>seq1 FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGPAHGTQWEIDIEFLGKDTTKVQFNYYTNGVGGHEKVISLGFDASKGFHTYAFDWQPGYIKWYVDGVLK >seq2 KASEDLVKKHAGVLGAILKKKGHHEAELKPLAQSHATKAHKNIFISEAIIHVLHSRHPGDFGADAQGAMNKALELFRKDIAAKYKELGY >seq3 TVEGAGSIAAATGFVKKDQLGKNEEGAPQEGILEDMPVDPDNEAYEMPSEEGYQDYEPEA >1PRC:L|PDBID|CHAIN|SEQUENCE ALLSFERKYRVRGGTLIGGDLFDFWVGPYFVGFFGVSAIFFIFLGVSLIGYAASQGPTWDPFAISINPPDLKYGLGAAPLLEGGFWQAITVCALGAFISWMLREVEISRKLGIGWHVPLAFCVPIFMFCVLQVFRPLL

Q1.1. Obtain the hydrophobicity profile for the sequences (Q1.fasta) and identify the ▢-helices and β-strands.

Hydrophobicity values: A: 13.85 D: 11.61 C: 15.37 E: 11.38 F: 13.93 G: 13.34 H: 13.82 I: 15.28 K: 11.58 L: 14.13 M: 13.86 N: 13.02 P: 12.35 Q: 12.61 R: 13.10 S: 13.39 T: 12.70 V: 14.56 W: 15.48 Y: 13.88

In [335...
```python
%matplotlib notebook
```

In [23]:
```python
import matplotlib.pyplot as plt
from matplotlib.figure import Figure
fig = Figure()


seq1="FDCAEYRSTNIYGYGLYEVSMKPAKNTGIVSSFFTYTGPAHGTQWEIDIEFLGKDTTKVQFNYYTNGVGGHEKVISLGFDASKGFHTYAFDWQPGYIKWYVDGVLK"

seq2="KASEDLVKKHAGVLGAILKKKGHHEAELKPLAQSHATKAHKNIFISEAIIHVLHSRHPGDFGADAQGAMNKALELFRKDIAAKYKELGY"

seq3="TVEGAGSIAAATGFVKKDQLGKNEEGAPQEGILEDMPVDPDNEAYEMPSEEGYQDYEPEA"

seq4="ALLSFERKYRVRGGTLIGGDLFDFWVGPYFVGFFGVSAIFFIFLGVSLIGYAASQGPTWDPFAISINPPDLKYGLGAAPLLEGGFWQAITVCALGAFISWMLREVEISRKLGIGWHVPLAFCVPIFMFCVLQV

hy_val={"A": 13.85, "D": 11.61 ,'C': 15.37 ,'E': 11.38, 'F': 13.93 ,'G': 13.34, 'H': 13.82, 'I': 15.28 ,'K': 11.58 ,'L': 14.13 ,
'M': 13.86 ,'N': 13.02 ,'P': 12.35 ,'Q': 12.61, 'R': 13.10, 'S': 13.39, 'T': 12.70, 'V': 14.56 ,'W': 15.48, 'Y': 13.88}
```

In [21]:
```python
def my_h_plot(seq):

    x=[i for i in range(1,len(seq)+1) ]
    y=list()
    for i in seq:

        y.append(hy_val[i])

    plt.figure()
    plt.plot(x,y,c="red")
    plt.xticks(x)
    plt.xlim(1,len(seq))
    # add a label to the x axis
    plt.xlabel('Amino acid sequence')
    # add a label to the y axis
    plt.ylabel('hydrophobicity index')
    # add a title
    plt.title('Relationship between sequence no and hydrophobicity profile.')
    return x,y
```
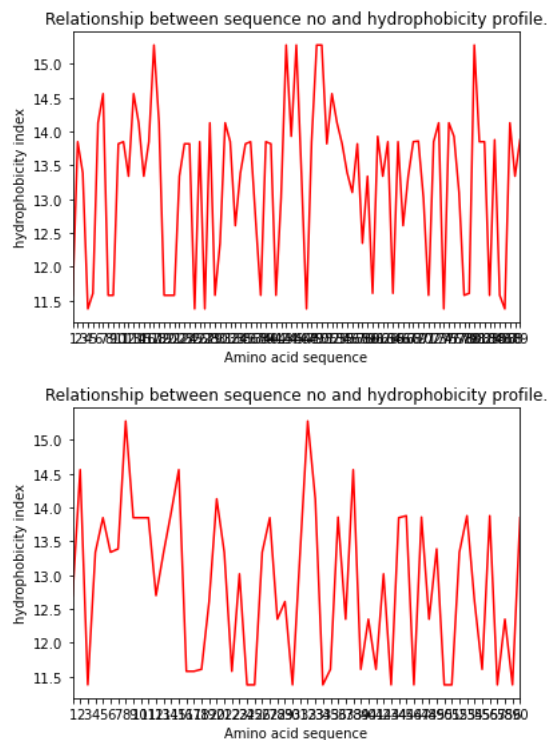
In [26]:
```python
x1,y1=my_h_plot(seq1);
x2,y2=my_h_plot(seq2);
x3,y3=my_h_plot(seq3);
```

Relationship between sequence no and hydrophobicity profile.

Relationship between sequence no and hydrophobicity profile.
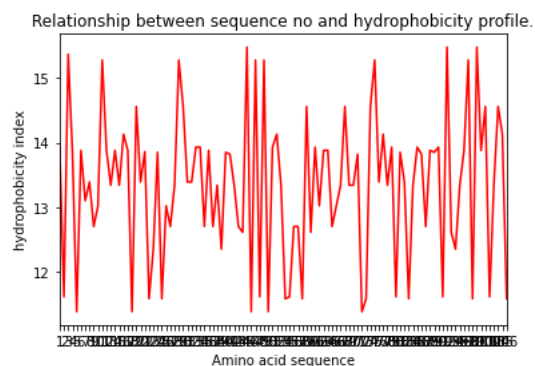
I have marked prominent alpha annd beta sheet in above three sequence.

in seq1 :alpha helix=[(5,12),{13,20),(25,32),(33,40),(60,66),(92,99)]

beta sheet=[(0,5),(44,49)]

in seq2 :alpha helix=[(1,8),(9,16),(32,39),(51,58)]

beta sheet=[(40,45),(82,87)]

in seq3 :alpha helix=[(3,10),(11,18),(19,26)]

beta sheet=[(28,33),(40,45)]

Q2. Calcualte the amphiphatic index for the helices and strands found in Q1. Use stretch lengths of 8 and 6 for ☐-helices and β-strands, respectively.

```python
In [52]:  def amphicity(x,y,B=None,A=None):

              #beta sheet

              for i in B:
                  s=i[0]
                  e=i[1]
                  ad1=[sum(y[i] for i in range(s,e+1) if i%2==0)]
                  ad2=[sum(y[i] for i in range(s,e+1) if i%2!=0)]
                  re=abs( (ad1[0]/3)-(ad2[0]/3) )
                  print("beta",i,re)

              #alpha sheet

              for i in B:
                  s=i[0]
                  e=i[1]
                  ad1=[sum(y[i] for i in range(s,e+1) if i%4==0 or i%4==1) ]
                  ad2=[sum(y[i] for i in range(s,e+1) if i%4==2 or i%4==3)]
                  re=abs( (ad1[0]/4)-(ad2[0]/4) )
                  print("alpha",i,re)
```

In [54]:
```python
# seq1 :
A_1=[(5,12),(13,20),(25,32),(33,40),(60,66),(92,99)]
B_1=[(0,5),(44,49)]
# seq2 :
A_2=[(1,8),(9,16),(32,39),(51,58)]
B_2=[(40,45),(82,87)]
# seq3 :
A_3=[(3,10),(11,18),(19,26)]
B_3=[(28,33),(40,45)]
print("for seq 1")
amphicity(x2,y2,A_2,B_2)
print("for seq 2")
amphicity(x2,y2,A_2,B_2)
print("for seq 3")
amphicity(x3,y3,A_3,B_3)
```

```
for seq 1
beta (1, 8) 0.06666666666666643
beta (9, 16) 0.6300000000000026
beta (32, 39) 0.11333333333333329
beta (51, 58) 0.283333333333335
alpha (1, 8) 0.0649999999999995
alpha (9, 16) 0.8525000000000009
alpha (32, 39) 1.2650000000000006
alpha (51, 58) 0.067499999999999
for seq 2
beta (1, 8) 0.06666666666666643
beta (9, 16) 0.6300000000000026
beta (32, 39) 0.11333333333333329
beta (51, 58) 0.283333333333335
alpha (1, 8) 0.0649999999999995
alpha (9, 16) 0.8525000000000009
alpha (32, 39) 1.2650000000000006
alpha (51, 58) 0.067499999999999
for seq 3
beta (3, 10) 0.28999999999999915
beta (11, 18) 0.7566666666666677
beta (19, 26) 0.3866666666666667
alpha (3, 10) 0.24249999999999972
alpha (11, 18) 0.24749999999999872
alpha (19, 26) 0.6850000000000005
```

Q3. Plot the hydrophobicity profile for the sequence (Q2.fasta) with window lengths 9 and 19 and list the transmembrane segments.

In [345...
```python
def my_h_r_plot(seq,w=None):

    x=[i for i in range(1,len(seq)+1) ]
    y=list()
    for i in seq:

        y.append(hy_val[i])

    yn=list()

    for i in range(w):
        yn.append(y[i])

    for i in range(w,len(seq)-w):

        p=(sum(y[i-w:i])+sum(y[i+1:i+w+1]))/(2*w)
        yn.append(p)

    for i in range(len(seq)-w,len(seq)):
        yn.append(y[i])


    plt.figure()
    plt.plot(x,yn,c="red")

#     plt.ylim(11,16)
#     axes = plt.gca()
#     axes.xaxis.label.set_size(40)
#     axes.yaxis.label.set_size(20)
    # add a label to the x axis
    plt.xlabel('Amino acid sequence')
    # add a label to the y axis
    plt.ylabel('hydrophobicity index')
    # add a title
    plt.title('Relationship between sequence no and hydrophobicity profile.')
```

In [346...
```python
#window_length 9,n=4 and  for 19 =n=8
my_h_r_plot(seq4,4)
my_h_r_plot(seq4,9)
```

Relationship between sequence no and hydrophobicity profile.

Relationship between sequence no and hydrophobicity profile.
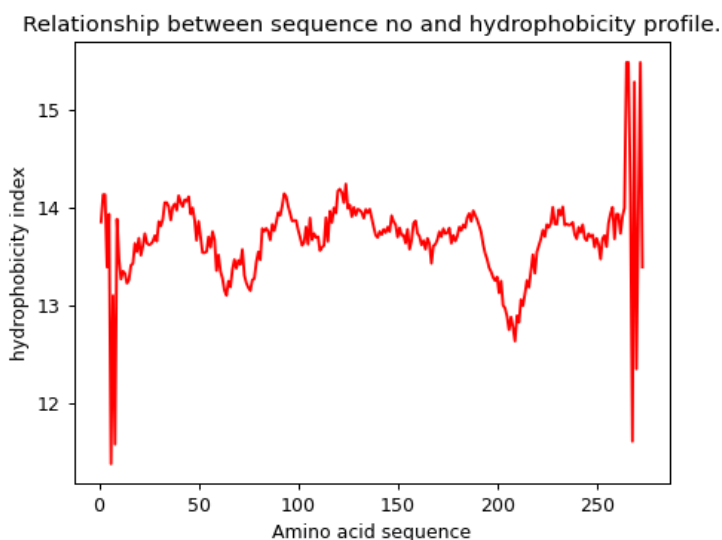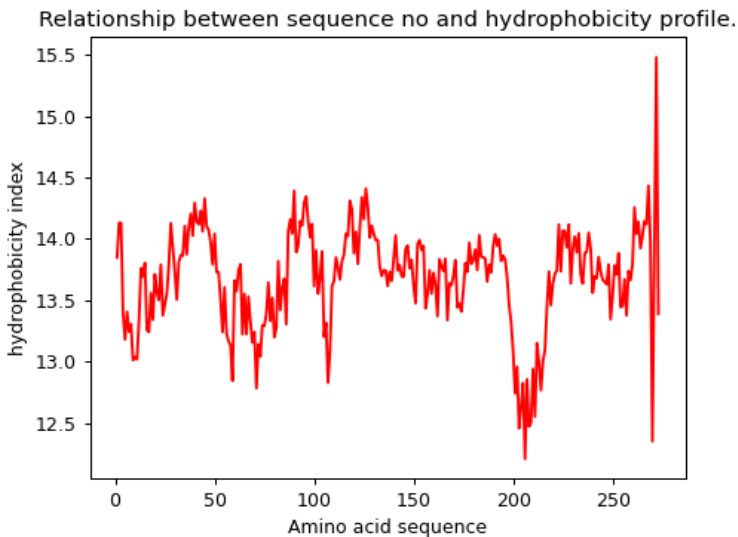
Q4. Use ScanProsite tool (https://prosite.expasy.org/scanprosite/ - select option 2), to search for the patterns a) [SV]-T-[VT]-DERK-{IL} and b) [FILV]Qxxx{RK}Gxxx[RK]xx[FILVWY] in in UniProtKB (Include Swiss-Prot, isoforms). List the number of matches for each pattern.

Hits for USERPAT1{[SV]-T-[VT]-DERK-{IL}} motif on all UniProtKB/Swiss-Prot (release 2021_02 of 07-Apr-21: 564638 entries) database sequences

link for all data matched

https://prosite.expasy.org/cgi-bin/prosite/ScanView.cgi?scanfile=316626646436.scan.gz&sig=[SV]-T-[VT]-[DERK](2)-{IL}

Hits for USERPAT1{[FILV]QXXX{RK}GXXX[RK]XX[FILVWY]} motif on all UniProtKB/Swiss-Prot (release 2021_02 of 07-Apr-21: 564638 entries) database sequences

link for all data matched https://prosite.expasy.org/cgi-bin/prosite/ScanView.cgi?scanfile=201303742718.scan.gz&sig=[FILV]Qxxx{RK}Gxxx[RK]xx[FILVWY]

5. Write a program to identify the patterns (refer Q4) in the sequence database (Q4.fasta). List the matches along with the sequence header and

