



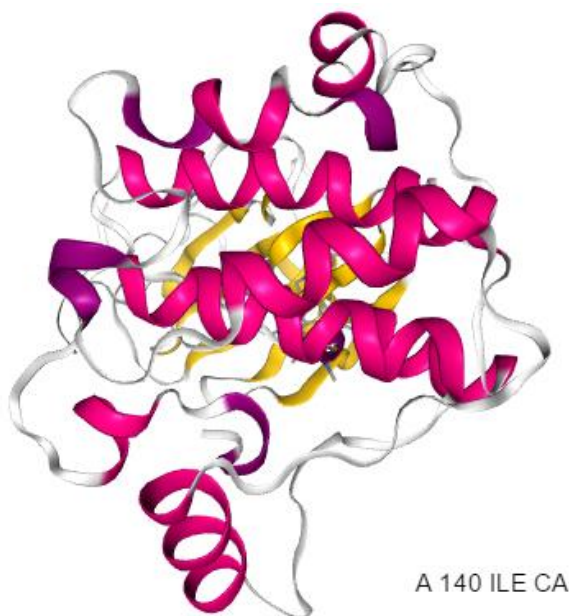
ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΧΗΜΕΙΑΣ

Άσκηση στο μάθημα της Βιοπληροφορικής

Καθηγήτρια Χολή-Παπαδοπούλου Θεοδώρα



Τατιάνα Καράμοβα

ΑΕΜ: 8503

Άσκηση Βιοπληροφορικής

Στην άσκηση αυτή εξετάζεται η πρωτεΐνη με την ονομασία: **Q15418**

Από Βάσεις δεδομένων πρωτεϊνικών ακολουθιών και βάσεις για την ανάλυση ακολουθιών είναι δυνατόν να ληφθούν οι παρακάτω πληροφορίες:

UniProtKB Από τη βάση δεδομένων UniProtKB παίρνουμε τις λειτουργικές πληροφορίες για τη συγκεκριμένη ριβοσωμική πρωτεΐνη:

The UniProt Knowledgebase (UniProtKB) is produced by the UniProt consortium and is the central hub for the collection of functional information on proteins with accurate, consistent and rich annotation. It consists of: UniProtKB/Swiss-Prot (manually-annotated records and curator-evaluated computational analysis) and UniProtKB/TrEMBL (computationally analyzed records awaiting manual annotation)

Η Q15418 είναι μια Ριβοσωμική Πρωτεΐνη που προέρχεται από το γένος **RPS6KA1**, από ανθρώπινο οργανισμό (*Homo Sapiens*) και είναι μια **Κινάση** που σημαίνει ότι φωσφορυλιώνει πρωτεΐνες ρυθμίζοντας ορισμένες λειτουργίες του κυττάρου. Επίσης, η συγκεκριμένη πρωτεΐνη μπορεί να βρεθεί μέσα στο κύτταρο στον πυρήνα (nucleus) και στο κυτταρόπλασμα (cytoplasm).

Πιο συγκεκριμένα εκτελεί τις παρακάτω λειτουργίες:

Serine/threonine-protein kinase that acts downstream of ERK (MAPK1/ERK2 and MAPK3/ERK1) signaling and mediates mitogenic and stress-induced activation of the transcription factors CREB1, ETV1/ER81 and NR4A1/NUR77, regulates translation through RPS6 and EIF4B phosphorylation, and mediates cellular proliferation, survival, and differentiation by modulating mTOR signaling and repressing pro-apoptotic function of BAD and DAPK1. In fibroblast, is required for EGF-stimulated phosphorylation of CREB1, which results in the subsequent transcriptional activation of several immediate-early genes. In response to mitogenic stimulation (EGF and PMA), phosphorylates and activates NR4A1/NUR77 and ETV1/ER81 transcription factors and the cofactor CREBBP. Upon insulin-derived signal, acts indirectly on the transcription regulation of several genes by phosphorylating GSK3B at 'Ser-9' and inhibiting its activity. Phosphorylates RPS6 in response to serum or EGF via an mTOR-independent mechanism and promotes translation initiation by facilitating assembly of the pre-initiation complex. In response to insulin, phosphorylates EIF4B, enhancing EIF4B affinity for the EIF3 complex and stimulating cap-dependent translation. Is involved in the mTOR nutrient-sensing pathway by directly phosphorylating TSC2 at 'Ser-1798', which potently inhibits TSC2 ability to suppress mTOR signaling, and mediates phosphorylation of RPTOR, which regulates mTORC1 activity and may promote rapamycin-sensitive signaling independently of the

Τατιάνα Καράμοβα

A.E.M.: 8503

PI3K/AKT pathway. Mediates cell survival by phosphorylating the pro-apoptotic proteins BAD and DAPK1 and suppressing their pro-apoptotic function. Promotes the survival of hepatic stellate cells by phosphorylating CEBPB in response to the hepatotoxin carbon tetrachloride (CCl₄). Mediates induction of hepatocyte proliferation by TGFA through phosphorylation of CEBPB (By similarity). Is involved in cell cycle regulation by phosphorylating the CDK inhibitor CDKN1B, which promotes CDKN1B association with 14-3-3 proteins and prevents its translocation to the nucleus and inhibition of G1 progression. Phosphorylates EPHA2 at 'Ser-897', the RPS6KA-EPHA2 signaling pathway controls cell migration.

Παρόμοιες πρωτεΐνες με τη συγκεκριμένη είναι και αυτές που προέρχονται από αρουραίο **Q63531** (KS6A1_RAT) και από ποντίκι **P18653** (KS6A1_MOUSE)

Η **καταλυτική δραστηριότητα** της πρωτεΐνης φαίνεται από την παρακάτω αντίδραση:



Για να είναι καταλυτικά ενεργή η πρωτεΐνη χρειάζεται έναν ακόμα παράγοντα μη πρωτεϊνικής προέλευσης, είτε οργανικό είτε ανόργανο. Στη συγκεκριμένη περίπτωση ο παράγοντας αυτός είναι το κατιόν μαγνησίου **Mg²⁺**.

Επεκτάσεις της λειτουργίας της συγκεκριμένης πρωτεΐνης είναι και οι εξής έννοιες:

Sites:

Sites					
Feature key	Position(s)	Description	Actions	Graphical view	Length
Binding site ⁱ	94	ATP PROSITE-ProRule annotation			1
Active site ⁱ	187	Proton acceptor By similarity			1
Binding site ⁱ	447	ATP PROSITE-ProRule annotation			1
Active site ⁱ	535	Proton acceptor By similarity			1

Όπου ενδεικτικά από το εργαλείο BLAST (UniProt) παίρνουμε τις πληροφορίες:

Binding site: This subsection of the 'Function' section describes the interaction between a single amino acid and another chemical entity. Priority is given to the annotation of physiological ligands.

Entry & position(s)	Q15418 [94 - 94]
Description	Ribosomal protein S6 kinase alpha-1, Homo sapiens
Feature key	Binding site

Τατιάνα Καράμοβα

A.E.M.: 8503

Feature identifier

```

      10      20      30      40      50
MPLAQLKEPW PLMELVPLDP ENGQTSGEAA GLQPSKDEGV LKEISITHHV
      60      70      80      90     100
KAGSEKADPS HFELLKVLGQ GSFGKVFLVR KVTRPD SGHL YAMKVLKKAT
      110     120     130     140     150
LKVRDRVRTK MERDILADVN HPFVVKLHYA FQTEGKLYLI LDFLRGGDLF
      160     170     180     190     200
TRLSKEVMFT EEDVKFYLAELALGLDHLHSLGIIYRDLKPENILLDEEGH
      210     220     230     240     250
IKLTDGFLSK EAIDHEKKAYSFCGTVEYMAPEVVNRQGHSHSADWWSYGV
      260     270     280     290     300
LMFEMLTGSL PFQGKDRKETMTLILKAKLGMPOFLSTEAQSLLRALFKRN
      310     320     330     340     350
PANRLGSGPD GAEEIKRHVFYSTIDWNKLYRREIKPPFKP AVAQPD DTFY
      360     370     380     390     400
FDTEFTSRTPKDSPGIPPSAGAHQLFRGFSEFVATGLMEDDGKPRAPQAPL
      410     420     430     440     450

HSVVQQLHGK NLVFS DGYVVKETIGVGSYS ECKRCVHKAT NMEYAVKVID
      460     470     480     490     500
KSKRDPSEEI EILLRYGQHPNIITLKDVYD DGKHVYLVTE LMRGGELLDK
      510     520     530     540     550
ILRQKFFSER EASFVLHTIGKTVEYLSHSGVVHRDLKPSN ILYVDESNGP
      560     570     580     590     600
ECLRICDFGF AKQLRAENGL LMTPCYTANFVAPEVLKRQGYDEGCDIWSL
      610     620     630     640     650
GILLYTMLAG YTPFANGPSDTPPEILTRISGKFTLSGGN WNTVSETAKD
      660     670     680     690     700
LVSKMLHVDHQRITAKQVLQHPWVTQKDKLPQSQLSHQDLQLVKGAMAA
      710     720     730
TYSALNSSKP TPQLKPIESSILAQRVRVKLPSTTL

```

Active Site: This subsection of the 'Function' section is used for enzymes and indicates the residues directly involved in catalysis.

