ABRF 2009

Protein Expression Research Group

Recombinant Protein Laboratory

February 6-7, 2009

St Jude Children's Research Hospital

Introduction:
Basics of Protein Expression

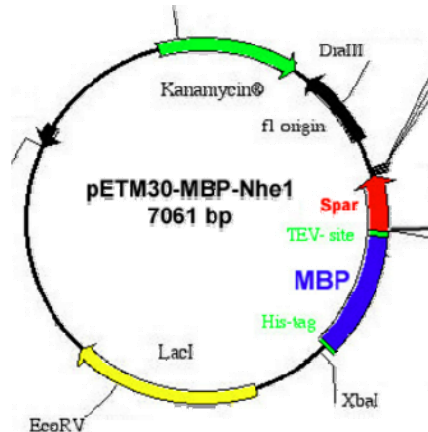
- **A. Choice of a Host System**
  1. Bacteria, yeast, insect and mammalian culture
  2. Equipment needed for core lab

- **B. Biology of Heterologous Protein Overexpression in Bacteria**
  1. Overview of the pET system
  2. The biggest failure in bacteria: insoluble protein
  3. Tips for shifting the balance to soluble expression

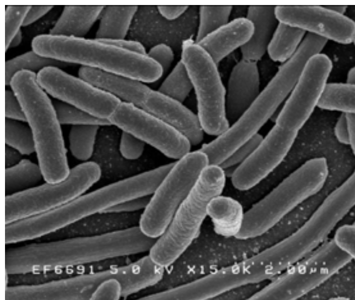
- **C. Details of Lab Work**
  1. Growth of *E. coli* in shake flasks
  2. Introduction to 10 L fermentors

Recombinant Expression: What do you need?

1. DNA of target of interest



2. Expression vector for host system
   - promoter
   - tags/fusions



3. Host cells

4. Cell growth system

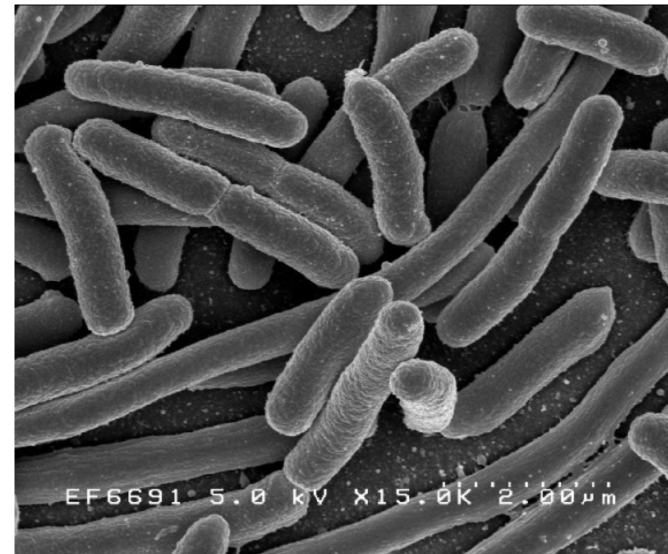
R.Page, ABRF 2007

Host Systems – *E. coli*

E. coli

Fast, cheap, "easy" (when it works)

- Easy cloning and genetic manipulation
- Inexpensive to culture
- Rapid growth, fast expression

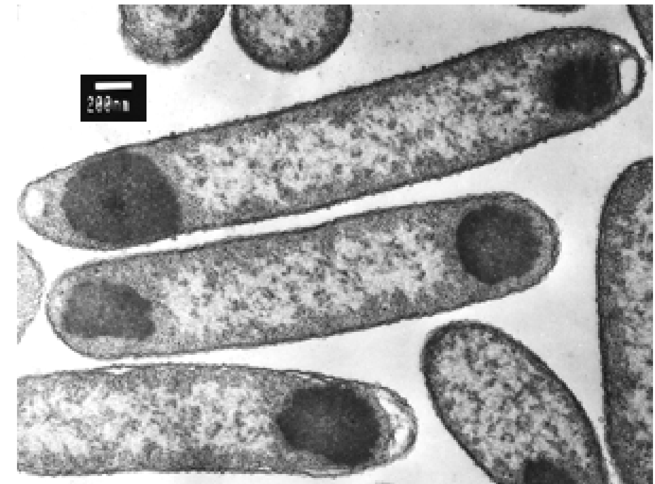


E. coli – Benefits

- Flexibility
  - Multiple promoters, tags, fusion partners, cleavage sites
- Generally works well for intracellular (reduced) proteins
- Strains available for oxidation of secreted proteins
- Can optimize for soluble exp. versus inclusion bodies
- Can optimize for high yield (shake flask or fermentation)
- Isotopic and seleno-Met labeling is straight forward
- Miniaturizable and scalable
  - 1-4 ml expression trials generally predictive of larger scale

E. coli – Disadvantages

- No eukaryotic post-translation modifications
  - glycosylation, phosphorylation, etc.
- Deficient in certain tRNAs common to eukaryotic genes
  - These “rare” codons can severely limit expression level
  - Can be corrected with strains supplemented w/ rare tRNAs
- More difficult for secreted / oxidized proteins
- Certain proteins can not fold and form inclusion bodies
  - Finding refolding conditions for insoluble proteins can be challenging
  - BUT when it works, this can offer very high yields



Host Systems - Yeast

Yeast – ***Pichia pastoris*** **or** ***Saccharomyces Cerevisiae***

Not quite as fast or easy, but still cheap

- Cloning in E. coli, then transfect DNA into yeast and screen for stable integration into yeast chromosome
- Time required for selection of high expressing clones
- Once established, rapid growth in shake flasks or fermentation

Yeast – Benefits and Issues

- Provides eukaryotic secretion and glycosylation
  - Glycosylation is less complex than mammalian, high mannose
- Requires use of yeast secretion signal peptides

Pichia pastoris

- Often used for secreted proteins / difficult to "crack"
- Promoter: α mating factor (MeOH induced), others

Saccharomyces Cerevisiae

- Used for both intracellular and secreted proteins
- Promoters: Gal1-10 promoter, others

