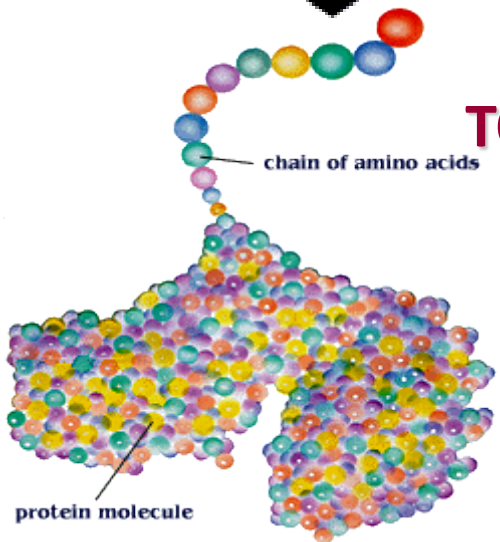
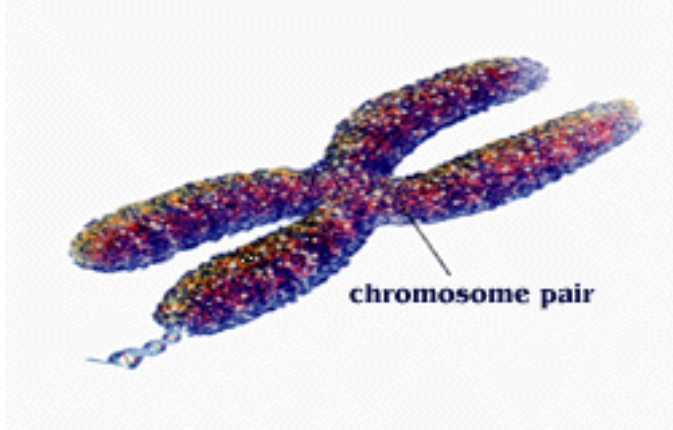


PROTEIN SEQUENCE ANALYSIS



TOOLS

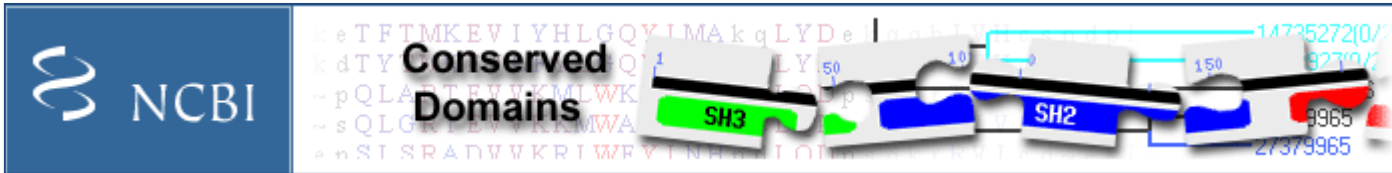
PROTEIN DATABASES

- PROTEIN SEQUENCE
- PROPERTIES
- MOTIF/DOMAIN
- FOLDING

Protein Sequence/Motif/Domain Databases



<http://www.uniprot.org/>



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



<http://www.ebi.ac.uk/Tools/pfa/iprscan/>

Protein Analysis Tools



<http://tw.expasy.org/>



UniProtKB ▾

P53 again

Advanced ▾



BLAST Align Retrieve/ID Mapping

Help Contact

Welcome to the new UniProt website! We hope you enjoy the new design. If you're not quite ready yet, you can still [go back to the old site](#).

The mission of UniProt is to provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information.

UniProtKB

Swiss-Prot
(546,439)



Manually annotated and reviewed.

TrEMBL
(83,955,074)



Automatically annotated and not reviewed.

UniRef

Sequence clusters



UniParc

Sequence archive



Proteomes



Supporting data

Literature citations



Taxonomy



Subcellular locations



Cross-ref. databases



Diseases

XXX

Keywords



News



Small is beautiful (and useful) | Evidences in the UniProtKB flat file format
[UniProt release 2014_09](#)

Ubiquitin caught at its own game | New human variant types available on the FTP site
[UniProt release 2014_08](#)

Lark or owl? PER3 is the answer | [UniProt release 2014_07](#)

[News archive](#)

Results


<http://www.uniprot.org/uniprot/?query=p53&sort=score>


Filter byⁱ

 Reviewed
(1,839)
Swiss-Prot

 Unreviewed
(14,987)
TrEMBL

Show help for UniProtKB


 Basket ▾

 Columns

 BLAST

 Align

 Download

 Add to basket

◀ 1 to 25 of 16,826 ▶

Show 25 ▾

Popular organisms

Human (1,056)

Mouse (803)

Rat (312)

Bovine (270)

Zebrafish (231)

Other organisms

Search terms

Filter "p53" as:
gene name (119)

gene ontology
(3,273)

organism (3,646)

protein family (630)

protein name (828)

strain (3,679)

taxonomy (3,680)

View by

Taxonomy

Keywords

Gene Ontology















Enzyme class

Pathway

UniRef

Your results in
sequence clusters with
identity of:

100%, 90% or 50%


<input type="checkbox"/>	Entry	Entry name ▾		Protein names ▾ 	Gene names ▾	Organism ▾	Length ▾ 
<input type="checkbox"/>	P04637	P53_HUMAN		Cellular tumor antigen p53	TP53, P53	Homo sapiens (Human)	393
<input type="checkbox"/>	P02340	P53_MOUSE		Cellular tumor antigen p53	Tp53, P53, Trp53	Mus musculus (Mouse)	387
<input type="checkbox"/>	P10361	P53_RAT		Cellular tumor antigen p53	Tp53, P53	Rattus norvegicus (Rat)	391
<input type="checkbox"/>	Q42578	PER53_ARATH		Peroxidase 53	PER53, P53, At5g06720, MPH15.8	Arabidopsis thaliana (Mouse-ear cress)	335
<input type="checkbox"/>	P25035	P53_ONCMY		Cellular tumor antigen p53	tp53, p53	Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri)	396
<input type="checkbox"/>	P79820	P53_ORYLA		Cellular tumor antigen p53	tp53, p53	Oryzias latipes (Medaka fish) (Japanese ricefish)	352
<input type="checkbox"/>	Q92143	P53_XIPMA		Cellular tumor antigen p53	tp53, p53	Xiphophorus maculatus (Southern platyfish) (Platypoecilus maculatus)	342
<input type="checkbox"/>	Q29537	P53_CANFA		Cellular tumor antigen p53	TP53, P53	Canis familiaris (Dog) (Canis lupus familiaris)	381
<input type="checkbox"/>	O09185	P53_CRIGR		Cellular tumor antigen p53	TP53, P53	Cricetulus griseus (Chinese hamster) (Cricetulus barabensis griseus)	393
<input type="checkbox"/>	Q8SPZ3	P53_DELLE		Cellular tumor antigen p53	TP53, P53	Delphinapterus leucas (Beluga whale)	387
<input type="checkbox"/>	P61260	P53_MACFU		Cellular tumor antigen p53	TP53, P53	Macaca fuscata fuscata (Japanese macaque)	393

Results

<http://www.uniprot.org/uniprot/?query=p53&sort=score>


Filter byⁱ

 Reviewed
(1,839)
Swiss-Prot

 Unreviewed
(14,987)
TrEMBL


Show help for UniProtKB


 Basket ▾

 Columns

 BLAST

 Align

 Download

 Add to basket

◀ 1 to 25 of 16,826 ▶

Show 25 ▾

Popular organisms

Human (1,056)

Mouse (803)

Rat (312)

Bovine (270)

Zebrafish (231)

Other organisms

Search terms

Filter "p53" as:
gene name (119)

gene ontology
(3,273)

organism (3,646)

protein family (630)

protein name (828)

strain (3,679)

taxonomy (3,680)

View by

Taxonomy

Keywords

Gene Ontology















Enzyme class

Pathway

UniRef

Your results in
sequence clusters with
identity of:

100%, 90% or 50%

<input type="checkbox"/>	Entry	Entry name ▾		Protein names ▾ 	Gene names ▾	Organism ▾	Length ▾ 
<input type="checkbox"/>	P04637	P53_HUMAN		Cellular tumor antigen p53	TP53, P53	Homo sapiens (Human)	393
<input type="checkbox"/>	P02340	P53_MOUSE		Cellular tumor antigen p53	Tp53, P53, Trp53	Mus musculus (Mouse)	387
<input type="checkbox"/>	P10361	P53_RAT		Cellular tumor antigen p53	Tp53, P53	Rattus norvegicus (Rat)	391
<input type="checkbox"/>	Q42578	PER53_ARATH		Peroxidase 53	PER53, P53, At5g06720, MPH15.8	Arabidopsis thaliana (Mouse-ear cress)	335
<input type="checkbox"/>	P25035	P53_ONCMY		Cellular tumor antigen p53	tp53, p53	Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri)	396
<input type="checkbox"/>	P79820	P53_ORYLA		Cellular tumor antigen p53	tp53, p53	Oryzias latipes (Medaka fish) (Japanese ricefish)	352
<input type="checkbox"/>	Q92143	P53_XIPMA		Cellular tumor antigen p53	tp53, p53	Xiphophorus maculatus (Southern platyfish) (Platypoecilus maculatus)	342
<input type="checkbox"/>	Q29537	P53_CANFA		Cellular tumor antigen p53	TP53, P53	Canis familiaris (Dog) (Canis lupus familiaris)	381
<input type="checkbox"/>	O09185	P53_CRIGR		Cellular tumor antigen p53	TP53, P53	Cricetulus griseus (Chinese hamster) (Cricetulus barabensis griseus)	393
<input type="checkbox"/>	Q8SPZ3	P53_DELLE		Cellular tumor antigen p53	TP53, P53	Delphinapterus leucas (Beluga whale)	387
<input type="checkbox"/>	P61260	P53_MACFU		Cellular tumor antigen p53	TP53, P53	Macaca fuscata fuscata (Japanese macaque)	393

Customize results table

Columns to be displayedⁱ

Reset to *default*

Save

Cancel

Drag and drop to re-order.

