Protein Expression and Purification

James Chou

BCMP201  Spring 2008
The *lac* repressor inhibits the expression of genes coding for proteins involved in the metabolism of lactose in bacteria; it is inactivated by IPTG.
# Bacterial Strains for Protein Expression

<table>
<thead>
<tr>
<th>Bacterial Strain</th>
<th>Features</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL21(DE3) (Novagen)</td>
<td>deficient in lon and omp-t proteas es</td>
<td>Non-toxic proteins</td>
</tr>
<tr>
<td>BL21(DE3)-pLysS (Novagen)</td>
<td>reduces basal level expression of the gene of interest</td>
<td>More toxic proteins</td>
</tr>
<tr>
<td>C41(DE3) (Lucigen)</td>
<td>prevents cell death associated with expression of many toxic proteins</td>
<td>Toxic membrane proteins</td>
</tr>
<tr>
<td>C43(DE3) (Lucigen)</td>
<td>Effective in expressing toxic and membrane proteins from all classes of organisms, including viruses, eubacteria, archaea, yeasts, plants, insects, and mammals</td>
<td>Toxic, difficult proteins</td>
</tr>
</tbody>
</table>
The Luria Broth (LB) rich medium

10 g tryptone (peptide digested from casein by trypsin)

5 g yeast extract

5 g NaCl

in 1 litre with distilled water
The minimal (M9) medium

Components for 1 L

6.8 g Na$_2$HPO$_4$
3.0 g KH$_2$PO$_4$
1.0 g NH$_4$Cl
0.5 g NaCl
4.0 g Glucose
2 mM MgSO$_4$
0.1 mM CaCl$_2$

ph ~ 7.2

sole source of N, use $^{15}$NH$_4$Cl for NMR studies

sole source of C, use $^{13}$C-glucoes for NMR studies

1mM of antibiotics, e.g., ampicillin, kanamycin

Supplement with 1 tablet
Centrum A to Zinc
The standard cell growth conditions

37 °C
flask volume > 4 x medium
shake ~300 rpm
OD$_{600}$ limited by O$_2$

Constant feed of O$_2$, NH$_4$OH, glucose
Can reach OD$_{600}$ ~ 50 !!
A typical cell growth curve

- **Lag phase**
- **Log phase**
- **Stationary phase**
- **Decline phase**

**E. coli growth curve**

- **IPTG induction**
IMPORTANT: always screen for expression conditions
Example of expression vector - His-tag (Novagen)
General protocol for purification of His-tagged proteins

Lyse cell

Spin the lysate at 15,000 g for 30 min

Load supernatant to Ni-NTA column

Wash with buffer containing ~10mM imidazole

Elute protein with 300 mM imidazole

Ni NTA resin

Imidazole, competitive binder of His tag
Native conditions

- Tris or phosphate buffer, pH 8
- 300 mM NaCl
- 10–20 mM imidazole

30–60 min (Batch or column format)

Denaturing conditions

- Phosphate buffer, pH 8
- 8 M urea or 6 M GuHCl (imidazole optional)

15–30 min (Batch or column format)
Further purify by size exclusion
Example of expression vector - GST fusion

- GST (26 kDa)
- Protein of interest

protease site

Vectors from Amershan

- pGEX-2TK
  - Thrombin (37 kDa)
- pGEX-5X-1
  - Factor Xa (48 kDa)
- pGEX-6P-1
  - Precision protease (46 kDa)
Glutathione affinity purification

Glutathione

![Glutathione structure](image)

Glutathione Sepharose

Glutathione S-transferase (GST)

![GST structure](image)

Equilibrate column with binding buffer

Apply sample wash with binding buffer

Elute with elution buffer

3 min

5–15 min

2 min

Waste

Collect

Collect fractions
A)  
**Column:** GSTrap FF 1 ml  
**Sample:** 8 ml cytoplasmic extract from *E. coli* expressing a GST fusion protein  
**Binding buffer:** PBS, pH 7.3  
**Elution buffer:** 50 mM Tris-HCl, pH 8.0 with 10 mM reduced glutathione  
**Flow:** 1 ml/min  
**Chromatographic procedure:**  
- 4 CV binding buffer, 8 ml sample,  
- 10 CV binding buffer, 5 CV elution buffer, 5 CV binding buffer (CV = column volume)  

