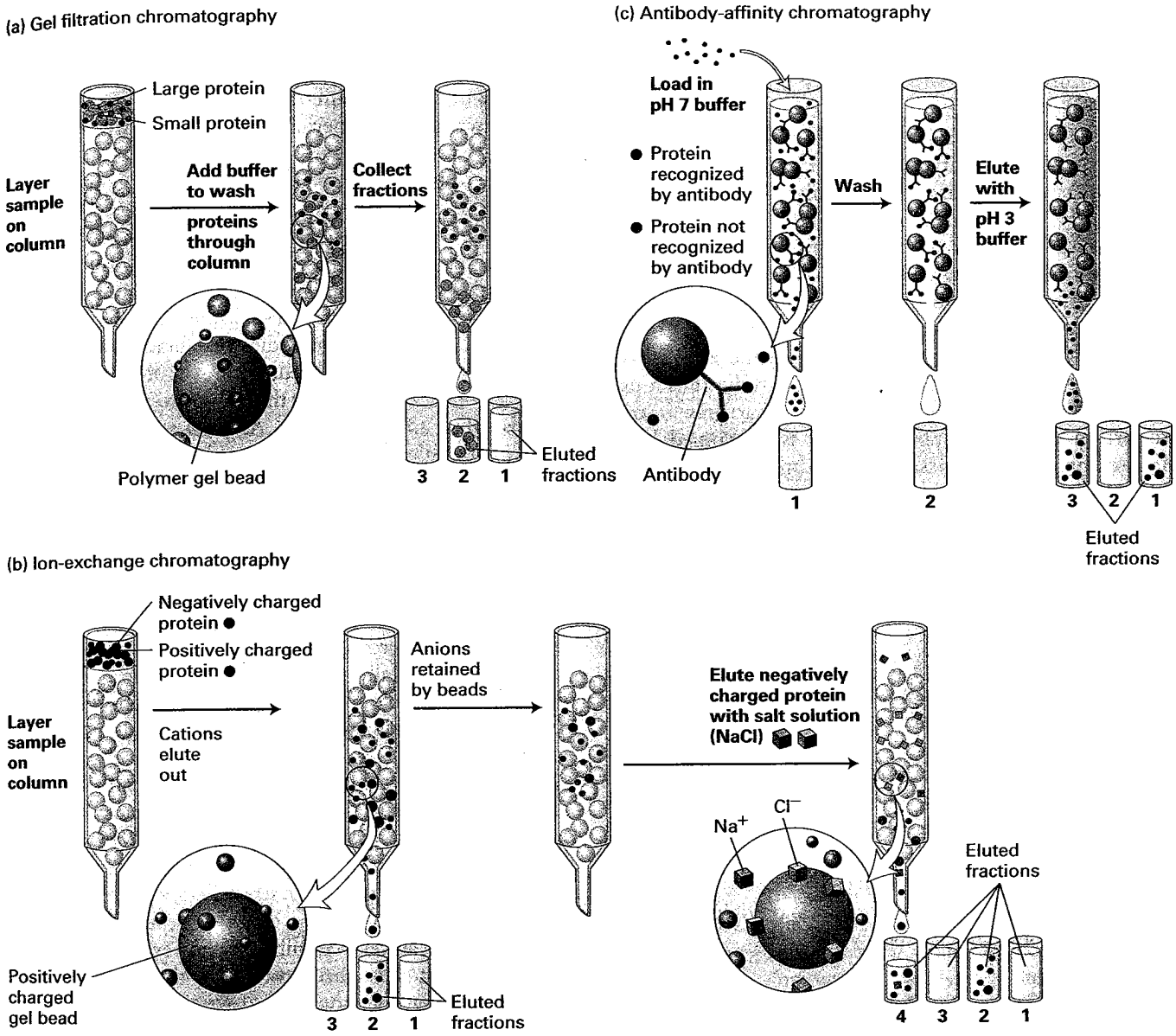


▲ **EXPERIMENTAL FIGURE 3-37 Three commonly used liquid chromatographic techniques separate proteins on the basis of mass, charge, or affinity for a specific binding partner.** (a) Gel filtration chromatography separates proteins that differ in size. A mixture of proteins is carefully layered on the top of a cylinder packed with porous beads. Smaller proteins travel through the column more slowly than larger proteins. Thus different proteins emerging in the eluate flowing out of the bottom of the column at different times (different elution volumes) can be collected in separate tubes, called *fractions.* (b) Ion-exchange chromatography separates proteins that differ in net charge in columns packed with special beads that carry either a positive charge (shown here) or a negative charge. Proteins having the same net charge as the beads are repelled and flow through the column, whereas proteins having the opposite charge bind to the beads more or less tightly, depending on their structures. Bound proteins—in this case, negatively charged—are subsequently eluted by passing a salt gradient (usually of NaCl or KCl) through the column. As the ions bind to the beads, they displace the proteins (more tightly bound proteins require higher salt concentration in order to be released). (c) In antibody-affinity chromatography, a mixture of proteins is passed through a column packed with beads to which a specific antibody is covalently attached. Only protein with high affinity for the antibody is retained by the column; all the nonbinding proteins flow through. After the column is washed, the bound protein is eluted with an acidic solution or some other solution that disrupts the antigen–antibody complexes; the released protein then flows out of the column and is collected.

retained on the column. The proteins bound to the affinity column are then eluted by adding an excess of a soluble form of the ligand or by changing the salt concentration or pH such that the binding to the molecule on the column is disrupted. The ability of this technique to separate particular proteins depends on the selection of appropriate binding partners that bind more tightly to the protein of interest than to other proteins.

Highly Specific Enzyme and Antibody Assays Can Detect Individual Proteins

The purification of a protein, or any other molecule, requires a specific assay that can detect the presence of the molecule of interest as it is separated from other molecules (e.g., in column or density-gradient fractions or gel bands or spots). 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 to minimize errors and the possibility that the protein of interest becomes 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 by it. 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, enzymatic activity assays are based on the ability to detect the loss of substrate or the formation of product. Some enzymatic assays utilize chromogenic substrates, which change color in 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; the rate of the reaction provides a measure of the quantity of enzyme present. Enzymes that catalyze chromogenic reactions can also be fused or chemically linked to an antibody and used to "report" the presence or location of an antigen to which the antibody binds (see below).

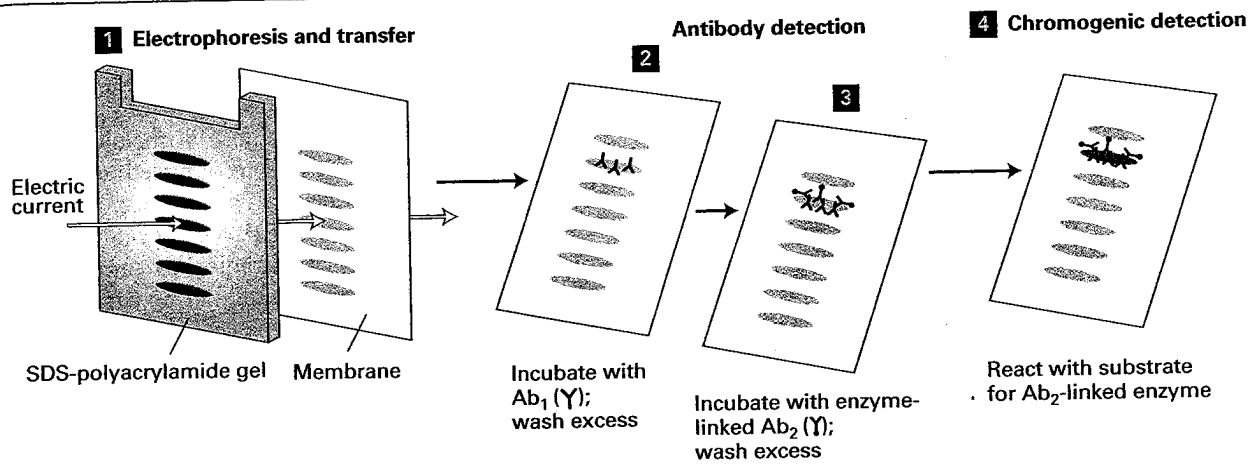
Antibody Assays As noted earlier, antibodies have the distinctive characteristic of binding tightly and specifically to antigens. As a consequence, preparations of antibodies that recognize a protein antigen of interest can be generated and used to detect the presence of the protein, either in a complex mixture of other proteins (finding a needle in a haystack, as it were) or in a partially purified preparation of a particular protein. The tight binding of the antibody to its antigen, and thus the presence of the antigen, can be visualized by labeling the antibody with an enzyme, a fluorescent molecule, or radioactive isotopes. For example, luciferase, an enzyme present in fireflies and some bacteria, can be linked to an antibody. In the presence of ATP and the substrate luciferin, luciferase catalyzes a light-emitting reaction. In either case, after the antibody binds to the protein of interest (the antigen) and unbound antibody is washed away, substrates of the linked enzyme are added and the appearance of color or emitted light is monitored. The intensity is proportional to the amount of enzyme-linked antibody, and thus antigen, in the sample. A variation of this technique, particularly useful in detecting specific proteins within living cells, makes use of *green fluorescent protein (GFP)*, a naturally fluorescent protein found in jellyfish (see Figure 9-12). Alternatively, after the first antibody binds to its target protein, a second, labeled antibody is used to bind to the complex of the first antibody and its target. This combination of two antibodies permits very high sensitivity in the detection of a target protein.

