BROWN

BROWN

Strategies for Maximizing Heterologous Protein Expression in *E. coli* with Minimal Cost

Rebecca Page

Department of Molecular Biology, Biochemistry and Cell Biology

March 31, 2007

BROWN

The Challenge

Getting soluble, active protein from *E. coli*.

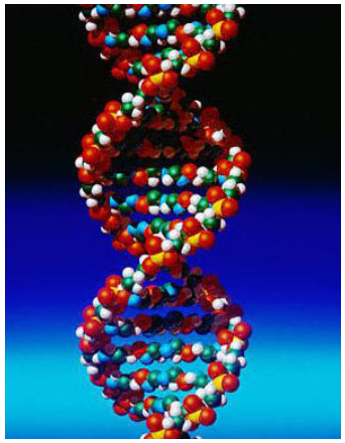
Overview

- **Introduction recombinant protein expression in *E. coli:*** its benefits and disadvantages
- **Protein Sequence:** predictions for soluble expression
- **Strains and tags used for heterologous protein expression**: what to use when
- **Expression optimization:** how to maximize your protein yield
- **Conclusions**

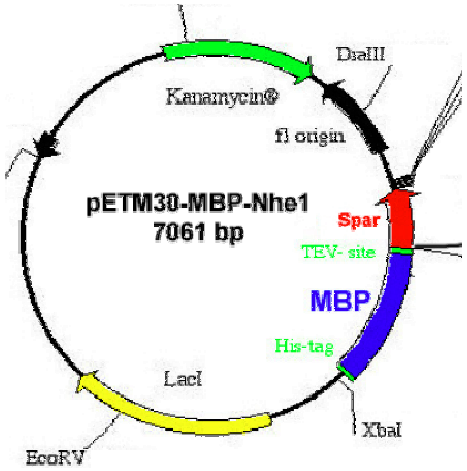
Overview

- **Introduction recombinant protein expression in *E. coli:*** its benefits and disadvantages
- **Protein Sequence:** predictions for soluble expression
- **Strains and tags used for heterologous protein expression**: what to use when
- **Expression optimization:** how to maximize your protein yield
- **Conclusions**

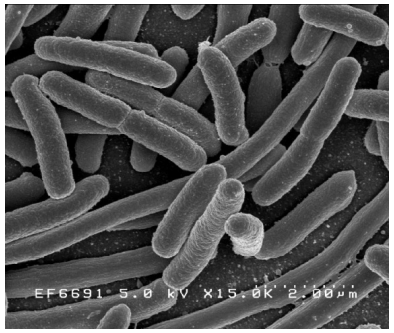
What is Needed?



1. DNA of target of interest



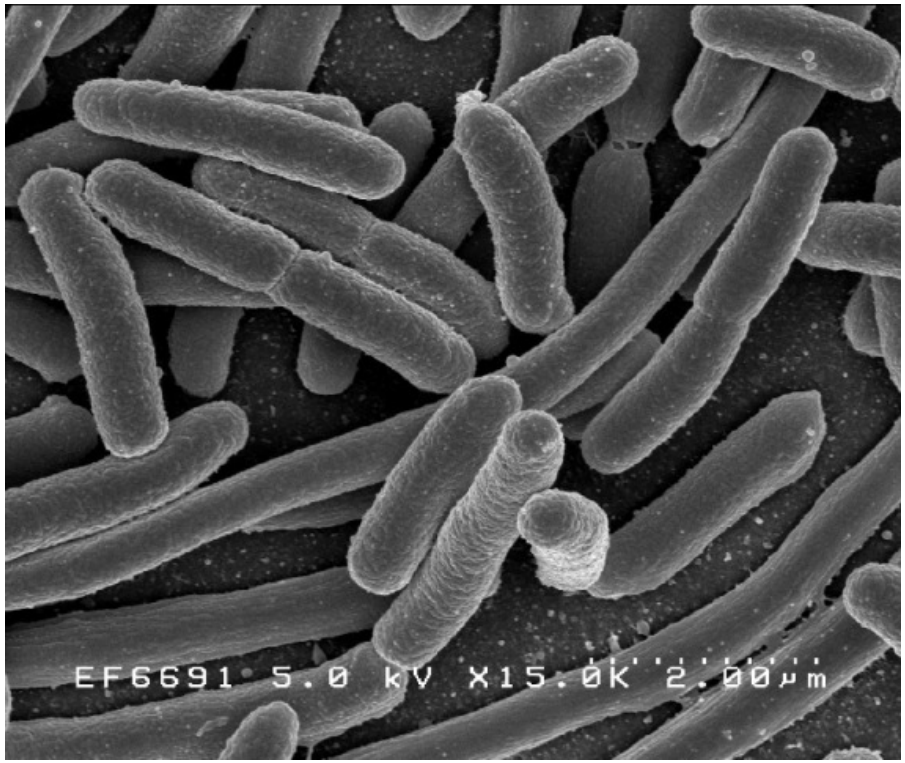
2. *E. coli* expression vector



3. *E. coli*

4. Shaker

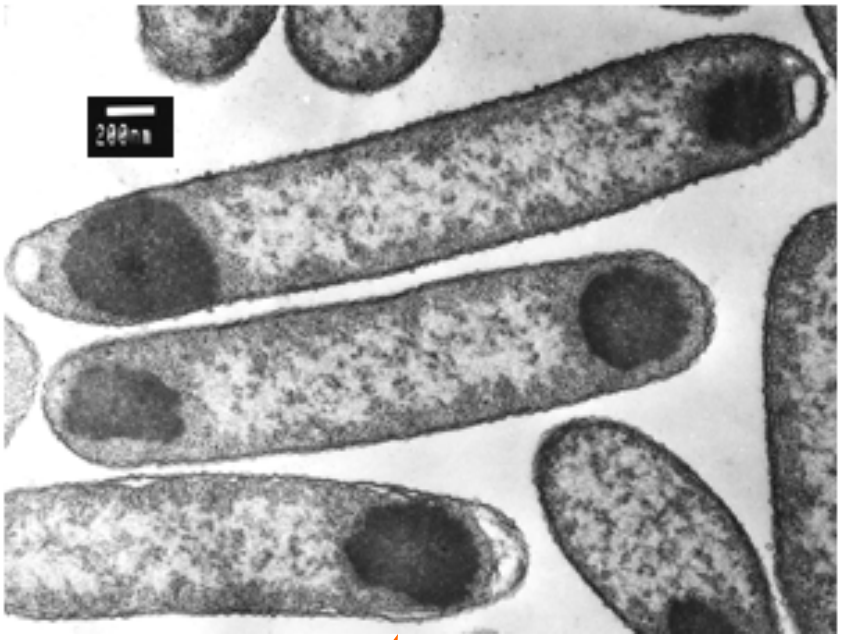
Benefits of *E. coli* expression



Credit: Rocky Mountain Laboratories, NIAID, NIH

- Easy to genetically manipulate
- Relatively inexpensive to culture
- Expression is fast
- Well-established labeling protocols for structural studies (stable Isotope; selenomethionine

Disadvantages



Inclusion Bodies

- **Incapable of producing eukaryotic post-translation modifications**
  - **glycosylation**
  - **phosphorylation**
- **Certain proteins can not fold and form inclusion bodies**
- **Deficient in certain tRNAs commonly found in eukaryotic genes**

Overview

- **Introduction recombinant protein expression in *E. coli:*** its benefits and disadvantages
- **Protein Sequence:** predictions for soluble expression
- **Strains and tags used for heterologous protein expression**: what to use when
- **Expression optimization:** how to maximize your protein yield
- **Conclusions**

Sequence Features Correlate with Expression

Sequence Features	Expression
Length/MW	✓ ✓
pI (% charged residues)	No ✓
GRAVY index (protein hydrophobicity)	No ✓
Contiguous Hydrophobic Residues	✓ ✓

- **Only small (23kD average) proteins would express solubly in the *absence* of a solubility enhancing tag**
- **MBP and Thx were the best N-terminal fusion tags**
- **MBP was the best C-terminal fusion tag**

Dyson, *et al* (2004) BMC Biotech

Canaves, Page, *et al* (2004) JMB

Sequence Features Correlate with Expression

Sequence Features	Expression
Length/MW	✓ ✓
pI (% charged residues)	No ✓
GRAVY index (protein hydrophobicity)	No ✓
Contiguous Hydrophobic Residues	✓ ✓

- **Only small (23kD average) proteins would express solubly in the *absence* of a solubility enhancing tag**
- **MBP and Thx were the best N-terminal fusion tags**
- **MBP was the best C-terminal fusion tag**
- **Design 1-3 construct/protein to maximize the likelihood of soluble protein expression**

Dyson, *et al* (2004) BMC Biotech

Canaves, Page, *et al* (2004) JMB

Overview

- **Introduction recombinant protein expression in *E. coli:*** its benefits and disadvantages
- **Protein Sequence:** predictions for soluble expression
- **Strains and tags used for heterologous protein expression**: solubility tags, protein targeting and vector design
- **Expression optimization:** how to maximize your protein yield
- **Conclusions**

Protein Fusion Tags

- **Facilitate Expression**
- **Facilitate Purification**

Common Protein Fusion Tags

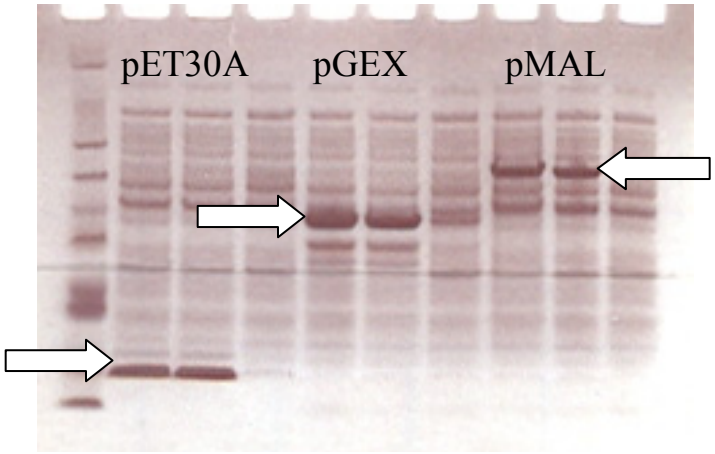
TAG	RES	SIZE (Kd)	USE	MATRIX	COMMENTS
$Thio_6$	6	0.5	E	N/A	Small; facilitates expression
Polyhistidine	2-10	~0.2-1	P	IMAC	Small; inexpensive
MBP	396	40	P; S	Amylose	↑ solubility; periplasm
GST	211	26	P; S	Glutathione	Dimerization
Intein CBD	51	5.6	P	Chitin	Purif. in absence of thiol reducing agents
Strep-tag II	8	1.2	P	Biotin	Downstream detection
NusA	495	54.9	S	N/A	↑ solubility
Trx	109	11.7	S	N/A	↑ Disulfide bond*
DsbA	208	23.1	S	N/A	↑ Disulfide bond*; periplasm
Ubiquitin (Ub)	128	14.7	S	N/A	↑ solubility
SUMO	101	11.6	S	N/A	↑ solubility
GFP	239	27	D	N/A	Visualization of *in vivo* soluble expression
Z-tag	91	10.6	S	Protein A	Solubility enhancing tag
GB1	85	9.7	S	IgG	NMR testing *prior* to tag removal
pSET	~40	4.5	S	N/A	↑ solubility by ↑ negative charge of tag

Which works best?