Entry Entry name  Protein names Gene names Organism Length

Add more columnsⁱ

▶ Expand all

Search:

▶ Names & Taxonomy

▶ Sequences

▶ Function

▶ Miscellaneous

▶ Interaction

▶ Expression

▶ Gene Ontology (GO)

▶ Pathology & Biotech

▶ Subcellular location

▶ PTM / Processing

▶ Structure

▶ Publications

▶ Date of

▶ Family & Domains

▶ Taxonomic lineage

▶ Taxonomy Id

Databases

▶ Sequence

▶ 3D structure

▶ Protein-protein

▶ Chemistry

▶ Protein family/group

▶ PTM

▶ Polymorphism

▶ 2D gel

▶ Proteomic

▶ Protocols and

▶ Genome annotation

▶ Organism-specific

▶ Phylogenomic

▶ Enzyme and

▶ Other

▶ Gene expression

▶ Family and domain

▼ Names & Taxonomy

- Entry name
- Protein names
- Gene names
 - Gene names (primary)
 - Gene names (synonym)
 - Gene names (ordered locus)
 - Gene names (ORF)
- Organism
 - Organism ID
 - Proteomes
 - Taxonomic lineage
 - Virus hosts



▼ Sequences

- Length
 - Fragment
 - Gene encoded by
 - Alternative products
 - Erroneous gene model prediction
 - Erroneous initiation
 - Erroneous termination
 - Erroneous translation
 - Frameshift
 - Mass spectrometry
 - Polymorphism
 - RNA editing
 - Sequence caution
 - Mass
 - Sequence
 - Alternative sequence
 - Natural variant
 - Non-adjacent residues
 - Non-standard residue
 - Non-terminal residue
 - Sequence conflict
 - Sequence uncertainty
 - Sequence version

▼ Function

- EC number
- Absorption
- Catalytic activity
- Cofactor
- Enzyme regulation
- Function[CC]¹
- Kinetics
- Pathway
- Redox potential
- Temperature dependence
- pH dependence
- Active site
- Binding site
- DNA binding
- Metal binding
- Nucleotide binding
- Site

▼ Miscellaneous

-  
- Annotation score
- Features
- Caution
- Miscellaneous[CC]²
- Keywords
- Matched text
- Protein existence
- Tools

▶ Interaction

▶ Expression

▶ Gene Ontology (GO)

▶ Pathology & Biotech

▶ Subcellular location

▶ PTM / Processing

▶ Structure

▶ Publications

▶ Date of

▶ Family & Domains

▶ Taxonomic lineage

▶ Taxonomy Id

Databases

▶ Sequence

▶ 3D structure

▶ Protein-protein interaction

▶ Chemistry

▶ Protein family/group

▶ PTM

▶ Polymorphism

▶ 2D gel

▶ Proteomic

▶ Protocols and materials

▶ Genome annotation

▶ Organism-specific

▶ Phylogenomic

▶ Enzyme and pathway

▶ Other

▶ Gene expression


▶ Family and domain

Submit a list of identifiers to retrieve the corresponding UniProt entries, or to map them from or to an external database

We are moving our web pages and services from HTTP to HTTPS. To transition to HTTPS, we will support separate HTTP and HTTPS services until June 20, 2018. From this date, the HTTP traffic will be automatically redirected to HTTPS. [More information or view this page using https](#)


UniProtKB results








Filter byⁱ

-  Reviewed (21) Swiss-Prot
-  Unreviewed (143) TrEMBL

Popular organisms

- Human (9)
- Fruit fly (6)
- Rat (5)
- Mouse (4)
- Zebrafish (3)
- Other organisms

Show only exact matches for p53 

Entry	Entry name	Protein names	Gene names	Organism	Length
<input type="checkbox"/> P04637	P53_HUMAN	 Cellular tumor antigen p53	TP53 P53	Homo sapiens (Human)	393
<input type="checkbox"/> P02340	P53_MOUSE	 Cellular tumor antigen p53	Tp53 P53, Trp53	Mus musculus (Mouse)	387
<input type="checkbox"/> P10361	P53_RAT	 Cellular tumor antigen p53	Tp53 P53	Rattus norvegicus (Rat)	391
<input type="checkbox"/> Q8IMZ4	Q8IMZ4_DROME	 p53 protein long form variant 1	p53 CG10873, CG31325, D-p53, Dm-P53, Dmel\CG33336 	Drosophila melanogaster (Fruit fly)	495
<input type="checkbox"/> Q9N6D8	Q9N6D8_DROME	 GH11591p	p53 CG10873, CG31325, D-p53, Dm-P53, Dmel\CG33336 	Drosophila melanogaster (Fruit fly)	385

Retrieve information on:

- P04637
- P02340
- P10361
- Q8IMZ4
- Q9N6D8

To improve security and privacy, we are moving our web pages and services from HTTP to HTTPS.
To give users of web services time to transition to HTTPS, we will support separate HTTP and HTTPS services until June 20, 2018.
From this date, the HTTP traffic will be automatically redirected to HTTPS.
[More information](#) or [view this page using https](#)


Retrieve/ID mapping

How to use this tool

Enter or upload a list of identifiers to do one of the following:

Retrieve the corresponding UniProt entries to download them or work with them on this website.

Convert identifiers which are of a different type to UniProt identifiers or vice versa and download the identifier lists.

1. Enter identifiers, separated by spaces or new lines, into the form field, for example: 





P31946 P62258

ALBU_HUMAN

EFTU_ECOLI

2. If you need to convert to another identifier type, select the source and target type from the dropdown menus.

3. Click the **Go** button.

 [Help](#)  [Help video](#)  [Other tutorials and videos](#)  [Downloads](#)

1. Provide your identifiers

P04637
P02340
P10361
Q8IMZ4
Q9N6D8

P04637
P02340
P10361
Q8IMZ4
Q9N6D8

OR upload your own file: 未選擇任何檔案

Run in a new window.

2. Select options

From

To

UniProtKB AC/ID ▾

UniProtKB ▾

To improve security and privacy, we are moving our web pages and services from HTTP to HTTPS.
 To give users of web services time to transition to HTTPS, we will support separate HTTP and HTTPS services until June 20, 2018.
 From this date, the HTTP traffic will be automatically redirected to HTTPS.
[More information](#) or [view this page using https](#)

UniProtKB results

[About Upload lists](#)


5 out of 5 UniProtKB AC/ID identifiers were successfully mapped to 5 UniProtKB IDs in the table below.

Filter byⁱ

Reviewed (3)
Swiss-Prot

Unreviewed (2)
TrEMBL

Popular organisms

Fruit fly (2)

Human (1)

Mouse (1)

Rat (1)

View by

	BLAST	Align	Download	Add to basket	Columns				
<input checked="" type="checkbox"/>	Your list:...	25DFE6	Entry	Entry name	Protein names	Gene names	Organism	Length	
<input type="checkbox"/>	P04637	P04637	P53_HUMAN		Cellular tumor antigen p53	TP53 P53	Homo sapiens (Human)	393	
<input type="checkbox"/>	P02340	P02340	P53_MOUSE		Cellular tumor antigen p53	Tp53 P53, Trp53	Mus musculus (Mouse)	387	
<input type="checkbox"/>	P10361	P10361	P53_RAT		Cellular tumor antigen p53	Tp53 P53	Rattus norvegicus (Rat)	391	
<input type="checkbox"/>	Q8IMZ4	Q8IMZ4	Q8IMZ4_DROME		p53 protein long form variant 1	p53 CG10873, CG31325, D-p53, Dm-P53, Dmel\CG33336	Drosophila melanogaster (Fruit fly)	495	
<input type="checkbox"/>	Q9N6D8	Q9N6D8	Q9N6D8_DROME		GH11591p	p53 CG10873, CG31325, D-p53, Dm-P53, Dmel\CG33336	Drosophila melanogaster (Fruit fly)	385	

1 to 5 of 5

Show 25

Unsaved columnsⁱ

Click on the **+** to add to 'Columns to be displayed' and to the 'Miscellaneous' section below for future use.

+ Your list:M20171228..567E7N **×**

Columns to be displayedⁱ

Drag and drop to re-order.

Reset to default **Save** Cancel

3

Entry Entry name Protein names Gene names Organism Length Active site Binding site Calcium binding 2
Catalytic activity Cofactor

Add more columnsⁱ

▶ Expand all

Search:

▼ Names & Taxonomy

- Entry name
- Gene names
- Gene names (ordered locus)
- Gene names (ORF)
- Gene names (primary)
- Gene names (synonym)
- Organism
- Organism ID
- Protein names
- Proteomes
- Taxonomic lineage
- Virus hosts



▼ Sequences

- Alternative products (isoforms)
- Alternative sequence
- Erroneous gene model prediction
- Fragment
- Gene encoded by
- Length
- Mass
- Mass spectrometry
- Natural variant
- Non-adjacent residues
- Non-standard residue
- Non-terminal residue
- Polymorphism
- RNA editing
- Sequence
- Sequence caution
- Sequence conflict
- Sequence uncertainty
- Sequence version

▼ Function

- Absorption
- Active site 1
- Binding site
- Calcium binding
- Catalytic activity
- Cofactor
- DNA binding
- EC number
- Enzyme regulation
- Function [CC]ⁱ
- Kinetics
- Metal binding
- Nucleotide binding
- Pathway
- pH dependence
- Redox potential
- Site
- Temperature dependence

▼ Miscellaneous

- Annotation
- Caution
- Features
- Isoform map:M20171228..25DFE6
- Isoform map:M20171228..25E4D1
- Isoform map:M20171228..5674BD
- Isoform map:M20171228..567E7N
- Keywords
- Matched text
- Miscellaneous [CC]ⁱ
- Protein existence
-  
- Tools
- UniParc
- Your list:M20171228..25DFE6
- Your list:M20171228..25E4D1
- Your list:M20171228..5674BD
- Your list:M20171228..567E7N

Entry	Entry name	Protein names	Gene names	Organism	Length	Active site	Binding site	Calcium binding	Catalytic activity	Cofactor
P04637	P53_HUMAN	Cellular tumor antigen p53	TP53 P53	Homo sapiens (Human)	393					Zn ²⁺ Note: Binds 1 zinc ion per subunit.
P02340	P53_MOUSE	Cellular tumor antigen p53	Tp53 P53, Trp53	Mus musculus (Mouse)	387					Zn ²⁺ Note: Binds 1 zinc ion per subunit.
P10361	P53_RAT	Cellular tumor antigen p53	Tp53 P53	Rattus norvegicus (Rat)	391					Zn ²⁺ By similarity Note: Binds 1 zinc ion per subunit. By similarity
Q8IMZ4	Q8IMZ4_DROME	p53 protein long form variant 1	p53 CG10873, CG31325, D-p53, Dm-P53, Dmel\CG33336	Drosophila melanogaster (Fruit fly)	495					
Q9N6D8	Q9N6D8_DROME	GH11591p	p53 CG10873, CG31325, D-p53, Dm-P53, Dmel\CG33336	Drosophila melanogaster (Fruit fly)	385					

Download Add to basket Columns

Download selected (0)
 Download all (5)

Format: **Tab-separated**

- FASTA (canonical)
- FASTA (canonical & isoform)
- Tab-separated**
- Text
- Excel
- GFF
- XML
- RDF/XML
- List

Preview

Download selected (0)
 Download all (5)

Format: **Tab-separated**

Compressed Uncompressed

Preview first 10¹

組合管理 新增資料夾

- > Streaming Video Recorder
- > Tencent Files
- > TRAVEL
- Tv30001
- Vaginal microbiome
- > Video Download Capture
- > WeChat Files
- 自訂 Office 範本
- 我的掃描
- > 刻錄菱袖
- > 掃描的文件
- > 音樂
- > 桌面

名稱	修改日期	類型	大小
4CSCP	2016/12/10 下午 ...	檔案資料夾	
Angiostrongylus_Sanger	2017/3/4 上午 08...	檔案資料夾	
Codon Usage	2014/12/27 下午 ...	檔案資料夾	
Desktop	2015/4/10 上午 0...	檔案資料夾	
Dropbox_backup 20170728	2017/7/28 上午 0...	檔案資料夾	
From installer	2017/7/19 下午 1...	檔案資料夾	
Manual	2014/1/6 下午 04...	檔案資料夾	
NGS	2016/4/17 下午 0...	檔案資料夾	
Parasite Systems Biology	2016/11/24 上午 ...	檔案資料夾	
PARASITE VIDEOS	2017/7/9 下午 05...	檔案資料夾	
Salary	2015/2/20 下午 0...	檔案資料夾	
sci_hub	2016/2/29 下午 0...	檔案資料夾	
STORY	2016/8/24 上午 1...	檔案資料夾	
temp	2016/5/21 下午 1	檔案資料夾	

檔案名稱(N): uniprot-yourlist%63AM20171228AAF7E4D2F1D05654627429E83DA5CCEE567E7N.tab

存檔類型(T): TAB 檔案 (.tab)

隱藏資料夾

存檔(S)

取消

Open excel and drag the file to it

The image shows a Windows File Explorer window on the left and a Microsoft Excel spreadsheet on the right. A red arrow points from a file named 'uniprot-yourlist%3AM20171228AAF87E4D2F1D05654627429E83DA5CCEE567E7N.tab' in the File Explorer to the Excel spreadsheet.

