

Before continuing with today's material, I want to say a little about the quiz last time. The point of the quiz of course was to bring it home to you that for this course you need to know the properties of the amino acids, because they to a first approximation comprise those of proteins.

The amino acids with side chains that ionize are 7: aspartic and glutamic acids, arginine, histidine and lysine, tyrosine and cysteine. Arginine is always protonated, since its pK_a is about 12.5, and lysine is almost always protonated, since the "normal" pK_a is about 10, but in special cases the pK_a can be as low as 8 and the unprotonated form be significant. Similarly, aspartic and glutamic acids are normally deprotonated, with a pK_a of 4, but not uncommonly the pK_a is higher and the protonated form is significant. The phenolic OH of tyrosine has a $pK_a \approx 10$, but rarely ionizes except when the protein is denatured in base, because tyrosine is otherwise a hydrophobic amino acid, likely to be buried in the interior of the protein – there are ways to estimate the % exposure of tyrosines. But the OH may well be involved in a hydrogen bond. Histidine, with a pK_a in proteins of about 7, and cysteine, with an average pK_a for the side chain of perhaps 8.5, are most likely to have both protonated and unprotonated states involved in enzyme mechanisms.

Of course α -amino groups and carboxyl groups also ionize, but there is only one of each per polypeptide chain – the others are in peptide linkage, non-ionizing but able to form hydrogen bonds which are the basis of secondary structure.

The remaining amino acids comprise those with polar but non-ionizing side chains – serine and threonine deprotonate only above pH 13, and asparagine and glutamine protonate the amide only below pH 1 – and the purely hydrophobic, all the others unless you make glycine a special class because it has no side chain at all. The polar but non-ionizing do participate significantly in hydrogen bonds, as do those with ionizing side chains, and potentially the indole nitrogen of tryptophan, which also protonates only below pH 1. For the rest, remember that CH_2 and CH_3 groups, and aromatic side chains, are as hydrophobic as salad oil and benzene. And never draw structures with five bonds to carbon!

The next few lectures are concerned with protein purification. You may want to do this to study the protein physically (what is its size, shape, 3-D structure, etc.) or chemically (what is its amino acid sequence, post-translational modification, etc.), both of which types of study require that the total protein present is substantially the one you want to study - at least 95% of the total. If you want to clone the gene or cDNA for the protein, you want either to determine at least some of the amino acid sequence, in order to design oligonucleotide probes for the cDNA sequence, or to make antibodies to the protein, to look for clones expressing it. In the latter case you may not have to purify it completely, because you can also select a monoclonal antibody for it; but this is work of another kind.

Study of its biological activity requires less purification, because you have only to remove other proteins which interfere with this study, but the only way to be sure that one is free of such interference is to remove all other proteins. Nowadays, you may want to purify the protein for therapeutical purposes, and need to have it very pure indeed. The usual way to determine if the protein is pure is gel electrophoresis, which will be covered much later.

In purifying a protein, we are trying to keep it, specifically, and throw away all other proteins, until we have **only** the protein we want. To measure whether progress is being made in purification you need to measure the '**specific activity**', the ratio of the amount of the protein you are interested in, as measured by some **specific** assay, to the amount of **all** protein present, as measured by a general assay of all protein present. You determine the efficiency of a given purification step by how much the specific activity is increased after it. And sometimes you need to measure the protein concentration because it is critical for the success of a given step - for instance, $(NH_4)_2SO_4$ precipitation won't work if the solution is too dilute. On the other hand, there are times when specific activity is less important than total activity. For instance, if you want to test whether a protein is being secreted into culture medium at a higher level, you care about total activity in the medium; increased specific activity might just mean that less *other* protein was being secreted.

Protein determination

We should therefore discuss how we measure protein in general. Most of you have probably done this at some time, but it is worth discussing various assays, since which you use may depend on a particular situation.

A protein assay should have the following characteristics, usually but not always in about this order of importance: it should be sensitive (i.e. measuring low levels of protein), quick, specific (i.e. not measuring other compounds present, nor affected negatively by their presence), non-destructive, and invariant from protein to protein (i.e. giving the same results per mg of protein present). No one assay is perfect with respect to all considerations - else it would have displaced all others - but which you use may depend on which of the above considerations is most important at the time. Note also that almost all assays do give different results with different proteins, because they are affected by the amino acid composition of the protein, and therefore cannot be said to be completely accurate unless they have been standardized using a solution of the protein you are interested in, whose concentration has been determined by the dry weight of protein per ml. And even that is not accurate if the protein contains a significant amount of non-protein material such as carbohydrate. As a standard for the method used during purification one usually uses bovine serum albumin, which is cheap and average in composition; one prepares a standard curve, measuring the absorbance of known amounts. However, it gives more color in the Coomassie Blue assay than the average of other proteins. For best accuracy one should dry down a measured volume of the standard solution to obtain the actual weight concentration.

