2

Protein Purification: Classical Approaches

INTRODUCTION

In Chap. 1 we discussed various approaches to obtaining starting material for the purification of proteins. Now we consider some of the standard approaches in purification. While many proteins are now purified using various types of affinity chromatography (considered in Chap. 3), the classical approaches described in this chapter are still in everyday use, either alone, in systems where the appropriate affinity chromatographic approaches have not been worked out, or usually in conjunction with affinity chromatography.

Before examining these "classical" approaches, however, two important general considerations must be discussed. The aim in protein purification is self-evident—preparation of a "pure" protein. The achievement depends, however, on the definition of purity. With the increasingly sensitive methods of detecting proteins that have been developed in recent years (discussed in detail in Chap. 4 in the section on electrophoresis) it has become considerably more difficult to prepare a "pure" protein. The main question is, pure enough for what purpose? The purity required for the accurate determination of a molecular weight may be quite different from that required for structural studies of the sequence of the polypeptide chain or for enzyme kinetic or ligand binding studies. Thus a pragmatic approach to the question of protein purity must be used, and this is discussed in these and other contexts in subsequent chapters. Since the determination of purity is usually based on one or more of the various approaches used to establish the molecular weight of a protein, further discussion is contained in Chap. 4.

During the course of protein purification, the specific activity is followed from step to step and a frequently used criterion of purity is the achievement of a constant

specific activity for several steps. This approach is particularly useful when constant specific activities are obtained for steps involving quite different physical bases for separation (such as molecular size and ionic properties). It is, however, advisable to determine the purity of the sample independently.

The second important consideration involves the yields from the purification scheme used and the amounts of "pure" protein that are required. For some purposes small amounts of highly purified material are desirable; on other occasions larger amounts may be required and judgments then have to be made as to whether particular steps in a purification scheme which may have low yields but good increases in specific activity are justified. As discussed previously, an informed choice based on a thorough understanding of the pitfalls of a particular experimental approach can allow the researcher to make such decisions.

PRELIMINARY FRACTIONATION PROCEDURES

The approaches that we consider in this section may yield only a few-fold purification; however, their use is not restricted to purification purposes alone. Early stages in most purification schemes have three other motives in addition to increased specific activity:

1. The rapid removal of proteolytic enzymes that might otherwise degrade the desired protein. Protease inhibitors may not always be sufficient to block the action of either specific or nonspecific proteases that may be present at the early stages of a purification or may be activated during a purification.
2. The concentration of starting material to more manageable volumes. In many of the procedures used, large volumes of material are not desirable: some of the precipitation methods described here are useful for effective and rapid concentration of the starting material—with the added advantage that they yield a purification as well.
3. The removal of material that may interfere with subsequent stages of the purification. In various procedures the desired protein is adhered to an immobile phase to allow contaminating proteins to be washed away. Whether this immobilization is by specific affinity, as in affinity chromatography, or by the general characteristics of the protein, as in ion-exchange or hydrophobic chromatography, it is often necessary to remove as much nonspecific protein as possible first so as to prevent interference with the immobilization.

Ammonium Sulfate Precipitation

Differential precipitation of proteins by ammonium sulfate is one of the most widely used preliminary purification procedures. It is based on the differing solubility proteins have in ammonium sulfate solutions and can result in a two- to fivefold increase in specific activity (in the case of glyceraldehyde-3-phosphate dehydrogenase from rabbit muscle, essentially homogeneous protein can be prepared simply by using a three-step ammonium sulfate precipitation procedure). Provided that appropri-

