Identify the Optimal Purification Strategies for Your Recombinant Protein Production

Liyan Pang, Ph.D.

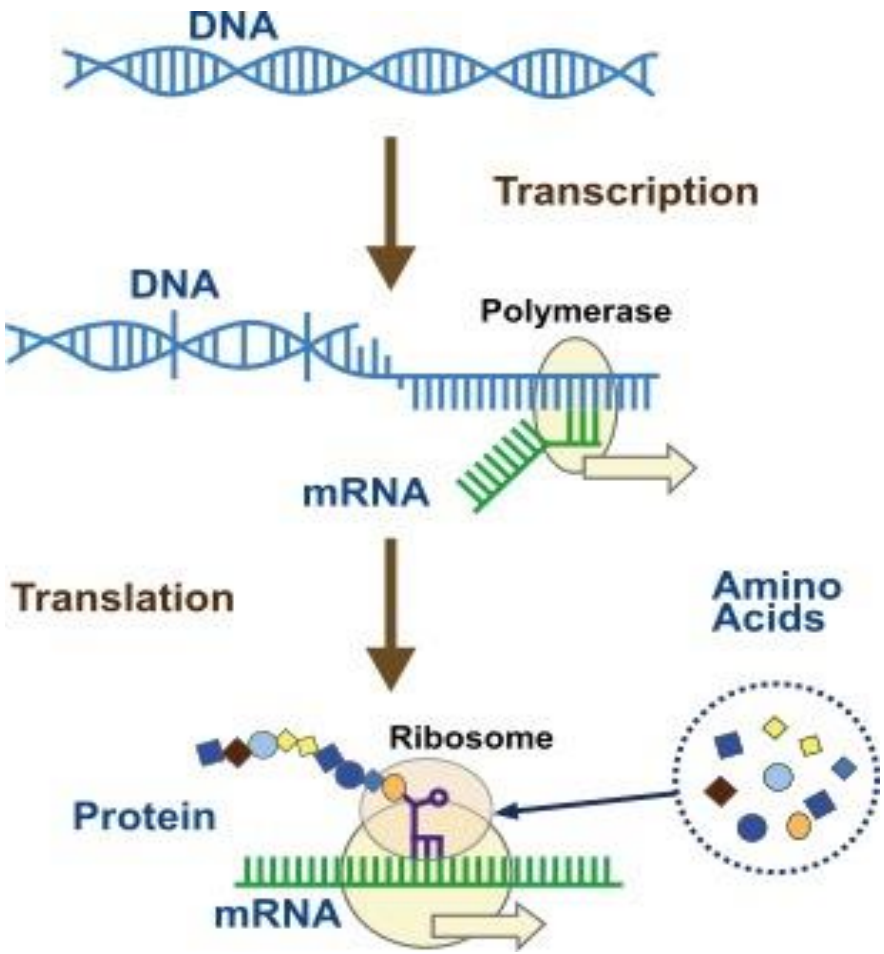
liyan.pang@genscript.com

GenScript
Make Research Easy

Table of Contents

Protein Production Workflow

Protein expression *in vivo*



Protein name

Obtain DNA

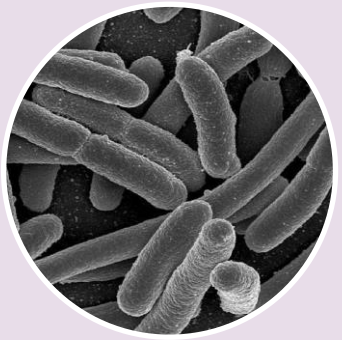
Create expression clone

Express the protein

Protein purification

Protein characterization

Expression Systems



Bacteria

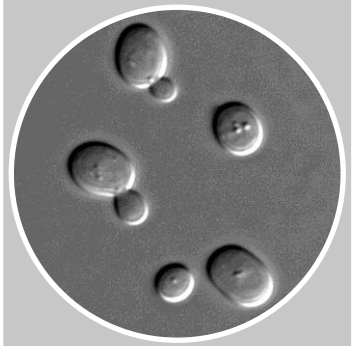
E. coli

1. "Work horse"
2. Well established
3. High expression
4. Simple genetics
5. Easy scale up
6. Speed
7. Costs
8. Equipment

Insect

Sf9, Sf21, S2, High-5

1. PTMs
2. Soluble proteins
3. High expressers



Yeast

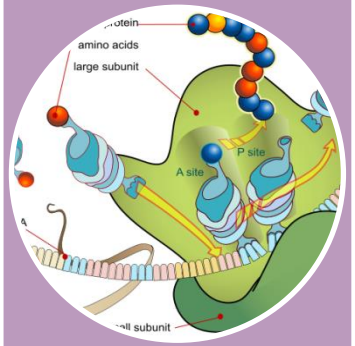
S. cerevisieae
P. pastoris

1. PTMs
2. Soluble proteins
3. High expresser

Mammalian

CHO, HEK, COS

1. PTMs
2. Soluble proteins
3. Low expresser
4. Expensive



Cell Free

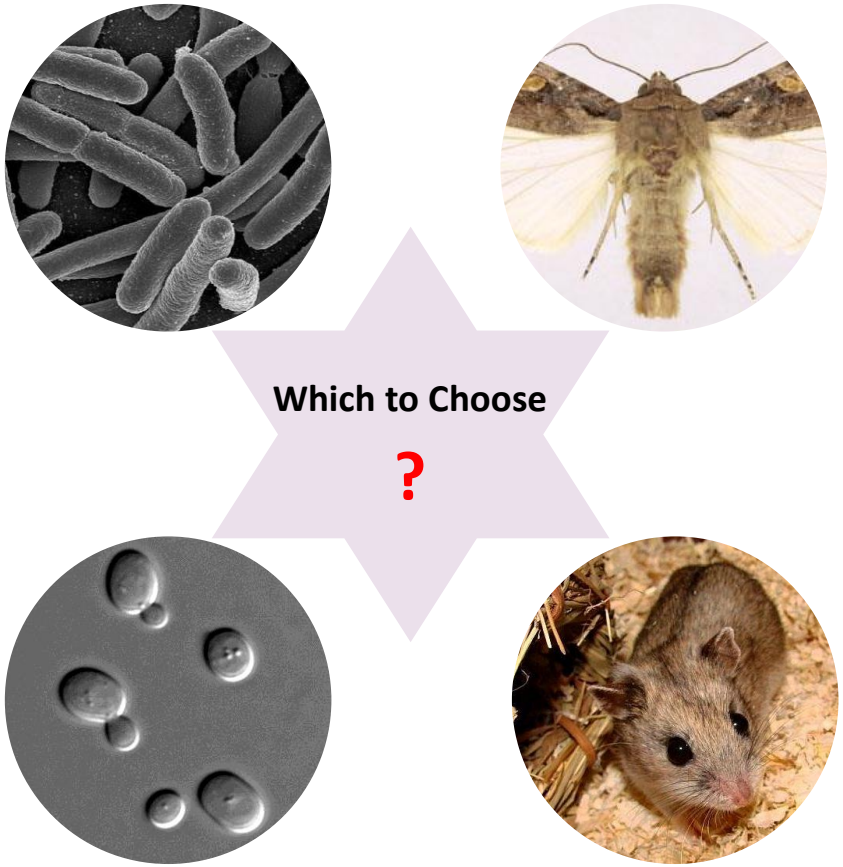
In vitro

1. Expensive
2. Not reproducible
3. Scalability issues

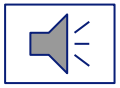
Which Expression System to Choose?

Factors to Consider:

- Protein property
  - MW
  - Disulfide bonds
  - Post-translational modifications
  - Homogeneity
- Intended applications
  - Structural biology
  - Functional assays
  - Therapeutic protein/vaccines
  - Antigens for Ab production
  - Protein-protein interactions
- Yield
- Cost



GenScript Protein Production Services



From sequence to purified protein - gene synthesis included!

