Supporting Information for

Novel nanoporous covalent organic frameworks for the selective extraction of endogenous peptides

Xiaofei Zhang,[#] Guangyan Qing,[#] Long Yu, Hongjian Kang, Cheng Chen, Xiuling Li,[*] and Xinmiao Liang[*]

Key Laboratory of Separation Science for Analytical Chemistry, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, 457 Zhongshan Road, Dalian 116023, P. R. China.
[#]*These authors contributed equally to this work.*
E-mail: lixiuling@dicp.ac.cn, liangxm@dicp.ac.cn,
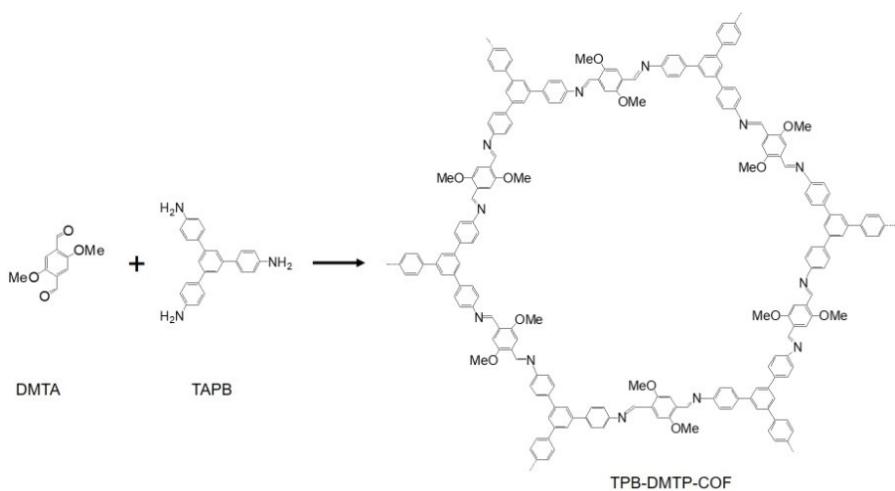
**Scheme S1.** Synthesis of TPB-DMTP-COF through the condensation of DMTA and TAPB.
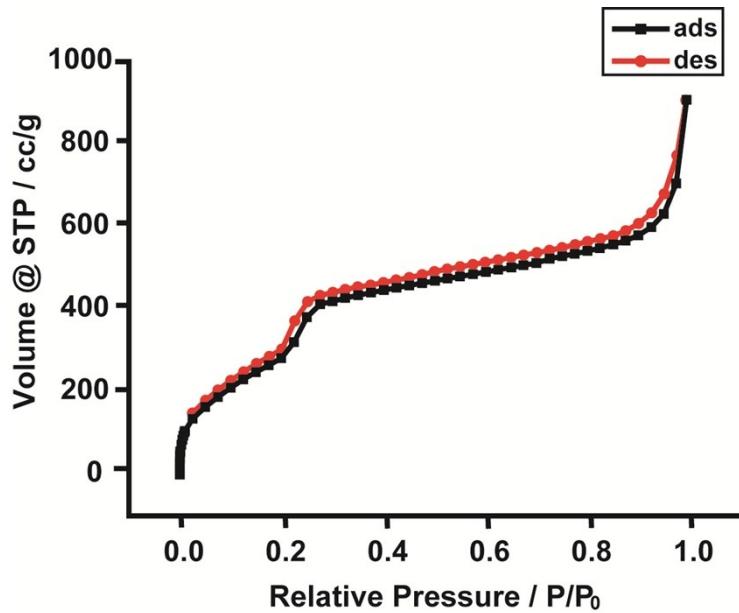
Figure S1. Nitrogen adsorption/desorption isotherm of TPB-DMTP-COF. The nitrogen adsorption/desorption isotherm revealed a type-IV curve, indicating the mesostructure of TPB-DMTP-COF.

Figure S2. Water contact angle of MCM-41 (A) and TPB-DMTP-COF (B).
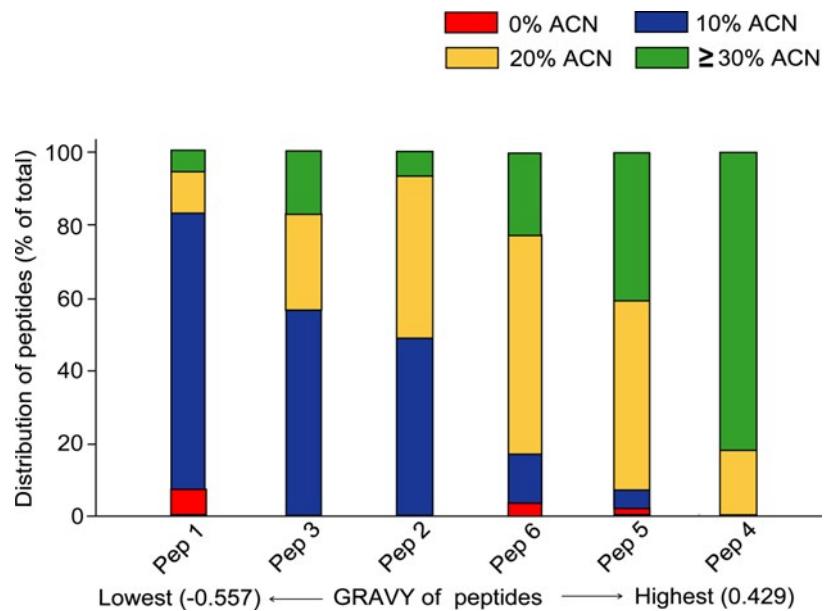
Figure S3. The relationship of GRAVY value and distribution of peptides.
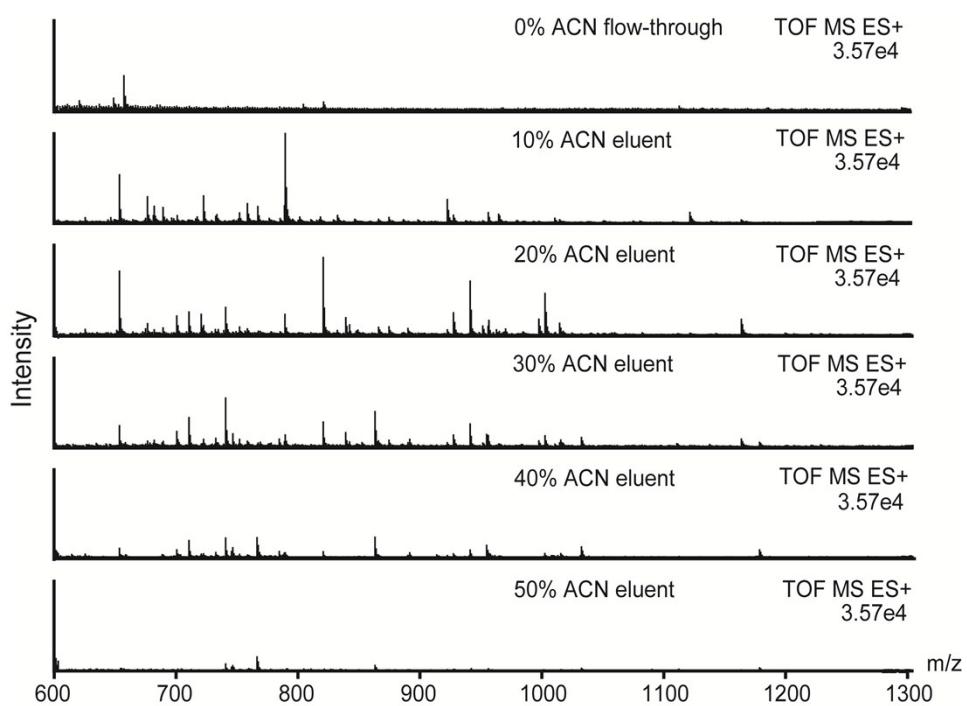
Figure S4. The distribution of BSA tryptic digests in fractions eluted with 0-50% ACN from TPB-DMTP-COF at pH 5.9. The majority of BSA tryptic peptides distributed in the eluents with 10-40% ACN.
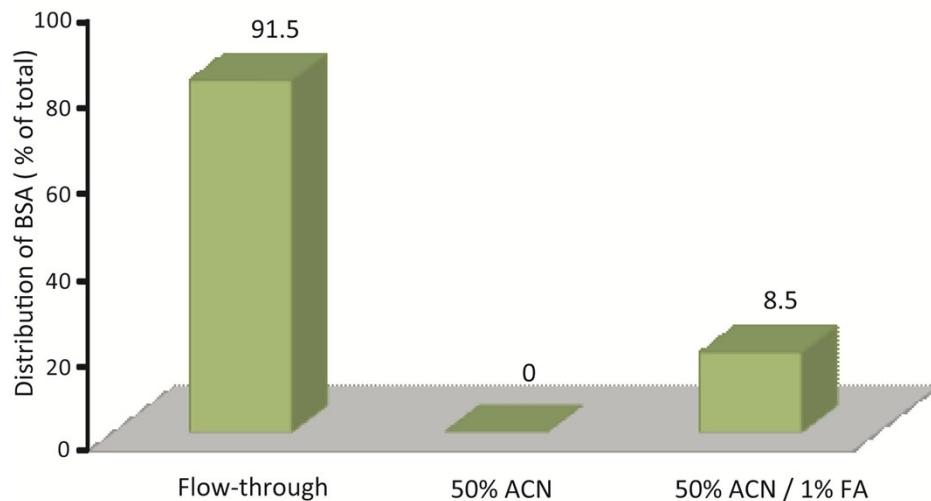
Figure S5. The distribution of BSA protein in flow-through and fractions eluted with 50% ACN and 50% ACN/1% FA from TPB-DMTP-COF. 91.5% of the BSA protein were detected in flow-through for the size-exclusion of TPB-DMTP-COF, only 8.5% of the BSA protein (adsorbed on the hydrophobic external surface of TPB-DMTP-COF) were detected in 50% ACN/1% FA.