To generate the antibodies, the intact protein or a fragment of the protein is injected into an animal (usually a rabbit, mouse, or goat). Sometimes a short synthetic peptide of 10–15 residues based on the sequence of the protein is used as the antigen to induce antibody formation. A synthetic peptide, when coupled to a large protein carrier, can induce an animal to produce antibodies that bind to that portion (the epitope) of the full-sized, natural protein. Biosynthetically or chemically attaching the epitope to an unrelated protein is called *epitope tagging*. As we'll see throughout this book, antibodies generated using either peptide epitopes or intact proteins are extremely versatile reagents for isolating, detecting, and characterizing proteins.

Detecting Proteins in Gels Proteins embedded within a one- or two-dimensional gel usually are not visible. The two general approaches for detecting proteins in gels are either to label or stain the proteins while they are still within the gel or to electrophoretically transfer the proteins to a membrane made of nitrocellulose or polyvinylidene difluoride and then detect them. Proteins within gels are usually stained with an organic dye or a silver-based stain, both detected with normal visible light, or with a fluorescent dye that requires specialized detection equipment. Coomassie blue is the most commonly used organic dye, typically used to detect ≈1000 ng of protein, with a lower limit of detection of ≈4–10 ng. Silver staining or fluorescence staining are more sensitive (lower limit of ≈1 ng). Coomassie and other stains can also be used to visualize proteins after transfer to membranes; however, the most common method to visualize proteins in these membranes is immunoblotting, commonly called Western blotting.

Western blotting combines the resolving power of gel electrophoresis and the specificity of antibodies. This multistep procedure is commonly used to separate proteins and then identify a specific protein of interest. As shown in Figure 3-38, two different antibodies are used in this method, one that is specific for the desired protein and a second that binds to the first and is linked to an enzyme or other molecule that permits detection of the first antibody (and thus the protein of interest to which it binds). Enzymes to which the second antibody is attached can either generate a visible colored product or, by a process called *chemiluminescence*, produce light that can readily be recorded by film or a sensitive detector. The two different antibodies, sometimes called a "sandwich," are used to amplify the signals and improve sensitivity. If an antibody is not available, but the gene encoding the protein is available and can be used to express the protein, recombinant DNA methods (Chapter 5) can incorporate a small peptide epitope (epitope tagging) into the normal sequence of the protein that can be detected by a commercially available antibody to that epitope.

Technique Animation: Immunoblotting



▲ **EXPERIMENTAL FIGURE 3-38 Western blotting (immunoblotting) combines several techniques to resolve and detect a specific protein.** Step 1: After a protein mixture has been electrophoresed through an SDS gel, the separated bands (or spots, for a two-dimensional gel) are transferred (blotted) from the gel onto a porous membrane from which it is not readily removed. Step 2: The membrane is flooded with a solution of antibody (Ab_1) specific for the desired protein and allowed to incubate for a while. Only the membrane-bound band containing this protein binds the antibody, forming a layer of antibody molecules (whose position cannot be seen at this point). Then the membrane is washed to remove unbound Ab_1. Step 3: The membrane is incubated with a second antibody (Ab_2) that specifically recognizes and binds to the first Ab_1. This second antibody is covalently linked to either an enzyme (e.g., alkaline phosphatase, which can catalyze a chromogenic reaction), radioactive isotope, or some other substance whose presence can be detected with great sensitivity. Step 4: Finally, the location and amount of bound Ab_2 are detected (e.g., by a deep-purple precipitate from chromogenic reaction), permitting the electrophoretic mobility (and therefore the mass) of the desired protein to be determined, as well as its quantity (based on band intensity).

Radioisotopes Are Indispensable Tools for Detecting Biological Molecules

A sensitive method for tracking a protein or other biological molecule is by detecting the radioactivity emitted from radioisotopes introduced into the molecule. At least one atom in a radiolabeled molecule is present in a radioactive form, called a **radioisotope.**

TABLE 3-1 Radioisotopes Commonly Used in Biological Research

ISOTOPE	HALF-LIFE
Phosphorus-32	14.3 days
Iodine-125	60.4 days
Sulfur-35	87.5 days
Tritium (hydrogen-3)	12.4 years
Carbon-14	5730.4 years

Radioisotopes Useful in Biological Research 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 probability of decay of the radioisotope, determined 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 (Table 3-1).

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

Labeled compounds in which a radioisotope replaces atoms normally present in the molecule have virtually the same chemical properties as the corresponding nonlabeled compounds. Enzymes, for instance, generally cannot distinguish between substrates labeled in this way and their nonlabeled substrates. The presence of such radioactive atoms is indicated with the isotope in brackets (no hyphen) as a prefix (e.g., [^{3}H]leucine). In contrast, labeling almost all biomolecules (e.g., protein or nucleic acid) with the radioisotope iodine-125 (^{125}I) requires the covalent addition of ^{125}I to a molecule that normally does not have iodine as part of its structure. Because this labeling procedure modifies the chemical structure, the biological activity of the labeled molecule may differ somewhat from that of the nonlabeled form. The presence of such radioactive atoms is indicated with the isotope as a prefix with a hyphen (no bracket) (e.g., ^{125}I-trypsin). Standard methods for labeling proteins with ^{125}I result in covalent attachment of the ^{125}I primarily to the aromatic rings of tyrosine side chains (mono- and diiodotyrosine).

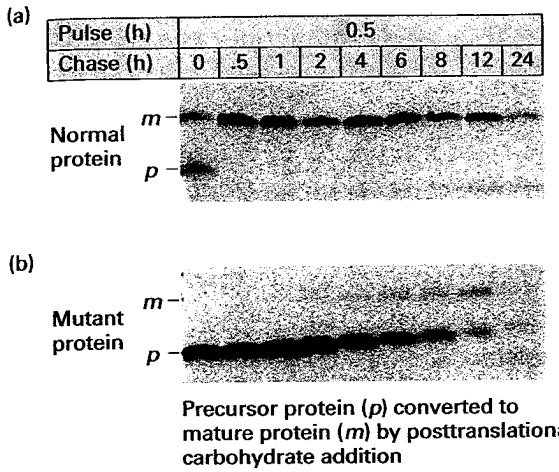
Labeling Experiments and Detection of Radiolabeled Molecules Whether labeled compounds are detected by **autoradiography**, a semiquantitative visual assay, or their radioactivity is measured in an appropriate "counter," a highly quantitative assay that can determine the amount of a radiolabeled compound in a sample, depends on the nature of the experiment. In some experiments, both types of detection are used.

In one use of autoradiography, a tissue, cell, or cell constituent is labeled with a radioactive compound, unassociated radioactive material is washed away, and the structure of the sample is stabilized either by chemically cross-linking the macromolecules ('fixation') or by freezing it. The sample is 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 and is usually detected by microscopy. Autoradiographic studies of whole cells were crucial in determining the intracellular sites where various macromolecules are synthesized and the subsequent movements of these macromolecules within cells. Various techniques employing fluorescent microscopy, which we describe in Chapter 9, have largely supplanted autoradiography for studies of this type. However, autoradiography is commonly used in various assays for detecting specific isolated DNA or RNA sequences (Chapter 5).

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 in the decay of the radioisotope; a phototube in the instrument detects and counts these light flashes. Phosphorimagers are used to detect radioactivity using a two-dimensional array detector, storing digital data on the number of decays in disintegrations per minute per small pixel of surface area. These instruments, which can be thought of as a kind of reusable electronic film, are commonly used to quantitate radioactive molecules separated by gel electrophoresis and are replacing photographic emulsions for this purpose.

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]methionine) for a few minutes, during which time the labeled amino acid mixes with the cellular pool of unlabeled amino acids and some of the labeled amino acid is biosynthetically incorporated into newly synthesized protein. Subsequently unincorporated radioactive amino acid is washed away from the cells. The cells are harvested, the mixture of cellular proteins is extracted from the cells (for example, by a detergent solution), and then separated 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 can be determined quantitatively with a counter. To detect only one specific protein, rather than all the proteins biosynthetically labeled this way, a specific antibody to the protein of interest can be used to precipitate the protein away from the other proteins in the sample (immunoprecipitation). The precipitate is then solubilized in a detergent, separating the antibody from the protein, and the sample is subjected to SDS-PAGE followed by autoradiography. In this type of experiment, a fluorescent compound that is activated by the radiation is often infused into the gel so that the light emitted can be used to detect the presence of the labeled protein, either using film or a two-dimensional electronic detector.

Pulse-chase experiments are particularly useful for tracing changes in the intracellular location of proteins or the modification of a protein or metabolite over time. In this experimental protocol, a cell sample is exposed to a radiolabeled compound that can be incorporated or otherwise attached to a cellular molecule of interest—the "pulse"—for a brief period of time, then washed with buffer to remove the unincorporated label, and finally incubated with an unlabeled form of the compound—the "chase" (Figure 3-39). Samples taken periodically during the chase period are assayed to determine the location or chemical form of the radiolabel as a function of time. Often, pulse-chase experiments, in which the protein is detected by autoradiography after immunoprecipitation and SDS-PAGE, are used to follow the rate of synthesis, modification, and degradation of proteins by adding radioactive amino acid precursors during the pulse and then detecting the amounts and characteristics of the radioactive protein during the chase. One can thus observe postsynthetic modifications of the protein that change its electrophoretic mobility and the rate of degradation of a specific protein. A classic use of the pulse-chase technique was in studies to elucidate the pathway traversed by secreted proteins from their site of synthesis in the endoplasmic reticulum to the cell surface (Chapter 14).



