GLOBAL EDITION

Introduction to Biotechnology

FOURTH EDITION

William J. Thieman • Michael A. Palladino

INTRODUCTION TO BIOTECHNOLOGY

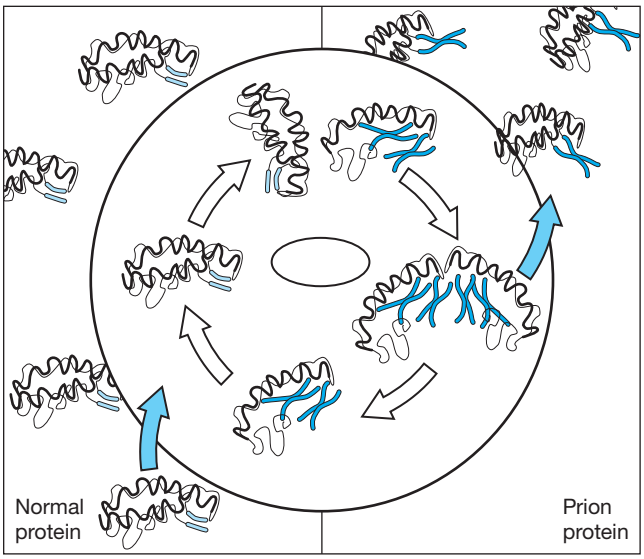

FIGURE 4.5 Prions Are Misfolded Proteins A misfolding of proteins that occurs in prion disorders can be duplicated in the lab to produce large quantities of prion proteins, which can then be studied to create diagnostic kits. Follow the arrows from the left as a prion protein attracts a normal protein and changes it into a prion, resulting in an accumulation of prions within the cell due to the inability of cellular enzymes to degrade the changed structure.

infectious material has a different structure and is resistant to proteases. A model of prion replication must explain both how prions propagate and why their spontaneous appearance is so rare. Manfred Eigen showed the heterodimer model (called PrPSc) to be an extraordinarily effective catalyst, increasing the rate of the conversion reaction of proteins by a factor of about 10^{15}. A gene for the normal protein has been identified: the *PRNP* gene. In all inherited cases of prion disease, there is a mutation in the *PRNP* gene. Many different PRNP mutations have been identified, and these proteins are more likely to fold into abnormal prions.

4.3 Protein Production

By now, two things should be evident: (1) Proteins are valuable, and (2) they are complex, fragile products. With these points in mind, we now examine the work of biotechnology in the production of proteins.

Producing a protein in the lab is a long, painstaking process, and at every stage there are many methods of production from which to choose. We refer to the two major phases used in producing proteins as **upstream processing** and **downstream processing.** Upstream processing includes the actual expression of the protein in the cell. During downstream processing, the protein is first separated from other parts of the cell and isolated from other proteins. Purity and functional abilities are then verified. Finally, a stable means of preserving the protein is developed. The choices made during upstream processing can simplify downstream processing.

Protein Expression: Upstream Processing

We begin a detailed discussion of protein processing by looking at the first decision made in upstream processing: selecting the cell to be used as a protein source. Microorganisms, fungi, plant cells, and animal cells all have unique qualities that make them good choices in the right circumstances.

Bacteria

Bacteria are an attractive protein source for several reasons. First, the fermentation processes of bacteria are well understood. Also, they can be cultured in large quantities in a short time. In industrial applications, this ability to generate the product on a large scale is often essential. Bacteria are also relatively easy to alter genetically.

Several methods of recombinant DNA technology can be used to increase the level of production of a bacterial protein. One is the introduction of additional copies of the relevant gene to the host cell. In most cases the relevant gene introduced into the organism is under the control of expression by a more powerful transcriptional promoter (see Chapter 3). You may wish to watch the animation "The *lac* Operon in *E. coli*" on the Companion Website for a refresher.

