



JOHANNES GUTENBERG  
UNIVERSITÄT MAINZ

# Protein domains

Miguel Andrade

Faculty of Biology,

Institute of Organismic Molecular Evolution,

Johannes Gutenberg University

Mainz, Germany

[andrade@uni-mainz.de](mailto:andrade@uni-mainz.de)

# Introduction

Protein domains are structural units (average 160 aa) that share:

Function

Folding

Evolution

Proteins normally are multidomain (average 300 aa)

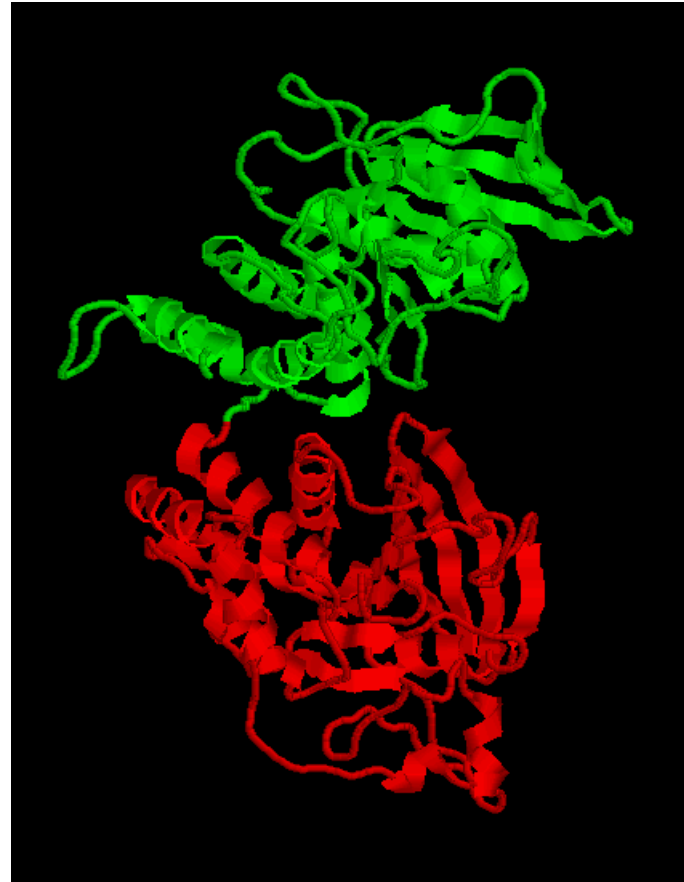


# Introduction

Protein domains are structural units (average 160 aa) that share:

Function  
Folding  
Evolution

Proteins normally are multidomain (average 300 aa)



# Domains

## Why to search for domains:

Protein structural determination methods such as X-ray crystallography and NMR have size limitations that limit their use.

Multiple sequence alignment at the domain level can result in the detection of homologous sequences that are more difficult to detect using a complete chain sequence.

Methods used to gain an insight into the structure and function of a protein work best at the domain level.

# Domain databases

# SMART

Peer Bork

<http://smart.embl.de/>

Manual definition of domain (bibliography)

Generate profile from instances of domain

Search for remote homologs (HMMer)

Include them in profile


Iterate until convergence

Schultz et al (1998) *PNAS*

...

Letunic et al (2014) *Nucleic Acids Research*

# Domain databases



Schultz et al. (1998) *Proc. Natl. Acad. Sci. USA* 95, 5857-5864  
Letunic et al. (2012) *Nucleic Acids Res* , doi:10.1093/nar/gkr931

HOME SETUP FAQ ABOUT GLOSSARY WHAT'S NEW FEEDBACK


**SMART MODE:**  
NORMAL  
GENOMIC

Simple  
Modular  
Architecture  
Research  
Tool


## Sequence analysis

You may use either a [Uniprot/Ensembl](#) sequence identifier (ID) / accession number (ACC) or the protein sequence itself to perform the SMART analysis service.

**Sequence ID or ACC**

Examples: #1, #2 

**Protein sequence**

Examples: #1, #2 


HMMER searches of the SMART database occur by default. You may also find:

[Outlier homologues](#) and homologues of known structure


## Architecture analysis

You can search for proteins with combinations of [specific domains](#) in different species or taxonomic ranges. You can input the domains directly into "Domain selection" box, or use "GO terms query" to get a list of domains.


**Domain selection**

Examples: #1, #2 

**GO terms query**

Examples: #1, #2 

**Taxonomic selection**

Select a taxonomic range via the selection box or type it into the text box below: 

All

Examples: #1, #2

You can try an [Advanced Query](#) if you're familiar with SQL.

# Domain databases

# SMART

## Domains detected by SMART

### SH3

Src homology 3 domains



SH3

SMART  
accession  
number:

SM00326

Description:

Src homology 3 (SH3) domains bind to target proteins through sequences containing proline and hydrophobic amino acids. Pro-containing polypeptides may bind to SH3 domains in 2 different binding orientations.

Interpro  
abstract  
([IPR001452](#)):

SH3 (src Homology-3) domains are small protein modules containing approximately 50 amino acid residues [([PUBMED:15335710](#)), ([PUBMED:11256992](#))]. They are found in a great variety of intracellular or membrane-associated proteins [([PUBMED:1639195](#)), ([PUBMED:14731533](#)), ([PUBMED:7531822](#))] for example, in a variety of proteins with enzymatic activity, in adaptor proteins, such as fodrin and yeast actin binding protein ABP-1.

The SH3 domain has a characteristic fold which consists of five or six beta-strands arranged as two tightly packed anti-parallel beta sheets. The linker regions may contain short helices. The surface of the SH3-domain bears a flat, hydrophobic ligand-binding pocket which consists of three shallow grooves defined by conservative aromatic residues in which the ligand adopts an extended left-handed helical arrangement. The ligand binds with low affinity but this may be enhanced by multiple interactions. The region bound by the SH3 domain is in all cases proline-rich and contains PXXP as a core-conserved binding motif. The function of the SH3 domain is not well understood but they may mediate many diverse processes such as increasing local concentration of proteins, altering their subcellular location and mediating the assembly of large multiprotein complexes [([PUBMED:7953536](#))].

