Supplement information

One-Pot Synthesis of Hydrophilic Poly (glycerol methacrylate) Chitosan for Highly Selective Enrichment of Glycopeptides

J Jianzheng Jie, [b] *Dan Liu* [a] *and Xiajuan Zou* [a]*

[a] Medical and Healthy Analysis Center, Peking University, Xueyuan Road 38, Haidian District, Beijing 100191, China,

[b] Department of Gastrointestinal surgery, China-Japan Friendship Hospital, 2 Yinghua Dongjie, Chaoyang District, Beijing100029, China.

2. EXPERIMENTAL SECTION

2.1. Reagents and Materials

Glycidyl methacrylate (GMA, for GC, ⩾97.0%，Fluke, Buchs, Switzerland), 2,5-dihydroxybenzoic acid (DHB), tris(2-carboxyethyl)phosphine (TCEP), iodoacetamide, trifluoroacetic acid (TFA), acetonitrile (ACN) (HPLC grade), acetic anhydride, d_6-acetic anhydride, horseradish peroxidase (HRP, 98%) and bovine serum albumin (BSA, 98%) were purchased from Sigma Aldrich (St. Louis, MO, USA), Intravenous immunoglobulin G (IgG) was from Wuhan institute of biological products Co. Ltd, PNGase F was purchased from New England Biolabs (Ipswich, MA). Chitosan (Mw=500,000g/mol) was obtained from Qingdao Hecreat Bio-tech Company (Qingdao, China), other chemical reagents were analytical grade. Mice were purchased from Peking University Health Science Center (Beijing, China). The study strictly abided by the “International Ethics Standards for Human Medical Research” and was approved by the Ethics Review Committee of Qinghai Provincial People’s Hospital.

2.2. Synthesis of poly (glycerol methacrylate) Chitosan (PGMA@CS) nanospheres.

1.0 g chitosan was dissolved in diluted acetic acid with strong agitation. 2.0 mL GMA, 0.070g ammonium persulfate and 0.070g sodium thiosulfate were added. The reaction

mixture was heated for 2 h at 70 ºC. The polymerization and grafting reaction were finalized by raising the temperature to 80 ºC for an additional 2 h. Then, 40 mL of 2 M sodium carbonate solution were added and heated at 65 ºC for 6 h with stirring. The final PGMA@CS product was vacuum-filtered and washed with deionized water until pH neutral, dried at room temperature and ground into final powder with a mortar.

2.3. Material Characterization.

JSM-5600LV instrument (JEOL Company, Japan) was employed for the scanning electron microscope (SEM) analysis. Fourier transformed infrared spectroscopy (FT-IR) characterization has been performed on NEXUS-470 (nicolet, USA). Elemental analyses were performed on Vario EL III (Elementar, Hanau, Germany). Zeta potential analysis was performed using a Malvern Zetasizer Nano ZS (Malvern, U.K.) at 25 °C. Thermogravimetric analysis (TGA) was detected with Q600SDT instrument under a nitrogen atmosphere (100 mL min^{-1}) at a heating rate of 10 °C min^{-1} from 20 to 600 °C (Thermal Analysis, USA). The contact angles were measured by OCA20 (Krüss, Hamburg, German). The nitrogen adsorption and desorption isotherms were measured by ASAP2020M (Micromeritics, USA). The range of relative pressures was between 0 and 1. The Brunauer-Emmett-Teller (BET) method was used to calculate the specific surface areas with adsorption data in a relative pressure range from 0.0799 to 0.1998.

2.4. Trypsin digestion of standard glycoprotein and protein mixture extracted from mouse liver.

For standard protein, BSA, HRP and IgG was dissolved in a solution containing ammonium bicarbonate (50 mM, pH 8.3) to a final concentration of 1 mg / mL and heated at 95 ºC for 5 mins. TCEP (10 mM) was added and incubated at 67 ºC for about 15 min. The reduced protein was then alkylated with iodoacetamide (10 mM) at RT in the dark for 45 min. Sequencing grade trypsin (V5111, Promega) was added to the protein solution and the digestion performed overnight at 37 ºC and was stopped with 2 % TFA. The tryptic digests were stored at -20 ºC.

The mouse liver tissues were lysed in a homogenization buffer, consisting of 8 M urea, 1% Triton X-100 v/v, 65 mM DTT, 1mM EDTA, 0.5 mM EGTA, 1 mM PMSF, and protease inhibitor cocktail (cOmplete, Mini, EDTA-free, EASYpa; Roche, Basel,

Switzerland) in a ratio of 1:500 (mass ratio). The protein concentration was determined by Bradford assay. For the preparation of mouse liver digestion using Tube-gel method, 50 μg mouse liver proteins, 7.1 μL water, 22.5 μL of Tris-HCl buffer (pH 8.8), 17.5 μL of 30% acrylamide solution, 1 μL 10% SDS and 1.5 μL of 10% ammonium persulfate were added to a 1.5 mL Eppendorf tube. The mixture was vortexed for 0.5 min. Then 0.25 μL TEMED was added.

