EXPRESSION & PURIFICATION OF RECOMBINANT PROTEINS

E. coli is a popular and well understood system for heterologous protein expression.

Expression options:

- **Direct expression.** *E. coli* cytoplasm is a reducing environment - difficult to ensure proper disulphide bonds formation.
- **Fusion expression.** Ensures good translation initiation. Can overcome insolubility and/or instability problems with small peptides. Has purification advantages based on affinity chromatography.
- **Secretion** a fusion alternative when proteins are fused to peptides or proteins targetted for secretion. Periplasm offers a more oxidising environment, where proteins tend to fold better. Major drawbacks: limited capacity for secretion (0.1-0.2% total cell protein compared to 10% produced intracellularly) and **inability for posttranslational modifications** of proteins.

Disadvantages

- **Insolubility of heterologous proteins produced in *E.coli* - one of the main problems**
- **Inclusion bodies**. Dense particles, containing precipitated proteins. Their formation depends on protein synthesis rate, growth conditions. Advantages: proteolysis resistant, big yield, relatively pure, easy to separate. Disadvantages: inactive product requires in vitro refolding and renaturation
- **Refolding of recombinant proteins**

Solubilisation. High temperatures, detergents, high concentration of inorganic salts or organic solvents all used. The most commonly used organic solutes such as urea or guanidine-HCl often used in the presence of reducing agents (mercaptoethanol or DTT).Solubilised proteins purified by ion-exchange chromatography or other conventional methods, prior to refolding.

Refolding. If no S-S bonds present - remove denaturing agent to allow protein to fold correctly. If S-S bonds present - their formation can be accomplished: by air oxidation, catalysed by trace

metal ions; by a mixture of reduced and oxidised thiol compounds - oxidised DTT, reduced DTT; GSSG/GSH; cystine and cysteine, cystamine and cysteamine.

Isolation and characterisation of correctly folded proteins.

- Biological activity.
- Purity. SDS-PAGE.
- Chromatography - reversed phase or ion-exchange.
- N-terminus determination by sequencing. Peptide mapping.

YEAST SYSTEMS FOR HETEROLOGOUS EXPRESSION: *SACCHAROMYCES CEREVISIAE*

Eukaryote, unicellular, GRAS (**G**enerally **R**egarded **A**s **S**afe), capable of performing post-translational modifications. Excellent recombinant technology: vectors, markers, methods for transformation and gene manipulation, homologous recombination of cloned sequences by single cross over (insertion) and double cross over

Intracellular expression - higher protein yields, but more difficult extraction and purification. Additional potential problems with:

- co- and post-translational processing of proteins at N- and C-termini.
- proteolytic degradation
- addition of tags might result in aggregation and insolubility

Secretion

The yeast secretory pathway **is very similar to that in higher eukaryotes.** N-terminal signal sequences for co-translational translocation of screted proteins into the ER are removed by a signal peptidase. Examples of popular signal sequences used for secretion of heterologous proteins -these of Pho5, Suc2 and the a -factor.

Specific problems with secretion of heterologous proteins

Hyperglycosylation can inhibit reactivity with AB, or render proteins immunogenic (a problem for the production of therapeutic glycoproteins). The obvious solutions: glycosylation mutants

(*mnn1, mnn9*) or elimination of potential sites for glycosylation. Alternatively use other yeast species like *P. pastoris*. The cell wall permeability can be a limiting factor. Some cell wall mutants have higher cell wall porosity and release, as a result, heterologous proteins better.

Folding of secreted proteins in the ER and involves accessory proteins such as BiP (the product of *KAR2*), and PDI (protein disulphide isomerase). Overexpression of these genes has been beneficial in some cases.

Proteolytic processing could be limited by insufficient amounts of required processing enzymes, and in particular the products of *SEC11, KEX2, STE13* and *KEX1* in cases of multicopy expression of proteins. Again might need to overexpress some of these genes.

Modification by N-linked (to asparagine) and O-linked (to serine/threonine) glycosylation. Hyperglycosylation (outer chain extension) in the yeast Golgi is not typical of mammalian cells. Yeast proteins only modified by mannosylation (no other sugars).