Another microbial protein factory is the Gram-positive bacterium *Bacillus subtilis*, an inhabitant of the upper layers of the soil that has the capacity to secrete proteins in the gram per liter range. The engineering of *B. subtilis* into a next-generation super-secreting cell factory has occurred. The analysis of protein secretion in *Bacillus* was initiated in the 1980s with classical molecular genetics approaches, providing valuable insights into cellular components involved in this process. The many applications of *B. subtilis* are related to the high-level secretion of proteins that has focused major research interests on the **secretome**, which includes both the protein secretion machinery and the secreted proteins. A very important outcome of the secretome analyses was the definition of secretion signals, the so-called signal peptides, for all secreted proteins of *B. subtilis*. The efficient production of secreted enzymes by microorganisms like *B. subtilis* is one of the key drivers of today's successes in the protein industry.

In some instances, the target gene (in the form of **cDNA)** for the desired protein product is attached directly to a complete or partial *E. coli* gene. In these

TABLE 4.3 **Advantages and Disadvantages of Recombinant Protein Production in *E. coli***

Advantages	Disadvantages
E. coli genetics are well understood	Foreign proteins produced as inclusion bodies must be refolded
Almost unlimited quantities of proteins can be generated	Proteins cannot be folded in ways needed for many proteins active in mammalian systems
Fermentation technology is well understood	Some proteins are inactive in humans

cases, the genetically engineered *E. coli* produce the desired protein, but it is in the form of a **fusion protein.** In fusion proteins**,** a target protein is fused to a bacterial protein; therefore, an additional step is required to break the two apart. The fused bacterial protein is usually an enzyme that will bind to its substrate and can be attached to a purification column (see description of affinity columns, in the text that follows). The majority of proteins synthesized naturally by *E. coli* are intracellular (within the cell). In most cases, the resultant foreign protein accumulates in the cell's cytoplasm in the form of insoluble clumps called **inclusion bodies,** which must be purified (and usually refolded) from the other cell proteins before they can be used. You may wish to visit "Use of GFP Fusions for Protein Localization" on the Companion Website for an example.

There are some limitations to the use of microorganisms in protein production. All bacteria, including *E. coli,* are prokaryotic. Prokaryotes are unable to carry out certain processes, such as glycosylation. For this reason, some proteins can be produced only by eukaryotic cells, as seen in Table 4.3.

Although it is possible to conduct the entire protein production process in a small flask in the laboratory, genetically engineered microorganisms can also be grown in large-scale fermenters (anaerobic) or bioreactors (aerobic) (Figure 4.6).

Computers monitor the environment in bioreactors, keeping oxygen levels and temperature ideal for cell growth. Cell growth is monitored carefully, because when the phase of growth is highest, the promoter must be activated to stimulate foreign gene expression. Activating a gene in a recombinant organism requires correct timing. It must be done after the organism has completed synthesizing important natural proteins needed for its metabolism.

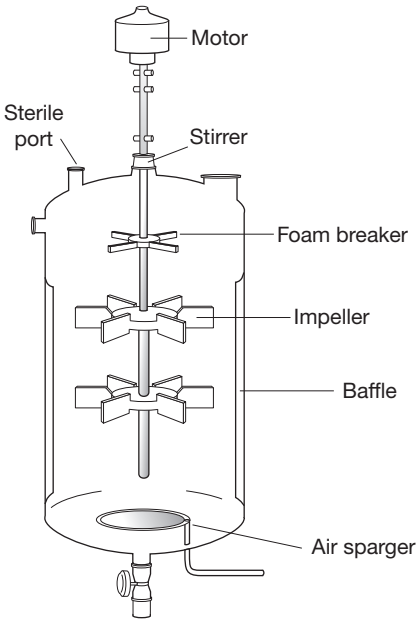

FIGURE 4.6 **Bioreactor** Large-volume cell culture is commonly performed in a self-contained, closed (sterile) system until the products are harvested. Through sterile ports, workers can adjust pH, gas concentrations, and other variables based on input from internal sensors. Sterile bags as large as 2,000 liters can be used on rocking tables as alternatives to stainless steel bioreactors.