Calculation of ratio between external surface area and total surface area of TPB-DMTP-COF

The external surface area of TPB-DMTP-COF particles was estimated via assuming the particles are solid spheres and the external surface area of 1g TPB-DMTP-COF was the specific surface area of the solid spheres. Assumed the specific surface area is "S", partical size is "d", quantity of particles in 1g TPB-DMTP-COF is "n", density of particles is "ρ", the surface area equation is as follows:

$$1 = \frac{n\rho\pi d^3}{6}$$

$$n = \frac{6}{\rho \pi d^3}$$

$$S = n\pi d^2 = \frac{6}{\rho \pi d^3} \cdot \pi \cdot d^2 = \frac{6}{\rho d}$$

Assumed the density of TPB-DMTP-COF particles is 1 kg m^{-3}, particle size measured with TEM is around 100 nm, the external surface area of 1 g TPB-DMTP-COF particles is approximately 60 m^2, which accounts for 7.3 % of the total surface area (826.5 m^2) of 1 g TPB-DMTP-COF acquired from N$_2$ sorption–desorption isotherms.
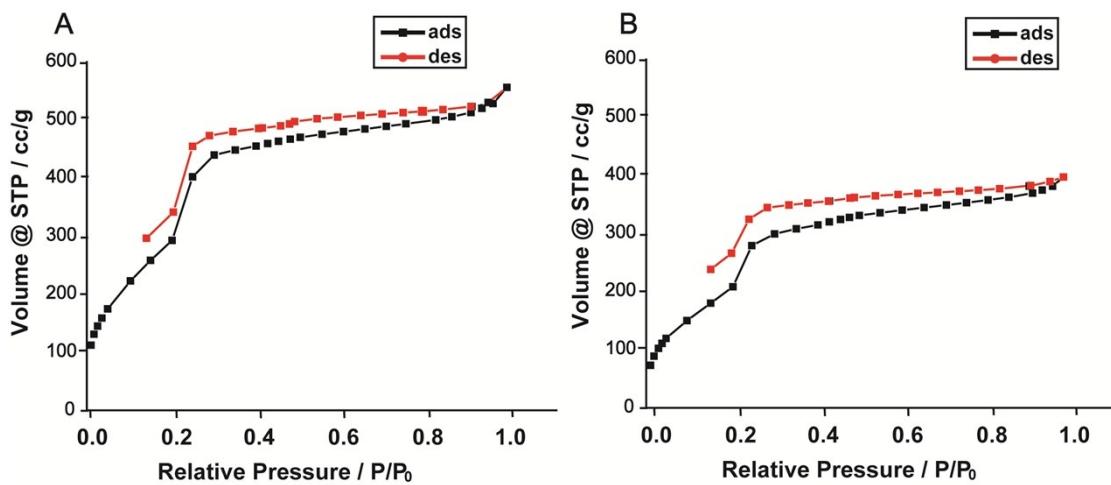
Figure S6. Nitrogen adsorption/desorption isotherm of TPB-DMTP-COF after incubation with BSA protein (A) and BSA tryptic peptides (B). Measurements of the extent of N$_2$ adsorption indicated that, the pore volume of TPB-DMTP-COF was 8.516e-01 cc/g, the pore volumes of TPB-DMTP-COF after incubation with BSA protein and BSA tryptic peptides were 8.602 e-01 cc/g and 6.150 e-01 cc/g, respectively. The pore volume of PB-DMTP-COF decreased by about 27.8% after the adsorption of the BSA tryptic peptides, but those were almost unchanged after the adsorption of BSA.
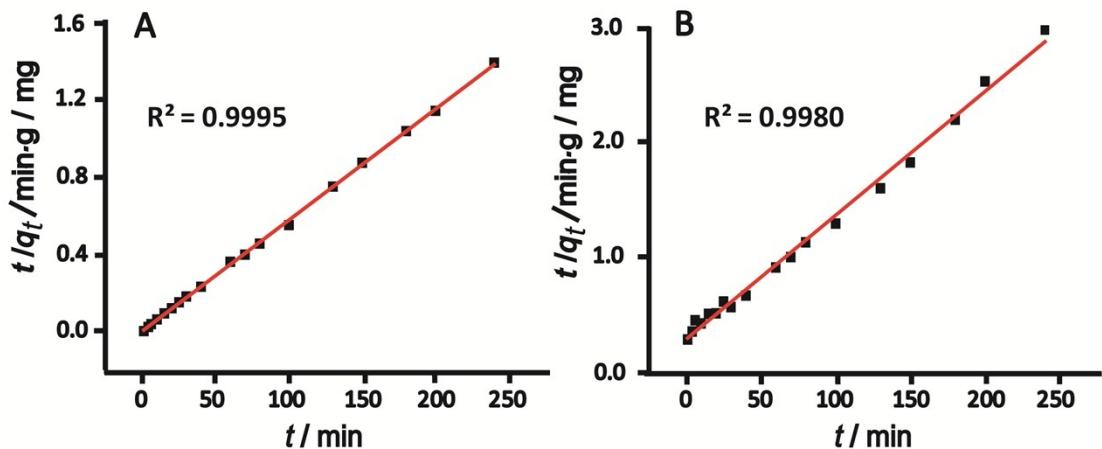
Figure S7. Linear fitting curves with the adsorption kinetic data of BSA tryptic digests (A) and BSA protein (B) on TPB-DMTP-COF according to pseudo-second-order rate equation;

