Electronic Supplementary Information (ESI)

Metal-Organic Frameworks for Efficient Enrichment of Peptides with Simultaneous Exclusion of Proteins from Complex Biological Samples

Zhi-Yuan Gu, Ying-Jun Chen, Jun-Qing Jiang, Xiu-Ping Yan*

Research Center for Analytical Sciences, and Key Laboratory of Functional Polymer Materials, Ministry of Education, College of Chemistry, Nankai University, Tianjin 300071, China

*To whom correspondence should be addressed. E-mail: xpyan@nankai.edu.cn.

1. Chemicals and Materials

All chemicals were at least of analytical grade and used without further treatment. $Al(NO_3)_3 \cdot 9H_2O$, $Cr(NO_3)_3 \cdot 9H_2O$, terephthalic acid, CrO_3, trimesic acid, and ethanol were from Sinopharm Chemical Reagent Co. Ltd. (Shanghai, China). Hydrofluoric acid (40.0%) and *N, N*-dimethylformamide (DMF) were from Tianjin Standard Science and Technology Co., Ltd.. α-cyano-4-hydroxycinnamic acid (CHCA) and tryptic digest of bovine serum albumin (BSA) were purchased from Bruker Daltonics (Bremen, Germany). Ultrapure water (18.2MΩ cm) was obtained from a WaterPro Water Purification System (Labconco Corporation, Kansas City, MO). Human plasma sample was obtained from volunteers, provided by a local Hospital.

2. Synthesis, Activation and Characterization of MIL-53, MIL-100 and MIL-101

MIL-53 was synthesized according to Loiseau et al.[S1] Typically, 1300 mg of $Al(NO_3)_3 \cdot 9H_2O$ and 288 mg of terephthalic acid were mixed with 5 mL of ultrapure water. The obtained mixture was transferred to a Teflon-lined bomb. Then, the Teflon-lined bomb was sealed, placed in an oven, and left at 220 °C for 3 days. The white crystalline solid was thus obtained. After washing with water, the solid was purified and activated upon heating in air at 330 ^oC for 3 days. According to Loiseau et al.[S1], the unreacted terephthalic acid species and the terephthalic acid molecules contained in the pores of MIL-53 were expelled after the activation procedure.

MIL-100 was synthesized according to Latroche et al.[S2] Typically, 500 mg of CrO_3, 1050 mg of trimesic acid and 1.0 mL of HF (5M) were mixed with 24 mL of ultrapure water. The obtained mixture was stirred a few minutes and transferred to a Teflon-lined bomb. Then, the Teflon-lined bomb was sealed, placed in an oven, and left at 220 °C for 4 days. The green crystalline solid was thus obtained. After thorough washing with DMF, the solid was emerged in ethanol in 1 h, and collected by centrifugation at 8000 rpm for 5 min. The procedure was repeated 3 times to eliminate the unreacted terephthalic acid from MIL-100. The solid was obtained by centrifugation at 8000 rpm for 5 min and then evacuated in vacuum under 150 ^oC for 12 h.

MIL-101 was synthesized according to Férey et al.[S3] Typically, $Cr(NO_3)_3 \cdot 9H_2O$ (800 mg, 2.0 mmol), terephthalic acid (332 mg, 2.0 mmol) and HF (0.4 mL, 2.0 mmol) were mixed with ultrapure water (9.5 mL). The obtained mixture was transferred to a Teflon-lined bomb. Then, the Teflon-lined bomb was sealed, placed in an oven, and left at 200 °C for 8 h. The green crystalline solid was thus obtained. After thorough washing with DMF, the solid was emerged in ethanol in 1 h, and collected by centrifugation at 8000 rpm for 5 min. The procedure was repeated 3 times to eliminate the unreacted terephthalic acid from MIL-101. The solid was obtained by centrifugation at 8000 rpm for 5 min and then evacuated in vacuum under 150 ^oC for 12 h.

XRD, SEM and N_2 adsorption were employed to characterize the prepared MIL-53, MIL-100 and MIL-101. The XRD patterns were recorded with a D/max-2500 diffractometer (Rigaku, Japan) using $Cu_{K\alpha}$ radiation (λ=1.5418 Å). The XRD patterns of synthesized MIL-53, MIL-100 and MIL-101 were in accordance with the simulated ones (Figure S2). BET surface area was measured on a TriStar 3000 sorptometer (Micromeritics, Norcross, GA) using nitrogen adsorption at 77 K in the range $0.02 \leq P/P_0 \leq 0.20$, respectively. The determined BET surface area of MIL-53, MIL-100 and MIL-101 was 1038, 1595 and 2907 $m^2\ g^{-1}$, respectively. The scanning electron microscopy (SEM) images were recorded on a Shimadzu SS-550 scanning electron microscope at 15.0 kV.

