Electronic Supplementary Information

A Facilely Synthesized Amino-functionalized Metal-organic Framework for Highly Specific and Efficient Enrichment of Glycopeptides

Yi-Wei Zhang, Ze Li, Qiang Zhao, Ying-Lin Zhou, Hu-Wei Liu and Xin-Xiang Zhang*

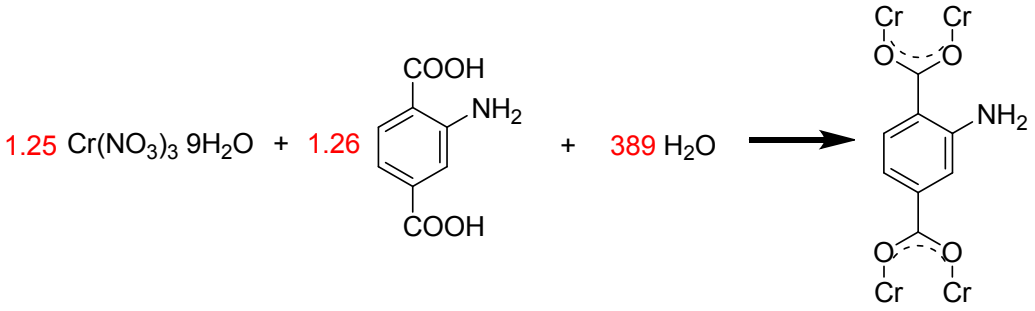
Beijing National Laboratory for Molecular Sciences (BNLMS), MOE Key Laboratory of Bioorganic Chemistry and Molecular Engineering, College of Chemistry, Peking University, Beijing 100871, P.R. China

Experimental Section

Chemicals and Materials

Horseradish peroxidase (HRP), bovine serum albumin (BSA), and 2, 5-dihydroxyl benzoic acid (DHB), dithiothreitol (DTT) were purchased from sigma-Aldrich (St, Louis, MO, USA). Secreted form mouse IgG (anti-α-fetoprotein (AFP) monoclonal) was obtained from Biocell CO. LTD (Zhengzhou, China). 2-aminoterephthalic acid and iodoacetamide (IAA) were from Alfa Aesar China Chemical Co., Ltd. (Tianjin, China). Sequencing grade modified trypsin was purchased from Promega. Peptide-N-glycosidase (PNGase F) was from New England Biolabs (Ipswich, MA, USA). HPLC grade methanol and acetonitrile (ACN) were from Fisher Scientific (Fair Lawn, NJ, USA). Formic acid (FA) was from TCI (Tokyo, Japan). Chromic nitrate hydrate, urea, ammonium bicarbonate (NH_4HCO_3) and other analytical grade reagents were obtained from Beijing Chemicals (Beijing, China). Water was purified by Milli-Q pure water system (Millipore, Bedford, MA, USA).

Synthesis of MIL-101(Cr)-NH_2



Scheme S1 Synthesis of MIL-101(Cr)-NH_2

MIL-101(Cr)-NH_2 was synthesized according to *Jiang et al.*[S1] with a little modifications. Scheme S1 was schematic representation of the synthesis procedure. In detail, chromic nitrate hydrate (500 mg, 1.25 mmol) and 2-aminoterephthalic acid (230 mg, 1.26 mmol) were dispersed in deionized water (7 mL, 389 mmol). After stirred for 3 h at room temperature, the suspension was heated under autogeneous pressure in a Teflon-lined stainless steel autoclave at 130 ℃ for 24 h. The autoclave then cooled down to room temperature. The resulting green precipitate was collected by centrifugation and washed with ethanol for five times to remove the excess regents, followed by drying at 80 ℃.

Characterization

Fourier-transformed infrared spectroscope (FT-IR) characterization was performed on a Bruker VERTEX 70/70v FT-IR spectrometer to confirm the existence of 2-aminoterephthalic acid (NH_2bdc) skeleton in the structure. In the FT-IR spectrum (Fig. S1), the double peaks at 3421 cm^{-1} and 3384 cm^{-1} ascribed to the asymmetrical and symmetrical stretching vibration absorption of the amine groups. The peak at 1582 cm^{-1} corresponded to the N-H bending vibration while the 1338 cm^{-1} and 1257 cm-1 peaks represented for C-N stretching of aromatic amines. The peaks at 1499 cm^{-1} and 1430 cm^{-1} illustrated the -(O-C-O)- stretching vibration in the MIL-101(Cr)-NH_2 skeleton. X-ray diffraction (XRD) pattern of the dried powders was carried out on a Rigaku

D/MAX-PC 2500 diffractometer with monochromatic Cu Kα radiation (λ = 1.5406 Å) at an accelerating potential of 40 kV and a tube current of 300 mA. Powder XRD (Fig. S2) indicated that the resulting powder was a typical MIL-101(Cr) structure with a pattern similar to the simulated. The rather broad Bragg reflections were consequence of the small particle sizes, as supported by transmission electron microscopy (TEM) images (Fig. S3), which revealed the uniform and narrow size distribution of particles with about 25 nm. The TEM images were taken on a FEI Tecnai G2 T20 transmission electron microscope. Thermal stability was measured on a simultaneous thermal analysis apparatus (Q600SDT TGA-DTA-DSC) in air at a heat rate of 5 ℃ / min up to 800 ℃. Thermogravimetric analysis showed the nanoparticles were stable up to 260 ℃ (Fig. S4). The nitrogen adsorption-desorption isotherm was recorded by ASAP 2020M apparatus at 77.2 K. The range of relative pressures was between 0 and 1. The BET surface area was calculated over the range of relative pressures between 0.05 and 0.20. N_2 adsorption experiment illustrates the amazing BET surface area of 2187.4 m²/g and Langmuir surface area of 3129.0 m²/g. The sharp uptake under low pressure ($P/P_0=10^{-5}$ to 0.1) in N_2 adsorption-desorption isotherms (Fig. S5) demonstrated the microporous feature of the material while another uptake occurred near $P/P_0=1.0$ attributed to the textural pores created by nanoparticles aggregation.

Preparation of protein digests

Glycoproteins were dissolved in 50 mM NH_4HCO_3 solution (HRP 2 mg/mL and IgG 4 mg/mL, respectively.) and denatured by adding 8 M urea, 10 mM DTT for 45 min at 60 ℃. Then 100 mM IAA was added and reacted at 37℃ for 45 min in dark. The solution was diluted until the urea concentration reached 0.8 M and treated with trypsin at 37℃ (enzyme/protein ratio of 1:20

for HRP and enzyme/protein ratio of 1:25 for IgG) for 18 h. The tryptic digests were stored at -20 ℃ before use.

Enrichment of glycopeptides

MIL-101(Cr)-NH_2 was dispersed in ACN contain 0.1% FA. Then 10 μL tryptic digest of HRP (50 pmol) or IgG (200 pmol) was added into 90 μL 5 mg/mL MIL-101(Cr)-NH_2 suspension and shaken at room temperature for 5 min. After capture, the slurry was centrifuged at 10000 rpm for 3 min to remove the suspensions. Then the MOFs were washed twice with 80% ACN containing 1% H_3PO_4 to remove the most non-glycopeptides. Finally, 30 μL 30% ACN containing 0.1% FA were added to elute the glycopeptides and collected for analysis by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS).

