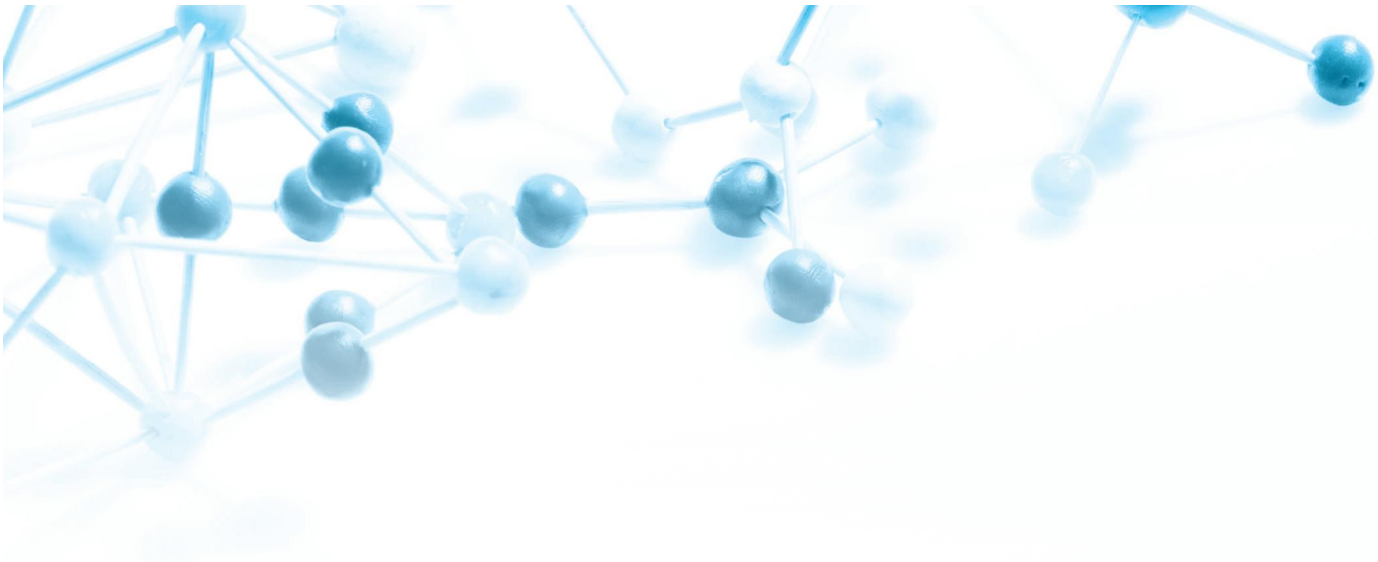


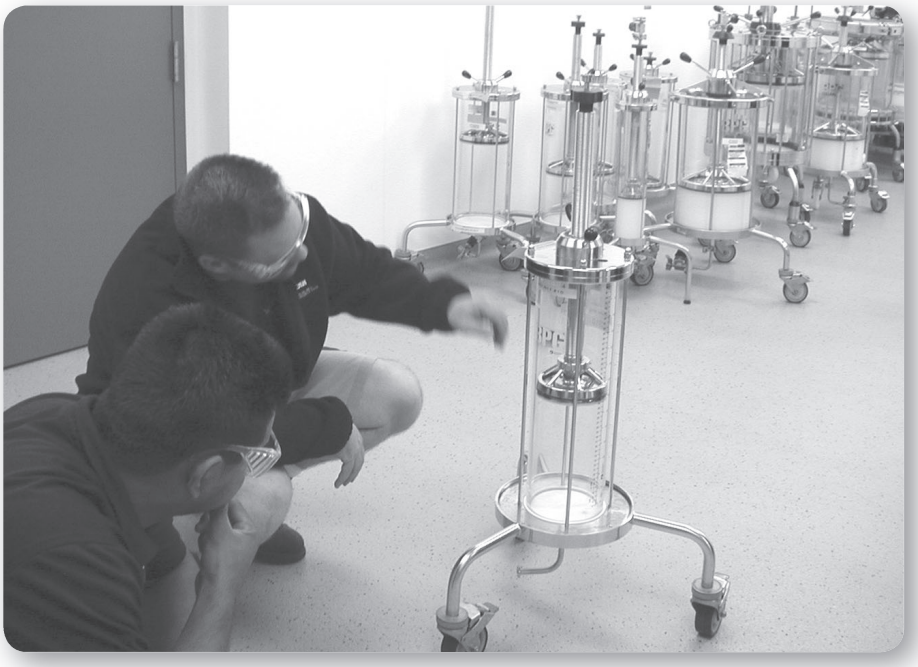
Proteins as Products

From Chapter 4 of *Introduction to Biotechnology*, Third Edition. William J. Thieman, Michael A. Palladino. Copyright © 2013 by Pearson Education, Inc. All rights reserved.

Proteins as Products

After completing this chapter, you should be able to:

- Explain the uses of some biotechnologically produced enzymes in industry.
- Provide three examples of medical applications of proteins.
- List common household products that may include manufactured proteins as ingredients.
- Discuss the advantages and disadvantages of bacterial, fungal, plant, and animal sources of protein expression.
- Explain why *Escherichia coli* is frequently used for protein production, and describe *E. coli*'s limitations.
- Explain why protein glycosylation may determine the choice of a protein expression system.
- Describe a general scheme for protein purification of a protein like hemoglobin.
- Explain how a target protein is separated from other cell proteins given a specific purification sequence.
- Discuss the benefits of being able to predict protein structure from the DNA sequence (proteomics).



Lab Supervisor Preparing a Purification Column to Separate a Commercial Biological Product.

Michael A. Palladino.

FORECASTING THE FUTURE

Biotechnologically produced medications (or drugs) are all proteins. Until recently, the chemistry of the purified protein molecule dictated the method of drug delivery (injection, absorption, ingestion, or other means). Unfortunately, these delivery methods can lead to serious systemic and autoimmune side effects. Interferon alpha-2b, traditionally given by injection and approved for treatment of melanoma, hepatitis C, and conditions caused by the human papillomavirus (HPV), is a good example of a medicine with a traditionally problematic delivery method. Recently, to address this problem, the company Helix BioPharma has developed a multilamellar nanovesicle (a multilayered submicroscopic particle that can be easily assimilated by the body) that enables a timed, controlled release of the protein applied in a skin cream. This topical delivery is adapted to large hydrophilic molecules such as interferon alpha-2b, which helps avoid the serious side effects caused by more traditional delivery methods. In the future, the design of delivery methods that are adapted to the chemistry of the protein biomolecule may be as important as the discovery of the biomolecule itself.

Tropical rainforests, the deepest reaches of the ocean, boiling geysers in Yellowstone National Park, and whale skeletons—these are on the frontier of the scientific quest for proteins. **Proteins** are large molecules required for the structure, function, and regulation of living cells. Each protein molecule has a unique function in the biochemical reactions that sustain life. As researchers explore the proteins that occur in nature, they unlock secrets that govern growth, speed of chemical decomposition, and protection from disease.

The applications of proteins are as numerous as proteins themselves. Consider whale skeletons: during the natural decomposition process, bones are often colonized by bacteria, some of which have evolved especially to digest the fatty residue remaining on them. The proteins that the bacteria produce to break down the fats are adapted to the frigid waters of the deep sea. Researchers recognized that a substance with the ability to dissolve fats at cold temperatures would make a great additive for commercial laundry detergents.

Even after a protein is discovered in nature and an application is matched to its characteristics, a great deal of ingenuity is required to produce the protein of the necessary quality and quantity required. For example, if we plan to mass-produce a new cold-water detergent, we cannot rely on whale skeletons; rather, we must find another source of the proteins needed. Fortunately, biotechnology facilitates the production of virtually all proteins. We focus on those production processes in this chapter.

In 2000, the National Institutes of Health launched the Protein Structure Initiative (PSI), a 10-year, $600 million effort to identify the structure of human proteins. The PSI is a federal university and industry effort aimed at reducing costs and lessening the time it takes to determine three-dimensional protein structures. The initial goal of the PSI was to make the structures of most proteins easily obtainable from knowledge of their corresponding DNA sequences. In the current phase, which began in 2010, investigators are using high-throughput structure determination, which was successfully developed during the earlier phases of the PSI, to study a broad range of important biological and biomedical problems.

This massive effort to further understand the structure of proteins helped researchers understand their function. More than 1,200 superfamilies (groups of related protein structures) of proteins have been identified so far, and the relationship between a protein's amino acid sequence to its structure is now well understood. The public database that is part of the initiative currently holds more than 33,000 protein sequences. The goal of the initiative is to be able to model unknown protein structures based on structural comparisons to those stored in the database. For example, the Human Genome Project produced the 3 billion base-pair sequences of human DNA, but the predicted 100,000 proteins resulting from this DNA will be hard to determine without the effort of scientists supported by major governmental initiatives like this.

We begin with a quick survey of the many applications of proteins in a variety of industries. Then we look at the nature of protein structures, paying special attention to the process of protein folding. With that as a foundation, we delve into some of the details of protein processing, beginning with the methods of expressing proteins. We then learn how expressed proteins are purified and examine the processes used to analyze and verify the final product. Although there is no one best method for processing proteins, several generally useful techniques are available. In this chapter, we look at those techniques, keeping in mind that the specifics of protein processing vary from case to case.

1 Proteins as Biotechnology Products

The use of proteins in manufacturing processes is a time-tested technology. For example, two of the oldest food-processing endeavors depend on proteins: beer brewing and winemaking. Fermentation depends not

only on yeast-produced enzymes but also on others added directly to the batch. Cheese making is another food-processing industry that has always used proteins and, thanks to bioengineering, the protein source now used is from engineered bacteria rather than the calves' stomachs that originally provided the cheese-making enzyme (Figure 1).

Even though the value of proteins in manufacturing had long been evident, we were not able to further this knowledge until the 1970s, when recombinant DNA technology was first developed and it became possible to produce specific proteins on demand. Since that time, the production of proteins has been the driving force behind the development of new products in a wide variety of industries.

Other industries are also benefiting from the ready availability of bioengineered proteins. Many of these applications depend on the power of a group of proteins called **enzymes** to speed up chemical reactions. Enzymes serve myriad purposes, such as making detergents work better, increasing the flow of oil in drilling operations, and cleaning contact lenses. Enzymes are used in manufacturing to break down large molecules—a process called **depolymerization.** These enzymes include glycosidases (carbohydrases) like **amylase,** which breaks down starch; **proteases,** which break down other proteins; and **lipases,** which break down fats (Table 1). Such enzymes (like chymosin, for cheese) are used in food and beverage production and for bulk processing in industries (see Figure 1).

