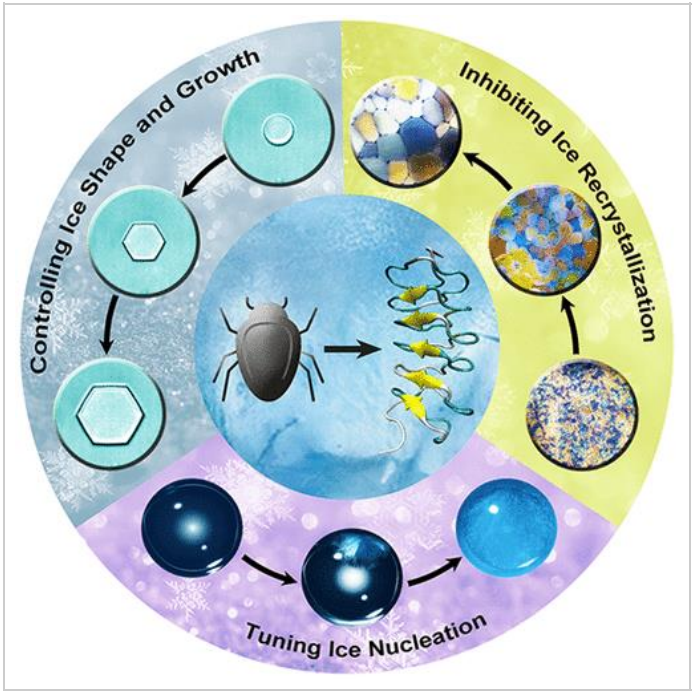
BIOANALYTICAL PROTEOMICS AND INTERACTOMICS
2018-19

Prof. Elisa Michelini

1. In silico cloning, recombinant expression and preliminary investigation on a new anti-freeze protein

Introduction

Ice formation, mainly consisting of ice nucleation, ice growth, and ice recrystallization, is ubiquitous and crucial in wide-ranging fields from cryobiology to atmospheric physics. Despite active research for more than a century, the mechanism of ice formation is still far from satisfactory. Meanwhile, nature has unique ways of controlling ice formation and can provide resourceful avenues to unravel the mechanism of ice formation [Acc Chem Res. 2018 Apr 17. doi: 10.1021/acs.accounts.7b00528.Bioinspired Materials for Controlling Ice Nucleation, Growth, and Recrystallization. He Z, Liu K, Wang J].



For instance, antifreeze proteins (AFPs) protect living organisms from freezing damage via controlling ice formation, for example, tuning ice nucleation, shaping ice crystals, and inhibiting ice growth and recrystallization.

In addition, AFP mimics can have applications in cryopreservation of cells, tissues, and organs, food storage, and anti-icing materials. Therefore, continuous efforts have been made to understand the mechanism of AFPs and design AFP inspired materials.

You have just obtained the sequence (cDNA) of a new antifreeze protein from an Antartic fish .

These proteins possess the ability to inhibit the formation of ice and are therefore essential to the survival of many marine teleost fishes that routinely encounter sub-zero temperatures.

ORIGIN

```
   1 agaagtctca gctacagctt tcacttcatt ctccgctaat taattaatta ttaattaatt
  61 aagtctcagc cacagctatg aattcagcta ttttaactgg tttccttttc gtcctccttt
 121 gtgtcgacaa catgacttca gccggctcgg gtaaatccgt ggtggccaac cagctgatcc
 181 ccataaatac tgccctgact cggatcatga tgaaggcgga gttggtcgcc ccaatgggca
 241 tccccgccga ggacattccc cgactagtca gtctgcaagt caacagggca gtgccgatgg
 301 gcacaaccct catgccagac atggtgaaaa cgtaccaacc agcgaagtaa ttctgagggt
 361 gccaaggagt ttcttcccaa aaccagaaga agaaatgccc cctctcacaa tcaaccttgt
 421 ttttgtcaga aacccaagtc tgtccggatg ttaactgaac atgtcaaaac ctgtggagac
 481 tatgttgaga tttgatggtc tgaaaagata aagcatataa ataaaatttt gcccaaaaaa
 541 aaaaaaaaaa
```

CDS 78..350

The aim of this tutorial is to clone this new sequence into a plasmid for heterologous expression and metabolic biotinylation.

