eine Proteindatenbank

14.07.2005
Stefan Stiller

Swiss-Prot - Inhalt

Übersicht
- → Betreiber & Gründung
- → UniProt: Swiss-Prot vs. TrEMBL
- → Einträge
  - → Inhalt
- → Wachstum

Organisation
- → Taxonomie
- → Vernetzung

Benutzung
- → Werkzeuge
- → Daten Einbringen

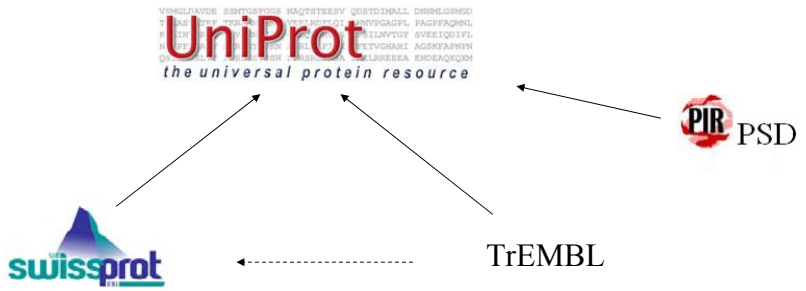
Swiss-Prot - Übersicht

-Swiss-Prot ist eine kommentierte Proteinsequenzdatenbank
-enthält 186882 Einträge (5. Juli 2005)
-Teil der UniProt Proteinsequenzdatenbank
-Inhalte werden geprüft und manuell kommentiert
-steht allen Personen öffentlich zur Verfügung
-alle Proteinsequenzen sind als precursor gespeichert, d.h. vor posttranslationaler Modifikation

Swiss-Prot - Betreiber & Gründung

-1986 gegründet und seit 1987 gemeinsam unterhalten vom Schweizer Institut für Bioinformatik (SIB) und dem Europäischen Bioinformatik Institut (EBI)

Swiss-Prot - UniProt: Swiss-Prot vs. TrEMBL



-geprüfte und manuell kommentierte Einträge

-Computer-kommentierte PSD
-Translationen aller codierenden Regionen aus EMBL
-hinzugefügte Proteinsequenzen, die noch nicht in Swiss-Prot integriert wurden
→ Qualität hängt vom Verfasser ab

Swiss-Prot – Einträge → Inhalt

- Swiss-Prot unterscheidet sich von reinen PSD durch:
  a) Dateninhalt
  b) minimale Redundanz
  c) Integration von anderen Datebanken

Swiss-Prot – Einträge → Inhalt

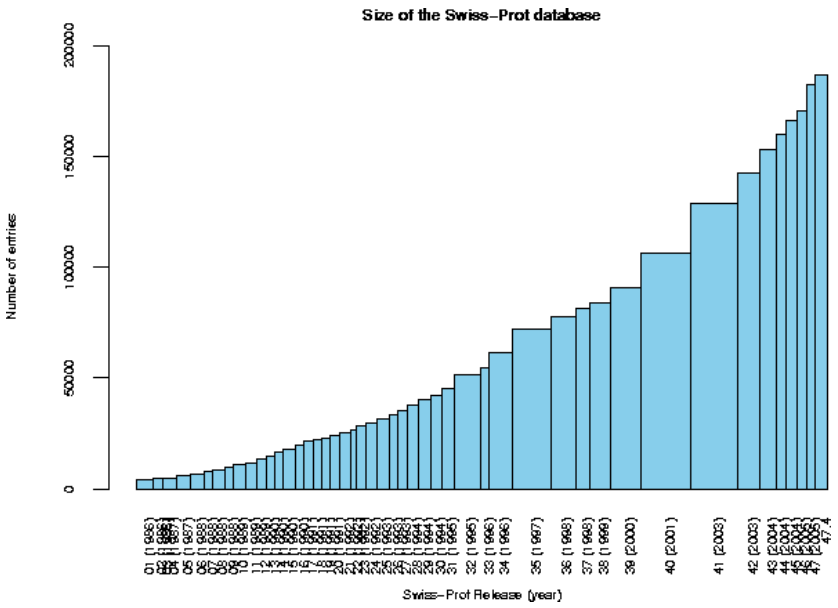
- Swiss-Prot unterscheidet sich von reinen PSD durch:
  a) Dateninhalt :
     - Kerndaten:
       - Sequenzdaten
       - zitierte Informationen
       - taxonomische Daten
     - Beschreibungen:
       - Funktion
       - Posttranslationale Modifikation
       - Domänen
       - sek. & qua. Struktur
       - Ähnlichkeiten
       - Krankheitsbezüge

