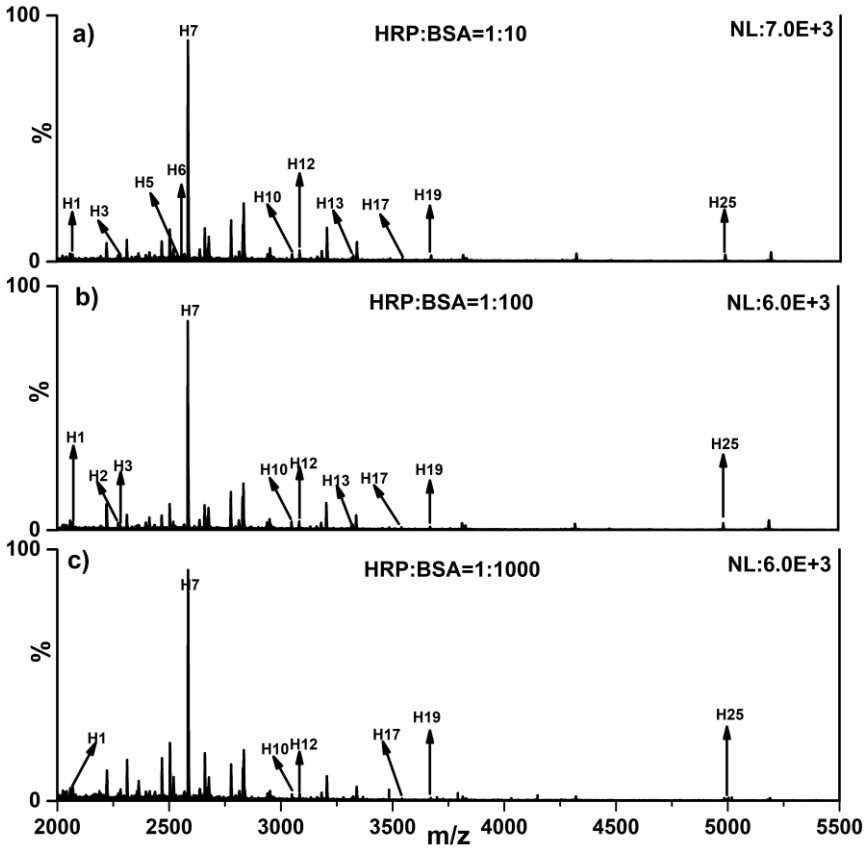


Fig. S7 MALDI-TOF mass spectra of glycopeptides enriched by Hf-BTB nanosheets from a peptide mixture of HRP and BSA at molar ratios of (a) 1:10, (b) 1:100, (c) 1:1000 (all unmarked peaks are from non-glycopeptides; H1: No. 1 glycopeptide from HRP; NL: Normalized Level). Glycopeptide peaks identified are marked with the symbol "HX" (X = Arabic numerals). Detected glycopeptides with m/z values and amino acid sequences are listed in Table S2.