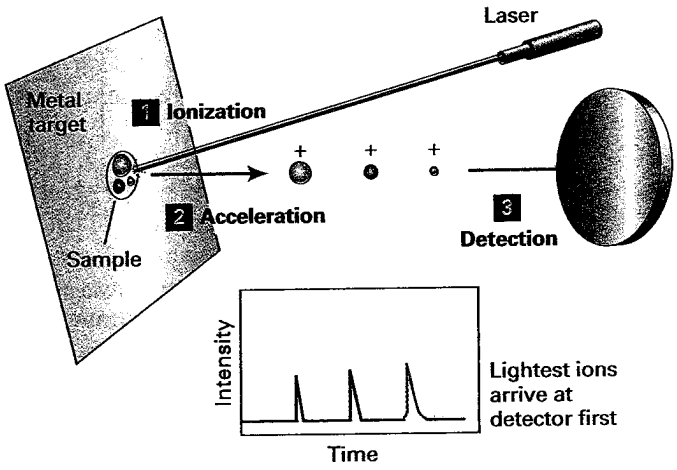
▲ **EXPERIMENTAL FIGURE 3-39 Pulse-chase experiments can track the pathway of protein modification or movement within cells.** (a) To follow the fate of a specific newly synthesized protein in a cell, cells were incubated with [^{35}S]methionine for 0.5 hr (the pulse) to label all newly synthesized proteins, and the radioactive amino acid not incorporated into the cells was then washed away. The cells were further incubated (the chase) for varying times up to 24 hours, and samples from each time of chase were subjected to immunoprecipitation to isolate one specific protein (here, the low-density lipoprotein receptor). SDS-PAGE of the immunoprecipitates followed by autoradiography permitted visualization of the one specific protein, which is initially synthesized as a small precursor (p) and then rapidly modified to a larger mature form (m) by addition of carbohydrates. About half of the labeled protein was converted from p to m during the pulse, the rest was converted after 0.5 hour of chase. The protein remains stable for 6–8 hours before it begins to be degraded (indicated by reduced band intensity). (b) The same experiment was performed in cells in which a mutant form of the protein is made. The mutant p form cannot be properly converted to the m form, and it is more quickly degraded than the normal protein. [Adapted from K. F. Kozarsky, H. A. Brush, and M. Krieger, 1986, J. Cell *Biol.* **102**(5):1567–1575.]

Mass Spectrometry Can Determine the Mass and Sequence of Proteins

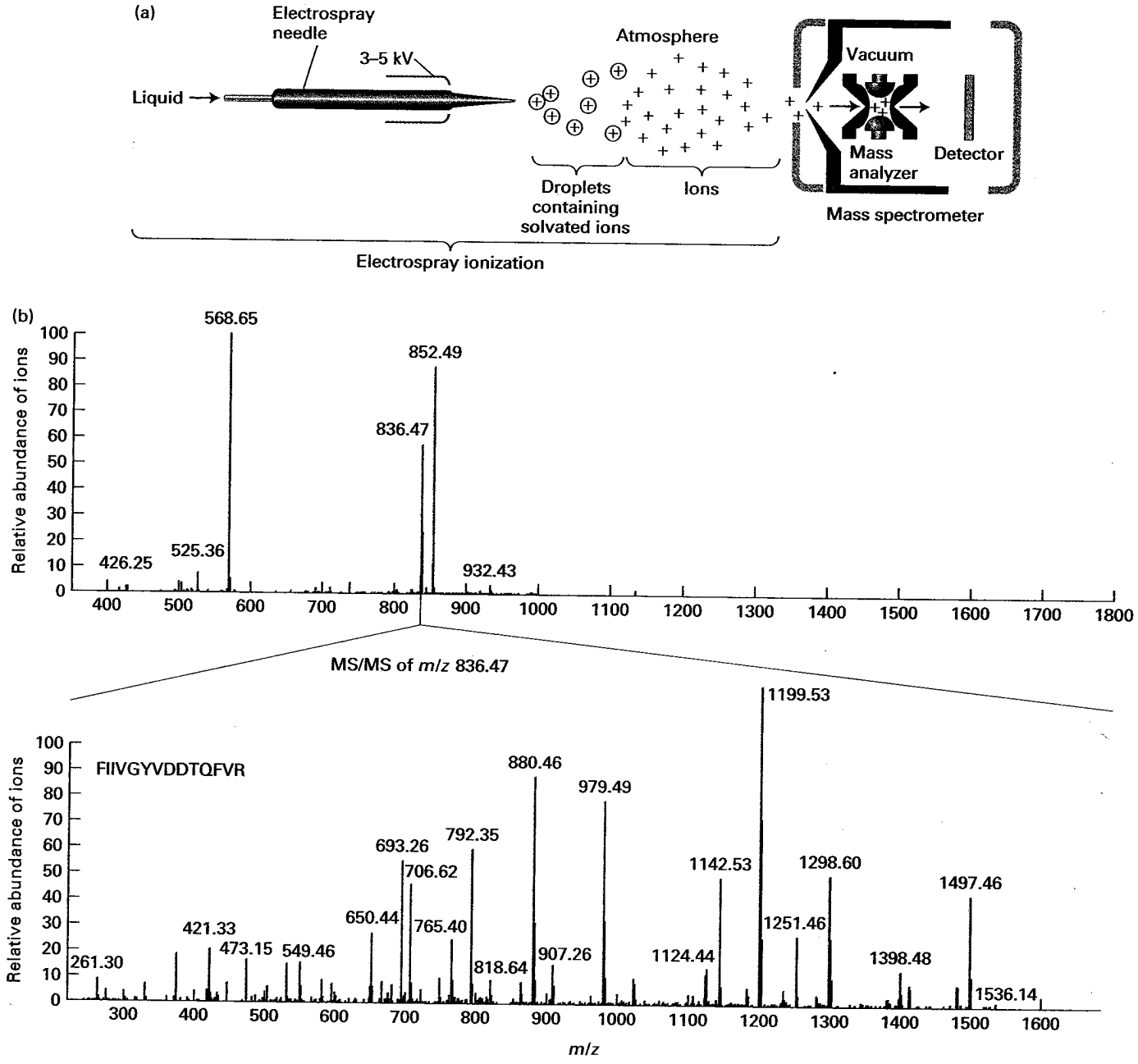
Mass spectrometry (MS) is a powerful technique for characterizing proteins. MS is particularly useful in determining the mass of a protein or fragments of a protein. With such information in hand, it is also possible to determine part of or all the protein's sequence. This method permits the very highly accurate direct determination of the ratio of the mass (m) of a charged molecule (ion) to its charge (z), or m/z. Techniques are then used to deduce the absolute mass of the ion. There are four key features of all mass spectrometers. The first is an ion source, from which charge, usually in the form of protons, is transferred to the peptide or protein molecules. The formation of these ions occurs in the presence of a high electric field that then directs the charged molecular ions into the second key component, the mass analyzer. The mass analyzer, which is always in a high vacuum chamber, physically separates the ions on the basis of their differing mass-to-charge (m/z) ratios. The mass-separated ions are subsequently directed to strike a detector, the third key component, which provides a measure of the relative abundances of each of the ions in the sample. The fourth essential component is a computerized data system that is used to calibrate the instrument; acquire, store, and process the resulting data; and often direct the instrument automatically to collect additional specific types of data from the sample, based on the initial observations. This type of automated feedback is used for the tandem MS (MS/MS) peptide-sequencing methods described below.

The two most frequently used methods of generating ions of proteins and protein fragments are (1) matrix-assisted laser desorption/ionization (MALDI) and (2) electrospray (ES). In MALDI (Figure 3-40) the peptide or protein sample is mixed with a low-molecular-weight, UV-absorbing organic acid (the matrix) and then dried on a metal target. Energy from a laser ionizes and vaporizes the sample producing singly charged molecular ions from the constituent molecules. In ES (Figure 3-41a), the sample of peptides or proteins in solution is converted into a fine mist of tiny droplets by spraying through a narrow capillary at atmospheric pressure. The droplets are formed in the presence of a high electric field, rendering them highly charged. The droplets evaporate in their short flight (mm) to the entrance of the mass spectrometer's analyzer, forming multiply charged ions from the peptides and proteins. The gaseous ions are sampled into the analyzer region of the MS, where they are then accelerated by electric fields and separated by the mass analyzer on the basis of their m/z.

The two most frequently used mass analyzers are time-of-flight (TOF) instruments and ion traps. TOF instruments exploit the fact that the time it takes an ion to pass through the length of the analyzer before reaching the detector is



▲ **EXPERIMENTAL FIGURE 3-40 Molecular mass can be determined by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry.** In a MALDI-TOF mass spectrometer, pulses of light from a laser ionize a protein or peptide mixture that is absorbed on a metal target (step 1). An electric field accelerates the ions in the sample toward the detector (steps 2 and 3). The time to the detector is proportional to the square root of the mass-to-charge (m/z) ratio. For ions having the same charge, the smaller ions move faster (shorter time to the detector). The molecular weight is calculated using the time of flight of a standard.