---  

B)  
**Lane 1:** Low Molecular Weight (LMW) Calibration kit, reduced, Amersham Biosciences (10 µl prepared for silver stain)  
**Lane 2:** Cytoplasmic extract of *E. coli* expressing GST fusion protein, 1 g cell paste/10 ml (5 µl sample from collect. fraction + 35 µl sample cocktail -> 10 µl applied)  
**Lane 3:** GST fusion protein eluted from GSTrap FF 1 ml (5 µl sample from collect. fraction + 35 µl sample cocktail -> 10 µl applied)
Further purify by size exclusion
Example of expression vector - MBP fusion

<table>
<thead>
<tr>
<th>MBP (45 kDa)</th>
<th>Protein of interest</th>
</tr>
</thead>
</table>

Tobacco etch virus (TEV) protease cleavage site

Maltose Binding Protein (MBP)
Purification of MBP-fusion by amylose/dextrin column

Fig 1. MBPTrap HP 1 ml and 5 ml columns give fast and convenient affinity purifications of recombinant proteins tagged with maltose binding protein.

Sepharose™ High Performance based on a 34 µm bead-sized matrix.
Columns: MBPTrap HP 1 ml
MBPTrap HP 5 ml

Sample: MBP2*-β-galactosidase (M, ~158 000) in E. coli lysate

Sample volumes: 2 ml (1 ml column)
10 ml (5 ml column)

Flow rates: 1.0 ml/min (0.5 ml/min during sample loading and wash)
(1 ml column)
5.0 ml/min (2.5 ml/min during sample loading and wash)
(5 ml column)

Binding buffer: 20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 7.4

Elution buffer: 10 mM maltose in binding buffer

System: ÄKTAexplorer
Introduction to Chromatography

Gel filtration  Hydrophobic interaction  Ion exchange  Affinity  Reversed phase
The Theory of Chromatography

Plate Theory - Martin & Synge 1954

View column as divided into a number \( N \) of adjacent *imaginary* segments called theoretical plates within each theoretical plate. Complete equilibration of analytes between stationary and mobile phase occurs.

Sample injection

Mobile Phase \( \leftrightarrow \) Stationary Phase
Predicting elution profile using the plate theory

\[ V_i = \text{volume of phase } i \]
\[ X_i = \text{analyte conc. in phase } i \]
\[ N = \text{number of plates} \]
\[ \text{d}X_{\text{Stationary}} = K \text{d}X_{\text{Mobile}} \]
\[ \Delta V = \text{volume of solution from one plate to the next} \]

\[ \Delta (\text{mass}) = \left( X_{M(p-1)} - X_{M(p+1)} \right) \Delta V \]

\[ \ldots \]

Elution profile is a Gaussian function when \( N \) is large
Bandwidth of the elution profile

\[ t_R = \text{retension time of protein, the time between sample injection and an analyte peak reaching a detector at the end of the column} \]

\[ t_M = \text{retension time of the mobile phase (for reference)} \]

\[ N = 5.55 \frac{t_r^2}{w_{1/2}^2}, \quad w_{1/2} \text{ is width of the elution peak at half-height} \]
Predictions from the Plate Theory

\[ N = 5.55 \frac{t_r^2}{w_{1/2}^2} \]

Greater separation occurs:
– with greater number of theoretical plates (N)
– as plate thickness becomes smaller

The width of bands increases as their retention volume increases

The thinner the plate, the narrower the eluted peak
Column Dimensions

Column length $L = n \times$ thickness of the plates

$N_{\text{max}} = 0.4 \times \frac{L}{dp}$

$N$ - maximum column efficiency (# of plates)
$L$ - column length
$dp$ - particle size

So, the smaller the particle size the higher the efficiency!
Ion exchange chromatography
Common functional groups used on ion exchangers

Weak:
- DEAE-anion exchanger

Strong:
- Q-anion exchanger
- CM-cation exchanger
- S-cation exchanger
Common resin material