Host Systems – Baculovirus / Insect cells

More time & labor intensive than *E. coli,* but…

Provides soluble expression for Many proteins that do not express well (soluble) in E. coli

- Cloning in E. coli, transfect insect cells to prepare virus, infect insect cells for expression
- Post-trasnslational modifications closer to mammalian
  - Secretion, glycosylation and phosphorylation
- Works well for secreted, membrane, and intracellular proteins

Baculovirus / Insect cells

- Making the virus: Two major systems in use
  - both start with cloning in *E. coli*
  1. Transfer vector method: co-transfect expression plasmid with 2nd plasmid containing required viral genes, co-transfected cells produce virus, which is then amplified
  2. Bac-to-bac: transform E. coli containing bacmid DNA, target gene transposed into bacmid, prep bacmid DNA and directly infect insect cells which produce virus, amplify virus
- 3 cell types generally used: Sf9, Sf21, High-Five
- Late-very promoter (p10 + polyhedrin) drives expression in last stages of virus cycle
- Infection cycle is lytic – recombinant protein is maximally expressed in last phases before cell lysis

Host Systems – Mammalian Cells

More time & labor intensive than *E. coli,* but…

Provides most authentic secretion, glycosylation, phosphorylation, and other post-translational modification

- Cloning in *E. coli*, then direct transfection of cells
- Transient transfection provides rapid (~5 days) batch expression
- Selection of stable clones (expression construct integrated into chromosome) can take months but provides higher yield, reproducible production, and scalabilty

Host Systems – Mammalian Cells

- Most often used for secreted or membrane proteins
  - Biomass / volume is low, so yields of intracellular proteins are low compared to other systems
- Several promoters available:
  - Constitutive: CMV, SV40, human EF-1, human UbC, others…
  - Inducible: tTA/Tet, GLVP/TAXI (GAL4-PR-LBD), GAL4-E1b, VP16-EcR-GR (VgEcR), others…
- Several cell types commonly used:
  - COS (monkey), CHO (hamster), HEK293 (human)

Newer Expression Systems

Insect Cells: Direct Transfection

- Similar to direct transfection of mammalian cells
- Can used for either transient transfection or stable integration
- Uses insect rather than viral promoters
  - e.g. hr5/IE1, COPIA 5 LTR, hsp70
- Used with Sf9, Sf21, Hi5, or drosophila S2 cells

Novagen’s InsectDirect system:

http://www.emdbiosciences.com/html/NVG/insectdirectindex.html

Invitrogen’s InsectSelect system:

http://www.invitrogen.com/content/sfs/manuals/insectselpiz_man.pdf

Newer Expression Systems

"BacMam"

- Dead-end "transduction" of mammalian cells by baculovirus
  - Virus enters cells but does not replicate, does not lyse cells
- Can be used for transient expression,
  - Small scale testing, or over-expression in "assay-ready" plates
- OR can be used to select for stable integration

- Advantages:
  - For screening, maintain just one cell line, multiple virus stocks
  - Can transduce some cells that are difficult to transfect
- Disadvantages:
  - Requires high MOI, thus large amounts of virus
  - Has not worked well in suspension so limited protein yield

Newer Expression Systems

Mammalian Cells: Adeno- and Lenti-viruses

- Analogous to baculovirus, but both adenovirus and lentivirus truly infect most mammalian cell lines (replicative)
- Can be used for transient transfection or stable integration

- Advantages:
  - Can be used to infect most mammalian cells
- Disadvantages:
  - Because viruses infect human cells, must be used in ≥ BSL-2

Invitrogen's ViralPower™ Lentivirus Expression System

http://www.invitrogen.com/content/sfs/manuals/virapower_lentiviral_system_man.pdf

TaKaRa's Adenovirus Expression kit:

http://bio.takara.co.jp/BIO_EN/Catalog_d.asp?C_ID=C1274

Newer Expression Systems

In vitro transcription / translation

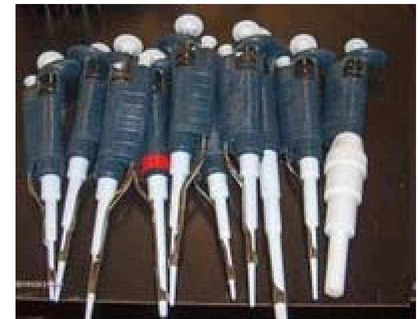
- “Host-less” – completely cell free
- Several lysate source systems commercially available
  - E. coli, yeast, wheat germ, insect cell, rabbit reticulocyte
- Degree of eukaryotic-like post-translational modification depends on system / lysate used

- Advantages:
  - Small scale, parallel, rapid testing
  - Toxic proteins tolerated (no living cells to kill)
  - Some systems readily amenable to isotopic labeling
- Disadvantages
  - Ease of use / low yields / high scale-up costs have been barriers, *but this it changing…*

Equipment Needed for a Core Lab

- Basic
- Cloning
- Expression
- Purification

Equipment Needed for a Core Lab



Basic Equipment:

- Ultra pure water system
- Pipetors, vortex, microcentrifuges
- Tabletop centrifuges (refrigerated preferred)
- Preparative centrifuge (e.g. RC5 or Avanti J series)
- Benchtop platform shaker
- Autoclave
- Buffer prep: pH meter, stirrers, etc.

