

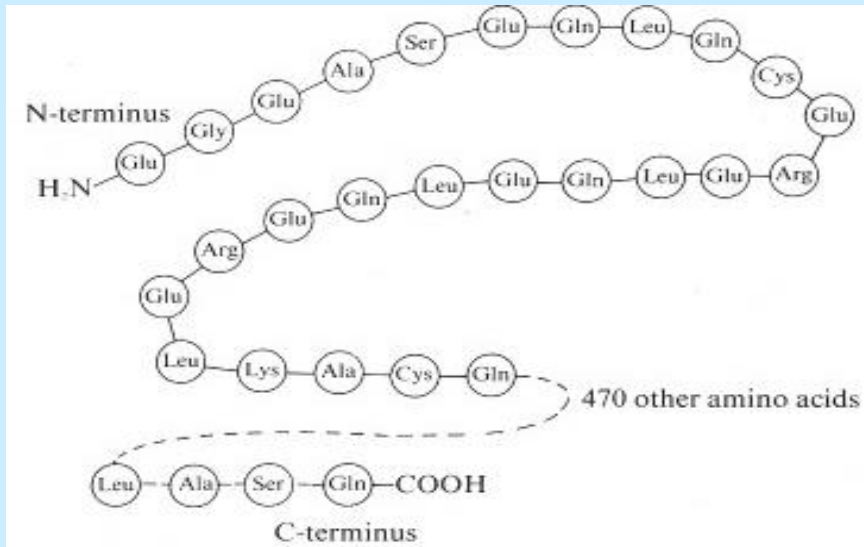
- **Πηγές δεδομένων για πρωτεΐνες**

Υπάρχει ένα εύρος πηγών πληροφοριών για τις πρωτεΐνες, οι οποίες χωρίζονται ανάλογα με την πληροφορία που παρέχουν, δηλ. τον τύπο δεδομένων των πρωτεϊνών: βάσεις δεδομένων με αλληλουχίες, βάσεις δεδομένων με δευτεροταγή δεδομένα και βάσεις δεδομένων με δομές πρωτεϊνών.

- Ο τύπος της βάσης δεδομένων ορίζεται ανάλογα με το επίπεδο της πληροφορίας για τις πρωτεΐνες που έχουν αποθηκευμένο:

<u>Πληροφορία πρωτεΐνης</u>	<u>Τύπος βάσης δεδομένων</u>
Primary structure (sequence)	primary
Secondary structure (motif)	secondary
Tertiary structure (domain-module)	structure

# Επίπεδα της πρωτεϊνικής δομής

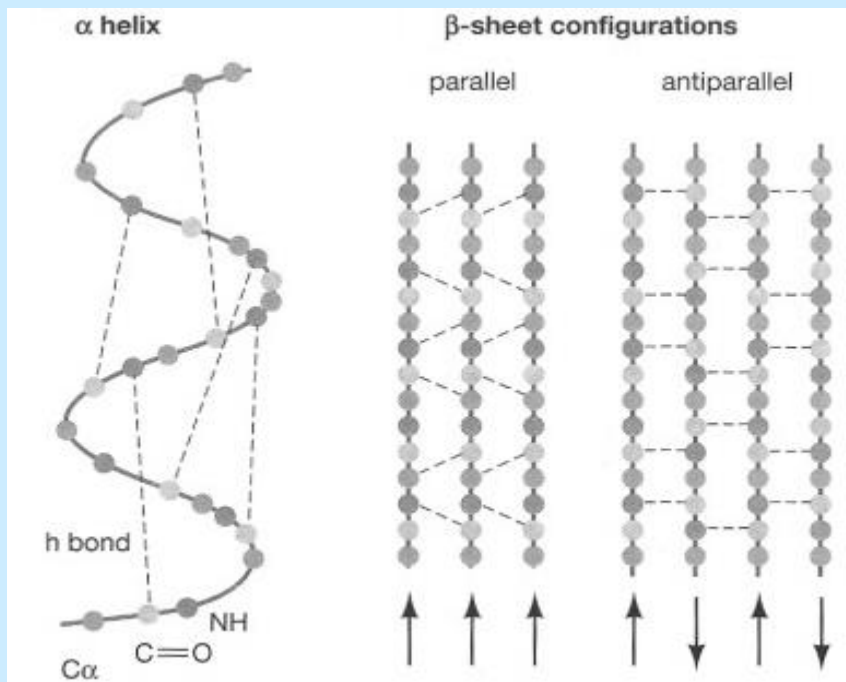


## Πρωτοταγής δομή

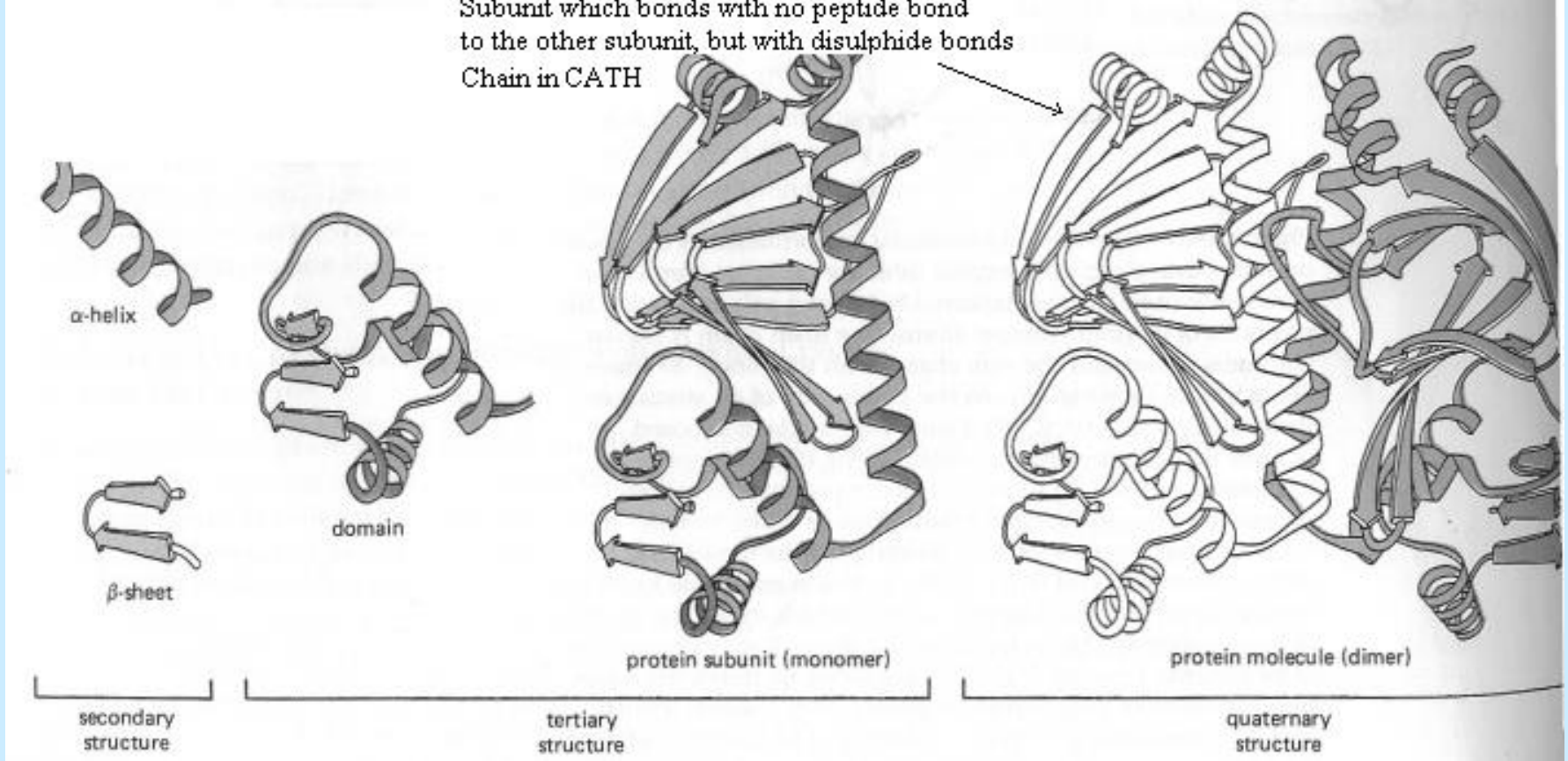
Πρωτοταγής δομή είναι η γραμμική αλληλουχία αμινοξέων σε ένα μόριο πρωτεΐνης.

## Δευτεροταγής δομή

Αμινοξέα, τα οποία είναι δίπλα το ένα στο άλλο στην αλυσίδα, αλληλεπιδρούν (αλληλεπιδράσεις δεσμών υδρογόνου) με πολλαπλούς τρόπους: αυτό έχει σαν συνέπεια το ένα μέρος του γραμμικού πολυπεπτιδίου να γίνει ελικοειδές, σχηματίζοντας μια α-έλικα (α-helix) και το άλλο μέρος να σχηματίσει ένα πλέγμα με γραμμές που έχουν μία κατεύθυνση (β-sheet). Αυτές οι αλληλεπιδράσεις δημιουργούν την δευτεροταγή δομή.



Subunit which bonds with no peptide bond  
to the other subunit, but with disulphide bonds  
Chain in CATH



### Τριτοταγής δομή

Τριτοταγής δομή είναι η ολική δίπλωση μιας πρωτεϊνικής αλληλουχίας, η οποία σχηματίζεται από το πακετάρισμα των δευτεροταγών (δισουλφιδικοί θειούχοι δεσμοί ανάμεσα σε αμινοξέα που περιέχουν θείο) σταθεροποιούν το μόριο στην τρισδιάστατη δομή του).

### Τεταρτοταγής δομή

Τεταρτοταγής δομή είναι η διάταξη διαφορετικών πολυπεπτιδικών αλυσίδων (υποομάδες-subunits) σε ένα μόριο πρωτεΐνης.

### Domain

Ο τομέας (domain) μιας πρωτεΐνης είναι ο συνδυασμός από  $\alpha$ -helices και  $\beta$ -sheets, τα οποία ενώνονται μεταξύ τους για να σχηματίσουν συμπαγές διπλωμένες σφαιρικές μονάδες.

## Μοτίβα (motifs)

- Συγκεκριμένοι συνδυασμοί των  $\alpha$ -helices και  $\beta$ -sheets που δημιουργού μία σφαιρική δομή και υπάρχουν επαναλαμβανόμενες σε πολλές πρωτεΐνες που δεν σχετίζονται μεταξύ τους ονομάζονται μοτίβα (motifs).
- Για παράδειγμα, το beta-alpha-beta μοτίβο, το οποίο συναντάται σε πολλές διαφορετικές πρωτεΐνες.



