

Supporting Information S2. The benchmark dataset S_K used to train and test the model for predicting the possibility of hydroxylation at Lys site. It contains 142 positive samples and 980 negative samples, of which were extracted from the 164 hydroxylation proteins. See the main text for further explanation.

(1) List of the 142 peptide samples in the positive subset

Sample #	Protein ID	Site	Sequences
1	O18495	40	KSVGKFYYKH K YYIKAAWQIG
2	O18495	44	KFYYKHKYYI K AAWQIGKHAL
3	O18495	27	IQQSEAGWLR K AAKSVGKFYY
4	O18495	30	SEAGWLRKAA K SVGKFYYKHK
5	O18495	34	WLRKAAKSVG K FYYKHKYYIK
6	O18495	38	AAKSVGKFYY K HKYYIKAAWQ
7	P01170	120	RKAGCKNFYW K GFTSCCSTFG
8	P02452	265	LPGTAGLPGM K GHRGFSGLDG
9	P02452	1108	ETGEQQDRGI K GHRGFSGLQG
10	P02453	276	HRGFSGLDGA K GDAGPAGPKG
11	P02453	285	AKGDAGPAGP K GEPGSPGENG
12	P02453	708	ANGAPGNDGA K GDAGAPGAPG
13	P02453	780	PPGPAGAPGD K GEAGPSGPAG
14	P02453	861	PIGNVGAPGP K GARGSAGPPG
15	P02453	933	PPGPPGPAGE K GAPGADGPAG
16	P02453	1095	PAGPQQPRGD K GETGEQQDRG
17	P02454	575	QAGVMGFPGP K GTAGEPGKAG
18	P02454	698	NNGAPGNDGA K GDTGAPGAPG
19	P02457	419	PQGPPGAPGP K GNSGEPGAPG
20	P02457	494	FPGADGIAGP K GPPGERGSPG
21	P02457	740	AKGDRGDPGP K GADGAPGKDG
22	P02459	190	AAQMAGGFDE K AGGAQMGMQ
23	P02459	287	FPGTPGLPGV K GHRGYPGLDG
24	P02459	299	HRGYPGLDGA K GEAGAPGVKG
25	P02459	308	AKGEAGAPGV K GESGSPGENG
26	P02459	374	GPGFPGAPGA K GEAGPTGARG
27	P02459	419	NPGTDGIPGA K GSAGAPGIAG
28	P02459	452	PQGATGPLGP K GQTGEPGIAG
29	P02459	464	QTGEPGIAGF K GEQGPKEPG
30	P02459	470	IAGFKGEQGP K GEPPGAPQG
31	P02459	527	FPGQDGLAGP K GAPGERGPSG
32	P02459	542	ERGPSGLAGP K GANGDPGRPG
33	P02459	608	QPGVMGFPGP K GANGEPGKAG
34	P02459	620	ANGEPGKAGE K GLPGAPGLRG

35	P02459	1130	ETGEAGERGL K GHRGFTGLQG
36	P02461	263	PAGIPGFPGM K GHRGFDGRNG
37	P02461	284	EKGETGAPGL K GENGLPGENG
38	P02461	860	PAGPPGPQGV K GERGSPGGPG
39	P02461	977	PRGSPGPQGV K GESGKPGANG
40	P02461	1094	APGPQGPRGD K GETGERGAAG
41	P02461	1106	ETGERGAAGI K GHRGFPGNPG
42	P02745	33	EDLCRAPDGK K GEAGRPGRRG
43	P02745	48	RPGRRGRPGL K GEQGEFGAPG
44	P02745	67	PGIRTGIQGL K GDQGEPPGSG
45	P02745	100	PLGARGIPGI K GTKGSPGNIK
46	P02747	75	IPGIRGPKGQ K GEPGLPGHPG
47	P04142	47	MGRNIRDGIV K AGPAIEVLGS
48	P04258	95	PRGFPGPPGM K GPAGMPGFPG
49	P04258	107	PAGMPGFPGM K GHRGFDGRNG
50	P04258	119	HRGFDGRNGE K GEFGAPGLKG
51	P04258	938	PPGPQGPRGD K GETGERGAMG
52	P04258	950	ETGERGAMGI K GHRGFPGNPG
53	P12108	181	LCPTNCPGPG K GPQGLQGLKG
54	P12111	2103	VKGSRGFPGE K GEVGEIGLDG
55	P12111	2209	GFGRRGPPGA K GNKGPGQPFG
56	P12111	2212	RRGPPGAKGN K GGPGQPGFEG
57	P12111	2322	ERGFPGYPGP K GNPGEPLNG
58	P12111	2337	EPGLNGTTGP K GIRRRGNSG
59	P19999	44	ACGRDGRDGP K GEKGEPGQGL
60	P19999	47	RDGRDGPKE K GEFGQGLRGL
61	P19999	79	SVGAPGSQGP K GQKDRGDSR
62	P19999	82	APGSQGPKGQ K GDRGDSRAIE
63	P20908	627	ARGMPGQTGP K GDRGFDGLAG
64	P20908	642	FDGLAGLPGE K GHRGDPGPGSG
65	P20908	708	VTGMDGQPGP K GNVGPQGEFG
66	P20908	744	PQGAIGPPGE K GPLGKPGPLPG
67	P20908	774	HPGKEGPPGE K GGQGPPGPQG
68	P20908	795	PIGYPGPRGV K GADGIRGLKG
69	P20908	804	VKGADGIRGL K GTKGEKGEDG
70	P20908	807	ADGIRGLKGT K GEKGEDGFPG
71	P20908	810	IRGLKGTKGE K GEDGFPGFKG
72	P20908	819	EKGEDGFPGF K GDMGIKDRG
73	P20908	846	PRGEDGPEGP K GRGGPNGDPG
74	P20908	864	DPGPLGPPGE K GKLGVPGPLPG
75	P20908	882	LPGYPGRQGP K GSIGFPGFPG
76	P20908	897	FPGFPGANGE K GGRGTPGKPG
77	P23805	63	QDGRECPHGE K GDPGSPGPAG
78	P23805	87	RPGWVGPIGP K GDNGFVGEFG

79	P23805	99	DNGFVGEPPG K GDTGPRGPPG
80	P23805	135	SMGPPGTTPGP K GETGPKGGVG
81	P23805	141	TPGPKGETGP K GGVGAPGIQG
82	P23805	159	IQGFPGPSGL K GEKGAPGETG
83	P23805	162	FPGPSGLKGE K GAPGETGAPG
84	P23805	198	PSGARGPPGL K GDRGDPGETG
85	P23805	210	DRGDPGETGA K GESGLAEVNA
86	P26368	15	DEFERQLNEN K QERDKENRHR
87	P26368	276	LPNYLNDDQV K ELLTSFGPLK
88	P30754	96	AGGMPGMPGP K GHRGFRGLSG
89	P30754	108	HRGFRGLSGS K GEQKSGNQG
90	P30754	183	SPGFPGTPGS K GDRGQSGIKG
91	P30754	192	SKGDRGQSGI K GAQGLQGPVG
92	P30754	261	TAGSPGQAGA K GDGGPTGEQG
93	P30754	279	EQGRPGAPGV K GSSGPPGDVG
94	P30754	342	ASGAMGIPGE K GPSGEPGAKG
95	P30754	351	EKGPSGEPGA K GPTGDTGRQG
96	P30754	546	PTGSPGVAGA K QGGPPGPAG
97	P30754	567	LVGLPGERGP K GVGGSKGSRG
98	P30754	573	ERGPKGVGGS K SRGDIGPRG
99	P30754	612	LPGPSGLAAS K GERGDMGSPG
100	P30754	657	PPGDAGPAGT K GDIGFPGERG
101	P30754	738	PAGTAGKQGV K GARGSPGLVG
102	P30754	765	SDGEPGRDGT K GERGEDGPPG
103	P30754	810	ETGPMGGQGM K GDGGPPGPSG
104	P30754	927	SQGPAGIMGM K GEAGGKGAKG
105	P30754	933	IMGMKGEAGG K AKGDKGWTG
106	P30754	936	MKGEAGGKG K AGDKGWTGLPG
107	P30754	939	EAGGKGAKGD K GTGLPGLQG
108	P30943	41	GATSVNMAMD K SAKAPVITIF
109	P35248	86	LPGPRGPVGP K GENGSAGEPG
110	P35248	98	ENGSAAGEPGP K GERGLVGPPG
111	P83414	21	VGRNIRDGIV K AGPAIAVLGQ
112	P83415	21	MGRNIRDGVI K AAPAIEVLGQ
113	P84698	18	TTCANGWECC K GYPCVNKACS
114	P84698	25	ECCKGYPCVN K ACSGCTHGKK
115	Q00433	56	SGAPAGAAMD K SAKAPQITIF
116	Q02388	2625	DVGFMGPRGL K GERGVKGACG
117	Q02388	2631	PRGLKGERGV K GACGLDGEKG
118	Q05707	1476	PPGGPGLRGP K GQQGEPGPKG
119	Q05707	1485	PKGQQGEPGP K GPDGPRGEIG
120	Q05707	1523	IQGMPPGMPGE K GEKGDGTGLPG
121	Q05707	1526	MPGMPGEKGE K GDTGLPGPQG
122	Q05707	1601	ALGPPGVPGA K GERGERGDLQ

123	Q05707	1698	TPGERGLTGI K GEKGNPGVGT
124	Q05707	1701	ERGLTG I KGE K GNPGVGTQGP
125	Q15848	33	QGPVLLPLP K GACTGWMAGI
126	Q15848	65	RDGRDGT P GE K GEKGD P GLIG
127	Q15848	68	RDGTPGE K GE K GD P GLIG P KG
128	Q15848	77	EKGDPGLIG P K GDIGETGV P G
129	Q15848	101	PRGF P GIQGR K GE P GE G AYVY
130	Q3Y5Z3	28	EDNMEDPPLP K GACAGWMAGI
131	Q3Y5Z3	60	RDGRDGT P GE K GEKGD P GLVG
132	Q3Y5Z3	63	RDGTPGE K GE K GD P GLVG P KG
133	Q3Y5Z3	72	EKGDPGLVG P K GD T GETG I TG
134	Q3Y5Z3	96	PRGF P GTPGR K GE P GESAYVY
135	Q4ZJN1	73	GPGKDGIRGE K GE P GADGRVE
136	Q4ZJN1	127	ETGPQ G QKGD K GEV G PTG P EG
137	Q60994	68	RDGRDGT P GE K GEKGDAGLLG
138	Q60994	71	RDGTPGE K GE K GDAGLLG P KG
139	Q60994	80	EKG D AGLLG P K GETGDVGM T G
140	Q60994	104	PRGF P GTPGR K GE P GEAA V Y
141	Q9Y383	266	EKLRRSRSHS K NP K RSRSREH
142	Q9Y383	269	RRSRSHSKNP K RSRSREHRRH

