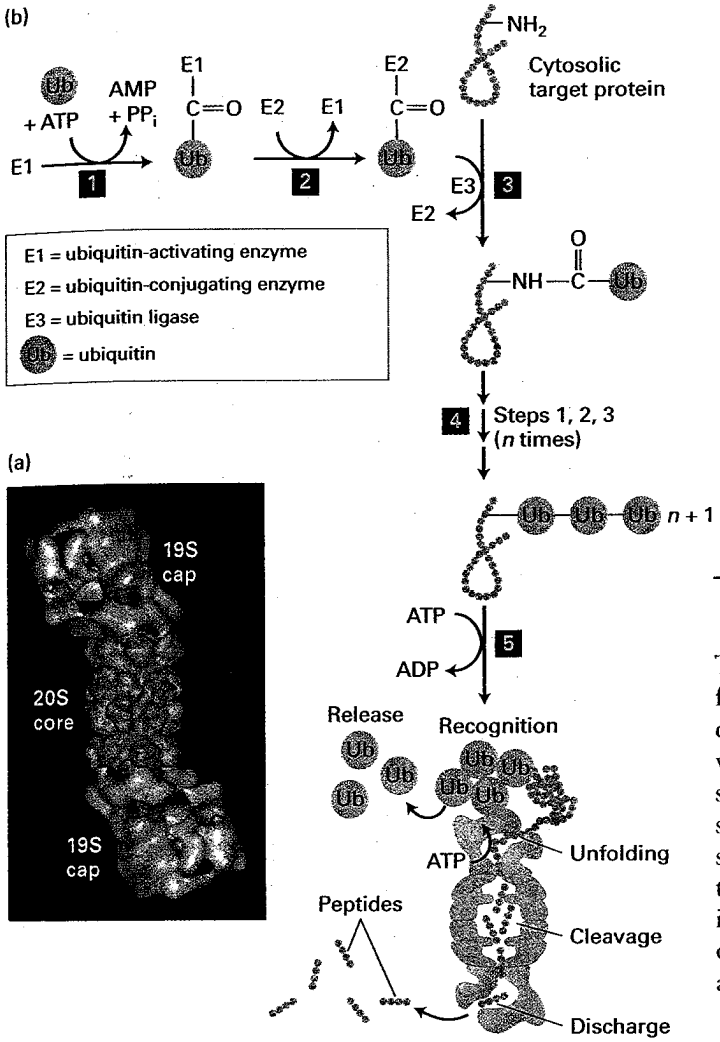
Animation: The Proteasome



◀ **FIGURE 3-29 Ubiquitin- and proteasome-mediated proteolysis.** (a) Computer-generated image reveals that a proteasome has a cylindrical structure with a 19S cap (blue) at each end of a 20S core. Proteolysis of ubiquitin-tagged proteins occurs within the inner chamber of the core. (b) Proteins are targeted for proteasomal degradation by polyubiquitination. Enzyme E1 is activated by attachment of a ubiquitin (Ub) molecule (step 1) and then transfers this Ub molecule to a cysteine residue in E2 (step 2). Ubiquitin ligase (E3) transfers the bound Ub molecule on E2 to the side-chain $—NH_2$ of a lysine residue in a target protein (step 3). Additional Ub molecules are added to the target protein by repeating steps 1–3, forming a polyubiquitin chain (step 4). The polyubiquitinylated target is recognized by the proteasome cap, which uses ATP hydrolysis to drive removal of the Ub groups, unfolding, and transfer of the unfolded protein into the proteolysis chamber in the core, from which the short peptide digestion fragments are later released (step 5) [Part (a) from W. Baumeister et al., 1998, *Cell* **92**:357; courtesy of W. Baumeister.]

called autophagy (see Figure 9-2)—and toward extracellular proteins taken up by the cell. Lysosomes will be discussed at length in later chapters. Here we will focus on cytoplasmic protein degradation by proteasomes.

The Proteasome Is a Complex Molecular Machine Used to Degrade Proteins

Proteasomes are very large macromolecular machines consisting of ≈50 protein subunits and having a mass of 2–2.4 × 10^6 Da. They have a cylindrical, barrel-like catalytic core called the *20S proteasome* (where *S* is a Svedberg unit based on the sedimentation properties of the particle and is proportional to its size). Bound to the ends of this core are one or two cap complexes that regulate proteasomal activity. There are approximately 30,000 proteasomes in a typical mammalian cell. There are multiple forms of proteasomes. The best studied of these is the 26S proteasome (Figure 3-29a), which has a catalytic core approximately 14.8 nm tall and 11.3 nm in diameter and a 19S cap regulatory particle at each end.

There are several distinct regulatory cap complexes with different activities. The 19S cap has 16–18 protein subunits, 6 of which can hydrolyze ATP (i.e., they are ATPases) to provide the energy needed to unfold protein substrates and selectively transfer them into the inner chamber of the proteasome. Genetic studies in yeast have shown that cells cannot survive without functional proteasomes, thus demonstrating their importance. Furthermore, proper proteasomal activity is so important that cells will expend as much as 30 percent of the energy needed to synthesize a protein to degrade it in a proteasome.

The proteasomal catalytic core comprises two inner rings, with six proteolytic active sites facing toward the inner chamber of the ≈1.7-nm-diameter barrel, and two outer rings that control substrate access. Proteasomes can degrade most proteins thoroughly because they have active sites (two each) that cleave after hydrophobic residues, acidic residues, and basic residues. Polypeptide substrates must enter the chamber via a regulated aperture at the center of the outer rings. In the 26S proteasome, the opening of the aperture, which is narrow and often allows the entry of only unfolded proteins, is controlled by ATPases in the 19S cap. The short peptide products of proteasomal digestion (2–24 residues long) exit the chamber and are further degraded rapidly by cytosolic peptidases, eventually being converted to individual amino acids. Some have quipped that a proteasome is a "cellular chamber of doom" in which proteins suffer a "death of a thousand cuts."