location of the matches in the sequence.

```
In [291…  import regex as re

          pattern1="[SV]T[VT][DERK]{2}[^IL]"
          pattern2="[FILV]Q...[^RK]G...[RK]..[FILVWY]"
          file1 = open("Q4.fasta")
          l=list()
          for i in file1:
              l.append(i.rstrip())


          for i in range(len(l)):
              if(re.findall(pattern1,l[i])):
                  print("this seq match to pattern ",re.findall(pattern1,l[i]),"in","\n",l[i-1],"\n",l[i])
```

```
this seq match to pattern  ['STTERV'] in
 >4A0C_2|Chains C,E|CULLIN-4B|HOMO SAPIENS (9606)
 MHHHHHHVDENLYFQGGGGRGSAKKLVIKNFKDKPKLPENYTDETWQKLKEAVEAIQNSTSIKYNLEELYQAVENLCSYKISANLYKQLRQICEDHIKAQIHQFREDSLDSVLFLKKIDRCWQNHCRQMIMIRSIFLFL
DRTYVLQNSMLPSIWDMGLELFRAHIISDQKVQNKTIDGILLLIERERNGEAIDRSLLRSLLSMLSDLQIYQDSFEQRFLEETNRLYAAEGQKLMQEREVPEYLHHVNKRLEEEADRLITYLDQTTQKSLIATVEKQLL
GEHLTAILQKGLNNLLDENRIQDLSLLYQLFSRVRGGVQVLLQQWIEYIKAFGSTIVINPEKDKTMVQELLDFKDKVDHIIDICFLKNEKFINAMKEAFETFINKRPNKPAELIAKYVDSKLRAGNKEATDEELEKMLD
KIMIIFRFIYGKDVFEAFYKKDLAKRLLVGKSASVDAEKSMLSKLKHECGAAFTSKLEGMFKDMELSKDIMIQFKQYMQNQNVPGNIELTVNILTMGYWPTYVPMEVHLPPEMVKLQEIFKTFYLGKHSGRKLQWQSTL
GHCVLKAEFKEGKKELQVSLFQTLVLLMFNEGEEFSLEEIKQATGIEDGELRRTLQSLACGKARVLAKNPKGKDIEDGDKFICNDDFKHKLFRIKINQIQMKETVEEQASTTERVFQDRQYQIDAAIVRIMKMRKTLSH
NLLVSEVYNQLKFPVKPADLKKRIESLIDRDYMERDKENPNQYNYIA
this seq match to pattern  ['STTERV'] in
 >4A0C_2|Chains C,E|CULLIN-4B|HOMO SAPIENS (9606)
 MHHHHHHVDENLYFQGGGGRGSAKKLVIKNFKDKPKLPENYTDETWQKLKEAVEAIQNSTSIKYNLEELYQAVENLCSYKISANLYKQLRQICEDHIKAQIHQFREDSLDSVLFLKKIDRCWQNHCRQMIMIRSIFLFL
DRTYVLQNSMLPSIWDMGLELFRAHIISDQKVQNKTIDGILLLIERERNGEAIDRSLLRSLLSMLSDLQIYQDSFEQRFLEETNRLYAAEGQKLMQEREVPEYLHHVNKRLEEEADRLITYLDQTTQKSLIATVEKQLL
GEHLTAILQKGLNNLLDENRIQDLSLLYQLFSRVRGGVQVLLQQWIEYIKAFGSTIVINPEKDKTMVQELLDFKDKVDHIIDICFLKNEKFINAMKEAFETFINKRPNKPAELIAKYVDSKLRAGNKEATDEELEKMLD
KIMIIFRFIYGKDVFEAFYKKDLAKRLLVGKSASVDAEKSMLSKLKHECGAAFTSKLEGMFKDMELSKDIMIQFKQYMQNQNVPGNIELTVNILTMGYWPTYVPMEVHLPPEMVKLQEIFKTFYLGKHSGRKLQWQSTL
GHCVLKAEFKEGKKELQVSLFQTLVLLMFNEGEEFSLEEIKQATGIEDGELRRTLQSLACGKARVLAKNPKGKDIEDGDKFICNDDFKHKLFRIKINQIQMKETVEEQASTTERVFQDRQYQIDAAIVRIMKMRKTLSH
NLLVSEVYNQLKFPVKPADLKKRIESLIDRDYMERDKENPNQYNYIA
this seq match to pattern  ['STTERV'] in
 >4A0K_1|Chain A|CULLIN-4A|HOMO SAPIENS (9606)
 MHHHHHHVDEENLYFQGGGGRGSKKLVIKNFRDRPRLPDNYTQDTWRKLHEAVRAVQSSTSIRYNLEELYQAVENLCSHKVSPMLYKQLRQACEDHVQAQILPFREDSLDSVLFLKKINTCWQDHCRQMIMIRSIFLF
LDRTYVLQNSTLPSIWDMGLELFRTHIISDKMVQSKTIDGILLLIERERSGEAVDRSLLRSLLGMLSDLQVYKDSFELKFLEETNCLYAAEGQRLMQEREVPEYLNHVSKRLEEEGDRVITYLDHSTQKPLIACVEKQL
LGEHLTAILQKGLDHLLDENRVPDLAQMYQLFSRVRGGQQALLQHWSEYIKTFGTAIVINPEKDKDMVQDLLDFKDKVDHIVEVCFQKNERFVNLMKESFETFINKRPNKPAELIAKHVDSKLRAGNKEATDEELERTL
DKIMILFRFIHGKDVFEAFYKKDLAKRLLVGKSASVDAEKSMLSKLKHECGAAFTSKLEGMFKDMELSKDIMVHFKQHMQNQSDSGPIDLTVNILTMGYWPTYTPMEVHLTPEMIKLQEVFKAFYLGKHSGRKLQWQTT
LGHAVLKAEFKEGKKEFQVSLFQTLVLLMFNEGDGFSFEEIKMATGIEDSELRRTLQSLACGKARVLIKSPKGKEVEDGDKFIFNGEFKHKLFRIKINQIQMKETVEEQVSTTERVFQDRQYQIDAAIVRIMKMRKTLG
HNLLVSELYNQLKFPVKPGDLKKRIESLIDRDYMERDKDNPNQYHVA
this seq match to pattern  ['STVRRR'] in
 >5F0J_3|Chain C|Sorting nexin-3|Homo sapiens (9606)
 GAMGSMAETVADTRRLITKPQNLNDAYGPPSNFLEIDVSNPQTVGVGRGRFTTYEIRVKTNLPIFKLKESTVRRRYSDFEWLRSELERESKVVVPPLPGKAFLRQLPFRGDDGIFDDNFIEERKQGLEQFINKVAGHP
LAQNERCLHMFLQDEIIDKSYTPSKIRHA
this seq match to pattern  ['STVRRR'] in
 >5F0L_3|Chain C|Sorting nexin-3|Homo sapiens (9606)
 GAMGSMAETVADTRRLITKPQNLNDAYGPPSNFLEIDVSNPQTVGVGRGRFTTYEIRVKTNLPIFKLKESTVRRRYSDFEWLRSELERESKVVVPPLPGKAFLRQLPFRGDDGIFDDNFIEERKQGLEQFINKVAGHP
LAQNERCLHMFLQDEIIDKSYTPSKIRHA
this seq match to pattern  ['STVRRR'] in
 >5F0M_3|Chain C|Sorting nexin-3|Homo sapiens (9606)
 GAMGSMAETVADTRRLITKPQNLNDAYGPPSNFLEIDVSNPQTVGVGRGRFTTYEIRVKTNLPIFKLKESTVRRRYSDFEWLRSELERESKVVVPPLPGKAFLRQLPFRGDDGIFDDNFIEERKQGLEQFINKVAGHP
LAQNERCLHMFLQDEIIDKSYTPSKIRHA
this seq match to pattern  ['STVRRR'] in
 >5F0P_3|Chain C|Sorting nexin-3|Homo sapiens (9606)
 GAMGSMAETVADTRRLITKPQNLNDAYGPPSNFLEIDVSNPQTVGVGRGRFTTYEIRVKTNLPIFKLKESTVRRRYSDFEWLRSELERESKVVVPPLPGKAFLRQLPFRGDDGIFDDNFIEERKQGLEQFINKVAGHP
LAQNERCLHMFLQDEIIDKSYTPSKIRHA
this seq match to pattern  ['VTVKED'] in
 >5N69_1|Chains A,B|Myosin-7|Bos taurus (9913)
 MVDAEMAAFGEAAPYLRKSEKERLEAQTRPFDLKKDVFVPDDKEEFVKATILSREGGKVTAETEHGKTVTVKEDQVLQQNPPKFDKIEDMAMLTFLHEPAVLYNLKERYASWMIYTYSGLFCVTINPYKWLPVYNAEV
VAAYRGKKRSEAPPHIFSISDNAYQYMLTDRENQSILITGESGAGKTVNTKRVIQYFAVIAAIGDRSKKEQATGKGTLEDQIIQANPALEAFGNAKTVRNDNSSRFGKFIRIHFGATGKLASADIETYLLEKSRVIFQL
KAERDYHIFYQILSNKKPELLDMLLITNNPYDYAFISQGETTVASIDDAEELMATDNAFDVLGFTTEEKNSMYKLTGAIMHFGNMKFKLKQREEQAEPDGTEEADKSAYLMGLNSADLLKGLCHPRVKVGNEYVTKGQN
VQQVVYAKGALAKAVYERMFNWMVTRINATLETKQPRQYFIGVLDIAGFEIFDFNSFEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEWEFIDFGMDLQACIDLIEKPMGIMSILEEECMFPKATDMTFKAKLFDN
HLGKSSNFQKPRNIKGKPEAHFSLIHYAGTVDYNIIGWLQKNKDPLNETVVDLYKKSSLKMLSSLFANYAGFDTPIEKGKGKAKKGSSFQTVSALHRENLNKLMTNLRSTHPHFVRCIIPNETKSPGVIDNPLVMHQLR
CNGVLEGIRICRKGFPNRILYGDFRQRYRILNPAAIPEGQFIDSRKGAEKLLGSLDIDHNQYKFGHTKVFFKAGLLGLLEEMRDERLSRIITRIQAQSRGVLSRMEFKKLLERRDSLLIIQWNIRAFMGVKNWP
this seq match to pattern  ['VTVKED'] in
 >5N69_1|Chains A,B|Myosin-7|Bos taurus (9913)
 MVDAEMAAFGEAAPYLRKSEKERLEAQTRPFDLKKDVFVPDDKEEFVKATILSREGGKVTAETEHGKTVTVKEDQVLQQNPPKFDKIEDMAMLTFLHEPAVLYNLKERYASWMIYTYSGLFCVTINPYKWLPVYNAEV
VAAYRGKKRSEAPPHIFSISDNAYQYMLTDRENQSILITGESGAGKTVNTKRVIQYFAVIAAIGDRSKKEQATGKGTLEDQIIQANPALEAFGNAKTVRNDNSSRFGKFIRIHFGATGKLASADIETYLLEKSRVIFQL
KAERDYHIFYQILSNKKPELLDMLLITNNPYDYAFISQGETTVASIDDAEELMATDNAFDVLGFTTEEKNSMYKLTGAIMHFGNMKFKLKQREEQAEPDGTEEADKSAYLMGLNSADLLKGLCHPRVKVGNEYVTKGQN
VQQVVYAKGALAKAVYERMFNWMVTRINATLETKQPRQYFIGVLDIAGFEIFDFNSFEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEWEFIDFGMDLQACIDLIEKPMGIMSILEEECMFPKATDMTFKAKLFDN
HLGKSSNFQKPRNIKGKPEAHFSLIHYAGTVDYNIIGWLQKNKDPLNETVVDLYKKSSLKMLSSLFANYAGFDTPIEKGKGKAKKGSSFQTVSALHRENLNKLMTNLRSTHPHFVRCIIPNETKSPGVIDNPLVMHQLR
CNGVLEGIRICRKGFPNRILYGDFRQRYRILNPAAIPEGQFIDSRKGAEKLLGSLDIDHNQYKFGHTKVFFKAGLLGLLEEMRDERLSRIITRIQAQSRGVLSRMEFKKLLERRDSLLIIQWNIRAFMGVKNWP
this seq match to pattern  ['VTVKED'] in
 >5N6A_1|Chain A|Myosin-7|Bos taurus (9913)
 MVDAEMAAFGEAAPYLRKSEKERLEAQTRPFDLKKDVFVPDDKEEFVKATILSREGGKVTAETEHGKTVTVKEDQVLQQNPPKFDKIEDMAMLTFLHEPAVLYNLKERYASWMIYTYSGLFCVTINPYKWLPVYNAEV
VAAYRGKKRSEAPPHIFSISDNAYQYMLTDRENQSILITGESGAGKTVNTKRVIQYFAVIAAIGDRSKKEQATGKGTLEDQIIQANPALEAFGNAKTVRNDNSSRFGKFIRIHFGATGKLASADIETYLLEKSRVIFQL
KAERDYHIFYQILSNKKPELLDMLLITNNPYDYAFISQGETTVASIDDAEELMATDNAFDVLGFTTEEKNSMYKLTGAIMHFGNMKFKLKQREEQAEPDGTEEADKSAYLMGLNSADLLKGLCHPRVKVGNEYVTKGQN
VQQVVYAKGALAKAVYERMFNWMVTRINATLETKQPRQYFIGVLDIAGFEIFDFNSFEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEWFIDFGMDLQACIDLIEKPMGIMSILEEECMFPKATDMTFKAKLFDN
HLGKSSNFQKPRNIKGKPEAHFSLIHYAGTVDYNIIGWLQKNKDPLNETVVDLYKKSSLKMLSSLFANYAGFDTPIEKGKGKAKKGSSFQTVSALHRENLNKLMTNLRSTHPHFVRCIIPNETKSPGVIDNPLVMHQLR
CNGVLEGIRICRKGFPNRILYGDFRQRYRILNPAAIPEGQFIDSRKGAEKLLGSLDIDHNQYKFGHTKVFFKAGLLGLLEEMRDERLSRIITRIQAQSRGVLSRMEFKKLLERRDSLLIIQWNIRAFMGVKNWP
this seq match to pattern  ['VTVKED'] in
 >5TBY_1|Chains A,B|Myosin-7|Homo sapiens (9606)
 MGDSEMAVFGAAAPYLRKSEKERLEAQTRPFDLKKDVFVPDDKQEFVKAKIVSREGGKVTAETEYGKTVTVKEDQVMQQNPPKFDKIEDMAMLTFLHEPAVLYNLKDRYGSWMIYTYSGLFCVTVNPYKWLPVYTPEV
VAAYRGKKRSEAPPHIFSISDNAYQYMLTDRENQSILITGESGAGKTVNTKRVIQYFAVIAAIGDRSKKDQSPGKGTLEDQIIQANPALEAFGNAKTVRNDNSSRFGKFIRIHFGATGKLASADIETYLLEKSRVIFQL
KAERDYHIFYQILSNKKPELLDMLLITNNPYDYAFISQGETTVASIDDAEELMATDNAFDVLGFTTEEKNSMYKLTGAIMHFGNMKFKLKQREEQAEPDGTEEADKSAYLMGLNSADLLKGLCHPRVKVGNEYVTKGQN
VQQVIYATGALAKAVYERMFNWMVTRINATLETKQPRQYFIGVLDIAGFEIFDFNSFEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEWTFIDFGMDLQACIDLIEKPMGIMSILEEECMFPKATDMTFKAKLFDN
HLGKSSNFQKPRNIKGKPEAHFSLIHYAGIVDYNIIGWLQKNKDPLNETVVDLYKKSSLLLSLTFANYAGADPIEKGKGKAKKGSSFQTVSALHRENLNKLMTNLRSTHPHFVRCIIPNETKSPGVMDNPLVMHQLR
CNGVLEGIRICRKGFPNRILYGDFRQRYRILNPAAIPEGQFIDSRKGAEKLLSSLDIDHNQYKFGHTKVFFKAGLLGLLEEMRDERLSRIITRIQAQSRGVLARMEYKKLLERRDSLLVIQWNIRAFMGVKNWPWMKLY
FKIKPLLKSAEREKEMASMKEEFTRLKEALEKSEARRKELEEKMVSLLQEKNDLQLQVQAEQDNLADAEERCDQLIKNKIQLEAKVKEMNERLEDEEEMNAELTAKKRKLEDECSELKRDIDDLELTLAKVEKEKHATE
NKVKNLTEEMAGLDEIIAKLTKEKKALQEAHQQALDDLQAEEDKVNTLTKAKVKLEQQVDDLEGSLEQEKKVRMDLERAKRKLEGDLKLTQESIMDLENDKQQLDERLKKKDFELNALNARIEDEQALGSQLQKKLKEL
QARIEELEEELEAERTARAKVEKLRSDLSRELEEISERLEEAGGATSVQIEMNKKREAEFQKMRRDLEEATLQHEATAAALRKKHADSVAELGEQIDNLQRVKQKLEKEKSEFKLELDDVTSNMEQIIKAKANLEKMCR
TLEDQMNEHRSKAEETQRSVNDLTSQRAKLQTENGELSRQLDEKEALISQLTRGKLTYTQQLEDLKRQLEEEVKAKNALAHALQSARHDCDLLREQYEEETEAKAELQRVLSKANSEVAQWRTKYETDAIQRTEELEEA
KKKLAQRLQEAEEAVEAVNAKCSSLEKTKHRLQNEIEDLMVDVERSNAAAAALDKKQRNFDKILAEWKQKYEESQSELESSQKEARSLSTELFKLKNAYEESLEHLETFKRENKNLQEEISDLTEQLGSSGKTIHELEK
VRKQLEAEKMELQSALEEAEASLEHEEGKILRAQLEFNQIKAEIERLAEKDEEMEQAKRNHLRVVDSLQTSLDAETRSRNEALRVKKKMEGDLNEMEIQLSHANRMAAEAQKQVKSLQSLLKDTQIQLDDAVRANDDL
KENIAIVERRNNLLQAELEELRAVVEQTERSRKLAEQELIETSERVQLLHSQNTSLINQKKKMDADLSQLTEVEEAVQECRNAEEKAKKAITDAAMMAEELKKEQDTSAHLERMKKNMEQTIKDLQHRLDEAEQIALK
GGKKQLQKLEARVRELENELEAEQKRNAESVKGMRKSERRIKELTYQTEEDRKNLLRLQDLVDKLQLKVKAYKRQAEEAEEQANTNLSKFRKVQHELDEAEERADIAESQVNKLRAKSRDIGTKGLNEE
this seq match to pattern  ['VTVKED'] in
```

```
>5TBY_1|Chains A,B|Myosin-7|Homo sapiens (9606)
MGDSEMAVFGAAAPYLRKSEKERLEAQTRPFDLKKDVFVPDDKQEFVKAKIVSREGGKVTAETEYGKTVTVKEDQVMQQNPPKFDKIEDMAMLTFLHEPAVLYNLKDRYGSWMIYTYSGLFCVTVNPYKWLPVYTPEV
VAAYRGKKRSEAPPHIFSISDNAYQYMLTDRENQSILITGESGAGKTVNTKRVIQYFAVIAAIGDRSKKDQSPGKGTLEDQIIQANPALEAFGNAKTVRNDNSSRFGKFIRIHFGATGKLASADIETYLLEKSRVIFQL
KAERDYHIFYQILSNKKPELLDMLLITNNPYDYAFISQGETTVASIDDAEELMATDNAFDVLGFTSEEKNSMYKLTGAIMHFGNMKFKLKQREEQAEPDGTEEADKSAYLMGLNSADLLKGLCHPRVKVGNEYVTKGQN
VQQVIYATGALAKAVYERMFNWMVTRINATLETKQPRQYFIGVLDIAGFEIFDFNSFEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEWEFIDFGMDLQACIDLIEKPMGIMSILEEECMFPKATDMTFKAKLFDN
HLGKSANFQKPRNIKGKPEAHFSLIHYAGIVDYNIIGWLQKNKDPLNETVVGLYQKSSLKLLSTLFANYAGADAPIEKGKGKAKKGSSFQTVSALHRENLNKLMTNLRSTHPHFVRCIIPNETKSPGVMDNPLVMHQLR
CNGVLEGIRICRKGFPNRILYGDFRQRYRILNPAAIPEGQFIDSRKGAEKLLSSLDIDHNQYKFGHTKVFFKAGLLGLLEEMRDERLSRIITRIQAQSRGVLARMEYKKLLERRDSLLVIQWNIRAFMGVKNWPWMKLY
FKIKPLLKSAEREKEMASMKEEFTRLKEALEKSEARRKELEEKMVSLLQEKNDLQLQVQAEQDNLADAEERCDQLIKNKIQLEAKVKEMNERLEDEEEMNAELTAKKRKLEDECSELKRDIDDLELTLAKVEKEKHATE
NKVKNLTEEMAGLDEIIAKLTKEKKALQEAHQQALDDLQAEEDKVNTLTKAKVKLEQQVDDLEGSLEQEKKVRMDLERAKRKLEGDLKLTQESIMDLENDKQQLDERLKKKDFELNALNARIEDEQALGSQLQKKLKEL
QARIEELEEELEAERTARAKVEKLRSDLSRELEEISERLEEAGGATSVQIEMNKKREAEFQKMRRDLEEATLQHEATAAALRKKHADSVAELGEQIDNLQRVKQKLEKEKSEFKLELDDVTSNMEQIIKAKANLEKMCR
TLEDQMNEHRSKAEETQRSVNDLTSQRAKLQTENGELSRQLDEKEALISQLTRGKLTYTQQLEDLKRQLEEEVKAKNALAHALQSARHDCDLLREQYEEETEAKAELQRVLSKANSEVAQWRTKYETDAIQRTEELEEA
KKKLAQRLQEAEEAVEAVNAKCSSLEKTKHRLQNEIEDLMVDVERSNAAAAALDKKQRNFDKILAEWKQYEESQSELESSQKEARSLSTELFKLKNAYEESLEHLETFKRENKNLQEEISDLTEQLGSSGKTIHELEK
VRKQLEAEKMELQSALEEAEASLEHEEGKILRAQLEFNQIKAEIERKLAEKDEEMEQAKRNHLRVVDSLQTSLDAETRSRNEALRVKKKMEGDLNEMEIQLSHANRMAAEAQVKSLQSLLKDTQIQLDDAVRANDDL
KENIAIVERRNNLLQAELEELRAVVEQTERSRKLAEQELIETSERVQLLHSQNTSLINQKKKMDADLSQLQTEVEEAVQECRNAEEKAKKAITDAAMMAEELKKEQDTSAHLERMKKNMEQTIKDLQHRLDEAEQIALK
GGKKQLQKLEARVRELENELEAEQKRNAESVKGMRKSERRIKELTYQTEEDRKNLLRLQDLVDKLQLKVKAYKRQAEEAEEQANTNLSKFRKVQHELDEAEERADIAESQVNKLRAKSRDIGTKGLNEE
this seq match to pattern  ['VTVKED'] in
 >6FSA_1|Chains A,B|Myosin-7|Bos taurus (9913)
MVDAEMAAFGEAAPYLRKSEKERLEAQTRPFDLKKDVFVPDDKEEFVKATILSREGGKVTAETEHGKTVTVKEDQVLQQNPPKFDKIEDMAMLTFLHEPAVLYNLKERYASWMIYTYSGLFCVTINPYKWLPVYNAEV
VAAYRGKKRSEAPPHIFSISDNAYQYMLTDRENQSILITGESGAGKTVNTKRVIQYFAVIAAIGDRSKKEQATGKGTLEDQIIQANPALEAFGNAKTVRNDNSSRFGKFIRIHFGATGKLASADIETYLLEKSRVIFQL
KAERDYHIFYQILSNKKPELLDMLLITNNPYDYAFISQGETTVASIDDAEELMATDNAFDVLGFTTEEKNSMYKLTGAIMHFGNMKFKLKQREEQAEPDGTEEADKSAYLMGLNSADLLKGLCHPRVKVGNEYVTKGQN
VQQVVYAKGALAKAVYERMFNWMVTRINATLETKQPRQYFIGVLDIAGFEIFDFNSFEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEWEFIDFGMDLQACIDLIEKPMGIMSILEEECMFPKATDMTFKAKLFDN
HLGKSSNFQKPRNIKGKPEAHFSLIHYAGTVDYNIIGWLQKNKDPLNETVVDLYKKSSLKMLSSLFANYAGFDTPIEKGKGKAKKGSSFQTVSALHRENLNKLMTNLRSTHPHFVRCIIPNETKSPGVIDNPLVMHQLR
CNGVLEGIRICRKGFPNRILYGDFRQRYRILNPAAIPEGQFIDSRKGAEKLLGSLDIDHNQYKFGHTKVFFKAGLLGLLEEMRDERLSRIITRIQAQSRGVLSRMEFKKLLERRDSLLIIQWNIRAFMGVKNWPWMKLY
FKIKPLLKSAETEKEIALMKEEFGRLKEALEKSEARRKELEEKMVSLLQEKNDLQLQVQAEQDNLADAEEERCDQLIKNKIQLEAKVKEMTERLEDEEEMNAELTAKKRKLEDECSELKRDIDDLELTLAKVEKEKHATE
NKVKNLTEEMAGLDEIIAKLTKEKKALQEAHQQALDDLQAEEDKVNTLTKAKVKLEQHVDDLEGSLEQEKKVRMDLERAKRKLEGDLKLTQESIMDLENDKQQLDERLKKKDFELNALNARIEDEQALGSQLQKKLKEL
QARIEELEEELEAERTARAKVEKLRSDLSRELEEISERLEEAGGATSVQIEMNKKREAEFQKMRRDLEEATLQHEATAAALRKKHADSVAELSEQIDNLQRVKQKLEKEKSEFKLELDDVTSNMEQIIKAKANLEKMCR
TLEDQMNEHRSKAEETQRSVNDLTSQRAKLQTENGELSRQLDEKEALISQLTRGKLTYTQQLEDLKRQLEEEVKAKNALAHALQSARHDCDLLREQYEEETEAKAELQRVLSKANSEVAQWRTKYETDAIQRTEELEEA
KKKLAQRLQDAEEAVEAVNAKCSSLEKTKHRLQNEIEDLMVDVERSNAAAAALDKKQRNFDKILAEWKQYEESQSELESSQKEARSLSTELFKLKNAYEESLEHLETFKRENKNLQEEISDLTEQLGSSGKTIHELEK
VRKQLEAEKLELQSALEEAEASLEQEEGKILRAQLEFNQIKAEMERKLAEKDEEMEQAKRNHLRVVDSLQTSLDAETRSRNEALRVKKKMEGDLNEMEIQLSHANRLAAEAQKQVKSLQSLLKDTQIQLDDAVRANDDL
KENIAIVERRNNLLQAELEELRAVVEQTERSRKLAEQELIETSERVQLLHSQNTSLINQKKKMEADLSQLQTEVEEAVQECRNAEEKAKKAITDAAMMAEELKKEQDTSAHLERMKKNMEQTIKDLQHRLDEAEQIALK
GGKKQLQKLEARVRELENELEAEQKRNAESVKGMRKSERRIKELTYQTEEDRKNLLRLQDLVDKLQLKVKAYKRQAEEAEEQANTNLSKFRKVQHELDEAEERADIAESQVNKLRAKSRDIGTKGLNEE
this seq match to pattern  ['VTVKED'] in
 >6FSA_1|Chains A,B|Myosin-7|Bos taurus (9913)
MVDAEMAAFGEAAPYLRKSEKERLEAQTRPFDLKKDVFVPDDKEEFVKATILSREGGKVTAETEHGKTVTVKEDQVLQQNPPKFDKIEDMAMLTFLHEPAVLYNLKERYASWMIYTYSGLFCVTINPYKWLPVYNAEV
VAAYRGKKRSEAPPHIFSISDNAYQYMLTDRENQSILITGESGAGKTVNTKRVIQYFAVIAAIGDRSKKEQATGKGTLEDQIIQANPALEAFGNAKTVRNDNSSRFGKFIRIHFGATGKLASADIETYLLLEKSRVIFQL
KAERDYHIFYQILSNKKPELLDMLLITNNPYDYAFISQGETTVASIDDAEELMATDNAFDVLGFTTEEKNSMYKLTGAIMHFGNMKFKLKQREEQAEPDGTEEADKSAYLMGLNSADLLKGLCHPRVKVGNEYVTKGQN
VQQVVYAKGALAKAVYERMFNWMVTRINATLETKQPRQYFIGVLDIAGFEIFDFNSFEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEWEFIDFGMDLQACIDLIEKPMGIMSILEEECMFPKATDMTFKAKLFDN
HLGKSSNFQKPRNIKGKPEAHFSLIHYAGTVDYNIIGWLQKNKDPLNETVVDLYKKSSLKMLSSLFANYAGFDTPIEKGKGKAKKGSSFQTVSALHRENLNKLMTNLRSTHPHFVRCIIPNETKSPGVIDNPLVMHQLR
CNGVLEGIRICRKGFPNRILYGDFRQRYRILNPAAIPEGQFIDSRKGAEKLLGSLDIDHNQYKFGHTKVFFKAGLLGLLEEMRDERLSRIITRIQAQSRGVLSRMEFKKLLERRDSLLIIQWNIRAFMGVKNWPWMKLY
FKIKPLLKSAETEKEIALMKEEFGRLKEALEKSEARRKELEEKMVSLLQEKNDLQLQVQAEQDNLADAEEERCDQLIKNKIQLEAKVKEMTERLEDEEEMNAELTAKKRKLEDECSELKRDIDDLELTLAKVEKEKHATE
NKVKNLTEEMAGLDEIIAKLTKEKKALQEAHQQALDDLQAEEDKVNTLTKAKVKLEQHVDDLEGSLEQEKKVRMDLERAKRKLEGDLKLTQESIMDLENDKQQLDERLKKKDFELNALNARIEDEQALGSQLQKKLKEL
QARIEELEEELEAERTARAKVEKLRSDLSRELEEISERLEEAGGATSVQIEMNKKREAEFQKMRRDLEEATLQHEATAAALRKKHADSVAELSEQIDNLQRVKQKLEKEKSEFKLELDDVTSNMEQIIKAKANLEKMCR
TLEDQMNEHRSKAEETQRSVNDLTSQRAKLQTENGELSRQLDEKEALISQLTRGKLTYTQQLEDLKRQLEEEVKAKNALAHALQSARHDCDLLREQYEEETEAKAELQRVLSKANSEVAQWRTKYETDAIQRTEELEEA
KKKLAQRLQDAEEAVEAVNAKCSSLEKTKHRLQNEIEDLMVDVERSNAAAAALDKKQRNFDKILAEWKQYEESQSELESSQKEARSLSTELFKLKNAYEESLEHLETFKRENKNLQEEISDLTEQLGSSGKTIHELEK
VRKQLEAEKLELQSALEEAEASLEQEEGKILRAQLEFNQIKAEMERKLAEKDEEMEQAKRNHLRVVDSLQTSLDAETRSRNEALRVKKKMEGDLNEMEIQLSHANRLAAEAQKQVKSLQSLLKDTQIQLDDAVRANDDL
KENIAIVERRNNLLQAELEELRAVVEQTERSRKLAEQELIETSERVQLLHSQNTSLINQKKKMEADLSQLQTEVEEAVQECRNAEEKAKKAITDAAMMAEELKKEQDTSAHLERMKKNMEQTIKDLQHRLDEAEQIALK
GGKKQLQKLEARVRELENELEAEQKRNAESVKGMRKSERRIKELTYQTEEDRKNLLRLQDLVDKLQLKVKAYKRQAEEAEEQANTNLSKFRKVQHELDEAEERADIAESQVNKLRAKSRDIGTKGLNEE
this seq match to pattern  ['VTVKED'] in
 >6X5Z_3|Chains D,G,J|Myosin-7|Bos taurus (9913)
MVDAEMAAFGEAAPYLRKSEKERLEAQTRPFDLKKDVFVPDDKEEFVKATILSREGGKVTAETEHGKTVTVKEDQVLQQNPPKFDKIEDMAMLTFLHEPAVLYNLKERYASWMIYTYSGLFCVTINPYKWLPVYNAEV
VAAYRGKKRSEAPPHIFSISDNAYQYMLTDRENQSILITGESGAGKTVNTKRVIQYFAVIAAIGDRSKKEQATGKGTLEDQIIQANPALEAFGNAKTVRNDNSSRFGKFIRIHFGATGKLASADIETYLLEKSRVIFQL
KAERDYHIFYQILSNKKPELLDMLLITNNPYDYAFISQGETTVASIDDAEELMATDNAFDVLGFTTEEKNSMYKLTGAIMHFGNMKFKLKQREEQAEPDGTEEADKSAYLMGLNSADLLKGLCHPRVKVGNEYVTKGQN
VQQVVYAKGALAKAVYERMFNWMVTRINATLETKQPRQYFIGVLDIAGFEIFDFNSFEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEWEFIDFGMDLQACIDLIEKPMGIMSILEEECMFPKATDMTFKAKLFDN
HLGKSSNFQKPRNIKGKPEAHFSLIHYAGTVDYNIIGWLQKNKDPLNETVVDLYKKSSLKMLSSLFANYAGFDTPIEKGKGKAKKGSSFQTVSALHRENLNKLMTNLRSTHPHFVRCIIPNETKSPGVIDNPLVMHQLR
CNGVLEGIRICRKGFPNRILYGDFRQRYRILNPAAIPEGQFIDSRKGAEKLLGSLDIDHNQYKFGHTKVFFKAGLLGLLEEMRDERLSRIITRIQAQSRGVLSRMEFKKLLERRDSLLIIQWNIRAFMGVKNWPWMKLY
FKIKPLLKSAETEKEIA
this seq match to pattern  ['VTVKED'] in
 >6X5Z_3|Chains D,G,J|Myosin-7|Bos taurus (9913)
MVDAEMAAFGEAAPYLRKSEKERLEAQTRPFDLKKDVFVPDDKEEFVKATILSREGGKVTAETEHGKTVTVKEDQVLQQNPPKFDKIEDMAMLTFLHEPAVLYNLKERYASWMIYTYSGLFCVTINPYKWLPVYNAEV
VAAYRGKKRSEAPPHIFSISDNAYQYMLTDRENQSILITGESGAGKTVNTKRVIQYFAVIAAIGDRSKKEQATGKGTLEDQIIQANPALEAFGNAKTVRNDNSSRFGKFIRIHFGATGKLASADIETYLLEKSRVIFQL
KAERDYHIFYQILSNKKPELLDMLLITNNPYDYAFISQGETTVASIDDAEELMATDNAFDVLGFTTEEKNSMYKLTGAIMHFGNMKFKLKQREEQAEPDGTEEADKSAYLMGLNSADLLKGLCHPRVKVGNEYVTKGQN
VQQVVYAKGALAKAVYERMFNWMVTRINATLETKQPRQYFIGVLDIAGFEIFDFNSFEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEWEFIDFGMDLQACIDLIEKPMGIMSILEEECMFPKATDMTFKAKLFDN
HLGKSSNFQKPRNIKGKPEAHFSLIHYAGTVDYNIIGWLQKNKDPLNETVVDLYKKSSLKMLSSLFANYAGFDTPIEKGKGKAKKGSSFQTVSALHRENLNKLMTNLRSTHPHFVRCIIPNETKSPGVIDNPLVMHQLR
CNGVLEGIRICRKGFPNRILYGDFRQRYRILNPAAIPEGQFIDSRKGAEKLLGSLDIDHNQYKFGHTKVFFKAGLLGLLEEMRDERLSRIITRIQAQSRGVLSRMEFKKLLERRDSLLIIQWNIRAFMGVKNWPWMKLY
FKIKPLLKSAETEKEIA
this seq match to pattern  ['VTVKED'] in
 >6X5Z_3|Chains D,G,J|Myosin-7|Bos taurus (9913)
MVDAEMAAFGEAAPYLRKSEKERLEAQTRPFDLKKDVFVPDDKEEFVKATILSREGGKVTAETEHGKTVTVKEDQVLQQNPPKFDKIEDMAMLTFLHEPAVLYNLKERYASWMIYTYSGLFCVTINPYKWLPVYNAEV
VAAYRGKKRSEAPPHIFSISDNAYQYMLTDRENQSILITGESGAGKTVNTKRVIQYFAVIAAIGDRSKKEQATGKGTLEDQIIQANPALEAFGNAKTVRNDNSSRFGKFIRIHFGATGKLASADIETYLLEKSRVIFQL
KAERDYHIFYQILSNKKPELLDMLLITNNPYDYAFISQGETTVASIDDAEELMATDNAFDVLGFTTEEKNSMYKLTGAIMHFGNMKFKLKQREEQAEPDGTEEADKSAYLMGLNSADLLKGLCHPRVKVGNEYVTKGQN
VQQVVYAKGALAKAVYERMFNWMVTRINATLETKQPRQYFIGVLDIAGFEIFDFNSFEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEWEFIDFGMDLQACIDLIEKPMGIMSILEEECMFPKATDMTFKAKLFDN
HLGKSSNFQKPRNIKGKPEAHFSLIHYAGTVDYNIIGWLQKNKDPLNETVVDLYKKSSLKMLSSLFANYAGFDTPIEKGKGKAKKGSSFQTVSALHRENLNKLMTNLRSTHPHFVRCIIPNETKSPGVIDNPLVMHQLR
CNGVLEGIRICRKGFPNRILYGDFRQRYRILNPAAIPEGQFIDSRKGAEKLLGSLDIDHNQYKFGHTKVFFKAGLLGLLEEMRDERLSRIITRIQAQSRGVLSRMEFKKLLERRDSLLIIQWNIRAFMGVKNWPWMKLY
FKIKPLLKSAETEKEIA
this seq match to pattern  ['VTVKED'] in
 >7JH7_2|Chains F,G,H|Myosin-7|Sus scrofa (9823)
MVDAEMAAFGEAAPYLRKSEKERLEAQTRPFDLKKDVYVPDDKEEFVKAKILSREGGKVTAETEHGKTVTVKEDQVLQQNPPKFDKIEDMAMLTFLHEPAVLYNLKERYASWMIYTYSGLFCVTINPYKWLPVYNAEV
VAAYRGKKRSEAPPHIFSISDNAYQYMLTDRENQSILITGESGAGKTVNTKRVIQYFAVIAAIGDRSKKEQTPGKGTLEDQIIQANPALEAFGNAKTVRNDNSSRFGKFIRIHFGATGKLASADIETYLLEKSRVIFQL
KAERDYHIFYQILSNKKPELLDMLLITNNPYDYAFISQGETTVASIDDAEELMATDNAFDVLGFTSEEKNSMYKLTGAIMHFGNMKFKLKQREEQAEPDGTEEADKSAYLMGLNSADLLKGLCHPRVKVGNEYVTKGQN
VQQVMYATGALAKAVYEKMFNWMVTRINTTLETKQPRQYFIGVLDIAGFEIFDFNSFEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIEWEFIDFGMDLQACIDLIEKPMGIMSILEEECMFPKATDMTFKAKLYDN
HLGKSSNFQKPRNIKGRPEAHFALIHYAGTVDYNIIGWLQKNKDPLNETVVDLYKKSSLKLLSNLFANYAGADTPVEKGKGKAKKGSSFQTVSALHRENLNKLMTNLRSTHPHFVRCIIPNETKSPGVIDNPLVMHQLR
CNGVLEGIRICRKGFPNRILYGDFRQRYRILNPAAIPEGQFIDSRKGAEKLLGSLDIDHNQYKFGHTKVFFKAGLLGLLEEMRDERLSRIITRIQAQSRGVLSRMEFKKLLERRDSLLIIQWNIRAFMGVKNWPWMKLY
FKIKPLLKSAETEKEMATMKEEFGRLKEALEKSEARRKELEEKMVSLLQEKNDLQLQVQAEQDNLADAEEERCDQLIKNKIQLEAKVKEMTERLEDEEEMNAELTAKKRKLEDECSELKRDIDDLELTLAKVEKEKHATE
NKVKNLTEEMAGLDEIIAKLTKEKKALQEAHQQALDDLQAEEDKVNTLTKAKVKLEQHVDDLEGSLEQEKKVRMDLERAKRKLEGDLKLTQESIMDLENDKQQLDERLKKKDFELNALNARIEDEQALGSQLQKKLKEL
QARIEELEEELEAERTARAKVEKLRSDLSRELEEISERLEEAGGATSVQIEMNKKREAEFQKMRRDLEEATLQHEATAAALRKKHADSVAELGEQIDNLQRVKQKLEKEKSEFKLSLDDVTSNMEQIIKAKANLEKMCR
TLEDQMNEHRSKAEETQRSVNDLTSQRAKLQTENGELSRQLDEKEALISQLTRGKLTYTQQLEDLKRQLEEEVKAKNALAHALQSARHDCDLLREQYEEETEAKAELQRVLSKANSEVAQWRTKYETDAIQRTEELEEA
KKKLAQRLQDAEEAVEAVNAKCSSLEKTKHRLQNEIEDLMVDVERSNAAAAALDKKQRNFDKILAEWKQKYEESQSELESSQKEARSLSTELFKLKNAYEESLEHLETFKRENKNLQEEISDLTEQLGSSGKTIHELEK
VRKQLEAEKLELQSALEEAEASLEHEEGKILRAQLEFNQIKAEMERKLAEKDEEMEQAKRNHLRVVDSLQTSLDAETRSRNEALRVKKKMEGDLNEMEIQLSHANRMAAEAQKQVKSLQSLLKDTQIQLDDAVRANDDL
KENIAIVERRNNLLQAELEELRAVVEQTERSRKLAEQELIETSERVQLLHSQNTSLINQKKKMEADLSQLQTEVEEAVQECRNAEEKAKKAITDAAMMAEELKKEQDTSAHLERMKKNMEQTIKDLQHRLDEAEQIALK
GGKKQLQKLEARVRELENELEAEQKRNAESVKGMRKSERRIKELTYQTEEDRKNLLRLQDLVDKLQLKVKAYKRQAEEAEEQANTNLSKFRKVQHELDEAEERADIAESQVNKLRAKSRDIGTKGLNEE
this seq match to pattern  ['STVEKT'] in
 >1A8J_1|Chains H,L|IMMUNOGLOBULIN LAMBDA LIGHT CHAIN DIMER (MCG)|Homo sapiens (9606)
QSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1A8J_1|Chains H,L|IMMUNOGLOBULIN LAMBDA LIGHT CHAIN DIMER (MCG)|Homo sapiens (9606)
QSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
```

```
this seq match to pattern  ['STVEKT'] in
 >1ADQ_2|Chain L|IGM-LAMBDA RF-AN FAB (LIGHT CHAIN)|Homo sapiens (9606)
 YVLTQPPSVSVAPGQTARITCGGNNIGSKSVHWYQQKPGQAPVLVVYDDSDRPPGIPERFSGSNSGNTATLTISRVEAGDEADYYCQVWDSSSDHAVFGGGTKLTVLGQPKAAPSVTLFPPSSEELQANKATLVCLIS
DFYPGAVTVAWKADGSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHKSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEEN'] in
 >1AIV_1|Chain A|OVOTRANSFERRIN|Gallus gallus (9031)
 APPKSVIRWCTISSPEEKKCNNLRDLTQQERISLTCVQKATYLDCIKAIANNEADAISLDGGQAFEAGLAPYKLKPIAAEVYEHTEGSTTSYYAVAVVKKGTEFTVNDLQGKTSCHTGLGRSAGWNIPIGTLLHRGAI
EWEGIESGSVEQAVAKFFSASCVPGATIEQKLCRQCKGDPKTKCARNAPYSGYSGAFHCLKDGKGDVAFVKHTTVNENAPDQKDEYELLCLDGSRQPVDNYKTCNWARVAAHAVVARDDNKVEDIWSFLSKAQSDFGVD
TKSDFHLFGPPGKKDPVLKDLLFKDSAIMLKRVPSLMDSQLYLGFEYYSAIQSMRKDQLTPSPRENRIQWCAVGKDEKSKCDRWSVVSNGDVECTVVDETKDCIIKIMKGEADAVALDGGLVYTAGVCGLVPVMAERYD
DESQCSKTDERPASYFAVAVARKDSNVNWNNLKGKKSCHTAVGRTAGWVIPMGLIHNRTGTCNFDEYFSEGCAPGSPPNSRLCQLCQGSGGIPPEKCVASSHEKYFGYTGALRCLVEKGDVAFIQHSTVEENTGGKNKA
DWAKNLQMDDFELLCTDGRRANVMDYRECNLAEVPTHAVVVRPEKANKIRDLLERQEKRFGVNGSEKSKFMMFESQNKDLLFKDLTKCLFKVREGTTYKEFLGDKFYTVISSLKTCNPSDILQMCSFLEGK
this seq match to pattern  ['STVEKT'] in
 >1AQK_1|Chain L|FAB B7-15A2|Homo sapiens (9606)
 QNVLTQPPSVSGAPGQRVTISCTGSNSNIGAGFTVHWYQHLPGTAPKLLIFANTNRPSGVPDRFSGSKSGTSASLAITGLQAEDEADYYCQSYDSSLSARFGGGTRLTVLGQPKAAPSVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADSSPVNAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHKSYSCQVTHEGSTVEKTVAPAECS
this seq match to pattern  ['STVEKT'] in
 >1BJM_1|Chains A,B|LOC - LAMBDA 1 TYPE LIGHT-CHAIN DIMER|Homo sapiens (9606)
 QSVLTQPPSASGTPGQRVTISCSGSSSNIGENSVTWYQHLSGTAPKLLIYEDNSRASGVSDRFSASKSGTSASLAISGLQPEDETDYYCAAWDDSLDVAVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1BJM_1|Chains A,B|LOC - LAMBDA 1 TYPE LIGHT-CHAIN DIMER|Homo sapiens (9606)
 QSVLTQPPSASGTPGQRVTISCSGSSSNIGENSVTWYQHLSGTAPKLLIYEDNSRASGVSDRFSASKSGTSASLAISGLQPEDETDYYCAAWDDSLDVAVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1DCL_1|Chains A,B|MCG|Homo sapiens (9606)
 PSALTQPPSASGSLGQSVTISCTGTSSNVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1DCL_1|Chains A,B|MCG|Homo sapiens (9606)
 PSALTQPPSASGSLGQSVTISCTGTSSNVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVDDD'] in
 >1DPU_1|Chain A|REPLICATION PROTEIN A (RPA32) C-TERMINAL DOMAIN|Homo sapiens (9606)
 ANSQPSAGRAPISNPGMSEAGNFGGNSFMPANGLTVAQNQVLNLIKACPRPEGLNFQDLKNQLKHMSVSSIKQAVDFLSNEGHIYSTVDDDHFKSTDAE
this seq match to pattern  ['STVKRS'] in
 >1E8I_1|Chains A,B|EARLY ACTIVATION ANTIGEN CD69|HOMO SAPIENS (9606)
 VSSCSEDWVGYQRKCYFISTVKRSWTSAQNACSEHGATLAVIDSEKDMNFLKRYAGREEHWVGLKKEPGHPWKWSNGKEFNNWFNVTGSDKCVFLKNTEVSSMECEKNLYWICNKPYK
this seq match to pattern  ['STVKRS'] in
 >1E8I_1|Chains A,B|EARLY ACTIVATION ANTIGEN CD69|HOMO SAPIENS (9606)
 VSSCSEDWVGYQRKCYFISTVKRSWTSAQNACSEHGATLAVIDSEKDMNFLKRYAGREEHWVGLKKEPGHPWKWSNGKEFNNWFNVTGSDKCVFLKNTEVSSMECEKNLYWICNKPYK
this seq match to pattern  ['STVDKP'] in
 >1FM2_2|Chain B|GLUTARYL 7-AMINOCEPHALOSPORANIC ACID ACYLASE|Brevundimonas diminuta (293)
 SNSWAVAPGKTANGNALLLQNPHLSWTTDYFTYYEAHLVTPDFEIYGATQIGLPVIRFAFNQRMGITNTVNGMVGATNYRLTLQDGGYLYDGQVRPFERRQASYRLRQADGSTVDKPLEIRSSVHGPVFERADGTAVA
VRVAGLDRPGMLEQYFDMITAHSFDDYEAAMARMQVPTFNIVYADREGTINYSFNGVAPKRAEGDIAFWQGNVPGDSSRYLWTETHPLDDLPRVTNPPGGFVQNSNDPPWTPTWPVTYCPANHPSYLAPQTPHSLRAQQ
SVRLMSENDDLTLERFMALQFSHRAVMADRTLPDLIPAALIDPDPEVQAAARLLAAWDRDFTSDSRAALLFEEWARLFAGQNFAGQAAFATPWSLDKPVSTPYGVRDPKAAVDQLRTAIANTKRKYGAIDRPFGDASRM
ILNDVNVPGAAGYGNLGSFRVFTWSDPDENGIRTPVHGETWVAMIEFSTPVRAYGLMSYGNSRQPGTTHYSDQIERVSRADFRELLLRREQVEAAVQERTPFNF
this seq match to pattern  ['STVKRS'] in
 >1FM5_1|Chain A|EARLY ACTIVATION ANTIGEN CD69|Homo sapiens (9606)
 MASGQYNCPGQYTFSMPSDSHVSSCSEDWVGYQRKCYFISTVKRSWTSAQNACSEHGATLAVIDSEKDMNFLKRYAGREEHWVGLKKEPGHPWKWSNGKEFNNWFNVTGSDKCVFLKNTEVSSMECEKNLYWICNKPY
K
this seq match to pattern  ['STVEDK'] in
 >1FNT_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|PROTEASOME ACTIVATOR PROTEIN PA26|Trypanosoma brucei (5691)
 MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGSVDAESGKTKGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGTDHMVS
this seq match to pattern  ['STVEDK'] in
 >1FNT_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|PROTEASOME ACTIVATOR PROTEIN PA26|Trypanosoma brucei (5691)
 MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGSVDAESGKTKGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGTDHMVS
this seq match to pattern  ['STVEDK'] in
 >1FNT_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|PROTEASOME ACTIVATOR PROTEIN PA26|Trypanosoma brucei (5691)
 MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGSVDAESGKTKGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGTDHMVS
this seq match to pattern  ['STVEDK'] in
 >1FNT_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|PROTEASOME ACTIVATOR PROTEIN PA26|Trypanosoma brucei (5691)
 MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGSVDAESGKTKGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGTDHMVS
this seq match to pattern  ['STVEDK'] in
 >1FNT_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|PROTEASOME ACTIVATOR PROTEIN PA26|Trypanosoma brucei (5691)
 MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGSVDAESGKTKGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGTDHMVS
this seq match to pattern  ['STVEDK'] in
 >1FNT_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|PROTEASOME ACTIVATOR PROTEIN PA26|Trypanosoma brucei (5691)
 MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGSVDAESGKTKGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGTDHMVS
this seq match to pattern  ['STVEDK'] in
 >1FNT_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|PROTEASOME ACTIVATOR PROTEIN PA26|Trypanosoma brucei (5691)
 MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGSVDAESGKTKGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGTDHMVS
this seq match to pattern  ['STVEDK'] in
 >1FNT_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|PROTEASOME ACTIVATOR PROTEIN PA26|Trypanosoma brucei (5691)
 MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGSVDAESGKTKGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGTDHMVS
this seq match to pattern  ['STVEDK'] in
 >1FNT_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|PROTEASOME ACTIVATOR PROTEIN PA26|Trypanosoma brucei (5691)
 MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGSVDAESGKTKGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGTDHMVS
this seq match to pattern  ['STVEDK'] in
 >1FNT_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|PROTEASOME ACTIVATOR PROTEIN PA26|Trypanosoma brucei (5691)
 MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGSVDAESGKTKGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGTDHMVS
this seq match to pattern  ['STVEDK'] in
 >1FNT_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|PROTEASOME ACTIVATOR PROTEIN PA26|Trypanosoma brucei (5691)
 MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGSVDAESGKTKGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGTDHMVS
```

```
this seq match to pattern  ['STVEDK'] in
>1FNT_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|PROTEASOME ACTIVATOR PROTEIN PA26|Trypanosoma brucei (5691)
 MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGSVDAESGKTKGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGTDHMVS
this seq match to pattern  ['STVEDK'] in
>1FNT_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|PROTEASOME ACTIVATOR PROTEIN PA26|Trypanosoma brucei (5691)
 MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGSVDAESGKTKGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGTDHMVS
this seq match to pattern  ['STVRRC'] in
>1FO9_1|Chain A|ALPHA-1,3-MANNOSYL-GLYCOPROTEIN BETA-1,2-N-ACETYLGLUCOSAMINYLTRANSFERASE|Oryctolagus cuniculus (9986)
 SRLQELAVIPILVIACDRSTVRRCLDKLLHYRPSAELFPIIVSQDCGHEETAQVIASYGSAVTHIRQPDLSNIAVQPDHRKFQGYYKIARHYRWALGQIFHNFNYPAAVVVEDDLEVAPDFFEYFQATYPLLKADPSL
WCVSAWNDNGKEQMVDSSKPELLYRTDFFPGLGWLLLAELWAELEPKWPKAFWDDWMRRPEQRKGRACVRPEISRTMTFGRKGVSHGQFFDQHLKFIKLNQQFVPFTQLDLSYLQQEAYDRDFLARVYGAPQLQVEKVR
TNDRKELGEVRVQYTGRDSFKAFAKALGVMDDLKSGVPRAGYRGIVTFLFRGRRVHLAPPQTWDGYDPSWT
this seq match to pattern  ['STVRRC'] in
>1FOA_1|Chain A|ALPHA-1,3-MANNOSYL-GLYCOPROTEIN BETA-1,2-N-ACETYLGLUCOSAMINYLTRANSFERASE|Oryctolagus cuniculus (9986)
 SRLQELAVIPILVIACDRSTVRRCLDKLLHYRPSAELFPIIVSQDCGHEETAQVIASYGSAVTHIRQPDLSNIAVQPDHRKFQGYYKIARHYRWALGQIFHNFNYPAAVVVEDDLEVAPDFFEYFQATYPLLKADPSL
WCVSAWNDNGKEQMVDSSKPELLYRTDFFPGLGWLLLAELWAELEPKWPKAFWDDWMRRPEQRKGRACVRPEISRTMTFGRKGVSHGQFFDQHLKFIKLNQQFVPFTQLDLSYLQQEAYDRDFLARVYGAPQLQVEKVR
TNDRKELGEVRVQYTGRDSFKAFAKALGVMDDLKSGVPRAGYRGIVTFLFRGRRVHLAPPQTWDGYDPSWT
this seq match to pattern  ['STVEEE'] in
>1GPJ_1|Chain A|Glutamyl-tRNA reductase|Methanopyrus kandleri (2320)
 MEDLVSVGITHKEAEVEELEKARFESDEAVRDIVESFGLSGSVLLQTSNRVEVYASGARDRAEELGDLIHDDAWVKRGSEAVRHLFRVASGLESMMVGEQEILRQVKKAYDRAARLGTLDEALKIVFRRAINLGKRAR
EETRISEGAVSIGSAAVELAERELGSLHDKTVLVVGAGEMGKTVAKSLVDRGVRAVLVANRTYERAVELARDLGGEAVRFDELVDHLARSDVVVSATAAPHPVIHVDDVREALRKRDRRSPILIIDIANPRDVEEGVEN
IEDVEVRTIDDLRVIARENLERRRKEIPKVEKLIEEELSTVEEELEKLKERRLVADVAKSLHEIKDRELERALRRLKTGDPENVLQDFAEAYTKRLINVLTSAIMELPDEYRRAASRALRRASELNG
this seq match to pattern  ['STVDKV'] in
>1IQ7_1|Chain A|Ovotransferrin|Gallus gallus (9031)
 ENRIQWCAVGKDEKSKCDRWSVVSNGDVECTVVDETKDCIIKIMKGEADAVALDGGLVYTAGVCGLVPVMAERYDDESQCSKTDERPASYFAVAVARKDSNVNWNNLKGKKSCHTAVGRTAGWVIPMGLIHNRTGTCN
FDEYFSEGCAPGSPPNSRLCQLCQGSGGIPPEKCVASSHEKYFGYTGALRCLVEKGDVAFIQHSTVEENTGGKNKADWAKNLQMDDFELLCTDGRRANVMDYRECNLAEVPTHAVVVRPEKANKIRDLLERQEKRFGVN
GSEKSKFMMFESQNKDLLFKDLTKCLFKVREGTTYKEFLGDKFYTVISSLKTCNPSDILQMCSFLEGK
this seq match to pattern  ['STVDKV'] in
>1JT1_1|Chain A|FEZ-1, class B3 metallo-beta-lactamase|Fluoribacter gormanii (464)
 AYPMPNPFPPFRIAGNLYYVGTDDLASYLIVTPRGNILINSDLEANVPMIKASIKKLGFKFSDTKILLISHAHFDHAAGSELIKQQTKAKYMVMDEDVSVILSGGKSDFHYANDSSTYFTQSTVDKVLHDGERVELGG
TVLTAHLTPGHTRGCTTWTMKLKDHGKQYQAVIIGSIGVNPGYKLVDNITYPKIAEDYKHSIKVLESMRCDIFLGSHAGMFDLKNKYVLLSKGQNNPFVDPTGCKNYIEQKANDFYTELKKQETG
this seq match to pattern  ['STVEKT'] in
>1JVK_1|Chains A,B|IMMUNOGLOBULIN LAMBDA LIGHT CHAIN|Homo sapiens (9606)
 QTALTQPASVSGSPGQSITVSCTGVSSIVGSYNLVSWYQQHPGKAPKLLTYEVNKRPSGVSDRFSGSKSGNSASLTISGLQAEDEADYYCSSYDGSSTSVVFGGGTKLTVLGQPKAAPSVTLFPPSSEELQANKATLV
CLISDFYPGAVTVAWKADSSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTAC
this seq match to pattern  ['STVEKT'] in
>1JVK_1|Chains A,B|IMMUNOGLOBULIN LAMBDA LIGHT CHAIN|Homo sapiens (9606)
 QTALTQPASVSGSPGQSITVSCTGVSSIVGSYNLVSWYQQHPGKAPKLLTYEVNKRPSGVSDRFSGSKSGNSASLTISGLQAEDEADYYCSSYDGSSTSVVFGGGTKLTVLGQPKAAPSVTLFPPSSEELQANKATLV
CLISDFYPGAVTVAWKADSSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTAC
this seq match to pattern  ['STVDKP'] in
>1JVZ_2|Chain B|cephalosporin acylase beta chain|Brevundimonas diminuta (293)
 SNSWAVAPGKTANGNALLLQNPHLSWTTDYFTYYEAHLVTPDFEIYGATQIGLPVIRFAFNQRMGITNTVNGMVGATNYRLTLQDGGYLYDGQVRPFERRQASYRLRQADGSTVDKPLEIRSSVHGPVFERADGTAVA
VRVAGLDRPGMLEQYFDMITAHSFDDYEAAMARMQVPTFNIVYADREGTINYSFNGVAPKRAEGDIAFWQGNVPGDSSRYLWTETHPLDDLPRVTNPPGGFVQNSNDPPWTPTWPVTYCPANHPSYLAPQTPHSLRAQQ
SVRLMSENDDLTLERFMALQFSHRAVMADRTLPDLIPAALIDPDPEVQAAARLLAAWDRDFTSDSRAALLFEEWARLFAGQNFAGQAAFATPWSLDKPVSTPYGVRDPKAAVDQLRTAIANTKRKYGAIDRPFGDASRM
ILNDVNVPGAAGYGNLGSFRVFTWSDPDENGIRTPVHGETWVAMIEFSTPVRAYGLMSYGNSRQPGTTHYSDQIERVSRADFRELLLRREQVEAAVQERTPFNF
this seq match to pattern  ['STVDKV'] in
>1K07_1|Chains A,B|FEZ-1 beta-lactamase|Fluoribacter gormanii (464)
 AYPMPNPFPPFRIAGNLYYVGTDDLASYLIVTPRGNILINSDLEANVPMIKASIKKLGFKFSDTKILLISHAHFDHAAGSELIKQQTKAKYMVMDEDVSVILSGGKSDFHYANDSSTYFTQSTVDKVLHDGERVELGG
TVLTAHLTPGHTRGCTTWTMKLKDHGKQYQAVIIGSIGVNPGYKLVDNITYPKIAEDYKHSIKVLESMRCDIFLGSHAGMFDLKNKYVLLSKGQNNPFVDPTGCKNYIEQKANDFYTELKKQETG
this seq match to pattern  ['STVDKV'] in
>1K07_1|Chains A,B|FEZ-1 beta-lactamase|Fluoribacter gormanii (464)
 AYPMPNPFPPFRIAGNLYYVGTDDLASYLIVTPRGNILINSDLEANVPMIKASIKKLGFKFSDTKILLISHAHFDHAAGSELIKQQTKAKYMVMDEDVSVILSGGKSDFHYANDSSTYFTQSTVDKVLHDGERVELGG
TVLTAHLTPGHTRGCTTWTMKLKDHGKQYQAVIIGSIGVNPGYKLVDNITYPKIAEDYKHSIKVLESMRCDIFLGSHAGMFDLKNKYVLLSKGQNNPFVDPTGCKNYIEQKANDFYTELKKQETG
this seq match to pattern  ['STVDKP'] in
>1KEH_1|Chain A|precursor of cephalosporin acylase|Brevundimonas diminuta (293)
 EPTSTPQAPIAAYKPRSNEILWDGYGVPHIYGVDAPSAFYGYGWAQARSHGDNILRLYGEARGKGAEYWGPDYEQTTVWLLTNGVPERAQQWYAQQSPDFRANLDAFAAGINAYAQQNPDDISPEVRQVLPVSGADVV
AHAHRLMNFLYVASPGRTLGEGDPPDLADQGANSWAVAPGKTANGNALLLQNPHLSWTTDYFTYYEAHLVTPDFEIYGATQIGLPVIRFAFNQRMGITNTVNGMVGATNYRLTLQDGGYLYDGQVRPFERRQASYRLRQ
ADGSTVDKPLEIRSSVHGPVFERADGTAVAVRVAGLDRPGMLEQYFDMITAHSFDDYEAAMARMQVPTFNIVYADREGTINYSFNGVAPKRAEGDIAFWQGNVPGDSSRYLWTETHPLDDLPRVTNPPGGFVQNSNDPP
WTPTWPVTYCPANHPSYLAPQTPHSLRAQQSVRLMSENDDLTLERFMALQFSHRAVMADRTLPDLIPAALIDPDPEVQAAARLLAAWDRDFTSDSRAALLFEEWARLFAGQNFAGQAAFATPWSLDKPVSTPYGVRDPK
AAVDQLRTAIANTKRKYGAIDRPFGDASRMILNDVNVPGAAGYGNLGSFRVFTWSDPDENGIRTPVHGETWVAMIEFSTPVRAYGLMSYGNSRQPGTTHYSDQIERVSRADFRELLLRREQVEAAVQERTPFNF
this seq match to pattern  ['STVKDG'] in
>1KVD_2|Chains B,D|SMK TOXIN|Pichia farinosa (4920)
 GEATTIWGVGADEAIDKGTPSKNDLQNMSADLAKNGFKGHQGVACSTVKDGNKDVYMIKFSLAGGSNDPGGSPCSDD
this seq match to pattern  ['STVKDG'] in
>1KVD_2|Chains B,D|SMK TOXIN|Pichia farinosa (4920)
 GEATTIWGVGADEAIDKGTPSKNDLQNMSADLAKNGFKGHQGVACSTVKDGNKDVYMIKFSLAGGSNDPGGSPCSDD
this seq match to pattern  ['STVKDG'] in
>1KVE_2|Chains B,D|SMK TOXIN|Pichia farinosa (4920)
 GEATTIWGVGADEAIDKGTPSKNDLQNMSADLAKNGFKGHQGVACSTVKDGNKDVYMIKFSLAGGSNDPGGSPCSDD
this seq match to pattern  ['STVKDG'] in
>1KVE_2|Chains B,D|SMK TOXIN|Pichia farinosa (4920)
 GEATTIWGVGADEAIDKGTPSKNDLQNMSADLAKNGFKGHQGVACSTVKDGNKDVYMIKFSLAGGSNDPGGSPCSDD
this seq match to pattern  ['STVDKV'] in
>1L9Y_1|Chains A,B|FEZ-1 b-lactamase|Fluoribacter gormanii (464)
 AYPMPNPFPPFRIAGNLYYVGTDDLASYLIVTPRGNILINSDLEANVPMIKASIKKLGFKFSDTKILLISHAHFDHAAGSELIKQQTKAKYMVMDEDVSVILSGGKSDFHYANDSSTYFTQSTVDKVLHDGERVELGG
TVLTAHLTPGHTRGCTTWTMKLKDHGKQYQAVIIGSIGVNPGAKLVDNITYPKIAEDYKHSIKVLESMRCDIFLGSHAGMFDLKNKYVLLSKGQNNPFVDPTGCKNYIEQKANDFYTELKKQETG
this seq match to pattern  ['STVDKV'] in
>1L9Y_1|Chains A,B|FEZ-1 b-lactamase|Fluoribacter gormanii (464)
 AYPMPNPFPPFRIAGNLYYVGTDDLASYLIVTPRGNILINSDLEANVPMIKASIKKLGFKFSDTKILLISHAHFDHAAGSELIKQQTKAKYMVMDEDVSVILSGGKSDFHYANDSSTYFTQSTVDKVLHDGERVELGG
TVLTAHLTPGHTRGCTTWTMKLKDHGKQYQAVIIGSIGVNPGAKLVDNITYPKIAEDYKHSIKVLESMRCDIFLGSHAGMFDLKNKYVLLSKGQNNPFVDPTGCKNYIEQKANDFYTELKKQETG
this seq match to pattern  ['STVEKT'] in
>1LGV_1|Chains A,B|IMMUNOGLOBULIN LAMBDA LIGHT CHAIN|Homo sapiens (9606)
 QTALTQPASVSGSPGQSITVSCTGVSSIVGSYNLVSWYQQHPGKAPKLLTYEVNKRPSGVSDRFSGSKSGNSASLTISGLQAEDEADYYCSSYDGSSTSVVFGGGTKLTVLGQPKAAPSVTLFPPSSEELQANKATLV
CLISDFYPGAVTVAWKADSSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTAC
this seq match to pattern  ['STVEKT'] in
>1LGV_1|Chains A,B|IMMUNOGLOBULIN LAMBDA LIGHT CHAIN|Homo sapiens (9606)
 QTALTQPASVSGSPGQSITVSCTGVSSIVGSYNLVSWYQQHPGKAPKLLTYEVNKRPSGVSDRFSGSKSGNSASLTISGLQAEDEADYYCSSYDGSSTSVVFGGGTKLTVLGQPKAAPSVTLFPPSSEELQANKATLV
CLISDFYPGAVTVAWKADSSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTAC
this seq match to pattern  ['STVEKT'] in
>1LHZ_1|Chains A,B|IMMUNOGLOBULIN LAMBDA LIGHT CHAIN|Homo sapiens (9606)
 QTALTQPASVSGSPGQSITVSCTGVSSIVGSYNLVSWYQQHPGKAPKLLTYEVNKRPSGVSDRFSGSKSGNSASLTISGLQAEDEADYYCSSYDGSSTSVVFGGGTKLTVLGQPKAAPSVTLFPPSSEELQANKATLV
CLISDFYPGAVTVAWKADSSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTAC
this seq match to pattern  ['STVEKT'] in
>1LHZ_1|Chains A,B|IMMUNOGLOBULIN LAMBDA LIGHT CHAIN|Homo sapiens (9606)
 QTALTQPASVSGSPGQSITVSCTGVSSIVGSYNLVSWYQQHPGKAPKLLTYEVNKRPSGVSDRFSGSKSGNSASLTISGLQAEDEADYYCSSYDGSSTSVVFGGGTKLTVLGQPKAAPSVTLFPPSSEELQANKATLV
CLISDFYPGAVTVAWKADSSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTAC
this seq match to pattern  ['STVEKT'] in
```

```
>1LIL_1|Chains A,B|LAMBDA III BENCE JONES PROTEIN CLE|Homo sapiens (9606)
YEVTQPPSLSVSPGQTARITCSGEKLGDAYVCWYQQRPGQSPVVVIYQDNRRPSGIPERFSGSSSGNTATLTISGTQTLDEADYYCQVWDSNASVVFGGGTKLTVLGQPKAAPSVTLFPPSSEELQANKATLVCLISD
FYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1LIL_1|Chains A,B|LAMBDA III BENCE JONES PROTEIN CLE|Homo sapiens (9606)
YEVTQPPSLSVSPGQTARITCSGEKLGDAYVCWYQQRPGQSPVVVIYQDNRRPSGIPERFSGSSSGNTATLTISGTQTLDEADYYCQVWDSNASVVFGGGTKLTVLGQPKAAPSVTLFPPSSEELQANKATLVCLISD
FYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1MCB_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1MCB_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1MCC_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1MCC_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1MCE_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1MCE_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1MCF_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1MCF_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1MCH_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1MCH_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1MCJ_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1MCJ_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1MCL_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1MCL_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1MCO_1|Chain L|IGG1 MCG INTACT ANTIBODY (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATEVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1MCR_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1MCR_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1MCS_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1MCS_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1NL0_1|Chain L|anti-factor IX antibody, 10C12, chain L|Homo sapiens (9606)
QSVLTQPPSVSAAPGQKVTISCSGSTSNIGNNYVSWYQQHPGKAPKLMIYDVSKRPSGVPDRFSGSKSGNSASLDISGLQSEDEADYYCAAWDDSLSEFLFGTGTKLTVLGQPKAAPSVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHKSYSCQVTHEGSTVEKTVAPT
this seq match to pattern  ['STVEEN'] in
 >1OVT_1|Chain A|OVOTRANSFERRIN|Gallus gallus (9031)
APPKSVIRWCTISSPEEKKCNNLRDLTQQERISLTCVQKATYLDCIKAIANNEADAISLDGGQAFEAGLAPYKLKPIAAEVYEHTEGSTTSYYAVAVVKKGTEFTVNDLQGKTSCHTGLGRSAGWNIPIGTLLHRGAI
EWEGIESGSVEQAVAKFFSASCVPGATIEQKLCRQCKGDPKTKCARNAPYSGYSGAFHCLKDGKGDVAFVKHTTVNENAPDQKDEYELLCLDGSRQPVDNYKTCNWARVAAHAVVARDDNKVEDIWSFLSKAQSDFGVD
TKSDFHLFGPPGKKDPVLKDLLFKDSAIMLKRVPSLMDSQLYLGFEYYSAIQSMRKDQLTPSPRENRIQWCAVGKDEKSKCDRWSVVSNGDVECTVVDETKDCIIKIMKGEADAVALDGGLVYTAGVCGLVPVMAERYD
DESQCSKTDERPASYFAVAVARKDSNVNWNNLKGKKSCHTAVGRTAGWVIPMGLIHNRTGTCNFDEYFSEGCAPGSPPNSRLCQLCQGSGGIPPEKCVASSHEKYFGYTGALRCLVEKGDVAFIQHSTVEENTGGKNKA
DWAKNLQMDDFELLCTDGRRANVMDYRECNLAEVPTHAVVVRPEKANKIRDLLERQEKRFGVNGSEKSKFMMFESQNKDLLFKDLTKCLFKVREGTTYKEFLGDKFYTVISSLKTCNPSDILQMCSFLEGK
this seq match to pattern  ['STVEKT'] in
 >1Q1J_1|Chains L,M|Fab 447-52D, light chain|Homo sapiens (9606)
QSVLTQPPSVSAAPGQKVTISCSGSSSNIGNNYVLWYQQFPGTAPKLLIYGNNKRPSGIPDRFSGSKSGTSATLGITGLQTGDEADYFCATWDSGLSADWVFGGGTKLTVLSQPKAAPSVTLFPPSSEELQANKATLV
CLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTE
```

```
this seq match to pattern  ['STVEKT'] in
 >1Q1J_1|Chains L,M|Fab 447-52D, light chain|Homo sapiens (9606)
 QSVLTQPPSVSAAPGQKVTISCSGSSSNIGNNYVLWYQQFPGTAPKLLIYGNNKRPSGIPDRFSGSKSGTSATLGITGLQTGDEADYFCATWDSGLSADWVFGGGTKLTVLSQPKAAPSVTLFPPSSEELQANKATLV
 CLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTE
this seq match to pattern  ['STVERE'] in
 >1R5M_1|Chain A|SIR4-interacting protein SIF2|Saccharomyces cerevisiae (4932)
 SESNKAGEDGASTVERETQEDDTNSIDSSDDLDGFVKILKEIVKLDNIVSSTWNPLDESILAYGEKNSVARLARIVETDQEGKKYWKLTIIAELRHPFALSASSGKTTNQVTCLAWSHDGNSIVTGVENGELRLWNKT
 GALLNVLNFHRAPIVSVKWNKDGTHIISMDVENVTILWNVISGTVMQHFELKETGGSSINAENHSGDGSLGVDVEWVDDDKFVIPGPKGAIFVYQITEKTPTGKLIGHHGPISVLEFNDTNKLLLSASDDGTLRIWHGG
 NGNSQNCFYGHSQSIVSASWVGDDKVISCSMDGSVRLWSLKQNTLLALSIVDGVPIFAGRISQDGQKYAVAFMDGQVNVYDLKKLNSKSRSLYGNRDGILNPLPIPLYASYQSSQDNDYIFDLSWNCAGNKISVAYSLQ
 EGSVVAIPG
this seq match to pattern  ['STVRED'] in
 >1RP1_1|Chain A|PANCREATIC LIPASE RELATED PROTEIN 1|Canis lupus familiaris (9615)
 KEVCYEQIGCFSDAEPWAGTAIRPLKVLPWSPERIGTRFLLYTNKNPNNFQTLLPSDPSTIGASNFQTDKKTRFIIHGFIDKGEENWLLDMCKNMFKVEEVNCICVDWKKGSQTSYTQAANNVRVVGAQVAQMLSMLS
 ANYSYSPSQVQLIGHSLGAHVAGEAGSRTPGLGRITGLDPVEASFQGTPEEVRLDPTDADFVDVIHTDAAPLIPFLGFGTSQQMGHLDFFPNGGEEMPGCKKNALSQIVDLDGIWEGTRDFVACNHLRSYKYYSESILN
 PDGFASYPCASYRAFESNKCFPCPDQGCPQMGHYADKFAVKTSDETQKYFLNTGDSSNFARWRYGVSITLSGKRATGQAKVALFGSKGNTHQFNIFKGILKPGSTHSNEFDAKLDVGTIEKVKFLWNNNVVNPTFPKVG
 AAKITVQKGEEKTVHSFCSESTVREDVLLTLTPC
this seq match to pattern  ['STVEEN'] in
 >1RYX_1|Chain A|Ovotransferrin|Gallus gallus (9031)
 APPKSVIRWCTISSPEEKKCNNLRDLTQQERISLTCVQKATYLDCIKAIANNEADAISLDGGQAFEAGLAPYKLKPIAAEVYEHTEGSTTSYYAVAVVKKGTEFTVNDLQGKTSCHTGLGRSAGWNIPIGTLLHRGAI
 EWEGIESGSVEQAVAKFFSASCVPGATIEQKLCRQCKGDPKTKCARNAPYSGYSGAFHCLKDGKGDVAFVKHTTVNENAPDQKDEYELLCLDGSRQPVDNYKTCNWARVAAHAVVARDDNKVEDIWSFLSKAQSDFGVD
 TKSDFHLFGPPGKKDPVLKDLLFKDSAIMLKRVPSLMDSQLYLGFEYYSAIQSMRKDQLTPSPRENRIQWCAVGKDEKSKCDRWSVVSNGDVECTVVDETKDCIIKIMKGEADAVALDGGLVYTAGVCGLVPVMAERYD
 DESQCSKTDERPASYFAVAVARKDSNVNWNNLKGKKSCHTAVGRTAGWVIPMGLIHNRTGTCNFDEYFSEGCAPGSPPNSRLCQLCQGSGGIPPEKCVASSHEKYFGYTGALRCLVEKGDVAFIQHSTVEENTGGKNKA
 DWAKNLQMDDFELLCTDGRRANVMDYRECNLAEVPTHAVVVRPEKANKIRDLLREQEKRFGVNGSEKSKFMMFESQNKDLLFKDLTKCLFKVREGTTYKEFLGDKFYTVISSLKTCNPSDILQMCSFLEGK
this seq match to pattern  ['STVEKT'] in
 >1RZF_L|Chain L|Fab E51 light chain|Homo sapiens (9606)
 QSILTQPPSVSAAPGQKVTISCSGSSSNIGNNDVSWYQQFPGTVPKLVIYENNERPSGIPDRFSGSKSGTSATLGITGLQTGDEADYYCGTWDSSLSAVVFGGGSKVTVLGQPKAAPSVTLFPPSSEELQANKATLVC
 LISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPT
this seq match to pattern  ['STVKKA'] in
 >1S5J_1|Chain A|DNA polymerase I|Sulfolobus solfataricus (2287)
 VVRREWLEEAQENKIYFLLQVDYDGKKGKAVCKLFDKETQKIYALYDNTGHKPYFLVDLEPDKVGKIPKIVRDPSFDHIETVSKIDPYTWNKFKLTKIVVRDPLAVRRLRNDVPKAYEAHIKYFNNYMYDIGLIPGMP
 YVVKNGKLESVYLSLDEKDVEEIKKAFADSDEMTRQMAVDWLPIFETEIPKIKRVAIDIEVYTPVKGRIPDSQKAEFPIISIALAGSDGLKKVLVLNRNDVNEGSVKLDGISVERFNTEYELLGRFFDILLEYPIVLTF
 NGDDFDLPYIYFRALKLGYFPEEIPIDVAGKDEAKYLAGLHIDLYKFFFNKAVRNYAFEGKYNEYNLDAVAKALLGTSKVKVDTLISFLDVEKLIEYNFRDAEITLQLTTFNNDLTMKLIVLFSRISRLGIEELTRTEI
 STWVKNLYYWEHRKRNWLIPLKEEILAKSSNIRTSALIKGKGYKGAVVIDPPAGIFFNITVLDFASLYPSIIRTWNLSYETVDIQQCKKPYEVKDETGEVLHIVCMDRPGITAVITGLLRDFRVKIYKKKAKNPNNSEE
 QKLLYDVVQRAMKVFINATYGVFGAETFPLYAPRVAESVTALGRYVITSTVKKAREEGLTVLYGDTDSLFLLNPPKNSLENIIKWVKTTFNLDLEVDKTYKFVAFSGLKKNYFGVYQDGKVDIKGMLVKKRNTPEFVKK
 VFNEVKELMISINSPNDVKEIKRKIVDVVKGSYEKLKNKGYNLDELAFKVMLSKPLDAYKKNTPQHVKAALQLRPFGVNVLPRDIIYYVKVRSKDGVKPVQLAKVTEIDAEKYLEALRSTFEQILRAFGVSWDEIAATM
 SIDSFFSYPSKGNS
this seq match to pattern  ['STVEDE'] in
 >1UYP_1|Chains A,B,C,D,E,F|BETA-FRUCTOSIDASE|THERMOTOGA MARITIMA (243274)
 LFKPNYHFFPITGWMNDPNGLIFWKGKYHMFYQYNPRKPEWGNICWGHAVSDDLVHWRHLPVALYPDDETHGVFSGSAVEKDGKMFLVYTYYRDPTHNKGEKETQCVVMSENGLDFVKYDGNPVISKPPEEGTHAFRD
 PKVNRSNGEWRMVLGSGKDEKIGRVLLYTSDDLFHWKYEGAIFEDETTKEIECPDLVRIGEKDILIYSITSTNSVLFSMGELKEGKLNVEKRGLLDHGTDFYAAQTFFGTDRVVVIGWLQSWLRTGLYPTKREGWNGVM
 SLPRELYVENNELKVKPVDELLALRKRKVFETAKSGTFLLDVKENSYEIVCEFSGEIELRMGNESEEVVITKSRDELIVDTTRSGVSGGEVRKSTVEDEATNRIRAFLDSCSVEFFFNDSIAFSFRIHPENVYNILSVK
 SNQVKLEVFELENIWL
this seq match to pattern  ['STVEDE'] in
 >1UYP_1|Chains A,B,C,D,E,F|BETA-FRUCTOSIDASE|THERMOTOGA MARITIMA (243274)
 LFKPNYHFFPITGWMNDPNGLIFWKGKYHMFYQYNPRKPEWGNICWGHAVSDDLVHWRHLPVALYPDDETHGVFSGSAVEKDGKMFLVYTYYRDPTHNKGEKETQCVVMSENGLDFVKYDGNPVISKPPEEGTHAFRD
 PKVNRSNGEWRMVLGSGKDEKIGRVLLYTSDDLFHWKYEGAIFEDETTKEIECPDLVRIGEKDILIYSITSTNSVLFSMGELKEGKLNVEKRGLLDHGTDFYAAQTFFGTDRVVVIGWLQSWLRTGLYPTKREGWNGVM
 SLPRELYVENNELKVKPVDELLALRKRKVFETAKSGTFLLDVKENSYEIVCEFSGEIELRMGNESEEVVITKSRDELIVDTTRSGVSGGEVRKSTVEDEATNRIRAFLDSCSVEFFFNDSIAFSFRIHPENVYNILSVK
 SNQVKLEVFELENIWL
this seq match to pattern  ['STVEDE'] in
 >1UYP_1|Chains A,B,C,D,E,F|BETA-FRUCTOSIDASE|THERMOTOGA MARITIMA (243274)
 LFKPNYHFFPITGWMNDPNGLIFWKGKYHMFYQYNPRKPEWGNICWGHAVSDDLVHWRHLPVALYPDDETHGVFSGSAVEKDGKMFLVYTYYRDPTHNKGEKETQCVVMSENGLDFVKYDGNPVISKPPEEGTHAFRD
 PKVNRSNGEWRMVLGSGKDEKIGRVLLYTSDDLFHWKYEGAIFEDETTKEIECPDLVRIGEKDILIYSITSTNSVLFSMGELKEGKLNVEKRGLLDHGTDFYAAQTFFGTDRVVVIGWLQSWLRTGLYPTKREGWNGVM
 SLPRELYVENNELKVKPVDELLALRKRKVFETAKSGTFLLDVKENSYEIVCEFSGEIELRMGNESEEVVITKSRDELIVDTTRSGVSGGEVRKSTVEDEATNRIRAFLDSCSVEFFFNDSIAFSFRIHPENVYNILSVK
 SNQVKLEVFELENIWL
this seq match to pattern  ['STVEDE'] in
 >1UYP_1|Chains A,B,C,D,E,F|BETA-FRUCTOSIDASE|THERMOTOGA MARITIMA (243274)
 LFKPNYHFFPITGWMNDPNGLIFWKGKYHMFYQYNPRKPEWGNICWGHAVSDDLVHWRHLPVALYPDDETHGVFSGSAVEKDGKMFLVYTYYRDPTHNKGEKETQCVVMSENGLDFVKYDGNPVISKPPEEGTHAFRD
 PKVNRSNGEWRMVLGSGKDEKIGRVLLYTSDDLFHWKYEGAIFEDETTKEIECPDLVRIGEKDILIYSITSTNSVLFSMGELKEGKLNVEKRGLLDHGTDFYAAQTFFGTDRVVVIGWLQSWLRTGLYPTKREGWNGVM
 SLPRELYVENNELKVKPVDELLALRKRKVFETAKSGTFLLDVKENSYEIVCEFSGEIELRMGNESEEVVITKSRDELIVDTTRSGVSGGEVRKSTVEDEATNRIRAFLDSCSVEFFFNDSIAFSFRIHPENVYNILSVK
 SNQVKLEVFELENIWL
this seq match to pattern  ['STVEDE'] in
 >1UYP_1|Chains A,B,C,D,E,F|BETA-FRUCTOSIDASE|THERMOTOGA MARITIMA (243274)
 LFKPNYHFFPITGWMNDPNGLIFWKGKYHMFYQYNPRKPEWGNICWGHAVSDDLVHWRHLPVALYPDDETHGVFSGSAVEKDGKMFLVYTYYRDPTHNKGEKETQCVVMSENGLDFVKYDGNPVISKPPEEGTHAFRD
 PKVNRSNGEWRMVLGSGKDEKIGRVLLYTSDDLFHWKYEGAIFEDETTKEIECPDLVRIGEKDILIYSITSTNSVLFSMGELKEGKLNVEKRGLLDHGTDFYAAQTFFGTDRVVVIGWLQSWLRTGLYPTKREGWNGVM
 SLPRELYVENNELKVKPVDELLALRKRKVFETAKSGTFLLDVKENSYEIVCEFSGEIELRMGNESEEVVITKSRDELIVDTTRSGVSGGEVRKSTVEDEATNRIRAFLDSCSVEFFFNDSIAFSFRIHPENVYNILSVK
 SNQVKLEVFELENIWL
this seq match to pattern  ['STVEDE'] in
 >1UYP_1|Chains A,B,C,D,E,F|BETA-FRUCTOSIDASE|THERMOTOGA MARITIMA (243274)
 LFKPNYHFFPITGWMNDPNGLIFWKGKYHMFYQYNPRKPEWGNICWGHAVSDDLVHWRHLPVALYPDDETHGVFSGSAVEKDGKMFLVYTYYRDPTHNKGEKETQCVVMSENGLDFVKYDGNPVISKPPEEGTHAFRD
 PKVNRSNGEWRMVLGSGKDEKIGRVLLYTSDDLFHWKYEGAIFEDETTKEIECPDLVRIGEKDILIYSITSTNSVLFSMGELKEGKLNVEKRGLLDHGTDFYAAQTFFGTDRVVVIGWLQSWLRTGLYPTKREGWNGVM
 SLPRELYVENNELKVKPVDELLALRKRKVFETAKSGTFLLDVKENSYEIVCEFSGEIELRMGNESEEVVITKSRDELIVDTTRSGVSGGEVRKSTVEDEATNRIRAFLDSCSVEFFFNDSIAFSFRIHPENVYNILSVK
 SNQVKLEVFELENIWL
this seq match to pattern  ['STVERS'] in
 >1VQT_1|Chain A|Orotidine 5'-phosphate decarboxylase|Thermotoga maritima (2336)
 MGSDKIHHHHHHMTPVLSLDMEDPIRFIDENGSFEVVKVGHNLAIHGKKIFDELAKRNLKIILDLKFCDIPSTVERSIKSWDHPAIIGFTVHSCAGYESVERALSATDKHVFVVVKLTSMEGSLEDYMDRIEKLNKLG
 CDFVLPGPWAKALREKIKGKILVPGIRMEVKADDQKDVVTLEEMKGIANFAVLGREIYLSENPREKIKRIKEMRL
this seq match to pattern  ['STVEKT'] in
 >1W72_5|Chains L,M|HYB3 LIGHT CHAIN|HOMO SAPIENS (9606)
 SYVLTQPPSVSVAPGQTARITCGGNNIGSRSVHWYQQKPGQAPVLVVYDDSDRPSGIPERFSGSNSGNMATLTISRVEAGDEADYYCQVWDSRTDHWVFGGGTDLTVLGQPKAAPSVTLFPPSSEELQANKATLVCLI
 SDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAP
this seq match to pattern  ['STVEKT'] in
 >1W72_5|Chains L,M|HYB3 LIGHT CHAIN|HOMO SAPIENS (9606)
 SYVLTQPPSVSVAPGQTARITCGGNNIGSRSVHWYQQKPGQAPVLVVYDDSDRPSGIPERFSGSNSGNMATLTISRVEAGDEADYYCQVWDSRTDHWVFGGGTDLTVLGQPKAAPSVTLFPPSSEELQANKATLVCLI
 SDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAP
this seq match to pattern  ['STVERR'] in
 >1WF5_1|Chain A|sidekick 2 protein|Homo sapiens (9606)
 GSSGSSSGRSAHLRVRQLPHAPEHPVATLSTVERRAINLTWTKPFDGNSPLIRYILEMSENNAPWTVLLASVDPKATSVTVKGLVPARSYQFRLCAVNDVGKGQFSKDTERVSLPESGPSSG
this seq match to pattern  ['STVDDD'] in
 >1Z1D_1|Chain A|Replication protein A 32 kDa subunit|Homo sapiens (9606)
 GSHMANSQPSAGRAPISNPGMSEAGNFGGNSFMPANGLTVAQNQVLNLIKACPRPEGLNFQDLKNQLKHMSVSSIKQAVDFLSNEGHIYSTVDDDHFKSTDAE
this seq match to pattern  ['STVEDK'] in
 >1Z7Q_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|proteasome activator protein PA26|Trypanosoma brucei (5691)
 MPPKRAALIQNLRDSYTETESSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
 PIGMYALREYLSARSTVEDKLLGVDAESGKTKGGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGSDHMVS
this seq match to pattern  ['STVEDK'] in
 >1Z7Q_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|proteasome activator protein PA26|Trypanosoma brucei (5691)
 MPPKRAALIQNLRDSYTETESSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
 PIGMYALREYLSARSTVEDKLLGVDAESGKTKGGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGSDHMVS
```

```
this seq match to pattern  ['STVEDK'] in
 >1Z7Q_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|proteasome activator protein PA26|Trypanosoma brucei (5691)
  MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGVDAESGKTKGGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGSDHMVS
this seq match to pattern  ['STVEDK'] in
 >1Z7Q_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|proteasome activator protein PA26|Trypanosoma brucei (5691)
  MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGVDAESGKTKGGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGSDHMVS
this seq match to pattern  ['STVEDK'] in
 >1Z7Q_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|proteasome activator protein PA26|Trypanosoma brucei (5691)
  MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGVDAESGKTKGGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGSDHMVS
this seq match to pattern  ['STVEDK'] in
 >1Z7Q_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|proteasome activator protein PA26|Trypanosoma brucei (5691)
  MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGVDAESGKTKGGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGSDHMVS
this seq match to pattern  ['STVEDK'] in
 >1Z7Q_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|proteasome activator protein PA26|Trypanosoma brucei (5691)
  MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGVDAESGKTKGGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGSDHMVS
this seq match to pattern  ['STVEDK'] in
 >1Z7Q_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|proteasome activator protein PA26|Trypanosoma brucei (5691)
  MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGVDAESGKTKGGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGSDHMVS
this seq match to pattern  ['STVEDK'] in
 >1Z7Q_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|proteasome activator protein PA26|Trypanosoma brucei (5691)
  MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGVDAESGKTKGGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGSDHMVS
this seq match to pattern  ['STVEDK'] in
 >1Z7Q_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|proteasome activator protein PA26|Trypanosoma brucei (5691)
  MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGVDAESGKTKGGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGSDHMVS
this seq match to pattern  ['STVEDK'] in
 >1Z7Q_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|proteasome activator protein PA26|Trypanosoma brucei (5691)
  MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGVDAESGKTKGGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGSDHMVS
this seq match to pattern  ['STVEDK'] in
 >1Z7Q_15|Chains c,d,e,f,g,h,i,j,k,l,m,n,o,p|proteasome activator protein PA26|Trypanosoma brucei (5691)
  MPPKRAALIQNLRDSYTETSSFAVIEEWAAGTLQEIEGIAKAAAEAHGVIRNSTYGRAQAEKSPEQLLGVLQRYQDLCHNVYCQAETIRTVIAIRIPEHKEEDNLGVAVQHAVLKIIDELEIKTLGSGEKSGSGGAPT
PIGMYALREYLSARSTVEDKLLGVDAESGKTKGGGSQSPSLLLELRQIDADFMLKVELATTHLSTMVRAVINAYLLNWKKLIQPRTGSDHMVS
this seq match to pattern  ['STVDKY'] in
 >1ZTM_1|Chains A,B,C|Fusion glycoprotein|Human parainfluenza virus 3 (11216)
  QIDITKLQHVGVLVNSPKGMKISQNFETRYLILSLIPKIEDSNSCGDQQIKQYKRLLDRLIIPLYDGLRLQKDVIVSNQESNENTDPSTKRFFGGVIGTIALGVATSAQITAAVALVEAKQARSDIEKLKEAIRDTNK
AVQSVQSSIGNLIVAIKSVQDYVNKEIVPSIARLGCEAAGLQLGIALTQHYSELTNIFGDNIGSLQEKGIKLQGIASLYRTNITEIFTTSTVDKYDIYDLLFTESIKVRVIDVDLNDYSITLQVRLPLLTRLLNTQIYR
VDSISYNIQNREWYIPLPSHIMTKGAFLGGADVKECIEAFSSYICPSDPGFVLNHEMESCLSGNISQCPRTVVKSDIVPRYAFVNGGVVANCITTTCTCNGIGNRINQPPDQGVKIITHKECNTIGINGMLFNTNKEGT
LAFYTPNDITLNNSVALDPIDISIELNKAKSDLEESKEWIRRSNQKLDSIGNWHQSSTTGGPLVPRGSHHHHHH
this seq match to pattern  ['STVDKY'] in
 >1ZTM_1|Chains A,B,C|Fusion glycoprotein|Human parainfluenza virus 3 (11216)
  QIDITKLQHVGVLVNSPKGMKISQNFETRYLILSLIPKIEDSNSCGDQQIKQYKRLLDRLIIPLYDGLRLQKDVIVSNQESNENTDPSTKRFFGGVIGTIALGVATSAQITAAVALVEAKQARSDIEKLKEAIRDTNK
AVQSVQSSIGNLIVAIKSVQDYVNKEIVPSIARLGCEAAGLQLGIALTQHYSELTNIFGDNIGSLQEKGIKLQGIASLYRTNITEIFTTSTVDKYDIYDLLFTESIKVRVIDVDLNDYSITLQVRLPLLTRLLNTQIYR
VDSISYNIQNREWYIPLPSHIMTKGAFLGGADVKECIEAFSSYICPSDPGFVLNHEMESCLSGNISQCPRTVVKSDIVPRYAFVNGGVVANCITTTCTCNGIGNRINQPPDQGVKIITHKECNTIGINGMLFNTNKEGT
LAFYTPNDITLNNSVALDPIDISIELNKAKSDLEESKEWIRRSNQKLDSIGNWHQSSTTGGPLVPRGSHHHHHH
this seq match to pattern  ['STVDKY'] in
 >1ZTM_1|Chains A,B,C|Fusion glycoprotein|Human parainfluenza virus 3 (11216)
  QIDITKLQHVGVLVNSPKGMKISQNFETRYLILSLIPKIEDSNSCGDQQIKQYKRLLDRLIIPLYDGLRLQKDVIVSNQESNENTDPSTKRFFGGVIGTIALGVATSAQITAAVALVEAKQARSDIEKLKEAIRDTNK
AVQSVQSSIGNLIVAIKSVQDYVNKEIVPSIARLGCEAAGLQLGIALTQHYSELTNIFGDNIGSLQEKGIKLQGIASLYRTNITEIFTTSTVDKYDIYDLLFTESIKVRVIDVDLNDYSITLQVRLPLLTRLLNTQIYR
VDSISYNIQNREWYIPLPSHIMTKGAFLGGADVKECIEAFSSYICPSDPGFVLNHEMESCLSGNISQCPRTVVKSDIVPRYAFVNGGVVANCITTTCTCNGIGNRINQPPDQGVKIITHKECNTIGINGMLFNTNKEGT
LAFYTPNDITLNNSVALDPIDISIELNKAKSDLEESKEWIRRSNQKLDSIGNWHQSSTTGGPLVPRGSHHHHHH
this seq match to pattern  ['STVEKT'] in
 >1ZVO_1|Chains A,B|myeloma immunoglobulin D lambda|Homo sapiens (9606)
  QSVLTQPPSASGTPGQRVTISCFGSSSNIGRYYVYWYQQLPGTTPKLLIYKDNQRPSGVPDRFSGSKSGTSASLAISGLRSEDEADYYCAAWDDSLWVFGGGTTLTVLSQPKAAPSVTLFPPSSEELQANKATLVCLI
SDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >1ZVO_1|Chains A,B|myeloma immunoglobulin D lambda|Homo sapiens (9606)
  QSVLTQPPSASGTPGQRVTISCFGSSSNIGRYYVYWYQQLPGTTPKLLIYKDNQRPSGVPDRFSGSKSGTSASLAISGLRSEDEADYYCAAWDDSLWVFGGGTTLTVLSQPKAAPSVTLFPPSSEELQANKATLVCLI
SDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVRRC'] in
 >2APC_1|Chain A|Alpha-1,3-mannosyl-glycoprotein 2-beta-N-acetylglucosaminyltransferase|Oryctolagus cuniculus (9986)
  AVIPILVIACDRSTVRRCLDKLLHYRPSAELFPIIVSQDCGHEETAQVIASYGSAVTHIRQPDLSNIAVQPDHRKFQGYYKIARHYRWALGQIFHNFNYPAAVVVEDDLEVAPDFFEYFQATYPLLKADPSLWCVSAW
NDNGKEQMVDSSKPELLYRTDFFPGLGWLLLAELWAELEPKWPKAFWDDWMRRPEQRKGRACVRPEISRTMTFGRKGVSHGQFFDQHLKFIKLNQQFVPFTQLDLSYLQQEAYDRDFLARVYGAPQLQVEKVRTNDRKE
LGEVRVQYTGRDSFKAFAKALGVMDDLKSGVPRAGYRGIVTFLFRGRRVHLAPPQTWDGYDPSWT
this seq match to pattern  ['STVEKT'] in
 >2B0S_1|Chain L|Fab 2219, light chain|Homo sapiens (9606)
  QSVLTQPPSASGTPGQRISISCSGTSSNVENNYVYWYQHLPGTAPKLLIYRNDHRSSGIPDRFSASKSGTSASLAISGLRPEDEGDYYCAAWDDSRGGPDWVFGGGTKLTVLAQPKAAPSVTLFPPSSEELQANKATL
VCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHKSYSCQVTHEGSTVEKTVAPT
this seq match to pattern  ['STVEKT'] in
 >2B1A_1|Chain L|Fab 2219, light chain|Homo sapiens (9606)
  QSVLTQPPSASGTPGQRISISCSGTSSNVENNYVYWYQHLPGTAPKLLIYRNDHRSSGIPDRFSASKSGTSASLAISGLRPEDEGDYYCAAWDDSRGGPDWVFGGGTKLTVLAQPKAAPSVTLFPPSSEELQANKATL
VCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHKSYSCQVTHEGSTVEKTVAPT
this seq match to pattern  ['STVEKT'] in
 >2B1H_1|Chain L|Fab 2219, light chain|Homo sapiens (9606)
  QSVLTQPPSASGTPGQRISISCSGTSSNVENNYVYWYQHLPGTAPKLLIYRNDHRSSGIPDRFSASKSGTSASLAISGLRPEDEGDYYCAAWDDSRGGPDWVFGGGTKLTVLAQPKAAPSVTLFPPSSEELQANKATL
VCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHKSYSCQVTHEGSTVEKTVAPT
this seq match to pattern  ['STVKDT'] in
 >2BB0_1|Chains A,B|Imidazolonepropionase|Bacillus subtilis (1423)
  MPKQIDTILINIGQLLTMESSGPRAGKSMQDLHVIEDAVVGIHEQKIVFAGQKGAEAGYEADEIIDCSGRLVTPGLVDPHTHLVFGGSREKEMNLKLQGISYLDILAQGGGILSTVKDTRAASEEELLQKAHFHLQRM
LSYGTTTAEVKSGYGLEKETELKQLRVAKKLHESQPVDLVSTFMGAHAIPPEYQNDPDDFLDQMLSLLPEIKEQELASFADIFTETGVFTVSQSRRYLQKAAAEAGFGLKIHADEIDPLGGAELAGKLKAVSADHLVGTS
DEGIKKLAEAGTIAVLLPGTTFYLGKSTYARARAMIDEGVCVSLATDFNPGSSPTENIQLIMSIAALHLKMTAEEIWHAVTVNAAYAIGKGEEAGQLKAGRSADLVIWQAPNYMYIPYHYGVNHVHQVMKNGTIVVNRE
GAILG
this seq match to pattern  ['STVKDT'] in
```

```
>2BB0_1|Chains A,B|Imidazolonepropionase|Bacillus subtilis (1423)
 MPKQIDTILINIGQLLTMESSGPRAGKSMQDLHVIEDAVVGIHEQKIVFAGQKGAEAGYEADEIIDCSGRLVTPGLVDPHTHLVFGGSREKEMNLKLQGISYLDILAQGGGILSTVKDTRAASEEELLQKAHFHLQRM
LSYGTTTAEVKSGYGLEKETELKQLRVAKKLHESQPVDLVSTFMGAHAIPPEYQNDPDDFLDQMLSLLPEIKEQELASFADIFTETGVFTVSQSRRYLQKAAEAGFGLKIHADEIDPLGGAELAGKLKAVSADHLVGTS
DEGIKKLAEAGTIAVLLPGTTFYLGKSTYARARAMIDEGVCVSLATDFNPGSSPTENIQLIMSIAALHLKMTAEEIWHAVTVNAAYAIGKGEEAGQLKAGRSADLVIWQAPNYMYIPYHYGVNHVHQVMKNGTIVVNRE
GAILG
this seq match to pattern  ['STVEDV'] in
 >2BB5_1|Chains A,B|Transcobalamin II|Homo sapiens (9606)
 EMCEIPEMDSHLVEKLGQHLLPWMDRLSLEHLNPSIYVGLRLSSLQAGTKEDLYLHSLKLGYQQCLLGSAFSEDDGDCQGKPSMGQLALYLLALRANCEFVRGHKGDRLVSQLKWFLEDEKRAIGHDHKGHPHTSYYQ
YGLGILALCLHQKRVHDSVVDKLLYAVEPFHQGHHSVDTAAMAGLAFTCLKRSNFNPGRRQRITMAIRTVREEILKAQTPEGHFGNVYSTPLALQFLMTSPMPGAELGTACLKARVALLASLQDGAFQNALMISQLLPV
LNHKTYIDLIFPDCLAPRVMLEPAAETIPQTQEIISVTLQVLSLLPPYRQSISVLAGSTVEDVLKKAHELGGFTYETQASLSGPYLTSVMGKAAGEREFWQLLRDPNTPLLQGIADYRPKDGETIELRLVSW
this seq match to pattern  ['STVEDV'] in
 >2BB5_1|Chains A,B|Transcobalamin II|Homo sapiens (9606)
 EMCEIPEMDSHLVEKLGQHLLPWMDRLSLEHLNPSIYVGLRLSSLQAGTKEDLYLHSLKLGYQQCLLGSAFSEDDGDCQGKPSMGQLALYLLALRANCEFVRGHKGDRLVSQLKWFLEDEKRAIGHDHKGHPHTSYYQ
YGLGILALCLHQKRVHDSVVDKLLYAVEPFHQGHHSVDTAAMAGLAFTCLKRSNFNPGRRQRITMAIRTVREEILKAQTPEGHFGNVYSTPLALQFLMTSPMPGAELGTACLKARVALLASLQDGAFQNALMISQLLPV
LNHKTYIDLIFPDCLAPRVMLEPAAETIPQTQEIISVTLQVLSLLPPYRQSISVLAGSTVEDVLKKAHELGGFTYETQASLSGPYLTSVMGKAAGEREFWQLLRDPNTPLLQGIADYRPKDGETIELRLVSW
this seq match to pattern  ['STVEKT'] in
 >2DD8_2|Chain L|IGG Light Chain|Homo sapiens (9606)
 SYELTQPPSVSVAPGKTARITCGGNNIGSKSVHWYQQKPGQAPVLVVYDDSDRPSGIPERFSGSNSGNTATLTISRVEAGDEADYYCQVWDSSSDYVFGTGTKVTVLGQPKANPTVTLFPPSSEEFQANKATLVCLIS
DFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVKRK'] in
 >2DVV_1|Chain A|Bromodomain-containing protein 2|Homo sapiens (9606)
 GSHMEQLKHCNGILKELLSKKHAAYAWPFYKPVDASALGLHDYHDIIKHPMDLSTVKRKMENRDYRDAQEFAADVRLMFSNCYKYNPPDHDVVAMARKLQDVFEFRYAKMPD
this seq match to pattern  ['STVKRK'] in
 >2E3K_1|Chains A,B,C,D|Bromodomain-containing protein 2|Homo sapiens (9606)
 GSHMEQLKHCNGILKELLSKKHAAYAWPFYKPVDASALGLHDYHDIIKHPMDLSTVKRKMENRDYRDAQEFAADVRLMFSNCYKYNPPDHDVVAMARKLQDVFEFRYAKMPD
this seq match to pattern  ['STVKRK'] in
 >2E3K_1|Chains A,B,C,D|Bromodomain-containing protein 2|Homo sapiens (9606)
 GSHMEQLKHCNGILKELLSKKHAAYAWPFYKPVDASALGLHDYHDIIKHPMDLSTVKRKMENRDYRDAQEFAADVRLMFSNCYKYNPPDHDVVAMARKLQDVFEFRYAKMPD
this seq match to pattern  ['STVKRK'] in
 >2E3K_1|Chains A,B,C,D|Bromodomain-containing protein 2|Homo sapiens (9606)
 GSHMEQLKHCNGILKELLSKKHAAYAWPFYKPVDASALGLHDYHDIIKHPMDLSTVKRKMENRDYRDAQEFAADVRLMFSNCYKYNPPDHDVVAMARKLQDVFEFRYAKMPD
this seq match to pattern  ['STVKRK'] in
 >2E3K_1|Chains A,B,C,D|Bromodomain-containing protein 2|Homo sapiens (9606)
 GSHMEQLKHCNGILKELLSKKHAAYAWPFYKPVDASALGLHDYHDIIKHPMDLSTVKRKMENRDYRDAQEFAADVRLMFSNCYKYNPPDHDVVAMARKLQDVFEFRYAKMPD
this seq match to pattern  ['STVDDW'] in
 >2E7N_1|Chain A|Bromodomain-containing protein 3|Homo sapiens (9606)
 GSSGSSGGKLSEHLRYCDSILREMLSKKHAAYAWPFYKPVDAEALELHDYHDIIKHPMDLSTVKRKMDGREYPDAQGFAADVRLMFSNCYKYNPPDHEVVAMARKLQDVFEMRFAKM
this seq match to pattern  ['STVDDW'] in
 >2ES7_1|Chains A,B,C,D|putative thiol-disulfide isomerase and thioredoxin|Salmonella typhimurium (99287)
 MANDTPFSALWQRLLTRGWQPVEASTVDDWIKRVGDGVILLSSDPRRTPEVSDNPVMIAELLREFPQFDWQVAVADLEQSEAIGDRFNVRRFPATLVFTDGKLRGALSGIHPWAELLTLMRSIVDTPAAQETVQLEHH
HHHH
this seq match to pattern  ['STVDDW'] in
 >2ES7_1|Chains A,B,C,D|putative thiol-disulfide isomerase and thioredoxin|Salmonella typhimurium (99287)
 MANDTPFSALWQRLLTRGWQPVEASTVDDWIKRVGDGVILLSSDPRRTPEVSDNPVMIAELLREFPQFDWQVAVADLEQSEAIGDRFNVRRFPATLVFTDGKLRGALSGIHPWAELLTLMRSIVDTPAAQETVQLEHH
HHHH
this seq match to pattern  ['STVDDW'] in
 >2ES7_1|Chains A,B,C,D|putative thiol-disulfide isomerase and thioredoxin|Salmonella typhimurium (99287)
 MANDTPFSALWQRLLTRGWQPVEASTVDDWIKRVGDGVILLSSDPRRTPEVSDNPVMIAELLREFPQFDWQVAVADLEQSEAIGDRFNVRRFPATLVFTDGKLRGALSGIHPWAELLTLMRSIVDTPAAQETVQLEHH
HHHH
this seq match to pattern  ['STVDDW'] in
 >2ES7_1|Chains A,B,C,D|putative thiol-disulfide isomerase and thioredoxin|Salmonella typhimurium (99287)
 MANDTPFSALWQRLLTRGWQPVEASTVDDWIKRVGDGVILLSSDPRRTPEVSDNPVMIAELLREFPQFDWQVAVADLEQSEAIGDRFNVRRFPATLVFTDGKLRGALSGIHPWAELLTLMRSIVDTPAAQETVQLEHH
HHHH
this seq match to pattern  ['STVEKT'] in
 >2FB4_1|Chain L|IGG1-LAMBDA KOL FAB (LIGHT CHAIN)|Homo sapiens (9606)
 QSVLTQPPSASGTPGQRVTISCSGTSSNIGSSTVNWYQQLPGMAPKLLIYRDAMRPSGVPDRFSGSKSGASASLAIGGLQSEDETDYYCAAWDVSLNAYVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEEV'] in
 >2FH6_1|Chain A|pullulanase|Klebsiella aerogenes (548)
 CDNSSSSSTSGSPGSPGNPGNPGNPGTPGTPDPQDVVVRLPDVAVPGEAVQASARQAVIHLVDIAGITSSTPADYATKNLYLWNNETCDALSAPVADWNDVSTTPTGSDKYGPYWVIPLTKESGCINVIVRDGTNKLIDSD
LRVSFSDFTDRTVSVIAGNSAVYDSRADAFRAAFGVALADAHWVDKTTLLWPGGENKPIVRLYYSHSSKVAADSNGEFSDKYVKLTPTTVNQQVSMRFPHLASYPAFKLPDDVNVDELLQGETVAIAAESDGILSSATQ
VQTAGVLDDTYAAAAEALSYGAQLTDSGVTFRVWAPTAQQVELVIYSADKKVIASHPMTRDSASGAWSWQGGSDLKGAFYRYAMTVYHPQSRKVEQYEVTDPYAHSLSTNSEYSQVVDLNDSALKPEGWDGLTMPHAQK
TKADLAKMTIHESHIRDLSAWDQTVPAELRGKYLALTAQESNMVQHLKQLSASGVTHIELLPVFDLATVNEFSDKVADIQQPFSRLCEVNSAVKSSEFAGYCDSGSTVEEVLTQLKQNDSKDNPVQVQALNTLVAQTDSY
NWGYDPFHYTVPEGSYATDPEGTARIKEFRTMIQAIKEFRTMIQAIKQDLGMNVIMDVVYNHTNAAGPTDRTSVLDKIVPWYYQRLNETTGSVESATCCSDSAPEHRMFAKLIADSLAVWTTDYKIDGFRFDLMLYHPKAQILSAWERI
KALNPDIYFFGEGWDSNQSDRFEIASQINLKGTGIGTFSDRLRDAVRGGGPFDSGDALRQNQGVGSGAGVLPNELTTLSDDQARHLADLTRLGMAGNLADFVLIDKDGAVKRGSEIDYNGAPGGYAADPTEVVNYVSKH
DNQTLWDMISYKAAQEADLDTRVRMQAVSLATVMLGQGIAFDQQGSELLRSKSFTRDSYDSGDWFNRVDYSLQDNNYNVGMPRSSDDGSNYDIIARVKDAVATPGETELKQMTAFYQELTALRKSSPLFTLGDGATVMK
RVDFRNTGADQQTGLLVMTIDDGMQAGASLDSRVDGIVVAINAAPESRTLQDFAGTSLQLSAIQQAAGDRSLASGVQVAAEDGSVTLPAWSVAVLELPQGESQGAGLPVSSK
this seq match to pattern  ['STVEEV'] in
 >2FH8_1|Chain A|pullulanase|Klebsiella aerogenes (548)
 CDNSSSSSTSGSPGSPGNPGNPGNPGTPGTPDPQDVVVRLPDVAVPGEAVQASARQAVIHLVDIAGITSSTPADYATKNLYLWNNETCDALSAPVADWNDVSTTPTGSDKYGPYWVIPLTKESGCINVIVRDGTNKLIDSD
LRVSFSDFTDRTVSVIAGNSAVYDSRADAFRAAFGVALADAHWVDKTTLLWPGGENKPIVRLYYSHSSKVAADSNGEFSDKYVKLTPTTVNQQVSMRFPHLASYPAFKLPDDVNVDELLQGETVAIAAESDGILSSATQ
VQTAGVLDDTYAAAAEALSYGAQLTDSGVTFRVWAPTAQQVELVIYSADKKVIASHPMTRDSASGAWSWQGGSDLKGAFYRYAMTVYHPQSRKVEQYEVTDPYAHSLSTNSEYSQVVDLNDSALKPEGWDGLTMPHAQK
TKADLAKMTIHESHIRDLSAWDQTVPAELRGKYLALTAQESNMVQHLKQLSASGVTHIELLPVFDLATVNEFSDKVADIQQPFSRLCEVNSAVKSSEFAGYCDSGSTVEEVLTQLKQNDSKDNPQVQALNTLVAQTDSY
NWGYDPFHYTVPEGSYATDPEGTARIKEFRTMIQAIKQDLGMNVIMDVVYNHTNAAGPTDRTSVLDKIVPWYYQRLNETTGSVESATCCSDSAPEHRMFAKLIADSLAVWTTDYKIDGFRFDLMLYHPKAQILSAWERI
KALNPDIYFFGEGWDSNQSDRFEIASQINLKGTGIGTFSDRLRDAVRGGGPFDSGDALRQNQGVGSGAGVLPNELTTLSDDQARHLADLTRLGMAGNLADFVLIDKDGAVKRGSEIDYNGAPGGYAADPTEVVNYVSKH
DNQTLWDMISYKAAQEADLDTRVRMQAVSLATVMLGQGIAFDQQGSELLRSKSFTRDSYDSGDWFNRVDYSLQDNNYNVGMPRSSDDGSNYDIIARVKDAVATPGETELKQMTAFYQELTALRKSSPLFTLGDGATVMK
RVDFRNTGADQQTGLLVMTIDDGMQAGASLDSRVDGIVVAINAAPESRTLQDFAGTSLQLSAIQQAAGDRSLASGVQVAAEDGSVTLPAWSVAVLELPQGESQGAGLPVSSK
this seq match to pattern  ['STVEEV'] in
 >2FHB_1|Chain A|pullulanase|Klebsiella aerogenes (548)
 CDNSSSSSTSGSPGSPGNPGNPGNPGTPGTPDPQDVVVRLPDVAVPGEAVQASARQAVIHLVDIAGITSSTPADYATKNLYLWNNETCDALSAPVADWNDVSTTPTGSDKYGPYWVIPLTKESGCINVIVRDGTNKLIDSD
LRVSFSDFTDRTVSVIAGNSAVYDSRADAFRAAFGVALADAHWVDKTTLLWPGGENKPIVRLYYSHSSKVAADSNGEFSDKYVKLTPTTVNQQVSMRFPHLASYPAFKLPDDVNVDELLQGETVAIAAESDGILSSATQ
VQTAGVLDDTYAAAAEALSYGAQLTDSGVTFRVWAPTAQQVELVIYSADKKVIASHPMTRDSASGAWSWQGGSDLKGAFYRYAMTVYHPQSRKVEQYEVTDPYAHSLSTNSEYSQVVDLNDSALKPEGWDGLTMPHAQK
TKADLAKMTIHESHIRDLSAWDQTVPAELRGKYLALTAQESNMVQHLKQLSASGVTHIELLPVFDLATVNEFSDKVADIQQPFSRLCEVNSAVKSSEFAGYCDSGSTVEEVLTQLKQNDSKDNPQVQALNTLVAQTDSY
NWGYDPFHYTVPEGSYATDPEGTARIKEFRTMIQAIKQDLGMNVIMDVVYNHTNAAGPTDRTSVLDKIVPWYYQRLNETTGSVESATCCSDSAPEHRMFAKLIADSLAVWTTDYKIDGFRFDLMLYHPKAQILSAWERI
KALNPDIYFFGEGWDSNQSDRFEIASQINLKGTGIGTFSDRLRDAVRGGGPFDSGDALRQNQGVGSGAGVLPNELTTLSDDQARHLADLTRLGMAGNLADFVLIDKDGAVKRGSEIDYNGAPGGYAADPTEVVNYVSKH
DNQTLWDMISYKAAQEADLDTRVRMQAVSLATVMLGQGIAFDQQGSELLRSKSFTRDSYDSGDWFNRVDYSLQDNNYNVGMPRSSDDGSNYDIIARVKDAVATPGETELKQMTAFYQELTALRKSSPLFTLGDGATVMK
RVDFRNTGADQQTGLLVMTIDDGMQAGASLDSRVDGIVVAINAAPESRTLQDFAGTSLQLSAIQQAAGDRSLASGVQVAAEDGSVTLPAWSVAVLELPQGESQGAGLPVSSK
this seq match to pattern  ['STVEEV'] in
 >2FHC_1|Chain A|pullulanase|Klebsiella aerogenes (548)
 CDNSSSSSTSGSPGSPGNPGNPGNPGTPGTPDPQDVVVRLPDVAVPGEAVQASARQAVIHLVDIAGITSSTPADYATKNLYLWNNETCDALSAPVADWNDVSTTPTGSDKYGPYWVIPLTKESGCINVIVRDGTNKLIDSD
LRVSFSDFTDRTVSVIAGNSAVYDSRADAFRAAFGVALADAHWVDKTTLLWPGGENKPIVRLYYSHSSKVAADSNGEFSDKYVKLTPTTVNQQVSMRFPHLASYPAFKLPDDVNVDELLQGETVAIAAESDGILSSATQ
VQTAGVLDDTYAAAAEALSYGAQLTDSGVTFRVWAPTAQQVELVIYSADKKVIASHPMTRDSASGAWSWQGGSDLKGAFYRYAMTVYHPQSRKVEQYEVTDPYAHSLSTNSEYSQVVDLNDSALKPEGWDGLTMPHAQK
TKADLAKMTIHESHIRDLSAWDQTVPAELRGKYLALTAQESNMVQHLKQLSASGVTHIELLPVFDLATVNEFSDKVADIQQPFSRLCEVNSAVKSSEFAGYCDSGSTVEEVLTQLKQNDSKDNPQVQALNTLVAQTDSY
NWGYDPFHYTVPEGSYATDPEGTARIKEFRTMIQAIKQDLGMNVIMDVVYNHTNAAGPTDRTSVLDKIVPWYYQRLNETTGSVESATCCSDSAPEHRMFAKLIADSLAVWTTDYKIDGFRFDLMLYHPKAQILSAWERI
KALNPDIYFFGEGWDSNQSDRFEIASQINLKGTGIGTFSDRLRDAVRGGGPFDSGDALRQNQGVGSGAGVLPNELTTLSDDQARHLADLTRLGMAGNLADFVLIDKDGAVKRGSEIDYNGAPGGYAADPTEVVNYVSKH
DNQTLWDMISYKAAQEADLDTRVRMQAVSLATVMLGQGIAFDQQGSELLRSKSFTRDSYDSGDWFNRVDYSLQDNNYNVGMPRSSDDGSNYDIIARVKDAVATPGETELKQMTAFYQELTALRKSSPLFTLGDGATVMK
RVDFRNTGADQQTGLLVMTIDDGMQAGASLDSRVDGIVVAINAAPESRTLQDFAGTSLQLSAIQQAAGDRSLASGVQVAAEDGSVTLPAWSVAVLELPQGESQGAGLPVSSK
this seq match to pattern  ['STVEEV'] in
```

```
>2FHF_1|Chain A|pullulanase|Klebsiella aerogenes (548)
CDNSSSSSTSGSPGSPGNPGNPGTPGTPDPQDVVVRLPDVAVPGEAVQASARQQAVIHLVDIAGITSSTPADYATKNLYLWNNETCDALSAPVADWNDVSTTPTGSDKYGPYWVIPLTKESGCINVIVRDGTNKLIDSD
LRVSFSDFTDRTVSVIAGNSAVYDSRADAFRAAFGVALADAHWVDKTTLLWPGGENKPIVRLYYSHSSKVAADSNGEFSDKYVKLTPTTVNQQVSMRFPHLASYPAFKLPDDVNVDELLQGETVAIAAESDGILSSATQ
VQTAGVLDDTYAAAAEALSYGAQLTDSGVTFRVWAPTAQQVELVIYSADKKVIASHPMTRDSASGAWSWQGGSDLKGAFYRYAMTVYHPQSRKVEQYEVTDPYAHSLSTNSEYSQVVDLNDSALKPEGWDGLTMPHAQK
TKADLAKMTIHESHIRDLSAWDQTVPAELRGKYLALTAQESNMVQHLKQLSASGVTHIELLPVFDLATVNEFSDKVADIQQPFSRLCEVNSAVKSSEFAGYCDSGSTVEEVLTQLKQNDSKDNPQVQALNTLVAQTDSY
NWGYDPFHYTVPEGSYATDPEGTARIKEFRTMIQAIKQDLGMNVIMDVVYNHTNAAGPTDRTSVLDKIVPWYYQRLNETTGSVESATCCSDSAPEHRMFAKLIADSLAVWTTDYKIDGFRFDLMLYHPKAQILSAWERI
KALNPDIYFFGEGWDSNQSDRFEIASQINLKGTGIGTFSDRLRDAVRGGGPFDSGDALRQNQGVGSGAGVLPNELTTLSDDQARHLADLTRLGMAGNLADFVLIDKDGAVKRGSEIDYNGAPGGYAADPTEVVNYVSKH
DNQTLWDMISYKAAQEADLDTRVRMQAVSLATVMLGQGIAFDQQGSELLRSKSFTRDSYDSGDWFNRVDYSLQDNNYNVGMPRSSDDGSNYDIIARVKDAVATPGETELKQMTAFYQELTALRKSSPLFTLGDGATVMK
RVDFRNTGADQQTGLLVMTIDDGMQAGASLDSRVDGIVVAINAAPESRTLQDFAGTSLQLSAIQQAAGDRSLASGVQVAADGSVTLPAWSVAVLELPQGESQGAGLPVSSK
this seq match to pattern  ['STVEKT'] in
>2FL5_1|Chains A,C,E,L|Immunoglobulin Igg1 Lambda Light Chain|Homo sapiens (9606)
SYELKQPPSVSVSPGQTARITCSGDVLPKKYAYWYQERSGQAPVLVVYEDSGRPSEIPERFSGSSSGTKATLTISGAQVEDEADYYCYSDISNGYPLFGGGTKLSVGQPKAAPSVTLFPPSSEELQANKATLVCLISD
FYPGAVTVAWKADSSPIKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPT
this seq match to pattern  ['STVEKT'] in
>2FL5_1|Chains A,C,E,L|Immunoglobulin Igg1 Lambda Light Chain|Homo sapiens (9606)
SYELKQPPSVSVSPGQTARITCSGDVLPKKYAYWYQERSGQAPVLVVYEDSGRPSEIPERFSGSSSGTKATLTISGAQVEDEADYYCYSDISNGYPLFGGGTKLSVGQPKAAPSVTLFPPSSEELQANKATLVCLISD
FYPGAVTVAWKADSSPIKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPT
this seq match to pattern  ['STVEKT'] in
>2FL5_1|Chains A,C,E,L|Immunoglobulin Igg1 Lambda Light Chain|Homo sapiens (9606)
SYELKQPPSVSVSPGQTARITCSGDVLPKKYAYWYQERSGQAPVLVVYEDSGRPSEIPERFSGSSSGTKATLTISGAQVEDEADYYCYSDISNGYPLFGGGTKLSVGQPKAAPSVTLFPPSSEELQANKATLVCLISD
FYPGAVTVAWKADSSPIKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPT
this seq match to pattern  ['STVKDT'] in
>2FL5_1|Chains A,C,E,L|Immunoglobulin Igg1 Lambda Light Chain|Homo sapiens (9606)
SYELKQPPSVSVSPGQTARITCSGDVLPKKYAYWYQERSGQAPVLVVYEDSGRPSEIPERFSGSSSGTKATLTISGAQVEDEADYYCYSDISNGYPLFGGGTKLSVGQPKAAPSVTLFPPSSEELQANKATLVCLISD
FYPGAVTVAWKADSSPIKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPT
this seq match to pattern  ['STVKDT'] in
>2G3F_1|Chains A,B|Imidazolonepropionase|Bacillus subtilis (1423)
MPKQIDTILINIGQLLTMESSGPRAGKSMQDLHVIEDAVVGIHEQKIVFAGQKGAEAGYEADEIIDCSGRLVTPGLVDPHTHLVFGGSREKEMNLKLQGISYLDILAQGGGILSTVKDTRAASEEELLQKAHFHLQRM
LSYGTTTAEVKSGYGLEKETELKQLRVAKKLHESQPVDLVSTFMGAHAIPPEYQNDPDDFLDQMLSLLPEIKEQELASFADIFTETGVFTVSQSRRYLQKAAEAGFGLKIHADEIDPLGGAELAGKLKAVSADHLVGTS
DEGIKKLAEAGTIAVLLPGTTFYLGKSTYARARAMIDEGVCVSLATDFNPGSSPTENIQLIMSIAALHLKMTAEEIWHAVTVNAAYAIGKGEEAGQLKAGRSADLVIWQAPNYMYIPYHYGVNHVHQVMKNGTIVVNRE
GAILG
this seq match to pattern  ['STVKDT'] in
>2G3F_1|Chains A,B|Imidazolonepropionase|Bacillus subtilis (1423)
MPKQIDTILINIGQLLTMESSGPRAGKSMQDLHVIEDAVVGIHEQKIVFAGQKGAEAGYEADEIIDCSGRLVTPGLVDPHTHLVFGGSREKEMNLKLQGISYLDILAQGGGILSTVKDTRAASEEELLQKAHFHLQRM
LSYGTTTAEVKSGYGLEKETELKQLRVAKKLHESQPVDLVSTFMGAHAIPPEYQNDPDDFLDQMLSLLPEIKEQELASFADIFTETGVFTVSQSRRYLQKAAEAGFGLKIHADEIDPLGGAELAGKLKAVSADHLVGTS
DEGIKKLAEAGTIAVLLPGTTFYLGKSTYARARAMIDEGVCVSLATDFNPGSSPTENIQLIMSIAALHLKMTAEEIWHAVTVNAAYAIGKGEEAGQLKAGRSADLVIWQAPNYMYIPYHYGVNHVHQVMKNGTIVVNRE
GAILG
this seq match to pattern  ['STVKRK'] in
>2G4A_1|Chain A|Bromodomain-containing protein 2|Homo sapiens (9606)
EQLKHCNVILKELLSKKHAAYAWPFYKPVDASALGLHDYHDIIKHPMDLSTVKRKMENRDYRDAQEFAADVRLMFSNCYKYNPPDHDVVAMARKLQDVFEFRYAKMPDLEHHHHHH
this seq match to pattern  ['STVEKT'] in
>2G75_2|Chains B,D|IGG Light Chain|Homo sapiens (9606)
SYELTQPPSVSVAPGKTARITCGGNNIGSKSVHWYQQKPGQAPVLVVYDDSDRPSGIPERFSGSNSGNTATLTISRVEAGDEADYYCQVWDSSSDYVFGTGTKVTVLGQPKANPTVTLFPPSSEEFQANKATLVCLIS
DFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
>2G75_2|Chains B,D|IGG Light Chain|Homo sapiens (9606)
SYELTQPPSVSVAPGKTARITCGGNNIGSKSVHWYQQKPGQAPVLVVYDDSDRPSGIPERFSGSNSGNTATLTISRVEAGDEADYYCQVWDSSSDYVFGTGTKVTVLGQPKANPTVTLFPPSSEEFQANKATLVCLIS
DFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVKDE'] in
>2GAN_1|Chains A,B|182aa long hypothetical protein|Pyrococcus horikoshii (70601)
MEGVKKIKNPSTVKDELLELMFRIYRSTNGKYPALEWVKRKPNPNDFNGFREVYEPFLKFRLSQEFDELYTYQKDNRIIGTIALVYKRIKEKGIWWVPEELMNEKVGLIEFFVVDPEFQGKGIGSTLLEFAVKRLRSL
GKDPYVVTFPNLEAYSYYYMKKGFREIMRYKEFVILKFNHKKFQLEHHHHHH
this seq match to pattern  ['STVKDE'] in
>2GAN_1|Chains A,B|182aa long hypothetical protein|Pyrococcus horikoshii (70601)
MEGVKKIKNPSTVKDELLELMFRIYRSTNGKYPALEWVKRKPNPNDFNGFREVYEPFLKFRLSQEFDELYTYQKDNRIIGTIALVYKRIKEKGIWWVPEELMNEKVGLIEFFVVDPEFQGKGIGSTLLEFAVKRLRSL
GKDPYVVTFPNLEAYSYYYMKKGFREIMRYKEFVILKFNHKKFQLEHHHHHH
this seq match to pattern  ['STVEKT'] in
>2H32_2|Chain B|Immunoglobulin omega chain|Homo sapiens (9606)
SVTHVFGSGTQLTVLSQPKATPSVTLFPPSSEELQANKATLVCLMNDFYPGILTVTWKADGTPITQGVEMTTPSKQSNNKYAASSYLSLTPEQWRSRRSYSCQVMHEGSTVEKTVAPAECS
this seq match to pattern  ['STVEKT'] in
>2H3N_2|Chains B,D|Ig lambda-5|Homo sapiens (9606)
VTHVFGSGTQLTVLSQPKATPSVTLFPPSSEELQANKATLVCLMNDFYPGILTVTWKADGTPITQGVEMTTPSKQSNNKYAASSYLSLTPEQWRSRRSYSCQVMHEGSTVEKTVAPA
this seq match to pattern  ['STVEKT'] in
>2H3N_2|Chains B,D|Ig lambda-5|Homo sapiens (9606)
VTHVFGSGTQLTVLSQPKATPSVTLFPPSSEELQANKATLVCLMNDFYPGILTVTWKADGTPITQGVEMTTPSKQSNNKYAASSYLSLTPEQWRSRRSYSCQVMHEGSTVEKTVAPA
this seq match to pattern  ['STVEDF'] in
>2IDR_1|Chains A,B|Eukaryotic translation initiation factor 4E-1|Triticum aestivum (4565)
AHPLENAWTFWFDNPQGKSRQVAWGSTIHPIHTFSTVEDFWGLYNNIHNPSKLNVGADFHCFKNKIEPKWEDPICANGGKWTISCGRGKSDTFWLHTLLAMIGEQFDFGDEICGAVVSVRQKQERVAIWTKNAANEAA
QISIGKQWKEFLDYKDSIGFIVHEDAKRSDKGPKNRYTV
this seq match to pattern  ['STVEDF'] in
>2IDR_1|Chains A,B|Eukaryotic translation initiation factor 4E-1|Triticum aestivum (4565)
AHPLENAWTFWFDNPQGKSRQVAWGSTIHPIHTFSTVEDFWGLYNNIHNPSKLNVGADFHCFKNKIEPKWEDPICANGGKWTISCGRGKSDTFWLHTLLAMIGEQFDFGDEICGAVVSVRQKQERVAIWTKNAANEAA
QISIGKQWKEFLDYKDSIGFIVHEDAKRSDKGPKNRYTV
this seq match to pattern  ['STVEDF'] in
>2IDV_1|Chain A|Eukaryotic translation initiation factor 4E-1|Triticum aestivum (4565)
AHPLENAWTFWFDNPQGKSRQVAWGSTIHPIHTFSTVEDFWGLYNNIHNPSKLNVGADFHCFKNKIEPKWEDPISANGGKWTISCGRGKSDTFWLHTLLAMIGEQFDFGDEICGAVVSVRQKQERVAIWTKNAANEAA
QISIGKQWKEFLDYKDSIGFIVHEDAKRSDKGPKNRYTV
this seq match to pattern  ['STVEKT'] in
>2IG2_1|Chain L|IGG1-LAMBDA KOL FAB (LIGHT CHAIN)|Homo sapiens (9606)
QSVLTQPPSASGTPGQRVTISCSGTSSNIGSSTVNWYQQLPGMAPKLLIYRDAMRPSGVPDRFSGSKSGASASLAIGGLQSEDETDYYCAAWDVSLNAYVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKA'] in
>2J28_23|Chain O|50S RIBOSOMAL PROTEIN L18|ESCHERICHIA COLI (562)
MDKKSARIRRATRARRKLQELGATRLVVHRTPRHIYAQVIAPNGSEVLVAASTVEKAIAEQLKYTGNKDAAAAVGKAVAERALEKGIKDVSFDRSGFQYHGRVQALADAAREAGLQF
this seq match to pattern  ['STVDDT'] in
>2J42_1|Chain A|C2 TOXIN COMPONENT-II|CLOSTRIDIUM BOTULINUM (1491)
MLVSKFENSVKNSNKNYFTINGLMGYYFENDFFNLNIISPTLDGNLTFSKEDINSILGNKIIKSARWIGLIKPSITGEYILSTNSPNCRVELNGEIFNLSLNTSNTVNLIQGNVYDIRIEQLMSENQLLKNYEGIKLY
WETSDIIKEIIPSEVLLKPNYSNTNEKSKFIPNNTLFSNAKLKANANRDTDRDGIPDEWEINGYTVMNQKAVAWDDKFAANGYKKYVSNPFKPCTANDPYTDFEKVSGQIDPSVSMVARDPMISAYPIVGVQMERLVVS
KSETITGDSTKSMSKSTSHSSTNINTVGAEVSGSLQLAGGIFPVFSMSASANYSHTWQNTSTVDDTTGESFSQGLSINTAESAYINPNIRYYNTGTAPVYNVTPTTTIVIDKQSVATIKGQESLIGDYLNPGGTYPIIG
EPPMALNTMDQFSSRLIPINYNQLKSIDNGGTVMLSTSQFTGNFAKYNSNGNLVTDGNNWGPYLGTIKSTTASLTLSLPDQTTQVAVVAPNFSDPEDKTPRLTLEQALVKAFRLEKKNGKFYFHGMEISANQKIQVFLD
RNTNVDFENQLKNTANKDIMNCIIKRNMNILVKVITFKENISSINIINDTNFGVESMTGLSKRIKGNDGIYRASTKSFSFKSKEIKYPEGFYRMRFVIQSYEPFTCNFKLFNNLIYSNSFDIGYYDEFFYFYCNGSKSF
FDISCDIINSINRLSGVFLIELDKLII
this seq match to pattern  ['STVEKT'] in
>2J6E_3|Chains L,M|IGM|HOMO SAPIENS (9606)
MAGFPLLLTLLTHCAGSWAQSVLTQPPSASGTPGQRVTISCSGSSSNIGSNYVYWYQQLPGTAPKLLIYRNNQRPSGVPDRFSGSKSGTSASLAISGLRSEDEADYYCATWDDSLSAVIFGGGTKLTVLGQPKAAPSV
TLFPPSSEELQANKATLVCLISDFFPGAVTVAWKADGAPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTEC
this seq match to pattern  ['STVEKT'] in
>2J6E_3|Chains L,M|IGM|HOMO SAPIENS (9606)
```