The gel was fixed with 50% methanol, 12% acetic acid for 30 mins at room temperature and cut into small pieces. The gel pieces were dehydrated with ACN and reduced with 10 mM TCEP for 10 min at 67 °C, then alkylated with 15 mM iodoacetamide at RT in the dark for 45 min. Gel pieces were washed with 50% ACN / 50mM NH_4HCO_3 buffer and dehydrated with ACN. The gel was rehydrated with 100 μL of 10 ng/μL trypsin in 25 mM NH_4HCO_3 buffer at 4 °C for 2 h and then incubated at 37 °C overnight. After enzyme digestion, the peptides mixture was extracted with ACN / 5% FA (2:1) twice and vacuum-dried.

2.5. Glycopeptides enrichment with PGMA@ CS nanospheres.

For glycopeptide enrichment: PGMA@CS nanospheres (0.4 mg) were firstly washed with water and dispersed in 200 mL loading buffer (ACN/H_2O/TFA = 89:8:3, v/v/v) containing a certain amount of HRP digests, or IgG digests. The entire suspension mixture was incubated on an Eppendorf Mixer at 900 rpm at 30 °C for 20 mins. After centrifugation at 12000 rpm for 2 mins, the supernatant was discarded, followed by rinsing with loading buffer (3 × 100 μL) to remove non-glycopeptides. The glycopeptides enriched on the PGMA@ CS was analyzed by MALDI-TOF-MS.

For glycopeptides enrichment from complex sample, 50 μg digests extracted from mouse liver were redissolved in 400 μL of loading buffer (ACN/H_2O/TFA = 89:8:3, v/v/v) and the entire suspension mixture was incubated on an Eppendorf Mixer at 900 rpm at 30 °C for 20 mins. After centrifugation at 12000 rpm for 2 mins, the supernatant was discarded, followed by rinsing with loading buffer (3 × 200 μL) to remove non-glycopeptides.

2.6. Deglycosylation of N-linked Glycopeptides by PNGase F.

The captured glycopeptides nanomaterials were dispersed in the solution of NH_4HCO_3

(25 mM), PNGase F was added and incubated at 37 °C for overnight. The reaction solutions were directly spotted on the target plate for MALDI-TOF-MS analysis.

For glycopeptides enrichment from mouse liver protein, the obtained peptides-loaded PGMA@CS was redispersed in NH_4HCO_3 (25 mM), 50 units of PNGase F was added to the solution and incubated at 37 °C for overnight to remove the glycan moieties. The nanospheres were eluted with 50μL of 50% ACN-0.5%TFA for 15 min on an Eppendorf Mixer at 1200rpm at 30 °C twice. After centrifuged at 12000 rpm for 2 min, the supernatant containing deglycosylated peptides was collected and vacuum-dried. Finally, it was redissolved in 10 μL of 0.1% FA solution for nanoLC-MS/MS analysis.

2.7. Evaluating Binding Capacity of PGMA@CS for glycopeptides Enrichment. Different amount of PGMA@CS (5 –50 μg) was added to a fixed amount of human IgG digest (3 μg), after the enrichment, the eluted fraction (0.5 μL from 20 μL total) was analyzed with MALDI-TOF MS. When the signals of six selected glycopeptides reached the maximum, the total amount of glycopeptides were bonded onto the PGMA@CS. The binding capacity was calculated by the amount of human IgG digest (3 μg) to PGMA@CS.

2.8. Recovery estimation of PGMA@CS nanospheres for glycopeptides enrichment.

The recovery yields of glycopeptides were also investigated by stable isotope acetylate labeling approaches according to a previously reported procedure with some modification. Two equivalent protein digests of human IgG (3 μg) were first labeled with light (+42.011 Da) and heavy acetylate isotopes (+45.029 Da), respectively. 1 μL of 0.1 M acetic anhydride, and d_6-acetic anhydride were added to the tubes containing the IgG digest, respectively. 10μL of 50 mM $Na_2B_2O_9$ was added to each tube, sonicated for 20 min. Then, 5μL of 0.1 M glycine was added to react with the excessive acetic anhydride and d_6-acetic anhydride. The reaction solutions were vacuum-dried. The heavy-tagged human IgG digest was enriched with PGMA@CS according to above-mentioned procedure and the resulting eluted fraction was spiked into light-tagged human IgG digest. The combined mixture was re-enriched with PGMA@CS, and the eluant was deglycosylated, followed by MALDI-TOF MS analysis. The

recovery was calculated by the peak intensity ration of heavy isotope-labeled deglycopeptides to the light isotope-labeled deglycopeptides.

