

Hydrophobic chromatography, metal ion chromatography, gel filtration, hplc**Hydrophobic chromatography**

Just as proteins can stick to an ionic adsorbent by charge interaction, they can stick to a hydrophobic material by hydrophobic interaction with nonpolar regions of their surfaces. This was discovered during early work on affinity chromatography, when control adsorbents with hydrophobic spacer arms but no specific ligand at their end turned out to adsorb proteins too. The hydrophobic groups used include alkyl (octyl-Sepharose), phenyl (phenyl-Sepharose) and alkylamino chains. The capacity is high, and adsorption is strongest at high salt concentration, so a sample may be applied immediately on redissolution after $(\text{NH}_4)_2\text{SO}_4$ precipitation, though in that case the exact salt concentration might be difficult to repeat. Proteins are eluted by decreasing the salt concentration, or in extreme cases by adding ethylene glycol or even ethanol to a very low ionic strength buffer. Resolution is not as good as in ion exchange chromatography, because α changes slowly, and the mechanism is similar to $(\text{NH}_4)_2\text{SO}_4$ precipitation, so it may not purify much more if that method has been used carefully. It is best used fairly early in a purification procedure.

A variant of hydrophobic chromatography is "T-gel", agarose which has been treated successively with divinylsulfone and β -mercaptoethanol, producing a $-\text{CH}_2\text{CH}_2\text{SO}_2\text{CH}_2-\text{CH}_2\text{SCH}_2\text{CH}_2\text{SH}$ side chain which for unknown reasons has high and fairly specific affinity for immunoglobulins.

Agarose and even cellulose act as hydrophobic adsorption materials at high enough salt concentration, a little below the $(\text{NH}_4)_2\text{SO}_4$ concentration required to precipitate the protein, which may be applied in this concentration of $(\text{NH}_4)_2\text{SO}_4$, or precipitated in presence of agarose, onto the agarose rather than as a precipitate, and poured onto the top of an agarose column and eluted with a decreasing salt gradient.

Metal-binding chromatography

This can be a general approach for proteins with exposed histidines, cysteines or carboxyl groups near each other, but in practice it is mainly for cloned proteins. If you are working with a cloned protein, you can add to the cloned gene a short sequence which facilitates purification. The method which is widely used is to add a sequence of six or so histidines at the C-terminus; you can buy vectors with the DNA sequence for this present. These bind well to divalent cations of transition metals such as nickel. A column is prepared by attaching nitrilotriacetic acid to a solid support. This binds nickel ions tightly; the resin is washed with 5 mM imidazole to remove unbound nickel. The fusion protein, perhaps denatured in 6M urea to ensure that the hexa-His sequence is exposed, is applied to the column. The column is washed with dilute imidazole, then more concentrated imidazole to elute the desired protein. An example of this is in a paper by Witzgall et al., *Anal. Biochem.* **223**:291-298 (1994), who in addition attached at the N-terminal end of the protein a DNA-binding domain from the yeast transcription factor GAL4 and an eleven amino acid peptide from bacteriophage T7 to which there are commercially available antibodies; this gave them a way to visualize the fusion protein on western blots. Other examples of genetically modifying the cloned protein for purification are using a $(\text{his-asp})_n$ sequence to bind to nickel, or a poly-arginine sequence to make the protein cationic. Added C-terminal sequences can often be removed by an appropriate carboxypeptidase - A for histidines, B for arginines - if this does not chew on into the native protein.

Some variations are: mercuric ions bound tightly to immobilized sulfhydryl groups, which can bind proteins by their exposed sulfhydryl groups - elution is with excess free SH compound

such as mercaptoethanol; and Fe^{+++} bound to iminodiacetic acid, which binds phosphoproteins by the phosphate groups.

Other general adsorbents

These include mixed-function groups attached to agarose, such as sulfanilic acid and arginine, which have shown good capacity and some selectivity; phenylboronate-agarose, whose $-\text{B}(\text{OH})_3$ group forms reversible covalent bonds with vicinal diols, and thus selectively adsorbs glycoproteins, or other proteins which bind tightly a ligand with a carbohydrate moiety such as NAD^+ or AMP.

Column dimensions

The manufacturer generally specifies the binding capacity of the material, though after it is swelled up and separated from its original bottle you will lose track of this. Large proteins cannot penetrate as many pores as small ones, so the capacity per unit volume or per weight of adsorbent will be considerably less. For a simple stepwise elution procedure, especially concentrating the protein from a dilute solution, one can adsorb on as much as half the column, since little further fractionation is expected; for careful separations with gradients the total column volume should be 5 to 20x the volume it takes to adsorb the protein.