TAG	RES	SIZE (Kd)	USE	MATRIX	COMMENTS
$Thio_6$	6	0.5	E	N/A	Small; facilitates expression
Polyhistidine	2-10	~0.2-1	P	IMAC	Small; inexpensive
MBP	396	40	P; S	Amylose	↑ solubility; periplasm
GST	211	26	P; S	Glutathione	Dimerization
Intein CBD	51	5.6	P	Chitin	Purif. in absence of thiol reducing agents
Strep-tag II	8	1.2	P	Biotin	Downstream detection
NusA	495	54.9	S	N/A	↑ solubility
Trx	109	11.7	S	N/A	↑ Disulfide bond*
DsbA	208	23.1	S	N/A	↑ Disulfide bond*; periplasm
Ubiquitin (Ub)	128	14.7	S	N/A	↑ solubility
SUMO	101	11.6	S	N/A	↑ solubility
GFP	239	27	D	N/A	Visualization of *in vivo* soluble expression
Z-tag	91	10.6	S	Protein A	Solubility enhancing tag
GB1	85	9.7	S	IgG	NMR testing *prior* to tag removal
pSET	~40	4.5	S	N/A	↑ solubility by ↑ negative charge of tag

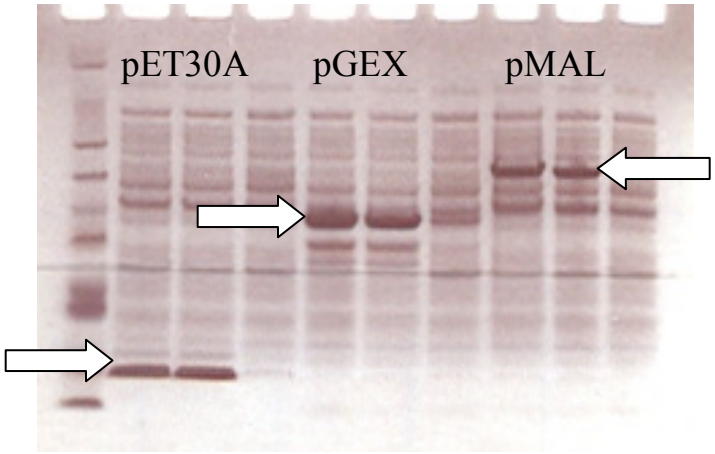
Solubility Tags : Case Study

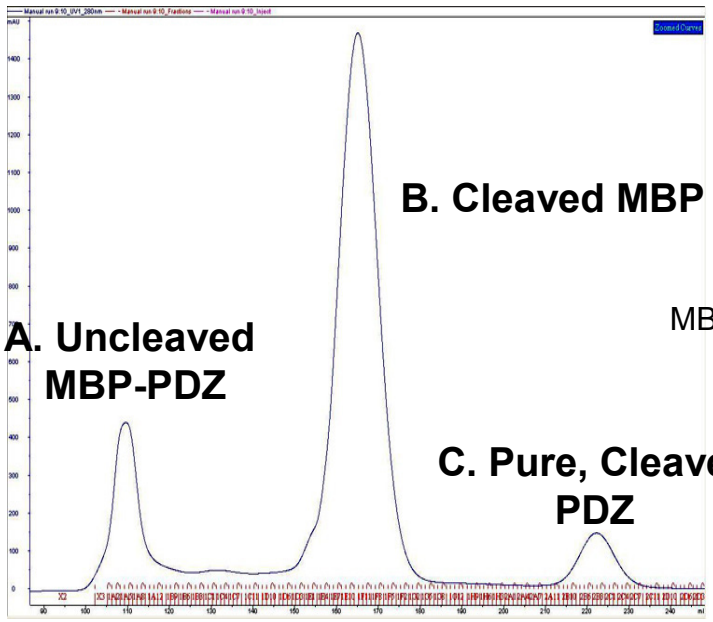
- **PDZ domain**
- **Sub-cloned various constructs into 5 vectors:**
  - **Untagged**
  - **6thio_6his**
  - **6his_Ztag**
  - **6his_GST**
  - **6his_MBP**
- **Expression outstanding in all cases**

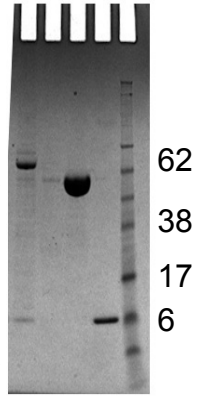


Solubility Tags : Case Study

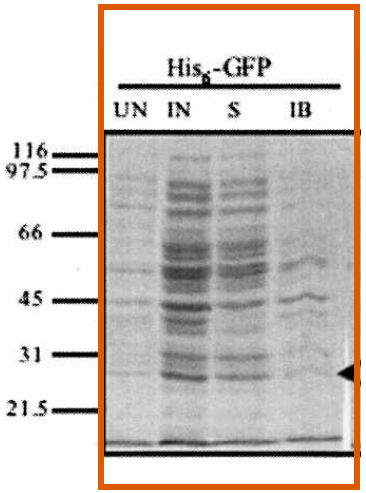
- **PDZ domain**
- **Sub-cloned various constructs into 5 vectors:**
  - **Untagged**
  - **6thio_6his**
  - **6his_Ztag**
  - **6his_GST**
  - **6his_MBP**
- **Expression outstanding in all cases**
- **Soluble only when fused to MBP**







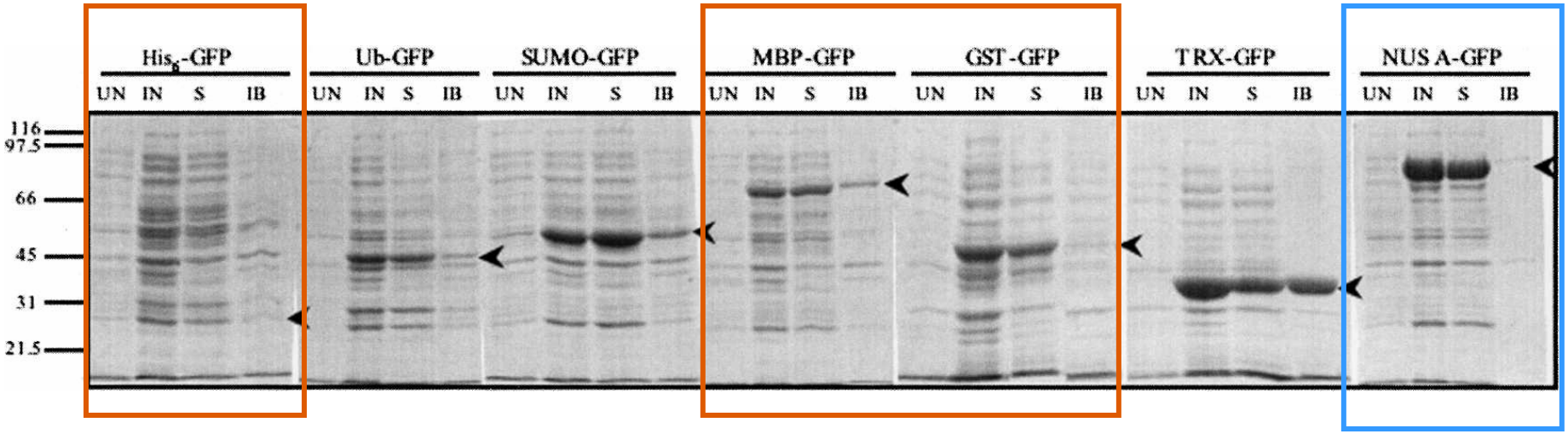
Solubility Tags Work



- **NusA, N-utilyzing substance A protien**
  - discovered through a systematic search for *E. coli* proteins that have the highest potential for solubility when overexpressed.
- **Trx, thioredoxin**
  - appears to catalyze the formation of disulfide bonds in the cytoplasm of *trxB* mutants
- **GST, glutathione-S-transferase**
- **MBP, maltose-binding-protein**
- **UB, ubiquitin**
- **SUMO, small ubiquitin-like modifier**

Marblestone, *et al* (2006) Protein Science

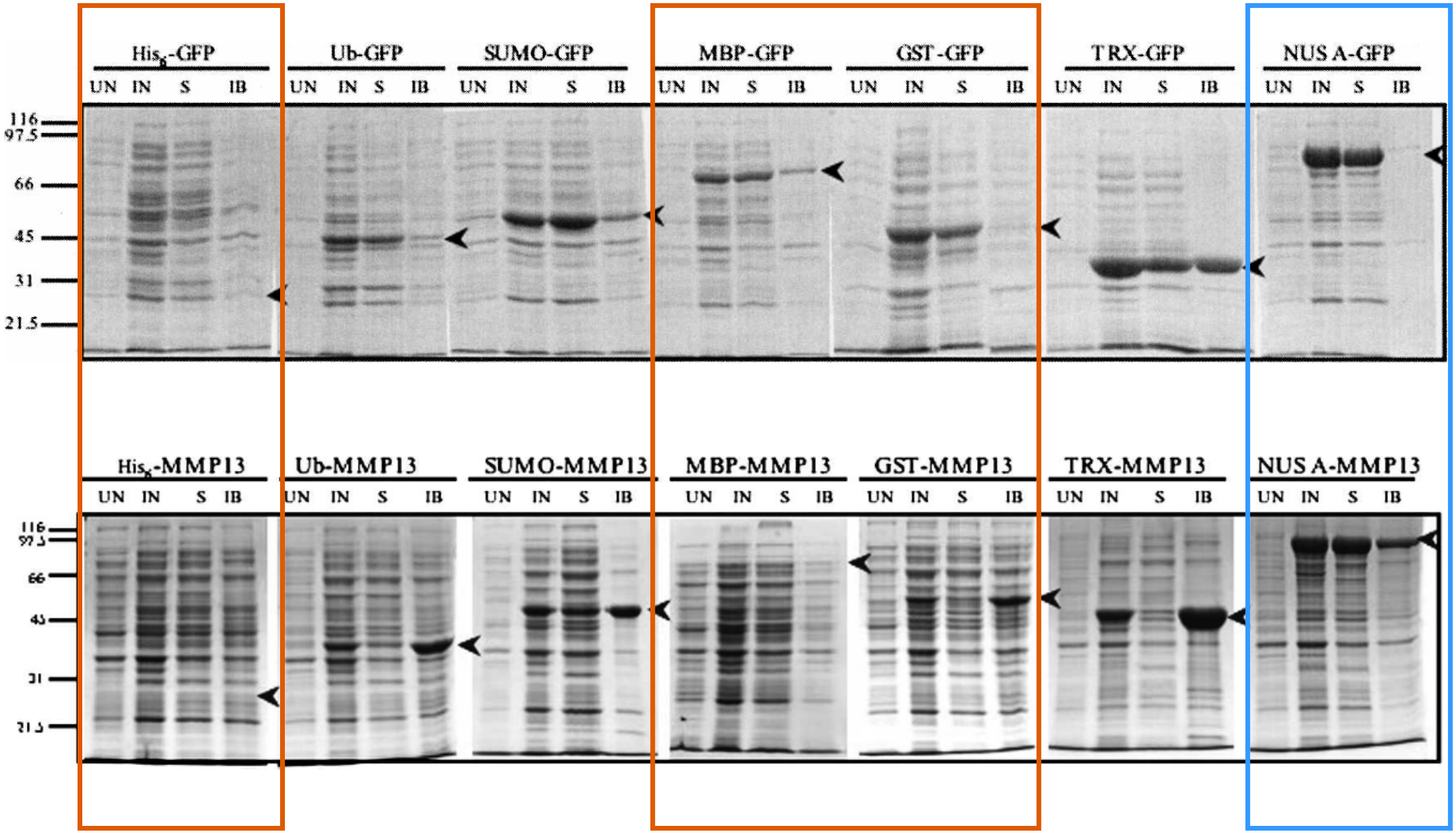
Solubility Tags Work



- **NusA, N-utilyzing substance A protien**
  - discovered through a systematic search for *E. coli* proteins that have the highest potential for solubility when overexpressed.
- **Trx, thioredoxin**
  - appears to catalyze the formation of disulfide bonds in the cytoplasm of *trxB* mutants
- **GST, glutathione-S-transferase**
- **MBP, maltose-binding-protein**
- **UB, ubiquitin**
- **SUMO, small ubiquitin-like modifier**

