Základy praktické Bioinformatiky

PETRA MATOUŠKOVÁ

2023/2024

2/10

H-index (může se lišit)

Scopus

Skálová, Lenka

View potential author matches

Author ID: 56344040000

http://orcid.org/0000-0003-0741-3871

Affiliation(s):

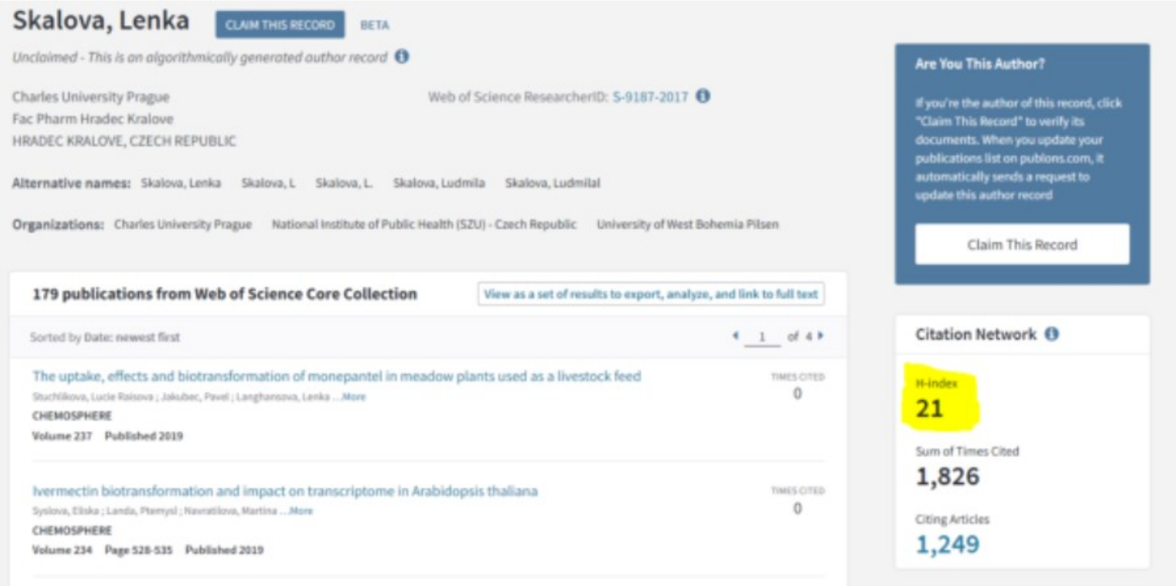
Charles University, Prague Praha, Czech Republic View more

Other name formats: Skalova, L. | Skalova, Lenka | Skálová, L.

Subject area: Pharmacology, Toxicology and Pharmaceutics | Chemistry | Environmental Science | Medicine | Biochemistry, Genetics and Molecular Biology | Agricultural and Biological Sciences | Veterinary | View all

Documents by author	Total citations	*h*-index:
160	1920 by 1320 documents	22
Analyze author output	View citation overview	View *h*-graph

WoS



Skalova, Lenka CLAIM THIS RECORD BETA

Unclaimed - This is an algorithmically generated author record

Charles University Prague
Fac Pharm Hradec Kralove
HRADEC KRALOVE, CZECH REPUBLIC

Web of Science ResearcherID: S-9187-2017

Alternative names: Skalova, Lenka | Skalova, L | Skalova, L. | Skalova, Ludmila | Skalova, Ludmilal

Organizations: Charles University Prague | National Institute of Public Health (SZU) - Czech Republic | University of West Bohemia Pilsen

179 publications from Web of Science Core Collection

View as a set of results to export, analyze, and link to full text

Sorted by Date: newest first — 1 of 4

The uptake, effects and biotransformation of monepantel in meadow plants used as a livestock feed — TIMES CITED 0
Stuchlikova, Lucie Raisova ; Jakubec, Pavel ; Langhansova, Lenka ...More
CHEMOSPHERE
Volume 237 Published 2019

Ivermectin biotransformation and impact on transcriptome in Arabidopsis thaliana — TIMES CITED 0
Syslova, Eliska ; Landa, Premysl ; Navratilova, Martina ...More
CHEMOSPHERE
Volume 234 Page 528-535 Published 2019

Are You This Author?

If you're the author of this record, click "Claim This Record" to verify its documents. When you update your publications list on publons.com, it automatically sends a request to update this author record

Claim This Record

Citation Network

H-index
21

Sum of Times Cited
1,826

Citing Articles
1,249

Základy praktické bioinformatiky

Téma 2/10

Proteinová bioinformatika I

Cíle:

Student bude schopen vyhledat a stáhnout požadovanou sekvenci proteinu. Na základě sekvence bude schopen určit (nalézt nebo spočítat) fyzikálně-biochemické vlastnosti proteinu.

Rešeršní projekt: NQO1

Vyhledejte dostupné informace o NQO1, získejte co nejvíce literárních, sekvenčních, případně i strukturních a dalších údajů o tomto genu/proteinu.

Nalezněte podobné sekvence a porovnejte je na proteinové úrovni.

- **NAD(P)H:chinonoxidoreduktasa / reference**

„Proteinová bioinformatika I“

Vyhledávání AMK sekvencí

Analýza vlastností sekvencí (aminokyselinové složení, molekulová hmotnost, isoelektrický bod…)

Štěpení proteasami

Analýza hydrofobních segmentů, transmembránových úseků

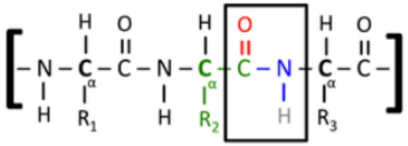
Predikce sekundární struktury

3D-struktura, vizualizace

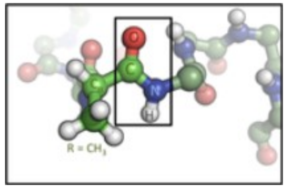
Vyhledání a porovnání podobných sekvencí

Evoluční příbuznost sekvencí

…



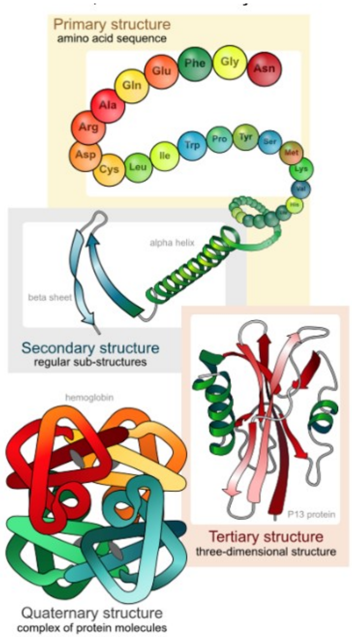
Proteiny

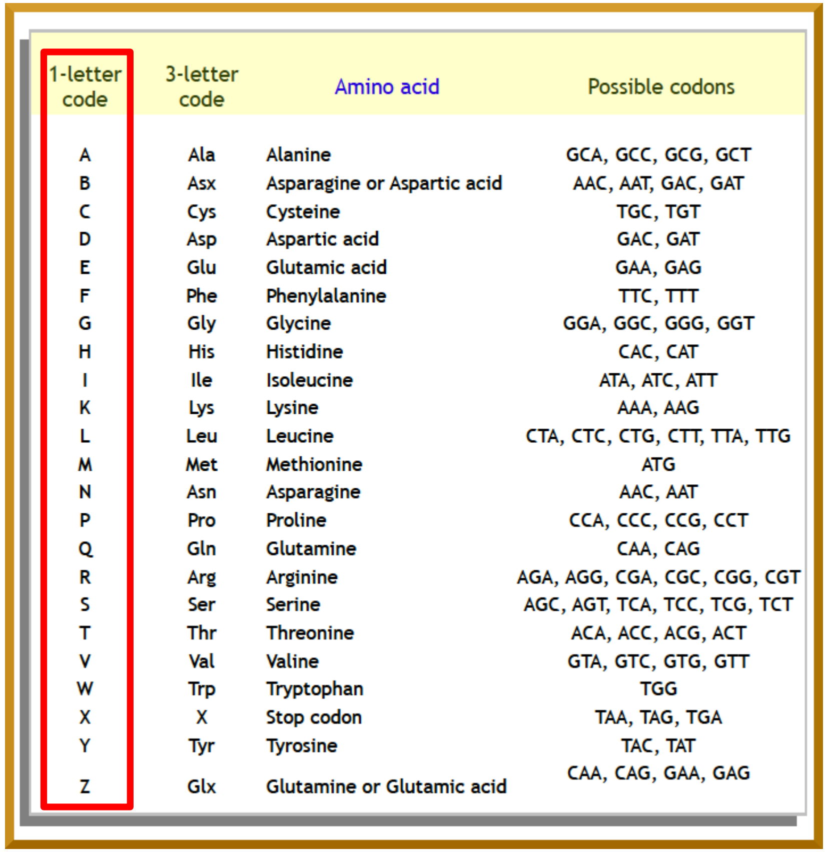


20 Aminokyselin – primární struktura:

(Frederick Sanger-1958 Nobelova cena za sekvenování insulinu)

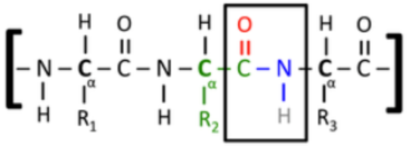
Sekundární struktura
Terciární struktura
Kvartérní struktura



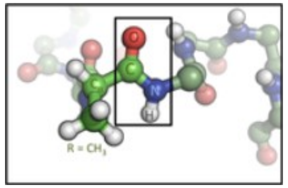


1-letter code	3-letter code	Amino acid	Possible codons
A	Ala	Alanine	GCA, GCC, GCG, GCT
B	Asx	Asparagine or Aspartic acid	AAC, AAT, GAC, GAT
C	Cys	Cysteine	TGC, TGT
D	Asp	Aspartic acid	GAC, GAT
E	Glu	Glutamic acid	GAA, GAG
F	Phe	Phenylalanine	TTC, TTT
G	Gly	Glycine	GGA, GGC, GGG, GGT
H	His	Histidine	CAC, CAT
I	Ile	Isoleucine	ATA, ATC, ATT
K	Lys	Lysine	AAA, AAG
L	Leu	Leucine	CTA, CTC, CTG, CTT, TTA, TTG
M	Met	Methionine	ATG
N	Asn	Asparagine	AAC, AAT
P	Pro	Proline	CCA, CCC, CCG, CCT
Q	Gln	Glutamine	CAA, CAG
R	Arg	Arginine	AGA, AGG, CGA, CGC, CGG, CGT
S	Ser	Serine	AGC, AGT, TCA, TCC, TCG, TCT
T	Thr	Threonine	ACA, ACC, ACG, ACT
V	Val	Valine	GTA, GTC, GTG, GTT
W	Trp	Tryptophan	TGG
X	X	Stop codon	TAA, TAG, TGA
Y	Tyr	Tyrosine	TAC, TAT
Z	Glx	Glutamine or Glutamic acid	CAA, CAG, GAA, GAG

SEKVENCE⇨STRUKTURA⇨FUNKCE

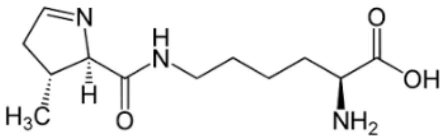


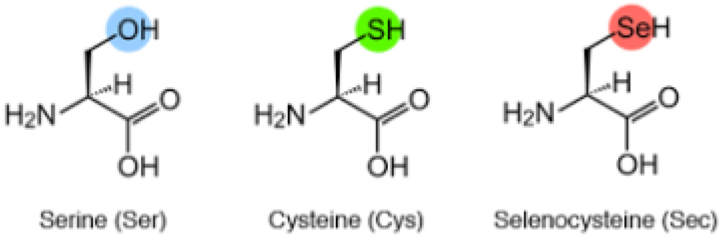
Proteiny



20 Aminokyselin – primární struktura:

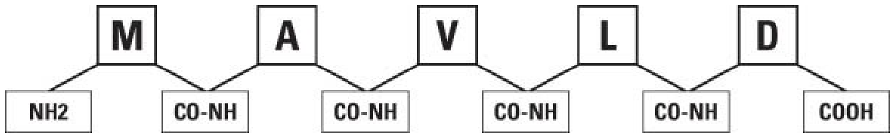
J	Xle	Isoleucine/Leucine
O	Pyl	Pyrrolysine
U	Sec	Selenocysteine
X	Xaa	Any residue

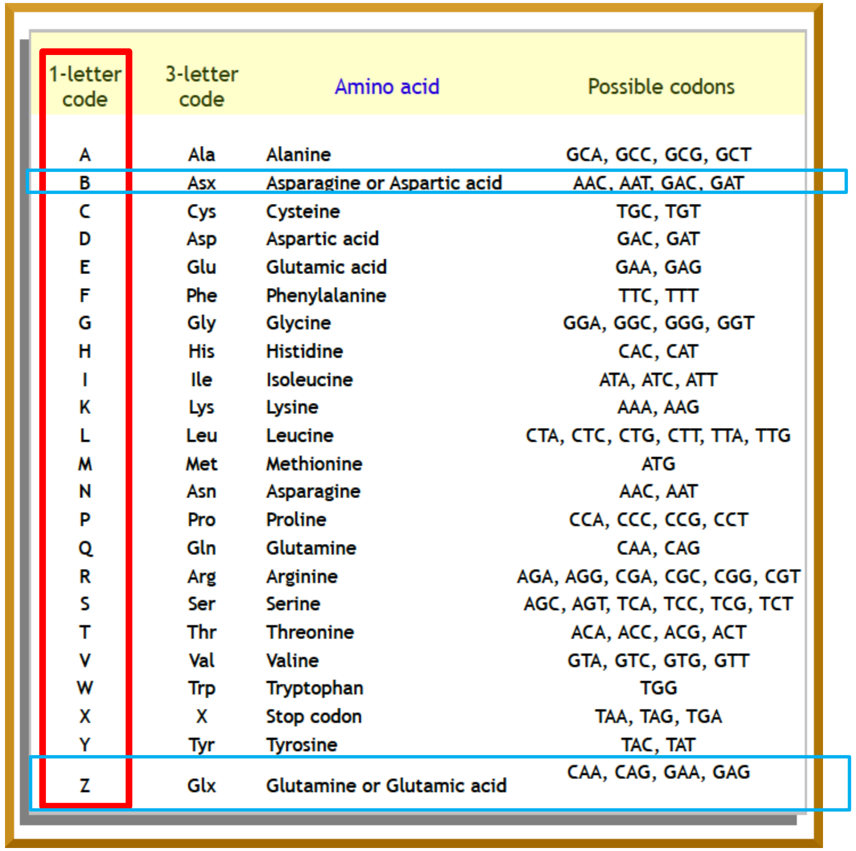




Serine (Ser) Cysteine (Cys) Selenocysteine (Sec)

N-konec → C-konec



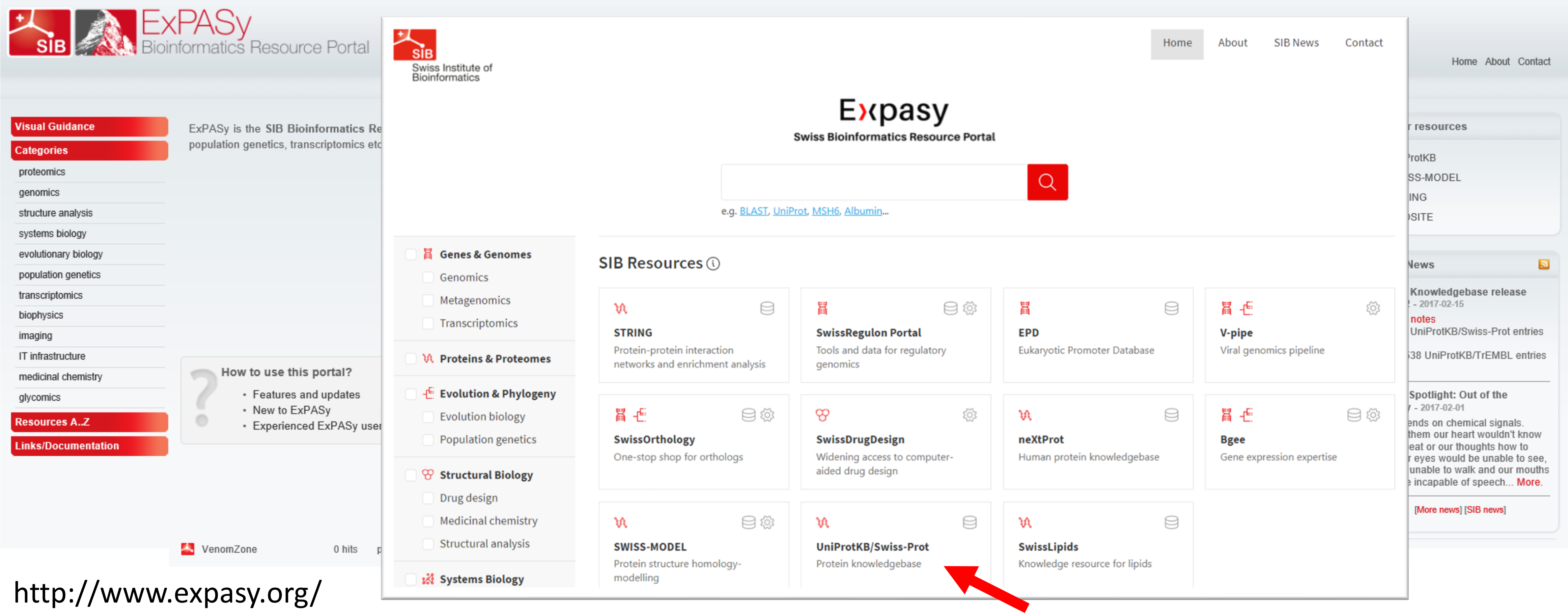


1-letter code	3-letter code	Amino acid	Possible codons
A	Ala	Alanine	GCA, GCC, GCG, GCT
B	Asx	Asparagine or Aspartic acid	AAC, AAT, GAC, GAT
C	Cys	Cysteine	TGC, TGT
D	Asp	Aspartic acid	GAC, GAT
E	Glu	Glutamic acid	GAA, GAG
F	Phe	Phenylalanine	TTC, TTT
G	Gly	Glycine	GGA, GGC, GGG, GGT
H	His	Histidine	CAC, CAT
I	Ile	Isoleucine	ATA, ATC, ATT
K	Lys	Lysine	AAA, AAG
L	Leu	Leucine	CTA, CTC, CTG, CTT, TTA, TTG
M	Met	Methionine	ATG
N	Asn	Asparagine	AAC, AAT
P	Pro	Proline	CCA, CCC, CCG, CCT
Q	Gln	Glutamine	CAA, CAG
R	Arg	Arginine	AGA, AGG, CGA, CGC, CGG, CGT
S	Ser	Serine	AGC, AGT, TCA, TCC, TCG, TCT
T	Thr	Threonine	ACA, ACC, ACG, ACT
V	Val	Valine	GTA, GTC, GTG, GTT
W	Trp	Tryptophan	TGG
X	X	Stop codon	TAA, TAG, TGA
Y	Tyr	Tyrosine	TAC, TAT
Z	Glx	Glutamine or Glutamic acid	CAA, CAG, GAA, GAG

Databáze

PRIMÁRNÍ	**SEKUNDÁRNÍ**
Ukládání primárních dat (sekvence, struktury…)	Shromažďování, analýza a propojování dat
Přímé vkládání	Primární databáze, literatura (po analýze a interpretaci primárních dat)
Genbank, DDBJ, ENA Protein Data Bank GEO…	SwissProt, InterPro, Ensembl, OMIM, STRING…

Proteinové databáze: UniProt



ExPASy is the SIB Bioinformatics Re… population genetics, transcriptomics etc

How to use this portal?
- Features and updates
- New to ExPASy
- Experienced ExPASy user

STRING Protein-protein interaction networks and enrichment analysis

SwissRegulon Portal Tools and data for regulatory genomics

EPD Eukaryotic Promoter Database

V-pipe Viral genomics pipeline

SwissOrthology One-stop shop for orthologs

SwissDrugDesign Widening access to computer-aided drug design

neXtProt Human protein knowledgebase

Bgee Gene expression expertise

SWISS-MODEL Protein structure homology-modelling

UniProtKB/Swiss-Prot Protein knowledgebase

SwissLipids Knowledge resource for lipids

http://www.expasy.org/

Možnosti databáze UniProt

http://www.uniprot.org/

Vyhledávání sekvencí, vlastnosti sekvencí

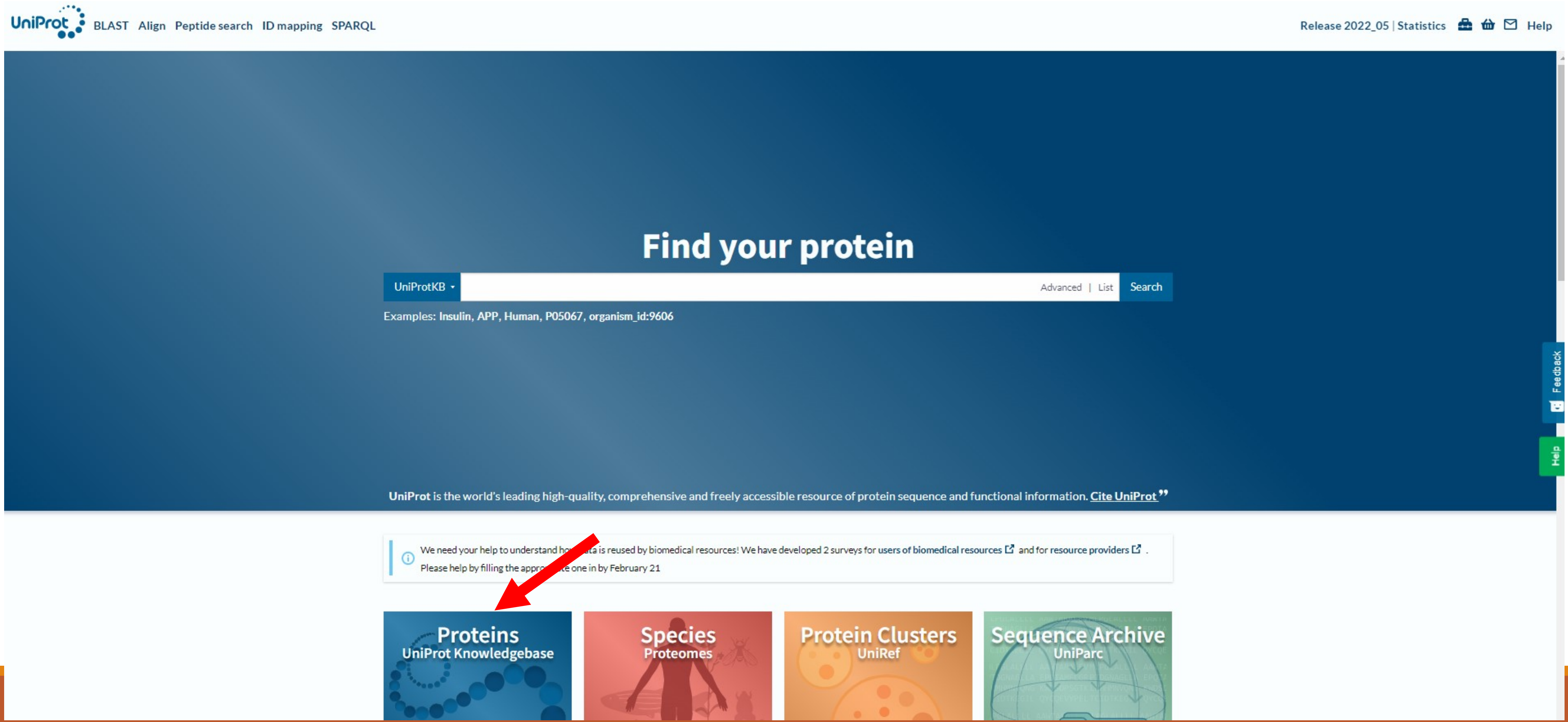
Funkce proteinů, interakce, dráhy

Propojení s nemocemi, mutace

Propojení do dalších databází

Proteinové databáze: UniProt

http://www.uniprot.org/



Find your protein

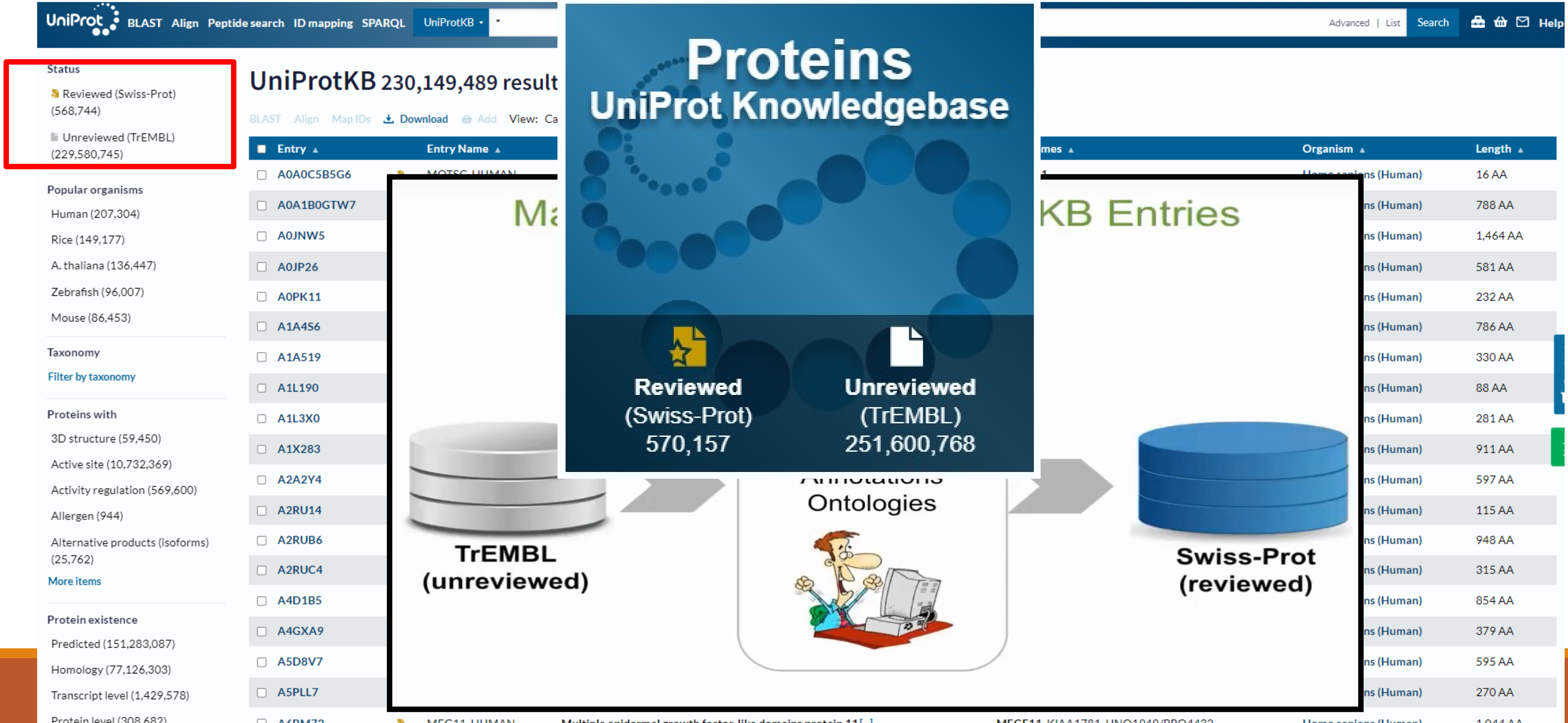
Examples: Insulin, APP, Human, P05067, organism_id:9606