Sepharose -- crosslinked, beaded-form of agarose
# Sepharose ion exchangers

Table 9. Characteristics of Q, SP, DEAE and CM Sepharose Fast Flow.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of gel</td>
<td>strong anion</td>
<td>strong cation</td>
<td>weak anion</td>
<td>weak cation</td>
</tr>
<tr>
<td>Total ionic capacity</td>
<td>180-250</td>
<td>180-250</td>
<td>110-160</td>
<td>90-130</td>
</tr>
<tr>
<td>(µmol/ml gel)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recommended working flow</td>
<td>100-300</td>
<td>100-300</td>
<td>100-300</td>
<td>100-300</td>
</tr>
<tr>
<td>rate range (cm/h)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approx. mean particle</td>
<td>90</td>
<td>90</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>size (µm)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Particle size range</td>
<td>45-165</td>
<td>45-165</td>
<td>45-165</td>
<td>45-165</td>
</tr>
<tr>
<td>(µm) working pH range*</td>
<td>2-12</td>
<td>4-13</td>
<td>2-9</td>
<td>6-10</td>
</tr>
<tr>
<td>pH stability**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>short term</td>
<td>1-14</td>
<td>3-14</td>
<td>1-14</td>
<td>2-14</td>
</tr>
<tr>
<td>long term</td>
<td>2-12</td>
<td>4-13</td>
<td>2-13</td>
<td>4-13</td>
</tr>
</tbody>
</table>
MonoBeads Exchanger

10 μm beaded hydrophilic polystyrene/divinyl benzene resin, pH stable, porous

Mono Q - substituted with quaternary amine groups to yield the strong anion exchanger

Mono S - substituted with methyl sulphonate groups to yield the strong cation exchanger
pl of the protein and pH of the mobile phase is important

Example

<table>
<thead>
<tr>
<th>Protein</th>
<th>pl</th>
<th>pH 4.8 cm</th>
<th>pH 7.2 cm</th>
<th>pH 8 cm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carbonic Anhydrase</td>
<td>7.0</td>
<td>+16.5</td>
<td>-0.4</td>
<td>-2.7</td>
</tr>
<tr>
<td>Carboxypeptidase B</td>
<td>6.2</td>
<td>+12.0</td>
<td>-3.3</td>
<td>-6.3</td>
</tr>
<tr>
<td>Chymotrypsin</td>
<td>8.0</td>
<td>+9.0</td>
<td>+2.7</td>
<td>0.0</td>
</tr>
<tr>
<td>Lysozyme</td>
<td>9.8</td>
<td>+14.1</td>
<td>+7.9</td>
<td>+6.9</td>
</tr>
</tbody>
</table>
Column: SOURCE 30S, a) 7.5 mm i.d. x 50 mm (2.2 ml)
b) 200 mm i.d. x 50 mm (1.57 l)
Sample: Mixture of chymotrypsinogen, cytochrome C and lysozyme
Sample load: 0.32 mg/ml bed volume
Eluent A: 20 mM sodium phosphate, pH 6.8
Eluent B: 20 mM sodium phosphate + 0.5 M NaCl, pH 6.8
Flow rate: 300 cm/h ; a) 2.2 ml/min b) 1.57 l/min
Gradient: 0-100% B; 20 column volumes
System: a) FPLC System
b) BioProcess Engineering System
Size exclusion chromatography

1. Spherical particles of gel filtration medium are packed into a column.
Larger molecules move faster through the column
Example: desalting of a protein sample

Sample: (His)$_6$ protein eluted from HiTrap™ Chelating HP with sodium phosphate 20 mM, sodium chloride 0.5 M, imidazole 0.5 M, pH 7.4

Column: HiTrap Desalting 5 ml

Buffer: Sodium phosphate 20 mM, sodium chloride 0.15 M, pH 7.0

void volume $V_0$, total column volume $V_t$
Optimize resolution

Factors affecting resolution

- ratio of starting sample volume to the column volume
- flow rate
- column dimension
- particle size, pore size, packing density
Smaller sample volume -> better separation

