Protein Production and Purification

Dr Ruth Rose
Protein Production Facility

Fogg Building, 4.33, Mile End Campus
020 7882 7603 or ext 4620
http://www.sbcs.qmul.ac.uk/research/facilities/protein-facility/

SBCS-proteinfacility@qmul.ac.uk
Why would you want to study proteins?

- What does the protein do?
- Where is the protein?
  - In the cell (membrane, cytoplasm)
  - Which tissues (muscle, brain)
- Protein engineering
  - Can you make a protein do what it does normally but better?
- Antibodies are proteins too
- Structural studies
  - Design drugs more effectively
- Does the protein move?
  - *E.g.* BACE1 and CCS move together down a neuron
From DNA to proteins

---

**Starting material**
- mRNA/cDNA
- Synthesised DNA
- Plasmid DNA
- PCR product
- Annealed oligos
- Gibson assembly

**Expression system**
- E. coli
- Yeast
- Insect cells
- Mammalian cells

**Purification**
- Affinity chromatography (His, MBP, GST, protein A/G)
- Size exclusion
- Hydrophobic interaction (HIC)

**End use**
- Pull downs
- Structural investigations
- Antibody generation/ELISA
- In vivo studies
- Sequence analysis

Will the tag need to be removed?
Key decisions

• **Talk to your supervisor and future proof your plans**
  Cloning can take a long time so try and minimise the number of steps and the number of times you have to reclone

• **Availability**
  Many people will use certain vectors or host systems for historical or financial reasons – respect these but ask around…. Someone might have the plasmid of your dreams

• **Communication**
  If you get stuck – don’t stress
  **ASK** – someone will be able to give you some ideas
Starting material

**Plasmid DNA**

Traditional sub-cloning
Restriction enzymes to excise the fragment
Ligase to anneal into complimentary vector
PCR template
Incorporate new restriction enzymes for traditional cloning or flanking region for InFusion cloning or recombineering

**mRNA/cDNA**

PCR template
Incorporate new restriction enzymes or flanking region for InFusion cloning or recombineering
Blunt/TA cloning = blue/white screening
Insert PCR directly into a plasmid (e.g. pGEM T-easy, pbluescript, TOPO)
Starting material

**PCR product**
- PCR template cloning
  - Incorporate new restriction enzymes or flanking region for InFusion cloning or recombineering
- Blunt/TA cloning
  - Insert PCR directly into a plasmid (e.g. pGEM T-easy, pbluescript, TOPO)

**Annealed oligos**
- Gibson assembly
  - Overlapping sequence regions allow construction of larger sequence
- Insertion of short sequences
  - Create sticky overhands or use blunt end sequences
Full cloning scheme

1. Cell harvest
2. Resuspension
   Lysis
   Neutralization
3. Column binding
4. Column washing
5. DNA elution
6. Pure DNA

Send for sequencing

Plate onto agar plate containing antibiotic, add IPTG and X-gal for blue white screen

Pick individual colonies

Isolate plasmid DNA

miniprep

Culture colonies
## Vectors

<table>
<thead>
<tr>
<th>Vector</th>
<th>Promoter</th>
<th>Tag</th>
<th>Terminal</th>
<th>Advantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>pGEM T-easy/ pBluescript/ PCR clone/ TOPO TA</td>
<td>T7</td>
<td>none</td>
<td></td>
<td>Blunt or TA overhang. Good for PCR product</td>
</tr>
<tr>
<td>pOPIN</td>
<td>CMV/T7</td>
<td>Various</td>
<td>N or C</td>
<td>Can be used for bacterial, insect and mammalian</td>
</tr>
<tr>
<td>pET</td>
<td>T7</td>
<td>6xHis</td>
<td>N- C-</td>
<td>Quick, flexible, wide range of sites in mcs.</td>
</tr>
<tr>
<td>pGEX</td>
<td>Tac</td>
<td>GST</td>
<td>N-</td>
<td>Promotes solubility, tag cleavage options</td>
</tr>
<tr>
<td>pMAL</td>
<td>Tac</td>
<td>MBP</td>
<td>N- or C-</td>
<td>Promotes solubility</td>
</tr>
<tr>
<td>Many e.g. pcDNA, pTriEx</td>
<td>CMV</td>
<td>Myc or FLAG</td>
<td>N- or C-</td>
<td>Used for multiple hosts, suitable for IP, WB etc.</td>
</tr>
</tbody>
</table>

- Many more tags are available to promote solubility, stability or purification
- Can include epitope tags for western blotting
- check that your sequencing company stocks suitable primers
Before you go any further….