The experimental kinetic data were fitted with the pseudo-first-order (Lagergren, 1) and pseudo-second-order rate equations (2), which were extensively used adsorption kinetic models in the solid - liquid phase system.

$$\frac{dq_t}{dt} = k_1 \cdot (q_e - q_t) \qquad 1$$

$$\frac{dq_t}{dt} = k_2 \cdot (q_e - q_t)^2 \qquad 2$$

After integral of the formula, and assumed the marginal conditions: when t=0, qt=0, the equation could be changed as follows:

$$ln(q_e - q_t) = lnq_e - k_1 t \qquad 3$$

$$\frac{t}{q_t} = \frac{t}{q_e} + \frac{1}{k_2 q_e^2} \qquad 4$$

Where q_e (mg/g) and q_t (mg/g) are the adsorption amount at equilibrium state and at time t (min), respectively; k_1 (min^{-1}) and k_2 (g/(mg·min)) are the rate constants of two models, respectively. After linear fitting of the kinetic data according to the linear function of $ln(q_e - q_t)$ and t (or t/q_t and t), it was realized that the data fitting better to the pseudo-second-order model (Fig. S6) than the pseudo-first-order model (Fig. S8), for the former's

linearly dependent coefficient $R^2 > 0.99$. The adsorption rate constant k_2 calculated from pseudo-second-order rate equation is 2.44×10^{-3} g/(mg·min) for BSA tryptic peptides and 6.80×10^{-4} g/(mg·min) for BSA protein on TPB-DMTP-COF.

,

Figure S8. Distributions of 6 peptides in eluents with 10-50% ACN at different pH values after extraction with TPB-DMTP-COF.
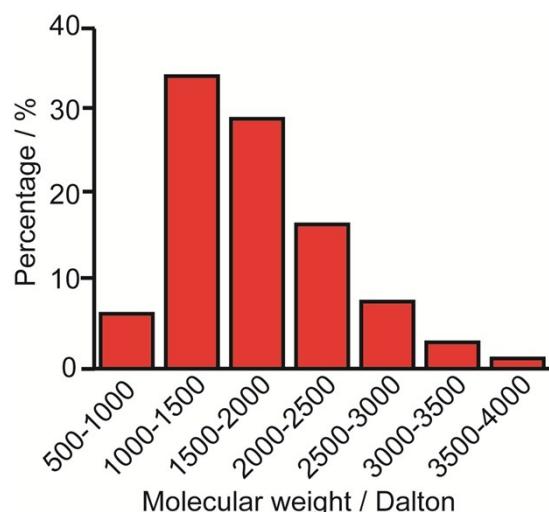
Figure S9. The molecular weight distribution for all of the identified endogenous peptides from human serum.
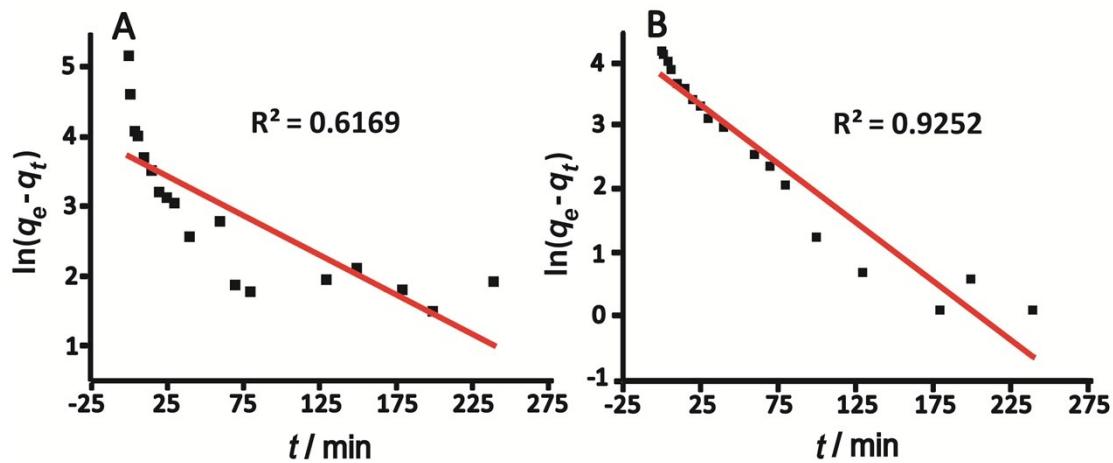
Figure S10. The linear fitting curves with the adsorption kinetic data of tryptic digests of BSA (A) and BSA protein (B) on TPB-DMTP-COF according pseudo-first-order rate equation.
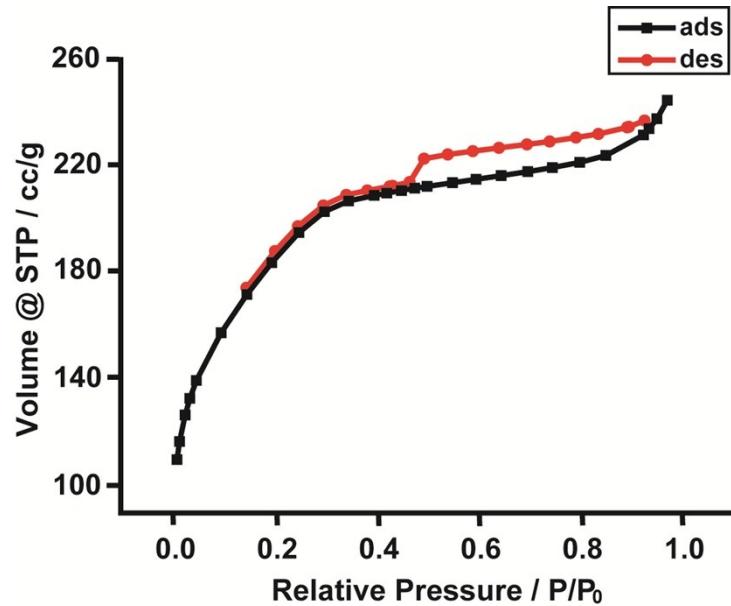
Figure S11. Nitrogen adsorption/desorption isotherm of MCM-41.
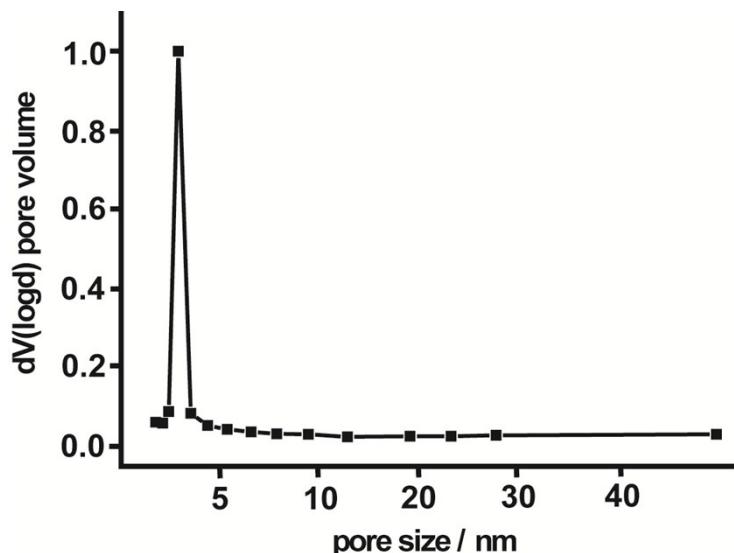
Figure S12. Pore size distribution of MCM-41.

