

## Chromatography (and batch adsorption)

### General concepts

All adsorption and chromatographic methods depend on distribution of the material being purified (here protein) between a stationary (usually solid, sometimes adsorbed liquid) phase and a moving (here liquid) phase. We speak of a distribution coefficient  $\alpha = [\text{adsorbed material}]/[\text{total material}]$  (I say "material" in order to keep the definition general enough to include gas chromatography, but of course for us this is protein). The free (mobile) concentration is then  $[\text{total}](1-\alpha)$ . The amounts of free and adsorbed, or at least non-mobile, protein are the concentrations times the volumes of the mobile and non-mobile phases,  $V_0$  or  $V_S$  and  $V_t - V_0 = V_a$ ; note that the non-mobile phase always contains a good deal of solvent if the adsorptive material is made up of porous particles, as for our purposes it nearly always is. For calculations concerning the protein of interest it is immaterial whether the protein is actively adsorbed or passively entrapped in the pores of the material - in gel filtration it is only passively entrapped - but this can become important in assessing how well the adsorbed protein is separated from other proteins which are entrapped but not actively adsorbed.

### Batch adsorption

This is a simple technique intermediate between precipitation and chromatographic methods; it is like precipitation in that the solid phase is generally collected by centrifugation, or even by filtration which can reduce further the volume of trapped solvent containing unbound proteins - but like chromatography in that the basic mechanism is adsorption on the surface of the solid phase. The adsorption must be very tight, as shown by the following calculations: the conc. of free protein is  $p$ , so the total amount of free protein is  $p(V_S + V_a)$ ; the conc. of adsorbed protein is  $q$ , so the amount of adsorbed protein is  $q V_a$ . The fraction of protein bound,  $f = q V_a / p(V_a + V_S) + q V_a = V_a \alpha / [V_a + (1 - \alpha)V_S]$ . From this it can be seen that a high value for  $f$  requires a high value for  $\alpha$ , and  $f$  is also made smaller if  $V_a$  is very small compared to  $V_S$ , since even a very small value for  $(1 - \alpha)$  can make the denominator larger than the numerator if  $V_S$  is very large compared to  $V_a$ . And you want  $V_a$  to be very small compared to  $V_S$  in order to minimize the amount of unadsorbed but trapped protein carried along with the adsorbed protein. For instance, at  $V_a/V_S = 0.1$ , for  $\alpha = 0.9$   $f =$  only 0.45; for  $\alpha = 0.98$   $f =$  0.82; for  $\alpha = 0.99$   $f =$  0.99. Scopes states that "with most adsorbents,  $K_p$  values smaller than  $10^{-6}$  M are needed to give  $\alpha$  values of 0.98 or greater." Effective adsorbents, except affinity adsorbents, are likely to adsorb many other proteins and thus require that a substantial amount be used. And since separation is only between two phases, adsorbed and supernatant, resolution is generally much less than in chromatography. Nevertheless, batch adsorption, which is simpler to set up than chromatography, especially for large volumes, is generally used in two cases: 1) with inorganic adsorbents which are flocculent and do not give a usable flow rate in column chromatography; 2) with affinity adsorbents which have such high and such specific adsorption of the protein of interest that no further purification is achieved by using a column set-up. One can collect the adsorbent by centrifugation (relatively small amounts) or vacuum filtration (large amounts) and wash by resuspension in the starting buffer to remove remaining unadsorbed protein, then resuspend in an eluting buffer and again separate adsorbent and supernatant solution, this time with the protein in the supernatant.

The most commonly used adsorbent is calcium phosphate, originally in a gel form but now usually in one of two crystalline forms, brushite and hydroxylapatite, which can be used in columns; Biorad and other suppliers sell hydroxylapatite, but it is easy to make your own if you have a peristaltic pump. You pump 1M solutions of  $\text{CaCl}_2$  and  $\text{K}_2\text{HPO}_4$ , with a little  $\text{K}_3\text{PO}_4$  to keep the solution basic, at the same rate into a large vessel with very slow stirring. The form originally produced is brushite, which is converted into hydroxylapatite by incubation with base, for a few days at room temperature or for two hours at  $100^\circ$ . For column use you want as large crystals as possible, which is somewhat

uncontrolled; probably presence of a few nuclei in the original solution, to generate comparatively few large crystals rather than many little crystals, is what does it (some people say to use the poorest quality  $\text{CaCl}_2$  you can get, desiccant grade; the insoluble material serves as nuclei and the soluble junk is washed away anyway. You let the large crystals settle and pour off the fines). You treat the material gently in order not to break up the crystals.