2.9. Enrichment of N-glycopeptides by TiO_2. TiO_2 purification of glycopeptides was performed according to the method of Larsen et al. with minor changes [1]. The 50 μg digests extracted from mouse liver were used to purify. The bound peptides on TiO_2 were eluted with 100 μL of 1% $NH_3 \cdot H_2O$ for 20 min on an Eppendorf Mixer at 1200 rpm at 30 °C. After centrifuged at 12000 rpm for 2 min, the supernatant containing glycosylated peptides was collected and vacuum-dried, then 50 units of PNGase F in 20μL 50mM NH_4HCO_3 was added and incubated at 37 °C for overnight, and then vacuum-dried. It was redissolved in 10 μL of 0.1% FA solution for LC-MS/MS analysis.

2.10. MALDI-TOF MS process

0.5 μL of glycopeptides was loaded onto a stainless steel target, and 0.5 μL of a mixture of 20 mg/ mL DHB in 50% (v/v) ACN and 1% H_3PO_4 was added as matrix. MALDI-TOF mass spectra were acquired on an AXIMA-CFP plus (KRATOS Analytical, Shimadzu Group Company, Japan) mass spectrometer equipped with a nitrogen laser (337.1 nm). Mass spectra were obtained in positive ion and linear mode with an acceleration voltage of 20kV and an average of over 200 laser shots. Mass spectrometric data analysis was performed using Launchpad V 2. 4 Kompact MALDI software. Data analysis was carried out using Kompact MALDI software with default parameters. Each spectrum was externally calibrated with insulin (5734.62). Each spectrum was internally calibrated with HRP digestion fragment ions at m/z 1842.0, 3353.5 and 4984.2.

2.11. Reversed phase nano-liquid chromatography tandem mass spectrometry (nano-LC-MS/MS)

The deglycosylated peptides (10 μL) were loaded on a C18 pre-column (Thermo Scientific) and separated by nano-LC-MS/MS using an Easy-LC nano-HPLC (Thermo Scientific). For a gradient separation, H_2O/FA (99.9:0.1) was used as the mobile phase A while ACN/FA (99.9:0.1) was mobile phase B. At first, for phase B, from 5% to 30% for 90 min, from 30% to 50% for 10 min, then from 50% to 100% for 10 min, and held

at 100% for 10 min. The flow rate was 300 nL/min. Mass spectrometric analysis was performed using an LTQ Orbitrap Velos pro (Thermo Scientific, Bremen, Germany). The spray voltage was operated at 2.2 kV with the ion transfer capillary at 300 ℃. The MS/MS spectra were obtained in a data-dependent collision induced dissociation (CID) mode, and the full MS was acquired from m/z 350 to 2000 with resolution 60, 000. The top 15 most intense ions were selected to for MS/MS. Parameters for acquiring CID were as follows: activation time = 10 ms, normalized energy = 35, Q-activation = 0.25. The dynamic exclusion was set as follows: repeat count 1, duration 30 s, exclusion list size 500 and an exclusion duration 30 s.

2.12. Database search and data analysis

Proteome Discoverer v1.4.1.14 (Thermo Scientific) was used to search all of the obtained MS/MS spectra with UniProt Mouse database (53290 sequences). The search parameters were set as follows: fixed modification of cysteine residues (+57.021 Da), some variable modifications were set as: deamidation on asparagine and glutamine (Asp and Asn, + 0.9858 Da), oxidation on methionine (Met, + 15.9949 Da). The proteolytic enzyme was set as trypsin with no more than two missed cleavage sites allowed. The mass tolerance of the precursor ion and the fragment ion were set to 10 ppm and 0.8 Da, respectively. False discovery rates were obtained using Percolator selecting identification with a q-value equal and less than 0.01.

All the identified deamidation sites according to the consensus sequence of N-glycosylation (N-X-[S/T], X can be any amino acids except proline) were verified as glycosylation sites. The identified glycoproteins must contain no less than one glycosylation site defined above.

2.13. Quantification and Standard Curve Calculation.

30 pmol of HRP digest were enriched with PGMA@CS according to above-mentioned procedure. To calculate standard curves, the eluant was applied onto a MALDI sample plate in a concentration range from 140 to 1.4 fmol/μL and 2000-100 fmol/μL diluted in 10 mg/ mL CHCA in 50% (v/v) ACN and 0.1% TFA. The spots were air-dried and followed by MALDI-TOF MS analysis. All data points for standard curve calculation

are averages of at least three replicates. Relative standard deviations (RSD) and linearity (R^2) were calculated with Microsoft Excel 2010.

Table S1 Elemental analysis of chitosan and PGMA@CS

	C (%)	H (%)	N (%)
chitosan	40.00	7.23	7.18
PGMA@CS	51.51	6.81	2.45

Table S2. Detailed information of the glycopeptides enriched by PGMA@CS from HRP digest, N# denotes the N-linked glycosylation site.