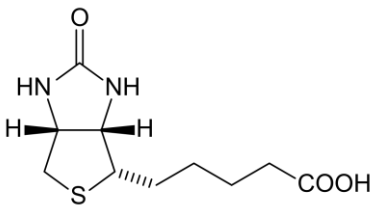
IN VIVO BIOTINYLATION SYSTEM TO STUDY PROTEIN FUNCTION

The biotin/avidin system is a very useful **affinity-based purification method**, well known in the biochemistry field. Compared to other affinity-based method, it has two major advantages:

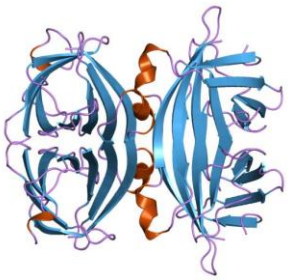
- Biotin/avidin binding is the **strongest non-covalent interaction** known in nature ($Kd=10^{-15}M$)
- There are **few naturally biotinylated proteins in mammalian cells**.

During biotinylation, **BirA ligase** from *E. coli* **adds a biotin moiety to the lysine residue of the 23 AA tag added either at the N-terminal or C-terminal** end of the modified target protein.

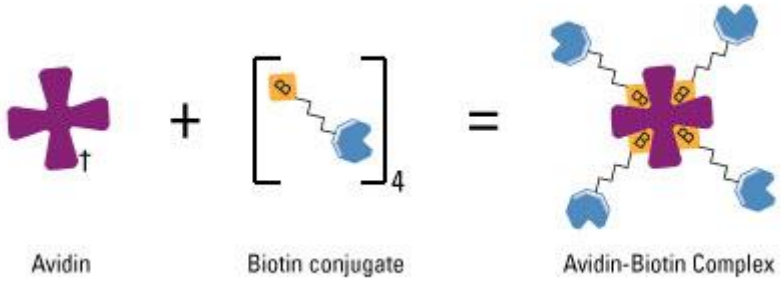
A one-step purification step is sufficient for protein complex identification by MS. Alternatively biotinylated proteins can be cross linked to DNA for ChIPseq analysis.



BIOTIN



STREPTAVIDIN

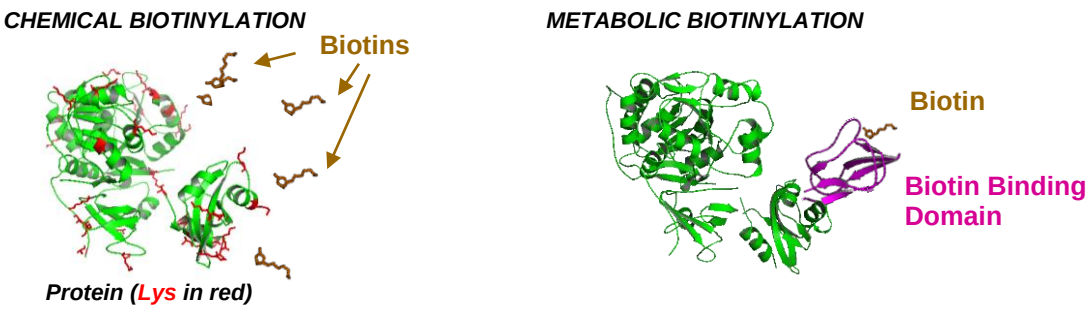


Streptavidin can bind up to four molecules of biotin simultaneously with a high degree of affinity and specificity.

