

Online Supporting Information S6. Seven negative subsets for studying Lys-methylation. Each subset contains 217 negative samples randomly taken from the 1,518 negative samples in [Online Supporting Information S4](#) except for the 5th subset, which only contains 216 samples. None of the samples in one subset occurs in any other subset.

(1) Subset 1 (217 samples)

Sample #	Protein ID	Site	Sequences
1	P05141	49	QITADKQYKGI
2	P05141	94	FAFKDKYKQIF
3	P05141	96	FKDKYKQIFLG
4	P05141	105	LGGVDKRTQFW
5	P05141	245	MQSGRKGTDIM
6	P05141	272	GKAFFKGAWSN
7	Q9Y232	73	DKRKNKKGKTE
8	Q9Y232	74	KRKNKKGKTEY
9	Q9Y232	157	ALVIGKDHESK
10	Q9Y232	162	KDHESKNSQLF
11	P00042	11	KGSEKKGATLF
12	P00042	17	GATLFKTRCLQ
13	P00042	48	GRQSGKAEGYS
14	P00042	60	TDANIKKAVEW
15	P26358	58	LQTEIKNQLCD
16	P26358	83	YLAKVKSLLNK
17	P26358	177	AKGPAKRKPQE
18	P26358	198	IKEEDKDQDEK
19	Q5VTE0	41	KRTIEKFEKEA
20	Q5VTE0	129	EAGISKNGQTR
21	P0CE47	238	ERGIKVGEEV
22	O75367	117	ELLAKKRGSKG
23	P62807	12	SAPAPKKGSKK
24	P09988	123	VTIQPKDMQLA
25	P69150	19	AKAPRKQLASK
26	Q55BP0	30	SSFPSKPVNEV
27	Q55BP0	65	DLLIKKLPFQR
28	P61830	126	QKKDIKLARRL
29	P62803	32	IQGITKPAIRR
30	P04637	164	AMAIYKQSQHM
31	P04637	386	KKLMFKTEGPD
32	Q86L05	43	KNYDTKDKRF
33	Q86L05	53	FSGQIKIGTVT
34	Q86L05	97	IGPKNKKAIKK
35	Q86L05	105	IKKLSKKYDAF
36	Q86L05	147	EDMASKINDVK
37	Q86L05	152	KINDVKSTVKF
38	P14909	34	VEKTKKIKIID
39	P14909	80	DELREKIAQYL
40	P14909	94	YGTDVKKEEVI
41	P14909	105	VTPGAKPALFL

42	P14909	142	KLLGGKPIYAN
43	P14909	194	PNDVKKIVDIS
44	P14909	257	GYIVAKREIIQ
45	P14909	285	QKAAVKAFDTF
46	P14909	313	YDELTKVKGVE
47	P39462	136	IVPHYKMYKL
48	P39462	263	PKALAKQGKYV
49	P22498	76	QKMGLKIARLN
50	P22498	138	IFKDLKSRGLY
51	P22498	256	KSVSKKPVGII
52	P22498	445	RFGLLKVDYNT
53	P0C835	53	KYFRHKLPPDY
54	P80053	16	YTQOVKKLYKV
55	P80053	54	RGSDGKLTTFM
56	P80053	170	IKITGKVDFAV
57	P13123	52	EKDAPKELLDM
58	P61991	40	DEGGGKTGRGA
59	P26358	284	DEKDEKKHRSQ
60	P26358	291	HRSQPKDLAAK
61	O75530	398	VEDPHKAKCTT
62	O75530	400	DPHKAKCTTLT
63	Q5VTE0	371	AHIACKFAELK
64	P81542	20	DHAGHKVYAPA
65	P81542	32	PPVKEKALGIH
66	P00219	20	DHSGHKIYGPV
67	O43524	242	NPDGGKSGKAP
68	Q08369	311	PLAMRKEGIQT
69	P16403	110	FKLNKKAASGE
70	P16403	117	ASGEAKPKVKK
71	P16403	130	GTKPKKPVGAA
72	P16403	139	AAKKPKKAAGG
73	P10412	22	KTPVKKKARKS
74	P10412	34	GAAKRKASGPP
75	P10412	121	AKPKAKKAGAA
76	P10412	136	PAGAAKPKKA
77	P10412	148	GAATPKKSACK
78	P16401	21	EKSPAKKATK
79	P16401	23	SPAKKATKKA
80	P16401	35	GAGAAK RKATG
81	P16401	154	KKAVK KTPKKA
82	O75367	157	KKGARKSKKKQ
83	O75367	236	NTLEKKGKKEF
84	P40283	13	KPAEKKPASEK
85	P40283	33	KAPAEKPKKAG
86	P40283	52	AGGDKKKMKK
87	Q41811	60	TRGVRKIFLEN
88	O14979	322	QQQQQKGGRGA
89	Q14103	231	CFITFKEEEPV
90	P16862	76	VEAHLKELNSI
91	P16862	197	LNRPQKAIAVM
92	P16862	525	AVNENKIVRKP
93	P16862	548	EAIQAKDFKRA

94	O00159	24	PHRPC K LALGS
95	O00159	274	KGQCA K VSSIN
96	O00159	285	DKSDW K VVRKA
97	O00159	486	DLVEE K FKGII
98	O00159	488	VEEK F KGIISI
99	O00159	595	ELSD K KRPETV
100	O00159	605	VATQ F KMSLLQ
101	O00159	618	EILQ S KEPAYV
102	P13538	85	SMNPP K YDKIE
103	P13538	207	ASGE K KKEEQS
104	P13538	260	FGAT G KLASAD
105	P13538	354	KTAI Y KLTGAV
106	Q96RI1	36	SMKPA K GVLTE
107	Q96RI1	196	ECRL R KCKEMG
108	Q96RI1	251	VTST T KSCREK
109	Q96RI1	289	TNKIL K EEFSA
110	Q96RI1	335	QIALL K GSAVE
111	Q96RI1	474	RVND H KFTPLL
112	P11940	30	AMLYE K FSPAG
113	P11940	113	KSID N KALYDT
114	P11940	157	ERAIE K MNGML
115	P11940	277	RQTEL K RKFEQ
116	P06400	8	PKTP R KTAATA
117	P06400	80	WLTWE K VSSVD
118	P06400	122	FTEL Q KNIEIS
119	P06400	143	IDTST K VDNAM
120	P06400	164	FALF S KLERTC
121	P06400	228	LDYFI K LSPPM
122	P06400	327	EIY L KNKDL
123	P06400	341	FLDHD K TLQTD
124	P0CX43	91	SVDDL K KLNKN
125	P47914	47	NMRFA K KHNKK
126	P47914	55	NKKGL K KMQAN
127	Q969Q0	6	MVNVP K TRRTF
128	Q969Q0	58	KPIFR K KAKTT
129	Q969Q0	59	PIFR K KAKTTK
130	Q969Q0	89	AIKR C KHFELG
131	P0CX27	76	RLECV K CKTRA
132	P0CX27	89	TLKR C KHFELG
133	P02393	91	GLKEA K DKVDG
134	P02393	114	EAE E A K KQLVE
135	P07472	98	APATL K EGMSK
136	Q12962	212	YGINV K KPHYF
137	Q04207	122	GIQCV K KRDLE
138	Q04207	221	CDKV Q KEDIEV
139	Q04207	301	HRIE E KRKRTY
140	Q04207	425	SAPVP K STQAG
141	Q53688	12	RLQL N KNFNFG
142	E5KWG9	29	TASHD K VWPIL
143	E5KWG9	135	VSELV K LPSIP
144	E5KWG9	278	IASTP K GEYHK
145	E5KWG9	283	KGEY H KIASYY

146	E5KWG9	292	YYQWQKRVEKL
147	B6A879	93	TTADV KPTENI
148	B6A879	173	DGGGSKEEERG
149	B6A879	290	LRELQKKAHAM
150	Q9L5X8	126	FAHGLKPIVVI
151	Q9L5X8	438	MAPDGKGRVRM
152	Q6NC90	9	HIDPTKEVFAQ
153	Q6NEZ3	66	YKFNQKSNFSA
154	Q6NEZ3	108	GYKNAKLTKSS
155	Q6NEZ3	120	RINVGKNVKLD
156	Q6NEZ3	178	TFIPEKTILTN
157	Q6NEZ3	341	VMPGEKVSVSA
158	Q9UKV3	103	LVKRLKGALML
159	P18281	16	QGAALKHAETV
160	P18281	34	ENVTVKKVDRS
161	P18281	52	KPHELKHAETV
162	Q9SJU4	246	YDVAEKVWAEV
163	Q9ZU52	120	LYQSTKDGKTF
164	Q9ZU52	123	STKDGKTFVDC
165	Q9ZU52	142	GIKVDKGLSPL
166	Q9ZU52	171	SAEYKQGARF
167	P62157	22	FSLFDKDG DGT
168	P11118	76	TLMSRKMHD TD
169	P07463	76	SLMARKMKEQD
170	P04573	43	MDMLTKLGQTY
171	P04573	133	VEEAMKEADED
172	P00889	208	MDLIAKLPCVA
173	P00063	16	PAAGEKIFKTK
174	P00063	107	LIAYLKQSTAA
175	P00078	19	AARGEKLFKGR
176	P00077	19	AAKGEKIFKGR
177	P00077	22	GEKIFKGRAAQ
178	P00077	98	FAGIKKPQERA
179	P00043	60	YTDANKKKGVE
180	P00076	7	DAER GKLFES
181	P00076	65	EETLHKFLENP
182	P00076	78	YVPGTKMAFAG
183	P00047	13	PGDASKGANLF
184	P00047	97	GLKKNKDRNDL
185	Q71V39	179	VSAYIKKIGYN
186	P02993	172	FEEIKKEVSAY
187	P02994	177	TSNFIKKVGYN
188	P02994	217	WEKETKAGVVK
189	Q96KQ7	139	MTGAGKSPPSV
190	P53991	49	TTDMSKRTVPT
191	P53991	55	RTVPTKLEEGE
192	P53991	69	NTYSNKAPFKA
193	P53991	81	VRSVEKITGPK
194	P53991	164	DPETGKEDPAK
195	P11416	94	FVCQDKSSGYH
196	P11416	166	DRNKKKKEAPK
197	P11416	360	LLEALKVYVRK