Entry & position(s) [Q15418](#)[187 - 187]

Description Ribosomal protein S6 kinase alpha-1, Homo sapiens

Feature key Active site

Feature identifier

```

      10      20      30      40      50
MPLAQLKEPW PLMELVPLDP ENGQTSGEAA GLQPSKDEGV LKEISITHHV
      60      70      80      90     100
KAGSEKADPS HFELLKVLGQ GSFGKVFLVR KVTRPD SGHL YAMKVLKKAT
      110     120     130     140     150
LKVRDRVRTK MERDILADVN HPFVVKLHYA FQTEGKLYLI LDFLRGGDLF
      160     170     180     190     200
TRLSKEVMFT EEDVKFYLAELALGLDHLHSLGIIYRDLKPENILLDEEGH
      210     220     230     240     250
IKLTDGFLSK EAIDHEKKAYSFCGTVEYMAPEVVNRQGHSHSADWWSYGV
      260     270     280     290     300
LMFEMLTGSL PFQGKDRKETMTLILKAKLGMPOFLSTEAQSLLRALFKRN
      310     320     330     340     350
PANRLGSGPD GAEEIKRHVFYSTIDWNKLYRREIKPPFKP AVAQPD DTFY

```

Τατιάνα Καράμοβα

A.E.M.: 8503


```

      360      370      380      390      400
FDTEFTSRTP KDSPGIPPSA GAHQLFGRFS FVATGLMEDD GKPRAPQAPL
      410      420      430      440      450
HSVVQQLHGK NLVFSDBGYVV KETIGVGSYS ECKRCVHKAT NMEYAVKVID
      460      470      480      490      500
KSKRDPSEEI EILLRYGQHP NIITLKDVYD DGKHVYLVTE LMRGGELLDK
      510      520      530      540      550
ILRQKFFSER EASFVLHTIG KTVEYLHSQG VVHRDLKPSN ILYVDESGNP
      560      570      580      590      600
ECLRICDFGF AKQLRAENGL LMTPCYTANF VAPEVLKRQG YDEGCDIWSL
      610      620      630      640      650
GILLYTMLAG YTPFANGPSD TPEEILTRIG SGKFTLSGGN WNTVSETAKD
      660      670      680      690      700
LVSKMLHVDP HQRLTAKQVL QHPWVTQKDK LPQSQLSHQD LQLVKGAMAA
      710      720      730
TYSALNSSKP TPQLKPIESS ILAQRRVRKL PSTTL

```

Regions:

Regions

Feature key	Position(s)	Description	Actions	Graphical view	Length
Nucleotide binding [†]	68 – 76	ATP  PROSITE-ProRule annotation			9
Nucleotide binding [†]	424 – 432	ATP  PROSITE-ProRule annotation			9

Nucleotide binding: This subsection of the ‘Function’ section describes a region in the protein which binds nucleotide phosphates. It always involves more than one amino acid and includes all residues involved in nucleotide-binding.

Entry & position(s)	Q15418 [68 - 76]
Description	Ribosomal protein S6 kinase alpha-1, Homo sapiens
Feature key	Nucleotide binding
Feature identifier	

```

      10      20      30      40      50
MPLAQLKEPW PLMELVPLDP ENGQTSGEAA GLQPSKDEGV LKEISITHHV
      60      70      80      90     100
KAGSEKADPS HFELLKVL LGQ GSFGKV FLVR KVTRPD SGHL YAMKVLKKAT
      110     120     130     140     150
LKVRDRVRK MERDILADV HPFVVKLHYA FQTEGKLYLI LDFLRGGDLF
      160     170     180     190     200
TRLKSEVMFT EEDVKFYLA LALGLDHLHS LGIIYRDLKP ENILLDEEGH
      210     220     230     240     250
IKLTDFGLSK EAIDHEKKAY SFCGTVEYMA PEVVNRQGH HSADWWSYGV
      260     270     280     290     300
LMFEMLTGSL PFQGKDRKET MTLILKAKLG MPQFLSTEAQ SLLRALFKRN
      310     320     330     340     350
PANRLGSGPD GAEEIKRHVF YSTIDWNKLY RREIKPPFKP AVAQPD DTFY
      360     370     380     390     400
FDTEFTSRTP KDSPGIPPSA GAHQLFGRFS FVATGLMEDD GKPRAPQAPL
      410     420     430     440     450

```

Τατιάνα Καράμοβα

A.E.M.: 8503

```

HSVYQQLHGK NLVFSDBGYV KETIGVGSYS ECKRCVHKAT NMEYAVKVID
  460          470          480          490          500
KSKRDPSEEI EILLRYGQHP NIITLKDVID DGKHVYLVTE LMRGGELLDK
  510          520          530          540          550
ILRQKFFSER EASFVLHTIG KTVEYLHSQG VVHRDLKPSN ILYVDESGNP
  560          570          580          590          600
ECLRICDFGF AKQLRAENGL LMTPCYTANF VAPEVLKRQG YDEGCDIWSL
  610          620          630          640          650
GILLYTMLAG YTPFANGPSD TPEEILTRIG SGKFTLSGGN WNTVSETAKD
  660          670          680          690          700
LVSKMLHVDP HQRLTAKQVL QHPWVTQKDK LPQSQLSHQD LQLVKGAMAA
  710          720          730
TYSALNSSKP TPQLKPIESS ILAQRRVRKL PSTTL

```

EBI:

Από τις βάση δεδομένων **EBI** (European Bioinformatics Institute) λαμβάνουμε πληροφορίες για την Οντολογία γονιδίων (Genome Ontology) δηλαδή για το σύνολο των λειτουργικών γονιδιακών ιδιοτήτων.

ATP binding:

GO:0005524 ATP binding

QuickGO [Web Services](#) [Dataset](#) [Term Basket: 0](#) [i](#)

Term Information **Ancestor Chart** [Child Terms](#) [Protein Annotation](#) [Co-occurring Terms](#) [Change Log](#)

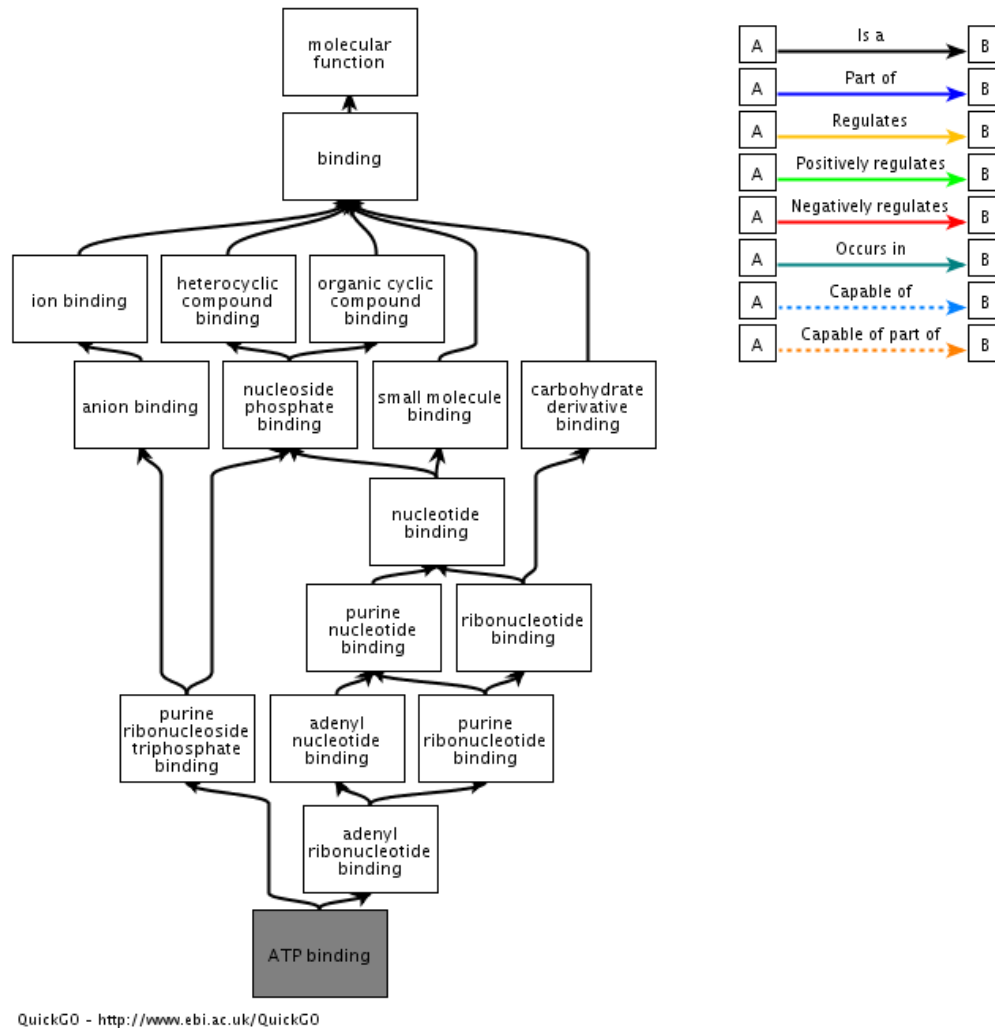
ID GO:0005524
Name ATP binding
Ontology Molecular Function
Definition Interacting selectively and non-covalently with ATP, adenosine 5'-triphosphate, a universally important coenzyme and enzyme regulator.
GONUTS [GO:0005524 Wiki Page](#)

