

## SHORT LINEAR MOTIFS

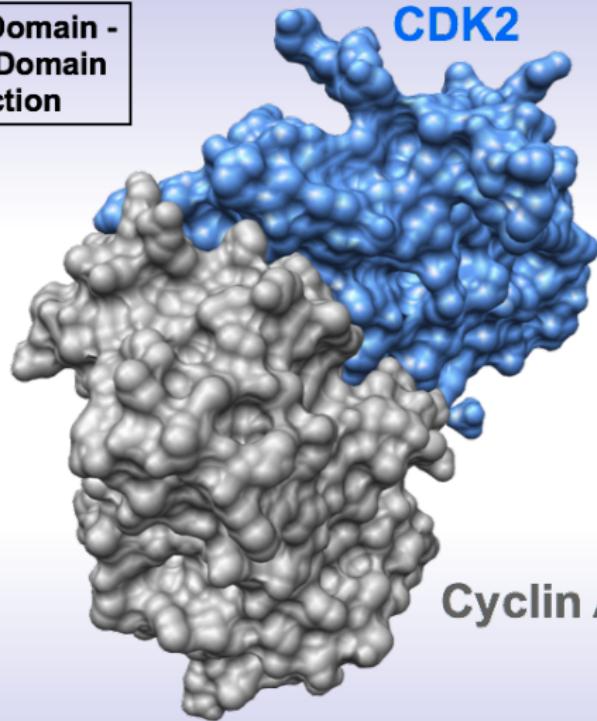
---

Holger Dinkel

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

# IMPORTANCE OF SHORT LINEAR MOTIFS

Globular Domain -  
Globular Domain  
Interaction



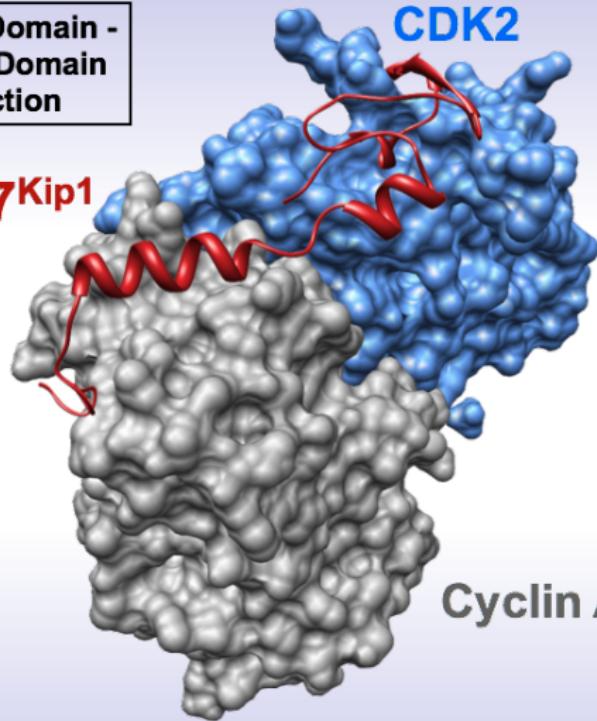
Cyclin A



# IMPORTANCE OF SHORT LINEAR MOTIFS

Globular Domain -  
Disordered Domain  
Interaction

Globular Domain -  
Globular Domain  
Interaction



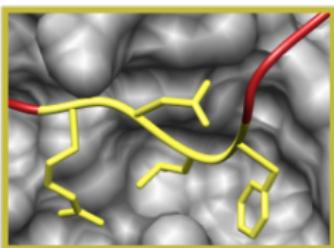
Cyclin A

# IMPORTANCE OF SHORT LINEAR MOTIFS

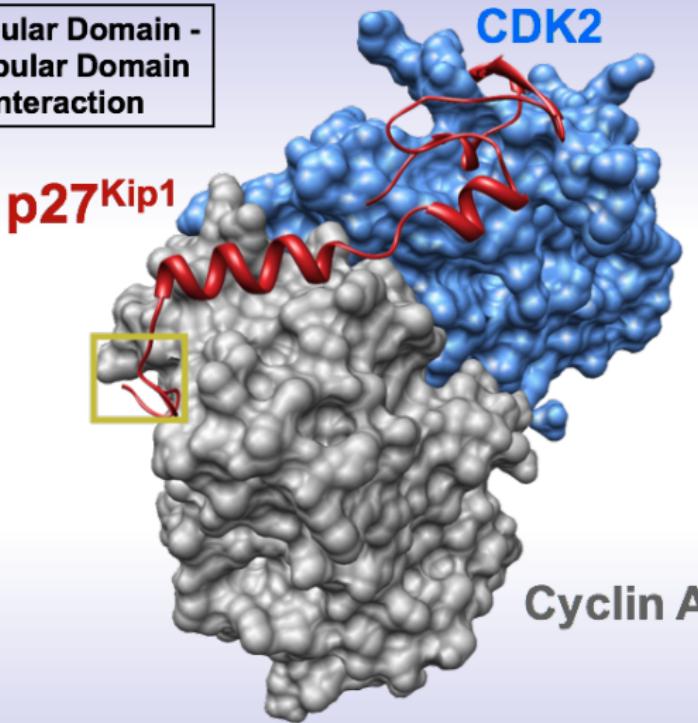
Globular Domain -  
Globular Domain  
Interaction