Fungi

Fungi are the source of a wide range of proteins used in products as diverse as animal feed and beer. In addition to naturally occurring proteins, many species of fungi are good hosts for engineered proteins. Unlike bacteria, fungi are eukaryotic and capable of some post-translational modification (like folding human proteins correctly and other modifications) and are used for that synthesis, as illustrated in Table 4.4.

Plants

Plant cells can also be used for protein expression. In fact, plants are an abundant source of naturally occurring, biologically active molecules, and 85

TABLE 4.4 **Some Recombinant Proteins from Fungi**

Protein	Fungi
Human interferon	*A. niger*, *A. nidulans*
Human lactoferrin	*A. oryzae*, *A. niger*
Bovine chymosin	*A. niger*, *A. nidulans*
Aspartic proteinase	*A. oryzae*
Triglyceride lipase	*A. oryzae*

percent of all current drugs originated in plants. One example of a plant-derived protein produced on an industrial scale is the **proteolytic** (protein-degrading) enzyme **papain.** Papain, or vegetable pepsin, is a protease used as a meat-tenderizing agent. It digests the collagen present in connective tissue and blood vessels that makes meat tough.

Plants can be genetically modified to produce specific proteins that do not occur naturally. This process encourages rapid growth and reproductive rates in plants, which can be a distinct advantage. For example, tobacco, the first plant to be genetically engineered, can produce a million seeds from a single plant. As a non-food plant, it makes a good choice for biotech protein production. Once the genetic material is integrated, a million new "plant protein factories" can fill the fields.

There are also disadvantages to using plants as protein producers. Not all proteins can be expressed in plants, and, because they have tough cell walls, the process of extracting proteins from them can be time-consuming and difficult. Finally, although plant cells can often properly glycosylate proteins, the process is slightly different from that of animal cells. This may rule out using plants as biofactories for the expression of some proteins. We will discuss transgenic plants in greater detail later in Chapter 6.

Mammalian cell culture systems

It is possible to culture animal cells, growing them in a medium until it is time to harvest the proteins. This process is challenging because the nutritional requirements of mammalian animal cells are complex. Mammalian cells also grow relatively slowly, and the opportunity for mammalian cell cultures to become contaminated is greater than that of other culture systems. Despite these issues, mammalian cells are still the best choice for proteins destined to be used in humans.

Animal bioreactor production systems

Cells in culture are not the only option in using animal cells; sometimes living animals are protein producers. Consider, for example, the technique used to harvest monoclonal antibodies. Monoclonal antibodies react against only one target, making them valuable in diagnostic and therapeutic applications (see Chapter 7). Antibodies are proteins produced in reaction to **antigens** (usually an invading virus or bacteria). Antibodies can combine with and neutralize an antigen, protecting the organism. The production of antibodies is part of the immune response that helps living things resist infectious disease. When the production of a monoclonal antibody is the goal, mice are injected with an antigen. Then, the mouse's antibody-producing tissue is fused with cancer cells (to make them immortal). When fluid from the tumor that results is collected, the monoclonal antibodies can be purified from it.

Another method of animal bioreactor protein production uses the milk or eggs from transgenic animals (animals that contain genes from other organisms). These animal products contain the proteins from the recombinant gene that was introduced and can be purified from the milk or egg proteins. In 2009, the FDA approved the first human drug produced from a goat: ATryn, which treats a rare bleeding disorder in humans (see Chapter 7).

Insect systems

Insect systems are another avenue of protein production from animal cells. **Baculoviruses** (viruses that infect insects) are used as vehicles to insert DNA, causing the desired proteins to be produced by the insect cells. However, there are instances in which the post-translational modification of proteins is slightly different in insects than it is in other eukaryotes. Some covalent modifications, such as glycosylation and disulfide-bond formation, are frequently present in eukaryotic proteins and are often essential for correct protein folding and activity. Insect expression systems are commonly used when small quantities of proteins are needed in research.

