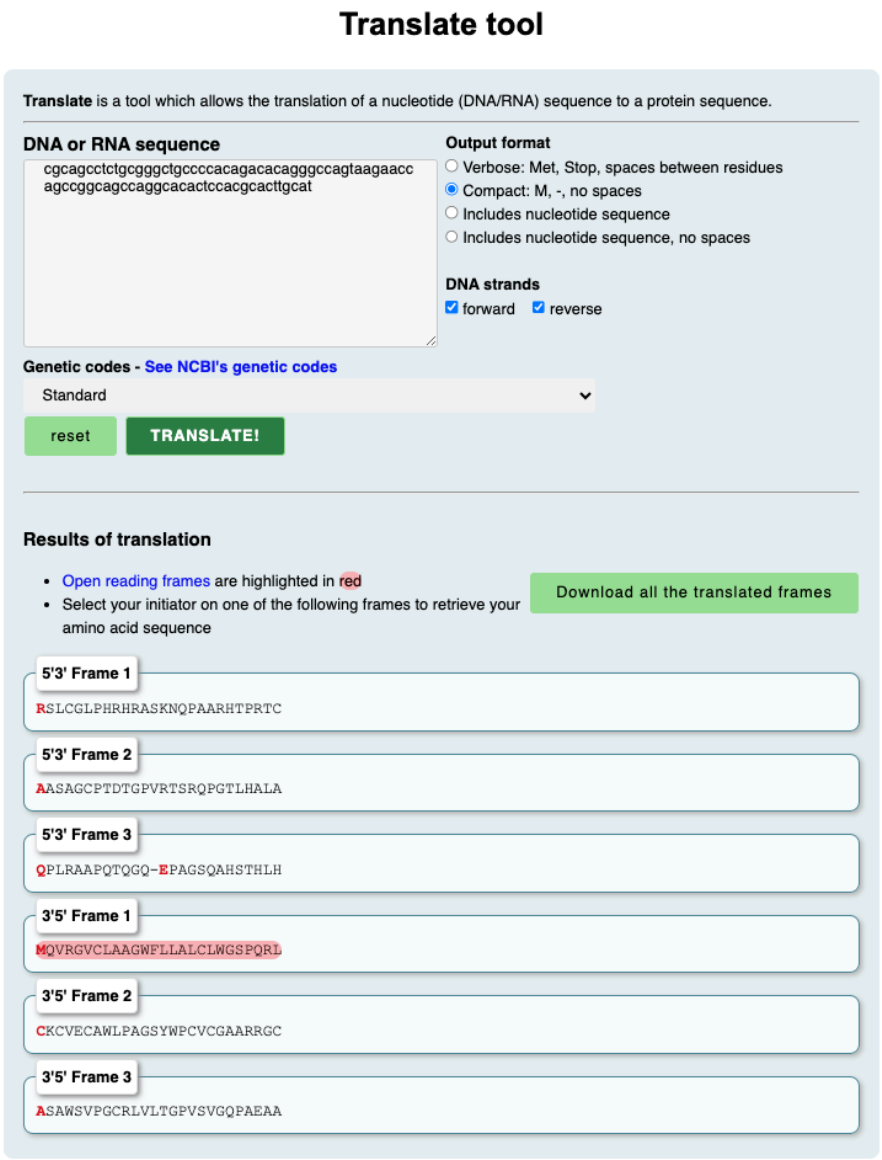
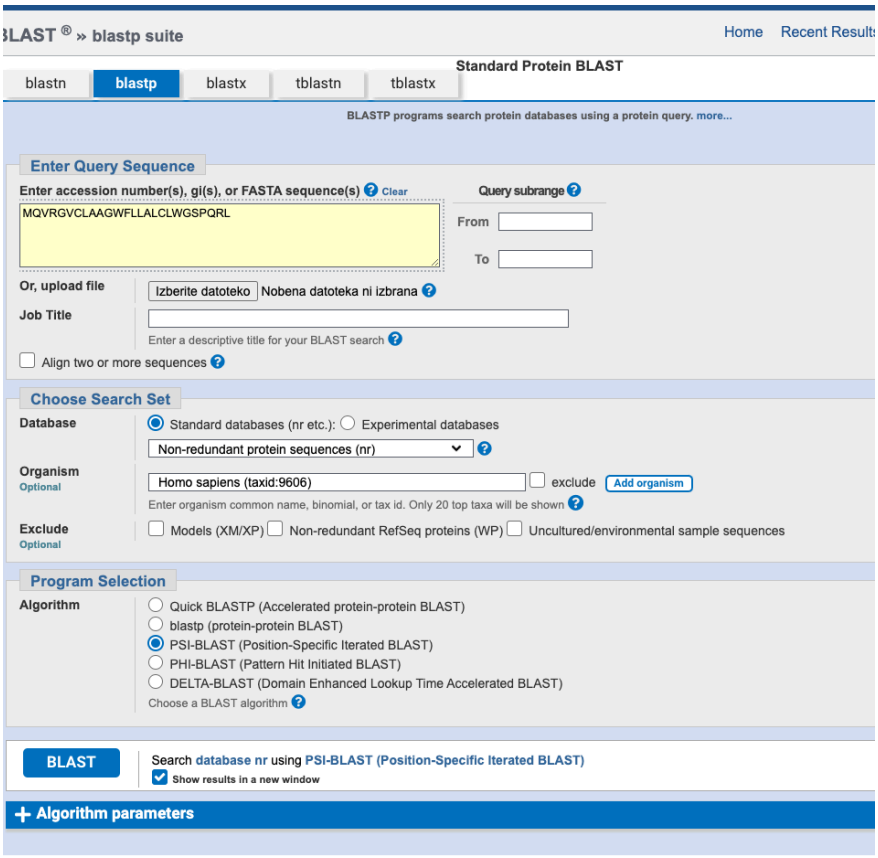
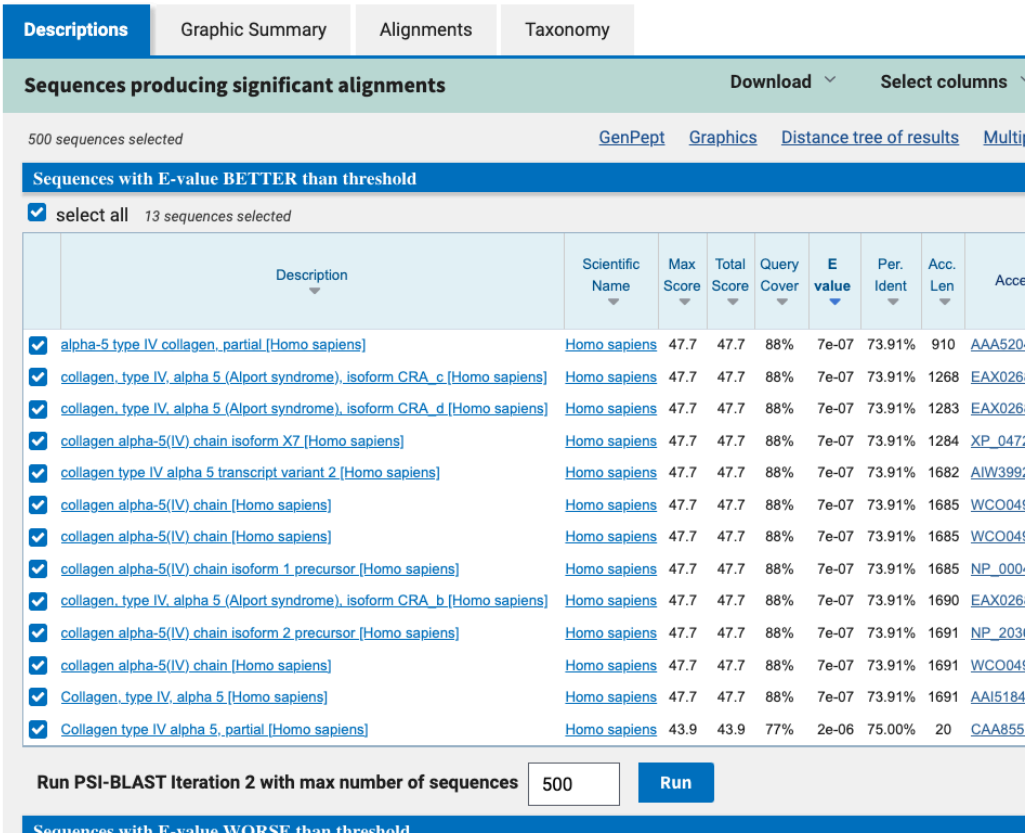
Zaposlili so te v biokemijskem laboratoriju, kjer se ukvarjajo z raziskovanjem nekega sindroma. Oboleli s tem sindromom se že od zgodnjega otroštva trpijo za kroničnimi boleznimi ledvic. Nekateri s časom oz. po propadu delovanja ledvic se celo soočajo z izgubo sluha. Hočejo preveriti tvoje biokemijsko znanje, in sicer moraš ugotoviti, mutacija katerega proteina povzroča ta sindrom. Trenutno preučujejo ta sindrom na miši, zato ti dajo mišjo mRNA za del tega proteina. Uspešno si ga posekvenciral in to je tvoj rezultat.
del mRNA pri miši: cgcag cctctgcggg ctgccccaca gacacagggc cagtaagaac cagccggcag ccaggcacac tccacgcact tgcat