Annotation Guidance **Cross-Ontology Relations** [Cross-references](#)

Usage of this term is subject to the following annotation guidelines:
Annotation Guidelines
[Binding term considerations](#) [Link](#)

Και από το Ancestor Chart παίρνουμε το σχήμα:

Τατιάνα Καράμοβα
Α.Ε.Μ.: 8503



Παρακάτω φαίνεται η πρωτοταγής ακολουθία της πολυπεπτιδικής αλυσίδας της συγκεκριμένης πρωτεΐνης:

```
>sp|Q15418|KS6A1_HUMAN Ribosomal protein S6 kinase alpha-1 OS=Homo sapiens GN=RPS6KA1 PE=1 SV=2
MPLAQLKEPWPLMELVPLDPENGQTSGEELQPSKDEGVLKEISITHHVKAGSEKADPS
HFELLKVLGQGSFGKVFVLRKVTRPDSGHLYAMKVLKKATLKVRDRVTRKMERDILADVN
HPFVVKLHYAFQTEGKLYLILDFLRGGDLFTRLKSKEVMFTEEDVKFYLAELALGLDHLHS
LGIIYRDLKPENILLDEEGHIKLTDFGLSKEAIDHEKKAYSFCGTVEYMAPEVVNRQGH
HSADWWSYGVLMEMLTGSLPFQGKDRKETMTLILKAKLGMPQFLSTEAQSLLRALFKRN
PANRLGSGPDGAEEIKRHVFYSTIDWNKLYRREIKPPFKPAVAQPDDTFYFDTEFTSRT
KDSPGIPPSAGAHQLFRGFSFVATGLMEDDGKPRAPQAPLHSSVQQHLHGKNLVFSDGYV
KETIGVGSYSECKRCVHKATNM EYAVKVIDKSKRDPSEEIEILLRYGQHPNIITLKDVID
DGKHVYLVTLMRGGELLDKILRQKFFSEREASFVLHTIGKTVEYLHSQGVVHRDLKPSN
ILYVDESGNPECLRICDFGFAKQLRAENGLLMTPCYTANFVAPEVLKRQGYDEGCDIWSL
GILLYTMLAGYTPFANGPSDTP E EILTRIGSGKFTLSGGNWNVTSETAKDLVSKMLHVD
HQRLLTAKQVLQHPVWTQKDKLPQS QLSHQDLQLVKGAMAATYSALNSSKPTPQLKPIESS
ILAQRVRVKLPSTTL
```

Από τα Εργαλεία (Tools) από τον ιστότοπο <http://www.expasy.org/proteomics> παίρνουμε ορισμένες πληροφορίες:

Τατιάνα Καράμοβα

A.E.M.: 8503

2ZIP-Server: Prediction of leucine zipper domains

The leucine zipper is a dimerisation domain occurring mostly in regulatory and thus in many oncogenic proteins. 2ZIP combines a standard coiled coil prediction algorithm with an approximate search for the characteristic leucine repeat. No further information from homologues is required for prediction. This approach improves significantly over existing methods, especially in that the coiled coil prediction turns out to be highly informative and avoids large numbers of false positives

Trying to run ncoils...

Command line options: ./ncoils -f -win 28

The amino acid sequence in /863059117.fasta is predicted NOT to contain a Leucine Zipper domain

(No coiled coil segments detected by ncoils)

Verbose:

Description line:

>sp|Q15418|KS6A1_HUMAN

1) number of potential LEUCINE ZIPPERS: 0

2) COILED COILS which do NOT correspond to a Leucine Zipper none

3) LEUCINE REPEATS which do NOT correspond to a Leucine Zipper none

Τελικά, από αυτό το εργαλείο συμπεραίνουμε ότι δεν υπάρχει η χαρακτηριστική συσπειρωμένη ακολουθία λυσίνης.

Big-PI: Predict GPI modification sites

Use of the prediction function for METAZOA

None potential GPI-modification site was found.

Among all positions checked, sequence position 719 had the best score.

Total Score.....: **-46.61**
(PValue = 2.039882e-01)

Components of the Score Function:

Profile Score.....	-4.84
Term 0 Contents and Windows of DE in Region [-11..1].....	0.00
Term 1 Hydrophilicity of N-terminal Region [-11..1].....	0.00
Term 2 Penalty for low Profile Score in Region [0..2].....	-4.00
Term 3 Volume Limitation [-1..+2].....	-0.94
Term 4 Volume Compensation (-1, 1, 2).....	-1.27
Term 5 Volume Compensation (-1, 2).....	-1.99
Term 6 Backbone Flexibility [-1..2].....	0.00
Term 7Propeptide Length.....	-4.00
Term 8 Hydrophilicity of Spacer Region [3..8].....	-0.02
Term 9 Volume Limitation [3..8].....	-1.24
Term 10 Penalty for charged AAs in Spacer Region [3..10]....	-0.00
Term 11 Backbone Flexibility [3..8].....	0.00

A.E.M.: 8503

Term 12	Penalty for low Profile Score in Region [10..end]...	0.00
Term 13	Hydrophobicity of Tail [10..end].....	-3.49
Term 14	Hydrophobicity of Tail [26..end].....	-0.82
Term 15	Even Distribution of Hydrophobicity [9..end].....	-12.00
Term 16	Penalty for polar Windows in Region [10..end].....	0.00
Term 17	Penalty for SGC-Windows in Region [10..end].....	0.00
Term 18	LVI Contents [10..end].....	-12.00
Term 19	Penalty for FYHW - Sections in Region [10..end].....	0.00
Term 20	Penalty for Windows with small Volume [10..end].....	0.00
Profile	independent Score.....	-37.77

CFSSP: Protein secondary structure prediction

This server predicts secondary structure of protein from the amino acid sequence. In this server, Chou & Fasman algorithm has been implemented.

