Electronic Supplementary Information

Facile Synthesis of Core-shell Structured Magnetic Covalent Organic Framework Composite Nanospheres for Selective Enrichment of Peptides with Simultaneous Exclusion of Proteins

Chaohong Gao,[†,§] *Guo Lin,*[†,§] *Zhixian Lei,*[†] *Qiong Zheng,*[†] *Jiashi Lin,*[‡] *and Zian Lin*[*†]

Ministry of Education Key Laboratory of Analytical Science of Food Safety and Biology, Fujian Provincial Key Laboratory of Analysis and Detection Technology for Food Safety, College of Chemistry, Fuzhou University, Fuzhou, Fujian, 350116, China.

- First corresponding author: **Zi-an Lin;**
- Second corresponding author: **Jiashi Lin**
- Postal address: **College of Chemistry, Fuzhou University, Fuzhou, Fujian, 350108, China**
- Fax: **86-591-22866165**
- E-mail: **zianlin@fzu.edu.cn (Z.A. Lin);**

The first two authors contributed equally to this paper

MS analysis and database searching.

Mass spectrometry (MS) analysis was conducted on a HPLC-Q-TOF MS containing a 1260 series HPLC system with a binary SL pump and an Agilent 6520 Q-TOF with dual electrospray ionization source (ESI). The separation of all samples were collected on an Agilent Poroshell 120 EC-C18 column (2.7 μm, 3.0 mm × 50 mm, Agilent). The flow rate was set as 0.3 mL min^{-1}. Solvent A was composed of aqueous solution containing 0.1% formic acid and solvent B consisted of ACN containing 0.1% formic acid. The gradient elution program was as follows: in the original 3 min, maintained at 3% B, then linear gradient to 40% B within 20 min, and keep rising to 80% B within 3 min and maintaining for 2 min. After a back wash step was performed with 80 % B for 3 min, the column was equilibrated with 3 % B within 4 min. The column temperature was held at 40 °C and the sample injection volume was 10 μL. The MS system was dried by Nitrogen at 350 ℃ with a flow rate of 10 L min^{-1}. Both full scan MS data and MS/MS data were obtained at m/z 300–2000 and 100–3000, respectively. Scan rates for MS and MS/MS data were tuned to 3 spectra s^{-1}. The voltage was set as 4 kV for the MS capillary and the voltage of fragmentor was set as 175 V. In the MS2 experiments, the collision energy was set up in terms of formula, where the top three highest intensity peaks in each mass spectra were chosen for collision-induced dissociation. Isolation width for MS^2 was ± 4 amu.

All the LC-MS/MS raw data were measured on a Spectrum Mill versionA.03.03 software (Agilent Technologies). The mass tolerances were 100 ppm for parent ions and 200 amu for fragment ions. Trypsin restriction was set with two missed cleavages. Carboxymethylation was set as the static modification; oxidized methionine was set as the variable modifications.

