Polymer Separation by Size Exclusion Chromatography

See **II/CHROMATOGRAPHY/Size Exclusion Chromatography of Polymers**

Protein Separation

R. K. Scopes, La Trobe University, Melbourne, Victoria, Australia

Introduction

Proteins are the essence of life processes. DNA contains the coded information for life, and is analogous to computer software, but it is the proteins that are analogous to the hardware, that actually carry out the job. Proteins have many roles, from catalysts (enzymes), through proteins that bind and interact with other molecules to control their behaviour, to structural and storage proteins which, although less functional, are nevertheless just as essential. Some proteins are 'solid', e.g. the proteins in our skin; others are soluble, such as those in our blood.

As with most separation procedures, those designed for proteins are designed to deal with a complex mixture of similar components, and separation often depends on slight and subtle differences between these components. Moreover, whereas some proteins may comprise a substantial proportion of the starting mixture, others may make up only a tiny fraction. The situation is very much like mining for minerals: first, select a source that is particularly enriched in the component you want, then work on removing all those you do not want. However, protein purifiers have one advantage over miners: nobody has succeeded in finding the philosopher's stone to turn base metals into gold, but molecular biologists do have the equivalent – the ability to greatly enrich the starting material with the desired protein. Consequently when talking about protein purification today, it is necessary to include a discussion of techniques for production of the protein by recombinant procedures. We also refer to the overall separation processes as 'purification', since the object is usually to obtain a homogeneous preparation of one single protein type, a 'pure' protein (even if this aim is not always quite achieved).

Separation procedures depend on differences in properties between the components, and fortunately proteins do come in a wide range of shapes and sizes. The properties that are exploited are solubility, ionic charge, size and shape, surface features and natural biological interactions. With recombinant techniques it is possible to modify the protein's structure so as to greatly simplify its purification.

The Starting Material

As indicated previously, the starting material should be rich in the protein of interest. That is not always possible, of course; the proteins may only come from a particular source, such as blood, mosquito larvae or potatoes (to name just a few possible sources). But even then it is possible to use molecular biology techniques to amplify the amount in either the natural source, or in an unnatural one, provided the gene (DNA) that encodes the protein, is known. Techniques for isolating the gene are outside the scope of this article, as are the details of the methods used for production of the recombinant protein, but both are becoming increasingly easy, and are often the first consideration when wanting to purify a protein. If the gene route is feasible, it will nearly always be the best way. Take the example of blood. Handling of human blood is now tightly controlled because of the dangers of viruses and because of ethical concerns, so if a blood protein can be produced in a fermentor full of yeast, or even in a cow's udder, these problems are circumvented. It is also much easier to produce a few grams of bacteria than a few grams of mosquito larvae.

Apart from simplifying and standardizing the raw material, the other two advantages of recombinant protein production are that (1) the amount of the desired protein, i.e. as a percentage of the protein in the starting material, will nearly always be much higher than in the natural source, and (2) simple techniques for modifying the protein by genetic means can make it very easy to separate from all the other proteins.

In most cases the desired protein will be soluble in aqueous media. If not, there are both advantages and disadvantages: the advantages include the ability to remove all the water-soluble proteins by simple extraction, which is a major separation process in itself. But the disadvantages include the problems of deciding what to do next, and how to separate the many insoluble proteins from each other. The answer is to get them into solution, and often this involves detergents, or other agents that may actually disrupt the natural biological state of the proteins. The natural state is called 'native' protein, whereas the disrupted state is called 'denatured'. It is sometimes possible to 'renature' a protein from the denatured state, and this may be necessary when recombinant expression of otherwise soluble proteins results in an insoluble, denatured product (inclusion bodies).

Solubility

Separation by solubility characteristics is the oldest technique in protein purification. Changes that result in proteins becoming less soluble in aqueous media include addition of salts, miscible organic solvents and organic polymers, and adjustment of pH. The aim is to make some proteins less soluble than others by varying these conditions, so the proteins can be separated from each other by centrifugation. It is not possible to predict which proteins will be affected, so it is essential to have a method for detecting the protein that is to be isolated, so that its position is known after any separation procedure. Typical precipitants include ammonium sulfate as a salt, ethyl alcohol and acetone as miscible solvents, polyethylene glycol as an organic polymer, and pH adjustment with a weak acid to about pH 5, which coagulates complex proteins.

