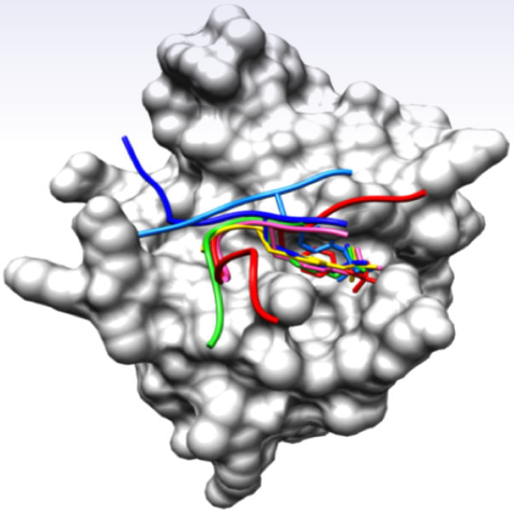


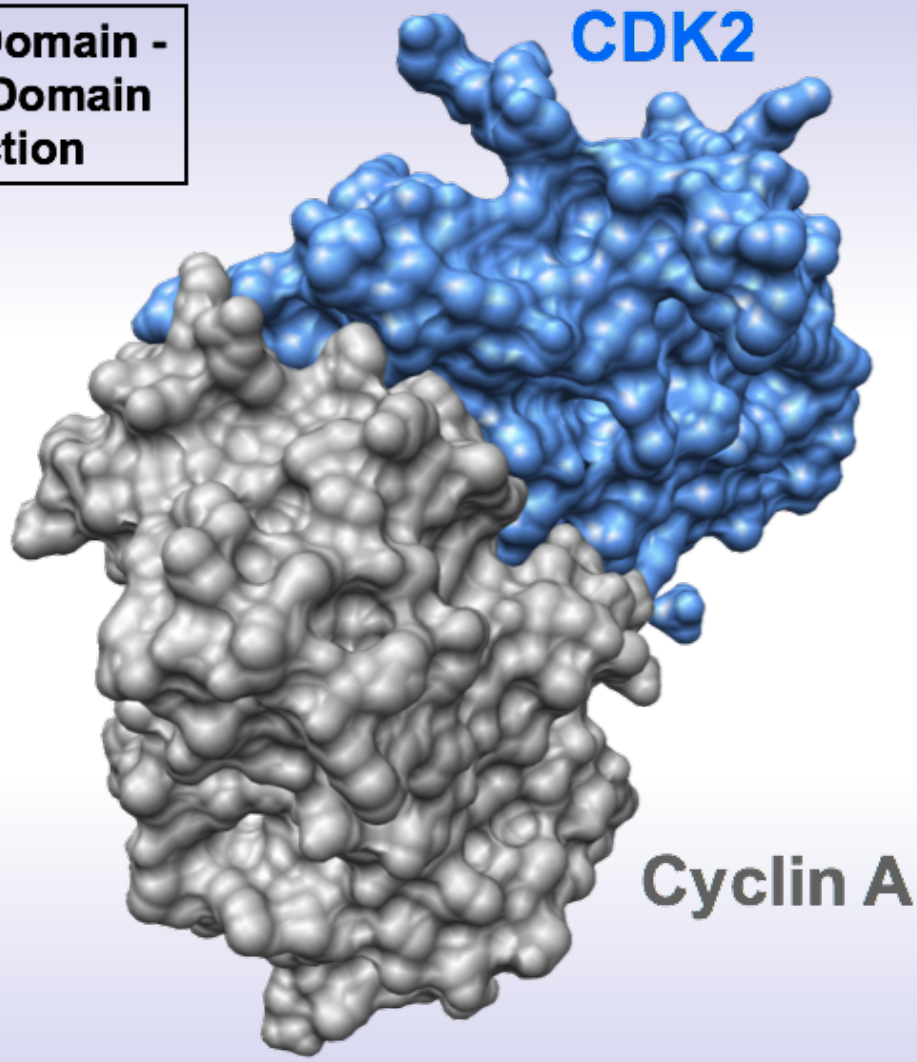
Short Linear Motifs

Holger Dinkel

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

Importance of Short Linear Motifs

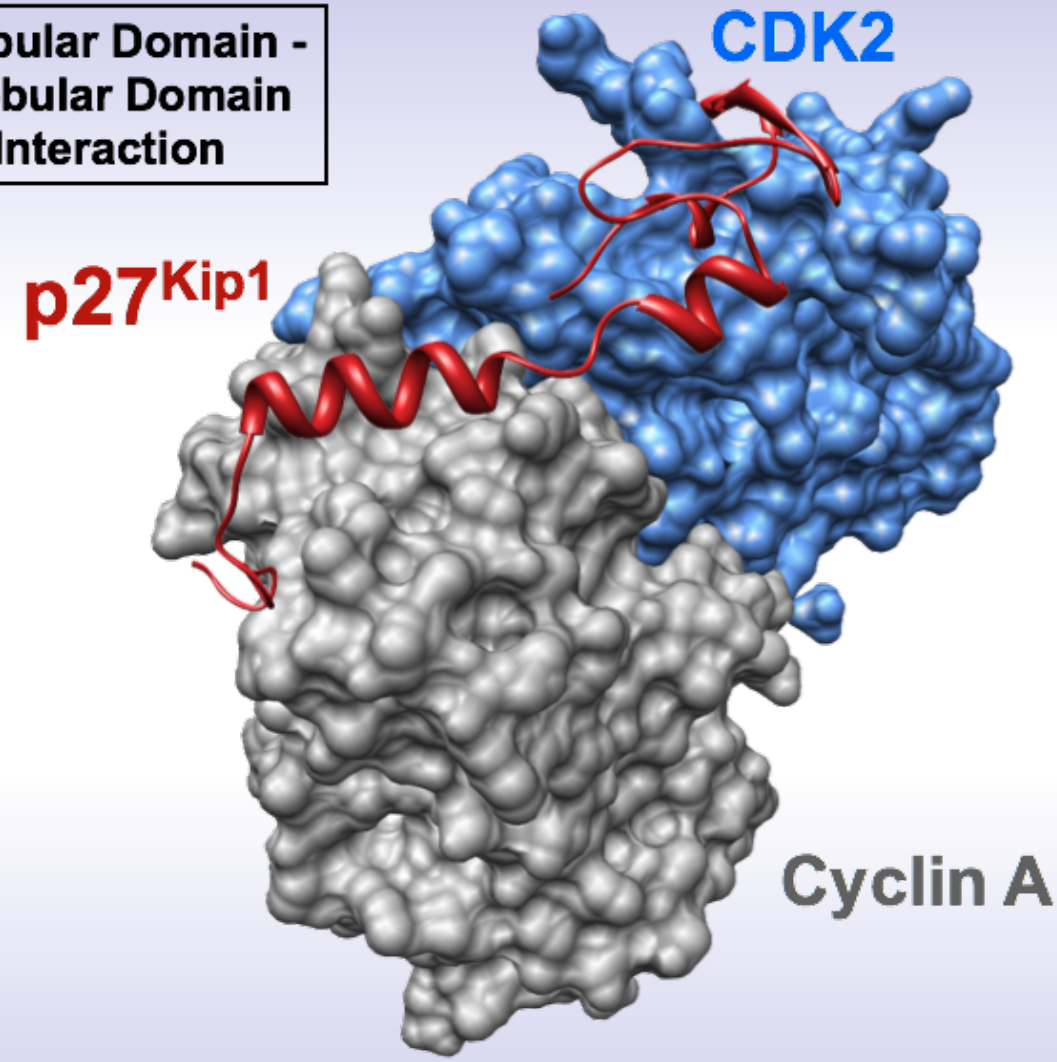
Globular Domain - Globular Domain Interaction