Equipment Needed for - Cloning

IF you will clone your own expression vectors:

Basic:

- Agarose (slab) and acrylamide gel aparatus, power supplies
- Heat blocks, water bath, ice buckets / labtop coolers
- DNA purification kits (miniprep, PCR clean-up, gel extraction)
- Vacuum pump or vacuum supply, with trap
- 37º incubator (plates), 37º shaker-incubator (tubes and flasks)

Instruments:

- PCR machine
- Spectrophotometer
- Access to DNA sequencing

Equipment Needed for - Expression

E. coli Expression

- 37º incubators (plates and tubes / flasks)
- Refrigerated shaker-incubator (tubes and flasks)
- Baffled flasks for cell growth
- Fermentor - optional

Optional Small Scale Parallel Optimization

- Basic 96- or 24-well plate capabilities
- Multi-channel pipetors
- Rapid / short stroke shaker-incubator
- Plate vacuum manifold

Equipment Needed for - Expression

Baculovirus / Insect cell or Mammalian

Cell culture facility:

- Laminar flow tissue culture hood
- Incubators for dishes, plates, T-flasks
- Microscopes
- Incubators for shake flasks or spinner flasks
  - 27º C for insect cells
  - 37º C with controlled CO_2 for mammalian
- Optional Wave or traditional bioreactors

Optional Small Scale Parallel Optimization

- Basic 24-well plate capabilities, plate vacuum manifold
- Plate shaker in appropriate incubator

Equipment Needed for - Purification

- High speed centrifuge (harvesting cells, clearing lysate)
- Cell disruption: sonicator, French press, microfluidizer, or N_2 cavitation bomb
- SDS-PAGE & Western Blot aparatus & power supplies
- UV-VIS Spectrometer
- Chromatography Equipment:

Fancy:
Integrated Workstation
e.g. Akta Explorer,
Waters, BioRad, …

Not so Fancy:
Peristaltic pumps
Gradient mixer
UV detector (flow cell)
Fraction collector

- **A. Choice of a Host System**
  1. Bacteria, yeast, insect and mammalian culture
  2. Equipment needed for core lab

- **B. Biology of Heterologous Protein Overexpression in Bacteria**
  1. Overview of the pET system
  2. The biggest failure in bacteria: insoluble protein
  3. Tips for shifting the balance to soluble expression

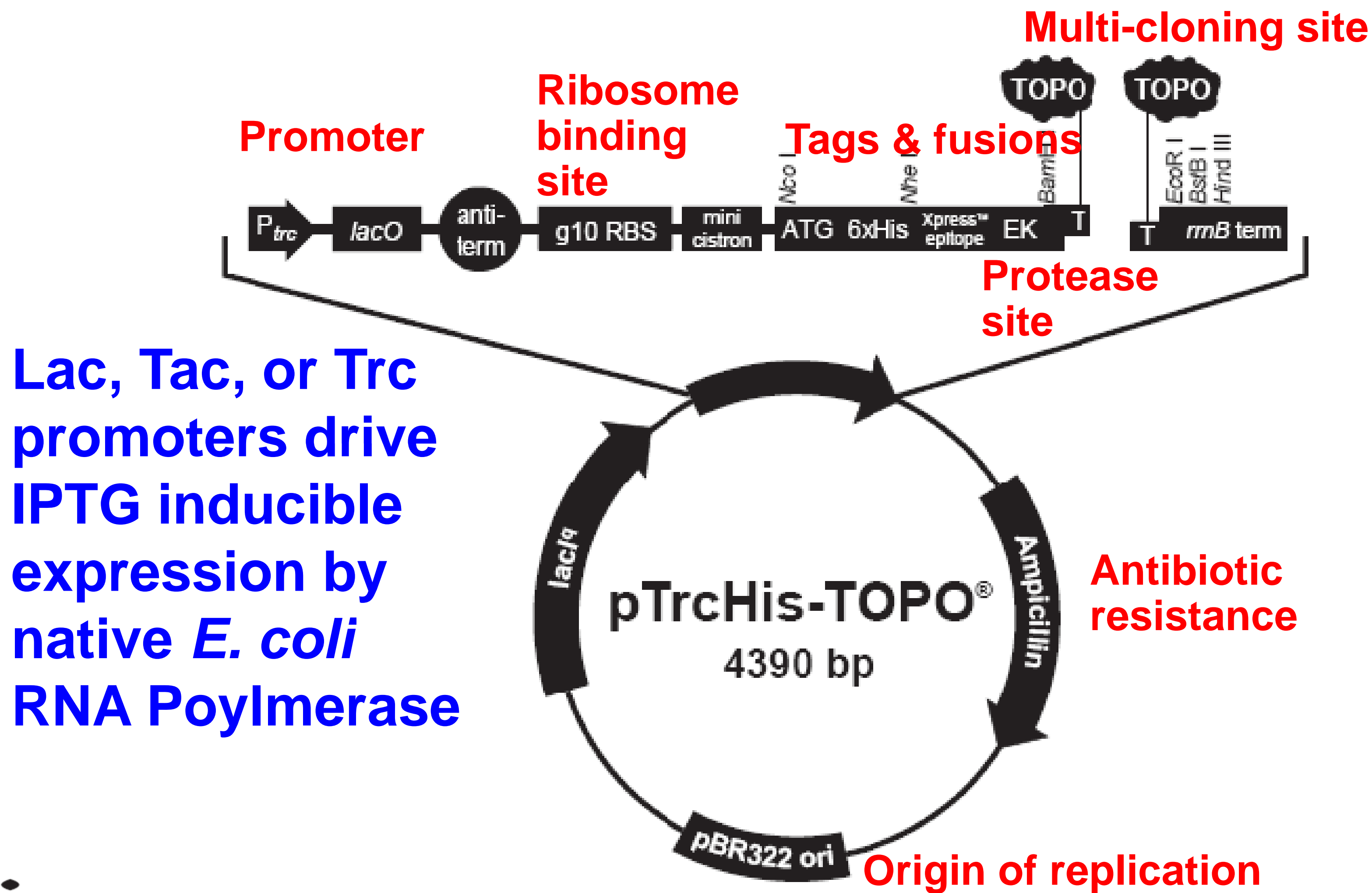
- **C. Details of Lab Work**
  1. Growth of *E. coli* in shake flasks
  2. Introduction to 10 L fermentors

E. coli Expression
Anatomy of an Expression Vector

Critical Vector Elements:

- Origin of replication
- Antibiotic resistance
- Promoter
- Ribosome binding site
- Multi-cloning site
- Various tags & fusion partners
- Protease cleavage site

E. coli Expression
Anatomy of an Expression Vector



Lac, Tac, or Trc promoters drive IPTG inducible expression by native *E. coli* RNA Poylmerase

Invitrogen™ life technologies

Overview of pET System

Critical Vector Elements:

Origin of replication = Ori f1

Antibiotic = Amp or Kan

Promoter = T7-lac

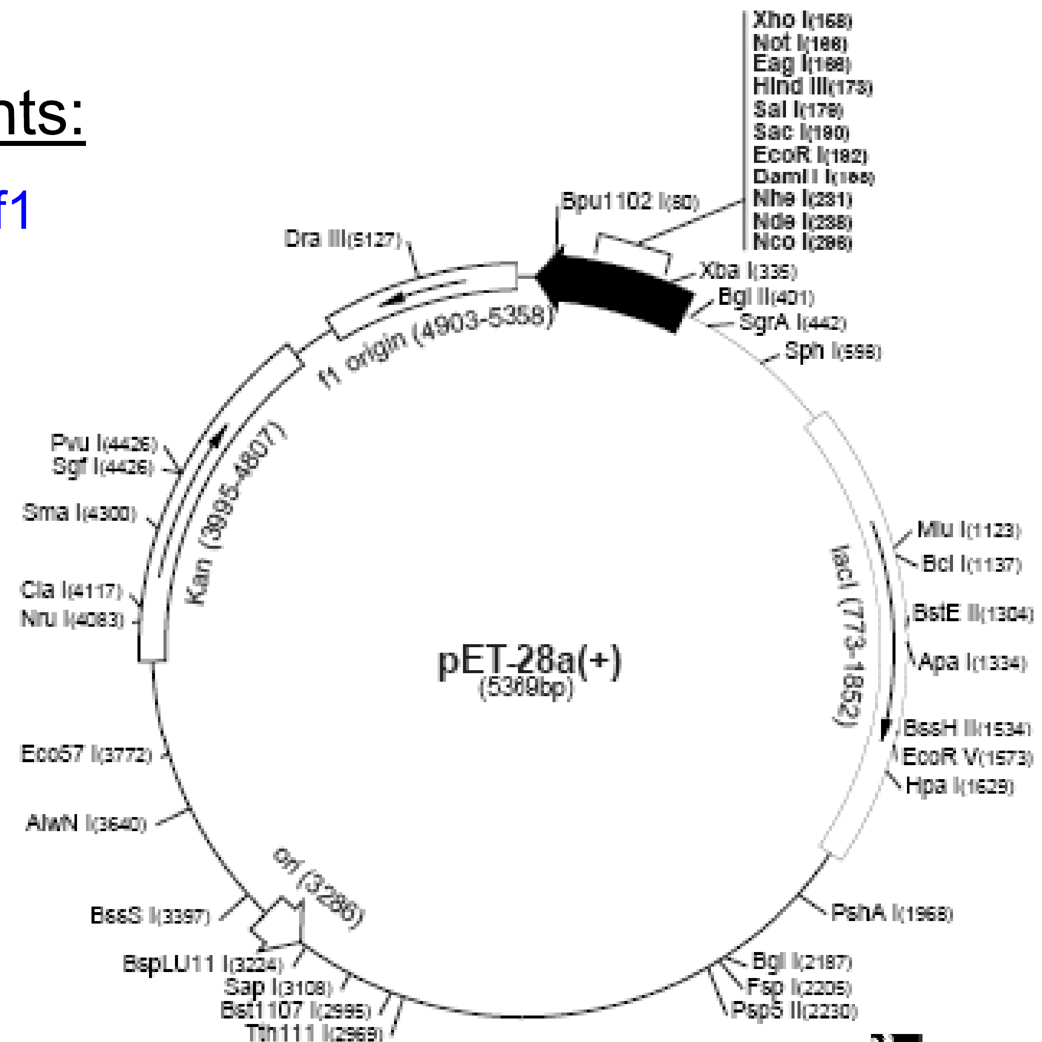
Ribosome binding site

Multi-cloning site

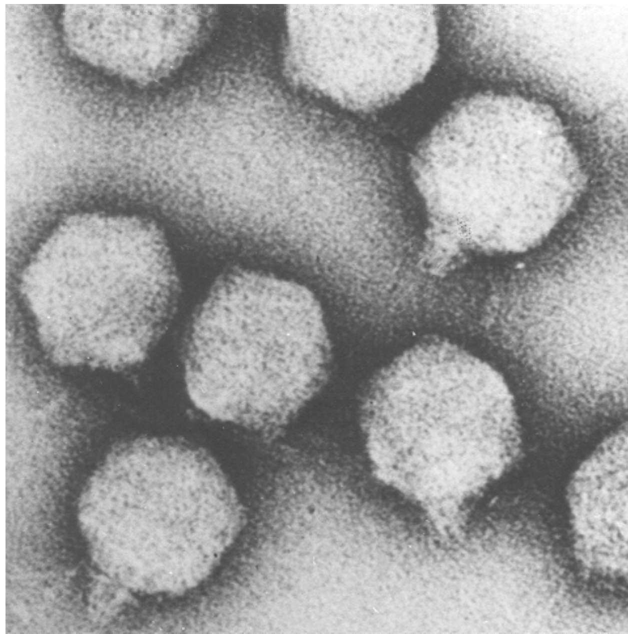
Various tags &

fusion partners

Various protease sites



T7 Bacteriophage and Polymerase



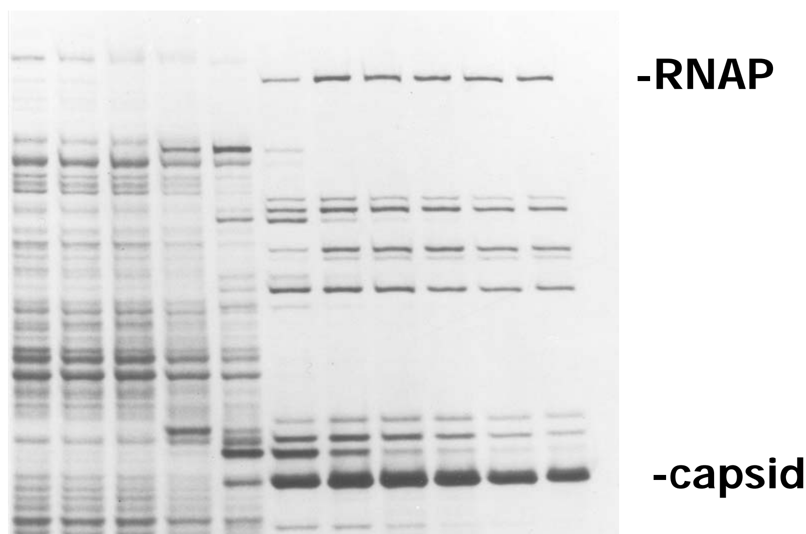
Electron micrograph of T7 phage particles