UniProt is the world's leading high-quality, comprehensive and freely accessible resource of protein sequence and functional information. Cite UniProt

We need your help to understand ho [illegible] ta is reused by biomedical resources! We have developed 2 surveys for users of biomedical resources and for resource providers . Please help by filling the appro [illegible] te one in by February 21

Proteinové databáze: UniProt

http://www.uniprot.org/



UniProt BLAST Align Peptide search ID mapping SPARQL UniProtKB Advanced | List Search Help

Status

Reviewed (Swiss-Prot) (568,744)

Unreviewed (TrEMBL) (229,580,745)

Popular organisms

Human (207,304)

Rice (149,177)

A. thaliana (136,447)

Zebrafish (96,007)

Mouse (86,453)

Taxonomy

Filter by taxonomy

Proteins with

3D structure (59,450)

Active site (10,732,369)

Activity regulation (569,600)

Allergen (944)

Alternative products (isoforms) (25,762)

More items

Protein existence

Predicted (151,283,087)

Homology (77,126,303)

Transcript level (1,429,578)

Protein level (308,682)

UniProtKB 230,149,489 result

BLAST Align Map IDs Download Add View: Ca

Entry	Entry Name	mes	Organism	Length
A0A0C5B5G6	MOTSC_HUMAN		Homo sapiens (Human)	16 AA
A0A1B0GTW7			ns (Human)	788 AA
A0JNW5			ns (Human)	1,464 AA
A0JP26			ns (Human)	581 AA
A0PK11			ns (Human)	232 AA
A1A4S6			ns (Human)	786 AA
A1A519			ns (Human)	330 AA
A1L190			ns (Human)	88 AA
A1L3X0			ns (Human)	281 AA
A1X283			ns (Human)	911 AA
A2A2Y4			ns (Human)	597 AA
A2RU14			ns (Human)	115 AA
A2RUB6			ns (Human)	948 AA
A2RUC4			ns (Human)	315 AA
A4D1B5			ns (Human)	854 AA
A4GXA9			ns (Human)	379 AA
A5D8V7			ns (Human)	595 AA
A5PLL7			ns (Human)	270 AA
A6BM72	MEGF11_HUMAN	Multiple epidermal growth factor-like domains protein 11 [...] MEGF11 KIAA1781 UNQ1949/PRO4432	Homo sapiens (Human)	1,044 AA

Proteins

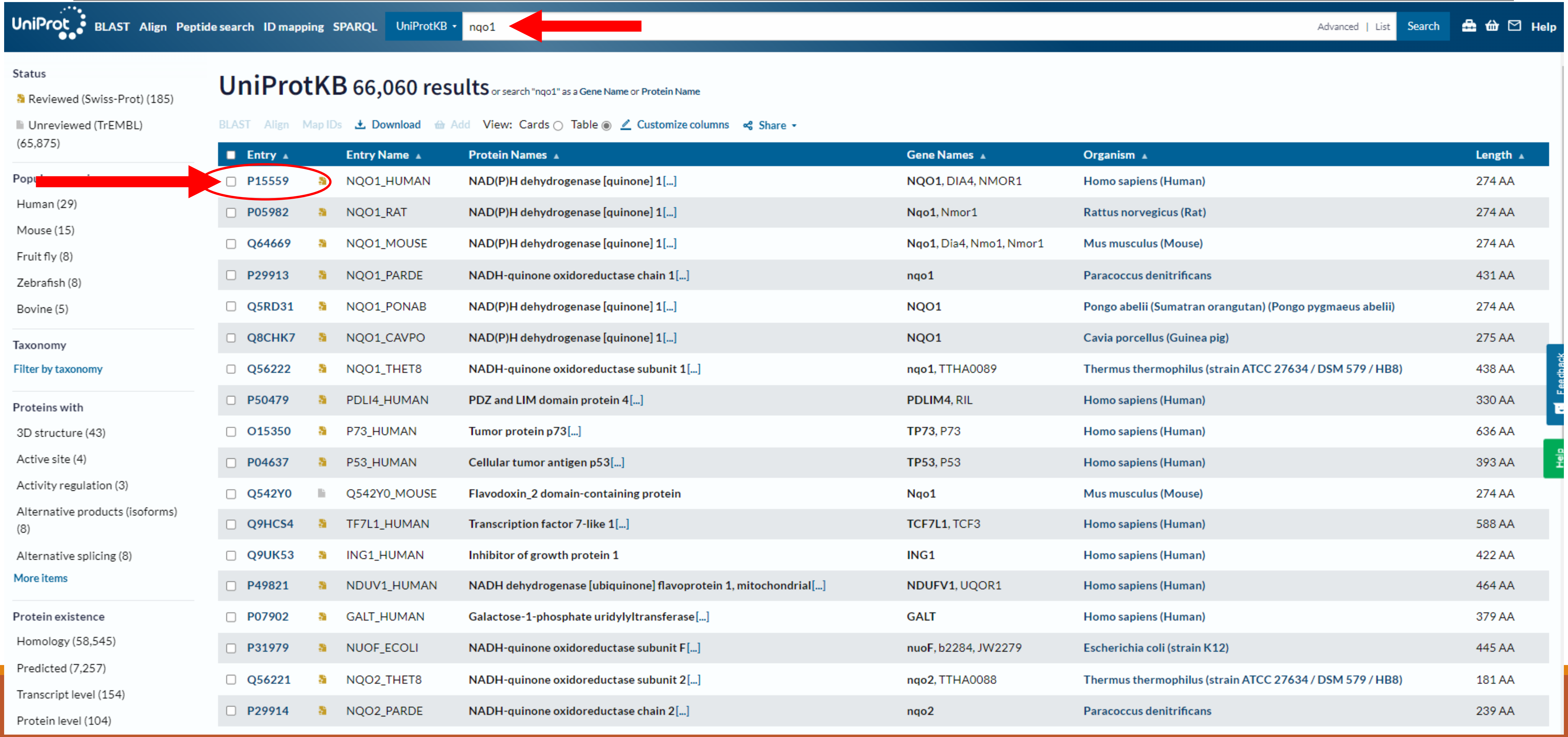
UniProt Knowledgebase

Reviewed (Swiss-Prot) 570,157

Unreviewed (TrEMBL) 251,600,768

Proteinové databáze: UniProt

http://www.uniprot.org/



nqo1

Status

Reviewed (Swiss-Prot) (185)

Unreviewed (TrEMBL) (65,875)

Popu

Human (29)

Mouse (15)

Fruit fly (8)

Zebrafish (8)

Bovine (5)

Taxonomy

Filter by taxonomy

Proteins with

3D structure (43)

Active site (4)

Activity regulation (3)

Alternative products (isoforms) (8)

Alternative splicing (8)

More items

Protein existence

Homology (58,545)

Predicted (7,257)

Transcript level (154)

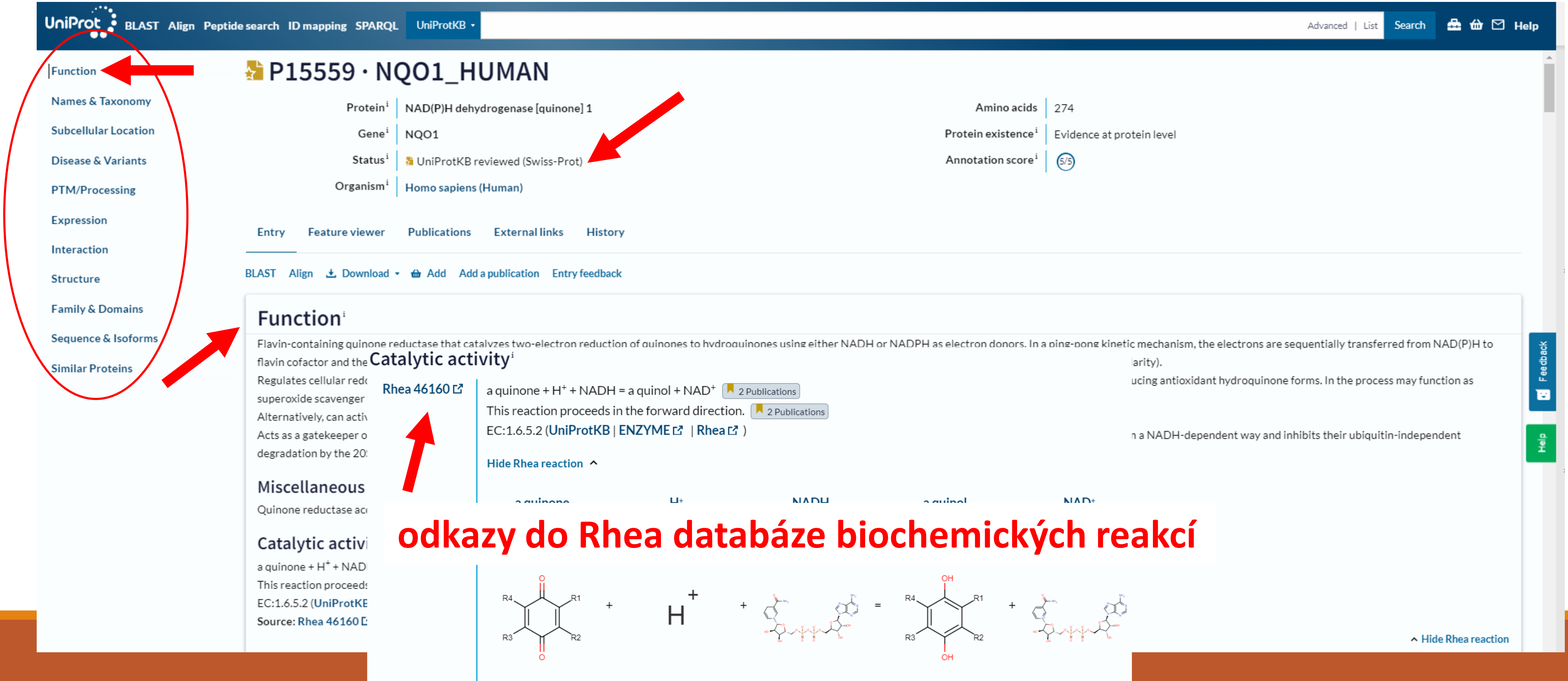
Protein level (104)

UniProtKB 66,060 results or search "nqo1" as a Gene Name or Protein Name

BLAST Align Map IDs Download Add View: Cards ○ Table ● Customize columns Share

☐ Entry	Entry Name	Protein Names	Gene Names	Organism	Length
☐ P15559	NQO1_HUMAN	NAD(P)H dehydrogenase [quinone] 1[...]	NQO1, DIA4, NMOR1	Homo sapiens (Human)	274 AA
☐ P05982	NQO1_RAT	NAD(P)H dehydrogenase [quinone] 1[...]	Nqo1, Nmor1	Rattus norvegicus (Rat)	274 AA
☐ Q64669	NQO1_MOUSE	NAD(P)H dehydrogenase [quinone] 1[...]	Nqo1, Dia4, Nmo1, Nmor1	Mus musculus (Mouse)	274 AA
☐ P29913	NQO1_PARDE	NADH-quinone oxidoreductase chain 1[...]	nqo1	Paracoccus denitrificans	431 AA
☐ Q5RD31	NQO1_PONAB	NAD(P)H dehydrogenase [quinone] 1[...]	NQO1	Pongo abelii (Sumatran orangutan) (Pongo pygmaeus abelii)	274 AA
☐ Q8CHK7	NQO1_CAVPO	NAD(P)H dehydrogenase [quinone] 1[...]	NQO1	Cavia porcellus (Guinea pig)	275 AA
☐ Q56222	NQO1_THET8	NADH-quinone oxidoreductase subunit 1[...]	nqo1, TTHA0089	Thermus thermophilus (strain ATCC 27634 / DSM 579 / HB8)	438 AA
☐ P50479	PDLI4_HUMAN	PDZ and LIM domain protein 4[...]	PDLIM4, RIL	Homo sapiens (Human)	330 AA
☐ O15350	P73_HUMAN	Tumor protein p73[...]	TP73, P73	Homo sapiens (Human)	636 AA
☐ P04637	P53_HUMAN	Cellular tumor antigen p53[...]	TP53, P53	Homo sapiens (Human)	393 AA
☐ Q542Y0	Q542Y0_MOUSE	Flavodoxin_2 domain-containing protein	Nqo1	Mus musculus (Mouse)	274 AA
☐ Q9HCS4	TF7L1_HUMAN	Transcription factor 7-like 1[...]	TCF7L1, TCF3	Homo sapiens (Human)	588 AA
☐ Q9UK53	ING1_HUMAN	Inhibitor of growth protein 1	ING1	Homo sapiens (Human)	422 AA
☐ P49821	NDUV1_HUMAN	NADH dehydrogenase [ubiquinone] flavoprotein 1, mitochondrial[...]	NDUFV1, UQOR1	Homo sapiens (Human)	464 AA
☐ P07902	GALT_HUMAN	Galactose-1-phosphate uridylyltransferase[...]	GALT	Homo sapiens (Human)	379 AA
☐ P31979	NUOF_ECOLI	NADH-quinone oxidoreductase subunit F[...]	nuoF, b2284, JW2279	Escherichia coli (strain K12)	445 AA
☐ Q56221	NQO2_THET8	NADH-quinone oxidoreductase subunit 2[...]	nqo2, TTHA0088	Thermus thermophilus (strain ATCC 27634 / DSM 579 / HB8)	181 AA
☐ P29914	NQO2_PARDE	NADH-quinone oxidoreductase chain 2[...]	nqo2	Paracoccus denitrificans	239 AA

Proteinové databáze: UniProt



P15559 · NQO1_HUMAN

Protein: NAD(P)H dehydrogenase [quinone] 1

Gene: NQO1

Status: UniProtKB reviewed (Swiss-Prot)

Organism: Homo sapiens (Human)

Amino acids: 274

Protein existence: Evidence at protein level

Annotation score: 5/5

Function

Flavin-containing quinone reductase that catalyzes two-electron reduction of quinones to hydroquinones using either NADH or NADPH as electron donors. In a ping-pong kinetic mechanism, the electrons are sequentially transferred from NAD(P)H to flavin cofactor and the … larity).

Regulates cellular red… ucing antioxidant hydroquinone forms. In the process may function as superoxide scavenger…

Alternatively, can acti…

Acts as a gatekeeper o… n a NADH-dependent way and inhibits their ubiquitin-independent degradation by the 20…

Miscellaneous

Quinone reductase ac…

Catalytic activi…

a quinone + H$^+$ + NAD…

This reaction proceeds…

EC:1.6.5.2 (UniProtKE…

Source: Rhea 46160

Catalytic activity

Rhea 46160

a quinone + H$^+$ + NADH = a quinol + NAD$^+$ 2 Publications

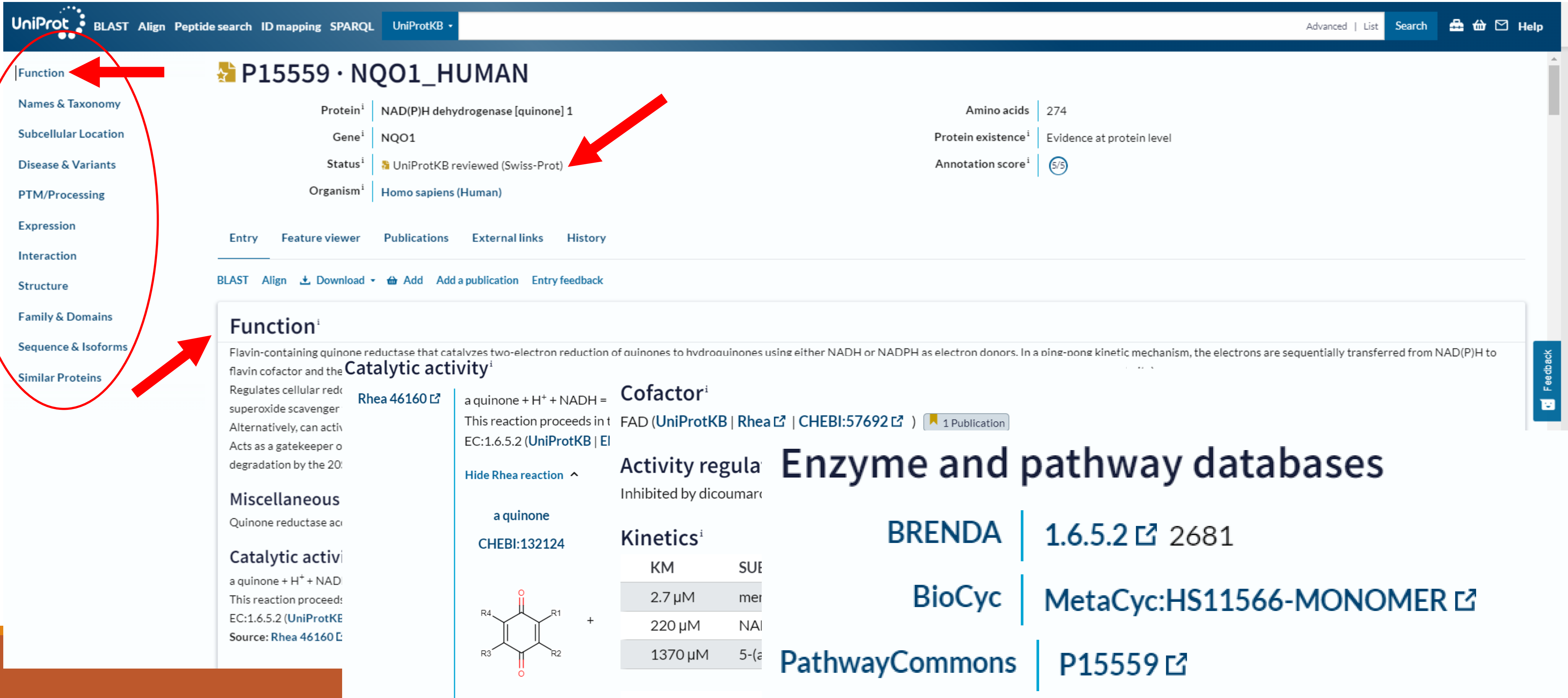
This reaction proceeds in the forward direction. 2 Publications

EC:1.6.5.2 (UniProtKB | ENZYME | Rhea)

Hide Rhea reaction

odkazy do Rhea databáze biochemických reakcí

Proteinové databáze: UniProt



P15559 · NQO1_HUMAN

Protein	NAD(P)H dehydrogenase [quinone] 1
Gene	NQO1
Status	UniProtKB reviewed (Swiss-Prot)
Organism	Homo sapiens (Human)

Amino acids	274
Protein existence	Evidence at protein level
Annotation score	5/5

Function

Flavin-containing quinone reductase that catalyzes two-electron reduction of quinones to hydroquinones using either NADH or NADPH as electron donors. In a ping-pong kinetic mechanism, the electrons are sequentially transferred from NAD(P)H to flavin cofactor and the

Regulates cellular red
superoxide scavenger
Alternatively, can activ
Acts as a gatekeeper o
degradation by the 20

Miscellaneous

Quinone reductase ac

Catalytic activi

a quinone + H^+ + NAD
This reaction proceed
EC:1.6.5.2 (UniProtKE
Source: Rhea 46160

Catalytic activity

Rhea 46160

a quinone + H^+ + NADH =
This reaction proceeds in t
EC:1.6.5.2 (UniProtKB | El

Hide Rhea reaction

a quinone
CHEBI:132124

Cofactor

FAD (UniProtKB | Rhea | CHEBI:57692) 1 Publication

Activity regula

Inhibited by dicoumar

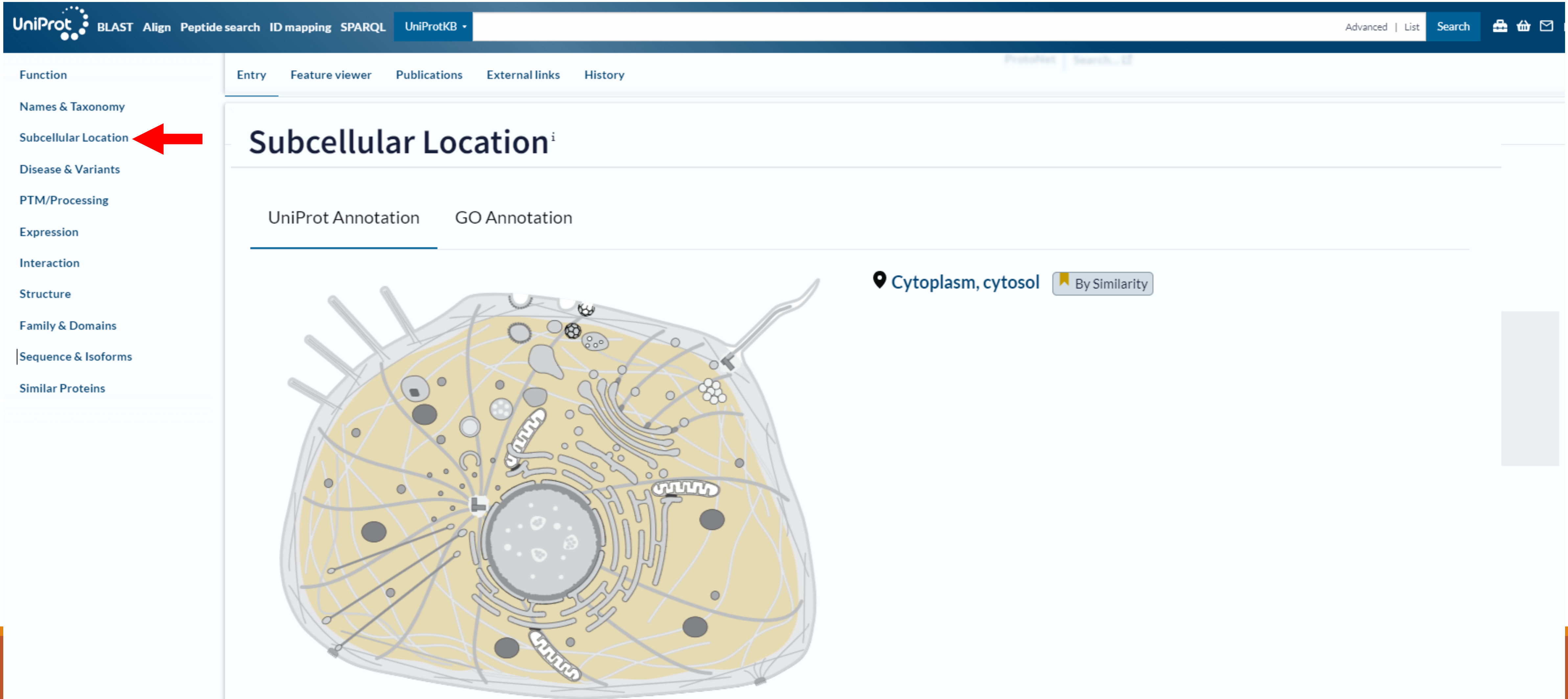
Kinetics

KM	SUE
2.7 μM	mer
220 μM	NAI
1370 μM	5-(a

Enzyme and pathway databases

BRENDA	1.6.5.2 2681
BioCyc	MetaCyc:HS11566-MONOMER
PathwayCommons	P15559

Proteinové databáze: UniProt

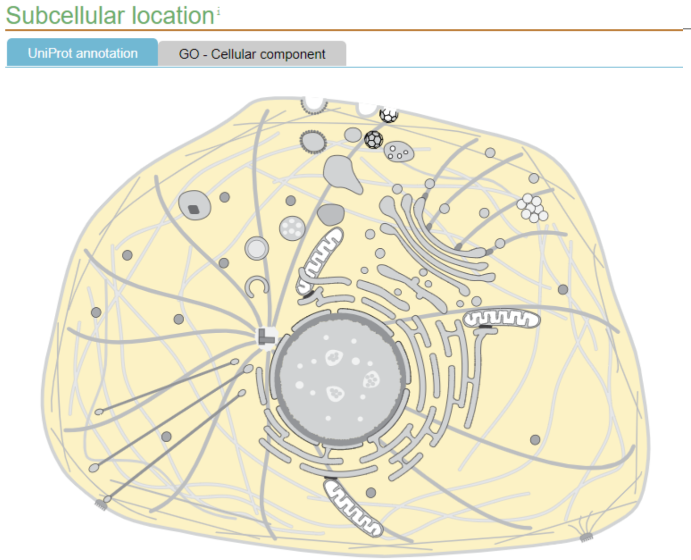


Subcellular Location

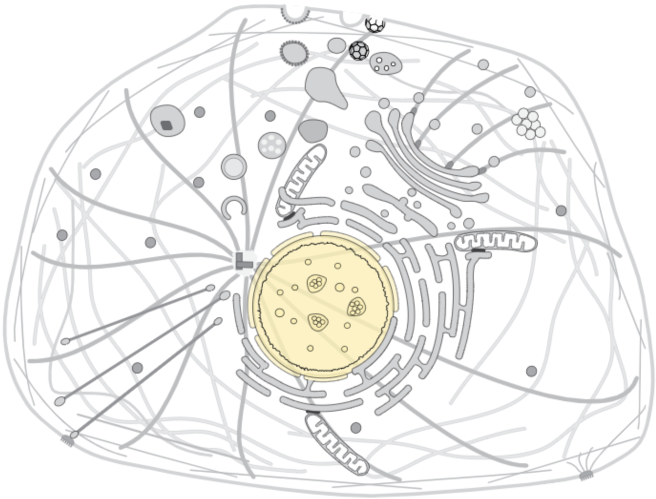
UniProt Annotation | GO Annotation

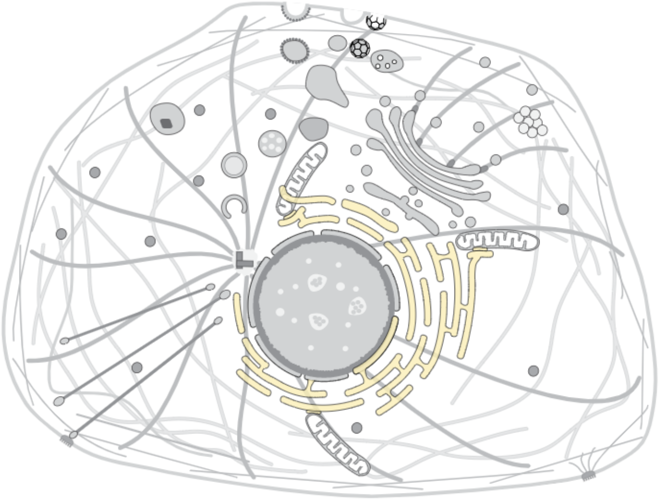
Cytoplasm, cytosol — By Similarity

UNIPROT / subcelulární lokalizace

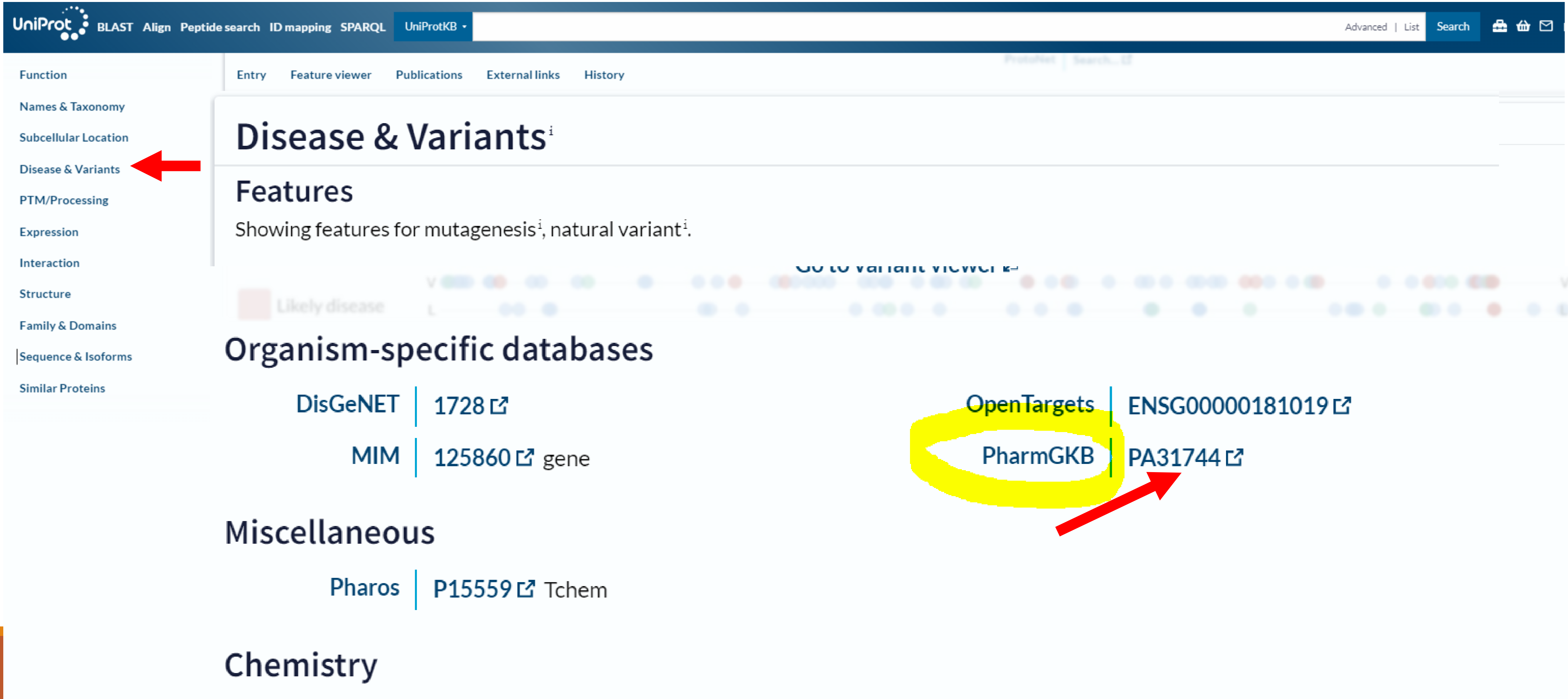


Subcellular location





Proteinové databáze: UniProt



Disease & Variants

Features

Showing features for mutagenesis, natural variant.