MAGFPLLLTLLTHCAGSWAQSVLTQPPSASGTPGQRVTISCSGSSSNIGSNYVYWYQQLPGTAPKLLIYRNNQRPSGVPDRFSGSKSGTSASLAISGLRSEDEADYYCATWDDSLSAVIFGGGTKLTVLGQPKAAPSV
TLFPPSSEELQANKATLVCLISDFFPGAVTVAWKADGAPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTEC
this seq match to pattern ['STVEKT'] in
>2JB5_2|Chain L|FAB FRAGMENT MOR03268 LIGHT CHAIN|HOMO SAPIENS (9606)
DIALTQPASVSGSPGQSITISCTGTSSDVGSNNYVSWYQQHPGKAPKLMIYGGSNRPSGVSNRFSGSKSGNTASLTISGLQAEDEADYYCRSWDSNLSYSVFGGGTKLTVLGQPKAAPSVTLFPPSSEELQANKATLV
CLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTEA
this seq match to pattern ['STVEKT'] in
>2JB6_1|Chains A,L|FAB FRAGMENT MOR03268 LIGHT CHAIN|HOMO SAPIENS (9606)
DIALTQPASVSGSPGQSITISCTGTSSDVGSNNYVSWYQQHPGKAPKLMIYGGSNRPSGVSNRFSGSKSGNTASLTISGLQAEDEADYYCRSWDSNLSYSVFGGGTKLTVLGQPKAAPSVTLFPPSSEELQANKATLV
CLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTEA
this seq match to pattern ['STVEKT'] in
>2JB6_1|Chains A,L|FAB FRAGMENT MOR03268 LIGHT CHAIN|HOMO SAPIENS (9606)
DIALTQPASVSGSPGQSITISCTGTSSDVGSNNYVSWYQQHPGKAPKLMIYGGSNRPSGVSNRFSGSKSGNTASLTISGLQAEDEADYYCRSWDSNLSYSVFGGGTKLTVLGQPKAAPSVTLFPPSSEELQANKATLV
CLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTEA
this seq match to pattern ['VTTERY', 'STVKEF'] in
>2JE8_1|Chains A,B|BETA-MANNOSIDASE|BACTEROIDES THETAIOTAOMICRON (226186)
MGQGNDTSEVMLLDTGWEFSQSGTEKWMPATVPGTVHQDLISHELLPNPFYGMNEKKIQWVENEDWEYRTSFIVSEEQLNRDGIQLIFEGLDTYADVYLNGSLLLKADNMFVGYTLPVKSVLRKGENHLYIYFHSPIR
QTLPQYASNGFNYPADNDHHEKHLSVFSRKAPYSYGWDWGIRMVTSGVWRPVTLRFYDIATISDYYVRQLSLTDENARLSNELIVNQIVPQKIPAEVRVNVSLNGTTVTEVKQQVTLQPGINHITLPAEVTNPVRWMPN
GWGTPTLYDFSAQIACGDRIVAEQSHRIGLRTIRVVNEKDKDGESFYFEVNGIPMFAKGANYIPQDALLPNVTTERYQTLFRDMKEANMNMVRIWGGGTYENNLFYDLADENGILVWQDFMFACTPYPSDPTFLKRVEA
EAVYNIRRLRNHASLAMWCGNNEILEALKYWGFEKKFTPEVYQGLMHGYDKLFRELLPSTVKEFDSDRFYVHSSPYLANWGRPESWGTGDSHNWGVWYGKKPFESLDTDLPRFMSEFGFQSFPEMKTIAAFAAPEDYQI
ESEVMNAHQKSSIGNSLIRTYMERDYIIPESFEDFVYVGLVLQGQGMRHGLEAHRRNRPYCMGTLYWQLNDSWPVVSWSSIDYYGNWKALHYQAKRAFAPVLINPIQQNDSLSVYLISDRLDTMEQMTLEMKVVDFDGK
TLGKKIQVHSLEVPANTSKCVYRAKLDGWLTPEDCRRSFLKLILKDKSGHQVAESVHFFRKTKDLQLPPTSVSYQMKQTDGKCELTLFSSMLAKDIFIETPLQGARYSDNFFDLLPGERKKVIITSPRIKKGEELPVNI
KHIRETYKEHHHHHH
this seq match to pattern ['VTTERY', 'STVKEF'] in
>2JE8_1|Chains A,B|BETA-MANNOSIDASE|BACTEROIDES THETAIOTAOMICRON (226186)
MGQGNDTSEVMLLDTGWEFSQSGTEKWMPATVPGTVHQDLISHELLPNPFYGMNEKKIQWVENEDWEYRTSFIVSEEQLNRDGIQLIFEGLDTYADVYLNGSLLLKADNMFVGYTLPVKSVLRKGENHLYIYFHSPIR
QTLPQYASNGFNYPADNDHHEKHLSVFSRKAPYSYGWDWGIRMVTSGVWRPVTLRFYDIATISDYYVRQLSLTDENARLSNELIVNQIVPQKIPAEVRVNVSLNGTTVTEVKQQVTLQPGINHITLPAEVTNPVRWMPN
GWGTPTLYDFSAQIACGDRIVAEQSHRIGLRTIRVVNEKDKDGESFYFEVNGIPMFAKGANYIPQDALLPNVTTERYQTLFRDMKEANMNMVRIWGGGTYENNLFYDLADENGILVWQDFMFACTPYPSDPTFLKRVEA
EAVYNIRRLRNHASLAMWCGNNEILEALKYWGFEKKFTPEVYQGLMHGYDKLFRELLPSTVKEFDSDRFYVHSSPYLANWGRPESWGTGDSHNWGVWYGKKPFESLDTDLPRFMSEFGFQSFPEMKTIAAFAAPEDYQI
ESEVMNAHQKSSIGNSLIRTYMERDYIIPESFEDFVYVGLVLQGQGMRHGLEAHRRNRPYCMGTLYWQLNDSWPVVSWSSIDYYGNWKALHYQAKRAFAPVLINPIQQNDSLSVYLISDRLDTMEQMTLEMKVVDFDGK
TLGKKIQVHSLEVPANTSKCVYRAKLDGWLTPEDCRRSFLKLILKDKSGHQVAESVHFFRKTKDLQLPPTSVSYQMKQTDGKCELTLFSSMLAKDIFIETPLQGARYSDNFFDLLPGERKKVIITSPRIKKGEELPVNI
KHIRETYKEHHHHHH
this seq match to pattern ['STVKDK'] in
>2JQ3_1|Chain A|Apolipoprotein C-III|Homo sapiens (9606)
SEAEADASLLSFMQGYMKHATKTAKDALSSVQESQVAQQARGWVTDGFSSLKDYWSTVKDKFSEFWDLDPEVRPTSAVAA
this seq match to pattern ['STVREQ'] in
>2KC8_1|Chain A|Toxin relE|Escherichia coli (83333)
GSHMAYFLDFDERALKEWRKLGSTVREQLKKKLVEVLESPRIEANKLRGMPDCYKIKLRSSGYRLVYQVIDEKVVVFVISVGKAEASEVYSEAVKRIL
this seq match to pattern ['STVREQ'] in
>2KC9_1|Chain A|Toxin relE|Escherichia coli (83333)
GSHMAYFLDFDERALKEWRKLGSTVREQLKKKLVEVLESPRIEANKLRGMPDCYKIKLRSSGYRLVYQVIDEKVVVFVISVGKAEASEVYSEAVKRIL
this seq match to pattern ['STVKEQ'] in
>2LDX_1|Chains A,B,C,D|APO-LACTATE DEHYDROGENASE|Mus musculus (10090)
STVKEQLIQNLVPEDKLSRCKITVVGVGDVGMACAISILLKGLADELALVDADTDKLRGEALDLQHGSLFLSTPKIVFGKDYNVSANSKLVIITAGARMVSGQTRLDLLQRNVAIMKAIVPGVIQNSPDCKIIVVTNP
VDILTYVVWKISGFPVGRVIGSGCNLDSARFRYLIGEKLGVNPTSCHGWVLGEHGDSSVPIWSGVNVAGVTLKSLNPAIGTDKNKQHWKNVHKQVVEGGYEVLDMKGYTSWAIGLSVTDLARSILKNLKRVHPVTTLVK
GFHGIKEEVFLSIPCVLGESGITDFVKVNMTAEEEGLLKKSADTLWNMQKNLEL
this seq match to pattern ['STVKEQ'] in
>2LDX_1|Chains A,B,C,D|APO-LACTATE DEHYDROGENASE|Mus musculus (10090)
STVKEQLIQNLVPEDKLSRCKITVVGVGDVGMACAISILLKGLADELALVDADTDKLRGEALDLQHGSLFLSTPKIVFGKDYNVSANSKLVIITAGARMVSGQTRLDLLQRNVAIMKAIVPGVIQNSPDCKIIVVTNP
VDILTYVVWKISGFPVGRVIGSGCNLDSARFRYLIGEKLGVNPTSCHGWVLGEHGDSSVPIWSGVNVAGVTLKSLNPAIGTDKNKQHWKNVHKQVVEGGYEVLDMKGYTSWAIGLSVTDLARSILKNLKRVHPVTTLVK
GFHGIKEEVFLSIPCVLGESGITDFVKVNMTAEEEGLLKKSADTLWNMQKNLEL
this seq match to pattern ['STVKEQ'] in
>2LDX_1|Chains A,B,C,D|APO-LACTATE DEHYDROGENASE|Mus musculus (10090)
STVKEQLIQNLVPEDKLSRCKITVVGVGDVGMACAISILLKGLADELALVDADTDKLRGEALDLQHGSLFLSTPKIVFGKDYNVSANSKLVIITAGARMVSGQTRLDLLQRNVAIMKAIVPGVIQNSPDCKIIVVTNP
VDILTYVVWKISGFPVGRVIGSGCNLDSARFRYLIGEKLGVNPTSCHGWVLGEHGDSSVPIWSGVNVAGVTLKSLNPAIGTDKNKQHWKNVHKQVVEGGYEVLDMKGYTSWAIGLSVTDLARSILKNLKRVHPVTTLVK
GFHGIKEEVFLSIPCVLGESGITDFVKVNMTAEEEGLLKKSADTLWNMQKNLEL
this seq match to pattern ['STVKEQ'] in
>2LDX_1|Chains A,B,C,D|APO-LACTATE DEHYDROGENASE|Mus musculus (10090)
STVKEQLIQNLVPEDKLSRCKITVVGVGDVGMACAISILLKGLADELALVDADTDKLRGEALDLQHGSLFLSTPKIVFGKDYNVSANSKLVIITAGARMVSGQTRLDLLQRNVAIMKAIVPGVIQNSPDCKIIVVTNP
VDILTYVVWKISGFPVGRVIGSGCNLDSARFRYLIGEKLGVNPTSCHGWVLGEHGDSSVPIWSGVNVAGVTLKSLNPAIGTDKNKQHWKNVHKQVVEGGYEVLDMKGYTSWAIGLSVTDLARSILKNLKRVHPVTTLVK
GFHGIKEEVFLSIPCVLGESGITDFVKVNMTAEEEGLLKKSADTLWNMQKNLEL
this seq match to pattern ['STVKET'] in
>2MCG_1|Chains 1,2|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
QSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern ['STVKET'] in
>2MCG_1|Chains 1,2|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
QSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern ['STVRRR'] in
>2MXC_1|Chain A|Sorting nexin-3|Homo sapiens (9606)
MAHHHHHHVGTAETVADTRRLITKPQNLNDAYGPPSNFLEIDVSNPQTVGVGRGRFTTYEIRVKTNLPIFKLKESTVRRRYSDFEWLRSELERESKVVVPPLPGKAFLRQLPFRGDDGIFDDNFIEERKQGLEQFINK
VAGHPLAQNERCLHMFLQDEIIDKSYTPSKIRHA
this seq match to pattern ['STVKEN', 'STVKEQ'] in
>2OAJ_1|Chain A|Protein SNI1|Saccharomyces cerevisiae (4932)
NKNKIFSLAETNKYGMSSKPIAAAFDFTQNLLAIATVTGEVHIYGQQQVEVVIKLEDRSAIKEMRFVKGIYLVVINAKDTVYVLSLYSQKVLTTVFVPGKITSIDTDASLDWMLIGLQNGSMIVYDIDRDQLSSFKLD
NLQKSSFFPAARLSPIVSIQWNPRDIGTVLISYEYVTLTYSLVENEIKQSFIYELPPFAPGGDFSEKTNEKRTPKVIQSLYHPNSLHIITIHEDNSLVFWDANSGHMIMARTVFETEINVPQPDYIRDSSTNAAKISKV
YWMCENNPEYTSLLISHKSISRGDNQSLTMIDLGYTPRYSITSYEGMKNYYANPKQMKIFPLPTNVPIVNILPIPRQSPYFAGCHNPGLILLILGNGEIETMLYPSGIFTDKASLFPQNLSWLRPLATTSMAASVPNKL
WLGALSAAQNKDYLLGGVRTKRQKLPAEYGTAFITGHSNGSVRIYDASHGDIQDNASFEVNLSRTLNKAKELAVDKISFAAETLELAVSIETGDVVLFKYEVNQFYSVENRPESGDLEMNFRRFSLNNTNGVLVDVRD
RAPTGVRQGFMPSTAVHANKGKTSAINNSNIGFVGIAYAAGSLMLIDRRGPAIIYMENIREISGAQSACVTCIEFVIMEYGDDGYSSILMVCGTDMGEVITYKILPASGGKFDVQLMDITNVTSKGPIHKIDAFSKETK
SSCLATIPKMQNLSKGLCIPGIVLITGFDDIRLITLGKSKSTHKGFKYPLAATGLSYISTVEKNNDRKNLTVIITLEINGHLRVFTIPDFKEQMSEHIPFPIAAKYITESSVLRNGDIAIRVSEFQASLFSTVKEQDTL
APVSDTLYINGIRIPYRPQVNSLQWARGTVYCTPAQLNELLGGVNRPASKYKESIIAEGSFSERSSDDN
this seq match to pattern ['STVKET'] in
>2OLD_1|Chains A,B|Bence Jones KWR Protein - Immunoglobulin Light Chain|Homo sapiens (9606)
QSALPQPASVSGSPGQSITISCTGTSSDVGGYDLVSWYQHHPGGAPKLIIYEVTNRPSGVSDRFSGSKSGNTASLTISGLQAEDEADYYCSSYASGSTPRIFGGGTRLTVLGQPKAAPSVTLFPPSSEELQANKATLV
CLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern ['STVKET'] in
>2OLD_1|Chains A,B|Bence Jones KWR Protein - Immunoglobulin Light Chain|Homo sapiens (9606)
QSALPQPASVSGSPGQSITISCTGTSSDVGGYDLVSWYQHHPGGAPKLIIYEVTNRPSGVSDRFSGSKSGNTASLTISGLQAEDEADYYCSSYASGSTPRIFGGGTRLTVLGQPKAAPSVTLFPPSSEELQANKATLV
CLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern ['STVKET'] in
>2OMB_1|Chains A,B,C,D|Bence Jones KWR Protein - Immunoglobulin Light Chain|Homo sapiens (9606)
QSALPQPASVSGSPGQSITISCTGTSSDVGGYDLVSWYQHHPGGAPKLIIYEVTNRPSGVSDRFSGSKSGNTASLTISGLQAEDEADYYCSSYASGSTPRIFGGGTRLTVLGQPKAAPSVTLFPPSSEELQANKATLV
CLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern ['STVKET'] in
>2OMB_1|Chains A,B,C,D|Bence Jones KWR Protein - Immunoglobulin Light Chain|Homo sapiens (9606)
QSALPQPASVSGSPGQSITISCTGTSSDVGGYDLVSWYQHHPGGAPKLIIYEVTNRPSGVSDRFSGSKSGNTASLTISGLQAEDEADYYCSSYASGSTPRIFGGGTRLTVLGQPKAAPSVTLFPPSSEELQANKATLV
CLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS

```
this seq match to pattern  ['STVEKT'] in
 >2OMB_1|Chains A,B,C,D|Bence Jones KWR Protein - Immunoglobulin Light Chain|Homo sapiens (9606)
 QSALPQPASVSGSPGQSITISCTGTSSDVGGYDLVSWYQHHPGGAPKLIIYEVTNRPSGVSDRFSGSKSGNTASLTISGLQAEDEADYYCSSYASGSTPRIFGGGTRLTVLGQPKAAPSVTLFPPSSEELQANKATLV
 CLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >2OMB_1|Chains A,B,C,D|Bence Jones KWR Protein - Immunoglobulin Light Chain|Homo sapiens (9606)
 QSALPQPASVSGSPGQSITISCTGTSSDVGGYDLVSWYQHHPGGAPKLIIYEVTNRPSGVSDRFSGSKSGNTASLTISGLQAEDEADYYCSSYASGSTPRIFGGGTRLTVLGQPKAAPSVTLFPPSSEELQANKATLV
 CLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >2OMN_1|Chains A,B|Bence Jones KWR Protein - Immunoglobulin Light Chain|Homo sapiens (9606)
 QSALPQPASVSGSPGQSITISCTGTSSDVGGYDLVSWYQHHPGGAPKLIIYEVTNRPSGVSDRFSGSKSGNTASLTISGLQAEDEADYYCSSYASGSTPRIFGGGTRLTVLGQPKAAPSVTLFPPSSEELQANKATLV
 CLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVEKT'] in
 >2OMN_1|Chains A,B|Bence Jones KWR Protein - Immunoglobulin Light Chain|Homo sapiens (9606)
 QSALPQPASVSGSPGQSITISCTGTSSDVGGYDLVSWYQHHPGGAPKLIIYEVTNRPSGVSDRFSGSKSGNTASLTISGLQAEDEADYYCSSYASGSTPRIFGGGTRLTVLGQPKAAPSVTLFPPSSEELQANKATLV
 CLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern  ['STVDDD'] in
 >2PI2_1|Chains A,B,C,D|Replication protein A 32 kDa subunit|Homo sapiens (9606)
 MWNSGFESYGSSSYGGAGGYTQSPGGFGSPAPSQAEKKSRARAQHIVPCTISQLLSATLVDEVFRIGNVEISQVTIVGIIRHAEKAPTNIVYKIDDMTAAPMDVRQWVDTDDTSSENTVVPPETYVKVAGHLRSFQNK
 KSLVAFKIMPLEDMNEFTTHILEVINAHMVLSKANSQPSAGRAPISNPGMSEAGNFGGNSFMPANGLTVAQNQVLNLIKACPRPEGLNFQDLKNQLKHMSVSSIKQAVDFLSNEGHIYSTVDDDHFKSTDAE
this seq match to pattern  ['STVDDD'] in
 >2PI2_1|Chains A,B,C,D|Replication protein A 32 kDa subunit|Homo sapiens (9606)
 MWNSGFESYGSSSYGGAGGYTQSPGGFGSPAPSQAEKKSRARAQHIVPCTISQLLSATLVDEVFRIGNVEISQVTIVGIIRHAEKAPTNIVYKIDDMTAAPMDVRQWVDTDDTSSENTVVPPETYVKVAGHLRSFQNK
 KSLVAFKIMPLEDMNEFTTHILEVINAHMVLSKANSQPSAGRAPISNPGMSEAGNFGGNSFMPANGLTVAQNQVLNLIKACPRPEGLNFQDLKNQLKHMSVSSIKQAVDFLSNEGHIYSTVDDDHFKSTDAE
this seq match to pattern  ['STVDDD'] in
 >2PI2_1|Chains A,B,C,D|Replication protein A 32 kDa subunit|Homo sapiens (9606)
 MWNSGFESYGSSSYGGAGGYTQSPGGFGSPAPSQAEKKSRARAQHIVPCTISQLLSATLVDEVFRIGNVEISQVTIVGIIRHAEKAPTNIVYKIDDMTAAPMDVRQWVDTDDTSSENTVVPPETYVKVAGHLRSFQNK
 KSLVAFKIMPLEDMNEFTTHILEVINAHMVLSKANSQPSAGRAPISNPGMSEAGNFGGNSFMPANGLTVAQNQVLNLIKACPRPEGLNFQDLKNQLKHMSVSSIKQAVDFLSNEGHIYSTVDDDHFKSTDAE
this seq match to pattern  ['STVDDD'] in
 >2PI2_1|Chains A,B,C,D|Replication protein A 32 kDa subunit|Homo sapiens (9606)
 MWNSGFESYGSSSYGGAGGYTQSPGGFGSPAPSQAEKKSRARAQHIVPCTISQLLSATLVDEVFRIGNVEISQVTIVGIIRHAEKAPTNIVYKIDDMTAAPMDVRQWVDTDDTSSENTVVPPETYVKVAGHLRSFQNK
 KSLVAFKIMPLEDMNEFTTHILEVINAHMVLSKANSQPSAGRAPISNPGMSEAGNFGGNSFMPANGLTVAQNQVLNLIKACPRPEGLNFQDLKNQLKHMSVSSIKQAVDFLSNEGHIYSTVDDDHFKSTDAE
this seq match to pattern  ['STVRKA'] in
 >2QA2_1|Chain A|Polyketide oxygenase CabE|Streptomyces (1883)
 MAHHHHHHHRSDASVIVVGAGPAGLMLAGELRLGGVDVMVLEQLPQRTGESRGLGFTARTMEVFDQRGILPAFGPVETSTQGHFGGRPVDFGVLEGAHYGVKAVPQSTTESVLEEWALGRGAELLRGHTVRALTDEGD
 HVVVEVEGPDGPRSLTTRYVVGCDGGRSTVRKAAGFDFPGTSASREMFLADIRGCEITPRPIGETVPLGMVMSAPLGDGVDRIIVCERGAPARRRTGPPPYQEVAAAWQRLTGQDISHGEPVWVSAFGDPARQVSAYRR
 GRVLLAGDSAHVHLPAGGQGMNVSVQDSVNLGWKLAAVVSGRAPAGLLDTYHEERHPVGRRLLMNTQAQGMLFLSGDEMQPLRDVLSELIRYDEVSRHLAGMVSGLDIRYEVDGGDHPLLGMRMPHQELVRAHGKTSTT
 ELLHPARGVLLDIADDAEVREAATGWSDRVDIVTASLHDAPPQGPLSDARAVLVRPDGYVAWISPGSRAGLTEALDRWFGPAR
this seq match to pattern  ['STVEKT'] in
 >2RCJ_2|Chains C,D,G,H,K,L,O,P,S,T|IgA1 heavy chain|Homo sapiens (9606)
 ESVLTQPPSASGTPGQRVTISCTGTSSGSITVNWYQQLPGMAPKLLIYRDAMRPSGVPTRFSGSKSGTSASLAISGLEAEDESDYYCASWNSSDNSYVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVCLI
 SDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPQPVTVSWNSGALTSGVHTFPAVLQSSGLY
 SLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKRVEPLSLPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPQVKFNWYVDGVQVHNAKTKPREQQYNSTYRVVSVLTVLHQNWLDGKEYKCKVSNKALPAPIEKTI
 SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSL
this seq match to pattern  ['STVEKT'] in
 >2RCJ_2|Chains C,D,G,H,K,L,O,P,S,T|IgA1 heavy chain|Homo sapiens (9606)
 ESVLTQPPSASGTPGQRVTISCTGTSSGSITVNWYQQLPGMAPKLLIYRDAMRPSGVPTRFSGSKSGTSASLAISGLEAEDESDYYCASWNSSDNSYVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVCLI
 SDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPQPVTVSWNSGALTSGVHTFPAVLQSSGLY
 SLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKRVEPLSLPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPQVKFNWYVDGVQVHNAKTKPREQQYNSTYRVVSVLTVLHQNWLDGKEYKCKVSNKALPAPIEKTI
 SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSL
this seq match to pattern  ['STVEKT'] in
 >2RCJ_2|Chains C,D,G,H,K,L,O,P,S,T|IgA1 heavy chain|Homo sapiens (9606)
 ESVLTQPPSASGTPGQRVTISCTGTSSGSITVNWYQQLPGMAPKLLIYRDAMRPSGVPTRFSGSKSGTSASLAISGLEAEDESDYYCASWNSSDNSYVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVCLI
 SDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPQPVTVSWNSGALTSGVHTFPAVLQSSGLY
 SLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKRVEPLSLPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPQVKFNWYVDGVQVHNAKTKPREQQYNSTYRVVSVLTVLHQNWLDGKEYKCKVSNKALPAPIEKTI
 SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSL
this seq match to pattern  ['STVEKT'] in
 >2RCJ_2|Chains C,D,G,H,K,L,O,P,S,T|IgA1 heavy chain|Homo sapiens (9606)
 ESVLTQPPSASGTPGQRVTISCTGTSSGSITVNWYQQLPGMAPKLLIYRDAMRPSGVPTRFSGSKSGTSASLAISGLEAEDESDYYCASWNSSDNSYVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVCLI
 SDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPQPVTVSWNSGALTSGVHTFPAVLQSSGLY
 SLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKRVEPLSLPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPQVKFNWYVDGVQVHNAKTKPREQQYNSTYRVVSVLTVLHQNWLDGKEYKCKVSNKALPAPIEKTI
 SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSL
this seq match to pattern  ['STVEKT'] in
 >2RCJ_2|Chains C,D,G,H,K,L,O,P,S,T|IgA1 heavy chain|Homo sapiens (9606)
 ESVLTQPPSASGTPGQRVTISCTGTSSGSITVNWYQQLPGMAPKLLIYRDAMRPSGVPTRFSGSKSGTSASLAISGLEAEDESDYYCASWNSSDNSYVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVCLI
 SDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPQPVTVSWNSGALTSGVHTFPAVLQSSGLY
 SLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKRVEPLSLPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPQVKFNWYVDGVQVHNAKTKPREQQYNSTYRVVSVLTVLHQNWLDGKEYKCKVSNKALPAPIEKTI
 SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSL
this seq match to pattern  ['STVEKT'] in
 >2RCJ_2|Chains C,D,G,H,K,L,O,P,S,T|IgA1 heavy chain|Homo sapiens (9606)
 ESVLTQPPSASGTPGQRVTISCTGTSSGSITVNWYQQLPGMAPKLLIYRDAMRPSGVPTRFSGSKSGTSASLAISGLEAEDESDYYCASWNSSDNSYVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVCLI
 SDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPQPVTVSWNSGALTSGVHTFPAVLQSSGLY
 SLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKRVEPLSLPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPQVKFNWYVDGVQVHNAKTKPREQQYNSTYRVVSVLTVLHQNWLDGKEYKCKVSNKALPAPIEKTI
 SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSL
this seq match to pattern  ['STVEKT'] in
 >2RCJ_2|Chains C,D,G,H,K,L,O,P,S,T|IgA1 heavy chain|Homo sapiens (9606)
 ESVLTQPPSASGTPGQRVTISCTGTSSGSITVNWYQQLPGMAPKLLIYRDAMRPSGVPTRFSGSKSGTSASLAISGLEAEDESDYYCASWNSSDNSYVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVCLI
 SDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPQPVTVSWNSGALTSGVHTFPAVLQSSGLY
 SLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKRVEPLSLPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPQVKFNWYVDGVQVHNAKTKPREQQYNSTYRVVSVLTVLHQNWLDGKEYKCKVSNKALPAPIEKTI
 SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSL
this seq match to pattern  ['STVEKT'] in
 >2RCJ_2|Chains C,D,G,H,K,L,O,P,S,T|IgA1 heavy chain|Homo sapiens (9606)
 ESVLTQPPSASGTPGQRVTISCTGTSSGSITVNWYQQLPGMAPKLLIYRDAMRPSGVPTRFSGSKSGTSASLAISGLEAEDESDYYCASWNSSDNSYVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVCLI
 SDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPQPVTVSWNSGALTSGVHTFPAVLQSSGLY
 SLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKRVEPLSLPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPQVKFNWYVDGVQVHNAKTKPREQQYNSTYRVVSVLTVLHQNWLDGKEYKCKVSNKALPAPIEKTI
 SKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSL
```