The File Explorer window shows the following files in the '桌面' (Desktop) folder:

- test.pds
- TRA203870.docx
- Tv_Protein Kinase.xlsx
- TV_varscan_snpeff.xlsx
- Tv_WGS_QC_summary170528.xlsx
- Tv_WGS_QC_summary170531.xlsx
- TV.xlsx
- TvWGS_mapped reads.pptx
- uniprot-yourlist%3AM20171228AAF87E4D2F1D05654627429E83DA5CCEE567E7N.tab
- WeChat
- WhatsApp
- 文件 - 捷徑
- 台灣人體微生物相研究及發展計畫構想書.docx
- 交通車.jpg
- 百度网盘
- 吳穎奕.pdf
- 沖繩也有溫泉？美麗海水族館周邊&那霸市內優質溫泉度假飯店精選！ - 溫泉達人.pdf
- 私校退撫儲蓄試算.xls
- 擬聘請專家情況表_鄧致剛.docx
- 擬聘請專家情況表_辛致偉.docx
- 旅平卡DM.pdf
- 泰式料理即食包.xlsx
- 記事本
- 退休金.xlsx
- 退撫儲蓄試算表.xls
- 高通量定序在生物醫學研究之應用及市場.pdf
- 基因銀行.pptx
- 教師參與系統院務行政服務(含輔導)綜合表現評核表.docx
- 處次丙申.pptx
- 遠端桌面重新導向印表機文件.pdf
- 遠端桌面連線
- 影音先鋒
- 影像中心

The Excel spreadsheet shows the following data:

Entry	Entry nam	Status	Protein nai	Gene nam	Organism	Length	Active site	Binding sit	Calcium bi	Catalytic a	Cofa
P04637	P53_HUM	reviewed	Cellular tur	TP53	P53	Homo sapi	393				COF
P02340	P53_MOU	reviewed	Cellular tur	TP53	P53	Mus musc	387				COF
P10361	P53_RAT	reviewed	Cellular tur	TP53	P53	Rattus nor	391				COF
Q8IMZ4	Q8IMZ4_D	unreviewec	p53	protein p53	CG10f	Drosophila	495				COF
Q9N6D8	Q9N6D8_L	unreviewec	GH11591p	p53	CG10f	Drosophila	385				

P04637 - P53_HUMAN

Protein	Cellular tumor antigen p53
Gene	TP53
Organism	<i>Homo sapiens (Human)</i>
Status	Reviewed - ●●●●● - Experimental evidence at protein level ⁱ

 Display None

- [BLAST](#)
- [Align](#)
- [Format](#)
- [Add to basket](#)
- [History](#)

- [Comment \(0\)](#)
- [Feedback](#)
- [Help video](#)

- FUNCTION
- NAMES & TAXONOMY
- SUBCELLULAR LOCALIZATION
- PATHOLOGY & BIOTECHNOLOGY
- PTM / PROCESSING
- EXPRESSION
- INTERACTION
- STRUCTURE
- FAMILY & DOMAINS**
- SEQUENCES (9)
- CROSS-REFERENCES
- PUBLICATIONS
- ENTRY INFORMATION
- MISCELLANEOUS

Functionⁱ

Acts as a tumor suppressor in many tumor types; induces growth arrest or apoptosis depending on the physiological circumstances and cell type. Involved in cell cycle regulation as a trans-activator that acts to negatively regulate cell division by controlling a set of genes required for this process. One of the activated genes is an inhibitor of cyclin-dependent kinases. Apoptosis induction seems to be mediated either by stimulation of BAX and FAS antigen expression, or by repression of Bcl-2 expression. In cooperation with mitochondrial PPIF is involved in activating oxidative stress-induced necrosis; the function is largely independent of transcription. Induces the transcription of long intergenic non-coding RNA p21 (lincRNA-p21) and lincRNA-Mkl1. LincRNA-p21 participates in TP53-dependent transcriptional repression leading to apoptosis and seem to have to effect on cell-cycle regulation. Implicated in Notch signaling cross-over. Prevents CDK7 kinase activity when associated to CAK complex in response to DNA damage, thus stopping cell cycle progression. Isoform 2 enhances the transactivation activity of isoform 1 from some but not all TP53-inducible promoters. Isoform 4 suppresses transactivation activity and impairs growth suppression mediated by isoform 1. Isoform 7 inhibits isoform 1-mediated apoptosis. 11 Publications

Cofactorⁱ

Binds 1 zinc ion per subunit.

Sites

Feature key	Position(s)	Length	Description	Graphical view	Feature identifier	Actions
Site ⁱ	120 – 120	1	Interaction with DNA			
Metal binding ⁱ	176 – 176	1	Zinc			
Metal binding ⁱ	179 – 179	1	Zinc			
Metal binding ⁱ	238 – 238	1	Zinc			
Metal binding ⁱ	242 – 242	1	Zinc			

Display

Entry

Publications

Feature viewer

Feature table

All None

Function

Names & Taxonomy

Subcellular location

Pathology & Biotech

PTM / Processing

Expression

Interaction

Structure

Family & Domains

Sequences (9)

Similar proteins

Cross-references

Entry information

Miscellaneous

[▲ Top](#)

Sequence similaritiesⁱ

Belongs to the p53 family. [Curated](#)

Phylogenomic databases

eggNOG ⁱ	ENOG410IITK. Eukaryota. ENOG410ZSWV. LUCA.
GeneTree ⁱ	ENSGT00390000015092.
HOVERGEN ⁱ	HBG005201.
InParanoid ⁱ	P04637.
KO ⁱ	K04451.
OMA ⁱ	PATSWPL.
OrthoDB ⁱ	EOG091G0XY5.
PhylomeDB ⁱ	P04637.
TreeFam ⁱ	TF106101.

Family and domain databases

CDD ⁱ	cd08367. P53. 1 hit.
Gene3D ⁱ	2.60.40.720. 1 hit. 4.10.170.10. 1 hit.
InterPro ⁱ	View protein in InterPro IPR008967. p53-like_TF_DNA-bd. IPR012346. p53/RUNT-type_TF_DNA-bd_sf. IPR011615. p53_DNA-bd. IPR036674. p53_tetramer_sf. IPR010991. p53_tetrameristn. IPR013872. p53_transactivation_domain. IPR002117. p53_tumour_suppressor.
PANTHER ⁱ	PTHR11447. PTHR11447. 1 hit.
Pfam ⁱ	View protein in Pfam PF00870. P53. 1 hit. PF08563. P53_TAD. 1 hit. PF07710. P53_tetramer. 1 hit.
PRINTS ⁱ	PR00386. P53SUPPRESSR.
SUPFAM ⁱ	SSF47719. SSF47719. 1 hit. SSF49417. SSF49417. 1 hit.
PROSITE ⁱ	View protein in PROSITE PS00348. P53. 1 hit.



PROSITE documentation PDOC00301 [for PROSITE entry PS00348]

p53 family signature

[Description](#) [Technical section](#) [References](#) [Copyright](#) [Miscellaneous](#)

Description

The p53 tumor antigen [1,2,3,4,5] is a protein found in increased amounts in a wide variety of transformed cells. It is also detectable in many proliferating nontransformed cells, but it is undetectable or present at low levels in resting cells. It is frequently mutated or inactivated in many types of cancer. p53 seems to act as a tumor suppressor in some, but probably not all, tumor types. p53 is probably involved in cell cycle regulation, and may be a trans-activator that acts to negatively regulate cellular division by controlling a set of genes required for this process.

p53 is a phosphoprotein of about 390 amino acids which can be subdivided into four domains: a highly charged acidic region of about 75 to 80 residues, a hydrophobic proline-rich domain (position 80 to 150), a central region (from 150 to about 300), and a highly basic C-terminal region. The sequence of p53 is well conserved in vertebrate species; attempts to identify p53 in other eukaryotic phylum has so far been unsuccessful.

The p53 protein belongs to a family [6] that also includes:

- p51 (p63 or Ket), a transcriptional activator.
- p73, a transcriptional activator that participates in the apoptotic response to DNA damage.

As a signature pattern for this family we selected a conserved stretch of 13 residues located in the central region of the protein. This region, known as domain IV in [3], is involved (along with an adjacent region) in the binding of the large T antigen of SV40. In man this region is the focus of a variety of point mutations in cancerous tumors.

Last update:

December 2001 / Pattern and text revised.

Technical section

PROSITE method (with tools and information) covered by this documentation:

P53, PS00348; p53 family signature (PATTERN)

- Consensus pattern:
M-C-N-S-S-C-[MV]-G-G-M-N-R-R
- Sequences in UniProtKB/Swiss-Prot known to belong to this class: 40
 - detected by PS00348: 40 (true positives)
 - undetected by PS00348: 0 (false negative or 'partial')
- Other sequence(s) in UniProtKB/Swiss-Prot detected by PS00348:
NONE.
- Retrieve an alignment of UniProtKB/Swiss-Prot true positive hits:
[Clustal format, color, condensed view](#) / [Clustal format, color](#) / [Clustal format, plain text](#) / [Fasta format](#)
- Retrieve the sequence logo from the alignment
- Taxonomic tree view of all UniProtKB (Swiss-Prot + TrEMBL) entries matching PS00348
- Retrieve a list of all UniProtKB (Swiss-Prot + TrEMBL) entries matching PS00348
- Scan UniProtKB (Swiss-Prot and/or TrEMBL) entries against PS00348
- View ligand binding statistics of PS00348
- Matching PDB structures: 1GZH 1HU8 1KZY 1TSR ... [ALL]



PROSITE documentation PDOC00301 [for PROSITE entry PS00348]

p53 family signature

[Description](#) [Technical section](#) [References](#) [Copyright](#) [Miscellaneous](#)

Description

The p53 tumor antigen [1,2,3,4,5] is a protein found in increased amounts in a wide variety of transformed cells. It is also detectable in many proliferating nontransformed cells, but it is undetectable or present at low levels in resting cells. It is frequently mutated or inactivated in many types of cancer. p53 seems to act as a tumor suppressor in some, but probably not all, tumor types. p53 is probably involved in cell cycle regulation, and may be a trans-activator that acts to negatively regulate cellular division by controlling a set of genes required for this process.

p53 is a phosphoprotein of about 390 amino acids which can be subdivided into four domains: a highly charged acidic region of about 75 to 80 residues, a hydrophobic proline-rich domain (position 80 to 150), a central region (from 150 to about 300), and a highly basic C-terminal region. The sequence of p53 is well conserved in vertebrate species; attempts to identify p53 in other eukaryotic phylum has so far been unsuccessful.

The p53 protein belongs to a family [6] that also includes:

- p51 (p63 or Ket), a transcriptional activator.
- p73, a transcriptional activator that participates in the apoptotic response to DNA damage.

As a signature pattern for this family we selected a conserved stretch of 13 residues located in the central region of the protein. This region, known as domain IV in [3], is involved (along with an adjacent region) in the binding of the large T antigen of SV40. In man this region is the focus of a variety of point mutations in cancerous tumors.

Last update:

December 2001 / Pattern and text revised.

Technical section

PROSITE method (with tools and information) covered by this documentation:

P53, PS00348; p53 family signature (PATTERN)

- Consensus pattern:
M-C-N-S-S-C-[MV]-G-G-M-N-R-R
- Sequences in UniProtKB/Swiss-Prot known to belong to this class: 40
 - detected by PS00348: 40 (true positives)
 - undetected by PS00348: 0 (false negative or 'partial')
- Other sequence(s) in UniProtKB/Swiss-Prot detected by PS00348:
NONE.
- Retrieve an alignment of UniProtKB/Swiss-Prot true positive hits:
[Clustal format, color, condensed view](#) / [Clustal format, color](#) / [Clustal format, plain text](#) / [Fasta format](#)
- Retrieve the sequence logo from the alignment
- Taxonomic tree view of all UniProtKB (Swiss-Prot + TrEMBL) entries matching PS00348
- Retrieve a list of all UniProtKB (Swiss-Prot + TrEMBL) entries matching PS00348
- Scan UniProtKB (Swiss-Prot and/or TrEMBL) entries against PS00348
- View ligand binding statistics of PS00348
- Matching PDB structures: 1GZH 1HU8 1KZY 1TSR ... [ALL]



Database of protein domains, families and functional sites

PROSITE consists of documentation entries describing protein domains, families and functional sites as well as associated patterns and profiles to identify them [[More...](#) / [References](#) / [Commercial users](#)].