Organism-specific databases

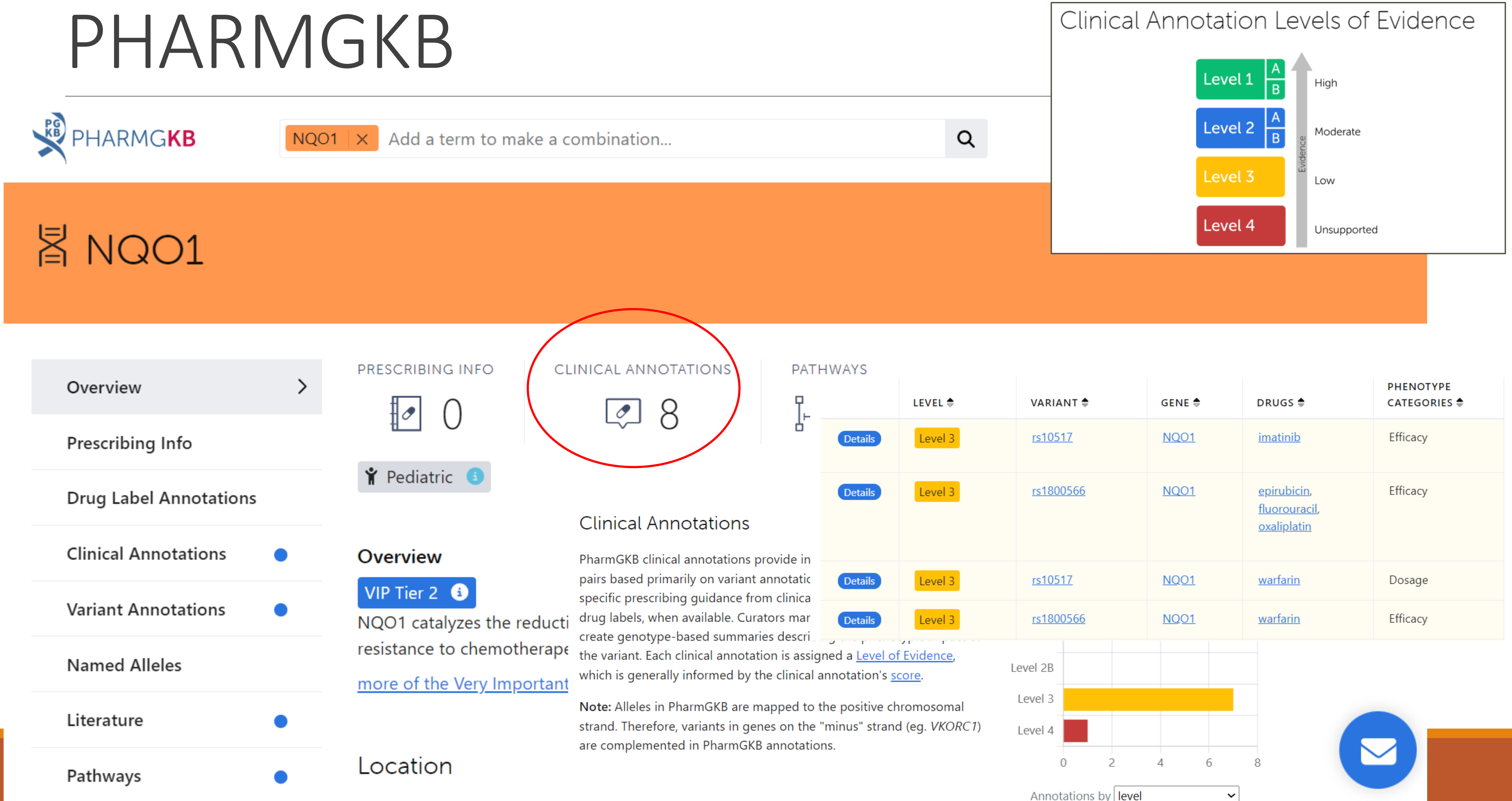
DisGeNET	1728	OpenTargets	ENSG00000181019
MIM	125860 gene	PharmGKB	PA31744

Miscellaneous

Pharos	P15559 Tchem

Chemistry

PHARMGKB



Clinical Annotation Levels of Evidence

	LEVEL	VARIANT	GENE	DRUGS	PHENOTYPE CATEGORIES
Details	Level 3	rs10517	NQO1	imatinib	Efficacy
Details	Level 3	rs1800566	NQO1	epirubicin, fluorouracil, oxaliplatin	Efficacy
Details	Level 3	rs10517	NQO1	warfarin	Dosage
Details	Level 3	rs1800566	NQO1	warfarin	Efficacy

Overview

NQO1 catalyzes the reducti
resistance to chemotherape

more of the Very Important

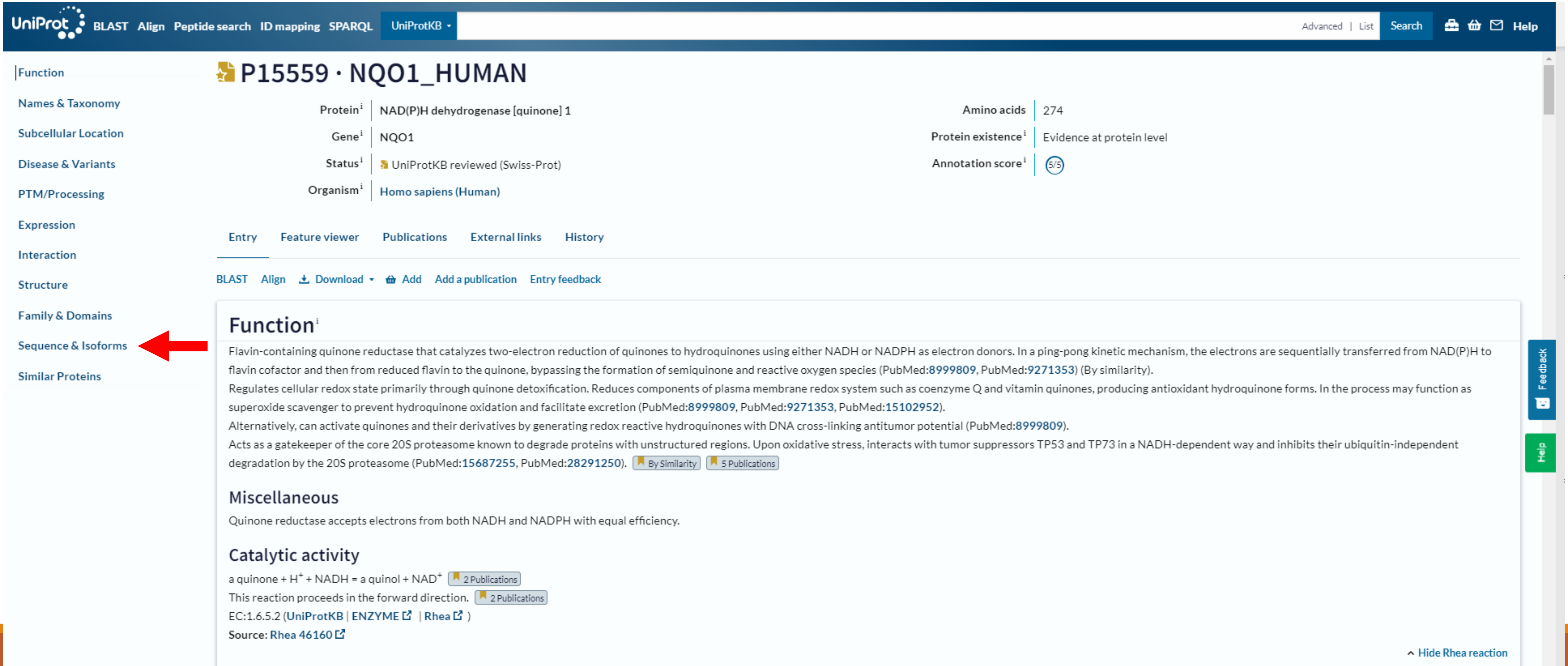
Clinical Annotations

PharmGKB clinical annotations provide in pairs based primarily on variant annotatic specific prescribing guidance from clinica drug labels, when available. Curators mar create genotype-based summaries descri the variant. Each clinical annotation is assigned a Level of Evidence, which is generally informed by the clinical annotation's score.

Note: Alleles in PharmGKB are mapped to the positive chromosomal strand. Therefore, variants in genes on the "minus" strand (eg. *VKORC1*) are complemented in PharmGKB annotations.

Location

Proteinové databáze: UniProt



P15559 · NQO1_HUMAN

Protein	NAD(P)H dehydrogenase [quinone] 1
Gene	NQO1
Status	UniProtKB reviewed (Swiss-Prot)
Organism	Homo sapiens (Human)

Amino acids	274
Protein existence	Evidence at protein level
Annotation score	5/5

Function

Flavin-containing quinone reductase that catalyzes two-electron reduction of quinones to hydroquinones using either NADH or NADPH as electron donors. In a ping-pong kinetic mechanism, the electrons are sequentially transferred from NAD(P)H to flavin cofactor and then from reduced flavin to the quinone, bypassing the formation of semiquinone and reactive oxygen species (PubMed:8999809, PubMed:9271353) (By similarity).

Regulates cellular redox state primarily through quinone detoxification. Reduces components of plasma membrane redox system such as coenzyme Q and vitamin quinones, producing antioxidant hydroquinone forms. In the process may function as superoxide scavenger to prevent hydroquinone oxidation and facilitate excretion (PubMed:8999809, PubMed:9271353, PubMed:15102952).

Alternatively, can activate quinones and their derivatives by generating redox reactive hydroquinones with DNA cross-linking antitumor potential (PubMed:8999809).

Acts as a gatekeeper of the core 20S proteasome known to degrade proteins with unstructured regions. Upon oxidative stress, interacts with tumor suppressors TP53 and TP73 in a NADH-dependent way and inhibits their ubiquitin-independent degradation by the 20S proteasome (PubMed:15687255, PubMed:28291250). By Similarity 5 Publications

Miscellaneous

Quinone reductase accepts electrons from both NADH and NADPH with equal efficiency.

Catalytic activity

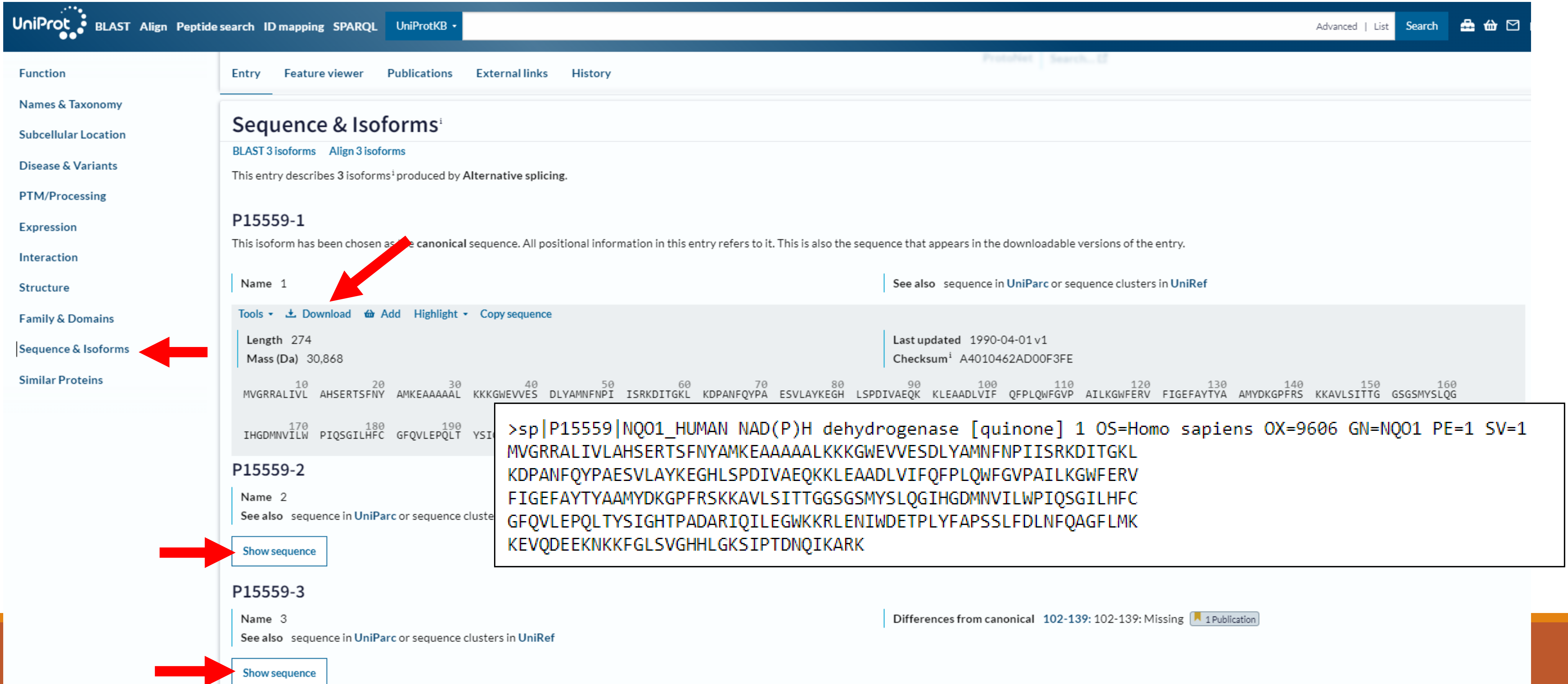
a quinone + H^+ + NADH = a quinol + NAD^+ 2 Publications

This reaction proceeds in the forward direction. 2 Publications

EC:1.6.5.2 (UniProtKB | ENZYME | Rhea)

Source: Rhea 46160

Proteinové databáze: UniProt



Sequence & Isoforms

This entry describes **3** isoforms produced by **Alternative splicing**.

P15559-1

This isoform has been chosen as the **canonical** sequence. All positional information in this entry refers to it. This is also the sequence that appears in the downloadable versions of the entry.

Name 1

See also sequence in UniParc or sequence clusters in UniRef

Length 274
Mass (Da) 30,868

Last updated 1990-04-01 v1
Checksum A4010462AD00F3FE

```
        10         20         30         40         50         60         70         80         90        100        110        120        130        140        150        160
MVGRRALIVL AHSERTSFNY AMKEAAAAAL KKKGWEVVES DLYAMNFNPI ISRKDITGKL KDPANFQYPA ESVLAYKEGH LSPDIVAEQK KLEAADLVIF QFPLQWFGVP AILKGWFERV FIGEFAYTYA AMYDKGPFRS KKAVLSITTG GSGSMYSLQG

       170        180        190
IHGDMNVILW PIQSGILHFC GFQVLEPQLT YSI
```

```
>sp|P15559|NQO1_HUMAN NAD(P)H dehydrogenase [quinone] 1 OS=Homo sapiens OX=9606 GN=NQO1 PE=1 SV=1
MVGRRALIVLAHSERTSFNYAMKEAAAAALKKKGWEVVESDLYAMNFNPIISRKDITGKL
KDPANFQYPAESVLAYKEGHLSPDIVAEQKKLEAADLVIFQFPLQWFGVPAILKGWFERV
FIGEFAYTYAAMYDKGPFRSKKAVLSITTGGSGSMYSLQGIHGDMNVILWPIQSGILHFC
GFQVLEPQLTYSIGHTPADARIQILEGWKKRLENIWDETPLYFAPSSLFDLNFQAGFLMK
KEVQDEEKNKKFGLSVGHHLGKSIPTDNQIKARK
```

P15559-2

Name 2
See also sequence in UniParc or sequence cluste

P15559-3

Name 3
See also sequence in UniParc or sequence clusters in UniRef

Differences from canonical 102-139: 102-139: Missing

FASTA (a RAW) formát

název odvozen:

FASTA = populární program pro porovnávání sekvencí a prohledávání databází

W.R. Pearson a D.J. Lipman 1988

následující formát sekvence:

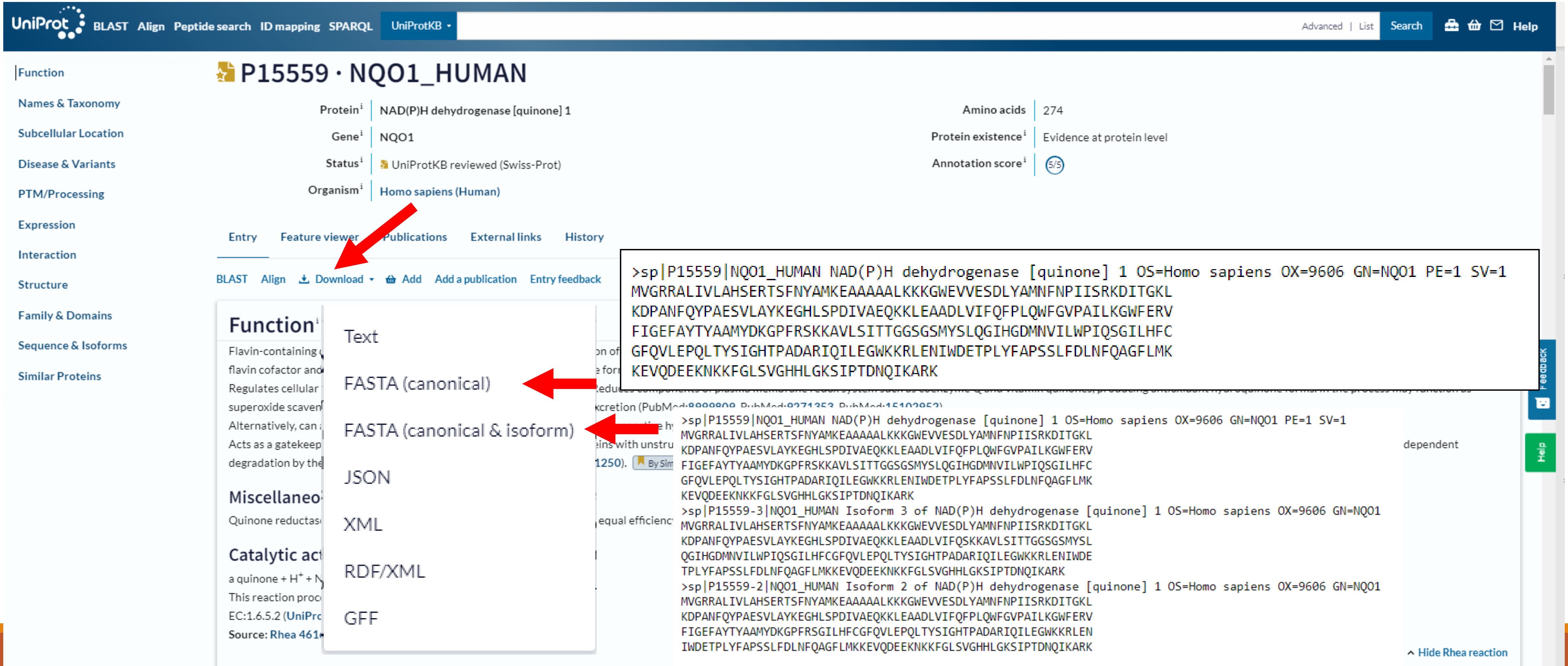
```
>NQO1_homo
MVGRRALIVLAHSERTSFNYAMKEAAAAA
LKKKGWEVVESDLYAMNFNPIISRKDITG
KLKDPANFQYPA
```

```
MVGRRALIVLAHSERTSFNYAMKEAAAAA
LKKKGWEVVESDLYAMNFNPIISRKDITG
KLKDPANFQYPA
```

1. „definiční“ řádek (hlavička) začíná **>** a obsahuje unikátní identifikátor, následovaný krátkým popisem (nemusí být).

2. řádek – vlastní sekvence (DNA/protein- v jedno-písmenném kódování) až po další symbol >, určující začátek další sekvence.

Proteinové databáze: UniProt



P15559 · NQO1_HUMAN

Protein: NAD(P)H dehydrogenase [quinone] 1
Gene: NQO1
Status: UniProtKB reviewed (Swiss-Prot)
Organism: Homo sapiens (Human)
Amino acids: 274
Protein existence: Evidence at protein level
Annotation score: 5/5

```
>sp|P15559|NQO1_HUMAN NAD(P)H dehydrogenase [quinone] 1 OS=Homo sapiens OX=9606 GN=NQO1 PE=1 SV=1
MVGRRALIVLAHSERTSFNYAMKEAAAAALKKKGWEVVESDLYAMNFNPIISRKDITGKL
KDPANFQYPAESVLAYKEGHLSPDIVAEQKKLEAADLVIFQFPLQWFGVPAILKGWFERV
FIGEFAYTYAAMYDKGPFRSKKAVLSITTGGSGSMYSLQGIHGDMNVILWPIQSGILHFC
GFQVLEPQLTYSIGHTPADARIQILEGWKKRLENIWDETPLYFAPSSLFDLNFQAGFLMK
KEVQDEEKNKKFGLSVGHHLGKSIPTDNQIKARK
```

```
>sp|P15559|NQO1_HUMAN NAD(P)H dehydrogenase [quinone] 1 OS=Homo sapiens OX=9606 GN=NQO1 PE=1 SV=1
MVGRRALIVLAHSERTSFNYAMKEAAAAALKKKGWEVVESDLYAMNFNPIISRKDITGKL
KDPANFQYPAESVLAYKEGHLSPDIVAEQKKLEAADLVIFQFPLQWFGVPAILKGWFERV
FIGEFAYTYAAMYDKGPFRSKKAVLSITTGGSGSMYSLQGIHGDMNVILWPIQSGILHFC
GFQVLEPQLTYSIGHTPADARIQILEGWKKRLENIWDETPLYFAPSSLFDLNFQAGFLMK
KEVQDEEKNKKFGLSVGHHLGKSIPTDNQIKARK
>sp|P15559-3|NQO1_HUMAN Isoform 3 of NAD(P)H dehydrogenase [quinone] 1 OS=Homo sapiens OX=9606 GN=NQO1
MVGRRALIVLAHSERTSFNYAMKEAAAAALKKKGWEVVESDLYAMNFNPIISRKDITGKL
KDPANFQYPAESVLAYKEGHLSPDIVAEQKKLEAADLVIFQSKKAVLSITTGGSGSMYSL
QGIHGDMNVILWPIQSGILHFCGFQVLEPQLTYSIGHTPADARIQILEGWKKRLENIWDE
TPLYFAPSSLFDLNFQAGFLMKKEVQDEEKNKKFGLSVGHHLGKSIPTDNQIKARK
>sp|P15559-2|NQO1_HUMAN Isoform 2 of NAD(P)H dehydrogenase [quinone] 1 OS=Homo sapiens OX=9606 GN=NQO1
MVGRRALIVLAHSERTSFNYAMKEAAAAALKKKGWEVVESDLYAMNFNPIISRKDITGKL
KDPANFQYPAESVLAYKEGHLSPDIVAEQKKLEAADLVIFQFPLQWFGVPAILKGWFERV
FIGEFAYTYAAMYDKGPFRSGILHFCGFQVLEPQLTYSIGHTPADARIQILEGWKKRLEN
IWDETPLYFAPSSLFDLNFQAGFLMKKEVQDEEKNKKFGLSVGHHLGKSIPTDNQIKARK
```

Procvičování v UniProt

Najděte si „svou vlastní“ sekvenci

-podívejte se na: lokalizaci, počet isoforem…

- uložte si FASTA formát

Procvičování v UniProt

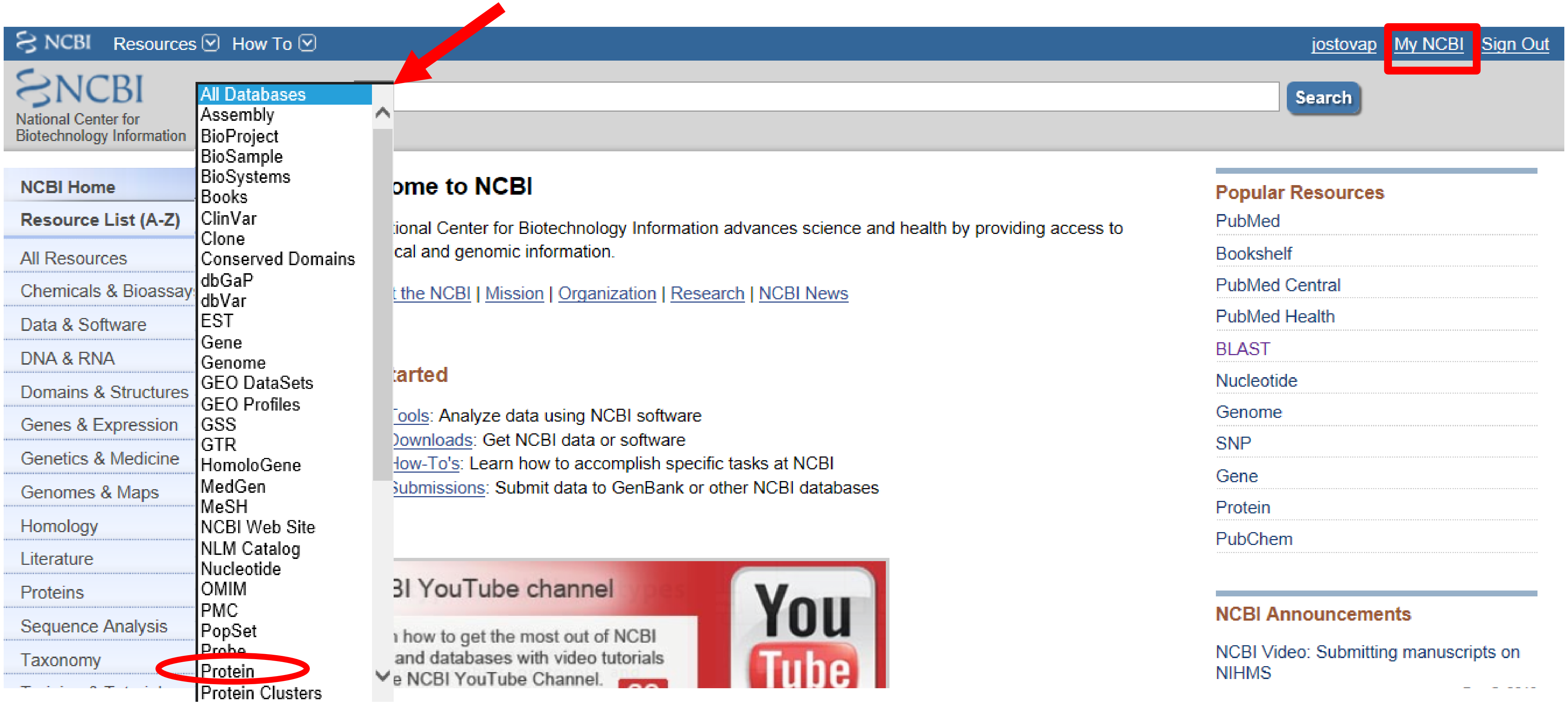
Najděte přístupový kód lidského cytochromu P450 isoformy 3A4 (CYP3A4)

Jakou má funkci a kde je lokalizován enzym Q9C0B1?