Deglycosylation of N-linked glycopeptides by PNGase F

The eluted glycopeptides after enrichment were dissolved in 17 μL deionized water, followed by 2 μL 10×G7 Reaction Buffer and 1 μL PNGase F added to the solution then incubated at 37 ℃ for 18 h. The reaction solutions were directly spotted on the target plate for MALDI-TOF-MS analysis or analysis by LC-MS/MS.

Mass spectrometry analysis

MALDI-TOF-MS Analysis

A mixture of 30 mg/mL DHB in 50% (v/v) ACN was prepared as the matrix. Then, 1 μL aliquot of the eluate and 1 μL matrix solution were sequentially dropped onto the ground steel plate, dried at room temperature for MS analysis. All MALDI-MS spectra were acquired in reflection

and positive mode on an Ultraflex MALDI TOF/TOF mass spectrometer (Bruker Daltonics, Bremen, Germany), equipped with pulsed-nitrogen laser (337 nm). The accelerating voltage was set at 25 kV, 70 ns extraction delay. Laser fluence was optimized at 20%. Each mass spectrum was generated by an average of 400 laser shots.

LC-MS/MS Analysis

All LC-MS/MS analyses were performed on a Thermo Q-Exactive Orbitrap mass spectrometer with a Thermo Ultimate 3000 UHPLC (Thermo, San Jose, CA, USA). The deglycosylated glycopeptides were loaded on a C18 column. 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, 5 % B was hold for 10 min, then from 5% to 35% for 120 min, from 35% to 80% for 10 min and held at 80% for 6 min. The flow rate was 0.2 mL/min. Full mass scan was acquired from 300 to 2000 m/z with resolution 70000 at R=400 m/z. The MS/MS spectra were obtained in data-dependent $ddMS^2$ mode.

Database search and data analysis

The LC-MS/MS raw file was searched against the Swiss-Prot database using MASCOT software (version 2.3.02). The search parameters were set as follows: fixed modification of cysteine residues (+57 Da), variable modification of methionine oxidation (+16 Da), N-terminal acetylation, and deamidation (N), at most two missed tryptic cleavage sites, 20 ppm error tolerance in MS and 0.1 Da error tolerance in MS/MS. The resulting data files were exported with the cut off set to 0.05. According to the consensus N-X-S/T (X≠P) sequon of N-

glycosylation, the remaining peptide sequences were additionally filtered to remove non-motif containing peptides.

Results

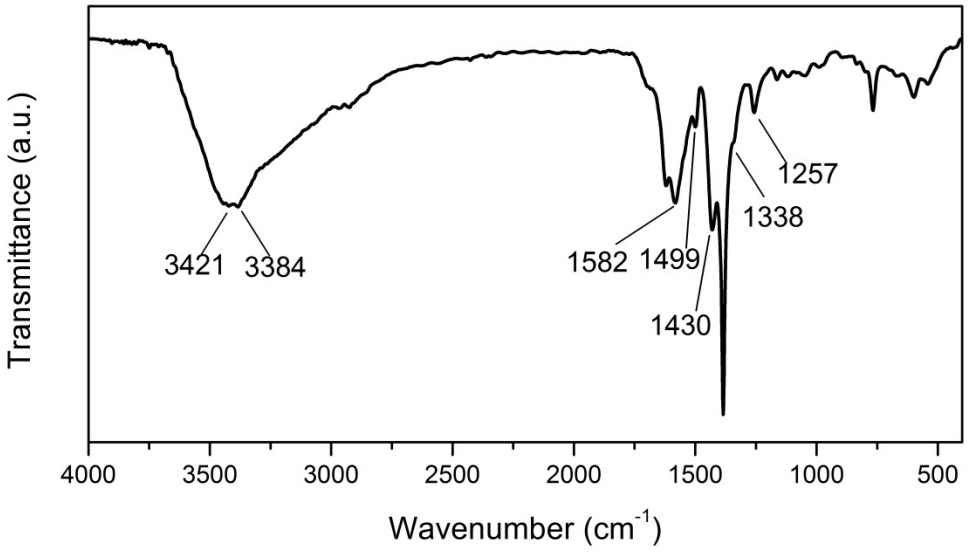


Figure S1 FT-IR spectrum of the MIL-101(Cr)-NH_2

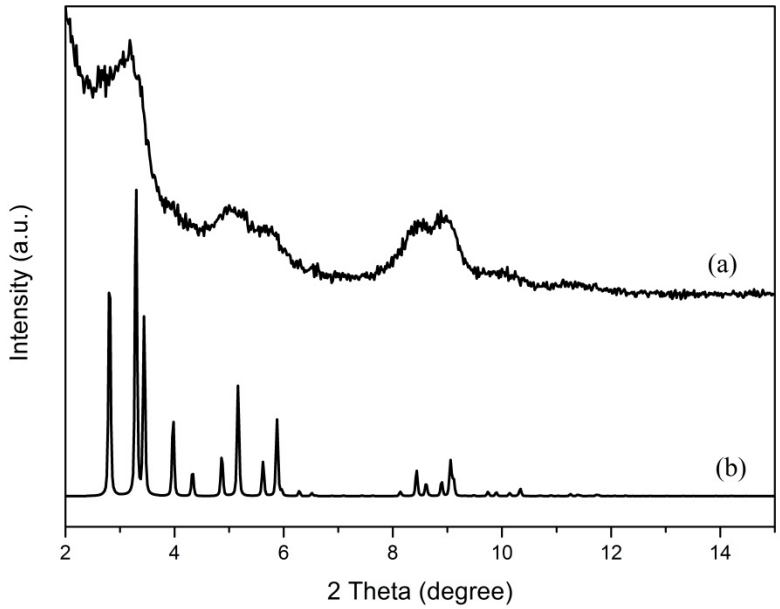


Figure S2 Powder XRD pattern of MIL-101(Cr)-NH_2 (a) and the simulated PXRD pattern calculated from the MIL-101(Cr) crystal structure (b)

