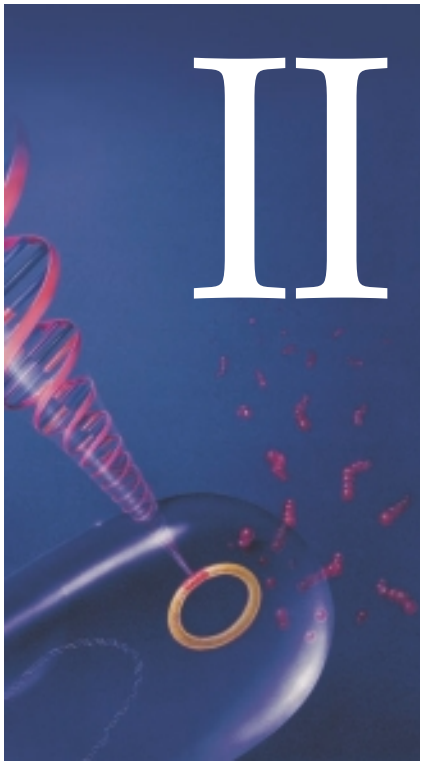


II MOLECULAR BIOTECHNOLOGY OF MICROBIAL SYSTEMS

With the advent of recombinant DNA technology, many of the properties of microorganisms that might be useful in a variety of applications can be more readily exploited. In part II, we examine some of the uses for genetically engineered microbial systems.

Currently, bacteria are being genetically manipulated to act as biological factories for the production of pharmaceutical proteins, nucleic acid therapeutic agents, restriction endonucleases, chemical compounds, amino acids, antibiotics, and biopolymers. In some applications, cloned genes have been introduced into bacterial host cells to create novel biosynthetic pathways that produce novel metabolites. Genes and DNA fragments from pathogenic organisms have been isolated and used as probes for the diagnosis of disease in both animals and humans. In other instances, isolated genes and DNA fragments have been used to produce safer and more efficacious vaccines.

Genetic manipulation of microbial systems also entails enhancing the natural ability of certain bacterial strains to carry out specific biological processes. For example, researchers have developed bacterial strains that can degrade environmental pollutants, improve the growth of plant crops, degrade cellulosic biomass into utilizable low-molecular-weight compounds, and prevent the proliferation of specific insect pests.

It is often assumed that the growth of large quantities of microbes is a routine procedure. Successful large-scale production of proteins synthesized

by recombinant microorganisms, however, requires that many different factors be controlled during both the growth phase of the microorganism and the purification process to ensure that high yields of a pure product are obtained.

9 Molecular Diagnostics

THE SUCCESS OF MODERN MEDICINE and agriculture often depends on the ability of workers in these fields to detect the presence of specific viruses, bacteria, fungi, parasites, proteins, and small molecules in humans, animals, plants, water, and soil. For example, the prevention, control, or treatment of infectious disease is generally facilitated by the early and accurate identification of the causative pathogenic organism. Many of these detection procedures require the growth in culture of the potential pathogen and then the analysis of a spectrum of physiological properties that facilitate its identification. Although tests of this type are effective and reasonably specific, they are often slow and costly. These constraints apply to the identification of both bacterial and parasitic (Table 9.1) organisms. In addition, if the pathogenic organism does not grow well or cannot be cultivated, the opportunity to detect the disease-causing organism is severely limited. For example, *Chlamydia trachomatis,* an obligately intracellular bacterium, causes a sexually transmitted disease prevalent in North America and Europe. Clinical diagnosis of chlamydial infection is difficult, because long-term cell culture is required. Frequently, false-negative results (i.e., the diagnosis of the absence of the organism is erroneous) are obtained, and consequently, adequate treatment procedures are not implemented. Certainly, if growth were required for detection, then at best only a few of all known pathogenic organisms could ever be routinely identified. To overcome this major constraint, molecular diagnostic procedures using either immunological or DNA detection methodologies have been devised.

In general, any useful detection strategy must be specific, sensitive, and simple. Specificity means that the assay must yield a positive response for only the target organism or molecule. Sensitivity means that the diagnostic test must identify very small amounts of the target organism or molecule, even in the presence of other potentially interfering organisms or substances. Simplicity is required for the test to be run efficiently, effectively, and inexpensively on a routine basis.

TABLE 9.1 A comparison of some of the methods used to diagnose parasite infection

Method	Advantages	Disadvantages
Microscopic examination	Simple Direct detection of parasite Differentiates morphologically distinct organisms	Slow, laborious, and tedious Low sensitivity Cannot discriminate between similar organisms Requires a high skill level
In vitro culture and mouse inoculation	Detects only viable parasites Measures virulence and infectivity	Slow and expensive Different strains show a range of responses Parasite may lose its viability in the specimen Uses animals
Detection of antibodies in serum	Simple and fast Automatable Can be used to screen a large number of samples	Not always specific Does not distinguish between active and latent infections
DNA hybridization and PCR	Fast, sensitive, and specific Detects parasite directly Can distinguish different species Independent of previous infections Parasites need not be viable Automatable	Expensive and multistep Does not distinguish between live and dead organisms Possible false positives and false negatives

Adapted from Weiss, *Clin. Microbiol Rev.* **8:**113–130, 1995.

It is estimated that worldwide sales of immunodiagnostics accounted for approximately $7.7 billion in 1999, and this figure continues to increase by 5 to 10% per year. The market for DNA-based diagnostic procedures was around $500 million in 1999 and is increasing at around 20 to 30% per year, so that in 2004 it was worth approximately $2 billion. In this chapter, the principles behind some of these molecular diagnostic procedures and the use of these procedures for a variety of applications are discussed.

Immunological Diagnostic Procedures

Many immunological detection methods are sensitive, specific, and simple. They can be used for a wide range of applications, including drug testing, assessment and monitoring of various cancers, detection of specific metabolites, pathogen identification, and monitoring infectious agents. However, there are limitations. For example, if the target is a protein, then the use of antibodies requires that the genes contributing to the presence of the target site be expressed and that the target site not be masked or blocked in any way that would prevent the binding of the antibody.

In principle, traditional diagnostic procedures for infectious agents rely on either a discrete set of traits characteristic of the pathogenic agent or, preferably, one unique, readily distinguishable feature. The clinical microbiologist searches for the smallest number of biological characteristics that

can, with complete certainty, reveal the presence and precise identity of a pathogenic agent. For example, some infectious agents produce distinctive biochemical molecules. The problem is how to determine when the identifying component is present in a biological sample. Often, such a marker molecule can be identified directly in a specialized biochemical assay that is very specific for the marker molecule. The problem with this approach is that it can potentially lead to a proliferation of highly individualized detection systems for different pathogenic organisms. A standardized method of identifying any key marker molecule, regardless of its chemical nature, is preferred. Because antibodies bind with high specificity to discrete target sites (antigens), assays based solely on identifying specific antibody–antigen complexes have abolished the need to devise a unique identification procedure for each particular marker molecule.

ELISA

There are a number of different ways to determine whether an antibody has bound to its target antigen. The enzyme-linked immunosorbent assay (ELISA) is one method, and it is frequently used for diagnostic detection. The ELISA procedure may be either indirect (Fig. 9.1A) or direct (Fig. 9.1B). A generalized indirect ELISA protocol (Fig. 9.1A) has the following steps.

1. Bind the sample being tested for the presence of a specific molecule or organism to a solid support, such as a plastic microtiter plate, which usually contains 96 sample wells. Wash the support to remove unbound molecules.
2. Add a marker-specific antibody (primary antibody directed against the target antigen) to the bound material, and then wash the support to remove unbound primary antibody.
3. Add a second antibody (secondary antibody) that binds specifically to the primary antibody and not to the target molecule. Bound (conjugated) to the secondary antibody is an enzyme, such as alkaline phosphatase, peroxidase, or urease, that can catalyze a reaction that converts a colorless substrate into a colored product. Wash the mixture to remove any unbound secondary antibody–enzyme conjugate.
4. Add the colorless substrate.
5. Observe or measure the amount of colored product.

If the primary antibody does not bind to a target site in the sample, the second washing step removes it. Consequently, the secondary antibody–enzyme conjugate has nothing to bind to and is removed during the third washing step, and the final mixture remains colorless. Conversely, if the target site is present in the sample, then the primary antibody binds to it, the secondary antibody binds to the primary antibody, and the attached enzyme catalyzes the reaction to form an easily detected colored product. Since secondary antibodies that are complexed with an enzyme are available commercially, each new diagnostic test requires only a unique primary antibody. In addition, several secondary antibody molecules, each with several enzyme molecules attached, bind to one primary antibody molecule, thereby amplifying the intensity of the signal.

With a direct ELISA protocol (Fig. 9.1B), a monoclonal antibody specific for the target antigen is first bound to the surface of the microtiter plate. To

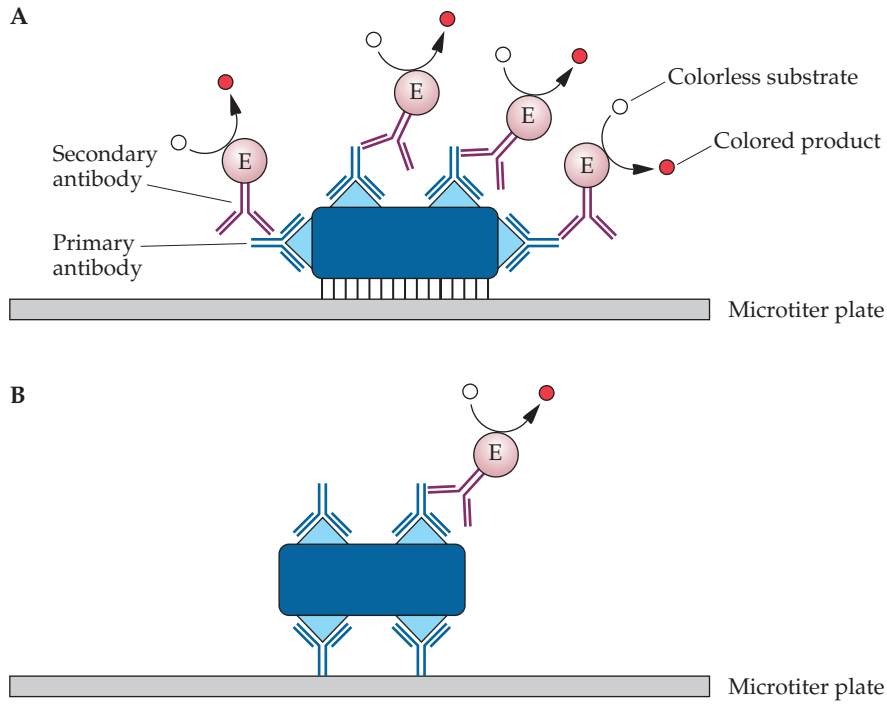


FIGURE 9.1 Generalized ELISA protocol for detecting a target antigen. The primary antibody is often obtained from rabbits that have been immunized with the target antigen, while the secondary antibody is from goats immunized with rabbit antibodies. The enzyme (E) is conjugated to the secondary antibody. **(A)** Indirect ELISA; **(B)** direct ELISA.

assess the amount of a particular antigen in a sample, the sample is added to the well of the microtiter plate and allowed to interact with the bound antibody. This is followed by a wash to remove any unbound molecules. Then, the primary antibody and the secondary antibody conjugated to an enzyme are added, as described above, before the presence of bound antigen is visualized.

The principal feature of an ELISA system is the specific binding of the primary antibody to the target site. If the target molecule is, for example, a protein, then a purified preparation of this protein is generally used to generate the antibodies that will be used to detect the target. The resulting antibody mixture, which is found in the serum (antiserum) of an inoculated animal, usually a rabbit, contains a number of different antibodies that would each bind to a different antigenic determinant (epitope) on the target molecule. Such a mixture of antibodies is called a polyclonal preparation. For some diagnostic assays, the use of polyclonal antibodies has two drawbacks: (1) the amounts of the different antibodies within a polyclonal preparation may vary from one batch to the next, and (2) polyclonal antibodies cannot be used to distinguish between two similar targets, e.g., when the difference between the pathogenic form (target) and the nonpathogenic one (nontarget) is a single determinant. However, these problems can be overcome, because it is now possible to generate an antibody preparation that is directed against a single antigenic determinant, namely, a monoclonal antibody. Also, despite these drawbacks, diagnostic assays employing polyclonal antibodies are widely used for a variety of purposes.

Monoclonal Antibodies

In mammals, a complex set of cellular systems has evolved to protect the body from toxic substances and invasion by infectious agents. As part of the defensive response, cells of the lymphatic system can be induced to produce specific proteins (antibodies) that bind to foreign substances (antigens) and—with the help of other immune system proteins, including the complement system—neutralize their biological impact. In response to an immunological challenge, each antibody-producing cell synthesizes and secretes a single antibody that recognizes with high affinity a discrete region (epitope, or antigenic determinant) of the immunizing antigen. Because an antigen generally has several different epitopes, normally several cells of the immune system each produce a different antibody against one of the many epitopes of the antigen. Such a set of antibodies, all of which react with the same antigen, is designated a polyclonal antibody (Fig. 9.2).

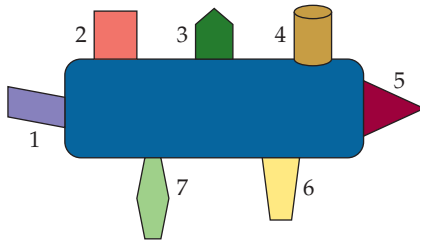


FIGURE 9.2 Schematic representation of a target antigen. The surface of the antigen depicted has seven (numbered 1 to 7) different antigenic determinants (epitopes). When this antigen is used to immunize an animal, each antigenic determinant elicits the synthesis of a different antibody. Together, the different antibodies that interact with an antigen constitute a polyclonal antibody directed against that antigen.

Early in the 20th century, although the polyclonal nature of antibodies was not appreciated, it was realized that antibody specificity could be used to prevent infections. Later, antibodies were used as diagnostic agents to determine the presence of toxic substances in clinical samples. Unfortunately, the effectiveness of a polyclonal antibody preparation varies from batch to batch because, in some immunizations, certain antigenic determinants of a particular antigen are strong stimulators of antibody-producing cells, whereas at other times, the immune system responds more actively to different epitopes of the same antigen. Also, individual animals often respond differently to a particular antigen. This variation can affect the abilities of different preparations to neutralize antigens because different epitopes have different potencies (stimulating abilities). Hence, one batch of polyclonal antibody may have a low level of antibody molecules directed against a major epitope and not be as effective as a previous antibody preparation.

Consequently, a fundamental objective for the applied use of antibodies, as diagnostic agents or as components of therapeutic agents, was to discover how to create a cell line that could be grown in culture and that would produce a single type of antibody molecule (monoclonal antibody) with a high affinity for a specific target antigen. Such a cell line would provide a consistent and continuous source of identical antibody molecules. Unfortunately, the B lymphocytes (B cells) that synthesize antibodies do not reproduce in culture. However, it was envisioned that a hybrid cell type could be created to solve this problem. This hybrid would have the B-cell genetic components for producing antibodies and the cell division functions of a compatible cell type to enable the cells to grow in culture. It was known that normal B lymphocytes sometimes become cancer cells (myelomas) that acquire the ability to grow in culture while retaining many of the attributes of B cells. Thus, myeloma cells, especially those that did not produce antibody molecules, became candidates for fusion with antibody-producing B cells. In the mid-1970s, these ideas became reality.

Formation and Selection of Hybrid Cells

The initial step leading to the production of a hybrid cell line that produces a single antibody entails the inoculation of mice with an antigen. After

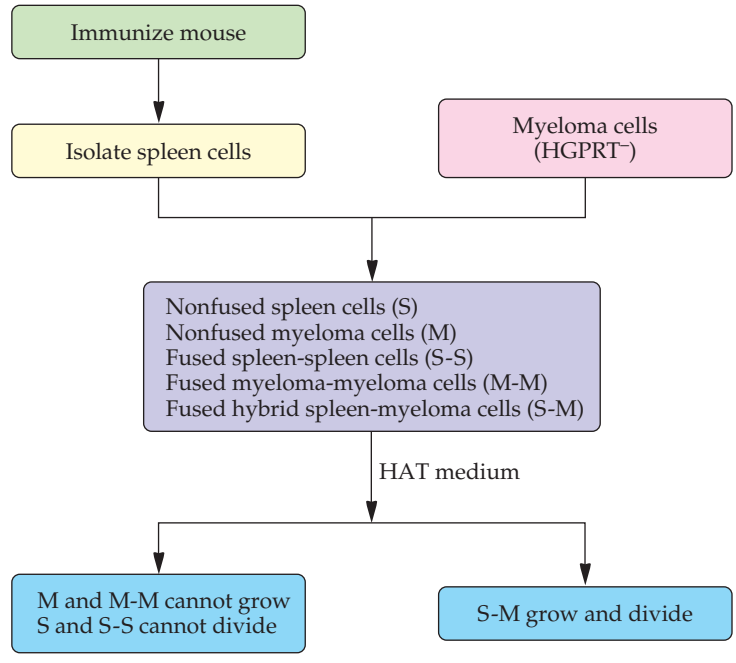


FIGURE 9.3 The HAT procedure for selecting hybrid spleen–myeloma (hybridoma) cells.

several inoculations over a period of a few weeks, the animals are tested, generally using an ELISA or similar system, to determine whether they have developed an immune response. If they have, they are killed and their spleens are removed, washed, minced, and gently agitated to release individual cells, some of which are antibody-producing B cells. The splenic cell suspension is mixed with a suspension of myeloma cells that are genetically defective for the enzyme hypoxanthine-guanine phosphoribosyltransferase ($HGPRT^-$). The combined cell suspensions are mixed with 35% polyethylene glycol for a few minutes and then transferred to a growth medium containing hypoxanthine, aminopterin, and thymidine (HAT medium).

The polyethylene glycol treatment facilitates fusion between cells. Nevertheless, the fusion events are rare and random. There will be myeloma cells, spleen cells, myeloma–spleen fusion cells, myeloma–myeloma fusion cells, and spleen–spleen fusion cells in the mixture. The HAT medium, however, allows only the myeloma–spleen fusion cells to grow, because none of the other cell types can proliferate in this medium. Unfused spleen cells and spleen–spleen fusion cells cannot grow in any culture medium. The $HGPRT^-$ myeloma and the myeloma–myeloma fusion cells cannot use hypoxanthine as a precursor for the biosynthesis of the purines guanine and adenine, which are, of course, essential for nucleic acid synthesis. However, they have a second, naturally occurring pathway for purine biosynthesis that utilizes the enzyme dihydrofolate reductase. Therefore, aminopterin is included in the medium because it inhibits dihydrofolate reductase activity. Hence, $HGPRT^-$ myeloma and myeloma–myeloma fusion cells are unable to synthesize purines in HAT medium, so they die (Fig. 9.3).

The spleen–myeloma fusion cells survive in HAT medium because the spleen cell contributes a functional HGPRT, which can utilize the exogenous hypoxanthine in the medium even though purine production by means of dihydrofolate reductase is blocked by aminopterin, and because the cell division functions of the myeloma cell are active. Thymidine is provided to overcome the block in pyrimidine production that is caused by the inhibition of dihydrofolate reductase by aminopterin. About 10 to 14 days after the fusion treatment, only spleen–myeloma fusion cells have survived and grown in the HAT medium. These cells are then distributed into the wells of plastic microtiter plates and grown on complete culture medium without HAT.

Identification of Specific Antibody-Producing Hybrid Cell Lines

The next task is to identify those hybrid cells that produce antibody against the immunizing antigen. One common screening procedure uses the culture medium, which contains secreted antibodies. The medium is collected from the wells that have growing cells and is added to a well of another microtiter plate that has been precoated with the target antigen. If the culture medium contains an antibody (primary antibody) that recognizes an epitope of the antigen, it will bind to the antigen and not be washed away during a subsequent washing step. A second antibody (secondary antibody) that is specific for mouse antibodies is added to the wells of the test plate. It will bind to any primary antibody that is bound to the antigen.

Before its use in the immunoassay, the secondary antibody is conjugated to an enzyme that can convert a colorless substrate to a colored compound. The presence of color in one of the test wells indicates that the original culture medium contained an antibody that was specific for the antigen (Fig. 9.4). If the culture medium does not contain an antibody that binds to the antigen, then the first wash will remove the primary antibody. Therefore, when the secondary antibody is added, it has nothing to bind to and is removed by the second washing step. In a well where such a sequence of events occurs, the substrate solution remains colorless.

Those wells of the original microtiter plate whose media give a positive (color) response in the immunoassay may contain a mixture of cell fusions. These cells are therefore diluted with culture medium and seeded into fresh wells to establish cell lines from single cells (clones). After the clones have been cultured, their media are tested again to determine which cell lines (hybrid spleen–myeloma cells, or hybridomas) produce monoclonal antibody molecules that recognize the target antigen. If more than one specific hybridoma is isolated, further tests are conducted to determine whether the different clones produce antibody against the same antigenic determinant. Each monoclonal antibody-producing clone can be maintained, more or less indefinitely, in culture. In addition, samples can be frozen in liquid nitrogen to provide a source of cells for future use.

Because a monoclonal antibody binds to a single discrete site, the specificity of an ELISA protocol can be considerably enhanced by using a monoclonal rather than a polyclonal antibody. Many monoclonal antibodies have been developed for use as immunodiagnostic agents for a variety of compounds and pathogenic organisms (Table 9.2). As an alternative to the

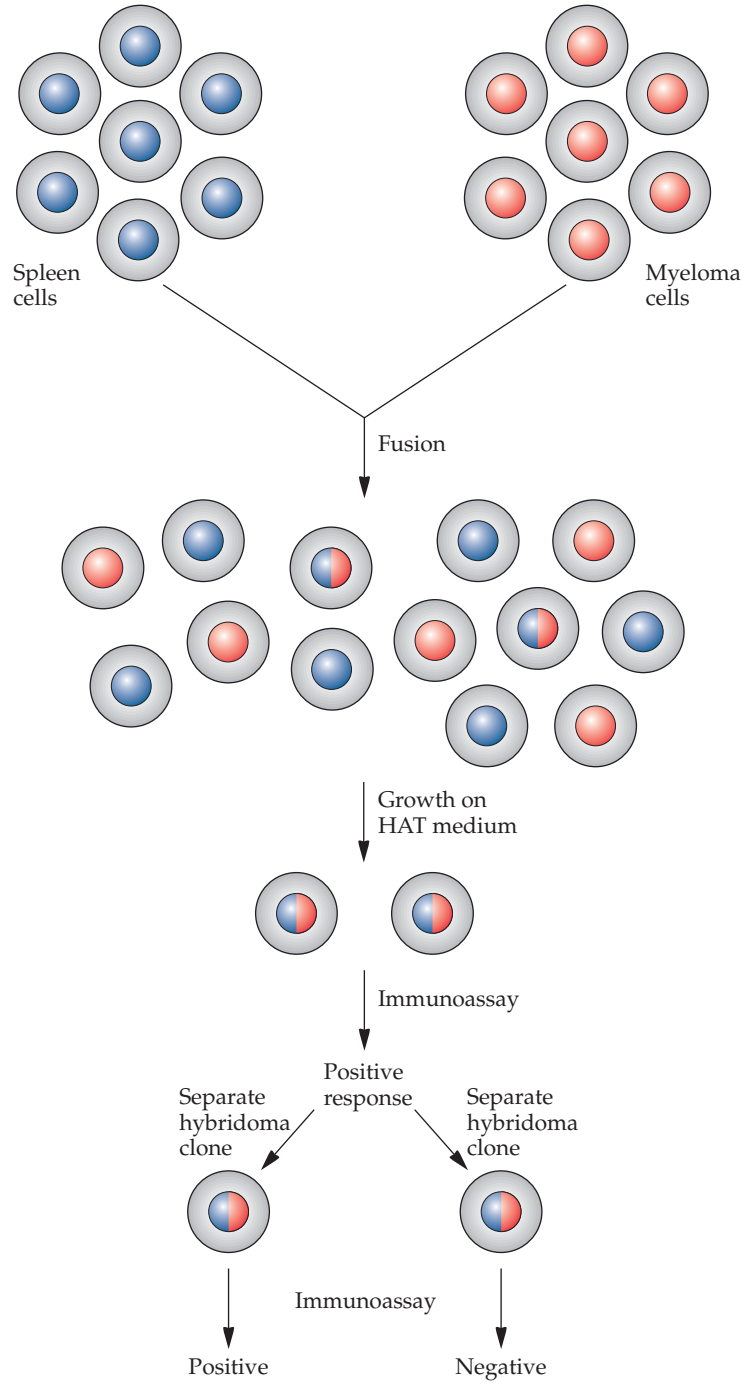


FIGURE 9.4 Screening for the production of a monoclonal antibody. Spleen cells from a mouse that was immunized with a specific antigen are isolated and fused in culture with myeloma cells that do not produce antibodies. Fused cells are selected for the ability to grow on HAT medium, which contains hypoxanthine, aminopterin, and thymidine. Cells that produce a specific antibody to the immunizing antigen (hybridomas) are identified by an immunoassay and individually subcultured. A hybridoma, which grows in culture and secretes a single type of antibody molecules, is the source of a monoclonal antibody.

isolation and synthesis of monoclonal antibodies in hybridoma cells in culture, monoclonal antibodies and parts of antibodies (Fab or Fv fragments) directed against a target antigen may be selected and produced in *Escherichia coli* (see chapter 10).

TABLE 9.2 Targets for diagnostic monoclonal antibodies

Polypeptide hormones
Chorionic gonadotropin
Growth hormone
Luteinizing hormone
Follicle-stimulating hormone
Thyroid-stimulating hormone
Prolactin
Tumor markers
Carcinoembryonic antigen
Prostate-specific antigen
Interleukin-2 receptor
Epidermal growth factor receptor
Cytokines
Interleukins 1–8
Colony-stimulating factor
Drug monitoring
Theophylline
Gentamicin
Cyclosporin
Miscellaneous targets
Thyroxine
Vitamin B_{12}
Ferritin
Fibrin degradation products
Tau protein
Infectious disease
Chlamydia
Herpes
Rubella
Hepatitis B
Legionella
Human immunodeficiency virus

Biofluorescent and Bioluminescent Systems

Proteins that naturally fluoresce or luminesce, or that can be easily induced to do so, may be used as biological reporters in a variety of ways. For example, genes encoding these bioreporter proteins may be used to engineer cells to produce a measurable signal in response to a particular chemical or physical agent in their environment. In one version of this system, a gene for a fluorescent or luminescent protein is placed under the control of a promoter that responds to certain environmental signals so that when the promoter is activated, a fluorescent or luminescent signal is produced (Fig. 9.5).

Colored Fluorescent Proteins

Green fluorescent protein. The 238-amino-acid-long photoprotein green fluorescent protein, isolated from the jellyfish *Aequorea victoria,* fluoresces green when it is exposed to ultraviolet light. While many fluorescent dyes are phototoxic, the incorporation of green fluorescent protein into cells allows intact living cells to be monitored in real time. The use of this reporter molecule has revolutionized fluorescence microscopy. Among its many uses, researchers have used green fluorescent protein to monitor tumor cells in gene therapy protocols, to assess the responses of specific cell types to various therapeutic drugs and treatments, to monitor protein-protein interactions, and to monitor the fates of individual proteins in different therapies.

Red fluorescent protein. Following the discovery and subsequent successful employment of green fluorescent protein, scientists began to search, both in nature and by directed mutagenesis, for other colored fluorescent proteins. Having multiple colored fluorescent proteins would enable several biological processes to be monitored at the same time. For the practical use of these proteins, it is essential that they be both as stable and as bright as possible. One problem with many of the naturally occurring colored fluorescent proteins is that they often have a tendency to form homodimers or homotetramers. Such multimeric structures can adversely influence the subcellular localization of the proteins, potentially leading to intracellular aggregation and other artifacts. One group of researchers isolated a gene for a red fluorescent protein from *Discosoma* coral and, by means of multiple random mutations, generated a mutant that existed exclusively as a monomer instead of as a tetramer. With each iterative cycle of random mutagenesis, proteins that yielded red fluorescence that was both bright and stable were selected. The production of monomeric red fluorescent protein required 33 mutations. This success notwithstanding, monomeric red fluorescent protein had several drawbacks compared to the native tetrameric form of the protein, including decreased brightness and reduced photostability.

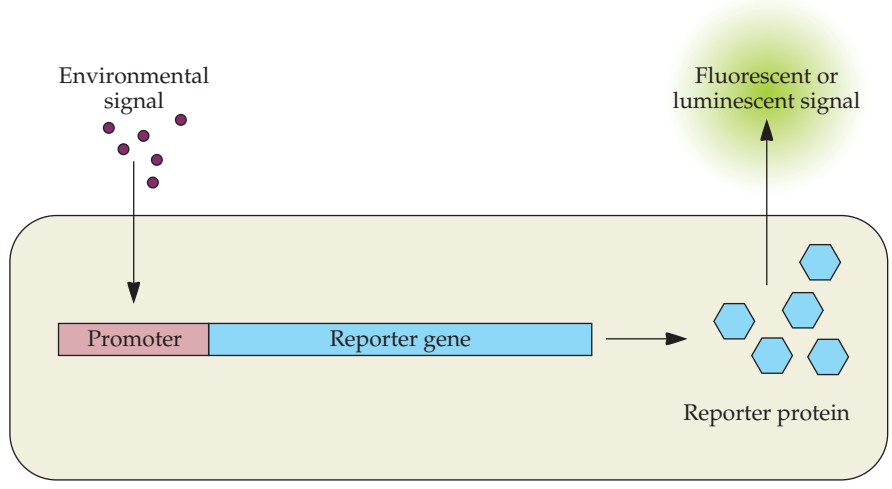


FIGURE 9.5 Schematic representation of a bacterial cell, engineered to respond to a particular environmental compound, producing a detectable fluorescent or bioluminescent signal.

To address some of the above-mentioned problems with monomeric red fluorescent protein, as well as to expand the repertoire of available fluorescent proteins, additional rounds of mutagenesis were performed. First, DNA encoding 7 amino acids from the N terminus of green fluorescent protein was added to the gene for red fluorescent protein (Fig. 9.6). Then, DNA for the 6 amino acids from the green fluorescent protein C terminus was added to the gene for the red fluorescent protein. This construct then became the starting point for several additional rounds of random mutagenesis and directed evolution. Eventually, seven different monomeric colored fluorescent proteins were produced (Fig. 9.7). It is argued that there is no one best colored fluorescent protein. Some are brighter than others, some are more photostable, and some are more sensitive to changes in pH. Thus, various fluorescent proteins may be used for different applications. Moreover, using two or three of these proteins, it is possible to label several different cellular components or cell types at once, thereby increasing the utility of this approach.

Luciferase

The luciferase enzyme, which catalyzes a light-emitting reaction, may be produced by a variety of different organisms, including bacteria, algae, fungi, jellyfish, insects, shrimp, and squid. Luciferase genes from bacteria are termed *lux* genes, while those from other organisms—the most widely studied and utilized being the firefly—are termed *luc* genes. The *lux* system includes five genes, *luxCDABE,* and produces a peak of light at 490 nm. In some applications, all five *lux* genes are utilized as a means of monitoring the presence and concentrations of various compounds in the environment, such as organic compounds, including phenol, salicylate, benzene, trichloroethylene, ammonia, xylene, toluene, and ethylbenezene, or metals, including cobalt, copper, iron, lead, mercury, nickel, and zinc. When all five genes are used, the light-generating system does not require the addition of any other compounds. Therefore, following the addition of a contaminant-containing sample to bacterial cells carrying *luxCDABE,* a quantifiable

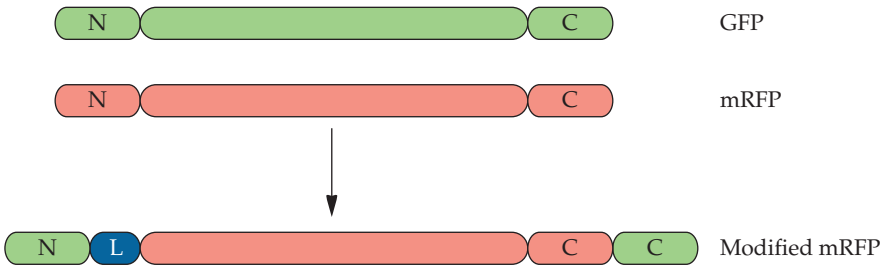
amount of bioluminescence is produced within a period of a few minutes to no more than a few hours. In some cases, the reporter system includes only the *luxAB* genes. While *luxA* and *luxB* are together responsible for generating the light signal, this system requires that the substrate decanal be added during the assay procedure.

The firefly luciferase-catalyzed reaction results in the production of light at 550 to 575 nm. Moreover, the system requires the addition of the low-molecular-weight organic compound luciferin as a substrate for the light reaction. Typically, *luc* genes are used in conjunction with eukaryotic cells.

Microbial Biosensors

There is a need for methods that can easily and rapidly detect the large numbers of potentially toxic compounds that contaminate the environment. Once the contaminated sites have been identified and their range has been delineated, there are a number of highly sophisticated analytical techniques available to identify and quantify specific pollutants. Bacteria that are constitutively bioluminescent (i.e., unlike the situation mentioned above, the bioluminescence does not need to be induced) are good candidates for pollutant detectors. In the presence of pollutants, the bioluminescence decreases, providing a clear indication of the presence of the pollutants. Naturally bioluminescent bacteria, such as the marine bacterium *Vibrio fischeri,* require saline conditions and a particular pH range and are therefore not useful for testing terrestrial groundwater. However, structural genes encoding enzymes that lead to bioluminescence *(luxCDABE)* may be inserted into random sites in the chromosomal DNA of a soil bacterium, such as *Pseudomonas fluorescens*. These genes do not contain a transcriptional promoter, so after insertion into the chromosomal DNA of *P. fluorescens,* the only luminescent colonies (visualized in a darkroom) are those in which the *lux* genes are inserted downstream from a constitutive *P. fluorescens* promoter (without disrupting any important bacterial genes). The cells that luminesce to the greatest extent and have a growth rate similar to that of the wild-type strain are selected for testing with various environmental pollutants. To screen water samples for the presence of various pollutants (both metals and organic compounds), a suspension of bioluminescent *P. fluorescens* is mixed with the solution being tested, and after a 15-minute incubation together, the luminescence of the suspension

FIGURE 9.6 Construction of a modified monomeric red fluorescent protein (modified mRFP). The regions of the gene encoding the N and C termini of the green fluorescent protein (GFP) were spliced onto the gene for mRFP following the removal of the portion of the mRFP gene encoding the N terminus and the addition of an oligonucleotide encoding a short linker peptide (L).



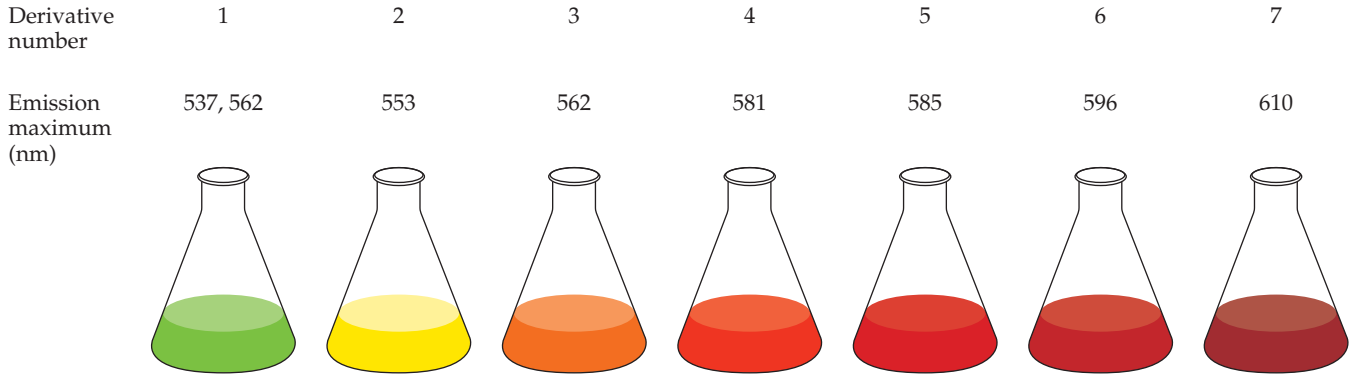


FIGURE 9.7 Various colored monomeric fluorescent proteins derived from monomeric red fluorescent protein showing their emission wavelengths and color maxima following excitation. The colors have been called, from derivative 1 through 7, honeydew, banana, orange, tomato, tangerine, strawberry, and cherry, respectively. Adapted from Shaner et al., *Nat. Biotechnol.* **22:**1567–1572, 2004.

is measured in a luminometer (Fig. 9.8). When a test sample contains a low to moderate level of certain pollutants, the cell luminescence is inhibited, presumably because the pollutant directly interferes with bacterial metabolism. Since this procedure is rapid, simple, and inexpensive, it is a good first screen for assessing the presence of pollution at a particular site. After a positive response with a bacterial biosensor, the actual pollutants can be determined by other methods.

In the United States, it has been estimated that there are approximately 87,000 different chemical compounds that need to be tested for estrogenic activity, i.e., steroid-like activity that can disrupt the endocrine system in vertebrates. While a number of different methods exist that could be used to test these compounds, they are too slow for this sort of large-scale screening. Therefore, scientists have developed a simple and sensitive system for the rapid initial screening of these 87,000 compounds. With this system, yeast *(Saccharomyces cerevisiae)* cells have been genetically engineered to produce measurable quantities of light in the presence of extremely low levels of estrogenic compounds (Fig. 9.9). Using this method, in which luminescence is induced, to test for estrogenic compounds, light production could be detected in as little as 1 hour. Moreover, following 6 hours of incubation, the assay attained maximum bioluminescence when the engineered yeast cells were exposed to as little as 5×10^{-11} M 17β-estradiol, a common estrogen. Of course, many estrogenic compounds required higher concentrations in order to be detected. The main drawback of this approach is that the yeast cell wall and transport system facilitate the entry of some compounds into the cell and inhibit the uptake of other compounds. This can skew the results and in some instances may suggest that a compound is not estrogenic when it is unable to efficiently enter yeast cells. Nevertheless, this technique is likely to identify a large number of estrogenic compounds that were deemed nonestrogenic until they were tested with this protocol.