Protein Purification Methods: Downstream Processing

Once a protein is produced, downstream processing begins (**Figure 4.7**). First, the protein must be harvested. If the protein is intracellular, the entire cell is harvested; if it is extracellular, the protein is excreted into the culture medium that is collected. Harvesting, though, is just the beginning of downstream processing. Next, the real work begins: The protein must be purified. This is the process of separating the target protein from the complex mixture of biological molecules in which it was produced.

Purity in this context is a relative term. Generally, the FDA requires that a sample be composed of 99.99 percent of the target protein. Separating proteins from all other cellular contents is not easy, and isolating the target protein from the other proteins in the sample can be even more difficult. To understand the process of purifying proteins, we look at some steps commonly followed in purification.

Step one: Preparing an extract for purification

If the protein is intracellular, the first task is **cell lysis,** disrupting the cell wall to release the protein. There are many methods for doing this: freezing and thawing (which disrupts cell membranes and releases cell contents), detergents (used to dissolve cell walls),

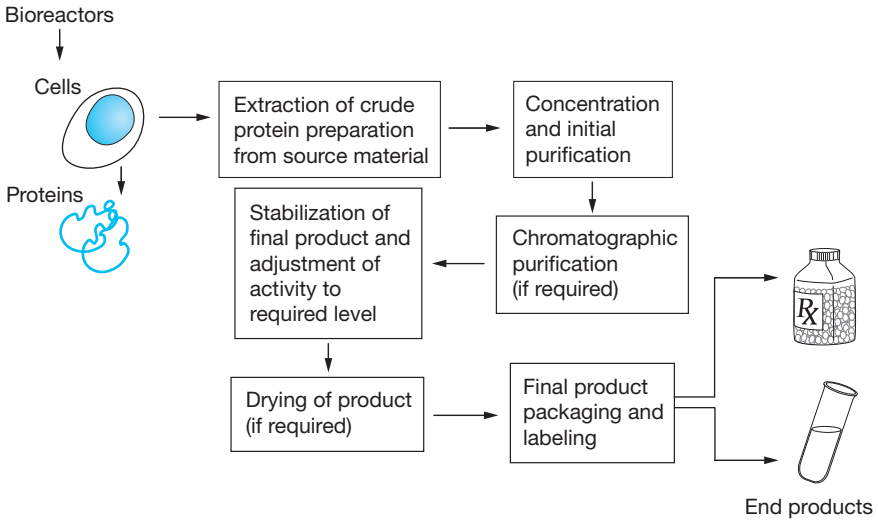

FIGURE 4.7 Basic Steps in Bioprocessing Purifications can be accomplished from raw materials or from bioreactors. The steps in the process must be devised and are often unique (and patentable).

and mechanical methods (ultrasonics or grinding with tiny glass beads). Given the fragility of proteins, freeing them from the cell without degrading them entirely is challenging. The disruption process releases the protein of interest as well as the entire intracellular content of the cell.

After the cells have been ruptured, organic alcohols or salts may be added to the mixture. Both of these take advantage of the hydrophobic orientation of the proteins by attracting water from the proteins, causing them to coalesce. These agents increase the interactions between the protein molecules to separate them from the mixture.

Step two: Stabilizing proteins in solution

Next, the proteins must be stabilized. Recall that it is important to maintain the bioactivity of the protein and that proteins are relatively fragile molecules. As a consequence, precautions must be taken to protect the protein during the purification process.

Maintaining a low temperature is crucial to protecting proteins, so most purifications must occur at low temperatures. Heat as moderate as room temperature limits the activity of proteins. Maintaining the proper pH for the activity of a protein is also important, and most active proteins are suspended in buffering agents to preserve maximal function.