After centrifuging the precipitate, it can be redissolved in a buffer that lacks the precipitant. **Table 1** illustrates an idealized analysis of the results of an ammonium sulfate fractionation of an extract containing an enzyme. Before the next separation procedure, it may be necessary to remove excess precipitant by dialysis or gel filtration.

Ionic Charge

All proteins have charges on them as a result of amino acid side chains such as aspartate, glutamate, histidine, lysine and arginine. The net charge on a given protein depends on its exact composition, and on the pH. Consequently at a given pH different proteins will have different net charges, and a shift in pH will change this value for each protein, though for all them it will become more negative for a higher pH, and more positive for a lower pH.

Although it is possible to exploit the charge differences by electrophoretic techniques, this has rarely been fully successful in preparative (as opposed to analytical) separations of proteins. The main problems have been in the design of reliable and safe equipment; there are some useful systems, but usually they will only be used when all else has failed. Ionic charge separations are carried out by ion exchange chromatography, which has been the most successful and widely used method of protein separation since its introduction some 40 years ago. Anion exchange columns have positive charges that attract negatively charged proteins, and at neutral pH most native proteins are negatively charged. For the minority of positively charged (high isoelectric point) proteins, cation exchangers are used. Samples are applied in a buffer that has low salt content (low ionic

Table 1 Results of a typical ammonium sulfate fractionation procedure in purifying an enzyme. Percentages are given as a percentage of saturation with ammonium sulfate. Specific activity is in units of enzyme activity per milligram of protein. Although the degree of purification is not high (2.1-fold), it is useful to note that much of the nonprotein material stays in the supernatant, and the precipitated fraction containing the enzyme can be dissolved in a much smaller volume than the starting extract. Results are also given for a second trial, in which recovery of activity has been sacrificed for a higher degree of purification

Fraction	*Volume (mL)*	*Amount of protein (mg)*	*Amount of enzyme (units)*	*Specific activity*	*Purification (x-fold)*	Recovery *(%)*
Extract	500	6000	380	0.063	1	100
0–50%	20	750	20			
50–60%	45	2400	310	0.130	2.1	82
60–70%	35	1600	70			
70% supernatant	540	1000	0			
0–55%	40	1500	90			
55–60%	35	1300	220	0.169	2.7	53
60% supernatant	535	3100	70			

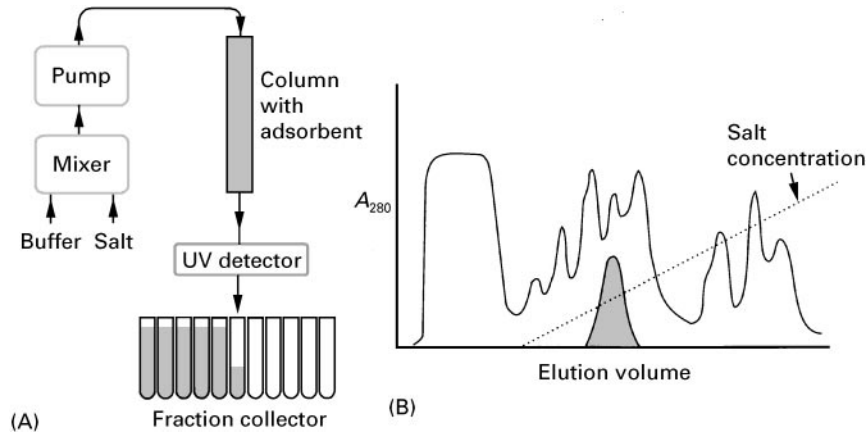

Figure 1 Example of protein separation using anion exchange chromatography. (A) A diagram of the basic principles of the equipment. The amount of protein emerging from the column is monitored by detecting the absorbance of light in the ultraviolet region (280 nm or 215 nm) due to proteins, and separate fractions are collected automatically. (B) The elution pattern. Some proteins did not bind to this adsorbent, being positively charged at this pH. The elution position of a specific protein is indicated by the shaded peak.

