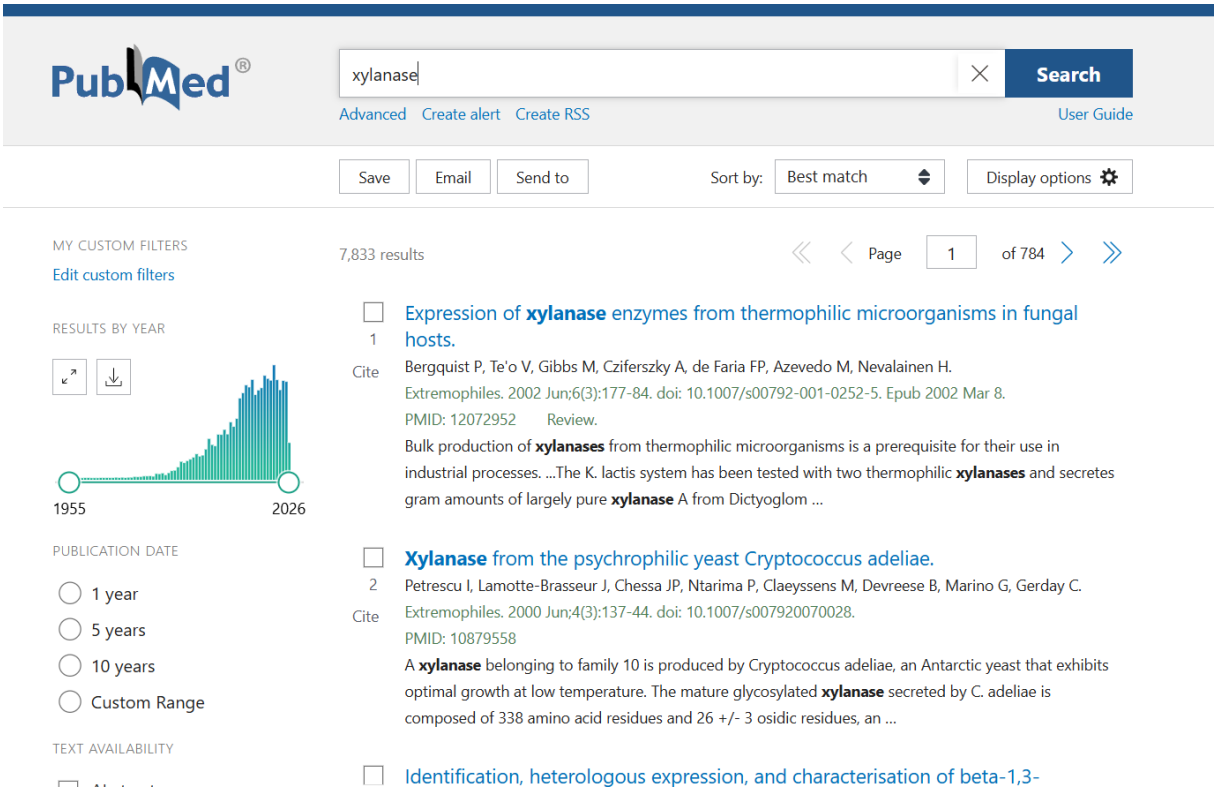
Rešitve nalog

1.

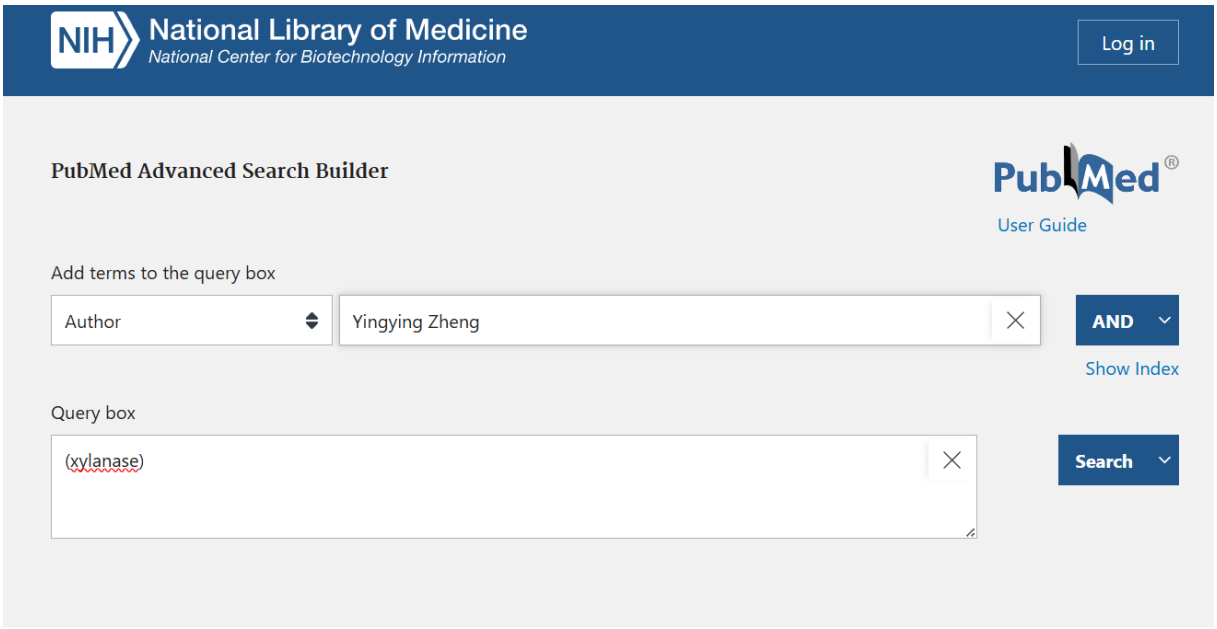
Za iskanje člankov uporabimo pubmed, kjer vstavimo xylanase v iskalno vrstico in dobimo

a) 7833 rezultatov (Slika1)



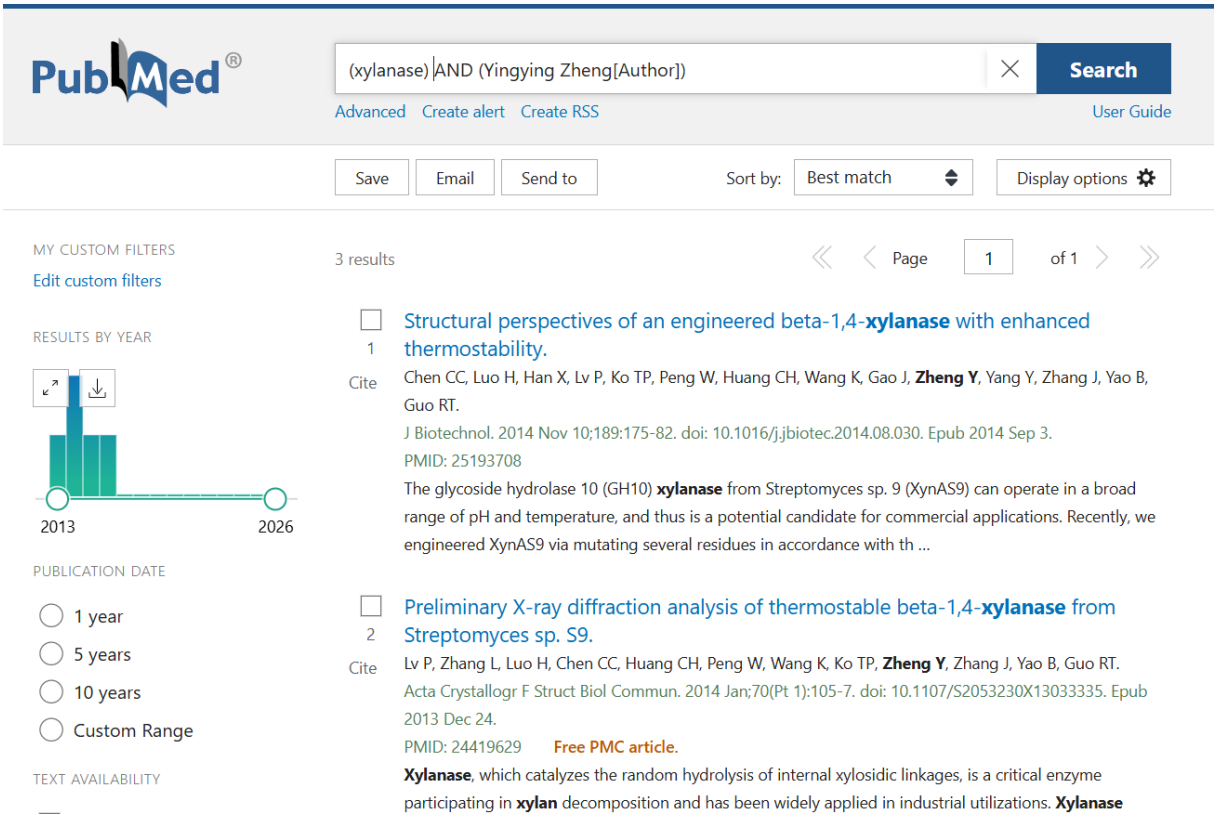
Slika 1: Iskanje v Pubmedu člankov, ki se navezujejo na ksilanazo.

Potem pa v iskanje dodamo še avtorja in to tako da gremo pod advanced in ponovno vpišemo xylanase potem pa izberemo term »Author » in zraven dopišemo »Yingying Zheng«. (Slika2)



Slika 2: Prikazuje vpis v advanced kjer združimo xylanase in Yingying Zheng.