MAMMALIAN CELL LINES EXPRESSION SYSTEMS

- Two modes of expression - transient and stable.
- Cell lines used. Three cell types are dominant in transient expression: human embryonic kidney (HEK), COS and baby hamster kidney (BHK), whilst CHO (Chinese hamster ovary) cells are used predominantly for stable expression.
- Mammalian expression vectors. Eukaryotic origin of replication is from an animal virus: e.g. Simian virus 40 (SV40). Popular markers for selection are the bacterial gene Neor (encodes neomycin phosphotransferase), which confers resistance to G418 (Geneticin), and the gene, encoding dihydropholate reductase (*DHFR*). When *DHFR* is used, the recipient cells must have a defective *DHFR* gene, which makes them unable to grow in the presence of methotrexate (MTX), unlike transfected cells with a functional *DHFR* gene. Promoter sequences that drive expression of both marker and cloned heterologous gene, and the transcription termination (polyadenilation signals) are usually from animal viruses (human CMV, SV40, herpes simplex virus) or mammalian genes (bovine growth hormone, thymidine kinase).

Advantages:

- There are no examples of higher eukaryotic proteins, which could not be made in detectable levels, and in a form identical to the natural host (that includes all types of post-translational modifications).

Disadvantages:

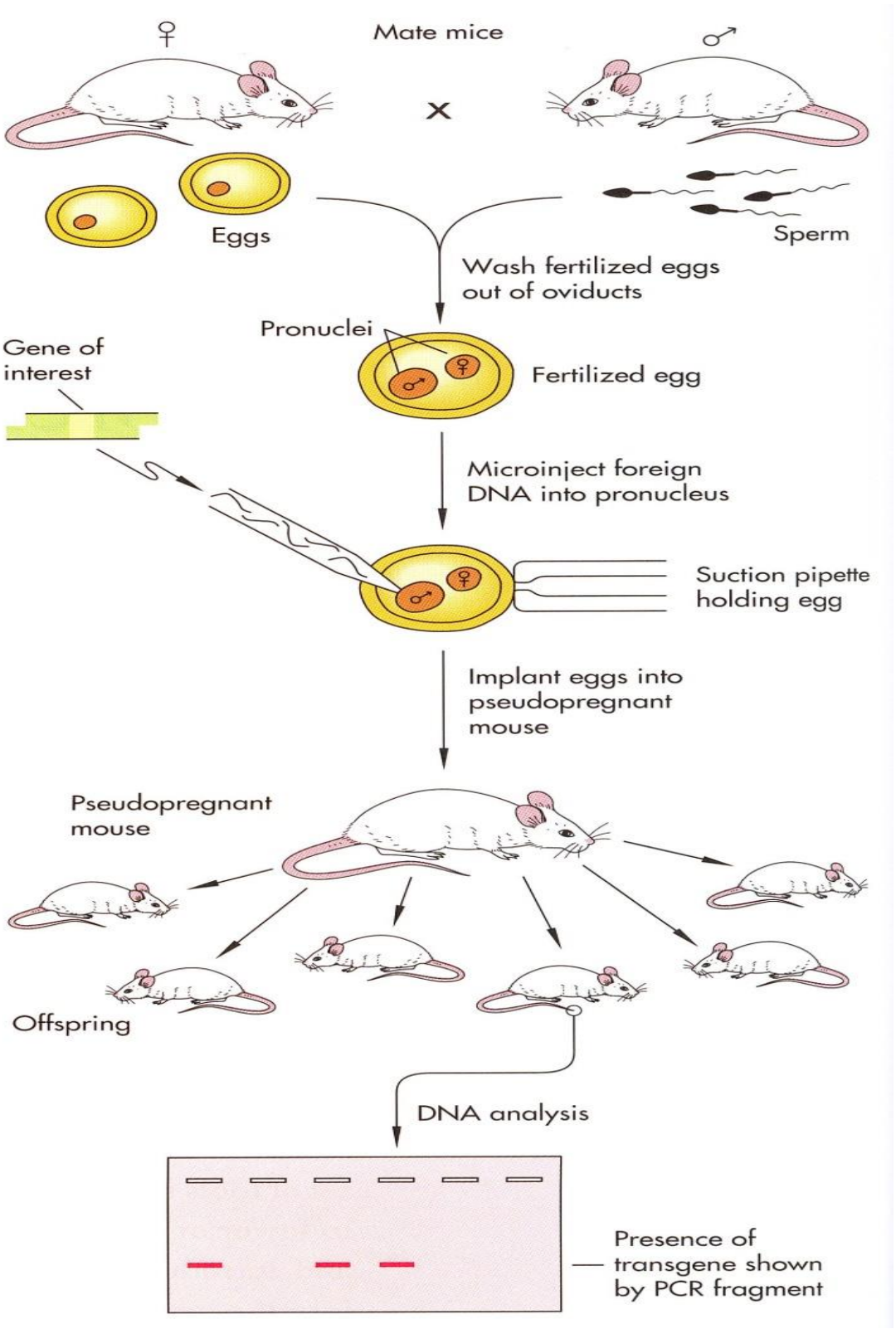
- Cultures characterised by lower cell densities and lower growth rates. Maintenance and growing very expensive. Gene manipulations are very difficult. Mammalian cells might contain oncogenes or viral DNA, so recombinant protein products must be tested more extensively.