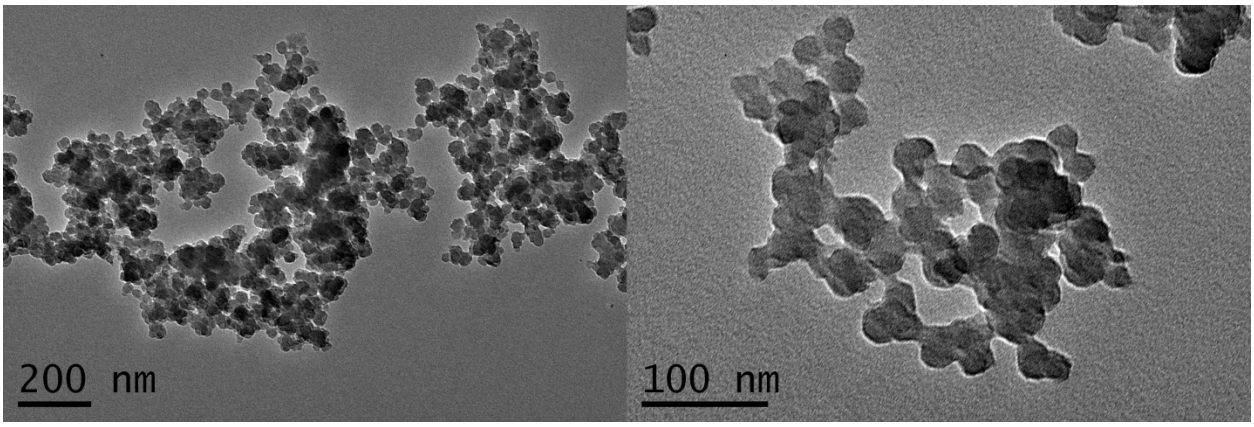


Figure S3 TEM images of MIL-101(Cr)-NH_2

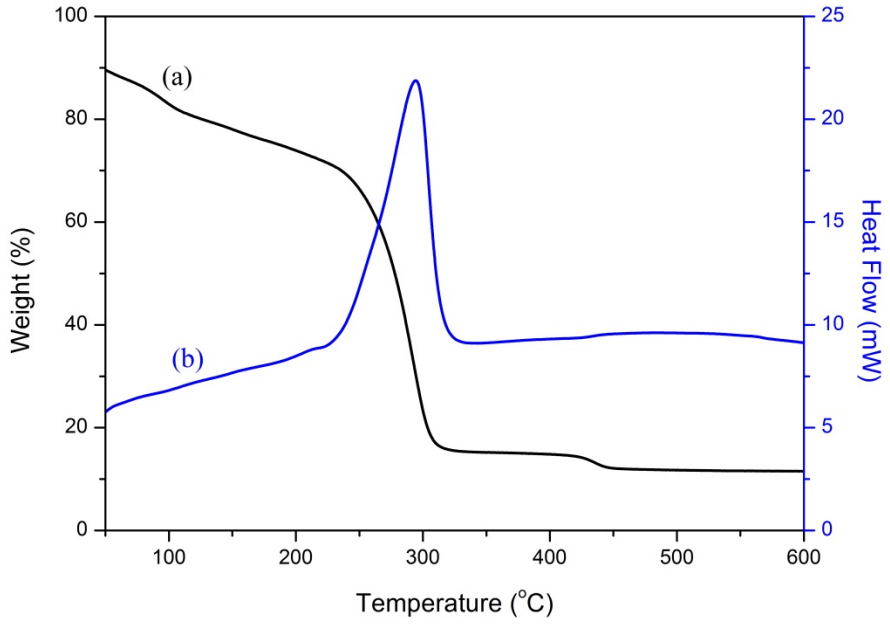


Figure S4 The thermogravimetric curve (a) and differential thermal analysis spectrum (b) for MIL-101(Cr)-NH_2 powder. TG curve showed the start decomposition temperature was about 260 ℃. The sharp endothermic peak 300 ℃ in heat flow spectrum was due to the rapid decomposition of the resulting powders.

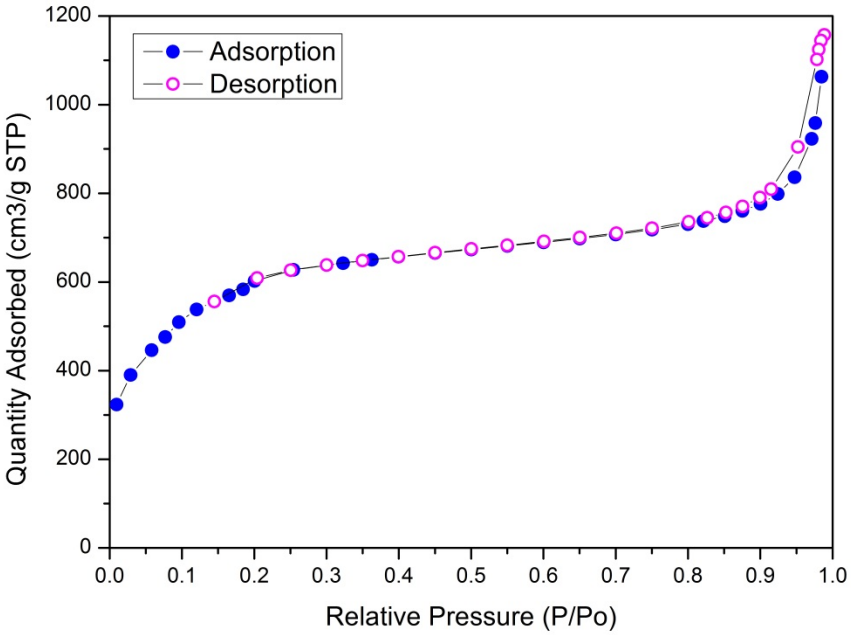


Figure S5 N_2 adsorption-desorption isotherms at 77.2 K

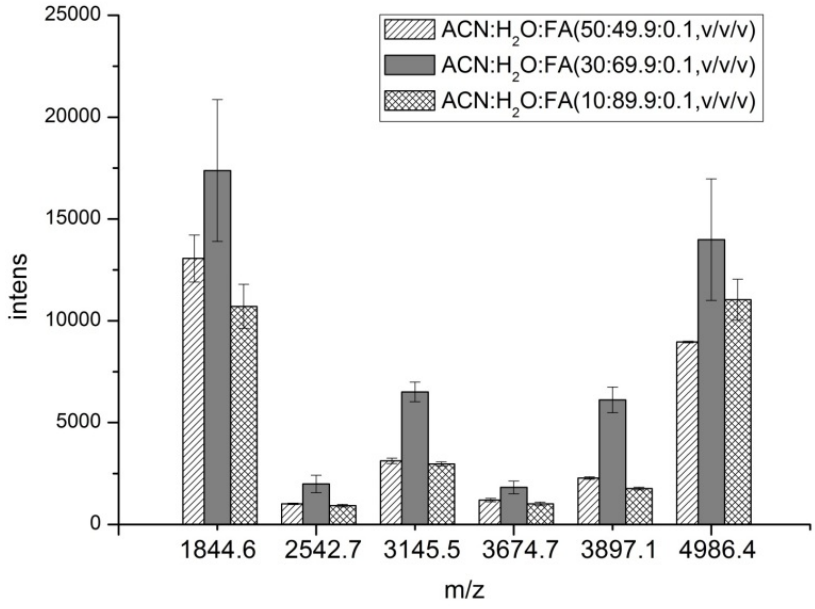


Figure S6 The influence of final acetonitrile concentration on peak intensity of six chosen glycopeptides enriched by MIL-101(Cr)-NH_2; Sample: HRP tryptic digests, 1.0 pmol/μL

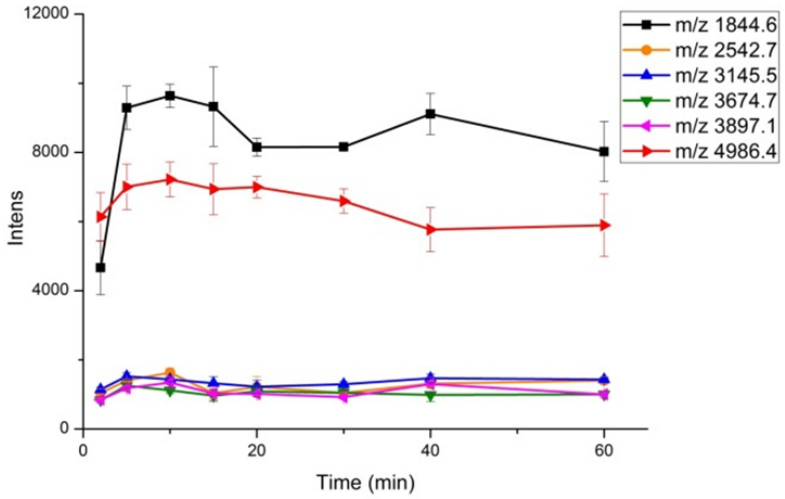


Figure S7 The influence of incubation time on peak intensity of six chosen glycopeptides enriched by MIL-101(Cr)-NH_2; Sample: HRP tryptic digests, 1.0 pmol/μL

