

# Tools & Databases of Short Linear Motifs

---

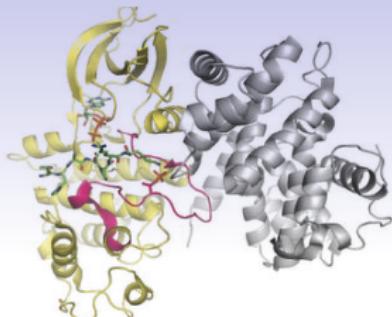
Holger Dinkel

EMBO Practical Course:

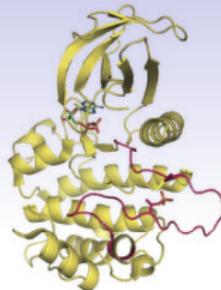
“Computational Analysis of Protein-Protein Interactions:  
Sequences, Networks and Diseases”

Rome, 08.11.2018

# PROTEIN PHOSPHORYLATION SITES



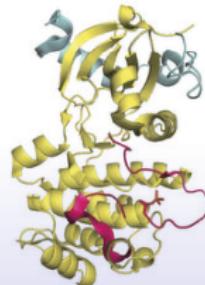
Cdk1/cyclin B



Plk1



Aurora A/TPX2



Aurora B/INCENP

*"Spatial exclusivity combined with positive and negative selection of phosphorylation motifs is the basis for context-dependent mitotic signaling"; ALEXANDER ET AL.; (SCI. SIG 2011)*

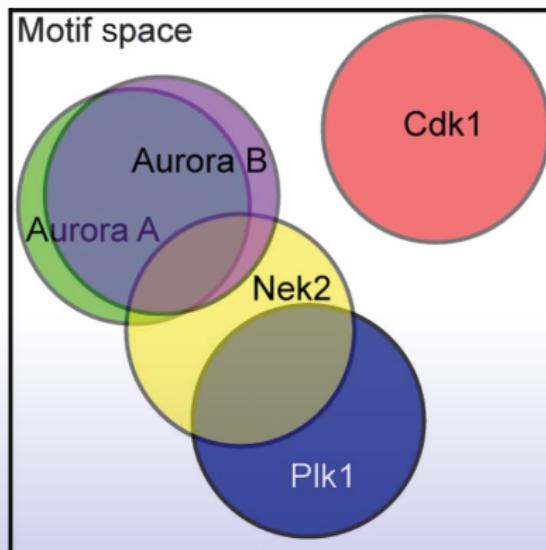
## PROTEIN PHOSPHORYLATION SITES

Kinase	-3	-2	-1	0	1	2	3
Cdk1	.	.	.	p[ST]	P	.	[KR]
Plk1	.	[DEN]	.	p[ST]	[ILMVFWY]	.	.
Nek2	[FML]	[!P]	[!P]	p[ST]	[ILMV]	.	.
AuroraA	R	[KR]	.	p[ST]	[!P]	.	.
AuroraB	.	R	[KR]	p[ST]	[!P]	.	.

*"Spatial exclusivity combined with positive and negative selection of phosphorylation motifs is the basis for context-dependent mitotic signaling"; ALEXANDER ET AL.; (SCI. SIG 2011)*

## PROTEIN PHOSPHORYLATION SITES

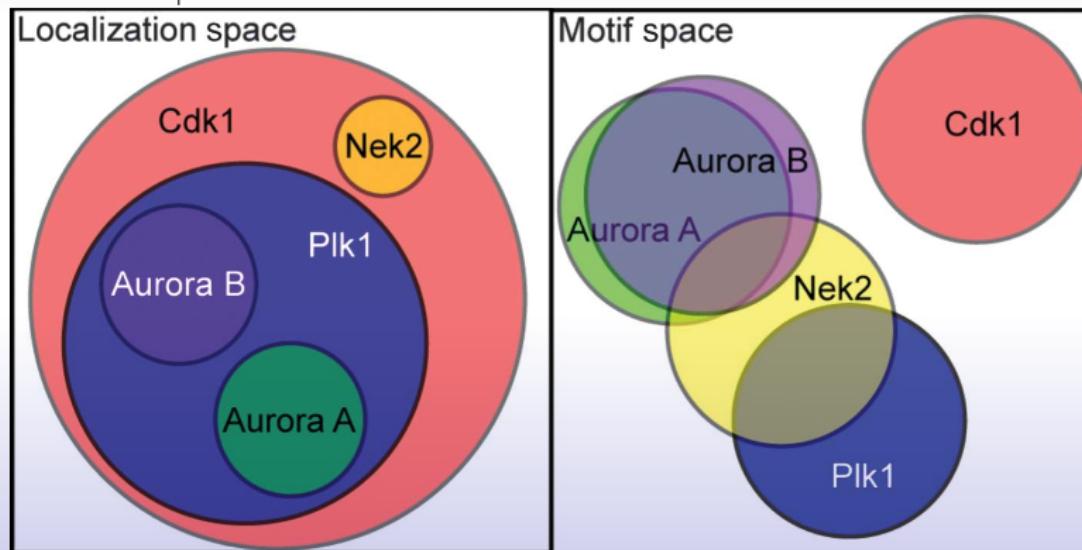
Kinase	-3	-2	-1	0	1	2	3
Cdk1	.	.	.	p[ST]	P	.	[KR]
Plk1	.	[DEN]	.	p[ST]	[ILMVFWY]	.	.
Nek2	[FML]	[!P]	[!P]	p[ST]	[ILMV]	.	.
AuroraA	R	[KR]	.	p[ST]	[!P]	.	.
AuroraB	.	R	[KR]	p[ST]	[!P]	.	.



"Spatial exclusivity combined with positive and negative selection of phosphorylation motifs is the basis for context-dependent mitotic signaling"; ALEXANDER ET AL.; (SCI. SIG 2011)

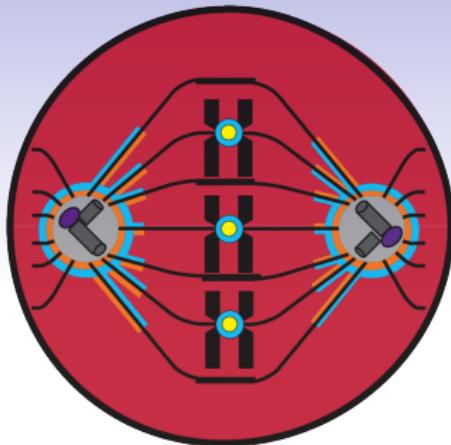
# PROTEIN PHOSPHORYLATION SITES

Kinase	-3	-2	-1	0	1	2	3
Cdk1	.	.	.	p[ST]	P	.	[KR]
Plk1	.	[DEN]	.	p[ST]	[ILMVFWY]	.	.
Nek2	[FML]	[!P]	[!P]	p[ST]	[ILMV]	.	.
AuroraA	R	[KR]	.	p[ST]	[!P]	.	.
AuroraB	.	R	[KR]	p[ST]	[!P]	.	.



"Spatial exclusivity combined with positive and negative selection of phosphorylation motifs is the basis for context-dependent mitotic signaling"; ALEXANDER ET AL.; (SCI. SIG 2011)

# PROTEIN PHOSPHORYLATION SITES



Kinase localization in Metaphase:

**Cdk1** whole cell

**Plk1** kinetochores

**Aurora A** centrosomes & microtubules

**Aurora B** centromeres & spindle

**Nek2** centrosomes

*"Spatial exclusivity combined with positive and negative selection of phosphorylation motifs is the basis for context-dependent mitotic signaling"; ALEXANDER ET AL.; (SCI. SIG 2011)*

## Phospho.ELM

Database of experimentally verified phosphorylation sites in eukaryotic proteins.

Current release contains 8,718 protein entries covering more than 42,500 instances. (Instances are fully linked to literature references.)

# Phospho.ELM

a database of S/T/Y phosphorylation sites

## Statistics:

Instances	42,575
Kinases	310
Reference	3,672
Sequences	11,223
Substrates	8,718

[Home](#) [PhosphoBlast](#) [Contribute](#) [Download](#) [Help](#) [Links](#) [About](#)

## SEARCH

- for phosphorylation sites in proteins using protein name or gene name  
(eg. Paxillin, Shc, MAPK)

- by UniPROT accession or Ensembl identifier:  
(eg. P12931 or P55211)

- by selected kinase (List):

 None

- by selected phospho-peptide binding domain (List):

 None

- Choose which organisms to include

- All
- Caenorhabditis
- Drosophila
- Vertebrates

- Do not show high throughput data

- Output as Comma-Separated-Values (.csv)

# Phospho.ELM

a database of S/T/Y phosphorylation sites

Statistics:

Instances	42,575
Kinases	310
Reference	3,672
Sequences	11,223
Substrates	8,718

[Home](#) [PhosphoBlast](#) [Contribute](#) [Download](#) [Help](#) [Links](#) [About](#)

Substrate: p53 (Cellular tumor antigen p53)

Seq-ID: P04637 [*Homo sapiens*]
 

External Source(s): PHOSIDA

MINT Interaction(s): [\[show\]](#)GO-Terms: [\[show\]](#)

Conservation:

● Click on table headers for sorting

Res.	Pos.	Sequence	Kinase	PMID	Src	Cons.	ELM	Binding Domain	SMART/Pfam	IUPRED score	PDB	P3D Acc.
S	9	MEEPQSDP <b>S</b> VEPPLSQETF	-	11875057	LTP	0.75		-	P53_TAD	0.94	-	-
S	15	QSDPSVEPPL <b>S</b> QETFSDLWKL	DNA-PK	10446957	LTP	1.00	MOD_PIKK_1	-	P53_TAD	0.66	-	-
S	15	QSDPSVEPPL <b>S</b> QETFSDLWKL	ATM	11875057	LTP	1.00	MOD_PIKK_1	-	P53_TAD	0.66	-	-
T	18	PSVEPPLSQ <b>E</b> TFSDLWKLPE	CK1_group	10606744	LTP	1.00	MOD_CK1_1	-	P53_TAD	0.66	-	-
T	18	PSVEPPLSQ <b>E</b> TFSDLWKLPE	TTK	19332559	LTP	1.00	MOD_CK1_1	-	P53_TAD	0.66	-	-
T	18	PSVEPPLSQ <b>E</b> TFSDLWKLPE	VRK1	10951572	LTP	1.00	MOD_CK1_1	-	P53_TAD	0.66	-	-
T	18	PSVEPPLSQ <b>E</b> TFSDLWKLPE	VRK1	15542844	LTP	1.00	MOD_CK1_1	-	P53_TAD	0.66	-	-
S	20	VEPPLSQET <b>F</b> SDLWKLPPENN	-	15254178	LTP	0.95		-	P53_TAD	0.58	-	-
S	20	VEPPLSQET <b>F</b> SDLWKLPPENN	-	15489221	LTP	0.95		-	P53_TAD	0.58	-	-
S	20	VEPPLSQET <b>F</b> SDLWKLPPENN	-	10801407	LTP	0.95	-	P53_TAD	0.58	-	-	
S	20	VEPPLSQET <b>F</b> SDLWKLPPENN	-	12111733	LTP	0.95	-	P53_TAD	0.58	-	-	



**Links to:**

- STRING
- NetworKin
- Phosida
- Phospho3D

**Display:**

- MINT interactions
- GO-Terms

**Substrate:****Caspase 9** (Cysteine protease)**Seq-ID:****P55211** [*Homo sapiens*]**Interaction Network(s):****External Source(s):****PHOSIDA**[\[hide\]](#)

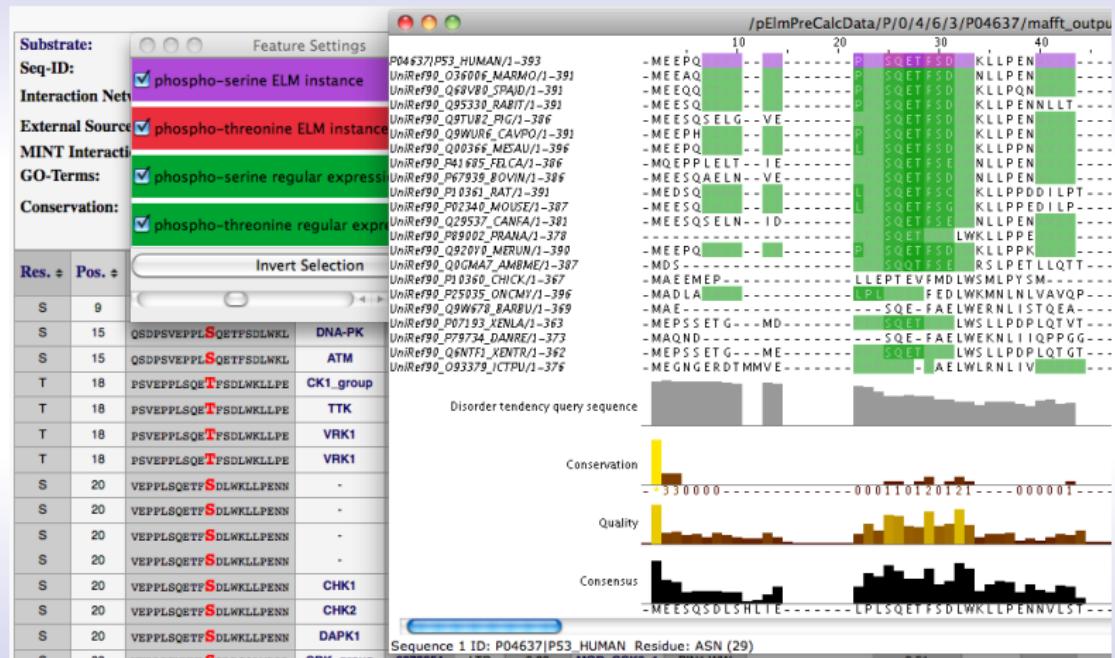
MINT-15372 APAF\_HUMAN

MINT-18815 CASP3\_HUMAN

MINT-25026 XIAP\_HUMAN

[\[hide\]](#)**MINT Interaction(s):****Molecular Function**cysteine-type endopeptidase activity,  
protein binding,  
enzyme activator activity**GO-Terms:**

Precalculated conservation scores for the phosphorylation sites are presented using **Jalview**



PhosphoSitePlus<sup>®</sup>

with grant support from

from

Cell Signaling  
TECHNOLOGY<sup>®</sup>

ABOUT PHOSPHOSITE | USING PHOSPHOSITE | CURATION PROCESS | CONTACT

Home

PhosphoSitePlus<sup>®</sup> (PSP) is an online systems biology resource providing comprehensive information and tools for the study of protein post-translational modifications (PTMs) including phosphorylation, ubiquitination, acetylation and methylation. See [About PhosphoSite](#) above for more information.

Please cite the following reference for this resource: Hornbeck PV, et al. (2012) *Nucleic Acids Res.* 2012 40:D261-70. [reprint]

## A PROTEIN MODIFICATION RESOURCE

### PROTEIN OR SUBSTRATE SEARCH

Protein Name:

**SEARCH**

### ADVANCED SEARCH AND BROWSE OPTIONS

[Protein, Sequence, or Reference Search](#)

[Site Search](#)

[Comparative Site Search](#)

[Browse MS2 Data By Disease](#)

[Browse MS2 Data by Cell Line](#)

[Browse MS2 Data by Tissue](#)

### DOWNLOADS, LINKS & APPLICATIONS

[Reprints, References, Supplemental Tables](#)

[Downloadable Datasets](#)

[Motif Analysis Tools](#)

### WHAT'S NEW

Aug 2014 [Download PTM-VarMut dataset](#): Overlap of disease missense mutations & genetic variants, with their corresponding PTMs and flanking sequences.

Jul 2012 [Download Datasets of Regulatory or Disease-Associated Sites](#).

Dec 2011 [Download "PhosphoSitePlus: a comprehensive resource..."](#) in January 2012 issue of *Nucleic Acids Research*.

Jul 2011 [Multiple Sequence Alignment \(MSA\)](#) added to the Protein Page.

Jul 2011 [Download PyMOL & Chimera Scripts](#) from the Structure Viewer window.

### Phosphorylation Site Statistics

Non-redundant sites:	239,738
Non-redundant proteins:	19,680
Sites curated from literature:	136,109
All sites using site-specific (SS) methods:	12,528
All sites using discovery-mode MS (MS) methods:	127,064
Sites using both SS and MS methods:	6,010
MS sites observed at CST:	151,472
Number of curated papers:	16,428

### Other Modification Site Statistics

Acetylation:	27,657	Caspase cleavage:	481
Di-methylation:	2,555	Methylation:	163
Mono-methylation:	4,992	O-GalNAc:	2,118
O-GlcNAc:	1,390	Succinylation:	4,657
Sumoylation:	816	Tri-methylation:	321
Ubiquitination:	51,255		

PhosphoSite, created by Cell Signaling Technology is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. Information about permissions beyond the scope of this license are available at <http://www.phosphosite.org/staticContact.do>.

[Home](#) | [Curator Login](#)

With enhanced literature mining using Lingumatics I2E

Produced by 3rd Millennium | Design by Digizyme

©2003-2013 Cell Signaling Technology, Inc.

Tools & Databases of Short Linear Motifs

8 / 21

**PhosphoSitePlus®**

with grant support from

[Home](#) [ABOUT PHOSPHOSITE](#) [USING PHOSPHOSITE](#) [CURATION PROCESS](#) [CONTACT](#)

Advanced Search / Browse Functions:

Search Results for: p53

Modification-specific Antibodies Available from Cell Signaling Technology®

Protein-specific Antibodies or siRNA Available from Cell Signaling Technology®

Displaying 1-64 of 64 records. << Previous | Next >>

**DOWNLOAD**

Protein	GeneSymb	ACC#	Organism	MW (Da)	Modifications(show legend)	
p53	TP53	P04637 P04649 P10361 Q93330 P13481	human mouse rat rabbit monkey	43,653 43,559 43,451 43,455 43,696	H-m1, K-ac, K-m1, K-m2, K-sm, K-ub, R-m1, S-gl, S-p, T-p, Y-p	
S3BP1	tumor protein p53 binding protein 1	TP53BP1	Q12888 XP_215812	human mouse	213,574 211,340 212,859	D-ca, K-ac, K-m1, K-ub, R-m1, S-p, T-p, Y-p
S3BP2	Apoptosis-stimulating of p53 protein 2	TP53BP2	Q13625 XP_215812	human mouse	125,616 125,301 125,312	K-ub, S-gl, S-p, T-p, Y-p
AIFM2		AIFM2	Q9BR08 Q8BU44	human mouse	40,527 40,635	K-ac, K-ub, S-p, T-p, Y-p
ANO9	tumor protein p53 inducible protein 5	ANO9	P86044 XP_574586	human mouse	90,367 87,180 98,746	S-p, T-p, Y-p
CDIP		CDIP1	Q9H305 Q9DB75 Q5UZU8	human mouse rat	21,892 21,835 21,858	K-ub, T-p
CYFIP2	p53 inducible protein	CYFIP2	Q96F07 Q5SGX6 D3ZK82	human mouse rat	148,398 145,659 68,079	K-ac, K-ub, S-p, T-p, Y-p
EFEMP2	mutant p53 binding protein 1	EFEMP2	Q9WVJ9	human	49,403 49,425	Y-p
E124	tumor protein p53 inducible protein 8	E124	Q14681 Q61070 Q4KHK7	human mouse rat	38,065 38,033 38,093	K-ub, S-p, T-p
ENCL	tumor protein p53 inducible protein 10	ENCL	Q37509 Q2V9T0	human mouse	66,130 66,113 66,196	K-ub, S-p, T-p, Y-p
GADD45GIP1		GADD45GIP1	Q9TAE8 Q9C595 Q5W42	human mouse rat	25,384 23,320 26,467	K-ub, S-p, T-p, Y-p
IQCBI1	p53 and DNA damage-regulated IQ motif protein	IQCBI1	Q9BPF00	mouse	68,734	K-m2, K-ub, S-p, T-p
IRSp53	Insulin receptor substrate p53	IRSp53	BA1AP2	human mouse rat	60,868 59,237 59,183	K-ac, K-ub, S-gl, S-p, T-p, Y-p
JMY	junction mediating and regulatory protein, p53 cofactor	JMY	Q9QXM1	mouse	111,445 110,586	K-ub, S-p, T-p, Y-p
L GALSTB		L GALSTB	P47929 Q54974	human mouse	15,075 15,173	Y-p
LITAF	tumor protein p53 inducible protein 7	LITAF	Q99732	human	17,107	K-ub, S-p, T-p, Y-p
MAD1L1	tumor protein p53 inducible protein 9	MAD1L1	Q9YSD9 Q5WVXB	human mouse	83,057 83,541	K-ac, K-ub, S-p, T-p, Y-p
MARKA		MARKA	Q8TC05	human	80,735 79,680	