3. Enrichment of BSA Tryptic Digest and MALDI-TOF MS Analysis

Specific MOF (MIL-53, MIL-101 or MIL-100) suspension (100 μL, 10 mg mL^{-1}) was added into 1 mL of BSA tryptic digest solution (2 fmol μL^{-1}) and vortexed at 25 ^oC for 1 h. The slurry was centrifuged at 12000 rpm for 3 min. After the supernatant was removed, the peptides adsorbed on the MOF particles were eluted by washing with 50% acetonitrile aqueous solution containing 0.1% TFA (10 μL). After centrifugation at 12000 rpm for 3 min, 1 μL of the supernatant was then mixed with 1 μL of pre-made α-cyano-4-hydroxycinnamic acid (CHCA) solution (10 mg mL^{-1} in 50% acetonitrile aqueous solution containing 0.1% TFA) on the MALDI plate and applied for MALDI-TOF MS analysis.

Peptides eluted from MOFs after the enrichment were analyzed with an Autoflex III MALDI-TOF mass spectrometer (Bruker, Bremen, Germany) equipped with a 337-nm nitrogen laser, an ion source, delayed-extraction electronics, a high-resolution timed ion selector, and a 2-GHz digitizer. One microliter of peptide sample was typically mixed with equal volume of pre-made CHCA matrix solution, deposited onto the MTP 384 stainless steel target and allowed to dry at room temperature.

The standard protocol was as follows. Each spectrum was the result of 800 laser shots, per *m*/*z* segment per sample, delivered in eight sets of 100 shots (at frequency of 100 Hz and the laser energy level of 30%) to each of eight different locations on the surface of the matrix spot. Spectra were acquired in reflection mode under 19 kV of ion accelerating and with deflection of mass ions of 900 *m*/*z*. The data was controlled by preparing three parallel samples and depositing each sample three spots onto the targets. Spectra were accepted if the sample had at least 2 of the 3 spotting replicates summing up 800 shots in at least 2 of the 3 run replicates per sample and further confirmed by visual inspection. The entire acquiring program was automated using the AutoXecute function of instrument.

Online search for the obtained peptides were using the following parameters: database, SwissProt; digest used, Trypsin; maximum of missed cleavages, 1; mass tolerance, ± 90 ppm. Mascot from Matrix Science Ltd. (London, U.K.) was used to search all of the mass spectra.

4. Selective Extraction of Peptides from Human Plasma and MALDI-TOF MS Analysis

Caution: *For all blood and blood-derived samples, always observe precautions: Handle all biological samples as a potential source of pathogens/hazards, use the appropriate protective attire (lab coats, safety glasses etc.) and dispose of all (bio)hazardous materials in an appropriate manner.*

Specific MOF (MIL-53, MIL-101, or MIL-100) suspension (250 μL, 10 mg mL^{-1}) was added into diluted human plasma (250 μL plasma in 750 μL ultrapure water) and vortexed at 25 ^{o}C for 1 h. The slurry was centrifuged at 12000 rpm for 3 min. After the supernatant was removed, the resulting MOF particles were washed with ultrapure water for three times. The peptides adsorbed in the MOF particles were eluted by washing with 50% acetonitrile aqueous solution (20 μL, 1:1). After centrifugation at 12000 rpm for 3 min, 1 μL of the supernatant was then mixed with 1 μL of pre-made CHCA solution on the MALDI plate and applied for MALDI-TOF MS analysis. To compare with the direct analysis of human plasma and to obtain clear mass spectra, human plasma was diluted for direct analysis (100 μL plasma in 4900 μL ultrapure water). 1 μL of the diluted human plasma sample was then mixed with 1 μL of pre-made CHCA solution on the MALDI plate and applied for MALDI-TOF analysis.

The standard protocol was as follows. Each spectrum was the result of 800 laser shots, per *m*/*z* segment per sample, delivered in eight sets of 100 shots (at a frequency of 25 Hz and the laser energy level of 40%) to each of eight different locations on the surface of the matrix spot. Separate spectra were obtained for two restricted mass-to-charge (*m*/*z*) ranges, corresponding to peptides with molecular weights of 1-10 kDa and proteins with molecular weights of 10-90 kDa, under specifically optimized instrument settings. Spectra were acquired in linear mode under 20 kV of ion accelerating and with deflection of mass ions of 700 *m*/*z* (1 to 10 kDa segment) or 10000 *m*/*z* (10 to 90 kDa segment). Delayed-extraction was maintained for 80 (1 to 10 kDa) or 330 ns (10 to 90 kDa) to give appropriate time lag focusing after each laser shot. The data was controlled by three parallel samples and depositing each sample three spots onto the targets. Spectra were accepted if the sample had at least 2 of the 3 spotting replicates summing up 800 shots in at least 2 of the 3 run replicates per sample and further confirmed by visual inspection. The entire acquiring program was automated using the AutoXecute function of the instrument.

