Electronic Supporting Information for:

Unprecedented High Efficient Capture of Glycopeptides by Fe_3O_4@Mg-MOF-74 core-shell nanoparticles

Jie Li,[a] Jiaxi Wang,[b] Yun Ling [a,*], Zhenxia Chen [a], Mingxia Gao,[b] Xiangmin Zhang,[b] and Yaming Zhou [a,*]

[a] Shanghai Key Laboratory of Molecular Catalysis and Innovative Materials, Department of Chemistry, Fudan University, Shanghai, 200433, China.

[b] Department of Chemistry and Institutes of Biomedical Sciences, Fudan University, Shanghai 200433, China.

To whom correspondence should be addressed:

yunling@fudan.edu.cn (Prof. Ling Y.),

ymzhou@fudan.edu.cn (Prof. Zhou YM);

Experimental Section

Materials and chemicals: Horse-radish peroxidase (HRP), 2,5-dihydroxyterephthalic acid (H_4DOT), ammonium bicarbonate (NH_4HCO_3), bovine serum albumin (BSA), trypsin (from bovine pancreas), trifluoroacetic acid (TFA) and 2,5-dihydroxybenzoic acid (DHB) were purchased from Sigma Chemical. Acetonitrile (ACN) was purchased from Merck (Darmstadt, Germany). Human serum was obtained from Zhongshan hospital (Shanghai, China). Peptide-N-glycosidase (PNGase F) was obtained from New England Biolabs. Milli-Q water was used in all experiment process. All other reagents were obtained from Sinopharm Chemical Reagent (Shanghai, China).

Synthesis of Fe_3O_4@Mg-MOF-74 nanoparticles: The magnetite particles (Fe_3O_4) were hydrothermally synthesized.[1] The obtained Fe_3O_4 nanoparticles (0.05g) and $Mg(NO_3)_2$ were re-dispersed in the solution of DMF (67.5 mL), ethanol (4.5 mL) and water (4.5 mL) by ultrasonication for 30 min. After that, H_4DOT (0.250 g, 1.26 mmol) was added and sonicated for another 30 min. The mixture was sealed in a Teflon-lined stainless-steel autoclave and heated at 125 °C about 5 h. Finally, the resulting nanoparticles were collected by magnetic separation and washed with DMF and ethanol (3 × 10 mL), respectively. The obtained products were dried in vacuum at 50 ºC for further use.

Synthesis of Fe_3O_4@Mg-MOF-74-II nanoparticles: The procedure for preparation of the composite is similar to the synthesis of Fe_3O_4@Mg-MOF-74 except that the 3,3'-dihydroxy-[1,1'-biphenyl]-4,4'-dicarboxylic acid was used as the ligand.

Synthesis of Fe_3O_4@DOT: The procedure is similar to that of Fe_3O_4@Mg-MOF-74 without $Mg(NO_3)_2$•$6H_2O$.

Synthesis of Fe_3O_4@Mg/DOT: The procedure is similar to that of Fe_3O_4@Mg-MOF-74 except that the reaction temperature is settled to be 90 ºC, where Mg-MOF-74 could not be generated.

Characterization: Transmission electron microscopy (TEM) images and energy dispersive Xray (EDX) spectra were taken on a JEOL 2011 microscope (Japan) operated at 200 kV. Fourier-transform infrared (FT-IR) spectra were tested on a Nicolet Fourier spectrophotometer. Powder X-ray Powder diffraction (XRD) patterns were collected on a Bruker D8 X-ray diffractometer with Cu Kα radiation. Nitrogen sorption isotherms were measured at 77 K on ASAP 2020 analyzer. The Brunauer-Emmett-Teller (BET) method was utilized to calculate the specific surface area.

Sample preparation. The HRP standard glycoprotein was dissolved in 25 mM NH_4HCO_3 buffer, the solution was heated at 100 °C for 10 min. After the solution cooled to 25 °C, trypsin was added into the solution at an enzyme/substrate ratio of 1:50 (w/w) at 37 °C with overnight shaking. The treating process of human serum is the same as previously reported method.[2] Briefly, the serum was thawed on the ice, and 1 μL of serum sample was diluted with 20 μL ultrapure water. The mixture was treated with 10 mM DTT at 60 °C for 30 min and alkylated with 20 mM IAA at 37 °C for 1 h in the dark. In order to obtain the serum protein sample, acetone solution was added with a protein solution/acetone ratio of 1:6 (v/v) at -20 °C overnight. Prior to digestion, the solution was diluted with 50 mM NH_4HCO_3 until the protein final concentration reaches 2 mg/mL. Then, the solution was digested with trypsin (protein: enzyme = 30:1, w/w) for 16 h under shaking.

Glycopeptides enrichment: To evaluate the efficiency of materials for enrichment of glycopeptides, the composites (10 mg/ml, 50 μL) were first washed with the loading buffer (90% ACN and 1% TFA aqueous solution (v/v)) three times, followed by HRP digestion or

peptides mixture were added. After incubated for 45 min at 37 °C, the supernatant and materials were separated with the help of magnet, then the materials were rinsed with loading buffer (200 μL) for 3 times. Subsequently, 10 μL of 30 % ACN and 0.1% TFA aqueous solution (v/v) was added into and vibrated for 15 min to elute the enriched glycopeptides.

Optimizing of enrichment conditions: The amount of the composite and the time of reaction were optimized, respectively. Different concentrations of ACN and TFA solution were applied to the loading buffer. All these optimizing procedure were investigated with trypsin digest of 10^{-6} M HRP.

MALDI-TOF-MS analysis. The glycopeptides sample elution (1 μL) was deposited on the target and then another DHB aqueous solution was introduced as a matrix. MALDI-TOF-MS experiments were performed in positive ion mode on a 5800 Proteomics Analyzer (Applied Biosystems, Framingham, MA, USA) with the Nd-YAG laser at 383 nm, a repetition rate of 200 Hz and an acceleration voltage of 20 kV.

Nano-Liquid chromatography tandem mass spectrometry (Nano-LC−MS/MS) analysis of glycopeptides. Glycopeptides enriched from human serum were re-suspended with 12 μL solvent A (A: water with 0.1% formic acid; B: ACN with 0.1% formic acid) and then separated by nanoLC and analyzed by on-line electrospray tandem mass spectrometry. The experiments were performed on a Nano Aquity UPLC system (Waters Corporation, Milford, MA) connected to a quadrupole-Orbitrap mass spectrometer (Q-Exactive, Thermo Fisher Scientific, Bremen, Germany) equipped with an online nano-electrospray ion source. 8 μL peptide sample was loaded onto the trap column (Thermo Scientific Acclaim PepMap C18, 100μm x 2cm), with a flow of 10μL/min for 3 min and subsequently separated on the analytical column (Acclaim PepMap C18, 75μm x 50cm) with a linear gradient, from 2% B to 45% B in 75 min. The column was re-equilibrated at initial conditions for 15 min. The

column flow rate was maintained at 300 nL/min and column temperature was maintained at 40 °C. The electrospray voltage of 1.8 kV versus the inlet of the mass spectrometer was used.

The Q-Exactive mass spectrometer was operated in the data-dependent mode to switch automatically between MS and MS/MS acquisition. Survey full-scan MS spectra (m/z 350-1600) were acquired with a mass resolution of 70K, followed by twenty sequential high energy collisional dissociation (HCD) MS/MS scans with a resolution of 17.5K. For MS, the automatic gain control (AGC) was set to 1000 000 ions, with maximum accumulation times of 20 ms. For MS/MS, precursor ions were activated using 30% normalized collision energy, an isolation window of 2 m/z and the AGC was set to 100 000 ions, with maximum accumulation time of 120 ms. Single charge state was rejected and dynamic exclusion was used with one microscan and 20 s exclusion duration.

