Lodish H, Berk A, Zipursky SL, et al. Molecular Cell Biology. 4th edition. New York: W. H. Freeman; 2000.

Section 3.5 Purifying, Detecting, and Characterizing Proteins

A protein must be purified before its structure and the mechanism of its action can be studied. However, because proteins vary in size, charge, and water solubility, no single method can be used to isolate all proteins. To isolate one particular protein from the estimated 10,000 different proteins in a cell is a daunting task that requires methods both for separating proteins and for detecting the presence of specific proteins.

Any molecule, whether protein, carbohydrate, or nucleic acid, can be separated from other molecules based on large differences in some physical characteristic. Although the sequence of amino acids in a protein uniquely determines its function, the most useful physical characteristic for separation of proteins is *size,* defined as either length or mass. In this section, we briefly outline different techniques for separating proteins based on their size and other properties. These techniques also apply to the separation of nucleic acids and other biomolecules. We then consider general methods for detecting, or *assaying,* specific proteins, including the use of radioactive compounds for tracking biological activity. Finally, we discuss several techniques for characterizing a protein's mass, sequence, and three-dimensional structure.

Proteins Can Be Removed from Membranes by Detergents or High-Salt Solutions

Because water-soluble globular proteins have many exposed hydrophilic groups, they maintain their native conformation and remain individually suspended in an aqueous medium when separated from cells. In contrast, when transmembrane proteins are separated from membranes, their exposed hydrophobic regions interact, causing the protein molecules to aggregate and precipitate from aqueous solutions. Such proteins can be solubilized by detergents, which have affinity both for hydrophobic groups and for water.

Detergents are amphipathic molecules that disrupt membranes by intercalating into phospholipid bilayers and solubilizing lipids and proteins. The hydrophobic part of a detergent molecule is attracted to hydrocarbons and mingles with them readily; the hydrophilic part is strongly attracted to water. Some detergents are natural products, but most are synthetic molecules developed for cleaning and for dispersing mixtures of oil and water (Figure 3-38). Ionic detergents, such as sodium deoxycholate and sodium dodecylsulfate (SDS), contain a charged group; nonionic detergents, such as Triton X-100 and octylglucoside, lack a charged group.

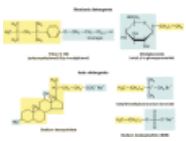


Figure 3-38

Structures of five common detergents. The bile salt sodium deoxycholate is a natural product; the others are synthetic ones. The hydrophobic portion of each molecule is shown in yellow; the hydrophilic portion, in blue.

At very low concentrations, detergents dissolve in pure water as isolated molecules. As the concentration increases, the molecules begin to form *micelles.* These are small, spherical aggregates in which hydrophilic parts of the molecules face outward and the hydrophobic parts cluster in the center (see Figure 2-20). The *critical micelle concentration (CMC)* at which micelles form is characteristic of each detergent and is a function of the structures of its hydrophobic and hydrophilic parts.

Ionic detergents bind to the exposed hydrophobic regions of membrane proteins as well as to the hydrophobic core of water-soluble proteins. Because of their charge, these detergents also disrupt ionic and hydrogen bonds. At high concentrations, for example, sodium dodecylsulfate completely denatures proteins by binding to every side chain. *Nonionic detergents* act in different

ways at different concentrations. At high concentrations (above the CMC), they solubilize biological membranes by forming mixed micelles of detergent, phospholipid, and integral membrane proteins (Figure 3-39). At low concentrations (below the CMC), these detergents may bind to the hydrophobic regions of most membrane proteins, making them soluble in aqueous solution. In this case, although mixed micelles are not formed, the solubilized protein will not aggregate during subsequent purification steps.

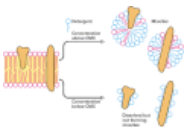


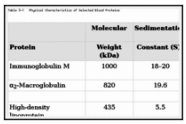
Figure 3-39

Solubilization of integral membrane proteins by nonionic detergents. At a concentration higher than its critical micelle concentration (CMC), a detergent solubilizes lipids and integral membrane proteins, forming mixed micelles containing detergent, protein, (more...)

Most peripheral proteins are bound to specific integral membrane proteins by ionic or other weak interactions. Generally they can be removed from the membrane by solutions of high ionic strength (high salt concentrations), which disrupt ionic bonds, or by chemicals that bind divalent cations such as Mg^{2+}. Unlike integral proteins most peripheral proteins are soluble in aqueous solution and are not solubilized by nonionic detergents.

