Supporting Information

Click synthesis of hydrophilic maltose-functionalized iron oxide magnetic nanoparticles based on dopamine anchors for highly selective enrichment of glycopeptides

Changfen Bi,[†,‡] Yingran Zhao,[†,‡] Lijin Shen,[§] Kai Zhang,[§] Xiwen He,[†,‡] Langxing Chen*[,†,‡] and Yukui Zhang[†,‡,#]

[†]Research Center for Analytical Sciences, College of Chemistry, Tianjin Key Laboratory of Biosensing and Molecular Recognition, State Key Laboratory of Medicinal Chemical Biology, Nankai University, Tianjin 300071, China

Fax: (+86) 22-2350-2458, E-mail: lxchen@nankai.edu.cn

[‡]Collaborative Innovation Center of Chemical Science and Engineering (Tianjin), Tianjin 300071, China

[§] Department of Biochemistry and Molecular Biology &Tianjin Key Laboratory of Medical Epigenetics, Tianjin Medical University, Tianjin, 300070, China

[#]Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116011, China

Cell culture and protein extraction

Human renal mesangial cells (HRMC, kindly donated by Dr. Mingzhen Li (Metabolic Diseases Hospital, Tianjin Medical University, Tianjin, China)) were grown in Dulbeco's modified eagle medium (Thermo Fisher) supplemented with 1% penicillin/streptomycin and 10% fetal bovine serum in a humidified atmosphere with 5% CO_2 in air. After collection, HRMC cells were washed twice with cold phosphate-buffer saline (PBS). The cell layers were further solubilized in RIPA buffer (0.1%SDS, 1%Triton X-100 and 1% deoxycholate) containing proteinase inhibitors (sodium butyrate and 1mmol/L phenylmethylsulfonyl fluoride (PMSF, Sigma-Aldrich)). Then total cell undissolved substance were ultrasound. The crude extract was finally clarified by centrifugation at 16, 000×g at 4 °C for 10 min. The concentration of total proteins was determined by BCA protein assay kit (Pierce BCA protein assay kit, Thermo Scientific).

LC-MS/MS analysis

The glycopeptide enrichment was desalted using a µ-C18 Ziptip and dissolved in 10 µL of HPLC buffer A (0.1% (v/v) formic acid in water). 5 µL sample was injected into a Nano-LC system (EASY-nLC 1000, Thermo Fisher Scientific, Waltham, MA). Chromatography was performed using an EASY-Spray Nano-LC source with a 15 cm × 50 µm inner diameter column packed with 2 µm C18 particles. The flow rate was 300 nl/min, a 45-min linear gradient from 2 to 35% HPLC buffer B (0.1% formic acid in ACN) was developed, eventually the organic content was increased to 50% over 10 min. The HPLC elute was electrosprayed directly into an Orbitrap Q-Exactive mass spectrometer (Thermo Fisher Scientific, Waltham, MA). The source was operated at 1.8 kV. For full MS survey scan, automatic gain control (AGC) target was 3e6, scan range was from 350 to 1750 with the resolution of 70,000. The 10 most intense peaks with charge state 2 and above were selected for fragmentation by higher-energy collision dissociation (HCD) with normalized collision energy of 27 %. The MS2 spectra were acquired with 17, 500 resolution.

MS/MS Data Analysis

Raw file was searched against the Uniport-Human protein sequence database using the PD search engine (version 2.1.0, Thermo Fisher Scientific) with an overall false discovery rate (FDR) for peptides of less than 1%. Trypsin was specified as digesting enzyme. A maximum of 2 missing cleavages were allowed. Mass tolerances for precursor ions were set at ±10 ppm for precursor ions and ±0.02 Da for MS/MS. Oxidation of methionine, acetylation on protein N-terminal, and asparagine deamination were fixed as variable modifications. Carbamidomethylation on Cys was specified as fixed modification.