Jakou má funkci enzym O95251 a kolik má isoforem?

Procvičování v UniProt

Proteinové databáze: NCBI



ome to NCBI

ional Center for Biotechnology Information advances science and health by providing access to
cal and genomic information.

t the NCBI | Mission | Organization | Research | NCBI News

tarted

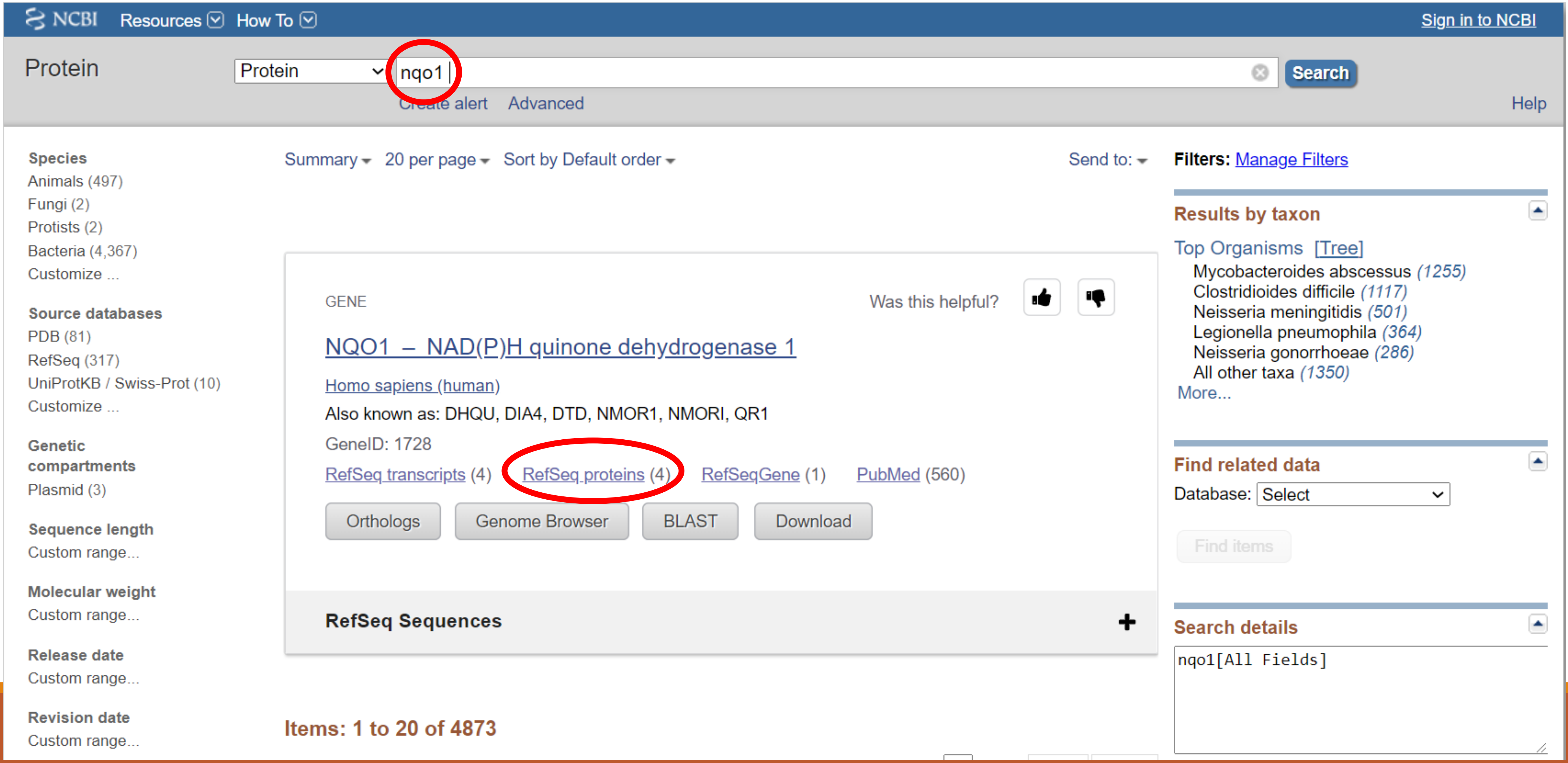
ools: Analyze data using NCBI software
ownloads: Get NCBI data or software
ow-To's: Learn how to accomplish specific tasks at NCBI
ubmissions: Submit data to GenBank or other NCBI databases

BI YouTube channel

how to get the most out of NCBI
and databases with video tutorials
e NCBI YouTube Channel.

NCBI Video: Submitting manuscripts on NIHMS

Proteinové databáze: NCBI



Protein

Protein nqo1

Species
Animals (497)
Fungi (2)
Protists (2)
Bacteria (4,367)
Customize ...

Source databases
PDB (81)
RefSeq (317)
UniProtKB / Swiss-Prot (10)
Customize ...

Genetic compartments
Plasmid (3)

Sequence length
Custom range...

Molecular weight
Custom range...

Release date
Custom range...

Revision date
Custom range...

GENE

Was this helpful?

NQO1 – NAD(P)H quinone dehydrogenase 1

Homo sapiens (human)

Also known as: DHQU, DIA4, DTD, NMOR1, NMORI, QR1

GeneID: 1728

RefSeq transcripts (4) RefSeq proteins (4) RefSeqGene (1) PubMed (560)

RefSeq Sequences

Items: 1 to 20 of 4873

Filters: Manage Filters

Results by taxon

Top Organisms [Tree]
Mycobacteroides abscessus *(1255)*
Clostridioides difficile *(1117)*
Neisseria meningitidis *(501)*
Legionella pneumophila *(364)*
Neisseria gonorrhoeae *(286)*
All other taxa *(1350)*
More...

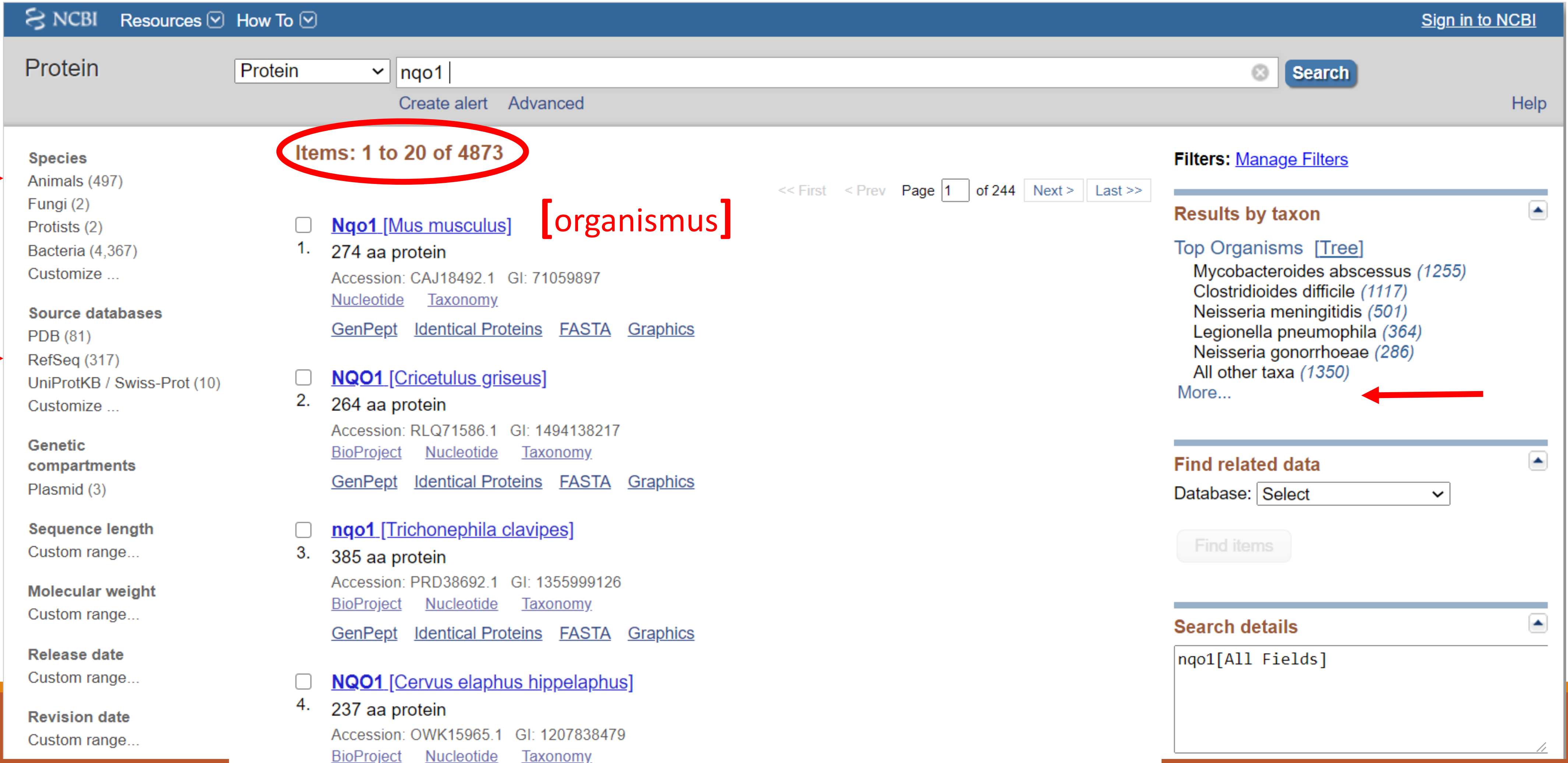
Find related data

Database: Select

Search details

nqo1[All Fields]

Proteinové databáze: NCBI



NCBI Resources How To Sign in to NCBI

Protein

Protein ▾ nqo1 Search

Create alert Advanced Help

Species
Animals (497)
Fungi (2)
Protists (2)
Bacteria (4,367)
Customize ...

Source databases
PDB (81)
RefSeq (317)
UniProtKB / Swiss-Prot (10)
Customize ...

Genetic compartments
Plasmid (3)

Sequence length
Custom range...

Molecular weight
Custom range...

Release date
Custom range...

Revision date
Custom range...

Items: 1 to 20 of 4873

<< First < Prev Page 1 of 244 Next > Last >>

[organismus]

☐ 1. Nqo1 [Mus musculus]
274 aa protein
Accession: CAJ18492.1 GI: 71059897
Nucleotide Taxonomy
GenPept Identical Proteins FASTA Graphics

☐ 2. NQO1 [Cricetulus griseus]
264 aa protein
Accession: RLQ71586.1 GI: 1494138217
BioProject Nucleotide Taxonomy
GenPept Identical Proteins FASTA Graphics

☐ 3. nqo1 [Trichonephila clavipes]
385 aa protein
Accession: PRD38692.1 GI: 1355999126
BioProject Nucleotide Taxonomy
GenPept Identical Proteins FASTA Graphics

☐ 4. NQO1 [Cervus elaphus hippelaphus]
237 aa protein
Accession: OWK15965.1 GI: 1207838479
BioProject Nucleotide Taxonomy

Filters: Manage Filters

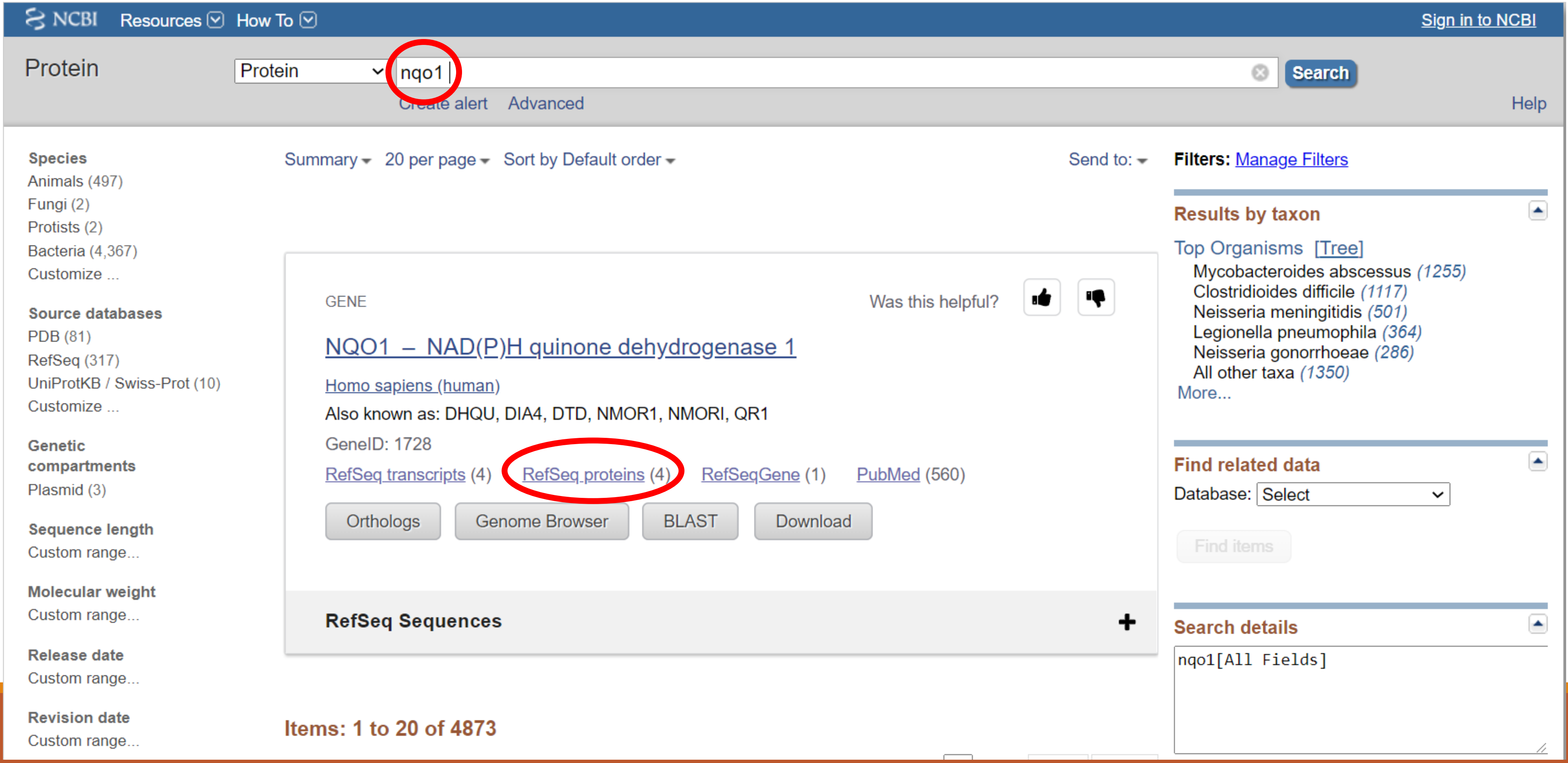
Results by taxon

Top Organisms [Tree]
Mycobacteroides abscessus *(1255)*
Clostridioides difficile *(1117)*
Neisseria meningitidis *(501)*
Legionella pneumophila *(364)*
Neisseria gonorrhoeae *(286)*
All other taxa *(1350)*
More...

Find related data
Database: Select ▾
Find items

Search details
nqo1[All Fields]

Proteinové databáze: NCBI



NCBI Resources How To Sign in to NCBI

Protein

Protein nqo1 Search

Create alert Advanced Help

Species
Animals (497)
Fungi (2)
Protists (2)
Bacteria (4,367)
Customize ...

Source databases
PDB (81)
RefSeq (317)
UniProtKB / Swiss-Prot (10)
Customize ...

Genetic compartments
Plasmid (3)

Sequence length
Custom range...

Molecular weight
Custom range...

Release date
Custom range...

Revision date
Custom range...

Summary 20 per page Sort by Default order Send to:

GENE Was this helpful?

NQO1 – NAD(P)H quinone dehydrogenase 1

Homo sapiens (human)

Also known as: DHQU, DIA4, DTD, NMOR1, NMORI, QR1

GeneID: 1728

RefSeq transcripts (4) RefSeq proteins (4) RefSeqGene (1) PubMed (560)

Orthologs Genome Browser BLAST Download

RefSeq Sequences +

Items: 1 to 20 of 4873

Filters: Manage Filters

Results by taxon

Top Organisms [Tree]
Mycobacteroides abscessus *(1255)*
Clostridioides difficile *(1117)*
Neisseria meningitidis *(501)*
Legionella pneumophila *(364)*
Neisseria gonorrhoeae *(286)*
All other taxa *(1350)*
More...

Find related data

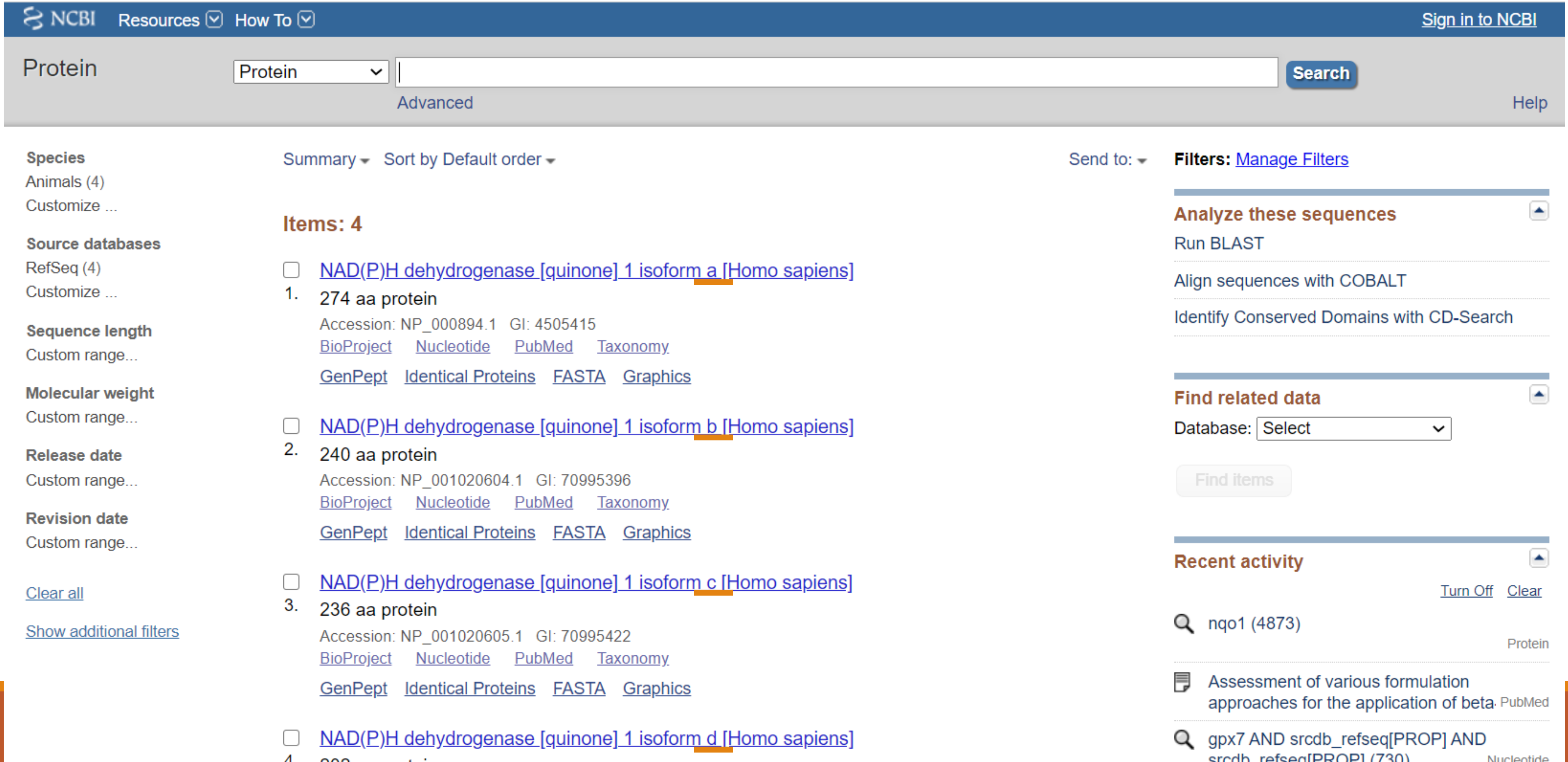
Database: Select

Find items

Search details

nqo1[All Fields]

Proteinové databáze: NCBI



Items: 4

☐ 1. NAD(P)H dehydrogenase [quinone] 1 isoform a [Homo sapiens]
274 aa protein
Accession: NP_000894.1 GI: 4505415
BioProject Nucleotide PubMed Taxonomy
GenPept Identical Proteins FASTA Graphics

☐ 2. NAD(P)H dehydrogenase [quinone] 1 isoform b [Homo sapiens]
240 aa protein
Accession: NP_001020604.1 GI: 70995396
BioProject Nucleotide PubMed Taxonomy
GenPept Identical Proteins FASTA Graphics

☐ 3. NAD(P)H dehydrogenase [quinone] 1 isoform c [Homo sapiens]
236 aa protein
Accession: NP_001020605.1 GI: 70995422
BioProject Nucleotide PubMed Taxonomy
GenPept Identical Proteins FASTA Graphics

☐ NAD(P)H dehydrogenase [quinone] 1 isoform d [Homo sapiens]

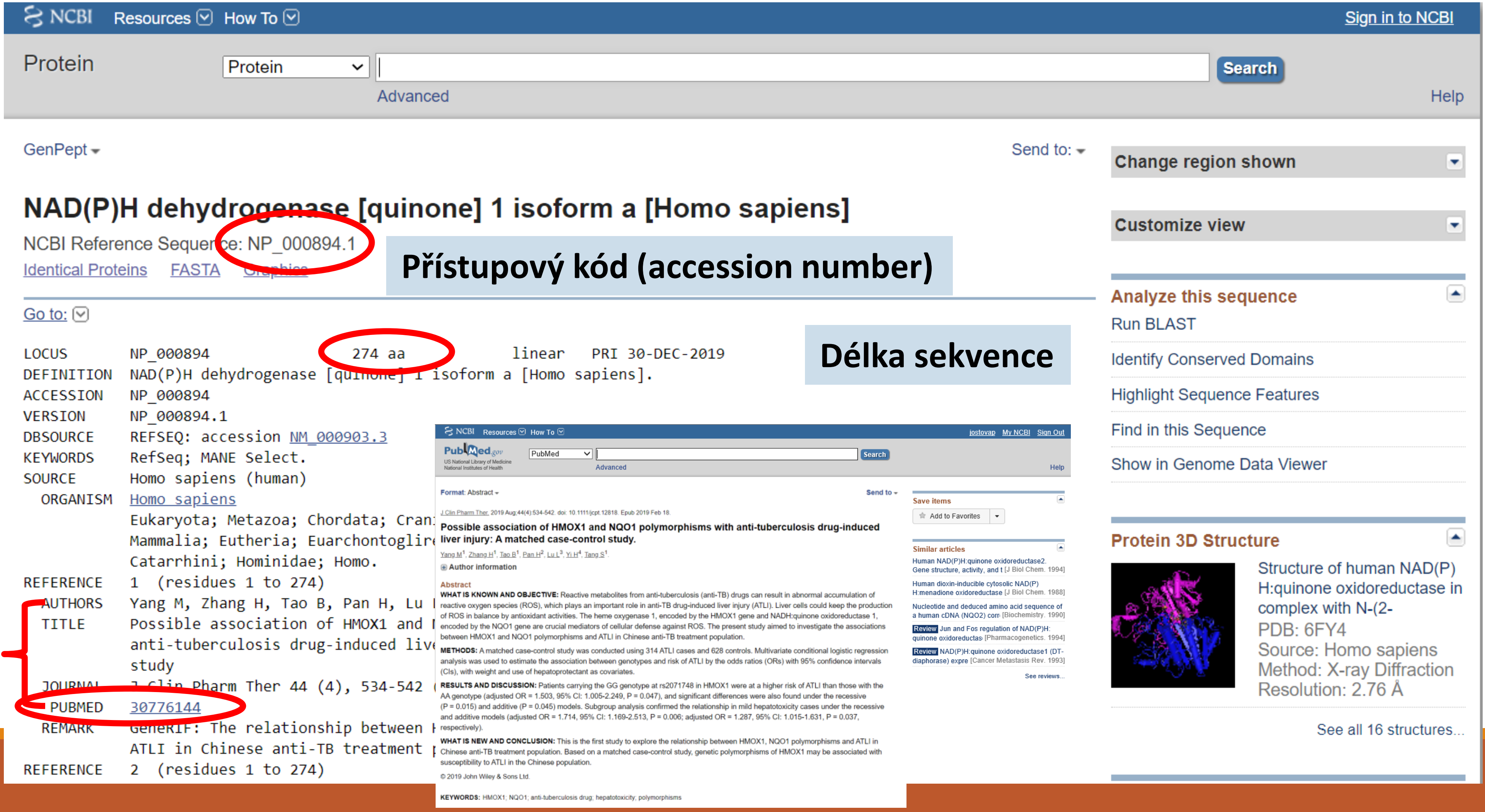
Recent activity:

nqo1 (4873) Protein

Assessment of various formulation approaches for the application of beta PubMed

gpx7 AND srcdb_refseq[PROP] AND

Proteinové databáze: NCBI



NAD(P)H dehydrogenase [quinone] 1 isoform a [Homo sapiens]

NCBI Reference Sequence: NP_000894.1

Přístupový kód (accession number)

Délka sekvence

```
LOCUS       NP_000894               274 aa            linear   PRI 30-DEC-2019
DEFINITION  NAD(P)H dehydrogenase [quinone] 1 isoform a [Homo sapiens].
ACCESSION   NP_000894
VERSION     NP_000894.1
DBSOURCE    REFSEQ: accession NM_000903.3
KEYWORDS    RefSeq; MANE Select.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Cran
            Mammalia; Eutheria; Euarchontoglir
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (residues 1 to 274)
  AUTHORS   Yang M, Zhang H, Tao B, Pan H, Lu
  TITLE     Possible association of HMOX1 and
            anti-tuberculosis drug-induced liv
            study
  JOURNAL   J Clin Pharm Ther 44 (4), 534-542
   PUBMED   30776144
  REMARK    GeneRIF: The relationship between
            ATLI in Chinese anti-TB treatment
REFERENCE   2  (residues 1 to 274)
```

Structure of human NAD(P)H:quinone oxidoreductase in complex with N-(2-
PDB: 6FY4
Source: Homo sapiens
Method: X-ray Diffraction
Resolution: 2.76 Å

J Clin Pharm Ther. 2019 Aug;44(4):534-542. doi: 10.1111/jcpt.12818. Epub 2019 Feb 18.