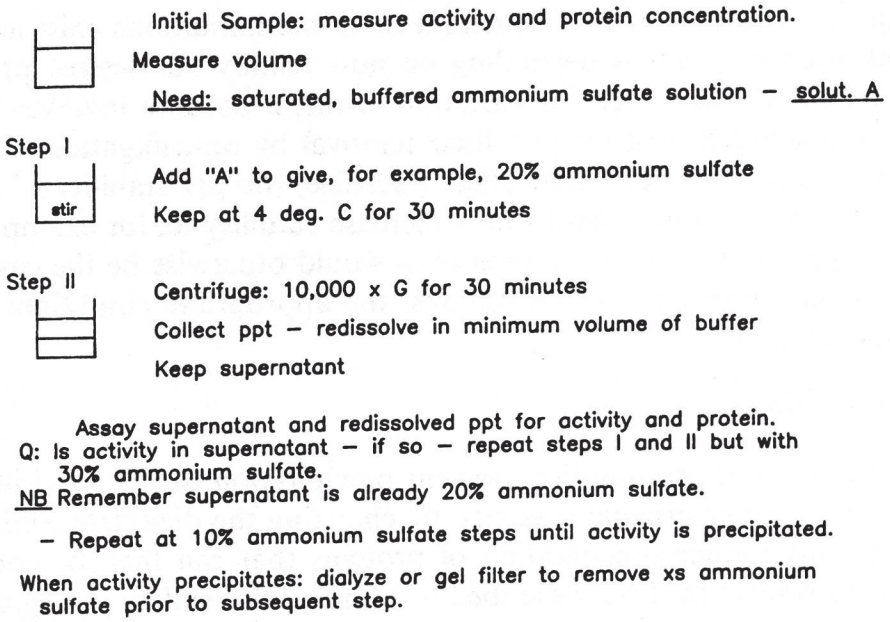

Figure 2-1 Outline of protocol for ammonium sulfate precipitation.

ately buffered ammonium sulfate solutions are used to protect the desired activity, recoveries approaching 100% can be expected. A typical protocol (as outlined in Fig. 2-1) consists of adding ammonium sulfate to give a certain percentage saturation, followed by a period of time for proteins to precipitate and a centrifugation step to collect the precipitate.

Once it is known in what range of ammonium sulfate concentrations the desired protein precipitates, the initial solution can be adjusted to a concentration sufficiently below this so that none (or very little—there are always judgments to be made) of the desired protein precipitates and the undesired protein can be removed by centrifugation. The ammonium sulfate concentration is then raised to a level sufficient to precipitate all (or most) of the desired protein while leaving in solution other undesirable proteins, and the precipitate retained for further purification. The appropriate concentration ranges are conveniently ascertained by screening a range of concentrations for small samples and determining the activity of the desired protein in the supernatant after centrifugation. Once this has been done, the appropriate concentration ranges can easily be chosen. It is important that when scaling up the total protein concentration in the sample is similar to that in the trial since the solubility of most proteins in ammonium sulfate is quite dependent on the total protein concentration.

Isoelectric Precipitation

Essentially similar in practice to ammonium sulfate precipitation, this approach is based on the fact that most proteins precipitate when there is no overall charge on the molecule—that is, at the isoelectric pH—since charge–charge repulsions tend to keep proteins in solution. Because proteins in general have fairly unique isoelectric points this procedure can give good, quick separation of unwanted proteins. In practice, the pH dependence of the stability of the desired protein can be a

determining factor in the method's usefulness. Some limitations exist for the effective concentration of proteins depending on how readily the desired protein, once precipitated, can be redissolved. A variation of this procedure involves the pH denaturation of unwanted proteins and their removal by centrifugation, an approach that can be assisted by factors that affect (increase) the pH stability of the desired protein. Substrates or other ligands may increase stability to, for example, low pH, thereby allowing a lower pH to be used than would otherwise be the case. As with ammonium sulfate precipitation procedures, the appropriate conditions are established on a small scale.

Solvent Precipitation

As with isoelectric precipitation, solvent precipitation can be used in two basic ways. Ethanol or other organic reagents, by changing the dielectric constant of the solvent, frequently induce precipitation of proteins that can then be collected and treated in ways similar to those described for ammonium sulfate precipitation.

Polyethylene glycol, of a variety of polymer sizes, is commonly used in fractional precipitation procedures. Any of the readily soluble polyethylene glycols can be employed: Those of higher molecular weight are frequently useful in concentration schemes. The dilute protein solution is placed in dialysis tubing and surrounded with dry polyethylene glycol, which absorbs water through the semipermeable membrane and concentrates the dialysis tube contents.

In the second approach, unwanted proteins in a mixture might be specifically inactivated and denatured by an organic solvent, thus allowing the contaminating protein to be removed. During the purification of Jack Bean α-mannosidase, contaminating β-*N*-acetylhexosaminidase is removed by specific inactivation with pyridine followed by centrifugation of the precipitated contaminant.

Heat Precipitation

Finally, we consider precipitation of contaminating proteins by heat denaturation. Different proteins have different stabilities at elevated temperatures, and if the desired protein has a greater heat stability than contaminating proteins, incubation at elevated temperatures for periods of time varying from a few minutes to a few hours often precipitates unwanted proteins, which can then be removed by centrifugation. As with pH-induced denaturation, the stability of the desired protein at elevated temperatures may in some cases be enhanced by the presence of substrates or other specific ligands.