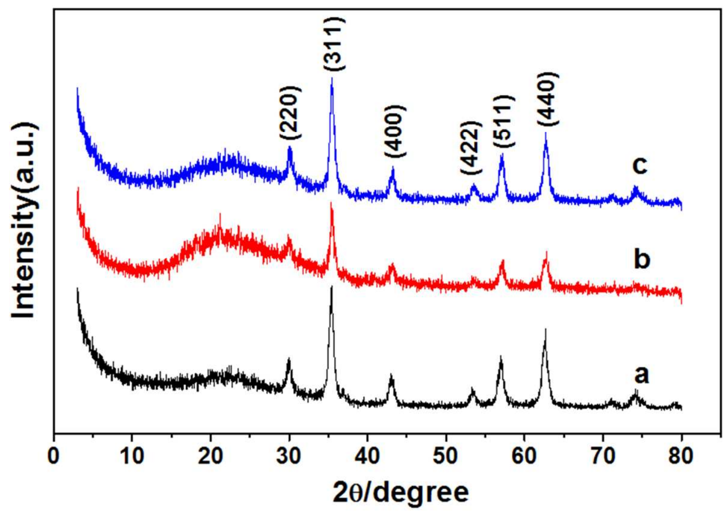

Figure S1 XRD patterns of Fe_3O_4 (a), Fe_3O_4-DA-N_3 (b) and Fe_3O_4-DA-Maltose (c) NPs.

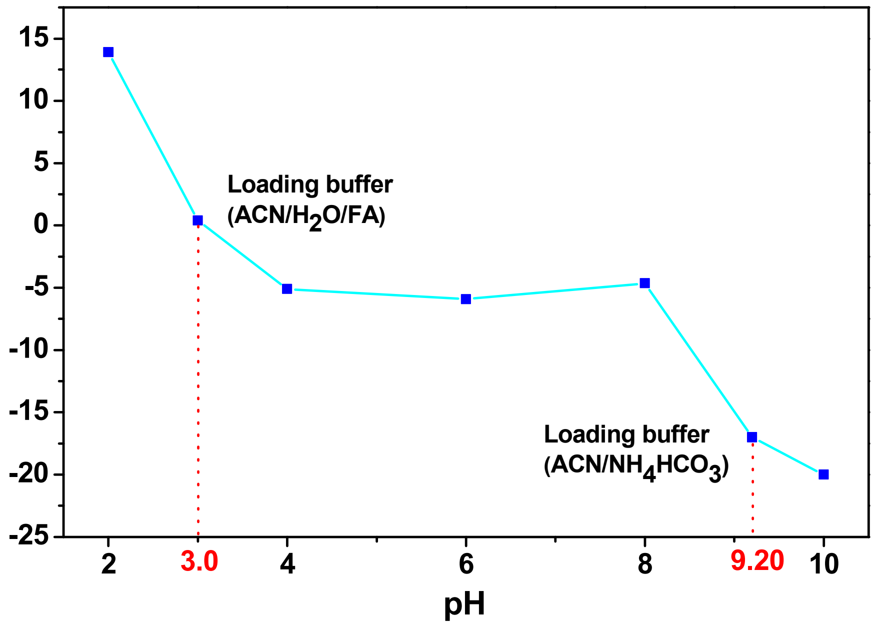

Figure S2 Dependence of zeta spectra of Fe_3O_4-DA-Maltose NPs on the change of pH.