Possible association of HMOX1 and NQO1 polymorphisms with anti-tuberculosis drug-induced liver injury: A matched case-control study.

Yang M[1], Zhang H[1], Tao B[1], Pan H[2], Lu L[3], Yi H[4], Tang S[1].

Author information

Abstract

WHAT IS KNOWN AND OBJECTIVE: Reactive metabolites from anti-tuberculosis (anti-TB) drugs can result in abnormal accumulation of reactive oxygen species (ROS), which plays an important role in anti-TB drug-induced liver injury (ATLI). Liver cells could keep the production of ROS in balance by antioxidant activities. The heme oxygenase 1, encoded by the HMOX1 gene and NADH:quinone oxidoreductase 1, encoded by the NQO1 gene are crucial mediators of cellular defense against ROS. The present study aimed to investigate the associations between HMOX1 and NQO1 polymorphisms and ATLI in Chinese anti-TB treatment population.

METHODS: A matched case-control study was conducted using 314 ATLI cases and 628 controls. Multivariate conditional logistic regression analysis was used to estimate the association between genotypes and risk of ATLI by the odds ratios (ORs) with 95% confidence intervals (CIs), with weight and use of hepatoprotectant as covariates.

RESULTS AND DISCUSSION: Patients carrying the GG genotype at rs2071748 in HMOX1 were at a higher risk of ATLI than those with the AA genotype (adjusted OR = 1.503, 95% CI: 1.005-2.249, P = 0.047), and significant differences were also found under the recessive (P = 0.015) and additive (P = 0.045) models. Subgroup analysis confirmed the relationship in mild hepatotoxicity cases under the recessive and additive models (adjusted OR = 1.714, 95% CI: 1.169-2.513, P = 0.006; adjusted OR = 1.287, 95% CI: 1.015-1.631, P = 0.037, respectively).

WHAT IS NEW AND CONCLUSION: This is the first study to explore the relationship between HMOX1, NQO1 polymorphisms and ATLI in Chinese anti-TB treatment population. Based on a matched case-control study, genetic polymorphisms of HMOX1 may be associated with susceptibility to ATLI in the Chinese population.

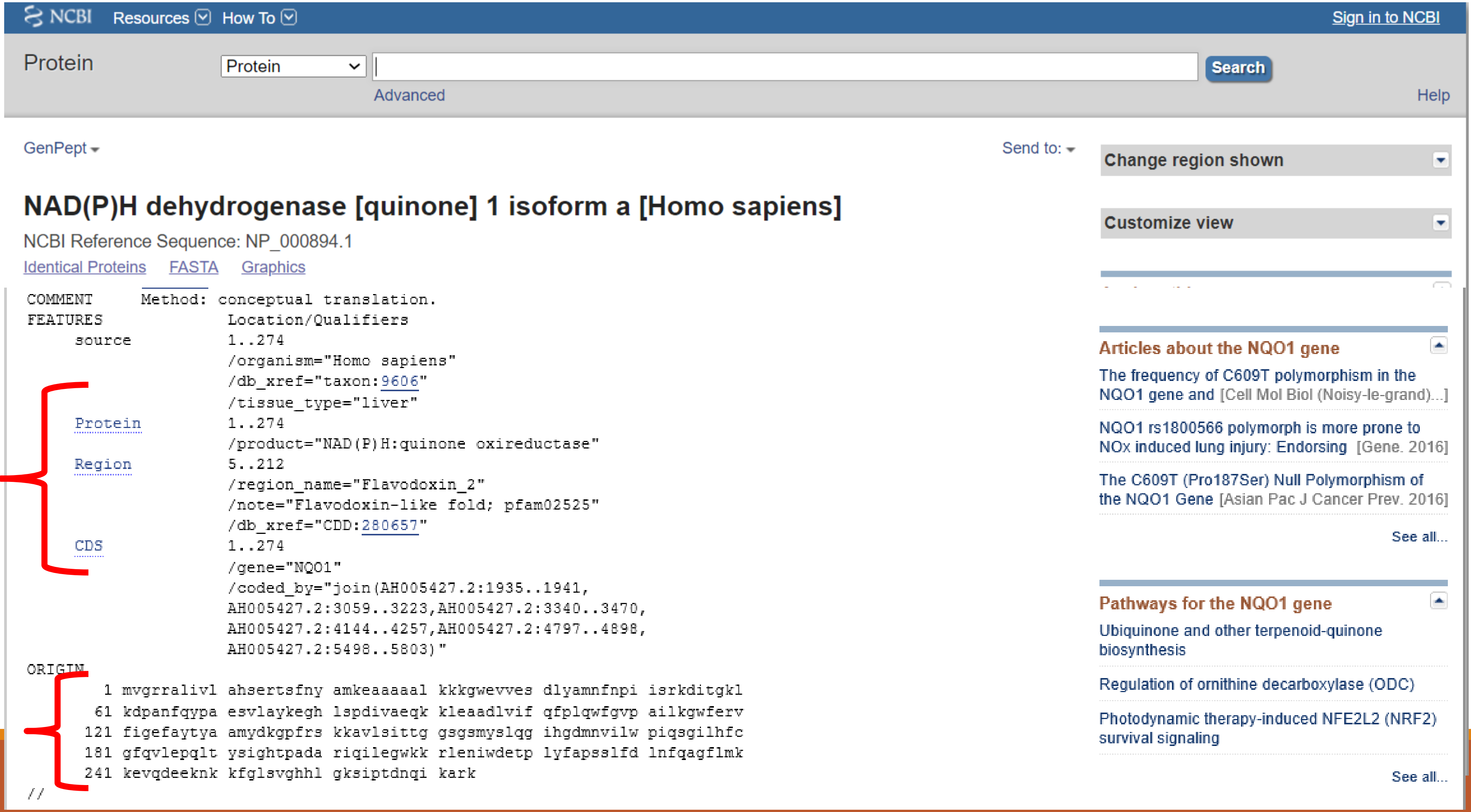
© 2019 John Wiley & Sons Ltd.

KEYWORDS: HMOX1; NQO1; anti-tuberculosis drug; hepatotoxicity; polymorphisms

Similar articles

- Human NAD(P)H:quinone oxidoreductase2. Gene structure, activity, and t [J Biol Chem. 1994]
- Human dioxin-inducible cytosolic NAD(P)H:menadione oxidoreductase [J Biol Chem. 1988]
- Nucleotide and deduced amino acid sequence of a human cDNA (NQO2) corr [Biochemistry. 1990]
- Review Jun and Fos regulation of NAD(P)H:quinone oxidoreductas [Pharmacogenetics. 1994]
- Review NAD(P)H:quinone oxidoreductase1 (DT-diaphorase) expre [Cancer Metastasis Rev. 1993]

Proteinové databáze: NCBI



NAD(P)H dehydrogenase [quinone] 1 isoform a [Homo sapiens]

NCBI Reference Sequence: NP_000894.1

Identical Proteins FASTA Graphics

```
COMMENT     Method: conceptual translation.
FEATURES             Location/Qualifiers
     source          1..274
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /tissue_type="liver"
     Protein         1..274
                     /product="NAD(P)H:quinone oxireductase"
     Region          5..212
                     /region_name="Flavodoxin_2"
                     /note="Flavodoxin-like fold; pfam02525"
                     /db_xref="CDD:280657"
     CDS             1..274
                     /gene="NQO1"
                     /coded_by="join(AH005427.2:1935..1941,
                     AH005427.2:3059..3223,AH005427.2:3340..3470,
                     AH005427.2:4144..4257,AH005427.2:4797..4898,
                     AH005427.2:5498..5803)"
ORIGIN
        1 mvgrralivl ahsertsfny amkeaaaaal kkkgwevves dlyamnfnpi isrkditgkl
       61 kdpanfqypa esvlaykegh lspdivaeqk kleaadlvif qfplqwfgvp ailkgwferv
      121 figefaytya amydkgpfrs kkavlsittg gsgsmyslqg ihgdmnvilw piqsgilhfc
      181 gfqvlepqlt ysightpada riqilegwkk rleniwdetp lyfapsslfd lnfqagflmk
      241 kevqdeeknk kfglsvghhl gksiptdnqi kark
//
```

Articles about the NQO1 gene

The frequency of C609T polymorphism in the NQO1 gene and [Cell Mol Biol (Noisy-le-grand)...]

NQO1 rs1800566 polymorph is more prone to NOx induced lung injury: Endorsing [Gene. 2016]

The C609T (Pro187Ser) Null Polymorphism of the NQO1 Gene [Asian Pac J Cancer Prev. 2016]

See all...

Pathways for the NQO1 gene

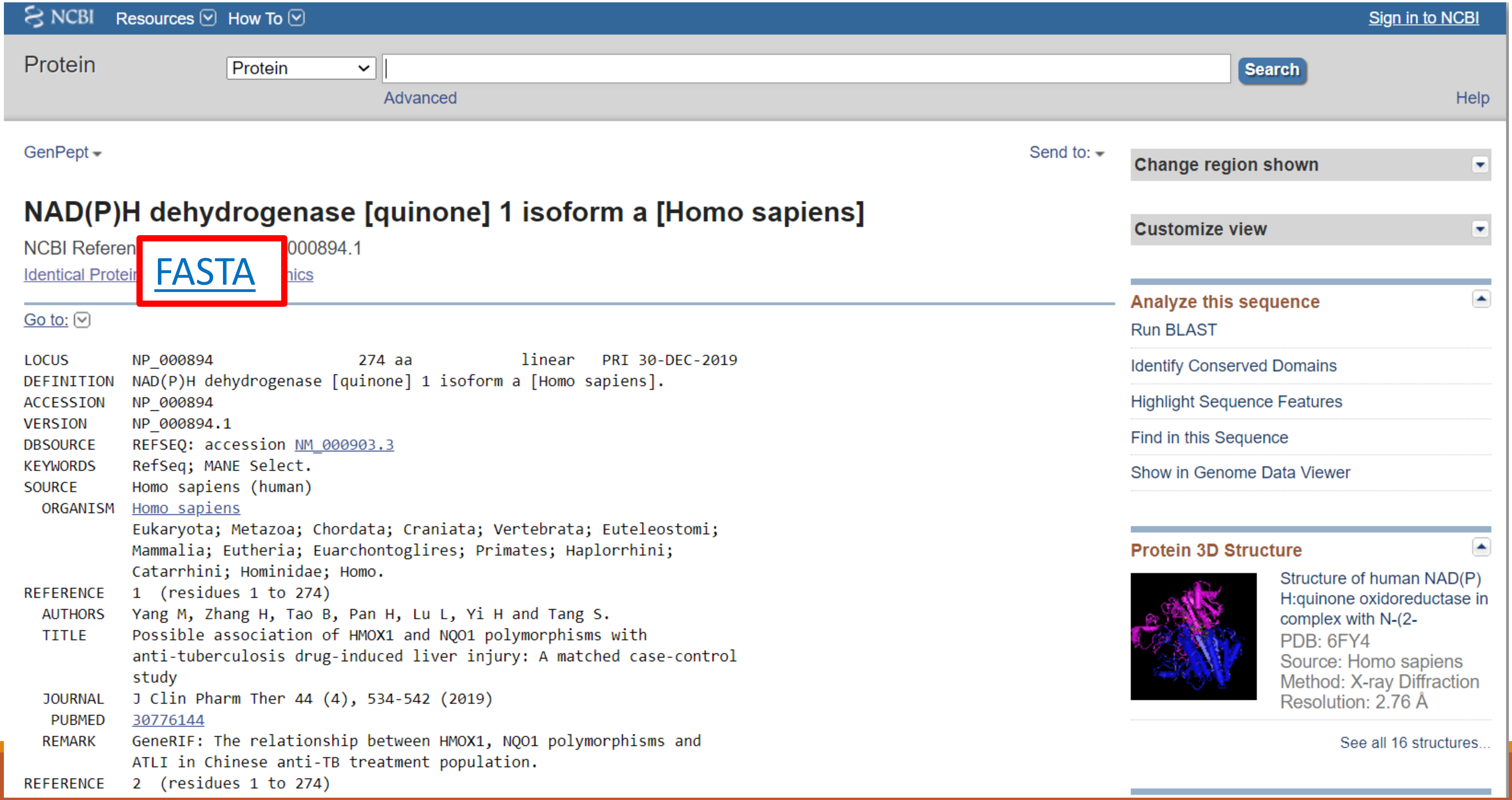
Ubiquinone and other terpenoid-quinone biosynthesis

Regulation of ornithine decarboxylase (ODC)

Photodynamic therapy-induced NFE2L2 (NRF2) survival signaling

See all...

Proteinové databáze: NCBI



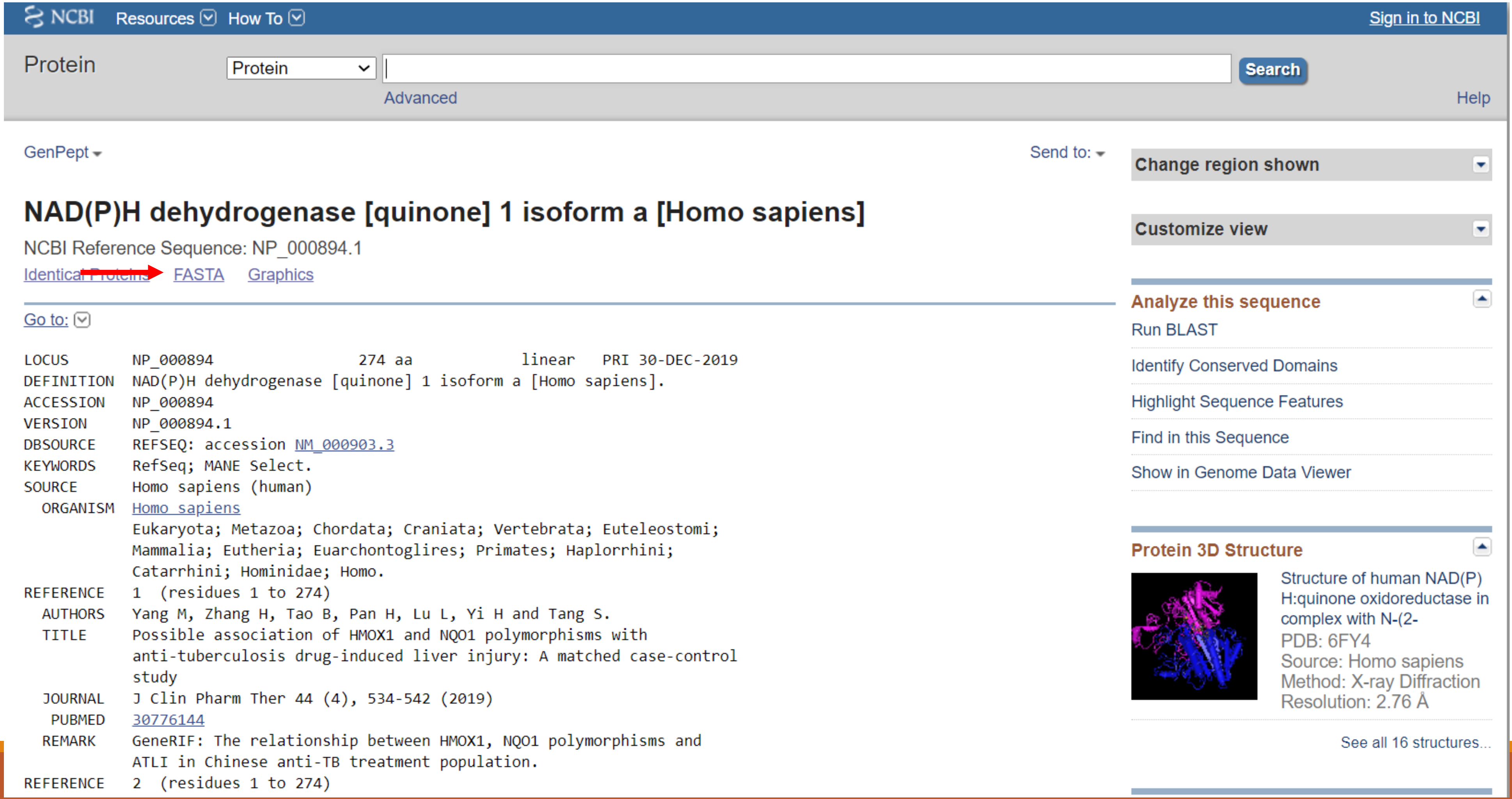
NAD(P)H dehydrogenase [quinone] 1 isoform a [Homo sapiens]

NCBI Referen 000894.1

```
LOCUS       NP_000894               274 aa            linear   PRI 30-DEC-2019
DEFINITION  NAD(P)H dehydrogenase [quinone] 1 isoform a [Homo sapiens].
ACCESSION   NP_000894
VERSION     NP_000894.1
DBSOURCE    REFSEQ: accession NM_000903.3
KEYWORDS    RefSeq; MANE Select.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (residues 1 to 274)
  AUTHORS   Yang M, Zhang H, Tao B, Pan H, Lu L, Yi H and Tang S.
  TITLE     Possible association of HMOX1 and NQO1 polymorphisms with
            anti-tuberculosis drug-induced liver injury: A matched case-control
            study
  JOURNAL   J Clin Pharm Ther 44 (4), 534-542 (2019)
   PUBMED   30776144
  REMARK    GeneRIF: The relationship between HMOX1, NQO1 polymorphisms and
            ATLI in Chinese anti-TB treatment population.
REFERENCE   2  (residues 1 to 274)
```

Structure of human NAD(P) H:quinone oxidoreductase in complex with N-(2-
PDB: 6FY4
Source: Homo sapiens
Method: X-ray Diffraction
Resolution: 2.76 Å

Proteinové databáze: NCBI



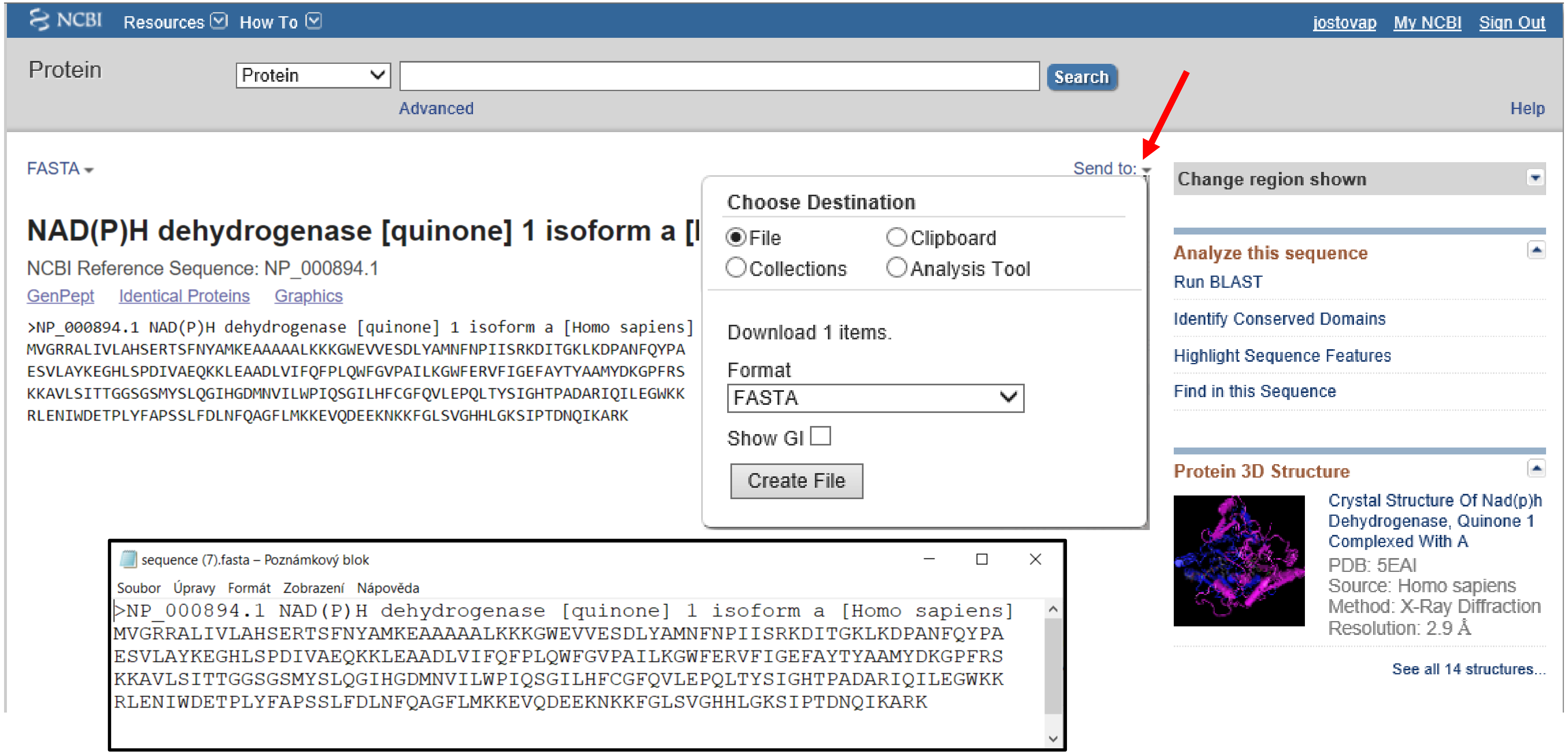
NAD(P)H dehydrogenase [quinone] 1 isoform a [Homo sapiens]

NCBI Reference Sequence: NP_000894.1

```
LOCUS       NP_000894                274 aa            linear   PRI 30-DEC-2019
DEFINITION  NAD(P)H dehydrogenase [quinone] 1 isoform a [Homo sapiens].
ACCESSION   NP_000894
VERSION     NP_000894.1
DBSOURCE    REFSEQ: accession NM_000903.3
KEYWORDS    RefSeq; MANE Select.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (residues 1 to 274)
  AUTHORS   Yang M, Zhang H, Tao B, Pan H, Lu L, Yi H and Tang S.
  TITLE     Possible association of HMOX1 and NQO1 polymorphisms with
            anti-tuberculosis drug-induced liver injury: A matched case-control
            study
  JOURNAL   J Clin Pharm Ther 44 (4), 534-542 (2019)
   PUBMED   30776144
  REMARK    GeneRIF: The relationship between HMOX1, NQO1 polymorphisms and
            ATLI in Chinese anti-TB treatment population.
REFERENCE   2  (residues 1 to 274)
```

Structure of human NAD(P) H:quinone oxidoreductase in complex with N-(2-
PDB: 6FY4
Source: Homo sapiens
Method: X-ray Diffraction
Resolution: 2.76 Å

Proteinové databáze: NCBI



NAD(P)H dehydrogenase [quinone] 1 isoform a [H

NCBI Reference Sequence: NP_000894.1

```
>NP_000894.1 NAD(P)H dehydrogenase [quinone] 1 isoform a [Homo sapiens]
MVGRRALIVLAHSERTSFNYAMKEAAAAALKKKGWEVVESDLYAMNFNPIISRKDITGKLKDPANFQYPA
ESVLAYKEGHLSPDIVAEQKKLEAADLVIFQFPLQWFGVPAILKGWFERVFIGEFAYTYAAMYDKGPFRS
KKAVLSITTGGSGSMYSLQGIHGDMNVILWPIQSGILHFCGFQVLEPQLTYSIGHTPADARIQILEGWKK
RLENIWDETPLYFAPSSLFDLNFQAGFLMKKEVQDEEKNKKFGLSVGHHLGKSIPTDNQIKARK
```

Choose Destination

● File ○ Clipboard
○ Collections ○ Analysis Tool

Download 1 items.