CHROMATOGRAPHY METHODS

Gel Filtration

A large variety of gel filtration media are available and all work primarily on the basis of an *exclusion limit*, which is generally defined as the protein size that cannot penetrate the bead space of the material and thus is excluded from the column

matrix. Proteins larger than this size co-chromatograph through the column and elute in the *void volume* (V_0) of the column. Other material falls into two classes:

1. Material smaller than the exclusion limit which does not physically interact with the matrix material. Such material can be considered as having "normal" gel filtration behavior, and its *elution volume* (V_e) depends on the size of the material relative to the pore size of the matrix.
2. Material that interacts with the matrix material. Any physical interaction (causes of which are considered in the context of the nature of the matrix material) causes a retardation of the chromatographed material greater than what would be expected to occur simply by its ability to penetrate matrix space, and thus such material elutes at an anomalous elution volume. If the interaction with the matrix is of sufficient magnitude, the interacting material may elute at a volume larger than the normal *total elution volume* (V_t) of the column, which is the volume taken to elute molecules from the column having sizes similar to the bulk solvent volume.

Experimental Determination of Chromatography Parameters. The three parameters V_0, V_e, and V_t are used to describe the behavior of a particular molecule on a gel filtration column and must be determined experimentally. Three types of chromatography experiments can be envisaged:

1. In the ideal case, the sample size loaded onto the column is very small compared to the volume of the packed matrix material. In this instance (Fig. 2-2A), the elution volume, V_e, is simply the volume of eluent collected from the *start* of loading the sample to the midpoint of the sample elution.

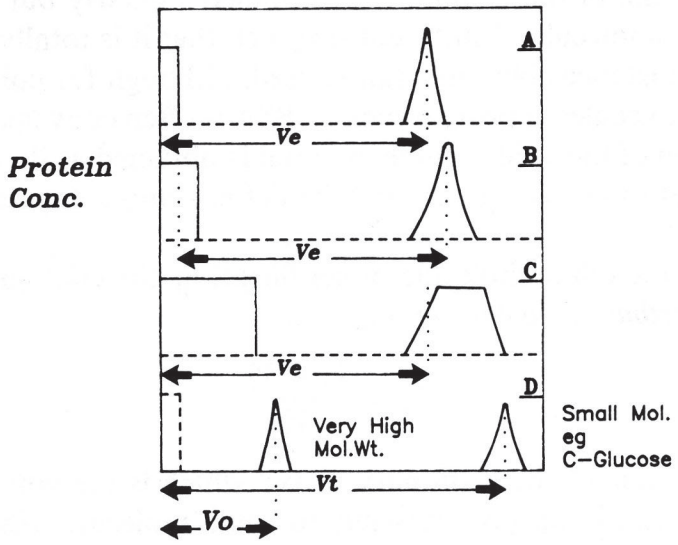

Figure 2-2 Elution profiles for gel filtration with sample sizes of 2% V_t (A), 10% V_t (B), and 25% V_t (C). (D) Summary of various gel filtration parameters.

2. When the sample size is not negligible compared to the bed volume of the column, the elution volume is usually calculated from the *midpoint* of the sample loading to the midpoint of the elution profile (Fig. 2-2B).

3. If the sample size is so large that a simple elution peak is not obtained (Fig. 2-2C), the elution volume is calculated from the *start* of the sample loading to the midpoint of the ascending side of the elution profile.

We should consider practical limits to the sample size that can be used in gel filtration chromatography. The *separation volume* (V_{sep}) between two peaks A and B can be defined as

$$V_{sep} = V_{e_B} - V_{e_A} \tag{2-1}$$

If a sample eluted from a column behaved ideally, the maximum sample size could be as great as V_{sep}. However, as the sample size is increased the size of the eluted peaks increases, and for resolution of the peaks the sample size should always be smaller than V_{sep}.

One problem is that under many experimental situations V_{sep} is not known, so as a general rule the sample size should be kept as small as is practical—in the range 2 to 5% of the column bed volume. For desalting applications, however, where the matrix has usually been chosen such that the desired protein elutes in the void volume while the elution volume of the "salt" approaches the total column elution volume, it is possible to use much larger sample sizes—in the region of 20% of V_t—and achieve effective desalting with minimal sample dilution.