PROSITE is complemented by [ProRule](#), a collection of rules based on profiles and patterns, which increases the discriminatory power of profiles and patterns by providing additional information about functionally and/or structurally critical amino acids [[More...](#)].

Release 2017_12 of 20-Dec-2017 contains 1797 documentation entries, 1309 patterns, 1202 profiles and 1223 ProRule.

Search

e.g. PDOC00022, PS50089, SH3, zinc finger

Browse

- by documentation entry
- by ProRule description
- by taxonomic scope
- by number of positive hits

Quick Scan mode of ScanProsite

Quickly find matches of your protein sequences to PROSITE signatures (max. 10 sequences). [\[?\]](#) [Examples](#)

Enter UniProtKB accessions or identifiers or PDB identifiers or sequences in FASTA format

Exclude motifs with a high probability of occurrence from the scan

For more scanning options go to [ScanProsite](#)

Other tools

- [PRATT](#) - allows to interactively generate conserved patterns from a series of unaligned proteins.
- [MyDomains - Image Creator](#) - allows to generate custom domain figures.





Prosite search results

Search in PROSITE for: casein kinase

(Release 20.76, of 17-Oct-2011)

Enter search terms:

casein kinase

Prefix and append wildcard '*' to words.

By default, this search engine searches for complete words only. If you did not find what you would try to do a substring match, you should perform a new search and select 'prefix and append wildcard '*' to words'.

Number of documents in PROSITE containing the search term:2

- [PDOC00006](#) Casein kinase II phosphorylation site
- [PDOC00845](#) Casein kinase II regulatory subunit signature

Casein kinase II phosphorylation site

Description:

Casein kinase II (CK-2) is a protein serine/threonine kinase whose activity is independent of cyclic nucleotides and calcium. CK-2 phosphorylates many different proteins. The substrate specificity [1] of this enzyme can be summarized as follows:

- (1) Under comparable conditions Ser is favored over Thr.
- (2) An acidic residue (either Asp or Glu) must be present three residues from the C-terminal of the phosphate acceptor site.
- (3) Additional acidic residues in positions +1, +2, +4, and +5 increase the phosphorylation rate. Most physiological substrates have at least one acidic residue in these positions.
- (4) Asp is preferred to Glu as the provider of acidic determinants.
- (5) A basic residue at the N-terminal of the acceptor site decreases the phosphorylation rate, while an acidic one will increase it.

Note:

This pattern is found in most of the known physiological substrates.

Last update:

May 1991 / Text revised.

Technical section:

PROSITE method (with tools and information) covered by this documentation:

CK2_PHOSPHO_SITE, PS00006; **Casein kinase II phosphorylation site** (PATTERN with a high probability of occurrence!)

Consensus pattern: [ST]-x(2)-[DE][S or T is the phosphorylation site]

- [Scan Swiss-Prot/TrEMBL entries against PS00006](#)
- [view ligand binding statistics](#)

Casein kinase II regulatory subunit signature

Description:

Casein kinase II (CK-2) [1] is an ubiquitous eukaryotic serine/threonine protein kinase which is found both in the cytoplasm and the nucleus and whose substrates are numerous. It generally phosphorylates Ser or Thr at the N-terminal of stretch of acidic residues (see <PDOC00006>). CK-2 exists as an heterotetramer composed of two catalytic subunits (α) and two regulatory subunits (β). In most species there are two closely related isoforms of the catalytic subunit: α and α' . Some species, such as fungi and plants, express two forms of regulatory subunits: β and β' . The β subunit plays a complex role in regulating the basal catalytic activity of the α subunit. The β subunit is a highly conserved protein of about 25 Kd that contains, in its central section, a cysteine-rich motif that could be involved in binding a metal such as zinc [2]. We have used this region as a signature pattern.

Last update:

December 2004 / Pattern and text revised.

Technical section:

PROSITE method (with tools and information) covered by this documentation:

CK2_BETA, PS01101; **Casein kinase II regulatory subunit signature** (PATTERN)

Consensus pattern:

C-P-x-[LIVMYAT]-x-C-x(5)-[LI]-P-[LIVMCA]-G-x(9)-V-[KRM]-x(2)-C-[PA]-x-C

Sequences known to belong to this class detected by the pattern: ALL

Other sequence(s) detected in Swiss-Prot: NONE

- Retrieve an alignment of Swiss-Prot true positive hits:
[Clustal format, color, condensed view](#) / [Clustal format, color](#) / [Clustal format, plain text](#) / [Fasta format](#)
- [Retrieve the sequence logo from the alignment](#)
- [Taxonomic tree view of all Swiss-Prot/TrEMBL entries matching PS01101](#)
- [Retrieve a list of all Swiss-Prot/TrEMBL entries matching PS01101](#)
- [Scan Swiss-Prot/TrEMBL entries against PS01101](#)
- [view ligand binding statistics](#)



<http://prosite.expasy.org/>

Database of protein domains, families and functional sites

PROSITE consists of [documentation entries](#) describing protein domains, families and functional sites as well as associated [patterns](#) and [profiles](#) to identify them [[More details](#) / [References](#) / [Commercial users](#)].

PROSITE is complemented by [ProRule](#), a collection of rules based on profiles and patterns, which increases the discriminatory power of profiles and patterns by providing additional information about functionally and/or structurally critical amino acids [[More details](#)].

Release 20.76, of 17-Oct-2011 (1627 documentation entries, 1308 patterns, 946 profiles and 944 ProRule)

PROSITE access

e.g: PDOC00022, PS50089, SH3, zinc finger

Browse:

add wildcard **

- [by documentation entry](#)
- [by ProRule description](#)
- [by taxonomic scope](#)
- [by number of positive hit](#)

PROSITE tools

Scan a sequence against PROSITE patterns and profiles - quick scan

(Output includes graphical view and feature detection)



Enter your sequence or a UniProtKB (Swiss-Prot or TrEMBL) ID or AC [[help](#)]:

exclude [patterns with a high probability of occurrence](#)

- [ScanProsite](#) - advanced scan
- [PRATT](#) - allows to interactively generate conserved patterns from a series of unaligned proteins.
- [MyDomains - Image Creator](#) ^{new} - allows to generate custom domain figures.



Sequences (9)¹

Sequence status¹: Complete.

This entry describes 9 isoforms¹ produced by **alternative promoter usage** and **alternative splicing**. [Align](#)

Isoform 1 (identifier: **P04637-1**) [UniParc] [FASTA](#) [Add to Basket](#)

Also known as: p53, p53alpha

This isoform has been chosen as the 'canonical' sequence. All positional information in this entry refers to it. This is also the sequence that appears in the downloadable versions of the entry.

[View](#)

```
>sp|P04637|P53_HUMAN Cellular tumor antigen p53 OS=Homo sapiens GN=TP53 PE=1 SV=4
MEEPQSDPSVEPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDIEQWFTEDPGP
DEAPRMPEAAAPPVAPAPAAPTAAPAPAPSWPLSSSVPSQKTYQGSYGFRLGFLHSGTAK
SVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRRAAIYKQSHMTEVRRCPHHE
RCSDSGLAPPQHLIRVEGNLRVEYLDNRNTRFRHSVVVYPPEVEGSDCTTIHYNYMCNS
SCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRCACPGDRRTEENLRKKGEPHHELP
PGSTKRALPNNTSSSPQPKKPLDGEYFTLQIRGRERFEMFRELNEALELKDAQAGKEPG
GSAHSSHLKSKKGQSTSRHKLMFKTEGPDSD
```

Copy this sequence to your computer

FS	DLWKLLEN	VLSPLPSQAM	DDLMLSPDDI	50
AA	PPVAPAPAAP	TPAAPAPAPS	WPLSSSVPSQ	100
AK	SVTCTYSPAL	NKMFCQLAKT	CPVQLWVDST	150
PPGTRVRRAAI	YKQSHMTEVRRCPHHE	EVRRCRPHHE	RCSDSGLAP	200
LRVEYLDNRN	TFRHSVVVYP	EPPEVGS DCT	TIHYNYMCNS	250
ILTIITLED	SGNLLGRNSF	EVRVACACPG	DRRTEENLR	300
PGSTKRALPN	NTSSSPQPKK	KPLDGEYFTL	QIRGRERFEM	350
KDAQAGKEPG	GSAHSSHLK	SKKGQSTSRH	KKLMFKTEGP	393

Length: 393
Mass (Da): 43,653
Last modified: November 24, 2009 - v4
Checksum: AD5C149FD8106131

[GO](#)

Isoform 2 (identifier: **P04637-2**) [UniParc] [FASTA](#) [Add to Basket](#)

Also known as: I9RET, p53beta

The sequence of this isoform differs from the canonical sequence as follows:

332-341: IRGRERFEMF → DQTSFQKENC

342-393: Missing.

Note: Expressed in quiescent lymphocytes. Seems to be non-functional. May be produced at very low levels due to a premature stop codon in the mRNA, leading to nonsense-mediated mRNA decay.

[Show](#) »

Length: 341
Mass (Da): 37,826
Checksum: 2C5FE7A14A575E43

[GO](#)

Isoform 3 (identifier: **P04637-3**) [UniParc] [FASTA](#) [Add to Basket](#)

Also known as: p53gamma

The sequence of this isoform differs from the canonical sequence as follows:

332-346: IRGRERFEMFRELNE → MLLDLRWYFLINSS

347-393: Missing.

Note: Expressed in quiescent lymphocytes. Seems to be non-functional. May be produced at very low levels due to a premature stop codon in the mRNA, leading to nonsense-mediated mRNA decay.

[Show](#) »

Length: 346
Mass (Da): 38,501
Checksum: 6F18E09F8CD9129F

[GO](#)

Example

cellular tumor antigen p53 isoform a [Homo sapiens]

```
>sp|P04637|P53_HUMAN Cellular tumor antigen p53 OS=Homo sapiens GN=TP53 PE=1 SV=4
MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIEQWFTEDPGP
DEAPRMPEAAPPVAPAPAAPTPAAPAPAPSWPLSSSVPSQKTYQGSYGFRLGFLHSGTAK
SVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAMAIYKQSQHMTFVRRCPHHE
RCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVVPYEPPEVGSDCCTTIHYNMCNS
SCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCACPGRDRRTEENLRKKGEPHHELP
PGSTKRALPNNTSSSPQPKKPLDGEYFTLQIRGRERFEMFRELNEALELKDAQAGKEPG
GSAHSSHLKSKKGQSTSRHKKLMFKTEGPDS
```

How to use this tool

The Basic Local Alignment Search Tool (BLAST) finds regions of local similarity between sequences, which can be used to infer functional and evolutionary relationships between sequences as well as help identify members of gene families.

1. Enter either a protein or nucleotide sequence or a UniProt identifier (e.g. P00750 or A4_HUMAN or UPI0000000001) into the form field.
2. Optionally, change the program parameters with the dropdown menus under the form.
3. Click the *Run BLAST* button.

