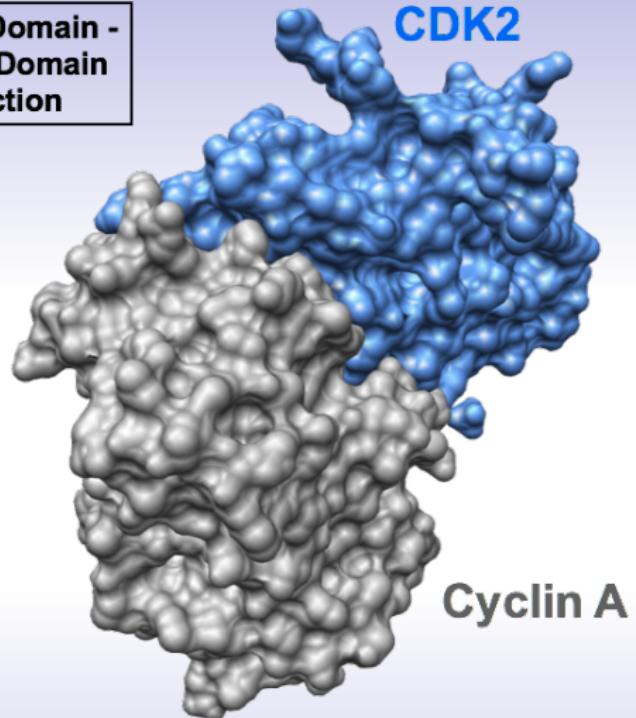


SHORT LINEAR MOTIFS

Holger Dinkel

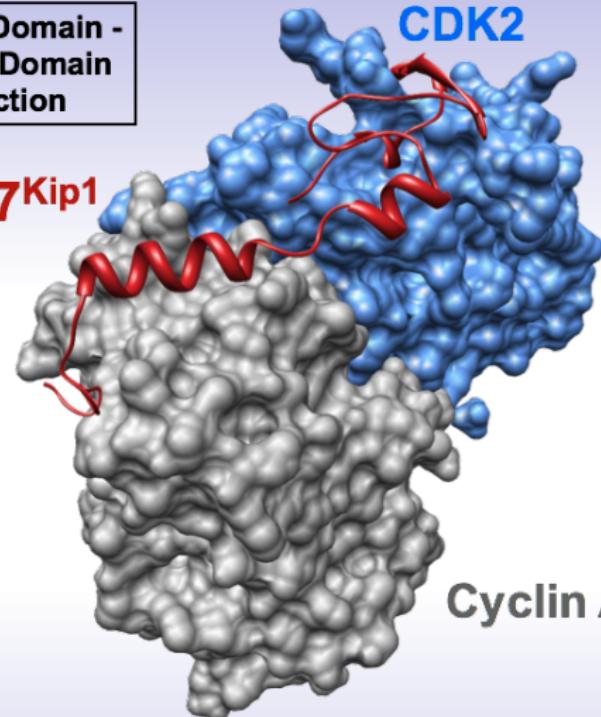
EMBO Practical Course “Computational analysis of
protein-protein interactions – From sequences to
networks”

**Globular Domain -
Globular Domain
Interaction**



Globular Domain -
Disordered Domain
Interaction

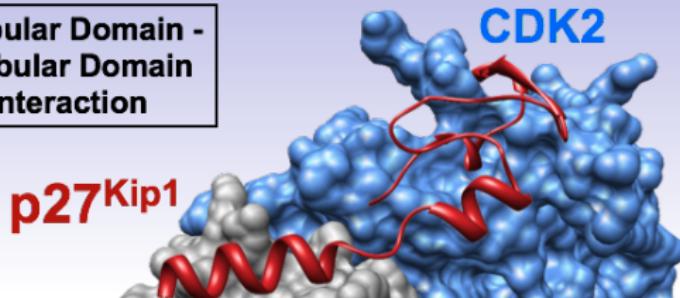
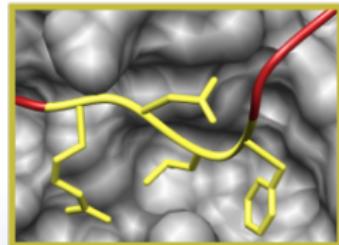
Globular Domain -
Globular Domain
Interaction



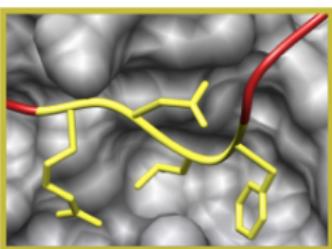
Globular Domain -
Globular Domain
Interaction

