Supporting Information

Streptomyces erythraeus Trypsin for Proteomics Applications

Jianying Z. Kiser[†], Marc Post[⊥], Benlian Wang[†§], Masaru Miyagi[*,†,‡,§]

[†]Case Center for Proteomics and Bioinformatics, [‡]Department of Pharmacology, [§]Department of Ophthalmology and Visual Sciences, Case Western Reserve University, 10900 Euclid Ave., Cleveland, OH, USA
[⊥]USB Research and Development, Affymetrix, Inc., 26111 Miles Rd. Cleveland, OH, USA

Supplemental Figures and Tables:

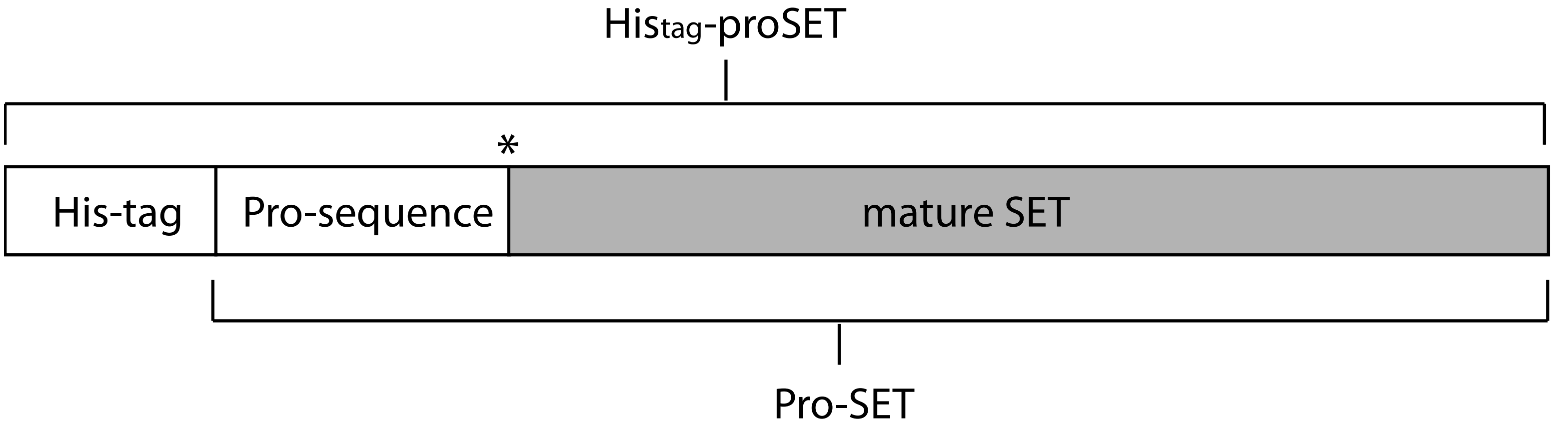

Figure S-1. Synthetic His-tag-proSET gene. The His tag-proSET gene consists of three distinct sequences: the His tag, the pro-peptide sequence (12 AA), and the mature SET sequence. The asterisk indicates the site for cleavage by chymotrypsin.