  → periodische Updates der Beschreibungen, durch wissenschaftl.Artikel und externe Experten

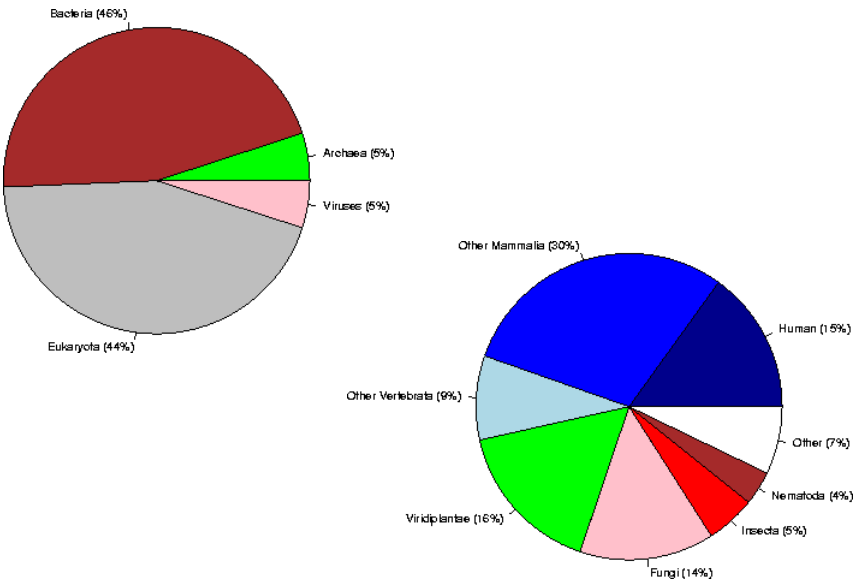
Swiss-Prot – Einträge → Inhalt

- Swiss-Prot unterscheidet sich von reinen PSD durch:
  b) minimale Redundanz
     - redundante Einträge werden vermieden, bei Unstimmigkeiten werden diese in einer Tabelle aufgeführt
  c) Integration von anderen Datebanken
     - durch Integration von anderen Datenbanken werden Referenzen zu DNA-Sequenzen und tertiären Proteinstrukturen angeboten