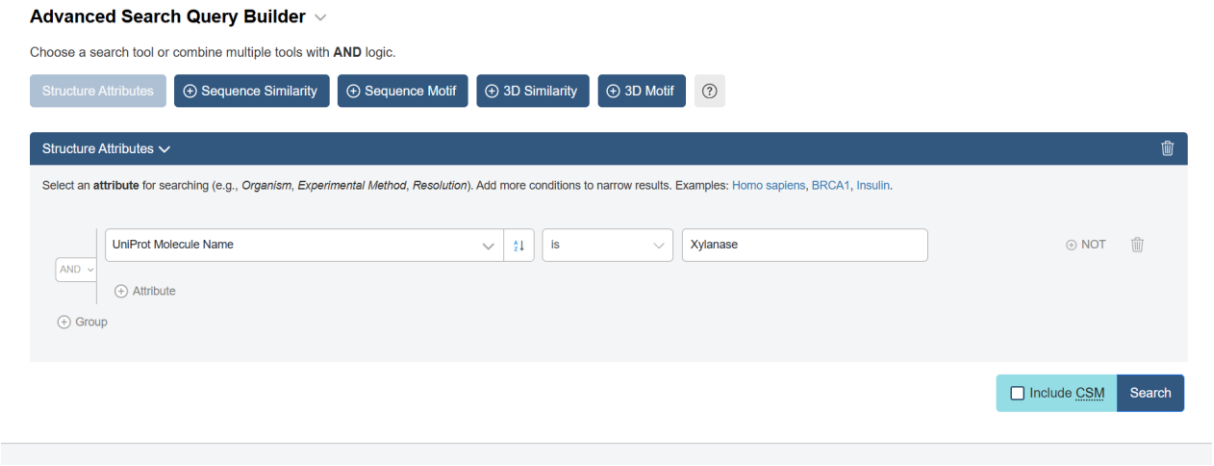
b) Kot rezultat iskanja dobimo 3 rezultate (Slika 3)



Slika 3: Rezultat našega iskanja so trije članki, za katere veljajo postavljeni pogoji

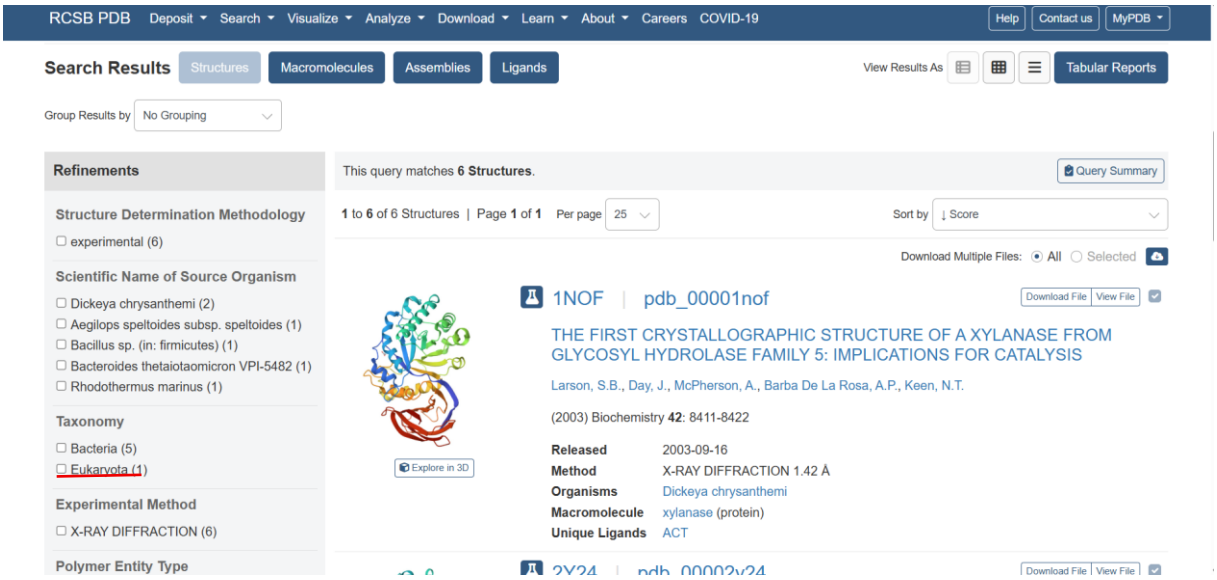
2.

Za iskanje v PDB smo morali iti pod »Advanced search« kjer smo izbrali atribut »UniProt Molecule Name« in potem vpisali xylanase. (Slika 4)



Slika 4: Advanced search z uniprot molecule name.

Dobimo 6 rezultatov. (Slika 5)



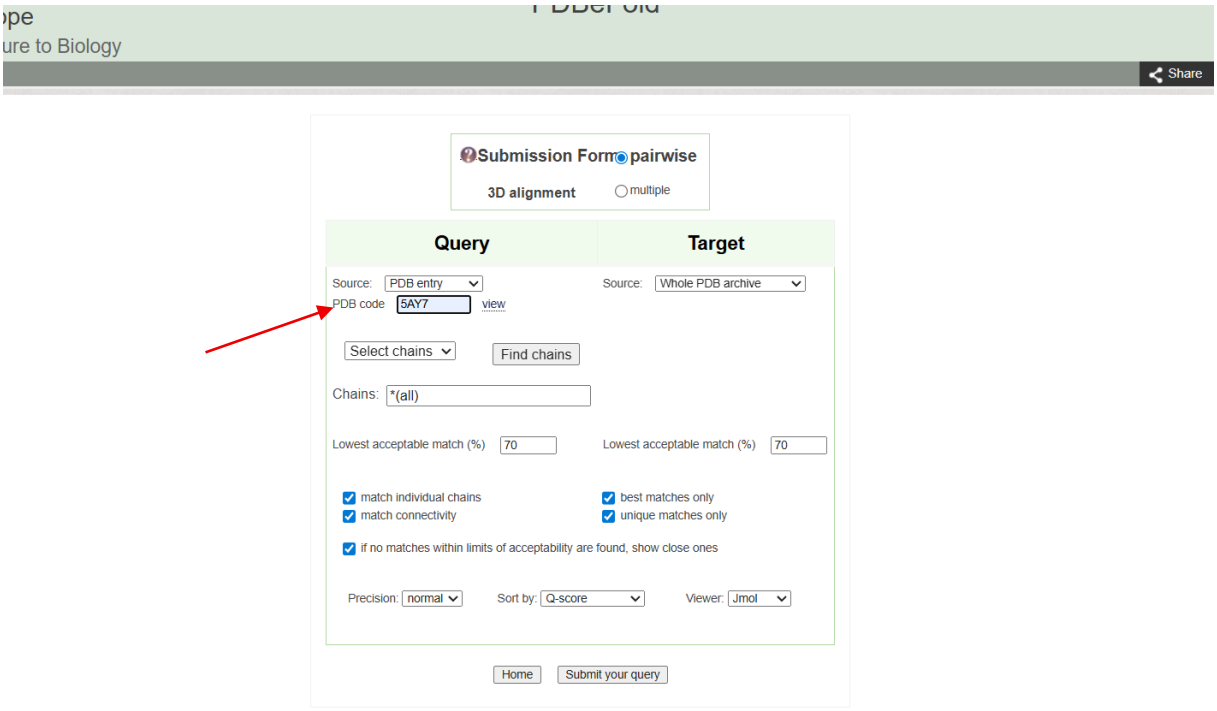
Slika 5: Rezultati pridobljeni iz Slike 4.

Potem pa na levi strani (z rdečo podčrtano na Sliki 5) izberemo »Eukaryota« kot zahteva naloga, da iščemo le evkariontske proteine. Dobimo le en zadetek. (Slika 6)

Slika 6: Pridobljen rezultat ko vklopimo filter za evkarionte.

a)Iz dobljenega rezultata (rdeči kvadrat na Sliki 6) vidimo, da je koda proteina 5AY7

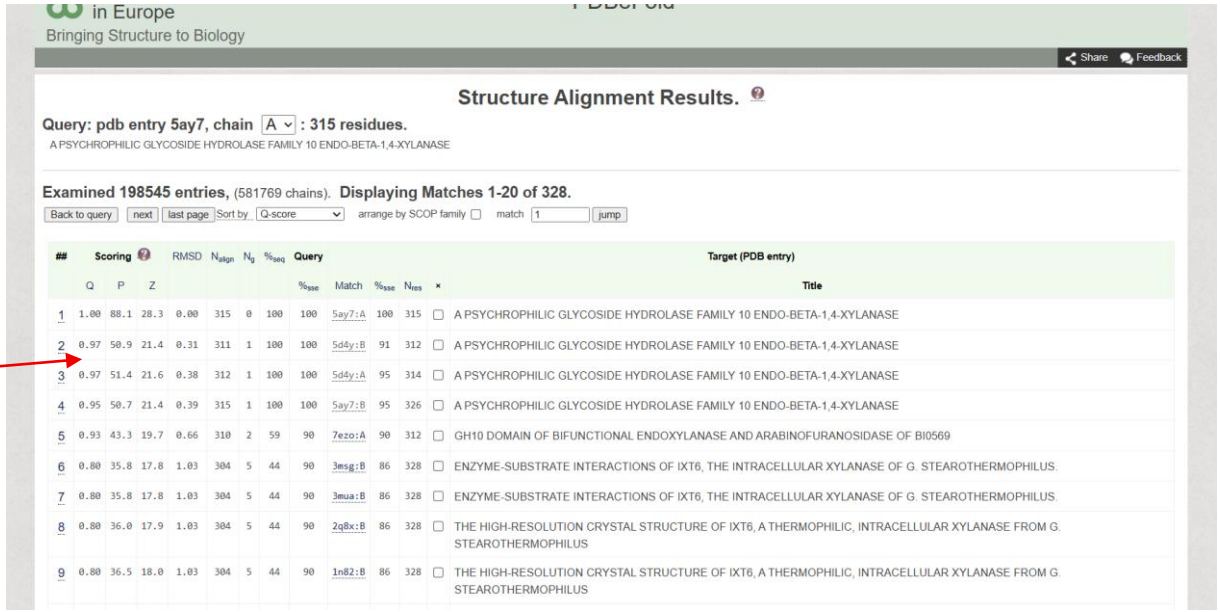
Da najdemo strukturno podoben protein gremo v PDBefolod in tam vpišemo PDB kodo. (Slika 7)



c3) 14 Apr 2014

Slika 7: Rdeča puščica prikazuje mesto vpisa kode v PDBefold.