At the same time that some groups of scientists are working to develop and perfect cells as biosensors, others have focused their efforts on automating these systems. Such an automated system might include genetically engineered cells that emit blue-green light (~490 nm) in response to specific

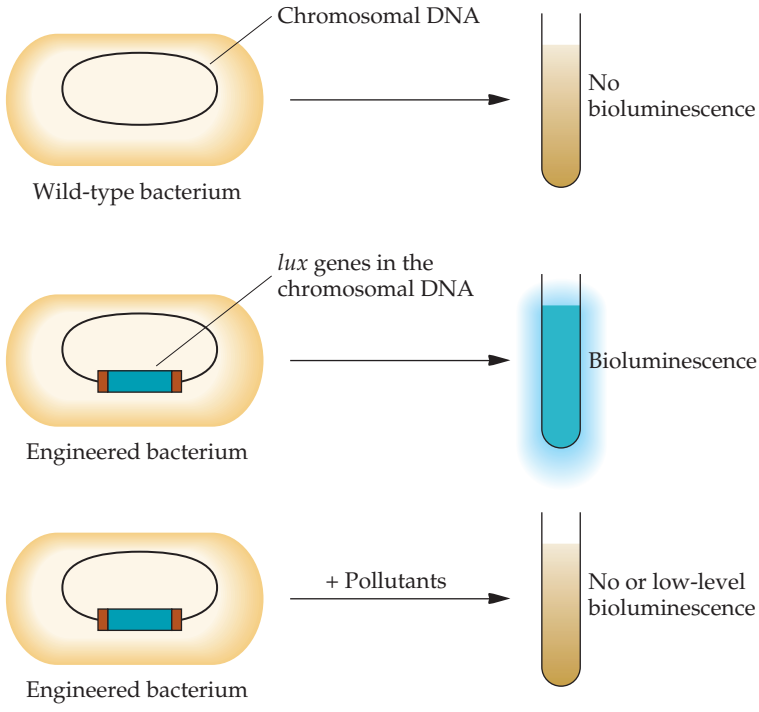


FIGURE 9.8 Assaying for the presence of pollutants with genetically engineered bioluminescent *P. fluorescens*.

compounds, an environment that sustains the cells, a light-tight enclosure, and an integrated circuit for light detection and signal processing. While simple prototypes have been constructed, the difficulties associated with utilizing living cells remain major impediments.

Nucleic Acid Diagnostic Systems

The genetic material of an organism contains the essential information that contributes to its various features and characteristics. For example, bacterial pathogenicity may be due to the presence of a specific gene or set of genes. Similarly, alteration of a gene may cause an inherited genetic disease in humans. In theory, the sequence of nucleotides that contributes to a particular biological characteristic is a distinctive signature that, if detectable, can be used as a definitive diagnostic determinant.

Nucleic acid hybridization is the basis for rapid and reliable assays. The physical basis of these systems is precise nucleotide base pairing and hydrogen bonding between one string of nucleotides and a complementary nucleotide sequence. A general laboratory nucleic acid hybridization scheme is as follows.

1. Bind single-stranded DNA (the target) to a membrane support.
2. Add single-stranded labeled DNA (the probe) under appropriate conditions of temperature and ionic strength to promote base pairing between the probe and the target DNAs.
3. Wash the support to remove excess unbound labeled probe DNA.
4. Detect the hybrid sequences that form between the probe and target DNA.

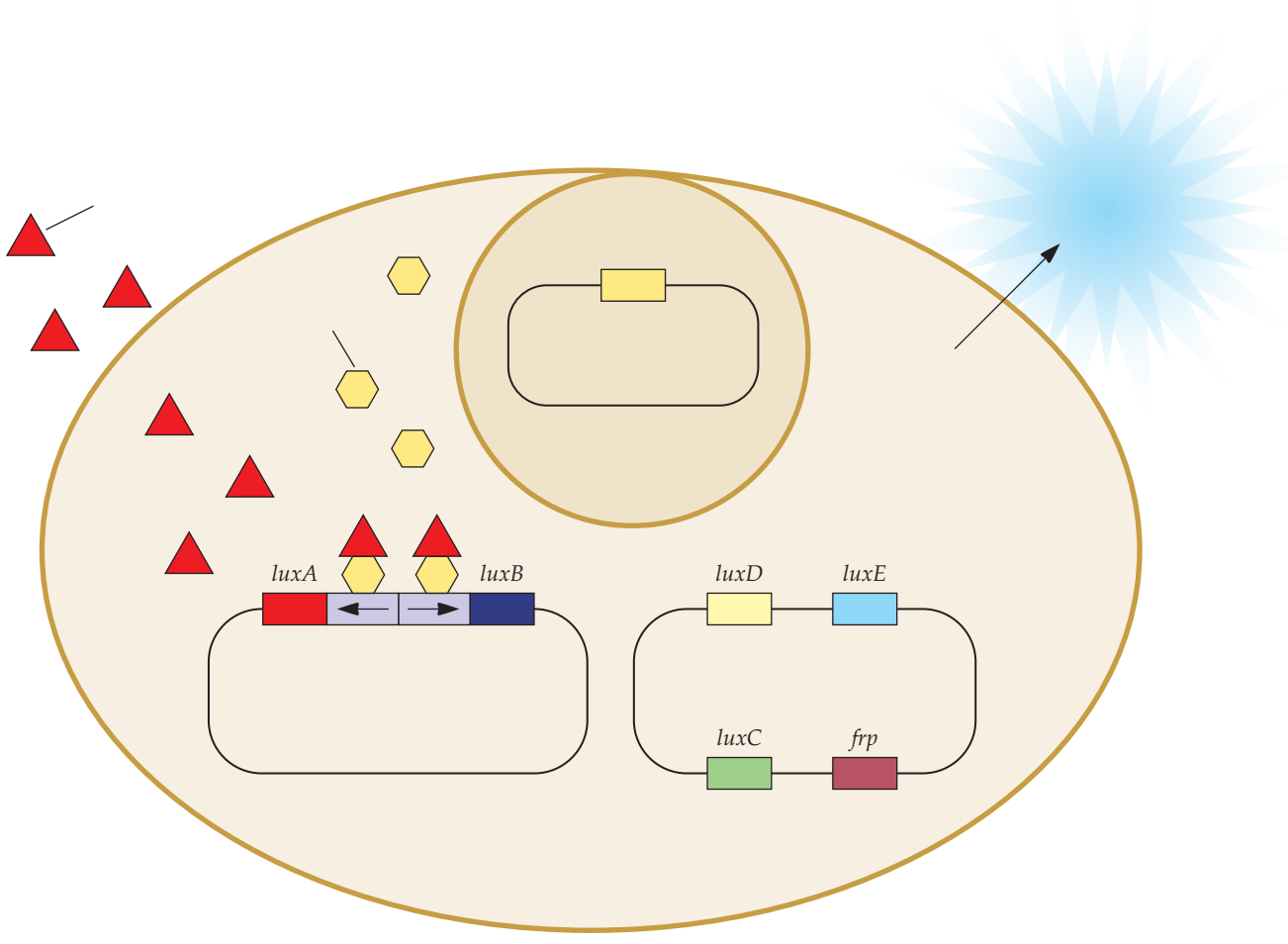


FIGURE 9.9 Schematic representation of a strain of the yeast *Saccharomyces cerevisiae* that has been engineered to detect low levels of estrogenic substances in the environment. Each yeast cell contains the human estrogen receptor (hER) gene integrated into its chromosomal DNA, one plasmid (pUTK404) that contains constitutively expressed bacterial *luxCDE* genes and the flavin oxidoreductase *(frp)* gene from the bacterium *Vibrio harveyi,* and one plasmid (pUTK407) that contains *luxA* and *luxB* genes under the regulatory control of estrogen response elements (ERE) and constitutive bacterial promoters (not shown). Following the binding of the hER protein to an estrogenic compound, the complex binds to an ERE and activates *luxAB* transcription. Expression of the *luxAB, luxCDE,* and *frp* genes leads to the production of a measurable bioluminescent signal. Adapted from Sanseverino et al., *Appl. Environ. Microbiol.* **71:**4455–4460, 2005.

A nucleic acid hybridization diagnostic test has three critical elements: probe DNA, target DNA, and signal detection. This type of detection system can be both extremely specific and highly sensitive.

Hybridization Probes

To be effective, a nucleic acid hybridization probe must have a high degree of specificity. In other words, the probe must hybridize exclusively to the selected target nucleic acid sequence. False positives (i.e., responses in the

absence of the target sequence) and false negatives (i.e., no response when the target is present) severely undermine the utility of a diagnostic procedure. Probes can be specific at different organismic levels. They can distinguish between two or more species, determine particular strains within a given species, or identify differences between genes. Depending on the requirements of the test protocol, probes can be DNA or RNA, long (>100 nucleotides) or short (<50 nucleotides), and chemically synthesized, cloned intact genes, or isolated regions of a gene.

Sequences that might make effective probes can be isolated in a number of ways. For example, the DNA from a pathogenic organism can be cut with a restriction endonuclease and cloned into a plasmid vector. Recombinant plasmids are screened with the genomic DNA from both pathogenic and nonpathogenic strains. Those plasmids that contain sequences that hybridize only to the pathogenic strain form the basis for species-specific, and even strain-specific, probes (Fig. 9.10). Additional hybridization tests with DNA from a wide range of organisms are then conducted to ensure that the candidate probe sequences do not cross-hybridize. Each potential probe is also tested under simulated sample conditions, including the presence of mixed cultures, to determine its level of sensitivity. It is important to note that knowledge of the genomic sequence of a large number of bacterial pathogens (currently several hundred) has facilitated the identification of unique stretches of DNA that could be used as probes.

The ability to perform nucleic acid probe diagnostic assays directly on available samples without either additional culturing or time-consuming extraction procedures is extremely desirable, especially with clinical specimens. Researchers have successfully used probes that hybridize to target DNA from fecal samples, urine, blood, throat washes, and tissue samples without extensive DNA purification. If a target sequence is rare in the working sample, the polymerase chain reaction (PCR) can be used to amplify it.

Diagnosis of Malaria

An example of a diagnostic protocol that utilizes a DNA probe as a means of detection is the procedure developed for the detection of *Plasmodium falciparum*. This parasite causes malaria, a disease that affects about one-third of the world's population. The parasite infects and destroys red blood cells, leading to fever and, in severe cases, damage to the brain, kidneys, and other organs. Sensitive, simple, and inexpensive methods are required to identify the source(s) of the parasite in various localities, to assess the progress of eradication programs, and to facilitate early treatment. Currently, malarial infections are diagnosed by either microscopic examination of blood smears or immunological detection of parasite antigens, effective but labor-intensive and time-consuming processes, especially given the large numbers of samples that need to be examined. Although immunological procedures for *Plasmodium* detection, such as ELISAs, are rapid and amenable to automation, they do not always discriminate between current and past infections, because they are designed simply to detect anti-*Plasmodium* antibodies in the blood of affected individuals.

A DNA diagnostic procedure that selectively measures only current infections, i.e., the presence of DNA-containing organisms, was developed

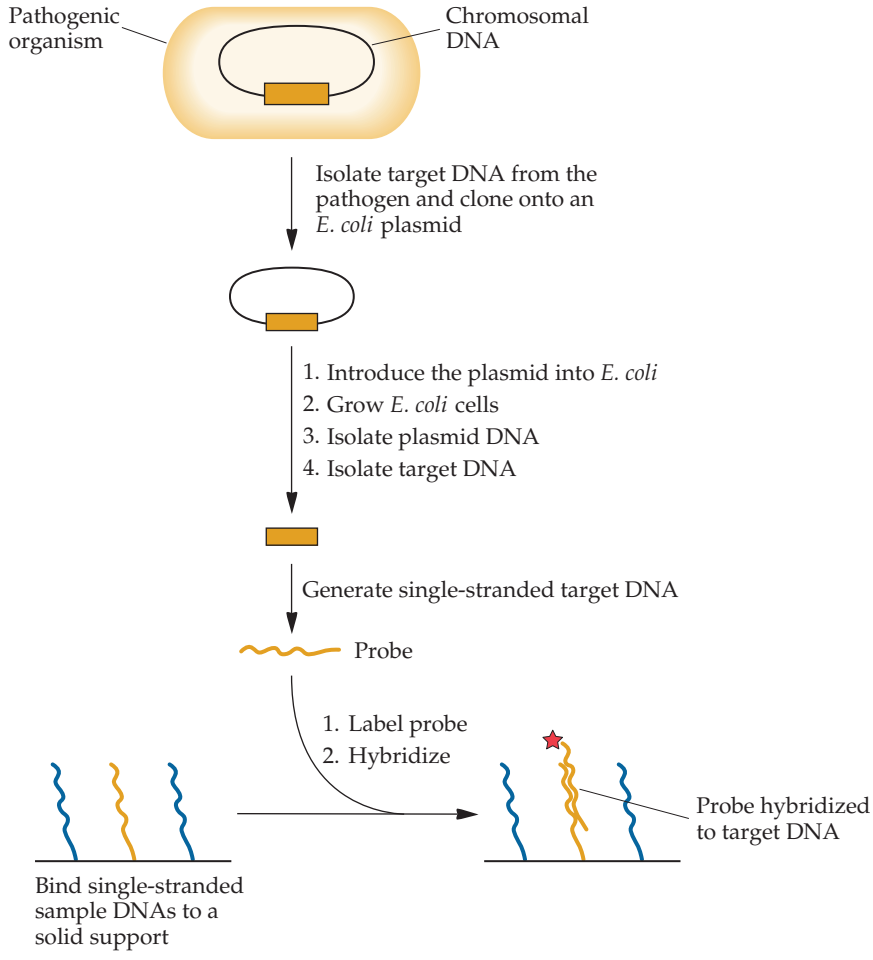


FIGURE 9.10 Overview of the development and use of a DNA hybridization probe.

by using highly repeated DNA (DNA that is present in many copies) from *P. falciparum*. First, a genomic library of the parasite DNA was screened with labeled whole-genome parasite DNA. The most intensely labeled hybridizing colonies were selected because they were expected to contain repetitive DNA. The DNA from each of the selected colonies was then tested for its ability to hybridize with DNA from several other *Plasmodium* species that do not cause malaria. The DNA sequence that was chosen as a specific probe hybridized with *P. falciparum* but not with *Plasmodium vivax, Plasmodium cynomolgi,* or human DNA, despite the fact that *P. vivax* causes a less severe form of malaria. This probe can detect as little as 10 picograms of purified *P. falciparum* DNA or 1 nanogram (ng) of *P. falciparum* DNA in blood.

More than 100 different DNA diagnostic probes have been isolated and characterized for the detection of various pathogenic strains of bacteria, viruses, and parasites. For example, probes have been developed for the diagnosis of human bacterial infections caused by *Legionella pneumophila* (respiratory failure), *Salmonella enterica* serovar Typhi (food poisoning), *Campylobacter hyointestinalis* (gastritis), and enterotoxigenic *E. coli* (gastroenteritis). Clearly, this is just the "tip of the iceberg," because in principle, nearly all pathogenic organisms can be detected by this procedure.

Detection of *T. cruzi*

The protozoan parasite *Trypanosoma cruzi* is the causative agent of Chagas disease. In this disease, the parasites invade the liver, spleen, lymph nodes, and central nervous system, where they multiply and destroy the infected cells. *T. cruzi* is quite common in Latin America. It is spread by insects and is responsible for approximately 50,000 deaths per year. Diagnosis of acute Chagas disease is usually made by microscopic examination of a fresh blood sample. Alternatively, a test that takes a longer time but ensures that the parasite has not been overlooked entails feeding a patient's blood to uninfected insects and then examining with a microscope the contents of the insects' intestines for parasites about 30 to 40 days later. Both of these tests are laborious, time-consuming, and costly. Chagas disease can also be diagnosed by immunological tests; however, these tests are notorious for producing false-positive responses. As a possible alternative to these less than satisfactory procedures, several PCR-based assays have been developed. At present, PCR assays for Chagas disease are used as adjuncts to the traditional diagnostic procedures that are currently in widespread use.

In one of the PCR-based assay procedures, a 188-base-pair (bp) DNA fragment that is present in multiple copies in the *T. cruzi* genome but is absent from the genomic DNA of several related parasites is the target sequence. The presence of the amplified 188-bp DNA fragment is readily detected by polyacrylamide gel electrophoresis. In general, with minor variations in the methodology, such as the primer sequences, PCR can facilitate the detection of a wide range of bacteria, viruses, and parasites. Currently, there are several PCR-based diagnostic kits that have been

MILESTONE

Detection of Sickle Cell β^S-Globin Allele by Hybridization with Synthetic Oligonucleotides

B. J. Conner, A. A. Reyes, C. Morin, K. Itakura, R. L. Teplitz, and R. B. Wallace

Proc. Natl. Acad. Sci. USA **80:**278–282, 1983

The development of techniques in the early 1980s for rapid, efficient, and inexpensive chemical synthesis of DNA oligonucleotides opened up the possibility of using radiolabeled oligonucleotides as hybridization probes for the detection of a variety of DNAs, including mutations in human genes. At the time, relatively few human genes or even cDNAs had been isolated. Thus, it was not always easy to find a homologous hybridization probe for a particular human gene. Moreover, even if the relevant cDNAs had been isolated, it was not possible to discriminate by DNA hybridization between wild-type and mutant human genes that had only a single-base-pair difference when the hybridization probe was large (≥100 bp). However, Conner et al. synthesized specific oligonucleotides that could recognize either the wild-type or the mutant DNA. Moreover, oligonucleotides complementary to the sequences of both the wild type and the mutant were synthesized so that it could be determined whether a person was heterozygous, as well as whether a person was homozygous. For sickle-cell anemia, two 19-base-long oligonucleotides were used, one complementary to the normal β-globin gene (β^A) and the other complementary to the sickle-cell gene (β^S). DNA from normal individuals ($\beta^A\beta^A$) hybridized only with the β^A probe, DNA from individuals with sickle-cell anemia ($\beta^S\beta^S$) hybridized only with the β^S probe, and DNA from heterozygous individuals ($\beta^A\beta^S$) hybridized with both probes. This model system was the first demonstration of the feasibility of determining genotypes by DNA hybridization and opened up the possibility of detecting a range of human genetic disorders by hybridization with oligonucleotide probes. This possibility has been realized with the development of a large number of DNA-based gene mutation tests. Although the original strategy has been largely supplanted by newer techniques, such as PCR and OLA, this work was important in establishing that single DNA base pair mutations could be accurately and easily detected.

approved for use by the U.S. Food and Drug Administration for the detection and quantitation of human immunodeficiency virus, *Mycobacterium tuberculosis* (the causative agent of tuberculosis), and *C. trachomatis*.

Nonradioactive Hybridization Procedures

In many research laboratories, nucleic acid hybridization is routinely detected by labeling the probe with a radioactive isotope, commonly phosphorus-32. High specific activity ensures an excellent signal-to-noise ratio. In a standard detection system, a radiolabeled probe is mixed with target DNA that is bound to a membrane support. After the support is washed free of nonhybridized probe DNA, the presence of radioactivity is determined by laying the membrane on X-ray film (autoradiography).

However, phosphorus-32 is short-lived, is potentially dangerous, and requires special laboratory equipment for handling and safe disposal, so nonradioactive systems for indicating hybrid DNA formation have also been developed. The nonradioactive detection systems achieve signal amplification by enzymatic conversion of either chromogenic or chemiluminescent substrates. Chromogenic substrates change color and chemiluminescent substrates give off light when they are converted into a specific product by an appropriate enzyme. The signal is detected, in most of these systems, by incorporating biotin-labeled nucleotides into the DNA probe and following a more or less standard procedure:

1. The biotin-labeled probe is hybridized to the target DNA (Fig. 9.11A).
2. Either avidin, a chicken egg white protein, or streptavidin, a bacterial analogue of avidin, is added (Fig. 9.11B).
3. A biotin-labeled enzyme, such as alkaline phosphatase or peroxidase, is added (Fig. 9.11C).
4. Depending on which biotin-labeled enzyme was used in the previous step, either a chromogenic or a chemiluminescent substrate is added, and either the color change or the light produced as a consequence of the conversion of substrate into product is measured (Fig. 9.11D).

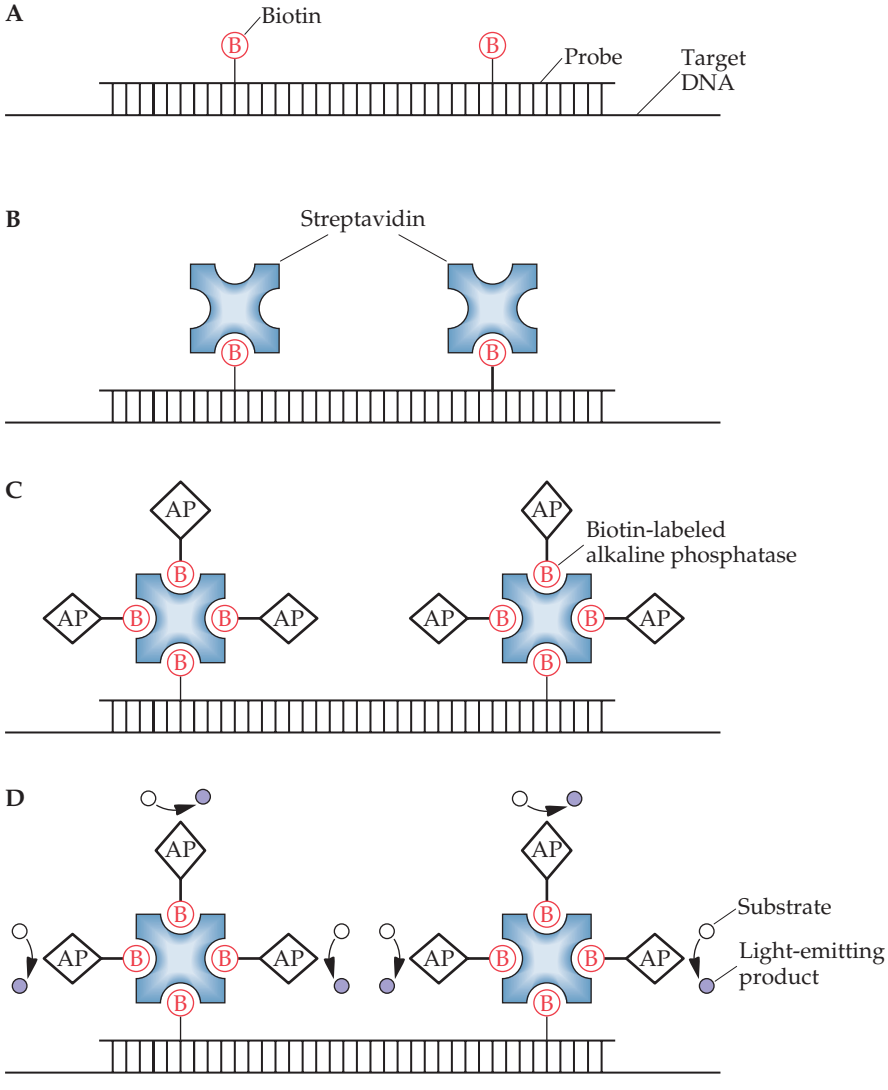
Alternatively, following hybridization with a biotin-labeled probe in step 2 above, a streptavidin–enzyme complex with an available biotin-binding site can be added.

Both avidin and streptavidin bind very tightly (K_d [dissociation constant] = ~10^{-15}) to biotin; in addition, each of these proteins has four separate biotin-binding sites, so a single molecule of avidin or streptavidin can bind both a biotin-labeled enzyme and a biotin-labeled probe. Enzymatic activity is not impaired by biotin labeling or binding to streptavidin. In chromogenic detection systems, the action of the enzyme on the substrate creates a colored insoluble dye that remains at the site of the hybrid DNA. In chemiluminescent systems, enzymatic alteration of the substrate generates a product that emits light at the site of the hybrid DNA.

Nonradioactive systems have other advantages: biotin-labeled DNA is stable for at least 1 year at room temperature, devices that detect chemiluminescence are as sensitive as those that detect radioactive signals, and detection of the emitted light with either X-ray film or a luminometer, or scoring of a color change, can be completed within a few hours. The use of

chemiluminescence, which is more sensitive than chromogenic dyes, is becoming the detection signal system of choice for many nucleic acid probe-based diagnostic assays. For PCR-based assays, the amplification product can be labeled by a fluorescent dye that is bound to the 5′ end of each primer. A fluorescent compound emits light of a longer wavelength after it absorbs light of a shorter wavelength. Fluorescein, which appears green under certain wavelengths of light, and rhodamine, which appears red, are often used for this purpose. After PCR amplification of a target DNA with fluorescence-labeled primers, the primers are separated from the amplification product and the presence of the label is detected (Fig. 9.12). If the target DNA is not present in the sample, then no fluorescent product will be

FIGURE 9.11 Chemiluminescent detection of target DNA. **(A)** A biotin-labeled probe is bound to the target DNA. **(B)** Streptavidin is bound to the biotin molecules. **(C)** Biotin-labeled alkaline phosphatase binds to the streptavidin. **(D)** Alkaline phosphatase converts the substrate into a light-emitting product. B, biotin; AP, alkaline phosphatase.

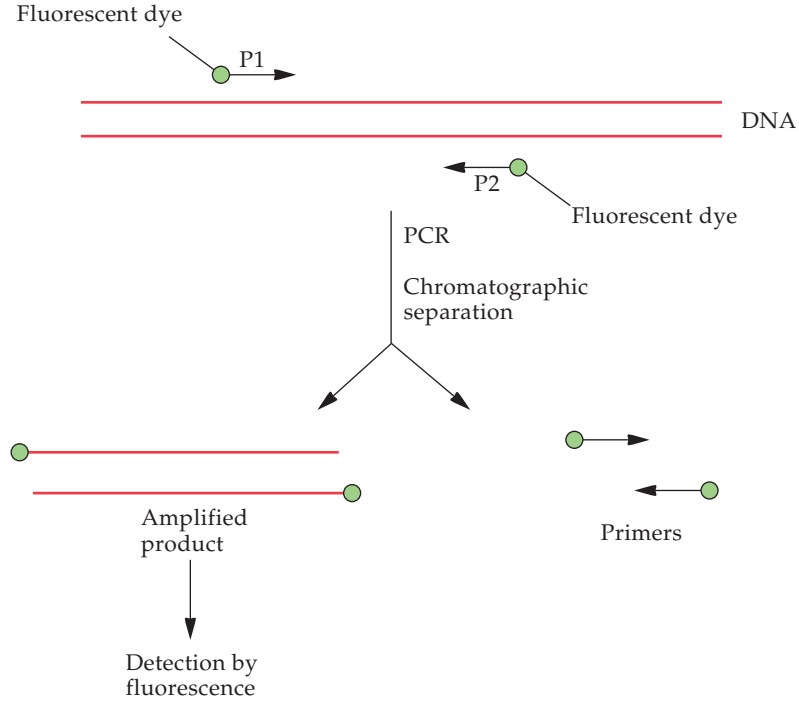


observed. This system is not only sensitive, it is also quite rapid, since it is not necessary to run a gel to separate the amplified target DNA.

Molecular Beacons

A novel nonradioactive method for detecting specific sequences of nucleic acids involves using "molecular beacon" probes (Fig. 9.13). A typical molecular beacon probe is 25 nucleotides long. The 15 nucleotides in the middle are complementary to the target DNA and are designed so that this single-stranded molecule does not does not form a structure in which these nucleotides base pair with one another. However, the 5 nucleotides at each end are complementary to each other and not to the target DNA. A fluorescent molecule (fluorophore) is attached to the 5′ end, and a nonfluorescent molecule (quencher) that can absorb the energy emitted by the fluorophore before it fluoresces is attached to the 3′ end. In solution at room temperature, the conformation of the molecular beacon ensures that the fluorophore and quencher are close to one another, and the fluorophore is quenched (does not fluoresce). On the other hand, when the 15 middle nucleotides of the molecular beacon probe hybridize to a target DNA or RNA sequence, the fluorophore and quencher are separated from each other and the fluorophore is not quenched, i.e., it fluoresces. With this procedure, care must be taken to maintain the reaction mixture at near-ambient temperatures, since high temperatures can also cause the nucleotides in the intrastrand (hairpin) stem of the molecular beacon to become unpaired, with the result that the molecule fluoresces. For this procedure to be effective, all 15 nucleotides in the molecular beacon probe must be perfectly complementary to the target

FIGURE 9.12 Use of fluorescent dyes that are attached to primers for detecting amplified PCR products. The primers are marked P1 and P2.



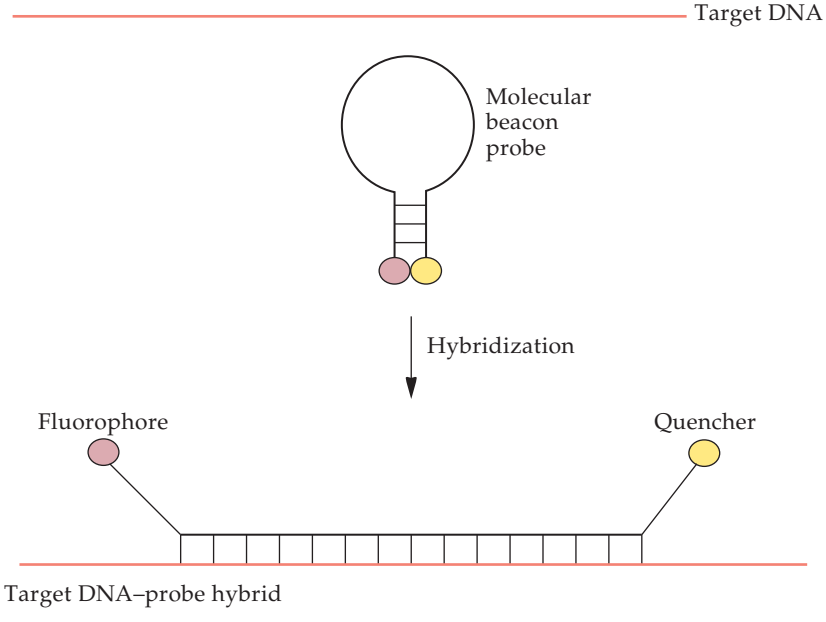
DNA or RNA. The sensitivity of this procedure can be improved dramatically if the target DNA is first amplified by PCR.

A number of variations of the basic molecular beacon protocol have been developed. For example, combinations of molecular beacon probes may be used simultaneously provided that each one is complementary to a different target DNA and contains a different-color light-emitting fluorophore (Fig. 9.14). For example, when one wants to determine the genotype of an individual, two different molecular beacon probes are added to a biological sample, such as blood, that contains DNA from that individual. The first molecular beacon probe is labeled with one of the fluorophores (e.g., fluorescein), and all 15 probe nucleotides are exactly complementary to the wild-type sequence (Fig. 9.15). One nucleotide difference is sufficient to prevent hybridization. The second molecular beacon probe is labeled with a different fluorophore (e.g., Texas red), and the 15 probe nucleotides are complementary to the sequence from the mutant form (Fig. 9.15). Following hybridization, the appearance of green fluorescence indicates a homozygous normal genotype, red fluorescence indicates a homozygous mutant genotype, and green and red fluorescence indicates a heterozygous genotype.

DNA Fingerprinting

The DNA from a biological sample left at the scene of a crime can be analyzed and compared with the DNAs of likely suspects. A match between evidence and a particular individual is helpful to the prosecution. In addition, DNA comparisons are used to determine whether individuals have

FIGURE 9.13 Hybridization of a molecular beacon probe to target DNA. If the target sequence is present, the unpaired single-stranded portion of the molecular beacon base pairs with its complementary sequence, thereby separating the fluorophore and quencher moieties. As a result, the fluorescence of the fluorophore is not quenched. Adapted from Tyagi and Kramer, *Nat. Biotechnol.* **14:**303–308, 1996.



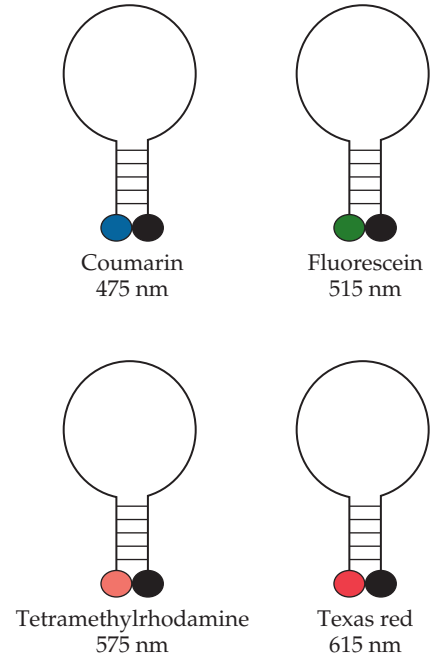
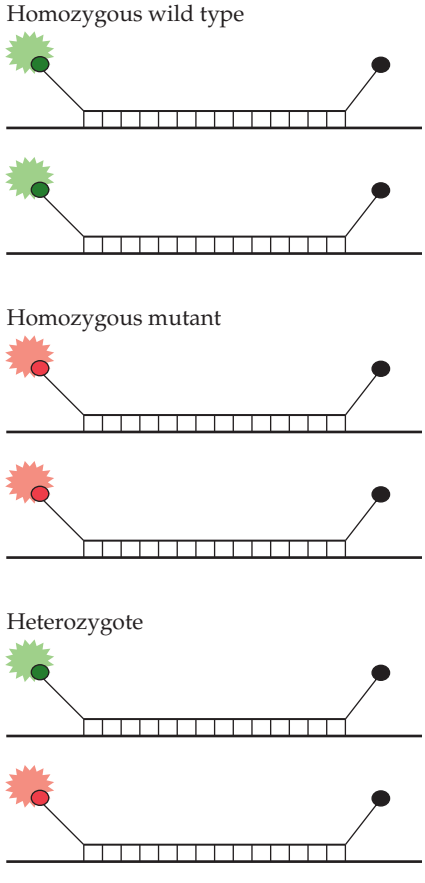


FIGURE 9.14 Molecular-beacon probes with different fluorophores. Beneath each molecular beacon probe is the name of the fluorophore and the wavelength of its maximum emission. Each fluorophore is represented (on the left) by the color of its fluorescence.

FIGURE 9.15 Use of two different molecular beacon probes to distinguish between different genotypes. Once the probe is bound to target DNA, the fluorescence of the fluorophore is no longer quenched.



been wrongly convicted of a crime. In other instances, DNA analyses help determine paternity and identify victims of disasters. Distinguishing individuals with DNA analysis is called DNA fingerprinting (DNA typing). One approach for determining DNA relationships among humans relies on DNA hybridization to undegraded minisatellite DNA. The probes for this type of analysis consists of human minisatellite DNAs, sequences that occur throughout the human genome and consist of tandemly repeated sequences (Fig. 9.16). The lengths of the repeats range from 9 to 40 bp, and the numbers of repeats in the minisatellites range from about 10 to 30. A minisatellite DNA sequence at a specific chromosome location can have different lengths in different individuals. This variability is due to either a gain or a loss of tandem repeats, probably during DNA replication. These changes do not have any biological effect because minisatellite DNA does not encode any proteins. Unrelated individuals generally have minisatellites that differ in length, but children inherit one set of minisatellite DNA sequences from each parent. For minisatellite DNA typing, the sample DNA is digested with a restriction enzyme, and the fragments are separated on an agarose gel and transferred by blotting them onto a nylon membrane. The membrane is hybridized sequentially with four or five separate labeled minisatellite DNA probes, each of which recognizes a distinct DNA sequence. After each hybridization reaction, the bands in which the probe has bound to the digested DNA sample are visualized by autoradiography, and the banding pattern for each sample is noted (Fig. 9.17). Before the next probe is used, the first probe is completely removed

(stripped) from the membrane. Since each hybridization and autoradiography step can take up to 10 to 14 days, the entire process may take many weeks, and even several months.

A minisatellite DNA pattern (fingerprint) represents the repertoire of the lengths of some of these sequences in an individual. Because of the extensive variability in human minisatellite DNA sequences, the chance of finding two individuals in the population with the same DNA fingerprint is about 1 in 10^5 to 1 in 10^8. Therefore, individuals' DNA banding patterns based on minisatellite DNA sequences are almost as unique as their fingerprints.