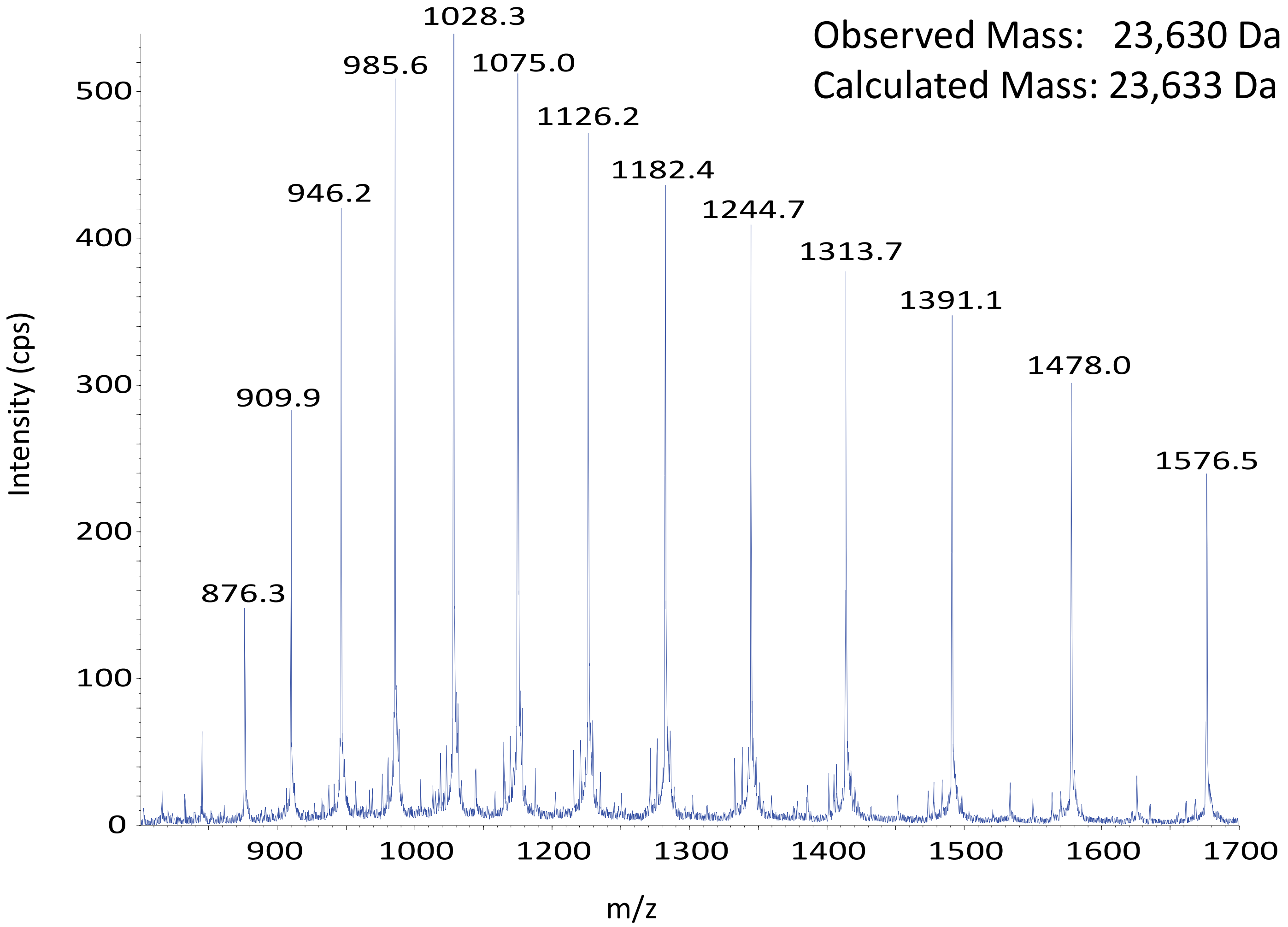

Figure S-2. Electrospray mass spectrum of rSET. rSET was dissolved in 0.1% formic acid/50% acetonitrile and introduced into a QStar quadrupole/time-of-flight mass spectrometer (Applied Biosystem-MDS Sciex, Foster City, CA) equipped with a nano-electrospray ion source at a flow rate of 5 μL/min.

