Lectures 1: Recombinant Proteins

Lectures:
Students should:

- Understand the basic features of recombinant expression systems used in modern research.
- Be able to describe why different systems are chosen for different proteins and provide examples.
- Be able to describe the structure and function of antibodies and engineered forms
- Understand the basics of how monoclonal antibodies are developed.
- Understand the how and why monoclonal antibodies are "humanised"
- Be able to use examples to describe the importance of antbodies as therapeutic agents.

Insulin Glargine- made in yeast

Epoetin alfa- made in mammalian cells

Pegfilgrastim- made in bacteria

Top Selling Therapeutic Proteins in 2010 ($US billion)

# (09)	Name	Target/mechanism	Type	Company	Indication	Sales
1 (1)	Etanercept	TNFα	Fc fusion TNFR2 ECD	Amgen Wyeth	Immune diseases	7.287
2 (3)	Bevacizumab	VEGF	Humanized IgG	Genentech Roche Chugai	Cancer	6.973
3 (4)	Rituximab	CD20	Chimeric IgG	Genentech Biogen-IDEC Roche	Cancer	6.859
4 (5)	Adalimumab	TNFα	Human IgG	Abbott Eisai	Immune diseases	6.548
5 (2)	Infliximab	TNFα	Chimeric IgG	Centocor (J&J) Schering-Plough Mitsubishi Tanabe	Immune diseases	6.520
6 (7)	Trastuzumab	Her2	Humanized IgG	Genentech Chugai Roche	Cancer	5.859
7 (8)	Insulin glargine	Insulin receptor	Modified insulin	Sanofi-Aventis	Diabetes	4.834
8 (6)	Epoetin alfa	EPO-R	Human EPO	Amgen Ortho Biotech Kyowa Hakko Kirin	Anemia	4.590
9 (9)	Pegfilgrastim	G-CSF receptor	PEGhuman G-CSF	Amgen	Neutropenia	3.558
10 (11)	Ranibizumab	VEGF	Humanized Fab	Genentech Novartis	AMD	3.106
11 (10)	Darbepoetin alfa	EPO-R	Modified human EPO	Amgen Kyowa Hakko Kirin	Anemia	2.995
12 (12)	Interferon beta-1a (Avonex)	Interferon beta receptor	Human interferon beta-1a	Biogen Idec	Multiple sclerosis	2.518
13 (13)	Interferon beta-1a (Rebif)	Interferon beta receptor	Human interferon beta-1a	Merck Serono	Multiple sclerosis	2.297
14 (17)	Insulin aspart	Insulin receptor	Modified insulin	Novo Nordisk	Diabetes	2.198

Why do we need to use different host organisms?

- cost
- ethics
- folding and modifications…making the correct protein
    - Insulin: di-sulfide bonds, (yeast glycosylate but it's a different process)
    - Pegfilgrastim: normally glycosylated in humans but not in bacteria however wasn't important for the function

Why make proteins recombinantly?

- can't get it from endogenous sources
- for efficient and selective purification
- quality control (large batches and they need to be consistent)
- to optimize activity/efficacy

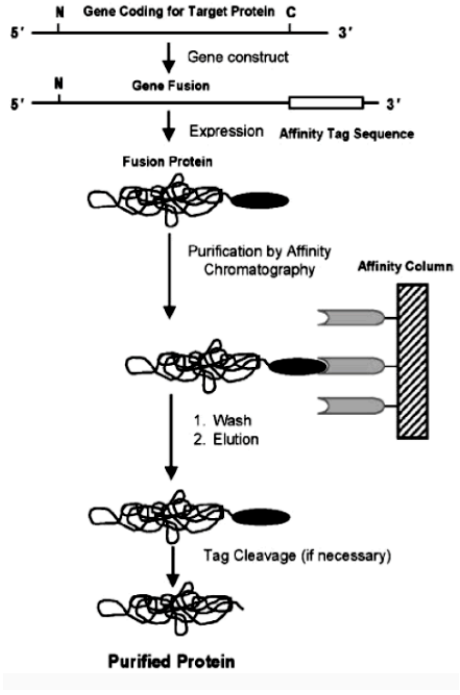
Heterologous hosts for protein expression

- Bacteria
- Yeast (Pichia sp, Sachromyces cerevisiae)
- Plants
- Baculovirus
    - virus infects insect cells and makes proteins
- Cultured mammalian cells
- Animals
    - e.g. goats- purified from goat milk

<table>
<tr><td>Bacterial Expression</td><td>
<table>
<tr><th>Pros</th><th>Cons</th></tr>
<tr><td>• widely used
• easy manipulation
• rapid growth ∴ cheap
• many commercial vectors and tags to add to proteins to enhance purification</td><td>• no post-translational modifications
  ○ because theres no secretory pathway (no membrane bound organelles)
• most proteins expressed are non-soluble
  ○ large quantities may disrupt folding and result in formation of fusion bodies
• high endotoxin content
  ○ need efficient purification processes</td></tr>
</table>
</td></tr>
<tr><td>Expression as fusion proteins</td><td>
Most target proteins are not simple to <u>purify</u> from a cell

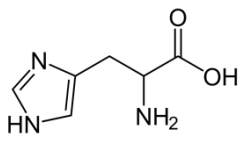
• <u>Alternate approach</u>
  ○ genetically fuse the gene encoding the target protein with a gene encoding a purification tag.
    ▪ Fusion proteins enhance solubility

  ○ When the chimeric protein is expressed, the tag allows for specific capture of the fusion protein. This will allow the purification of virtually any protein without any prior knowledge of its biochemical properties.