[www.phosphosite.org/proteinAction.do?id=465&showAll=true](http://www.phosphosite.org/proteinAction.do?id=465&showAll=true)



Home

with grant support from



from

Cell Signaling  
TECHNOLOGY®

ABOUT PHOSPHOSITE | USING PHOSPHOSITE | CURATION PROCESS | CONTACT

Advanced Search /  
Browse Functions:

Protein Page:

p53 (human)

**Overview**

**p53** is a transcription factor and major tumor suppressor that plays a major role in regulating cellular responses to DNA damage and other genomic aberrations. Activation of p53 can lead to either cell cycle arrest and DNA repair or apoptosis. More than 50 percent of human tumors contain a mutation or deletion in the TP53 gene. p53 is modified post-translationally at multiple sites. DNA damage induces phosphorylation of p53 at S15, S20 and S37, reducing its interaction with the nucleophiles HDM2; HDM2 inhibits p53 accumulation by targeting it for ubiquitination and proteasomal degradation. Phosphorylated by many kinases including Chk2 and Chk1 at S20, enhancing its tetramerization, stability and activity. The phosphorylation by CK2 at S192 is increased in human tumors and has been reported to influence the growth suppressor function, DNA binding and transcriptional activation of p53. Phosphorylation of p53 at S46 regulates the ability of p53 to induce apoptosis. The acetylation of p53 appears to play a positive role in the accumulation of p53 during the stress response. Following DNA damage, p53 becomes acetylated at K382, enhancing its binding to DNA. Deacetylation of p53 can occur through interaction with SIRT1, a deacetylase that may be involved in cellular aging and the DNA damage response. p53 regulates the transcription of a set of genes encoding endosomal proteins that regulate endosomal functions. These include STEAP3 and CHMP4C, which enhance exosome production, and CAV1 and CHMP4C, which produce a more rapid endosomal clearance of the EGFR from the plasma membrane. DNA damage regulates a p53-mediated secretory pathway, involving the recruitment of some proteins such as Hsp90, SEC61, SEC63, MKEF-A, and VAPs, and inhibiting the secretion of others including CTSL and IGFBP-2. Two alternatively spliced human isoforms have been reported. Isoform 2 is 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. Note: This description may include information from UniProtKB.

**Protein type:** DNA binding protein; Nuclear receptor co-regulator; Motility/polarity/chemotaxis; Transcription factor; Activator protein; Tumor suppressor

**Cellular Component:** PML body; transcription factor TFIID complex; protein complex; nuclear matrix; mitochondrion; endoplasmic reticulum; replication fork; cytosol; nucleoplasm; nuclear body; mitochondrial matrix; cytoplasm; nuclear chromatin; nucleolus; chromatin; nucleus

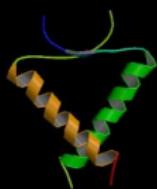
**Molecular Function:** identical protein binding; protease binding; zinc ion binding; protein phosphatase 2A binding; p53 receptor binding; receptor tyrosine kinase binding; transcription factor activity, sequence-specific binding; protein phosphatase binding; protein kinase binding; histone deacetyltransferase binding; protein binding; copper ion binding; histone deacetylase regulator activity; enzyme binding; DNA binding; protein heterodimerization activity; chaperone binding; ubiquitin protein ligase binding; damaged DNA binding; chromatin binding; transcription factor activity; ATP binding

**Biological Process:** central nervous system development; viral reproduction; positive regulation of apoptosis; multicellular organismal development; positive regulation of transcription, DNA-dependent; T cell differentiation in the thymus; gastrulation; determination of adult life span; DNA damage response; signal transduction by p53 class mediator resulting in cell cycle arrest; response to antibiotic; regulation of apoptosis; cellular response to glucose starvation; protein localization; negative regulation of neuroblast proliferation; base-excision repair; transforming growth factor beta receptor signaling pathway; protein complex assembly; cell cycle arrest; ER overload response; response to X-ray; somogenesis; release of cytochrome c from mitochondria; chromatin assembly; cell aging; rRNA transcription; positive regulation of peptidyl-tyrosine phosphorylation; negative regulation of DNA replication; negative regulation of fibroblast proliferation; embryonic organ development; positive regulation of transcription from RNA polymerase II promoter; regulation of mitochondrial membrane permeability; negative regulation of transcription; DNA-dependent; regulation of tissue remodeling; negative regulation of apoptosis; G1 DNA damage checkpoint; DNA damage response; signal transduction by p53 class mediator; apoptosis; negative regulation of transcription; RNA polymerase II promoter; DNA strand break repair; regulation of cell proliferation; positive regulation of protein oligomerization; positive regulation of histone deacetylation; DNA damage response; signal transduction by p53 class mediator resulting in transcription of p21 class mediator; regulation of transcription; DNA-dependent; T cell proliferation during immune response; double-strand break repair; positive regulation of neuron apoptosis; response to gamma radiation; cell differentiation; DNA damage response; signal transduction by p53 class mediator resulting in induction of apoptosis; protein tetramerization; Notch signaling pathway; in utero embryonic development; multicellular organism growth; B cell lineage commitment; cell proliferation; neuron apoptosis; T cell lineage commitment; negative regulation of helicase activity; nucleotide-excision repair; protein import into nucleus; translocation; DNA strand renaturation; Ras protein signal transduction; negative regulation of cell growth; negative regulation of transforming growth factor beta receptor signaling pathway; blood coagulation; response to DNA damage stimulus

**Reference #:** P04637 (UniProtKB)

**Select Structure to View Below**

p53



TA1U - A/C=324-356 (human)

