Supporting Information

Maltose-Functionalized Hydrophilic Magnetic Nanoparticles with Polymer Brushes for Highly Selective Enrichment of N-linked Glycopeptides

Changfen Bi[†], Yulu Liang[‡,§], Lijin Shen[∥], Shanshan Tian[∥], Kai Zhang[∥], Yiliang Li[*,†], Xiwen He[‡], Langxing Chen[*,‡,§], and Yukui Zhang[‡,⊥]

[†]Tianjin Key Laboratory of Radiation Medicine and Molecular Nuclear Medicine, Institute of Radiation Medicine, Peking Union Medical College & Chinese Academy of Medical Sciences, Tianjin 300192, China

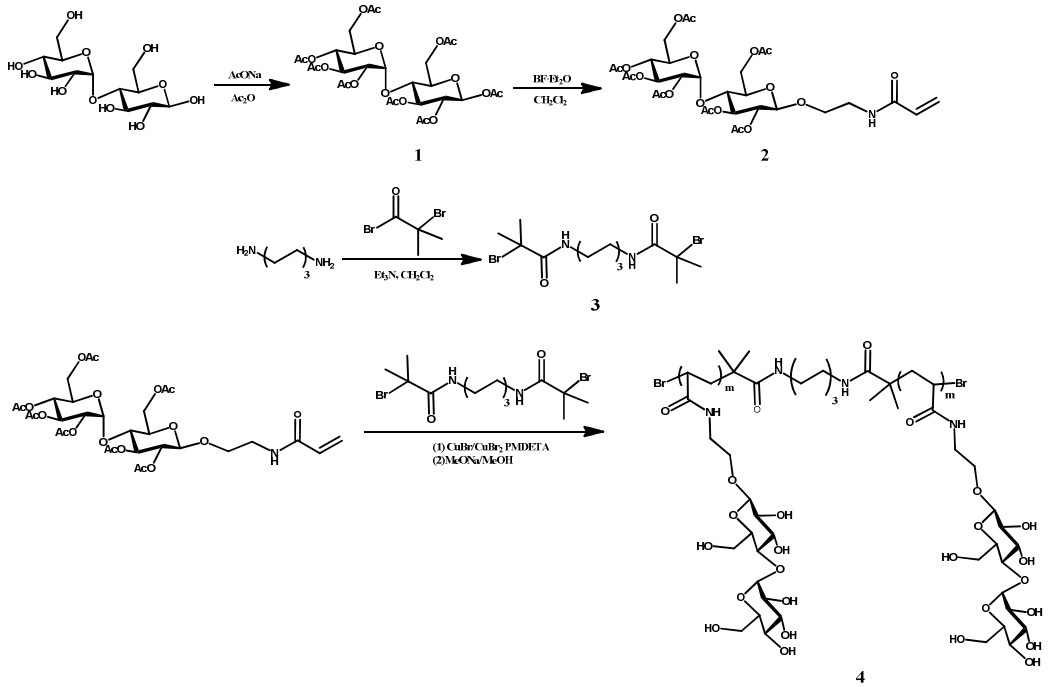
[‡]Research Center for Analytical Sciences, College of Chemistry, Tianjin Key Laboratory of Biosensing and Molecular Recognition, State Key Laboratory of Medicinal Chemical Biology, Nankai University, Tianjin 300071, China

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

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

[∥]2011 Collaborative Innovation Center of Tianjin for Medical Epigenetics, Tianjin Medical University, Tianjin 300070, China

[⊥]Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China



Scheme S1. Schematic illustration of maltose polymer brush with ATRP polymerization and off-protection group.

1. Synthesis of per-O-acetylated maltose (1)

Into a solution of D-(+)-maltose (10 g, 28.6 mmol) in acetic anhydride (140 mL) was added NaAc (10 g, 121.9 mmol). The resulting mixture was heated to 140 °C and stirred for 12 h. The reaction mixture was cooled to room temperature and water was added. The product was extracted with ethyl acetate, washed with water, brine and dried over anhydrous Na_2SO_4. The solvent were removed in vacuo and the residue was purified by recrystallization from methanol to give a white solid as product.^{1}H NMR (300 M Hz, $CDCl_3$) δ 5.73 (d, 1H), 5.40 (t, 1H), 5.33 (d, 1H), 5.27 (d, 1H), 5.05 (t, 1H), 4.87 (m, 1H), 4.85 (dd, 1H), 4.45 (dd, 1H), 4.25 (t, 1H), 4.21 (t, 1H), 4.07~4.01 (m, 2H), 3.94~3.92 (m, 1H), 3.86~3.82 (m, 1H), 2.14~2.01 (m, 24H). LR MS (ESI) m/z Calcd for $C_{28}H_{38}O_{19}$ $[M+Na]^+$:701.2, found 701.3.

2. Synthesis of (2-acrylamidoethoxy) 2,3,6-tri-O-(2,3,4,6-tetra-O-acetyl)-β-D-glucopyranosyl) -β- D-glucopyranoside (2)

This compound was prepared according to a literature[1]: Per-O-acetylated maltose[2] (3 g, 4.42 mmol) and N-(2-hydroxyethyl) acrylamide (0.76 g, 6.60 mmol) were dissolved in dry CH_2Cl_2 (40 mL). Boron trifluorideetherate (1 mL, 8.84 mmol) was added dropwise at 0 °C. Stirring was continued for 2 h at 0 °C and an additional 18 h at room temperature.The solution was poured on ice water and extracted twice with CH_2Cl_2. The combined organic phases were washed with water, brine and dried over anhydrous Na_2SO_4. The solvent were removed in vacuo and the residue was purified by column chromatography (SiO_2, 0-10 % MeOH:CH_2Cl_2) to give a white solid as product.^{1}H NMR (300 M Hz, $CDCl_3$) δ 6.29~6.23(m, 1H), 6.16~6.11 (m, 1H), 5.61 (dd, 1H), 5.42~5.19 (m, 3H), 5.06~4.99 (m, 1H), 4.86~4.75 (m, 2H), 4.55~4.47 (m, 2H), 4.25~3.45 (m, 11H), 2.11 (s, 3H), 2.08 (s, 3H), 2.01~1.97 (m, 15H). LR MS (ESI) m/z Calcd for $C_{31}H_{43}NO_{19}$ $[M+Na]^+$:756.2, found 756.4.