5. Selective Extraction of Peptides from Human Urine and MALDI-TOF MS Analysis

MOF (MIL-53, MIL-101, or MIL-100) suspension (250 μL, 10 mg mL^{-1}) was added into 1 mL of urine sample and vortexed at 25 ^{o}C for 1 h. The slurry was centrifuged at 12000 rpm for 3 min. After the supernatant was removed, the resulting MOF particles were washed with ultrapure water three times. The peptides adsorbed in MOF particles were eluted by washing with 50% acetonitrile aqueous solution (20 μL, 1:1). After the centrifugation of 12000 rpm for 3 min, 1 μL of the supernatant was then mixed with 1 μL of pre-made CHCA solution on the MALDI plate and applied for

MALDI-TOF MS analysis. In order to compare with the direct analysis of human urine sample, 1 μL of the human urine sample was then mixed with 1 μL of pre-made CHCA solution on the MALDI plate and applied for MALDI-TOF analysis.

The standard protocol was as follows. Each spectrum was the result of 800 laser shots, per *m*/*z* segment per sample, delivered in eight sets of 100 shots (at frequency of 25 Hz and the laser energy level of 40%) to each of eight different locations on the surface of the matrix spot. Separate spectra were obtained for two restricted mass-to-charge (*m*/*z*) ranges, corresponding to peptides with molecular mass of 1-10 kDa and proteins with molecular mass of 10-90 kDa, under specifically optimized instrument settings. Spectra were acquired in linear mode under 20 kV of ion accelerating and with deflection of mass ions of 700 *m*/*z* (1 to 10 kDa segment) or 10000 *m*/*z* (10 to 90 kDa segment). Delayed-extraction was maintained for 80 (1 to 10 kDa) or 330 ns (10 to 90 kDa) to give appropriate time lag focusing after each laser shot. The data was controlled by three parallel samples and depositing each sample three spots onto the targets. Spectra were accepted if the sample had at least 2 of the 3 spotting replicates summing up 800 shots in at least 2 of the 3 run replicates per sample and further confirmed by visual inspection. The entire acquiring program was automated using the instrument's AutoXecute function.

Efficient enrichment and selective extraction of peptides (1-10 kDa) on MIL-53, MIL-100 and MIL-101 was observed (Figure S9). No protein peaks on the MW range of 10-90 kDa were observed for both direct analysis of urine and after the enrichment with MIL-53, MIL-100, and MIL-101 (The mass spectra were not shown).