strength), and proteins of the opposite charge to the column are adsorbed. Proteins are generally eluted by increasing the ionic strength with gradual salt addition to the buffer, so that proteins of successive adsorbing strengths elute as the salt concentration goes up. This enables relatively high resolution of separate protein components, giving a high degree of purification. There are many commercial adsorbents and equipment to use them with. The type to be used will depend very much on the purpose; 'high performance' systems are expensive to operate and may not be suitable for commercial production of the protein, but most suited to research and development. Speed may be desirable, or may be of little concern; there are materials for all needs. An example of the separation of a fairly complex mixture of proteins on a 'moderate performance' adsorbent is shown in **Figure 1**.

Size and Shape

The sizes of protein molecules vary over a large range, but most have a molecular weight 20 000–200 000 Da (native protein). This translates to a size, assuming a perfect sphere, of roughly 3–8 nm in diameter. Many proteins are not spherical, however, so their longer dimensions will be greater than these figures. Separation by size is carried out by a process called gel filtration or size exclusion chromatography. The principle of gel filtration is shown in **Figure 2**. Beads that are porous, with pores of similar size to the proteins, are packed into a column. The largest proteins cannot penetrate the beads because the pores are too small, so they flow quickly around the outside of the beads, and emerge first. They do not emerge until after a volume (the void volume, V_0) that is equivalent to the volume of liquid outside the beads has passed through the column. This volume is typically about 30% of the total column volume, V_t. The smallest proteins are able to penetrate the pores in the beads. While in the beads they are isolated from the flow of liquid (which cannot pass through the beads because of capillary forces), and so do not move down the column. However, they do spend some time outside the beads, and eventually, after a full column volume has passed through, they emerge from the column at V_t. The gel

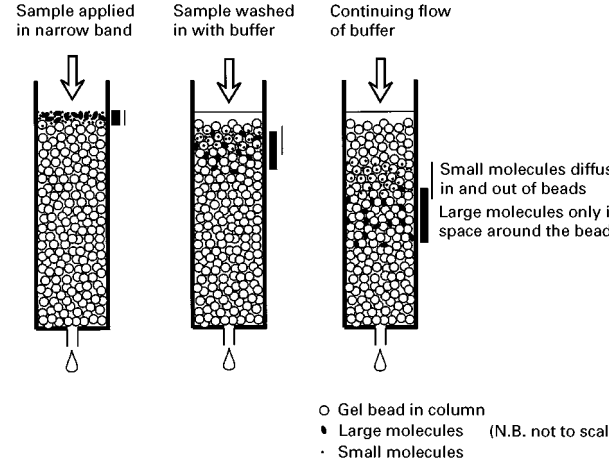

Figure 2 Principles of gel filtration. A mixture of proteins is applied to the column. Only the smallest molecules can penetrate the beads, while the larger molecules pass round the outside of the beads. Intermediate-sized molecules can partially penetrate the beads. As the separation continues, the larger molecules run out ahead of the smaller ones.