Inhibitors of proteasome function can be used therapeutically. Because of the global importance of proteasomes for cells, continuous, complete inhibition of proteasomes kills cells. However, partial, discontinuous proteasome inhibition has been introduced as an approach to cancer chemotherapy. To survive and grow, cells normally

require the robust activity of a regulatory protein called *$NF_\kappa B$*, as well as other similar "pro-survival" proteins. In turn, $NF_\kappa B$ can function fully and promote survival only when its inhibitor, $I_\kappa B$, is disengaged and degraded by proteasomes (Chapter 16). Partial inhibition of proteasomal activity by a small-molecule inhibitor drug results in increased levels of $I_\kappa B$ and, consequently, reduced $NF_\kappa B$ activity (loss of pro-survival activity). Cells subsequently die by a mechanism called **apoptosis** (programmed cell death, Chapter 21). Because at least some types of tumor cells are more sensitive to being killed by proteasome inhibitors than normal cells are, *controlled* administration of proteasome inhibitors (at levels that kill the cancer cells but not normal cells) has proved to be an effective therapy for at least one type of lethal cancer, multiple myeloma. ▮

Ubiquitin Marks Cytosolic Proteins for Degradation in Proteasomes

If proteasomes are to rapidly degrade only those proteins that are either defective or scheduled to be removed, they must be able to distinguish between those proteins that need to be degraded from most of the proteins that don't. To solve this problem, cells identify proteins that should be degraded by covalently attaching multiple copies of a 76-residue polypeptide called **ubiquitin** that is highly conserved from yeast to humans. A complex sensing system has evolved to determine which proteins are to be degraded, and then a three-step process is used to polyubiquitinylate the target proteins. The 19S regulatory cap of the 26S proteasome then recognizes the ubiquitin-labeled proteins, and unfolds and transports them into the proteasome for degradation. The ubiquitination process (Figure 3-29b) involves:

1. Activation of *ubiquitin-activating enzyme (E1)* by the addition of a ubiquitin molecule, a reaction that requires ATP

2. Transfer of this ubiquitin molecule to a cysteine residue in *ubiquitin-conjugating enzyme (E2)*

3. Formation of an isopeptide bond between the carboxyl terminus of the ubiquitin bound to E2 and the amino group of the side chain of a lysine residue in the target protein, a reaction catalyzed by *ubiquitin-protein ligase (E3)*. Subsequent ligase reactions covalently attach additional ubiquitins to the side chain of lysine 48 of the previously added ubiquitin to generate a linear polymer of ubiquitins, or a polyubiquitin-modified target protein.

Specificity of Degradation Targeting of specific proteins is primarily achieved through the substrate specificity of the E3 ligase. There are hundreds of E3 ligases in mammalian cells, ensuring that the wide variety of proteins to be polyubiquitinylated can be modified when necessary.

An example of the control of the activity of a key cellular protein by the ubiquitin-proteasome system is the regulated degradation of proteins called **cyclins**, which control the cell cycle (Chapter 20). Cyclins contain the internal sequence Arg-X-X-Leu-Gly-X-Ile-Gly-Asp/Asn (X can be any amino acid), which is recognized by specific ubiquitinylating enzyme complexes. At a specific time in the cell cycle, each cyclin is phosphorylated by a cyclin kinase. This phosphorylation is thought to cause a conformational change that exposes the recognition sequence to the ubiquitinylating enzymes, leading to polyubiquitination and proteasomal degradation.

Multifunctional Ubiquitin Tagging Some ubiquitination performs cell functions other than the degradation of a targeted protein. Examples of alternative ubiquitination schemes include (1) the covalent addition of a single ubiquitin molecule (monoubiquitination) to a lysine on a target protein, (2) the addition of multiple single ubiquitins (multiubiquitination), (3) linking the ubiquitin to the N-terminus of the target protein, and (4) polyubiquitination in which the ubiquitins are linked to one another via their Lys-63 residue instead of at the Lys-48 position. These modifications can influence the trafficking (sorting) of proteins within a cell (e.g., internalization from the cell surface), control DNA repair and regulation of transcription, and undoubtedly perform numerous other functions yet to be discovered. Cells also have a variety of deubiquitinylating enzymes that can remove ubiquitins from the target proteins and thus introduce the possibility in some cases of reversing the regulation caused by the initial ubiquitination.

KEY CONCEPTS OF SECTION 3.4

Regulating Protein Function I: Protein Degradation

- Proteins may be regulated at the level of protein synthesis, protein degradation, or the intrinsic activity of proteins through noncovalent or covalent interactions.
- The life span of intracellular proteins is largely determined by their susceptibility to proteolytic degradation.
- Many proteins are marked for destruction with a polyubiquitin tag and then degraded within proteasomes, large cylindrical complexes with multiple proteases in their interiors (see Figure 3-29).
- Variations in the nature of the covalent attachment of ubiquitin to proteins are involved in cellular functions other than proteasome-mediated degradation, such as changes in the location or activity of proteins.

3.5 Regulating Protein Function II: Noncovalent and Covalent Modifications

The intrinsic activities of proteins are modulated by both noncovalent and covalent changes in the protein. Noncovalent modifications usually involve the binding or dissociation of a molecule and a consequent change in the conformation of the protein. Often, in such cases, protein activation involves the release or rearrangement of an inhibitory subunit or domain.

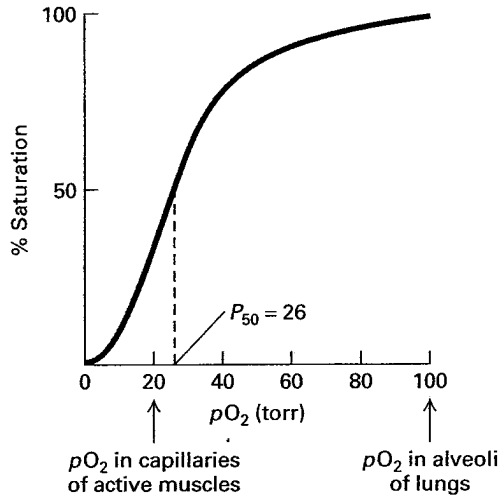
Covalent modifications include hydrolysis of the polypeptide chain or addition of a molecule to the side chain of one or more residues or to the N- or C-terminus of the protein. Such modifications can cause a conformational change in the protein that can alter its activity (form is function). Covalent modifications can also modify the shape of a protein without changing the conformation of the polypeptide and its side chains, for example, by adding a charge or bulky group that can alter the ability of the protein to bind to other molecules. Lastly, covalent modifications can direct the protein to particular locations in a cell (e.g., the cytoplasmic surface of the plasma membrane).

Many noncovalent and covalent modifications are reversible, thus allowing the activity of an individual protein to be enhanced or suppressed multiple times during the lifetime of the protein. Others, such as proteolysis, are irreversible and can be superseded only by degradation of the modified protein and synthesis of a replacement. In the case of enzymes, these regulatory modifications alter K_m, V_{max}, or both. Nature has devised many different strategies for noncovalent and covalent regulation of activity. Here we discuss some common mechanisms for regulating protein function; additional examples will be described in other chapters.