Format: FASTA

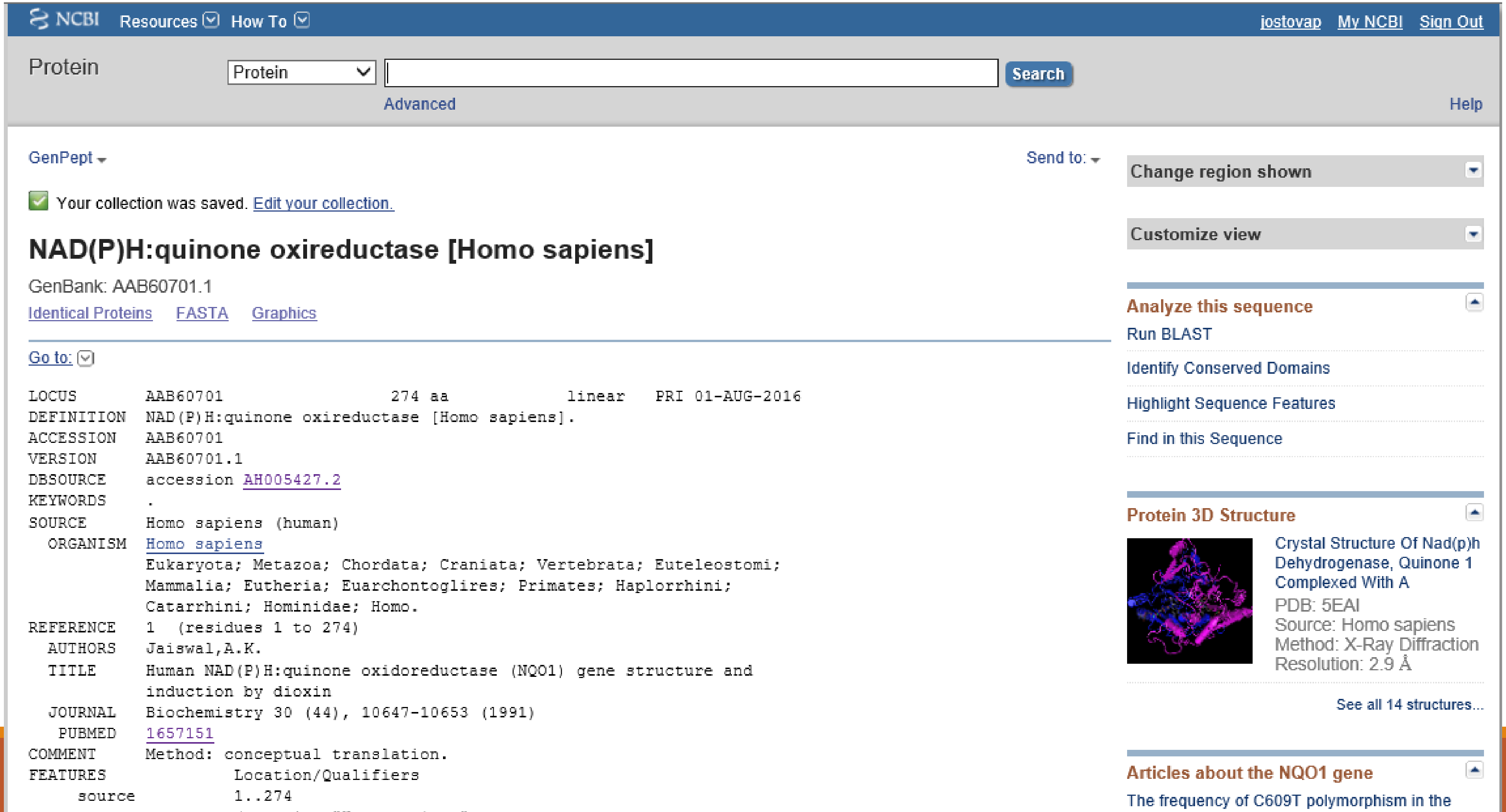
Show GI ☐

Create File

Crystal Structure Of Nad(p)h Dehydrogenase, Quinone 1 Complexed With A
PDB: 5EAI
Source: Homo sapiens
Method: X-Ray Diffraction
Resolution: 2.9 Å

```
>NP_000894.1 NAD(P)H dehydrogenase [quinone] 1 isoform a [Homo sapiens]
MVGRRALIVLAHSERTSFNYAMKEAAAAALKKKGWEVVESDLYAMNFNPIISRKDITGKLKDPANFQYPA
ESVLAYKEGHLSPDIVAEQKKLEAADLVIFQFPLQWFGVPAILKGWFERVFIGEFAYTYAAMYDKGPFRS
KKAVLSITTGGSGSMYSLQGIHGDMNVILWPIQSGILHFCGFQVLEPQLTYSIGHTPADARIQILEGWKK
RLENIWDETPLYFAPSSLFDLNFQAGFLMKKEVQDEEKNKKFGLSVGHHLGKSIPTDNQIKARK
```

Proteinové databáze: NCBI → **My NCBI**



Your collection was saved. Edit your collection.

NAD(P)H:quinone oxireductase [Homo sapiens]

GenBank: AAB60701.1

Identical Proteins FASTA Graphics

Go to:

```
LOCUS       AAB60701                 274 aa            linear   PRI 01-AUG-2016
DEFINITION  NAD(P)H:quinone oxireductase [Homo sapiens].
ACCESSION   AAB60701
VERSION     AAB60701.1
DBSOURCE    accession AH005427.2
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
            Catarrhini; Hominidae; Homo.
REFERENCE   1  (residues 1 to 274)
  AUTHORS   Jaiswal,A.K.
  TITLE     Human NAD(P)H:quinone oxidoreductase (NQO1) gene structure and
            induction by dioxin
  JOURNAL   Biochemistry 30 (44), 10647-10653 (1991)
   PUBMED   1657151
COMMENT     Method: conceptual translation.
FEATURES             Location/Qualifiers
     source          1..274
```

Analyze this sequence

Run BLAST

Identify Conserved Domains

Highlight Sequence Features

Find in this Sequence

Protein 3D Structure

Crystal Structure Of Nad(p)h Dehydrogenase, Quinone 1 Complexed With A

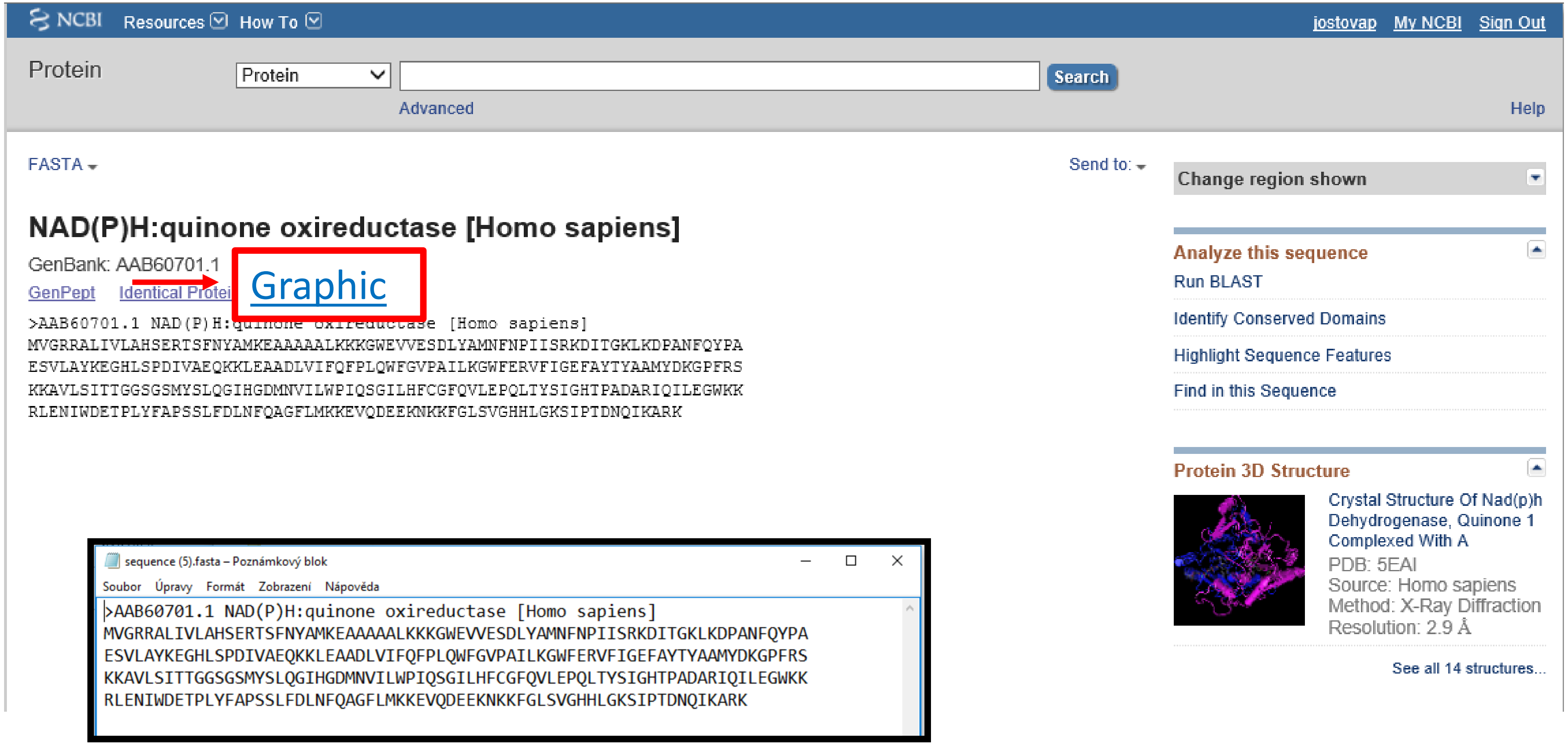
PDB: 5EAI
Source: Homo sapiens
Method: X-Ray Diffraction
Resolution: 2.9 Å

See all 14 structures...

Articles about the NQO1 gene

The frequency of C609T polymorphism in the

Proteinové databáze: NCBI



NAD(P)H:quinone oxireductase [Homo sapiens]

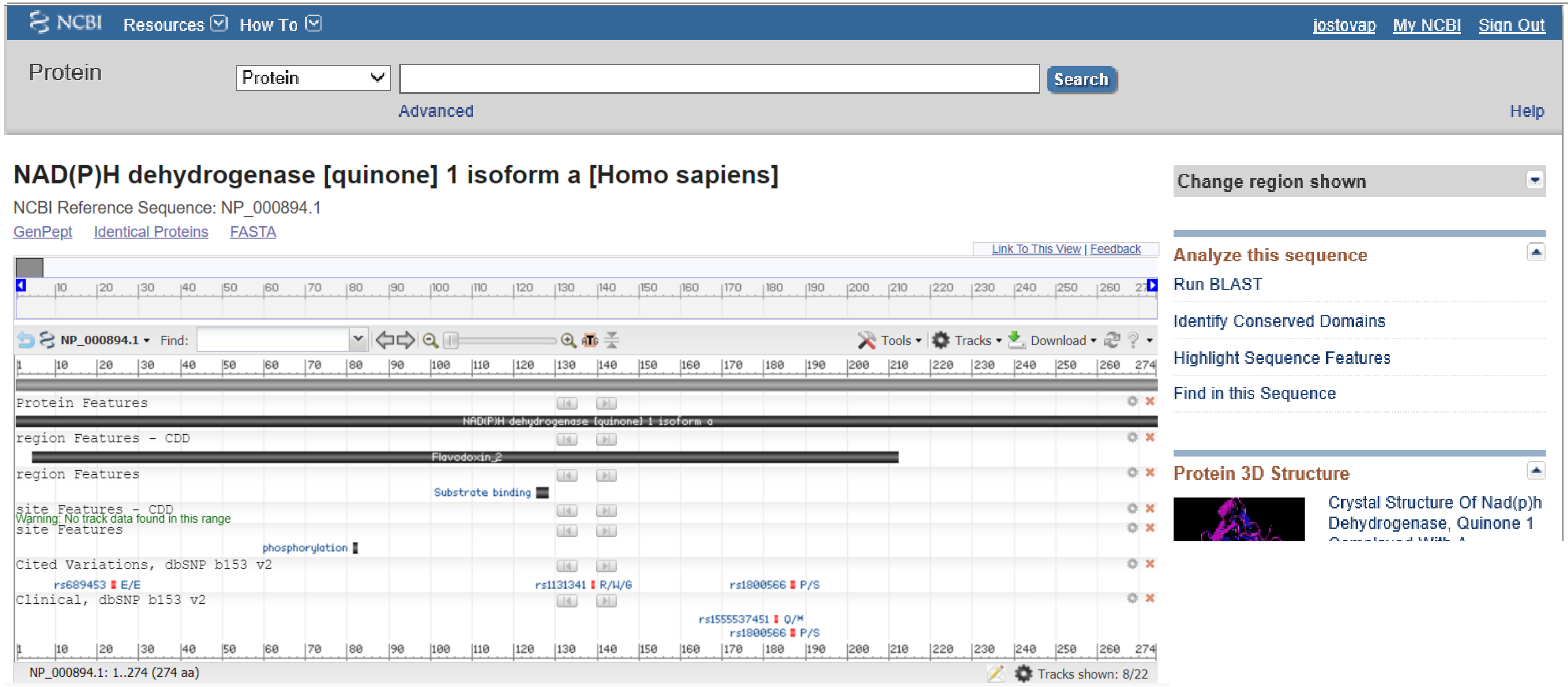
GenBank: AAB60701.1

```
>AAB60701.1 NAD(P)H:quinone oxireductase [Homo sapiens]
MVGRRALIVLAHSERTSFNYAMKEAAAAALKKKGWEVVESDLYAMNFNPIISRKDITGKLKDPANFQYPA
ESVLAYKEGHLSPDIVAEQKKLEAADLVIFQFPLQWFGVPAILKGWFERVFIGEFAYTYAAMYDKGPFRS
KKAVLSITTGGSGSMYSLQGIHGDMNVILWPIQSGILHFCGFQVLEPQLTYSIGHTPADARIQILEGWKK
RLENIWDETPLYFAPSSLFDLNFQAGFLMKKEVQDEEKNKKFGLSVGHHLGKSIPTDNQIKARK
```

Crystal Structure Of Nad(p)h Dehydrogenase, Quinone 1 Complexed With A
PDB: 5EAI
Source: Homo sapiens
Method: X-Ray Diffraction
Resolution: 2.9 Å

```
>AAB60701.1 NAD(P)H:quinone oxireductase [Homo sapiens]
MVGRRALIVLAHSERTSFNYAMKEAAAAALKKKGWEVVESDLYAMNFNPIISRKDITGKLKDPANFQYPA
ESVLAYKEGHLSPDIVAEQKKLEAADLVIFQFPLQWFGVPAILKGWFERVFIGEFAYTYAAMYDKGPFRS
KKAVLSITTGGSGSMYSLQGIHGDMNVILWPIQSGILHFCGFQVLEPQLTYSIGHTPADARIQILEGWKK
RLENIWDETPLYFAPSSLFDLNFQAGFLMKKEVQDEEKNKKFGLSVGHHLGKSIPTDNQIKARK
```

Proteinové databáze: NCBI



NAD(P)H dehydrogenase [quinone] 1 isoform a [Homo sapiens]

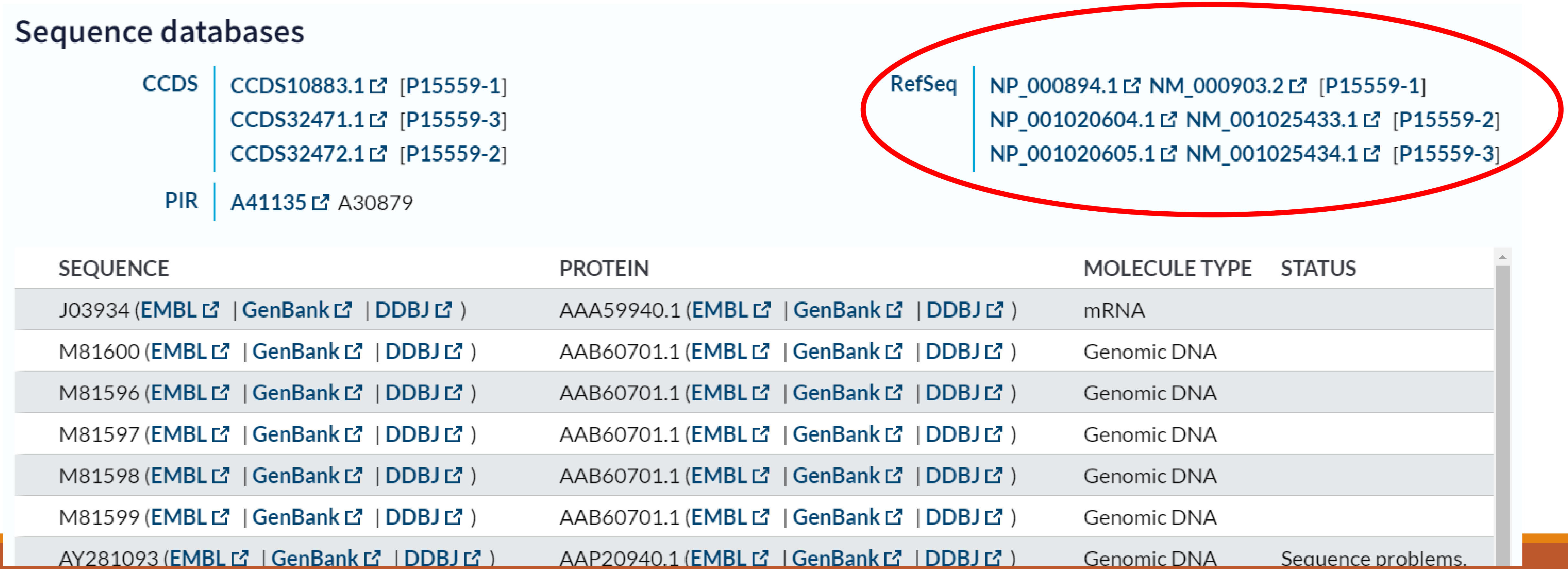
NCBI Reference Sequence: NP_000894.1

Warning: No track data found in this range

NP_000894.1: 1..274 (274 aa)

Proteinové databáze: UNIPROT →NCBI

Sequence & Isoforms - posunout se níže na:



Sequence databases

CCDS	CCDS10883.1 [P15559-1] CCDS32471.1 [P15559-3] CCDS32472.1 [P15559-2]
PIR	A41135 A30879
RefSeq	NP_000894.1 NM_000903.2 [P15559-1] NP_001020604.1 NM_001025433.1 [P15559-2] NP_001020605.1 NM_001025434.1 [P15559-3]

SEQUENCE	PROTEIN	MOLECULE TYPE	STATUS
J03934 (EMBL | GenBank | DDBJ)	AAA59940.1 (EMBL | GenBank | DDBJ)	mRNA	
M81600 (EMBL | GenBank | DDBJ)	AAB60701.1 (EMBL | GenBank | DDBJ)	Genomic DNA	
M81596 (EMBL | GenBank | DDBJ)	AAB60701.1 (EMBL | GenBank | DDBJ)	Genomic DNA	
M81597 (EMBL | GenBank | DDBJ)	AAB60701.1 (EMBL | GenBank | DDBJ)	Genomic DNA	
M81598 (EMBL | GenBank | DDBJ)	AAB60701.1 (EMBL | GenBank | DDBJ)	Genomic DNA	
M81599 (EMBL | GenBank | DDBJ)	AAB60701.1 (EMBL | GenBank | DDBJ)	Genomic DNA	
AY281093 (EMBL | GenBank | DDBJ)	AAP20940.1 (EMBL | GenBank | DDBJ)	Genomic DNA	Sequence problems.

Vyzkoušejte si NCBI proteinovou databázi

Najděte si „svou vlastní“ sekvenci

-vyzkoušejte hledat „přímo“ v NCBI
-vyzkoušejte „projít“ skrz „Sequence databases“ a „Refseq“ v Uniprot databázi

„Proteinová bioinformatika“

Vyhledávání AMK sekvencí

Analýza vlastností sekvencí (aminokyselinové složení, molekulová hmotnost, isoelektrický bod…)

Štěpení proteasami

Analýza hydrofobních segmentů, transmembránových úseků

Predikce sekundární struktury

3D-struktura, vizualizace

Vyhledání a porovnání podobných sekvencí

Evoluční příbuznost sekvencí

…

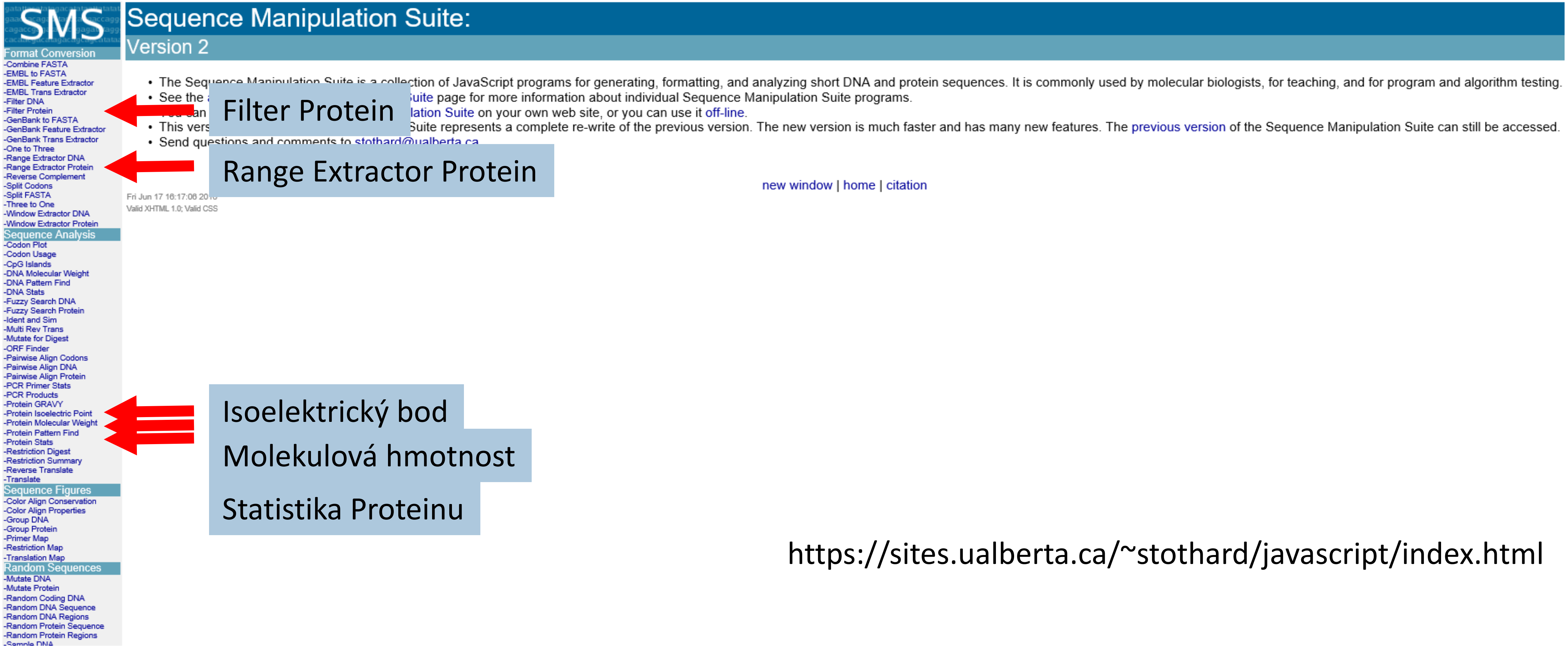
Analýza vlastností proteinu/sekvence

Molekulární hmotnost

Isoelektrický bod

Složení AMK

Analýza vlastností proteinu/sekvence



SMS

Format Conversion
- -Combine FASTA
- -EMBL to FASTA
- -EMBL Feature Extractor
- -EMBL Trans Extractor
- -Filter DNA
- -Filter Protein
- -GenBank to FASTA
- -GenBank Feature Extractor
- -GenBank Trans Extractor
- -One to Three
- -Range Extractor DNA
- -Range Extractor Protein
- -Reverse Complement
- -Split Codons
- -Split FASTA
- -Three to One
- -Window Extractor DNA
- -Window Extractor Protein

Sequence Analysis
- -Codon Plot
- -Codon Usage
- -CpG Islands
- -DNA Molecular Weight
- -DNA Pattern Find
- -DNA Stats
- -Fuzzy Search DNA
- -Fuzzy Search Protein
- -Ident and Sim
- -Multi Rev Trans
- -Mutate for Digest
- -ORF Finder
- -Pairwise Align Codons
- -Pairwise Align DNA
- -Pairwise Align Protein
- -PCR Primer Stats
- -PCR Products
- -Protein GRAVY
- -Protein Isoelectric Point
- -Protein Molecular Weight
- -Protein Pattern Find
- -Protein Stats
- -Restriction Digest
- -Restriction Summary
- -Reverse Translate
- -Translate

Sequence Figures
- -Color Align Conservation
- -Color Align Properties
- -Group DNA
- -Group Protein
- -Primer Map
- -Restriction Map
- -Translation Map

Random Sequences
- -Mutate DNA
- -Mutate Protein
- -Random Coding DNA
- -Random DNA Sequence
- -Random DNA Regions
- -Random Protein Sequence
- -Random Protein Regions
- -Sample DNA

Sequence Manipulation Suite:

Version 2

- The Sequence Manipulation Suite is a collection of JavaScript programs for generating, formatting, and analyzing short DNA and protein sequences. It is commonly used by molecular biologists, for teaching, and for program and algorithm testing.
- See the uite page for more information about individual Sequence Manipulation Suite programs.
- You can lation Suite on your own web site, or you can use it off-line.
- This vers Suite represents a complete re-write of the previous version. The new version is much faster and has many new features. The previous version of the Sequence Manipulation Suite can still be accessed.
- Send questions and comments to stothard@ualberta.ca

Fri Jun 17 16:17:08 20
Valid XHTML 1.0; Valid CSS

new window | home | citation

Filter Protein

Range Extractor Protein

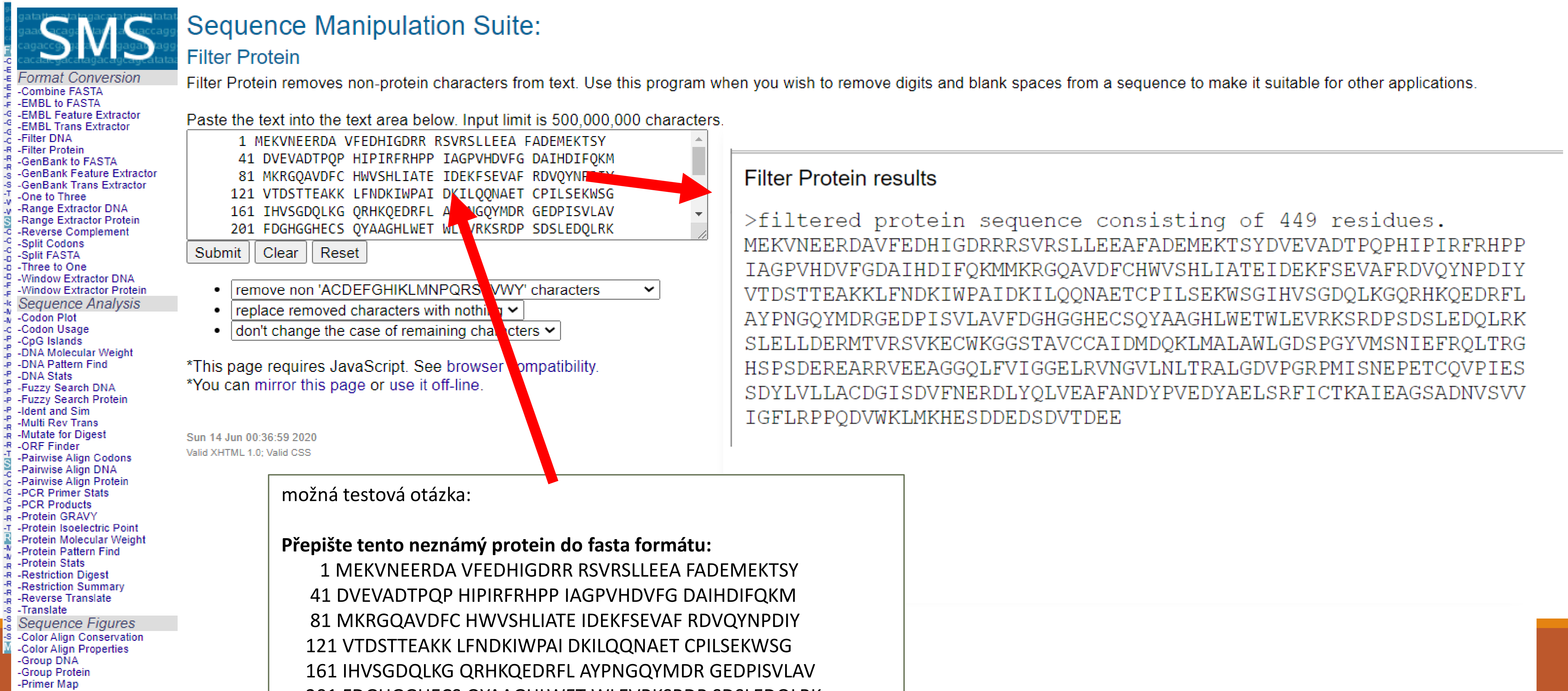
Isoelektrický bod

Molekulová hmotnost

Statistika Proteinu

https://sites.ualberta.ca/~stothard/javascript/index.html

Analýza vlastností proteinu/sekvence



Sequence Manipulation Suite:

Filter Protein

Filter Protein removes non-protein characters from text. Use this program when you wish to remove digits and blank spaces from a sequence to make it suitable for other applications.