3. Synthesis of bis (2-bromoisobutyryl) hexanediamide (3)

Hexamethylenediamine (0.428 mg, 3.68 mmol) and Et_3N (1.22 mL, 8.82 mmol) were dissolved in dry CH_2Cl_2 (30 mL). 2-bromoisobutyryl bromide (1.69 g, 7.35 mmol) in CH_2Cl_2 (10 mL) was added slowly into the reaction mixture at 0 °C. After 2 h at 0 °C, the reaction was left to reach room temperature and stirred for 8 h. The reaction mixture was then filtered to remove all solids. After evaporation of the solvent, the residue was purified by column chromatography [SiO_2, 0-10 % ethyl acetate: petroleum (b.p. 60-90 °C)] to give a white solid as product. ^{1}H NMR (300 M Hz, $CDCl_3$) δ 6.76 (br, 2H), 3.29~3.23 (m, 4H), 1.96 (s, 12H),

1.56~1.52 (m, 4H), 1.40~1.35 (m, 4H).

4. Synthesis of maltose polymer brush (4) by atom transfer radical polymerization (ATRP)

(2-acrylamido) ethoxy-2,3,6-tri-*O*-(2,3,4,6-tetra-*O*-acetyl)-*β*-D-glucopyranosyl) -*β*-D-glucopyranoside (1.1 g, 20 eq), bis(2-bromoisobutyryl) hexanediamide (30 mg, 1 eq), PMDETA (39.2 μL) were added to a 50 mL Schlenk flask. The mixture was deoxygenated under vacuum and backfilled with argon three times. $CuBr/CuBr_2$ (20 mg) was added to the mixture and three vacuum-argon cycles were performed. The reaction system was stirred at 60 °C for 24 h. Low-resolution mass spectrometry revealed that raw material has reacted. The solvent was evaporated under reduced pressure, and the residue was dissolved in ethyl acetate (150 mL). The organic phase was washed with saturated brine (40 mL×3) and then dried over anhydrous Na_2SO_4. The crude mixture was filtered and concentrated. After the residue was dissolved in dry MeOH, 0.596 M NaOMe/MeOH was added at 0 °C. The resulting mixture was allowed to stir at room temperature 1 h. The reaction mixture was neutralized by addition of Amberlite IR 120 H^+ until pH 6 and filtered off the resin. The filtrate was concentrated under vacuum to give the product as a white solid. ^{1}H NMR (300 MHz, CD_3OH) δ 8.14 (s, 1H), 5.16~5.10 (m, 2H), 3.82~3.80 (m, 8H), 3.67~3.57 (m 10H), 3.46~3.35 (m, 6H), 3.28~3.26 (m, 2H), 1.96 (s, 1H), 1.90 (s, 1H).

5. 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).

6. LC-MS/MS analysis

The deglycosylated peptides enriched from HRMC tryptic digest was desalted using a μ-C18 Ziptip and dissolved in 10 μL of HPLC buffer A (0.1 % (v/v) formic acid in water). 5μL 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 CH_3CN) 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.

7. 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 asparagines deamination were fixed as variable modifications. Carbamidomethylation on Cys was specified as fixed modification.

References

[1] Belardi B.; O'Donoghue G. P.; Smith A. W.; Groves J. T.; Bertozzi C. R. Investigating Cell Surface Galectin-Mediated Cross-Linking on Glycoengineered Cells. *J. Am. Chem. Soc*. **2012**, 134, 9549-9552.

[2] Virgil P.; Pawaret L.; Sun H. J.; Oleg K.; Christopher D. N.; Tam M. T.; Annabelle B.; Daniela A. W.; Mihai P.; Zhang S. D.; Neha P. K.; Kevin V.; Diana M.; Eric D. J.; Daniel A. H.; Darrin J. P.; Chen Y. C.; Yoann M. C.; Tze C. S.; Milan B.-B.; Sabine A.; René R.; Hans-J. G.; Paul A. H. Modular Synthesis of Amphiphilic Janus Glycodendrimers and Their Self-Assembly into Glycodendrimersomes and Other Complex Architectures with Bioactivity to Biomedically Relevant Lectins. *J. Am. Chem. Soc*. **2013**, 135, 9055-9077.

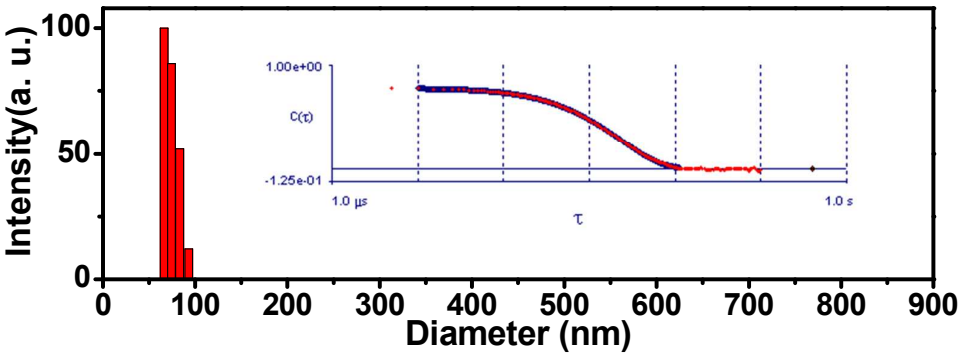


Figure S1 The hydrodynamic diameter histograms of Fe_3O_4-PEI-pMaltose NPs

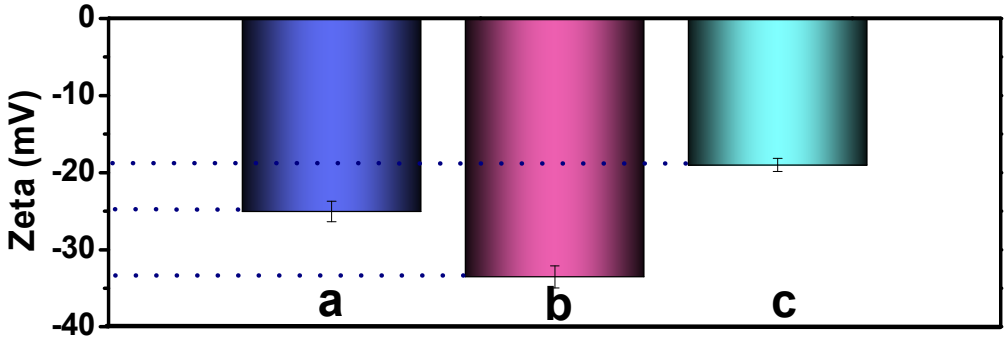


Figure S2 Zeta potential of Fe_3O_4-PEI (a), Fe_3O_4-PEI-COOH (b) and Fe_3O_4-PEI-pMaltose (c) NPs at pH 10.5.

