

Affinity chromatography

This refers to the use as adsorbents in chromatography of materials with groups which are supposed to be specific ligands of the protein being purified. Within the general term are often included other methods which depend on a specific interaction, but not that which is functional for the protein, binding to known ligands of the protein in question, usually small molecules which have been attached to a solid support. These others include: dye ligand chromatography, the use of a number of synthetic dyes which have by a combination of pure chance and systematic if not particularly rational search been shown to bind a variety of proteins tightly; lectin chromatography, lectins being plant (usually) proteins which bind specific sugars and may thus be used on immobilization for chromatography of glycoproteins; and immunoabsorbent chromatography, the use of antibodies specific to the protein being purified. Most of the principles I shall mention apply, more or less, to all of these.

The **matrix**, or support, is usually agarose or a cross-linked derivative, because it is very porous and admits large proteins to the pores, but has good strength and stability, and is reasonably derivatizable. In general any matrix useful for ion exchange or gel filtration chromatography is also good for affinity chromatography. It has been observed (Narayanan, Knoch and Crane, *J. Chromatogr.* **503**:93-102 [1990] - Jane Crane is a graduate of our Biochemistry program) that there is an optimum pore size for the matrix in affinity chromatography, because of two opposing physical effects: the pore size must be large enough to admit the protein being chromatographed (or immobilized), but on the other hand large pores mean less total surface area and thus lower adsorptive capacity. Just what the optimal pore size is depends on the protein, but a pore diameter of 200 Å is generally good, particularly for proteins $\geq 150,000$ kDa.

The attachment usually proceeds by treating the matrix with a reactive compound which either leaves reactive groups which can be displaced by amino groups of the spacer or ligand, or leaves another reactive group on itself. In the first category are cyanogen bromide, which leaves a variety of compounds, of which cyanate esters are probably the most important; glutaraldehyde, whose free aldehyde group forms Schiff bases with amino groups, which then are reduced with sodium borohydride or cyanoborohydride to create stable N-alkyl linkages; carbonyldiimidazole, yielding an N-carboxyimidazole ester; tosyl chloride, yielding a tosyl ester which can be displaced by an amine, leaving the amine attached directly to the sugar ring of the matrix, more stable than other bonds. In the second category are diglycidyl ethers = bisoxiranes, a chain with an ethylene oxide ring on either end, one to react with the matrix, the other with the ligand; dichlorotriazine and related compounds; and divinylsulfone.

The ligand is usually attached with a **spacer arm** between it and the matrix, to assure that the ligand will be fully accessible to the desired protein. One may either attach the linker arm to the ligand, then react this extended ligand with the activated matrix (6-aminohexylAMP is a good example), or react a nonspecific spacer such as 1,6-diaminohexane with the activated matrix, then couple the ligand to it, often by activation of a carboxyl group with a carbodiimide and displacement of the carbodiimide by the linker amine (or use ϵ -aminocaproate as the spacer, activate its carboxyl group with carbodiimide, and react ligand amine with that). There are more possibilities than there is time to discuss. Of course some of the activating compounds above, such as bisoxiranes and divinylsulfone, generate a spacer arm themselves. Hydrophobic chromatography was discovered as a result of systematic study of spacer arms, which in many cases turned out to adsorb proteins even without a specific ligand on their ends. The next step was to make hydrophilic spacer groups, for example with an

amide link in them, but often these turned out to make poorer affinity materials than those with alkyl spacers; the hydrophobic spacers were contributing to the specific affinity binding. Scopes expands on the need for nonspecific interactions in addition to the specific interactions, to add up to a dissociation constant $\leq 10^{-6}$ M, needed because the concentration of bound ligands able to bind the desired protein is often quite low, 0.01 mM or so, 1-2% of the total bound ligands.

How the spacer arm is attached to the ligand is important, as it should not interfere with ligand binding to the protein. For instance, NAD⁺ and AMP were attached to matrices in various ways, through the phosphate, the ribose hydroxyls, C-8 of the adenine and N-6 of the adenine; the last proved to be the most useful, and X-ray crystallography eventually showed that the adenine nucleotide binding pocket, conserved across a wide range of proteins, has N-6 pointing up out of it, where an attachment does not interfere with binding. But usually the best way to attach a ligand has to be worked out by trial and error, synthesizing small test molecules with alkyl groups attached to the ligand in various ways and determining which bind best to the protein.

Many materials for affinity chromatography are commercially available, complete (if they bind many proteins or a very important one - for example 5'-AMP-agarose), with spacer arms, or just activated. This in particular avoids the use of nasty chemicals such as cyanogen bromide and diglycidyl ethers, but allows one to attach one's own ligand (if it has an amino group which can be modified without loss of binding to the protein). Often one uses cyanogen bromide activated agarose for small scale testing of ligands, then switch to another means of attachment to make a material which will last longer for regular use.

