PROTEIN PURIFICATION

References:

Books:


Journal Articles:


Methods of Separation: See Scopes

1. Precipitation
   a. ammonium sulfate
   b. organic extraction
   c. polymer extraction

2. Chromatography
   a. ion exchange (adapted from Ion Exchange Chromatography: Principles and Methods)

Separation in ion exchange chromatography depends upon the reversible adsorption of charged solute molecules to immobilized ion exchange groups of opposite charge. The solute is typically a protein, but nucleic acids (e.g. oligonucleotides), or even nucleotides (e.g. GTP, GDP) can be separated on ion exchange columns. Most ion exchange experiments are performed in five main steps. These steps are illustrated schematically below.

The first stage is equilibration in which the ion exchanger is brought to a starting state in terms of pH and ionic strength, which allows the binding of the desired solute molecules. The exchanger groups are associated at this time with exchangeable counter-ions (usually simple anions or cations, such as chloride or sodium).

The second stage is sample application and adsorption, in which solute molecules carrying the appropriate charge displace counter-ions and bind reversibly to the gel. Unbound substances can be washed out from the exchanger bed using starting buffer.

In the third stage, substances are removed from the column by changing to elution conditions unfavorable for ionic bonding of the solute molecules. This normally involves increasing the ionic strength of the eluting buffer or changing its pH. In Figure 1, desorption is achieved by the introduction of an increasing salt concentration gradient and solute molecules are released from the column in the order of their strengths of binding, the most weakly bound substances being eluted first.

The fourth and fifth stages are the removal from the column of substances not eluted under the previous experimental conditions and re-equilibration at the starting conditions for the next purification.
Typically the ion exchange column is stripped with a buffer containing high salt concentration to elute any material retained on the column. The matrix is then re-equilibrated in the starting buffer of low ionic strength.

Separation is obtained since different substances have different degrees of interaction with the ion exchanger due to differences in their charges, charge densities and distribution of charge on their surfaces. These interactions can be controlled by varying conditions such as ionic strength and pH. The differences in charge properties of biological compounds are often considerable, and since ion exchange chromatography is capable of separating species with very minor differences in properties, e.g. two proteins differing by only one charged amino acid, it is a very powerful separation technique.

The choice of an anion or cation exchange column to separate your protein of interest should be based on the pI of the protein (if known) and the pH range of its stability. When dealing with an unknown protein, this often is determined empirically.

From Freifelder

![Diagram of net charge on protein vs pH](image)

**Figure 8-23**
The net charge of a protein as a function of pH, showing the pH ranges for stability and for binding to cation and to anion exchangers. This protein would be chromatographed on an anion exchanger using a buffer adjusted to pH 6.3–8.4.

<table>
<thead>
<tr>
<th><strong>Anion Exchangers</strong></th>
<th><strong>Functional Group</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Diethylaminoethyl (DEAE)</td>
<td>-O-CH₂-CH₃-N⁺H(CH₃CH₃)₂</td>
</tr>
<tr>
<td>Quaternary aminoethyl (QAE)</td>
<td>-O-CH₂-CH₂-N⁺(C₂H₅)₂-CH₂-CHOH-CH₃</td>
</tr>
<tr>
<td>Quaternary ammonium (Q)</td>
<td>-O-CH₂-CHOH-CH₂-O-CH₂-CHOH-CH₂-N⁺(CH₃)₃</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Cation exchangers</strong></th>
<th><strong>Functional Group</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Carboxymethyl (CM)</td>
<td>-O-CH₂-COO⁻</td>
</tr>
<tr>
<td>Sulphopropyl (SP)</td>
<td>-O-CH₂-CHOH-CH₂-O-CH₂-CH₂SO₃⁻</td>
</tr>
<tr>
<td>Methyl sulphonate (S)</td>
<td>-O-CH₂-CHOH-CH₂-O-CH₂-CHOH-CH₂SO₃⁻</td>
</tr>
</tbody>
</table>
b. Gel Filtration  (adapted from Gel Filtration: Principles and Methods)

In gel filtration, molecules in solution are separated according to differences in their sizes as they pass through a column packed with a chromatographic medium which is a gel. The principle of gel filtration is shown in the figure to the left (taken from Freifelder)

Proteins to be separated are applied in a small volume to the column and eluted with buffer. Very large proteins that cannot penetrate into the pores of the gel will elute in the void volume (Vo) (see figure on next page). Small molecules that freely permeate the gel will elute when a volume of buffer equivalent to the total capacity of the column (Vt) has passed through the column. Intermediate sized molecules will fractionate between these two extremes. If there is some type of interaction (ionic or hydrophobic) between the gel filtration matrix and the protein, such a protein will be retained on the matrix and elute later than predicted by its molecular weight.

The sample size applied to the column must be small relative to Vt. The smaller the size of the sample, the better resolution. In practice, 1-2% of the column volume is an optimal ratio.

