

Q1. Project Focus: Human APP

I will be looking into the amyloid beta precursor protein in humans because it has been associated with the buildup of beta-amyloid plaques seen in Alzheimers. The accession number for the A isoform is NP_000475.1.

According to Wikipedia, “**Amyloid-beta precursor protein (APP)** is an integral membrane protein expressed in many tissues and concentrated in the synapses of neurons. It functions as a cell surface receptor^[5] and has been implicated as a regulator of synapse formation,^[6] neural plasticity,^[7] antimicrobial activity,^[8] and iron export”.

Q2. Blastp of Human APP in tsa_nr database

Job Title	NP_000475:amyloid-beta precursor protein isoform...		
RID	43DJHJ7Z013	Search expires on 05-14 14:42 pm	Download All ▾
Program	BLASTP ?	Citation ▾	
Database	tsa_nr	See details ▾	
Query ID	NP_000475.1		
Description	amyloid-beta precursor protein isoform a precursor [Homo ...		
Molecule type	amino acid		
Query Length	770		
Other reports	Distance tree of results	Multiple alignment	MSA viewer ?

Filter Results

Organism

only top 20 will appear

☐ exclude

Type common name, binomial, taxid or group name

+ Add organism

Percent Identity

E value

Query Coverage

to

to

to

Filter

Reset

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	amyloid beta A4 protein isoform a precursor [Macaca mulatta]	Macaca mulatta	1602	1602	100%	0.0	99.48%	770	AFH29688.1
<input checked="" type="checkbox"/>	amyloid beta A4 protein isoform a precursor [Macaca mulatta]	Macaca mulatta	1596	1596	100%	0.0	99.35%	769	AFI37037.1
<input checked="" type="checkbox"/>	amyloid beta A4 protein isoform a precursor [Callithrix jacchus]	Callithrix jacchus	1596	1596	100%	0.0	99.09%	770	JAB11620.1
<input checked="" type="checkbox"/>	TPA: amyloid-beta A4 protein isoform a precursor [Sus scrofa]	Sus scrofa	1575	1575	100%	0.0	97.79%	770	HCZ80958.1
<input checked="" type="checkbox"/>	amyloid beta A4 protein isoform g [Heterocephalus glaber]	Heterocephalus glaber	1563	1563	100%	0.0	97.01%	770	JAO02449.1
<input checked="" type="checkbox"/>	amyloid beta A4 protein [Castor canadensis]	Castor canadensis	1558	1558	100%	0.0	97.14%	769	JAV38683.1
<input checked="" type="checkbox"/>	amyloid beta A4 protein isoform b precursor [Macaca mulatta]	Macaca mulatta	1557	1557	100%	0.0	97.27%	751	AFH29687.1
<input checked="" type="checkbox"/>	Putative conserved secreted mucin [Desmodus rotundus]	Desmodus rotundus	1555	1555	100%	0.0	96.36%	770	JAA48996.1
<input checked="" type="checkbox"/>	amyloid beta A4 protein isoform b precursor [Callithrix jacchus]	Callithrix jacchus	1549	1549	100%	0.0	96.75%	751	JAB11622.1
<input checked="" type="checkbox"/>	amyloid beta A4 protein isoform h precursor [Callithrix jacchus]	Callithrix jacchus	1546	1546	100%	0.0	96.75%	752	JAB11621.1
<input checked="" type="checkbox"/>	TPA: amyloid-beta A4 protein isoform h precursor [Sus scrofa]	Sus scrofa	1528	1528	100%	0.0	95.45%	752	HDB78849.1
<input checked="" type="checkbox"/>	Putative conserved secreted mucin [Desmodus rotundus]	Desmodus rotundus	1509	1509	100%	0.0	94.03%	752	JAA48941.1
<input checked="" type="checkbox"/>	amyloid beta A4 protein isoform i precursor [Callithrix jacchus]	Callithrix jacchus	1500	1500	100%	0.0	94.42%	733	JAB40955.1
<input checked="" type="checkbox"/>	TPA: amyloid-beta A4 protein isoform i precursor [Sus scrofa]	Sus scrofa	1485	1485	100%	0.0	93.25%	733	HDC75628.1
<input checked="" type="checkbox"/>	TPA: amyloid-beta A4 protein isoform a precursor [Sus scrofa]	Sus scrofa	1413	1413	94%	0.0	94.09%	705	HDA60619.1
<input checked="" type="checkbox"/>	TPA: amyloid-beta A4 protein isoform c precursor [Sus scrofa]	Sus scrofa	1404	1404	100%	0.0	89.74%	714	HDC52635.1
<input checked="" type="checkbox"/>	TPA: amyloid-beta A4 protein isoform h precursor [Sus scrofa]	Sus scrofa	1365	1365	94%	0.0	91.62%	687	HDA60620.1
<input checked="" type="checkbox"/>	hypothetical protein [Micurus spixii]	Micurus spixii	1344	1344	100%	0.0	87.45%	753	LAB26479.1
<input checked="" type="checkbox"/>	protein AMBP [Cerrophidion petalcalensis]	Cerrophidion petalcalensis	1334	1334	100%	0.0	86.80%	753	MDQ4152784.1
<input checked="" type="checkbox"/>	amyloid beta A4 protein-like protein [Agkistrodon contortrix contortrix]	Agkistrodon contortrix contortrix	1333	1333	100%	0.0	86.69%	750	JAV51273.1
<input checked="" type="checkbox"/>	hypothetical protein [Cerrophidion tzotzilorum]	Cerrophidion tzotzilorum	1332	1332	100%	0.0	87.05%	752	MDQ4159560.1
<input checked="" type="checkbox"/>	amyloid beta A4 protein-like protein [Micurus fulvius]	Micurus fulvius	1332	1332	98%	0.0	87.27%	750	JAI10297.1
<input checked="" type="checkbox"/>	amyloid beta A4 protein-like protein [Boiga irregularis]	Boiga irregularis	1332	1332	100%	0.0	86.42%	749	JAG68744.1
<input checked="" type="checkbox"/>	Amyloid beta A4 protein-like [Crotalus adamanteus]	Crotalus adamanteus	1328	1328	100%	0.0	86.30%	750	AFJ49397.1
<input checked="" type="checkbox"/>	amyloid-beta precursor protein precursor [Falco peregrinus]	Falco peregrinus	1309	1309	97%	0.0	91.49%	751	LAH05057.1
<input checked="" type="checkbox"/>	amyloid-beta precursor protein precursor [Ciconia boyciana]	Ciconia boyciana	1308	1308	97%	0.0	91.76%	751	LAG58143.1
<input checked="" type="checkbox"/>	amyloid-beta precursor protein precursor [Antigone vipio]	Antigone vipio	1306	1306	97%	0.0	91.49%	751	LAH73172.1
<input checked="" type="checkbox"/>	amyloid-beta precursor protein precursor [Grus japonensis]	Grus japonensis	1305	1305	97%	0.0	91.36%	751	LAF65940.1
<input checked="" type="checkbox"/>	amyloid-beta precursor protein precursor [Accipiter gentilis]	Accipiter gentilis	1305	1305	97%	0.0	91.09%	751	LAI61108.1
<input checked="" type="checkbox"/>	protein AMBP [Cerrophidion godmani]	Cerrophidion godmani	1283	1283	100%	0.0	84.72%	734	MDQ4165573.1
<input checked="" type="checkbox"/>	amyloid-beta precursor protein precursor [Ciconia boyciana]	Ciconia boyciana	1261	1261	97%	0.0	89.36%	733	LAG58135.1
<input checked="" type="checkbox"/>	amyloid-beta precursor protein precursor [Accipiter gentilis]	Accipiter gentilis	1257	1257	97%	0.0	88.70%	733	LAI61116.1
<input checked="" type="checkbox"/>	TPA: amyloid-beta A4 protein isoform i precursor [Sus scrofa]	Sus scrofa	1255	1255	82%	0.0	94.65%	643	HDA89464.1
<input checked="" type="checkbox"/>	TPA: amyloid-beta A4 protein isoform i precursor [Sus scrofa]	Sus scrofa	1254	1254	83%	0.0	94.07%	632	HDA89465.1
<input checked="" type="checkbox"/>	TPA: amyloid-beta A4 protein isoform a precursor [Sus scrofa]	Sus scrofa	1227	1227	82%	0.0	93.27%	618	HDA60617.1
<input checked="" type="checkbox"/>	amyloid beta A4 protein-like protein [Callorhinchus milii]	Callorhinchus milii	1204	1204	100%	0.0	76.51%	753	AFO96226.1
<input checked="" type="checkbox"/>	TPA: amyloid-beta A4 protein isoform h precursor [Sus scrofa]	Sus scrofa	1179	1179	83%	0.0	90.47%	642	HDA60616.1
<input checked="" type="checkbox"/>	TPA: amyloid-beta A4 protein isoform b precursor [Sus scrofa]	Sus scrofa	1162	1162	76%	0.0	95.61%	591	HDB03229.1