1. Kateri protein je homologen temu pri človeku? Napiši UniProt ID. Pomagaj si z Expasy Translate Tool. Glede na to, da primerjaš mišji in človeški protein, katera vrsta blasta je najprimernejša?





Uporabi se PSI-BLAST, ker omogoča detekcijo bolj oddaljenih homolov.



Descriptions | Graphic Summary | Alignments | Taxonomy

Sequences producing significant alignments Download Select columns

500 sequences selected GenPept Graphics Distance tree of results Multi

Sequences with E-value BETTER than threshold

select all *13 sequences selected*

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Acce
alpha-5 type IV collagen, partial [Homo sapiens]	Homo sapiens	47.7	47.7	88%	7e-07	73.91%	910	AAA520
collagen, type IV, alpha 5 (Alport syndrome), isoform CRA_c [Homo sapiens]	Homo sapiens	47.7	47.7	88%	7e-07	73.91%	1268	EAX026
collagen, type IV, alpha 5 (Alport syndrome), isoform CRA_d [Homo sapiens]	Homo sapiens	47.7	47.7	88%	7e-07	73.91%	1283	EAX026
collagen alpha-5(IV) chain isoform X7 [Homo sapiens]	Homo sapiens	47.7	47.7	88%	7e-07	73.91%	1284	XP_047
collagen type IV alpha 5 transcript variant 2 [Homo sapiens]	Homo sapiens	47.7	47.7	88%	7e-07	73.91%	1682	AIW399
collagen alpha-5(IV) chain [Homo sapiens]	Homo sapiens	47.7	47.7	88%	7e-07	73.91%	1685	WCO04
collagen alpha-5(IV) chain [Homo sapiens]	Homo sapiens	47.7	47.7	88%	7e-07	73.91%	1685	WCO04
collagen alpha-5(IV) chain isoform 1 precursor [Homo sapiens]	Homo sapiens	47.7	47.7	88%	7e-07	73.91%	1685	NP_000
collagen, type IV, alpha 5 (Alport syndrome), isoform CRA_b [Homo sapiens]	Homo sapiens	47.7	47.7	88%	7e-07	73.91%	1690	EAX026
collagen alpha-5(IV) chain isoform 2 precursor [Homo sapiens]	Homo sapiens	47.7	47.7	88%	7e-07	73.91%	1691	NP_203
collagen alpha-5(IV) chain [Homo sapiens]	Homo sapiens	47.7	47.7	88%	7e-07	73.91%	1691	WCO04
Collagen, type IV, alpha 5 [Homo sapiens]	Homo sapiens	47.7	47.7	88%	7e-07	73.91%	1691	AAI5184
Collagen type IV alpha 5, partial [Homo sapiens]	Homo sapiens	43.9	43.9	77%	2e-06	75.00%	20	CAA855

