Lecture 1 – Introduction and Review

Objectives:

- Understand the basic features of recombinant expression systems used in modern research
    - Recombinant proteins; state of the art, problems, new developments
    - Antibodies, design production, engineering
    - Personalised medicine: gene therapy approaches now and beyond
- Be able to describe why different systems are chosen for different proteins and provide examples

Summary:

- Why make proteins recombinantly?
    - Unavailable from endogenous sources
    - For more efficient and selective purification
    - Quality control → large batches
    - Optimise activity/efficacy → by adapting the protein to be a better therapeutic protein
- Possible heterologous hosts for protein expression
    - Bacteria; Yeast (Pichia sp.); Plants; Baculovirus; Cultured mammalian cells; Animals
- Bacteria
    - Pros: Widely used; Easy to manipulate; Rapid growth; Cheap; Many commercial vectors and tags to add to protein codes to enhance purification
    - Cons: No post-translational modifications; High endotoxin content; *Many proteins expressed are insoluble* → Creates inclusion bodies = dense particles of insoluble proteins; the formation depends on protein synthesis rate and growth conditions; usually not quite the right folding conformation (can refold BUT poor success rate)
    - Site of expression
        - Directly in the cytoplasm → E. coli cytoplasm is a reducing environment = no proper disulphide bonds
            - Site of production of 10% of the total cell protein
        - Secretion in the periplasm or medium (a layer around the cytoplasm) → the target protein is fused to peptides/proteins targeted for secretion; this area is more oxidising so some proteins can be folded better
            - Bacteria have limited capacity for secretion (0.1 – 0.2% total cell protein)
- Expression of a fusion protein
    - Lots of proteins are made in a cell, but usually only one is wanted. Even though in most cases the protein of interest is far over expressed, to aid purification a unique tag will be added to the protein code, which can be identified while part of the protein but also can be cleaved once the proteins are purified
        - Given the knowledge of the tags proteins biochemistry is known, a source to bind to it, or to which it will bind can be used to isolate the protein → this also means that nothing needs to be known about the target protein
    - Common tags
        - Glutathione S-transferase (GST)
        - Maltose binding protein
        - His-tag → separated by protease cleavage site for on-column cleavage
    - Advantages
        - Improve protein yield
        - Prevent proteolysis
        - Facilitate protein refolding
        - Increase solubility
        - Ease of purification
    - Disadvantages
        - Lower protein yields → cleavage may not be complete or cleaves within the protein
        - Alteration in biological activity
        - Cleavage/removing the fusion partner may require expensive proteases
- Insulin
    - Uses
        - Single gene hormone, secreted by beta cells of islet cells of Langerhans of the pancreas
        - Its secretion is stimulated by glucose, amino acids and fatty acids

        - Binds to insulin receptors on cells and stimulates uptake and storage of fuels
    - Production
        - Trafficked through the secretory pathway: ER → Golgi → secretory vesicles
        - Disulphide bonds are present, NO glycosylation
        - Proteolytic processing → remove signal sequence and C-peptide → leaves A and B chains
    - Initially obtained insulin from pigs → injected into patients several times/day in crystalline form
        - Protamine → reduced insulin solubility → slows absorption → prolonged action
        - Zinc → prolongs effects
        - However, patients began to develop immunity to porcine and bovine insulin
            - Due to the impurities and the pro-insulin forms
        - Demand from animals also not meeting supplies
    - Bacterial recombinant insulin **= Humulin** *(Gentech)*
        - Using E. coli made the A chain in one batch of bacteria and the B chain in another → purified → oxidised together
        - BUT, no secretory pathway for processing and no S-S capacity in the cytosol
    - Yeast (eukaryotic) recombinant insulin *(Novo Nodisk)*
        - Use S. cerevisae for its secretory pathway, BUT it lacks processing enzymes for removing the C-peptide
        - The signal sequence is cleaved, this molecule is then secreted
        - Instead of the C-peptide, is has 2x arginine and a lysine residue → cleaved
            - First arginine is cleaved by **trypsin** AND then the Arg + Lys on the *B chain* is cleaved by **carboxypeptidase**
- Erythropoietin (EPO)
    - Features
        - Glycoprotein hormone involved in haematopoiesis
        - Synthesised in the kidney → stimulates RBC production and maturation in bone marrow
        - Used as a treatment for anaemia (in all KD/dialysis pt's, pt's on chemo and with HIV)
    - Structure
        - Glycosylated protein → required for EPO action → binding to EPOR
            - Three N- and one O- glycosylation sites (40% of total mass)
        - Given glycosylation → requires mammalian cell culture
    - Glycosylation
        - Most proteins synthesised in the rER are glycosylated by the addition of a common N-linked oligosaccharide
        - Major synthetic function of the ER/golgi + most proteins made in the ER are glycosylated
        - ~50% of eukaryotic proteins are glycosylated, for their trafficking, folding, stability and function
        - ~700 mammalian enzymes are involved in various glycosylation patterns (glycans) on proteins
- Mammalian tissue culture
    - **Requires:** Dulbecco's modified Eagle's media (glucose, vitamins, salts, phenol red) + extra factors (foetal calf serum [growth hormones]) → grown at 37degress under 5% CO2
        - Production can be modified using serum free media and defined growth factors → $$
    - **Cells cultured:** primary and transformed (immortal) lines
        - Must use mammalian specific plasmids
- CHO (Chinese hamster ovary cells) cells in protein expression
    - Defector standard of good manufacturing practice (GMP)-certified production
    - Able to use serum-free media
    - Can produce 10g/L of protein
    - High-density large-scale fed-batch cultivations are developed and scale-up technology is well established
    - **CONS:** low doubling time (14-23h), low cell concentrations
- Atryn = recombinant anti-thrombin alpha = anti-coagulant
    - Expressed under the control of beta-casein promoter → transgenic expression in goat milk (easy)

Characteristics	Bacteria	Yeast	Mammalian
Cell growth	Rapid ~30minutes	Rapid ~90minutes	Slow ~24h
Cost	Low	Low	High
Expression	High	Medium-High	Low-Medium
Protein folding	Not reliable	Usually reliable	Very reliable
S-S	Poor	Good (secreted)	Good (secreted)
Glycosylation	None	Simple N-linked	Complex
Phosphorylation	No	Yes	Yes