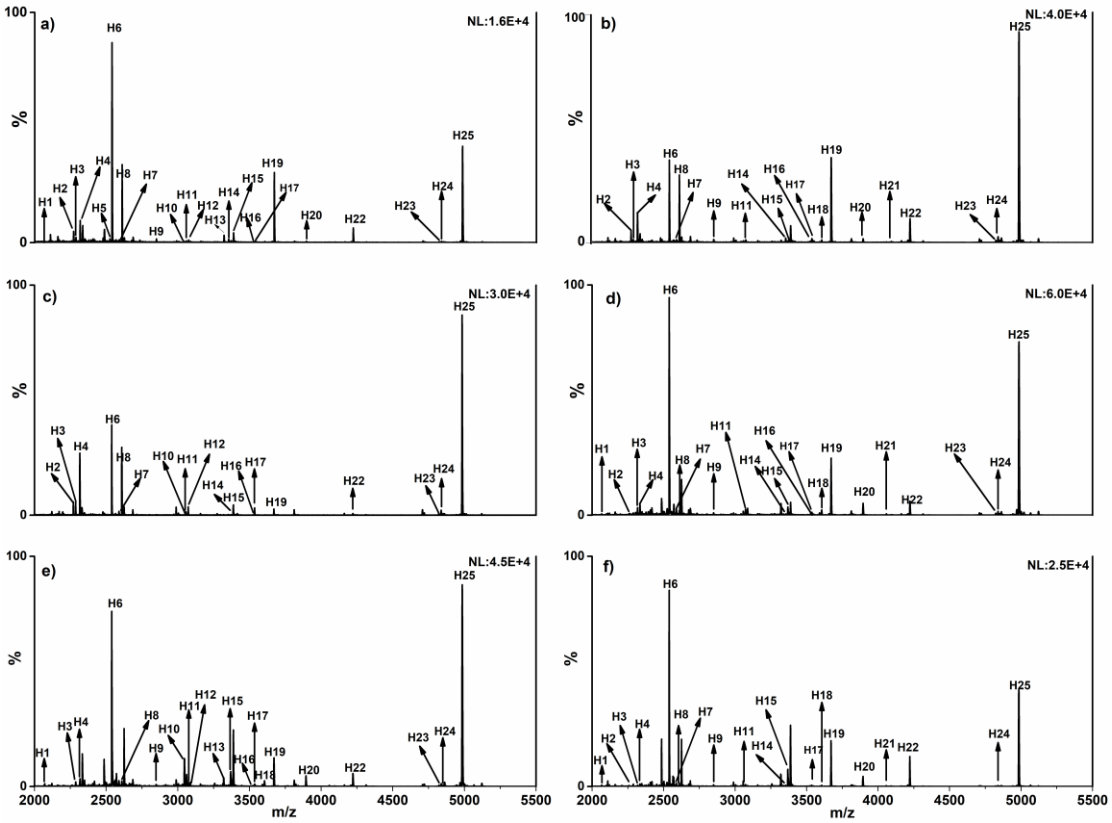


Fig. S8 Cycling performance of the Hf-BTB nanosheets: (a) first time, (b) second time, (c) third time, (d) fourth time, (e) fifth time and (f) after enrichment with Hf-BTB nanosheets stored for two weeks.

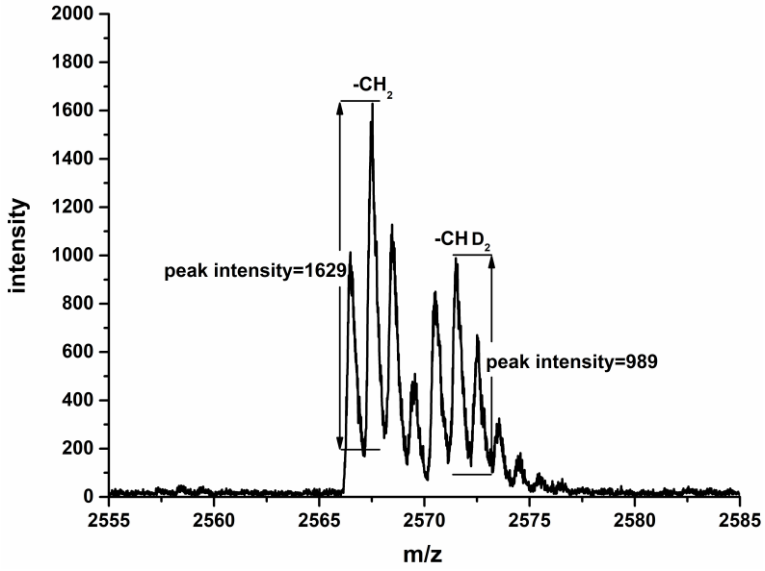
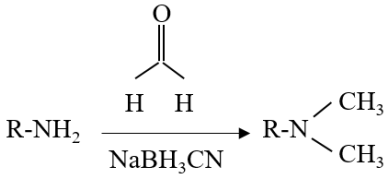
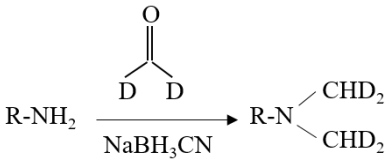


Fig. S9 Recovery of light and heavy dimethyl-labeled deglycosylated peptides from HRP tryptic digest. The labeling reactions were as follows; recovery was calculated from measurements of a 1:1 mixture of H-labelled and D-labelled m/z = 2541 peaks.[S5]