198	P27064	227	AEAIFK S QAET
199	P0A7J7	92	KSGSGK P NKDK
200	Q5SLP6	9	VVAVV K LQLPA
201	Q5SLP6	16	QLPAG K ATPAP
202	Q5SLP6	70	FTFVT K TPPAS
203	P0CX53	41	GLSPK K VGEDI
204	P62986	114	NCRK K KCGHTN
205	P55072	20	AILKQ K NRPNR
206	P55072	190	EGEPI K REDEE
207	P55072	211	IGGCR K QLAQI
208	P55072	231	HPALF K AIGVK
209	P05547	21	DIDR K KAEVRK
210	P05547	26	KAEVR K RLEEQ
211	P05547	46	MTPER K KKLRL
212	O95785	54	YLPVT K EGPRD
213	O95785	128	PRLLE K HQAQR
214	O95785	277	CSIYF K QKEHL
215	O95785	517	DFPLS K PLLHG
216	O95785	702	KHEER K CPYCP
217	O95785	754	FSFQ K KKKVA

(2) Subset 2 (217 samples)

Sample #	Protein ID	Site	Sequences
1	P05141	10	AVSFA K DFLAG
2	P26358	160	PRITR K STRQT
3	P26358	179	GPAKR K PQEEES
4	P26358	203	KDQDE K RRRVT
5	O75530	89	CVNSL K EDHNQ
6	O75530	184	TMQCI K HYVGH
7	O75530	293	PFISQ K IHFDP
8	O75530	322	DLILS K SCENA
9	Q5VTE0	146	YTLGV K QLIVG
10	Q5VTE0	179	VSTYI K KKIGYN
11	Q5VTE0	215	WFKGW K VTRKD
12	Q5VTE0	255	LQDVY K IGGIG
13	P0CE47	249	EIVGI K ETQKS
14	P0CE47	300	KPGTI K PHTKF
15	P0CE47	304	IKPHT K FESEV
16	P0CE47	314	VYILS K DEGGR
17	O75367	8	SRGG K KKSTKT
18	P62807	13	APAP K KGSKKA
19	P62807	24	VTKAQ K KDGKK
20	P62807	29	KKDGG K KRRSR
21	P62807	31	DGKKR K KRSRKE
22	P62807	86	LAHYN K RSTIT
23	Q55BN9	64	SELLI K KLFPQ
24	Q55BP0	64	SDLLI K KLFPQ
25	P62803	80	EHAKR K TVTAM
26	Q86L05	90	DIEAL K KIGPK
27	Q86L05	95	KKIGP K NKKAI

28	Q7M1B9	18	AEKVLKSKTPV
29	P14909	36	KTKKIKIIDFG
30	P14909	263	REIIQKMGILA
31	P14909	301	MVSLFKKRRDV
32	P14909	315	ELTKVKGVEVS
33	P14909	321	GVEVSKPNGAF
34	P22498	72	HDNAQKMGLKI
35	P22498	102	NFDESKQDVTE
36	Q9Y232	204	PKSPVKSRTAV
37	Q9Y232	348	DIVVRKQDGFT
38	Q9Y232	360	ILLSTKSSENN
39	P80053	20	VKKLYKVGELL
40	P80053	97	MIMTWKNSLLL
41	P80053	109	PYGGGKGGVRV
42	P80053	137	IQAIYKYLSE
43	P80053	313	PKVKAKLIVEG
44	P61991	21	DISKIKKVWRV
45	P26358	302	RRPEEKKEPEKV
46	O75530	335	CWKPGKMEDDI
47	O75530	344	DIDKIKPSESN
48	Q5VTE0	376	KFAELKEKIDR
49	P55907	74	HPASEKKADPV
50	P81543	20	THEGHKVYGPV
51	P49949	20	EHQGHKVYGPV
52	O43524	195	CVPYFKDKGDS
53	Q08369	328	NLNKSKTPAGP
54	Q08369	402	SAVCSKLSPOG
55	P10412	17	PAPAEKTPVKK
56	P10412	127	KAGAAKAKKPA
57	P16401	22	KSPAKKATKK
58	P16401	26	KKKATKKAAGA
59	P16401	37	GAAKRKATGPP
60	P16401	143	TPKKAKKAAGA
61	P16401	150	AAGAKKAVKKT
62	O75367	235	GNTLEKKGKKE
63	P40283	43	GKKLPKEAGAG
64	P40283	111	LARYNKKPTIT
65	Q41811	17	GKGARKRHRKV
66	O14979	76	GGAAIKGGRRR
67	O14979	142	KINASKNQDD
68	O14979	149	QQDDGKMFIGG
69	Q14103	67	ESEGAIDASK
70	Q14103	111	WDTTKDLKDY
71	Q14103	129	VDCTLKLDPIT
72	Q14103	146	GFVLFKESESV
73	Q14103	176	RAKAMKTKEPV
74	Q14103	178	KAMKTKEPVKK
75	Q14103	255	SKCEIKVAMSK
76	P16862	69	DEKINKNVEAH
77	P16862	323	IEELLKTNRIS
78	P16862	336	QYERMKHLNIC
79	P16862	536	LMESVKLTKAV

80	P16862	539	SVKLT K AVAEA
81	P16862	588	KDKRL K IAIVN
82	O00159	527	HFLTH K LADQR
83	P13538	370	LKFKQ K QREEQ
84	Q96RI1	132	SAGRI K GDEL C
85	Q96RI1	172	KNAVY K CKNGG
86	Q96RI1	353	EIFNK K LPSGH
87	Q96RI1	434	LQKLC K IHQPE
88	P11940	104	GNIFI K NLDKS
89	P11940	213	KDLFG K FGPAL
90	P11940	254	DEMNG K ELNGK
91	P06400	65	LCQK L KIPDHV
92	P06400	130	EISVH K FFNLL
93	P06400	192	SALVL K VSWIT
94	P06400	279	IEVLC K EHECN
95	P06400	319	VENLS K RYEEI
96	P06400	427	IGYIF K EKFAK
97	P06400	429	YIFKE K EFAKAV
98	Q86L05	156	VKSTV K FQLKK
99	Q86L05	207	KTLYV K TSMGP
100	P0CX43	39	LQVGL K NYDPQ
101	P0CX43	97	KLNKN K KLIKK
102	P0CX43	98	LNKN K KLIKKL
103	P0CX43	102	KKLIK K LSKKY
104	P47914	22	HRNGI K KPRSQ
105	P47914	23	RNGIK K KPRSQR
106	P47914	48	MRF A KKH N KKG
107	P47914	92	KGVS R KLDRLA
108	P47914	124	CRP K AKAKA
109	P47914	149	PAQAP K RTOAP
110	Q969Q0	44	RRYDR K QSGYG
111	Q969Q0	100	GDKKR K GQVIQ
112	P0CX27	78	ECVKC K TRAQL
113	P02393	6	MSSIT K EQVVE
114	P02393	72	AAGAN K IGVIK
115	P0A7K2	30	SAMEE K FGVSA
116	P0A7K2	109	AEAL K KALEEA
117	P07472	110	DGDEA K TKLEE
118	Q12962	42	TAAEN K ASPAG
119	Q12962	213	GINVK K PHYFT
120	Q04207	56	STD T TK T HPTI
121	Q04207	93	HELVG K DCRDG
122	Q53688	185	TLMSW K NPPSY
123	Q53688	417	DLRYY K ISPDQ
124	E5K W G9	296	QKRVE K LLRS D
125	B6A879	11	TYTSS K AMSDI
126	B6A879	83	ADEIN K IGNPT
127	B6A879	374	YVLLI K ELRSA
128	B6A879	552	LFSE A KGHFIS
129	Q9L5X8	119	RFV T Q K AFAHG
130	Q9L5X8	287	GLGEL K ISDTI
131	Q9L5X8	341	LERLE K ELVHN

132	D0ZRV8	13	LIELVKKVNRN
133	D0ZRV8	194	WKAQQKMNEQR
134	Q6NEZ3	80	MPESVKAGTGF
135	Q6NEZ3	105	NLSGYKNAKLT
136	Q6NEZ3	111	NAKLTKSSIRI
137	Q6NEZ3	202	TTNADKPFATV
138	Q6NEZ3	369	APEDVKYIDGT
139	Q6NEZ3	407	WKNESKKNES
140	Q6NEZ3	461	YSNSLKDLVSI
141	D1BQI7	38	YGRQLKTRIPP
142	Q9UKV3	143	KOYLEKQOELL
143	Q9UKV3	223	DEKPRKGERRS
144	Q9UKV3	315	EIPRVKPEEMM
145	P18281	59	AETVDKSGPAI
146	Q9SJU4	15	KASPVKSDWVK
147	Q9SJU4	20	KSDWVKGQSL
148	Q9SJU4	131	STTDGKMMVDV
149	Q9SJU4	132	TTDGKMMVDVL
150	Q9SJU4	348	QNTCLKTWGGK
151	Q9ZU52	271	PGAEHKNKASP
152	Q9ZU52	348	QGKPEKIEASQ
153	P62157	14	QIAEFKEAFSL
154	P07463	14	QIAEFKEAFAL
155	P07463	22	FALFDKDGDT
156	P07463	78	MARKMKEQDSE
157	P04573	67	GPKGDKKNIGP
158	P04573	68	PKGDKKNIGPE
159	P04573	112	DFDEFKFIMQK
160	P00889	14	RLFGAKNASCL
161	P00889	34	SSTNLKDILAD
162	P00889	233	GAIDSKLDWSH
163	P00889	317	LTQLQKEVGKD
164	P00889	327	DVSDEKLRDYI
165	P00889	366	REFALKHLPD
166	P00889	393	LLEQKAKNPW
167	P00063	21	KIFKTKCAQCH
168	P00078	61	GYAYSKANAES
169	P00078	97	SFAGMKKQER
170	P00077	33	CHTGAKGGANG
171	P00077	97	SFAGIKKQER
172	P00043	7	PAPYEKGSEKK
173	P00076	8	AERGGKLFESR
174	P00076	72	LENPKKYVPGT
175	P00076	88	GIKAKKDRQDI
176	P00047	87	FIPGTKMAFGG
177	P02993	244	SRPTEKPLRLP
178	P02994	129	EAGISKDGQTR
179	P02994	154	IVAVNKMDSVK
180	P02994	170	FQEIVKETSNF
181	P02994	184	VGYNPKTVPFV
182	Q96KQ7	179	PEGQPKVHRAR
183	Q96KQ7	253	EVTLTGDPGS