Courtesy of Bill Studier, *BNL*

- **Bacteriophage T7 infects bacteria like a virus infects cells**
- **T7 takes over the *E. coli* cell upon infection**
  - T7 DNA is translated to make T7 proteins, including T7 polymerase
  - T7 proteins direct host resources to production of new T7 phage
  - **Can produce 250 new T7 phage particles in 13 min at 37°C!**

T7 Bacteriophage and Polymerase

Time course of protein synthesis during T7 infection



Pulse labeling with [^{35}S]methionine

- **T7 directs all gene expression to its own DNA**
- **T7 RNA polymerase highly processive**
- **T7 mRNAs have strong upstream translation signals**

Courtesy of Bill Studier, *BNL*

Using T7 Polymerase for Recombinant Expression

pET system harnesses the power of T7 polymerase to drive expression of your protein

- Host E. coli cells contain T7 polymerase gene under IPTG-inducible Lac promoter (DE3 lysogen)
- Your gene of interest is under control of T7/lac promoter
- IPTG triggers expression of T7 Polymerase which drives overexpression of your protein

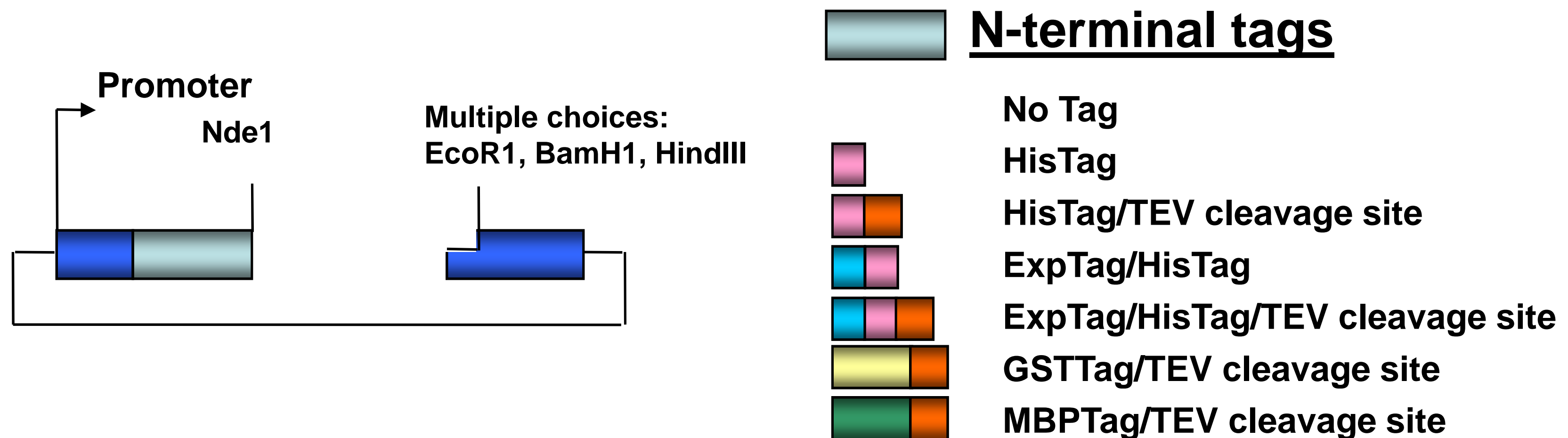
- Target mRNA can accumulate to levels that saturate ribosomes
- Low basal/high induced expression; tolerates toxic target proteins
- Target protein can accumulate to more than half of total cell protein

Novagen, pET System Manual, 11th Ed.
http://www.emdbiosciences.com/docs/docs/PROT/TB055.pdf

Commercial vs. Custom Vectors

Flexible and Customizable Vector Design

- Several tags, fusions proteins, cleavage sites available
- Or build your own custom set with your favorites
- Keep a consistent multi-cloning site for parallel cloning



A word about cleavage sites / enzymes…

Protease Cleavage Enzymes / Sites

Included in vectors to allow removal of tags / fusion partners

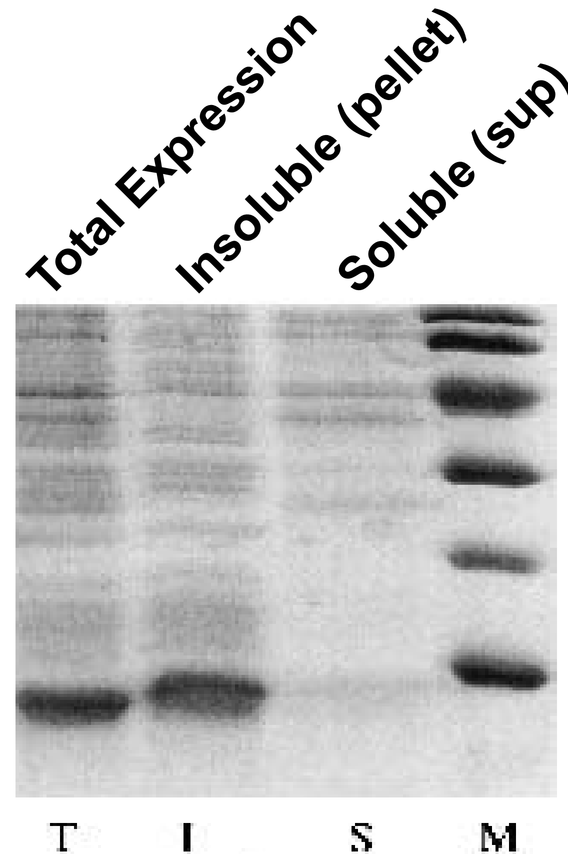
Enzyme	Recognition Site	Comments
Thrombin	LVPR/GS	Active, less specific
Factor Xa	IEGR/	Less specific
Enterokinase	DDDDK/	Very specific Leaves native N-term!
PreScission™	LEVLFQ/GP	Very specific
TEV	ENLYFQ/G(S)	Very specific
TVMV	ETVRFQ/G(S)	Very specific
Ubq/SUMO Pro.	Ubq/SUMO fusion	

Ian Hunt, *Protein Exp. & Purif.*, **40:** 1–22 (2005)

Biggest Pitfall: Insoluble Protein

The Problem:

Overexpression of some (many) proteins produces insoluble protein, sometimes in inclusion bodies.