RAPD

Not only are DNA banding patterns important for forensic analyses, they are also useful in distinguishing among different plant cultivars. Random amplified polymorphic DNA (RAPD) markers may be used for this purpose. With this procedure, an arbitrary oligonucleotide primer, usually 9 to 10 bp long, that does not contain any palindromic sequences and has a G+C content of 50 to 80% is added to a sample of plant chromosomal DNA; virtually any oligonucleotide sequence will suffice. Because of its short sequence, the added oligonucleotide will pair with the chromosomal DNA at many sites, sometimes including opposite strands on the target DNA. When the 3' ends of the oligonucleotides on opposite strands of the DNA face each other, the DNA in between can be amplified (Fig. 9.18). Although the sequence of each primer is known, it is not known which oligonucleotide, if any, will be effective in priming the PCR. Whenever a primer can hybridize to both strands of the target DNA in the proper orientation and the two sites are about 100 to 3,000 bp from each other, the intervening DNA region will be amplified via PCR. The DNA fragments of characteristic size that are produced can be visualized following polyacrylamide gel electrophoresis. The number of amplified DNA fragments in a sample is dependent on the primer and the genomic DNA used. Each time that the same primer is used with the same target DNA, the amplified products will be the same. A single nucleotide substitution in a primer will result in a complete change in the RAPD pattern. Thus, the RAPD fingerprints of different plant cultivars can be compared when the same set of oligonucleotide primers is used. To fingerprint the DNAs of two very similar plant strains or cultivars, it is often necessary to perform the RAPD procedure with a number of different arbitrary primers with known

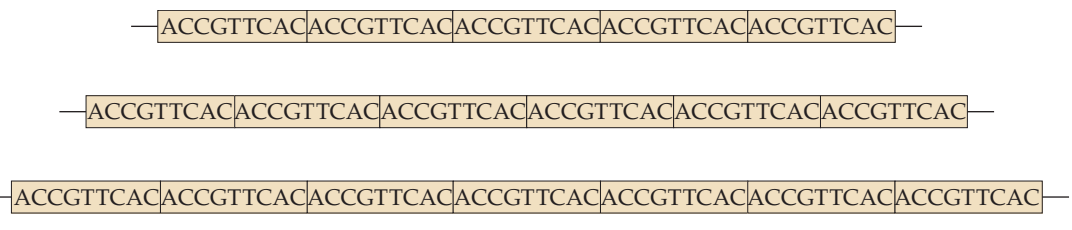
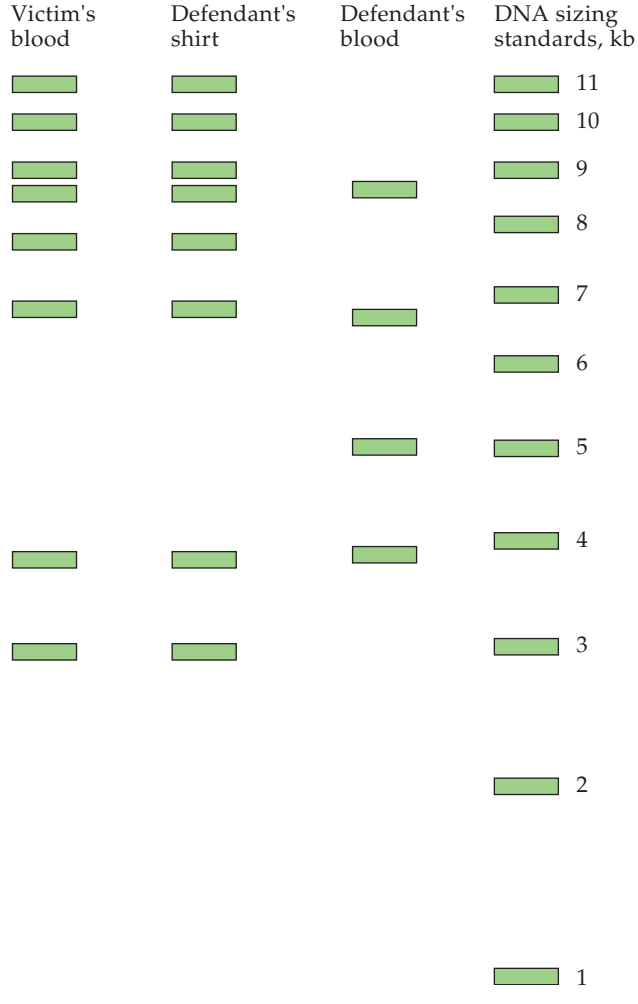


FIGURE 9.16 Schematic representation of human minisatellite DNA. Only one DNA strand is shown. In this example, the repeating unit is 9 bp, and there are 5, 6, and 7 repeating units per cluster (although 10 to 30 repeating units are more common).

sequences until differences are revealed (Fig. 9.19). Like other molecular markers, RAPDs can be used to characterize whole genomes, individual chromosomes, or, less commonly, specific genes. Although the procedure was originally developed for plants, it is also useful in the characterization of microorganisms.

In comparison with other procedures for characterizing complex DNA, the RAPD procedure has a number of advantages. (1) The same (universal) set of oligonucleotide primers can be used for all plant species. (2) No genomic libraries, radioactivity, Southern transfers, or DNA hybridization reactions are required, so a large number of samples may be easily and rapidly characterized. (3) The process can be automated. Moreover, with conventional PCR analysis it is necessary to know the sequence of a specific gene or gene segment that is the target for amplification. On the other hand, amplification in RAPD analysis occurs anywhere in a genome where

FIGURE 9.17 Southern blot of a forensic DNA sample. The DNA samples from the victim, the defendant's shirt, and the defendant were treated with the same restriction enzyme. Here, the banding pattern of the DNA extracted from the blood on the defendant's shirt is identical to the victim's DNA banding pattern and different from the defendant's pattern. The sizes of the DNA molecules in these bands are estimated by comparison with the positions of the sizing standards.



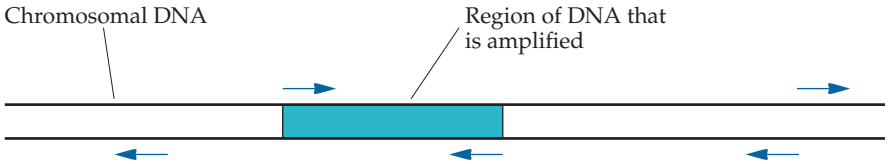
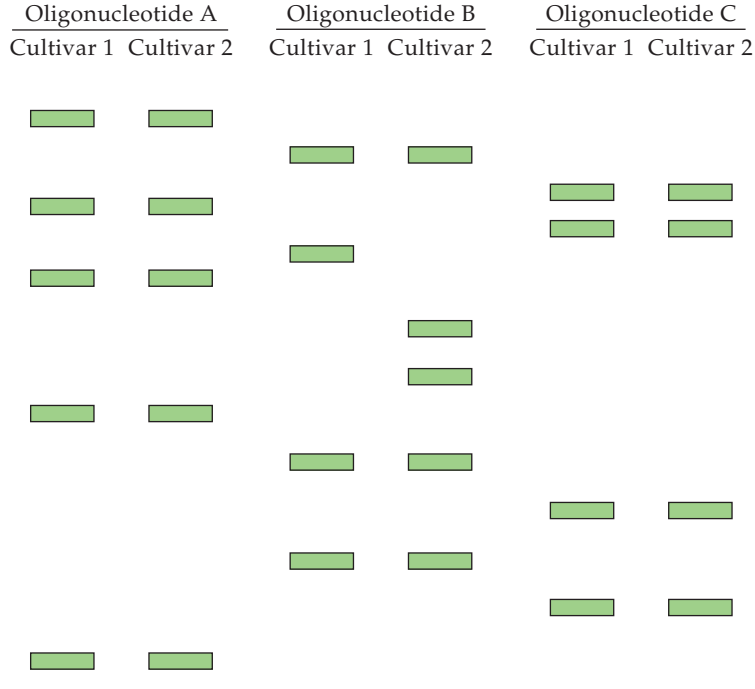


FIGURE 9.18 Binding of a single random oligonucleotide (arrow) to the chromosomal DNA of an animal, plant, or microbe. When two of the oligonucleotides on opposite strands are oriented facing one another and are 100 to 3,000 bp apart, the intervening DNA is amplified by PCR.

there are two sequences complementary to the primer that are within the length limits of the PCR.

With this technology, scientists were able to distinguish six inbred maize lines from each other, and maize hybrids produced by genetic crossings of these inbred lines were shown to have the PCR products of their parental inbred lines. RAPD markers have also been used to screen different strains of the fungus *Leptosphaeria maculans*, which is the causal agent of blackleg disease in crucifers. Differences between avirulent (non-disease-causing) and virulent (disease-causing) strains could be distinguished on the basis of specific RAPD markers, making it easier for scientists to produce an avirulent strain that could be used as a biological control agent that helps to prevent blackleg disease.

FIGURE 9.19 Ethidium bromide-stained bands following polyacrylamide gel electrophoresis of PCR-amplified plant DNA. Three separate oligonucleotide primers were used to amplify fragments from each of the two cultivars. Cultivars 1 and 2 show identical patterns of bands with oligonucleotides A and C. However, they have different patterns when oligonucleotide B is used; hence, oligonucleotide B can be used to distinguish between cultivars 1 and 2.



Real-Time PCR

By labeling the DNA that is amplified in a PCR with a fluorescent dye and monitoring the fluorescence that results when the dye bound to double-stranded DNA is irradiated with light of a certain wavelength, it is possible to "watch" the production of PCR products. Moreover, this approach allows one to quantify the amount of a specific DNA fragment in the starting material. Labeling the DNA is achieved using any one of a variety of protocols. In the simplest approach to labeling DNA, researchers add dyes that bind to double-stranded DNA and emit fluorescence, and the fluorescence intensity increases in proportion to the concentration of double-stranded DNA (Fig. 9.20).

Real-time PCR may be described as occurring in four phases (Fig. 9.21). In the first, or linear, phase (Fig. 9.21, phase 1), which generally takes about 10 to 15 cycles, fluorescence emission at each cycle has not yet risen above the background level. In the early exponential phase (Fig. 9.21, phase 2), the amount of fluorescence reaches a threshold at which it is significantly higher than the background. The cycle at which this occurs is known as the threshold cycle (C_T, or CP, depending upon the manufacturer of the PCR equipment). The C_T value is inversely correlated with the amount of target DNA in the original sample. During the exponential phase (Fig. 9.21, phase 3), the amount of product doubles in each cycle under ideal conditions, while in the plateau stage (Fig. 9.21, phase 4), the reaction components become limited and measurements of the fluorescence intensity are no longer useful. To quantitate the amount of target DNA in a sample, a standard curve is first generated by serially diluting a sample with a known number of copies of the target DNA, and assuming all samples are amplified with equal efficiency, the C_T values for each dilution are plotted against the starting amount of sample (Fig. 9.22). The number of copies of a target DNA in a sample can be determined by obtaining the C_T value for the sample and extrapolating the starting amount from the standard curve. In addition, since during the exponential phase the DNA doubles with each cycle, a sample that has four times the number of starting copies of the target sequence compared to another sample would require two fewer cycles of amplification to generate the same number of product strands.

Among its many other uses, real-time PCR has been used to monitor *Cryptosporidium parvum* (a waterborne protozoan parasite that is the causative agent of a range of human diseases, including persistent diarrhea and severe infections, in infected individuals). This approach is likely to replace the more imprecise and time-consuming traditional methods of monitoring *C. parvum* infections, such as histological staining. Similarly, other researchers have reported using real-time PCR to quantitate *S. enterica* contamination in food samples. In this case, food samples (chicken and mung beans were tested) were rinsed with 100 to 250 mL of water or with a physiological saline solution. The liquid was filtered to remove particulate matter and then filtered to capture the *Salmonella* sp. cells. The cells were removed from the filter membrane, lysed, and subjected to real-time PCR. In this case, the entire procedure took only approximately 3 hours and was able to detect and quantitate cell numbers as low as 7×10^2 colony-forming units (i.e., cells) per 100 mL of liquid. Compared to the existing methodology, real-time PCR offers a dramatic improvement in both the sensitivity of detection and the time that it takes to complete the analysis.

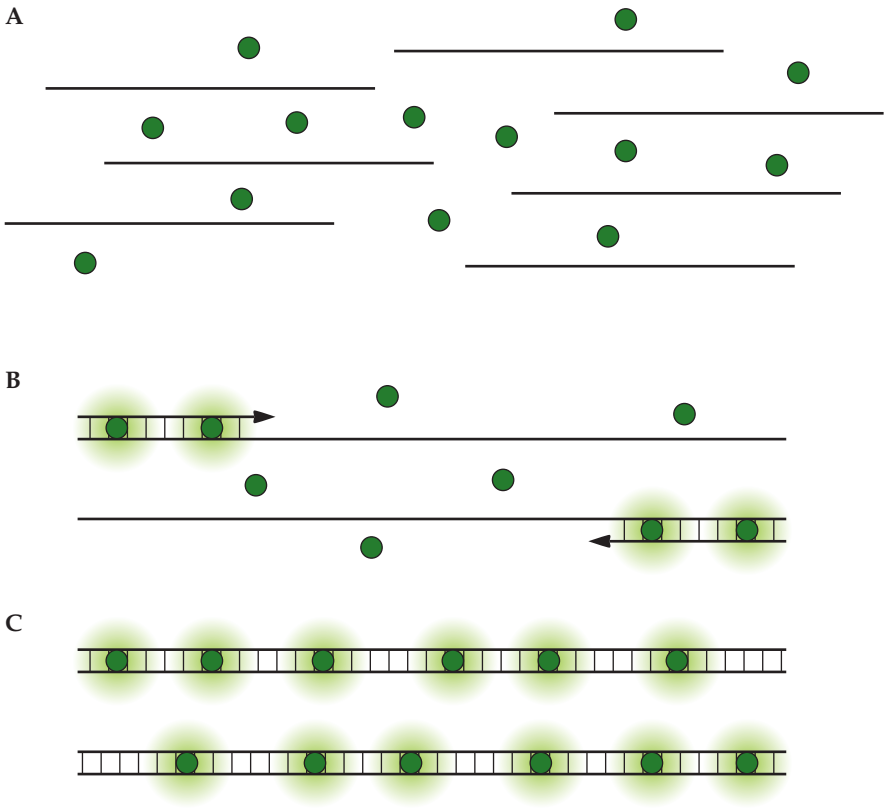


FIGURE 9.20 The fluorescent dye SYBR green does not bind to single-stranded DNA **(A)**, binds to double-stranded DNA as it is synthesized **(B)**, and is bound to the double-stranded amplified DNA **(C)**. Only the bound DNA fluoresces.

In addition to its use in the measurement of pathogenic agents in the environment, a variant of real-time PCR may be used to quantitate the levels of a variety of mRNAs in different eukaryotic tissues or prokaryotic cells. In this case, since the initial target is RNA and not DNA, a reverse transcription (RT) step is needed before the real-time PCR. In the first step of real-time RT-PCR, the mRNA sample is reverse transcribed to generate complementary DNA (cDNA). This may be done in the same tube as the subsequent PCR, or the RT reaction and PCR may be carried out in separate tubes. Many of the more traditional methods of monitoring gene expression, including Northern hybridization, ribonuclease (RNase) protection assays, and RNA dot blot hybridizations, are both limited in sensitivity and difficult to quantify. However, with the increasingly popular technique of real-time RT-PCR, it is possible to detect and quantify mRNA levels that are about 10,000- to 100,000-fold lower than those measurable by traditional techniques. With real-time RT-PCR, even a single copy of a transcript may be detected. The very low levels of RNA that are required for this procedure make it the method of choice for monitoring mRNA levels.

Immunoquantitative Real-Time PCR

The detection limits of many commercially available immunological methods for measuring levels of pathogenic microorganisms are often insufficient to

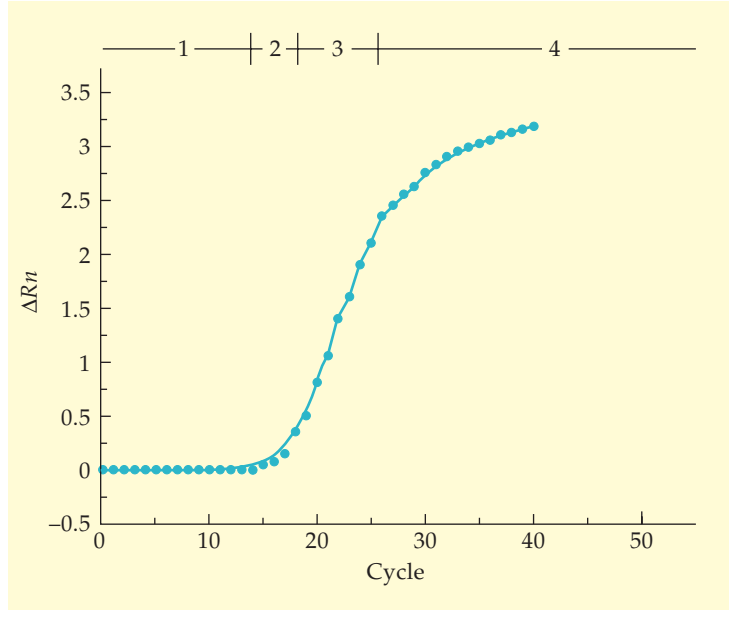
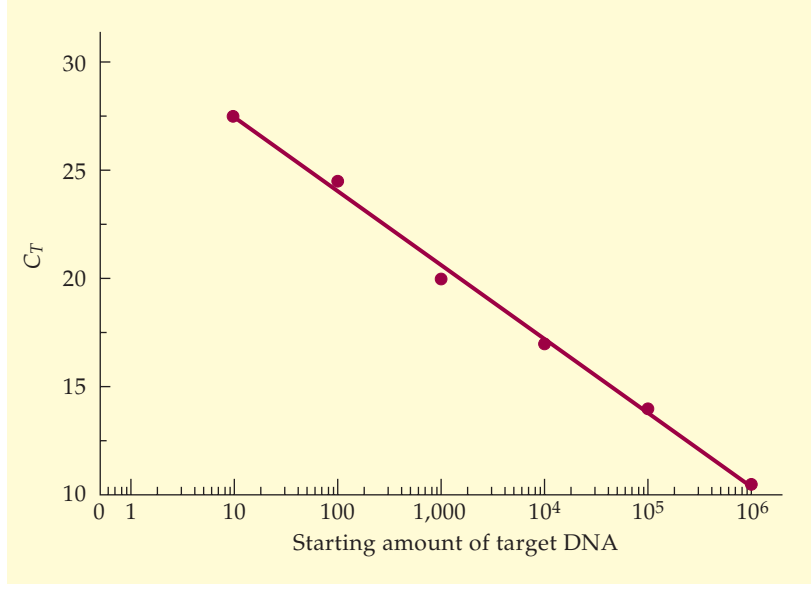


FIGURE 9.21 A plot of ΔRn (normalized fluorescence) versus cycle number in a real-time PCR experiment. Four phases of PCR are shown. (1) A linear phase, where fluorescence emission is not yet above background level. (2) An early exponential phase, where the fluorescence intensity becomes significantly higher than the background. The cycle at which this occurs is generally known as C_T. (3) An exponential phase, where the amount of product doubles in each cycle. (4) A plateau phase, where reaction components are limited and amplification slows down.

FIGURE 9.22 Plot of C_T versus the starting amount of a target nucleotide sequence. Fluorescence detection is linear over several orders of magnitude.



perceive low but still potentially dangerous levels of these organisms. Notwithstanding the high degree of specificity that antibodies provide, it would be advantageous to be able to increase the sensitivity of various immunological assay procedures. One way to do this is to develop a protocol that combines the specificity of antibodies with the sensitivity of PCR. Figure 9.23 compares an ELISA-type protocol with an immunoquantitative real-time PCR procedure. With the ELISA method (Fig. 9.23A), the first antibody is coupled to the surface of a microtiter plate. The added antigen binds to the first antibody. When the second antibody is added, it binds to a different epitope on the antigen. The bound antigen is visualized by the action of alkaline phosphatase, bound to the second antibody, which turns a colorless substrate into a colored product. With the immunoquantitative real-time PCR procedure (Fig. 9.23B), instead of alkaline phosphatase, a streptavidin-biotin complex links the second antibody to a 246-bp DNA fragment with a known sequence. Once the immunological complex has formed, it may be visualized and quantified by performing real-time PCR in the well of the microtiter plate, thereby significantly amplifying the signal from the immunological complex. In fact, it has been estimated that this procedure is approximately 1,000-fold more sensitive than an ELISA.

Ancestry Determination

By examining a number of different single-nucleotide polymorphisms (SNPs) (i.e., minor variations in DNA sequence) in an individual and comparing the pattern of the SNPs to those of other individuals in the population, it is possible to infer information regarding an individual's ancestry. For an analysis of an individual's ancestry, three different types of DNA can be examined: autosomal DNA (which includes all of a person's DNA except for the X and Y chromosomes and mitochondrial DNA), which originates from a combination of a person's parents' DNA; paternal DNA (i.e., the Y chromosome), which is passed on from father to son; and maternal DNA (i.e., mitochondrial DNA and the X chromosome), which is passed on from a mother to all of her children.

To perform an analysis of an individual's ancestry (or for paternity testing or forensic analysis), DNA is typically extracted and purified from buccal swabs (i.e., from cheek cells) or from blood samples. Stretches of DNA are then amplified by PCR using primers that target specific regions of the genome. The DNA samples are then separated by size on a small column by a technique known as capillary electrophoresis (this method has generally replaced gel electrophoresis, which was previously used to separate these small DNA fragments but was slower and less amenable to automation). If the PCR primers are labeled with fluorescent dyes before the PCR amplification reaction, fluorescent samples are eluted from the capillary column in a characteristic pattern of bands (in much the same way that DNA bands form a pattern unique to a particular fragment following gel electrophoresis). The sizes of the amplified DNA bands in specific regions are determined and are referred to as alleles. With autosomal DNA, each person should have two alleles at each site, one from the father and one from the mother. Each known allele has a determined frequency of occurrence in the general population and among various ethnic groups. After testing for approximately 150 to 350 different SNPs, the frequency of certain

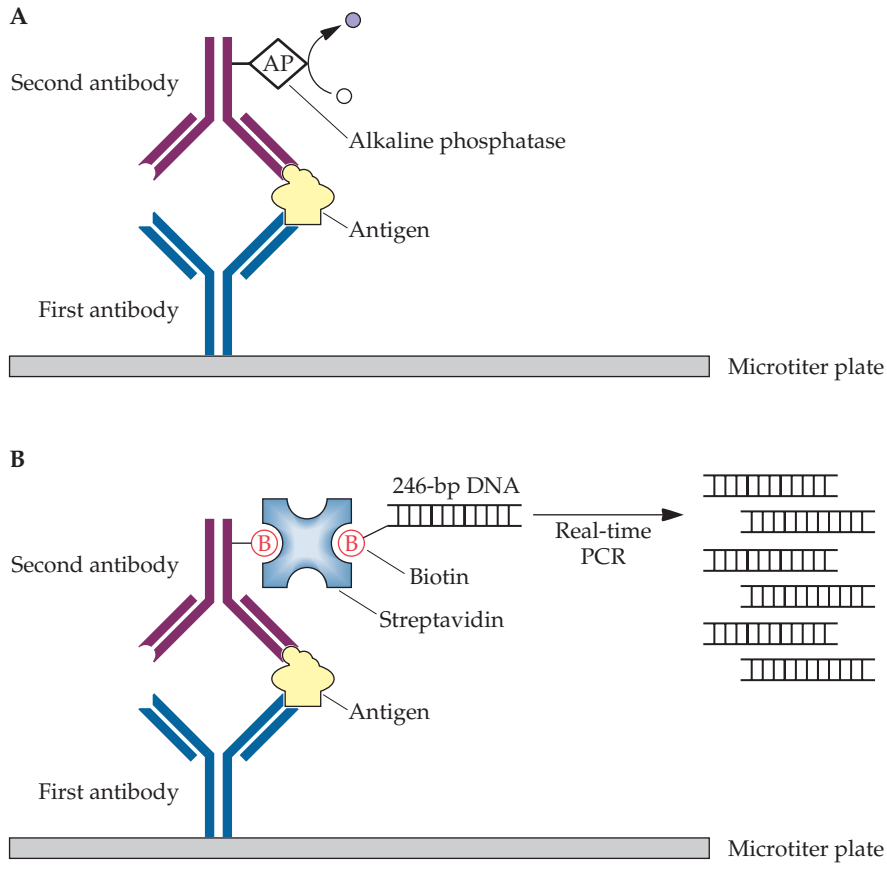


FIGURE 9.23 Antigen detection by ELISA with the substrate color change detected spectrophotometrically **(A)** and by immunoquantitative real-time PCR with the amplified DNA detected by measuring the fluorescence of a dye bound to the double-stranded DNA **(B).**

alleles in an individual is compared to the frequency of those alleles in various ethnic groups. This provides an indication of the ethnic background or ancestry of that individual. For example, this type of testing may indicate that an individual is genetically 70% Northern European, 17% Middle Eastern, and 13% Native American.

Since mitochondrial DNA changes only very little over many generations, characterization of mitochondrial DNA is an ideal means of tracking migrations over many hundreds of generations of human genetic history. Analysis of mitochondrial DNA indicates that the observed genetic variation in human populations may be divided into divisions of ethnically similar individuals called haplotypes. Table 9.3 summarizes the currently accepted mitochondrial DNA haplotypes and the groups associated with these haplotypes. The root of all human lineages is the L groups in Africa. All other groups diverged from these groups after early humans began to migrate out of Africa around 150,000 years ago.

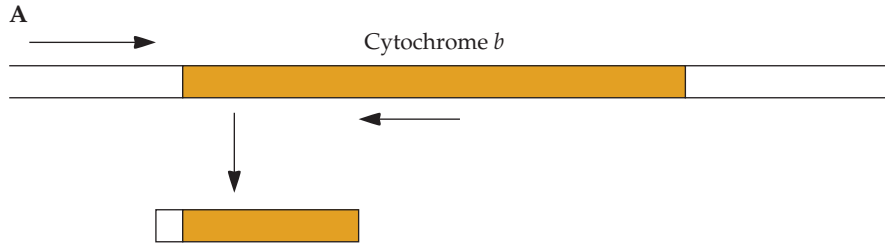
Analysis of parental (Y chromosome) haplotypes has been used to examine the claim that all members of a priestly line of Jewish males called "Kohanim" are descended from the family of the biblical Aaron, the brother of Moses. According to Jewish tradition, membership in this priestly line may be acquired only by males whose biological fathers are Kohanim.

When Y-linked genetic markers were examined among Jews claiming to be Kohanim, as well as men who either were not Jewish or did not claim to be Kohanim, the limited variation in the markers examined was entirely consistent with a 3,300-year-old origin of this priestly line in a single male or a small group of related males. Interestingly, these same Y chromosome genetic markers are found among the Lemba, a South African tribal group claiming paternal kinship with the Jews of Yemen.

By mid-2008, there were nearly 30 companies marketing genetic ancestry kits; each test costs from $100 to $900. Moreover, in the few years that they have been available, more than 500,000 people have purchased these tests, and there is every indication that the demand will continue to grow. Some geneticists, anthropologists, and epidemiologists have publicly expressed concern that naïve laypersons, anxious for definitive information about their personal genetic ancestry, often misinterpret the results of these tests. For example, the fact that a particular allele or haplotype is

TABLE 9.3 Human mitochondrial haplotypes and the groups associated with these haplotypes

Haplotype group	Associated traits
A	Originated in Asia ~60,000 years ago; currently found widely in Asia; a precursor of Native Americans
B	Originated in Asia ~50,000 years ago; subgroup B2 is one precursor of Native Americans
C	Originated in Asia ~60,000 years ago; includes the Siberian region of northern Asia and is a precursor of Native Americans
CZ	Found in modern Eurasian populations in northern and eastern Asia, including Siberia
D	Originated in Asia ~60,000 years ago; along with groups A, B, and C, this group is thought to have produced Native Americans; currently found in northern and eastern Asia
E	Not a well-characterized group; currently found among some people in Argentina
F	Originated in eastern Asia; from haplotype group R1; currently found in China and Japan
G	Currently found in northeastern Siberia; also found among indigenous people of Kamchatka
H	Common in the Middle East and northern Africa; ancestor to about half of all Europeans; a prominent subgroup of HV
HV	Originated ~20,000 years ago; a progenitor of groups H and V; commonly found in modern western Europeans
I	Originated ~30,000 years ago; currently found in both southern Europe and northern Africa
J	Originated in Mesopotamia ~10,000 years ago; currently found in Russia and eastern Europe
JT	Derived from group R and a progenitor of groups J and T
K	Originated ~18,000 years ago in Eurasia; currently found in some parts of western Europe
L1	Originated ~150,000 years ago in Africa; this haplotype group represents the group from which all of humankind is thought to descend; currently found in West and Central Africa
L2	Originated from haplotype group L1 in Africa ~70,000 years ago; currently commonly found in sub-Saharan Africa and among American blacks
L3	Originated from haplotype group L1; gave rise to haplotype groups M and N; currently commonly found in East Africa
M	Originated from haplotype group L3 ~80,000 years ago; this group is thought to have migrated into Eurasia ~60,000 years ago; currently found throughout southern Asia
M1	Believed to be the result of migration from North Africa and parts of Asia to sub-Saharan Africa
N	Originated from haplotype group L3 ~80,000 years ago; this haplotype group is believed to be the progenitor of nearly all European haplotype groups; groups B, U, F, HV, H, and V all descend from haplotype group N
N1a	Currently found in Iran and other parts of western and central Asia
N1b	Common in Near and Middle East regions of Asia, as well as among Ashkenazi Jewish people
Pre-HV	Widely represented in the Middle East and parts of eastern Africa; the ancestor of haplotype groups HV, H, and V
Q	Currently found in the southern Pacific region, especially New Guinea and Melanesia
R	Descended from haplotype group N; currently found throughout Asia and Eastern Europe



B

5′ ATGACCAACA TTCGAAAATC CCACCCACTA ATAAAAATTA TAAACAACTC 3′

FIGURE 9.24 **(A)** PCR amplification of a portion of mammalian mitochondrial DNA. **(B)** The consensus sequence of the first 50 nucleotides of the cytochrome *b* gene. Minor variations from the consensus sequence are characteristic of specific genera and species of various mammals. Adapted from Hsieh et al., *Forensic Sci. Int.* **122:**7–18, 2001.

quite common in certain populations does not mean that its presence is a definitive diagnostic of membership in those populations. There is often high genetic diversity within populations, and gene flow can readily occur between populations. Thus, a particular allele could have been inherited from a population in which it is less common rather than from the population in which it appears to be "diagnostic," leading some individuals to mistakenly believe that they are genetically part of a particular ethnic or racial group.

Animal Species Determination

In many parts of the world, numerous large mammalian species have been hunted to the verge of extinction because of the trade in their skin, bones, horns, or other body parts. As part of an effort to stem the illegal traffic in wild-animal remains and enforce international conventions designed to prevent this trade, a number of laboratories have been set up to determine the species of origin from powdered animal remains (so prepared because of their supposed medicinal properties). Currently, the method of choice for animal species determination involves DNA typing. More specifically, using PCR, a portion of the animal's cytochrome *b* gene, which is found in the mitrochondrial DNA, is amplified and then sequenced (Fig. 9.24). This locus was chosen because it is both sufficiently conserved so that it is present in all mammals and sufficiently polymorphic that members of different, but closely related, species can be distinguished. The primers used in this method amplify a 486-bp DNA fragment, which is sufficient for DNA sequencing and to identify most mammals despite the fact that the sample DNA is often somewhat degraded. To ensure that high-quality data are obtained, it is necessary to (1) start with 20 to 50 ng of DNA template and (2) use a control DNA sample (usually mouse or cow DNA) that is treated in parallel to the target DNA.

Automated DNA Analysis

More than 1,400 different organisms (bacteria, fungi, viruses, and protozoa) have been identified as being pathogenic to humans. Moreover, new patho-

gens are constantly being identified and characterized. To limit the damage from either a natural epidemic or a bioterrorist attack, it is necessary to rapidly identify the organism(s) that is the source of the infectious outbreak so that appropriate public health measures may be instituted as rapidly as possible. In addition to human diseases, it is also important to rapidly identify the causative agents of outbreaks of animal and plant diseases.

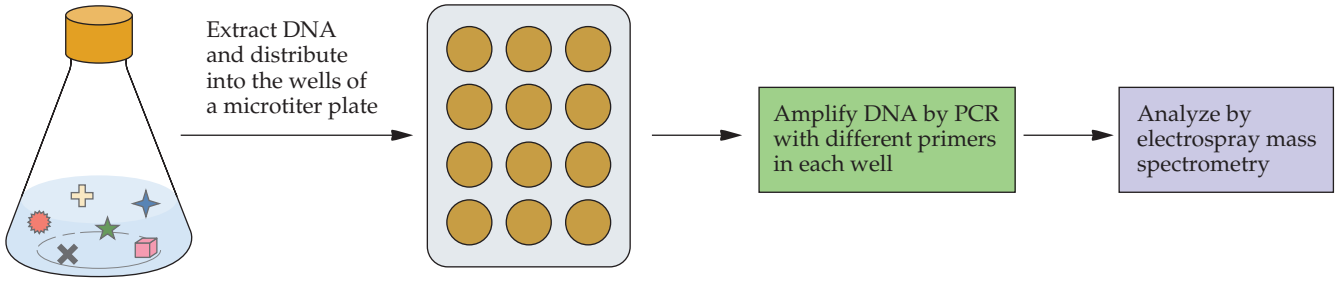
By performing a combination of PCR and electrospray ionization mass spectrometry (Fig. 9.25), it is possible to rapidly and accurately identify a wide range of human pathogens. Small aliquots of the DNA that is isolated from a test sample are placed into a number of different wells of a microtiter plate. Each well contains a pair of PCR primers that have been designed to amplify a gene product from a broad group of organisms within a selected domain of microbial life. For example, workers have reported using 12 primers that are targeted to universally conserved sequences and 6 primers that are targeted to broad divisions of microbial life (such as bacilli). The products of each of the PCRs are electrosprayed into a mass spectrometer, and the DNA base sequences of the various samples are determined. This technique allows scientists to very rapidly hone in on the nature of an infectious agent. Of course, this technique is facilitated by the fact that the genomic DNA sequences of a large number of microbes have already been determined. Moreover, with the very rapid progress that has been made recently in DNA sequence analysis, it is reasonable to expect that all known pathogens will be fully sequenced by 2015.

Molecular Diagnosis of Genetic Disease

The ability to diagnose the occurrence of specific inherited diseases in humans at the genetic level makes it possible for individuals to discover whether they or their offspring are at risk. DNA analysis can be used for the identification of carriers of hereditary disorders, for prenatal diagnosis of serious genetic conditions, and for early diagnosis before the onset of symptoms.

Tests at the DNA level are definitive for determining the existence of specific genetic mutations. Previously, genetic testing relied almost exclusively on biochemical assays that scored either the presence or the absence of a gene product. A DNA-based test does not, however, require expression for detection of the mutant gene, thereby making it theoretically possible to develop screening assays for all single-gene diseases.

FIGURE 9.25 Flowchart of an automated system to identify the pathogenic microbes in an environmental sample.



Screening for Cystic Fibrosis

Often, screening for genetic diseases can be rather complex. This reflects the fact that instead of a disease being the consequence of a single alteration to the wild-type DNA, as is the case with sickle-cell anemia (see below), many diseases are caused by any one of a large number of genetic alterations to the normal DNA for that gene. For example, cystic fibrosis, one of the most common lethal autosomal recessive disorders in Europeans and their descendants, with an incidence of approximately 1 in every 2,500 live births and a carrier frequency of approximately 1 in 29, is caused by mutations to the cystic fibrosis transmembrane conductance regulator (CFTR) gene that result in defects in chloride ion transport. There are currently nearly 1,400 known mutations to the CFTR gene that can result in the development of cystic fibrosis. Screening individuals who may be at risk for cystic fibrosis for 1,400 different mutations is a daunting task. Fortunately, some of the mutations that cause cystic fibrosis are much more common than others (Table 9.4). In fact, over 90% of cystic fibrosis patients carry at least one ΔF508 allele, and nearly 50% of cystic fibrosis cases are individuals who are homozygous for ΔF508. Despite the fact that separate tests are required for each mutation, it is estimated that screening individuals for ΔF508 and for the next 20 most common mutations should identify approximately 98% of cystic fibrosis-affected individuals and carriers.

Current diagnostic tests for cystic fibrosis include several different techniques. One of the most widely used methods is allele-specific oligonucleotide dot blots (also called allele-specific hybridization). With this technique, genomic DNA or cDNA from an individual is amplified by PCR and, following transfer to a membrane, is hybridized (separately) to labeled oligonucleotide probes for the mutant (usually ΔF508) and wild-type genes (Fig. 9.26). In this way, it is possible to distinguish between normal individuals, cystic fibrosis carriers, and cystic fibrosis-affected individuals (Fig. 9.27). With this technique, the probe or the probe–target complex may be labeled in a variety of ways, including the use of radioactivity, enzymes that produce color change when acting on certain substrates (see the discussion of the ELISA procedure above), and fluorescent dyes. This technique may be automated and is currently commercially

TABLE 9.4 The most common mutations of the CFTR protein that lead to cystic fibrosis

Mutation designation	Amino acid change to the CFTR protein
ΔF508	Deletion of phenylalanine at position 508
G542X	Replacement of glycine at position 542 by a stop codon
W1282X	Replacement of tryptophan at position 1282 by a stop codon
N1303K	Replacement of asparagine at position 1303 by lysine
1717-1G>A	Replacement of glycine by alanine at the last nucleotide in the intron proceeding nucleotide 1717 in the cDNA
R553X	Replacement of arginine at position 553 by a stop codon
I148T	Replacement of isoleucine at position 148 by threonine
3120+1G>A	Replacement of glycine by alanine at the first nucleotide in the intron following nucleotide 3120 in the cDNA

Adapted from Eshaque and Dixon, *Biotechnol. Adv.* **24:**86–93, 2006.

Amino acids are numbered starting at the N-terminal end of the protein.

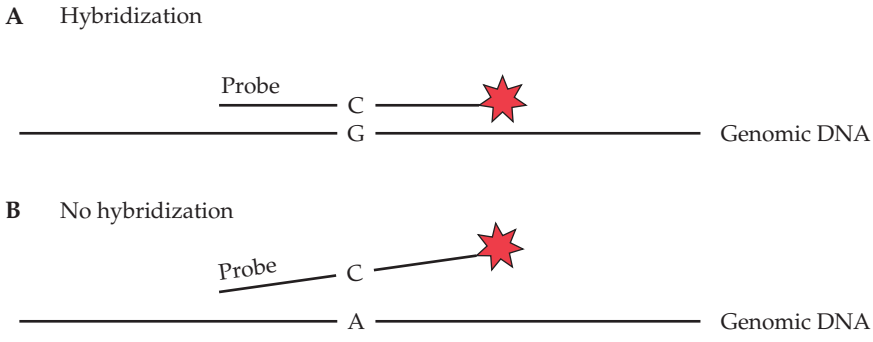


FIGURE 9.26 Labeled oligonucleotide probe hybridizes to a completely complementary genomic DNA but, under stringent conditions **(A)**, not to a DNA sequence in which one of the bases in the middle has been altered **(B)**. In this case, the probe was designed to hybridize to wild-type DNA. Other probes, designed to hybridize to the genomic DNA of known mutations, would not bind to the wild-type DNA.

available in a kit form that can detect 12 frequent and 17 rare cystic fibrosis mutations.

Another method that has been marketed as a kit is based on the PCR amplification of specific alleles. Using this protocol, several PCRs are performed simultaneously for each DNA sample—the primers anneal to different regions for different mutations. Following amplification, the presence of a DNA band of a particular size indicates that a specific mutation is present. In this case, different-size DNA fragments are typically separated either by gel electrophoresis or by capillary electrophoresis. This test is quite rapid, and it has the ability to detect a variety of different mutations. However, it does not distinguish between homozygotes and heterozygotes, so a positive response must be followed up by additional tests to determine whether a positive test is indicative of a cystic fibrosis gene carrier or affected individual.