Globular Domain -
Disordered Domain
Interaction

Globular Domain -
Short Linear Motif
Interaction



Cyclin A

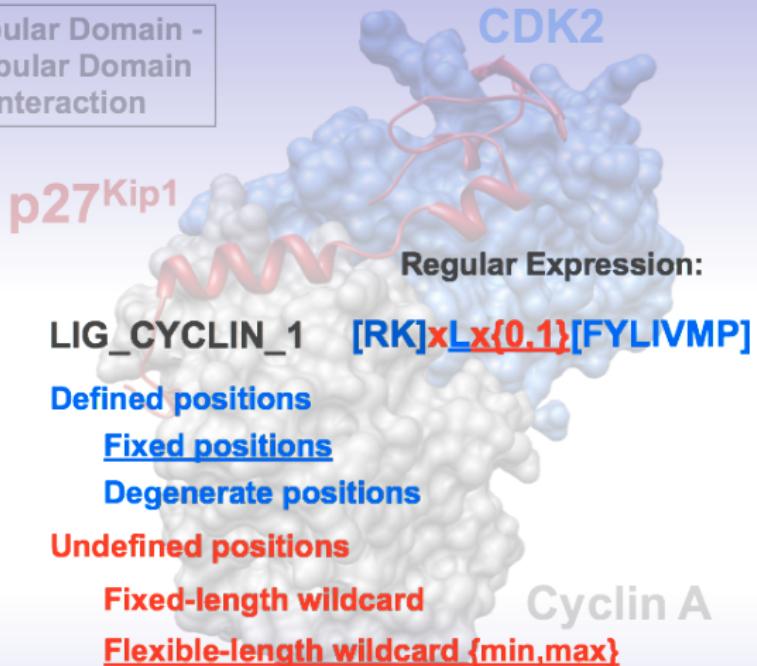


RNLF

Globular Domain -
Disordered Domain
Interaction

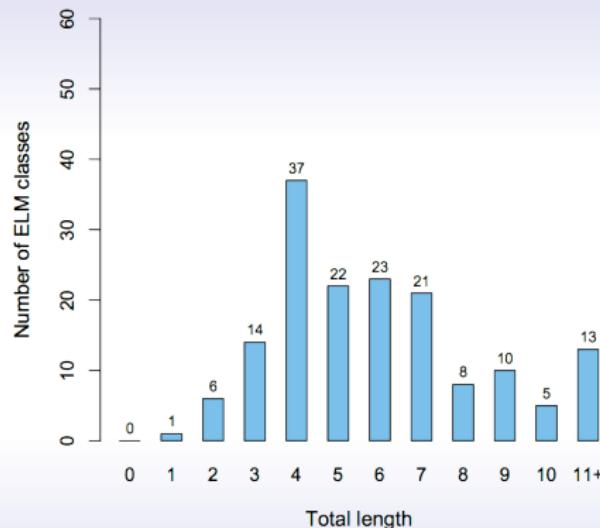
Globular Domain -
Globular Domain
Interaction

PDB 1JSU
Russo *et al.*, Nature, 1996;
382: 325-331.



LINEAR MOTIFS

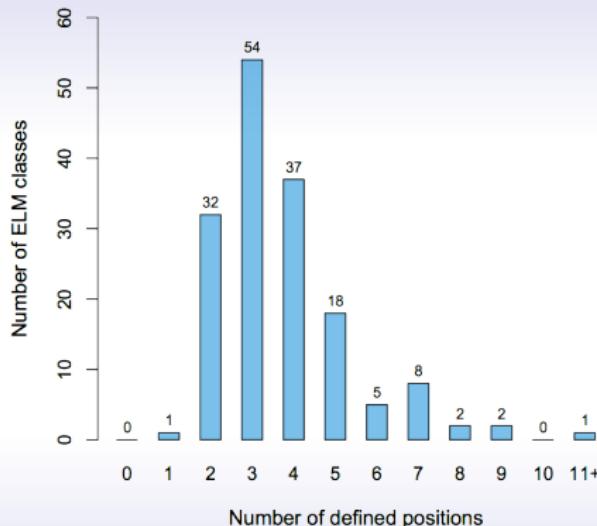
- are small.
- have few defined positions.
- mediate transient, low affinity interactions.



"Attributes of short linear motifs"; DAVEY, VAN ROEY, WEATHERITT, TOEDT, UYAR, ALtenBERG, BUDD, DIELLA, DINKEL & GIBSON; (MOL BIOSYST. 2011)

LINEAR MOTIFS

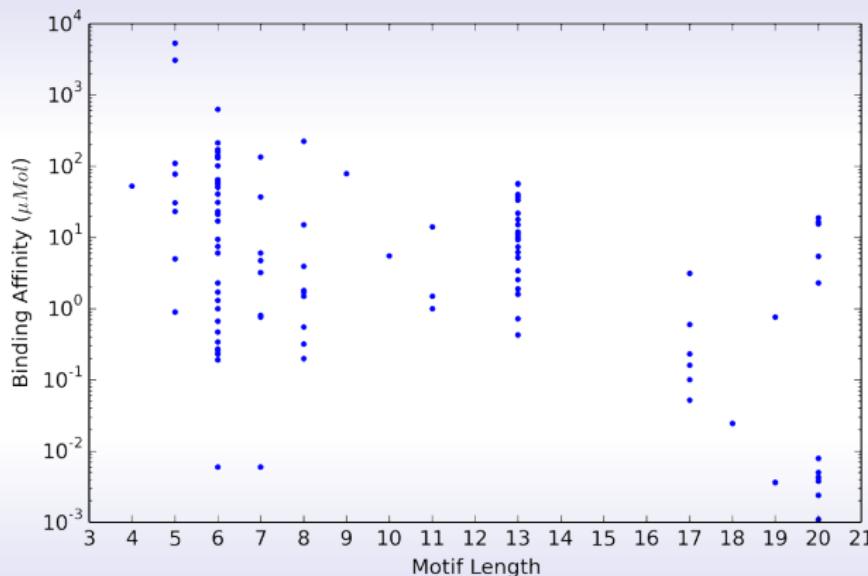
- are small.
- have few defined positions.
- mediate transient, low affinity interactions.