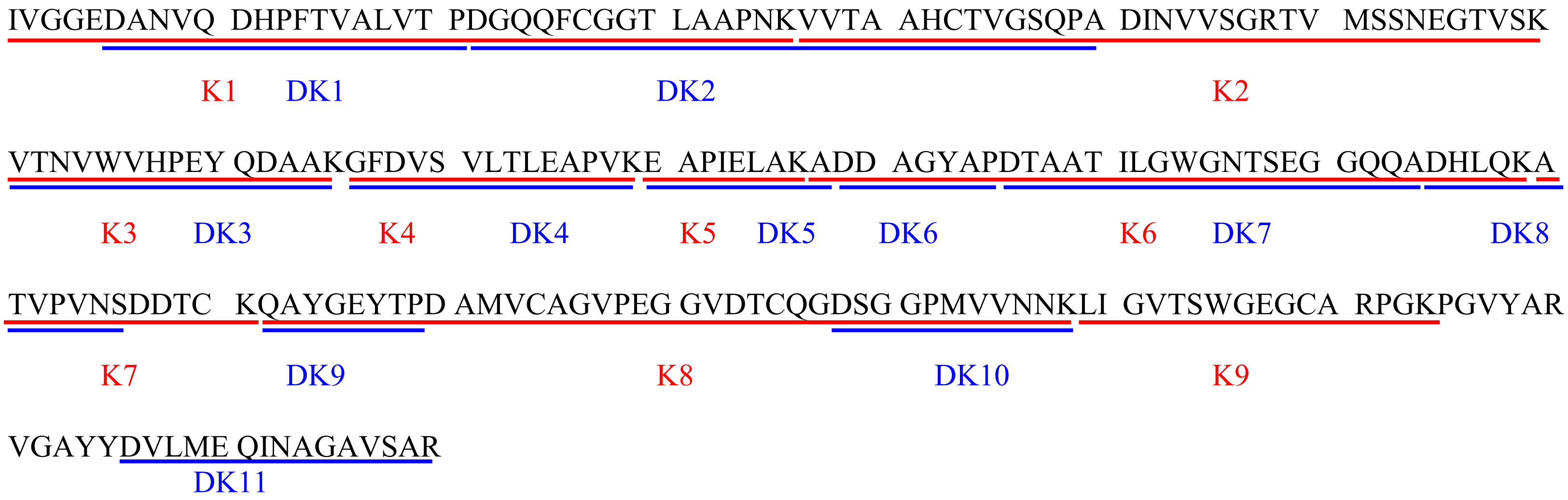

Figure S-3. Peptides identified by LC-MS/MS. The observed peptides are underlines with red (Lys-C peptides) or blue (Lys-C/Asp-N peptides) lines. Lys-C peptides and Lys-C/Asp-N peptides are referred to as K and KD peptides, respectively, and numbered from the N-terminus to the C-terminus. The coverage of the rSET amino acid sequence was 95% by LC-MS/MS analysis.

Table S-1. Peptides analyzed.

			Molecular Mass	
Peptide Name	Peptide Sequence	Position	Observed	Calculated
K1	IVGGEDANVQDHPFTVALVTPDGQQFCGGTLAAPNK	1-36	3723.76	3723.81
K2	VVTAAHCTVGSQPADINVVSGRTVMSSNEGTVSK	37-70	3457.64	3457.69
K3	VTNVWVHPEYQDAAK	71-85	1755.82	1755.86
K4	GFDVSVLTLEAPVK	86-99	1473.78	1473.81
K5	EAPIELAK	100-107	869.47	869.49
K6	ADDAGYAPDTAATILGWGNTSEGGQQADHLQK	108-139	3257.45	3257.49
K7	ATVPVNSDDTCK	140-151	1305.56	1305.59
K8	QAYGEYTPDAMVCAGVPEGGVDTCQGDSGGPMVVNNK	152-188	3829.64	3829.63
K9	LIGVTSWGEGCARPGK	189-204	1686.82	1686.85
DK1	DANVQDHPFTVALVTP	6-21	1722.86	1722.86
DK2	DGQQFCGGTLAAPNKVVTAAHCTVGSQPA	22-48	2827.33	2627.34
DK3	VTNVWVHPEYQDAAK	71-85	1755.86	1755.86
DK4	GFDVSVLTLEAPVK	86-99	1473.81	1473.81
DK5	EAPIELAKA	100-108	940.52	940.52
DK6	DDAGYAP	109-115	707.28	707.28
DK7	DTAATILGWGNTSEGGQQA	116-134	1875.87	1875.86
DK8	DHLQKATVPVNS	135-146	1307.68	1307.68
DK9	QAYGEYTP	152-159	927.40	927.40
DK10	DSGGPMVVNNK	178-188	1116.52	1116.52
DK11	DVLME QINAGAVSAR	216-230	1572.80	1572.79