Run PSI-BLAST Iteration 2 with max number of sequences 500 Run

Sequences with E-value WORSE than threshold

alpha-5 type IV collagen, partial [Homo sapiens]

GenBank: AAA52046.1

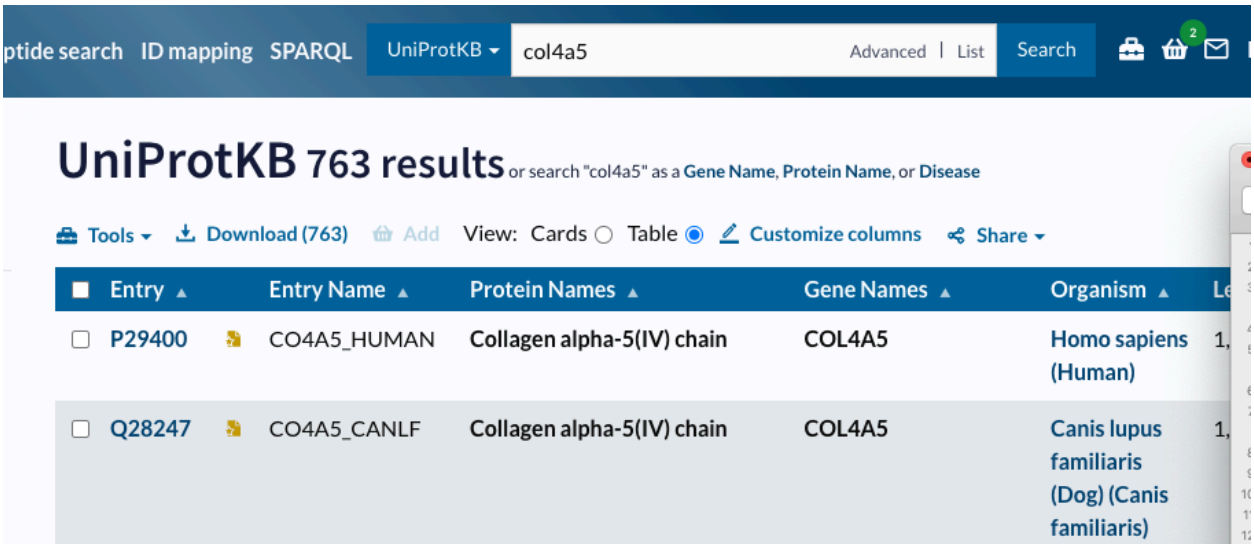
Identical Proteins FASTA Graphics

Go to:

```
LOCUS       AAA52046                  910 aa            linear   PRI 01-NOV-1994
DEFINITION  alpha-5 type IV collagen, partial [Homo sapiens].
ACCESSION   AAA52046
VERSION     AAA52046.1
DBSOURCE    locus HUMCOL4A5X accession M90464.1
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 910)
  AUTHORS   Zhou,J., Hertz,J.M., Leinonen,A. and Tryggvason,K.
  TITLE     Complete amino acid sequence of the human alpha 5 (IV) collagen
            chain and identification of a single-base mutation in exon 23
            converting glycine 521 in the collagenous domain to cysteine in an
            Alport syndrome patient
  JOURNAL   J. Biol. Chem. 267 (18), 12475-12481 (1992)
   PUBMED   1352287
COMMENT     On Oct 3, 1994 this sequence version replaced gi:180827.
            Method: conceptual translation.
FEATURES             Location/Qualifiers
     source          1..910
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="Xq22"
                     /tissue_type="kidney"
     Protein         1..910
                     /product="alpha-5 type IV collagen"
     sig_peptide     1..40
                     /note="G00-120-596; putative"
     Region          <452..>698
                     /region_name="gly_rich_SclB"
                     /note="LPXTG-anchored collagen-like adhesin Scl2/SclB;
                     NF038329"
                     /db_xref="CDD:468478"
     CDS             1..910
                     /gene="COL4A5"
                     /coded_by="M90464.1:203..2933"
                     /db_xref="GDB:G00-120-596"
ORIGIN
```