Hormones, which carry chemical messages, and **antibodies**, which protect the organism from disease, are two other groups of proteins produced commercially, primarily for the medical industry. Hormones also have agricultural uses. For instance, hormones can stimulate the rooting of plant cuttings and encourage more rapid growth of meat animals.

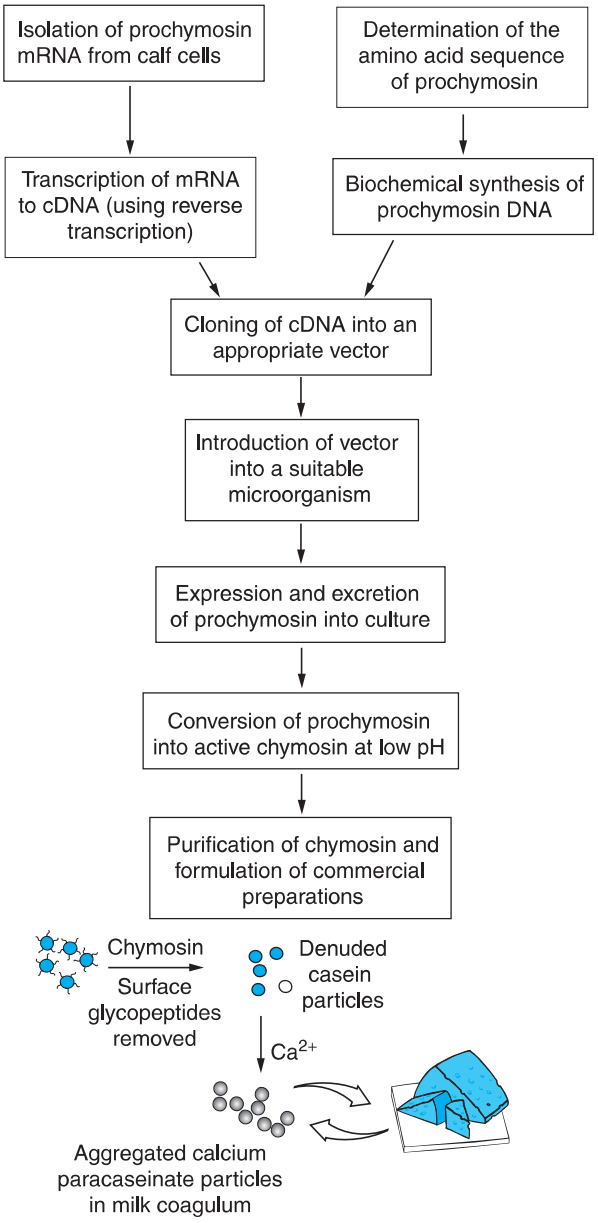


FIGURE 1 Cheese Production Casein, the primary ingredient in cheese, is the result of a chemical conversion that depends on chymosin. This enzyme has been obtained from the stomach walls of unweaned calves for centuries. Today 80 percent of the chymosin manufactured is produced (and purified) from genetically modified cells.

John E. Smith, *Biotechnology*, Third Edition, 1996, © J. E. Smith 1981, 1988, © Cambridge University Press 1996, reproduced with permission of Cambridge University Press and the author.

TABLE 1 SOME ENZYMES AND THEIR INDUSTRIAL APPLICATIONS

Enzyme	Application
Amylases	Digest starch in fermentation and processing
Proteases	Digest proteins for detergents, meat/leather, cheese, brewing/baking, animal/human digestive aids
Lipases	Digest lipids (fats) in dairy and vegetable oil products
Pectinases	Digest enzymes in fruit juice/pulp
Lactases	Digest milk sugar
Glucose isomerase	Produce high-fructose syrups
Cellulases/hemicellulases	Produce animal feeds, fruit juices, brewing converters
Penicillin acylase	Produces penicillin

Biotech Drugs and Other Medical Applications

Biotechnology proteins have revolutionized the health care and pharmaceutical industries in the past several decades. Many illnesses, from common conditions like diabetes to rare diseases like Gaucher's disease, can be treated by replacing missing proteins. In the case of diabetes, the missing protein is the hormone insulin. Insulin once had to be harvested from pigs and cows. This was less than ideal because human bodies often rejected this foreign protein. Researchers overcame the problem by turning to an unlikely source: the bacterium *E. coli*. By inserting human genes into *E. coli*, they created microscopic insulin factories. (We look at the remarkable use of genetically engineered organisms as a source of proteins more closely later in this chapter.) The U.S. Food and Drug Administration (FDA) approved this new insulin in 1982, making it the first recombinant DNA drug. The ability to produce an abundant supply of human insulin has improved the health and lives of millions of people. (Table 2 lists some other protein-based pharmaceutical products.)

Therapeutic proteins—such as monoclonal antibodies, blood proteins, and enzymes produced by living organisms to fight disease—can also be thought of as biotech drugs. Unlike other medicines, these drugs are not synthetically produced (i.e., chemically synthesized by adding one compound at a time) but are usually produced through microbial fermentation or by mammalian cell culture. Today, there are nearly 400 biotechnology medicines in the pipeline, and the majority are proteins. If even a small percentage of these drugs succeed, it will significantly add to the approximately 40 biotech drugs currently in use.

Producing biotech drugs is a complicated and time-consuming process. Researchers can spend many years just identifying the relevant therapeutic protein, determining its gene sequence, and working out a process to make sufficient quantities of the protein molecules using biotechnology. Once this method is determined, technicians can produce large batches of the protein products in bioreactors, under carefully controlled conditions, by growing host cells that have been transformed to contain the therapeutic gene (a bioreactor is a sterile production container designed to produce biological products). The cells are stimulated to produce the target proteins through precise culture conditions that include a balance of temperature, oxygen, acidity, and other variables. At the appropriate time, the proteins are isolated from the cultures, stringently tested at every step of purification (which we discuss later in this chapter), and formulated into pharmaceutically active products. Manufacturing technicians monitoring bioreactors must strictly comply with FDA regulations at all stages of the procedure.

Another dramatic example of the potential use of proteins in health care is in the treatment of damaged

TABLE 2 SOME PROTEIN-BASED PHARMACEUTICAL PRODUCTS (MOST PRODUCED AS RECOMBINANT PROTEINS)

Protein	Application
Erythropoietins	Treatment of anemia
Interleukins 1, 2, 3, 4	Treatment of cancer, AIDS; radiation- or drug-induced bone marrow suppression
Monoclonal antibodies	Treatment of cancer, rheumatoid arthritis; used for diagnostic purposes
Interferons (α, β, γ, including consensus)	Treatment of cancer, allergies, asthma, arthritis, and infectious disease
Colony-stimulating factors	Treatment of cancer, low blood cell count; adjuvant chemotherapy; AIDS therapy
Blood clotting factors	Treatment of hemophilia and related clotting disorders
Human growth factor	Treatment of growth deficiency in children
Epidermal growth factor	Treatment of wounds, skin ulcers, cancer
Insulin	Treatment of types 1 and 2 diabetes mellitus
Insulin-like growth factor	Treatment of type 1 diabetes mellitus
Tissue plasminogen factor	Treatment after heart attack, stroke
Tumor necrosis factor	Cancer treatment
Vaccines	Vaccination against hepatitis B, malaria, herpes

corneas. A study from researchers in Canada and Sweden has shown that biosynthetic corneas can help regenerate and repair damaged eye tissue and improve vision. The discovery is significant because this approach could help restore sight to millions of people who are waiting for donated human corneas for transplantation. In the study, each patient underwent surgery on one eye to remove damaged corneal tissue and replace it with the biosynthetic cornea, made from synthetically cross-linked recombinant human collagen. The protein was produced in yeast cells and chemically cross-linked for the experiments. Over 2 years of follow-up, the researchers observed that cells and nerves from the patients' own corneas had grown into the implant, resulting in a "regenerated" cornea that resembled normal, healthy tissue. The biosynthetic corneas also became sensitive to touch. Vision improved in many of the patients, and after contact lens fitting, vision was comparable to conventional corneal transplantation.

Using a new method for rapidly screening molecules associated with disease, Joshua LaBaer and colleagues from the Biodesign Institute at Arizona State University have identified a broad panel of 28 early predictors, or **biomarker proteins,** which may one day aid in the early diagnosis of breast cancer. Studies have shown that proteins produced by cancers can trigger the body to produce antibodies not found in healthy individuals. These "autoantibodies" can be measured in the blood and used to reveal the presence of cancer. In the breast cancer study, protein microarrays were used to display thousands of different potential biomarker proteins on a single microscope slide. A tiny drop of blood was added to the microarray to look for those proteins that are recognized by antibodies from cancer patients but not from the healthy women (see section 10). The initial results narrowed the number of potential biomarkers candidates from 5,000 to 761. Finally, these 761 proteins were tested in a blind study to find the final 28 protein biomarkers. Autoantibody biomarkers in patients can be readily used for the detection of many other cancers: ovarian, prostate, and lung, among others.

Food Processing

You can find plenty of examples of industrial uses of proteins in your kitchen. For decades the food processing industry has relied on enzymes to improve baby food, canned fruit, cheeses, baked goods, beer, desserts, and dietetic foods. The enhancements are quite varied.

In bread, for example, enzymes may be used to make starches easier for yeast to act on, allowing the dough to rise more quickly. As a result, the bread dough is easier to handle and the final texture is more consistent.

Textiles and Clothing

Other examples of industrial proteins are right in your closet. For more than a century, enzymes have been used in the textile industry to break down the starches used to "size" (stiffen) fabrics during the manufacturing process. Enzymes are also replacing harsh chemicals and processes used to lighten and soften fabrics.

One exciting example of the textile uses of proteins is illustrated by recent research into spider silk. Spider-silk fibers have astonishing mechanical properties: they have strength comparable to steel, toughness greater than Kevlar and are less dense than cotton or nylon. However, spiders don't spin nearly enough silk to be harvested for use in industrial products. Researchers in Germany have recently discovered that protein subunits are responsible for silk's amazing properties. These findings provide a clearer understanding of the mechanical nature of spider-silk fibers and may be useful for design of silk-like products. The transfer of spider-silk genes into host organisms that can produce the quantities of protein needed has already been completed and the new structural information should lead to new products soon.

YOU DECIDE

Testing for the Best Product: Who Should Pay?

It takes about $800 million to bring a drug to market. Included in this cost is the price of researching drugs that are not approved owing to an adverse reaction or ineffectiveness. Usually, the purification process has already been developed and the product is in human trials when this happens Although drug prices are often high, many people do not realize that these prices reflect, in part, the cost of research on products that were not approved. If we also add the ability to determine which drug is best for each patient (pharmacogenetics) to the cost of bringing a drug to market, the cost climbs higher. Because biotechnology companies are supposed to make a profit for their stockholders, it is more profitable for the company to have everyone buy their drug, even if the drug is not entirely suited for each individual. What do you think is the best solution: higher prices and better drugs, or lower prices and drugs that do not always work and may even have negative side effects? You decide.