Importance of Short Linear Motifs

Globular Domain - Globular Domain Interaction

Globular Domain - Disordered Domain Interaction

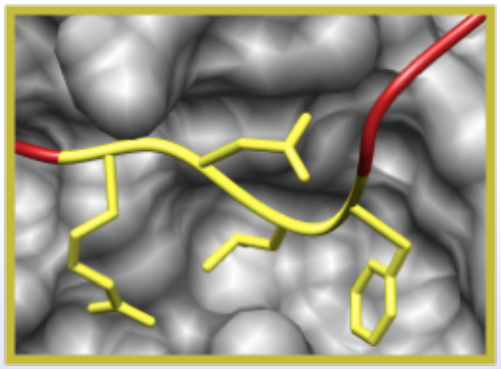


Importance of Short Linear Motifs

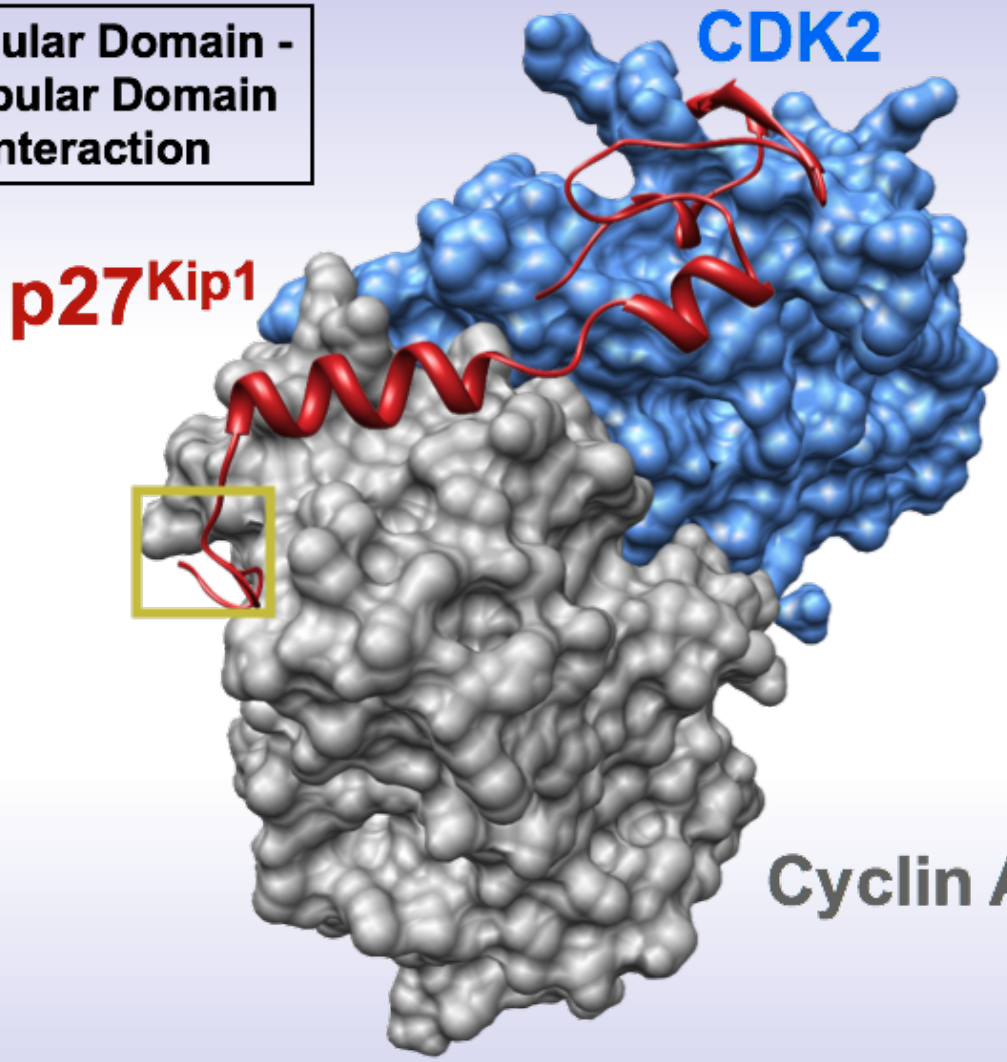
Globular Domain - Globular Domain Interaction

Globular Domain - Disordered Domain Interaction

Globular Domain - Short Linear Motif Interaction



RNLF



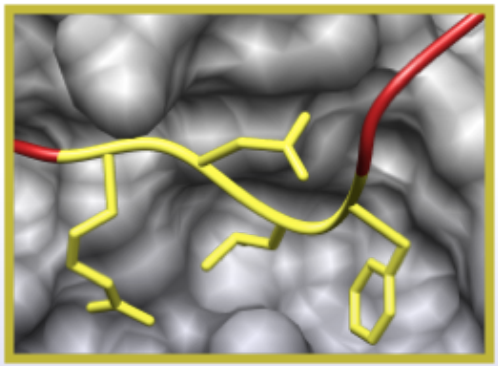
Importance of Short Linear Motifs

Globular Domain - Globular Domain Interaction

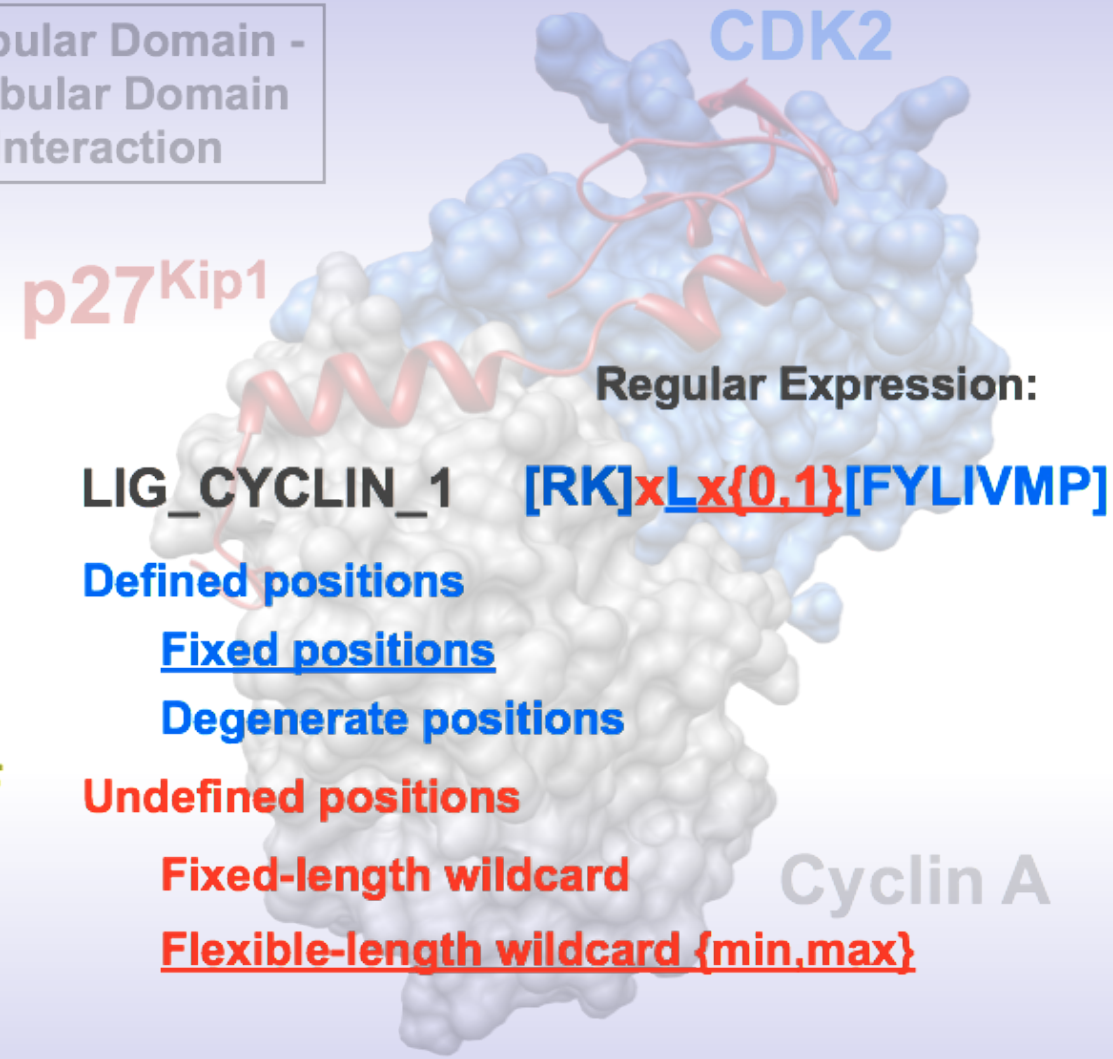
Globular Domain - Disordered Domain Interaction