- Οι (primary) πρωτοταγείς βάσεις δεδομένων περιέχουν την πρωτοταγή δομή της πρωτεΐνης, δηλ. τη σειρά αμινοξέων που συμβολίζονται με γράμματα στη σειρά.
- Στις δευτεροταγής βάσεις δεδομένων υπάρχουν patterns (συνηθισμένες εκφράσεις, «δαχτυλικά αποτυπώματα», μπλόκα, περιγράμματα), δηλ. οι δευτεροταγές δομές αντιστοιχούν σε περιοχές με τοπική κανονικότητα (*α-helices* και *b-strands*) που είναι διατηρημένα μοτίβα.
- Στις βάσεις δεδομένων δομής υπάρχουν τομείς (*modules*) που είναι τα αποτελέσματα της συσκευασίας των δευτεροταγών δομών μέσα σε μια δίπλωση (αυτόνομες διπλούμενες μονάδες).
  - Η πληροφορία υπάρχει ως ένα σύνολο ατομικών συντεταγμένων.

- **Πρωτοταγείς βάσεις δεδομένων για αλληλουχίες**
- Στην επιστημονική βιβλιογραφία, όταν οι πληροφορίες για αλληλουχίες άρχισαν να γίνονται πολλές, τότε ξεκίνησαν σε διάφορα μέρη του κόσμου πολλές μελέτες για τη εξέλιξη πρωτοταγών βάσεων δεδομένων. Οι πρωτοταγείς βάσεις δεδομένων με αλληλουχίες πρωτεϊνών είναι οι παρακάτω:
- PIR
- MIPS
- SWISS-PROT
- TrEMBL
- NRL-3D

## PIR

- Το Protein Information Resource (PIR) δημιουργήθηκε στο National Biomedical Research Foundation (NBRF) από την M. Dayhoff, σαν μια συλλογή αλληλουχιών για την διερεύνηση των εξελικτικών σχέσεων ανάμεσα στις πρωτεΐνες.
- Στην σημερινή της μορφή, η βάση δεδομένων διαχωρίζεται σε 4 τμήματα: PIR1-PIR4, που διαφέρουν όσον αφορά την ποιότητα των δεδομένων και το επίπεδο σχολιασμού που δίνεται. Το PIR1 περιέχει πλήρεις ταξινομημένες καταγραφές με σχόλια και το PIR4 περιέχει καταχωρήσεις που είναι εννοιολογικές μεταφράσεις των αλληλουχιών (not transcribed but translated-genetically engineered).



Welcome to PIR [Protein Information Resource] x

https://proteininformationresource.org


Συχνές επισκέψεις | Εκκινώντας | Suggested Sites | Web Slice Gallery | YPOMEDI

Protein Information Resource

Protein Search | Site Search

About PIR | Resources | Search/Analysis | Download | Support


### INTEGRATED PROTEIN INFORMATICS RESOURCE FOR GENOMIC, PROTEOMIC AND SYSTEMS BIOLOGY RESEARCH




The Universal Protein Resource (UniProt) provides the scientific community with a single, centralized, authoritative resource for protein sequences and functional information.

[UniProtKB](#) | [UniRef](#) | [UniParc](#)


Current release: 2019\_03



**PRO**  
Protein Ontology



**iPTMnet**  
Integrated Protein PTM Resource



**iProLINK**  
Literature Information & Knowledge

**O OTHER RESOURCE**

- [Reference Proteomes](#)
- [iProClass](#)
- [iProXpress](#)
- [RESID](#)

**P PEPTIDE SEARCH**

DATABASE: UniProtKB

Use single letter amino acid code

**T TEXT SEARCH**

DATABASE: iProClass

### Bioinformatics & Computational Biology Graduate Programs:

- [MS program at Georgetown University](#)
- [PhD, MS, PSM and Graduate Certificate programs at University of Delaware](#)

Home | About PIR | Databases | Search/Analysis | Download | Support

SITE MAP | TERMS OF USE

©2016 Protein Information Resource

**University of Delaware**  
15 Innovation Way, Suite 205  
Newark, DE 19711, USA

**Georgetown University Medical Center**  
3300 Whitehaven Street, NW, Suite 1200  
Washington, DC 20007, USA



Welcome to PIR [Protein Information Resource] Search & Tools [PIR - Protein Information Resource]

https://proteininformationresource.org/pirwww/search/

Συχνές επισκέψεις Ξεκινώντας Suggested Sites Web Slice Gallery YPOMEDI

**PIR** A UniProt Consortium Member  
Protein Information Resource

Protein Search Site Search

About PIR Resources Search/Analysis Download Support

**HOME / Search / Search & Analysis Tools**

**Text Search**

**Text Search:** Enter text or identifiers to search against iProClass for individual proteins, or PIRSF for protein families. An advanced search option is available for restriction of search terms, options, matching, and operators.

**Batch Retrieval:** Retrieve multiple entries from iProClass or multiple families from PIRSF database using a specific identifier, or a combination of different identifiers.

**Sequence Search**

**Peptide Match:** Find an exact match for a peptide sequence (3 to 30 amino acid long). Two UniProt databases can be used to perform the search: (1) [UniProtKB](#), which contains functional information on proteins, with accurate, consistent, and rich annotation; or (2) [UniRef100](#), which combines identical sequences and sub-fragments, from any organism, into a single entry.

**Sequence Alignment**

**Multiple Alignment:** Enter multiple sequences in FASTA format and/or multiple UniProtKB identifiers in the ID box to get the CLUSTALW alignment of the sequences along with a neighbor-joining tree and a PIR interactive tree and alignment viewer.

**Pairwise Alignment:** Insert two sequences using the single letter amino acid code or enter two UniProtKB identifiers. The results show the SSearch Smith-Waterman full-length alignments between the two sequences.

**Others**

**ID Mapping:** Use ID Mapping to see the correspondence between any two identifiers, and therefore facilitates the navigation among databases and analysis programs.

https://proteininformationresource.org/pirwww/search/pairwise.shtml

EN 11:23 AM 4/12/2019

## UNI-PROT

- Η βάση δεδομένων UNI-PROT δημιουργήθηκε από το τμήμα Ιατρικής Βιοχημείας στο Πανεπιστήμιο της Γενεύης και από το EMBL, τώρα η βάση δεδομένων διατηρείται σε συνεργασία με το Ελβετικό Ινστιτούτο Βιοπληροφορικής (SIB) και το Ευρωπαϊκό Ινστιτούτο Βιοπληροφορικής (EBI) του EMBL.
- Η βάση δεδομένων δίνει πολλά σχόλια, συμπεριλαμβάνοντας περιγραφές λειτουργιών της πρωτεΐνης και της δομής των τομέων της, των παραλλαγών της κ.ο.κ. Είναι ελάχιστα πλεονάζουσα και είναι διασυνδεόμενη με πολλές άλλες πηγές πληροφόρησης.

## TrEMBL

- Το 1996, η TrEMBL, δημιουργήθηκε ως συμπλήρωμα για τη Uni-Prot/SWISS-Prot. Περιέχει τις μεταφράσεις όλων των κωδικοποιημένων αλληλουχιών στο EMBL. Οι καταχωρήσεις εμπεριέχονται τελικά στη UNI-PROT αλλά με ανεπαρκείς αναλύσεις και σχόλια.
- Το UniProtKB / TrEMBL περιέχει υψηλής ποιότητας υπολογιστικά αναλυμένα αρχεία που εμπλουτίζονται με αυτόματο σχολιασμό και ταξινόμηση. Αυτές οι καταχωρήσεις UniProtKB / TrEMBL που δεν έχουν εξεταστεί διατηρούνται χωριστά από τις καταχωρήσεις χειροκίνητης ανασκόπησης του UniProtKB / Swiss-Prot, έτσι ώστε τα δεδομένα υψηλής ποιότητας του τελευταίου να μην τροποποιούνται με κανέναν τρόπο. Η αυτόματη επεξεργασία των δεδομένων επιτρέπει την ταχεία διάθεση των αρχείων στο κοινό.

## **NRL-3D**

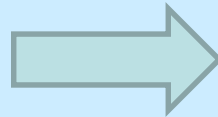
- Η βάση δεδομένων NRL-3D δημιουργήθηκε από το PIR, από αλληλουχίες που ανακτήθηκαν από τη Brookhaven Protein Bank (PBD). Στη PBD, οι πληροφορίες για αλληλουχίες είναι διαθέσιμες για ανάκτηση με λέξεις κλειδιά και για αναζήτηση ομοιοτήτων μεταξύ αλληλουχιών.

# Πρωτοταγείς βάσεις δεδομένων για αλληλουχίες

- Περιέχουν πειραματικά δεδομένα όπως η αλληλουχία νουκλεοτιδίων, η αλληλουχία πρωτεϊνών ή η μακρομοριακή δομή.
- Τα πειραματικά αποτελέσματα υποβάλλονται απευθείας στη βάση δεδομένων από τους ερευνητές και τα δεδομένα είναι ουσιαστικά αρχειακά.
- Κατά την εγγραφή τους δίνεται ένας μοναδικός αριθμός πρόσβασης (sequence ID) που δεν αλλάζει ποτέ: αποτελούν μέρος του επιστημονικού αρχείου.

# Καταγραφή σε βάσεις δεδομένων


- Είναι σημαντικό να καταγράφονται τόσο τα δεδομένα όσο και τα μετα-δεδομένα (δεδομένα από την επεξεργασία των δεδομένων) με συνεπή τρόπο.
- Ελάχιστα πρότυπα πληροφοριών: Σύνολα κατευθυντήριων γραμμών και μορφοτύπων για την αναφορά δεδομένων.



- Τα δεδομένα εύκολα επαληθεύσιμα, αναλύσιμα και ερμηνεύσιμα από την ευρύτερη επιστημονική κοινότητα.
- Διευκολύνουν τη μεταφορά δεδομένων από άρθρα περιοδικών (μη δομημένα δεδομένα) σε βάσεις δεδομένων (δομημένα δεδομένα).

# Εργαλεία για την καταγραφή

- **Ελεγχόμενα λεξιλόγια:** ορίζουν συγκεκριμένες λέξεις για να μειώσουν την αμφισημία και την επανάληψη.

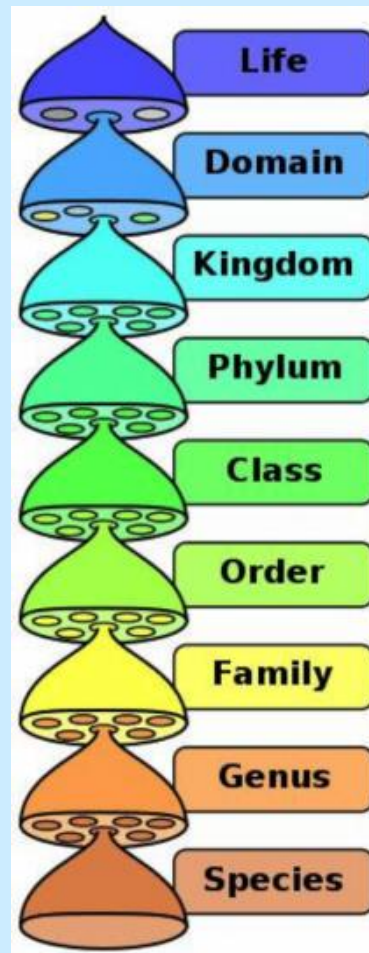
 βοηθά τους υπολογιστές και τους ανθρώπους να κατηγοριοποιούν τις πληροφορίες και βοηθάει στη μείωση του πλεονασμού και των σφαλμάτων.