+28Da, light-isotope labeled



+32Da, heavy-isotope labeled

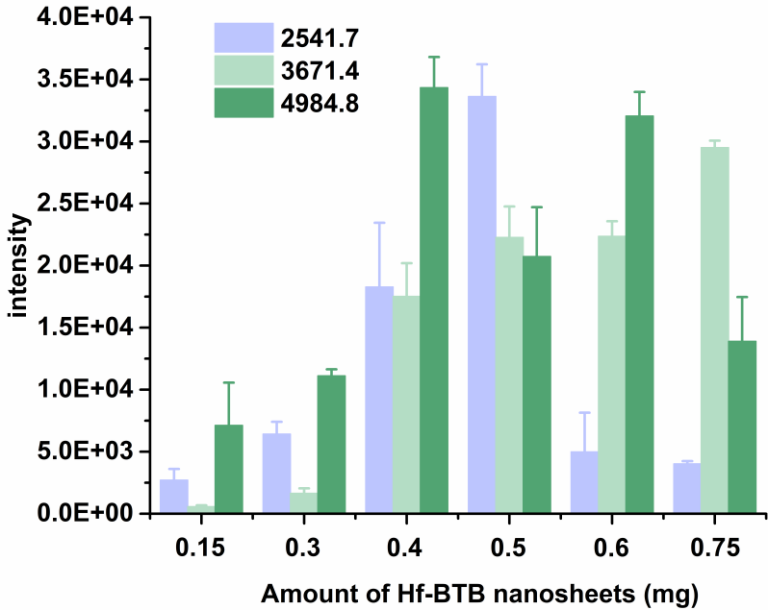


Fig. S10 Evaluation of the loading capacity of Hf-BTB nanosheets for glycopeptides.

Table S3. The MOF-based affinity materials used for enrichment of low abundance glycopeptides.

MOF-based materials	Sensitivity (LOD)	Selectivity (glycopeptides:BSA n:n)	Reference
MIL-101(Cr)-NH_2@PAMAM	1 fmol μL^{-1}	1:100	[S6]
MIL-101(Cr)-NH_2	0.2 fmol μL^{-1}	1:10	[S7]
LCD-MOFs	0.02 fmol μL^{-1}	N/A	[S8]
Fe_3O_4@PDA@UiO-66-NH_2	0.2 fmol μL^{-1}	1:100	[S9]
Fe_3O_4@Mg-MOF-74	0.5 fmol μL^{-1}	1:800	[S10]
Fe_3O_4@PDA@Zr-SO_3H	0.1 fmol μL^{-1}	1:100	[S11]
UiO-66-COOH	0.5 fmol μL^{-1}	1:20	[S12]
MG@Zn-MOFs	0.8 fmol μL^{-1}	1:800	[S4]
MIL-101(NH_2)@Au-Cys	2 fmol μL^{-1}	1:50	[S13]
Hf-BTB nanosheets	1 fmol μL^{-1}	1:1000	This work

N/A: not available

Table S4. Identified N-glycopeptides containing deamidated Asn from 2 μL human serum digest. N# denotes the N-linked glycosylation site.

No.	Peptide Sequence	Protein Group Accessions
1	FQN#ALLVR	P02768
2	EQLGEYKFQN#ALLVR	P02768
3	FKDLGEEN#FK	P02768
4	YICEN#QDSISSKLKECCEKPLLEKSHCIAEVEN#DEMPADLPSLAADFVESK	P02768
5	MPCAEDYLSVVLN#QLCVLHEK	P02768
6	VFDEFKPLVEEPQN#LIK	P02768
7	N#ECFLQHKDDNPN#LPR	P02768
8	QN#CELFEQLGEYK	P02768
9	RMPCAEDYLSVVLN#QLCVLHEK	P02768
10	KLVN#EVTEFAK	P02768
11	LVN#EVTEFAK	P02768
12	LVRPEVDVMCTAFHDN#EETFLK	P02768
13	YICEN#QDSISSK	P02768
14	N#HLQLEGLFFTN#GEHTSK	P04114
15	AALGKLPQQAN#DYLN#SFNWER	P04114
16	ALYWVN#GQVPDGVSK	P04114
17	TPGAAAN#LELIFVGPQHAGN#YR	P04217
18	EQAPHCICAN#GR	P01023
19	AIGYLN#TGYQR	P01023
20	ALLAYAFALAGN#QDK	P01023
21	HN#VYIN#GITYTPVSSTN#EK	P01023
22	N#QGN#TWLTAFVLK	P01023
23	AGDFLEAN#YMN#LQR	P01024
24	AKDQLTCN#KFDLK	P01024
25	KPVEEYAN#CHLAR	P02787