Expression System	Deliverables	Timeline	Price
BacPower™	3mg purified protein guaranteed	6 -8 weeks	Staring from $2,200
InsectPower™	1mg purified protein guaranteed	8 -10 weeks	Staring from $3,950
MamPower™	3mg purified recombinant protein or 50mg purified antibody guaranteed	8 -12 weeks	Staring from $8,499
YeastHigh™	Customizable production up to 2000L	8 -10 weeks	Quote

PROTential™ - Expression Evaluation & Optimization

Eliminate the guesswork from your protein production work

- Evaluate whether your target protein expresses in your chosen system
- Identify the best expression system for your target protein
- Optimize your protein expression

Name	Service Type	Expression System(s)	Price	Timeline
Standard	Protein Expression Evaluation	*E.coli* (SC1653-B)	Starting from $280	1-2 weeks
		Insect (SC1653-I)	Starting from $400	3-4 weeks
		Mammalian (SC1653-M)	Starting from $500	3-4 weeks
		All 3 Systems (SC1653-3S)	Starting from $950	3-4 weeks
Silver	Protein Expression Optimization	*E.coli* (SC1667)	Login to inquire	2-3 weeks
Gold	Protein Expression Optimization	*E.coli* (SC1668)	Login to inquire	4-8 weeks

GenScript's Solution for Expression Optimization

Case Study- Protein Expression Optimization

- **Challenges:**
  - Pilot purification (final yield 1mg/L)- ~ 28kDa protein;
  - Large amount of protein with large scale fermentation (1000L) and purification is needed.
- **Strategies:**
  1. Expression improvement:
     a. Promoter optimization
     b. Strain optimization
     c. pH optimization
     d. Temperature and induction optimization
     e. Inoculated quantity optimization
  2. Recovery rate improvement during purification - Purification condition optimization
  3. High-throughput screening with 48 conditions

Case Study- Protein Expression Optimization

- 12X as much protein expression yield as original

1 mg/L starting protocol

2 mg/L T7 promoter/induction condition optimization

5 mg/L phoA promoter/induction condition optimization

6 mg/L growth condition optimization (pH)

12 mg/L seeding density optimization

Table of Contents

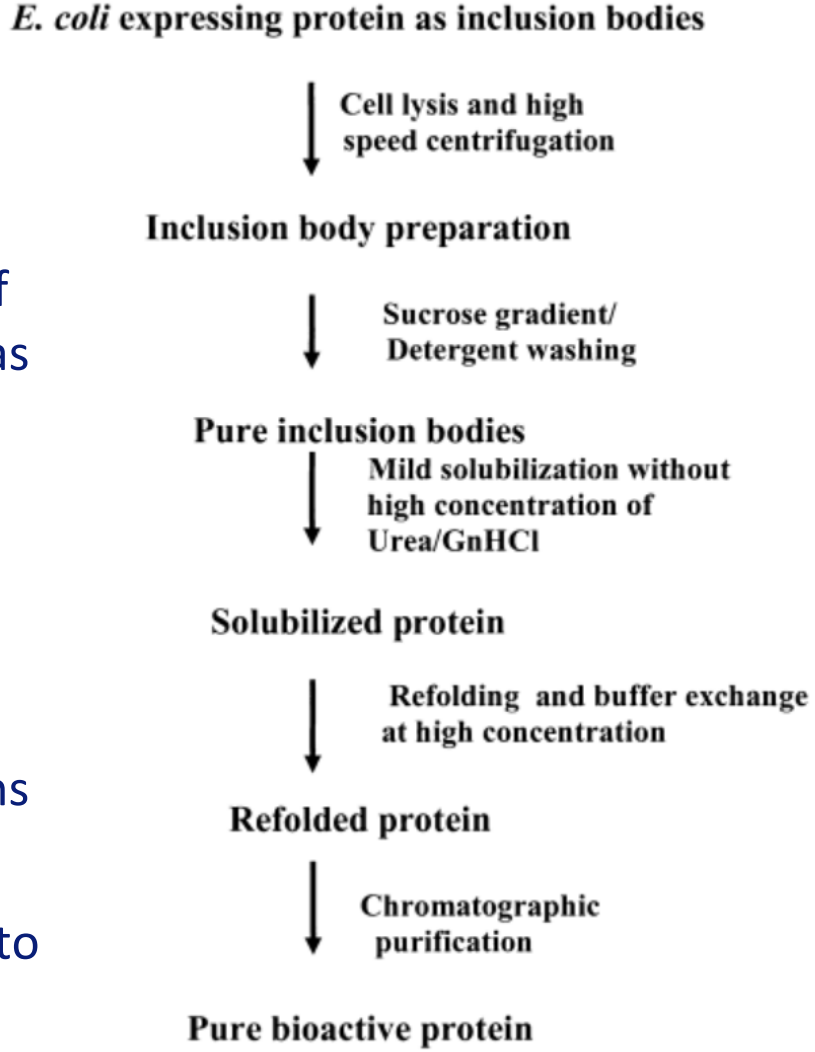
What to Consider for Protein Purification?

Goal: To apply the optimal protein purification strategy specific to your recombinant protein, with which the highest level of purification is reached in the fewest steps.

- Protein characteristics
- Soluble vs. inclusion bodies
- Protein purification methods
- Protein with specific ligands
- Tagged protein
- Easy-to-degrade protein
- Concentration - Ultrafiltration
- Endotoxin removal

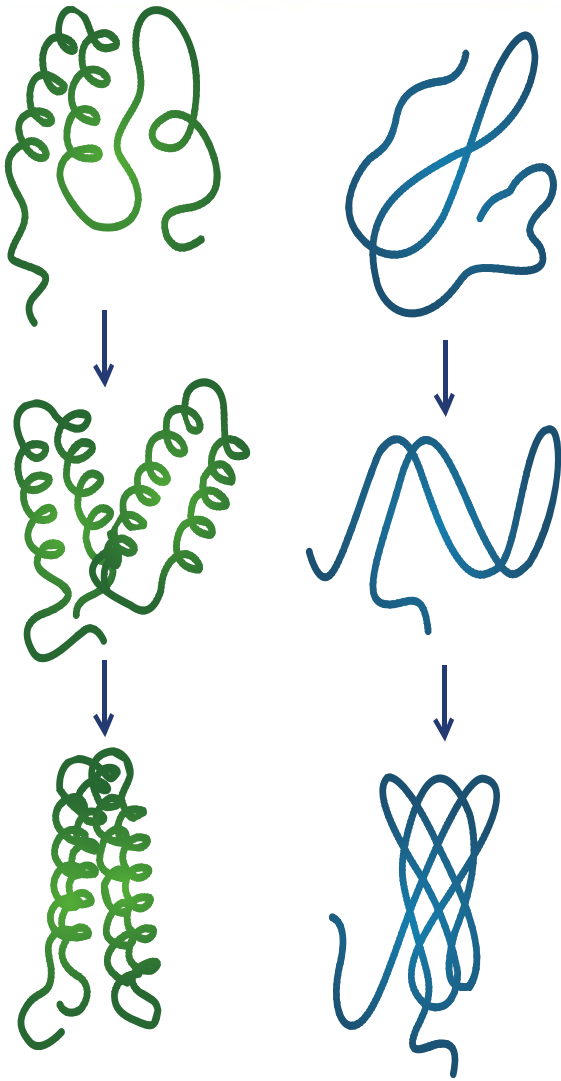
Protein Expression as Inclusion Bodies

- **What is an inclusion body?**
  - When E.coli is transformed to manufacture large amounts of recombinant protein, the protein sometimes forms dense aggregates of insoluble misfolded proteins, known as inclusion bodies.
- **Benefit**
  - allow high protein concentrations
  - protect sensitive proteins from proteolytic (enzymatic) degradation
  - protect the cell from any toxic proteins
- **Challenge**
  - to solubilise and refold this protein into its correct 'active' form



Protein Refolding Challenges

- All the information necessary for folding the peptide chain into its native structure is contained in the primary amino acid sequence of the peptide.
- There are vastly too many different possible conformations for a protein to fold by a random search.
- A new view of protein folding suggested that there is no single route, but a large ensemble of structures follow a many dimensional funnel to its native structure.
- Refolding conditions must be optimized for each individual protein.
- Important variables are:
  - buffer type
  - pH
  - ionic strength
  - Additives, often in combination (glycerol, redox reagents, saccharides, amino acids, metal ion, detergents, chaperones)



GenScript's FoldArt™ Technology Overview

- Evaluation of target proteins' biochemical and biophysical properties
- Refolding optimizations
  - Selection of particular refolding strategy based on protein's sequence and the structural properties.
  - Buffer screening: Solutions for the inclusion body will be diluted to 20 different refolding buffers to determine which parameters affect the refolding results.
- Denaturant removal

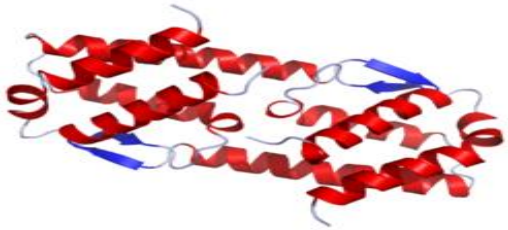
  Techniques: dilution, dialysis, diafiltration, gel filtration, and chromatography (ion exchange, size exclusion, and affinity)
- Validation