Globular Domain -  
Disordered Domain  
Interaction

Globular Domain -  
Short Linear Motif  
Interaction



RNLF



Cyclin A

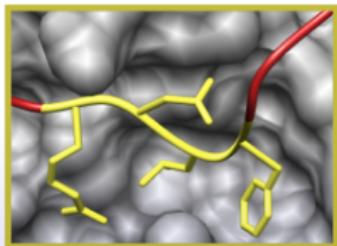


# IMPORTANCE OF SHORT LINEAR MOTIFS

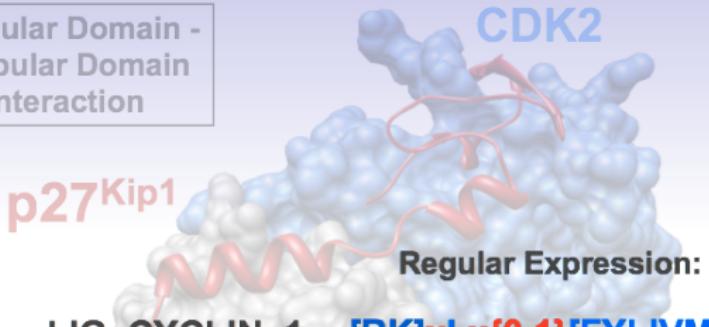
Globular Domain -  
Disordered Domain  
Interaction

Globular Domain -  
Globular Domain  
Interaction

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



RNLF



Regular Expression:

LIG\_CYCLIN\_1    **[RK]xLx{0,1}[FYLIVMP]**

Defined positions

Fixed positions

Degenerate positions

Undefined positions

Fixed-length wildcard

Flexible-length wildcard {min,max}

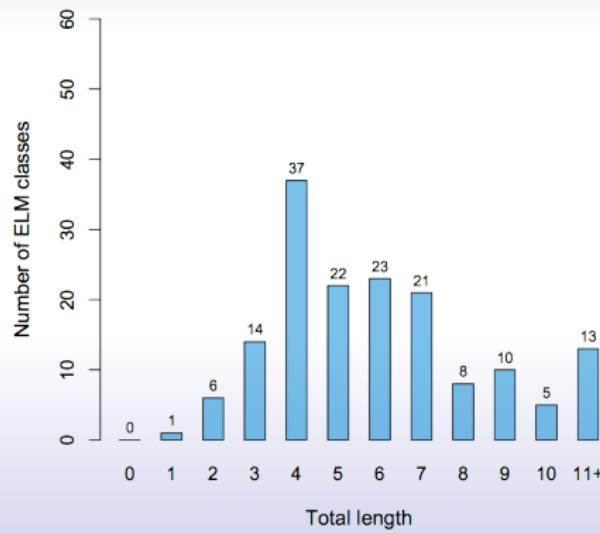
Cyclin A



# ATTRIBUTES OF SHORT LINEAR MOTIFS

## LINEAR MOTIFS

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

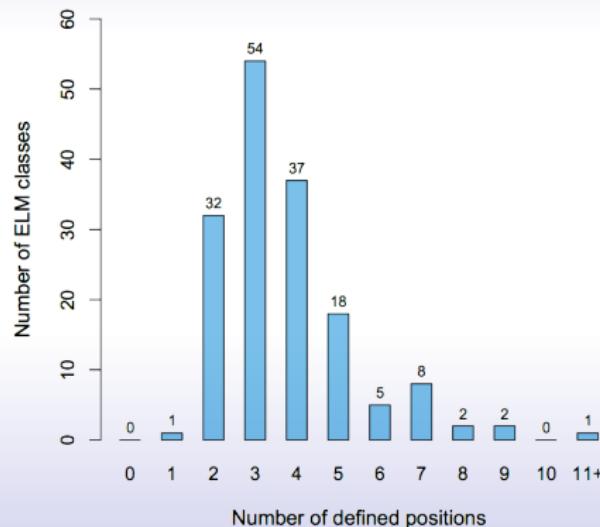


*"Attributes of short linear motifs"*; DAVEY, VAN ROEY, WEATHERITT, TOEDT, UYAR, ALtenBERG, BUDD, DIELLA, DINKEL & GIBSON; (MOL BIOSYST. 2011)