Potem kliknemo »Submit your query«. Iz dobljenih rezultatov vidimo, da je prvi na seznamu 5AY7, kar je ista struktura, ki smo jo vnesli. Izberemo naslednjo možnost, kot vidimo na Sliki 8 sta naslednja dva zadetka obadva 5D4Y eden je edino A eden pa B. Ni pomembno katerega izberemo oba sta isti protein z isto Q vrednostjo.

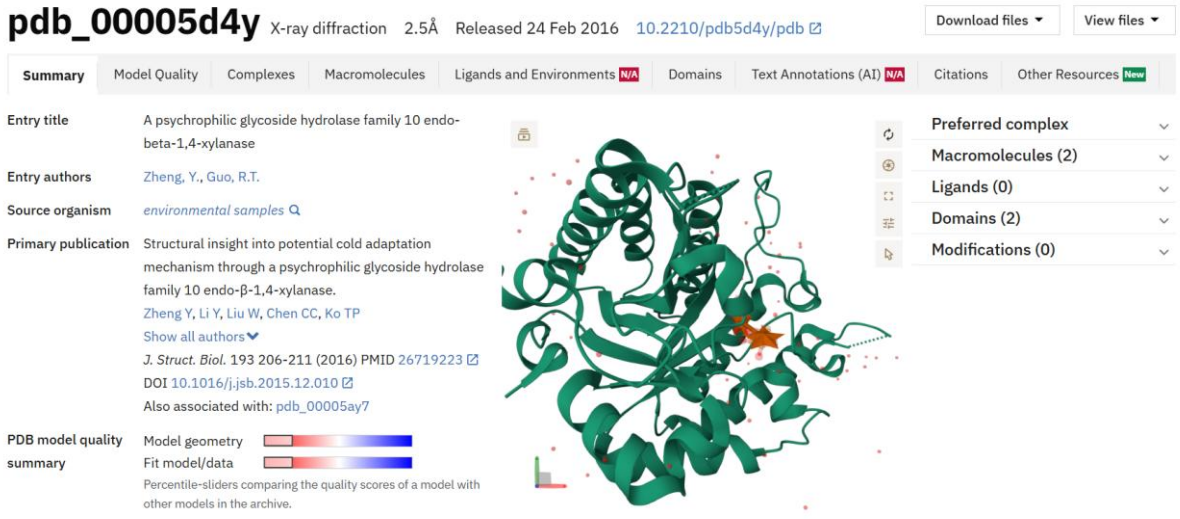


Slika 8: Dobljeni rezultati kjer vidimo, da je prvi zadetek 5AY7, drugi in tretji pa sta 5D4Y.

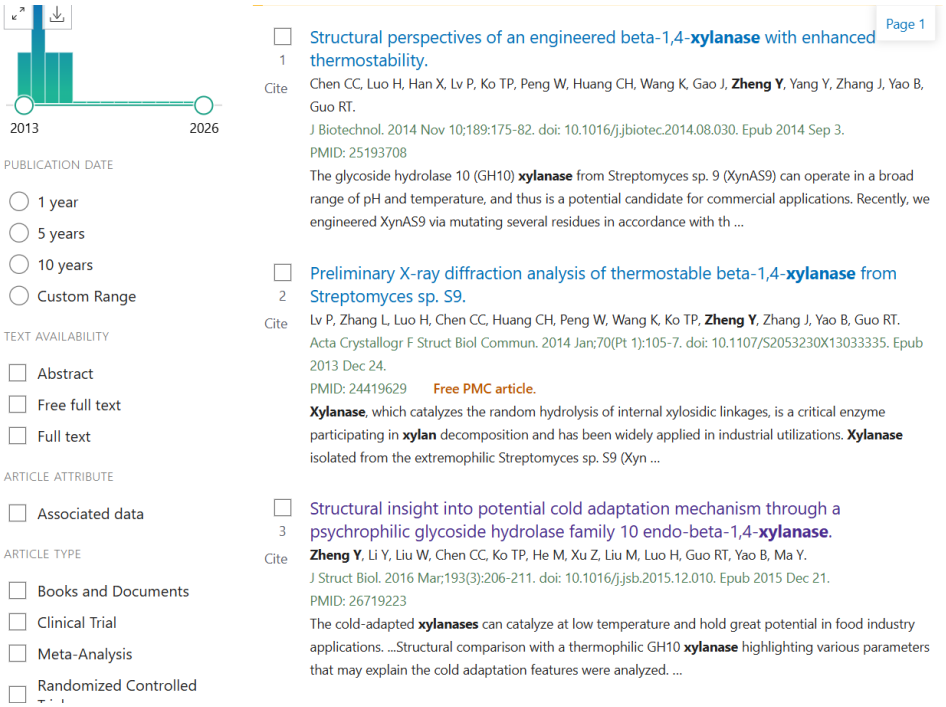
b) Odgovor vidimo v prvem stolpcu, da je Q vrednost 0.97

Da ugotovimo če je članek podoben prej najdenim člankom moramo odpreti protein 5D4Y tako da kliknemo na kodo v vrstici »Match«. Odpre se nam slika proteina in zraven lahko najdemo »Primary publication«, kjer nam navaja kje je bil članek prvo objavljen (Slika 9). Če ta zadetek primerjamo z zadetki ki smo jih dobili pri 1. b delu naloge, ugotovimo da je ena izmed opciji isti članek.

c) Naš članek je ena izmed opciji. Yingying Zheng se nahaja na prvem mestu. (Slika 10)



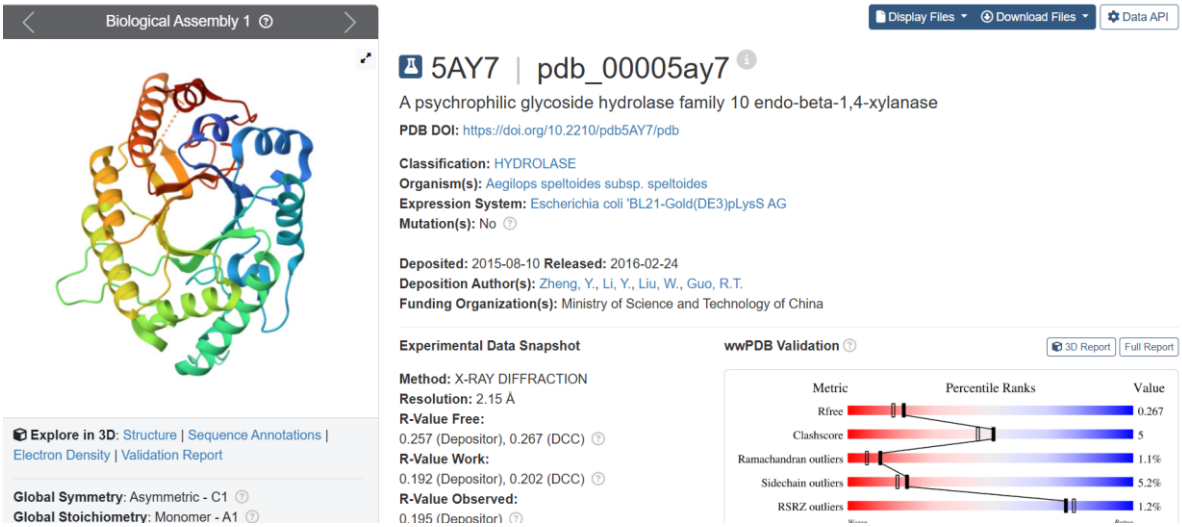
Slika 9: Pridobljene informacije o proteinu 5D4Y, med njimi tudi mesto publikacije.