Centrifugation Can Separate Particles and Molecules That Differ in Mass or Density

The first step in a typical protein-purification scheme is centrifugation. The principle behind centrifugation is that two particles in suspension (cells, organelles, or molecules), having different masses or densities will settle to the bottom of a tube at different rates. Remember, *mass* is the weight of a sample (measured in grams), whereas *density* is the ratio of its weight to volume (grams/liter). Proteins vary greatly in mass but not in density. The average density of a protein is 1.37 g/cm^3. Unless a protein has an attached lipid or carbohydrate, its density will not vary by more than 15 percent from this value. Table 3-1 lists the density and other physical characteristics of several blood proteins. Heavier or more dense molecules settle, or sediment, more quickly than lighter or less dense molecules.



Protein	Molecular Weight (kDa)	Sedimentation Constant (S)
Immunoglobulin M	1000	18-20
α_2-Macroglobulin	820	19.6
High-density lipoprotein	435	5.5

Table 3-1

Physical Characteristics of Selected Blood Proteins.

A centrifuge speeds sedimentation by subjecting particles in suspension to centrifugal forces as great as 600,000 times the force of gravity *g*. The centrifugal force is proportional to the rotation rate of the rotor (measured in revolutions per minute, or rpm) and the distance of the tube from the center of the rotor. Modern ultracentrifuges reach speeds of 60,000 rpm or greater and generate forces sufficient to sediment particles with masses greater than 10,000 daltons (Da). However, small particles with masses of 5 Da or less will not sediment uniformly even at such high rotor speeds.

Centrifugation is used for two basic purposes: (1) as a preparative technique to separate one type of material from others and (2) as an analytical technique to measure physical properties (e.g., molecular weight, density, shape, and equilibrium binding constants) of macromolecules. The *sedimentation constant, s,* of a protein equals its velocity in a centrifugal field divided by the centrifugal force. The value of *s* depends on a protein's density and shape, as well as the density and viscosity of the medium. Because the centrifugal force and the density and viscosity of the medium are all known, the radius and mass of a molecule can be calculated from measurements of its rate of movement in an analytical ultracentrifuge. The sedimentation constant is commonly expressed in svedbergs (S): 1 S = 10^{-13} seconds.

Differential Centrifugation

The most common initial step in protein purification is separation of soluble proteins from insoluble cellular material by *differential centrifugation* (Figure 3-40a). The centrifugal force and duration of centrifugation are adjusted to ensure that the insoluble materials sediment into a pellet. After a starting mixture of a cell homogenate is poured into a tube and spun in a centrifuge, cell organelles such as nuclei collect into a pellet, but the soluble proteins remain in the supernatant. The supernatant fraction still contains a large mixture of proteins, which can be collected by decanting the supernatant and then subjecting it to further purification methods.

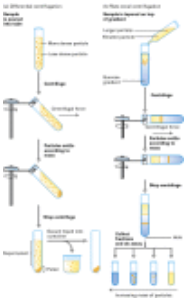


Figure 3-40

Two common centrifugation techniques for separating particles. (a) Differential centrifugation separates a mixture of particles (macromolecules, cell organ-elles, and cells) that differ in mass or density. The most dense particles collect at the bottom (more...)

Rate-Zonal Centrifugation

Based on differences in their mass, proteins can be separated by centrifugation through a solution, usually containing sucrose (an inert sugar), of increasing density called a *density gradient.* When mixtures of proteins are layered on top of a sucrose gradient in a tube and subjected to centrifugation, they migrate down the tube at a rate controlled by the factors that affect the sedimentation constant. The proteins start from a thin zone at the top of the tube and separate into bands, or zones (actually disks), of proteins of different masses. This density-gradient separation technique is called *rate-zonal centrifugation* (Figure 3-40b). Samples are centrifuged just long enough to separate the molecules of interest. If they are centrifuged for too short a time, the molecules will not separate sufficiently. If they are centrifuged much longer than necessary, all the molecules will end up in a pellet at the bottom of the tube.

Although the sedimentation rate is strongly influenced by particle mass, rate-zonal centrifugation is seldom effective in determining *precise* molecular weights because variations in shape also affect sedimentation rate. The exact effects of shape are hard to assess, especially for proteins and single-stranded nucleic acid molecules that can assume many complex shapes. Nevertheless, rate-zonal centrifugation has proved to be the most practical method for separating many different types of polymers and particles. A second density-gradient technique, called *equilibrium density-gradient centrifugation,* is used mainly to separate DNA or organelles (see Figure 5-24).

Electrophoresis Separates Molecules according to Their Charge:Mass Ratio

Electrophoresis is a technique for separating, or *resolving,* molecules in a mixture under the influence of an applied electric field. Dissolved molecules in an electric field move, or migrate, at a speed determined by their charge:mass ratio. For example, if two molecules have the same mass and shape, the one with the greater net charge will move faster toward an electrode. The separation of small molecules, such as amino acids and nucleotides, is one of the many uses of electrophoresis. In this case, a small drop of sample is deposited on a strip of filter paper or other porous substrate, which is then soaked with a conducting solution. When an electric field is applied at the ends of the strip, small molecules dissolved in the conducting solution move along the strip at a rate corresponding to the magnitude of their charge.