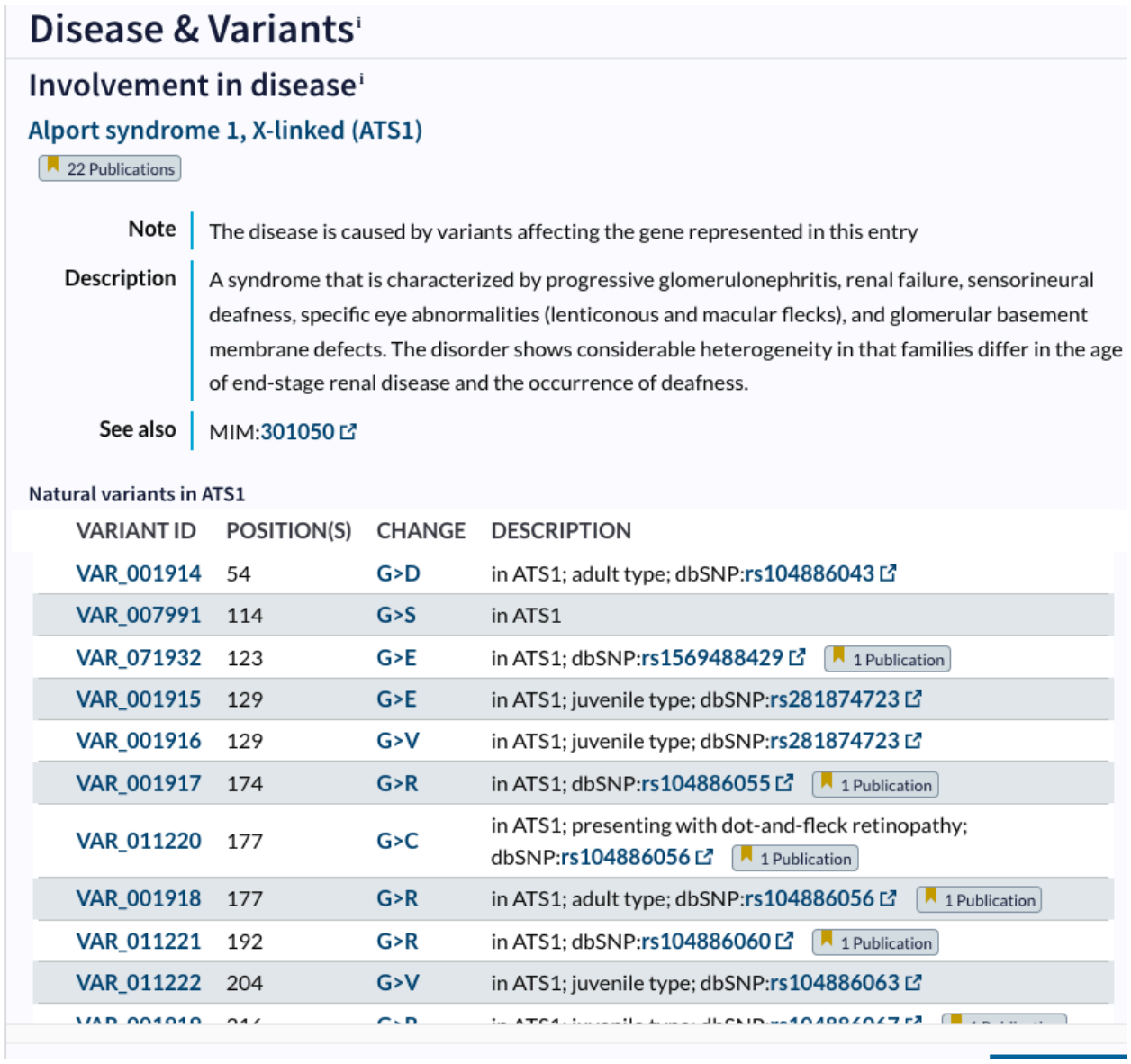
```
        1 mklrgvslaa glfllalslw ggpaeaaacy gcspgskcdc sgikgekger gfpgleghpg
       61 lpgfpgpegp pgprggkgdd gipgppgpkg irgppglpgf pgtpglpgmp ghdgapgpgg
      121 ipgcngtkge rgfpgspgfp glggppgppg ipgmkgepgs iimsslpgpk gnpgypgppg
      181 igglpgptgi pgpigppgpp glmgppgppg lpgpkgnmgl nfggpkgekg egglggppgp
      241 pggiseqkrp idvefgkgdg glpgdrgppg ppgirgppgp pggekgekge ggepgkrgkp
      301 gkdgenggpg ipglpgdpgy pgepgrdgek ggkgdtgppg ppglviprpg tgitigekgn
      361 iglpglpgek gergfpgigg ppglpgppga avmgppgppg fpgerggkgd egppgisipg
      421 ppgldggpga pglpgppgpa gphippsdei cepgppgppg spgdkglgge ggvkgdkgdt
      481 cfncigtgis gppggpglpg lpgppgslgf pggkgekgga gatgpkglpg ipgapgapgf
      541 pgskgepgdi ltfpgmkgdk gelgspgapg lpglpgtpgg dglpglpgpk gepggitfkg
      601 ergppgnpgl pglpgnigpm gppgfgppgp vgekgiggva gnpggpgipg pkgdpggtit
      661 gpgkpglpgn pgrdgdvglp gdpglpggpg lpgipgskge pgipgiglpg ppgpkgfpgi
      721 pgppgapgtp griglegppg ppgfpgpkge pgfalpgppg ppglpgfkga lgpkgdrgfp
      781 gppgppgrtg ldglpgpkgd vgpngppgpm gppglpgigv ggppgppgip gpigppglhg
      841 ipgekgdpgp pgldvpgppg ergspgipga pgpigppgsp glpgkagasg f
      901 mgppgppgpl
```

1..910
/gene="COL4A5"



Protein: kolagen 4 veriga 5
UniProt ID: P29400

2. Kateri sindrom preučuje ta laboratorij? Katera aminokislina (ime) najpogosteje mutira pri tem sindromu?



Disease & Variants[i]

Involvement in disease[i]

Alport syndrome 1, X-linked (ATS1)

22 Publications

Note The disease is caused by variants affecting the gene represented in this entry

Description A syndrome that is characterized by progressive glomerulonephritis, renal failure, sensorineural deafness, specific eye abnormalities (lenticonous and macular flecks), and glomerular basement membrane defects. The disorder shows considerable heterogeneity in that families differ in the age of end-stage renal disease and the occurrence of deafness.

