Supplemental Figure 1. Clustal W alignment of the protein sequence for 100 homologues of the SC PEP protein. Numbers listed in a column on the left correspond to the NCBI database accession number. The name of each protein corresponding to a number is given at the end of the multiple sequence alignment. The SC PEP protein sequence in the first listed (62738741). Perfectly conserved residues are colored in red, strongly similar residues are colored in green, and weakly similar residues are colored in blue. Statistics and Clustal W parameters used are listed at the end of the alignment. Residues mutated in this experiment are highlighted in pink at the beginning of each page. A full list of mutated residues is included at the end of this supplementary file. Residues in that list are specified by SC PEP numbering as well as the ClustalW alignment numbering. With the exception of K407S, I454L, V455I, P473S, R474K, I689V, and I880V, the wild-type residues were mutated to residues shown in the consensus sequence.

```
                     10        20        30        40        50        60        70        80        90
                      |         |         |         |         |         |         |         |         |
62738741     -------------------------------------------------------MKNRLWLAMAAPLALATPVAFAQTPP-----TLAK
87200538     -------------------------------------------------------------MPSIRPLLAASALACLAMS--------MT
103487461    --------------------------------------------------------MPAKRLSAPLALVALALTPTAAQAA-----AAAA
87200537     --------------------------------------------------------------MSEEAALPAASVATAGQP-------AAA
94496227     --------------------------------------------------------MRGFALILLSLWLAAAPAMLRAHPD------DAA
35211243     ------------------------------------------------------------MNALFGLCAVALAISLDASCA-----LAQN
32448009     ----------------------------------MVTPRSTMYDEIFNMTVSLRNDRLRKTLAAFSASVIAGLCISSLFHP-----AAKA
75906687     ------------------------------------------------------------------------------------------
35211187     ------------------------------------------------------------------------------------------
67921530     ------------------------------------------------------------------------------------------
67930945     -------------------------------------------------------------MKLNPLILITGCVFTMPGQ---------N
17131625     ------------------------------------------------------------------------------------------
86609965     ------------------------------------------------------------------------------------------
83859556     --------------------------------------------------------MRAALFAFASVLALTACSQPAAQTD-----APAE
28808634     ---------------------MAFTTRLMKLDSFCRYIVEFRCIAQKDNRKIKRNFMKKNTLAFATALALGSVLSVAHAN----------
75854458     --------------------------------------------------------MKKNTLAFATALALGSVLSVAHAN----------
86607173     ------------------------------------------------------------------------------------------
23128197     ------------------------------------------------------------------------------------------
91226380     --------------------------------------------------------MKKNTLALTTALALGSVLSVAHAN----------
88711708     ---------------------------------------------------------MKKVFIIPI--AVVLLTACNTNTK---------
110167745    ------------------------------------------------------------------------------------------
109899092    ---------------------------------------------------------MKLTLIAVS--ISAMLMGCSNANH--SANQTPA
86144168     ---------------------------------------------------------MKKLAYVG---ALTILFSACETNE---------
4973227      --------------------------------------------------------------------------MSGKAR----------
6048357      --------------------------------------------------------------------------MSGKAR----------
76875127     --------------------------------------------------------MFKQTLACA---VLVALSGCSEP----ANTATNK
```

```
216201       ----------------------------------------------------------------------------MSGKAR----------
88804657     ---------------------------------------------------------MKHTTAILC--LGLLLTACQNQTE---------
89517967     ----------------------------------------------------------------------MALVSCEDKK----------
90588230     ---------------------------------------------------------MKKTFLLMA--LTTAGISFAQIK----------
71143549     --------------------------------------------------------MKKQSLFLTSILVTSVLAGCTGN----SVEQTSD
83855898     ---------------------------------------------------------MKQLILVTI--TAATIFSCKTET----------
85710915     ---------------------------------------------------------MKYVLSAV---AVSVLVACSPAQNDSSQTAETQ
78368715     -------------------------------------------------------MALKTILSKQTLLLAGLGLTLMGCQY-----HAN-
56178266     ---------------------------------------------------------MKYLLGAV---AVSLLTACS-------QNAETE
91215481     ---------------------------------------------------------MKQIVLPLA--ALCLLASCNEGEK---------
78688320     -------------------------------------------------------MALRFRLAQQGLLVATLGMALAACQSGNTPEEAA-
24374294     -------------------------------------------------------MALRFRLAQQGLLVATLGMALAACQSGNTPEEAAA
456523       ---------------------------------------------------------MKYNKLSVAVAAFAFAAVSAQNS----------
77816782     -------------------------------------------------------MAVRFHLAKQGLLVATLGLALGACQSGNTPEQHSV
68547790     ---------------------------------------------------------MTSRISTLSLLTLGFALGLSGCHT----KDAT-
82744257     -------------------------------------------------------MAVRFHLAKQGLLVATLGLALGACQSGNTPEQHSV
86131249     ---------------------------------------------------------MKKLVLFAL--ATTAIISCNQDAE---------
88796133     ---------------------------------------------------------MKKKLIYAAVVAASAGVTFACSTS-----YSNN
42526704     ------------------------------------------------------------------------------------------
27367204     ------------------------------------------------------------------------------------------
37676919     ------------------------------------------------------------------------------------------
75855979     ------------------------------------------------------------------------------------------
28809898     ------------------------------------------------------------------------------------------
6460324      --------------------------------------------------------------------------MSKPYPQ-----SPRG
91223249     ------------------------------------------------------------------------------------------
56180108     ---------------------------------------------------------MTR---YSAFILCSTVLILSGCSS-----FEQE
55769741     ----------------------------------------------------------------MGSVAGDAARL---------------
47507507     ----------------------------PLSHLYLTAKRVLPVTVSARLLCHAVVFPRPSLCSASLSIPGTMPQQQK-------------
28502989     ------------------------------SHLYLTAKRVLPVTVSARLLCHAGVFPRPSLCSASLSISGTMPQQQK-------------
86144914     ---------------------------------------------------------------------------MSYLK----------
84386759     ---------------------------------------------------------------------------MSYLK----------
63100723     ---------------------------------------------------------------------------MA-------------
38649021     -----------------------------------------------------------------------MPQQQK-------------
59711729     ------------------------------------------------------------------------------------------
6755152      --------------------------------------------------------------------------MLS-------------
74215349     --------------------------------------------------------------------------MLS-------------
```

```
50514023      --------------------------------------------------------------------------MLS-------------
26345256      --------------------------------------------------------------------------MLS-------------
157879458     --------------------------------------------------------------------------MLS-------------
51592147      --------------------------------------------------------------------------MLS-------------
10835490      --------------------------------------------------------------------------MLS-------------
50514022      --------------------------------------------------------------------------MLS-------------
37928243      --------------------------------------------------------------------------MLS-------------
27065055      --------------------------------------------------------------------------MLS-------------
5103285       --------------------------------------------------------------------------MLS-------------
13786144      --------------------------------------------------------------------------MLS-------------
5107662       --------------------------------------------------------------------------MLS-------------
79382269      MLTAFASHARSHVFAFVTVPTITRRLRINILRQSPLSSSLLLNETFSNRPNSVSRRCYCSSSAIMGSSSVFGEQL---------------
27066372      --------------------------------------------------------------------------MLS-------------
19347837      --------------------------------------SLLLNETFSNRPNSVSRRCYCSSSAIMGSSSVFGEQL---------------
109072195     --------------------------------------------------------------------------MLS-------------
11691900      --------------------------------------------------------------------------MLS-------------
41349456      --------------------------------------------------------------------------MLS-------------
558596        --------------------------------------------------------------------------MLS-------------
53127346      --------------------------------------------------------------------------MQA-------------
904214        --------------------------------------------------------------------------MLS-------------
15217929      ----------------------------------------------------------------MGSLPAFEERL---------------
9558588       ----------------------------------------------------------------MGSLPAFEERL---------------
38345850      ----------------------------------------------------------------MGSLPAADKPLPA-------------
91082735      ---------------------------------------------------------------------------MS-------------
54650756      ------------------------------MHTIFRGPLLYRRICQIVVPKTLPRNYIHRELPSNLTRAMPSTLSAK-------------
73973662      ------------------------------------------------------------------------------------------
73973660      ------------------------------------------------------------------------------------------
47222200      ---------------------------------------------------------------------------MA-------------
73973664      ------------------------------------------------------------------------------------------
55235507      ----------------------------------------------------------------------------Q-------------
73973658      ------------------------------------------------------------------------------------------
71666762      ---------------------------------------------------------------------------MR-------------
49523575      ---------------------------------------------------------------------------MR-------------
71747962      ---------------------------------------------------------------------------MR-------------
33317309      ---------------------------------------------------------------------------MR-------------
47224662      ------------------------------------------------------------------------------------------
```

```
15291259      ---------------------------------------------------------------------MPSTLSAK-------------
73973656      ------------------------------------------------------------------------------------------

Prim.cons.    MLTAFASHARSHVFAFVTVPT22222R2332HLYL2A2RLL323232N2P3S32RMMMKKTLAA2GLLAAT2LLMLSAAQSGNTPEAAAA


                     100       110       120       130       140       150       160       170       180
                       |         |         |         |         |         |         |         |         |
62738741      DQAMPSLPP---------YPASPQV-PLVEDHFG-----EKVSDPWRWLEADVRTDAKVAAWVQAQSAYTAAYLKQLPE-RAALEKRMKA
87200538      AEAAPAAMK---------YPQTERG-TVVETAFG-----EKVADPYRWLEADVRVDPKVAAWVDAQSRFTDAYLKALPE-RPAFEQRLKT
103487461     ASAPAAALA---------YPDTARG-DTVDPQFG-----VDVADPYRWLEDDVRVNPEVAAWVEAQNRVTDAYLDTLPG-RDAFRARMTE
87200537      ARLFLPRGA---------YPETRRD-GLVEQVFG-----QRVADPYRWLEADPRGDAGVAGWIARQNALSADYLAKLPG-RERLATRIRA
94496227      IAASDPALR---------YPLARRQ-DIVEDHFG-----VPVADPYRWLENDLRADPAVRDWVARENALTRRYLDALPG-REALKARLQA
35211243      --LAASAPR-------LTYPPTERV-EQFDDYHG-----TRVADPYRWLEDP--NSPKTRAWVEAQNKVTFAYLEGIPY-REPLKERLTK
32448009      DEPTASSSR-----MKLEYPVSRRD-EVVDDYHG-----REVADPYRWLEDV--ESDETAAWVEAQNEVTQSYLQSLPQ-RESMRARLEA
75906687      --MPYSEKS-------LNYPLSHKI-DQVDDYHG-----TLVADPYRWLEDP--DSETTRAWIEAQNQVTFAYLGEVST-REKIQQRLNK
35211187      ----MTSP--------FTYPPSRRD-DQVDDYHG-----TPVADPYRWLEDA--ESPGTKAWVEAQNQVTFGFLETIPE-REAIRERLTH
67921530      -----MSPN-------FSYPSSPQQ-DVTDVYHG-----VEIKDSYRWLENP--DSEETKVWITAQNKLTFDYLSTISA-REKIKQRLTK
67930945      MPRLNYPAA----------PQS----GQTDAYHE-----VQIADPYRGLENA--DAPETRKWVEDENALTQSWLGKQPD-RAPIRKQLTK
17131625      --MPYSEKS-------LNYPLSHKI-DHVDDYHG-----TLVADPYRWLEDP--DSEKTRVWIEAQNQITFAYLGEVSI-REKIQQRLNK
86609965      --------M-------LNYPPSYPDPTVVDIYHG-----QAVPDPYRWLEDL--DSEQTRAWIEAQNHLTFNYLQQIPA-RQRIRERLTQ
83859556      SEASGPAMTQI---AAPVYPETRTV-EQVDVYASAAEGEVAVSDPYRWLEQDVRVSEDVANWVEAQTEVTNAYLDQLPG-RERIAARLAE
28808634      -EAAQSSVEQV-TPKAITYPTTQKV-DVVDDYFG-----TKMSDPYRWLEDD--LSPETAEWVKAQNAVTFDYLSKIPY-REQIEERITK
75854458      -EAAQSSVEQV-TPKAITYPTTQKV-DVVDDYFG-----TKVSDPYRWLEDD--LSLETAEWVKAQNALTFDYLSKIPY-REQIEERITK
86607173      --------M-------LNYPPSHPDPTVVDFYHG-----QAVPDPYRWLEDL--DSEQTRAWIEAQNRLTFDYLQRIPA-RQRLLERLRQ
23128197      --MPSSTKP-------ITYPSSHKS-NQIDNYHG-----TLVTDSYRWLEDP--DSEETRAWIEAQNQITFGYLSEIPA-REKIKQRLTK
91226380      -EAAQSSVEQV-SPKAITYPTTQKV-DIVDDYFG-----TKVSDPYRWLEDD--LSPETAEWVKAQNAVTFDYLAKIPY-REQIEERITK
88711708      --KDP---------ITVNYPTTTKA-DSVDTYFG-----TEVKDPYRWLEDD--RSPETEAWVKEQNSTTFGYLEKIPF-RDGLKKRLEK
110167745     --MKEKKQP-------LTYPITEKT-DTVENYHG-----VEVADPYRWLEDP--NLEKTKEWVKSQNEITFNYLAEISE-GETIKKRLTK
109899092     TMKTSITSNQL-QEMNLSYPKTQKG-DVIDTYFG-----KQVADPYRWLEDD--MSSETEQWVKAENAVTQAYLAQVPS-RDKLKERLKV
86144168      --KKSPE-------MAVTYPETKKV-DTVDTYFG-----TEVPDPYRWLEDD--RSAETEAWVKAENKATFGYLDKIPF-KDEIKQKLEK
4973227       ----------------LHYPVTRQG-EQVDHYFG-----QAVADPYRWLEDD--RSPETEAWVKAQNAVTQDYLAQIPY-RAAIKEKLAA
6048357       ----------------LHYPVTRQG-AQVDHYFG-----QAVADPYRWVEDD--RSPETEAWVKAQNAVTQDYLAQIPY-RAAIKEKLAA
76875127      EAKAPMS-----QANSVNYPVTKKG-NVVDNYFG-----ETLADPYRWLEDD--MSAETANWVKAQNKVTFSYLEQIPY-RDKLKQRLEK
216201        ----------------LHYPVTRQS-EQLDHYFG-----QAVADPYRWLEDD--RSPETEAWVKAQNRVTQDYLAQIPF-RDAIKGKLAT
88804657      --TN----------MAINYPETAKV-DTVDTYFG-----TEVADPYRWLEDD--RSPETEDWVRRQNAVTFGYLDGIPY-REALRQRLEK
```

```
89517967     --QDSKK-------MTVTYPTTQKV-DHVDNYHG-----TDVQDPYRWLEDD--RSAETGEWVKSQNKVTNSYLDQIRF-RESVKNRLTE
90588230     ---Y---------------PETKKG-ETVDVYFD-----TKVSDPYRWLEDD--KSAETGAWVKAENEVTYGYLDKIPF-REELKKRMEK
71143549     KVAAAQK-----QTQKISYPVTKKG-DVVDTFFG-----TDVADPYRWLEDD--RSEETGAWVKAENKVTFDYLSQIPY-RGQLKSRLAE
83855898     --KTDRT-------IAVTYPETKKV-DTVDVYFG-----NEVKDPYRWLEDD--RSKETEDWVKAQNQATFGYLDKIPF-REDLKNRLTE
85710915     QVEKQQAG-----SMQVKYPETRKG-DTVDTYFD-----TKVKDPYRWLEDD--RSAETEAWVEAQNEVTFNHLEKIPF-RDKIEKRLTE
78368715     --ETAAAQA-------AVQPQG-----VVETIHG-----VQVADPYRYLEE---ESDATHAWVKQQQQAGHDFLAQLEN-KQAIVDRITE
56178266     QEANTQS---------MSYPETRKG-DVVDTYFG-----TEVADPYRWLEDD--RSEETENWVKAQNEVTFAHLESIPY-RETIEKRLTE
91215481     --NQ----------ANLTYPETKQV-DTVDTYFD-----TEVKDPYRWLEDD--RNAETEDWVKAQNEVTFGYLDSIPF-RNKLKERLSK
78688320     --QQVSANK-------LVYPETRKD-NLVETIHG-----VAVADPYRHLEE---NTPETESWVKEQQAFGQAYLAKIPN-KQAVVDRITE
24374294     --QQISANK-------LVYPETEKD-ALVETIHG-----VAVADPYRHLEE---NTPETESWVKEQQAFGQAYLAKIPN-KQVVVDRITE
456523       ----N----------VLKYPETKKV-SHTDTYFG-----TQVSDPYRWLEDD--RAEDTKAWVQQEVKFTQDYLAQIPF-RDQLKKQLMD
77816782     EVQQASSNK-------FVYPETDKD-DLVETIHG-----ISVADPYRHLEA---NTPETEKWVKTQQAFGQEYLAKIPN-KQAVVERITE
68547790     --VDAPVAA-------IAYPASQTV-AISEQIHG-----VEVQDPYRWLEE---SSPETEDWVKRQQAFGEEYLAAIPN-KQVVVDRITE
82744257     EVQQASSNK-------FVYPETDKD-DLVETIHG-----ISVADPYRHLEA---NTPETEKWVKTQQAFGQEYLAKIPN-KQAVVERITE
86131249     --KQSTS-------MAVTYTETKKV-DTVDTYFG-----TEVQDPYRWLEDD--RSEETMDWVKRENNTTQDYLKNIPF-RAELKDRLAT
88796133     AEPAKSEMQHTAMPTVSSYPETKTVAQQ-DNYHG-----TTVSDPYRWLEE---ESEEVNAWVEAQNELARPYLAELSS-KALYKERLTA
42526704     -------------------MQYKKS-DVSDNYFG-----TIVPDPYRWLEDD--NAPEVIAWVKEENKKTEDFLSKISF-RGELKKRLEE
27367204     ----------------MQYIKTRKD-PIVDHYFG-----HAVADPYRWLEDD--RSAETTEWVSGQNAITFDYLGQIGF-RQQLREQIAA
37676919     ----------------MQYIKTRKD-PIVDHYFG-----HAVADPYRWLEDD--RSAETTEWVSGQNAITFDYLGQIGF-RQQLREQIAA
75855979     ----------------MHYPKTRKD-SVVDTYFG-----HDIADPYRWLEND--RSEETAQWVSGQNSVTFDFLGQIPY-RQQIRDLVAN
28809898     ----------------MHYPKTRKD-SVVDTYFG-----HDIADPYRWLEDD--RSEETAQWVSGQNSVTFDFLGQIPY-RQQIRDLVAN
6460324      DHVDVYQNA-----AGQEVRVP---------------------DPYRWLEDP--DSPETRRWVEAQNAVTEDFLAALPA-RAAYRERLTA
91223249     ----------------MHYPKTRKD-SVVDTYFG-----HDIADPYRWLEDD--RSEETAQWVSGQNSVTFDFLGQIPY-RQQIRDLVAN
56180108     KEQPK-----------GPYPETKTVEQQ-DDYHG-----TRVEDPYRWLEQ---DNQEVSNWVDAQNQLSTPLLQNLPA-FADIKDRMTE
55769741     -----------------SYPPTRRDNSVVDMYHG-----VPVADPYRWLEDP--ESEDTKEFVASQVELAESVLAGCFD-RENLRREVTR
47507507     ----------------IQYPEAARDNCAVDNYHG-----VKMCDPYRWLEDP--DSEQTKAFVEAQNKLTMPLLEQCPVR-RLFKDRMTE
28502989     ----------------IQYPEAARDNCAVDNYHG-----VKMCDPYRWLEDP--DSEQTKAFVEAQNKLTMPLLEQCPVR-RLFKDRMTE
86144914     ---------------EYQYPVTNKQ-IVSDDYFG-----QIIEDPYRWLEDD--RSDETAQWVASQNEVTFDYLAQIPY-RAELRERLAK
84386759     ---------------EYQYPITNKQ-IVSDDYFD-----HIVEDPYRWLEDD--RSDETAQWVASQNEVTFDYLAQIPY-SAELRERLAK
63100723     ----------------FQYPNVYRDESVVDDYHG-----CKIPDPYSWLEDP--DSEKTQAFVNAQNQLTLPFLEQCETR-DLFKERMTE
38649021     ----------------IQYPAAARDDCAVDNYHG-----VKMCDPYRWLEDP--DSEQTKAFVEAQNKLTMPFLEKCPVR-RLFKDRMTE
59711729     ----------------MIYPFTKQD-STVDHYFS-----HAIADPYRWLEDD--RSEETEAWVNKQNDVTFDYLSHIDF-RDDIRDLIAK
6755152      ----------------FQYPDVYRDETSVQEYHG-----HKICDPYSWLEDP--DSEQTKAFVEAQNKITVPFLEQCPIR-GLYKERMTE
74215349     ----------------FQYPDVYSDETSVQEYHG-----HKICDPYSWLEDP--DSEQTKAFVEAQNKITVPFLEQCPIR-GLYKERMTE
50514023     ----------------FQYPDVYRDETAIQDYHG-----HKVCDPYAWLEDP--DSEQTKAFVEAQNKITVPFLEQCPIR-GLYKERMTE
26345256     ----------------FQYPDVYRDETSVQEYHG-----HKICDPYSWLEDP--DSEQTKAFVEAQNKITVPFLEQCPIR-GLYKERMTE
```

```
157879458    ----------------FQYPDVYRDETAIQDYHG-----HKVCDPYAWLEDP--DSEQTKAFVEAQNKITVPFLEQCPIR-GLYKERMTE
51592147     ----------------FQYPDVYRDETAIQDYHG-----HKVCDPYAWLEDP--DSEQTKAFVEAQNKITVPFLEQCPIR-GLYKERMTE
10835490     ----------------FQYPDVYRDETAIQDYHG-----HKVCDPYAWLEDP--DSEQTKAFVEAQNKITVPFLEQCPIR-GLYKERMTE
50514022     ----------------FQYPDVYRDETAIQDYHG-----HKVCDPYAWLEDP--DSEQTKAFVEAQNKITVPFLEQCPIR-GLYKERMTE
37928243     ----------------FQYPDVYRDETAIQDYHG-----HKVCDPYAWLEDP--DSEQTKAFVEAQNKITVPFLEQCPIR-GLYKERMTE
27065055     ----------------FQYPDVYRDETAIQDYHG-----HKVCDPYAWLEDP--DSEQTKAFVEAQNKITVPFLEQCPIR-GLYKERMTE
5103285      ----------------FQYPDVYRDETAVQDYHG-----HKICDPYAWLEDP--DSEQTKAFVEAQNKITVPFLEQCPIR-GLYKERMTE
13786144     ----------------FQYPDVYRDETSVQDYHG-----HKICDPYAWLEDP--DSEQTKAFVEAQNKITVPFLEQCPIR-GLYKERMTE
5107662      ----------------FQYPDVYRDETAIQDYHG-----HKVXDPYAWLEDP--DSEQTKAFVEAQNKITVPFLEQXPIR-GLYKERMTE
79382269     -----------------QYPATRRDDSVVDDYHG-----VKIGDPYRWLEDP--DAEEVKEFVQSQVKLTDSVLEKCET-KEKLRQNITK
27066372     ----------------FQYPDVYRDETAIQDYHG-----HKVCDPYAWLEDP--DSEQTKAFVEAQNKITVPFLEQCPIR-GLYKERMTE
19347837     -----------------QYPATRRDDSVVDDYHG-----VKIGDPYRWLEDP--DAEEVKEFVQSQVKLTDSVLEKCET-KEKLRQNITK
109072195    ----------------LQYPDVYRDETAVQDYHG-----HKICDPYAWLEDP--DSEQTKAFVEAQNKITVPFLEQCPIR-GLYKERMTE
11691900     ----------------FQYPDVYRDETAVQDYHG-----HKICDPYAWLEDP--DSEQTKAFVEAQNKITVPFLEQCPIR-GLYKERMTE
41349456     ----------------LQYPDVYRDETAVQDYHG-----HKICDPYAWLEDP--DSEQTKAFVEAQNKITVPFLEQCPIR-GLYKERMTE
558596       ----------------FQYPDVYRDETAVQDYHG-----HKICDPYAWLEDP--DSEQTKAFVEAQNKITVPFLEQCPIR-GLYKERMTE
53127346     ----------------FQYPEVYRDEAAVLDYHG-----HQISDPYCWLEDP--DSEQTKAFVEAQNKLTVPFLEQCPVR-GLFKERMTE
904214       ----------------FQYPDVYRDETAIQDYHG-----HKICDPYAWLEDP--DSEQTKAFVEAQNKITVPFLEQCPIR-GLYKERMTE
15217929     -----------------QYPTARRDESVVEDYHG-----VKVSDPYRWLEDP--DAEEVKEFVEKQVQLSDSVLKSCET-KEKLHDKFTK
9558588      -----------------QYPTARRDESVVEDYHG-----VKVSDPYRWLEDP--DAEEVKEFVEKQVQLSDSVLKSCET-KEKLHDKFTK
38345850     ----------------LRYPPARRDDDIVDDYHG-----VTVPDPYRWMEEL--ESEEVKGFVDAQAAVAEAVLSTCDDHRVRLRGQLTA
91082735     ----------------FKYPDARRDETVKDNYFG-----TEITDPYRWLEDP--DSEETKKYVDGQNAVTRPYLDGCSF-KESIKKKITQ
54650756     ----------------LAYPVARKDESVAEDFHG-----TQIKDVYRWLEDP--DSTETEEFVNAQNSISRPFLENGEE-WKKLNTKLTK
73973662     ---------------------------------------------------------------------------------------MTE
73973660     ---------------------------------------------------------------------------------------MTE
47222200     ----------------FKYPTARRDEAKVDDYHG-----NKICDPYAWLEDP--DGAETMAFVEEQNKLTMPFLEQCGVR-DRFRQRLTE
73973664     ---------------------------------------------------------------------------------------MTE
55235507     ----------------FVYPAARRDESVVEEFHG-----VKIADPYRWLEDP--DAEETQAYVEKQNEISKPFLDTCPE-WKKLNEKLRK
73973658     ---------------------------------------------------------------------------------------MTE
71666762     ----------------SVYPLARRS-MAAYTMHN-----MTVPEPYDYLEDP--ENPETKTFVNEQNAFFEEYFASEAELRKKIFESISN
49523575     ----------------LAYPTARRS-AASYVLHN-----VTVPEPYDYLEDP--ENNETKSFVKAQNDLFGEYMRSTEALGNKLYNRISQ
71747962     ----------------LAYPTARRS-AASYVLHN-----VTVPEPYDYLEDP--ENNETKSFVKAQNDLFGEYMRSTEALGNKLYNRISQ
33317309     ----------------SVYPLARRS-MAAYTMHN-----MTVPEPYDYLEDP--ENPETKTFVNEQNAFFEEYFASEAELRKKIFESISN
47224662     -------------------------------------------------------------------------MASAEG--TLHSDNISD
15291259     ----------------LAYPVARKDESVAEDFHG-----TQIKDVYRWLEDP--DSTETEEFVNAQNSISRPFLENGEE-WKKLNTKLTK
73973656     ---------------------------------------------------------------------------------------MTE
```

```
                                                                                                    .