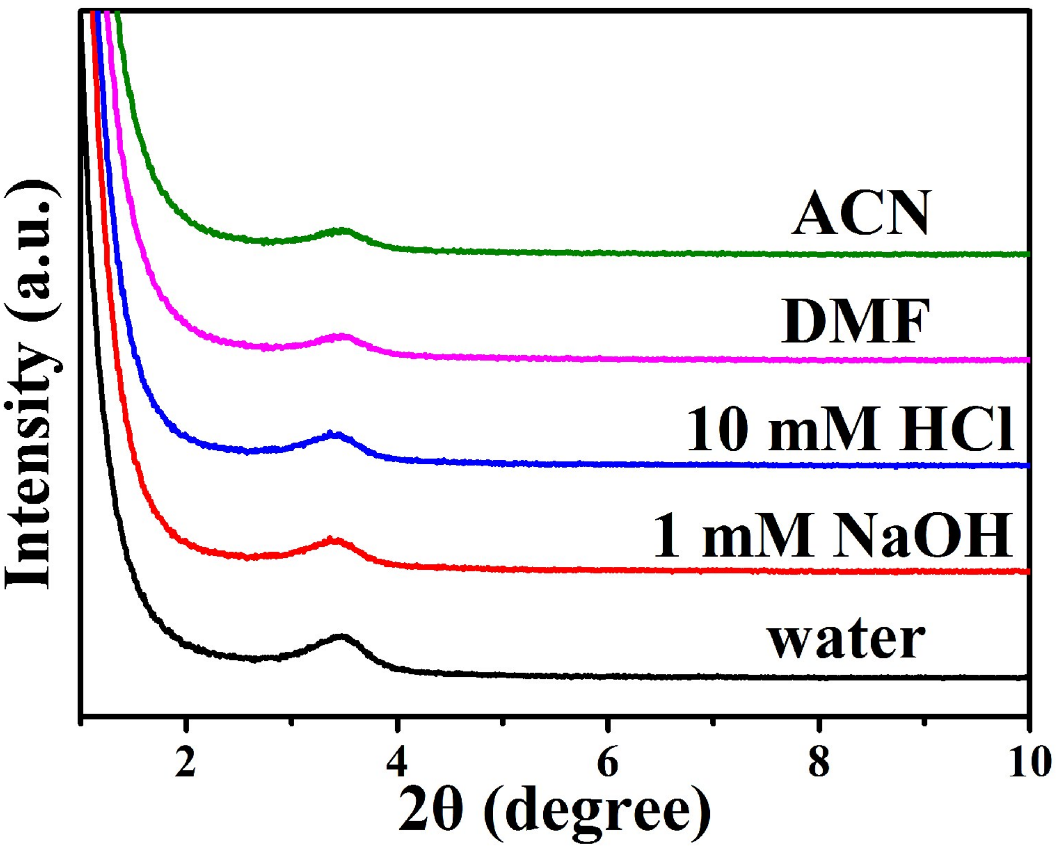

Fig. S1 XRD patterns of the Fe_3O_4@TbBd nanospheres treated with ACN, DMF, 10 mM HCl, 1mM NaOH and water for overnight.