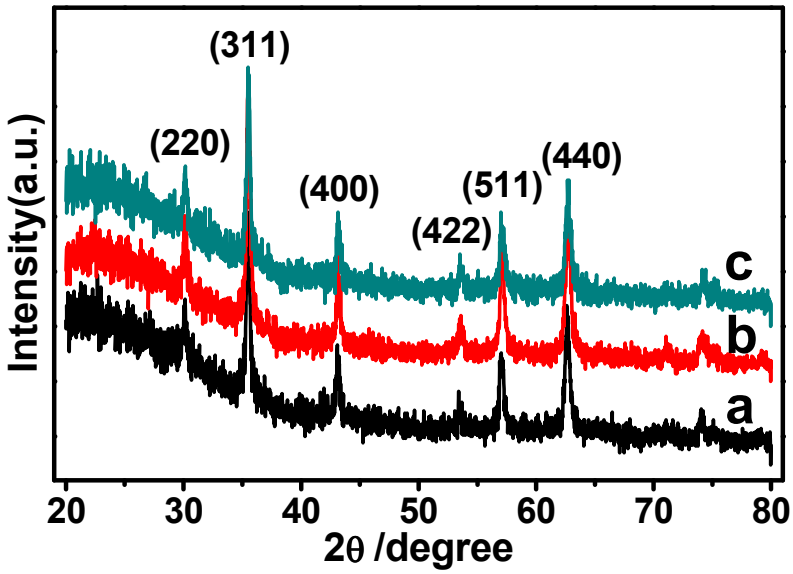


Figure S3 XRD patterns of Fe_3O_4-PEI (a), Fe_3O_4-PEI-COOH (b) and Fe_3O_4-PEI-pMaltose (c) NPs.

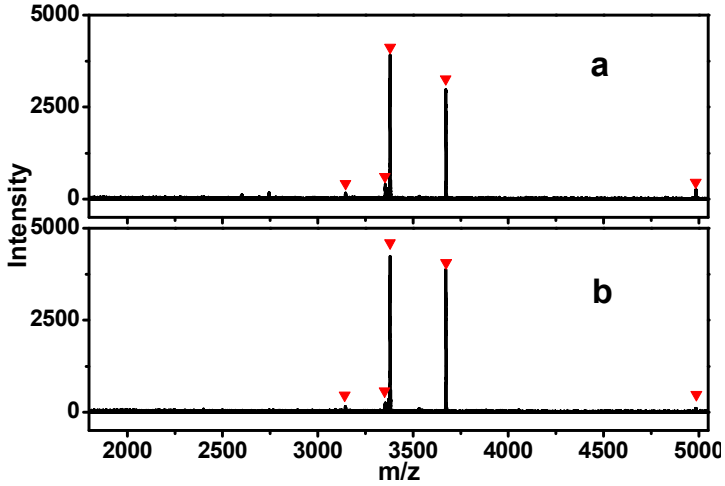


Figure S4 MALDI-TOF mass analysis of the glycopeptidesderived from HRP tryptic digest: (a) after treatment with Fe_3O_4-PEI-pMaltose NPs in the first time run, (b) after treatment with Fe_3O_4-PEI-pMaltose NPs in the sixth time run. The peaks of glycopeptides are marked with▼.

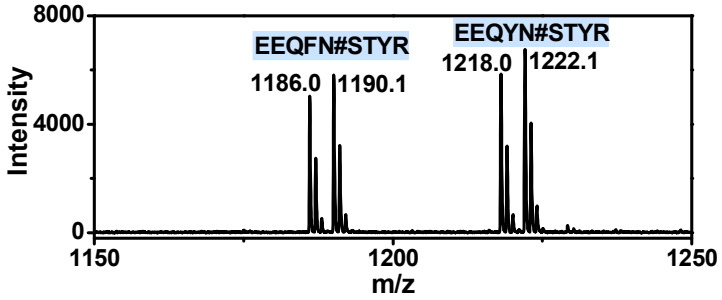


Figure S5 MALDI-TOF MS spectra of the isotope dimethylation labeled human IgG peptides after deglycosylation by PNGase F.

Table S1 Observed glycopeptides and glycan structures of human HRP digests enriched by Fe_3O_4-PEI-pMaltose NPs. The N-glycosylation sites are marked with N#.

Peak number	Observed m/z	Glycan composition	Amino acid sequence
1	2533.1	[HexNAc]1[Fuc]1	SFAN#STQTFFNAFVEAMDR
2	2612.2	[Hex]3[HexNAc]2[Xyl]1	MGN#ITPLTGTQGQIR
3	2851.6	[HexNAc]1[Fuc]1	GLIQSDQELFSSPN#ATDTIPLVR
4	3089.4	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	GLCPLNGN#LSALVDFDLR
5	3146.4	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	GLCPLNGN#LSALVDFDLR.Oxide
6	3189.2	[Hex]2[HexNAc]2[Fuc]1[Xyl]1	SFAN#STQTFFNAFVEAMDR
7	3209.6	[Hex]3[HexNAc]2[Xyl]1	SFAN#STQTFFNAFVEAMDR
8	3223.3	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	QLTPTFYDNSCPN#VSNIVR
9	3323.2	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	QLTPTFYDNSCPN#VSNIVR
10	3355.3	[Hex]2[HexNAc]2[Fuc]1[Xyl]1	SFAN#STQTFFNAFVEAMDR
11	3371.5	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	SFAN#STQTFFNAFVEAM*DR
12	3380.3	[Hex]2[HexNAc]2[Fuc]1	GLIQSDQELFSSPN#ATDTIPLVR
13	3527.7	[Hex]3[HexNAc]2[Xyl]1	GLIQSDQELFSSPN#ATDTIPLVR
14	3541.4	[Hex]3[HexNAc]2[Fuc]1	GLIQSDQELFSSPN#ATDTIPLVR
15	3673.6	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	GLIQSDQELFSSPN#ATDTIPLVR
16	3896.8	[Hex]3[HexNAc]2[Fuc]1[Xyl]1	LHFHDCFVNGCDASILLDN#TTSFR
17	4058.4	[Hex]3[HexNAc]2[Xyl]1	QLTPTFYDNSC(AAVESACPR)PN# VSNIVR-H_2O
18	4839.7	[Hex]3[HexNAc]2[Fuc]1[Xyl]1 [Hex]3[HexNAc]2[Xyl]1	LYN#FSNTGLPDPTLN#TTYLQTLR
19	4986.1	[Hex]3[HexNAc]2[Fuc]1[Xyl]1 [Hex]3[HexNAc]2[Fuc]1[Xyl]1	LYN#FSNTGLPDPTLN#TTYLQTLR

Hex=mannose, HexNAc= N-acetylglucosamine, Fuc= fucose, Xyl= xylose.