No.	Observed m/z	Glycan composition	Amino acid sequence
H1	1842.2	$XylMan_3FucGlcNAc_2$	NVGL*N*#R
H2	2532.8	FucGlcNAc	SFAN#STQTFFNAFVEAMDR
H3	2590.8	$XylMan_3FucGlcNAc_2$	PTLN#TTYLQTLR
H4	2610.8	$XylMan_3GlcNAc_2$	MGN#ITPLTGTQGQIR
H5	3074.5	FucGlcNAc	LHFHDCFVNGCDASILLDN#TTSFR
H6	3088.1	$XylMan_3FucGlcNAc_2$	GLCPLNGN#LSALVDFDLR
H7	3191.1	$XylMan_2FucGlcNAc_2$	SFAN#STQTFFNAFVEAMDR
H8	3207.1	$XylMan_3GlcNAc_2$	SFAN#STQTFFNAFVEAMDR
H9	3222.0	$Man_3FucGlcNAc_2$	SFAN#STQTFFNAFVEAMDR
H10	3320.9	$XylMan_3FucGlcNAc_2$	QLTPTFYDNSCPN#VSNIVR
H11	3353.1	$XylMan_3FucGlcNAc_2$	SFAN#STQTFFNAFVEAMDR
H12	3369.4	$XylMan_3FucGlcNAc_2$	SFAN#STQTFFNAFVEAM*DR
H13	3508.7	XylMan2FucGlcNAc2	GLIQSDQELFSSPN#ATDTIPLVR
H14	3525.6	XylMan3GlcNAc2	GLIQSDQELFSSPN#ATDTIPLVR
H15	3539.6	Man3FucGlcNAc2	GLIQSDQELFSSPN#ATDTIPLVR
H16	3605.3	XylMan3FucGlcNAc2	NQCRGLCPLNGN#LSALVDFDLR
H17	3671.5	XylMan3FucGlcNAc2	GLIQSDQELFSSPN#ATDTIPLVR
H18	3894.6	$XylMan_3FucGlcNAc_2$	LHFHDCFVNGCDASILLDN#TTSFR

H19	4056.8	$XylMan_3GlcNAc_2$	QLTPTFYDNSC(AAVESACPR)PN#VSNIVR-H_2O
H20	4222.4	XylMan3FucGlcNAc2	QLTPTFYDNSC(AAVESACPR)PN#VSNIVR
H21	4720.1	$Man_3FucGlcNAc_{2,}$ $Man_3FucGlcNAc_2$	LYN#FSNTGLPDPTLN#TTYLQTLR
H22	4821.7	$XylMan_2FucGlcNAc_2$, $XylMan_2GlcNAc_2$	LYN#FSNTGLPDPTLN#TTYLQTLR
H23	4838.1	$XylMan_3FucGlcNAc_{2,}$ $XylMan_3GlcNAc_2$	LYN#FSNTGLPDPTLN#TTYLQTLR
H24	4852.7	$Man_3FucGlcNAc_{2,}$ $XylMan_3FucGlcNAc_2$	LYN#FSNTGLPDPTLN#TTYLQTLR
H25	4984.5	$XylMan_3FucGlcNAc_{2,}$ $XylMan_3FucGlcNAc_2$	LYN#FSNTGLPDPTLN#TTYLQTLR

Table S3. Detailed information of the glycopeptides enriched by PGMA@CS from human IgG digest. N# denotes the N-linked glycosylation site.

No.	Observed m/z	Glycan composition	Amino acid sequence
I1	2235.9	[Hex]3[HexNAc]2[Fuc]1	EEQFN#STFR
I2	2266.7	[Hex]3[HexNAc]2[Fuc]1	EEQYN#STYR
I3	2397.8	[Hex]3[HexNAc]3[Fuc]1	EEQFN#STFR
I4	2412.8	[Hex]3[HexNAc]3[Fuc]1	EEQFN#STYR
I5	2430.0	[Hex]3[HexNAc]3[Fuc]1	EEQYN#STYR
I6	2469.9	[Hex]3[HexNAc]4	EEQFN#STYR
I7	2487.1	[Hex]3[HexNAc]4	EEQYN#STYR
I8	2560.3	[Hex]4[HexNAc]3[Fuc]1	EEQFN#STFR
I9	2592.3	[Hex]4[HexNAc]3[Fuc]1	EEQYN#STYR
I10	2601.3	[Hex]4[HexNAc]4[Fuc]1	EEQFN#STFR
I11	2617.0	[Hex]4[HexNAc]4[Fuc]1, or[Hex]4[HexNAc]4	EEQFN#STYR, or EEQFN#STFR