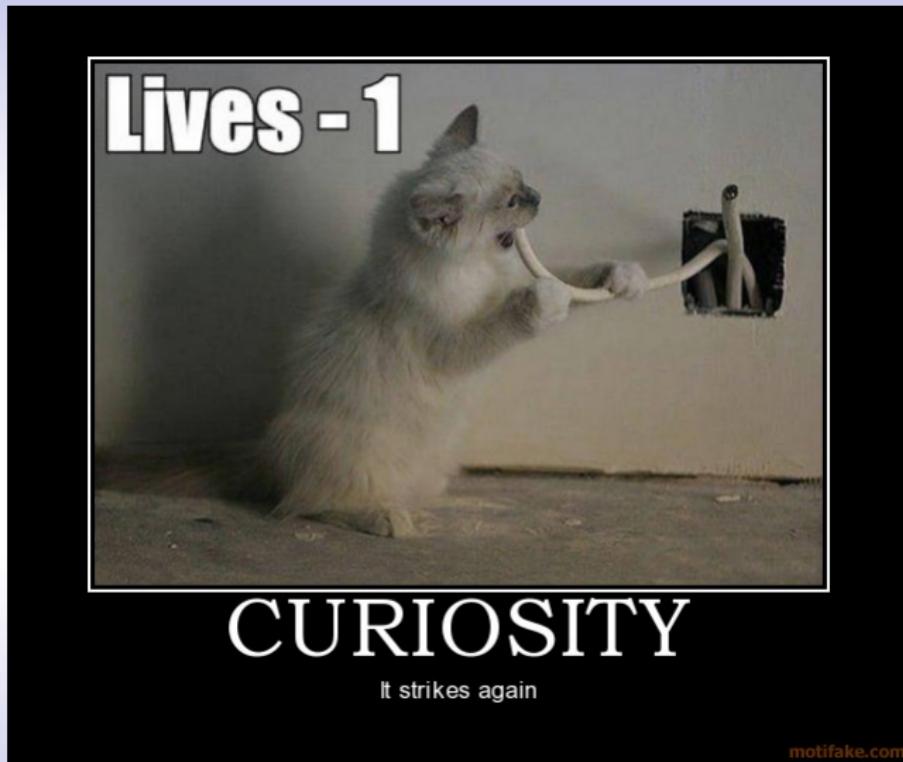
Open Viewer

## Modification Sites in Parent Protein, Orthologs, and Isoforms

## Show Multiple Sequence Alignment

SS	MS	human		mouse		rat		rabbit		monkey	
				▼ Show Isoforms							
6	0	P4	MEEPQSDPwVE	S4-p	MEESQSDISLE	S4-p	MEDQSDMSIE	S4	MEESQSDLSLE	P4	MEEPQSDPSIE
31	4	S6-p	MEEPQSDPwVEPP	S6-p	MEESQSDISLELP	S6-p	MEDQSDMSIELP	S6	MEESQSDLSLEPP	S6	MEEPQSDPSIEPP
34	3	S9-p	EEFwQDwTPEPLQ	S9-p	EEQwQDwTIELPLQ	S9-p	EDQwQDwIELPLwQ	S9	EEQSQDLSEPLLSQ	S9	EEFQSDPSIEPPLwQ
358	2	S15-p	PsVEPLLwQETwDL	S15-p	IsIELPLwQETwGL	S15-p	MsIELPLwQETwCwL	S15	LSLEPLLwQETwSDL	S15-p	PSIEPPPLwQETwESL
28	0	T18-p	EPFLwQETwDLwRL	T18-p	ELPLwQETwFaGLNL	T18-p	ELPLwQETwCwLwRL	T18	EPPLwQETwSDLwRL	T18	EPFLwQETwSDLwRL
110	1	S20-p	PLwQETwSDLwLLP	S20-p	PLwQETwGWLwLLP	S20-p	PLwQETwCwLwLLP	S20	PLwQETwSDLwLLP	S20	PLwQETwSDLwLLP
30	3	S33-p	LPEHNVLwPLPwQAH	S33	LPEHDLILPwPHCDL	S33	LPPDDILPwPHCDL	T33	LPENHVLwPLPSQAV	S33-p	LPENHVLwPLPSQAV
65	3	S37-p	NVLwPLPwQEDMLN	S34-p	PPEDILILPwPHCDL	S39-p	LPFTATGwPHMSDL	H37	HLTTTSLRwPVQDL	S37	NVLwPLPwQEVQDLN
85	2	S44-p	RHDDILLwPDDIEQW	L43	HCHDILLwPDDIEQW	L48	HSHEDDILLwPDDVRL	S45	PPVDDBLLwPDDVRL	S44	RVDDILLwPDDVRL
15	0	T55-p	DQDIEQWwEDDPGFD	-	qap	-	qap	H54	EDVAMWwEDPPEEGL	T55	DQDIEQWwEDDPGFD
2	0	D61	FEEDPGFDEAPRNPE	S55-p	EFEFFEGwEARLVRSG	E60	RELLEGGwPEALVRSG	E58	EWLHEEDwEDGLRVPA	D61	LTEDPGFDEAPRNSE
8	2	T81-p	RPRPAAPPTwPAPAPK	G75	DPYETEPwGPVAPAPK	A79	EPGEATAPwAPVAPAKA	A78	RPRPAAPPTwPAPAPK	T81	RPRPAAPPTwPAPAPK
0	2	S99-p	PLSSSVPSQRTYQGS	S93	PLSSVPwSQTYQGS	S97	PLSSSVPSQRTYQGS	S96	PLSSSVPSQRTYQHGS	S99	PLSSSVPSQRTYQHGS
1	2	K101-ub	SSSVPSQRTYQGSYG	E95	SSSVPSQRTYQHGSYG	E99	SSSVPSQRTYQHGSYG	K98	SSSVPSQRTYQHGSYG	K101	SSSVPSQRTYQHGSYG
1	0	S106-p	TSQTYQHGSYGFRLGF	H100	TSQTYQHGSYGFRLGF	H104	TSQTYQHGSYGFRLGF	H103	TSQTYQHGSYGFRLGF	S106	TSQTYQHGSYGFRLGF
0	1	H118-m1	YQGeYGFwLwGLwLQG	H114	YQGeYGFwLwGLwLQG	H118	YQGeYGFwLwGLwLQG	H117	YQGeYGFwLwGLwLQG	H118	YQGeYGFwLwGLwLQG
0	1	H115-m1	GPRwGENwGTRKSV	G119	GFMLGFwLQSGTAKSV	G113	GFMLGFwLQSGTAKSV	H112	GFMLGFwLQSGTAKSV	H115	GFMLGFwLQSGTAKSV
23	1	K128-ac	FLwSGTRwSVCTYs	K114-ac	FLwSGTRwSVCTYs	K118-ac	FLwSGTRwSVCTYs	K117	FLwSGTRwSVCTYs	K128	FLwSGTRwSVCTYs
1	19	K128-ub	FLAGSTGRwSVCTYs	K114	FLwSGTRwSVCTYs	K118	FLwSGTRwSVCTYs	K117	FLwSGTRwSVCTYs	K128	FLwSGTRwSVCTYs
1	0	Y126-p	RASVCTYwPALMRH	Y120	RASVCTYwSPPLNL	Y124	RASVCTYwSPPLNL	Y123	RASVCTYwSPPLNL	Y122	RASVCTYwSPPLNL
1	1	K137-ub	TISPLwHwPFCQLAK	K126	TISPLwHwPFCQLAK	K130	TISPLwHwPFCQLAK	K129	TISPLwHwPFCQLAK	K132	TISPLwHwPFCQLAK
1	0	K139-m1	AMFCQLAKTCPVQLW	K133	KLFCQLAKTCPVQLW	K137	KLFCQLAKTCPVQLW	K134	KLFCQLAKTCPVQLW	K139	KLFCQLAKTCPVQLW
3	1	S149-p	PVQLwWDwTPPFGR	E143	PVQLwVSwTPPFGR	S147	PVQLwVSwTPPFGR	S146	PVQLwVSwTPPFGR	S149	PVQLwVSwTPPFGR
1	1	S149-g1	PVQLwWDwTPPFGR	E143	PVQLwVSwTPPFGR	S147	PVQLwVSwTPPFGR	S146	PVQLwVSwTPPFGR	S149	PVQLwVSwTPPFGR
4	8	T158-p	VQLwWDwTPPFGRV	T144	VQLwVSwTPPFGRV	T148	VQLwVSwTPPFGRV	T147	VQLwVSwTPPFGRV	T158	VQLwWDwTPPFGRV
4	1	T155-p	DTTFPGwBVRMAM	S149-p	DTTFPGwBVRMAM	T153	DTTFPGwBVRMAM	T152	DTTFPGwBVRMAM	S155	DTTFPGwBVRMAM
4	1	K164-ac	VRMAMIXwQSONHTE	K158	VRMAMIXwQSONHTE	K162	VRMAMIXwQSONHTE	K161	VRMAMIXwQSONHTE	K164	VRMAMIXwQSONHTE
1	1	K164-ub	VRMAMIXwQSONHTE	K158	VRMAMIXwQSONHTE	K162	VRMAMIXwQSONHTE	K161	VRMAMIXwQSONHTE	K164	VRMAMIXwQSONHTE
2	0	S183-p	CPHHERCSDSDGLAF	S177	CPHHERCSDSDGLAF	S181	CPHHERCSDSDGLAF	S180	CPHHERCSDSDGLAF	S183	CPHHERCSDSDGLAF
0	1	R209-m1	RVEYLDDwHtFwHwM	R203	YPEXLEDwQTFTRHSV	R207	YAEYLDDwQTFTRHSV	R206	RAYYLDDwHtFwHwM	R209	RVEYLDDwHtFwHwM
1	0	T211-p	EXLDwHtFwHwVV	T205	EXLEDwHtFwHwVV	T209	EXLDwHtFwHwVV	T208	EXLDwHtFwHwVV	T211	EXSDDBRTwHwVV
0	1	R213-m1	LDDBRTwHwVVVY	R217	LDDBRTwHwVVVY	R211	LDDBRTwHwVVVY	R210	LDDBRTwHwVVVY	R213	SDDBRTwHwVVVY
4	0	S215-p	DrHtFwHwVVVYEP	S209	DrQTFTRHSwVVVYEP	S213	DrQTFTRHSwVVVYEP	S212	DrHtFTRHSwVVVYEP	S215	DRHtFTRHSwVVVYEP
1	0	Y228-p	RHSVVVYwPEPEVGs	Y214	RHSVVVYwPEPEVGs	Y218	RHSVVVYwPEPEVGs	Y217	RHSVVVYwPEPEVGs	Y228	RHSVVVYwPEPEVGs
0	1	C229	PPEVGSDCTTICWY	Y223-p	PPEVGSDCTTICWY	Y227	PPEVGSDCTTICWY	C226	PPEVGSDCTTICWY		

QUESTIONS?



Lives - 1

CURIOSITY

It strikes again

[motifake.com](http://motifake.com)

## The Eukaryotic Linear Motif resource for *Functional Sites in Proteins*

### The resource

is a collection of more than 260 thoroughly annotated motif classes with over 3000 annotated instances.

It is also a prediction tool to detect these motifs in protein sequences employing different filters to distinguish between **functional** and **non-functional** motif instances.

## The Eukaryotic Linear Motif resource for Functional Sites in Proteins

### The resource

is a collection of more than 260 thoroughly annotated motif classes with over 3000 annotated instances.

It is also a prediction tool to detect these motifs in protein sequences employing different filters to distinguish between **functional** and **non-functional** motif instances.

Functional sites	ELM classes	ELM instances	PDB structures	GO terms	PubMed links
Total	159	246	2702	348	549
By category					
LIG	137	Human	1594	Biological Process	283 From class
MOD	31	Mouse	253		1174
DEG	25	Rat	130		
DOC	22	Yeast	94	Cellular Compartment	119 From instance
TRG	20	Fly	90		1746
CLV	11	Other	541	Molecular Function	147

## ELM Class

Condensed information about a motif. Regular Expressions used to annotate the motif (eg.  
"[KR]xLx{0,1}[FYLIVMP]" for Cyclin motif)

**DOC\_CYCLIN\_1**

**Functional site class:** Cyclin recognition site  
**Functional site description:** Functional site that interacts with cyclins, and thereby increases the specificity of phosphorylation by cyclin/CDK complexes.  
**ELM with this motif:** [#DOC\\_CYCLIN\\_1](#)

**Description:** Substrate recognition site that interacts with cyclin and thereby increases phosphorylation by cyclin/cdk complexes. Predicted proteins should have a CDK phosphorylation site ([#MOD\\_CDCK\\_4](#)). Also used by cyclin/cdk inhibitors.