Elution from affinity columns may be with the free ligand in solution, which competes with the bound ligand for the protein's binding site; or may be a non-specific method such as high salt. Pure elution with free ligand may require an uneconomically high concentration; it is better to develop conditions where the protein is not quite eluted non-specifically, then pull it off with a relatively low concentration of free ligand. However, these conditions may also weaken binding of the free ligand! You want to weaken the **non**-specific binding forces, then compete for the specific binding site with free ligand; unfortunately you don't necessarily know whether the non-specific forces are polar (in which case high salt will weaken them) or hydrophobic (in which case you want **low** salt, or perhaps a decrease in temperature, or inclusion of ethylene glycol or ethanol in the buffer, or a detergent). With moderately general affinity materials such as immobilized 5'-AMP or NAD⁺ you want conditions as specific as possible for your protein. For instance, lactate dehydrogenase generally forms tight complexes with NAD⁺ and pyruvate; one therefore washes the column with NAD⁺ not quite high enough to elute the lactate dehydrogenase, then elutes with NAD⁺ at the same concentration plus a good concentration of pyruvate.

If you cannot elute your protein with a specific eluant, you may have to use strong non-specific means - high salt, high or low pH - but also cold distilled water, which minimizes hydrophobic interactions!

After you have eluted your protein, you generally clean the column by eluting with a strong non-specific eluant - KSCN, urea, sodium dodecyl sulfate - then wash this out, and store it with a preservative such as azide.

Dye ligand chromatography

Pharmacia Fine Chemicals attached a blue dye known as Cibacron Blue F3GA to a dextran of molecular weight 3 million, to serve as a visible marker for the void volume in gel filtration. People

fractionating their protein by gel filtration, with this marker, known as Blue Dextran, included, found their protein came out in the void volume. It turned out that this dye looks to proteins very much like AMP and generally binds to proteins which bind AMP; thus these proteins bind to the Blue Dextran and move through gel filtration columns in the void volume. After an intermediate phase in which the whole Blue Dextran was bound to supports and used for affinity chromatography, the dye was attached directly to agarose and other matrices and used as a general affinity material for such proteins - commercially attractive because many proteins bind adenine and guanine nucleotides, and experimentally attractive because it has a high capacity and high affinity for them. Later people began to try other such dyes; a red dye called Procion Red H-E3B seemed to show specificity for NADP-linked dehydrogenases, though with more experience it turned out not to be so specific. Some 40 such dyes have now been tried as ligands for protein purification - mostly by Scopes; his book has 15 pages on this method. He suggests trying 5 dyes of differing ability to bind proteins in general, in order to find out how tightly your protein sticks to such dyes, and then trying other members of the weakest class to which it binds, to find which one is best, i.e. binds your protein reasonably tightly but not too many other proteins. In many cases you also identify a material which binds many other proteins but not yours, and run your protein solution through that column first to remove proteins which otherwise would be competing with your protein for the material to which it binds. Such methods are particularly useful when you want to purify many proteins from the same source, and can fractionate them by a series of dye ligand columns; Scopes has thus purified all the glycolytic enzymes of *Zymomonas mobilis*.

Elution is as with 'true' affinity columns, either with specific ligands which compete with the dye for the protein's binding site, or with high salt concentration or high pH. If the dye ligand has been selected as not binding the protein very strongly, the chances are better that a natural ligand will elute it, without having to go to high salt or pH. In some cases the binding of the protein to the ligand is potentiated by presence of metal ion, and buffer without the metal ion is all that is needed to elute the protein. Thus though the method as such is not necessarily as specific as 'true' affinity chromatography using a natural ligand, optimization of the several variables involved can make it a very effective procedure, in some cases yielding a pure protein directly from crude extract (Scopes uses this procedure as the first step in purification. I have used it so in one case, where the crude extract - of tomato seeds - contained something which prevented redissolution of my enzyme after $(\text{NH}_4)_2\text{SO}_4$ precipitation.).

The dyes, which after all were originally synthesized to react with and dye fabrics such as cotton, have reactive groups such as chlorotriazine or divinylsulfone, and react directly with the agarose or other matrix, though elevated temperatures are often required, to which the matrix must be stable. He notes that in many cases what is commercially available is a mixture of related compounds; you might purify this mixture to define the reactive ligand better, and you might end up with a major but inactive component. Also, the dye may cease to be commercially available. Amicon sells a number of immobilized dye ligands under the general title Matrex.