Noncovalent Binding Permits Allosteric, or Cooperative, Regulation of Proteins

One of the most important mechanisms for regulating protein function is through allosteric interactions. Broadly speaking, **allostery** (from the Greek "other shape") refers to any change in a protein's tertiary or quaternary structure, or in both, induced by the noncovalent binding of a ligand. When a ligand binds to one site (A) in a protein and induces a conformational change and associated change in activity of a different site (B), the ligand is called an *allosteric effector* of the protein, while site A is called an *allosteric binding site,* and the protein is called an *allosteric protein.* By definition, allosteric proteins have multiple binding sites for either a single type of ligand or for multiple different ligands. The allosteric change in activity can be positive or negative, i.e., can induce an increase or a decrease in protein activity. Allosteric regulation is particularly prevalent in multimeric enzymes and other proteins where conformational changes in one subunit are transmitted to an adjacent subunit. *Cooperativity* is a term often used synonymously with allostery, and usually refers to the influence (positive or negative) that the binding of a ligand at one site has on the binding of another molecule of the same type of ligand at a different site.

Hemoglobin presents a classic example of positive cooperative binding in that the binding of a single ligand, oxygen, increases the affinity of the binding of the next oxygen molecule. Each of the four subunits in hemoglobin contains one heme molecule. The heme groups are the oxygen-binding components of hemoglobin (see Figure 3-13). The binding of oxygen to the heme molecule in one of the four hemoglobin subunits induces a local conformational change whose effect spreads to the other subunits, lowering the K_m (increasing the affinity) for the binding of additional oxygen molecules to the remaining hemes and yielding a sigmoidal oxygen-binding curve (Figure 3-30). Because of the sigmoidal shape of the oxygen-saturation curve, it takes only a fourfold increase in oxygen concentration for the percent saturation of the oxygen binding sites in hemoglobin to go from 10 to 90 percent. Conversely, if there were no cooperativity and the shape of the curve was typical of that for Michaelis-Menten-type binding, it would take an 81-fold increase in oxygen concentration to accomplish the same increase in loading. This cooperativity permits hemoglobin to take up oxygen very efficiently in the lungs where the oxygen concentration is high, and unload it in tissues where the concentration is low. Thus, cooperativity amplifies the sensitivity of a system to concentration changes in its ligands, providing in many cases selective evolutionary advantage.



▲ **EXPERIMENTAL FIGURE 3-30 Hemoglobin binds oxygen cooperatively.** Each tetrameric hemoglobin protein has four oxygen-binding sites; at saturation all the sites are loaded with oxygen. The oxygen concentration is commonly measured as the partial pressure (pO_2). P_{50} is the pO_2 at which half the oxygen-binding sites at a given hemoglobin concentration are occupied; it is somewhat analogous to the K_m for an enzymatic reaction. The large change in the amount of oxygen bound over a small range of pO_2 values permits efficient unloading of oxygen in peripheral tissues such as muscle. The sigmoidal shape of a plot of percent saturation versus ligand concentration is indicative of cooperative binding. In the absence of cooperative binding, a binding curve is a hyperbola, similar to the curves in Figure 3-22. [Adapted from L. Stryer, 1995, *Biochemistry,* 4th ed., W. H. Freeman and Company.]

Negative cooperativity often involves the end product of a multistep biochemical pathway, which binds to and reduces the activity of an enzyme that catalyzes an early, rate-controlling step for that pathway. In this way excessive buildup of the product is prevented. This kind of regulation of a metabolic pathway is also called *end-product inhibition* or *feedback inhibition.*

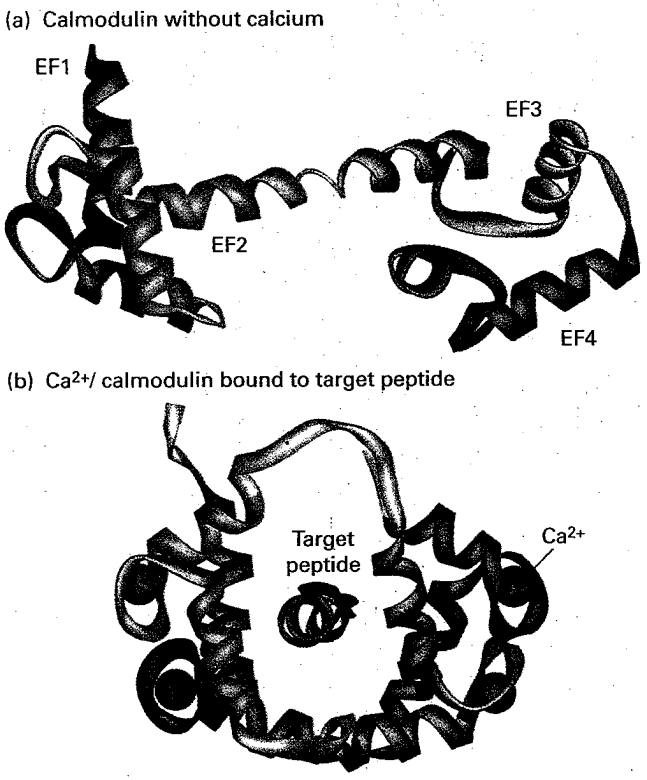
Noncovalent Binding of Calcium and GTP Are Widely Used As Allosteric Switches to Control Protein Activity

Unlike oxygen, which causes graded allosteric changes in the activity of hemoglobin, other allosteric effectors act as switches, turning the activity of many different proteins on or off. Two important allosteric switches that we will encounter many times throughout this book are Ca^{2+} and GTP.

Ca^{2+}/Calmodulin-Mediated Switching The concentration of Ca^{2+} free in the cytosol (not bound to molecules other than water) is kept very low ($\approx 10^{-7}$ M) by specialized membrane transport proteins that continually pump Ca^{2+} out of the cytosol. However, as we learn in Chapter 11, the cytosolic Ca^{2+} concentration can increase from 10- to 100-fold when Ca^{2+}-permeable channels in the cell surface membranes open and allow extracellular Ca^{2+} to flow into the cell. This rise in cytosolic Ca^{2+} is sensed by specialized Ca^{2+}-binding proteins, which alter cellular behavior by turning other proteins on or off. The importance of extracellular Ca^{2+} for cell activity was first documented by S. Ringer in 1883, when he discovered that isolated rat hearts suspended in a NaCl solution made with 'hard' (Ca^{2+}-rich) London tap water contracted beautifully, whereas they beat poorly and stopped quickly if distilled water was used.