Database search. Tandem mass spectra were extracted by Proteome Discoverer software (Thermo Fisher Scientific, version 1.4.0.288). Charge state deconvolution and deisotoping were not performed. All MS/MS samples were analyzed using Mascot (Matrix Science, London, UK; version 2.3.2). The database was the Human UniProtKB/Swiss-Prot database (Release 2015-03-11, with 20199 sequences). Raw files generated by the Orbitrap Fusion were searched directly using a 10 ppm precursor mass tolerance and a 50 mmu fragment mass tolerance. The enzyme specificity with trypsin was used. Up to two missed cleavages were allowed and peptides with at least 7 amino acids were retained. Carbamidomethyl on cysteine was set as a fixed modification. Oxidation on methionine and Deamidation on asparagine were set as variable modifications. Use the percolator algorithm to control peptide level false discovery rates (FDR) lower than 1%. The Asn modification that did not occur in the N-X-S/T (X≠P) sequon was eliminated to ensure the false positive rate below 1% for the identified glycosylation sites.

Fig. S1 TEM images of as-prepared Fe_3O_4@Mg-MOF-74 composites with the different concentration of the MOF precursor: (a) $Mg(NO_3)_2$•$6H_2O$ (7.3 mM), Ligand (4.4 mM); (b) $Mg(NO_3)_2$•$6H_2O$ (9 mM), Ligand (8 mM); (c) $Mg(NO_3)_2$•$6H_2O$ (18 mM), Ligand (16 mM); (d) $Mg(NO_3)_2$•$6H_2O$ (36 mM), Ligand (32 mM).

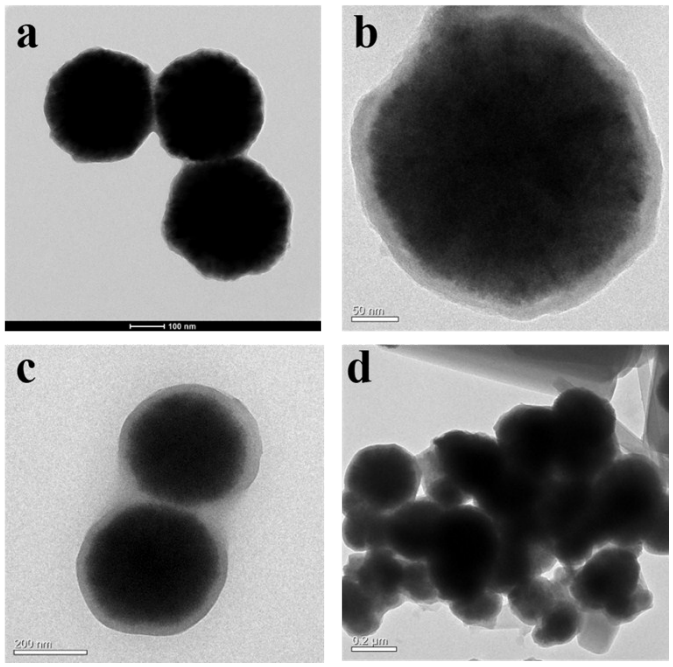

Fig. S2 The energy dispersive X-ray spectrometer (EDX) of Fe_3O_4@Mg-MOF-74 composites.

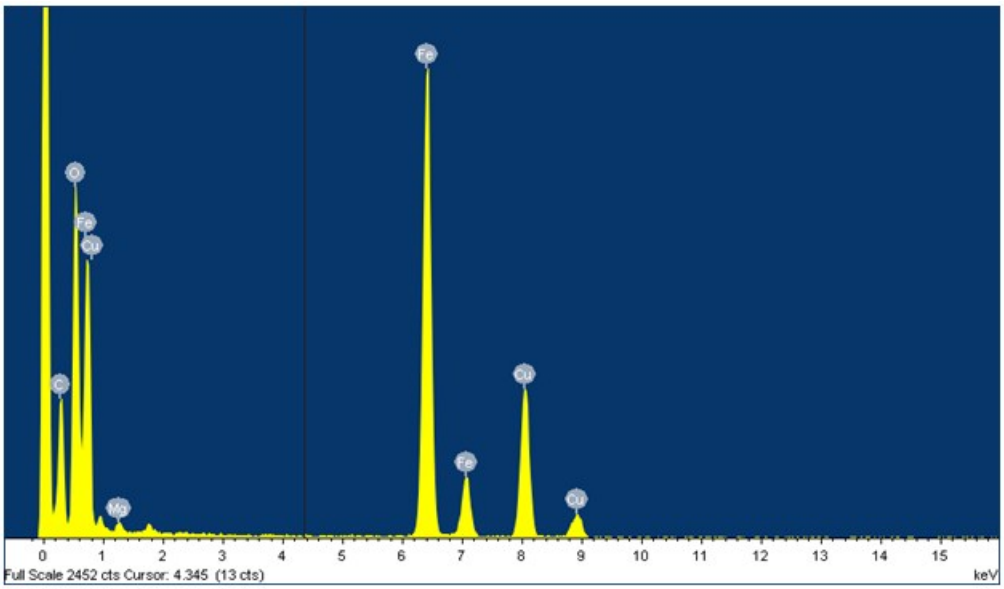

Fig. S3 FT-IR spectra of Fe_3O_4 (the band at 1611 cm^{-1} assigned to ν_{COO^-}), Fe_3O_4@Mg-MOF-74 and Mg-MOF-74.

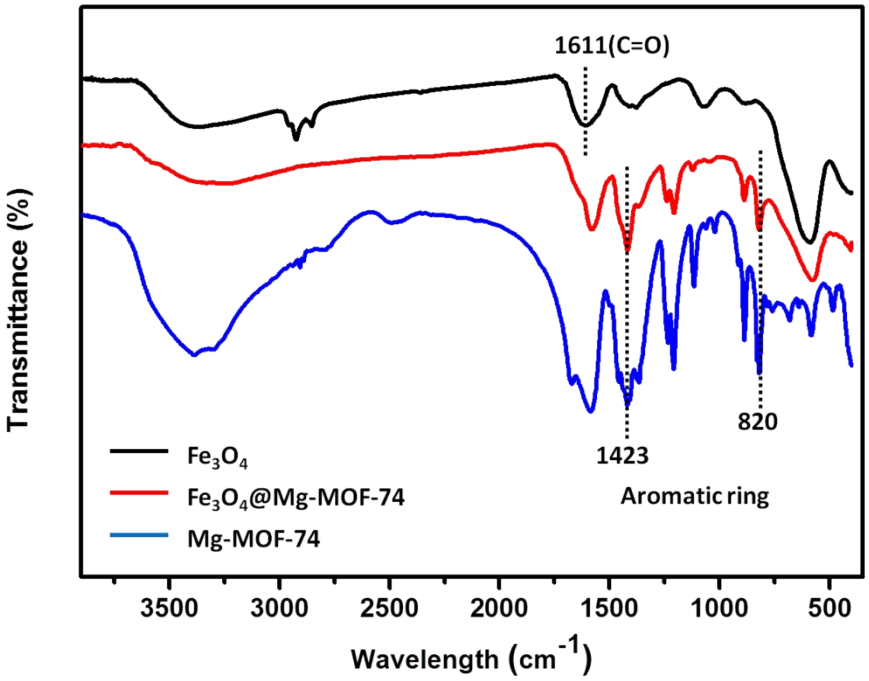

Fig. S4 The PXRD pattern of Fe_3O_4 and Fe_3O_4@Mg-MOF-74.