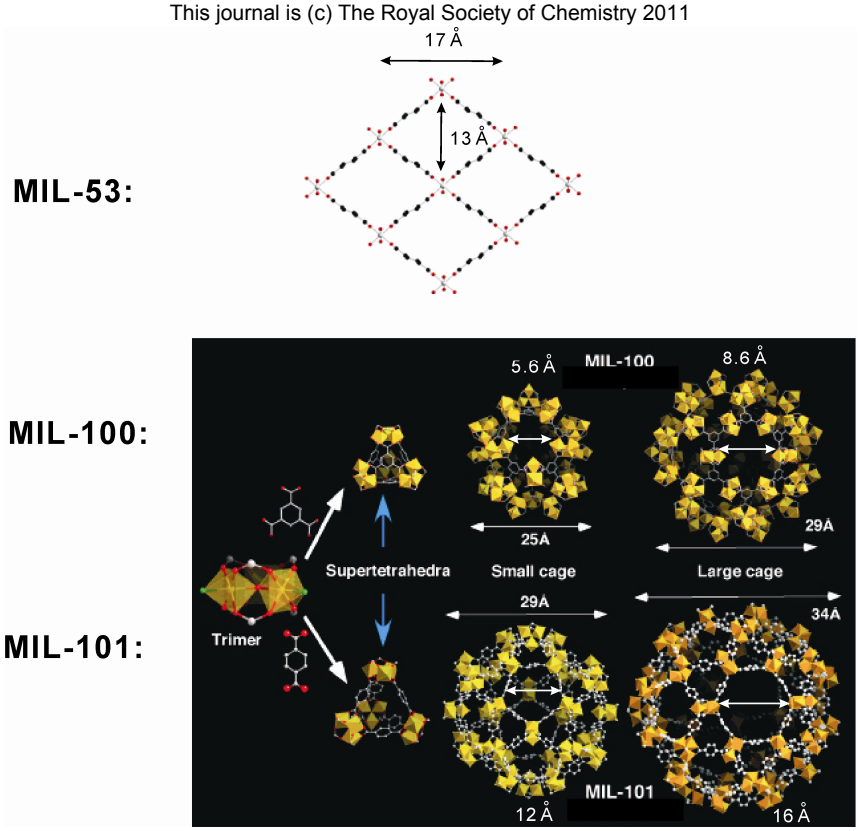

Figure S1. Structures of MIL-53, MIL-100 and MIL-101. MIL-53 has open frameworks with pore and window sizes of 1.7 nm × 1.3 nm.[S4] For MIL-100 and MIL-101, chromium trimers and the carboxylate linkers trimesate and terephthalate are combined to form supertetrahedra. Both MIL-100 and MIL-101 contain two-types of nanometer-sized cages and accessible windows (See also Table S1).[S5] This figure is a combination of Figure 1 in Ref. S5 and Figure 1 in Ref. S4. Copyright 2005 [S4] and 2008 [S5] American Chemical Society and with permission.

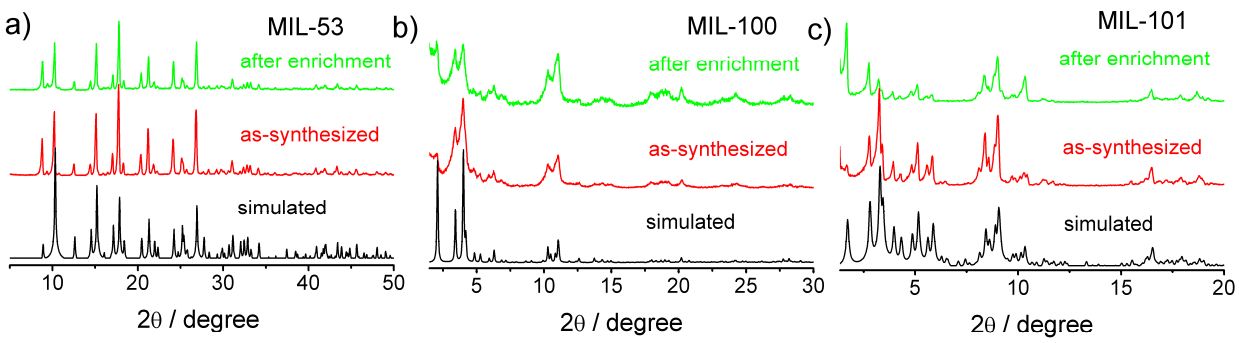

Figure S2. XRD patterns of (a) MIL-53, (b) MIL-100, and (c) MIL-101: simulated (black), as-synthesized (red), and after-enrichment (green).

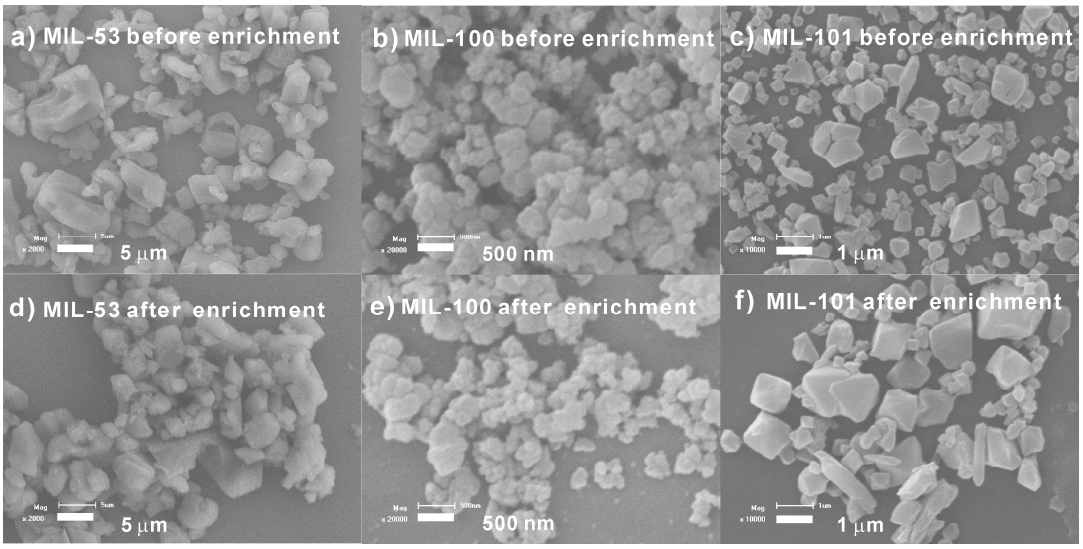

Figure S3. SEM images of the prepared MIL-53 (a,d), MIL-100 (b.e) and MIL-101 (c,f) before enrichment (a,b,c) and after enrichment (d,e,f) of the peptides from human plasma samples.

