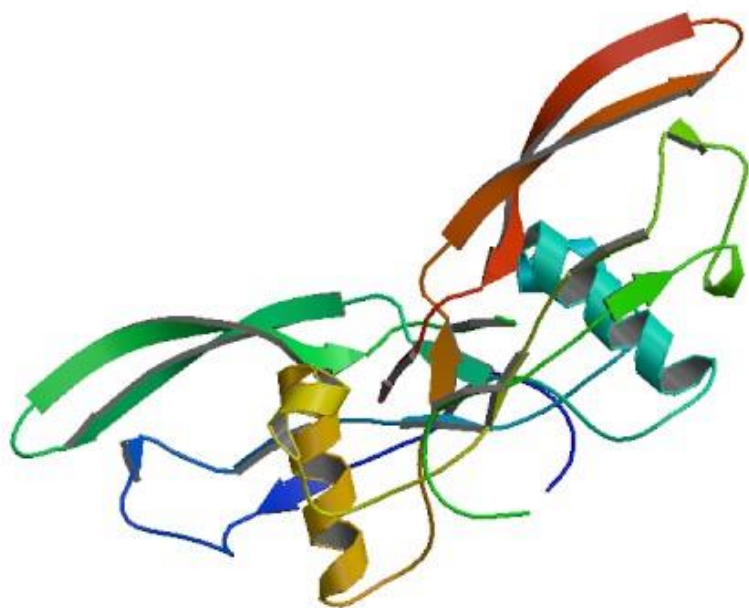


ΔΟΜΙΚΗ ΒΙΟΧΗΜΕΙΑ & ΣΤΟΙΧΕΙΑ ΒΙΟΠΛΗΡΟΦΟΡΙΚΗΣ

ΑΝΑΦΟΡΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ



Καθηγήτρια : κ. Χολή-Παπαδοπούλου Θεωδώρα

Ονοματεπώνυμο : Γκάναλη Βασιλική

Α.Ε.Μ : 8922

Ημερομηνία : 24/06/2019

Τίτλος : Οστική Μορφογενετική Πρωτεΐνη 3

Κωδικός : **P12645**

Στην εργασία αυτή μελετάται η **οστική μορφογενετική πρωτεΐνη 3 (BMP 3/ Bone Morphogenetic Protein 3)**.

Σε ένα αρχικό στάδιο επισκεπτόμαστε τη διεύθυνση της βάσης δεδομένων των πρωτεϊνικών ακολουθιών SPROT:

<http://www.expasy.ch/sprot/>.

Έπειτα διαλέγουμε την τράπεζα δεδομένων επιλέγοντας το σύνδεσμο : **List of UniProtKB/Swiss-Prot (reviewed) entries**. Στο πεδίο πάνω δεξιά πληκτρολογούμε το **BMP 3** και ακολούθως επιλέγουμε την πρώτη πρωτεΐνη με τον κωδικό καταχώρησης **P12645 (Homo Sapiens/Human)**. Μπαίνουμε στην σελίδα της πρωτεΐνης και λαμβάνουμε από εκεί τις παρακάτω πληροφορίες.

Πρωτεΐνη : Οστική Μορφογενετική Πρωτεΐνη 3

Γονίδιο : BMP 3

Οργανισμός : Homo Sapiens (Human)

Λειτουργία : Ρυθμίζει αρνητικά την οστική πυκνότητα. Επιπλέον ανταγωνίζεται την ικανότητα ορισμένων οστεογονικών BMPs να επάγουν διαφοροποίηση (δηλαδή την διαφοροποίηση των μυοκυττάρων σε οστεοκύτταρα) και οστεοποίηση οστεοπαράγοντων.

Η αμινοξική ακολουθία της πρωτεΐνης σε *FASTA format* :

```
>sp|P12645|BMP3_HUMAN Bone morphogenetic protein 3 OS=Homo sapiens OX=9606 GN=BMP3 PE=1 SV=2
MAGASRLFLWLGCFCVSLAQGERPKPPFPELRKAVPGDRTAGGGPDSELQPQDKVSEHM
LRLYDRYSTVQAARTPGSLEGGSQPWRPRLREGNTVRSFRAAAETLERKGLYIFNLTS
LTKSENILSATLYFCIGELGNISLSCPVS GGCSHHAQRKHIQIDLSAWTLKFSRNQSLL
GHLSVDMAKSHRDIMSWLSKDITQLLRKAKENEEFLIGFNITSKGRQLPKRRLPFPEPYI
LVYANDAAISEPESVSSSLQGHRNFPTGTVPKWDSHIRAALSIERRKKRSTGVLLPLQNN
ELPGA EYQYKKDEVWEERKPYKTLQAQAPEKSKNKKKQRKGPHRKSQTLQFDEQTLKKAR
RKQWIEPRNCARRYLKVD FADIGWSEWISPKSFDAYYCSGACQFPMPKSLKPSNHATIQ
SIVRAVG VVPGIPEPCCVPEKMSSLSILFFDENKNVVLKVYPNMTVESACR
```

Βάσεις δεδομένων της νουκλεοτιδικής αλληλουχίας :

- EMBL
- GenBank
- DDBJ

EMBL (Accession Number BC096269, Sequence Version BC096269.2, Keywords MGC.):

>ENA|BC096269|BC096269.2 Homo sapiens bone morphogenetic protein 3, mRNA (cDNA clone MGC:116934 IMAGE:40006302), complete cds.

```
TCAGCGTTGGAGTGGAGACGGCGCCCGCAGCGCCCTGCGCGGGTGAGGTCCGCGCAGCTG
CTGGGGAAGAGCCCACCTGTCAGGCTGCGCTGGGTCAGCGCAGCAAGTGGGGCTGGCCGC
TATCTCGCTGCACCCGGCCGCGTCCCGGGCTCCGTGCGCCCTCGCCCCAGCTGGTTTGA
```

GTTC AACCTCGGCTCCGCCGCCGGCTCCTTGCGCCTTCGGAGTGTCCTCGCAGCGACGCC
GGGAGCCGACGCGCCGCGCGGGTACCTAGCCATGGCTGGGGCGAGCAGGCTGCTCTTTCT
GTGGCTGGGCTGCTTCTGCGTGAGCCTGGCGCAGGGAGAGAGACCGAAGCCACCTTTCCC
GGAGCTCCGCAAAGCTGTGCCAGGTGACCGCACGGCAGGTGGTGGCCCGGACTCCGAGCT
GCAGCCGCAAGACAAGGTCTCTGAACACATGCTGCGGCTCTATGACAGGTACAGCACGGT
CCAGGCGGGCCGGACACCGGGCTCCCTGGAGGGAGGCTCGCAGCCCTGGCGCCCTCGGCT
CCTGCGCGAAGGCAACACGGTTTCGCAGCTTTTCGGGCGGCAGCAGCAGAACTCTTGAAAG
AAAAGGACTGTATATCTTCAATCTGACATCGCTAACCAAGTCTGAAAACATTTTGTCTGC
CACACTGTATTTCTGTATTGGAGAGCTAGGAAACATCAGCCTGAGTTGTCCAGTGTCTGG
AGGATGCTCCCATCATGCTCAGAGGAAACACATTCAGATTGATCTTTCTGCATGGACCCT
CAAATTCAGCAGAAACCAAAGTCAACTCCTTGGCCATCTGTCAGTGGATATGGCCAAATC
TCATCGAGATATTATGTCCTGGCTGTCTAAAGATATCACTCAATTCTTGAGGAAGGCCAA
AGAAAATGAAGAGTTCCTCATAGGATTTAACATTACGTCCAAGGGACGCCAGCTGCCAAA
GAGGAGGTTACCTTTTCCAGAGCCTTATATCTTGGTATATGCCAATGATGCCGCCATTTCT
TGAGCCAGAAAGTGTGGTATCAAGCTTACAGGGACACCGGAATTTTCCCACTGGAAGTGT
TCCCAAATGGGATAGCCACATCAGAGCTGCCCTTTCCATTGAGCGGAGGAAGAAGCGCTC
TACTGGGGTCTTGCTGCCTCTGCAGAACACGAGCTTCCTGGGGCAGAATACCAGTATAA
AAAGGATGAGGTGTGGGAGGAGAGAAAGCCTTACAAGACCCTTCAGGCTCAGGCCCCTGA
AAAGAGTAAGAATAAAAAGAAACAGAGAAAGGGGCCTCATCGGAAGAGCCAGACGCTCCA
ATTTGATGAGCAGACCCTGAAAAAGGCAAGGAGAAAGCAGTGGATTGAACCTCGGAATTG
CGCCAGGAGATACCTCAAGGTAGACTTTGCAGATATTGGCTGGAGTGAATGGATTATCTC
CCCCAAGTCCTTTGATGCCTATTATTGCTCTGGAGCATGCCAGTTCCCCATGCCAAAGTC
TTTGAAGCCATCAAATCATGCTACCATCCAGAGTATAGTGAGAGCTGTGGGGGTCGTTCC
TGGGATTCTTGAGCCTTGCTGTGTACCAGAAAAGATGTCCTCACTCAGTATTTTATTCTT
TGATGAAAATAAGAATGTAGTGCTTAAAGTATACCCTAACATGACAGTAGAGTCTTGCGC
TTGCAGATAACCTGGCAAAGAACTCATTTGAATGC

EMBL -mRNA Translation (Accession Number AAH96269, Sequence Version AAH96269.1, Keywords MGC.):

>ENA|AAH96269|AAH96269.1 Homo sapiens (human) bone morphogenetic protein 3

ATGGCTGGGGCGAGCAGGCTGCTCTTTCTGTGGCTGGGCTGCTTCTGCGTGAGCCTGGCG
CAGGGAGAGAGACCGAAGCCACCTTTCCCGGAGCTCCGCAAAGCTGTGCCAGGTGACCGC
ACGGCAGGTGGTGGCCCGGACTCCGAGCTGCAGCCGCAAGACAAGGTCTCTGAACACATG
CTGCGGCTCTATGACAGGTACAGCACGGTCCAGGCGGGCCGGACACCGGGCTCCCTGGAG
GGAGGCTCGCAGCCCTGGCGCCCTCGGCTCCTGCGCGAAGGCAACACGGTTTCGCAGCTTT
CGGGCGGCAGCAGCAGAACTCTTGAAAGAAAAGGACTGTATATCTTCAATCTGACATCG