184	Q96KQ7	288	VDSDSKSEVEA
185	Q96KQ7	344	RRKAKKKWRKD
186	P53991	101	IETEGKIPFWE
187	P53991	282	EQNNRKGKMY
188	P10587	33	DWSAKKLVWVP
189	P10587	67	QENGGKVTLSK
190	P10587	82	KMNPPKFSKVE
191	P10587	183	ESGAGKTENTK
192	P10587	189	TENTKKVIQYL
193	P11416	109	ACEGCKGFFRR
194	P11416	128	TCHRDKNCIIN
195	P11416	134	NCIINKVTRNR
196	P11416	171	KKEAPKPECSE
197	P11416	227	IDLWDKFSELS
198	P11416	234	SELSTKCIKT
199	P27064	164	RDKLNKYGRPL
200	P27064	316	FRVLAKALRMS
201	P27064	466	VWKEIKFEFEA
202	P04717	252	CEEMLKRAVFA
203	P04717	356	DDYIKKDRSRG
204	P04992	21	GVKEYKLTYYT
205	P04992	252	CEEMMKRAVFA
206	P04992	463	ACEVWKEIRFN
207	Q5SLP6	39	IMEFVKAFNAA
208	Q5SLP6	109	VLEIAKQKMPD
209	P0CX53	96	DRKKDKNVKHS
210	P0CX53	99	KDKNVKHSNGI
211	P62986	27	TIENVKAKIQD
212	P62986	29	ENVKAKIQDKE
213	P62986	112	AVNCRKKKCGH
214	P55072	136	FEVYLPYPFLE
215	O95785	105	IVRTMKPHAEL
216	O95785	373	YVQHAKLHMRE
217	O95785	479	VDYFGKAEPST

(3) Subset 3 (217 samples)

Sample #	Protein ID	Site	Sequences
1	P05141	33	PIERVKLLLQV
2	P05141	92	LNFAFKDKYKQ
3	P05141	166	LVKIYKSDGIK
4	Q9Y232	17	GKSRKKNWQYE
5	P00042	6	PAPFKKGSEKK
6	P00042	28	CHTVEKGGPHK
7	P00042	95	GLKKEKDRNDL
8	P26358	29	VRRRLKDLERD
9	P26358	45	ECVKEKLNLLH
10	P26358	67	CDLETCLRKEE
11	P26358	81	EGYLAVKSLI

12	P26358	130	ANSPPKPLSKP
13	P26358	148	SDGEAKPEPSP
14	O75530	70	RKSWGKWKWKS
15	O75530	72	SWGKWKWKS
16	O75530	77	KWKSCKYYSF
17	O75530	250	DLLGEKIMSCG
18	Q5VTE0	20	HVDSGKSTTTG
19	Q5VTE0	51	AAEMGKGSFKY
20	Q5VTE0	84	KFETSKYYVTI
21	Q5VTE0	154	IVGVNKMDSTE
22	Q5VTE0	172	YEEIVKEVSTY
23	P0CE47	137	IVFLNKCDMVD
24	P0CE47	283	LLRGIKREEIE
25	P0CE47	325	HTPFFKGYRPQ
26	O75367	96	LNQLLKGVTIA
27	O75367	142	KSPSQKPVSK
28	O75367	143	SPSQKPVSKK
29	O75367	147	KKPVSKKAGGK
30	O75367	152	KKAGGKKGARK
31	O75367	153	KAGGKKGARKS
32	P62807	16	PKKGSKAVTK
33	P62807	17	KKGSKAVTKA
34	O15819	126	VTIMPKDIHLA
35	Q55BN9	123	VTIMVKDIQLA
36	P61830	38	TGGVKKPHRYK
37	P04637	24	FSDLWKLLPEN
38	P04637	139	FCQLAKTCPVQ
39	Q86L05	46	DTKKDKRFSGQ
40	Q86L05	78	CDEATKIGAEF
41	P80322	37	GDISTKELGTI
42	P14909	31	ARNVEKTKKIK
43	P14909	137	YAEVVKLLGGK
44	P39462	313	LAEAGKVKPMI
45	Q9Y232	312	GVTASKRKFID
46	Q9Y232	314	TASKRKFIDDR
47	P0C835	6	MSSGKAVKVK
48	P0C835	48	DPETGKYFRHK
49	P80053	45	RIIQVKIQIRG
50	P80053	231	GYAGKFLSEM
51	P80053	265	KAIEIKEKTGS
52	P13123	22	TSKIKKVVRVG
53	P61991	13	YKGEEKQVDIS
54	P61991	53	EKDAPKELLQM
55	P39476	13	YKGEEKQVDIS
56	P26358	240	EERDEKEEKRL
57	P26358	243	DEKEEKRLRSQ
58	P26358	255	KEPTPKQKLKE
59	P26358	259	PKQKLKEEPDR
60	P26358	281	EDGDEKDEKKH
61	P26358	285	EKDEKHKRSQP
62	P26358	296	KDLAAKRRPEE
63	Q5VTE0	273	ETGVLPGMVV

64	P81539	13	SLPVV K EEQGV
65	P81539	19	EEQGV K IYGTY
66	P49949	73	HPASE K KADPI
67	O43524	270	GRAAK K KAALQ
68	O43524	455	TIQEN K PATFS
69	O43524	569	SLGSA K HQQQS
70	Q08369	320	QTRKR K PKNLN
71	Q08369	326	PKNLN K SKTPA
72	P16403	26	KKKAA K KAGGT
73	P16403	136	PVGAA K KPKKA
74	P16403	152	PKKSA K KTPKK
75	P16403	156	AKKTP K KAKKP
76	P10412	32	SAGAA K RKASG
77	P10412	139	AAK K PKATGA
78	P16401	67	LAAL K KALAAG
79	P16401	141	GATP K KAKKAA
80	O75367	160	ARKS K KQGEV
81	O75367	196	LFLG Q KLNLIH
82	O75367	250	VLELR K KNGPL
83	P40283	40	PKAG K KLPKEA
84	P40283	58	KKMK K SVETY
85	P40283	98	NDIFE K LAQEA
86	O14979	137	FAEGS K INASK
87	O14979	212	ELKEH K LDGKL
88	O14979	224	DPKRA K ALKGK
89	Q14103	242	KKIME K KYHNV
90	P16862	33	IDFYT K FLSLE
91	P16862	106	DILAV K DTLNA
92	P16862	144	GVTST K NAVST
93	P16862	276	CMEF K KREGRL
94	P16862	551	QAKDF K RAMSL
95	O00159	289	WKVVR K KALTVI
96	O00159	444	NYCNE K LQQLF
97	O00159	627	YVRCI K PNDAK
98	P13538	50	KESFV K GTIQS
99	P13538	56	G T IQS K EGGKV
100	P13538	206	AASGE K KKKEEQ
101	P13538	273	TYLLE K SRVTF
102	P13538	366	HYGNL K FKQKQ
103	Q96RI1	158	TCEG C KGFFRR
104	Q96RI1	174	AVYK C KNGGNC
105	Q96RI1	224	SKRLR K KNVKQH
106	Q96RI1	384	MFSFY K SIGEL
107	Q96RI1	390	SIGEL K MTQEE
108	P11940	80	DVIK G KPVRIM
109	P11940	129	NILS C KVVCDE
110	P11940	229	TDESG K SKGFG
111	P11940	231	ESGKS K GFGFV
112	P11940	333	EGGRS K GFGFV
113	P11940	348	PEEAT K AVTEM
114	P06400	154	SRL L KKYDVLV
115	P06400	447	GSQRY K LGVRL