Help Tutorials and Videos Downloads

BLAST

```
MEEPQSDPSVEPPLSQETFSDLWKLLENNVLSPLPSQAMDDLMSDDIEQWFTEDPGP
DEAPRMPPEAAPFVAPAPAAPTFAAPAPAPSNPLSSSVPSQKTYQGSYGFRLGFLHSGIAK
SVTCTYSPALNKMFCOLAKTCFVQLWVDSTPEPGRVVRAMAIYKQSOHMTVEVVRCPHHE
RCSDSDGLAPPQHLIRVEGNLRVEYLLDRNTFRHSVVPYEPPEVGSDCCTIIHYNMNCNS
SCMGGMNRRPILTIITILEDSSGNLLGRNSFEVRVCACPGDRDRTEENLRKKGEPPHHELP
FGSTKRALPNNTSSSPQPKKKFLDGEYFTLQIRGRERFEMFRELNEALELKDAQAGKEPG
G5RAHSSHLKSKKGQSTSRHKKLMFKTEGPDSD
```

 Target databaseⁱ

UniProtKB

 E-Thresholdⁱ

10

 Matrixⁱ

Auto

 Filteringⁱ

None

 Gappedⁱ

yes

 Hitsⁱ

250

 Run Blast in a separate window.

BLAST Job status: RUNNING

→ Running Blast job against UNIPROTKB for 12s

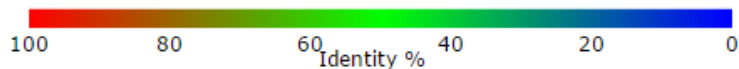
Job Information

Query sequence

```
MEEPQSDPSVEPPLSQETFSDLWKLLENNVLSPLPSQAMDDLMSDDIEQWFTEDPGP
DEAPRMPPEAAPFVAPAPAAPTFAAPAPAPSNPLSSSVPSQKTYQGSYGFRLGFLHSGIAK
SVTCTYSPALNKMFCOLAKTCFVQLWVDSTPEPGRVVRAMAIYKQSOHMTVEVVRCPHHE
RCSDSDGLAPPQHLIRVEGNLRVEYLLDRNTFRHSVVPYEPPEVGSDCCTIIHYNMNCNS
SCMGGMNRRPILTIITILEDSSGNLLGRNSFEVRVCACPGDRDRTEENLRKKGEPPHHELP
FGSTKRALPNNTSSSPQPKKKFLDGEYFTLQIRGRERFEMFRELNEALELKDAQAGKEPG
G5RAHSSHLKSKKGQSTSRHKKLMFKTEGPDSD
```

Job Identifier	920141034FC00T93RY (Jobs are stored for 7 days)
Program	blastp
Matrix	Blosum62
Threshold	10
Filtered	false
Gapped	true
Maximum number of hits reported	250

BLAST



Basket

Filter byⁱ

Reviewed (59)
Swiss-Prot

Unreviewed (180)
TrEMBL

With 3D structure (4)

Proteomes (100)

Organisms

Human (51)

Mouse (18)

Zebrafish (7)

Rat (4)

Bovine (3)

Other organisms

Go

Map To

UniProtKB

UniRef

UniParc

View by

Taxonomy

Text version

XML version

Demo

Help video

[← Edit and resubmit](#)

Order by: [Score](#) Taxonomy: [All](#)

Overview

[Show all 239](#)

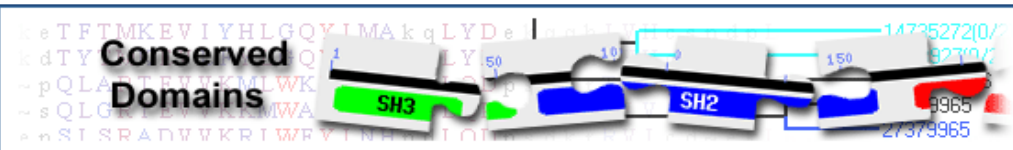
Entry	Protein names	Match hit	Identity
H2QC53	Cellular tumor antigen p53 (Pan troglodytes)		100.0%
K7PPA8	Cellular tumor antigen p53 (Homo sapiens)		100.0%
P04637	Cellular tumor antigen p53 (Homo sapiens)		100.0%
G3R2U9	Cellular tumor antigen p53 (Gorilla gorilla gorilla)		99.0%

Alignments

[Columns](#) [BLAST](#) [Align](#) [Download](#) [Add to basket](#)

1 to 25 of 239 Show [25](#)

Entry	Alignment overview	Info	Status
Query: B20141016F000TRFJRY			
H2QC53	H2QC53_PANTR - Cellular tumor antigen p53 - Pan troglodytes ... - View alignment 	E-value: 0.0 Score: 2,121 Ident.: 100.0%	
K7PPA8	K7PPA8_HUMAN - Cellular tumor antigen p53Homo sapiens (Human) - View alignment 	E-value: 0.0 Score: 2,121 Ident.: 100.0%	
P04637	P53_HUMAN - Cellular tumor antigen p53Homo sapiens (Human) - View alignment 	E-value: 0.0 Score: 2,121 Ident.: 100.0%	



Conserved Domains and Protein Classification

RESOURCES **SEARCH** HOW TO HELP NEWS FTP PUBLICATIONS DISCOVER

Resources

Conserved Domain Database (CDD)

CDD is a protein annotation resource that consists of a collection of well-annotated multiple sequence alignment models for ancient domains and full-length proteins. These are available as position-specific score matrices (PSSMs) for fast identification of conserved domains in protein sequences via RPS-BLAST. CDD content includes NCBI-curated domains, which use 3D-structure information to explicitly define domain boundaries and provide insights into **sequence/structure/function relationships**, as well as domain models imported from a number of external source databases (Pfam, SMART, COG, PRK, TIGRFAM).
[Search](#) | [How To](#) | [Help](#) | [News](#) | [FTP](#) | [Publications](#)

CD-Search & Batch CD-Search

CD-Search is NCBI's interface to searching the Conserved Domain Database with **protein or nucleotide query sequences**. It uses RPS-BLAST, a variant of PSI-BLAST, to quickly scan a set of pre-calculated position-specific scoring matrices (PSSMs) with a protein query. The **results** of CD-Search are presented as an annotation of protein domains on the user query sequence (**illustrated example**), and can be visualized as domain **multiple sequence alignments** with **embedded user queries**. High confidence associations between a query sequence and conserved domains are shown as **specific hits**. The **CD-Search Help** provides additional details, including information about **running CD-Search locally**.

Batch CD-Search serves as both a web application and a **script interface** for a conserved domain search on **multiple protein sequences**, accepting up to 100,000 proteins in a single job. It enables you to view a **graphical display** of the concise or full search result for any individual protein from your input list, or to **download** the results for the complete set of proteins. The **Batch CD-Search Help** provides additional details.

[CD-Search \(Help & FTP\)](#) | [Batch CD-Search \(Help\)](#) | [Publications](#)

CDART: Domain Architectures

Conserved Domain Architecture Retrieval Tool (CDART) performs similarity searches of the **Entrez Protein** database based on domain architecture, defined as the sequential order of conserved domains in protein queries. CDART finds protein similarities across significant evolutionary distances using sensitive domain profiles rather than direct sequence similarity. Proteins similar to the query are grouped and scored by architecture. You can search CDART directly with a query protein sequence, or, if a sequence of interest is already in the Entrez Protein database, simply retrieve the record, open its **"Links"** menu, and select **"Domain Relatives"** to see the precalculated CDART results (**illustrated example**). Relying on domain profiles allows CDART to be fast and, because it relies on annotated functional domains, informative.

[About](#) | [Search](#) | [Help](#) | [FTP](#) | [Publications](#)

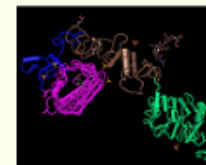
CDTree

CDTree is a helper application for your web browser that allows you to interactively view and examine conserved domain hierarchies curated at NCBI. CDTree works with Cn3D as its alignment viewer/editor, it is used in the CDD curation process and is a both **classification and research tool** for functional annotation and the study of protein and protein domain families.

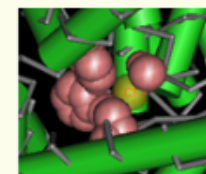
[About](#) | [Install](#) | [Publications](#)

Highlights

What is a conserved domain?



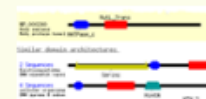
3-D structures and conserved core motifs:



Conserved features (binding and catalytic sites)



Proteins with Similar Domain Architectures



Search Methods: Quick Start Guide

Text Term Search

Retrieve [conserved domain](#) records that contain a term(s) of interest (e.g., chloride channel). Enter the terms in the query box at the top of this page or use the [Entrez Conserved Domains Database \(CDD\)](#) home page. See the help document for [search tips](#), including a list of available [search fields](#) and examples of their use.

Protein or Nucleotide Query Sequence

Enter a [protein or nucleotide query](#) as an [accession or GI](#) number (e.g., AAC50285 or 463989), or as a sequence in [FASTA](#) format, to identify the protein's conserved domains and therefore its putative function:

Search Database **CDD v3.12 - 46675 PSSMs** ▾

```
PGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRGRERFEMFRELN
EALELKDAQAGKEPG
GSRAHSSHLLKSKKGQSTSRHKKLMFKTEGPDSD
```

The text box above provides a short cut to the [CD-Search](#) tool, using its default parameters. The [help document](#) provides details about [database selection](#) and [search results](#). To view/change the default parameters, or to use [advanced search options](#), enter your query directly on the [CD-Search](#) home page. (Note: If a sequence of interest is already in the [Entrez Protein database](#), you can simply follow the ["Conserved Domains"](#) link for that sequence record to view its pre-computed CD-Search results.)

Batch of Protein Sequences

The [Batch CD-Search](#) tool allows the computation and download of conserved domain annotation for large sets of protein queries. [Input](#) up to 100,000 protein query sequences as a list of sequence identifiers and/or raw sequence data, then download [output](#) in a variety of formats (including tab-delimited text files) or [view the search results graphically](#). See the [help document](#) for additional details, including information on using Batch CD-Search for [scripted data downloads](#).

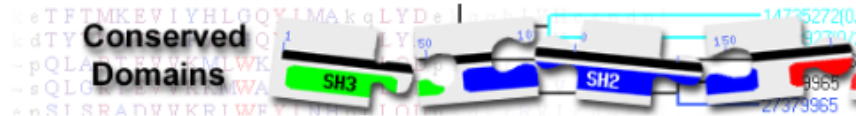
Direct fetch via UID

Retrieve a conserved domain record directly from the backend database by entering its unique identifier (UID), in the form of an [accession](#) (e.g., cd00400) or [PSSM ID](#) (e.g., 79359), in the text box below:

(Note: the "text term search" function also allows you to enter either of those unique identifiers (UIDs), but it first searches the Entrez indices for the UID, then retrieves the record. The "direct fetch via UID" option bypasses the Entrez indices and simply retrieves the specified record.)

Find proteins with similar domain architectures

Enter a [protein query](#) as an [accession or GI](#) number (e.g., AAC50285 or 463989), or as a sequence in [FASTA](#) format, on the [Conserved Domain Architecture Retrieval Tool \(CDART\)](#) page to find other proteins with similar domain architectures. (Note: If a sequence of interest is already in the [Entrez Protein database](#), you can simply select ["Domain Relatives"](#) in that sequence record's ["Links"](#) menu to find proteins with similar architecture.)

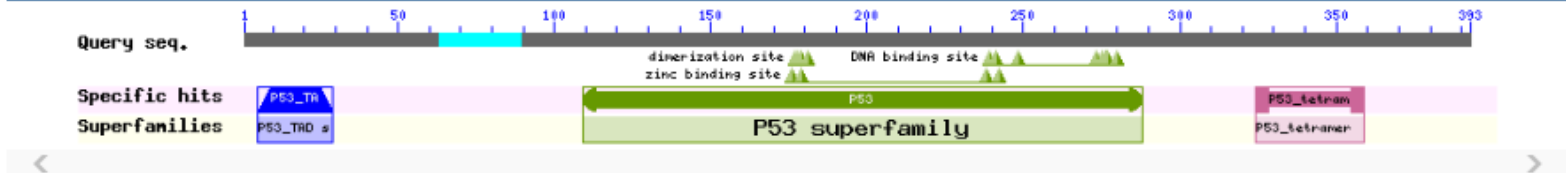


Conserved domains on [lc1|local_MEEPQSDPSV]

View Concise Results ?