To comment on those on the sheets I am handing out, and a few others: the **biuret** reaction, which measures the absorbance at 540 nm of Cu^{++} bound to peptide bonds, is the most invariant from protein to protein - because it is a reaction of the peptide bond, and therefore unaffected by amino acid composition, though the **weight** of protein will be affected by composition, polytyrosine obviously weighing more per peptide bond than polyglycine. It is also very insensitive, requiring at least 1 mg, and destructive, using strong base; it is therefore rarely used nowadays. However, it is more sensitive in the ultraviolet; the methods of Goa and Itzhaki & Gill take advantage of the fact that the absorbance of the biuret complex is greater in the near UV, 310 or 320 nm, but it's not a peak.

A variation of the biuret reaction is that of **Klungsoyr**. Cu^{++} complexed with protein in 0.1 M NaOH moves with it through a little Sephadex column, while free Cu^{++} is adsorbed at the top of the column. One can then measure the Cu^{++} carried through as its complex with diethyldithiocarbamate. This is as sensitive as the Lowry method - 0.05 to 0.4 mg - but as invariant as the biuret method. But it is not adapted to doing many samples routinely, since each must be chromatographed. (The Sephadex spin column methodology much used in molecular biology for small samples would be useful here.) It can be useful to give a fairly absolute measurement of the concentration of a pure protein, which can be translated into a molar concentration when you are chemically modifying the protein and want a ratio of moles modifier/mole protein. Measuring protein by a method depending on reaction with the side chains may be inappropriate if you are chemically modifying the side chains which react in the method.

The most frequently referred to paper in biochemistry - and in all science - is that by **Lowry**, Rosebrough, Farr and Randall, describing what is also called the phenol or Folin method, which uses a combination of the biuret absorbance of Cu^{++} bound to the peptide bond and Cu^+ produced by reduction of the Cu^{++} by tyrosine, cysteine and tryptophan; the Cu^+ is measured by its reduction of arsenomolybdate, as in the Nelson-Somogyi method for reducing sugars and the Fiske-SubbaRow method for inorganic phosphate. This method is fairly sensitive, measuring 0.05 - 0.4 mg protein, and only moderately variable from protein to protein; but a wide variety of compounds which may be present in a protein solution interfere, notably SH compounds such as 2-mercaptoethanol and compounds with the $\text{HNCH}_2\text{CH}_2\text{O}^-$ structure such as triethanolamine and morpholine buffers. The interference can generally be dealt with, as by first precipitating the protein with a combination of trichloroacetic acid (TCA) and deoxycholate, but this makes even slower an already relatively slow procedure, requiring 40 minutes' incubation besides pipetting time. It is also not strictly linear, though if only absorbances less than 0.4 are used this is not a problem.

A related procedure has been ballyhooed by Pierce Chemical Co. It is based on the chelation of Cu^+ by **bicinchoninic acid** and the absorption of this complex. The method is apparently somewhat less affected by interfering compounds than the Lowry method, but it seems to me that any compound capable of reducing Cu^{++} to

Cu^+ would still interfere strongly. The color development is probably faster than in the Lowry procedure, but generally incomplete; sensitivity is increased by incubating at a temperature such as 60° . This would tend to cause variability from assay to assay and require a standard curve each time.

The main current competition of the Lowry method, probably now much more used, is the 'dye-binding' or **Coomassie Blue** or Bradford method. BioRad sells the reagent made up, but unless money is no object it is cheaper to make up your own; however, a standard curve should be run regularly. The method depends on the fact that a solution of the dye Coomassie Blue G-250 in acid - phosphoric or perchloric, pH about 1 - is orange, λ_{max} 466 nm, but protein displaces two protons from the dye, generating a blue form which is measured at 595 nm. There is a good deal of background absorption by the reagent, which can cause variation in measurements, particularly if poorly matched cuvettes are used; usually one is adding a small amount of protein to a large amount of the reagent, so that errors due to variation in pipetting the reagent are small. The method is perhaps 4x more sensitive than the Lowry method, and much less affected by interfering substances, except for one important one: sodium dodecyl sulfate (SDS), which binds to and favors formation of a neutral green form, $\lambda_{\text{max}} = 650$ nm. It should be remembered that any compound present in high enough concentration to shift the pH into the range where the dye is blue anyway will also interfere; a large amount of a high concentration of $(\text{NH}_4)_2\text{SO}_4$ could for instance do this. The method is moderately variable from protein to protein; the paper by Compton and Jones showed that binding is mainly to arginine groups, to a lesser degree hydrophobic. It can be non-linear at very low protein concentration; the paper by Zor & Selinger shows that this is due to equilibration between the three ionic forms, and can be linearized and made more sensitive by reading both A₄₆₆ and A₅₉₅ and plotting A₅₉₅/A₄₆₆. This is more trouble, unless you have a spectrophotometer which reads at two wavelengths and give you the ratio automatically; but it expands the useful concentration range both down and up, and makes it more sensitive because the A₅₉₅/A₄₆₆ ratio changes more with protein concentration than the A₅₉₅ alone. The other disadvantage is that the reagent tends to stick to cuvettes, particularly good quartz cuvettes. If you can't afford a single use of disposable cuvettes, you can remove the blue stain with acetone (not for plastic cuvettes!), 0.1 M NaOH, or SDS.