# ATTRIBUTES OF SHORT LINEAR MOTIFS

## LINEAR MOTIFS

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

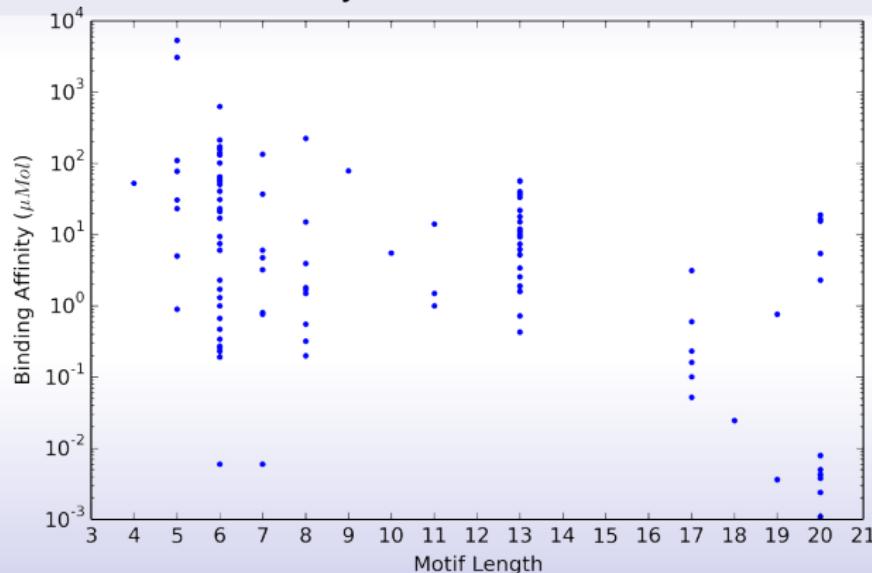


*"Attributes of short linear motifs"*; DAVEY, VAN ROEY, WEATHERITT, TOEDT, UYAR, ALtenBERG, BUDD, DIELLA, DINKEL & GIBSON; (MOL BIOSYST. 2011)

# ATTRIBUTES OF SHORT LINEAR MOTIFS

## LINEAR MOTIFS

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



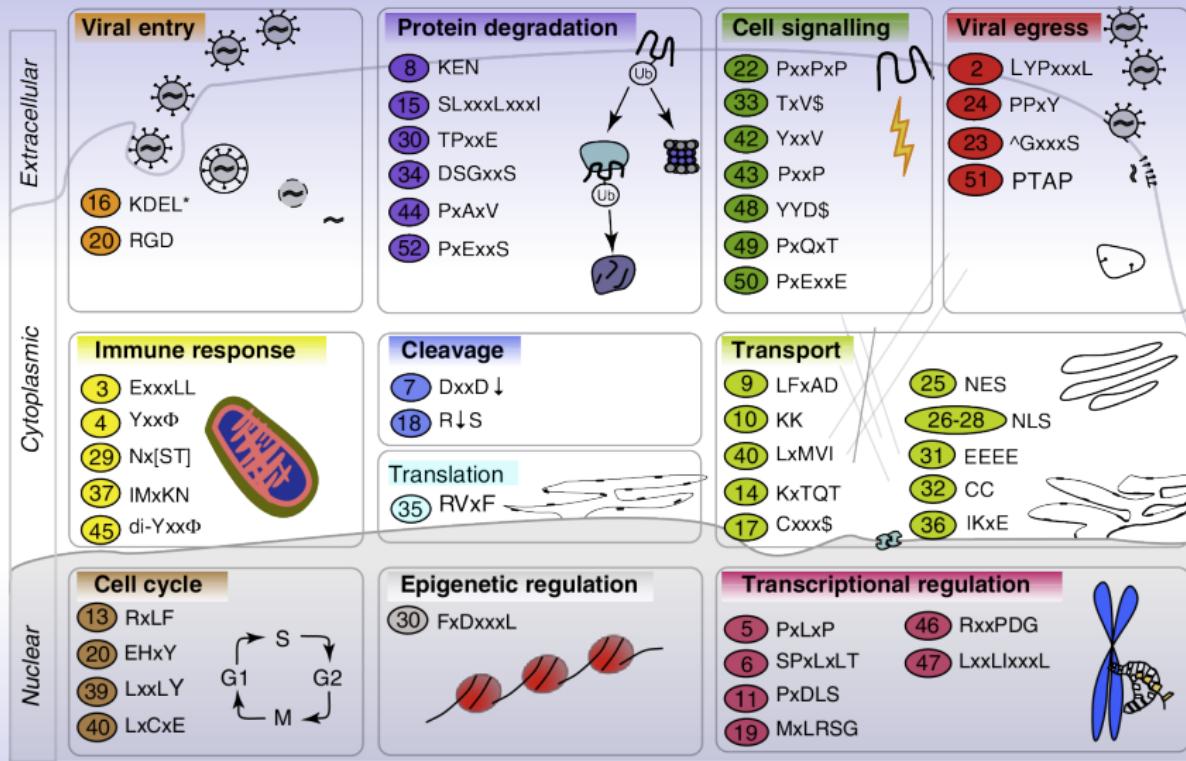
# PREVALENCE OF SHORT LINEAR MOTIFS

## 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	$_P Yxx[IV]$
WW	151 / 103	$PPxY$
PTB	142 / 133	$NPx_pY$



# IMPORTANCE OF SHORT LINEAR MOTIFS: VIRUSES



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

# IMPORTANCE OF SHORT LINEAR MOTIFS: DISEASES

See also <http://elm.eu.org/infos/diseases.html>

## LIDDLE'S-SYNDROME: WW-INTERACTION MOTIF

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



# IMPORTANCE OF SHORT LINEAR MOTIFS: DISEASES

See also <http://elm.eu.org/infos/diseases.html>

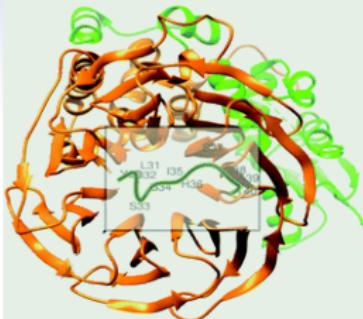
## BACILLUS ANTHRACIS “LETHAL FACTOR”

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.