Prim.cons.   AEAAASSNAQVAQPMAFQYPDTRRDEDVVDDYHGAAEGEHKVADPYRWLEDPVRDSEETKAWVEAQNKVTFPYLEQIPIRRELLKERLTE

                    190       200       210       220       230       240       250       260       270
                      |         |         |         |         |         |         |         |         |
62738741     LIDYERFGLPQR--------RGASVFYSWNSGLMNQSQLLVRPADAP--VGTKG-RVLLDPNTWAKDGATALDAWAASDDGRLLAYSVQD
87200538     LFDFERFGLPVK--------AGDLLFFRHNSGLQNQSVLYVRKADG---SGER--RVLIDPNGWAKDGATALDDWQPSPDGTKVAYSVQD
103487461    LYDYERFGLPTK--------AGARYFYTRNDGLQPQSVLYVREGLK----GEG--RVLIDPNLWARDGATALAEWEPSEDGKYLLYAVQD
87200537     LFDFERYGLPRK--------AGRSYFYTRNTGLQNQSALWVRRGLD----GEQ--RLLVDPNLWSADGSLALAQWEPSPSGRYLAIAEQE
94496227     LFAHGRYTVPRK--------AGDRYFYGYNRGLENQTPLYVREGLT----GRQ--RLLLDPNSWAQDGASALAEWTPSPDGRMLAYGVQD
35211243     IWNYEKYGIPNREG--------DRFFFTRNSGLQNQSVLYNAAKLG----DAP--RVLLDPNTLSKDGTVALTGIDITDDGRYLAYGTAA
32448009     LWDYSRTGLPAK--------RGETYFYTFNDGLQNQSVLYRTPVDVPPSQWEANREVLLDPNKLSEDGTMSLASWVPSEDGKYLAYAIAD
75906687     LWDYEKYGIPFKEG--------ENYFYFKNDGLQNQSVLYTLKSLD----SEP--RVLLDPNKLSDDGTVALSGLAISDNGKLLAYGIAT
35211187     IWNYERYGTPFREG--------GRYFLFKNTGLQNQSVLYTVPDLE----GEL--RLLLDPNALSADGTVALSGLAVSDDGKYLAYGTSA
67921530     LWDYEKYGIPFKKA--------DRYFYFKNNGLQNQSVLYTIKSLD----GDP--KILLDPNTITTDGTVALSVLSISENAQYLAYGLST
67930945     LWNYEKYGTLFK--------AGPHYFYSYNTGLQNQSVVFVADSPN----AKG--KELLDPNTHRKDGTAALSGWSVSRNGKLYGYAIAE
17131625     LWDYEKYGIPFKEG--------ENYFYFKNDGLQNQSVLYTLKSLD----SEP--RVLLDPNKLSDDGTVALSGLAISNNGKLLAYGIST
86609965     LWNYEKYSQPFKEG--------GRYFYFKNDGLQNQSVLYTQESLE----AEA--RVLLDPNTFSEDGTVALAGIAISRDGRYLAYGLSR
83859556     LWNYERYGLPTTRE----TEDGLRYFFSRNDGLQDQSVFMVQDGLE----GEA--RPLIDPNEWAADGTTALAGTVPSPDGTKLAYLIAD
28808634     LMDYEKQSQPFK--------EGRFTYFYKNDGLQNQDVLYRQLGDGDAE-------IFLDPNTFSEDGTTSLAGVSFSRDGSLVAYSISE
75854458     LMDYEKQSQPFK--------EGKFTYFYKNDGLQNQDVLYRQLGDGEAE-------IFLDPNTFSEDGTTSLAGVSFSRDGSLVAYSISE
86607173     LWNYEKYSQPFKEG--------NRYFYFKNDGLQNQSVLYTQESLE----GEA--RVLLDPNTLSEDGTVALSGIAISRDGRYLAYGLSR
23128197     LWDYEKYGIPFKEGKSLQDGSTERYFYFKNNGLQNQSVLYTLKTLD----NQP--KVLLDPNQLSEDGTIALSGLSISENGKLLAYGLST
91226380     LMDYEKQSQPFK--------EGKFTYFYKNDGLQNQDVLYRQLGDGEAE-------IFLDPNTFSEDGTTSLAGVSFSRDGSLVAYSISE
88711708     LWNYEKLGSPFK--------EGDYTYFYKNDGLQNQYVVYRQKDSKEP----E---VFLDPNTFSEDGTTSLAGLRFTKDGSKAAYLISE
110167745    IWDYEKYSVPFKEG--------DRYFYYKNDGLQNQSILYTLPTLD----AEP--KVLIDPNQFSEDGTVALSGIAISKDGKYIAYGISK
109899092    LLDYEKVGAPFK--------EGKYTYFFKNDGLQNQAVLYRQLDDGEAQ-------VFLDPNSFSDDGTTSLASIDFSKDGSLVTYLISE
86144168     IWNYEKIGAPFK--------EGDYTYFYKNDGLQNQYVIYRYKTGEDPSTA-E---VFLDPNTFAEDGTVSLGGLSFSDNGKLAAYSISE
4973227      SWNYAKEGAPFR--------EGRYHYFFKNDGLQNQNVLWRQKEGKPAE-------VFLDPNTLSPDGTTALDQLSFSRDGRILAYSLSL
6048357      SWNYAKEGAPFR--------EGRYHYFFKNDGLQNQNVLWRQKEGKPAE-------VFLDPNTLSPDGTTALDQLSFSRDGRILAYSLSL
76875127     LMNYEKIGAPFT--------EGDYTYFYKNDGLQNQYVLYRSKEGGDVE-------VFLDPNTFSDDGTTSMSGLSFSKDGSLLAYQISE
216201       SWNYAKEGAPFR--------EGRYHYFFKNDGLQNQNVLCGQLAGKPAE-------VFLDPNLLSPDGTTALDQLSFSRDGKTLAYSLSL
88804657     LWNYEKLGTPFK--------EGGKTYFYKNDGLQDQYVLYRQGETGDP----E---VFLDPNTFSEDGTTSLAGLSFTEDGSRVAYSISE
89517967     LWNYEKIGAPFI--------EGDYSYFYKNDGLQNQYVIYRFPTNGKMEDA-E---VFLDPNQFSKDGTTSLGGTSFTKDGSLFAYAISE
90588230     LWNYEKIGAPFK--------EGKFTYYYKNNGLQNQSVVYRKDQSGKE----E---VFLDPNTFSKDGTTSLGGLDFSKDGSKAAYAISE
71143549     LWNYEKVGAPFK--------KGDYSYFYKNDGLQNQYVVYRQKEGGTAE-------LFLDPNTFSEDGTTSMGQLSFSKDGSIAAYAISE
```

```
83855898      LWNYEKLGSPFK--------EGEYTYYFKNNGLQNQSVIYRYKSTESPENA-K---VFLDPNKFSEDGTTSLGGLNFSKDGSKAAYSISE
85710915      LWNYEKVSAPFK--------EGDYTYYYKNDGLQNQYVVYRKKNDSDEEE------VFLDPNTFSEDGTTSLAQLTFSEDGKLAAYSISE
78368715      LWNFEKVTAPNERG--------DNTFYYRNDGLQAQSVLY-VKNKA----GEE--KVALDPNSFSTDGTVALSGVSVSGDGKIIAYGVSK
56178266      LWNYEKISAPFK--------EGDYTYFYKNDGLQNQYVVYRQKGDEDPE-------VFLDPNTFSEDGTTSLAQLTFSDDGSLAAYSISE
91215481      VWNYEKIGAPFE--------RGKYTYFSKNDGLQNQYVFYRYKNEDSIENA-E---VFLDPNTFSDDGTTSLAGMSFTEDGETLAYSISE
78688320      LWNYEKISAPFENG--------DNQFYYRNDGLQAQSVLY-VKGKD----GVE--KPVLDPNGFSADGTVALSGVSVSNDGKILAYGVSQ
24374294      LWNYEKISAPFENG--------DNQFYYRNDGLQAQSVLY-VKGKD----GVE--KAVLDPNSFSADGTVALSGVSVSNDGKILAYGVSK
456523        IWNYEKISAPFK--------KGKYTYFSKNDGLQAQSVLYRKDAAGKTE-------VFLDPNKFSEKGTTSLASVSFNKKGTLVAYSISE
77816782      LWNYEKVSAPFEHG--------ENQFYYRNDGLQAQSVLY-VKGRD----GVE--KPILDPNKLSTNGTVALSGVSVSNDGKILAYGVSN
68547790      LWNFEKVGAPFEHG--------ENRFIFRNNGLQSQNVLY-VSEKG----GDE--RVLLDPNTLSADGTVALSGVSVSGDGKTLAYGVSK
82744257      LWNYEKVSAPFEHG--------ENQFYYRNDGLQAQSVLY-VKGRD----GVE--KPILDPNKLSTNGTVALSGVSVSNDGKILAYGVSN
86131249      LWNYEKVGSPFK--------EGDYTYFYKNDGLQNQYVIYRYKTGEDPSTA-T---VFLDPNTFAEDGTISLGGASFSEDGSILAYAISE
88796133      LWDYEKYSTPYMVN--------DKLFYTYNNGLQNQYVLYMADGLN----GKP--EVLIDPNTLSDDGTVSMASTELSPKASYLAYMLSD
42526704      IWDYEKRSGLFK--------AGNFYYFFRTEGLQNQSIMCRQSGNIKAESSPE---VFFDPNKLSADGTTALKNLAFSKDGKYMAYSVSG
27367204      SQDYEKSSQPFV--------HGEYTYFYKNDGLQNQSVLYRRKGDEAAQ-------VFLDPNGFSDDGTTSLSGVFFSKDDSLVAYTISE
37676919      SQDYEKSSQPFV--------HGEYTYFYKNDGLQNQSVLYRRKGDEAAQ-------VFLDPNGFSADGTTSLSGVFFSKDDSLVAYTISE
75855979      SQNYEKYSQPFL--------RGDYTYFYKNDGLQNQSVLYRRKGEGEAE-------VFLDPNTFSEEGTTSLGEVSFSKDYRLVAYSISE
28809898      SQNYEKYSQPFV--------RGDYTYFYKNDGLQNQSVLYRRKGEGEAE-------VFLDPNTFSEEGTTSLGEVSFSKDYRLVAYSISE
6460324       LWDYPRDGLPWE--------RGGRYFRTFNPGLLNQPVLQ--TADSPRGPWHE----LLDPNALSADGTVALMGASVSQDGTQLAYATQS
91223249      SQNYEKYSQPFV--------RGDYTYFYKNDGLQNQSVLYRRKGEGEAE-------VFLDPNTFSEEGTTSLGEVSFSKDYRLVAYSISE
56180108      LWNYEKVSTPFRYA--------DRYFYFANDGLQNQSVLYTMTDLN----SAP--DILINPNNFSDDGTVSLARINVSPDATYVAYATSD
55769741      LFDHPRHGAPFR--------RGNKYFYFHNSGLQAQSVLYVQDSLDG------EAEVLLDPNALSKDGTVALSTYSVSKDGKYIAYGLSE
47507507      LYDYPKYSCNFK--------KGKRYFYFYNSGLQNQRVLYVQDSLKD------EPRVFLDPNTFSEDGTVALQGYAFSDDGEYFAYGLSA
28502989      LYDYPKYSCNFK--------KGKRYFYFYNSGLQNQRVLYVQDSLKD------EPRVFLDPNTFSEDGTVALQGYAFSDDGEYFAYGLSA
86144914      AQDYKKSSQPFV--------RGDYTYFYKNDGLQNHSILYRQKEGLQVE-------VFLDPNTFSEDGTTSLGSVSFSKDYGLVAYSISE
84386759      AQDYKKSSQPFV--------RGEYTYFYKNDGLQNHSILYRHKEGQQEE-------IFLDPNTFSEDGTTSLGSVSFSKDYSLVAYSISE
63100723      LYDYPKYSCPFK--------RGDRYFHFYNTGLQNQSVMYVQDNLDA------EPTVFLDPNTFSEDGTVALRGYSFSEDGEFLAYGTSA
38649021      LYDYPKYSCNFK--------KGNRYFYFYNSGLQNQRVLYVQDSLKD------EPKVFLDPNTFSEDGTVALQGYAFSEDGEYFAYGLSA
59711729      GQDYQKTSQPFK--------RGEYTYFYQNDGLQNQSVLYRSKEGKDVE-------VFLDPNTFSEEGTTSLSVVSFSKDASLVAYSISE
6755152       LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLEG------EARVFLDPNTLSDDGTVALRGYAFSEDGEYFAYGLSA
74215349      LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLEG------EARVFLDPNTLSDDGTVALRGYAFSEDGEYFAYGLSA
50514023      LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLEG------EARVFLDPNILSDDGTVALRGYAFSEDGEYFAYGLSA
26345256      LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLEG------EARVFLDPNTLSDDGTVALRGYAFSEDGEYFAYGLSA
157879458     LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLEG------EARVFLDPNILSDDGTVALRGYAFSEDGEYFAYGLSA
51592147      LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLEG------EARVFLDPNILSDDGTVALRGYAFSEDGEYFAYGLSA
10835490      LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLEG------EARVFLDPNILSDDGTVALRGYAFSEDGEYFAYGLSA
```

```
50514022     LYDCPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLEG------EARVFLDPNILSDDGTVALRGYAFSEDGEYFAYGLSA
37928243     LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLEG------EARVFLDPNILSDDGTVALRGYAFSEDGEYFAYGLSA
27065055     LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLEG------EARVFLDPNILSDDGTVALRGYAFSEDGEYFAYGLSA
5103285      LYDYPKYSCNFK--------KGKRYFYFYNTGLQNQRVLYVQDSLEG------EARVCLDPNTLSDDGTVALRGYAFSEDGEYVAYGLSA
13786144     LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLEG------EARVFLDPNTLSDDGTVALRGYAFSEDGEYFAYGLSA
5107662      LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLEG------EARVFLDPNILSDDGTVALRGYAFSEDGEYFAYGLSA
79382269     LIDHPRYDSPFR--------QGDKYFYFHNTGLQAQSVLYMQDNLDA------EPEVLLDPNTLSDDGTVALNTFSVSEDAKYLAYGLSS
27066372     LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLEG------EARVFLDPNILSDDGTVALRGYAFSEDGEYFAYGLSA
19347837     LIDHPRYDSPFR--------QGDKYFYFHNTGLQAQSVLYMQDNLDA------EPEVLLDPNTLSDDGTVALNTFSVSEDAKYLAYGLSS
109072195    LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLEG------DARVFLDPNILSDDGTVALRGYAFSEDGEYFAYGLSA
11691900     LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLEG------EARVFLDPNILSDDGTVALRGYAFSEDGEYFAYGLSA
41349456     LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLEG------EARVFLDPNILSDDGTVALRGYAFSEDGEYFAYGLSA
558596       LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLEG------EARVFLDPNILSDDGTVALRGYAFSEDGEYFAYGLSA
53127346     LYDYPKYSCHFK--------KGKRYFHFYNTGLQNQRVLYVQDSLDA------DAKVFLDPNKLSDDGTVALRGYAFSEDGEYFAYGLSS
904214       LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLEG------EARVFLDPNILSDDGTVALRGYAFSEDGEYFAYGLSA
15217929     FIDYPRFDTPFK--------RGNSYFYFHNSGLQAQSVLHVQDDLES------EAEILLDPNTLSDDGTVSLNTLSISEDAKYLAYGLSS
9558588      FIDYPRFDTPFK--------RGNSYFYFHNSGLQAQSVLHVQDDLES------EAEILLDPNTLSDDGTVSLNTLSISEDAKYLAYGLSS
38345850     LFDHPRYRAPFK--------RAGSYFYLHNPGLQPHSALYVQHGLGGG----EEPDVLLDPNTFSDDATVSLAMFGVSHDGEHLAYGTSA
91082735     LWNYPKFSTPYR--------HGTKYYQYRNTGLQNQSVIYVQKDLA------SKAEIFLDPNTFSEDGTVALSGTAFSEDGQTFAYGLSS
54650756     LWNYPKYGCPMR--------YGNYYYYFMNTGLQNQSVMYQQKSLGDE----SESKVFLDPNTLSEDGTIALTQKAFSEDGKYMAYGLSE
73973662     LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLDG------EARVFLDPNLLSDDGTVALRGYAFSEDGEYFAYGLSA
73973660     LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLDG------EARVFLDPNLLSDDGTVALRGYAFSEDGEYFAYGLSA
47222200     LYDYPKYSCPYK--------RGSRYFYFHNEGLQNQDVLYVQDSLEA------PASVLFDPNKLSEDGTVALKMCRLSEECEYFAYGLSS
73973664     LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLDG------EARVFLDPNLLSDDGTVALRGYAFSEDGEYFAYGLSA
55235507     RWNYPKYSCPFK--------HGNKYFFFMNTGLQNQDVLYVQDKLD------GEPKVFLDPNTLSADGTIALVGSRFSDDGKLYAYGLSQ
73973658     LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLDG------EARVFLDPNLLSDDGTVALRGYAFSEDGEYFAYGLSA
71666762     SQDYPRTSNPSY--------INGHYYYYHNSGLQNQSVLMRATSLTD-----TAPSIFLDPNTMSSDGTTALKATAWSEDESMLAYSLSD
49523575     TFDFPRTSNPSF--------RNGRYYFYHNTGLQNQSVLKRATSLTD-----SNPTTFLDPNILSNDGTTALKATAWSEDESLFAYSASD
71747962     TFDFPRTSNPSF--------RNGRYYFYHNTGLQNQSVLKRATSLTD-----SNPTTFLDPNILSNDGTTALKATAWSEDESLFAYSASD
33317309     SQDYPRTSNPSY--------INGHYYYYHNSGLQNQSVLMRAMSLTD-----TAPSIFLDPNSMSSDGTTALKATAWSEDESMLAYSLSD
47224662     I------T---R--------RGSRYFHFYNTGLQNQSVMYVQESLDA------EPRVFLDPNTFSEDGTVALRGYAFSEDGEYLAYGTSA
15291259     LWNYPKYGCPMR--------YGNYYYYFMNTGLQNQSVMYQQKSLGDE----SESKVFLDPNTLSEDGTIALTQKAFSEDGKYMAYGLSE
73973656     LYDYPKYSCHFK--------KGKRYFYFYNTGLQNQRVLYVQDSLDG------EARVFLDPNLLSDDGTVALRGYAFSEDGEYFAYGLSA
                                      :   . **  :  .                   .:**    ..: ::     . .