The total volume, V_t, is usually obtained by loading a sample containing a small molecule (which does not physically interact with the matrix material) that can be conveniently monitored by absorbance or radioactivity and directly determining its elution position as described previously for V_e estimations.

The void volume of the column is determined similarly but by using a sample containing a macromolecule of sufficient size such that it is totally excluded from the matrix. In many instances blue dextran is used, although for columns employed to separate smaller molecules, a protein such as BSA is often convenient. As mentioned, the elution position of the void volume material is obtained as described for V_e determinations. Several other parameters can be defined and estimated once V_e, V_t, and V_0 are known.

The elution of a solute molecule in gel filtration chromatography can be characterized by a *distribution coefficient*, k_D:

$$k_D = \frac{V_e - V_0}{V_s} \tag{2-2}$$

where V_s is the volume of the stationary phase, which is the volume of solvent that can permeate the matrix and is accessible to small molecules (those which elute at V_t).

$$V_s = V_t - V_0 - V_{\text{gel matrix}} \tag{2-3}$$

In practice V_s is difficult to determine and is usually approximated by $V_t - V_0$, and k_D is replaced by K_{av},

$$K_{av} = \frac{V_e - V_0}{V_t - V_0} \tag{2-4}$$

and is *not* a true partition coefficient. These parameters are summarized in Fig. 2-2D.

Prior to considering the chemical nature of the various matrix materials available, we should discuss several other choices that have to be made concerning the practical setup of a gel filtration experiment. Many of the available matrices come in different particle sizes, from superfine to coarse. The smaller particles of the superfine grades give better physical packing of the matrix than does larger material, resulting in less zone broadening of peaks and consequentially, better resolution. The larger particle grades have considerably faster flow rates, however, which may be advantageous in working with unstable material or with rapid procedures such as desalting. The physical size of the column must also be chosen: Since the resolution of separated peaks increases as the square root of the column length, long columns in general give better separation than short columns but elute more slowly. The diameter of the column is important since narrow columns can hinder ideal passage of solvent through the column and wide columns give increased sample dilution. By far the most important choice regards the *sample viscosity*: High sample vicosity leads to distortion of elution peaks, which vary with the molecule size. The sample and buffer viscosity should not differ by more than a factor of 2, which, for most proteins, puts an upper limit for concentration of 50 to 70 mg/ml. It must be emphasized that many proteins undergo concentration-dependent aggregation, which can lead to anomalous gel filtration behavior not just due to viscosity problems, but also because of the molecular weight and size polydispersity that such a phenomenon can create.

Choice of Gel Filtration Matrix Material. Three basic types of matrix material have been used which differ somewhat in their physical and chemical properties. The most common are the cross-linked dextrans (e.g., Sephadex). This bead-formed gel is prepared by cross-linking dextran with epichlorohydrin. The resultant gel contains a large number of hydroxyl groups, which makes it quite hydrophilic and causes the gel to swell readily in water or electrolyte solutions. The porosity of the gel, and hence the useful fractionation range, is governed by the degree of cross-linking.

As discussed earlier, adsorption of material being chromatographed to the matrix leads to anomalous elution. Two principal types of adsorption must be considered: ionic and aromatic. With the cross-linked dextrans these effects are particularly noticeable on the highly cross-linked gels used to fractionate small molecules. The matrix material contains a low level of carboxyl groups, which at low ionic strength lead to the retardation of positively charged species and increased exclusion of negatively charged species. At ionic strengths above about 0.02, however, these effects become negligible with most proteins or peptides. A variety of aromatic compounds (such as purines, pyrimidines, dyes, and hydrophobic peptides) interact with the matrix material, causing additional retardation. These interactions can be suppressed by

using urea or phenol–acetic acid–water buffer systems for elution. However, such interactions are not always undesirable. Frequently, fairly similar aromatic compounds can be separated by making use of their interactions with the matrix, which can be modulated by changing the composition of the elution buffer. The addition of methanol or ethanol tends to increase the strength of these interactions, while altered ionic strength or pH can be used to weaken them. In essence this is hydrophobic chromatography, which is discussed in more detail later in this section.

The second type of matrix material commonly used consists of allyl dextran cross-linked with *N*,*N*′-methylene bisacrylamide, which gives a quite rigid gel structure having well-defined porosity. Due to the rigidity of the matrix, this type of material (e.g., Sephacryl) can easily be used with organic solvents with a much smaller effect on pore size (and hence distortion of the fractionation range) than with the Sephadex-like matrices. In general, the Sephacryl-like resins give better flow rates for equivalent fractionation ranges, but are only available for the separation of larger molecules [20,000 – 10^6 daltons (Da)].