To test conditions to use with calcium phosphate or any batchwise adsorbent, prepare a series of microtubes with a constant amount of protein solution (at whatever stage of purity you want to use the method). Add increasing amounts of a well-mixed suspension of the adsorbent - being sure it is well mixed, so that you are adding the same amount to each tube, is the hard part - and water to bring to a constant volume. Let sit a few minutes, then centrifuge down the adsorbent and assay the supernatant solution for activity. You may find that a certain amount of adsorbent can be added without adsorbing your protein (as others more strongly adsorbed compete for the gel), then a certain amount more adsorbs all of your protein. You can thus achieve better purification by adding the first amount of adsorbent, centrifuging down, then adding more to the supernatant solution. But you also test for conditions of elution.

Proteins are not eluted from hydroxylapatite by monovalent ions such as chloride (so they could probably be adsorbed from a chloride-containing eluate from an ion exchange column); they are weakly eluted by divalent anions such as succinate and dimethylglutarate, but more usually they are eluted with increasing concentrations of phosphate, or in extreme cases with pyrophosphate (hydroxylapatite cannot be reused after elution with pyrophosphate). Some basic proteins are eluted by the cations, monovalent cations reasonably well but  $\text{Ca}^{++}$  much better. There was a time when it was believed that one should elute hydroxylapatite columns stepwise rather than with a gradient, but I don't think that that is believed any more.

Other inorganic adsorbents used batchwise include alumina  $\text{C}_\gamma$ , the most stable form (achieved by prolonged storage) of aluminum hydroxide precipitated by adding base to a solution of an acidic aluminum salt such as aluminum sulfate; bentonite, which is a natural silica material, containing very fine particles which are difficult to sediment, and supposed to have particular affinity for ribonucleases, hence sometimes used in RNA preparation; titanium dioxide; zinc hydroxide, prepared like alumina  $\text{C}_\gamma$  by basification of a solution of a salt such as zinc acetate; even charcoal, and benzoic acid. Celite, the diatomaceous earth usually used as a filter aid (to increase flow), sometimes acts as a weak adsorbent of cationic proteins.

Perhaps the greatest virtue of batch adsorbents is their cheapness; you usually use them once and throw them away. Hydroxylapatite may be washed with a high concentration of phosphate, then with water (indeed for some purposes it should be so washed before the first use) for storage, but there is always some protein that doesn't wash off, and smelly bacteria grow happily on this unless the material is stored in presence of sodium azide, a fairly generally bacteriostatic compound, which however binds to hemoproteins (and hence should not be used in their solutions) and when poured down the drain reacts with copper pipes to produce insoluble, explosive copper azide; Dr. Kahn tells of a plumber being killed by an azide explosion at Albert Einstein Medical School when he whacked a pipe with his wrench.

## **Chromatography**

Here of course we refer to the use of a tube containing some adsorbent packing material; different proteins adsorb to the packing with different affinities, and are eluted out the end of the column in order of their affinities, those more loosely bound first. Column chromatography may be run with an eluting solution of constant composition, which is called isocratic chromatography, but this is in fact relatively rarely done for proteins, because for isocratic chromatography to be useful  $\alpha$  must be within relatively narrow limits, say between 0.3 and 0.8; if it is lower the adsorbed solute will not be well separated from the often large class of solutes which do not adsorb at all ( $R_f = 1 - \alpha$ , and while on a thin layer plate an  $R_f$  of 0.7 would mean a compound was well separated from those at the front, on a column

those which do not really adsorb will tail this far); on the other hand, a solute with  $\alpha = 0.8$ ,  $R_f = 0.2$ , is eluted when 5 column volumes of buffer have passed through, but generally in a very broad peak, not well resolved from other proteins of similar  $R_f$ . The value of  $\alpha$  tends to change rapidly with small changes in conditions, and it can be quite hard to find good conditions for isocratic chromatography of a given protein, though they can be very effective when found; one example is the separation of what are called  $\alpha$ - and  $\beta$ - trypsin on sulfopropyl-Sephadex in 0.1 M Tris Cl pH 7.09.  $\beta$ -trypsin is the complete single chain protein,  $\alpha$ -trypsin is a form still possessing all the amino acids, but having a single peptide bond cleaved, after lysine-131.