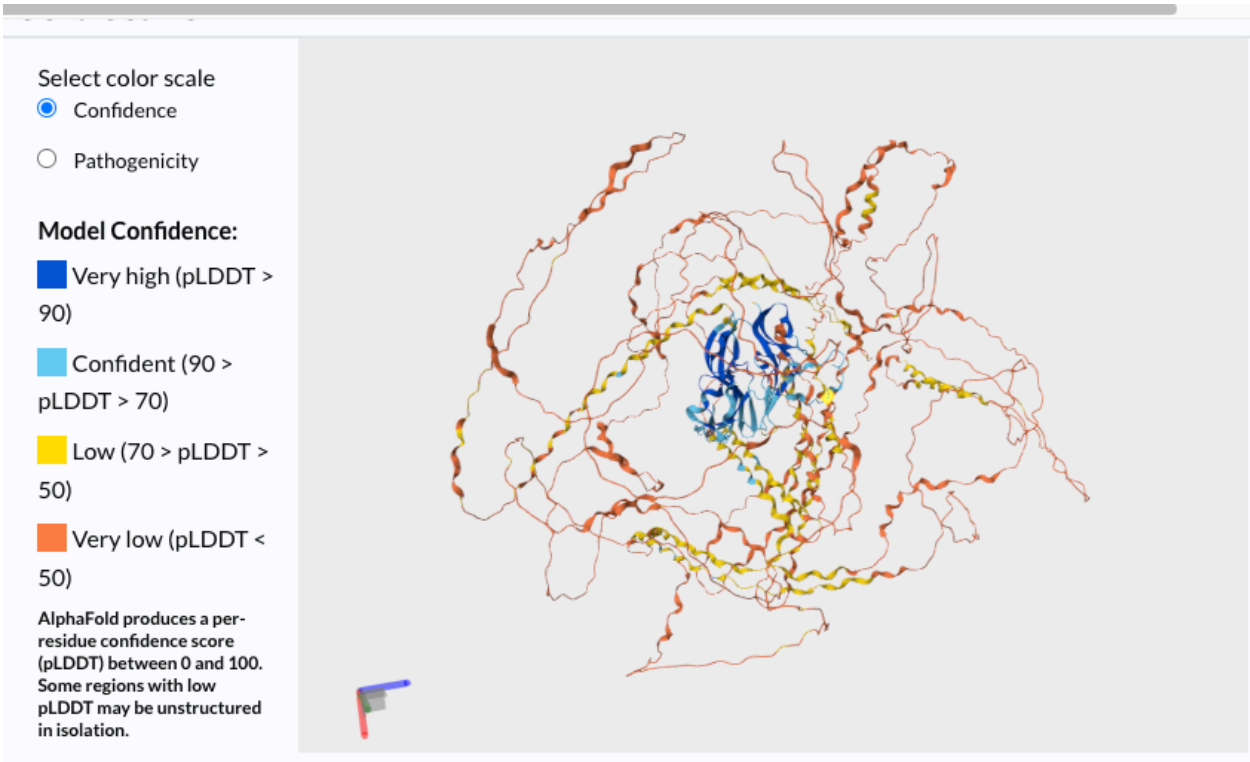
See also MIM:301050

Natural variants in ATS1

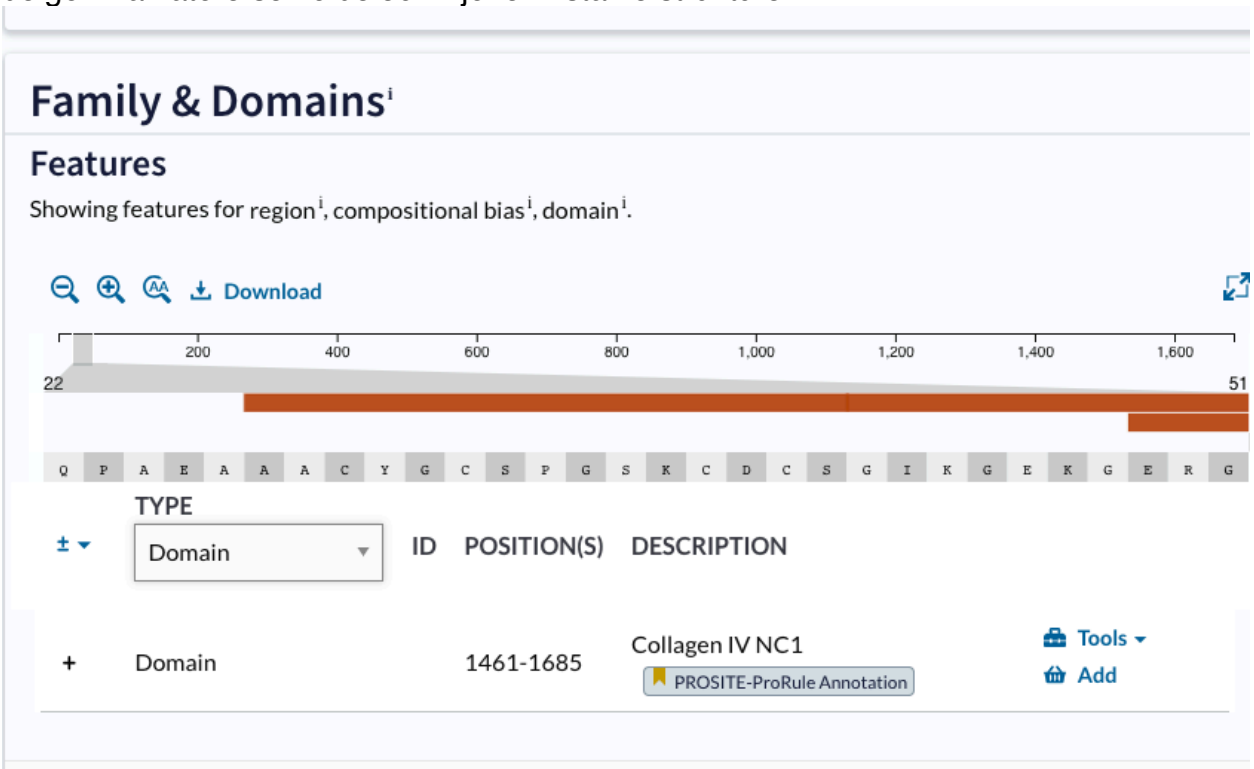
VARIANT ID	POSITION(S)	CHANGE	DESCRIPTION
VAR_001914	54	G>D	in ATS1; adult type; dbSNP:rs104886043
VAR_007991	114	G>S	in ATS1
VAR_071932	123	G>E	in ATS1; dbSNP:rs1569488429 1 Publication
VAR_001915	129	G>E	in ATS1; juvenile type; dbSNP:rs281874723
VAR_001916	129	G>V	in ATS1; juvenile type; dbSNP:rs281874723
VAR_001917	174	G>R	in ATS1; dbSNP:rs104886055 1 Publication