# Εργαλεία για την καταγραφή

- **Ταξινόμηση:** Μια ταξινόμηση είναι ένα σύστημα ταξινόμησης. Η παλαιότερη μορφή ταξινόμησης είναι η ταξινόμηση των οργανισμών Linnean. Τυπικά, οι ταξινομίες είναι ιεραρχικές: κάθε όρος «παιδιού» (πιο συγκεκριμένος όρος) έχει έναν «γονικό» όρο (γενικότερο όρο) και ένας γονικός όρος μπορεί να έχει ένα ή περισσότερα παιδιά.
- Οι ταξινομήσεις δεν πρέπει να είναι μόνο ιεραρχίες. Μπορεί επίσης να είναι δίκτυα, στα οποία ένας όρος «παιδιών» μπορεί να έχει έναν ή περισσότερους «γονείς». Το κύριο πλεονέκτημα της χρήσης ταξινόμησης σε μια μη ιεραρχική λίστα είναι ότι μπορείτε να βρείτε όλα τα σχολιασμένα ως υποκατηγορία του όρου αναζήτησης.

Linnean taxonomy : start off with living organism, move to domain, then kingdom, then phylum, class, order, family, genus and finally species.



# Εργαλεία για την καταγραφή

- **Gene Ontology**: Το αρχέτυπο παράδειγμα μιας οντολογίας στις επιστήμες της μοριακής ζωής είναι η Οντολογία των Γονιδίων (GO),
- δημιουργήθηκε και συντηρήθηκε από την Consortium Gene Ontology.
- περιγράφει τη λειτουργία και τον κυτταρικό εντοπισμό γονιδιακών προϊόντων σε όλα τα είδη
- χρησιμοποιείται για να περιγράψει τα γονίδια και τα προϊόντα τους σε μεγάλες δημόσιες βάσεις δεδομένων, συμπεριλαμβανομένων των UniProt, Ensembl
- Παρέχει ένα ισχυρό μέσο ανάλυσης συνόλων δεδομένων.

- **Η δομή των καταχωρήσεων του SWISS-PROT**
- Η UNI-PROT είναι η πιο δημοφιλής βάση δεδομένων για θέματα αναζήτησης λόγω της δομής της και της ποιότητας των σχολίων της. Περιέχει περίπου 100000 καταχωρήσεις από πάνω από περίπου 5000 διαφορετικά είδη.
- Μια καταχώρηση της UNI-PROT έχει την παρακάτω μορφή:

```

ID  OPSD_SHEEP      STANDARD;      PRT;    348 AA.
AC  P02700;
DT  21-JUL-1986 (Rel. 01, Created)
DT  01-FEB-1991 (Rel. 17, Last sequence update)
DT  05-JUL-2004 (Rel. 44, Last annotation update)
DE  Rhodopsin.
GN  Name=RHO;
OS  Ovis aries (Sheep).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
OC  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae
OC  Caprinae; Ovis.
OX  NCBI_TaxID=9940;
RN  [1]
RP  SEQUENCE.
RA  Pappin D.J.C., Elipoulos E., Brett M., Findlay J.B.C.;
RT  "A structural model for ovine rhodopsin.";
RL  Int. J. Biol. Macromol. 6:73-76(1984).
...
CC  -!- FUNCTION: Visual pigments are the light-absorbing molecules
CC      mediate vision. They consist of an apoprotein, opsin, covalen
CC      linked to cis-retinal.
CC  -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC  -!- TISSUE SPECIFICITY: Rod shaped photoreceptor cells which med
CC      vision in dim light.
CC  -!- MISCELLANEOUS: Maximal absorption at 495 nm.
CC  -!- SIMILARITY: Belongs to family 1 of G-protein coupled recepto
CC      Opsin subfamily.
DR  PIR; ; OOSH.
...
DR  Pfam; PF00001; 7tm_1; 1.
DR  PRINTS; PR00237; GPCRRHODOPSN.
DR  PRINTS; PR00238; OPSIN.
DR  PRINTS; PR00579; RHODOPSIN.
DR  PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR  PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
DR  PROSITE; PS00238; OPSIN; 1.
KW  Direct protein sequencing; G-protein coupled receptor; Glycoprote
KW  Lipoprotein; Palmitate; Phosphorylation; Photoreceptor;
KW  Retinal protein; Transmembrane; Vision.
FT  DOMAIN      1      36      Extracellular.
FT  TRANSMEM    37      61      1 (Potential).
FT  DOMAIN      62      73      Cytoplasmic.
FT  TRANSMEM    74      98      2 (Potential).
FT  DOMAIN      99     113      Extracellular.
...
FT  BINDING     296     296      Retinal chromophore.
FT  LIPID       322     322      S-palmitoyl cysteine (By similarity
FT  LIPID       323     323      S-palmitoyl cysteine (By similarity
FT  DISULFID    110     187      By similarity.
FT  MOD_RES     334     334      Phosphoserine (by RK).
FT  MOD_RES     335     335      Phosphothreonine (by RK).
...
SQ  SEQUENCE 348 AA; 38891 MW; AAFD6F0D6A8BAEE5 CRC64;
MNGTEGPNFY VPFSNKTGVV RSPFEAPQYY LAEPWQFSML AAYMFLLLIVL GFPINFLTL
VTVQHKKLRT PLNYILLNLA VADLFMVFGG FTTTLYTS LH GYFVFGPTGC NLEGFFATL
GEIALWLSLV LAIERYVVVC KPMSNFRFGE NHAIMGVAFT WVMALACAAP PLVGWSRYI
QGMQCSCGAL YFTLKPEINN ESFVIYMFVV HFSIPLIVIF FCYGQLVFTV KEAAAQQQE
ATTQKAEKEV TRMVIIMVIA FLICWLPYAG VAFYIFTHQG SDFGPIFMTI PAFFAKSSS
YNPVIYIMMN KQFRNCMLTT LCCGKNPLGD DEASTTVSKT ETSQVAPA

```

//

Μια καταχώρηση του SWISS-PROT έχει την παρακάτω μορφή:

Κάθε γραμμή είναι σημειωμένη με ένα διψήφιοκωδικό, που βοηθάει στην παρουσίαση της πληροφορίας με ένα δομημένο τρόπο.

- Η γραμμή **ID** μας πληροφορεί ότι το όνομα της καταχώρησης είναι OPSD\_SHEEP, μια πρωτεΐνη με 348 αμινοξέα. Ο τύπος του ID είναι PROTEIN\_SOURCE, δηλ. ο τύπος της πρωτεΐνης και η ονομασία του οργανισμού.
- Το **AC** (P02700) είναι ο κωδικός πρόσβασης που συνήθως παραμένει σταθερός μεταξύ των βάσεων δεδομένων. Το AC είναι απαραίτητο διότι μερικές φορές το ID αλλάζει.
- Τα πεδία **DT** δίνουν πληροφορίες πάνω στην ημερομηνία καταχώρησης της σειράς στη βάση δεδομένων, όπως και πότε τροποποιήθηκε.
- Τα πεδία **DE** μας πληροφορούν τις ονομασίες με τις οποίες οι πρωτεΐνες είναι γνωστές, π.χ. rhodopsin.

- Το πεδίο **GN** δίνει την ονομασία του γονίδιου
- Το πεδίο **OS** δίνει το είδος του οργανισμού
- Το πεδίο **OC** δίνει την ταξινόμηση του οργανισμού μέσα στα βιολογικά βασίλεια.
- Τα πεδία **RN, RP, RA, RL** δίνουν αναφορές
- Τα πεδία **CC** παρέχουν σχόλια για τη λειτουργία της πρωτεΐνης, τις μετά-μεταφραστικές αλλαγές (PTM), τον τύπο του ιστού, την υπό-κυτταρική τοποθεσία και την ομοιότητα ή τον δεσμό με συγκεκριμένες οικογένειες πρωτεϊνών.
- Τα πεδία **DR** παρέχουν συνδέσμους με άλλες βιομοριακές βάσεις δεδομένων για τη RHODOPSIN.



- Τα πεδία **KW** παρέχουν λέξεις κλειδιά
- Τα πεδία **FT** σχηματίζουν ένα πίνακα χαρακτηριστικών (Feature Table), με περιοχές στην αλληλουχία, που αντιστοιχούν σε transmembrane domains, ligand binding sites (ενεργό μέρος της πρωτεΐνης) κλπ.
- Τα πεδία **SQ** περιέχουν την καθαυτό σειρά. Ένας κωδικός που αποτελείται από ένα γράμμα του αμινοξέως και κάθε γραμμή περιέχει 60 αμινοξέα. Το **MW** δίνει το μοριακό βάρος.

## Σύνθετες βάσεις δεδομένων αλληλουχιών πρωτεϊνών

- Σε μία σύνθετη βάση δεδομένων συγχωνεύονται διάφορες πρωτογενείς πηγές δεδομένων, καθιστώντας έτσι πιο αποδοτική τη διαδικασία αναζήτησης σειρών.
- Η βάση αυτή είναι μη πλεονάζουσα (non-redundant – τα δεδομένα δεν επαναλαμβάνονται) και επομένως η διαδικασία αναζήτησης μίας σειράς δε χρειάζεται να πραγματοποιηθεί περισσότερες από μία φορές.
- Για τη δημιουργία των σύνθετων βάσεων δεδομένων χρησιμοποιούνται διάφορες στρατηγικές ανάλογα με τις πρωτογενείς πηγές δεδομένων που έχουν επιλεγεί σε κάθε περίπτωση και ανάλογα με τα κριτήρια που χρησιμοποιούνται για τη συγχώνευση των πηγών αυτών.
- Παραδείγματος χάριν, μία στρατηγική είναι η βάση δεδομένων να μην περιέχει πανομοιότυπες αλληλουχίες και οι αλληλουχίες να διαφέρουν μόνο κατά ένα αμινοξύ (residue – στην περίπτωση αυτή η βάση θα είναι πράγματι μη πλεονάζουσα – non-redundant).

Οι κύριες σύνθετες βάσεις δεδομένων είναι οι ακόλουθες:

	<u>NRDB</u>	<u>OWL</u>	<u>MIPSX</u>	<u>SP+TrEMBL</u>
Πρωτογενείς πηγές δεδομένων	PDB SWISS-PROT PIR GenPept SWISS-PROTupdate GenPeptupdate	SWISS-PROT PIR GenBank NRL-3D	PIR MIPSOwn MIPSTrn MIPSH PIRMOD NRL-3D SWISS-PROT EMTrans GBTrans Kabat PseqIP	SWISS-PROT TrEMBL

## NRDB

- Η NRDB (Non-Redundant Database ) δημιουργήθηκε στο NCBI. Πρόκειται για μία βάση δεδομένων με εκτενή περιγραφή (έχει όλο το annotation της πρωτεΐνης) η οποία ενημερώνεται συνεχώς.
- Για τη δημιουργία της αφαιρέθηκαν από τις πρωταρχικές πηγές δεδομένων μόνο τα πανομοιότυπα αντίτυπα των σειρών (επομένως η NRDB είναι μη-πανομοιότυπη, όχι όμως και μη-πλεονάζουσα). Αυτό σημαίνει ότι η βάση περιλαμβάνει πολλά αντίτυπα της ίδιας πρωτεΐνης, τα οποία προκύπτουν ως αποτέλεσμα είτε πολυμορφισμών είτε ελάχιστων λαθών κατά την αλληλούχηση (sequencing).
- Επιπλέον, λανθασμένες σειρές οι οποίες είχαν διορθωθεί στην SWISS-PROT τώρα επαναλαμβάνονται κατά τη διαδικασία της μετάφρασης από το DNA.

