

Searching PubMed

Retrieving Protein Sequences

Searching PubMed

Navigate to <http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed>

Type in protein name in the for Window

How to save?

- clipboard

- text file

How to narrow down?

- combination of words not a single word

- use Limits

How to start with a new protein?

- read Review articles

More tips about PubMed

Add initials to proper names: ex. Park YS

Write down the PubMed identifier (PMID number)

Check spelling and synonyms

Try the Related Papers button

Limits: no record before 1965 & no abstract before 1976

 [4: Choi YK, Park JS, Kong JS, Morio T, Park YS.](#)

[Related Articles, Links](#)



D-threo-tetrahydrobiopterin is synthesized via 1'-oxo-2'-D-hydroxypropyl-tetrahydropterin in Dictyostelium discoideum Ax2.

FEBS Lett. 2005 Jun 6;579(14):3085-9.

[PMID: 15896778](#) [PubMed - indexed for MEDLINE]

Retrieving protein sequences in NCBI

Navigate to Entrez <http://www.ncbi.nlm.nih.gov/>

Click “Entrez Home” in the right box (Hot Spots)

Type protein name in the blank box

Click searched proteins

Click proteins of your interest

Change Display for more information

Click boxes to retrieve multiple sequences

Display FASTA

Click boxes and choose text

Retrieving protein sequences in Expasy

Navigate to Expasy (<http://kr.expasy.org/sprot/>)

Type in protein name in the Search Window

Select organism

Click Accession Number to see detailed protein information

Write down Primary Accession Number for future use

Check important information on the protein

Sequence section in the last section

select FASTA format

>definition line

SEQUENCEINCAPITAL

save in text file

use Courier font for easy alignment

*RAW format: simply the sequence part of the FASTA format

Retrieve multiple sequences

Select sequences

Click boxes in the first column

Go to the bottom line

Click retrieve

Download or open FASTA sequences



Identifiers

Q65TN8
P00915
P22748
P20507
P14141
P17067

or

Batch retrieval tips
Enter or upload a list of UniProt identifiers to download corresponding entries, for example:
P00750 A4_HUMAN
UniRef500_P00750
[More...](#)

7 unique items available for download

UniProtKB (7)

› Download data [compressed](#) or [uncompressed](#)

FASTA
Sequence data in FASTA format.
[\[Download \(3 KB\) | Open \]](#)

GFF
Sequence features in GFF.
[\[Download \(20 KB\) | Open \]](#)

Assignment

1. Choose a Protein
2. Significance of the protein: read at least a review article
3. Protein sequences (10 organisms) in FASTA format

Protein Summary in A4 2 pages

A review article

A protein sequence file in text format