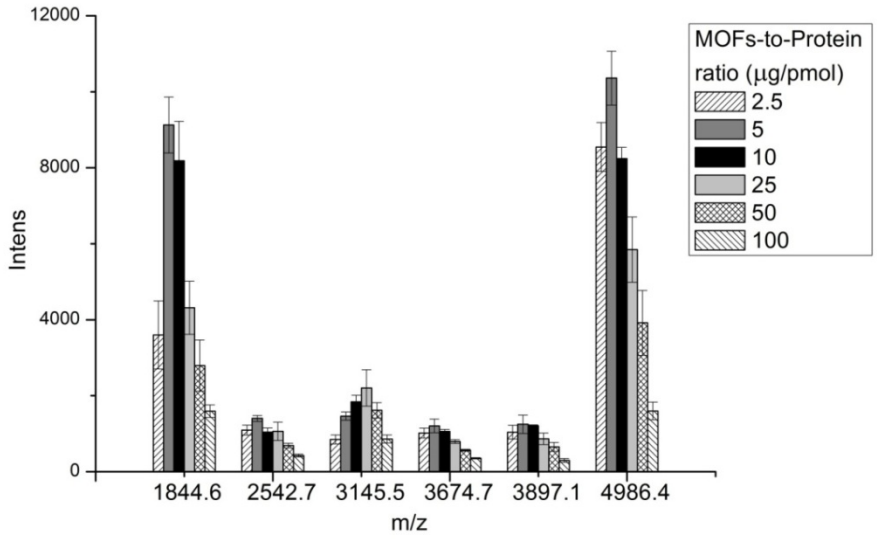


Figure S8 The influence of MOFs-to-protein ratio on peak intensity of six chosen glycopeptides enriched by MIL-101(Cr)-NH_2; Sample: HRP tryptic digests, 1.0 pmol/μL

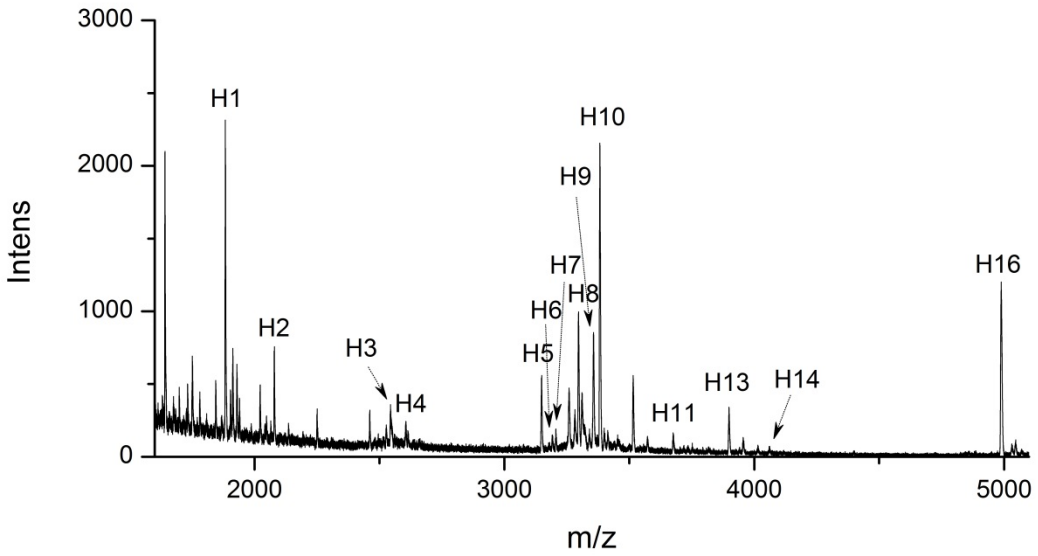


Figure S9 MALDI-TOF-MS spectra of the mixture of HRP (1.0 pmol/μL) and BSA tryptic digests (10.0 pmol/μL) after enriching and sequenced washing

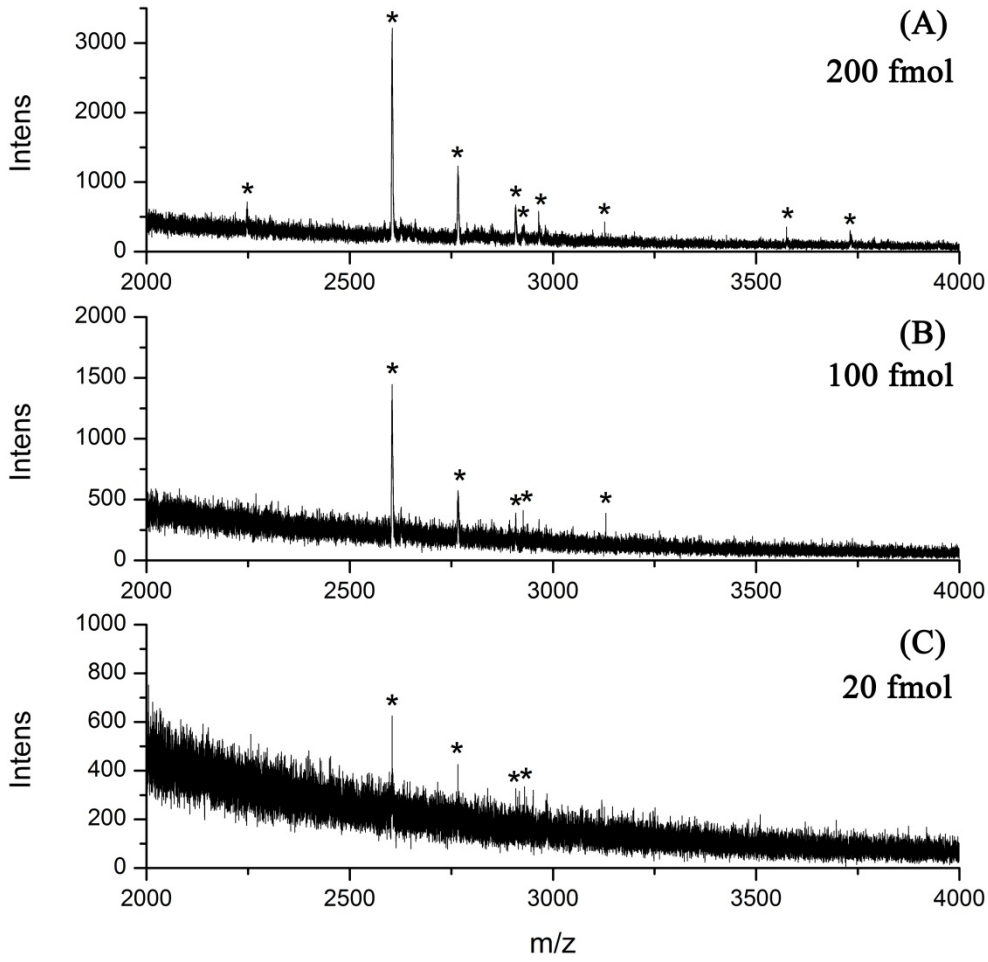


Figure S10 MALDI-TOF-MS spectra of IgG tryptic digest after enrichment by MIL-101(Cr)-NH_2 (A) 200 fmol (1 μL), (B) 100 fmol (1 μL), (C) 20 fmol (1 μL), * denotes glycopeptides