SDS-Polyacrylamide Gel Electrophoresis

Because many proteins or nucleic acids that differ in size and shape have nearly identical charge:mass ratios, electrophoresis of these macromolecules in solution results in little or no

separation of molecules of different lengths. However, successful separation of proteins and nucleic acids can be accomplished by electrophoresis in various *gels* (semisolid suspensions in water) rather than in a liquid solution. Electrophoretic separation of proteins is most commonly performed in *polyacrylamide gels.* These gels are cast between a pair of glass plates by polymerizing a solution of acrylamide monomers into polyacrylamide chains and simultaneously cross-linking the chains into a semisolid matrix. The *pore size* of a gel can be varied by adjusting the concentrations of polyacrylamide and the cross-linking reagent.

When a mixture of proteins is applied to a gel and an electric current applied, smaller proteins migrate faster than larger proteins through the gel. The rate of movement is influenced by the gel's pore size and the strength of the electric field. The pores in a highly cross-linked polyacrylamide gel are quite small. Such a gel could resolve small proteins and peptides, but large proteins would not be able to move through it.

In what is probably the most powerful technique for resolving protein mixtures, proteins are exposed to the ionic detergent SDS (sodium dodecylsulfate) before and during gel electrophoresis (Figure 3-41). SDS denatures proteins, causing multimeric proteins to dissociate into their subunits, and all polypeptide chains are forced into extended conformations with similar charge:mass ratios. SDS treatment thus eliminates the effect of differences in shape, so that chain length, which reflects mass, is the sole determinant of the migration rate of proteins in SDS-polyacrylamide electrophoresis. Even chains that differ in molecular weight by less than 10 percent can be separated by this technique. Moreover, the molecular weight of a protein can be estimated by comparing the distance it migrates through a gel with the distances that proteins of known molecular weight migrate.

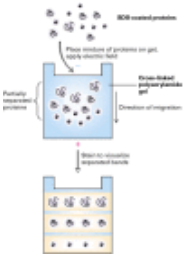


Figure 3-41

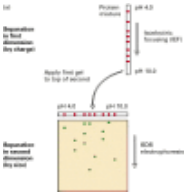
SDS-polyacrylamide gel electrophoresis, a common technique for separating proteins at good resolution. The protein mixture first is treated with SDS, a negatively charged detergent that binds to proteins. This binding dissociates multimeric proteins and (more...)

Two-Dimensional Gel Electrophoresis

Electrophoresis of all cellular proteins through an SDS gel can separate proteins having relatively large differences in molecular weight but cannot resolve proteins having similar molecular weights (e.g., a 41-kDa protein from a 42-kDa protein). To separate proteins of similar mass, another physical characteristic must be exploited. Most commonly, this is electric charge, which is determined by the number of acidic and basic residues in a protein. Two unrelated proteins having similar masses are unlikely to have identical net charges because their sequences, and thus the number of acid and basic residues, are different.

In two-dimensional electrophoresis, proteins are separated in two sequential steps: first by their charge and then by their mass. In the first step, a cell extract is fully denatured by high concentrations (8 M) of urea and then layered on a glass tube filled with polyacrylamide that is saturated with a solution of *ampholytes,* a mixture of polyanionic and polycationic molecules. When placed in an electric field, the ampholytes will separate and form a continuous gradient based on their net charge (Figure 3-42a). The most highly polyanionic ampholytes will collect at one end of the tube, and the most polycationic ampholytes will collect at the other end. This gradient of ampholytes establishes a pH gradient. Charged proteins will migrate through the gradient until they reach their pI, or isoelectric point, the pH at which the net charge of the protein is zero. This technique, called isoelectric focusing (IEF), can resolve proteins that differ by only one charge unit.

Figure 3-42



Two-dimensional gel electrophoresis, a technique for separatir on their charge and their mass. (a) Preparation of a two-dimen by isoelectric focusing (IEF) followed by SDS electrophoresis dimensional gel of protein (more...)

Proteins that have been separated on an IEF gel can then be separated in a second dimension based on their molecular weights. To accomplish this, the IEF gel is extruded from the tube and placed lengthwise on a second polyacrylamide gel, this time formed as a slab saturated with SDS. When an electric field is imposed, the proteins will migrate from the IEF gel into the SDS slab gel and then separate according to their mass. The sequential resolution of proteins by their charge and mass can achieve excellent separation of cellular proteins (Figure 3-42b). For example, two-dimensional gels have been very useful in studying the expression of various genes in differentiated cells because as many as 1000 proteins can be resolved simultaneously.