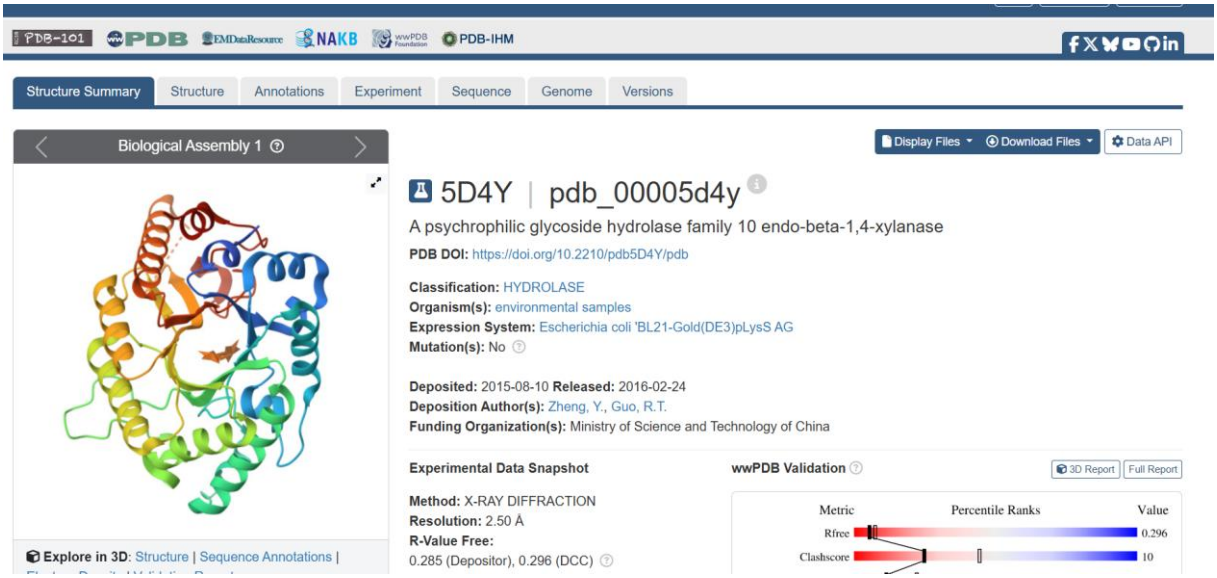
Slika 10: To so pridobljeni članki iz 1. b dela naloge, kjer lahko vidimo da se naslov, ki smo ga dobili od 5D4Y sklada z naslovom tretjega članka. Opazimo pa tudi da je Yingying Zheng napisan na prvem mestu, vendar je zapisano prvo priimek in potem kratica (vendar program že sam označi ime, ker smo vstavili filter pri iskanju).

3.

Informacije o encimih bo najlažje najti v PDB, ker imamo PDB id za oba.



Slika 11: Ko vpišemo kodo 5AY7 v iskalnik dobimo ven to sliko. Iz slike lahko razberemo, da pod »Organism(s)« piše »Aegilops speltoides subsp. speltoides«.



Slika 12: Ko vpišemo kodo 5D4Y v iskalnik dobimo ven to sliko. Iz slike lahko razberemo, da pod »Organism(s)« piše »environmental samples«.

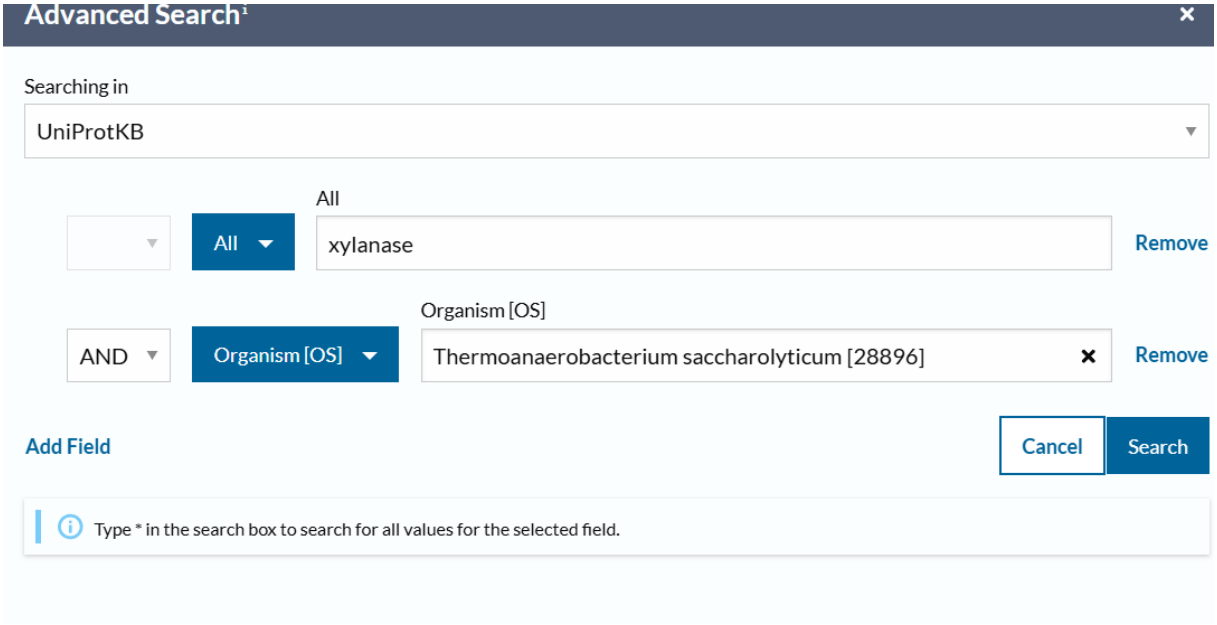
a) Encim 5AY7 je pridobljen iz »Aegilops speltoides subsp. Speltoides« (Slika 11). Encim 5D4Y pa je pridobljen iz »environmental samples« (Slika 12).

b) Pri obeh je ekspresijki sistem »Escherichia coli 'BL21-Gold(DE3)pLysS AG« (Slika 11 in 12). Zato vemo, da bo primeren ekspresijski sistem bakterijski.

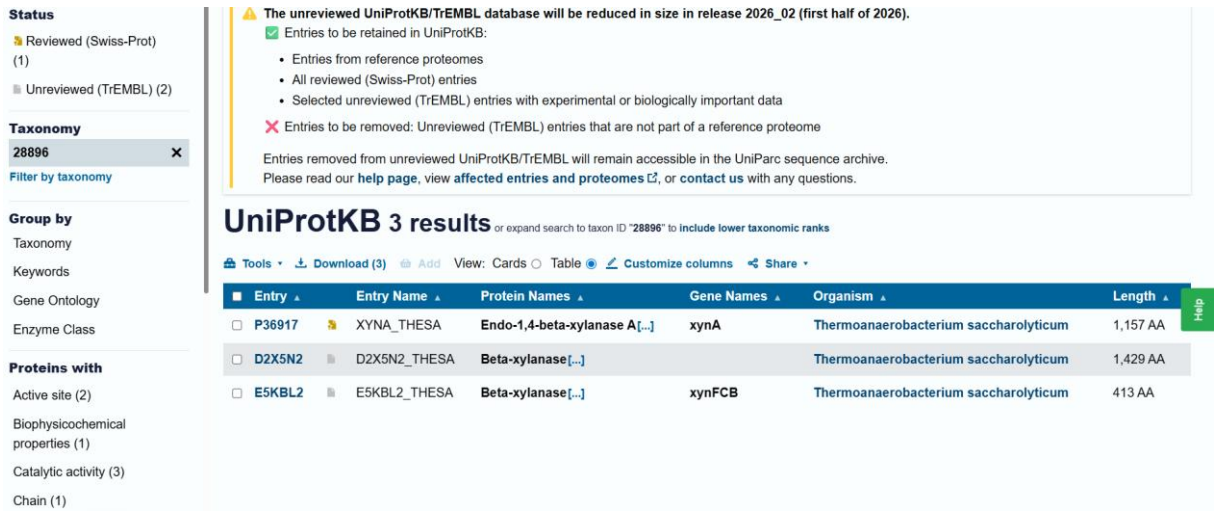
Slabosti so; ni kompleksnih posttranslacijskih modifikacij (npr. glikozilacije), težave z napačnim zvijanjem proteinov, tvorjenje netopnih agregatov, narobna tvorba disulfidnih vezi.

4.

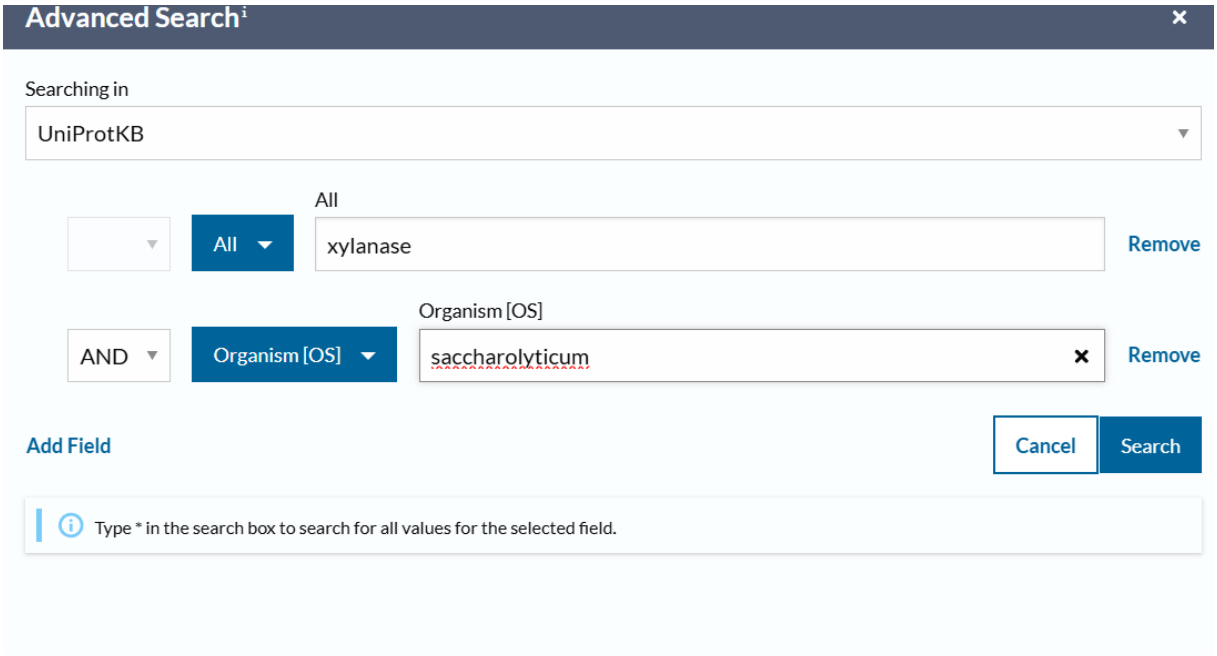
Da najdemo uniprot kodo moramo v UniProt, kjer v »Advanced Search« vpišemo »xylanase« in potem dodamo še en filter, ki ga nastavimo na »Organism [OS]« in tam potem poiščemo »saccharolyticum«. Tukaj lahko pride do dveh možnosti, če že v iskalniku izberemo »Thermoanaerobacterium saccharolyticum« potem dobimo samo en protein, ki je bil pregledan (Slika 13 in 14). Če pa v iskalnik vstavimo namesto »Thermoanaerobacterium saccharolyticum« samo »saccharolyticum« potem dobimo rezultate pod Slikama 15 in 16.