Prim.cons.   LWDYEKYSCPFKEGKSLQ22KGKRYFYFKNDGLQNQSVLYVQDSLEGAE3GEEEARVFLDPNTLSDDGTVALRGYAFSEDGEYLAYGLSA
                    280       290       300       310       320       330       340       350       360
```

```
                         |         |         |         |         |         |         |         |         |
62738741       GGSDWRTVKFVGVADGKPLADELKWVKFSGLAWLG-NDALLYSRFAEPKEG-------QAFQALNYNQTVWLHRLGTPQSADQPVFA--T
87200538       GGSDWRTLKVIDVASGQVLSDTVEHVKFSHIAWAG-NEAVVYSRFPAPKAG-------EAFQAVSSNQSVWLHKLGTPQSEDRLLHA--T
103487461      GGTDWRIVRVKDVATGQDLPDEVRWVKFSALDWAKDGSGFYYSRFPEPKEG-------EAFQSLNENHAVYFHRLGTPQSADVLIHA--T
87200537       AGSDWRTLRVVEVSSGRVLDERVDWANDTEIAWVG-DEGFLYSRFPAPGQG-------EDARAPRFGKAVWFHRVGTAQDRDEQVFA--T
94496227       AGSDWRTLRLIDVDSGRTLDDAVQWVKFSQTAWDGRSEGFFYSRFAAPGPG-------ETFRSTNLGQSLYYHRIGTSQGQDQLIYA--T
35211243       SGSDWQQWHVREVATGRDLPDVIKWVKFSEASWLKDGKGFFYSRYDEPNAA-------TQFQDVNYFQKLYFHRLSTPQAQDVLVYE--R
32448009       GGSDWRTWRVREVATGKDTDDLIQWSKFSGIAWTPDGDGFYYSRYAEPAEG-------EELTGTNDNQMMYLHELGTPQSEDKLVMK--R
75906687       SGSDWQEWKVIDVETGADFPDHLNWVKFSGASWTNDNQGFFYSRYDEPNEK-------TKLEDVNYYQKLYYHQLGTPQSEDILIYQ--R
35211187       SGSDWQQWRVLDIETGEDLPDMVQWVKFSGASWLKDGKGFFYSRYDEPNAA-------GQFQDVNYFQKLYFHRLGTPQDQDVLVYE--R
67921530       SGSDWVEWKVREIETGKDLSDHLKWTKFTGTSWTKDNQGFFYSRYNQPKEK-------TKLEDINHYQKLYYHKLGTSQDEDVLIYE--R
67930945       AGSDWSDCRIRSVADGKDLPDVIRWTKFSNMAWATDDGGFYYSRYPEPPAG-------ELLTAAALNQKVYFHRLGDPQTADRLVYE--R
17131625       SGSDWQEWKVVDVETGADFPDHLNWIKFSGASWTNDNQGFFYSRYDEPKEK-------TKLEDVNYYQKLYYHRLGTPQSEDVLIYQ--R
86609965       SGSDWQEWKVRDIETGEDLPDHLRWVKFSGASWTLDGQGFFYSRYDEPAPG-------SEYESVNYFQKLYYHRLGTSQSEDVLVYH--R
83859556       GGSDWRSIRVMDVETGETLSDEIEWVKFSPLSWAKDGSGFYYSRYPEPEEG-------EAFTALNLNQAIYFHELGTDQSEDVQIMA--D
28808634       GGSDWRKVIVLDTETKKPVGETLVDIKFSGISWLG-NQGFYYSSYDKP-KG-------SELSAKTDQHKLYFHKLGTKQSEDQLIFGG-F
75854458       GGSDWRKVIVLDTETKKPVGETLVDIKFSGISWLG-NQGFYYSSYDKP-QG-------SELSAKTDQHKLYFHKLGTKQSEDKLIFGG-F
86607173       SGSDWQEWKVRDIETGEDLPDHLRWIKFSGASWTLDGQGFFYSRYDEPAPG-------SEYESANYFQKLYYHRLGTPQSEDLLVYH--R
23128197       SGSDWEEWKVRDVETGEDLQDHLKWIKFSGASWTHNNQGFFYSRYDEPNEK-------TQLEDVNYYQKLYYHQLGKPQSEDVLIYH--R
91226380       GGSDWRKVIVLDTETKKPVGETLVDIKFSGISWLG-NQGFYYSSYDKP-KG-------SELSAKTDQHKLYFHKLGTKQSEDQLIFGG-V
88711708       GGSDWRKAIVIDAKSKEIVEDTLVDIKFSGVSWKG-NEGFYYSSYDKP-KG-------SELSAKTDQHKVYYHKLGTPQKEDKVIFGA-K
110167745      SGSDWQEWRIKNIDTGEHFPDVLQWIKFYIPTWKNDNQGLFYSRYEQPKE--------GKLKDTNYLHKVYYHSLGTSQDNDVLIYE--K
109899092      GGSDWRKAITIDVETMQPVSETLKDIKFSGISWLG-NEGFYYSSYDKP-EG-------SELSAKTDQHKLYFHRLNDKQTQDTLVFGG-T
86144168       GGSDWRKVLVMDAETKEITEDTLVDIKFSGISWRE-NDGFYYSSYDKP-DG-------SELSAKTDQHKVYYHKLNTAQKEDELIFGG-T
4973227        AGSDWREIHLMDVESKQPLETPLKDVKFSGISWLG-NEGFFYSSYDKP-DG-------SELSARTDQHKVYFHRLGTAQEDDRLVFGA-I
6048357        AGSDWREIHLMDVESKQPLETPLKDVKFSGISWLG-NEGFFYSSYDKP-DG-------SELSARTDQHKVYFHRLGTAQEDDRLVFGA-I
76875127       GGSDWRKIIIIDTQTKEQVEQALVDVKFSGIDWLG-NEGFYYSSYDKP-EG-------SELSAKTDQHKVYYHTLGSDQATDKLVFGA-T
216201         AGSDWREIHLMDVESKQPLETPLRDVKFSGISWLG-NEGFFYSSYDKP-DG-------SELSARTDQHKLYFHRLGTAQEEDRLVFGA-I
88804657       GGSDWRKVIVMDAASREIIEDTLVDIKFSGVSWRG-NEGFYYSSYDKP-EG-------SELSARTDQHKVYFHQLGQPQAGDQLIFGG-T
89517967       GGSDWRKIIVLNTNDLTQVGDTIKDVKFSGMSWKG-NDGFFYSSYDKP-DG-------SELSAMTDQHKLYYHKLGTPQNEDQVIFGG-T
90588230       GGSDWRKVIIIDALSKKVVEDTLVDVKFSGISWLG-NEGFYYSSYDKP-KG-------SELSAKTDQHKLYFHKLGTSQKEDKVIFG---
71143549       GGSDWRKIIIIDVETKKVLETPLIDVKFSDISWFK-NEGFYYSSYDKP-EG-------SELSAKTDQHKLYYHVLGQPQKEDKVIFGN-T
83855898       GGSDWRKVIVVDAETLERVEDTLQDIKFSGVSWNV-NEGFYYSSYDKP-KG-------SELSAKTDQHKLYYHKLGTSQKEDKLIFGG-T
85710915       GGSDWRKIIVIDAETKERLEDPLVDVKFSGISWKG-NEGFYYSSYDKP-EG-------SELSAKTDQHKLYYHELGTPQSDDELVFGG-T
78368715       SGSDWQQWQFVDVATGAKLGDELNWIKFSSAEWSKDNKGVYYSRYDAPEGG-------NLLADVNFNQKVYYHQLGTDQSQDKLIYE--R
56178266       GGSDWRKIRVIDAETKEQLEEPLVDVKFSGISWVG-NEGFYYSSYDKP-EG-------SELSAKTDQHKLYYHELGTDQSDDKLVFGG-T
```

```
91215481     GGSDWRKIIVLNTETMDRKEDTLVDVKFSGVSWKG-NEGFYYSSYDKP-EG-------SELSAKTDQHKLYFHTMGTPQSGDEVIFGA-T
78688320     SGSDWQQWQFIDIATGKKLKDELNWIKFSSAVWDKENQGVFYARYDAPAGG-------DLLADVNFNQKVYYHKLGTDQRQDLLIYE--R
24374294     SGSDWQQWQFIDIATGKKLNDELSWIKFSSAVWDKENQGVFYARYDAPAGG-------DLLADVNFNQKVYYHKLGTDQRQDLLIYE--R
456523       GGSDWNKIIILDAETKKQLDETLLDVKFSGISWLG-DEGFFYSSYDKPKEG-------SVLSGMTDKHKVYFHKLGTKQSQDELIIGG--
77816782     SGSDWQQWQFINVATGKKLSDELKWIKFSNAVWDKENQGVFYARYDAPAGG-------DLLADVNFNQKVYYHTLGTDQRQDLLIYE--R
68547790     SGSDWQEWAFIDVATGRKLADSLNWIKFSRAEWAHDNSGVYYSRYDAPAGG-------DILVDVNYNQKVYFHRIGDEQAKDTLVYE--R
82744257     SGSDWQQWQFINVATGKKLSDELKWIKFSNAVWDKENQGVFYARYDAPAGG-------DLLADVNFNQKVYYHTLGTDQRQDLLIYE--R
86131249     GGSDWRKILVMDTEKKEIVEDTLVDVKFSGMSWYK-NEGFYYSSYDKP-KG-------SELSAKTDQHKLYYHKLGTAQSTDKLIFGG-T
88796133     GGTDWKTIHIRETSNKSDLTDVIKGVKFSNVAWLPDESGFFYSRYPENDAG-------EFDDSQTVS--IYFHAIGTAQNEDTKVFS--F
42526704     SGSDWEEIFVFDAEKKADTGEHIHWVKFSNIAWY--KDGFFYSSYDTPDKG-------KSLTEKNEFQKLKYHKLGTKESDDLLIFED--
27367204     GGSDWRKIFVLNAETGEQIEPEIVDAKFTGVSWLG-NQGFYYSRYDKP-QG-------SELSARTEQHKLYYHQLGTPQSDDVLVFGE-A
37676919     GGSDWRKIFVIDAETGEQIEPEIVDAKFTGVSWLG-NQGFYYSRYDKP-QG-------SELSARTEQHKLYYHQLGTPQSDDVLVFGE-A
75855979     GGSDWRKIFVIDAETKEQLESEIVDAKFTSISWLG-SKGFYYSSYDKP-QG-------SELSARTEHHKLYYHELGTPQSEDKVIFGE-L
28809898     GGSDWRKIFVIDTETKEQLEPEIVDAKFTSISWLG-SKGFYYSSYDKP-QG-------SELSARTEHHKLYYHELGTPQSEDKVIFGE-L
6460324      GGSDWLTWQVRDVASGEDTGEPLRWSKFSGAAWLPDGSGFFYSAYDAPGEG-------EALTGANKNQRLMFHRLGTPQDADELVLE--R
91223249     GGSDWRKIFVIDTETKEQLEPEIVDAKFTSISWLG-SKGFYYSSYDKP-QG-------SELSARTEHHKLYYHELGTPQSEDKVIFGE-L
56180108     GGSDWTEIKVRNVESGKDTRDVIEGVKFTEVAWLPNERGFYYSRYPENKDG-------SFDDTKPVS--VYFHKLGTEQSDDELIFA--F
55769741     SGSDWVTIRVMNIADKQTLSDKLSWVKFSSISWTHDGKGFFYGRYPAPREVE---LDAGTETNINLNHEIYYHVVGSDQSEDILCWKDP-
47507507     SGSDWVTIKFMKVANQEELPDVLERVKFSCMSWTHDGVGMFYNCYPKQVGK-----SDGTETSANLNQKLFYHVLGTSQSDDVLCAEFP-
28502989     SGSDWVTIKFMKVANQEELPDVLERVKFSCMSWTHDGVGMFYNCYPKQVGK-----SDGTETSANLNQKLFYHVLGTSQSDDVLCAEFP-
86144914     GGSDWRKIFVIDTETKQQLEAEITDAKFTGISWLG-NRGFYYSSYDKP-DG-------SQLSARTEQHKLYFHELGTEQANDKVIFGE-N
84386759     GGSDWRKIFVIDTETKQQLEAEITDAKFTGISWLG-NRGFYYSSYDKP-DG-------SQLSARTEQHKLYFHELGTEQASDKVVFGE-N
63100723     SGSDWVEIRFLRVDGAVLLEDRLERVKFTCMSWTHDGKGLFYNSYPEQEGK-----SDGTETSTNLHQKLYYHVLGTPQSQDVLCAEFP-
38649021     SGSDWVTIKFMKVANQEELPDVLERVKFSCMSWTHDGVGIFYNCYPQQKGK-----SDGTETSTNLNQKLFYHVLGTSQSDDVLCAEFP-
59711729     AGSDWRKIIVLDAETKEQIEAPIVDAKFTGISWLG-SKGFYYSSYDKP-EG-------SELSARTEQHKLYFHEIGKPQADDVLVFGG-S
6755152      SGSDWVTIKFMKVDGAKELPDVLERVKFTCMAWTHDGKGMFYNSYPQQDGK-----SDGTETSTNLHQKLCYHVLGTDQSEDILCAEFP-
74215349     SGSDWVTIKFMKVDGAKELPDVLERVKFTCMAWTHDGKGMFYNSYPQQDGK-----SDGTETSTNLHQKLCYHVLGTDQSEDILCAEFP-
50514023     SGSDWVTIKFMKVDGAKELPDVLERVKFSCMAWTHDGKGMFYNAYPQQDGK-----SDGTETSTNLHQKLYYHVLGTDQSEDILCAEFP-
26345256     SGSDWVTIKFMKVDGAKELPDVLERVKFTCMAWTHDGKGMFYNSYPQQDGK-----SDGTETSTNLHQKLCYHVLGTDQSEDILCAEFP-
157879458    SGSDWVTIKFMKVDGAKELPDVLERVKFSCMAWTHDGKGMFYNAYPQQDGK-----SDGTETSTNLHQKLYYHVLGTDQSEDILCVEFP-
51592147     SGSDWVTIKFMKVDGAKELPDVLERVKFSCMAWTHDGKGMFYNAYPQQDGK-----SDGTETSTNLHQKLYYHVLGTDQSEDILCAEFP-
10835490     SGSDWVTIKFMKVDGAKELPDVLERVKFSCMAWTHDGKGMFYNAYPQQDGK-----SDGTETSTNLHQKLYYHVLGTDQSEDILCAEFP-
50514022     SGSDWVTIKFMKVDGAKELPDVLERVKFSCMAWTHDGKGMFYNAYPQQDGK-----SDGTETSTNLHQKLYYHVLGTDQSEDILCAEFP-
37928243     SGSDWVTIKFMKVDGAKELPDVLERVKFSCMAWTHDGKGMFYNAYPQQDGK-----SDGTETSTNLHQKLYYHVLGTDQSEDILCAEFP-
27065055     SGSDWVTIKFMKVDGAKELPDVLERVKFSCMAWTHDGKGMFYNAYPQQDGK-----SDGTETSTNLHQKLYYHVLGTDQSEDILCAEFP-
5103285      SGSDWVTIKFMKVDGAKELADVLERVKFSCMAWTHDGKGMFYNAYPQQDGK-----SDGTETSTNLHQKLCYHVLGTDQSEDILCAEFP-
```

```
13786144        SGSDWVTIKFMKVDGAKELPDVLERVKFTCMAWTHDGKGMFYNSYPQQDGK-----SDGTETSTNLHQKLCYHVLGTDQSEDVLCAEFP-
5107662         SGSDWVTIKFMKVDGAKELPDVLERVKFSXMAWTHDGKGMFYNAYPQQDGK-----SDGTETSTNLHQKLYYHVLGTDQSEDILCAEFP-
79382269        SGSDWVTIKLMKIEDKKVEPDTLSWVKFTGITWTHDSKGFFYGRYPAPKEGED--IDAGTETNSNLYHELYYHFIGTDQSQDILCWRDN-
27066372        SGSDWVTIKFMKVDGAKELPDVLERVKFSCMAWTHDGKGMFYNAYPQQDGK-----SDGTETSTNLHQKLYYHVLGTDQSEDILCAEFP-
19347837        SGSDWVTIKLMKIEDKKVEPDTLSWVKFTGITWTHDSKGFFYGRYPAPKEGED--IDAGTETNSNLYHELYYHFIGTDQSQDILCWRDN-
109072195       SGSDWVTIKFMKVDGAKELPDVLERVKFSCMAWTHDGKGMFYNSYPQQDGK-----SDGTETSTNLHQKLYYHVLGTDQSEDILCAEFP-
11691900        SGSDWVTIKFMKVDGAKELPDVLERVKFSCMAWTHDGKGMFYNSYPQQDGK-----SDGTETSTNLHQKLYYHVLGTDQSEDILCAEFP-
41349456        SGSDWVTIKFMKVDGAKELPDVLERVKFSCMAWTHDGKGMFYNSYPQQDGK-----SDGTETSTNLHQKLYYHVLGTDQSEDILCAEFP-
558596          SGSDWVTIKFMKVDGAKELPDVLERVKFSCMAWTHDGKGMFYNSYPQQDGK-----SDGTETSTNLHQKLYYHVLGTDQSEDILCAEFP-
53127346        SGSDWVTIKFMKVEGAEELPDTLERVKFSCMAWTHDGKGMFYNCYPKQDGK-----SDGTETSTNLHQKLHYHVLGTNQSEDILCAEFP-
904214          SGSDWVTIKFMKVDGAKELPDVLERVKFSCMAWTHDGKGMFYNSYPQQDGK-----SDGTETSTNLHQKLYYHVLGTDQSEDILCAEFP-
15217929        SGSDWVTIKVMKIEDKKVEPDSLSWVKFSGITWTHDGKGFFYSRYPAPREGEK--IDAGTETNSNLYHELYYHFLGTDQSEDVLCWRDQ-
9558588         SGSDWVTIKVMKIEDKKVEPDSLSWVKFSGITWTHDGKGFFYSRYPAPREGEK--IDAGTETNSNLYHELYYHFLGTDQSEDVLCWRDQ-
38345850        SGSDWVTIRVMRVRDRRHLHDEICWVKFSAIAWTRDGKGFFYSRFPAPKNDGAP-LGAGIKTSVNLNHEVYYHFLGTDQSEDLLCWEDP-
91082735        SGSDWLEIKFKDVETGKDYKEILKKVKFSPMTWMHDNKGFFYGAYLDQTGK-----ADGSETKTNENQKLYYHELGTDQSQDVVVVEF--
54650756        SGSDWIKILIRDAETGKDLSEVLEKVKFSEISWTKDNKGFFYGRYPDQDGK-----TDGSETKQNENQKLYYHRVGESQDKDTLVVEFP-
73973662        SGSDWVTIKFMKVDGAKELPDVLERVKFSCMAWTHDGKGMFYNSYPQQDGK-----SDGTETSTNLHQKLCYHVLGTDQSEDVLCAEFP-
73973660        SGSDWVTIKFMKVDGAKELPDVLERVKFSCMAWTHDGKGMFYNSYPQQDGKK---REAGTETSTNLHQKLCYHVLGTDQSEDVLCAEFP-
47222200        SGSDWVTIHFIKADDLSKLPDVLERVKFSCLAWTHDAKGIFYNSYPRQEGK-----TDGTETTANVNQKLYYHVIGTQQSEDVLVAEFP-
73973664        SGSDWVTIKFMKVDGAKELPDVLERVKFSCMAWTHDGKGMFYNSYPQQDGK-----SDGTETSTNLHQKLCYHVLGTDQSEDVLCAEFP-
55235507        SGSDWTKLKIRNVETGEDFPETIEHTKFVTASWTKDNKGFFYARYPVVAGK-----ADGSETAANENQKLYYHRVGESQDKDVLIAEFP-
73973658        SGSDWVTIKFMKVDGAKELPDVLERVKFSCMAWTHDGKGMFYNSYPQQDGKSDVNLESGTETSTNLHQKLCYHVLGTDQSEDVLCAEFP-
71666762        KGSDWQRIHVRRADTVEDTSDVIEWAKFTTIAWWHN-LGFFYTRYPALQGD----VDKGAETDAAQDAFICFHRIGRPQDEDVVILSVP-
49523575        KGSDWQHIHVRRADTAEDTEDVIEWAKFTGISWLHN-TGFFYTRFPALKGD----VDKGAETDAAKDPFVCFHRLGTKQDEDAVVLSLP-
71747962        KGSDWQHIHVRRADTAEDTEDVIEWAKFTGISWLHN-TGFFYTRFPALKGD----VDKGAETDAAKDPFVCFHRLGTKQDEDVVVLSLP-
33317309        KGSDWQRIHVRRADTVEDTSDVIEWAKFTAIAWWHN-LGFFYTRYPALQGD----VDKGAETDAAQDAFICFHRIGRPQDEDVVILSVP-
47224662        SGSDWVEIHFLRVEDALLLEDRLKRVKFSCMSWTHDGKGLFYNSYPDQEGK-----SDGTETSTNLHQKLYFHVLGTSQSEDILCAEFP-
15291259        SGSDWIKILIRDAETGKDLSEVLEKVKFSEISWTKDNKGFFYGRYPDQDGK-----TDGSETKQNENQKLYYHRVGESQDKDTLVVEFP-
73973656        SGSDWVTIKFMKVDGAKELPDVLERVKFSCMAWTHDGKGMFYNSYPQQDGK-----SDGTETSTNLHQKLCYHVLGTDQSEDVLCAEFP-
                 *:**                 :   :     *     .. *  :                        :  * :.  :  *
Prim.cons.      SGSDWVTIKVMDVETGKELPDVLEWVKFSGISWTHDGKGFFYSSYPKPDGGED2N2SDGTETSTNLHQKLYYHVLGTDQSEDVLCFEFPR



                         370       380       390       400       410       420       430       440       450
```

```
                         |         |         |         |         |         |         |         |         |
62738741        PELPKRGHGASVSSDGRWVVITSSEGTDPVNTVHVARVTNG-----KIGP---------VTALIPDLKAQWDFVDGVGDQLWFVSGDGAP
87200538        PDNPRLYHSAETTHDQRWLVVSTSTGSEKGNAVGLARIGG-------DWK---------VQPLVSTLADEWSLIAGIGDRLWFVTSKDAP
103487461       PDKPKLNNSALVTDDGDYLLVVSSEGTDERYGLTLHPLGR------PGAK---------PIVLVDDYANNWEYVTNAGTRFTFLTNKGAP
87200537        PDHPEWSHKALVTSDGRWAVVVSEISTDKRNAVHLIRLTGRE--R-GTWK---------AEALVPDIADHWKLVAGIGERLWFLTDRGAP
94496227        PDRPRLSHQAQVTADGRWLVVSSFEGIDPRREIHVAPLTG------GPIV---------LRRLVRGPAQDWRLIGSRGSTLYFLTDHRAP
35211243        KDQKEWGFGGEVSEDGRYLVINVSQGTDPKNRVFYKDLQN------NSGR---------VMELLPDADAAYTFIGNDGPRFWFVTDKEAP
32448009        PDHPKWGFGASVTDDGRYLVISNWKGTEPKTQVFIQDLTIED--A-P------------VRGLIMGFDADYSFVGSVGSTLYFLTDHEAP
75906687        LDQKEWGFNGVVTEDGCYLIISVWLGTDSRNLVFYKDLTN------PNTE---------VVELIDQFEADYSFIDNDESVFYFRTDLDAP
35211187        KDQKEWGFGGEVSEDGRYLVISVSQGTDPKNRVFYKDLAD------PASA---------VVELLPEADAAYEFIDNDGPLFWFTTDKDAP
67921530        PDEKEWGFSGRVTEDGQYLIISVWQGTDTKNLVFYKNLSQ------ENSK---------IIELINQFEADYSFIDNKGSKFYFRTDLNAP
67930945        PDRPAWSIDPVLTDDGRYLLLIMSSGIPGKNMLSFQDMQS------PDRR---------IVDLIPMETASYQPIEAVGSTLYVQTTDGAP
17131625        LDQKEWGFNGVVSEDGCYLIISVWLGTDSRNLVFYKDLTN------LHAE---------VVELINQFEADYSFIDNDESVFYFRTDWDAP
86609965        PDQKEWGFAGGVTEDGNYLIISVWRGTDPKNLIFYKDLRD------PQSP---------VVELIREFEAEYSFVGNDGSRFWLLTDCQAP
83859556        PDNPEVGWRGFVSDDGQYLIINSSTGTD-GNGVHILDLET------EGAE---------PVEIFEGFGNNHSYVGNDGETFWFQTDLDAS
28808634        EEEKYRYVGGYTSDDERYLFISASVSTS-GNKLFFKDLSKP------NSP---------LKTILDNTSSDTWVIDNQGTKLYLVTNLNAP
75854458        EEEKYRYVGGYTSDDERYLFISASVSTS-GNKLFFKDLSKP------NSP---------LKTILDNTSSDTWVIDNQGTKLYLVTNLNAP
86607173        PDQKEWGFAGGVTEDGDYLIISVWRGTDPKNLLFYKDLRD------PSSP---------VVELIREFQAEYAFVGNDGSRFWLLTDLQAP
23128197        PDQKEWGFSGGVTEDGHYLIISIWLGTDSKNLVFFKDLTN------PNAE---------VVELINQFEADYSFIENDDSFFYFRTDLNAP
91226380        KDEKYRYVGGYTSEDERYLFISASVSTS-GNKLFFKDLSKP------NSP---------LKTILDNTSSDTWVIDNQGTKLYLVTNLNAP
88711708        PEEKHRYIGAGVSEDNRYLTISASTSTS-GNKLFIKDLTKP------NSK---------LVTILDHTDTDSYIIENVGSKLYIVTNMDAP
110167745       PEQKEWSFNCHVTEDNKYLIITVWQSTERKNLVFYQDLSI------PNAP---------IVELISEFEAEYLLIDNYQNIFWFFTDLNAP
109899092       DEQKNRYVSGQVTEDNRYLLISAAVSTS-GNKLYLKDLSEP------DSP---------LVTILENTDSDTNLLDNEGSTLLFVTNLDAP
86144168        PEQKHRYIGAGTTEDNKYLILSASNATS-GNKLYIKDLSKP------NAD---------FVAIVDTEETNTSVIENVGSKLYIVTDKDAP
4973227         PAQHHRYVGATVTEDDRFLLISAANSTS-GNRLYVKDLSQE------NAP---------LLTVQGDLDADVSLVDNKGSTLYLLTNRDAP
6048357         PAQHHRYVGATVTEDDRFLLISAANSTS-GNRLYVKDLSQE------NAP---------LLTVQGDLDADVSLVDNKGSTLYLLTNRDAP
76875127        DEQKHRYVGADVTHDGRYLLISASTSTS-GNKLFIKDLTQK------DSE---------FVTIVGNTDSDTSVIDNEGSKLFLVTNLNAP
216201          PAQRHRYVGATVTEDDRYLLISAADSTS-GNRLYVKDLTRE------GAP---------LLTVQGDLAADVSLVDNKGSRLYLLTNRDAP
88804657        PEQKHRYIRASVTEDQRYLILSAANTTS-GNKLFIQDLQDP------GSG---------LKTIVGDEESDSYVIENSGSKLYIVTDRDAP
89517967        PDQKYRYVGGGVTDDQNYLVISASTSTN-GGKMFMKELNKP------NAP---------LVTVLDNFDTNTYLLHNEGSKLWLVTDYNAP
90588230        ADQKRRYVGGYVTEDNHYLVITAANSTY-GNELYIKDLKTP------NSP---------IITIVDNFNTDNSIIENEGTKLFIHTDYNAP
71143549        AEQKRRYVGGTVTDDNRFLLISGAISTS-GNDLYLKDLTKP------NSP---------LLTITDNFDADTYVIENEGDKLFLVTNLGAP