Local query sequence

Graphical summary show options ?



List of domain hits ?

Name	Accession	Description	Interval	E-value
[+] P53	cd08367	P53 DNA-binding domain; P53 is a tumor suppressor gene product; mutations in p53 or lack of ...	109-288	6.36e-96
<p>P53 DNA-binding domain; P53 is a tumor suppressor gene product; mutations in p53 or lack of expression are found associated with a large fraction of all human cancers. P53 is activated by DNA damage and acts as a regulator of gene expression that ultimately blocks progression through the cell cycle. P53 binds to DNA as a tetrameric transcription factor. In its inactive form, p53 is bound to the ring finger protein Mdm2, which promotes its ubiquitinylation and subsequent proteosomal degradation. Phosphorylation of p53 disrupts the Mdm2-p53 complex, while the stable and active p53 binds to regulatory regions of its target genes, such as the cyclin-kinase inhibitor p21, which complexes and inactivates cdk2 and other cyclin complexes.</p>				
<p>Pssm-ID: 176262 Cd Length: 179 Bit Score: 286.86 E-value: 6.36e-96</p> <pre> 10 20 30 40 50 60 70 80 lc1 local_MEEPQSDPSV 109 FRLGFLHSGTAKSVICTYSPALNKMFCQLAKTICPVQLWVDSTPPPGRVVRAMAIYKQSQHMTFVVRRCPHERCsDSGDL 188 Cdd:cd08367 1 FEVTLDESGVAKSSTWITYSPKLNKLFVKMAKTCPIQFKVNSPPPGLVVRAMLVYKDPEHVKEPVERCPNHRQG-DDGHT 79 90 100 110 120 130 140 150 160 lc1 local_MEEPQSDPSV 189 APPQHILIRVEgNLRVEYLLDRNTFRHSVVVPYEPPEVGSDCCTTIHYNYMCSNCSMGGMNRRPILTIITLEDSSGNLLGRN 268 Cdd:cd08367 80 APNSHVIRCE-NPQAEYVGDFTGRLSVVVPLEPPQVGSEYVTVLLQFMCQNSCPGGINRRRPIQLVFTLEDENGNVLGR 158 170 180 lc1 local_MEEPQSDPSV 269 SFEVRCACPRDRRTEEN 288 Cdd:cd08367 159 VIEVRCACPRDRKNEEKA 178 </pre>				
[+] P53_tetramer	pfam07710	P53 tetramerization motif;	324-359	1.34e-15
[+] P53_TAD	pfam08563	P53 transactivation motif; The binding of the p53 transactivation domain by regulatory ...	5-29	5.36e-07



P53 DNA-binding domain

P53 is a tumor suppressor gene product; mutations in p53 or lack of expression are found associated with a large fraction of all human cancers. P53 is activated by DNA damage and acts as a regulator of gene expression that ultimately blocks progression through the cell cycle. P53 binds to DNA as a tetrameric transcription factor. In its inactive form, p53 is bound to the ring finger protein Mdm2, which promotes its ubiquitinylation and subsequent proteosomal degradation. Phosphorylation of p53 disrupts the Mdm2-p53 complex, while the stable and active p53 binds to regulatory regions of its target genes, such as the cyclin-kinase inhibitor p21, which complexes and inactivates cdk2 and other cyclin complexes.

Links

?

Source: [pfam](#)

Taxonomy: [Fungi/Metazoa group](#)

PubMed: [10 links](#)

Book: [20 links](#)

Protein: [Representatives](#)
[Specific Protein](#)
[Related Protein](#)
[Related Structure](#)
[Architectures](#)

Superfamily: [cl14608](#)

BioSystems: [1056 links](#)

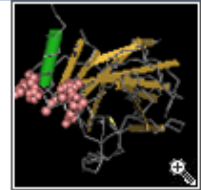
Conserved Features/Sites

[PubMed References](#)[Book References](#)[DNA binding](#)[zinc binding](#)[dimerization](#)

Feature 1: [DNA binding site \[nucleic acid binding site\]](#)

Evidence:

- **Structure:** 2AC0: Human p53 tetramer binds DNA, contacts at 4 Å
- [View structure with Cn3D](#)
- **Citation:** [PMID 16793544](#)
- **Structure:** 2GEQ: A mouse P53 core dimer binds DNA, contacts at 4 Å
- [View structure with Cn3D](#)
- **Citation:** [PMID 16717092](#)

[Download Cn3D for Viewing 3D Structure](#)[Scroll to Sequence Alignment Display](#)



▶ NCBI/BLAST Home

BLAST finds regions of similarity between biological sequences. [more...](#)

New DELTA-BLAST, a more sensitive protein-protein search

BLAST Assembled Genomes

Find Genomic BLAST pages:

Enter organism name or id—completions will be suggested

- [Human](#)
- [Mouse](#)
- [Rat](#)
- [Cow](#)
- [Pig](#)
- [Dog](#)
- [Rabbit](#)
- [Chimp](#)
- [Guinea pig](#)
- [Fruit fly](#)
- [Honey bee](#)
- [Chicken](#)
- [Zebrafish](#)
- [Clawed frog](#)
- [Arabidopsis](#)
- [Rice](#)
- [Yeast](#)
- [Microbes](#)

Basic BLAST

Choose a BLAST program to run.

- [nucleotide blast](#) Search a **nucleotide** database using a **nucleotide** query
Algorithms: blastn, megablast, discontinuous megablast
- [protein blast](#) Search **protein** database using a **protein** query
Algorithms: blastp, psi-blast, phi-blast, delta-blast
- [blastx](#) Search **protein** database using a **translated nucleotide** query
- [tblastn](#) Search **translated nucleotide** database using a **protein** query
- [tblastx](#) Search **translated nucleotide** database using a **translated nucleotide** query

Your Recent Results **New!**

[All Recent results...](#)

News

[Find Genomic BLAST pages](#)

You can now find Genomic BLAST pages using the search box from the BLAST homepage.

Thu, 02 Oct 2014 11:00:00 EST

[More BLAST news...](#)

Tip of the Day

[More tips...](#)

NCBI/BLAST/blastp suite

Standard Protein BLAST

blastn blastp **blastx** tblastn tblastx

Enter Query Sequence

BLASTp programs search protein databases using a protein query. [more...](#)

[Reset page](#) [Bookmark](#)

Enter accession number(s), gi(s), or FASTA sequence(s)

[Clear](#)

Query subrange

```
SVTCTYSPALNKMFCQLAKTCPVQLWVDS TPPPGTRVRAMAIYKQSQHMT EVVRRCPHHE  
RCSDS DGLAPPOHLIRVEGNLRVEYLD DRNTRFRHSVVVPEPPEVGS DCTTIHYNMYCNS  
SCHNGGNNRRPILTIITLEDSSGNLLGRNSFEVRYVCACPGDRRTEENLRKKGEPHHELP  
PGSTKRALPNNTSSSPQPKKKPLDGEYFTLQIRGRERFEMFRELNEALELKDAQAGKEPG  
GSAHSSHLKSKKGQSTSRHKKLMFKTEGPDSD
```

From

To

Or, upload file

選擇檔案 未選擇任何檔案

Job Title

Enter a descriptive title for your BLAST search

Align two or more sequences

Choose Search Set

Database

Non-redundant protein sequences (nr)

Organism
Optional

Enter organism name or id—completions will be suggested

Exclude

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown.

Exclude
Optional

Models (XM/XP) Uncultured/environmental sample sequences

Entrez Query
Optional

Enter an Entrez query to limit search

[YouTube](#) [Create custom database](#)

Program Selection

Algorithm

- blastp (protein-protein BLAST)
- PSI-BLAST (Position-Specific Iterated BLAST)
- PHI-BLAST (Pattern Hit Initiated BLAST)
- DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST)

Choose a BLAST algorithm

BLAST

Search database Non-redundant protein sequences (nr) using Blastp (protein-protein BLAST)

Show results in a new window

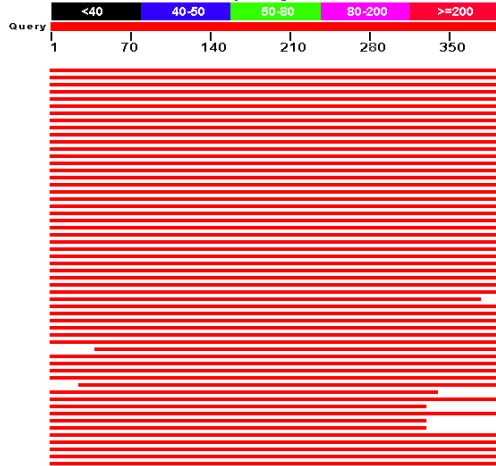
[Algorithm parameters](#)

Show Conserved Domains

Putative conserved domains have been detected, click on the image below for detailed results.



Color key for alignment scores



Sequences producing significant alignments:

Select: All None Selected 0

Alignments	Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/>	cellular tumor antigen p53 isoform a (Homo sapiens)	813	813	100%	0.0	100%	NP_000537.3
<input type="checkbox"/>	p53 antigen (Homo sapiens)	812	812	100%	0.0	99%	AA681242.1
<input type="checkbox"/>	P53(DCITL): cellular tumor antigen p53 (Drosophila ocellularis)	812	812	100%	0.0	99%	XP_004605859.1
<input type="checkbox"/>	p53 transformation suppressor (Homo sapiens)	812	812	100%	0.0	99%	CAA42633.1
<input type="checkbox"/>	tumor suppressor p53 (Homo sapiens)	812	812	100%	0.0	99%	A7861606.1
<input type="checkbox"/>	tumor protein p53 (L-Fraumeni syndrome) isoform CRA_c (Homo sapiens)	812	812	100%	0.0	99%	A1099162.1
<input type="checkbox"/>	p53 cellular tumor antigen (Homo sapiens)	811	811	100%	0.0	99%	AA599889.1
<input type="checkbox"/>	tumor suppressor gene p53 (Homo sapiens)	811	811	100%	0.0	99%	BA62715.1
<input type="checkbox"/>	tumor protein p53 (synthetic construct)	811	811	100%	0.0	99%	AA542852.1
<input type="checkbox"/>	tumor suppressor protein p53 (Homo sapiens)	811	811	100%	0.0	99%	AA528283.1
<input type="checkbox"/>	p53 transformation suppressor (Homo sapiens)	810	810	100%	0.0	99%	AA542833.1
<input type="checkbox"/>	tumor protein p53 (L-Fraumeni syndrome) (Homo sapiens)	810	810	100%	0.0	99%	AA542828.1
<input type="checkbox"/>	p53 transformation suppressor (Homo sapiens)	810	810	100%	0.0	99%	CAA42630.1
<input type="checkbox"/>	p53 transformation suppressor (Homo sapiens)	810	810	100%	0.0	99%	CAA42634.1
<input type="checkbox"/>	tumor suppressor protein p53 (Homo sapiens)	810	810	100%	0.0	99%	AA52835.1
<input type="checkbox"/>	p53 transformation suppressor (Homo sapiens)	809	809	100%	0.0	99%	CAA42632.1
<input type="checkbox"/>	p53 transformation suppressor (Homo sapiens)	809	809	100%	0.0	99%	AA681211.1
<input type="checkbox"/>	p53 antigen (Homo sapiens)	808	808	100%	0.0	99%	AA542639.1
<input type="checkbox"/>	tumor protein p53 (synthetic construct)	808	808	100%	0.0	99%	CAA42629.1
<input type="checkbox"/>	p53 transformation suppressor (Homo sapiens)	808	808	100%	0.0	99%	CAA42628.1
<input type="checkbox"/>	p53 transformation suppressor (Homo sapiens)	808	808	100%	0.0	99%	AA59987.1
<input type="checkbox"/>	oncogene protein p53 (Homo sapiens)	808	808	100%	0.0	99%	AA593598.1
<input type="checkbox"/>	tumor protein p53 (Homo sapiens)	808	808	100%	0.0	99%	AC025593.1
<input type="checkbox"/>	p53 transformation suppressor (Homo sapiens)	808	808	100%	0.0	99%	CAA42635.1