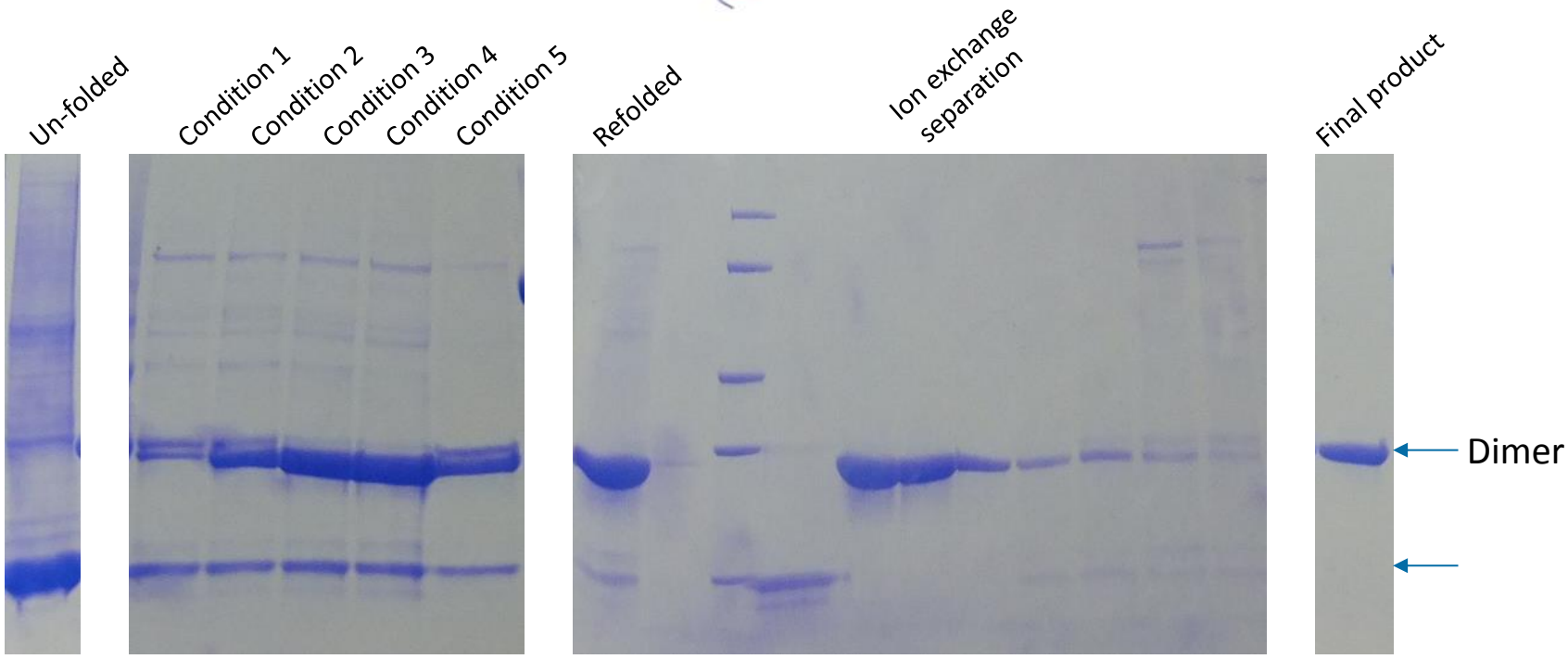
  Refolding results will be validated by SDS-PAGE, HPLC and/or functional assay.

Over 95% of the inclusion bodies can be solubilized and refolded by our proprietary FoldArt™ protein refolding technology

Case Study: Protein Refolding

- Human interleukin – 5: disulfide-bond linked homodimer as active form

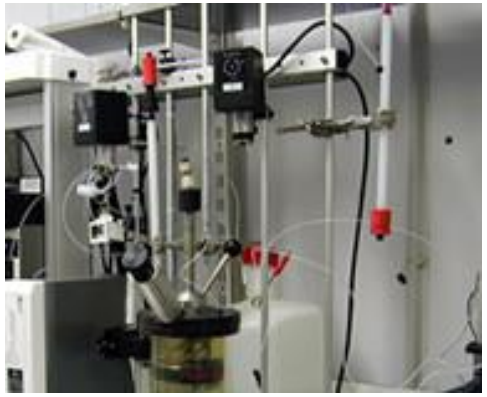




Protein Purification Methods

- Multiple choices on purification methods
  - Affinity column (GST, Ni-NTA, protein A/G/L resins, etc.)
  - Ion exchange (IEC)
  - Size exclusion (SEC)
  - Hydrophobic interaction chromatography (HIC)

- Double Tag strategy for big protein isolation



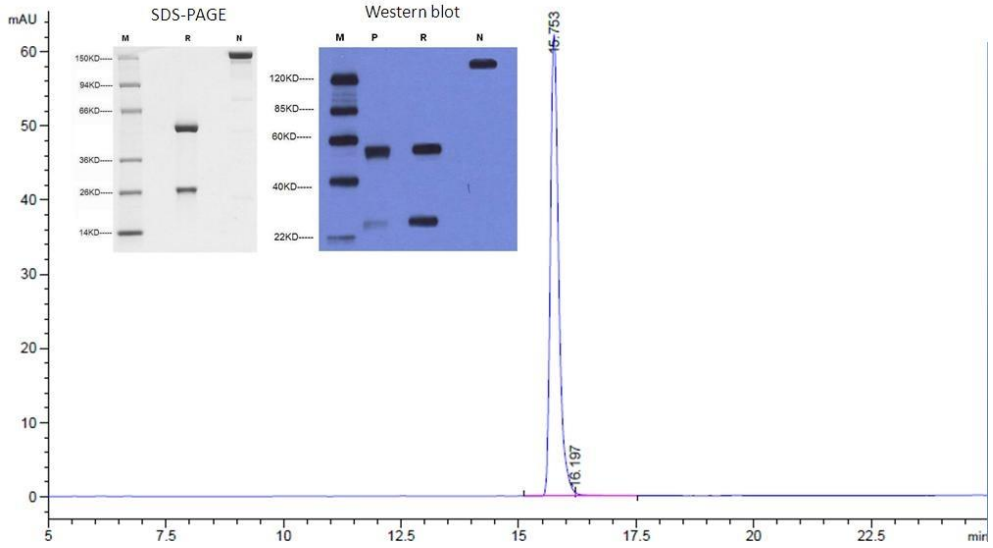
Select Optimal Protein Purification Method

- **Protein with special ligands**
  - IgG: Protein A or Protein G column;
  - Fab & ScFv: Protein G or Protein L column;
  - VhH/ScAb: Protein A column.
- **Protein with tags**
  - Small tags:
    - His tag: Nickel-NTA column;
    - FLAG tag (DYKDDDDK): Anti-DYKDDDK conjugated resin;
  - Big tags:
    - GST tag: Beads coated with Glutathione;
    - MBP tag: Amylose resin
- **Tag free protein or additional purification**
  - Tag free protein: first to precipitate out the protein by ammonia sulfate.
  - MW: proteins with distinct sizes can be separated by SEC.
  - pI: IEC can be used to separate proteins based on their charges under particular pH.
  - Protein with particular binding domain: Heparin column can be used to separate proteins with positive charged domains;
  - Hydrophobicity: last resort to try, separation based on the proteins' hydrophobicity.