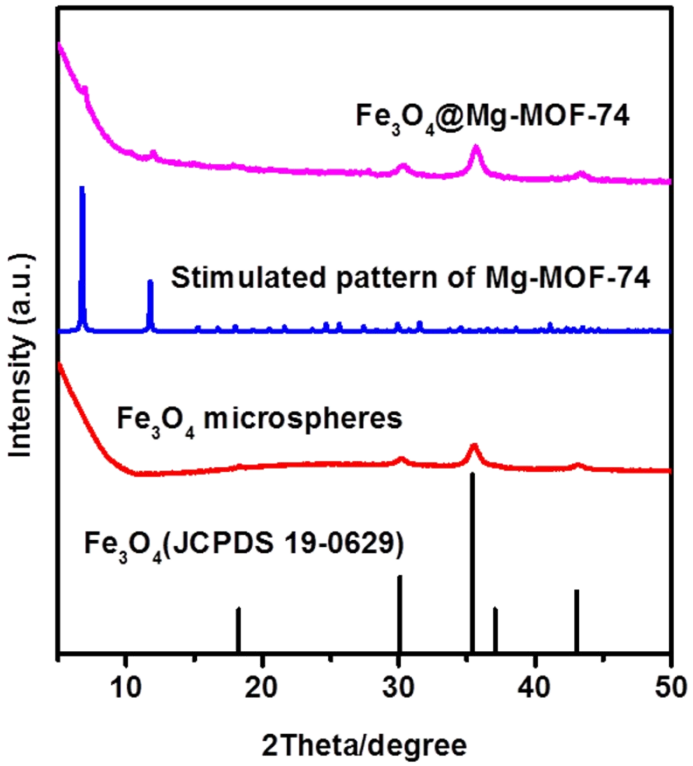

Fig. S5 (a) N_2 adsorption-desorption isotherms of Fe_3O_4@Mg-MOF-74 at 77 K. Pore size distributions calculated by H-K for the analysis of micropore (b) and BJH for the mesopore (c), respectively.

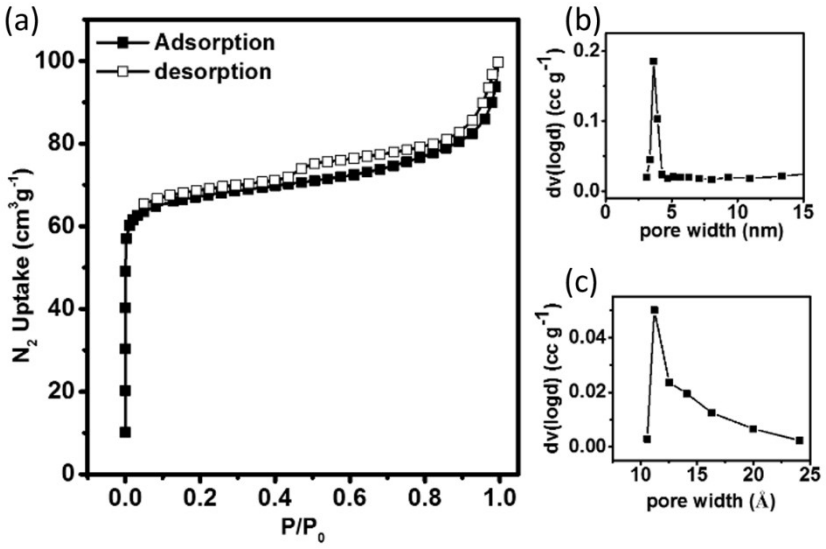

Fig. S6 The TEM and PXRD images of Fe_3O_4@Zn-MOF-74 (a, c) and Fe_3O_4@Mg-MOF-74-II particles (b, d).

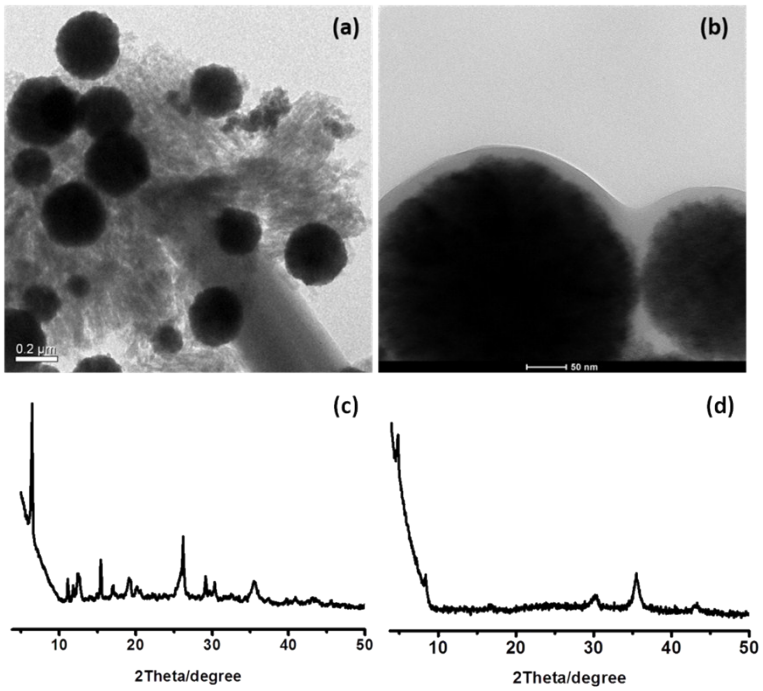

Fig. S7 MALDI-TOF-MS analysis of glycopeptides derived from HRP enriched by Fe_3O_4@Mg-MOF-74 composites with different buffer solution (a) 85% ACN 1% TFA; (b) 95% ACN 1% TFA; (c) 85% ACN 0.1% TFA and (d) 95% ACN 0.1% TFA.

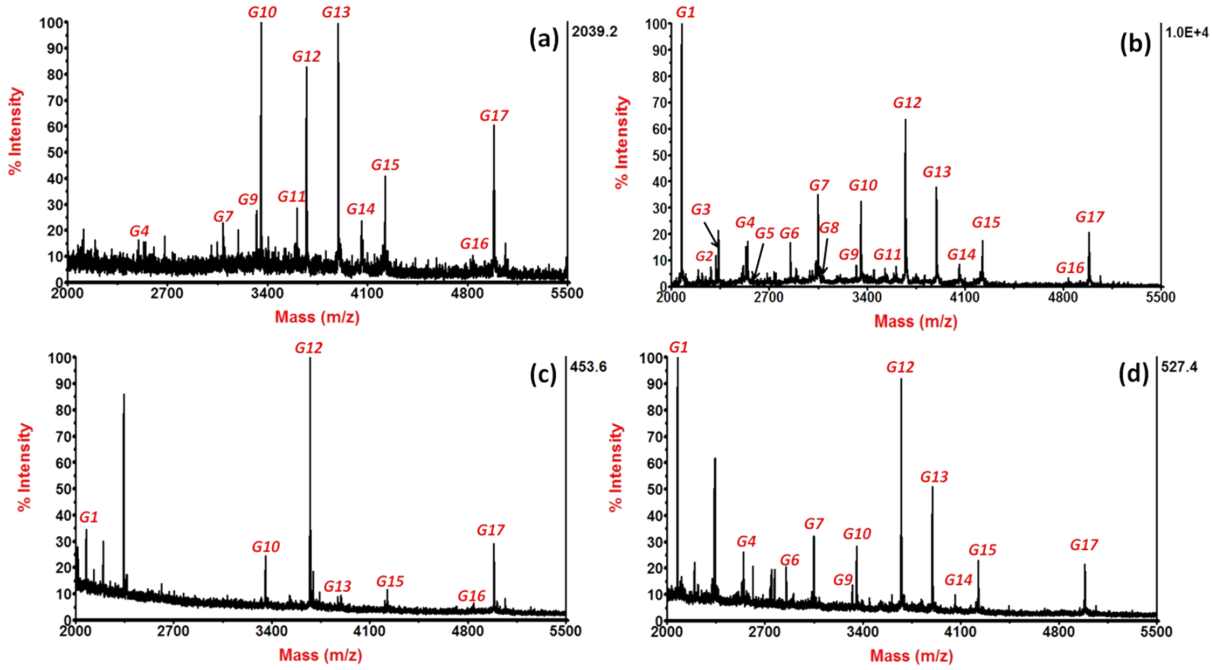

Fig. S8 The effect of different enrichment time (a) and amount of materials (b) on glycopeptides enriched from HRP tryptic digest at 37 °C.