Paste the text into the text area below. Input limit is 500,000,000 characters.

```
      1 MEKVNEERDA VFEDHIGDRR RSVRSLLEEA FADEMEKTSY
     41 DVEVADTPQP HIPIRFRHPP IAGPVHDVFG DAIHDIFQKM
     81 MKRGQAVDFC HWVSHLIATE IDEKFSEVAF RDVQYNPDIY
    121 VTDSTTEAKK LFNDKIWPAI DKILQQNAET CPILSEKWSG
    161 IHVSGDQLKG QRHKQEDRFL AYPNGQYMDR GEDPISVLAV
    201 FDGHGGHECS QYAAGHLWET WLEVRKSRDP SDSLEDQLRK
```

- remove non 'ACDEFGHIKLMNPQRSTVWY' characters
- replace removed characters with nothing
- don't change the case of remaining characters

*This page requires JavaScript. See browser compatibility.
*You can mirror this page or use it off-line.

Sun 14 Jun 00:36:59 2020
Valid XHTML 1.0; Valid CSS

Filter Protein results

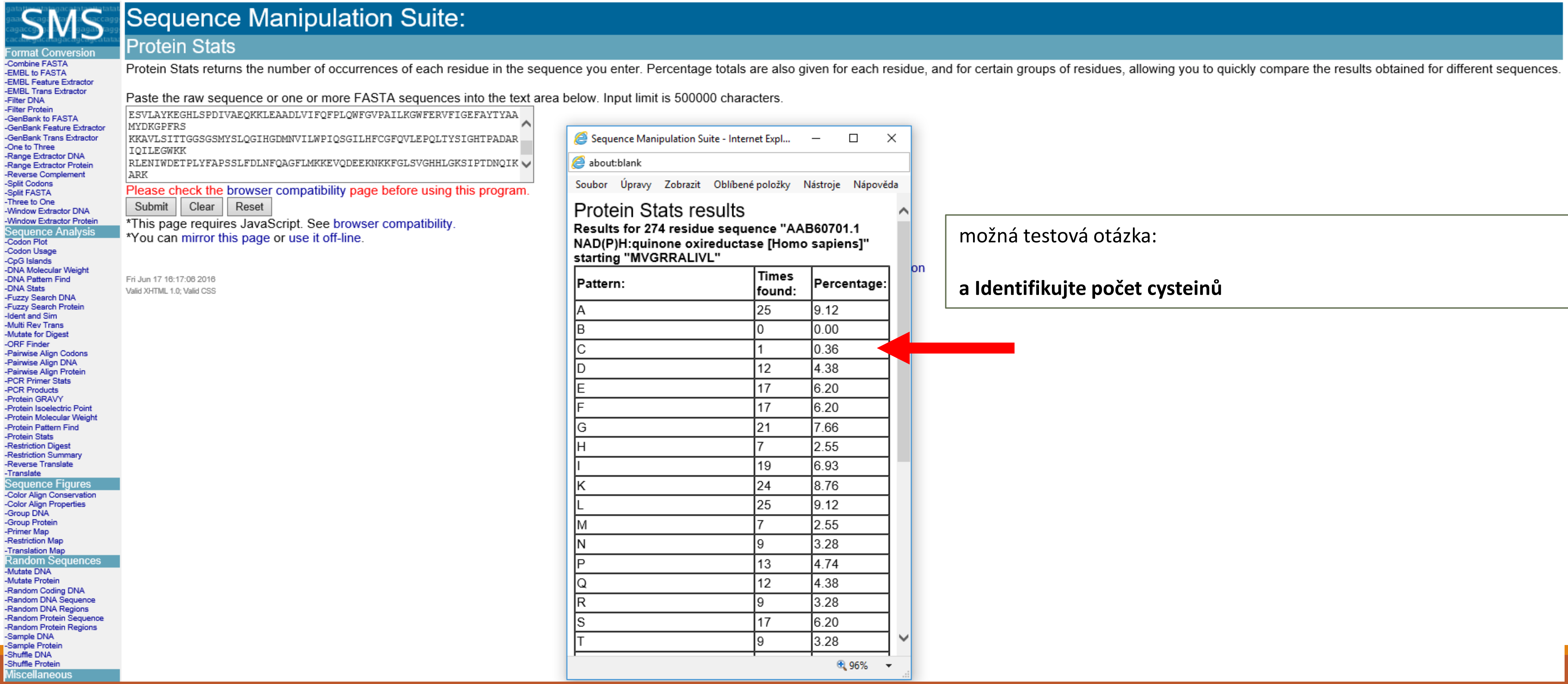
```
>filtered protein sequence consisting of 449 residues.
MEKVNEERDAVFEDHIGDRRRSVRSLLEEAFADEMEKTSYDVEVADTPQPHIPIRFRHPP
IAGPVHDVFGDAIHDIFQKMMKRGQAVDFCHWVSHLIATEIDEKFSEVAFRDVQYNPDIY
VTDSTTEAKKLFNDKIWPAIDKILQQNAETCPILSEKWSGIHVSGDQLKGQRHKQEDRFL
AYPNGQYMDRGEDPISVLAVFDGHGGHECSQYAAGHLWETWLEVRKSRDPSDSLEDQLRK
SLELLDERMTVRSVKECWKGGSTAVCCAIDMDQKLMALAWLGDSPGYVMSNIEFRQLTRG
HSPSDEREARRVEEAGGQLFVIGGELRVNGVLNLTRALGDVPGRPMISNEPETCQVPIES
SDYLVLLACDGISDVFNERDLYQLVEAFANDYPVEDYAELSRFICTKAIEAGSADNVSVV
IGFLRPPQDVWKLMKHESDDEDSDVTDEE
```

možná testová otázka:

Přepište tento neznámý protein do fasta formátu:

```
  1 MEKVNEERDA VFEDHIGDRR RSVRSLLEEA FADEMEKTSY
 41 DVEVADTPQP HIPIRFRHPP IAGPVHDVFG DAIHDIFQKM
 81 MKRGQAVDFC HWVSHLIATE IDEKFSEVAF RDVQYNPDIY
121 VTDSTTEAKK LFNDKIWPAI DKILQQNAET CPILSEKWSG
161 IHVSGDQLKG QRHKQEDRFL AYPNGQYMDR GEDPISVLAV
```

Analýza vlastností proteinu/sekvence



Sequence Manipulation Suite:

Protein Stats

Protein Stats returns the number of occurrences of each residue in the sequence you enter. Percentage totals are also given for each residue, and for certain groups of residues, allowing you to quickly compare the results obtained for different sequences.

Paste the raw sequence or one or more FASTA sequences into the text area below. Input limit is 500000 characters.

```
ESVLAYKEGHLSPDIVAEQKKLEAADLVIFQFPLQWFGVPAILKGWFERVFIGEFAYTYAA
MYDKGPFRS
KKAVLSITTGGSGSMYSLQGIHGDMNVILWPIQSGILHFCGFQVLEPQLTYSIGHTPADAR
IQILEGWKK
RLENIWDETPLYFAPSSLFDLNFQAGFLMKKEVQDEEKNKKFGLSVGHHLGKSIPTDNQIK
ARK
```

Please check the browser compatibility page before using this program.

*This page requires JavaScript. See browser compatibility.
*You can mirror this page or use it off-line.

Fri Jun 17 16:17:06 2016
Valid XHTML 1.0; Valid CSS

Protein Stats results

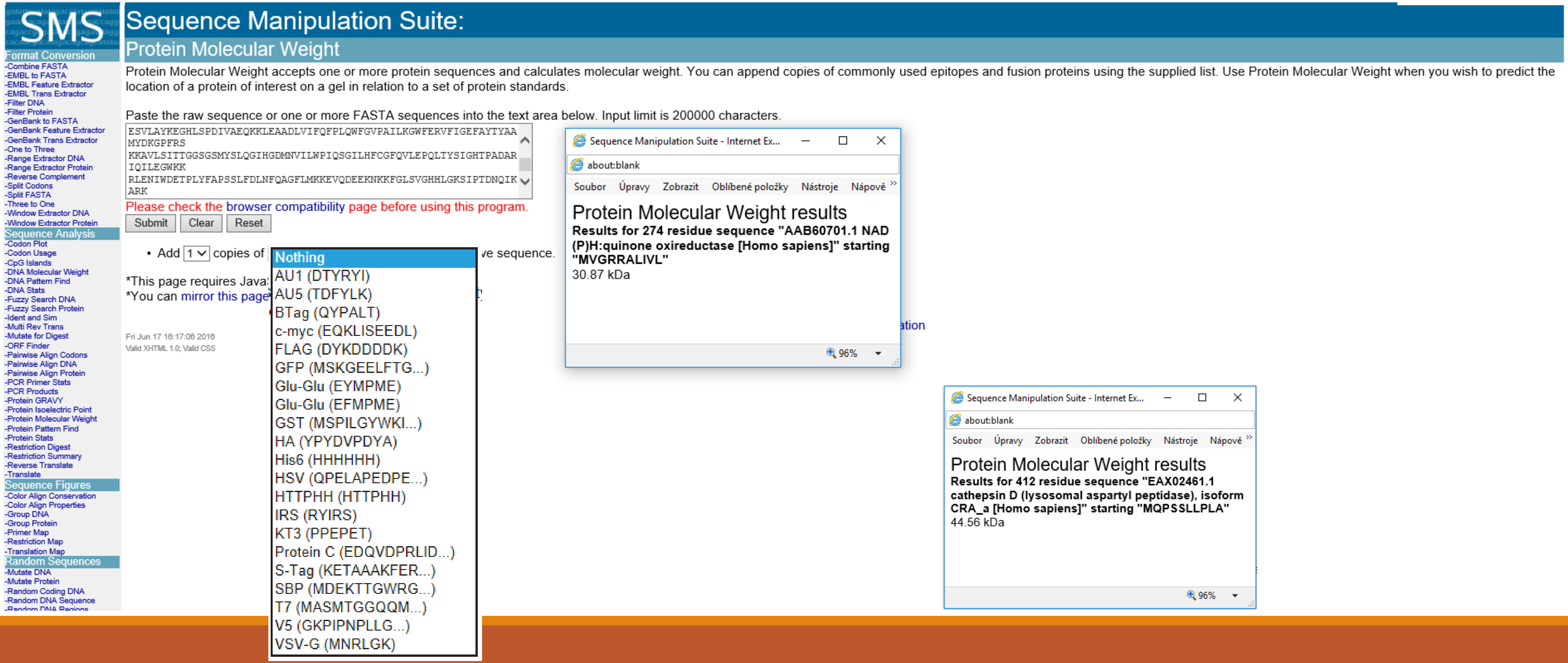
Results for 274 residue sequence "AAB60701.1 NAD(P)H:quinone oxireductase [Homo sapiens]" starting "MVGRRALIVL"

Pattern:	Times found:	Percentage:
A	25	9.12
B	0	0.00
C	1	0.36
D	12	4.38
E	17	6.20
F	17	6.20
G	21	7.66
H	7	2.55
I	19	6.93
K	24	8.76
L	25	9.12
M	7	2.55
N	9	3.28
P	13	4.74
Q	12	4.38
R	9	3.28
S	17	6.20
T	9	3.28

možná testová otázka:

a Identifikujte počet cysteinů

Analýza vlastností proteinu/sekvence



Sequence Manipulation Suite:

Protein Molecular Weight

Protein Molecular Weight accepts one or more protein sequences and calculates molecular weight. You can append copies of commonly used epitopes and fusion proteins using the supplied list. Use Protein Molecular Weight when you wish to predict the location of a protein of interest on a gel in relation to a set of protein standards.

Paste the raw sequence or one or more FASTA sequences into the text area below. Input limit is 200000 characters.

```
ESVLAYKEGHLSPDIVAEQKKLEAADLVIFQFPLQWFGVPAILKGWFERVFIGEFAYTYAA
MYDKGPFRS
KKAVLSITTGGSGSMYSLQGIHGDMNVILWPIQSGILHFCGFQVLEPQLTYSIGHTPADAR
IQILEGWKK
RLENIWDETPLYFAPSSLFDLNFQAGFLMKKEVQDEEKNKKFGLSVGHHLGKSIPTDNQIK
ARK
```

Please check the browser compatibility page before using this program.

- Add 1 copies of ve sequence.

*This page requires Java

*You can mirror this page

ation

Fri Jun 17 16:17:08 2016
Valid XHTML 1.0; Valid CSS

Protein Molecular Weight results

Results for 274 residue sequence "AAB60701.1 NAD (P)H:quinone oxireductase [Homo sapiens]" starting "MVGRRALIVL"

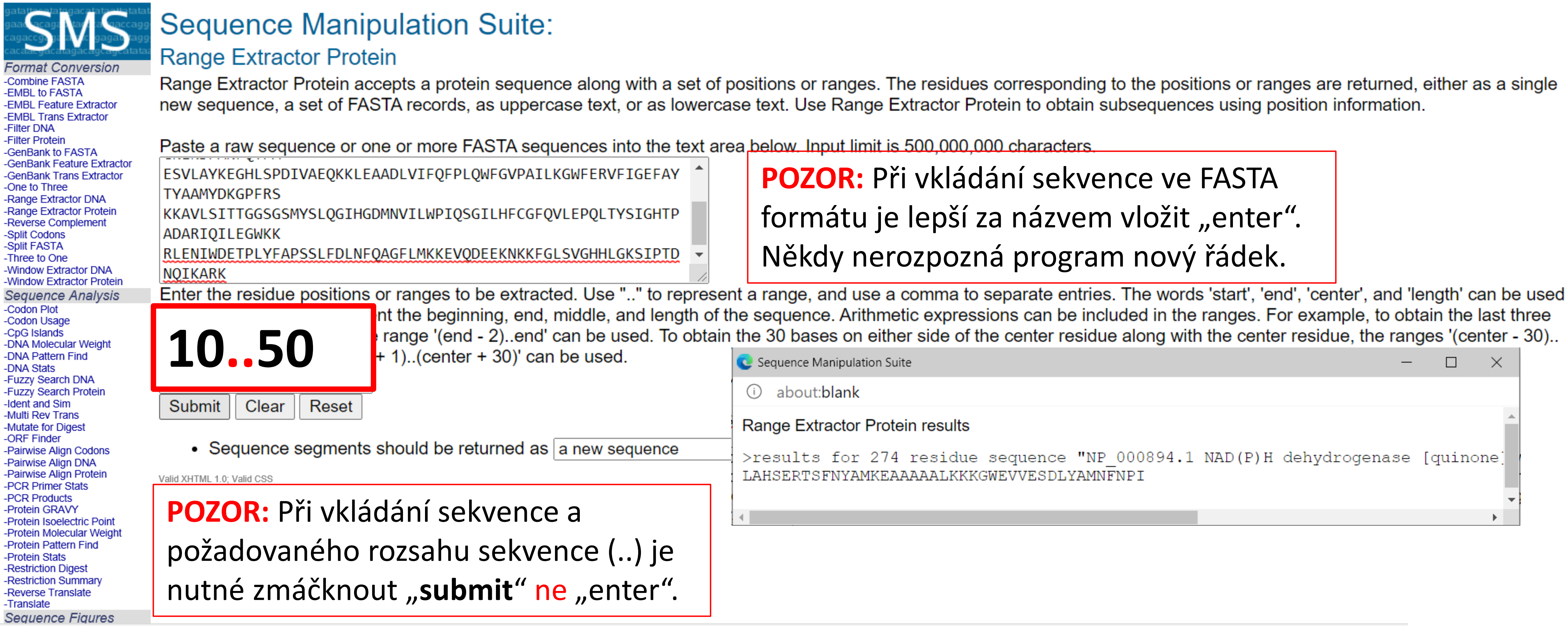
30.87 kDa

Protein Molecular Weight results

Results for 412 residue sequence "EAX02461.1 cathepsin D (lysosomal aspartyl peptidase), isoform CRA_a [Homo sapiens]" starting "MQPSSLLPLA"

44.56 kDa

Analýza vlastností proteinu/sekvence



Sequence Manipulation Suite:

Range Extractor Protein

Range Extractor Protein accepts a protein sequence along with a set of positions or ranges. The residues corresponding to the positions or ranges are returned, either as a single new sequence, a set of FASTA records, as uppercase text, or as lowercase text. Use Range Extractor Protein to obtain subsequences using position information.

Paste a raw sequence or one or more FASTA sequences into the text area below. Input limit is 500,000,000 characters.

```
ESVLAYKEGHLSPDIVAEQKKLEAADLVIFQFPLQWFGVPAILKGWFERVFIGEFAY
TYAAMYDKGPFRS
KKAVLSITTGGSGSMYSLQGIHGDMNVILWPIQSGILHFCGFQVLEPQLTYSIGHTP
ADARIQILEGWKK
RLENIWDETPLYFAPSSLFDLNFQAGFLMKKEVQDEEKNKKFGLSVGHHLGKSIPTD
NQIKARK
```

Enter the residue positions or ranges to be extracted. Use ".." to represent a range, and use a comma to separate entries. The words 'start', 'end', 'center', and 'length' can be used nt the beginning, end, middle, and length of the sequence. Arithmetic expressions can be included in the ranges. For example, to obtain the last three range '(end - 2)..end' can be used. To obtain the 30 bases on either side of the center residue along with the center residue, the ranges '(center - 30).. + 1)..(center + 30)' can be used.

10..50

- Sequence segments should be returned as a new sequence

Valid XHTML 1.0; Valid CSS

Range Extractor Protein results

```
>results for 274 residue sequence "NP_000894.1 NAD(P)H dehydrogenase [quinone]
LAHSERTSFNYAMKEAAAAALKKKGWEVVESDLYAMNFNPI
```

POZOR: Při vkládání sekvence ve FASTA formátu je lepší za názvem vložit „enter“. Někdy nerozpozná program nový řádek.

POZOR: Při vkládání sekvence a požadovaného rozsahu sekvence (..) je nutné zmáčknout „**submit**“ ne „enter“.

Vyzkoušejte si SMS suite

Analyzujte si „svou vlastní“ sekvenci

- kolik má váš protein prolinů?
- identifikujte 100. aminokyselinu
- jaká bude molekulová hmotnost vašeho proteinu s histidinovou kotvou (6xHis)?

„Proteinová bioinformatika“

Vyhledávání AMK sekvencí

Analýza vlastností sekvencí (aminokyselinové složení, molekulová hmotnost, isoelektrický bod…)

Štěpení proteasami

Analýza hydrofobních segmentů, transmembránových úseků

Predikce sekundární struktury

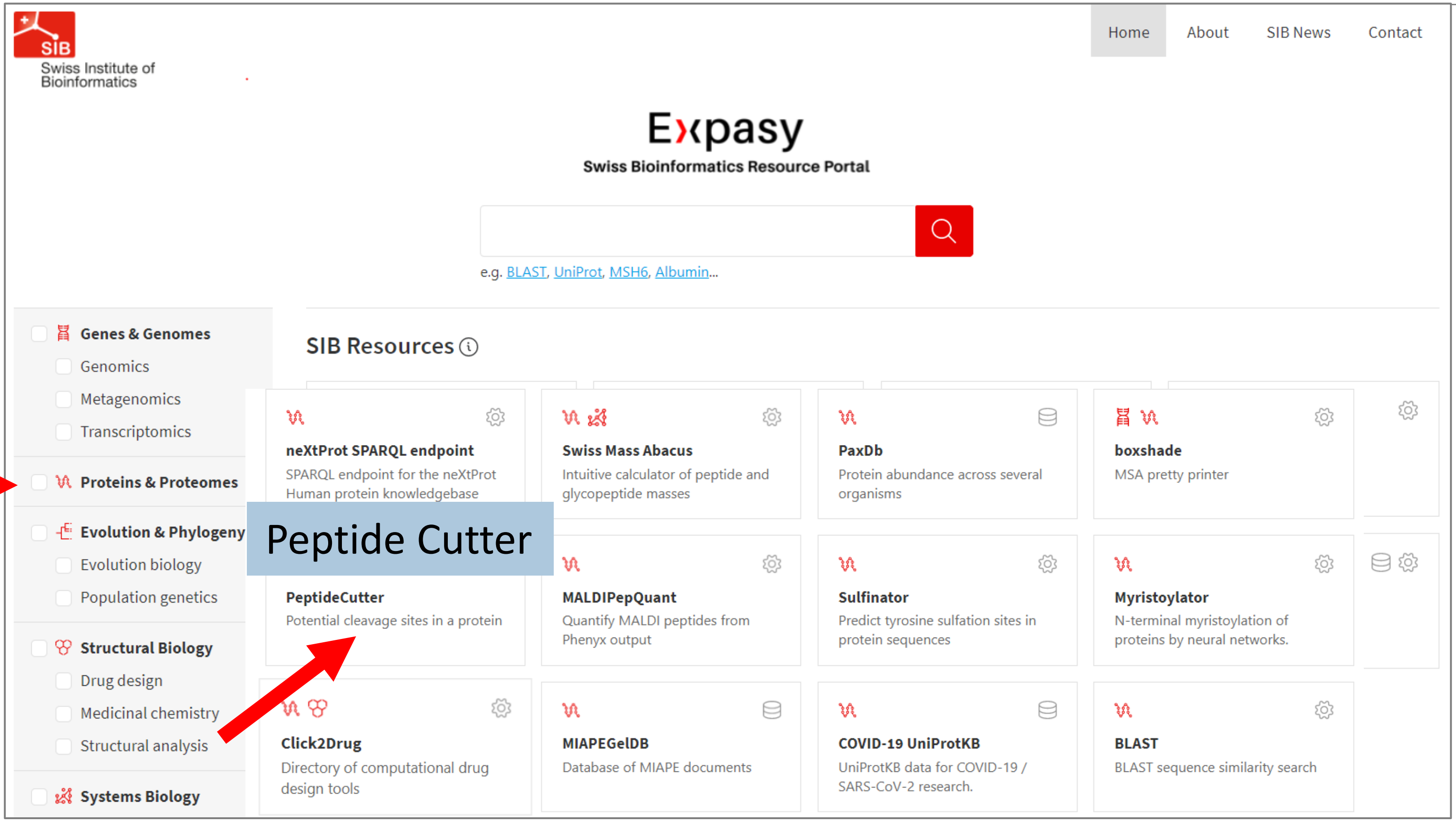
3D-struktura, vizualizace

Vyhledání a porovnání podobných sekvencí

Evoluční příbuznost sekvencí

…

Štěpení proteasami



SIB Swiss Institute of Bioinformatics

Expasy

Swiss Bioinformatics Resource Portal

e.g. BLAST, UniProt, MSH6, Albumin...

- ☐ **Genes & Genomes**
  - ☐ Genomics
  - ☐ Metagenomics
  - ☐ Transcriptomics
- ☐ **Proteins & Proteomes**
- ☐ **Evolution & Phylogeny**
  - ☐ Evolution biology
  - ☐ Population genetics
- ☐ **Structural Biology**
  - ☐ Drug design
  - ☐ Medicinal chemistry
  - ☐ Structural analysis
- ☐ **Systems Biology**

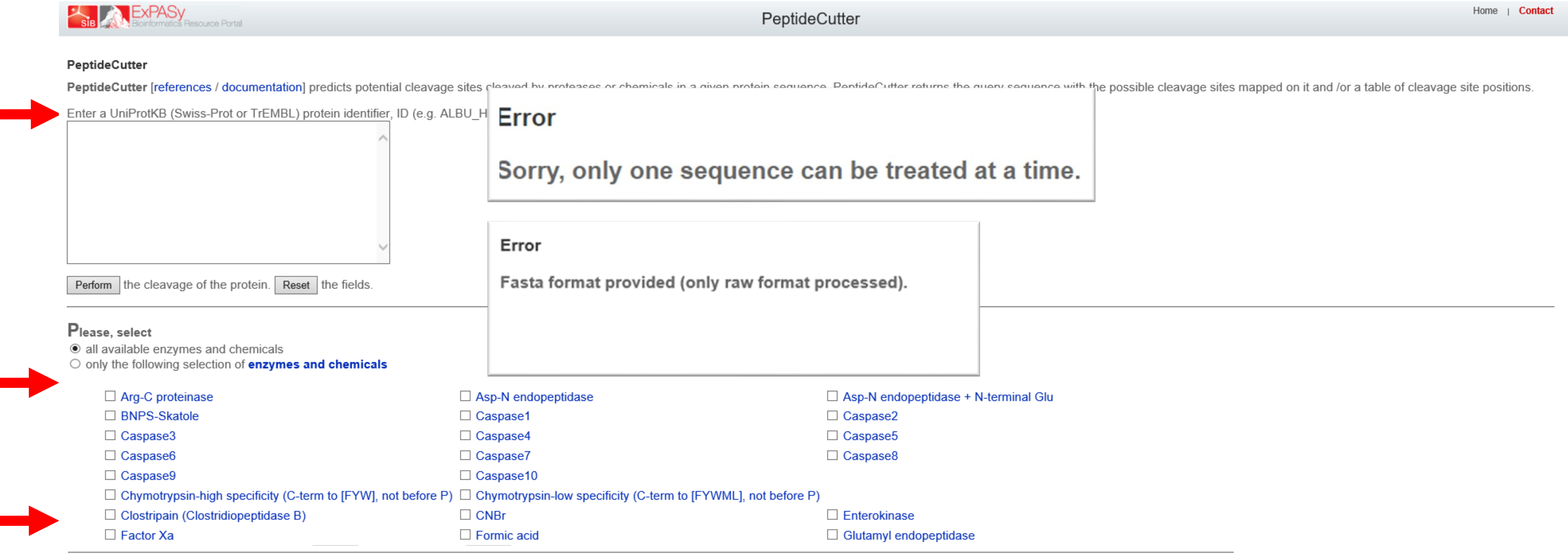
SIB Resources ⓘ

neXtProt SPARQL endpoint SPARQL endpoint for the neXtProt Human protein knowledgebase	**Swiss Mass Abacus** Intuitive calculator of peptide and glycopeptide masses	**PaxDb** Protein abundance across several organisms	**boxshade** MSA pretty printer
PeptideCutter Potential cleavage sites in a protein	**MALDIPepQuant** Quantify MALDI peptides from Phenyx output	**Sulfinator** Predict tyrosine sulfation sites in protein sequences	**Myristoylator** N-terminal myristoylation of proteins by neural networks.
Click2Drug Directory of computational drug design tools	**MIAPEGelDB** Database of MIAPE documents	**COVID-19 UniProtKB** UniProtKB data for COVID-19 / SARS-CoV-2 research.	**BLAST** BLAST sequence similarity search

Peptide Cutter

http://web.expasy.org/peptide_cutter/

Štěpení proteasami



PeptideCutter

PeptideCutter [references / documentation] predicts potential cleavage sites cleaved by proteases or chemicals in a given protein sequence. PeptideCutter returns the query sequence with the possible cleavage sites mapped on it and /or a table of cleavage site positions.