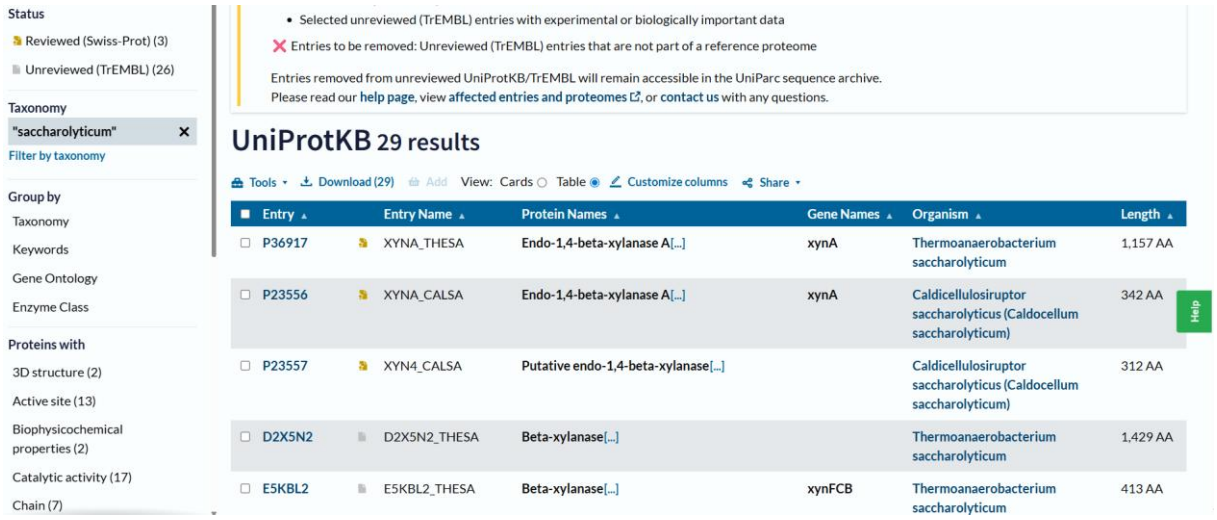
Slika 13: V uniprot vstavimo »xylanase« in »Thermoanaerobacterium saccharolyticum«.



Slika 14: Z iskalnikom iz Slike 13 dobimo tri možnosti, vendar je samo ena od njih »Reviewed« zatp je ta prava odločitev.



Slika 15: Če v uniprot iskalnik vstavimo »xylanase« in samo »saccharolyticum« dobimo rezultate na Sliki 16.



Slika 16: Tukaj dobimo 29 rezultatov, vendar so samo trije od njih »Reviewed«. Da najdemo pravega pa moramo upoštevati, da so vzeli termoanaerobno bakterijo, tako da je pravilna odločitev prvi protein.

a) Pri obeh možnostih pridemo do istega rezultata in to je P36917.

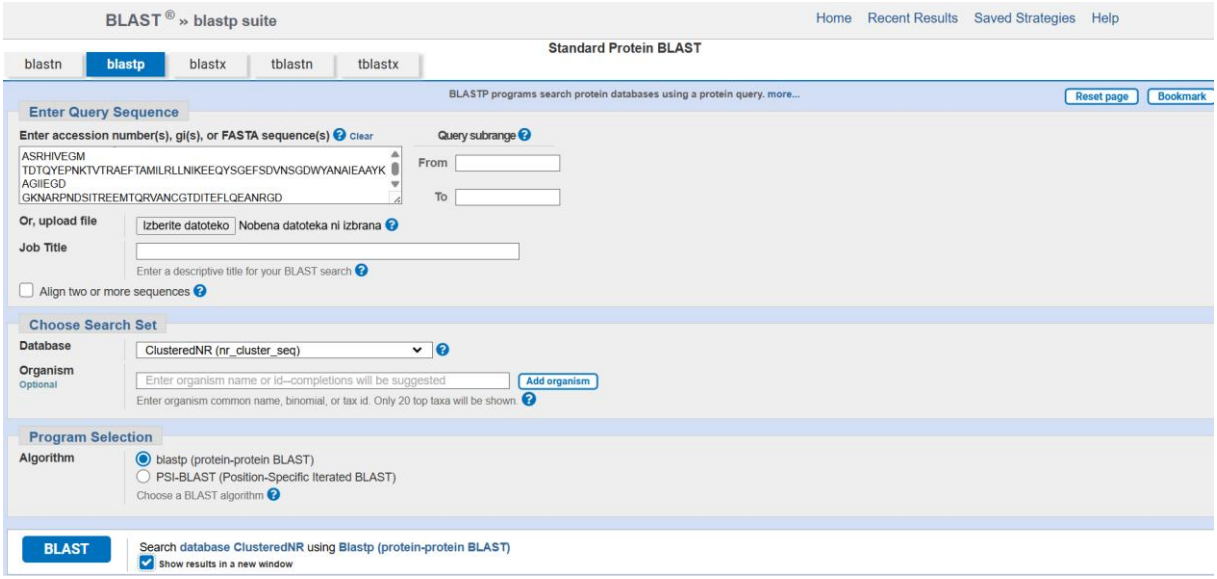
Da najdemo mesto mutacije na aktivnem mestu moramo prvo odpreti protein in potem iti pod »Function« del z »Features« kjer nam pišejo tri aaktivna mesta na ostankih 495, 537 in 600. Potem pa pogledamo še mutacije pod »Phenotypes & Variants« kjer vidimo da nastanejo tri pogoste mutacije na 537, 600 in 602. Če kliknemo na njih nam pokaže kateri aminokislinski ostanki se zamenjajo.

b) Mutacija se zgodi na aktivnem mestu 537 in 600, mutacija na 495 ni nikjer zabeležena zato sklepamo, da se ne zgodi ali pa še ni bila odkrita.

c) Pri 537 pride do mutacije, kjer se zamenja D → N, pri 600 pa se zamenja E → Q.

5.

Naloga nas spomni na program Blast oziroma natančneje Blastp. V katerega vstavimo dobljeno zaporedje in potem poiščemo podobne proteine. (Slika 17)

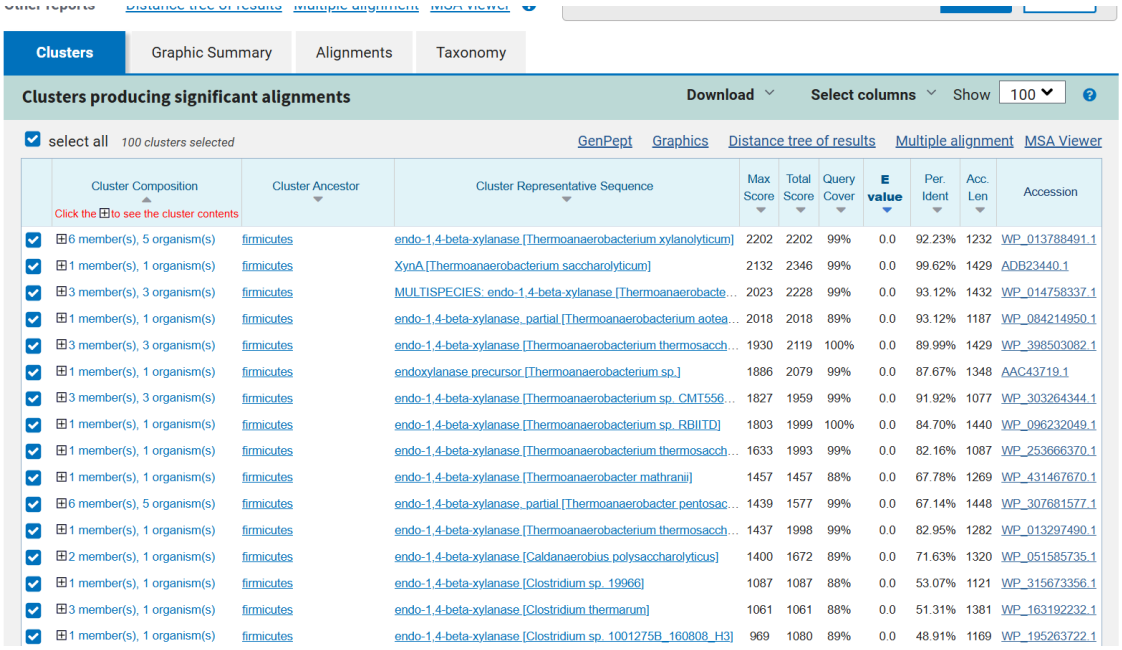


Slika 17: Odpremo blastp in vstavimo FASTA zaporedje ki nam je podano pod 5 nalogo.