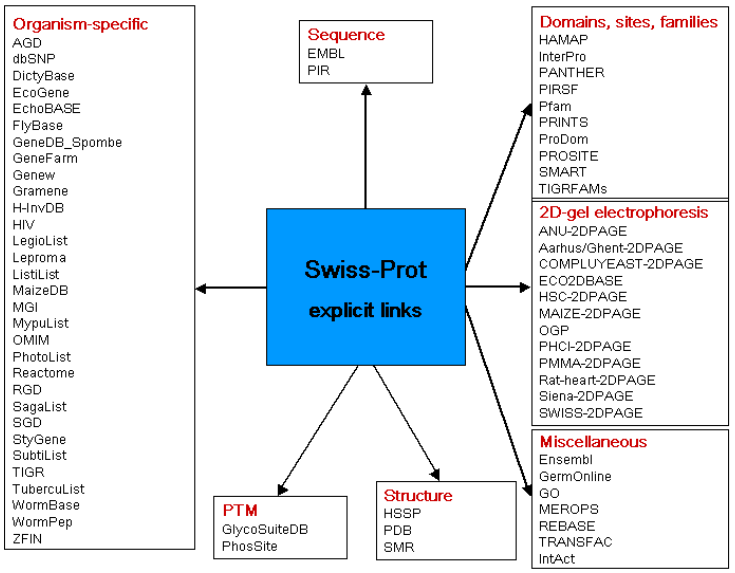
Swiss-Prot – Wachstum



Swiss-Prot – Organisation → Taxonomie



Swiss-Prot – Organisation → Vernetzung



Swiss-Prot – Benutzung

- Werkzeuge:
  - -SRS (Sequence Retrival System)
  - -Volltextsuche
  - -erweiterte Suche
  - -weitere Werkzeuge
- -Daten Einbringen

Swiss-Prot – Benutzung - SRS

Top Page | Query Form | Query Manager | View Manager | Databanks | Help

Select one or more databanks and continue

Continue | Reset

Sequence

☑ SWISS_PROT ☐ TREMBL

SeqRelated

☐ SWISS2D ☐ PROSITE ☐ PROSITEDOC ☐ ENZYME

Bookmark this link to return to your session: resume session

... tired of looking at all this data? Change their color! Try plain *If you find problems or have suggestions please mail the* SRS administrator

Swiss-Prot – Benutzung - SRS

Top Page | Query Form | Query Manager | View Manager | Databanks | Help

Search SWISS_PROT

Do Query | Reset | Combine searches with AND | ☑ Append wildcard '*' to words.

Info	Organism	Human
Info	Title	Galactosidase
Info	AllText	
Info	AllText	

Include fields in output: ID, AccNumber, Date, SubmissionDate, Description, GeneName, Keywords

Display in: ◉ list ○ table

Entry List in chunks of 100

Sequence Format * default *

Use view Short Description

Retrieve set of entry

* default * (* default *, fasta, gcg, pir, swiss, plain, pretty, string, ProteinChart)

Alternative Query Form *Separate multiple values by & (and), | (or),*

Databank	Fields in Group
SWISS_PROT	ID AccNumber Description GeneName Keywords Organism NCBI_TaxId Organelle_Plasmid Authors Title Patent PatentDate PubMed ReportType RefGroup DBxref CommentType Comment

Swiss-Prot – Benutzung - SRS

Top Page | Query Form | Query Manager | View Manager | Databanks | Help

Query **"[swiss_prot-Organism: Human*] & [swiss_prot-Title: Galactosidase*] "** found 3 entries

Perform operation on ◉ all but selected ○ selected *entries in chunks of* 100 • link • save • view with Short Description

RootLibs	*acc*	*des*	*sl*
SWISS_PROT:AGAL_HUMAN	P06280	Alpha-galactosidase A precursor (EC 3.2.1.22) (Melibiase) (Alpha-D-galactoside galactohydrolase) (Alpha-D-galactosidase A) (Agalsidase alfa).	429
SWISS_PROT:BGAL_HUMAN	P16278	Beta-galactosidase precursor (EC 3.2.1.23) (Lactase) (Acid beta-galactosidase).	677
SWISS_PROT:BGAM_HUMAN	P16279	Beta-galactosidase-related protein precursor (Beta-galactosidase-like protein) (S-Gal) (Elastin-binding protein) (EBP).	546

Swiss-Prot – Benutzung - SRS

UniProtKB/Swiss-Prot entry P16278

Printer-friendly view | Submit update | Quick BlastP search

[Entry info] [Name and origin] [References] [Comments] [Cross-references] [Keywords] [Features] [Sequence] [Tools]

Note: most headings are clickable, even if they don't appear as links. They link to the user manual or other documents.