Marblestone, *et al* (2006) Protein Science

Solubility Tags Work



Marblestone, *et al* (2006) Protein Science

Common Protein Fusion Tags: Targeted to the Cytosol

TAG	RES	SIZE (Kd)	USE	MATRIX	COMMENTS
$Thio_6$	6	0.5	E	N/A	Small; facilitates expression
Polyhistidine	2-10	~0.2-1	P	IMAC	Small; inexpensive
MBP	396	40	P; S	Amylose	↑ solubility; periplasm
GST	211	26	P; S	Glutathione	Dimerization
Intein CBD	51	5.6	P	Chitin	Purif. in absence of thiol reducing agents
Strep-tag II	8	1.2	P	Biotin	Downstream detection
NusA	495	54.9	S	N/A	↑ solubility
Trx	109	11.7	S	N/A	↑ Disulfide bond*
DsbA	208	23.1	S	N/A	↑ Disulfide bond*; periplasm
Ubiquitin (Ub)	128	14.7	S	N/A	↑ solubility
SUMO	101	11.6	S	N/A	↑ solubility
GFP	239	27	D	N/A	Visualization of *in vivo* soluble expression
Z-tag	91	10.6	S	Protein A	Solubility enhancing tag
GB1	85	9.7	S	IgG	NMR testing *prior* to tag removal
pSET	~40	4.5	S	N/A	↑ solubility by ↑ negative charge of tag

Protein Targeting

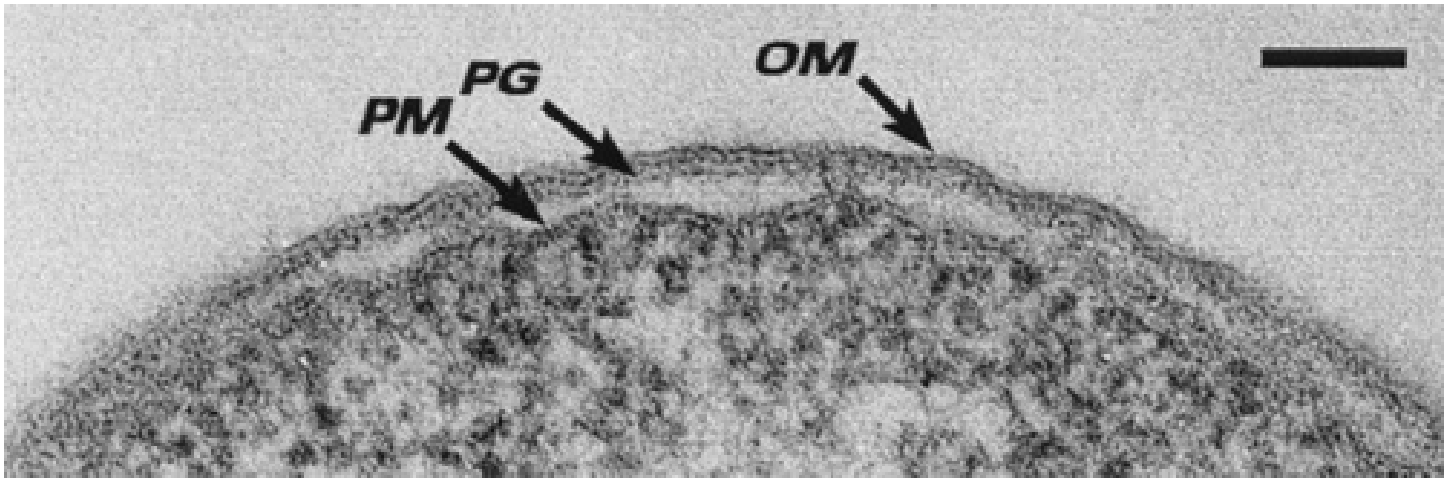
TAG	RES	SIZE (Kd)	USE	MATRIX	COMMENTS
$Thio_6$	6	0.5	E	N/A	Small; facilitates expression
Polyhistidine	2-10	~0.2-1	P	IMAC	Small; inexpensive
SignalSeq-MBP	396	40	P; S	Amylose	↑ solubility; periplasm
GST	211	26	P; S	Glutathione	Dimerization
Intein CBD	51	5.6	P	Chitin	Purif. in absence of thiol reducing agents
Strep-tag II	8	1.2	P	Biotin	Downstream detection
NusA	495	54.9	S	N/A	↑ solubility
Trx	109	11.7	S	N/A	↑ Disulfide bond*
DsbA	208	23.1	S	N/A	↑ Disulfide bond*; periplasm
Ubiquitin (Ub)	128	14.7	S	N/A	↑ solubility
SUMO	101	11.6	S	N/A	↑ solubility
GFP	239	27	D	N/A	Visualization of *in vivo* soluble expression
Z-tag	91	10.6	S	Protein A	Solubility enhancing tag
GB1	85	9.7	S	IgG	NMR testing *prior* to tag removal
pSET	~40	4.5	S	N/A	↑ solubility by ↑ negative charge of tag

Targeting to the Periplasm

Tags with Signal Sequences N-terminal to the Fusion Protein

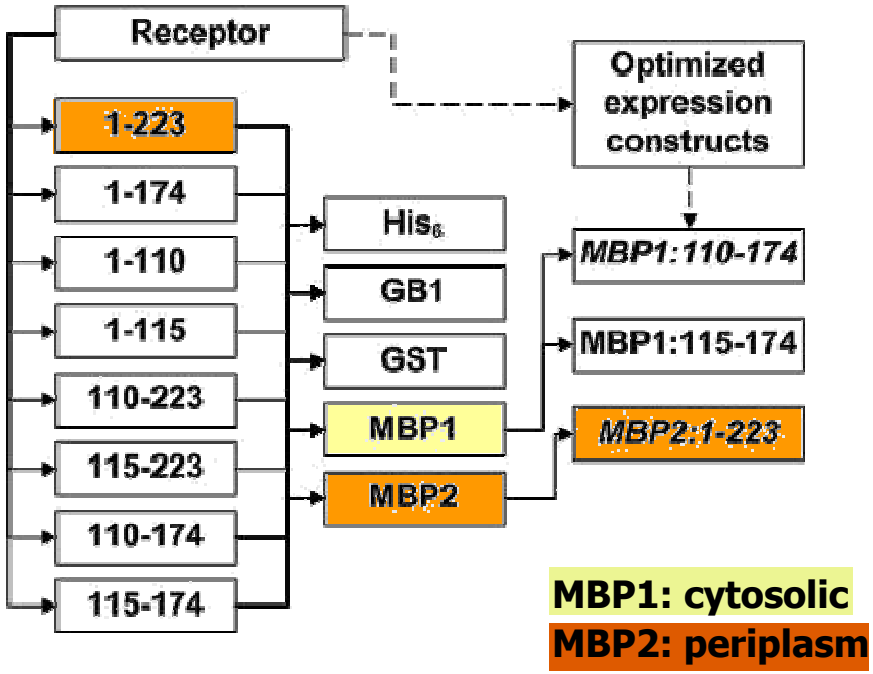
Result: Expressed Proteins targeted to the E. coli Periplasm

Sometimes the only way to get the protein expressed



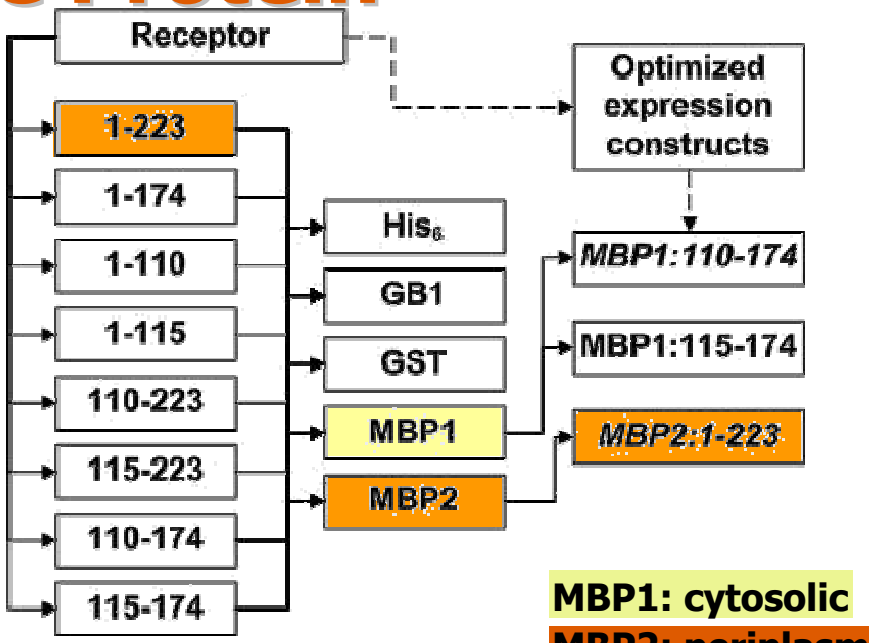
Lab Example: Membrane Protein

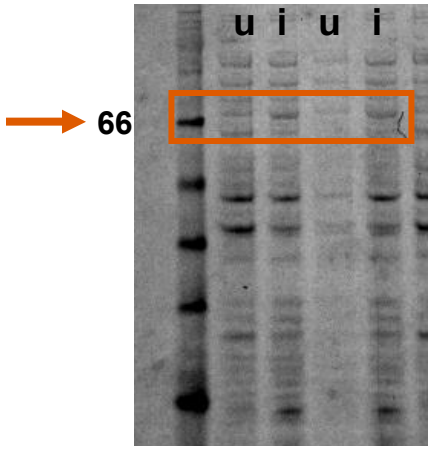
- **Receptor**
- **2 transmembrane helices**
- **Subcloned into multiple vectors with various fusion tags**

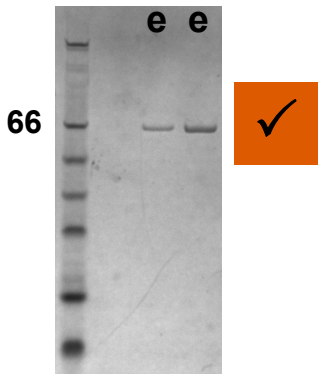


Targeting → No Overexpression but Only Soluble Protein

- Receptor
- 2 transmembrane helices
- Subcloned into multiple vectors with various fusion tags
- Weak <u>overexpression</u> for full-length protein (70 kD)
- Fully active, pure protein obtained <u>only</u> when expressed as an MBP fusion targeted to the periplasm with a signal sequence
- Yield: 1 mg/L cell culture







Common Protein Fusion Tags

TAG	RES	SIZE (Kd)	USE	MATRIX	COMMENTS
$Thio_6$	6	0.5	E	N/A	Small; facilitates expression
Polyhistidine	2-10	~0.2-1	P	IMAC	Small; inexpensive
MBP	396	40	P; S	Amylose	↑ solubility; periplasm
GST	211	26	P; S	Glutathione	Dimerization
Intein CBD	51	5.6	P	Chitin	Purif. in absence of thiol reducing agents
Strep-tag II	8	1.2	P	Biotin	Downstream detection
NusA	495	54.9	S	N/A	↑ solubility
Trx	109	11.7	S	N/A	↑ Disulfide bond*
DsbA	208	23.1	S	N/A	↑ Disulfide bond*; periplasm
Ubiquitin (Ub)	128	14.7	S	N/A	↑ solubility
SUMO	101	11.6	S	N/A	↑ solubility
GFP	239	27	D	N/A	Visualization of *in vivo* soluble expression
Z-tag	91	10.6	S	Protein A	Solubility enhancing tag
GB1	85	9.7	S	IgG	NMR testing *prior* to tag removal
pSET	~40	4.5	S	N/A	↑ solubility by ↑ negative charge of tag