I12	26333.4	[Hex]4[HexNAc]4[Fuc]1,or [Hex]4[HexNAc]4	EEQYN#STYR, or EEQFN#STYR,
I13	2649.6	[Hex]4[HexNAc]4	EEQYN#STYR
I14	2658.5	[Hex]3[HexNAc]5	EEQFN#STFR
I15	2690.6	[Hex]3[HexNAc]5	EEQYN#STYR
I16	2763.7	[Hex]4[HexNAc]4[Fuc]1	EEQFN#STFR
I17	2779.3	[Hex]4[HexNAc]4[Fuc]1,or[Hex]5[HexNAc]4	EEQFN#STYR, or EEQFN#STFR
I18	2795.9	[Hex]4[HexNAc]4[Fuc]1,or[Hex]5[HexNAc]4	EEQYN#STYR, or EEQFN#STYR
I19	2804.2	[Hex]3[HexNAc]5[Fuc]1	EEQFN#STFR
I20	2812.2	[Hex]4[HexNAc]4[Fuc]1	EEQYN#STYR
I21	2821.7	[Hex]3[HexNAc]5[Fuc]1, or [Hex]4[HexNAc]5	EEQFN#STYR, or EEQFN#STFR
I22	2836.1	[Hex]3[HexNAc]5[Fuc]1, or [Hex]4[HexNAc]5	EEQYN#STYR or EEQFN#STYR
I23	2853.3	[Hex]4[HexNAc]5	EEQYN#STYR
I24	2926.5	[Hex]5[HexNAc]4[Fuc]1	EEQFN#STFR
I25	2942.2	[Hex]5[HexNAc]4[Fuc]1	EEQFN#STYR
I26	2958.5	[Hex]5[HexNAc]4[Fuc]1	EEQYN#STYR
I27	2967.2	[Hex]4[HexNAc]5[Fuc]1	EEQFN#STFR
I28	2983.6	[Hex]4[HexNAc]5[Fuc]1, or [Hex]5[HexNAc]5	EEQFN#STYR, or EEQFN#STFR
I29	2999.1	[Hex]4[HexNAc]5[Fuc]1,or[Hex]5[HexNAc]5	EEQYN#STYR, or EEQFN#STYR
I30	3016.1	[Hex]5[HexNAc]5	
I31	3056.1	[Hex]4[HexNAc]4[Fuc]1[NeuAc]1	EEQFN#STFR
I32	3088.5	[Hex]4[HexNAc]4[Fuc]1[NeuAc]1	EEQYN#STYR
I33	3104.7	[Hex]5[HexNAc]4 [NeuAc]1	EEQYN#STYR
I34	3130.5	[Hex]5[HexNAc]5[Fuc]1	EEQFN#STFR
I35	3145.9	[Hex]5[HexNAc]5[Fuc]1	EEQFN#STYR
I36	3162.4	[Hex]5[HexNAc]5[Fuc]1	EEQYN#STYR
I37	3218.4	[Hex]5[HexNAc]4[Fuc]1[NeuAc]1	EEQFN#STFR
I38	3234.2	[Hex]5[HexNAc]4[Fuc]1[NeuAc]1	EEQFN#STYR
I39	3250.5	[Hex]5[HexNAc]4[Fuc]1[NeuAc]1	EEQYN#STYR

Table S4. Recovery of two deglycosylated peptides from human IgG digest by using PGMA@CS

Ratio(%)	EEQFN#STYR	EEQYN#STYR
D/H	100.2	97.79
D/H	99.93	95.85
D/H	99.89	95.74
Average recovery ± S. D	100.00±0.22	96.46±1.15

Table S5. N-glycopeptides of HRP digest enrichment with PGMA@CS by MALDI-TOF analysis: limits of detection, limits of quantification, and calibration curve results (N=3)

analyte	LLOD fmol/μL	LLOQ fmol/μL	RSD %	R^2
3353	0.5	1.4	2.4-15.0	0.9934
3671	0.5	1.4	1.5-10.2	0.9953
4222	0.5	1.4	2.47-10.77	0.9916

Table S8. Comparation of the detected glycopeptides and glycoproteins from mouse liver digest with published paper

	Ref. 18 (50 μg proteins)		Ref.19 (50 μg proteins)	PGMA@CS (45μg proteins)
	HAPD strategy	conventional method		
Glycopeptides	889	700	149	1416
glycoproteins	482	407	129	597
LC-MS/MS times(min)	140	140	127	120
ID glycopeptides	17.78/μg, 6.35/min	14/μg, 5/min	2.98/μg, 1.17/min	31.47/μg, 11.8/min