Biotin is widely used to conjugate proteins for biochemical assays. Thanks to its small size the biological activity of the protein will most likely be unaffected during biotinylation. Because both streptavidin and avidin bind biotin with high affinity (Kd of 10^{-14} mol/l to 10^{-15} mol/l) and specificity, biotinylated proteins of interest can be isolated from a sample by exploiting this highly stable interaction.

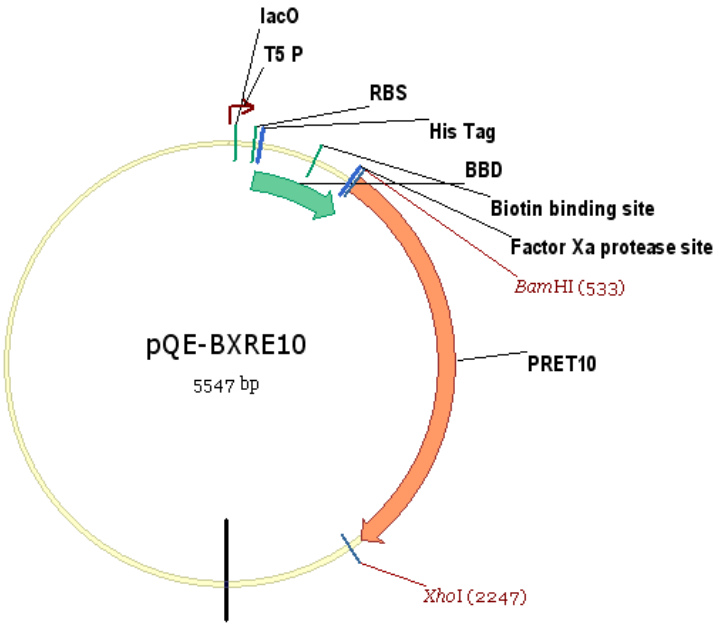
Streptavidin-biotin complex is highly resistant to organic solvents, detergents (e.g. SDS, Triton), denaturants (e.g. guanidinium chloride), proteolytic enzymes, high ranges of temperature and pH.

IN VIVO BIOTINYLATION VS. CHEMICAL BIOTINYLATION



The principle of in vivo biotinylation is based on coexpression of the **biotinylation enzyme BirA** from *Escherichia coli* together with the **protein of interest** fused to a short peptide, the **biotin binding domain**, shown to be an efficient target for BirA.