this seq match to pattern ['STVEKA'] in
>2RDO_14|Chain O|50S ribosomal protein L18|Escherichia coli (562)
MDKKSARIRRATRARRKLQELGATRLVVHRTPRHIYAQVIAPNGSEVLVAASTVEKAIAEQLKYTGNKDAAAAVGKAVAERALEKGIKDVSFDRSGFQYHGRVQALADAAREAGLQF
this seq match to pattern ['VTVEDS'] in
>2RDO_34|Chain 8|Ribosome recycling factor|Escherichia coli (562)
MISDIRKDAEVRMDKCVEAFKTQISKIRTGRASPSLLDGIVVEYYGTPTPLRQLASVTVEDSRTLKINVFDRSMSPAVEKAIMASDLGLNPNSAGSDIRVPLPPLTEERRKDLTKIVRGEAEQARVAVRNVRRDANDK
VKALLKDKEISEDDDRRSQDDVQKLTDAAIKKIEAALADKEAELMQF
this seq match to pattern ['VTTERY', 'STVKEF'] in
>2VJX_1|Chains A,B|BETA-MANNOSIDASE|BACTEROIDES THETAIOTAOMICRON (226186)
QGNDTSEVMLLDTGWEFSQSGTEKWMPATVPGTVHQDLISHELLPNPFYGMNEKKIQWVENEDWEYRTSFIVSEEQLNRDGIQLIFEGLDTYADVYLNGSLLLKADNMFVGYTLPVKSVLRKGENHLYIYFHSPIRQT
LPQYASNGFNYPADNDHHEKHLSVFSRKAPYSYGWDWGIRMVTSGVWRPVTLRFYDIATISDYYVRQLSLTDENARLSNELIVNQIVPQKIPAEVRVNVSLNGTTVTEVKQQVTLQPGINHITLPAEVTNPVRWMPNGW
GTPTLYDFSAQIACGDRIVAEQSHRIGLRTIRVVNEKDKDGESFYFEVNGIPMFAKGANYIPQDALLPNVTTERYQTLFRDMKEANMNMVRIWGGGTYENNLFYDLADENGILVWQDFMFACTPYPSDPTFLKRVEAEA
VYNIRRLRNHASLAMWCGNNEILEALKYWGFEKKFTPEVYQGLMHGYDKLFRELLPSTVKEFDSDRFYVHSSPYLANWGRPESWGTGDSHNWGVWYGKKPFESLDTDLPRFMSEFGFQSFPEMKTIAAFAAPEDYQIES
EVMNAHQKSSIGNSLIRTYMERDYIIPESFEDFVYVGLVLQGQGMRHGLEAHRRNRPYCMGTLYWQLNDSWPVVSWSSIDYYGNWKALHYQAKRAFAPVLINPIQQNDSLSVYLISDRLDTMEQMTLEMKVVDFDGKTL
GKKIQVHSLEVPANTSKCVYRAKLDGWLTPEDCRRSFLKLILKDKSGHQVAESVHFFRKTKDLQLPPTSVSYQMKQTDGKCELTLFSSMLAKDIFIETPLQGARYSDNFFDLLPGERKKVIITSPRIKKGEELPVNIKH
IRETYKEHHHHHH
this seq match to pattern ['VTTERY', 'STVKEF'] in
>2VJX_1|Chains A,B|BETA-MANNOSIDASE|BACTEROIDES THETAIOTAOMICRON (226186)
QGNDTSEVMLLDTGWEFSQSGTEKWMPATVPGTVHQDLISHELLPNPFYGMNEKKIQWVENEDWEYRTSFIVSEEQLNRDGIQLIFEGLDTYADVYLNGSLLLKADNMFVGYTLPVKSVLRKGENHLYIYFHSPIRQT
LPQYASNGFNYPADNDHHEKHLSVFSRKAPYSYGWDWGIRMVTSGVWRPVTLRFYDIATISDYYVRQLSLTDENARLSNELIVNQIVPQKIPAEVRVNVSLNGTTVTEVKQQVTLQPGINHITLPAEVTNPVRWMPNGW
GTPTLYDFSAQIACGDRIVAEQSHRIGLRTIRVVNEKDKDGESFYFEVNGIPMFAKGANYIPQDALLPNVTTERYQTLFRDMKEANMNMVRIWGGGTYENNLFYDLADENGILVWQDFMFACTPYPSDPTFLKRVEAEA
VYNIRRLRNHASLAMWCGNNEILEALKYWGFEKKFTPEVYQGLMHGYDKLFRELLPSTVKEFDSDRFYVHSSPYLANWGRPESWGTGDSHNWGVWYGKKPFESLDTDLPRFMSEFGFQSFPEMKTIAAFAAPEDYQIES
EVMNAHQKSSIGNSLIRTYMERDYIIPESFEDFVYVGLVLQGQGMRHGLEAHRRNRPYCMGTLYWQLNDSWPVVSWSSIDYYGNWKALHYQAKRAFAPVLINPIQQNDSLSVYLISDRLDTMEQMTLEMKVVDFDGKTL
GKKIQVHSLEVPANTSKCVYRAKLDGWLTPEDCRRSFLKLILKDKSGHQVAESVHFFRKTKDLQLPPTSVSYQMKQTDGKCELTLFSSMLAKDIFIETPLQGARYSDNFFDLLPGERKKVIITSPRIKKGEELPVNIKH
IRETYKEHHHHHH
this seq match to pattern ['VTTERY', 'STVKEF'] in
>2VL4_1|Chains A,B|BETA-MANNOSIDASE|BACTEROIDES THETAIOTAOMICRON (226186)
QGNDTSEVMLLDTGWEFSQSGTEKWMPATVPGTVHQDLISHELLPNPFYGMNEKKIQWVENEDWEYRTSFIVSEEQLNRDGIQLIFEGLDTYADVYLNGSLLLKADNMFVGYTLPVKSVLRKGENHLYIYFHSPIRQT
LPQYASNGFNYPADNDHHEKHLSVFSRKAPYSYGWDWGIRMVTSGVWRPVTLRFYDIATISDYYVRQLSLTDENARLSNELIVNQIVPQKIPAEVRVNVSLNGTTVTEVKQQVTLQPGINHITLPAEVTNPVRWMPNGW
GTPTLYDFSAQIACGDRIVAEQSHRIGLRTIRVVNEKDKDGESFYFEVNGIPMFAKGANYIPQDALLPNVTTERYQTLFRDMKEANMNMVRIWGGGTYENNLFYDLADENGILVWQDFMFACTPYPSDPTFLKRVEAEA
VYNIRRLRNHASLAMWCGNNEILEALKYWGFEKKFTPEVYQGLMHGYDKLFRELLPSTVKEFDSDRFYVHSSPYLANWGRPESWGTGDSHNWGVWYGKKPFESLDTDLPRFMSEFGFQSFPEMKTIAAFAAPEDYQIES
EVMNAHQKSSIGNSLIRTYMERDYIIPESFEDFVYVGLVLQGQGMRHGLEAHRRNRPYCMGTLYWQLNDSWPVVSWSSIDYYGNWKALHYQAKRAFAPVLINPIQQNDSLSVYLISDRLDTMEQMTLEMKVVDFDGKTL
GKKIQVHSLEVPANTSKCVYRAKLDGWLTPEDCRRSFLKLILKDKSGHQVAESVHFFRKTKDLQLPPTSVSYQMKQTDGKCELTLFSSMLAKDIFIETPLQGARYSDNFFDLLPGERKKVIITSPRIKKGEELPVNIKH
IRETYKEHHHHHH
this seq match to pattern ['VTTERY', 'STVKEF'] in
>2VL4_1|Chains A,B|BETA-MANNOSIDASE|BACTEROIDES THETAIOTAOMICRON (226186)
QGNDTSEVMLLDTGWEFSQSGTEKWMPATVPGTVHQDLISHELLPNPFYGMNEKKIQWVENEDWEYRTSFIVSEEQLNRDGIQLIFEGLDTYADVYLNGSLLLKADNMFVGYTLPVKSVLRKGENHLYIYFHSPIRQT
LPQYASNGFNYPADNDHHEKHLSVFSRKAPYSYGWDWGIRMVTSGVWRPVTLRFYDIATISDYYVRQLSLTDENARLSNELIVNQIVPQKIPAEVRVNVSLNGTTVTEVKQQVTLQPGINHITLPAEVTNPVRWMPNGW
GTPTLYDFSAQIACGDRIVAEQSHRIGLRTIRVVNEKDKDGESFYFEVNGIPMFAKGANYIPQDALLPNVTTERYQTLFRDMKEANMNMVRIWGGGTYENNLFYDLADENGILVWQDFMFACTPYPSDPTFLKRVEAEA
VYNIRRLRNHASLAMWCGNNEILEALKYWGFEKKFTPEVYQGLMHGYDKLFRELLPSTVKEFDSDRFYVHSSPYLANWGRPESWGTGDSHNWGVWYGKKPFESLDTDLPRFMSEFGFQSFPEMKTIAAFAAPEDYQIES
EVMNAHQKSSIGNSLIRTYMERDYIIPESFEDFVYVGLVLQGQGMRHGLEAHRRNRPYCMGTLYWQLNDSWPVVSWSSIDYYGNWKALHYQAKRAFAPVLINPIQQNDSLSVYLISDRLDTMEQMTLEMKVVDFDGKTL
GKKIQVHSLEVPANTSKCVYRAKLDGWLTPEDCRRSFLKLILKDKSGHQVAESVHFFRKTKDLQLPPTSVSYQMKQTDGKCELTLFSSMLAKDIFIETPLQGARYSDNFFDLLPGERKKVIITSPRIKKGEELPVNIKH
IRETYKEHHHHHH
this seq match to pattern ['VTTERY', 'STVKEF'] in
>2VMF_1|Chains A,B|BETA-MANNOSIDASE|BACTEROIDES THETAIOTAOMICRON (226186)
QGNDTSEVMLLDTGWEFSQSGTEKWMPATVPGTVHQDLISHELLPNPFYGMNEKKIQWVENEDWEYRTSFIVSEEQLNRDGIQLIFEGLDTYADVYLNGSLLLKADNMFVGYTLPVKSVLRKGENHLYIYFHSPIRQT
LPQYASNGFNYPADNDHHEKHLSVFSRKAPYSYGWDWGIRMVTSGVWRPVTLRFYDIATISDYYVRQLSLTDENARLSNELIVNQIVPQKIPAEVRVNVSLNGTTVTEVKQQVTLQPGINHITLPAEVTNPVRWMPNGW
GTPTLYDFSAQIACGDRIVAEQSHRIGLRTIRVVNEKDKDGESFYFEVNGIPMFAKGANYIPQDALLPNVTTERYQTLFRDMKEANMNMVRIWGGGTYENNLFYDLADENGILVWQDFMFACTPYPSDPTFLKRVEAEA
VYNIRRLRNHASLAMWCGNNEILEALKYWGFEKKFTPEVYQGLMHGYDKLFRELLPSTVKEFDSDRFYVHSSPYLANWGRPESWGTGDSHNWGVWYGKKPFESLDTDLPRFMSEFGFQSFPEMKTIAAFAAPEDYQIES
EVMNAHQKSSIGNSLIRTYMERDYIIPESFEDFVYVGLVLQGQGMRHGLEAHRRNRPYCMGTLYWQLNDSWPVVSWSSIDYYGNWKALHYQAKRAFAPVLINPIQQNDSLSVYLISDRLDTMEQMTLEMKVVDFDGKTL
GKKIQVHSLEVPANTSKCVYRAKLDGWLTPEDCRRSFLKLILKDKSGHQVAESVHFFRKTKDLQLPPTSVSYQMKQTDGKCELTLFSSMLAKDIFIETPLQGARYSDNFFDLLPGERKKVIITSPRIKKGEELPVNIKH
IRETYKEHHHHHH
this seq match to pattern ['VTTERY', 'STVKEF'] in
>2VMF_1|Chains A,B|BETA-MANNOSIDASE|BACTEROIDES THETAIOTAOMICRON (226186)
QGNDTSEVMLLDTGWEFSQSGTEKWMPATVPGTVHQDLISHELLPNPFYGMNEKKIQWVENEDWEYRTSFIVSEEQLNRDGIQLIFEGLDTYADVYLNGSLLLKADNMFVGYTLPVKSVLRKGENHLYIYFHSPIRQT
LPQYASNGFNYPADNDHHEKHLSVFSRKAPYSYGWDWGIRMVTSGVWRPVTLRFYDIATISDYYVRQLSLTDENARLSNELIVNQIVPQKIPAEVRVNVSLNGTTVTEVKQQVTLQPGINHITLPAEVTNPVRWMPNGW
GTPTLYDFSAQIACGDRIVAEQSHRIGLRTIRVVNEKDKDGESFYFEVNGIPMFAKGANYIPQDALLPNVTTERYQTLFRDMKEANMNMVRIWGGGTYENNLFYDLADENGILVWQDFMFACTPYPSDPTFLKRVEAEA
VYNIRRLRNHASLAMWCGNNEILEALKYWGFEKKFTPEVYQGLMHGYDKLFRELLPSTVKEFDSDRFYVHSSPYLANWGRPESWGTGDSHNWGVWYGKKPFESLDTDLPRFMSEFGFQSFPEMKTIAAFAAPEDYQIES
EVMNAHQKSSIGNSLIRTYMERDYIIPESFEDFVYVGLVLQGQGMRHGLEAHRRNRPYCMGTLYWQLNDSWPVVSWSSIDYYGNWKALHYQAKRAFAPVLINPIQQNDSLSVYLISDRLDTMEQMTLEMKVVDFDGKTL
GKKIQVHSLEVPANTSKCVYRAKLDGWLTPEDCRRSFLKLILKDKSGHQVAESVHFFRKTKDLQLPPTSVSYQMKQTDGKCELTLFSSMLAKDIFIETPLQGARYSDNFFDLLPGERKKVIITSPRIKKGEELPVNIKH
IRETYKEHHHHHH
this seq match to pattern ['VTTERY', 'STVKEF'] in
>2VO5_1|Chains A,B|BETA-MANNOSIDASE|BACTEROIDES THETAIOTAOMICRON (226186)
QGNDTSEVMLLDTGWEFSQSGTEKWMPATVPGTVHQDLISHELLPNPFYGMNEKKIQWVENEDWEYRTSFIVSEEQLNRDGIQLIFEGLDTYADVYLNGSLLLKADNMFVGYTLPVKSVLRKGENHLYIYFHSPIRQT
LPQYASNGFNYPADNDHHEKHLSVFSRKAPYSYGWDWGIRMVTSGVWRPVTLRFYDIATISDYYVRQLSLTDENARLSNELIVNQIVPQKIPAEVRVNVSLNGTTVTEVKQQVTLQPGINHITLPAEVTNPVRWMPNGW
GTPTLYDFSAQIACGDRIVAEQSHRIGLRTIRVVNEKDKDGESFYFEVNGIPMFAKGANYIPQDALLPNVTTERYQTLFRDMKEANMNMVRIWGGGTYENNLFYDLADENGILVWQDFMFACTPYPSDPTFLKRVEAEA
VYNIRRLRNHASLAMWCGNNEILEALKYWGFEKKFTPEVYQGLMHGYDKLFRELLPSTVKEFDSDRFYVHSSPYLANWGRPESWGTGDSHNWGVWYGKKPFESLDTDLPRFMSEFGFQSFPEMKTIAAFAAPEDYQIES
EVMNAHQKSSIGNSLIRTYMERDYIIPESFEDFVYVGLVLQGQGMRHGLEAHRRNRPYCMGTLYWQLNDSWPVVSWSSIDYYGNWKALHYQAKRAFAPVLINPIQQNDSLSVYLISDRLDTMEQMTLEMKVVDFDGKTL
GKKIQVHSLEVPANTSKCVYRAKLDGWLTPEDCRRSFLKLILKDKSGHQVAESVHFFRKTKDLQLPPTSVSYQMKQTDGKCELTLFSSMLAKDIFIETPLQGARYSDNFFDLLPGERKKVIITSPRIKKGEELPVNIKH
IRETYKEHHHHHH
this seq match to pattern ['VTTERY', 'STVKEF'] in
>2VO5_1|Chains A,B|BETA-MANNOSIDASE|BACTEROIDES THETAIOTAOMICRON (226186)
QGNDTSEVMLLDTGWEFSQSGTEKWMPATVPGTVHQDLISHELLPNPFYGMNEKKIQWVENEDWEYRTSFIVSEEQLNRDGIQLIFEGLDTYADVYLNGSLLLKADNMFVGYTLPVKSVLRKGENHLYIYFHSPIRQT
LPQYASNGFNYPADNDHHEKHLSVFSRKAPYSYGWDWGIRMVTSGVWRPVTLRFYDIATISDYYVRQLSLTDENARLSNELIVNQIVPQKIPAEVRVNVSLNGTTVTEVKQQVTLQPGINHITLPAEVTNPVRWMPNGW
GTPTLYDFSAQIACGDRIVAEQSHRIGLRTIRVVNEKDKDGESFYFEVNGIPMFAKGANYIPQDALLPNVTTERYQTLFRDMKEANMNMVRIWGGGTYENNLFYDLADENGILVWQDFMFACTPYPSDPTFLKRVEAEA
VYNIRRLRNHASLAMWCGNNEILEALKYWGFEKKFTPEVYQGLMHGYDKLFRELLPSTVKEFDSDRFYVHSSPYLANWGRPESWGTGDSHNWGVWYGKKPFESLDTDLPRFMSEFGFQSFPEMKTIAAFAAPEDYQIES
EVMNAHQKSSIGNSLIRTYMERDYIIPESFEDFVYVGLVLQGQGMRHGLEAHRRNRPYCMGTLYWQLNDSWPVVSWSSIDYYGNWKALHYQAKRAFAPVLINPIQQNDSLSVYLISDRLDTMEQMTLEMKVVDFDGKTL
GKKIQVHSLEVPANTSKCVYRAKLDGWLTPEDCRRSFLKLILKDKSGHQVAESVHFFRKTKDLQLPPTSVSYQMKQTDGKCELTLFSSMLAKDIFIETPLQGARYSDNFFDLLPGERKKVIITSPRIKKGEELPVNIKH
IRETYKEHHHHHH
this seq match to pattern ['VTTERY', 'STVKEF'] in
>2VOT_1|Chains A,B|BETA-MANNOSIDASE|BACTEROIDES THETAIOTAOMICRON (226186)
QGNDTSEVMLLDTGWEFSQSGTEKWMPATVPGTVHQDLISHELLPNPFYGMNEKKIQWVENEDWEYRTSFIVSEEQLNRDGIQLIFEGLDTYADVYLNGSLLLKADNMFVGYTLPVKSVLRKGENHLYIYFHSPIRQT
LPQYASNGFNYPADNDHHEKHLSVFSRKAPYSYGWDWGIRMVTSGVWRPVTLRFYDIATISDYYVRQLSLTDENARLSNELIVNQIVPQKIPAEVRVNVSLNGTTVTEVKQQVTLQPGINHITLPAEVTNPVRWMPNGW
GTPTLYDFSAQIACGDRIVAEQSHRIGLRTIRVVNEKDKDGESFYFEVNGIPMFAKGANYIPQDALLPNVTTERYQTLFRDMKEANMNMVRIWGGGTYENNLFYDLADENGILVWQDFMFACTPYPSDPTFLKRVEAEA
VYNIRRLRNHASLAMWCGNNEILEALKYWGFEKKFTPEVYQGLMHGYDKLFRELLPSTVKEFDSDRFYVHSSPYLANWGRPESWGTGDSHNWGVWYGKKPFESLDTDLPRFMSEFGFQSFPEMKTIAAFAAPEDYQIES
EVMNAHQKSSIGNSLIRTYMERDYIIPESFEDFVYVGLVLQGQGMRHGLEAHRRNRPYCMGTLYWQLNDSWPVVSWSSIDYYGNWKALHYQAKRAFAPVLINPIQQNDSLSVYLISDRLDTMEQMTLEMKVVDFDGKTL
GKKIQVHSLEVPANTSKCVYRAKLDGWLTPEDCRRSFLKLILKDKSGHQVAESVHFFRKTKDLQLPPTSVSYQMKQTDGKCELTLFSSMLAKDIFIETPLQGARYSDNFFDLLPGERKKVIITSPRIKKGEELPVNIKH
IRETYKEHHHHHH
this seq match to pattern ['VTTERY', 'STVKEF'] in
>2VOT_1|Chains A,B|BETA-MANNOSIDASE|BACTEROIDES THETAIOTAOMICRON (226186)
QGNDTSEVMLLDTGWEFSQSGTEKWMPATVPGTVHQDLISHELLPNPFYGMNEKKIQWVENEDWEYRTSFIVSEEQLNRDGIQLIFEGLDTYADVYLNGSLLLKADNMFVGYTLPVKSVLRKGENHLYIYFHSPIRQT
LPQYASNGFNYPADNDHHEKHLSVFSRKAPYSYGWDWGIRMVTSGVWRPVTLRFYDIATISDYYVRQLSLTDENARLSNELIVNQIVPQKIPAEVRVNVSLNGTTVTEVKQQVTLQPGINHITLPAEVTNPVRWMPNGW
GTPTLYDFSAQIACGDRIVAEQSHRIGLRTIRVVNEKDKDGESFYFEVNGIPMFAKGANYIPQDALLPNVTTERYQTLFRDMKEANMNMVRIWGGGTYENNLFYDLADENGILVWQDFMFACTPYPSDPTFLKRVEAEA
VYNIRRLRNHASLAMWCGNNEILEALKYWGFEKKFTPEVYQGLMHGYDKLFRELLPSTVKEFDSDRFYVHSSPYLANWGRPESWGTGDSHNWGVWYGKKPFESLDTDLPRFMSEFGFQSFPEMKTIAAFAAPEDYQIES
EVMNAHQKSSIGNSLIRTYMERDYIIPESFEDFVYVGLVLQGQGMRHGLEAHRRNRPYCMGTLYWQLNDSWPVVSWSSIDYYGNWKALHYQAKRAFAPVLINPIQQNDSLSVYLISDRLDTMEQMTLEMKVVDFDGKTL
GKKIQVHSLEVPANTSKCVYRAKLDGWLTPEDCRRSFLKLILKDKSGHQVAESVHFFRKTKDLQLPPTSVSYQMKQTDGKCELTLFSSMLAKDIFIETPLQGARYSDNFFDLLPGERKKVIITSPRIKKGEELPVNIKH
IRETYKEHHHHHH
this seq match to pattern ['VTTERY', 'STVKEF'] in

```
>2VQT_1|Chains A,B|BETA-MANNOSIDASE|BACTEROIDES THETAIOTAOMICRON (818)
QGNDTSEVMLLDTGWEFSQSGTEKWMPATVPGTVHQDLISHELLPNPFYGMNEKKIQWVENEDWEYRTSFIVSEEQLNRDGIQLIFEGLDTYADVYLNGSLLLKADNMFVGYTLPVKSVLRKGENHLYIYFHSPIRQT
LPQYASNGFNYPADNDHHEKHLSVFSRKAPYSYGWDWGIRMVTSGVWRPVTLRFYDIATISDYYVRQLSLTDENARLSNELIVNQIVPQKIPAEVRVNVSLNGTTVTEVKQQVTLQPGINHITLPAEVTNPVRWMPNGW
GTPTLYDFSAQIACGDRIVAEQSHRIGLRTIRVVNEKDKDGESFYFEVNGIPMFAKGANYIPQDALLPNVTTERYQTLFRDMKEANMNMVRIWGGGTYENNLFYDLADENGILVWQDFMFACTPYPSDPTFLKRVEAEA
VYNIRRLRNHASLAMWCGNNEILEALKYWGFEKKFTPEVYQGLMHGYDKLFRELLPSTVKEFDSDRFYVHSSPYLANWGRPESWGTGDSHNWGVWYGKKPFESLDTDLPRFMSEFGFQSFPEMKTIAAFAAPEDYQIES
EVMNAHQKSSIGNSLIRTYMERDYIIPESFEDFVYVGLVLQGQGMRHGLEAHRRNRPYCMGTLYWQLNDSWPVVSWSSIDYYGNWKALHYQAKRAFAPVLINPIQQNDSLSVYLISDRLDTMEQMTLEMKVVDFDGKTL
GKKIQVHSLEVPANTSKCVYRAKLDGWLTPEDCRRSFLKLILKDKSGHQVAESVHFFRKTKDLQLPPTSVSYQMKQTDGKCELTLFSSMLAKDIFIETPLQGARYSDNFFDLLPGERKKVIITSPRIKKGEELPVNIKH
IRETYKEHHHHHH
this seq match to pattern ['VTTERY', 'STVKEF'] in
>2VQT_1|Chains A,B|BETA-MANNOSIDASE|BACTEROIDES THETAIOTAOMICRON (818)
QGNDTSEVMLLDTGWEFSQSGTEKWMPATVPGTVHQDLISHELLPNPFYGMNEKKIQWVENEDWEYRTSFIVSEEQLNRDGIQLIFEGLDTYADVYLNGSLLLKADNMFVGYTLPVKSVLRKGENHLYIYFHSPIRQT
LPQYASNGFNYPADNDHHEKHLSVFSRKAPYSYGWDWGIRMVTSGVWRPVTLRFYDIATISDYYVRQLSLTDENARLSNELIVNQIVPQKIPAEVRVNVSLNGTTVTEVKQQVTLQPGINHITLPAEVTNPVRWMPNGW
GTPTLYDFSAQIACGDRIVAEQSHRIGLRTIRVVNEKDKDGESFYFEVNGIPMFAKGANYIPQDALLPNVTTERYQTLFRDMKEANMNMVRIWGGGTYENNLFYDLADENGILVWQDFMFACTPYPSDPTFLKRVEAEA
VYNIRRLRNHASLAMWCGNNEILEALKYWGFEKKFTPEVYQGLMHGYDKLFRELLPSTVKEFDSDRFYVHSSPYLANWGRPESWGTGDSHNWGVWYGKKPFESLDTDLPRFMSEFGFQSFPEMKTIAAFAAPEDYQIES
EVMNAHQKSSIGNSLIRTYMERDYIIPESFEDFVYVGLVLQGQGMRHGLEAHRRNRPYCMGTLYWQLNDSWPVVSWSSIDYYGNWKALHYQAKRAFAPVLINPIQQNDSLSVYLISDRLDTMEQMTLEMKVVDFDGKTL
GKKIQVHSLEVPANTSKCVYRAKLDGWLTPEDCRRSFLKLILKDKSGHQVAESVHFFRKTKDLQLPPTSVSYQMKQTDGKCELTLFSSMLAKDIFIETPLQGARYSDNFFDLLPGERKKVIITSPRIKKGEELPVNIKH
IRETYKEHHHHHH
this seq match to pattern ['VTTERY', 'STVKEF'] in
>2VQU_1|Chains A,B|BETA-MANNOSIDASE|BACTEROIDES THETAIOTAOMICRON (818)
QGNDTSEVMLLDTGWEFSQSGTEKWMPATVPGTVHQDLISHELLPNPFYGMNEKKIQWVENEDWEYRTSFIVSEEQLNRDGIQLIFEGLDTYADVYLNGSLLLKADNMFVGYTLPVKSVLRKGENHLYIYFHSPIRQT
LPQYASNGFNYPADNDHHEKHLSVFSRKAPYSYGWDWGIRMVTSGVWRPVTLRFYDIATISDYYVRQLSLTDENARLSNELIVNQIVPQKIPAEVRVNVSLNGTTVTEVKQQVTLQPGINHITLPAEVTNPVRWMPNGW
GTPTLYDFSAQIACGDRIVAEQSHRIGLRTIRVVNEKDKDGESFYFEVNGIPMFAKGANYIPQDALLPNVTTERYQTLFRDMKEANMNMVRIWGGGTYENNLFYDLADENGILVWQDFMFACTPYPSDPTFLKRVEAEA
VYNIRRLRNHASLAMWCGNNEILEALKYWGFEKKFTPEVYQGLMHGYDKLFRELLPSTVKEFDSDRFYVHSSPYLANWGRPESWGTGDSHNWGVWYGKKPFESLDTDLPRFMSEFGFQSFPEMKTIAAFAAPEDYQIES
EVMNAHQKSSIGNSLIRTYMERDYIIPESFEDFVYVGLVLQGQGMRHGLEAHRRNRPYCMGTLYAQLNDSWPVVSWSSIDYYGNWKALHYQAKRAFAPVLINPIQQNDSLSVYLISDRLDTMEQMTLEMKVVDFDGKTL
GKKIQVHSLEVPANTSKCVYRAKLDGWLTPEDCRRSFLKLILKDKSGHQVAESVHFFRKTKDLQLPPTSVSYQMKQTDGKCELTLFSSMLAKDIFIETPLQGARYSDNFFDLLPGERKKVIITSPRIKKGEELPVNIKH
IRETYKEHHHHHH
this seq match to pattern ['VTTERY', 'STVKEF'] in
>2VQU_1|Chains A,B|BETA-MANNOSIDASE|BACTEROIDES THETAIOTAOMICRON (818)
QGNDTSEVMLLDTGWEFSQSGTEKWMPATVPGTVHQDLISHELLPNPFYGMNEKKIQWVENEDWEYRTSFIVSEEQLNRDGIQLIFEGLDTYADVYLNGSLLLKADNMFVGYTLPVKSVLRKGENHLYIYFHSPIRQT
LPQYASNGFNYPADNDHHEKHLSVFSRKAPYSYGWDWGIRMVTSGVWRPVTLRFYDIATISDYYVRQLSLTDENARLSNELIVNQIVPQKIPAEVRVNVSLNGTTVTEVKQQVTLQPGINHITLPAEVTNPVRWMPNGW
GTPTLYDFSAQIACGDRIVAEQSHRIGLRTIRVVNEKDKDGESFYFEVNGIPMFAKGANYIPQDALLPNVTTERYQTLFRDMKEANMNMVRIWGGGTYENNLFYDLADENGILVWQDFMFACTPYPSDPTFLKRVEAEA
VYNIRRLRNHASLAMWCGNNEILEALKYWGFEKKFTPEVYQGLMHGYDKLFRELLPSTVKEFDSDRFYVHSSPYLANWGRPESWGTGDSHNWGVWYGKKPFESLDTDLPRFMSEFGFQSFPEMKTIAAFAAPEDYQIES
EVMNAHQKSSIGNSLIRTYMERDYIIPESFEDFVYVGLVLQGQGMRHGLEAHRRNRPYCMGTLYAQLNDSWPVVSWSSIDYYGNWKALHYQAKRAFAPVLINPIQQNDSLSVYLISDRLDTMEQMTLEMKVVDFDGKTL
GKKIQVHSLEVPANTSKCVYRAKLDGWLTPEDCRRSFLKLILKDKSGHQVAESVHFFRKTKDLQLPPTSVSYQMKQTDGKCELTLFSSMLAKDIFIETPLQGARYSDNFFDLLPGERKKVIITSPRIKKGEELPVNIKH
IRETYKEHHHHHH
this seq match to pattern ['VTTERY', 'STVKEF'] in
>2VR4_1|Chains A,B|BETA-MANNOSIDASE|BACTEROIDES THETAIOTAOMICRON (226186)
QGNDTSEVMLLDTGWEFSQSGTEKWMPATVPGTVHQDLISHELLPNPFYGMNEKKIQWVENEDWEYRTSFIVSEEQLNRDGIQLIFEGLDTYADVYLNGSLLLKADNMFVGYTLPVKSVLRKGENHLYIYFHSPIRQT
LPQYASNGFNYPADNDHHEKHLSVFSRKAPYSYGWDWGIRMVTSGVWRPVTLRFYDIATISDYYVRQLSLTDENARLSNELIVNQIVPQKIPAEVRVNVSLNGTTVTEVKQQVTLQPGINHITLPAEVTNPVRWMPNGW
GTPTLYDFSAQIACGDRIVAEQSHRIGLRTIRVVNEKDKDGESFYFEVNGIPMFAKGANYIPQDALLPNVTTERYQTLFRDMKEANMNMVRIWGGGTYENNLFYDLADENGILVWQDFMFACTPYPSDPTFLKRVEAEA
VYNIRRLRNHASLAMWCGNNEILEALKYWGFEKKFTPEVYQGLMHGYDKLFRELLPSTVKEFDSDRFYVHSSPYLANWGRPESWGTGDSHNWGVWYGKKPFESLDTDLPRFMSEFGFQSFPEMKTIAAFAAPEDYQIES
EVMNAHQKSSIGNSLIRTYMERDYIIPESFEDFVYVGLVLQGQGMRHGLEAHRRNRPYCMGTLYWQLNDSWPVVSWSSIDYYGNWKALHYQAKRAFAPVLINPIQQNDSLSVYLISDRLDTMEQMTLEMKVVDFDGKTL
GKKIQVHSLEVPANTSKCVYRAKLDGWLTPEDCRRSFLKLILKDKSGHQVAESVHFFRKTKDLQLPPTSVSYQMKQTDGKCELTLFSSMLAKDIFIETPLQGARYSDNFFDLLPGERKKVIITSPRIKKGEELPVNIKH
IRETYKEHHHHHH
this seq match to pattern ['VTTERY', 'STVKEF'] in
>2VR4_1|Chains A,B|BETA-MANNOSIDASE|BACTEROIDES THETAIOTAOMICRON (226186)
QGNDTSEVMLLDTGWEFSQSGTEKWMPATVPGTVHQDLISHELLPNPFYGMNEKKIQWVENEDWEYRTSFIVSEEQLNRDGIQLIFEGLDTYADVYLNGSLLLKADNMFVGYTLPVKSVLRKGENHLYIYFHSPIRQT
LPQYASNGFNYPADNDHHEKHLSVFSRKAPYSYGWDWGIRMVTSGVWRPVTLRFYDIATISDYYVRQLSLTDENARLSNELIVNQIVPQKIPAEVRVNVSLNGTTVTEVKQQVTLQPGINHITLPAEVTNPVRWMPNGW
GTPTLYDFSAQIACGDRIVAEQSHRIGLRTIRVVNEKDKDGESFYFEVNGIPMFAKGANYIPQDALLPNVTTERYQTLFRDMKEANMNMVRIWGGGTYENNLFYDLADENGILVWQDFMFACTPYPSDPTFLKRVEAEA
VYNIRRLRNHASLAMWCGNNEILEALKYWGFEKKFTPEVYQGLMHGYDKLFRELLPSTVKEFDSDRFYVHSSPYLANWGRPESWGTGDSHNWGVWYGKKPFESLDTDLPRFMSEFGFQSFPEMKTIAAFAAPEDYQIES
EVMNAHQKSSIGNSLIRTYMERDYIIPESFEDFVYVGLVLQGQGMRHGLEAHRRNRPYCMGTLYWQLNDSWPVVSWSSIDYYGNWKALHYQAKRAFAPVLINPIQQNDSLSVYLISDRLDTMEQMTLEMKVVDFDGKTL
GKKIQVHSLEVPANTSKCVYRAKLDGWLTPEDCRRSFLKLILKDKSGHQVAESVHFFRKTKDLQLPPTSVSYQMKQTDGKCELTLFSSMLAKDIFIETPLQGARYSDNFFDLLPGERKKVIITSPRIKKGEELPVNIKH
IRETYKEHHHHHH
this seq match to pattern ['STVEER'] in
>1AA0_1|Chain A|FIBRITIN|Enterobacteria phage T4 (10665)
VSGLNNAVQNLQVEIGNNSAGIKGQVVALNTLVNGTNPNGSTVEERGLTNSIKANETNIASVTQEVNTAKGNISSLQGDVQALQEAGYIPEAPRDGQAYVRKDGEWVLLSTFL
this seq match to pattern ['STVRRC'] in
>1FO8_1|Chain A|ALPHA-1,3-MANNOSYL-GLYCOPROTEIN BETA-1,2-N-ACETYLGLUCOSAMINYLTRANSFERASE|Oryctolagus cuniculus (9986)
LAVIPILVIACDRSTVRRCLDKLLHYRPSAELFPIIVSQDCGHEETAQVIASYGSAVTHIRQPDLSNIAVQPDHRKFQGYYKIARHYRWALGQIFHNFNYPAAVVVEDDLEVAPDFFEYFQATYPLLKADPSLWCVSA
WNDNGKEQMVDSSKPELLYRTDFFPGLGWLLLAELWAELEPKWPKAFWDDWMRRPEQRKGRACVRPEISRTMTFGRKGVSHGQFFDQHLKFIKLNQQFVPFTQLDLSYLQQEAYDRDFLARVYGAPQLQVEKVRTNDRK
ELGEVRVQYTGRDSFKAFAKALGVMDDLKSGVPRAGYRGIVTFLFRGRRVHLAPPQTWDGYDPSWT
this seq match to pattern ['STVRDV'] in
>1JKM_1|Chains A,B|BREFELDIN A ESTERASE|Bacillus subtilis (1423)
YTPPGRLGDESSGPRTDPRFSPAMVEALATFGLDAVAAAPPVSASDDLPTVLAAVGASHDGFQAVYDSIALDLPTDRDDVETSTETILGVDGNEITLHVFRPAGVEGVLPGLVYTHGGGMTILTTDNRVHRRWCTDLA
AAGSVVVMVDFRNAWTAEGHHPFPSGVEDCLAAVLWVDEHRESLGLSGVVVQGESGGGNLAIATTLLAKRRGRLDAIDGVYASIPYISGGYAWDHERRLTELPSLVENDGYFIENGGMALLVRAYDPTGEHAEDPIAWP
YFASEDELRGLPPFVVAVNELDPLRDEGIAFARRLARAGVDVAARVNIGLVHGADVIFRHWLPAALESTVRDVAGFAADRARLR
this seq match to pattern ['STVDKP'] in
>1JW0_2|Chain B|cephalosporin acylase beta chain|Brevundimonas diminuta (293)
SNSWAVAPGKTANGNALLLQNPHLSWTTDYFTYYEAHLVTPDFEIYGATQIGLPVIRFAFNQRMGITNTVNGMVGATNYRLTLQDGGYLYDGQVRPFERRQASYRLRQADGSTVDKPLEIRSSVHGPVFERADGTAVA
VRVAGLDRPGMLEQYFDMITAHSFDDYEAAMARMQVPTFNIVYADREGTINYSFNGVAPKRAEGDIAFWQGNVPGDSSRYLWTETHPLDDLPRVTNPPGGFVQNSNDPPWTPTWPVTYCPANHPSYLAPQTPHSLRAQQ
SVRLMSENDDLTLERFMALQFSHRAVMADRTLPDLIPAALIDPDPEVQAAARLLAAWDRDFTSDSRAALLFEEWARLFAGQNFAGQAAFATPWSLDKPVSTPYGVRDPKAAVDQLRTAIANTKRKYGAIDRPFGDASRM
ILNDVNVPGAAGYGNLGSFRVFTWSDPDENGIRTPVHGETWVAMIEFSTPVRAYGLMSGRQPGTTHYSDQIERVSRADFRELLLRREQVEAAVQERTPFNF
this seq match to pattern ['STVEKT'] in
>1MCD_1|Chains A,B|Immunoglobulin lambda-1 light chain|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern ['STVEKT'] in
>1MCI_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern ['STVEKT'] in
>1MCK_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
this seq match to pattern ['STVEKT'] in
>1MCN_1|Chains A,B|IMMUNOGLOBULIN LAMBDA DIMER MCG (LIGHT CHAIN)|Homo sapiens (9606)
PSALTQPPSASGSLGQSVTISCTGTSSDVGGYNYVSWYQQHAGKAPKVIIYEVNKRPSGVPDRFSGSKSGNTASLTVSGLQAEDEADYYCSSYEGSDNFVFGTGTKVTVLGQPKANPTVTLFPPSSEELQANKATLVC
LISDFYPGAVTVAWKADGSPVKAGVETTKPSKQSNNKYAASSYLSLTPEQWKSHRSYSCQVTHEGSTVEKTVAPTECS
```