The crystal structure of the SH3 domain of the cytoskeletal protein spectrin, and the solution structures of SH3 domains of phospholipase C (PLC-y) and phosphatidylinositol 3-kinase p85 alpha-subunit, have been determined [([PUBMED:1279434](#)), ([PUBMED:7684655](#)), ([PUBMED:7681365](#))]. In spite of relatively limited sequence similarity, their overall structures are similar. The domains belong to the alpha+beta structural class, with 5 to 8 beta-strands forming 2 tightly-packed, anti-parallel beta-sheets arranged in a barrel-like structure, and intervening loops sometimes forming helices. Conserved aliphatic and aromatic residues form a hydrophobic core (A11, L23, A29, V34, W42, L52 and V59 in PLC-y [([PUBMED:7681365](#))] and a hydrophobic pocket on the molecular surface (L12, F13, W53 and P55 in PLC-y). The conserved core is believed to stabilise the fold, while the pocket is thought to serve as a binding site for target proteins. Conserved carboxylic amino acids located in the loops, on the periphery of the pocket (D14 and E22), may be involved in protein-protein interactions via proline-rich regions. The N- and C-termini are packed in close proximity, indicating that they are independent structural modules.

GO function:

protein binding ([GO:0005515](#))

# Domain databases

# SMART

## Sequence analysis

You may use either a [Uniprot/Ensembl](#) sequence identifier (ID) / accession number (ACC) or the protein sequence itself to perform the SMART analysis service.

### Sequence ID or ACC

Examples: [#1](#), [#2](#)



### Protein sequence

Examples: [#1](#), [#2](#)

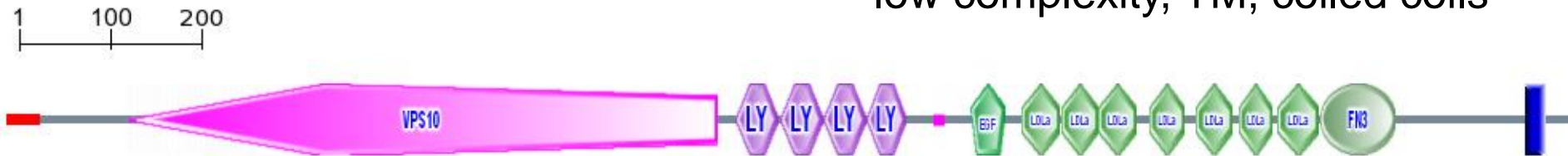




# Domain databases

# SMART

Extra features:  
Signal-peptide,  
low complexity, TM, coiled coils



Confidently predicted domains, repeats, motifs and features:

Name	Begin	End	E-value
signal peptide	1	36	-
VPS10	125	741	0.00e+00
LY	761	806	2.88e+00
LY	807	851	3.94e-04
LY	852	896	5.31e-10
LY	897	939	1.76e-15
low complexity	968	979	-
EGF	1006	1042	1.87e+01
LDLa	1059	1098	2.69e-10
LDLa	1100	1138	1.62e-13
EGF_like	1138	1177	5.24e+01
LDLa	1139	1178	1.46e-11
LDLa	1193	1230	2.07e-11
LDLa	1240	1278	2.91e-06
LDLa	1286	1321	3.21e-08
LDLa	1326	1369	1.27e-06
FN3	1370	1448	1.36e-03
transmembrane	1584	1606	-

#### Additional information

[Display](#) other IDs, orthology and alternative splicing data for this sequence.

#### Domain architecture analysis

This domain architecture was probably invented with the emergence of [Hydra viridis](#).

[Display](#) all proteins with similar domain [organisation](#).

[Display](#) all proteins with similar domain [composition](#).

# Domain databases

# SMART

The following proteins have the same domain **composition** as your query protein.

You can  of  or selected (below) proteins.

If you want only single domain sequences in the fasta file, type domain name here:

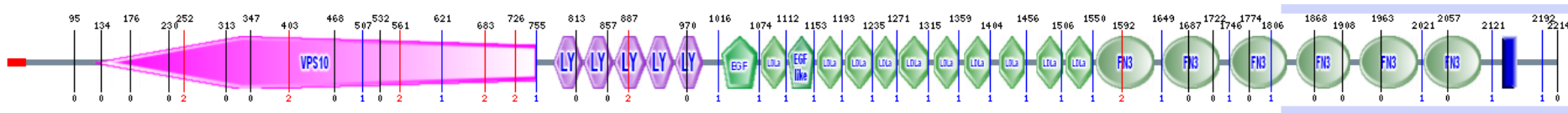
## Taxonomic tree of query results.

- Eukaryota (17)
  - Metazoa (17)
    - Arthropoda (5)

Protein	UPI000013D0B1 ( <a href="#">source</a> )
Description	Sortilin-related receptor precursor (Sorting protein-related receptor containing LDLR class A repeats) (SorLA) (SorLA-1) (Low-density lipoprotein receptor relative with 11 ligand-binding repeats) (LDLR relative with 11 ligand-binding repeats) (LR11).
Species	<i>Homo sapiens</i>
Domain architecture invented in	Eutheria
Representative of protein cluster	CLUST_UPI000013D0B1

1 100 200

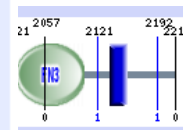
Due to overlapping domains, there are 4 representations of the protein



# Domain databases

# SMART

<b>Protein</b>	<b>EGF</b>
<b>Descrip</b>	Epidermal growth factor-like domain.
<b>Species</b>	
<b>Domain</b>	
<b>inverter</b>	
<b>Represent</b>	<b>SMART</b>
<b>cluster</b>	<b>accession</b>
1	SM00181
	<b>Description:</b>
<b>Due to o</b>	Epidermal growth factors and transforming growth factors belong to a general class of proteins that share a repeat pattern involving a number of conserved Cys residues. Growth factors are involved in cell recognition and division. The repeating pattern, especially of cysteines (the so-called EGF repeat), is thought to be important to the 3D structure of the proteins, and hence its recognition by receptors and other molecules. The type 1 EGF signature includes six conserved cysteines believed to be involved in disulphide bond formation. The EGF motif is found frequently in nature, particularly in extracellular proteins.
	<b>Interpro</b>
	<b>abstract</b>
	<b>(IPR006210):</b>
	<b>Family</b>
	<b>alignment:</b>
	View <input type="text" value="Alignment consensus sequence"/> or <input type="text" value="Family alignment in"/> <input type="text" value="CHROMA format"/>



There are **43703** EGF domains in 14525 proteins in SMART's nrdb database.