Case Study: Recombinant mAb Production

Scale-up Production of a Recombinant mAb (hIgG1/kappa)

- Quality Standards:
  - Purity over 95% by SDS-PAGE and HPLC, even up to 99% purity
  - Aggregation less than 5%
  - Endotoxin Level: 0.05-1 EU/mg
  - Concentration: 20 mg/ml
  - Host cell protein 100 p.p.m.
  - Host cell DNA: < 100 pg/dosage



SDS-PAGE, Western blot and HPLC analysis

Transient production
- 1g of purified Antibody

Industrial grade cell line development
- Cell line with productivity of gram level

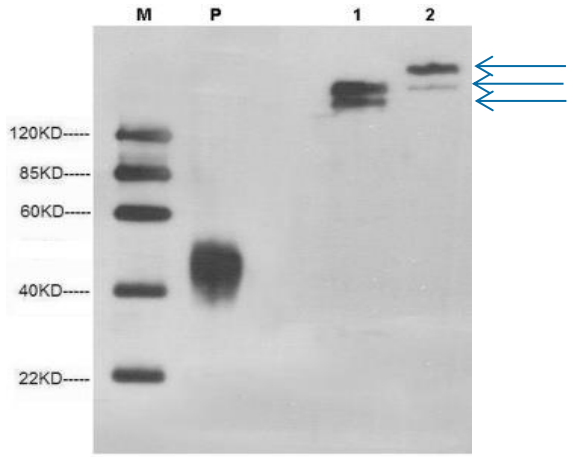
100L scale up production
- 40~50 g of purified antibody

Case Study: Prone-to-Degrade Protein

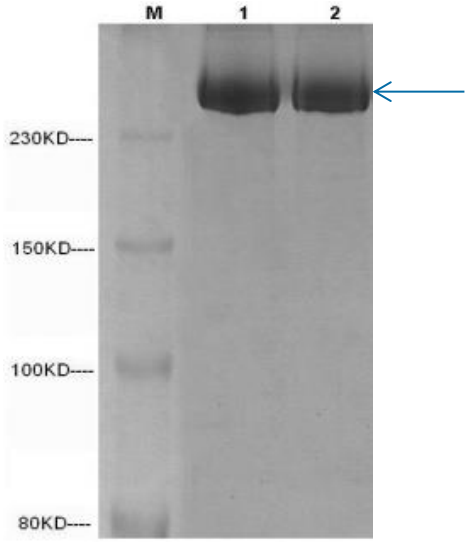
- Inconsistency of measured concentration
- Trouble shooting:
  - Transfection methods
  - Cell lysis
  - Purification
- Challenge: DNase is only partially responsible for the protein degradation. This protein itself is prone-to-degrade.
- Solutions:
  - Remove DNase
  - Add protease inhibitor to every step
  - Optimize buffer components
  - Add protein stabilizers
  - Lyophilization immediately after protein purification
  - Storage temperature

Western blot

Before



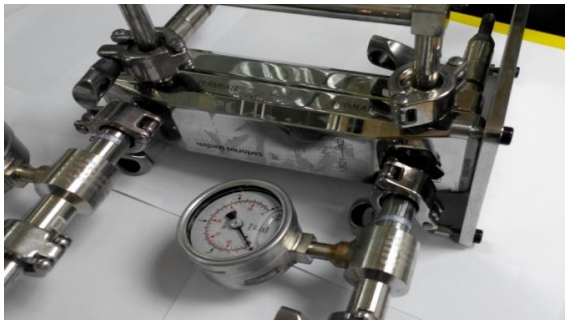
After



Protein Concentration - Ultrafiltration

Cross-flow filtration

Plate-and-Frame Module



Hollow-Fiber Module

Advantages

- High allowable work pressure
- Easy to clean
- Easy to replace membranes

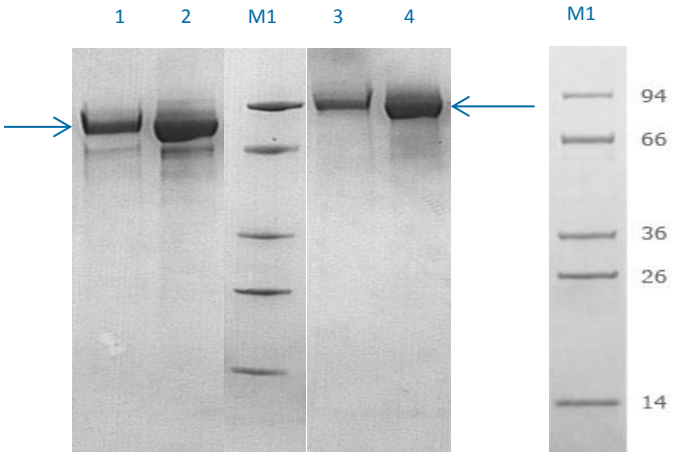
Dead-end filtration

Advantages

- Ultra-fast spin times (process 15mL samples as fast as 10-15 minutes)
- Ultra-high recoveries (typically >90%)
- Relative low costs

Case Study on Protein Concentration

Slice Cassettes; MWCO:30 KD; Membrane Area:0.1m²							
Sample	MW（KD）	Initial concentration（mg/ml）	Initial volume（L）	Final concentration(mg/ml)	Final volume（ml）	Tem (℃)	Recovery（%）
Protein S	79.5	0.05	5	1.000	243	RT	97.2
Protein Q	92.7	0.04	5	0.800	196	RT	98.0



M1: Protein Marker

Lane1:The **protein S** before Concentration (30ul)

Lane2:Concentrated **protein S** (3ul)

Lane3:The **protein Q** before Concentration (30ul)

Lane4:Concentrated **protein Q** (3ul)

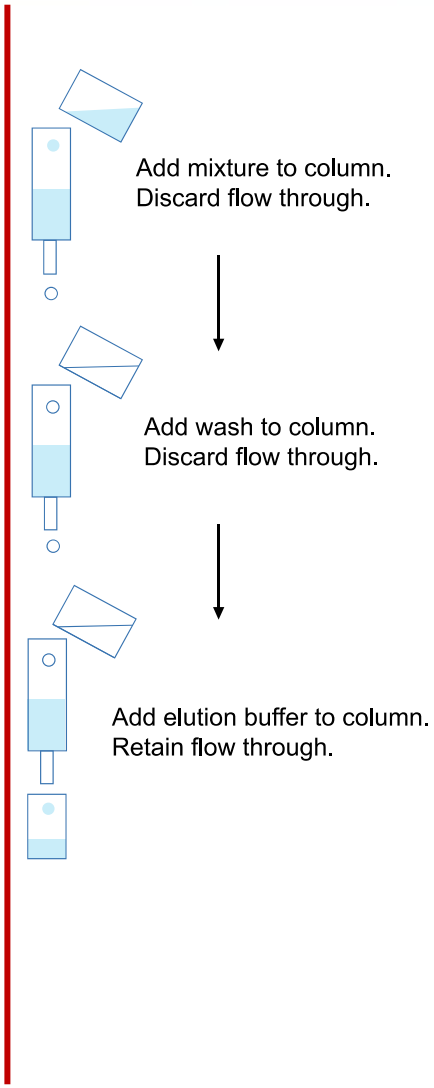
Table of Contents

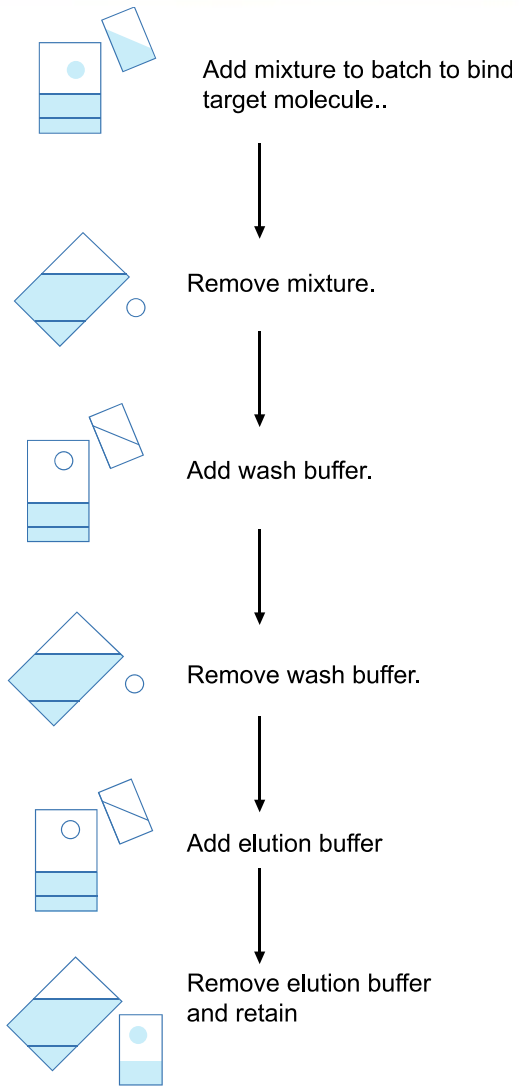
Factors to Consider When Using Protein Tags

- **Protein information:**
  - Fusion protein (PI, Molecular weight, concentration)
  - Tag type
  - Protease cleavage site
  - Tag-free protein (PI, Molecular weight)
- **Protease information:**
  - Protease activity
  - Protease reaction condition