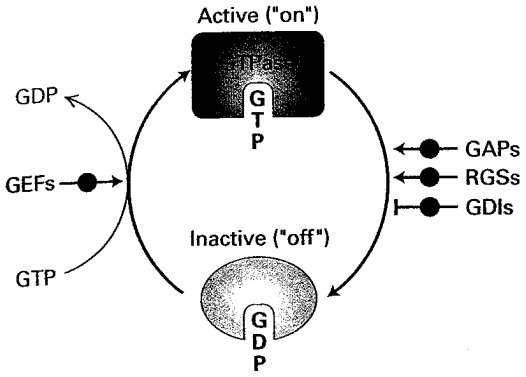
Many of the Ca^{2+}-binding proteins bind Ca^{2+} using the EF hand/helix-loop-helix structural motif discussed earlier (see Figure 3-9b). The prototype EF hand protein, **calmodulin,** is found in all eukaryotic cells and may exist as an individual monomeric protein or as a subunit of a multimeric protein. A dumbbell-shaped molecule, calmodulin contains four Ca^{2+}-binding EF hands with K_d's of $\approx 10^{-6}$ M. The binding of Ca^{2+} to calmodulin causes a conformational change that permits Ca^{2+}/calmodulin to bind to conserved sequences in various target proteins, thereby switching their activities on or off (Figure 3-31). Calmodulin and similar EF hand proteins thus function as *switch proteins,* acting in concert with changes in Ca^{2+} levels to modulate the activity of other proteins.

Switching Mediated by Guanine Nucleotide–Binding Proteins Another group of intracellular switch proteins constitutes the **GTPase superfamily.** As the name suggests, these proteins are enzymes, GTPases, that can hydrolyze GTP (guanosine triphosphate) to GDP (guanosine diphosphate). They include the monomeric Ras protein (see Figure 3-8) and the G_α subunit of the trimeric G proteins, both discussed at length in Chapter 15. Both Ras and G_α can bind to the plasma membrane, function in cell signaling, and play a key role in cell proliferation and differentiation. Other members of the GTPase superfamily function in protein synthesis, the transport of proteins between the nucleus and the cytoplasm, the formation of coated vesicles and their fusion with target membranes, and rearrangements of the actin cytoskeleton. The Hsp70 chaperone protein we encountered earlier is an example of an ATP/ADP switch, similar in many respects to a GTP/GDP switch.



▲ **FIGURE 3-31 Conformational changes induced by Ca^{2+} binding to calmodulin.** Calmodulin is a widely distributed cytosolic protein that contains four Ca^{2+}-binding sites, one in each of its EF hands. Each EF hand has a helix-loop-helix motif. At cytosolic Ca^{2+} concentrations above about 5×10^{-7} M, binding of Ca^{2+} to calmodulin changes the protein's conformation from the dumbbell-shaped, unbound form (a) to one in which hydrophobic side chains become more exposed to solvent. The resulting Ca^{2+}/calmodulin can wrap around exposed helices of various target proteins (b), thereby altering their activity.

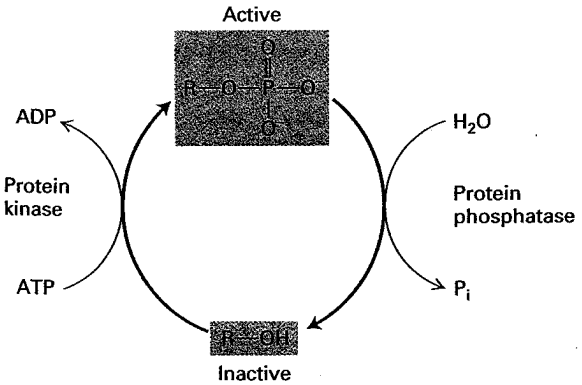
All the GTPase switch proteins exist in two forms, or conformations (Figure 3-32): (1) an active ("on") form with bound GTP that modulates the activity of specific target proteins to which they bind and (2) an inactive ("off") form with bound GDP, which is generated by the relatively slow hydrolysis of the GTP bound to the active form. The amount of time any given GTPase switch remains active depends on the rate of its GTPase activity. Thus the GTPase activity acts as a timer to control this switch. Cells contain a variety of proteins that can modulate the baseline (or intrinsic) rate of GTPase activity for any given GTPase switch. For example, GTPase activity can be enhanced by specific GTPase-activating proteins, called *GAPs,* or depressed by other proteins acting as allosteric regulators. After the switch has been turned off (GTP hydrolysis), it can be turned back on by GTP exchange factor (GEF), which replaces the bound GDP with a different GTP molecule from the surrounding fluid. Thus, cells can control when the switch is turned on and how long the switch remains on. We examine the role of various GTPase switch proteins in regulating intracellular signaling and other processes in several later chapters.



▲ **FIGURE 3-32 The GTPase switch.** Conversion of the active, GTP-bound GTPase into the inactive form by hydrolysis of GTP is accelerated by GAPs (GTPase-activating proteins) and RGSs (regulators of G protein signaling) and inhibited by GDIs (guanine nucleotide dissociation inhibitors). Reactivation by replacing GDP with GTP is promoted by GEFs (guanine nucleotide exchange factors).

Phosphorylation and Dephosphorylation Covalently Regulate Protein Activity

One of the most common mechanisms for regulating protein activity is *phosphorylation,* the addition of phosphate groups to hydroxyl groups on serine, threonine, or tyrosine residues. Protein **kinases** catalyze phosphorylation, and **phosphatases** catalyze *dephosphorylation.* The counteracting activities of kinases and phosphatases provide cells with a "switch" that can turn on or turn off the function of various proteins (Figure 3-33). Phosphorylation changes a protein's charge and generally leads to a conformational change; these effects can significantly alter ligand binding or other features of the protein, leading to an increase or decrease in its activity.



▲ **FIGURE 3-33 Regulation of protein activity by the kinase/phosphatase switch.** The cyclic phosphorylation and dephosphorylation of a protein is a common cellular mechanism for regulating protein activity. In this example, the target protein R is active (top) when phosphorylated and inactive (bottom) when dephosphorylated; some proteins have the opposite responses to phosphorylation.

Nearly 3 percent of all yeast proteins are protein kinases or phosphatases, indicating the importance of phosphorylation and dephosphorylation reactions even in simple cells. All classes of proteins—including structural proteins, scaffolds, enzymes, membrane channels, and signaling molecules—have members regulated by kinase/phosphatase switches. Different protein kinases and phosphatases are specific for different target proteins and can thus regulate a variety of cellular pathways, as discussed in later chapters. Some of these enzymes act on one or a few target proteins, whereas others have many targets. The latter are useful in integrating the activities of proteins that are coordinately controlled by a single kinase/phosphatase switch. Frequently, the target of the kinase (and phosphatase) is yet another kinase or phosphatase, creating a cascade effect. There are many examples of such kinase cascades, which permit amplification of a signal and many levels of fine-tuning control (see Chapter 15).