Table S2 Observed glycopeptides and glycan structures of human IgG digests enriched by Fe_3O_4-PEI-pMaltose NPs. The N-glycosylation sites are marked with N#.

Peak number	Observed m/z	Glycan composition	Amino acid sequence
1	2602.8	[Hex]3[HexNAc]4[Fuc]1	EEQFN#STFR
2	2634.7	[Hex]3[HexNAc]4[Fuc]1	EEQYN#STYR
3	2651.2	[Hex]4[HexNAc]4	EEQYN#STYR
4	2764.7	[Hex]4[HexNAc]4[Fuc]1	EEQFN#STFR
5	2796.7	[Hex]4[HexNAc]4[Fuc]1	EEQYN#STYR
6	2806.1	[Hex]3[HexNAc]5[Fuc]1	EEQYN#STFR
7	2819.8	[Hex]4[HexNAc]5	EEQFN#STFR
8	2837.7	[Hex]3[HexNAc]5[Fuc]1	EEQYN#STYR
9	2926.8	[Hex]5[HexNAc]4[Fuc]1	EEQFN#STFR
10	2958.7	[Hex]5[HexNAc]4[Fuc]1	EEQYN#STYR
11	2967.8	[Hex]4[HexNAc]5[Fuc]1	EEQFN#STFR
12	2982.2	[Hex]5[HexNAc]5	EEQFN#STFR
13	2999.8	[Hex]4[HexNAc]5[Fuc]1	EEQYN#STYR
14	3014.1	[Hex]5[HexNAc]5	EEQYN#STYR
15	3162.6	[Hex]5[HexNAc]5[Fuc]1	EEQYN#STYR
16	3217.8	[Hex]5[HexNAc]4[Fuc]1[NeuAc]1	EEQFN#STFR
17	3250.7	[Hex]5[HexNAc]4[Fuc]1[NeuAc]1	EEQYN#STYR

Hex= mannose, HexNAc= N-acetylhexosamine, Fuc= fucose, NeuAc= N-acetylneuraminic acid.

Table S3 The recovery yield of N-glycosylation motifs from tryptic digests of IgG enriched by Fe_3O_4-PEI-pMaltose NPs. L and H indicate the light- and heavy-tagged N-linked glycopeptide motifs, respectively.

Item	Peptide sequence	Recovery 1	Recovery 2	Recovery 3
L/H	EEQFN#STYR EEQYN#STYR	88.0%	92.8%	89.7%
H/L	EEQFN#STYR EEQYN#STYR	84.5%	89.0%	84.5%

Table S4 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-PEI-pMaltose NPs.

N# denotes the N-linked glycosylation site.

No	Protein Group Accessions	Peptide sequence
1	P22680	LSSASLN#IRTAKEDFTLHLEDGSYNIR
2	Q8NDV3-2	SVETLLN#QK
3	O60814	AMGIMN#SFVNDIFER
4		KEN#LSR
5	Q71DI3	FQSSAVMALQEASEAYLVGLFEDTN#LCAIHAK
6	Q9ULV0	N#PDILVGENDLTALSYLHEPAVLHN#LK
7	Q9UKT9-16	QN#HIYQQN#HMVLSR
8	Q8N6C5-2	GEDKPLQFLDATSIDDN#TSFFLNNVTYSDTGIYSCHYLLTWK
9	P68431	FQSSAVMALQEACEAYLVGLFEDTN#LCAIHAK
10	P82987-2	KPISWRHCLGPSCDSTYTSQTATN#K
11	H7C3P4	YPHNHHVVN#N#TLEGN#CSSK
12	E5RI99	TGVHHYSGN#NIELGTACGK
13	Q7L2E3-3	DKLVYVHTN#GPK
14	E9PEV0	LNMHMNVQN#GK
15	B4E241	VYVGNLGNN#GNK
16	F8VVM2	IQTQPGYAN#TLR
17	K7EM20	YLAEFATGN#DRK
18	V9GYN5	LGIYADVGN#K
19	G5E9G0	AHLMEIQVN#GGTVAEK
20	Q9Y5B9	GDKVDILYN#NIK
21	F5H4G7	KEAAWAITN#ATSGGTPEQIR
22	O60337-6	TVIEQVYAN#GIR
23	P13591-4	DGQLLPSSN#YSNIK
24	H0YEN5	AFVAIGDYN#GHVGLGVK
25	P11388	GFQQISFVN#SIATSK
26	Q09666	ISMPDIDLN#LK
27	Q96C36	IIASSPEMN#LPTVSALR
28	P06576	TVLIMELIN#NVAK
29	P13674-3	DMSDGFISN#LTIQR
30	P26599	AQAALQAVN#SVQSGNLALAASAAAVDAGMAMAGQSPVLR
31	Q9H9B4	SGDAPLTVN#ELGTAYVSATTGAVATALGLNALTK
32	P02751-2	HEEGHMLN#CTCFGQGR
33	P09382	FNAHGDAN#TIVCNSK
34	P34932	KFDEVLVN#HFCEEFGK
35	P00505	HFIEQGIN#VCLCQSYAK