Click on the following links for more information.

► [Evolution](#) (species in which this domain is found)

▼ [Structure](#) (3D structures containing this domain)

## 3D Structures of EGF domains in PDB

1a3p, 1adx, 1cqe, 1cqe, 1cvu, 1cvu, 1cww, 1cx2, 1cx2, 1cx2, 1cx2, 1ddx, 1ddx, 1ddx, 1ddx, 1diy, 1dqb, 1dx5, 1dx5, 1dx5, 1dx5, 1ebv, 1egf, 1epg, 1eph, 1epi, 1epj, 1eqg, 1eqg, 1eqh, 1eqh, 1esl, 1fe2, 1fjs, 1fsb, 1g1q, 1g1q, 1g1q, 1g1q, 1g1r, 1g1r, 1g1r, 1g1r, 1g1s, 1g1s, 1g1t, 1gk5, 1gl4, 1hae, 1haf, 1hcg, 1hre, 1hrf, 1ht5, 1ht5, 1ht8, 1ht8, 1igx, 1igz, 1ijq, 1ijq, 1iox, 1ip0, 1ivo, 1ivo, 1j9c, 1jbu, 1jl9, 1jl9, 1k36, 1k37, 1kig, 1kli, 1klj, 1kye, 1mox, 1mox, 1mq5, 1mq6, 1nql, 1p9j, 1pge, 1pge, 1pgf, 1pgf, 1pgg, 1pgg, 1prh, 1prh, 1pth, 1pth, 1pox, 1pox, 1pox, 1pox, 1q4g, 1q4g, 1qfk, 1rfn, 1tpg, 1u67, 1v3x, 1w7x, 1w8b, 1xdt, 1xfe, 1ygc, 1yo8, 1yuf, 1yug, 1z1y, 1z1y, 1z27, 1z3g, 1z3g, 1z6e, 1zaq, 2adx, 2ayl, 2ayl, 2bmg, 2bok, 2bq6, 2bq7, 2bqw, 2bz6, 2d1j, 2ddu, 2e26, 2fzz, 2g00, 2gd4, 2gd4, 2gy5, 2gy7, 2i9a, 2i9a, 2i9a, 2i9a, 2i9b, 2i9b, 2i9b, 2i9b, 2oye, 2oyu, 2p16, 2p3f, 2p3t, 2p3u, 2p93, 2p94, 2p95, 2pe4, 2pr3, 2puq, 2q1j, 2ra0, 2tgf, 3egf, 3pgh, 3pgh, 3pgh, 3pgh, 3tgf, 4cox, 4cox, 4cox, 4cox, 4tgf, 5cox, 5cox, 5cox, 5cox, 6cox, 6cox

# Domain databases

# PFAM

Erik Sonnhammer/Ewan Birney/Alex Bateman

<http://pfam.xfam.org/>



[HOME](#) | [SEARCH](#) | [BROWSE](#) | [FTP](#) | [HELP](#)  
| [ABOUT](#)

**Pfam 35.0 (November 2021, 19632 entries)**

The Pfam database is a large collection of protein families, each represented by **multiple sequence alignments** and **hidden Markov models (HMMs)**. [More...](#)

---

#### QUICK LINKS

[SEQUENCE SEARCH](#)

[VIEW A PFAM ENTRY](#)

[VIEW A CLAN](#)

[VIEW A SEQUENCE](#)

[VIEW A STRUCTURE](#)

#### YOU CAN FIND DATA IN PFAM IN VARIOUS WAYS...

Analyze your protein sequence for Pfam matches

View Pfam annotation and alignments

See groups of related entries

Look at the domain organisation of a protein sequence

Find the domains on a PDB structure

Sonnhammer et al (1997) *Proteins*

...

Mistry et al (2021) *Nucleic Acids Research*

# Domain databases

# PFAM

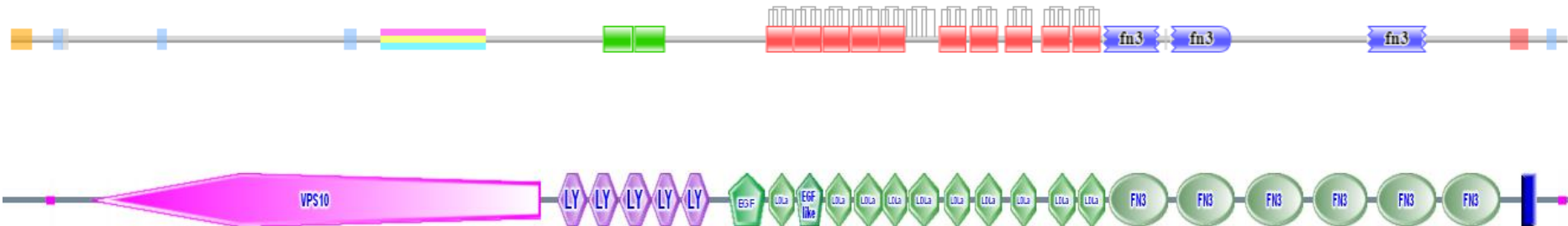
This is the summary of UniProt entry [SORL\\_HUMAN](#) (Q92673).

<b>Description:</b>	Sortilin-related receptor
<b>Source organism:</b>	<a href="#">Homo sapiens (Human)</a> (NCBI taxonomy ID <a href="#">9606</a> ) <a href="#">View Pfam proteome data.</a>
<b>Length:</b>	2214 amino acids

**Please note:** when we start each new Pfam data release, we take a copy of the UniProt sequence database. This snapshot of UniProt forms the basis of the overview that you see here. It is important to note that, although some UniProt entries may be removed *after* a Pfam release, these entries will not be removed from Pfam until the next Pfam data release.