Welcome to PIR [Protein Info...]

Search & Tools [PIR - Protein...]

MIPS: a database for prote...

Why is UniProtKB compo...

Google Translate

Non-redundant databases (NR...

← → ↻ 🏠

🔒 https://pubmlst.org/analysis/nrdb.shtml

⋮ 🛡️ ☆

🔍 📄 📑

⚙️ Συχνές επισκέψεις 📄 Ξεκινώντας 🌐 Suggested Sites 🌐 Web Slice Gallery 🌐 YPOMEDI

PubMLST Databases Downloads BIGSdb Contact Account

🔍 Google Custom Search 🔍 Search

### Non-redundant databases

This is a Javascript version of NRDB written by Keith Jolley. It was inspired by the original NRDB written by Warren Gish.

Paste in a FASTA file and any duplicate sequences will be identified and a new FASTA output generated containing the non-redundant sequences. You can choose to optionally renumber these.

Paste in FASTA file

Options

☐ Sequentially renumber

Prefix id with:

Action

Submit

EN 🚩 11:37 AM  
4/12/2019

## OWL

- Η OWL δημιουργήθηκε στο Πανεπιστήμιο του Leeds και είναι μία μη-πλεονάζουσα βάση δεδομένων. Οι πρωτογενείς πηγές δεδομένων της OWL έχουν ιεραρχηθεί και τους έχει αποδοθεί σειρά προτεραιότητας με βάση το επίπεδο περιγραφής (annotation) των πρωτεϊνών, καθώς και την αξιολόγηση των αλληλουχιών που περιλαμβάνουν.
- Η SWISS-PROT είναι η βάση με την υψηλότερη προτεραιότητα, επομένως οι υπόλοιπες πηγές συγκρίνονται με αυτή κατά τη διαδικασία της συγχώνευσης. Δεν περιλαμβάνει πανομοιότυπες αλληλουχίες, ούτε αλληλουχίες οι οποίες διαφέρουν μόνο κατά ένα αμινοξύ.

Welcome to PIR [Protein Info] X Search & Tools [PIR - Protein] X MIPS: a database for prote X Why is UniProtKB compo X Google Translate X OWL X

130.88.97.239/OWL/

Συχνές επισκέψεις m Εκκινώντας Suggested Sites Web Slice Gallery YPOMEDI

Home Databases Services and Tools EU Projects Education Research Group Videos Societies

OWL  
Composite Protein  
Sequence Database

OWL is a non-redundant composite of 4 publicly-available primary sources: [SWISS-PROT](#), [PIR \(1-3\)](#), [GenBank](#) (translation) and [NRL-3D](#). SWISS-PROT is the highest priority source, all others being compared against it to eliminate identical and trivially-different sequences. The strict redundancy criteria render OWL relatively "small" and hence efficient in similarity searches. [References](#).

**Direct OWL access**

- ♦ [By accession number](#)
- ♦ [By database code](#)
- ♦ [By text](#)
- ♦ [By sequence](#)
- ♦ [By title](#)
- ♦ [By author](#)
- ♦ [By query language](#)

**Available Documents**

- ♦ [Primary source contributions](#)
- ♦ [OWL contents](#)
- ♦ [OWL release notes](#)
- ♦ [How to obtain OWL](#)
- ♦ [DELPHOS manual](#)

**Disclaimer**

- ♦ [Here](#).

EN 11:38 AM 4/12/2019



## **MISPX**

- Η MIPSX δημιουργήθηκε στο Max-Planck Institut στο Martinsried. Στις πρωτογενείς πηγές της αποδίδεται προτεραιότητα βάσει της σειράς τους όπως αυτή παρουσιάζεται στον άνωθεν πίνακα.
- Σε περίπτωση που βρεθούν πανομοιότυπες αλληλουχίες είτε μέσα σε κάποια βάση, ή ανάμεσα σε δύο ή περισσότερες βάσεις, τότε οι αλληλουχίες αυτές διαγράφονται αφήνοντας μοναδικά αντίτυπά τους στην MIPSX.
- Επίσης διαγράφονται και οι αλληλουχίες οι οποίες είναι εξ' ολοκλήρου ενσωματωμένες μέσα σε άλλες αλληλουχίες (υπό-αλληλουχίες – subsequences).

## **SWISS-PROT+TrEMBL**

- Η SWISS-PROT+TrEMBL δημιουργήθηκε στο EBI και είναι μία βάση δεδομένων εκτενούς περιγραφής και μη-πλεονάζουσα.
- Εκτιμάται ότι λιγότερο από 30% των δεδομένων του ολικού συνόλου των δεδομένων της SWISS-PROT και της TrEMBL ήταν μη-μοναδικά.

Ανοίξτε το Internet Explorer για πρόσβαση στο SWISS-PROT, γράφοντας τη διεύθυνση <http://www.uniprot.org/>. Οπότε θα εμφανισθεί η παρακάτω σελίδα, κυλήστε την προς τα κάτω.

The screenshot shows the UniProt website homepage in Internet Explorer. The browser's address bar displays <https://www.uniprot.org>. The page features a blue header with the UniProt logo and a search bar. Below the header, a navigation bar includes links for BLAST, Align, Retrieve/ID mapping, Peptide search, Help, and Contact. The main content area is divided into several sections: UniProtKB (Swiss-Prot and TrEMBL), UniRef (Sequence clusters), UniParc (Sequence archive), and Proteomes (Proteome sets). A large purple section titled 'Supporting data' contains links for Literature citations, Cross-ref. databases, Taxonomy, Diseases, Subcellular locations, and Keywords. On the right, a 'News' section highlights 'Forthcoming changes' and 'Planned changes for UniProt', including 'UniProt release 2020\_01' and 'UniProt release 2019\_11'. A red banner at the top right of the Proteomes section reads 'View SARS-CoV-2 Proteins and Receptors'. The bottom right corner of the browser window shows a status bar with the text 'VODAFONE\_WIFI108 Πρόσβαση στο Internet'.

UniProt

UniProtKB

Advanced Search

BLAST Align Retrieve/ID mapping Peptide search Help Contact

The mission of UniProt is to provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information.

**UniProtKB**  
UniProt Knowledgebase

**Swiss-Prot (561,911)**  
Manually annotated and reviewed.  
Records with information extracted from literature and curator-evaluated computational analysis.

**TrEMBL (177,754,527)**  
Automatically annotated and not reviewed.  
Records that await full manual annotation.

**UniRef**  
Sequence clusters

**UniParc**  
Sequence archive

**Proteomes**  
Proteome sets

**Supporting data**

Literature citations  
Cross-ref. databases

Taxonomy  
Diseases

Subcellular locations  
Keywords

New UniProt portal for the latest SARS-CoV-2 coronavirus protein entries and receptors, updated independent of the general UniProt release cycle.  
**View SARS-CoV-2 Proteins and Receptors**

**News**

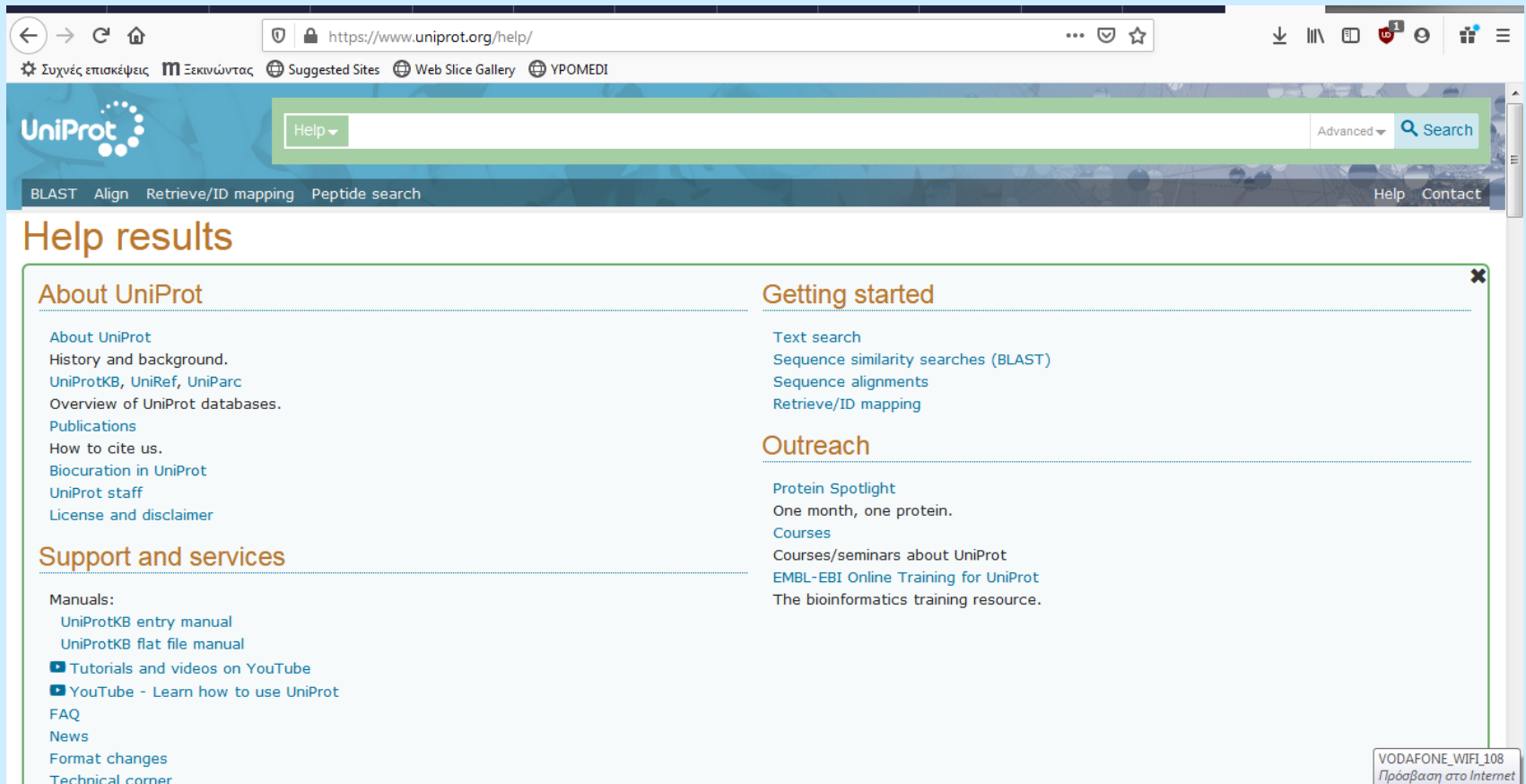
**Forthcoming changes**  
Planned changes for UniProt

**UniProt release 2020\_01**  
Coronavirus SARS-CoV-2 in UniProtKB | Changes to UniProt release cycle

**UniProt release 2019\_11**  
Thicker than water | Functional annotation of different gene products | Changes to FT and CC text format | Cross-references to RNaAct | Pr...

