H_4	LDH1
H_3M	LDH2
H_2M_2	LDH3
HM_3	LDH4
M_4	LDH5

Each isoenzyme promotes the same reaction but has different kinetic constants (K_m, V_{max}), thermal stability and electrophoretic mobility. The tissue distribution of isoenzymes within an organism is frequently different, for example, in humans LDH1 is the dominant isoenzyme in heart muscle but LDH5 is the most abundant form in liver and muscle. These differences are exploited in diagnostic enzymology to identify specific organ damage, for example following myocardial infarction, and thereby aiding clinical diagnosis and prognosis.

8.2.1 Post-translational modifications

Proteins are synthesised at the ribosome and as the growing polypeptide chain emerges from the ribosome it folds up into its native three-dimensional structure. However, this is often not the final active form of the protein. Many proteins undergo modifications once they leave the ribosome, where one or more amino acid side chains are modified by the addition of a further chemical group; this is referred to as post-translational modification. Such changes include extensive modifications of the protein structure, for example the addition of chains of carbohydrates to form glycoproteins (see Section 8.4.4), where in some cases the final protein consists of as much as over 40% carbohydrate. Less dramatic, but equally important modifications include the addition of a hydroxyl group to proline to produce hydroxyproline (found in the structure of collagen), or the phosphorylation of one or more amino acids (tyrosine, serine and threonine residues are all capable of being phosphorylated). Many cases are known, for example, where the addition of a single phosphate group (by enzymes known as kinases) can activate a protein molecule, and the subsequent removal of the phosphate group (by a phosphatase) can inactivate the molecule; protein phosphorylation reactions are a central part of intracellular signalling. Another example can be found in the post-translational modification of proline residues in the transcription factor HIF (the α subunit of the hypoxia-inducible factor), which is a key oxygen-sensing mechanism in cells. Many proteins therefore are not in their final active, biological form until post-translational modifications have taken place. Over 200 different post-translational modifications have been reported for proteins from microbial, plant and animal sources. Mass spectrometry is used to determine such modifications (see Section 9.5.5).

8.3 PROTEIN PURIFICATION

8.3.1 Introduction

At first sight, the purification of *one* protein from a cell or tissue homogenate that will typically contain 10 000–20 000 different proteins, seems a daunting task. However,

in practice, on average, only four different fractionation steps are needed to purify a given protein. Indeed, in exceptional circumstances proteins have been purified in a single chromatographic step. Since the reason for purifying a protein is normally to provide material for structural or functional studies, the final degree of purity required depends on the purposes for which the protein will be used, i.e. you may not need a protein sample that is 100% pure for your studies. Indeed, to define what is meant by a 'a pure protein' is not easy. Theoretically, a protein is pure when a sample contains only a single protein species, although in practice it is more or less impossible to achieve 100% purity. Fortunately, many studies on proteins can be carried out on samples that contain as much as 5–10% or more contamination with other proteins. This is an important point, since each purification step necessarily involves loss of some of the protein you are trying to purify. An extra (and unnecessary) purification step that increases the purity of your sample from, say, 90% to 98% may mean that you now have a more pure protein, but insufficient protein for your studies. Better to have studied the sample that was 90% pure and have enough to work on!

For example, a 90% pure protein is sufficient for amino acid sequence determination studies as long as the sequence is analysed quantitatively to ensure that the deduced sequence does not arise from a contaminant protein. Similarly, immunisation of a rodent to provide spleen cells for monoclonal antibody production (Section 7.2.2) can be carried out with a sample that is considerably less than 50% pure. As long as your protein of interest raises an immune response it matters not at all that antibodies are also produced against the contaminating proteins. For kinetic studies on an enzyme, a relatively impure sample can be used provided it does not contain any competing activities. On the other hand, if you are raising a monospecific polyclonal antibody in an animal (see Section 7.2.1), it is necessary to have a highly purified protein as antigen, otherwise immunogenic contaminating proteins will give rise to additional antibodies. Equally, proteins that are to have a therapeutic use must be extremely pure to satisfy regulatory (safety) requirements. Clearly, therefore, the degree of purity required depends on the purpose for which the protein is needed.