Entry information	
Entry name	**BGAL_HUMAN**
Primary accession number	**P16278**
Secondary accession numbers	None
Entered in Swiss-Prot in	Release 15, August 1990
Sequence was last modified in	Release 15, August 1990
Annotations were last modified in	Release 47, May 2005
Name and origin of the protein	
Protein name	**Beta-galactosidase [Precursor]**
Synonyms	**EC 3.2.1.23** **Lactase** **Acid beta-galactosidase**
Gene name	**Name: GLB1**
From	Homo sapiens (Human) [TaxID: 9606]
Taxonomy	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

References

[1] NUCLEOTIDE SEQUENCE [MRNA], AND PARTIAL PROTEIN SEQUENCE.
TISSUE=Testis;
PubMed=2511208 [NCBI, ExPASy, EBI, Israel, Japan]
Morreau H., Galjart N.J., Gillemans N., Willemsen R., van der Horst G.T.J., D'Azzo A.;
"Alternative splicing of beta-galactosidase mRNA generates the classic lysosomal enzyme and a beta-galactosidase-related protein.";
J. Biol. Chem. 264:20655-20663(1989).

Swiss-Prot – Benutzung - SRS

Comments

- ***FUNCTION***: Cleaves beta-linked terminal galactosyl residues from gangliosides, glycoproteins, and glycosaminoglycans.
- ***CATALYTIC ACTIVITY***: Hydrolysis of terminal non-reducing beta-D-galactose residues in beta-D-galactosides.
- ***SUBCELLULAR LOCATION***: Lysosomal.
- ***ALTERNATIVE PRODUCTS***:
  Display all isoform sequences in FASTA format
  - **Alternative splicing [2 named forms]**

Name	1
Isoform ID	P16278-1
This is the isoform sequence displayed in this entry.	

Name	2
Synonyms	Beta-galactosidase-related protein
Isoform ID	P16279-1
This isoform is stored in UniProtKB/Swiss-Prot entry P16279.	

- ***DISEASE***: Defects in GLB1 are the cause of GM1-gangliosidosis type I [MIM:230500]; also known as infantile GM1-gangliosidosis. This autosomal recessive disorder is characterized by the accumulation in visceral tissues, and ultimately excessive excretion in the urine, of beta-linked galactose-terminal oligosaccharides. Patients show central nervous system degeneration, and the coarse facial features, hepatosplenomegaly and skeletal dysmorphology reminiscent of Hurler syndrome. The infantile form is rapidly progressive leading to death usually between the first and second year.
- ***DISEASE***: Defects in GLB1 are the cause of GM1-gangliosidosis type II [MIM:230600]; also known as late infantile/juvenile type GM1-gangliosidosis. Patients do not display the skeletal changes associated with the infantile form, but they nonetheless excrete elevated amounts of beta-linked galactose-terminal oligosaccharides. Inheritance is autosomal recessive.
- ***DISEASE***: Defects in GLB1 are the cause of GM1-gangliosidosis type III [MIM:230650]; also known as adult or chronic GM1-gangliosidosis. Patients show mild skeletal abnormalities, dysarthria, gait disturbance, dystonia and visual impairment. Visceromegaly is absent. Inheritance is autosomal recessive.
- ***DISEASE***: Defects in GLB1 are the cause of mucopolysaccharidosis IV B (MPS4B) [MIM:253010]; also known as Morquio syndrome B. MPS4B is a rare autosomal recessive disorder characterized by severe bone deformities without CNS involvement.
- ***SIMILARITY***: Belongs to the glycosyl hydrolase 35 family [view classification].

Cross-references

	M27507; AAA51819.1; -; mRNA.	[EMBL / GenBank / DDBJ] [CoDingSequence]
	M34423; AAA51823.1; -; mRNA.	[EMBL / GenBank / DDBJ] [CoDingSequence]
EMBL	M22590; AAA51822.1; -; mRNA.	[EMBL / GenBank / DDBJ] [CoDingSequence]