Table S1. The detailed information of the 6 peptides in BSA digests.

Code	Sequence	MW(Da)	pI	GRAVY	Aliphatic index
Pep 1	KVPQVSTPTLVEVSR	1639.9	8.75	-0.557	70.00
Pep 2	AEFVEVTK	922.5	4.53	0.175	85.00
Pep 3	LVNELTEFAK	1163.6	4.53	0.130	117.00
Pep 4	LVTDLTK	789.5	5.84	0.429	152.86
Pep 5	LGEYGFQNALIVR	1479.7	6.00	0.292	120.00
Pep 6	AWSVAR	688.8	9.79	0.267	81.67

Table S2. Human serum endogenous peptides detected in this study using TPB-DMTP-COF

Sequence	MH+	pI	GRAVY
H.RIHWESASLL.R	1211.65	6.75	0.020
-.DIQM*TQSPSSVSA.S	1366.62	3.80	-0.277
K.SHALQLNNRQ.I	1180.62	9.49	-1.310
R.SLDRNLPSDSQDLGQHGLEED.F	2325.06	4.10	-1.371
T.SSTSYNRGDSTFES.K	1537.64	4.37	-1.379

Q.FTSSTSYNRGDS.T	1321.56	5.84	-1.250
S.GEGDFLAEGGGVR.G	1263.60	4.14	-0.339
N.TKSPLFM*GKVVNPTQK.-	1790.98	10.30	-0.475
A.VDSGNDVTDIAD.D	1220.53	3.52	-0.392
Y.GSGGGSYGSGGGGGGHGSY.G	1514.59	6.74	-0.726
N.TKSPLFMGKVVNPTQK.-	1774.99	10.30	-0.475
P.YSYSTTAVVTNPKE.-	1559.76	6.22	-0.614
S.TTAVVTNPKE.-	1059.57	5.88	-0.440
S.DLQAQSKGNPEQTPV.L	1611.80	4.37	-1.347
R.ASVGQDSPEPR.S	1142.54	4.37	-1.291
Y.NTGGNPTEDVSVNSRPF.R	1790.83	4.37	-1.047
L.PPTSAHGNVAEGETKPD.P	1706.80	4.83	-1.241
G.EIVMTQSPATLSVSPGERATLS.C	2274.17	4.53	0.096
S.APHGPGLIYRQPN.C	1419.75	8.80	-0.885
E.KNPLPSKETIEQEKQAGES.-	2113.08	4.95	-1.684
R.ILGGHLDAKGS.F	1067.58	6.74	0.118
S.SRQLGLPGPPDVPDHAAYHPF.R	2271.13	5.93	-0.671
K.SSSYSKQFTSSTSYNRGDST.F	2189.96	8.22	-1.380
C.GQDQVTVAM*TPR.S	1318.64	5.84	-0.525
E.GKKLVAASQAALGL.-	1326.81	10.00	0.707
A.EEAGARVQQNVPSGTDTGDPQSKPLG.D	2638.27	4.32	-1.154
N.HYTQKSLSLSPG.K	1317.68	8.60	-0.783
R.RAVPPNNSNAAEDDLPTVELQGVVPR.G	2758.41	4.32	-0.596
F.SLGSKINV.K	817.48	8.47	0.388
E.EGKKLVAASQAALGL.-	1455.85	8.69	0.427
-.DIQMTQSPSSLSA.S	1364.64	3.80	-0.308
V.PPNNSNAAEDDLPTVELQGVVPR.G	2432.20	3.92	-0.739
K.VNPFRPGDSEPPPAPGAQRAQ.M	2188.09	6.04	-1.205
V.SETESRGSESGIFTNTKESS.S	2132.96	4.49	-1.315
P.LFVSPPTHGRPVIASPSYPCHSAIPHAGASLPPP.P	3514.76	8.25	0.065
K.QFTSSTSYNRGDSTFES.K	1913.82	4.37	-1.218
S.SRQLGLPGPPDVPD.H	1447.75	4.53	-0.800
A.EDPQGDAAQKTDTSH.H	1599.69	4.22	-1.913
A.TKTAKDALSSVQESQVAQQA.R	2090.07	5.73	-0.660

A.EEAGARVQQNVPSGTDT.G	1758.82	4.14	-1.065
P.NSPLDEENLTQENQDRGTHVDLGLASAN.V	3037.41	4.01	-1.146
K.SLAELGGHLDQQVEEF.R	1771.85	4.00	-0.350
G.VLSSRQLGLPGPPDVPDHAA.Y	2026.07	5.21	-0.180
L.VAASQAALGL.-	900.52	5.49	1.430
K.VNPFRPGDSEPPPAPGAQ.R	1832.89	4.37	-1.061
D.SGEGDFLAEGGGVR.G	1350.63	4.14	-0.371
L.LSPYSYSTTAVVTNPKE.-	1856.93	6.22	-0.424
A.SVGQDSPEPR.S	1071.51	4.37	-1.600
G.EIVM*TQSPATLSVSPGERATLS.C	2290.18	4.53	0.096
R.TVVQPSVGAAA.G	999.55	5.18	1.000
G.EIVLTQSPATLSLSPGERATL.S	2183.19	4.53	0.210
K.NPLPSKETIEQEKQAGES.-	1984.98	4.49	-1.561
F.DTASTGKTFPG.F	1081.52	5.84	-0.736
A.EDPQGDAAQKTDTSHH.D	1736.75	4.64	-1.994
R.SVQLTEKR.M	960.55	8.46	-1.113
K.GNPEQTPVLKPEEEAPAPEVGASKPEGI.D	2870.44	4.21	-0.929
N.SQESVTEQD.S	1022.43	3.88	-1.733
R.ATASPRPSSGNIPSSPTASGGGSPTSPR.A	2581.27	12.00	-0.796
P.HGPGLIYRQPN.C	1251.66	8.75	-1.064
R.AVPPNNSNAAEDDLPTVELQGVVPR.G	2602.31	3.92	-0.440
R.QLGLPGPPDVPDHAAYHPF.R	2028.00	5.05	-0.463
R.ILGGHLDAKG.S	980.55	6.74	0.210
R.NVHSAGAAGSRMN.F	1271.59	9.76	-0.431
F.HVQPQPQPKPQVQ.L	1510.81	8.76	-1.739
F.MGKVVNPTQK.-	1101.61	10.00	-0.720
Y.LQGAKIPKPEASFSPR.R	1725.97	9.99	-0.713
E.TQEKNPLPSKETIEQEKQAGES.-	2471.23	4.65	-1.805
P.GVLSSRQLGLPGPPDVPDHAA.Y	2083.09	5.21	-0.191
E.EIVKEVSTYIK.K	1308.73	6.24	-0.018
Y.LQGAKIPKPEA.S	1151.68	8.59	-0.591
R.SLDRNLPSDSQDLGQHGLEEDF.M	2472.13	3.96	-1.182
F.AEEGKKLVAASQAALGL.-	1655.93	6.19	0.277
R.IHWESASL.L	942.47	5.24	0.113