CTAACCAAGTCTGAAAACATTTTGTCTGCCACACTGTATTTCTGTATTGGAGAGCTAGGA
AACATCAGCCTGAGTTGTCCAGTGTCTGGAGGATGCTCCCATCATGCTCAGAGGAAACAC
ATTCAGATTGATCTTTCTGCATGGACCCTCAAATTCAGCAGAAACCAAAGTCAACTCCTT
GGCCATCTGTCAGTGGATATGGCCAAATCTCATCGAGATATTATGTCCTGGCTGTCTAAA
GATATCACTCAATTCTTGAGGAAGGCCAAAGAAAATGAAGAGTTCCTCATAGGATTTAAC
ATTACGTCCAAGGGACGCCAGCTGCCAAAGAGGAGGTTACCTTTTCCAGAGCCTTATATC
TTGGTATATGCCAATGATGCCGCCATTTCTGAGCCAGAAAGTGTGGTATCAAGCTTACAG
GGACACCGGAATTTTCCCACTGGAAGTGTCCCAAATGGGATAGCCACATCAGAGCTGCC
CTTTCCATTGAGCGGAGGAAGAAGCGCTCTACTGGGGTCTTGCTGCCTCTGCAGAACAAAC
GAGCTTCCTGGGGCAGAATACCAGTATAAAAAGGATGAGGTGTGGGAGGAGAGAAAGCCT
TACAAGACCCTTCAGGCTCAGGCCCTGAAAAGAGTAAGAATAAAAAGAAACAGAGAAAG
GGGCCTCATCGGAAGAGCCAGACGCTCCAATTTGATGAGCAGACCCTGAAAAAGGCAAGG
AGAAAGCAGTGGATTGAACCTCGGAATTGCGCCAGGAGATACCTCAAGGTAGACTTTGCA
GATATTGGCTGGAGTGAATGGATTATCTCCCCAAGTCCTTTGATGCCTATTATTGCTCT
GGAGCATGCCAGTTCCCCATGCCAAAGTCTTTGAAGCCATCAAATCATGCTACCATCCAG
AGTATAGTGAGAGCTGTGGGGGTCGTTCTGGGATTCTTGAGCCTTGCTGTGTACCAGAA
AAGATGTCCTCACTCAGTATTTTATTCTTTGATGAAAATAAGAATGTAGTGCTTAAAGTA
TACCCTAACATGACAGTAGAGTCTTGCGCTTGCAGATAA

GenBank (Accession Number BC096269, Sequence Version BC096269.2, Keywords MGC.):

ORIGIN

1 tcagcgttgg agtggagacg gcgcccgcag cgccctgcgc gggtaggtc cgcgcagctg
61 ctggggaaga gccacctgt caggctgcgc tgggtcagcg cagcaagtgg ggctggccgc
121 tatctcgtg caccggccg cgtcccgggc tccgtgcgc ctcgcccag ctggtttgga
181 gttcaacct cggctccgc gccggctct tgcgcctcg gagtgtccg cagcgacgcc
241 gggagccgac gcgccgcgc ggtacctagc catggctggg gcgagcagcg tgcttttct
301 gtggctgggc tgcttctcg tgagcctggc gcagggagag agaccgaagc cactttccc
361 ggagctccgc aaagctgtgc caggtgaccg cacggcaggt ggtggcccgg actccgagct
421 gcagccgcaa gacaaggtct ctgaacacat gctgcggctc tatgacaggt acgcacggt
481 ccagggcgcc cggacaccgg gtcctctgga gggaggctcg cagccctggc gccctcggt
541 cctgcgcgaa ggcaacacgg ttcgcagctt tcggcgcgca gcagcagaaa ctctgaaag
601 aaaaggactg tatatctca atctgacatc gctaaccaag tctgaaaaca tttgtctgc
661 cacactgtat ttctgtatt gagagctagg aaacatcagc ctgagttgtc cagtgtctgg
721 aggatgctcc catcatgctc agaggaaaca cattcagatt gatctttctg catggaccct
781 caaattcagc agaaacaaa gtcaactcct tggccatctg tcagtggata tggccaaatc

841 tcatcgagat attatgtcct ggctgtctaa agatatact caattcttga ggaaggccaa
 901 agaaaatgaa gagttcctca taggatttaa cattacgtcc aagggacgcc agctgccaaa
 961 gaggaggta cctttccag agccttatat ctggtatat gccaatgatg ccgccatttc
 1021 tgagccagaa agtgtggtat caagcttaca gggacaccgg aattttccca ctggaactgt
 1081 tcccaaatgg gatagccaca tcagagctgc cctttccatt gagcggagga agaagcgctc
 1141 tactggggtc ttgctgcctc tgcagaacaa cgagcttctt ggggcagaat accagtataa
 1201 aaaggatgag gtgtgggagg agagaaagcc ttacaagacc cttcaggctc agggccctga
 1261 aaagagtaag aataaaaaga aacagagaaa ggggcctcat cggaagagcc agacgtcca
 1321 attgatgag cagaccctga aaaaggcaag gagaaagcag tgattgaac ctcggaattg
 1381 cgccaggaga tacctaagg tagactttgc agatattggc tggagtgaat ggattatctc
 1441 ccccaagtcc tttagtcct attattgctc tggagcatgc cagtcccca tgccaaagtc
 1501 ttgaagcca tcaaatcatg ctaccatcca gagtatagtg agagctgtgg gggtcgttc
 1561 tgggattcct gagccttget gtgtaccaga aaagatgtcc tcaactagta tttattctt
 1621 tgatgaaaat aagaatgtag tgctaaagt ataccctaac atgacagtag agtcttgcgc
 1681 ttgcagataa cctggcaaag aactcattg aatgc

//

GenBank -mRNA Translation (Accession Number AAH96270, Sequence Version AAH96270.1, Keywords MGC.):

ORIGIN

1 magasrllfl wlgcfcvsla qgerpkppfp elrkavpgdr tagggpdsel qpqdkvsehm
 61 lrlydrystv qartpgsle ggsqpwrprl lregntvrsf raanaetler kglyifnlts
 121 ltksenilsa tlyfcigelg nislscpvsg gcshhaqrkh iqidlswtl kfsrnqsll
 181 ghlsvdmaks hrdimswlsk ditqflrkak eneefligfn itskgrqlpk rrlpfpepyi
 241 lvyandaais epesvsslq ghrnfptgtv pkwdshiraa lsierkkrs tgvllplqnn
 301 elpgaeyqyk kdevweerkp yktlqaqape ksknkkkqrk gphrksqtlq fdeqtlkkr
 361 rkqwieprnc arrylkvdfa digwsewiis pksfdayyys gacqfmpkps lkpsnhatiq
 421 sivravgvvp gipepcvpe kmsslslff denknvvlkv ypnmtvesca cr

//

DDBJ (Accession Number BC096271, Sequence Version BC096271.2, Keywords MGC.):

BASE COUNT 417 a 452 c 478 g 368 t

ORIGIN

1 tcagcgttgg agtggagacg gcgcccgcag cgccctgcgc gggtagagtc cgcgagctg
 61 ctggggaaga gccacactgt caggctgcgc tgggtcagcg cagcaagtgg ggctggccgc
 121 tatctcgtg cacccggccg cgtcccgggc tccgtcgcgc ctcgcccag ctggtttgga

181 gttcaaccct cggctccgcc gccggctcct tgcgccttcg gagtgtcccg cagcgacgcc
 241 gggagccgac gcgccgcgcg ggtacctagc catggctggg gcgagcaggc tgctctttct
 301 gtggctgggc tgcttctgcg tgagcctggc gcagggagag agaccgaagc cacctttccc
 361 ggagctccgc aaagctgtgc caggtgaccg cacggcaggt ggtggcccgg actccgagct
 421 gcagccgcaa gacaaggtct ctgaacacat gctgcggctc tatgacaggt acagcacggt
 481 ccaggcggcc cggacaccgg gctccctgga gggaggctcg cagccctggc gccctcggct
 541 cctgcgcgaa ggcaacacgg ttcgcagctt tcgggcggca gcagcagaaa ctcttgaaag
 601 aaaaggactg tatatcttca atctgacatc gctaaccaag tctgaaaaca tttgtctgc
 661 cacactgtat ttctgtattg gagagctagg aaacatcagc ctgagttgtc cagtgtctgg
 721 aggatgctcc catcatgctc agaggaaaca cattcagatt gatctttctg catggaccct
 781 caaattcagc agaaacaaaa gtcaactcct tggccatctg tcagtggata tggccaaatc
 841 tcatcgagat attatgtcct ggctgtctaa agatatcact caattcttga ggaaggccaa
 901 agaaaatgaa gagtctctca taggatttaa cattacgtcc aaggagcgc agctgccaaa
 961 gaggagggtta cctttccag agccttatat ctgtgtatat gccaatgatg ccgccatttc
 1021 tgagccagaa agtgtggtat caagcttaca gggacaccgg aattttccca ctggaactgt
 1081 tcccaaatgg gatagccaca tcagagctgc ctttccatt gagcggagga agaagcgctc
 1141 tactggggtc ttgctgcctc tgcagaacaa cgagcttctc ggggcagaat accagtataa
 1201 aaaggatgag gtgtgggagg agagaaagcc ttacaagacc cttcaggctc aggccctga
 1261 aaagagtaag aataaaaaaga aacagagaaa ggggcctcat cggaagagcc agacgtcca
 1321 attgatgag cagaccctga aaaaggcaag gagaagcag tggattgaac ctcggaattg
 1381 cgccaggaga tacctcaagg tagactttgc agatattggc tggagtgaat ggattatctc
 1441 ccccaagtcc ttgatgcct attattgctc tggagcatgc cagttccca tgccaaagtc
 1501 ttgaagcca tcaaatcatg ctaccatcca gagtatagt agagctgtgg ggtcgttcc
 1561 tgggattcct gagccttgc tgtaccaga aaagatgtcc tcaactagta tttattctt
 1621 tgatgaaat aagaatgtag tgcttaaagt ataccctaac atgacagtag agtcttgcgc
 1681 ttgcagataa cctggcaaa gaaactcattg aatgc