Liquid Chromatography Resolves Proteins by Mass, Charge, or Binding Affinity

Liquid chromatography, a third commonly used technique to separate mixtures of proteins, nucleic acids, and other molecules, is based on the principle that molecules dissolved in a solution will interact (bind and dissociate) with a solid surface. If the solution is allowed to flow across the surface, then molecules that interact frequently with the surface will spend more time bound to the surface and thus move more slowly than molecules that interact infrequently with the surface. Liquid chromatography is performed in a column packed tightly with spherical beads. The nature of these beads determines whether separation of proteins depends on differences in mass, charge, or binding affinity.

Gel Filtration Chromatography

Proteins that differ in mass can be separated by gel filtration chromatography. In this technique, the column is composed of porous beads made from polyacrylamide, dextran (a bacterial polysaccharide), or agarose (a seaweed derivative). Proteins *flow around* the spherical beads in gel filtration chromatography. However, the surface of the beads is punctured by large holes, and proteins will spend some time within these holes. Because smaller proteins can penetrate into the beads more easily than larger proteins, they travel through a gel filtration column more slowly than larger proteins (Figure 3-43a). (In contrast, proteins migrate *through* the pores in an electrophoretic gel; thus smaller proteins move faster than larger ones.) The total volume of liquid required to elute a protein from the column depends on its mass: the smaller the mass, the greater the elution volume. By use of proteins of known mass, the elution volume can be used to estimate the mass of a protein in a mixture.

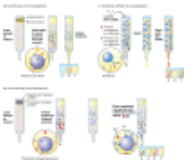


Figure 3-43

Three commonly used liquid chromatographic techniques. (a) Gel filtration chromatography separates proteins that differ in size. A mixture of proteins is carefully layered on the top of a glass cylinder packed with porous beads. Smaller proteins travel (more...)

Ion-Exchange Chromatography

In a second type of liquid chromatography, called *ion-exchange chromatography,* proteins are separated based on differences in their charge. This technique makes use of specially modified beads whose surfaces are covered by amino groups or carboxyl groups and thus carry either a positive charge (NH_3^+) or a negative charge (COO^-) at neutral pH.

The proteins in a mixture carry various net charges at any given pH. When a solution of a protein mixture flows through a column of positively charged beads, only proteins with a net negative

charge (acidic proteins) adhere to the beads; neutral and basic proteins flow unimpeded through the column (Figure 3-43b). The acidic proteins are then eluted selectively by passing a gradient of increasing concentrations of salt through the column. At low salt concentrations, protein molecules and beads are attracted by their opposite charges. At higher salt concentrations, negative salt ions bind to the positively charged beads, displacing the negatively charged proteins. In a gradient of increasing salt concentration, weakly charged proteins are eluted first and highly charged proteins are eluted last. Similarly, a negatively charged column can be used to retain and fractionate positively charged (basic) proteins.

Affinity Chromatography

A third form of chromatography, called *affinity chromatography,* relies on the ability of a protein to bind specifically to another molecule. Columns are packed with beads to which are covalently attached ligand molecules that bind to the protein of interest. Ligands can be enzyme substrates or other small molecules that bind to specific proteins. In a widely used form of this technique, *antibody-affinity chromatography,* the attached ligand is an antibody specific for the desired protein (Figure 3-43c). An affinity column will retain only the proteins that bind the ligand attached to the beads; the remaining proteins, regardless of their charge or mass, will pass through the column without binding to it. The proteins bound to the affinity column then are eluted by adding an excess of ligand or by changing the salt concentration or pH. Obviously, the ability of this technique to separate particular proteins depends on the selection of appropriate ligands.

Highly Specific Enzyme and Antibody Assays Can Detect Individual Proteins

Purification of a protein, or any other molecule, requires a specific assay that can detect the molecule of interest in column fractions or gel bands. An assay capitalizes on some highly distinctive characteristic of a protein: the ability to bind a particular ligand, to catalyze a particular reaction, or to be recognized by a specific antibody. An assay must also be simple and fast in order to minimize errors and the possibility that the protein of interest is denatured or degraded while the assay is performed. The goal of any purification scheme is to isolate sufficient amounts of a given protein for study; thus a useful assay must also be sensitive enough that only a small proportion of the available material is consumed. Many common protein assays require just 10^{-9} to 10^{-12} g of material.

Chromogenic and Light-Emitting Enzyme Reactions

Many assays are tailored to detect some functional aspect of a protein. For example, enzyme assays are based on the ability to detect the loss of substrate or the formation of product. Many enzyme assays utilize *chromogenic* substrates, which change color during the course of the reaction. (Some substrates are naturally chromogenic; if they are not, they can be linked to a chromogenic molecule.) Because of the specificity of an enzyme for its substrate, only samples that contain the enzyme will change color in the presence of a chromogenic substrate and other required reaction components; the rate of the reaction provides a measure of the quantity of enzyme present.

Such chromogenic enzymes also can be fused or chemically linked to an antibody and used to "report" the presence or location of the antigen. Alternatively *luciferase,* an enzyme present in fireflies and some bacteria, can be linked to an antibody. In the presence of ATP and luciferin, this enzyme catalyzes a light-emitting reaction. In either case, after the antibody binds to the protein of interest, substrates of the linked enzyme are added and the appearance of color or emitted light is monitored.