The PCR/OLA procedure (described in detail below) is also commonly used to detect cystic fibrosis mutations. This technique is considered to be highly accurate compared to many other protocols and has the highest detection rate of any of the diagnostic tests for this disease. Moreover, it is readily amenable to automation. Notwithstanding the success with all of the procedures mentioned above, researchers continue to refine and develop these and other approaches to the diagnosis of cystic fibrosis.

Sickle-Cell Anemia

Sickle-cell anemia is a genetic disease that is the result of a single-nucleotide change in the codon for the sixth amino acid of the β chain of the hemoglobin molecule. In individuals homozygous for the defect *(S/S)*, the shape of the red blood cells is irregular (sickle shaped) because the conformation of the hemoglobin molecule is distorted by a single amino acid change from glutamic acid to valine. The biological ramifications of this genetic alteration are severe anemia and progressive damage to the heart, lungs, brain, joints, and major organ systems. The anemia is caused by the inability of the mutated hemoglobin to carry sufficient oxygen. The life expectancy for *S/S* homozygotes is quite short. Heterozygous individuals *(A/S)* (genetic carriers) have normal-shape red blood cells and no symptoms unless they are subjected to extreme conditions, such as high altitude

or extremes of temperature, that lower the oxygen supply. If both parents are heterozygous, there is a 25% chance that a child of theirs will have sickle-cell anemia, i.e., will be an *S/S* homozygote. The sickle-cell anemia gene occurs with high frequency among black Africans and their descendants and in Hispanic populations. Carrier screening for the sickle-cell anemia gene is routinely conducted in the United States so that those individuals who are at risk for transmitting the gene to their offspring can be identified. One of the test systems is described below.

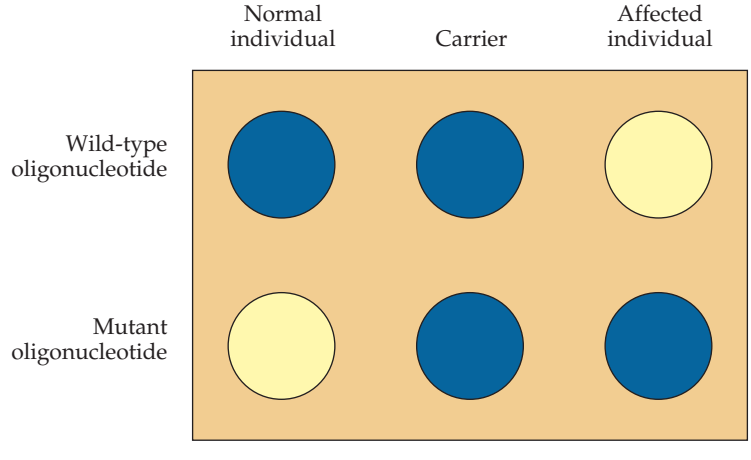
The single-nucleotide change in the β-globin gene that causes sickle-cell anemia by chance abolishes a CvnI restriction endonuclease site. This restriction enzyme recognizes the sequence CCTNAGG and cleaves the DNA between the C and the T. (The letter N indicates that any one of the four nucleotides can occupy this position.) In the normal gene, the DNA sequence is CCTGAGG, whereas in the sickle-cell anemia gene, the sequence is CCTGTGG. This difference forms the basis for a DNA diagnostic assay (Fig. 9.28A).

After two oligonucleotide primer sequences that flank the CvnI site are added, a small amount of sample DNA can be amplified by PCR (Fig. 9.28B). The amplified DNA is digested with CvnI (Fig. 9.28C), and the cleavage products are separated by gel electrophoresis and visualized by ethidium bromide staining of the DNA in the gel. If the CvnI site is present, a specific set of DNA fragments is observed (Fig. 9.28D). A different profile of DNA fragments occurs if the CvnI site is absent. By this procedure, the genetic makeup of a tested person can be determined quickly, directly, and easily. Moreover, because of the fortuitous loss of the CvnI site, this assay functions without the need for a target–probe hybridization reaction.

The PCR/OLA Procedure

Obviously, not all genetic changes that produce defective genes affect existing restriction endonuclease sites. Therefore, other strategies for detecting single-nucleotide changes are required. One of these procedures

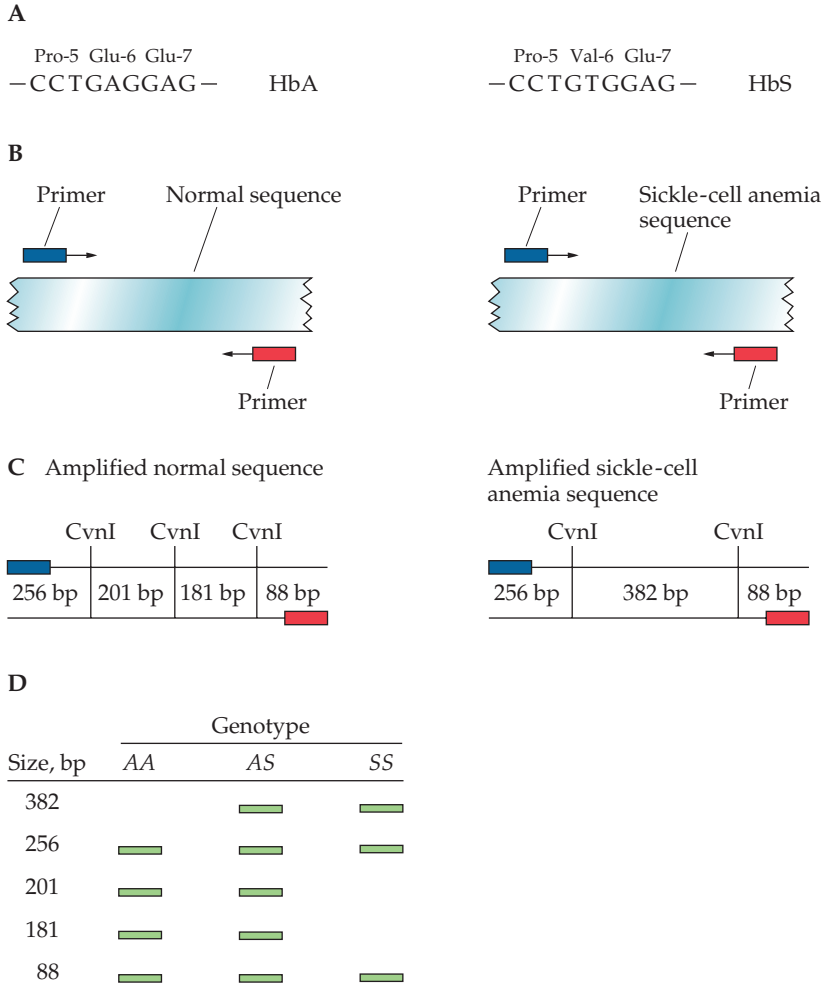
FIGURE 9.27 Allele-specific oligonucleotide dot blot to diagnose individuals who are either carriers of a mutant CFTR gene (heterozygotes) or affected by the disease (homozygotes). The dark dot blot indicates that the labeled oligonucleotide has bound to the individual's DNA.



combines PCR with an oligonucleotide ligation assay (OLA); not surprisingly, it is called PCR/OLA.

Let us assume that in a normal gene at a specific site (say, nucleotide number 106) the nucleotide pair is A·T; in the mutant form, the nucleotide pair at this site is G·C. Knowledge of the sequence of nucleotides on both sides of position 106 enables the design and use of two short (20-nucleotide) adjacent oligonucleotide sequences that are complementary to one of the two native DNA strands (Fig. 9.29). The essential feature of this pair of oligonucleotides is that one of them (probe X) has as its last base at the 3′

FIGURE 9.28 Detection of the sickle-cell anemia gene at the DNA level. **(A)** A portion of the sequence of the wild-type (HbA) and sickle-cell (HbS) human β-globin gene. The amino acids (numbered from the N-terminal end of the peptide chain) encoded by this portion of the DNA are shown above the DNA sequence. **(B)** PCR amplification of the portion of the β-globin gene containing the CvnI recognition site that is altered in the mutant gene. **(C)** CvnI digestion of the PCR products. The normal (wild-type) gene has three CvnI sites between the PCR primers, and the mutant gene has two. **(D)** Size distribution of fragments following gel electrophoresis of CvnI-digested PCR-amplified β-globin DNA. *AA*, homozygous condition for the normal β-globin gene; *AS*, heterozygous condition; *SS*, homozygous condition for the sickle-cell anemia β-globin gene.



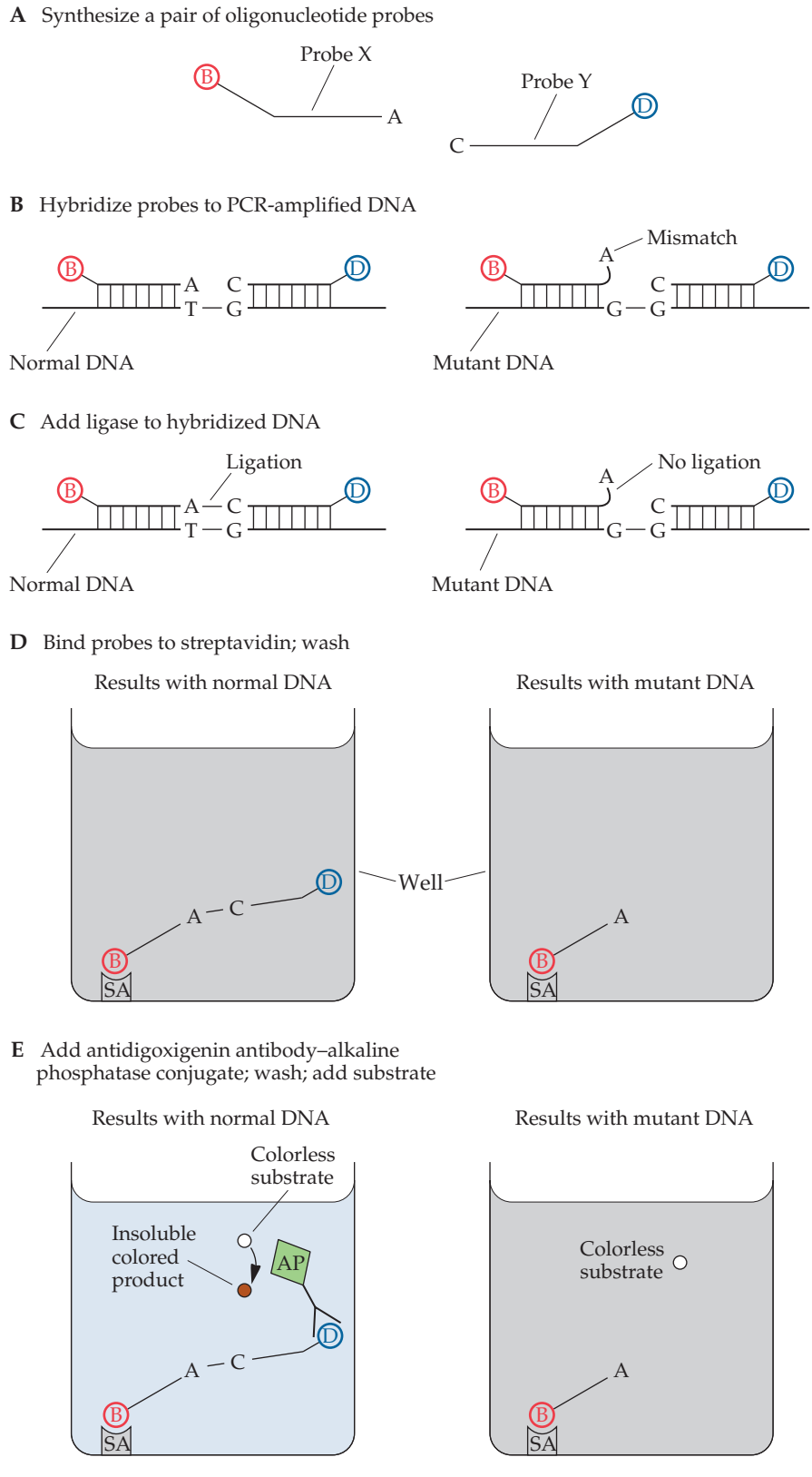


FIGURE 9.29 PCR/OLA procedure. B, biotin; D, digoxigenin; AP, alkaline phosphatase; SA, streptavidin; A, adenine; C, cytosine; G, guanine; T, thymine.

end the nucleotide that is complementary to the nucleotide at position 106 of the normal sequence. The other oligonucleotide (probe Y) starts at its 5′ end with a nucleotide that is complementary to the nucleotide immediately adjacent to position 106. When these two probes are hybridized with target DNA containing the normal sequence (which has been amplified by PCR), the nucleotide at the 3′ end of probe X base pairs with the target DNA, and probe Y is aligned so that its 5′ end lies next to the 3′ end of probe X. The addition of DNA ligase to the reaction covalently joins probe X and probe Y. By contrast, when these two probes are hybridized to mutant DNA in which the nucleotide at position 106 is altered, the nucleotide at the 3′ end of probe X is mismatched and is not able to pair with nucleotide 106 in the target DNA sequence; probe Y, however, is perfectly aligned. In this case, DNA ligase cannot join probe X and probe Y because of the single-nucleotide misalignment.

Other oligonucleotides (probes) can also be chemically synthesized to give a perfect base pair match when nucleotide 106 is mutated. Obviously, with this second set of probes, ligation occurs when they are hybridized to target DNA that contains the mutant nucleotide, whereas with normal target DNA, the single nucleotide pair mismatch prevents the ligation of the probes. In short, PCR/OLA is designed to distinguish between two possibilities: ligation and no ligation of two input probes.

To determine whether ligation has occurred between two indicator probes, probe X is labeled at its 5′ end with biotin and probe Y is labeled at its 3′ end with digoxigenin. The low-molecular-weight compound digoxigenin serves as an antibody-binding indicator. After the hybridization and ligation steps are carried out, the DNA is denatured to release the hybridized probes, and the entire mixture is transferred to a small plastic well that has been coated with streptavidin. The well is then washed to remove unbound material, so only the biotin-labeled probe DNA remains bound. Next, antidigoxigenin antibodies, which have been previously coupled to alkaline phosphatase, are added to the well. After an additional washing step to remove unbound conjugated antidigoxigenin antibodies, a colorless chromogenic substrate is added. The appearance of color in the well indicates that antidigoxigenin antibodies have bound to digoxigenin and that the digoxigenin-labeled probe was ligated to the biotinylated probe. If no color appears, then no ligation occurred.

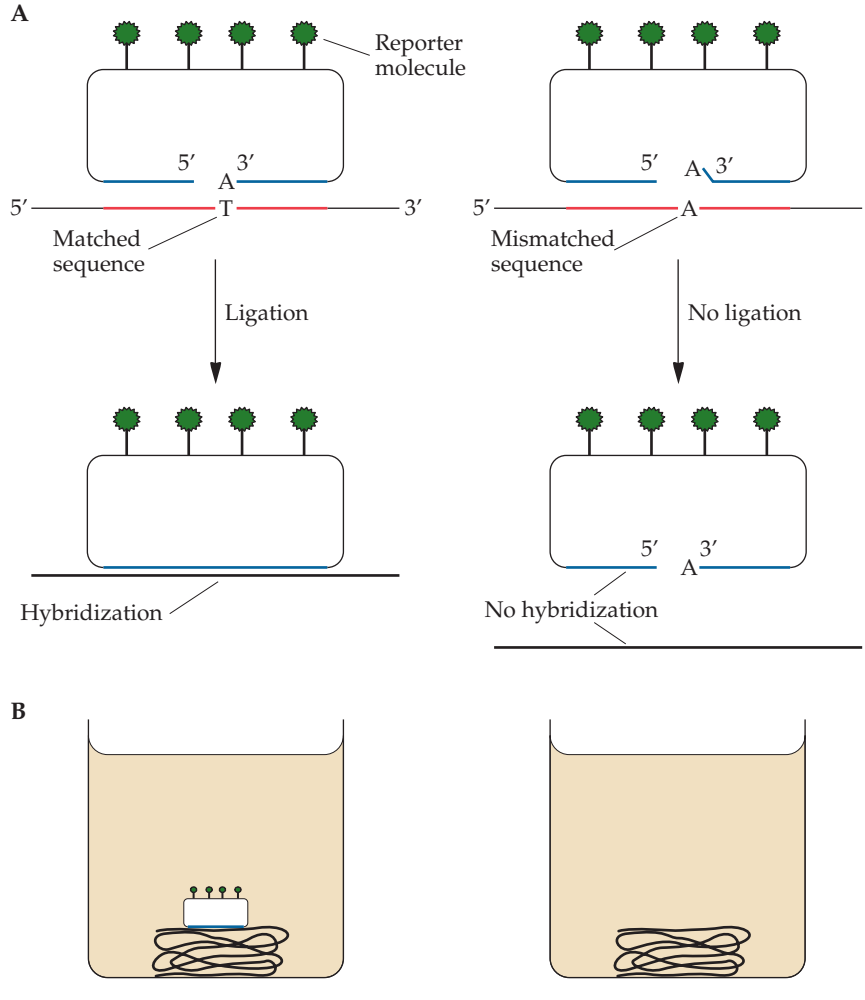
With two pairs of probes, it is possible to ascertain the genetic makeup of any tested individual at a particular site. For example, heterozygous individuals yield positive results with both pairs of probes. The DNA from people with two copies of the normal gene gives a positive response only with the set of probes that contains the nucleotide complementary to the nucleotide at the normal site. Finally, DNA from individuals with two altered gene copies will give a positive response only with the set of probes that is designed to detect the mutant site. To minimize the amount of the original sample DNA that is required for the assay, the segment of the target DNA sequence that contains the nucleotide site to be tested is amplified by PCR before the hybridization reaction.

Overall, the PCR/OLA system is rapid, sensitive, and highly specific. It has even been automated with a robotic workstation to carry out the steps of the assay procedures. Under these conditions, as many as 1,200 ligation reactions can be conducted per day.

The ligase chain reaction assay is a simpler, albeit less sensitive, variant of the PCR/OLA system. Sample DNA is mixed with an excess amount of

a pair of OLA indicator probes (as described above) in the presence of a heat-resistant DNA ligase. After an initial ligation reaction at 65°C, the temperature is raised to 94°C to denature the probe–target DNA hybrid and then lowered to 65°C to allow hybridization of the free, nonligated OLA indicator probes to the target DNA. The cycle is repeated 20 times. If the OLA indicator probes match the target DNA perfectly, then ligation will occur at 65°C during each cycle, and after 20 cycles, enough ligation product (probe X joined to probe Y) will accumulate to be observed by either gel electrophoresis or an ELISA detection system. If no ligation

FIGURE 9.30 Schematic representation of the functioning of a padlock probe. **(A)** When the bases at the 5′ and 3′ ends of the probes are completely paired to the target DNA, ligation can take place. When there is a single-base mismatch at the 3′ end of the probe, ligation cannot occur and the probe assumes a conformation that does not allow hybridization. **(B)** Under stringent conditions, the ligated probe remains bound to the target DNA, which is bound to the surface of a 96-well microtiter plate. The nonligated probe is removed during washing. The bound probe is detected by interaction with the reporter molecules. If the reporter is biotin, then avidin and a biotinylated enzyme, such as alkaline phosphatase, are added sequentially. A colored well indicates that the probe is present and bound to the target DNA or RNA.



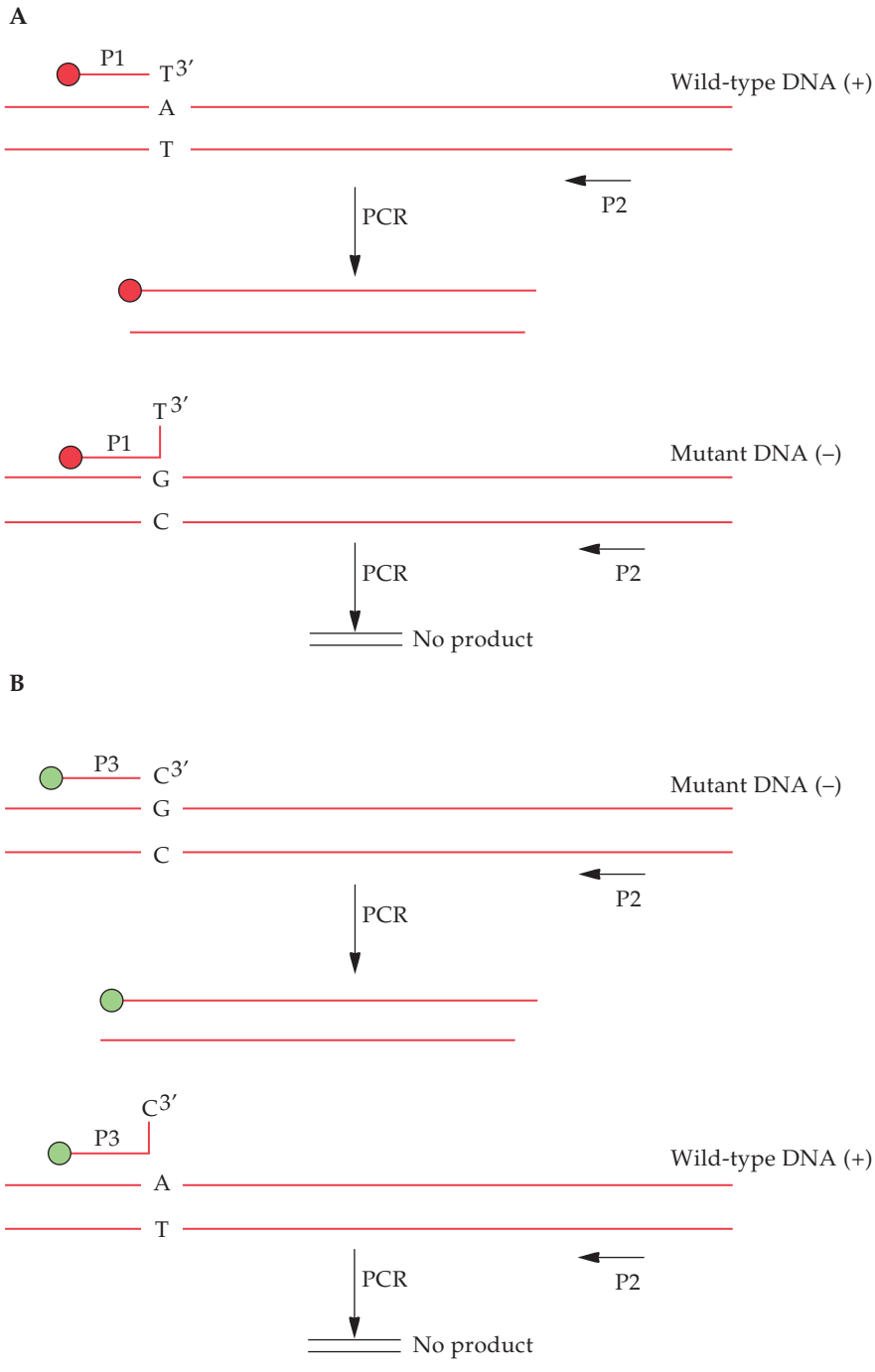


FIGURE 9.31 Detection of a single-base mutation with fluorescence-labeled PCR primers. **(A)** Primers P1 and P2 amplify DNA from the wild-type sequence. The same primers cannot amplify DNA from the mutant sequence because primer P1 is mismatched with this DNA. Primer P1 is labeled at its 5′ end with rhodamine (red). Primer P2 is unlabeled. **(B)** Primers P3 and P2 amplify DNA from the mutant but not the wild-type sequence. Primer P3 is labeled at its 5′ end with fluorescein (green). Primer P2 is unlabeled. The plus and minus signs denote wild-type and mutant sites, respectively. The genotypes 1/1, 1/2, and 2/2 produce PCR products that contain rhodamine only, rhodamine and fluorescein, and fluorescein only and that fluoresce red, yellow, and green, respectively.

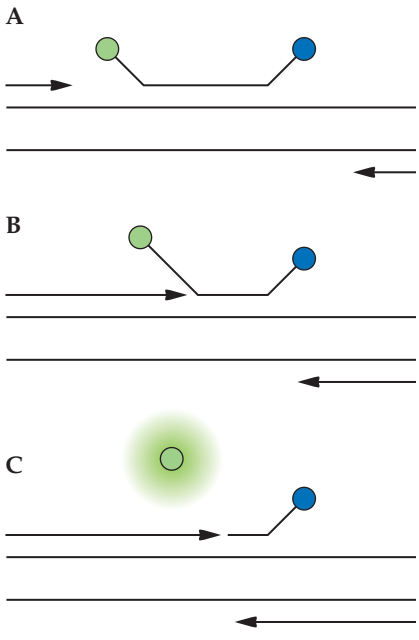


FIGURE 9.32 TaqMan assay. **(A)** A TaqMan probe that is complementary to the wild-type DNA is added prior to PCR amplification of the DNA sequence. The probe contains a fluorescent dye attached to its 5′ end (green) and a quencher attached to its 3′ end (blue). When the probe is intact, the quencher interacts with the fluorophore, quenching its fluorescence. The PCR primers are indicated by arrows. **(B)** In the extension phase of PCR, the TaqMan probe is displaced by the growing DNA strand. **(C)** Subsequently, the 5′ fluorescent dye is cleaved from the probe by the 5′ nuclease activity of the *Taq* polymerase, leading to a dramatic increase in the fluorescence of the reporter dye (shown as a starburst of color from the dye).

occurs because of a mismatch, then no joined probe product will be produced or detected.

Padlock Probes

A padlock probe is an oligonucleotide that is complementary to a target (DNA or RNA) sequence at its 5′ and 3′ ends but not in its middle region (Fig. 9.30). When a padlock probe hybridizes to its target sequence, the 5′ and 3′ ends of the probe come into close proximity with one another and the middle portion loops out. Following hybridization, if the probe is exactly complementary to the target sequence, the 5′ and 3′ ends of the probe can be joined to one another by DNA ligase. The fact that two sequences (actually two ends of the same oligonucleotide) must bind to the target ensures a high "specificity of detection." For DNA ligation to occur, both sequences must hybridize perfectly to the target; this makes it possible to easily detect allelic sequence variants. If there is a mismatch at either end, no ligation occurs. Following the ligation reaction, the probe–target hybrid can be detected because of reporter molecules, such as biotin or digoxigenin, that are attached to the middle (linker) portion of the padlock probe. Padlock probes typically have sequences approximately 15 to 20 nucleotides in length at the 5′ and 3′ ends that are complementary to the target sequence and a middle region of approximately 50 nucleotides. This procedure has become popular with researchers, as it is simpler, with fewer steps than the OLA procedure. In addition, the procedure requires one oligonucleotide compared to two for OLA, and it is amenable to automation.

Genotyping with Fluorescence-Labeled PCR Primers

PCR primers labeled with different fluorescent dyes can be used in the development of nonradioactive color-based detection systems. To distinguish between mutant and wild-type DNAs, PCR is performed with two different primers. One is exactly complementary to the wild-type DNA and is labeled at its 5′ end with rhodamine (red). The other is complementary to the mutant DNA and is labeled at its 5′ end with fluorescein (green) (Fig. 9.31). In both cases, amplification is programmed by a third, unlabeled primer that is complementary to the opposite strand. Since PCR amplification can occur only when the primer is exactly complementary to the target DNA, the presence of these three primers in the same reaction mixture will result in the amplification of either the wild-type or the mutant DNA or both, depending on which target DNAs are initially present to act as PCR templates. If an individual is homozygous for the wild-type DNA, after PCR and removal of unincorporated primer, the reaction mixture will fluoresce red; if he or she is homozygous for the mutant DNA, the reaction mixture will fluoresce green; and if he or she has both mutant and wild-type DNA (i.e., is heterozygous), the reaction mixture will fluoresce yellow. This assay can be automated and adapted for any single-nucleotide target site of any gene that has been sequenced. The problem with this technique is that it is limited to detecting an SNP. Analysis of multiple loci is not possible, since the presence of many different PCR primers in one reaction tube could lead to large numbers of cross-reactions among primer pairs, with a large number of non-specific PCR products being formed.

TaqMan Assay

The TaqMan protocol is used to check individuals for the presence of SNPs that are indicative of any of a variety of genetic diseases (Fig. 9.32). Made popular by one particular company, it is based on the 5′ nuclease activity of *Taq* polymerase, which is commonly used to amplify DNA in PCR applications. To simultaneously monitor wild-type and mutant alleles, two TaqMan probes are utilized. Each probe is exactly complementary to either the wild-type or the mutant DNA sequence, and each probe has a different fluorescent dye attached to its 5′ end. Intact probes, whether bound or unbound to cDNA, do not fluoresce because of the presence of the quencher molecule at the 3′ end of the probe. As PCR proceeds from primers flanking the probe hybridization site, the TaqMan probe is displaced by the growing DNA strand, and the 5′ nuclease activity of the *Taq* polymerase degrades the 5′ end of the TaqMan probe, thereby releasing the fluorescent dye and removing it from the proximity of the quencher molecule. Thus, only TaqMan probes that were previously bound to target DNA will be degraded and subsequently fluoresce. Any mismatched probes, due to mutations in the region where the TaqMan probe binds, will be displaced but not cleaved, so they will not fluoresce. By monitoring the fluorescence at two different wavelengths (one for each TaqMan probe), it is possible to distinguish the wild type, heterozygotes (carrying one mutant and one wild-type gene), and individuals that are homozygous for the target mutation. In fact, this technique may be used to assay for two or three mutations at the same time. The only requirements for the successful employment of the technique are that (1) the precise DNA sequences of the target DNAs must be known and (2) the fluorescent dyes must have well-separated, nonoverlapping fluorescence maxima.

SUMMARY

To be effective, a diagnostic test must be (1) specific for the target molecule, (2) sensitive enough to detect minute levels of the target, and (3) technically simple, with unequivocal results that can be obtained readily. There are two categories of molecular diagnostic techniques. One category relies on the specificity of an antibody for a particular antigen. The other uses nucleic acid hybridization or PCR to detect a specific nucleic acid sequence.

A common assay that uses antibodies is the ELISA. In this procedure, (1) a sample is bound to a solid support, (2) a primary antibody specific for the target antigen is added and binds to the target antigen, (3) a secondary antibody–enzyme conjugate that binds to the primary antibody is added, and (4) a colorless substrate that is transformed into a colored compound by the enzyme in the conjugate is added. The appearance of a color response in an ELISA indicates the presence of the target molecule in the sample.

ELISAs have been used for detecting various proteins, identifying viruses and bacteria, and determining the presence of low-molecular-weight compounds in a wide range of biological samples. To increase the specificity of the primary antibody and to ensure the reliability of the antibody preparation, monoclonal antibodies are often used for diagnostic ELISAs.

Nucleic acid hybridization can be a highly sensitive and specific method for detecting the presence of a nucleic acid sequence in a biological sample. This method has been used to develop diagnostic assays for disease-causing organisms in a clinical setting and other organisms in the environment.

Because a nucleic acid detection assay is directed toward a known DNA sequence, primers for PCR can be synthesized and then used to amplify the target sequence. The detection assay can be run in a nonradioactive system, such as the biotin–streptavidin–chemiluminescence protocol, or the amplified PCR product can be scored by gel electrophoresis. Also, a PCR product can be labeled with a fluorescent dye that is attached to the 5′ end of the primer.

One way to characterize forensic samples is by DNA fingerprinting. In this technique, human minisatellite DNA, which does not encode any proteins and is highly variable in sequence, is usually used as a hybridization probe. The extensive variability of human minisatellite DNA sequences means that each human being produces a unique set of hybridized DNA bands.

To characterize plant DNA, a set of arbitrary oligonucleotide primers can be used to amplify random segments of the plant DNA by PCR and, after electrophoresis, to produce a specific set of DNA bands. This procedure is called the RAPD procedure. Any particular set of primer sequences will produce a unique collection of amplified DNA fragments that is characteristic of the genomic DNA of a particular plant cultivar.

DNA diagnostic assays can also be used to detect the presence of a single-nucleotide change in a particular gene. One of these methods distinguishes between the ligation and the absence of ligation of two oligonucleotides. A single-nucleotide mismatch at the junction of the hybridized oligonucleotides prevents ligation. In general, the use of PCR increases the resolution of nucleic acid diagnostic tests and should also decrease the overall costs of these assays.

The development of molecular diagnostic assays is a growing and dynamic field. Although the technical details of various tests may differ, the general principles have been established. At present, PCR has contributed significantly to overcoming the problem of the limited availability of target DNA. The use of PCR for probe systems has eliminated most concerns about the sensitivity of the detection signal, with the result that nonradioactive chromogenic, chemiluminescent, or fluorescent systems can be used reliably for certain assays. Moreover, in a number of tests, PCR treatment and electrophoretic analysis are sufficient to determine the presence of either a genetic mutation or an infectious agent in the targeted sample. Undoubtedly, many novel DNA-based systems will be created for the diagnosis of most, if not all, of the common genetic, infectious, and malignant diseases.

REFERENCES

Banér, J., M. Nilsson, A. Isaksson, M. Mendel-Hartvig, D.-O. Antson, and U. Landegren. 2001. More keys to padlock probes: mechanisms for high-throughput nucleic acid analysis. *Curr. Opin. Biotechnol.* **12:**11–15.

Barker, R. H., L. Suebsaeng, W. Rooney, G. C. Alecrim, H. V. Dourado, and D. F. Wirth. 1986. Specific DNA probe for the diagnosis of *Plasmodium falciparum* malaria. *Science* **231:**1434–1436.

Bolnick, D. A., D. Fullwiley, T. Duster, R. S. Cooper, J. H. Fujimura, J. Kahn, J. S. Kaufman, J. Marks, A. Morning, A. Nelson, P. Ossorio, J. Reardon, S. M. Reverby, and K. TallBear. 2007. The science and business of genetic ancestry testing. *Science* **318:**399–400.

Boonham, N., J. Tomlinson, and R. Mumford. 2007. Microarrays for rapid identification of plant viruses. *Annu. Rev. Phytopathol.* **45:**307–328.

Bugawan, T. L., R. K. Saiki, C. H. Levenson, R. M. Watson, and H. A. Erlich. 1988. The use of non-radioactive oligonucleotide probes to analyze enzymatically amplified DNA for prenatal diagnosis and forensic HLA typing. *Bio/Technology* **6:**943–947.

Campbell, R. E., O. Tour, A. E. Palmer, P. A. Steinbach, G. S. Baird, D. A. Zacharias, and R. Y. Tsien. 2002. A monomeric red fluorescent protein. *Proc. Natl. Acad. Sci. USA* **12:**7877–7882.

Cao, W. 2004. Recent developments in ligase-mediated amplification and detection. *Trends Biotechnol.* **22:**38–44.

Carlson, D. P., C. Superko, J. Mackey, M. E. Gaskill, and P. Hansen. 1990. Chemiluminescent detection of nucleic acid hybridization. *Focus* **12:**9–12.

Caskey, C. T. 1987. Disease diagnosis by recombinant DNA methods. *Science* **236:**1223–1229.

Chehab, F. F., and Y. W. Kan. 1989. Detection of specific DNA sequences by fluorescence amplification: a color complementation assay. *Proc. Natl. Acad. Sci. USA* **86:**9178–9182.

Claustres, M. 2005. Frequency and nature of mutations and methods to detect them, p. 9–31. *In* G. R. Taylor and I. N. M. Day (ed.), *Guide to Mutation Detection*. John Wiley & Sons, Inc., New York, NY.

Debenham, P. G. 1992. Probing identity: the changing face of DNA fingerprinting. *Trends Biotechnol.* **10:**96–102.

Erlich, H. A., D. Gelfand, and J. J. Sninsky. 1991. Recent advances in the polymerase chain reaction. *Science* **252:**1643–1651.

Eshaque, B., and B. Dixon. 2006. Technology platforms for molecular diagnosis of cystic fibrosis. *Biotechnol. Adv.* **24:**86–93.

Gillam, I. C. 1987. Non-radioactive probes for specific DNA sequences. *Trends Biotechnol.* **5:**332–334.

Godiwala, N. T., A. Vandewalle, H. D. Ward, and B. A. Leav. 2006. Quantification of in vitro and in vivo *Cryptosporidium parvum* infection by using real-time PCR. *Appl. Environ. Microbiol.* **72:**4484–4488.

Hardenbol, P., J. Banér, M. Jain, M. Nilsson, E. A. Namsaraev, G. A. Karlin-Neumann, H. Fakhari-Rad, M. Ronaghi, T. D. Willis, U. Landegren, and R. W. Davis. 2003. Multiplexed genotyping with sequence-tagged molecular inversion probes. *Nat. Biotechnol.* **21:**673–678.

Hsieh, H.-M., H.-L. Chiang, L.-C. Tsai, S.-Y. Lai, N.-E. Huang, A. Linacre, and J. C.-I. Lee. 2001. Cytochrome *b* gene for species identification of conservation animals. *Forensic Sci. Int.* **122:**7–18.

Jeffreys, A. J., A. MacLeod, K. Tamaki, D. L. Neil, and D. G. Monckton. 1991. Minisatellite repeat coding as a digital approach to DNA typing. *Nature* **354:**204–209.

Kingsbury, D. T. 1987. DNA probes in the diagnosis of genetic and infectious diseases. *Trends Biotechnol.* **5:**107–111.

Klevan, L., and G. Gebeyehu. 1990. Biotinylated nucleotides for labeling and detecting DNA. *Methods Enzymol.* **184:**561–577.

Kohler, G., and C. Milstein. 1975. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **256:**495–497.

Kostrikis, L. G., S. Tyagi, M. M. Mhlanga, D. D. Ho, and F. R. Kramer. 1998. Spectral genotyping of human alleles. *Science* **279:**1228–1229.

Kuppuswamy, M. N., J. W. Hoffmann, C. K. Kasper, S. G. Spitzer, S. L. Groce, and S. P. Bajaj. 1991. Single nucleotide primer extension to detect genetic diseases: experimental application to hemophilia B (factor IX) and cystic fibrosis genes. *Proc. Natl. Acad. Sci. USA* **88:**1143–1147.