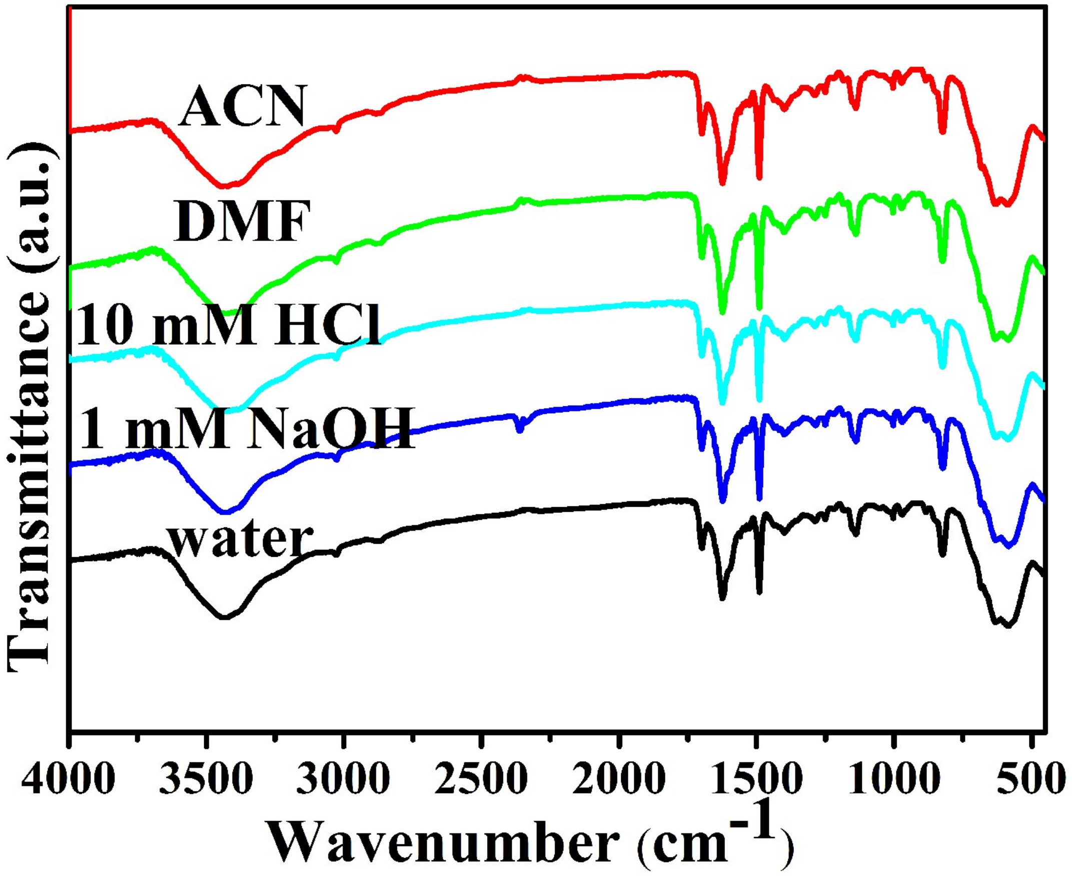

Fig. S2 FT-IR spectra of the Fe_3O_4@TbBd nanospheres treated with ACN, DMF, 10 mM HCl, 1mM NaOH and water for overnight.

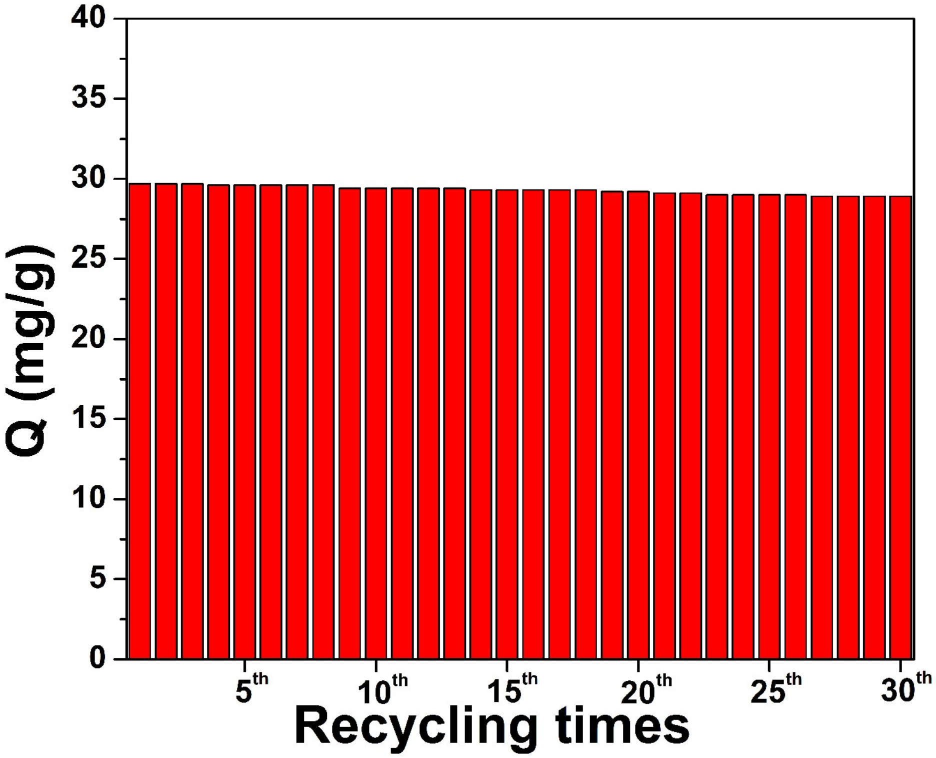

Fig. S3 Recycled use of the Fe_3O_4@TbBd nanospheres for FGFGF adsorption. Amount of Fe_3O_4@TbBd nanospheres: 0.3 mg; binding media: 300 μL eluent (50% ACN aqueous solution); incubation time: 10 min; C_{FGFGF}: 70 μg/mL.