26	LKCDEWSVN#SVGK	P02787
27	DSGFQMN#QLR	P02787
28	HSTIFEN#LANK	P02787
29	IMN#GEADAMSLDGGFVYIAGK	P02787
30	SMGGKEDLIWELLN#QAQEHFGKDK	P02787
31	LEALKEN#GGAR	P02647
32	LLDN#WDSVTSTFSK	P02647
33	GFSLDEATN#LNGGLLR	P19827
34	N#QVSLTCLVK	P01857
35	FN#WYVDGVEVHN#AK	P01857
36	TVLHQDWLN#GK	P01857
37	VVSVLTVLHQDWLN#GK	P01857
38	KYVLPN#FEVK	A0A0G2JL54
39	LVN#GQSHISLSK	A0A0G2JL54
40	LQHLEN#ELTHDIITK	P01009
41	VFSN#GADLSGVTEEAPLKLSK	P01009
42	LN#HQLEGLTFQMK	P06727
43	SELTQQLN#ALFQDK	P06727
44	RVEPYGEN#FNK	P06727
45	QGHN#SVFLIKGDK	P02790
46	QGHN#SVFLIK	P02790
47	LRTEGDGVYTLN#NEK	P00738
48	TEGDGVYTLNN#EK	P00738
49	HYQIN#QQWER	P02751
50	IYLYTLN#DNAR	P02751
51	KLIN#DYVK	P01011
52	LALDN#GGLAR	Q14624
53	SGTASVVCLLN#NFYPR	P01834
54	VVSVLTVVHQDWLN#GK	P01859

55	FN#WYVDGVEVHN#AK	P01859
56	TVVHQDWLN#GK	P01859
57	TVVHQDWLN#GKEYK	P01859
58	WYVDGVEVHN#AK	P01859
59	N#QVSLTCLVK	P01859
60	VCPFAGILEN#GAVR	P02749
61	HTLN#QIDEVK	P02765
62	N#TLYLQMN#SLR	A0A0B4J1X5
63	AIEDYIN#EFSVR	P02748
64	SIEVFGQFN#GKR	P02748
65	LIVHN#GYCDGR	P02753
66	FN#WYVDGVEVHN#AK	P01861
67	N#QVSLTCLVK	P01861
68	TVLHQDWLN#GK	P01861
69	VVSVLTVLHQDWLN#GK	P01861
70	N#QVSLTCLVK	P01860
71	TVLHQDWLN#GK	P01860
72	WYVDGVEVHN#AK	P01860
73	VVSVLTVLHQDWLN#GK	P01860
74	VMPICLPSKN#YAEVGR	P00739
75	KYVLPN#FEVK	P0C0L4
76	LVN#GQSHISLSK	P0C0L4
77	N#SLYLQMN#SLR	P01780
78	GLEWVAN#IK	P01780

According to the Table S4, there were 90 N-glycosylation sites of 78 glycopeptides from 29 different glycoproteins were finally identified. It was worth noting that hundreds of nonglycopeptides were also detected from the serum sample and excluded by MS database. We will try to further improve the selectivity for real biological samples in our future work.

References

[S1]. L. Cao, Z. Lin, F. Peng, W. Wang, R. Huang, C. Wang, J. Yan, J. Liang, Z. Zhang, T. Zhang, L. Long, J. Sun, W. Lin, *Angew. Chem. Int. Ed.* 2016, ***55,*** 4962-4966.

[S2]. Boersema, P. J.; Raijmakers, R.; Lemeer, S.; Mohammed, S.; Heck, A. J. R. *Nat. Protoc.* 2009, ***4,*** 484-494.

[S3]. B. Jiang, Y. Liang, Q. Wu, H. Jiang, K. Yang, L. Zhang, Z. Liang, X. Peng, Y. Zhang, *Nanoscale* 2014, ***6,*** 5616-5619.

[S4]. J. Wang, J. Li, Y. Wang, M. Gao, X. Zhang, P. Yang, *ACS Appl. Mater. Interfaces* 2016, ***8,*** 27482-27489.

[S5]. Y. Zhao, L. Zhang, Z. Chu, Z. Xiong, W. Zhang, *Anal. Methods* 2017, ***9,*** 443-449.

[S6]. Y. Wang, J. Wang, M. Gao, X. Zhang *Proteomics*, **2017,** 17, 1700005.

[S7]. Y. Zhang, Z. Li, Q. Zhao, Y. Zhou, H. Liu, X. Zhang *Chem. Commun.*, **2014,** 50, 11504-11506.

[S8]. Y. Ji, Z. Xiong, G. Huang, J. Liu, Z. Zhang, Z. Liu, et al. *Analyst*, **2014,** 139, 4987-4993.

[S9]. Y. Xie, C. Deng, *Sci Rep.*, **2017,** 7, 1162.

[S10]. Q. Liu, Y. Xie, C. Deng, Y. Li, *Talanta*, **2017,** 175, 477-482.

[S11]. Y. Xie, C. Deng, Y. Li, *J Chromatogr A.*, **2017,** 1508, 1-6.

[S12]. Q. Liu, Y. Xie, C. Deng, Y. Li, *Talanta*, **2017,** 175, 477-482.

[S13]. W. Ma, L. Xu, X. Li, S. Shen, M. Wu, Y. Bai, et al. *ACS Appl Mater Interfaces*, **2017,** 9, 19562-19568.