From: M.Hammarström et al., Protein Sci., **11**: 313-321 (2002).

Biggest Pitfall: Insoluble Protein

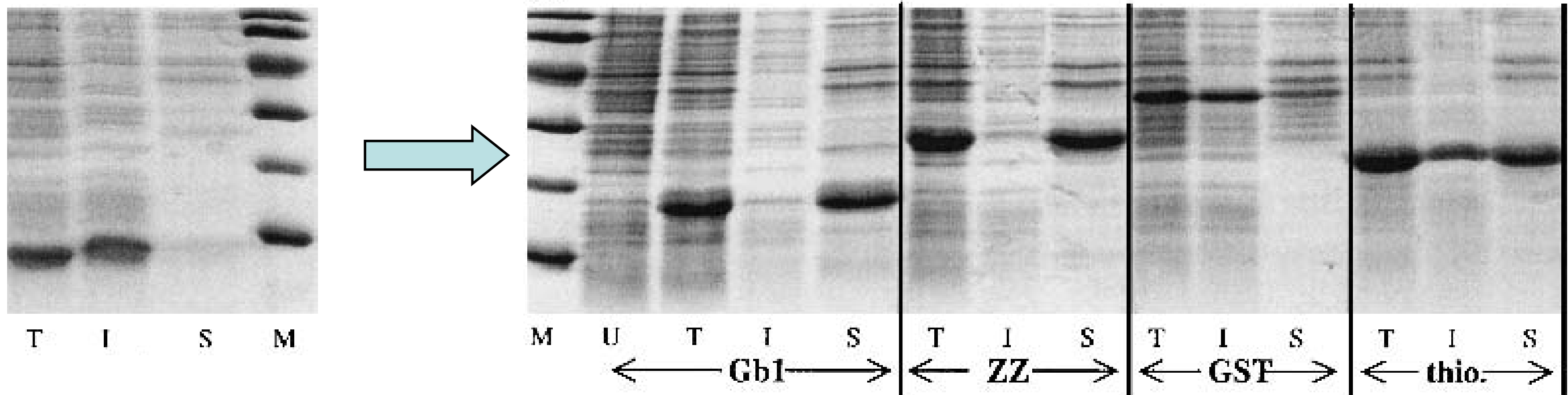
Possible Solutions:

1) Fusion Proteins
2) Expression Temperature
3) Cell Lines
4) Chaperone Co-expression
5) Co-expression / Purification with Co-factors
6) Lysis Method
7) Lower IPTG concentration (slow expression)

Insoluble Protein – Possible Solutions

1) Fusion proteins

In-line fusion of a soluble protein can sometimes help solubilize your protein.



From: M.Hammarström et al., "Rapid screening for improved solubility of small human proteins produced as fusion proteins in Escherichia coli," Protein Sci., **11**: 313-321 (2002).

Insolable Protein – Possible Solutions

2) Expression Temperature

- Reducing the temperature (and lengthening time) of induction can improve protein solubility
- Test several conditions:
  - 37º C for 2-4 hrs
  - 30º C for 4-6 hrs
  - 20º C for 16 hrs (O/N)
- To go even lower, try Stratagene's ArcticExpress cells
  - express cold shock proteins, can be induced at 10-13º C

Insoluble Protein – Possible Solutions

3) Cell Lines

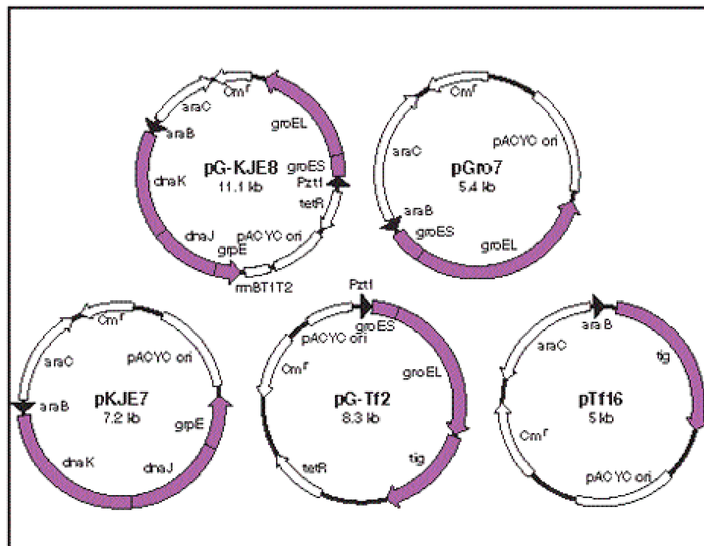
CELL LINE	COMMENTS
General Expression	
BL21	Deficient in ompT and lon proteases
BL21 DE3	T7-based expression
Expression Problems	
RIL/RP	Codon supplements (high AT content/High GC content, respectively)
RILP/Rosetta	Codon supplements (Codons for both high AT and GC content)
Solubility Problems	
Origami	Enhance disulfide bond formation (thx and glut. Reductase mutants)
Tuner	Can finely tune expression using IPTG (mutation in *lacZY*)
Arctic Express	Express Cpn60/Cpn10, cold-adapted chaperones (10 ºC expression)
Labeling	
B834	Selenomethionine labeling for crystallography (*Met* auxotroph)
Membrane/Toxic Prot.	
pLysS	Reduce basal expression by expression of lysozyme
C43	Facilitates soluble expression of toxic and integral membrane proteins