K peptides and DK peptides were analyzed by Q-star and FT-LTQ, respectively.
Cystein residues were carbamidomethylated.

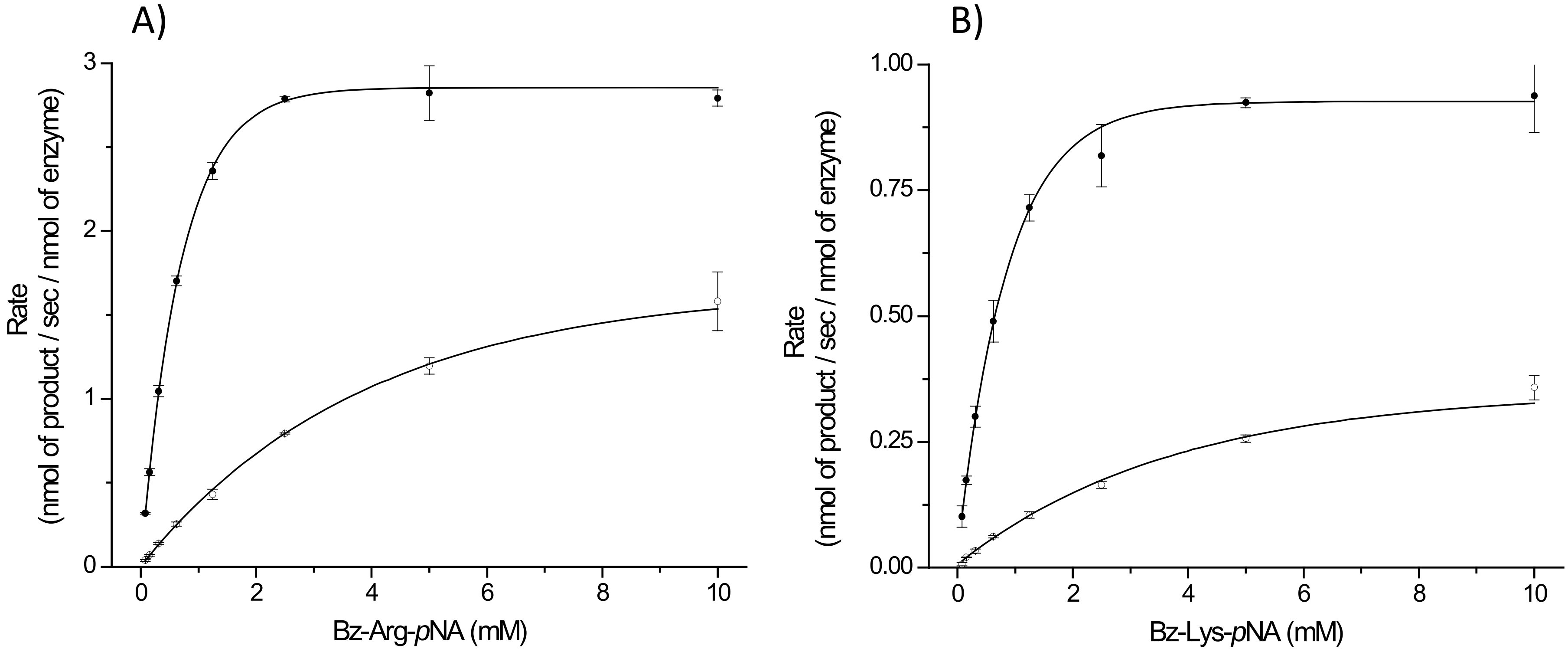

Figure S-4. Substrate concentration curves for rSET and modified porcine trypsin with Bz-Arg-*p*NA (A) and Bz-Lys-*p*NA (B). Various concentrations of Bz-Arg-*p*NA or Bz-Lys-*p*NA were incubated with rSET (-●-) or modified porcine trypsin (-○-) at 25 ºC and pH 8.0. The *p*-nitroaniline produced was quantified by VIS-spectroscopy (405 nm). Values represent the mean ± standard deviation of triplicate reactions.

Table S-2. BSA tryptic peptides identified by rSET and modified porcine trypsin in in-solution digestion experiment.
The peptides were produced by digesting BSA in solution by either rSET or modified porcine trypsin at 25 ºC for 5 h.