Naloga nas sprašuje če je to zaporedje podobno tistemu od XynA, ki ga opazimo na drugem mestu med zadetki. (Slika 18)

a) Opazimo tudi, da je E-value enaka 0, kar pomeni da je verjetnost da to zaporedje naključno najdemo v bazi zelo zelo mala oziroma 0. Zato vemo da je ta encim podoben XynA.

b) Odstotek identičnosti vidimo da je 99.62%, torej je večina zaporedja istega.

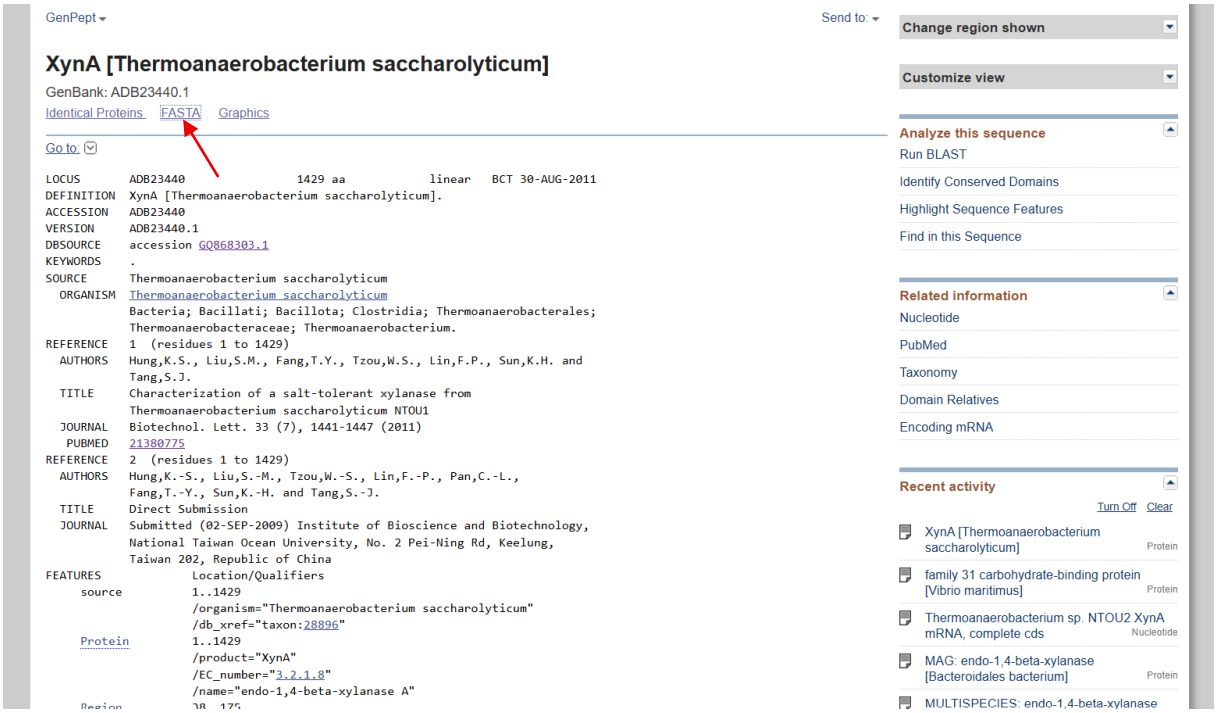


Cluster Composition (Click the ⊞ to see the cluster contents)	Cluster Ancestor	Cluster Representative Sequence	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
⊞6 member(s), 5 organism(s)	firmicutes	endo-1,4-beta-xylanase [Thermoanaerobacterium xylanolyticum]	2202	2202	99%	0.0	92.23%	1232	WP_013788491.1
⊞1 member(s), 1 organism(s)	firmicutes	XynA [Thermoanaerobacterium saccharolyticum]	2132	2346	99%	0.0	99.62%	1429	ADB23440.1
⊞3 member(s), 3 organism(s)	firmicutes	MULTISPECIES: endo-1,4-beta-xylanase [Thermoanaerobacte…	2023	2228	99%	0.0	93.12%	1432	WP_014758337.1
⊞1 member(s), 1 organism(s)	firmicutes	endo-1,4-beta-xylanase, partial [Thermoanaerobacterium aotea…	2018	2018	89%	0.0	93.12%	1187	WP_084214950.1
⊞3 member(s), 3 organism(s)	firmicutes	endo-1,4-beta-xylanase [Thermoanaerobacterium thermosacch…	1930	2119	100%	0.0	89.99%	1429	WP_398503082.1
⊞1 member(s), 1 organism(s)	firmicutes	endoxylanase precursor [Thermoanaerobacterium sp.]	1886	2079	99%	0.0	87.67%	1348	AAC43719.1
⊞3 member(s), 3 organism(s)	firmicutes	endo-1,4-beta-xylanase [Thermoanaerobacterium sp. CMT556…	1827	1959	99%	0.0	91.92%	1077	WP_303264344.1
⊞1 member(s), 1 organism(s)	firmicutes	endo-1,4-beta-xylanase [Thermoanaerobacterium sp. RBIITD]	1803	1999	100%	0.0	84.70%	1440	WP_096232049.1
⊞1 member(s), 1 organism(s)	firmicutes	endo-1,4-beta-xylanase [Thermoanaerobacterium thermosacch…	1633	1993	99%	0.0	82.16%	1087	WP_253666370.1
⊞1 member(s), 1 organism(s)	firmicutes	endo-1,4-beta-xylanase [Thermoanaerobacter mathranii]	1457	1457	88%	0.0	67.78%	1269	WP_431467670.1
⊞6 member(s), 5 organism(s)	firmicutes	endo-1,4-beta-xylanase, partial [Thermoanaerobacter pentosac…	1439	1577	99%	0.0	67.14%	1448	WP_307681577.1
⊞1 member(s), 1 organism(s)	firmicutes	endo-1,4-beta-xylanase [Thermoanaerobacterium thermosacch…	1437	1998	99%	0.0	82.95%	1282	WP_013297490.1
⊞2 member(s), 1 organism(s)	firmicutes	endo-1,4-beta-xylanase [Caldanaerobius polysaccharolyticus]	1400	1672	89%	0.0	71.63%	1320	WP_051585735.1
⊞1 member(s), 1 organism(s)	firmicutes	endo-1,4-beta-xylanase [Clostridium sp. 19966]	1087	1087	88%	0.0	53.07%	1121	WP_315673356.1
⊞3 member(s), 1 organism(s)	firmicutes	endo-1,4-beta-xylanase [Clostridium thermarum]	1061	1061	88%	0.0	51.31%	1381	WP_163192232.1
⊞1 member(s), 1 organism(s)	firmicutes	endo-1,4-beta-xylanase [Clostridium sp. 1001275B_160808_H3]	969	1080	89%	0.0	48.91%	1169	WP_195263722.1

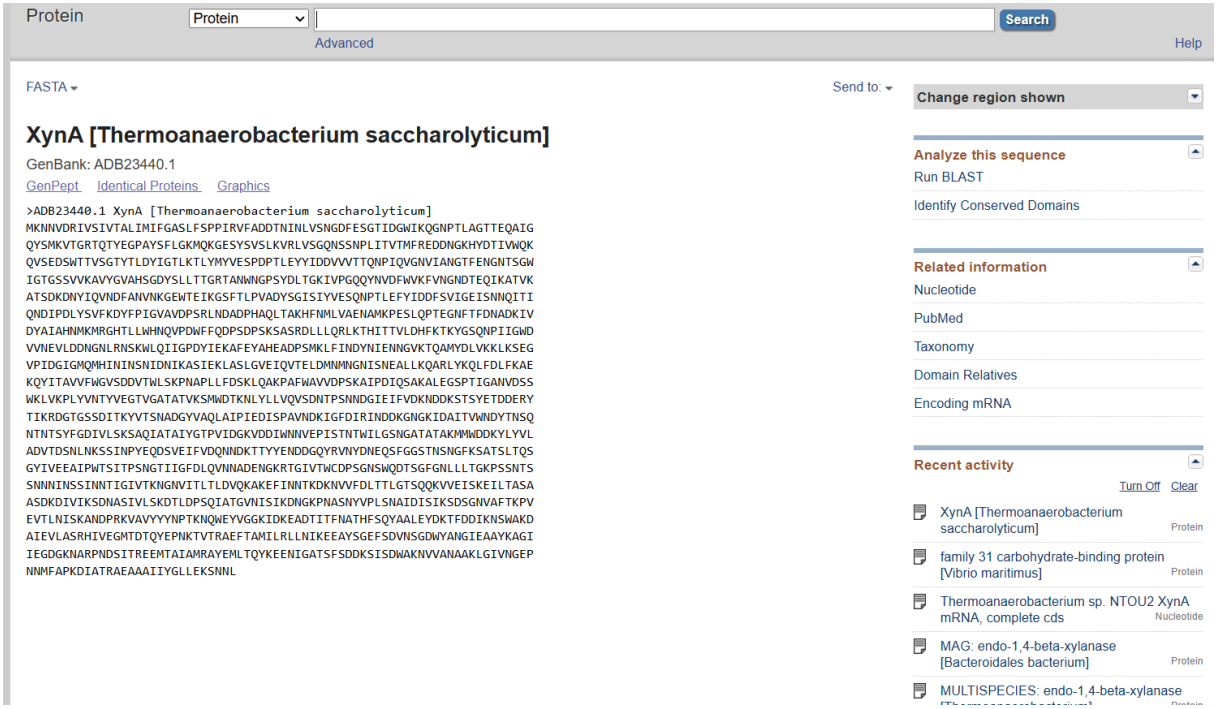
Slika 18: Dobljen rezultat z blastanjem, kjer lahko vidimo vse možne podobne proteine.