116	P0CX43	18	VKELLKYSNET
117	P0CX43	105	IKKLSKKYNAF
118	P47914	15	HNQSRKWHRNG
119	P47914	63	QANNAKAMSAR
120	P47914	73	RAEAIKALVKP
121	P47914	126	PKAKAKAKAKD
122	P47914	130	AKAKAKDQTKA
123	Q969Q0	13	RRTFCCKCGKH
124	Q969Q0	14	RTFCCKCGKHQ
125	Q969Q0	22	KHOPHKVTOYK
126	Q969Q0	27	KVTOYKKGKDS
127	Q969Q0	97	ELGGDKKRKGQ
128	P0CX27	6	MVNVPKTRKTY
129	P0CX27	9	VPKTRKTYCKG
130	P0CX27	13	RKTYCKGKTCR
131	P0CX27	29	KVTOYKAGKAS
132	P02393	31	KELEEKFGVSA
133	P02393	102	APSTLKEAVSK
134	P0A7K2	66	AAGANKVAVIK
135	P0A7K2	96	APAALKEGVSK
136	Q04207	62	THPTIKINGYT
137	Q53688	216	QESHSKILDLD
138	Q53688	250	SIIKNKIIIVE
139	Q53688	302	NFTAEEKISISE
140	Q53688	429	HVFNQKRRGKI
141	E5KWG9	16	YSIIEKTAQVV
142	E5KWG9	84	HGLTPKTDHPV
143	E5KWG9	258	FERLIKSAFYD
144	B6A879	148	HYVGDKTKPKL
145	B6A879	152	DKTKPKLSAYI
146	Q9L5X8	54	SNDIEKERGIT
147	Q9L5X8	63	ITILAKNTAIN
148	Q9L5X8	133	IVVINKIDRPG
149	Q9L5X8	254	KTRNGKVGTVM
150	Q9L5X8	427	NIGLRKGELKD
151	Q9L5X8	545	NALKGKQLTNV
152	D0ZRV8	14	IELVKKVNRNK
153	Q6NEZ3	92	ALIKIKDISVS
154	Q6NEZ3	208	PFATVKVDPAD
155	Q6NEZ3	238	ATVPEKLNEKM
156	Q6NEZ3	242	EKLNEKMDGKV
157	Q6NEZ3	311	TELDTKKPEDT
158	Q6NEZ3	403	YDPEWKNESKK
159	Q9UKV3	58	PSTSRKMAELE
160	Q9UKV3	70	VTLDGKPLQAL
161	Q9UKV3	91	QRGLAKSGQKS
162	P18281	71	EDVHVKKVDRG
163	Q9SJU4	10	TATMLKASPVK
164	Q9SJU4	58	ADELVKTAKTI
165	Q9SJU4	150	GIKVDKGLVPL
166	Q9SJU4	378	LAQLGKYTGEG
167	Q9ZU52	238	LEVAEKVWSEV

168	P11118	144	YEEFVKMMMSK
169	P04573	81	LTLCSKWVRQD
170	P04573	153	FMDLIKSKNA
171	P00044	11	AGSAKKGATLF
172	P00063	35	KGAGHKQGNL
173	P00043	11	EKGSEKKGANL
174	P00070	17	PTTGEKIFKTK
175	P02993	84	KFETAKYYVTI
176	P02993	171	RFEIIEKKEVSA
177	P02994	44	IEKFEKEAAEL
178	P02994	178	SNFIKKVGYNP
179	Q96KQ7	53	EETLPKATPDS
180	Q96KQ7	122	PSSPSKGGSCP
181	Q96KQ7	229	LAKRRKLNSGG
182	Q96KQ7	348	KKKWRKDSPWV
183	Q96KQ7	354	DSPWVKPSRKR
184	Q96KQ7	361	SRKRRKREPPR
185	P53991	73	NKAPFKAKVRS
186	P53991	251	GNSDAKLYDEE
187	P53991	262	LQAIKAYPGQ
188	P10587	85	PPKFSKVEDMA
189	P10587	142	IIDMYKGGKRH
190	P68696	81	RSIYGKKSAG
191	P68696	82	SIYGKKSAGV
192	P68696	94	TVKTSKSILVG
193	P68696	116	ANVVEKLADYL
194	P11416	155	DVGMSKESVRN
195	P11416	207	LCQLGKYTTNN
196	P27064	146	PTAYIKTFQGP
197	P27064	463	ACEVWKEIKFE
198	P04717	146	PYAYVKTFOGP
199	P0A7J7	81	AAVLLKKAAGI
200	P0A7J7	97	KPNKDKVGKIS
201	P0A7J7	113	EIAQTKAADMT
202	Q5SLP6	86	AAGLEKGAHKP
203	P0CX53	48	GEDIAKATKEF
204	P0CX53	57	EFKGIKVTVQL
205	P62986	11	KTLTGKTITLE
206	P62986	48	LIFAGKQLEDG
207	P62986	93	KYNCDKMICRK
208	P55072	8	SGADSKGDDL
209	P55072	18	STAILKQKNRP
210	P55072	45	SLSQPKMDELQ
211	P55072	60	DTVLLKGGKRR
212	P55072	62	VLLKGKRRREA
213	P55072	312	DAIAPKREKTH
214	P05547	11	AAEEAKKQDD
215	P05547	34	EEQSLKQKKG
216	P05547	63	AEELKKEQERK
217	O95785	339	IEEIQKQKQVP

(4) Subset 4 (217 samples)

Sample #	Protein ID	Site	Sequences
1	P05141	43	VQHASKQITAD
2	P05141	163	GDCLVKIYKSD
3	P05141	171	KSDGIKGLYQG
4	P05141	206	MLPDPKNTHIV
5	Q9Y232	116	RRHTEKQKEST
6	P00042	33	KGGPHKVGPNL
7	P26358	188	ESERAKSDESI
8	P26358	233	GTRTEKKEERD
9	Q5VTE0	100	HRDFIKNMITG
10	P0CE47	38	TTVLAKTYGGA
11	P0CE47	253	IKETQKSTCTG
12	O75367	7	SSRGGKKKSTK
13	O75367	121	KKRGSKGKLEA
14	O75367	137	PAKKAKSPSQK
15	P69150	54	LREIRKYQKST
16	P69150	123	VTIMTKDMQLA
17	P59226	57	IRKYQKSTELL
18	P62803	60	TRGVLKVFLN
19	P62803	78	YTEHAKRKTVT
20	P62803	92	VVYALKRQGRT
21	P04637	319	SSPQPKKKPLD
22	P04637	320	SPQPKKKPLDG
23	P04637	357	DAQAGKEPGGS
24	Q86L05	61	TVTCKPKLSVCV
25	Q86L05	133	LNKVGKFPDLL
26	Q7M1B9	57	RLTIAKVNTDD
27	P80322	6	SDDYVKARVMF
28	P80322	12	ARVMFKEEQIS
29	P14909	51	DLPTFKRIRDA
30	P14909	58	IRDAAKEALDQ
31	P14909	169	QSKISKRTKMI
32	P14909	242	VNGFSKTFST
33	P14909	354	KLIEEKGVVTI
34	P39462	22	EIGVCPKGPQ
35	P39462	73	HEIAGKIEEVG
36	P39462	239	RITESKGVDAV
37	P39462	259	LSVYPKALAKQ
38	P39462	315	EAGKVKPMITK
39	P39462	338	NLENFKAIGRQ
40	P22498	35	NTDWYKVVHDP
41	P22498	251	AYDGIKSVSKK
42	P0C835	9	GKKAVKVKTPA
43	P80053	17	TQQVKKLYKVG
44	P80053	207	KEAANKFIGGV
45	P80053	304	ENVINKFNAPK
46	P80053	415	VVRAMKARGIL
47	P13123	13	YKGEEKVDTS
48	P13123	21	DTSKIKKVVWRV

49	P13123	39	YDDNGKTGRGA
50	P13123	48	GAVSEKDAPKE
51	P26358	250	LRSQTKIPTPK
52	O75530	375	MDFWQKMLALG
53	Q5VTE0	392	LEDGPKFLKSG
54	P55907	75	PASEKKADPVN
55	O43524	259	MDNSNKYTKSR
56	Q08369	318	GIQTRKRKPKN
57	P16403	21	EKAPVKKKAAK
58	P16403	63	SLAALKKALAA
59	P16403	97	TLVQTKGTGAS
60	P16403	122	KPKVKKAGGTK
61	P16403	127	KAGGTKPKKPV
62	P16403	129	GGTKPKKPVGA
63	P10412	23	TPVKKKARKSA
64	P10412	129	GAAKAKKPAGA
65	P16401	133	AAKAKKPAGAT
66	O75367	226	ADIDLKDDLGN
67	O75367	239	EKKGGKEFVEA
68	P40283	36	AEKKPKAGKKL
69	P40283	57	KKMKKKSJET
70	P40283	69	KIYIFKVLKQV
71	P40283	142	VSEGTKAVTKF
72	Q41811	80	EHARRKTVTAM
73	O14979	209	KVLELKEHKLD
74	O14979	250	SEEQIKYFGA
75	O14979	269	LPMDTKTNERR
76	O14979	288	DEEPVKLLES
77	O14979	302	QIGSGKCEIKV
78	Q14103	98	QREEWKFIGG
79	Q14103	153	SESVDKVMQK
80	Q14103	158	KVMDQKEHKLN
81	Q14103	218	LPMDNKTNKRR
82	Q14103	341	QSGYGKVSRRG
83	P16862	58	DSISLKILLRP
84	P16862	225	RSALFKGCRAF
85	P16862	248	GPEYIKFHWI
86	P16862	432	CDIVSKHRSRG
87	P16862	585	KLPKDKRLKIA
88	O00159	281	SSINDKSDWKV
89	O00159	361	RKIIAKGEELL
90	O00159	455	IELTLKSEQEE
91	O00159	561	TGFLDKNNDLL
92	O00159	578	TMCSSKNPIMS
93	P13538	31	IEAQNKPFDK
94	P13538	88	PPKYDKIEDMA
95	P13538	208	SKEKKKEEQSG
96	P13538	297	IMSNKKPELID
97	P13538	386	TEVADKAAYLM
98	P13538	400	SAELLKALCYP
99	P11940	108	IKNLDKSIDNK
100	P11940	167	LLNDRKVFVGR

