Biochemistry Lab

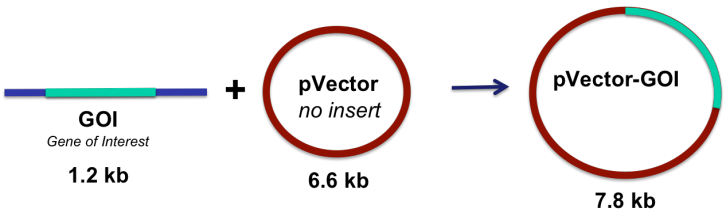
Systems for Recombinant Protein Expression
Lecture Notes Handout

Recombinant Protein -What is it?

- ✓ Expression of a gene from one organism expressed in another organism.
  - Heterologous Gene/Protein expression
- ✓ Often times but not always "tagged" with extra amino acids to identify the foreign protein and to assist in its purification
- ✓ Proteins used in research, clinical, and agriculture environments
- ✓ 100 μg of a purified expressed protein can sell for $500 to $2000.
- ✓ Insulin, interferon, theraputic antibodies are all examples

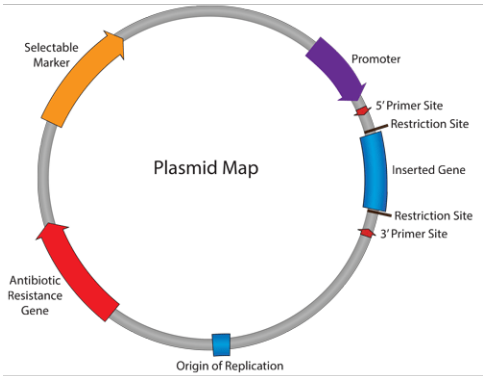
Creating a Recombinant Protein

- ➢ Recombine gene from one organism with the plasmid expression vector for another organism (host system)
  - ✓ i.e. cloned gene from human is cut into DNA to be transformed into bacteria



What is a plasmid

At their most basic level, plasmids are small circular pieces of DNA that replicate independently from the host's chromosomal DNA. They are mainly found in bacteria, but also exist naturally in archaea and eukaryotes such as yeast and plants. In nature, plasmids provide one or more functional benefits to the host such as resistance to antibiotics, degradative functions, and/or virulence. All natural plasmids contain an origin of replication (which controls the host range and copy number of the plasmid) and typically include a gene that is advantageous for survival, such as an antibiotic resistance gene. In contrast, plasmids utilized in the lab are usually artificial and designed to introduce foreign DNA into another cell. Minimally, lab-created plasmids have an origin of replication, selection marker, and cloning site. The ease of modifying plasmids and the ability of plasmids to self-replicate within a cell make them attractive tools for the life scientist or bioengineer. *(taken from plasmid 101)*

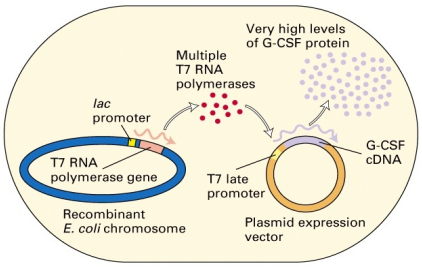


Critical Points for Expression Clone

- ➢ Framing – ensure that your gene's start site is in the correct reading frame
  - ✓ Many expression vectors have multiple variants for each of the three possible frames
- ➢ Restriction sites – blunt or sticky ends – PCR your needs into the clone, then digest.
- ➢ Ensure start site for plasmid is upstream of your gene
- ➢ If using a tag (more later) is it in frame and on the desired terminus of the protein

Plasmid Points

- ➢ ***<u>Important regions of the vector</u>***
  - ✓ Antibiotic resistance – selectable marker
    - For bacteria – to create more plasmid
    - For host system – if creating stable transfectants
  - ✓ Origin of Replication
    - Used by bacteria to make copies of plasmid
  - ✓ Second "helper" plasmid
  - ✓ For bacteria and T7
  - ✓ Typically carried in host competent cell
  - ✓ Used for high level of expression
  - ✓ Fusion tag
    - N or C terminus
    - Large or small size – does the tag alter the function of the protein
    - Purification or detection tag
  - ✓ Compartment sequences
    - Short sequences to send protein to membrane or organelle or even secreted into cell culture media
  - ✓ High or low expression – inducible or repressed
    Appropriate Termination of GOI
  - ✓ Proteolytic Cleavage site from fusion partner or tag