Secondary Structure:

[illegible]

[illegible]

Τατιάνα Καράμοβα

A.E.M.: 8503

```

Struc 421
HHEEHCCCCCCHHHHHHHHHHHEEEHHHTHCCCCCHHHHEEEHCTCCCEEHHEEHCTHHEEHHH
H 490

* * * * *

*
Query 491
LMRGGELLDKILRQKFFSEREASFVLHTIGKTVEYLHSQGVVHRDLKPSNILYVDESGNPECLRICDFG
F 560
Helix 491 HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
HHHHHHH 560
Sheet 491 E EEEEEEEEE EEEEEEEEEEEEE EEE
EEEEEE 560
Turns 491 TT T TT T TT TT T T
560
Struc 491
HHHTHHHHHEEHHEEHHHHHHHHHHEEEEEEHEEHHEHCCCTHHHHHHHHHTHEEHHTCCCTCEEEEEEEH
H 560

* * * * *

*
Query 561
AKQLRAENGLIMTPCYTANFVAPEVLKRQGYDEGCDIWSLGILLYTMLAGYTPFANGPSDTPPEILTRI
G 630
Helix 561 HHHHHHHHHHH HHHHHHHHHHH HHHHHHHHHHH
HHHHHH 630
Sheet 561 EEEEEEEEE EEEEEEEEEEEEEEEEE
EEEE 630
Turns 561 TTTT TT T TT T T 630
561
Struc 561
HHHHHHHTHEEEEEEEEEHHHHHHHTTCCCTCEEEEEEEEEHHHEEECCCTTCTCCCHHEEEEC
C 630

* * * * *

*
Query 631
SGKFTLSGGNWNTVSETAKDLVSKMLHVDPHQRLTAKQVLQHPWVTQKDKLPQSQLSHQDLQLVKGAMA
A 700
Helix 631 HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 700
Sheet 631 EEEEE EEEEEEEEEEEEE EEEEE
EEEEEE 700
Turns 631 TT TTT TT TT TT T 700
631
Struc 631
TTCCCCCTTCCCCCHHHHHHHHEHHHEEHCCCHHHHHHEEHHEEEHHHTHEHETHHHHHHEEEHHHHHH
C 700

* * *

Query 701 TYSALNSSKPTPQLKPIESSILAQRRVRKLPSTTL 735
Helix 701 HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH
735
Sheet 701 EEEEEEEEEEEEE 735
Turns 701 T TTTT 735
Struc 701 CCCCCCTCTCCTHHHHHHHTHHHHHEEEEEEEEEEECC 735

Total Residues: H: 540 E: 294 T: 110
Percent: H: 73.5 E: 40.0 T: 15.0

```

ColorSeq: Color Protein Sequence

Τατιάνα Καράμοβα

A.E.M.: 8503

Εδώ τα χρωματισμένα αμινοξικά κατάλοιπα είναι σε κάθε μια από τις παρακάτω περιπτώσεις υδρόφοβα και υδρόφιλα αντίστοιχα:

Hydrophobic sites

10	20	30	40	50	60	70
MPLAQLKEPWPLME	LVPLDPENGQTS	GEEAGLQPSKDE	GVLKEISITHHV	KAGSEKADPSHF	ELLKVLGQ	
GSFGKVF	LVRKVTRPDS	GHLYAMKVLK	KATLKVRDRV	RTKMERDILAD	VNHPFVVKLHY	AFQTEGKLYLI
LDFLRGGDL	FTRLSKEVM	FTTEEDVKFY	LAELALGLD	HLHSLGIIYR	DLKPENILL	DEEGHIKLTDFGLSK
E	AIDHEKKAY	SFCGTVEYMA	PEVVNRQGH	SHSADWWSY	GVLMFEMLT	GS
LPFQ	GKDRKET	MTLILKAKLG				
MPQFL	STEAQSLL	RALFKRNPAN	RLGSGPDG	AEEIKRHF	YFSTIDWNK	LYRREIKPPFKPAVAQPD
FT	EFTSRTPKD	SPGIPPS	AGAHQLFR	GSFVATGL	MEDDGKPR	APQAPLH
SVVQQL	HGKNLV	FS	SDGYVV			
KETIG	VG	SYSECKRC	VHKATNMEY	AVKVIDK	SKRDPSEE	IEILLRYGQHPNII
TLKD	VYDDGKH	VYLVTE				
LMRGGEL	LDKILRQ	KFFSERE	ASFVLHT	IGKTVEY	LHSQGVV	HRDLKPSN
ILYV	DESGN	PECLRIC	DFGF			
AKQLRA	ENGLLMT	PCYTANF	VAPVLRQ	GYDEGCD	IWSLGILL	YTMLAGYTPFANGP
SDTP	EEILTRIG					
SGKFTL	SGGNWNT	VSETAKD	LVSKMLH	VDPHQRL	TAKQVLQHP	WVTQKDKLPQ
SQLSHQDL	QLVKG	AMAA				
TYSAL	NSSKPTQL	KPIESS	ILAQRR	VRKLP	STTL	

Total number of **ALIVMW** in sequence : 230*Hidrophilic sites*

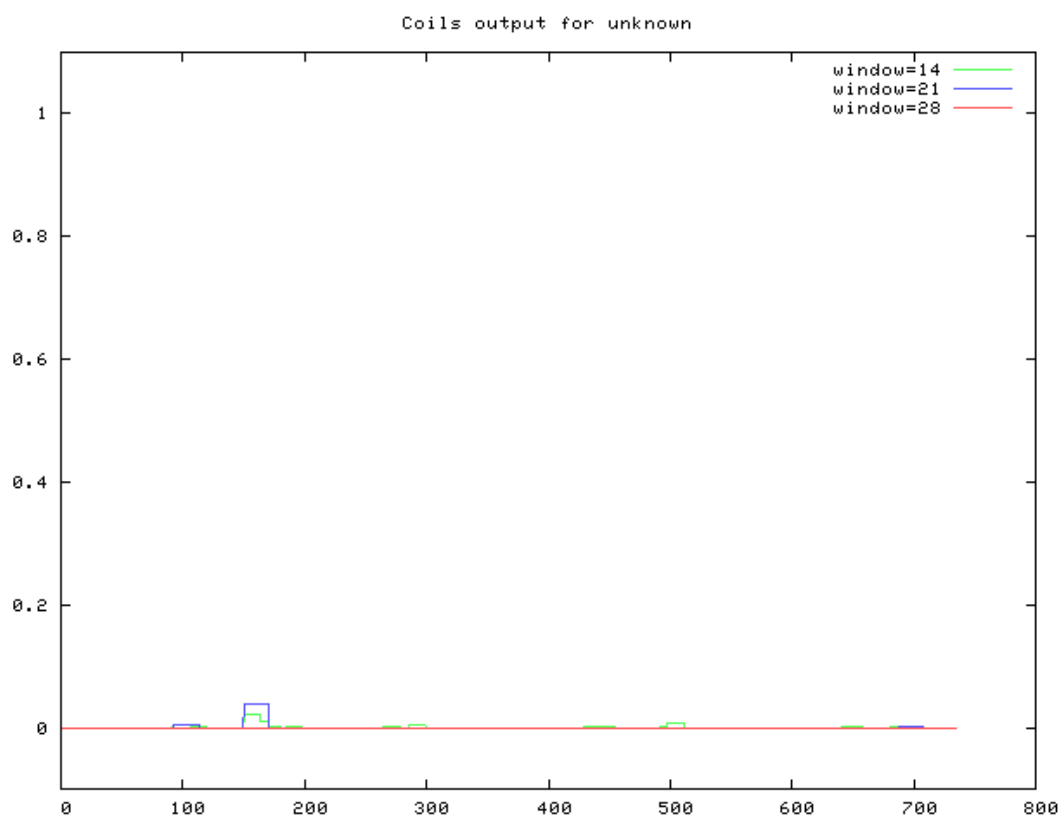
10	20	30	40	50	60	70
MPLAQLKEPWPLME	LVPLDPENGQTS	GEEAGLQPSKDE	GVLKEISITHHV	KAGSEKADPSHF	ELLKVLGQ	
GSFGKVF	LVRKVTRPDS	GHLYAMKVLK	KATLKVRDRV	RTKMERDILAD	VNHPFVVKLHY	AFQTEGKLYLI
LDFLRGGDL	FTRLSKEVM	FTTEEDVKFY	LAELALGLD	HLHSLGIIYR	DLKPENILL	DEEGHIKLTDFGLSK
E	AIDHEKKAY	SFCGTVEYMA	PEVVNRQGH	SHSADWWSY	GVLMFEMLT	GS
LPFQ	GKDRKET	MTLILKAKLG				
MPQFL	STEAQSLL	RALFKRNPAN	RLGSGPDG	AEEIKRHF	YFSTIDWNK	LYRREIKPPFKPAVAQPD
FT	EFTSRTPKD	SPGIPPS	AGAHQLFR	GSFVATGL	MEDDGKPR	APQAPLH
SVVQQL	HGKNLV	FS	SDGYVV			
KETIG	VG	SYSECKRC	VHKATNMEY	AVKVIDK	SKRDPSEE	IEILLRYGQHPNII
TLKD	VYDDGKH	VYLVTE				
LMRGGEL	LDKILRQ	KFFSERE	ASFVLHT	IGKTVEY	LHSQGVV	HRDLKPSN
ILYV	DESGN	PECLRIC	DFGF			
AKQLRA	ENGLLMT	PCYTANF	VAPVLRQ	GYDEGCD	IWSLGILL	YTMLAGYTPFANGP
SDTP	EEILTRIG					
SGKFTL	SGGNWNT	VSETAKD	LVSKMLH	VDPHQRL	TAKQVLQHP	WVTQKDKLPQ
SQLSHQDL	QLVK	GAMAA				
TYSAL	NSSKPTQL	KPIESS	ILAQRR	VRKLP	STTL	

Τατιάνα Καράμοβα
Α.Ε.Μ.: 8503

Total number of **DEKNQRST** in sequence : 321

COILS: Prediction of Coiled Coil Regions in Proteins

COILS is a program that compares a sequence to a database of known parallel two-stranded coiled-coils and derives a similarity score. By comparing this score to the distribution of scores in globular and coiled-coil proteins, the program then calculates the probability that the sequence will adopt a coiled-coil conformation



FindMod: Protein post-translational modification prediction

Predict potential protein post-translational modifications (PTM) and find potential single amino acid substitutions in peptides. The experimentally measured peptide masses are compared with the theoretical peptides calculated from a specified Swiss-Prot/TrEMBL entry or from a user-entered sequence, and mass differences are used to better characterise the protein of interest.