83855898        QEEKRRYVGGSVTEDGKYLIVSGSVSTS-GNDLRIKDLTKP------NSD---------FKTIISGYETDSYVIENEGSKLYIVTNLNAP
85710915        DAQKHRYVGGYVTDDNQFLVISAANSTS-GNKLFIKDLSKE------DAP---------LVTVVDDYDSDVSVLDSEDGKLYIETNREAP
78368715        PQNKDWGFGAEVSEAGDYLLISISQGTDSRNRFFYKSLID------ANAP---------VVELISELEAEYSFLGNDKSVFYFKTDLDAP
56178266        GAEKHRYVGGYVTDDDRYLMISAANSTS-GNKLFIKDLTEE------DSK---------LVTVLDHTDSDTRVLDNDGSKLYLVTNLDAP
```

```
91215481       EAEKHRYVSGNVTEDDRYLVISASISTS-GNKLMIKDLTNP------DSD---------FIEVVDNYDSDVYVIENKGSKLFMVTDKDAP
78688320       PQNKDWGFGIDVSEQGEYLLLSISQGTDKRNRFFYKSLFE------PKAQ---------VVELILNLEAEYEFLGNDGAVFYFKTDLDAP
24374294       PQNKDWGFGIDVSEKGEYLLLSISQGTDKRNRFFYKSLFE------PKAQ---------VVELILKLEAEYEFLGNDGTVFYFKTDLDAP
456523         DKFPRRYIGAYVTDDQRYLVVSAANATN-GNELYIKDLKN-------KTD---------FIPIITGFDSNVNVADTDGDTLYLFTDKDAP
77816782       PQNKDWGFKIDVSEQGEYLLLSISQGTDKRNRFFYKSLFE------PKSQ---------VVELILNLEAEYEFLGNDGSVFYFKTDLDAP
68547790       PDQPEWGFSSTVSDDGNFLLISVSVGTDSRNRFFYKPLKD------KSLK---------VVELMSELEAEYQFLGNEGPVFYFKTDLDAP
82744257       PQNKDWGFKIDVSEQGEYLLLSISQGTDKRNRFFYKSLFE------PKSQ---------VVELILNLEAEYEFLGNDGSVFYFKTDLDAP
86131249       AAEKHRYVGGSVTKDNRYLLVSARSSTS-GGKLFMKDLTKP------GSD---------FVTILGHEDSDSYVMENVGSKLFIATNLDAP
88796133       DGKPTWNPYPLVIQDGETLLISVFEGYE-ANGVYAKSLKE------KNSK---------LIPIFDKWDGRYDFIGEKEDTLFFTSTADAP
42526704       KDHPLRSFSASTTEDEKTLLLTAFEVGSEGNMLFVADLSEGL--PKCSHC---------FKQYNTHFNDSVWPLETENGFLYLLTNKQAP
27367204       EGELHRYVYGQTSEDDRYLVISGHESTS-GNRLFYVDLQSD------ERQ---------IYTLLDHVDSDTHLLDSTDREFLLYTNLDAP
37676919       EGELHRYVYGQTSEDDRYLVISGHESTS-GNRLFYVDLQSD------ERQ---------IYTLLDHVDSDTHLLDSTDREFLLYTNLDAP
75855979       DTQVHRYVSGTTTTDDRFLIISGAESTS-GNRLFYIDLQSE------SQA---------IVTLRDTTQGDTYLIDSQDETLLLYTNLDAP
28809898       GAQIHRYVSGTTTTDDRFLIISGAESTS-GNRLFYIDLQSD------SQA---------IVTLRDTTQGDTYLIDSQDETLLLYTNLDAP
6460324        PDQPDWGFAAEVTEDGAWLAVQVWLGTSPKNLLWVRPLGKEG--P-GSGD---------FQPLVNDFQAMYQLVGSDGNTLFLQTDEDAP
91223249       DTQVHRYVSGTTTTDDRFLIISGAESTS-GNRLFYIDLQSE------LQA---------IVTLRETTQGDTYLIDSQDATLLLYTNLDAP
56180108       DNQPELNPYAKVSDDGKYLLIDVFKGYD-ANAVYLRPLAQ------ENAP---------FKGLFTKWDGQYQYVTSQKGLLYFKTTNEAS
55769741       -EHPKYSFGASVTEDGKYIILGTYEGCDPVNKLYYCEICT-LPQGIEGFKETKG--MLPFVKLIDNFDAQYHVVANDGDEFTFLTNRNAP
47507507       -EEPKWMGGAEVTDDGQYVLLSIREGCDPVNRLWYCKLNK-N-TGITGTL--------PWVKLIDNFEAEYEYITNEGTIFTFKTNRNAP
28502989       -EEPKWMGGAEVTDDGQYVLLSIREGCDPVNRLWYCKLNK-N-TGITGTL--------PWVKLIDNFEAEYEYITNEGTIFTFKTNRNAP
86144914       NAEQHRYVSGYTTEDDRYLIILGRESTS-GNRLFYIDLNSP------EQQ---------LNTLIDHVDSDTYLIDNQDETFILYTNLDAP
84386759       NAEQHRYVSGYTTEDDHYLIILGRESTS-GNRLFYIDLNSP------EQS---------LNTLIDHVDSDTYLIDNQDEVFILYTNLDAP
63100723       -DEPKWMSGVEVSDDGRYVLLSIREGCDPVNRLWYCDLNE-EPQGITGLL--------PWVKLIDNFDAEYEYVTNEGTVFTFKTNLEAP
38649021       -EEPKWMGGAEVTDDGQYVLLSIREGCDPVNRLWYCKLDK-N-TGITGIL--------PWVKLIDNFDAEYEYITNEGTVFTFKTNRNAP
59711729       EEEKHRYVSGYTTEDDAYLVISASTSTS-GNKLFIEPLNRD------SSY---------RLTVLDHTDSDTYVIENEDTHLLMFTNLNAP
6755152        -DEPKWMGGAELSDDGRYVLLSIWEGCDPVNRLWYCDLQQ-EPNGITGIL--------KWVKLIDNFEGEYDYVTNEGTVFTFKTNRNSP
74215349       -DEPKWMGGAELSDDGRYVLLSIWEGCDPVNRLWYCDLQQ-EPNGITGIL--------KWVKLIDNFEGEYDYVTNEGTVFTFKTNRNSP
50514023       -DEPKWMGGAELSDDGRYVLLSIREGCDPVNRLWYCDLQQ-ESNGITGIL--------KWVKLIDNFEGEYDYVTNEGTVFTFKTNRHSP
26345256       -DEPKWMGGAELSDDGRYVLLSIWEGCDPVNRLWYCDLQQ-EPNGITGIL--------KWVKLIDNFEGEYDYVTNEGTVFTFKTNRNSP
157879458      -DEPKWMGGAELSDDGRYVLLSIREGCDPVNRLWYCDLQQ-ESNGITGIL--------KWVKLIDNFEGEYDYVTNEGTVFTFKTNRHSP
51592147       -DEPKWMGGAELSDDGRYVLLSIREGCDPVNRLWYCDLQQ-ESNGITGIL--------KWVKLIDNFEGEYDYVTNEGTVFTFKTNRHSP
10835490       -DEPKWMGGAELSDDGRYVLLSIREGTDPVNRLWYCDLQQ-ESNGITGIL--------KWVKLIDNFEGEYDYVTNEGTVFTFKTNRHSP
50514022       -DEPKWMGGAELSDDGRYVLLSIREGTDPVNRLWYCDLQQ-ESNGITGIL--------KWVKLIDNFEGEYDYVTNEGTVFTFKTNRHSP
37928243       -DEPKWMGGAELSDDGRYVLLSIREGCDPVNRLWYCDLQQ-ESNGITGIL--------KWVKLIDNFEGEYDYVTNEGTVFTFKTNRHSP
27065055       -DEPKWMGGAELSDDGRYVLLSIREGCDPVNRLWYCDLQQ-ESNGITGIL--------KWVKLIDNFEGEYDYVTNEGTVFTFKTNRHSP
5103285        -DEPKWMGGAELSDDGRYVLLSIREGCDPVNRLWYCDLHQ-EPNGITGIL--------KWVKLIDNFEGEYDYVTNEGTVFTFKTNRHSP
```

```
13786144       -DEPKWMGGAELSDDGRYVLLSIWEGCDPVNRLWYCDLQQ-GSNGINGIL--------KWVKLIDNFEGEYDYITNEGTVFTFKTNRNSP
5107662        -DEPKWMGGAELSDDGRYVLLSIREGXDPVNRLWYCDLQQ-ESNGITGIL--------KWVKLIDNFEGEYDYVTNEGTVFTFKTNRHSP
79382269       -ENPKYMFGAEVTDDGKYLIMSIGESCDPVNKLYYCDMTS-LSGGLESFRGSSS--FLPFIKLVDTFDAQYSVISNDETLFTFLTNKDAP
27066372       -DEPKWMGGAELSDDGRYVLLSIREGCDPVNRLWYCDLQQ-ESNGITGIL--------KWVKLIDNFEGEYDYVTNEGTVFTFKTNRHSP
19347837       -ENPKYMFGAEVTDDGKYLIMSIGESCDPVNKLYYCDMTS-LSGGLESFRGSSS--FLPFIKLVDTFDAQYSVISNDETLFTFLTNKDAP
109072195      -DEPKWMGGAELSDDGRYVLLSIREGCDPVNRLWYCDLQQ-ESNGITGIL--------KWVKLIDNFEGEYDYVTNEGTVFTFKTNRHSP
11691900       -DEPKWMGGAELSDDGRYVLLSIREGCDPVNRLWYCDLQQ-ESSGIAGIL--------KWVKLIDNFEGEYDYVTNEGAVFTFKTNRQSP
41349456       -DEPKWMGGAELSDDGRYVLLSIREGCDPVNRLWYCDLQQ-ESSGIAGIL--------KWVKLIDNFEGEYDYVTNEGTVFTFKTNRQSP
558596         -DEPKWMGGAELSDDGRYVLLSIREGCDPVNRLWYCDLQQ-ESSGIAGIL--------KWVKLIDNFEGEYDYVTNEGTVFTFKTNRQSP
53127346       -DEPKWMGGAEISDDGRYVLLSIREGCDPVNRLWYCDLQK-ESQGISGIL--------QWVKLIDNFEAEYEYVTNEGTVFTFKTNRHSP
904214         -DEPKWMGGAELSDDGCYVLLSIREGCDPVNRLWYCDLQQ-ESSGIAGIL--------KWVKLIDNFEGEYDYVTNEGAVFTFKTNRQSP
15217929       -DNPKHMFGSKVTDDGKYLIMSIEEGCDPVNKVYHCDLSL-LPKGLEGFRGSNT--LLPFVKLIDTFDAQYIAIANDETLFTFLTNKDAP
9558588        -DNPKHMFGSKVTDDGKYLIMSIEEGCDPVNKVYHCDLSL-LPKGLEGFRGSNT--LLPFVKLIDTFDAQYIAIANDETLFTFLTNKDAP
38345850       -DHPKYIYTPEVSEDGKYVILSVAETSEPVNKLYYCDLSA-LPDGLEGMKGNHGNAMLPFVKLVDEFEAYYALIANDDTQFTFLTNKNAP
91082735       -DDPHLRIGAHVSHCGKYLVITGTKGCK-NNLLYFAQLD--SGKITGKLK---------LTEVVTEFVADFEYITNDKNLFYFHTNKNAS
54650756       -EEPSWRIQSTVSDCGKYLILAIVKDCR-DNIVFYADLTP-GAEITSKLN---------VKKIVEKFEADYDYITNEGSKIFFRTNKNAP
73973662       -DEPKWMGGAELSDDGRYVLLSIREGCDPVNRLWYCDLQQ-ESNGISGKL-----SFKQWVKLIDNFEGEYDYVTNEGTVFTFKTNRHSP
73973660       -DEPKWMGGAELSDDGRYVLLSIREGCDPVNRLWYCDLQQ-ESNGISGMLGLSP-GILKWVKLIDNFEGEYDYVTNEGTVFTFKTNRHSP
47222200       -QHPKWHSSATISDDGRYAVLSITEGCEPVNQLWYCDLQQ-LPDGITGLL--------PWVKLVDNFEAQYSYITNEGTVFTFHSNLEAP
73973664       -DEPKWMGGAELSDDGRYVLLSIREGCDPVNRLWYCDLQQ-ESNGISGMLGLSP-GILKWVKLIDNFEGEYDYVTNEGTVFTFKTNRHSP
55235507       -EEPSWRLMPEVSDCGKYLMLFIMKGCK-DMLLYFSNLEK-AGTLESKLD---------FVKVVTEFDSDYDYVTNEDNIFSFRTNKVAP
73973658       -DEPKWMGGAELSDDGRYVLLSIREGCDPVNRLWYCDLQQ-ESNGISGFQ-STA-GILKWVKLIDNFEGEYDYVTNEGTVFTFKTNRHSP
71666762       -EHPQWNMGASVSDCHSYVIVVLFDGCEPHNLVWVAELPS-VEKGLGSEP-------LVFKKLVNEFAGRYTYLGNEGSTFYFVTTRDAP
49523575       -EHPHWGVSAEVSNCHSYLVVSITDGCEPKNLIWITKLPIDGAEKLSPST-------LTYNKLFNEFVGSFEYLGNDGTTFYFVTTRDAP
71747962       -EHPHWGVSAEVSNCHSYLVVSITDGCEPKNLIWITKLPIDGAEKLSPST-------LTYNKLFNEFVGSFEYLGNDGTTFYFVTTRDAP
33317309       -EHPQWNMGASVSDCHSYVIVVLFDGCEPHNLVWVAELPS-VEKGLGSEP-------LVFKKLVNEFAGRYTYLGNEGSTFYFVTTRDAP
47224662       -DHPKWMSGAEVSEDGRYVLLSIREGCDPVNRLWYCDLKT-IPQGITGLL--------PWVKLIDNFDAEYEYVTNEDTQFTFKTNLDAP
15291259       -EEPSWRIQSTVSDCGKYLILAIVKDCR-DNIVFYADLTP-GAEITSKLN---------VKKIVEKFEADYDYITNEGSKIFFRTNKNAP
73973656       -DEPKWMGGAELSDDGRYVLLSIREGCDPVNRLWYCDLQQ-ESNGISGIL--------KWVKLIDNFEGEYDYVTNEGTVFTFKTNRHSP
                                   :           .    :                                          : . :   :.
Prim.cons.     PDEPKWMGGAEVSDDGRYLLLSIREGTDPVNRLFYCDLQQPESNGITG2LGSS4NG2LKWVKLIDNFEAEYDYITNEGTVFTFKTNRDAP


                      460       470       480       490       500       510       520       530       540
```

```
                         |         |         |         |         |         |         |         |         |
62738741        LKKIVRVDLSGST---------PR-------------FDTVVPESKDNLES--VGIAGNRLFASYIHDAKSQVLAFDL-DGKPAGAVSLP
87200538        RKKVVMVDMSGAA---------PV-------------TTTVVPESDDVLES--AKVVGDRLVLGYLRDVKAELRLATL-DGKPAGTLALP
103487461       RGRLVSFDIRKP----------DK-------------LTELVAENPATLVG--ASRVGDRIILSYLGDAKSEARMVAL-NGEPIANINLA
87200537        NYHLVRVDLSRPQ---------EG-------------WQVVVPQRGNTLEG--ARMIGDRFLLSYLRDGQSVAVMTDR-KGRPGKAITLN
94496227        HRRVVTLDAANPR---------RR-------------AQEIVPERAQTLVG--GALVGDRLILASMNGAQTIAELVEL-DGRRVGDVPLP
35211243        RGRLVLIDTTRP----------LQ-------------LQEVVPQTDATLQS--ADIVGERLFLRYLKDARSQVKVFDL-KGKFLSEVAFP
32448009        RRRVISLDVAEHA---------KRTDDNVDEPADRAGWEEVIPQSEHVLEH--VSLLSGVFFANYLADALNQVERFSL-DGSPMGPLELP
75906687        RGRVIAIDIANP---------AKEI------------WREIIPQAEATLES--VNILNNQFIAGYLEDARSQVKIFDL-NGTLVRNVELP
35211187        RGRVVAIDITQP----------LH-------------LNELIPESEDTLQG--VSILDNKFFANYLKDAHTQVRIYDL-QGQYVGEVELL
67921530        KGKVIAIDIDNS---------QPEN------------WQEIIPESNEALKG--VGILNNQFVCDYLQDAKSAIKIHDL-QGNLLRNVSLP
67930945        RGRVIAIDLQKPV---------PS------------KWREIVPEAAETLES--VQMADGKLLLAYMKDAHAAARLVTT-EGKPVAEVAMP
17131625        RGRVIAIDTANP---------SQEI------------WREIIPQAEATLES--VNILNNQFIADYLEDARSQVKIFDL-NGTLIRDVELP
86609965        RRRLVAIDLEQP----------DR-------------VQEVIPEAEETLQG--VSLINNQFVAFYLKDAHTQIKTFAL-DGTYLGEIPLP
83859556        NGRIVSVSLSAP----------ED-------------LSDVVAEGDGPITG--ASHVGGHLIIETMRDVASAVSVYTP-AGDFVREVALP
28808634        NKRVVTVDASQPQP---------------------KNWKDLIPETKNVLR---VSTAGGNLFASYIVDAISMVKQYDM-NGKLIREIKLP
75854458        NKRVVTVDASQPQP---------------------KNWKDLIPETKNVLR---VSTAGGNLFASYIVDAISMVKQYDM-NGKLIREIKLP
86607173        RRRLVAIDLDNP----------GQ-------------LQEVIPEAEETLQG--VSLIHNQFVAFYLKDAHTQIRTFAL-DGTYLGEIPLP
23128197        RGRVIAIDTQNP---------ASEK------------WREIIPQSAETLES--VGILNNQFVADYLKDAHSEIKIFDL-KGGFIREVELP
91226380        NKRVVTVDASQPQP---------------------KNWKDLIPETKNVLR---VSTAGGNLFASYIVDAISMVKQYDM-NGKLIREIKLP
88711708        NQKIVTVDAANPTP---------------------EKWIDFIPETENVLS---PNTGGEYFFAEYMVDAISKVFQYDY-DGNLVREVKLP
110167745       KRRVIAIDINNPPSPSLVRGENQNK------------WQEIIPEATDALQG--IGTLNNQFVTFYLKDAHTQIKIFNL-DGSPVRNVELP
109899092       NKRVVSVDAGQPQP---------------------ENWQDFIAQTDNVLD---VSTGGGYIFASYMVDAISNVKQLDM-SGKLVREIALP
86144168        NRKIVTVDAANPTP---------------------ENWEDFIPETENVLS---PSTGGGYFFAEYMVDAVSQVKQYDY-EGKLVREVELP
4973227         NRRLVTVDAANPGP---------------------AHWRDLIPERHRVLT---VHSGTAYLFAEYMVDATAPVEQFDY-EGKRVREVALP
6048357         NRRLVTVDAANPGP---------------------AHWRDLIPERHRVLT---VHSGSGYLFAEYMVDATAPVEQFDY-EGKRVREVALP
76875127        NQKVVTVDAADPGP---------------------QNWQDFIPETKNVLK---LTKGGDTFFANYMVDAISKVKQYNK-KGELIRDITLP
216201          NRRLVTVEADNPGP---------------------EQWRDLIPERQQVLT---VHSGGGYLFAEYMVDATARVEQFDH-DGKRVREVGLP
88804657        NKRIVTVDASDPGP---------------------ENWVDFIPETENVLS---PNTGGGYFFAEYMVDAISKVLQYDY-QGNLVREVPLP
89517967        NKRIVSTDFSNPTQ---------------------ENWTDVISETENVLS---PSTGGGYIFAEYMVDAVSKIQQYNY-DGTLVREVKLP
90588230        NGRVVTVDFSNPKQ---------------------ENWKDFIKETENVLS---PSTGGGYFFANYTKDAVSLVLQYDY-NGKLVREIKLP
71143549        NKKIVTVNAKAPAS---------------------KNWTDFIAETDYVLS---ASTGGGFFFTEYMVDAISKVYQYDY-QGKQVREINLP
83855898        NKKIVTVDAENPSP---------------------ENWVDFIPETEHVLS---PSKAGGYFFAEYMVDAVSEVKQYDY-AGKLIREVKLP
85710915        NGRVVVVDAGAPQE---------------------ANWSDLIPETEHVLN---PTTGAGYIFAEYMVDAVSQVKQFDY-QGKEIRQIDLP
78368715        NGKIVAVDVNRP---------AKAN------------WKTVVPESSDPIAS--VAIINDHLVVSYLHDVLGKLTIFSM-NGVKRQDVDLP
56178266        NQRVATVDASNPTP---------------------ENWEDFIPETENVLS---VSTGAGYIFAEYMVDAVAKVKQYAY-NGDMVREVQLP
```

```
91215481     NKKIVTVEASAPQP---------------------ENWEDFIPETEHVLS---PSTGAGYFFAEYMVDAVSQVKQYNY-EGELVRDIELP
78688320     NGKVIAIDTRNS---------DKSQ------------WQTIIPESKDPINK--VAIINDHLVVSYLHDVLGQLSIFSM-GGQKRQDVTLP
24374294     NGKVIAIDTRNS---------DKSQ------------WQTIIPESKDPINK--VAIINDHLVVSYLHDVLGQLSIFSM-GGQKRQEVTLP
456523       NKRLVKTTIQNPKA---------------------ETWKDVIAETSEPLE---INTGGGYFFATYMKDAIDQVKQYDK-NGKLVRAIKLP
77816782     NGKVIAIDTRNS---------DKSQ------------WQTIIPESKDPINN--VAIINDHLVVSYLHDVLGQLSIYSM-GGQKRQDVALP
68547790     RGRIIAVDTRNP---------AKDN------------WRTLVPESQDPIAE--VSIISEHLVVSYLHDVLGQLSIYNM-DGVKREDVPLP
82744257     NGKVIAIDTRNS---------DKSQ------------WQTIIPESKDPINN--VAIINDHLVVSYLHDVLGQLSIYSM-GGQKRQDVALP
86131249     NMRVVTVDASNPTP---------------------ENWVDFIPETENVLS---PSTGGGSFFARYMVDAVSKVKQYDY-SGKLVRDVELP
88796133     TGKVVKVEFDNG-----------VQS-----------SETVIESTSDTLNS--VSLLGGKLFAQYLKDVKGQVSVFDL-SGKQIDDISFS
42526704     FYRVVKTSLNNISE---------------------KSIDEVIPQKDCLLSS--AALCGGKLLTVYLRDVQDEAFICGL-DGKNSTKINLP
27367204     NGKVVSVDIES------------------------GQWRDVIAEQDQPLE---VVAAGGYLFATYMVDALSKVVQYSY-QGKKVREITLP
37676919     NGKVVSVDIES------------------------GQWRDLIAEQDQPLD---AVAAGGYLFATYMVDVLSKVVQYSY-QGEKVREITLP
75855979     NGKVVSYNTRT------------------------AQWMDVIAEQEQPLE---ISKGGGYLFATYMVDVLSKVQQFNY-QGEWIRDVELP
28809898     NGKVVSYNTQT------------------------EQWADVIAEQEQPLE---ISKGGGYLFATYMVDVLSKVQQFTY-QGEWIRDVELP
6460324      LGKLMAWNIRTGE---------RR-----------D----LLPEGADKLEQ--VLTVPGGFLALTLHDASHRLTLYDR-NGERQREIELP
91223249     NGKVVSYNTQT------------------------EQWADVIAEQEQPLE---ISKGGGYLFATYMVDVLSKVQQFNY-QGEWIRDVQLP
56180108     TGRVVAVDPKQSE----------PEN-----------WTEIIAAKTHTLKD--VHAVNGQLFAHYLQSARSRVAIFDI-KGTMLQELALP
55769741     KNKLVRVDIKKPE-----------------------LWTDILPEHERDVLESADAVNGNQLLVCYMSDVKHILQIRDLVTGNLLHKLPL-
47507507     NYNLINIDFNNPEE---------------------SNWKALVPEHQKDVLEWVSCVHKKFLVLCYLHDVKNILQLHDLDSGSHLKTFPL-
28502989     NYNLINIDFNNPEE---------------------SNWKALVPEHQKDVLEWVSCVHKKFLVLCYLHDVKNILQLHDLDSGSHLKTFPL-
86144914     NGKVVSFDTRD------------------------GKWLEIIPEKPQPLD---ISTGGGYLFAHYMVDVVSKIEQLDY-QGNLVREIHLP
84386759     NGKVVSFDTSS------------------------QQWLDIIPEKPQPLD---ISTGGGYLFAHYMVDVVSKIEQLDY-QGNLVREIHLP
63100723     QYRLINIDFTQPSV---------------------SQWKELIPQHDKDVIVFATCTFSSFLFVCFLHDVKNVLKMFHLSSGEEIRTFPL-
38649021     NYKLINIDFNNPEE---------------------SNWKVLVPEHEKDVLEWVSCVHKKYLVLCCLHDVKNVLQLHDLASGSHLKTFPL-
59711729     NGKVVSYDVVT------------------------GEWTDIIPEQPQPLD---ISVGGGTLFATYMVDVLSQVKQYDF-SGNLLREIELP
6755152      NYRLINIDFTDPDE---------------------SKWKVLVPEHEKDVLEWVACVRSNFLVLCYLHDVKNILQLHDLTTGALLKTFPL-
74215349     NYRLINIDFTDPDE---------------------SKWKVLVPEHEKDVLEWVACVRSNFLVLCYLHDVKNILQLHDLTTGALLKTFPL-
50514023     NYRLINIDFTDPEE---------------------SKWKVLVPEHEKDVLEWVACVRSNFLVLCYLHDVKNTLQLHDLATGALLKIFPL-
26345256     NYRLINIDFTDPDE---------------------SKWKVLVPEHEKDVLEWVACVRSNFLVLCYLHDVKNILQLHDLTTGALLKTFPL-
157879458    NYRLINIDFTDPEE---------------------SKWKVLVPEHEKDVLEWVACVRSNFLVLCYLHDVKNTLQLHDLATGALLKIFPL-