Biochemistry Lab

Systems for Recombinant Protein Expression

Lecture Notes Handout

Antibiotics

- Addition of a plasmid must include a second gene coding for a protein adding antibiotic resistance
  - ✓ Only cells with plasmid can grow on selective medium (in the presence of an antibiotic)
  - ✓ Maintains pressure on the bacteria to keep the plasmids
- In general, antibiotics are agents that kill bacteria or inhibit their growth.
  - ✓ Kanamycin – inhibits 30S ribosome-protein expression
  - ✓ Ampicillin – inhibits bacterial cell wall synthesis
  - ✓ Carbencillin – longer lasting derivative of Amp
  - ✓ Chloramphenicol – Stops 50S ribosome peptidyl translocation
- Used as frozen 500 – 1000X stocks.
- Additional gene in plasmid must code for protein which degrades drug
  - ✓ Correct Multiple Cloning Site (MCS)

The Host with the Most!

- Choice of host depends on several needs
  - ✓ Costs and availability of supports (kits, vectors, reagents and expertise)
  - ✓ Yield – large mass of protein presents specific challenges for scale up (bacteria/yeast vs. mammalian)
  - ✓ How will the protein be used after expression?
  - ✓ Toxicity of protein on host cell
  - ✓ Post-translational modification?
  - ✓ Impact on host cell on protein solubility and hydrolysis
  - ✓ Location of protein within the host cell – membrane, inclusion body, secretion into media
  - ✓ Complexity of medium

Expression Systems (host cells)

- Bacteria, Yeast, Insect Cells, Mammalian Cells Plant Cells – protoplasts, Cell free "test tube" lysate translation
- Animals / Eggs (antibodies, serum proteins)

Why Chose *E.coli an* Expression System?

- Advantages –
  - ✓ First choice due to level of support (technical, literature, kits, plasmids, reagents)…
  - ✓ High yield and easier scale up (up to 100 mg per liter of culture)
  - ✓ Cost of media and reagents is relatively inexpensive
  - ✓ Many expression plasmids – many with gene expression regulation
  - ✓ Fast growth condition – one to two days from starter culture to cell pellet
  - ✓ Amenable to shaker flasks or fermentation
  - ✓ Up to 50% of protein can be GOI!
- Disadvantages
  - ✓ No post-translational modification
    - Lipid modification, phosphorylation, glycosylation, decarboxylation, acetylation
    - These may be critical requirements for functional protein
  - ✓ Large proteins may be difficult to express (plasmid size is limited to about 15 kb)
  - ✓ Membrane proteins may not fold or isolate well
    - Lipid membrane of bacteria – mostly phosphatidylethanolamine, animal cells – phosphatidylecholine and other p-lipids
  - Large number of proteins are toxic to E.coli
    - May need to significantly alter plating and other expression issues
  - ✓ Inclusion body packing of protein
    - Insoluble, aggregated and non-native protein
    - Often times takes place in high expressed or toxic proteins
  - ✓ Codon Use differences, false stops and low expression

Bacterial Expression

- Basic Process of Expression
  - ✓ Clone heterologous gene into *E.coli* vector
  - ✓ Transform plasmid (with antibiotic resistance) into appropriate strain of bacteria
    - Low efficiency method (competent chemical transformation)– if plenty of plasmid DNA is present

- High efficiency method- (electroporation) if low abundant or problems with selection of plasmid
- Bacterial strain for DNA purification (DH5a, XL1-Blue or others) which are low in recombinases (RecA-) - **not** appropriate for protein expression – so you must keep two strains available, one for DNA maintenance and another for expression (more on expression later)

✓ Expand Single Colony from plate into shaking suspension culture
- Large clumps or lawns will allow cells which do not carry resistance and thus your plasmid with GOI – a great recipe for disaster

✓ Expand culture to low density
✓ Induce cells while early log phase
✓ Harvest cells before density reaches critical stage
✓ Lyse cells and purify protein – what could be easier!

E.coli Expression Strains

➢ Key choice when starting expression
➢ Characteristics to consider:
- ✓ Low protease activity (*clpC*, *clpP*, *lon*, and *ompT* genes)
- ✓ Lower RNAse activity – enhances protein expression
- ✓ Level of tRNA for rare codons
  - BL21 codon plus, Rosetta – contain additional plamids for tRNA
- ✓ Appropriateness for induction
  - IPTG (LacZ gene) – most will work but must have lac permease gene in cell strain (LacZY) to allow the compound to enter the cell
  - Tuner strains from Novagen – have a mutated lac permease to allow a better control of IPTG entry and titration of induction
  - Arabinose (P_{BAD} – operator) need to have arabinose transporter expressed in cells (BL21-A1, KS272 or LMG194)