# IMPORTANCE OF SHORT LINEAR MOTIFS: CANCER

## $\beta$ -CATENIN

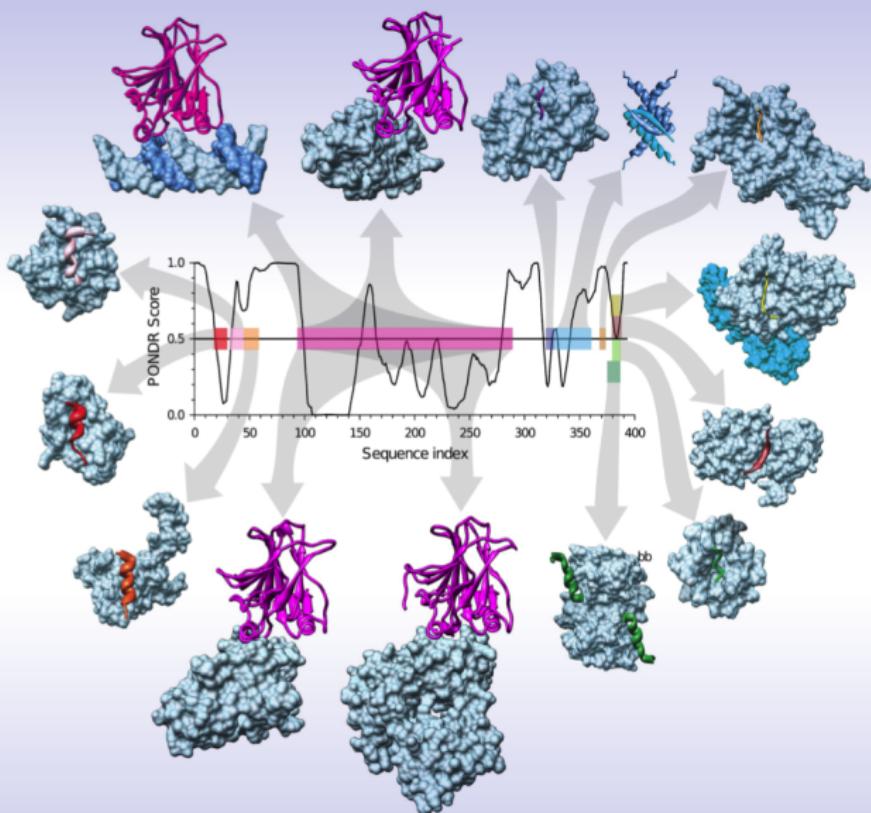


The most recurrently mutated experimentally validated SLiM in the COSMIC DB is the conserved proteasomal degradation motif in the highly disordered N-terminal region of  $\beta$ -Catenin which mediates binding to the WD40 repeat domain of the  $\beta$ -TRCP subunit of the SCF-betaTRCP E3 ubiquitin ligase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



*"Proteome-wide analysis of human disease mutations in short linear motifs: neglected players in cancer?"*; UYAR, WEATHERITT, DINKEL, DAVEY & GIBSON; (MOL. BIOSYST.; 2014)

# IMPORTANCE OF SHORT LINEAR MOTIFS: P53

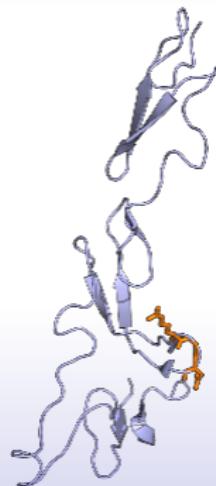


"Understanding protein non-folding"; UVERSKY & DUNKER; (BIOCHIMICA ET BIOPHYSICA ACTA 2010)

# MOTIFS IN EXTERNAL PROTEINS

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

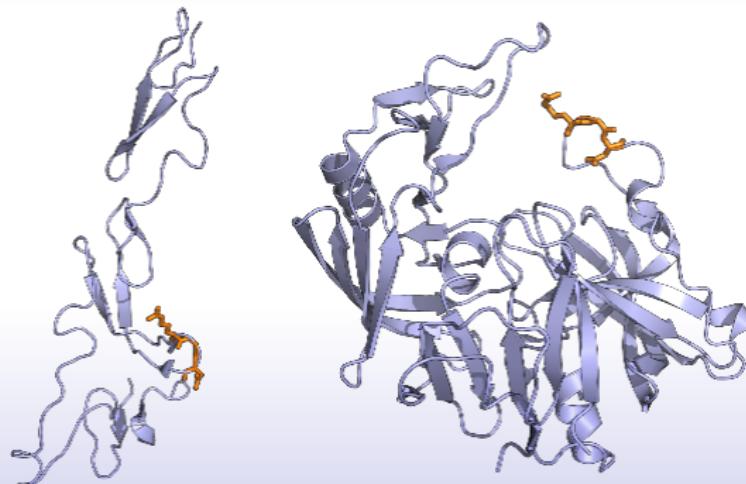
"How pathogens use linear motifs to perturb host cell networks"; VIA ET AL.; (TRENDS BIOCHEM SCI. 2015)