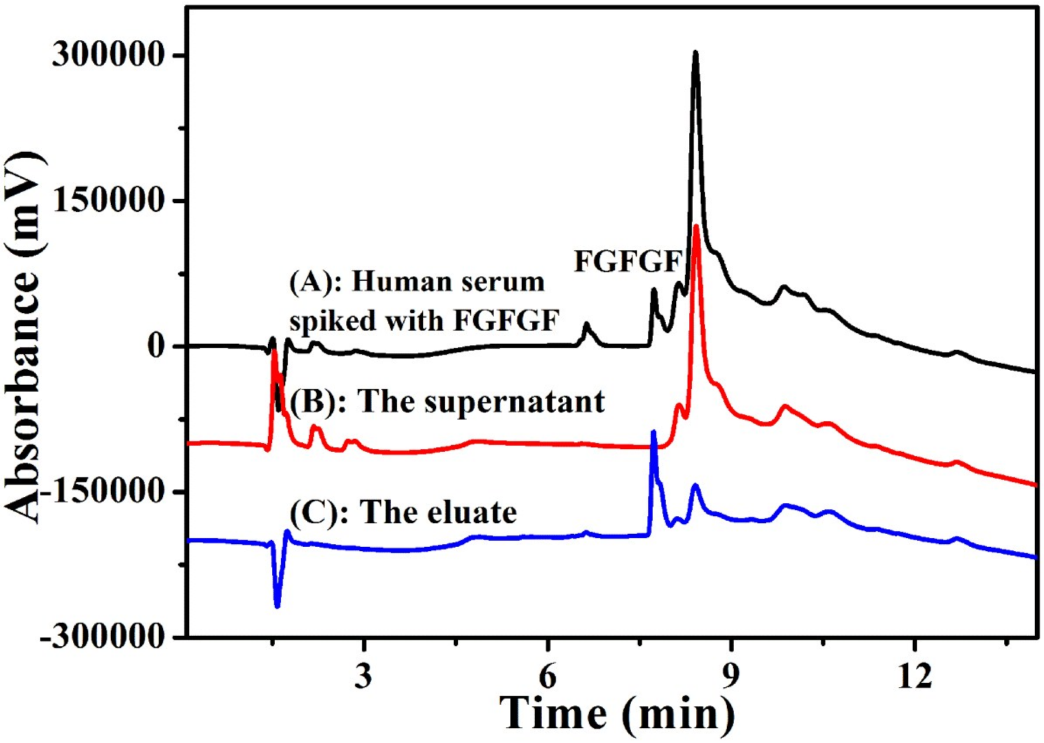

Fig. S4 Separation of 50-fold diluted human serum spiked with 25 μg/mL FGFGF before and after treatment with the Fe_3O_4@TbBd. (A) The human serum spiked with FGFGF, (B) the supernatant, and (C) the eluate. The following gradient program was used: 0-12 min, 3-90% B; 12-12.5 min, 90% B; 12.5-13 min, 90-13% B; 14 min, stop.

Table S1. Proteins in human serum were identified by HPLC-Q-TOF/MS before and after treatment with the Fe_3O_4@TbBd nanospheres.

Database Accession	Protein Name	Before enrichment			After enrichment		
		Distinct Peptides	% AA Coverage	Mean Peptide Spectral Intensity	Distinct Peptides	% AA Coverage	Mean Peptide Spectral Intensity
P02768	Serum albumin precursor	39	64	1.20e+005	20	25	1.98e+005
P01871	Ig mu chain C region	/	/	/	3	5	1.80e+004
P04220	Ig mu heavy chain disease protein	/	/	/	2	4	2.24e+004
P01857	Ig gamma-1 chain C region	5	19	7.31e+004	9	19	1.24e+005
P01859	Ig gamma-2 chain C region	4	13	3.41e+004	5	16	7.37e+004
P01861	Ig gamma-4 chain C region	3	11	4.29e+004	3	8	1.06e+005
P01860	Ig gamma-3 chain C region	3	9	4.06e+004	3	8	1.14e+004
P01834	Ig kappa chain C	3	50	5.68e+004	2	31	5.88e+004
P02647	Ig lambda chain C regions	3	11	1.57e+004	4	17	2.81e+004
P01842	Ig kappa chain C region	2	23	2.82e+004	\	\	\
P01876	Ig alpha-1 chain C region	/	/	/	2	5	3.10e+004
P01877	Ig alpha-2 chain C region	/	/	/	1	2	2.79e+004
P01023	Alpha-2-macroglobulin	/	/	/	3	2	1.03e+003

	precursor			
P20742	Pregnancy zone protein precursor	1	0	8.21e+004

Table S2. Serum peptides of human serum digest (5 ng/μL) identified by HPLC-Q-TOF/MS before and after treatment with the Fe_3O_4@TbBd nanospheres.