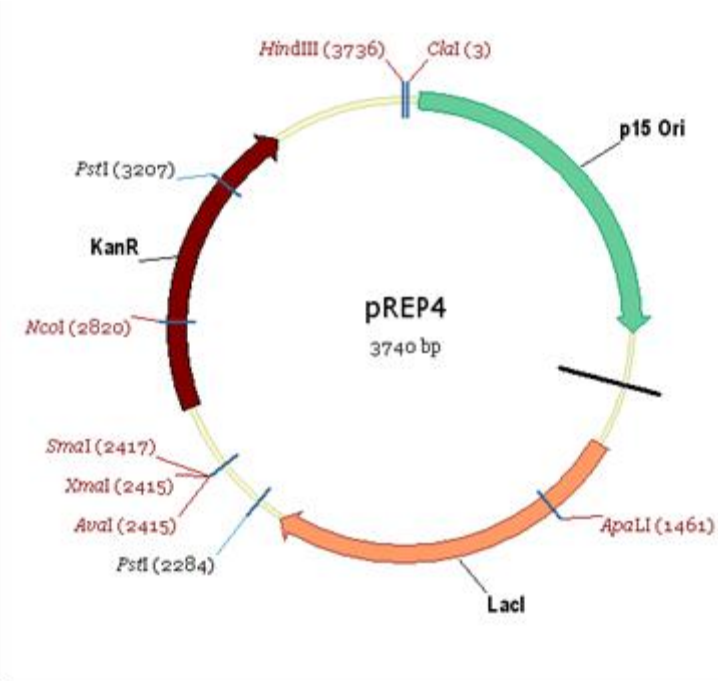
The expression plasmid pQEBXRE10 (3.4 kb) contains a 6xHis-tag coding sequence 5' to the multiple cloning site designed to generate N-terminal tagged proteins with the polyhistidine tag in order to purify the tagged protein using a Ni^{2+}-affinity chromatography column.
pQE-BXRE10 plasmid encoding for PpyRE10 *Photinus pyralis* mutant (5547 bp) carries an optimized version of the *Klebsiella pneumoniae* **oxaloacetate decarboxylase biotinyl-binding domain** with the target sequence for biotinylation 3' to the reporter sequence.
It has the following aminoacids:
MRGSHHHHHHVMARVDVSQLTAAAPAPAPAPAPASAPAAAAPAGAGTPVTAPLAG
TIWKVLASEGQTVAAGEVLLILEA**MKM**ETEIRAAQAGTVRGIAVKAGDAVAVGDTL
MTLATVRANSSTLAAVTTGSGSIEGRGSGS (underlined represents the target sequence containing lysine to which biotin prosthetic group is metabolically bound in the bacterial system). pQEXR10 is a high copy number plasmid because of the presence of relaxed ColE1 origin of replication. Other regulatory elements in the plasmid are: phage T5 promoter (recognized by the *Escherichia coli* RNA polymerase), two lac operator sequences (to increase lac repressor binding and ensure high repression of the strong T5 promoter) and two strong transcriptional terminators (t0 from phage lambda and T1 from *Escherichia coli* rrnB operon) for guaranteeing stability of the expression construct.
This is an IPTG inducible expression system where isopropyl β-D-1-thiogalactopyranoside binds the lac repressor stabilizing it in the non-DNA-binding conformation and leading to hyper-expression levels of the downstream gene.
Synthetic ribosomal binding site (RBSII) increases translation rates and β-lactamase gene (bla) confers resistance to ampicillin (as reported on QIAexpressionist, QIAGEN, June 2003).

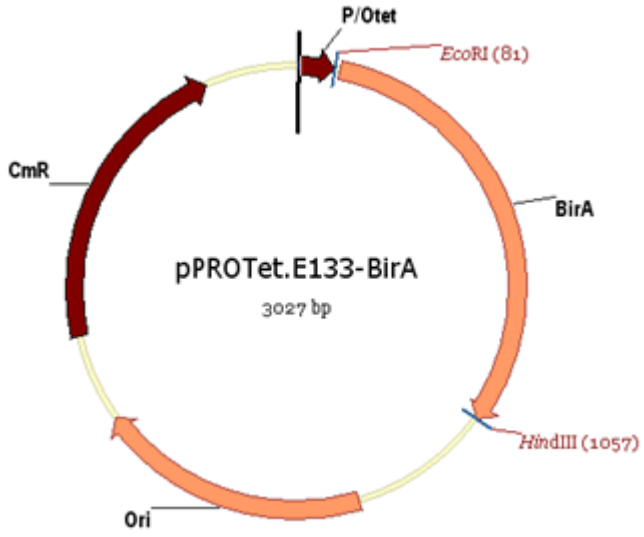


BirA cloning

The *Escherichia coli* **biotin holoenzyme synthetase**, **BirA**, catalyses transfer of biotin to the epsilon amino group of a specific lysine residue of the biotin carboxyl carrier protein (BCCP) subunit of acetyl-CoA carboxylase. It is a two-step reaction in which the enzyme activates biotin to form biotinyl 5' adenylate and then attach the biotin to biotin-accepting proteins. Biotin holoenzyme synthetase also functions as a biotin operon repressor [Beckett *et al.*, 1999]. BirA is encoded by the commercially available pDW363 plasmid (6.7 kbp).

pPROTet.E133 plasmid (2151 bp) carries the ColE1 high-copy origin of replication and the tetracycline-regulated PltetO^{-} promoter from the tet operon Tn10. Hence the sequence under its regulation is transcripted in presence of tetracycline or one of its analogues (such anhydrotetracycline, which is effective at lower concentrations and less toxic). It encodes for chloramphenicol resistance.