The most important other method is **ultraviolet absorbance**. Absorbance at 280 nm is quite variable from protein to protein (depending on the content of tyrosine and tryptophan), subject to interference in very crude extracts (though nucleic acid content can be corrected for by A₂₆₀), and only middling sensitive (an average protein has an E₂₈₀ of about 1 for a 1 mg/ml solution). But the method is the quickest possible (pipet it into the cuvette and read) and non-destructive ; it is therefore the best method when you have lots of fractions to read, as from column chromatography, and don't care too much about the exact concentration, only where there's protein and where there isn't.

UV absorbance can also be read in the far ultraviolet, typically at 205 nm, where the peptide bond absorbs and a 1 mg/ml solution has an absorbance of about 30; the variation depends mostly on the tyr + tryp content and thus the extinction coefficient can be corrected for this, $E_{205} = 27.0 + 120(A_{280}/A_{205})$. Some far-UV monitors such as the LKB Uvicord read at 206 nm. The problem here is that nearly everything else, except $\text{SO}_4^{=}$ and ClO_4^- , also absorbs at this wavelength, and you have to have the best available quartz cuvettes, very clean, and a good deuterium lamp to read at this wavelength. Since the sample usually must be diluted to decrease absorbance by the buffer, the increase in sensitivity is not fully realized. Where this method, and one other I'll mention, is most useful is in column or HPLC chromatography of peptides from enzymatic or chemical cleavage, being separated for sequence determination. Such peptides, particularly small ones, may not contain tyrosine or tryptophan, and thus you want to monitor fractions by a method not depending on the amino acid composition. (The biuret method of course is too insensitive.)

The other method for this situation is basic hydrolysis (to free amino acids) followed by **quantitative ninhydrin** reaction. This is quite sensitive and insensitive to amino acid composition except for proline (yellow product rather than purple), but of course it's a lot of trouble for a large number of fractions. Look for it in Methods in Enzymology vol. 11, paper by C.H.W. Hirs.

Another reagent which made a splash for a while is **fluorescamine**, a reagent developed at Hoffman-LaRoche which gives a fluorescent product with primary amines. This can be very sensitive, but of course cannot be used in

amine buffers such as Tris, and in my experience fluorescent impurities are a problem. It and similar reagents such as *o*-phthalaldehyde find their best use as replacements for ninhydrin in measuring free amino acids in **amino acid analysis** (the determination of the amino acid composition of a protein), which is really the most absolute determination of protein concentration - you can add up the amounts of all the amino acids present and calculate the protein concentration - but too expensive and equipment-intensive for routine use. Where amino acid analysis is important is in measuring chemical modification of the protein; it is generally easy to measure the amount of modification, more difficult to measure the molar concentration of the protein exactly in order to determine moles modification/mole protein.

A couple of older methods should be mentioned. **Turbidimetry** - the optical density (not absorbance) of a cloudy suspension of protein denatured with TCA or other acids - used to be used, but hardly ever is since the development of the Lowry method. It is linear over only a fairly narrow range, dependent on very careful control of conditions, and not very sensitive. I probably mention it only because we did it in a lab when I was a graduate student. **Kjeldahl** nitrogen determination - reduction of essentially all nitrogen in a sample to ammonia, which is distilled and measured by titration or the Nessler reaction - is ancient and highly non-specific, but useful under one condition (not likely to be of interest to people in this course): when the sample is highly insoluble, as for protein content of seeds. One could certainly use basic hydrolysis and ninhydrin here.

The article by Stoschak in *Methods in Enzymology* vol. 182 includes a method she devised using colloidal gold. I have no experience with this; it seems to be extremely sensitive but also dependent on careful control of conditions.

To summarize: one normally uses the Lowry or Coomassie Blue method on each defined stage of a purification, though one probably uses UV absorbance on column chromatographic fractions. But it is good to know a lot of possible methods, because you may have reason to use another under some specific condition, such as when you need to know the molar concentration of your protein, or when you have modified amino acids that react in your usual method.