**Pattern:** [KR].L.([0,1])[FYLIVMP]  
**Pattern Probability:** 0.0053239  
**Present in taxon:** Eukaryota  
**Interaction Domain:** [#Cyclin\\_N \(PF000134\) Cyclin, N-terminal domain \(Stoichiometry: 1 : 1\)](#)

**PDB Structure:** [1H24](#)



## ELM Class

Condensed information about a motif. Regular Expressions used to annotate the motif (eg.  
"[KR]xLx{0,1}[FYLIVMP] for Cyclin motif)

■ 24 instances for DOC\_CYCLIN\_1  
(click table headers for sorting; Notes column: ▲=Number of Beitrags, ▼=Number of Interactions)

Protein Name	Gene Name	Start	End	Subsequence	Logic	#Ev.	Organism	Notes
JRB_HUMAN	JRB1	873	877	DKPFYFLKLLPDKQPKR	TP	3	Homo sapiens (Human)	1H25 ▲
JQUWJR_CHICK	JCDH1-A	394	398	KLQDSEKRYVFLKAMPPRER	FP	1	Gallus gallus (Chicken)	
JPMYTL_HUMAN	JPMYTL	486	489	QDPFPRFLKLLPDKLPLD	TP	1	Homo sapiens (Human)	
JEF21_HUMAN	JEF21	90	94	LQDFPFYFLKLLPDKQPKR	TP	3	Homo sapiens (Human)	1H24
JCONIC_HUMAN	JCDKNIC	31	34	VLQPFTRKFLKLLPDKQPKR	TP	1	Homo sapiens (Human)	
JRUX_DROSOPHILA	Jrux	248	251	PTKAKRCKFLKLLPDKQPKR	TP	1	Drosophila melanogaster (Fruit Fly)	
JEF22_HUMAN	JEF22	87	91	AQRLPAPFLKLLPDKQPKR	TP	1	Homo sapiens (Human)	
JEF23_HUMAN	JEF23	134	138	QDGPFLPFLKLLPDKQPKR	TP	1	Homo sapiens (Human)	
JAKA12_MOUSE	Jakap12	501	504	EVGDPYFLKLLPDKQPKR	TP	1	Mus musculus (House mouse)	1▲
JCDC6_HUMAN	JCDC6	94	98	RRRTLAKFLKLLPDKQPKR	TP	2	Homo sapiens (Human)	2CCH ▲
JCDN1A_HUMAN	JCDN1A	19	22	HPCSKRAFLKLLPDKQPKR	TP	4	Homo sapiens (Human)	1B ▲
JCDN1A_HUMAN	JCDN1A	155	159	KHSTTPTPFLKLLPDKQPKR	TN	1	Homo sapiens (Human)	
JORCS6_YEAST	JORCS6	178	182	RSPPTPFLKLLPDKQPKR	TP	1	Saccharomyces cerevisiae (Baker's yeast)	
JPS3_HUMAN	JTP53	381	385	QDGFPRFLKLLPDKQPKR	TP	5	Homo sapiens (Human)	1H26
JRBL1_HUMAN	JRBL1	658	661	HPNSKEAFKLLPDKQPKR	TP	3	Homo sapiens (Human)	1H28
JRBL2_HUMAN	JRBL2	680	684	PPAPTTTFLKLLPDKQPKR	TP	1	Homo sapiens (Human)	
JHRA_HUMAN	JHRA	629	633	KAASLAKFLKLLPDKQPKR	TP	1	Homo sapiens (Human)	

## DOC\_CYCLIN\_1

Functional site class:

Cyclin recognition site  
Functional site that interacts with cyclin, and thereby increases the specificity of phosphorylation by cyclin/CDK complexes.

ELM with this model:

■ DOC\_CYCLIN\_1

Description:

Substrate recognition site that interacts with cyclin and thereby increases phosphorylation by cyclin/cdk complexes. Predicted protein should have a CDK phosphorylation site (■ MOD\_CDk\_4). Also used by cyclin/cdk inhibitors.

Pattern:

(KR) . L . (D, I) [FYLLVPR]

0.0053239

Present in taxon:

■ Eukaryota

Interaction Domain:

■ Cyclin\_N (PF001934) Cyclin, N-terminal domain (Stoichiometry: 1 : 1)

PDB Structure: 1JTA4



## ELM Instance

An experimentally verified instance of an ELM class in a particular sequence.

## ELM Class

Condensed information about a motif. Regular Expressions used to annotate the motif (eg. "[KR]xLx{0,1}[FYLIVMP] for Cyclin motif)

### Instance

Sequence	Start	End	Subsequence	Logic	PDB	Organism	Length
...Q99741 CDCB_HUMAN	94	98	HEEILV <b>K</b> E <b>L</b> IVMP	TP		Homo sapiens (Human)	560

### Instance evidence

Evidence class	PMID	Method	BioSource	PubMed	Logic	Reliability	Notes
experimental	MI0114	x-ray crystallography	In vitro	Cheng,2006	support	certain	InteractionDetection FeatureDetection
experimental	MI0096	pull down	In vivo/In vitro	Petersen,1999	support	certain	InteractionDetection

This ELM instance is part of the following switching mechanism(s) annotated at the [ASwifts:ELM](#) resource:



## DOC\_CYCLIN\_1

Functional site class: Cyclin recognition site

Functional site description: Functional site that interacts with cyclins, and thereby increases the specificity of phosphorylation by cyclin/CDK complexes.

ELM with this model: #DOC\_CYCLIN\_1

Description:

Substrate recognition site that interacts with cyclin and thereby increases phosphorylation by cyclin/cdk complexes. Predicted proteins should have a CDK phosphorylation site ([MOD\\_CDk\\_4](#)). Also used by cyclin/cdk inhibitors.

Pattern:

[KR] . L . [D,I] [FYLLVRF]

Pattern Probability: 0.0053239

Present in taxon: Eukaryota

Interaction Domain: Cyclin\_N (PF000134) Cyclin, N-terminal domain (Stoichiometry: 1 : 1)

PDB Structure: 1IR4



## ELM Instance

An experimentally verified instance of an ELM class in a particular sequence.

- Experimental Evidences
- Methods
- References
- Interactions

## ELM Class

Condensed information about a motif. Regular Expressions used to annotate the motif (eg. "[KR]xLx{0,1}[FYLIVMP] for Cyclin motif)

### Instance

Sequence	Start	End	Subsequence	Logic	PDB	Organism	Length
CDCB_HUMAN	94	98	HEEILAEKQVYDQNLTKR	TP	2CCH	Homo sapiens (Human)	560

### Instance evidence

Evidence class	PMID	Method	BioSource	PubMed	Logic	Reliability	Notes
experimental	MI0114	x-ray crystallography	In vitro	Cheng,2006	support	certain	InteractionDetection FeatureDetection
experimental	MI0096	pull down	In vivo/In vitro	Petersen,1999	support	certain	InteractionDetection

This ELM instance is part of the following switching mechanism(s) annotated at the [Δswitches.ELM](#) resource:



## DOC\_CYCLIN\_1

Functional site class: Cyclin recognition site

Functional site description: Functional site that interacts with cyclins, and thereby increases the specificity of phosphorylation by cyclin/CDK complexes.

ELM with this model: [#DOC\\_CYCLIN\\_1](#)

Description:

Substrate recognition site that interacts with cyclin and thereby increases phosphorylation by cyclin/cdk complexes. Predicted proteins should have a CDK phosphorylation site ([MOD\\_CDk\\_4](#)). Also used by cyclin/cdk inhibitors.

Pattern:

[KR] . L . [D,I] [FYLLVRF]

Pattern Probability: 0.0053239

Present in taxon: Eukaryota

Interaction Domain: Cyclin\_N (PF000134) Cyclin, N-terminal domain (Stoichiometry: 1 : 1)

PDB Structure: [1H24](#)

PDB Structure: [1H24](#)



## ELM Instance

An experimentally verified instance of an ELM class in a particular sequence.

### Experimental Evidences

- Methods
- References
- Interactions

## ELM Class

Condensed information about a motif. Regular Expressions used to annotate the motif (eg. "[KR]xLx{0,1}[FYLIVMP] for Cyclin motif)

### Instance

Sequence	Start	End	Subsequence	Logic	PDB	Organism	Length
...Q99741 CDCB_HUMAN	94	98	HEEILV <b>K</b> E <b>L</b> IVMP	TP	2CCH	Homo sapiens (Human)	560

### Instance evidence

Evidence class	PMID	Method	BioSource	PubMed	Logic	Reliability	Notes
experimental	MI0114	x-ray crystallography	In vitro	Cheng,2006	support	certain	InteractionDetection FeatureDetection
experimental	MI0096	pull down	In vivo/In vitro	Petersen,1999	support	certain	InteractionDetection

This ELM instance is part of the following switching mechanism(s) annotated at the [Δswitches.ELM](#) resource:



## DOC\_CYCLIN\_1

Functional site class: Cyclin recognition site

Functional site description: Functional site that interacts with cyclins, and thereby increases the specificity of phosphorylation by cyclin/CDK complexes.

ELM with this model: [#DOC\\_CYCLIN\\_1](#)

Description:

Substrate recognition site that interacts with cyclin and thereby increases phosphorylation by cyclin/cdk complexes. Predicted proteins should have a CDK phosphorylation site ([MOD\\_CDK\\_4](#)). Also used by cyclin/cdk inhibitors.

Pattern:

[KR] . L . [D,I] [FYLLVRF]

Pattern Probability: 0.0053239

Present in taxon: Eukaryota

Interaction Domain: Cyclin\_N (PF000134) Cyclin, N-terminal domain (Stoichiometry: 1 : 1)

PDB Structure: [1ITa4](#)



## ELM Instance

An experimentally verified instance of an ELM class in a particular sequence.

- Experimental Evidences
- Methods
- References
- Interactions

## ELM Class

Condensed information about a motif. Regular Expressions used to annotate the motif (eg. "[KR]xLx{0,1}[FYLIVMP] for Cyclin motif)

### Instance

Sequence	Start	End	Subsequence	Logic	PDB	Organism	Length
...Q99741 CDCB_HUMAN	94	98	HEEILV <b>K</b> E <b>L</b> IVMP	TP	2CCH	Homo sapiens (Human)	560

### Instance evidence

Evidence class	PMID	Method	BioSource	PubMed	Logic	Reliability	Notes
experimental	MI0114	x-ray crystallography	In vitro	Cheng,2006	support	certain	InteractionDetection FeatureDetection
experimental	MI0096	pull down	In vivo/In vitro	Petersen,1999	support	certain	InteractionDetection

This ELM instance is part of the following switching mechanism(s) annotated at the [Δswitches.ELM](#) resource:



## DOC\_CYCLIN\_1

Functional site class: Cyclin recognition site

Functional site description: Functional site that interacts with cyclins, and thereby increases the specificity of phosphorylation by cyclin/CDK complexes.

ELM with this model: #DOC\_CYCLIN\_1

Description:

Substrate recognition site that interacts with cyclin and thereby increases phosphorylation by cyclin/cdk complexes. Predicted proteins should have a CDK phosphorylation site ([MOD\\_CDk\\_4](#)). Also used by cyclin/cdk inhibitors.