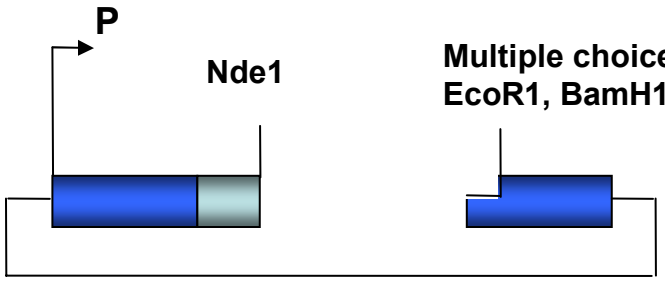
Common Protein Fusion Tags

TAG	RES	SIZE (Kd)	USE	MATRIX	COMMENTS
6xhis_Thio$_6$	6	0.5	E, P	N/A	Small; facilitates expression
Polyhistidine	2-10	~0.2-1	P	IMAC	Small; inexpensive
6xhis_MBP	396	40	P; S	Amylose	↑ solubility; periplasm
6xhis_GST	211	26	P; S	Glutathione	Dimerization
Intein CBD	51	5.6	P	Chitin	Purif. in absence of thiol reducing agents
Strep-tag II	8	1.2	P	Biotin	Downstream detection
6xhis_NusA	495	54.9	S, P	N/A	↑ solubility
6xhis_Trx	109	11.7	S, P	N/A	↑ Disulfide bond*
6xhis_DsbA	208	23.1	S, P	N/A	↑ Disulfide bond*; periplasm
6xhis_Ubiquitin (Ub)	128	14.73	S, P	N/A	↑ solubility
6xhis_SUMO	101	11.6	S, P	N/A	↑ solubility
6xhis_GFP	239	27	D	N/A	Visualization of *in vivo* soluble expression
6xhis_Z-tag	91	10.6	S, P	Protein A	Solubility enhancing tag
6xhis_GB1	85	9.7	S, P	IgG	NMR testing *prior* to tag removal
6xhis_pSET	~40	4.5	S, P	N/A	↑ solubility by ↑ negative charge of tag

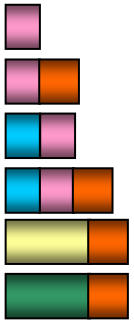
Vector Design

Maximize Efficiency: Parallel Cloning

'Click' Vector Design



N-terminal tags



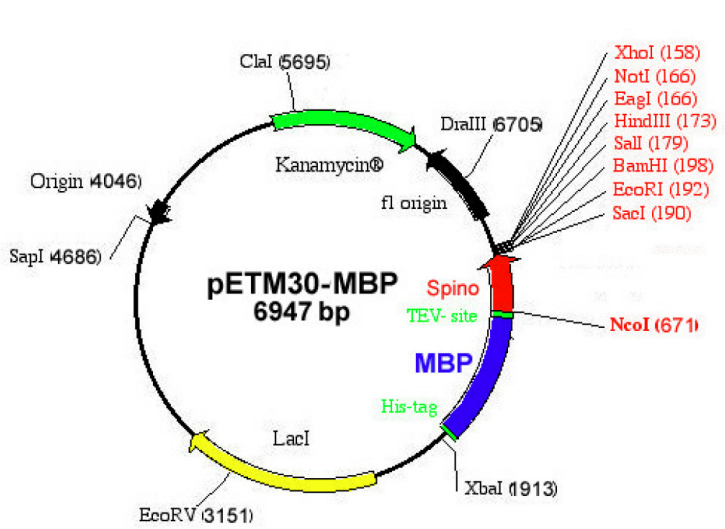
No Tag
HisTag
HisTag/TEV cleavage site
ExpTag/HisTag
ExpTag/HisTag/TEV cleavage site
GSTTag/TEV cleavage site
MBPTag/TEV cleavage site

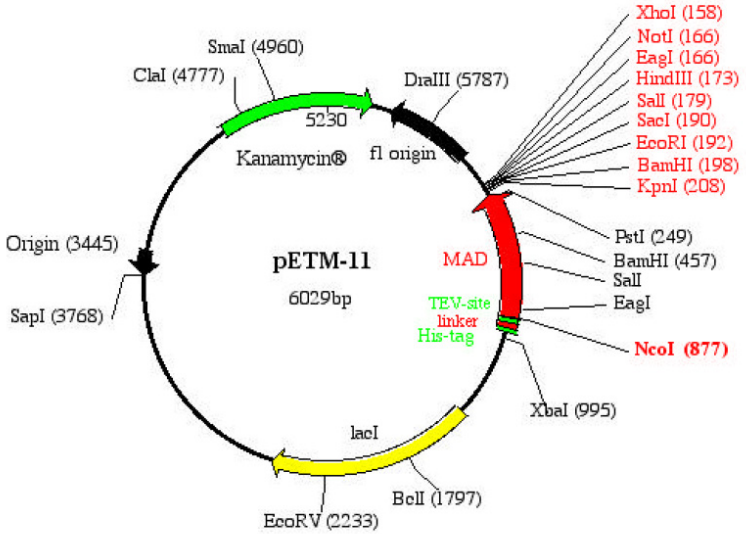
pBADM-60	map	seq	araBAD	Amp	N-NusA N-His	TEV	pUC	A. Geerlof
pBADM-60 (+)	map	seq	araBAD	Amp	N-NusA N-His	TEV	pUC	A. Geerlof
pBAT4	map	seq	T7	Amp	none	none	pUC	M. Hyvönen
pBAT5	map	seq	T7	Amp	none	none	pUC	M. Hyvönen
pETM-10	map	seq	T7-lac	Kan	N-His	none	pBR322	G. Stier

- **Sticky-end cloning into vectors with identical MCS sites**
  - **1 set of primers for gene amplification and sequencing**
  - **Only one digest needed to go into range of vectors**
  - **Affordable and great for undergraduates**
- **EMBL and 'in-house'**
- **Different promotors (pBAD/T7), origins of replication, antibiotic selection**

http://www.embl.de/ExternalInfo/protein_unit/draft_frames/index.html

Anatomy of the Vectors

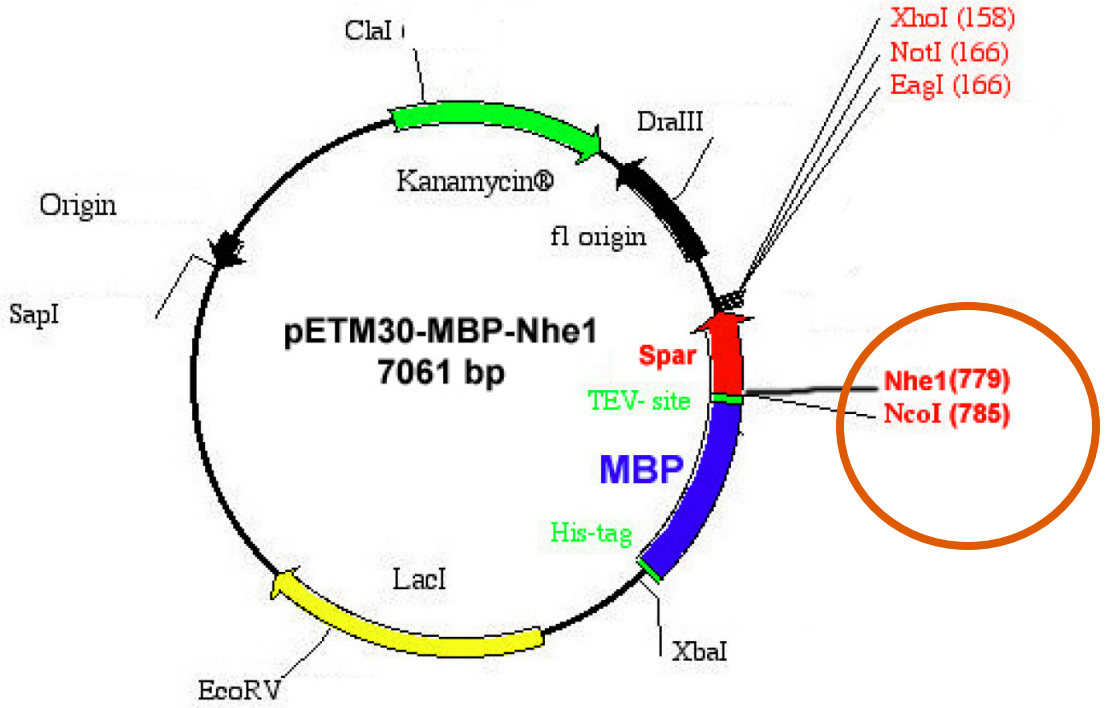




- **Identical MCS sites**
  - **1 N-terminal**
  - **7 C-terminal**
- **Nearly every vector includes a his-tag (purification) and Tobacco Etch Mosaic Virus Protease (TEV) cleavage site (tag removal)**
- **'Stuffer' sequences to simplify digestion (this is especially when working with undergraduates in the laboratory!)**

http://www.embl.de/ExternalInfo/protein_unit/draft_frames/index.html

Expand the Vector Set



- **Add a second digestion site at the N-terminus**
- **Additional N-terminal protein fusion tags**

Overview

- **Introduction recombinant protein expression in *E. coli:*** its benefits and disadvantages
- **Protein Sequence:** predictions for soluble expression
- **Strains and tags used for heterologous protein expression**: bacterial strains
- **Expression optimization:** how to maximize your protein yield
- **Conclusions**

Protein Expression Cell Lines

- Minimize Proteolysis
- Maximize Expression
- Minimize Leaky Expression
- Facilitate Disulfide Bond Formation
- Facilitate Folding (and thus solubility)

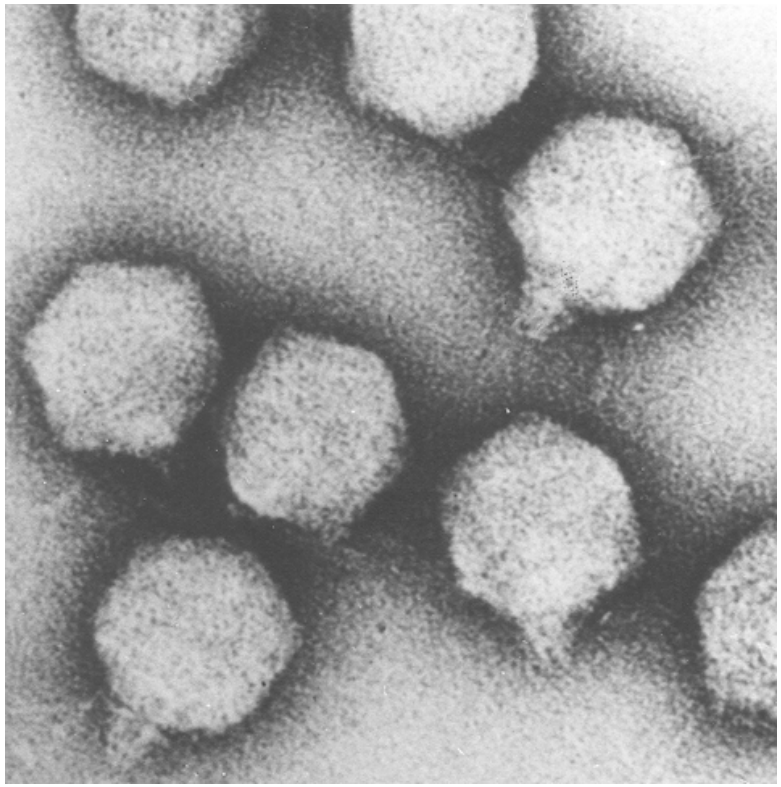
Bacterial Cell Lines

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 ^{o}C expression)
Labeling	
B834	Selenomethionine labeling for crystallography (*Met* auxotroph)
Membrane/Toxic Prot.	
pLysS	Reduce basal expression by expression of lysozyme
C43	BL21 derivatives that facilitate the soluble expression of toxic and especially integral membrane proteins

Bacterial 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 ^{o}C expression)
Labeling	
B834	Selenomethionine labeling for crystallography (*Met* auxotroph)
Membrane/Toxic Prot.	
pLysS	Reduce basal expression by expression of lysozyme
C43	BL21 derivatives that facilitate the soluble expression of toxic and especially integral membrane proteins