Download GenPept Graphics

cellular tumor antigen p53 isoform a (Homo sapiens)
 Sequence ID: np_000537.3 Length: 393 Number of Matches: 1
 See 9 more titles

Range 1: 1 to 393 GenPept Graphics
 Score: 813 bits(2101) 0.0
 Expect Method: 0.0
 Compositional matrix adjust. 393/393(100%) 393/393(100%) 0/393(0%)

Query 1 MEEPGDPSVEPPLSQETFSDLKLLPENNVSLPLPSQAWDQLMSPODIQWFTEDPG
 HEERQSDPSVEPPLSQETFSDLKLLPENNVSLPLPSQAWDQLMSPODIQWFTEDPG
 HEERQSDPSVEPPLSQETFSDLKLLPENNVSLPLPSQAWDQLMSPODIQWFTEDPG 60

Subject 1 DEAPRPEAAPVAPAPAAPTAAPAPAPSIPLSSVPSQNTYQSYVFRLOFLHSGTAK
 DEAPRPEAAPVAPAPAAPTAAPAPAPSIPLSSVPSQNTYQSYVFRLOFLHSGTAK
 DEAPRPEAAPVAPAPAAPTAAPAPAPSIPLSSVPSQNTYQSYVFRLOFLHSGTAK 120

Query 61 SVTCTYSPALNWF CQLAKTCPQLVSDTPPPPTVRVAHATVYQSQHITVARRCPHE
 SVTCTYSPALNWF CQLAKTCPQLVSDTPPPPTVRVAHATVYQSQHITVARRCPHE
 SVTCTYSPALNWF CQLAKTCPQLVSDTPPPPTVRVAHATVYQSQHITVARRCPHE 180

Query 181 RCSDGDLAPPQHLIRVEGILRVEYLDNRITFMSVVVYEPPEVSDCTTIDYNYKIG
 RCSDGDLAPPQHLIRVEGILRVEYLDNRITFMSVVVYEPPEVSDCTTIDYNYKIG
 RCSDGDLAPPQHLIRVEGILRVEYLDNRITFMSVVVYEPPEVSDCTTIDYNYKIG 240

Query 241 SCHGWRRPILTIITLEDSSGILGRNSFEVRCACPRDRTEELNKKVEPHELP
 SCHGWRRPILTIITLEDSSGILGRNSFEVRCACPRDRTEELNKKVEPHELP
 SCHGWRRPILTIITLEDSSGILGRNSFEVRCACPRDRTEELNKKVEPHELP 300

Query 301 POSTKRALPHTSSSPQPKKPLDGEYFTLQIRGRERFHPRELNEALELKDAGAKEPG
 POSTKRALPHTSSSPQPKKPLDGEYFTLQIRGRERFHPRELNEALELKDAGAKEPG
 POSTKRALPHTSSSPQPKKPLDGEYFTLQIRGRERFHPRELNEALELKDAGAKEPG 360

Query 361 OSRAHSHLKSXKQSTSRKXLIHFATEGPSD 393
 OSRAHSHLKSXKQSTSRKXLIHFATEGPSD 393

Related Information

- Gene - associated gene details
- UniGene - clustered expressed sequence tags
- PubChem BioAssay - bioactivity screening
- Map Viewer - aligned genomic context
- Identical Proteins - Proteins identical to the subject



[proSITE](#) is a database of protein families and domains. It consists of biologically significant sites, patterns and profiles that help to reliably identify to which known protein family a new sequence belongs. PROSITE is based at the Swiss Institute of Bioinformatics (SIB), Geneva, Switzerland.



[HAMAP](#) stands for High-quality Automated and Manual Annotation of Proteins. HAMAP profiles are manually created by expert curators. They identify proteins that are part of well-conserved protein families or subfamilies. HAMAP is based at the SIB Swiss Institute of Bioinformatics, Geneva, Switzerland.



[Pfam](#) is a large collection of multiple sequence alignments and hidden Markov models covering many common protein domains. Pfam is based at the Wellcome Trust Sanger Institute, Hinxton, UK.



[PRINTS](#) is a compendium of protein fingerprints. A fingerprint is a group of conserved motifs used to characterise a protein family or domain. PRINTS is based at the University of Manchester, UK.



[ProDom](#) protein domain database consists of an automatic compilation of homologous domains. Current versions of ProDom are built using a novel procedure based on recursive PSI-BLAST searches. ProDom is based at PRABI Villeurbanne, France.



[SMART](#) (a Simple Modular Architecture Research Tool) allows the identification and annotation of genetically mobile domains and the analysis of domain architectures. SMART is based at EMBL, Heidelberg, Germany.



[TIGRFAMs](#) is a collection of protein families, featuring curated multiple sequence alignments, hidden Markov models (HMMs) and annotation, which provides a tool for identifying functionally related proteins based on sequence homology. TIGRFAMs is based at the J. Craig Venter Institute, Rockville, MD, US.



[PIRSF](#) protein classification system is a network with multiple levels of sequence diversity from superfamilies to subfamilies that reflects the evolutionary relationship of full-length proteins and domains. PIRSF is based at the Protein Information Resource, Georgetown University Medical Centre, Washington DC, US.



[SUPERFAMILY](#) is a library of profile hidden Markov models that represent all proteins of known structure. The library is based on the SCOP classification of proteins: each model corresponds to a SCOP domain and aims to represent the entire SCOP superfamily that the domain belongs to. SUPERFAMILY is based at the University of Bristol, UK.



[CATH-Gene3D](#) database describes protein families and domain architectures in complete genomes. Protein families are formed using a Markov clustering algorithm, followed by multi-linkage clustering according to sequence identity. Mapping of predicted structure and sequence domains is undertaken using hidden Markov models libraries representing CATH and Pfam domains. CATH-Gene3D is based at University College, London, UK.



[PANTHER](#) is a large collection of protein families that have been subdivided into functionally related subfamilies, using human expertise. These subfamilies model the divergence of specific functions within protein families, allowing more accurate association with function, as well as inference of amino acids important for functional specificity. Hidden Markov models (HMMs) are built for each family and subfamily for classifying additional protein sequences. PANTHER is based at University of Southern California, CA, US.

← → ↻ 🏠 安全 | <https://www.ebi.ac.uk/interpro/search/sequence-search> ☆

應用程式 8 日曆 百度 NAS IE CGU FUN MEET NCBI SHOP 劇 BANK Travel SMS2 從IE匯入 INVEST [OFFICIAL] KeepVid JMI

EMBL-EBI Services Research Training About us EMBL-EBI Hinxton

InterPro

Protein sequence analysis & classification

Examples: IPR020405, kinase, P51587, PF02932, GO:0007165

Home Search Release notes Download About InterPro Help Contact

By sequence By domain architecture

InterProScan sequence search

This form allows you to scan your sequence for matches against the InterPro protein signature databases, using InterProScan tool.

Enter or paste a protein sequence in FASTA format (complete or not - e.g. `PMPIGSKERPTFFEIFKTRCNKADLGPISLN`), with a maximum length of 40,000 amino acid long.


Please note that you can only scan one sequence at a time.

Analyse your protein sequence

▸ **Advanced options**

Submit Clear Example protein sequence

InterProScan



InterProScan is a sequence analysis application (nucleotide and protein sequences) that combines different protein signature recognition methods into one resource.

[More about InterProScan.](#)

? Need more help?

If you need more info on InterProScan, you can either look at the:

InterProScan sequence search

This form allows you to scan your sequence for matches against the InterPro protein signature databases, using InterProScan tool.

Enter or paste a protein sequence in FASTA format (complete or not - e.g. PMPIGSKERPTFFEIFKTRCNKADLGPISLN), with a maximum length of 40,000 amino acid long.

Please note that you can only scan one sequence at a time.

Analyse your protein sequence

▼ Advanced options

Select the applications to run: [Uncheck all](#) [Select all](#)

Member databases

Families, domains, sites & repeats

CDD HAMAP PANTHER PfamA PIRSF PRINTS ProDom Prosite-Profiles

SMART TIGRFAM Prosite-Patterns

Structural domains

Gene3d SFLD SUPERFAMILY

Other sequence features

Coils MobiDB Lite Phobius SignalP TMHMM

Submit

Clear

Example protein sequence

InterProScan sequence search

This form allows you to scan your sequence for matches against the InterPro protein signature databases, using InterProScan tool.

Enter or paste a protein sequence in FASTA format (complete or not - e.g. PMPIGSKERPTFFEIFKTRCNKADLGPISLN), with a maximum length of 40,000 amino acid long.

Please note that you can only scan one sequence at a time.

Analyse your protein sequence

```
MITIDGNGAVASVAFRTSEVIAIYPITPSSSTMAEQADAWAGNGLKNVWGDTPRVVEMQSEAGAIATVHGALQTGAL
STSFTSSQGLLLMIPTLYKLAGELTPFVLHVAARTVATHALSIFGDHSDVMAVRQTGCAMLCAANVQEAQDFALISQI
ATLKSRVPIHFFDFGFRSHEINKIVPLADDTILDMPQVEIDHRARALNPEHPVIRGTSANPDYFQSREATNPWY
NAVYDHVEQAMNDFSAATGRQYQPFYYGHPQAERVIILMGSAGTCEEVDELLTRGEKVGVLKVRLYRPFSAK
HLLQALPGSVRSVAVLDRTKEPGAQAEPLYLDVMTALAEAFNNGERETLPRVIGGRYGLSSKEFGPDCVLAVFAEL
NAAKPKARFTVGIYDDVTNLSLPLPENTLPNSAKLEALFYGLGSDGSVSATKNNIKIIGNSTPWYAQQYFVYDSKKA
```

Advanced options

Select the applications to run:

Member databases

Families, domains, sites & repeats

CDD HAMAP PANTHER PfamA PIRSF PRINTS ProDom Prosite-Profiles

SMART TIGRFAM Prosite-Patterns

Structural domains

Gene3d SFLD SUPERFAMILY

Other sequence features

Coils MobiDB Lite Phobius SignalP TMHMM

Your job is currently running... please be patient

The result of your job will appear in this browser window. This page refreshes automatically every 30 seconds.

You may bookmark this page to view your results later if you wish. Results are stored for 7 days.