Because of its high matrix density (and consequent carboxyl group density) these matrices have more pronounced ionic adsorption properties than the simple dextrans. In general, higher-ionic-strength buffers are therefore used with this type of material to help suppress such effects.

Finally, various derivatives of agarose have been used. The gel structure of agarose-based gels is stabilized by hydrogen bonding rather than chemical cross-linking but is quite stable under most conditions. The porosity is governed by the concentration of agarose in the material. The open structure of the agarose-based matrix makes this type of material (e.g., Sepharose) most suitable for the fractionation of very large macromolecules, although matrices with high agarose contents (up to approximately 6%) can be used with proteins in the range 10,000 Da and upward.

Such resins do contain low levels of carboxyl and sulfate groups, which can cause retardation of basic proteins, although as discussed for the other resins, these effects can be minimized by using elution buffers of reasonable ionic strength ($I >$ 0.02). The thermal and chemical stability of agarose gels can be increased (with negligible effect on porosity) by chemical cross-linking with 2,3-dibromopropanol. The enhanced stability of the resultant material allows alkaline hydrolysis (under reducing conditions) to be used to remove sulfate groups, giving a gel with a very low content of ionic groups and consequent elimination of most ionic adsorption effects. The basic structures of these various resins is given in Fig. 2-3.

Ion-Exchange Chromatography

Ion-exchange chromatography is based on the simple concept that at a given pH most proteins have a charge (either overall negative or positive, depending on the p*I* of the protein) and hence are attracted to (i.e., interact with) an opposite charge. Different proteins have differing amounts of charge and hence adhere more or less tightly to the opposite charge compared to other proteins. This interaction causes a retardation in chromatography provided that the matrix material has the

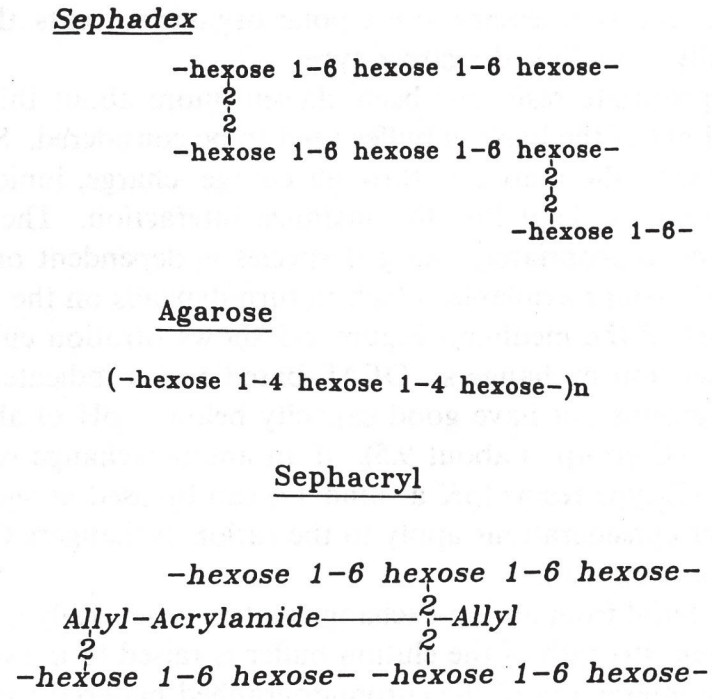

Figure 2-3 Schematic outline of the matrix structures of Sephadex, Sephacryl, and Agarose resins.

appropriate charge. In essence, the various matrices we have discussed for gel filtration chromatography are the basis of ion-exchange matrices: The matrix is derivatized to give it the desired anion- or cation-exchange properties. The commonly used functional groups are shown in Fig. 2-4. The basic properties of the support matrix are as discussed previously and should be selected based on the size of the proteins

Type	Functional Group	Counter–ion	Comment
DEAE	diethylamino ethyl $-O-CH_2CH_2N^+(H)(C_2H_5)(C_2H_5)$	Cl^-	Weak anion exchange
QAE	diethyl–(2–hydroxypropyl)–amino ethyl $-O-CH_2CH_2N^+(C_2H_5)(C_2H_5)-CH_2CH(OH)CH$	Cl^-	Strong anion exchange
CM	Carboxymethyl $-O-CH_2COO^-$	Na^+	Weak cation exchange
SP	Sulfopropyl $-O-CH_2CH_2CH_2SO_3^-$	Na^+	Strong cation exchange

Figure 2-4 Structure and properties of ion-exchange groups.

to be fractionated. If it is necessary to use polar organic solvents, the matrix should be of the chemically cross-linked agarose type.