Genomic clone






Q3. Novel Gene Info

```
>AF096226.1 amyloid beta A4 protein-like protein [Callorhinchus milii]
MFLQPAARLLLLLSAIGTVTVSPSALTAEVPTDGGTGLLAAEPQIAMFCDKLNMHMNVQTGKWISDPSGT
NSCFGTKEGILQYCQEVYDDLQITNVVEANQPVTIQNWCKKGRKQCKGHPHIVIPIYRCLVGEFVSDALLV
PDKCKFLHREKMDTCENHLYWHTVAKEACGDKTMNLKDYGMLLPCGIDQFRGVEFVCCPISEENEKVDVS
DVDEEDSYIWWGGADA EYSDGGDKVIEEPMDEEEV DIEDDDDEEEDEDEDDEDDQYEDPTEHTTSSTTT
TTEAIEEVVREVCSEQAETGPCRAMIPRWYFDIGEGKCAQFIYGGCGGNRNNFDSE EYCLSVCSVIPAT
AASTPDAIDQYLETPGDENEHAYFQKAKERQEARHRERMSQIMREWEEAERQARNLHKADKKAVIQRFQE
MVESLEQEAASERQQLVETHMARVEAMLNDRRRRIALENYLAALQADPPRPRHVLNALKKYIRAEQKDRQH
TLKHFEHVRMVDPKKAAQIKSQVMTHLRVIDERMNQSLSLLYKVPAAVEEIQDEVDELLQKEQSYIDDMV
ANSVSDTRVSYGNDALVPSLSETKTTIELLPDQGEFTLDDLQPLHPFVVDSPVNTENEVEPVDPARPAAD
RGLTTRPGSGLTGIKTQEIAEVKMETKFRQDSGYEVHHQKL VFFPEDVGSNKGAIIGLMVGGVVIATVIV
ITLVMLKKKQYTSIHGVI EVDAAVTPEERHLSKMQQNGYENPTYKFFEQMQN
```

This is the FASTA sequence for the amyloid beta A4 protein-like protein from *Callorhinchus milii*, the Australian Ghost Shark

Q4: Proving Novelty

The output below shows that this protein is indeed novel as the top hit does not match 100%.