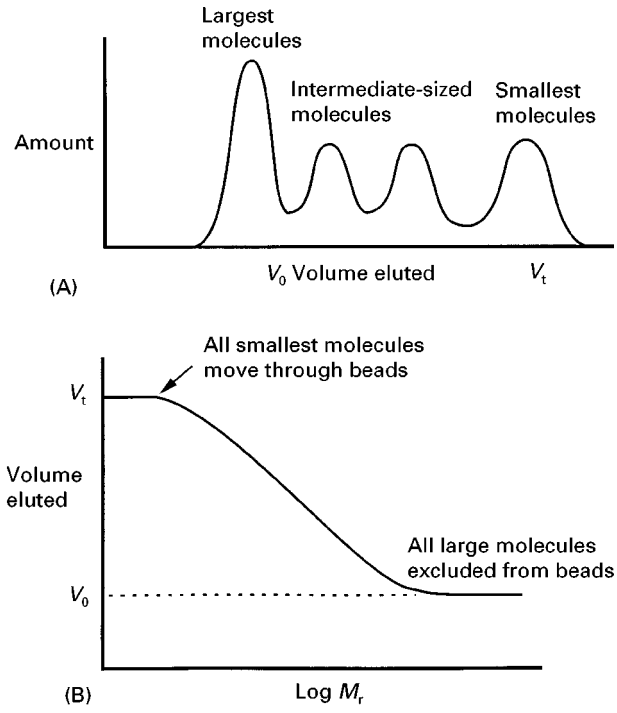

Figure 3 (A) The elution pattern of a mixture of proteins from a gel filtration column. (B) Plot of the 'elution volume' against the logarithm of the molecular weight. All small molecules emerge at V_t, all very large molecules earlier at V_0, and the values for intermediate-sized molecules fall on a straight line over about a 10-fold increase in molecular weight range. By comparing the results with the elution volumes of proteins of known molecular weight, the molecular weight of an unknown protein can be determined.

beads have a range of pore sizes, so that intermediate-sized proteins can spend some time inside the beads, but not as much as the smallest proteins. Consequently the emergence volume of each protein is related to size, as shown in **Figure 3**.

This is also the principle used to separate large molecules from small molecules in general, and is very useful for removing salts from protein solutions. In this case the beads have pores so small that even the smallest proteins cannot penetrate the beads, but salts can.

Gel filtration is a very gentle method; the proteins remain in solution at all times, and need not be exposed to any extremes of pH or salt concentrations. Resolution depends on the relative sizes of the proteins in the starting mixture. Gel filtration will commonly be used at a later stage of purification, when the quantities of sample have been reduced.

There are also membranes available which have pores of suitable size to separate proteins. A force, such as gas pressure, is used to push the liquid, plus smaller molecules through the membrane. The largest molecules are retained, and concentrated behind the membrane. This has application in the relatively crude separation of large proteins from small ones.

Separation by size is the commonest technique in the analysis of proteins, and is described in a later section. Using slab gels, which have pore sizes similar to the beads described previously, proteins are forced by electrophoresis to separate according to their size, the larger ones having more difficulty in finding their way through the gel, and so moving more slowly. This method has been applied preparatively, but only on a relatively small scale and mainly with unfolded, denatured proteins. Normally, one wants to isolate the protein in its native state, as not all proteins can be re-folded *in vitro*.

Other Surface Features of Proteins

The charged amino acids in a protein are responsible for the properties that allow separation on the basis of overall electrical charge. Other amino acids, especially those exposed on the surface of the molecule, confer other properties that can be exploited in separation methods. In particular, hydrophobic side chains would prefer not to be in contact with water. These amino acid side chains like to stick to each other, or to some surface that is also hydrophobic, and will do so in preference to being surrounded by water.

An adsorbent that consists of fat-like molecules attached to the insoluble beads is called a hydrophobic adsorbent, and will attract those proteins that have a larger proportion of hydrophobic side chains on their surface. Hydrophobic interactions are strengthened by inclusion in the solution of high concentrations of certain salts such as sulfates. A typical procedure would be to add up to 1 M sodium sulfate to the protein mixture, then run it through a column of hydrophobic adsorbent. Many proteins will not bind; these have few hydrophobic side chains. Then the salt concentration is gradually decreased, and proteins successively elute. The most hydrophobic proteins remain even a low salt concentration, and need other solutes such as chaotropic salts (e.g. thiocyanates), nonionic detergents, or high concentrations of glycols to elute them.

The most widely used method that exploits hydrophobic interactions is generally known as high performance liquid chromatography (HPLC). In this application proteins are adsorbed onto a material containing an aliphatic chain, such as C8 or C18, in conditions that result in partial unfolding of the native state of the protein so that internal hydrophobic regions are exposed. An increasing gradient of organic solvent then weakens the interactions between the protein and the adsorbent. Each protein is eluted