51592147     NYRLINIDFTDPEE---------------------SKWKVLVPEHEKDVLEWVACVRSNFLVLCYLHDVKNTLQLHDLATGALLKIFPL-
10835490     NYRLINIDFTDPEE---------------------SKWKVLVPEHEKDVLEWVACVRSNFLVLCYLHDVKNTLQLHDLATGALLKIFPL-
50514022     NYRLINIDFTDPEE---------------------SKWKVLVPEHEKDVLEWVACVRSNFLVLCYLHDVKNTLQLHDLATGALLKIFPL-
37928243     NYRLINIDFTDPEE---------------------SKWKVLVPEHEKDVLEWVACVRSNFLVLCYLHDVKNTLQLHDLATGALLKIFPL-
27065055     NYRLINIDFTDPEE---------------------SKWKVLVPEHEKDVLEWVACVRSNFLVLCYLHDVKNTLQLHDLATGALLKIFPL-
5103285      NYRLINIDFTDPEE---------------------SRWKVLVPEHEKDVLEWVACVRSNFLVLCYLHDVKNTLQLHDMATGALLKTFPL-
```

```
13786144      NYRLINIDFTDPDE---------------------SKWKVLVPEHEKDVLEWVACVRSNFLVLCYLRNVKNILQLHDLTTGALLKTFPL-
5107662       NYRLINIDFTDPEE---------------------SKWKVLVPEHEKDVLEWVACVRSNFLVLCYLHDVKNTLQLHDLATGALLKIFPL-
79382269      KYKLVRVDLKEPN-----------------------SWTDVVEEHEKDVLASACAVNGNHLVACYMSDVKHILQIRDLKSGSLLHQLPL-
27066372      NYRLINIDFTDPEE---------------------SKWKVLVPEHEKDVLEWVACVRSNFLVLCYLHDVKNTLQLHDLATGALLKIFPL-
19347837      KYKLVRVDLKEPN-----------------------SWTDVVEEHEKDVLASACAVNGNHLVACYMSDVKHILQIRDLKSGSLLHQLPL-
109072195     NYRVINIDFRDPEE---------------------SKWKVLVPEHEKDVLEWVACVRSNFLVLCYLHDVKNILQLHDLTTGALLKTFPL-
11691900      NYRVINIDFRDPEE---------------------SKWKVLVPEHEKDVLEWIACVRSNFLVLCYLHDVKNILQLHDLTTGALLKTFPL-
41349456      NYRVINIDFRDPEE---------------------SKWKVLVPEHEKDVLEWIACVRSNFLVLCYLHDVKNILQLHDLTTGALLKTFPL-
558596        NYRVINIDFWDPEE---------------------SKWKVLVPEHEKDVLEWIACVRSNFLVLCYLHDVKNILQLHDLTTGALLKTFPL-
53127346      NYRLINIDFSDPEE---------------------SKWKVLIPEHERDVLEWVACVRSNFLVLCYLHDVKNILQLHDLATGAHLKTFPL-
904214        NYRVINIDFRDPEE---------------------SKWKVLVPEHEKDVLEWIACVRSNFLVLCYLHDVKNILQLHDLTTGALLKTFPL-
15217929      KYKVVRVDLKEPS-----------------------SWTDVIAEHEKDVLSTASAVNGDQLVVSYMSDVKHILQIRDLKSGSLLHGLPV-
9558588       KYKVVRVDLKEPS-----------------------SWTDVIAEHEKDVLSTASAVNGDQLVVSYMSDVKHILQIRDLKSGSLLHGLPV-
38345850      KYKLSRIDVNEPH-----------------------SWMDILPEDEKAVLESACAVHGDKLLVNYLSDVKYVLQMRSLVTGELLHDIPI-
91082735      NYRIVIIDFDNPKE---------------------SEWKDLISEHPKDVLDWAHAINENMLVVCYLQDVKNIMQLYDIKSSNKLHDFKL-
54650756      NYQVIAIDFNNSAE---------------------DKWETLIAEHKSDVLDWVKCVDNDKLLVCYIRDVKSVLQVNSLKDGTLLREFDL-
73973662      NYRLINIDFTDPEE---------------------SKWKVLVPEHEKDVLEWVACVRSNFLVLCYLHDVRNILQLHDLATGALLKTFPL-
73973660      NYRLINIDFTDPEE---------------------SKWKVLVPEHEKDVLEWVACVRSNFLVLCYLHDVRNILQLHDLATGALLKTFPL-
47222200      RYRLINIDIQKPER---------------------QHWTTIIPQHDKDVMGFVSCVNQRHLLVNYLHDVKDILQVCELSTGRRLRDLPL-
73973664      NYRLINIDFTDPEE---------------------SKWKVLVPEHEKDVLEWVACVRSNFLVLCYLHDVRNILQLHDLATGALLKTFPL-
55235507      NYRIVNIDFEQPEM---------------------EHWKTLVPEHPKNVLDWTTCVNKDRIVLGYIDDVKSLLVVHSLADGSFVSKFPL-
73973658      NYRLINIDFTDPEE---------------------SKWKVLVPEHEKDVLEWVACVRSNFLVLCYLHDVRNILQLHDLATGALLKTFPL-
71666762      RKKIVSIDIHTGQ-------------------------ETVIVEQQRSVLSQAALVKN-TLLLAYLEDVKDVFYYCRLEDPT-LNAIPL-
49523575      RKKIISVDIATGK-------------------------EQIVVGEREAVLDHAALVKN-TLILVYLEDVKSTLYYCCLDKPE-LKKITI-
71747962      RKKIISVDIATGK-------------------------EQIVVGEREAVLDHAALVKN-TLILVYLEDVKSTLYYCCLDKPE-LKKITI-
33317309      RKKIVSIDIHTGQ-------------------------ETVIVEQQRSVLSQAALVKK-TLLLAYLEDVKDVFYYCRLEDPT-LNAIPL-
47224662      RYRLINIDFACPDP---------------------SSWKELLPQHDKDVI------------------A---------------------
15291259      NYQVIAIDFNNSAE---------------------DKWETLIAEHKSDVLDWVKCVDNDKLLVCYIRDVKSVLQVNSLKDGTLLREFDL-
73973656      NYRLINIDFTDPEE---------------------SKWKVLVPEHEKDVL----------------------------------------
                .:                                    .:      :
Prim.cons.    NYRVVNIDFTNPEEPSLVRGE2K2Q3DNVDEPADRSKWKDLIPEHEKDLLEWVACVRGNFLVACYLHDVKSILQLHDLATGALLREFPLP


                     550       560       570       580       590       600       610       620       630
```

```
                     |         |         |         |         |         |         |         |         |
62738741     GIGSASGLSGRPGDRHAYLSFSSFTQPATVLALDP-ATAKT--------TPWEPVHLT---FDPADFRVEQVFYPSKDGTKVPMFIVRRK
87200538     GIGSIGGVVGEPGDPQGHFAFSGFTQPATIYAFDAGDAASA--------KVWAAPKLT---FDPARFETRQVFYPSKDGTRIPMFVVRRK
103487461    DIGAASGFGGKSSDPETFYAFSSFARPTTIYRFDT-ETGNS--------EIFAEPRLT---FNPADFSVEQRFYKSKDGTEVPMFLVMKK
87200537     GIGTASGFGGRPGDTETFYQFTSFNMPPAVYRMDL-RTGAV--------TPFAVPRMA---FDPADYDVEQRQFTSKDGTKVPIYIVRKR
94496227     GFGTAAGFGGRSGDPETFFSFSGFVTPASIYRFDT-ATRQF--------QLFAQPDLP---FDPNDYGIEQRLYPSKDGTMIPLTILRKR
35211243     GLGTVTGFGGKRTDTETFYAFTSFTTPTTIYRYDI-PTAKS--------TVLFQPKVD---FDPTAYTTEQVFFNSKDGTRIPMFITYKK
32448009     GKGSVGGLGGRQDAKETFFSFTNYVTPPSIHRVDV-ATGKS--------ELAIMPEVA---FDVSQYVTEQVFCTSKDGTKVPILITRHK
75906687     GLGAVDGFGGKRGDTETFYKFTSFTTPGTIYRYNL-VTGKS--------EVFRETNVD---FNPDNYETKQVFYQSQDGTQVPMFITHKK
35211187     GIGSAGGFGGKRTDTETFYAFTSFTTPTTIYRYDI-PTAKS--------TVLFQPKVD---FDPTAYTTEQVFFNSKDGTRIPMFITYKK
67921530     GIGSVGGFDGKKEDTETFYSFTSFTTPSTVYRYDM-ITGES--------EIFCQPNVD---FNPHNYEIKQVFYKSKDGTKIPMFITHKK
67930945     GLGTAIWSPARVKDKEMFYGFMTYLAPTTMYRLDV-ETGRS--------EVARPVKLP---FDPSGFETTQVFYPSKDGSRIPMFLTRRK
17131625     GLGAVGGFGGKRDDTETFYKFTSFTTPGTIYRYNL-VTGKS--------EVFRETIVD---FNPDNYETKQVFYQSKDGTQVPMFITHKK
86609965     GLGSASGFGGKRYDTETFYTFTSFTTPPTIYRYDF-TSGTS--------TLFRQPQVD---FDPQAYEVQQVFYTSKDGTRIPMFLVHRR
83859556     GLGVASGFGGDPQRSETFYSYESLNQPATLYRYDV-ETGES--------EVFRAPELT---FNPDDYVVSQTFYESTGGARIPMFVAHHR
28808634     DIGSAYGFSGKKDETEVYYSFTNYKMPSTTYRLNI-KDGDS--------EVYYKSKAP---FDPALYDSRQVFYTSKDGTKVPMIITYKK
75854458     DIGSAYGFSGKKEETEVYYSFTNYKMPSTTYRLNI-KGGDS--------EVYYKSKAP---FDPALYDSRQVFYTSKDGTKVPMIITYKK
86607173     GLGSASGFGGKRHDTETFYTFTSFTTPPTIYRYDF-TSGRS--------TLFRQPQVD---FDPQAYEVQQVFYASQDGTRIPMFLVHRR
23128197     GLGSAGGFGGKRHDTETFYSYTSFTTPGTIYRYDM-ITGKS--------TVFRQPEVD---FNPNDYETKQIFYQSKDGTRVPMFITHKK
91226380     DIGSAYGFSGKKEETEVYYSFTNYKMPSTTYRLNI-KDGDS--------EVYYKSKAP---FDPVLYESRQVFYTSKDGTKVPMIITYKK
88711708     GVGSASGFGGKKEDKEFYFSFTNYNTPGSSYKYNV-ESGEY--------EQYWKPDID---FNPNDYESHQVFYNSKDGTKIPMIITHKK
110167745    GIGSVVGFYGRRHDTSTFYSYVSFTTPSTIYHYDM-VSGES--------KIYRQSNVD---FNPNQFETKQVFYSSKDGTSIPMFITHKK
109899092    GVGTASGFDGKKDQQTLYYSFTNYKTPSTIFSLNV-ESGES--------SVYLKSKAN---FDSDAYESKQVFYTSKDGTKVPMIITHKK
86144168     GIGSAGGFGAKKEDKELYYSFTNYTSPGSIYKYDI-ESGTS--------ELFRKPAIN---FDPENYESKQVFYTSKDGTKVPMIITFKK
4973227      GLGSVSGFNGKHDDPALYFGFENYAQPPTLYRFEP-KSGAI--------SLYRASAAP---FKPEDYVSEQRFYQSKDGTRVPLIISYRK
6048357      GLGSVSGFNGKHDDPALYFGFENYAQPPTLYRFEP-KSGAI--------SLYSASAAP---FKPEDYVSEQRFYQSKDGTRVPLIISYRK
76875127     GVGTAGGFSGKKEQTTLYYSFTNYKTPGTTYTFDV-TTGES--------GVYRKSGVD---FNSDDYTSEQVFYNSKDGTKVPMIITYKN
216201       GLGSVSGFNGKQDDPALYFGFENYAQPPTLYKFEP-NSGAI--------SLYRASAAP---FKPEDYVSEQRFYRSKDGTRVPLIISYRK
88804657     GVGSASGFGGKKEEKEFYFYFTNYNTPGSSYKYNV-ETGEY--------ELYWKPDID---FNPDDYVSNQVFYESADGTRIPMIITHKK
89517967     GIGNVGGFGAKEKDKTLYYSFTNYTTPGSTYLYDI-ATGES--------KLYRKPEIQ---FNSDEYESKQIFYTSKDGTKIPMIISHKK
90588230     AVGTAGGFGAKKEEKILYYSFTNYTTPGSIFSFEP-KSGKS--------EIYQKPKVD---FKSEDYESKQVFYTSKDGTKVPMIITYKK
71143549     GVGSASALEGDKDESTLYYSFSNYKTPGTIYSYNI-DKGNS--------DVYRKSGAK---FDSDAYESKQVFYPSKDGTKVPMIITYKK
83855898     GVGTVGGFGAKKEDKELYFSFTNYVTPGSIYKYDI-EDGNS--------ELYVKPEID---FNPDHYKSEQVFFNSKDGTKIPMIITYKK
85710915     GVGSAGGFGGKDDAETLYYSFTNYSTPQTIYALNP-DTGES--------TVYEESGAD---FDSSNYESKQVFYTSKDGTQVPMIITYKK
78368715     GKGKIAGPYGKRNKDYFYYTFNSYVQPQTTYKFDF-KTGES--------SLYAKPKVA---FNPDNYVSEQVFYTSKDGTRVPMMISYKK
56178266     GVGSIGGFSGKKEADELYYTFTNYSTPSTIYKFDP-DQGGS--------EVYAESGAD---FDSANYESKQVFYTSKDGTEVPMIITYKK
```

```
91215481      GVGSVGGFGGKKEAKDIYFSFTNYTTPGTIYKFTP-ENGTY--------VVYQKPDID---FDTEAYESKQVFYPSKDGTKIPMIITHKK
78688320      GKGNVAGPFGKASKDYFYYVFNSYIQPETTYKFDF-KTAES--------TVVAKPQVS---FNPDDYVSEQVFYTSKDGTRVPMMVSYKK
24374294      GKGNVAGPFGKASKDYFYYVFNSYIQPETTYKFDF-KTAES--------TVVAKPQVS---FNPDNYVSEQVFYTSKDGTRVPMMLSYKK
456523        GSGNASGFGGEKTEKDLYYSFTNYITPPTIFKYNV-TTGNS--------EVYQKPKVK---FNPENYVSEQVFYTSSDGTKIPMMISYKK
77816782      GLGNVTGPFGKASKDYFYYVFNSYIQPETTYKFDF-KTGGS--------TVIAKPKVS---FNPDNYISEQVFYTSKDGTRVPMMLSYKK
68547790      GKGKIAGPYGKASKPYFYYIFNSYVQPQTIYKYDL-NKGTN--------EVFSAPKVS---FNPDDYVSEQVFYTSKDGTRIPMLVSYKK
82744257      GLGNVTGPFGKASKDYFYYVFNSYIQPETTYKFDF-KTGGS--------TVIAKPKVS---FNPDNYISEQVFYTSKDGTRVPMMLSYKK
86131249      GVGSVGGFGAEKDEKELYYSFSNYKTPGSIYKYDI-ASGTS--------ELFIKPDID---FDPEAYESKQVFYNSKDGTKVPMIITYKK
88796133      DIGSVNGFYGDETADTTFYKLTGFTNPGQVYAYDV-QSGES--------TLFKRIDTG---INYDDYVTKQVFYTSKDGTKVPMFIVHKK
42526704      ANGSISFSGTRKNEDSLFFNFTSYTTPNKIIRYDI-KTNSL--------TDFFVPAIP---INTGDFKCEQVFFKSKDGTKIPMHIVSKK
27367204      GEGTATGLEGKKSQTTLYYTFTNYVTPPTIFSLDV-NSGES--------QVFQESKAP---FDRTQFESQQVFYPSKDGTLIPMIISYKK
37676919      GEGTATGLEGKKSQTTLYYTFTNYVTPPTIFSLDV-NSGES--------QVFQESKAP---FDRHQYESQQVFYPSKDGTLIPMIISYKK
75855979      GEGTAYGLAGKKEETTLYYTFTNYVTPPTIFSFDV-ESGES--------TLFQESKAP---FDRSEYESKQVFYTSKDGTQVPMIISYKK
28809898      GEGTAYGLAGKKEETTLYYTFTNYVTPPTIFSFDV-ESGAS--------TLYQESKAP---FDRNEYESKQVFYTSKDGTQVPMIISYKK
6460324       ALGTVS-VSAEQDSSEVFVAFTSFLVPSRPYRLKL-PGGEL--------EPLADPALD---FDAATYEVTQEFAMSKDGTRVPMFIVARK
91223249      GEGTAYGLAGKKEETTLYYTFTNYVTPPTIFSFDV-ESGES--------TLFQESKAP---FDRNEYESKQVFYTSKDGTQVPMIISYKK
56180108      GIGSISGFEGSHKAPDVFFSFKSFTEPGKIYRYQP-DQQRT--------TLWYETKVP---AELDQYTTEQVVYNSADGTEVPMFLVHDK
55769741      EIGSVSEISCRREDMDVFIGFTSFLSPGIIYRCNLTSAIPE-------MKIFREISVP--GFDRTNFEVKQIFVNSKDGTKIPMFIMSKR
47507507      NVGSIVGYSGQKKDSEIFYQFTSFLSPGIVFHCDLTKEELN-------PTVYREVSVK--GFDPSDYQTIQVFYPSKDGTMIPMFIVHKK
28502989      NVGSIVGYSGQKKDSEIFYQFTSFLSPGIVFHCDLTKEELN-------PTVYREVSVK--GFDPSDYQTIQVFYPSKDGTMIPMFIVHKK
86144914      GLGTASGLGGKNEQTQLYYTFTNYVTPPTIFSFDV-ESGSS--------EIYQRSESP---FESDQFESKQVSYTSKDGTKVPMLISYKK
84386759      SLGTASGLGGKNEQTQLHYTFTNYVTPPTIFSFDV-ESGSS--------EIYQRSESP---FESDQFESKQVFYTSKDGTQVPMLISYKK
63100723      DVGSIVGFTGRKKDSEIFYSFTSFLSPAIIYHCDLTMEPLQ-------PHVFREVTVK--GFNPADYQTTQVFYPSKDGTQIPMFIVHKK
38649021      DVGSVVGYSGQKKDSEIFYQFTSFLSPGIIFHCDLTKEELK-------PTVFREVSVK--GFDPSDYQTVQVFYPSKDGTMIPMFIVHKK
59711729      GEGQATGFSGKREQTDLFFTFTNYVTPPTIYQFDV-KTGET--------ELYLASASP---FDSEKFESKQVFYTSKDGTQIPMLISYRK
6755152       DVGSVVGYSGRKKDSEIFYQFTSFLSPGVIYHCDLTKEELE-------PMVFREVTVK--GIDAADYQTIQIFYPSKDGTKIPMFIVHKK
74215349      DVGSVVGYSGRKKDSEIFYQFTSFLSPGVIYHCDLTKEELE-------PMVFREVTVK--GIDAADYQTIQIFYPSKDGTKIPMFIVHKK
50514023      EVGSVVGYSGQKKDTEIFYQFTSFLSPGIIYHCDLTKEELE-------PRVFREVTVK--GIDASDYQTVQIFYPSKDGTKIPMFIVHKK
26345256      DVGSVVGYSGRKKDSEIFYQFTSFLSPGVIYHCDLTKEELE-------PMVFREVTVK--GIDAADYQTIQIFYPSKDGTKIPMFIVHKK
157879458     EVGSVVGYSGQKKDTEIFYQFTSFLSPGIIYHCDLTKEELE-------PRVFREVTVK--GIDASDYQTVQIFYPSKDGTKIPMFIVHKK
51592147      EVGSVVGYSGQKKDTEIFYQFTSFLSPGIIYHCDLTKEELE-------PRVFREVTVK--GIDASDYQTVQIFYPSKDGTKIPMFIVHKK
10835490      EVGSVVGYSGQKKDTEIFYCFTSFLSPGIIYHCDLTKEELE-------PRVFREVTVK--GIDASDYQTVQIFYPSKDGTKIPMFIVHKK
50514022      EVGSVVGYSGQKKDTEIFYQFTSFLSPGIIYHCDLTKEELE-------PRVFREVTCK--GIDASDYQTVQIFYPSKDGTKIPMFIVHKK
37928243      EVGSVVGYSGQKKDTEIFYQFTSFLSPGIIYHCDLTKEELE-------PRVFREVTVK--GIDASDYQTVQIFYPSKDGTKIPMFIVHKK
27065055      EVGSVVGYSGQKKDTEIFYQFTSFLSPGIIYHCDLTKEELE-------PRVFREVTVK--GIDASDYQTVQIFYPSKDGTKIPMFIVHKK
5103285       EVGSVVGYSGQKKDTEIFYQFTSFLSPGIIYHCDLTKEELE-------PRVFREVTVK--GIDASDYQTVQIFYPSKDGTKIPMFIVHKK
```

```
13786144     DVGSVVGYSGRKKDSEIFYQFTSFLSPGVIYHCDLTREELE-------PRVFREVTVK--GIDASDYQTIQVFYPSKDGTKIPMFIVHKK
5107662      EVGSVVGYSGQKKDTEIFYQFTSFLSPGIIYHCDLTKEELE-------PRVFREVTVK--GIDASDYQTVQIFYPSKDGTKIPMFIVHKK
79382269     DIGSVSDVSARRKDNTFFFSFTSFLTPGVIYKCDLANESPE-------VKVFREVTVP--GFDREAFQAIQVFYPSKDGTKIPMFIVAKK
27066372     EVGSVVGYSGQKKDTEIFYQFTSFLSPGIIYHCDLTKEELE-------PRVFREVTVK--GIDASDYQTVQIFYPSKDGTKIPMFIVHKK
19347837     DIGSVSDVSARRKDNTFFFSFTSFLTPGVIYKCDLANESPE-------VKVFREVTVP--GFDREAFQAIQVFYPSKDGTKIPMFIVAKK
109072195    EVGSIVGYSGQKKDTEIFYQFTSFLSPGIIYHCDLTKEELE-------PRVFREVTVK--GIDASDYQTVQIFYLSKDGTKIPMFIVHKK
11691900     DVGSIVGYSGQKKDTEIFYQFTSFLSPGIIYHCDLTKEELE-------PRVFREVTVK--GIDASDYQTVQIFYPSKDGTKIPMFIVHKK
41349456     DVGSIVGYSGQKKDTEIFYQFTSFLSPGIIYHCDLTKEELE-------PRVFREVTVK--GIDASDYQTVQIFYPSKDGTKIPMFIVHKK
558596       DVGSIVGYSGQKKDTEIFYQFTSFLSPGIIYHCDLTKEELE-------PRVFREVTVK--GIDASDYQTVQIFYPSKDGTKIPMFIVHKK
53127346     DVGSIVGYSGQKKDNEIFYQFTSFLSPGIIYHCDLTKEELE-------PRVFREVTVK--GFDPSVYQTIQVFYPSKDGTKIPMFIIHKK
904214       DVGSIVGYSGQKKDTEIFYQFTSFLSPGIIYHCDLTKEELE-------PRVFREVTVK--GIDASDYQTVQLFYPSKDGTKIPMFIVHKK
15217929     DIGSVCGVFARRKDTTFFFRFTSFLTPGVIYICDLSHEAPE-------VTVFREIGVP--GFDRTAFQVTQVFYPSKDGTDIPMFIVARK
9558588      DIGSVCGVFARRKDTTFFFRFTSFLTPGVIYICDLSHEAPE-------VTVFREIGVP--GFDRTAFQVTQVFYPSKDGTDIPMFIVARK
38345850     DIGSVNGISGRRDNSEVFIEFASFLTPGIIYRCDVSKETPE-------MNIYREISVG--GFDRTDFEAKQVFYPSKDGTKIPMFIVSKK
91082735     DVGTISAISGKKYHKEMFFSFCSFLTPNIIYKVDFDQGSIK-------ETLFHETKVG--DFESSKYETKQVFYKSKDGTEIPMFIINKK
54650756     DIGTIVGTSGEKKYSEIFYNFSSFLNPGSIYRYDFKTPDKS-------PSVFREIKLNLEGFRREDYAVEQIFYKSKDGTKVPMFIIRKK
73973662     EVGSIVGYSGQKKDTEIFYQFTSFLSPGIIYHCDLTKEELE-------PRVFREVTVK--GIDASDYQTVQIFYPSKDGTKIPMFIVHKK
73973660     EVGSIVGYSGQKKDTEIFYQFTSFLSPGIIYHCDLTKEELE-------PRVFREVTVK--GIDASDYQTVQIFYPSKDGTKIPMFIVHKK
47222200     KVGTVCGVSCKKKHSEFFYKFTSFTTPAIIYHCDLSELNPE-------PKVFREVEVK--GVKQEDFQTNQVFYASKDGTKIPMFLVHSQ
73973664     EVGSIVGYSGQKKDTEIFYQFTSFLSPGIIYHCDLTKEELE-------PRVFREVTVK--GIDASDYQTVQIFYPSKDGTKIPMFIVHKK
55235507     EIGTVAGFSGKKKYSEIFYHFVSFLTPGIIYHYDFEGKTEAAGGGAMEPTVFREVKI--EDFDNSRYAVEQIFYHSKDGEKVPMFIVQRK
73973658     EVGSIVGYSGQKKDTEIFYQFTSFLSPGIIYHCDLTKEELE-------PRVFREVTVK--GIDASDYQTVQIFYPSKDGTKIPMFIVHKK
71666762     PIGTITSFFSDRKKDFVSFKITSFLLPGRSFFLDINDPQSS-------LRVFKDDTVE--GLLVDDFVTEQTFYNSSDGVRIPMFIVYRK
49523575     PIGAISSLFADRKVDLVSFKVTSFLLPGRSFVLDINDPEGS-------LRVFKDDNIE--GLSADDFITEQKFYNSADGTRIPMFIIHRK
71747962     PIGAISSLFADRKVDLVSFKVTSFLLPGRSFVLDINDPEGS-------LRVFKDDNIE--GLSADDFITEQKFYNSADGTRIPMFIIHRK
33317309     PIGTITSFFSDRKKDFVSFKITSFLLPGRSFFLDINDPQSS-------LRVFKDDTVE--GLLVDDFVTEQTFYNSSDGVRIPMFIVYRK
47224662     ----------------------------IIYHCDLTKEPLQ-------PHIFREVTVK--GFNPSDYQTTQIFYPSKDGTQIPMFIVHKK
15291259     DIGTIVGTSGEKKYSEIFYNFSSFLNPGSIYRYDFKTPDKS-------PSVFREIKLNLEGFRREDYAVEQIFYKSKDGTKVPMFIIRKK
73973656     ------G---------------------IIYHCDLTKEELE-------PRVFREVTVK--GIDASDYQTVQIFYPSKDGTKIPMFIVHKK
                                                                                 :   *    * .*  :*: :   .