Western Blotting

One of the most powerful methods for detecting a particular protein in a complex mixture combines the superior resolving power of gel electrophoresis, the specificity of antibodies, and

the sensitivity of enzyme assays. Called Western blotting, or immunoblotting, this three-step procedure is commonly used to separate proteins and then identify a specific protein of interest. As shown in Figure 3-44, two different antibodies are used in this method, one specific for the desired protein and the other linked to a reporter enzyme.

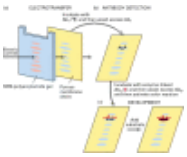


Figure 3-44

Western blotting, or immunoblotting. (a) A protein mixture is electrophoresed through an SDS gel, and then transferred from the gel onto a membrane. (b) The membrane is flooded with a solution of antibody (Ab_1) specific for the desired protein. Only the (more...)

Radioisotopes Are Indispensable Tools for Detecting Biological Molecules

Since World War II, when radioactive materials first became widely available as byproducts of work in nuclear physics, chemists and biologists have fashioned an almost limitless variety of radioactive chemicals. Today, radioactively labeled precursors of macromolecules greatly simplify many standard biochemical assays and significantly enhance the ability of researchers to follow biochemical events in whole cells as well as in cell extracts. Almost all experimental biology depends on the use of radioactive compounds.

At least one atom in a radiolabeled molecule is present in a radioactive form, called a radioisotope (Table 3-2). The presence of a radioisotope does not change the chemical properties of a molecule. For example, enzymes, both in vivo and in vitro, catalyze reactions involving labeled substrates just as readily as those involving nonlabeled substrates. Because radioisotopes emit easily detected particles, the fate of radiolabeled molecules can be traced in cells and cellular extracts.

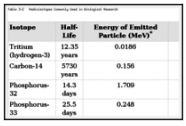


Table 3-2

Radioisotopes Commonly Used in Biological Research.

Characteristics of Different Radiolabels

The choice of which labeled compound to use in a particular experiment involves several considerations. Some labeled compounds, for instance, are not suitable for studies with whole cells because they do not enter cells. One prominent example is ATP, as well as most other phosphorylated compounds (e.g., glucose 6-phosphate). Although ^{32}P-labeled ATP can contribute phosphorus-32 during RNA and DNA synthesis in a cell-free system, it cannot do so with whole cells, because it never gets into the cells. On the other hand, labeled orthophosphate ($^{32}PO_4^{3-}$) in the medium does enter both bacterial cells and animal cells, and then is incorporated into phosphorylated proteins, nucleotides, and eventually into cellular RNA and DNA.

Hundreds of biological compounds (e.g., amino acids, nucleosides, and numerous metabolic intermediates) labeled with various radioisotopes are commercially available. These preparations vary considerably in their *specific activity,* which is the amount of radioactivity per unit of material, measured in disintegrations per minute (dpm) per millimole. The specific activity of a labeled compound depends on the ratio of unstable potentially radioactive atoms to stable nonradioactive atoms. It also depends on the probability of decay of the radioisotope, indicated by its *half-life,* which is the time required for half the atoms to undergo radioactive decay. In general, the shorter the half-life of a radioisotope, the higher its specific activity (see Table 3-2).

The specific activity of a labeled compound must be high enough so that sufficient radioactivity is incorporated into cellular molecules to be accurately detected. For example, methionine and cysteine labeled with sulfur-35 (^{35}S) are widely used to label cellular proteins because

preparations of these amino acids with high specific activities (>10^{15} dpm/mmol) are available. Likewise, commercial preparations of ^{3}H-labeled nucleic acid precursors have much higher specific activities than the corresponding ^{14}C-labeled preparations. In most experiments, the former are preferable because they allow RNA or DNA to be adequately labeled after a shorter time of incorporation or require a smaller cell sample.

Various phosphate-containing compounds in which every phosphorus atom is the radioisotope phosphorus-32 are readily available. Because of their high specific activity, ^{32}P-labeled nucleotides are routinely used to label nucleic acids in cell-free systems. The radioisotope iodine-125 (^{125}I), which also is available in almost pure form, can be covalently linked to a protein or nucleic acid to yield preparations with a high specific activity. Such attachment of iodine-125 can be achieved enzymatically or chemically and generally does not drastically alter the properties of a macromolecule.

Labeling Experiments and Detection of Radiolabeled Molecules

Depending on the nature of an experiment, labeled compounds are detected by autoradiography, a semiquantitative visual assay, or their radioactivity is measured in an appropriate "counter," a highly quantitative assay that can determine the concentration of a radiolabeled compound in a sample. In some experiments, both types of detection are used.