## Pfam domains

This image shows the arrangement of the Pfam domains that we found on this sequence. Clicking on a domain will take you to the page describing that Pfam entry. The table below gives the domain boundaries for each of the domains. [More...](#)

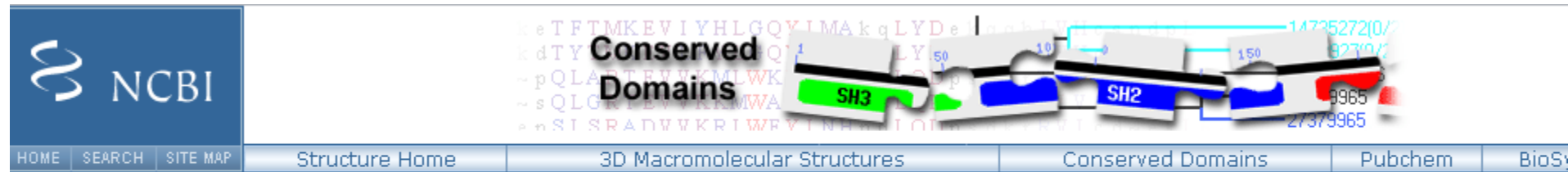


# Domain databases

## CDD

Stephen Bryant

<http://www.ncbi.nlm.nih.gov/cdd>



NCBI

HOME SEARCH SITE MAP

Structure Home 3D Macromolecular Structures Conserved Domains Pubchem BioS

Search for [Conserved Domains](#) within a protein sequence

Enter **Protein** Query as Accession, Gi, or Sequence in [FASTA format](#) [?](#)

Submit

Reset

### OPTIONS

Search against database [?](#): CDD -- 34177 PSSMs [?](#)

Expect Value [?](#) threshold: 0.01 [?](#)

Apply low-complexity filter [?](#)

Force live search [?](#)

Maximum number of hits [?](#) 250 [?](#)

Result mode  Concise [?](#)  Full [?](#)