Prim.cons.   GVGSVVGFSGKKKDTEIFY2FTSFLTPGTIYHCDLTKEGLSAGGGAMEPRVFREVTVKLEGFDPSDYQTEQVFYPSKDGTKIPMFIVHKK


                    640       650       660       670       680       690       700       710       720
```

```
                      |         |         |         |         |         |         |         |         |
62738741     DAKG---PLPTLLYGYGGFNVALTPWF------------------------SAGFMTWID-SGGAFALANLRGGGEYGDAWHDAGRRDKK
87200538     DLAG---PLPTILYGYGGFNISVLPAF------------------------SAGRMAWLD-AGGAFAVANIRGGGEYGEAWHLAGKGPTK
103487461    GLDRS-KGSPTLLYGYGGFNVSLTPGF------------------------SPTRLAWVD-KGGVLAIANLRGGGEYGKAWHDAGRLANK
87200537     VLAAADKPLPTLLYGYGGFDISLTPAY------------------------SPVRMAWLE-AGGAFALANIRGGGEFGRSWYEAGRRENK
94496227     ALADSATAAPTILYGYGGFNISLTPGY------------------------SATRMAWLE-QGGVYAIANLRGGGEFGRAWHDAGRGANK
35211243     GTPRN-ALNPTYLYGYGGFNVSITPTF------------------------SPANLLWLE-MGGLYAVPNLRGGGEYGEDWHQAGTRLNK
32448009     DAPMD-GSNRTLLYAYGGFNISLTPSY------------------------SPGIAGWLD-AGGVYAVANLRGGGEYGREWHEAGMQLKK
75906687     GIQLD-GNNPTYLYAYGGFNVSLTPNF------------------------SVSMLVWME-MGGVYAMPNIRGGGEYGEEWHQAGMKDKK
35211187     GTPRN-ALNPTYLYGYGGFNVSITPTF------------------------SPSNLLWLE-MGGLYAVPNLRGGGEYGEDWHQAGTRLNK
67921530     GLKID-GNNPTYLYAYGGFNVSLTPNF------------------------SISNLVWME-MGGVYAVPNLRGGGEYGEEWHQAGMKDKK
67930945     GLKLD-GRNPALLYGYGGFDVPNTPAF------------------------SPTVAEWIR-MGGVFASACLRGGSEYGEEWHQAGMRAKK
17131625     GIQLN-GNNPTYLYAYGGFNVSLTPSF------------------------SVSMLVWME-MGGVYAMPNIRGGGEYGEEWHQAGMKDKK
86609965     GLTRT-GDHPTLLYGYGGFGISLTPSF------------------------SVGLVAWLE-MGGVYAQPNLRGGGEYGEEWHQAGTKLNK
83859556     DVDPN-GVRPTLIYGYGGFAITRRPDF------------------------DVRRLQWME-MGGVYAIANLRGGSAYGRDWHDAGRLANK
28808634     GTPMD-GCAPTILYGYGGFNISLTPSF------------------------NPTRAAWLE-LGGVYAVANIRGGGEYGKEWHNAGIQMQK
75854458     GTPMD-GSAPTILYGYGGFNISLTPSF------------------------SPTRAAWLE-LGGVYAVANIRGGGEYGKEWHNAGIQMQK
86607173     GLART-GDHPTLLYGYGGFGISLTPSF------------------------SVGLVAWLE-MGGVYAQPSLRGGGEYGEAWHQAGTKLNK
23128197     GIKLD-GNNPTYLYAYGGFNVSMTPSF------------------------SVSLLVWME-MGGVYAMPNIRGGGEYGEEWHQAGMKEKK
91226380     GAPMD-GSAPTILYGYGGFNISLTPSF------------------------SPTRAAWLE-LGGVYAVANIRGGGEYGKEWHNAGIQMQK
88711708     GLELN-GENPTILYGYGGFNISLTPSF------------------------SIGNAVWME-QGGIYAVPNLRGGGEYGKKWHDAGTKTKK
110167745    GVKLD-GNNPTILYGYGGFNISLTPNF------------------------SISRLIWLE-MGGVYAVPNIRGGGEYGEGWHQAGIKQQK
109899092    GLKLD-GTNPTMLYGYGGFNISLQPAF------------------------SSVNAAWLE-QGGVYAVPNLRGGGEYGKAWHDAGTKLQK
86144168     GLERN-GKNPTILYGYGGFNISLTPSF------------------------SITNAIWME-MGGVYAVPNLRGGGEYGKEWHDAGTKMQK
4973227      GLKLD-GSNPTILYGYGGFDVSLTPSF------------------------SVSVANWLD-LGGVYAVANLRGGGEYGQAWHLAGTQQNK
6048357      GLKLD-GSNPTILYGYGGFDVSLTPSF------------------------SVSVANWLD-LGGVYAVANLRGGGEYGQAWHLAGTQQNK
76875127     GIKLD-GSNPTILYGYGGFNISLTPSF------------------------SPTTAAWLE-QGGVYAVANIRGGGEYGKEWHNAGTKLQK
216201       GLKLD-GSNPTILYGYGGFDVSLTPSF------------------------SVSVANWLD-LGGVYAVANLRGGGEYGQAWHLAGTRMNK
88804657     GLEKN-GDNPTILYGYGGFNVSLTPGF------------------------STSNAVWLE-QGGIYAVPNIRGGGEYGKKWHDAGTKTSK
89517967     GLVLD-GKNPTILYGYGGFNVSLNPSF------------------------SVSRAAWME-MGGIYAVANLRGGGEYGKKWHDAGTKMQK
90588230     GTKLD-GKNPTILYGYGGFNISLTPSF------------------------SIANAVWME-NGGVYAVANLRGGGEYGKKWHDAGTKLQK
71143549     GIELN-GKNPTILYGYGGFNISLTPRF------------------------SVSRAVWLE-QGGIYAVANLRGGGEYGKTWHKAGTQLDK
83855898     GTELN-GKNPTILYGYGGFNISLTPSF------------------------SIANAVWME-QGGIYAVPNLRGGGEYGKAWHDAGTKLQK
85710915     GTELD-GSNPTILYGYGGFNISLTPSF------------------------SIANAAWLE-MGGVYAVANLRGGGEYGKAWHDAGTKMQK
78368715     GLTLN-GENPTLLYAYGGFSISLTPRF------------------------SPANIAWMD-MGGVYAVPNLRGGAEYGESWHQAGMFDKK
56178266     GTKLD-GSNPTILYGYGGFNISLTPSF------------------------SIANAAWLE-MGGVYAVANLRGGGEYGKDWHKAGTKMQK
```

```
91215481     GIELD-GTHPTMLYGYGGFNISLTPSF------------------------STANTVWLE-NGGVYAVANLRGGGEYGKEWHNAGIKMKK
78688320     GLVKN-GQNPTLLYAYGGFAISMTPRF------------------------SPANIAWLD-MGGIYAVPSLRGGADYGESWHQAGMFDKK
24374294     GLVRN-GQNPTLLYAYGGFAISMTPRF------------------------SPANIAWLD-MGGIYAVPSLRGGADYGESWHQAGMFDKK
456523       GLKKD-GKNPTILYSYGGFNISLQPAF------------------------SVVNAIWME-NGGIYAVPNIRGGGEYGKKWHDAGTKMQK
77816782     GLVKN-GQNPTLLYAYGGFAISMTPRF------------------------SPANIAWLD-MGGIYAVPSLRGGAEYGESWHQAGMFDKK
68547790     GLKKD-GKNPTLLYAYGGFSISMTPRF------------------------SPATIAWLD-MGGIYAVPALRGGSEYGEEWHKAGMFGKK
82744257     GLVKN-GQNPTLLYAYGGFAISMTPRF------------------------SPANIAWLD-MGGIYAVPSLRGGAEYGESWHQAGMFDKK
86131249     GTELN-GKNPTILYAYGGFNVSLTPSF------------------------SIANAVWME-QGGVYAVPNLRGGGEYGKKWHDAGTKMQK
88796133     GLKLD-GNNKTLLYGYGGFNISLKPGY------------------------SVSRMVWVE-QGNVLAIANLRGGGEYGQQWHKAGTKLNK
42526704     DIKLD-GSNPTIMYGYGGFAISLPPAF------------------------SAARMAFLE-KGGIFACVNLRGGLEYGEAWHSAGKKMKK
27367204     GIELN-GKNPTILYGYGGFDVSLTPAF------------------------SGMVASWLE-LGGVYAVANLRGGGEYGKAWHNAGTQLQK
37676919     GIELN-GKNPTILYGYGGFDVSLTPAF------------------------SGMVASWLE-LGGVYAVANLRGGGEYGKAWHNAGTQLQK
75855979     GIALD-GSAPTMLYGYGGFNISLTPMF------------------------SGNVANWLE-LGGIYAVANMRGGGEYGKAWHNAGTQQQK
28809898     GISLD-GSAPTMLYGYGGFNISLTPMF------------------------SGNVANWLE-LGGIYAVANMRGGGEYGKAWHNAGTQQQK
6460324      DAPRD-GSNRTLLYGYGGFSISLTPAF------------------------SASRLAWLE-RGGVFVQANLRGGGEYGEAWHEAGTLGRK
91223249     GIALD-GRAPTMLYGYGGFNISLTPMF------------------------SGNVSNWLE-LGGIYAVANMRGGGEYGKAWHNAGTQQQK
56180108     NIELN-GENPTLLYGYGGFNISITPDY------------------------DTTRLVWLE-MGGVLAIPNLRGGGEFGEMWHQQGTKENK
55769741     DIELDG-SHPTLLYGYGGFNISLTPSF------------------------SVSRVVLCKNMGFVVCVANIRGGGEYGEEWHKAGARAMK
47507507     GIALDN-SHPAFLYGYGGFNISITPSY------------------------SVSRLIFMRHLGGILAVANIRGGGEYGETWHKAGSLGNK
28502989     GIALDN-SHPAFLYGYGGFNISITPSY------------------------SVSRLIFMRHLGGILAVANIRGGGEYGETWHKAGSLGNK
86144914     GLVLD-GNNPTMLYAYGGFNVSLTPSF------------------------SGTVGSWLE-LGGVYAVPNLRGGGEYGKAWHKAGTQQQK
84386759     GLVLD-GNNPTMLYAYGGFNVSLTPSF------------------------SGTVGSWLE-LGGVYAVPNLRGGGEYGKAWHKAGTQQQK
63100723     GIKMDG-SHPAFLYGYGGFNISITPSY------------------------SVSRLIFIRHLGGVLAVANIRGGGEYGETWHKGGMLANK
38649021     GITLDK-SHPAFLYGYGGFNISITPSY------------------------SVSRLIFVRHLGGILAVANIRGGGEYGETWHKAGILGNK
59711729     GIELD-GSNPTILYGYGGFNVSLTPNF------------------------SGVMASWLE-LGGVYAIPNIRGGGEYGKEWHKAGTQQQK
6755152      GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGVLAVANIRGGGEYGETWHKGGILANK
74215349     GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGVLAVANIRGGGEYGETWHKGGILANK
50514023     GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGVLAVANIRGGGEYGETWHKGGILANK
26345256     GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGVLAVANIRGGGEYGETWHKGGILANK
157879458    GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGVLAVANIRGGGEYGETWHKGGILANK
51592147     GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGVLAVANIRGGGEYGETWHKGGILANK
10835490     GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGVLAVANIRGGGEYGETWHKGGILANK
50514022     GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGVLAVANIRGGGEYGETWHKGGILANK
37928243     GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGVLAVANIRGGGEYGETWHKGGILANK
27065055     GIKLDG-SHPAFLYGFGGFNISITPNY------------------------SVSRLIFVRHMGGVLAVANIRGGGEYGETWHKGGILANK
5103285      GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVCRLIFVRHMGGVLAVANIRGGGEYGETWHKGGILANK
```

```
13786144        GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGVLAVANIRGGGEYGETWHKGGILANK
5107662         GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGVLAVANIRGGGEYGETWHKGGILANK
79382269        DIKLDG-SHPCLLYAYGGFNISITPSF------------------------SASRIVLSKHLGVVFCFANIRGGGEYGEEWHKAGSLAKK
27066372        GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGVLAVANIRGGGEYGETWHKGGILANK
19347837        DIKLDG-SHPCLLYAYGGFNISITPSF------------------------SASRIVLSKHLGVVFCFANIRGGGEYGEEWHKAGSLAKK
109072195       GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGILAVANIRGGGEYGETWHKGGILANK
11691900        SIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGILAVANIRGGGEYGETWHKGGILANK
41349456        GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGILAVANIRGGGEYGETWHKGGILANK
558596          GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGILAVANIRGGGEYGETWHKGGILANK
53127346        GIKLDG-SHPAFLYGYGGFNISITPSY------------------------SVSRLIFVRHLGGVLAVANIRGGGEYGETWHKGGILANK
904214          SIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGILAVANIRGGGEYGETWHKGGILANK
15217929        DIKLDG-SHPCLLYAYGGFSISMTPFF------------------------SATRIVLGRHLGTVFCFANIRGGGEYGEEWHKSGALANK
9558588         DIKLDG-SHPCLLYAYGGFSISMTPFF------------------------SATRIVLGRHLGTVFCFANIRGGGEYGEEWHKSGALANK
38345850        SIVLDG-SHPTLLYGYGGFGMNMTPHF------------------------SVTRIVLMRNLGFVSCIANIRGGGEYGEDWHKAGSLANK
91082735        GLVNDG-SKPCLLYGYGGFNVNLTPSF------------------------GVSRLVFIENFDGVYALANIRGGGEYGDNWHNGGRFGKK
54650756        RDSVE--PRPCLLYGYGGFNISMLPSF------------------------GLSGLMFIDTFDGVLAYPNLRGGGEYGEKWHNGGRLLNK
73973662        GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGVLAVANIRGGGEYGETWHKGGILANK
73973660        GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGVLAVANIRGGGEYGETWHKGGILANK
47222200        GLKKDG-SHPVFLYGYGGFEASIQPYYNHVLTQLTSSTFTGFDRSSWLPVTSVAYLLFVRHLGGILAVANIRGGGEYGLTWHKGGTLGNK
73973664        GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGVLAVANIRGGGEYGETWHKGGILANK
55235507        QEKKE--HKPCLLYGYGGFNICVQPSF------------------------SITGLVFIDSFDGILAYPNIRGGGEYGERWHNAGRLLKK
73973658        GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGVLAVANIRGGGEYGETWHKGGILANK
71666762        GSVSS--ESPLLLYGYGGFNIPLTPAF------------------------SSSRMVFLRDLGGVLAVPNIRGGGEYGEEWHDAGRRVCK
49523575        GIVTS--ESPVLLYGYGGFNISLTPSF------------------------SSSRVVFLQHLRGVLAVPNIRGGGEYGEAWHNAGRLTSK
71747962        GIVTS--ESPVLLYGYGGFNISLTPSF------------------------SSSRVVFLQHLRGVLAVPNIRGGGEYGEAWHNAGRLTSK
33317309        GSVSS--ESPLLLYGYGGFNIPLTPAF------------------------SSSRMVFLRDLGGVLAVLNIRGGGEYGEEWHDAGRRACK
47224662        GIELDG-SHPGFLYGYGGFNISITPSY------------------------SVSRLIFVCHLGGVLAVANIRGGGEYGETWHKAGMLANK
15291259        RDSVE--PRPCLLYGYGGFNISMLPSF------------------------GLSGLMFIDTFDGVLAYPNLRGGGEYGEKWHNGGRLLNK
73973656        GIKLDG-SHPAFLYGYGGFNISITPNY------------------------SVSRLIFVRHMGGVLAVANIRGGGEYGETWHKGGILANK
                            :*.:***     * :                        .                  :***  :*  *:  *    *
Prim.cons.      GIKLDGGS2PT2LYGYGGFNISLTPSFNHVLTQLTSSTFTGFDRSSWLPVTSVSRLIWLEHMGGVYAVANIRGGGEYGETWHKAGTLANK


                       730       740       750       760       770       780       790       800       810
```

```
                      |         |         |         |         |         |         |         |         |
62738741     QNVFDDFIAAGEWLIANGVTPRHGLAIEGGSNGGLLIGAVTNQRPDLFAAASPAVGVMDMLRFDQFTAG-RYWVDDYGYPEK-EADWRVL
87200538     QNVFDDFIAAGEWLKANGVTSANGLAVEGGSNGGLLVGAVVNQRPDLFAAAVPAVGVMDMLRFDKFTAG-REWVFDYGYPEK-EEDWRRL
103487461    QNVFDDFIAAGEYLIAEGITGKGQLAIEGGSNGGLLVGAVTNQRPDLFAAALPAVGVMDMLRFDRFTAG-RYWVDDYGYPSK-EADFRNL
87200537     QNSFDDFIAAGEFLIREGIAGKGQLAIQGASNGGLLVGAVVNQRPDLFAAANPDVGVMDMLRFDRFTSG-RFWVDDYGRPDR-EEDWRTL
94496227     QNVFDDFIAAAEYLKANGFTPPDGLAIEGRSNGGLLVGAVVNQRPDLFAAALPAVGVMDMLRFDRFTAG-RYWTDDYGSPAD-RRAFPLL
35211243     QNVFDDFLAAAEYLIANKYTSPEKLAIGGGSNGGLLVGAAMTQRPELFAAALPAVGVMDMLRFEQFTIG-WAWVSDYGSSQD-PEQFQAL
32448009     QNVFDDFIAAAEHLIDMGLTSRERLGVRGGSNGGLLIGAVMTQRPDLFGACLPAVGVMDMLRYHKFTIG-WAWVSEFGSSDD-ETQIDNL
75906687     QNVFDDFIAAAEWLMANNYTKPEKLAIAGGSNGGLLVGACMTQRPELFGAALPAVGVMDMLRFHKFTIG-WAWTAEYGSPDN-PQEFPAI
35211187     QNVFDDFLAAAEYLIANKYTSPEKLAIGGGSNGGLLVGAAMTQRPELFAAALPAVGVMDMLRFEQFTIG-WAWVSDYGSSQD-PEQFQAL
67921530     QNVFDDFIAAAEWLIKNNYTSSQKLAIGGGSNGGLLVGACLTQKPKLYAAALPAVGVLDMLRFHQFTIG-WAWCPEYGSSEN-EEEFKTL
67930945     QNVFNDFIAAAEWLVANRYTSTPKLAIMGGSNGGLLVGAVLNQRPDLFGAAVAQVGVMDMLRFQKFGFG-TQWVGEYGSPEN-PEDFKVL
17131625     QNVFDDFIAAAEWLIANNYTKPEKLAIAGGSNGGLLVGACMTQRPELFGAALPAVGVMDMLRFHKFTIG-WAWTAEYGSPDN-PQEFPAL
86609965     QKVFDDFIAAAEWLIAHGYTNPAKLAISGGSNGGLLVGACLVQRPELFAAALPAVGVFDMLRFHKFTIG-WAWISEYGSPEN-PEEFKAL
83859556     QNVFDDFINAGHALIDMGWTSPDHLAIHGRSNGGLLVGAVANQAPDLFAAALPAVGVMDMLRFNQFTAG-RFWVDDYGSPQD-PEMFDVL
28808634     QNVFDDFIAAAEYLIDEKVTSSNKLAINGGSNGGLLVGAVMTQRPELFKVALPAVGVLDMLRYHTFTAG-AGWAYDYGTAEQSKEMFQYL
75854458     QNVFDDFIAAAEYLIDEKVTSSNKLAINGGSNGGLLVGAVMTQRPELFKVALPAVGVLDMLRYHTFTAG-AGWAYDYGTAEQSKEMFQYL
86607173     QKVFDDFIAAAEWLMANGYTNPSKLAIAGGSNGGLLVGACLTQRPDLFAAALPAVGVFDMLRFHKFTIG-WAWISEYGSPED-PEEFKAL
23128197     QNVFDDFIGAAEWLIANKYTKTEKLAIAGGSNGGLLVGACITQRPDLFGAAIPSVGVMDMLRFHKFTIG-WAWTSEFGSADN-SEEFPAL
91226380     QNVFDDFIAAAEYLIDEKVTSSNKLAINGGSNGGLLVGAVMTQRPELFKVALPAVGVLDMLRYHTFTAG-AGWAYDYGTAEQSKEMFQYL
88711708     QNVFDDFIAAGEYLIANKYTSKDYLAIRGGSNGGLLVGATMTQRPDLIKVALPAVGVLDMLRYHTFTAG-AGWAYDYGTAEESEEMFNYL
110167745    QNVFDDFISAAEWLIENNWSSSPKLAITGASNGGLLVGACITQRPELFGAALPAVGVMDMLRFHKFTIG-WAWTAEYGSPDD-PEEFKAL
109899092    QNVFDDFIAAAEYLIEQNYTSPDYLAIRGGSNGGLLVGATMLQRPDLFKVALPAVGVLDMLRYHTFTAG-AGWAYDYGTSEDSQEMFDYL
86144168     QNVFDDFIAAAEYLISENYTSSDYLAVRGGSNGGLLVGATMTQRPDLMKVALPAVGVMDMLRYHTFTAG-AGWAYDYGTSEDSKEMFEYI
4973227      QNVFDDFIAAAEYLKAEGYTRTDRLAIRGGSNGGLLVGAVMTQRPDLMRVALPAVGVLDMLRYHTFTAG-AGWAYDYGTSADSEAMFDYL
6048357      QNVFDDFIAAAEYLKAEGYTRTDRLAIRGGSNGGLLVGAVMTQRPDLMRVALPAVGVLDMLRYHTFTAG-AGWAYDYGTSADSEAMFDYL
76875127     QNVFDDFIAAAEYLQSKKYTSKQRLALRGGSNGGLLVGAVMTQRPDLFQVALPAVGVLDMLRYHTFTAG-AGWAYDYGTSEQSKEMFTYI
216201       QNVFDDFIAAAEYLKAEGYTRTDRLAIRGGSNGGLLVGAVMTQRPDLMRVACQAVGVLDMLRYHTFTAG-AGWAYDYGTSADSEAMFDYL
88804657     QNVFDDFIAAAEYLIAEGYTSAGKLAIRGGSNGGLLVGAVMTQRPDLMQVALPAVGVLDMLRYHTFTAG-AGWSYDYGTSEESEEMFEYL
89517967     QNVFDDFIAAGEWLKDNDYTSTKYLAIQGGSNGGLLVGATMTQRPDLASVAFPAVGVLDMLRYNKFTAG-AGWAYDYGTAEDNPEMFEYL
90588230     QNVFDDFIAAAEYLIAQKYTSSDYLAIRGGSNGGLLVGATMTQRPDLMKVALPAVGVMDMLRYHTFTAG-AGWAYDYGTAQDSKEMFEYL
71143549     QNVFDDFIAAGEYLIKQDYTSSDYLAINGGSNGGLLVGAVMTQRPDLMQVALPAVGVLDMLRYHTFTAG-AGWAYDYGTAEQSEEMFTYL
83855898     QNVFNDFIAAAEYLIEKNYTSKEYLAIRGGSNGGLLVGATMTQRPDLMQVALPAVGVMDMLRYHTFTAG-AGWAYDYGTAEDSDEMFQYL
85710915     QNVFDDFIAAAEYLIDEDYTSKEKLAIRGGSNGGLLVGATMTQRPDLFAVALPAVGVLDMLRYHTFTAG-AGWAYDYGTAEDSKEMFQYL
78368715     QNVFDDYFAAAEYLIDNKYTNQSKLGAYGRSNGGLLMGAAVTQRPDLFAAILPAVGVLDMLRFHKFTIG-WAWTSEYGSADN-AEQFPAL
56178266     QNVFDDFIAAAEYLIEKDYTSPERLAIRGGSNGGLLVGAVMTQRPELFAVALPAVGVLDMLRYHTFTAG-AGWAYDYGTANDSKEMFEYL
```

```
91215481      QNVFDDFIAAAEYLIAENYTSSNKLAIRGGSNGGLLVGATMTQRPDLMQVALPAVGVLDMLRYHTFTSG-AGWAYDYGTSQDSSEMFEYL
78688320      QNVFDDYFAAAEYLINEKYTNSTKLGAYGRSNGGLLMGAAVTQRPELFAAVLPAVGVLDMLRFHKFTIG-WAWTSEYGSADN-AEQFPAL
24374294      QNVFDDYFAAAEYLINEKYTNSTKLGAYGRSNGGLLMGAALTQRPELFAAVLPAVGVLDMLRFHKFTIG-WAWTSEYGSADN-AEQFPAL
456523        KNVFNDFIAAGEYLQKNGYTSKEYMALSGRSNGGLLVGATMTMRPDLAKVAFPGVGVLDMLRYNKFTAG-AGWAYDYGTAEDSKEMFEYL
77816782      QNVFDDYFAAAEYLINEKYTNTSKLGAYGRSNGGLLMGAAVTQRPELFAAVLPAVGVLDMLRFHKFTIG-WAWTSEYGSADN-AEQFPAL
68547790      QNVFDDYFAAAEYLISENYTNSSKLGAYGRSNGGLLMGAAVTQRPDLFAAILPAVGVLDMLRFQKFTIG-WAWTAEYGSADN-AEDFPYL
82744257      QNVFDDYFAAAEYLINEKYTNTSKLGAYGRSNGGLLMGAAVTQRPELFAAVLPAVGVLDMLRFHKFTIG-WAWTSEYGSADN-AEQFPAL
86131249      QNVFDDFIAAAEYLIDNKYTSKKYLAIRGGSNGGLLVGATMTQRPDLMQVALPAVGVMDMLRYHTFTAG-AGWAYDYGTAEDSKEMFEYL
88796133      QNVFDDFIAAGEYLVESGYTNPQKMGIQGGSNGGLLVGATLTQRPDLFAAALPAVGVLDMLRYHTPSANARAWSSDYGLSEN-NDEFEAL
42526704      QNVFDDFIAAGEYLIEHKYTSSKKLAIQGGSNGGLLIGAVTNQRPDLFAVAIPQVGVLDMLRYQHFTIG-WAWVDEYGSSEDSKEMFEYL
27367204      QNVFDDFVAAAEYLIEQQYTSSDYLAIRGGSNGGLLVGACMTQRPELFKVALPAVGVLDMLRYHTFTSG-EGWKYDYGTSAQSEEMFQYL
37676919      QNVFDDFIAAAEYLIEQQYTSSDFLAIRGGSNGGLLVGACMTQRPDLFKVALPAVGVLDMLRYHTFTSG-EGWKYDYGTSAQSEEMFQYL
75855979      QNVFDDFIAAAEYLIENDYTSSDRLAIRGGSNGGLLVGACMTQRPELFKVALPAVGVLDMLRYHTFTSG-EGWKYDYGTSEQSEAMFQYL
28809898      QNVFDDFIAAAEYLIENDYTSSERLAIRGGSNGGLLVGACMTQRPELFKVALPAVGVLDMLRYHTFTSG-EGWKYDYGTSEQSEAMFQYL
6460324       QNVFDDFAACAEHLARQGWTRPERLAIQGGSNGGLLVGASITQRPELFGAAVAQVGVLDMLRYHLFTIG-WAWASDYGRSDD-PEMFATL
91223249      QNVFDDFIAAAEYLIENDYTSSDRLAIRGGSNGGLLVGACMTQRPELFKVALPAVGVLDMLRYHTFTSG-EGWKYDYGTSSQSEAMFQYL
56180108      QNVFDDFIAAAEWLINNNYTNKNKLAIQGRSNGGLLVGATMTQRPELFAAALPAVGVMDMLRYQLPSANARGWGSEFGLSEN-KKDFKTL
55769741      QNCFDDFIACAELLISAGYTSYRQLCIEGGSNGGLLIAACVNQRPDLFGCALAHVGVMDMLRFHKFTIG-HAWTTDYGCSDN-EEEFHWL
47507507      QNCFDDFQCAAEYLVKEGYSSAKNITINGGSNGGLLVAACTNQRPDLFGCTIAQVGVMDMLKFHKFTIG-HAWTTDFGCSDN-KEHFDWL
28502989      QNCFDDFQCAAEYLVKEGYSSAKNITINGGSNGGLLVAACTNQRPDLFGCTIAQVGVMDMLKFHKFTIG-HAWTTDFGCSDN-KEHFDWL
86144914      QNVFDDFIAAAEFLIEENYTSSDKLAIRGGSNGGLLVGACMTQRPELFQVALPAVGVLDMLRYHTFTSG-EGWAYDYGTSAQNKEMFEYL
84386759      QNVFDDFIAAAEFLIEENYTCSDTLAIRGGSNGGLLVGACMTQRPELFQVALPAVGVLDMLRYHTFTSG-EGWAYDYGTSAQNKEMFEYL
63100723      QNCFTDFQCAAEYLIKEGYTSPKKLTINGGSNGGLLVAACVNQRPDLFGCAVAQVGVMDMLKFHKFTIG-HAWTTDFGCSEI-KEQFDWL
38649021      QNCFDDFQCAAEYLVKEGYASAEKITINGGSNGGLLVAACANQRPELFGCAIAQVGVMDMLKFHKFTIG-HAWTTDYGCSDN-KEHFEWL
59711729      QNVFDDFIAAAEFLIAEKYTDSSRLAIRGGSNGGLLVGACMTQRPELFKVALPAVGVLDMLRYHTFTSG-EGWAYDYGTSEQSEEMFNYL
6755152       QNCFDDFQCAAEYLIKEGYTSPKRLTINGGSNGGLLVAACANQRPDLFGCVIAQVGVMDMLKFHKFTIG-HAWTTDYGCSDT-KQHFEWL
74215349      QNCFDDFQCAAEYLIKEGYTSPKRLTINGGSNGGLLVAACANQRPDLFGCVIAQVGVMDMLKFHKFTIG-HAWTTDYGCSDT-KQHFEWL
50514023      QNCFDDFQCAAEYLIKEGYTSPKRLTINGGSNGGLLVATCANQRPDLFGCVIAQVGVMDMLKFHKYTIG-HAWTCDYGCSDS-KQHFEWL
26345256      QNCFDDFQCAAEYLIKEGYTSPKRLTINGGSNGGLLVAACANQRPDLFGCVIAQVGVMDMLKFHKFTIG-HAWTTDYGCSDT-KQHFEWL
157879458     QNCFDDFQCAAEYLIKEGYTSPKRLTINGGSNGGLLVATCANQRPDLFGCVIAQVGVMDMLKFHKYTIG-HAWTTDYGCSDS-KQHFEWL
51592147      QNCFDDFQCAAEYLIKEGYTSPKRLTINGGSNGGLLVATCANQRPDLFGCVIAQVGVMDMLKFHKYTIG-HAWTTDYGCSDS-KQHFEWL
10835490      QNCFDDFQCAAEYLIKEGYTSPKRLTINGGSNGGLLVATCANQRPDLFGCVIAQVGVMDMLKFHKYTIG-HAWTTDYGCSDS-KQHFEWL
