

REŠITVE SEMINARJA S45

NALOGA 1:

1. Uporabimo bazo PubMed (<https://pubmed.ncbi.nlm.nih.gov/>). Gremo pod advanced iskanje in naredimo prikazano na sliki:

kliknemo add in search. Dobimo podatek, da je 2,751 takšnih znanstvenih člankov.



Slika 1: Število znanstvenih člankov.

2. Vtipkamo v search "Dihydrofolate reductase" AND cancer, "Dihydrofolate reductase" AND antibiotic, "Dihydrofolate reductase" AND malaria. Dobimo rezultate: 2,240 ; 815 ; 906. Pravilen odgovor je a).
3. V bazi UniProt (<https://www.uniprot.org/>) vpišemo podano kodo P00374 in gremo pod zavihek Structure in pod Source filtriramo samo PDB in vidimo, da protein pripada več PDB kod, saj je bil isti protein večkrat eksperimentalno določen v kompleksu z različnimi ligandi ali pri različnih pogojih. Zato ni ene same PDB kode, ampak več možnih struktur, npr. 1DHF, 1DRF in 1HFR...

NALOGA 2:

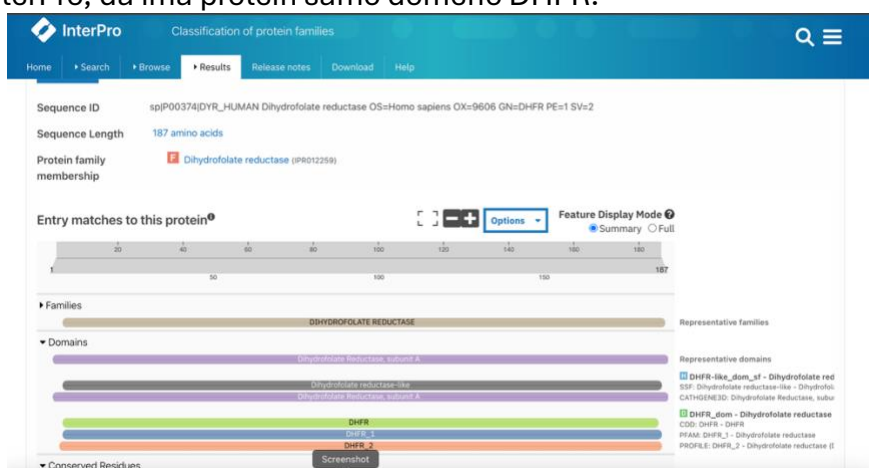
1. Uporabimo bazo UniProt vpišemo podano kodo P00374 in v zavihku Function vidimo rešitev, ki je sledeča: Dihidrofolat reduktaza (DHFR) je encim, ki katalizira redukcijo 7,8- dihidrofolata v 5,6,7,8 -tetrahidrofolat ob prisotnosti kofaktorja NADPH. Gre za ključni encim v metabolizmu folatov. Sodeluje pri de novo biosintezi timidilata v jedru in mitohondrijih. Katalizira esencialno reakcijo za de novo sintezo glicina in purinov ter za sintezo prekursorjev DNA.
2. Uporabimo bazo UniProt in v zavihku Function najdemo podatek, da veže obe mRNA.
3. Uporabimo bazo UniProt in v zavihku Function najdemo podatek, da je kofaktor NADPH, ki ima 5 vezavnih mest.

4. V zavihku PTM/Processing filtriramo v Type po Modified.Vidimo, da je na 55. aminokislinskem mestu lizin modificiran. Na tem mestu sta dve posttranslacijski modifikaciji: **ubikvitinacija (ubiquitinated lysine) in sumoilacija (sumoylated lysine)**. In če kliknemo mali klicaj (!) zraven statusa lahko se zasedimo, da je ubikvitinacija na mestu 55 je bolj potrjena, saj ima zlati status (gold). To nakazuje, da je bila ta modifikacija opažena v več podatkovnih zbirkah z visoko stopnjo zaupanja (nizka stopnja napačne lokalizacije, ≤ 1 % FLR). Sumoilacija na istem mestu je manj zanesljivo potrjena, ker ima bronasti status (bronz). To pomeni, da je bila opažena v manj podatkih oziroma z nižjo stopnjo zaupanja (do 5 % FLR).