c) Za poravnavo dveh zaporedji uporabljamo ali Needle ali Water, ker sta ta dva encima skoraj da identična je boljša poravnava Needle. Needle je globalna poravnava, ki poravna celoti obeh zaporedji in zato se uporablja ko sta zaporedji podobno dolgi (ti dve sta) in ko sta tudi si med seboj dokaj podobni (tudi sta). Water pa je lokalna poravnava, ki poravna zaporedje z delom zaporedja, ki najbolj ustreza (npr. eno domeno), zato ni najbolj primerna za te dve zaporedji, ki sta si skoraj identični

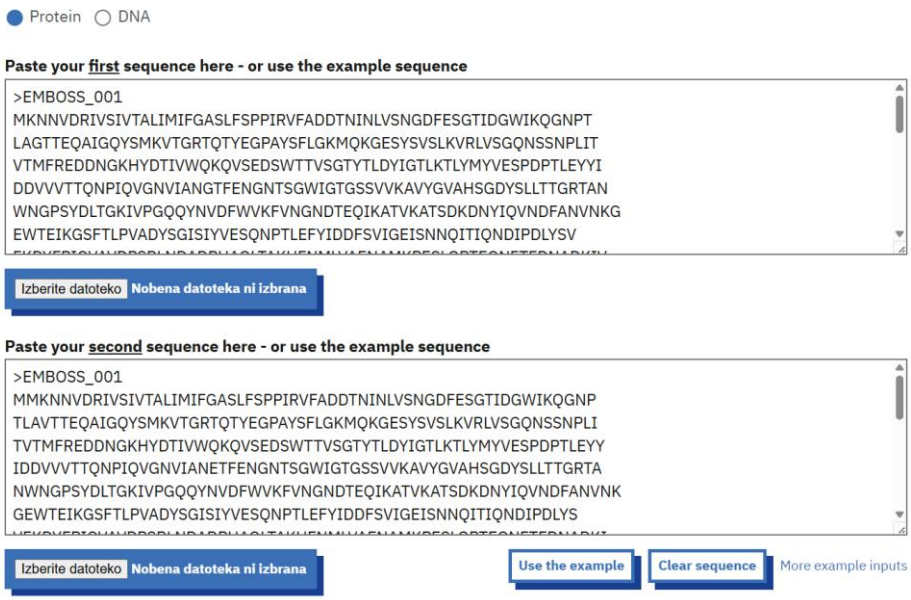
Da odkrijemo score poravnave moramo odpreti Needle, kam moramo prilepiti obe zaporedji. Prvo zaporedje je podano že v nalogi, drugo zaporedje pa dobimo če kliknemo pri »Accession« na kodo ADB23440.1. In potem kliknemo na »FASTA« (Slika 19), dobimo celo zaporedje tega proteina (Slika 20), ki ga kopiramo.



Slika 19: Informacije, ki se nam odprejo če kliknemo na ADB23440.1. Potem pa nadaljujemo tako, da kliknemo na FASTA.



Slika 20: Dobljeno FASTA zaporedje, ki ga potem kopiramo in prilepimo v Needle.



Slika 21: Tako zgleda ko vstavimo obe zaporedji v Needle.

d) Dobimo poravnavo teh dveh zaporedji in če pogledamo pod »Score« vidim, da je odgovor na vprašanje 5899. (Slika 22)

Download

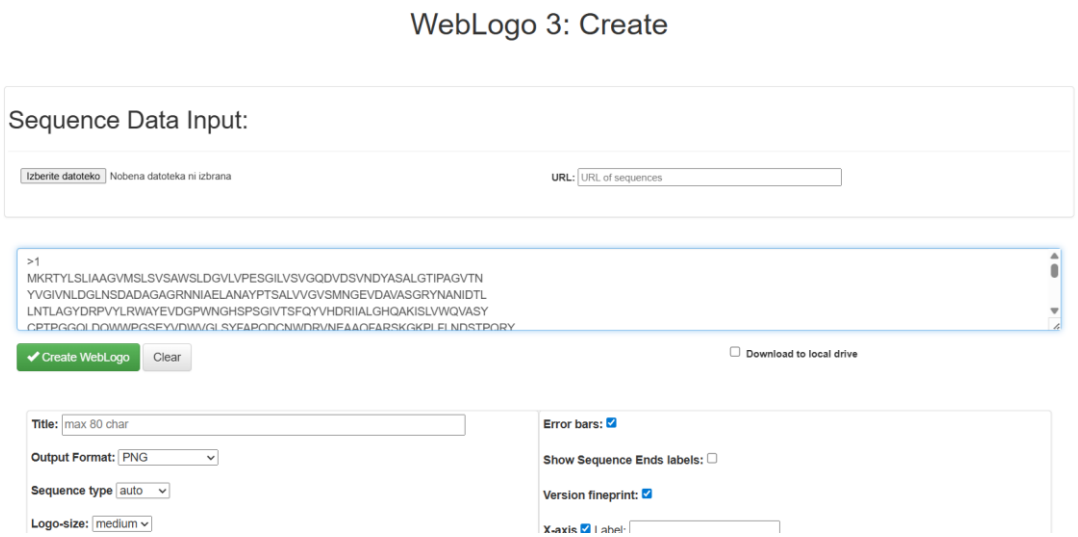
```
# Commandline: needle
#    -auto
#    -stdout
#    -asequence emboss_needle-I20260512-094155-0410-10521866-p1m.asequence
#    -bsequence emboss_needle-I20260512-094155-0410-10521866-p1m.bsequence
#    -datafile EBLOSUM62
#    -gapopen 10.0
#    -gapextend 0.5
#    -endopen 10.0
#    -endextend 0.5
#    -aformat3 pair
#    -sprotein1
#    -sprotein2
# Align_format: pair
# Report_file: stdout
########################################

#=======================================
#
# Aligned_sequences: 2
# 1: EMBOSS_001
# 2: EMBOSS_001
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 1431
# Identity:    1148/1431 (80.2%)
# Similarity:  1154/1431 (80.6%)
# Gaps:         256/1431 (17.9%)
# Score: 5899.0
#
```

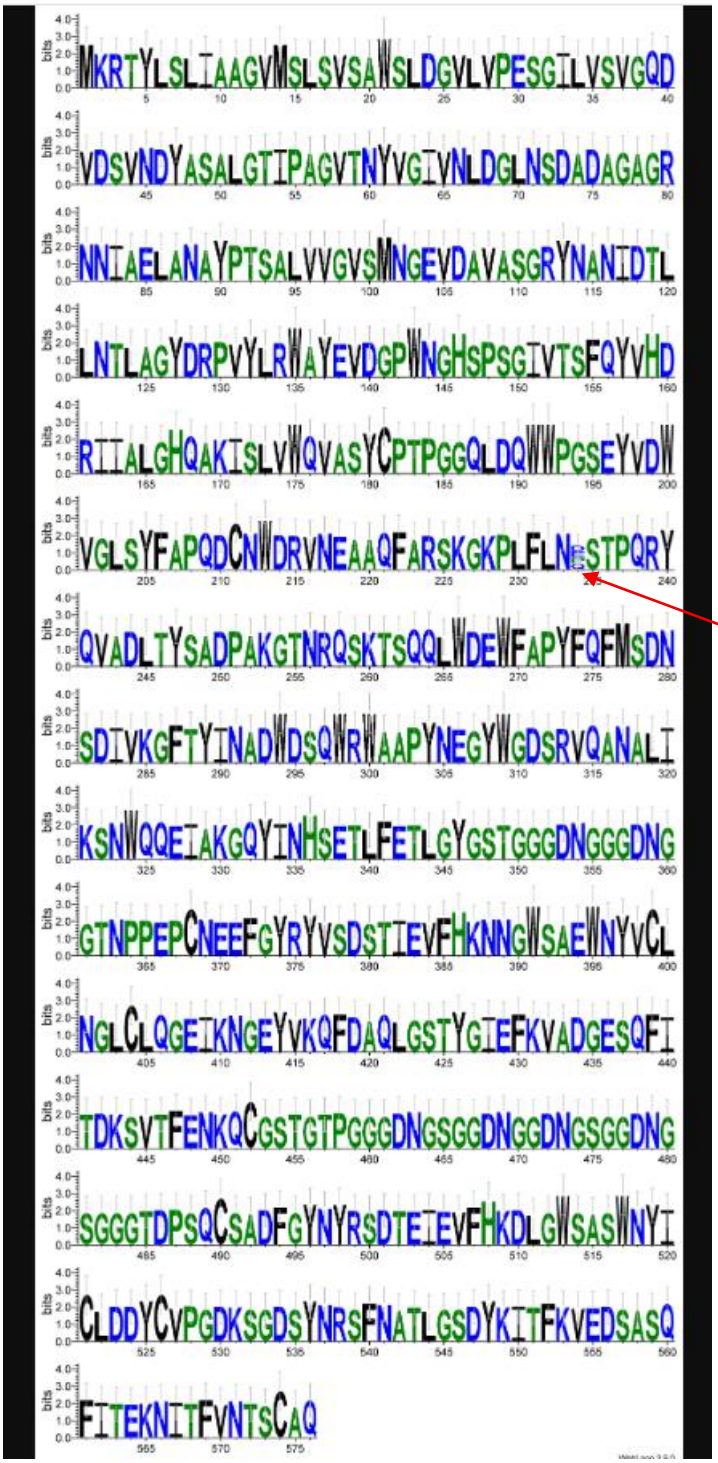
Slika 22: Dobljena poravnava, vidimo da sta zaporedji res skoraj identični (kar 80%), vendar nas zanima Score.