T7 pET *E. coli* Expression System



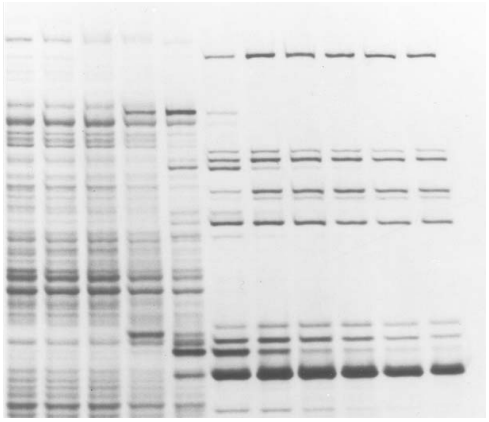
Electron micrograph of T7 phage particles

- **Bacteriophage T7 efficiently takes over the *E. coli* cell upon infection**
  - **T7 phage particle delivers T7 DNA to the *E. coli* cell**
  - **Efficient mechanisms direct the resources of the cell toward the production of new T7 phage**
  - **Can produce 250 new T7 phage particles in 13 min at 37°C**

Courtesy of Bill Studier, *BNL*

T7 Directs All Gene Expression in the Cell to its Own DNA

Time course of protein synthesis during T7 infection



-RNAP

-capsid

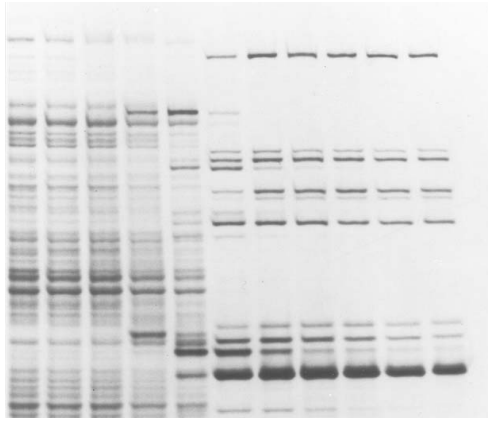
Pulse labeling with [^{35}S]methionine

- **T7 RNA polymerase highly processive**
- **T7 mRNAs have strong upstream translation signals**

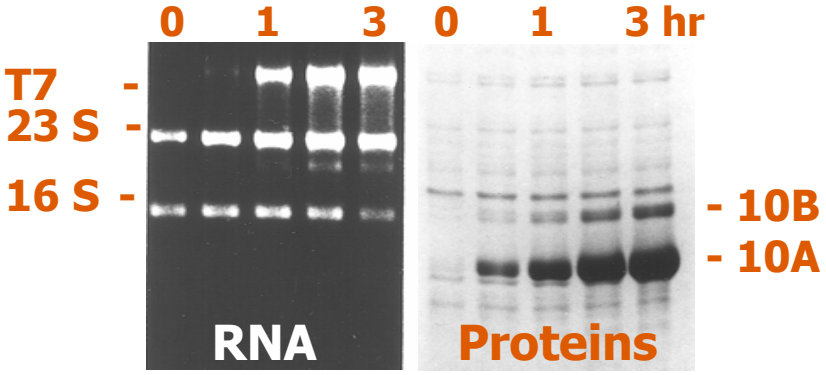
Courtesy of Bill Studier, *BNL*

T7 Directs All Gene Expression in the Cell to its Own DNA

Time course of protein synthesis during T7 infection



Pulse labeling with [^{35}S]methionine



- **T7 RNA polymerase highly processive**
- **T7 mRNAs have strong upstream translation signals**
- **Expression induced by IPTG (or lactose for autoinduction)**
- **Target mRNA can accumulate to levels that saturate ribosomes**
- **Tolerates toxic target proteins**
- **Low basal/high induced expression**
- **Target protein can accumulate to more than half of total cell protein**

Courtesy of Bill Studier, *BNL*

Codon Plus Cell Lines

CELL LINE	COMMENTS
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)

Especially useful for the expression of eukaryotic/human proteins

Table 1. Arg, Gly, Ile, Leu and Pro codon usage in *E. coli*

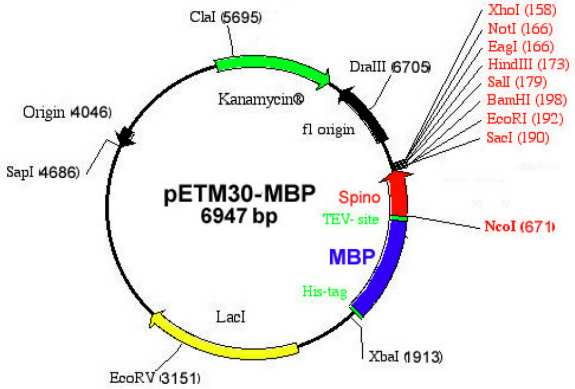
amino acid	codon	fraction in all genes	fraction in Class II
Arg	AGG	0.022	0.003
Arg	AGA	0.039	0.006
Arg	CGG	0.098	0.008
Arg	CGA	0.065	0.011
Arg	CGU	0.378	0.643
Arg	CGC	0.398	0.330
Gly	GGG	0.151	0.044
Gly	GGA	0.109	0.020
Gly	GGU	0.337	0.508
Gly	GGC	0.403	0.428
Ile	AUA	0.073	0.006
Ile	AUU	0.507	0.335
Ile	AUC	0.420	0.659
Leu	UUG	0.129	0.034
Leu	UUA	0.131	0.055
Leu	CUG	0.496	0.767
Leu	CUA	0.037	0.008
Leu	CUU	0.104	0.056
Leu	CUC	0.104	0.080
Pro	CCG	0.525	0.719
Pro	CCA	0.191	0.153
Pro	CCU	0.159	0.112
Pro	CCC	0.124	0.016

Novagen/Stratagene

Factorial Screening

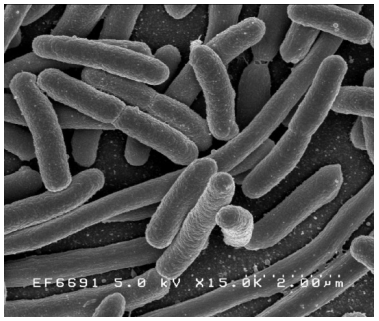
3 protein constructs

x



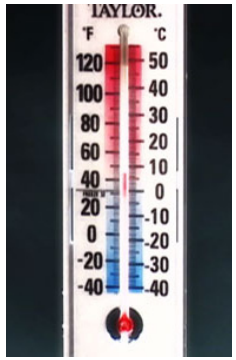
6 expression vectors

x



3 expression strains

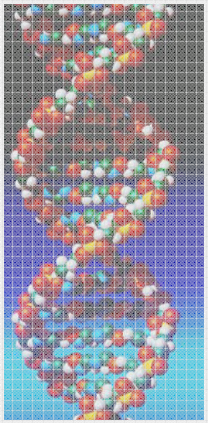
x



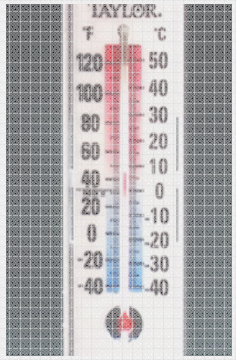
2 temps

102 Experiments!

Factorial Screening



3 protein constructs



2 temps

Army of Undergraduates

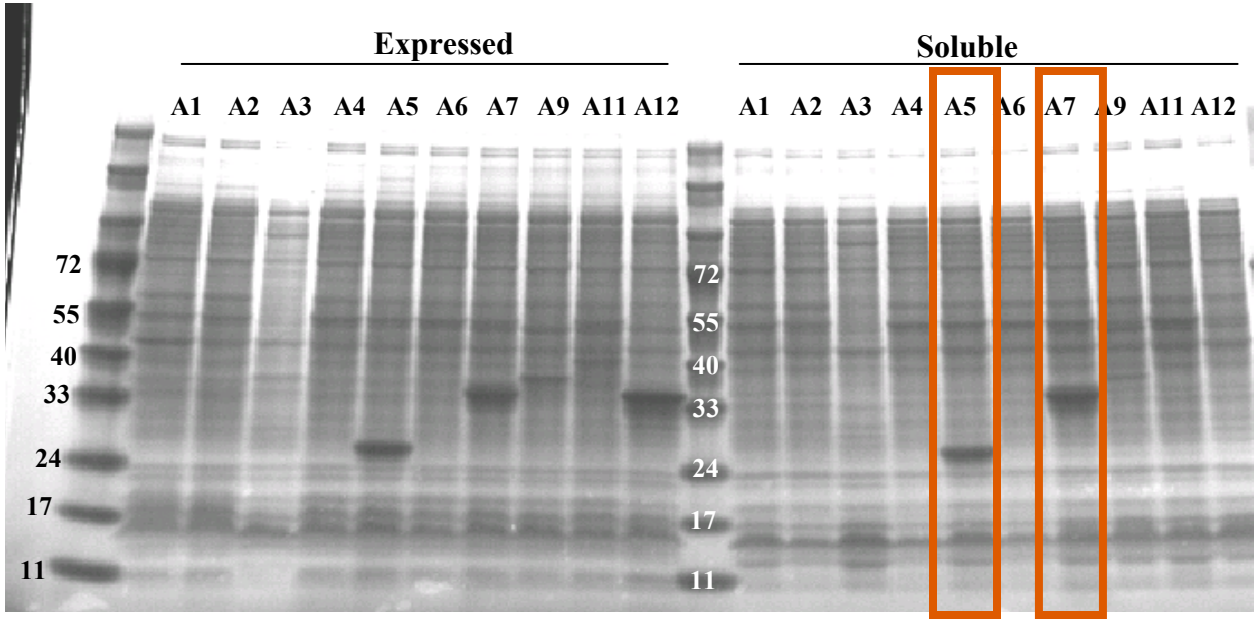
Micro-expression Prior to Scale-up

- **Vertiga Expression Device**
- **8 x 24-well blocks**
- **5 ml media/well**
- **3-D high velocity, shaking leads optical cell densities of 15-25.**
- **Temperature controlled (4-37 °C)**

THOMSON INSTRUMENT COMPANY

Page, et al. (2004) *BioTechniques,* 37:364-370

Vertiga Micro-Expression

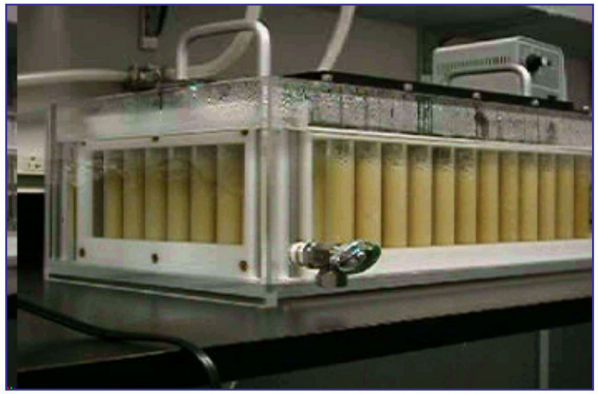


- **Only <u>soluble</u> constructs proceed to macroexpression**

Overview

- **Introduction recombinant protein expression in *E. coli:*** its benefits and disadvantages
- **Protein Sequence:** predictions for soluble expression
- **Strains and tags used for heterologous protein expression**: what to use when
- **Expression optimization:** how to maximize your protein yield: expression, chaperones, lysis and proteolysis
- **Conclusions**

Flasks and Media: Increasing Cell Culture Densities



Flasks: Ultra-Yield Flasks

- **Ultra-Yield Flasks**
  - **Plastic**
  - **Double-baffled flasks for increased aeration**
  - **62% of proteins expressed to higher ODs using these flasks than Fernbach when expressed at 18º C (no difference at 37º C)**
  - **A second study of 12 proteins found 100% increased at higher ODs (Pfizer; 18º C)**