Proteolytic Cleavage Irreversibly Activates or Inactivates Some Proteins

Unlike phosphorylation, which is reversible, the activation or inactivation of protein function by proteolytic cleavage is an irreversible mechanism for regulating protein activity. For example, many polypeptide hormones, such as insulin, are synthesized as long precursors, and prior to secretion from cells some of their peptide bonds must be hydrolyzed for them to fold properly. In some cases, a single long precursor *prohormone* polypeptide can be cleaved into several distinct active hormones. To prevent the pancreatic serine proteases from inappropriately digesting proteins before they reach the small intestines, they are synthesized as *zymogens,* inactive precursor proteins. Cleavage of a peptide bond near the N-terminus of trypsinogen (the zymogen of trypsin) by a highly specific protease in the small intestine generates a new N-terminal residue (Ile-16), whose amino group can form an ionic bond with the carboxylic acid side chain of an internal aspartic acid. This causes a conformation change that opens the substrate-binding site, activating the enzyme. The active trypsin can then activate trypsinogen, chymotrypsinogen, and other zymogens. Similar, but more elaborate, protease cascades (one protease activating inactive precursors of others) that can amplify an initial signal play important roles in several systems, such as the blood-clotting cascade. The importance of carefully regulating such systems is clear—inappropriate clotting could fatally clog the circulatory system, while insufficient clotting could lead to uncontrolled bleeding.

An unusual and rare type of proteolytic processing, termed *protein self-splicing,* takes place in bacteria and some eukaryotes. This process is analogous to editing film: an internal segment of a polypeptide is removed and the ends of the polypeptide are rejoined (ligated). Unlike other forms of proteolytic processing, protein self-splicing is an autocatalytic process, which proceeds by itself without the participation of enzymes. The excised peptide appears to eliminate itself from the protein by a mechanism similar to that used in

the processing of some RNA molecules (Chapter 8). In vertebrate cells, the processing of some proteins includes self-cleavage, but the subsequent ligation step is absent. One such protein is Hedgehog, a membrane-bound signaling molecule that is critical to a number of developmental processes (Chapter 16).

Higher-Order Regulation Includes Control of Protein Location and Concentration

All the regulatory mechanisms heretofore described affect a protein locally at its site of action, turning its activity on or off. Normal functioning of a cell, however, also requires the segregation of proteins to particular compartments such as the mitochondria, nucleus, and lysosomes. In regard to enzymes, compartmentation not only provides an opportunity for controlling the delivery of substrate or the exit of product, but also permits competing reactions to take place simultaneously in different parts of a cell. We describe the mechanisms that cells use to direct various proteins to different compartments in Chapters 12 and 13.

KEY CONCEPTS OF SECTION 3.5

Protein Regulation II: Noncovalent and Covalent Modifications

- In allostery, the noncovalent binding of one ligand molecule, the allosteric effector, induces a conformational change that alters a protein's activity or affinity for other ligands. The allosteric effector can be identical in structure to or different from the other ligands, whose binding it affects. The allosteric effector can be a substrate, activator, or inhibitor.
- In multimeric proteins, such as hemoglobin, that bind multiple identical ligand molecules (e.g., oxygen), the binding of one ligand molecule may increase or decrease the binding affinity for subsequent ligand molecules. This type of allostery is known as cooperativity.
- Several allosteric mechanisms act as switches, turning protein activity on and off in a reversible fashion.
- Two classes of intracellular switch proteins regulate a variety of cellular processes: (1) Ca^{2+}-binding proteins (e.g., calmodulin) and (2) members of the GTPase superfamily (e.g., Ras), which cycle between active GTP-bound and inactive GDP-bound forms (see Figure 3-32).
- The phosphorylation and dephosphorylation of hydroxyl groups on serine, threonine, or tyrosine residue side chains by protein kinases and phosphatases provide reversible on/off regulation of numerous proteins.
- Many types of covalent and noncovalent regulation are reversible, but some forms of regulation, like proteolytic cleavage, are irreversible.
- Higher-order regulation includes compartmentation of proteins and control of protein concentration.

3.6 Purifying, Detecting, and Characterizing Proteins

A protein often must be purified before its structure and the mechanism of its action can be studied in detail. 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 particular type of cell is a daunting task that requires methods both for separating proteins and for detecting the presence of specific proteins.

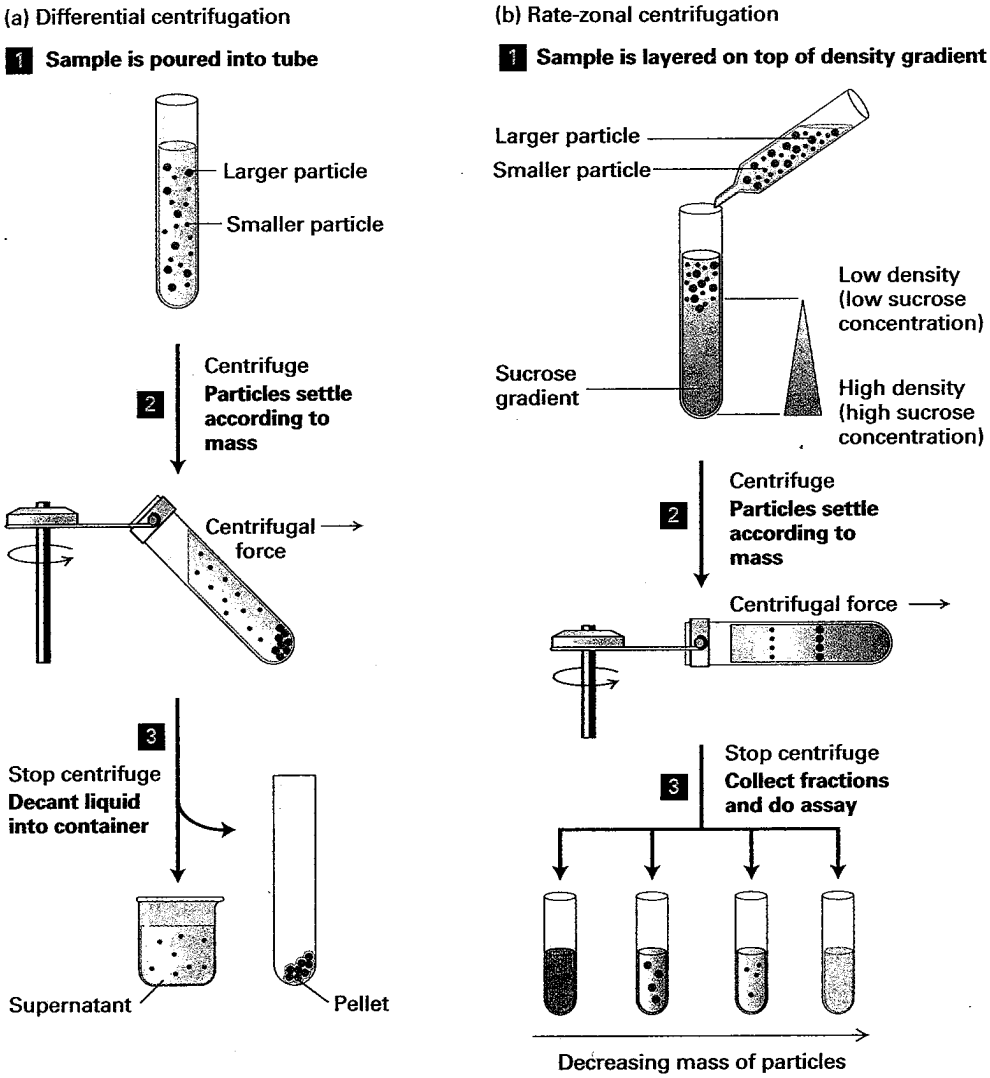
Any molecule, whether protein, carbohydrate, or nucleic acid, can be separated, or *resolved,* from other molecules on the basis of their differences in one or more physical or chemical characteristics. The larger and more numerous the differences between two proteins, the easier and more efficient their separation. The two most widely used characteristics for separating proteins are *size,* defined as either length or mass, and *binding affinity* for specific ligands. In this section, we briefly outline several important techniques for separating proteins; these separation techniques are also useful for the separation of nucleic acids and other biomolecules. (Specialized methods for removing membrane proteins from membranes are described in Chapter 10 after the unique properties of these proteins are discussed.) We then consider the use of radioactive compounds for tracking biological activity. Finally, we consider several techniques for characterizing a protein's mass, sequence, and three-dimensional structure.