Peptide Sequence[a]	Position	Calculated Mass	Retention Time[b] (min)	Observed Mass[c] in rSET digest	Observed Mass[c] in porcine trypsin digest
R.FKDLGEEHFK.G	35 - 44	1248.614	20.0	1248.617	1248.617
K.DLGEEHFK.G	37 - 44	973.451	20.0	973.452	973.452
K.LVNELTEFAK.T	66 - 75	1162.623	28.6	1162.623	1162.622
K.LVNELTEFAKTCVADESHAGCEK.S	66 - 88	2607.195	29.7	2607.198	2607.198
K.TCVADESHAGCEK.S	76 - 88	1462.582	14.0	1462.587	1462.588
K.TCVADESHAGCEKSLHTLFGDELCK.V	76 - 100	2863.258	26.7	2863.265	2863.266
K.SLHTLFGDELCK.V	89 - 100	1418.686	28.0	1418.688	1418.689
K.VASLRETYGDMADCCEK.Q	101 - 117	2003.839	24.0	N.I.	2003.842
K.LKPDPNTLCDEFKADEK.K	139 - 155	2018.962	25.5	2018.964	2018.964
K.YLYEIAR.R	161 - 167	926.486	25.1	926.488	926.488
K.YLYEIARR.H	161 - 168	1082.587	22.2	1082.589	1082.590
R.RHPYFYAPELLYYANK.Y	168 - 183	2044.021	30.6	2044.025	2044.026
R.HPYFYAPELLYYANK.Y	169 - 183	1887.920	33.0	1887.926	1887.927
K.YNGVFQECCQAEDK.G	184 - 197	1746.698	23.4	1746.702	1746.701
K.YNGVFQECCQAEDKGACLLPK.I	184 - 204	2486.103	28.0	2486.107	2486.107
K.GACLLPKIETMR.E	198 - 209	1387.732	26.8	1387.735	N.I.
R.ALKAWSVAR.L	233 - 241	1000.582	21.1	N.I.	1000.584
R.LSQKFPK.A	242 - 248	846.496	15.8	846.500	846.500
K.FPKAEFVEVTK.L	246 - 256	1293.697	24.8	1293.698	1293.699
K.AEFVEVTK.L	249 - 256	921.481	21.8	921.483	921.483
K.AEFVEVTKLVTDLTK.V	249 - 263	1691.935	29.2	1691.941	1691.941
K.ECCHGDLLECADDR.A	264 - 280	2112.878	18.3	2112.885	2112.884
K.YICDNQDTISSK.L	286 - 297	1442.635	19.9	1442.635	1442.635
K.YICDNQDTISSKLK.E	286 - 299	1683.814	22.4	1683.818	1683.817
K.LKECCDKPLLEK.S	298 - 309	1531.774	17.9	1531.777	1531.777
K.ECCDKPLLEK.S	300 - 309	1290.595	17.7	1290.599	1290.599
K.DAFLGSFLYEYSR.R	347 - 359	1566.735	27.7	1566.744	1566.744
R.RHPEYAVSVLLR.L	360 - 371	1438.805	24.7	1438.807	1438.807
R.LAKEYEATLEECCAK.D	372 - 386	1813.823	22.8	N.I.	1813.826
K.EYEATLEECCAK.D	375 - 399	3037.242	27.0	3037.247	N.I.
K.DDPHACYSTVFDKLK.H	387 - 401	1794.825	27.0	N.I.	1794.829