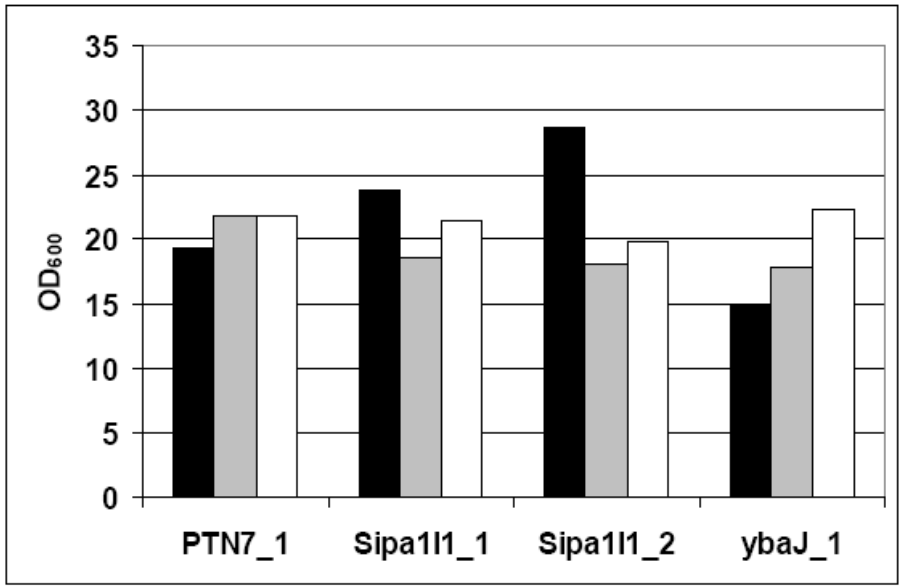
Flasks: Ultra-Yield Flasks

- **Ultra-Yield Flasks**
  - **Plastic**
  - **Double-baffled flasks for increased aeration**
  - **62% of proteins expressed to higher ODs using these flasks than Fernbach when expressed at 18º C (no difference at 37º C)**
  - **A second study of 12 proteins found 100% increased at higher ODs (Pfizer; 18º C)**

- **We use these flasks exclusively because of the increased aeration, they are light and easy to handle, and they don't break!**

Media: Auto-Induction: Simply Inoculate and Grow to Saturation

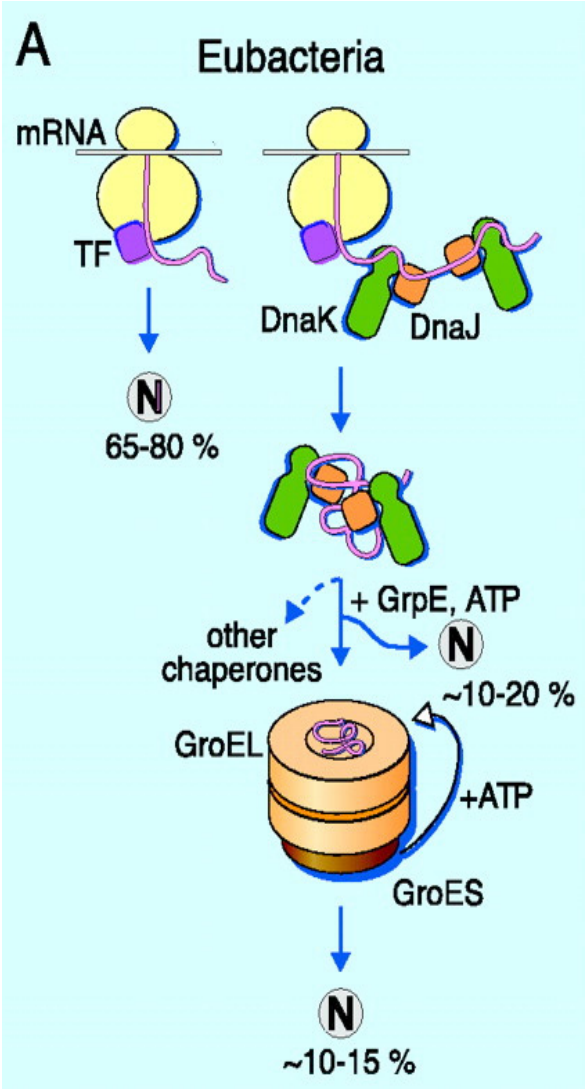
- **T7-based expression**
- **Combination of glucose and lactose**
- **Depletion of glucose → metabolism lactose → allolactose, a natural induction agent for the *lacUV5* promoter**
- **Compatible with SeMet and NMR Isotopic Labeling**
- **However, in our hands, we have found that the high level of expression can cause some proteins to express <u>insolubly</u>**



- **NO IPTG**
- **NO monitoring of ODs**
- **ODs as high as 30-40 with overnight expression**

Novagen

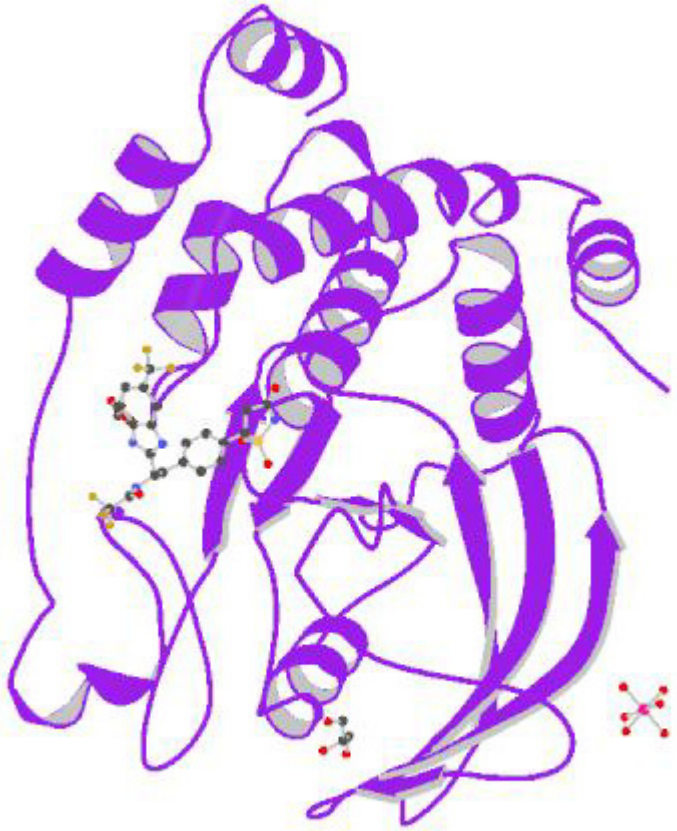
Chaperones



- **TF, Trigger Factor**
  - **Binds to and holds nascent proteins in a state that is competent to fold upon release from the ribosome**
- **DNaK/DNaJ/GrpE**
  - **Heat shock chaperones; ATP-dependent cycles of substrate binding and release**
- **GroEL/GroES (Cpn60/10)**
  - **Two homopentameric rings the bind and sequester unfolded proteins to allow folding with out non-productive intermolecular contacts (preventing aggregation). ATP-dependent.**

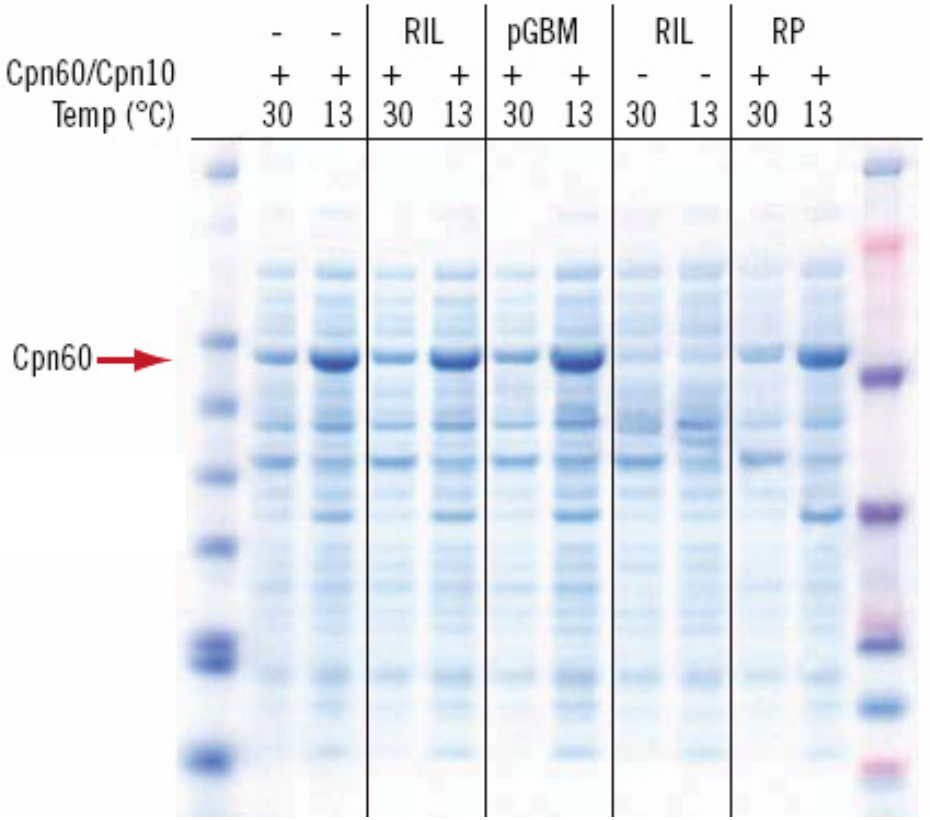
Lab Example: Maximizing Soluble PP1 Expression in *E. coli*

- **Protein Phosphatase 1**
- **Binds 2 metals at its active site**
- **Expression in *E. coli* is insoluble**
- **Increase solubilty**
  - **Add metals (supplement media with 1 mM $MnCl_2$)**
  - **Slow expression**
    - **induce with 0.1 mM IPTG**
    - **expressed at low temperature**



Expression Cells with Chaperones: Arctic Express

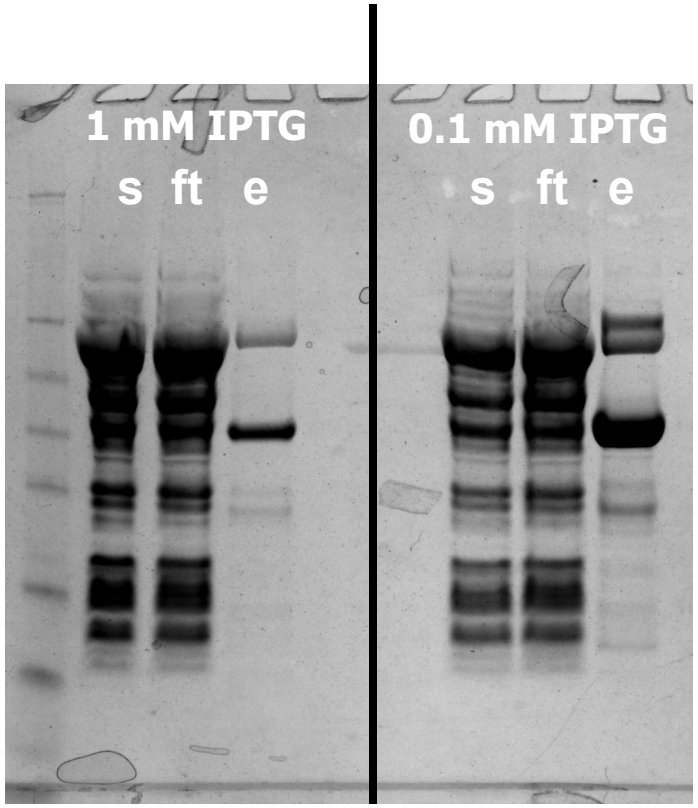
- **Cell Line that express the cold adapted chaperones Cpn60 and Cpn10 at 4-12 ºC**