(2) List of the 980 peptide samples in the negative subset

Sample #	Protein ID	Site	Sequences
1	O18495	4	ITAKMQMMQM K ATILIVLVAL
2	O18495	51	YYIKAAWQIG K HALGDMTDEE
3	O18495	67	MTDEEFQDFM K EVEQAREEEL
4	P01170	61	SAGLLSQEWS K RAVEELLAQM
5	P01170	111	STNNLPPRE R KAGCKNFYWK G
6	P01170	115	LPPRE R KAGC K NFYWK G FTSC
7	P02452	54	GLRYHDRV V K PEPCRICVCD
8	P02452	67	PCRICVCD N G VLCDDVICDE
9	P02452	79	LCDDVICDE T K NC P GAEV P EG
10	P02452	114	DQETT G VE G P K GD T GPRGPRG
11	P02452	170	APQLSYGYDE K STGGISVP G P
12	P02452	228	GPRGPPGPP G K NGDDGEAG K P
13	P02452	237	GKNGDDGEAG K PGR P GERG P P
14	P02452	277	HRGFSGLDGA K GDAGPAG P KG
15	P02452	286	AKGDAGPAG P K GE P GS P GENG
16	P02452	352	PPGF P GAVGA K GEAG P QGP R G
17	P02452	397	NPGADGQ P GA K GANGAPGIAG
18	P02452	430	PQGP G PPGP K NSGEPGAP G
19	P02452	442	NSGEPGAPGS K GD T GAKGE P G
20	P02452	448	APGSKGD T GA K GE P GPVGV Q G

21	P02452	468	GPPGPAGEEG K RGARGEPPGPT
22	P02452	505	FPGADGVAGP K GPAGERGSPG
23	P02452	520	ERGSPGPAGP K GSPGEAGRPG
24	P02452	538	RPGEAGLPGA K GLTGSPGSPG
25	P02452	552	GSPGSPGPDG K TGPPGPAGQD
26	P02452	586	QAGVMGFPGP K GAAGEPGKAG
27	P02452	594	GPKGAAGEPG K AGERGVPPGP
28	P02452	612	GPPGAVGPAG K DGEAGAQGPP
29	P02452	657	GPAGPPGEAG K PGEQGVPGDL
30	P02452	709	ANGAPGNDGA K GDAGAPGAPG
31	P02452	742	ERGAAGLPGP K GDRGDAGPKG
32	P02452	751	PKGDRGDAGP K GADGSPGKDG
33	P02452	759	GPKGADGSPG K DGVRGLTGPI
34	P02452	781	PPGPAGAPGD K GESGPSGPAG
35	P02452	826	PPGADGQPGA K GEPGDAGAKG
36	P02452	835	AKGEPGDAGA K GDAGPPGPAG
37	P02452	862	PIGNVGAPGA K GARGSAGPPG
38	P02452	903	GPPGPPGPAG K EGGKGPRGET
39	P02452	907	PPGPAGKEGG K GPRGETGPAG
40	P02452	934	PPGPPGPAGE K GSPGADGPAG
41	P02452	984	GLPGPSGEPG K QGSPGASGER
42	P02452	1033	SPGRDGSPGA K GDRGETGPAG
43	P02452	1062	GAPGPVGPAG K SGDRGETGPA
44	P02452	1096	PAGPQGPRGD K GETGEQGDRG
45	P02452	1152	GPPGSAGAPG K DGLNGLPGPI
46	P02452	1208	FSFLPQPPQE K AHDGGYYRA
47	P02452	1236	DRDLEVDTTL K SLSQQIENIR
48	P02452	1253	ENIRSPEGS R KNPARTCRDLK
49	P02452	1263	KNPARTCRDL K MCHSDWKSGE
50	P02452	1270	RDLKMCHSDW K SGEYWIDPNQ
51	P02452	1288	PNQGCNLDAI K VFCNMETGET
52	P02452	1310	VYPTQPSVAQ K NWYISKNP KD
53	P02452	1316	SVAQKNWYIS K NP KD KRHVWF
54	P02452	1319	QKNWYISKNP K DKRHWVFGES
55	P02452	1321	NWYISKNP KD K RHWVFGESMT
56	P02452	1371	EASQNITYHC K NSVAYMDQQT
57	P02452	1385	AYMDQQTGNL K KALLLQGSNE
58	P02452	1386	YMDQQTGNL K KALLLQGSNEI
59	P02452	1424	GCTSHTGAWG K TVIEYKTTKT
60	P02452	1430	GAWGKTVIEY K TTKTSRLPII
61	P02452	1433	GKTVIEYKTT K TSRLPIIDVA
62	P02453	54	GLRYHDRDVW K PVPCQICVCD
63	P02453	79	LCDDVICDEL K DCPNAKVPTD
64	P02453	85	CDELKDCPN K AVPTDECCPVC

65	P02453	170	APQLSYGYDE K STGISVPGPM
66	P02453	351	PPGFPGAVGA K EGGPGQGPRG
67	P02453	429	PQGPPSGPPGP K NSGEPGAPG
68	P02453	447	APGSKGDTGA K GEPGPTGIQG
69	P02453	467	GPPGPAGEEG K RGARGEPPGA
70	P02453	504	FPGADGVAGP K GPAGERGAPG
71	P02453	519	ERGAPGPAGP K GSPGEAGRPG
72	P02453	750	PKGDRGDAGP K GADGAPGKDG
73	P02453	758	GPKGADGAPG K DGVRGLTGPI
74	P02453	902	GPPGPPGPAG K EGSKGPRGET
75	P02453	906	PPGPAGKEGS K GPRGETGPAG
76	P02453	1151	GPPGSAGSPG K DGLNGLPGPI
77	P02453	1207	LSFLPQPPQE K AHDGGYYRA
78	P02453	1309	VYPTQPSVAQ K NWYISKNPKE
79	P02453	1315	SVAQKNWYIS K NPKEKRHVWY
80	P02453	1318	QKNWYISKNP K EKRRHVWYGES
81	P02453	1320	NWYISKNPKE K RHVWYGESMT
82	P02454	45	GLRVPNGETW K PDVCLICICH
83	P02454	67	GTAVCDGVL C KEDLDCPNPQK
84	P02454	77	KEDLDCPNPQ K REGCECCPFCP
85	P02454	104	DAEVIGVEGP K GDPGPQGPRG
86	P02454	160	ASQMSYGYDE K SAGVSVPGPM
87	P02454	266	HRGFSGLDGA K GDTGPAGPKG
88	P02454	275	AKGDTGPAGP K GEPGSPGENG
89	P02454	341	PPGFPGAAGA K GEAGPQGARG
90	P02454	419	PQGPPSGAPGP K NSGEPGAPG
91	P02454	431	NSGEPGAPGN K GDTGAKGEPG
92	P02454	437	APGNKGDTGA K GEPGPAGVQG
93	P02454	457	GPPGPAGEEG K RGARGEPPGPS
94	P02454	583	GPKGTAGEPG K AGERGVPGPP
95	P02454	601	GPPGAVGPAG K DGEAGAQQAP
96	P02454	770	PPGPAGAPGD K GETGPSGPAG
97	P02454	815	PPGADGQPGA K GEPGDTGVKG
98	P02454	824	AKGEPGDTGV K GDAGPPGPAG
99	P02454	851	PIGNVGAPGP K SRGAAGPPG
100	P02454	892	GPPGPPGPVG K EGGKGPRGET
101	P02454	896	PPGPVGKEGG K GPRGETGPAG
102	P02454	1022	SPGRDGAPGA K GDRGETGPAG
103	P02454	1051	GAPGPVGPAG K NGDRGETGPA
104	P02454	1097	ETGEQGDRGI K GHRGFSGLQG
105	P02454	1197	FSFLPQPPQE K SQDGGYYRA
106	P02454	1277	PNQGCNLDAI K VYCNMETGQT
107	P02454	1299	VFPTQPSVPQ K NWYISPNPKE
108	P02454	1308	QKNWYISPNP K EKKHVWFGES

109	P02454	1310	NWYISPNPKE K KHVWFGESMT
110	P02454	1311	WYISPNPKE K KHVWFGESMTD
111	P02454	1374	AYMDQQTGNL K KSLLLQGSNE
112	P02454	1375	YMDQQTGNL K KSLLLQGSNEI
113	P02454	1413	GCTSHTGTWG K TVIEYKTTKT
114	P02454	1419	GTWGKTVIEY K TTKTSRLPII
115	P02457	43	CVQDGLTYND K DVWKPEPCQI
116	P02457	47	GLTYNDKD V WKPEPCQICVCD
117	P02457	107	YPESAGVEGP K GDTGPRGDRG
118	P02457	160	APQMSYGYDE K SAGVAVPGPM
119	P02457	217	GPRGPAGPP K NGDDGEAGKP
120	P02457	226	GKNGDDGEAG K PGRPGQRGPP
121	P02457	266	HRGFSGLDGA K GQPGPAGPKG
122	P02457	275	AKGQPGPAG P KGEPGSPGENG
123	P02457	341	PPGFPGAAGA K GETGPQGARG
124	P02457	386	NPGADGQPGA K GATGAPGIAG
125	P02457	509	ERGSPGAVGP K GSPGEAGRPG
126	P02457	575	QAGVMGFPG P KGAAGEPGKPG
127	P02457	583	GPKGAAGEPG K PGERGAPGPP
128	P02457	601	GPPGAVGAAG K DGEAGAQQPP
129	P02457	646	GPAGPPGEAG K PGEQGVPGNA
130	P02457	731	ERGAAGLPGA K GDRGDPGPKG
131	P02457	748	GPKGADGAPG K DGLRGLTGPI
132	P02457	770	PPGPAGAPGD K GEAGPPGPAG
133	P02457	815	PPGADGQPGA K GETGDAGAKG
134	P02457	824	AKGETGDAGA K GDAGPPGPAG
135	P02457	850	GPAGVGAPGP K GARGSAGPPG
136	P02457	891	GLPGPPGPAG K GSKGPRGETG
137	P02457	894	GPPGPAGKGS K GPRGETGPAG
138	P02457	921	PAGPPGPPGE K GSPGADGPIG
139	P02457	1020	APGRDGAAGP K GDRGETGPAG
140	P02457	1095	ETGEQQDRGM K GHRGFSGLQG
141	P02457	1139	GPPGSAGAAG K DGLNGLPGPI
142	P02457	1240	ENIRSPEGTR K NPARTCRDLK
143	P02457	1250	KNPARTCRDL K MCHGDWKSSE
144	P02457	1257	RDLKMCHGDW K SGEYWIDPNQ
145	P02457	1275	PNQGCNLDAI K VYCNMETGET
146	P02457	1297	VYPTQATIAQ K NWYLSKNPKE
147	P02457	1303	TIAQKNWYLS K NPKEKKHVWF
148	P02457	1306	QKNWYLSKN P KEKKHVWFGET
149	P02457	1308	NWYLSKNPKE K KHVWFGETMS
150	P02457	1309	WYLSKNPKE K KHVWFGETMSD
151	P02457	1358	EATQNVTYHC K NSVAYMDHDT
152	P02457	1372	AYMDHDTGNL K KALLLQGAN