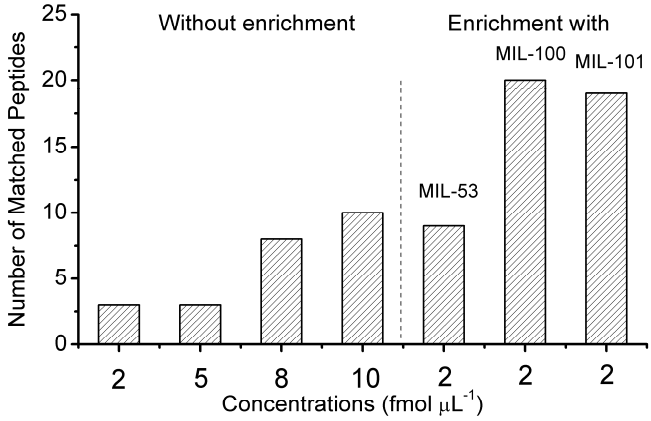

Figure S4. Number of matched peptides obtained by BSA tryptic digest before enrichment at concentrations of 2, 5, 8, and 10 fmol μL^{-1} and after enrichment with MIL-53, MIL-100 and MIL-101 at concentration of 2 fmol μL^{-1}.

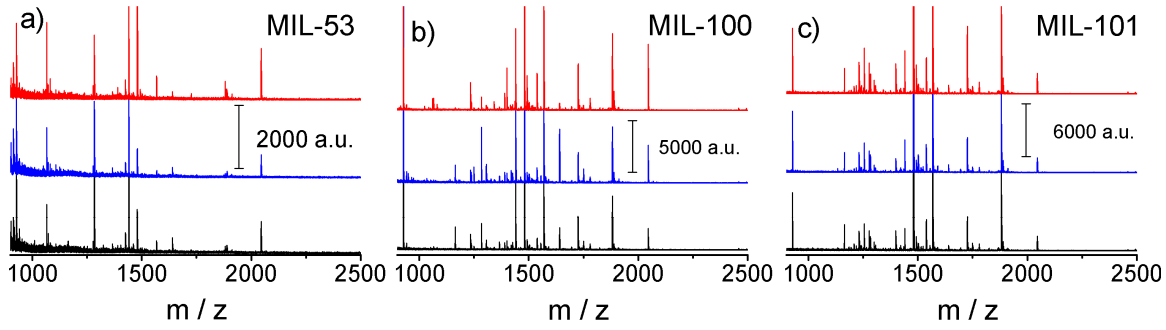

Figure S5. Reproducible mass spectra of the BSA tryptic digest (2 fmol μL^{-1}) after enrichment with MIL-53 (a), MIL-100 (b), and MIL-101 (c).

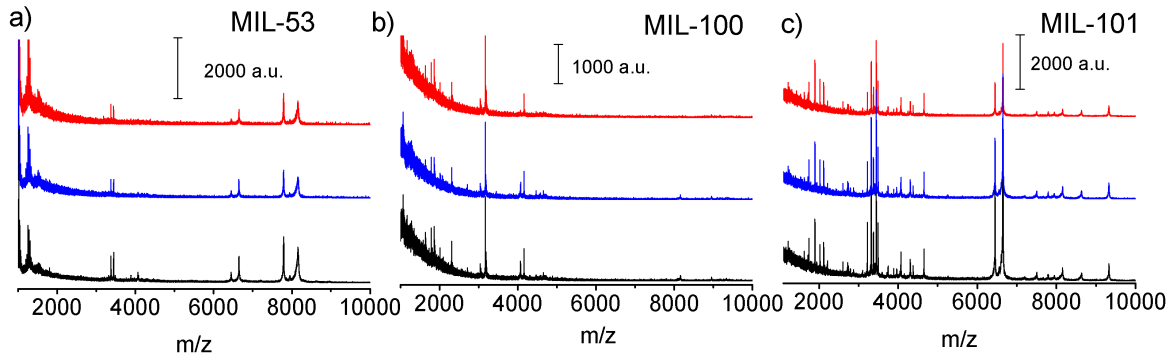

Figure S6. Reproducible mass spectra of the human plasma after enrichment with MIL-53 (a), MIL-100 (b), and MIL-101 (c).