Marchler-Bauer et al (2015) *Nucleic Acids Res*



# Domain databases

SORLA/SORL1 from *Homo sapiens*

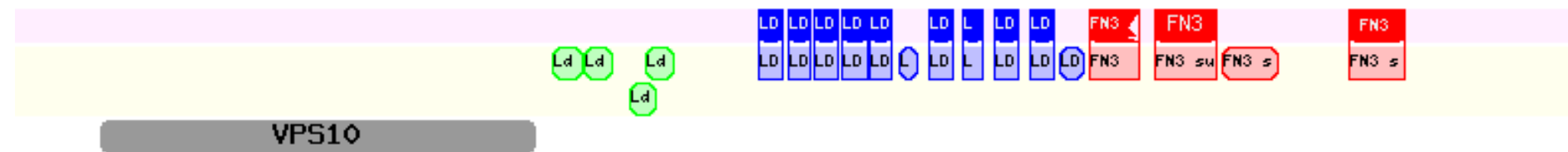
## SMART



## PFAM

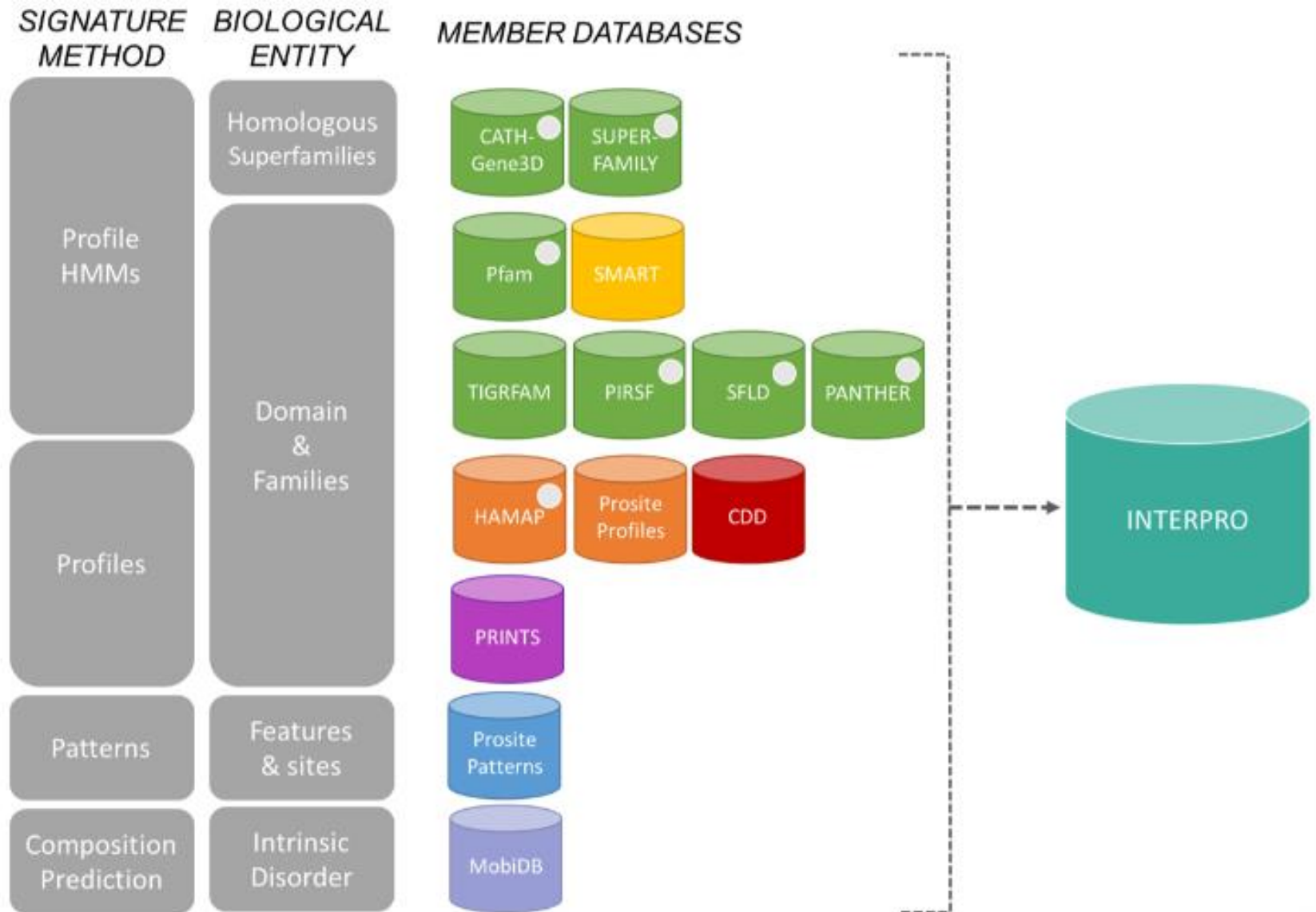


## CDD





# InterPro

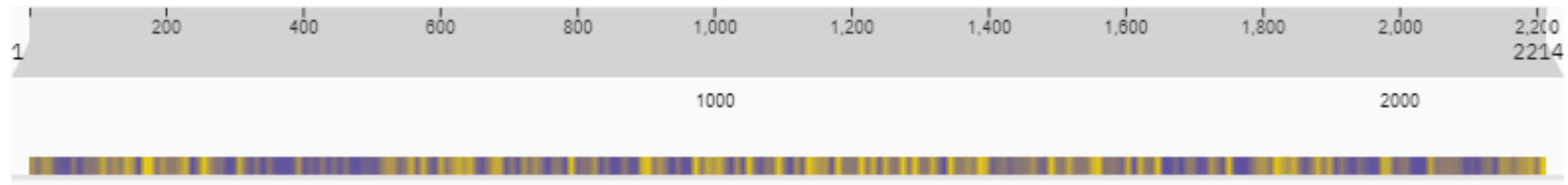


# InterPro

## SORLA/SORL1 from *Homo sapiens*

<https://www.ebi.ac.uk/interpro/protein/reviewed/Q92673/>

Entry matches to this protein ⓘ



▼ Domain



IPR003961  
SM00060  
PSS0853  
PF00041  
cd00063

IPR006581  
SM00602

IPR031777  
PF15901

IPR031778  
PF15902

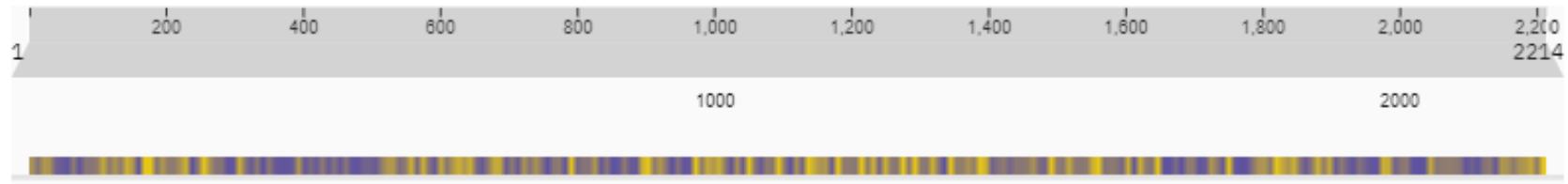


# InterPro

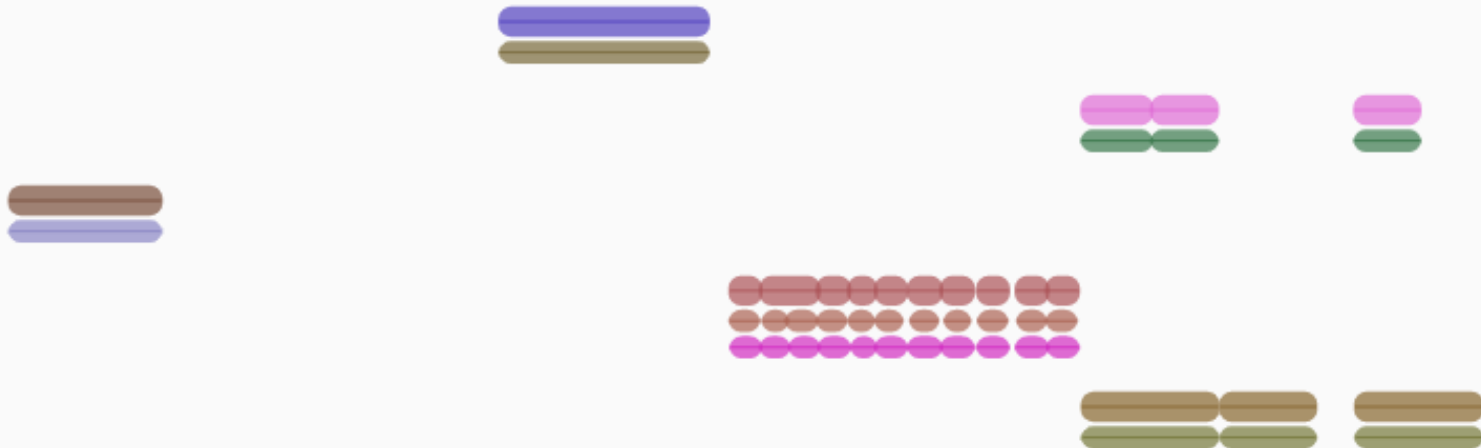
## SORLA/SORL1 from *Homo sapiens*

<https://www.ebi.ac.uk/interpro/protein/reviewed/Q92673/>

Entry matches to this protein ⓘ



▼ Homologous Superfamily



**H** IPR011042  
G3DSA:2.120.10.30

**H** IPR013783  
G3DSA:2.60.40.10