▲ **EXPERIMENTAL FIGURE 3-41 Molecular mass of proteins and peptides can be determined by electrospray ionization ion-trap mass spectrometry.** (a) Electrospray (ES) ionization converts proteins and peptides in a solution into highly charged gaseous ions by passing the solution through a needle (forming the droplets) that has a high voltage across it (charging the droplets). Evaporation of the solvent produces gaseous ions that enter a mass spectrometer. The ions are analyzed by an ion-trap mass analyzer that then directs ions to the detector. (b) *Top panel:* Mass spectrum of a mixture of three major and several minor peptides is presented as the relative abundance of the ions striking the detector (y axis) as a function of the mass-to-charge (*m/z*) ratio (x axis). *Bottom panel:* In an MS/MS instrument (such as the ion trap shown in part (a), a specific peptide ion can be selected for fragmentation into smaller ions that are then analyzed and detected. The MS/MS spectrum (also called the product-ion spectrum) provides detailed structural information about the parent ion, including sequence information for peptides. Here, the ion with a *m/z* of 836.47 was selected, fragmented and the *m/z* mass spectrum of the product ions was measured. Note there is no longer an ion with an *m/z* of 836.47 present because it was fragmented. From the varying sizes of the product ions, the understanding that peptide bonds are often broken in such experiments, the known *m/z* values for individual amino acid fragments, and database information, the sequence of the peptide, FIIVGYVDDTQFVR, can be deduced. [Part (a) based on a figure from S. Carr; part (b), unpublished data from S. Carr.]

proportional to the square root of *m/z* (smaller ions move faster than larger ones with the same charge, see Figure 3-40). In ion-trap analyzers, tunable electric fields are used to capture, or 'trap,' ions with a specific *m/z* and to sequentially pass the trapped ions out of the analyzer onto the detector (see Figure 3-41a). By varying the electric fields, ions with a wide range of *m/z* values can be examined one by one, producing a mass spectrum, which is a graph of *m/z* (x axis) versus relative abundance (y axis) (Figure 3-41b, *top panel*).

In tandem, or MS/MS, instruments, any given parent ion in the original mass spectrum (Figure 3-41b, *top panel*) can be mass-selected, broken into smaller ions by collision with an inert gas, and then the *m/z* and relative abundances of the resulting fragment ions measured (Figure 3-41b, *bottom panel*), all within the same machine in about 0.1 s per selected parent ion. This second round of fragmentation and analysis permits the sequences of short peptides (<25 amino acids) to be determined, because collisional fragmentation occurs primarily at peptide bonds, so the differences in masses between ions correspond to the in-chain masses of the individual amino acids, permitting deduction of the sequence in conjunction with database sequence information (Figure 3-41b, *bottom panel*).

Mass spectrometry is highly sensitive, able to detect as little as 1×10^{-16} mol (100 attomoles) of a peptide or 10×10^{-15} mol (10 femtomoles) of a protein of 200,000 MW. Errors in mass measurement accuracy are dependent upon the specific mass analyzer used, but are typically ≈0.01 percent for peptides and 0.05–0.1 percent for proteins. As described in the proteomics section that follows, it is possible to use MS to analyze complex mixtures of proteins as well as purified proteins. Most commonly, protein samples are digested by proteases, and the peptide digestion products are subjected to analysis. An especially powerfully application of MS is to take a complex mixture of proteins from a biological specimen, digest it with trypsin or other proteases, partially separate the components using liquid chromatography (LC), and then transfer the solution flowing out of the chromatographic column directly into an ES tandem mass spectrometer. This technique, called *LC-MS/MS*, permits the nearly continuous analysis of a very complex mixture of proteins.

The abundances of ions determined by mass spectrometry in any given sample are relative, not absolute, values. Therefore, if one wants to use MS to compare the amounts of a particular protein in two different samples (e.g., from a normal versus a mutant organism), it is necessary to have an internal standard in the samples whose amounts do not differ between the two samples. One then determines the amounts of the protein of interest relative to that of the standard in each sample. This permits quantitatively accurate intersample comparisons of protein levels.

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

The classic method for determining the amino acid sequence of a protein is Edman degradation. In this procedure, the free amino group of the N-terminal amino acid of a polypeptide is labeled, and the labeled amino acid is 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.

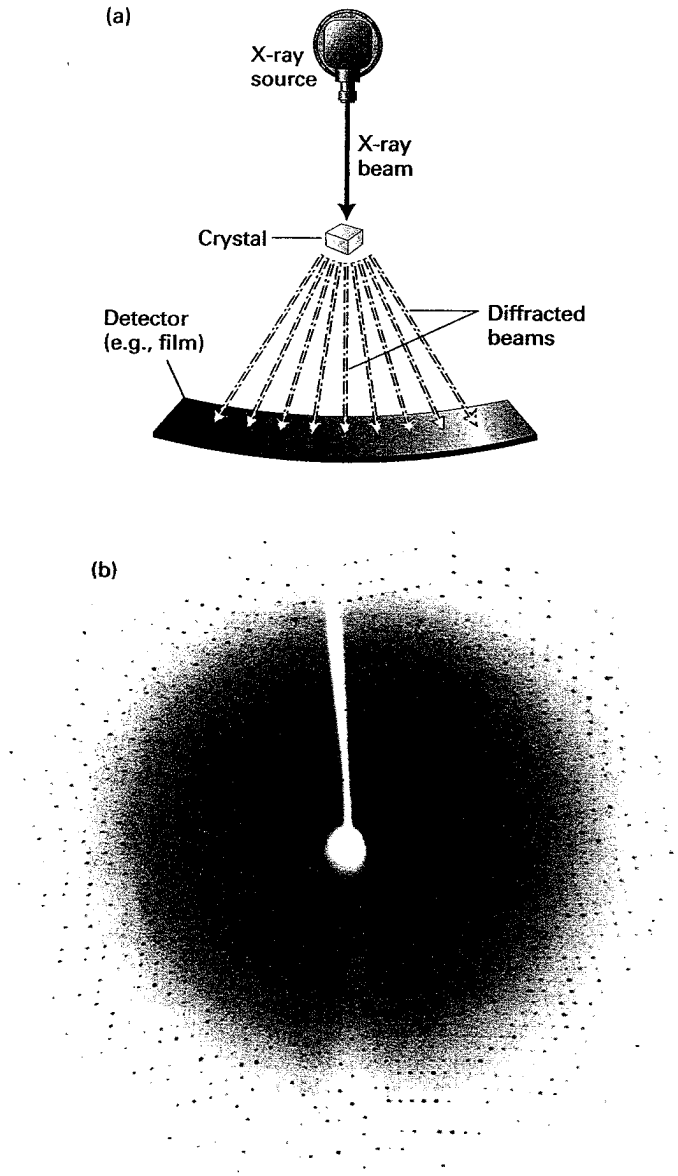
Before about 1985, biologists commonly used the Edman chemical procedure for determining protein sequences. Now, however, complete protein sequences usually are determined primarily by analysis of genome sequences. The complete genomes of several organisms have already been sequenced, and the database of genome sequences from humans and numerous model organisms is expanding rapidly. As discussed in Chapter 5, the sequences of proteins can be deduced from DNA sequences that are predicted to encode proteins.

A powerful approach for determining the primary structure of an isolated protein combines MS and the use of sequence databases. First, the peptide "mass fingerprint" of the protein is obtained by MS. A *peptide mass fingerprint* is the list of the molecular weights of peptides that are generated from the protein by digestion with a specific protease, such as trypsin. The molecular weights of the parent protein and its proteolytic fragments are then used to search genome databases for any similarly sized protein with identical or similar peptide mass maps. Mass spectrometry can also be used to directly sequence peptides using MS/MS, as described above.

Protein Conformation Is Determined by Sophisticated Physical Methods

In this chapter, we have emphasized that protein function is dependent on 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. In this technique, beams of x-rays are passed through a protein crystal in which millions of protein molecules are precisely aligned with one another in a rigid crystalline array. The wavelengths of x-rays are about 0.1–0.2 nm, short enough to determine the positions of individual atoms in the protein. The electrons in the atoms of the crystal scatter the x-rays, which produce a diffraction pattern of discrete spots when they are intercepted by photographic film or an electronic detector (Figure 3-42). Such patterns are extremely complex—composed of as many as 25,000 diffraction spots, or reflections, whose measured intensities vary depending on the distribution of the electrons, which is, in turn, determined by the atomic structure and three-dimensional conformation of the protein. Elaborate calculations and modifications of the protein (such as the binding of heavy metals) must be made to interpret the diffraction pattern and calculate the distribution of electrons (called the *electron density map*). With the three-dimensional electron density map in hand, one then "fits" a molecular model of the protein to match the electron density,



▲ **EXPERIMENTAL FIGURE 3-42 X-ray crystallography provides diffraction data from which the three-dimensional structure of a protein can be determined.** (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 waves, which form periodic arrangements of diffraction spots, is recorded on an x-ray film or with a solid-state electronic detector. (b) X-ray diffraction pattern for a protein crystal collected on a solid-state detector. From complex analyses of patterns of spots like this one, the location of the atoms in a protein can be determined. [Part (a) adapted from L. Stryer, 1995, *Biochemistry,* 4th ed., W. H. Freeman and Company, p. 64; part (b) courtesy of J. Berger.]

and it is these models that one sees in the various diagrams of proteins throughout this book (e.g., Figure 3-8). The process is analogous to reconstructing the precise shape of a rock from the ripples that it creates when thrown into a pond. Although sometimes the structures of portions of the protein cannot be clearly defined, using x-ray crystallography researchers are systematically determining the structures of representative types of most proteins. To date, the detailed three-dimensional structures of more than 10,000 proteins have been established.