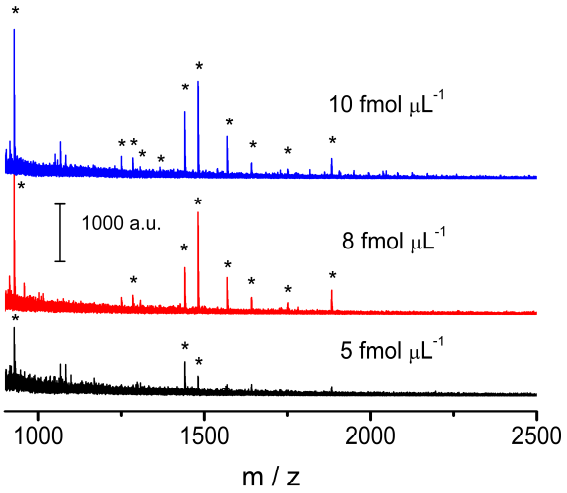

Figure S7. Mass spectra of BSA tryptic digest (5, 8, and 10 fmol μL^{-1}) obtained by MALDI-TOF-MS without enrichment. Asterisks indicate the matched peaks.

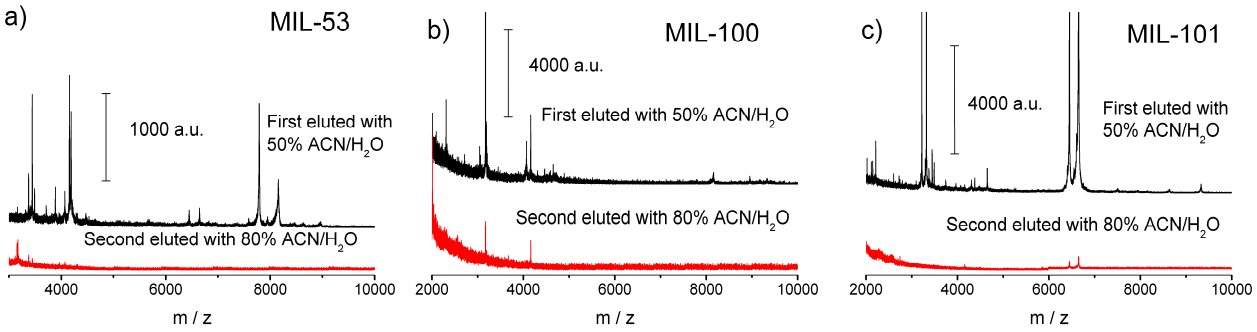

Figure S8. To show the elution efficiency, second elution step with 80% ACN/H_2O (red line) was performed after elution with 50% ACN/H_2O (black line) for human plasma sample using MIL-53 (a), MIL-100 (b), and MIL-101 (c).

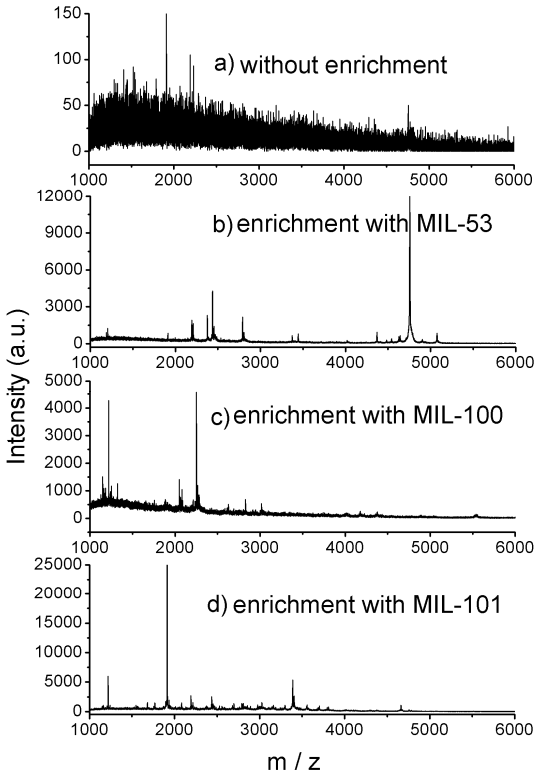

Figure S9. Mass spectra of human urine sample obtained by MALDI-TOF MS before enrichment (a) and after enrichment with MIL-53 (b), MIL-100 (c) and MIL-101 (d).

Table S1. Characteristics of MIL-53, MIL-100 and MIL-101.

MOFs	BET surface area ($m^2 g^{-1}$)[a]	Pore size (nm)[b]	Pore window size (nm)[b]	Particle size[c]
MIL-53	1038	1.7 × 1.3	1.7 × 1.3	2 ± 1 μm
MIL-100	1595	2.5 2.9	0.56 0.86	200 ± 100 nm
MIL-101	2907	2.9 3.4	1.2 1.6	400 ± 200 nm

[a] Measured from N_2 adsorption; [b] Data from reference S1-S3 (see also Figure S1); [c] Measured from SEM graphs (Figure S3).