**H** IPR015943  
G3DSA:2.130.10.10

**H** IPR036055  
SSF57424  
G3DSA:4.10.400.10

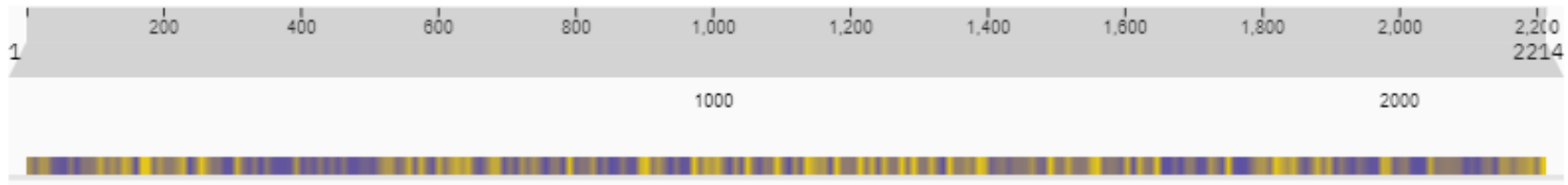
**H** IPR036116  
SSF49265

# InterPro

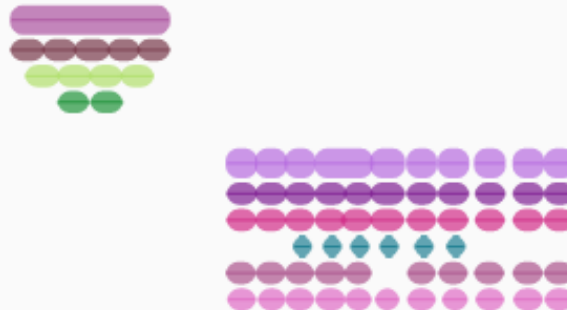
## SORLA/SORL1 from *Homo sapiens*

<https://www.ebi.ac.uk/interpro/protein/reviewed/Q92673/>

Entry matches to this protein ⓘ



▼ Repeat



**R** IPR000033  
SM00135  
PS51120  
PF00058

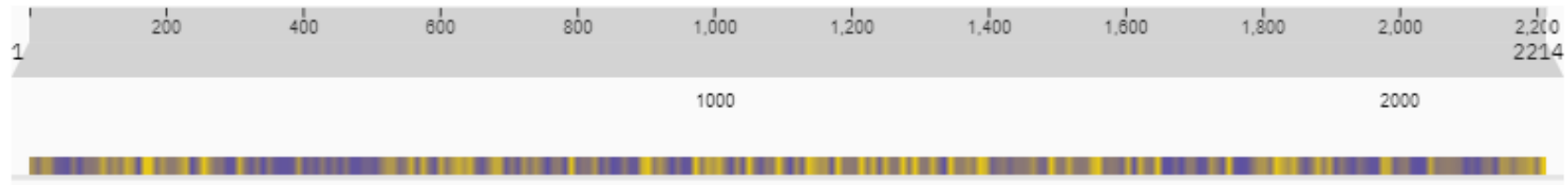
**R** IPR002172  
SM00192  
PS50068  
PR00261  
PF00057  
cd00112

# InterPro

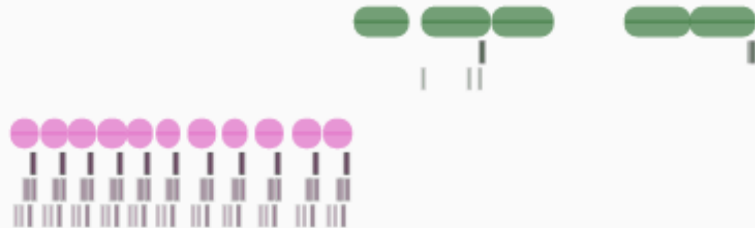
## SORLA/SORL1 from *Homo sapiens*

<https://www.ebi.ac.uk/interpro/protein/reviewed/Q92673/>

Entry matches to this protein ⓘ



▼ Residues



cd00063  
Cytokine receptor motif  
Interdomain contacts

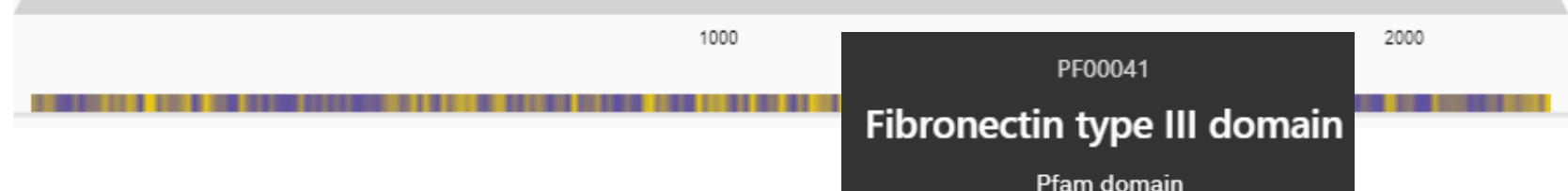
cd00112  
D-X-S-D-E motif  
Calcium-binding site  
Putative binding surface