More usually, the protein of interest is initially fully adsorbed,  $\alpha > 0.99$ , and is then eluted by a change in conditions. For gas chromatography this may be by a change in temperature, but for liquid chromatography it is usually a change in composition of the solvent. Changing it abruptly is called stepwise elution, because if you plotted eluant composition, say salt concentration, vs. fraction number, it would look like a series of steps. More usually the composition is changed continuously, by the continuous mixing of solvents in which the protein would have high and low values of  $\alpha$ ; this is called gradient elution. It allows one to cover a wide range of conditions, among which should be the best for separating your protein, in a single experiment. In the simplest form the gradient rises at a steady rate and the plot of composition vs. fraction is a straight line; this is called a linear gradient. You have two bottles of equal cross-sectional area, one (with a stir bar in it) connected to the column, the second connected to the first, so that as two drops are delivered to the column, the level in each vessel goes down by one drop. The first bottle contains the low salt buffer, the second the high salt buffer. Having the high salt solution in a vessel of larger or smaller cross-sectional area results in convex-upwards and convex-downwards gradients respectively (this can be remembered by reasoning that if the high salt solution's bottle is wider, more final solution will be delivered to the column, therefore the area under the plot of concentration will be larger). I say low and high salt for simplicity, but the same arrangement can be used to deliver a gradient of any solute.

The value of  $\alpha$  tends to decrease as protein concentration increases, because of competition of the various proteins present for the range of binding sites, which are not homogeneous in composition. Consequently in isocratic chromatography and with stepwise elution the leading edge of the 'peak' of protein passing down the column tends to be sharp, because protein which gets out ahead goes into an area of higher  $\alpha$ , and is therefore slightly more adsorbed and drops back to the peak. But the trailing edge of the peak trails out, for the same reason; protein back there in a region of lower concentration has a higher  $\alpha$  and consequently drops further back. But with a gradient of eluant composition, the value of  $\alpha$  is continually dropping, so that protein behind the peak is in a region of lower  $\alpha$  and is less adsorbed; it therefore spends more of its time moving and tends to catch up with the main peak. Gradient elution therefore tends to sharpen peaks from both sides. Note that if the protein is initially completely adsorbed, it can be adsorbed from a large volume of solution, and then released in a small volume by a large stepwise increase in the eluting property of the eluant. When I prepared  $\alpha$ - and  $\beta$ -trypsin as described above, on a large column, the volume of each pool of peak fractions was close to one liter. By dropping the pH to 3.0, which was done anyway to minimize self-degradation, the protein could be completely adsorbed on a short fat column of SP-Sephadex; it could then be eluted in a volume of 60-70 ml 0.5 M Tris Cl pH 7.6. Any chromatographic method can also be used as a concentration procedure.

[The value of  $\alpha$  depends on the average dissociation constant of protein from the adsorbent,  $K_p$ , the total protein concentration  $p_t$ , and the concentration of binding sites,  $m_t$ . Since  $\alpha = q/p_t$ , where  $q$  is the concentration of bound protein,  $q = p_t \alpha$ . The concentrations of free sites  $m$  and free protein  $p$  are the total concentrations minus the bound protein  $q$ ,  $(m_t - p_t \alpha)$  and  $(p_t - p_t \alpha) = p_t (1-\alpha)$ .  $K_p = m \cdot p / q = (m_t - p_t \alpha)(1-\alpha)/\alpha$ . Solving this for  $\alpha$  sets up a quadratic equation,  $p_t \alpha^2 - \alpha(m_t + p_t + K_p) + m_t = 0$ . Scopes gives a table of values of  $\alpha$  resulting from various assumptions about  $K_p$ ,  $p_t$ , and  $m_t$ . For instance, if  $K_p = 0.1$  mM,  $p_t = 0.5$  mM, and  $m_t = 1.0$  mM - quite possible for an ion exchanger -  $\alpha =$

0.85. Lowering  $K_p$  to 0.01 mM raises  $\alpha$  to 0.98, while lowering  $m_t$  to 0.1 mM lowers  $\alpha$  to 0.16, which is useless. A high concentration of adsorption sites is important as well as a low  $K_p$  (though less so if the concentration of protein which will bind is low, as in affinity chromatography of a relatively crude mixture).]