±	Modified residue (large scale data)	ID	POSITION(S)	All	DESCRIPTION
+	Modified residue (large scale data)	4	PKIDE		Phosphoserine Combined Sources
+	Modified residue (large scale data)	19	PTMeXchange		Sumoylated lysine Silver Combined Sources
+	Modified residue (large scale data)	47	PTMeXchange		Sumoylated lysine Silver Combined Sources
+	Modified residue (large scale data)	47	PTMeXchange		Ubiquitinated lysine Gold Combined Sources
+	Modified residue (large scale data)	55	PTMeXchange		Sumoylated lysine Bronze Combined Sources
+	Modified residue (large scale data)	55	PTMeXchange		Ubiquitinated lysine Gold Combined Sources
+	Modified residue (large scale data)	56	PTMeXchange		Sumoylated lysine Silver Combined Sources
+	Modified residue (large scale data)	56	PTMeXchange		Ubiquitinated lysine Gold Combined Sources
+	Modified residue (large scale data)	64	PTMeXchange		Sumoylated lysine Gold Combined Sources

Slika 2: Posttranslacijske modifikacije.

5. V bazi UniProt v zavihku *Family & Domains* je navedena ena domena DHFR, ki obsega območje od 4. do 165. aminokislinskega ostanka. Pravtako lahko vidimo v bazi InterPro, da ima protein samo domeno DHFR.



Slika 3: Prikaz podatkov v InterPro.

6. V bazi UniProt poiščemo človeški DHFR (koda P00374) in mišji DHFR (mus musculus) tega najdemo pod kodo P00375. Skopiramo sekvenco obeh posebej v fasta formatu v program ProtParam dostopen preko spletnega brskalnika (<https://web.expasy.org/protparam/>).

Za človeka dobimo sledeče:

Expasy ProtParam

Home Documentation Reference Contact

ProtParam - Results

User-provided sequence:

```

10      20      30      40      50      60
MVGSLNCIVA VSQNMGIGKN GDLWPPLRN EFRYFQMTT TSSVEGKQNL VIMGKKTWFS
70      80      90      100     110     120
IPEKNRPLKG RINLVLSREL KEPPQGAHFL SRSLDDALKL TEQPELANKV DMVWIVGGSS
130     140     150     160     170     180
VYKEAMNHPG HCLKFVTRIM QDFESDTFFP EIDLEKYKLL PEYPGVLSDV QEEKGIKYKF
EYVEKND

```

[Documentation / Reference]

Number of amino acids: 187
Theoretical pI: 6.85
Molecular weight: 21452.73

Amino acid composition:

Ala (A)	5	2.7%
Arg (R)	8	4.3%
Asn (N)	10	5.3%

Slika 4: Zadetek za amiokislinsko zaporedje človeškega DHFR.

Za miš dobimo sledeče:

Expasy ProtParam

Home Documentation Reference Contact

ProtParam - Results

User-provided sequence:

```

10      20      30      40      50      60
MVRPLNCIVA VSQNMGIGKN GDLWPPLRN EFRYFQMTT TSSVEGKQNL VIMGRKTWFS
70      80      90      100     110     120
IPEKNRPLKD RINIVLSREL KEPPRGAHFL AKSLDDALRL IEQPELASKV DMVWIVGGSS
130     140     150     160     170     180
VYQEAQNQPG HLRLFVTRIM QEFESDTFFP EIDLGKYKLL PEYPGVLSEV QEEKGIKYKF
EYVEKND

```

[Documentation / Reference]

Number of amino acids: 187
Theoretical pI: 8.56
Molecular weight: 21606.05

Amino acid composition:

Ala (A)	6	3.2%
Arg (R)	11	5.9%

Slika 5: Zadetek za aminokislinsko zaporedje mišjega DHFR.

Razlika med tvojima pI vrednostma je opazna (človek ~6,85 vs. miš ~8,56) in vedno razlik v aminokislinskem zaporedju, predvsem na površini proteina. In pri mišjem DHFR vidimo nižje, da je total number of positively charged residues (Arg +Lys) **27**. Total number of negatively charged residues (Asp + Glu) **25**. Torej zaradi več pozitivno nabitih

aminokislinskih ostankov se lahko pl pomakne višje. Pri človeškem pa zasledimo, da sta obe vrednosti enaki in sicer **25**.

Naloga 4:

Tukaj lahko v UniProtu poiščemo mišji DHFR, gremo pod zavihek Sequence in pod Sequence databases poiščemo link do GenBanka z želenim zaporedjem mRNA.

The screenshot shows the UniProt interface with the 'Sequence databases' section expanded. It displays a table with columns for 'NUCLEOTIDE SEQUENCE', 'PROTEIN SEQUENCE', 'MOLECULE TYPE', and 'STATUS'. The table lists several entries, including EMBL, GenBank, and DDBJ links for both nucleotide and protein sequences. A green box highlights the GenBank link for the mRNA sequence.

Slika 6: Prikaz linkov do GenBanka.