Job Title	AFO96226:amyloid beta A4 protein-like protein...		
RID	43DXX8JV013	Search expires on 05-14 14:48 pm	Download All 
Program	BLASTP 	Citation 	
Database	nr	See details 	
Query ID	AFO96226.1		
Description	amyloid beta A4 protein-like protein [Callorhinchus milii]		
Molecule type	amino acid		
Query Length	753		
Other reports	Distance tree of results	Multiple alignment	MSA viewer 

Filter Results

Organism

only top 20 will appear

☐ exclude

Type common name, binomial, taxid or group name

+ Add organism

Percent Identity

E value

Query Coverage

to

to

to

Filter

Reset



Descriptions	Graphic Summary	Alignments	Taxonomy					
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
amyloid beta (A4) precursor protein a isoform X4 [Callorhinchus milii]	Callorhinchus milii	1479	1479	93%	0.0	99.86%	707	XP_007901748.1
amyloid beta (A4) precursor protein a isoform X1 [Callorhinchus milii]	Callorhinchus milii	1431	1431	93%	0.0	97.31%	689	XP_007901746.1
amyloid beta (A4) precursor protein a isoform X1 [Rhincodon typus]	Rhincodon typus	1184	1184	88%	0.0	90.27%	666	XP_048458906.1
amyloid beta (A4) precursor protein a isoform X1 [Mobula hypostoma]	Mobula hypostoma	1295	1295	96%	0.0	89.96%	751	XP_062906350.1
amyloid-beta A4 protein-like isoform X1 [Carcharodon carcharias]	Carcharodon carcharias	1353	1353	99%	0.0	89.95%	749	XP_041067979.1
amyloid-beta precursor protein-like isoform X1 [Pristis pectinata]	Pristis pectinata	1314	1314	96%	0.0	89.68%	751	XP_051869955.1
hypothetical protein [Chiloscyllium punctatum]	Chiloscyllium punctatum	1324	1324	97%	0.0	89.34%	748	GCC35796.1
amyloid-beta precursor protein-like isoform X1 [Hypanus sabinus]	Hypanus sabinus	1307	1307	96%	0.0	89.16%	753	XP_059824912.1
amyloid-beta A4 protein-like isoform X1 [Scyliorhinus canicula]	Scyliorhinus canicula	1365	1365	99%	0.0	89.14%	742	XP_038657332.1
amyloid-beta precursor protein-like isoform X1 [Stegostoma tigrinum]	Stegostoma tigrinum	1306	1306	99%	0.0	89.14%	747	XP_048396842.1
amyloid beta (A4) precursor protein a isoform X1 [Chiloscyllium plagiosum]	Chiloscyllium plagiosum	1337	1337	99%	0.0	88.87%	748	XP_043557332.1
amyloid-beta precursor protein-like isoform X3 [Leucoraja erinacea]	Leucoraja erinacea	1285	1285	96%	0.0	88.75%	754	XP_055501201.1
amyloid-beta A4 protein-like isoform X2 [Scyliorhinus canicula]	Scyliorhinus canicula	1352	1352	99%	0.0	88.74%	739	XP_038657333.1
amyloid-beta precursor protein isoform X3 [Amblyraja radiata]	Amblyraja radiata	1281	1281	96%	0.0	88.63%	754	XP_032888281.1
amyloid beta (A4) precursor protein a isoform X1 [Hemiscyllium ocellatum]	Hemiscyllium ocellatum	1333	1333	99%	0.0	88.25%	748	XP_060689823.1
amyloid-beta A4 protein-like isoform X2 [Carcharodon carcharias]	Carcharodon carcharias	1306	1306	99%	0.0	87.67%	731	XP_041067980.1
amyloid beta (A4) precursor protein a isoform X2 [Mobula hypostoma]	Mobula hypostoma	1249	1249	96%	0.0	87.62%	733	XP_062906351.1
amyloid-beta precursor protein-like isoform X2 [Pristis pectinata]	Pristis pectinata	1268	1268	96%	0.0	87.35%	733	XP_051869956.1
hypothetical protein Y1Q_0003230 [Alligator mississippiensis]	Alligator mississippiensis	842	842	60%	0.0	87.34%	743	KYO22727.1
amyloid-beta precursor protein isoform X1 [Dermochelys coriacea]	Dermochelys coriacea	842	842	60%	0.0	87.34%	752	XP_038244078.1
amyloid-beta precursor protein isoform X2 [Alligator mississippiensis]	Alligator mississippiensis	841	841	60%	0.0	87.34%	756	XP_006264186.2
amyloid-beta precursor protein isoform X1 [Caretta caretta]	Caretta caretta	841	841	60%	0.0	87.34%	752	XP_048724019.1

Q5. Multiple Sequence Alignment

Below is the MSA for the amyloid beta A4 protein-like protein sequences of 7 different species

CLUSTAL multiple sequence alignment by MUSCLE (3.8)
Link: muscle-I20240529-195849-0477-60507466-p1m

