

## Protein Purification: Precipitation

Methods for protein separation may be divided broadly into those which divide the protein between **two phases**, usually but not always a solid (precipitate) and liquid (supernatant), and those which separate proteins by different **rates of movement** through some material, usually a chromatographic column but also including electrophoresis. There are also methods which are essentially filtration, separation of proteins by whether they pass through very small holes, whether actually carried out on a filter or in a column (gel filtration). Scopes has three headings, precipitation, adsorption and solution methods, but in practice most of the latter are either two-phase or rate-of-movement methods - gel filtration is usually carried out in a column like adsorption chromatography and is a rate-of-movement method; ultrafiltration has two phases, and separation of two liquid phases is virtually continuous with precipitation since polyethylene glycol can give either a solid precipitate or a liquid lower phase depending on the protein concentration. Two-phase methods resolve proteins much less than rate of movement methods, but are easier to use on large amounts of material, so they are typically used early in a purification procedure, before going on to methods which are harder to use with large volumes.

We should consider briefly the interactions which lead to precipitation. Proteins might stick to each other through one of three forces: electrostatic, hydrophobic, and van der Waals. The last is difficult to distinguish from hydrophobic and operates over only a very short range; it may prevent molecules coming apart, but probably isn't important in their coming together. Electrostatic forces operate at long range, but between like molecules are *repulsive* rather than attractive, since like molecules have the same charge and repel each other. Only when the distribution of surface charge on molecules is very non-random, positive in some areas of the surface, negative in others, is there net electrostatic attraction; and this may lead only to associations of two proteins, not to large scale aggregates which require interactions between many molecules. What we observe is that proteins are most likely to aggregate and precipitate at their isoelectric points, where they bear no net charge and do not repel each other. However, they can be precipitated with non-protein polyanions such as polyacrylate, or protamine sulfate for cationic proteins and ribonucleoproteins.

The main mode of general protein-protein interaction is hydrophobic. Nonpolar patches of the protein surface are shielded by water molecules arranged in an ordered structure; when two non-polar patches come together, the water molecules are expelled and go to a free, less ordered state, which increases their entropy. This increase in the entropy of water molecules, as the number of them solvating hydrophobic surfaces decreases, is the main driving force for protein association. This is all the more effective when the proteins are denatured.

Protein precipitation by removal of the shell of hydrating water, as in  $(\text{NH}_4)_2\text{SO}_4$  or PEG precipitation, is generally reversible, but lyophilization or even solvent precipitation may be irreversible.

We have generally assumed that the proteins are soluble in our extract. However, as mentioned last time, maximum protein solubility is at salt concentration similar to that of the cytoplasm: 0.15 to 0.25 M for eukaryotic cells, 0.3 to 0.6 M in bacteria. T7 RNA polymerase is far more soluble at 0.30 M KCl than at 0.25 or 0.35 M. Proteins which are insoluble at very low salt, requiring 0.2 to 0.3 M to be soluble, are said to be **salted in**. They may precipitate out on desalting by dialysis or gel filtration later in the procedure; they could thus gum up a gel filtration column. Dialysis is described in Rosenberg, pp. 121-3

We should further contrast **negative** and **positive** precipitation methods. A negative method is one which leaves the desired protein active in solution; a positive method precipitates it. Thus negative methods can include selective denaturation procedures which would never yield an active precipitate.

Denaturation typically occurs by exposure to heat, extreme pH, or organic solvents; it could also include selective proteolysis, if the protein of interest is unusually resistant to a general protease which chews up other proteins present, but few take this chance. Heat, extreme pH and organic solvents all work together in denaturation, so when one is varied the other two should be kept very constant to assure reproducibility of the method. Extreme pH generally means low pH, since even denatured proteins usually are soluble at high pH. To test denaturation methods, one simply adjusts some of the protein through a range of temperature, pH or organic solvent concentration, setting aside small samples, centrifuges them, and assays for activity in the supernatant.

Low pH can also precipitate a protein **isoelectrically**, i.e. as the net charge goes to zero the protein, and others of the same pI, associate and precipitate, without necessarily denaturing. If the protein of interest can be precipitated fairly quantitatively without denaturing, and then can be redissolved at another pH and is active, one has a good means of purifying it. All precipitation methods, except those depending on ionic interaction with oppositely charged material, are most effective at the isoelectric point of the protein. About 4/5 of proteins are negatively charged at neutral pH and generally have pIs between 4 and 6; the other 1/5 are positively charged at neutral pH, have basic pIs and are not likely to be precipitable at their isoelectric point. Note that the isoelectric point of a protein is *unrelated* to the pH optimum of its biological activity.