Y.GSGGGSYGSGGGGGGGRGSY.G	1590.65	8.59	-0.775
K.SHALQLNN.R	896.46	6.46	-0.638
G.KKLVAASQAALGL.-	1269.79	10.00	0.792
R.NGFKSHALQLNNRQ.I	1626.85	11.00	-1.293
A.NSPTAGAAKSSPAAKPG.S	1511.78	10.00	-0.647
R.NVHSGSTF.F	848.39	6.74	-0.300
A.EEAGARVQQNVPSG.T	1441.70	4.53	-0.943
K.SSSYSKQFTSS.T	1208.54	8.31	-1.036
K.PKNPANPVQ.R	964.52	9.18	-1.467
A.ALLSPYSYSTTAVVTNPKE.-	2041.05	6.26	-0.084
K.THINIVVIGHVDSGK.S	1588.88	6.61	0.433
A.PHGPGLIYRQPN.C	1348.71	9.18	-1.108
C.HPNSPLDEENLTQEN.Q	1736.77	4.00	-1.887
T.ADSGEGDFLAEGGGV.R	1380.59	3.43	-0.160
Y.YLQGAKIPKPEAS.F	1401.77	8.50	-0.662
K.SSSYSKQFTSSTSYNRGDSTFESKS.Y	2768.23	8.15	-1.352
K.GSEM*VVAGKLQDR.G	1405.71	6.07	-0.354
T.SSTSYNRGDSTFESKSY.K	1915.83	5.79	-1.488
R.GGSGGSYGGGGSGGGYGGGSGSR.G	1791.73	8.59	-0.744
R.NLPSDSQDLGQHGLEED.F	1853.82	3.94	-1.400
Y.TIAALLSPYSYSTTAVVTNPKE.-	2326.22	5.88	0.182
C.GKPKNPANPVQ.R	1149.64	10.00	-1.591
E.EAPSLRPAPPPISGGGY.R	1665.86	6.10	-0.482
-.DIQM*TQSPSTL.S	1236.58	3.80	-0.446
I.HWESASLLR.S	1098.57	6.75	-0.478
R.GGSGGSHGGGSGFGGESGGSY.G	1714.67	5.24	-0.667
Q.EKNPLPSKETIEQEKQAGES.-	2242.12	4.65	-1.775
A.EDPQGDAAQKTDTSHHDQDHPTFN.K	2691.13	4.44	-2.025
R.SLDRNLPSDSQDLGQHG.L	1838.87	4.41	-1.300
G.DQTVSDNELQEMSNQGSKYVN.K	2386.05	3.92	-1.410
S.GGGGGGGLGSGGSIR.S	1145.57	9.75	-0.120
R.VQEKQHPVPPPAQNQN.Q	1810.92	6.72	-1.738
S.EM*VVAGKLQ.D	990.53	6.10	0.511
S.EEAERSDGDPVQPAVLQVHQTS.-	2392.15	4.17	-0.982

R.NVHSAGAAGSRM*N.F	1287.59	9.76	-0.431
K.KVPQVSTPTLVEVSR.N	1639.94	8.75	-0.067
K.QFTSSTSYNRGD.S	1362.59	6.09	-1.475
R.EGVQKEDIPPADLSDQVPDTESETRILL.Q	3094.54	3.85	-0.750
L.PGPPDVPDH.A	930.43	4.20	-1.422
G.DFLAEGGGVR.G	1020.51	4.37	-0.010
G.EGDFLAEGGGV.R	1050.47	3.57	0.046
G.KPKNPANPVQ.R	1092.62	10.00	-1.710
K.SPLFM*GKVVNPTQK.-	1561.84	10.00	-0.214
K.KKEEEEDEEDEEDEEEEDEEDEDEEEDDDD.E	3877.28	3.33	-3.526
S.IQDWVQKTIAEN.-	1444.74	4.37	-0.667
N.SGSSGPGSTGNRNPGSSGTGGTATWKPGSSGP.G	2847.29	11.00	-1.066
D.TNRASVGQDSPEPR.S	1513.74	5.73	-1.636
K.SSSYSKQFTSSTSYN.R	1673.73	8.22	-1.180
Q.FTSSTSYNRGDSTFES.K	1785.76	4.37	-1.075
L.SALEEYTKKLNTQ.-	1524.79	5.86	-1.223
V.NDNEEGFFS.A	1058.41	3.57	-1.456
Y.SYSTTAVVTNPKE.-	1396.70	5.94	-0.562
Y.STTAVVTNPKE.-	1146.60	5.94	-0.473
Y.TQKSLSLSPG.K	1017.56	8.41	-0.490
L.GLPGPPDVPDHAAY.H	1405.67	4.20	-0.507
Y.ALTSEEAERSDGDPVQPAVLQVHQTS.-	2764.34	4.17	-0.673
S.NAAEDDLPTVELQGVVPR.G	1922.98	3.92	-0.333
Q.SGNSQESVTEQDSKD.S	1610.68	4.12	-1.900
Y.NRGDSTFESKSY.K	1390.62	6.07	-1.742
Q.FTSSTSYNRGDST.F	1422.61	5.84	-1.208
V.LKPEEEAPAPEVGASKPEGI.D	2048.05	4.33	-0.770
G.EIVMTQSPATLSVSPGERATL.S	2187.13	4.53	0.138
-.EIVLTQSPGTLSLSPGERA.T	1955.04	4.53	-0.047
K.CCTESLVNR.R	1138.50	5.99	0.000
K.DDPDAPLQPVTPLQLFEGR.R	2108.07	3.84	-0.674
Y.ALTSEEAERSDGDPVQPAVLQVHQ.T	2576.26	4.17	-0.667
V.DTNRASVGQDSPEPR.S	1628.76	4.56	-1.760
K.MADEAGSEADHEGTHST.K	1744.67	4.17	-1.200

A.NTQPRGPPASSPAPAPK.F	1672.88	11.00	-1.312
R.NGFKSHALQLNNRQI.R	1739.93	11.00	-0.907
P.GSTGNRNPGSSGTGGTATWKPGSSGP.G	2375.09	11.00	-1.127
A.EEAGARVQQNVPSGTDTG.D	1815.85	4.14	-1.028
T.SESTAALGCLVK.D	1235.64	5.72	0.650
L.SPYSYSTTAVVTNPKE.-	1743.84	5.94	-0.688
R.NVHSGSTFF.K	995.46	6.74	0.044
G.DQTVSDNELQEMSNQG.S	1794.75	3.43	-1.519
A.DSGEGDFLAEGGGVR.G	1465.66	3.92	-0.580
P.SLTSDLQAQSKGNPEQTPV.L	1999.99	4.37	-0.984
A.NTQPRGPPASSPAPAPKF.S	1819.95	11.00	-1.083
A.KVEQAVETEPEPELRQQT.E	2111.06	4.33	-1.500
P.FRPGDSEPPPAPGAQRAQ.M	1877.93	6.07	-1.356
-.EIVLTQSPGTL.S	1157.64	4.00	0.464
S.TSYNRGDSTFES.K	1363.58	4.37	-1.475
P.GVLSSRQLGLPGPPDVPDHA.A	2012.06	5.21	-0.290
-.DIQM*TQSPSSLSA.S	1380.63	3.80	-0.308
S.PYSYSTTAVVTNPKE.-	1656.81	6.63	-0.680
R.SEETKENEGFTVTAEGK.G	1855.86	4.32	-1.394
K.KLVAASQAALG.L	1028.61	8.75	0.946
K.KLVAASQAALGL.-	1141.70	8.75	1.183
A.KDALSSVQESQVAQQA.R	1688.85	4.37	-0.606
N.SQESVTEQDSKD.S	1352.58	4.12	-1.983
R.SLAPYAQDTQEKLN.H	1577.78	4.37	-1.043
G.EIVLTQSPATL.S	1171.66	4.00	0.664
Y.LQGAKIPKPGLDHTEASFSPR.R	2249.21	8.60	-0.733
R.GSGGGSSGGSIGGRGSSSGGVK.S	1751.83	11.00	-0.441
R.SLDRNLPSDSQDLGQHGLEEDFM.L	2603.17	3.96	-1.048
K.SSSYSKQFTSSTSYNRGDSTFES.K	2553.11	5.79	-1.265
G.SPSGEVSHPRKT.R	1281.66	8.49	-1.467
G.KKLVAASQAALG.L	1156.70	10.00	0.542
K.DDPDAPLQPVTP.L	1264.61	3.42	-0.942
A.DQLRTQVNTQAEQL.R	1643.84	4.37	-1.200
N.SNAAEDDLPTVELQGVVPR.G	2010.01	3.92	-0.358