//

DDBJ -mRNA Translation (Accession Number AAH96271, Sequence Version AAH96271.1, Keywords MGC.):

BEGIN

1 MAGASRLLFL WLGCFCVSLA QGERPKPPFP ELRKAVPGDR TAGGGPDSEL QPQDKVSEHM
 61 LRLYDRYSTV QAARTPGSLE GGSQPWRPRL LREGNTVRSF RAAAAETLER KGLYIFNLTS
 121 LTKSENILSA TLYFCIGELG NISLSCPVS GCSHHAQRKH IQIDLSAWTL KFSRNQSQLL
 181 GHLSVDMAKS HRDIMSWLSK DITQFLRKAK ENEEFLIGFN ITSKGRQLPK RRLPFPEPYI
 241 LUYANDAAIS EPESVSSLQ GHRNFPTGTV PKWD SHIRAA LSIERRKKRS TGVLLPLQNN

301 ELPGAEYQYK KDEVWEERKP YKTLQAQAPE KSKNKKKQQRK GPHRKSQTLQ FDEQTLKKAR
361 RKQWIEPRNC ARRYLKVDFA DIGWSEWIIIS PKSFDAYYCS GACQFMPPKS LKPSNHATIQ
421 SIVRAVGVVP GIPEPCCVPE KMSSLSILFF DENKNVVLKV YPNMTVESCA CR

//

The Consensus CDS (CCDS) project (Sequence IDs included in CCDS 3588.1)

Blue highlighting indicates alternating exons.

Red highlighting indicates amino acids encoded across a splice junction.

Nucleotide Sequence (1419 nt):

ATGGCTGGGGCGAGCAGGCTGCTCTTTCTGTGGCTGGGCTGCTTCTGCGTGAGCCTGGCGCAGGGAGAGA
GACCGAAGCCACCTTTCCCGGAGCTCCGCAAAGCTGTGCCAGGTGACCGCACGGCAGGTGGTGGCCCGGA
CTCCGAGCTGCAGCCGCAAGACAAGGTCTCTGAACACATGCTGCGGCTCTATGACAGGTACAGCACGGTC
CAGGCGGCCCCGGACACCGGGCTCCCTGGAGGGAGGCTCGCAGCCCTGGCGCCCTCGGCTCCTGCGCGAAG
GCAACACGGTTTCGACGCTTTTCGGGCGGCAGCAGCAGAACTCTTGAAAGAAAAGGACTGTATATCTTCAA
TCTGACATCGCTAACCAAGTCTGAAAACATTTTGTCTGCCACACTGTATTTCTGTATTGGAGAGCTAGGA
AACATCAGCCTGAGTTGTCCAGTGTCTGGAGGATGCTCCCATCATGCTCAGAGGAAACACATTTCAGATTG
ATCTTTCTGCATGGACCCTCAAATTCAGCAGAAACCAAAGTCAACTCCTTGGCCATCTGTCTAGTGGATAT
GGCCAAATCTCATCGAGATATTATGTCTGGCTGTCTAAAGATATCACTCAACTCTTGAGGAAGGCCAAA
GAAAATGAAGAGTTCTCATAGGATTTAACATTACGTCCAAGGGACGCCAGCTGCCAAAGAGGAGGTTAC
CTTTTCAGAGCCTTATATCTTGGTATATGCCAATGATGCCGCCATTTCTGAGCCAGAAAGTGTGGTATC
AAGCTTACAGGGACACCGGAATTTTCCCACTGGAAGTGTCCCAATGGGATAGCCACATCAGAGCTGCC
CTTTCCATTGAGCGGAGGAAGAAGCGCTCTACTGGGGTCTTGTCTGCCTCTGCAGAACAACGAGCTTCCTG
GGGCAGAATACAGTATAAAAAGGATGAGGTGTGGGAGGAGAGAAAGCCTTACAAGACCCTTCAGGCTCA
GGCCCTGAAAAGAGTAAGAATAAAAAGAAACAGAGAAAGGGCCTCATCGGAAGAGCCAGACGCTCCAA
TTTGATGAGCAGACCCTGAAAAGGCAAGGAGAAAGCAGTGGATTGAACCTCGGAATTGCGCCAGGAGAT
ACCTCAAGGTAGACTTTGCAGATATTGGCTGGAGTGAATGGATTATCTCCCCCAAGTCCTTTGATGCCTA
TTATTGCTCTGGAGCATGCCAGTTCCCCATGCCAAAGTCTTTGAAGCCATCAAATCATGCTACCATCCAG
AGTATAGTGAGAGCTGTGGGGGTCGTTTCTGGGATTCTTGAGCCTTGCTGTGTACCAGAAAAGATGTCCT
CACTCAGTATTTTATTCTTTGATGAAAATAAGAATGTAGTGCTTAAAGTATACCCTAACATGACAGTAGA
GTCTTGCGCTTGACAGATAA

Translation (472 aa):

MAGASRLFLWLGCFCVSLAQGERPKPPFPELRKAVPGDRTAGGGPDSELQPQDKVSEHMLRLYDRYSTV
QAARTPGSLEGGSPWRPRLREGNTVRSFRAAAA^{TLERKGLYIFNLTSLT}SENILSATLYFCIGELG
NISLSCPVS^{GGCSHHAQRKHIQIDLSAWTLKFSRNQS}QLLGHLSVDMAKSHRDIMSWLSKDITQLLRKAK
ENEEFLIGFNITSKGRQLPKRRLPFPEPYILVYANDAAISEPESVSS^{LQGHRNFPTGTVPKWD}SHIRAA
LSIERRKKRSTGVLLPLQNNELPGA^{EYQYKKDEVWEERKPYKTLQA}QAPEKSKNKKKQQRKGPHRKSQTLQ
FDEQTLKKARRKQWIEPRNCARRYLKVD^{FADIGWSEWIIIS}PKSFDAYYCSGACQFMPPKSLKPSNHATIQ
SIVRAVGVVPGIPEPCCVPEKMSSLSILFFDENKNVVLKVYPNMTVESCA^{CR}

Ακολούθως μεταβαίνουμε στη σελίδα <https://www.expasy.org/proteomics> και επιλέγουμε το εργαλείο **Translate**

• nucleotide sequence translation • [\[more\]](#). Στη συνέχεια, εισάγουμε μια ακολουθία DNA στο πλαίσιο που εμφανίζεται (χωρίς αριθμούς και κενά) και πατάμε το κουμπί TRANSLATE SEQUENCE. Τα αποτελέσματα περιέχουν τρία πλαίσια για την κατεύθυνση 5'3' (μετάφραση του νοηματικού κλώνου) και τρία πλαίσια για την κατεύθυνση 3'5' (μετάφραση του αντινοηματικού κλώνου) με τα Ανοιχτά Πλαίσια Ανάγνωσης (Open Reading Frames) επισημασμένα με «κόκκινο». Επιλέγουμε το πλαίσιο (Frame) που περιέχει τα λιγότερα Stop-κωδικόνια:

5'3' Frame 1

Met AGASRLFLWLGCFCVSLAQGERPKPPFPELRKAVPGDRTAGGGPDSELQPQDKVSEH Met LRL
YDRYSTVQAARTPGSLEGGSPWRPRLREGNTVRSFRAAAAETLERKGLYIFNLTSLTKSENILSA
TLYFCIGELGNISLSCPVS GGCSHHAQRKHIQIDLSAWTLKFSRNQS QLLGHLSD Met AKSHRDI Met
SWLSKDITQLLRKAKENEEFLIGFNITSKGRQLPKRRLPFPEPYILVYANDAAISEPESVSSLQGHR
NFPTGTVPKWDSHIRAALSIERKKRSTGVLLPLQNNELPGA EYQYKKDEVWEERKPYKTLQAQAP
EKSKNKKKQRKGPHRKSQTLQFDEQTLKKARRKQWIEPRNCARRYLKVDFADIGWSEWIISPKSFD
AAYCSGACQFP Met PKSLKPSNHATIQSIVRAVGVP GIPEPCCVPEK Met SSLSILFFDENKNVVLKV
YPN Met TVESACR Stop

3'5' Frame 1

LSASARLYCHVRVYFKHYILIFIKE Stop NTE Stop GHLEFWYTARLRNPRNDPHSSHYTLDGS Met I Stop
WLQRLWHGELACSRAIIGIKGLGGDNPFPTANICKVYLEVSPGAIPRFNPLSPCLFQGLLIKLERLA
LP Met RPLSLFLFILTLFRGLSLKGLVRLSLLPHLILFILVFCPRKLVVLQRQQDPSRALLPPLNGKSS
DVAIPFGNSSSGKIPVSL Stop A Stop YHTFWLRNGGIIGIYQDIRLWKR Stop PPLWQLASLGRNVKSYE
ELFIFFGLPQELSDIFRQPGHNIS Met RFGHIH Stop Q Met AKELTLVSAEFEGPCRKINLNVFPLS Met Met
GASSRHWTQADVS Stop LSNTIEICGRQNVFRLG Stop RCQIEDIQSFSFKSFCCCRPKAANRVAFAQ
EPRAPGLRASLQGARCPRGLDRAVPVIEPQHVFRLDLVRLQLGVRATTCRAVTWHSFAELRERWLR
SLSLRQAHAEEAAQPQKEQPARPSH