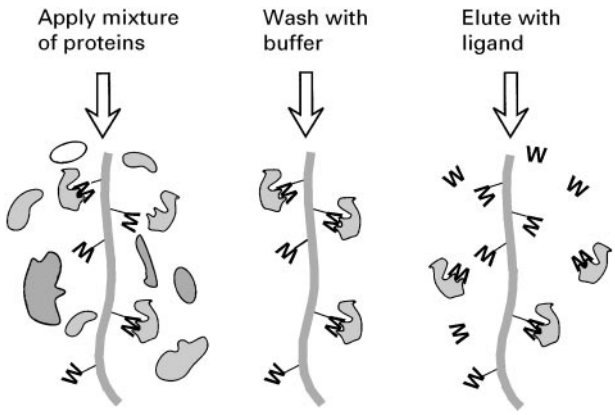

Figure 4 Principles of affinity chromatography. A ligand W that the desired protein binds to naturally is immobilized by attachment to column material. The protein mixture is applied, and the desired protein sticks to the immobilized ligand, while other proteins pass through and are washed away. Ideally only the proteins that bind to ligand W are left, though in practice there are usually others. Consequently a second 'affinity' step, involving washing the column with the natural, free ligand, is used to displace the specific proteins from the column.

in a narrow and very reproducible range of solvent concentrations, and provided that the proteins refold rapidly as they are eluted, this gives a very high resolution of components. This has been highly successful with small proteins and peptides, but is less useful for larger and more unstable protein components.

Bioaffinity

One surface feature that most proteins have is a binding site that corresponds to their natural physiological role. Enzymes, for example, bind to substrates known as ligands, which may be small molecules or parts of large ones. This binding is often very tight and specific, and can be made use of in separation procedures. The method is called 'affinity chromatography', and exploits the affinity of the protein for its natural ligand. The principle is illustrated in **Figure 4**. The adsorbent is designed so that it selects out only those proteins that bind to that particular ligand. The proteins may subsequently be eluted using some solute that disrupts the interaction, often just salt or more specifically with the natural ligand, free in solution.

Affinity Methods Associated with Recombinant Proteins

The most significant development in the past decade for purifying proteins has been the use of molecular biology techniques not only to express the desired protein in large amounts, but also to modify it in a way that makes its purification easy. This is achieved by making the protein longer with extra amino acid residues (polypeptide) that can be recognized by an affinity column. The product is called a fusion protein. There are many systems available, and the process is made easy by commercially available expression systems that automatically add the desired polypeptide. In some cases, the addition may be fewer than ten amino acids. In other cases it is a whole extra protein, which folds into an active shape while linked to the desired protein being expressed. Often the overall expression level is even higher with these 'fusion products' than with the unmodified protein, because the expression of the fusion portion has been optimized in designing the vector.

After purifying the fusion protein, it may be necessary to remove the fusion portion. This can be done with selective proteases, though these usually still leave a few extra amino acids on the end. One system involving an intein, or self-splicing protein, enables release of exactly the desired protein with no gain or loss of amino acid.

Integrated Purification Schemes

Purification of proteins, even when using the highly selective affinity methods, usually requires more than one step, with the extra steps carried out sequentially. The starting material is the most bulky stage, the crudest and least pure sample of the desired protein, and the first separation step needs to be able to deal with such material. A traditional scheme is to use salt (ammonium sulfate) fractionation first, then an ion exchange column, followed by gel filtration. Each of these steps exploits a different property of the proteins, so when combined they give a good chance of high selectivity. Repeated use of one particular procedure, e.g. ion exchange chromatography, will produce diminishing returns in terms of separation and recovery of protein.

An alternative first stage is to use an adsorbent that acts as a pseudo-affinity material and selects out the desired protein (along with many others), allowing nonprotein material to be removed. Any such adsorbent must be cheap to produce, because the crude extract can cause changes to its properties after a few uses, such as clogging or degradation. Dye adsorbents have been used for this purpose: brightly-coloured columns bind certain proteins selectively, and these can be washed off, then put through other steps such as ion exchange and/or gel filtration.

Table 2 Purification of an enzyme using three different procedures: dye adsorbent chromatography exploiting surface characteristics; ion exchange chromatography exploiting charge differences; and gel filtration to separate by size. The enzyme was xylulokinase from a thermophilic *Bacillus* sp. (author's unpublished results). Specific activity is in units of enzyme activity per milligram of protein. Nearly 200-fold purification was achieved, and the final enzyme was judged to be about 95% pure. This implies that the xylulokinase enzyme made up just 0.5% of the total protein in the original extract.