• <u>Common tags (which bind to the Affinity column):</u>
  ○ glutathione S-transferase
    ▪ In Biotechnology: fusion proteins of interest can undergo purification from the cells due to high glutathione affinity
    ▪ metabolic isozymes which catalyze the conjugation of the reduced form of glutathione (GSH) to xenobiotic substrates for the purpose of detoxification
    ▪ Glutathione is an essential metabolic molecule that is produced in the liver of humans and animals

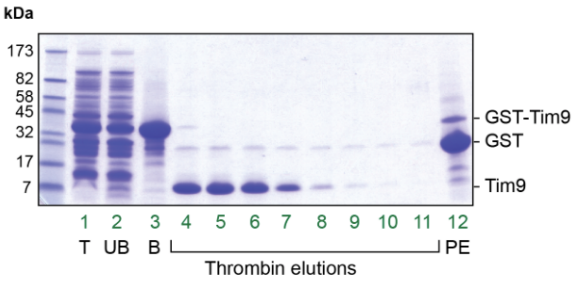
  ○ maltose binding protein
    ▪ binds to maltose

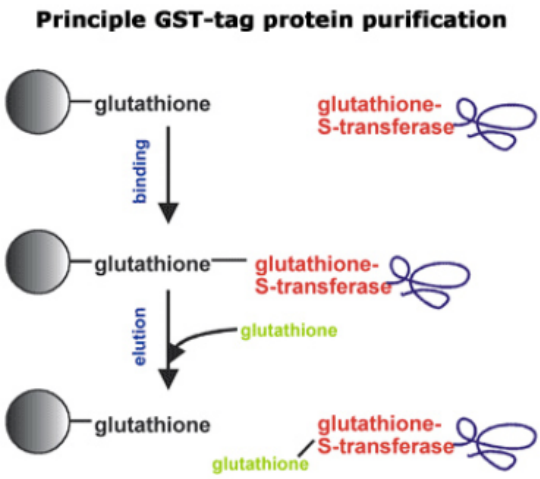
  ○ His-tag (separated by protease cleavage site for on-column cleavage)
    ▪ binds to nickel
    ▪ poly-Histidine tag: comprises 6 – 14 histidines and is typically fused to the N- or C-terminal end of a target protein
    ▪ hydrophilic and flexible
      ▪ ⇑ solubility of target proteins
      ▪ rarely interfere with protein's function


</td></tr>
</table>

Plasmid Vectors

e.g. for Glutathione S transferase (GST) tag to isolate and urify Tim9….sequence contained a sequence which could be cut by Thrombin (a protease) thus we could have purified Tim9





Pros and **Cons** for using **tags** in **fusion proteins**

Pros	Cons

• improve protein yield
• prevent proteolysis
• facilitate protein folding
• increase solubility (the fusion domain)
• ease of purification
	• lower protein yields
	  ○ cleavage domain may not be complete thus protease can't cleave
	• alteration in biological activity (few extra amino acids)
	• cleavage/removing the fusion partner may require expensive proteases
	  ○ *e.g.* Factor Xa→ enterokinase

Where to **target expression** of **recombinant proteins in bacteria?**

Direct Expression (*Cytosol*):

- *E. coli* cytoplasm is a **reducing** environment
  - ∴ cannot get proper disulfide S-S bonds

bacteria do have disulfide bond forming capabilities but only in the periplasm. We can't secrete that much protein into the periplasm thus we don't rely on bacterial disulfide bond formation

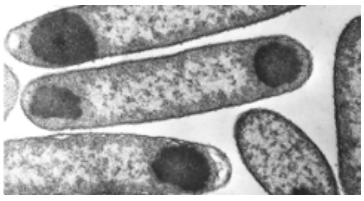
Secretion (*periplasm* or *medium*):

- fuse target protein with a peptide/protein which targets in for **secretion**
- periplasm offers a more **oxidaizing** environment
  ∴ **proteins fold better**

- Limitations:
  - limited capapcity for sercretion (0.1-0.2% total cell protein compared to 10% produced intracellularly)
  - limited capacity for post-traslational modifications of proteins

Inclusion Bodies in *E. coli*

- Dense particles which contain precipitated (**insoluble**)



- Formation depends on:
  - protein synthesis rate
  - growth conditions

- Protein **refolding** options exist but poor success rates