50514022      QNCFDDFQCAAEYLIKEGYTSPKRLTINGGSNGGLLVATCANQRPDLFGCVIAQVGVMDMLKFHKYTIG-HAWTTDYGCSDS-KQHFEWL
37928243      QNCFDDFQCAAEYLIKEGYTSPKRLTINGGANGGLLVATCANQRPDLFGCVIAQVGVMDMLKFHKYTIG-HAWTTDYGCSDS-KQHFEWL
27065055      QNCFDDFQCAAEYLIKEGYTSPKRLTINGGSNGGLLVATCANQRPDLFGCVIAQVGVMDMLKFHKYTIG-HAWTTDYGCSDS-KQHFEWL
5103285       QNCFDDFQCAAEYLIKEGYTSPKRLTINGGSNGGLLVATCANQRPDLFGCVIAQVGVMDMLKFHKYTIG-HAWTTDYGCSDN-KQHFEWL
```

```
13786144        QNCFDDFQCAAEYLIKEGYTTSKRLTINGGSNGGLLVAACANQRPDLFGCVIAQVGVMDMLKFHKFTIG-HAWTTDYGCSDS-KQHFEWL
5107662         QNCFDDFQCAAEYLIKEGYTSPKRLTINGGSNGGLLVATCANQRPDLFGCVIAQVGVMDMLKFHKYTIG-HAWTTDYGXSDS-KQHFEWL
79382269        QNCFDDFISGAEYLVSAGYTQPSKLCIEGGSNGGLLVGACINQRPDLYGCALAHVGVMDMLRFHKFTIG-HAWTSDYGCSEN-EEEFHWL
27066372        QNCFDDFQCAAEYLIKEGYTSPKRLTINGGSNGGLLVATCANQRPDLFGCVIAQVGVMDMLKFHKYTIG-HAWTTDYGCSDS-KQHFEWL
19347837        QNCFDDFISGAEYLVSAGYTQPSKLCIEGGSNGGLLVGACINQRPDLYGCALAHVGVMDMLRFHKFTIG-HAWTSDYGCSEN-EEEFHWL
109072195       QNCFDDFQCAAEYLIKEGYTSPKRLTINGGSNGGLLVAACANQRPDLFGCVIAQVGVMDMLKFHKYTIG-HAWTTDYGCSDS-KQHFEWL
11691900        QNCFDDFQCAAEYLIKEGYTSPKRLTINGGSNGGLLVAACANQRPDLFGCVIAQVGVMDMLKFHKYTIG-HAWTTDYGCSDS-KQHFEWL
41349456        QNCFDDFQCAAEYLIKEGYTSPKRLTINGGSNGGLLVAACANQRPDLFGCVIAQVGVMDMLKFHKYTIG-HAWTTDYGCSDS-KQHFEWL
558596          QNCFDDFQCAAEYLIKEGYTSPKRLTINGGSNGGLLVAACANQRPDLFGCVIAQVGVMDMLKFHKYTIG-HAWTTDYGCSDS-KQHFEWL
53127346        QNCFDDFQYAAKYLIREGYTAPKKLTINGGSNGGLLVAACANQRPDLFGCVIAQVGAMDMLKFHKYTIG-HAWTTDYGCSDH-KEQFEWL
904214          QNCFDDFQCAAEYLIKEGYTSPKRLTINGGSNGGLLVAACANQRPDLFGCVIAQVGVMDMLKFHKYTIG-HAWTTDYGCSDS-KQHFEWL
15217929        QNCFDDFISGAEYLVSAGYTQPRKLCIEGGSNGGILVGACINQRPDLFGCALAHVGVMDMLRFHKFTIG-HAWTSEFGCSDK-EEEFHWL
9558588         QNCFDDFISGAEYLVSAGYTQPRKLCIEGGSNGGILVGACINQRPDLFGCALAHVGVMDMLRFHKFTIG-HAWTSEFGCSDK-EEEFHWL
38345850        QNCFDDFIAAGEFLVSAGYTNPSRLCVEGASNGGLLVAACINQRPDLFGCALAHVGVMDMLRFHKFTIG-RAWTCDFGCSEK-EEEFHWL
91082735        QNCFDDFQYAAKYLVENKYTKVDKLIIQGGSNGGLLVAACINQAPELFGAAICQVGVLDMLRFHKFTIG-YAWKSDYGCSEE-QEQFEYL
54650756        QNVFDDFQAAAEYLIENKYTTKDRLAIQGGSNGGLLVGSCINQRPDLFGAAVAQVGVMDMLRFHKFTIG-HAWCSDYGNPSE-KEHFDNL
73973662        QNCFDDFQHAAEYLIKEGYTSPKRLTINGGSNGGLLVAACANQRPDLFGCVIAQVGVMDMLKFHKYTIG-HAWTTDYGCSDS-KQHFEWL
73973660        QNCFDDFQHAAEYLIKEGYTSPKRLTINGGSNGGLLVAACANQRPDLFGCVIAQVGVMDMLKFHKYTIG-HAWTTDYGCSDS-KQHFEWL
47222200        QKCFDDFLCAAEYLIQEKYTVASRIAINGASNGGLLVAACVNQRPDLFGCAVAEVGVMDMLKFHKFTIG-HAWTTDYGCSDN-PEQFEWL
73973664        QNCFDDFQHAAEYLIKEGYTSPKRLTINGGSNGGLLVAACANQRPDLFGCVIAQVGVMDMLKFHKYTIG-HAWTTDYGCSDS-KQHFEWL
55235507        QNVFDDFQHAAQYLVEHGYTTHDKIAIQGGSNGGLLVGACINQRPDLFGAAIAQVGVMDMLRFHKFTIG-RAWVSDYGDMNE-KEHFENL
73973658        QNCFDDFQHAAEYLIKEGYTSPKRLTINGGSNGGLLVAACANQRPDLFGCVIAQVGVMDMLKFHKYTIG-HAWTTDYGCSDS-KQHFEWL
71666762        QNCFTDFIEGAKFLHRQGYGSPQTTAIMGGSNGGLLVAAVANQAPELFRCVVCQVGVLDMYKFHKFTIG-HAWKSDYGDPEK-EEDFRVL
49523575        QNCFTDFIEAAKFLHGNGYGNPATTAIMGGSNGGLLVAAVANQAPELFRCVICQVGVLDMYKFHKFTIG-HAWRSDFGDPDK-EEDFKVL
71747962        QNCFTDFIEAAKFLHGNGYGNPATTAIMGGSNGGLLVAAVANQAPELFRCVICQVGVLDMYKFHKFTIG-HAWRSDFGDPDK-EEDFKVL
33317309        QNCFTDFIEGAKFLHRQGYGSPQTTAIMGGSNGGLLVAAVANQAPELFRCVVCRVGVLDMYKFHKFTIG-HAWKSDYGDPEK-EEDFRVL
47224662        QNCFTDFQCAAEYLIKEGYTSPSKLTINGGSNGGLLVAACVNQRPELFGCAVAQVGVMDMLKFHKFTIG-HAWTTDFGCSDI-KEQFDCL
15291259        QNVFDDFQAAAEYLIENKYTTKDRLAIQGGSNGGLLVGSCINQRPDLFGAAVAQVGVMDMLRFHKFTIG-HAWCSDYGNPSE-KEHFDNL
73973656        QNCFDDFQHAAEYLIKEGYTSPKRLTINGGSNGGLLVAACANQRPDLFGCVIAQVGVMDMLKFHKYTIG-HAWTTDYGCSDS-KQHFEWL
                :: * *:   .. *              * :***:*:.:     *.*       **.:** ::.    .   *  ::*           :
Prim.cons.      QNVFDDFIAAAEYLIKEGYTSPKRLAINGGSNGGLLVGACMNQRPDLFGCALPAVGVMDMLRFHKFTIGAHAWTTDYGCSD2SKEMFEWL


                       820       830       840       850       860       870       880       890       900
```

```
                       |         |         |         |         |         |         |         |         |
62738741      RRYSPYHNVRSG--------VDYPAILVTTADTDDRVVPGHSFKYTAALQTAAIGPKP-------HLIRIETRAGHGSGKPIDKQ-----
87200538      RAYSPYHNIASG--------KDYPAILVTTADTDDRVVPGHSFKYAAALQAASIGSKP-------HLIRIETRAGHGSGKPVAKL-----
103487461     LSYSPYHNIRSG--------VAYPAVLVTTADTDDRVVPGHSFKYTAALQHAKAGSKP-------HLIRIETRAGHGSGKPTDKI-----
87200537      RAYSPYHNIATG--------KPYPAILVTTADNDDRVVPAHSFKYVAALQAGDIGEKP-------HLLRVESRAGHGAGKPVDKV-----
94496227      YGYSPYHNIAGG--------SDYPAILVSTADTDDRVVPAHSFKYAAALQAADLGARP-------RLLRVESRAGHGAGKPVDKL-----
35211243      YAYSPLHNLKAG--------TRYPATLVTTADTDDRVVPGHSYKFTAALQAAQAGEGP-------VLIRIETKAGHGAGKPTTKL-----
32448009      LSYSPLHNLKPG--------TCYPATMVTTADRDDRVVPGHSFKFAAALQAAQSCDNP-------TLIRIETRAGHGAGTPTSKK-----
75906687      YAYSPLHNLKSG--------TAYPATLITTADHDDRVVPAHSFKFAAALQTAHNGNAP-------VLIRIETKAGHGAGKPTAKI-----
35211187      YAYSPLHNLKAG--------TRYPATLVTTADTDDRVVPGHSYKFTAALQAAQAGEGP-------VLIRIETKAGHGAGKPTTKL-----
67921530      LAYSPLHNLTAN--------TAYPATMITTADHDDRVVPAHSFKFAAALQNAHDGEKP-------VLIRIETKAGHGAGKPTTKL-----
67930945      RAYSPLHNIRAG--------TEYPAVLVTTSDHDDRVMPGHSLKYTATLQQAQKGPAP-------ILLRVETRAGHGAGKPTAKQ-----
17131625      YAYSPLHNLKPG--------TAYPATLITTADHDDRVVPAHSFKFAAALQTAHNGNAP-------VLIRIETKAGHGAGKPTAKI-----
86609965      YAYSPLHNLKPG--------TAYPATLITTADHDDRVVPAHSFKFAAALQVAQGGSQP-------ILIRIDTKAGHGAGKPTSKL-----
83859556      YGYSPYHNIPET--------GEYPATLITTADTDDRVVPGHSFKYAAALQAAQTGDAP-------TLIRIETRAGHGAGTPVSKL-----
28808634      KGYSPVHNVKAG--------VEYPATLITTGDHDDRVVPAHSYKFAAELQSKQAGSNP-------TLIRIETNAGHGAGTPTRKI-----
75854458      KGYSPVHNVKAG--------VEYPATLITTGDHDDRVVPAHSYKFAAELQSKQAGSNP-------TLIRIETNAGHGAGTPTRKI-----
86607173      YAYSPLHNLKPG--------TAYPATLITTADHDDRVVPAHSFKFAAALQAAQGGSQP-------ILIRIDTKAGHGAGKPTAKL-----
23128197      YAYSPLHNIKPD--------TAYPATLITTADHDDRVVPAHSFKFAAALQEAHAGDAP-------VLIRIETKAGHGAGKPTAKI-----
91226380      KGYSPVHNVKAG--------VEYPATLITTGDHDDRVVPAHSYKFAAELQSKQAGSNP-------TLIRIETNAGHGAGTPTRKI-----
88711708      KGYSPVHNVKEG--------TEYPATLVTTGDHDDRVVPAHSFKFAAELQSKQAGPNP-------TLIRIETNAGHGAGTPVSKT-----
110167745     YAYSPLHNLKPK--------TSYPPTFITTADHDDRVVPAHSFKFISTLQEVHIGDHP-------VLIRIETKAGHGAGKPTTKI-----
109899092     MTYSPVQNVREG--------VNYPATLITTGDHDDRVVPAHSFKFAAQLQAKNTGPNP-------MLIRIETNAGHGAGTPISKT-----
86144168      KGYSPVHNVKTG--------TEYPATLVTTGDHDDRVVPAHSFKFAAELQEKQAGDAP-------VLIRIETDAGHGAGTPVSKQ-----
4973227       KGYSPLHNVRPG--------VSYPSTMVTTADHDDRVVPAHSFKFAATLQADNAGPHP-------QLIRIETNAGHGAGTPVAKL-----
6048357       KGYSPLHNVRPG--------VSYPSTMVTTADHDDRVVPAHSFKFAATLQADNAGPHP-------QLIRIETNAGHGAGTPVAKL-----
76875127      KNYSPLHNVKAG--------IEYPATMITTGDHDDRVVPSHSFKFAAQLQAKQAGTNP-------TLIRIETNAGHGAGTPTSKI-----
216201        KGYSPLHSVRAG--------VSYPSTLVTTADHDDRVVPAHSFKFAATLQADDAGPHP-------QLIRIETNAGHGAGTPVAKL-----
88804657      KGYSPLHNIREG--------VAYPATLVTTADHDDRVVPAHSFKFAATLQEKHAGESP-------VLIRIETNAGHGAGTPISKT-----
89517967      KGYSPLHNIQEG--------VSYPATMVTTGDHDDRVVPAHSFKFAAELQSKQAGDAP-------TLIRIETDAGHGAGKSTAQV-----
90588230      KGYSPVQNVKKG--------TKYPATMVTTGDHDDRVVPAHSFKFAAELQDKQAGENP-------VLIRIDVKAGHGAGKSVAAT-----
71143549      KGYSPVHNVKAG--------VSYPATMVTTGDHDDRVVPAHSFKFAAELQAKQAGNAP-------TLIRIETNAGHGAGTPVSKT-----
83855898      KGYSPVHNVKEG--------VSYPATMVTTGDHDDRVVPAHSFKYAAELQDKQAGNAP-------TLIRIETNAGHGAGTPVSKT-----
85710915      LNYSPVHNVEQG--------VEYPATLITTGDHDDRVVPAHSFKFAAELQDKAGGDAP-------QLIRIETNAGHGAGTPVSKT-----
78368715      LAYSPYHNISE---------RDYPATMVMTADHDDRVVPLHSFKFGALLQDRQTGDAP-------IIMRIESKAGHGAGKPTAMK-----
56178266      LGYSPVHNVEEG--------VAYPATLITTGDHDDRVVPAHSFKFAAELQDKAGGENP-------QLIRIETNAGHGAGTPVSKT-----
```

```
91215481        KGYSPLHSIKDG--------TEYPATLVTTGDHDDRVVPAHSFKFAAELQSKQAGGAP-------TLIRIETDAGHGAGKPTSKI-----
78688320        LAYSPYHNVKT---------QAYPATMVMTADHDDRVVPLHSFKFAAMMQEMQQGDKP-------VIMRIESNAGHGAGKPTAMK-----
24374294        LAYSPYHNVKA---------QAYPATMVMTADHDDRVVPLHSFKFAAMMQEKQQGDKP-------VIMRIESNAGHGAGKPTSMK-----
456523          KSYSPVHNVKAG--------TCYPSTMVITSDHDDRVVPAHSFKFGSELQAKQSCKNP-------ILIRIETNAGHGAGRSTEQV-----
77816782        LAYSPYHNVKA---------QSYPATMVMTADHDDRVVPLHSFKFAAMLQDKQQGDKP-------VIMRIESNAGHGAGKPTAMK-----
68547790        LAYSPYHNVKA---------QTYPATMVMTADHDDRVVPLHSFKFGAMLQAKQQGQAP-------VIMRIESKAGHGAGKPTAMQ-----
82744257        LAYSPYHNVKA---------QSYPATMVMTADHDDRVVPLHSFKFAAMLQDKQQGDKP-------VIMRIESNAGHGAGKPTAMK-----
86131249        KGYSPVHNVKAG--------TSYPATMVTTGDHDDRVVPAHSFKFAAELQEKQAGDAP-------VLIRIETDAGHGAGTPVAKT-----
88796133        YAYSPVHNTKPG--------TCYPATLVTTGDHDDRVVPWHSYKFAAQLQADQGCDNP-------VLLRVETRAGHGAGTPTWMR-----
42526704        YAYSPLHNVKEG--------VNYPSIMVCTGDHDDRVVPAHSFKYAQALHDTYKGENP-------ILIRITEKAGHGAGKPTAKI-----
27367204        LGYSPVHNVKEG--------TAYPATLVTTADHDDRVVPAHSYKFIAELQEKHQGANP-------VLIRIDVNAGHGAGMPMSKM-----
37676919        LGYSPVHNVKEG--------TAYPATLVTTADHDDRVVPAHSYKFIAELQEKHQGANP-------VLIRIDVNAGHGAGMPMSKM-----
75855979        LGYSPVHNVKEN--------VQYPATLVTTADHDDRVVPAHSYKFIAELQDKQQGANP-------VLIRIDVNAGHGAGMPLSKQ-----
28809898        LGYSPVHNVKEG--------VQYPATLVTTADHDDRVVPAHSYKFIAELQDKQQGENP-------VLIRIDVNAGHGAGMPLSKQ-----
6460324         HAYSPLHNLKEG--------TRYPATLITTGDHDDRVVPAHSYKFAAELQRVQAGSAP-------TLIRIQTRAGHGAGKPTALV-----
91223249        LGYSPVHNVKEN--------VQYPATLVTTADHDDRVVPAHSYKFIAELQDKQQGANP-------VLIRIDVNAGHGAGMPLSKQ-----
56180108        YAYSPVHNTESG--------TCYPATLITTGDHDNRVVPWHSYKFAAALQRDQSCDQP-------ILLNVETRAGHGAGTPTWMR-----
55769741        IKYSPLHNVRRPWEQSFVNCCQYPAIMLLTADHDDRVVPLHSLKLLATLQYVLCTSIEDTPQVNPIIGRIDVKSGHGAGRPTKKM-----
47507507        IKYSPLHNIRVPEKDG----IQYPSMLLLTADHDDRVVPLHSLKFIASLQHIVGRSPN---QTNPLLIHVDTKAGHGAGKPTAKV-----
28502989        IKYSPLHNIRVPEKDG----IQYPSMLLLTADHDDRVVPLHSLKFIASLQHIVGRSPN---QTNPLLIHVDTKAGHGAGKPTAKV-----
86144914        LGYSPVHNVVRG--------TDYPATLVTTADHDDRVVPAHSYKFISELQDKHEGGAP-------VMIRIDVNAGHGAGMPLSKA-----
84386759        QGYSPVHNVVRG--------VDYPATLVTTADHDDRVVPAHSYKFIAELQDKHEGGVP-------VMIRIDVNAGHGAGMPLSKA-----
63100723        IKYSPLHNIQVPEGDG----VQYPAVLLLTGDHDDRVVPLHSLKYIATLQNVIGQCPG---QKNPLFIYIDTKSGHGAGKPTSKV-----
38649021        IKYSPLHNIKVPEKDG----IQYPSMLLLTADHDDRVVPLHSLKFIASLQHIAGQSPN---QTNPLLIHVDTKAGHGAGKPTAKV-----
59711729        LNYSPVHNVKEG--------VEYPATLVTTADHDDRVVPAHSYKFISELQAKQSGDNP-------VLIRIDVNAGHGAGMPISKS-----
6755152         LKYSPLHNVKLPEADD----IQYPSMLLLTADHDDRVVPLHSLKFIATLQYIVGRSRK---QSNPLLIHVDTKAGHGAGKPTAKV-----
74215349        LKYSPLHNVKLPEADD----IQYPSMLLLTADHDDRVVPLHSLKFIATLQYIVGRSRK---QSNPLLIHVDTKAGHGAGKPTAKV-----
50514023        IKYSPLHNVKLPEADD----IQYPSMLLLTADHDDRVVPLHSLKFIATLQYIVGRSRK---QNNPLLIHVDTKAGHGAGKPTAKV-----
26345256        LKYSPLHNVKLPEADD----IQYPSMLLLTADHDDRVVPLHSLKFIATLQYIVGRSRK---QSNPLLIHVDTKAGHGAGKPTAKV-----
157879458       IKYSPLHNVKLPEADD----IQYPSMLLLTADHNDRVVPLHSLKFIATLQYIVGRSRK---QNNPLLIHVDTKAGHGAGKPTAKV-----
51592147        IKYSPLHNVKLPEADD----IQYPSMLLLTADHDDRVVPLHSLKFIATLQYIVGRSRK---QNNPLLIHVDTKAGHGAGKPTAKV-----
10835490        IKYSPLHNVKLPEADD----IQYPSMLLLTADHDDRVVPLHSLKFIATLQYIVGRSRK---QNNPLLIHVDTKAGHGAGMPTAKV-----
50514022        IKYSPLHNVKLPEADD----IQYPSMLLLTADHDDRVVPLHSLKFIATLQYIVGRSRK---QNNPLLIHVDTKAGHGAGKPTAKV-----
37928243        IKYSPLHNVKLPEADD----IQYPSMLLLTADHDDRVVPLHSLKFIATLQYIVGRSRK---QNNPLLIHVDTKAGHGAGKPTAKV-----
27065055        IKYSPLHNVKLPEADD----IQYPSMLLLTADHDDRVVPLHSLKFIATLQYIVGRSRK---QNNPLLIHVDTKAGHGAGKPTAKV-----
5103285         IKYSPLHNVKLPEADD----IQYPSMLLLTADHDDRVVPLHSPKFIATLQHLVGRSRK---QNNPLLIHVDTKAGHGAGKPTAKV-----
```

```
13786144     LKYSPLHNVKLPEADD----IQYPSMLLLTADHDDRVVPLHSLKFIATLQYIVGRSRK---QSNPLLIHVDTKAGHGPGKPTAKV-----
5107662      IKYSPLHNVKLPEADD----IQYPSMLLLTADHDDRVVPLHSLKFIATLQYIVGRSRK---QNNPLLIHVDTKAGHGAGKPTAKV-----
79382269     IKYSPLHNVKRPWEQQTDHLVQYPSTMLLTADHDDRVVPLHSLKLLATLQHVLCTSLDNSPQMNPIIGRIEVKAGHGAGRPTQKM-----
27066372     IKYSPLHNVKLPEADD----IQYPSMLLLTADHADRVVPLHSLKFIATLQYIVGRSRK---QNNPLLIHVDTKAGHGAGKPTAKV-----
19347837     IKYSPLHNVKRPWEQQTDHLVQYPSTMLLTADHDDRVVPLHSLKLLATLQHVLCTSLDNSPQMNPIIGRIEVKAGHGAGRPTQKM-----
109072195    VKYSPLHNVKLPEADD----IQYPSMLLLTADHDDRVVPLHSLKFIATLQYIVGRSRK---QSNPLLIHVDTKAGHGAGKPTAKV-----
11691900     VKYSPLHNVKLPEADD----IQYPSMLLLTADHDDRVVPLHSLKFIATLQYIVGRSRK---QSNPLLIHVDTKAGHGAGKPTAKV-----
41349456     VKYSPLHNVKLPEADD----IQYPSMLLLTADHDDRVVPLHSLKFIATLQYIVGRSRK---QSNPLLIHVDTKAGHGAGKPTAKV-----
558596       VKYSPLHNVKLPEADD----IQYPSMLLLTADHDDRVVPLHSLKFIATLQYIVGRSRK---QSNPLLIHVDTKAGHGAGKPTAKV-----
53127346     CKYSPLHNVKLPEEDG----IQYPATLLLTADHDDRVVPLHSLKFIATLQYVVGRSRK---QTNPLLIHVDTKAGHGAGKPTAKV-----
904214       VKYSPLHNVKLPEADD----IQYPSMLLLTADHDDRVVPLHSLKFIATLQYIVGRSRK---QSNPLLIHVDTKAGHGAGKPTAKV-----
15217929     IKYSPLHNVKRPWEQKTDLFFQYPSTMLLTADHDDRVVPLHSYKLLATMQYELGLSLENSPQTNPIIARIEVKAGHGAGRPTQKM-----
9558588      IKYSPLHNVKRPWEQKTDLFFQYPSTMLLTADHDDRVVPLHSYKLLATMQYELGLSLENSPQTNPIIARIEVKAGHGAGRPTQKMCNVCE
38345850     IKYSPLHNVRRPWEKG-HRRQQYPSTMLLTADHDDRVVPSHTLKFLATMQHVLCTSVKESPQTNPIVARIDRKSGHGCGRSTQKI-----
91082735     YKYSPLHNIRVPQNGGQY-----PATLLLTADHDDRVVPLHSLKFIAELQNKIGRLPT---QKNPLMIRIETRAGHGAGKPTSKI-----
54650756     YKFSPLHNVHTPKGAETEY----PSTLILTADHDDRVSPLHSLKFIAALQEAVRDSEF---QKNPVLLRVYQKAGHGAGKPTSKR-----
73973662     IKYSPLHNVRLPEADD----TQYPSMLLLTADHDDRVVPLHSLKFIATLQYVVGRSRK---QSNPLLIHVDTKAGHGAGKPTAKV-----
73973660     IKYSPLHNVRLPEADD----TQYPSMLLLTADHDDRVVPLHSLKFIATLQYVVGRSRK---QSNPLLIHVDTKAGHGAGKPTAKV-----
47222200     IKYSPLHNLPQPPYSG----PAYPAVLLLTADHDDRVVPLHTLKYCAALQHGVGSSPA---QRQPLMVRVDTRSGHGAGKPTSKV-----
73973664     IKYSPLHNVRLPEADD----TQYPSMLLLTADHDDRVVPLHSLKFIATLQYVVGRSRK---QSNPLLIHVDTKAGHGAGKPTAKV-----
55235507     LRYSPLHNVRTPTSEKDQY----PATLVLTADHDDRVSPLHSLKFVAALHDAIKDSEH---QKNPLLLRVYSKAGHGMGKPTAKK-----
73973658     IKYSPLHNVRLPEADD----TQYPSMLLLTADHDDRVVPLHSLKFIATLQYVVGRSRK---QSNPLLIHVDTKAGHGAGKPTAKV-----
71666762     QKYSPLHNIKSGIKYP--------AILVVTGDHDDRVVPLHSLKYVATLQHMNPTEGG------PFLARIEVAAGHGAGKPTSKI-----
49523575     QKYSPLHNIKSGVKYP--------AILVVTGDHDDRVVPLHSLKYIAALQHENPTEGG------PFLARVEVAAGHGAGKPTSKI-----
71747962     QKYSPLHNIKSGVKYP--------AILVVTGDHDDRVVPLHSLKYIAALQHENPTEGG------PFLARVEVAAGHGAGKPTSKI-----
33317309     QQYSPLHNIKSGIKYP--------AILVVTGDHDDRVVPLHSLKYVATLQHMNPNEGG------PFLARIEVAAGHGAGKPTSKI-----
47224662     MKYSPLHNIHVPEGNG----VQYPAVLLLTGDHDDRVVPLHSLKYIATLQHIVGRSPK---QTNPLFILVDTKSGHGAGKPTSKV-----
15291259     YKFSPLHNVHTPKGAKRN---------------THRL-----------------------------------------------------
73973656     IKYSPLHNVRLPEADD----TQYPSMLLLTADHDDRVVPLHSLKFIATLQYVVGRSRK---QSNPLLIHVDTKAGHGAGKPTAKV-----
               :** :.                          .*:
Prim.cons.   IKYSPLHNVKLGEADDTD32TQYPATLVTTADHDDRVVPLHSLKFIATLQYKQGGSRPNSPQSNPLLIRIETKAGHGAGKPTAKVCNVCE


                    910       920       930       940       950
```

```
                      |         |         |         |         |
62738741     ---IEETADVQAFLAHFTG-----LTPRPWSSVDKLAAALEHHHHHH-------
87200538     ---IAENADVYAFVAHWTG-----LTPKE-------------------------
103487461    ---IAEAADKYAFAAKWTG-----LDVE--------------------------
87200537     ---IGAGADVMAFLAYWTG-----LSL---------------------------
94496227     ---IDEYADSYAFAAHFTG-----LAIAPRPAAAPRSAAGQPAHVMAPIVAGGQ
35211243     ---IEEAADRWAFLVANLR-----MQLKF-------------------------
32448009     ---IDEYADLWSFLLENLK-----------------------------------
75906687     ---IEEAADKWAFLVRALA-----VEV---------------------------
35211187     ---IEEAADRWAFLVANLR-----MQLKF-------------------------
67921530     ---IEEIADKWAFLVDNLA-----FNVL--------------------------
67930945     ---IDEAADILTFLKAALK-----VN----------------------------
17131625     ---IEEAADKWAFLVRTLA-----VEV---------------------------
86609965     ---IEEAADRWAFLVQVLG-----IPWDGTGS----------------------
83859556     ---IEEAADRWAFIAYHTG-----LEIED-------------------------
28808634     ---IETNADIYSFALFNMG-----IEKLQ-------------------------
75854458     ---IETNADIYSFALFNMG-----IEKLQ-------------------------
86607173     ---IEETADRWAFLVQVLG-----IQAG--------------------------
23128197     ---IEEAADKWAFLVRTLD-----VEV---------------------------
91226380     ---IETNADIYSFALFNMG-----IEKLQ-------------------------
88711708     ---IEQYADIFGFTLFNMG-----FEELPNQRVLKEFKD---------------
110167745    ---IAEITDEFAFLLRNLK-----IELPENFGN---------------------
109899092    ---IEQYADIFGFTLYNMG-----IKVQ--------------------------
86144168     ---IEQTADIFAFTLYNMG-----FDVLPEKTKS-EIKG---------------
4973227      ---IEQSADIYAFTLYEMG-----YRELPRQP----------------------
6048357      ---IEQSADIYAFTLYEMG-----YRELPRQP----------------------
76875127     ---IDLYADMYGFTLYNMG-----VKSL--------------------------
216201       ---IEQSADIYAFTLFEMG-----YRQLPRQP----------------------
88804657     ---IEQYADIFGFTFYNMG-----YEELPNQAALKDFKE---------------
89517967     ---IQEQTDIYSFALFNMG-----FEELPELSTRVKM-----------------
90588230     ---IQENVDIQAFTLYNMG-----FKALPKK-----------------------
71143549     ---IEQYADIYAFTLFNMG-----FTALPK------------------------
83855898     ---IEQYADIFGFTLYNMG-----YDELPVKK---QFKD---------------
85710915     ---IEQYSDIFGFTLYHMG-----YQSLPNE-----------------------
78368715     ---IDEFADIYSFLMHSFG-----LAVPSKIGG---------------------
56178266     ---IEQYSDIFGFTLYHMG-----FEELPE------------------------
```

```
91215481     ---IQEYADIFAFTFYNMG-----YTEL--------------------------
78688320     ---IDEFADIYSFLWQSFG-----LTLPQTLAK---------------------
24374294     ---IDEFADIYSFLWQSFG-----LTLPQTLAK---------------------
456523       ---VAENADLLSFALYEMG-----IKSLK-------------------------
77816782     ---IDEFADIYSFLWQSFG-----LTLPQTIAK---------------------
68547790     ---IAEFADIYAFLWQNFG-----LTLPTEL-----------------------
82744257     ---IDEFADIYSFLWQSFG-----LKLPQTIAK---------------------
86131249     ---IEQYADIFGFTLYNMG-----YEVLPEKVGN-VMKK---------------
88796133     ---IEGYADQWAFLESALE-----------------------------------
42526704     ---IEETADIYAFIFKQTG-----HII---------------------------
27367204     ---IDLSADMYAFTLFNMG-----ITALK-------------------------
37676919     ---IDLSADIYAFTLFNMG-----ITALK-------------------------
75855979     ---IDLTTDVYAFTLYNMG-----IESI--------------------------