"Attributes of short linear motifs"; DAVEY, VAN ROEY, WEATHERITT, TOEDT, UYAR, ALtenberg, BUDD, DIELLA, DINKEL & GIBSON; (MOL BIOSYST. 2011)

LINEAR MOTIFS

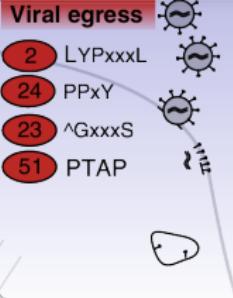
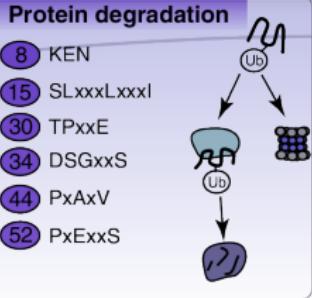
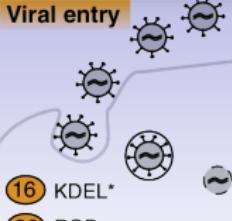
- are small.
- have few defined positions.
- mediate transient, low affinity interactions.



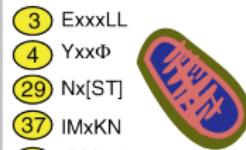
DOMAIN FREQUENCIES FROM PFAM (HUMAN PROTEOME):

Domain Family	Frequency [Domains / Proteins]	Pattern of recognized motif
PDZ	573 / 342	$[ST]x[ACVILF]_{-COOH}$
SH3	451 / 382	$PxxP$
SH2	237 / 219	${}_pYxx[IV]$
WW	151 / 103	$PPxY$
PTB	142 / 133	$NPx{}_pY$

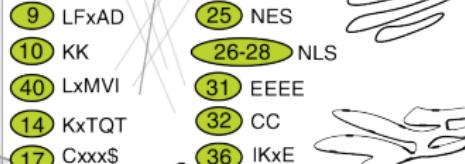
Extracellular



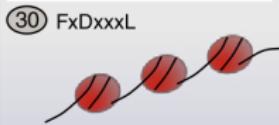
Cyttoplasmic

Immune response**Cleavage**

- 7 Dxxd ↓
18 R↓S

Translation**Transport**

Nuclear

Cell cycle**Epigenetic regulation****Transcriptional regulation**

- 5 PxLxP
6 SPxLxLT
11 PxDSL
19 MxLRSG
- 46 RxxPDG
47 LxxLxxxxL



"How viruses hijack cell regulation"; DAVEY, TRAVÉ & GIBSON; (TIBS 2010)

LIDDLE'S-SYNDROME: WW-INTERACTION MOTIF

has been implicated with autosomal dominant activating mutations in the WW interaction motif in the β - and γ -subunits of the epithelial sodium channel **ENAC**. These mutations abrogate the binding to the ubiquitin ligase **NEDD4-2**, ultimately resulting in increased Na^+ reabsorption, plasma volume extension and hypertension.

See <http://elm.eu.org/infos/diseases.html> for a list of motifs implicated in various diseases.

LIDDLE'S-SYNDROME: WW-INTERACTION MOTIF

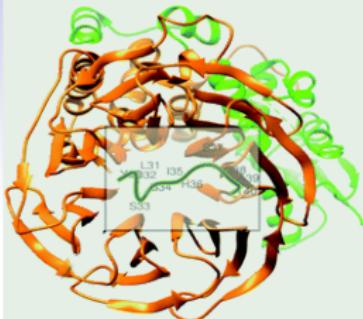
has been implicated with autosomal dominant activating mutations in the WW interaction motif in the β - and γ -subunits of the epithelial sodium channel **ENAC**. These mutations abrogate the binding to the ubiquitin ligase **NEDD4-2**, ultimately resulting in increased Na^+ reabsorption, plasma volume extension and hypertension.

BACILLUS ANTHRACIS “LETHAL FACTOR”

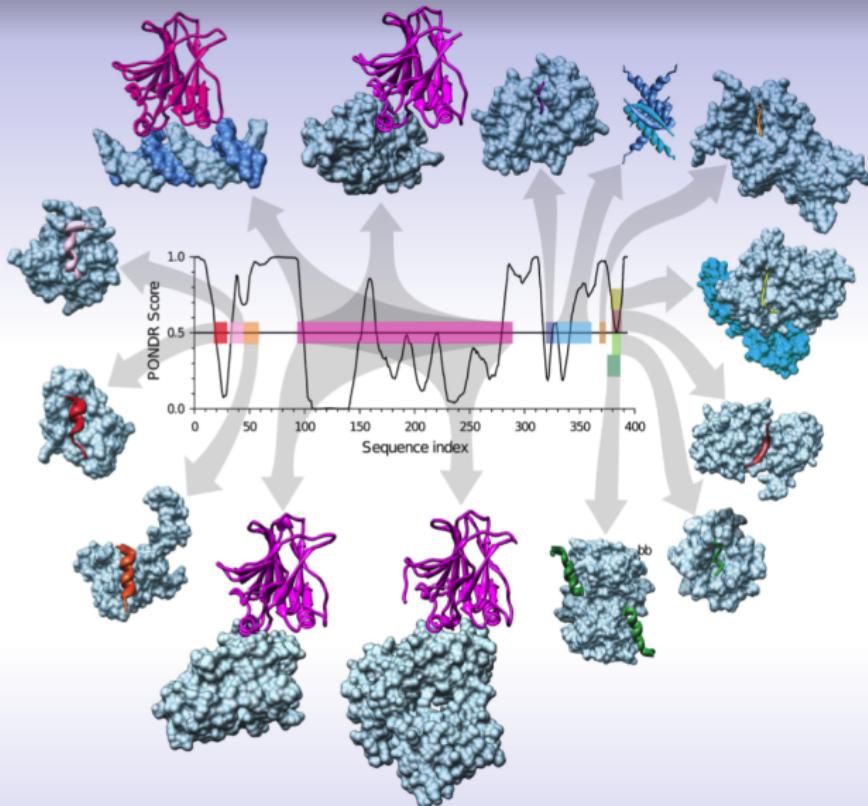
The protein **LEF_BACAN** is a metalloprotease that specifically targets mitogen-activated protein kinase kinases (MKKs). which are important regulators of signal transduction as they phosphorylate and thus activate specific MAPKs (such as ERK1, ERK2, p38 or JNK). *Bacillus anthracis*’ “lethal factor” cleaves its MKK substrates within or close to the MAPK docking sites, thus effectively preventing the MKK to dock to its MAPK.