8.3.2 The determination of protein concentration

The need to determine protein concentration in solution is a routine requirement during protein purification. The only truly accurate method for determining protein concentration is to acid hydrolyse a portion of the sample and then carry out amino acid analysis on the hydrolysate (see Section 8.4.2). However, this is relatively time-consuming, particularly if multiple samples are to be analysed. Fortunately, in practice, one rarely needs decimal place accuracy and other, quicker methods that give a reasonably accurate assessment of protein concentrations of a solution are acceptable. Most of these (see below) are colorimetric methods, where a portion of the protein solution is reacted with a reagent that produces a coloured product. The amount of this coloured product is then measured spectrophotometrically and the amount of colour related to the amount of protein present by appropriate calibration. However, none of these methods is absolute,

since, as will be seen below, the development of colour is often at least partly dependent on the amino acid composition of the protein(s). The presence of prosthetic groups (e.g. carbohydrate) also influences colorimetric assays. Many workers prepare a standard calibration curve using bovine serum albumin (BSA), chosen because of its low cost, high purity and ready availability. However, it should be understood that, since the amino acid composition of BSA will differ from the composition of the sample being tested, any concentration values deduced from the calibration graph can only be approximate.

Ultraviolet absorption

The aromatic amino acid residues tyrosine and tryptophan in a protein exhibit an absorption maximum at a wavelength of 280 nm. Since the proportions of these aromatic amino acids in proteins vary, so too do extinction coefficients for individual proteins. However, for most proteins the extinction coefficient lies in the range 0.4–1.5; so for a complex mixture of proteins it is a fair approximation to say that a solution with an absorbance at 280 nm (A_{280}) of 1.0, using a 1 cm pathlength, has a protein concentration of approximately 1 mg cm^{-3}. The method is relatively sensitive, being able to measure protein concentrations as low as 10 μg cm^{-3}, and, unlike colorimetric methods, is non-destructive, i.e. having made the measurement, the sample in the cuvette can be recovered and used further. This is particularly useful when one is working with small amounts of protein and cannot afford to waste any. However, the method is subject to interference by the presence of other compounds that absorb at 280 nm. Nucleic acids fall into this category having an absorbance as much as 10 times that of protein at this wavelength. Hence the presence of only a small percentage of nucleic acid can greatly influence the absorbance at this wavelength. However, if the absorbances (A) at 280 and 260 nm wavelengths are measured it is possible to apply a correction factor:

$$\text{Protein (mg cm}^{-3}) = 1.55\ A_{280} - 0.76A_{260}$$

The great advantage of this protein assay is that it is non-destructive and can be measured continuously, for example in chromatographic column effluents.

Even greater sensitivity can be obtained by measuring the absorbance of ultraviolet light by peptide bonds. The peptide bond absorbs strongly in the far ultraviolet, with a maximum at about 190 nm. However, because of the difficulties caused by the absorption by oxygen and the low output of conventional spectro-photometers at this wavelength, measurements are usually made at 205 or 210 nm. Most proteins have an extinction coefficient for a 1 μg cm^{-3} solution of about 30 at 205 nm and about 20 at 210 nm. Clearly therefore measuring at these wavelengths is 20 to 30 times more sensitive than measuring at 280 nm, and protein concentration can be measured to less than 1 μg cm^{-3}. However, one disadvantage of working at these lower wavelengths is that a number of buffers and other buffer components commonly used in protein studies also absorb strongly at this wavelength, so it is not always practical to work at this lower wavelength.

Nowadays all purpose-built column chromatography systems (e.g. fast protein liquid chromatography and high-performance liquid chromatography (HPLC)) have

in-line variable wavelength ultraviolet light detectors that monitor protein elution from columns.