VAR_011220	177	G>C	in ATS1; presenting with dot-and-fleck retinopathy; dbSNP:rs104886056 1 Publication
VAR_001918	177	G>R	in ATS1; adult type; dbSNP:rs104886056 1 Publication
VAR_011221	192	G>R	in ATS1; dbSNP:rs104886060 1 Publication
VAR_011222	204	G>V	in ATS1; juvenile type; dbSNP:rs104886063

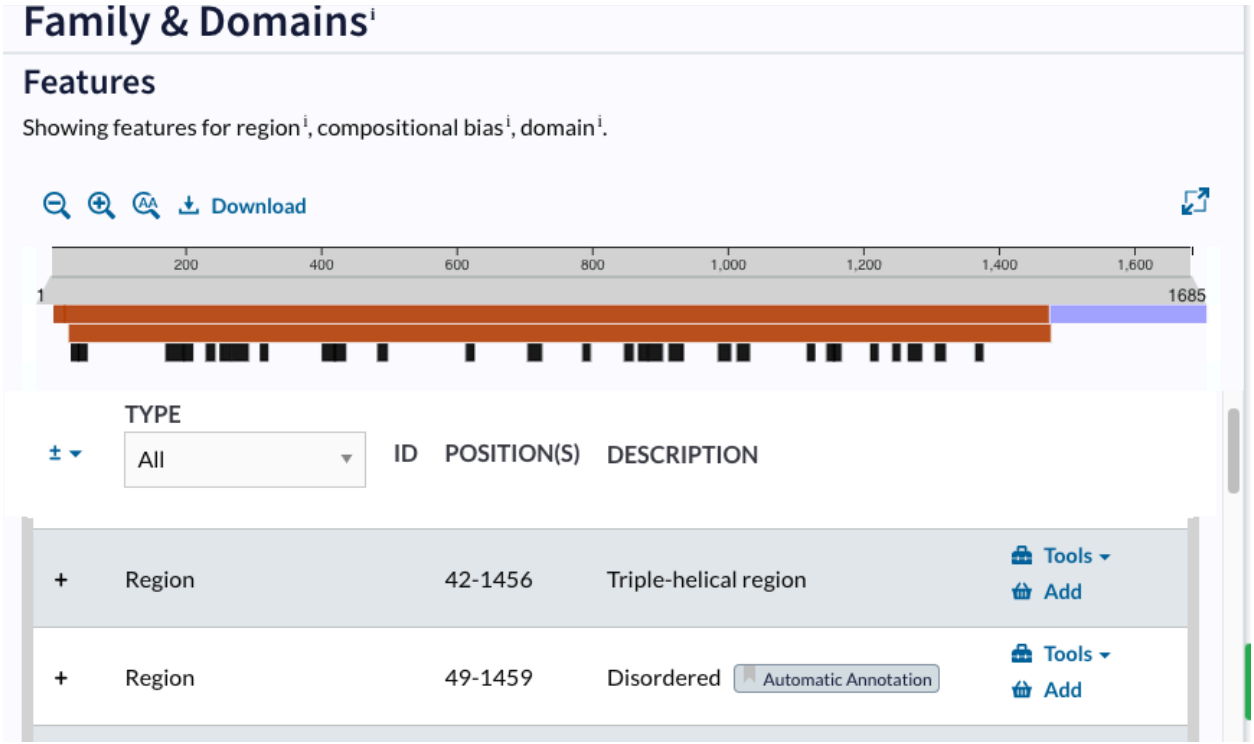
Odgovor: Alportov sindrom 1. Glicin.

3. Ali ima protein po celotni dolžini urejeno strukturo? Utemelji.



4. Poimenuj glavno domeno in glavno regijo. Koliko aminokislinskih ostankov so dolge? Za katero so že določili njeno kristalno strukturo?