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, cell fragments, organelles, or molecules) with 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. Unless a protein has an attached lipid or carbohydrate, its density will not vary by more than 15 percent from 1.37 g/cm^3, the average protein density. Heavier or more dense molecules settle, or sediment, more quickly than lighter or less dense molecules.

A centrifuge speeds sedimentation by subjecting particles in suspension to centrifugal forces as great as 1,000,000 times the force of gravity *g,* which can sediment particles as small as 10 kDa. Modern ultracentrifuges achieve these forces by reaching speeds of 150,000 revolutions per minute (rpm) or greater. However, small particles with masses of 5 kDa or less will not sediment uniformly even at such high 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

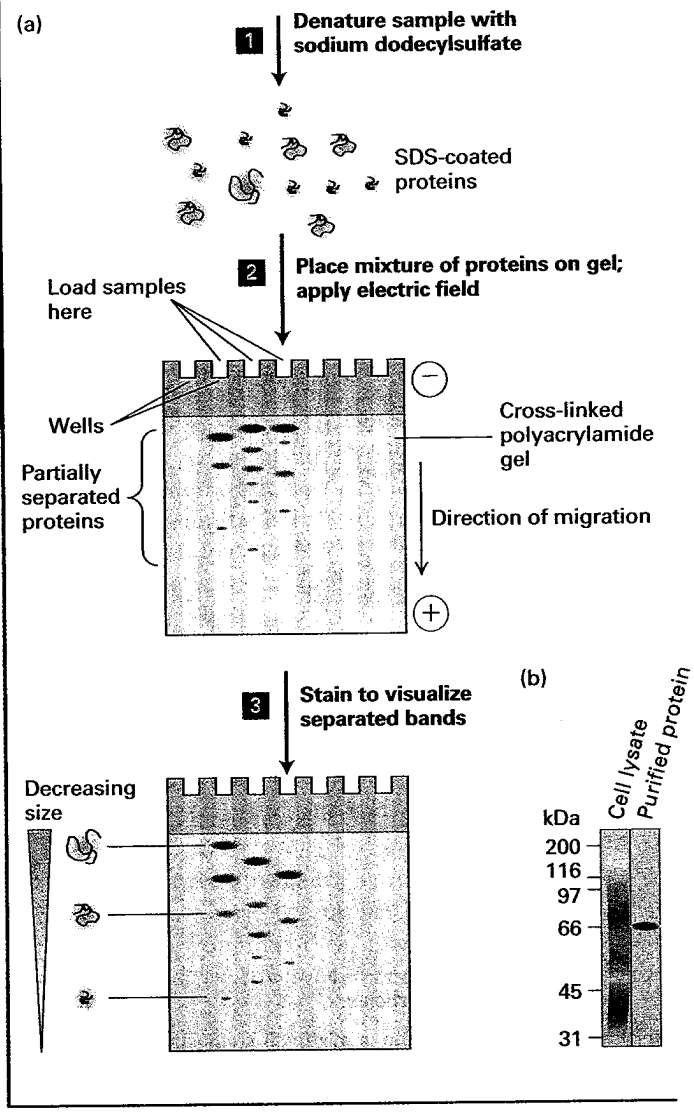


▲ **EXPERIMENTAL FIGURE 3-34 Centrifugation techniques separate particles that differ in mass or density.** (a) In differential centrifugation, a cell homogenate or other mixture is spun long enough to sediment the larger particles (e.g., cell organelles, cells), which collect as a pellet at the bottom of the tube (step 2). The smaller particles (e.g., soluble proteins, nucleic acids) remain in the liquid supernatant, which can be transferred to another tube (step 3). (b) In rate-zonal centrifugation, a mixture is spun (step 2) just long enough to separate molecules that differ in mass but may be similar in shape and density (e.g., globular proteins, RNA molecules) into discrete zones within a density gradient commonly formed by a concentrated sucrose solution. Fractions are removed from the bottom of the tube and subjected to testing (assayed).

constant *s* of a protein is a measure of its sedimentation rate. The sedimentation constant is commonly expressed in svedbergs (S), where a typical, large protein complex is about 3–5S, while a eukaryotic ribosome is 80S.

Differential Centrifugation The most common initial step in protein purification for cells or tissues is the separation of water-soluble proteins from insoluble cellular material by *differential centrifugation*. A starting mixture, commonly a cell homogenate (mechanically broken calls), is poured into a tube and spun at a rotor speed and for a period of time that forces cell organelles such as nuclei and large unbroken cells or large cell fragments to collect as a pellet at the bottom; the soluble proteins remain in the supernatant (Figure 3-34a). The supernatant fraction then is poured off, and either it or the pellet can be subjected to other purification methods to separate the many different proteins that they contain.