VODAFONE\_WIFI108  
Πρόσβαση στο Internet

Μπορείς να μάθεις όλες τις λεπτομέρειες για τη βάση δεδομένων SWISS-PROT επιλέγοντας το “UniProtKB Manual”, τότε θα εμφανισθεί η παρακάτω σελίδα



The screenshot shows the UniProtKB Manual page in a web browser. The browser's address bar displays 'https://www.uniprot.org/help/'. The page features a blue header with the UniProt logo and a search bar. Below the header, there are navigation links for 'BLAST', 'Align', 'Retrieve/ID mapping', and 'Peptide search'. The main content area is titled 'Help results' and is divided into three columns. The left column contains links for 'About UniProt' (including history, databases, and publications) and 'Support and services' (including manuals, tutorials, and FAQ). The middle column contains links for 'Getting started' (including text search and BLAST) and 'Outreach' (including protein spotlight and courses). The right column contains a 'Help' link and a 'Contact' link. A small notification in the bottom right corner indicates 'VODAFONE\_WIFI\_108' and 'Πρόσβαση στο Internet'.

UniProt

Help

Advanced Search

BLAST Align Retrieve/ID mapping Peptide search Help Contact

## Help results

### About UniProt

- About UniProt
- History and background.
- UniProtKB, UniRef, UniParc
- Overview of UniProt databases.
- Publications
- How to cite us.
- Biocuration in UniProt
- UniProt staff
- License and disclaimer

### Support and services

- Manuals:
  - UniProtKB entry manual
  - UniProtKB flat file manual
- Tutorials and videos on YouTube
- YouTube - Learn how to use UniProt
- FAQ
- News
- Format changes
- Technical corner

### Getting started

- Text search
- Sequence similarity searches (BLAST)
- Sequence alignments
- Retrieve/ID mapping

### Outreach

- Protein Spotlight
- One month, one protein.
- Courses
- Courses/seminars about UniProt
- EMBL-EBI Online Training for UniProt
- The bioinformatics training resource.


VODAFONE\_WIFI\_108  
Πρόσβαση στο Internet

Αρχείο Επεξεργασία Προβολή Ιστορικό Σελιδοδείκτες Εργαλεία Βοήθεια

RHO - Rhodopsin - Ovis ari... CATH Superfamily 4.10.840... \* in UniProtKB \* in UniProtKB

www.uniprot.org/uniprot/P02700

Πιο συχνά αναγνωσμέ... Ξεκινώντας Θέση ειδικότητας αιμα... Suggested Sites Web Slice Gallery



UniProt

BLAST Align Retrieve/ID mapping

## P02700 - OPS

Protein  
Gene  
Organism  
Status

Display None

- ☒ Function
- ☒ Names & Taxonomy
- ☒ Subcellular location
- ☐ Pathology & Biotech
- ☒ PTM / Processing
- ☒ Expression
- ☒ Interaction

### UniProtKB

Protein knowledgebase

### UniRef

Sequence clusters

### UniParc

Sequence archive

### Proteomes

Protein sets from fully sequenced genomes

### Supporting data

Select one of the options below to target your search:

- Literature citations
- Taxonomy
- Keywords
- Subcellular locations
- Cross-referenced databases
- Human diseases
- Annotation programs

### Help

Help pages, FAQs, UniProtKB manual, documents, news archive, etc.

Advanced Search

Help Contact

Basket

Feedback Help video

### Function

Photoreceptor required for image-forming vision at low light intensity. Required for photoreceptor cell viability after birth. Light-induced isomerization of 11-cis to all-trans retinal triggers a conformational change leading to G-protein activation and release of all-trans retinal (By similarity).

**Absorption**  
Abs(max)=495 nm

**Sites**

Feature key	Position(s)	Length	Description	Graphical view	Feature identifier	Actions
Metal binding <sup>i</sup>	201 - 201	1	Zinc <a href="#">By similarity</a>			

EN 11:27 PM 2/25/2015



Για να δείτε τη δομή της καταχώρησης για την πρωτεΐνη rhodopsin, πληκτρολογήστε rhodopsin στο πεδίο δίπλα στο ID και επιλέξτε το “Search”.

The screenshot shows the UniProtKB search results for 'rhodopsin'. The browser address bar shows the URL: <https://www.uniprot.org/uniprot/?query=rhodopsin&sort=score>. The search bar contains 'rhodopsin' and the 'Search' button is highlighted. The page title is 'UniProtKB results'. A banner explains that UniProtKB consists of two sections: Reviewed (Swiss-Prot) - Manually annotated and Unreviewed (TrEMBL) - Computationally analyzed. Below this, there are filters for 'Filter by' (Reviewed (512), Unreviewed (49,168)) and 'Popular organisms' (Human (111), Fruit fly (99)). A table of results is displayed, showing the first four entries. The table has columns for Entry, Entry name, Protein names, Gene names, Organism, and Length. The first entry is P02699, OPSD\_BOVIN, Rhodopsin, RHO, Bos taurus (Bovine), 348. The second entry is P08100, OPSD\_HUMAN, Rhodopsin, RHO OPN2, Homo sapiens (Human), 348. The third entry is P28327, GRK1\_BOVIN, Rhodopsin kinase GRK1, GRK1 RHOK, Bos taurus (Bovine), 561. The fourth entry is Q15835, GRK1\_HUMAN, Rhodopsin kinase GRK1, GRK1 RHOK, Homo sapiens (Human), 561. A notification in the bottom right corner says 'VODAFONE\_WIFI\_108 Πρόσβαση στο Internet'.

UniProtKB results

UniProtKB consists of two sections:

- Reviewed (Swiss-Prot) - Manually annotated**  
Records with information extracted from literature and curator-evaluated computational analysis.
- Unreviewed (TrEMBL) - Computationally analyzed**  
Records that await full manual annotation.

The UniProt Knowledgebase (UniProtKB) is the central hub for the collection of functional information on proteins, with accurate, consistent and rich annotation. In addition to capturing the core data mandatory for each UniProtKB entry (mainly, the amino acid sequence, protein name or description, taxonomic data and citation information), as much annotation information as possible is added.

Help UniProtKB help video Other tutorials and videos Downloads

Filter by:

- Reviewed (512) Swiss-Prot
- Unreviewed (49,168) TrEMBL

Popular organisms:

- Human (111)
- Fruit fly (99)

Entry	Entry name	Protein names	Gene names	Organism	Length
P02699	OPSD_BOVIN	Rhodopsin	RHO	Bos taurus (Bovine)	348
P08100	OPSD_HUMAN	Rhodopsin	RHO OPN2	Homo sapiens (Human)	348
P28327	GRK1_BOVIN	Rhodopsin kinase GRK1	GRK1 RHOK	Bos taurus (Bovine)	561
Q15835	GRK1_HUMAN	Rhodopsin kinase GRK1	GRK1 RHOK	Homo sapiens (Human)	561

1 to 25 of 49,680 Show 25

VODAFONE\_WIFI\_108  
Πρόσβαση στο Internet

# Επιλέγουμε το OPSD\_SHEEP

← → ↻ 🏠 <https://www.uniprot.org/uniprot/?query=rhodopsin&sort=score> ⚙️ 📄 📁 📧 📧 📧 📧 📧 📧

⚙️ Συχνές επισκέψεις 📄 Εκκινώντας 🌐 Suggested Sites 🌐 Web Slice Gallery 🌐 YPOMEDI

**View by**

- Results table
- Taxonomy
- Keywords
- Gene Ontology
- Enzyme class
- Pathway

**UniRef**

Your results in sequence clusters with identity of: 100%, 90% or 50%

**Demo**

- Help video

<input type="checkbox"/>	Q18700	OPSD_PIG		Rhodopsin	RHO RHO1	Sus scrofa (Pig)	348
<input type="checkbox"/>	P22328	OPSD_CHICK		Rhodopsin	RHO	Gallus gallus (Chicken)	351
<input type="checkbox"/>	P18545	CNRG_HUMAN		Retinal rod rhodopsin-sensitive cGM...	PDE6G PDEG	Homo sapiens (Human)	87
<input type="checkbox"/>	P15425	CYPR_DROME		Peptidyl-prolyl cis-trans isomerase...	ninaA CG3966	Drosophila melanogaster (Fruit fly)	237
<input type="checkbox"/>	Q68J47	OPSD_LOXAF		Rhodopsin	RHO RH1	Loxodonta africana (African elephant)	348
<input type="checkbox"/>	Q08K01	OPSD_FELCA		Rhodopsin	RHO OPS	Felis catus (Cat) (Felis silvestris catus)	348
<input checked="" type="checkbox"/>	P02700	OPSD_SHEEP		Rhodopsin	RHO	Ovis aries (Sheep)	348
<input type="checkbox"/>	O43924	PDE6D_HUMAN		Retinal rod rhodopsin-sensitive cGM...	PDE6D PDED	Homo sapiens (Human)	150
<input type="checkbox"/>	O01668	OPS6_DROME		Opsin Rh6	Rh6 CG5192	Drosophila melanogaster (Fruit fly)	369
<input type="checkbox"/>	P49912	OPSD_RABIT		Rhodopsin	RHO	Oryctolagus cuniculus (Rabbit)	348
<input type="checkbox"/>	Q8WP15	GRK7_PIG		Rhodopsin kinase GRK7	GRK7	Sus scrofa (Pig)	553
<input type="checkbox"/>	Q8HY69	OPSD_SMICR		Rhodopsin	RHO RH1	Sminthopsis crassicaudata (Fat-tailed dunnart) (Phascogale crassicaudata)	348

◀ 1 to 25 of 49,680 ▶ Show 25 ▼

**Tools** **Core data** **Supporting data** **Information**

BLAST Protein knowledgebase (UniProtKB) Literature citations About UniProt

VODAFONE\_WIFI\_108  
Πρόσβαση στο Internet



Τότε θα εμφανιστεί ολόκληρη η εκχώρηση του OPSD\_SHEEP. Προχωρήστε και στο τέλος της σελίδας για να δείτε όλες τις πληροφορίες.

Browser address bar: <https://www.uniprot.org/uniprot/P02700>

Navigation: BLAST Align Retrieve/ID mapping Peptide search Help Contact

## UniProtKB - P02700 (OPSD\_SHEEP)

[Basket](#)

**Display**

Entry Publications Feature viewer Feature table

**Protein** Rhodopsin  
**Gene** RHO  
**Organism** *Ovis aries* (Sheep)  
**Status** Reviewed - Annotation score: ●●●●● - Experimental evidence at protein level<sup>i</sup>

**Function<sup>i</sup>**

Photoreceptor required for image-forming vision at low light intensity. Required for photoreceptor cell viability after birth (By similarity). Light-induced isomerization of 11-cis to all-trans retinal triggers a conformational change that activates signaling via G-proteins. Subsequent receptor phosphorylation mediates displacement of the bound G-protein alpha subunit by the arrestin SAG and terminates signaling (By similarity). [By similarity](#)

**Sites**

Feature key	Position(s)	Description	Actions	Graphical view	Length
Site <sup>i</sup>	113	Plays an important role in the conformation switch to the active conformation <a href="#">By similarity</a>			1
Metal binding <sup>i</sup>	201	Zinc <a href="#">By similarity</a>			1
Metal binding <sup>i</sup>	279	Zinc <a href="#">By similarity</a>			1

**GO - Molecular function<sup>i</sup>**

- 11-cis retinal binding [Source: UniProtKB](#)

VODAFONE\_WIFI\_108  
Πρόσβαση στο Internet

Τότε θα εμφανιστεί ολόκληρη η εκχώρηση του OPSD\_SHEEP. Προχωρήστε και στο τέλος της σελίδας για να δείτε όλες τις πληροφορίες.