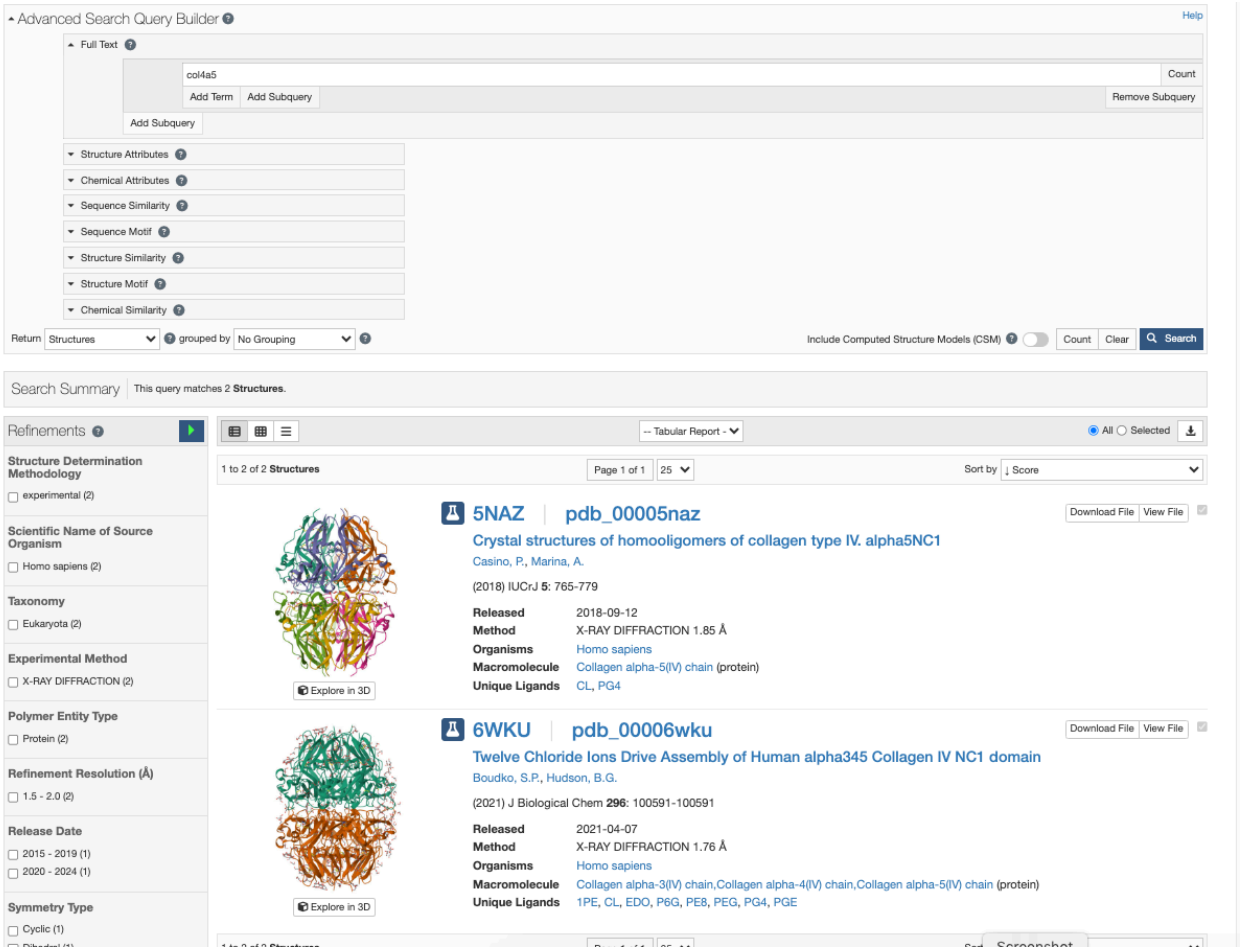


Odgovor: Domena: Collagen IV NC1: (1685 – 1461) + 1 = 225 ak

Regija: Triple-helical region: (1456 – 42) + 1 = 1415 ak

Za domeno:

5. Poišči delno strukturo tega proteina, ki je bila prva določena. Zapiši njen PDB ID.

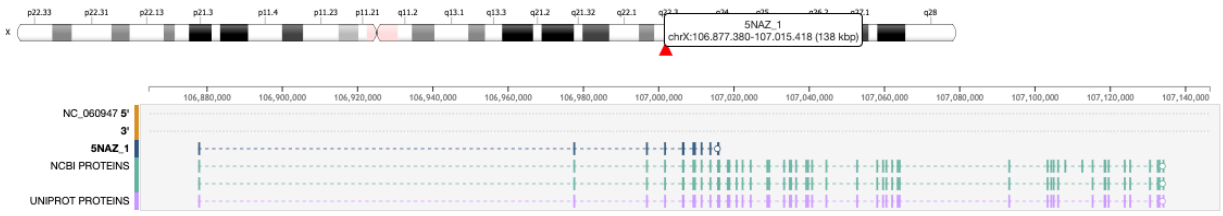


Odgovor: 5NAZ

6. Koliko podenot ima in s katero metodo so potrdili biološko stestavo oligomera?

Odgovor: 6, gelska filtracija

7. Na katerem mestu v človeškem genomu se nahaja zapis za ta protein (kromosom in dolžina gena)?



Odgovor: X kromosom, 138 kbp

8. Katera aminokislinska ostanka sta najpogostejša v tem proteinu (zapiši delež obeh)? Ali je to smiselno? Utemelji.

Number of amino acids: 1685

Molecular weight: 161043.89
Theoretical pI: 7.71

Amino acid composition: CSV format

Ala (A)	47	2.8%
Arg (R)	37	2.2%
Asn (N)	32	1.9%
Asp (D)	54	3.2%
Cys (C)	20	1.2%
Gln (Q)	73	4.3%
Glu (E)	61	3.6%
Gly (G)	478	28.4%
His (H)	13	0.8%
Ile (I)	69	4.1%
Leu (L)	112	6.6%
Lys (K)	80	4.7%
Met (M)	26	1.5%
Phe (F)	41	2.4%
Pro (P)	391	23.2%
Ser (S)	64	3.8%
Thr (T)	38	2.3%
Trp (W)	5	0.3%
Tyr (Y)	13	0.8%
Val (V)	31	1.8%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Odgovor: Glicin – 28,4 % in prolin – 23,2 %.