Nickerson, D. A., R. Kaiser, S. Lappin, J. Stewart, L. Hood, and U. Landegren. 1990. Automated DNA diagnostic using an ELISA-based oligonucleotide ligation assay. *Proc. Natl. Acad. Sci. USA* **87:**8923–8927.

Nivens, D. E., T. E. McKnight, S. A. Moser, S. J. Osbourn, M. L. Simpson, and G. S. Sayler. 2004. Bioluminescent bioreporter integrated circuits: potentially small, rugged and inexpensive whole-cell biosensors for remote environmental monitoring. *J. Appl. Microbiol.* **96:**33–46.

Persing, D. H., T. F. Smith, F. C. Tenover, and T. J. White (ed.). 1993. *Diagnostic Molecular Microbiology: Principles and Applications.* American Society for Microbiology, Washington, DC.

Pollard-Knight, D., A. C. Simmonds, A. P. Schaap, H. Akhavan, and M. A. W. Brady. 1990. Nonradioactive DNA detection on Southern blots by enzymatically triggered chemiluminescence. *Anal. Biochem.* **185:**353–358.

Rafalski, J. A., and S. V. Tingey. 1993. Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. *Trends Genet.* **9:**275–279.

Rajkovic, A., B. El Moualij, M. Uyttendaele, P. Brolet, W. Zorzi, E. Heinen, E. Foubert, and J. Debevere. 2006. Immunoquantitative real-time PCR for detection and quantification of *Staphylococcus aureus* enterotoxin B in foods. *Appl. Environ. Microbiol.* **72:**6593–6599.

Sanseverino, J., R. K. Gupta, A. C. Layton, S. S. Patterson, S. A. Ripp, L. Saidak, M. L. Simpson, T. W. Schultz, and G. S. Sayler. 2005. Use of *Saccharomyces cerevisiae* BLYES expressing bacterial bioluminescence for rapid detection of estrogenic compounds. *Appl. Environ. Microbiol.* **71:**4455–4460.

Sayler, G. S., and A. C. Layton. 1990. Environmental application of nucleic acid hybridization. *Annu. Rev. Microbiol.* **44:**625–648.

Shaner, N. C., R. E. Campbell, P. A. Steinbach, B. N. G. Giepmans, A. E. Palmer, and R. Y. Tsien. 2004. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat. Biotechnol.* **22:**1567–1572.

Tyagi, S., and F. R. Kramer. 1996. Molecular beacons: probes that fluoresce upon hybridization. *Nat. Biotechnol.* **14:**303–308.

Tyagi, S., D. P. Bratu, and F. R. Kramer. 1998. Multicolor molecular beacons for allele discrimination. *Nat. Biotechnol.* **16:**49–53.

Waldmann, T. A. 1991. Monoclonal antibodies in diagnosis and therapy. *Science* **252:**1657–1662.

Weiss, J. B. 1995. DNA probes and PCR for diagnosis of parasitic infections. *Clin. Microbiol. Rev.* **8:**113–130.

Weitz, H. J., J. M. Ritchie, D. A. Bailey, A. M. Horsburgh, K. Killham, and L. A. Glover. 2001. Construction of a modified mini-Tn*5 lux*CDABE transposon for the development of bacterial biosensors for ecotoxicity testing. *FEMS Microbiol. Lett.* **197:**159–165.

White, T. J., R. Madej, and D. H. Persing. 1992. The polymerase chain reaction: clinical applications. *Adv. Clin. Chem.* **29:**161–196.

Wolffs, P. F. G., K. Glencross, R. Thibaudeau, and M. W. Griffiths. 2006. Direct quantitation and detection of salmonellae in biological samples without enrichment using two-step filtration and real-time PCR. *Appl. Environ. Microbiol.* **72:**3896–3900.

Wong, M. L., and J. F. Medrano. 2005. Real-time PCR for mRNA quantitation. *BioTechniques* **39:**75–85.

Yu, K. F., A. Van Deynze, and K. P. Pauls. 1993. Random amplified polymorphic DNA (RAPD) analysis, p. 287–301. *In* B. R. Glick and J. E. Thompson (ed.), *Methods in Plant Molecular Biology and Biotechnology.* CRC Press, Inc., Boca Raton, FL.

REVIEW QUESTIONS

1. Briefly describe how the change in the human β-globin gene that gives rise to sickle-cell anemia can be detected by using PCR.

2. Describe and discuss the PCR/OLA detection protocol.

3. What is an ELISA? How does it work?

4. Describe several types of nonradioactive DNA labels. What are the advantages of nonradioactive detection procedures?

5. You have been given the task of developing a simple, sensitive, and reproducible diagnostic procedure for a double-stranded DNA virus that is devastating a local cattle population. Because effective treatment of this disease depends on early and correct diagnosis, you need to be able to detect the very low levels of this virus that are present in infected animals before the onset of disease symptoms. Briefly explain how you would proceed and why you have chosen a particular course of action.

6. For diagnostic assays, what is meant by sensitivity, specificity, and simplicity?

7. How is Chagas disease currently diagnosed? How might the existing procedures be improved?

8. What is a molecular beacon probe, and how does it work?

9. What is DNA fingerprinting, and how is it used to characterize traces of DNA in forensic samples?

10. What is the RAPD procedure, and how can it be used to characterize plant cultivars?

11. What is a padlock probe, and how is it used?

12. What are monoclonal antibodies? How are they different from polyclonal antibodies?

13. Briefly, explain how the HAT selection for hybridomas works.

14. How can molecular beacon probes be used to (1) detect several genes in the same sample and (2) characterize an individual's genotype for a particular genetic disease, such as sickle-cell anemia?

15. Why is it difficult to screen an individual's chromosomal DNA to assess whether he or she carries a mutation of the CFTR gene that leads to cystic fibrosis?

16. Why is it useful to simultaneously employ several different-color fluorescent proteins?

17. How would you develop microbial biosensors to detect environmental contaminants?

18. What is real-time PCR? What is it used for? How does it work?

19. What is immunoquantitative real-time PCR, and how does it work?

20. What is a padlock probe, and how does it work?

21. What is the TaqMan assay procedure, and how can it be used to assay SNPs?

10

Protein Therapeutics

Prior to the development of recombinant DNA technology, most human protein pharmaceuticals were available in only limited quantities, they were extremely costly to produce, and, in a number of cases, their biological modes of action were not well characterized. When recombinant DNA technology was first developed, it was heralded as a means of producing a whole range of possible human therapeutic agents in sufficient quantities for both efficacy testing and eventual human use. This forecast has turned out to be true. Today, the "genes" (mostly complementary DNAs [cDNAs]) for several thousand different proteins that are potential human therapeutic agents have been cloned. Most of these sequences have been expressed in mammalian as well as bacterial host cells, and currently more than 500 are undergoing clinical testing with human subjects for the treatment of various diseases. More than 250 of these "biotechnology drugs" have been approved for use in the United States or the European Union (Table 10.1). However, it will be several years before many of the other proteins are commercially available, because medical products must first be tested rigorously in animals and then undergo thorough human trials, which can last for several years, before being approved for general use. However, the financial incentive for pharmaceutical companies is considerable. It has been estimated that in 2006 the annual global market for human recombinant protein drugs was about $60 billion. Ten "blockbuster" drugs constitute nearly half of these sales. For example, in 2006, rituximab (Rituxan), a monoclonal antibody used to treat individuals with non-Hodgkin lymphoma, generated nearly $4 billion in sales, while various forms of recombinant human insulin generated around $2.5 billion.

The development of preventive procedures and treatments for human diseases was the outstanding contribution of medicine and science to human well-being in the 20th century. This process, however, is a continuous one. So-called old diseases (e.g., tuberculosis) can reappear if preventive measures are relaxed or if antibiotic-resistant organisms arise. The idea of using antibodies as therapeutic agents has come to fruition in the past several years, and specific antibodies are being tested to attack toxins, bacteria, viruses, and even cancer cells. An antibody may be viewed as a

TABLE 10.1 Examples of recombinant proteins that have been approved for human use in the United States or the European Union

Alglucosidase α	Granulocyte-macrophage colony-stimulating factor	Interferon analogues
Anakinra	Hirudin	Interleukin-2
Antihemophilic factor	Human growth hormone	Interleukin-2 analogues
Darbepoetin α	Human growth hormone analogue	Interleukin-11
Dibotermin	Hyaluronidase	Interleukin-11 analogue
DNase I	Insulin	Keratinocyte growth factor
Drotrecogin α	Insulin analogue	Laronidase
Erythropoietin	Insulin-like growth factor 1	Novel erythropoiesis-stimulating protein
Factor VIIa	Interferon-α2a	Osteogenic protein
Factor VIII	Interferon-α2b	Platelet-derived growth factor
Factor IX	Interferon-αN3	Stem cell factor
Follicle-stimulating hormone	Interferon-β1a	Tissue plasminogen activator
α-Galactosidase	Interferon-β1b	Thyrotropin-α
Galsulfase	Interferon-γ1b	Truncated tissue plasminogen activator
Glucagon	Interferon-N	
β-Glucocerebrosidase analogue		

target-seeking missile or as a magic bullet that either can directly neutralize an offending agent or, if equipped with a warhead or poison arrow, can destroy a specific target cell.

Pharmaceuticals

Isolation of Interferon cDNAs

A number of different strategies have been used to isolate either the genes or cDNAs for human proteins. In some cases, the target protein is isolated and a portion of the amino acid sequence is determined. From this information, a DNA coding sequence is deduced. The appropriate oligonucleotide is synthesized and used as a DNA hybridization probe to isolate the gene or cDNA from either a genomic or a cDNA library. Alternatively, antibodies are raised against the purified protein and used to screen a gene expression library. For human proteins that are synthesized primarily in a single tissue, a cDNA library from the messenger RNA (mRNA) of that tissue is enriched for the target DNA sequence. For example, the major protein synthesized by the islets of Langerhans of the pancreas is insulin; 70% of the mRNA fraction isolated from these cells encodes insulin.

Before the completion of sequencing of the human genome in 2001, it was often necessary to devise innovative approaches to isolate human genes or cDNAs, especially when the proteins encoded were found in very low concentrations or when the site of synthesis was not known. The human interferon (IFN) proteins, which include IFN-α, IFN-β, and IFN-γ, are naturally occurring proteins, each one with somewhat different biological activity. When the IFN cDNAs were initially isolated in the early 1980s, very little was known about the encoded proteins (IFN was originally thought to be a single protein), so a novel scheme had to be devised to overcome the scarcity of both the mRNAs and the proteins. Also, at the time, *Escherichia coli* expression vectors for eukaryotic cDNAs were not readily available, so it was necessary to devise an indirect scheme to iso-

late IFN cDNA. The isolation of IFN cDNAs included the following steps (Fig. 10.1).

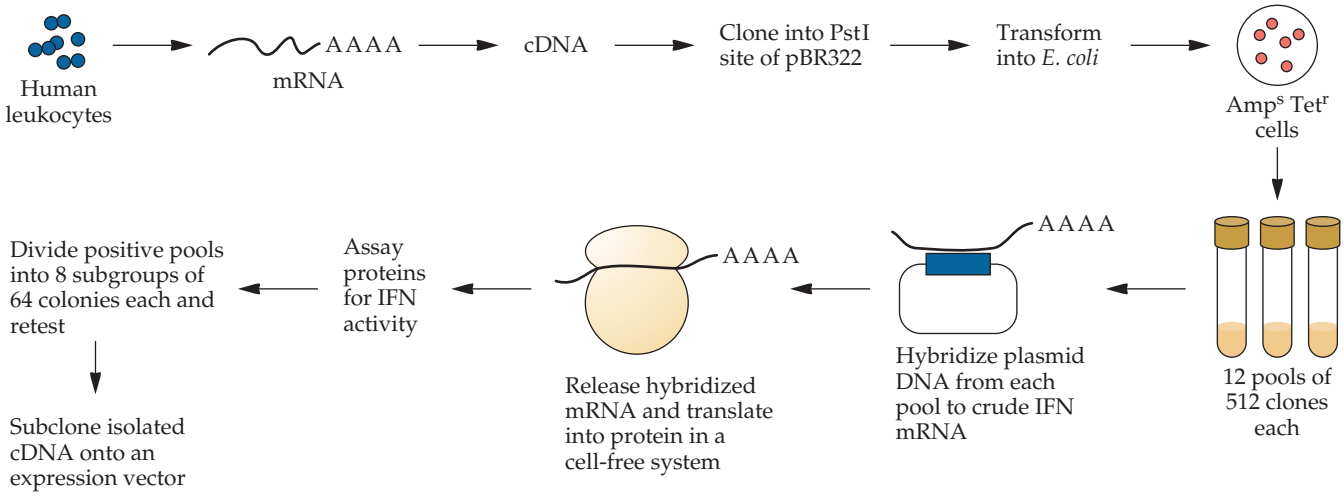
1. Size-fractionated mRNA was isolated from human leukocytes, reverse transcribed, and inserted into the PstI site of plasmid pBR322.
2. The approximately 6,000 clones that were produced following transformation of *E. coli* were divided into 12 pools of 512 clones each. Pools of clones, rather than individual clones, were tested to speed up the identification process.
3. The plasmid DNA from each pool was hybridized to a crude IFN mRNA preparation.
4. The input mRNA that hybridized to the plasmid DNA was separated from the cloned DNA–mRNA hybrids and translated in a cell-free protein synthesis system.
5. Each translation mixture was then assayed for IFN antiviral activity. The pools that showed IFN activity contained at least one clone with a cDNA that hybridized to IFN mRNA.
6. Positive pools were divided into eight subgroups of 64 clones each and retested (i.e., steps 3 to 5 were repeated). This subgrouping process was repeated until a clone with the complete cDNA for a human IFN was identified.

Subsequently, whenever large quantities of the IFN were required, the IFN cDNAs could be subcloned into an *E. coli* expression vector and expressed at high levels.

Human Interferons

After the isolation of the first IFN gene, researchers found that there are a number of different IFNs. On the basis of chemical and biological properties, the IFNs can be classified, as noted above, into three different groups: IFN-α, IFN-β, and IFN-γ. The proteins IFN-α and IFN-β are synthesized in cells that have been exposed to viruses or viral RNA; IFN-γ is synthesized in response to cell growth-stimulating agents. IFN-α is encoded by a family

FIGURE 10.1 Overview of the protocol used to isolate IFN cDNA.



of 13 different (but similar) genes, IFN-β is encoded by two genes, and IFN-γ is encoded by a single gene. The IFN-α subtypes have different specificities. For example, the antiviral activities of IFN-α2 and IFN-α1 are approximately the same when assessed with a virus-challenged bovine cell line, but IFN-α2 is seven times more effective than IFN-α1 when human cells are treated with virus. IFN-α2 is 30 times less effective than IFN-α1 when mouse cells are used in this assay.

Interferon gene shuffling. Several research groups have attempted to engineer IFNs with combined properties based on different members of the IFN-α gene family that vary in the extents and specificities of their antiviral activities. Theoretically, this can be achieved by splicing a portion of one IFN-α gene with a DNA sequence from a different IFN-α gene to create, after translation, a hybrid protein that exhibits novel properties, i.e., properties different from either of the contributing genes.

In one study, hybrid genes from IFN-α2 and IFN-α3 were constructed in an effort to create proteins with novel IFN activities. Comparison of the sequences of the two IFN-α cDNAs indicated that they had common restriction sites at positions 60, 92, and 150. Digestion of both cDNAs at these sites and ligation of the DNA fragments yielded a number of hybrid derivatives of the original genes (Fig. 10.2). These hybrids were expressed in *E. coli,* and the resultant proteins were purified and examined for various biological functions. When tested for the extent of protection of mammalian cells in culture against viral infection, some of the hybrid IFNs were found to have greater activity than the parental molecules. In addition, many of the hybrid IFNs induced test cells to synthesize (2′-5′)-oligoisoadenylate synthetase. This enzyme generates (2′-5′)-linked oligonucleotides, which in turn activate a latent cellular endoribonuclease that cleaves viral mRNA. Other hybrid IFNs had an antiproliferative activity against various human cancers that was greater than that of either of the parental molecules. More recently, additional hybrid IFN molecules have been generated by a variation of the above-mentioned procedure. In this case, the entire IFN-α cDNA family was PCR amplified and then digested with DNase into small DNA fragments (~50 to 60 nucleotides long) before the fragments were shuffled and amplified by PCR (Fig. 10.3). This procedure works because the PCR mixture contains many overlapping single-stranded DNAs that can act as PCR primers (see "Chemical Synthesis of DNA" in chapter 4). Following testing of the many shuffled IFN cDNAs, it is possible to select hybrid IFNs with vastly improved antiviral or antiproliferative activities. In fact, some hybrid IFNs have recently undergone successful clinical trials (Box 10.1) and have been approved for use as human therapeutic agents. The strategy for creating hybrid IFNs can also be applied to other gene families whose products have therapeutic potential.

Longer-acting interferons. Hepatitis C virus infection is one of the most common causes of liver disease, which affects nearly 200 million people worldwide. Many of these individuals eventually develop either cirrhosis of the liver or hepatocellular carcinoma. Therapeutic agents that maximize early antiviral response and maintain viral suppression throughout the course of therapy have the best chance of achieving lasting eradication of the virus from an infected individual. One effective treatment for hepatitis C includes the combined use of the antiviral chemical compound ribavirin with IFN-α. Longer-acting IFNs are needed to minimize the side effects

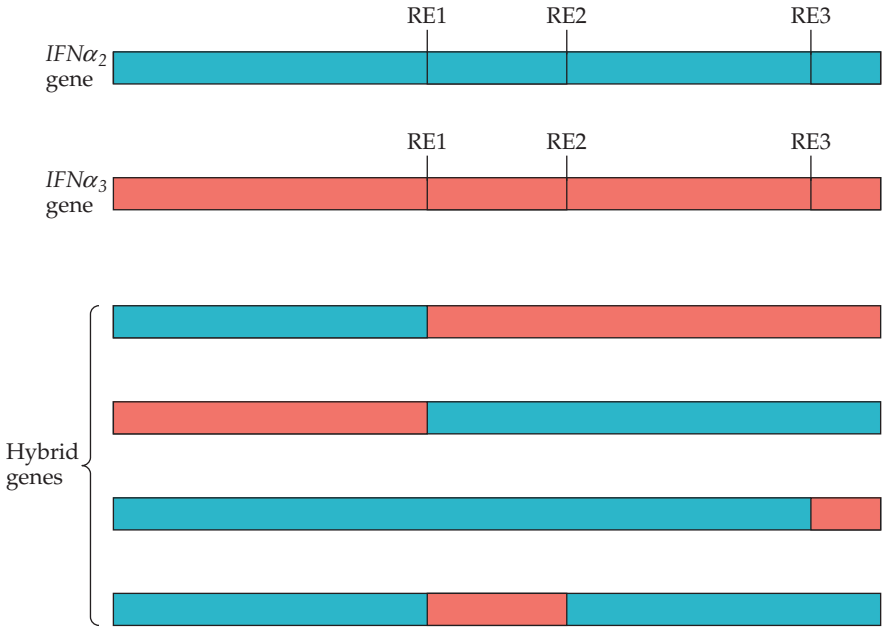


FIGURE 10.2 Structure of the IFN-α2 and IFN-α3 genes and four hybrid genes. Comparison of the sequences of the IFN-α2 and IFN-α3 genes shows shared restriction enzyme sites (RE1, RE2, and RE3). Digestion of the genes at the indicated restriction sites and ligation of the resulting fragments generate a number of different hybrid IFN genes, of which four possibilities are shown.

from IFN treatment, lower the required dosage, and decrease the required frequency of the treatments. One approach to creating long-acting IFNs includes PEGylation. PEGylation entails covalently attaching polyethylene glycol (PEG) to proteins. The binding is typically achieved by incubation of a reactive derivative of PEG with the target protein molecule. PEGylation increases the size of IFN in solution, thereby prolonging its circulatory time by reducing its renal clearance. A simpler means of generating longer-acting IFNs is to fuse an IFN gene with the gene for a stable protein, such as human serum albumin, that, after translation, produces a stable hybrid protein. This combination has been called the albumin–interferon hybrid molecule (Zalbin, formerly Albuferon), and it retains all of the biological activity of the native IFN molecule (Fig. 10.4). Native IFN levels in the blood of a treated patient typically decrease rapidly, so that 2 days after administration, they are undetectable. On the other hand, with the albumin-interferon hybrid molecule, the drug (in this case, the fusion protein) in serum remains at a therapeutically effective level for a much longer time, so that it needs to be administered no more than once every 2 weeks. The initial clinical trials of the albumin-interferon hybrid molecule have all been positive. Phase III clinical trials of the albumin-interferon hybrid molecule began in late 2006. If these trials are successful, then the albumin-interferon hybrid molecule may be available for general use some time in 2010.

Human Growth Hormone

Human growth hormone (somatotropin) is a 191-amino-acid pituitary protein with a molecular mass of 22,125 daltons (Da) that stimulates the

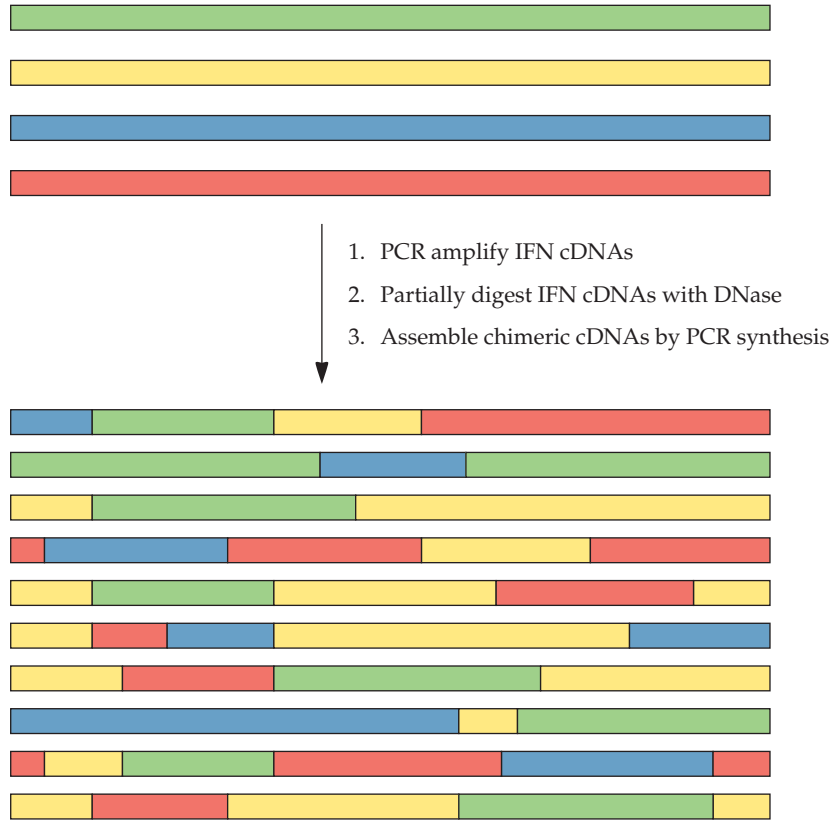


FIGURE 10.3 Construction of hybrid IFN-α genes. The resultant IFN-α gene-shuffled libraries are tested for antiproliferative and antiviral activities.

production of insulin-like growth factor 1. Insulin-like growth factor 1 is an essential component of the promotion of growth in children, and in adults, it controls metabolism. Human growth hormone was one of the first therapeutic proteins in the world to be approved for human use. The recombinant form of the protein is produced in *E. coli* and is identical to native pituitary-derived human growth hormone. Infants and children who lack sufficient endogenous levels of human growth hormone, patients with chronic renal insufficiency (defective kidneys), and individuals with Turner syndrome respond to treatment with growth hormone, which stimulates tissue and bone growth, increases protein synthesis and mineral retention, and decreases body fat storage.

The first recombinant growth hormone was called somatrem (Protropin); it was produced and marketed by Genentech beginning in 1985. It had an amino acid sequence that was identical to that of human growth hormone, except that there was an extra methionine residue at the N-terminal end of the peptide chain (which was thought to prolong its half-life). It was discontinued in the late 1990s.

Treatment of children with human growth hormone typically entails daily injections during the years when the child is growing. The cost of the treatment varies depending on the country and the size of the child but is generally approximately $10,000 to $30,000 per year. In addition, in 2004, the U.S. Food and Drug Administration (FDA) approved the use of recombinant human growth hormone for individuals whose short stature was

BOX 10.1

Clinical Trials

After the discovery of a new drug or course of treatment, and before it is made available to the public, it is essential that extensive studies and analysis of its safety and efficacy be conducted and then reviewed by an impartial agency. Although a large number of countries have developed their own approaches to test new therapeutics, the "gold standard" for clinical trials is the set of requirements established by the FDA. This process is briefly described here.

The preclinical phase of therapeutic drug development (i.e., the initial stage of the process of bringing a new therapeutic agent to market) entails thorough and extensive laboratory research on the mode of action, structure, and other biochemical and physical properties of a potential new drug. Scientists working at universities, research institutes, and drug and biotechnology companies are continually discovering and testing new molecules, as well as new uses for known compounds. However, it is impossible to know with any certainty which avenues of research will eventually bear fruit. Once a promising result has been obtained in the laboratory, and it has been shown to be reproducible, sufficient quantities of a highly purified version of the potential therapeutic compound must be produced so that it can be tested on small animals, such as mice. If the animal tests are positive and there is no evidence of any serious side effects, the organization seeking to commercialize the research files an "investigational new drug" application with the FDA. This is an application to begin the process of clinical trials. Based on the preclinical research data that are provided, about 85% of these applications are approved.

Clinical trials are conducted in three distinct phases (described below), generally requiring a total of about 7 to 9 years at a cost of approximately $75 million to $100 million to complete. At each stage, various compounds are dropped from consideration based on the results obtained. Eventually, approximately 20% of the compounds that looked promising based on preclinical results will, after a careful review of all the data, finally be approved. This slow and expensive process is claimed to be "the most effective method ever devised to assess the efficacy of a treatment."

The three phases of the FDA review process are as follows.

Phase I: With between 10 and 100 healthy people, the safety of the drug and, starting with very low doses, the highest dosages that can be administered are assessed. When there is a chance that serious side effects may result, individuals affected with the disorder that the drug is designed to alleviate may be used.

Phase II: With 50 to 500 affected patients, the optimal dosing regimen is determined. A control group is used so that it is possible to clearly distinguish between the effects of the drug and the natural remission of the disease. The use of a control group also helps to delineate real from apparent side effects of the treatment.

Phase III: Depending upon the disease, approximately 300 to 30,000 patients who have the disease are tested. After it is established that the drug is not harmful and the optimal dosing regimen has been determined, the effectiveness of the treatment needs to be proven.

The requirement for careful and thorough clinical trials ensures both the safety and efficacy of approved drugs. However, since the costs of both the preclinical research and the clinical trials are borne by pharmaceutical companies, this system makes it difficult for small companies that discover a new product to eventually bring that product to market without the involvement of a large corporation with significant financial resources. Furthermore, the high cost of clinical trials and the low probability of a new drug's being approved mean that it is unlikely that therapeutic agents will even be considered for clinical trials unless there is a strong possibility that there will be significant financial gains from the sale of that agent. This financial disincentive may discourage research on therapeutic agents for diseases that either affect only a relatively small number of people or affect only populations in poor, underdeveloped countries.

caused by a variety of medical conditions other than human growth hormone deficiency.

The strategy of designing a protein by either functional domain shuffling or directed mutagenesis can be used to augment or constrain its mode of action. For example, native human growth hormone binds to both growth hormone and prolactin receptors that occur on a number of different cell types. To avoid unwanted side effects during therapy, it is desirable that human growth hormone bind only to growth hormone receptors. Because the segment of the growth hormone molecule that binds to the growth hormone receptor overlaps but is not identical to the portion of the

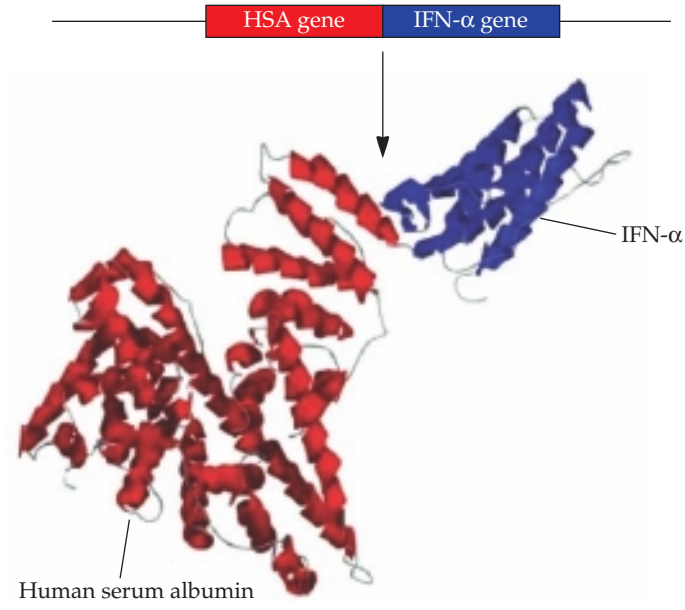
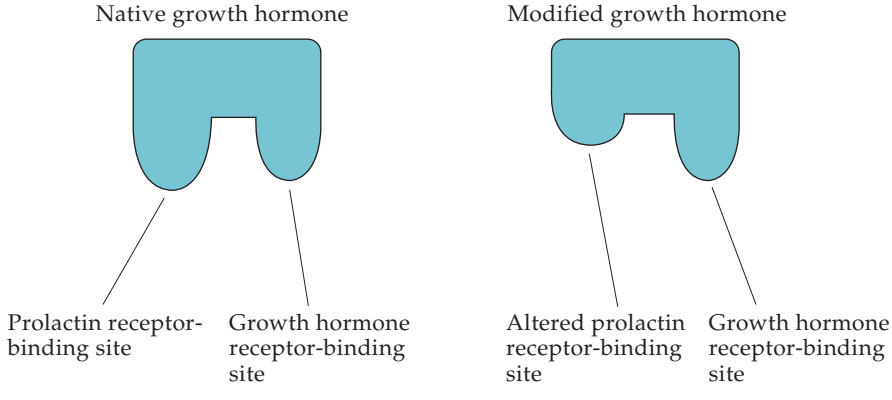


FIGURE 10.4 Schematic representation of the synthesis of the albumin–interferon fusion protein (Zalbin, formerly Albuferon), which includes human serum albumin (HSA) (red) at the N terminus and human IFN-α2b (blue) at the C terminus. Modified from http://www.hgsi.com/albinterferon-alfa-2b.html with permission.

molecule that binds to the prolactin receptor, it should be possible to selectively decrease the binding to the prolactin receptor.

Site-specific mutagenesis of the cloned human growth hormone cDNA was used to change some of the amino acid side chains that act as ligands for Zn^{2+} (i.e., His-18, His-21, and Glu-174), because the ion is required for the high-affinity binding of human growth hormone to the prolactin receptor (Fig. 10.5). As hoped, these modifications yielded human growth hormone derivatives that bound to the growth hormone receptor but not to the prolactin receptor. These derivatives are being tested for safety and efficacy in humans.

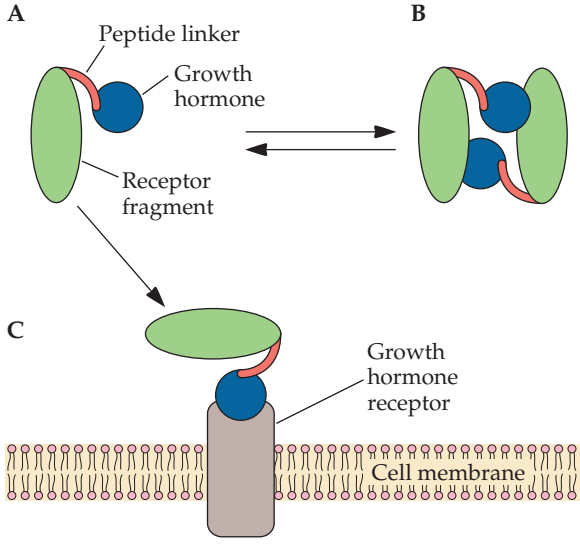
FIGURE 10.5 Schematic representation of native and modified human growth hormone. Oligonucleotide-directed mutagenesis was used to alter human growth hormone so that it no longer bound to the prolactin receptor but retained its specificity for the growth hormone receptor.



As a consequence of its relatively short half-life in plasma, human growth hormone therapy currently requires subcutaneous injection once a day. This treatment is both inconvenient and expensive. Therefore, it would be advantageous to have a long-lasting form of human growth hormone. To this end, the extracellular domain of the human growth hormone receptor was fused to human growth hormone using a 20-amino-acid-long linker peptide consisting of four repeats of the amino acids Gly_4Ser (Fig. 10.6). This construct has a very strong tendency to dimerize as the growth hormone moiety from one molecule binds with the receptor portion of another molecule. When this growth hormone construct was tested in rats, a single injection promoted growth for 10 days (compared to the usual requirement in rats for daily injections). It is thought that the dimerization of the growth hormone construct stabilizes human growth hormone in vivo so that it is cleared from plasma approximately 300 times more slowly than free human growth hormone. Under these conditions, the active monomeric form (Fig.10.6A) is slowly released from the inactive dimeric growth hormone (Fig.10.6B), allowing it to bind to the growth hormone receptor (Fig.10.6C). This experiment is certainly intriguing. It remains to be determined whether humans respond in a similar manner to the dimerized complex.

Another method that has been devised to prolong the active lifetime of human growth hormone includes fusing the coding sequences for the C-terminal end of human growth hormone (~22 kDa) with the N-terminal end of human serum albumin (~67 kDa). This fusion protein is called Albutropin (Fig. 10.7); it has a molecular mass of ~89 kDa and is produced by a strain of yeast that has been genetically modified so that the proteins that it produces have a minimal number of posttranslational modifications. The stabilization of the human growth hormone portion of Albutropin reflects the stability of human serum albumin, which has a half-life in serum of about 19 days. Albutropin has been shown to be effective in both

FIGURE 10.6 Derivatization of growth hormone by coupling it to a portion of the growth hormone receptor using a 20-amino-acid peptide. **(A)** Monomeric derivative; **(B)** dimeric derivative; **(C)** monomeric derivative bound to a growth hormone receptor.



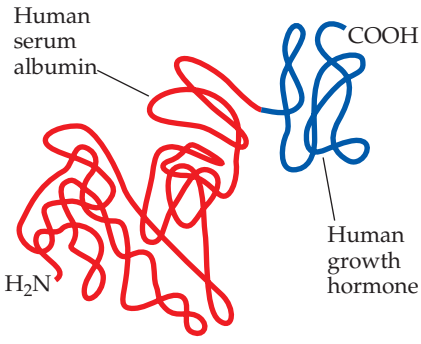


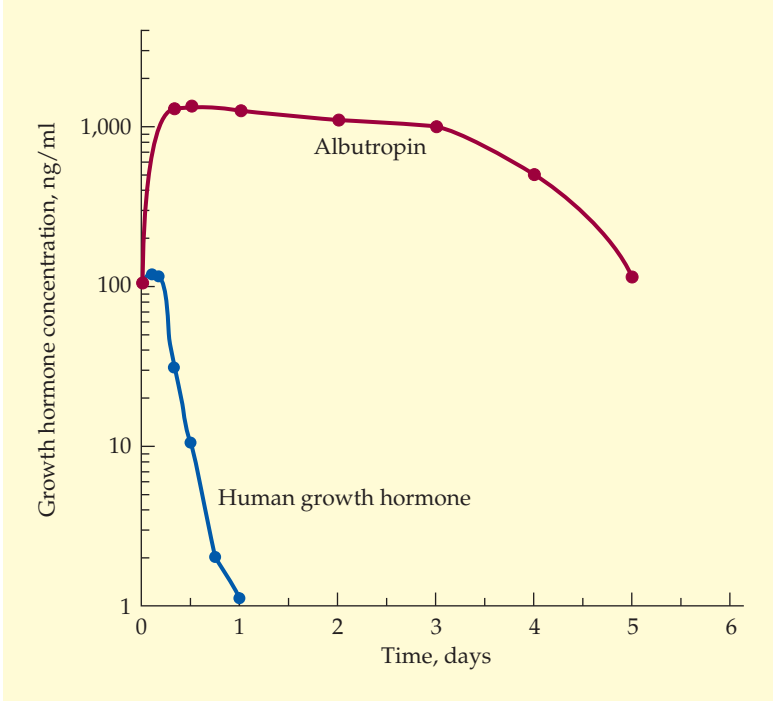
FIGURE 10.7 Schematic representation of the fusion protein Albutropin, which includes human serum albumin (red) at the N terminus and human growth hormone (blue) at the C terminus.

rats and monkeys, in which high levels of the protein in serum were observed 5 days after it was administered (Fig. 10.8). Moreover, Albutropin has successfully completed phase I clinical trials.

Tumor Necrosis Factor Alpha

While a number of studies have clearly shown that tumor necrosis factor alpha (TNF-α) is a potent antitumor agent, it has not been widely used in this capacity because of its severe toxicity. If TNF-α could be delivered directly to its site of action, i.e., the tumor, then lower doses could be used and the unwanted side effects would be diminished. To develop a version of TNF-α with tumor specificity, DNA encoding the peptide Cys-Asn-Gly-Arg-Cys-Gly (which targets a tumor cell surface protein) was fused to TNF-α DNA. The fusion protein contained a 6-amino-acid extension at its N-terminal end (Fig. 10.9). In mice, the cytotoxic activities of Cys-Asn-Gly-Arg-Cys-Gly–TNF-α and TNF-α were identical, indicating that the additional amino acids did not prevent protein folding, combining of three subunits to form a trimer, or binding to receptors. However, the modified version of TNF-α was 12 to 15 times more effective at inhibiting tumor growth than the unmodified form. Moreover, a higher percentage of mice with lymphoma survived after treatment with the modified factor (Fig. 10.10). In addition, all the mice that were treated with the modified factor and survived for 30 days survived a second and third challenge with mouse lymphoma cells. These data indicate that there is a significant benefit, at least in mice, to fusing TNF-α with a short targeting peptide. Nevertheless, this work must be regarded as preliminary until its efficacy is demonstrated in humans.