In a perfect world the shape of the column wouldn't make any difference to the elution, which in theory depends only on volumes of column material and buffers. But the world isn't that perfect; a given distance of diffusion or irregularity of elution makes much more difference to a short fat column than to a long thin column (see handout). But a short fat column runs much faster. Thus again for simple stepwise elution, including concentration, a short fat column is used, sometimes as wide as high; for careful separations with an extended gradient a long thin column is used, though this is not so critical as in gel filtration.

Ultrafiltration

This is the filtration of a protein solution through a membrane with pores small enough to retain the protein of interest. It is a two-phase method - what is retained and what passes through the filter - but I mention it here in preparation for gel filtration. To happen in a reasonable time it requires either gas pressure on the solution above the filter (for large volumes) or increase of gravitational force by centrifugation (for small volumes). It is usually used just to concentrate a dilute protein solution, such as a crude culture broth, and sometimes for dialysis by diluting the concentrated solution with dilute buffer and concentrating again until practically all small molecules have been flushed through the membrane. It is sometimes used as a purification method, with membranes which retain the protein of interest but pass through smaller ones; this is particularly appropriate if the protein is very large, both because greater purification is achieved and because the filtration will be faster through the larger pores involved. The biggest problem is clogging of the pores by protein accumulating on the membrane surface; Amicon makes apparatuses which provide for continuous stirring while the protein is filtered under pressure. Use of this as a purification method is most appropriate when you want to concentrate the solution anyway.

Gel filtration

I have mentioned this procedure many times, you probably have at least some knowledge of it, and you would surely have used it by this time in a purification.

The principle by now seems simple. The gel filtration material (Scopes complains that it doesn't have to be a gel and is not really filtration, but it's too late to attach a different name) is porous, with pores the size of protein molecules. Large molecules, too large to enter any of the pores, pass down the column in the space between the gel particles, $V_0 = 0.3 - 0.38$ of the total volume of the column, and therefore are eluted **first**, when that volume has passed through; very small molecules enter all the pores, and therefore spend much of their time not moving and elute only when a volume V_t virtually the entire volume occupied by the bed has passed through; intermediate size molecules enter some of the pores, and are eluted somewhere in between. I sometimes compare the situation to search of a neighborhood by policemen in a squad car, rapid but of course incomplete, to a search entering every room, which would take a lot longer. Here we want to separate fat policemen from thin policemen according to what narrow doors they can pass through. Gel filtration of course is also used to separate proteins from salts such as ammonium sulfate, using a small-pored gel such as Sephadex G-25 or BioGel P10 which excludes all proteins; it is much faster than dialysis, and whether it dilutes more or less depends on how much expansion you allow in dialysis; but it is less applicable to really large volumes, and requires a little more attention.

The first gel filtration materials were the Sephadexes, cross-linked dextrans sold by Pharmacia; the larger pore sizes are very easily crushed by gravity and flow and little used nowadays, because of the availability of stronger materials which do not crush easily. See Table 6.1 from Scopes (handout). Note that having a wide fractionation range, as with many of the more modern materials, is not necessarily an advantage; the fractionation occurs within 55% of the column volume, and if a wider range of molecular weights are fractionated, any given pair, say proteins of mol. wt. 100,000 and 150,000, will be closer together on a material of wider fractionation range. You might use a wide fractionation range material to give an initial cut, eliminating much larger and much smaller proteins, and then use a narrower range material for best separation of these two. The best separation occurs for molecules eluting at about 0.6 column volume, but the peaks get broader the later they come off, and if separation from a slightly smaller protein is most wanted, you may want to use a smaller-pored column from which your protein elutes earlier (Fig. 6.8).

Gel filtration is often used to determine the native molecular weight of a protein. The column is standardized using proteins of known molecular weight, and elution volume plotted vs. log mol. wt., which should give a straight line. The elution volume is then plotted on the line and the molecular weight found. Properly one should use the elution volume at half way up the leading edge of the peak – the left side on a plot of A_{280} vs column fraction – because this is unaffected by the size of the sample applied; but usually one applies a very small sample and uses the peak fraction. The value determined can be affected by the shape of the protein – cigar-shaped proteins will enter smaller pores than spheres of the same mol. wt., frisbee-shaped proteins will only enter larger pores – but in practice most proteins are good approximations of spheres.

There are no peak sharpening mechanisms on a gel filtration column, so for maximum resolution you want to apply your sample in as small a volume as possible, short of having viscosity effects - generally at a protein concentration between 10 and 30 mg/ml. Scopes

suggests figuring the column size as follows: if m is the amount of protein in mg, the column diameter in centimeters is given by the cube root of m divided by 10, and the length should be 30x the diameter. A heavy solution, such as an ammonium sulfate precipitate, may run ahead between the beads rather than equilibrating with them, leading to a diffuse leading edge. Scopes suggests flowing the sample upward from the bottom, which can be done either using a pump or having the effluent line below the lowest point of the application line, to avoid this; but in my experience materials such as Sephadex G-75 may shrink in the high salt concentration, leaving a pool of sample at the bottom of the column which doesn't run in. So I would do just the reverse of what he suggests: I would apply the sample to the top of the column in a 'right side up' position, let it run in slowly, then after it is all in the gel invert the column and elute it upwards. Running a column from the bottom upwards is good for the higher Sephadexes and other gels subject to crushing, because the flow tends to oppose gravity rather than working with it in crushing the gel. But with the modern cross-linked agaroses and dextrans this is hardly a problem anyway. Note also that resolution can be lost **after** the column, if there is too large a space below the bed where the solution is funneled into tubing, or if the tubing going to the fraction collector is too long or wide; you want to use the smallest tubing that isn't itself a limitation on the flow rate of the column.