Insolube Protein – Possible Solutions

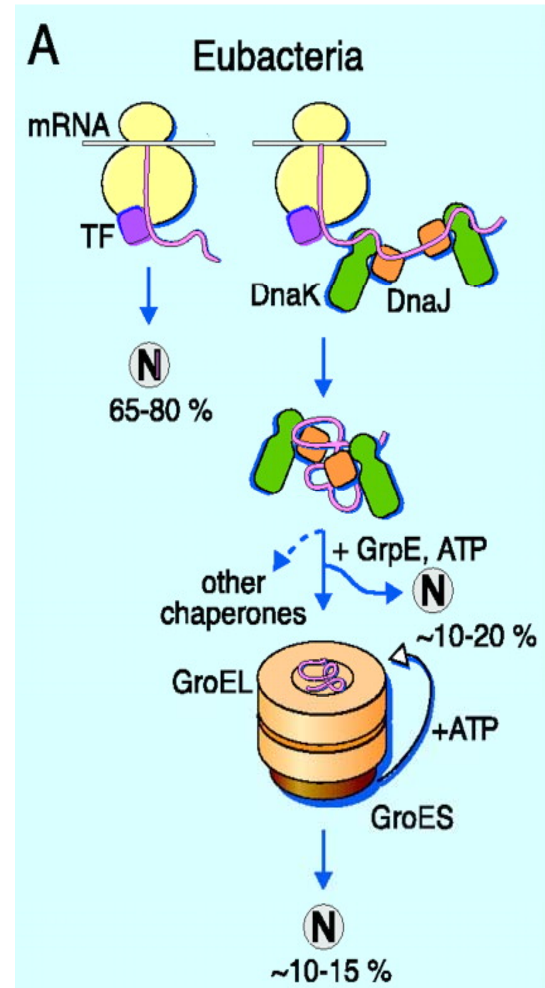
4) Chaperone Co-expression

Co-expression with chaperones can improve protein solubility

- Trigger Factor, DNaK/DNaJ/GrpE, GroEL/GroES (Ta Ka Ra)
- Cpn60/10 (Stratagene)