# MOTIFS IN EXTERNAL PROTEINS

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

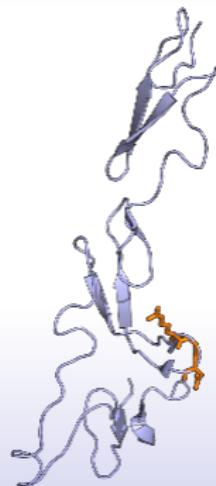
"How pathogens use linear motifs to perturb host cell networks"; VIA ET AL.; (TRENDS BIOCHEM SCI. 2015)



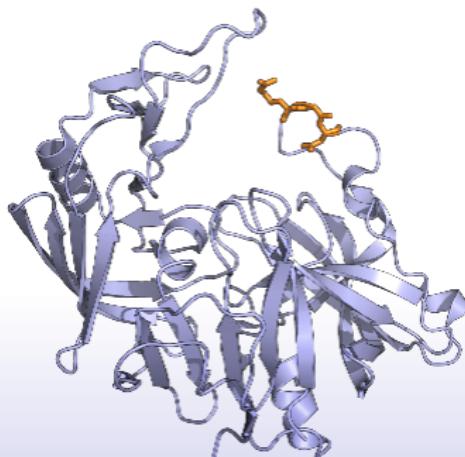
# MOTIFS IN EXTERNAL PROTEINS

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).

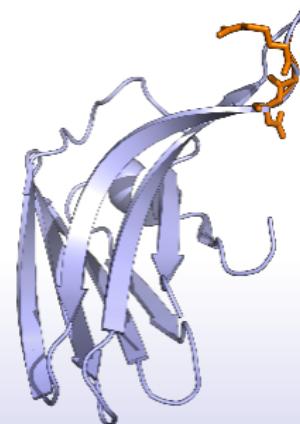
(Note how these extracellular motifs are located in loops of folded regions.)



Human Del-1 EGF2 domain



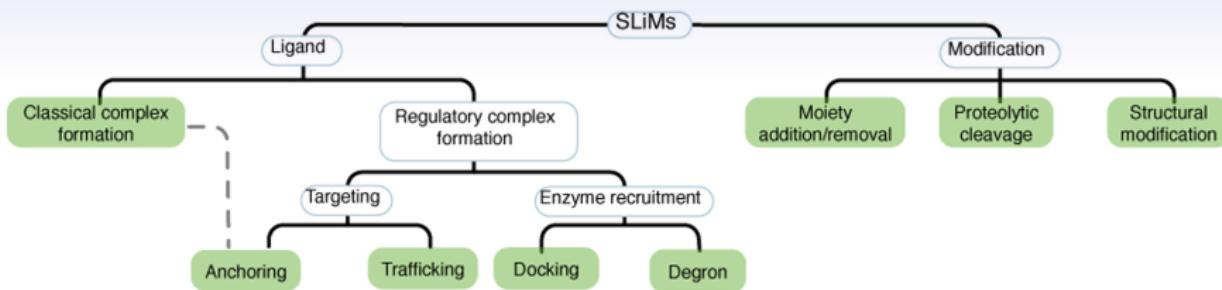
*C. albicans* secreted aspartic proteinase



*P. tritici-repentis* toxin Ptr ToxA

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

# CLASSIFICATION OF MOTIFS



# MOTIF CLASSES: MODIFICATION SITES

## 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



# MOTIF CLASSES: MODIFICATION SITES

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



# MOTIF CLASSES: MODIFICATION SITES

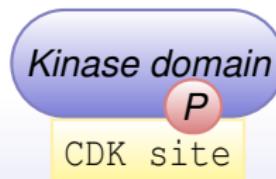
## DESCRIPTION:

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



# MOTIF CLASSES: MODIFICATION SITES

## 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*



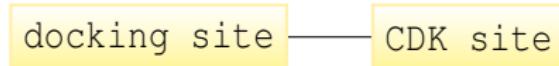
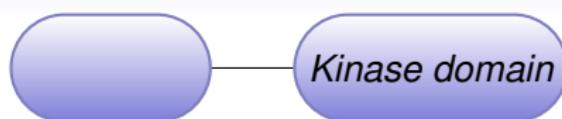
# MOTIF CLASSES: DOCKING MOTIFS