In [292…
```python
import regex as re

pattern1="[SV]T[VT][DERK]{2}[^IL]"
pattern2="[FILV]Q...[^RK]G...[RK]..[FILVWY]"
file1 = open("Q4.fasta")
l=list()
for i in file1:
    l.append(i.rstrip())
```

```python
for i in range(len(l)):
    if(re.findall(pattern2,l[i])):
        print("this seq match to pattern ",re.findall(pattern2,l[i]),"in","\n",l[i-1],"\n",l[i])
```

```
this seq match to pattern  ['LQIEKEGAIHREEL'] in
 >4FXG_2|Chains B,E|Complement C4-A alpha chain|Homo sapiens (9606)
 NVNFQKAINEKLGQYASPTAKRCCQDGVTRLPMMRSCEQRAARVQQPDCREPFLSCCQFAESLRKKSRDKGQAGLQRALEILQEEDLIDEDDIPVRSFFPENWLWRVETVDRFQILTLWLPDSLTTWEIHGLSLSKTK
GLCVATPVQLRVFREFHLHLRLPMSVRRFEQLELRPVLYNYLDKNLTVSVHVSPVEGLCLAGGGGLAQQVLVPAGSARPVAFSVVPTAAAAVSLKVVARGSFEFPVGDAVSKVLQIEKEGAIHREELVYELNPLDHRGR
TLEIPGNSDPNMIPDGDFNSYVRVTASDPLDTLGSEGALSPGGVASLLRLPRGCGEQTMIYLAPTLAASRYLDKTEQWSTLPPETKDHAVDLIQKGYMRIQQFRKADGSYAAWLSRDSSTWLTAFVLKVLSLAQEQVGG
SPEKLQETSNWLLSQQQADGSFQDPCPVLDRSMQGGLVGNDETVALTAFVTIALHHGLAVFQDEGAEPLKQRVEASISKANSFLGEKASAGLLGAHAAAITAYALSLTKAPVDLLGVAHNNLMAMAQETGDNLYWGSVT
GSQSNAVSPTPAPRNPSDPMPQAPALWIETTAYALLHLLLHEGKAEMADQASAWLTRQGSFQGGFRSTQDTVIALDALSAYWIASHTTEERGLNVTLSSTGRNGFKSHALQLNNRQIRGLEEELQFSLGSKINVKVGGN
SKGTLKVLRTYNVLDMKNTTCQDLQIEVTVKGHVEYTMEANEDYEDYEYDELPAKDDPDAPLQPVTPLQLFEG
this seq match to pattern  ['LQIEKEGAIHREEL'] in
 >4FXG_2|Chains B,E|Complement C4-A alpha chain|Homo sapiens (9606)
 NVNFQKAINEKLGQYASPTAKRCCQDGVTRLPMMRSCEQRAARVQQPDCREPFLSCCQFAESLRKKSRDKGQAGLQRALEILQEEDLIDEDDIPVRSFFPENWLWRVETVDRFQILTLWLPDSLTTWEIHGLSLSKTK
GLCVATPVQLRVFREFHLHLRLPMSVRRFEQLELRPVLYNYLDKNLTVSVHVSPVEGLCLAGGGGLAQQVLVPAGSARPVAFSVVPTAAAAVSLKVVARGSFEFPVGDAVSKVLQIEKEGAIHREELVYELNPLDHRGR
TLEIPGNSDPNMIPDGDFNSYVRVTASDPLDTLGSEGALSPGGVASLLRLPRGCGEQTMIYLAPTLAASRYLDKTEQWSTLPPETKDHAVDLIQKGYMRIQQFRKADGSYAAWLSRDSSTWLTAFVLKVLSLAQEQVGG
SPEKLQETSNWLLSQQQADGSFQDPCPVLDRSMQGGLVGNDETVALTAFVTIALHHGLAVFQDEGAEPLKQRVEASISKANSFLGEKASAGLLGAHAAAITAYALSLTKAPVDLLGVAHNNLMAMAQETGDNLYWGSVT
GSQSNAVSPTPAPRNPSDPMPQAPALWIETTAYALLHLLLHEGKAEMADQASAWLTRQGSFQGGFRSTQDTVIALDALSAYWIASHTTEERGLNVTLSSTGRNGFKSHALQLNNRQIRGLEEELQFSLGSKINVKVGGN
SKGTLKVLRTYNVLDMKNTTCQDLQIEVTVKGHVEYTMEANEDYEDYEYDELPAKDDPDAPLQPVTPLQLFEG
this seq match to pattern  ['LQIEKEGAIHREEL'] in
 >4FXK_2|Chain B|Complement C4-A Alpha chain|Homo sapiens (9606)
 NVNFQKAINEKLGQYASPTAKRCCQDGVTRLPMMRSCEQRAARVQQPDCREPFLSCCQFAESLRKKSRDKGQAGLQRALEILQEEDLIDEDDIPVRSFFPENWLWRVETVDRFQILTLWLPDSLTTWEIHGLSLSKTK
GLCVATPVQLRVFREFHLHLRLPMSVRRFEQLELRPVLYNYLDKNLTVSVHVSPVEGLCLAGGGGLAQQVLVPAGSARPVAFSVVPTAAAAVSLKVVARGSFEFPVGDAVSKVLQIEKEGAIHREELVYELNPLDHRGR
TLEIPGNSDPNMIPDGDFNSYVRVTASDPLDTLGSEGALSPGGVASLLRLPRGCGEQTMIYLAPTLAASRYLDKTEQWSTLPPETKDHAVDLIQKGYMRIQQFRKADGSYAAWLSRDSSTWLTAFVLKVLSLAQEQVGG
SPEKLQETSNWLLSQQQADGSFQDPCPVLDRSMQGGLVGNDETVALTAFVTIALHHGLAVFQDEGAEPLKQRVEASISKANSFLGEKASAGLLGAHAAAITAYALSLTKAPVDLLGVAHNNLMAMAQETGDNLYWGSVT
GSQSNAVSPTPAPRNPSDPMPQAPALWIETTAYALLHLLLHEGKAEMADQASAWLTRQGSFQGGFRSTQDTVIALDALSAYWIASHTTEERGLNVTLSSTGRNGFKSHALQLNNRQIRGLEEELQFSLGSKINVKVGGN
SKGTLKVLRTYNVLDMKNTTCQDLQIEVTVKGHVEYTMEANEDYEDYEYDELPAKDDPDAPLQPVTPLQLFEG
this seq match to pattern  ['LQIEKEGAIHREEL'] in
 >4XAM_2|Chains C,E|Complement C4-A|Homo sapiens (9606)
 ALEILQEEDLIDEDDIPVRSFFPENWLWRVETVDRFQILTLWLPDSLTTWEIHGLSLSKTKGLCVATPVQLRVFREFHLHLRLPMSVRRFEQLELRPVLYNYLDKNLTVSVHVSPVEGLCLAGGGGLAQQVLVPAGSA
RPVAFSVVPTAAAAVSLKVVARGSFEFPVGDAVSKVLQIEKEGAIHREELVYELNPLDHRGRTLEIPGNSDPNMIPDGDFNSYVRVTASDPLDTLGSEGALSPGGVASLLRLPRGCGEQTMIYLAPTLAASRYLDKTEQ
WSTLPPETKDHAVDLIQKGYMRIQQFRKADGSYAAWLSRDSSTWLTAFVLKVLSLAQEQVGGSPEKLQETSNWLLSQQQADGSFQDPCPVLDRSMQGGLVGNDETVALTAFVTIALHHGLAVFQDEGAEPLKQRVEASI
SKANSFLGEKASAGLLGAHAAAITAYALSLTKAPVDLLGVAHNNLMAMAQETGDNLYWGSVTGSQSNAVSPTPAPRNPSDPMPQAPALWIETTAYALLHLLLHEGKAEMADQASAWLTRQGSFQGGFRSTQDTVIALDA
LSAYWIASHTTEERGLNVTLSSTGRNGFKSHALQLNNRQIRGLEEELQFSLGSKINVKVGGNSKGTLKVLRTYNVLDMKNTTCQDLQIEVTVKGHVEYTMEANEDYEDYEYDELPAKDDPDAPLQPVTPLQLFEG
this seq match to pattern  ['LQIEKEGAIHREEL'] in
 >4XAM_2|Chains C,E|Complement C4-A|Homo sapiens (9606)
 ALEILQEEDLIDEDDIPVRSFFPENWLWRVETVDRFQILTLWLPDSLTTWEIHGLSLSKTKGLCVATPVQLRVFREFHLHLRLPMSVRRFEQLELRPVLYNYLDKNLTVSVHVSPVEGLCLAGGGGLAQQVLVPAGSA
RPVAFSVVPTAAAAVSLKVVARGSFEFPVGDAVSKVLQIEKEGAIHREELVYELNPLDHRGRTLEIPGNSDPNMIPDGDFNSYVRVTASDPLDTLGSEGALSPGGVASLLRLPRGCGEQTMIYLAPTLAASRYLDKTEQ
WSTLPPETKDHAVDLIQKGYMRIQQFRKADGSYAAWLSRDSSTWLTAFVLKVLSLAQEQVGGSPEKLQETSNWLLSQQQADGSFQDPCPVLDRSMQGGLVGNDETVALTAFVTIALHHGLAVFQDEGAEPLKQRVEASI
SKANSFLGEKASAGLLGAHAAAITAYALSLTKAPVDLLGVAHNNLMAMAQETGDNLYWGSVTGSQSNAVSPTPAPRNPSDPMPQAPALWIETTAYALLHLLLHEGKAEMADQASAWLTRQGSFQGGFRSTQDTVIALDA
LSAYWIASHTTEERGLNVTLSSTGRNGFKSHALQLNNRQIRGLEEELQFSLGSKINVKVGGNSKGTLKVLRTYNVLDMKNTTCQDLQIEVTVKGHVEYTMEANEDYEDYEYDELPAKDDPDAPLQPVTPLQLFEG
this seq match to pattern  ['LQIEKEGAIHREEL'] in
 >5JPM_2|Chains B,E|Complement C4-A|Homo sapiens (9606)
 NVNFQKAINEKLGQYASPTAKRCCQDGVTRLPMMRSCEQRAARVQQPDCREPFLSCCQFAESLRKKSRDKGQAGLQRALEILQEEDLIDEDDIPVRSFFPENWLWRVETVDRFQILTLWLPDSLTTWEIHGLSLSKTK
GLCVATPVQLRVFREFHLHLRLPMSVRRFEQLELRPVLYNYLDKNLTVSVHVSPVEGLCLAGGGGLAQQVLVPAGSARPVAFSVVPTAAAAVSLKVVARGSFEFPVGDAVSKVLQIEKEGAIHREELVYELNPLDHRGR
TLEIPGNSDPNMIPDGDFNSYVRVTASDPLDTLGSEGALSPGGVASLLRLPRGCGEQTMIYLAPTLAASRYLDKTEQWSTLPPETKDHAVDLIQKGYMRIQQFRKADGSYAAWLSRDSSTWLTAFVLKVLSLAQEQVGG
SPEKLQETSNWLLSQQQADGSFQDPCPVLDRSMQGGLVGNDETVALTAFVTIALHHGLAVFQDEGAEPLKQRVEASISKANSFLGEKASAGLLGAHAAAITAYALSLTKAPVDLLGVAHNNLMAMAQETGDNLYWGSVT
GSQSNAVSPTPAPRNPSDPMPQAPALWIETTAYALLHLLLHEGKAEMADQASAWLTRQGSFQGGFRSTQDTVIALDALSAYWIASHTTEERGLNVTLSSTGRNGFKSHALQLNNRQIRGLEEELQFSLGSKINVKVGGN
SKGTLKVLRTYNVLDMKNTTCQDLQIEVTVKGHVEYTMEANEDYEDYEYDELPAKDDPDAPLQPVTPLQLFEG
this seq match to pattern  ['LQIEKEGAIHREEL'] in
 >5JPM_2|Chains B,E|Complement C4-A|Homo sapiens (9606)
 NVNFQKAINEKLGQYASPTAKRCCQDGVTRLPMMRSCEQRAARVQQPDCREPFLSCCQFAESLRKKSRDKGQAGLQRALEILQEEDLIDEDDIPVRSFFPENWLWRVETVDRFQILTLWLPDSLTTWEIHGLSLSKTK
GLCVATPVQLRVFREFHLHLRLPMSVRRFEQLELRPVLYNYLDKNLTVSVHVSPVEGLCLAGGGGLAQQVLVPAGSARPVAFSVVPTAAAAVSLKVVARGSFEFPVGDAVSKVLQIEKEGAIHREELVYELNPLDHRGR
TLEIPGNSDPNMIPDGDFNSYVRVTASDPLDTLGSEGALSPGGVASLLRLPRGCGEQTMIYLAPTLAASRYLDKTEQWSTLPPETKDHAVDLIQKGYMRIQQFRKADGSYAAWLSRDSSTWLTAFVLKVLSLAQEQVGG
SPEKLQETSNWLLSQQQADGSFQDPCPVLDRSMQGGLVGNDETVALTAFVTIALHHGLAVFQDEGAEPLKQRVEASISKANSFLGEKASAGLLGAHAAAITAYALSLTKAPVDLLGVAHNNLMAMAQETGDNLYWGSVT
GSQSNAVSPTPAPRNPSDPMPQAPALWIETTAYALLHLLLHEGKAEMADQASAWLTRQGSFQGGFRSTQDTVIALDALSAYWIASHTTEERGLNVTLSSTGRNGFKSHALQLNNRQIRGLEEELQFSLGSKINVKVGGN
SKGTLKVLRTYNVLDMKNTTCQDLQIEVTVKGHVEYTMEANEDYEDYEYDELPAKDDPDAPLQPVTPLQLFEG
this seq match to pattern  ['LQIEKEGAIHREEL'] in
 >5JPN_2|Chain B|Complement C4-A|Homo sapiens (9606)
 NVNFQKAINEKLGQYASPTAKRCCQDGVTRLPMMRSCEQRAARVQQPDCREPFLSCCQFAESLRKKSRDKGQAGLQRALEILQEEDLIDEDDIPVRSFFPENWLWRVETVDRFQILTLWLPDSLTTWEIHGLSLSKTK
GLCVATPVQLRVFREFHLHLRLPMSVRRFEQLELRPVLYNYLDKNLTVSVHVSPVEGLCLAGGGGLAQQVLVPAGSARPVAFSVVPTAAAAVSLKVVARGSFEFPVGDAVSKVLQIEKEGAIHREELVYELNPLDHRGR
TLEIPGNSDPNMIPDGDFNSYVRVTASDPLDTLGSEGALSPGGVASLLRLPRGCGEQTMIYLAPTLAASRYLDKTEQWSTLPPETKDHAVDLIQKGYMRIQQFRKADGSYAAWLSRDSSTWLTAFVLKVLSLAQEQVGG
SPEKLQETSNWLLSQQQADGSFQDPCPVLDRSMQGGLVGNDETVALTAFVTIALHHGLAVFQDEGAEPLKQRVEASISKANSFLGEKASAGLLGAHAAAITAYALSLTKAPVDLLGVAHNNLMAMAQETGDNLYWGSVT
GSQSNAVSPTPAPRNPSDPMPQAPALWIETTAYALLHLLLHEGKAEMADQASAWLTRQGSFQGGFRSTQDTVIALDALSAYWIASHTTEERGLNVTLSSTGRNGFKSHALQLNNRQIRGLEEELQFSLGSKINVKVGGN
SKGTLKVLRTYNVLDMKNTTCQDLQIEVTVKGHVEYTMEANEDYEDYEYDELPAKDDPDAPLQPVTPLQLFEG
this seq match to pattern  ['LQIEKEGAIHREEL'] in
 >5JTW_2|Chains B,E|Complement C4-A|Homo sapiens (9606)
 ALEILQEEDLIDEDDIPVRSFFPENWLWRVETVDRFQILTLWLPDSLTTWEIHGLSLSKTKGLCVATPVQLRVFREFHLHLRLPMSVRRFEQLELRPVLYNYLDKNLTVSVHVSPVEGLCLAGGGGLAQQVLVPAGSA
RPVAFSVVPTAAAAVSLKVVARGSFEFPVGDAVSKVLQIEKEGAIHREELVYELNPLDHRGRTLEIPGNSDPNMIPDGDFNSYVRVTASDPLDTLGSEGALSPGGVASLLRLPRGCGEQTMIYLAPTLAASRYLDKTEQ
WSTLPPETKDHAVDLIQKGYMRIQQFRKADGSYAAWLSRDSSTWLTAFVLKVLSLAQEQVGGSPEKLQETSNWLLSQQQADGSFQDPCPVLDRSMQGGLVGNDETVALTAFVTIALHHGLAVFQDEGAEPLKQRVEASI
SKANSFLGEKASAGLLGAHAAAITAYALSLTKAPVDLLGVAHNNLMAMAQETGDNLYWGSVTGSQSNAVSPTPAPRNPSDPMPQAPALWIETTAYALLHLLLHEGKAEMADQASAWLTRQGSFQGGFRSTQDTVIALDA
LSAYWIASHTTEERGLNVTLSSTGRNGFKSHALQLNNRQIRGLEEELQFSLGSKINVKVGGNSKGTLKVLRTYNVLDMKNTTCQDLQIEVTVKGHVEYTMEANEDYEDYEYDELPAKDDPDAPLQPVTPLQLFEG
this seq match to pattern  ['LQIEKEGAIHREEL'] in
 >5JTW_2|Chains B,E|Complement C4-A|Homo sapiens (9606)
 ALEILQEEDLIDEDDIPVRSFFPENWLWRVETVDRFQILTLWLPDSLTTWEIHGLSLSKTKGLCVATPVQLRVFREFHLHLRLPMSVRRFEQLELRPVLYNYLDKNLTVSVHVSPVEGLCLAGGGGLAQQVLVPAGSA
RPVAFSVVPTAAAAVSLKVVARGSFEFPVGDAVSKVLQIEKEGAIHREELVYELNPLDHRGRTLEIPGNSDPNMIPDGDFNSYVRVTASDPLDTLGSEGALSPGGVASLLRLPRGCGEQTMIYLAPTLAASRYLDKTEQ
WSTLPPETKDHAVDLIQKGYMRIQQFRKADGSYAAWLSRDSSTWLTAFVLKVLSLAQEQVGGSPEKLQETSNWLLSQQQADGSFQDPCPVLDRSMQGGLVGNDETVALTAFVTIALHHGLAVFQDEGAEPLKQRVEASI
SKANSFLGEKASAGLLGAHAAAITAYALSLTKAPVDLLGVAHNNLMAMAQETGDNLYWGSVTGSQSNAVSPTPAPRNPSDPMPQAPALWIETTAYALLHLLLHEGKAEMADQASAWLTRQGSFQGGFRSTQDTVIALDA
LSAYWIASHTTEERGLNVTLSSTGRNGFKSHALQLNNRQIRGLEEELQFSLGSKINVKVGGNSKGTLKVLRTYNVLDMKNTTCQDLQIEVTVKGHVEYTMEANEDYEDYEYDELPAKDDPDAPLQPVTPLQLFEG
this seq match to pattern  ['LQIEKEGAIHREEL'] in
 >6YSQ_1|Chains A,B|Complement C4-B,Complement C4-B|Homo sapiens (9606)
 MRLLWGLIWASSFFTLSLQKPRLLLFSPSVVHLGVPLSVQLQDVPRGQVVKGSVFLRNPSRNNVPCSPKVDFTLSSERDFALLSLQVPLKDAKSCGLHQLLRGPEVQLVAHSPWLKDSLSRTTNIQGINLLFSSRR
GHLFLQTDQPIYNPGQRVRYRVFALDQKMRPSTDTITVMVENSHGLRVRKKEVVMPSSIFQDDFVIPDISEPGTWKISARFSDGLESNSSTQFEVKKYVLPNFEVKITPGKPYILTVPGHLDEMQLDIQARYIYGKPVQ
GVAYVRFGLLDEDGKKTFFRGLESQTKLVNGQSHISLSKAEFQDALEKLNMGITDLQGLRLYVAAAIIESPGGEMEEAELTSWYFVSSPFSLDLSKTKRHLVPGAPFLLQALVREMSGSPASGIPVKVSATVSSPGSVP
EVQDIQQNTDGSGQVSIPIIIPQTISELQLSVSAGSPHPAIARLTVAAPPSGGPGFLSIERPDSRPPRVGDTLNLNLRAVGSGATFSHYYYMILSRGQIVFMNREPKRTLTSVSVFVDHHLAPSFYFVAFYYHGDHPVA
NSLRVDVQAGACEGKLELSVDGAKQYRNGESVKLHLETDSLALVALGALDTALYAAGSKSHKPLNMGKVFEAMNSYDLGCGPGGGDSALQVFQAAGLAFSDGDQWTLSRKRLSCPKEKTTRKKRALEILQEEDLIDEDD
IPVRSFFPENWLWRVETVDRFQILTLWLPDSLTTWEIHGLSLSKTKGLCVATPVQLRVFREFHLHLRLPMSVRRFEQLELRPVLYNYLDKNLTVSVHVSPVEGLCLAGGGGLAQQVLVPAGSARPVAFSVVPTAAAAVS
LKVVARGSFEFPVGDAVSKVLQIEKEGAIHREELVYELNPLDHRGRTLEIPGNSDPNMIPDGDFNSYVRVTASDPLDTLGSEGALSPGGVASLLRLPRGCGEETMIYLAPTLAASRYLDKTEQWSTLPPETKDHAVDLI
QKGYMRIQQFRKADGSYAAWLSRDSSTWLTAFVLKVLSLAQEQVGGSPEKLQETSNWLLSQQQADGSFQDPCPVLDRSMQGGLVGNDETVALTAFVTIALHHGLAVFQDEGAEPLKQRVEASISKANSFLGEKASAGLL
GAHAAAITAYALSLTKAPVDLLGVAHNNLMAMAQETGDNLYWGSVTGSQSNAVSPTPAPRNPSDPMPQAPALWIETTAYALLHLLLHEGKAEMADQASAWLTRQGSFQGGFRSTQDTVIALDALSAYWIASHTTEERGL
NVTLSSTGRNGFKSHALQLNNRQIRGLEEELQFSLGSKINVKVGGNSKGTLKVLRTYNVLDMKNTTCQDLQIEVTVKGHVEYTMEANEDYEDYEYDELPAKDDPDAPLQPVTPLQLFEGRRNRNRRRREAPKVVEEQESRV
HYTVCIWRNGKVGLSGMAIADVTLLSGFHALRADLEKLTSLSDRYVVSHFETEGPHVLLYFDSVPTSRECVGFEAVQEVPVGLVQPASATLYDYYNPERRCSVFYGAPSKSRLLATLCSAEVCQCAEGKCPRQRRALERG
LQDEDGYRMKFACYYPRVEYGFQVKVLREDSRAAFRLFETKITQVLHFTKDVKAAANQMRNFLVRASCRLRLEPGKEYLIMGLDGATYDLEGHPQYLLDSNSWIEEMPSERLCRSTRQRAACAQLNDFLNEFGTQGCQV
this seq match to pattern  ['LQIEKEGAIHREEL'] in
 >6YSQ_1|Chains A,B|Complement C4-B,Complement C4-B|Homo sapiens (9606)
```

```
   MRLLWGLIWASSFFTLSLQKPRLLLFSPSVVHLGVPLSVGVQLQDVPRGQVVKGSVFLRNPSRNNVPCSPKVDFTLSSERDFALLSLQVPLKDAKSCGLHQLLRGPEVQLVAHSPWLKDSLSRTTNIQGINLLFSSRR
GHLFLQTDQPIYNPGQRVRYRVFALDQKMRPSTDITVMVENSHGLRVRKKEVYMPSSIFQDDFVIPDISEPGTWKISARFSDGLESNSSTQFEVKKYVLPNFEVKITPGKPYILTVPGHLDEMQLDIQARYIYGKPVQ
GVAYVRFGLLDEDGKKTFFRGLESQTKLVNGQSHISLSKAEFQDALEKLNMGITDLQGLRLYVAAAIIESPGGEMEEAELTSWYFVSSPFSLDLSKTKRHLVPGAPFLLQALVREMSGSPASGIPVKVSATVSSPGSVP
EVQDIQQNTDGSGQVSIPIIIPQTISELQLSVSAGSPHPAIARLTVAAPPSGGPGFLSIERPDSRPPRVGDTLNLNLRAVGSGATFSHYYYMILSRGQIVFMNREPKRTLTSVSVFVDHHLAPSFYFVAFYYHGDHPVA
NSLRVDVQAGACEGKLELSVDGAKQYRNGESVKLHLETDSLALVALGALDTALYAAGSKSHKPLNMGKVFEAMNSYDLGCGPGGGDSALQVFQAAGLAFSDGDQWTLSRKRLSCPKEKTTRKKRALEILQEEDLIDEDD
IPVRSFFPENWLWRVETVDRFQILTLWLPDSLTTWEIHGLSLSKTKGLCVATPVQLRVFREFHLHLRLPMSVRRFEQLELRPVLYNYLDKNLTVSVHVSPVEGLCLAGGGGLAQQVLVPAGSARPVAFSVVPTAAAAVS
LKVVARGSFEFPVGDAVSKVLQIEKEGAIHREELVYELNPLDHRGRTLEIPGNSDPNMIPDGDFNSYVRVTASDPLDTLGSEGALSPGGVASLLRLPRGCGEETMIYLAPTLAASRYLDKTEQWSTLPPETKDHAVDLI
QKGYMRIQQFRKADGSYAAWLSRDSSTWLTAFVLKVLSLAQEQVGGSPEKLQETSNWLLSQQQADGSFQDPCPVLDRSMQGGLVGNDETVALTAFVTIALHHGLAVFQDEGAEPLKQRVEASISKANSFLGEKASAGLL
GAHAAAITAYALSLTKAPVDLLGVAHNNLMAMAQETGDNLYWGSVTGSQSNAVSPTPAPRNPSDPMPQAPALWIETTAYALLHLLLHEGKAEMADQASAWLTRQGSFQGGFRSTQDTVIALDALSAYWIASHTTEERGL
NVTLSSTGRNGFKSHALQLNNRQIRGLEEELQFSLGSKINVKVGGNSKGTLKVLRTYNVLDMKNTTCQDLQIEVTVKGHVEYTMEANEDYEDYEYDELPAKDDPDAPLQPVTPLQLFEGRRNRRRREAPKVVEEQESRV
HYTVCIWRNGKVGLSGMAIADVTLLSGFHALRADLEKLTSLSDRYVSHFETEGPHVLLYFDSVPTSRECVGFEAVQEVPVGLVQPASATLYDYYNPERRCSVFYGAPSKSRLLATLCSAEVCQCAEGKCPRQRRALERG
LQDEDGYRMKFACYYPRVEYGFQVKVLREDSRAAFRLFETKITQVLHFTKDVKAAANQMRNFLVRASCRLRLEPGKEYLIMGLDGATYDLEGHPQYLLDSNSWIEEMPSERLCRSTRQRAACAQLNDFLNEFGTQGCQV
```