Detergents

As noted earlier, when enzymatic ingredients are added to detergents, they do a better job of cleaning and are more biodegradable. Laundry detergents take advantage of the specific roles of proteases, lipases, and amylases to dissolve stains in cooler water. If you look at the labels of many laundry stain removers, you will see enzymes listed as the first and sometimes only active ingredients.

Bioremediation: Treating Pollution with Proteins

In addition to reducing the quantity of pollutants produced during industrial processes, proteins can be used to clean up harmful wastes. Organic wastes from feedlots, homes, and businesses are a growing threat to the environment, especially aquatic ecosystems. Enzymes can be used to digest such organic wastes before they cause trouble.

Another promising new application for proteins is in neutralizing heavy metal pollutants such as mercury and cadmium. These dangerous elements can persist in the environment, causing harm to organisms throughout the food chain. These metals resist enzymatic breakdown, but this does not mean that proteins cannot be used to neutralize them. Some microorganisms have a sticky coat of **metallothioneins,** proteins that actually capture heavy metals. In this case, the pollutant is not dismantled or digested but simply made less dangerous. When toxic metals are bound to bacteria, they are less likely to be absorbed by plants and animals.

Researchers are currently using the power of genetic engineering to create new, better biological tools to attack and destroy toxic substances. Because the research process sometimes depends on randomly shuffling genes in the bacteria, the enzymes produced by the rearranged genes can be more or less reactive than those naturally produced. In a sense, scientists are accelerating the process of random mutation and evolution in the hope of discovering new, more efficient pollution-eating proteins.

2 Protein Structures

Ribosomes are factories that produce proteins. To understand the processes of expressing and harvesting proteins, we must look at the molecular structure of proteins more closely.

Proteins are complex molecules built of chains of amino acids. Like all molecules, proteins have specific molecular weights. They also have an electrical charge that cause them to interact with other molecules. This ability to interact is the key to the biological activity of proteins. Consider, for example, the way the chemical structure and electrical charge of an amino acid can influence its interactions with water: The molecules will be either **hydrophilic** (water loving, attracted to water molecules) or **hydrophobic** (water hating, as if the water molecules and amino acids repelled one another).

Structural Arrangement

Proteins are capable of four levels of structural arrangement. These are the primary, secondary, tertiary, and quaternary structures of proteins (see **Figure 2**). The exact arrangement that a protein has depends on the specific chemical sequence of its amino acids and the types of side groups that are present.

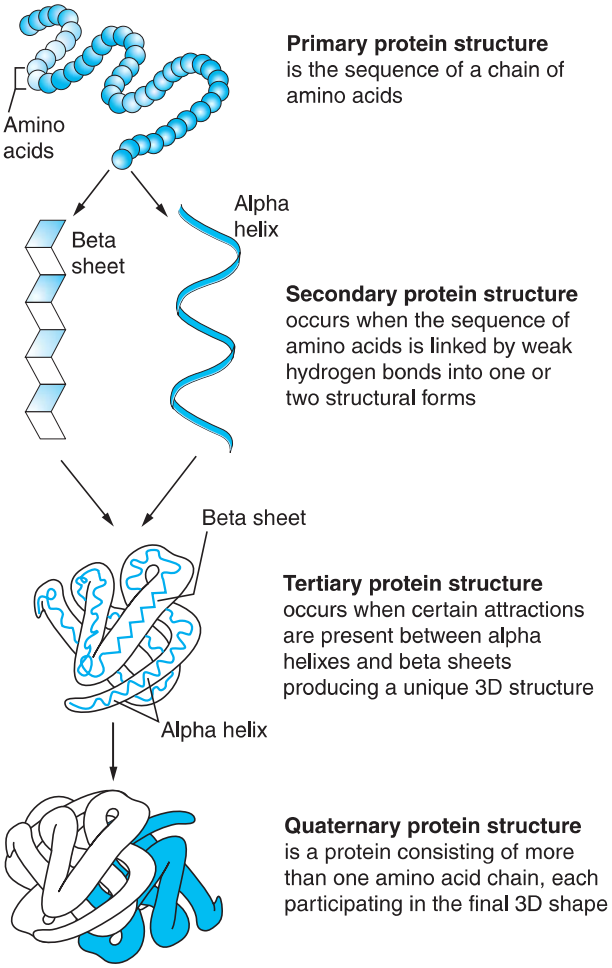


FIGURE 2 The Four Levels of Protein Structure The proper folding of proteins is necessary for full functional capability. Purification methods must guarantee that proper folding is retained.

Primary structures

The 20 commonly occurring amino acids are the building blocks that make up proteins. Ten to 10,000 amino acids can be linked together in a head-to-tail fashion to form the sequence of a protein. The sequence in which amino acids are linked together at the ribosome is known as the *primary structure*. Altering a single amino acid in the sequence can mean that the protein loses all function. Genetic diseases are often the result of these protein mutations.

Secondary structures

Secondary protein structures occur when chains of amino acids fold or twist at specific points, forming new shapes due to the formation of hydrogen bonds between hydrophobic amino acids. The most common shapes, alpha helices and beta sheets, are described in the section on protein folding. In the alpha-helix arrangement, the amino acids form a right-handed spiral. The hydrogen bonds stabilize the structure, linking an amino acid's nitrogen atom to the oxygen atom of another amino acid. Because the links occur at regular intervals, a spiraling chain is formed. In the beta-sheet structure, the hydrogen bonds also link the nitrogen and oxygen atoms; however, because the atoms belong to amino acid chains that run side by side, an essentially flat sheet is formed. The sheets can either be "parallel" (if the chains all run in the same direction) or "antiparallel" (in which case the chains alternate in direction). One of the fundamental elements of protein structure, the beta turn, occurs when a single chain loops back on itself to form an antiparallel beta sheet. Both the alpha-helix and beta-sheet structures exist because they are the most stable structures the protein can assume.

Tertiary structures

Tertiary protein structures are three-dimensional polypeptides (large molecules made up of many similar, smaller molecules) that are formed when secondary structures are cross-linked. The bonds that hold together tertiary structures occur between amino acids capable of forming secondary covalent bonds (like cysteine, which can form disulfide bonds with another adjacent cysteine, cross-linking the protein into a unique shape). There is a SH side group of cysteines that can cross-link to form disulfide bridges.) The tertiary structure of a protein determines its function, such as binding a cellular receptor or catalyzing a chemical reaction.

No matter what secondary and tertiary structure a protein takes, it is important to remember that structures are fragile. Hydrogen bonds can be broken easily, damaging a valuable protein. Anyone who has heated an egg has realized that the fluid egg proteins (albumin) quickly change shape (egg white) as the hydrogen and cross-linking tertiary bonds are broken by heat. Most proteins used in the lab are kept on ice to maintain their fragile structure and prevent them from being destroyed by the heat in the lab.

Quaternary structures

Quaternary protein structures are unique, globular, three-dimensional complexes built of several polypeptides. Hemoglobin, which carries oxygen in the blood, is an example of a protein with quaternary structure; it is composed of two tertiary proteins linked together.

Protein Folding

Everything that is important about a protein—its structure, its function—depends on folding. Folding describes how different strands of amino acids take their shape; for example, sickle cell anemia results from a misfolding due to a single amino acid replacement at a strategic location in the primary structure. If a protein folds incorrectly, not only will the desired function of the protein be lost but the resulting misfolded protein can be detrimental. For example, the plaque that forms in Alzheimer's disease accumulates because a misfolded protein cannot be broken down by the enzymes present in brain cells.

The first breakthrough in understanding the fundamental forms of proteins came in 1951, when Pauling and Corey described the alpha helix and beta sheet, which are the most common components of protein structure. Both structures result from hydrogen bonding that ties together the chain of amino acids. Not only Alzheimer's disease but cystic fibrosis, "mad cow" disease (bovine spongiform encephalitis, or BSE), many forms of cancer, and some heart attacks have all been linked to clumps of incorrectly folded proteins. Because protein folding occurs naturally, it is easy to see why one of the biggest challenges biotechnology faces is understanding and controlling the protein folding in the manufacturing process.

Glycosylation

More than 100 **posttranslational modifications** (like glycosylation) occur within a eukaryotic cell. In **glycosylation,** carbohydrate units (sugar molecules) are added to specific locations on proteins (see **Figure 3**). This change can have a significant effect on a protein's activity: it can increase solubility and orient proteins into membranes and extend the active life of a molecule in an organism.

Glycoproteins can be used in the treatment of disease, as scientists with the Scripps Research Institute (reported in the June 10, 2010, edition of the journal

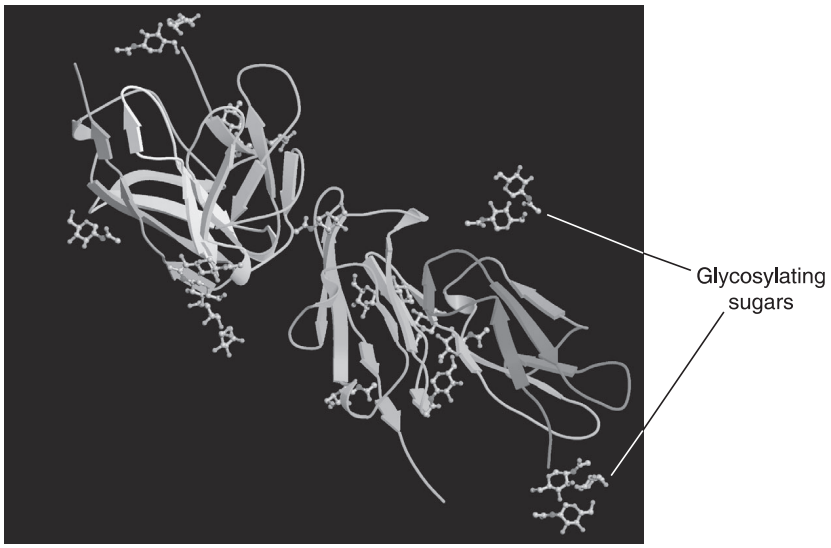


FIGURE 3 Three-Dimensional Protein Structure with Glycosylated Side Chains Glycosylation occurs within eukaryotic cells and probably extends the life of the protein by protecting it from natural cellular mechanisms that destroy proteins.

The Protein Data Bank/RCSB.

Blood) have discovered. This discovery represents a new way to target and destroy a type of cancerous cell. A glycoprotein can be combined with a nanoparticle (a small synthetic molecule) loaded with a chemotherapy drug, resulting in a new way to target and destroy cancer cells. By targeting B lymphoma cancers with chemotherapeutic-loaded nanoparticles, the effective dose of the drug is increased while simultaneously protecting normal tissues (as described in *Forecasting the Future* earlier in this chapter). It's clear that the findings of the Scripps scientists could lead to the development of other novel drug therapies based on glycosylation to treat leukemia, lymphomas, and related cancers.