➢ Key choice when starting expression
➢ Characteristics to consider:
- ✓ Cys-Cys disulfide bridges – multiple bridges require special strains which have mutations in thioredoxin reductase and glutathione reductase (involved with –SH redox) to enhance disulfide bond formation – Origami strain (also low in proteases)
- ✓ Toxicity of recombinant protein – will need to use C41(DE3) or C43(DE3) strains
  - Decreased cell death and allow membrane protein expression from a number of organisms
- ✓ Vector specific strains – provide "helper" genes
  - pLYSE pQE strains

IPTG_Induction

➢ Lac Operon – set of bacterial genes used by bacteria to transport and metabolize lactose when the sugar is present in media as a carbon source
➢ IPTG – mimics effect of lactose
- ✓ Is not metabolized and is longer lasting
- ✓ Provides strong induction of Lac Operon

➢ IPTG Induction Vectors/Promoters
- ✓ Replace the three genes with GOI

Expression Vectors for *E.coli*

➢ pET - Strong expression of GOI – up to 50% of total cell protein
- ✓ Dose-responsive promoter (IPTG) tunable strain may be used in concert
- ✓ Uses strong promoter of T7 RNA polymerase to drive expression of second promoter in the pET vector
- ✓ Expression of pET is indirectly tied to IPTG and MUST have either a second plasmid or a host cell already containing the T7 RNA pol gene (DE3)
- ✓ This provides a strong control of expression – not leaky

➢ pBAD
- ✓ Induction by addition of arabinose
- ✓ Sugar binds to AraC protein
- ✓ Complex binds and initiates RNA polymerase
- ✓ Permits a dose response

- ✓ Good choice for toxic proteins
- ✓ Control expression for insoluble – inclusion body formation

➢ pQE – Qiagen 6xHis tagged construct
- ✓ Low copy plasmid with a T5 promoter
- ✓ Two lac operon sequence for repressor binding
- ✓ These can be leaky promoters (as they are so strong) and thus inhibition may need pREP4 (a second plasmid coding for lac repressor protein)
- ✓ IPTG inducable
- ✓ Some allow two genes
  - Bis-cistronic

➢ pGEX– Amersham– GE HealthSciences
- ✓ IPTG inducable
- ✓ Fusion protein – Glutathione S Transferase (GST)
- ✓ Many different vector variations
- ✓ Large fusion protein sometimes allows for good expression
- ✓ Easy to purify
- ✓ GST may need to be removed

Inclusion Bodies

➢ Region of bacteria which can fill with insoluble protein
➢ Over expression of proteins or toxic proteins induce formation of inclusion bodies
- ✓ Aggregates may be mostly the expressed protein
  - disulfide bonds incorrectly formed as the protein is expressed at a high rate may be cause
- ✓ Isolation of inclusion bodies requires lysing bacteria but not inclusion bodies, several centrifugation steps followed by a difficult lysing of body

➢ Over expression of proteins or toxic proteins induce formation of inclusion bodies
- ✓ Denaturation and refolding of proteins can result in high yield of native protein
- ✓ Requires strong chaotropic detergents – urea, guanidine chloride
- ✓ Renaturation is an art form and no standard protocol exists
- ✓ Isolating non-functioning protein may work well for some uses – antigen injection for antibody production

Cloning & Transforming in Yeast Cells

Pichia Pastoris

➢ Yeast are single celled eukaryotes
➢ Behave like bacteria, but have key advantages of eukaryotes
➢ P. pastoris is a methylotrophic yeast that can use methanol as its sole carbon source (using alcohol oxidase)
➢ Has a very strong promoter for the alcohol oxidase (AOX) gene (~30% of protein produced when induced)

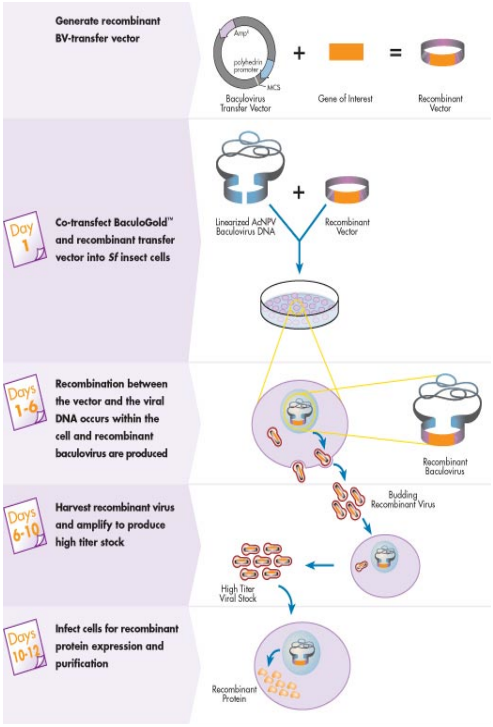
Yeast as an expression system

➢ *Pichia Pastoris* or *Saccharomyces Cervecisiae*
➢ Reasonable cost – media and scale up are straight forward
➢ Takes longer to prepare and screen for expression than bacteria
➢ Glycosylation is possible but not as complex as human cells –
- ✓ Modification can be heterogenious and interfere with purification