No.	Mw	Amino acid sequence[a]	GRAVY[b]	Before enrichment [c]	After enrichment
1	2778.359	(R)LVRPEVDVMcTAFHDNEETFLKK(Y)	-0.45	√	×
2	2650.264	(R)LVRPEVDVMcTAFHDNEETFLK(K)	-0.29	√	×
3	2585.118	(K)VHTEccHGDLLEcADDRADLAK(Y)	-0.84	√	×
4	2518.214	(R)MPcAEDYLSVVLNQLcVLHEK(T)	0.17	√	×
5	2490.285	(K)ALVLIAFAQYLQQcPFEDHVK(L)	0.39	√	×
6	2260.023	(K)EFNAETFTFHADIcTLSEK(E)	-0.4	√	×
7	2139.027	(R)TPEVTcVVVDVSHEDPEVK(F)	-0.27	√	×
8	2135.969	(K)VDNALQSGNSQESVTEQDSK(D)	-1.29	√	×
9	2086.838	(K)VHTEccHGDLLEcADDR(A)	-1.14	√	×
10	2045.095	(K)VFDEFKPLVEEPQNLIK(Q)	-0.35	√	√
11	1932.037	(K)SLHTLFGDKLcTVATLR(E)	0.35	√	√
12	1910.932	(R)RPcFSALEVDETYVPK(E)	-0.54	√	×
13	1875.927	(K)VYAcEVTHQGLSSPVTK(S)	-0.14	√	√
14	1797.895	(K)SGTASVVcLLNNFYPR(E)	0.23	√	√
15	1793.991	(R)VVSVLTVVHQDWLNGK(E)	0.51	×	√
16	1711.759	(R)SYScQVTHEGSTVEK(T)	-1.05	√	×
17	1677.802	(K)FNWYVDGVEVHNAK(T)	-0.46	×	√
18	1657.753	(K)QNcELFEQLGEYK(F)	-1.35	√	×
19	1639.938	(K)KVPQVSTPTLVEVSR(N)	-0.07	√	√
20	1612.785	(K)LLDNWDSVTSTFSK(L)	-0.32	×	√
21	1552.598	(K)ccAAADPHEcYAK(V)	-0.98	√	×
22	1546.797	(K)LKEccEKPLLEK(S)	-1.24	√	×

23	1511.843	(K)VPQVSTPTLVEVSR(N)	0.21	√	√
24	1498.578	(K)TcVADESAENcDK(S)	-1.37	√	×
25	1443.642	(K)YIcENQDSISSK(L)	-1.15	√	×
26	1434.533	(R)ETYGEMADccAK(Q)	-1.13	√	×
27	1423.71	(R)STSESTAALGcLVK(D)	0.29	√	×
28	1416.828	(K)MVSGFIPLKPTVK(M)	0.65	×	√
29	1386.715	(K)VSFLSALEEYTK(K)	0.16	√	√
30	1371.567	(K)AAFTEccQAADK(A)	-0.51	√	×
31	1342.635	(K)AVMDDFAAFVEK(C)	0.58	√	√
32	1321.678	(K)STSGGTAALGcLVK(D)	0.56	√	√
33	1287.651	(K)GPSVFPLAPcSR(S)	0.12	√	√
34	1272.673	(R)VTAAPQSVcALR(A)	0.59	×	√
35	1255.643	(K)AIGYLNTGYQR(Q)	-0.5	×	√
36	1249.636	(K)LIcQATGFSPR(Q)	0.14	×	√
37	1235.688	(K)DLATVYVDVLK(D)	0.83	√	×
38	1230.709	(R)QGLLPVLESFK(V)	0.43	×	√
39	1226.605	(R)FKDLGEENFK(A)	-1.28	√	×
40	1213.632	(R)WLQGSQELPR(E)	-1.11	×	√
41	1186.647	(K)GPSVFPLAPSSK(S)	0.09	√	√
42	1186.498	(K)ScDTPPPcPR(C)	-1.99	√	×
43	1161.63	(K)NQVSLTcLVK(G)	0.4	√	√
44	1149.615	(K)LVNEVTEFAK(T)	0.17	√	√
45	1141.694	(K)KLVAASQAALGL(-)	1.18	×	√
46	1138.498	(K)ccTESLVNR(R)	-0.71	√	×
47	1128.699	(K)KQTALVELVK(H)	0.23	√	√
48	1074.543	(K)LDELRDEGK(A)	-1.69	√	×
49	1031.519	(K)LSPLGEEMR(D)	-0.53	√	√
50	1017.536	(K)SLHTLFGDK(L)	-0.23	√	×