Once the appropriate resin has been chosen (more about this later) only the ionic strength and pH of the loading buffer need to be considered. Since the interaction of a protein with the matrix is through charge–charge, ionic strength of the loading buffer should be kept low to maximize interaction. The capacity of the column to bind the appropriately charged species is dependent on the number of oppositely charged groups available, which in turn depends on the pK values of the groups and the pH of the medium. Figure 2-5 shows titration curves for some of the commonly used ion exchangers. DEAE-based resins indicate the presence of multiple charged groups but have good capacity below a pH of about 8.5 (the pK of the normal DEAE group is about 9.5). If an anion-exchange resin is needed at higher pH, the QAE-type resins (pK around 12) can be used at significantly higher pH values. Similar considerations apply to the cation exchangers CM- (pK around 3.5) and SP- (pK around 2.0).

Elution of material from an ion-exchange matrix is generally achieved in one of two ways. The ionic strength of the elution buffer is raised to a level that decreases the charge–charge interaction of the chromatographed material with the matrix, or the pH of the eluent is changed so that the charge of the adhered protein is altered such that it no longer interacts with the matrix. The pH must be decreased with anion-exchange material but increased with cation-exchange material. In some cases a combination of these two effects is used. The change is usually produced by running a gradient of increased-ionic-strength buffer (or the appropriate pH gradient) through

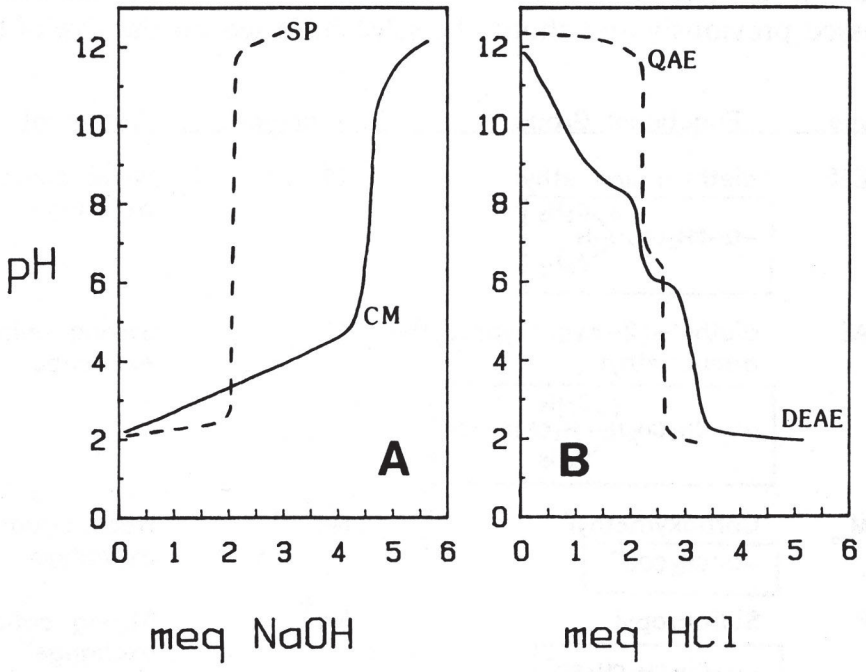

Figure 2-5 Titration curves: (A) CM (—) and SP (– – –) Sephadex; (B) DEAE (—) and QAE-Sephadex (– – –).

the column and monitoring the eluent for protein, activity, and so on, to locate the desired protein. Separation is achieved at two levels: First, not all proteins adhere to the column during the adsorption phase of the experiment. Second, as the elution gradient proceeds, different proteins elute based on the avidity of their interaction with the matrix; weakly bound proteins (i.e., those with the lowest charge density under the initial adsorption phase) are eluted first, while highly charged proteins require more drastic changes in pH or ionic strength.