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

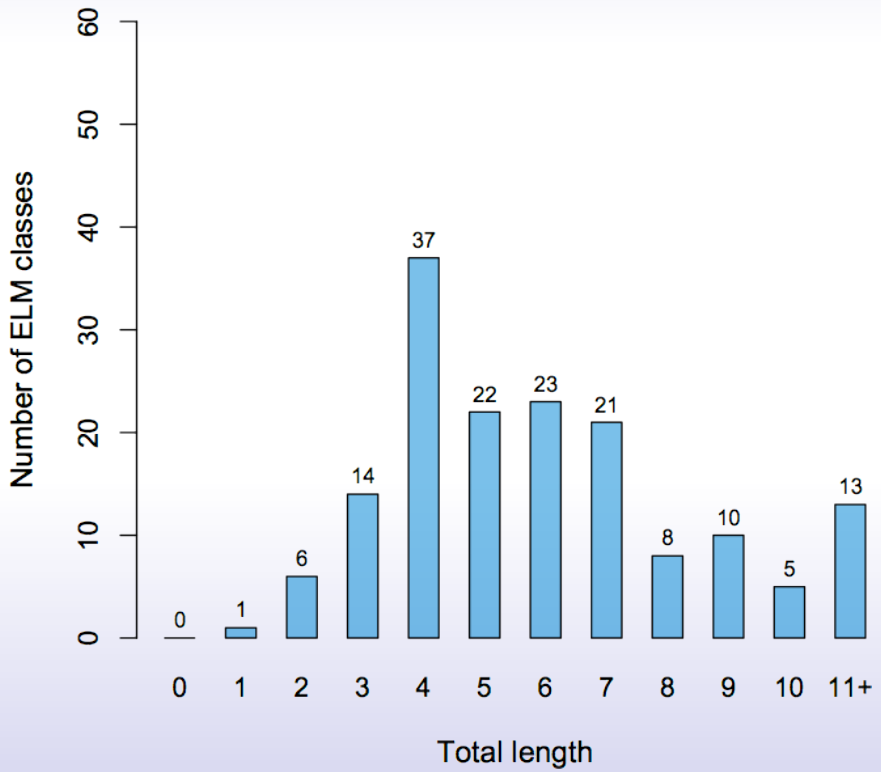
Fixed-length wildcard

Flexible-length wildcard {min,max}

Attributes of Short Linear Motifs

Linear Motifs

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

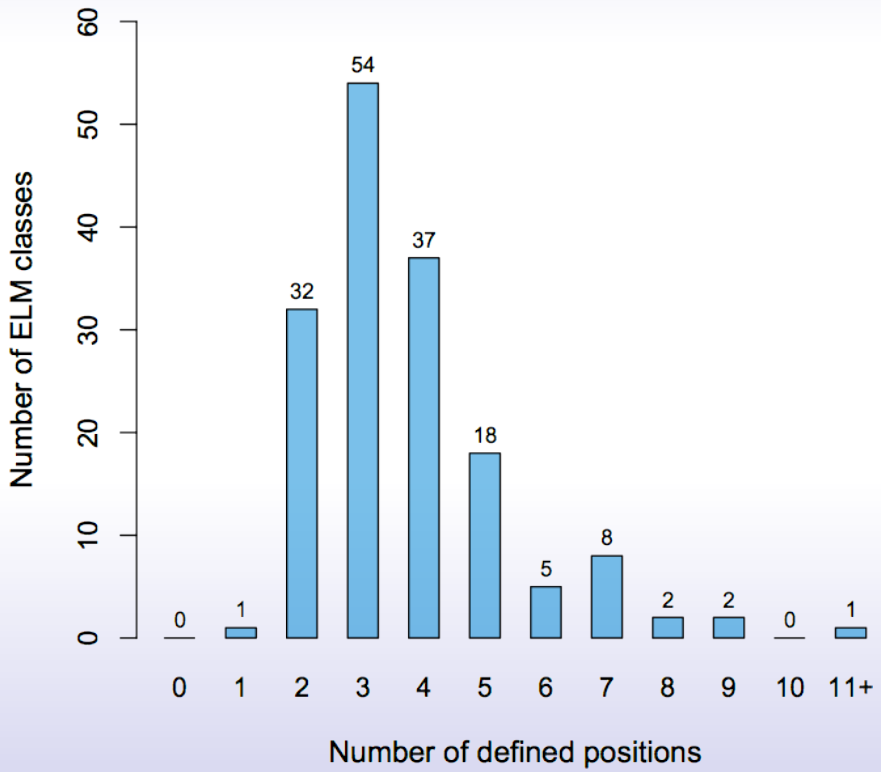


"Attributes of short linear motifs"; Davey, van Roey, Weatheritt, Toedt, Uyar, Altenberg, Budd, Diella, Dinkel & Gibson; (Mol Biosyst. 2011)

Attributes of Short Linear Motifs

Linear Motifs

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

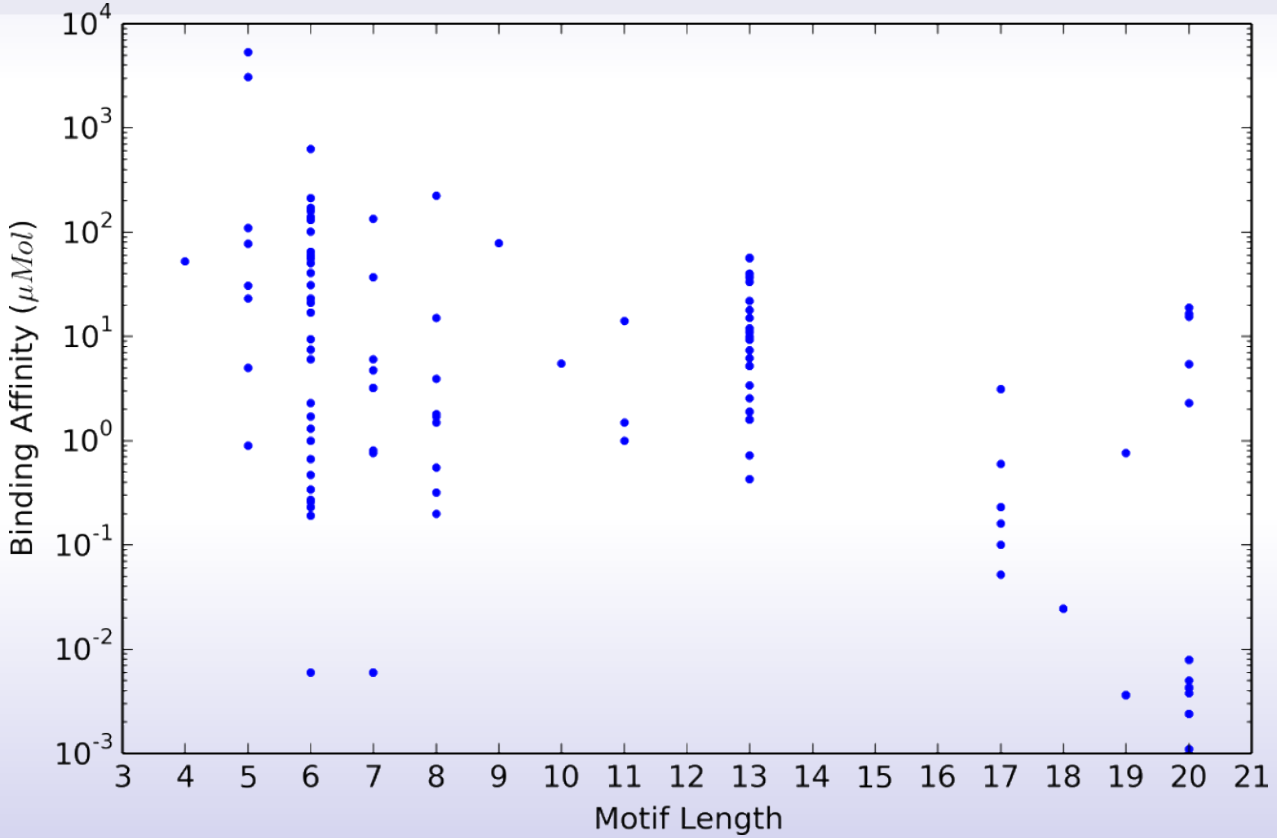


"Attributes of short linear motifs"; Davey, van Roey, Weatheritt, Toedt, Uyar, Altenberg, Budd, Diella, Dinkel & Gibson; (Mol Biosyst. 2011)