The commonest general process of protein precipitation is **salting out** at high concentration of a salt, usually  $(\text{NH}_4)_2\text{SO}_4$ . It used to be believed that the effect was due to competition with the protein for water molecules, allowing charged groups to interact. It is now believed that the precipitation is due rather to removal of 'bound' water molecules from hydrophobic surfaces of the protein, so that they associate by hydrophobic interaction, which is known to be stronger at high salt. Solubility in  $(\text{NH}_4)_2\text{SO}_4$  *decreases* at increasing temperature, as expected for a hydrophobic effect; this is used in crystallizing proteins, bringing them to an  $(\text{NH}_4)_2\text{SO}_4$  concentration where they are just soluble at 4° and letting them warm up to room temp. so that they will precipitate slowly, though usually in very small crystals, too small for X-ray crystallography.

Solubility decreases in presence of other proteins - i.e. a protein precipitates at a lower  $(\text{NH}_4)_2\text{SO}_4$  concentration in a crude extract, at high protein concentration, than when it is relatively pure. Solubility follows the equation  $\log \text{solubility (mg/ml)} = A - m(\text{salt conc.})$  A is a constant dependent on temperature and pH, while m is independent of these. This equation also makes the point that salting out only **reduces** the solubility;  $(\text{NH}_4)_2\text{SO}_4$  is most useful at protein concentrations above about 0.5 mg/ml. There is a procedure called **reverse dialysis**, in which a dilute solution of the protein is put inside a dialysis bag, and solid  $(\text{NH}_4)_2\text{SO}_4$  outside; about 3/4 of the water inside the bag diffuses out to dissolve the  $(\text{NH}_4)_2\text{SO}_4$ , reducing the volume so that the  $(\text{NH}_4)_2\text{SO}_4$  which diffuses in can precipitate the more concentrated protein. Concentration, but not precipitation, can also be done with PEG or dry Sephadex outside the bag.

Proteins are generally least soluble at their isoelectric point, and in some cases, such as rabbit muscle glyceraldehyde-3-phosphate dehydrogenase, precipitation is achieved by adjusting the pH at constant  $(\text{NH}_4)_2\text{SO}_4$  concentration. But this is only useful when the isoelectric point is unusual; as mentioned, most proteins are isoelectric somewhere between 3.5 and 5.

$(\text{NH}_4)_2\text{SO}_4$  is the salt usually used for salting out, because of its high solubility (about 3.6 M) and high ionic strength (which is proportional to the square of the charge on the ion, so that the ionic strength of 1M  $(\text{NH}_4)_2\text{SO}_4$  is 3 times that of 1M NaCl). Neither ion associates much with proteins, which is good since such association usually destabilizes proteins. Its solubility changes little with temperature, it is cheap, and the density of even a concentrated solution is less than that of protein, so that protein can be centrifuged down from concentrated solutions. One generally uses "enzyme-grade"  $(\text{NH}_4)_2\text{SO}_4$  crystallized from EDTA to minimize effect of contaminating heavy metals. The volume of

the solution increases as  $(\text{NH}_4)_2\text{SO}_4$  is added; the solubility is 533 g/L solution, but 761 g/L original solution. One thing one should remember is that because ammonia is a weaker base than sulfuric is an acid, the pH tends toward about 5.3, base (usually ammonium hydroxide) must be added to hold the pH at 7.0. Alternatively, if your protein doesn't mind pH 5.3, it may be least soluble at that pH, which may be near its isoelectric point.

The concentration is frequently expressed as per cent saturation, partly because lower concentrations may be achieved by adding saturated solution to the original protein solution. The table I am giving out tells how to get from any 5% step of saturation to any other, i.e. how much to add per 100 ml solution. Proteins generally precipitate over about a 15% range of saturation, and you may not achieve much purification if your protein precipitates in the same range as many others; but the two other purposes of  $(\text{NH}_4)_2\text{SO}_4$  precipitation are **concentration** - to diminish the volume in which the protein is precipitation are **concentration** - to diminish the volume in which the protein is dissolved, for instance after column chromatography - and for **storage**, proteins frequently are particularly stable as  $(\text{NH}_4)_2\text{SO}_4$  precipitates.