G.EIVLTQSPATLSLSPGERA.T	1969.06	4.53	0.068
P.PLPPPPPPPP.P	1102.63	5.96	-1.109
P.NNSNAAEDDLPTVELQGVVPR.G	2238.10	3.92	-0.657
K.SYKMADEAGSEADHEGTHST.K	2122.86	4.49	-1.320
K.M*ADEAGSEADHEGTHST.K	1760.67	4.17	-1.200
V.SLGSPSGEVSHPRKT.R	1538.79	8.49	-1.000
S.SRQLGLPGPPDVPDHAA.Y	1726.89	5.19	-0.635
D.PEVRPTSAVAA.-	1097.60	6.43	0.100
A.EEAGARVQQNVPSGTDTGDPQ.S	2155.98	3.92	-1.291
K.VNPFRPGDSEPPPAPGAQRAQMG.R	2376.15	6.04	-1.035
A.TPLPPTSAHGNVAEGETKPDPDVTER.C	2715.34	4.50	-1.123
R.VAPEEHPVLLTEAPLNPK.A	1954.06	4.75	-0.267
-.EIVLTQSPGTLS.L	1244.67	4.00	0.358
A.DSGEGDFLAEGGGV.R	1309.56	3.43	-0.300
N.YKILQADQEL.-	1220.65	4.37	-0.530
E.GDFLAEGGGVR.G	1077.53	4.37	-0.046
R.TVVQPSVGAAAG.P	1056.57	5.18	0.883
Y.GGGEEASGSGGGYGGGSGK.S	1527.63	4.53	-0.905
K.AVDTNRASVGQDSPEPR.S	1798.87	4.56	-1.200
K.KEEEEDEEDEEDEEEEEDEEDEDEEEDDDDE.-	3878.29	3.13	-3.513
R.QLGLPGPPDVPDHAAY.H	1646.82	4.20	-0.425
D.FLAEGGGVR.G	905.48	6.00	0.378
T.ADSGEGDFLAEGGGVR.G	1536.69	3.92	-0.431
S.SYSKQFTSSTSYNRGDST.F	2015.89	8.22	-1.444
K.SSSYSKQFTSSTSYNRGDSTFESKSY.K	2931.29	8.11	-1.350
N.AAEDDLPTVELQGVVPR.G	1808.94	3.92	-0.147
-.NIEM*TQSPSSL.S	1222.56	4.00	-0.455
C.EVQLLESGGGLVQPGG.S	1539.80	3.80	0.088
S.PAAKPGSTPSRPS.S	1252.67	11.01	-1.131
A.ATASRGASQAGAPQGRVPEARPN.S	2249.15	11.70	-0.904
-.NIEMTQSPSSLSASVG.D	1607.76	4.00	-0.063
F.AEEGKKLVAASQAALG.L	1542.85	6.19	0.056
K.REEAPSLRPAPPPISGGGY.R	1951.00	6.14	-0.853
A.EDPQGDAAQKTDTSHHD.Q	1851.77	4.49	-2.082

H.NHYTQKSLSLSPG.K	1431.73	8.60	-0.992
R.NGFKSHALQLNNRQIR.G	1896.03	12.01	-1.131
R.GGSGGGGGGSSGGRGSGGGSSGGSIGGR.G	2079.91	12.00	-0.618
G.EIVMTQSPATLSVSPGERA.T	1973.00	4.53	-0.011
L.GLPGPPDVPDHAAYHPF.R	1786.85	5.05	-0.535
-.DIQMTQSPSSL.S	1206.57	3.80	-0.455
K.GSEMVVAGKLQ.D	1118.59	6.00	0.309
I.AALLSPYSYSTTAVVTNPKE.-	2112.08	6.26	0.010
K.SYKM*ADEAGSEADHEGTHST.K	2138.86	4.49	-1.320
V.SETESRGSESGIFTNTKES.S	2045.93	4.49	-1.342
N.GQPENNYKTTPPVLDSDG.S	1931.90	4.03	-1.444
R.EEAPSLRPAPPPISGGGY.R	1794.90	4.53	-0.650
S.SYSKQFTSST.S	1135.53	8.31	-1.050
F.ISAAAVPPGSLLSGP.G	1336.72	5.52	0.913
-.NIEM*TQSPSSLS.A	1309.60	4.00	-0.483
I.LGGHLDAKG.S	867.47	6.74	-0.267
A.VPPNNSNAAEDDLPTVELQGVVPR.G	2531.27	3.92	-0.533
K.NMITGTSQADCAVLIVAAGVGEFEAGISK.N	2909.43	4.14	0.693
-.DIQMTQSPSTLSA.S	1378.65	3.80	-0.300
S.SRQLGLPGPPDVPDHA.A	1655.85	5.19	-0.788
A.VDTNRASVGQDSPEPR.S	1727.83	4.56	-1.388
-.DIQMTQSP.S	919.42	3.80	-0.900
N.SGALTSGVHTFPAVLQ.S	1584.84	6.46	0.644
Q.DEPPQSPW.D	955.42	3.67	-2.125
V.SLGSPSGEVSHPR.K	1309.65	6.47	-0.800
H.YTQKSLSLSPG.K	1180.62	8.59	-0.564
K.DDPDAPLQPVTPLQLFEGRRN.R	2378.21	4.23	-0.991
M.NFRPGVLS.S	889.49	9.75	0.000
Q.SGNSQESVTEQD.S	1280.52	3.88	-1.692
R.QLGLPGPPDVPDHA.A	1412.72	4.20	-0.521
R.TLEIPGNSDPNMIPDGDFNSYVR.V	2551.18	3.84	-0.717
N.NNYKILQADQEL.-	1448.74	4.37	-1.025
G.DQTVSDNELQEMSNQGSKYVNKEIQNAVN.G	3282.53	4.18	-1.276
V.QAAVGTSAAPVPSDNH.-	1521.73	5.08	-0.250

