1. **Design 50-100 amino acids long artificial protein/peptide either manually or use the generator of random sequence**

(ExPASy tool RandSeq, using average amino acid composition)

**Perform physical-chemical analysis**

**Isoelectric point, molecular mass, solubility**

**Kyte-doolittle and Hopp-Woods profiles**

**Tendency to be unstructured**

**Hydrophobic cluster analysis**

ProtParam

ProtScale

PrDOS protein disorder prediciton server

Hydrophobic cluster analysis

http://mobyle.rpbs.univ-paris-diderot.fr/cgibin/portal.py?form=HCA#jobs::HCA.U29000760905027

1. **Compare your sequence against the whole PDB database**

**Compare your sequence with UNIPROT database http://www.uniprot.org/blast/**

**How many similar sequences you obtained**

1. **Can you calculate probability you find the exactly the same sequence within whole known sequences of proteins? (Use UNIPROT database)**
2. **Take your peptide and try to find out if the peptides contains any known functional motif at Prosite database**
3. **Run 3 different (most distinct) secondary structure predictions on your sequence and compare them. What is the expected accuracy?**
4. **Define a PROSITE motif which contain conserved linear DTG amino acids motif and having from both termini 3 random amino acids**
5. **O jakou třídu enzymů se jedná podle následujícího Prosite profilu?**

[LIVMFGAC]-[LIVMTADN]-[LIVFSA]-D-[ST]-G-[STAV]-[STAPDENQ]-{GQ}-[LIVMFSTNC]-{EGK}-[LIVMFGTA]

[D is the active site residue]

1. **Go to ELM** [**http://elm.eu.org/news.html**](http://elm.eu.org/news.html) **functional sites in proteins database and put there the following sequence Q9NP70 (UNIPROT database) . Analyze predicted function motifs and list all of them separately**