Human
MLPGLA--LLLL-----AAWTARALEVPTDGNAGLL-AEPQIAMFCGRLNMHMINVQN

EasternCoralSnake
MLPHLA--LLLL-----GASWTAGALEVPTDGNAGLL-AEPQVAMFCGKSNMHMNVQN
Copperhead
MLPHLA--FLLL-----VASWTAGALEVPTDGNAGLL-AEPQVAMFCDKLNMHMNVQN
BrownTreeSnake
MLPHLA--FLLL-----GASWTAGALEVPTDGNAGLL-AEPQVAMFCGKSNMHMNVQN
AustralianGhostshark
MFLQPAARLLLLSAIGTVTVSPSALTAEVPTDGGTGLLAAEPQIAMFCDKLNMHMNVQT
GreatWhiteShark
MIPHLPGRLRML-LVAAVIVLPSLCQALEVPTDGGTGLLAAEPQIAMFCGKLNMHMNVQS
SmallSpottedCatshark
MIPHLPCRLRML--LAAAIVLSSLCRALEVPTDGGTGMLAAEPQIAMFCGKLNMHMNVQS
: . : : . *****.:*: *****:*****. *****.

Human
GKWSDPSGKTCTCIDTKEGILQYCQEVYPELQITNVVEANQPVTIQNWCKRGRKQCKTHP
EasternCoralSnake
GKWESDPSGKTSCITTKEGILQYCQQVYPELQITNVVEANQPVTIQNWCKQGRKQCRSHP
Copperhead
GKWESDPSGKTSCIATKEGILQYCQQVYPELQITNVVEANQPVTIENWCKQGRKQCRSHP
BrownTreeSnake
GKWESDPSGKTSCIATKEGILQYCQQVYPELQITNVVEANQPVTIQNWCKQGRKQCRSHP
AustralianGhostshark
GKWISDPSGTNSCFGTKEGILQYCQEVYPDLQITNVVEANQPVTIQNWCKKGRKQCKGHP
GreatWhiteShark
GKWVSDPSGKTSCFGTKEGILQYCQEVYPDLQITNVVEANQPVTIQNWCKKGRKQCKGHP
SmallSpottedCatshark
GKWVSDPSGTNSCFGTKEGILQYCQEVYPDLQITNVVEANQPVTIQNWCKKGRKQCKGHP
*** *****.:*: *****:***:*****:*****:*****.*****. **

Human
HFVIPYRCLVGEFVSDALLVPDKCKFLHQERMDVCETHLHWHTVAKETCSEKSTNLHDYG
EasternCoralSnake
YIVVPYRCLVGEFVSDALLVPDKCKFLHQERMDICETHLHWHTVAKESCSEKGMNLHDYG
Copperhead
YIVVPYRCLVGEFVSDALLVPDKCKFLHQERMDICETHLHWHTVAKEFCSEKGMNLHDYG
BrownTreeSnake
YIVVPYRCLVGEFVSDALLVPDKCKFLHQERMDICETHLHWHTVAKESCSEKGMNLHDYG
AustralianGhostshark
HIVIPYRCLVGEFVSDALLVPDKCKFLHREKMDTCENHLYWHTVAKEACGDKTMNLKDYG
GreatWhiteShark
HIVVPYRCLVGEFVSDALLVPDKCKFLHREKVDTCESHLWHTVAKETCGDKIMNLHDYG
SmallSpottedCatshark
HIVVPYRCLVGEFVSDALLVPDKCKFLHREKMDTCESHLWHTVAKETCGDKIMNLHDYG
.:*:*****:*****.*: * *.*:***** *: * **:

Human

MLLPCGIDKFRGVEFVCCPLAEESDNVDSADAEEDDSVWVGADTDYADGSEDKVVEVA
EasternCoralSnake
MLLPCGIDKFRGVEFVCCLVDESNDVDSAEAEEDDSVWVGADADYADGSYDKVTEEQ
Copperhead
MLLPCGIDKFRGVEFVCCLVADESDNTDSTAEEDDSVWVGADADYADGSYDKVTEEQ
BrownTreeSnake
MLLPCGIDKFCGVEFVCCLVADESDNTDSAEAEEDDSVWVGADADYADGSYDKVTEEQ
AustralianGhostshark
MLLPCGIDQFRGVEFVCCPISEENEKVDSVDVDEEDSYIWWGADA EYSDGG-DKVIEEE
GreatWhiteShark
MLLPCGIDEFRGVEFVCCPIPEENDKIDSDMDEEDSDVWVGDDADYADGS-DKVIEEQ
SmallSpottedCatshark
MLLPCGIDEFRGVEFVCCPIPEENDKIDS-DMDE-DSDVWVGDDADYADGS-DKIIEEQ
*****:* ***** : :*.:: *. : :* ** :**** *:*:**.* **: *