Affinity Chromatography with Tagged Proteins

- Affinity tags
  - His
  - GST
  - MBP
  - Flag
  - SUMO
- Matrices & elution conditions
- Methods
  - Column
  - Batch
- Tag removal by site-specific protease
  - TEV
  - Enterokinase
  - SUMO protease

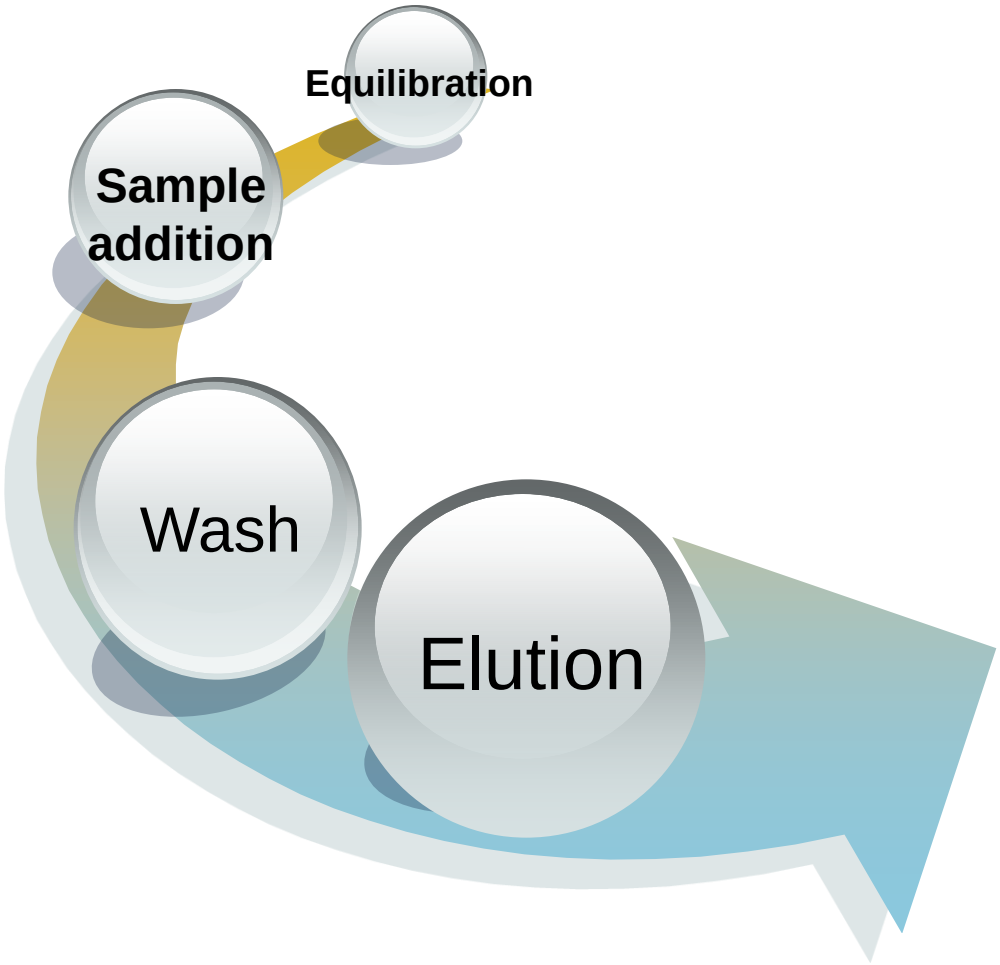


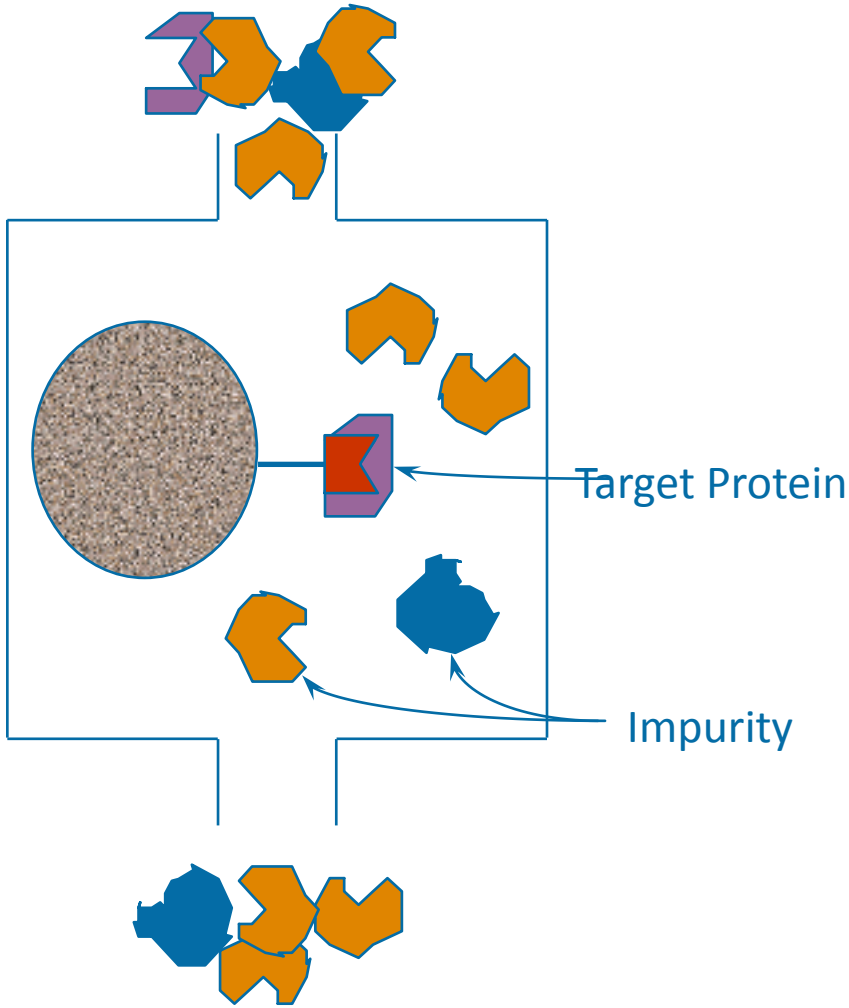


Commonly Used Protein Tags

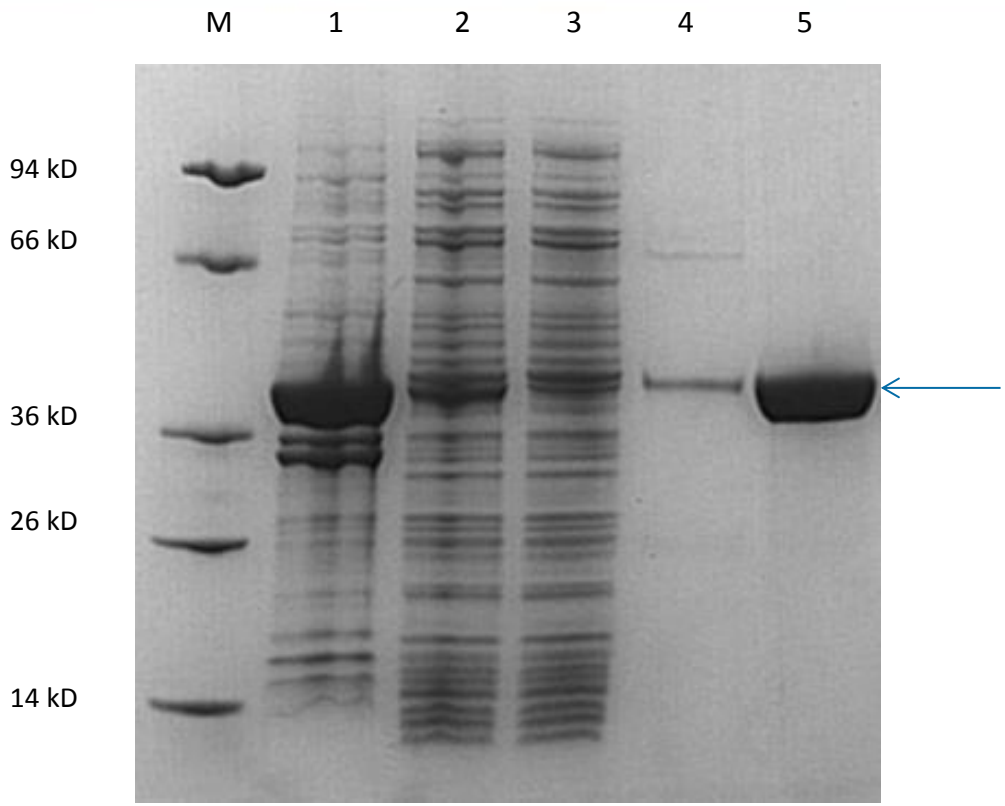
Tags	Amino Acids	Size (KDa)	Matrix	Comments
His	6-8	~1	Nickel	Most commonly used
FLAG	8	~1	Anti-FLAG	Specific
MBP	396	40	Amylose	↑ solubility; periplasm
GST	211	26	Glutathione	Dimerization
NusA	495	54.9	N/A	↑ solubility
Trx	109	11.7	N/A	↑ Disulfide bond
DsbA	208	23.1	N/A	↑ Disulfide bond; periplasma
SUMO	101	11.6	N/A	↑ solubility
Z-tag	91	10.6	Protein A	↑ solubility