Table S1 Observed glycopeptides and glycan structures of mouse IgG tryptic digests enriched by MIL-101(Cr)-NH_2. Hex, HexNAc, Fuc and NeuAc are the abbreviations of hexose, N-acetylhexosamine, fucose and N-acetylneuraminic acid, respectively. ***N***# denotes the N-linked glycosylation sites

Peak Number	Observed m/z	Theoretic glycan	Glycan structure	Peptide sequence

		mass		
I1	2251.5	1113.4	[Hex]3[HexNAc]3	EEQF***N***#STFR
I2	2374.0	1234.4	[Hex]5[HexNAc]2	EEQF***N***#STFR
I3	2401.1	1259.5	[Hex]3[HexNAc]3[Fuc]1	EEQF***N***#STFR
I4	2415.8	1275.5	[Hex]4[HexNAc]3	EEQF***N***#STFR
I5	2457.2	1316.5	[Hex]3[HexNAc]4	EEQF***N***#STFR
I6	2603.4	1462.5	[Hex]3[HexNAc]4[Fuc]1	EEQF***N***#STFR
I7	2619.6	1478.5	[Hex]4[HexNAc]4	EEQF***N***#STFR
I8	2659.6	1519.6	[Hex]3[HexNAc]5	EEQF***N***#STFR
I9	2765.0	1624.6	[Hex]4[HexNAc]4[Fuc]1	EEQF***N***#STFR
I10	2821.7	1681.6	[Hex]4[HexNAc]5	EEQF***N***#STFR
I11	2908.8	1769.6	[Hex]4[HexNAc]4[NeuAc]1	EEQF***N***#STFR
I12	2926.7	1786.7	[Hex]5[HexNAc]4[Fuc]1	EEQF***N***#STFR
I13	2965.4	1827.7	[Hex]4[HexNAc]5[Fuc]1	EEQF***N***#STFR
I14	2982.8	1843.7	[Hex]5[HexNAc]5	EEQF***N***#STFR
I15	3063.1	1915.7	[Hex]4[HexNAc]4[Fuc]1[NeuAc]1	EEQF***N***#STFR
I16	3126.1	1989.7	[Hex]5[HexNAc]5[Fuc]1	EEQF***N***#STFR
I17	3461.5	2321.9	[Hex]4[HexNAc]6[Fuc]1[NeuAc]1	EEQF***N***#STFR
I18	3508.6	2368.8	[Hex]5[HexNAc]4[Fuc]1[NeuAc]2	EEQF***N***#STFR
I19	3580.5	2442.9	[Hex]6[HexNAc]5[Fuc]1[NeuAc]1	EEQF***N***#STFR
I20	3730.2	2587.9	[Hex]5[HexNAc]5[NeuAc]2	EEQF***N***#STFR
I21	3786.6	2646.0	[Hex]6[HexNAc]6[Fuc]1[NeuAc]1	EEQF***N***#STFR
I22	3824.4	2684.9	[Hex]10[HexNAc]3[Fuc]1[NeuAc]1	EEQF***N***#STFR

Table S2 Observed glycopeptides and glycan structures of HRP tryptic digests enriched by MIL-101(Cr)-NH_2. Hex, HexNAc, Fuc and Xyl are the abbreviations of hexose, N-acetylhexosamine, fucose and xylose, respectively. ***N***# denotes the N-linked glycosylation sites

Peak Number	Observed m/z	Theoretic glycan mass	Glycan structure	Peptide sequence
H1	1844.6	1188.4	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	^{184}NVGL***N***#R^{189}
H2	2069.5	1188.4	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	^{42}PNVS***N***#IVR^{49}
H3	2542.7	1188.4	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	^{282}SSP***N***#ATDTIPLVR^{294}
H4	2613.1	1188.4	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	^{225}PTL***N***#TTYLQTLR^{236}
H5	3145.5	1188.4	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	^{237}GLCPLNG***N***#LSALVDFDLR^{254}
H6	3190.2	367.1	[HexNAc]1[Fuc]1	^{69}LHFHDCFVNGCDASILLD***N***#TTSFR^{92}
H7	3206.7	1042.4	[Hex]3[HexNAc]2[Xyl]1	^{295}SFA***N***#STQTFFNAFVEAMDR^{313}
H8	3338.9	1188.4	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	^{31}QLTPTFYDNSCP***N***#VSNIVR^{49}

H9	3353.3	1188.4	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	^{295}SFA**N**#STQTFFNAFVEAMDR^{313}
H10	3387.6	2226.8	[Hex]6[HexNAc]4[Fuc]2[Xyl]1	^{180}DSFRNVGL***N***#R^{189}
H11	3675.7	1188.4	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	^{272}GLIQSDQELFSSP***N***#ATDTIPLVR^{294}
H12	3752.2	1042.4	[Hex]3[HexNAc]2[Xyl]1	^{69}LHFHDCFVNGCDASILLD***N***#TTSFRTEK^{95}
H13	3897.1	1188.4	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	^{69}LHFHDCFVNGCDASILLD***N***#TTSFRTEK^{95}
H14	4059.7	1042.4	[Hex]3[HexNAc]2[Xyl]1	^{31}QLTPTFYDNSC(AAVESACPR)P***N***#VSNIV R^{49}-H2O
H15	4840.8	1188.4	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	^{214}LY**N**#FSNTGLPDPTL***N***#TTYLQTLR^{236}
		1042.4	[Hex]3[HexNAc]2[Xyl]1	
H16	4987.8	1188.4	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	^{214}LY***N***#FSNTGLPDPTL***N***#TTYLQTLR^{236}
		1188.4	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	

Table S3 Observed glycopeptides of human serum tryptic digests enriched by MIL-101(Cr)-NH_2. OS=OrganismName, GN=GeneName, PE=ProteinExistence, SV=SequenceVersion. N denotes the N-linked glycosylation sites

Protein Accession Number	Protein Description	Peptide Sequence
A1AG1_HUMAN	Alpha-1-acid glycoprotein 1 OS=Homo sapiens GN=ORM1 PE=1 SV=1	WFYIASAFRNEEYNK
A1AG1_HUMAN	Alpha-1-acid glycoprotein 1 OS=Homo sapiens GN=ORM1 PE=1 SV=1	QDQCIYNTTYLNVQRENGTISR
A1AT_HUMAN	Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3	YLGNATAIFFLPDEGK
A1AT_HUMAN	Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3	QLAHQSNSTNIFFSPVSIATAFAMLSLGTK
A1AT_HUMAN	Alpha-1-antitrypsin OS=Homo sapiens GN=SERPINA1 PE=1 SV=3	KLSSWVLLMKYLGNATAIFFLPDEGK
A1BG_HUMAN	Alpha-1B-glycoprotein OS=Homo sapiens GN=A1BG PE=1 SV=3	PLANVTLTCQAHLETPDFQLFK
A1BG_HUMAN	Alpha-1B-glycoprotein OS=Homo sapiens GN=A1BG PE=1 SV=3	EGDHEFLEVPEAQEDVEATFPVHQPGNYSCSYR
A2MG_HUMAN	Alpha-2-macroglobulin OS=Homo sapiens GN=A2M PE=1 SV=2	TEVSSNHVLIYLDKVSNQTLSLFFTVLQDVPVR
A2MG_HUMAN	Alpha-2-macroglobulin OS=Homo sapiens GN=A2M PE=1 SV=2	SNHVSRTEVSSNHVLIYLDKVSNQTLSLFFTVLQDVPVR
A2MG_HUMAN	Alpha-2-macroglobulin OS=Homo sapiens GN=A2M PE=1 SV=2	GCVLLSYLNETVTVSASLESVR
A2MG_HUMAN	Alpha-2-macroglobulin OS=Homo sapiens GN=A2M PE=1 SV=2	GCVLLSYLNETVTVSASLESVRGN
A2MG_HUMAN	Alpha-2-macroglobulin OS=Homo sapiens GN=A2M PE=1 SV=2	VIFIRGNEANYYSNATTDEHGLVQFSINTTNVMGTSLTVR
A2MG_HUMAN	Alpha-2-macroglobulin OS=Homo sapiens GN=A2M PE=1 SV=2	VSNQTLSLFFTVLQDVPVR