Attributes of Short Linear Motifs

Linear Motifs

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

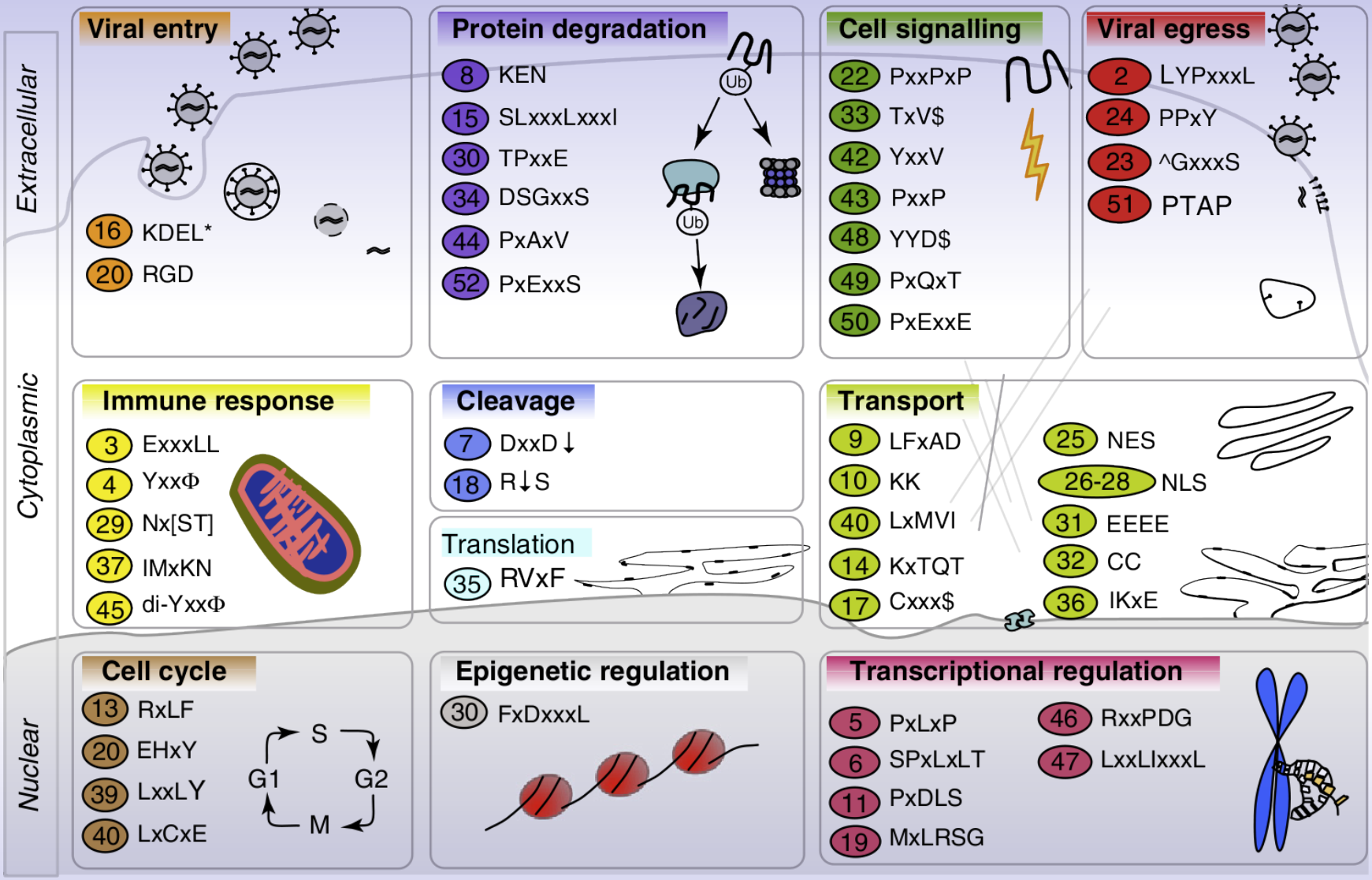


Prevalence of Short Linear Motifs

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	$NPx_{p}Y$

Importance of Short Linear Motifs: Viruses



"How viruses hijack cell regulation"; Davey, Travé & Gibson; (TiBS 2010)

Importance of Short Linear Motifs: 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.

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

Importance of Short Linear Motifs: 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 (one of the three proteins composing the anthrax toxin) 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.

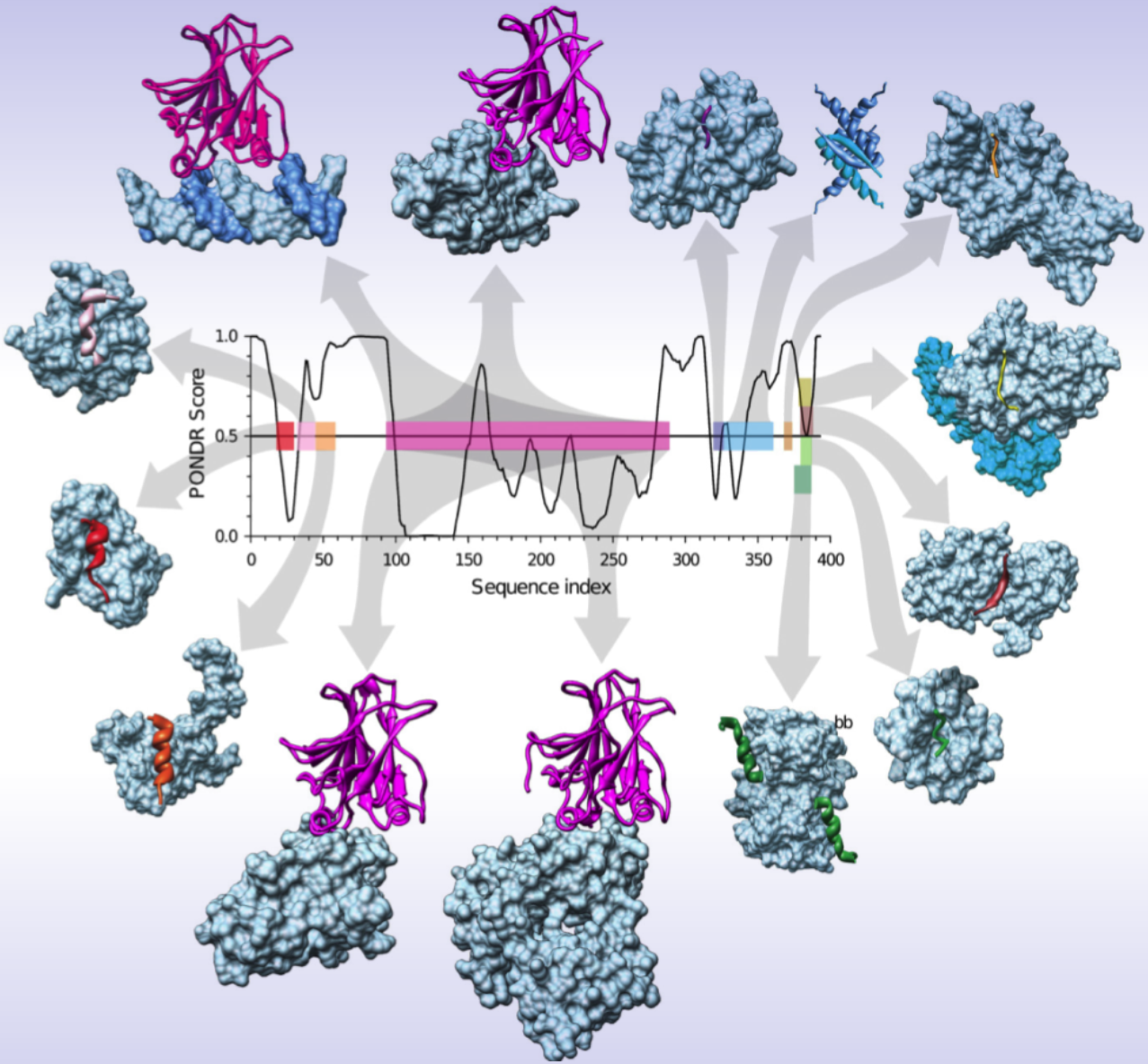
Importance of Short Linear Motifs: Cancer

β–Catenin

The most recurrently mutated experimentally validated motif in the COSMIC DB is the conserved proteasomal degradation motif (DEG_SCF_TRCP1_1)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)

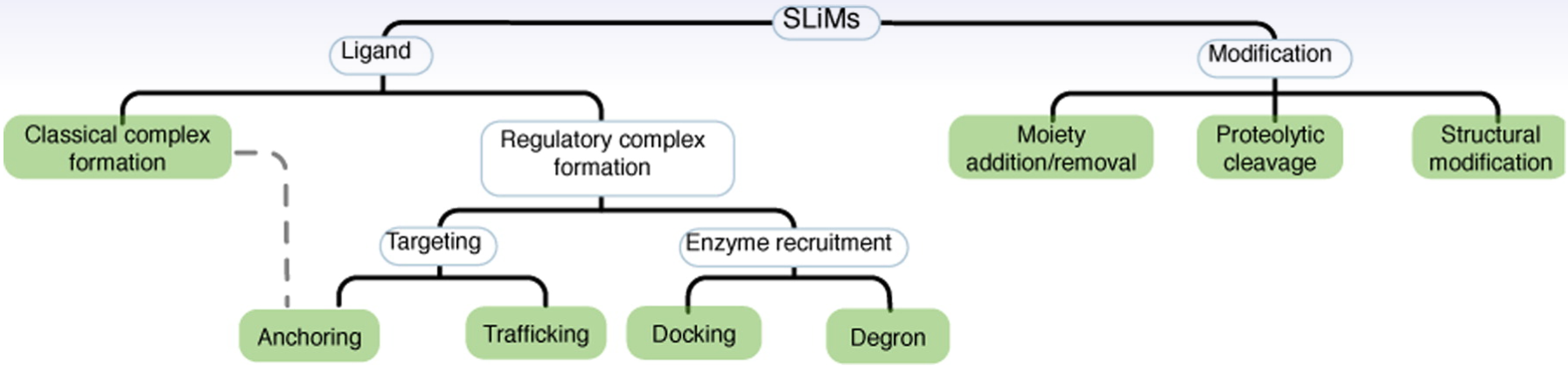
"Proteome-wide analysis of human disease mutations in short linear motifs: neglected players in cancer?"; Uyar, Weatheritt, Dinkel, Davey & Gibson; (Mol. Biosyst.; 2014)