6. Identify the beta barrel membrane proteins with the following pattern:

[K,R,H,Q,F,E]-x-G-[I,V,L,F,A,C]-x-[I,V,L,F,M,Y,W]-x-[I,V,L,F,W] Use: http://www.bioinformatics.org/sms2/protein_pattern.html and http://prosite.expasy.org/scanprosite/ Hint: Modify the patterns according to the input format of the server.

after seraching databse i get all sequence that i ahve kept in file anme search.txt

below i write code to dipaly all matched data

In [353...
```python
import regex as re


pattern="[KRHQNST].G[IVLFMYWAC].[IVLFMYW].[IVLFMYW]"
file1 = open("search.txt")
l=list()
for i in file1:
    l.append(i.rstrip())
l
```

Out[353...
```
['Protein Pattern Find results',
 'Results for 297 residue sequence "sp|O18423|TXL_EISFE Lysenin OS=Eisenia fetida OX=6396 PE=1 SV=1" starting "MSAKAAEGYE"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 487 residue sequence "sp|P66948|BEPA_ECOLI Beta-barrel assembly-enhancing protease OS=Escherichia coli (strain K12) OX=83333
GN=bepA PE=1 SV=1" starting "MFRQLKKNLV"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 346 residue sequence "sp|P0A910|OMPA_ECOLI Outer membrane protein A OS=Escherichia coli (strain K12) OX=83333 GN=ompA PE=1 SV
=1" starting "MKKTAIAIAV"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 367 residue sequence "sp|P06996|OMPC_ECOLI Outer membrane porin C OS=Escherichia coli (strain K12) OX=83333 GN=ompC PE=1 SV=
1" starting "MKVKVLSLLV"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 536 residue sequence "sp|P04062|GLCM_HUMAN Lysosomal acid glucosylceramidase OS=Homo sapiens OX=9606 GN=GBA PE=1 SV=3" starti
ng "MEFSSPSREE"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 810 residue sequence "sp|P0A940|BAMA_ECOLI Outer membrane protein assembly factor BamA OS=Escherichia coli (strain K12) OX=83
333 GN=bamA PE=1 SV=1" starting "MAMKKLLIAS"',
 '',
 '>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=422 end=430',
 'NTGSFNFGI',
 '',
 '',
 '',
 'Results for 113 residue sequence "sp|P0A937|BAME_ECOLI Outer membrane protein assembly factor BamE OS=Escherichia coli (strain K12) OX=83
333 GN=bamE PE=1 SV=1" starting "MRCKTLTAAA"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 326 residue sequence "sp|P9WIU5|ARFA_MYCTU Peptidoglycan-binding protein ArfA OS=Mycobacterium tuberculosis (strain ATCC 2561
8 / H37Rv) OX=83332 GN=arfA PE=1 SV=1" starting "MASKAGLGQT"',
 '',
 'no matches found for this sequence.',
```