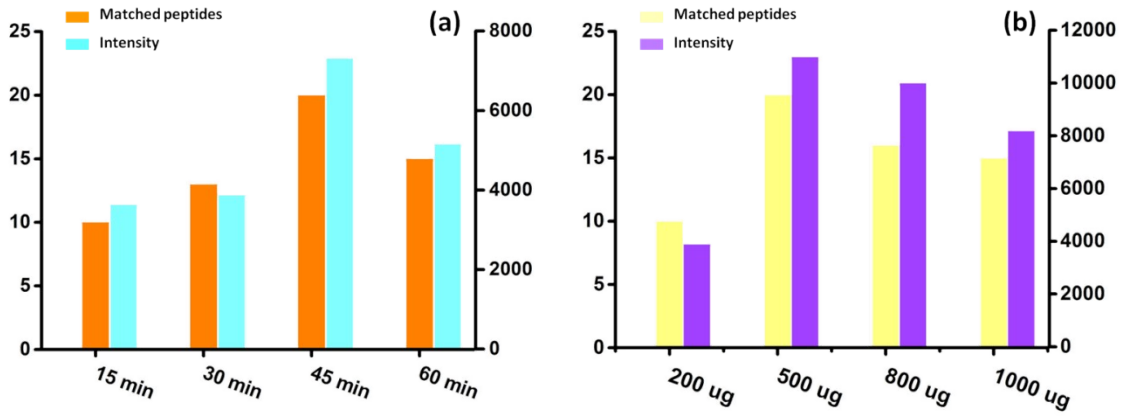

Fig. S9 MALDI-TOF-MS of tryptic digested HRP (10^{-6} M): after enrichment by (a) Fe_3O_4, (b) Fe_3O_4@ligand, (c) Fe_3O_4@Mg/ligand and (d) Fe_3O_4@Mg-MOF-74-II.

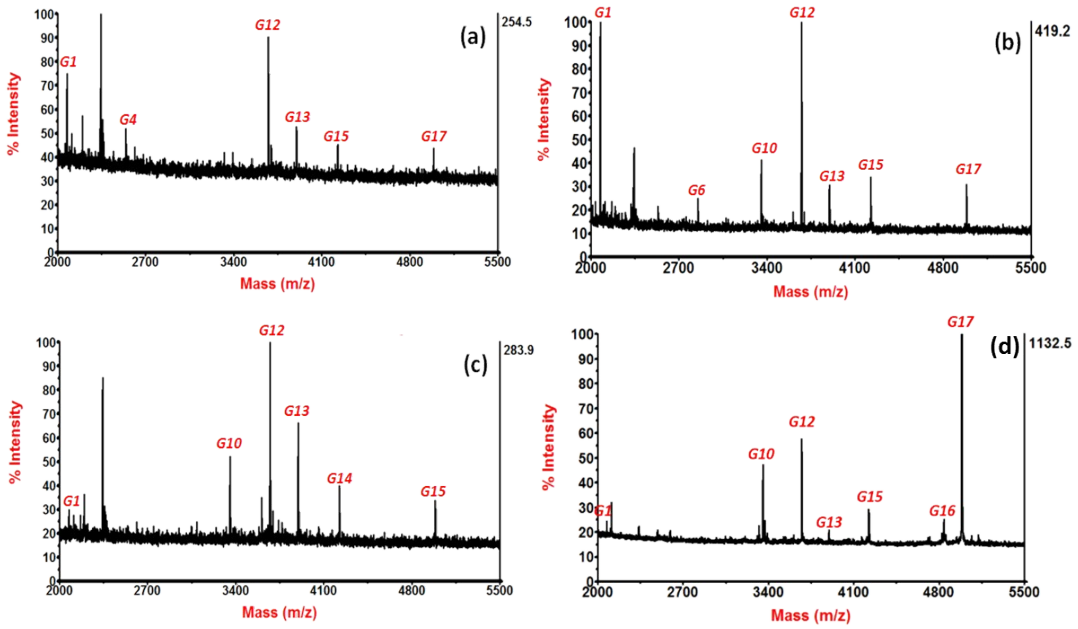

Fig. S10 MALDI-TOF-MS for glycopeptides enriched by Fe_3O_4@Mg-MOF-74 from HRP digests with different concentrations: (a) 100 fmol/μL, (b) 10 fmol/μL, (c) 1 fmol/μL and (d) 0.5 fmol/μL.

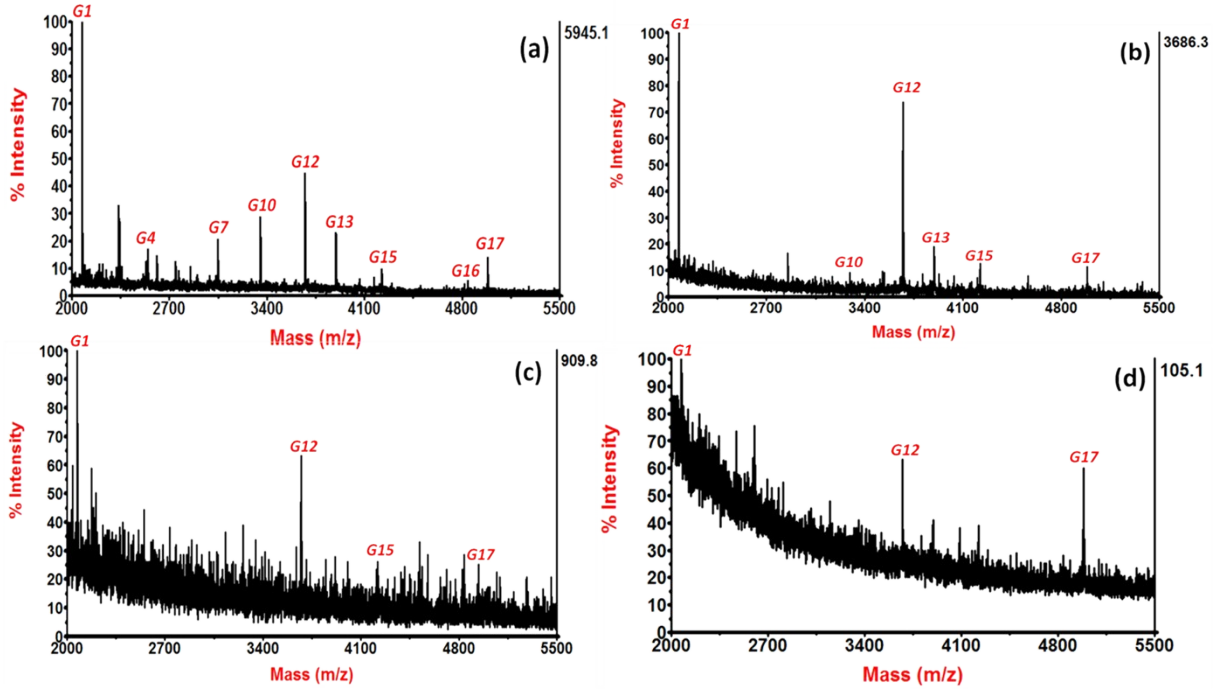

Fig. S11 The cycling performance of the Fe_3O_4@Mg-MOF-74: (a) the first time, (b) the second time, (c) the third time, (d) the fourth time, (e) the fifth time and (f) after enrichment with the Fe_3O_4@Mg-MOF-74 placed for three months.