Natural proteases that can digest the target proteins in a preparation are another threat. Protease inhibitors and antimicrobials can be added to prevent the protein molecules from being dismantled but must be removed later, as must any additive used in the purification process. Still another potential problem is mechanical destruction by foaming or shearing of the proteins into useless fragments. Once again, additives can help prevent foaming and shearing from destroying the protein, but the additives must be removed later.

As we've seen, some purification methods are powerful enough to damage the target protein. It takes a balancing act to extract and purify proteins successfully. Although it is essential that the protein be purified, it is equally important that the protein maintain its biological activity.

Step three: Separating the components in the extract

The last step in the purification process can be the most important. Similarities between proteins permit us to separate them from material such as lipids (fats), carbohydrates, and nucleic acids, which are also released when a cell is disrupted. Differences between individual proteins are then used to separate the target proteins from others. Following are several methods used for protein separation.

Protein precipitation

Proteins often have hydrophilic amino acids on their surfaces that attract and interact with water molecules. That characteristic is used as the basis for separating proteins from other substances in the extract. Salts, most commonly ammonium sulfate, can be added to the protein mixture to **precipitate** the proteins (to cause them to settle out of solution). Ammonium sulfate precipitation is frequently a first step in protein purification, resulting in a protein precipitate that is relatively stable. Problems associated with ammonium sulfate precipitation make it a poor choice in some industrial situations, however. Ammonium sulfate is highly reactive when it contacts stainless steel, for example, and many industrial purification facilities are made of stainless steel. Other solvents frequently used

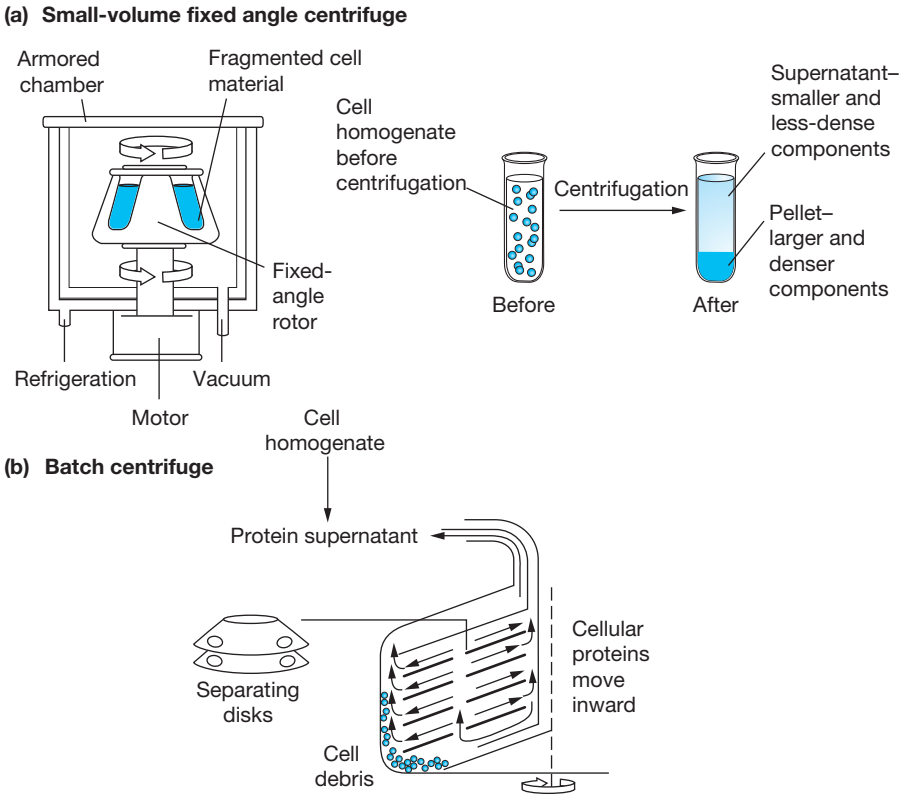

FIGURE 4.8 Fixed-Angle and Batch Centrifugation Fixed-angle centrifuges (a) can develop extremely high gravitational forces (*g* forces) but are limited to smaller quantities and must be run separately for each batch. Batch centrifuges (b) were developed to allow continuous flow of materials and separation of cell debris from cell proteins. The continuous pressure of the inflow and the centrifugal force can be adjusted to maximize the separation and outflow of the protein supernatant.