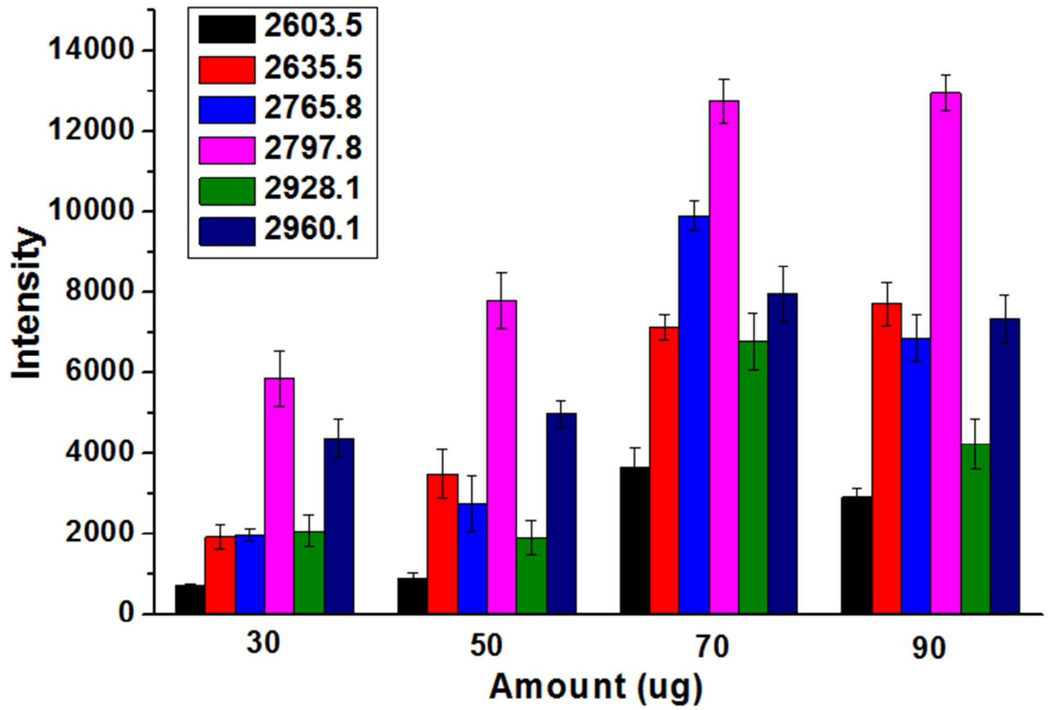

Figure S3 The amount of MNPs influencing intensity of six selected glycopeptides from tryptic digests of human IgG (3 μg) after enrichment using Fe_3O_4-DA-Maltose.

Table S1 Observed glycopeptides and glycan structures of human HRP digests enriched by Fe_3O_4-DA-Maltose NPs. The N- glycosylation sites are marked with N#. Hex= hexose, HexNAc= N-acetylhexosamine, Fuc= fucose, NeuAc= N-acetylneuraminic acid.

Peak number	Observed m/z	Glycan composition	Amino acid sequence
H1	2613.4	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	PTLN#TTYLQTLR
H2	2853.2	[HexNAc]1[Fuc]1	GLIQSDQELFSSPN#ATDTIPLVR
H3	3323.4	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	QLTPTFYDNSCPN#VSNIVR
H4	3355.4	[Hex]2[HexNAc]2[Fuc]1[Xyl]1	SFAN#STQTFFNAFVEAMDR
H5	3673.8	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	GLIQSDQELFSSPN#ATDTIPLVR
H6	3899.4	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	LHFHDCFVNGCDASILLDN#TTSFR
H7	4059.5	[Hex]3[HexNAc]2[Xyl]1	QLTPTFYDNSC(AAVESACPR)PN#VSNIVR -H_2O
H8	4223.8	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	QLTPTFYDNSC(AAVESACPR)PN#VSNIVR
H9	4986.1	[Hex]2[HexNAc]2[Fuc]1[Xyl]1 [Hex]2[HexNAc]2[Fuc]1[Xyl]1	LYN#FSNTGLPDPTLN#TTYLQTLR

Table S2 Observed glycopeptides and glycan structures of human IgG digests enriched by Fe_3O_4-DA-Maltose NPs. The N- glycosylation sites are marked with N#. Hex= hexose, HexNAc= N-acetylhexosamine, Fuc= fucose, NeuAc= N-acetylneuraminic acid.

