

S1-09-Morski_adhesivi-resitev

Datum izdelave: 2026-05-07

Koda seminarja: S1

Naloga 1

a) Zaporedje vstaviš v blastn (nič ne spreminjaš prejšnjih nastavitev), prvi rezultat ima 100% identičnost (accession code DQ000154.1, *Perna viridis* foot protein 1 variant 1 (fp-1) mRNA, complete cds). Za izračun razmerij dolžin v števec vstaviš dolžino vstavljenega zaporedja (angl. Query Length: 1204), v imenovalec pa dolžino celotnega kodirajočega zaporedja ($1700\text{bp} - 14 = 1686$), rezultat je 0,71.

b) Pod uniprot vpišeš ime proteina, in filter organism(*Perna viridis*) -> PTM/processing -> 4-hidroksi prolin, 7-hidroksitriptofan, kovalentno povezan preko C atoma z manozo 7-hidroksitriptofan, 3',4'-dihidroksifenilalanil. Za podatke o izoelektrični točki je možno pridobiti z vnosom ak zaporedja v ProtParam. Ker je izoelektrična točka 11,20, bo imel protein pozitiven naboj, torej ga lahko izoliramo z kationsko izmenjevalno kromatografijo, pH območje je tudi primerno, saj se razlikuje več kot za eno enoto pH od I_p , torej se protein ne bo oborjeval, območje tudi ne pripada nobenemu od ekstremnih pH območij, ki bi pozročala oborjevanje proteina.

c) Del podatkov je možno najti v UniProt -> Family & Domains. Če klikneš pod link od Interpro, bo anotirana le kolagenska trojna alfa vijačnica (ak 459-510), zanj značilno ponavljanje G-X-Y, pri čemer je na drugem in tretjem mestu lahko katerakoli aminokislina, pogosto pa najdemo prolin oz. hidroksiprolin, ima 20 ponovitev. Do podrobnosti o kolagenski domeni je možno priti s klikom na Collagen - Collagen triple helix repeat, ki je nahaja pod naslovom representative domains (seveda še vedno znotraj InterPro).

Podatke o domenah X 10 AA A-P-P-P-A-W-T-A-W-K (ak 41-410) in 27 X 10 AA A-T-P-K-P-W-T-A-W-K (ak 61-440) dobiš s klikom na PubMed datoteko, ki je navedena zraven ob npr 7-hidroksitriptofan. Članek se imenuje lycosylated hydroxytryptophan in a mussel adhesive protein from *Perna viridis*. J Biol Chem. 2009 Aug 28;284(35):23344-52. doi: 10.1074/jbc.M109.022517. Za slednji dve domeni je značilnih 42 decapeptidnih ponovitev z paralelnimi beta trakovi in ponavljajočim se

zaporedjem ATPKPW1TAW2K in APPPAW1TAW2K. Modificirani aminokislinski ostanki bi lahko vpljivali na Ip, saj deprotonirane hidroksidne skupine pripomorejo k dodatno negativno nabitem proteinu.

Glede na to, da niti domene X 10 AA A-P-P-P-A-W-T-A-W-K niti doene 27 X 10 AA A-T-P-K-P-W-T-A-W-K ni bilo možno najti v Interpro, ta pot iskanja odpade. Sami imeni teh domen oz dolgim ponovotvam so preveč obskurne da bi karkoli našli, ko njihovo ime v brskalnih dodamo. Iskanje v navedenem članku je še najhitrejše in sigurna pot do zanesljivih podatkov.

d) Za iskanje ostalih proteinov v ime proteina vpišeš samo "foot protein", kot organizem pa samo "mussel", ročno anotirana proteina bosta še dva :Adhesive plaque matrix protein (ime genov FP1 in PF2). Podatki o funkcijah modificiranih proteinov so v discussion od članka z poduprašanja c), pod oddelkom Razprava.

Vloge PTM-jev od ajdenih proteinov:

Protein z genom PF1- PTM:(3R,4S)-3,4-dihidroksiprolin, 3',4'-dihidroksifenilalanil(DOPA), 4-hidroksiprolin

protein z genom PF2- PTM: 3',4'-dihidroksifenilalanil(DOPA), disulfidni mostički med cisteini.

pri avtomatsko anotiranih proteinih bodo med PTM navedee samo signalne sekvnce, ni jih potrebno navesti.

DOPA je bila eksperimentalno določena kot PTM ki pripomore k adhesivni funkciji proteina, modifikacije na Foot protein 1 pa so bile tudi določene, da prispevajo k adhesijski moči(rezultati članka, ki je naveden pri večini PTM-jev od tega proteina) disulfidni mostički pripomorejo k trdnosti samega proteina.

Naloga 2

a) Pod UniProt structure klikneš prvi model (Alphafold DB, piše Foldseek zraven ikonice za naložit datoteko) in si ga naložiš. Nato odpreš samo strukturo v ChimeraX (znotraj aplikacije odpreš naloženo datoteko z ikonico Open), izbrišeš dele structure ki niso navedeni v navodilih (sel @A: 1-15, del sel, sel @A: 433-561, del sel) in shraniš datoteko kot pdb datoteko. Osrednjo strukturo sedaj vstaviš v PDBeFold, kot vir (Source) daš Coordinate file in naložiš osrednjo domeno, ki si jo shraniš kot pdb kodo. Ne spreminjaš nebenih nastavitvev.

Rezultati iskanja na PDBeFold naj bi dali samo en rezultat: POLYGALACTURONASE FROM ERWINIA CAROTOVORA SSP. CAROTOVORA (PDB ID: 1bhe), določena je bila po eksperimentalni metodi XRD, ločljivost strukture je 1.9Å.

b) V UniProt-u sedaj filtriraš zadetke za samo organizem *Pectobacterium carotovorum* in ime proteina Pectate lyase, pod sSimilar proteins 50% katere koli pektin liaze najdeš vseh 5 podobnih proteinov (Pectate lyase 1,2,3,B,C), ko narediš multible align in izbereš aktivno mesto enega od proteinov, vidiš da imajo vsi enako aktivno mesto, R 239. V signalni sekvenci se bo razlikovala le pektat liaza C, ki ima krajšo signalno sekvenco.

c) Rezultati poravnave z blastp in blastx nakazujejo navajajo protein z visoko identičnostjo iz samo rodu enterobacteria, torej ima verjetno hoologe v istem rodu kamor sama pripada, kar je smiselno. Če se oseba odločila za nblast (genebank id: D00217.1), bodo sicer rezultati samo od iste vrste, potemtakem predvidevam da bo posegla po drugi vrsti lokale blast poravnave.