See <http://elm.eu.org/infos/diseases.html> for a list of motifs implicated in various diseases.

β -CATENIN



The most recurrently mutated experimentally validated SLiM in the COSMIC DB is the conserved proteasomal degradation motif in the highly disordered N-terminal region of β -Catenin which mediates binding to the WD40 repeat domain of the β -TRCP subunit of the SCF-betaTRCP E3 ubiquitin ligase complex. (more than 1700 mutation entries for this motif derived from 1692 unique samples based on 256 different publications)



"Understanding protein non-folding"; UVERSKY & DUNKER; (BIOCHIMICA ET BIOPHYSICA ACTA 2010)

SLiMs play even a role in extracellular proteins as highlighted by these solved structures of three different RGD-containing proteins. (The occurrence of identical motifs on completely unrelated scaffolds supports the hypothesis they represent convergent evolution events).

(Note how these extracellular motifs are located in loops of folded regions.)

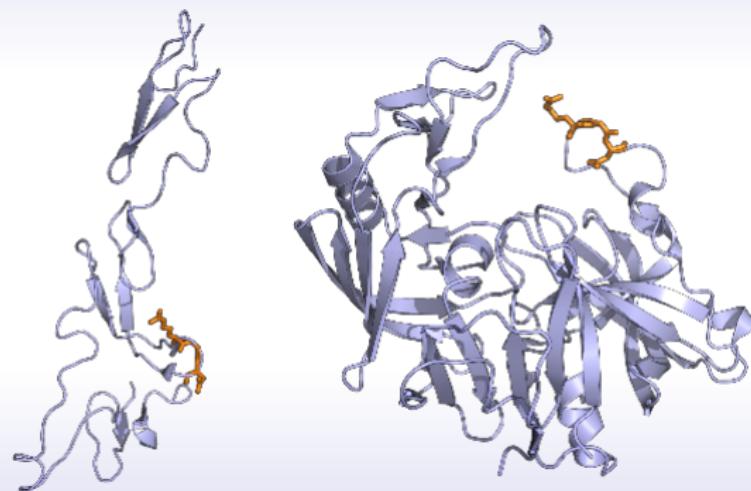


Human Del-1 EGF2 domain

"How pathogens use linear motifs to perturb host cell networks"; VIA ET AL.; (TRENDS BIOCHEM SCI. 2015)

SLiMs play even a role in extracellular proteins as highlighted by these solved structures of three different RGD-containing proteins. (The occurrence of identical motifs on completely unrelated scaffolds supports the hypothesis they represent convergent evolution events).

(Note how these extracellular motifs are located in loops of folded regions.)



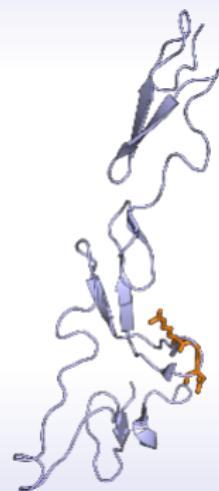
Human Del-1 EGF2 domain

C. albicans secreted aspartic proteinase

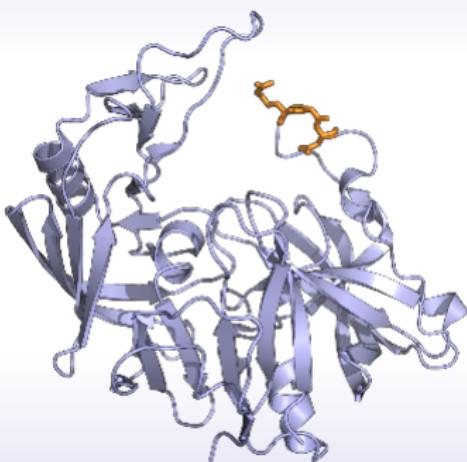
"How pathogens use linear motifs to perturb host cell networks"; VIA ET AL.; (TRENDS BIOCHEM SCI. 2015)

SLiMs play even a role in extracellular proteins as highlighted by these solved structures of three different RGD-containing proteins. (The occurrence of identical motifs on completely unrelated scaffolds supports the hypothesis they represent convergent evolution events).

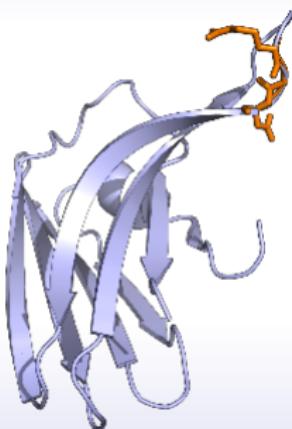
(Note how these extracellular motifs are located in loops of folded regions.)