36	H0YIM7	TTASTQMN#VQSSR
37	B4DWL3	ALQATVGN#SYK
38	K7EMG4	ASVDVFGN#R
39	P60709; P68032	IWHHTFYN#ELR
40	Q9NTI5-2	TLFSVINN#GHNQK
41	Q9H8M5-3	VYGQNINN#ETWSR
42	Q07954	LYWISSGN#HTINR
43	O75533	VPELNVQN#GVLK
44	Q07954	KLNLDGSN#YTLLK
45	Q02750-2	ISELGAGN#GGVVFK
46	B1ALC0	GFESPSDN#SSAMLLQWHEK
47	Q99536	VVTYGMAN#LLTGPK
48	P17480-2	FSQELLSN#GELNHLPLK
49	B4DJV2	DYIWNTLN#SGR
50	P30048-2	GLFIIDPN#GVIK
51	P07437	MAVTFIGN#STAIQELFK
52	P46977	TILVDNN#TWN#NTHISR
53	P02790	SWPAVGN#CSSALR
54	P16615	TGTLTTN#QMSVCR
55	O15230	VWQGHN#ASFCPHGYGCR
56	Q14974	VLANPGN#SQVAR
57	K7EP71	YHVLHIN#TTK
58	O60568	EQYIHEN#YSR
59	Q5VW52	IIEGHYN#GEQLGKPK
60	P43243	GNLGAGN#GNLQGPR
61	P29317	TASVSIN#QTEPPK
62	Q9HD45	KLEIGFN#GNR
63	P38646	QAVTNPN#NTFYATK
64	P62269	YSQVLAN#GLDNK
65	Q14697	SLLLSVN#AR
66	Q9H501	DVASEVN#LTAYKPK
67	P62158	VFDKDGN#GYISAAELR
68	P50542-2	LGATLAN#GNQSEEAVAAYR
69	P05023-3	IVEIPFN#STNK
70	P11166	VIEEFYN#QTWVHR
71	P22626	GFGDGYN#GYGGGPGGGNFGGSPGYGGGR
72	P62241	IIDVVYN#ASNNELVR
73	P11142	LLQDFFN#GK
74	Q09666	ISMPDVN#LNLK
75	Q99650-2	SVNILFN#LTHR
76	Q9Y5L4	VQIAVAN#AQELLQR
77	C9J931	ALAESWN#AAFLESSAK
78	P62316-2	REEEEFN#TGPLSVLTQSVK

79	P08133-2	SEIDLLN#IR
80	Q9Y6M1-1	TVNELQN#LTSAEVIVPR
81	O75165	FLATPPN#GNFADAVFR
82	O94874-3	TYDLPGN#FLTQALTQR
83	P05141; P12235	GLYQGFN#VSVQGIIIYR
84	P61006	FSEDAFN#STFISTIGIDFK
85	Q6DRA6	AMGIMN#SFLN#DIFER
86	B1AK13	LLEAGN#FICQALNR
87	P50395	NTNDAN#SCQIIIPQNQVNR
88	B4DEH8	SIYVGN#VDYGATAEELEAHFHGCGSVN#R
89	P08559	NFYGGN#GIVGAQVPLGAGIALACK
90	O75477	IYFGSN#IPNMFVDSSCALK
91	Q08945	ITVPGN#FQGHSGAQCITCSYK
92	P49746	LGFLGN#QSQGCLPAR
93	P49411	DKPHVN#VGTIGHVDHGK
94	P20700	LSSEMN#TSTVNSAR
95	P09874	KPPLLN#NADSVQAK
96	O60911	MIELHN#GEYSQGK
97	Q9H173	LINKFN#SSSSSLEEK
98	O94826	QAYTGN#NSSQIQAAMK
99	P67809	RPQYSN#PPVQGEVMEGADNQGAGEQGRPVR
100	O15173	ILLAVN#GK
101	P08238	KHLEIN#PDHPIVETLR
102	Q9UQ80	LVKPGN#QNTQVTEAWNK
103	Q07954	GVTHLN#ISGLK
104	E9PJ35	IAPASN#VSHTVVLRPLK
105		FASFIN#KVR
106	E9PL22	VFGSQN#LTTVK
107	P04264; P35908	FLEQQN#QVLQTK
108	J3KPX7	VLSRPN#AQELPSMYQR
109	P14866-2	AITHLN#NNFMFGQK
110	P20700	SLETEN#SALQLQVTER
111	Q07666-3	KDDEEN#YLDLFSHK
112	Q5T4L4	LVQSPN#SYFMDVK
113	V9GYF5	SVSTTN#IAGHFNDESPLGLR
114	Q16706	LLAENN#EIISNIR
115	Q9BXT5	ISPIDN#YIVLHQEYK
116	P39656	YSQTGN#YELAVALSR
117	O00468-6	NELMLN#SSLMR
118	Q14669-4	YFSSEN#MDGSNPALNVLQR
119	Q9Y5B9	GNENAN#GAPAITLLIR
120	B7ZLZ2	LLPAFN#TTSGLPYPR
121	Q6ZMG9	FWLPHN#VTWADLK

122	O60568	IFQNLN#GALDEVVLK
123	B4DDS3	DYYPIN#ESLASLPLR
124	E9PL22	VLQLIN#DNTATALSYGVFR
125	P04843	VHYEN#N#SPFLTITSMTR
126	H7C3P4	TPMTN#SSIQFLDN#AFR
127	P60709	DLYAN#TVLSGGTTMYPGIADR
128	P52907	DHYSN#GFCTVYAK
129	P12235	YFAGN#LASGGAAGATSLCFVYPLDFAR
130	P04040	LVNAN#GEAVYCK
131	Q07954	WTGHN#VTVVQR
132	J3KQE0	ALSPN#STISSAPK
133	P14625	DISTN#YYASQK
134	Q8N766-4	FINYN#QTVSR
135	P57105	ILSVN#GQDLK
136	O14773-2	LFGGN#FAHQASVAR
137	P13929-3	LAQSN#GWGVMVSHR
138		WDTAN#NPLYK
139	P61106	KIYQN#IQDGSLDLNAAESGVQHKPSAPQGGR
140	Q9Y277	LSQNN#FALGYK
141	Q9BX74	LSITN#ETFR
142	Q8WVX9	IIAIN#SELTQPK
143	Q13586-2	LAVTN#TTMTGTVLK
144	Q9Y5B9	DLGFN#GAPYR
145	P08238; P07900	ADLIN#NLGTIAK
146		AVIDN#IR
147	P62805	VFLEN#VIR
148	Q13637	VHLPN#GSPIPAVLLANK
149	P04843	VHYEN#NSPFLTITSMTR
150	E9PMX3	GEFYN#EASNLQVAIR
151	P60709	DLYAN#TVLSGGTTMYPGIADR
152	P49748-2	LWISN#GGLADIFTVFAK
153	P61106	IYQN#IQDGSLDLN#AAESGVQHKPSAPQGGR
154	Q8N6C5-2	GIGN#YSCSYR
155	P55072	AIAN#ECQAN#FISIK
156	P14866-2	ASLN#GADIYSGCCTLK
157	Q8WVV9-4	AALN#GADIYAGCCTLK
158	P08648	HPGN#FSSLSCDYFAVN#QSR
159	P16615	YETN#LTFVGCVGMLDPPR
160	Q96AY3	YHYN#GTFEDGKK
161	P11021	EFFN#GKEPSR
162	Q96AY3	YHYN#GTFEDGK
163	P04406	VGVN#GFGR
164	Q9Y2C2	KLGN#MTVTVK