When adding  $(\text{NH}_4)_2\text{SO}_4$  you should have good stirring and slow addition, to prevent local high concentrations of the salt and consequent precipitation of proteins before they should. Testing the concentration at which your protein precipitates is somewhat laborious, if you either bring individual samples to various % saturation by adding individually weighed amounts of  $(\text{NH}_4)_2\text{SO}_4$ , or bring one sample to a given level, remove a sample, measure the volume, and add more  $(\text{NH}_4)_2\text{SO}_4$ . You must then correct your activity determinations for the increase in volume of the solution as  $(\text{NH}_4)_2\text{SO}_4$  is added. A better way is by **back-extraction**; you precipitate at 90 or 100% saturated, pipet the suspension into a number of small centrifuge tubes, spin down the precipitate, and stir up the pellets (which are identical) in solutions of decreasing  $(\text{NH}_4)_2\text{SO}_4$  concentration (prepared by mixing saturated  $(\text{NH}_4)_2\text{SO}_4$  and buffer). After mixing well you spin down the remaining precipitate, what isn't resolubilized, and measure activity in the supernatant solution. This allows you to measure activity in a constant sample size, and avoids individual weighing of  $(\text{NH}_4)_2\text{SO}_4$ . However, you must remember that the pellet was 90 or 100% saturated, so that the protein will stay in the pellet at a lower concentration of added solution than if you had reached that % saturation 'on the way up', by adding solid to the original solution. This can be minimized by making the volume of extracting solution large in comparison to that of the pellet. This method of fractionation often achieves better fractionation, the protein going from completely precipitated to completely dissolved over a narrower  $(\text{NH}_4)_2\text{SO}_4$  range than achieved by adding solid  $(\text{NH}_4)_2\text{SO}_4$ , and in some cases the protein will crystallize out from an  $(\text{NH}_4)_2\text{SO}_4$  solution in which it had just been extracted from an amorphous precipitate. [Even greater fractionation can be achieved by carrying out back-extraction in a column; the precipitated protein is mixed with Celite, a diatomaceous earth, as a flow aid, and poured into a column; the column is eluted with a gradient of decreasing  $(\text{NH}_4)_2\text{SO}_4$  concentration.]

Another method is **solvent precipitation**. When large amounts of a water-miscible solvent such as ethanol or acetone are added to a protein solution, proteins precipitate out. The conventional wisdom is that this is due to decrease of the dielectric constant, which would make interactions between charged groups on the surface of proteins stronger. However, Van Oss has found that ethanol does not decrease the dielectric constant of water much, indeed 20% EtOH at  $-5^\circ$  has the same dielectric constant as water at  $20^\circ$ . He finds that ethanol associates with water much more strongly than do proteins, so that its real effect is to dehydrate protein surfaces, which then associate by van der Waals forces, at least if they are isoelectric or reasonably close to it. Removal of water molecules from around charged groups would also deshield them and allow charge interactions to occur more strongly, if you have areas of opposite charge on the surfaces of two proteins. Salts tend to bind to protein surfaces and make them less isoelectric, and therefore tend to mess up ethanol precipitation, which should be carried out at low salt.

In practice, one usually carries out solvent precipitation at low temperature: the protein is at 0° and the solvent colder, -20° in an ice-salt bath, because proteins tend to denature at higher temperatures - though if sufficient control can be achieved and your protein is more stable than others, this can be selective and achieve greater purification. I once found I could purify phosphotransacetylase efficiently at +6° to +8°, and set about holding this temperature with a freezing benzene bath, in a cold room. My liver recoils at the thought. One adds the solvent slowly, with good mixing. One would do a test precipitation with a small amount of solution of the protein, taking out very small samples at various amounts of solvent added, centrifuging, and assaying the supernatant to find out when the protein was precipitated. Then one would assay the precipitate to find out if the protein was precipitated without denaturation and is active upon redissolution. Redissolved proteins perhaps should be dialyzed to remove traces of the solvent, as the traces may affect behavior in other methods. You must record the volume of solvent added, as the volume of the solution will be less than the sum of the individual volumes; for instance, adding 50 ml of ethanol to 100 ml of extract yields a solution we would call 33% ethanol, but only 140 ml of it.

A related method is precipitation with polyethylene glycol, at low concentrations, 5 to 15%. It probably works the same way, by competing with the protein for water, but is less likely to inactivate the protein and does not require such low temperatures. It tends to give an oily precipitate, and may simply give a second, protein-rich liquid phase. I'll talk about liquid-liquid phase sep. later.

[On the other hand, Klivanov - at MIT, the great proponent of working with proteins in non-aqueous solvents - has found that proteins are soluble in pure dimethyl sulfoxide, and while they don't have their native conformation in DMSO - enzymes are not active in it - they return to the active conformation upon dilution in water. He dissolves lyophilized proteins in straight DMSO, and can carry out ion exchange chromatography in it, or precipitate proteins by adding solvents such as ethyl acetate. The advantage of this picturesque procedure is that proteases, which require water as the other substrate, are not active in DMSO; so this might be useful if you had severe protease problems in crude extract. Reference: Chang, Hen & Klivanov, *Biochem. Biophys. Res. Commun.* **176**:1462-1468 (1991).