### **Ion exchange chromatography**

This is of course the commonest type of protein chromatography, and much of the theory has been tested using it. The first ion exchange materials were essentially polystyrene plastics with ionic monomers such as dimethylaminostyrene copolymerized into the plastic, referred to as ion exchange resins. These are useful for chromatography of small molecules such as amino acids, but not usually for proteins, for two reasons: 1) they are rigid hydrophobic spheres which cannot be penetrated by proteins, which means their capacity is low; and many proteins also adsorb to them by hydrophobic binding, which becomes stronger at higher salt concentration. The first successful support (or matrix) for ion exchange chromatography of proteins was cellulose, and this is still widely used, particularly in the microgranular and spherical forms. Then Sephadex, which is a dextran polymer, was used, but the higher porosity variety is subject to considerable shrinkage in high salt, and must be taken out of the column to swell again. More recent ion exchange materials are based on cross-linked agarose (Sephacrose), polyacrylamide (BioGel, from BioRad) or other synthetic polymers (Trisacryl, from LKB). These have high porosity, consequently high capacity even for large proteins, and good mechanical stability, i.e. resistance to crushing.

Perhaps four-fifths of proteins are negatively charged at neutral pH, and consequently the most used ionic adsorbent group is diethylaminoethyl, or DEAE for short. The DEAE group has an intrinsic  $pK_a$  around 9.0, but there are some doubly substituted groups,  $(C_2H_5)_2NCH_2CH_2N(C_2H_5)CH_2CH_2-$ , which have a second  $pK_a$  around 5.6, and the ionic effects of nearby charged groups lower the  $pK_a$  of some singly substituted groups, so in practice DEAE materials have some buffer capacity over the entire range from pH 5 to 10. Quarternary nitrogen groups, such as triethylaminoethyl (TEAE) and QAE, which is TEAE with a hydroxyl on one ethyl, can't deprotonate, and consequently are useful at high pH or if you're eluting by pH change and don't want the adsorbent to buffer.

The most commonly used cation exchange group is carboxymethyl (CM), but like acetic acid it protonates below pH 5; at lower pH, where more proteins are cationic (but fewer are stable) sulfopropyl (SP; there used to be sulfoethyl = SE) is used. Phosphocellulose is also available and probably acts as a primitive affinity adsorbent for enzymes whose substrates are sugar phosphates.

Ion exchange adsorbents are usually eluted by means of a gradient or steps of increasing KCl or NaCl concentrations, in presence of a constant concentration of buffer. The charged form of the buffer should be of the same charge as the ion exchange material, i.e. use Tris or another amine with DEAE, phosphate or another acid with CM. This is because if the ionic form of the buffer adsorbs ionically to the adsorbent, it is removed from solution, and by the Henderson-Hasselbalch equation the pH will change. If total buffer concentration is low the pH may change considerably. However, this is mainly important at low buffer concentration, say 5 mM, and plenty of people have gotten away with eluting a DEAE column with phosphate buffer. The pH close to the adsorbent is different from the bulk pH anyway, because of Donnan effects: DEAE, for instance, will attract and concentrate  $OH^-$  ions, repel and deplete  $H^+$  ions, so that the pH will be as much as 1 pH unit higher in its vicinity; this may have consequences for the stability of the protein, especially on cation exchangers since denaturation is more likely at pH 4.5 (bulk pH 5.5) than at pH 9 (bulk pH 8). One could elute by changing the pH - down for a DEAE column, up for a CM column - until the protein has lost enough of its net charge to elute, but in practice the pH gradient tends to be quite irregular, changing sharply as protein is eluted, so that proteins come off in groups, unless your protein is quite strongly adsorbed and a buffer of 0.1 M or higher can be used. I have used pH elution of QAE-Sephadex effectively, using imidazole and piperazine Cl buffers which have been acidified to pH 3, then back-titrated to pH 7.3 and 5.3 respectively so that ionic strength remains nearly constant. The buffer  $pK_a$ s involved are 7.0 and 5.6, so there is buffering

throughout this range. Elution by pH change is likely to achieve better fractionation below pH 5.5 and above pH 9, where many protein side chains titrate and properties change rapidly.