Table S2. The results for BSA digested by trypsin with and without MOF enrichment.[a]

No.	Calculated m/z	Sequence	Signal-to-noise ratio			
			No enrichment	MIL-53	MIL-100	MIL-101
1	926.4861	**K.YLYEIAR.R**	3.0	27.0	242.0	86.4
2	1162.6234	**K.LVNELTEFAK.T**	-	-	35.8	29.4
3	1248.6139	**R.FKDLGEEHFK.G**	-	-	5.8	4.3
4	1282.7034	**R.HPEYAVSVLLR.L**	-	35.3	46.7	27.1
5	1304.7088	**K.HLVDEPQNLIK.Q**	-	-	14.5	7.1
6	1398.6853	**K.TVMENFVAFVDK.C**	-	-	14.8	23.2
7	1418.6864	**K.SLHTLFGDELCK.V**	-	-	8.6	4.7
8	1438.8045	**R.RHPEYAVSVLLR.L**	4.1	132.9	185.3	57.1
9	1478.7881	**K.LGEYGFQNALIVR.Y**	6.3	102.9	339.7	613.9
10	1553.6457	**K.DDPHACYSTVFDK.L**	-	-	13.1	5.6
11	1566.7354	**K.DAFLGSFLYEYSR.R**	-	15.2	292.2	465.1
12	1575.7603	**K.LKPDPNTLCDEFK.A**	-	-	6.9	3.8
13	1638.9305	**R.KVPQVSTPTLVEVSR.S**	-	6.3	42.9	13.6
14	1723.8273	**R.MPCTEDYLSLILNR.L**	-	-	86.7	75.2
15	1748.6553	**K.ECCHGDLLECADDR.A**	-	-	16.7	10.8
16	1879.9138	**R.RPCFSALTPDETYVPK.A**	-	12.6	137.5	177.1
17	1887.9195	**R.HPYFYAPELLYYANK.Y**	-	8.2	14.3	24.1
18	1906.9135	**K.LFTFHADICTLPDTEK.Q**	-	-	3.7	4.0
19	2044.0206	**R.RHPYFYAPELLYYANK.Y**	-	20.9	66.2	41.5
20	2491.2364	**K.GLVLIAFSQYLQQCPFDEHVK.L**	-	-	4.0	-
Peptides matched			3	9	20	19
Searched peptides			-	21	42	53
Protein score			-	87	224	161
Sequence coverage			-	15%	39%	35%

[a] Search parameters: database, SwissProt; digest used, Trypsin; maximum of missed cleavages, 1; mass tolerance, ±90 ppm. Mascot from Matrix Science Ltd. (London, U.K.) was used to search all of the mass spectra. See the following matched sequences for details. The protein scores greater than 70 obtained using these parameters are significant.

Appendix I. Matched sequence for the enriched peptides from BSA tryptic digest on MIL-53

Matched peptides shown in **Bold Red**

```
  1 MKWVTFISLL LLFSSAYSRG VFRRDTHKSE IAHRFKDLGE EHFKGLVLIA
 51 FSQYLQQCPF DEHVKLVNEL TEFAKTCVAD ESHAGCEKSL HTLFGDELCK
101 VASLRETYGD MADCCEKQEP ERNECFLSHK DDSPDLPKLK PDPNTLCDEF
151 KADEKKFWGK YLYEIARRHP YFYAPELLYY ANKYNGVFQE CCQAEDKGAC
201 LLPKIETMRE KVLTSSARQR LRCASIQKFG ERALKAWSVA RLSQKFPKAE
251 FVEVTKLVTD LTKVHKECCH GDLLECADDR ADLAKYICDN QDTISSKLKE
301 CCDKPLLEKS HCIAEVEKDA IPENLPPLTA DFAEDKDVCK NYQEAKDAFL
351 GSFLYEYSRR HPEYAVSVLL RLAKEYEATL EECCAKDDPH ACYSTVFDKL
401 KHLVDEPQNL IKQNCDQFEK LGEYGFQNAL IVRYTRKVPQ VSTPTLVEVS
451 RSLGKVGTRC CTKPESERMP CTEDYLSLIL NRLCVLHEKT PVSEKVTKCC
501 TESLVNRRPC FSALTPDETY VPKAFDEKLF TFHADICTLP DTEKQIKKQT
551 ALVELLKHKP KATEEQLKTV MENFVAFVDK CCAADDKEAC FAVEGPKLVV
601 STQTALA
```

Appendix II. Matched sequence for the enriched peptides from BSA tryptic digest on MIL-100