153	P02457	1373	YMDHDTGNLKKALLLQGANEI
154	P02457	1420	GKTVIEYKTTKTSRLPIIDLA
155	P02459	30	VLRCHGQDVQKAGSCVQDGQR
156	P02459	44	CVQDGQRYNDKDVWKPEPCRI
157	P02459	48	GQRYNDKDVWKPEPCRICVCD
158	P02459	73	LCDDIICEDMKDCLSPETPFG
159	P02459	103	LPTASGQPGPKGQKGEPGDIK
160	P02459	106	ASGQPGPKGQKGEPGDIKDIV
161	P02459	113	KGQKGEPGDIKDIVGPKGPPG
162	P02459	119	PGDIKDIVGPKGPPGPQGAPG
163	P02459	140	EQGPRGDRGDKGEKGAPGPRG
164	P02459	143	PRGDRGDKGEKGAPGPRGRDG
165	P02459	250	GPRGPPGPPGKPGDDGEAGKP
166	P02459	259	GKPGDDGEAGKPGKSGERGPP
167	P02459	262	GDDGEAGKPGKSGERGPPGPQ
168	P02459	490	GAPGPAGEEGKRGARGEPPGA
169	P02459	574	GRPGDAGPQGVGPSGAPGED
170	P02459	616	GPKGANGEPKGAGEKGLPGAP
171	P02459	634	GAPGLRGLPGKDGETGAAGPP
172	P02459	679	GPPGPPGEGGKPGDQGVPGEA
173	P02459	731	LPGTPTDGPKGAAAGPAGPPG
174	P02459	764	ERGAAGIAGPKGDRGDVGEKG
175	P02459	773	PKGDRGDVGEKGPEGAPGKDG
176	P02459	781	GEKGPEGAPKDGGRGLTGPI
177	P02459	803	PPGPAGANGEKGEVGPPGPAG
178	P02459	848	PPGADGQPGAKEQGEAGQKG
179	P02459	857	AKGEQGEAGQKGDAGAPGPQG
180	P02459	884	PQGPTGVTGPKGARGAQGPPG
181	P02459	925	GPPGPPGPSGKDGPKGARGDS
182	P02459	929	PPGPSGKDGPKGARGDSGPPG
183	P02459	956	LQGPAGPPGEKGEPGDDGPSG
184	P02459	1006	GLPGPSGEPGKQGAPGASGDR
185	P02459	1055	PPGRDGAAGVKGDRGETGAVG
186	P02459	1084	GSPGPAGPIGKQGDRGEAGAQ
187	P02459	1118	MPGPQGPRGDKGETGEAGERG
188	P02459	1174	GPPGPVGPSGKDGANIPGPI
189	P02459	1231	SAFAGLGQREKGPDPLOYMRA
190	P02459	1260	QHDAEVDATLKSLNNQIESLR
191	P02459	1277	ESLRSPEGSRKNPARTCRDLK
192	P02459	1287	KNPARTCRDLKLCHPEWKSGD
193	P02459	1294	RDLKLCHPEWKSGDYWIDPNQ
194	P02459	1312	PNQGCTLDAMKVFCNMETGET
195	P02459	1333	CVYPNPASVPKKNWWSSKSKD
196	P02459	1334	VYPNPASVPKKNWWSSKSKDK

197	P02459	1340	SVPKKNWWSS SK SKDKKHIWFG
198	P02459	1342	PKKNWWSSKS K DKKHIWFGET
199	P02459	1344	KNWWSSSKSD K KHIWFGETIN
200	P02459	1345	NWWSSSKSKD K KHIWFGETING
201	P02459	1394	EGSQNITYHC K NSIAYLDEAA
202	P02459	1408	AYLDEAAAGNL K KALLIQGSND
203	P02459	1409	YLDEAAAGNL K KALLIQGSNDV
204	P02459	1435	GNSRFTYTVL K DGCTKHTGKW
205	P02459	1440	TYTVLKDGC T KHTGKWGKTMI
206	P02459	1444	LKDGC T KHTGKWGKTMIEYRS
207	P02459	1447	GCTKHTGKW G KTMIEYRSQKT
208	P02459	1456	GKTMIYRSQ K TSRLPIIDIA
209	P02461	7	FSMMMMSFVQ K GSWLLLALLH
210	P02461	46	GQSYADRDVW K PEPCQICVCD
211	P02461	108	PPNGQGPQGP K GDPGPPGIPG
212	P02461	161	YSPQYDSYDV K SGVAVGGLAG
213	P02461	229	GAIGPSGPAG K DGESGRPGRP
214	P02461	251	ERGLPGPPGI K GPAGIPGFPG
215	P02461	275	HRGFDGRNGE K GETGAPGLKG
216	P02461	350	TAGFPGPS P AKGEVGPAGSPG
217	P02461	395	PPGINGSPPG K GEMGPAGIPG
218	P02461	436	GLRGGAGEPG K NGAKGEPGPR
219	P02461	440	GAGEPGKNGA K GEPGPRGERG
220	P02461	461	EAGIPGVPGA K GEDGKDGSPPG
221	P02461	466	GVPGA K GEDGKDGSPPGEPGAN
222	P02461	503	PAGPNGIPGE K GPAGERGAPG
223	P02461	550	GSPGGPGSDG K PGPPGSQGES
224	P02461	584	QPGVMGFPG P KGNDGAPGKNG
225	P02461	592	GPKGNDGAPG K NGERGGPGGP
226	P02461	610	GGPGPQGP P KNGETGPQGPP
227	P02461	629	PPGPTGPGGD K GDTGPPGPQG
228	P02461	655	GTGGPPGENG K PGEPPGPKGDA
229	P02461	662	ENGKPGEPPG K GDAGAPGAPG
230	P02461	674	DAGAPGAPGG K GDAGAPGERG
231	P02461	707	GAGPPGPEGG K GAAAGPPGPPG
232	P02461	740	ERGGLGSPGP K GDKGEPGGPG
233	P02461	743	GLGSPGPKGD K GEPGGPGADG
234	P02461	757	GGPGADGVPG K DGPRGPTGPI
235	P02461	779	PPGPAGQPGD K GEGGAPGLPG
236	P02461	824	APGQNGEPGG K GERGAPGEKG
237	P02461	833	GKGERGAPGE K GEGGPPGVAG
238	P02461	901	GPPGPSGSPG K DGPPGPAGNT
239	P02461	923	APGSPGVSGP K GDAGQPGEKG
240	P02461	932	PKGDAGQPGE K GSPGAQGGPPG

241	P02461	982	GPQGVKGESG K PGANGLSGER
242	P02461	1031	LPGRDGSPGG K GDRGENGSPG
243	P02461	1060	GPPGPVGPAG K SGDRGESGPA
244	P02461	1150	GPVGPSGPPG K DGTSGHPGPI
245	P02461	1212	AAAIAGIGGE K AGGFAPYYGD
246	P02461	1228	PYYGDEPMD F KINTDEIMTSL
247	P02461	1239	INTDEIMTSL K SVNGQIESLI
248	P02461	1256	ESLISPDGSR K NPARNCRDLK
249	P02461	1266	KNPARNCRDL K FCHPELKSGE
250	P02461	1273	RDLKFCHPEL K SGEYWVDPNQ
251	P02461	1286	EYWVDPNQGC K LDAIKVFCNM
252	P02461	1291	PNQGCKLDAI K VFCNMETGET
253	P02461	1313	ISANPLNVPR K HWWTDSSAEK
254	P02461	1323	KHWWTDSSAE K KHVWFGESMD
255	P02461	1324	HWWTDSSAE K KHVWFGESMDG
256	P02461	1373	RASQNITYHC K NSIAYMDQAS
257	P02461	1387	AYMDQASGN V KALKLMGSNE
258	P02461	1388	YMDQASGN V KALKLMGSNEG
259	P02461	1391	QASGN V KAL K LMGSNEGEFK
260	P02461	1401	KLMGSNEGE F KAEGNSKFTYT
261	P02461	1407	EGEFKAEGNS K FTYTVLEDGC
262	P02461	1419	TYTVLEDGCT K HTEGWSKTVF
263	P02461	1426	GCTKHTGEWS K TVFEYRTRKA
264	P02461	1435	SKTVFEYRTR K AVRLPIVDIA
265	P02745	32	TEDLCRAPDG K KGEAGRPGRR
266	P02745	81	GEPGPSGNPG K VGYPGPSGPL
267	P02745	103	ARGIPGIKGT K GSPGNIKDQP
268	P02745	110	KGTKGSPGNI K DQPRPAFSAI
269	P02745	195	RSLGFCDTTN K GLFQVVSGGM
270	P02745	219	LQQGDQVWVE K DPKKGHIYQG
271	P02745	222	GDQVWVEKDP K KGHIYQGSEA
272	P02745	223	DQVWVEKDP K KGHIYQGSEAD
273	P02747	14	GPSSLPHLGL K LLLLLLLLLPL
274	P02747	47	GMPGLPGAPG K DGYDGLPGPK
275	P02747	57	KDGYDGLPGP K GEPGIPAIPG
276	P02747	72	IPAIPGIRGP K GQKGEPGLPG
277	P02747	86	GEPGLPGHPG K NGPMGPPGMP
278	P02747	115	PGEPEGEGRY K QKFQSVFTVT
279	P02747	117	EPGEEGRY K QKFQSVFTVTRQ
280	P02747	157	PQGDYDTSTG K FTCKVPGLYY
281	P02747	161	YDTSTGKFT C KVPGLYYFVYH
282	P02747	188	LCVLLYRSGV K VVTFCGHTSK
283	P02747	198	KVVTFCGHTS K TNQVNSGGVL
284	P04142	5	IKAFNMNFA K ILSFVFALVL