Στη συνέχεια, επιστρέφουμε στη διεύθυνση της βάσης δεδομένων **CCDS** και αυτή τη φορά επιλέγουμε από την νουκλεοτιδική αλληλουχία τη μεταφράσιμη αλληλουχία (δηλαδή τα νουκλεοτίδια που είναι επισημασμένα με «μπλε» χρώμα) μαζί με 7 νουκλεοτίδια πριν το μεταφράσιμο τμήμα (**5'GCAGCAG 3'**). Στο επόμενο βήμα επαναλαμβάνεται η διαδικασία της μετάφρασης με τη βοήθεια του εργαλείου [Translate](#) • nucleotide sequence translation • [\[more\]](#) από την κατηγορία Proteomics. Τα αποτελέσματα δίνονται παρακάτω :

Nucleotide Sequence :

GCAGCAGAACTCTTGAAAGAAAAGGACTGTATATCTTCAATCTGACATCGCTAACCAAGTCTGAAAACATTTTGTCTGCCACACTGTA
TTTCTGTATTGGAGAGCTAGGAAACATCAGCCTGAGTTGTCCAGTGTCTGGAGGATGCTCCCATCATGCTCAGAGGAAACACATTGAGA
TTGATCTTTCTGCATGGACCCTCAAATTCAGCAGAAACCAAGTCAACTCCTTGGCCATCTGTCACTGAGTGGATATGGCCAAATCTCATCGA
GATATTATGTCTGGCTGTCTAAAGATATCACTCAACTCTTGAGGAAGGCCAAAGAAAATGAAGAGTTCTCATAGGATTTAACATTAC
GTCCAAGGGACGCCAGCTGCCAAAGAGGAGGTTACCTTTTCCAGAGCCTTATATCTTGGTATATGCCAATGATGCCGCCATTTCTGAGC
CAGAAAGTGTGGTATCAAGCTTACAGGGACACCGGAATTTTCCCACTGGA ACT GTTCCCAAATGGGATAGCCACATCAGAGCTGCCCTT
TCCATTGAGCGGAGGAAGAAGCGCTCTACTGGGTCTTGTCTGCTCTGCAGAACACGAGCTTCCTGGGGCAGAATACCAGTATAAAAA
GGATGAGGTGTGGGAGGAGAGAAAGCCTTACAAGACCCTTCAGGCTCAGGCCCTGAAAAGAGTAAGAATAAAAAGAAACAGAGAAAGG
GGCCTCATCGGAAGAGCCAGACGCTCCAATTTGATGAGCAGACCCTGAAAAGGCAAGGAGAAAGCAGTGGATTGAACCTCGGAATTGC
GCCAGGAGATACCTCAAGGTAGACTTTGCAGATATTGGCTGGAGTGAATGGATTATCTCCCCAAGTCCTTTGATGCCTATTATTGCTC
TGGAGCATGCCAGTTCCCCATGCCAAAG

Translation :

5'3' Frame 1

AAETLERKGLYIFNLTSLTKSENILSATLYFCIGELGNISLSCPVS GGCSHHAQRKHIQIDLSAWTLKF
SRNQS QLLGHLSD Met AKSHRDI Met SWLSKDITQLLRKAKENEEFLIGFNITSKGRQLPKRRLPFPE
PYILVYANDAAISEPESVSSLQGHRNFPTGTVPKWDSHIRAALSIERKKRSTGVLLPLQNNELPGA
EYQYKKDEVWEERKPYKTLQAQAPEKSKNKKKQRKGPHRKSQTLQFDEQTLKKARRKQWIEPRNC
ARRYLKVDFADIGWSEWIISPKSFD AAYCSGACQFP Met PK

3'5' Frame 1

LWHGELACSRAIIGIKGLGGDNPFPTANICKVYLEVSPGAIPRFNPLSPCLFQGLLIKLERLALP Met
RPLSLFLFILTLFRGLSLKGLVRLSLLPHLILFILVFCPRKLVVLQRQQDPSRALLPPLNGKSSDVAI
PFGNSSSGKIPVSL Stop A Stop YHTFWLRNGGIIGIYQDIRLWKR Stop PPLWQLASLGRNVKSYEELFI
FFGLPQELSDIFRQPGHNIS Met RFGHIH Stop Q Met AKELTLVSAEFEGPCRKINLNVFPLS Met Met GAS
SRHWTQADVS Stop LSNTIEICGRQNVFRLG Stop RCQIEDIQSFSFKSFCC

Τέλος, επιστρέφουμε στη διεύθυνση της βάσης δεδομένων **CCDS** και αυτήν τη φορά επιλέγουμε από τη νουκλεοτιδική ακολουθία μόνο τη μεταφράσιμη αλληλουχία. Επαναλαμβάνουμε τη διαδικασία μετάφρασης με τη βοήθεια του εργαλείου [Translate](#) • nucleotide sequence translation • [\[more\]](#) από την κατηγορία Proteomics και τα αποτελέσματα φαίνονται παρακάτω :

Nucleotide Sequence :

```
AAACTCTTGAAAGAAAAGGACTGTATATCTTCAATCTGACATCGCTAACCAAGTCTGAAAACATTTTGTCTGCCACACTGTATTTCTGT  
ATTGGAGAGCTAGGAAACATCAGCCTGAGTTGTCCAGTGTCTGGAGGATGCTCCCATCATGCTCAGAGGAAACACATTCAGATTGATCT  
TTCTGCATGGACCCTCAAATTCAGCAGAAACCAAAGTCAACTCCTTGGCCATCTGTCTAGTGGATATGGCCAAATCTCATCGAGATATTA  
TGTCTTGGCTGTCTAAAGATATCACTCAACTCTTGAGGAAGGCCAAAGAAAATGAAGAGTTTCTCATAGGATTTAACATTACGTCCAAG  
GGACGCCAGCTGCCAAAGAGGAGGTTACCTTTTCCAGAGCCTTATATCTTGGTATATGCCAATGATGCCGCCATTTCTGAGCCAGAAAG  
TGTGGTATCAAGCTTACAGGGACACCGGAATTTTCCACTGGAAGTGTCCCAAATGGGATAGCCACATCAGAGCTGCCCTTTCCATTG  
AGCGGAGGAAGAAGCGCTCTACTGGGGTCTTGCTGCCTCTGCAGAACACGAGCTTCCTGGGGCAGAATACCAGTATAAAAAGGATGAG  
GTGTGGGAGGAGAGAAAGCCTTACAAGACCCTTCAGGCTCAGGCCCTGAAAAGAGTAAGAATAAAAAGAAACAGAGAAAGGGGCCTCA  
TCGGAAGAGCCAGACGCTCCAATTTGATGAGCAGACCCTGAAAAGGCAAGGAGAAAGCAGTGGATTGAACCTCGGAATTGCGCCAGGA  
GATACCTCAAGGTAGACTTTGCAGATATTGGCTGGAGTGAATGGATTATCTCCCCCAAGTCCTTTGATGCCTATTATTGCTCTGGAGCA  
TGCCAGTTCCCCATGCCAAAG
```

Translation :

5'3' Frame 3

```
TLERKGLYIFNLTSLTkSENILSATLYFCIGELGNISLSCPVSggGCSHHAQRKHIQIDLSAWTLKFSRN  
QSQLLGHLsVD Met AKSHRDI Met SWLSKDIQLLRKAKENEEFLIGFNITSKGRQLPKRRLPFPEPYI  
LVYANDAAISEPESVSSSLQGHrNFPTGTVPKWDSHIRAALSIERKKRSTGVLLPLQNNELPGAey  
QYKKDEVWEERKPYKTLQAQAPeKSKNKKKQKRGPHRKSQTLQFDEQTLKKARRKQWIEPRNCAR  
RYLKVDfADIGWSEWIISPKSFDAYYCSGACQFP Met PK
```

3'5' Frame 1

```
LWHGELACSRAIIGIKGLGGDNPFTPANICKVYLEVSPGAIPRFNPLLSPCLFQGLLIKLERLALP Met  
RPLSLFLFILTLFRGLSLKGLVRLSLLPHLILFILVFCPRKLVLVLRQDPSRALLPPLNGKGSsDVAI  
PFGNSSSGKIPVSL Stop A Stop YHTFWLRNGGIIGIYQDIRLWKR Stop PPLWQLASLGRNVKSYEELFI  
FFGLPQELSDIFRQPGHNIS Met RFGHIH Stop Q Met AKELTLVSAEFEGPCRKINLNVFPLS Met Met GA  
SSRHWTtQADVS Stop LSNTEIQCGRQNVFRLG Stop RCQIEDIQSFSFKS
```

Έπειτα, κάνουμε την πολλαπλή ομοιοπαράθεση (multiple sequence alignment) με τον εξής τρόπο: ξαναμπαίνουμε στη διεύθυνση <https://www.expasy.org/>, επιλέγουμε την κατηγορία Proteomics και το εργαλείο **Multalin** • Multiple sequence alignment • [\[more\]](#). Στο παράθυρο που εμφανίζεται μεταφέρουμε διαφορετικές μεταφρασμένες αλληλουχίες νουκλεοτιδίων για τη σύγκριση της ομοιότητά τους. Πρώτα εισάγουμε ολόκληρη τη νουκλεοτιδική ακολουθία της BMP 3 σε μορφή FASTA, μετά-ακολουθία που αποτελείται από τη μεταφράσιμη αλληλουχία μαζί με 8 μεταφρασμένα νουκλεοτίδια πριν το μεταφράσιμο τμήμα (sequence 1). Τέλος, εισάγουμε μόνο τη μεταφράσιμη ακολουθία (sequence 2) και πατάμε το SUMBIT.

Παρακάτω παρουσιάζονται τα αποτελέσματα που παίρνουμε. Τα αμινόξεα που εμφανίζονται «κόκκινα» στην ομοπαράθεση είναι ταυτόσημα, τα «πράσινα» αμινοξέα μπορεί να είναι ταυτόσημα ή παρόμοιου φορτίου, αλλά δεν βρίσκονται σε όλες τις ακολουθίες που αναλύθηκαν, ενώ τα «μαύρα» δεν παρουσιάζουν καμία ομοιότητα μεταξύ τους.