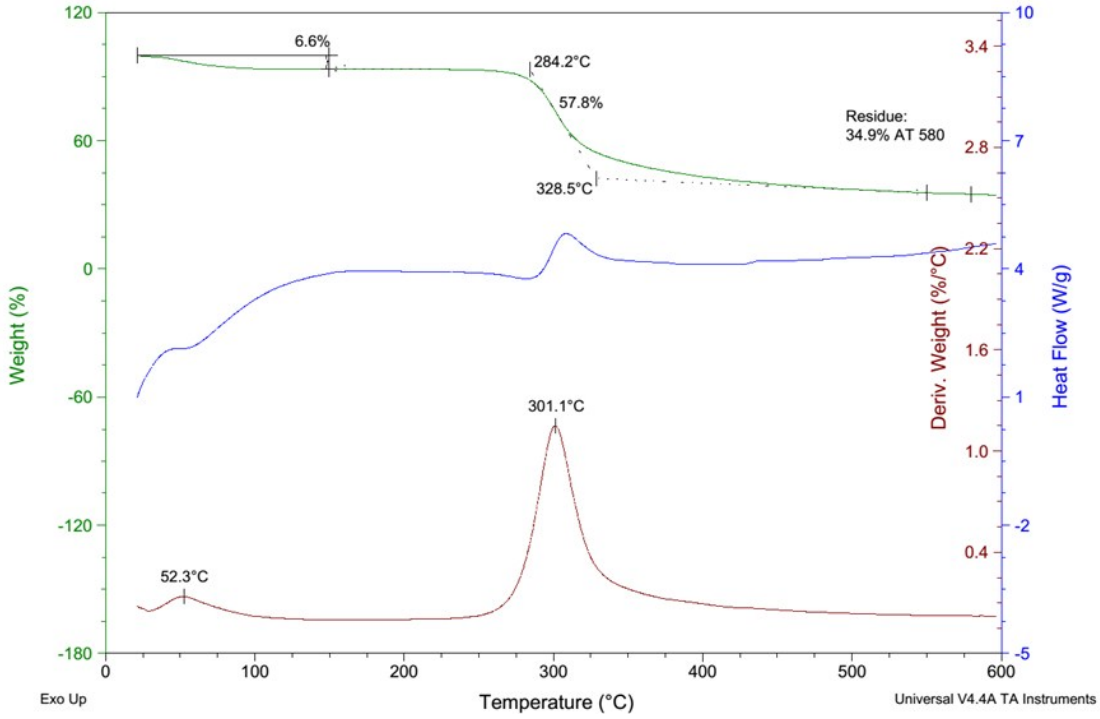


Fig. S1. The thermogravimetric analysis of CS

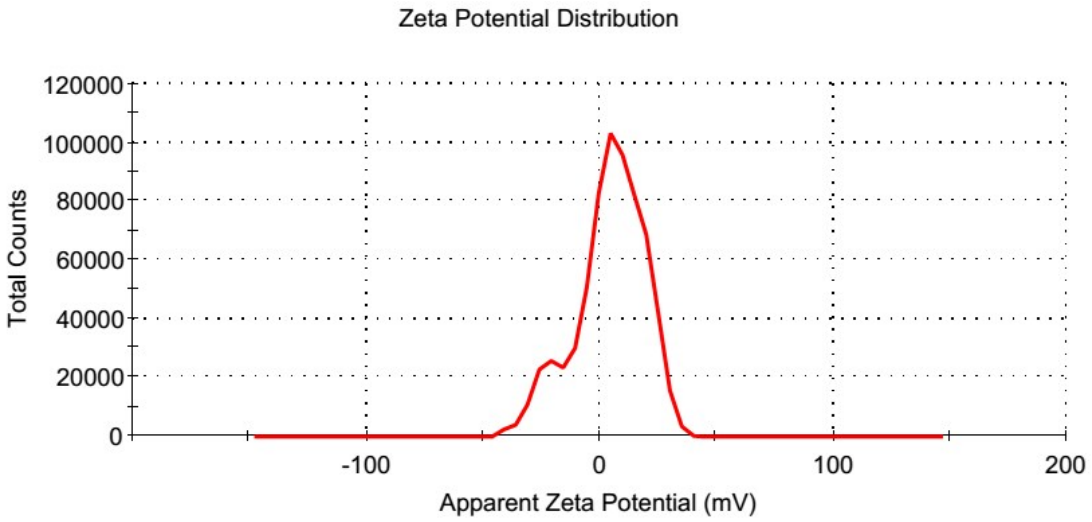


Fig. S2. The Zeta potentials of PGMA@CS