285	P04142	29	MTSAAPEPRW K IFKKIEKMGR
286	P04142	32	AAPEPRWKIF K KIEKMGRNIR
287	P04142	33	APEPRWKIF K KIEKMGRNIRD
288	P04142	36	PRWKIFKKIE K MGRNIRDGIV
289	P04142	59	GPAIEVLGSA K AIGKKGIACA
290	P04142	63	EVLGSAKAIG K KGIAKASGLV
291	P04258	8	EYEEYEAYDV K SGVAGGGIAG
292	P04258	73	GPPGAIGPSG K DGESGRPGRP
293	P04258	128	EKGEPGAPGL K GENGVPGEDG
294	P04258	239	PPGSDGSPGG K GEMGPAGIPG
295	P04258	280	GQGAAGEPG K NGAKGDPGPR
296	P04258	284	AAGEPGKNGA K GDPGPRGERG
297	P04258	305	EAGSPGIAGP K GEDGKDGS PG
298	P04258	310	GIAGPKGEDG K DGSPGEPGAN
299	P04258	347	PAGANGLPGE K GPPGDRGGPG
300	P04258	394	GSPGGPGSNG K PGPPGSQGET
301	P04258	454	GGPGPQGPAG K NGETGPQGPP
302	P04258	473	PPGPTGPSGD K GDTGPPGPQG
303	P04258	499	GTSGPPGENG K PGEPGPKGEA
304	P04258	506	ENGKPGEPGP K GEAGAPGIPG
305	P04258	518	EAGAPGIPGG K GDSGAPGERG
306	P04258	584	ERGGPGGPGP K GDKGEPGSSG
307	P04258	587	GPGGPGPKGD K GEPGSSGVDG
308	P04258	601	GSSGVDGAPG K DGPRGPTGPI
309	P04258	623	PPGPAGQPGD K GESGAPGVPG
310	P04258	668	APGQNGEPGA K GERGAPGEKG
311	P04258	677	AKGERGAPGE K GEGPPGAAG
312	P04258	704	PAGPPGPQGV K GERGSPGGPG
313	P04258	745	GPPGSSGAPG K DGPPGPPGSN
314	P04258	767	APGSPGISGP K GDSGPPGERG
315	P04258	821	ARGSPGPQGI K GENGKPGPSG
316	P04258	826	GPQGIKGENG K PGPSGQNGER
317	P04258	875	LPGRDGAPGA K GDRGENSGPG
318	P04258	904	GPPGPVGPAG K SGDRGETGPA
319	P04258	994	GPVGPSPGPPG K DGASGHPGPI
320	P12108	51	VPGADGIDGD K GSPGAPGSPG
321	P12108	63	SPGAPGSPGA K GEPGAPGPDG
322	P12108	77	GAPGPDGPPG K PGLDGLTGAK
323	P12108	87	KPGLDGLTG K SGRGPWGGQG
324	P12108	99	SRGPWGGQGL K GQPGLPGPPG
325	P12108	137	LPGEIGVPGP K GDPGPDGPRG
326	P12108	154	GPRGPPGPPG K PGPPGHIQGV
327	P12108	190	PKGPQGLQGL K GHRGRPGALG
328	P12108	207	GALGEPGQQG K QGPKGDVGVS

329	P12108	211	EPGQQGKQGP K GDVGVSGEQG
330	P12108	241	QRGYPGMAGP K GETGPAGYKG
331	P12108	250	PKGETGPAGY K GMVGTIGAAG
332	P12108	268	AAGRPGREGP K GPPGDPGEKG
333	P12108	277	PKGPPGDPGE K GELGGRGIRG
334	P12108	295	IRGPQGDIGP K GDMGLPGIDG
335	P12108	306	GDMGLPGIDG K DGTPGIPGVK
336	P12108	316	KDGTPGIPGV K GTAGQPGRPG
337	P12108	343	QAGLPGQPGS K GGPGDKGEVG
338	P12108	349	QPGSKGGPGD K GEVGARGQQG
339	P12108	385	DAGTAGVPGL K GDRGERGPVG
340	P12108	406	APGEAGQSGP K GEQGPPGIPG
341	P12108	424	IPGPQGLPGV K GDKGSPGKTG
342	P12108	427	PQGLPGVKGD K GSPGKTGPKG
343	P12108	432	GVKGDKGSPG K TGPKGSTGDP
344	P12108	436	DKGSPGKTGP K GSTGDPGVHG
345	P12108	451	DPGVHGLAGV K GEKGESGEPG
346	P12108	454	VHGLAGVKGE K GESGEPGPKG
347	P12108	463	EKGESGEPGP K GQQGIQGELG
348	P12108	529	GDQHIIDVVL K MMQEQLAEVA
349	P12108	543	EQLAEVAVSA K RAALGGVGAM
350	P12108	595	AAGQIGNIGP K GKRGEKGERG
351	P12108	597	GQIGNIGPKG K RGEKGERGDT
352	P12108	601	NIGPKGKRGE K GERGDTGRGH
353	P12108	634	PGIPGHALAG K DGERGPPGVP
354	P12111	3	PLHRHKRMMR K HRHLPLVAVF
355	P12111	33	THAQQQQADV K NGAAADIIFL
356	P12111	67	LVREFLYDVV K SLAVGENDFH
357	P12111	99	TEFLLNTYRT K QEVLSHISNM
358	P12111	120	SYIGGTNQTG K GLEYIMQSHL
359	P12111	132	LEYIMQSHLT K AAGSRAGDGV
360	P12111	155	VIVVLTDGHS K DGLALPSAEL
361	P12111	166	DGLALPSAEL K SADNVNFAIG
362	P12111	186	GVEDADEGAL K EIASEPLNMH
363	P12111	233	PERAGDTETL K DITAQDSADI
364	P12111	271	ILDFLVNLL K LPIGTQQIRV
365	P12111	302	TMFSLDTYST K AQVLGAVKAL
366	P12111	310	STKAQVLGAV K ALGFAGGELA
367	P12111	370	DEIRYGVVAL K QASVFSFGLG
368	P12111	414	EFRSFGDLQ K LLPYIVGVAQ
369	P12111	430	VGVAQRHIVL K PPTIVTQVIE
370	P12111	443	TIVTQVIEVN K RDIVFLVDGS
371	P12111	470	NFNAIRDFIA K VIQRLEIGQD
372	P12111	505	PEFYFNTHPT K REVITAVRKM

373	P12111	514	TKREVITAVR K MKPLDGSALY
374	P12111	516	REVITAVR K M K PLDGSALYTG
375	P12111	551	AGYRAAEGIP K LLVLITGGKS
376	P12111	560	PKLLVLITGG K SLDEISQPAQ
377	P12111	573	DEISQPAQEL K RSSIMAFaIG
378	P12111	585	SSIMAFaIGN K GADQAELEEI
379	P12111	637	LSGTPEVHSN K RDIIFLLDGS
380	P12111	652	FLLDGSANVG K TNFPYVRDFV
381	P12111	699	TEFSLNTYQT K SDILGHLRQL
382	P12111	783	TFCVGASQAN K AELEQIAFNP
383	P12111	835	EEVPLAQPE S K RDILFLFDGS
384	P12111	861	QFPVVRDFLY K IIDELNVKPE
385	P12111	869	LYKIIDELNV K PEGTRIAVAQ
386	P12111	885	IAVAQYSDDV K VESRFDEHQS
387	P12111	896	VESRFDEHQS K PEILNLVKRM
388	P12111	904	QSKPEILNLV K RMKIKTGKAL
389	P12111	907	PEILNLVKRM K IKTGKALNLG
390	P12111	909	ILNLVKRM K IKTGKALNLGYA
391	P12111	912	LVKRMKIKT G KALNLGYALDY
392	P12111	930	LDYAQRYIFV K SAGSRIEDGV
393	P12111	964	DRVDGPASNL K QSGVVPFIFQ
394	P12111	976	SGVVPFIFQA K NADPAELEQI
395	P12111	1001	AFILAAESLP K IGDLHPQIVN
396	P12111	1014	DLHPQIVNLL K SVHNGAPAPV
397	P12111	1028	NGAPAPVS G E K DVVFLLDGSE
398	P12111	1048	EGVRSGFPLL K EFVQRVVESL
399	P12111	1088	PEFYLN S YMN K QDVVNAVRQL
400	P12111	1156	DDVRNPSVVV K RGGAVPIGIG
401	P12111	1231	QPLPSPGVGG K RDVVFLIDGS
402	P12111	1281	VAVIQFSDDP K VEFLLNAHSS
403	P12111	1292	VEFLLNAHSS K DEVQNAVQRL
404	P12111	1305	VQNAVQRLRP K GGRQINVGNA
405	P12111	1325	ALEYVSRNIF K RPLGSRIEEG
406	P12111	1347	PQFLVLISS G K SDDEVDDPAV
407	P12111	1360	DEVDDPAVEL K QFGVAPFTIA
408	P12111	1380	ARNADQEELV K ISLSPEYVFS
409	P12111	1403	TFRELPSLEQ K LLTPITTLTS
410	P12111	1418	ITTLTSEQIQ K LLASTRYPPP
411	P12111	1472	IVRRLNIGPS K VRVGVVQFSN
412	P12111	1491	SNDVFPEFY L K TYRSQAPVLD
413	P12111	1518	LRGGSPLNT G K ALEFVARNLF
414	P12111	1530	LEFVARNLFV K SAGSRIEDGV
415	P12111	1551	PQHLVLVLGG K SQDDVSRAQ
416	P12111	1635	VDTPPPSRPE K KKADIVFLLD