Workflow of Tagged Protein Purification



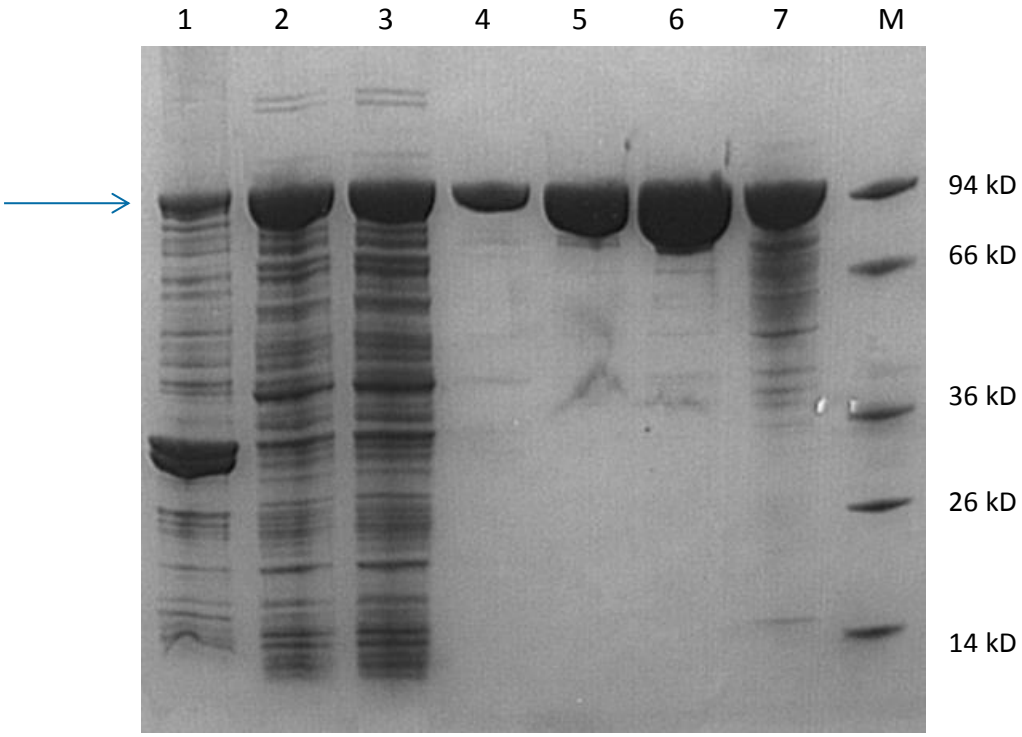


Case Study: One Step His-tag Nickel Purification



Lane:
1: Precipitation after cell lysate centrifugation
2: Supernatant after cell lysate centrifugation
3: Flow through
4: Wash with 50 mM imidazole
5: Elute with 500 mM imidazole

Case Study: One Step MBP-Amylose Purification



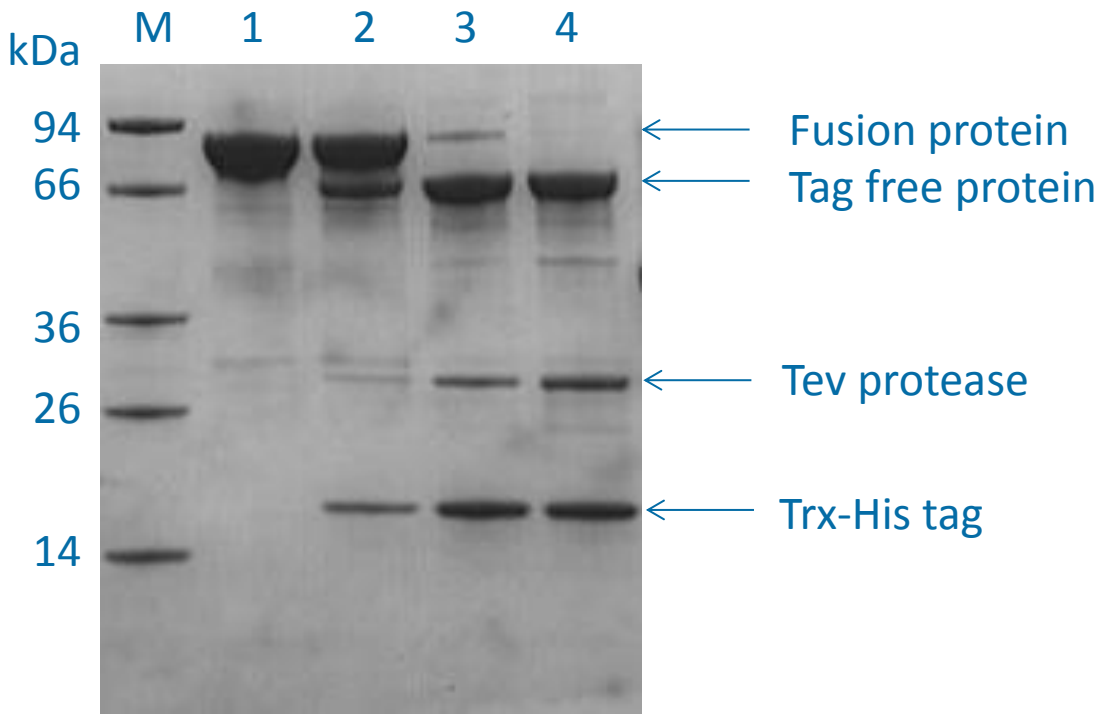
Lane:
1: Precipitation after cell lysate centrifugation
2: Supernatant after cell lysate centrifugation
3: Flow through
4-6: Elute with amylose gradient
7: Resin

Commonly Used Proteases

Protease	Recognition sequence	Amino acid residues	Temperature range (℃)	pH Range	Molecular weight
Tev	ENLYFQ^G	G	4~30	5.5~8.5	28KDa
EK	DDDDK^	None	4~37	6~9	22.7KDa
Thrombin	LVPR^GS	GS	4~37	~7	36KDa
3C	LEVLFQ^GP	GP	4~15	7~8.5	42KDa
SUMO	Small ubiquitin-like modifier	None	4~30	5.5~9.5	25KDa,28KDa

Case Study: Trx Tag Removal by TEV Protease

4°C, overnight



SDS-PAGE analysis of Tev protease digestion of Trx-His protein followed by Coomassie Blue Staining

Table of Contents

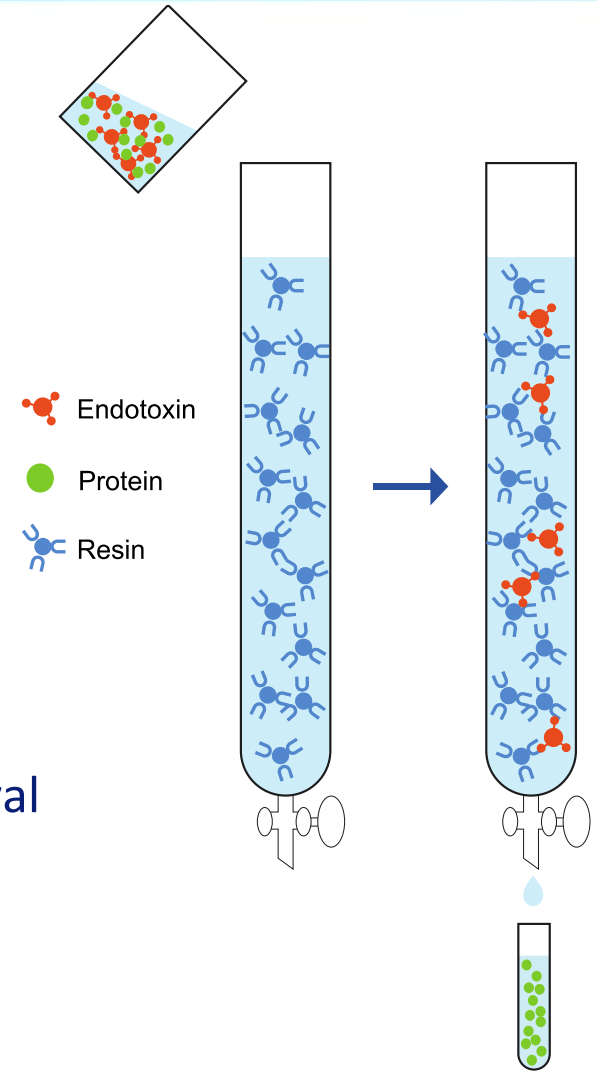
Endotoxin Removal

- **What are endotoxins?**

  - Endotoxins, also known as lipopolysaccharides (LPS), are large molecules found in the outer membrane of Gram-negative bacteria, which elicit strong immune responses in animals.