Matched peptides shown in **Bold Red**

```
  1 MKWVTFISLL LLFSSAYSRG VFRRDTHKSE IAHRFKDLGE EHFKGLVLIA
 51 FSQYLQQCPF DEHVKLVNEL TEFAKTCVAD ESHAGCEKSL HTLFGDELCK
101 VASLRETYGD MADCCEKQEP ERNECFLSHK DDSPDLPKLK PDPNTLCDEF
151 KADEKKFWGK YLYEIARRHP YFYAPELLYY ANKYNGVFQE CCQAEDKGAC
201 LLPKIETMRE KVLASSARQR LRCASIQKFG ERALKAWSVA RLSQKFPKAE
251 FVEVTKLVTD LTKVHKECCH GDLLECADDR ADLAKYICDN QDTISSKLKE
301 CCDKPLLEKS HCIAEVEKDA IPENLPPLTA DFAEDKDVCK NYQEAKDAFL
351 GSFLYEYSRR HPEYAVSVLL RLAKEYEATL EECCAKDDPH ACYSTVFDKL
401 KHLVDEPQNL IKQNCDQFEK LGEYGFQNAL IVRYTRKVPQ VSTPTLVEVS
451 RSLGKVGTRC CTKPESERMP CTEDYLSLIL NRLCVLHEKT PVSEKVTKCC
501 TESLVNRRPC FSALTPDETY VPKAFDEKLF TFHADICTLP DTEKQIKKQT
551 ALVELLKHKP KATEEQLKTV MENFVAFVDK CCAADDKEAC FAVEGPKLVV
601 STQTALA
```

Appendix III. Matched sequence for the enriched peptides from BSA tryptic digest on MIL-101

Matched peptides shown in **Bold Red**

```
  1 MKWVTFISLL LLFSSAYSRG VFRRDTHKSE IAHRFKDLGE EHFKGLVLIA
 51 FSQYLQQCPF DEHVKLVNEL TEFAKTCVAD ESHAGCEKSL HTLFGDELCK
101 VASLRETYGD MADCCEKQEP ERNECFLSHK DDSPDLPKLK PDPNTLCDEF
151 KADEKKFWGK YLYEIARRHP YFYAPELLYY ANKYNGVFQE CCQAEDKGAC
201 LLPKIETMRE KVLTSSARQR LRCASIQKFG ERALKAWSVA RLSQKFPKAE
251 FVEVTKLVTD LTKVHKECCH GDLLECADDR ADLAKYICDN QDTISSKLKE
301 CCDKPLLEKS HCIAEVEKDA IPENLPPLTA DFAEDKDVCK NYQEAKDAFL
351 GSFLYEYSRR HPEYAVSVLL RLAKEYEATL EECCAKDDPH ACYSTVFDKL
401 KHLVDEPQNL IKQNCDQFEK LGEYGFQNAL IVRYTRKVPQ VSTPTLVEVS
451 RSLGKVGTRC CTKPESERMP CTEDYLSLIL NRLCVLHEKT PVSEKVTKCC
501 TESLVNRRPC FSALTPDETY VPKAFDEKLF TFHADICTLP DTEKQIKKQT
551 ALVELLKHKP KATEEQLKTV MENFVAFVDK CCAADDKEAC FAVEGPKLVV
601 STQTALA
```

References:

(S1) Loiseau, T.; Serre, C.; Huguenard, C.; Fink, G.; Taulelle, F.; Henry, M.; Bataille, T.; Férey, G. *Chem. Eur. J.* **2004**, *10*, 1373-1382.

(S2) Latroche, M.; Surblé, S.; Serre, C.; Mellot-Draznieks, C.; Llewellyn, P. L.; Lee, J.-H.; Chang, J.-S.; Jhung, S. H.; Férey, G. *Angew. Chem. Int. Ed.* **2006**, *45*, 8227-8231.

(S3) Férey, G.; Mellot-Draznieks, C.; Serre, C.; Millange, F.; Dutour, J.; Surble, S.; Margiolaki, I. *Science* **2005**, *309*, 2040-2042.

(S4) Bourrelly, S.; Llewellyn, P. L.; Serre, C.; Millange, F.; Loiseau, T.; Férey, G. *J. Am. Chem. Soc.* **2005**, *127*, 13519-13521.
(S5) Llewellyn, P. L.; Bourrelly, S.; Serre, C.; Vimont, A.; Daturi, M.; Hamon, L.; De Weireld, G.; Chang, J.-S.; Hong, D.-Y.; Kyu Hwang, Y.; Hwa Jhung, S.; Férey, G. *Langmuir* **2008**, *24*, 7245-7250.