Protein Engineering by Directed Molecular Evolution

Biotechnology often relies on protein engineering At times it is necessary to introduce specific, predefined alterations in the amino acid sequence. This can be done with directed molecular evolution technology. A major biotech company, for example, induces mutations randomly into genes and then selects the bacterium with the protein product (enzyme) that has the highest activity. In this way, the company has been able to produce organisms (and industrial enzymes) that tolerate a cyanide concentration of more than 1.0 mole (M) per liter, which is too high for most bacteria to survive. No "natural" environment has experienced cyanide at this level. The resulting selected organisms can be used to remediate cyanide contamination resulting from mining and other industrial waste accumulation. A mutation resulting in this type of organism is not possible in a natural selection event because the natural environment has not changed enough to select for these types of bacterial survivors. Directed molecular evolution requires introducing specific changes in the nucleotide sequences of a particular gene, resulting in new arsenic-metabolizing enzymes, as seen in **Figure 4** on the next page.

Such newly modified genes can then be introduced into a host cell, where the required amino acid sequence is produced by the host system. This technique allows researchers to create proteins with specific enhancements. Unlike naturally occurring mutation, directed molecular evolution focuses only on mutations that occur in a specific gene and selects the best proteins from that gene, irrespective of the potential benefits it may have for the original organism. For example, when *E. coli* manufactures human insulin, there is no benefit to the bacteria. For more information about directed molecular evolution, visit Maxygen Corporation's website at www.maxygen.com.

Enzymes have been modified over millions of years of evolution to catalyze specific chemical reactions in cells. Within the last decade, scientists have used directed evolution and rational design (designing a protein to fit a surface) with varying degrees of success to improve activity, stability, and selectivity of native enzymes. Much progress has been made in David Baker's lab at the University of Washington using a rational design program called Rosetta. The team designed an enzyme not found in nature by modeling the active site part of the enzyme and then finding a protein scaffold structure to attach to the part. After testing 84 constructs for activity, they selected 50 and inserted their DNA sequence into *E. coli* for expression. The final chosen enzyme functioned, but not as well as the native enzymes. This required them to mutate the active site many times to

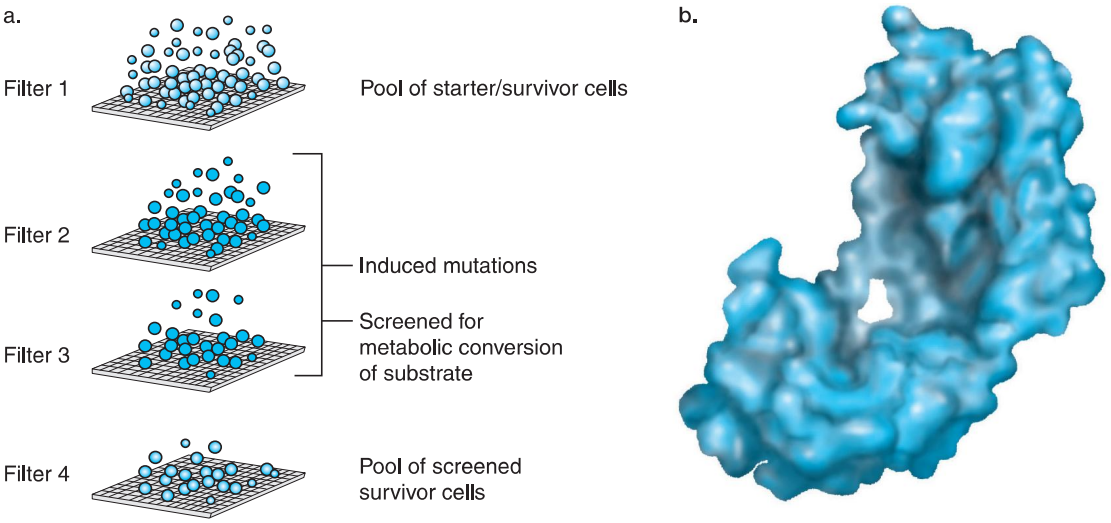


FIGURE 4 Directed Molecular Evolution Technology a. Genes that express proteins can be subjected to mutational events, generating a diverse group of novel gene sequences. Organisms with mutated genes are screened for unique properties. After an improvement in protein function, the process can be repeated until the maximum function is obtained. Unlike natural selection, this process focuses on the properties of the protein, not the organism, and can achieve changes that may never occur in nature. It may not have an evolutionary advantage, but the proteins it produces will have more commercial value. b. Scientists have discovered a protein used by bacteria to improve protein folding. The "spy" protein (shown) reduces the misfolding that commonly occurs when bacteria produce recombinant proteins for drug or industrial use. The research team used "directed evolution" to successively select bacteria for their abilities to refold proteins that protect them against antibiotics.

achieve better activity. This research shows that the day when proteins can be designed rationally is yet to come, but progress is being made.

In addition to naturally occurring and mutation-produced proteins, biotechnology is also creating entirely new protein molecules. These molecules, designed and built in the laboratory, indicate that it might be possible to invent proteins that are tailored for specific applications. Faulty protein folding is often expressed as a disease caused by infectious protein particles, called **prions** (Figure 5), which attract normal cell proteins and induce changes in their structure, leading to the accumulation of useless proteins that damage cells. Prion can occur in sheep and goats (scrapie) and cows (bovine spongiform encephalitis, or "mad cow" disease). Human forms of these brain-destroying diseases include kuru and **transformable spongiform encephalitis (TSE).** All these diseases involve changes in the conformation of the prion precursor protein, a protein normally found in mammalian neurons as a membrane glycoprotein. The inability to detect the disease until the infected animal is either sick or dead has complicated control of these diseases. Biotechnology research has sought to form synthetic infectious particles that can be studied, which will aid in developing detection and control.

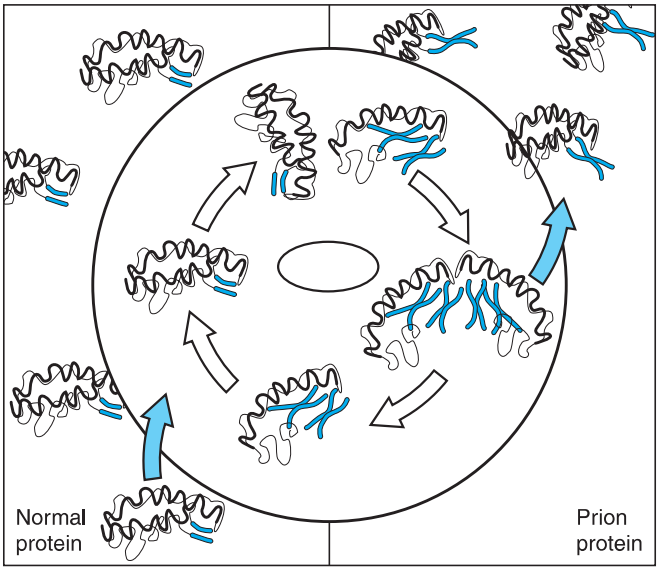


FIGURE 5 Prions Are Misfolded Proteins A misfolding of proteins that occurs in prion disorders can be duplicated in the lab to produce large quantities of prion proteins, which can then be studied to create diagnostic kits. Follow the arrows from the left as a prion protein attracts a normal protein and changes it into a prion, resulting in an accumulation of prions within the cell.

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.

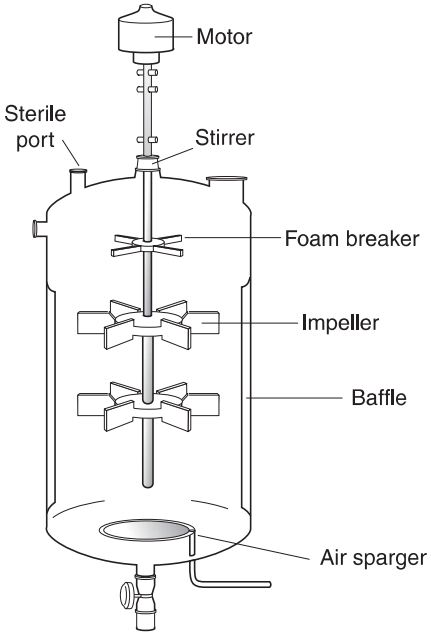


FIGURE 6 Bioreactor Large-volume cell culture is commonly performed in a self-contained, closed (sterile) system until the products are harvested. Through sterile ports, workers can adjust pH, gas concentrations, and other variables based on input from internal sensors. Sterile bags as large as 2,000 liters can be used on rocking tables as alternatives to stainless steel bioreactors.

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.

The bacterial species most commonly used to produce genetically engineered proteins is *E. coli*. Because early research into bacterial genetics focused on *E. coli* as a model system, we now understand the genetic characteristics of *E. coli* reasonably well.

In some instances, the foreign gene (in the form of **cDNA or complementary DNA**) for the desired protein product is attached directly to a complete or partial *E. coli* gene. In these 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, below). 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 from the other cell proteins before they can be used.

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 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 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 pro-

TABLE 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

moter 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.

Fungi

Fungi are the source of a wide range of proteins used in products as diverse as animal feed and beer. Naturally existing proteins found in some fungi are nutritious and used as foods 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 posttranslational modification (like folding human proteins correctly) and are used for that synthesis, as illustrated in Table 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 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. Enzymes produced by plants will undoubtedly also be used in the near future for increased drug production.

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 nonfood 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.

TABLE 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*

Mammalian cell culture systems

Sometimes 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 if not the only 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. Antibodies are proteins produced in reaction to **antigens** (usually an invading viruses 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. The mouse either secretes the desired antibody or 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.

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 posttranslational modification of proteins is slightly different in insects than it is in mammals; therefore the use of insect expression systems may be of limited value. For the time being, insect expression systems are primarily used when small quantities of proteins are needed in research.

Protein Purification Methods: Downstream Processing

Once a protein is produced, downstream processing begins (Figure 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 it.

Step one: Preparing an extract for purification

It's helpful to note the relative volume of the extract. Often, the medium, or culture filtrate, harvested from a large fermenter or bioreactor is enough to fill a swimming pool, and the target protein represents less than 1 percent of that pool. Even if the protein is being expressed on a much smaller scale, finding the essential protein can seem like finding a needle in a haystack.

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), 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 and 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.