Human Del-1 EGF2 domain

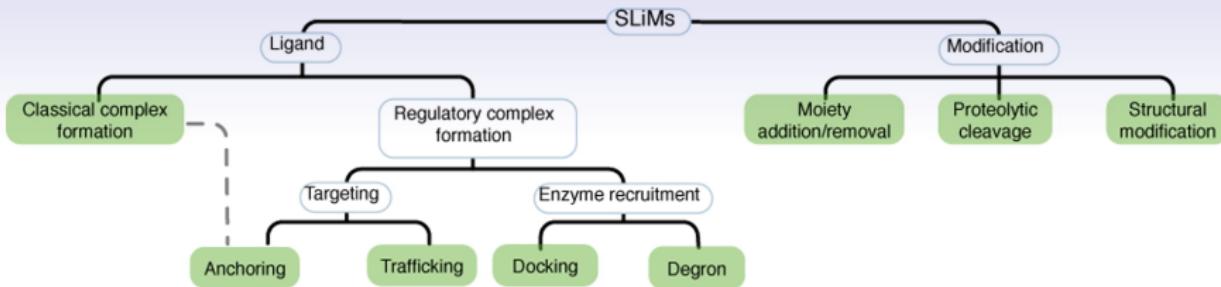


C. albicans secreted aspartic proteinase



P. tritici-repentis toxin Ptr ToxA

"How pathogens use linear motifs to perturb host cell networks"; VIA ET AL.; (TRENDS BIOCHEM SCI. 2015)



DESCRIPTION:

Modification Motifs mediate specific binding to the active site of a modifying enzyme to allow subsequent catalytic post-translational modification of the target site.

EXAMPLE:

NAME MOD_CDK_1

REGEx $xxx([ST])Px[KR]$

Kinase domain

CDK site

DESCRIPTION:

Modification Motifs mediate specific binding to the active site of a modifying enzyme to allow subsequent catalytic post-translational modification of the target site.

EXAMPLE:

NAME MOD_CDK_1

REGEx $xxx([ST])Px[KR]$



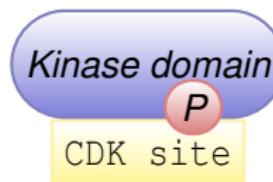
DESCRIPTION:

Modification Motifs mediate specific binding to the active site of a modifying enzyme to allow subsequent catalytic post-translational modification of the target site.

EXAMPLE:

NAME MOD_CDK_1

REGEx $xxx([ST])Px[KR]$



DESCRIPTION:

Modification Motifs mediate specific binding to the active site of a modifying enzyme to allow subsequent catalytic post-translational modification of the target site.

EXAMPLE:

NAME MOD_CDK_1

REGEx $xxx([ST])Px[KR]$

Kinase domain



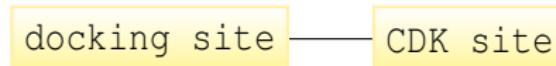
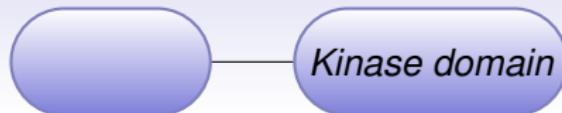
DESCRIPTION:

Docking motifs recruit enzymes via a surface that is distinct from the active site.

EXAMPLE:

NAME DOC_CYCLIN_1

REGEX $[RK]xLx\{0,1\}[LFY]$



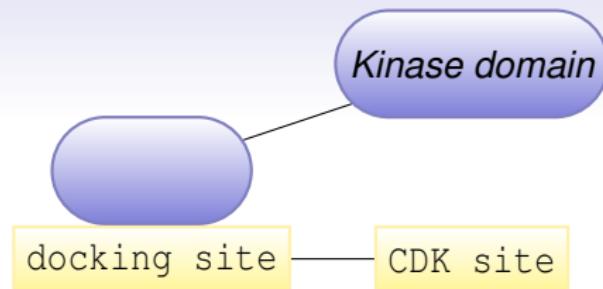
DESCRIPTION:

Docking motifs recruit enzymes via a surface that is distinct from the active site.

EXAMPLE:

NAME DOC_CYCLIN_1

REGEX $[RK]xLx\{0,1\}[LFY]$



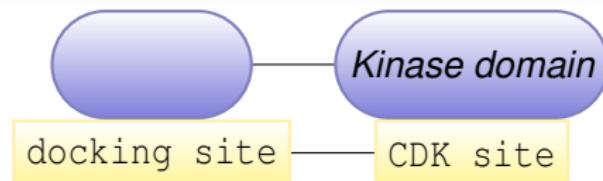
DESCRIPTION:

Docking motifs recruit enzymes via a surface that is distinct from the active site.

EXAMPLE:

NAME DOC_CYCLIN_1

REGEx $[RK]xLx\{0,1\}[LFY]$



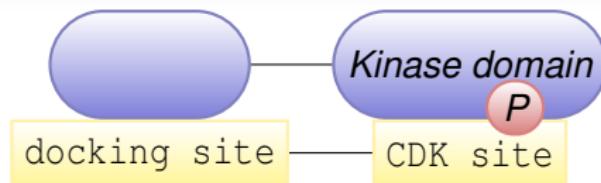
DESCRIPTION:

Docking motifs recruit enzymes via a surface that is distinct from the active site.

EXAMPLE:

NAME DOC_CYCLIN_1

REGEx $[RK]xLx\{0,1\}[LFY]$



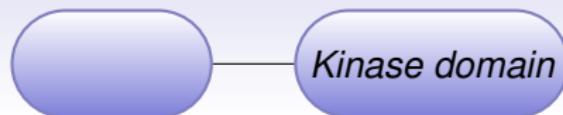
DESCRIPTION:

Docking motifs recruit enzymes via a surface that is distinct from the active site.