Cryoelectron Microscopy Although some proteins readily crystallize, obtaining crystals of others—particularly large multisubunit proteins and membrane-associated proteins—requires a time-consuming trial-and-error effort to find just the right conditions, if they can be found at all. (Growing crystals suitable for structural studies is as much an art as a science.) There are several ways to determine the structures of such difficult-to-crystallize proteins. One is cryoelectron microscopy. 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 of the protein are taken at various angles and 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. Recent advances in cryoelectron microscopy permit researchers to generate molecular models that can help provide insight into how the protein functions. The use of cryoelectron microscopy and other types of electron microscopy for visualizing cell structures is discussed in Chapter 9.

NMR Spectroscopy The three-dimensional structures of small proteins containing as many as 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 nuclear spin states of different atoms are measured. The spin state of any atom is influenced by neighboring atoms in adjacent residues, with closely spaced residues having a greater influence than distant residues. From the magnitude of the effect, the distances between residues can be calculated by a triangulation-like process; these distances are then used to generate a model of the three-dimensional structure of the protein.

Although NMR does not require the crystallization of a protein, a definite advantage, this technique is limited to proteins smaller than about 20 kDa. However, NMR analysis can also be applied to isolated protein domains, which can often be obtained as stable structures and tend to be small enough for this technique.

KEY CONCEPTS OF SECTION 3.6

Purifying, Detecting, and Characterizing Proteins

- Proteins can be separated from other cell components and from one another on the basis of differences in their physical and chemical properties.
- Various assays are used to detect and quantify proteins. Some assays use a light-producing reaction to generate a readily detected signal. Other assays produce an amplified colored signal with enzymes and chromogenic substrates.
- Centrifugation separates proteins on the basis of their rates of sedimentation, which are influenced by their masses and shapes.

- Electrophoresis separates proteins on the basis of their rates of movement in an applied electric field. SDS-polyacrylamide gel electrophoresis (PAGE) can resolve polypeptide chains differing in molecular weight by 10 percent or less (see Figure 3-35). Two-dimensional gel electrophoresis provides additional resolution by separating proteins first by charge (first dimension) and then by mass (second dimension).
- Liquid chromatography separates proteins on the basis of their rates 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, including antibody-based affinity chromatography (see Figure 3-37).
- Antibodies are powerful reagents used to detect, quantify, and isolate proteins.
- Immunoblotting, also called *Western blotting,* is a frequently used method to study specific proteins that exploits the high specificity and sensitivity of protein detection by antibodies and the high-resolution separation of proteins by SDS-PAGE.
- Radioisotopes play a key role in the study of proteins and other biomolecules. They can be incorporated into molecules without changing the chemical composition of the molecule, or as add-on tags. They can be used to help detect the synthesis, location, processing, and stability of proteins.
- Autoradiography is a semiquantitative technique for detecting radioactively labeled molecules in cells, tissues, or electrophoretic gels.
- Pulse-chase labeling can determine the intracellular fate of proteins and other metabolites (see Figure 3-39).
- Mass spectrometry is a very sensitive and highly precise method of detecting, identifying, and characterizing proteins and peptides.
- Three-dimensional structures of proteins are obtained by x-ray crystallography, cryoelectron microscopy, and NMR spectroscopy. X-ray crystallography provides the most detailed structures but requires protein crystallization. Cryoelectron microscopy is most useful for large protein complexes, which are difficult to crystallize. Only relatively small proteins are amenable to NMR analysis.

3.7 Proteomics

For most of the twentieth century, the study of proteins was restricted primarily to the analysis of individual proteins. For example, one would study an enzyme by determining its enzymatic activity (substrates, products, rate of reaction, requirement for cofactors, pH, etc.), its structure, and its mechanism of action. In some cases, the relationships between a few enzymes that participate in a metabolic pathway might also be studied. On a broader scale, the localization and activity of an enzyme would be examined in the context of a cell or tissue. The effects of mutations, diseases, or drugs on the expression and activity of the enzyme might also be the subject of investigation. This multipronged approach provided deep insight into the function and mechanisms of action of individual proteins or relatively small numbers of interacting proteins. However, this one-by-one approach to studying proteins does not efficiently provide insights into a global picture of what is happening in the proteome of a cell, tissue, or entire organism.

Proteomics Is the Study of All or a Large Subset of Proteins in a Biological System

The advent of genomics (sequencing of genomic DNA and its associated technologies, such as simultaneous analysis of the levels of all mRNAs in cells and tissues) clearly showed that a global, or systems, approach to biology could provide unique and highly valuable insights. Many scientists recognized that a global analysis of the proteins in biological systems had the potential for equally valuable contributions to our understanding. Thus, a new field was born—**proteomics.** Proteomics is the systematic study of the amounts, modifications, interactions, localization, and functions of all or subsets of proteins at the whole-organism, tissue, cellular, and subcellular levels.

A number of broad questions are addressed in proteomic studies:

- In a given sample (whole organism, tissue, cell, subcellular compartment), what fraction of the whole proteome is expressed (i.e., which proteins are present)?
- Of those proteins present in the sample, what are their relative abundances?
- What are the relative amounts of the different splice forms and chemically modified forms (e.g., phosphorylated, methylated, fatty acylated) of the proteins?
- Which proteins are present in large multiprotein complexes, and which proteins are in each complex? What are the functions of these complexes and how do they interact?
- When the state (e.g., growth rate, stage of cell cycle, differentiation, stress level) of a cell changes, do the proteins in the cell or secreted from the cell change in a characteristic (*fingerprint*-like) fashion? Which proteins change and how (relative amounts, modifications, splice forms, etc.)? (This is a form of *protein expression profiling* that complements the *transcriptional (mRNA) profiling* discussed in Chapter 7.)
- Can such fingerprint-like changes be used for diagnostic purposes? For example, do certain cancers or heart disease cause characteristic changes in blood proteins? Can the proteomic fingerprint help determine if a given cancer is resistant or sensitive to a particular chemotherapeutic drug? Proteomic fingerprints can also be the starting point for studies of the mechanisms underlying the change of state. Proteins (and other biomolecules) that show changes that are diagnostic of a particular state are called *biomarkers.*
- Can changes in the proteome help define targets for drugs or suggest mechanisms by which that drug might induce toxic side effects? If so, it might be possible to engineer modified versions of the drug with fewer side effects.

These are just a few of the questions that can be addressed using proteomics. The methods used to answer these questions are as diverse as the questions themselves, and their numbers are growing rapidly.

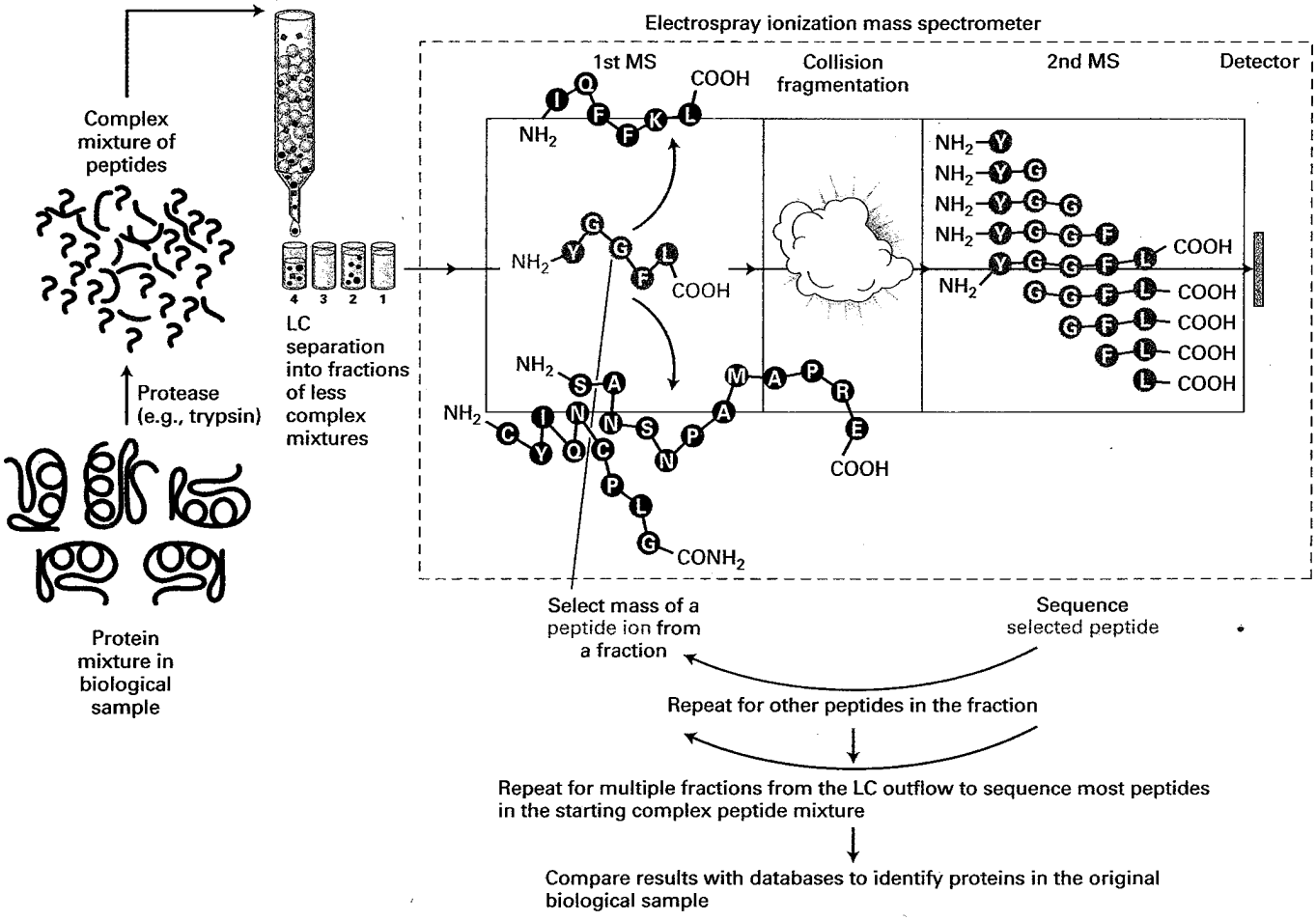
Advanced Techniques in Mass Spectrometry Are Critical to Proteomic Analysis