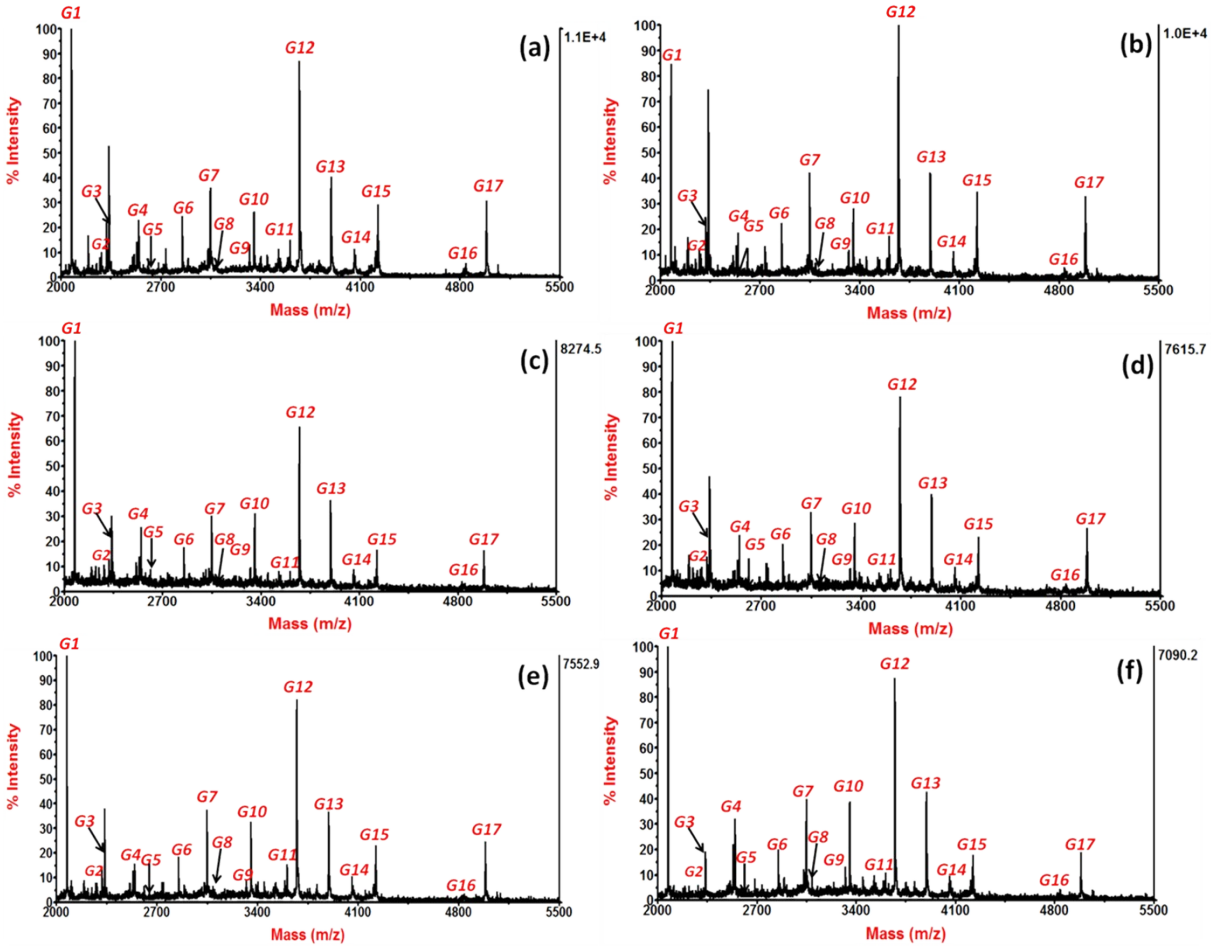

Table S1. Detailed information of the observed glycopeptides in HRP tryptic digest**.** N# denotes the N-linked glycosylation site.

No.	Observed m/z	Glycan composition	Amino acid sequence
1	2074	XylMan3GlcNAc2	PN#VSNIVRRR
2	2290	XylMan2GlcNAc2	SILLDN#TTSFR
3	2321	Man2GlcNAc2	MGN#ITPLTGTQGQIR
4	2543	XylMan3FucGlcNAc2	SSPN#ATDTIPLVR
5	2612	XylMan3GlcNAc2	MGN#ITPLTGTQGQIR
6	2850	FucGlcNAc	GLIQSDQELFSSPN#ATDTIPLVR
7	3048	XylMan2GlcNAc2	SFAN#STQTFFNAFVEAMDR
8	3087	XylMan3FucGlcNAc2	GLCPLNGN#LSALVDFDLR
9	3323	XylMan3FucGlcNAc2	QLTPTFYDNSCPN#VSNIVR
10	3354	XylMan3FucGlcNAc2	SFAN#STQTFFNAFVEAMDR
11	3606	XylMan3FucGlcNAc2	NQCRGLCPLNGN#LSALVDFDLR
12	3672	XylMan3FucGlcNAc2	GLIQSDQELFSSPN#ATDTIPLVR
13	3894	XylMan3FucGlcNAc2	LHFHDCFVNGCDASILLDN#TTSFR
14	4057	XylMan3GlcNAc2	QLTPTFYDNSC(AAVESACPR)PN#VSNIVR-H2O
15	4222	XylMan3FucGlcNAc2	QLTPTFYDNSC(AAVESACPR)PN#VSNIVR
16	4839	XylMan3FucGlcNAc2, XylMan3GlcNAc2	LYN#FSNTGLPDPTLN#TTYLQTLR
17	4984	XylMan3FucGlcNAc2, XylMan3FucGlcNAc2	LYN#FSNTGLPDPTLN#TTYLQTLR

Table S2. Identified N-glycoeptides containing deamidated Asn within N-X-S/T (X≠P) from 1μL human serum digest. N# denotes the N-linked lycosylation site.

No.	Peptide Sequence	Protein Group Accessions
1	AALAAFNAqnN#GSNFQLEEISR	P02765
2	LEPVHLQLqcMSqEqLAqVAAN#ATK	Q96PD5
3	LEPVHLQLqcMSQEQLAQVAAN#ATK	Q96PD5
4	LDAPTNLqFVN#ETDSTVLVR	P02751
5	ADGTVNqIEGEATPVN#LTEPAK	P05090
6	ADGTVNqIEGEATPVN#LTEPAKLEVK	P05090
7	LSVDKDqYVEPEN#VTIqcDSGYGVVGPQSITcSGnR	P04003
8	LSVDKDqYVEPEN#VTIQcDSGYGVVGPQSITcSGnR	P04003
9	LDPVSLqTLQTWN#TSYPK	Q99784
10	GLTFQqN#ASSmcVPDQDTAIR	P01871
11	ALPQPqN#VTSLLGcTH	P02790
12	GLTFQqN#ASSMcVPDQDTAIR	P01871
13	GGETAqSADPQWEQLNnKN#LSMPLLPADFHK	P05546
14	GGETAqSADPQWEQLnnKN#LSmPLLPADFHK	P05546
15	SVQEIqATFFYFTPN#KTEDTIFLR	P02763
16	SVQEIqATFFYFTPN#KTEDTIFLR	P19652
17	YEEQqLEIQN#SSR	Q8IZF2
18	SVqEIqATFFYFTPN#KTEDTIFLR	P02763
19	SVqEIqATFFYFTPN#KTEDTIFLR	P19652
20	QQqHLFGSN#VTDcSGnFcLFR	P02787
21	VTqVYAEN#GTVLQGSTVASVYK	P27169
22	EEqFN#STFR	P01859
23	QNqcFYN#SSYLNVQR	P19652
24	NPPMGGNVVIFDTVITNqEEPYQN#HSGR	P02745
25	VLTLNLDQVDFqHAGN#YScVASNVQGK	P07333
26	qVHFFVN#ASDVDNVK	Q96IY4
27	mALSWVLTVLSLLPLLEAqIPLcAnLVPVPITN#ATLDqITGK	P02763
28	GNEANYYSN#ATTDEHGLVqFSIN#TTNVMGTSLTVR	P01023
29	VGQLQLSHN#LSLVILVPqNLK	P05155
30	ANLTNFPEN#GTFVVnIAQLSQDDSGR	P01833
31	LAGKPTHVN#VSVVmAEVDGTcY	P01876
32	STGKPTLYN#VSLVMSDTAGTcY	P01871
33	LAGKPTHVN#VSVVMAEVDGTcY	P01876
34	QQQHLFGSN#VTDcSGnFcLFR	P02787
35	QQQHLFGSN#VTDcSGNFcLFR	P02787
36	ERSWPAVGN#cSSALR	P02790
37	ITYSIVQTN#cSK	P01042
38	FEVDSPVYN#ATWSASLK	P04114
39	ELHHLQEQN#VSNAFLDK	P00450
40	GFGVAIVGN#YTAALPTEAALR	Q96PD5
41	LGSFEGLVN#LTFIHLQHNR	P51884
42	KLPPGLLAN#FTLLR	P02750
43	TKPREEQFN#STFR	P01859
44	LSLLEEPGN#GTFTVILNQLTSR	P01833
45	VGQLQLSHN#LSLVILVPQNLK	P05155
46	LNPTVTYGN#DSFSAK	P05362
47	NAHGEEKEN#LTAR	Q06033
48	HGIQYFNN#NTqHSSLFMLNEVK	P01042
49	VYKPSAGN#nSLYR	P02749