The screenshot shows the UniProt entry for P02700 (Rhodopsin). The left sidebar contains navigation links: Entry, Publications, Feature viewer, and Feature table. The main content area is divided into several sections:

- GO - Molecular function<sup>i</sup>**
  - 11-cis retinal binding (Source: UniProtKB)
  - G protein-coupled photoreceptor activity (Source: UniProtKB)
  - metal ion binding (Source: UniProtKB-KW)

Complete GO annotation on QuickGO ...
- GO - Biological process<sup>i</sup>**
  - absorption of visible light (Source: AgBase)
  - protein-chromophore linkage (Source: UniProtKB-KW)
  - rhodopsin mediated signaling pathway (Source: UniProtKB)
  - visual perception (Source: UniProtKB-KW)

Complete GO annotation on QuickGO ...
- Keywords<sup>i</sup>**

Molecular function	G-protein coupled receptor, Photoreceptor protein, Receptor, Retinal protein, Transducer
Biological process	Sensory transduction, Vision
Ligand	Chromophore, Metal-binding, Zinc
- Names & Taxonomy<sup>i</sup>**

Protein names <sup>i</sup>	Recommended name: <b>Rhodopsin</b>
Gene names <sup>i</sup>	Name: <b>RHO</b>
Organism <sup>i</sup>	<i>Ovis aries</i> (Sheep)
Taxonomic identifier <sup>i</sup>	9940 [NCBI]

At the bottom right, a status bar indicates "VODAFONE\_WIFI\_108" and "Πρόσβαση στο Internet".

Τότε θα εμφανιστεί ολόκληρη η εκχώρηση του OPSD\_SHEEP. Προχωρήστε και στο τέλος της σελίδας για να δείτε όλες τις πληροφορίες.

The screenshot displays the UniProt entry for Rhodopsin (P02700). The 'Names & Taxonomy' section is highlighted with a red oval. The 'Subcellular location' section is also visible, showing the protein's location in the photoreceptor outer segment.

**Names & Taxonomy**

Protein names	Recommended name: <b>Rhodopsin</b>
Gene names	Name: <b>RHO</b>
Organism	<i>Ovis aries</i> (Sheep)
Taxonomic identifier	9940 [NCBI]
Taxonomic lineage	Eukaryota > Metazoa > Chordata > Craniata > Vertebrata > Euteleostomi > Mammalia > Eutheria > Laurasiatheria > Artiodactyla > Ruminantia > Pecora > Bovidae > Caprinae > <i>Ovis</i>
Proteomes	UP000002356 Component: Unplaced

**Subcellular location**

UniProt annotation GO - Cellular component

**Other locations**

- Membrane **2 Publications**; Multi-pass membrane protein
- photoreceptor outer segment **2 Publications**

**Note:** Synthesized in the inner segment (IS) of rod photoreceptor cells before vectorial transport to disk membranes in the rod outer segment (OS) photosensory cilia. **By similarity**

Τότε θα εμφανιστεί ολόκληρη η εκχώρηση του OPSD\_SHEEP. Προχωρήστε και στο τέλος της σελίδας για να δείτε όλες τις πληροφορίες.

www.uniprot.org/uniprot/P02700

Πιο συχνά αναγνώσιμ... Εικονίδιος... Θέση ειδικότητας αιμα... Suggested Sites Web Slice Gallery

### Display

None

- ☒ Function
- ☒ Names & Taxonomy
- ☒ Subcellular location
- ☐ Pathology & Biotech
- ☒ PTM / Processing
- ☒ Expression
- ☒ Interaction
- ☒ Structure
- ☒ Family & Domains
- ☒ Sequence
- ☒ Cross-references
- ☒ Publications
- ☒ Entry information
- ☒ Miscellaneous
- ☒ Similar proteins

▲ Top

### Family and domain databases

Gene3D <sup>i</sup>	4.10.840.10. 1 hit.
InterPro <sup>i</sup>	IPR000276. GPCR_Rhodpsn. IPR017452. GPCR_Rhodpsn_7TM. IPR001760. Opsin. IPR027430. Retinal_BS. IPR000732. Rhodopsin. IPR019477. Rhodopsin_N. [Graphical view]
Pfam <sup>i</sup>	PF00001. 7tm_1. 1 hit. PF10413. Rhodopsin_N. 1 hit. [Graphical view]
PRINTS <sup>i</sup>	PR00237. GPCRRHODOPSN. PR00238. OPSIN. PR00579. RHODOPSIN.
PROSITE <sup>i</sup>	PS00237. G_PROTEIN_RECEP_F1_1. 1 hit. PS00262. G_PROTEIN_RECEP_F1_2. 1 hit. PS00238. OPSIN. 1 hit. [Graphical view]

### Sequence<sup>i</sup>

Sequence status<sup>i</sup>: Complete.

P02700-1 [UniParc] [FASTA](#) [Add to Basket](#)

« Hide

**Length:** 348  
**Mass (Da):** 38,892  
**Last modified:** February 1, 1991 - v2



www.uniprot.org/uniprot/P02700

uniprot

Πιο συχνά αναγνωσμέ...

Εκκινώντας

Θεση ειδικότητας αιμα...

Suggested Sites

Web Slice Gallery

Display

None

☒Function

☒Names & Taxonomy

☒Subcellular location

☐Pathology & Biotech

☒PTM / Processing

☒Expression

☒Interaction

☒Structure

☒Family & Domains

☒Sequence

☒Cross-references

☒Publications

☒Entry information

☒Miscellaneous

☒Similar proteins

▲Top

1. "A structural model for ovine rhodopsin."  
Pappin D.J.C., Elipoulos E., Brett M., Findlay J.B.C.  
Int. J. Biol. Macromol. 6:73-76(1984)  
Cited for: PROTEIN SEQUENCE.

2. "Isolation and characterization of the CNBr peptides from the proteolytically derived N-terminal fragment of ovine opsin."  
Brett M., Findlay J.B.C.  
Biochem. J. 211:661-670(1983) [PubMed] [Europe PMC] [Abstract]  
Cited for: PROTEIN SEQUENCE OF 1-111 AND 144-239.

3. "Primary structure of C-terminal functional sites in ovine rhodopsin."  
Findlay J.B.C., Brett M., Pappin D.J.C.  
Nature 293:314-316(1981) [PubMed] [Europe PMC] [Abstract]  
Cited for: PROTEIN SEQUENCE OF 240-348.

4. "Sequence variability in the retinal-attachment domain of mammalian rhodopsins."  
Pappin D.J.C., Findlay J.B.C.  
Biochem. J. 217:605-613(1984) [PubMed] [Europe PMC] [Abstract]  
Cited for: RETINAL-BINDING SITE AT LYS-296.

5. "Phosphorylation of ovine rhodopsin. Identification of the phosphorylated sites."  
Thompson P., Findlay J.B.C.  
Biochem. J. 220:773-780(1984) [PubMed] [Europe PMC] [Abstract]  
Cited for: PHOSPHORYLATION AT SER-334; THR-335; THR-336; SER-338 AND SER-343.

Entry information<sup>i</sup>

Entry name<sup>i</sup>

OPSD\_SHEEP

Accession<sup>i</sup>

Primary (citable) accession number: **P02700**

Entry history<sup>i</sup>

Integrated into UniProtKB/Swiss-Prot: July 21, 1986

# FASTA Μορφή



A screenshot of a web browser displaying a FASTA file from the Uniprot database. The browser's address bar shows the URL <https://www.uniprot.org/uniprot/P02700.fasta>. Below the address bar, there is a navigation bar with icons for settings, search, and suggested sites. The main content area displays the FASTA file's header and sequence. The header line is `>sp|P02700|OPSD_SHEEP Rhodopsin OS=Ovis aries OX=9940 GN=RHO PE=1 SV=2`. The sequence is a single line of amino acid residues: `MNGTEGPNFYVPFSNKTGVVRSPFEAPQYYLAEPWFQFSLAAYMFLILVLGFPINFLTLTYVTVQHKKLRITPLNYILLNLAVADLFMVFGGFTTLYISLHGYFVFGPTGCNLEGFFATLGGIALLWSLVVLAIERVYVVCKPMSNFRFGENHAIMGVAFTWVMALACAAPPLVGWSRYIPQGMQCSGALYFTLKPEINNESFVIYMFVVHFSIPLIVIFFCYGQLVFTVKEAAQQQESATTQKAEKEVTRMVIIMVIAFLICWLPYAGVAFYIFTHQGSDFGPIFMTIPAFFAKSSSVYNPVIYIMMNKQFRNCMLTTLCCGKNPLGDDEASTTVSKTETSQVAPA`.

```
>sp|P02700|OPSD_SHEEP Rhodopsin OS=Ovis aries OX=9940 GN=RHO PE=1 SV=2
MNGTEGPNFYVPFSNKTGVVRSPFEAPQYYLAEPWFQFSLAAYMFLILVLGFPINFLTLTY
VTVQHKKLRITPLNYILLNLAVADLFMVFGGFTTLYISLHGYFVFGPTGCNLEGFFATLGG
IALLWSLVVLAIERVYVVCKPMSNFRFGENHAIMGVAFTWVMALACAAPPLVGWSRYIP
QGMQCSGALYFTLKPEINNESFVIYMFVVHFSIPLIVIFFCYGQLVFTVKEAAQQQES
ATTQKAEKEVTRMVIIMVIAFLICWLPYAGVAFYIFTHQGSDFGPIFMTIPAFFAKSSSV
YNPVIYIMMNKQFRNCMLTTLCCGKNPLGDDEASTTVSKTETSQVAPA
```