A2MG_HUMAN	Alpha-2-macroglobulin OS=Homo sapiens GN=A2M PE=1 SV=2	GNEANYYSNATTDEHGLVQFSINTTNVMGTSLTVR
AACT_HUMAN	Alpha-1-antichymotrypsin OS=Homo sapiens GN=SERPINA3 PE=1 SV=2	YTGNASALFILPDQDK
AACT_HUMAN	Alpha-1-antichymotrypsin OS=Homo sapiens GN=SERPINA3 PE=1 SV=2	NVIFSPLSISTALAFLSLGAHNTTLTEILK
AACT_HUMAN	Alpha-1-antichymotrypsin OS=Homo sapiens GN=SERPINA3 PE=1 SV=2	FNLTETSEAEIHQSFQHLLR
AACT_HUMAN	Alpha-1-antichymotrypsin OS=Homo sapiens GN=SERPINA3 PE=1 SV=2	TLNQSSDELQLSMGNAMFVK
AFAM_HUMAN	Afamin OS=Homo sapiens GN=AFM PE=1 SV=1	HNFSHCCSKVDAQRR
AFAM_HUMAN	Afamin OS=Homo sapiens GN=AFM PE=1 SV=1	FNETTEKSLKMVQQECK
ANT3_HUMAN	Antithrombin-III OS=Homo sapiens GN=SERPINC1 PE=1 SV=1	AAINKWVSNK
ANT3_HUMAN	Antithrombin-III OS=Homo sapiens GN=SERPINC1 PE=1 SV=1	LGACNDTLQQLMEVFK
ANT3_HUMAN	Antithrombin-III OS=Homo sapiens GN=SERPINC1 PE=1 SV=1	LGACNDTLQQLMEVFKFDTISEK
ANT3_HUMAN	Antithrombin-III OS=Homo sapiens GN=SERPINC1 PE=1 SV=1	ENAEQSRAAINKWVSNK
APOB_HUMAN	Apolipoprotein B-100 OS=Homo sapiens GN=APOB PE=1 SV=1	NLTDFAEQYSIQDWAKRMK
APOB_HUMAN	Apolipoprotein B-100 OS=Homo sapiens GN=APOB PE=1 SV=1	HLRVNQNLVYESGSLNFSK
APOB_HUMAN	Apolipoprotein B-100 OS=Homo sapiens GN=APOB PE=1 SV=1	PTVSSSMEFKYDFNSSMLYSTAK
APOB_HUMAN	Apolipoprotein B-100 OS=Homo sapiens GN=APOB PE=1 SV=1	AEEEMLENVSLVCPK
APOB_HUMAN	Apolipoprotein B-100 OS=Homo sapiens GN=APOB PE=1 SV=1	QVLFLDTVYGNCSTHFTVK
APOB_HUMAN	Apolipoprotein B-100 OS=Homo sapiens GN=APOB PE=1 SV=1	TIHDLHLFIENIDFNK
APOB_HUMAN	Apolipoprotein B-100 OS=Homo sapiens GN=APOB PE=1 SV=1	ELCTISHIFIPAMGNITYDFSFK
APOB_HUMAN	Apolipoprotein B-100 OS=Homo sapiens GN=APOB PE=1 SV=1	LNGESNLRFNSSYLQGTNQITGR
APOD_HUMAN	Apolipoprotein D OS=Homo sapiens GN=APOD PE=1 SV=1	CPNPPVQENFDVNK
APOD_HUMAN	Apolipoprotein D OS=Homo sapiens GN=APOD PE=1 SV=1	CIQANYSLMENGKIK
APOD_HUMAN	Apolipoprotein D OS=Homo sapiens GN=APOD PE=1 SV=1	ADGTVNQIEGEATPVNLTEPAKLEVK
APOE_HUMAN	Apolipoprotein E OS=Homo sapiens GN=APOE PE=1 SV=1	VQAAVGTSAAPVPSDNH
APOH_HUMAN	Beta-2-glycoprotein 1 OS=Homo sapiens GN=APOH PE=1 SV=3	PSAGNNSLYR
APOH_HUMAN	Beta-2-glycoprotein 1 OS=Homo sapiens GN=APOH PE=1 SV=3	LGNWSAMPSCKASCK
APOH_HUMAN	Beta-2-glycoprotein 1 OS=Homo sapiens GN=APOH PE=1 SV=3	DTAVFECLPQHAMFGNDTITCTTHGNWTK
AT1B3_HUMAN	Sodium/potassium-transporting ATPase subunit beta-3 OS=Homo sapiens GN=ATP1B3 PE=1 SV=1	KSLNQSLAEWK