## 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]



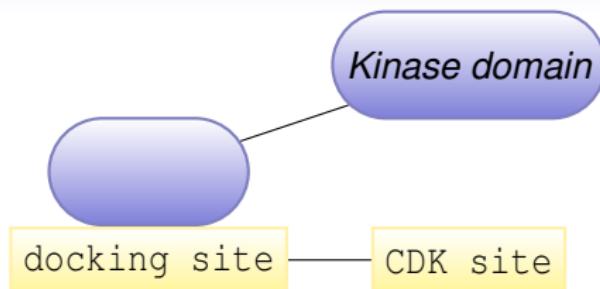
# MOTIF CLASSES: DOCKING MOTIFS

## 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]



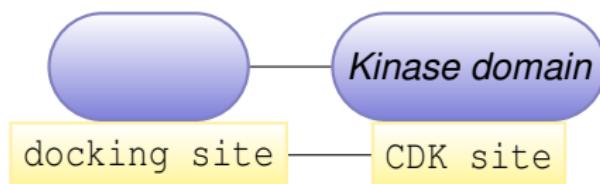
# MOTIF CLASSES: DOCKING MOTIFS

## 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]



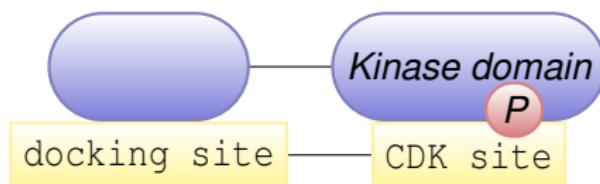
# MOTIF CLASSES: DOCKING MOTIFS

## 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]



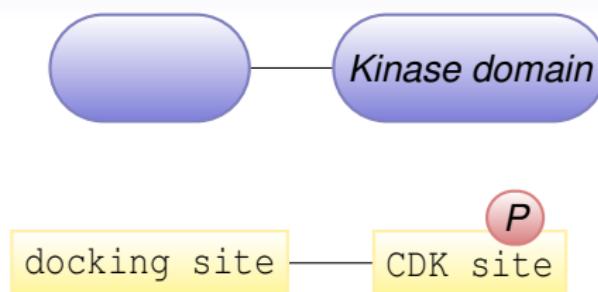
# MOTIF CLASSES: DOCKING MOTIFS

## 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]



# MOTIF CLASSES: CLEAVAGE MOTIFS

## 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 —



# MOTIF CLASSES: CLEAVAGE MOTIFS

## DESCRIPTION:

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

## EXAMPLE:

NAME CLV\_Separin\_Metazoa

REGEx  $E[IMPVL][MLVP]Rx$



# MOTIF CLASSES: CLEAVAGE MOTIFS

## DESCRIPTION:

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

## EXAMPLE:

NAME CLV\_Separin\_Metazoa

REGEx  $E[IMPVL][MLVP]Rx$



# MOTIF CLASSES: CLEAVAGE MOTIFS

## 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



# MOTIF CLASSES: DEGRADATION MOTIFS

## 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*



# MOTIF CLASSES: DEGRADATION MOTIFS

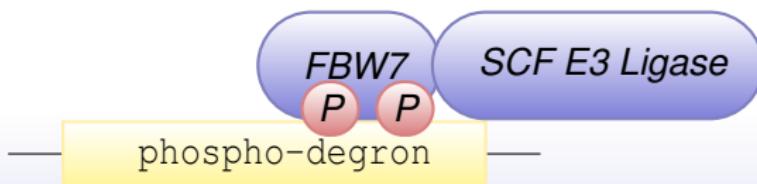
## DESCRIPTION:

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

## EXAMPLE:

NAME DEG\_SCF\_TRCP1\_1

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



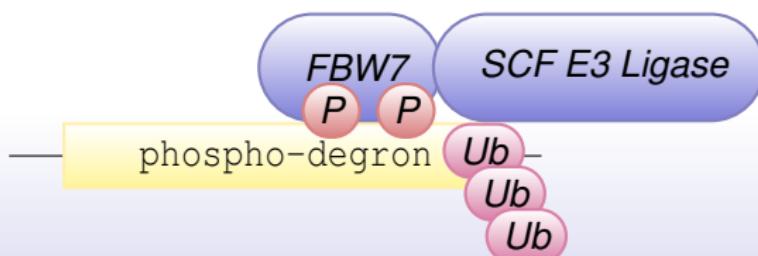
# MOTIF CLASSES: DEGRADATION MOTIFS

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



# MOTIF CLASSES: DEGRADATION MOTIFS

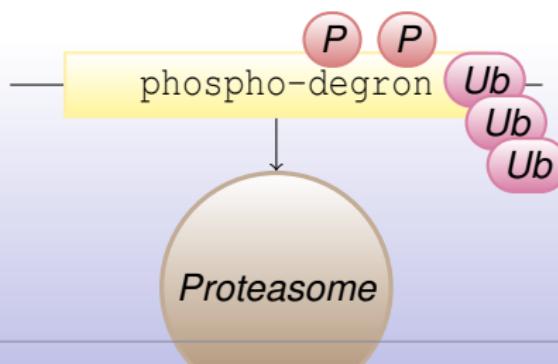
## DESCRIPTION:

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

## EXAMPLE:

NAME DEG\_SCF\_TRCP1\_1

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



# MOTIF CLASSES: TARGETING/ANCHORING MOTIFS

## 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** [^DE](K[RK]|RK)[KRP][KR][^DE]

*Importin  $\alpha$*

NLS



# MOTIF CLASSES: TARGETING/ANCHORING MOTIFS

## 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** [^DE](K[RK]|RK)[KRP][KR][^DE]



# MOTIF CLASSES: TARGETING/ANCHORING MOTIFS

## DESCRIPTION:

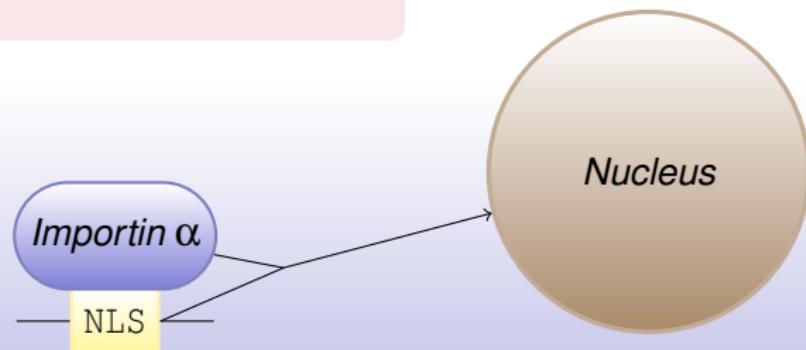
**TARGETING** motifs allow a protein to bind to the transport machinery that relocalizes it to a particular sub-cellular location.