K.LKHLVDEPQNLIK.Q	400 - 412	1545.888	23.1	1545.892	N.I.
K.HLVDEPQNLIK.Q	402 - 412	1304.709	24.0	1304.711	1304.711
K.HLVDEPQNLIKQNCDQFEK.L	402 - 420	2354.133	26.0	2354.139	2354.138
K.QNCDQFEKLGEYGFQNALIVR.Y	413 - 433	2528.212	33.6	2528.220	2528.220
K.LGEYGFQNALIVR.Y	421 - 433	1478.788	31.9	1478.791	1478.792
R.KVPQVSTPTLVEVSR.S	437 - 451	1638.931	25.2	1638.932	1638.933
K.VPQVSTPTLVEVSR.S	438 - 451	1510.836	27.7	1510.839	1510.839
R.CCTKPESERMPCTEDYLSLILNR.L	460 - 482	2871.302	32.6	2871.313	2871.315
R.MPCTEDYLSLILNR.L	469 - 482	1723.827	30.7	1723.836	1723.836
R.LCVLHEKTPVSEK.V	483 - 495	1538.813	19.4	1538.815	1538.815
K.CCTESLVNR.R	499 - 507	1137.491	19.2	1137.493	1137.495
R.RPCFSALTPDETYVPK.A	508 - 523	1879.914	27.2	1879.919	1879.919
R.RPCFSALTPDETYVPKAFDEK.L	508 - 528	2470.184	29.2	2470.195	2470.189
K.AFDEKLFTFHADICTLPDTEK.Q	524 - 544	2497.184	32.2	2497.187	2497.187
K.LFTFHADICTLPDTEK.Q	529 - 544	1906.914	30.6	1906.918	1906.919
K.LFTFHADICTLPDTEKQIK.K	529 - 547	2276.151	29.4	2276.154	2276.155
K.KQTALVELLK.H	548 - 557	1141.707	25.7	1141.709	1141.709
K.QTALVELLK.H	549 - 557	1013.612	30.0	1013.614	1013.614
K.ATEEQLK.T	562 - 568	817.418	27.3	817.422	817.422
K.TVMENFVAFVDK.C	569 - 580	1398.685	25.7	1398.694	1398.690
K.CCAADDKEACFAVEGPK.L	581 - 597	1926.791	22.2	1926.796	1926.796
K.EACFAVEGPK.L	588 - 597	1106.507	22.5	1106.509	N.I.
K.LVVSTQTALA.-	598 - 607	1001.576	27.5	1001.578	1001.578

[a] Cystein residues were carbamidomethylated.

[b] Retention times shown are from the LC-MS/MS data for rSET digest unless for peptides solely identified in modified porcine trypsin digest.

[c] N.I. indicates peptides that were not identified by Mascot database search with the LC-MS/MS data.

Table S-3. BSA tryptic peptides identified by rSET and modified porcine trypsin in in-gel digestion experiment.
The peptides were produced by digesting BSA in gel by either rSET or modified porcine trypsin at 25 ºC for 5 h.