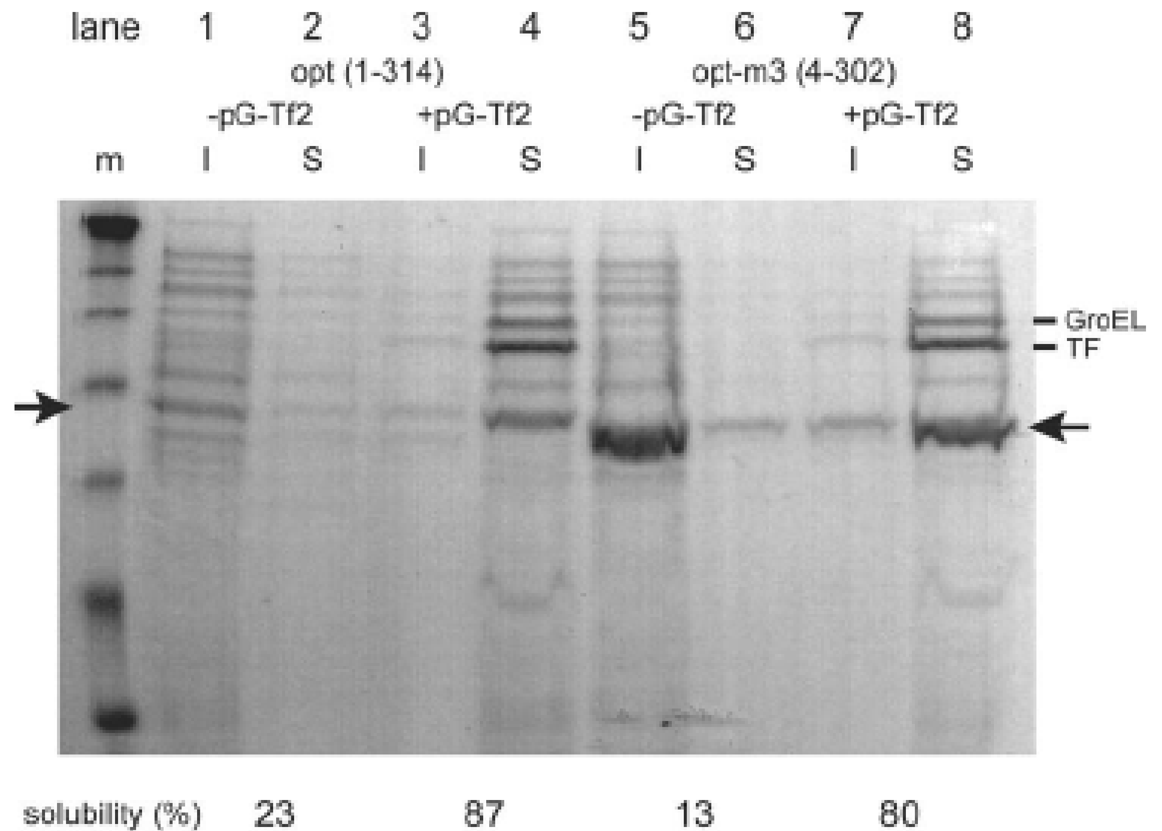
Maps of Takara's Chaperone Plasmids



Insoluble Protein – Possible Solutions

4) Chaperone Co-expression

Co-expression with GroEL and TF greatly improved soluble expression of heIF2α



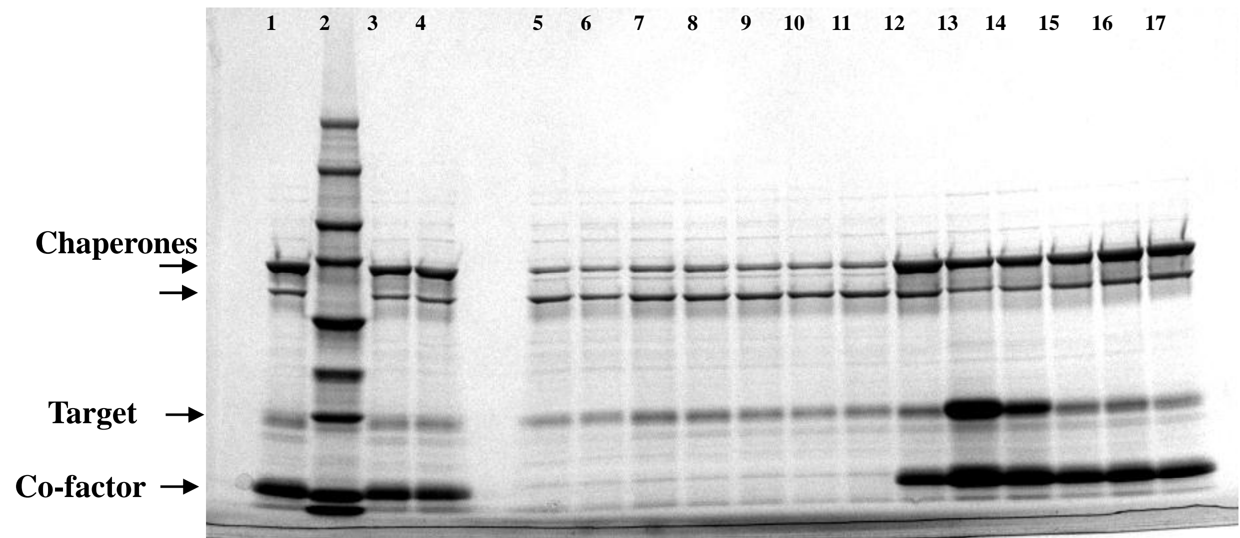
T.Ito, G.Wagner, Journal of Biomolecular NMR **28:** 357–367 (2004)

Insoluble Protein – Possible Solutions

5) Co-expression / Purification with Co-factors

- Other proteins that bind target protein
- Small molecule ligands or inhibitors
- Co-expressing a phosphatase to keep a kinase soluble

Co-expression with both a small protein co-factor AND an appropriate small molecule ligand gives improved soluble expression



Insoluble Protein – Possible Solutions

6) Lysis Method

Lysis Method	COMMENTS
Freeze-Thaw; Enzymatic	Methods can be used separately; usually more effective when used in combination with one another
	Can become expensive for large culture volumes (lysozyme/DNaseI)
	Gentle; but incomplete lysis can be a problem
Sonication	Effective method for cell lysis
	Generates a large amount of heat in small area and can result in denaturation and precipitation of an otherwise soluble protein
French Press	Gentle lysis, but slow for even small (30 ml) volumes
Homogenization	External enzymes not required
	Wide range of volumes (10 – 1000s mls)
	Fast, 3L / hour
	30,000 PSI so suitable for *E. coli* and yeast lysis

Rebecca Page, Brown Univ.

Insoluble Protein – Possible Solutions

7) Lower IPTG concentration

This example combines several stategies:

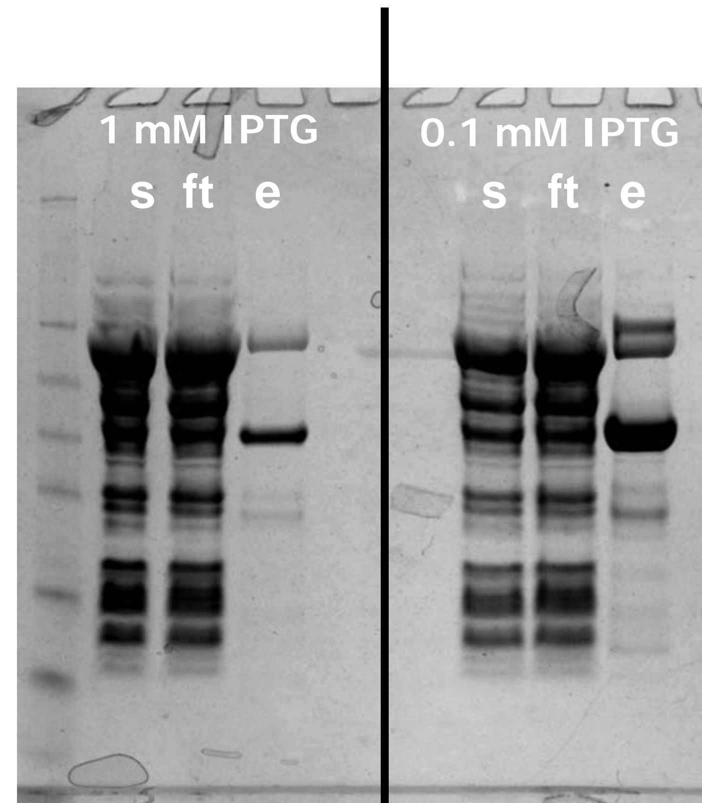
1) Lower Temp

(grow @ 24 ºC, induce @10 ºC)

2) Chaperone co-expression

(Cold shock protein Cpn60)

3) Lower IPTG concentration



Result: More soluble, active PP1 with 0.1 mM IPTG than 1 mM

Overcoming Insoluble Protein

Possible Solutions for *E. coli*:

1) Fusion Proteins
2) Expression Temperature
3) Cell Lines
4) Chaperone Co-expression
5) Co-expression / Purification with Co-factors
6) Lysis Method
7) Lower IPTG concentration (slow expression)
8) Others: Protein construct, codon usage, etc…

- **A. Choice of a Host System**
  1. Bacteria, yeast, insect and mammalian culture
  2. Equipment needed for core lab

- **B. Biology of Heterologous Protein Overexpression in Bacteria**
  1. Overview of the pET system
  2. The biggest failure in bacteria: insoluble protein
  3. Tips for shifting the balance to soluble expression

- **C. Details of Lab Work**
  1. Growth of *E. coli* in shake flasks
  2. Introduction to 10 L fermentors