Job ID: jprscan5-S20171228-005706-0136-81379822-p2m

Submit

Clear

Example protein sequence

Overview

Similar proteins

Structures

Filter view on

Entry type

- H** Homologous superfamily
- F** Family
- D** Domains
- R** Repeats
- S** Site

Status

- ?** Unintegrated

Per-residue features


- Residue annotation

Colour by

- domain relationship
- source database

Submitted

Length 1,174 amino acids

Export 

TSV
XML
JSON
GFF3
SVG
FASTA

Protein family membership

F Pyruvate-flavodoxin oxidoreductase (IPR011895)

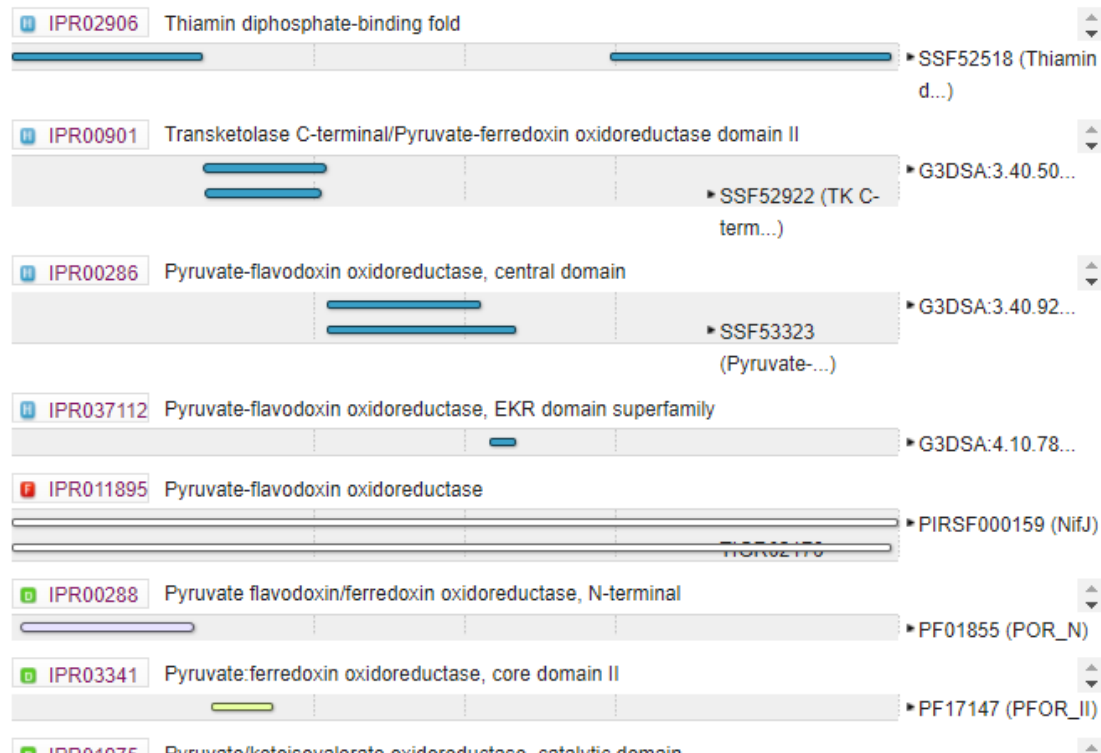
Homologous superfamilies



Domains and repeats



Detailed signature matches



Protein Sequence Analysis Tools



ExPASy
Bioinformatics Resource Portal

Visual Guidance

Categories

proteomics

protein sequences and identification
mass spectrometry and 2-DE data
protein characterisation and function
families, patterns and profiles
post-translational modification
protein structure
protein-protein interaction
similarity search/alignment

genomics

structural bioinformatics

systems biology

phylogeny/evolution

population genetics

transcriptomics

biophysics

imaging

IT infrastructure

drug design

Resources A..Z

Links/Documentation

ExPASy Molecular Biology Server

Expert Protein Analysis System is the new SIB Bioinformatics Resource Portal which provides access to scientific databases and software tools in different areas of life sciences including proteomics, genomics, phylogeny, systems biology, population genetics, transcriptomics etc

Tools

-  **SWISS-MODEL Workspace** • structure homology-modeling • [\[more\]](#)
-  **SwissDock** • protein ligand docking server • [\[more\]](#)
-  **2ZIP** • Prediction of leucine zipper domains • [\[more\]](#)
-  **3of5** • find user-defined patterns in protein sequences • [\[more\]](#)
-  **AACompldent** • protein identification by aa composition • [\[more\]](#)
-  **AACompSim** • amino acid composition comparison • [\[more\]](#)
-  **Agadir** • Prediction of the helical content of peptides • [\[more\]](#)
-  **ALF** • simulation of genome evolution • [\[more\]](#)
-  **Alignment tools** • Four tools for multiple alignments • [\[more\]](#)
-  **AllAll** • protein sequences comparisons • [\[more\]](#)
-  **APSSP** • Advanced Protein Secondary Structure Prediction • [\[more\]](#)
-  **Ascalaph** • Molecular modeling software • [\[more\]](#)
-  **big-PI** • predict GPI modification sites • [\[more\]](#)
-  **Biochemical Pathways** • Biochemical Pathways • [\[more\]](#)
-  **BLAST** • sequence similarity search • [\[more\]](#)
-  **BLAST (UniProt)** • BLAST search on the UniProt web site • [\[more\]](#)
-  **BLAST - NCBI** • Biological sequence similarity search • [\[more\]](#)
-  **BLAST - PBIL** • BLAST search on protein sequence databases • [\[more\]](#)
-  **Blast2Fasta** • Blast to Fasta conversion • [\[more\]](#)
-  **boxshade** • MSA pretty printer • [\[more\]](#)
-  **CFSSP** • Protein secondary structure prediction • [\[more\]](#)
-  **ChloroP** • chloroplast transit peptides & cleavage sites • [\[more\]](#)
-  **Click2Drug** • Directory of computational drug design tools • [\[more\]](#)
-  **ClustalO (UniProt)** • Align two or more protein sequences • [\[more\]](#)
-  **ClustalW** • Multiple sequence alignment • [\[more\]](#)
-  **ClustalW - PBIL** • Multiple sequence alignment program • [\[more\]](#)
-  **ClustalW2** • Multiple sequence alignment program • [\[more\]](#)
-  **Coiled-Coils prediction** • Prediction of coiled coils regions • [\[more\]](#)
-  **COILS** • Prediction of Coiled Coil Regions in Proteins • [\[more\]](#)

-  **ColorSeq** • Color Protein Sequence • [\[more\]](#)
-  **Compute pI/MW** • theoretical pI and Mw computation • [\[more\]](#)
-  **CPHmodels** • Protein homology modeling • [\[more\]](#)
-  **CSS-Palm** • Prediction of palmitoylation sites in proteins • [\[more\]](#)
-  **DAS-TMfilter** • Prediction of transmembrane regions • [\[more\]](#)
-  **Decrease redundancy** • Sequence redundancy reduction • [\[more\]](#)
-  **DIALIGN** • Local multiple sequence alignment • [\[more\]](#)
-  **DictyOGlyc** • GlcNAc O-glycosylation sites in D.discoideum • [\[more\]](#)
-  **DisEMBL** • Prediction of disordered protein regions • [\[more\]](#)
-  **DLP-SVM** • Domain linker predictor • [\[more\]](#)
-  **Dotlet** • sequence similarity plots • [\[more\]](#)
-  **EasyProt** • graphical platform for proteomics analysis • [\[more\]](#)
-  **ELM** • Eukaryotic Linear Motifs • [\[more\]](#)
-  **EMBnet services** • bioinformatics tools, databases and courses • [\[more\]](#)
-  **EMBOSS translation tools** • sequence translation tools • [\[more\]](#)
-  **epestfind** • Identification of PEST motifs • [\[more\]](#)
-  **FASTA/SSEARCH/GGSEARCH/GLSEARCH** • Sequence similarity searching of protein db • [\[more\]](#)
-  **FindMod** • protein post-translational modification prediction • [\[more\]](#)
-  **FindPept** • peptide identification from unspecific cleavage • [\[more\]](#)
-  **FingerPRINTScan** • scan sequences against PRINTS • [\[more\]](#)
-  **FUGUE** • Sequence-structure homology recognition • [\[more\]](#)
-  **GENIO/logo** • RNA/DNA & Amino Acid Sequence Logos • [\[more\]](#)
-  **Geno3D** • Protein molecular modelling • [\[more\]](#)
-  **GlobPlot** • Protein disorder/globularity/domain predictor • [\[more\]](#)
-  **GlycanMass** • oligosaccharide structure mass calculation • [\[more\]](#)
-  **GlycoMod** • oligosaccharide structure prediction • [\[more\]](#)
-  **Glycoviewer** • visualize a set of glycan structures • [\[more\]](#)
-  **GOR** • Protein secondary structure prediction • [\[more\]](#)
-  **GPI-SOM** • identify GPI-anchor signals • [\[more\]](#)
-  **GPMaw lite** • protein physical and chemical parameters • [\[more\]](#)
-  **Graphical Codon Usage Analyser** • Codon bias • [\[more\]](#)
-  **HAMAP** • UniProtKB family classification and annotation • [\[more\]](#)
-  **HAMAP-Scan** • Scan protein sequences against the HAMAP families • [\[more\]](#)
-  **HCA (hydrophobic cluster analysis)** • hydrophobic cluster analysis • [\[more\]](#)
-  **HCD/CID spectra merger** • combine HCD and CID MS/MS spectra • [\[more\]](#)
-  **HeliQuest** • alpha-helical properties screening • [\[more\]](#)
-  **HHpred** • Homology detection & structure prediction • [\[more\]](#)
-  **HMMTOP** • Prediction of transmembranes helices and topology • [\[more\]](#)
-  **HTMSRAP** • Helical TransMembrane Prediction • [\[more\]](#)
-  **ImageMaster / Melanie** • software for 2-D PAGE analysis • [\[more\]](#)
-  **InsPecT** • MS/MS tool to identify modified peptides • [\[more\]](#)
-  **InterProScan** • Family and domain datatabase search • [\[more\]](#)
-  **ISMARA** • genome-wide expression data modeling • [\[more\]](#)
-  **IsotopIdent** • theoretical isotopic distribution • [\[more\]](#)
-  **Jmol** • 3D molecule visualization tool • [\[more\]](#)
-  **Jpred** • Secondary Structure Prediction Server • [\[more\]](#)
-  **Kalign - EBI** • Fast and accurate multiple sequence alignment • [\[more\]](#)
-  **Kalign - SBC** • Fast and accurate multiple sequence alignment • [\[more\]](#)
-  **LALIGN** • Pairwise alignment • [\[more\]](#)
-  **LipoP** • predict lipoproteins and signal peptides • [\[more\]](#)
-  **LiveBench** • Continuous Benchmarking of Structure Prediction Se • [\[more\]](#)
-  **MAFFT - CBRC** • Multiple sequence alignment • [\[more\]](#)
-  **MAFFT - EBI** • Multiple sequence alignment • [\[more\]](#)
-  **Make2D-DB II** • package to build web-based proteomics database • [\[more\]](#)
-  **MakeMultimer.py** • Reconstruction of multimeric molecules in crystals • [\[more\]](#)
-  **MALDI PepQuant** • quantify MALDI peptides • [\[more\]](#)
-  **MARCOIL** • coiled-coils prediction • [\[more\]](#)
-  **MarvinSpace** • Visualization, modeling and interactive rendering • [\[more\]](#)
-  **Mascot** • protein identification from mass spectrometry data • [\[more\]](#)
-  **MassSearch** • sequence search by mass spectrometry weights • [\[more\]](#)

Example

cellular tumor antigen p53 isoform a [Homo sapiens]

```
>sp|P04637|P53_HUMAN Cellular tumor antigen p53 OS=Homo sapiens GN=TP53 PE=1 SV=4
MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIEQWFTEDPGP
DEAPRMPEAAPPVAPAPAAPTPAAPAPAPSWPLSSSVPSQKTYQGSYGFRLGFLHSGTAK
SVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAMAIYKQSQHMTEVVRRCPHHE
RCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVVPYEPPEVGSDCCTTIHYNMCNS
SCMGGMNRRLPILTIITLEDSSGNLLGRNSFEVRVCACPGRDRRTEENLRKKGEPHHELP
PGSTKRALPNNTSSSPQPKKPLDGEYFTLQIRGRERFEMFRELNEALELKDAQAGKEPG
GSAHSSHLKSKKGQSTSRHKKLMFKTEGPDS
```