TRANSGENICS

Transgenics are genetically modified organisms with DNA from another source inserted into their genomeA large number of transgenic animals have been created viz; Mice, Cows, Pigs Sheep, Goats, Fish, Frogs , Insects. Currently, no transgenic animal or animal product is approved by the FDA or USDA for human consumption

Some of the goals of transgenic animal creation are:

- Research into animal and human disease
- Improve livestock animals
- Use of animals as bioreactors


♀
Mate mice
♂
×
Eggs
Sperm
Wash fertilized eggs out of oviducts
Pronuclei
Gene of interest
Fertilized egg
Microinject foreign DNA into pronucleus
Suction pipette holding egg
Implant eggs into pseudopregnant mouse
Pseudopregnant mouse
Offspring
DNA analysis
Presence of transgene shown by PCR fragment

Transgenic cattle

Dairy cows carrying extra copies of two types of casein genes produce 13% more milk protein

Not only will this make the milk more nutritious, it would allow for less milk to make more cheese

Currently the milk from these animals is under FDA review

The important difference between this & other transgenics is that the DNA added is not foreign

Transgenic Fish

Tilapia

Salmon/trout

Catfish--Can grow up to 6 times faster than wildtype fish.Most have extra copies of growth hormone (GH) gene

AFPs(anti freeze proteins) lower the freezing temperature of blood & fluidsTrout normally do not survive in water below –0.6°C.Transgenic trout containing an AFP gene & promoter can survive in waters as cold as –1.2°C

Transgenic plants

Transgenic plants have genes inserted into them that are derived from another species. The inserted genes can come from species within the same kingdom (plant to plant) or between kingdoms (for example, bacteria to plant). In many cases the inserted DNA has to be modified slightly in order to correctly and efficiently expressin the host organism. Transgenic plants are used to express proteins like the cry toxins from *B. thuringiensis*, herbicide resistant genes, antibodies and antigensfor vaccinations. A study led by the European Food Safety Authority (EFSA) found also viral genes in transgenic plants.

The first genetically modified crop approved for sale in the U.S. was the *FlavrSavr* tomato, which had a longer shelf life. It is no longer on the market.In November 2014, the USDA approved a GM potato that prevents bruising.

In February 2015 Arctic Apples were approved by the USDA,[1] becoming the first genetically modified apple approved for US sale. Gene silencing was used to reduce the expression of polyphenol oxidase (PPO), thus preventing enzymatic browning of the fruit after it has been sliced open. The trait was added to Granny Smith and Golden Delicious varieties.[86][88]The trait includes a bacterial antibiotic resistance gene that provides resistance to the antibiotic kanamycin. The genetic engineering involved cultivation in the presence of kanamycin, which allowed only resistant cultivars to survive. Humans consuming apples do not acquire kanamycin resistance. The FDA approved the apples in March 2015.

Nutrition

Edible oils

Some GM soybeans offer improved oil profiles for processing or healthier eating. Camelina sativa has been modified to produce plants that accumulate high levels of oils similar to fish oils.

Vitamin enrichment

Golden rice, developed by the International Rice Research Institute (IRRI), provides greater amounts of Vitamin A targeted at reducing Vitamin A deficiency.

Researchers vitamin-enriched corn derived from South African white corn variety M37W, producing a 169-fold increase in Vitamin A, 6-fold increase in Vitamin C and doubled

concentrations of folate. Modified Cavendish bananas express 10-fold the amount of Vitamin A as unmodified varieties.

Toxin reduction

A genetically modified cassava under development offers lower cyanogen glucosides and enhanced protein and other nutrients (called BioCassava).

In November 2014, the USDA approved a potato, developed by J.R. Simplot Company, that prevents bruising and produces less acrylamide when fried. The modifications prevent natural, harmful proteins from being made via RNA interference. They do not employ genes from non-potato species. The trait was added to the Russet Burbank, Ranger Russet and Atlantic varieties.

Stress resistance

Plants engineered to tolerate non-biological stressors such as droughtfrost,[102][103] high soil salinityand nitrogen starvationwere in development. In 2011, Monsanto's DroughtGard maize became the first drought-resistant GM crop to receive US marketing approval.