EXAMPLE:

NAME DOC_CYCLIN_1

REGEX $[RK]xLx\{0,1\}[LFY]$



DESCRIPTION:

Proteolytic processing of proteins into smaller polypeptides by protease-catalyzed hydrolysis of specific peptide bonds

EXAMPLE:

NAME CLV_Separin_Metazoa

REGEX $E[IMPVL][MLVP]Rx$

Separase

Cleavage site

DESCRIPTION:

Proteolytic processing of proteins into smaller polypeptides by protease-catalyzed hydrolysis of specific peptide bonds

EXAMPLE:

NAME CLV_Separin_Metazoa

REGEX $E[IMPVL][MLVP]Rx$



DESCRIPTION:

Proteolytic processing of proteins into smaller polypeptides by protease-catalyzed hydrolysis of specific peptide bonds

EXAMPLE:

NAME CLV_Separin_Metazoa

REGEx $E[IMPVL][MLVP]Rx$



DESCRIPTION:

Proteolytic processing of proteins into smaller polypeptides by protease-catalyzed hydrolysis of specific peptide bonds

EXAMPLE:

NAME CLV_Separin_Metazoa

REGEx $E[IMPVL][MLVP]Rx$

Separase



DESCRIPTION:

Degradation motifs (Degrons) recognized by E3 Ubiquitin Ligase complexes priming proteins for degradation, regulating protein half-life.

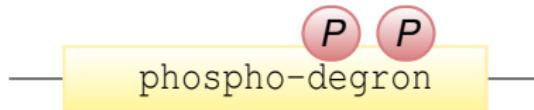
EXAMPLE:

NAME DEG_SCF_TRCP1_1

REGEx $D(S)GXX([ST])$

FBW7

SCF E3 Ligase



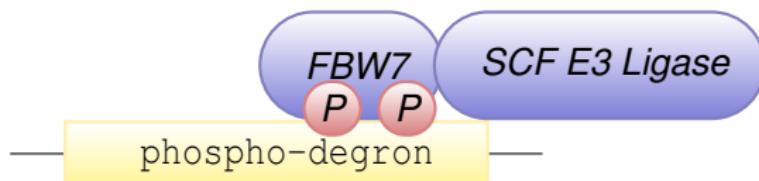
DESCRIPTION:

Degradation motifs (Degrons) recognized by E3 Ubiquitin Ligase complexes priming proteins for degradation, regulating protein half-life.

EXAMPLE:

NAME DEG_SCF_TRCP1_1

REGEx $D(S)GXX([ST])$



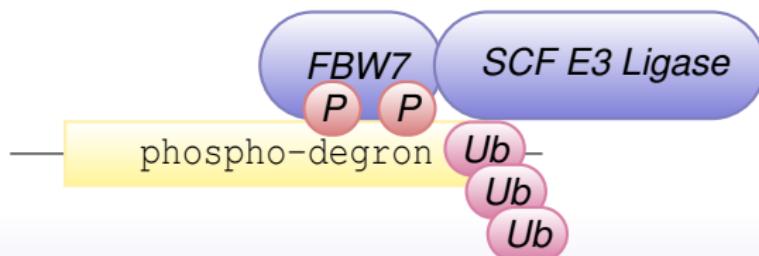
DESCRIPTION:

Degradation motifs (Degrons) recognized by E3 Ubiquitin Ligase complexes priming proteins for degradation, regulating protein half-life.

EXAMPLE:

NAME DEG_SCF_TRCP1_1

REGEx $D(S)Gxx([ST])$



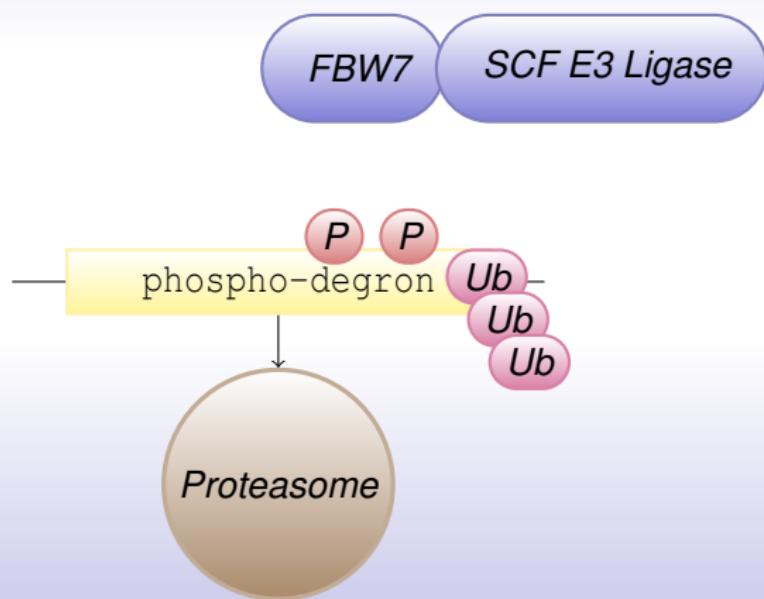
DESCRIPTION:

Degradation motifs (Degrons) recognized by E3 Ubiquitin Ligase complexes priming proteins for degradation, regulating protein half-life.

EXAMPLE:

NAME DEG_SCF_TRCP1_1

REGEx $D(S)Gxx([ST])$



DESCRIPTION:

TARGETING motifs allow a protein to bind to the transport machinery that relocalizes it to a particular sub-cellular location.

ANCHORING motifs are recognized by biomolecules specific to a sub-cellular location and thereby retain the motif-containing protein at that location.

EXAMPLE:

NAME TRG_NLS_MonoCore_2

REGEx $[\wedge DE](K[RK]|RK)[KRP][KR][\wedge DE]$

Importin α

NLS

DESCRIPTION:

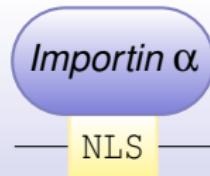
TARGETING motifs allow a protein to bind to the transport machinery that relocalizes it to a particular sub-cellular location.