Swiss-Prot – Benutzung - SRS

Sequence information

Length: **677 AA** [This is the length of the unprocessed precursor] | Molecular weight: **76091 Da** [This is the MW of the unprocessed precursor] | CRC64: **780987BD5B04CF12** [This is a checksum on the sequence]

```
        10         20         30         40         50         60
MPGFLVRILL LLLVLLLLGP TRGLRNATQR MFEIDYSRDS FLKDGQPFRY ISGSIHYSRV

        70         80         90        100        110        120
PRFYWKDRLL KMKMAGLNAI QTYVPWNFHE PWPGQYQFSE DHDVEYFLRL AHELGLLVIL

       130        140        150        160        170        180
RPGPYICAEW EMGGLPAWLL EKESILLRSS DPDYLAAVDK WLGVLLPKMK PLLYQNGGPV

       190        200        210        220        230        240
ITVQVENEYG SYFACDFDYL RFLQKRFRHH LGDDVVLFTT DGAHKTFLKC GALQGLYTTV

       250        260        270        280        290        300
DFGTGSNITD AFLSQRKCEP KGPLINSEFY TGWLDHWGQP HSTIKTEAVA SSLYDILARG

       310        320        330        340        350        360
ASVNLYMFIG GTNFAYWNGA NSPYAAQPTS YDYDAPLSEA GDLTEKYFAL RNIIQKFEKV

       370        380        390        400        410        420
PEGPIPPSTP KFAYGKVTLE KLKTVGAALD ILCPSGPIKS LYPLTFIQVK QHYGFVLYRT

       430        440        450        460        470        480
TLPQDCSNPA PLSSPLNGVH DRAYVAVDGI PQGVLERNNV ITLNITGKAG ATLDLLVENM

       490        500        510        520        530        540
GRVNYGAYIN DFKGLVSNLT LSSNILTDWT IFPLDTEDAV RSHLGGWGHR DSGHHDEAWA

       550        560        570        580        590        600
HNSSNYTLPA FYMGNFSIPS GIPDLPQDTF IQFPGWTKGQ VWINGFNLGR YWPARGPQLT

       610        620        630        640        650        660
LFVPQHILMT SAPNTITVLE LEWAPCSSDD PELCAVTFVD RPVIGSSVTY DHPSKPVEKR

       670
LMPPPPQKNK DSWLDHV
```

P16278 in FASTA format

Swiss-Prot – Volltextsuche

Swiss-Prot/TrEMBL

Full text search

The search is performed on the current Swiss-Prot and TrEMBL releases, as well as the weekly updates.
(UniProtKB/Swiss-Prot Release 47.4 of 05-Jul-2005, 186882 entries; UniProtKB/TrEMBL Release 30.4 of 05-Jul-2005, 1837312 entries).
The index has been created by the Glimpse search engine.

Enter search terms:

☐ Prefix and append wildcard '*' to words.
Search in: ☑ Swiss-Prot ☐ TrEMBL
submit clear

Swiss-Prot – erweiterte Suche

Swiss-Prot/TrEMBL Advanced Search

This search program uses SRS to perform queries. Simpler forms are available to search by description or by full text. Available connectors within a field are "&" (and), "|" (or) and "!" (but not). You can prefix your search terms by ! to specify "not" (this is not possible in SRS). Example queries:

- To retrieve all AP1 complex proteins from mouse (AP1S1, AP1G1, etc. but not MIAP1, IQGAP1, ...), specify *Gene Name: ap1**, *Organism: Mus*, and deselect *"Append and prefix * to query terms"*.
- To retrieve the three human beta-adrenergic receptor proteins in Swiss-Prot, but not the beta-adrenergic receptor kinases, specify *Description: beta&adrenergic&receptor!kinase*, *Organism: Homo sapiens*, and select *"Append and prefix * to query terms"*

Search ☑ Swiss-Prot ☑ TrEMBL
Description
AND Gene name
AND Organism or choose from the list
or choose from the list
Human
Baker's yeast
E. coli
Mouse
Rat
Fruit fly
☑ Append and prefix * to query terms
detailed view of 100 results
Submit Query Reset

This tool can be used to create links to Swiss-Prot/TrEMBL by using the URL of the results page.

Swiss-Prot – weitere Werkzeuge