[A variation on this is the so-called Chromatofocusing technique, in which the adsorbent, polyethyleneimine-agarose, buffers evenly over a wide range; as the column is eluted using the acid form of an ampholyte (called Polybuffer and sold by Pharmacia at a good price) a pH gradient develops in the column, at low ionic strength, and proteins are claimed to emerge in sharp, highly resolved peaks at or slightly above their isoelectric points. But the people I know who've tried it couldn't get it to work as reported.]

Some proteins are specifically stabilized by  $K^+$  or  $NH_4^+$ , but I've never heard of one being stabilized by  $Na^+$ ; so I would tend to use KCl rather than NaCl, or  $NH_4Cl$  if  $pH \leq 7$  (it will evaporate as  $NH_3$  as the pH goes higher). Also, protein ion exchange chromatography, unlike that of small molecules, has rarely used the different affinity of different small counterions. This is useful mainly for proteins which are only weakly adsorbed even at the lowest usable buffer concentration (5 mM) with chloride as counter ion; if this is the case, it may be possible to use DEAE formate or acetate instead (by deprotonating the DEAE in base, washing away excess base, titrating back to the desired pH with formic or acetic acid, and applying the protein in 5 mM buffer with formate or acetate as anion).

One initially tests one's protein by applying a small sample to a small DEAE column in say 20 mM Tris Cl pH 8.0 and seeing whether it is adsorbed. If not, some purification will have been achieved, since most proteins are; you would then test adsorption to CM-cellulose in MES buffer at pH 6.5. If it doesn't adsorb to this either, you have probably purified considerably, but could try a still lower pH (or higher pH with DEAE). If it did adsorb to DEAE at pH 8.0, you might try lower pHs where fewer other proteins will adsorb. It is of course always necessary to ensure that the salt concentration of the protein sample is low and reproducible; this means desalting an  $(NH_4)_2SO_4$  precipitate by dialysis or gel filtration before applying to the column. For this reason, a good sequence of operations may be  $(NH_4)_2SO_4$  precipitation to concentrate, gel filtration to fractionate by molecular weight and incidentally desalt, then ion exchange which can concentrate again the sample diluted by gel filtration.

When satisfied or exhausted with testing conditions for adsorption, one tests elution. Usually one initially elutes with a gradient over a wide range of concentration, say from 0 to 1 M salt, in the buffer with which the column was originally equilibrated. Having gotten a rough idea where the protein comes off, one then uses a narrower gradient. For instance, one might apply the protein in 20 mM buffer, wash with this buffer until all nonadsorbed protein comes through ( $A_{280}$  drops close to 0), then wash with 20 mM buffer + 0.1 M salt until the  $A_{260}$  again drops close to 0, then use a gradient from 0.1 to 0.2 M salt to elute your protein with good resolution from most other proteins, even those that elute in this range but not exactly at the salt conc. where yours elutes.

Scopes also cites 'affinity elution' - a specific ligand of the protein binds to it and changes its charge, or more rarely its conformation, enough so that its  $\alpha$  changes enough for it to elute. Since most polyions are anions, this is mostly with cationic proteins bound to CM-cellulose &c., which become less cationic when an anionic ligand binds. The great advantage of course is that such a procedure is much more specific for your protein than changing the chloride concentration.

Note that conditions for elution may not correspond exactly to those for adsorption, i.e. a protein may not elute at a salt concentration at which it would not have stuck well if you were trying to get it to adsorb. This is called hysteresis, and probably results from nonideality or non-equilibrium processes - we rarely run columns slowly enough that proteins, which diffuse slowly, are truly at equilibrium with the exchanger as they move down it.