Determination of Adsorption and Elution Conditions. During the initial stages of establishing a protein purification it is necessary to establish: (1) what type of ion exchanger should be used, (2) what conditions are necessary for adsorption, and (3) what conditions are necessary for elution. In general, conditions where the wanted protein adheres to the matrix should be established rather than conditions where other proteins adhere but not the wanted protein, since in the former case separation is achieved at both the loading and elution stages. In the absence of prior knowledge about the molecular properties of the protein it is convenient to screen a wide range of pH values rapidly with a particular resin type using the simple mixing and centrifugation procedure outlined in Fig. 2-6. Activity measurements on the supernatant allow one to establish adsorption (and, of course, elution) conditions rapidly.

An alternative approach for establishing optimal separation conditions for closely related molecules such as lactate dehydrogenase isoenzymes involves electrophoretic titration curves. This depends on the normal charge on the protein and its isoelectric

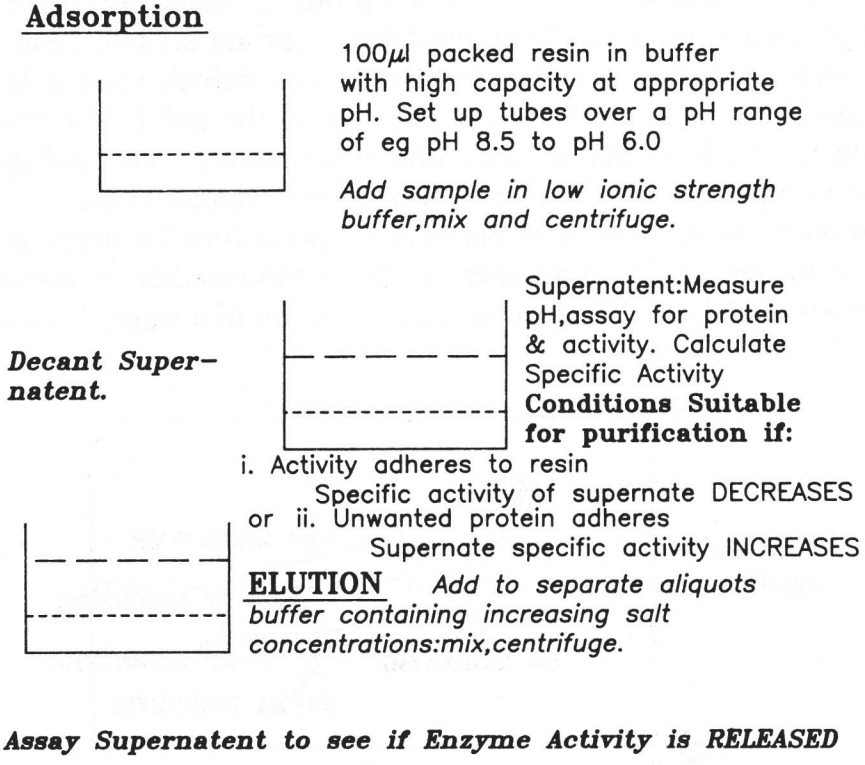

Figure 2-6 Experimental determination of conditions for adsorption and elution of material using ion-exchange resins.

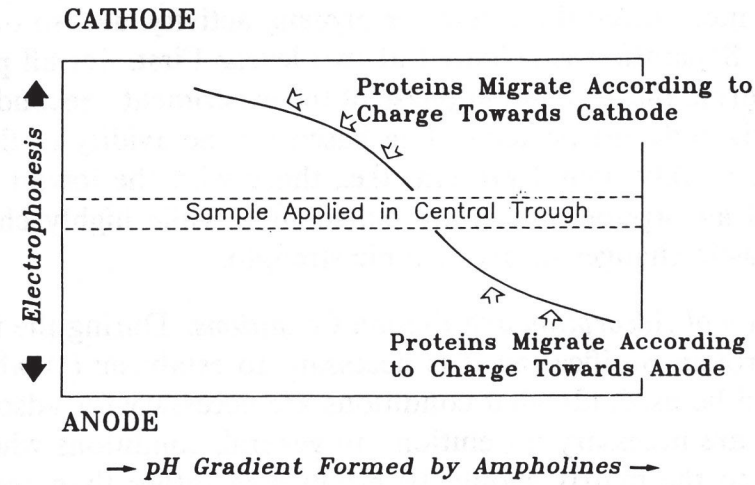

Figure 2-7 Outline of experimental determination of electrophoretic titration curves.

point (p*I*). Electrophoresis is carried out in a vertical plane using a large-pore gel matrix such as agarose or a low-percentage acrylamide which has a preformed *horizontal* pH gradient generated from the appropriate ampholines. As indicated in Fig. 2-7, the sample containing the mixture of proteins is added to a central horizontal well and electrophoresis is begun.