FIGURE 10.8 Intravenous concentration in monkeys of either human growth hormone or Albutropin following subcutaneous injection.



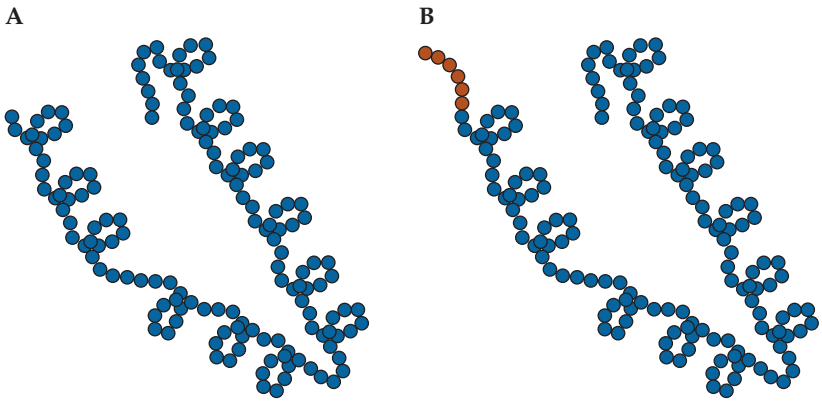


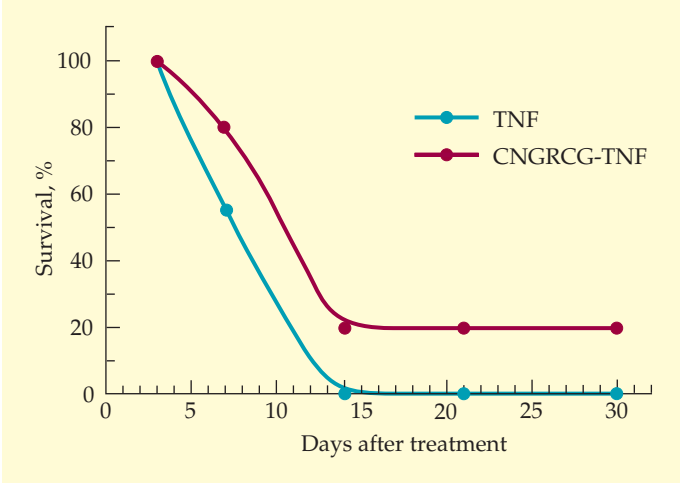
FIGURE 10.9 Schematic representation of TNF-α (blue) without **(A)** and with **(B)** a 6-amino-acid peptide (red) fused to its N terminus. The protein structure shown is hypothetical; only the numbers of amino acid residues (shown as circles) are accurately depicted.

Enzymes

DNase I

Cystic fibrosis is one of the most common fatal hereditary diseases among Europeans and their descendants, with approximately 30,000 diagnosed cases in the United States and another 23,000 cases in Canada and Europe. It is estimated that a mutant cystic fibrosis gene is carried by 1 in 29 Europeans, 1 in 65 African Americans, and 1 in 150 Asians. Individuals with cystic fibrosis are highly susceptible to bacterial infections in their lungs. Antibiotic treatment of patients who have these recurring infections eventually leads to the selection of antibiotic-resistant bacteria. The presence of bacteria, some alive and some lysed, contributes to the accumulation of a thick mucus in the lungs of these patients, making breathing very difficult and acting as a source for further infection. The thick mucus in the

FIGURE 10.10 Survival of lymphoma-bearing mice following treatment with 3 μg of either TNF-α or Cys-Asn-Gly-Arg-Cys-Gly–TNF-α (CNGRCG–TNF) as a function of the number of days after treatment.



lungs is the result of the combination of the alginate that is secreted by the living bacteria, the DNA that is released from lysed bacterial cells, and degenerating leukocytes that accumulate in response to the infection, as well as filamentous actin derived from the cytoskeletons of damaged epithelial cells (Fig. 10.11). To address this problem, scientists at the U.S. biotechnology company Genentech isolated the gene for the human enzyme deoxyribonuclease I (DNase I) and subsequently expressed the gene in Chinese hamster ovary (CHO) cells in culture. DNase I can hydrolyze long polymeric DNA chains into much shorter oligonucleotides. The purified enzyme was delivered in an aerosol mist to the lungs of patients with cystic fibrosis. The DNase I decreased the viscosity and adhesivity of the mucus in the lungs and made it easier for these patients to breathe. While this treatment is not a cure for cystic fibrosis, it nevertheless relieves the most severe symptom of the disease in most patients. The enzyme was approved for use by the FDA in 1994; it had sales of approximately $100 million in 2000.

The monomeric form of actin binds very tightly to DNase I (inhibitor constant $[K_i]$ = ~1 nM) and inhibits its ability to cleave DNA (Fig. 10.12). This interaction limits the effectiveness of DNase I as a therapeutic agent. On the basis of X-ray crystallographic studies, it was possible to predict which amino acid residues of DNase I interacted with actin and were therefore possible targets for change by directed mutagenesis. For example, changing amino acid 144 from alanine to arginine or amino acid 65 from tyrosine to arginine decreased the binding of DNase I to actin up to 10,000-fold. In addition, the actin-resistant mutants had 10- to 50-fold more DNase I activity than the native enzyme. It is not known whether any additional benefit might be realized by combining the amino acid changes from several actin-resistant mutants. The clinical efficacy of a DNase I mutant enzyme that does not bind actin still remains to be demonstrated.

Alginate Lyase

Alginate is a polysaccharide polymer that is produced by a wide range of seaweeds and both soil and marine bacteria. Alginate is composed of chains of the sugars β-D-mannuronate and α-L-guluronate. The properties of a particular alginate depend on the relative amounts and distribution of these two saccharides. For example, stretches of α-L-guluronate residues form both interchain and intrachain cross-links by binding calcium ions, and the β-D-mannuronate residues bind other metal ions. The cross-linked alginate polymer forms an elastic gel. In general, the structure of an alg-

FIGURE 10.11 Schematic representation of a portion of a human lung occluded by a combination of live alginate-secreting bacterial cells, lysed bacterial cells, and leukocytes and their released DNA. This matrix may be digested by alginate lyase or DNase I.

inate polymer is related to its viscosity, which is in turn directly proportional to its molecular size.

The excretion of alginate by mucoid strains of *Pseudomonas aeruginosa* that infect the lungs of patients with cystic fibrosis significantly contributes to the viscosity of the mucus in the airways. Once mucoid strains of *P. aeruginosa* have become established in the lungs of cystic fibrosis patients, it is almost impossible to eliminate them by antibiotic treatment. This is because the bacteria form biofilms in which the alginate prevents the antibiotics from coming into contact with the bacterial cells. In one experiment, it was shown that the addition of alginate lyase, which can liquefy bacterial alginate, together with or prior to antibiotic treatment, significantly decreased the number of bacteria found in biofilms (Fig. 10.13). This result suggests that, in addition to the DNase I treatment, depolymerization of the alginate would help clear blocked airways of individuals with cystic fibrosis.

An alginate lyase gene has been isolated from a *Flavobacterium* species, a gram-negative soil bacterium that is a strong producer of this enzyme. A *Flavobacterium* clone bank was constructed in *E. coli* and screened for alginate lyase-producing clones by plating the entire clone bank onto solid medium containing alginate. Following growth, colonies that produced alginate lyase formed a halo around the colony when calcium was added to the plate (Fig. 10.14). In the presence of calcium, all of the alginate in the medium, except in the immediate vicinity of an alginate lyase-positive clone, becomes cross-linked and opaque. Since hydrolyzed alginate chains do not form cross-links, the medium surrounding an alginate lyase-positive clone is transparent. Analysis of a cloned DNA fragment from one of the positive colonies revealed an open reading frame encoding a polypeptide with a molecular mass of approximately 69,000 Da. Detailed biochemical

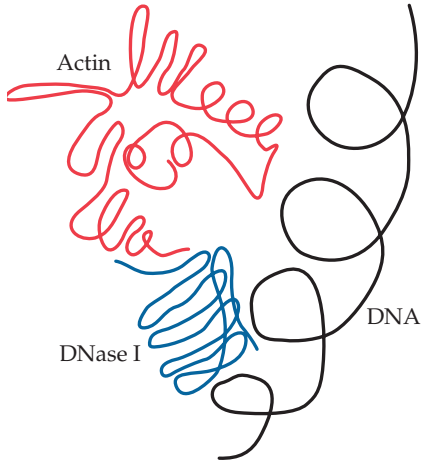
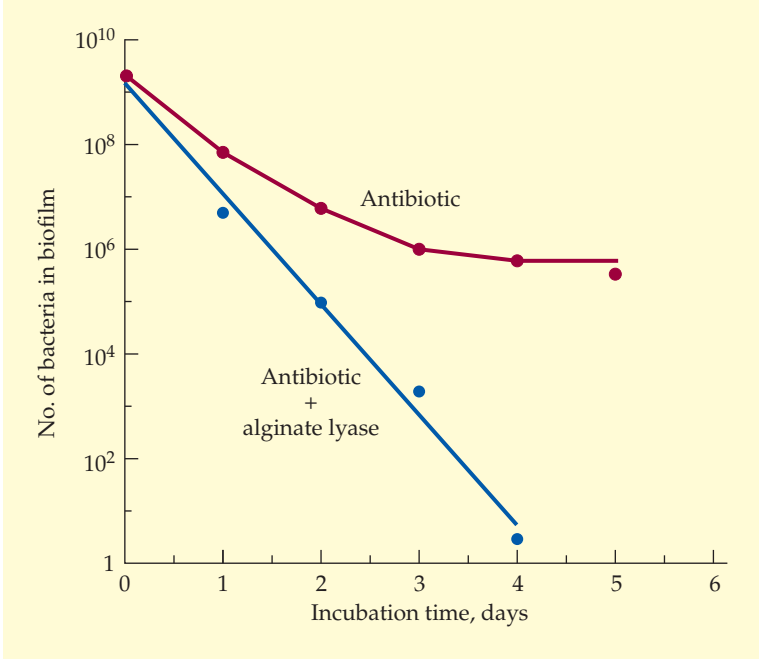


FIGURE 10.12 Schematic representation of the ternary complex of human DNase I, actin, and DNA.

FIGURE 10.13 Time courses of the killing of bacteria in a biofilm with and without the addition of alginate lyase.



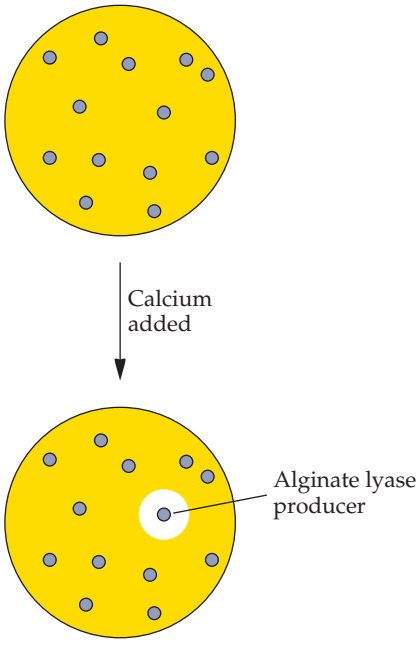


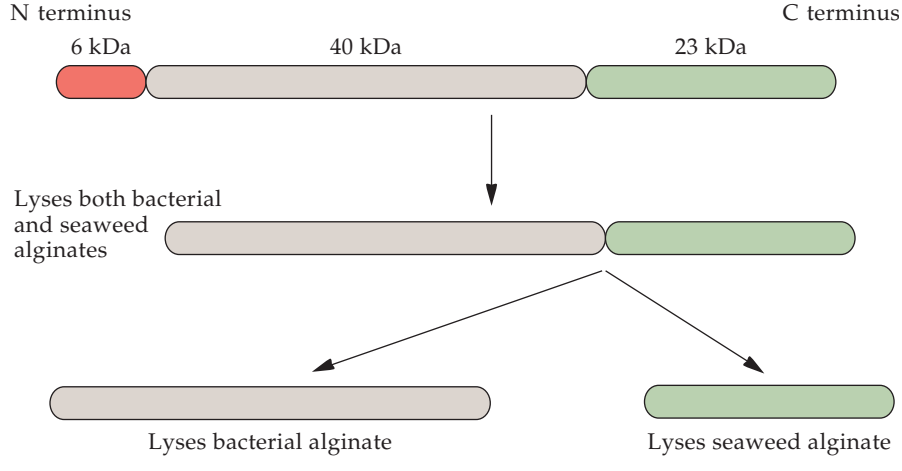
FIGURE 10.14 Schematic representation of the detection of an alginate lyase-producing clone from a clone bank of a *Flavobacterium* sp. in *E. coli*. The alginate that is present in the growth medium is digested by alginase secreted by an *E. coli* clone. The alginate in the vicinity of such a colony is not cross-linked when calcium is added and instead produces a clear zone (halo) surrounding the colony.

and genetic studies indicated that this polypeptide is a precursor of the three different alginate lyases produced by the *Flavobacterium* sp. (Fig. 10.15). After the 69,000-Da precursor is produced, a proteolytic enzyme cleaves off an N-terminal peptide of about 6,000 Da. The 63,000-Da protein can lyse both bacterial and seaweed alginates. Cleavage of the 63,000-Da protein yields a 23,000-Da enzyme that depolymerizes seaweed alginate and a 40,000-Da enzyme that is effective against bacterial alginate. To produce large amounts of the 40,000-Da enzyme, the DNA corresponding to the enzyme was amplified by the polymerase chain reaction (PCR) and then inserted into a *Bacillus subtilis* plasmid vector fused to a *B. subtilis* α-amylase leader peptide to direct the secretion of the protein and placed under the transcriptional control of a penicillinase gene promoter (Fig. 10.16). Transformation of *B. subtilis* cells with this construct yielded colonies with large halos on solid medium containing alginate after calcium was added. When these transformants were grown in liquid medium, the recombinant alginate lyase was secreted into the culture broth. Further tests showed that the enzyme efficiently liquefied alginates that were produced by mucoid strains of *P. aeruginosa* that had been isolated from the lungs of patients with cystic fibrosis. Additional studies are necessary to determine whether recombinant alginate lyase is an effective therapeutic agent.

Phenylalanine Ammonia Lyase

The human genetic disease phenylketonuria results from the impaired functioning of the enzyme phenylalanine hydroxylase. In the United States, approximately 1 of every 12,000 newborns has phenylketonuria. When phenylalanine hydroxylase, which oxidizes phenylalanine to tyrosine, is deficient, the normal cognitive development of an individual is impaired and mental retardation ensues due to a buildup of phenylalanine. Following diagnosis of phenylketonuria, either prenatally or shortly after birth, the

FIGURE 10.15 Processing of the recombinant *Flavobacterium* alginate lyase protein precursor in *E. coli*. A 6-kDa peptide is removed from the N terminus of the 69-kDa precursor to yield a 63-kDa protein that can depolymerize alginate from both seaweed and bacteria. A second cleavage event converts the 63-kDa protein into a 23-kDa protein that is active against seaweed alginate and a 40-kDa protein that hydrolyzes bacterial alginate.



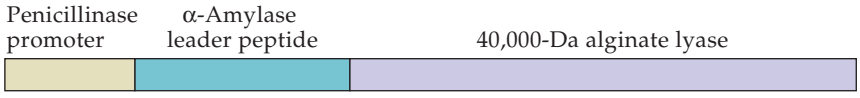


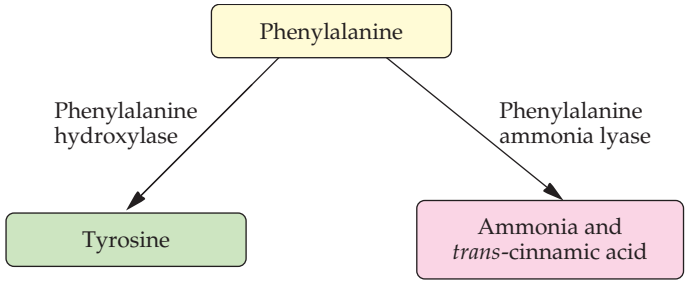
FIGURE 10.16 DNA construct encoding the 40,000-Da alginate lyase. The leader peptide from a *B. subtilis* α-amylase gene is fused to the N terminus of the alginate lyase coding sequence. The construct is under the transcriptional control of a *B. subtilis* penicillinase gene expression system.

treatment entails a controlled semisynthetic diet with low levels of phenylalanine through infancy and possibly for life. A possible alternative treatment would be the administration of the enzyme phenylalanine hydroxylase. Unfortunately, phenylalanine hydroxylase is a multienzyme complex that is not very stable and requires a cofactor for activity. On the other hand, phenylalanine ammonia lyase, which converts phenylalanine to ammonia and *trans*-cinnamic acid (Fig. 10.17), is a stable enzyme that does not require a cofactor and could potentially prevent the accumulation of phenylalanine in phenylketonuria patients. To test this concept, the gene for phenylalanine ammonia lyase from the yeast *Rhodosporidium toruloides* was cloned and overexpressed in *E. coli*. Preclinical studies were conducted with mice that were defective in producing phenylalanine ammonia lyase and therefore accumulated phenylalanine. With these mice, plasma phenylalanine levels were lowered when phenylalanine ammonia lyase was injected intravenously or encapsulated enzyme was administered orally. Thus, at least in mice, phenylalanine ammonia lyase is an effective substitute for phenylalanine hydroxylase, and the orally delivered enzyme is sufficiently stable to survive the mouse gastrointestinal tract and still function. Although this report is preliminary, a combination of oral enzyme therapy with phenylalanine ammonia lyase and a less stringent low-phenylalanine diet might serve to improve the quality of life of individuals affected with this disease.

α_1-Antitrypsin

The processing of a number of different pathogenic bacterial or viral precursor proteins by human proteases occurs when the protease recognizes the amino acid sequence –Arg-X-Lys/Arg-Arg↓–, with peptide bond cleavage on the C-terminal side of the C-terminal Arg (as indicated by the arrow), where X is any of the 20 common amino acids. Since this processing step is common to several infectious agents, a therapeutic agent

FIGURE 10.17 Products of the conversion of phenylalanine by phenylalanine hydroxylase and phenylalanine ammonia lyase.



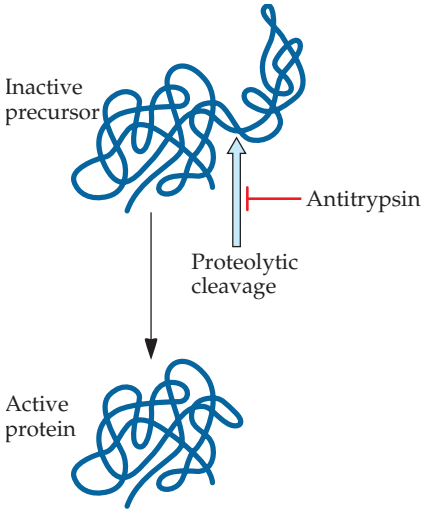


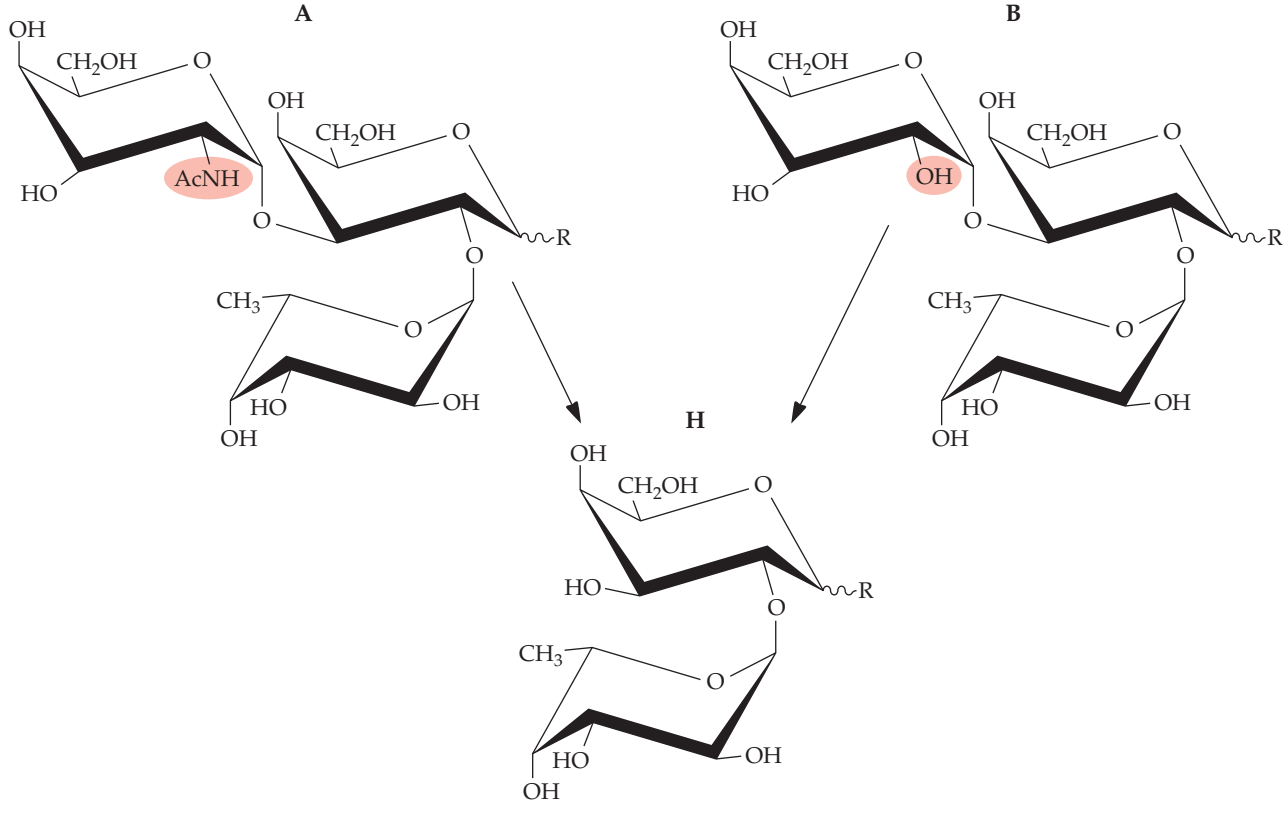
FIGURE 10.18 Schematic representation of α_1-antitrypsin inhibiting the proteolytic cleavage of pathogenic precursor proteins by human proteases.

that targeted the processing enzyme and blocked its activity might act as a broad-spectrum antibacterial and antiviral agent (Fig. 10.18). When a variant of human α_1-antitrypsin was genetically engineered and tested in tissue culture experiments, the protein blocked the processing of human immunodeficiency virus (HIV) type 1 glycoprotein gp160, as well as measles virus protein F_0, and consequently, in both cases, the production of infectious viruses. When the α_1-antitrypsin variant was added to cell cultures, it blocked the production of human cytomegalovirus, a major cause of illness and death in organ transplant recipients and AIDS patients. The α_1-antitrypsin variant is both potent and selective. Against human cytomegalovirus, it is at least 10-fold more effective than any currently used viral inhibitory agent. Its efficacy has been demonstrated in cell culture, but it remains to be determined if the strategy is effective with whole animals.

Glycosidases

The ABO blood group system is based upon the presence or absence of specific carbohydrate residues on the surfaces of erythrocytes, endothelial cells, and some epithelial cells. The monosaccharide that determines blood group A is a terminal α-1,3-linked *N*-acetylgalactosamine, while the corresponding monosaccharide of blood group B is α-1,3-linked galactose (Fig. 10.19). Group O cells lack both of these monosaccharides at the ends of

FIGURE 10.19 Digestion of the monosaccharides that determine blood groups A and B to obtain the H antigen (i.e., blood group O) with specific glucosidases. AcNH stands for an acetyl moiety covalently bound to a nitrogen atom.



their oligosaccharide chains and instead contain α-1,2-linked fucose, which is designated the H antigen. Plasma from blood group A individuals contains antibodies against the B antigen, blood group B individuals have antibodies against the A antigen, and blood group O individuals have antibodies against both the A and B antigens. In practice, this means that individuals with either anti-A or anti-B antibodies cannot safely receive a blood transfusion containing the incompatible antigen, since this is likely to cause a severe immune response (Table 10.2). As a consequence, blood group AB individuals are said to be universal recipients while those from blood group O are universal donors. Thus, when a blood transfusion is required, it is advantageous to have a large supply of plasma that is from blood group O (e.g., in an emergency situation, there may not be sufficient time to check a patient's blood group). Fortunately, digestion of blood cells from either type A or B with specific glycosidases can cause types A, B, and AB to be converted into type O (Fig. 10.19). These enzymes were found following an extensive screening process of 2,500 fungal and bacterial isolates. Eventually, an active α-*N*-acetylgalactosamidase, which converts group A to group O, was found in the gram-negative bacterium *Elizabethkingia meningosepticum* and one with α-galactosidase A, which converts group B to group O, was found in *Bacteroides fragilis* (also a gram-negative bacterium). The genes were isolated, and the proteins were characterized. Both of the enzymes have high specificity for cleaving the appropriate monosaccharide under conditions that maintain the integrity and functioning of the treated red blood cells. Moreover, each enzyme could readily be removed from the treated red blood cells following treatment. While this is a very recent and still preliminary experiment, if this novel approach works effectively in a clinical setting, then it should become a boon for all types of blood transfusions.

Lactic Acid Bacteria

Lactic acid bacteria are widely used in the production and preservation of fermented foods, and many have been given the designation "generally regarded as safe" within the food industry. Many of these organisms are members of the indigenous microflora of the human gut and have been recognized for their health-promoting properties. Some strains of lactic acid bacteria, notably lactobacilli, are used in probiotic products. A probiotic is a live microorganism that is claimed to confer a health benefit by altering the indigenous microflora of the intestinal tract. Lactic acid bacteria have also been used to treat several gastrointestinal disorders, including lactose intolerance, traveler's diarrhea, antibiotic-associated diarrhea, infections caused

TABLE 10.2 Compatible and incompatible blood groups

Donor blood type	Recipient blood type			
	A	B	AB	O
A	Compatible	Incompatible	Compatible	Incompatible
B	Incompatible	Compatible	Compatible	Incompatible
AB	Incompatible	Incompatible	Compatible	Incompatible
O	Compatible	Compatible	Compatible	Compatible

Individuals from one blood group may safely receive a blood transfusion from individuals from a compatible blood group but not from someone from an incompatible blood group.

by various bacterial and viral pathogens, and immunopathological disorders, such as Crohn disease and ulcerative colitis.

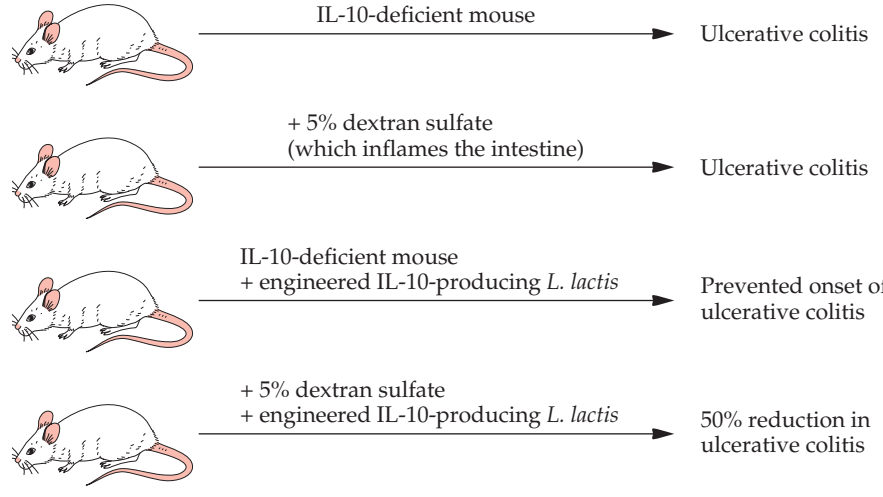
In the past few years, lactic acid bacteria have been used as a host system to express various foreign genes with the idea that these bacteria facilitate the delivery of the proteins encoded by the genes to the human gut. In particular, *Lactococcus lactis* has been developed as a host for this purpose. *L. lactis* is a nonpathogenic, noninvasive, noncolonizing gram-positive bacterium that is often used in the production of fermented foods. Moreover, *L. lactis* has been used for many years as a human probiotic.

Interleukin-10

Ulcerative colitis and Crohn disease, both diseases of the intestinal tract, affect approximately 1 in every 500 to 1,000 people in the developed countries of the world. Ulcerative colitis is associated with excess type 2 T helper cell cytokines, including interleukin-4 and interleukin-5, whereas in Crohn disease, type 1 T helper cell cytokines, including TNF-α, IFN-α, and interleukin-2, are overproduced. The treatment for Crohn disease often includes trying to lower the levels of cytokines, especially TNF-α. One approach has been the administration of antibodies against TNF-α. Other workers have targeted interleukin-10 as a means of controlling Crohn disease because it modulates the regulatory T cells that control inflammatory responses to intestinal antigens. However, interleukin-10 is not clinically acceptable because it needs to be administered by either frequent injections or rectal enemas. To overcome this problem, the bacterium *L. lactis* was engineered to synthesize and secrete interleukin-10.

Experiments were performed with mice to test whether interleukin-10-secreting *L. lactis* could be used to treat inflammatory bowel disease (Fig. 10.20). First, interleukin-10-secreting *L. lactis* was fed to mice with ulcerative colitis that had been induced by 5% dextran sulfate in their drinking water. Second, strains of mice that are genetically incapable of synthesizing interleukin-10 and provide an animal model for ulcerative colitis were tested. In both of these cases, the engineered *L. lactis* significantly alleviated

FIGURE 10.20 Schematic representation of the effects of intestinal interleukin-10 (IL-10)-secreting bacteria on inflammatory bowel disease in mice.



the symptoms of the disease, establishing that this approach works in principle. However, these mouse models for inflammatory bowel disease are not identical to the disease in humans, and a large number of questions remain before the treatment is used with humans.

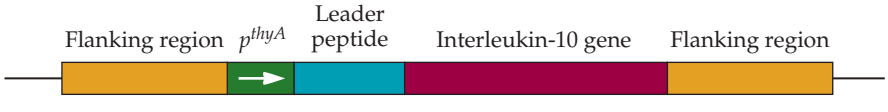
One concern about the use of an interleukin-10-secreting *L. lactis* strain as a therapeutic approach is the possibility that the genetically modified bacterium will be released to the environment. If this were to happen, the plasmid carrying the interleukin-10 gene and any plasmid-borne antibiotic resistance marker genes could be spread to other bacteria in the environment. To prevent this from occurring, a synthetic human interleukin-10 gene that replaced the *L. lactis* thymidylate synthase gene, *thyA,* which is essential for the growth of the bacterium, was inserted into the bacterial chromosome of *L. lactis* by homologous recombination (Fig. 10.21). This strain produced interleukin-10 and grew well in the laboratory when either thymidine or thymine was added to the medium. However, when it was deprived of thymidine and thymine, the viability of the bacterium declined by several orders of magnitude. When this modified bacterium was tested in pigs, whose digestive tract is similar to that of humans, it thrived and actively produced interleukin-10. In addition, laboratory experiments demonstrated that the modified *L. lactis* was extremely unlikely to acquire a thymidylate synthase gene from other bacteria in the environment, confirming both the safety and efficacy of this approach.

Recently, clinical trials with this *L. lactis* strain were initiated. To date, 10 patients with Crohn disease have been treated. So far, a significant decrease in disease activity has been observed, with only minor adverse events. Moreover, bacteria isolated from the patients' feces were not able to grow without the addition of thymidine. In other words, the engineered *L. lactis* did not acquire a thymidylate synthase gene, indicating that the containment strategy was effective. Thus, initial indications are that this strategy appears to be working as well in humans as it did with small animals.

Leptin

It has been estimated that approximately 30% of the North American and 20% of the European populations are overweight. Moreover, North Americans annually spend tens of billions of dollars on various weight reduction schemes, most of which are unsuccessful. However, real weight reduction may be obtained by administration of the protein leptin. Leptin, the product of the *obese (ob)* gene, is a 167-amino-acid protein with a molecular mass of approximately 16 kDa. Leptin is synthesized as a precursor with a 21-amino-acid-long signal peptide that is removed when leptin is secreted. Treatment with recombinant leptin can reduce food intake and

FIGURE 10.21 The genetic construct integrated into the chromosomal DNA of *L. lactis* in place of its thymidylate synthase gene. The promoter (p^{thyA}) is from the thymidylate synthase gene. The interleukin-10 gene was chemically synthesized so that its codon usage was optimized for *L. lactis,* thereby ensuring a high level of protein expression.



correct metabolic perturbations in (homozygous) leptin-deficient mice. Leptin also helps to overcome human congenital leptin deficiency. However, when it is introduced subcutaneously, leptin is not particularly effective in obese patients unless their serum leptin concentrations reach levels 20- to 30-fold higher than normal. This response has been attributed to the inefficient transport of leptin across the blood-brain barrier. To overcome this problem, a scheme for the intranasal delivery of leptin has been devised.

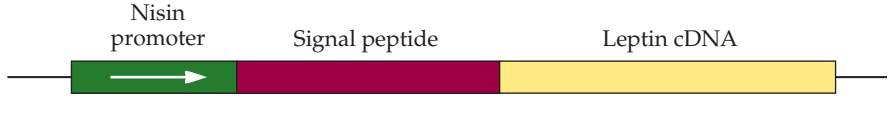
When leptin is produced in *E. coli*, it typically forms insoluble inclusion bodies that must be solubilized and renatured before the active protein is generated. This is a time-consuming, inefficient, and expensive process. In one study, the 462-bp cDNA for human leptin without its signal peptide was cloned and expressed under the control of the nisin promoter in *L. lactis* (Fig. 10.22). Nisin is a 34-amino-acid-residue polycyclic peptide that has antibacterial activity and is used as a food preservative. In *L. lactis*, leptin was produced efficiently without the formation of an inclusion body and was secreted from the recombinant bacteria. Intranasal administration of the leptin-producing *L. lactis* strain significantly reduced food intake and body weight in obese mice. This approach opens up the possibility that, if delivered properly, leptin might act as an effective weight loss treatment in humans.

An HIV Inhibitor

Worldwide, the predominant mode of HIV transmission is by heterosexual contact. One possible way to protect women, who currently comprise about half of all new cases of HIV/AIDS, against HIV infection is a topical microbicide, delivered by a live vaginal *Lactobacillus* strain, that prevents HIV infection directly at mucosal surfaces. This strategy seems reasonable because naturally occurring vaginal *Lactobacillus* strains play a protective role in preventing urogenital infections.

The compound cyanovirin N, isolated from the cyanobacterium *Nostoc ellipsosporum*, blocks several steps of HIV infection, preventing virus entry into human cells. Consequently, cyanovirin N is a candidate for a topical microbiocide to prevent HIV infections. To ensure that cyanovirin N would be expressed at a sufficiently high level in a vaginal strain of *Lactobacillus jensenii*, the gene was chemically synthesized to reflect the codon usage found in the bacterium. Typically, the GC content of lactobacilli is about 36%. In addition, during the synthesis of the gene, proline 51 was replaced by a glycine residue to stabilize the cyanovirin N, and four amino acids were added to the N terminus to ensure proper cleavage of the signal sequence (Fig. 10.23). The modified cyanovirin N gene was fused to a strong and constitutive *Lactobacillus* promoter. The final construct was integrated into the chromosomal DNA of a strain of *L. jensenii* and, when it was tested for efficacy, was found to be highly effective at preventing HIV infec-

FIGURE 10.22 Genetic construct used to secrete leptin from *L. lactis*.



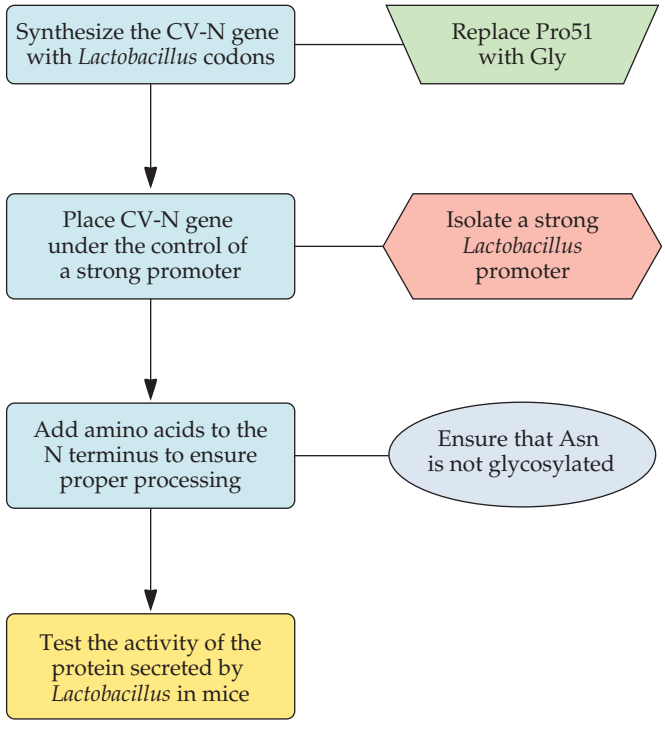


FIGURE 10.23 Flowchart of the scheme used to develop a *Lactobacillus* strain that produces and secretes cyanovirin N (CV-N).

tions in mice. Under these conditions, about 4 μg of cyanovirin N per mL was released into the culture medium.

Monoclonal Antibodies

About 100 years ago, horses were inoculated with the bacterium *Corynebacterium diphtheriae,* which causes diphtheria in humans. The resulting crude horse antiserum was used to treat this often fatal childhood disease. In those days, mortality sometimes reached 45%. *C. diphtheriae* infects the throat or tonsils and produces an exotoxin that is lethal to human cells. This exotoxin enters the bloodstream and damages organs that are distant from the primary site of infection. The administration of horse antiserum containing antibodies against the exotoxin provided passive immunity, protecting the patient from a fatal outcome when the antiserum was given within the first few days after the onset of infection.

Unfortunately, this kind of antibody therapy carries considerable risk and is not widely used today. Patients often develop antibodies against the foreign proteins of either whole or partially purified horse antiserum. After a second treatment, the sensitized patient may go into anaphylactic shock and die. As a result, the use of antibodies as therapeutic agents was considered too dangerous for patients and was used only rarely.