[illegible]

Τατιάνα Καράμοβα

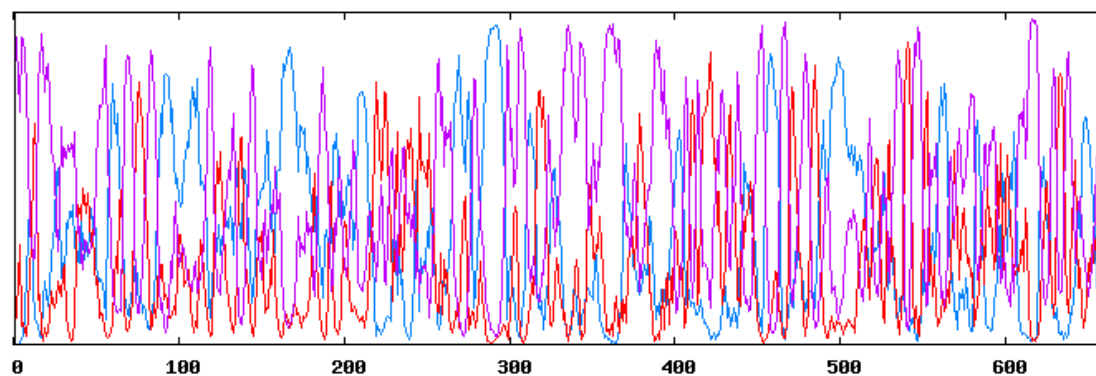
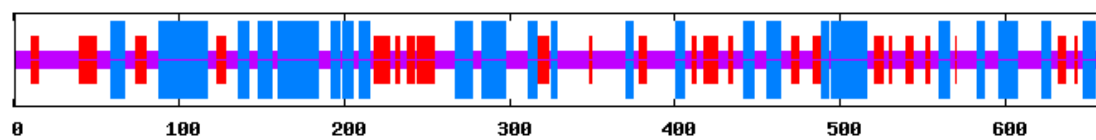
A.E.M.: 8503

Sequence length : 735

GOR4 :

Alpha helix	(Hh)	:	267	is	36.33%
3 ₁₀ helix	(Gg)	:	0	is	0.00%
Pi helix	(Ii)	:	0	is	0.00%
Beta bridge	(Bb)	:	0	is	0.00%
Extended strand	(Ee)	:	124	is	16.87%
Beta turn	(Tt)	:	0	is	0.00%
Bend region	(Ss)	:	0	is	0.00%
Random coil	(Cc)	:	344	is	46.80%
Ambiguous states (?)		:	0	is	0.00%

Otherstates	:	0	is	0.00%
-------------	---	---	----	-------

Prediction result file (text): [\[GOR4\]](#)

Για να εξετάσουμε τη δευτεροταγή διαμόρφωση της πρωτεΐνης μας μπορούμε από τον ιστότοπο της PDB να βρούμε όλες τις απαραίτητες πληροφορίες, ακόμη και 3D διαγράμματα με τις α-έλικες και τα β-πτυχωτά:

Τατιάνα Καράμοβα

Α.Ε.Μ.: 8503

Structureⁱ

Secondary structure

1

Legend: ■ Helix ■ Turn ■ Beta strand ■ PDB Structure known for this area[Show more details](#)

3D structure databases

Select the link destinations: ● PDB ⁱ ● RCSB PDB ⁱ ● PDBj ⁱ	PDB entry	Method	Resolution (Å)	Chain	Positions	PDBsum
	2WNT	X-ray	2.40	A/B	413-719	[>]
	2Z7Q	X-ray	2.00	A	33-353	[>]
	2Z7R	X-ray	2.00	A	33-353	[>]
	2Z7S	X-ray	2.10	A	33-353	[>]
	3RNY	X-ray	2.70	A/B	411-735	[>]
	3TEI	X-ray	2.40	B	712-735	[>]
	4H3P	X-ray	2.30	B/E	712-735	[>]
	4NIF	X-ray	2.15	A/D	411-735	[>]
	5CSF	X-ray	2.40	C	683-735	[>]
	5CSI	X-ray	2.13	C	689-735	[>]
	5CSJ	X-ray	2.70	C	696-735	[>]
	5CSN	X-ray	2.95	C	683-720	[>]
ProteinModelPortal ⁱ	Q15418.					
SMR ⁱ	Q15418.					
ModBase ⁱ	Search...					
MobiDB ⁱ	Search...					

Miscellaneous databases

EvolutionaryTraceⁱ Q15418.

Για τις θέσεις 413-719 έχουμε:

Function and Biology

[Details](#)

Reaction catalysed:

ATP + a protein = ADP + a phosphoprotein.

Biochemical function:

- ATP binding [↗](#)

Biological process:

- protein phosphorylation [↗](#)

Cellular component:

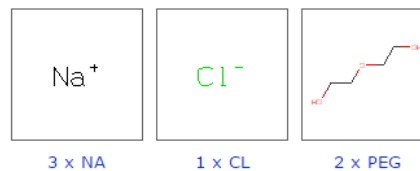
- not assigned

Sequence domains:

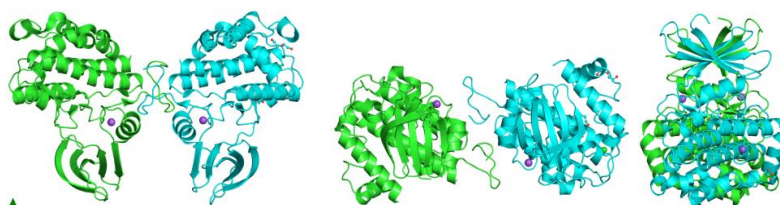
- Protein kinase, ATP binding site [↗](#)
- Protein kinase domain [↗](#)
- Serine/threonine-protein kinase, active site [↗](#)
- Protein kinase-like domain [↗](#)

Ligands and Environments

3 bound ligands:



No modified residues



Σχεδιασμός Primers

Στη συγκεκριμένη άσκηση υποδεικνύεται η διαδικασία κλωνοποίησης του γονιδίου της πρωτεΐνης TGFB1 P01137 (TGFB1_HUMAN) στον πλασμιδιακό φορέα (PAN5)

Ξεκινάμε με την εύρεση της νουκλεοτιδικής ακολουθίας της πρωτεΐνης. Έτσι, πηγαίνουμε στον ιστότοπο <http://www.expasy.org/> όπου στην αναζήτηση πληκτρολογούμε το όνομα της πρωτεΐνης και επιλέγουμε το αποτέλεσμα της αναζήτησης με βάση το UniProt. Όταν επιλέξουμε την πρωτεΐνη μπαίνουμε στην ιστοσελίδα <http://www.uniprot.org/uniprot/P01137> όπου μπορούμε να βρούμε όλες τις απαραίτητες πληροφορίες για το συγκεκριμένο μακρομόριο.

Πιο αναλυτικά, από το PTM / Processing επιλέγουμε την αλυσίδα που σχετίζεται με τη βασική λειτουργία της πρωτεΐνης και εντοπίζουμε την επιθυμητή ακολουθία:

10	20	30	40	50
MPPSGRLRL	LLLPLLWLV	LTPGRPAAGL	STCKTIDMEL	VKRKRIBAIR
60	70	80	90	100
GQILSKRLA	SPPSQGEVPP	GPLPEAVLAL	YNSTRDRVAG	ESAEPEPEPE
110	120	130	140	150
ADYYAKEVTR	VLMVETHNEI	YDKFKQSTHS	IYMFNTSEL	REAVPEPVLL
160	170	180	190	200
SRAELRLRL	KLKVEQHVEL	YQKYSNNSWR	YLSNRLAPS	DSPFWLSFDV
210	220	230	240	250
TGVVRQWLSR	GGEIEGFRLS	AHCSCDSRDN	TLQVDINGFT	TGRRGDLATI
260	270	280	290	300
HGMNRPFLLL	MATPLERAQH	LQSSRHRRAL	DTNYCFSSTE	KNCCVRQLYI
310	320	330	340	350
DFRKDLGWKW	IHEPKGYHAN	FCLGPCPYIW	SLDTQYSKVL	ALYNQHNPQA
360	370	380	390	
SAAPCCVPQA	LEPLPIVYV	GRKPKVEQLS	NMIVRSCKCS	

Όπου η υπογραμμισμένη είναι εκείνη που μας ενδιαφέρει.