	10	20	30	40	50	60
BMP3_HUMAN	MAGASRLFLWLGCFCVSLAQGERPKPPFELRKAVPGDRTAGGGPDSELQPQDKVSEHM					
sequence1	-----					
sequence2	-----					
Consensus						
Prim.cons.	MAGASRLFLWLGCFCVSLAQGERPKPPFELRKAVPGDRTAGGGPDSELQPQDKVSEHM					
	70	80	90	100	110	120
BMP3_HUMAN	LRLYDRYSTVQAARTPGSLEGGSQWRPRLREGNTVRSFRAAAATLERKGLYIFNLTS					
sequence1	-----FRAAAATLERKGLYIFNLTS					
sequence2	-----TLERKGLYIFNLTS					
Consensus	fraaaatLERKGLYIFNLTS					
Prim.cons.	LRLYDRYSTVQAARTPGSLEGGSQWRPRLREGNTVRSFRAAAATLERKGLYIFNLTS					
	130	140	150	160	170	180
BMP3_HUMAN	LTKSENILSATLYFCIGELGNISLSCPVSGGCSHHAQRKHIQIDLSAWTLKFSRNQSQLL					
sequence1	LTKSENILSATLYFCIGELGNISLSCPVSGGCSHHAQRKHIQIDLSAWTLKFSRNQSQLL					
sequence2	LTKSENILSATLYFCIGELGNISLSCPVSGGCSHHAQRKHIQIDLSAWTLKFSRNQSQLL					
Consensus	LTKSENILSATLYFCIGELGNISLSCPVSGGCSHHAQRKHIQIDLSAWTLKFSRNQSQLL					
Prim.cons.	LTKSENILSATLYFCIGELGNISLSCPVSGGCSHHAQRKHIQIDLSAWTLKFSRNQSQLL					
	190	200	210	220	230	240
BMP3_HUMAN	GHLSVDMAKSHRDIMSWLSKDITQLLRKAKENEEFLIGFNITSKGRQLPKRRLPFPEPYI					
sequence1	GHLSVDMAKSHRDIMSWLSKDITQLLRKAKENEEFLIGFNITSKGRQLPKRRLPFPEPYI					
sequence2	GHLSVDMAKSHRDIMSWLSKDITQLLRKAKENEEFLIGFNITSKGRQLPKRRLPFPEPYI					
Consensus	GHLSVDMAKSHRDIMSWLSKDITQLLRKAKENEEFLIGFNITSKGRQLPKRRLPFPEPYI					
Prim.cons.	GHLSVDMAKSHRDIMSWLSKDITQLLRKAKENEEFLIGFNITSKGRQLPKRRLPFPEPYI					
	250	260	270	280	290	300
BMP3_HUMAN	LVYANDAAISEPESVSSLQGHRNFTGTVPKWDSHIRAALSIERRKKRSTGVLLPLQNN					
sequence1	LVYANDAAISEPESVSSLQGHRNFTGTVPKWDSHIRAALSIERRKKRSTGVLLPLQNN					
sequence2	LVYANDAAISEPESVSSLQGHRNFTGTVPKWDSHIRAALSIERRKKRSTGVLLPLQNN					
Consensus	LVYANDAAISEPESVSSLQGHRNFTGTVPKWDSHIRAALSIERRKKRSTGVLLPLQNN					
Prim.cons.	LVYANDAAISEPESVSSLQGHRNFTGTVPKWDSHIRAALSIERRKKRSTGVLLPLQNN					
	310	320	330	340	350	360
BMP3_HUMAN	ELPGAQYQYKKDEVWEERKPYKTLQAQAEKSKNKKKQRKGPHRKSQTLQFDEQTLKKAR					
sequence1	ELPGAQYQYKKDEVWEERKPYKTLQAQAEKSKNKKKQRKGPHRKSQTLQFDEQTLKKAR					
sequence2	ELPGAQYQYKKDEVWEERKPYKTLQAQAEKSKNKKKQRKGPHRKSQTLQFDEQTLKKAR					
Consensus	ELPGAQYQYKKDEVWEERKPYKTLQAQAEKSKNKKKQRKGPHRKSQTLQFDEQTLKKAR					
Prim.cons.	ELPGAQYQYKKDEVWEERKPYKTLQAQAEKSKNKKKQRKGPHRKSQTLQFDEQTLKKAR					

	370	380	390	400	410	420
BMP3_HUMAN	RKQWIEPRNCARRYLKVDFADIGWSEWIISPKSFDAYYCSGACQFMPKSLKPSNHATIQ					
sequence1	RKQWIEPRNCARRYLKVDFADIGWSEWIISPKSFDAYYCSGACQFMPK-----					
sequence2	RKQWIEPRNCARRYLKVDFADIGWSEWIISPKSFDAYYCSGACQFMPK-----					
Consensus	RKQWIEPRNCARRYLKVDFADIGWSEWIISPKSFDAYYCSGACQFMPK					
Prim.cons.	RKQWIEPRNCARRYLKVDFADIGWSEWIISPKSFDAYYCSGACQFMPKSLKPSNHATIQ					

	430	440	450	460	470
BMP3_HUMAN	SIVRAVGVPGIPEPCCVPEKMSSLSILFFDENKNVVLKVYPNMTVESCACR				
sequence1	-----				
sequence2	-----				
Consensus					
Prim.cons.	SIVRAVGVPGIPEPCCVPEKMSSLSILFFDENKNVVLKVYPNMTVESCACR				

Alignment length : 472

Residues conserved for 50 % and less than 90 % (lower-case letters) : 7 is 1.48 %

V conserved positions (!) : 0 is 0.00 %

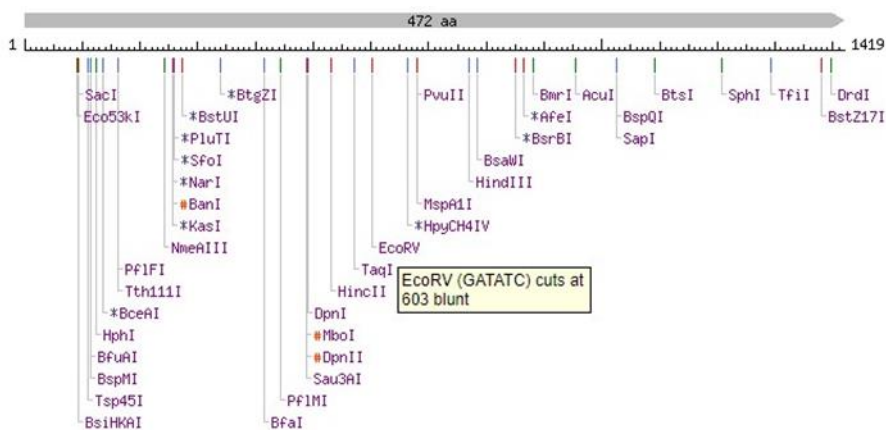
Στην επόμενη εργασία μελετώνται τα ένζυμα περιορισμού αλληλουχίας του DNA της πρωτεΐνης BMP 3. Μπαίνουμε στη διεύθυνση <http://nc2.neb.com/NEBcutter2/> του εργαλείου **NEBcutter V2.0**, όπου εμφανίζεται ένα παράθυρο στο οποίο εισάγουμε τη νουκλεοτιδική αλληλουχία της πρωτεΐνης από τη βάση δεδομένων **CCDS**, κάτω από το παράθυρο επιλέγουμε το NEB enzymes και πατάμε Submit :

Στη συνέχεια, επιλέγουμε μερικά ένζυμα από το εργαλείο **NEBcutter V2.0** με βάση τον αριθμό θέσεων κοπής της DNA ακολουθίας: **BamHI**, **EcoRI**, **KpnI**, **SmaI** – δεν κόβουν **καμία** φορά την DNA αλληλουχία : **EcoRV**, **HindIII**, **KasI** – κόβουν την αλληλουχία μόνο **1 φορά** : **MseI**, **RsaI** – κόβουν την αλληλουχία **2 φορές** : **HinfI**, **SfaNI** – κόβουν την αλληλουχία 3 φορές. Το εργαλείο αυτό μας δίνει τα εξής αποτελέσματα:

EcoRV: κόβει στη θέση 603 της νουκλεοτιδικής αλληλουχίας και δίνει τυφλά άκρα στο DNA. Οι θέσεις περιορισμού του ενζύμου είναι:

5' GAT|ATC 3'

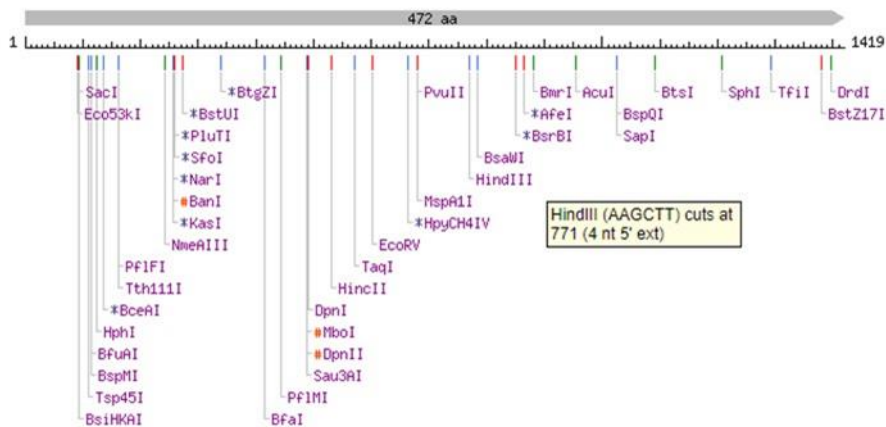
3' CTA|TAG 5'



HindIII: κόβει στη θέση 771 της νουκλεοτιδικής αλληλουχίας και δίνει 5' προεξέρχοντα άκρα στο DNA. Οι θέσεις περιορισμού του ενζύμου είναι:

5' A|AGCTT 3'