Lowry (Folin–Ciocalteau) method

In the past this has been the most commonly used method for determining protein concentration, although it is tending to be replaced by the more sensitive methods described below. The Lowry method is reasonably sensitive, detecting down to 10 μg cm^{-3} of protein, and the sensitivity is moderately constant from one protein to another. When the Folin reagent (a mixture of sodium tungstate, molybdate and phosphate), together with a copper sulphate solution, is mixed with a protein solution, a blue-purple colour is produced which can be quantified by its absorbance at 660 nm. As with most colorimetric assays, care must be taken that other compounds that interfere with the assay are not present. For the Lowry method this includes Tris, zwitterionic buffers such as Pipes and Hepes, and EDTA. The method is based on both the Biuret reaction, where the peptide bonds of proteins react with Cu^{2+} under alkaline conditions producing Cu^{+}, which reacts with the Folin reagent, and the Folin–Ciocalteau reaction, which is poorly understood but essentially involves the reduction of phosphomolybdotungstate to hetero-polymolybdenum blue by the copper-catalysed oxidation of aromatic amino acids. The resultant strong blue colour is therefore partly dependent on the tyrosine and tryptophan content of the protein sample.

The bicinchoninic acid method

This method is similar to the Lowry method in that it also depends on the conversion of Cu^{2+} to Cu^{+} under alkaline conditions. The Cu^{+} is then detected by reaction with bicinchoninic acid (BCA) to give an intense purple colour with an absorbance maximum at 562 nm. The method is more sensitive than the Lowry method, being able to detect down to 0.5 μg protein cm^{-3}, but perhaps more importantly it is generally more tolerant of the presence of compounds that interfere with the Lowry assay, hence the increasing popularity of the method.

The Bradford method

This method relies on the binding of the dye Coomassie Brilliant Blue to protein. At low pH the free dye has absorption maxima at 470 and 650 nm, but when bound to protein has an absorption maximum at 595 nm. The practical advantages of the method are that the reagent is simple to prepare and that the colour develops rapidly and is stable. Although it is sensitive down to 20 μg protein cm^{-3}, it is only a relative method, as the amount of dye binding appears to vary with the content of the basic amino acids arginine and lysine in the protein. This makes the choice of a standard difficult. In addition, many proteins will not dissolve properly in the acidic reaction medium.

Kjeldahl analysis

This is a general chemical method for determining the nitrogen content of any compound. It is not normally used for the analysis of purified proteins or for monitoring column fractions but is frequently used for analysing complex solid samples and microbiological samples for protein content. The sample is digested by boiling

Example 1 **PROTEIN ASSAY**

Question A series of dilutions of bovine serum albumin (BSA) was prepared and 0.1 cm^3 of each solution subjected to a Bradford assay. The increase in absorbance at 595 nm relative to an appropriate blank was determined in each case, and the results are shown in the table.

Concentration of BSA ($mg\,cm^{-3}$)	A_{595}
1.5	1.40
1.0	0.97
0.8	0.79
0.6	0.59
0.4	0.37
0.2	0.17

A sample (0.1 cm^3) of a protein extract from *E. coli* gave an A_{595} of 0.84 in the same assay. What was the concentration of protein in the *E. coli* extract?

Answer If a graph of BSA concentration against A_{595} is plotted it is seen to be linear. From the graph, at an A_{595} of 0.84 it can be seen that the protein concentration of the *E. coli* extracted is 0.85 $mg\,cm^{-3}$.

with concentrated sulphuric acid in the presence of sodium sulphate (to raise the boiling point) and a copper and/or selenium catalyst. The digestion converts all the organic nitrogen to ammonia, which is trapped as ammonium sulphate. Completion of the digestion stage is generally recognised by the formation of a clear solution. The ammonia is released by the addition of excess sodium hydroxide and removed by steam distillation in a Markham still. It is collected in boric acid and titrated with standard hydrochloric acid using methyl red–methylene blue as indicator. It is possible to carry out the analysis automatically in an autokjeldahl apparatus. Alternatively, a selective ammonium ion electrode may be used to directly determine the content of ammonium ion in the digest. Although Kjeldahl analysis is a precise and reproducible method for the determination of nitrogen, the determination of the protein content of the original sample is complicated by the variation of the nitrogen content of individual proteins and by the presence of nitrogen in contaminants such as DNA. In practice, the nitrogen content of proteins is generally assumed to be 16% by weight.

8.3.3 Cell disruption and production of initial crude extract

The initial step of any purification procedure must, of course, be to disrupt the starting tissue to release proteins from within the cell. The means of disrupting the tissue will depend on the cell type (see Cell disruption, below), but thought must first be given to the composition of the buffer used to extract the proteins.