417	P12111	1636	DTPPPSRPEK K KADIVFLLDG
418	P12111	1637	TTPPPSRPEK K KADIVFLLDGS
419	P12111	1694	NSDPTDEFFL K DfstKRQIID
420	P12111	1699	DEFFLKDFST K RQIIDAINKV
421	P12111	1708	TKRQIIDAIN K VVYKGGRHAN
422	P12111	1712	IIDAINKV V K GGRHANTKVG
423	P12111	1720	VYKGGRHANT K VGLEHLRVNH
424	P12111	1754	PQIAFVITGG K SVEDAQDVSL
425	P12111	1772	VSLALTQRGV K VFAVGVRNID
426	P12111	1788	VRNIDSEEV G KIASNSATAFR
427	P12111	1833	TLCPGVTDAA K ACNLDVILGF
428	P12111	1856	SRDQNVFVAQ K GFESKVDAIL
429	P12111	1861	VFVAQKGFES K VDAILNRISQ
430	P12111	1913	FDEYQPEMLE K FRNMRSQHPY
431	P12111	1931	HPYVLTEDTL K VYLNKFRQSS
432	P12111	1936	TEDTLKVYLN K FRQSSPDSVK
433	P12111	1946	KFRQSSPDSV K VVIHFTDGAD
434	P12111	2027	LAEQLDNIAE K ACCGVPCKCS
435	P12111	2035	AEKACCGVP C KCSGQRGDRGP
436	P12111	2052	DRGPIGSIG P KGIPGEDGYRG
437	P12111	2094	FQGCPCQRGV K GSRGFPGKEG
438	P12111	2121	LDGLDGEDGD K GLPGSSGEKG
439	P12111	2130	DKGLPGSSGE K GPNPGRRGDKG
440	P12111	2139	EKGPNPGRRGD K GPRGEKGERG
441	P12111	2145	RRGDKGPRGE K GERGDVGIRG
442	P12111	2170	PGQDSQERGP K GETGDLGPMG
443	P12111	2196	GVPGGPGET G KNGGFRRGPP
444	P12111	2272	ERGRTGPLGR K GEPGEPGPKG
445	P12111	2281	RKGEPGEPGP K GGIGNRGPRG
446	P12111	2309	DGVGSEGRRG K KGERGFPGYP
447	P12111	2310	GVGSEGRRG K KGERGFPGYPG
448	P12111	2355	NSGPPGIVGQ K GDPGYPGPAG
449	P12111	2367	DPGYPGPAG P KGNRGDSIDQC
450	P12111	2384	IDQCALIQSI K DKCPCCYGPL
451	P12111	2386	QCALIQSI K DKCPCCYGPLEC
452	P12111	2464	VTTEIRFADS K RKSVLLDKIK
453	P12111	2466	TEIRFADSKR K SVLLDKIKNL
454	P12111	2472	DSKRKSVLLD K IKNLQVALTS
455	P12111	2474	KRKSVLLDKI K NLQVALTSKQ
456	P12111	2483	IKNLQVALTS K QQSLETAMSF
457	P12111	2500	AMSFVARNT F KRVRNGFLMRK
458	P12111	2510	KRVRNGFLMR K VAVFFSNTPT
459	P12111	2532	ASPQLREAVL K LSDAGITPLF
460	P12111	2639	TTTLFQFNEM K KYIAYLVRQL

461	P12111	2640	TTLFQFNEMK K YIAYLVRQLD
462	P12111	2656	VRQLDMSPDP K ASQHFARVAV
463	P12111	2684	SVDNASMPVP K VEFSLTDYGS
464	P12111	2695	VEFSLTDYGS K EKLVDFLSRG
465	P12111	2697	FSLTDYGS K EKLVDFLSRGMT
466	P12111	2737	FESAPNPRDL K IVVLMMLTGEV
467	P12111	2763	EEAQRVILQA K CKGYFFVVLG
468	P12111	2765	AQRVILQAK K GYFFVVLGIG
469	P12111	2777	YFFVVLGIGR K VNIKEVYTFA
470	P12111	2781	VLGIGRKVN K EVYTFASEPN
471	P12111	2796	FASEPNDVFF K LVDKSTELNE
472	P12111	2800	PNDVFFKLVD K STELNEEPLM
473	P12111	2834	NAFYLSPDIR K QCDWFQGDQP
474	P12111	2846	CDWFQGDQPT K NLVKFGHKQV
475	P12111	2850	QGDQPTKNLV K FGHKQVNVPN
476	P12111	2854	PTKNLVKFGH K QVNVPNNVTS
477	P12111	2875	SPTSNPVTTTT K PVTTTTKPVT
478	P12111	2881	VTTTKPVTTTT K PVTTTTTKPVT
479	P12111	2888	TTTKPVTTTT K PVTTTTTKPVT
480	P12111	2895	TTTKPVTTTT K PVTIINQPSV
481	P12111	2906	PVTIINQPSV K PAAAKPAPAK
482	P12111	2911	NQPSVKPAA K PAPAKPVAAK
483	P12111	2916	KPAAAKPAP K PVAAKPVATK
484	P12111	2921	KPAPAKPVAA K PVATKMATVR
485	P12111	2926	KPVAAKPVAT K MATVRPPVAV
486	P12111	2937	MATVRPPVAV K PATAAKPVAA
487	P12111	2943	PVAVKPATA K PVAAKPAAVR
488	P12111	2948	PATAAKPVAA K PAAVRPPAAA
489	P12111	2961	AVRPPAAAA K PVATKPEVPR
490	P12111	2966	AAAAAKPVAT K PEVPRPQAAK
491	P12111	2976	KPEVPRPQAA K PAATKPATTK
492	P12111	2981	RPQAAKPAAT K PATTKPMVKM
493	P12111	2986	KPAATKPAT K PMVKMSREVQ
494	P12111	2990	TKPATTKPMV K MSREVQVFEI
495	P12111	3006	QVFEITENSA K LHWERAEPFG
496	P12111	3035	TSAMDQSLVL K QNLTVTDRVI
497	P12111	3077	RATYHGSFST K KSQPPPPQPA
498	P12111	3078	ATYHGSFST K KSQPPPPQPAR
499	P12111	3113	PLALTETD I C K LPKDEGTCRD
500	P12111	3116	LTETDICKLP K DEGTCRDFIL
501	P12111	3127	DEGTCRDFIL K WYYDPNTKSC
502	P12111	3135	ILKWYYDPNT K SCARFWYGGC
503	P12111	3151	WYGGCGGNEN K FGSQKECEKV
504	P12111	3156	GGNENKFGS K QKECEKVCAPVL

505	P12111	3160	NKFGSQKECEKVCAPVLAKPG
506	P12111	3168	CEKVCAPVLAKEPGVISVMGTT
507	P19999	28	SGSQTCETLKTCSVIACGRD
508	P19999	63	GLRGLQGPPGKLGPPGSVGAP
509	P19999	94	DRGDSRAIEVKLANMEAEINT
510	P19999	106	ANMEAEINTLKSKELELTNKLH
511	P19999	108	MEAEINTLKSKELELTNKLHAF
512	P19999	114	TLKSKELELTNKLHAFSMGKKS
513	P19999	122	TNKLHAFSMGKKS GKKFFVTN
514	P19999	123	NKLHAFSMGKKS GKKFFVTNH
515	P19999	126	HAFSMGKKS GKKFFVTNHERM
516	P19999	127	AFSMGKKS GKKFFVTNHERMP
517	P19999	140	VTNHERMPFSKV KALCSELRG
518	P19999	142	NHERMPFSKV KALCSELRGTV
519	P19999	162	VAIPRNAEENKAIQEVAKTSA
520	P19999	169	EENKAIQEVAKTSAFLGITDE
521	P19999	199	TGGRLTYSNWKDEPNDHGSG
522	P19999	200	GGRLTYSNWKDEPNDHGSGE
523	P20908	8	VDMMDVHTRWKARSALRPGAP
524	P20908	44	SRAAQPADLLKVLD FHNLPDG
525	P20908	57	DFHNLPDGITKT TGFCATRRS
526	P20908	69	TGFCATRRSSKGP DVAYRVTK
527	P20908	79	KGP DVAYRVTKDAQLSAPTQKQ
528	P20908	88	TKDAQLSAPTQQLYPASAFPE
529	P20908	107	PEDFSILT TVKAKKGSQAFLV
530	P20908	109	DFSILT TVKAKKGSQAFLVSI
531	P20908	110	FSILT TVKAKKGSQAFLVSIY
532	P20908	146	PVFLYEDHTGKPGPEDYPLFR
533	P20908	164	LFRGINLSDGKWHRIALSVHK
534	P20908	174	KWHRIALSVHKKNVT LILDCK
535	P20908	175	WHRIALSVHKKNVT LILDCKK
536	P20908	184	KKNVT LILDCKKKTTKFLDRS
537	P20908	185	KNVT LILDCKKKTTKFLDRSD
538	P20908	186	NVT LILDCKKKTTKFLDRSDH
539	P20908	189	LILDCKKKTTKFLDRSDHPMI
540	P20908	288	PYYEDPEDLGKEPTPSKKPVE
541	P20908	294	EDLGKEPTPSKKPVEAAKETTT
542	P20908	295	DLGKEPTPSKKPVEAAKETTE
543	P20908	301	TPSKKPVEAAKETTEVPEELT
544	P20908	329	PMPETSEGAGKEEDVGIGDYD
545	P20908	452	YEGIGGPRGEGQKGEP AIIIE
546	P20908	455	IGGPRGEGQKGEP AIIIEPGM
547	P20908	535	RFGGGGDAGSKGPMVSAQESQ
548	P20908	582	PVGPPGSGGLKGEPGDVGPQG