Old volumes of *Methods in Enzymology* mention somewhat specific precipitants such as Zn<sup>++</sup> reacting with imidazole groups, Hg<sup>++</sup> and Cu<sup>++</sup> with SH, Pb<sup>++</sup>, Fe<sup>+++</sup>, Ba<sup>++</sup>, all at mildly alkaline pH, acids such as trichloroacetic, phosphotungstic and sulfosalicylic at acid pH. The acids, however, usually denature proteins irreversibly, though they are sometimes used to precipitate all protein from a dilute solution for measurement, e.g. by the Lowry method, or before SDS gel electrophoresis. Zn<sup>++</sup> precipitation has had a renaissance, particularly for precipitation of proteins in culture broths, excreted by cells, bacteria, yeast or Chinese hamster ovary. Reference: Zaworski & Gill, *Anal. Biochem.* **173**:440-4 (1988). The broth is adjusted to pH 5, 1 M ZnCl<sub>2</sub> added to get concentrations from 0.1 to 50 mM, the pH adjusted back to 7.0, and precipitated proteins centrifuged down. The protein is redissolved by suspension in 0.1 to 0.25 mM EDTA. Precipitation varies from protein to protein, which makes it selective - recombinant porcine urokinase was essentially completely precipitated at 1 mM Zn, while BSA and interleukin-1β were not precipitated at all at 5 mM and only 75-85% at 25 mM. Some yeast proteins were not precipitated at even 50 mM, which purifies them considerably. But the best use of this would be to concentrate from dilute solution a protein precipitated with only a little Zn.]

I mentioned earlier adding an oligohistidine tail to the recombinant gene for the protein. This can also be used for precipitation, if the protein is at least dimeric, by linking protein molecules together with a bis-zinc or bis-nickel reagent. Van Dam et al., *Biotechnol. Appl. Biochem.* **11**: 492-502 (1989) used bis-copper compounds. A longer chain between the metal ions is better. In principle, any oligomeric protein can be precipitated using a specific bis-functional ligand, for instance bis-NAD<sup>+</sup> for NAD<sup>+</sup>-dependent dehydrogenases

Enzymes acting on insoluble polymeric substrates can often be purified by batch affinity precipitation. They are adsorbed on the substrate, which is then centrifuged down, and either the protein

is eluted with high salt concentrations, or it is simply allowed to digest the substrate into soluble products which are dialyzed away. Amylases, yeast cell wall-degrading enzymes and elastases have been purified in this way.

### Liquid phase partitioning

When two polymers, typically dextran (an  $\alpha$ -1,6-linked glucose polymer, with 5%  $\alpha$ -1,3 links, produced by the bacterium *Leuconostoc*) and polyethylene glycol are dissolved in water in appropriate proportions, say 6% PEG and 8% dextran overall, two phases develop, a dextran phase on the bottom and a PEG-rich phase on top. This can be described by a phase diagram, a plot of concentration of PEG and dextran, with a curved line across it; any total composition corresponding to a point above the curve will yield two phases, with compositions indicated by the ends on the curve of a 'tie line' through the point. (None of my references indicate *which* of the many possible tie lines intersecting with the curve will indicate the correct compositions) Other points on the tie line will yield the same phase compositions, but different amounts of them. Proteins distribute themselves between the two phases. PEG and high concentrations of salts (phosphate, citrate) can similarly generate two phases. A whole book has been written about this by Albertsson, *Partition of Cell Particles and Macromolecules*, 3rd edition 1986, John Wiley; a review is by Huddleston & Lyddiatt, *Applied Biochem. & Biotechnol.* 26:249 (1990). Most proteins are preferentially soluble in the lower, dextran-rich phase, with distribution coefficients ( $K = [\text{in PEG}]/[\text{in dextran}]$ ) as low as 0.01 for phosphofructokinase, some almost equal (0.58-0.86 for ovalbumin), some high (1.9 to 42 for phycoerythrin);  $K$  generally increases with increasing molecular weight for a PEG-dextran system, but goes down for a PEG-salt system. The behavior as a function of phase composition follows the empirical equation  $\ln K = A(w_t - w_b) + b(w_t - w_b)^2$  or in a linear form  $\ln K/(w_t - w_b) = A + b(w_t - w_b)$ , where  $w_t$  and  $w_b$  are weight % PEG in the upper and lower phases and  $A$  and  $b$ , intercept and slope of the linear form, are empirical constants characteristic of the protein. Decreasing the mol. wt. of one polymer puts more of the protein into that phase, an effect which increases with molecular weight of the protein (little change in  $K$  for cytochrome *c.*) At high concentrations all salts tend to put proteins into the PEG phase vs the dextran phase, but in PEG/salt systems proteins go to the lower phase.