Peptide Sequence[a]	Position	Calculated Mass	Retention Time[b] (min)	Observed Mass[c]	
				in rSET digest	in porcine trypsin digest
R.DTHKSEIAHR.F	25 - 34	1192.590	21.1	N.I.	1192.603
R.FKDLGEEHFK.G	35 - 44	1248.610	17.5	1248.618	1248.619
K.DLGEEHFK.G	37 - 44	973.451	17.2	973.454	N.I.
K.LVNELTEFAK.T	66 - 75	1162.620	25.2	1162.627	1162.627
K.LVNELTEFAKTCVADESHAGCEK	66 - 88	2607.195	26.0	2607.201	2607.204
K.TCVADESHAGCEK.S	76 - 88	1462.582	24.5	1462.589	1462.590
K.TCVADESHAGCEKSLHTLFGDELCK.V	76 - 100	2863.258	23.4	2863.266	2863.266
K.SLHTLFGDELCK.V	89 - 100	1418.690	24.5	1418.689	1418.691
K.VASLRETYGDMADCCEK.Q	101 - 117	2019.834	18.1	2019.838	N.I.
R.ETYGDMADCCEK.Q	106 - 117	1493.511	15.2	1493.518	N.I.
R.NECFLSHKDDSPDLPK.L	123 - 138	1900.863	19.8	1900.870	N.I.
K.DDSPDLPK.L	131 - 138	885.400	17.7	885.412	885.413
K.LKPDPNTLCDEFK.A	139 - 151	1575.760	22.8	N.I.	1575.763
K.LKPDPNTLCDEFKADEK.K	139 - 155	2018.960	22.4	2018.965	2018.966
K.FWGKYLYEIAR.R	157 - 167	1444.750	27.7	1444.755	1444.756
K.YLYEIAR.R	161 - 167	926.490	22.2	926.470	926.480
R.RHPYFYAPELLYYANK.Y	168 - 183	2044.020	26.8	2044.025	2044.029
K.YNGVFQECCQAEDK.G	184 - 197	1746.700	20.6	1746.702	N.I.
K.YNGVFQECCQAEDKGACLLPK.I	184 - 204	2486.103	24.8	2486.108	2486.108
R.CASIQKFGER.A	223 - 232	1194.582	16.1	1194.586	1194.586
R.ALKAWSVAR.L	233 - 241	1000.582	18.3	N.I.	1000.585
K.AWSVAR.L	236 - 241	688.366	17.7	688.368	688.368
R.LSQKFPK.A	242 - 248	846.496	14.3	846.501	846.501
K.AEFVEVTK.L	249 - 256	921.480	19.1	921.483	921.484
K.AEFVEVTKLVTDLTK.V	249 - 263	1691.935	30.9	1691.941	1691.941
K.LVTDLTK.V	257 - 263	788.460	18.3	788.467	N.I.
K.VHKECCHGDLLECADDR.A	264 - 280	2112.878	16.1	2112.885	N.I.
K.ECCHGDLLECADDR.A	267 - 280	1748.660	19.4	1748.660	N.I.
K.ECCHGDLLECADDRADLAK.Y	267 - 285	2246.935	21.1	2246.940	N.I.
K.YICDNQDTISSK.L	286 - 297	1442.630	17.7	1442.638	1442.638
K.YICDNQDTISSKLK.E	286 - 299	1683.814	19.1	1683.818	1683.819
K.LKECCDKPLLEK.S	298 - 309	1531.770	15.4	1531.779	1531.777
K.ECCDKPLLEK.S	300 - 309	1290.595	15.6	1290.599	1290.600
K.SHCIAEVEK.D	310 - 318	1071.502	14.0	1071.507	1071.507
K.SHCIAEVEKDAIPENLPPLTADFAEDK.D	310 - 336	3008.444	28.1	3008.453	N.I.
K.DAIPENLPPLTADFAEDKDVCK.N	319 - 340	2457.173	28.4	2457.179	N.I.
K.NYQEAKDAFLGSFLYEYSR.R	341 - 359	2300.075	30.9	2300.083	2300.086
K.DAFLGSFLYEYSR.R	347 - 359	1566.740	33.8	1566.744	1566.745