to promote protein precipitation include ethanol, isopropanol, acetone, and diethyl ether. Just like ammonium sulfate, these solvents cause protein precipitation by removing water from between the protein molecules.

Filtration (size-based) separation methods

There are a variety of ways to separate molecules based on size and density. **Centrifugation** separates samples by spinning them at high speed. With this process, proteins can often be isolated in a single layer or separated from heavier cell components. Small-volume centrifuges are capable of processing only a few liters at each run. Large reactors can process hundreds or thousands of liters (see **Figure 4.8**). Industrial-scale centrifugation is normally achieved using continuous-flow centrifuges that allow continuous processing of the contents of a bioreactor.

Filters of various sizes and types can also be used to separate protein from other molecules in the mixture. In this process, known as **membrane filtration,** thin membranes of nylon or other engineered substances with varying pore sizes are used to filter out all of the cellular debris from a solution. First, **microfiltration** removes the precipitates and bacteria. **Ultrafiltration** then uses filters that can catch molecules such as proteins and nucleic acids. Some ultrafiltration processes can even separate large proteins from smaller ones. (Refer to www.amicon.com to view some of these devices.) One of the main shortcomings of membrane filtration systems is their tendency to clog easily. On the plus side, the use of these filtration systems takes less time than centrifugation.

Diafiltration and **dialysis** are filtration methods that rely on the chemical concept of equilibrium, the migration of dissolved substances from areas of higher concentration to areas of lower concentration. As shown in **Figure 4.9**, dialysis depends on the ability of some molecules to pass through semipermeable membranes (osmosis) while others are halted or slowed because of their size. Dialysis is often required to remove the smaller salts, solvents, and other additives used earlier in the purification. The salts are then replaced with buffering agents that help stabilize the proteins during the remainder of the process. Diafiltration adds a filtering component to dialysis.

Chromatography

The initial steps in any purification process liberate a protein from the cell, remove undesired contaminants and particulates, and concentrate the proteins. **Chromatography** methods allow the sorting of proteins by size or by how they cling to, or separate from, other substances. In chromatography, long glass tubes are filled with microscopic resin beads and a buffered solution. The protein extract is then added and flows through the resin beads in a glass column. Depending on the resin used, the protein either sticks to the beads or passes through the column while the beads act as a filtration system.

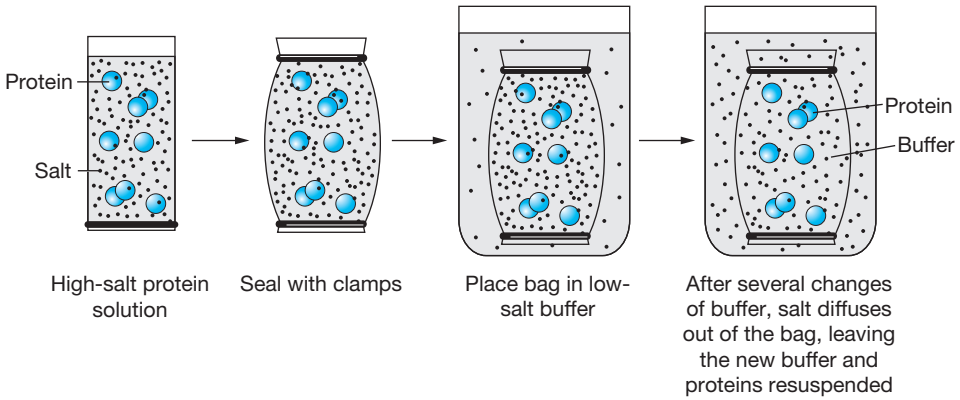

FIGURE 4.9 Dialysis This procedure can be used to remove salt by the process of diffusion and replace it with a buffer that is better suited for protein stability.