# InterPro

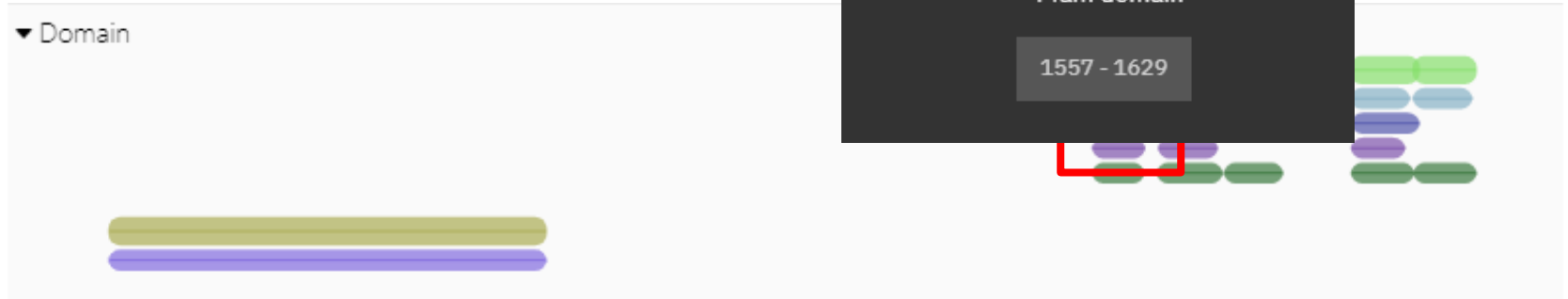
## SORLA/SORL1 from *Homo sapiens*

<https://www.ebi.ac.uk/interpro/protein/reviewed/Q92673/>

Entry matches to this protein <sup>i</sup>



PF00041  
**Fibronectin type III domain**  
Pfam domain  
1557 - 1629



**D** IPR003961  
SM00060  
P550853  
PF00041  
cd00063

**D** IPR006581  
SM00602

**D** IPR031777  
PF15901

**D** IPR031778  
PF15902

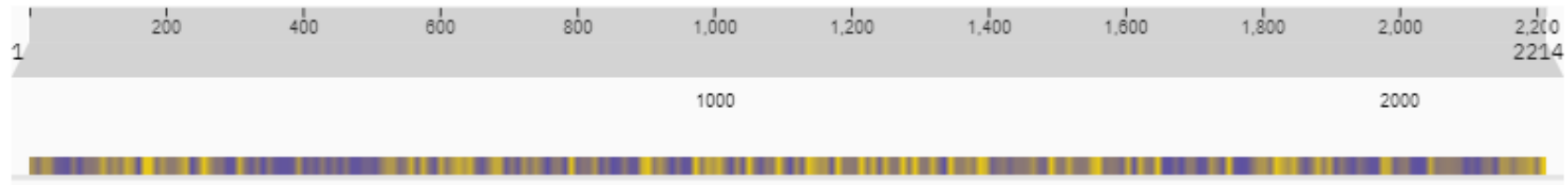
1380	CIPNRWKCDR	ENDCGDWSDE	KDCGDSHILP	FSTPGPSTCL	PNYYRCSSGT	CVMDTIWCDG
1440	YRDCADGSDE	EACPLLANVT	AASTPTQLGR	CDRFEFECHQ	PKTCIPNWKR	CDGHQDCQDG
1500	RDEANCPHTS	TLTCMSREFQ	CEDGEACIVL	SERCDGFLDC	SDESDEKACS	DELTIVYKVQN
1560	LQWTADFSGD	VLIWMRPKK	MPSASCVYNV	YYRVVGESIW	KILETHSNKI	NTVLKVLKPD
1620	TTYQVKVQVQ	CLSKAHTND	FVTLRTPEGL	PDAPRNLQLS	LPREAEGVIV	GHWAPPIHTH
1680	GLIREYIVEY	SRSGSKMWAS	QRAASNFTET	KNLLVNTLYT	VRVAAVTSRG	IGNWSDSKSI

# InterPro

## SORLA/SORL1 from *Homo sapiens*

<https://www.ebi.ac.uk/interpro/protein/reviewed/Q92673/>

Entry matches to this protein ⓘ



▼ Domain



IPR003961  
SM00060  
PF00041  
cd00063

IPR006581  
SM00602

IPR031777  
PF15901

IPR031778  
PF15902



# InterPro

Home / Browse / By Entry / Pfam / PF00041 / Overview

## Pfam Fibronectin type III domain PF00041

Pfam entry ⓘ

Overview	
Proteins	251k
Domain Architectures	18k
Taxonomy	21k
Proteomes	5k
Structures	303
Signature	
AlphaFold	126k
Alignment	
Curation	

Member database	Pfam ⓘ
Pfam type	Domain
Short name	<i>fn3</i>
Set	cl0159

### References

[1]. Primary structure of human fibronectin: differential splicing may generate at least 10 polypeptides from a single gene. Kornblihtt AR, Umezawa K, Vibe-Pedersen K, Baralle FE. *EMBO J.* 4, 1755-9, (1985). [View article](#) ⓘ PMID: 2992939 ⓘ

[2]. Tracing the spread of fibronectin type III domains in bacterial glycohydrolases. Little E, Bork P, Doolittle RF. *J. Mol. Evol.* 39, 631-43, (1994). [View article](#) ⓘ PMID: 7528812 ⓘ

[Add your annotation](#) ▾

### Integrated to

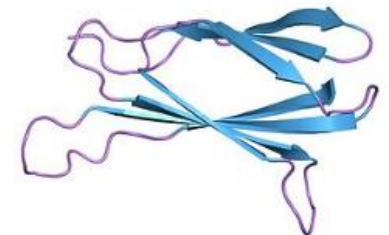
> IPR003961

[3]. Structural design and molecular evolution of a cytokine receptor superfamily. Bazan JF. *Proc. Natl. Acad. Sci. U.S.A.* 87, 6934-8, (1990). [View article](#) ⓘ PMID: 2169613 ⓘ

### Fibronectin type III domain ⓘ [Wikipedia](#)

The **Fibronectin type III domain** is an evolutionarily conserved protein domain that is widely found in animal proteins. The fibronectin protein in which this domain was first identified contains 16 copies of this domain. The domain is about 100 amino acids long and possesses a beta sandwich structure. Of the three fibronectin-type domains, type III is the only one without disulfide bonding present. Fibronectin domains are found in a wide variety of extracellular proteins. They are widely distributed in animal species, but also found sporadically in yeast, plant and bacterial proteins.