50	TVLTPATnHMGN#VTFTIPAnR	P01024
51	HIPGLIHN#mTAR	Q16610
52	TVLTPATnHmGN#VTFTIPANR	P01024
53	VLYLAAYN#cTLRPVSK	Q9UGM5
54	YPHKPEIN#STTHPGADLQENFcR	P00734
55	FSDGLESN#SSTQFEVKK	P0C0L5
56	FSDGLESN#SSTQFEVK	P0C0L5
57	FSDGLESN#SSTQFEVKK	P0C0L4
58	FSDGLESN#SSTQFEVK	P0C0L4
59	VYKPSAGN#NSLYR	P02749
60	KLHINHNN#LTESVGPLPK	P51884
61	VTQVYAEN#GTVLQGSTVASVYK	P27169
62	LPPGLLAN#FTLLR	P02750
63	KQVHFFVN#ASDVDNVK	Q96IY4
64	FVGTPEVN#QTTLYQR	P01033
65	LLDLSGNN#LTHLPK	P40197
66	VSTVYANN#GSVLQGTSVASVYHGK	Q15166
67	AALAAFnAqNN#GSNFqLEEISR	P02765
68	LYLGSNN#LTALHPALFQN#LSK	P22792
69	SLPNFPN#TSATAN#ATGGR	Q6UXB8
70	AALAAFnAQNN#GSNFQLEEISR	P02765
71	DIVEYYN#DSN#GSHVLqGR	P25311
72	DIVEYYN#DSN#GSHVLQGR	P25311
73	GLTFQQN#ASSmcVPDQDTAIR	P01871
74	QVFPGLnYcTSGAYSN#ASSTDSASYYPLTGDTR	P04114
75	LYLGSNN#LTALHPALFQN#LSK	P22792
76	SWPAVGN#cSSALR	P02790
77	SEGSSVN#LSPPLEQcVPDR	P00734
78	ALPQPQN#VTSLLGcTH	P02790
79	GLTFQQN#ASSMcVPDQDTAIR	P01871
80	LSDLSIN#STEcLHVHcR	P05156
81	LQNNENN#IScVER	Q03591
82	FVEGSHN#STVSLTTK	P04114
83	KKEDALN#ETR	P10909
84	LHINHNN#LTESVGPLPK	P51884
85	DLDMFIN#ASK	O75882
86	DFYVDEN#TTVR	P29622
87	YAEDKFN#ETTEK	P43652
88	AFGSNPN#LTK	P22792
89	IYSGILN#LSDITK	P03952
90	QESMNSN#VSVYQPPR	P13598
91	QVHFFVN#ASDVDNVK	Q96IY4
92	EGHFYYN#ISEVK	P55058
93	EHEAQSN#ASLDVFLGHTNVEELMK	P00736
94	VYSGILN#QSEIK	P03951
95	EVFVHPN#YSK	P04070
96	VVLHPN#YSqVDIGLIK	P00738
97	SLGNVN#FTVSAEALESqELcGTEVPSVPEHGRK	P01023
98	LSALDnLLN#HSSMFLK	Q9Y4L1
99	LYLGSnN#LTALHPALFQN#LSK	P22792
100	LHINHnN#LTESVGPLPK	P51884
101	AVLQLnEEGVDTAGSTGVTLN#LTSKPIILR	P08185
102	MVSHHN#LTTGATLInEqWLLTTAK	P00739
103	MVSHHN#LTTGATLInEqWLLTTAK	P00738
104	MVSHHN#LTTGATLInEQWLLTTAK	P00739
105	MVSHHN#LTTGATLInEQWLLTTAK	P00738

106	VASVInINPN#TTHSTGScR	P13473
107	IIVPLnNREN#ISDPTSPLR	P01591
108	GAFISN#FSmTVDGK	P19823
109	MDGASN#VTcInSR	P08603
110	MDGASN#VTcINSR	P08603
111	SLGNVN#FTVSAEALESQELcGTEVPSVPEHGRK	P01023
112	LQAPLN#YTEFQKPIcLPSK	P03952
113	LETTVN#YTDSQRPIcLPSK	P03951
114	KFVQGN#STEVAcHPGYGLPK	Q02985
115	MVSHHN#LTTGATLINEQWLLTTAK	P00739
116	VVLHPN#YSQVDIGLIK	P00738
117	VVLHPN#YSQVDIGLIKLK	P00738
118	MVSHHN#LTTGATLINEQWLLTTAK	P00738
119	MVSHHN#LTTGATLINEQWLLTTAK	P00739
120	MVSHHN#LTTGATLINEQWLLTTAK	P00738
121	SPDVIN#GSPISQK	P08603
122	KYNSQN#QSNNQFVLYR	P01042
123	LNAENN#ATFYFK	P01042
124	GAFISN#FSMTVDGK	P19823
125	TPLTAN#ITK	P01877
126	ALGFEN#ATQALGR	Q08380
127	IDSTGN#VTNELR	O75882
128	DIENFN#STQK	P43652
129	IVGGTN#SSWGEWPWQVSLQVK	P03952
130	WSDIWN#ATK	P06276
131	LVPHMN#VSAVEK	Q96KN2
132	SLEAIN#GSGLQMGLQR	P09172
133	LPYQGN#ATMLVVLMEK	Q9UK55
134	LHEITN#ETFR	Q6EMK4
135	LPTQN#ITFqTESSVAEqEAEFQSPK	Q14624
136	LPTQN#ITFqTESSVAEQEAEFQSPK	Q14624
137	YNSQN#qSNNQFVLYR	P01042
138	YGNPN#ETQN#N#STSWPVFK	P06276
139	NLFLN#HSEN#ATAK	P00739
140	NLFLN#HSEN#ATAK	P00738
141	LNAENN#ATFYFK	P01042
142	SEGTN#STLTLSPVSFEnEHSYLcTVTcGHK	P19320
143	EGYSN#ISYIVVnHqGISSR	P49908
144	EGYSN#ISYIVVnHQGISSR	P49908
145	YGNPN#ETQNN#STSWPVFK	P06276
146	ISEEN#ETTcYmGK	P08603
147	ISEEN#ETTcYMGK	P08603
148	VIDFN#cTTSSVSSALAnTK	P04196
149	VIDFN#cTTSSVSSALANTK	P04196
150	AELSN#HTRPVILVPGcLGNqLEAK	P04180
151	AELSN#HTRPVILVPGcLGNQLEAK	P04180
152	EPGSN#VTMSVDAEcVPMVR	Q08380
153	YLPVN#SSLLTSDcSER	Q9Y6R7

154	FVQGN#STEVAcHPGYGLPK	Q02985
156	NLFLN#HSENATAK	P00739
157	NLFLN#HSENATAK	P00738
158	EDALN#ETR	P10909
159	YNSQN#QSNNQFVLYR	P01042
160	EEQYN#STYR	P01857
161	SLTFN#ETYQDISELVYGAK	P01008
162	AFITN#FSMIIDGMTYPGIIK	Q14624
163	EEQFN#STYR	P01861
164	EEQYN#STFR	P01860
165	TMFPN#LTDVR	P06681
166	MFSQN#DTR	P02750
167	EGYSN#ISYIVVNHQGISSR	P49908
168	EEQFN#STFR	P01859
169	GVNFN#VSK	P03952
170	DTFVN#ASR	P05155
171	VQPFN#VTQGK	P13473
172	QGGVN#ATQVLIQHLR	Q13790
173	ETFFN#LSKR	Q9UK55
174	ETFFN#LSK	Q9UK55
175	FQLLN#FSSSELK	Q9BY67
176	YTGN#ASALFILPDqDKMEEVEAmLLPETLKR	P01011
177	SLGnVN#FTVSAEALESqELcGTEVPSVPEHGR	P01023
178	SLGnVN#FTVSAEALESqELcGTEVPSVPEHGRK	P01023
179	SLGnVN#FTVSAEALESQELcGTEVPSVPEHGR	P01023
180	VNQnLVYESGSLN#FSK	P04114
181	THTN#ISESHPN#ATFSAVGEASIcEDDWnSGER	P01871
182	YDFN#SSmLYSTAK	P04114
183	YTGN#ASALFILPDQDKmEEVEAMLLPETLK	P01011
184	EWDN#TTTEcR	P20851
185	FLNN#GTcTAEGK	P05156
186	GTAN#TTTAGVPcQR	P26927
187	YTGN#ASALFILPDQDK	P01011
188	YLGN#ATAIFFLPDEGK	P01009
189	GHVN#ITR	P00734
190	WVSN#KTEGR	P01008
191	AFEN#VTDLQWLILDHNLLENSK	P51884
192	VPGN#VTAVLGETLK	P01833
193	VLSN#NSDANLELINTWVAK	P05155
194	DFVN#ASSKYEITTIHNLFR	P05546
195	DFVN#ASSK	P05546
196	VTQN#LTLIEESLTSEFIHDIDR	P36955
197	IYSN#HSALESLALIPLQAPLK	P55058
198	TLN#qSSDELQLSmGNAMFVK	P01011
199	TLN#qSSDELQLSMGNAMFVK	P01011
200	LQN#LTLPTN#ASIK	Q13201
201	LQnNENN#IScVER	Q03591