Human

EEE-EVAEVEEEEEADDDDEDEDGDEVEEEAE EPYEEATERTTSIATTTTTTTTESVEEVVR
EasternCoralSnake
LAEGDTT DVEDENTDDDDDDDE--DGXEVTEDQYQEAIERTTTSIA-TTTTTTESVEEVVR
Copperhead
LAEGDTTDIEDENTDDDDDEDD--EAEVTEDEQYQEATERTTTSIA-TTTTTTESVEEVVR
BrownTreeSnake
LAEGDTT DVEDENTDDDDXDXD--EAEVTEDEQYQEATERTTTSIA-TTTTTTESVEEVVR
AustralianGhostshark
PMD-----EEEEVDIEDDDDEEDEDDEDDQYEDPTEHTTS---STTTTTEAIEEVVR
GreatWhiteShark
PME-----EESDIDDEDDDDL--DGEEDQYEDPTEHTTS---STTTTTEAIEEVVR
SmallSpottedCatshark
PME-----EESDIDDDLGDG-----DEDQYEDPTEHTTS---STTTTTEAIEEVVR
: *::: * : :: *::. *.*** :*****:*****

Human

EVCSEAETGPCRAMISRWFYDVTEGKCAPFFYGGCGNRNNFDTEEYCMVCGSAMSQS
EasternCoralSnake
EVCSEAETGPCRAMISRWFYDVTEGKCAPFFYGGCGNRNNFDTEEYCMVCGSV----
Copperhead
EVCSEAETGPCRAMISRWFYDVTEGKCAPFFYGGCGNRNNFDTEEYCMVCGSV----
BrownTreeSnake
EVCSEAETGPCRAMISRWFYDVTEGKCAPFFYGGCGNRNNFDTEEYCMVCGSV----
AustralianGhostshark
EVCSEAETGPCRAMIPRWYFDIGEGKCAQFIYGGCGNRNNFDSE EYCLSVCSSV----
GreatWhiteShark
EVCSEAETGPCQAMIPRWYFDISEGKCAQFIYGGCGNRNNFASEEYCLAVCNSV----
SmallSpottedCatshark
EVCSEAETGPCQAMIPRWYFDISEGKCAQFIYGGCGNRNNFASEEYCLAVCNSV----
*****.***.*****: ***** *:***** :*****:***.*.

Human

LLKTTQEPLARDPVKLPPTTAASTPDAVDKYLETPGDENEHAHFQKAKERLEAKHRERMSQ

EasternCoralSnake

-----IPTTAASTPDAVDKYLETPGDENEHSHFQKAKERLEAKHRERMSQ

Copperhead

-----IPTTAASTPDAVDRYLETTPGDENEHSHFQKAKERLEAKHRERMSQ

BrownTreeSnake

-----IPTTAASTPDAVDKYLETPGDENEHSHFQKAKERLEAKHRERMSQ

AustralianGhostshark

-----IPATAASTPDAIDQYLETTPGDENEHAYFQKAKERQEARHRERMSQ

GreatWhiteShark

-----IPTTAASTPDAVDKYLETPGDENEHAYFQKAKERLEAKHRERMSK

SmallSpottedCatshark

-----LPTTAASTPDAVDKYLETPGDENEHAYFQKAKERLEAKHRERMSK

:*:*****:*.*****:;***** **.******:

Human

VMREWEAAERQAKNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDR

EasternCoralSnake

VMREWEAAEHQAKNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDR

Copperhead

VMREWEAAEHQAKNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDR

BrownTreeSnake

VMREWEAAEHQAKNLPKADKKAVIQHFQEKVESLEQEAAANERQQLVETHMARVEAMLNDR

AustralianGhostshark

IMREWEAAERQARNLHKADKKAVIQRFQEMVESLEQEAAASERQQLVETHMARVEAMLNDR

GreatWhiteShark

IMREWEAAERQAKNLPKADKKAVIQRFQEMVESLEQEAAASERQQLVETHMARVEAMLNDR

SmallSpottedCatshark

IMREWEAAERQAKNLPKADKKAVIQRFQEMVESLEQEAAASERQQLVETHMARVEAMLNDR

:*****.*** ** *****.*** *****.*****

Human

RRLALENYITALQAVPPRPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRS

EasternCoralSnake

RRVALENYITALQAIPPKPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRS

Copperhead

RRVALENYITALQAVPPKPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRS

BrownTreeSnake

RRVALENYITALQAVPPKPRHVFNMLKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIRS

AustralianGhostshark

RRIALENYLAALQADPPRPRHVLNALKKYIRAEQKDRQHTLKHFEHVRMVDPKKAAQIKS

GreatWhiteShark

RRIALENYLAALQADPPRPRHVLNALKKYIRAEQKDRQHTLKHFEHVRMVDPKKAAQIKS

SmallSpottedCatshark

RRIALENYLAALQADPPRPRHVLNALKKYVRAEQKDRQHTLKHFEHVRMVDPKKAAQIKS

.***.:***** **.*****:* *****:*****.*****.*

Human

QVMTHLRVIYERMNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLANMISEPRISY

EasternCoralSnake

QVMTHLRVTYERMNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLANMISEPRISY

Copperhead

QVMTHLRVTYERMNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLANMISEPRISY

BrownTreeSnake

QVMTHLRVTYERMNQSLSLLYNVPAAVEEIQDEVDELLQKEQNYSDVLANMISEPRISY

AustralianGhostshark

QVMTHLRVIDERMNQSLSLLYKVPAAVEEIQDEVDELLQKEQSYIDDMVANSVSDTRVSY

GreatWhiteShark

QVMTHLRVIDERMNQSLSLLYKVPAAVEEIQDEVDELLQKEQSYIDDMVANSVSDTRISY

SmallSpottedCatshark

QVMTHLRVIDERMNQSLSLLYKVPAAVEEIQDEVDELLQKEQSYIDDMVANSVSDTRISY

***** *****:*.*****.* **::* :*:*:*

Human

GNDALMPSLTETKTTVELLPVNGEFLDDLQPWHSFGADSVANTENEVEPVDPAPADR

EasternCoralSnake

GNDALMPSLTETKTTVERLPVDGEFLDDLQPWHPFAVDSVPANTENEVEPVDPAPADR

Copperhead

GNDALMPSLTETKTTVEHLPVDGEFLDDLQPWHPFAVDSVPANTENEVEPVDPAPADR

BrownTreeSnake

GNDALMPSLTETK-TVEHLPVDGEFLDDLQPWHPFAVDSVPANTENEVEPVDPAPADR

AustralianGhostshark

GNDALVPSLSETKTTIELLPDQGEFTLDDLQPLHPFVVDSSIPANTENEVEPVDPAPADR

GreatWhiteShark

GNDALVPSLSETKTTIELLPDDGEFNLDDLQPLHPFVIDSSIPANTENEVEPVDPAPADR

SmallSpottedCatshark

GNDALVPSLSETKTTIELLPDDGEFNLDDLQPLHPFVIDSSIPANTENEVEPVDPAPADR

*****:***:*** *: * ** :***.***** *. * **:*.*****.*

Human

GLTTRPGSGLTNIKTTEEISEVKMDAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVG

EasternCoralSnake

GLTTRPGSGLTNVKTEETSELKMDAEYRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVG

Copperhead

GLTTRPGSGLTNVKTEETSELKMDAEYRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVG

BrownTreeSnake

GLTTRPGSGLTNVKTEETSELKMDAEYRHDSGFVHHQKLVFFAEDVGSNKGAIIGLMVG

AustralianGhostshark

GLTTRPGSGLTGIKTQEI AEVKMETKFRQDSGYEVHHQKLVFFPEDVGSNKGAIIGLMVG

GreatWhiteShark

GLTTRPGSGLTGIKTEGIAEVKMETEFRQDSGYEVHHQKLVFFPEDVGSNKGAIIGLMVG

SmallSpottedCatshark
GLTTRPGSGLTGIKTEEIAEVKMETEFRQDSGYEVHHQKLVFFPEDVGSNKGAIIGLMVG
*****.:**: :*:**::*:**:*:*****.*****

Human
GVVIATVIVITLVMLKKKQYTSIIHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQM
EasternCoralSnake
GVVIATVIVITLVMLKKKQYTSIIHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQM
Copperhead
GVVIATVIVITLVMLKKKQYTSIIHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQM
BrownTreeSnake
GVVIATVIVITLVMLKKKQYTSIIHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQM
AustralianGhostshark
GVVIATVIVITLVMLKKKQYTSIIHGVIEVDAAVTPEERHLSKMQQNGYENPTYKFFEQM
GreatWhiteShark
GVVIATVIVITLVMLKKKQYTSIIHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQM
SmallSpottedCatshark
GVVIATVIVITLVMLKKKQYTSIIHGVVEVDAAVTPEERHLSKMQQNGYENPTYKFFEQM

*****.:*****

Human	QN
EasternCoralSnake	QN
Copperhead	QN
BrownTreeSnake	QN
AustralianGhostshark	QN
GreatWhiteShark	QN
SmallSpottedCatshark	QN
	**