Task 1

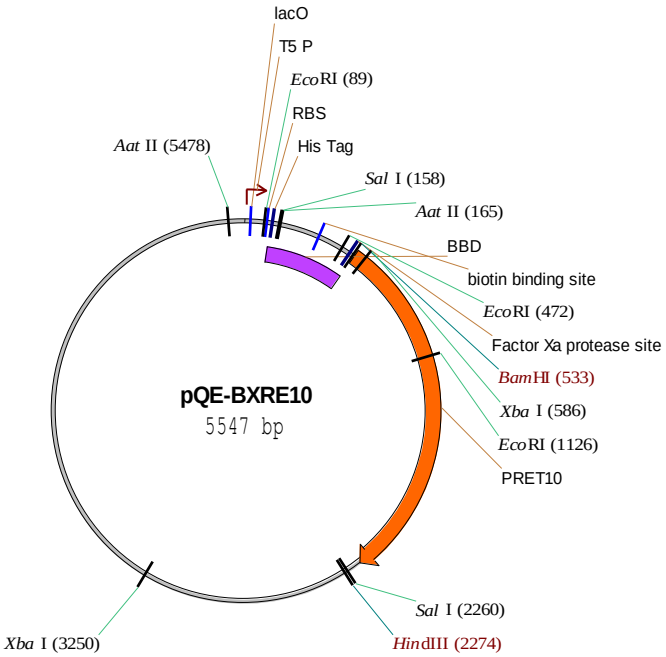
- Download Vector NTI Advance® 11.5 from website
http://www.invitrogen.com/site/us/en/home/Products-and-Services/Applications/Cloning/vector-nti-software/vector-nti-advance-software/try-vector-nti-advance-for-free.html

- Import sequence into Vector NT database
- add sequence reference points (CDS) to map

Task 2

Clone cDNA encoding for the protein into a new plasmid for bacterial expression and biotinylation pQEBXRE10 and replace gene PRET10 (red arrow).

Select cloning strategy with suitable restriction enzymes:
- identify enzymes that cut once upstream and downstream the sequence to be replaced
-check that enzymes do not cut inside the insert
-design PCR primers (with suitable restriction sites) using Vector NTI or on line tools



Primers features

1. primers should be 17-28 bases in length;
2. base composition should be 50-60% (G+C);
3. Tm between 55-70^{o}C are preferred;
4. 3'-ends of primers should not be complementary (ie. base pair), as otherwise primer dimers will be synthesised preferentially to any other product;
5. primer self-complementarity (ability to form 2^{o} structures such as hairpins) should be avoided;
6. runs of three or more Cs or Gs at the 3'-ends of primers may promote mispriming at G or C-rich sequences (because of stability of annealing), and should be avoided.

Once primers have been designed the experimental procedure involves:
-Amplification (PCR) of target sequence
-Restriction digestion of amplified sequence and vector with same enzymes
-Ligase reaction of insert and vector with a DNA ligase
- Transformation into competent bacterial cells (DNA plasmid are transformed into E.coli cells and are able to self-replicate)
-Next step is to verify that cells have taken the correct DNA. This can be done by isolating the plasmid DNA from single clones and by performing restriction digestions

Task 3. Perform Blast search with both nucleic acid sequence and aminoacid sequence using AlignX (if available with trial version) and BLAST.

Task 4. Identify the closest antifreeze proteins(s) on the basis of nucleic acid and protein sequences.

Task 5: perform reverse translation using different codon usage parameters. Select an on line tool (e.g. http://www.ebi.ac.uk/Tools/st/emboss_backtranambig)