549	P20908	605	GVQGPYPGAG K PGRRGRAGSD
550	P20908	687	EPGPRGLLGP K GPPGPPGPPG
551	P20908	749	GPPGEKGPLG K PGLPGMPGAD
552	P20908	767	GADGPPGHPG K EGPPGEKGGQ
553	P20908	825	FPGFKGDMGI K GDRGEIGPPG
554	P20908	866	GPLGPPGEKG K LGVPLPGYP
555	P20908	905	GEKGGRGTPG K PGPRGQRGPT
556	P20908	929	GERGPRGITG K PGPKGNSGGD
557	P20908	933	PRGITGKPGP K GNSGGDGPAG
558	P20908	963	PQGPTGFPGP K GPPGPPGKDG
559	P20908	971	GPKGPPGPPG K DGLPGHPGQR
560	P20908	989	GQRGETGFQG K TGPPGPPGVV
561	P20908	1034	GEQGLPGLAG K EGTKGDPGPA
562	P20908	1038	LPGLAGKEGT K GDPGPAGLPG
563	P20908	1049	GDPGPAGLPG K DGPPGLRGFP
564	P20908	1074	LPGPVGALGL K NEGPPGPPG
565	P20908	1119	PQGPPGPAGE K GAPGEKGPQG
566	P20908	1125	PAGEKGAPGE K GPQGPAGRDG
567	P20908	1158	PVGPPGEDGD K GEIGEPGQKG
568	P20908	1167	DKGEIGEPGQ K GSKGDKGEQG
569	P20908	1170	EIGEPGQKGS K GDKGEQGPPG
570	P20908	1173	EPGQKGSKGD K GEQGPPGPTG
571	P20908	1212	PRGQQGLFGQ K GDEGPRGFPG
572	P20908	1239	LQGLPGPPGE K GETGDVGQMG
573	P20908	1284	GIGNPGAVGE K GEPGEAGEPG
574	P20908	1305	LPGEggppGP K GERGEKGESG
575	P20908	1311	PPGPKGERGE K GESGPSGAAG
576	P20908	1326	PSGAAGPPGP K GPPGDDGPKG
577	P20908	1335	PKGPPGDDGP K GSPGPVGFPG
578	P20908	1365	PAGQDGP PGD K GDDGEPGQTG
579	P20908	1391	GEPGPSGPPG K RGPPGPAGPE
580	P20908	1407	PAGPEGRQGE K AKGEAGLEG
581	P20908	1410	PEGRQGEKG K GEAGLEGPPG
582	P20908	1421	GEAGLEGPPG K TGPIGPQGAP
583	P20908	1433	GPIGPQGAPG K PGPDGLRGIP
584	P20908	1473	PMGPPGLPGL K GDSGPKGEKG
585	P20908	1479	LPGLKGDSGP K GEKGHPGLIG
586	P20908	1482	LKGDSGPKGE K GHPGLIGLIG
587	P20908	1500	LIGPPGEQGE K GDRGLPGPQG
588	P20908	1515	LPGPQGSSGP K GEQGITGPSG
589	P20908	1542	PPGLPGPPGP K AKGSSGPTG
590	P20908	1545	LPGPPGPKG K GSSGPTGPKG
591	P20908	1554	AKGSSGPTGP K GEAGHPGPPG
592	P20908	1619	EEIFGSLNSL K LEIEQMKRPL

593	P20908	1626	NSLKLEIEQM K RPLGTQQNPA
594	P20908	1640	GTQQNPARTC K DLQLCHPDFP
595	P20908	1668	PNQGCSRDSF K VYCNFTAGGS
596	P20908	1685	AGGSTCVFPD K KSEGARITSW
597	P20908	1686	GGSTCVFPD K KSEGARITSWP
598	P20908	1697	SEGARITSWP K ENPGSWFSEF
599	P20908	1708	ENPGSWFSEF K RGKLLSYVDA
600	P20908	1711	GSWFSEFKRG K L LSYVDAEGN
601	P20908	1762	WQDAATGSYD K ALRFLGSNDE
602	P20908	1792	IRALVDGCAT K KGYQKTVLEI
603	P20908	1793	RALVDGCAT K KGYQKTVLEID
604	P20908	1797	DGCATKKGYQ K TVLEIDTPKV
605	P20908	1806	QKTVLEIDTP K VEQVPIVDIM
606	P20908	1826	MFNDFGEASQ K FGFEVGPACF
607	P23805	29	LGAEMTTFSQ K ILANACTLVM
608	P23805	122	GPAGREGPSG K QGSMPGPTP
609	P23805	222	ESGLAEVNAL K QRV TILDGHL
610	P23805	243	RRFQNAFSQY K KAVLFPDGQA
611	P23805	244	RFQNAFSQY K KAVLFPDGQAV
612	P23805	257	LFPDGQAVGE K IFKTAGAVKS
613	P23805	260	DGQAVGEK I FKTAGAVKSYSD
614	P23805	266	EK I FKTAGAV K SYSDAEQLCR
615	P23805	279	SDAEQLCREA K GQLASPRSSA
616	P23805	303	AVTQMVRAQE K NAYLSMNDIS
617	P23805	355	ENCVEIFPDG K WNDVPCSKQL
618	P23805	363	DGKWNDVPCS K QLLVICEFFE
619	P26368	20	QLNENKQERD K ENRHRKRSHS
620	P26368	26	QERDKENRHR K RSRSRSRSD
621	P26368	38	SHRSRSRDR K RRSRSRDRRN
622	P26368	63	SASRDRRRRS K PLTRGAKEEH
623	P26368	70	RRSKPLTRGA K EEHGGLIRSP
624	P26368	84	GGLIRSPRHE K KKKVRKYWDV
625	P26368	85	GLIRSPRHE K KKKVRKYWDVP
626	P26368	86	LIRSPRHE K KKVRKYWDVPP
627	P26368	87	IRSPRHE K KKKVRKYWDVPPP
628	P26368	90	PRHE K KKKVRKYWDVPPPGFE
629	P26368	108	GFEHITPMQY K AMQAAGQIPA
630	P26368	195	PVLAVQINQD K NFAFLEFRSV
631	P26368	225	DG I IFQGQSL K IRRP HDYQPL
632	P26368	260	VSTVVPDSA H KLFIGGLPNYL
633	P26368	286	KELLTSFGPL K AFNLVKDSAT
634	P26368	292	FGPLKAFNLV K DSATGLSKGY
635	P26368	300	LVKDSATGLS K GYAFCEYVDI
636	P26368	328	AGLNGMQLGD K KLLVQRASVG

637	P26368	329	GLNGMQLGDK K LLVQRASVGA
638	P26368	340	LLVQRASVGA K NATLVSPFST
639	P26368	408	IVEDVRDECS K YGLVKSIEIP
640	P26368	413	RDECSKYGLV K SIEIPRPVDG
641	P26368	431	VDGVEVPGCG K IFVEFTSVFD
642	P26368	444	VEFTSVFDCQ K AMQGLTGRKF
643	P26368	453	QKAMQGLTGR K FANRVVVTKY
644	P26368	462	RKFANRVVV T KYCDPDSYHRR
645	P30754	50	GDSGPMGP I G K RGPPGPAGIA
646	P30754	62	GPPGPAGIAG K SGDDGRDGEP
647	P30754	113	GLSGSKGEQ G K SGNQGPDGGP
648	P30754	146	GQTGERGRDG K SGLPGLRGVD
649	P30754	299	GAPGHAGEAG K RGSPGSPGPA
650	P30754	383	GNPGSDGRPG K DGRPGIRGKD
651	P30754	392	GKDGRPGIRG K DGKQGEQGPQ
652	P30754	395	GRPGIRGKD G K QGEQGPQGPQ
653	P30754	425	GPPGARGE P G K NGAPGEPGAH
654	P30754	443	GAHQEQGDAG K DGETGAAGPP
655	P30754	488	GAQGTPGEAG K TGERGAVGAT
656	P30754	584	GSRGDIGPRG K AGERGKDGGER
657	P30754	590	GPRGKAGERG K DGERGERGEN
658	P30754	674	GERGTRGATG K QGARGPRGLA
659	P30754	686	GARGPRLAG K RGLRGAGGSR
660	P30754	734	GPSGPAGTAG K QGVKGARGSP
661	P30754	749	GARGSPGLVG K QGDRGSDGEP
662	P30754	851	GAPGQEGAPG K DGLPGLAGRP
663	P30754	986	GGRGSQGPPG K DGQPGPSGRV
664	P30943	4	AI IKASMM S A K IIAFSAVVAT
665	P30943	44	SVNMAMDK S A K APVITIFDHR
666	P30943	61	FDHRGCSRAP K EYTGAKAGGK
667	P30943	67	SRAPKEYTG A K AGGKDDEMMV
668	P30943	71	KEYTGAKAGG K DDEMMVKAQS
669	P30943	78	AGGKDDEMMV K AQSVKIEVST
670	P30943	83	DEMMVKAQSV K IEVSTGTAE
671	P30943	101	AEGVLATSLA K MTKKTMKALS
672	P30943	104	VLATSLAKMT K KTMKALSTAL
673	P35248	23	QPLGDLGAEM K TLSQRSITNT
674	P35248	62	RDGREGPRGE K GDPGLPGPMG
675	P35248	118	GSPGISGPAG K EGPSGKQGN
676	P35248	124	GPAGKEGPS G K QGNIGPQGKP
677	P35248	133	GKQGNIGPQ G K PGPKGEAGPK
678	P35248	137	NIGPQGKPG P K GEAGPKGEVG
679	P35248	143	KPGPKGEAG P K GEVGAPGMQG
680	P35248	158	APGMQGSAG A K GPAGPKGERG

681	P35248	164	SAGAKGPAGP K GERGAPGEQG
682	P35248	203	APGSRGPPGL K GDRGAPGDRG
683	P35248	215	DRGAPGDRGI K GESGLPDSAA
684	P35248	236	LRQQMEALNG L QRLEAAFSR
685	P35248	248	QRLEAAFSRY K KAALFPDQGS
686	P35248	249	RLEAAFSRY K KAALFPDQSV
687	P35248	262	LFPDQGSVGD K IFRAANSEEP
688	P35248	277	ANSEEPFEDA K EMCRQAGGQL
689	P35248	308	AVQQLVTAHS K AAFLSMTDVG
690	P35248	322	LSMTDVGTEG K FTYPTGEALV
691	P35248	362	EIFTNGQWND K ACGEQRLVIC
692	P83414	1	KEIKKFVKWK K WKVFKKIEKV
693	P83414	3	IKKFVKWKK K VFKKIEKVGR
694	P83414	6	FVKWKKWKV F KKIEKVGRNIR
695	P83414	7	VKWKKWKV F KKIEKVGRNIRD
696	P83414	10	KKWKVFKKIE K VGRNIRDGIV
697	P83415	3	IKKFVKWRRW K VFKKIEKMGR
698	P83415	6	FVKWRRWKV F KKIEKMGRNIR
699	P83415	7	VKWRRWKV F KKIEKMGRNIRD
700	P83415	10	RRWKVFKKIE K MGRNIRDGVI
701	P83415	33	APAIEVLGQ A KKAQGLVEIAP
702	P84698	34	NKACSGCTHG K KGHTCGSCAK
703	Q00433	4	ALTKAFMMF A KTLASLAVIGS
704	Q00433	59	PAGAAMDKSA K APQITIFDHR
705	Q00433	76	FDHRGCSRAP K ESTGGKAGGQ
706	Q00433	82	SRAPKESTGG K AGGQDDEMMV
707	Q00433	93	AGGQDDEMMV K VASTKVTVSE
708	Q00433	98	DEMMVKVAST K VTVSESDAAK
709	Q00433	108	KVTVSESDA A KKLQEFITFEK
710	Q00433	109	VTVSESDA A KKLQEFITFEKG
711	Q00433	118	KKLQEFITFE K GIDGPFTSKN
712	Q00433	127	EKGIDGPFTS K NNKSTFPGDI
713	Q02388	113	VIRAIRELSY K GGNTRTGAAI
714	Q02388	142	LPQLARPGVP K VCILITDGKS
715	Q02388	151	PKVCILITDG K SQDLVDTAAQ
716	Q02388	164	DLVDTAAQRL K GQGVKLFAVG
717	Q02388	169	AAQRLKGQGV K LFAVGIKNAD
718	Q02388	176	QGVKLFAVG I KNADPEELKRV
719	Q02388	184	GIKNADPEEL K RVASQPTSDF
720	Q02388	263	TAASGPVTGY K VQYTPLTGLG
721	Q02388	459	RRETGLEPP Q KVVLPSDVTRY
722	Q02388	727	VSWHSAHGPE K SQLVSGEATV
723	Q02388	1253	RPEPCPVYCP K GQKGEPGEMG
724	Q02388	1256	PCPVYCPKG Q KGEPGEMGLRG