Q6. Phylogenetic Tree

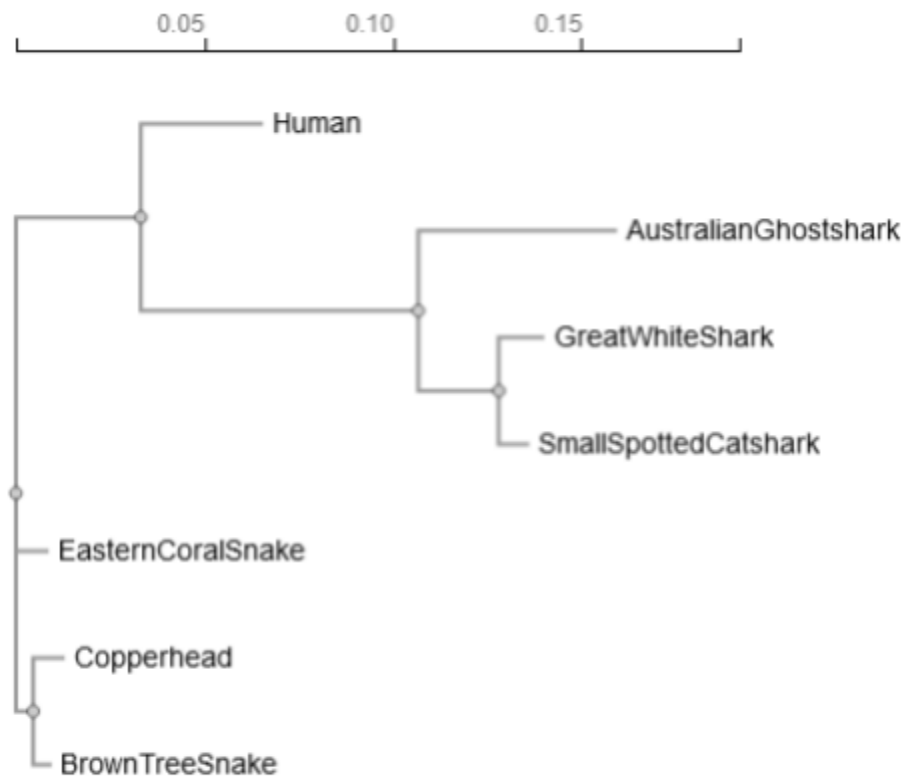


Fig.1 Phylogenetic tree for the MSA in Q5

Q7. Heatmap

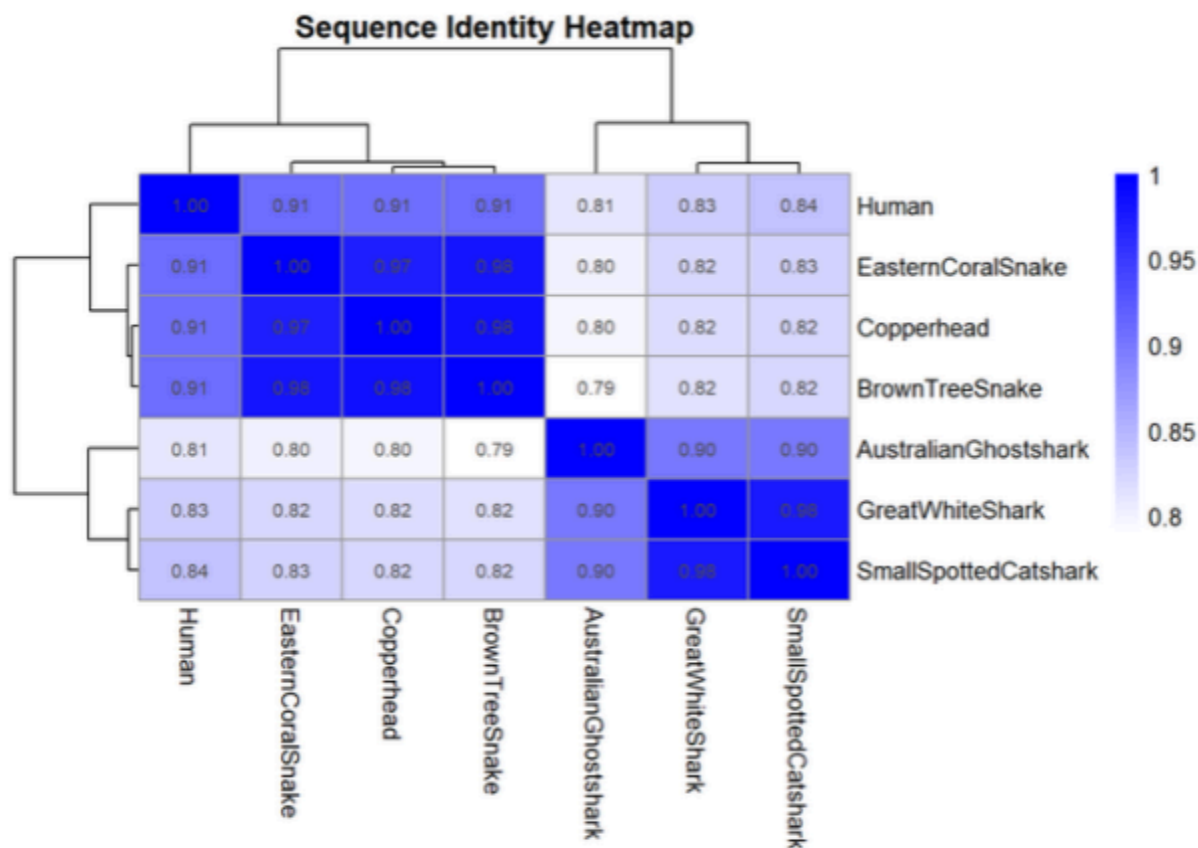


Fig 2. Sequence Identity Heatmap generated from the MSA in Q5

The heatmap shows 2 clusters: humans with snakes and sharks with each other

Q8. Blast PDB search of Consensus Sequence (EasternCoralSnake)

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/>	Chain A_Amyloid-beta precursor protein [Homo sapiens]	Homo sapiens	911	911	61%	0.0	92.45%	770	8OTF_A
<input checked="" type="checkbox"/>	Chain A_Amyloid beta A4 protein [Homo sapiens]	Homo sapiens	683	683	45%	0.0	95.60%	342	5BUO_A
<input checked="" type="checkbox"/>	Chain B_Kunitz Protease Inhibitory Domain of Protease Nexin II [Homo sapiens]	Homo sapiens	134	134	7%	2e-37	98.25%	57	1ZJD_B

Chain A. Amyloid-beta precursor protein (Homo sapiens)

Ab typell filament from Guam ALS/PDC

PDB: 8OTF

Source: Homo sapiens

Method: Electron Microscopy

Resolution: 3.3 Å

E-value = 0.0

Seq Identity = 92.45%

Chain A. Amyloid-beta A4 protein (Homo sapiens)

[A receptor molecule](#)

PDB: 5BUO

Source: Homo sapiens

Method: X-ray Diffraction

Resolution: 2.31 Å

E-value = 0.0

Seq Identity = 95.60%

Chain B. Kunitz Protease Inhibitory Domain of Protease Nexin II (Homo sapiens)

[Crystal Structure of the Catalytic Domain of Coagulation Factor XI in Complex with Kunitz
Protease Inhibitor Domain of Protease Nexin II](#)