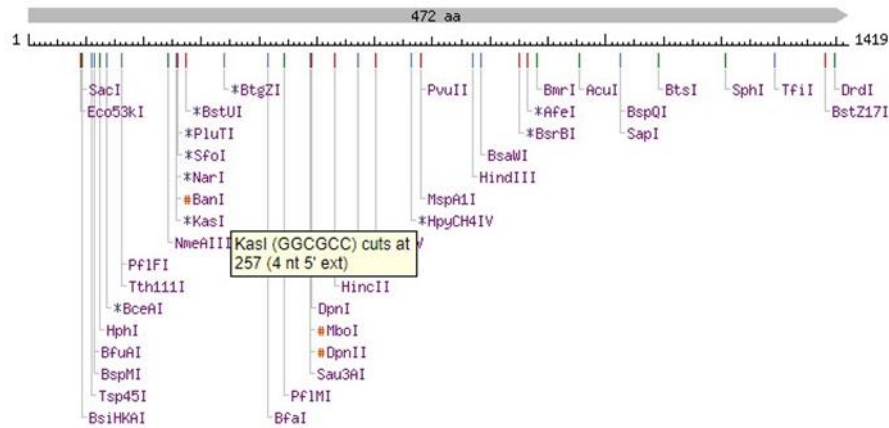
3' TTCGA|A 5'



KasI: κόβει στη θέση 257 της νουκλεοτιδικής αλληλουχίας και δίνει 5' προεξέρχοντα άκρα στο DNA. Οι θέσεις περιορισμού του ενζύμου είναι:

5' G|GCGCC 3'

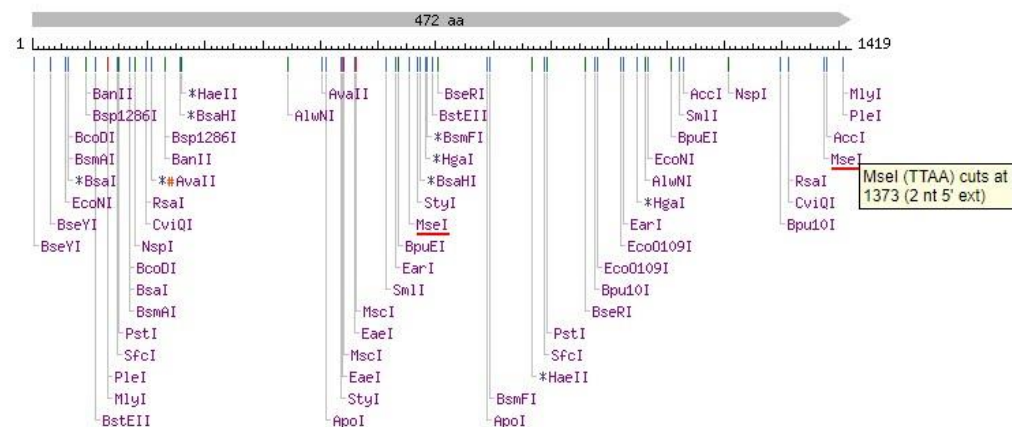
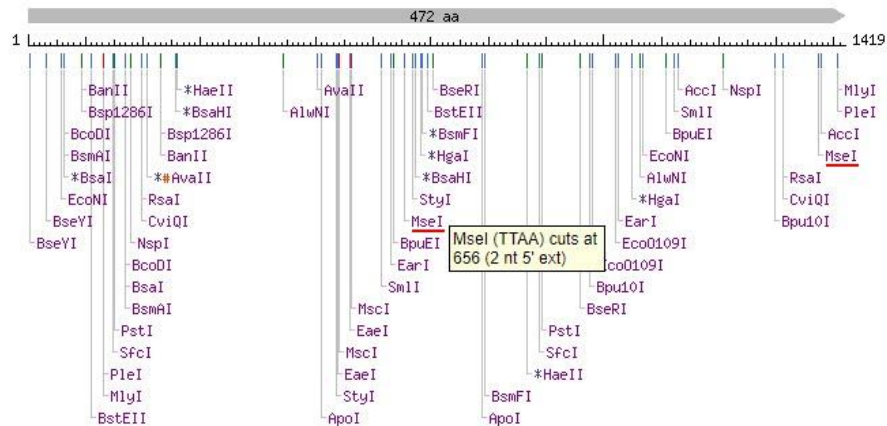
3' CCGCG|G 5'



MseI: κόβει στις θέσεις 656 και 1373 της νουκλεοτιδικής αλληλουχίας και δίνει 5' προεξέρχοντα άκρα στο DNA. Οι θέσεις περιορισμού του ενζύμου είναι:

5' T|TAA 3'

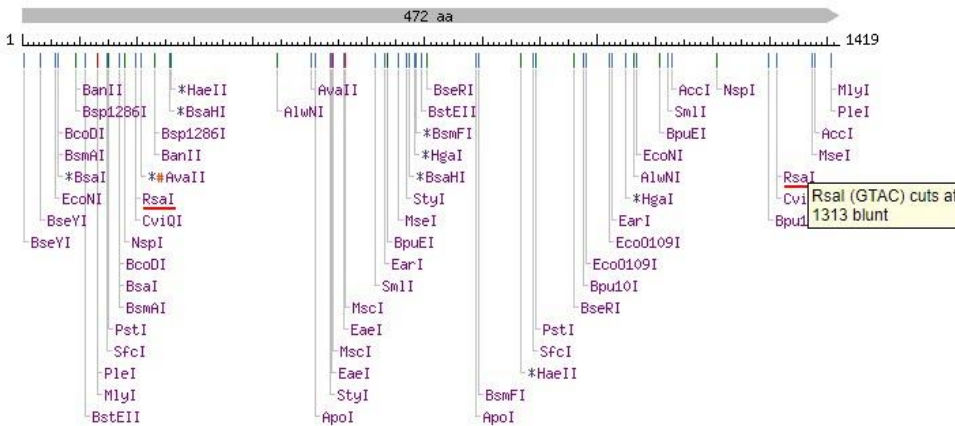
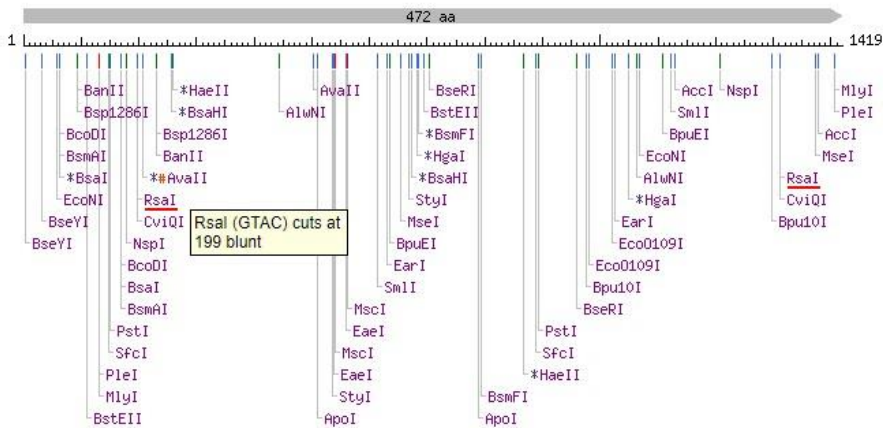
3' AAT|T 5'



RsaI: κόβει στις θέσεις 199 και 1313 της νουκλεοτιδικής αλληλουχίας και δίνει τυφλά άκρα στο DNA. Οι θέσεις περιορισμού του ενζύμου είναι:

5' GT|AC 3'

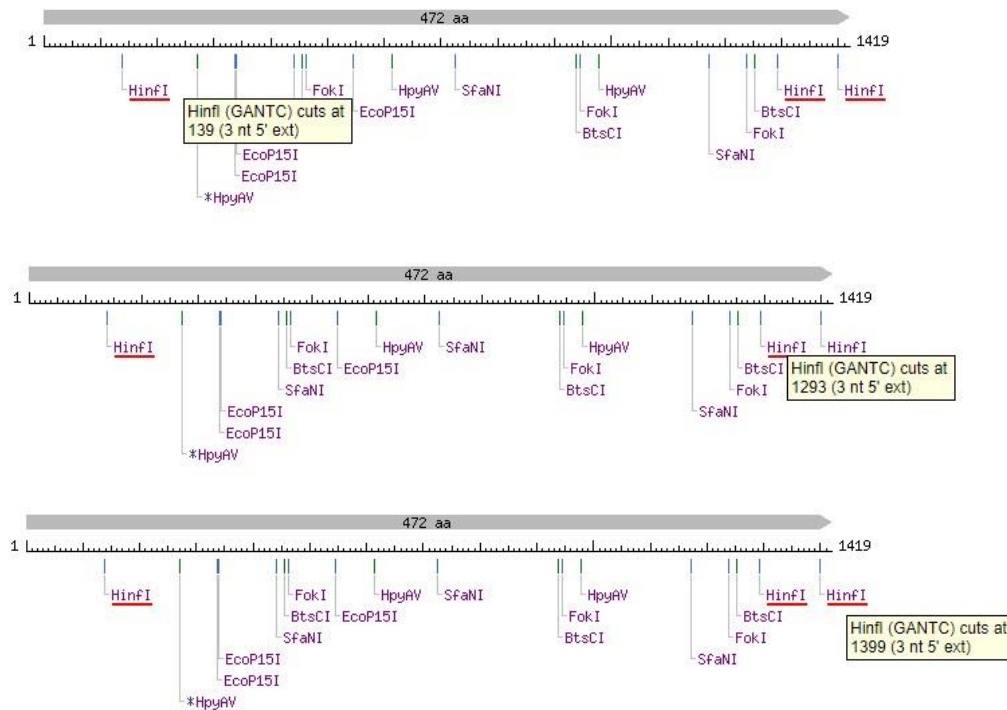
3' CA|TG 5'



HinfI: κόβει στις θέσεις 139, 1293 και 1399 της νουκλεοτιδικής αλληλουχίας και δίνει 5' προεξέρχοντα άκρα στο DNA. Οι θέσεις περιορισμού του ενζύμου είναι:

5' G|ANTC 3'

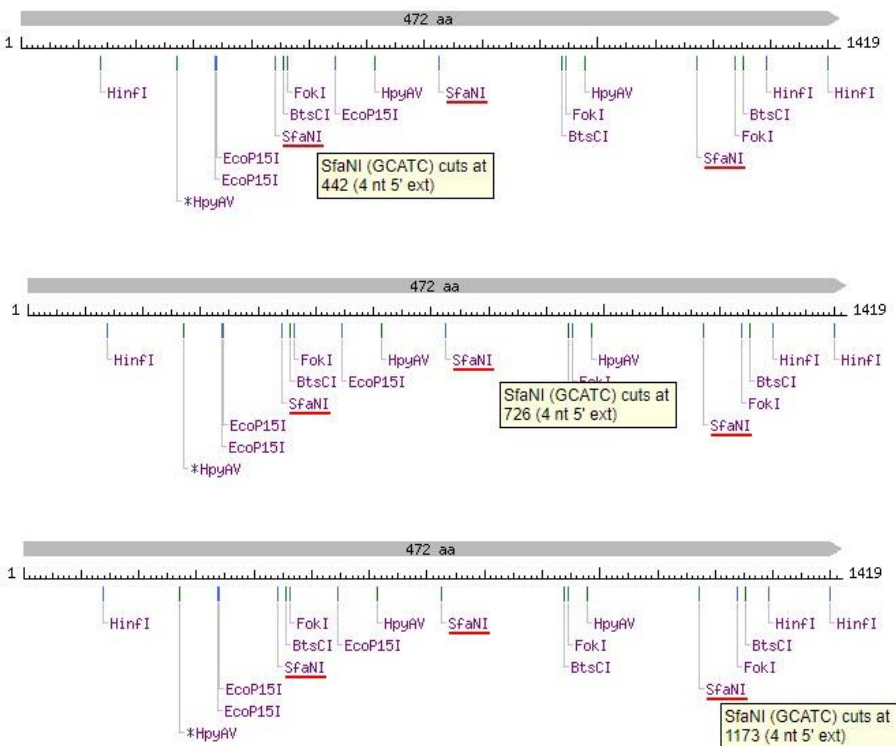
3' CTNA|G 5'



SfaNI: κόβει στις θέσεις 442, 726 και 1173 της νουκλεοτιδικής αλληλουχίας και δίνει 5' προεξέρχοντα άκρα στο DNA. Οι θέσεις περιορισμού του ενζύμου είναι:

5' GCATC(N)5| 3'

3' CGTAG(N)9| 5'



Ακολουθώντας, μπαίνουμε στην ιστοσελίδα <http://molbiol-tools.ca/PCR.htm>, επιλέγουμε το εργαλείο Sequence Extractor και στο πάνω παράθυρο εισαγάγουμε τη νουκλεοτιδική ακολουθία της πρωτεΐνης από τη βάση δεδομένων CCDS (ακολουθία DNA), ενώ στο κάτω παράθυρο – τις αλληλουχίες που αναγνωρίζει το συγκεκριμένο ένζυμο περιορισμού, ακολουθούμενες από κενό και μετά το όνομα του ενζύμου (επιλέξαμε τα ένζυμα που παρουσιάστηκαν πριν).

Παρακάτω παρουσιάζονται τα αποτελέσματα ανάλυσης ακολουθίας του DNA, όπου φαίνονται οι θέσεις κοπής από τα ένζυμα περιορισμού που επιλέξαμε.

Sequence Extractor - Results

Length: 1419 bp.

- Mouse over items on the map to view additional information. Click on two restriction sites or two compatible PCR primers to generate a product (requires JavaScript).

RestStart

PCRStart

```

                                     CfoI
                                     HhaI
M A G A S R L L F L W L G C F C V S L A Q G E R P K P
ATGGCTGGGGCGAGCAGGCTGCTCTTTCTGTGGCTGGGCTGCTTCTGCGTGAGCCTGGCGCAGGGAGAGAGACCGAAGCC
^10      ^20      ^30      ^40      ^50      ^60      ^70
TACCGACCCCGCTCGTCCGACGAGAAAGACACCGACCCGACGAAGACGCACTCGGACCGCGTCCCTCTCTCTGGCTTCGG
-----
                                     5' GANTC 3' HinfI
                                     3' CTNAG 5' HinfI
                                     HinfI
      AluI
HpaII
MspI
P F P E L R K A V P G D R T A G G G P D S E L Q P Q D
ACCTTTCCCGGAGCTCCGCAAAAGCTGTGCGCAGGTGACCGCACGGCAGGTGGTGGCCCGGACTCCGAGCTGCAGCCGCAAG
^90      ^100     ^110     ^120     ^130     ^140     ^150
TGGAAAGGGCCTCGAGGCGTTTCGACACGGTCCACTGGCGTGCCGTCCACCACCGGGCCTGAGGCTCGACGTCGGCGGTTG
-----
      5' GTAC 3' RsaI
      3' CATG 5' RsaI
                                     HpaII
                                     MspI
      RsaI
      HaeIII
      HpaII
      MspI
K V S E H M L R L Y D R Y S T V Q A A R T P G S L E
ACAAGGTCTCTGAACACATGCTGCGGCTCTATGACAGGTACAGCACGGTCCAGGCGGCCCGGACACCGGGCTCCCTGGAG
^170     ^180     ^190     ^200     ^210     ^220     ^230
TGTTCCAGAGACTTGTGTACGACGCCGAGATACTGTCCATGTCTGTGCCAGGTCCGCCGGGCCTGTGGCCCGAGGGACCTC
-----
      5' GGCGCC 3' KasI
      3' CCGCGG 5' KasI
      CfoI
      HhaI
      BstUI
      CfoI
      NarI
      HhaI
      AluI
G G S Q P W R P R L L R E G N T V R S F R A A A A E T
GGAGGCTCGCAGCCCTGGCGCCCTCGGCTCCTGCGCGAAGGCAACACGGTTCGCAGCTTTCGGGCGGCAGCAGCAGAAAC
^250     ^260     ^270     ^280     ^290     ^300     ^310
CCTCCGAGCGTCGGGACCGCGGGAGCCGAGGACGCGCTTCCGTTGTGCCAAGCGTCGAAAGCCCGCGTCGTCGTCCTTTC
-----
      MboII
L E R K G L Y I F N L T S L T K S E N I L S A T L Y F
TCTTGAAAGAAAAGGACTGTATATCTTCAATCTGACATCGCTAACCAGTCTGAAAACATTTTGTCTGCCACACTGTATT
^330     ^340     ^350     ^360     ^370     ^380     ^390
AGAACTTCTTTTCTGACATATAGAAGTTAGACTGTAGCGATTGGTTCAGACTTTTGTAAAACAGACGGTGTGACATAA
-----
```