202	FLnN#GTcTAEGK	P05156
203	YK N#nSDISSTR	P01871
204	GLN#VTLSSTGRnGFK	P0C0L5
205	GLN#VTLSSTGRnGFK	P0C0L4
206	MLN#TSSLLEQLnEQFNWVSR	P10909
207	LGN#WSAmPScK	P02749
208	FLN#DTmAVYEAK	P29622
209	TLN#QSSDELQLSMGNAmFVK	P01011
210	VSN#VScQASVSR	P55058
211	WVN#YScLDQAR	Q76LX8
212	ISN#SSDTVEcEcSENWK	O75882
213	LGN#WSAMPScK	P02749
214	GLN#VTLSSTGR	P0C0L5
215	GLN#VTLSSTGR	P0C0L4
216	YK N#NSDISSTR	P01871
217	MLN#TSSLLEQLNEQFNWVSR	P10909
218	LAN#LTQGEDQYYLR	P10909
219	GLN#LTEDTYKPR	Q08380
220	HAN#WTLTPLK	P27169
221	FLN#DTMAVYEAK	P29622
222	AVN#ITSENLIDDVVSLIR	P02748
223	VVN#STTGPGEHLR	P07996
224	VLN#FTTK	P13671
225	VQN#MSQSIEVLDR	Q99784
226	TWN#QSIALR	P12259
227	AFN#STLPTMAQMEK	P16070
228	LFN#VTPQDEQK	O75144
229	KIVLDPSGSMNIYLVLDGSDSIGASN#FTGAK	P00751
230	IVLDPSGSMNIYLVLDGSDSIGASN#FTGAK	P00751
231	RNPPMGGNVVIFDTVITNQEEPYQN#HSGR	P02745
232	NPPMGGNVVIFDTVITNQEEPYQN#HSGR	P02745
233	AVLQLNEEGVDTAGSTGVTLN#LTSKPIILR	P08185
234	FN#SSYLQGTnQITGR	P04114
235	HN#STGcLR	P10909
236	WN#NTGcQALPSQDEGPSK	P01833
237	RN#HScEPcQTLAVR	P00748
238	FN#LTETSEAEIHQSFQHLLR	P01011
239	EN#GTISR	P02763
240	FN#SSYLQGTNQITGR	P04114
241	EN#LTAPGSDSAVFFEQGTTR	P00450
242	EN#GTVSR	P19652
243	DN#YTKAEEILSR	P06276
244	YN#FSFR	P01031
245	GN#ETLHYETFGK	P13598
246	DN#YTDLVAIQNK	P14151
247	EN#ISDPTSPLR	P01591
248	KN#QSVNVFLGHTAIDEMLK	Q9NZP8

249	DTTVLVSPSSILEEGSSVN#MTcLSqGFPAPK	P19320
250	LVLSSEKTVLTPATNHMGN#VTFTIPANR	P01024
251	LSLHRPALEDLLLGSEAN#LTcTLTGLR	P01876
252	LSLHRPALEDLLLGSEAN#LTcTLTGLR	P01877
253	MPSQAPTGNFYPQPLLN#SSMcLEDSR	O75882
254	IQSPLFTLDANADIGN#GTTSANEAGIAASITAK	P04114
255	ADGTVNQIEGEATPVN#LTEPAK	P05090
256	ADGTVNQIEGEATPVN#LTEPAKLEVK	P05090
257	DVQIIVFPEDGIHGFN#FTR	P43251
258	MAWPEDHVFISTPSFN#YTGR	P04180
259	LVLSSEKTVLTPATnHMGN#VTFTIPANREFK	P01024
260	VTVQSLLTVETLEHN#QTYEcR	P07333
261	SVQEIQATFFYFTPN#KTEDTIFLR	P02763
262	SVQEIQATFFYFTPN#KTEDTIFLR	P19652
263	SVQEIQATFFYFTPN#KTEDTIFLR	P02763
264	SVQEIQATFFYFTPN#KTEDTIFLR	P19652
265	LDAFFALEGFPTEPN#SSSR	P08294
266	LSHNELADSGIPGnSFN#VSSLVELDLSYNK	P51884
267	ADTHDEILEGLNFN#LTEIPEAQIHEGFQELLR	P01009
268	TLYETEVFSTDFSN#ISAAK	P05543
269	VNQNLVYESGSLN#FSKLEIqSqVDSQHVGHSVLTAK	P04114
270	LSVDKDQYVEPEN#VTIQcDSGYGVVGPQSITcSGnR	P04003
271	LSVDKDQYVEPEN#VTIQcDSGYGVVGPQSITcSGNR	P04003
272	LQAILGVPWKDKN#cTSR	P01019
273	VNQNLVYESGSLN#FSK	P04114
274	IYPGVDFGGEELN#VTFVK	P03952
275	LGTSLSSGHVLMN#GTLK	P80108
276	AAPAPQEATATFN#STADR	P13598
277	AAPAPQEATATFN#STADREDGHR	P13598
278	LDPVSLQTLQTWN#TSYPK	Q99784
279	GPSTPLPEDPNWN#VTEFHTTPK	P15144
280	TVLTPATNHMGN#VTFTIPAnR	P01024
281	ADTHDEILEGLnFN#LTEIPEAqIHEGFQELLR	P01009
282	LNAAKALPQPQN#VTSLLGcTH	P02790
283	TVLTPATNHMGN#VTFTIPANR	P01024
284	FNPGAESVVLSN#STLK	Q13201
285	SVVAPATDGGLN#LTSTFLR	P41222
286	AYLLPAPPAPGN#ASESEEDR	P17936
287	IQSPLFTLDAnADIGN#GTTSAnEAGIAASITAK	P04114
288	IQSPLFTLDAnADIGN#GTTSANEAGIAASITAK	P04114
289	AALAAFNAQNN#GSNFQLEEISR	P02765
290	LDAPTNLQFVN#ETDSTVLVR	P02751
291	QSVPAHFVALN#GSK	P06681
292	EYESYSDFERN#VTEK	P07358
293	EHEGAIYPDN#TTDFqR	P00450
294	SLDFNTLVDN#ISVDPETGDLWVGcHPNGMK	P27169
295	VASVININPN#TTHSTGScR	P13473