725	Q02388	1295	PQGPPGSATA K GERGFPGADG
726	Q02388	1325	NPGTPGAPGL K GSPGLPGPRG
727	Q02388	1346	DPGERGPRGP K GEPGAPGQVI
728	Q02388	1367	GGEGLPLPGR K GDPGPSGPPG
729	Q02388	1399	GPPGLPGTAM K GDKGDRGERG
730	Q02388	1402	GLPGTAMKGD K GDRGERGPPG
731	Q02388	1443	GPQGPVGPPG K KGEKGDSEGD
732	Q02388	1444	PQGPVGPPG K KGEKGDSEDDA
733	Q02388	1447	PVGPPGKKGE K GDSDDGAPGL
734	Q02388	1479	PRGPPGAIGP K GDRGFPGPLG
735	Q02388	1494	FPGPLGEAGE K GERGPPGPAG
736	Q02388	1518	LPGVAGRPGA K GPEGPPGPTG
737	Q02388	1533	PPGPTGRQGE K GEPGRPGDPA
738	Q02388	1553	AVVGPAVAGP K GEKGDVGPAG
739	Q02388	1556	GPAVAGPKGE K GDVGPAGPRG
740	Q02388	1588	GLVLPGDGP K GDPGDRGPIG
741	Q02388	1615	PPGDSGPPGE K GDPGRPGPPG
742	Q02388	1639	PRGRDGEVGE K GDEGPPGDPG
743	Q02388	1653	GPPGDPGLPG K AGERGLRGAP
744	Q02388	1672	APGVRGPVGE K GDQGDPEGDG
745	Q02388	1693	RNGSPGSSGP K GDRGEPGPPG
746	Q02388	1718	LVDTGPGARE K GEPGDRGQEG
747	Q02388	1733	DRGQEGPRGP K GDPGLPGAPG
748	Q02388	1769	DPGVRGPAGE K GDRGPPGLDG
749	Q02388	1786	GLDGRSGLDG K PGAAGPSGPN
750	Q02388	1801	GPSGPNGAAG K AGDPGRDGLP
751	Q02388	1831	GPSGPPGLPG K PGEDGKPGLN
752	Q02388	1837	GLPGKPGEDG K PGLNGKNGEP
753	Q02388	1843	GEDGKPGLN K NGEPGDPGED
754	Q02388	1856	EPGDPGEDGR K GEKGDSGASG
755	Q02388	1859	DPGEDGRKGE K GDSGASGREG
756	Q02388	1874	ASGREGRDGP K GERGAPGILG
757	Q02388	1912	FPGVPGGTGP K GDRGETGSKG
758	Q02388	1921	PKGDRGETGS K GEQGLPGERG
759	Q02388	1952	VDRLLETAGI K SALREIVET
760	Q02388	1981	LPVPERRRG K GDSGEQGPPG
761	Q02388	1992	GDSGEQGPPG K EGPIGFPPER
762	Q02388	2005	PIGFPPERGL K GDRGDPGPQG
763	Q02388	2038	GPSGLAGEPG K PGIPGLPGRA
764	Q02388	2066	RPGERGERGE K GERGEQGRDG
765	Q02388	2093	TPGPPGPPGP K VSVDEPGPGL
766	Q02388	2113	LSGEQGPPGL K AKGEPGSNG
767	Q02388	2116	EQGPPGLKGA K GEPGSNGDQG
768	Q02388	2128	EPGSNGDQGP K GDRGVPGIKG

769	Q02388	2137	PKGDRGVPGI K GDRGEPGPRG
770	Q02388	2166	GERGMAGPEG K PGLQGPRGPP
771	Q02388	2206	PAGPQGPSGL K GEPGETGPPG
772	Q02388	2255	GLPGQVGETG K PGAPGRDGAS
773	Q02388	2267	GAPGRDGASG K DGDRGSPGV
774	Q02388	2289	SPGLPGPVGP K GEPGPTGAPG
775	Q02388	2309	GQAVVGLPGA K GEKGAPGGLA
776	Q02388	2312	VVGLPGA K GEKGAPGGLAGDL
777	Q02388	2329	AGDLVGEPGA K GDRGLPGPRG
778	Q02388	2341	DRGLPGPRGE K GEAGRAGEPG
779	Q02388	2359	EPGDPGEDGQ K GAPGPKGFKG
780	Q02388	2365	EDGQKGAPGP K GFKGDPGVGV
781	Q02388	2368	QKGAPGPKGF K GDPGVGVPGS
782	Q02388	2388	SPGPPGPPGV K GDLGLPGLPG
783	Q02388	2457	SVGPPGASGL K GDKGDPGVGL
784	Q02388	2460	PPGASGLKGD K GDPGVGLPGP
785	Q02388	2507	PPGSRGERGE K GDVGSAGLKG
786	Q02388	2516	EKGDVGSAGL K GDKGDSAVIL
787	Q02388	2519	DVGSAGLKG K GDSAVILGPP
788	Q02388	2535	ILGPPGPRGA K GDMGERGPRG
789	Q02388	2550	ERGPRGLDGD K GPRGDNGDPG
790	Q02388	2562	PRGDNGDPGD K GSKGEPGDKG
791	Q02388	2565	DNGDPGDKGS K GEPGDKGSAG
792	Q02388	2571	DKGSKGEPGD K GSAGLPGLRG
793	Q02388	2603	GIPGDPGSPG K DGVPGIRGEK
794	Q02388	2613	KDGVPGIRGE K GDVGFMGPRG
795	Q02388	2640	VKGACGLDGE K GDKGEAGPPG
796	Q02388	2643	ACGLDGEKGD K GEAGPPGRPG
797	Q02388	2658	PPGRPGLAGH K GEMGEPGVPG
798	Q02388	2675	GVPGQSGAPG K EGLIGPKGDR
799	Q02388	2682	APGKEGLIG K GDRGFDGQPG
800	Q02388	2694	DRGFDGQPGP K GDQGEKGERG
801	Q02388	2700	QPGPKGDQGE K GERGTPGIGG
802	Q02388	2742	PRGPEGLQGQ K GERGPPGERV
803	Q02388	2780	RPGPAGPRGE K GEAALTEDDI
804	Q05707	2	RMKRQFIKMM K IFQRKMRYWL
805	Q05707	7	FIKMMKIFQR K MRYWLLPPFL
806	Q05707	51	ISHDSIQISW K APRGKFGGYK
807	Q05707	56	IQISWKAPRG K FGGYKLLVTP
808	Q05707	61	KAPRGKFGGY K LLVTPPTSGGK
809	Q05707	71	KLLVTPPTSGG K TNQLNLQNTA
810	Q05707	83	NQLNLQNTAT K AI IQGLMPDQ
811	Q05707	104	NYTVQIIAYN K DKESKPAQGGQ
812	Q05707	106	TVQIIAYNKD K ESKPAQGGQFR

813	Q05707	109	IIAYNKDKESKPAQGGQFRIKD
814	Q05707	118	SKPAQGGQFRIKDLEKRKDPKP
815	Q05707	122	QGQFRIKDLEKRKDPKPRVKV
816	Q05707	124	QFRIKDLEKRKDPKPRVKVVD
817	Q05707	127	IKDLEKRKDPKPRVKVVDGRN
818	Q05707	131	EKRKDPKPRVKVVDGRNGSRP
819	Q05707	148	GSRPSSPEEVKFVCQTPAIAD
820	Q05707	194	LVTAFDVGSEKTRIGLAQYSG
821	Q05707	218	IEWHLNAFSTKDEVIEAVRNL
822	Q05707	231	VIEAVRNLPHYKGGNTLTGLAL
823	Q05707	250	ALNYIFENSFKPEAGSRTGVS
824	Q05707	261	PEAGSRTGVSIGILITDGKS
825	Q05707	270	SKIGILITDGKSQDDIIPPSR
826	Q05707	295	SGVELFAIGVKNADVNELQEI
827	Q05707	346	SRVEEQDREIKASAHAITGPP
828	Q05707	382	NWTHAPGNVEKYRVVYYPTRG
829	Q05707	394	RVVYYPTRGGKPDEVVVDGTV
830	Q05707	410	VDGTVSSTVLKNLMSLTEYQI
831	Q05707	462	YDVTENSMRVKWDAVPGASGY
832	Q05707	488	PLTEGLAGDEKEMKIGETHTD
833	Q05707	491	EGLAGDEKEMKIGETHTDIEL
834	Q05707	591	EVDPIITTFPLKGLTPLTEYTI
835	Q05707	655	HPLSADEGLHKLMWIPVYGGK
836	Q05707	665	KLMWIPVYGGKTEEVVLKEEQ
837	Q05707	672	YGGKTEEVVLKEEQDSHVIEG
838	Q05707	754	GDETTSSLRVKWDISDSDVQQ
839	Q05707	796	VPGSQNNLLLKPLLPDTEYKV
840	Q05707	805	LKPLLPDTEYKVTVTPIYTDG
841	Q05707	825	GEGVSVSAPGKTLPSGPGQNL
842	Q05707	857	ITWDPSSSPVKGYRIVYKPV
843	Q05707	864	SPVKGYRIVYKPVSVPGPTLE
844	Q05707	898	NLLSGMDYNVKIFASQASGFS
845	Q05707	916	GFSDALTGMVKTLFLGVTNLQ
846	Q05707	928	LFLGVTNLQAKHVEMTSLCAH
847	Q05707	960	VVIESLQDRQKQESTVGGGTT
848	Q05707	984	FYGLQPDSEYKISVYTKLQEI
849	Q05707	990	DSEYKISVYTKLQEI EGPSVS
850	Q05707	1004	IEGPSVSIMEKTQSLPTRPPT
851	Q05707	1023	PTFPPTIPPAKEVCKAAKADL
852	Q05707	1027	PTIPPAKEVCKAAKADLVFMV
853	Q05707	1030	PPAKEVCKAAKADLVFMVDGS
854	Q05707	1050	SWSIGDENFNKIISFLYSTVG
855	Q05707	1064	FLYSTVGALNKIGTDGTQVAM
856	Q05707	1086	QFTDDPRTEFKLNAYKTKETL