Rate-Zonal Centrifugation On the basis of differences in their masses, proteins can be separated by centrifugation through a solution of increasing density called a *density gradient*. A concentrated sucrose solution is commonly used to form density gradients. When a protein mixture is layered on top of a sucrose gradient in a tube and subjected to centrifugation, each protein in the mixture migrates down the tube at a rate controlled by the factors that affect the sedimentation constant. All the proteins start from a thin zone at the top of the tube and separate into bands, or zones (actually, disks), of proteins of



◀ **EXPERIMENTAL FIGURE 3-35 SDS-polyacrylamide gel electrophoresis (SDS-PAGE) separates proteins primarily on the basis of their masses.** (a) Initial treatment with SDS, a negatively charged detergent, dissociates multimeric proteins and denatures all the polypeptide chains (step 1). During electrophoresis, the SDS-protein complexes migrate through the polyacrylamide gel (step 2). Small complexes are able to move through the pores faster than larger ones. Thus the proteins separate into bands according to their sizes as they migrate. The separated protein bands are visualized by staining with a dye (step 3). (b) Example of SDS-PAGE separation of all the proteins in a whole-cell lysate (detergent solublized cells): (*left*) the many separate stained proteins, appearing almost as a continuum; (*right*) a protein purified from the lysate by a single step of antibody-affinity chromatography. The proteins were visualized by staining with a silver-based dye. [Part (b) modified from B. Liu and M. Krieger, 2002, *J. Biol. Chem.* **277**(37):34125–34135.]

different masses. In this separation technique, called *rate-zonal centrifugation*, samples are centrifuged just long enough to separate the molecules of interest into discrete zones (Figure 3-34b). If a sample is centrifuged for too short a time, the different protein molecules will not separate sufficiently. If a sample is centrifuged much longer than necessary, all the proteins 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 or other molecules, such as 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, lipoproteins that carry lipids through the circulatory system, or organelles (see Figure 9-26).

Electrophoresis Separates Molecules on the Basis of Their Charge-to-Mass Ratio

Electrophoresis is a technique for separating molecules in a mixture under the influence of an applied electric field and is one of the most frequently used techniques to study proteins and nucleic acids. Dissolved molecules in an electric field move, or migrate, at a speed determined by their charge-to-mass (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 of the opposite polarity.

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 similar to the congealed gelatin found in desserts) rather than in a liquid solution. Electrophoretic separation of proteins is most commonly performed in polyacrylamide gels. When a mixture of proteins is placed in a gel and an electric current is applied, smaller proteins migrate faster through the gel than do larger proteins because the gel acts as a sieve, with smaller species able to maneuver more rapidly through the pores in the gel than larger species. The shape of a molecule can also influence its rate of migration (long asymmetric molecules migrate more slowly than spherical ones of the same mass).

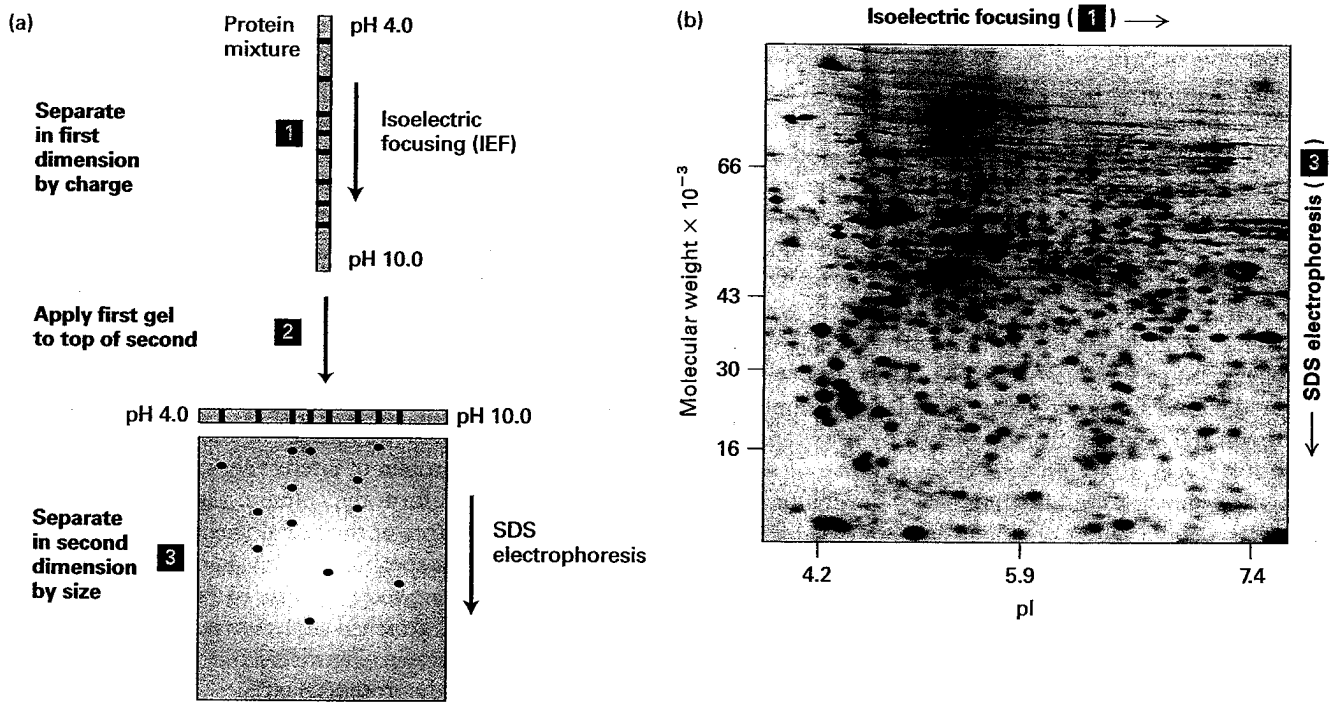
Gels are cast between a pair of glass plates by polymerizing a solution of acrylamide monomers into polyacrylamide chains and simultaneously cross-linking the chains into a semisolid matrix. The pore size of a gel can be varied by adjusting the concentrations of polyacrylamide and the cross-linking reagent. The rate at which a protein moves through a gel is influenced by the gel's pore size and the strength of the electric field. By suitable adjustment of these parameters,

proteins of widely varying sizes can be resolved (separated from one another) by polyacrylamide gel electrophoresis (PAGE).