# FASTA Μορφή

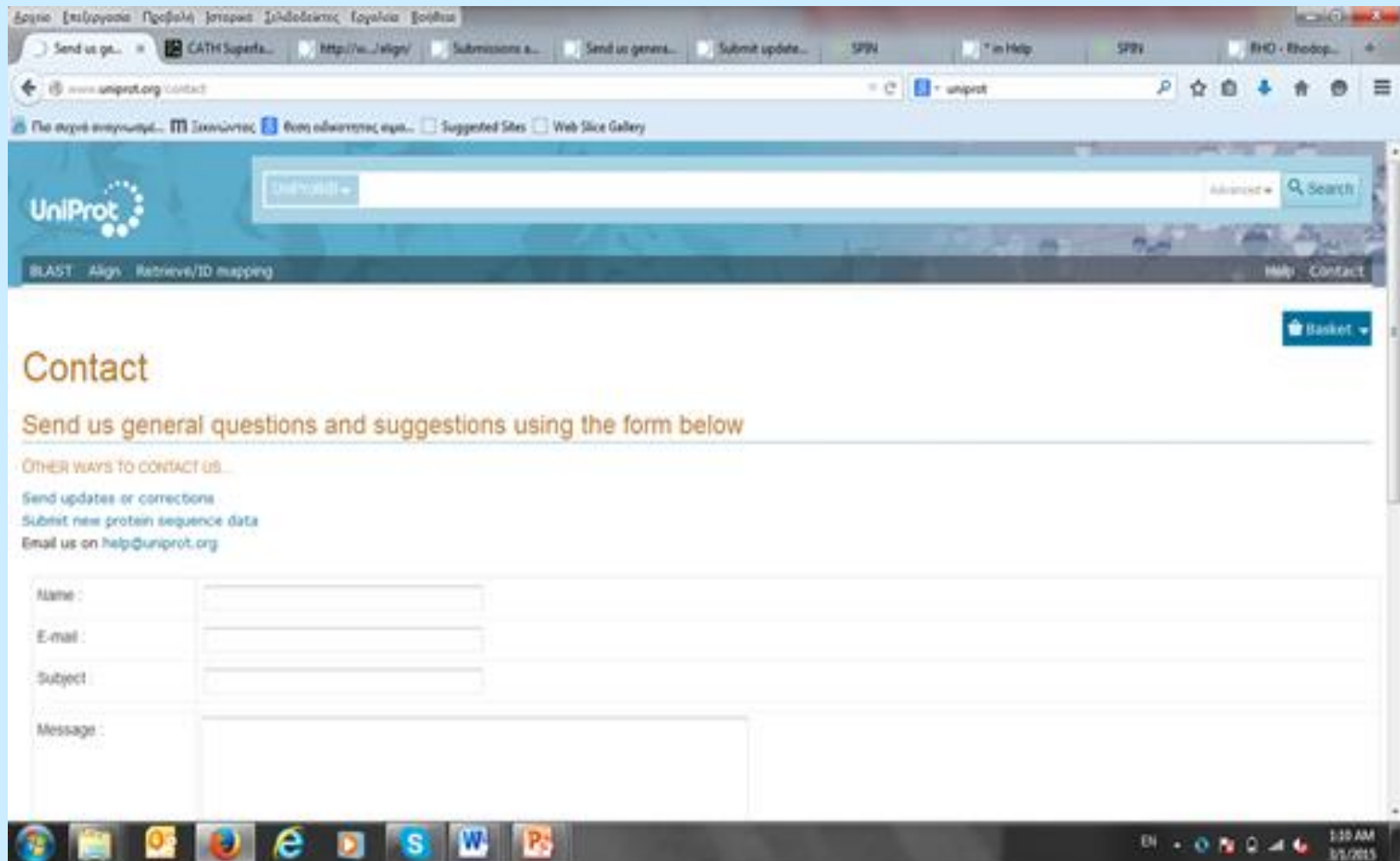
Η αλληλουχία σε μορφή FASTA format αποτελείται από:

- Μία γραμμή που ξεκινάει με το σημείο ">", ακολουθούμενο από τον κωδικό ταυτοποίησης της αλληλουχίας.
- Προαιρετικά ακολουθείται από την περιγραφή με μορφή κειμένου της αλληλουχίας.
- Μία ή περισσότερες γραμμές με την αλληλουχία.
- Ένα αρχείο σε μορφή FASTA μπορεί να αποτελείται από περισσότερες από μία αλληλουχίες.
- Η μορφή FASTA μερικές φορές αναφέρεται και ως μορφή "Pearson" (το όνομα του δημιουργού του προγράμματος FASTA και της αντίστοιχης μορφής).



# Ενημέρωση δεδομένων στο SWISS-PROT

Αν θέλεις να ενημερώσεις τα δεδομένα που σχετίζονται με μια ήδη υπάρχουσα πρωτεΐνη, επιλέγεις “Contact” από την αρχική σελίδα της Uni-prot.



The screenshot shows the UniProt website's 'Contact' page. The browser's address bar displays 'http://www.uniprot.org/contact'. The page features the UniProt logo and a search bar at the top. Below the logo, there are links for 'BLAST', 'Align', 'Retrieve/ID mapping', 'Help', and 'Contact'. A 'Basket' button is visible in the top right corner. The main heading is 'Contact', followed by the instruction 'Send us general questions and suggestions using the form below'. Underneath, there is a section titled 'OTHER WAYS TO CONTACT US...' with links for 'Send updates or corrections', 'Submit new protein sequence data', and 'Email us on help@uniprot.org'. A contact form is located at the bottom, with fields for 'Name', 'E-mail', 'Subject', and a larger 'Message' text area. The Windows taskbar at the bottom shows various application icons and the system clock indicating 1:10 AM on 3/1/2015.

UniProt

BLAST Align Retrieve/ID mapping Help CONTACT

Basket

## Contact

Send us general questions and suggestions using the form below

OTHER WAYS TO CONTACT US...

[Send updates or corrections](#)  
[Submit new protein sequence data](#)  
[Email us on help@uniprot.org](#)

Name :

E-mail :

Subject :

Message :

Στη συνέχεια επιλέγουμε “Send updates or corrections”.

The screenshot shows a web browser window with the UniProt website. The browser's address bar displays [www.uniprot.org/update](http://www.uniprot.org/update). The page features the UniProt logo and a search bar at the top. Below the logo, there are navigation links: BLAST, Align, Retrieve/ID mapping, Help, and Contact. A 'Basket' button is visible on the right side. The main heading is 'Contact', followed by the subheading 'Submit updates or corrections to UniProt'. The text on the page states: 'We are **actively seeking** any type of updates and/or corrections whether they have been published or not. Please see the guidelines for submitting updates.' Below this, there is a section titled 'OTHER WAYS TO CONTACT US...' with links to 'Send us general questions and suggestions', 'Submit new protein sequence data', and 'Email us on [help@uniprot.org](mailto:help@uniprot.org)'. At the bottom, there is a contact form with fields for 'Name', 'E-mail', and 'Subject'. The 'Subject' field is pre-filled with 'UniProt entry update request'. The Windows taskbar at the bottom shows various application icons and the system clock indicating 1:08 AM on 3/5/2015.

UniProt

Submit updates or corrections to UniProt

We are **actively seeking** any type of updates and/or corrections whether they have been published or not.  
Please see the guidelines for submitting updates.

OTHER WAYS TO CONTACT US...

Send us general questions and suggestions  
Submit new protein sequence data  
Email us on [help@uniprot.org](mailto:help@uniprot.org)

Name:   
E-mail:   
Subject: UniProt entry update request

Αφού συμπληρώσεις τη φόρμα επικοινωνίας επιλέγεις  
“SUBMIT”.

The image shows a screenshot of the UniProt website's contact form. The browser window has a single tab titled 'www.uniprot.org Contact'. The address bar shows 'http://www.uniprot.org'. The page has a light blue header with navigation links: 'Επιλογές', 'Προβλή', 'Στοιχεία', 'Εργαλεία', and 'Βοήθεια'. Below the header, there are several links: 'Send us gen...', 'CATH Superfa...', 'http://www.uniprot.org', 'Submissions a...', 'Send us gen...', 'Submit update...', 'SPIN', 'in Help', 'SPIN', and 'Bio - Rhodop...'. The main content area is a contact form with a 'Subject' field, a 'Message' text area, and a 'Submit' button. Below the form, there are four columns of links: 'Tools' (BLAST, Align, Retrieve/ID mapping), 'Core data' (Protein knowledgebase (UniProtKB), Sequence clusters (UniRef), Sequence archive (UniParc), Proteomes), 'Supporting data' (Literature citations, Taxonomy, Keywords, Subcellular locations, Cross-referenced databases, Diseases), and 'Information' (About UniProt, Help, FAQ, UniProtKB manual, Technical corner, Annotation programs). At the bottom, there is a copyright notice: '© 2002–2013 UniProt Consortium | License & Disclaimer', and logos for EMBL-EBI, PIR, and SIB. The Windows taskbar at the bottom shows various application icons and the system clock indicating 1:11 AM on 3/5/2013.

www.uniprot.org Contact

Send us gen... CATH Superfa... http://www.uniprot.org Submissions a... Send us gen... Submit update... SPIN in Help SPIN Bio - Rhodop...

Παράθεμα επικοινωνίας... Στοιχεία... Εργαλεία... Βοήθεια

Subject

Message

Submit

**Tools**

- BLAST
- Align
- Retrieve/ID mapping

**Core data**

- Protein knowledgebase (UniProtKB)
- Sequence clusters (UniRef)
- Sequence archive (UniParc)
- Proteomes

**Supporting data**

- Literature citations
- Taxonomy
- Keywords
- Subcellular locations
- Cross-referenced databases
- Diseases

**Information**

- About UniProt
- Help
- FAQ
- UniProtKB manual
- Technical corner
- Annotation programs

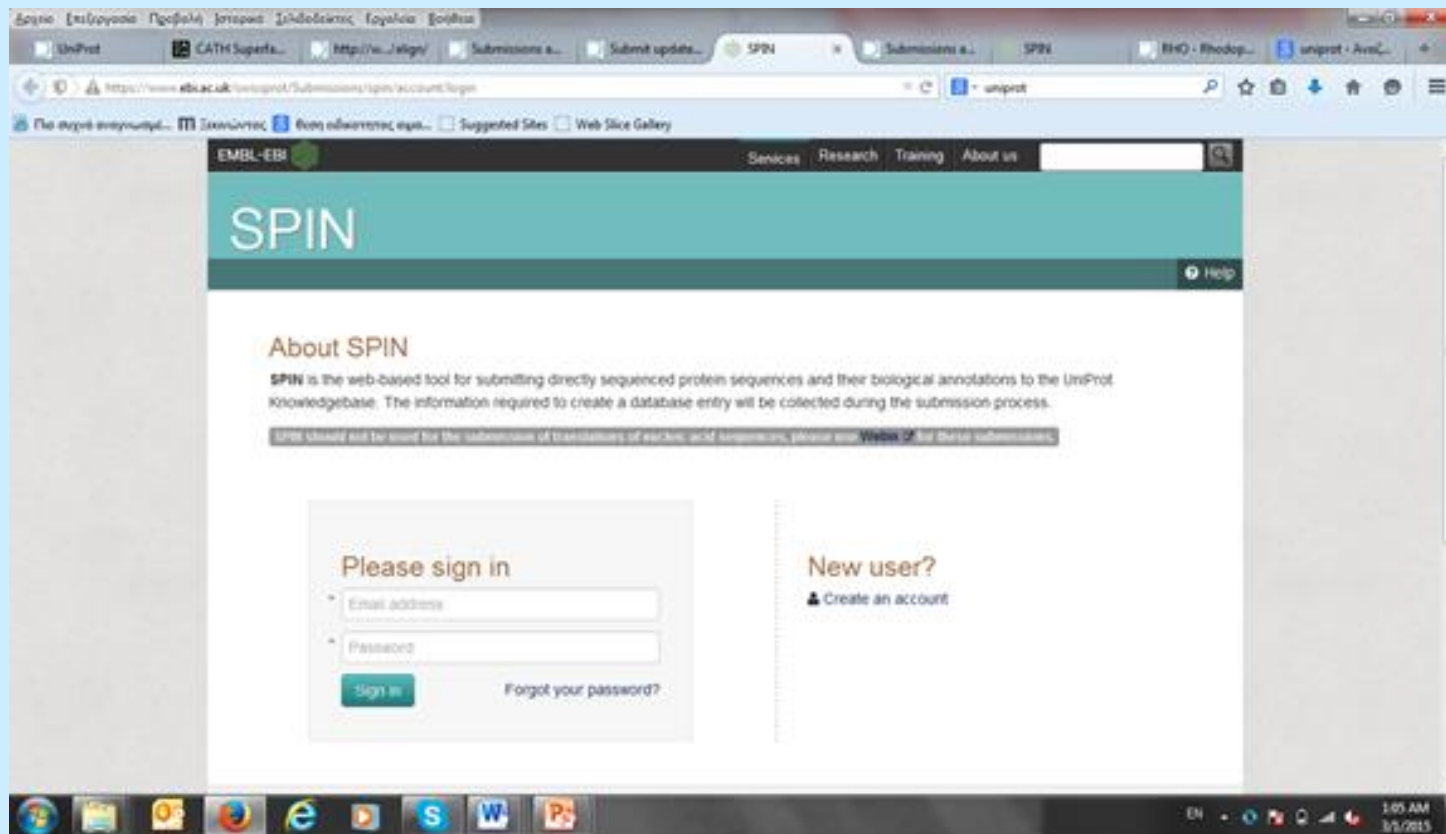
© 2002–2013 UniProt Consortium | License & Disclaimer