101	P11940	221	PALSVK V MTDE
102	P11940	246	HEDAQ K AVDEM
103	P11940	312	DERLR K EFSPP
104	P11940	371	ALAQR K EERQA
105	P11940	512	TVPOY K YAAGV
106	P06400	94	GGYIQ K KKELW
107	P06400	236	PPMLL K EPYKT
108	P06400	417	KESIL K RVKDI
109	Q86L05	195	LVSL L KKGWQN
110	P0CX43	54	FSGSL K LPNCP
111	P0CX43	92	VDDL K KLKNK
112	P0CX43	133	LSKAG K FPTPV
113	P0CX43	160	IKFQL K KVLCL
114	P0CX43	196	VSLL K KNWQNV
115	P47914	87	KPKIP K GVSRK
116	P47914	106	HPKL G KRARAR
117	P47914	114	RARIA K GLRLC
118	Q969Q0	30	QYKKG K DSLYA
119	Q969Q0	64	KAKTT K KIVLR
120	Q969Q0	65	AKTT K KIVLRL
121	Q969Q0	80	PNCRS K RMLAI
122	P0CX27	32	QYKAG K ASLFA
123	P0CX27	63	FHKKA K TTKKV
124	P0CX27	66	KAKTT K KVVLR
125	P0CX27	98	LGGE K KQKGOA
126	P02393	26	LSEFI K ELEEK
127	P07472	103	KEGMS K EDGDE
128	Q12962	175	ALQH C KMKGTA
129	Q04207	37	MRFRY K CEGRS
130	Q04207	123	IQC V KKRDLEQ
131	Q04207	315	KSIM K KSPFNG
132	Q53688	140	DTVIS K GLLKI
133	Q53688	144	SKGL L KIVKDG
134	Q53688	265	FQEEL K LNSDG
135	Q53688	379	PEAYT K LQOYM
136	E5K W G9	68	FTML P KGLDPY
137	E5K W G9	94	VGSL L KEVHEN
138	E5K W G9	126	FPSAE K LGKVS
139	E5K W G9	168	GIDYS K RMTNL
140	B6A879	210	GDIGN K SKKVQ
141	B6A879	298	HAMGH K LDLAF
142	B6A879	402	SGVKA K LKKS
143	B6A879	565	TPRTV K QKGEY
144	Q9L5X8	22	HVDHG K TTLVD
145	Q9L5X8	165	DQLDF K VVYAS
146	Q9L5X8	236	TRGSV K ANQQV
147	Q9L5X8	359	TDDPD K FRVSG
148	D0ZRV8	76	IETLQ K SIDNE
149	Q6NEZ3	90	FDALI K IKDIS
150	Q6NEZ3	280	DTSGS K TITAR
151	Q6NEZ3	312	ELDT K KPEDTD
152	Q6NEZ3	408	KNESK K PNESY

153	Q9UKV3	41	EGRWRKLPKMP
154	Q9UKV3	81	RVTDLKAALEQ
155	Q9UKV3	113	LENLQKHSTPH
156	Q9UKV3	138	QNSFIKQYLEK
157	Q9UKV3	220	DSDDDEKPRKGE
158	Q9UKV3	349	QEEARKSHLAR
159	P18281	23	AETVDKSAPQI
160	Q9SJU4	358	KEENVKAAQDI
161	Q9SJU4	369	LLARAKANSLA
162	Q9SJU4	389	ESEEAKEGMFV
163	Q9ZU52	53	LVKTAKSIASP
164	Q9ZU52	273	AEHKNKASPET
165	Q9ZU52	345	RTWQGKPEKIE
166	Q9ZU52	354	IEASQKALLVR
167	Q9ZU52	361	LLVRAKANSLA
168	P62157	78	MARKMKDSE
169	P04573	50	GQTYTKRETEA
170	P04573	156	LIKKSKNALKE
171	P00889	107	QKMLPKAKGGE
172	P00889	321	QKEVGKDVSD
173	P00044	60	TDANIKKNVLW
174	P00044	61	DANIKKNVLWD
175	P00078	22	GEKLFKGRAAQ
176	P00077	61	GFAYSKANADS
177	P00043	96	GLKKAADRNDL
178	P00070	64	AGNKNKAVIWE
179	P00047	61	YTDANKQAGIT
180	Q71V39	215	WFKGWKVERKE
181	P02993	222	ERKEGKADGKT
182	P02994	84	KFETPKYQVTV
183	Q96KQ7	28	ALLLEKETRGA
184	Q96KQ7	200	PPVPEKRPPEI
185	Q96KQ7	219	VHSLGKVTSDL
186	Q96KQ7	226	TSDLAKRRKLN
187	P53991	169	KEDPAKGLCS
188	P53991	194	TGPTGKVLLLP
189	P10587	50	EAASIKEEKGD
190	P10587	66	LQENGGKVTLS
191	P10587	77	KDDIQKMNPPK
192	P10587	136	PIYSEKIIDMY
193	P11416	193	IEKVRKAHQET
194	P11416	238	TKCIIKTVEFA
195	P27064	21	GVKDYKLTYYT
196	P27064	128	NVFGFKALRAL
197	P04717	81	SLDRYKGRCE
198	P04717	316	FRVLAKALRLS
199	P04992	305	VIDRQKNHGIH
200	P0A7J7	45	KAFNAKTDSIE
201	P0A7J7	82	AVLLKKAAGIK
202	Q5SLP6	80	SYLIRKAAGLE
203	Q5SLP6	111	EIAKQKMPDLN
204	P0CX53	54	ATKEFKGIKVT

205	P0CX53	63	VTVQLKIQNRQ
206	P62986	63	DYNIQKESTLH
207	P62986	113	VNCRKKKCGHT
208	P55072	63	LLKGKKRREAV
209	P55072	81	TCSDEKIRMNR
210	P55072	164	RAVEFKVVETD
211	P55072	277	PEIMSKLAGES
212	P55072	288	ESNLRKAFEEA
213	P05547	20	DDIDRKKAEVR
214	P05547	55	RLLLRKKAEE
215	P05547	99	LRAIKEYFDH
216	O95785	341	EIQKLKQVPGD
217	O95785	743	SAEEVKAIERR

(5) Subset 5 (216 samples)

Sample #	Protein ID	Site	Sequences
1	Q9Y232	27	EGPTQKFLFKR
2	Q9Y232	84	YLVRWKGYDSE
3	Q9Y232	118	HTEKQKESTLT
4	Q9Y232	185	SLSSRKNMDLA
5	P00042	10	KKGSEKKGATL
6	P26358	88	KSLLNKDLSE
7	P26358	194	SDESIKEEDKD
8	O75530	20	MPAAKKQKLSS
9	O75530	79	KSKKCKYSFKC
10	O75530	83	CKYSFKCVNSL
11	O75530	275	MMNAIKESYDY
12	Q5VTE0	30	GHLIYKCGGID
13	Q5VTE0	62	AWVLDKKAER
14	P0CE47	188	AEWEAKILELA
15	P0CE47	264	VEMFRKLLDEG
16	O75367	73	ARDNKKGRVTP
17	P62807	21	KKAVTKAQKKD
18	P62807	28	QKKGDKKRKRS
19	P61830	116	AAIHAKRVTIQ
20	P61830	122	RVTIQKDKIKL
21	P62803	9	RGKGGKGLGKG
22	P62803	45	RRGGVKRISGL
23	P04637	120	HSGTAKSVTCT
24	P04637	351	EALELKDAQAG
25	Q86L05	38	LQINLKNYDTK
26	Q86L05	44	NYDTKKDKRFS
27	Q86L05	59	IGTVTKPKLSV
28	Q86L05	91	IEALKKIGPKN
29	Q86L05	101	NKKAIKLSKK
30	Q86L05	106	KKLSKKYDAFL
31	Q7M1B9	20	KVLKSKTPVVV
32	Q7M1B9	44	APILDKLAGFY
33	Q7M1B9	82	TLVIFKDGREV

34	P80322	104	FRVLDKNGDGF
35	P80322	112	DGFIDKDEFRA
36	P14909	23	TLLYKEIARN
37	P14909	172	ISKRTKMIVFN
38	P14909	302	VSLFKKRRDVM
39	P22498	124	DEYANKDALNH
40	P22498	328	TRTVVKRTEKG
41	P22498	451	VDYNTKRLYWR
42	P0C835	11	KAVKVKTPAGK
43	P0C835	24	ELVPEKVVWALA
44	P80053	117	VRVDPKKL TRE
45	P80053	166	LDEYIKITGKV
46	P13123	19	EVDTSKIKKVW
47	P61991	9	VKFKYKGEEKQ
48	P61991	22	ISKIKKVWRVG
49	P13125	9	VRFKYKGEEKE
50	P13125	52	EKDAPKELMDM
51	P39476	19	EVDISKIKKVW
52	P26358	257	PTPKQKLKEEP
53	O75530	342	EDDIDKIKPSE
54	O75530	408	TLTHHKCGAAI
55	Q5VTE0	290	VTTEVKSVMH
56	Q5VTE0	313	VGFKVKNVSVK
57	Q5VTE0	330	VAGDSKNDPPM
58	Q5VTE0	386	RRSGKKLEDGP
59	P49949	74	PASEKKADPIN
60	O43524	663	NFTGAKQASSQ
61	Q08369	245	CGLYHKMNGIN
62	Q08369	366	EMRPIKTEPGL
63	P16403	64	LAALKKALAAA
64	P16403	106	ASGSFKLNKKA
65	P16403	119	GEAKPKVKKAG
66	P16403	140	AKKPKKAAGGA
67	P10412	21	EKTPVKKKARK
68	P10412	119	GEAKPKAKKAG
69	P10412	122	KPKAKKAGAAK
70	P16401	17	PAPVEKSPAKK
71	P16401	149	KAAGAKKAVKK
72	P16401	153	AKKAVKKT PPK
73	O75367	161	RKSKKKQGEVS
74	O75367	167	QGEVSKAASAD
75	P40283	8	PKAEKKPAEKK
76	P40283	18	KPASEKPVEEK
77	P40283	39	KPKAGKLPKE
78	P40283	54	GDKKKKMKKKS
79	P40283	64	SVETYKIYIFK
80	P40283	72	IFKVLKQVHPD
81	P40283	105	AQEASKLARYN
82	O14979	221	KLIDPKRAKAL
83	O14979	234	KEPPKVFVGG
84	Q14103	72	KIDASKNEEDE
85	Q14103	110	SWDTTKDLKD