AT1B3_HUMAN	Sodium/potassium-transporting ATPase subunit beta-3 OS=Homo sapiens GN=ATP1B3 PE=1 SV=1	PYTLEEQKNLTVCPDGALFEQK
ATS20_HUMAN	A disintegrin and metalloproteinase with thrombospondin motifs 20 OS=Homo sapiens GN=ADAMTS20 PE=2 SV=2	KLEDTNCSQVQKPPTHK
ATS20_HUMAN	A disintegrin and metalloproteinase with thrombospondin motifs 20 OS=Homo sapiens GN=ADAMTS20 PE=2 SV=2	CPQGQFSINLSGTGMKISSTAK
C1R_HUMAN	Complement C1r subcomponent OS=Homo sapiens GN=C1R PE=1 SV=2	MLLTFHTDFSNEENGTIMFY
C1R_HUMAN	Complement C1r subcomponent OS=Homo sapiens GN=C1R PE=1 SV=2	EHEAQSNASLDVFLGHTNVEELMK
C1R_HUMAN	Complement C1r subcomponent OS=Homo sapiens GN=C1R PE=1 SV=2	RSYPPDLRCNYSIRVER
CERU_HUMAN	Ceruloplasmin OS=Homo sapiens GN=CP PE=1 SV=1	ELHHLQEQNVSNAFLDK
CERU_HUMAN	Ceruloplasmin OS=Homo sapiens GN=CP PE=1 SV=1	AGLQAFFQVQECNKSSSK
CERU_HUMAN	Ceruloplasmin OS=Homo sapiens GN=CP PE=1 SV=1	EHEGAIYPDNTTDFQRADDK
CERU_HUMAN	Ceruloplasmin OS=Homo sapiens GN=CP PE=1 SV=1	ENLTAPGSDSAVFFEQGTTR
CERU_HUMAN	Ceruloplasmin OS=Homo sapiens GN=CP PE=1 SV=1	EFYLFPTVFDENESLLLEDNIR
CERU_HUMAN	Ceruloplasmin OS=Homo sapiens GN=CP PE=1 SV=1	DVDKEFYLFPTVFDENESLLLEDNIR
CFAH_HUMAN	Complement factor H OS=Homo sapiens GN=CFH PE=1 SV=4	SPDVINGSPISQK
CFAH_HUMAN	Complement factor H OS=Homo sapiens GN=CFH PE=1 SV=4	DGRWQSIPLCVEKIPCSQPPQIEHGTINSSR
CFAH_HUMAN	Complement factor H OS=Homo sapiens GN=CFH PE=1 SV=4	WDPEVNCSMAQIQLCPPPPQIPNSHNMTTTLNYR
CFAH_HUMAN	Complement factor H OS=Homo sapiens GN=CFH PE=1 SV=4	SPYEMFGDEEVMCLNGNWTEPPQCK
CLUS_HUMAN	Clusterin OS=Homo sapiens GN=CLU PE=1 SV=1	EILSVDCSTNNPSQAK
CLUS_HUMAN	Clusterin OS=Homo sapiens GN=CLU PE=1 SV=1	EDALNETRESETKLK
CLUS_HUMAN	Clusterin OS=Homo sapiens GN=CLU PE=1 SV=1	ELPGVCNETMMALWEECK
CLUS_HUMAN	Clusterin OS=Homo sapiens GN=CLU PE=1 SV=1	MLNTSSLLEQLNEQFNWVSR
CO3_HUMAN	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2	TVLTPATNHMGNVTFTIPANREFK
CO3_HUMAN	Complement C3 OS=Homo sapiens GN=C3 PE=1 SV=2	SYTVAIAGYALAQMGRLKGPLLNK
CO4A_HUMAN	Complement C4-A OS=Homo sapiens GN=C4A PE=1 SV=1	FSDGLESNSSTQFEVKK
CO4A_HUMAN	Complement C4-A OS=Homo sapiens GN=C4A PE=1 SV=1	GLNVTLSSTGR
CO4A_HUMAN	Complement C4-A OS=Homo sapiens GN=C4A PE=1 SV=1	GLNVTLSSTGRNGFKSHALQLNNR
FETUA_HUMAN	Alpha-2-HS-glycoprotein OS=Homo sapiens GN=AHSG PE=1 SV=1	VCQDCPLLAPLNDTR
FETUA_HUMAN	Alpha-2-HS-glycoprotein OS=Homo sapiens GN=AHSG PE=1 SV=1	AALAAFNAQNNGSNFQLEEISR

FETUA_HUMAN	Alpha-2-HS-glycoprotein OS=Homo sapiens GN=AHSG PE=1 SV=1	VCQDCPLLAPLNDTRVVHAAK
FINC_HUMAN	Fibronectin OS=Homo sapiens GN=FN1 PE=1 SV=3	DQCIVDDITYNVNDTFHK
FINC_HUMAN	Fibronectin OS=Homo sapiens GN=FN1 PE=1 SV=3	HEEGHMLNCTCFGQGR
FINC_HUMAN	Fibronectin OS=Homo sapiens GN=FN1 PE=1 SV=3	HEEGHMLNCTCFGQGRGR
FINC_HUMAN	Fibronectin OS=Homo sapiens GN=FN1 PE=1 SV=3	LDAPTNLQFVNETDSTVLVR
FINC_HUMAN	Fibronectin OS=Homo sapiens GN=FN1 PE=1 SV=3	GGNSNGALCHFPFLYNNHNYTDCTSEGRR
GP115_HUMAN	Probable G-protein coupled receptor 115 OS=Homo sapiens GN=GPR115 PE=2 SV=2	SRAAENASLGPTNGSK
GP115_HUMAN	Probable G-protein coupled receptor 115 OS=Homo sapiens GN=GPR115 PE=2 SV=2	SRAAENASLGPTNGSKLMNR
GP115_HUMAN	Probable G-protein coupled receptor 115 OS=Homo sapiens GN=GPR115 PE=2 SV=2	GFHINHNTSEKSLNFSMSMNNTTEDILGMVQIPR
GP115_HUMAN	Probable G-protein coupled receptor 115 OS=Homo sapiens GN=GPR115 PE=2 SV=2	SSETTSGNIAFIVELLKNISTDLSDNVTR
GP115_HUMAN	Probable G-protein coupled receptor 115 OS=Homo sapiens GN=GPR115 PE=2 SV=2	NISTDLSDNVTREK
GP115_HUMAN	Probable G-protein coupled receptor 115 OS=Homo sapiens GN=GPR115 PE=2 SV=2	CEGPCISSSNCSQPCAK
GP115_HUMAN	Probable G-protein coupled receptor 115 OS=Homo sapiens GN=GPR115 PE=2 SV=2	NASSDLLQSVNLFAR
HEMO_HUMAN	Hemopexin OS=Homo sapiens GN=HPX PE=1 SV=2	GHGHRNGTGHGNSTHHGPEYMR
HEMO_HUMAN	Hemopexin OS=Homo sapiens GN=HPX PE=1 SV=2	SWPAVGNCSSALRWLGR
HPT_HUMAN	Haptoglobin OS=Homo sapiens GN=HP PE=1 SV=1	KQLVEIEKVVLHPNYSQVDIGLIK
HPT_HUMAN	Haptoglobin OS=Homo sapiens GN=HP PE=1 SV=1	NLFLNHSENATAK
HRG_HUMAN	Histidine-rich glycoprotein OS=Homo sapiens GN=HRG PE=1 SV=1	HSHNNNSSDLHPHK
IGHA1_HUMAN	Ig alpha-1 chain C region OS=Homo sapiens GN=IGHA1 PE=1 SV=2	PALEDLLLGSEANLTCTLTGLR
IGHA2_HUMAN	Ig alpha-2 chain C region OS=Homo sapiens GN=IGHA2 PE=1 SV=3	TPLTANITKSGNTFR
IGHA2_HUMAN	Ig alpha-2 chain C region OS=Homo sapiens GN=IGHA2 PE=1 SV=3	PALEDLLLGSEANLTCTLTGLR
IGHA2_HUMAN	Ig alpha-2 chain C region OS=Homo sapiens GN=IGHA2 PE=1 SV=3	HYTNPSQDVTVPCPVPPPPPCCHPR
IGHG1_HUMAN	Ig gamma-1 chain C region OS=Homo sapiens GN=IGHG1 PE=1 SV=1	EEQYNSTYR
IGHG1_HUMAN	Ig gamma-1 chain C region OS=Homo sapiens GN=IGHG1 PE=1 SV=1	FNWYVDGVEVHNAKTKPREEQYNSTYR
IGHG2_HUMAN	Ig gamma-2 chain C region OS=Homo sapiens GN=IGHG2 PE=1 SV=2	EEQFNSTFR
IGHG3_HUMAN	Ig gamma-3 chain C region OS=Homo sapiens GN=IGHG3 PE=1 SV=2	EEQYNSTFR
IGHM_HUMAN	Ig mu chain C region OS=Homo sapiens GN=IGHM PE=1 SV=3	NNSDISSTRGFPSVLR