Ali pa gremo direktno v GenBank in napišemo v iskalno vrstico Dihydrofolate reductase in označimo še v filtrih ob strani organizem, mRNA. In spodaj vidimo, da se prevede v protein zaporedje od 56 mesta do 619 v sledeče aminokislinsko zaporedje spodaj.

```

CDS
-----
56..619
/gene="dhfr"
/codon_start=1
/product="dihydrofolate reductase"
/protein_id="AAA37523.1"
/translation="MVRPLNCIVAVSQNMGIGKNGDLPWPPLRNEFKYFQRMTTSSV
EGKQNLVIMGRKTWFSIPEKNRPLKDRINIVLSRELKEPPRGAHFLAKSLDDALRLIE
QPELASKVDMVWIVGGSSVYQEAMNQPHLRLFVTRIMQEFESDTFFPEIDLKGYKLL
PEYPGVLSEVQEEKGIKYKFEVYEKKD"
-----

```

Slika 7: Rezultat iz GenBanka.

Naloga 4:

V zgornjih nalogah si že poiskal mišji in človeški DHFR v UniProtu upam, da si si pustil še odprte njune sekvence ;)

Uporabimo Blast (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). In sicer blastp kamor prilepimo aminokislinsko zaporedje človeškega DHFR, database non-redundant protein sequences, organism: mus musculus in pa kasneje damo še Plasmodium falciparum. Dobimo sledeče rezultate:

Other reports [Distance tree of results](#) [Multiple alignment](#) [MSA viewer](#) [Filter](#) [Reset](#)

Compare these results against the new Clusted nr database [BLAST](#)

Descriptions [Graphic Summary](#) [Alignments](#) [Taxonomy](#)

Sequences producing significant alignments [Download](#) [Select columns](#) [Show](#) 100 [?](#)

select all 11 sequences selected [GenPept](#) [Graphics](#) [Distance tree of results](#) [Multiple alignment](#) [MSA Viewer](#)

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> dihydrofolate reductase [Mus musculus]	Mus musculus	353	353	100%	8e-126	89.84%	187	NP_034179.1
<input checked="" type="checkbox"/> unnamed protein product [Mus musculus]	Mus musculus	352	352	100%	5e-125	89.30%	187	CAA24112.1
<input checked="" type="checkbox"/> Chain A, Dihydrofolate reductase [Mus musculus]	Mus musculus	351	351	99%	6e-125	89.78%	186	2FZJ_A
<input checked="" type="checkbox"/> Chain A, Dihydrofolate reductase [Mus musculus]	Mus musculus	350	350	99%	2e-124	89.25%	186	3K45_A
<input checked="" type="checkbox"/> dihydrofolate reductase [Mus musculus]	Mus musculus	349	349	100%	4e-124	88.77%	187	CAA39544.1
<input checked="" type="checkbox"/> Chain A, Dihydrofolate reductase [Mus musculus]	Mus musculus	349	349	99%	5e-124	89.25%	186	1U70_A
<input checked="" type="checkbox"/> dihydrofolate reductase, partial [Mus musculus]	Mus musculus	145	145	42%	4e-45	87.18%	78	AAA37525.1
<input checked="" type="checkbox"/> dihydrofolate reductase, partial [Mus musculus]	Mus musculus	61.2	61.2	15%	1e-12	92.86%	28	AAA37637.1
<input checked="" type="checkbox"/> unnamed protein product, partial [Mus musculus]	Mus musculus	49.3	49.3	13%	5e-08	92.00%	25	CAA24111.1
<input checked="" type="checkbox"/> dihydrofolate reductase, partial [Mus musculus]	Mus musculus	49.3	49.3	13%	5e-08	91.67%	24	AAA37524.1
<input checked="" type="checkbox"/> unnamed protein product, partial [Mus musculus]	Mus musculus	49.3	49.3	13%	5e-08	91.67%	28	CAB43539.2

LibreOffice

Slika 8: Podobnost s mišjo.

Sequences producing significant alignments [Download](#) [Select columns](#) [Show](#) 100 [?](#)

select all 100 sequences selected [GenPept](#) [Graphics](#) [Distance tree of results](#) [Multiple alignment](#) [MSA Viewer](#)