Advances in proteomics technologies (e.g., mass spectrometry) are having a profound effect on the types of questions that can be practically studied. For many years, two-dimensional gel electrophoresis has allowed researchers to separate, display, and characterize a mixture of proteins (Figure 3-36). The spots on a two-dimensional gel can be excised, the protein fragmented by proteolysis (e.g. by trypsin digestion), and the fragments identified by MS. An alternative to this two-dimensional gel method is *high throughput LC-MS/MS*. Figure 3-43 outlines the general LC-MS/MS approach in which a complex mixture of proteins is digested with a protease, the myriad resulting peptides are fractionated by LC into multiple, less complex fractions, which are slowly but continuously injected by electrospray ionization into a tandem mass spectrometer. The fractions are then sequentially subjected to multiple cycles of MS/MS until sequences of many of the peptides are determined and used to identify from databases the proteins in the original biological sample.

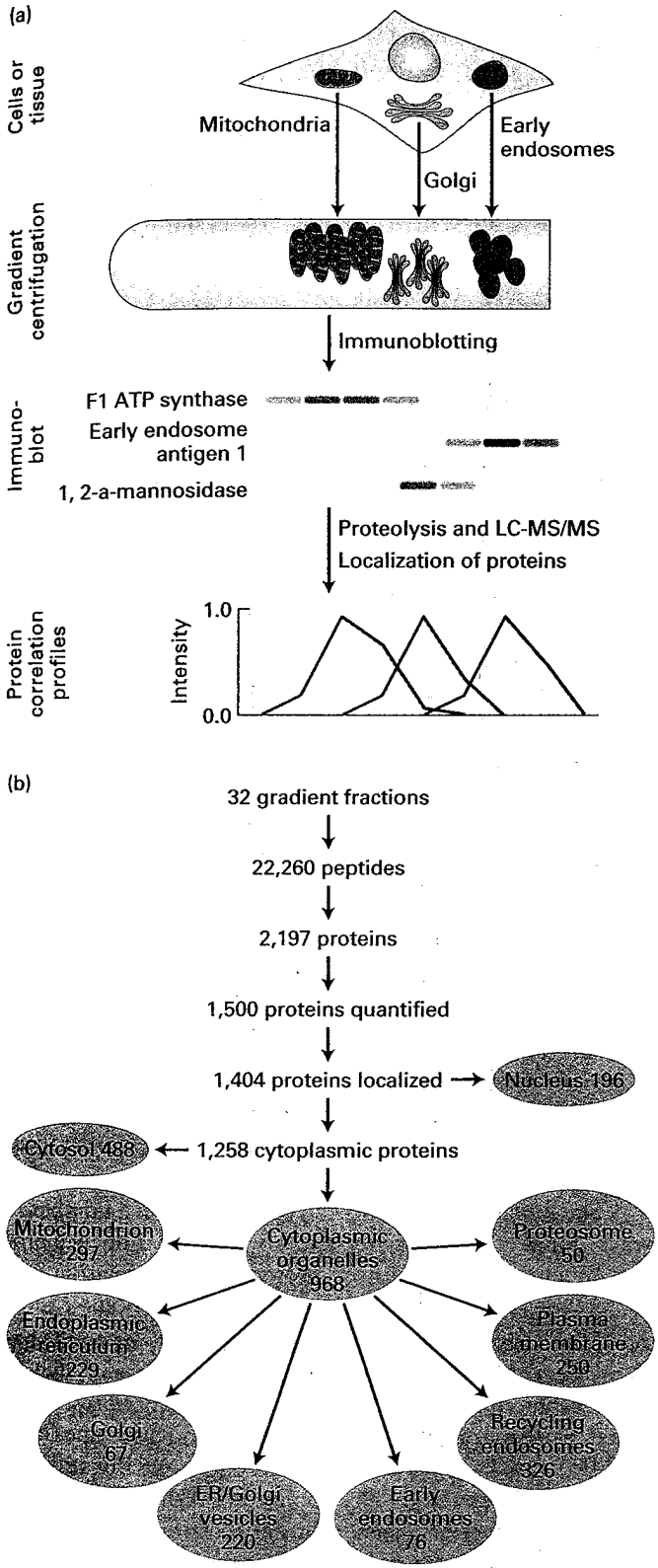
An example of the use of LC-MS/MS to identify many of the proteins in each organelle is seen in Figure 3-44. Cells from murine (mouse) liver tissue were mechanically broken to release the organelles, and the organelles were partially separated by density-gradient centrifugation. The locations of the organelles in the gradient were determined using immunoblotting with antibodies that recognize previously identified, organelle-specific proteins. Fractions from the gradient

MEDIA CONNECTIONS

Podcast: Use of Mass Spectrometry in Cell Biology



▲ **EXPERIMENTAL FIGURE 3-43 LC-MS/MS is used to identify the proteins in a complex biological sample.** A complex mixture of proteins in a biological sample (e.g., isolated preparation of Golgi organelles) is digested with a protease, the mixture of resulting peptides is fractionated by liquid chromatography (LC) into multiple, less complex, fractions, which are slowly but continuously injected by electrospray ionization into a tandem mass spectrometer. The fractions are then sequentially subjected to multiple cycles of MS/MS until masses and sequences of many of the peptides are determined and used to identify the proteins in the original biological sample through comparison with protein databases. [Based on a figure provided by S. Carr.]



◀ **EXPERIMENTAL FIGURE 3-44 Density-gradient centrifugation and LC-MS/MS can be used to identify many of the proteins in organelles.** (a) The cells in liver tissue were mechanically broken to release the organelles, and the organelles were partially separated by density-gradient centrifugation. The locations of the organelles—which were spread out through the gradient and somewhat overlapped with one another—were determined using immunoblotting with antibodies that recognize previously identified, organelle-specific proteins. Fractions from the gradient were subjected to proteolysis and LC-MS/MS to identify the peptides, and hence the proteins, in each fraction. Comparisons with the locations of the organelles in the gradient (called *protein correlation profiling*) permitted assignment of many individual proteins to one or more organelles (organelle proteome identification). (b) The hierarchical breakdown of data derived from the procedures in part (a). Note that not all proteins identified could be assigned to organelles and some proteins were assigned to more than one organelle. [From L. J. Foster et al., 2006, *Cell* **125**(1):187–199.]

were subjected to LC-MS/MS to identify the proteins in each fraction, and the distributions in the gradient of many individual proteins were compared with the distributions of the organelles. This permitted assignment of many individual proteins to one or more organelles (organelle proteome profiling).

Proteomics combined with molecular genetics methods are currently being used to identify all protein complexes in a eukaryotic cell, the yeast *Saccharomyces cerevisiae*. Approximately 500 complexes have been identified, with an average of 4.9 distinct proteins per complex, and these in turn are involved in at least 400 complex-to-complex interactions. Such systematic proteomic studies are providing new insights into the organization of proteins within cells and how proteins work together to permit cells to live and function.