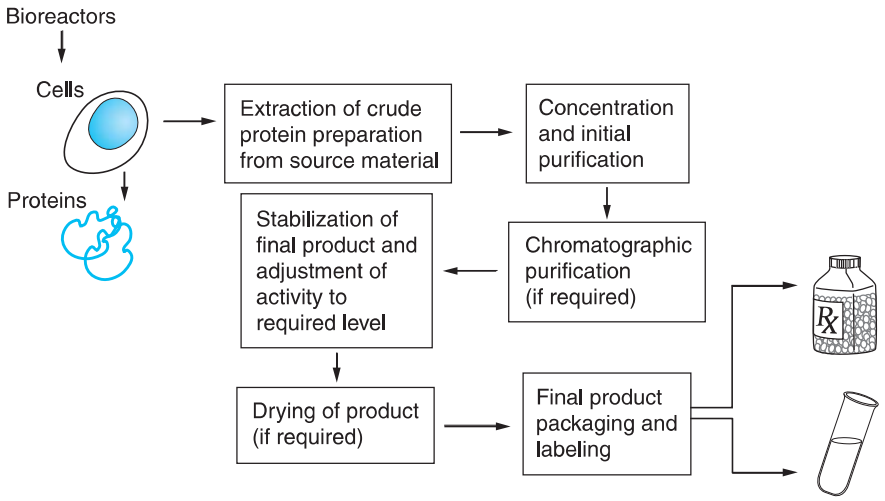


FIGURE 7 Basic Steps in Bioprocessing Purifications can be accomplished from raw materials or from bioreactors. The steps in the process must be devised and are often unique (and patentable).

From *Protein Biotechnology*, by Gary Walsh and Dennis Headon, Figure 3.1, p. 42. Copyright © 1994, John Wiley & Sons, Limited. Reproduced with permission.

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 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 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 9**, dialysis depends on the ability of some molecules to pass through semipermeable membranes 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.

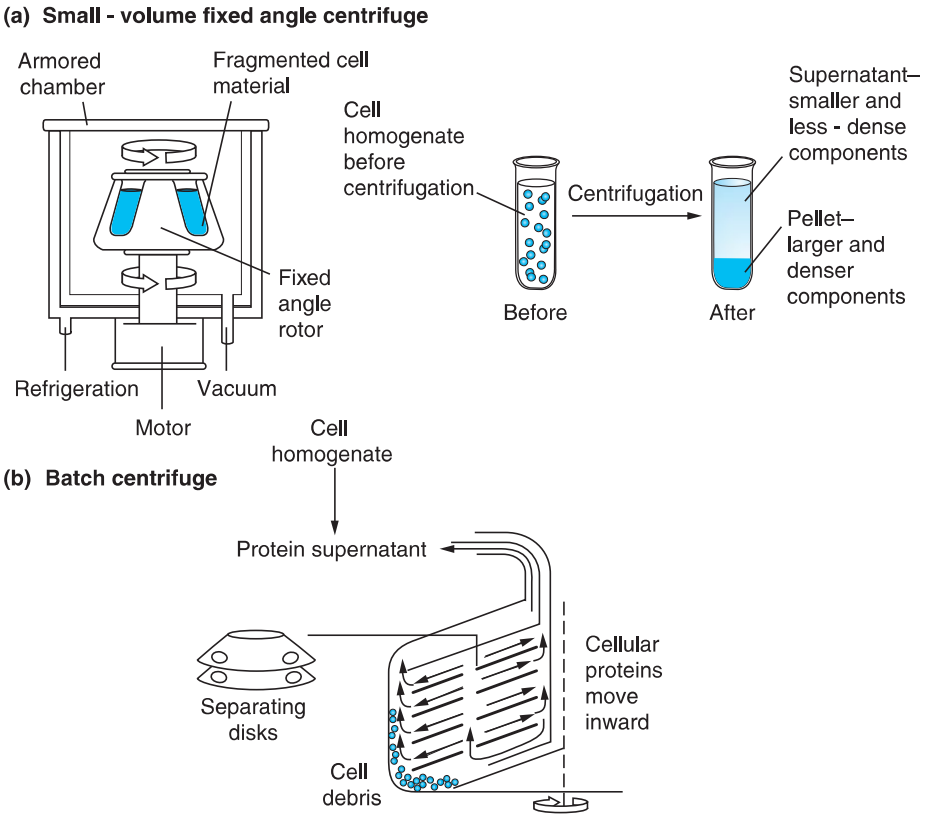


FIGURE 8 Fixed-Angle and Batch Centrifugation Fixed-angle centrifuges (a) can develop extremely high gravitational forces (*g* forces) but are limited to smaller quantities and must be run separately for each batch. Batch centrifuges (b) were developed to allow continuous flow of materials and separation of cell debris from cell proteins. The continuous pressure of the inflow and the centrifugal force can be adjusted to maximize the separation an outflow of the protein supernatant.

Chromatography methods allow us to sort 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.

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

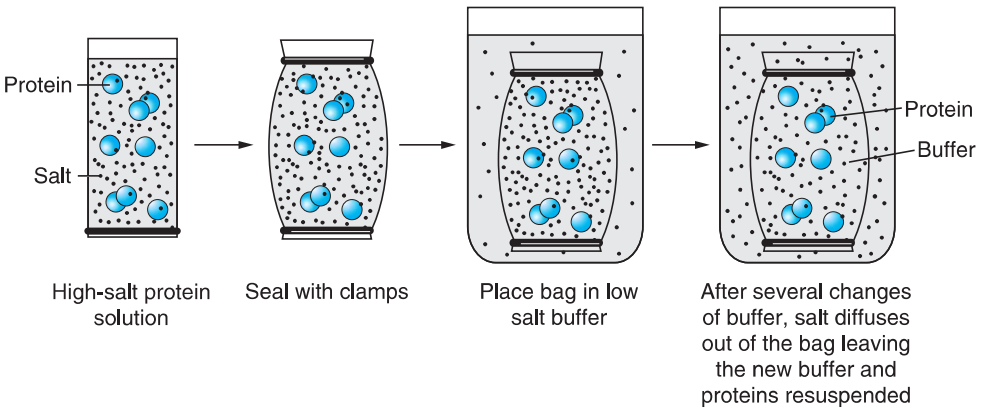


FIGURE 9 Dialysis Dialysis can be used to remove salt (large dots in figure), by the process of diffusion, and replace it with a buffer that is better suited for protein stability.

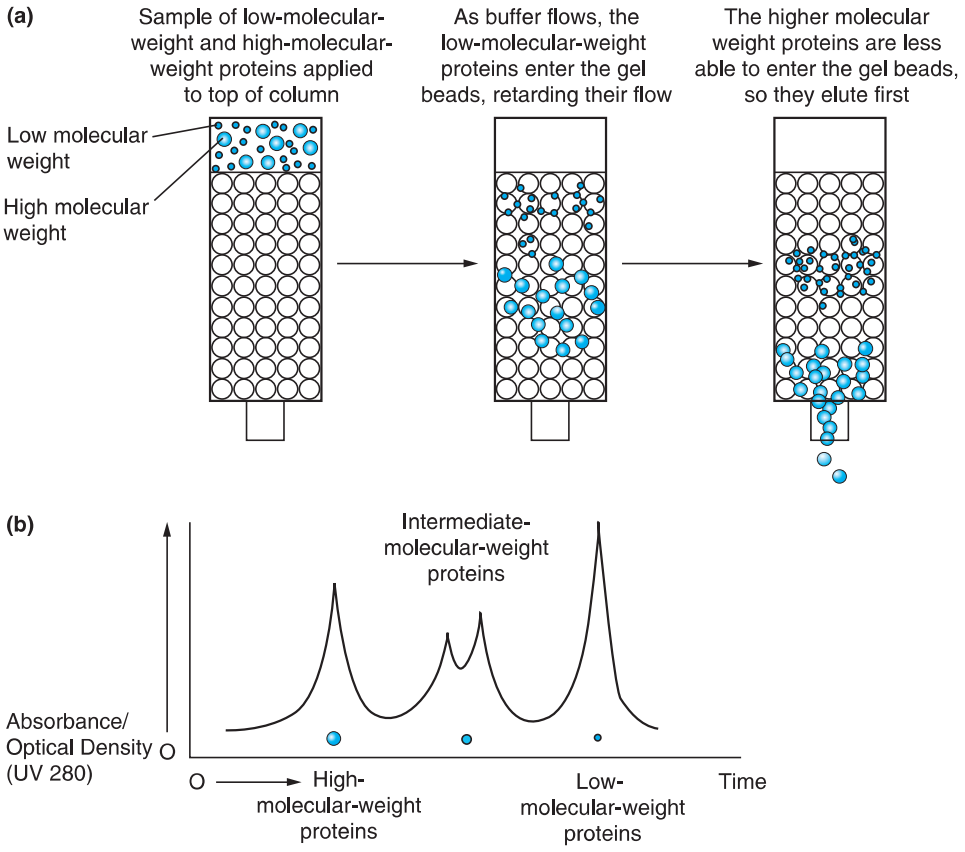


FIGURE 10 Protein Purification by Size Exclusion Chromatography (a) Low- and high-molecular-weight proteins travel through a size-exclusion chromatography column. (b) The diagram shows how the low and high molecular weight proteins appear as they come off the column when monitored for proteins (ultraviolet [UV] 280 absorption). 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.

From *Proteins to PCR* by David W. Burden, Figure 5.4, p. 99. Copyright © 1995 by Springer-Verlag. Reprinted with permission.

beads, as shown in Figure 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 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 12). After the proteins have bound to the resin beads, a buffer solution is used to wash out the unbound molecules. Finally, special buffer solutions are used to cause desorption (to break the ligand bonds) of the retained proteins. Fusion proteins, as mentioned earlier, can be used in affinity chromatography because the substrate (ligand) of the bacterial protein can be part of the affinity column, attracting the fused protein (bacterial enzyme protein) to the column. Affinity chromatography may shorten the purification process by reducing the number of steps.

As we have seen, amino acids are either attracted to or repelled by water molecules. In **hydrophobic interaction chromatography (HIC),** proteins are sorted on the basis of their repulsion of water. The column beads in HIC are coated with hydrophobic molecules, and the hydrophobic amino acids in a protein are attracted to the similar chemicals in the beads, shown in Figure 13.