A.EEGKKLVAASQAALGL.-	1584.90	6.24	0.181
L.SSRQLGLPGPPDVPDHA.A	1742.88	5.19	-0.788
-.EIVLTQSPG.T	943.51	4.00	0.222
V.LLTEAPLNPK.A	1095.64	6.00	-0.160
G.EGDFLAEGGGVR.G	1206.57	4.14	-0.333
K.TETQEKNPLPSKETIEQEKQAGES.-	2701.31	4.48	-1.829
N.AVEEVTQNNFR.L	1306.64	4.53	-0.882
A.EDPQGDAAQKTDTSHHDQ.D	1979.83	4.34	-2.161
L.DRNLPSDSQDLGQHG.L	1638.75	4.41	-1.673
S.GNSQESVTEQDSKD.S	1523.65	4.12	-1.979
A.EVSADQVATV.M	1018.51	3.67	0.420
R.MNFRPGVLS.S	1020.53	9.50	0.211
N.PFRPGDSEPPPAPGAQRAQ.M	1974.98	6.50	-1.368
S.PLFMGKVVNPTQK.-	1458.81	10.02	-0.169
R.TGKEKVTSGSTTTTR.R	1553.81	9.99	-1.213
N.SGALTSGVHTFPAVLQS.S	1671.87	6.46	0.559
A.EEAGARVQQNVPSGTDTGDPQSKP.L	2468.16	4.32	-1.392
R.GGSGGSHGGGSGFGGESG.G	1407.55	5.24	-0.639
S.STSYNRGDSTFES.K	1450.61	4.37	-1.423
A.EDPQGDAAQKTDTS.H	1462.63	3.84	-1.821
R.SLDRNLPSDSQDLGQHGLEEDFM*.L	2619.16	3.96	-1.048
P.LPPTSAHGNVAEGE.T	1378.66	4.81	-0.543
T.SDLQAQSKGNPEQTPVLKPEEEAPAPEVGASKPEGI.D	3727.87	4.31	-1.011
G.DQTVSDNELQEM*SNQGSKYVNKEIQNAVN.G	3298.53	4.18	-1.276
-.DIQMTQSPS.S	1006.45	3.80	-0.889
T.FTYAKVVALIAIIVM*GLVKLCQEIC.P	2868.57	8.03	1.764
A.LLSPYSYSTTAVVTNPKE.-	1970.01	6.22	-0.189
V.SGSTGQWHSESGSFRPDSPGSGN.A	2320.98	5.30	-1.361
R.ELLESYIDGR.I	1194.60	4.14	-0.540
D.SGEGDFLAEGGGV.R	1194.53	3.57	-0.054
A.EEAGARVQQNVPSGTDTGD.P	1930.87	4.12	-1.158
K.SSSYSKQFTSSTSY.N	1559.69	8.22	-1.014
K.LVAASQAALGL.-	1013.60	5.52	1.646
R.NLPSDSQDLGQHGLEEDFM.L	2131.92	3.77	-1.005

-.DIQMTQSPSTL.S	1220.58	3.80	-0.446
-.EIVLTQSPGTLSLSPGERATL.S	2169.18	4.53	0.105
R.ILGGHLDAK.G	923.53	6.74	0.278
S.SSYSKQFTSSTSYNRGDSTFES.K	2466.07	5.79	-1.286
R.M*NFRPGVLS.S	1036.52	9.50	0.211
P.PASSPAPAPKFSPVTPKFTPVAS.K	2281.22	10.02	-0.104
S.CPGSM*DEVLASLRHGRAPL.R	2082.00	6.74	-0.063
S.KFSPGAPGGSGSQPNQKLGHPEA.L	2248.11	8.60	-1.113
S.GEGDFLAEGGGV.R	1107.50	3.57	0.008
L.GLPGPPDVPDH.A	1100.54	4.20	-0.855
R.SGGGGGGGLGSGGSIR.S	1232.60	9.46	-0.163
R.PGVLSSRQL.G	956.55	10.18	0.022
K.TAKDALSSVQESQVAQQA.R	1860.93	4.37	-0.478
C.HPNSPLDEEN.L	1151.50	4.13	-2.090
T.SEEAERSDGDPVQPAVLQVHQTS.-	2479.17	4.17	-0.974
T.SIQDWVQKTIAEN.-	1531.78	4.37	-0.677
C.EVQLLESGGGLVQPGGSLRLS.C	2096.14	4.53	0.138
D.WVQKTIAEN.-	1088.57	6.00	-0.611
R.NVHSAGAAGSRM.N	1157.55	9.76	-0.175
R.TATSEYQTFFNPR.T	1561.73	5.66	-1.031
H.TASYGMSLI.V	942.47	5.18	0.889
-.DIQMTQSPSSVSA.S	1350.62	3.80	-0.277
T.NRASVGQDSPEPR.S	1412.69	6.07	-1.708
S.SSYSKQF.T	846.40	8.31	-1.186
G.SEMVVAGKLQ.D	1061.57	5.72	0.380
F.TSSTSYNRGDSTFESKSY.K	2016.88	5.73	-1.444
E.GKKLVAASQAALG.L	1213.73	10.00	0.469
A.ATASRGASQAGAPQGRV.P	1584.82	12.00	-0.465
V.AASQAALGL.-	801.45	5.57	1.122
R.IHWESASLLR.S	1211.65	6.75	0.020
R.GPPASSPAPAPK.F	1076.57	8.75	-0.708
K.VIHDNFGIVEGLMTTVHAITATQK.T	2595.36	5.99	0.429
K.KVEEEDEEEEEEEEEEEEEEDE.-	2799.05	3.43	-3.168
R.GGSGGSHGGGSGFGGESGGS.Y	1551.61	6.61	0.433

Y.AASSLQSGVPSR.F	1159.61	3.52	-0.392
K.SPLFMGKVVNPTQK.-	1545.85	9.79	-0.133
A.GTELVNFLSYFVELGTQPATQ.-	2314.16	6.61	0.433
P.KNPANPVQ.R	867.47	3.84	-0.674
K.TIGGGDDSFNTFFSETGAGK.H	2007.89	5.99	0.429
N.FLSYFVELGTQPATQ.-	1700.85	4.03	-0.445
A.VDSGNDVTDIADD.G	1335.55	10.30	-0.475
R.IHWESASLL.R	1055.55	4.32	-1.394
H.RIHWESASLLR.S	1367.75	5.93	-0.671
R.GGSGGSHGGGSGFGGE.S	1263.50	4.32	-1.394
L.YLYEIAR.R	927.49	8.60	-0.783
L.GLPGPPDVPDHA.A	1171.57	10.00	0.792
C.DIQM*TQSPSSLSA.S	1380.63	6.74	-0.726
Q.SGNSQESVTEQDSKDSTY.S	1961.82	3.33	-3.526
K.KKNEPEDEEEEEEEEDEDEEEEDED.E	3126.12	5.99	0.429
A.GTELVNFLSYFVELGTQPAT.Q	2186.10	4.00	0.213
Q.FTSSTSYNRGDSTFESKSY.K	2163.95	3.57	0.008
K.GSEM*VVAGKLQ.D	1134.58	6.61	0.433
L.SRSGGGGGGGGLGSGGSIRSSY.S	1812.86	8.59	-0.775
R.KKVEEEDEEEEEEEEEEEEED.E	2798.05	8.15	-1.352
R.AFVHWYVGEGMEEGEFSEAR.E	2330.02	3.84	-0.674
A.EDPQGDAAQKTD.T	1274.55	3.65	-3.186
V.VIGHVDSGK.S	911.50	9.46	-0.163
G.NSQESVTEQDSKDSTY.S	1817.77	4.37	-1.075
R.GGSGGSHGGGSGFGGESGGSYGGGEEA.S	2214.87	6.61	0.433
Y.GGGEEASGSGGGYGGGSGKSSHS.-	1925.79	5.79	-1.265
S.SYSKQFTSSTSYNRGDSTFES.K	2379.04	10.00	0.542
K.TFPGFFSPMLGEFV.S	1575.76	5.99	0.429
R.DAHKSEVAHRFKDL.G	1652.87	6.92	-1.114
Y.YLQGAKIPKPEASFSPR.R	1889.03	9.70	-0.747
R.DAHKSEVAHRFKDLG.E	1709.87	6.92	-1.067
R.MNFRPGVLSSRQLGLPGPPDVPDHAAYHPF.R	3272.64	6.69	-0.407
G.DQTVSDNELQEM*SNQG.S	1810.74	3.43	-1.519
S.STSYNRGDSTFESKSY.K	1828.80	5.79	-1.531