However, with the development of hybridoma technology, antibodies are once again seen as potential therapeutic agents. One reason for the renewed interest in therapeutic antibodies is that it is now possible to engi-

TABLE 10.3 Some therapeutic monoclonal antibodies that have been approved for human use in either the United States or the European Union

Approval date	Antibody	Drug name	Antibody type	Therapeutic use
1986	Muromomab	Orthoclone	Murine	Prevention of acute kidney transplant rejection
1994	Abciximab	ReoPro	Chimeric	Prevention of blood clots
1997	Daclizumab	Zenapax	Humanized	Prevention of acute kidney transplant rejection
1998	Rituximab	Rituxan	Chimeric	Treatment of non-Hodgkin lymphoma
1998	Infliximab	Remicade	Chimeric	Treatment of Crohn disease, psoriasis, rheumatoid arthritis
1998	Basiliximab	Simulect	Chimeric	Preventation of transplantation rejection
1998	Palivizumab	Synagis	Humanized	Treatment of viral infections in children
1998	Trastuzumab	Herceptin	Humanized	Treatment of metastatic breast cancer
2000	Gemtuzumab	Mylotarg	Humanized	Treatment of acute myeloid leukemia
2001	Alemtuzumab	Leukosite	Humanized	Treatment of chronic lymphocytic leukemia
2002	Adalimumab	Humira	Human	Treatment of rheumatoid arthritis
2002	Ibritumomab	Zevalin	Chimeric	Treatment of non-Hodgkin lymphoma
2003	Efalzumab	Raptiva	Humanized	Treatment of severe plaque psoriasis
2003	Omalizumab	Xolair	Humanized	Treatment of severe persistent asthma
2003	Tositumomab	Bexxar	Murine + iodine-131	Treatment of non-Hodgkin lymphoma
2004	Cetuximab	Erbitux	Chimeric	Treatment of various cancers
2004	Natalizumab	Tysabri	Humanized	Treatment of multiple sclerosis
2004	Bevacizumab	Avastin	Humanized	Treatment of various cancers
2006	Panitumumab	Vectibix	Human	Treatment of colorectal cancer
2009?	Denosumab		Human	Treatment of osteoporosis

In addition to the monoclonal antibodies listed here, a number of monoclonal antibodies have been approved for diagnostic and imaging purposes. Phase III clinical trials of denosumab were successfully completed in 2008, and in early 2009 the manufacturer applied for FDA approval of denosumab.

neer antibodies with a greatly reduced level of immunogenicity in humans. In addition, this technique can be used to maintain a continuous supply of pure monospecific antibody. However, the problems of cross-reactivity leading to an immune response and anaphylaxis have not been completely overcome. Thus, the recipient might still produce antibodies to a monoclonal antibody that carries mouse (murine) determinants. To avoid this problem, human monoclonal antibodies with both specific immunotherapeutic properties and lowered potential for immunogenicity have been produced. In fact, a number of monoclonal antibodies have been approved for treating human diseases (Table 10.3, Box 10.2, and Box 10.3).

Structure and Function of Antibodies

An antibody molecule (immunoglobulin) consists of two identical light (L) protein chains and two identical heavy (H) protein chains held together by both hydrogen bonding and precisely localized disulfide linkages. The N-terminal regions of the L and H chains together form the antigen recognition site of each antibody. Antibody genes can be readily manipulated because the various functions of an antibody molecule are confined to discrete domains (regions) (Fig. 10.24). The sites that recognize and bind antigens consist of three complementarity-determining regions (CDRs) that lie

BOX 10.2

Trastuzumab: the First Humanized Monoclonal Antibody Approved for the Treatment of Breast Cancer

In 25 to 30% of women with aggressive metastatic breast cancer, there is a genetic alteration in the *HER2* gene that results in the production of an increased amount of human epidermal growth factor receptor 2 (HER2) protein on the surface of the tumor. Overexpression of the HER2 protein can readily be determined by using an immunohistochemistry-based assay. Some years ago, researchers at Genentech isolated a mouse monoclonal antibody with high affinity for the HER2 protein and then (using procedures similar to those described in this chapter) humanized it. The humanized anti-HER2 monoclonal antibody, trastuzumab (Herceptin), contains human FRs and mouse CDRs and is produced commercially using mammalian (CHO) cells grown in suspension culture as the host for the expression of the antibody. Antibodies produced in CHO cells are glycosylated similarly to bona fide human antibodies. After humanization, trastuzumab bound to the HER2 protein with a dissociation constant of approximately 5×10^{-9} M, indicating that the high level of specificity for the substrate had been maintained through the process of humanization.

In the laboratory, and then in initial clinical trials with more than 800 patients, trastuzumab mediated antibody-dependent cellular cytotoxicity (i.e., it told the immune system to target the cancerous cells) and inhibited the proliferation of human tumor cells that overexpressed HER2 (i.e., it stopped the cancerous cells from growing). Trastuzumab was most effective when it was administered together with some of the chemicals that are currently used for the treatment (chemotherapy) of breast cancer, provided that the breast cancer was at a later stage of development. In two large clinical trials that included over 3,700 women, those who received trastuzumab and chemotherapy had a 52% higher chance that the cancer would not return than those who were treated with chemotherapy alone. Trastuzumab is provided by the manufacturer as a sterile white to pale yellow powder containing 440 mg per vial, and after reconstitution, it is typically administered intravenously over a period of 30 minutes and is taken weekly for 52 weeks. Since a small number of individuals treated with trastuzumab develop heart problems, it is necessary to carefully monitor the cardiac functions of all patients on this therapy, especially older patients and those with a family history of heart problems. In the relatively short time that it has been available, trastuzumab has become a blockbuster drug, with annual sales above $1 billion. In 2006, in the United States, trastuzumab treatment for one individual cost approximately $40,000 for the year.

within the variable (V_H and V_L) regions at the N-terminal ends of the two H and two L chains. The CDRs are the part of an antibody molecule with the greatest variability in amino acid sequence. In addition to the variable regions, each L chain contains one constant region, or domain (C_L), and each H chain has three constant regions, or domains (C_{H1}, C_{H2}, and C_{H3}). When antibodies are digested with the proteolytic enzyme papain, three fragments are released: two identical (Fab) fragments, each of which contains an intact L chain linked by a disulfide bond to the C_{H1} region of the H chain, and one Fc fragment, which consists of two H chain fragments, each containing the C_{H2} and C_{H3} domains and joined by a disulfide bond (Fig. 10.24). The Fab fragment retains the antigen-binding activity. In fact, the N-terminal half of the Fab fragment, which is called the Fv fragment, contains all of the antigen-binding activity of the intact antibody molecule (Fig. 10.24). The amino acid sequence of this portion of the antibody varies considerably from one molecule to another. Each of the constant and variable regions consists of approximately 110 amino acid residues. A complete antibody molecule has a molecular mass of approximately 150 kDa, a Fab fragment is around 50 kDa, and an Fv fragment is about 25 kDa.

In an intact antibody molecule, the Fc portion elicits several immunological responses after antigen–antibody binding occurs.

- The complement cascade is activated. The components of this system break down cell membranes, activate phagocytes, and gen-

erate signals to mobilize other components of the immunological response system.

- Antibody-dependent cell-mediated cytotoxicity (ADCC), which is the result of the binding of the Fc portion of the antibody to an Fc receptor of an ADCC effector cell, is produced. The bound effector cell releases substances that lyse the foreign cell to which the Fab portion of the antibody molecule is bound.
- After the Fab region binds to a soluble antigen, the Fc portion of an antibody can be bound to Fc receptors of phagocytic cells, which engulf and destroy the antibody–antigen complex.

Preventing Rejection of Transplanted Organs

In the 1970s, passive immunization was reconsidered as a way of preventing immunological rejection of a transplanted organ. The rationale was to administer to patients a specific antibody that would bind to certain lymphocytes and diminish the immune response directed against the transplanted organ. The mouse monoclonal antibody OKT3 was approved

BOX 10.3

Rituximab and Ibritumomab: Therapeutic Monoclonal Antibodies That Treat Non-Hodgkin Lymphoma

Non-Hodgkin lymphoma is a malignant growth of B or T cells of the lymph system. It has been estimated by the American Cancer Society that in 2007 alone approximately 63,000 new cases of non-Hodgkin lymphoma were diagnosed, resulting in approximately 19,000 deaths. In fact, about 5 million people worldwide have non-Hodgkin lymphoma, 5 to 10% of these people die every year, and the incidence of the disease is growing. It is the fifth most common cancer (although there are about 29 different lymphomas in this category), with an individual's chance of developing the disease in their lifetime being about 1 in 50.

There are a variety of treatments for patients with non-Hodgkin lymphoma, including radiation therapy, chemotherapy, immunotherapy, bone marrow transplantation, and "watch and wait" for slowly growing cases. In 1997, the FDA approved the use of rituximab (Rituxan) for the treatment of non-Hodgkin lymphoma. Rituximab is a genetically engineered chimeric (murine/human) monoclonal antibody directed against the CD20 antigen (a protein on the surfaces of B lymphocytes). Following binding of the antibody to CD20, the body's defenses attack and kill the antibody-marked B cells. Stem cells in bone marrow lack CD20, so they are uninhibited by this treatment. Healthy B cells can regenerate from those stem cells, after the completion of the course of rituximab treatment (given once a week for 4 to 8 weeks), and return to normal levels within several months. In 2006, the FDA approved the use of rituximab in combination with CHOP (cyclophosphamide, doxorubicin, vincristine, and prednisone) and other anthracycline-based chemotherapy regimens. In addition, the use of rituximab in combination with the chemical compound methotrexate was approved for the treatment of moderately to severely active rheumatoid arthritis in patients who had been refractory to other treatments.

Notwithstanding some severe side effects in some patients, rituximab has been enormously successful. Hundreds of thousands of people worldwide who did not respond well to conventional chemotherapy have been successfully treated with rituximab. In fact, while the incidence of non-Hodgkin lymphoma continues to increase, since the introduction of rituximab, mortality from the disease in the United States has declined at a rate of approximately 2.3% a year. In 2002, the FDA approved the use of ibritumomab tiuxetan (Zevalin) together with rituximab. Ibritumomab is also a monoclonal antibody that targets B cells. However, ibritumomab is linked to a chemical chelator molecule (tiuxetan) that binds tightly to radioactive indium-111 or yttrium-90. Thus, a therapeutic regimen with ibritumomab tiuxetan targets tumor cells with a high dose of radiation. In late 2007, treatment with ibritumomab tiuxetan was priced at approximately $24,000 per month, with treatments typically lasting 1 or 2 months. Treatment with ibritumomab tiuxetan is quite toxic, and around half of the treated individuals experience side effects. Therefore, ibritumomab tiuxetan is approved only for patients who have failed to respond to other treatments.

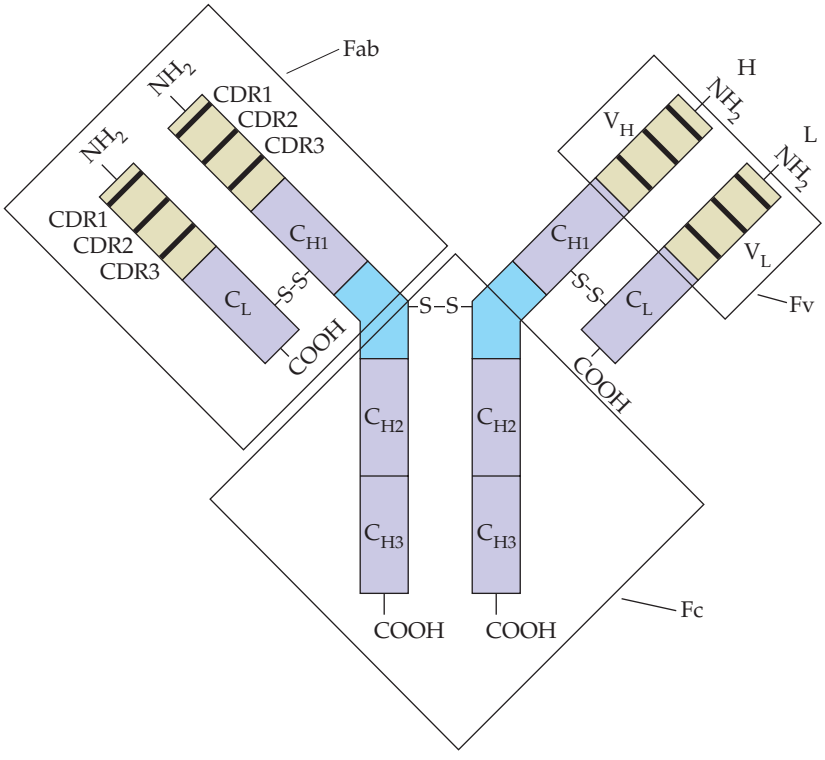


FIGURE 10.24 Structure of an antibody molecule. The H and L chains contain variable regions (V_L and V_H) with their CDRs (CDR1, CDR2, and CDR3) and constant domains (C_L, C_{H1}, C_{H2}, and C_{H3}). The Fv, Fab, and Fc portions of an antibody molecule are delineated. The N-terminal (NH_2) and C-terminal (COOH) ends of each polypeptide chain are indicated.

in 1986 by the FDA for use as an immunosuppressive agent after organ transplantation in humans (Table 10.3). Lymphocytes that differentiate in the thymus are called T cells. Various members of the T-cell population act as immunological helper and effector cells and are responsible for organ rejection. The OKT3 monoclonal antibody binds to a cell surface receptor called CD3, which is present on all T cells. As a result, a full immunological response is blocked, and the transplanted organ is not rejected. Immunosuppression by this means was reasonably effective, although as anticipated, because the antibody was from a mouse, there were some side effects, including fever and rash formation.

Recombinant Antibodies

Hybrid Human–Mouse Monoclonal Antibodies

The modular nature of antibody functions has made it possible to convert a mouse monoclonal antibody into one that has some human segments but still retains its original antigen-binding specificity. This hybrid molecule is called a chimeric antibody (Fig. 10.25), or, with more human sequences, a "humanized" antibody (Fig. 10.26). The difference between a chimeric and a humanized mouse monoclonal antibody depends on which portions of the mouse antibody are removed. The first portion of a mouse monoclonal antibody that was targeted for replacement with a human sequence was

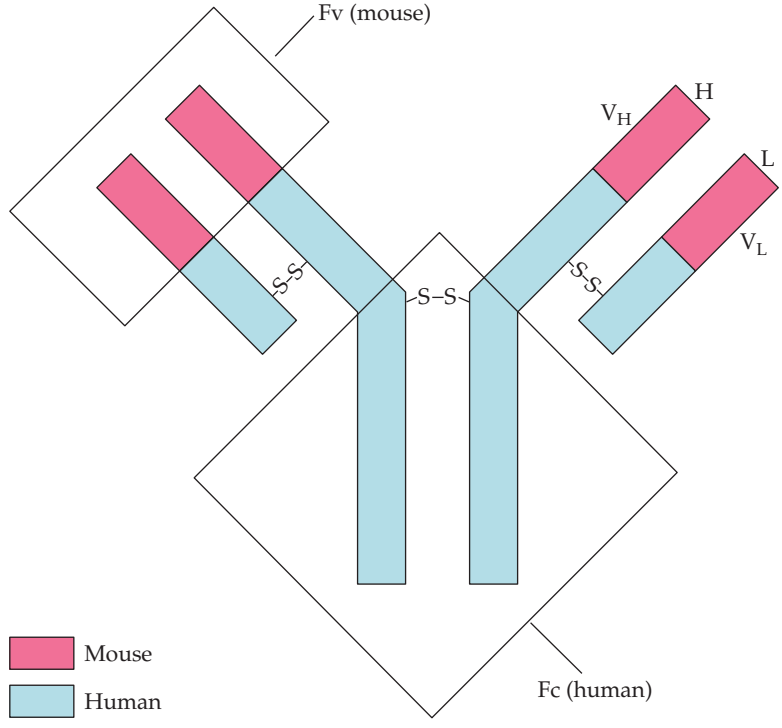


FIGURE 10.25 Genetically engineered chimeric antibody. The V_L and V_H DNA regions from the immunoglobulin L and H genes that encode part of a mouse monoclonal antibody were substituted for the V_L and V_H DNA regions of a human immunoglobulin molecule. The product of the constructed gene is a chimeric (partially humanized) immunoglobulin with the antigen-binding specificity of the mouse monoclonal antibody and both lowered immunogenicity in humans and human Fc effector capabilities.

the mouse Fc fragment. The mouse Fc fragment was chosen because it functions poorly as an effector of immunological responses in humans. It is also the most likely fragment to elicit the production of human antibodies. To diminish immunogenicity and to introduce human Fc effector capabilities, the DNA coding sequences for the Fv regions of both the L and the H chains of a human immunoglobulin were substituted for the Fv DNA sequences for the L and H chains from a specific mouse monoclonal antibody (Fig. 10.25). This replacement of Fv coding regions can be accomplished by using either oligonucleotides with in vitro DNA replication or cloned DNA segments. The DNA constructs for both chimeric chains were cloned into an expression vector and transfected into cultured B lymphocytes, from which the chimeric antibody was collected. Chimeric antibodies are composed of approximately 70% human and 30% mouse DNA sequences.

When a chimeric antibody that contained the binding site from a mouse monoclonal antibody directed against the surfaces of human colon cancer cells was tested in patients with colorectal cancer, it remained in the blood system about six times longer than the complete mouse monoclonal antibody, thereby extending the period of effectiveness. Only 1 patient of the 10 developed a mild immunological reaction to the chimeric antibody. However, in this trial, no antitumor effects were observed, an outcome that

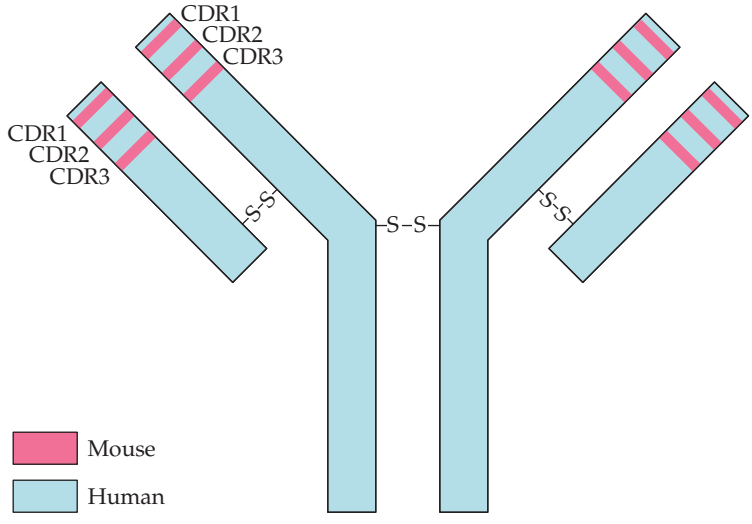


FIGURE 10.26 Genetically engineered humanized antibody. The CDRs (CDR1, CDR2, and CDR3) from the genes for H and L immunoglobulin chains of a mouse monoclonal antibody replace the CDRs of the genes for a human antibody. The product of this constructed gene is an immunoglobulin with the antigen-binding specificity of the mouse monoclonal antibody and all the other properties of a human antibody molecule.

may be due to the low dose levels or to the advanced stage of the cancer in the subjects.

The "humanizing" of mouse and rat monoclonal antibodies has been taken one step further than the formation of chimeric molecules by substituting into human antibodies only the CDRs of the rodent monoclonal antibodies (Fig. 10.26). Humanized antibodies consist of approximately 95% human and 5% mouse DNA sequences. Because these engineered human antibodies have antigen-binding affinities similar to those of the original rodent monoclonal antibodies, they may be more effective therapeutic agents.

The humanizing of rodent monoclonal antibodies may be performed as follows. Starting with a rodent hybridoma cell line, cDNAs for the L and H chains are isolated. The variable regions of these cDNAs are amplified by PCR. The oligonucleotide primers that are used for this amplification are complementary to the sequences at the 5′ and 3′ ends of the DNA encoding the variable regions. From the nucleotide sequences of the cDNAs for the L and H regions (V_L and V_H), it is possible to delineate the limits of the CDRs. It is usually straightforward to determine where the CDRs begin and end, because these regions are highly variable in sequence while the sequences of the framework regions (FRs) are relatively conserved. On the basis of the sequences of the DNAs encoding the rodent CDRs, six pairs of oligonucleotide PCR primers are synthesized. Each pair of primers is designed to initiate the synthesis of the DNA for one of the six rodent CDRs—three from the L chain and three from the H chain. In addition, each primer includes an extra 12 nucleotides at its 5′ end, complementary to the flanking regions within the human framework DNA into which the DNA for the rodent CDRs is targeted (Fig. 10.27). Oligonucleotide-directed mutagenesis is then used to replace, one at a time, the complete DNA sequence for each

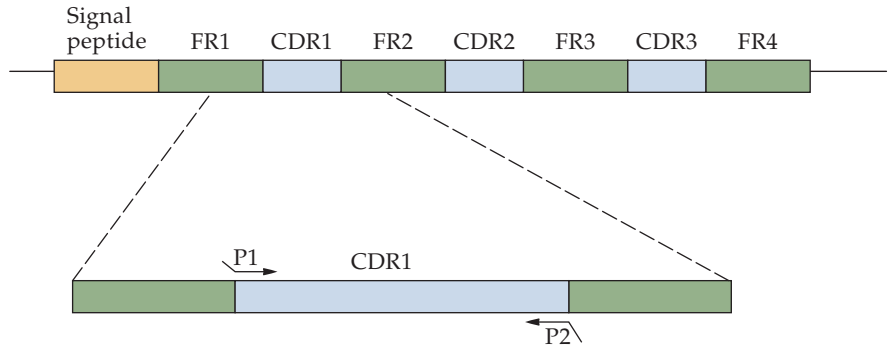


FIGURE 10.27 PCR amplification of CDR1 from a rodent monoclonal antibody L chain cDNA. The PCR primers P1 and P2 contain oligonucleotides complementary to the rodent CDR1 DNA. In addition, P1 and P2 each contain 12 nucleotides at their 5′ ends that are complementary to the FRs of human monoclonal L chain cDNAs. Using six separate pairs of oligonucleotide primers—three for the V_L region and three for the V_H region—each of the rodent CDRs is separately amplified by PCR. Then, by PCR, the amplified rodent CDRs are spliced into human antibody genes in place of the resident CDRs. This grafting is made possible by the presence of DNA complementary to the human FRs on the amplified rodent CDR DNAs.

of the human CDRs with the amplified DNA for the rodent CDRs. Thus, it is necessary to carry out six cycles of oligonucleotide-directed mutagenesis, one cycle to replace each CDR. This procedure, in effect, "grafts" the rodent CDRs onto the human antibody framework. The humanized variable-region cDNAs are then cloned into expression vectors, which are then introduced into appropriate host cells, usually either *E. coli* or mammalian cells, for the production of antibodies.

To date, more than 50 different monoclonal antibodies have been humanized. While this technology is clearly effective and widely applicable, it is nevertheless time-consuming and expensive. Probably (as described below) in the future other strategies will be used to produce human antibodies and antibody fragments, such as (1) phage display combinatorial libraries that are constructed from mRNA from human B cells from nonimmunized donors and (2) transgenic mice that express the entire human antibody repertoire.

Human Monoclonal Antibodies

Although most of the immunotherapeutic agents that have been developed have been effective, there are drawbacks to the use of monoclonal antibodies that contain nonhuman sequences. For example, if multiple treatments are required, which is often the case, it is desirable that the antibody contain no or only a very limited amount of nonhuman sequences to prevent immunological cross-reactivity and sensitization of the patient. Unfortunately, it is very difficult to create human monoclonal antibodies for a number of reasons. The human chromosomes of fused human lymphocyte–mouse myeloma cells during hybridoma formation are unstable, and cells that produce a human monoclonal antibody are rarely formed. To date, no human myeloma cell line has been discovered that can replace the mouse myeloma cell line in this procedure. Even if it were possible to form human hybridoma cell lines, it is contrary to accepted norms of medical research to inject humans with a specific antigen for nontherapeutic pur-

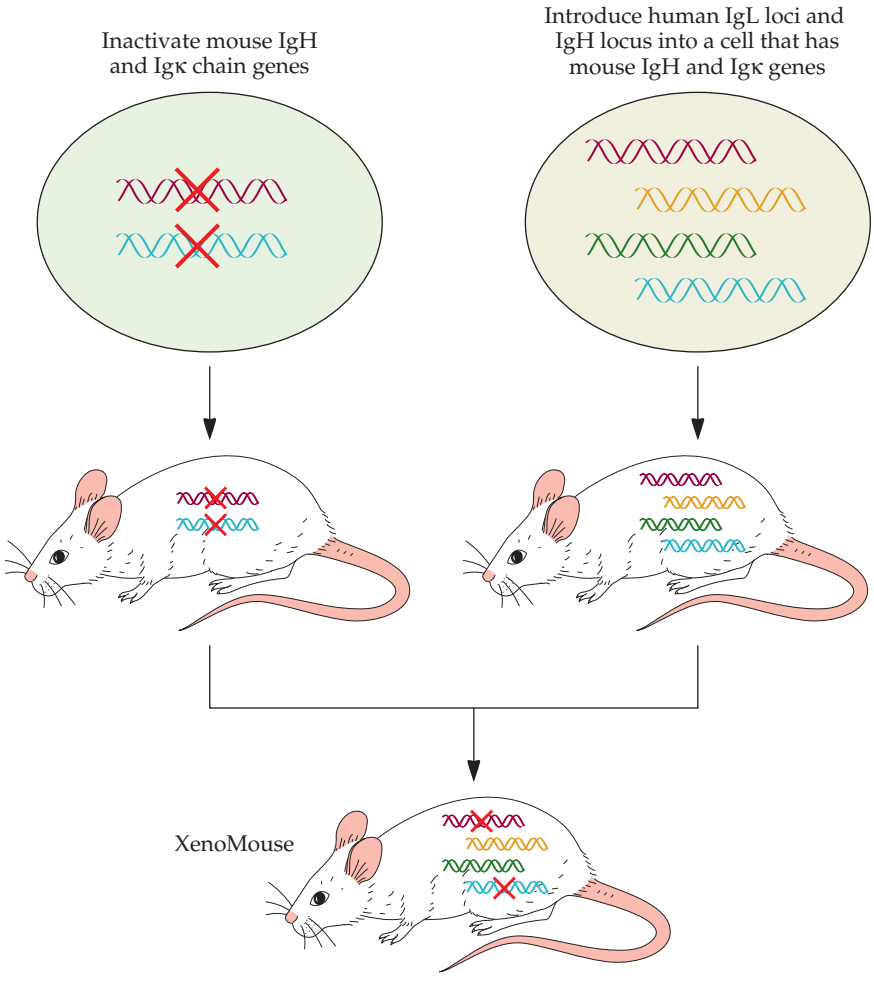


FIGURE 10.28 Generation of a XenoMouse. Mouse antibody genes are inactivated by specific deletions in embryonic stem cells, which are subsequently used to generate transgenic mice unable to make antibodies. The human genes encoding immunoglobulin light and heavy chains are introduced on a YAC into mouse embryonic stem cells. These cells are used to generate transgenic mice able to synthesize both mouse and human antibodies. The mice generated from these two types of manipulation are cross-bred, and mice that can synthesize only human immunoglobulins are selected, immunized, and used to make hybridomas producing human antibodies.

poses and to perform a partial splenectomy to collect antibody-producing cells. Therefore, it has been necessary to devise other approaches for obtaining human monoclonal antibodies.

To address this need, researchers constructed a "XenoMouse" in which (1) the mouse antibody production machinery was inactivated and (2) all of the human immunoglobulin loci (both light and heavy chains) are integrated into a mouse chromosome (Fig. 10.28). The human heavy chain genes and the human κ and λ light chain genes (where κ and λ are different classes of light chain genes) were cloned onto yeast artificial chromosomes (YACs) that can carry very large amounts of foreign DNA. The YACs with the human immunoglobulin genes were then introduced into mouse embryonic stem cells by fusing YAC-containing yeast spheroplasts with the

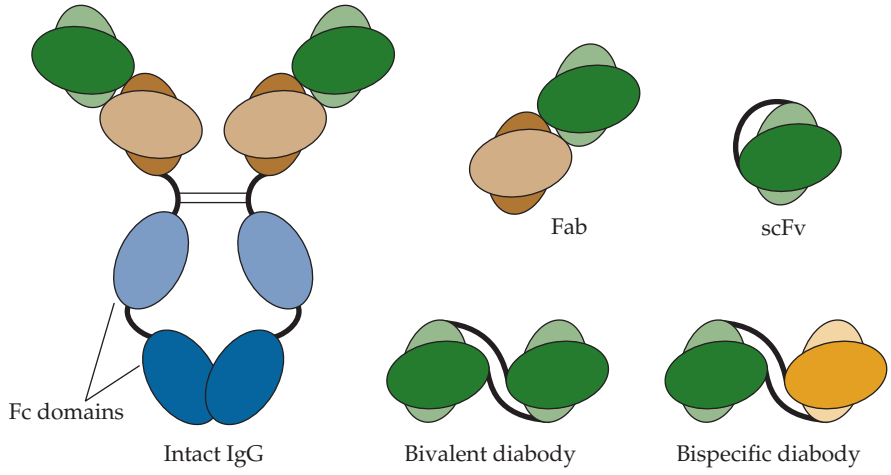


FIGURE 10.29 Schematic representation of active antibodies and antibody fragments.

embryonic stem cells. This procedure yields a large number of embryonic stem cells in which all of the introduced human immunoglobulin genes have become stably integrated into the chromosomal DNA. These transfected cells were used to generate mice containing human immunoglobulin gene loci. Cross-breeding of two mouse lines, one carrying both mouse and human immunoglobulin genes and the other carrying the deleted mouse immunoglobulin genes, produced a mouse strain (called XenoMouse) that expresses only human immunoglobulins. It is now possible, after immunization of a XenoMouse with a particular antigen, to produce a fully human immunoglobulin. The large human antibody repertoire in the XenoMouse has enabled researchers to produce a number of fully human antibodies, many of which are currently at various stages of clinical development. For example, the first fully human monoclonal antibody produced using this technology is panitumumab, which was approved in September 2006 for the treatment of certain forms of colorectal cancer (Table 10.3).

Antibody Fragments

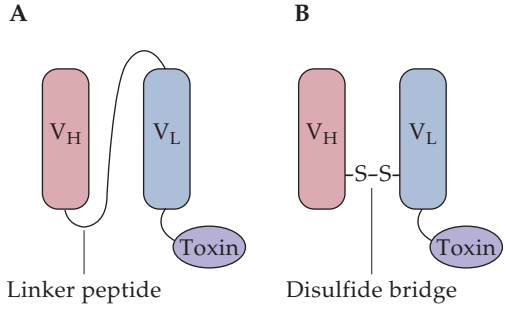
Naturally occurring antibodies are highly specific targeting reagents that provide animals with a powerful means of defending themselves against a wide range of pathogenic organisms and toxins. Immunoglobulin G (IgG) is the main antibody found in mammalian serum, and it is the native form that is almost exclusively used in therapeutic antibodies (Fig. 10.29). The fact that IgG molecules have two identical sites that bind to two identical antigens (i.e., they are bivalent) generally increases their effectiveness in vivo. While the Fc portion of the IgG molecule is important in recruiting cytotoxic effector functions through complement or interaction with specific receptors, Fc-mediated effects are not necessary for all applications and may even sometimes be undesirable. Over the past several years, by manipulating portions of the IgG light and heavy chain cDNAs, researchers have constructed a variety of IgG derivatives or fragments that may be used instead of whole antibody molecules (Fig. 10.29). Some of these molecules, because of their small size, bind more efficiently to targets that are inaccessible to conventional whole antibodies. Others have multiple sites

for binding to the same antigen, while others have binding specificities for two or more target antigens.

Initially, based on Fab and Fv fragments of antibodies, antigen-binding single protein chains (scFv) consisting of only V_L and V_H domains were developed. Single-chain antibodies may be used for a variety of therapeutic and diagnostic applications in which Fc effector functions are not required and when small size is an advantage. Single-chain antibodies have a molecular mass of approximately 27 kDa compared with approximately 150 kDa for IgG molecules. Because of their small size, single-chain antibodies can penetrate and distribute in large tumors more readily than intact antibodies. In addition, a protein-coding sequence can be linked to a single-chain antibody sequence to create a dual-function molecule that can both bind to a specific target and deliver a toxin or some other specific activity to a cell (Fig. 10.30).

Computer simulations of the three-dimensional structure of a potential single-chain antibody showed that the V_L and V_H domains have to be separated by a linker peptide to assume the correct conformation for antigen binding. On the basis of this design constraint, DNA constructs of V_L and V_H sequences from a cDNA of a cloned monoclonal antibody were each ligated to a chemically synthesized DNA linker fragment in the order V_L–linker–V_H. After expression in *E. coli*, the single-chain protein was purified, and both its affinity and specificity were found to be equivalent to those of the original intact monoclonal antibody. Moreover, instead of linking the V_H and V_L chains with a short peptide, amino acids in the FR can be modified to form a disulfide linkage between the two peptides (Fig. 10.30B). The effectiveness of this disulfide-stabilized Fv molecule (V_L-S-S-V_H) coupled to a cancer cell toxin was compared with that of an scFv molecule coupled to the same toxin. The disulfide-stabilized and scFv immunotoxins had the same activity and specificity. However, the former molecule was severalfold more stable than the latter. This suggests that disulfide-stabilized Fv molecules may be more useful than scFv molecules in some therapeutic applications. The two types of molecules have been used in different ways. For example, by altering the number of amino acids in the linker in an scFv molecule (usually to five or fewer amino acids) it is possible to direct the self-assembly of these molecules into either bivalent dimers, called diabodies (Fig. 10.29), trimers (triabodies), or tetramers (tetrabodies). Shortening the linker affects not only the multimerization, but also the stability of the molecule, with molecules with a shorter linker

FIGURE 10.30 Schematic representation of an scFv immunotoxin **(A)** and a disulfide-stabilized Fv immunotoxin **(B).**



MILESTONE

Synthesis in *E. coli* of a Polypeptide with Human Leukocyte Interferon Activity

S. NAGATA, H. TAIRA, A. HALL, L. JOHNSRUD, M. STREULI, J. ECSÖDI, W. BOLL, K. CANTELL, and C. WEISSMANN

Nature **284:**316–320, 1980

The late 1970s and early 1980s were a time of tremendous excitement in molecular biotechnology. The promise of this new technology was being touted to both the public and large institutional investors. One of the products of the new biotechnology that captured the imagination of a large number of people was IFN, which at the time was seen by many as a possible miracle cure for a wide range of diseases caused by viruses and cancer. Thus, the isolation of a human IFN cDNA and its subsequent expression in *E. coli* were reported in newspapers and magazines around the world.

Several features of IFN made it particularly difficult to synthesize and isolate a cDNA encoding the polypeptide. First, although IFN had been purified more than 80,000-fold, only minuscule amounts were available, so researchers did not even have an accurate estimate of its molecular mass. Second, unlike many other proteins, IFN did not have a chemical or biological activity that was easy to monitor. At the time, its activity was measured by the reduction in the cytopathic effect of an animal virus on cells in culture, which was an extremely complex and time-consuming process. Third, unlike insulin, researchers had no idea if there was one particular human cell that produced high levels of IFN and therefore could serve as a source of mRNA that was enriched for IFN mRNA. These problems notwithstanding, a cDNA encoding IFN was eventually isolated and characterized. Since that time, researchers have discovered that there are several different types of IFNs. Unfortunately, IFN is not the panacea that was dreamed of by both investors and the press. However, the genes for several IFNs have been isolated, and clinical trials have shown that they are effective treatments for a variety of viral diseases.

tending to be more stable. In addition, it is possible to combine two different antigen specificities into a single bispecific diabody (Fig. 10.29).

Most recombinant antibody–toxin combinations (immunotoxins) have been constructed using *Pseudomonas* exotoxin A, which is a 66-kDa protein with three separate domains. Domain I is responsible for cell binding, domain II for translocation of the protein into the cell, and domain III for ADP-ribosylation (Fig. 10.31A). Other protein toxins that have been used include bacterially derived diphtheria toxin and the plant-derived toxin ricin. An immunotoxin is generally synthesized by replacing the N-terminal domain of the toxin, e.g., *Pseudomonas* exotoxin A (domain I), with the single-chain antibody sequence, thereby creating molecules very similar in size to the original toxin with the ability to bind, enter, and kill a specific cell (Fig. 10.31B). A number of immunotoxins that have antitumor activity in vitro and in animal models have been constructed. These include antibodies directed against the p55 subunit of the interleukin-2 receptor, the transferrin receptor, carbohydrate antigens, the epidermal growth factor receptor, and some cancer cell surface proteins. Toxin molecules may also be directed to cancer cells by using a dispecific diabody that is engineered to bind to a surface-specific tumor-associated antigen and then to a toxin molecule, thereby directing the toxin molecule to the tumor (Fig. 10.32). A number of different engineered immunotoxins are currently in clinical trials.