Results from the Superdex 200 HR 10/30 column
Longer column increases resolution

Resolution in chromatography

\[ R_S = \frac{V_{r1} - V_{r2}}{(w_1 + w_2)/2} \]

\( R_S \) increases as \( \sqrt{\text{column bed height}} \)
Faster flow rate decreases resolution

Results from Amershan HiLoad 16/60 Superdex 30 column
Commonly used columns in protein purification

Amersham HiLoad Superdex 75/200 prep column
Hydrophobic interaction between polypeptide and the stationary phase

Reversed phase chromatography (native)
Modify hydrophobic interactions with salts

Hofmeister series: the effect of anions and cations on protein precipitation

Increase hydrophobic interaction

Anions: $\text{PO}_4^{3-}, \text{SO}_4^{2-}, \text{CH}_3\text{COO}^-, \text{Cl}^-, \text{Br}^-, \text{NO}_3^-, \text{ClO}_4^-, \text{I}^-, \text{SCN}^-$

Cations: $\text{NH}_4^+, \text{Rb}^+, \text{K}^+, \text{Na}^+, \text{Cs}^+, \text{Li}^{2+}, \text{Mg}^{2+}, \text{Ca}^{2+}, \text{Ba}^{2+}$

decrease hydrophobic interaction

Relative effects of some salts on protein precipitation.

$\text{Na}_2\text{SO}_4 > \text{K}_2\text{SO}_4 > (\text{NH}_4)_2\text{SO}_4 > \text{Na}_2\text{HPO}_4 > \text{NaCl} > \text{LiCl} .... > \text{KSCN}$
Yellow: hydrophobic residues

Proteins separated in order of increasing surface hydrophobicity

Column: Phenyl Sepharose High Performance packed in Tricorn 10/100 column
Sample: Cytochrome c, RNAse A, lysozyme, α-chymotrypsin
Start buffer: 1.7 M ammonium sulfate, 0.02 M Tris-HCl, pH 7.5
Elution buffer: 0.02 M Tris-HCl, pH 7.5
Gradient: 0–100% elution buffer in 10 CV
Flow: 1 ml/min, 76 cm/h
Reversed phase chromatography (denatured)

Advantages
- resolution (pure)
- can purify anything

Disadvantages
- harsh solvent condition
- denature protein
Reversed phase HPLC columns are packed with silica beads

- Si - OH
- Si - OH

Pore size (affect binding capacity)
- 100 Å for small peptide
- 300 Å for proteins

Bead size
- 3 - 5 µm for small prep (slow flow)
- > 15 µm for large prep (fast flow)

\[
\text{Si-OH} + \text{Cl-Si-(CH}_2\text{)}_{17}\text{-CH}_3 \rightarrow \text{Si-O-Si-(CH}_2\text{)}_{17}\text{-CH}_3 + \text{HCl}
\]
Modify hydrophobicity of the column by varying the hydrocarbon length

- Residual silanol group
- Ether; source of silanols
- Octadecyl group
- C₂ capping group
Properties of the mobile phase for eluting the protein

The lower the polarity of the mobile phase, the greater its eluting strength. The non-polar organic solvents are used to elute proteins.

The most widely used organic solvents (no UV absorption)

- acetonitrile
- methanol
- isopropanol (viscous, high back pressure)
Strong acids are also used in reversed phase HPLC

Denature the protein at ~ pH 2-3 to expose all hydrophobic groups.

The retention of polypeptides can be modified by mobile phase pH due to ionizable groups in the peptide.
  - suppress the ionization the acidic groups in protein

The most widely used acid
  - trifluoroacetic acid (TFA)
  - ortho-phosphoric acid

Base is normally not used because the silica matrix breaks down above pH 8.
Example of reversed phase HPLC

Sample: Synthetic peptide: NVILTKPEVSEGTEVT
Column: Sephasil Peptide C18 5 µm ST 4.6/100
Eluent A: 0.06% trifluoroacetic acid (TFA), pH 2.5
Eluent B: 84% acetonitrile in 0.055% TFA
Gradient: (A) 0 - 60% (B) 20 - 35%
Flow: 1 ml/min