- **Detection**
  - Gel clot method
  - Chromogenic method

- **Removal methods:**
  - Polymyxin B (PMB) – affinity based ToxinEraser™ (L00338) by GenScript allows highly efficient removal of endotoxin down to 0.1 EU/ml
  - Size exclusion chromatography (SEC)
  - Ion exchange chromatography (IEC)



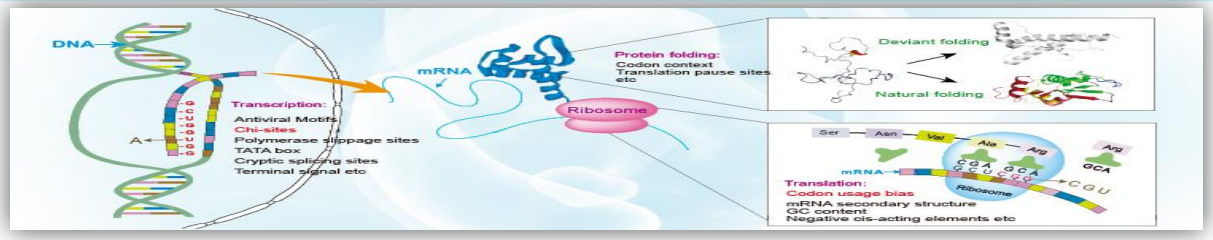
Case Study: Endotoxin Removal by PMB

GenScript Endotoxin Removal Services: Endotoxin <= 1 EU/ug; <= 0.1 EU/ug; <= 0.01 EU/ug

	Protein 1	Protein 2	Protein 3
Volume of resins	3 ml	3 ml	3 ml
Volume of sample	15 ml	15 ml	15ml
Initial endotoxin	500,000 - 2,000,000 EU/ml	> 40,000,000 EU/ml	> 40,000,000 EU/ml
Final concentration	2.2 mg/ml	1 mg/ml	0.8 mg/ml
Final endotoxin	64 – 128 EU/ml	20 – 40 EU/ml	12.5 – 25 EU/ml
Final endotoxin	0.029 –0.058 EU/μg	0.02 –0.04 EU/μg	0.016 – 0.032 EU/μg

Table of Contents

Advantages of GenScript Protein Services



- **Core in-house technologies for expression optimization & production efficiency.**

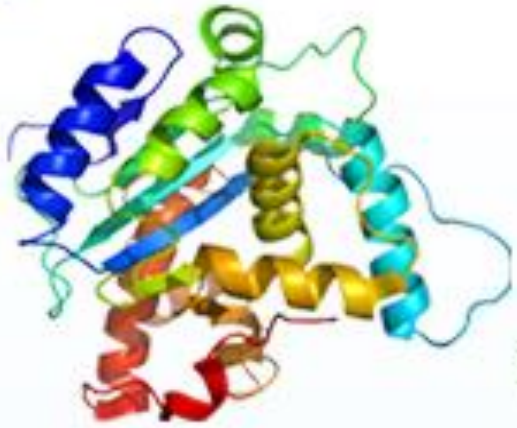
- OptimumGene™ – expression system specific codon optimization
- BacPower™ – increase bacterial soluble expression
- FoldArt™ – ensure appropriate protein refolding
- YeastHigh™ – high copy-number gene selection technique
- BacuVance™ – for protein secretion from baculovirus-infected insect cells
- MamPower™ – technology licensed from NRC for rapid recombinant protein production with high yield
- DoubleTag strategy – for big protein isolation

Advantages of GenScript Protein Services

◆ **One-stop service from sequence to purified proteins with large capacity.**

Guaranteed Protein Expression Package

- **3 mg** purified soluble protein from **$2,200**
- Subcloning
- Transformation
- Expression
- Refolding



Protein Expression and Purification Services

- Bacteria
- Yeast
- Baculovirus/insect cells
- Mammalian

Bioprocessing Services

- Mammalian protein expression services
- Stable cell line development & protein production

Large-scale Protein Production Services

- Bacterial fermentation up to 1,000 L
- Yeast fermentation up to 500 L
- Baculovirus/insect cell production up to grams
- Mammalian cell production up to grams

OptimumGene™ Gene Design Service

——Minimum 10-fold increase in protein express level

Advantages of GenScript Protein Services

- Flexible production scales
- Fast turn-around time (from sequence to purified protein in as little as 4 weeks)

Capacity:

Bacteria	Yeast	Baculovirus	Mammalian
1,000 L	500 L	100 L	500 L
Fermentor	Fermentor	Wave™ Mixer	Wave™ Mixer Sartorious bioreactor Hyclone SUB bioreactor

Variety of GenScript Protein Services

Protein Expression & Purification

Bacteria, Insect, Mammalian, Yeast, Customized protein services, Fermentation, Transient expression, Stable cell lines

Expression Evaluation & Optimization

Small scale expression testing and optimization in bacterial, insect and mammalian expression systems

Large Scale Production

Scale up production in Bacteria, Yeast, Insect and Mammalian cells, Large-scale Recombinant Mab Production