In principle separation depends only on the volume of gel the sample passes through, so that a short fat column could be used as well as a long thin column; but in practice the effects of uneven flow, as well as diffusive broadening, are much worse in a short fat column, as shown in Fig. 6.9.

The column material is supposed to be inert toward the protein, but there are various exceptions. Most column materials have a few free carboxyl groups, due to oxidation of carbohydrate or hydrolysis of acrylamides; chromatography should therefore be run in a high enough buffer concentration to prevent ionic adsorption to the column material, generally 50 mM, unless you are also desalting to a lower ionic strength. Polyacrylamide materials such as BioGels, and even cellulose and agarose at very high salt concentration, can adsorb protein by hydrophobic interactions; if this happens one should chromatograph at lower ionic strength or on a dextran polymer. On the other hand, proteins whose natural role is to bind to polysaccharide materials, such as cellulases and amylases, will probably bind to dextrans and be retarded, eluting much later than expected from their molecular weight, even after the salt fraction. Specific interactions of this sort, a sort of low-grade affinity chromatography, can of course be taken advantage of in a purification - elute from polyacrylamide at the protein's proper molecular weight, then separate it from other proteins of the same mol. wt. by chromatography on Sephadex. I've just been purifying an enzyme using Sepharose CL-6B for gel filtration, and found that most of the impurities at the end are higher molecular weight than my enzyme; I suspect they interacted with the Sepharose enough to be retarded to where my smaller protein elutes. So we are rechromatographing on BioGel P100, which is a different matrix (polyacrylamide). We haven't run the SDS gel yet, but it seems to help.

HPLC of proteins

HPLC now stands for "high performance liquid chromatography" though HP could also be said to stand for "high pressure". It was first introduced for rapid chromatographic

separation of small molecules, using small columns of very small particles 5 or 10 μm in diameter. The smaller the particles, the greater the number of "theoretical plates", the more rapid the attainment of diffusional equilibrium between stationary and moving phases, and therefore the better the separation for a given length of column (read about this in Scopes, pp. 105-111. In previous years I have had Dr. Ward lecture on protein hplc, but he would spend two full lectures on the theory and not get to the practice.) These columns consequently yield good separation and high resolution even though run rapidly. One advantage of fast operation is that they can be run at room temperature without denaturing the protein, because the protein comes off in 5 to 60 min. However, although the *linear* rate of flow is high, say 20 cm/hr, because the diameter is small, say 7.5 mm, the optimal *volume* rate is low, say 0.15 ml/min. Consequently the columns are usually run at a volume rate much greater than the optimal for reaching diffusional equilibrium. Optimum flow rate decreases with bead size - Scopes says increases inversely, which is confusing - and pressure to achieve the same flow rate increases with $1/r^2$ of the bead, so that halving the bead size leads to an eight-fold increase in the pressure necessary to run the column at a given rate. Thus the columns for full hplc are constructed of stainless steel; Pharmacia has a system with plastic columns of intermediate size, called Fast Protein Liquid Chromatography, useful for slightly larger quantities. It is run at room temperature, but doesn't take long, so denaturation in the warm isn't usually a problem.

For small organic molecules the particles could be based on strong inorganic materials such as silica, and organic solvents could be used; however, high pressure was required to push solvent through columns filled with small particles. For protein chromatography it was necessary to develop materials both strong enough to stand the pressure and porous enough to have a high surface area for adsorption, or for gel filtration. These materials are very expensive, and the operator becomes paranoid about not having the column clogged by precipitated protein; the sample is usually passed through a 0.2 μm filter just before going onto the column.

The bottom line is that hplc is a high resolution, but low capacity method. Its main uses, unless you have thousands of dollars for preparative columns, are as a final step in purification to remove the very last impurity, or for separation of peptides after proteolytic digestion of the protein, when biological activity no longer matters and organic solvents and 0.1% trifluoroacetic acid can be used, often in reversed phase mode, more appropriate because in denatured peptides all the hydrophobic groups are exposed. Such separation of peptides can be used as a "fingerprint" - how many peptides are produced by tryptic digestion, which one changes mobility in a mutant protein - but is most important for sequencing the protein.