Pattern:

[KR] . L . [D,E] [FYLLVRF]

Pattern Probability: 0.0053239

Present in taxon: Eukaryota

Interaction Domain: Cyclin\_N (PF000134) Cyclin, N-terminal domain (Stoichiometry: 1 : 1)

PDB Structure: [3I7A4](#)



## ELM Instance

An experimentally verified instance of an ELM class in a particular sequence.

- Experimental Evidences
- Methods
- References
- Interactions

## ELM Class

Condensed information about a motif. Regular Expressions used to annotate the motif (eg. "[KR]xLx{0,1}[FYLIVMP] for Cyclin motif)

■ Instance							
Sequence	Start	End	Subsequence	Logic	PDB	Organism	Length
D(Q99741) CDCB_HUMAN	94	98	RRRLTLLGGWVQNGLQKAA	TP		Homo sapiens (Human)	560

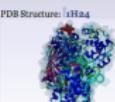
■ Instance evidence							
Evidence class	PSMI	Method	BioSource	PubMed	Logic	Reliability	Notes
experimental	MI0114	x-ray crystallography	in vitro	5 Cheng, 2006	support	certain	InteractionDetection FeatureDetection
computational	MI0001	multi-class	In silico prediction	10000 Redondo, 2000	support	certain	InteractionPrediction
computational	MI0001	multi-class	In silico prediction	10000 Redondo, 2000	support	certain	InteractionPrediction

This ELM instance is part of the following mediation configuration(s) presented in the *Mediations ELM* page:



DOC CYCLIN 1

<b>Functional site class:</b>	Cyclin recognition site
<b>Functional site description:</b>	Functional site that interacts with cyclin, and thereby increases the specificity of phosphorylation by cyclin/CDK complexes.
<b>ELM with this model:</b>	<a href="#">#DOC_CYCLIN_1</a>
<b>Description:</b>	Substrate recognition site that interacts with cyclin and thereby increases phosphorylation by cyclin/cdk complexes. Predicted proteins should have a CDK phosphorylation site ( <a href="#">#MOD_CDK_x</a> ). Also used by cyclin/cdk inhibitors.
<b>Pattern:</b>	[RK]...L...[D E][VFLIVMP]
<b>Pattern Probability:</b>	0.0053239
<b>Present in taxon:</b>	Eukaryota
<b>Interaction Domain:</b>	# Cyclin_N (PF00134) Cyclin_N-terminal domain (Stoichiometry: 1)
	PDB Structure: <a href="#">1H24</a>



## ELM Instance

An experimentally verified instance of an ELM class in a particular sequence.

- Experimental Evidences
  - Methods
  - References
  - Interactions

*"iELM – a web server to explore short linear motif-mediated interactions."*; WEATHERITT, JEHL, DINKEL & GIBSON ; (NUCLEIC ACIDS RES. 2012) Table 8 Database



# The Eukaryotic Linear Motif resource for Functional Sites in Proteins

[ELM Home](#)   [ELM Prediction](#)   [ELM DB](#)   [ELM Candidates](#)   [ELM](#)
[ELM classes](#)
[ELM instances](#)
[ELM pdb structures](#)
[ELM binding domains](#)

## SEARCH the ELM

The ELM relational database is curated from the literature. The site contains one to many ELM pathways validated motif instances matching the website according to the following:

[239 annotated ELM classes](#)
[2,675 experimentally validated ELM methods](#)
[100 ELM pathogenic abuse](#)
[358 solved PDB structures](#)
[118 globular ELM binders](#)
[1,031 interactions mediated by motifs](#)
[836 regulatory switches mediated by curated ELM instances](#)
[784 pathways from KEGG involving linear motifs annotated in 832 Sequences](#)
[219 viral instances interfering with host cellular processes](#)
[11 ELM related diseases annotated as being caused by aberrant motif function](#)
[2 examples where pathogens abuse motifs to deregulate host cells](#)
[Search ELM Instances and Classes](#)



kelch

ELM CLASS: DEG\_Kelch\_actinfilin\_1

ELM CLASS: DEG\_Kelch\_Keap1\_1

ELM CLASS: DEG\_Kelch\_Keap1\_2

ELM CLASS: DEG\_Kelch\_KLHL3\_1

INSTANCE: P42260 GRK2\_RAT [881:885] DEG\_Kelch\_actinfilin\_1

INSTANCE: Q14494 NF2L1\_HUMAN [231:236] DEG\_Kelch\_Keap1\_1

INSTANCE: Q16236 NF2L2\_HUMAN [77:82] DEG\_Kelch\_Keap1\_1

INSTANCE: P20482 CNC\_DROME [458:463] DEG\_Kelch\_Keap1\_1

INSTANCE: Q13501 SQSTM\_HUMAN [347:352] DEG\_Kelch\_Keap1\_1

INSTANCE: Q96HS1-1 PGAMS\_HUMAN [77:82] DEG\_Kelch\_Keap1\_1

INSTANCE: O14920 IKKB\_HUMAN [34:39] DEG\_Kelch\_Keap1\_1

INSTANCE: Q5JTC6 AMER1\_HUMAN [286:291] DEG\_Kelch\_Keap1\_1

INSTANCE: Q86YC2 PALB2\_HUMAN [89:94] DEG\_Kelch\_Keap1\_1

INSTANCE: Q13402 MYO7A\_HUMAN [1636:1641] DEG\_Kelch\_Keap1\_1

INSTANCE: Q12830 BPTF\_HUMAN [729:734] DEG\_Kelch\_Keap1\_1

■ DEG\_SPOP\_SBC\_1,

■ DOC\_GSK3\_Axin\_1,

■ LIG\_CID\_NIM\_1,

■ LIG\_GBD\_WASP\_1,

■ LIG\_Mtr4\_Air2\_1,

■ LIG\_Mtr4\_Trf4\_1,

■ LIG\_Mtr4\_Trf4\_2,

■ LIG\_Pex14\_3,

■ LIG\_Pex14\_4,

■ LIG\_RPA\_C\_Funaf,

■ LIG\_RPA\_C\_Insects,

■ LIG\_RPA\_C\_Plants,

■ LIG\_RPA\_C\_Vert,

■ MOD\_SUMO\_rev\_2

[«MOD\\_WntLipid»](#)[»TRG\\_Cilium\\_Arf4\\_1»](#)

## TRG\_AP2beta\_CARGO\_1

**Accession:** [ELME000247](#)

**Functional site class:** AP-2 beta2 appendage CCV component motifs

**Functional site description:** Several motifs are responsible for the binding of accessory endocytic proteins to the beta2-subunit appendage of the adaptor protein complex AP-2 as part of their recruitment to the site of clathrin coated vesicle (CCV) formation. Proteins binding the platform subdomain have been found to be cargo family specific (for example can load all GPCRs, or all LDL receptor family members) clathrin adaptors. Accessory proteins which help in CCV formation bind the sandwich subdomain site or the alpha ear domain.

**ELM Description:** Motif binding as a helix in a depression on the top surface of the AP-2 beta appendage platform subdomain. The pattern [ED]x{1,2}Fxx[FL]xxxR is conserved in beta Arrestins, ARH and Epsin-1, -2 of vertebrates. It is also found in homologues of other metazoans, but the pattern is sometimes not matched exactly, meaning that the ELM regular expression will not provide a match. In other lineages, if there is an equivalent motif, the pattern is likely to have diverged.

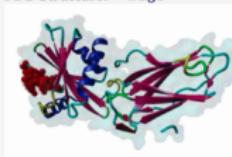
**Pattern:** [DE].{1,2}F[^P][^P][FL][^P][^P][^P]R

**Pattern Probability:** 0.0000182

**Present in taxon:** [Metazoa](#)

**Interaction Domain:** B2-adapt-app\_C (PF09066) Beta2-adaptin appendage, C-terminal sub-domain (Stoichiometry: 1 : 1)

PDB Structure: [2G30](#)




[Search](#) | [ELMs](#) | [Instances](#) | [Candidates](#) | [Links](#) | [About](#) | [News](#) | [Help](#) | [Diseases](#)

## Search ELM Instances

Full-Text Search (to show all instances, enter 'all' or '')

Filter by instance Logic true positive | Filter by organism Homo sapiens

export 58 instances as: [fasta](#) | [tsv](#)

■ 58 Instances for search term 'ap2':

(click table headers for sorting)