In autoradiography, a cell or cell constituent is labeled with a radioactive compound and then overlaid with a photographic emulsion sensitive to radiation. Development of the emulsion yields small silver grains whose distribution corresponds to that of the radioactive material (Figure 3-45a). Autoradiographic studies of whole cells have been crucial in determining the intracellular sites where various macromolecules are synthesized and their subsequent movements within cells. For example, when cells are incubated for a short time with [^{3}H]thymidine, a unique DNA precursor, most of the radioactivity is localized to the nucleus, identifying the nucleus as the major site of DNA synthesis (Figure 3-45b). Even after cells are incubated for prolonged periods with [^{3}H]thymidine, virtually all the radioactivity remains in the nucleus, indicating that the DNA remains there. Similarly, the site of synthesis of RNA is revealed by incubating cells for 1 minute with [^{3}H]uridine, a unique RNA precursor; in this case, all the autoradiographic grains are found over the nucleus. After a longer period of incorporation, however, many autoradiographic grains are located over the cytoplasm, indicating that RNA, in contrast to DNA, is transported from the nucleus.

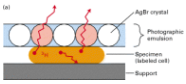


Figure 3-45

Autoradiography. (a) A radiation-sensitive photographic emulsion containing silver salts (AgBr) is placed over tritium-labeled cells attached to a glass slide (for the light microscope) or to a carbon-coated grid (for the electron microscope). The cell (more...)

In autoradiographic studies, the ability of the experimenter to localize the site at which the radioisotope is incorporated is affected by the energy of the particles emitted during radioactive disintegrations. For example, the β particles emitted by phosphorus-32 are so energetic that the streaks they make on a photographic emulsion can be as long as 1 mm, much longer than the diameter of individual cells. In contrast, the β particles emitted by tritium create tracks on a photographic emulsion that are only about 0.47 μm long; thus ^{3}H-labeled structures can be located within cells to an accuracy of about 0.5–1.0 mm, or about one-fifth the diameter of the nucleus of mammalian cells. Because tritium emits the least-energetic particles of all the common radioisotopes, it is highly preferred for locating labeled compounds or structures within cells.

Quantitative measurements of the amount of radioactivity in a labeled material are performed with several different instruments. A *Geiger counter* measures ions produced in a gas by the β

particles or γ rays emitted from a radioisotope. In a *scintillation counter,* a radiolabeled sample is mixed with a liquid containing a fluorescent compound that emits a flash of light when it absorbs the energy of the β particles or γ rays released during decay of the radioisotope; a phototube in the instrument detects and counts these light flashes. *Phosphorimagers* are used to detect radiolabeled compounds on a surface, storing digital data on the number of decays in dpm per small pixel of surface area. These instruments commonly are used to quantitate radioactive molecules separated by gel electrophoresis and are replacing photographic film for this purpose.

The usual experimental protocol for determining the cellular location of a particular molecule has three steps:

1. A radioactive precursor is incubated with whole cells or cell-free extracts.
2. The cellular constituents then are isolated and purified in various ways.
3. The radioactivity of the various fractions is measured with a counter.

For example, to identify the site of RNA synthesis, cells can be incubated for a short period with [^{3}H]uridine and then subjected to a fractionation procedure to separate the various organelles (Chapter 5). The specific activity of the nuclear fraction (dpm/mg protein) is found to be much higher than that of any other organelle fraction, thus confirming the nucleus as the site of RNA synthesis.

A combination of labeling and biochemical techniques and of visual and quantitative detection methods is often employed in labeling experiments. For instance, to identify the major proteins synthesized by a particular cell type, a sample of the cells is incubated with a radioactive amino acid (e.g., ^{35}S-labeled methionine) for a few minutes. The mixture of cellular proteins then is resolved by gel electrophoresis, and the gel is subjected to autoradiography or phosphorimager analysis. The radioactive bands correspond to newly synthesized proteins, which have incorporated the radiolabeled amino acid. Alternatively, the proteins can be resolved by liquid chromatography, and the radioactivity in the eluted fractions determined quantitatively with a counter.

Researchers often use the pulse-chase technique in labeling experiments. In this protocol, a cell sample is exposed to a radiolabeled compound for only a brief period of time (the “pulse”), then washed with buffer to remove the label, and finally incubated with a nonlabeled form of the compound (the “chase”). Pulse-chase experiments are particularly useful for tracing changes in the intracellular location of proteins or the transformation of a metabolite into others over time.

Protein Primary Structure Can Be Determined by Chemical Methods and from Gene Sequences

The primary structure of a protein is characterized in two ways: by its overall amino acid composition and by its precise amino acid sequence. The amino acid composition of a protein gives the same information as an elemental analysis of a molecule—the types of amino acids present and their abundance but not their linear order. In contrast, the sequence of a protein is like a fingerprint; it uniquely establishes the identity of a protein—the linear order of amino acids. The composition of a protein is easily calculated from the sequence. Two proteins can differ in their sequence but nonetheless have identical amino acid compositions. Composition and sequence are determined by chemical methods based on the ability to cleave the peptide backbone at the peptide bond.