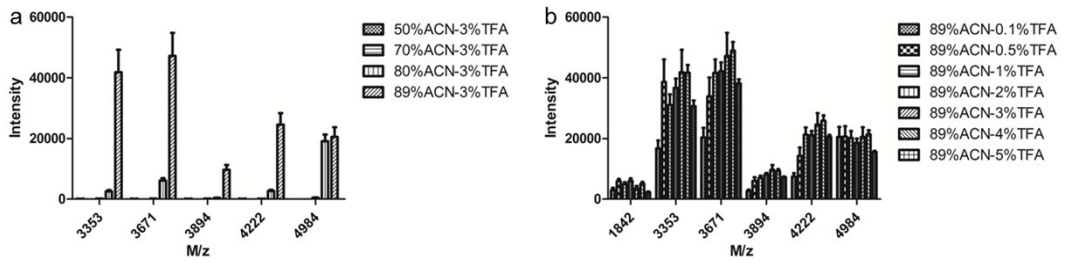


Fig. S3. Optimum of the binding condition of PGMA@CS to glycopeptides

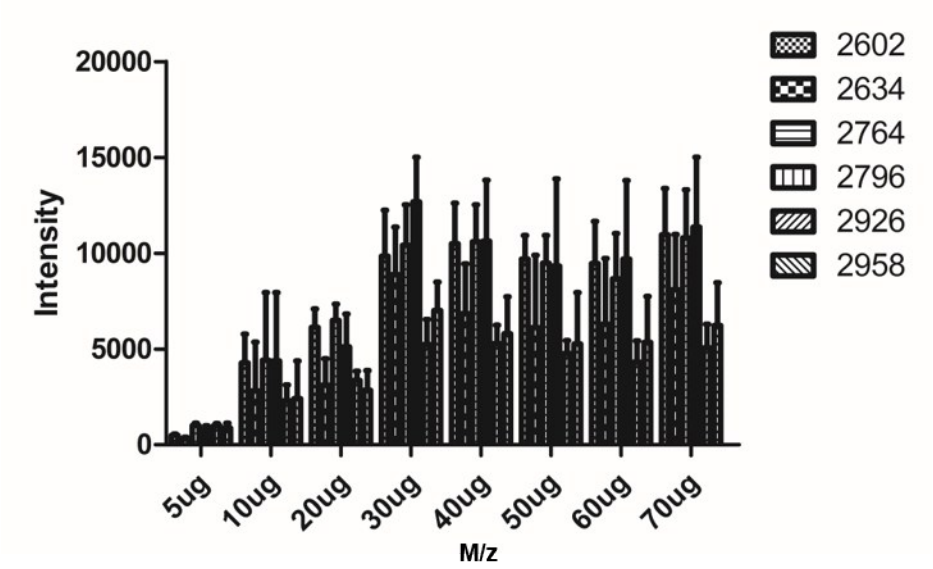
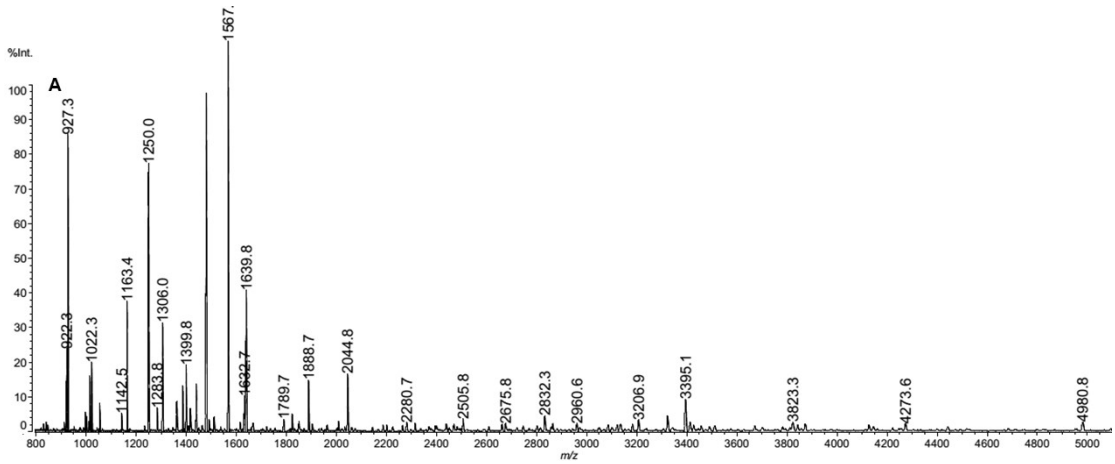