Στη συνέχεια επιστρέφουμε πίσω στη σελίδα της UniProtKB και κάτω από το Sequence βρίσκουμε το Sequence Databases όπου υπάρχει ένα λινκ για το CCDS όπου μπορεί να βρεθεί η αντιστοιχία των καταλοίπων των αμινοξέων της πολυπεπτιδικής αλυσίδας με τις τριπλέτες βάσεων που τα κωδικοποιεί. Έτσι, μπορεί κανείς εύκολα να βρεί τη νουκλεοτιδική αλληλουχία. Αυτή λοιπόν φαίνεται παρακάτω

Τατιάνα Καραμόβα

A.E.M.: 8503

Nucleotide Sequence (1173 nt):

ATGCCGCCCTCCGGGCTGCGGCTGCTGCGCTGCTACCGCTGCTGTGGCTACTGGTGCTGACGCCTG
 GCCGGCCGGCCGCGGGACTATCCACCTGCAAGACTATCGACATGGAGCTGGTGAAGCGGAAGCGCATCGA
 GGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCC
 GGCCCGCTGCCCGAGGCCGTGCTCGCCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGCAG
 AACCAGAGCCCCGAGCCTGAGGCCGACTACTACGCCAAGGAGGTCACCCGCGTGCTAATGGTGGAAACCCA
 CAACGAAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC
 CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTCAAGTTAAAG
 TGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGATACCTCAGCAACCGGCTGCT
 GGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTGATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGT
 GGAGGGGAAATTGAGGGCTTTTCGCTTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAG
 TGGACATCAACGGGTTCACTACCGGCCGCGGAGGTGACCTGGCCACCATTGATGGCATGAACCGGCTTT
 CCTGCTTCTCATGGCCACCCCGCTGGAGAGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG
 GACACCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCA
 AGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCC
 CTACATTTGGAGCCTGGACACGCAGTACAGCAAGTCTGCGCCCTGTACAACCAGCATAACCCGGGCGCC
 TCGGCGGGCGCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCCGCAAGC
 CCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCTCCTGCAAGTGCAGCTGA

:

GCCCTG

GACACCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAG
 CTGTACATTGACTTCCGCA
 AGGACCTCGGCTGGAAGTGGATCCACGAGCCCAAGGGCTACCATGCCAACT
 TCTGCCTCGGGCCCTGCCC
 CTACATTTGGAGCCTGGACACGCAGTACAGCAAGTCTGCGCCCTGTACAA
 CCAGCATAACCCGGGCGCC
 TCGGCGGGCGCCGTGCTGCGTGCCGCAGGCGCTGGAGCCGCTGCCCATCGTG
 TACTACGTGGGCCGCAAGC
 CCAAGGTGGAGCAGCTGTCCAACATGATCGTGCGCTCCTGCAAGTGCAGCT
 GA

Τατιάνα Καράμοβα

A.E.M.: 8503

Nucleotide Sequence (1173 nt):

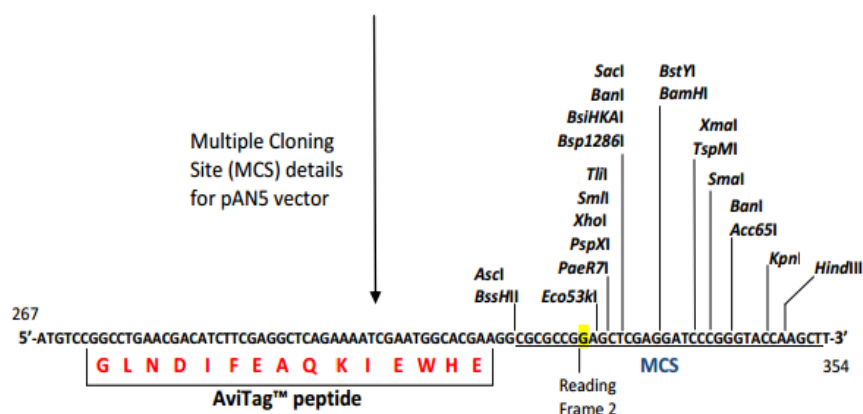
ATGCCGCCCTCCGGGCTGCGGCTGCTGCCGCTGCTGCTACCGCTGCTGTGGCTACTGGTGCTGACGCCTG
 GCCGGCCGGCCGCGGGACTATCCACCTGCAAGACTATCGACATGGAGCTGGTGAAGCGGAAGCGCATCGA
 GGCCATCCGCGGCCAGATCCTGTCCAAGCTGCGGCTCGCCAGCCCCCGAGCCAGGGGGAGGTGCCGCC
 GGCCCGCTGCCGAGGCCGTGCTCGCCCTGTACAACAGCACCCGCGACCGGGTGGCCGGGGAGAGTGACAG
 AACCGGAGCCCGAGCCTGAGGCCGACTACTACGCCAAGGAGGTACCCCGCGTGTAAATGGTGGAAACCA
 CAACGAAATCTATGACAAGTTCAAGCAGAGTACACACAGCATATATATGTTCTTCAACACATCAGAGCTC
 CGAGAAGCGGTACCTGAACCCGTGTTGCTCTCCCGGGCAGAGCTGCGTCTGCTGAGGCTCAAGTTAAAG
 TGGAGCAGCACGTGGAGCTGTACCAGAAATACAGCAACAATTCTGGCGATACCTCAGCAACCGGCTGCT
 GGCACCCAGCGACTCGCCAGAGTGGTTATCTTTTGTATGTCACCGGAGTTGTGCGGCAGTGGTTGAGCCGT
 GGAGGGGAAATTGAGGGCTTTTCGCTTAGCGCCCACTGCTCCTGTGACAGCAGGGATAACACACTGCAAG
 TGGACATCAACGGGTTCACTACCGGCCGCCGAGGTGACCTGGCCACCATTATGGCATGAACCGGCCCTTT
 CCTGCTTCTCATGGCCACCCCGCTGGAGAGGGGCCAGCATCTGCAAAGCTCCCGGCACCGCCGAGCCCTG
 GACACC AACTAT TGC TTC AGCTCCACGGAG AAC TGCTGCGTG CAGCTGTACATT GAC TTC
 GAC CTC GGC TGG TGGATCCACGAGCCC GGC TACCATGCCAACTTC TGCTC GGG CCG TGCC
 CTACATT TGG AGC CTG GACACG CAGTACAGC GTCTG GCCCTGTACAACCAGCAT AAC CCG GGC GCC
 TCGGCGGCGCGTGTGCTGCGTCCG CAGGCGCTG GAG CCG CTGCCATCGTGTACTACGTG GGC CGCAAGC
 CCAAGGTG GAG CAGCTG TCCAACATGATCGTGCCTCC TGCAAG TGCAGCTGA

Translation (390 aa):

MPPSGLRLLP LLLP LLLWLLV LTPGRPAAGLSTCKTIDMELVKRKRIE AIRGQILSKRLRLASPPSQGEVPP
 GPLPEAVLALYNSTRDRVAGESAEPEPEPEADYYAKEVTRVLMVETHNEIYDKFKQSTHSIYMFNTSEL
 REAVPEPVLLSRAELRLLRLKLKVEQHVELYQKYSNNSWRYLSNRL LAPSDSPEWLSFDVTGVVRQWLSR
 GGEIEGFRLSAHCSCDSRDNTLQVDINGFTTGRRGLATI HGMNRPFL LLMATPLERAQHLQSSRRHRRAL
 DTNYCFSSTE NCCV QLYIDE DLGM KIEEP GYHANFCLGPCPYIMSLDTQVS VLALYNQHNPGA
 SAAPCCVPQALEPLPTVYVVG RPKVEQLSNMIVRSCKCS

Για την Κλωνοποίηση

Κάνουμε αναζήτηση πληροφοριων του φορέα κλωνοποίησης PAN5 στο <https://www.google.gr> με τις λέξεις Avidity Pan5. Επιλέγουμε το 2^ο αποτέλεσμα της αναζήτησης όπου εμφανίζεται ο χάρτης πλασμιδίου και άλλες γενικές πληροφορίες.