S.YSTTAVVTNPKE.-	1309.66	6.22	-0.542
C.SVMHEALHNHYTQKSLSLSPG.K	2336.14	6.78	-0.567
R.NVNFQKAIN.E	1047.56	8.75	-0.511
T.GIFTDQVLSVLKGEE.-	1634.86	4.43	0.207
R.SSKITHRIHWESASLLR.S	2021.10	6.74	-0.300
R.PGVLSSRQLGLPGPPDVPDHAAYHPF.R	2724.39	6.61	0.433
R.MNFRPGVLSSRQLGLPGPPDVPDHA.A	2657.36	4.00	0.213
N.DSGPRRYTIAALLSPYSYSTTAVVTNPKE.-	3157.61	6.50	-0.428
C.FAEEGKKLVAASQAALGL.-	1803.00	6.61	0.433
R.DAHKSEVAHRFKDLGEEN.F	2082.00	8.43	-0.435
G.DQTVSDNELQEM*SNQGSKYVNKEIQN.A	3014.36	5.39	-1.472
K.CDSSPDSAEDVR.K	1337.53	4.18	-1.519
A.DEAGSEADHEGTHSTKRGHAKSRPV.R	2659.26	5.94	-0.473
R.GDSTFESKSY.K	1120.48	6.69	-0.407
R.QLGLPGPPDVPDHAA.Y	1483.76	3.92	-0.440
C.EVQLVESGGGLVQPG.G	1468.76	5.93	-0.671
K.SYKMADEAGSEADHEGTHSTK.R	2250.96	8.22	-1.014
S.SYSKQFTSSTSYN.R	1499.67	3.33	-3.526
N.SQESVTEQDSKDSTY.S	1703.73	8.22	-1.239
S.SSYSKQFTSST.S	1222.56	3.92	-1.773
T.ECCHGDLLECADDR.A	1749.66	5.88	-0.440
I.THRIHWESASLLR.S	1605.86	4.17	-1.200
F.TSSTSYNRGDSTFES.K	1638.69	8.59	-0.744
I.THRIHWESASLL.R	1449.76	5.39	-1.472
N.IFFMSKVTNPKQA.-	1510.81	4.18	-1.276
C.SVM*HEALHNHYTQKSLSLSPG.K	2352.14	8.11	-1.350
M.ADEAGSEADHEGTHST.K	1613.63	6.62	-0.308
F.TDQVLSVLKGEE.-	1317.69	8.46	-1.113
C.EVQLVESGGGLVQPGGSLRLS.C	2082.14	4.17	-1.394
G.NSQESVTEQDSKD.S	1466.62	6.61	0.433
N.FLSYFVELGTQPAT.Q	1572.79	6.75	-0.478
K.SSSYSKQFT.S	1034.48	6.61	0.433
R.TAFGGRRAVPPNNSNAAEDDLPTVELQGVVPR.G	3347.71	4.17	-0.673
V.SETESRGSESGIFTNTKESSSHHPGIAEFPSRG.K	3505.62	6.61	0.433

R.M*NFRPGVLSSRQLGLPGPPDVPDHAA.Y	2744.39	8.31	-1.089
R.DAHKSEVAHRF.K	1296.65	5.39	-1.472
R.ILGGHLDAKGSFPWQ.A	1625.84	4.78	-0.528
L.SSRQLGLPGPPDVPDHAAYHPF.R	2358.16	5.44	-1.127
N.LDVSLQLPSR.S	1127.64	6.92	-1.067
Y.GGGSGSGGGSGGGYGGGSGGGHSGGSGGGHSGGSGGNY.G	2855.11	9.76	-0.431
R.YEGSYALTSEEAERSDGDPVQPAVLQVHQTS.-	3363.56	6.61	0.433
A.DEAGSEADHEGTHST.K	1542.59	11.00	-1.127
S.SSYSKQFTSSTSYNRGDST.F	2102.93	4.37	-0.010
C.DIQM*TQSPSSL.S	1222.56	3.96	-1.182
E.LLRTLNQPDSQLQLTTGNGLFLSE.G	2658.41	6.61	0.433
Y.FVELGTQPATQ.-	1190.61	10.00	0.707
K.HFSVEGQLEFR.A	1348.67	4.08	-0.800
R.QLGLPGPPDVPDHAAYHPFR.R	2184.10	4.17	-1.607
C.EVQLVESGGGLVQPGG.S	1525.79	3.80	-0.455
R.IVEGSDAEIGM*SPWQVMLFRKSPQELL.C	3076.59	4.00	-0.118
G.VLSSRQLGLPGPPDVPDHA.A	1955.03	6.61	0.433
K.SLEDKTERELLESYIDGR.I	2153.08	5.24	-0.639
G.DQTVSDNELQEM*SNQGSKYVN.K	2402.04	6.61	0.433
C.DIQMTQSPSSLSA.S	1364.64	5.44	-1.127
G.DSTFESKSY.K	1063.46	4.41	0.019
S.SYSKQFTSSTSYNRGDSTFESKSY.K	2757.23	6.69	-0.407
S.RQLGLPGPPDVPDHAAYHPF.R	2184.10	3.84	-0.717
G.VLSSRQLGLPGPPDVPDHAAYHPF.R	2570.31	8.11	-1.396
G.SPMYSIITPNILRLE.S	1746.94	5.19	-0.635
K.HNDDEQYAWESSAGGSFTVR.T	2255.96	4.32	-1.394
-.EIVLTQSPGTLSL.S	1357.76	3.92	-0.533
N.FRPGVLSSRQLGLPGPPDVPDHAAYHPF.R	3027.56	3.92	-0.431
K.SSSYSKQFTS.S	1121.51	5.98	-0.288
C.DIQM*TQSPSS.L	1109.48	5.94	0.313
K.MADEAGSEADHEGTHSTK.R	1872.77	6.92	-0.379
C.DIQMTQSPS.S	1006.45	8.15	-1.352
R.NVHSGSTFFKYYLQGAKIPKPEA.S	2582.34	3.43	-0.160
A.EDPQGDAAQKTDT.S	1375.60	4.32	-0.596

R.M*NFRPGVLSSRQLGLPGPPDVPDHAAYHPF.R	3288.64	5.98	-0.288
P.HFFFPK.S	822.43	5.88	0.182
P.GVLSSRQLGLPGPPDVPDHAAYHPF.R	2627.33	8.60	-0.733

References

1 H. Xu, J. Gao, D. Jiang, Nat. Chem., 2015, 7, 905–912.