'Results for 291 residue sequence "sp|P9WHT9|PSB_MYCTU Proteasome subunit beta OS=Mycobacterium tuberculosis (strain ATCC 25618 / H37Rv) O
X=83332 GN=prcB PE=1 SV=1" starting "MTWPLPDRLS"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 2290 residue sequence "sp|P03304|POLG_EMCV Genome polyprotein OS=Encephalomyocarditis virus OX=12104 PE=1 SV=1" starting "MAT
TMEQETC"',
'',
'>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=28 end=36',
'RNGFYLLKY',
'',
'>match number 2 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=257 end=265',
'HAGGLLVFM',
'',
'',
'',
'Results for 2292 residue sequence "sp|Q66765|POLG_EMCVR Genome polyprotein OS=Encephalomyocarditis virus (strain Rueckert) OX=650129 PE=1
SV=1" starting "MATTMEQETC"',
'',
'>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=28 end=36',
'RNGFYLLKY',
'',
'>match number 2 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=259 end=267',
'HAGGLLVFM',
'',
'',
'',
'Results for 2293 residue sequence "sp|P12296|POLG_ENMGO Genome polyprotein OS=Mengo encephalomyocarditis virus OX=12107 PE=1 SV=3" starti
ng "MATTMEQEIC"',
'',
'>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=28 end=36',
'RNGFYLLKY',
'',
'>match number 2 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=259 end=267',
'HAGSLLVFM',
'',
'',
'',
'Results for 2181 residue sequence "sp|Q155Z9|POLG_SVV1 Genome polyprotein OS=Seneca Valley virus (isolate -/United States/SSV-001/2002) O
X=686944 PE=1 SV=1" starting "MQNSHFSFDT"',
'',
'>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=95 end=103',
'NNGNMTFNY',
'',
'>match number 2 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=277 end=285',
'HQGALLVAM',
'',
'>match number 3 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=847 end=855',
'KGGKVSFVL',
'',
'>match number 4 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=1450 end=1458',
'TPGLIILAV',
'',
'',
'',
'Results for 2292 residue sequence "sp|P17594|POLG_EMCVD Genome polyprotein OS=Encephalomyocarditis virus (strain emc-d diabetogenic) OX=1
2106 PE=3 SV=2" starting "MATTMEQEIC"',
'',
'>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=28 end=36',
'RNGFYLLKY',
'',
'>match number 2 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=259 end=267',
'HAGSLLVFM',
'',
'>match number 3 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=922 end=930',
'NHGGLEIRL',
'',
'',
'',
'Results for 2292 residue sequence "sp|P17593|POLG_EMCVB Genome polyprotein OS=Encephalomyocarditis virus (strain emc-b nondiabetogenic) O
X=12105 PE=3 SV=1" starting "MATTMEQEIC"',
'',
'>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=28 end=36',
'RNGFYLLKY',
'',
'>match number 2 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=259 end=267',
'HAGSLLVFM',
'',
'>match number 3 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=922 end=930',
'NHGGLEIRL',
'',
'',
'',
'Results for 2295 residue sequence "sp|C0MHL9|POLG_SAFV Genome polyprotein OS=Saffold virus OX=434309 PE=1 SV=1" starting "MACKHGYPLL"',
'',
'>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=265 end=273',
'HAGSLLVFM',
'',
'>match number 2 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=535 end=543',
'TKGKFLISY',
'',
'>match number 3 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=959 end=967',
'RSGRLLVCM',
'',
'',
'',

```
'Results for 426 residue sequence "sp|P41800|MMM1_YEAST Maintenance of mitochondrial morphology protein 1 OS=Saccharomyces cerevisiae (str
ain ATCC 204508 / S288c) OX=559292 GN=MMM1 PE=1 SV=2" starting "MTDSENESTE"',
'',
'>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=300 end=308',
'KPGIAALPI',
'',
'',
'',
'Results for 2124 residue sequence "sp|B8XTP8|POLG_COSAA Genome polyprotein OS=Cosavirus A (isolate Human/Pakistan/0553/-) OX=1554483 PE=3
SV=1" starting "MGANNSKESV"',
'',
'>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=191 end=199',
'HNGILGLFL',
'',
'>match number 2 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=457 end=465',
'TRGKLLISY',
'',
'>match number 3 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=520 end=528',
'SDGWITVWL',
'',
'>match number 4 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=1593 end=1601',
'NVGNMPFVM',
'',
'',
'',
'Results for 245 residue sequence "sp|P0AC02|BAMD_ECOLI Outer membrane protein assembly factor BamD OS=Escherichia coli (strain K12) OX=83
333 GN=bamD PE=1 SV=1" starting "MTRMKYLVAA"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 93 residue sequence "sp|P87108|TIM10_YEAST Mitochondrial import inner membrane translocase subunit TIM10 OS=Saccharomyces cer
evisiae (strain ATCC 204508 / S288c) OX=559292 GN=TIM10 PE=1 SV=1" starting "MSFLGFGGGQ"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 747 residue sequence "sp|P06971|FHUA_ECOLI Ferrichrome outer membrane transporter/phage receptor OS=Escherichia coli (strain
K12) OX=83333 GN=fhuA PE=1 SV=2" starting "MARSKTAQPK"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 485 residue sequence "sp|Q51487|OPRM_PSEAE Outer membrane protein OprM OS=Pseudomonas aeruginosa (strain ATCC 15692 / DSM 226
44 / CIP 104116 / JCM 14847 / LMG 12228 / 1C / PRS 101 / PAO1) OX=208964 GN=oprM PE=1 SV=2" starting "MKRSFLSLAV"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 287 residue sequence "sp|P30656|PSB5_YEAST Proteasome subunit beta type-5 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S
288c) OX=559292 GN=PRE2 PE=1 SV=3" starting "MQAIADSFSV"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 818 residue sequence "sp|Q9STE8|TC753_ARATH Protein TOC75-3, chloroplastic OS=Arabidopsis thaliana OX=3702 GN=TOC75-3 PE=1 SV
=1" starting "MAAFSVNGQL"',
'',
'>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=435 end=443',
'NEGGIIVEI',
'',
'',
'',
'Results for 87 residue sequence "sp|O74700|TIM9_YEAST Mitochondrial import inner membrane translocase subunit TIM9 OS=Saccharomyces cerev
isiae (strain ATCC 204508 / S288c) OX=559292 GN=TIM9 PE=1 SV=1" starting "MDALNSKEQQ"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 261 residue sequence "sp|P25043|PSB2_YEAST Proteasome subunit beta type-2 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S
288c) OX=559292 GN=PUP1 PE=1 SV=1" starting "MAGLSFDNYQ"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 593 residue sequence "tr|L8X6U0|L8X6U0_THACA Mitochondrial distribution and morphology protein 34 OS=Thanatephorus cucumeris
(strain AG1-IA) OX=983506 GN=MDM34 PE=3 SV=1" starting "MDSNVYAFPP"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 103 residue sequence "tr|A0A0N7MLG8|A0A0N7MLG8_9SACH Mitochondrial import inner membrane translocase subunit OS=Lachancea que
becensis OX=1654605 GN=LAQU0_S05e00980g PE=3 SV=1" starting "MALSSIFGGA"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 597 residue sequence "tr|A0A4T0RA24|A0A4T0RA24_9BASI Mitochondrial distribution and morphology protein 34 OS=Wallemia mellico
la OX=1708541 GN=MDM34 PE=3 SV=1" starting "MAFKFQWPDF"',
```

```
                      '',
                      'no matches found for this sequence.',
                      '',
                      '',
                      '',
                      'Results for 472 residue sequence "tr|A0A1E3JCQ6|A0A1E3JCQ6_9TREE Maintenance of mitochondrial morphology protein 1 OS=Cryptococcus depaup
eratus CBS 7855 OX=1295532 GN=MMM1 PE=3 SV=1" starting "MADLFSSTHW"',
                      '',
                      'no matches found for this sequence.',
                      '',
                      '',
                      '',
                      'Results for 89 residue sequence "tr|A0A397IV93|A0A397IV93_9GLOM Mitochondrial import inner membrane translocase subunit OS=Diversispora e
pigaea OX=1348612 GN=Glove_140g159 PE=3 SV=1" starting "MSSLSTKSGN"',
                      '',
                      'no matches found for this sequence.',
                      '',
                      '',
                      '',
                      'Results for 430 residue sequence "tr|A0A679NPU9|A0A679NPU9_GIBZA Mitochondrial distribution and morphology protein 10 OS=Gibberella zeae
OX=5518 GN=MDM10 PE=3 SV=1" starting "MDYVHSAFYE"',
                      '',
                      'no matches found for this sequence.',
                      '',
                      '',
                      '',
                      'Results for 87 residue sequence "tr|A0A7H9HRS1|A0A7H9HRS1_9SACH Mitochondrial import inner membrane translocase subunit OS=Torulaspora gl
obosa OX=48254 GN=HG537_0C06300 PE=3 SV=1" starting "MSSFSPADLS"',
                      '',
                      'no matches found for this sequence.',
                      '',
                      '',
                      '',
                      'Results for 104 residue sequence "tr|A0A022XM38|A0A022XM38_TRISD Mitochondrial import inner membrane translocase subunit OS=Trichophyton
soudanense CBS 452.61 OX=1215331 GN=H105_06003 PE=3 SV=1" starting "MSLLNPFSSS"',
                      '',
                      'no matches found for this sequence.',
                      '',
                      '',
                      '',
                      'Results for 82 residue sequence "tr|R9AF93|R9AF93_WALI9 Mitochondrial import inner membrane translocase subunit OS=Wallemia ichthyophaga
(strain EXF-994 / CBS 113033) OX=1299270 GN=J056_004685 PE=3 SV=1" starting "MDEATQRELN"',
                      '',
                      'no matches found for this sequence.',
                      '',
                      '',
                      '',
                      'Results for 74 residue sequence "tr|A0A4P9XJL8|A0A4P9XJL8_9FUNG Mitochondrial import inner membrane translocase subunit (Fragment) OS=Tha
mnocephalis sphaerospora OX=78915 GN=THASP1DRAFT_11661 PE=3 SV=1" starting "KKQEVMDQVR"',
                      '',
                      'no matches found for this sequence.',
                      '',
                      '',
                      '',
                      'Results for 103 residue sequence "tr|A0A397JJG6|A0A397JJG6_9GLOM Mitochondrial import inner membrane translocase subunit OS=Diversispora
epigaea OX=1348612 GN=Glove_22g228 PE=3 SV=1" starting "MSNFSSDYGL"',
                      '',
                      'no matches found for this sequence.',
                      '',
                      '',
                      '',
                      'Results for 572 residue sequence "tr|A0A0K8L768|A0A0K8L768_9EURO Mitochondrial distribution and morphology protein 34 OS=Aspergillus udag
awae OX=91492 GN=MDM34 PE=3 SV=1" starting "MAFNFNWSPL"',
                      '',
                      'no matches found for this sequence.',
                      '',
                      '',
                      '',
                      'Results for 487 residue sequence "tr|A0A0M8NU54|A0A0M8NU54_9EURO Mitochondrial distribution and morphology protein 10 OS=Penicillium nord
icum OX=229535 GN=MDM10 PE=3 SV=1" starting "MLDFMDYIQL"',
                      '',
                      'no matches found for this sequence.',
                      '',
                      '',
                      '',
                      'Results for 457 residue sequence "tr|A0A0P1KMV5|A0A0P1KMV5_9SACH Mitochondrial distribution and morphology protein 10 OS=Lachancea quebec
ensis OX=1654605 GN=MDM10 PE=3 SV=1" starting "MIDYMEHVLR"',
                      '>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=186 end=194',
                      'NEGLLGYRI',
                      '',
                      '>match number 2 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=445 end=453',
                      'RPGKLGIQL',
                      '',
                      '',
                      '',
                      'Results for 392 residue sequence "tr|W7EFP5|W7EFP5_COCVI Mitochondrial distribution and morphology protein 10 OS=Bipolaris victoriae FI3
OX=930091 GN=MDM10 PE=3 SV=1" starting "MLDFMESVQH"',
                      '',
                      'no matches found for this sequence.',
                      '',
                      '',
                      '',
                      'Results for 473 residue sequence "tr|A0A4Q1BQ83|A0A4Q1BQ83_TREME Mitochondrial distribution and morphology protein 10 OS=Tremella mesente
rica OX=5217 GN=MDM10 PE=3 SV=1" starting "MIPFSTLLLR"',
                      '',
                      'no matches found for this sequence.',
                      '',
```

```
'',
'',
'Results for 96 residue sequence "tr|A0A1E3IUQ0|A0A1E3IUQ0_9TREE Mitochondrial import inner membrane translocase subunit OS=Cryptococcus d
epauperatus CBS 7855 OX=1295532 GN=L204_05429 PE=3 SV=1" starting "MSSFFSSSAN"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 105 residue sequence "tr|A0A1S1VEF0|A0A1S1VEF0_9PEZI Mitochondrial import inner membrane translocase subunit OS=Colletotrichu
m incanum OX=1573173 GN=CSPAE12_11427 PE=3 SV=1" starting "MDSDSVKKAV"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 92 residue sequence "tr|A0A1E4TFS8|A0A1E4TFS8_9ASCO Mitochondrial import inner membrane translocase subunit OS=Tortispora cas
einolytica NRRL Y-17796 OX=767744 GN=CANCADRAFT_25083 PE=3 SV=1" starting "MSAISSLFGK"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 82 residue sequence "tr|A0A3M2RZM6|A0A3M2RZM6_9HYPO Mitochondrial import inner membrane translocase subunit OS=Fusarium kuros
hium OX=2010991 GN=CDV36_009857 PE=3 SV=1" starting "MDSSQVKQAV"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 494 residue sequence "tr|W7ESU6|W7ESU6_COCVI Maintenance of mitochondrial morphology protein 1 OS=Bipolaris victoriae FI3 OX=
930091 GN=MMM1 PE=3 SV=1" starting "MAQDVPTPVP"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 103 residue sequence "tr|L8WKT2|L8WKT2_THACA Mitochondrial import inner membrane translocase subunit OS=Thanatephorus cucumer
is (strain AG1-IA) OX=983506 GN=AG1IA_08553 PE=3 SV=1" starting "MSDFFKTNLG"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 447 residue sequence "tr|A0A0P1KQV1|A0A0P1KQV1_9SACH Maintenance of mitochondrial morphology protein 1 OS=Lachancea quebecens
is OX=1654605 GN=MMM1 PE=3 SV=1" starting "MNYLNGTGST"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 593 residue sequence "tr|F4S4Q0|F4S4Q0_MELLP Maintenance of mitochondrial morphology protein 1 OS=Melampsora larici-populina
(strain 98AG31 / pathotype 3-4-7) OX=747676 GN=MMM1 PE=3 SV=1" starting "MVQAANKTPP"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 447 residue sequence "tr|A0A1J8Q3R3|A0A1J8Q3R3_9AGAM Mitochondrial distribution and morphology protein 10 OS=Rhizopogon vesic
ulosus OX=180088 GN=MDM10 PE=3 SV=1" starting "MHPFASYVLR"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 82 residue sequence "tr|A0A0S6XHC3|A0A0S6XHC3_9FUNG Mitochondrial import inner membrane translocase subunit OS=fungal sp. No.
11243 OX=1603295 GN=ANO11243_032780 PE=3 SV=1" starting "MASSVEQFKD"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 93 residue sequence "tr|A0A0S6XLR7|A0A0S6XLR7_9FUNG Mitochondrial import inner membrane translocase subunit OS=fungal sp. No.
11243 OX=1603295 GN=ANO11243_048140 PE=3 SV=1" starting "MDFSSSSATD"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 450 residue sequence "tr|A0A5N6WXF8|A0A5N6WXF8_9EURO Mitochondrial distribution and morphology protein 10 OS=Aspergillus serg
ii OX=1034303 GN=MDM10 PE=3 SV=1" starting "MLDFMDYIQL"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 433 residue sequence "tr|I2G2S5|I2G2S5_USTH4 Mitochondrial distribution and morphology protein 12 OS=Ustilago hordei (strain
Uh4875-4) OX=1128400 GN=MDM12 PE=3 SV=1" starting "MSLDLDWSLL"',
'',
'no matches found for this sequence.',
'',
'',
'',
'Results for 85 residue sequence "tr|A0A6A6E252|A0A6A6E252_9PEZI Mitochondrial import inner membrane translocase subunit OS=Zopfia rhizoph
ila CBS 207.26 OX=1314779 GN=K469DRAFT_579234 PE=3 SV=1" starting "MNADLSKLSD"',
'',
'no matches found for this sequence.',
'',
```

```
 '',
 '',
 'Results for 295 residue sequence "tr|A0A397IYJ5|A0A397IYJ5_9GLOM Maintenance of mitochondrial morphology protein 1 OS=Diversispora epigae
a OX=1348612 GN=MMM1 PE=3 SV=1" starting "MMSFVQEISP"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 94 residue sequence "tr|A0A163BMD4|A0A163BMD4_DIDRA Mitochondrial import inner membrane translocase subunit OS=Didymella rabi
ei OX=5454 GN=ST47_g7006 PE=3 SV=1" starting "MDGLGGGLAN"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 100 residue sequence "tr|A0A4U9EKF2|A0A4U9EKF2_GIBZA Mitochondrial import inner membrane translocase subunit OS=Gibberella ze
ae OX=5518 GN=MDC_LOCUS91003 PE=3 SV=1" starting "MSAQQLNIDN"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 87 residue sequence "tr|A0A0P1KYC9|A0A0P1KYC9_9SACH Mitochondrial import inner membrane translocase subunit OS=Lachancea queb
ecensis OX=1654605 GN=LAQU0_S22e00650g PE=3 SV=1" starting "MSSITQNELA"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 471 residue sequence "tr|A0A0K8L4N9|A0A0K8L4N9_9EURO Mitochondrial distribution and morphology protein 10 OS=Aspergillus udag
awae OX=91492 GN=MDM10 PE=3 SV=1" starting "MLDFMDYIQL"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 476 residue sequence "tr|A0A022Y1S2|A0A022Y1S2_TRISD Maintenance of mitochondrial morphology protein 1 OS=Trichophyton soudan
ense CBS 452.61 OX=1215331 GN=MMM1 PE=3 SV=1" starting "MSSSENASCP"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 506 residue sequence "tr|A0A1E3IN29|A0A1E3IN29_9TREE Mitochondrial distribution and morphology protein 10 OS=Cryptococcus dep
auperatus CBS 7855 OX=1295532 GN=MDM10 PE=3 SV=1" starting "MISFSAFILR"',
 '>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=189 end=197',
 'TPGNVFFTL',
 '',
 '',
 '',
 'Results for 276 residue sequence "tr|A0A0C9MEN7|A0A0C9MEN7_9FUNG Mitochondrial distribution and morphology protein 12 OS=fungal sp. No.11
243 OX=1603295 GN=MDM12 PE=3 SV=1" starting "MSIEINWEAL"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 136 residue sequence "tr|C4Y4D1|C4Y4D1_CLAL4 Mitochondrial import inner membrane translocase subunit OS=Clavispora lusitaniae
(strain ATCC 42720) OX=306902 GN=CLUG_02503 PE=3 SV=1" starting "MLYDLTLRFQ"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 596 residue sequence "tr|A0A4T0NNN3|A0A4T0NNN3_9BASI Mitochondrial distribution and morphology protein 34 OS=Wallemia mellico
la OX=1708541 GN=MDM34 PE=3 SV=1" starting "MAFKFQWPDF"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 566 residue sequence "tr|A0A5N6X001|A0A5N6X001_9EURO Mitochondrial distribution and morphology protein 34 OS=Aspergillus serg
ii OX=1034303 GN=MDM34 PE=3 SV=1" starting "MAFNFNWSPL"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 246 residue sequence "tr|R9AEJ4|R9AEJ4_WALI9 Mitochondrial distribution and morphology protein 12 OS=Wallemia ichthyophaga (s
train EXF-994 / CBS 113033) OX=1299270 GN=MDM12 PE=3 SV=1" starting "MSVDIDYTKL"',
 '',
 '>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=140 end=148',
 'SPGFMSLPV',
 '',
 '',
 '',
 'Results for 422 residue sequence "tr|W6YAW4|W6YAW4_COCCA Mitochondrial distribution and morphology protein 12 OS=Bipolaris zeicola 26-R-1
3 OX=930089 GN=MDM12 PE=3 SV=1" starting "MSIDINWDTI"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 109 residue sequence "tr|C4YHK0|C4YHK0_CANAW Mitochondrial import inner membrane translocase subunit OS=Candida albicans (str
ain WO-1) OX=294748 GN=CAWG_03547 PE=3 SV=1" starting "MAFWNSSSSS"',
 '',
```

```
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 389 residue sequence "tr|A0A1E4RZU6|A0A1E4RZU6_CYBJN Mitochondrial distribution and morphology protein 10 OS=Cyberlindnera ja
dinii (strain ATCC 18201 / CBS 1600 / BCRC 20928 / JCM 3617 / NBRC 0987 / NRRL Y-1542) OX=983966 GN=MDM10 PE=3 SV=1" starting "MDYILQSFY
R"',
 '',
 '>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=155 end=163',
 'NNGTMTFYW',
 '',
 '',
 '',
 'Results for 476 residue sequence "tr|A0A022Y7A1|A0A022Y7A1_TRISD Mitochondrial distribution and morphology protein 10 OS=Trichophyton sou
danense CBS 452.61 OX=1215331 GN=MDM10 PE=3 SV=1" starting "MLDFMDYIQL"',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 438 residue sequence "tr|A0A0K8LMH9|A0A0K8LMH9_9EURO Mitochondrial distribution and morphology protein 12 OS=Aspergillus udag
awae OX=91492 GN=MDM12 PE=3 SV=1" starting "MSIDVNWRSA"',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 491 residue sequence "tr|A0A5N6WN12|A0A5N6WN12_9EURO Maintenance of mitochondrial morphology protein 1 OS=Aspergillus sergii
OX=1034303 GN=MMM1 PE=3 SV=1" starting "MTFQQNEPSA"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 470 residue sequence "tr|A0A1X7R723|A0A1X7R723_9SACH Mitochondrial distribution and morphology protein 10 OS=Kazachstania sau
lgeensis OX=1789683 GN=MDM10 PE=3 SV=1" starting "MLDFMDYVLR"',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 447 residue sequence "tr|A0A4V4MTY2|A0A4V4MTY2_9BASI Mitochondrial distribution and morphology protein 10 OS=Wallemia mellico
la OX=1708541 GN=MDM10 PE=3 SV=1" starting "MINYAEYILR"',
 '',
 '>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=348 end=356',
 'NQGAASLAI',
 '',
 '',
 '',
 'Results for 92 residue sequence "tr|W6YKR2|W6YKR2_COCCA Mitochondrial import inner membrane translocase subunit OS=Bipolaris zeicola 26-R
-13 OX=930089 GN=COCCADRAFT_83123 PE=3 SV=1" starting "MNSLGPSDPK"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 727 residue sequence "tr|A0A1E3J463|A0A1E3J463_9TREE Mitochondrial distribution and morphology protein 34 OS=Cryptococcus dep
auperatus CBS 7855 OX=1295532 GN=MDM34 PE=3 SV=1" starting "MSFVFPSWST"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 315 residue sequence "tr|A0A4Z0A2T1|A0A4Z0A2T1_9AGAM Maintenance of mitochondrial morphology protein 1 OS=Hericium alpestre O
X=135208 GN=MMM1 PE=3 SV=1" starting "MGNNYVFSMH"',
 '',
 '>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=270 end=278',
 'RKGTWKVVL',
 '',
 '',
 '',
 'Results for 84 residue sequence "tr|A0A316YLW8|A0A316YLW8_9BASI Mitochondrial import inner membrane translocase subunit OS=Acaromyces ing
oldii OX=215250 GN=FA10DRAFT_266578 PE=3 SV=1" starting "MASGLGESDQ"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 371 residue sequence "tr|A0A421JA48|A0A421JA48_9ASCO Maintenance of mitochondrial morphology protein 1 OS=Meyerozyma sp. JA9
OX=2028340 GN=MMM1 PE=3 SV=1" starting "MSRDVTRPDE"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 400 residue sequence "tr|A0A4Z1JFG0|A0A4Z1JFG0_9HELO Mitochondrial distribution and morphology protein 10 OS=Botrytis ellipti
ca OX=278938 GN=MDM10 PE=3 SV=1" starting "MLEFMDYVQF"',
 '',
 '>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=39 end=47',
 'TPGGLRFNL',
 '',
 '',
 '',
 'Results for 451 residue sequence "tr|A0A066WF73|A0A066WF73_TILAU Mitochondrial distribution and morphology protein 12 OS=Tilletiaria anom
ala (strain ATCC 24038 / CBS 436.72 / UBC 951) OX=1037660 GN=MDM12 PE=3 SV=1" starting "MSLDLDWELF"',
 '',
 'no matches found for this sequence.',
 '',
```

```
 '',
 '',
 'Results for 612 residue sequence "tr|W7EBM9|W7EBM9_COCVI Mitochondrial distribution and morphology protein 34 OS=Bipolaris victoriae FI3
OX=930091 GN=MDM34 PE=3 SV=1" starting "MAFNFNWSPL"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 85 residue sequence "tr|A0A0B4F2C5|A0A0B4F2C5_METAF Mitochondrial import inner membrane translocase subunit (Fragment) OS=Met
arhizium anisopliae (strain ARSEF 549) OX=1276135 GN=MAN_03196 PE=3 SV=1" starting "MDSSTVKQSV"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 436 residue sequence "tr|A0A6H0XRF6|A0A6H0XRF6_9PEZI Mitochondrial distribution and morphology protein 12 OS=Peltaster fructi
cola OX=286661 GN=MDM12 PE=3 SV=1" starting "MAIQGVVELV"',
 '',
 '>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=184 end=192',
 'NLGYFHLPL',
 '',
 '',
 '',
 'Results for 459 residue sequence "tr|C7GWY1|C7GWY1_YEAS2 Mitochondrial distribution and morphology protein 34 OS=Saccharomyces cerevisiae
(strain JAY291) OX=574961 GN=MDM34 PE=3 SV=1" starting "MSFRFNEAVF"',
 '',
 '>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=111 end=119',
 'NNGSFTIPI',
 '',
 '',
 '',
 'Results for 352 residue sequence "tr|A0A1E3IDZ3|A0A1E3IDZ3_9TREE Mitochondrial distribution and morphology protein 12 OS=Cryptococcus dep
auperatus CBS 7841 OX=1295531 GN=MDM12 PE=3 SV=1" starting "MSLDIRWDLL"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 88 residue sequence "tr|W1QBV2|W1QBV2_OGAPD Mitochondrial import inner membrane translocase subunit OS=Ogataea parapolymorpha
(strain ATCC 26012 / BCRC 20466 / JCM 22074 / NRRL Y-7560 / DL-1) OX=871575 GN=HPODL_05323 PE=3 SV=1" starting "MSEIDPKLLQ"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 480 residue sequence "tr|A0A0B0E367|A0A0B0E367_NEUCS Mitochondrial distribution and morphology protein 10 OS=Neurospora crass
a OX=5141 GN=MDM10 PE=3 SV=1" starting "MREFMNYITN"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 334 residue sequence "tr|W1QEI9|W1QEI9_OGAPD Maintenance of mitochondrial morphology protein 1 OS=Ogataea parapolymorpha (str
ain ATCC 26012 / BCRC 20466 / JCM 22074 / NRRL Y-7560 / DL-1) OX=871575 GN=MMM1 PE=3 SV=1" starting "MSIESPDYSD"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 99 residue sequence "tr|A0A146FL93|A0A146FL93_ASPKA Mitochondrial import inner membrane translocase subunit OS=Aspergillus ka
wachii OX=1069201 GN=RIB2604_02102970 PE=3 SV=1" starting "MGIFGGSSSP"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 676 residue sequence "tr|A0A074SYM9|A0A074SYM9_9AGAM Mitochondrial distribution and morphology protein 34 OS=Rhizoctonia sola
ni 123E OX=1423351 GN=MDM34 PE=3 SV=1" starting "MSFQFEWPRF"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 108 residue sequence "tr|A0A317VII2|A0A317VII2_9EURO Mitochondrial import inner membrane translocase subunit OS=Aspergillus h
eteromorphus CBS 117.55 OX=1448321 GN=BO70DRAFT_365049 PE=3 SV=1" starting "MALFGSDAPA"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 96 residue sequence "tr|A0A0N1P1Z5|A0A0N1P1Z5_9EURO Mitochondrial import inner membrane translocase subunit OS=Phialophora at
tinorum OX=1664694 GN=AB675_9359 PE=3 SV=1" starting "MDDMESQPQI"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 434 residue sequence "tr|A0A1T3C9Q0|A0A1T3C9Q0_9HYPO Mitochondrial distribution and morphology protein 12 OS=Trichoderma guiz
houense OX=1491466 GN=MDM12 PE=3 SV=1" starting "MSIDLNWETL"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 415 residue sequence "tr|A0A068AZ58|A0A068AZ58_NEUCS Maintenance of mitochondrial morphology protein 1 OS=Neurospora crassa O
X=5141 GN=MMM1 PE=3 SV=1" starting "MADVCPSRSE"',
 '',
```

```
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 89 residue sequence "tr|A0A318YCX2|A0A318YCX2_ASPNB Mitochondrial import inner membrane translocase subunit OS=Aspergillus ne
oniger (strain CBS 115656) OX=1448310 GN=BO87DRAFT_342634 PE=3 SV=1" starting "MDAQTQVDIS"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 467 residue sequence "tr|A0A1E3IFQ2|A0A1E3IFQ2_9TREE Maintenance of mitochondrial morphology protein 1 OS=Cryptococcus depaup
eratus CBS 7841 OX=1295531 GN=MMM1 PE=3 SV=1" starting "MADLFSSTHW"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 509 residue sequence "tr|A0A1E3IV39|A0A1E3IV39_9TREE Mitochondrial distribution and morphology protein 10 OS=Cryptococcus dep
auperatus CBS 7841 OX=1295531 GN=MDM10 PE=3 SV=1" starting "MIPFSAFILR"',
 '',
 '>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=192 end=200',
 'TPGNVFFTL',
 '',
 '',
 '',
 'Results for 403 residue sequence "tr|M3JX54|M3JX54_CANMX Maintenance of mitochondrial morphology protein 1 OS=Candida maltosa (strain Xu3
16) OX=1245528 GN=MMM1 PE=3 SV=1" starting "MTNVIIRTET"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 89 residue sequence "tr|A0A317WHZ7|A0A317WHZ7_9EURO Mitochondrial import inner membrane translocase subunit OS=Aspergillus sc
lerotioniger CBS 115572 OX=1450535 GN=BO94DRAFT_535789 PE=3 SV=1" starting "MDAQTQVDVT"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 82 residue sequence "tr|A0A1Y1XUV4|A0A1Y1XUV4_9FUNG Mitochondrial import inner membrane translocase subunit OS=Basidiobolus m
eristosporus CBS 931.73 OX=1314790 GN=K493DRAFT_318620 PE=3 SV=1" starting "MSQEFSEQDQ"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 107 residue sequence "tr|V2Y0I7|V2Y0I7_MONRO Mitochondrial import inner membrane translocase subunit OS=Moniliophthora roreri
 (strain MCA 2997) OX=1381753 GN=Moror_25 PE=3 SV=1" starting "MSDFFKNPLS"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 454 residue sequence "tr|A0A0N0NQR8|A0A0N0NQR8_9EURO Maintenance of mitochondrial morphology protein 1 OS=Phialophora attinor
um OX=1664694 GN=MMM1 PE=3 SV=1" starting "MASTLTFTQG"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 371 residue sequence "tr|M3HT68|M3HT68_CANMX Mitochondrial distribution and morphology protein 10 OS=Candida maltosa (strain
Xu316) OX=1245528 GN=MDM10 PE=3 SV=1" starting "MGTKDVSLQD"',
 '',
 '>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=78 end=86',
 'KNGTMIVYL',
 '',
 '',
 '',
 'Results for 563 residue sequence "tr|A0A4Z1EFD8|A0A4Z1EFD8_9HELO Mitochondrial distribution and morphology protein 34 OS=Botrytis tulipae
 OX=87230 GN=MDM34 PE=3 SV=1" starting "MAFNFNWSPL"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 439 residue sequence "tr|A0A7H9HUZ3|A0A7H9HUZ3_9SACH Maintenance of mitochondrial morphology protein 1 OS=Torulaspora globosa
 OX=48254 GN=MMM1 PE=3 SV=1" starting "MNDLDLSDSL"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 97 residue sequence "tr|A0A1D9QJG9|A0A1D9QJG9_SCLS1 Mitochondrial import inner membrane translocase subunit OS=Sclerotinia sc
lerotiorum (strain ATCC 18683 / 1980 / Ss-1) OX=665079 GN=sscle_14g098380 PE=3 SV=1" starting "MDSLNNAFAS"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 291 residue sequence "tr|A0A0C9Y863|A0A0C9Y863_9AGAR Maintenance of mitochondrial morphology protein 1 OS=Laccaria amethystin
a LaAM-08-1 OX=1095629 GN=MMM1 PE=3 SV=1" starting "MPNNYVFSLQ"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 104 residue sequence "tr|F2Q3P0|F2Q3P0_TRIEC Mitochondrial import inner membrane translocase subunit OS=Trichophyton equinum
```

```
(strain ATCC MYA-4606 / CBS 127.97) OX=559882 GN=TEQG_07715 PE=3 SV=1" starting "MSLSNPFSSS"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 355 residue sequence "tr|A0A1B7SJT8|A0A1B7SJT8_9ASCO Mitochondrial distribution and morphology protein 10 OS=Ogataea polymorp
ha OX=460523 GN=MDM10 PE=3 SV=1" starting "MTVSSKNTEN"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 481 residue sequence "tr|A0A317VB63|A0A317VB63_9EURO Maintenance of mitochondrial morphology protein 1 OS=Aspergillus sclerot
ioniger CBS 115572 OX=1450535 GN=MMM1 PE=3 SV=1" starting "MSFQPNEPVP"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 431 residue sequence "tr|A0A1D9Q492|A0A1D9Q492_SCLS1 Mitochondrial distribution and morphology protein 12 OS=Sclerotinia scle
rotiorum (strain ATCC 18683 / 1980 / Ss-1) OX=665079 GN=MDM12 PE=3 SV=1" starting "MSIDLNWETL"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 82 residue sequence "tr|A0A4V6TQG4|A0A4V6TQG4_9BASI Mitochondrial import inner membrane translocase subunit OS=Wallemia melli
cola OX=1708541 GN=E3Q18_02679 PE=3 SV=1" starting "MDEATQRELN"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 549 residue sequence "tr|E6ZUB6|E6ZUB6_SPORE Mitochondrial distribution and morphology protein 10 OS=Sporisorium reilianum (s
train SRZ2) OX=999809 GN=MDM10 PE=3 SV=1" starting "MHDFVSHILR"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 367 residue sequence "tr|F2PVT4|F2PVT4_TRIEC Mitochondrial distribution and morphology protein 12 OS=Trichophyton equinum (st
rain ATCC MYA-4606 / CBS 127.97) OX=559882 GN=MDM12 PE=3 SV=1" starting "MSIDINWEAA"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 638 residue sequence "tr|R9AC33|R9AC33_WALI9 Mitochondrial distribution and morphology protein 34 OS=Wallemia ichthyophaga (s
train EXF-994 / CBS 113033) OX=1299270 GN=MDM34 PE=3 SV=1" starting "MLREFHHGRG"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 1082 residue sequence "tr|A0A177UU34|A0A177UU34_9BASI Mitochondrial distribution and morphology protein 34 OS=Tilletia caries
OX=13290 GN=MDM34 PE=3 SV=1" starting "MSFRFDWPEF"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 430 residue sequence "tr|A0A4Q4WFG1|A0A4Q4WFG1_9PEZI Maintenance of mitochondrial morphology protein 1 OS=Monosporascus sp. m
g162 OX=1081914 GN=MMM1 PE=3 SV=1" starting "MAASDMCPTK"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 120 residue sequence "tr|A0A0A2L1F7|A0A0A2L1F7_PENIT Mitochondrial import inner membrane translocase subunit OS=Penicillium i
talicum OX=40296 GN=PITC_016530 PE=3 SV=1" starting "MSSIFGSGSA"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 544 residue sequence "tr|A0A1Y1UI98|A0A1Y1UI98_9TREE Mitochondrial distribution and morphology protein 34 OS=Kockovaella impe
ratae OX=4999 GN=MDM34 PE=3 SV=1" starting "MSFQFPDWST"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 579 residue sequence "tr|A0A1S8B958|A0A1S8B958_9PEZI Mitochondrial distribution and morphology protein 34 OS=Diplodia seriata
OX=420778 GN=MDM34 PE=3 SV=1" starting "MAFNFNWSPL"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 590 residue sequence "tr|A0A4U9EWY3|A0A4U9EWY3_GIBZA Mitochondrial distribution and morphology protein 34 OS=Gibberella zeae
OX=5518 GN=MDM34 PE=3 SV=1" starting "MAFNFNWSPL"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 439 residue sequence "tr|A0A7G3ZJ32|A0A7G3ZJ32_9SACH Maintenance of mitochondrial morphology protein 1 OS=Torulaspora globosa
```

```
    OX=48254 GN=MMM1 PE=3 SV=1" starting "MSDLGLSDSV"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 1087 residue sequence "tr|E6ZTQ6|E6ZTQ6_SPORE Mitochondrial distribution and morphology protein 34 OS=Sporisorium reilianum
(strain SRZ2) OX=999809 GN=MDM34 PE=3 SV=1" starting "MSFNFKWPTF"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 93 residue sequence "tr|A0A366QWY0|A0A366QWY0_9HYPO Mitochondrial import inner membrane translocase subunit OS=Fusarium coffe
atum OX=231269 GN=FIESC28_10192 PE=3 SV=1" starting "MSAQQLSIDN"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 361 residue sequence "tr|A0A1Y1WWG5|A0A1Y1WWG5_9FUNG Maintenance of mitochondrial morphology protein 1 OS=Basidiobolus merist
osporus CBS 931.73 OX=1314790 GN=MMM1 PE=3 SV=1" starting "MSSEWGDYLS"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 91 residue sequence "tr|A0A0A2LBN8|A0A0A2LBN8_PENIT Mitochondrial import inner membrane translocase subunit OS=Penicillium it
alicum OX=40296 GN=PITC_087780 PE=3 SV=1" starting "MEQQLDLTKL"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 106 residue sequence "tr|A0A7H8R0K1|A0A7H8R0K1_9EURO Mitochondrial import inner membrane translocase subunit OS=Talaromyces r
ugulosus OX=121627 GN=TRUGW13939_07034 PE=3 SV=1" starting "MSFFGSSSST"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 568 residue sequence "tr|A0A2G5BC32|A0A2G5BC32_COERN Maintenance of mitochondrial morphology protein 1 OS=Coemansia reversa
(strain ATCC 12441 / NRRL 1564) OX=763665 GN=MMM1 PE=3 SV=1" starting "MPTAVQSTFG"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 746 residue sequence "tr|V2XMY8|V2XMY8_MONRO Mitochondrial distribution and morphology protein 34 OS=Moniliophthora roreri (s
train MCA 2997) OX=1381753 GN=MDM34 PE=3 SV=1" starting "MSFTFNWPRF"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 98 residue sequence "tr|A0A1Y1XVB0|A0A1Y1XVB0_9FUNG Mitochondrial import inner membrane translocase subunit OS=Basidiobolus m
eristosporus CBS 931.73 OX=1314790 GN=K493DRAFT_318508 PE=3 SV=1" starting "MSEYASDFSI"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 94 residue sequence "tr|A0A2G5BAH0|A0A2G5BAH0_COERN Mitochondrial import inner membrane translocase subunit OS=Coemansia reve
rsa (strain ATCC 12441 / NRRL 1564) OX=763665 GN=COEREDRAFT_81840 PE=3 SV=1" starting "MSSLGSSTFS"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 384 residue sequence "tr|A0A4U0UCD9|A0A4U0UCD9_9PEZI Mitochondrial distribution and morphology protein 12 OS=Hortaea thailand
ica OX=706561 GN=MDM12 PE=3 SV=1" starting "MSVEIDWQAL"',
 '',
 '>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=141 end=149',
 'NLGYFHLPL',
 '',
 '',
 '',
 'Results for 510 residue sequence "tr|A0A177V897|A0A177V897_9BASI Mitochondrial distribution and morphology protein 12 OS=Tilletia caries
OX=13290 GN=MDM12 PE=3 SV=1" starting "MSVELDWTLL"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 601 residue sequence "tr|A0A6A6ZLA9|A0A6A6ZLA9_9PLEO Mitochondrial distribution and morphology protein 34 OS=Ophiobolus disse
minans OX=1469910 GN=MDM34 PE=3 SV=1" starting "MAFNFNWSPL"',
 '',
 '>match number 1 to "[KRHQNST].G.[IVLFMYWAC].[IVLFMYW].[IVLFMYW]" start=433 end=441',
 'RHGSLDVGV',
 '',
 '',
 '',
 'Results for 88 residue sequence "tr|A0A5N6VDX0|A0A5N6VDX0_9EURO Mitochondrial import inner membrane translocase subunit OS=Aspergillus tr
ansmontanensis OX=1034304 GN=BDV41DRAFT_569727 PE=3 SV=1" starting "MEQQLDVSKL"',
 '',
 'no matches found for this sequence.',
 '',
 '',
```

```
'',
'Results for 969 residue sequence "tr|A0A2S5BCG9|A0A2S5BCG9_9BASI Mitochondrial distribution and morphology protein 34 OS=Rhodotorula taiw
anensis OX=741276 GN=MDM34 PE=3 SV=1" starting "MSFNFEWPEF"',
'',
'no matches found for this sequence.',
'',
'',
'Results for 435 residue sequence "tr|A0A318ZVJ9|A0A318ZVJ9_ASPNB Mitochondrial distribution and morphology protein 12 OS=Aspergillus neon
iger (strain CBS 115656) OX=1448310 GN=MDM12 PE=3 SV=1" starting "MSIEVDWGAA"',
'',
'no matches found for this sequence.',
'',
'',
'Results for 366 residue sequence "tr|A0A074RZ23|A0A074RZ23_9AGAM Mitochondrial distribution and morphology protein 12 OS=Rhizoctonia sola
ni 123E OX=1423351 GN=MDM12 PE=3 SV=1" starting "MSVDLAWDSL"',
'',
'no matches found for this sequence.',
'',
'',
'Results for 540 residue sequence "tr|A0A1T3CMM4|A0A1T3CMM4_9HYPO Mitochondrial distribution and morphology protein 34 OS=Trichoderma guiz
houense OX=1491466 GN=MDM34 PE=3 SV=1" starting "MAFNFNWSPL"',
'',
'no matches found for this sequence.',
'',
'',
'Results for 471 residue sequence "tr|A0A0D0YLB6|A0A0D0YLB6_9TREE Maintenance of mitochondrial morphology protein 1 OS=Cryptococcus gattii
VGIV IND107 OX=1296105 GN=MMM1 PE=3 SV=1" starting "MYEPFSPNLT"',
'',
'no matches found for this sequence.',
'',
'',
'Results for 392 residue sequence "tr|A0A6A7B9D8|A0A6A7B9D8_9PLEO Mitochondrial distribution and morphology protein 10 OS=Plenodomus trach
eiphilus IPT5 OX=1408161 GN=MDM10 PE=3 SV=1" starting "MLDFMDNVQN"',
'',
'no matches found for this sequence.',
'',
'',
'Results for 411 residue sequence "tr|A0A179ICC7|A0A179ICC7_CORDF Maintenance of mitochondrial morphology protein 1 OS=Cordyceps confragos
a OX=2714763 GN=MMM1 PE=3 SV=1" starting "MSTGTCPARS"',
'',
'no matches found for this sequence.',
'',
'',
'Results for 417 residue sequence "tr|D5G665|D5G665_TUBMM Mitochondrial distribution and morphology protein 12 OS=Tuber melanosporum (stra
in Mel28) OX=656061 GN=MDM12 PE=3 SV=1" starting "MSIEVNWETL"',
'',
'no matches found for this sequence.',
'',
'',
'Results for 470 residue sequence "tr|W1QCQ0|W1QCQ0_OGAPD Mitochondrial distribution and morphology protein 34 OS=Ogataea parapolymorpha
(strain ATCC 26012 / BCRC 20466 / JCM 22074 / NRRL Y-7560 / DL-1) OX=871575 GN=MDM34 PE=3 SV=1" starting "MSFQINWEAI"',
'',
'no matches found for this sequence.',
'',
'',
'Results for 97 residue sequence "tr|A0A2T2P8E8|A0A2T2P8E8_CORCC Mitochondrial import inner membrane translocase subunit OS=Corynespora ca
ssiicola Philippines OX=1448308 GN=BS50DRAFT_481075 PE=3 SV=1" starting "MDSLGASGMS"',
'',
'no matches found for this sequence.',
'',
'',
'Results for 95 residue sequence "tr|A0A2S7Q3K8|A0A2S7Q3K8_9HELO Mitochondrial import inner membrane translocase subunit OS=Rutstroemia s
p. NJR-2017a BVV2 OX=2070413 GN=CJF31_00010768 PE=3 SV=1" starting "MDSLNNTFGS"',
'',
'no matches found for this sequence.',
'',
'',
'Results for 517 residue sequence "tr|A0A6A5QNQ2|A0A6A5QNQ2_AMPQU Maintenance of mitochondrial morphology protein 1 OS=Ampelomyces quisqua
lis OX=50730 GN=MMM1 PE=3 SV=1" starting "MANPSPVSAP"',
'',
'no matches found for this sequence.',
'',
'',
'Results for 94 residue sequence "tr|A0A6A5QIC6|A0A6A5QIC6_AMPQU Mitochondrial import inner membrane translocase subunit OS=Ampelomyces qu
isqualis OX=50730 GN=BDU57DRAFT_557509 PE=3 SV=1" starting "MDALGGGLAN"',
'',
'no matches found for this sequence.',
'',
'',
'Results for 371 residue sequence "tr|D5GJQ6|D5GJQ6_TUBMM Mitochondrial distribution and morphology protein 10 OS=Tuber melanosporum (stra
in Mel28) OX=656061 GN=MDM10 PE=3 SV=1" starting "MLTYMDYLQN"',
'',
'no matches found for this sequence.',
'',
'',
```

```
 '',
 'Results for 93 residue sequence "tr|A0A024RYK1|A0A024RYK1_HYPJR Mitochondrial import inner membrane translocase subunit OS=Hypocrea jecor
ina (strain ATCC 56765 / BCRC 32924 / NRRL 11460 / Rut C-30) OX=1344414 GN=M419DRAFT_12855 PE=3 SV=1" starting "MNAAQTGIEN"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 2920 residue sequence "tr|A0A024SMC6|A0A024SMC6_HYPJR Mitochondrial distribution and morphology protein 12 OS=Hypocrea jecori
na (strain ATCC 56765 / BCRC 32924 / NRRL 11460 / Rut C-30) OX=1344414 GN=MDM12 PE=3 SV=1" starting "MASNLTTSKP"',
 '',
 'no matches found for this sequence.',
 '',
 '',
 '',
 'Results for 564 residue sequence "tr|A0A1V6TBD8|A0A1V6TBD8_9EURO Mitochondrial distribution and morphology protein 34 OS=Penicillium stec
kii OX=303698 GN=MDM34 PE=3 SV=1" starting "MAFNFNWSPL"',
 '',
 'no matches found for this sequence.',
 ...]
```