• Plan your cloning thoroughly before you start
  – Design your primers: Will your protein be in frame? Check reverse primer is reverse compliment
  – Does your gene need optimising for the organism you are going to be using?
  – Draw a map
  – Sequencing primers
  – Do you have all of the enzymes, vectors and cell lines to hand?
  – Translate the sequence….will it produce what you are expecting?
Getting to Know Your Protein

• Translated sequence, including tag
• MW = size in Daltons
  – If you need to cleave the tag, is it sufficiently different in size from the tag?
• pI = pH where protein has no electric charge. Value can affect solubility
• Extinction coefficient
• Instability index = value <40 classifies protein as stable
• GRAVY = more negative the less soluble
Some Helpful Websites

- Analysing primers: IDT DNA
- For drawing maps: http://tools.neb.com/NEBcutter2/
- Analysing sequences:
  http://pga.mgh.harvard.edu/web_apps/web_map/start
- Aligning sequences: http://multalin.toulouse.inra.fr/multalin/
- SnapGene, Enzyme X, vector builders

- Helpful information: ThermoScientific Molecular Biology solutions (molecular biology tools and product guide)

  New England Biolabs Catalogue
  pET manual (Novagen),
  Molecular Biology: Sambrook et al.
## Expression systems

<table>
<thead>
<tr>
<th>Host cell</th>
<th>e.g.</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli</strong></td>
<td>BL21(DE2)pLysS, Rosetta, SHufFLe</td>
<td>Cheap, quick, high cell densities, relatively easy, very easy to insert a vector, stable transformant</td>
<td>No post-translational modifications</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Alter time post induction to control protein production.</td>
<td></td>
</tr>
<tr>
<td><strong>Yeast</strong></td>
<td><em>Saccharomyces cerevisae</em></td>
<td>Cheap, quick, high cell densities, relatively easy, once transformant created it is relatively stable</td>
<td>Post-translational modifications, not as fast as <em>E. coli</em>. Takes time to introduce a vector</td>
</tr>
</tbody>
</table>
# Expression systems

<table>
<thead>
<tr>
<th>Host cell</th>
<th>e.g.</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insect cells</td>
<td>Sf9, Sf21</td>
<td>Grow relatively easily, high yield of proteins that have been post-translationally modified</td>
<td>Cloning and introduction via baculovirus (requires optimisation!)</td>
</tr>
<tr>
<td>Mammalian carcinoma</td>
<td>HeLa, HEK, COS, CHO, Expi293</td>
<td>Grow relatively easily, once transformant created it is relatively stable</td>
<td>Cell metabolism not strictly in native state, can be tricky to introduce vector</td>
</tr>
<tr>
<td>Mammalian primary culture</td>
<td>Neurons, breast tissue, stomach</td>
<td>Wild type cells – as close to the native state as possible</td>
<td>Extremely difficult to culture, requires extraction from animal/patient (ethics), short life time, expensive, skilled</td>
</tr>
</tbody>
</table>
Optimising Protein Production - Example

a. Western Blot of his-tagged *MthCbiA1* over-produced in BL21(DE3)pLysS

b. SDS-PAGE gel of his-tagged *MthCbiA1* over-produced in Rosetta(DE3)pLysS

c. SDS-PAGE gel of MBP: *MthCbiA1*
Variables that affect protein production

- DNA
  - Codon usage
  - mRNA structure
  - G + C content
  - Regulatory motifs
  - Repeats

- Vector
  - Replication origin
  - Promoters
  - RBS
  - Regulatory elements
  - Terminators

- Strain
  - Drug resistance
  - Protease deficiency
  - Redox environment
  - Recombinases
  - Polymerases

- Production
  - Temperature
  - Carbon source
  - Nutrients
  - Aeration
  - pH
Purifying Protein

- Use minimal purification steps to avoid loss or denaturation of protein.