3' CTACG 5' N 5NNNN SfaNI

AluI DdeI DdeI

C I G E L G N I S L S C P V S G G C S H H A Q R K H

TCTGTATTGGAGAGCTAGGAAACATCAGCCTGAGTTGTCCAGTGTCTGGAGGATGCTCCCATCATGCTCAGAGGAAACAC

^410 ^420 ^430 ^440 ^450 ^460 ^470

AGACATAACCTCTCGATCCTTTGTAGTCGGACTCAACAGGTCACAGACCTCTACGAGGTTAGTACGAGTCTCCTTTGTG

MboI NdeII HincII HaeIII

I Q I D L S A W T L K F S R N Q S Q L L G H L S V D M

ATTGAGATTGATCTTTCTGCATGGACCCCTCAAATTCAGCAGAAACCAAGTCAACTCCTTGGCCATCTGTCTAGTGGATAT

^490 ^500 ^510 ^520 ^530 ^540 ^550

TAAGTCTAACTAGAAAGACGTACCTGGGAGTTTAAGTCGTCTTTGGTTTCAGTTGAGGAACCGGTAGACAGTCACCTATA

5' GATATC 3' EcoRV

3' CTATAG 5' EcoRV

HaeIII TaqI EcoRV HaeIII

A K S H R D I M S W L S K D I T Q L L R K A K E N E E

GGCCAAATCTCATCGAGATATTATGTCTGGCTGTCTAAGATATCACTCAACTCTTGAGGAAGGCCAAGAAAATGAAG

^570 ^580 ^590 ^600 ^610 ^620 ^630

CCGGTTTAGAGTAGCTCTATAATACAGGACCGACAGATTTCTATAGTGAAGTGAAGAACTCCTTCCGGTTTCTTTACTTC

5' TTAA 3' MseI

3' AATT 5' MseI

MseI AluI PvuII

MboII

F L I G F N I T S K G R Q L P K R R L P F P E P Y I

AGTTCTCATAGGATTTAACATTACGTCCAGGGACGCCAGCTGCCAAAGAGGAGTTACCTTTTCCAGAGCCTTATATC

^650 ^660 ^670 ^680 ^690 ^700 ^710

TCAAGGAGTATCCTAAATTGTAATGCAGGTCCCTGCGGTGCGCGTTTCTCTCCAATGAAAAGGTCTCGGAATATAG

5' AAGCTT 3' HindIII

3' CTCGAA 5' HindIII

AluI HpaII

AluI HindIII MspI

L V Y A N D A A I S E P E S V V S S L Q G H R N F P T

TTGGTATATGCCAATGATGCCGCCATTCTGAGCCAGAAAGTGTGGTATCAAGCTTACAGGGACACCGGAATTTTCCAC

^730 ^740 ^750 ^760 ^770 ^780 ^790

AACCATATACGGTTACTACGGCGGTAAAGACTCGGTCTTTACACCATAGTTCGAATGTCCCTGTGGCCTTAAAAGGGTG

MboII CfoI HhaI

AluI

G T V P K W D S H I R A A L S I E R R K K R S T G V L

TGGAAGTGTTCCTAAATGGGATAGCCACATCAGAGCTGCCCTTTCCATTGAGCGGAGGAAGAAGCGCTCTACTGGGGTCT

^810 ^820 ^830 ^840 ^850 ^860 ^870

ACCTTGACAAGGGTTTACCCTATCGGTGTAGTCTCGACGGGAAGGTAAGTTCGCTCTCTTCGCGAGATGACCCAGA

PstI AluI

L P L Q N N E L P G A E Y Q Y K K D E V W E E R K P

TGCTGCCTCTGCAGAACACGAGCTTCTGGGGCAGAATACCAGTATAAAAAGGATGAGGTGTGGGAGGAGAGAAAGCCT

^890 ^900 ^910 ^920 ^930 ^940 ^950

ACGACGGAGACGTCTTGTGCTCGAAGGACCCGCTTATGGTCATATTTTCTCTACTCCACACCCTCCTCTCTTTTCGGA

HaeIII DdeI HaeIII

Y K T L Q A Q A P E K S K N K K K Q R K G P H R K S Q

TACAAGACCTTCAGGCTCAGGCCCTGAAAAGAGTAAGAATAAAAAAGAAACAGAGAAAGGGGCCTCATCGGAAGAGCCA

^970 ^980 ^990 ^1000 ^1010 ^1020 ^1030

ATGTTCTGGGAAGTCCGAGTCCGGGACTTTTCTCATTCTTATTTTCTTTGTCTCTTTCCCGGAGTAGCCTTCTCGGT

MboII CfoI
 HhaI
 T L Q F D E Q T L K K A R R K Q W I E P R N C A R R Y
 GACGCTCCAATTTGATGAGCAGACCCTGAAAAAGGCAAGGAGAAAGCAGTGGATTGAACCTCGGAATTGCGCCAGGAGAT
 ^1050 ^1060 ^1070 ^1080 ^1090 ^1100 ^1110
 CTGCGAGGTTAAACTACTCGTCTGGGACTTTTCCGTTCCTCTTCGTCACCTAACTTGGAGCCTTAACGCGGTCTCTCTA

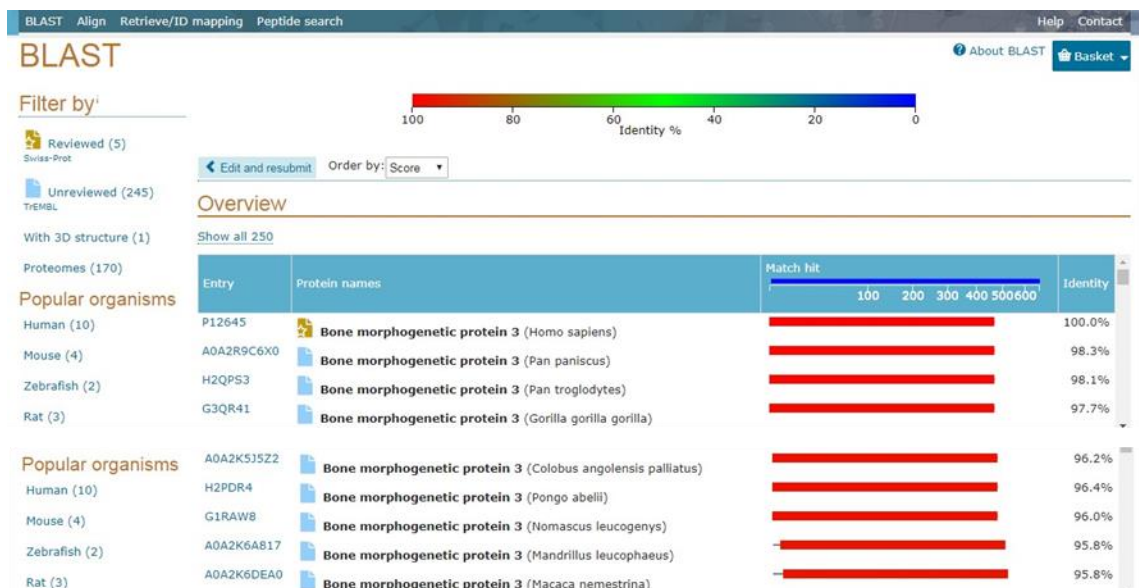
 L K V D F A D I G W S E W I I S P K S F D A Y Y C S
 ACCTCAAGGTAGACTTTGCAGATATTGGCTGGAGTGAATGGATTATCTCCCCAAGTCCCTTTGATGCCTATTATTGCTCT
 ^1130 ^1140 ^1150 ^1160 ^1170 ^1180 ^1190
 TGGAGTTCCTCTGAAACGCTCTATAACCGACCTCACTTACCTAATAGAGGGGGTTTCAGGAACTACGGATAATAACGAGA

 faNI SphI AluI
 G A C Q F P M P K S L K P S N H A T I Q S I V R A V G
 GGAGCATGCCAGTTCCCATGCAAGTCTTTGAAGCATCAATCATGCTACCATCCAGAGTATAGTGAGAGCTGTGGG
 ^1210 ^1220 ^1230 ^1240 ^1250 ^1260 ^1270
 CCTCGTACGGTCAAGGGGTACGGTTTCAGAACTTCGGTAGTTAGTACGATGGTAGGTCTCATATCACTCTCGACACCC

 5' GATC 3' HinfI 5' GTAC 3' RsaI
 3' CTNAG 5' HinfI 3' CATG 5' RsaI
 DdeI
 HinfI RsaI DdeI
 V V P G I P E P C C V P E K M S S L S I L F F D E N K
 GGTCTGTTCTCGGATTCCTGAGCCTTGCTGTACCCAGAAAGATGTCCTCACTCAGTATTTATTCTTTGATGAAAAATA
 ^1290 ^1300 ^1310 ^1320 ^1330 ^1340 ^1350
 CCAGCAAGGACCCTAAGGACTCGGAACGACACATGGTCTTTTCTACAGGAGTGAGTCATAAAATAAGAACTACTTTTAT

 RestEnd
 PCREnd
 5' TTAA 3' MseI 5' GATC 3' HinfI
 3' AATT 5' MseI 3' CTNAG 5' HinfI
 CfoI
 MseI HinfI HhaI
 N V V L K V Y P N M T V E S C A C R Z
 AGAATGTAGTGCTTAAAGTATACCCCTAACATGACAGTAGAGTCTTTCGCTTGCAGATAA
 ^1370 ^1380 ^1390 ^1400 ^1410
 TCTTACATCAGCAATTTCATATGGGATTGTACTGTCTCAGAACGCGAACGTCTATT

Εάν θέλουμε να γίνει σύγκριση της πρωτοταγούς ακολουθίας της P12645 με άλλες όμοιες ή ομόλογες σε διαφορετικά ποσοστά κάθε φορά πρωτεΐνες, επιλέγουμε την ένδειξη BLAST στο πάνω αριστερό μέρος της σελίδας της πρωτεΐνης <https://www.uniprot.org/uniprot/P12645>. Παρακάτω παρουσιάζονται τα αποτελέσματα σύγκρισης με μία συγκεκριμένη σειρά αρχίζοντας από την πρωτεΐνη με την υψηλότερη ομολογία / ομοιότητα (μέχρι και την πρωτεΐνη με ποσοστό ομοιότητας 87,4%).



Proteomes (170)	A0A2K5KT92	Bone morphogenetic protein 3 (Cercopithecus atys)	95.3%
Popular organisms	A0A096NF75	Bone morphogenetic protein 3 (Papio anubis)	95.3%
Human (10)	A0A0D9QWS2	Bone morphogenetic protein 3 (Chlorocebus sabaeus)	95.1%
Mouse (4)	A0A2K6L6F4	Bone morphogenetic protein 3 (Rhinopithecus bieti)	95.1%
Zebrafish (2)	A0A2K6QIG5	Bone morphogenetic protein 3 (Rhinopithecus roxellana)	94.9%
Rat (3)	A0A2K5F957	Bone morphogenetic protein 3 (Aotus nancymaae)	94.9%
Proteomes (170)	A0A2K5RTJ7	Bone morphogenetic protein 3 (Cebus capucinus imitator)	94.1%
Popular organisms	A0A2R8N1P9	Bone morphogenetic protein 3 (Callithrix jacchus)	93.6%
Human (10)	A0A2K6UVL2	Bone morphogenetic protein 3 (Saimiri boliviensis boliviensis...)	93.0%
Mouse (4)	L5L066	Bone morphogenetic protein 3 (Pteropus alecto)	87.9%
Zebrafish (2)	A0A452SLB5	Bone morphogenetic protein 3 (Ursus americanus)	87.2%
Rat (3)	D2GUF8	Uncharacterized protein (Alluropoda melanoleuca)	87.4%

Τέλος, ανοίγουμε στο πρόγραμμα **RasMol** το αρχείο PDB της πρωτεΐνης, το οποίο κατεβάζουμε από την ιστοσελίδα της βάσης δεδομένων Protein Data Bank in Europe (<https://www.ebi.ac.uk/pdbe/entry/pdb/2QCQ>). Στο παράθυρο με τις εντολές σε μορφή κειμένου, πληκτρολογούμε τις παρακάτω εντολές, ώστε να αλλάξουμε το **background** σε «λευκό», το **αμινοξύ κυστεΐνη** της πρωτεΐνης να χρωματίσουμε σε «μωβ» και να φαίνεται το συγκεκριμένο αμινοξύ μόνο σε μορφή **ribbons**, ενώ ολόκληρη η πρωτεΐνη να είναι σε μορφή **wireframe**:

- **set background white**
- **select cys**
- **purple ribbons**

```

RasMol Command Line
RasMol Molecular Renderer
Roger Sayle, August 1995
Copyright (C) Roger Sayle 1992-1999
Version 2.7.5.2 May 2011
Copyright (C) Herbert J. Bernstein 1998-2011
*** See "help notice" for further notices ***
RasMol> set background white
RasMol> select cys
84 atoms selected!
RasMol> color purple
RasMol> ribbons
RasMol>

```