ANCHORING motifs are recognized by biomolecules specific to a sub-cellular location and thereby retain the motif-containing protein at that location.

EXAMPLE:

NAME TRG_NLS_MonoCore_2

REGEx $[\wedge DE](K[RK]|RK)[KRP][KR][\wedge DE]$



DESCRIPTION:

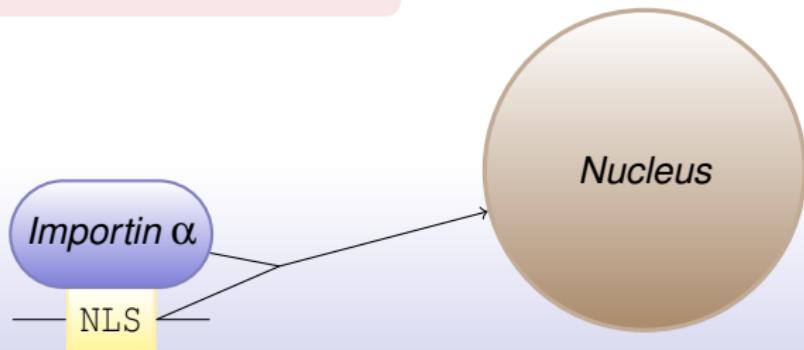
TARGETING motifs allow a protein to bind to the transport machinery that relocalizes it to a particular sub-cellular location.

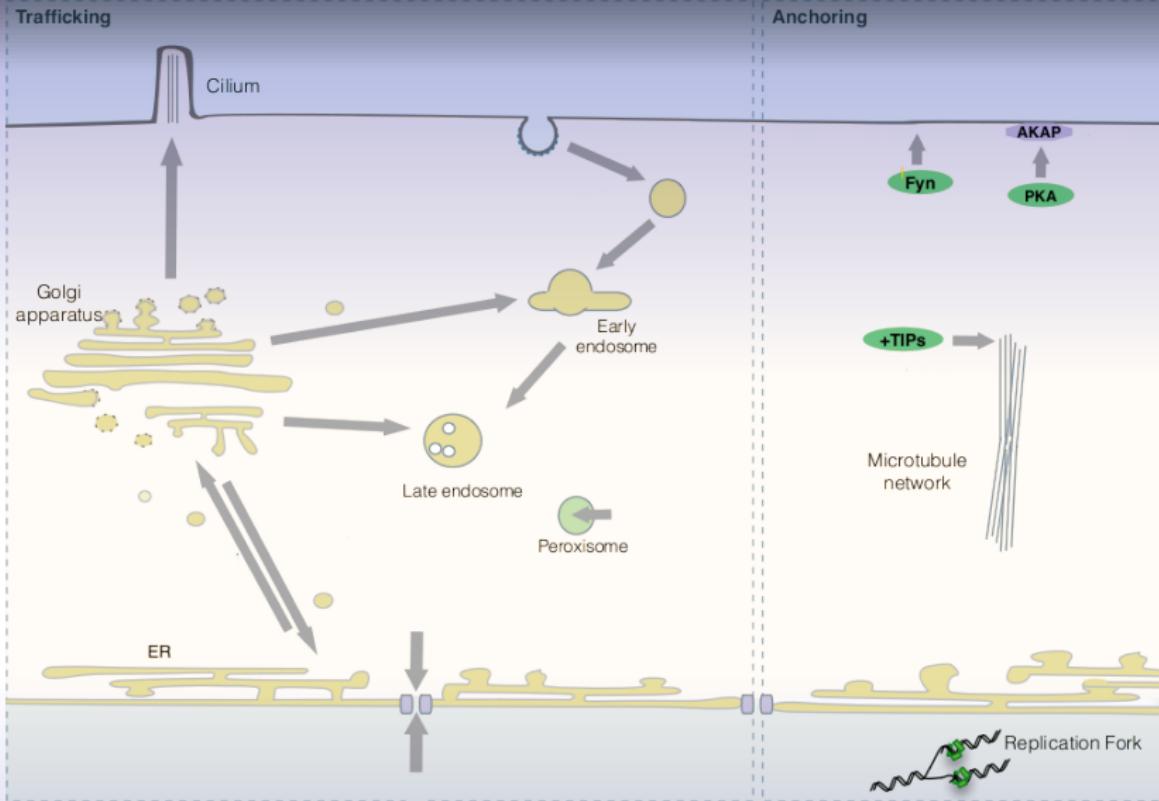
ANCHORING motifs are recognized by biomolecules specific to a sub-cellular location and thereby retain the motif-containing protein at that location.

EXAMPLE:

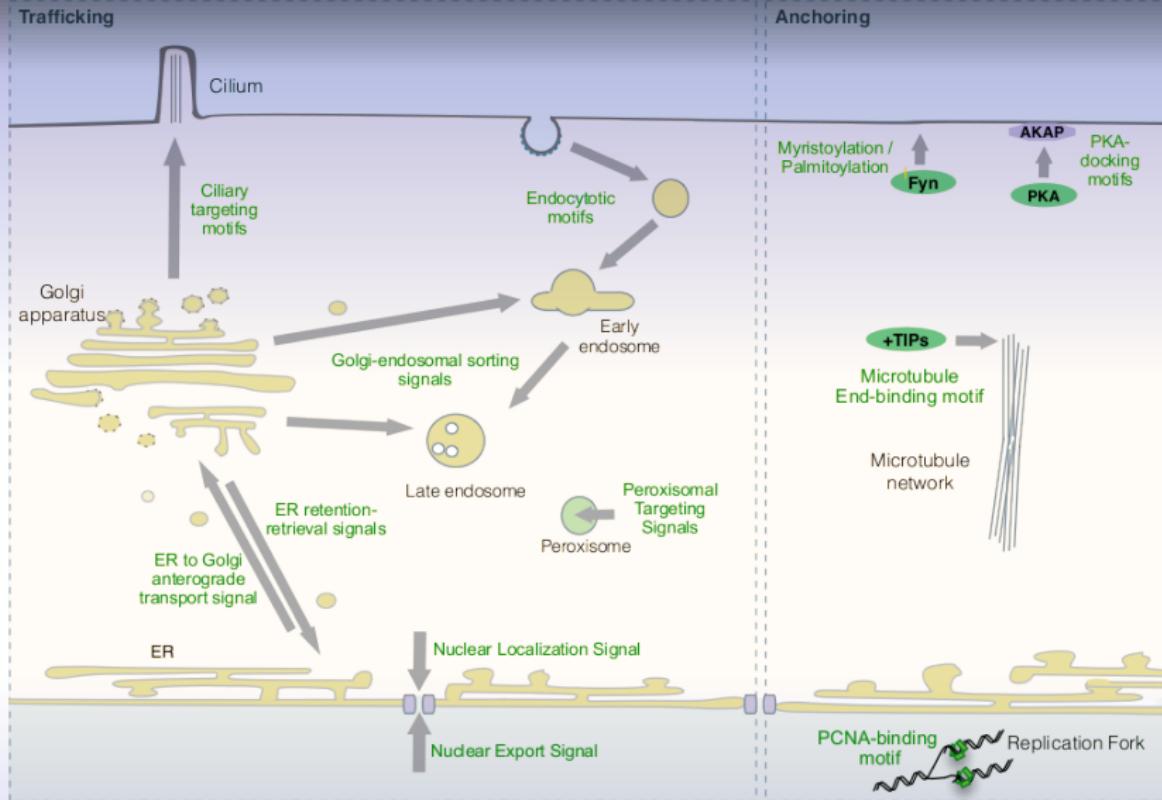
NAME TRG_NLS_MonoCore_2

REGEx $[\wedge DE](K[RK]|RK)[KRP][KR][\wedge DE]$





"Short linear motifs: Ubiquitous and functionally diverse protein interaction modules directing cell regulation"; VAN ROEY, UYAR, WEATHERITT, DINKEL, SEILER, BUDD, GIBSON & DAVEY; (CHEM. REVIEWS; 2014)



"Short linear motifs: Ubiquitous and functionally diverse protein interaction modules directing cell regulation"; VAN ROEY, UYAR, WEATHERITT, DINKEL, SEILER, BUDD, GIBSON & DAVEY; (CHEM. REVIEWS; 2014)



PI3K-Akt signaling pathway

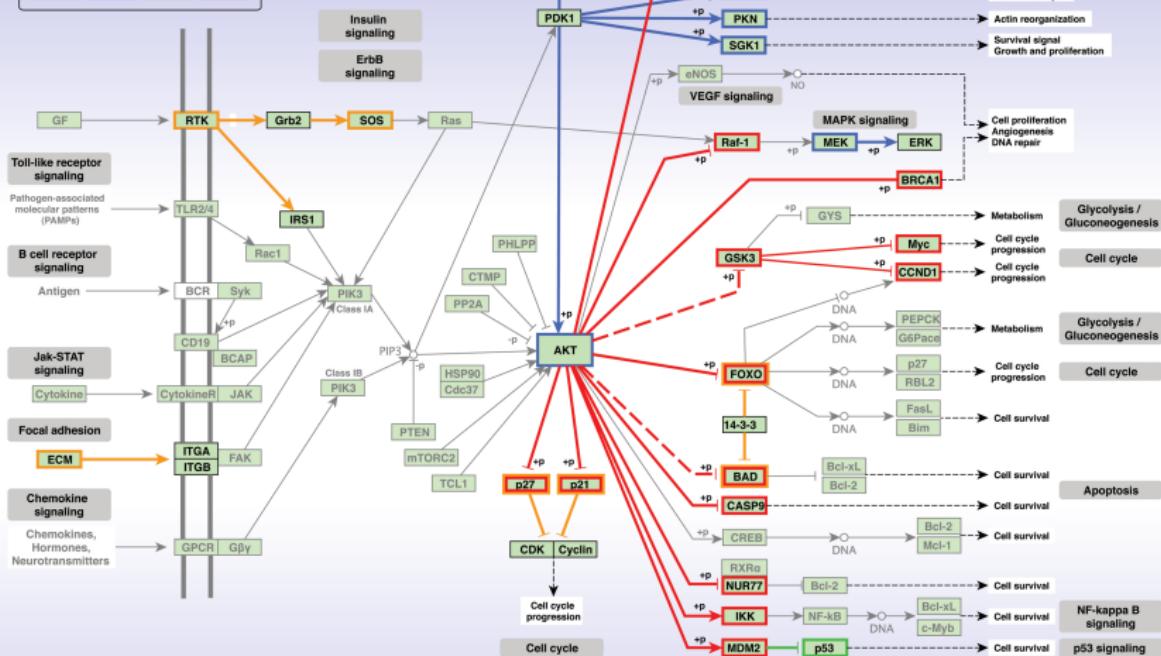
ELM types highlighted in color:

DEG

DOC

LIG

MOD



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

ELM is a **REPOSITORY** of more than 240 thoroughly annotated motif classes with over 2700 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.

SHORT LINEAR MOTIFS

- small, versatile modules which mediate transient interactions
- important regulators of cellular processes.
- “kidnapped” by viruses
- play an important role in diseases
- collected in the Eukaryotic Linear Motif Resource (ELM)



I mustache you a
Question

BUT I'M SHAVING IT

for later.