296	EHEGAIYPDN#TTDFQR	P00450
297	AQLLQGLGFN#LTER	P08185
298	DIVEYYNDSN#GSHVLQGR	P25311
299	AAIPSALDTN#SSK	Q08380
300	TVIRPFYLTN#SSGVD	Q08380
301	SQILEGLGFN#LTELSESDVHR	P29622
302	NPVGLIGAEN#ATGETDPSHSK	P43251
303	VVTVAALGTN#ISIHKDEIGK	Q9Y6R7
304	NPNNDQVFPN#GTLAPSIPIWGGSWR	Q9Y6R7
305	IIVPLNNREN#ISDPTSPLR	P01591
306	SMVDFMNTDN#FTSHR	Q99784
307	nqALN#LSLAYSFVTPLTSMVVTKPDDQEQSQVAEKPMEGESR	Q14624
308	nGTGHGN#STHHGPEYMR	P02790
309	nLFLN#HSEnATAK	P00739
310	nLFLN#HSEnATAK	P00738
311	nN#ATVHEQVGGPSLTSDLQAQSK	P04004
312	N#LSmPLLPADFHK	P05546
313	N#TTcqDLQIEVTVK	P0C0L5
314	N#TTcqDLQIEVTVK	P0C0L4
315	N#HScEPcQTLAVR	P00748
316	N#VScPQLEVPVcPSGFQLScK	P04275
317	N#TTcQDLQIEVTVK	P0C0L5
318	N#TTcQDLQIEVTVK	P0C0L4
319	N#HScSEGQISIFR	O75882
320	ncGVN#cSGDVFTALIGEIASPnYPKPYPENSR	P09871
321	N#LTDFAEQYSIQDWAK	P04114
322	N#NSDISSTR	P01871
323	N#LSMPLLPADFHK	P05546
324	N#LTTSLTESVDR	P80108
325	N#VSVEALASALQLLAR	Q13790
326	N#VSEDLPLPTFSPTLLGDSR	P11597
327	N#ATLVNEADKLR	P15144
328	N#GSGAVFPVAGADVQTLR	O00391
329	N#LTALPPDLPK	P07359
330	RNPPmGGNVVIFDTVITNQEEPYQN#HSGR	P02745
331	NPPmGGNVVIFDTVITNQEEPYQN#HSGR	P02745
332	TmFPN#LTDVR	P06681
333	TVLTPATNHmGN#VTFTIPANR	P01024
334	mVSHHN#LTTGATLINEqWLLTTAK	P00739
335	mVSHHN#LTTGATLINEqWLLTTAK	P00738
336	mDGASN#VTcINSR	P08603
337	mVSHHN#LTTGATLINEQWLLTTAK	P00739
338	mVSHHN#LTTGATLINEQWLLTTAK	P00738
339	mLN#TSSLLEqLNEQFNWVSR	P10909
340	mLN#TSSLLEQLNEQFNWVSR	P10909
341	QVFPGLNYcTSGAYSN#ASSTDSASYYPLTGDTR	P04114
342	VYLQGLIDcYLFGN#SSTVLEDSK	P35542

343	LKELPGVcN#ETMMALWEEcKPcLK	P10909
344	TELFSSScPGGIMLN#ETGqGYQR	O95445
345	TELFSSScPGGIMLN#ETGQGYQR	O95445
346	DTAVFEcLPqHAMFGN#DTITcTTHGN#WTK	P02749
347	DTAVFEcLPqHAmFGN#DTITcTTHGN#WTK	P02749
348	DTAVFEcLPQHAMFGN#DTITcTTHGN#WTK	P02749
349	WDPEVNcSmAQIQLcPPPPqIPnSHN#MTTTLNYR	P08603
350	WDPEVNcSmAQIQLcPPPPQIPnSHN#MTTTLNYR	P08603
351	DTAVFEcLPQHAmFGN#DTITcTTHGN#WTK	P02749
352	WDPEVNcSMAQIQLcPPPPQIPnSHN#MTTTLNYR	P08603
353	DAGVVcTN#ETR	Q08380
354	ELPGVcN#ETMmALWEEcKPcLK	P10909
355	ELPGVcN#ETMMALWEEcKPcLK	P10909
356	KEHETcLAPELYNGN#YSTTQK	P05160
357	KEHETcLAPELYnGN#YSTTQK	P05160
358	LTDTIcGVGN#MSAN#ASDQER	P06681
359	LTDTIcGVGN#mSAN#ASDQER	P06681
360	RGPEcSQN#YTTPSGVIK	O14786
361	EHETcLAPELYNGN#YSTTQK	P05160
362	EHETcLAPELYnGN#YSTTQK	P05160
363	AATcINPLN#GSVcERPAN#HSAK	O75882
364	QDQcIYN#TTYLnVQR	P02763
365	QDQcIYN#TTYLNVQR	P02763
366	QNQcFYN#SSYLNVQR	P19652
367	LGAcN#DTLqQLMEVFK	P01008
368	TLFcnASKEWDN#TTTEcR	P20851
369	LGAcN#DTLQQLmEVFK	P01008
370	LGAcN#DTLQQLMEVFK	P01008
371	TLFcN#ASK	P20851
372	WEYcN#LTR	Q16609
373	DKIcDLLVANNHFAHFFAPQN#LTNMNK	P19827
374	LGHcPDPVLVNGEFSSSGPVN#VSDK	P20851
375	LGHcPDPVLVnGEFSSSGPVN#VSDK	P20851
376	VTAcHSSQPN#ATLYK	P05543
377	KVScPIMPcSN#ATVPDGEccPR	P07996
378	IPcSqPPqIEHGTIN#SSR	P08603
379	KVcqDcPLLAPLN#DTR	P02765
380	SScGKEN#TSDPSLVIAFGR	P11279
381	GIcN#SSDVR	O75882
382	IPcSQPPQIEHGTIN#SSR	P08603
383	DQcIVDDITYNVN#DTFHK	P02751
384	VScPIMPcSN#ATVPDGEccPR	P07996
385	KVcQDcPLLAPLN#DTR	P02765
386	VcqDcPLLAPLN#DTR	P02765
387	IcDLLVANnHFAHFFAPQN#LTNMNK	P19827
388	GcVLLSYLN#ETVTVSASLESVR	P01023
389	NcGVN#cSGDVFTALIGEIASPnYPKPYPENSR	P09871

390	NcGVN#cSGDVFTALIGEIASPNYPKPYPENSR	P09871
391	FcRDN#YTDLVAIQNK	P14151
392	KcGN#cSLTTLKDEDFcK	P49908
393	IcDLLVANNHFAHFFAPQN#LTNmNK	P19827
394	IcDLLVANNHFAHFFAPQN#LTNMNK	P19827
395	GcSSSTSVLLTLDNNVVN#GSSPAIR	P07996
396	GcSSSTSVLLTLDNnVVN#GSSPAIR	P07996
397	VcQDcPLLAPLN#DTR	P02765
398	GcScFSDWQGPGcSVPVPAN#QSFWTR	O75882
399	SPYEMFGDEEVMcLNGN#WTEPPQcK	P08603
400	SPYEMFGDEEVMcLNGN#WTEPPQcK	Q03591
401	SPYEMFGDEEVMcLnGN#WTEPPqcK	P08603
402	SPYEMFGDEEVMcLnGN#WTEPPqcK	Q03591
403	SPYEMFGDEEVMcLnGN#WTEPPQcK	P08603
404	SPYEMFGDEEVMcLnGN#WTEPPQcK	Q03591
405	GLQPTLTNPGEcRPN#FTcAcR	P04275
406	FSLLGHASIScTVEN#ETIGVWRPSPPTcEK	P04003
407	LEPVHLQLQcMSQEQLAQVAAN#ATK	Q96PD5
408	LEPVHLQLQcmSQEQLAQVAAN#ATK	Q96PD5
409	cIqAN#YSLMEnGK	P05090
410	cATPHGDN#ASLEATFVK	P23142
411	cIQAN#YSLMEnGK	P05090
412	cIQAN#YSLMEnGKIK	P05090
413	cIQAN#YSLMENGK	P05090
414	cFLGN#GTGYR	Q04756
415	cGN#cSLTTLKDEDFcK	P49908
416	cGN#cSLTTLKDEDFcKR	P49908
417	cSDGWSFDATTLDDN#GTmLFFK	P02790
418	cSDGWSFDATTLDDN#GTMLFFK	P02790

References:

(S1) J. Liu, Z. K. Sun, Y. H. Deng, Y. Zou, C. Y. Li, X. H. Guo, L. Q. Xiong, Y. Gao, F. Y. Li and D. Y. Zhao, *Angew. Chem. Int. Ed.*, 2009, **48**, 5875 –5879.

(S2) J. X. Wang, J. Li, Y. N. Wang, M. X. Gao, X. M. Zhang and P. Y. Yang, *ACS Appl. Mater. Interfaces*, 2016, **8**, 27482−27489.