Immunoabsorbent chromatography

This is simply using antibodies to your protein as the specific immobilized ligands. In principle this is both general and the last word in specificity and tight binding, but of course there are drawbacks. First of all, in order to prepare the antibodies you may have to purify the protein first, **without** this technique, in order to have something with which to immunize the rabbit; but this can be done with a few hundred μg of purified protein, purified by two-dimensional gel electrophoresis and injected as mashed gel. It can be done with incompletely pure protein, but you then run the risk that you will continue to coisolate the impurities present, to which antibodies have also been raised. But you can make

monoclonal antibodies to the impure preparation, then test various clones to find one which reacts with your protein. Or you may use a synthetic peptide, from the gene sequence, as antigen. The antibody fraction is purified somewhat, by precipitation with $(\text{NH}_4)_2\text{SO}_4$ or polyethylene glycol or chromatography on "T-gel" or immobilized Protein A from *Staphylococcus*, then immobilized. Getting your protein to stick is not hard, but releasing it in an active, native form is. Generally either very low pH (2 to 3, glycine buffer) or chaotropic salts such as thiocyanate or iodide are needed, and may leave the protein denatured. One idea is to run the eluate directly onto a gel filtration column, so that the protein is separated from the eluting salt as fast as possible. The ideal eluant loosens up the protein or the antibody enough to weaken the specific interactions, but does not completely denature it. Some examples are 2M guanidine HCl, 0.6 M K thiocyanate, and dilute ammonia. An elegant approach is to make antibodies not to the exact protein wanted, say a human protein, but to the same protein from another source, say horse; the antibody-antigen interactions will then be less perfect and the antigen more readily eluted. But the capacity of the column may be less, and some epitopes of the protein may be exactly the same so that your protein still binds tightly. If you are making a monoclonal antibody you may be able to choose one with a moderate dissociation constant, say 10^{-6} - 10^{-8} M, rather than the tightest available, 10^{-12} M. Monoclonal antibodies also have advantages in having only a single dissociation constant, rather than a range as with a polyclonal mixture, and in making higher capacity columns, since all the antibody present is specific for your protein. The cell line can be maintained or frozen, so that the same antibody can be produced at a later date, rather than dying with a specific rabbit. However, there is more work involved.

Antigen-antibody reactions are slow, so that for getting the antigen to stick to the column you typically apply a volume less than that of the column and let it sit overnight to adsorb as fully as possible. Elution may also be slow, particularly with a mild, non-denaturing eluant.

Some general comments

It should be apparent that we have been moving from general techniques able to handle large amounts of material cheaply to specific techniques best applied on a very small scale, though able to give completely pure protein. Industry, however, may be able to run even such techniques on a fairly large scale, particularly when an important criterion is convincing the FDA that their product is *completely* pure and therefore safe to license for sale; the \$1500 price for one treatment with recombinant tissue plasminogen activator what drives this. On the other hand, the choice of techniques may be changed by how workable they are on a large scale; for instance, cross-flow filtration may be a better way to separate and concentrate cells than centrifugation, and aqueous phase partition may be a better way to remove cell debris, because large scale centrifuges are harder to keep running and require a lot of power.

An article by Leser and Asenjo (*J. Chromatogr.* **584**:43-57 [1992]), though entitled "Rational design of purification processes for recombinant proteins", has useful comments on all industrial processes and indeed all protein purification. They have five rules, some of which should be self-evident by now:

"Rule 1: Choose separation processes based on different physical, chemical or biochemical properties." Repeating the same process doesn't gain much, though sometimes it makes a later step more efficient.

"Rule 2: Separate the most plentiful impurities first." This means especially non-protein impurities such as cell debris and small molecules.

"Rule 3: Choose those processes that will exploit the differences in the physicochemical properties of the product and impurities in the most efficient manner." This is easiest when you know the properties of the purified protein and are designing a large scale process, which you want to be as efficient as possible. To purify a recombinant protein you'd also like to know properties of the commonest proteins of the host cell, and how to remove them efficiently.

"Rule 4: Use a high-resolution step as soon as possible." This is less obvious, but eliminates as many impurities as possible at an early stage. They mean affinity chromatography where possible, otherwise probably ion exchange chromatography.

"Rule 5: Do the most arduous step last." This means removal of the last few per cent impurities. High resolution gel filtration is often the best step here.

A further point they make for industrial processes is that if the product is to be used in humans, the whole process has to be approved by the FDA and can't be changed later, so you had better make it as efficient as possible to begin with. The need for this caution is shown in the case of the fatal disease caused by some preparations of L-tryptophan used as dietary supplement. The Japanese producers had upped the production of tryptophan by so much using a recombinant organism that they skipped a step previously used in purification. Unfortunately this meant they didn't get rid of a toxic by-product removed by the previous process.