➢ Less support for expression than bacteria
➢ Pastoris expression induced by methanol
➢ Cloning in GOI is straight forward – many cDNA genes available

Expression in Baculovirus

➢ SF9 cells/Insect cells
- ✓ From the gut of moth
- ✓ Grown in suspension at mild temps
- ✓ Allow for scale up
- ✓ Often used for difficult to express clones in bacteria



- ✓ Also used for modified proteins
- ✓ Expression level is variable between infections

➢ Making the Virus
- ✓ Clone GOI into transfer plasmid vector
- ✓ Second DNA (linear) includes viral recombination region
- ✓ Both are infected into SF9 cells where they recombine and are packaged as viral particals
- ✓ Viral particals are tested for infection levels (MOI)
- ✓ Infect Sf-9 for protein expression

Types of Insect cell lines

cells	Doubling time	Cell appearance	Medium	Origin	Type of culture
Sf 9	72 hrs	Spherical, granular, regular in size, firm attachment to surface	TNM-FH	IPLBSF-21 cell lines of the fall army worm spodoptera frugiperda	Grow well as monolayer and suspension
Sf 21	24 hrs	Spherical, granular, different in size, firm attachment to surface	TNM-FH	IPLBSF-21 cell lines of the fall army worm spodoptera frugiperda	Grow well as monolayer and suspension
High-five	18 hrs	Spherical, granular, regular in size, loose attachment to surface	Express five SFM	Ovarian cells of cabbage looper	Grow well as monolayer, also as suspension

Insect Medium

- ➢ Grace's Insect medium- unsupplemented but contains L-glutamine
- ➢ Grace's Insect medium supplemented-contains additional TC yeastolate & Lactalbumin hydrolysate
- ➢ Trichoplusia ni Medium formulation hink (TNM-FH)- contains 10% FBS

Requirements for proper cell culture

- ➢ Temperature- Optimal range is 27-28 C
- ➢ pH- Optimal range is 6.1 to 6.4
- ➢ Aeration-Requires passive 02 diffusion for optimal growth & recombinant protein expression
- ➢ Osmolality- Optimum is 345-380 mOsm/kg
- ➢ FBS- Working with suspension culture it is advisable to use (10-20% FBS) to gave protection from cellular shear forces

Types of cell culturing

- ➢ Suspension culture
- ➢ Monolayer culture
  - ✓ Three methods to dislodge monolayers in adherent cell culture
    - Sloughing
    - -Trypsinization
    - -Tapping the layer until monolayer loosens

Expression in Baculovirus

- ➢ More time and labor intensive
- ➢ More costs – reasonable support
- ➢ Advantages:
  - ✓ Provides soluble post-translationally modified proteins

- ✓ Virus can handle large genes
- ✓ Works well for membrane or cytoplasmic proteins
- ✓ Can infect with multiple virus to express multi-subunit proteins
- ✓ High levels of expression possible 1-5 mg of protein / liter
- ✓ Scale up possible
- ✓ Low level of risk as virus is restricted to specific insect cell

Mammalian Cell Expression

- Much more expensive and time intensive than E. coli
- Much lower potential yield than yeast or bacterial expression

However:

- Most authentic secretion, glycosyation, phosphorylation and other post-translational modification
- Nearly the only way to fully express membrane proteins from mammalian cells
- Best structural and functional features and similar to cognate native forms
- Infection with virus or transient transfection
  - ✓ Selection must occur for producing large scale amounts of protein.

Scale-able

- Suspension cells include COS (monkey), CHO (hamster), HEK293 (Human)
  - ✓ Easy to transfect and select
  - ✓ Grow in relatively simple media and can be conditioned to large flasks
- Monolayer cells – fibroblasts but other cells will work

Scale-able

- Suspension - After stable transfection, can move to shaker flasks, wave flasks or bioreactors

Scale-able

- Monolayers – limited to roller bottles or cell farms

Potential Problems

- Cost and time – both are very high
  - ✓ Labor to create cell lines
  - ✓ Cost of infection or transfection
  - ✓ Often low yield
- Over expression of protein can be toxic to cell
- Interference with native processes or proteins

Biochemistry Lab

Systems for Recombinant Protein Expression

Lecture Notes Handout