CLV  
LIG  
MOD  
TRG

ELM identifier	Sequence	Start	End	Subsequence	Instance Logic	#Evidence	PDB	Organism
TRG_LysEnd_APsAcLL_1	OPRD_HUMAN	241	246	GLMLLRLRSVRLLSGSKEKD	true positive	8	---	Homo sapiens (Human)
TRG_AP2beta_CARGO_1	ARRB1_HUMAN	385	395	TNDDIVFEDFARQRLKGMK	true positive	5	2IV8	Homo sapiens (Human)
TRG_LysEnd_APsAcLL_1	HG2A_HUMAN	19	24	DQKPVMDQQRDLISNNQLP	true positive	5	---	Homo sapiens (Human)
LIG_AP2alpha_2	EPS15_HUMAN	672	674	DPFTATSSTDPPFAANNSIT	true positive	4	---	Homo sapiens (Human)
LIG_AP2alpha_2	EPS15_HUMAN	692	694	SVETLKHNDPPFAAPGTVVAA	true positive	4	---	Homo sapiens (Human)
LIG_AP2alpha_2	EPS15_HUMAN	709	711	VAASDSATDPFAASVFGNESF	true positive	4	---	Homo sapiens (Human)
LIG_AP2alpha_2	EPS15_HUMAN	737	739	TLSKVNNEDPFRSATSSSVS	true positive	4	---	Homo sapiens (Human)
TRG_AP2beta_CARGO_1	EPN1_HUMAN	377	386	FDTEPDEFSDFDRRLTALPT	true positive	4	---	Homo sapiens (Human)
TRG_LysEnd_APsAcLL_1	ATP7A_HUMAN	1483	1488	SVVTSEPRKHSLLVGDFRED	true positive	4	---	Homo sapiens (Human)
LIG_SxIP_EBH_1	CLAP2_HUMAN	492	502	ASAOKRSKIPRPSQGCSEAS	true positive	3	---	Homo sapiens (Human)
LIG_SxIP_EBH_1	CLAP2_HUMAN	515	525	LSVARSSRIPRPPSVSQCSR	true positive	3	---	Homo sapiens (Human)
TRG_LysEnd_APsAcLL_1	BCAM_HUMAN	604	609	HSGSEQEPQTGLLMMGASGG	true positive	3	---	Homo sapiens (Human)
TRG_LysEnd_APsAcLL_1	NPC1_HUMAN	1271	1276	KSCATEERYKGTERERLLNF	true positive	3	---	Homo sapiens (Human)
LIG_APCC_KENbox_2	CKAP2_HUMAN	80	84	KLTKTKMADKENMKRPAESKN	true positive	2	---	Homo sapiens (Human)
LIG_MAPK_1	MP2K1_HUMAN	3	11	MPKKKPTPIQLNPADGSAV	true positive	2	---	Homo sapiens (Human)
LIG_MAPK_1	MP2K4_HUMAN	40	48	SSMQGKRKAALKLNFPNPPFK	true positive	2	---	Homo sapiens (Human)
TRG_AP2beta_CARGO_1	ARH_HUMAN	256	266	DDGLDEAFSRLAQSRTNPQV	true positive	2	2G30	Homo sapiens (Human)
TRG_LysEnd_APsAcLL_1	CD44_HUMAN	708	713	GEASKS0EMVHLNVNESSET	true positive	2	---	Homo sapiens (Human)
LIG_AP2alpha_1	AMPH_HUMAN	324	328	QENIISFEDNFVPEISVTT	true positive	1	1KY7	Homo sapiens (Human)
LIG_AP2alpha_2	EP15R_HUMAN	599	601	RGSFGAMDPPFKNKALLFSN	true positive	1	---	Homo sapiens (Human)
LIG_AP2alpha_2	EP15R_HUMAN	618	620	NTNTQELHPDPFQTEDPFKSD	true positive	1	---	Homo sapiens (Human)



## Diseases mediated by short linear motifs

Several diseases are known which are caused by one or more mutations in linear motifs mediating important interactions. Below you find a selection of such diseases; for linear motifs abused by viruses, see the dedicated [Viruses](#) page. For a large-scale analysis on disease-causing mutations see [\[Proteome-wide analysis of human disease mutations in short linear motifs: neglected players in cancer? Uyar B, et al., 2014\]](#)

### Noonan Syndrome

The developmental disorder "Noonan Syndrome" can be caused by mutations in [DRaf-1](#) which abrogate the interaction with 14-3-3 proteins mediated by corresponding motifs and thereby deregulate the Raf-1 kinase activity [[Pandit et al., 2007](#)]. The [DRaf-1](#) sequence features two [LIG\\_14-3-3\\_1](#) binding sites, which are annotated at [256-261](#) and [618-623](#).

### Noonan-like Syndrome

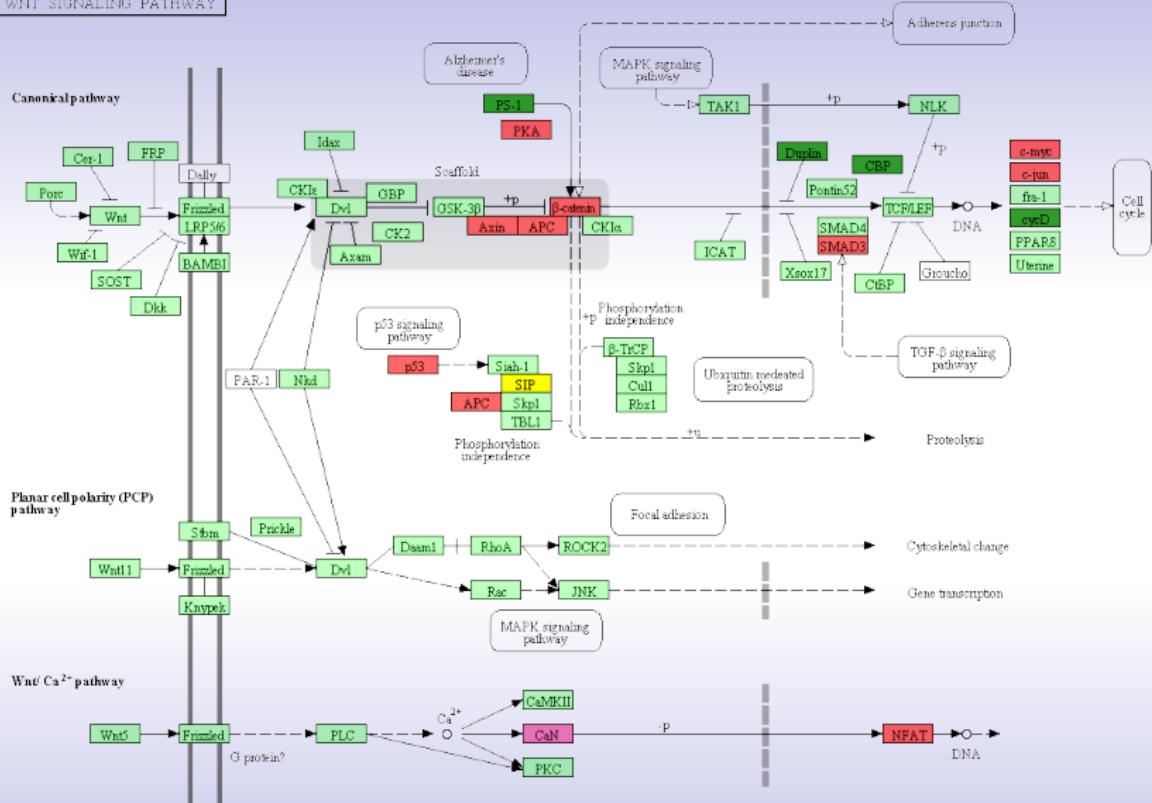
A S->G mutation at position 2 creates a novel [MOD\\_NMyristoyl](#) site (irreversible modification) resulting in aberrant targeting of SHOC2 to the plasma membrane and impaired translocation to the nucleus upon growth factor stimulation [[Cordedu et al., 2007](#)].

### Usher's Syndrome

"Usher's Syndrome" is the most frequent cause of hereditary deaf-blindness in humans [[Eudy and Sumegi, 1999](#)], affecting one child in 25 000. This disease can be caused by mutations in either PDZ domains in [Harmonin](#) or the corresponding PDZ interaction motifs in the [SANS](#) protein (annotated at [456-461](#)) [[Weil et al., 2003](#), [Kalay et al., 2005](#)].

Another example implicating PDZ domains is "*familial hypomagnesemia with hypercalciuria and nephrocalcinosis*" (FHHN), an autosomal recessive wasting disorder of renal Mg<sup>2+</sup> and Ca<sup>2+</sup> that leads to progressive kidney failure. Here, motifs mediating interaction to PDZ domains are mutated in [Claudin 16](#), abolishing important interactions to the scaffolding protein [ZO-1](#) resulting in lysosomal mislocalization of the protein [[Müller et al., 2003](#), [Müller et al., 2006](#)].

## WNT SIGNALING PATHWAY





[ELM Home](#) [ELM Prediction](#) [ELM DB](#) [ELM Candidates](#) [ELM Information](#) [ELM downloads](#)
[Help](#)

## Functional site prediction

### Protein sequence

Enter Uniprot identifier or accession number: (auto-completion)  
e.g. **EPN1\_HUMAN, P04637, TAU\_HUMAN, [RANDOM]**

**EPN1**  
CARP\_CRYPA [P11838] Cyphonotria parasitica  
EPD1\_CARAU [P13506] Carassius auratus  
**EPN1\_ARATH** [Q8VY07] Arabidopsis thaliana  
**EPN1\_HUMAN** [Q9Y6I3] Homo sapiens  
**EPN1\_MOUSE** [Q80VP1] Mus musculus  
**EPN1\_RAT** [O88339] Rattus norvegicus  
F2QLC2\_PICP7 [F2QLC2] Komagataella pastoris  
KOKY34\_WICCF [KOKY34] Wickerhamomyces cferrill  
A0A024R4S1\_HUMAN [A0A024R4S1] Homo sapiens  
K7EMP4\_HUMAN [K7EMP4] Homo sapiens  
W8B7F4\_CERCA [W8B7F4] Ceratitis capitata  
A8X4H2\_CAEBR [A8X4H2] Caenorhabditis briggsae  
QB1I71\_CAEEL [QB1I71] Caenorhabditis elegans  
QB1I70\_CAEEL [QB1I70] Caenorhabditis elegans  
cytosol  
peroxisome  
glycosome  
glyoxosome

(ASTA format):

```
SEIMSMNIWKRLNDHGKNWRHVYKAMTL
DLRDEDRLLREERAHALKTTEKLQAQTATA
QLALSLSLSREEHDEKEERIIRGGDLRLQM
IDFWGGPPVPPAADPWGGPAPTPASGDP
FAFSDPNGGSPAKPSTNGTAAAGGFDT
ISPPPAATPTPTPPTRKTPESFLGPNA
RLSPVPPPVGAPPTYISPLGGGGPLPP
```

### Taxonomic Context

Type in species name (auto-completion):