Συνεχίσουμε με την εύρεση των περιοριστικών ενζύμων που θα χρησιμοποιήσουμε στην κλωνοποίηση του γονιδίου στον φορέα PAN5.

Από το <https://www.google.gr> αναζητούμε την ιστοσελίδα NEBcutter και τοποθετούμε σε αυτήν τη νουκλεοτιδική ακολουθεία που βρήκαμε παραπάνω και πατάμε Submit.

Τατιάνα Καράμοβα

A.E.M.: 8503

Sequence Name:

TGFB1

Target Sequence:

```

1      10      20      30      40      50      60      70      80      90
1      GCCCTGGACACCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCC
101    ACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCTACATTGGAGCCTGGACACGCAGTACAGCAAGGTCTGGCCCTGTACAACCA
201    GCATAACCCGGGCGCTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCGCAAGCCCAAGGTGGAGCAG
301    CTGTCCAACATGATCGTGCGCTCCTGCAAGTGCAGCTGA

```

Primer Size (bases)			Primer Tm (°C)			Primer %GC		
Min	Opt	Max	Min	Opt	Max	Min	Opt	Max
18	20	27	57	60	63	20	50	80

Region to Amplify		Experimental Conditions	
From	To	Salt conc.	Primer conc.
1 bp	339 bp	50 mM	50 nM

Fwd Primer 5' Additions:

Proteolytic Cleavage Site:

None

Restriction Enzyme Site:

XhoI

Rev Primer 5' Additions:

Proteolytic Cleavage Site:

None

Restriction Enzyme Site:

HindIII

Prepare Primers For:

☒ Traditional Cloning
(no TOPO® or Gateway® additions)

☐ Gateway® Cloning

☐ Directional TOPO® Cloning

☐ Multisite Gateway® Cloning

☐ B4-B1 Element

☐ B1-B2 Element

☐ B2-B3 Element

Στο πεδίο Region to Amplify επιλέγουμε από την αρχή μέχρι το τέλος του γονιδίου. Συνεχίζοντας επιλέγουμε από τις λίστες των ενζύμων περιορισμού εκείνα τα ένζυμα που ενώ κόβουν τον πλασμιδιακό φορέα, δεν κόβουν το γονίδιο που θέλουμε να κλωνοποιήσουμε. Αφού τα επιλέξουμε, πατάμε Submit και στα αποτελέσματα βρίσκονται πιθανοί εκκινητές με τις θερμοκρασίες υβριδισμού τους. Έτσι έχουμε:

Sequence Name:

TGFB1

Target Sequence:

```

1      10      20      30      40      50      60      70      80      90
1      GCCCTGGACACCAACTATTGCTTCAGCTCCACGGAGAAGAACTGCTGCGTGCGGCAGCTGTACATTGACTTCCGCAAGGACCTCGGCTGGAAGTGGATCC
101    ACGAGCCCAAGGGCTACCATGCCAACTTCTGCCTCGGGCCCTGCCCTACATTGGAGCCTGGACACGCAGTACAGCAAGGTCTGGCCCTGTACAACCA
201    GCATAACCCGGGCGCTCGGCGGCGCCGTGCTGCGTGCCGAGGCGCTGGAGCCGCTGCCCATCGTGTACTACGTGGGCGCAAGCCCAAGGTGGAGCAG
301    CTGTCCAACATGATCGTGCGCTCCTGCAAGTGCAGCTGA

```

Rank: 1 Product Length: 339 Product Region: 1 - 339				
Primer Name	%GC	Strand	Size (bases)	Tm (°C)
<input checked="" type="checkbox"/> TGFB1 1 F	55.00	FWD	20	60.38
<input checked="" type="checkbox"/> TGFB1 1 R	55.56	REV	18	60.45
5' Addition *		Primer Sequence		
<input checked="" type="checkbox"/> C T C G A G		<input checked="" type="checkbox"/> G C C C T G G A C A C C A A C T A T T G		
<input checked="" type="checkbox"/> A A G C T T		<input checked="" type="checkbox"/> T C A G C T G C A C T T G C A G G A		

Παρόμοια διαδικασία ακολουθούμε και για τον σχεδιασμό Primers, την Κλωνοποίηση και την Δημιουργία Εκκινητών για την πρωτεΐνη BMP2:

Σχεδιασμό Primers

Από το PTM / Processing επιλέγουμε την αλυσίδα που σχετίζεται με τη βασική λειτουργία της πρωτεΐνης και εντοπίζουμε την επιθυμητή ακολουθεία:

```

      10      20      30      40      50
MVAGTRCLLA LLLPQVLLGG AAGLVPELGR RKFAAASSGR PSSQPSDEVL
      60      70      80      90     100
SEFELRLLSM FGLKQRPTPS RDAVPPYML DLYRRHSGQP GSPAPDHRLE

```

Τατιάνα Καράμοβα
A.E.M.: 8503

```

      110      120      130      140      150
RAASRANTVR SFHHEESLEE LPETSGKTTR RFFFNLSIP TEEFITSDEL
      160      170      180      190      200
QVFREQMQDA LGNNSSFHHR INIYEIIPKA TANSKFPVTR LLDTRLVNQN
      210      220      230      240      250
ASRWESFDVT PAVMRWTAQG HANHGFEVEV AHLEEKQGVV KRVHVRISRL
      260      270      280      290      300
HQDEHSWSQI RPLLVTFGHD GKGHPLHKRE KRQAKHKQRK RLKSSCKRHP
      310      320      330      340      350
LYVDFSDVGV NDWIVAPPGY HAFYCHGECF FPLADHLNST NHAIVQTLVN
      360      370      380      390
SVNSKIPKAC CVPTELSAIS MLYLDENEKV VLKNYQDMVV EGCGR

```

Όπου η υπογραμμισμένη είναι εκείνη που μας ενδιαφέρει.

Στη συνέχεια επιστρέφουμε πίσω στη σελίδα της UniProtKB και κάτω από το Sequence βρίσκουμε το Sequence Databases όπου υπάρχει ένα λινκ για το CCDS όπου μπορεί να βρεθεί η αντιστοιχία των καταλοίπων των αμινοξέων της πολυπεπτιδικής αλυσίδας με τις τριπλέτες βάσεων που τα κωδικοποιεί. Έτσι, μπορεί κανείς εύκολα να βρεί τη νουκλεοτιδική αλληλουχία. Αυτή λοιπόν φαίνεται παρακάτω:

Nucleotide Sequence (1191 nt):

ATGGTGGCCGGGACCCGCTGTCTTCTAGCGTTGCTGCTTCCCCAGGTCCTCTGGGCGGCGGGCTGGCC
TCGTTCCGGAGCTGGGCCGAGGAAGTTCGCGGCGGCGTCTCGGGCCGCCCTCATCCCAGCCCTCTGA
CGAGGTCCTGAGCGAGTTCGAGTTGCGGCTGCTCAGCATGTTCCGGCTGAAACAGAGACCCACCCCGAGC
AGGGACGCGGTGGTGCCCCCTACATGCTAGACCTGTATCGCAGGCACTCAGGTCAGCCGGGCTACCCG
CCCCAGACCACCGGTTGGAGAGGGCAGCCAGCCAGCCAACTGTGCGCAGCTTCCACCATGAAGAATC
TTTGGAAGAACTACCAGAAACGAGTGGGAAAACAACCCGGAGATTCTTCTTTAATTTAAGTTCTATCCCC
ACGGAGGAGTTTATCACCTCAGCAGAGCTTCAGGTTTTCCGAGAACAGATGCAAGATGCTTTAGGAAACA
ATAGCAGTTTCCATCACCGAATTAATATTTATGAAATCATAAACCTGCAACAGCCAACTCGAAATCCC
CGTGACCAGACTTTTGGACACCAGGTTGGTGAATCAGAATGCAAGCAGGTGGGAAAGTTTGTATGTCACC
CCCGCTGTGATGCGGTGGACTGCACAGGGACACGCCAACCATGGATTCTGTGGTGGAAAGTGGCCCACTTGG
AGGAGAAACAAGGTGTCTCCAAGAGACATGTTAGGATAAGCAGGTCTTTGCACCAAGATGAACACAGCTG
GTCACAGATAAGGCCATTGCTAGTAACCTTTTGCCCATGATGGAAAAGGGCATCTCTCCACAAAAGAGAA
AAACGTCAAGCCAAACACAAACAGCGGAAACGCCTTAAGTCCAGCTGTAAGAGACACCCCTTTGTACGTGG
ACTTCAGTGACGTGGGGTGGAAATGACTGGATTGTGGCTCCCCCGGGGTATCACGCCTTTTACTGCCACGG
AGAATGCCCTTTTCTCTGGCTGATCATCTGAACCTCCACTAATCATGCCATTGTTAGACGTTGGTCAAC
TCTGTTAACTCTAAGATTCTTAAGGCATGCTGTGTCCGACAGAACTCAGTGCTATCTCGATGCTGTACC
TTGACGAGAATGAAAAGGTTGTATTAAAGAACTATCAGGACATGGTTGTGGAGGGTTGTGGGTGTCGCTA
G