Proteins themselves, and other materials such as cell wall fragments, can form a phase when present at high enough concentration. This suggests two consequences: both PEG and  $(\text{NH}_4)_2\text{SO}_4$  precipitation amount to an extreme case of phase separation, in which the protein forms the other phase but is so concentrated that it precipitates; and the method is particularly useful as a first step in purification, separating protein from cell debris etc. more easily than by centrifugation. One can make a crude homogenate, at high cell concentration with some salt, and extract with a PEG solution; proteins go into the top phase, usually 85-99%, while insoluble cell debris stays in the bottom phase. The top phase is then extracted with a high-salt lower phase, such as 11% sodium citrate pH 7.0; again the yield is typically 85-99%.

[A further improvement is PEG with affinity ligands attached, so that proteins binding to the ligand go into the upper phase (under conditions where not much bulk protein would go into that phase). The upper phase is then extracted with a high salt solution, which not only generates a lower phase into which proteins go, but loosens the specific binding to the affinity ligand. It should then be possible to clean up the upper phase for re-use. It can also be run, with greater resolution, as column chromatography, preequilibrating the column with the dextran phase and running the PEG phase past it. With an affinity ligand this is essentially affinity chromatography as previously described, except that the protein is never adsorbed, only partitioned. However, equilibration is slow and the advantage of high capacity is lost. Another possibility is counter-current distribution, essentially a whole series of phase separations, a very powerful if time-consuming procedure.]

A recent paper (Dennison & Lovrien, *Protein Expression and Purification* 11:149-161 (1997), drawing on earlier work by Lovrien, describes another phase separation technique: while *t*-butanol is

normally completely miscible with water, addition of enough salt, usually  $(\text{NH}_4)_2\text{SO}_4$ , causes separation into two phases; and if protein is also present, it tends to precipitate out as a *third* phase between the lower aqueous and upper *t*-BuOH phase; they call this "three phase partitioning". *t*-BuOH and  $(\text{NH}_4)_2\text{SO}_4$  act in similar ways - they list six effects of  $(\text{NH}_4)_2\text{SO}_4$ , ionic strength effects, what they call kosmotropy (I don't know what this is), cavity surface tension enhancement, osmotic dehydration, exclusion crowding, and conformational shrinkage due to sulfate binding to positively charged sites on proteins. They think that *t*-BuOH also binds and tightens conformation, as well as being a kosmotrope in solution. Also, of course, lipids tend to go into the *t*-BuOH layer, so this is a good way to solubilize membrane-bound enzymes, as originally suggested by Morton, using *n*-BuOH, in the 1950s. Dr. Ward has tried this method with green fluorescent protein, it works like a charm, and he frequently thanks me for it.

They suggest proceeding as follows: (i) add sulfate, to half to 3/4 the concentration necessary to salt out by itself; (ii) adjust the pH - the pH should be 2 to 4 pH units *below* the pI of the protein sought, but of course this may not be possible if it has an acidic pI; however, the sulfate protects against acid denaturation. BSA is precipitated quantitatively below pH 5.2, but it is stable down to 3.0. The pH of best effect should be explored. (iii) In optimizing the method you can play with temperature, though this does not have to be done in the cold as with ethanol or acetone precipitation; (iv) add *t*-BuOH, 0.2 to 0.5 ml per ml of aqueous solution. The phases tend to separate out even without centrifugation, but low speed centrifugation speeds this.

The precipitated protein is greatly concentrated - often 100-fold - usually increases considerably in specific activity, and in some cases the yield is much greater than 100%, presumably due to separation from an inhibitor (their best case is yeast invertase, 500-1000% yield with 75x increase in specific activity, but invertases are often accompanied by inhibitors). It is usually collected by pipetting off the upper and lower layers. The precipitate then can be dissolved by adding water, or if necessary buffer at a higher pH. In some cases the precipitate contains so little sulfate that one can proceed directly to ion exchange chromatography; free sulfate is easily dialyzed away, bound sulfate isn't. The proteins are much more easily separated from *t*-BuOH and  $(\text{NH}_4)_2\text{SO}_4$  than they are from dextran and PEG, and the costs are less (*t*-BuOH is 10% the cost of dextran, 50% that of PEG).

The paper goes on to discuss the thermodynamic effects of sulfate in protein precipitation.