86	Q14103	170	KVIDP K RAKAM
87	Q14103	182	TKEPV K KIFVG
88	Q14103	197	DTPEEK I REYF
89	Q14103	221	DNKTN K RRGFC
90	Q14103	237	EEEPV K KIMEK
91	Q14103	243	KIMEK K YHNVG
92	Q14103	251	NVGLS K CEIKV
93	P16862	275	RCMEF K KREGR
94	P16862	369	LDRIC K AIDYV
95	P16862	529	NKIVR K PLMES
96	P16862	583	EPKLP K DKRLK
97	O00159	151	KTEAT K RLLQF
98	O00159	224	SYLLE K SRVVH
99	O00159	336	TENQL K YLTRL
100	O00159	404	RSLAS K DVESP
101	O00159	477	QYFNN K IICDL
102	O00159	571	LFRNL K ETMCS
103	P13538	23	LRKSE K ERIEA
104	P13538	108	VLYNL K ERYAA
105	P13538	214	EEQSG K MOGTL
106	Q96RI1	188	MYMRR K CQECR
107	Q96RI1	227	LRKNV K QHADQ
108	Q96RI1	317	LVEFT K KLPGF
109	Q96RI1	318	VEFT K KLPGFQ
110	P11940	196	TNVYI K NFGED
111	P11940	259	KELNG K QIYVG
112	P11940	324	TITSA K VMMEG
113	P11940	361	RIVAT K PLYVA
114	P06400	63	TALCQ K LKIPD
115	P06400	95	GYIQ K KELWG
116	P06400	153	MSRLL K KYDVL
117	P06400	265	SARIA K QLEND
118	P06400	329	IYLKN K DL DAR
119	P06400	412	CTVNP K ESILK
120	P06400	462	MESML K SEEER
121	P0CX43	78	DVDRA K SCGVD
122	P0CX43	95	LKKLN K NKKLI
123	P0CX43	118	SEVL I KQV PRL
124	P0CX43	156	VRST I KFQLKK
125	P47914	52	KKHN K KGLKKM
126	P47914	56	KKGL K KMQANN
127	P47914	77	IKALV K PKEVK
128	P47914	134	AKDOT K AQAAA
129	Q969Q0	98	LGGDK K RRKGQV
130	P0CX27	24	KHTQH K VTOYK
131	P0CX27	46	RRYDR K QSGFG
132	P0CX27	60	KPVF H KAKTT
133	P02393	66	FDVIL K AAGAN
134	P02393	115	AEEAK K Q LVEA
135	P0A7K2	108	DAEAL K KALEE
136	P07472	67	TSAGE K KVNVI
137	P07472	73	KVNVI K VVREI

138	Q12962	108	ANGDV K PVVSS
139	Q12962	177	QHCKM K GTASG
140	Q04207	79	ISLV T KDPPHR
141	Q53688	66	ELGGE K EYRRL
142	Q53688	176	YDTLQ K QNYTL
143	Q53688	362	VKECD K TNEIE
144	Q53688	371	IEEAT K RNPEA
145	Q53688	445	STHDT K FSEDV
146	E5KVG9	129	AEKLG K VSELV
147	B6A879	43	VIFDG K VYTNT
148	B6A879	531	SLSMG K NGFVL
149	Q9L5X8	431	RKGEL K DMAPD
150	D0ZRV8	19	KVNRN K VPTPM
151	Q6NEZ3	123	VGKNV K LDGNQ
152	Q6NEZ3	352	SLLPN K APIRV
153	Q6NEZ3	455	QMGVY K YSNSL
154	D1BQI7	90	MASIV K ALSrk
155	Q9UKV3	95	AKSGQ K SALVK
156	Q9UKV3	268	ADRNL K TEEEE
157	Q9UKV3	302	EAPIL K EFKEE
158	Q9UKV3	325	MDERP K TRSQE
159	Q9SJU4	293	ASYTL K LLRNR
160	Q9ZU52	8	SASFV K PNTLS
161	Q9ZU52	370	LAQLG K YSAEG
162	Q9ZU52	382	NEDAK K GMFVK
163	P62157	76	TMMAR K MKDTD
164	P04573	64	EARGP K GDKKN
165	P04573	154	MDLIK K SKNAL
166	P00889	43	ADLIP K EQARI
167	P00889	103	IPECQ K MLPKA
168	P00889	352	HAVLR K TDPRY
169	P00889	450	PLERP K SMSTD
170	P00063	61	YSAAN K NMAVN
171	P00078	84	LENPK K FMPGT
172	P00078	90	FMPGT K MSFAG
173	P00043	18	GANLF K TRCLQ
174	P00047	19	GANLF K TRCAQ
175	P00047	81	LENPK K FIPGT
176	Q71V39	84	KFETT K YYITI
177	Q71V39	172	YDEIV K EVSAy
178	P02994	51	AAELG K GsfKY
179	P02994	159	KMDSV K WDESR
180	Q96KQ7	131	CPSRA K MSMTG
181	P53991	122	TKINS K GKEVP
182	P53991	291	MYIQD K VEEYA
183	P53991	315	YFCGL K GMPG
184	P07041	123	VTIQS K DIQLA
185	P10587	53	SIKEE K GDEVT
186	P10587	144	DMYKG K KRHEM
187	P10587	188	KTENT K KVIQY
188	P10587	204	SSHKG K KDTSI
189	P10587	205	SHKG K KDTSIT

190	P11416	262	QITLLKAACLD
191	P11416	365	KVYVRKRRPSR
192	P27064	32	PEYETKDTDIL
193	P27064	201	GLDFTKDDENV
194	P27064	252	CEEMMKRAIFA
195	P04717	32	PDYQTKDTDIL
196	P04717	450	IREACKWSPPEL
197	P04992	81	SLDRYKGHRYR
198	P0A7J7	10	VQAYVKLQVAA
199	P0A7J7	51	TDSIEKGLPIP
200	P0A7J7	100	KDKVGKISRAQ
201	P0CX53	40	LGLSPKKGED
202	P0CX53	86	VITALKEPPRD
203	P55072	217	QLAQIKEMVEL
204	P05547	35	EQSLKKQKKG
205	P05547	37	SLKKQKKGFM
206	P05547	38	LKKQKKGFMTP
207	P05547	47	TPERKKKLRL
208	P05547	48	PERKKKLRL
209	P05547	74	AGERRKIIDQR
210	P05547	112	QIESDKYDVEL
211	P05547	122	LEIIRKDYEIN
212	P05547	139	NDLRGKFIKPT
213	O95785	279	IYFKQKEHLL
214	O95785	332	QASREKIIIEE
215	O95785	357	RLQCPKCVFG
216	O95785	756	FQKKKKVANS

(6) Subset 6 (217 samples)

Sample #	Protein ID	Site	Sequences
1	P05141	23	AAAIKSTAVAP
2	Q9Y232	31	QKLFLKRNNS
3	Q9Y232	69	ERIVDKRKNK
4	Q9Y232	71	IVDKRKNKKG
5	Q9Y232	147	NSNFSKTSPK
6	Q9Y232	151	SKTSPKALVIG
7	Q9Y232	175	SQKFRKNTAPS
8	Q9Y232	191	NMDLAKSGIK
9	Q9Y232	200	KILVPKSPVKS
10	P26358	39	DSLTEKECVKE
11	O75530	76	GKWKSKCKYS
12	O75530	106	FNWHSKEGDPL
13	Q5VTE0	212	NMPWFKGWKVT
14	Q5VTE0	244	TRPTDKPLRLP
15	P0CE47	90	HADYVKNMITG
16	P0CE47	177	RGSALKALEGD
17	P0CE47	358	PGDNIKMVVTL
18	O75367	12	KKKSTKTSRSA
19	O75367	116	PELLAKKRGSK

20	O75367	134	TPPPAKKAKSP
21	O75367	148	KPVSKKAGGKK
22	P62807	25	TKAQKKGKKR
23	P62807	44	SVYVYKVLKQV
24	P62807	117	VSEGTKAVTKY
25	P62807	121	TKAVTKYTSSK
26	P62803	6	MSGRGKGGKGL
27	P62803	17	GKGGAKRHRKV
28	P04637	291	EENLRKKGEPH
29	P04637	292	ENLRKKGEPHH
30	Q86L05	98	GPKNKKAIKKL
31	Q86L05	130	GPGLNKVGKFP
32	Q7M1B9	15	SDFAEKVLKSK
33	P14909	33	NVEKTKKIKII
34	P14909	95	GTDVKKEEVIV
35	P14909	193	SPNDVKKIVDI
36	P14909	220	FVYEGKMRSTL
37	P14909	337	VSKILKTSGFD
38	P14909	344	SGFDVKSLAIK
39	P39462	24	GVPKPKGPQVL
40	P39462	86	VVGYSKGD LVA
41	P39462	252	LNNSEKTL SVY
42	P39462	266	LAKQGKYVMVG
43	P22498	304	TRGNEKIVRDD
44	P22498	374	YDVLTKYWNRY
45	Q9Y232	219	SESPEKLDPVE
46	Q9Y232	284	LAVNGKGTSPF
47	P0C835	34	APKGRKGVKIG
48	P80053	56	SDGKCLKTFMGW
49	P80053	118	RVDPKCLTREE
50	P80053	130	EQLSRKYIQAI
51	P80053	279	YPEGRKVTNEE
52	P80053	309	KFNAPKVKAKL
53	P80053	311	NAPKVKAKLIV
54	P80053	360	EWANNKMGEII
55	P80053	371	SDEEAKKLIVD
56	P13123	28	VWRVGMVSFT
57	P61991	19	QVDISKIKKVV
58	P61991	28	VWRVGMISFT
59	Q5VTE0	311	DNVGFVKNV S
60	P55907	21	EHSGHKVYGPV
61	O43524	606	VMGHEKFPSDL
62	Q08369	255	NRPLIKPQRRL
63	Q08369	322	RKRKPKNLNKS
64	Q08369	421	SQASSKQDSWN
65	P16403	52	AVAASKERSGV
66	P16403	81	NNSRIKLG LKS
67	P16403	90	KSLVSKGTLVQ
68	P16403	149	GATPKKSAKKT
69	P16403	153	KKSAKTPKKA
70	P10412	137	AGAAKPKKAT
71	P10412	140	AKKPKKATGAA