BioProcess

Upstream & Downstream Process development, GLP-compliant Bioprocess Services

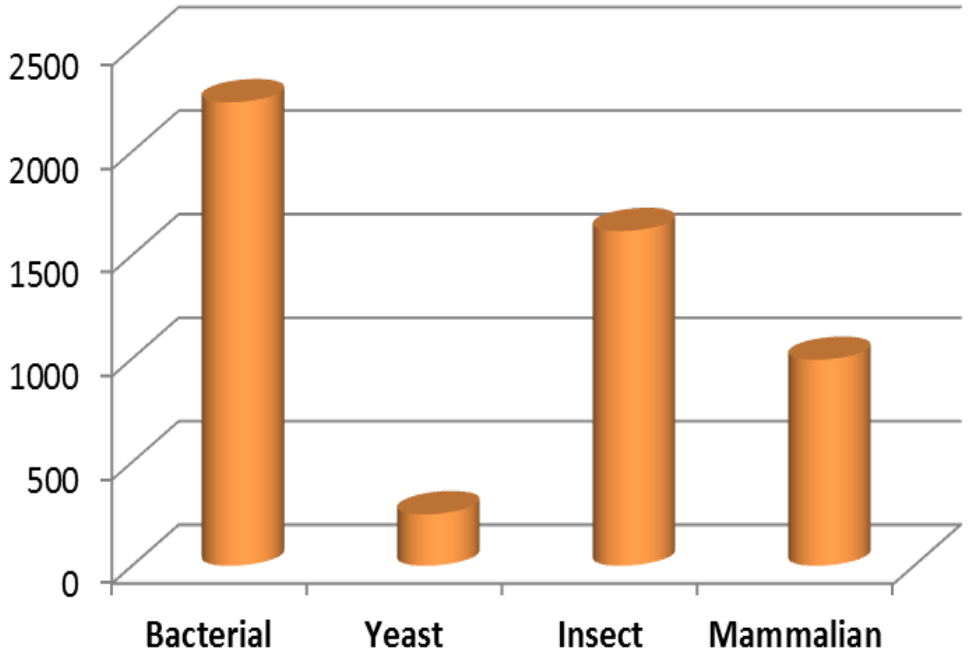
Other Protein Services

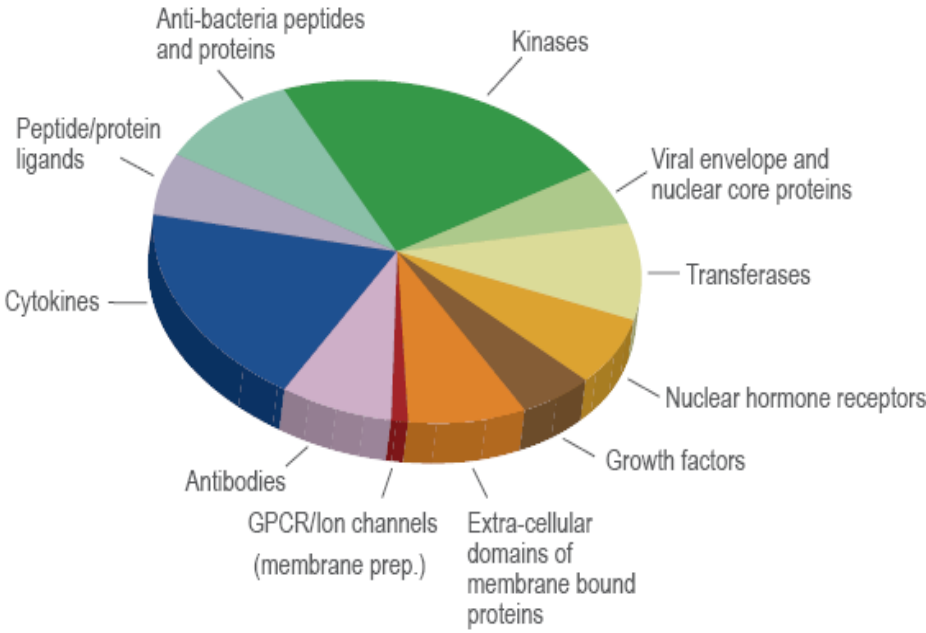
Structural Biology, HTP Variants, Chemical Protein Synthesis, Endotoxin removal, Codon optimization, Custom protein purification, Protein characterization, Protein Refolding

Resources and Technical References

PSORT II, WoLF PSORT II, Case Studies, Downloads, FAQs, ProtBank™ Protein Database Protein News Page, Publications, Service Selection Guide

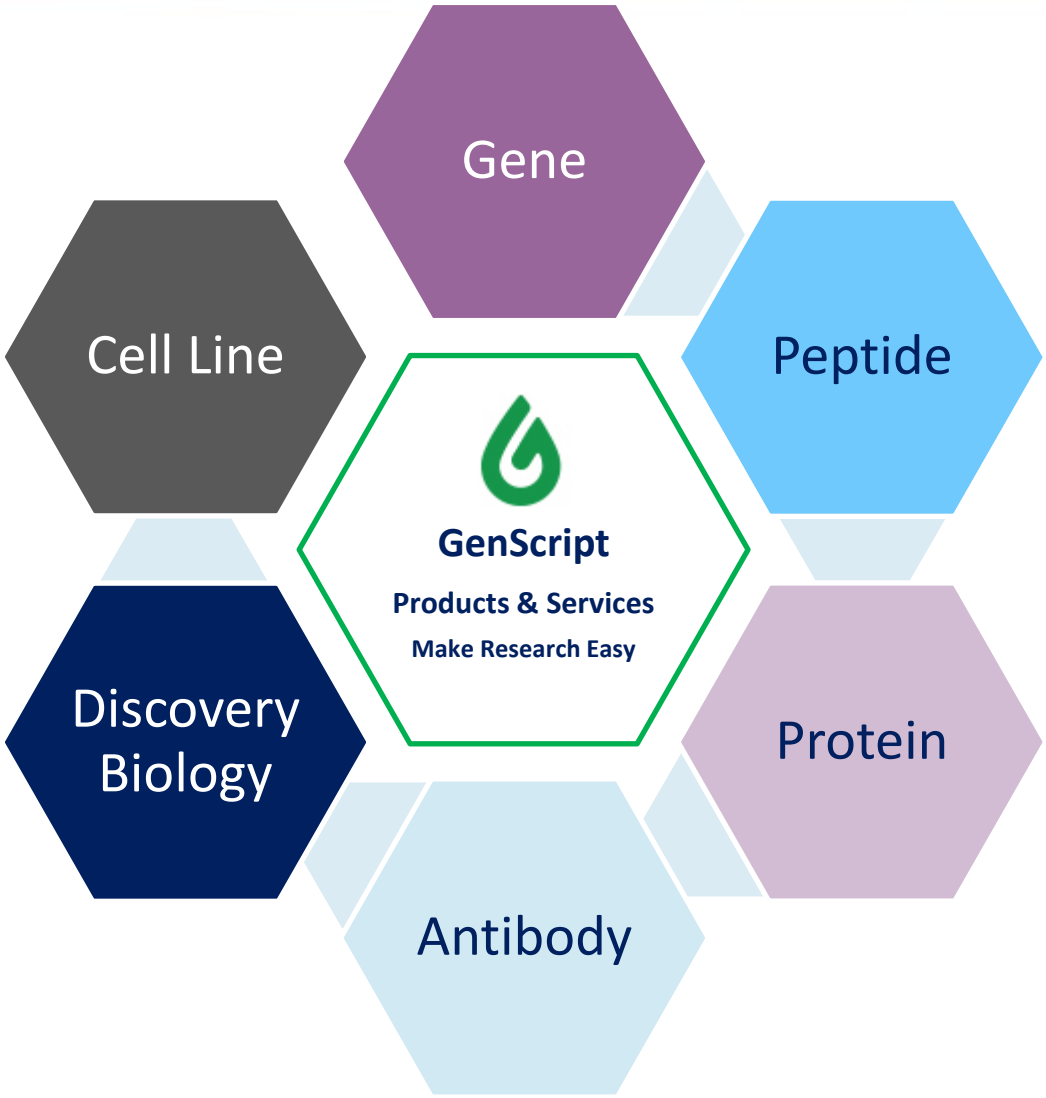
GenScript's Experience in Protein Expression & Purification





GenScript has delivered over **5,000** proteins in four expression systems. Statistics showed **95%** success rate for all protein projects.

About GenScript



Thank you for your participation

We wish you all success in your Research

Register for other webinars in the GenScript Webinar Series @ http://www.genscript.com/webinars.html

Anti-idiotypic antibodies – A powerful tool for antibody drug studies

Need a reliable tool for your antibody drug PK/PD or immunogenicity studies? Anti-idiotype antibodies (anti-IDs) serve as powerful tools for such studies, as they are able to differentiate antibody drug from endogenous antibodies and allow tracking of antibody drug in biological fluid. However, it has been challenging to develop highly specific and sensitive anti-IDs due to the low percentage of anti-IDs relative to other naturally-occurring antibodies in immunized animals. This webinar will take an in-depth look into the field of anti-IDs, including the obstacles and solutions surrounding the generation of highly sensitive, highly specific anti-IDs, as well as their use in vaccine development, PK/PD and immunogenicity studies.

November 20, 2014/ 8:00 am

Register now

November 20, 2014/ 2:00 pm

Register now

Presented by: Michelle Parker, Ph.D.

Peptide design strategy: basics, optimization, and application

Is your peptide sequence optimally designed for your unique research application? Peptide design effects the overall stability and usability of your peptide in experiments, as well as the turnaround time for your custom peptide. Is this session, we will review the basics of peptide design strategy including sequence design, modification selection, and purity choice. Case studies, best practices, and design tools will be introduced. In addition, we will explore specific trends in peptide design as it applies to research applications including drug discovery, drug delivery, structural biology, and biomaterials.

November 25, 2014/ 8:00 am

Register now

November 25, 2014/ 2:00 pm

Register now

Presented by: Tiffany Gupton Campolongo, Ph.D.

Fusion partner for recombinant soluble protein production in *E. coli*

Are you working on an E. coli expression project and looking for ways to generate soluble protein? If so, this interactive webinar may be for you. Target protein solubility is one of the most important end goals of a recombinant protein production campaign yet it remains a significant bottleneck. While there are various ways to optimize expression in order to produce soluble protein, a fusion partner based approach can be very useful. This webinar reviews the practical aspects of a fusion partner strategy. Topics include but are not limited to - the need for protein solubility, the relationship between solubility and protein functionality, use of soluble fusion partners, examples of commonly used partners, identifying novel fusion partners and pros and cons of a fusion partner approach.

December 3, 2014/ 8:00 am

Register now

December 3, 2014/ 2:00 pm

Register now

Presented by: Keshav Vasanthavada