IGHM_HUMAN	Ig mu chain C region OS=Homo sapiens GN=IGHM PE=1 SV=3	VDHRGLTFQQNASSMCVPDQDTAIR
IGHM_HUMAN	Ig mu chain C region OS=Homo sapiens GN=IGHM PE=1 SV=3	PTLYNVSLVMSDTAGTCY
IGHM_HUMAN	Ig mu chain C region OS=Homo sapiens GN=IGHM PE=1 SV=3	THTNISESHPNATFSAVGEASICEDDWNSGER
ITIH1_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H1 OS=Homo sapiens GN=ITIH1 PE=1 SV=3	ICDLLVANNHFAHFFAPQNLTNMNK
ITIH2_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H2 OS=Homo sapiens GN=ITIH2 PE=1 SV=2	GAFISNFSMTVDGK
ITIH2_HUMAN	Inter-alpha-trypsin inhibitor heavy chain H2 OS=Homo sapiens GN=ITIH2 PE=1 SV=2	VVNNSPQPQNVVFDVQIPK
KLKB1_HUMAN	Plasma kallikrein OS=Homo sapiens GN=KLKB1 PE=1 SV=1	TSTRIVGGTNSSWGEWPWQVSLQVK
KLKB1_HUMAN	Plasma kallikrein OS=Homo sapiens GN=KLKB1 PE=1 SV=1	IYPGVDFGGEELNVTFVK
KLOTB_HUMAN	Beta-klotho OS=Homo sapiens GN=KLB PE=1 SV=1	AIWSKNPNFTPVNESQLFLYDTFPK
KLOTB_HUMAN	Beta-klotho OS=Homo sapiens GN=KLB PE=1 SV=1	NVSSTNGSSDSYIFLEK
KLOTB_HUMAN	Beta-klotho OS=Homo sapiens GN=KLB PE=1 SV=1	GWLSITLGSHWIEPNR
KLOTB_HUMAN	Beta-klotho OS=Homo sapiens GN=KLB PE=1 SV=1	YGGWKNDTIIDIFNDYATYCFQMFGDR
KLOTB_HUMAN	Beta-klotho OS=Homo sapiens GN=KLB PE=1 SV=1	MGQNVSLNLREALNWIK
KLOTB_HUMAN	Beta-klotho OS=Homo sapiens GN=KLB PE=1 SV=1	SGNDTYGAAHNLLVAHALAWR
KNG1_HUMAN	Kininogen-1 OS=Homo sapiens GN=KNG1 PE=1 SV=2	YNSQNQSNNQFVLYR
KV104_HUMAN	Ig kappa chain V-I region CAR OS=Homo sapiens PE=1 SV=1	ASQNISSWLAWYQQKPGK
PLMN_HUMAN	Plasminogen OS=Homo sapiens GN=PLG PE=1 SV=2	GNVAVTVSGHTCQHWSAQTPHTHNR
PON1_HUMAN	Serum paraoxonase/arylesterase 1 OS=Homo sapiens GN=PON1 PE=1 SV=2	VVAEGFDFANGINISPDGKYVYIAELLAHK
THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	GHVNITR
THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	YPHKPEINSTTHPGADLQENFCR
THRB_HUMAN	Prothrombin OS=Homo sapiens GN=F2 PE=1 SV=2	NFTENDLLVRIGK
TRFE_HUMAN	Serotransferrin OS=Homo sapiens GN=TF PE=1 SV=2	ILRQQQHLFGSNVTDCSGNFCLFR

Appendix I. Matched sequence of BSA tryptic digest by 80% ACN containing 1% H_3PO_4 washing for the first time

Matched peptides shown in ***bold red***.

1	MKWVTFISLL	LLFSSAYSRG	VFRRDTHKSE	IAHR**FKDLGE**	**EHFK**GLVLIA
51	FSQYLQQCPF	DEHVK**LVNEL**	**TEFAK**TCVAD	ESHAGCEK**SL**	**HTLFGDELCK**
101	VASLRETYGD	MADCCEKQEP	ERNECFLSHK	DDSPDLPK**LK**	**PDPNTLCDEF**
151	**K**ADEK**KFWGK**	**YLYEIAR**RHP	YFYAPELLYY	ANKYNGVFQE	CCQAEDKGAC
201	LLPKIETMRE	KVLASSARQR	LRCASIQKFG	ERALK**AWSVA**	**RLSQKFPKAE**
251	**FVEVTK**LVTD	LTKVHKECCH	GDLLECADDR	ADLAK**YICDN**	**QDTISSKLKE**
301	**CCDKPLLEK**S	HCIAEVEKDA	IPENLPPLTA	DFAEDKDVCK	NYQEAK**DAFL**
351	**GSFLYEYSRR**	**HPEYAVSVLL**	**R**LAK**EYEATL**	**EECCAKDDPH**	**ACYSTVFDK**L
401	K**HLVDEPQNL**	**IKQNCDQFEK**	**LGEYGFQNAL**	**IVR**YTR**KVPQ**	**VSTPTLVEVS**
451	**R**SLGKVGTRC	CTKPESER**MP**	**CTEDYLSLIL**	**NRLCVLHEKT**	**PVSEK**VTK**CC**
501	**TESLVNRRPC**	**FSALTPDETY**	**VPKAFDEKLF**	**TFHADICTLP**	**DTEK**QIK**KQT**
551	**ALVELLK**HKP	K**ATEEQLKTV**	**MENFVAFVDK**	CCAADDKEAC	FAVEGPKLVV
601	STQTALA				

Appendix II. Matched sequence of BSA tryptic digest by 80% ACN containing 1% H_3PO_4 washing for the second time

Matched peptides shown in ***bold red***.

1	MKWVTFISLL	LLFSSAYSRG	VFRRDTHKSE	IAHR**FKDLGE**	**EHFK**GLVLIA
51	FSQYLQQCPF	DEHVK**LVNEL**	**TEFAK**TCVAD	ESHAGCEK**SL**	**HTLFGDELCK**
101	VASLRETYGD	MADCCEKQEP	ERNECFLSHK	DDSPDLPK**LK**	**PDPNTLCDEF**
151	**K**ADEKKFWGK	**YLYEIARRHP**	**YFYAPELLYY**	**ANK**YNGVFQE	CCQAEDKGAC
201	LLPKIETMRE	KVLASSARQR	LRCASIQKFG	ERALK**AWSVA**	**R**LSQKFPK**AE**
251	**FVEVTK**LVTD	LTKVHK**ECCH**	**GDLLECADDR**	ADLAK**YICDN**	**QDTISSKLKE**
301	**CCDKPLLEK**S	HCIAEVEKDA	IPENLPPLTA	DFAEDKDVCK	NYQEAK**DAFL**
351	**GSFLYEYSRR**	**HPEYAVSVLL**	**R**LAK**EYEATL**	**EECCAKDDPH**	**ACYSTVFDK**L
401	K**HLVDEPQNL**	**IK**QNCDQFEK	**LGEYGFQNAL**	**IVR**YTR**KVPQ**	**VSTPTLVEVS**
451	**R**SLGKVGTRC	CTKPESER**MP**	**CTEDYLSLIL**	**NR**LCVLHEK**T**	**PVSEK**VTK**CC**
501	**TESLVNRRPC**	**FSALTPDETY**	**VPK**AFDEK**LF**	**TFHADICTLP**	**DTEK**QIKKQT
551	ALVELLKHKP	KATEEQLKTV	MENFVAFVDK	CCAADDKEAC	FAVEGPKLVV
601	STQTALA				

References

S1. D. Jiang, L. L. Keenan, A. D. Burrows and K. J. Edler, *Chem. Commun.*, 2012, **48**, 12053-12055.