During electrophoresis the proteins move either toward the cathode or the anode or, if the pH is at their isoelectric point, they do not move at all. The rate of movement depends on the pH relative to the p*I* of the proteins. After electrophoresis is terminated the proteins are stained (for activity if appropriate; see later) and the titration curves examined. A typical set of curves for lactate dehydrogenase isoenzymes is shown schematically in Fig. 2-8. From the results the pH that gives the largest separation on the basis of charge can easily be evaluated. This pH gives optimal separation during elution from the appropriate ion-exchange resin.

In addition to being a suitable purification procedure for many proteins, ion-exchange chromatography has a number of other attributes that are outlined in Table 2-1. Particularly useful are the potential concentration of a wanted protein during a

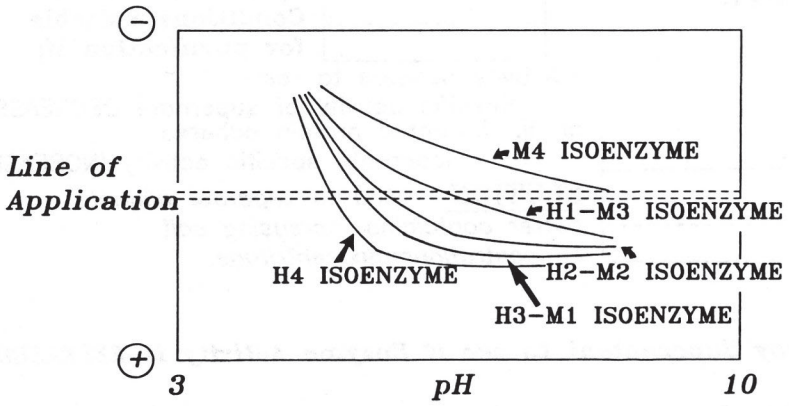

Figure 2-8 Electrophoretic titration curves for lactate dehydrogenase isoenzymes.

TABLE 2-1 Other uses of ion-exchange resins

Use	Resin of choice	Comments
Protein concentration	Any	Use steep salt gradient or stepwise elution
Demetallizing	Chelex	Dialyze protein vs. Chelex in buffer
Electrolyte removal	Mixed bed	
Peptide separation	Cation exchange	Gradient elution
Nucleotide separation	DEAE, QAE	Use gradient elution with QAE

purification procedure and the removal of metals from metalloproteins during the preparation of apoenzymes. In protein concentration it is often convenient to use a stepwise elution procedure rather than gradient elution.

Although we have discussed gel filtration and ion exchange in terms of column chromatography, both approaches are readily adaptable to thin-layer chromatography, which is particularly useful when a two-dimensional separation involving electrophoresis in addition to gel filtration (for example) is used. Ion-exchange methods are also particularly suitable for batchwise procedures since nonadsorbed material can easily be removed by washing and centrifugation prior to elution.

Hydrophobic Chromatography

Although the use of hydrophobic chromatography in protein purification has been popularized only recently, the idea owes its genesis both to gel filtration and affinity chromatography (Chap. 3). The matrix employed is usually based on agarose that has been derivatized in aprotic solvents with epoxides (which have relatively large alkyl chains). A generalized formula for the derivatives is

$$\text{agarose—O—CH}_2\text{—CH(OH)—CH}_2\text{—O—R} \qquad (2\text{-}5)$$

where R represents the alkyl chain and usually contains between 5 and 12 carbons. Any protein with some external hydrophobic characteristics tends to interact with such a matrix and be retarded relative to proteins lacking such characteristics. In general, the capacity of such columns for protein increases with increasing hydrophobicity of the substituent, with increasing degree of substitution, and with increasing ionic strength. The latter characteristic is quite distinct from the charge–charge interactions described earlier for ion-exchange chromatography, and leads to the principal method of elution from such a matrix: The ionic strength of the loading buffer is kept high and elution is achieved using a decreasing-ionic-strength gradient. Because the porosity of the matrix is decreased as the hydrophobicity of the substituent increases, generally a lower degree of substitution is employed, which is compensated for by using a higher initial ionic strength to maximize capacity and adsorption. In circumstances where adsorption is particularly tight (i.e., long alkyl