Fig. S4. The binding capacity of PGMA@CS toward glycopeptides



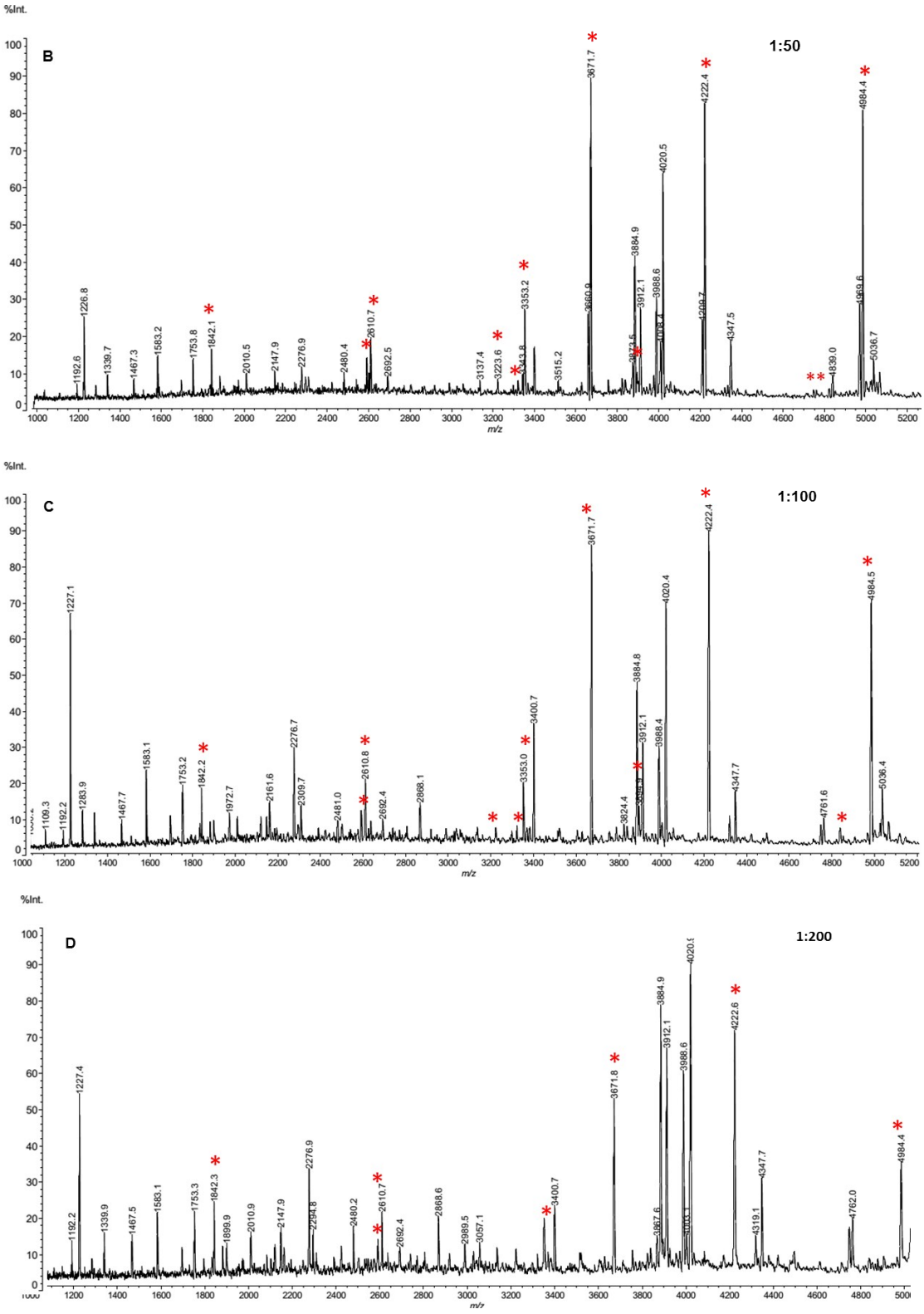


Fig. S5. MALDI-TOF MS spectra obtained from peptide mixture of HPR and BSA with different molar ratios. A) before enrichment of HRP:BSA (1:50)； after enrichment B) HRP:BSA (1:50); C) HRP:BSA (1:100) and D) HRP:BSA (1:200).

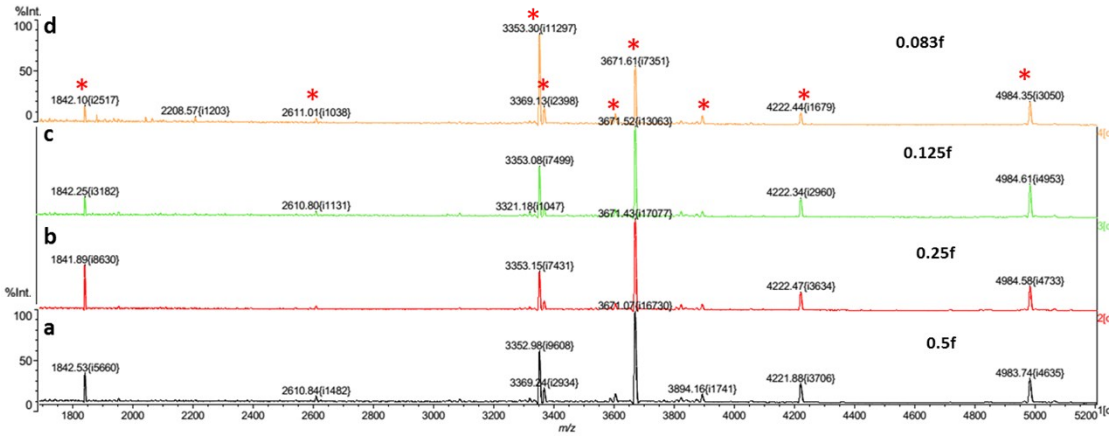


Fig. S6. The detection limit of PGMA@CS toward glycopeptides

Reference

1 G. Palmisano, B. L. Parker, K. Engholm-Keller, S. E. Lendal, K. Kulej, M. Schulz, V. Schwammle, M. E. Graham, H. Saxtorph, S. J. Cordwell, M. R. Larsen, *Mol. Cell. Proteomics*, 2012, **11**, 1191-202.