EMBL-EBI PIR SIB

www.uniprot.org/contact#

1:11 AM 3/5/2013

# Καταγραφή μιας νέας πρωτεΐνης στο SWISS-PROT



Στην προηγούμενη επιλογή “Contact” επιλέγουμε «Submit new protein sequence data» και μεταφερόμαστε στην αρχική σελίδα της SPIN ιστοσελίδα του EMBL-EBI

The screenshot shows a web browser window displaying the EMBL-EBI SPIN website. The browser's address bar shows the URL <https://www.ebi.ac.uk/submitproteins/spin/account/create>. The website has a teal header with the 'SPIN' logo and a navigation menu with links for 'Services', 'Research', 'Training', and 'About us'. Below the header, the main content area is titled 'Create a new account' and contains a registration form. The form includes fields for 'Name' (split into 'First name' and 'Surname'), 'Email address', 'Password', 'Confirm password', 'Address', 'Country' (a dropdown menu currently showing 'Ελλάδα'), and 'Telephone'. The browser's taskbar at the bottom shows various application icons, and the system clock in the bottom right corner indicates the time is 1:17 AM on 3/4/2015.

EMBL-EBI SPIN

Create a new account

Name \*  First name  Surname

Email address \*

Password \*

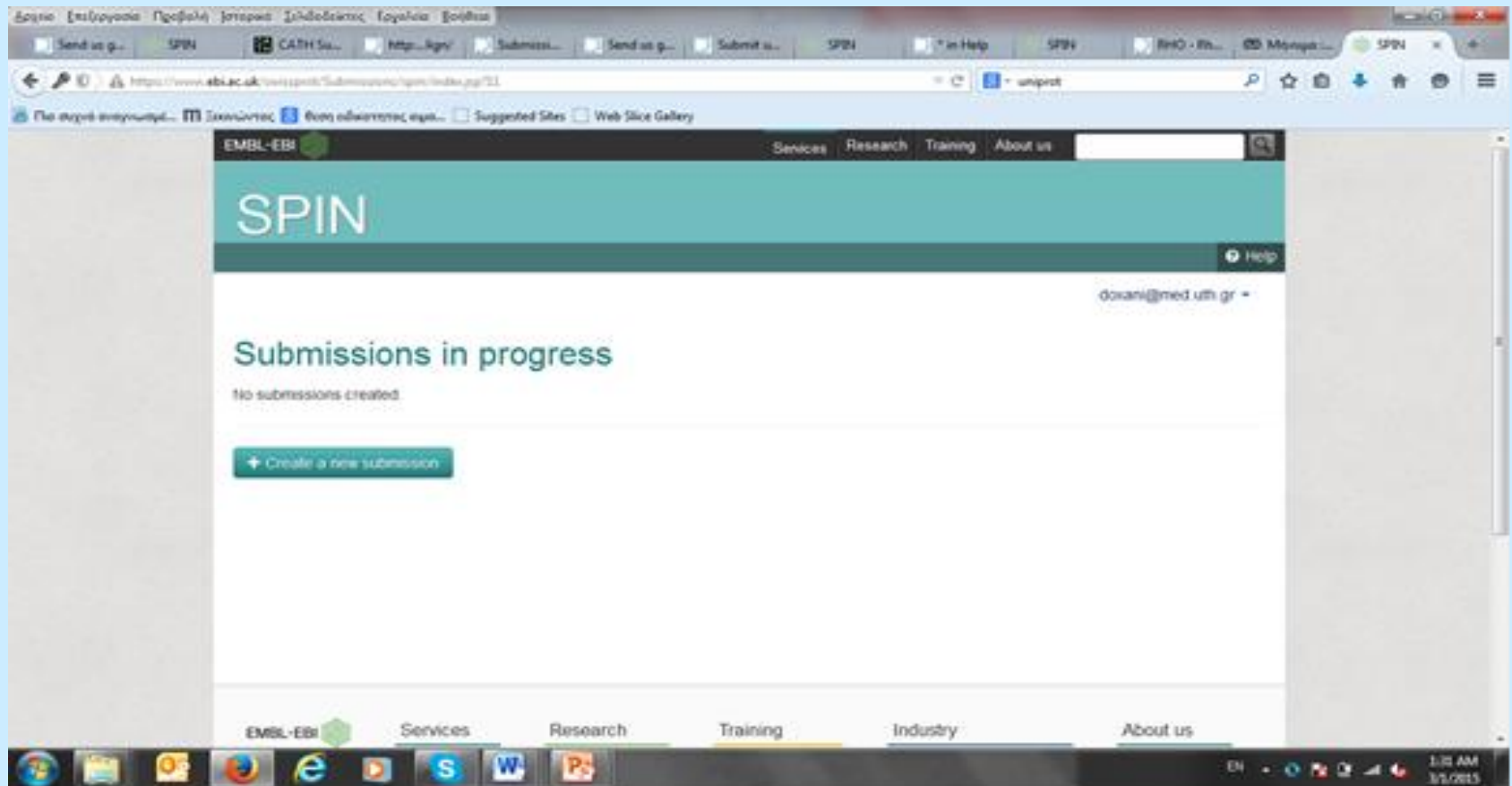
Confirm password \*

Address \*

Country \* Ελλάδα

Telephone \*

Αφού δημιουργήσουμε λογαριασμό συνεχίζουμε



Επιλέγοντας το “Create new submissions”,  
Συμπληρώνουμε και υποβάλλουμε την φόρμα.

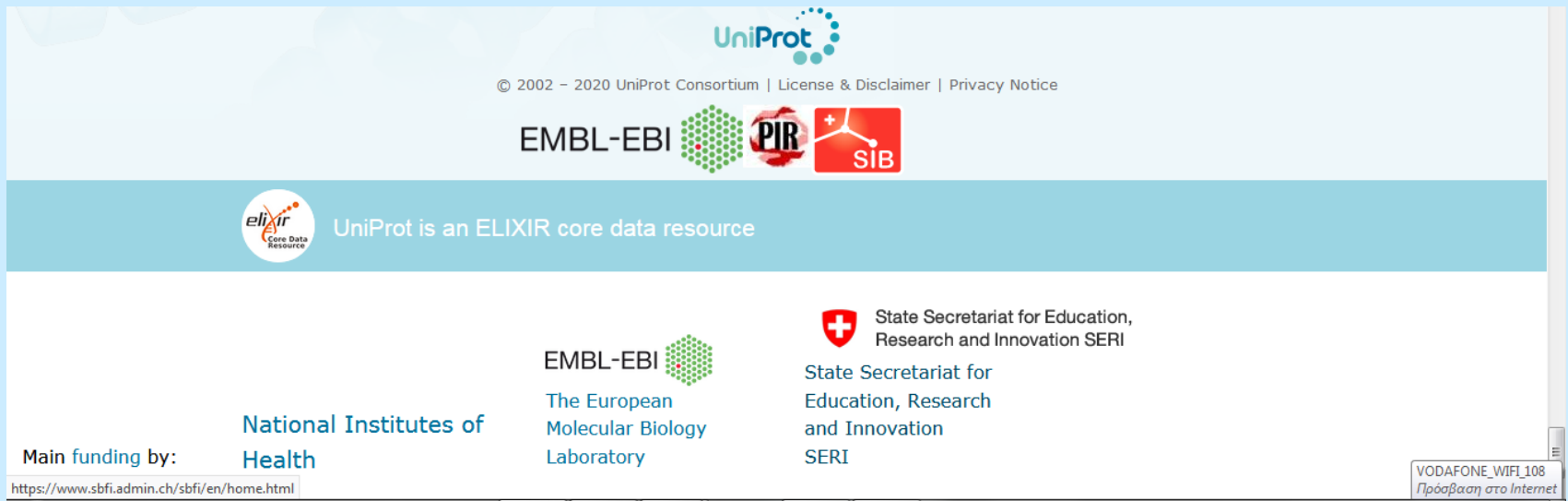
The screenshot shows a web browser window with the URL <https://www.ebi.ac.uk/submit/spin/submissions/create/1>. The page is titled "SPIN" and is part of the EMBL-EBI submission system. The user is logged in as `doxani@med.uth.gr`. The form is titled "Submission details" and contains the following fields:

- Protein name**: A text input field with a placeholder "e.g. Hemoglobin".
- Sequencing method**: A dropdown menu with "Method" selected and "EmuSeq" chosen.
- Sequence**: A section with three radio buttons:
  - ☒ I have a complete sequence
  - ☐ I have fragment(s) of a sequence
  - ☐ Upload from file...

The Windows taskbar at the bottom shows the time as 1:10 AM on 3/3/2015.



# Χρηματοδοτείται από 3 ιδρύματα




The screenshot displays the UniProt website. At the top, the UniProt logo is centered, with a copyright notice below it: "© 2002 – 2020 UniProt Consortium | License & Disclaimer | Privacy Notice". Below this, the logos for EMBL-EBI, PIR, and SIB are shown. A light blue banner features the ELIXIR logo and the text "UniProt is an ELIXIR core data resource". The footer lists the main funding partners: the National Institutes of Health, EMBL-EBI (The European Molecular Biology Laboratory), and the State Secretariat for Education, Research and Innovation (SERI) in Switzerland. A status bar at the bottom right shows "VODAFONE\_WIFI\_108" and "Πρόσβαση στο Internet".

UniProt

© 2002 – 2020 UniProt Consortium | License & Disclaimer | Privacy Notice


EMBL-EBI PIR SIB

 UniProt is an ELIXIR core data resource

Main funding by:

National Institutes of Health

EMBL-EBI  
The European  
Molecular Biology  
Laboratory

 State Secretariat for Education,  
Research and Innovation SERI

State Secretariat for  
Education, Research  
and Innovation  
SERI

<https://www.sbf.admin.ch/sbf/en/home.html>

VODAFONE\_WIFI\_108  
Πρόσβαση στο Internet