165	Q96C19	ADLN#QGIGEPQSPSR
166	Q07954	MHLN#GSNVQVLHR
167		ALLN#QLR
168	E9PKZ0	ASGN#YATVISHNPETK
169	P62424	AGVN#TVTTLVENK
170	Q9Y276	RPLN#SVVLQQGLADR
171	P14625	LIIN#SLYK
172	Q00610-2	LLYN#NVSNFGR
173	Q9NS18	IFVN#GTFIGGATDTHR
174	Q09666	VDIN#APDVDVHGPDWHLK
175	F5H801	SSEN#GVDYVIMGMPHR
176	P11142	TVTN#AVVTVPAYFNDSQR
177	P68363; Q9BQE3	VGIN#YQPPTVVPGGDLAK
178	B7Z230	YHYN#GTFLDGTLFDSSHNR
179	P08574	AANN#GALPPDLSYIVR
180	Q96AY3	YHYN#GSLMDGTLFDSSYSR
181	Q6P2Q9	GPGN#PVPGPLAPLPDYMSEEK
182	O75976	GILN#ATISVAEINHPVTTYK
183	B7Z230	YHYN#ASLLDGTLLDSTWNLGK
184	P06576	EGN#DLYHEMIESGVINLK
185	Q8NBJ5	AMN#TSQVEALGIQMLPGYR
186	P05556	KEN#SSEICSN#NGECVCGQCVCR
187	P05556	KEN#SSEICSNNGECVCGQCVCR
188	P09382	DSN#NLCLHFNPR
189	P61224; F5GZG1	QWN#NCAFLESSAK
190	P09874	SAN#YCHTSQGDPIGLILLGEVALGNMYELK
191	Q92626	VNN#CSDMCFHQK
192	P55072	MTN#GFSGADLTEICQR
193	P38646	ASN#GDAWVEAHGK
194	Q9UJ14	TAN#HSAPSLENSVQPGK
195	H7BZF2	LHN#AIEGGTQLSR
196	J3QRP6	GPN#GYGFHLHGEK
197	P08670	FAN#YIDK
198	B4DFA2	VVN#GPEILNK
199	Q08380	GLN#LTEDTYKPR
200	H0YFI1	ALN#GAEPNYHSLPSAR
201	P69849	LEN#ITTGTYTIHAQK
202	P14866-2	NPN#GPYPYTLK
203	P02545-6	NSN#LVGAAHEELQQSR
204	P62244	WQN#NLLPSR
205	Q8NBK3-2	FVN#STGYLTEAEK
206	Q07954	MTN#GAMNVEIGNPTYK
207	J3QLE5	GEN#LVSMTVEGPPPK

208	P08559	MVN#SNLASVEELK
209	P45880-2	VNN#SSLIGVGYTQTLRPGVK
210	P54886-2	NLN#GTLHELLR
211	P21796	VNN#SSLIGLGYTQTLKPGIK
212	P14625	HNN#DTQHIWESDSNEFSVIADPR
213	P31943	SNN#VEMDWVLK
214	P23284	DTN#GSQFFITTVK
215	P26006	VWN#STFIEDYR
216	Q01081-4	WFN#GQPIHAELSPVTDFR
217	Q70UQ0-4	FQN#ITDFWK
218	C9JRH2	DNN#GVIGLLEPMK
219	Q99798	HPN#GTQETILLNHTFNETQIEWFR
220	P06576	EGN#DLYHEMIESGVINLK
221	Q8NBJ5	AMN#TSQVEALGIQMLPGYR
222	P63244	GHN#GWVTQIATTPQFPDMILSASR
223	P50897-2	FLN#DSIVDPVDSEWFGFYR
224	Q71DI3	FQSSAVMALQEASEAYLVGLFEDTN#LCAIHAK
225	P13929-3	HIADLAGNPDLILPVPAFNVIN#GGSHAGNK
226	P07437; P68371	SGPFGQIFRPDNFVFGQSGAGN#NWAK
227	O95831-3	SIYFQPPSFYVSAQDLPHIEN#GGVAVLTGK
228	P10809	KPLVIIAEDVDGEALSTLVLN#R
229	P10809	QSKPVTTPEEIAQVATISAN#GDK
230	Q5T6W5	IITITGTQDQIQNAQYLLQN#SVK
231	P78357	GLLTFVDHLPVTQVVIGDTN#R
232	H0YGE9	LN#ETHIFN#GSNWIMLIYK
233	Q92896	GN#ITEYQCHQYITK
234	P05556	EN#SSEICSN#NGECVCGQCVCR
235	P05556	EN#SSEICSNNGECVCGQCVCR
236	P22087	AN#CIDSTASAEAVFASEVK
237	Q7Z7K6-3	SN#PGGFGIAPHCLDEGTVR
238	Q9Y639-3	AN#ATIEVK
239	J3KS10	AN#GTTVHVGIHPSK
240	P18084	SN#LTVLR
241	Q9UJ14	RN#ESHLIDFR
242	Q99519	VN#LTLR
243	Q6DKI1	IN#QLIR
244	P06756-3	AN#TTQPGIVEGGQVLK
245	Q07954	FN#STEYQVVTR
246	P12694	NN#GYAISTPTSEQYR
247	P62249	VN#GRPLEMIEPR
248	P51149	NN#IPYFETSAK
249	P62851	LN#NLVLFDK
250	Q9Y5M8	GN#SLTLIDLPGHESLR

