Linear motifs and PTM sites

(http://elm.eu.org , http://phospho.elm.eu.org/, http://www.phosphosite.org, http://switches.elm.eu.org/)

Exercise 1.

Motifs and phosphorylation sites in P53

Search for the the human p53 protein in the ELM and the Phospho.ELM databases. (You can restrict the search for the nucleus and cytoplasm)

- How many motif hits are listed in ELM? How many of them are experimentally verified for the human p53? Can you find motif hits that are not yet verified for this particular protein but are plausible?
- Can you find any data indicating that p53 is sumylated?
- Is there any overlap of sumoylation and other PTMs?
- How many phosphorylation sites can you find in Phospho.ELM for p53?
- How many different kinases are involved?
- Find the modification in the Phosphositeplus database.
- What other types of modifications can you find?
- Find the graphic with modifications.
- Are PTMs mostly in the ~100AA disordered termini or in the folded DNA-binding domain?
- Some PTMs sites can't be clicked on in the graphic.
- Why might that be? To find out, track down where the PTM evidence comes from for the Cysteine residue C229.

Exercise 2.

Search for the 'MDM4_HUMAN' protein and find the 'USP binding motif' DOC_USP7_1 motif.

1. How many motif hits you can find?
2. Try to assess the biological significance of these motifs (based on the structural filter).

Exercise 3.

Search for the Cyclin dependent kinase inhibitor 1B' CDN1B_HUMAN (use the accession P46527)

protein in the Phospho.ELM database (second box)

1. How many positions are annotated? How many of these have multiple citations?
2. Which one are from high throughput experiments?
3. Which positions can be found in a known structural entry? What is accessibility of these positions? Visualize these positions in the structure.
4. Can you find positions that are conserved but disordered?
5. What kind of other information can you find about this protein in the PhosphoSitePlus database? (you can use the crosslink)

Exercise 4.

Find human beta catenin in the Uniprot database (`CTNB1_HUMAN`)

- Go down to the "natural variants" section. You can find several positions that are mutated in various types of cancer.
- Find more information about this protein in the ELM and Switches ELM (http://switches.elm.eu.org/) servers. How many switches have been identified for this protein ? What is the function of these switches.
- Characterize the frequently mutated positions structurally and functionally.
- Can you explain how mutations in these positions can lead to cancer?
- Can you find a structure corresponding to the mutated region? Model the mutation in the structure.

Exercise 5.

Discovery of linear motif sites based on their conservation pattern

- Find the amino acid sequence for protein **Q9UK97 in the uniprot database**
- Where can you find disordered binding regions in this protein using the ANCHOR server?
- Use the **SlimPrints server for this protein.**

http://bioware.ucd.ie/~compass/biowareweb/Server_pages/slimprints.php

- Which is the most significant hit**?**
- Check the sequence alignment. How conserved is the motif and its sequential neighborhood?
- Can you find this motif in other proteins**?**

Use the slimsearch server:

http://bioware.ucd.ie/~compass/biowareweb/Server_pages/slimsearch3.php

- Compare the domain composition of this hits.

Exercise 6.

PEX19_HUMAN

Compare isoforms 1 and 5 for PEX19 protein

(You can select both isoforms and put them in the basket in the Uniprot page and select them to align)

Where do they differ?

Can you find in this region experimentally verified motifs?
What is the biological function of this motif?

Interpret the difference between the two isoforms!