The classic method for determining the amino acid sequence of a protein involves *Edman degradation* (Figure 3-46). In this procedure the amino group at the **N**-terminus of a polypeptide is labeled and its amino acid then cleaved from the polypeptide and identified by high-pressure liquid chromatography. The polypeptide is left one residue shorter, with a new amino acid at the **N**-terminus. The cycle is repeated on the ever shortening polypeptide until all the residues have been identified.

Figure 3-46

Chemical determination of the sequence of a protein by Edma
which involves a repetitive three-step procedure. In the first st
polypeptide Nterminus is reacted with phenylisothiocyanate (I
second step, the N-terminal amino (more...)

Before about 1985, biologists commonly used the Edman chemical procedure for determining protein sequences. However, recombinant DNA techniques developed in the 1970s and 1980s permit the detection and cloning of the mRNA or gene encoding a specific protein. From the sequence of its mRNA or gene, the protein's amino acid sequence can be deduced. Because sequencing of mRNA and DNA generally is faster than chemical sequencing of proteins, this approach is now the most popular way to determine protein sequences, especially of large proteins. As we describe in Chapter 7, the complete genome sequences of several organisms have already been determined, and the database of genome sequences from humans and numerous model organisms is expanding rapidly.

Time-of-Flight Mass Spectrometry Measures the Mass of Proteins and Peptides

A powerful technique for measuring the mass of molecules like proteins and peptides is *mass spectrometry* (Figure 3-47). Mass spectrometry requires a method for ionizing the sample, usually a mixture of peptides or proteins, accelerating the molecular ions, and then detecting the ions. In a laser desorption mass spectrometer, the proteins are mixed with an organic acid and then dried on a metal target. Light from a laser ionizes the proteins, which "fly" down a tube to a detector. Their time of flight is inversely proportional to their mass and directly proportional to the charge on the protein. As little as 1 femtomole of proteins as large as 200,000 MW can be measured with an error of 0.1 percent.

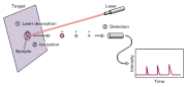


Figure 3-47

Mass measurements by time-of-flight mass spectrometry. Pulses of light from a laser ionize a protein or peptide mixture that is absorbed on a metal target. An electric field accelerates the molecules in the sample toward the detector. The time to the (more...)

One powerful use of mass spectrometers is to identify a protein from its *peptide mass fingerprint.* A peptide mass fingerprint is a compilation of the molecular weights of peptides that are generated by a specific protease. The molecular weights of the parent protein and its proteolytic fragments are used to search genome databases for any similarly sized protein with identical or similar peptide mass maps. With the increasing availability of genome sequences, this approach has almost eliminated the need to chemically sequence a protein to determine its primary structure.

Peptides with a Defined Sequence Can Be Synthesized Chemically

Synthetic peptides that are identical with peptides synthesized in vivo are useful experimental tools in studies of proteins and cells. For example, short synthetic peptides of 10–15 residues can function as antigens to trigger production of antibodies in animals. A synthetic peptide can trick the animal into producing antibodies that bind the full-sized, natural protein antigen. As we'll see throughout this book, antibodies are extremely versatile reagents for isolating proteins from mixtures by affinity chromatography (see Figure 3-43c), separating and detecting proteins by Western blotting (see Figure 3-44), and localizing proteins in cells by microscopic techniques described in Chapter 5. Synthetic peptides also have been helpful in elucidating the rules that

determine the secondary and tertiary structure of proteins. By systematically varying the sequence of synthetic peptides, researchers have studied the influence of various amino acids on protein conformation.

Peptides are routinely synthesized in a test tube from monomeric amino acids by condensation reactions that form peptide bonds. Peptides are constructed sequentially by coupling the C-terminus of a monomeric amino acid with the N-terminus of the growing peptide, as outlined in Figure 3-48. To prevent unwanted reactions involving the amino groups and carboxyl groups of the side chains during the coupling steps, a protecting (blocking) group is attached to the side chains. Without these protecting groups, branched peptides would be generated. In the last steps of synthesis, the side chain–protecting groups are removed and the peptide is cleaved from the resin on which synthesis occurs.

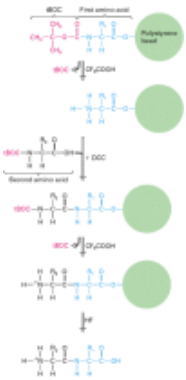


Figure 3-48

Solid-phase peptide synthesis. The first amino acid (blue) of the desired peptide is attached at its carboxyl end by esterification to a polystyrene bead. The amino group of this amino acid is blocked by the attachment of a tertbutyloxycarbonyl (tBOC) (more...)