72	P10412	149	AATPKKSAKKT
73	P16401	140	AGATPKKAKKA
74	P16401	144	PKKAKKAAGAK
75	P16401	157	VKKTPKKAKKP
76	O75367	159	GARKSKKKQGE
77	P40283	7	APKAEKKPAEK
78	P40283	23	KPVEEKSKAEK
79	P40283	25	VEEKSKAEKAP
80	P40283	28	KSKAEKAPA EK
81	O14979	216	HKLDGK LIDPK
82	O14979	233	GKEPPK KVFVG
83	O14979	289	E E P V K L L E S R
84	Q14103	173	DPKRAKAMKTK
85	P16862	83	LNSITKTQDWR
86	P16862	415	IFIPEK PATSS
87	P16862	580	DHNEPKLPKDK
88	P16862	641	RSLNWKDMLGW
89	O00159	200	SSRFGKYMDVQ
90	O00159	269	YLYLVKGQCAK
91	O00159	357	ALTHRKIIAKG
92	O00159	396	TWLVGKINRSL
93	O00159	512	LTFLEKLEDTV
94	O00159	632	KPNDAKQPGRF
95	P13538	20	APYLRKSEKER
96	P13538	60	SKEGGKVTVKT
97	P13538	64	GKVTVKTEGGE
98	P13538	74	ETLTVKEDQVF
99	P13538	191	KTVNTRVIOY
100	P13538	237	AFGNAKTVRND
101	P13538	249	SSRFGKFIRIH
102	P13538	368	GNLKFKQKQRE
103	Q96RI1	286	QEITNKILKEE
104	Q96RI1	352	AEIFNKKLPSG
105	Q96RI1	431	LDVLQK LCKIH
106	P11940	78	NFDVIK GKPVR
107	P11940	95	DPSLRKSGVGN
108	P11940	138	DENGSKYGFV
109	P11940	268	VGRAQK KVERQ
110	P11940	269	GRAQK KVERQT
111	P06400	96	YIQKKKELWGI
112	P06400	136	FFNLLKEIDTS
113	P06400	202	TFLLAKEVLQ
114	P06400	289	NIDEVKNVYFK
115	P06400	294	KNVYFKNFIPF
116	P06400	359	QRTPRKSNLDE
117	Q86L05	161	KFQLK KVLCLA
118	Q86L05	202	GWQNIKTLYVK
119	P0CX43	14	VREHVKELLKY
120	P0CX43	207	GSLVVKSSMGP
121	P47914	38	KGVDPKFLRNM
122	P47914	82	KPKEVKPKIPK
123	P47914	103	YIAHPKLGKRA

124	P47914	122	RLCRPKAKAKA
125	Q969Q0	28	VTQYKKGKDSL
126	P0CX27	61	PVFHKKAKTTK
127	P0CX27	67	AKTTKKVVLRL
128	P0CX27	97	ELGGEKKQKGQ
129	P02393	58	PAEEEEKTEFDV
130	P0A7K2	52	EAAEEKTEFDV
131	P0A7K2	85	GLKEAKDLVES
132	P07472	112	DEAKTKLEEAG
133	Q04207	28	IIEQPKQGM
134	Q04207	303	IEEKRKRTYET
135	Q53688	147	LLKIVKDGDEY
136	Q53688	158	FLEYFKWKLPL
137	Q53688	160	EYFKWKLPLTE
138	Q53688	248	LRSIINKKIII
139	Q53688	358	ANQIVKECDKT
140	Q53688	390	PAVYAKAYEDT
141	E5KWG9	191	GVFEEKGVRAI
142	E5KWG9	211	SEELLKFCERS
143	E5KWG9	229	SWDSPKINRFT
144	B6A879	121	PDVNFKEDGSQ
145	B6A879	291	RELQKKAHAMG
146	B6A879	445	NLYSPKDPGEQ
147	Q9L5X8	28	TTLVDKLLQOS
148	Q9L5X8	480	HYGPHKGGNIG
149	Q9L5X8	499	ANAAGKALTNA
150	Q6NC90	132	RLKPLKPGKGF
151	Q6NC90	135	PLKPGKGFGEI
152	D0ZRV8	36	RLRVRKYRDPQ
153	D0ZRV8	53	LPESLKALLAY
154	D0ZRV8	190	HLPQWKAQOKM
155	Q6NEZ3	145	INDHLKASLEG
156	Q9UKV3	213	SISEEKGDSDD
157	Q9UKV3	305	ILKEFKEEGEE
158	Q9SJU4	61	LVKTAKTIASP
159	Q9SJU4	186	GARFAKWRTVV
160	Q9SJU4	203	SALAVKEAAWG
161	Q9ZU52	288	TLTMLKRRVPP
162	P62157	31	GTITTKELGTV
163	P04573	14	LGPEEKDECMK
164	P00889	49	EQARIKTFRQQ
165	P00889	109	MLPKAKGGEEP
166	P00889	139	VSWLSKEWAKR
167	P00889	382	VAQLYKIVPNV
168	P00078	98	FAGMKKPQERA
169	P00043	61	TDANKKKGVEW
170	P00043	62	DANKKKGVEWS
171	P00070	31	CHTVEKGAGHK
172	P00070	96	FPGLKQPQERA
173	P00047	35	QGGANKIGPNL
174	P00047	47	GLFGRKTGSVE
175	Q71V39	212	NMPWFKGWKVE

176	P02993	212	RLPWYKGNIE
177	P02994	210	NAPWYKWEKE
178	P02994	222	KAGVVKGKTL
179	Q96KQ7	341	SSGRKAKKKW
180	Q96KQ7	358	VKPSRKRRKRE
181	P53991	38	PVAATKASTAV
182	P53991	170	EDPAKKGLCSN
183	P07041	116	CAIHAKRVTIQ
184	P10587	11	LSDDEKFLFVD
185	P10587	72	KVTLKDDIQK
186	P68696	91	GVITVKTSKI
187	P11416	86	LPRIYKPCFVC
188	P11416	147	YCRLQKCFDVG
189	P11416	164	RNDRNKKKKEA
190	P27063	6	PQTETKASVGF
191	P27064	81	SLDRYKGRCYG
192	P27064	177	CTIKPKLGLSA
193	P27064	183	LGLSAKNYGRA
194	P27064	334	GTVVVKLEGER
195	P27064	356	DDFVEKDRSRG
196	P27064	450	IREASKWSPEL
197	P04717	8	PQTETKAKVGF
198	P04992	146	PPAYVKTFQGP
199	P0A7J7	72	FTFVTKTPPAA
200	Q5SLP6	90	EKGAHKPGREK
201	P0CX53	31	AALAPKIGPLG
202	P0CX53	93	PPRDRKKDKNV
203	P0CX53	94	PRDRKKDKNVK
204	P0CX53	119	RQMRDKSFGRT
205	P0CX53	146	CRVDFKNPHDI
206	P62986	6	MQIFVKTLTGK
207	P55072	112	DVKYGKRIHVL
208	P55072	148	YRPIRKGDIFL
209	P55072	295	FEEAEKNAPAI
210	P05547	12	AEEAKKQDDI
211	P05547	68	KEQERKAGERR
212	P05547	84	RCGQPKNLDGA
213	O95785	179	VEDTPKTLDMA
214	O95785	385	PGQTTKEPFGG
215	O95785	610	AALELKQAFRE
216	O95785	641	IVLVAKLGPQV
217	O95785	697	PAMALKHEERK

(7) Subset 7 (217 samples)

Sample #	Protein ID	Site	Sequences
1	P05141	63	VVRIPKEQGV
2	P05141	147	AADVGKAGAER
3	P05141	268	RDEGGKAFK
4	Q9Y232	76	KNKKGKTEYLV