251	P31930	NN#GAGYFLEHLAFK
252	H0Y6T7	AN#NSWFQSILR
253	P26599	NN#QFQALLQYADPVSAQHAK
254	O75629	LN#ITNIWVLDYFGGPK
255	P68104	VETGVLKPGMVVTFAPVN#VTTEVK
256	P62937	HTGPGILSMANAGPNTN#GSQFFICTAK
257	Q9H488	EGNPFGPFWDQFHVSFN#K
258	K7EM56	DMIILPEMVGSMVGVYN#GK
259	E5RJU9	VLTDSGSLDSTIPGIEN#TITVTTEQLTTASFPVGSK
260	Q13148	FGGNPGGFGNQGGFGN#SR
261	P60709	VAPEEHPVLLTEAPLN#PK
262	Q15293	VVRPDSELGERPPEDN#QSFQYDHEAFLGK
263	O43707	LSGSNPYTTVTPQIIN#SK
264	O96019-2	SEASLHPVLMSEAPWN#TR
265	P53621	SGAWDESGVFIYTTSN#HIK
266	Q8TCJ2	VMSWWDYGYQIAGMAN#R
267	P46977	VMSWWDYGYQITAMAN#R
268	P55072	IHVLPIDDTVEGITGN#LFEVYLKPYFLEAYRPIR
269	O95470	TPEIVAPQSAHAAFN#K
270	O95477	YPTPGEAPGVVGNFN#K
271	Q53EP0	VQDILSGIEKPQVSN#IQAR
272	P07954-2	VAALTGLPFVTAPN#KFEALAAHDALVELSGAMN#TTACSLMK
273	P07954-2	VAALTGLPFVTAPN#KFEALAAHDALVELSGAMNTTACSLMK
274	Q9UBV2	LASEQQHSAQAMFN#LGYMHEK
275	P21796	KLETAVNLAWTAGN#SNTR
276	Q9Y5B9	APGEQTVPALNLQN#AFR
277	P60709; P68032	SYELPDGQVITIGN#ER
278	P38646	GVPQIEVTFDIDAN#GIVHVSAK
279	Q14974	LLETTDRPDGHQN#NLR
280	O14828-2	TAAANAAAGAAEN#AFR
281	P22626	GGGGNFGPGPGSN#FR
282	P11021	VTHAVVTVPAYFN#DAQR
283	P61106	IYQNIQDGSLDLN#AAESGVQHKPSAPQGGR
284	M0R116; P05023-3	QGAIVAVTGDGVN#DSPALK
285	Q9UBV2	MYSEGSDIVPQSN#ETALHYFK
286	Q00610-2	VGEQAQVVIIDMN#DPSNPIR
287	P55145	DRDVTFSPATIEN#ELIK
288	P05387	YVASYLLAALGGN#SSPSAK
289	Q9Y224	HILGFDTGDAVLN#EAAQILR
290	Q96CW1-2	RNELFLDVLESVN#LLMSPQGQVLSAHVSGR
291	P07437; P68371	SGPFGQIFRPDN#FVFGQSGAGN#NWAK
292	P38646	GAVVGIDLGTTN#SCVAVMEGK
293	Q9BUV8-4	RKEEPPQPQLAN#GALK

294	P08670	ISLPLPNFSSLN#LR
295	O60716-24	HYEDGYPGGSDN#YGSLSR
296	Q6P2Q9	HGDEIITSTTSN#YETQTFSSK
297	Q9ULH0-3	GHADIVHLLLQN#GAK
298	O75947-2	KYPYWPHQPIEN#L
299	P11498	IVGDLAQFMVQN#GLSR
300	P07942	MEMPSTPQQLQN#LTEDIR
301	Q12931-2	YSNFVSFPLYLN#GR
302	P16615	TVEEVLGHFGVN#ESTGLSLEQVK
303	I3NI02	ILDIANMLGLSN#TVMR
304	H3BQA8	YLASGAIDGIIN#IFDIATGK
305	F5H4J2	SPSPSSPAAVN#HHSSSDISPVSN#ESTSSSPGK
306	P63244	TNHIGHTGYLN#TVTVSPDGSLCASGGK
307	O15230	EQVLPAGQIVN#CDCSAAGTQGNACR
308	P18850	DHLLLPATTHN#K
309	E7EP04	LHPDKNPNNPN#AHGDFLK
310	P31943; P55795	HTGPNSPDTAN#DGFVR
311	Q9UBS4	KGEGLPNFDNN#NIK
312	E9PCX7	MATQASTLYSN#NITK
313	P23284	HYGPGWVSMAN#AGK
314	P68363; Q9BQE3	NLDIERPTYTN#LNR
315	Q15717	SEAEEAITSFN#GHKPPGSSEPITVK
316	B4DQ14	NVNNEVHFFEN#NNFNTIANK
317	P47755	EGAAHAFAQYN#LDQFTPVK
318	P14625	EEEAIQLDGLN#ASQIR
319	D6RIU4	VFPYISVMVNN#GSLSYDHSK
320	R4GMR2	AASNIIFSNGN#LDPWAGGGIR
321	P55145	DVTFSPATIEN#ELIK
322	P04908; Q16777	VTIAQGGVLPN#IQAVLLPK
323	P26599	IAIPGLAGAGN#SVLLVSNLNPER
324	Q9BT09	VVMDIPYELWN#ETSAEVADLKK
325	H7C3P4	YPHNHHVVNN#TLEGN#CSSK
326	Q8TCJ2	TTLVDNNTWN#N#SHIALVGK
327	P53621	LGEVALLQGN#HQIVEMCYQR
328	P07437; P68371	LTTPTYGDLN#HLVSATMSGVTTCLR
329	P53621	LGEVALLQGN#HQIVEMCYQR
330	Q9NZM1-2	HSATTVFGAN#TPIVSCNFDR
331	P26599	LSLDGQNIYN#ACCTLR
332	E5RGV5	HAAAALAAMN#GR
333	P01130-2	LHSISSIDVN#GGNRK
334	Q9BXJ5	ILMNEGGHYN#ASSGK
335	P01130-2	LHSISSIDVN#GGNR
336	P31930	EVESMGAHLN#AYSTR