- Binding affinity
  - Nickel, Maltose, Glutathione, Protein G/A

- Size
  - Dialysis
  - Gel filtration

- Charge
  - Ion exchange
Separated cell extract containing your protein of interest = His-protein1

Add high concentration of imidazole, to displace histidine residues

Wash column to remove non-specific binding

Wash

• Method is quick, effective, cheap and easy.
• Protein is largely suitable for enzyme assays but purity must be checked beforehand.

Elution fractions

50 kDa ➔

20 kDa ➔

80 kDa ➔

s/n ladder pellet F/T wash

His-protein1

Ni - imidazole

Ni - imidazole

Ni - imidazole

Ni - imidazole

• Method is quick, effective, cheap and easy.
• Protein is largely suitable for enzyme assays but purity must be checked beforehand.

SBCS-proteinfacility@qmul.ac.uk
Purifying Protein – Size

Protein solution

Small molecules enter the aqueous space within the beads

Porous bead, made of polymer such as dextra, agarose or polyacrylamide

Large molecules can’t enter the beads so move through more quickly

- Gel filtration chromatography
- Separate according to size
- Accounts for native state of protein complex
- Large molecules flow more rapidly through column
- Need FPLC
Purifying Protein – FPLC

- **FPLC** = Fast Protein Liquid Chromatography

Calculating protein concentration:

\[
\frac{A_{280} \times \text{mw}}{\text{Extinction coefficient} \times \text{pathlength}}
\]
Purifying Protein – Tag Cleavage

- Range of proteases available
  - e.g. PreScission™, Factor Xa, Thrombin, TEV
- How will you separate protease from your protein?
- Optimise cleavage
- Which buffer does cleavage take place in?
  Will your protein cope?

![SDS-PAGE gel of MBP cleavage from MthCbiA1]
Preparing your sample for use – buffer exchange

- G25 Sepharose/PD-10 column
  - Work in same way as gel filtration
  - Gravity flow column
  - Quick, easy
  - Expensive, diluting method

- Dialysis
  - Small molecules transverse semi-permeable membrane
  - Cheap, variety of membranes
  - Slow, very diluting method
Purifying Protein – Difficult Proteins Need a Different Approach

- Buffers – MOPS, PIPES, HEPES, Phosphate...
- pH – check pl
- Solubility – detergents, different tag, lysis method
- Multimer stability – salt concentration
- Poor stability – 4 °C, metals
- Protein degradation – protease inhibitors
- Contaminant – different purification method e.g. ion exchange
Protein Production Summary

• Plan your cloning thoroughly before you start
• Check the nature of the protein that you will create
• Optimise protein production
• Monitor every step throughout (SDS-PAGE)
• Select minimal number of purification methods
• Optimise tag cleavage where necessary
From DNA to proteins

**starting material**
- mRNA/cDNA
- Synthesised DNA
- Plasmid DNA
- PCR product
- Annealed oligos

**expression system**
- E. coli
- Yeast
- Insect cells
- Mammalian cells

**purification**
- Affinity chromatography (His, MBP, GST, protein A/G)
- Size exclusion
- Hydrophobic interaction (HIC)

**End use**
- Pull downs
- Structural investigations
- Antibody generation/ELISA
- In vivo studies
- Sequence analysis

Will the tag need to be removed?
Why study proteins?

- Possibilities are endless and dependent on your experimental design and outcomes

Gel electrophoresis

Cell imaging

Yeast assays

Animal models

X-ray Crystallography

Assays

Why study proteins?
What else is available in the facility?

• SNP analysis
  • Using subcloning
• Western blotting
• Antibody purification
• Anaerobic experimentation
If in doubt…. Ask!

Dr Ruth Rose
Protein Production Facility

Fogg Building, 4.33, School of Biological and Chemical Sciences, Mile End Campus
020 7882 7603 or ext 4620

http://www.sbcs.qmul.ac.uk/research/Research%20facilities/Protein%20facility/23820.html