PDB: 1ZJD

Source: Homo sapiens

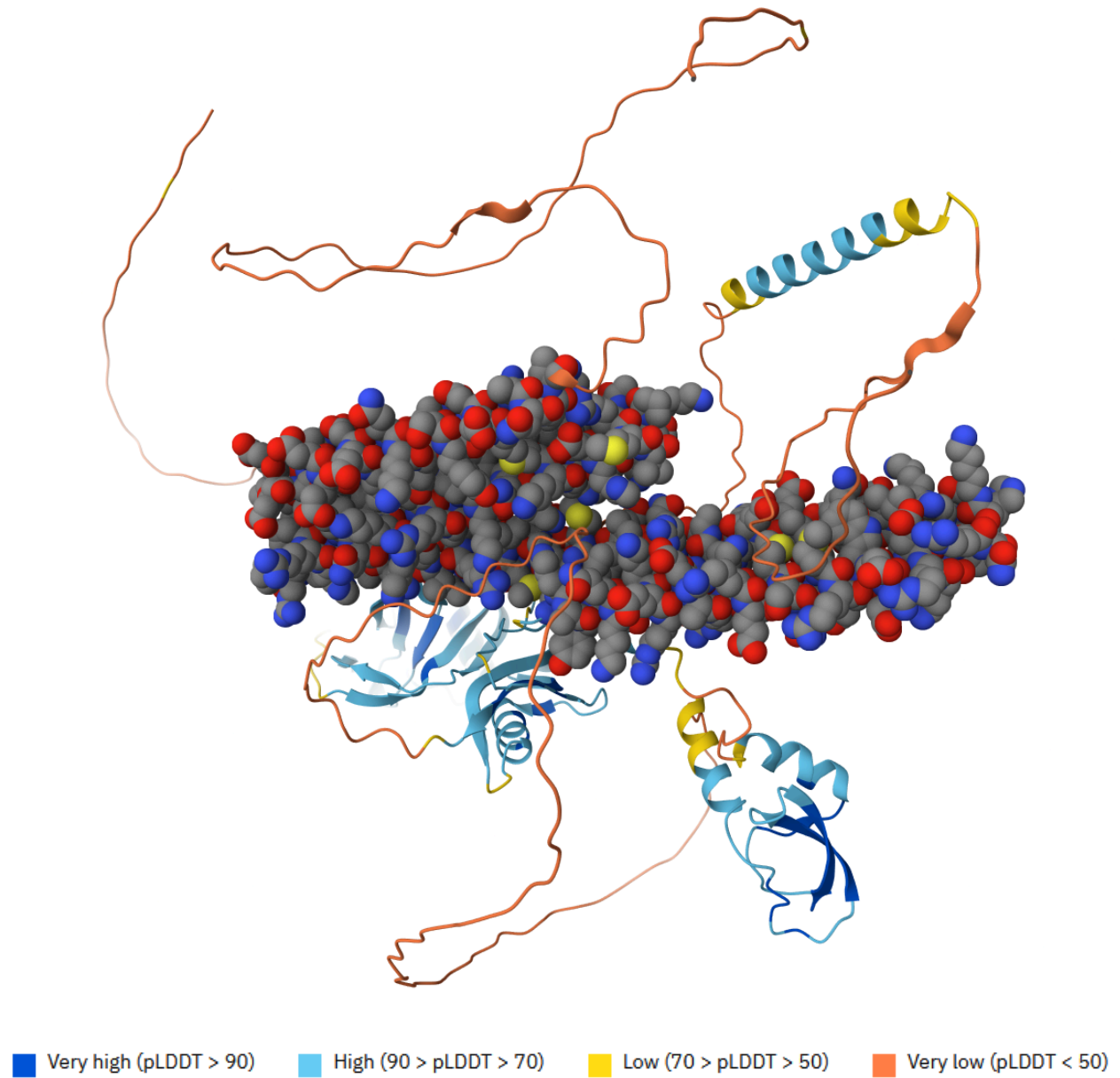
Method: X-ray Diffraction

Resolution: 2.6 Å

E-value = 2e-37

Seq Identity = 98.25%

Q9. AlphaFold of Novel Protein



Q10. ChEMBL Search for Novel Protein

<input type="checkbox"/>	#	E-Value	Positives %	Identities %	Score (bits)	Score	Length	ChEMBL ID	Name
<input type="checkbox"/>	1.	0	86.5	77.6	1226.08	3171	770	CHEMBL2487	Amyloid-beta protein
<input type="checkbox"/>	2.	0	87	77.8	1224.54	3167	770	CHEMBL3638365	Amyloid-beta protein
<input type="checkbox"/>	3.	0	86.9	77.8	1223.38	3164	770	CHEMBL4523942	Amyloid-beta protein
<input type="checkbox"/>	4.	6.2e-12	63.6	37.7	66.6254	161	304	CHEMBL3713062	Tissue fatty acid pathway inhibitor
<input type="checkbox"/>	5.	3.6e-11	61.4	41.4	64.3142	155	306	CHEMBL4523142	Tissue fatty acid pathway inhibitor

Target associated assays and ligand efficiency data are non listed as of 6/3/2024. The only information listed is that this protein is mostly found in eukaryotic cells in mammalia like primates and rodents. Most of the targets are unclassified, but there are many classified as enzymes and epigenetic regulators.

ChEMBL Job ID:

BIOLOGICAL_SEQUENCE_SEARCH-pqZ2iRXxWDD58C1DxSEAaLyAa-Pv9pRprQDJrBscVQw=