Stratagene

PP1 in Arctic Express Cells

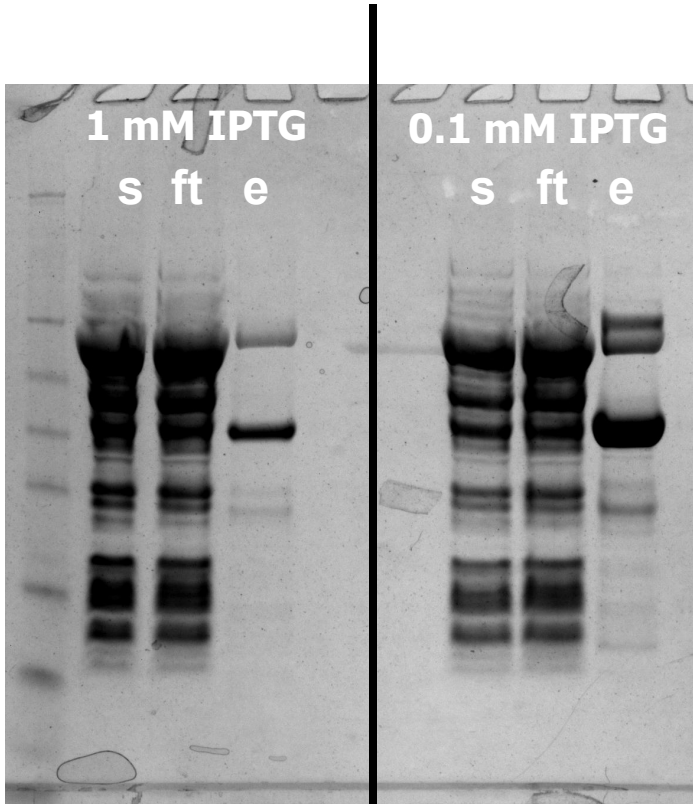
- **Protocol**
  - **Grow 24 ºC**
  - **Transfer to 10 ºC**
  - **Induce expression with IPTG**
  - **Express for 24-48 hours at 10 ºC**



More soluble, highly active (pNPP assay) PP1 with 0.1 mM IPTG

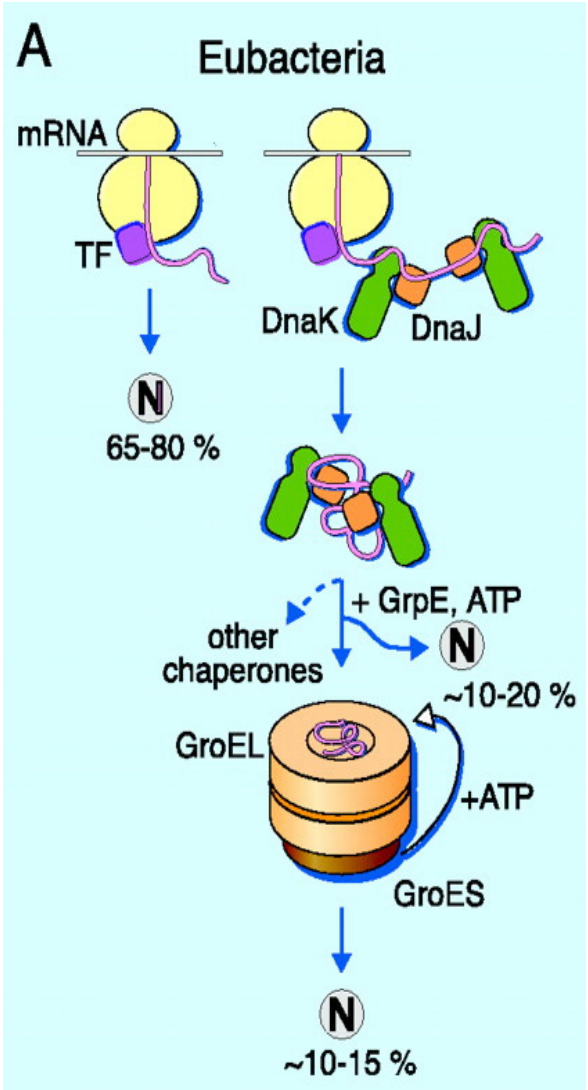
PP1 in Arctic Express Cells

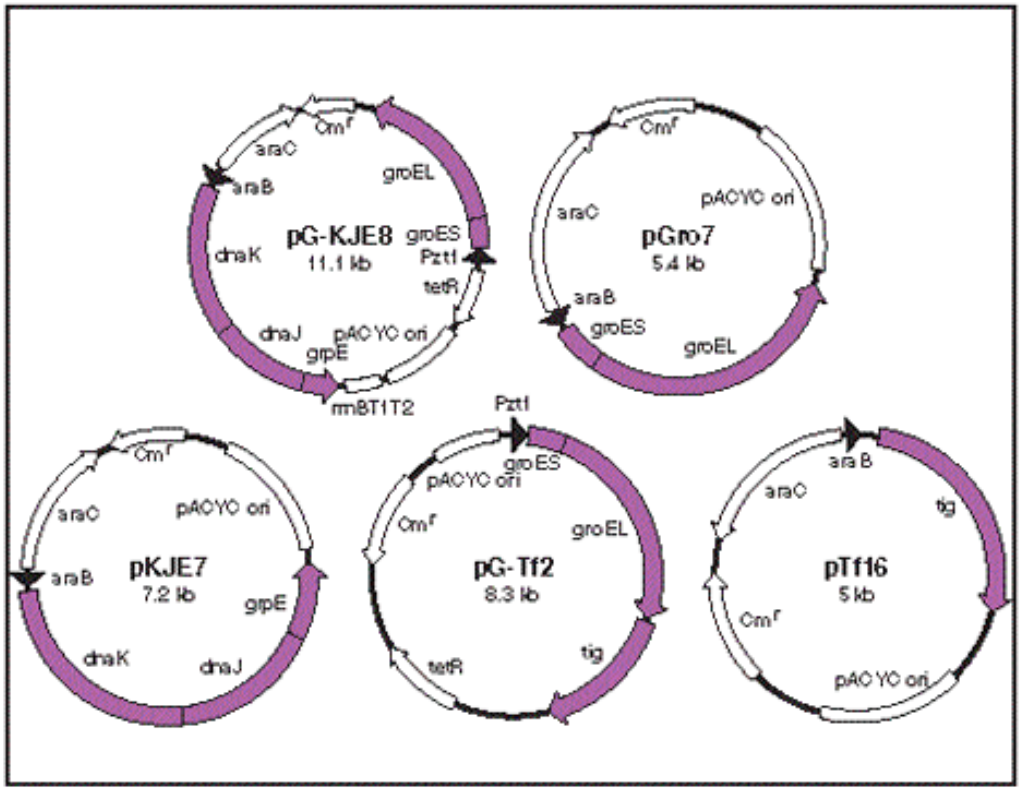
- **Protocol**
  - Grow 24 ºC
  - Transfer to 10 ºC
  - Induce expression with IPTG
  - Express for 24-48 hours at 10 ºC
- **Issues**
  - Cpn60 can co-purify with target protein
  - It is difficult (impossible) to dissociate target protein from chaperone



More soluble, highly active (pNPP assay) PP1 with 0.1 mM IPTG

Vectors with Chaperones: Co-expression

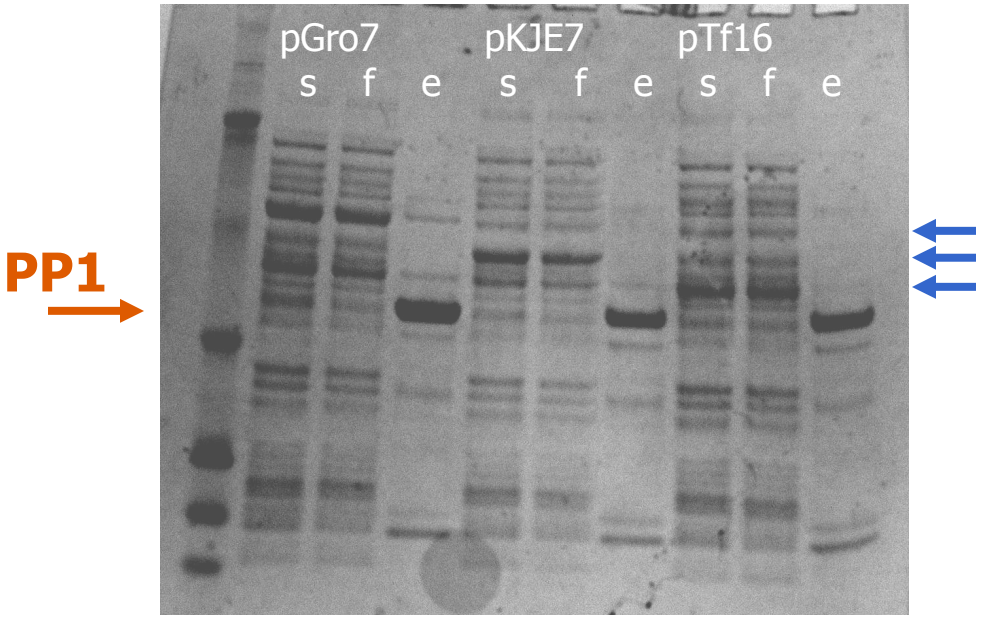




Maps of Takara's Chaperone Plasmids

Co-express PP1 with Chaperones

- **Co-expression of PP1 with**
  - **GroEL-GroES (pGro7)**
  - **DnaK/DnaJ/GrpE (pKJE7)**
  - **Trigger Factor (pTf16)**
- **Even higher yields of soluble PP1 protein**



10-fold increase in the amount of soluble protein produced using *chaperones*

Improving Soluble Cell Lysis

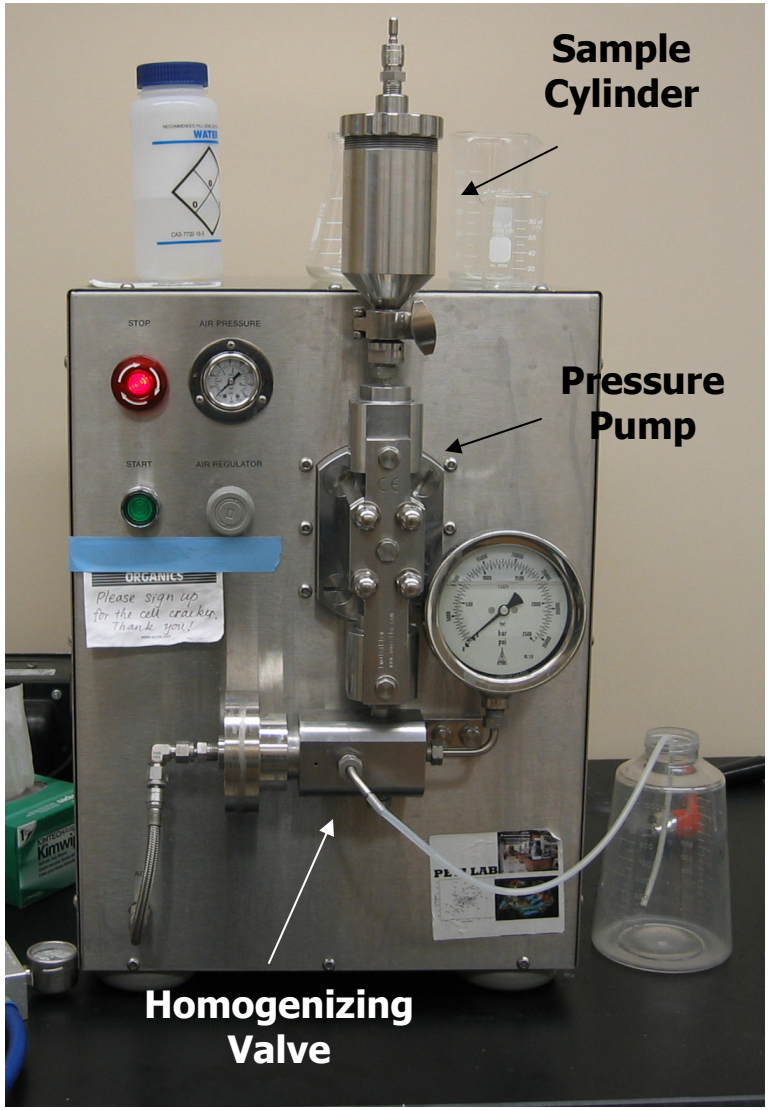
Lysis Method	COMMENTS
Freeze-Thaw; Enzymatic	
	Methods can be used separately; usually more effective when used in combination with one another
	Can become expensive for large culture volumes (lysozyme/DNaseI)
	Gentle; but incomplete lysis can be a problem
Sonication	
	Effective method for cell lysis
	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