Herbicides

Glyphosate

As of 1999 the most prevalent GM trait was glyphosate-resistance. Glyphosate, (the active ingredient in Roundup and other herbicide products) kills plants by interfering with the shikimate pathway in plants, which is essential for the synthesis of the aromatic amino acids phenylalanine, tyrosine and tryptophan. The shikimate pathway is not present in animals, which instead obtain aromatic amino acids from their diet. More specifically, glyphosate inhibits the enzyme 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS).

This trait was developed because the herbicides used on grain and grass crops at the time were highly toxic and not effective against narrow-leaved weeds. Thus, developing crops that could withstand spraying with glyphosate would both reduce environmental and health risks, and gives an agricultural edge to the farmer

Some micro-organisms have a version of EPSPS that is resistant to glyphosate inhibition. One of these was isolated from an *Agrobacterium* strain CP4 (CP4 EPSPS) that was resistant to glyphosate. The CP4 EPSPS gene was engineered for plant expression by fusing the 5' end of the

gene to a chloroplast transit peptide derived from the petunia EPSPS. This transit peptide was used because it had shown previously an ability to deliver bacterial EPSPS to the chloroplasts of other plants. This CP4 EPSPS gene was cloned and transfected into soybeans.

The plasmid used to move the gene into soybeans was PV-GMGTO4. It contained three bacterial genes, two CP4 EPSPS genes, and a gene encoding beta-glucuronidase (GUS) from *Escherichia coli* as a marker. The DNA was injected into the soybeans using the particle acceleration method. Soybean cultivar A54O3 was used for the transformation.

Pest resistance

Insects

Tobacco, corn, rice and many other crops have been engineered to express genes encoding for insecticidal proteins from Bacillus thuringiensis (Bt). Papaya, potatoes, and squash have been engineered to resist viral pathogens such ascucumber mosaic virus which, despite its name, infects a wide variety of plants. The introduction of Bt crops during the period between 1996 and 2005 has been estimated to have reduced the total volume of insecticide active ingredient use in the United States by over 100 thousand tons. This represents a 19.4% reduction in insecticide use.

In the late 1990s, a genetically modified potato that was resistant to the Colorado potato beetle was withdrawn because major buyers rejected it, fearing consumer opposition.

By-products

Drugs

In 2012, the FDA approved the first plant-produced pharmaceutical, a treatment for Gaucher's Disease. Tobacco plants have been modified to produce therapeutic antibodies

Biofuel

Algae is under development for use in biofuels. Modified jatropha offers improved qualities for fuel. Syngenta has USDA approval to market a maize trademarked Enogen that has been genetically modified to convert its starch to sugar for ethanol. In 2013, the Flemish Institute for Biotechnology was investigating poplar trees genetically engineered to contain less lignin to ease conversion into ethanol. Lignin is the critical limiting factor when using wood to make bio-

ethanol because lignin limits the accessibility of cellulose microfibrils to depolymerization by enzymes.

Bioremediation

Scientists at the University of York developed a weed (*Arabidopsis thaliana*) that contains genes from bacteria that can clean TNT and RDX-explosive soil contaminants. 16 million hectares in the USA (1.5% of the total surface) are estimated to be contaminated with TNT and RDX. However *A. thaliana* was not tough enough for use on military test grounds.

Genetically modified plants have been used for bioremediation of contaminated soils. Mercury, selenium and organic pollutants such as polychlorinated biphenyls (PCBs).

Marine environments are especially vulnerable since pollution such as oil spills are not containable. In addition to anthropogenic pollution, millions of tons of petroleum annually enter the marine environment from natural seepages. Despite its toxicity, a considerable fraction of petroleum oil entering marine systems is eliminated by the hydrocarbon-degrading activities of microbial communities. Particularly successful is a recently discovered group of specialists, the so-called hydrocarbonoclastic bacteria (HCCB) that may offer useful genes.

Example of metabolic engineering in plants

Transgenic carrots have been used to produce the drug Taliglucerase alfa which is used to treat Gaucher's disease. In the laboratory, transgenic plants have been modified to increase photosynthesis (currently about 2% at most plants versus the theoretic potential of 9–10%). This is possible by changing the rubisco enzyme (i.e. changing C3 plants into C4 plants), by placing the rubisco in acarboxysome, by adding CO_2 pumps in the cell wall, by changing the leaf form/size. Plants have been engineered to exhibit bioluminescence that may become a sustainable alternative to electric lighting.