Protein Conformation Is Determined by Sophisticated Physical Methods

In this chapter we have emphasized that protein function is derived from protein structure. Thus, to figure out how a protein works, its three-dimensional structure must be known. Determining a protein's conformation requires sophisticated physical methods and complex analyses of the experimental data. We briefly describe three methods used to generate three-dimensional models of proteins.

X-Ray Crystallography

The use of x-ray crystallography to determine the three-dimensional structures of proteins was pioneered by Max Perutz and John Kendrew in the 1950s. To date, the detailed three-dimensional structures of more than 8000 proteins have been established by this technique in which beams of x-rays are passed through a crystal of protein. The wavelengths of x-rays are about 0.1–0.2 nanometer (nm), short enough to resolve the atoms in the protein crystal. Atoms in the protein crystal scatter the x-rays, which produce a diffraction pattern of discrete spots when they are intercepted by photographic film (Figure 3-49). Such patterns are extremely complex; as many as 25,000 diffraction spots can be obtained from a small protein. Elaborate calculations and modifications of the protein (such as binding of heavy metals) must be made to interpret the diffraction pattern and to solve the structure of the protein. The process is analogous to reconstructing the precise shape of a rock from the ripples it creates in a pond.

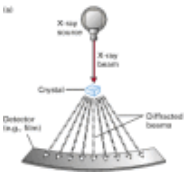


Figure 3-49

X-ray diffraction. (a) Basic components of an x-ray crystallographic determination. When a narrow beam of x-rays strikes a crystal, part of it passes straight through and the rest is scattered (diffracted) in various directions. The intensity of the diffracted (more...)

Cryoelectron Microscopy

Although some proteins readily crystallize, obtaining crystals of others—particularly large multisubunit proteins—requires a time-consuming trial-and-error effort to find just the right

conditions. Low-resolution views of such difficult-to-crystallize proteins can be obtained by electron microscopy (see Figure 4-34). In this technique, a protein sample is rapidly frozen in liquid helium to preserve its structure and then examined in the frozen, hydrated state in a *cryoelectron microscope.* Pictures are recorded on film using a low dose of electrons to prevent radiation-induced damage to the structure. Sophisticated computer programs analyze the images and reconstruct the protein's structure in three dimensions. With recent advances in cryoelectron microscopy, this technique can generate molecular models that compare with those generated by x-ray crystallography.

NMR Spectroscopy

The three-dimensional structures of small proteins containing up to about 200 amino acids can be studied with nuclear magnetic resonance (NMR) spectroscopy. In this technique, a concentrated protein solution is placed in a magnetic field and the effects of different radio frequencies on the resonances of different atoms are measured. However, the behavior of any atom is influenced by neighboring atoms in adjacent residues; the closely spaced residues are more perturbed than distant residues. From the magnitude of the effect, the distances between residues can be calculated; these distances then are used to generate a model of the three-dimensional structure of the protein. Although NMR does not require crystallization of a protein, a definite advantage, this technique is limited to proteins smaller than about 20 kDa. However, NMR analysis also can be applied to protein domains, which tend to be small enough for this technique and often can be obtained as stable structures.

SUMMARY

- Proteins can be isolated based on differences in their physical and chemical properties. Centrifugation, electrophoresis, and chromatography are the most common techniques for purifying and analyzing proteins.
- Centrifugation separates proteins based on their rate of sedimentation, which is influenced by their mass and shape.
- Gel electrophoresis separates proteins based on their rate of movement in an applied electric field. SDSpolyacrylamide gel electrophoresis can resolve polypeptide chains differing in molecular weight by 10 percent or less (see Figure 3-41).
- Liquid chromatography separates proteins based on their rate of movement through a column packed with spherical beads. Proteins differing in mass are resolved on gel filtration columns; those differing in charge, on ion-exchange columns; and those differing in ligand-binding properties, on affinity columns (see Figure 3-43).
- Various assays are used to detect, identify, and quantify proteins. The most sensitive assays use a lightproducing reaction or radioactivity to generate a signal. Other assays produce an amplified colored signal with enzymes and chromogenic substrates.
- Antibodies are powerful reagents used to detect, quantify, and isolate proteins. They are used in affinity chromatography and combined with gel electrophoresis in Western blotting, a powerful method for separating and detecting a protein in a mixture (see Figure 3-44).
- Autoradiography is a semiquantitative technique for detecting radioactively labeled molecules in cells, tissues, or electrophoretic gels (see Figure 3-45).
- Three-dimensional structures of proteins are obtained by x-ray crystallography, NMR spectroscopy, and cryoelectron microscopy. X-ray crystallography provides the most detailed structures, but requires protein crystallization. Cryoelectron microscopy is particularly useful for large protein complexes, which are difficult to crystallize. Only relatively small proteins are amenable to NMR analysis.

Bookshelf ID: NBK21589