Isoelectric focusing is often used in quality control during purification to identify two similar proteins that are difficult to separate by other means. Each protein has a specific number of charged amino acids on its surface in specific places. Because of this unique combination of charged groups, each protein has a unique electronic signature known as its **isoelectric point (IEP),** where the charges on the protein match the pH of the solution. The IEP can be used to separate

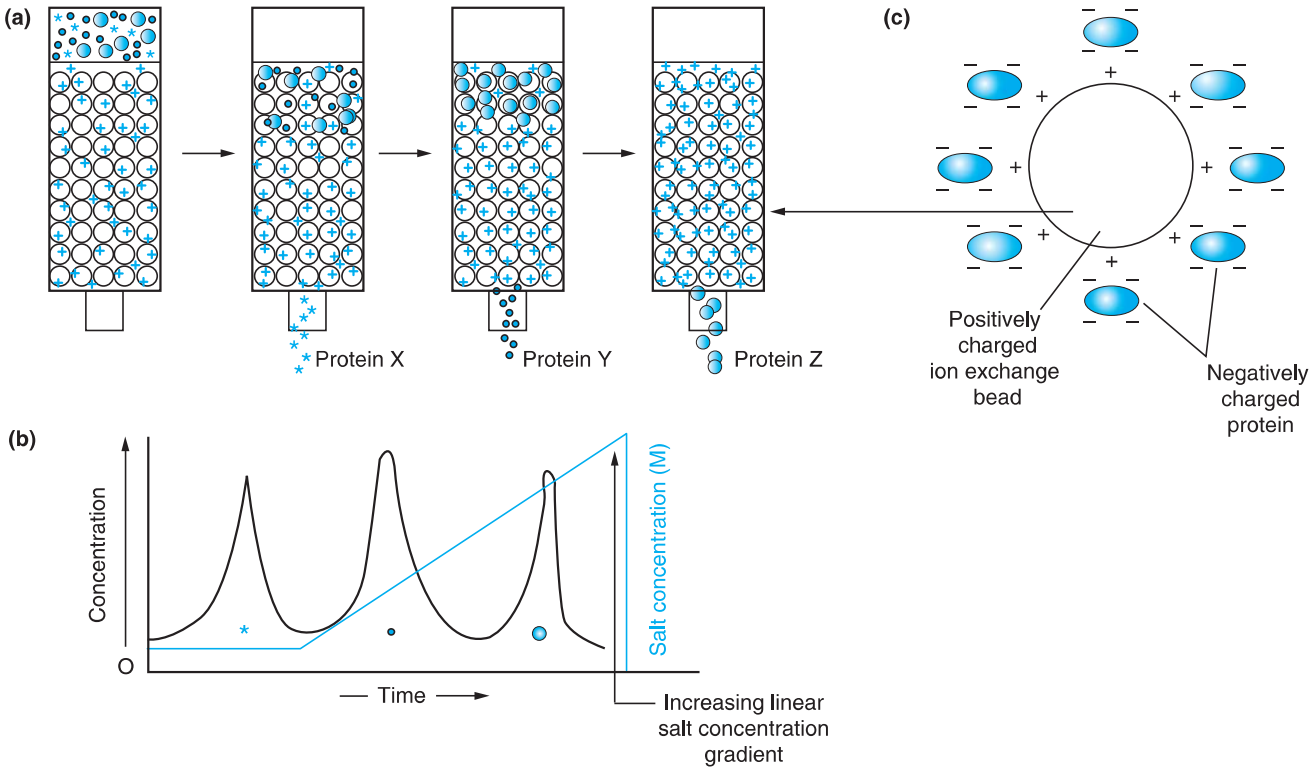


FIGURE 11 **Ion-Exchange Chromatography** (a) Charged amino acids bind to ionic resin beads. Increasing the ionic strength of the buffer displaces proteins (based on their binding strength) after they have bound to the column resin. (b) Protein X is released at the lowest concentration of the displacing salt gradient; protein Y has a higher binding strength and is released second; protein Z has the highest binding strength and requires a high salt concentration to displace it from the ion exchange beads of the column. (c) Anion-exchange resin is positive; cation-exchange resins are negatively charged.

From *Proteins to PCR* by David W. Burden, Figure 554, p. 102. Copyright © 1995 by Springer-Verlag. Reprinted with permission.

similar proteins from one another. Isoelectric focusing is the first dimension of **two-dimensional electrophoresis.**

Two-dimensional electrophoresis separates proteins based on their electrical charge and size. It is essentially the combination of two methods, IEP and gel electrophoresis. In this technique, researchers introduce a solution of cell proteins onto a specially prepared strip of polymer. When the strip is exposed to an electrical current, each protein in the mixture settles into a layer according to its charge. Next, the strip is aligned with a gel and again exposed to electrical current. As the proteins migrate through the gel, they separate according to their molecular weight.

Analytic methods

High-performance liquid chromatography (HPLC) adds a twist to the previously described chromatography

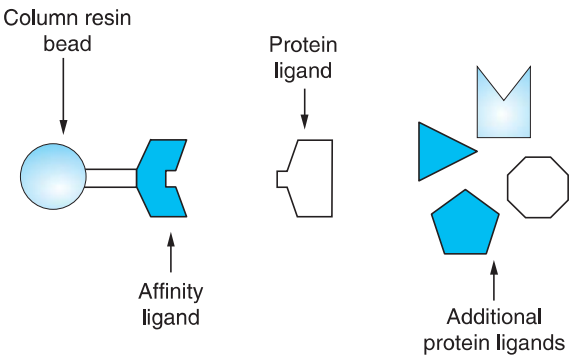


FIGURE 12 **Affinity Chromatography** Affinity ligands are designed to bind specifically to unique three-dimensional chemical components of the protein being purified. The non-bound proteins wash through the column. Increasing the ionic strength of the buffer can displace the bound protein (after preferential binding) and regenerate the affinity column. The displaced (pure) protein can be collected and concentrated.

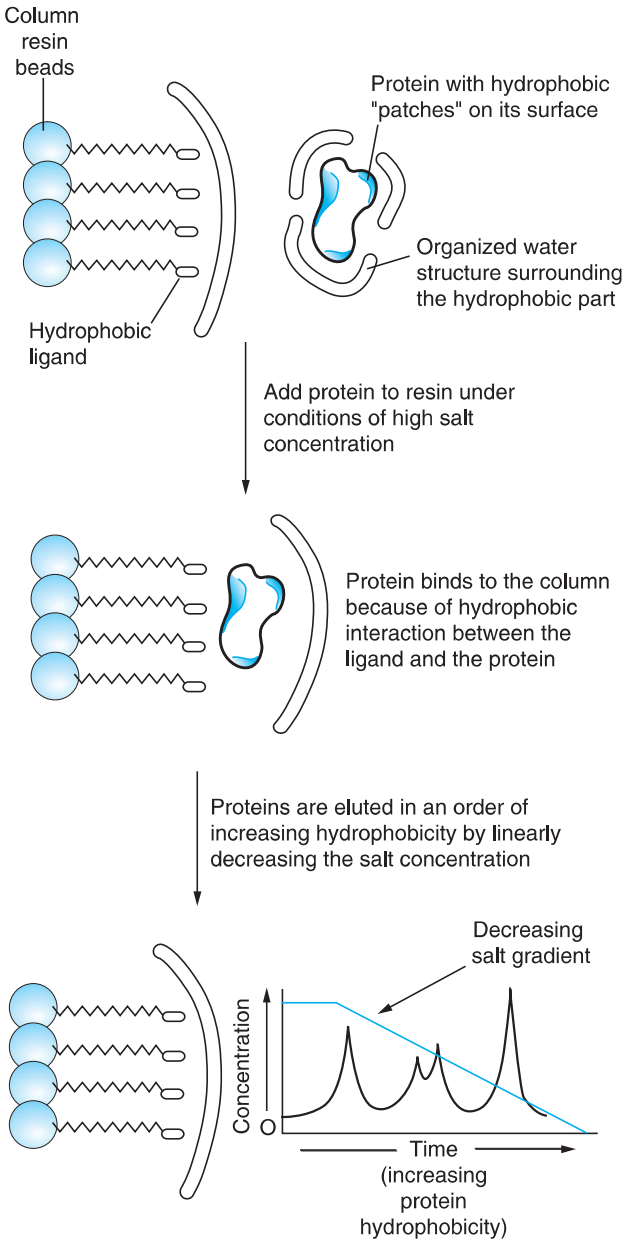


FIGURE 13 Hydrophobic Interaction Chromatography Increasing the nonpolar concentration of the buffer can cause the hydrophobic portions of proteins to combine with a hydrophobic ion-exchange-column resin. Further reducing the ionic concentration displaces the protein from the column resin and replaces its attachment with a nonpolar solvent. Fractions can be collected and protein concentration determined based on detection by spectrophotometric analysis at UV 280 nanometers (nm).

From *Proteins to PCR* by David W. Burden, Figure 5.6, p. 104. Copyright © 1995 by Springer-Verlag. Reprinted with permission.

methods, which depend on gravity flow or very low pressure pumps to move the extract through the columns. Such low-flow methods can take several hours to process a single sample. In contrast, HPLC systems use greater pressure to force the extract through the column in a shorter time. HPLC systems have limitations, however. Less protein is separated, so the technique is more useful in analytical situations than in mass production.

Mass spectrometry (mass spec) is a highly sensitive method used to identify small differences between proteins. In fact, it is used frequently on the outflow of HPLC systems. All mass spectrometers do three things: suspend the sample molecules into a charged gas phase, separate the molecules based on their mass-to-charge ratios by acceleration down a narrow tube, and finally detect the separated ions. A sample as small as one picogram (one billionth of a gram) can be analyzed by this process, which is illustrated in **Figure 14**. A definitive readout is produced, which indicates the identity and size of most of the proteins or fragments analyzed.

An important application of this process is in protein sequencing. A larger protein can be digested into fragments (peptides) and analyzed to determine the amino acid sequence. Mass spectrometry is now the preferred method and, in biotech companies, has largely replaced the slower Edmond end-group analysis method used for amino acid sequencing. Mass spectrometry can detect the difference between isomers of the same protein, and its capabilities are improving faster than most lab instruments in terms of critical analysis.

Verification

At each step of the purification process, it is important to verify that the target protein has not been lost and that concentration efforts have been successful. **SDS-PAGE** (polyacrylamide gel electrophoresis) is often used for verification. In this method, a detergent called sodium dodecyl sulfate (SDS) is added to a sample of the protein mixture and the mixture is heated. The sulfate charges are distributed evenly along the denatured (heated) protein, making separation dependent on the size of the protein. After this treatment, the protein sample is loaded onto a special gel matrix (PAGE) where it forms a single band at a specific location depending on its molecular size and mass, as seen in **Figure 15**.

By adding a dye that combines with proteins, such as **Coomassie stain,** a colored band results, and a known size marker can be compared with the stained sample. When the sample and the known marker are equivalent, we have evidence that the protein of interest is indeed being concentrated. Since this test is run at each step in the purification process, the colored band should become increasingly intense, proving that the protein has not been lost.