In the most powerful technique for resolving protein mixtures, proteins are exposed to the ionic detergent SDS (sodium dodecylsulfate) before and during gel electrophoresis (Figure 3-35). SDS denatures proteins, in part because it binds to hydrophobic side chains, destabilizing the hydrophobic interactions in the core of a protein that contribute to its stable conformation. (SDS treatment is usually combined with heating in the presence of reducing agents that break disulfide bonds.) As a consequence, multimeric proteins 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 in native structures; therefore, chain length, which corresponds to mass, is the principal determinant of the migration rate of proteins in *SDS-polyacrylamide electrophoresis (SDS-PAGE)*. Even chains that differ in molecular weight by less than 10 percent can be resolved by this technique. Moreover, the molecular weight of a protein can be estimated by comparing the distance that it migrates through a gel with the distances that proteins of known molecular weight migrate (there is roughly a linear relationship between migration distance and the log of the molecular weight). Proteins within the gels can be extracted for further analysis (e.g., identification by methods described below).

Two-Dimensional Gel Electrophoresis Electrophoresis of all cellular proteins by SDS-PAGE can separate proteins having relatively large differences in mass but cannot readily resolve proteins having similar masses (e.g., a 41-kDa protein versus a 42-kDa protein). To separate proteins of similar masses, another physical characteristic must be exploited. Most commonly, this characteristic is electric charge, which is determined by the pH and the relative number of the protein's positively and negatively charged groups, which is in turn dependent on the pK_a's of the ionizable groups (see Chapter 2). Two unrelated proteins having similar masses are unlikely to have identical net charges because their sequences, and thus the number of acidic and basic residues, are different.

In two-dimensional electrophoresis, proteins are separated sequentially, first by their charges and then by their masses (Figure 3-36a). In the first step, a cell or tissue extract is fully denatured by high concentrations (8 M) of urea and then layered on a gel strip that contains a continuous pH gradient. SDS cannot be used, because its binding changes the charge of the protein. The gradient is formed by ampholytes, a mixture of polyanionic and polycationic molecules, that are cast into the gel, with the most acidic ampholyte at one end and the most basic ampholyte at the



▲ **EXPERIMENTAL FIGURE 3-36 Two-dimensional gel electrophoresis separates proteins on the basis of charge and mass.** (a) In this technique, proteins are first separated into bands on the basis of their charges by isoelectric focusing (step 1). The resulting gel strip is applied to an SDS-polyacrylamide gel (step 2), and the proteins are separated into spots by mass (step 3). (b) In this two-dimensional gel of a protein extract from cultured cells, each spot represents a single polypeptide. Polypeptides can be detected by dyes, as here, or by other techniques such as autoradiography. Each polypeptide is characterized by its isoelectric point (pI) and molecular weight. [Part (b) courtesy of J. Celis.]

opposite end. A charged protein will migrate through the gradient until it reaches its **isoelectric point (pI)**, 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. Proteins that have been separated on an IEF gel can then be separated in a second dimension on the basis of their molecular weights. To accomplish this separation, the IEF gel strip is placed lengthwise on one outside edge of a sheetlike (two-dimensional, or slab) polyacrylamide gel, this time 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 masses.

The sequential resolution of proteins by charge and mass can achieve excellent separation of cellular proteins (Figure 3-36b). For example, two-dimensional gels have been very useful in comparing the proteomes in undifferentiated and differentiated cells or in normal and cancer cells because as many as 1000 proteins can be resolved as individual spots simultaneously. Sophisticated methods have been developed to permit the comparison of complex patterns of proteins in two-dimensional gels from related, but distinct, samples (e.g., tissue from a normal versus a mutant individual) to permit identification of differences in the types or amounts of proteins in the samples (see section on proteomics, below).

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

A third common technique for separating mixtures of proteins or fragments of proteins, as well as other molecules, is based on the principle that molecules dissolved in a solution can differentially interact (bind and dissociate) with a particular solid surface, depending on the physical and chemical properties of the molecule and the 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 flow past the surface more slowly than molecules that interact infrequently with it. In this technique, called **liquid chromatography** (LC), the sample is placed on top of a tightly packed column of spherical beads held within a glass or plastic cylinder. The sample then flows down the column, usually driven by gravitational or hydrostatic forces alone or with the assistance of a pump, and small aliquots of fluid flowing out of the column, called *fractions*, are collected sequentially for subsequent analysis for the presence of the proteins of interest. The nature of the beads in the column determines whether the separation of proteins depends on differences in mass, charge, or binding affinity.

Gel Filtration Chromatography Proteins that differ in mass can be separated on a column composed of porous beads made from polyacrylamide, dextran (a bacterial polysaccharide), or agarose (a seaweed derivative)—a technique called gel filtration chromatography. Although proteins flow around the spherical beads in gel filtration chromatography, they spend some time within the large depressions that cover a bead's surface. Because smaller proteins can penetrate into these depressions more readily than larger proteins can, they travel through a gel filtration column more slowly than larger proteins (Figure 3-37a). (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 (or separate and remove) a protein from a gel filtration column depends on its mass: the smaller the mass, the more time it is trapped on the beads, the greater the elution volume. By use of proteins of known mass as standards to calibrate the column, the elution volume can be used to estimate the mass of a protein in a mixture. A protein's shape as well as its mass can influence the elution volume.

Ion-Exchange Chromatography In ion-exchange chromatography, a second type of liquid chromatography, proteins are separated on the basis of differences in their charges. 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 positively charged (basic) proteins flow unimpeded through the column (Figure 3-37b). The acidic proteins are then eluted selectively from the column by passing a solution of increasing concentrations of salt (a salt gradient) 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 bound proteins, those with relatively low charge, are eluted first and highly charged proteins are eluted last. Similarly, a negatively charged column can be used to retain and fractionate basic (positively charged) proteins.

Affinity Chromatography The ability of proteins to bind specifically to other molecules is the basis of affinity chromatography. In this technique, ligand or other molecules that bind to the protein of interest are covalently attached to the beads used to form the column. Ligands can be enzyme substrates, inhibitors or their analogues, or other small molecules that bind to specific proteins. In a widely used form of this technique—*antibody-affinity*, or *immunoaffinity, chromatography*—the attached molecule is an antibody specific for the desired protein (Figure 3-37c). (We discuss antibodies as tools to study proteins next).

An affinity column in principle will retain only those proteins that bind the molecule attached to the beads; the remaining proteins, regardless of their charges or masses, will pass through the column because they do not bind. However, if a retained protein is in turn bound to other molecules, forming a complex, then the entire complex is