IMPORTANCE OF SHORT LINEAR MOTIFS: P53



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

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

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

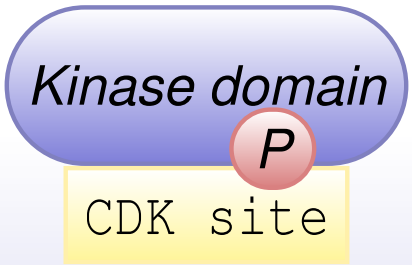
Description:

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

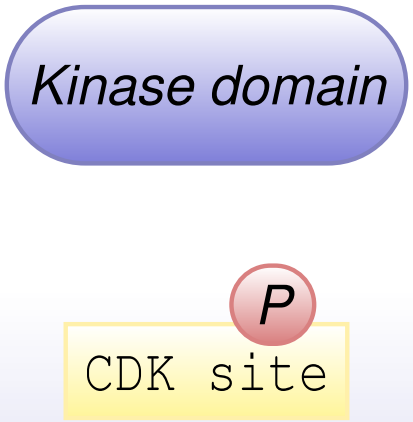
Description:

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



Motif Classes: Docking Motifs

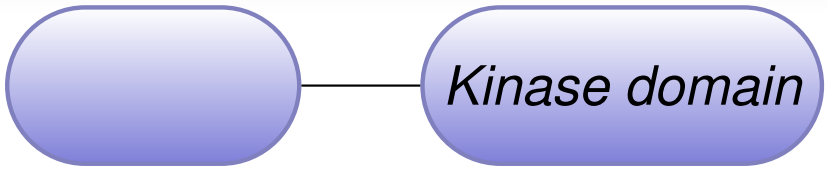
Description:

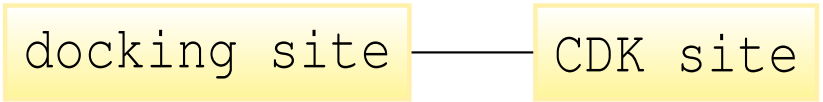
Docking motifs recruit enzymes via a surface that is distinct from the active site.

Example:

Name DOC_CYCLIN_1

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





Motif Classes: Docking Motifs

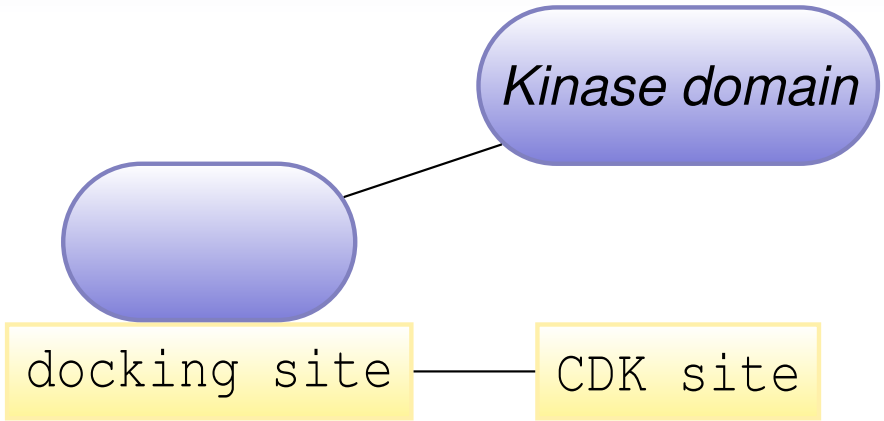
Description:

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

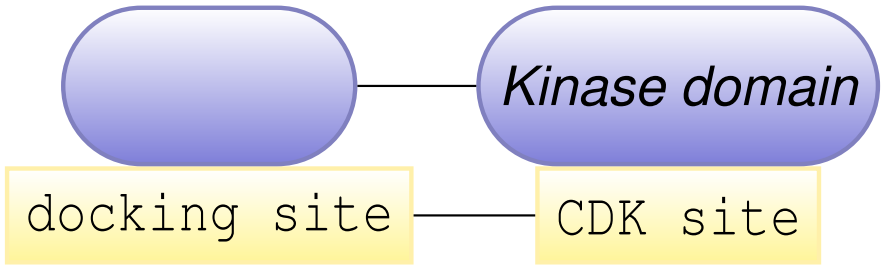
Description:

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

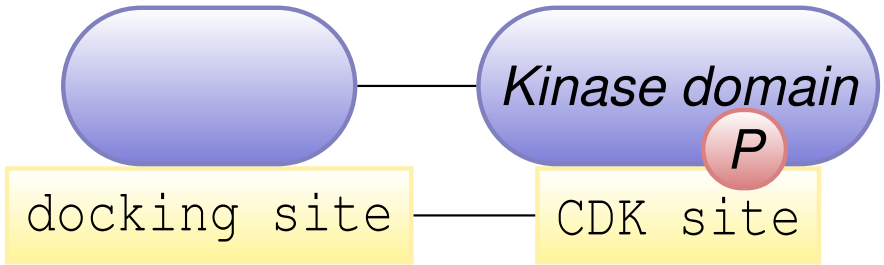
Description:

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

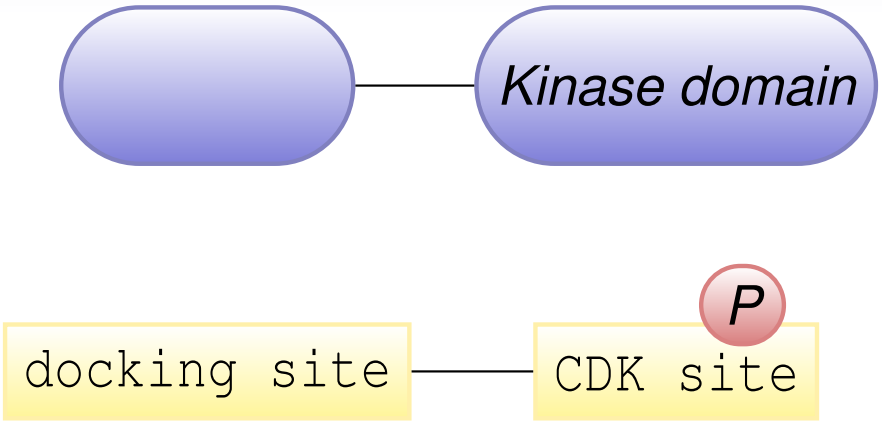
Description:

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 $E[IMPVL][MLVP]Rx$

Motif Classes: Cleavage Motifs

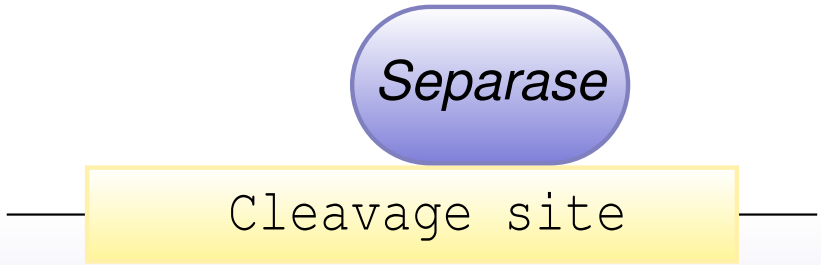
Description:

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

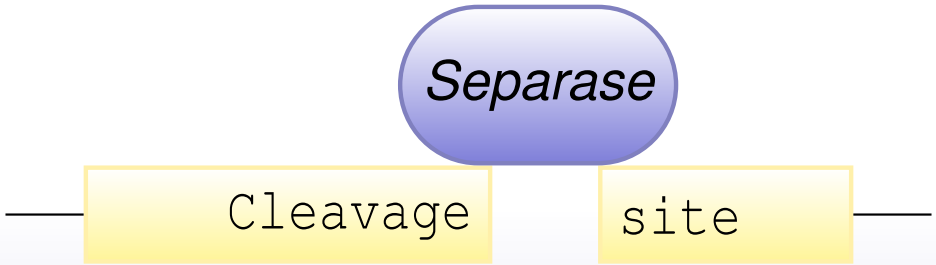
Description:

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

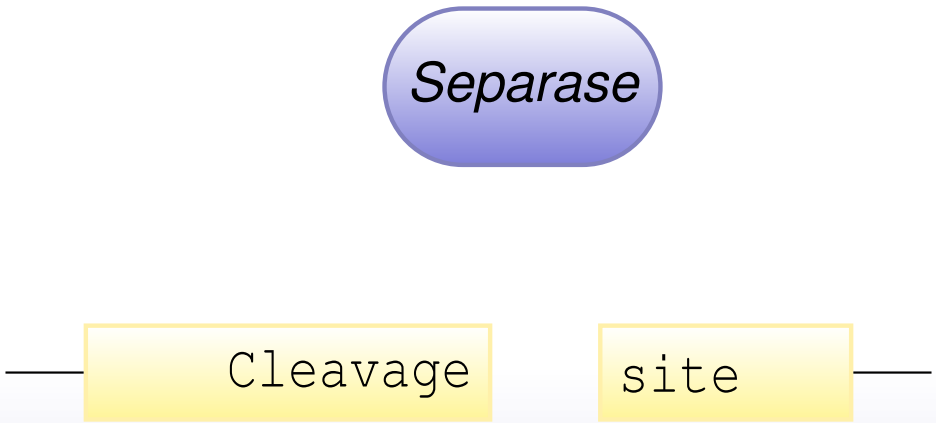
Description:

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

Example:

Name CLV_Separin_Metazoa

RegEx $E[IMPVL][MLVP]Rx$



Motif Classes: Degradation Motifs

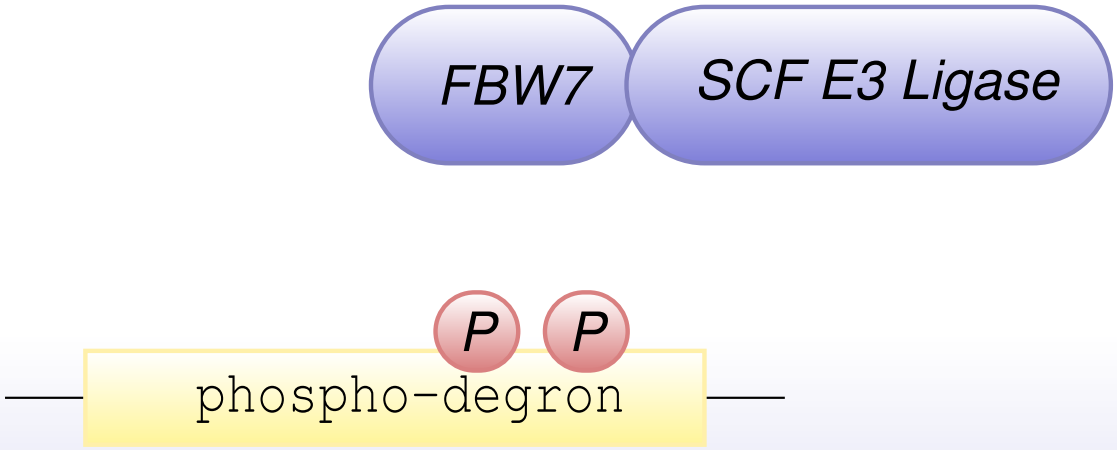
Description:

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

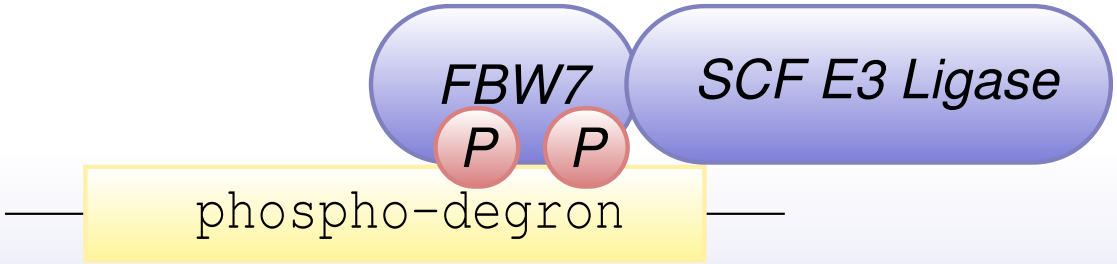
Description:

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

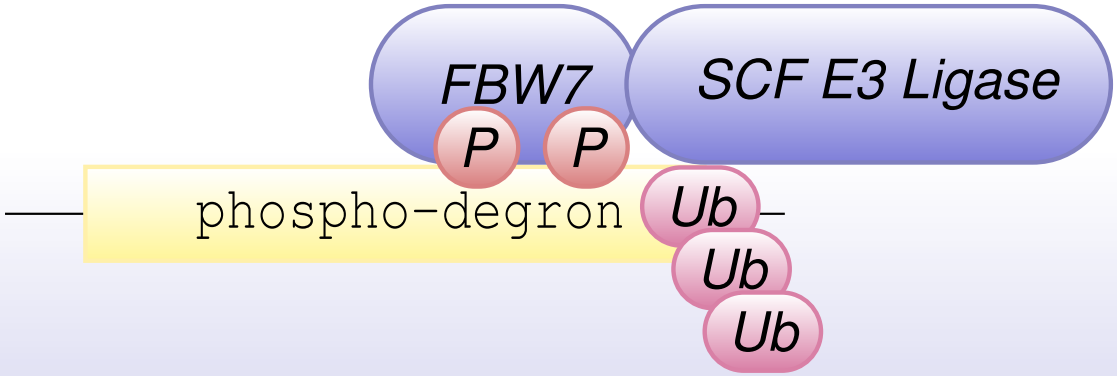
Description:

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: Degradation Motifs

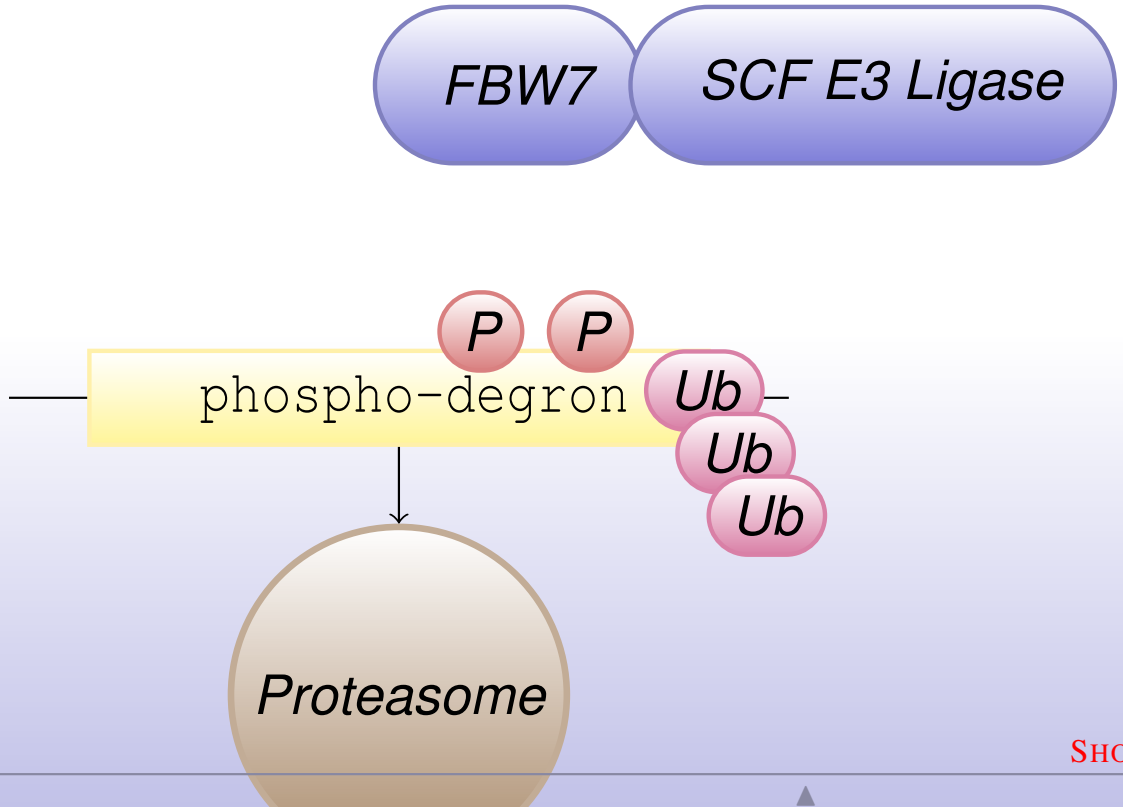
Description:

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:

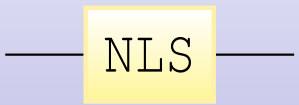
Targeting motifs allow a protein to bind to the transport machinery that relocalizes it to a particular sub-cellular location.

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]



Motif Classes: Targeting/Anchoring Motifs

Description:

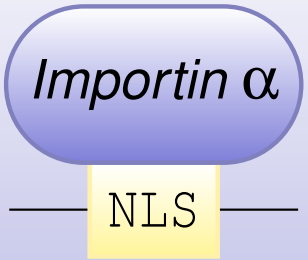
Targeting motifs allow a protein to bind to the transport machinery that relocalizes it to a particular sub-cellular location.