28809898     ---IDLTTDVYAFTLYNMG-----IESI--------------------------
6460324      ---IEEAADIWAFLEEVLG-----G-----------------------------
91223249     ---IDLTTDVYAFTLHNMG-----IESI--------------------------
56180108     ---IEEHAENWAFLYKHLG-----MNSEVVQSNASE------------------
55769741     ---IDEVADRYSFMANMLD-----ASWTE-------------------------
47507507     ---IEEVSDMFAFIAQCLG-----LQWIE-------------------------
28502989     ---IEEVSDMFAFIAQCLG-----LQWIE-------------------------
86144914     ---IDLTADIYAFTLFNMG-----IESLDSI-----------------------
84386759     ---LDLTADIYAFTLFNMG-----IESLDSL-----------------------
63100723     ---IQEVADTYAFIARCLN-----LSWLE-------------------------
38649021     ---IEEVSDMFAFIAQCLN-----LQWIE-------------------------
59711729     ---MDLMADVYAFTLSNMK-----VNPFE-------------------------
6755152      ---IEEVSDMFAFIARCLN-----IEWIQ-------------------------
74215349     ---IEEVSDMFAFIARCLN-----IEWIQ-------------------------
50514023     ---IEEVSDMFAFIARCLN-----IDWIP-------------------------
26345256     ---IEEVSDMFAFIAAVPEHRVDPVNGISSSSHQTQKTSRAFTPS---------
157879458    ---IEEVSDMFAFIARCLN-----IDWIPXGP----------------------
51592147     ---IEEVSDMFAFIARCLN-----IDWIP-------------------------
10835490     ---IEEVSDMFAFIARCLN-----IDWIP-------------------------
50514022     ---IEEVSDMFAFIARCLN-----IDWIP-------------------------
37928243     ---IEEVSDMFAFIARCLN-----IDWIP-------------------------
27065055     ---IEEVSDMFAFIARCLN-----IDWIP-------------------------
5103285      ---IEEVSDMFAFIARCLN-----IDWIQ-------------------------
```

```
13786144     ---IEEVSDMFAFIARCLN-----IEWIQ-------------------------
5107662      ---IEEVSDMFAFIARCLN-----IDWIP-------------------------
79382269     ---IDEAADRYSFMAKMVN-----ASWTE-------------------------
27066372     ---IEEVSDMFAFIARCLN-----IDWIP-------------------------
19347837     ---IDEAADRYSFMAKMVN-----ASWTE-------------------------
109072195    ---IEEVSDMFAFIARCLN-----IDWIP-------------------------
11691900     ---IEEVSDMFAFIARCLN-----VDWIP-------------------------
41349456     ---IEEVSDMFAFIARCLN-----VDWIP-------------------------
558596       ---IEEVSDMFAFIARCLN-----VDWIP-------------------------
53127346     ---IEEVSDMFAFIARCLN-----LDWIE-------------------------
904214       ---IEEVSDMFAFIARCLN-----IDWIP-------------------------
15217929     ---IDEAADRYSFMAKMVD-----ASWID-------------------------
9558588      LQQIDEAADRYSFMAKMVD-----ASWID-------------------------
38345850     ---IDEAADRYAFAAKTMG-----ISWID-------------------------
91082735     ---IEEVTDTFCFISRALN-----LTFSS-------------------------
54650756     ---IEEATDILTFLSKSLN-----VDIVNV------------------------
73973662     ---IEEVSDMFAFIARCLD-----IDWIQ-------------------------
73973660     ---IEEVSDMFAFIARCLD-----IDWIQ-------------------------
47222200     ---ILEDTDIFSFIAETLG-----LSW---------------------------
73973664     ---IEEVSDMFAFIARCLD-----IDWIQ-------------------------
55235507     ---IEEATDILTFMYKTLG-----LKLSF-------------------------
73973658     ---IEEVSDMFAFIARCLD-----IDWIQ-------------------------
71666762     ---LREAGDIYTFIAKNIN-----ASWKE-------------------------
49523575     ---MQESADIYTFIAKNTN-----AQWTD-------------------------
71747962     ---MQESADIYTFIAKNTN-----AQWTD-------------------------
33317309     ---LREAGDIYTFIAKNIN-----ASWKE-------------------------
47224662     ---IQEVADTYAFIAKCLN-----ISWVE-------------------------
15291259     ------------------------------------------------------
73973656     ---IEEVSDMFAFIARCLD-----IDWIQ-------------------------

Prim.cons.   LQQIEEVADIFAFIARNLGHRVDPIDWIP2QAK2KEFK23333H322PIVAGGQ
```

Alignment data :
Alignment length : 954

CLUSTALW options used :
endgaps=1

```
Identity (*) : 46 is 4.82 %
Strongly similar (:) : 47 is 4.93 %
Weakly similar (.) : 29 is 3.04 %
Different : 832 is 87.21 %
```

```
gapdist=8
gapext=0.2
gapopen=10.0
hgapresidues=GPSNDQERK
matrix=gonnet
maxdiv=30
outorder=input
type=PROTEIN
```

Names of proteins corresponding to numbers used in alignment:

```
>gi|62738741|pdb|1YR2|A Chain A prolyl oligopeptidase [Novosphingobium capsulatum]
>gi|87200538|ref|YP_497795.1| Prolyl oligopeptidase [Novosphingobium aromaticivorans DSM 12444]
>gi|103487461|ref|YP_617022.1| Prolyl oligopeptidase [Sphingopyxis alaskensis RB2256]
>gi|87200537|ref|YP_497794.1| Prolyl oligopeptidase [Novosphingobium aromaticivorans DSM 12444]
>gi|94496227|ref|ZP_01302805.1| Prolyl oligopeptidase [Sphingomonas sp. SKA58]
>gi|35211243|dbj|BAC88622.1| prolyl endopeptidase [Gloeobacter violaceus PCC 7421]
>gi|32448009|emb|CAD77529.1| prolyl endopeptidase [Rhodopirellula baltica SH 1]
>gi|75906687|ref|YP_320983.1| Peptidase S9, prolyl oligopeptidase active site region [Anabaena variabilis ATCC 29413]
>gi|35211187|dbj|BAC88566.1| prolyl endopeptidase [Gloeobacter violaceus PCC 7421]
>gi|67921530|ref|ZP_00515048.1| Prolyl oligopeptidase [Crocosphaera watsonii WH 8501]
>gi|67930945|ref|ZP_00524112.1| Prolyl oligopeptidase [Solibacter usitatus Ellin6076]
>gi|17131625|dbj|BAB74232.1| prolyl endopeptidase [Nostoc sp. PCC 7120]
>gi|86609965|ref|YP_478727.1| prolyl oligopeptidase [Synechococcus sp. JA-2-3B'a(2-13)]
>gi|83859556|ref|ZP_00953076.1| prolyl endopeptidase [Oceanicaulis alexandrii HTCC2633]
>gi|28808634|dbj|BAC61811.1| prolyl endopeptidase [Vibrio parahaemolyticus RIMD 2210633]
>gi|75854458|ref|ZP_00762150.1| COG1505: Serine proteases of the peptidase family S9A [Vibrio sp. Ex25]
>gi|86607173|ref|YP_475936.1| prolyl oligopeptidase [Synechococcus sp. JA-3-3Ab]
>gi|23128197|ref|ZP_00110050.1| COG1505: Serine proteases of the peptidase family S9A [Nostoc punctiforme PCC 73102]
>gi|91226380|ref|ZP_01261179.1| prolyl endopeptidase [Vibrio alginolyticus 12G01]
>gi|88711708|ref|ZP_01105796.1| Prolyl endopeptidase [Flavobacteriales bacterium HTCC2170]
```

```
>gi|110167745|gb|ABG52285.1| prolyl oligopeptidase. Serine peptidase. MEROPS family S09A [Trichodesmium erythraeum IMS101]
>gi|109899092|ref|YP_662347.1| Prolyl oligopeptidase [Pseudoalteromonas atlantica T6c]
>gi|86144168|ref|ZP_01062505.1| Prolyl endopeptidase [Flavobacterium sp. MED217]
>gi|4973227|gb|AAD34991.1| prolyl endopeptidase [Aeromonas punctata subsp. punctata]
>gi|6048357|gb|AAF02211.1| prolyl endopeptidase [Aeromonas punctata subsp. punctata]
>gi|76875127|emb|CAI86348.1| Prolyl endopeptidase [Pseudoalteromonas haloplanktis TAC125]
>gi|216201|dbj|BAA03105.1| prolyl endopeptidase [Aeromonas hydrophila]
>gi|88804657|ref|ZP_01120177.1| Prolyl endopeptidase [Robiginitalea biformata HTCC2501]
>gi|89517967|gb|EAS20623.1| prolyl endopeptidase serine protease [Flavobacteria bacterium BBFL7]
>gi|90588230|ref|ZP_01243882.1| Prolyl oligopeptidase [Flavobacterium johnsoniae UW101]
>gi|71143549|gb|AAZ24022.1| prolyl endopeptidase [Colwellia psychrerythraea 34H]
>gi|83855898|ref|ZP_00949427.1| Prolyl endopeptidase [Croceibacter atlanticus HTCC2559]
>gi|85710915|ref|ZP_01041976.1| Prolyl endopeptidase [Idiomarina baltica OS145]
>gi|78368715|ref|ZP_00838926.1| Prolyl oligopeptidase [Shewanella sp. PV-4]
>gi|56178266|gb|AAV80988.1| Prolyl endopeptidase [Idiomarina loihiensis L2TR]
>gi|91215481|ref|ZP_01252452.1| Prolyl endopeptidase [Psychroflexus torquis ATCC 700755]
>gi|78688320|ref|ZP_00853025.1| Prolyl oligopeptidase [Shewanella sp. ANA-3]
>gi|24374294|ref|NP_718337.1| prolyl endopeptidase [Shewanella oneidensis MR-1]
>gi|456523|emb|CAA45213.1| proline-specific endopeptidase; prolyl endopeptidase [Elizabethkingia meningoseptica]
>gi|77816782|ref|ZP_00815936.1| Prolyl oligopeptidase [Shewanella putrefaciens CN-32]
>gi|68547790|ref|ZP_00587318.1| Prolyl oligopeptidase [Shewanella amazonensis SB2B]
>gi|82744257|ref|ZP_00906837.1| Prolyl oligopeptidase [Shewanella sp. W3-18-1]
>gi|86131249|ref|ZP_01049848.1| Prolyl endopeptidase [Cellulophaga sp. MED134]
>gi|88796133|ref|ZP_01111815.1| Prolyl endopeptidase [Alteromonas macleodii 'Deep ecotype']
>gi|42526704|ref|NP_971802.1| prolyl endopeptidase [Treponema denticola ATCC 35405]
>gi|27367204|ref|NP_762731.1| Serine protease of the peptidase family S9A [Vibrio vulnificus CMCP6]
>gi|37676919|ref|NP_937315.1| serine protease [Vibrio vulnificus YJ016]
>gi|75855979|ref|ZP_00763616.1| COG1505: Serine proteases of the peptidase family S9A [Vibrio sp. Ex25]
```

```
>gi|28809898|dbj|BAC62839.1| prolyl endopeptidase [Vibrio parahaemolyticus RIMD 2210633]
>gi|6460324|gb|AAF12044.1|AE002079_9 prolyl endopeptidase [Deinococcus radiodurans R1]
>gi|91223249|ref|ZP_01258515.1| prolyl endopeptidase [Vibrio alginolyticus 12G01]
>gi|56180108|gb|AAV82830.1| Prolyl endopeptidase [Idiomarina loihiensis L2TR]
>gi|55769741|ref|XP_549860.1| putative prolyl endopeptidase [Oryza sativa (japonica cultivar-group)]
>gi|47507507|gb|AAH71008.1| PREP protein [Xenopus laevis]
>gi|28502989|gb|AAH47161.1| PREP protein [Xenopus laevis]
>gi|86144914|ref|ZP_01063246.1| prolyl endopeptidase [Vibrio sp. MED222]
>gi|84386759|ref|ZP_00989784.1| prolyl endopeptidase [Vibrio splendidus 12B01]
>gi|63100723|gb|AAH95363.1| Zgc:110670 [Danio rerio]
>gi|38649021|gb|AAH63222.1| Prolyl endopeptidase [Xenopus tropicalis]
>gi|59711729|ref|YP_204505.1| prolyl endopeptidase [Vibrio fischeri ES114]
>gi|6755152|ref|NP_035286.1| prolyl endopeptidase [Mus musculus]
>gi|74215349|dbj|BAE41885.1| unnamed protein product [Mus musculus]
>gi|50514023|pdb|1VZ3|A Chain A, Prolyl Oligopeptidase From Porcine Brain, T597c Mutant
>gi|26345256|dbj|BAC36278.1| unnamed protein product [Mus musculus]
>gi|157879458|pdb|1O6G|A Chain A, Prolyl Oligopeptidase From Porcine Brain, D641n Mutant With Bound Peptide Ligand
>gi|51592147|ref|NP_001004050.1| prolyl endopeptidase [Sus scrofa]
>gi|10835490|pdb|1E5T|A Chain A, Prolyl Oligopeptidase From Porcine Brain, Mutant
>gi|50514022|pdb|1VZ2|A Chain A, Prolyl Oligopeptidase From Porcine Brain, Y73cV427CC255T Mutant
>gi|37928243|pdb|1UOQ|A Chain A, Prolyl Oligopeptidase From Porcine Brain, S554a Mutant With Bound Peptide Ligand Glu-Phe-Ser-Pro
>gi|27065055|pdb|1H2Y|A Chain A, Prolyl Oligopeptidase From Porcine Brain, Y473f Mutant With Covalently Bound Inhibitor Z-Pro-Prolinal
>gi|5103285|dbj|BAA78907.1| prolyl oligopeptidase [Bos taurus]
>gi|13786144|ref|NP_112614.1| prolyl endopeptidase [Rattus norvegicus]
>gi|5107662|pdb|1QFM|A Chain A, Prolyl Oligopeptidase From Porcine Muscle
>gi|79382269|ref|NP_177741.3| prolyl oligopeptidase [Arabidopsis thaliana]
```

```
>gi|27066372|pdb|1O6F|A Chain A, Prolyl Oligopeptidase From Porcine Brain, D641a Mutant With Bound Peptide Ligand Suc-Gly-Pro-replaced original
>gi|19347837|gb|AAL86330.1| putative prolyl endopeptidase [Arabidopsis thaliana]
>gi|109072195|ref|XP_001087580.1| PREDICTED: similar to prolyl endopeptidase [Macaca mulatta]
>gi|11691900|dbj|BAB19053.1| prolyl oligopeptidase [Homo sapiens]
>gi|41349456|ref|NP_002717.3| prolyl endopeptidase [Homo sapiens]
>gi|558596|emb|CAA52605.1| prolyl oligopeptidase [Homo sapiens]
>gi|53127346|emb|CAG31056.1| hypothetical protein [Gallus gallus]
>gi|904214|dbj|BAA04661.1| prolyl endopeptidase [Homo sapiens]
>gi|15217929|ref|NP_173463.1| prolyl oligopeptidase, putative / prolyl endopeptidase, putative / post-proline cleaving enzyme, putative [Arabidopsis thaliana]
>gi|9558588|gb|AAF88151.1|AC026234_2 Contains similarity to a rPOP protein from Rattus norvegicus gi|3043760 and is a member of the prolyl oligopeptidase family PF|00326.  ESTs gb|AA651190, gb|H36145 come from this gene. [Arabidopsis thaliana]
>gi|38345850|emb|CAE01849.2| OSJNBa0084K11.14 [Oryza sativa (japonica cultivar-group)]
>gi|91082735|ref|XP_972959.1| PREDICTED: similar to CG5355-PA [Tribolium castaneum]
>gi|54650756|gb|AAV36957.1| LP07359p [Drosophila melanogaster]
>gi|73973662|ref|XP_539066.2| PREDICTED: similar to prolyl endopeptidase isoform 1 [Canis familiaris]
>gi|73973660|ref|XP_868346.1| PREDICTED: similar to prolyl endopeptidase isoform 5 [Canis familiaris]
>gi|47222200|emb|CAG11079.1| unnamed protein product [Tetraodon nigroviridis]
>gi|73973664|ref|XP_854289.1| PREDICTED: similar to Prolyl endopeptidase (Post-proline cleaving enzyme) (PE) isoform 2 [Canis familiaris]
>gi|55235507|gb|EAA14977.2| ENSANGP00000016749 [Anopheles gambiae str. PEST]-been replaced
>gi|73973658|ref|XP_868343.1| PREDICTED: similar to prolyl endopeptidase isoform 4 [Canis familiaris]
>gi|71666762|ref|XP_820337.1| prolyl oligopeptidase, putative [Trypanosoma cruzi strain CL Brener]
>gi|49523575|emb|CAD42967.1| prolyl oligopeptidase [Trypanosoma brucei]
>gi|71747962|ref|XP_823036.1| prolyl oligopeptidase [Trypanosoma brucei TREU927]
>gi|33317309|gb|AAQ04681.1|AF452421_1 80 kDa prolyl oligopeptidase [Trypanosoma cruzi]
>gi|47224662|emb|CAG03646.1| unnamed protein product [Tetraodon nigroviridis]
```

>gi|15291259|gb|AAK92898.1| GH13952p [Drosophila melanogaster]
>gi|73973656|ref|XP_868339.1| PREDICTED: similar to Prolyl endopeptidase (Post-proline cleaving enzyme) (PE) isoform 3 [Canis familiaris]

SC PEP PYMOL Numbering	Clustal W alignment numbering
M102L	M178
K103T	K179
R110K	R186
F227Y	F315
I284L	I381
V301L	V398
N303Q	N400
K305S	K407
V323I	V434
I343L/V	I454
V344I	V455
V346I	V457
S351P	S462
P353S	P473
R354K	R474
I371V	I506
I381L	I516

I406V	**I542**
V434I	**V570**
L452V	**L597**
F459Y	**F607**
V474I	**V622**
R481K	**R629**
V500I	**V651**
M511L	**M686**
I514V	**I689**
I581V	**I757**
Y669F	**Y855**
I687V	**I880**

Supplemental Figure 2a. List of 1st generation variants with their corresponding mutations. Variant numbers are listed in the 1st column and mutations are listed along the row. A "1" in a given cell highlighted in purple for enhanced identification indicates that the mutation was incorporated into that variant.

Variant #	M102L	K103T	R110K	F227Y	I284L	V301L	N303Q	K305S	V323I	I343L	I343V	V344I	V346I	S351P	P353S	R354K	I371V	I381L	I406V	V434I	L452V	F459Y	V474I	R481K	V500I	M511L	I514V	I581V	Y669F	I687V
6609							1			1																			1	
6610																			1										1	1
6611				1			1																					1		
6612									1		1	1																		
6613					1								1													1				
6614		1			1		1																							
6615					1	1																		1						
6617			1	1				1																						
6618						1									1					1										
6619									1				1	1																
6620			1								1					1														
6621	1													1	1															
6622										1		1																		1
6623																						1			1		1			
6624				1													1						1							
6625	1					1															1									
6626								1				1													1					
6627		1						1	1																					
6628			1																		1		1							
6629										1								1								1				
6630		1																		1				1						
6631																	1										1	1		
6632	1										1											1								
6633	1											1											1					1		
6634		1	1										1									1								
6635						1								1				1									1			
6636														1		1				1										1
6637								1					1					1						1						
6638				1	1														1		1									
6639																	1								1	1				1
6640															1						1						1		1	
6641							1				1									1					1					
6642										1									1			1						1		
6643									1						1	1	1													
6644																							1	1		1			1	
6645			1									1						1		1									1	
6646								1											1	1						1	1			
6647					1										1							1	1							1
6648				1						1			1		1										1					
6649						1							1			1												1	1	
6650	1			1												1								1			1			
6651							1		1												1			1						1
6652		1									1							1					1		1					
6653											1			1							1					1		1		
6654		1										1		1			1		1											
6655	1		1		1				1	1																				
6656						1	1	1									1					1								

Supplemental Figure 2b. List of 2^{nd} generation variants with their corresponding mutations. Variant numbers are listed in the 1^{st} column and mutations are listed along the row. A “1” in a given cell highlighted in purple for enhanced identification indicates that the mutation was incorporated into that variant.

Variant #	M102L	K103T	R110K	F227Y	I284L	V301L	N303Q	K305S	V323I	I343L	I343V	V344I	V346I	S351P	P353S	R354K	I371V	I381L	I406V	V434I	L452V	F459Y	V474I	R481K	V500I	M511L	I514V	I581V	Y669F	I687V
10184																						1						1		
10185		1												1			1					1					1	1		
10186									1																			1		
10187									1								1					1						1		
10188											1			1							1	1								
10189		1									1			1			1				1	1				1		1		
10190									1		1			1														1		
10191		1							1													1						1		
10192																			1	1										
10193															1		1					1								
10194		1						1	1								1					1						1		
10195		1												1					1									1		
10196									1										1			1				1		1		
10197					1																	1								
10198								1											1			1				1				
10199		1							1			1		1			1		1			1						1		
10200					1										1							1						1		
10201																					1	1				1		1		
10202												1		1																
10203		1												1					1			1								
10204														1			1		1			1						1		
10205																			1							1	1	1		
10206								1											1							1		1		
10207											1							1												
10208																		1				1								
10209	1													1	1		1					1						1		
10210																										1		1		
10211					1																	1				1		1		
10212									1												1									
10213								1														1								
10214	1										1			1			1					1						1		
10215								1																				1		
10216		1									1							1										1		
10217									1	1																				
10218		1									1							1				1								
10219								1											1			1				1	1	1		
10220									1					1	1													1		
10221									1		1											1						1		
10222						1																						1		
10223	1											1										1								
10224					1										1		1					1	1					1		
10225		1									1																			
10226									1					1								1						1		
10227												1						1												
10228							1				1											1								
10229	1					1											1				1	1						1		
10230					1																							1		
10231														1	1							1						1		