Gel filtration is frequently used to exchange buffers in samples – “desalting”. Gel filtration resins come in a variety of pore sizes. A resin with a 10,000 molecular weight cut-off will exclude most macromolecules which will elute in the void volume. Small ions such as salts will permeate the gel. When desalting, the sample size can be up to 10% of Vt.
c. Affinity Chromatography

The principle of affinity chromatography is shown below. A biospecific ligand is covalently attached to a chromatographic bed material. The protein of interest binds to the ligand. Contaminating proteins will not be retained by the column and can be removed by subsequent washes. Elution of the bound protein relies on selective desorption of the protein. This is usually accomplished by adding a solution of free ligand, but other methods can be used. For example, many receptor-ligand interactions are disrupted at low pH.

Affinity chromatography is the most powerful technique for many-fold purification in a single step, but requires a high affinity interaction. Purification schemes for recombinant proteins typically use some form of affinity chromatography – e.g. glutathione-S-transferase binding to glutathione.
d. hydrophobic interaction - separates on the basis of hydrophobicity. Based on interactions between solvent-accessible non-polar groups (hydrophobic patches) on the surface of biomolecules and the hydrophobic ligands (alkyl or aryl groups) covalently attached to the gel matrix.

e. dye ligand chromatography: the structure of dyes mimics that of nucleotides and other small molecules; can determine empirically if your protein binds preferentially to a particular dye. Many different resins available with different characteristics.

f. hydroxylapatite – a calcium phosphate resin that has some characteristics of an ion exchange column, but often gives a different binding and elution profile than does conventional ion exchange. Proteins are applied in low phosphate concentrations and eluted with a gradient of increasing phosphate conc.

3. Differential centrifugation

4. Electrophoresis

5. Membrane protein extraction
   a. ionic strength
   b. high pH
   c. detergents
   d. mild proteolysis

Purification of Recombinant Proteins:

Good sources of information are the handbooks and websites of the vendors who market products for protein expression and purification.

1. *E. coli*
   Advantages: Cheap, simple to set up, large-scale culture easy
   Disadvantage: Eukaryotic proteins often are insoluble in bacterial expression systems (can sometimes be refolded after purification under denaturing conditions); proteins lack post-translational modifications
   Advantage for all listed below: Eukaryotic posttranslational modifications will be added to proteins.

2. Yeast:
   Advantage: Cheap, reasonably simple to set up, large-scale culture easy
   Disadvantages: Difficult to disrupt the cell wall to make cell extracts; may be difficult to keep protease activity under control; not widely used for protein purification

3. Baculovirus/insect cells:
   Advantages: Eukaryotic post-translational modifications (but glycosylation may be different than in mammalian cells); cells can be grown in suspension culture so large-scale purification is feasible. Can express multiple proteins together using multiple viruses or a single virus with two promoters to drive expression of two different recombinant proteins.
Disadvantages: Making recombinant Baculovirus is more complex than constructing mammalian or bacterial expression vectors; need tissue culture facility, more expensive

4. Mammalian cells:
Advantages: Plasmid construction is simple
Disadvantages: More difficult to scale up; need tissue culture facility, more expensive

Tagging recombinant proteins to facilitate purification:

1. Histidine tags (six to ten histidines): Based on metal chelate affinity chromatography. A metal chelating ligand is immobilized on a column and charged with Ni$^{2+}$ ions. The tag of six consecutive histidine residues will bind to the column with high affinity. Proteins are eluted with a gradient of imidazole or by dropping the pH.

2. Antibody epitopes (FLAG, myc, HA, etc.) – bind to immobilized monoclonal antibody

Fusion proteins:

1. Glutathione S-transferase (GST) - binds to immobilized glutathione

2. Maltose binding protein (MBP) – binds to immobilized maltose

Protease cleavage sites within fusion proteins
a. thrombin
b. TEV (Lee et al.)

Once your favorite recombinant protein is purified, you may want to get rid of the GST- or MBP- fusion protein. Many vectors have built-in proteolytic cleavage sites. The fusion protein is cleaved by the protease (thrombin or TEV). The cleaved GST or MBP can be removed from your recombinant protein by another round of affinity purification.

Co-expression of proteins:

1. E. coli

a. N-myristoyltransferase (NMT) and substrate (ARF or Gi subunit) see Duronio et al. and Mumby and Linder

Uses two plasmids – one under kanamycin selection; the other under ampicillin selection. NMT is cloned into one of the plasmids; a protein substrate is cloned into the other. Both promoters are induced by IPTG. NMT folds into a native conformation in the bugs and is active as an enzyme. The substrate protein is synthesized, its initiator methionine is cleaved, and is myristoylated by NMT. Very efficient myristoylation of Gi – up to 95% of the protein is modified. Less efficient for ARF, usually 10% or less of the total ARF is myristoylated.
b. MEK1 or MEKK1 and downstream effector kinase (Khokhlatchev et al.) Uses the same two plasmid system as in (a), but in this case you are coexpressing protein kinases. MEK1 will phosphorylate the downstream target kinase in the *E. coli*, allowing you to purify the phosphorylated, active form of that kinase.

c. Endothelial nitric oxide synthase and chaperones groEL and groES (Martasek). Endothelial nitric oxide synthase will not fold properly in *E. coli* when expressed by itself, but will when coexpressed with the chaperones groEL and groES. See reference for details about the plasmids.

2. Baculovirus/insect cells

   a. coinfect with multiple viruses (Kozasa et al.)
   b. one virus expressing 2 proteins