Improving Soluble Cell Lysis

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
High Volume Homogenization	*Avestin Emulsiflex-C3*
	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

Avestin Emulsiflex-C3

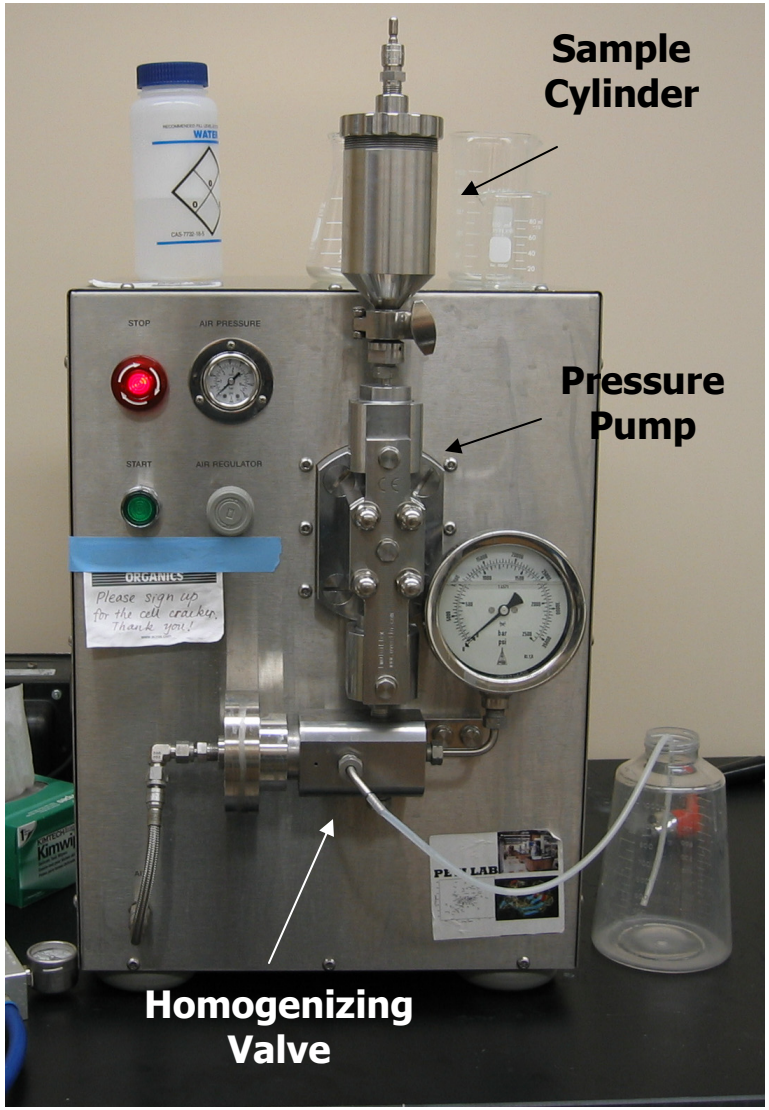
- 3 passes of cell suspension through the Emulsiflex for complete lysis
- For a 100 ml sample, 30 minutes from start to finish!
- We keep samples on ice to prevent heat denaturation (can also add a heat exchanger)
- Easy to use (even for undergraduates!)



www.avestin.com

Avestin Emulsiflex-C3

- 3 passes of cell suspension through the Emulsiflex for complete lysis
- For a 100 ml sample, 30 minutes from start to finish!
- We keep samples on ice to prevent heat denaturation (can also add a heat exchanger)
- Easy to use (even for undergraduates!)
- We've had a number of proteins that were insoluble when lysed using sonication that were soluble when lysed with the Emulsiflex!



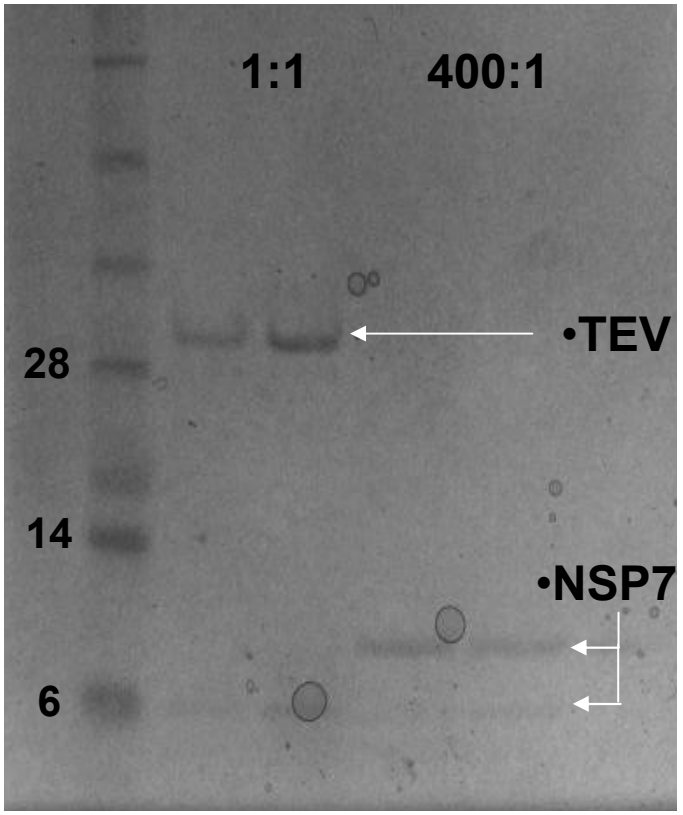
www.avestin.com

Removing Protein Fusion Tags

TAG	RES	SIZE	USE	MATRIX	COMMENTS
$Thio_6$_TEV	6	0.5	E	N/A	Small; facilitates expression
Polyhis_TEV	6	.6	P	IMAC	Small; inexpensive
MBP_TEV	396	40	P; S	Amylose	↑ solubility; periplasm
GST_TEV	211	26	P; S	Glutathione	Dimerization
Intein CBD	51	5.6	P	Chitin	Purif. in absence of thiol reducing agents
Strep-tag II_TEV	8	1.2	P	Biotin	Downstream detection
NusA_TEV	495	54.9	S	N/A	↑ solubility
Trx_TEV	109	11.7	S	N/A	↑ Disulfide bond*
DsbA_TEV	208	23.1	S	N/A	↑ Disulfide bond*; periplasm
Ubiquitin _TEV	128	14.7	S	N/A	↑ solubility
SUMO_TEV	101	11.6	S	N/A	↑ solubility
GFP_TEV	239	27	D	N/A	Visualization of *in vivo* soluble expression
Z-tag_TEV	91	10.6	S	Protein A	Solubility enhancing tag
GB1_TEV	85	9.7	S	IgG	NMR testing *prior* to tag removal
pSET_TEV	~40	4.5	S	N/A	↑ solubility by ↑ negative charge of tag

TEV Protease: Advantages

- **Highly specific with little to no non-specific cleavage**
- **Readily expressed and can be purified in the laboratory**
- **His-tagged for easy removal following cleavage**
- **Stored at -80º C for months without losing activity**
- **S219V mutant 100-fold more resistant to auto-inactivation than wild-type**
- **pRK603 TEV vector available for co-expression and *intracellular* cleavage of the tag to minimize purification of target protein**



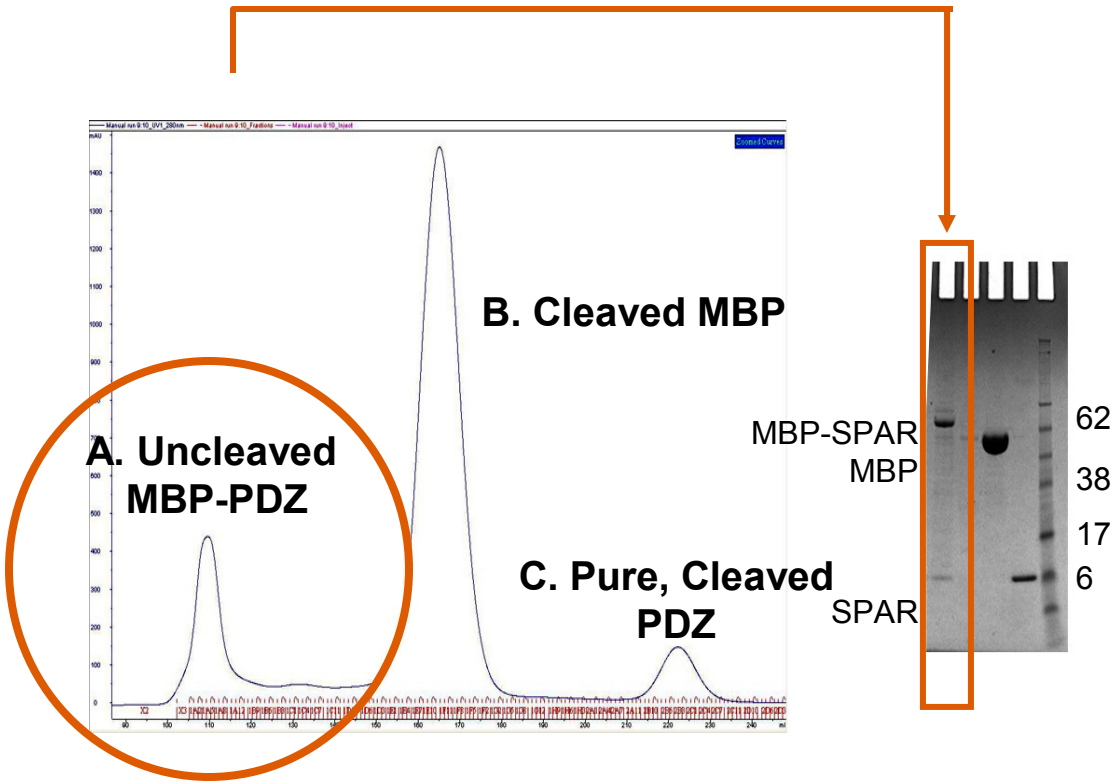
D. Waugh, NIH

TEV Protease: Practical Considerations

- **Active over a broad range of pH values**
- **Smaller proteins cleave readily to completion**

TEV Protease: Practical Considerations

- **Active over a broad range of pH values**
- **Smaller proteins cleave readily to completion**
- **MBP fusion proteins**
  - **require much higher concentrations (10-fold) of TEV for efficient cleavage**



Overview

- **Introduction recombinant protein expression in *E. coli:*** its benefits and disadvantages
- **Protein Sequence:** predictions for soluble expression
- **Strains and tags used for heterologous protein expression**: what to use when
- **Expression optimization:** how to maximize your protein yield
- **Conclusions**

Conclusions

Soluble protein expression is highly protein dependent:

there is still no magic bullet

- 'Click'-type vector system
  - save time
  - save money
- Experiment with your cell line: "think outside the -80"
- Change your habits:
  - Flasks
  - Media
  - Chaperones
  - Gentle Lysis
  - Cleavage
- Use a targeted approach
  - Test the 'best' combinations before trying them all

Acknowledgements

BROWN

Antoni Tortajada*
David Critton
Breann Brown
Kyle Skottke
Michael Hadley*
Igor Nederlof
Geoffery Stetson
Ojus Doshi

Tingting Ju*
Matt Kelker*
Barbara Dancheck
Mike Ragusa
Jebecka Hudak
Rene Kessler
Wolfgang Peti

2005 Rhode Island Foundation
Medical Research Grant

BROWN
2005 Richard B. Salomon
Faculty Research Award

BROWN
2006 Brown University
Seed Funds