857	Q05707	1091	PRTEFKLNAY K TKETLLDAIK
858	Q05707	1093	TEFKLNAYKT K ETLLDAIKHI
859	Q05707	1101	KTKETLLDAI K HISYKGGNTK
860	Q05707	1106	LLDAIKHISY K GGNTKTGKAI
861	Q05707	1111	KHISYKGGNT K TGKAICYVRD
862	Q05707	1114	SYKGGNTKT G KAIKYVRDTLF
863	Q05707	1117	GGNTKTGKA I KYVRDTLFTAE
864	Q05707	1136	AESGTRRGIP K VIVVITDGRS
865	Q05707	1152	TDGRSQDDVN K ISREMQLDGY
866	Q05707	1183	DYSELVSI G S K PSARHVFFVD
867	Q05707	1199	VFFVDDFDAF K KIEDELITFV
868	Q05707	1200	FFVDDFDAF K KIEDELITFVC
869	Q05707	1222	TASATCPVVH K DGIDLAGFKM
870	Q05707	1231	HKDGIDLAGF K MMEMFGLVEK
871	Q05707	1241	KMMEMFGLVE K DFSSVEGVSM
872	Q05707	1266	FNVFPCYQLH K DALVSQPTRY
873	Q05707	1311	EPFALWEILN K NSDPLVGVIL
874	Q05707	1326	LVGVILDNGG K TLTYFNVDQS
875	Q05707	1351	TVTFEGPEIR K IFYGSFHKLH
876	Q05707	1359	IRKIFYGSF H K LHIVVSETLV
877	Q05707	1370	LHIVVSETLV K VVIDCKQVGE
878	Q05707	1376	ETLVKVVIDC K QVGEKAMNAS
879	Q05707	1381	VVIDCKQVGE K AMNASANITS
880	Q05707	1399	ITSDGVEVL G KMVRSRGPGGN
881	Q05707	1431	VCSTSWANTD K CCELPGLRDD
882	Q05707	1558	GSPGQRGLP G KDGSSGPPGPP
883	Q15848	134	TIPNMPIRFT K IFYNQQNHYD
884	Q15848	149	QQNHYDGST G KFHCNIPGLYY
885	Q15848	169	YFAYHITVYM K DVKVSLFKKD
886	Q15848	172	YHITVYMKDV K VSLFKKDKAM
887	Q15848	177	YMKDVKVSLF K KDKAMLFTYD
888	Q15848	178	MKDVKVSLF K KDKAMLFTYDQ
889	Q15848	180	DVKVSLFKKD K AMLFTYDQYQ
890	Q3Y5Z3	129	TVPNVPIRFT K IFYNQQNHYD
891	Q3Y5Z3	144	QQNHYDGT T G KFLCNIPGLYY
892	Q3Y5Z3	164	YFSYHITVYL K DVKVSLYKND
893	Q3Y5Z3	167	YHITVYLKDV K VSLYKNDKAL
894	Q3Y5Z3	172	YLKDVKVSLY K NDKALLFTHD
895	Q3Y5Z3	175	DVKVSLYKND K ALLFTHDQFQ
896	Q3Y5Z3	187	LLFTHDQFQD K NVDQASGSVL
897	Q3Y5Z3	202	ASGSVLLYLE K GDQVWLQVYE
898	Q4ZJN1	49	LPGRDGRDGA K GDKGDAGEPG
899	Q4ZJN1	52	RDGRDGA K GDAGEPGHPG
900	Q4ZJN1	66	GEPGHPGGP G KDGIRGEKGEP

901	Q4ZJN1	85	EPGADGRVEAKGIKGDPGSRG
902	Q4ZJN1	88	ADGRVEAKGIKGDPGSRGSPG
903	Q4ZJN1	99	GDPGSRGSPGKHGPKGSIGPT
904	Q4ZJN1	103	SRGSPGKHGPKGSIGPTGEQG
905	Q4ZJN1	124	LPGETGPQGGQKGDKEVGP TG
906	Q4ZJN1	148	LMGSTGPLGPKGLPGPMGP IG
907	Q4ZJN1	159	GLPGPMGP IGKPGPRGEAGPM
908	Q4ZJN1	184	EPGVRGMRGWKGD RGEKGKVG
909	Q4ZJN1	190	MRGWKGD RGEKGKVGEAPLVP
910	Q4ZJN1	192	GWKGD RGEKGKVGEAPLV PK S
911	Q4ZJN1	201	GKVGEAPLV PKSAFTVGLTVI
912	Q4ZJN1	213	AFTVGLTVISKFPPPD APIKF
913	Q4ZJN1	222	SKFPPPD APIKF D KILYNELN
914	Q4ZJN1	225	PPPD APIKF D KILYNELNHYN
915	Q4ZJN1	240	ELNHYNVATGKFTCHVAGVYY
916	Q4ZJN1	268	FSRNVQVSLVKNGVKVLHTKD
917	Q4ZJN1	272	VQVSLVKNGVKVLHTKDSYMS
918	Q4ZJN1	277	VKNGVKVLHTKDSYMSSEDQA
919	Q4ZJN1	296	QASGGIVQELKLGDEVWMQVT
920	Q60994	36	ELAPALVPPPKGTCAGWMAGI
921	Q60994	152	QQNHYDGSTGKFYCNIPGLYY
922	Q60994	172	YFSYHITVYMKDVKVSLFKKD
923	Q60994	175	YHITVYMKDVKVSLFKKDKAV
924	Q60994	180	YMKDVKVSLFKKDKAVLFTYD
925	Q60994	181	MKDVKVSLFKKDKAVLFTYDQ
926	Q60994	183	DVKVSLFKKDKAVLFTYDQYQ
927	Q60994	195	VLFTYDQYQEKNV DQASGSVL
928	Q9SE35	2	LPFIATLKMMKLTAIFPLLFT
929	Q9SE35	36	AAANLSTEDSKSAQLISADSS
930	Q9SE35	110	EEEELRILSSKSGSYYSYGT
931	Q9SE35	111	EEELRILSSKSGSYYSYGTK
932	Q9SE35	121	KSGSYYSYGT KSGSYSGYST
933	Q9SE35	122	SGSYYSYGTK KSGSYSGYSTK
934	Q9SE35	132	KSGSYSGYST KKSASRRILSS
935	Q9SE35	133	SGSYSGYSTKKSASRRILSSK
936	Q9SE35	143	KSASRRILSSKSGSYSGYST
937	Q9SE35	144	SASRRILSSKSGSYSGYSTK
938	Q9SE35	154	KSGSYSGYST KSGSRRILSS
939	Q9SE35	155	SGSYSGYSTK KSGSRRILSSK
940	Q9SE35	165	KSGSRRILSSKSGSYSGSKG
941	Q9SE35	166	SGSRRILSSKSGSYSGSKGS
942	Q9SE35	174	SKKSGSYSGSKGSKRRILSSK
943	Q9SE35	177	SGSYSGSKGSKRRILSSKKSG
944	Q9SE35	184	KGSKRRILSSKSGSYSGSKG

945	Q9SE35	185	GSKRRILSSK K SGSYSGSKGS
946	Q9SE35	193	SKKSGSYSGS K GSKRRNLSSK
947	Q9SE35	196	SGSYSGSKGS K RRNLSSKKSG
948	Q9SE35	203	KGSKRRNLSS K KSGSYSGSKG
949	Q9SE35	204	GSKRRNLSSK K SGSYSGSKGS
950	Q9SE35	250	SKKSGSYSGS K GSKRRILSGG
951	Q9SE35	253	SGSYSGSKGS K RRILSGGLRG
952	Q9Y383	29	RDGDTTRQRI K FSDDRVCKSH
953	Q9Y383	37	RIKFSDDRV C KSHLLNCCPHD
954	Q9Y383	61	GTRMDLGECL K VHDLALRADY
955	Q9Y383	76	ALRADYEIAS K EQDFFFELDA
956	Q9Y383	105	ADCDRTEVA K KRLAETQEEI
957	Q9Y383	106	DCDRTEVA K KRLAETQEEIS
958	Q9Y383	122	QEEISAEVAA K AERVHELNEE
959	Q9Y383	135	RVHELNEEIG K LLAKVEQLGA
960	Q9Y383	139	LNEEIGKLLA K VEQLGAEGNV
961	Q9Y383	154	GAEGNVEES Q KVMDEVEKARA
962	Q9Y383	161	ESQKVMDEVE K KARAKKREAEE
963	Q9Y383	165	VMDEVEKARA K KREAEEVYRN
964	Q9Y383	166	MDEVEKARAK K KREAEEVYRNS
965	Q9Y383	186	SMPASSFQQ Q KLRVCEVCSAY
966	Q9Y383	213	DRRLADHFG G KLHLGFIEIRE
967	Q9Y383	224	LHLGFIEIRE K LEELKRVVAE
968	Q9Y383	229	IEIREKLEEL K RVVAEKQEKR
969	Q9Y383	235	LEELKRVVAE K QEKRNQERLK
970	Q9Y383	238	LKRVVAEK Q EKRNQERLKRRE
971	Q9Y383	245	KQEKRNQERL K RREEREREER
972	Q9Y383	257	REEREREERE K LRRSRSHSKN
973	Q9Y383	289	HRSRSMSRE R KRRTRSKSREK
974	Q9Y383	295	SRERKRRT R SKSREKRRHRS
975	Q9Y383	299	KRRTRSKS R EKRHRHRSRSSS
976	Q9Y383	332	SSRDRSRERS K RRSSKERFRD
977	Q9Y383	337	SRERSKRRSS K ERFRDQDLAS
978	Q9Y383	365	RDRSPRDRD R KDKKRSYESAN
979	Q9Y383	367	RSPRDRDR K DKKRSYESANGR
980	Q9Y383	368	SPRDRDR K DKKRSYESANGRS