A specific detection method for proteins separated by SDS-PAGE is Western blotting (similar to Southern blotting). In Western blotting, protein is transferred

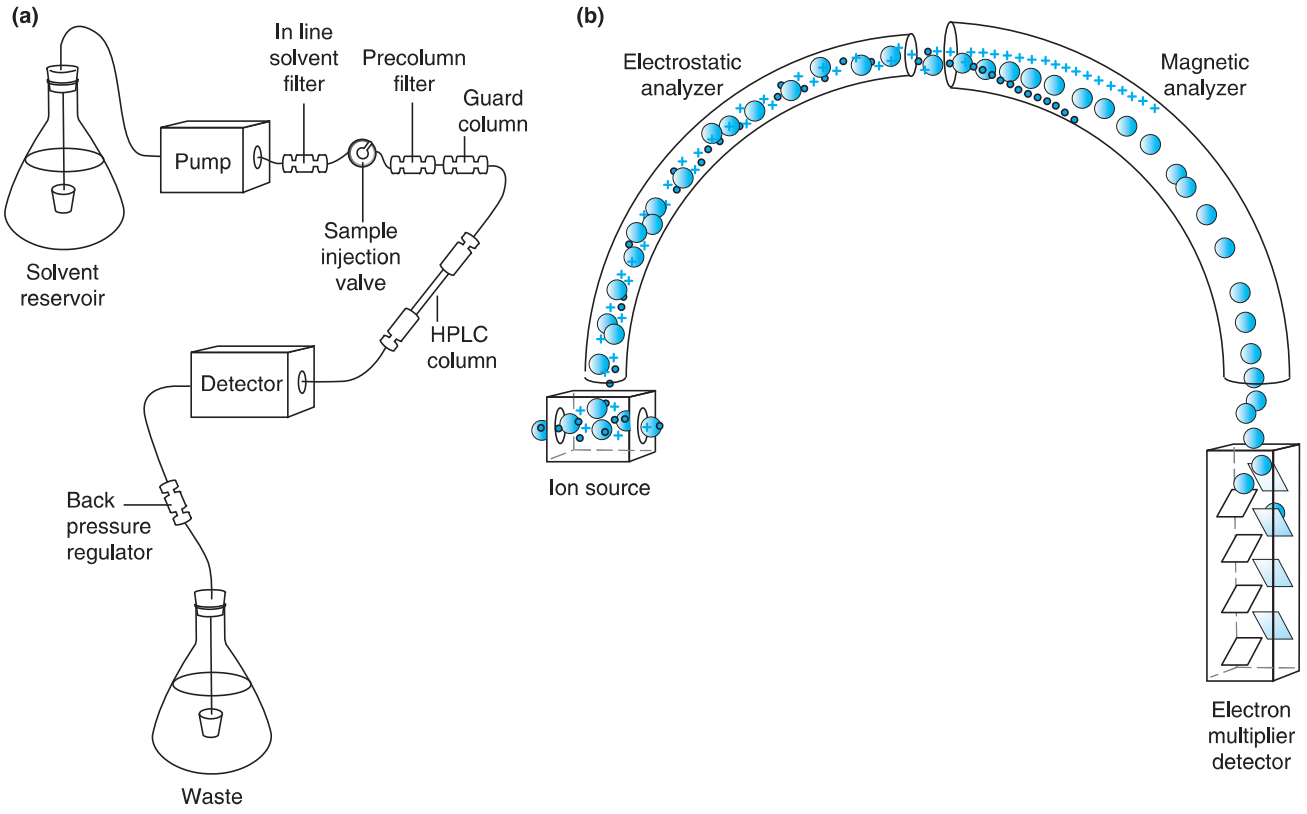


FIGURE 14 **High-Performance Liquid Chromatography and Mass Spectrometry** High-performance liquid chromatography is often coupled with mass spectrometry to analyze proteins. (a) HPLC uses noncompressible resin beads under very high pressures to separate proteins that are often similar in size. (b) Mass spectrometry follows this initial separation to detect subtle differences in ionized and accelerated proteins that will be analyzed (by distance or time) as they travel down a vacuum-filled tube.

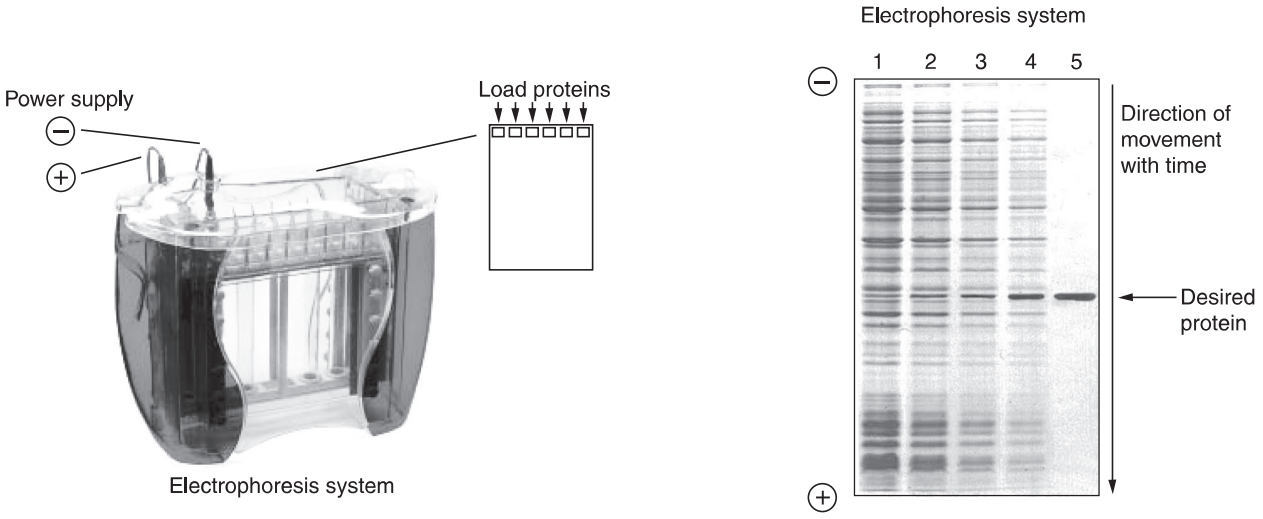


FIGURE 15 **SDS-PAGE** Polyacrylamide gel electrophoresis of proteins that have uniform charges (produced by heating in sodium dodecyl sulfate) can be separated by size, based on their migration in a vertical electrical field. This procedure commonly follows each major step in purification to verify that the band of desired protein is becoming more concentrated (and has not been lost in the separation). Notice that the 65-kilodalton (kD) protein has the highest concentration in the final step (lane 2) of this procedure (lane 1) when proteins of known size are used for comparison. The smallest proteins will reach the + pole first.

Amersham Biosciences.

CAREER PROFILE

Positions in Protein Production

Manufacturing assistant (MA) is an entry-level position in protein production that requires a certificate or associate's degree in biotechnology, biology, microbiology, biochemistry, or biomanufacturing. Employers expect at least 2 years of experience in a related area such as an MA's work in a biomanufacturing plant using bioreactors or fermenters. Some fermenters are used to produce food products or enzymes for industrial processes. The duties of the job vary depending on the type of manufacturing being done. Some MAs are involved in manufacturing medical devices or media for research and the development of products (usually proteins). MAs usually work in a clean-room environment. They weigh and measure chemicals and raw materials used in the manufacturing process. In aseptic fill positions, MAs set up and operate equipment that fills and packages purified products. Purification of products for fill (of containers) is a common function and requires an understanding of the purification equipment and how to troubleshoot it. MAs must maintain accurate and detailed records and have a working knowledge of government regulations that apply to their production activities.

Manufacturing technicians are usually promoted from MA positions after they demonstrate experience and aptitude. Their salaries are higher than those of MAs. This position may require a bachelor's degree in a life science (in addition to experience as an MA). An MA in the same company may advance to this position by returning to school to obtain a bachelor's degree. Companies often value this type of education and may arrange work schedules or even provide subsidies for further education.

Bioprocessing engineer is the most common title in manufacturing for a doctoral-level scientist with a degree in bioengineering and a thorough understanding of the processes and sequences used to produce and purify products in biotechnology.

from a gel to a nitrocellulose membrane and detected with a specific antibody that recognizes that protein by its unique structure. An enzyme attached to the antibody can be used to convert a fluorescent substrate so as to produce a permanent record of the detection. In this procedure, an electric current is applied to the gel. The separated proteins then migrate through the gel and onto the membrane in the same pattern as they separate on the SDS-PAGE. All sites on the membrane can then be blocked in such a way that antibody (serum) will not bind to them nonspecifically to detect the antigen blotted on the membrane; a primary antibody (serum) is then added at an appropriate dilution and incubated with the membrane. If there are any antibodies present that are directed against one or more of the blotted antigens, those antibodies will bind to the protein(s) while other antibodies will be washed away at the end of the incubation. To detect the antibodies that have bound, anti-immunoglobulin antibodies coupled to a reporter group, such as the enzyme alkaline phosphatase, are added. Finally, after excess second antibody is washed free of the blot, a substrate is added that precipitates upon reaction with the conjugate, resulting in a visible band where the primary antibody bound to the protein (see Figure 15).

Antibody detection of a specific protein can also be carried out using an enzyme-linked immunosorbent assay (ELISA). The most commonly used ELISA requires two antibodies: one to capture the unique protein and one attached to an enzyme to produce a color reaction (see Figure 12). ELISA occurs on a multiwell plate using affinity chromatography. The first antibody to the unique protein is plated on to a multiwell ELISA plate, the protein is added, and—after a series of wash and blocking steps—the second antibody (enzyme attached) is added. The addition of a substrate permits a color reaction to occur if the antibody has bound. ELISA plates are designed to capture many important proteins currently under research.

Preserving Proteins

After the target protein has been isolated, collected, and purified, it must be saved in a manner that will preserve its activity until it can be used. One means of preserving the protein is **lyophilization,** or freeze-drying. In this process, a purified liquid protein solution is frozen. A vacuum is then used to hasten the evaporation of water from the fluid. In lyophilization, ice crystals are converted directly to water vapor without melting into liquid water first. The containers of freeze-dried material are sealed after the water is removed, leaving the dried proteins behind. Lyophilization has been demonstrated to maintain protein structure, making it a commonly chosen method for the preservation of biotechnologically derived proteins. Many freeze-dried proteins may be stored at room temperature for prolonged periods of time.

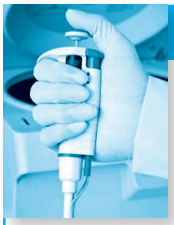
Scaling Up Protein Purification

Protein purification protocols are usually designed in the laboratory on a small scale. These techniques at this level of productivity are feasible if the product is needed in only small amounts. Orders of monoclonal antibodies, for example, are in the range of grams per year; a relatively small demand. A single laboratory-scale bioreactor can usually produce adequate quantities. However, other proteins are required in much larger quantities, which means that the production methods must be scaled up.