Peak number	Observed m/z	Glycan composition	Amino acid sequence
I1	2603.5	[Hex]3[HexNAc]4[Fuc]1	EEQFN#STFR
I2	2635.5	[Hex]3[HexNAc]4[Fuc]1	EEQYN#STYR
I3	2652.5	[Hex]4[HexNAc]4	EEQYN#STYR
I4	2765.8	[Hex]4[HexNAc]4[Fuc]1	EEQFN#STFR
I5	2778.2	[Hex]5[HexNAc]4	EEQFN#STFR
I6	2797.8	[Hex]4[HexNAc]4[Fuc]1	EEQYN#STYR
I7	2806.9	[Hex]3[HexNAc]5[Fuc]1	EEQFN#STFR
I8	2819.4	[Hex]4[HexNAc]5	EEQFN#STFR
I9	2838.9	[Hex]3[HexNAc]5[Fuc]1	EEQYN#STYR
I10	2928.1	[Hex]5[HexNAc]4[Fuc]1	EEQFN#STFR
I11	2960.1	[Hex]5[HexNAc]4[Fuc]1	EEQYN#STYR
I12	2982.7	[Hex]5[HexNAc]5	EEQFN#STFR
I13	3001.1	[Hex]4[HexNAc]5[Fuc]1	EEQYN#STYR
I14	3014.5	[Hex]5[HexNAc]5	EEQYN#STYR
I15	3058.4	[Hex]4[HexNAc]4[Fuc]1[NeuAc]1	EEQFN#STFR
I16	3130.2	[Hex]5[HexNAc]5[Fuc]1	EEQFN#STFR
I17	3162.4	[Hex]5[HexNAc]5[Fuc]1	EEQYN#STYR

Table S3 List of identified glycoproteins and peptides sequence from tryptic digest of proteins sample extracted from human renal mesangial cells after enrichment by Fe_3O_4-DA-Maltose NPs. N# denotes the N-linked glycosylation site.

No	Protein	Peptide sequence
1	E9PFH4	HTNPIVEN#GQTHPcQK
2	P27797	HEQN#IDcGGGYVK
3	P52597	HSGPN#SADSANDGFVR
4	E9PKE3	N#QTAEKEEFEHQQK
5	Q5T3Q7	AQHSEN#DLEEVGK
6	P23284	DTN#GSQFFITTVK
7	B4DGN5	IIAEGAN#GPTTPEADK
8	F8VWV4	GHLEN#NPALEK
9	Q13247-2	DADDAVYELN#GK
10	P10809	VGGTSDVEVN#EK
11	Q07065	IETNEN#NLESAK
12	E5RGH4	HTGPN#SPDTANDGFVR
13	P62269	YSQVLAN#GLDNK
14	P14625;P08238	ELISN#ASDALDK
15	Q9HDC9-2	LEN#GEIETIAR
16	P09651-3	YHTVN#GHNcEVR
17	K7ENJ4	VVDALGN#AIDGK
18	P51991-2	YHTIN#GHNcEVK
19	H0YI43	EGN#GTVMGAELR
20	P22626-2	YHTINGHN#AEVR
21	Q5HY54	VHGPGIQSGTTN#KPNK
22	C9JI87	SEN#GLEFTSSGSANTETTK
23	J3QQQ9	AN#GTTVHVGIHPSK
24	J3QR48	VLANPGN#SQVAR
25	Q5ST81	EVDEQMLN#VQNK
26	P05023-2	VDN#SSLTGESEPQTR
27	P11021;E9PKE3	VEILAN#DQGNR
28	P08670	VELQELN#DR
29	P08670	DGQVIN#ETSQHHDDLE
30	H0YEL5	KIENVPTGPN#NKPK
31	P08238	EQVAN#SAFVER
32	O95202	VAEVEGEQVDN#K