Size-exclusion chromatography (SEC) uses gel beads as a filtering system. Larger protein molecules quickly work their way around the gel beads while smaller molecules pass through more slowly because they are able to slip through tiny holes in the beads, as shown in Figure 4.10. The gels are available in a variety of pore sizes, and the necessary gel for proper separation depends on the molecular weight of the contaminants or proteins being separated. This method can make only preliminary separations, however, and can pose problems in industrial settings because it requires very large columns.

Ion-exchange (IonX) chromatography relies on an electrostatic charge (like static cling) to bind proteins to resin beads in a column. While the charged proteins cling to the resin, other contaminants pass through and out of the column, as shown in Figure 4.11. The proteins can then be eluted (released from the resin) by changing the electrostatic charge; this is done by rinsing the column with salt solutions of increasing concentrations. The bound protein is then released from its attachment (detected by viewing under UV 280) and collected.

Affinity chromatography relies on the ability of most proteins to bind specifically and reversibly to uniquely shaped compounds called **ligands.** Ligands are small molecules that bind to a particular large molecule in a protein. Think of ligands fitting with a unique protein molecule the way a key fits a lock (Figure 4.12). After the proteins have bound to the

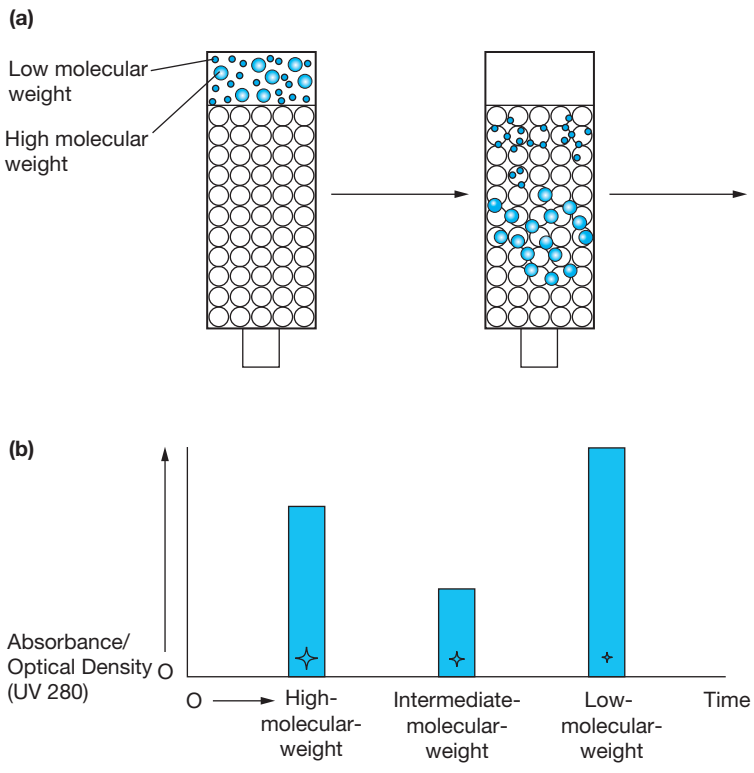

FIGURE 4.10 Protein Purification by Size-Exclusion Chromatography (a) Low- and high-molecular-weight proteins travel through a size-exclusion chromatography column. (b) The diagram shows how the low and high molecular weight proteins appear as they come off the column when monitored for proteins (ultraviolet [UV] 280 nm absorption by aromatic amino acids used for detection). Notice that the high-molecular-weight proteins move quickly through the buffer, whereas the low-molecular-weight proteins are slowed by the matrix of the column resin. Resins may be purchased with many different pore sizes.