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:

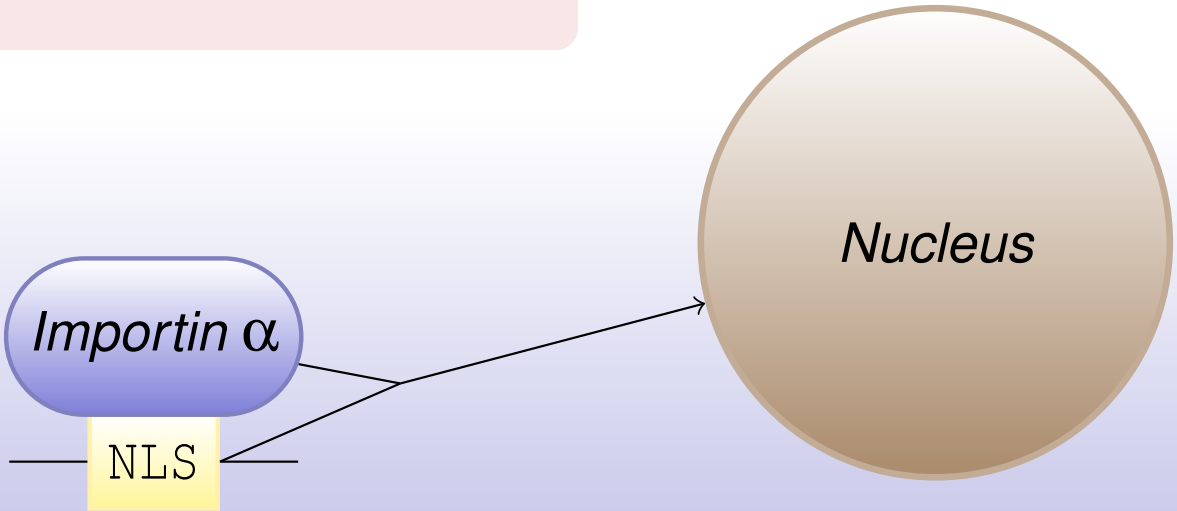
Targeting motifs allow a protein to bind to the transport machinery that relocalizes it to a particular sub-cellular location.

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

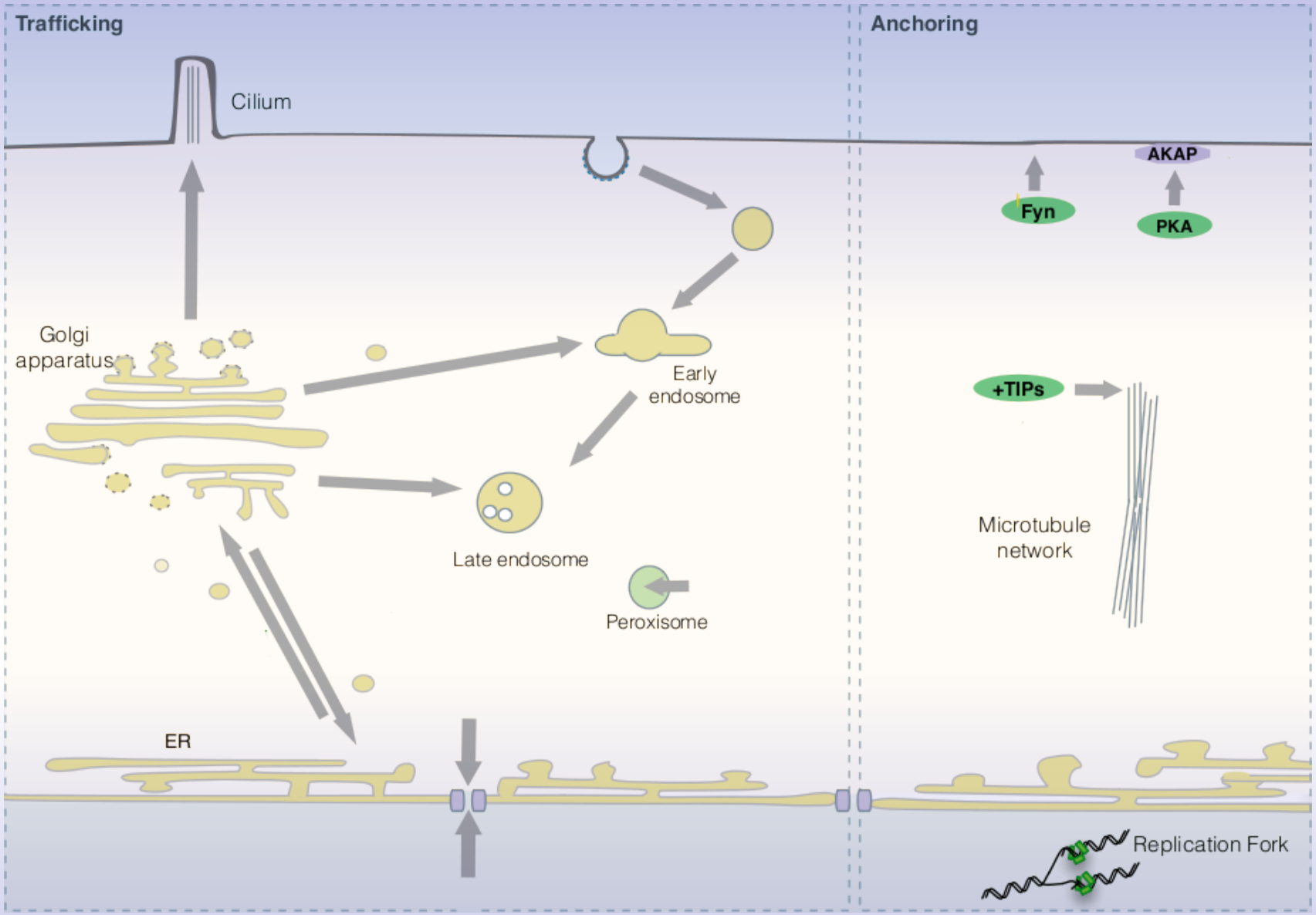
Example:

Name TRG_NLS_MonoCore_2

RegEx [^DE](K[RK]|RK)[KRP][KR][^DE]