KEY CONCEPTS OF SECTION 3.7

Proteomics

- Proteomics is the systematic study of the amounts (and changes in the amounts), modifications, interactions, localization, and functions of all or subsets of all proteins in biological systems at the whole-organism, tissue, cellular, and subcellular levels.
- Proteomics provides insights into the fundamental organization of proteins within cells, and how this organization is influenced by the state of the cells (e.g., differentiation into distinct cell types; response to stress, disease, and drugs).
- A wide variety of methods are used for proteomic analyses, including two-dimensional gel electrophoresis, density-gradient centrifugation, and mass spectroscopy (MALDI-TOF and LC-MS/MS).
- Proteomics has helped begin to identify the proteomes of organelles ("organelle proteome profiling") and the organization of individual proteins into multiprotein complexes that interact in a complex network to support life and cellular function

Perspectives for the Future

Impressive expansion of the computational power of computers is at the core of advances in determining the three-dimensional structures of proteins. For example, vacuum tube computers running on programs punched on cards were used to solve the first protein structures on the basis of x-ray crystallography. In the future, researchers aim to predict the structures of proteins, using only amino acid sequences

deduced from gene sequences. This computationally challenging problem requires supercomputers or large clusters of computers working in synchrony. Currently, only the structures of very small domains containing 100 residues or fewer can be predicted at a low resolution. However, continued developments in computing and models of protein folding, combined with large-scale efforts to solve the structures of all protein structural motifs by x-ray crystallography, will allow the prediction of the structures of larger proteins. With an exponentially expanding database of structurally defined motifs, domains, and proteins, scientists will be able to identify the motifs in an unknown protein, match the motif to the sequence, and use this to predict the three-dimensional structure of the entire protein.

New combined approaches will also help in determining high-resolution structures of molecular machines. Although these very large macromolecular assemblies usually are difficult to crystallize and thus to solve by x-ray crystallography, they can be imaged in a cryoelectron microscope at liquid helium temperatures and high electron energies. From millions of individual "particles," each representing a random view of the protein complex, the three-dimensional structure can be built. Because subunits of the complex may already be solved by crystallography, a composite structure consisting of the x-ray-derived subunit structures fit to the EM-derived model will be generated. An example of this approach for an electron-transport "supercomplex" is described in Chapter 12.

Methods for rapid structure determination combined with identification of novel substrates and inhibitors will help determine the structures of enzyme–substrate complexes and transition states, and thus help provide detailed information regarding the mechanisms of enzyme catalysis.

The rapid development of new technologies can be expected to help solve some of the still outstanding problems in proteomics. It should soon be possible to identify and sequence intact proteins in complex mixtures using MS techniques without first digesting the samples into peptides. An ongoing problem in proteomic analysis of complex mixtures is that it is difficult to detect and identify protein fragments from samples whose concentrations in the sample differ by more than 1000-fold: some samples, such as blood plasma, contain proteins whose concentrations vary over a 10^{11}-fold range. Routine analysis of specimens with such diverse concentrations should dramatically improve the mechanistic and diagnostic value of blood plasma proteomics.

Key Terms

α helix *66*
activation energy *79*
active site *80*
allostery *89*
amyloid filament *77*
autoradiography *100*
β sheet *66*
β turn *66*
chaperones *76*
conformations *63*
cooperativity *89*
domain *70*
electrophoresis *94*
homology *72*
K_m *80*
ligand *78*
liquid chromatography *96*
molecular machine *64*
motif *68*
motor proteins *85*
peptide bond *65*
polypeptides *66*
primary structure *66*
proteasomes *64*
protein *66*
proteome *64*
proteomics *105*
quaternary structure *72*
rate-zonal centrifugation *93*
secondary structure *66*
tertiary structure *67*
ubiquitin *88*
V_{max} *80*
Western blotting *98*
x-ray crystallography *103*

Review the Concepts

1. The three-dimensional structure of a protein is determined by its primary, secondary, and tertiary structures. Define the *primary, secondary,* and *tertiary structures.* What are some of the common secondary structures? What are the forces that hold together the secondary and tertiary structures?

2. Proper folding of proteins is essential for biological activity. Describe the roles of molecular chaperones and chaperonins in the folding of proteins.

3. Enzymes can catalyze chemical reactions. How do enzymes increase the rate of a reaction? What constitutes the active site of an enzyme? For an enzyme-catalyzed reaction, what are K_m and V_{max}? For enzyme X, the K_m for substrate A is 0.4 mM and for substrate B is 0.01 mM. Which substrate has a higher affinity for enzyme X?

4. Motor proteins convert energy into a mechanical force. Describe the three general properties characteristic of motor proteins.

5. Proteins are degraded in cells. What is ubiquitin, and what role does it play in tagging proteins for degradation? What is the role of proteasomes in protein degradation? How might proteasome inhibitors serve as chemotherapeutic (cancer-treating) agents?

6. The function of proteins can be regulated in a number of ways. What is cooperativity, and how does it influence protein function? Describe how protein phosphorylation and proteolytic cleavage can modulate protein function.

7. A number of techniques can separate proteins on the basis of their differences in mass. Describe the use of two of these techniques, centrifugation and gel electrophoresis. The blood proteins transferrin (MW 76 kDa) and lysozyme (MW 15 kDa) can be separated by rate-zonal centrifugation or SDS-polyacrylamide gel electrophoresis. Which of the two proteins will sediment faster during centrifugation? Which will migrate faster during electrophoresis?

8. Chromatography is an analytical method used to separate proteins. Describe the principles for separating proteins by gel filtration, ion-exchange, and affinity chromatography.

9. Various methods have been developed for detecting proteins. Describe how radioisotopes and autoradiography can be used for labeling and detecting proteins. How does Western blotting detect proteins?

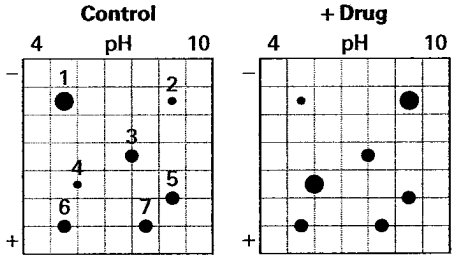
10. Physical methods are often used to determine protein conformation. Describe how x-ray crystallography, cryo-electron microscopy, and NMR spectroscopy can be used to determine the shape of proteins.

11. Mass spectrometry is a powerful tool in proteomics. What are the four key features of a mass spectrometer? Describe briefly how MALDI and two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) could be used to identify a protein expressed in cancer cells but not in normal healthy cells.

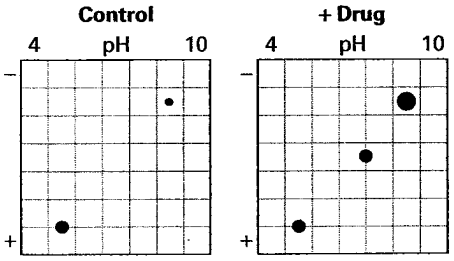
Analyze the Data

Proteomics involves the global analysis of protein expression. In one approach, all the proteins in control cells and treated cells are extracted and subsequently separated using two-dimensional gel electrophoresis. Typically, hundreds or thousands of protein spots are resolved and the steady-state levels of each protein are compared between control and treated cells. In the following example, only a few protein spots are shown for simplicity. Proteins are separated in the first dimension on the basis of charge by isoelectric focusing (pH 4–10) and then separated by size by SDS-polyacrylamide gel electrophoresis. Proteins are detected with a stain such as Coomassie blue and assigned numbers for identification.

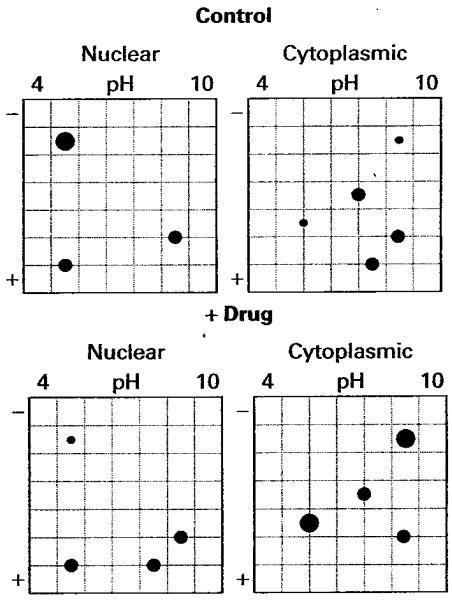
a. Cells are treated with a drug ("+ Drug") or left untreated ("Control"), and then proteins are extracted and separated by two-dimensional gel electrophoresis. The stained gels are shown below. What do you conclude about the effect of the drug on the steady-state levels of proteins 1–7?



b. You suspect that the drug may be inducing a protein kinase and so repeat the experiment in part (a) in the presence of ^{32}P-labeled inorganic phosphate. In this experiment the two-dimensional gels are exposed to x-ray film to detect the presence of ^{32}P-labeled proteins. The x-ray films are shown below. What do you conclude from this experiment about the effect of the drug on proteins 1–7?



c. To determine the cellular localization of proteins 1–7, the cells from part (a) were separated into nuclear and cytoplasmic fractions by differential centrifugation. Two-dimensional gels were run, and the stained gels are shown below. What do you conclude about the cellular localization of proteins 1–7?



d. Summarize the overall properties of proteins 1–7, combining the data from parts (a), (b), and (c). Describe how you could determine the identity of any one of the proteins.

References

General References

Berg, J. M., J. L. Tymoczko, and L. Stryer. 2007. *Biochemistry,* 6th ed. W. H. Freeman and Company.

Nelson, D. L., and M. M. Cox. 2005. *Lehninger Principles of Biochemistry,* 4th ed. W. H. Freeman and Company.