Enter a UniProtKB (Swiss-Prot or TrEMBL) protein identifier, ID (e.g. ALBU_H

Error

Sorry, only one sequence can be treated at a time.

Error

Fasta format provided (only raw format processed).

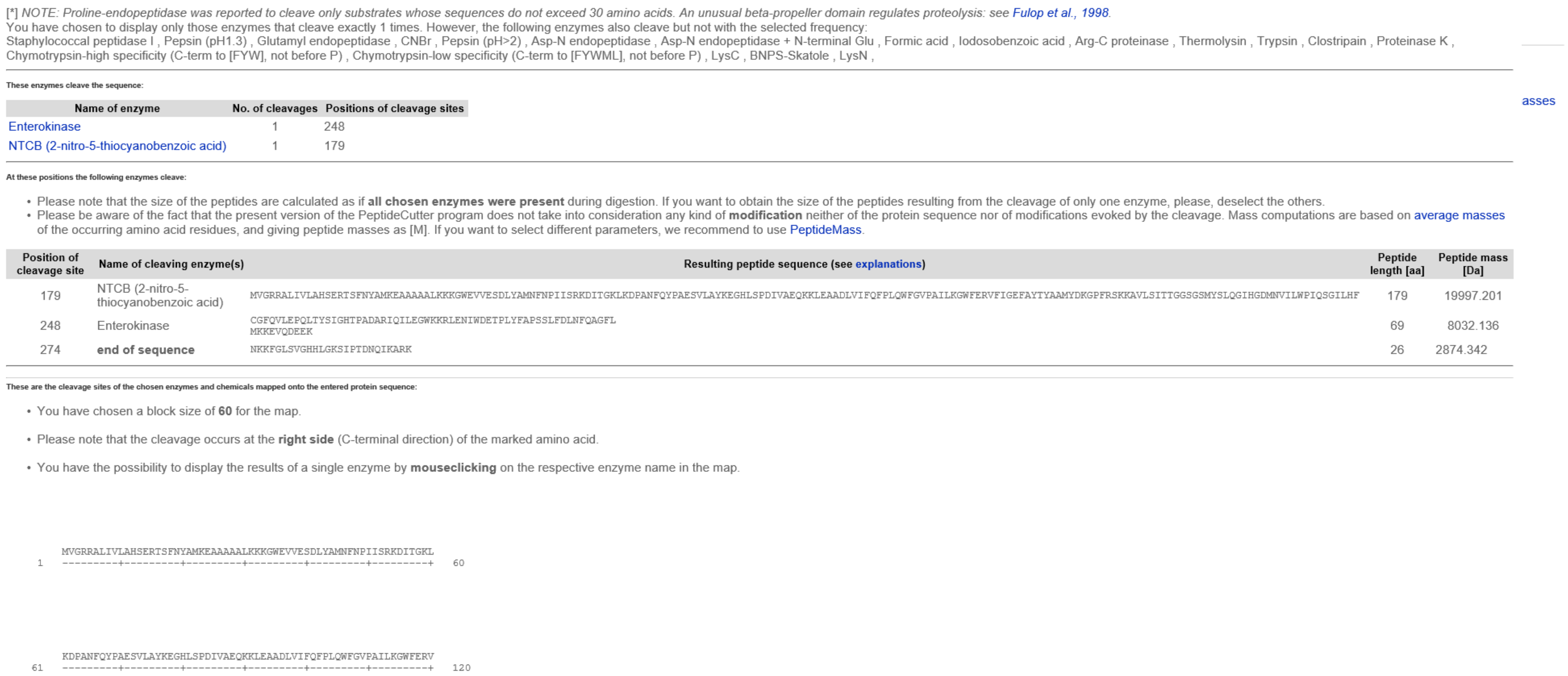
Perform the cleavage of the protein. Reset the fields.

Please, select

● all available enzymes and chemicals
○ only the following selection of **enzymes and chemicals**

☐ Arg-C proteinase
☐ Asp-N endopeptidase
☐ Asp-N endopeptidase + N-terminal Glu
☐ BNPS-Skatole
☐ Caspase1
☐ Caspase2
☐ Caspase3
☐ Caspase4
☐ Caspase5
☐ Caspase6
☐ Caspase7
☐ Caspase8
☐ Caspase9
☐ Caspase10
☐ Chymotrypsin-high specificity (C-term to [FYW], not before P)
☐ Chymotrypsin-low specificity (C-term to [FYWML], not before P)
☐ Clostripain (Clostridiopeptidase B)
☐ CNBr
☐ Enterokinase
☐ Factor Xa
☐ Formic acid
☐ Glutamyl endopeptidase

Štěpení proteasami-PeptideCutter



[*] *NOTE: Proline-endopeptidase was reported to cleave only substrates whose sequences do not exceed 30 amino acids. An unusual beta-propeller domain regulates proteolysis: see Fulop et al., 1998.*
You have chosen to display only those enzymes that cleave exactly 1 times. However, the following enzymes also cleave but not with the selected frequency:
Staphylococcal peptidase I , Pepsin (pH1.3) , Glutamyl endopeptidase , CNBr , Pepsin (pH>2) , Asp-N endopeptidase , Asp-N endopeptidase + N-terminal Glu , Formic acid , Iodosobenzoic acid , Arg-C proteinase , Thermolysin , Trypsin , Clostripain , Proteinase K , Chymotrypsin-high specificity (C-term to [FYW], not before P) , Chymotrypsin-low specificity (C-term to [FYWML], not before P) , LysC , BNPS-Skatole , LysN ,

These enzymes cleave the sequence:

Name of enzyme	No. of cleavages	Positions of cleavage sites
Enterokinase	1	248
NTCB (2-nitro-5-thiocyanobenzoic acid)	1	179

asses

At these positions the following enzymes cleave:

- Please note that the size of the peptides are calculated as if **all chosen enzymes were present** during digestion. If you want to obtain the size of the peptides resulting from the cleavage of only one enzyme, please, deselect the others.
- Please be aware of the fact that the present version of the PeptideCutter program does not take into consideration any kind of **modification** neither of the protein sequence nor of modifications evoked by the cleavage. Mass computations are based on average masses of the occurring amino acid residues, and giving peptide masses as [M]. If you want to select different parameters, we recommend to use PeptideMass.

Position of cleavage site	Name of cleaving enzyme(s)	Resulting peptide sequence (see explanations)	Peptide length [aa]	Peptide mass [Da]
179	NTCB (2-nitro-5-thiocyanobenzoic acid)	MVGRRALIVLAHSERTSFNYAMKEAAAAALKKKGWEVVESDLYAMNFNPIISRKDITGKLKDPANFQYPAESVLAYKEGHLSPDIVAEQKKLEAADLVIFQFPLQWFGVPAILKGWFERVFIGEFAYTYAAMYDKGPFRSKKAVLSITTGGSGSMYSLQGIHGDMNVILWPIQSGILHF	179	19997.201
248	Enterokinase	CGFQVLEPQLTYSIGHTPADARIQILEGWKKRLENIWDETPLYFAPSSLFDLNFQAGFL MKKEVQDEEK	69	8032.136
274	**end of sequence**	NKKFGLSVGHHLGKSIPTDNQIKARK	26	2874.342

These are the cleavage sites of the chosen enzymes and chemicals mapped onto the entered protein sequence:

- You have chosen a block size of **60** for the map.
- Please note that the cleavage occurs at the **right side** (C-terminal direction) of the marked amino acid.
- You have the possibility to display the results of a single enzyme by **mouseclicking** on the respective enzyme name in the map.

```
     MVGRRALIVLAHSERTSFNYAMKEAAAAALKKKGWEVVESDLYAMNFNPIISRKDITGKL
  1  ---------+---------+---------+---------+---------+---------+   60



     KDPANFQYPAESVLAYKEGHLSPDIVAEQKKLEAADLVIFQFPLQWFGVPAILKGWFERV
 61  ---------+---------+---------+---------+---------+---------+   120
```

Štěpení proteasami-PeptideCutter

PeptideCutter

PeptideCutter [references / documentation] predicts potential cleavage sites cleaved by proteases or chemicals in a given protein sequence. PeptideCutter returns the query sequence with the possible cleavage sites mapped on it and /or a table of cleavage site positions.

Enter a UniProtKB (Swiss-Prot or TrEMBL) protein identifier, ID (e.g. ALBU_HUMAN), or accession number, AC (e.g. P04406), **or** an amino acid sequence (e.g. 'SERVELAT'):

```
MAARRALIVLAHSEKTSFNYAMKEAAVEALKKRGWEVLESDLYAMNFNPIIS
RNDITGEL
KDSKNFQYPSESSLAYKEGRLSPDIVAEHKKLEAADLVIFQFPLQWFGVPAI
LKGWFERV
LVAGFAYTYAAMYDNGPFQNKKTLLSITTGGSGSMYSLQGVHGDMNVILWPI
QSGILRFC
GFQVLEPQLVYSIGHTPPDARMQILEGWKKRLETVWEETPLYFAPSSLFDLN
FQAGFLMK
KEVQEEQKKNKFGLSVGHHLGKSIPADNQIKARK
```

Hledání nejdelšího peptidu po štěpení

the cleavage of the protein. the fields.

Please, select

○ all available enzymes and chemicals
● only the following selection of **enzymes and chemicals**

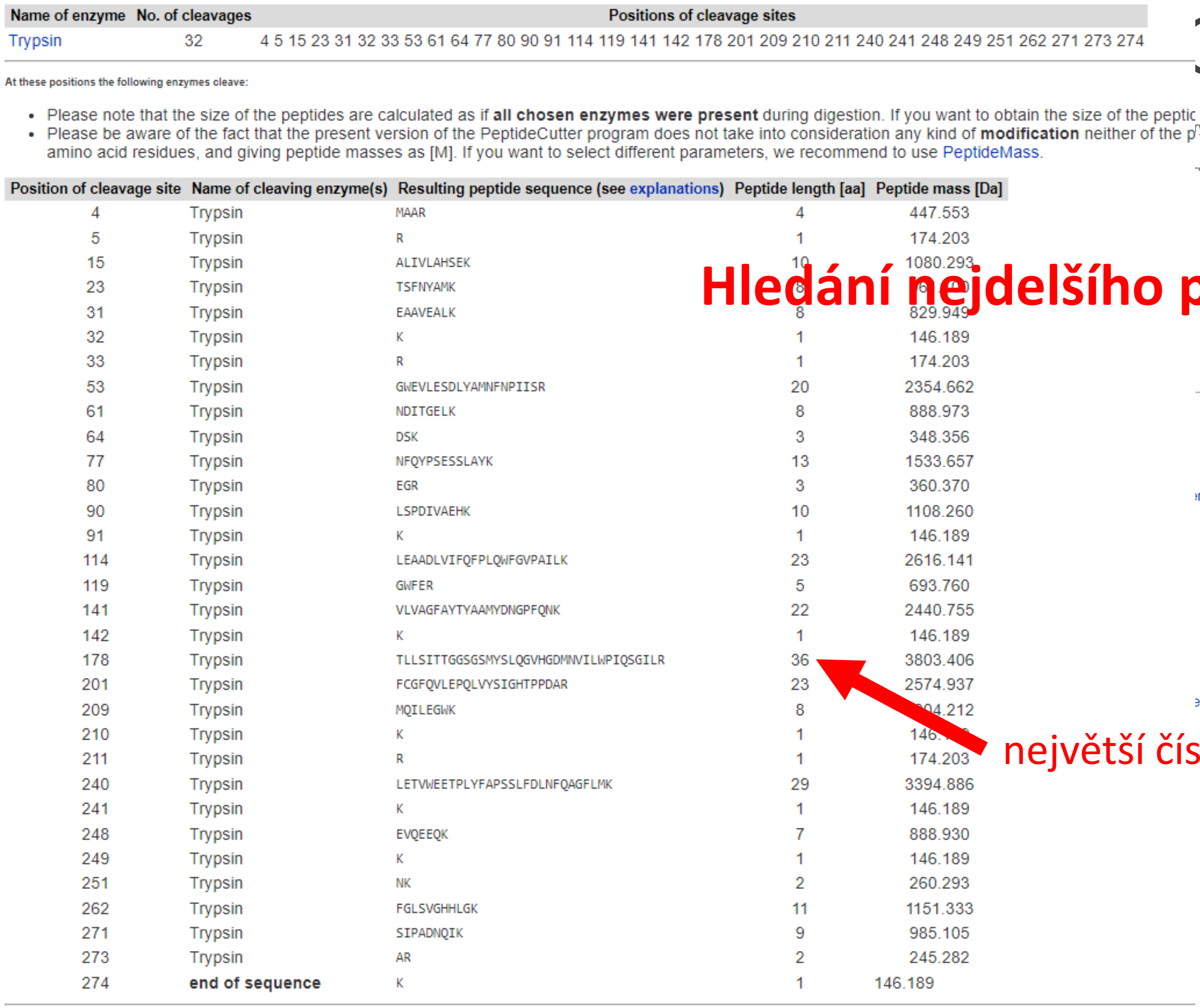
☐ Arg-C proteinase
☐ Asp-N endopeptidase
☐ Asp-N endopeptidase + N-terminal Glu
☐ BNPS-Skatole
☐ Caspase1
☐ Caspase2
☐ Caspase3
☐ Caspase4
☐ Caspase5
☐ Caspase6
☐ Caspase7
☐ Caspase8
☐ Caspase9
☐ Caspase10
☐ Chymotrypsin-high specificity (C-term to [FYW], not before P)
☐ Chymotrypsin-low specificity (C-term to [FYWML], not before P)
☐ Clostripain (Clostridiopeptidase B)
☐ CNBr
☐ Enterokinase
☐ Factor Xa
☐ Formic acid
☐ Glutamyl endopeptidase
☐ GranzymeB
☐ Hydroxylamine
☐ Iodosobenzoic acid
☐ LysC
☐ LysN
☐ NTCB (2-nitro-5-thiocyanobenzoic acid)
☐ Neutrophil elastase
☐ Pepsin (pH1.3)
☐ Pepsin (pH>2)
☐ Proline-endopeptidase
☐ Proteinase K
☐ Staphylococcal peptidase I
☐ Tobacco etch virus protease
☐ Thermolysin
☐ Thrombin
☑ Trypsin

Please indicate the way you would like the cleavage sites to be displayed

☑ Map of cleavage sites. Please select the number of amino acid within one block: 60
☑ Table of sites, sorted alphabetically by enzyme and chemical name
☑ Table of sites, sorted sequentially by amino acid number

ptideCutter

Hledání nejdelšího peptidu po štěpení



Name of enzyme	No. of cleavages	Positions of cleavage sites
Trypsin	32	4 5 15 23 31 32 33 53 61 64 77 80 90 91 114 119 141 142 178 201 209 210 211 240 241 248 249 251 262 271 273 274

At these positions the following enzymes cleave:

- Please note that the size of the peptides are calculated as if **all chosen enzymes were present** during digestion. If you want to obtain the size of the peptic
- Please be aware of the fact that the present version of the PeptideCutter program does not take into consideration any kind of **modification** neither of the p amino acid residues, and giving peptide masses as [M]. If you want to select different parameters, we recommend to use PeptideMass.

Position of cleavage site	Name of cleaving enzyme(s)	Resulting peptide sequence (see explanations)	Peptide length [aa]	Peptide mass [Da]
4	Trypsin	MAAR	4	447.553
5	Trypsin	R	1	174.203
15	Trypsin	ALIVLAHSEK	10	1080.293
23	Trypsin	TSFNYAMK	8	[illegible]
31	Trypsin	EAAVEALK	8	829.949
32	Trypsin	K	1	146.189
33	Trypsin	R	1	174.203
53	Trypsin	GWEVLESDLYAMNFNPIISR	20	2354.662
61	Trypsin	NDITGELK	8	888.973
64	Trypsin	DSK	3	348.356
77	Trypsin	NFQYPSESSLAYK	13	1533.657
80	Trypsin	EGR	3	360.370
90	Trypsin	LSPDIVAEHK	10	1108.260
91	Trypsin	K	1	146.189
114	Trypsin	LEAADLVIFQFPLQWFGVPAILK	23	2616.141
119	Trypsin	GWFER	5	693.760
141	Trypsin	VLVAGFAYTYAAMYDNGPFQNK	22	2440.755
142	Trypsin	K	1	146.189
178	Trypsin	TLLSITTGGSGSMYSLQGVHGDMNVILWPIQSGILR	36	3803.406
201	Trypsin	FCGFQVLEPQLVYSIGHTPPDAR	23	2574.937
209	Trypsin	MQILEGWK	8	[illegible]04.212
210	Trypsin	K	1	146.[illegible]
211	Trypsin	R	1	174.203
240	Trypsin	LETVWEETPLYFAPSSLFDLNFQAGFLMK	29	3394.886
241	Trypsin	K	1	146.189
248	Trypsin	EVQEEQK	7	888.930
249	Trypsin	K	1	146.189
251	Trypsin	NK	2	260.293
262	Trypsin	FGLSVGHHLGK	11	1151.333
271	Trypsin	SIPADNQIK	9	985.105
273	Trypsin	AR	2	245.282
274	**end of sequence**	K	1	146.189

největší číslo

ry sequence with the possible cleavage sites mapped on it and /or a table of cleavage site

"):

erminal Glu

enzoic acid)

Vyzkoušejte si PeptideCutter

Analyzujte si „svou vlastní“ sekvenci

- Štěpí nějaký enzym pouze jedenkrát? (-vyzkoušejte štěpení všemi (default nastavení)
  (nebo: Vyberte pouze enzym, který štěpí jedenkrát.)
- Jak dlouhý je nejdelší vzniklý peptid po štěpení trypsinem? (omezte štěpení pouze na Trypsin, zaškrtněte si tabulku s „počtem aminokyselin“)

„Proteinová bioinformatika“

Vyhledávání AMK sekvencí

Analýza vlastností sekvencí (aminokyselinové složení, molekulová hmotnost, isoelektrický bod…)

Štěpení proteasami

Analýza hydrofobních segmentů, transmembránových úseků

Predikce sekundární struktury

3D-struktura, vizualizace

Vyhledání a porovnání podobných sekvencí

Evoluční příbuznost sekvencí

…

shrnutí I - Rešeršní projekt: NQO1

Vyhledejte dostupné informace o NQO1, získejte co nejvíce literárních, sekvenčních, případně i strukturních a dalších údajů o tomto genu/proteinu.

Nalezněte podobné sekvence a porovnejte je na proteinové úrovni.

- **NAD(P)H:chinonoxidoreduktasa / reference**
- **Protein: NP_000894, P15559 / 274 AMK; 30,8 kDa; trypsin štěpí 33x (59AMK nejdelší peptid)…**

DÚ2: Vyhledávání proteinů

Pracujte s „vaším“ genem/proteinem/enzymem (př. NQO1)

1) Vyhledejte „referenční“ NCBI a UniProt přístupové kódy

2) Najděte a uložte si „FASTA“ formát sekvence a uložte si ji (jako text-ne jako výstřižek)

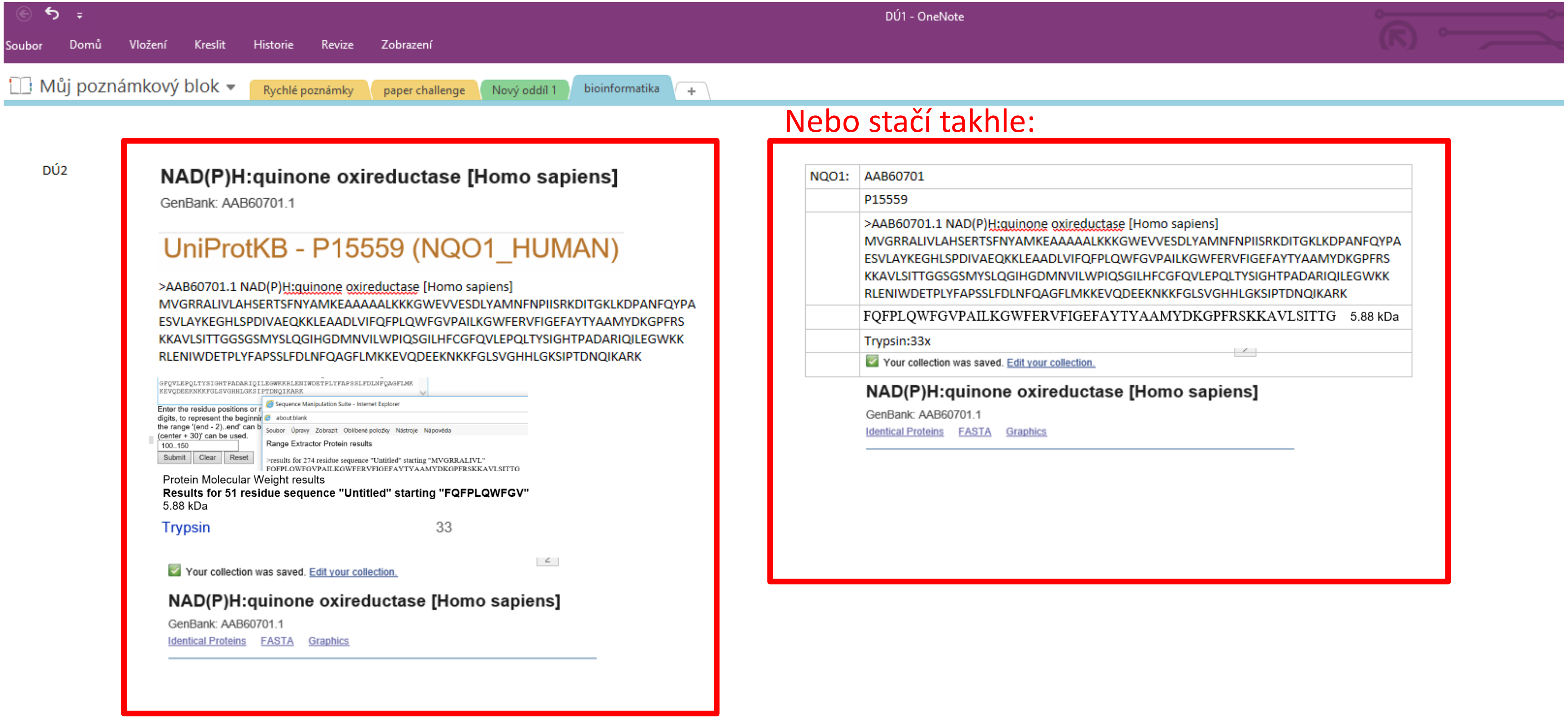
3) Vyberte část proteinu od **10. do 50.** aminokyseliny

4) Spočítejte molekulovou hmotnost tohoto úseku proteinu

5) Zjistěte kolikrát bude vaše **celá** sekvence štěpena trypsinem

- vše zpracujte do bloku OneNote, Word, Pdf

Literární rešerše: DÚ2-příklad řešení



Nebo stačí takhle:

DÚ2

NAD(P)H:quinone oxireductase [Homo sapiens]

GenBank: AAB60701.1

UniProtKB - P15559 (NQO1_HUMAN)

>AAB60701.1 NAD(P)H:quinone oxireductase [Homo sapiens]
MVGRRALIVLAHSERTSFNYAMKEAAAAALKKKGWEVVESDLYAMNFNPIISRKDITGKLKDPANFQYPA
ESVLAYKEGHLSPDIVAEQKKLEAADLVIFQFPLQWFGVPAILKGWFERVFIGEFAYTYAAMYDKGPFRS
KKAVLSITTGGSGSMYSLQGIHGDMNVILWPIQSGILHFCGFQVLEPQLTYSIGHTPADARIQILEGWKK
RLENIWDETPLYFAPSSLFDLNFQAGFLMKKEVQDEEKNKKFGLSVGHHLGKSIPTDNQIKARK

```
GFQVLEPQLTYSIGHTPADARIQILEGWKKRLENIWDETPLYFAPSSLFDLNFQAGFLMK
KEVQDEEKNKKFGLSVGHHLGKSIPTDNQIKARK
```

Enter the residue positions or r… digits, to represent the beginni… the range '(end - 2)..end' can b… (center + 30)' can be used.

100..150

Range Extractor Protein results

```
>results for 274 residue sequence "Untitled" starting "MVGRRALIVL"
FQFPLQWFGVPAILKGWFERVFIGEFAYTYAAMYDKGPFRSKKAVLSITTG
```

Protein Molecular Weight results
Results for 51 residue sequence "Untitled" starting "FQFPLQWFGV"
5.88 kDa

Trypsin 33

✅ Your collection was saved. Edit your collection.

NAD(P)H:quinone oxireductase [Homo sapiens]

GenBank: AAB60701.1

Identical Proteins FASTA Graphics

NQO1:	AAB60701
	P15559
	>AAB60701.1 NAD(P)H:quinone oxireductase [Homo sapiens] MVGRRALIVLAHSERTSFNYAMKEAAAAALKKKGWEVVESDLYAMNFNPIISRKDITGKLKDPANFQYPA ESVLAYKEGHLSPDIVAEQKKLEAADLVIFQFPLQWFGVPAILKGWFERVFIGEFAYTYAAMYDKGPFRS KKAVLSITTGGSGSMYSLQGIHGDMNVILWPIQSGILHFCGFQVLEPQLTYSIGHTPADARIQILEGWKK RLENIWDETPLYFAPSSLFDLNFQAGFLMKKEVQDEEKNKKFGLSVGHHLGKSIPTDNQIKARK
	FQFPLQWFGVPAILKGWFERVFIGEFAYTYAAMYDKGPFRSKKAVLSITTG 5.88 kDa
	Trypsin:33x
	✅ Your collection was saved. Edit your collection.

NAD(P)H:quinone oxireductase [Homo sapiens]

GenBank: AAB60701.1

Identical Proteins FASTA Graphics