337	G5E9G0	KAHLMEIQVN#GGTVAEK
338	Q9Y5B9	MQGSLEAHVN#GFR
339	P26006	TSIPTINMEN#K
340	O43909	KSDTQNLLYN#VSTGR
341	P46977	TILVDNNTWN#NTHISR
342	P11021	SQIFSTASDN#QPTVTIK
343	P20700	IESLSSQLSN#LQK
344	B4DZI8	VWQLGSSSPN#FTLEGHEK
345	Q8TCJ2	TTLVDNNTWN#NSHIALVGK
346	P45880-2; Q5JSD2	LTFDTTFSPN#TGK
347	F5H6I7	SMLQATAEAN#NLAAAASAK
348	Q92922	KVEHEISEGN#VATAAAAALASAATK
349	P40926	FVFSLVDAMN#GK
350	O60814	AMGIMNSFVN#DIFER
351	O14957	LILDWVPYIN#GK
352	F2Z389	N#ISGHIYN#QNVSQK
353	P62314	N#GTQVHGTITGVDVSMN#THLK
354	P22626	N#MGGPYGGGN#YGPGGSGGSGGYGGR
355	Q13740-2	N#ATVVWMK
356	P14625	N#LLHVTDTGVGMTR
357	H0YKJ9	N#QSVPLSCCR
358	Q14974	N#YAEYQVCLAAVGLVGDLCR
359	Q14118	N#CSTITLQN#ITR
360	F5H4F9	N#CLTNFHGMDLTR
361	P62241	N#CIVLIDSTPYR
362	P14854	N#CWQNYLDFHR
363	E9PND2	N#LDSTTVAVHGEEIYCK
364	P31930	N#ALVSHLDGTTPVCEDIGR
365	M0R061	N#AFYIGSYQQCINEAQR
366	P68104	N#MITGTSQADCAVLIVAAGVGEFEAGISK
367	P09651-3	N#QGGYGGSSSSSSYGSGR
368	O00159-2	N#GHLAVVAPR
369	F2Z389	N#ISGHIYNQNVSQK
370	Q07954	N#STTLVMHMK
371	H7C084	N#GQVIGIGAGQQSR
372	P67809; P16989-2	N#GYGFINR
373	Q8IWA5-3	N#DGSAERPYFMSSTLK
374	B4DT72	N#YLHYSLYDQAEK
375	P22626	N#MGGPYGGGNYGPGGSGGSGGYGGR
376	Q13740-2	N#ATVVWMK
377	P62314	N#GTQVHGTITGVDVSMNTHLK
378	Q00839-2	N#GQDLGVAFK
379	Q14573	N#GGGVGDILR

380	O75367-2	N#GPLEVAGAAVSAGHGLPAK
381	P50897-2	N#HSIFLADINQER
382	Q3T906	N#LSHLPTFSSPAIESHIHR
383	Q9UHD9	N#PEISHLLNNPDIMR
384	Q00839-2	N#FILDQTNVSAAAQR
385	P11142	N#QVAMNPTNTVFDAK
386	X6RDD7	N#GSFQPTNEMMLK
387	Q9NVP1	N#GTGVLILSPTR
388	E5RHF4	N#GGVVTTAFYDPR
389	P11142	N#SLESYAFNMK
390	A8MXL6	N#GGQILIADLR
391	P38646	N#AVITVPAYFNDSQR
392	P07437; P68371; Q9BUF5	N#SSYFVEWIPNNVK
393	P54819-3	N#GFLLDGFPR
394	H3BND8	N#SSLAEFVQSLSQTMGFPQDQIR
395	O14656	N#LIDYFVPFLPLEYK
396	P23284	HYGPGWVSMAN#AGK
397	Q71DI3	FQSSAVMALQEASEAYLVGLFEDTN#LCAIHAK
398	P68431	FQSSAVMALQEACEAYLVGLFEDTN#LCAIHAK
399	O60814	AMGIMN#SFVNDIFER
400	Q8TCJ2	VMSWWDYGYQIAGMAN#R
401	Q86Y82	LMNDFSAALN#NFQAVQR
402	O60814	AMGIMNSFVN#DIFER
403	Q15233	MGQMAMGGAMGIN#NR
404	Q5JRX3-3	MTAQELAN#GIPDSGHLYASIR
405	Q07954	MHLN#GSNVQVLHR
406	H7C3P4	MLVANIDLGPTILDIAGYDLN#K
407	Q7Z304	MTTGYCIECDFEENHLCGFVNRWN#PN#VN#WFVGGGSIR
408	Q16643	VASASAGSCDVPSPFN#HRPGSHLDSHR
409	P34897-3	AALEALGSCLN#NK
410	C9J9K3	YVDIAIPCN#NK
411	P11717	MSVINFECN#K
412	P50148; K7EL62	IIYSHFTCATDTEN#IR
413	P28288-2	SGANVLICGPN#GCGK
414	B7Z658	SFSGVLDCGN#CSR
415	R4GN99	VIPSFMCQAGDFTNHN#GTGGK
416	O75131	EALAQCVLAEIPQQVVGYFN#TYK
417	P14625	LTESPCALVASQYGWSGN#MER
418	P14625	LTESPCALVASQYGWSGN#MER
419	P11717	YYINVCRPLNPVPGCN#R
420	Q07954	GTNVCAVAN#GGCQQLCLYR
421	H0YEV6	GLNECLHN#NGGCSHICTDLK

422	P25398	QAHLCVLASN#CDEPMYVK
423	Q07954	LTSCATN#ASIN#GDEAR
424	Q07954	FGTCSQLCN#NTK
425	P21912	EGICGSCAMNIN#GGNTLACTR
426	H0YKJ9	QLHCCGIHN#YSDWENTDWFK
427	Q8IY17-2	SDCTYIVLN#GR
428	P05556	NPCTSEQN#CTSPFSYK
429	Q9HAV0	KACNDATLVQITSN#MDSVGR
430	P16219	IGCFALSEPGN#GSDAGAASTTAR
431	H0YGW7	ICIVGPN#GVGK
432	Q07954	TCVSN#CTASQFVCK
433	Q9Y277	VCN#YGLTFTQK
434	P53582	SCCTSVN#EVICHGIPDR
435	E9PF10	VPLPPELVEQFGHMQCN#CMMGVFPPISR
436	Q9BVK6	FTFTSHTPGEHQICLHSN#STK
437	Q7Z7H5-3	FTFTSHTPGDHQICLHSN#STR
438	P68431	FQSSAVMALQEACEAYLVGLFEDTN#LCAIHAK
439	P13639	LMEPIYLVEIQCPEQVVGGIYGVLN#R
440	Q07954	SGFHTVPGQPGCQDIN#ECLR
441	P05386	ALANVNIGSLICN#VGAGGPAPAAGAAPAGGPAPSTAAAPAEEK
442	H0YKQ8	SSFAPPHFPVCMLGN#R
443	Q969P0-3	IGPGEPLELLCN#VSGALPPAGR
444	P07437; P68371	EIVHIQAGQCGN#QIGAK
445	Q8WWI5-3	CAPVN#ISCYAK
446	P05556	CHEGN#GTFECGACR
447	P60709	CPEALFQPSFLGMESCGIHETTFN#SIMK (Oxidation)
448	P60709	CPEALFQPSFLGMESCGIHETTFN#SIMK
449	P18084	CHAGYIGDNCN#CSTDISTCR