6.

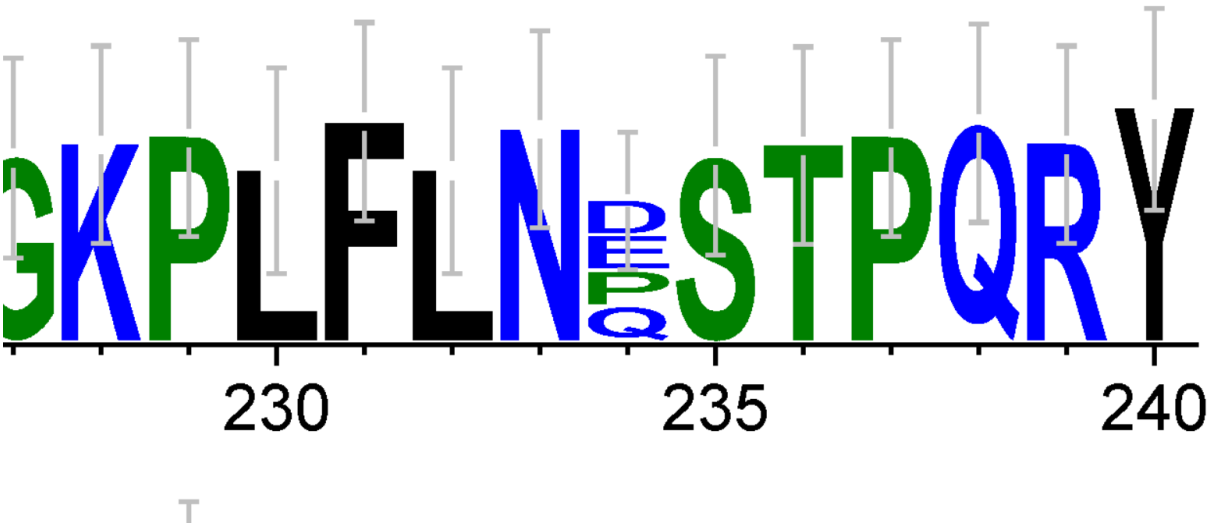
Odpremo Weblogo in vanj prilepimo 4 zaporedja, ki so nam jih posredovali. (Slika 23)



Slika 23: V WebLogo vstavimo vsa štiri zaporedja, da dobimo prikaz vseh aminokislinskih ostankov in kje se zaporedja razlikujejo.



Slika 24: Vizualni prikaz vseh štirih zaporedji v WebLogu. Rdeča puščica prikazuje edino mesto v zaporedju kjer opazimo da je več črk na istem aminokislinskem zaporedju.

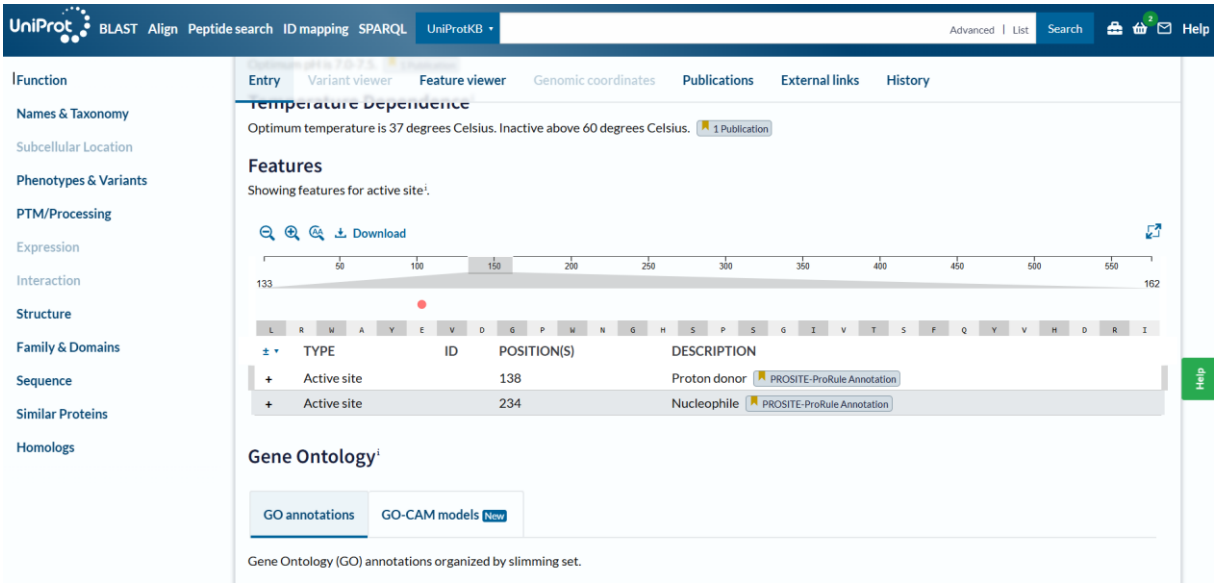


Slika 25: Ko približamo opazimo da pride do spremembe ostankov na mestu 234.

a) Na Sliki 25 vidimo, da edina razlika med zaporedji je na aminokislinskem ostanku 234.

Da bi ugotovili funkcijo tega ostanka gremo na UniProt, kjer vpišemo v iskalnik D5MP61. Potem gremo pod »Function« in spodaj pod »Features«, opazimo da je 234 aminokislinski ostanek aktivno mesto proteina. (Slika 26)

b) Aktivno mesto



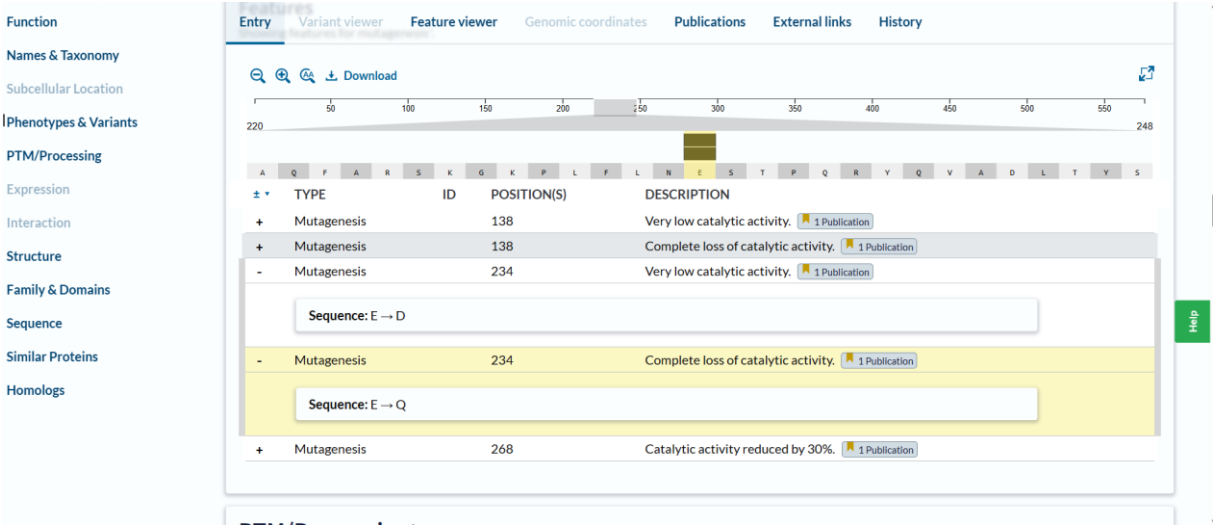
Slika 26: Ko vpišemo uniprot kodo v iskalnik se nam odpre ta protein

Če pogledamo v Weblogo vidimo, da je aminokislinski ostanek v zaporedjih E, D, Q ali P (Slika 25). Če kliknemo na aktivno mesto v uniprotu vidimo da je na tistem mestu E. Če gremo pod »Phenotypes & Variants« vidimo da se na 234 zgodita dve mutaciji in to v D in Q. (Slika 27)

c) Če je v zaporedju na mestu 234 E potem je katalitična aktivnost normalna oziroma nespremenjena. Če se tam nahaja D potem to povzroči nizko katalitično aktivnost. Q na mestu 234 pa povzroči popolno izgubo katalitične aktivnosti (Slika 27). Tudi če zamenjam E s P se bo

izgubila katalitična aktivnost, saj E (glutamat) ima negativen naboj med tem ko P (prolin) nima naboja. To pomeni da se izgubijo pomembne elektrostatske interakcije, kar povzroči izgubo katalitičnega mesta.

d)



Slika 27: Vidimo vse možne mutacije, mi se osredotočimo na tiste na aminokislinskem ostanku 234. Zgodita se dve mutaciji E → D in E →Q.

d) Sprememba v P bo spremenila strukturo, saj kot sem pojasnil že v c) delu se spremeni naboj kar podre interakcije, kar vodi do spremembe strukture proteina.