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> bifunctional dihydrofolate reductase-thymidylate synthase, partial [Plasmodium falciparum]	Plasmodium falciparum	89.0	89.0	92%	2e-21	30.20%	222	YDC15383.1
<input checked="" type="checkbox"/> dihydrofolate reductase-thymidylate synthase, partial [Plasmodium falciparum]	Plasmodium falciparum	89.0	89.0	92%	2e-21	30.20%	220	AAG10113.1
<input checked="" type="checkbox"/> dihydrofolate reductase-thymidylate synthase, partial [Plasmodium falciparum]	Plasmodium falciparum	89.0	89.0	92%	2e-21	30.20%	220	AAG10129.1
<input checked="" type="checkbox"/> bifunctional dihydrofolate reductase-thymidylate synthase, partial [Plasmodium falciparum]	Plasmodium falciparum	89.0	89.0	92%	2e-21	30.20%	228	WHA04249.1
<input checked="" type="checkbox"/> bifunctional dihydrofolate reductase-thymidylate synthase, partial [Plasmodium falciparum]	Plasmodium falciparum	89.0	89.0	92%	3e-21	30.20%	235	WHA04250.1
<input checked="" type="checkbox"/> dihydrofolate reductase, partial [Plasmodium falciparum]	Plasmodium falciparum	88.2	88.2	91%	3e-21	30.50%	200	WNV48385.1
<input checked="" type="checkbox"/> bifunctional dihydrofolate reductase-thymidylate synthase, partial [Plasmodium falciparum]	Plasmodium falciparum	88.6	88.6	92%	3e-21	30.20%	220	WHA04252.1
<input checked="" type="checkbox"/> bifunctional dihydrofolate reductase-thymidylate synthase, partial [Plasmodium falciparum]	Plasmodium falciparum	88.2	88.2	92%	5e-21	30.20%	228	WHA04243.1
<input checked="" type="checkbox"/> bifunctional dihydrofolate reductase-thymidylate synthase, partial [Plasmodium falciparum]	Plasmodium falciparum	87.8	87.8	92%	7e-21	30.20%	232	WHA04244.1
<input checked="" type="checkbox"/> dihydrofolate reductase, partial [Plasmodium falciparum]	Plasmodium falciparum	86.7	86.7	90%	1e-20	30.30%	211	QJX15324.1
<input checked="" type="checkbox"/> dihydrofolate reductase, partial [Plasmodium falciparum]	Plasmodium falciparum	88.2	88.2	92%	1e-20	30.20%	296	QOP39095.1
<input checked="" type="checkbox"/> dihydrofolate reductase-thymidylate synthase, partial [Plasmodium falciparum]	Plasmodium falciparum	86.7	86.7	92%	1e-20	29.70%	220	AAG10126.1
<input checked="" type="checkbox"/> dihydrofolate reductase-thymidylate synthase, partial [Plasmodium falciparum]	Plasmodium falciparum	86.7	86.7	92%	1e-20	30.20%	220	AAG10109.1
<input checked="" type="checkbox"/> Chain A, Bifunctional dihydrofolate reductase-thymidylate synthase [Plasmodium falciparum]	Plasmodium falciparum	87.4	87.4	92%	2e-20	30.20%	280	1J3K_A
<input checked="" type="checkbox"/> dihydrofolate reductase-thymidylate synthase, partial [Plasmodium falciparum]	Plasmodium falciparum	86.3	86.3	92%	2e-20	29.70%	220	AAG10106.1
<input checked="" type="checkbox"/> dihydrofolate reductase, partial [Plasmodium falciparum]	Plasmodium falciparum	86.3	86.3	91%	2e-20	30.00%	223	ANG56392.1
<input checked="" type="checkbox"/> dihydrofolate reductase-thymidylate synthase, partial [Plasmodium falciparum]	Plasmodium falciparum	88.2	88.2	92%	2e-20	30.20%	347	ADC94836.1
<input checked="" type="checkbox"/> dihydrofolate reductase, partial [Plasmodium falciparum]	Plasmodium falciparum	85.9	85.9	90%	2e-20	29.80%	211	OLX15326.1

Screenshot

Slika 9: Podobnost z plazmodijem.

Človeški in mišji DHFR sta homologna proteina, kar je razvidno iz visoke stopnje identičnosti (~89 %), velikega pokritja zaporedja (100 %) in zelo nizke E-vrednosti (8e-126). To kaže na skupen evolucijski izvor in ohranjeno funkcijo encima.

DHFR iz plazmodija pa je ni tako zelo podoben človeškemu (**Percent identity: ~30%, Query cover: ~90%, E-value: ~2e-21**).

Razlike v zaporedju imajo pomemben vpliv na delovanje zdravil. Antimalarična zdravila, ki ciljajo DHFR, lahko selektivno zavirajo Plasmodium falciparum DHFR zaradi njegovih strukturnih razlik, medtem ko imajo manjši učinek na človeški in mišji DHFR, ki sta si med seboj zelo podobna. Zato ta zdravila ne delujejo enako na vse organizme, temveč imajo večjo afiniteto do parazitskega DHFR, vendar lahko v manjši meri vplivajo tudi na človeški DHFR.