33	P08670	LQDEIQN#MK
34	P62701	DAN#GNSFATR
35	Q00839-2	N#GQDLGVAFK
36	P06748-3	VDN#DENEHQLSLR
37	P83881	LECVEPN#CR
38	B4E241	VYVGNLGNN#GNK
39	P19338	N#DLAVVDVR
40	P08670	FADLSEAAN#R
41	P62158	DGN#GYISAAELR
42	P67809	GAEAAN#VTGPGGVPVQGSK
43	P07942	AAQN#SGEAEYIEK
44	P04040	LVNAN#GEAVYcK
45	Q01081	NPQN#SSQSADGLR
46	H0YI43	EGN#GTVMGAEIR
47	Q5ST81	EVDEQMLNVQN#K
48	B5MCW2	NN#ASTDYDLSDK
49	P11021;E9PKE3	VEILAN#DQGN#R
50	O94906-2	LEEANGNTQMVEK
51	Q96GQ7	EMQQSEAQIN#TAK
52	P67809	EDGNEEDKEN#QGDETQGQQPPQR
53	P13591-1	LEGQMGEDGN#SIK
54	P11021	ITITN#DQNR
55	P55884	N#GDYLcVK
56	P35637-2	APKPDGPGGGPGGSHMGGN#YGDDR
57	Q00610-2	NN#RPSEGPLQTR
58	P31040-3	GEGGILIN#SQGER
59	P38646	VLEN#AEGAR
60	A8MUD9	TTHFVEGGDAGN#REDQINR
61	Q71U36-2	DVN#AAIATIK
62	B5MD17	N#SDEADLVPAK
63	Q6UVK1	N#GLAGDTETFR
64	F8VQQ8	TN#EAQAIETAR
65	F5H8J2	NN#FEGEVTK
66	J3QSB4	STESLQAN#VQR
67	P51991-2	YHTIN#GHN#CEVK
68	H0YI18	N#GESSELDLQGIR
69	P07305-2	VGEN#ADSQIK

70	P62701	DAN#GN#SFATR
71	P10809	N#AGVEGSLIVEK
72	D6RGK8	DETN#YGIPQR
73	P04406	VGVN#GFGR
74	O00160	EEASN#ILLNK
75	Q7KZF4	N#LPGLVQEGEPFSEEATLFTK
76	P08670	FLEQQN#K
77	P10809	VTDALN#ATR
78	Q00610-2	IYIDSN#NNPER
79	P08670	QDVDN#ASLAR
80	H0YEC5	EN#PLLPEEEEQR
81	J3QR53	lTN#VPAcVLITPGDSK
82	H0YJC0	LSDGFN#GADLR
83	Q9H2G2	EVIN#EVEK
84	H7C314	TVQGSGHQEHIN#IHK
85	Q9Y2Q5-3	N#GNQAFNEDNLK
86	Q12972-2	GLLGLPEEETELDN#LTEFNTAHNK
87		LN#LN#EVEK
88	J3KQI1	DAN#LGLK
89	P20700	EYEAALN#SK
90	P10809	LVQDVANN#TnEEAGDGTTTATVLAR
91	P05387	VISELN#GK
92	Q8N1T3-3	TLKAEQAEYEMEGIEWEPIKYFN#N#K
93	Q5H928	LVAGEMGQN#EPDQGGQR
94		GN#SATMVRVVPYAAIQFSAHEEYK
95	Q12906-5	SIGTAN#RPMGAGEALR
96	Q6UXH1-6	KNEN#CYN#TPGSYVCVCPDGFEETEDACVPPAEAGEWHGCPPHR
97	P46779-4	YN#GLIHR
98	H7C367	MEELHN#QEVQK
99	H3BSW0	TIDLSN#NK
100	P04406	VIPELN#GK
101	C9J3L8	GEDFPAN#NIVK
102	A6NGH7	LN#LENR
103	B4DS77-2	DLN#IVVHVQHYENMDTR
104	P38646	DDIEN#MVK
105	Q9NTI2-3	MDTTSDNFGYN#LLTFIILYNNLIPISLLVTLEVVK
106	A6NCW0	FLQEQN#K

107	P41218	N#LVNNLR
108	Q9Y2Q5-3	N#GN#QAFNEDNLK
109	P04264;P35908	IEISELN#R
110	H3BMH2	N#EFNLESK
111	B7ZAR1	TSLGPN#GLDK
112	F8VV53	N#GSLICTASK
113	H7C5S7	LN#EIVGNEDTVSR
114	P41218	N#LVN#NLR
115	O00116	FGGLAAGEDN#GQR