R.RHPEYAVSVLLR.L	360 - 371	1438.800	21.6	1438.808	1438.808
R.LAKEYEATLEECCAK.D	372 - 386	1813.823	20.0	1813.829	1813.829
K.EYEATLEECCAKDDPHACYSTVFDK.L	375 - 399	3037.242	23.8	3037.250	N.I.
K.DDPHACYSTVFDK.L	387 - 399	1553.650	21.6	1553.649	N.I.
K.DDPHACYSTVFDKLK.H	387 - 401	1794.820	23.8	N.I.	1794.830
K.LKHLVDEPQNLIK.Q	400 - 412	1545.888	20.2	1545.892	N.I.
K.HLVDEPQNLIK.Q	402 - 412	1304.710	21.1	1304.712	1304.713
K.HLVDEPQNLIKQNCDQFEK.L	402 - 420	2354.133	22.8	2354.138	2354.138
K.QNCDQFEKLGEYGFQNALIVR.Y	413 - 433	2528.212	25.5	2528.220	2528.221
K.LGEYGFQNALIVR.Y	421 - 433	1478.790	28.1	1478.770	1478.770
K.LGEYGFQNALIVRYTR.K	421 - 436	1899.000	27.4	N.I.	1899.008
R.KVPQVSTPTLVEVSR.S	437 - 451	1638.930	22.2	1638.934	1638.934
K.VPQVSTPTLVEVSR.S	438 - 451	1510.836	24.4	1510.839	1510.840
R.SLGKVGTR.C	452 - 459	816.482	18.2	N.I.	816.486
K.VGTRCCTKPESER.M	456 - 468	1578.724	15.4	N.I.	1578.730
R.CCTKPESER.M	460 - 468	1165.490	17.3	1165.489	N.I.
R.CCTKPESERMPCTEDYLSLILNR.L	460 - 482	2871.302	28.7	2871.312	2871.315
R.MPCTEDYLSLILNR.L	469 - 482	1723.827	33.1	1723.836	1723.836
R.LCVLHEK.T	483 - 489	897.470	15.3	897.477	897.479
R.LCVLHEKTPVSEK.V	483 - 495	1538.813	17.0	1538.815	1538.817
K.CCTESLVNR.R	499 - 507	1137.490	16.8	1137.494	1137.494
K.CCTESLVNRRPCFSALTPDETYVPK.A	499 - 523	2999.394	24.9	N.I.	2999.401
R.RPCFSALTPDETYVPK.A	508 - 523	1879.910	23.9	1879.917	1879.919
R.RPCFSALTPDETYVPKAFDEK.L	508 - 528	2470.184	25.6	2470.187	2470.189
K.LFTFHADICTLPDTEK.Q	529 - 544	1906.910	26.9	1906.918	1906.922
K.LFTFHADICTLPDTEKQIK.K	529 - 547	2276.151	25.8	2276.157	2276.158
K.KQTALVELLK.H	548 - 557	1141.710	22.5	1141.708	1141.709
K.QTALVELLK.H	549 - 557	1013.612	26.4	1013.615	1013.615
K.QTALVELLKHKPK.A	549 - 561	1503.914	20.2	N.I.	1503.917
K.HKPKATEEQLK.T	558 - 568	1307.720	21.1	1307.726	1307.724
K.TVMENFVAFVDK.C	569 - 580	1398.685	31.1	1398.691	N.I.
K.CCAADDKEACFAVEGPK.L	581 - 597	1926.791	19.6	1926.796	1926.797
K.EACFAVEGPK.L	588 - 597	1106.507	19.9	1106.509	N.I.
K.LVVSTQTALA.-	598 - 607	1001.576	24.4	1001.578	1001.579

[a] Cystein residues were carbamidomethylated.

[b] Retention times shown are from the LC-MS/MS data for rSET digest unless for peptides solely identified in modified porcine trypsin digest.

[c] N.I. indicates peptides that were not identified by Mascot database search with the LC-MS/MS data.