Purification step	*Volume (mL)*	*Amount of protein (mg)*	*Amount of enzyme (units)*	*Specific activity*	*Purification (x-fold)*	Recovery *(%)*
Extract	50	650	290	0.45	1	100
Off dye adsorbent	29	36	208	5.8	13	72
Off anion exchange	15	7.5	150	20	44	52
$(NH_4)_2SO_4$ precipitated	0.65	2.2	150	66	147	52
Gel filtration	2.0	1.4	120	86	190	41

The final stage in a protein purification will often be a high resolution step such as HPLC. Any asymmetry of the final peak might indicate that impurities are still present.

A purification procedure ideally results in a good yield of a 'pure' protein. If the final product is an enzyme it is relatively easy to determine how much remains at the final stage, compared with the amount in the crude extract, by measuring its activity.

Sacrifices of recovery for purity (or vice versa) during the procedure will need to be made, depending on the relative importance of these two factors. **Table 2** illustrates a typical purification of an enzyme, with a respectable recovery of activity and an end product that was analytically close to homogeneous.

Analytical Separations

At various stages, but more particularly at the end, one needs to know how close to purity the protein fraction is. Of many different analytical techniques that have been used, one is universally employed. This is known as SDS-PAGE, which stands for sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The gel (polyacrylamide) has pores that are similar in size to the protein molecules, and the mixture of proteins is forced through the gel by an electric field. First the proteins are fully denatured by treatment with the anionic detergent sodium dodecyl sulfate, which binds to all the proteins and confers on them a negative charge proportional to their size. The smaller proteins move through the gel relatively easily, but the larger ones are retarded by the lack of pores large enough to move through. Consequently there is a separation of components according to size, in a logarithmic way similar to that observed for gel filtration (see Figure 3). The individual components remain in tight bands; as a consequence proteins that differ in size by only 3–5% can be completely separated. This sharp resolution is thus able to detect small amounts of impurity, and a judgement can be made on the absolute purity of the sample. One further feature is that, as with gel filtration, the distance moved by an unknown protein can be compared with known standards and the unknown molecular weight determined to within about 5% accuracy. The amount of protein required is only about 1 μg or less, depending on the detection (staining) method. Capillary electrophoresis systems can be even more sensitive and accurate for analytical work.

See also: **II/Affinity Separation.** Theory and Development of Affinity Chromatography. **Chromatography:** Protein Separation. **Chromatography: Liquid:** Column Technology; Large-Scale Liquid Chromatography; Mechanisms: Ion Chromatography. **Electrophoresis:** One-dimensional Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis; Proteins, Detection of; Two-dimensional Polyacrylamide Gel Electrophoresis.

Further Reading

Coligan J, Dunn B, Ploegh H, Speicher D and Wingfield P (eds) (1995) *Current Protocols in Protein Science*. New York: John Wiley.

Deutscher MP (ed.) (1990) *Methods in Enzymology*, vol. 182: *Guide to Protein Purification*. San Diego: Academic Press.

Harris ELV and Angal S (eds) (1990) *Protein Purification Applications*; *A Practical Approach*. Oxford: IRL Press.

Janson J-C and Ryden L (eds) (1989) *Protein Purification*: *Principles*, *High Resolution Methods*, *and Applications*. New York: VCH Publishers.

Kenney A and Fowell S (eds) (1992) *Practical Protein Chromatography*. New Jersey: Humana Press.

Price NC (ed.) (1995) *Proteins Labfax*. Oxford: Bios Scientific Publishers. San Diego: Academic Press.

Scopes RK (1993) *Protein Purification*, *Principles and Practice*, 3rd edn. New York: Springer-Verlag.

Scopes RK (1997) Protein purification in the nineties. *Biotechnology and Applied Biochemistry* 23:197–204.

Wheelwright SM (1991) *Protein Purification*; *Design and Scale up of Downstream Processing*. Munich: Carl Hanser Verlag.