Scaling up is not always easy to accomplish. Laboratory methods that may work on a small scale may not always be adaptable to large-scale production. Furthermore, changes in the purification process can invalidate earlier laboratory-scale clinical studies. For example, when the FDA approves a bioengineered protein, it also approves the process for producing it. To change the process, it may be necessary to seek FDA approval once again. For this reason, bioprocess engineers are involved in the early stages of purification and work to ensure that it will be possible to scale up the process later on.

Postpurification Analysis Methods

During research, it is often helpful to look closely at a purified protein. The goal may be to better understand the molecular structure of a specific protein or to alter its structure to change the protein's function. The following two methods are used in the postpurification analysis of proteins.



TOOLS OF THE TRADE

Piecing Together the Human Proteome with Protein

Proteomes, the collection of proteins associated with a specific life function, have become more important since the discovery of the human genome. As mentioned earlier, the cost of bringing a drug to market is about $500 million and takes between 5 and 8 years. Because most of the drugs produced by the biotechnology industry are proteins (such as growth factors, antibodies, and synthetic hormones) that replace missing or nonfunctional proteins in humans, companies are particularly interested in developing less expensive and faster methods of detecting how proteins function in disease detection and treatment—specifically, new protein microarrays. Like DNA microarrays, these miniature devices detect proteins associated with disease and those present in abnormal concentrations. As new protein structures are identified, new microarrays are constructed.

Protein microarrays have many uses. Most current biotech drugs function at the protein level, interacting with receptors, triggering events, and targeting other proteins in cells of the body. The detection of biomarker proteins using microarrays has improved the monitoring of drug action in many diseases. For example, we have all experienced the benefit of vaccines: proteins that stimulate our immune systems to recognize disease organisms without contracting the disease. Antibody production in response to vaccination is an example of the effect of a protein biomarker and can indicate positive immunity.

Protein microarrays are usually constructed of glass slides coated with a material that binds proteins. Attached to each slide is an antibody that is specific to the protein to be detected as well as a signaling mechanism that indicates capture of the protein. Cambridge Antibody Technology (in the United Kingdom) has a library of more than 100 billion antibodies collected from the blood of healthy individuals. Packard BioScience (in Connecticut) has developed a protein microarray that allow the attached protein to maintain its three-dimensional shape while embedded in the coating material. Ciphergen Biosystems (in California) has adapted mass spectroscopy to perform rapid on-chip detection and analysis of proteins directly from biological samples. So, what is left to do?

There are countless proteins from numerous diseases yet to be discovered. The ability to diagnose a disease and determine the most effective treatment will depend on the ability to purify proteins and develop antibodies that are related to the disease. The opportunity to change human disease conditions that were previously incurable depends on discovering and purifying the important proteins of life's proteome. Take the time to access the websites of some of the companies that develop these devices to keep up with this important technology (e.g., www.ciphergen.com, www.packardbioscience.com, or www.biacore.com).

Protein sequencing

Recall from our discussion of protein structures that each protein has a specific sequence of amino acids, known as the primary sequence. To understand a protein completely, it is important to determine its primary sequence. Automated protein sequencers make this task possible (see Figure 14). In the mass spectrometry method, peptide masses are identified by their unique signatures (retention times). By converting peptides into ions and subjecting them to acceleration in a vacuum, it is possible to identify many unique proteins in a matter of minutes.

X-ray crystallography

X-ray crystallography is used to determine the complex tertiary and quaternary structures of proteins. The method requires pure crystals of proteins that have been carefully dehydrated from solutions. Bombarded with x-rays, a pure protein creates specific shadow patterns based on its configuration. Computer analysis of these patterns allows the generation of "ribbon" diagrams, which not only describe the structure of the proteins but also make it possible to plan potential modifications of the protein molecule to improve function. Protein crystallography is usually a requirement for approval by the FDA because it verifies that the process produces the product being approved.

4 Proteomics

Many diseases are the result of flaws in protein expression, but not all of those diseases can be understood simply in terms of genetic mutation. Because proteins undergo posttranslational modification, the puzzle can be far more complicated. A new scientific discipline, **proteomics,** is dedicated to understanding the complex relationship of disease and protein expression.

In proteomics, **proteomes** (the PROTEin complement to a genOME) are compared between healthy and diseased states. The variations of protein expression are then correlated to the onset or progression of a specific disease. The goal of proteomic research is the discovery of protein markers that can be used in new diagnostic methods and the development of targeted drugs for treating disease. For example, scientists at the University of California, Davis, purified the protein produced by BRCA2, an oncogene linked to breast cancer. The researchers were then able to synthesize the protein to study BRCA2's role in DNA repair. BRAC2 is susceptible to the drug trastuzumab (Herceptin; Genentech, San Francisco), which has increased the breast cancer survival rate to nearly 70 percent. In this way, only patients with BRAC2 biomarkers receive trastuzumab, saving discomfort and unnecessary treatment of all breast cancer patients.

Although the BRCA2 gene was discovered in 1994, purifying the protein made by the gene proved difficult, since it is very large, does not express well, and degrades easily. The U.C. Davis researchers tested many different cell lines and succeeded in introducing a BRCA2 gene into a human cell line and expressing it as a whole protein. Other researchers used genetic engineering techniques to manufacture the human protein in yeast. They then tested the purified protein for its function in repairing damaged DNA. Experiments with the BRCA2 protein confirmed that it plays a role in repairing damaged DNA, and when it is damaged in breast cancer, the DNA fails to function. Research continues on the function of BRCA2 and other proteins involved in breast cancer, but this progress would not have been possible without proteomics.

Proteomics applies several techniques described earlier: two-dimensional electrophoresis is used to separate proteins, mass spectrometry verifies the protein's identity, and the protein is characterized using amino acid sequencing. Some of this labor-intensive work may be assisted by automation in the future. A **protein microarray** is a set of proteins immobilized on a surface—usually a glass slide—which has been coated with a reagent that will indicate binding by color change (**Figure 16**). The capacity of these arrays has increased but does not rival that of DNA arrays because of the difficulties of production. A single DNA

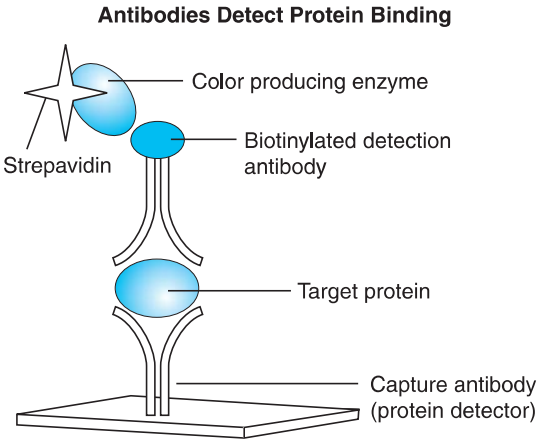


FIGURE 16 Protein Microarrays protein microarrays that uniquely bind to single proteins are being perfected. When a protein binds, it releases a fluorescent signal. The ability to detect the presence of unique proteins (proteomics) will permit proper decisions in disease diagnosis and therapy to be made.

from "Antibodies Detect Protein Building" figure 4.17, p. 44 from *Genomics & Proteomics* November/December 2001. Copyright © 2003 *Genomics & Proteomic,* a publication of Reed Business Information, a division of Reed Elsevier Inc. All rights reserved. Reprinted by permission.

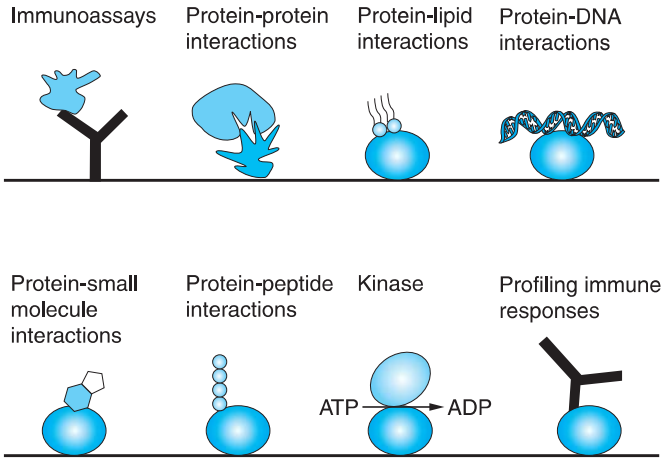


FIGURE 17 Protein Microarrays Can Detect More Than Just Proteins The capacity of protein microarrays has increased. Functional protein microarrays have recently been applied to the discovery of protein interactions (i.e., protein-protein, protein-lipid, protein-DNA, protein-drug, and protein-protein interactions), expanding knowledge of the significance of protein interactions to normal cell function. Binding antibodies or other proteins to the matrix allows detection when the binding occurs.

array can monitor the expression of an entire genome. Functional protein microarrays have recently been applied to the discovery of protein interactions: protein-protein, protein-lipid, protein-DNA, protein-drug, and protein-protein interactions (**Figure 17**).

MAKING A DIFFERENCE

As we have seen, disease diagnosis and treatment have been vastly improved by protein purification technology. Despite this, there are still problems to be solved: biological fluids (plasma, sera, urine, and saliva) have a wide range of different protein concentrations, and high concentrations of some proteins mask the presence of others—like disease-specific proteins called biomarkers. These low-concentration biomarker proteins are often degraded by natural enzymes or the purification process itself. Ceres Nanosciences of Virginia is applying nanotechnology to production purification to address this problem. The Nanotrap uses porous nanoparticles to attract specific biomarkers The nanoparticles exclude other proteins owing to the small size of the pores on the surface that they must penetrate. Nanotrap technology adds another dimension to the protein purification process, permitting the detection of proteins that "mark" the presence of a disease. Nanotechnology is making a difference in disease treatment.

QUESTIONS & ACTIVITIES

Answers can be found at the end of the chapter.

1. What can the public database on proteins tell us?
2. How does directed molecular evolution technology differ from mutations that occur naturally?
3. How has an understanding of protein structure benefited from the results of the Protein Structure Initiative?
4. Why are proteins being primarily searched through cDNA products?
5. If organisms produce proteins on their own, why should companies be allowed to patent proteins?
6. If you are purifying a small protein from an SEC column, which fractions do you want to collect?
7. Do proteins strongly bound to an IonX column elute first or last from the column?
8. Is affinity chromatography more or less selective at separating proteins than IonX chromatography?
9. List the protein separation methods primarily used in analytic (rather than production) methods.
10. If you carry out an SDSPAGE analysis after each step in a protein purification sequence and found that the last step resulted in the lack of a band at the location you expected, what would this mean?

Visit www.pearsonhighered.com/biotechnology

To download learning objectives, chapter summary, "Keeping Current" web links, glossary, flashcards, and jpegs of figures from this chapter.