Homo sapiens

### Motif Probability Cutoff:

- **ELM database update**
- The following elm classes have been added to the database:
- [DEG\\_Kelch\\_actinfilin\\_1](#),
- [DEG\\_Kelch\\_Keap1\\_1](#),
- [DEG\\_Kelch\\_Keap1\\_2](#),
- [DEG\\_Kelch\\_KLHL3\\_1](#),
- [DEG\\_Nend\\_Nbox\\_1](#),
- [DEG\\_Nend\\_UBRbox\\_1](#),
- [DEG\\_Nend\\_UBRbox\\_2](#),
- [DEG\\_Nend\\_UBRbox\\_3](#),
- [DEG\\_Nend\\_UBRbox\\_4](#),
- [DEG\\_SPOP\\_SBC\\_1](#),
- [DOC\\_GSK3\\_Axin\\_1](#),
- [LIG\\_CID\\_NIM\\_1](#),
- [LIG\\_GBD\\_WASP\\_1](#),
- [LIG\\_Mtr4\\_Air2\\_1](#),
- [LIG\\_Mtr4\\_Trf4\\_1](#),
- [LIG\\_Mtr4\\_Trf4\\_2](#),
- [LIG\\_Pex14\\_3](#),
- [LIG\\_Pex14\\_4](#),
- [LIG\\_RPA\\_C\\_Fungi](#),
- [LIG\\_RPA\\_C\\_Insects](#),
- [LIG\\_RPA\\_C\\_Plants](#),
- [LIG\\_RPA\\_C\\_Vert](#),
- [MOD\\_SUMO\\_rev\\_2](#)

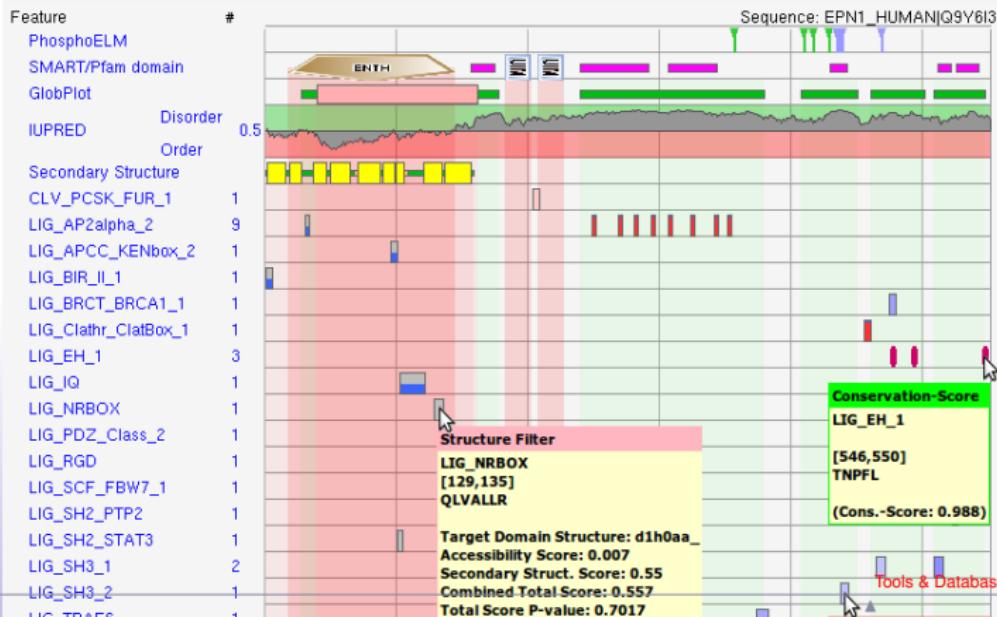
Many new instances: 8

## Summary of features reported by the ELM resource.

### KEY

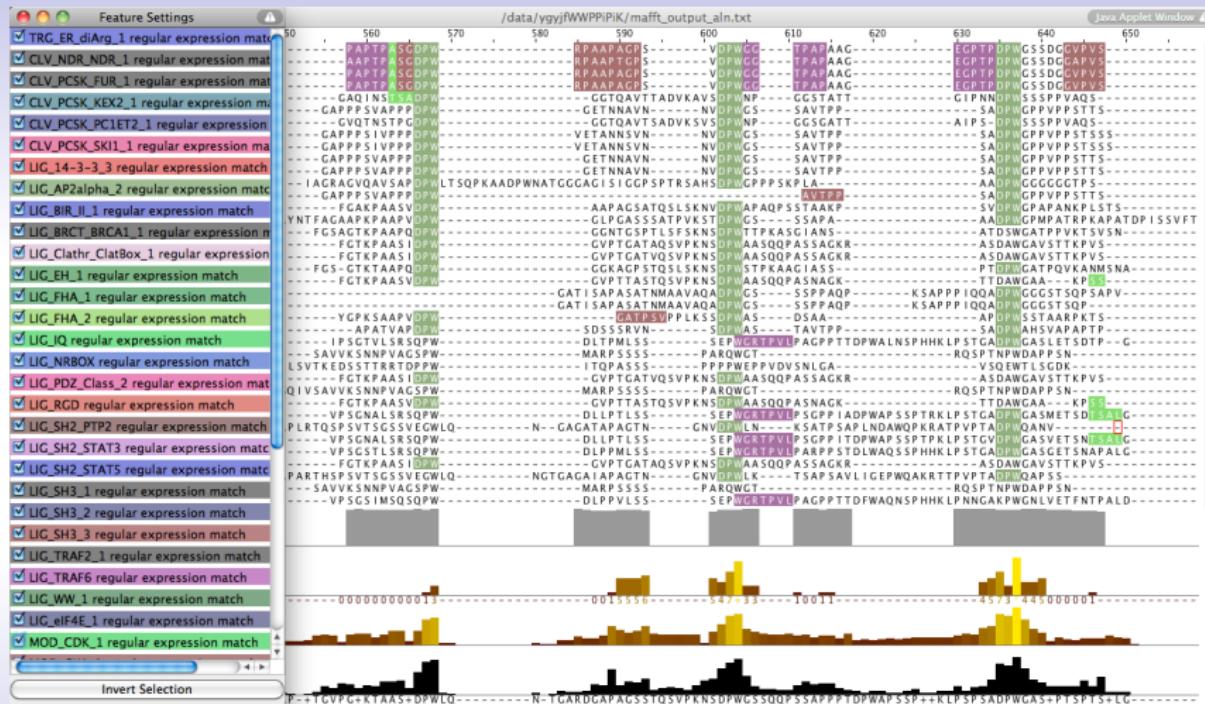
DOMAINS:	Smart/Pfam domain	Signal peptide (pred.)	Low-complexity region	Coiled-coil (pred.)	TM helix (pred.)
GLOBPLOT:	GlobDom		Disorder		
2D STRUCT:	Strand	Helix	Loop	3/10 Helix	
MOTIFS:	Favourable Context	Sparse/Smart filtered	Neutral	Annotated:	TP  FP  TN  FN  <  <  Assigned by homology
CONSCORE:	low Conservation	medium Conservation	high Conservation		

(Mouseover the matches for more details )



# VIEW CONSERVATION IN JALVIEW

ELM



Invert Selection

# Questions?



# CURIOSITY

Do you really want to know?

[fakeposters.com](http://fakeposters.com)

## PROTEIN VISUALIZATION (PROVIZ)

**ProViz** <http://proviz.ucd.ie/> is a tool to visualize biological data allowing the investigation of functional and evolutionary protein features. The tool is designed to be an intuitive and accessible resource to allow users with limited bioinformatic skills to rapidly access and visualise data pertinent to their research.

*"ProViz-a web-based visualization tool to investigate the functional and evolutionary features of protein sequences."*; JEHL P, MANGUY J, SHIELDS DC, HIGGINS DG, DAVEY NE.; (NUCLEIC ACIDS RES. 2016 APR 16)

# PROTEIN VISUALIZATION (PROVIZ)



"ProViz-a web-based visualization tool to investigate the functional and evolutionary features of protein sequences."; JEHL P, MANGUY J, SHIELDS DC, HIGGINS DG, DAVEY NE.; (NUCLEIC ACIDS RES. 2016 APR 16)

Tools & Databases of Short Linear Motifs

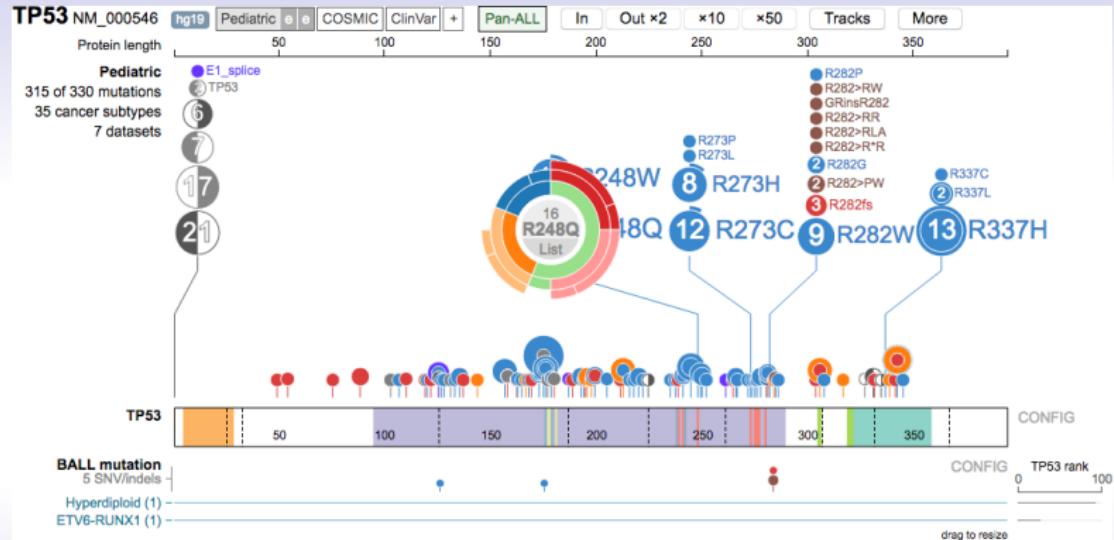
## PROTEIN VISUALIZATION (PROVIZ)



"ProViz-a web-based visualization tool to investigate the functional and evolutionary features of protein sequences.;" JEHL P, MANGUY J, SHIELDS DC, HIGGINS DG, DAVEY NE.; (NUCLEIC ACIDS RES. 2016 APR 16) [Tools & Databases](#)

## Tools & Databases of Short Linear Motifs

# PECAN / PROTEINPAINT – VISUALIZATION OF PEDIATRIC CANCER MUTATIONS



# Questions?



CURIOSITY KILLED THE CAT

Good boy curiosity.....  
Good boy!!!

[motifake.com](http://motifake.com)