Fibronectin type III domain

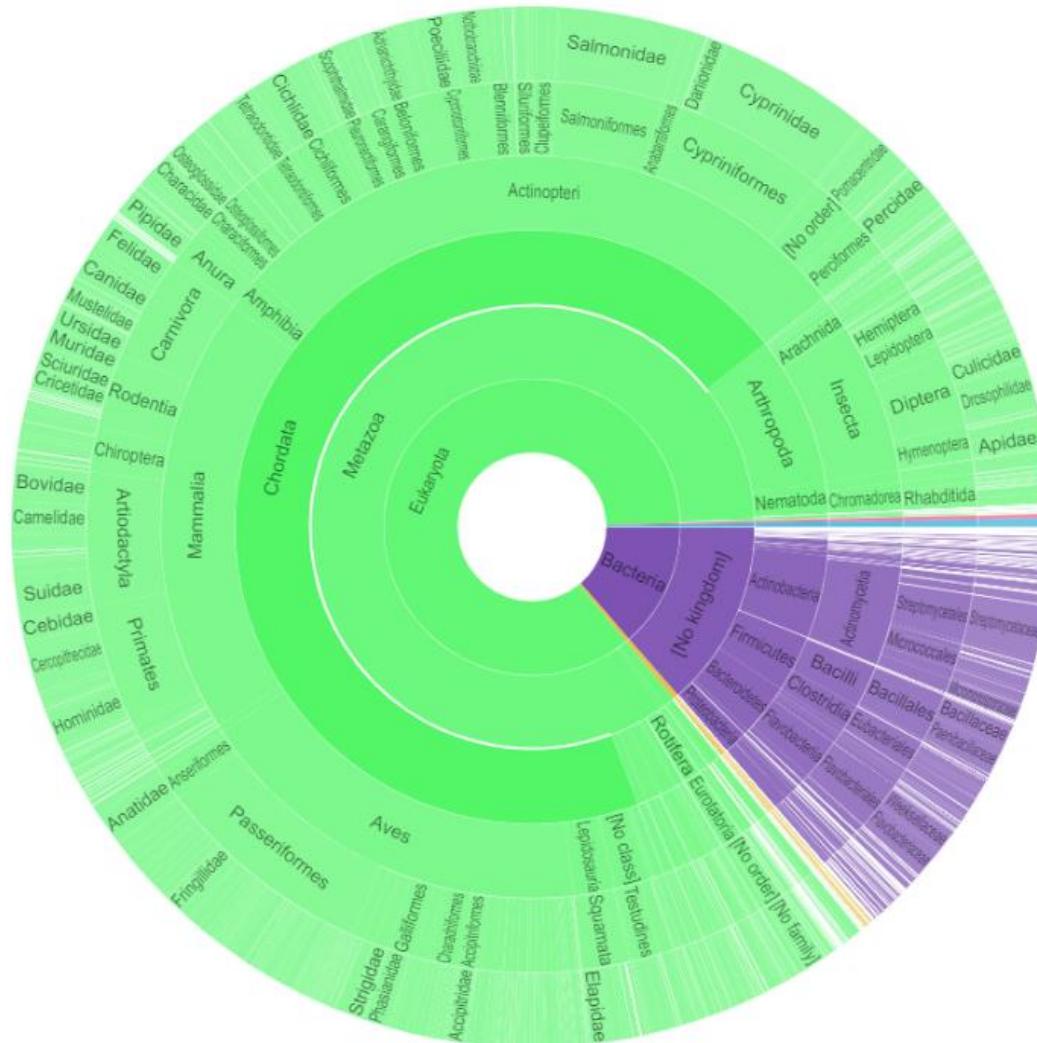




# InterPro

- Domain Architectures 18k
- Taxonomy 21k**
- Proteomes 5k
- Structures 303
- Signature
- AlphaFold 126k
- Alignment
- Curation

The number of species for this sunburst is 12594. The depth of the visualisation has been limited. You can modify this with the controller in the right side. however, please note this might affect the performance in your browser.



## Legends

- bacteria
- viruses
- archaea
- eukaryota
- Other

## Weight Segments by

Number of sequences

Font Size  
14

## Sunburst Depth (6 rings)

2  8

## Selected Taxon

### Name

*Chordata*

### Number of sequences

175913

### Number of species

1737

### Lineage

root; Eukaryota; Metazoa; Chordata;

# InterPro

## Pfam Fibronectin type III domain PF00041

Pfam entry ⓘ



This entry matches these structures:

Overview

Proteins 251k

Domain Architectures 18k

Taxonomy 21k

Proteomes 5k

**Structures 303**

Signature

AlphaFold 126k

Alignment

Curation

1 - 20 of 303 structures



Export



ACCESSION	NAME	SOURCE DATABASE	STRUCTURE	MATCHES
1a22	HUMAN GROWTH HORMONE BOUND TO SINGLE RECEPTOR	PDB		
1axi	STRUCTURAL PLASTICITY AT THE HGH:HGHBP INTERFACE	PDB		
1bj8	THIRD N-TERMINAL DOMAIN OF GP130, NMR, MINIMIZED AVERAGE STRUCTURE	PDB		
1bpv	TITIN MODULE A71 FROM HUMAN CARDIAC MUSCLE, NMR, 50 STRUCTURES	PDB		
1bqu	CYTOKYNE-BINDING REGION OF GP130	PDB		

# Exercise 1

## Find structures in PDB for human myosin X

Find this sequence in InterPro.

<https://www.ebi.ac.uk/interpro/protein/reviewed/Q9HD67/>

[Home](#) / [Browse](#) / [By Protein](#) / [Reviewed](#) / [Q9HD67](#) / [Overview](#)

UniProtKB/Swiss-Prot protein <sup>i</sup>

Q9HD67

### Overview

Entries 18

Structures 7

Sequence

Similar Proteins 88

AlphaFold 1

Short name

*MYO10\_HUMAN*

Length

2058 amino acids

Species

[Homo sapiens \(Human\)](#)

Proteome

UP000005640

Function <sup>i</sup>

Myosins are actin-based motor molecules with ATPase activity. Unconventional myosins serve in intracellular movements. MYO10 binds to actin filaments and actin bundles and functions as a plus end-directed motor. Moves with higher velocity and takes l...

...

# Exercise 1

## Find structures in PDB for human myosin X









- Which domains of myosin X are covered by the solved structures?
  
  
  
  
  
  
  
  
  
  
- Is there a part of the protein for which there are no known structures? Does it have predicted domains?

# Exercise 2

## Analyse domain predictions

Slide down to see the details of the structures.

This protein matches these structures:

ACCESSION	NAME	SOURCE DATABASE	STRUCTURE	MATCHES
2lw9	NMR solution structure of Myo10 anti-CC	PDB		
3au4	Structure of the human myosin-X MyTH4-FERM cassette bound to its specific cargo, DCC	PDB		
3au5	Structure of the human myosin-X MyTH4-FERM cassette	PDB		
3pzd	Structure of the myosin X MyTH4-FERM/DCC complex	PDB		
5i0h	Crystal structure of myosin X motor domain in pre-powerstroke state	PDB		

# Exercise 2

## Analyse domain predictions

- Examine the structure of 3pzd

How do the domain predictions fit the structure?

- Chain B in this structure is a small peptide.

Which domain in Myosin X is interacting with this peptide?

# Exercise 3

## AlphaFold prediction

- There is a predicted structure

[Home](#) / [Browse](#) / [By Protein](#) / [Reviewed](#) / [Q9HD67](#) / [Overview](#)

UniProtKB/Swiss-Prot protein <sup>i</sup>

Q9HD67

<b>Overview</b>	Short name	<i>MYO10_HUMAN</i>
Entries 18	Length	2058 amino acids
Structures 7	Species	<a href="#">Homo sapiens (Human)</a>
Sequence	Proteome	UP000005640
Similar Proteins 88	Function <sup>i</sup>	Myosins are actin-based motor molecules with ATPase activity. Unconventional myosins serve in intracellular movements. MYO10 binds to actin filaments and actin bundles and functions as a plus end-directed motor. Moves with higher velocity and takes l...
<b>AlphaFold 1</b>		

# Exercise 3

## AlphaFold prediction

- There is a predicted structure
- Download the PDB file and load it in Chimera
- Select the central region without PDB information (Select/Atom specifier), inverse the selection, and delete everything else (Actions/Atoms/Delete).  
Describe the structure predicted for this region and how this could affect structure determination.
- Examine the PH domains. How many domains do you see? Is there anything particular about them?



# Exercise 3

## AlphaFold prediction