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

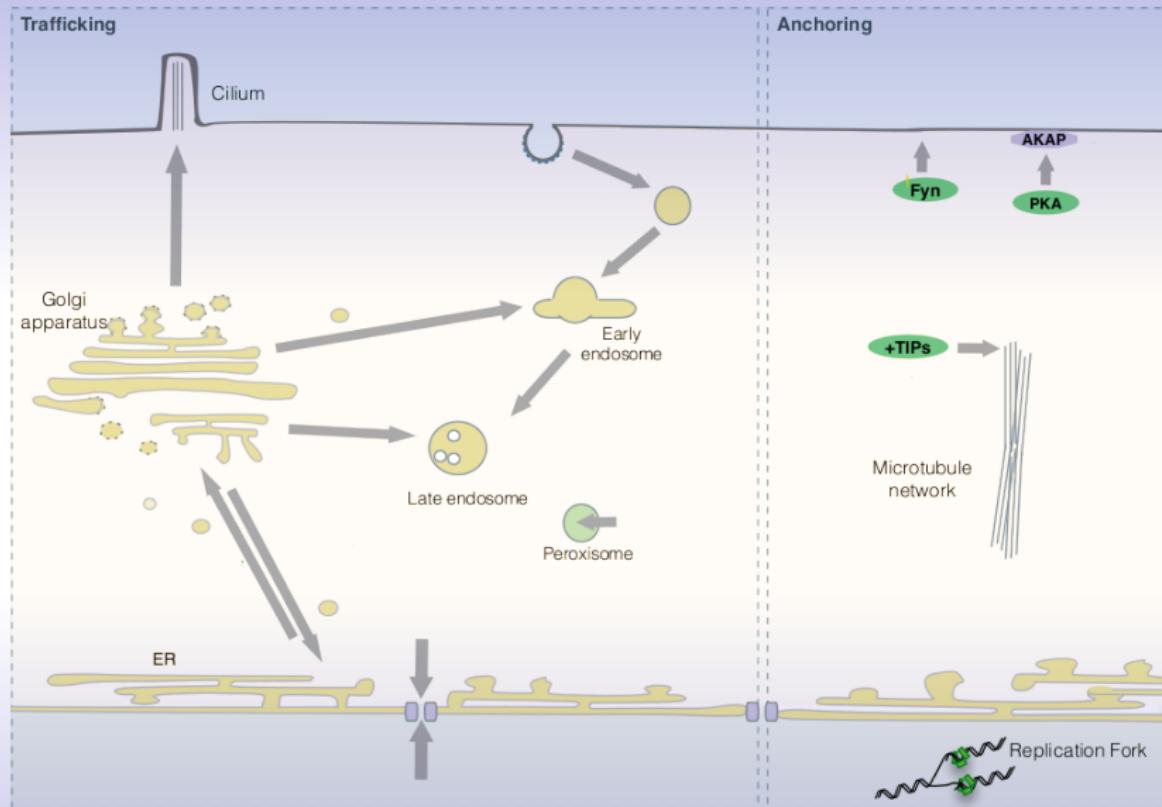
## EXAMPLE:

NAME TRG\_NLS\_MonoCore\_2

REGEx  $[\text{^DE}](K[RK]|RK)[KRP][KR][\text{^DE}]$



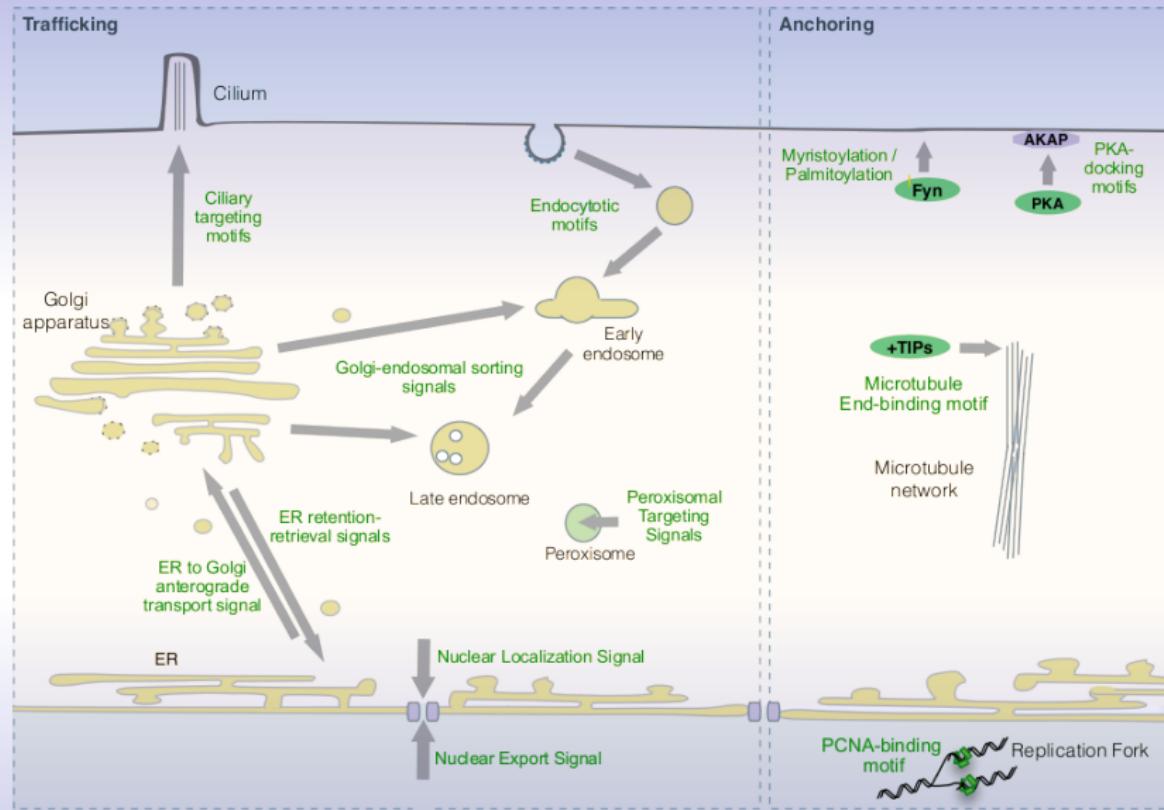
# MOTIF CLASSES: TARGETING/ANCHORING MOTIFS



*"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

# MOTIF CLASSES: TARGETING/ANCHORING MOTIFS



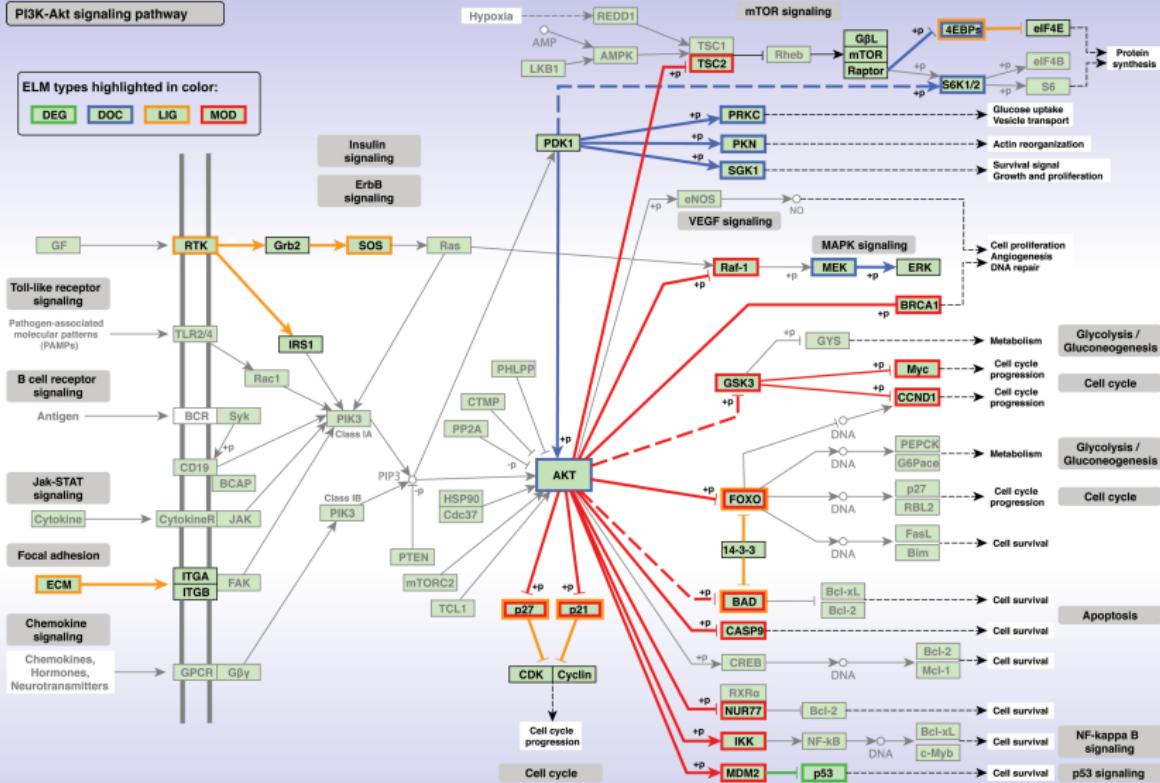
"Short linear motifs: Ubiquitous and functionally diverse protein interaction modules directing cell regulation"; VAN ROEY, UYAR, WEATHERITT, DINKEL, SEILER, BUDD, GIBSON & DAVEY; (CHEM. REVIEWS; 2014)

SHORT LINEAR MOTIFS

# VISUALIZING MOTIF-MEDIATED INTERACTIONS



# VISUALIZING MOTIF-MEDIATED INTERACTIONS



# OUTLOOK: THE ELM RESOURCE

## 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.



# SUMMARY

## 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)



# QUESTIONS?