51	1013.599	(K)LVAASQAALGL(-)	1.65	√	√
52	1000.604	(K)QTALVELVK(H)	0.69	√	√
53	990.51	(K)AGVETTTPSK(Q)	-0.63	√	×
54	984.488	(K)TYETTLEK(C)	-1.31	√	√
55	960.563	(K)FQNALLVR(Y)	0.61	√	√
56	951.442	(K)DLGEENFK(A)	-1.46	√	×
57	940.448	(K)DDNPNLPR(L)	-2.24	√	×
58	933.519	(K)LcTVATLR(E)	1.1	√	√
59	931.546	(K)TPLTATLSK(S)	0.11	×	√
60	927.493	(K)YLYEIAR(R)	-0.07	√	√
61	900.53	(K)VSVFVPPR(D)	0.86	×	√
62	880.441	(K)AEFAEVSK(L)	-0.14	√	√
63	838.503	(K)ALPAPIEK(T)	0.16	√	√
64	835.434	(K)DTLMISR(T)	0.1	×	√
65	824.488	(K)GLPAPIEK(T)	-0.11	√	√
66	789.472	(K)LVTDLTK(V)	0.43	√	√
67	775.446	(R)GFPSVLR(G)	0.5	×	√
68	772.439	(K)AAcLLPK(L)	0.95	√	√
69	706.355	(K)cASLQK(F)	-0.52	√	√
70	698.358	(K)SEVAHR(F)	-1	√	×
71	674.347	(K)TPVSDR(V)	-1.15	×	√
72	673.378	(K)AWAVAR(L)	0.7	×	√

[a]The red mark in database sequence represents hydrophobic group-containing amino acids;
[b] The red mark in database sequence represents the observed hydrophobic peptides.
[c] The symbol of"×"and "√" represent with observing of peptides and without observing of peptides, respectively.

Table S3. Serum peptides of human serum digest (0.5 ng/μL) identified by HPLC-Q-TOF/MS before and after treatment with the Fe_3O_4@TbBd nanospheres.

No.	Mw	M/Z	Amino acid sequence[a]	GRAVY[b]	Before enrichment[c]	After enrichment
1	1910.932	673.98	(R)RPcFSALEVDETYVPK(E)	-0.54	√	×
2	1342.635	672.32	(K)AVMDDFAAFVEK(C)	0.575	√	×
3	1149.615	575.81	(K)LVNEVTEFAK(T)	0.17	√	√
4	1138.498	570.25	(K)ccTESLVNR(R)	-0.71	√	×
5	1013.599	507.80	(K)LVAASQAALGL(-)	1.65	×	√
6	1000.604	501.30	(K)QTALVELVK(H)	0.69	√	√
7	960.563	481.28	(K)FQNALLVR(Y)	0.61	√	√
8	927.493	464.75	(K)YLYEIAR(R)	-0.071	√	√
9	880.441	441.22	(K)AEFAEVSK(L)	-0.14	√	×
10	789.472	395.74	(K)LVTDLTK(V)	0.43	√	√
11	772.439	387.22	(K)AAcLLPK(L)	0.95	×	√

12	706.355	354.18	(K)cASLQK(F)	-0.52	√	×
Number of peptides					**10**	**7**
Number of hydrophobic peptides					**5**	**6**

[a] The red mark in database sequence represents hydrophobic group-containing amino acids;
[b] The red mark in database sequence represents the observed hydrophobic peptides.
[c] The symbol of“×”and “√” represent with observing of peptides and without observing of peptides, respectively.

Table S4. List of human serum digest (0.5 ng/μL) identified by HPLC-Q-TOF/MS before and after treatment with the Fe_3O_4@TbBd nanospheres.

	S/N (signal-to-noise)	
m/z	**Before enrichment**	**After enrichment**
672.32	14	
637.47	7	
575.81	7	51
570.25	3	
507.80		15
501.30	10	28
481.28	16	33
464.75	8	25
441.22	7	
395.74	11	30
387.22		14
354.18	9	
Number of peptides	10	7
Number of hydrophobic peptides	5	6

[a] The red mark in the value of m/z represents the observed hydrophobic peptides.