Post-translational modification[i]

Prolines at the third position of the tripeptide repeating unit (G-X-Y) are hydroxylated in some or all of the chains.

Type IV collagens contain numerous cysteine residues which are involved in inter- and intramolecular disulfide bonding. 12 of these, located in the NC1 domain, are conserved in all known type IV collagens.

The trimeric structure of the NC1 domains is stabilized by covalent bonds between Lys and Met residues. By Similarity

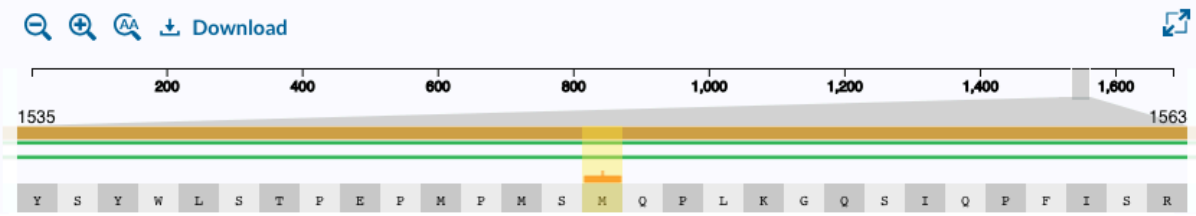
Protein je ena od verig kolagena IV, za katere je značilno, da tvorijo trojne vijačnice. Te vsebujejo predbsem ponovitve glicin-X-Y, kjer sta X in Y najpogosteje prolin in hidroksiprolin.

9. Za stabilizacijo trojne vijačnice, ki jo tvorijo ti proteini, se med njimi tvorijo tudi disulfidni mostički. Koliko jih tvori 1 podenota tega proteina? Na katerih mestih se nahajajo ti aminokislinski ostanki?

PTM/Processing[i]

Features

Showing features for signal[i], chain[i], glycosylation[i], disulfide bond[i], cross-link[i].



±	TYPE: Disulfide bond	ID	POSITION(S)	DESCRIPTION
+	Disulfide bond		451	Interchain PROSITE-ProRule Annotation
+	Disulfide bond		481	Interchain PROSITE-ProRule Annotation
+	Disulfide bond		484	Interchain PROSITE-ProRule Annotation
+	Disulfide bond		1476↔1567	Or C-1476 with C-1564 PROSITE-ProRule Annotation
+	Disulfide bond		1509↔1564	Or C-1509 with C-1567 PROSITE-ProRule Annotation
+	Disulfide bond		1521↔1527	PROSITE-ProRule Annotation
+	Disulfide bond		1586↔1681	Or C-1586 with C-1678 PROSITE-ProRule Annotation
+	Disulfide bond		1620↔1678	Or C-1620 with C-1681 PROSITE-ProRule Annotation
+	Disulfide bond		1632↔1638	PROSITE-ProRule Annotation

Odgovor: 3. Mesta: 451, 461, 484

10. Dve izmed pogostejših mutacij sta označeni z ID: VAR_001923 in VAR_001928. Oceni katera ima hujše posledice za človeka. Utemelji.

Disease & Variants

Involvement in disease

Alport syndrome 1, X-linked (ATS1)

22 Publications

Note The disease is caused by variants affecting the gene represented in this entry

Description A syndrome that is characterized by progressive glomerulonephritis, renal failure, sensorineural deafness, specific eye abnormalities (lenticonous and macular flecks), and glomerular basement membrane defects. The disorder shows considerable heterogeneity in that families differ in the age of end-stage renal disease and the occurrence of deafness.

See also MIM:301050

Natural variants in ATS1

VARIANT ID	POSITION(S)	CHANGE	DESCRIPTION
VAR_011229	319	G>R	in ATS1; juvenile type; dbSNP:rs104886085 1 Publication
VAR_001923	325	G>E	in ATS1; dbSNP:rs104886091 1 Publication
VAR_001924	325	G>R	in ATS1; juvenile and adult types; dbSNP:rs104886088 3 Publications
VAR_007992	331	G>V	in ATS1; dbSNP:rs104886092
VAR_001925	365	G>E	in ATS1; juvenile type; dbSNP:rs104886096
VAR_001926	365-367	*missing*	in ATS1; juvenile type
VAR_001927	371	G>E	in ATS1; juvenile type; dbSNP:rs104886097
VAR_001928	374	G>A	in ATS1; dbSNP:rs104886108
VAR_001929	383	G>D	in ATS1; juvenile type; dbSNP:rs104886105
VAR_001930	400	G>E	in ATS1; adult type; dbSNP:rs104886107 1 Publication
VAR_001931	406	G>V	in ATS1; adult type; dbSNP:rs104886100 1 Publication

Odgovor: Hujša bo mutacija VAR_001923, ker pride do substitucije glicina v glutamat, ki je velika in pozitivno nabita aminikislina, ki bolj zmoti strukturo trojne vijačnice, kot substutucija v alanin v primeru VAR_001928.