CAAGCCAAACACAAACAGCGGAAACGCCTTAAGTCCAGCTGTAAGAGACAC
CCTTTGTACGTGG
ACTTCAGTGACGTGGGGTGGAAATGACTGGATTGTGGCTCCCCCGGGGTATC
ACGCCTTTTACTGCCACGG
AGAATGCCCTTTTCTCTGGCTGATCATCTGAACCTCCACTAATCATGCCAT
TGTTTCAGACGTTGGTCAAC
TCTGTTAACTCTAAGATTCTTAAGGCATGCTGTGTCCGACAGAACTCAGT
GCTATCTCGATGCTGTACC
TTGACGAGAATGAAAAGGTTGTATTAAAGAACTATCAGGACATGGTTGTGG
AGGGTTGTGGGTGTCGCTA
G

Έχουμε λοιπόν, τα εξής αποτελέσματα:

Τατιάνα Καράμοβα

A.E.M.: 8503



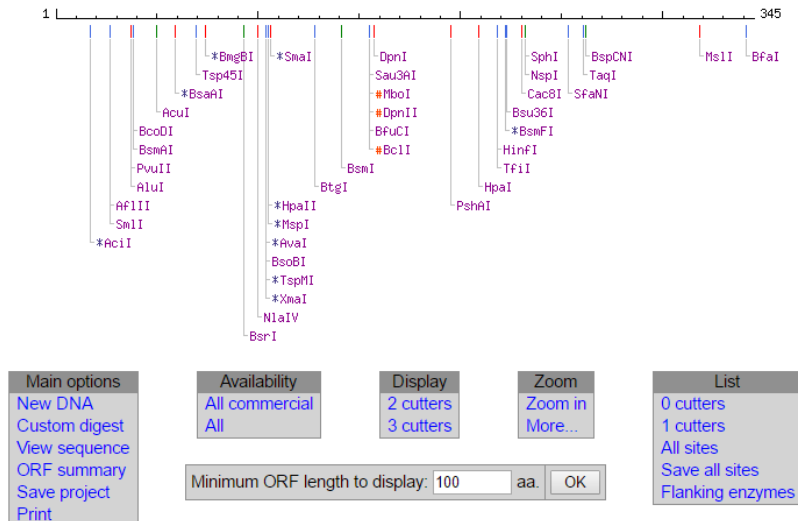
Linear Sequence: unnamed sequence

Help

Comments

Display: - NEB single cutter restriction enzymes
 - Main non-overlapping, min. 100 aa ORFs
 GC=49%, AT=51%

Cleavage code	Enzyme name code
blunt end cut	Available from NEB
5' extension	Has other supplier
3' extension	Not commercially available
cuts 1 strand	*: cleavage affected by CpG meth.
	: cleavage affected by other meth.
	(enz.name): ambiguous site



Από εδώ επιλέγουμε τη λίστα με τα 0 cutters τα οποία αναφέρονται σε ένζυμα που δεν κόβουν πουθενά το γονίδιο. Είναι επιθυμητό να βρεθούν ένζυμα τα οποία να κόβουν το πλασμίδιο αλλά να μην κόβουν το γονίδιο, επομένως θα πρέπει να βρεθούν 2 τέτοια ένζυμα τα οποία να υπάρχουν στον χάρτη του πλασμιδίου και στη λίστα των 0 cutters.

Έτσι, έχουμε τα δύο ένζυμα που μας ενδιαφέρουν:

- XhoI CTCGAG
- HindIII AAGCTT

Δημιουργία εκκινήτων

Αναζητούμε στο <https://www.google.gr> τις λέξεις κλειδιά life technologies primers design και επιλέγουμε το πρώτο αποτέλεσμα της αναζήτησης. Στο πεδίο Sequence Name συμπληρώνουμε ένα όνομα, ως Application επιλέγουμε το PCR: cloning και στο Researcher Name γράφουμε ένα όνομα. Πατάμε Submit και εμφανίζεται το παρακάτω σχήμα:

Τατιάνα Καράμοβα

A.E.M.: 8503

Sequence Name:

BMP2

Target Sequence:

```

1      10      20      30      40      50      60      70      80      90
1  CAAGCCAAACACAAACAGCGGAAACGCCTTAAGTCCAGCTGTAAGAGACACCCCTTTGTACGTGGACTTCAGTGACGTGGGGTGAATGACTGGATTGTGG
101 CTCCCCCGGGGTATCAGCCCTTTACTGCCACGGAGAAATGCCCTTTTCCTCTGGCTGATCATCTGAACCTCCACTAATCATGCCATTGTTGAGACGTTGGT
201 CAACTCTGTTAACTCTAAGATTCTTAAGGCATGCTGTGTCCCGACAGAACTCAGTGTATCTCGATGCTGTACCTTGACGAGAATGAAAAGGTTGTATTA
301 AAGAACTATCAGGACATGGTTGTGGAGGGTTGTGGGTGTCGCTAG

```

Primer Size (bases)			Primer Tm (°C)			Primer %GC		
Min	Opt	Max	Min	Opt	Max	Min	Opt	Max
18	20	27	57	60	63	20	50	80

Region to Amplify		Experimental Conditions	
From	To	Salt conc.	Primer conc.
1 bp	345 bp	50 mM	50 nM

Fwd Primer 5' Additions:

Proteolytic Cleavage Site:

None

Restriction Enzyme Site:

XhoII

Rev Primer 5' Additions:

Proteolytic Cleavage Site:

None

Restriction Enzyme Site:

HindIII

Prepare Primers For:

- ☒ Traditional Cloning
 (no TOPO® or Gateway® additions)
☐ Gateway® Cloning
☐ Directional TOPO® Cloning

Multisite Gateway® Cloning

- ☐ B4-B1 Element
☐ B1-B2 Element
☐ B2-B3 Element

Στο πεδίο Region to Amplify επιλέγουμε από την αρχή μέχρι το τέλος του γονιδίου. Συνεχίζοντας επιλέγουμε από τις λίστες των ενζύμων περιορισμού εκείνα τα ένζυμα που ενώ κόβουν τον πλασμιδιακό φορέα, δεν κόβουν το γονίδιο που θέλουμε να κλωνοποιήσουμε. Αφού τα επιλέξουμε, πατάμε Submit και στα αποτελέσματα βρίσκονται πιθανοί εκκινητές με τις θερμοκρασίες υβριδισμού τους. Έτσι έχουμε:

Sequence Name:





BMP2

Target Sequence:

```

1      10      20      30      40      50      60      70      80      90
1  CAAGCCAAACACAAACAGCGGAAACGCCTTAAGTCCAGCTGTAAGAGACACCCCTTTGTACGTGGACTTCAGTGACGTGGGGTGAATGACTGGATTGTGG
101 CTCCCCCGGGGTATCAGCCCTTTACTGCCACGGAGAAATGCCCTTTTCCTCTGGCTGATCATCTGAACCTCCACTAATCATGCCATTGTTGAGACGTTGGT
201 CAACTCTGTTAACTCTAAGATTCTTAAGGCATGCTGTGTCCCGACAGAACTCAGTGTATCTCGATGCTGTACCTTGACGAGAATGAAAAGGTTGTATTA
301 AAGAACTATCAGGACATGGTTGTGGAGGGTTGTGGGTGTCGCTAG

```

Rank: 1 Product Length: 345 Product Region: 1 - 345					
Primer Name		%GC	Strand	Size (bases)	Tm (°C)
<input type="checkbox"/> BMP2 1 F		50.00	FWD	20	62.73
<input type="checkbox"/> BMP2 1 R		63.16	REV	19	61.53
5' Addition *		Primer Sequence			
 R G A T C Y		 C A A G C C A A A C A C A A A C A G C G			
 A A G C T T		 C T A G C G A C A C C C A C A A C C C			

Highlight Target Sequence