Web Sites

Entry site into proteins, structures, genomes, and taxonomy: http://www.ncbi.nlm.nih.gov/Entrez/

The protein 3D structure database: http://www.rcsb.org/

Structural classifications of proteins: http://scop.berkeley.edu/

Sites containing general information about proteins: http://www.expasy.ch/; http://www.proweb.org/; http://scop.berkeley.edu/intro.html

PROSITE Database of protein families and domains: http://www.expasy.org/prosite/

Domain organization of proteins and large collection of multiple sequence alignments: http://www.sanger.ac.uk/Software/Pfam/

Hierarchical Structure of Proteins and Protein Folding

Branden, C., and J. Tooze. 1999. *Introduction to Protein Structure.* Garland.

Brodsky, J. L., and G. Chiosis. 2006. Hsp70 molecular chaperones: emerging roles in human disease and identification of small molecule modulators. *Curr. Top. Med. Chem.* 6(11):1215–1225.

Bukau, B, J. Weissman, and A. Horwich. 2006. Molecular chaperones and protein quality control. *Cell* 125(3):443–451.

Cohen, F. E. 1999. Protein misfolding and prion diseases. *J. Mol. Biol.* 293:313–320.

Coulson, A. F., and J. Moult. 2002. A unifold, mesofold, and superfold model of protein fold use. *Proteins* 46:61–71.

Daggett, V, and A. R. Fersht. 2003. Is there a unifying mechanism for protein folding? *Trends Biochem. Sci.* 28(1):18–25.

Dobson, C. M. 1999. Protein misfolding, evolution, and disease. *Trends Biochem. Sci.* 24:329–332.

Gimona, M. 2006. Protein linguistics—a grammar for modular protein assembly? *Nat. Rev. Mol. Cell Biol.* 7(1):68–73.

Gough, J. 2006. Genomic scale sub-family assignment of protein domains. *Nucl. Acids Res.* 34(13):3625–3633.

Koonin, E. V., Y. I. Wolf, and G. P. Karev. 2002. The structure of the protein universe and genome evolution. *Nature* 420:218–223.

Lesk, A. M. 2001. *Introduction to Protein Architecture.* Oxford.

Lin, Z, and H. S. Rye. 2006. GroEL-mediated protein folding: making the impossible, possible. *Crit. Rev. Biochem. Mol. Biol.* 41(4):211–239.

Orengo, C. A., D. T. Jones, and J. M. Thornton. 1994. Protein superfamilies and domain superfolds. *Nature* 372: 631–634.

Patthy, L. 1999. *Protein Evolution.* Blackwell Science.

Rochet, J.-C., and P. T. Landsbury. 2000. Amyloid fibrillogenesis: themes and variations. *Curr. Opin. Struc. Biol.* 10:60–68.

Vogel, C, and C. Chothia. 2006. Protein family expansions and biological complexity. *PLoS Comput. Biol.* 2(5):e48.

Yaffe, M. B. 2006. "Bits" and pieces. *Sci STKE.* Jun 20 (340):pe28.

Young, J. C., et al. 2004. Pathways of chaperone-mediated protein folding in the cytosol. *Nat. Rev. Mol. Cell Biol.* 5:781–791.

Protein Function

Dressler, D. H., and H. Potter. 1991. *Discovering Enzymes.* Scientific American Library.

Fersht, A. 1999. *Enzyme Structure and Mechanism,* 3d ed. W. H. Freeman and Company.

Jeffery, C. J. 2004. Molecular mechanisms for multitasking: recent crystal structures of moonlighting proteins. *Curr. Opin. Struc. Biol.* 14(6):663–668.

Marnett, A. B., and C. S. Craik. 2005. Papa's got a brand new tag: advances in identification of proteases and their substrates. *Trends Biotechnol.* 23(2):59–64.

Polgar, L. 2005. The catalytic triad of serine peptidases. *Cell Mol. Life Sci.* 62(19–20):2161–2172.

Radisky, E. S., et al. 2006. Insights into the serine protease mechanism from atomic resolution structures of trypsin reaction intermediates. *Proc. Nat'l Acad. Sci. USA* 103(18):6835–6840.

Schenone, M, B. C. Furie, and B. Furie. 2004. The blood coagulation cascade. *Curr. Opin. Hematol.* 11(4):272–277.

Schramm, V. L. 2005. Enzymatic transition states and transition state analogues. *Curr. Opin. Struc. Biol.* 15(6):604–613.

Regulating Protein Function I: Protein Degradation

Glickman, M. H., and A. Ciechanover. 2002 The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol. Rev.* 82(2):373–428.

Goldberg, A. L. 2003. Protein degradation and protection against misfolded or damaged proteins. *Nature* 426:895–899.

Goldberg, A. L, S. J. Elledge, and J. W. Harper. 2001. The cellular chamber of doom. *Sci. Am.* 284(1):68–73.

Groll, M., and R. Huber. 2005. Purification, crystallization, and x-ray analysis of the yeast 20S proteasome. *Meth. Enzymol.* 398:329–336.

Kisselev, A. F., A. Callard, and A. L. Goldberg. 2006. Importance of the different proteolytic sites of the proteasome and the efficacy of inhibitors varies with the protein substrate. *J. Biol. Chem.* 281(13):8582–8590.

Rechsteiner, M., and C. P. Hill. 2005. Mobilizing the proteolytic machine: cell biological roles of proteasome activators and inhibitors. *Trends Cell Biol.* 15(1):27–33.

Zhou, P. 2006. REGgamma: a shortcut to destruction. *Cell* 124(2):256–257.

Zolk, O., C. Schenke, and A. Sarikas. 2006. The ubiquitin-proteasome system: focus on the heart. *Cardiovasc. Res.* 70(3):410–421.

Regulating Protein Function II: Noncovalent and Covalent Modifications

Bellelli, A., et al. 2006. The allosteric properties of hemoglobin: insights from natural and site directed mutants. *Curr. Prot. Pep. Sci.* 7(1):17–45.

Burack, W. R., and A. S. Shaw. 2000. Signal transduction: hanging on a scaffold. *Curr. Opin. Cell Biol.* 12:211–216.

Halling, D. B, P. Aracena-Parks, and S. L. Hamilton. 2006. Regulation of voltage-gated Ca^{2+} channels by calmodulin. *Sci STKE* Jan 17 (318):er1.

Horovitz, A., et al. 2001. Review: allostery in chaperonins. *J. Struc. Biol.* 135:104–114.

Kern, D., and E. R. Zuiderweg. 2003. The role of dynamics in allosteric regulation. *Curr. Opin. Struc. Biol.* 13(6):748–757.

Lane, K. T., and L. S. Beese. 2006. Thematic review series: lipid post-translational modifications. Structural biology of protein farnesyltransferase and geranylgeranyltransferase type I. *J. Lipid Res.* 47(4):681–699.

Lim, W. A. 2002. The modular logic of signaling proteins: building allosteric switches from simple binding domains. *Curr. Opin. Struc. Biol.* 12:61–68.

Martin, C., Y. Zhang. 2005. The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* 6(11):838–849.

Sawyer, T. K., et al. 2005. Protein phosphorylation and signal transduction modulation: chemistry perspectives for small-molecule drug discovery. *Med. Chem.* 1(3):293–319.

Xia, Z., and D. R. Storm. 2005. The role of calmodulin as a signal integrator for synaptic plasticity. *Nat. Rev. Neurosci.* 6(4):267–276.

Yap, K. L., et al. 1999. Diversity of conformational states and changes within the EF-hand protein superfamily. *Proteins* 37:499–507.

Purifying, Detecting, and Characterizing Proteins

Domon, B., and R. Aebersold. 2006. Mass spectrometry and protein analysis. *Science* 312(5771):212–217.

Encarnacion, S., et al. 2005. Comparative proteomics using 2-D gel electrophoresis and mass spectrometry as tools to dissect stimulons and regulons in bacteria with sequenced or partially sequenced genomes. *Biol. Proc. Online* 7:117–135.

Hames, B. D. *A Practical Approach.* Oxford University Press. A methods series that describes protein purification methods and assays.

Patton, W. F. 2002. Detection technologies in proteome analysis. *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.* 771(1–2):3–31.

White, I. R., et al. 2004. A statistical comparison of silver and SYPRO Ruby staining for proteomic analysis. *Electrophoresis* 25(17):3048–3054.

Proteomics

Foster, L. J., et al. 2006. A mammalian organelle map by protein correlation profiling. *Cell* 125(1):187–199.

Fu, Q., and J. E. Van Eyk. 2006. Proteomics and heart disease: identifying biomarkers of clinical utility. *Expert Rev. Proteomics* 3(2):237–249.

Gavin. A. C., et al. 2006. Proteome survey reveals modularity of the yeast cell machinery. *Nature* 440(7084):631–636.

Kislinger, T., et al. 2006. Global survey of organ and organelle protein expression in mouse: combined proteomic and transcriptomic profiling. *Cell* 125(1):173–186.

Kolker, E, R. Higdon, and J. M. Hogan. 2006. Protein identification and expression analysis using mass spectrometry. *Trends Microbiol.* 14(5):229–235.

Krogan, N. J., et al. 2006. Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae. Nature* 440(7084):637–643.

Ong, S. E., and M. Mann. 2005. Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol.* 1(5):252–262.

Rifai, N., M. A. Gillette, and S. A. Carr. 2006. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat. Biotech.* 24(8):971–983.