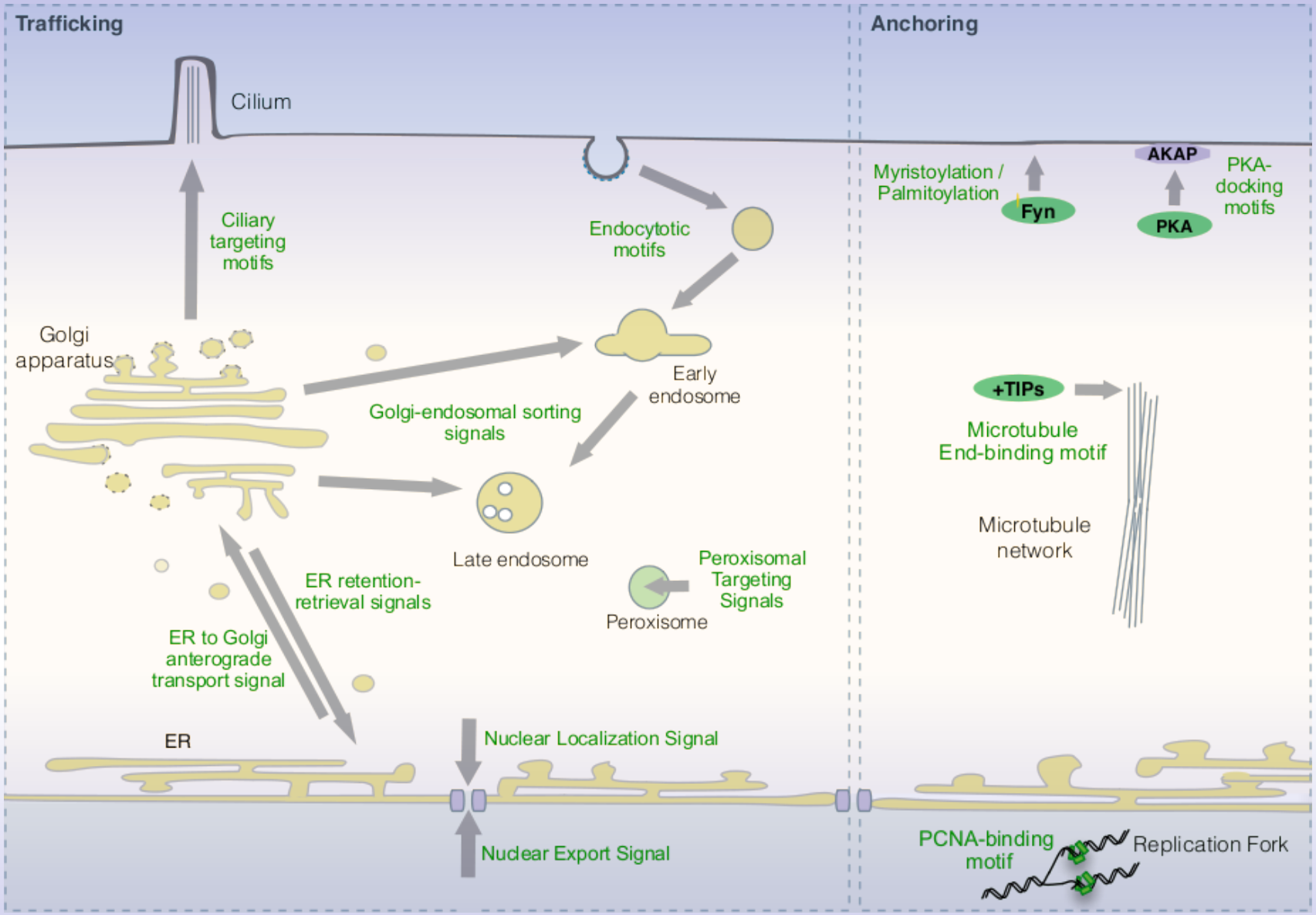


Motif Classes: Targeting/Anchoring Motifs



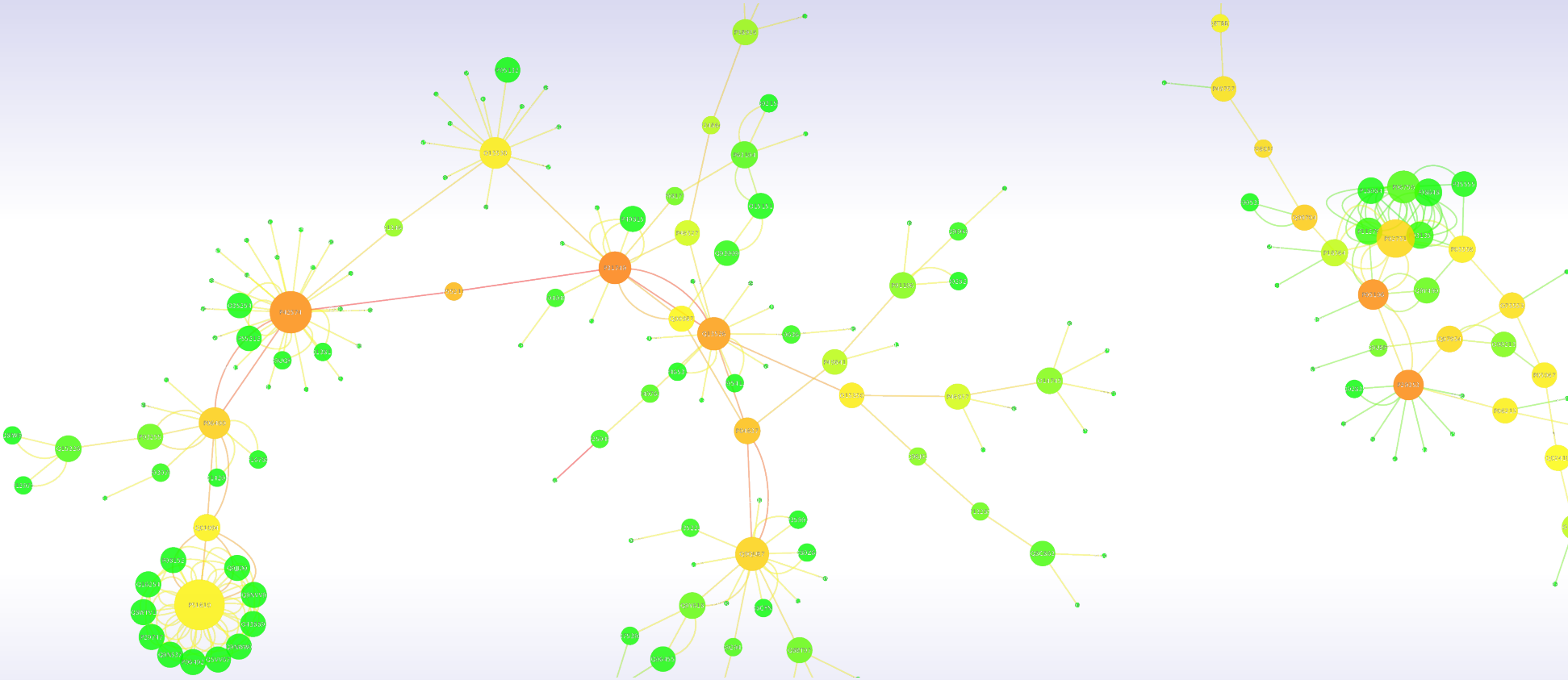
"Short linear motifs: Ubiquitous and functionally diverse protein interaction modules directing cell regulation"; van Roey, Uyar, Weatheritt, Dinkel, Seiler, Budd, Gibson & Davey; (Chem. Reviews; 2014)

MOTIF CLASSES: TARGETING/ANCHORING MOTIFS

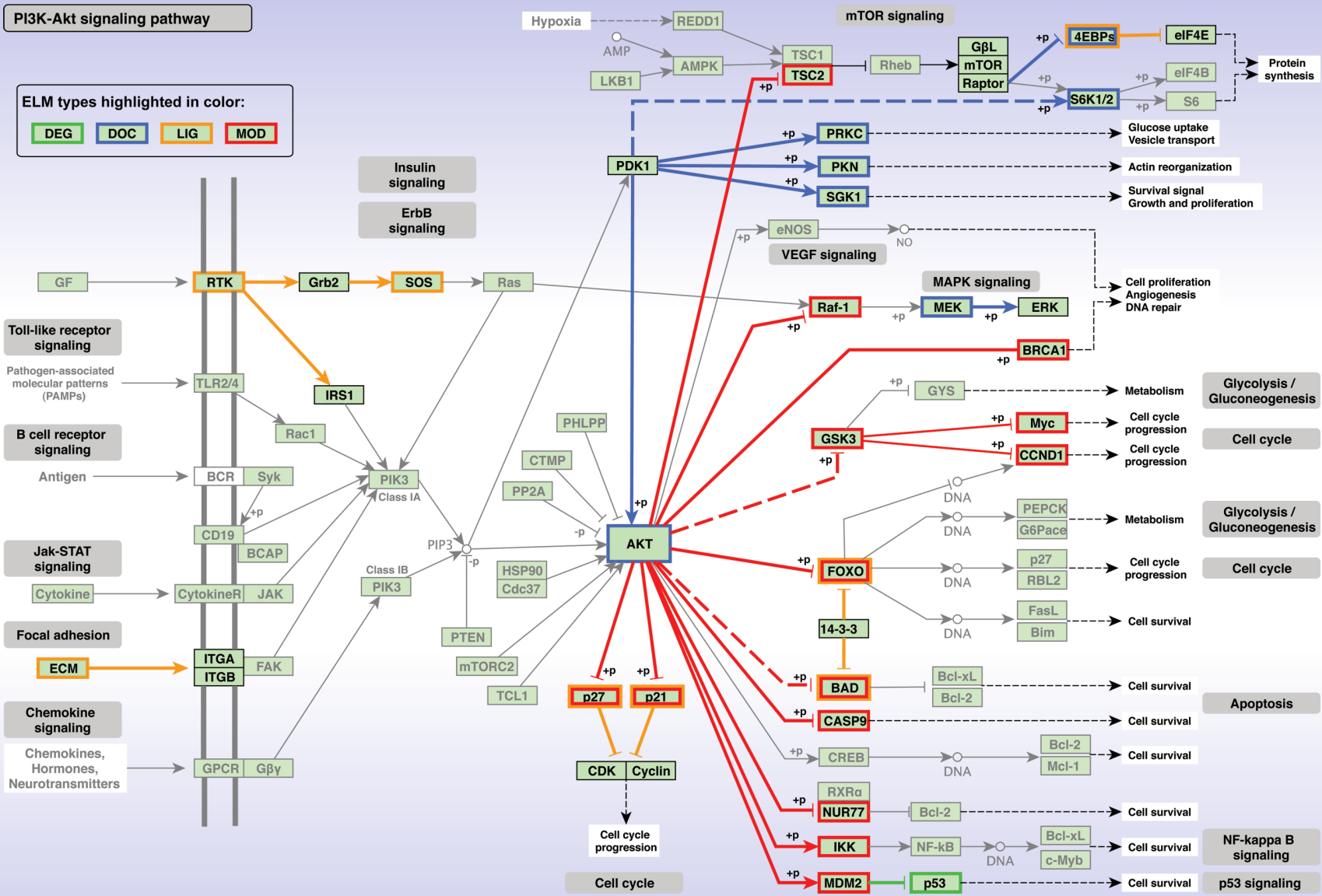


"Short linear motifs: Ubiquitous and functionally diverse protein interaction modules directing cell regulation"; VAN ROEY, UYAR, WEATHERITT, DINKEL, SEILER, BUDD, GIBSON & DAVEY; (CHEM. REVIEWS; 2014)

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?