5	Q9Y232	195	AKSGIKILVPK
6	P00042	85	YIPGTKMAFGG
7	P00042	92	AFGGLKKEKDR
8	P26358	43	EKECVKEKLN
9	P26358	134	PKPLSKPRTPR
10	O75530	22	AAKKQKLSSE
11	O75530	74	GKGKWKSKKCK
12	O75530	211	LLSVSKDHALR
13	O75530	332	AIVCWKPGKME
14	Q5VTE0	44	IEKFEKEAAEM
15	Q5VTE0	64	VLDKCLKAERER
16	P0CE47	25	HVDHGKTTLTA
17	P0CE47	295	GQVLAKPGTIK
18	O75367	9	RGGKKKSTKTS
19	O75367	34	LRYIKKGHPKY
20	O75367	72	AARDNKGRVT
21	O75367	135	PPPAKKAKSPS
22	P62807	6	MPEPAKSAPAP
23	P68432	116	CAIHAKRVTIM
24	P09988	116	CAIHGKRVTIQ
25	Q55BN9	65	ELLIKKLFPQR
26	P61830	57	IRRFQKSTELL
27	P61830	123	VTIQKKDIKLA
28	P62803	13	GKGLGKGGAKR
29	P04637	101	SVPSQKTYQGS
30	P04637	132	SPALNKMFCQL
31	P04637	321	PQPKKKPLDGE
32	Q86L05	18	VSQLFKEAQES
33	Q86L05	24	EAQESKRGFLE
34	Q86L05	102	KKAIKKLSKKY
35	P80322	44	LGTIMKRLGMS
36	P80322	153	DYEEWKEIIQE
37	P14909	166	DDLQSKISKRT
38	P14909	281	TSFVQKAAVKA
39	P14909	334	FPNVSKILKTS
40	P14909	349	KSLAIKLIIEK
41	P39462	31	PQVLIKVEAAG
42	P39462	61	EDLGVKLPVTL
43	P39462	172	SLDPTKTLV
44	P39462	192	AVQIAKAVSGA
45	P39462	320	KPMITKTMKLE
46	P39462	323	ITKTMKLEEAN
47	P22498	219	GYVGVKSGFPP
48	P22498	255	IKSVSKKPVGI
49	Q9Y232	388	AADDSKLVLLS
50	Q9Y232	417	LTDDRKRESTK
51	Q9Y232	422	KRESTKMAEAI
52	P0C835	37	GRKGVKIGLFK
53	P0C835	42	KIGLFDKDPETG
54	P80053	202	VATIAKEAANK
55	P80053	239	SEMGAKIVGVS
56	P80053	247	GVSDSKGGVIN

57	Q5VTE0	385	DRRSGKKLEDG
58	P00219	10	NYRTSKPVVGD
59	O43524	176	ESSPDKRLTLS
60	O43524	197	PYFKDKGDSNS
61	O43524	360	SPSVSKPCTVE
62	P16403	17	APPAEKAPVKK
63	P16403	22	KAPVKKKAAKK
64	P16403	23	APVKKKAAKKA
65	P16403	46	SELITKAVAAS
66	P16403	85	IKLGLKSLVSK
67	P16403	109	SFKLNKKAASG
68	P16403	121	AKPKVKKAGGT
69	P16403	137	VGAAKPKKAA
70	P16403	148	GGATPKKSACK
71	P16403	157	KKTPKAKKPA
72	P10412	117	ASGEAKPKAKK
73	P16401	55	AVAASKERNGL
74	O75367	189	TVLSTKSLFLG
75	P40283	12	KKPAEKKPASE
76	P40283	34	APAEEKPKAGK
77	P40283	51	GAGGDKKKMK
78	P40283	53	GGDKKKMKKK
79	P40283	56	KKKKMKKSVE
80	Q42681	64	ELVIRKLPFOR
81	Q41811	13	GKGLKGARKR
82	O14979	180	VDCTIKTDPVT
83	O14979	204	AASVDKVLELK
84	O14979	227	RAKALKGKEPP
85	O14979	229	KALKGKEPPKK
86	O14979	306	GKCEIKVAQPK
87	O14979	311	KVAQPKVYRQ
88	Q14103	114	TKKDLKDYFSK
89	Q14103	161	DQKEHKLNGKV
90	Q14103	165	HKLNGKVIDPK
91	Q14103	238	EEPVKKIMEKK
92	Q14103	260	KVAMSKQYQQ
93	P16862	25	SVQSYKAAIDF
94	P16862	66	LRPDEKINKNV
95	P16862	150	NAVSTKPTPPP
96	P16862	168	SGLSSKVHSYT
97	P16862	438	HRSRGKRTTIV
98	O00159	146	ESGAGKTEATK
99	O00159	188	AFGNAKTLRND
100	O00159	209	VQFDFKGPVVG
101	O00159	535	DQRTRKSLGRG
102	O00159	594	SELSDKKRPET
103	P13538	147	LAYRGKQRQEA
104	P13538	148	AYRGKQRQEAP
105	P13538	186	ESGAGKTVNTK
106	P13538	296	QIMSNKKPELI
107	P13538	349	FSADEKTAIYK
108	Q96RI1	33	EMMSMKPAKGV

109	Q96RI1	120	EMPVTKKPRMG
110	Q96RI1	198	RLRKCKEMGML
111	Q96RI1	218	TEIQCKSKRLR
112	Q96RI1	256	KSCREKTELTP
113	Q96RI1	413	DRQYIKDREAV
114	Q96RI1	420	REAVEKLQEPL
115	P11940	177	RFKSRKEREAE
116	P11940	188	LGARAKEFTNV
117	P11940	208	DDERLKDILFGK
118	P11940	279	TELKRKFEQMK
119	P11940	284	KFEQMKQDRIT
120	P06400	240	LKEPYKTAVIP
121	P06400	420	ILKRVKDIGYI
122	Q86L05	160	VKFQLKKVLCL
123	Q86L05	196	VSLLLKQWQNI
124	P0CX43	24	YSNETKKRNFL
125	P0CX43	25	SNETKKRNFL
126	P0CX43	101	NKKLIKLSKK
127	P0CX43	106	KKLSKKYNAFI
128	P0CX43	147	DDLYGKVTDVR
129	P0CX43	195	FVSLLLKNWQNI
130	P47914	33	RYESLKGVDPK
131	P47914	51	AKKHNNKGLKK
132	P47914	79	ALVKPKKVKPK
133	P47914	84	KEVKPKIPKGV
134	Q969Q0	17	CKKCGKHQPHK
135	Q969Q0	61	FRKKAKTTKKI
136	P0CX27	15	TYCKGKTCRKH
137	P0CX27	86	AQLTLKRCKHF
138	P0CX27	100	GEKKQKGOALQ
139	P02393	107	KEAVSKEEAEE
140	P0A7K2	71	KVAVIKAVRGA
141	P0A7K2	101	KEGVSKDDAEA
142	P07472	68	SAGEKKVNVIK
143	P07472	87	GLKEAKAAVDG
144	Q12962	161	SLAAQKFISDI
145	Q04207	195	NTAELKICRVN
146	Q04207	218	FLLCDKVQKED
147	Q04207	314	FKSIMKKSPFN
148	Q53688	256	IIIVEKILGFQ
149	B6A879	107	ILVEDKYTEET
150	B6A879	212	IGNKSKKVQEV
151	B6A879	213	GNKSKKVQEVW
152	B6A879	567	RTVKQKGEYAA
153	B6A879	573	GEYAAKNKLG
154	Q9L5X8	8	TPQIDKLRNIA
155	Q9L5X8	543	TVNALKGKQLT
156	Q6NEZ3	41	DTQAPKPVNTV
157	Q6NEZ3	62	AWGLYKFNQKS
158	Q6NEZ3	226	ESASIKQDVQI
159	Q6NEZ3	332	ATSLLKPAKVM
160	Q9UKV3	44	WRKLPKMPEAV

161	Q9UKV3	238	QARAAKLSEGS
162	Q9UKV3	294	DDEGQKSREAP
163	P18281	47	LEEVAKPHELK
164	Q9SJU4	82	NATCGKRLASI
165	Q9SJU4	147	IVPGIKVDKGL
166	Q9SJU4	268	EGILLKPSMVT
167	Q9SJU4	353	KTWGGKEENVK
168	Q9ZU52	50	SDELVKTAKSI
169	P11118	87	TEEEIKEAFRV
170	P00889	76	GMRGMKGLVYE
171	P00889	193	GIHRTKYWELI
172	P00889	215	PCVAAKIYRNL
173	P00063	19	GEKIFKTKCAQ
174	P00063	30	CHTVDKGAGHK
175	P00063	87	YIPGTMVFPG
176	P00077	16	PGDAAKGEKIF
177	P00043	12	KGSEKKGANLF
178	P00043	94	FGGLKKAKDRN
179	P00076	22	CHSAQKGVNST
180	P00076	52	YSNANKNAAIV
181	P00076	87	AGIKAKKDRQD
182	P00070	62	YSAGNKNKAVI
183	P00047	95	FGGLKKNKDRN
184	Q71V39	219	WKVERKEGNAS
185	P02993	44	IEKFEKEAQEM
186	P02993	51	AQEMKGSFKY
187	P02994	214	YKGWEKETKAG
188	P02994	224	GVVKGKTLLEA
189	Q96KQ7	343	GRRKAKKKWRK
190	Q96KQ7	345	RKAKKKWRKDS
191	Q96KQ7	368	EPPRAKEPRGV
192	P53991	6	MASLRKPSNHA
193	P53991	234	NVPSYKFTGLF
194	P53991	285	NRKGGKMYIQD
195	P10587	17	FLFVDKNFVNN
196	P10587	145	MYKGGKRHEMP
197	P10587	202	VASSHKGK KDT
198	P11416	118	RRSIQKNMVYT
199	P11416	165	NDRNKKKKEAP
200	P11416	167	RNKKKKKEAPKP
201	P27064	161	QVERDKLNKYG
202	P27064	305	VIDRQKNHGMH
203	P04717	227	AEAIYKSOAET
204	P04717	355	RDDYIKKDRSR
205	P04992	236	ETGEIKGHYLK
206	P04992	241	KGHYLKATAGT
207	P0A7J7	87	KAAGIKSGSGK
208	P0A7J7	95	SGKPNKDKVGK
209	P0CX53	130	LASVTK EILGT
210	P62986	88	RQLAQKYNCDK
211	P55072	109	PCPDVKYGKRI
212	P05547	13	EEAKKKQDDID

213	P05547	56	LLLRK K AAEEL
214	P05547	62	AAEEL K KEQER
215	O95785	550	QLGRN K STVHP
216	O95785	753	RFSF Q KKKKV
217	O95785	755	SF Q KK K KKVAN