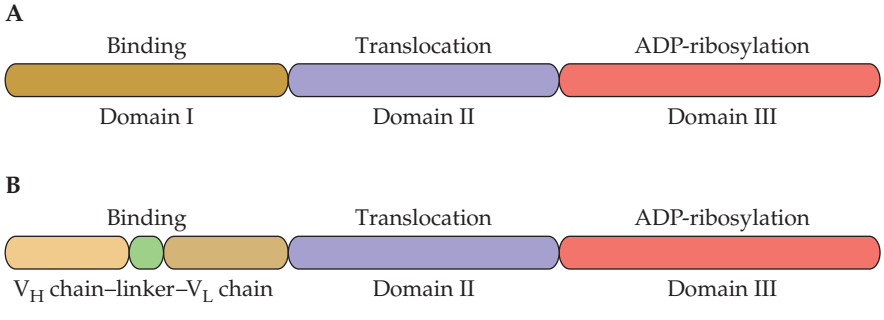
It may be possible to create peptides that are smaller than scFvs and still retain the ability to bind to a specific antigen. The rationale for developing smaller antibody–toxin complexes is that they are more likely to penetrate a tumor and more completely stop tumor growth. Recently, a group of researchers constructed a short peptide (28 amino acids long) that retained the binding specificity of the monoclonal antibody from which it

was derived. It is well established that antibody-binding specificity resides within the six hypervariable loops called CDRs, three from the variable region of the heavy chain and three from the variable region of the light chain (Fig. 10.24 and 10.33). In all antibody molecules, the CDRs are flanked by FRs. Moreover, it was speculated, at least for some antibodies, that the major portion of the antigen-binding site might reside primarily within two CDRs, one from the heavy chain and the other from the light chain. To test this possibility, starting with genes for the variable portion of a monoclonal antibody against a surface protein from Epstein-Barr virus, eight different peptide combinations were synthesized. Each peptide contained at least one CDR3 loop (known to be the major antigen-contacting segment), as well as one other CDR loop and a linker peptide (usually an FR spacer). All eight of these peptides were tested in vitro for the ability to compete with the parental antibody for binding to the Epstein-Barr virus (thought to be the causative agent of Burkitt lymphoma and other cancers) surface protein. One of the peptide combinations, V_HCDR1–V_HFR2–V_LCDR3, appeared to be promising (Fig. 10.33). Next, this peptide was coupled to a toxin molecule, colicin Ia, and the combination was tested both with cells in culture and with mice. In mice, the peptide–colicin adduct efficiently traveled through the circulation and then found and killed the tumor cells expressing the target antigen. Colicin Ia by itself does not affect these tumors to any significant extent. Also, the original monoclonal antibody is unable to penetrate into the tumor. On the other hand, the peptide–colicin adduct accumulated at the cores of the targeted tumors. This very exciting work is at an early stage of development, so a large number of issues remain to be addressed before it can become an effective human therapeutic measure. Nevertheless, the demonstration that a small peptide can mimic the binding specificity of an entire antibody molecule and successfully deliver a cellular toxin to targeted cells may provide the basis for a whole new approach for treating tumors.

Combinatorial Libraries of Antibody Fragments

Hybridoma cells, like most other animal cells in culture, grow relatively slowly, do not attain high cell densities, and require complex and expensive growth media. The cost of monoclonal antibody production is an impediment to their more widespread use as therapeutic agents. To circumvent this problem, attempts have been made to genetically engineer bacteria, plants, and animals to act as "bioreactors" for the production of

FIGURE 10.31 Domain structure of *Pseudomonas* exotoxin A **(A)** and a single-chain antibody–*Pseudomonas* exotoxin A **(B).** The functions of the various domains are shown.



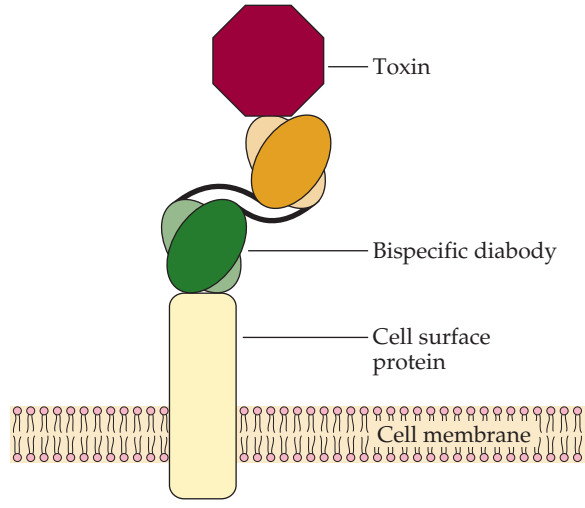


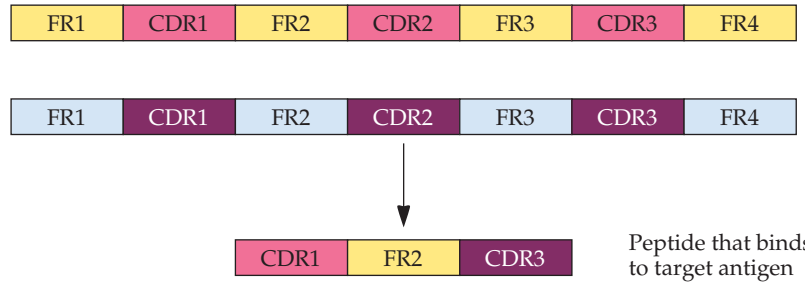
FIGURE 10.32 Schematic representation of the binding of a diabody to a protein molecule on the surface of a cancerous cell, as well as the binding of a toxin protein molecule to the other portion of the diabody.

monoclonal antibodies. For effective delivery and function of some immunotherapeutic agents, only the antigen-binding region of an antibody (the Fab or Fv fragment) is required. In other words, the Fc portion of an antibody is dispensable for some applications.

An elaborate series of manipulations makes it possible to select, as well as produce, functional antibodies in *E. coli* (Fig. 10.34).

1. cDNA is synthesized from mRNA isolated from mouse antibody-producing cells (B lymphocytes).
2. The H and L chain sequences in the cDNA preparation are amplified separately by PCR.
3. Each amplified cDNA preparation is treated with a specific set of restriction endonucleases and cloned into a bacteriophage λ vector. The cDNA sequences of the H and L chains each have distinctive restriction endonuclease recognition sites, an arrangement that facilitates the directional cloning of each sequence into a separate bacteriophage λ vector. At this stage of the process, many different H and L chain sequences are cloned (Fig. 10.35A and B).
4. The cDNAs of one H and one L chain are cloned into a single "combinatorial" vector, thereby enabling the bacteriophage to coexpress

FIGURE 10.33 Organization of V_H and V_L regions of a monoclonal antibody and the development of a peptide, based on a portion of the CDR and FR of the V_H and V_L regions of the antibody molecule, with a similar binding specificity.

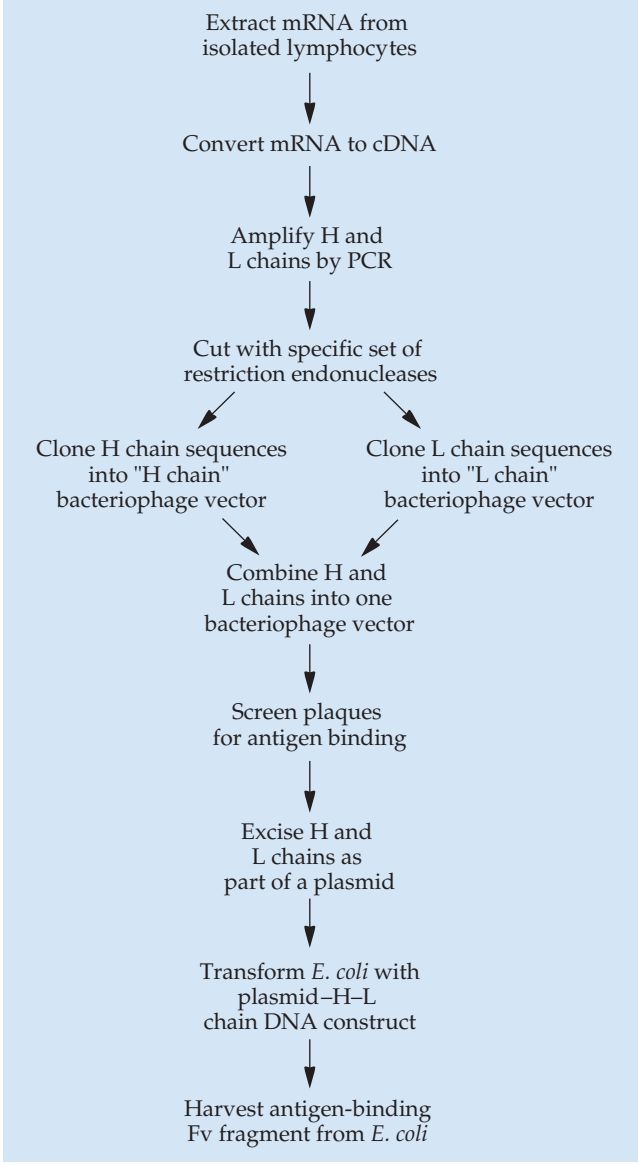


both chains, thus forming an assembled antibody Fv fragment (Fig. 10.35C).

5. The H and L chains are expressed during the lytic cycle of bacteriophage λ, so that the library of combinatorial bacteriophage clones can be screened for the presence of antigen-binding activity.

The step in which L and H chain cDNAs are combined on one vector creates a vast array of diverse antibody genes, some of which encode unique target-binding sites whose isolation would never have been possible by standard hybridoma procedures. The mammalian antibody repertoire has the potential to produce approximately 10^6 to 10^8 different antibodies. A phage library contains approximately this number of clones, so one combi-

FIGURE 10.34 Procedure to create a combinatorial library of the V_L and V_H regions of antibody chains in *E. coli*. Note that the H and L chains are amplified in separate PCRs.



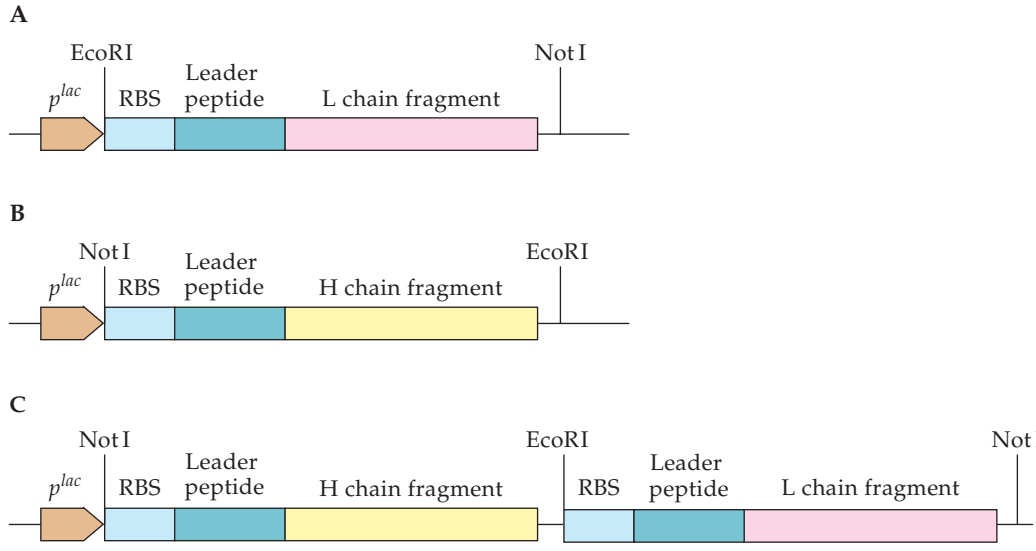


FIGURE 10.35 DNA constructs of an Fv combinatorial gene library cloned into bacteriophage λ DNA. **(A and B)** Portions of the cDNAs of the L (A) and H (B) chains are separately cloned into bacteriophage λ vectors. **(C)** Each of these libraries is digested with EcoRI, and then the fragments from the H chain library are ligated to the fragments from the L chain library, thereby creating a combinatorial library that contains all possible combinations of L and H chain fragments. p^{lac}, the *E. coli lac* promoter; RBS, ribosome-binding site.

natorial library can be expected to produce as many different antibodies (in this case, Fv molecules) as any mammal. In addition, once an initial combinatorial library has been constructed, it is possible to shuffle the L and H chains to obtain Fv molecules that recognize unusual epitopes, and even greater variation may be achieved by random mutagenesis of the DNAs in the combinatorial library (see chapter 8). Because millions of bacteriophage plaques can be screened in a relatively short period, the identification of Fv molecules with the desired specificity takes only about 7 to 14 days. By contrast, screening hybridoma cell lines is a slow, time-consuming process.

Because they lyse bacterial host cells, bacteriophage λ vectors are not particularly useful for the production of large quantities of protein. To overcome this drawback, the bacteriophage λ vector was engineered so that the H and L chain DNA sequences were inserted into a site that was flanked by plasmid DNA sequences. This plasmid DNA, containing an H and L chain DNA combination, can be excised from the bacteriophage λ vector and transformed into *E. coli* (Fig. 10.34). As part of a plasmid, large numbers of Fv fragments can be produced in *E. coli* cells.

As an alternative to the use of bacteriophage λ, filamentous bacteriophages, such as M13 and fd, have been used for the production of combinatorial libraries (Fig. 10.36). In these cases, the antibody fragment is synthesized as part of a fusion protein that is located on the outer surface of the bacteriophage. A combinatorial library of antibody fragments displayed on the surface of a filamentous bacteriophage can be screened by an enzyme-linked immunosorbent assay-like system. Briefly, samples (aliquots) of the library are added to the wells of a multiwell plate that are coated with the target antigen (Fig. 10.37). The wells are rinsed thoroughly to remove any unbound bacteriophage. Next, an antibody that binds to the

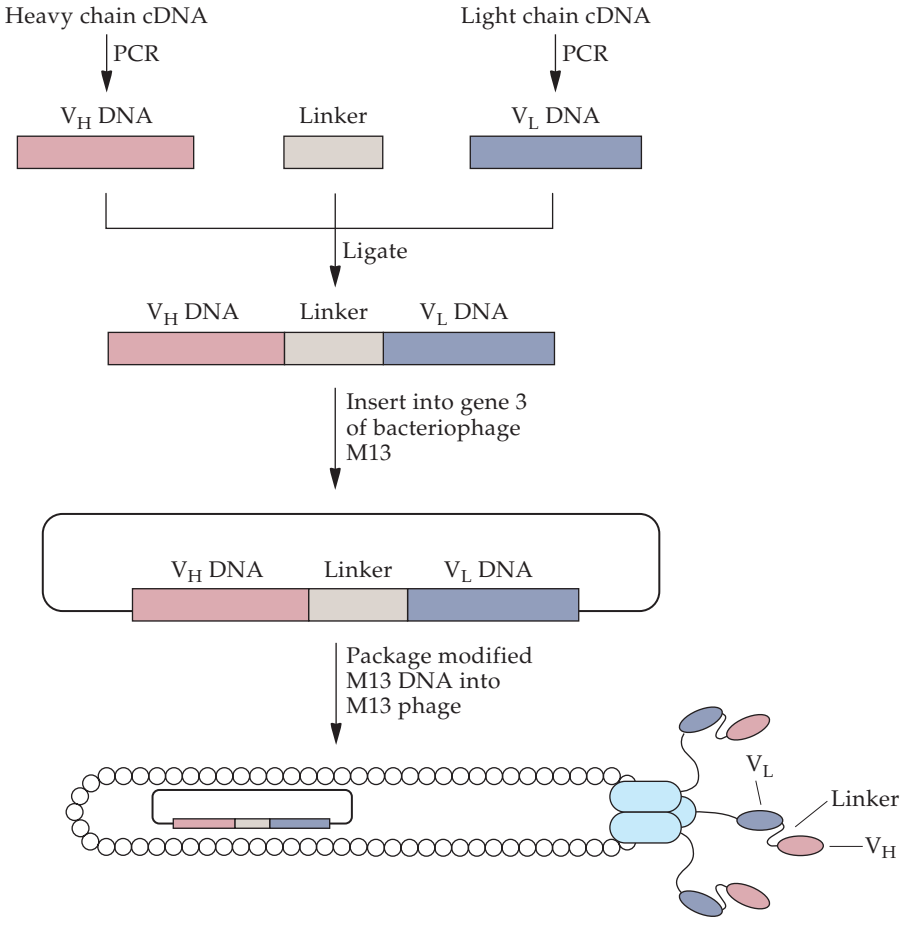


FIGURE 10.36 Formation of an Fv antibody combinatorial library in the filamentous bacteriophage M13. Following extraction of mRNA and its conversion to cDNA, the cDNAs for the V_L and V_H regions are amplified (separately) by PCR and then ligated to DNA that encodes a short linker peptide. Each single-chain antibody–DNA construct is cloned into gene 3 of bacteriophage M13. There are three copies of the phage gene 3 protein, which is a phage surface protein, per M13 bacteriophage.

bacteriophage coat protein and is conjugated with an enzyme is added to each well. The wells are rinsed to remove any unbound antibody–enzyme complex. The phage particles bound to the target antigen are recognized by the antibody–enzyme complex. A chromogenic substrate that is cleaved by the bound enzyme is then used to determine which wells contain a phage carrying antibodies to the target antigen. This approach is easier than using plaque assays with bacteriophage λ to select and subsequently purify a bacteriophage producing an antibody fragment that binds to a specific antigen. Once a desired antibody fragment-producing bacteriophage has been isolated, using either bacteriophage λ or M13, the DNA can be isolated and subcloned into an expression vector. These procedures are used to produce mouse, chimeric, or humanized antibodies.

A Combinatorial Library of Full-Length Antibodies

Until recently, all of the combinatorial libraries of antibodies included either single-chain antibodies or Fab fragments and not full-length antibodies.

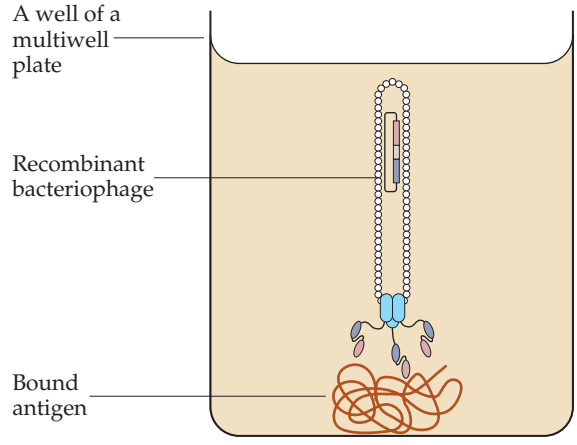


FIGURE 10.37 Immunological screening of a bacteriophage M13 combinatorial library.

However, for many applications it is advantageous for therapeutic antibodies to be full length. With this in mind, researchers cloned and expressed complete antibody molecules (using either two separate vectors, one encoding a light chain and the other encoding a heavy chain, or a single operon controlling the expression of both the light and heavy chain on a single vector) in *E. coli*.

The process of selecting antibodies with specified affinities would be greatly facilitated if an *E. coli* library could be screened directly for binding to various antigens. To do this, a combinatorial library of full-length antibodies was generated and engineered so that the antibodies were secreted into the periplasm between the inner and outer membranes. In addition, prior to the expression of the antibody molecules, the host *E. coli* was engineered to express a fusion protein that became anchored within the inner membrane (lipoprotein fragment) and also contained a portion of a protein from the *Staphylococcus aureus* protein A that binds specifically to Fc regions (Fig. 10.38). When an IgG molecule is secreted into the periplasm, the fusion protein binds to the Fc region. The tightly bound IgG–fusion protein complex remains intact when the *E. coli* cells are treated with EDTA and lysozyme to remove a portion of the outer membrane. Then, a fluorescently labeled target antigen is added to detect those cells that have expressed an antibody directed against that target antigen. By cell sorting, fluorescently labeled cells are selected. Then, the plasmid DNA that encodes the selected IgG molecules is isolated and expressed in an *E. coli* strain that does not synthesize the membrane-anchored fusion protein. This simple, yet powerful, technique simplifies the selection and production in *E. coli* of full-length monoclonal antibodies.

Shuffling CDR Sequences

Very large libraries of single-chain antibodies can be the sources of a wide range of highly specific human antibodies so that one does not have to resort to using mice for the initial monoclonal antibody. Theoretically, such libraries contain a greater diversity of antibodies than is normally found in the human immune system. To construct a library of this sort, B cells and other antibody-producing cells, from different, nonimmunized individuals

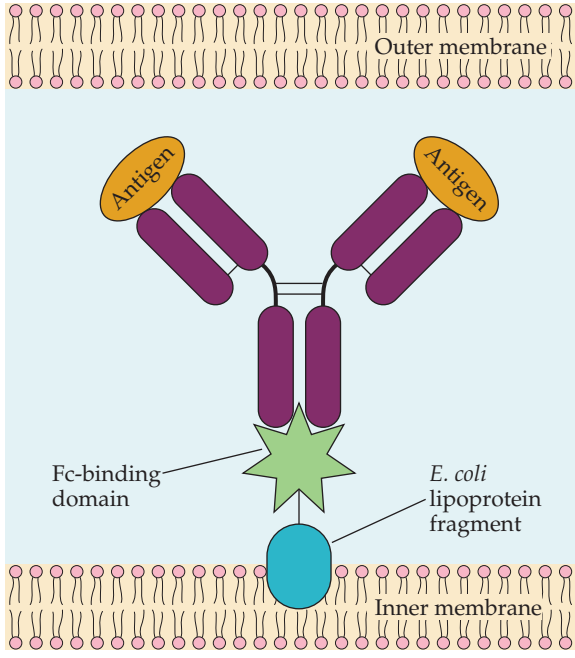


FIGURE 10.38 Schematic representation of the selection of a full-length antibody that binds to a fluorescently labeled target antigen. The Fc portion of the antibody is bound by a fragment of *S. aureus* protein A fused to a portion of an *E. coli* lipoprotein anchored in the inner membrane.

and from different tissues and organs, are collected and pooled before the mRNA is isolated (Fig. 10.39). The isolated mRNA is used to program the synthesis of cDNA, which then becomes a template for the specific PCR amplification of each CDR (separately). The amplified CDRs are mixed with oligonucleotides encoding FRs, a linker, and DNA sequences encoding the variable L and H domains. Overlap extension PCR (see chapter 4) is used to order, join, and amplify H and L antibody chain genes. In addition to being entirely human, the 2×10^9 different single-chain antibodies that have been produced in this way have a very wide range of specificities, reflecting the fact that the CDR sequences have been incorporated in random order and from a variety of sources, i.e., they have been shuffled. There are single-chain antibodies against β-galactosidase, the β subunit of cholera toxin, fluorescein isothiocyanate, human cell surface antigen, human leptin, human prostrate-specific antigen, and streptavidin, among others. Moreover, the dissociation constants (K_ds) for the interaction of the selected antibodies with their target antigens ranged from 0.9×10^{-9} to 420×10^{-9} M^{-1} (where tight binding is represented by a low number). For antibodies, K_ds of around 0.1×10^{-8} to 10×10^{-8} M^{-1} reflect a high degree of specificity for the target antigen. The fact that a very large library of single-chain antibodies, all with high affinity for their target antigens, could easily be generated by this procedure means that the goal of being able to select virtually any antibody from a nonimmunized library has been achieved.

Chemically Linked Monoclonal Antibodies

Drugs that are very effective when tested in cell culture are often much less potent in a whole organism. This difference in potency is typically due to

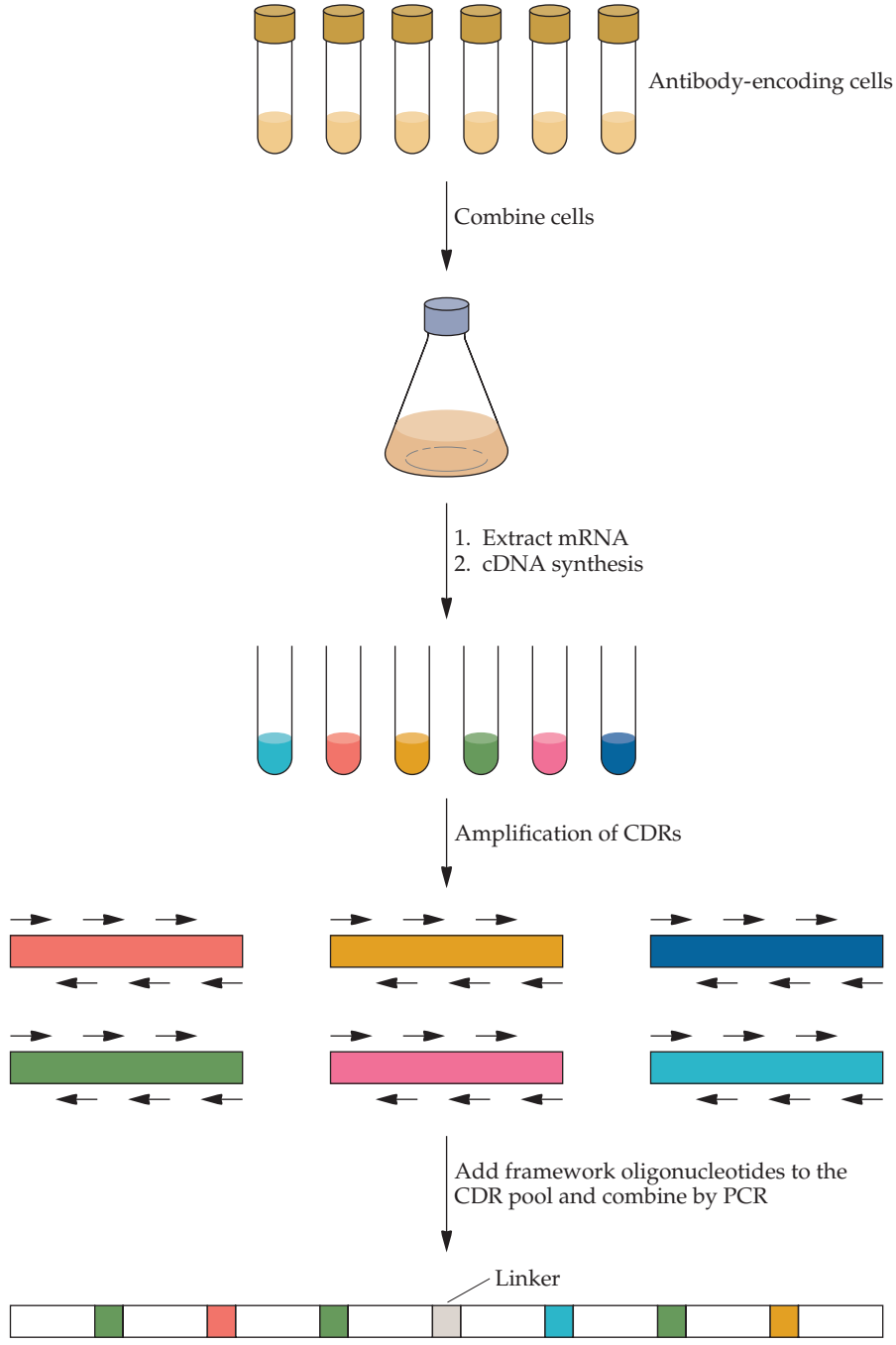


FIGURE 10.39 Construction of a large library of single-chain antibodies. B cells from several nonimmunized individuals were collected and pooled, the mRNA was isolated and used to program the synthesis of cDNA, oligonucleotide primers containing DNA sequences that included small portions of the FR sequence were added to the cDNA preparation, and all six CDRs were amplified separately by PCR. The amplified CDRs from all six PCRs were mixed together with oligonucleotides encoding the FRs and the linker, and genes encoding the variable L and H domains were synthesized by overlap extension PCR. One of the many possible single-chain antibodies that were synthesized is shown.

MILESTONE

Construction of a Retrovirus Packaging Mutant and Its Use To Produce Helper-Free Defective Retrovirus

R. MANN, R. C. MULLIGAN, and D. BALTIMORE
Cell **33:**153–159, 1983

Human gene therapy has always been a controversial subject. However, most scientists agreed with T. Friedmann and R. Roblin, who in 1972 (*Science* **175:**949–955) stated, "In our view, gene therapy may ameliorate some human genetic diseases in the future. For this reason, we believe that research directed at the development of techniques for gene therapy should continue." The essential features of human gene therapy are the delivery of a remedial gene and its expression in a cell type or tissue that cures a disease without risk to either those administering the therapeutic agent or those receiving it. Initially, vectors derived from human viruses seemed likely to be the principal mode for delivering remedial genes because they have specialized mechanisms for entering specific cells. In particular, of the various potential viral vectors, those based on retroviruses were considered the most promising. However, a native retrovirus is an infectious agent that can cause cell damage and, in some instances, induce cancer. The most significant advance that made human gene therapy possible was the development of a system for packaging a remedial gene into a noninfectious virus particle that retains the capability of attaching to its host cell.

Mann et al. constructed the first retroviral packaging cell line. In essence, they integrated a viral genome from which they had removed a DNA segment that contained the packaging signal into a chromosome of a cell line. Under these conditions, the cell line produced noninfectious virus particles. However, after transfection of these cells with a DNA construct that had a packaging signal and a remedial gene but no retroviral genes, the construct was packaged into virus particles that in turn could be used to deliver a remedial gene to a particular cell type. This clever strategy was immediately adopted by many other researchers who were working on human gene therapy. Over the years, the original retrovirus cell-packaging line has been enhanced, and the concept has been successfully applied to other viral vectors.

In 1985, as a result of the work of Mann et al. and others, W. F. Anderson, who has been a persistent advocate for human gene therapy, noted, "It now appears that effective delivery-expression systems are becoming available that will allow reasonable attempts at human gene therapy." On 22 May 1989, Anderson and his colleagues initiated the first clinical trial using a gene therapy strategy. To date, gene therapy in general has not been particularly successful; however, as more information accumulates from ongoing studies, it is inevitable that it will become the standard mode of treatment for many diseases.

the drug not being able to reach the targeted site in the whole animal at a concentration sufficient to be effective. Increasing the dose of a drug is not the answer to this problem, because high drug concentrations often have deleterious side effects. A number of different strategies are used to enhance the delivery of a drug to its target site. (1) Drugs may be encapsulated in liposomes, i.e., particles in which the drug is surrounded by a specific lipid surface, that can be targeted to certain organs. (2) Certain toxin genes may be incorporated into tumor-infiltrating lymphocytes. These cells can deliver the incorporated toxin directly to the site of a tumor. (3) A drug can be coupled to a monoclonal antibody that is specific for proteins found only on the surfaces of certain cells, e.g., tumor cells (Fig. 10.40). (4) A prodrug is an inert form of a drug that requires a specific enzyme to be activated. To ensure that the drug is released only in the vicinity of the target cells, the activating enzyme is coupled to a monoclonal antibody directed against specific cell surface antigens (Fig. 10.40).

For this type of therapeutic system to be effective, the monoclonal antibody or single-chain antibody that is complexed with the prodrug-converting enzyme must be available in quantity in a relatively pure form, bind to a protein that is highly specific to the target cell, be stable under physiological conditions but cleared rapidly from circulation, and, when necessary, be able to penetrate into tumor masses so that all of the cells can

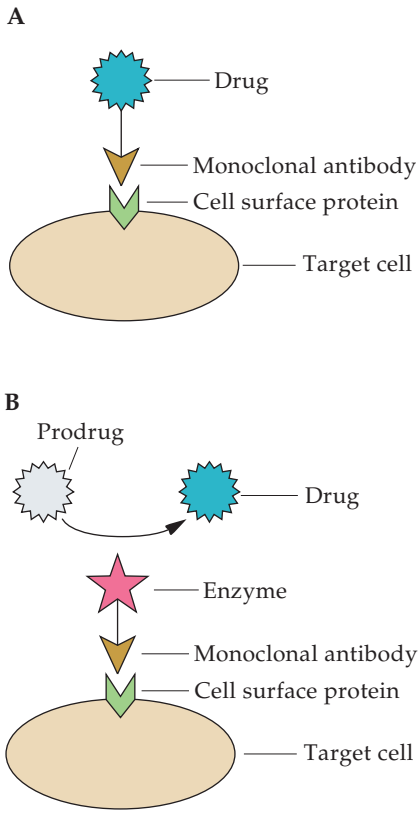


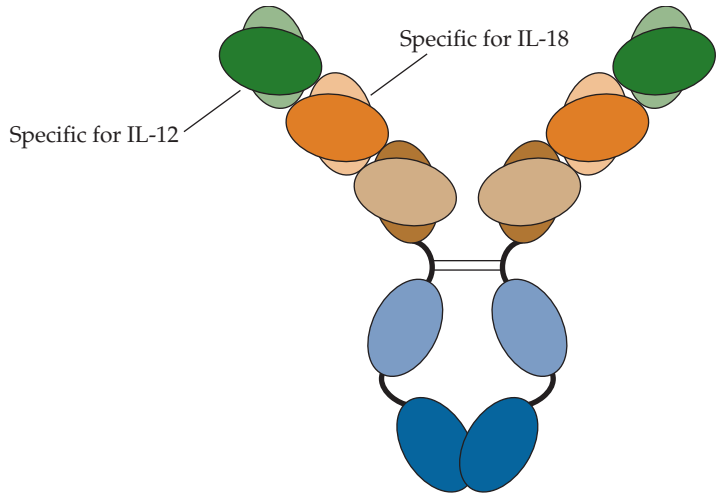
FIGURE 10.40 Schematic representation of a monoclonal antibody-based drug delivery system. **(A)** The drug is coupled to a monoclonal antibody. **(B)** An enzyme that converts an inactive prodrug to an active drug is attached to a monoclonal antibody. The active drug is formed only in the immediate vicinity of the target cells. In both cases, the monoclonal antibody binds to a specific protein on the surface of the target cell.

be exposed to the drug. With this approach, only specifically targeted cells are exposed to the drug, permitting the use of a much lower concentration than if it were administered directly.

Dual-Variable-Domain Antibodies

In some instances, an antibody (particularly one that is conjugated to a toxin or radiochemical) is able to destroy a tumor or pathogen cell. In such cases, it is often advantageous to use antibody fragments, since the Fc portion of the molecule is not only not needed, it may impede or prevent the rest of the molecule from binding to relatively inaccessible antigens. Despite the usefulness of antibody fragments in a variety of applications, a major limitation of using them as therapeutic agents is that, since they lack the Fc portion of the molecule, they are unable to mount a complete immune response. To increase the utility of naturally existing antibodies, as well as to ensure that they are effective initiators of a complete immune response, researchers have created what they have termed "dual-variable-domain immunoglobulins" (Fig. 10.41). These constructs are essentially IgG molecules containing two tandem Fv regions, each with a different specificity. Dual-variable-domain immunoglobulins are bispecific and tetravalent, and they consist entirely of human immunoglobulin sequences. That is, each molecule contains four Fv regions, two identical Fv regions directed against one antigen and two Fv regions directed against another antigen. A dual-variable-domain immunoglobulin specific for both interleukin-12 and interleukin-18 produced in CHO cells showed binding to both of these cytokines, binding each with an affinity similar to that of the original monospecific antibody. In addition, in a biological assay, using a severe combined immunodeficient mouse model engrafted with human peripheral blood mononuclear cells, the interleukin-12/interleukin-18 antibody was as effective at inhibiting induced IFN-γ production as was a combination of the original two monospecific antibodies. Using this strategy, it should be possible to generate full-size bispecific antibodies for a variety of therapeutic purposes.

FIGURE 10.41 Dual-variable-domain immunoglobulin directed against both interleukin-12 (IL-12) and interleukin-18.



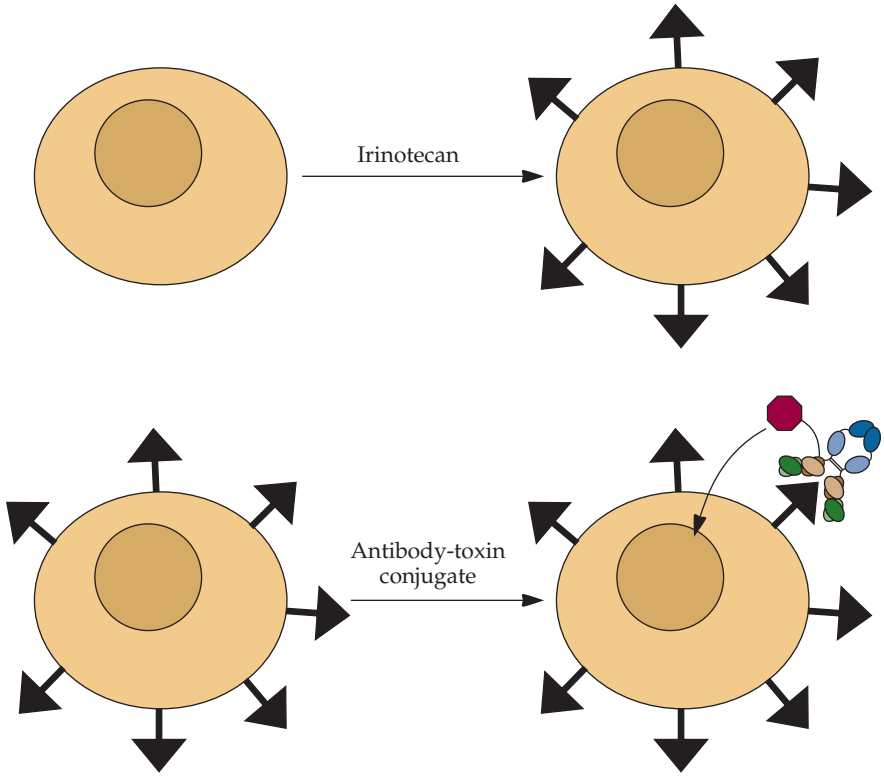


FIGURE 10.42 Targeting tumor cells for destruction by monoclonal antibody–toxin conjugates. Tumor cells are first treated with the chemotherapeutic agent irinotecan, which induces the synthesis of a unique cell surface protein. Then, a monoclonal antibody directed against the cell surface protein and conjugated to a toxin molecule is added. After binding of the antibody to the cell surface, the toxin is internalized, thereby killing the tumor cell.

Anticancer Antibodies

A number of therapeutic antibodies that are directed against protein antigenic determinants on the surfaces of cancer cells have been selected because the proteins are overexpressed compared to those on noncancerous cells. Unfortunately, (1) antibodies directed against these proteins may also bind to some noncancerous cells expressing the same or a similar antigen and (2) this approach presents researchers with only a limited number of targets for therapeutic antibodies. One way to select for additional cell surface targets would be to identify proteins whose expression is selectively induced in tumor cells exposed to chemotherapeutic drugs. When colorectal cancer cells were treated with the drug irinotecan, which is a topoisomerase inhibitor and is commonly used to treat this type of cancer, several newly synthesized proteins were found on the surfaces of those cells. (Topoisomerases are enzymes that unwind DNA during either DNA replication or mRNA transcription.) The new cell surface proteins were expressed early, prior to any major effects of the chemotherapeutic compound on cell viability. Monoclonal antibodies directed against one newly synthesized cell surface protein (called LY6D/E48) were generated, and then the antibodies were complexed with the cellular toxin auristatin E. The antibody–toxin conjugate was then used to treat tumor cells that were

first treated with irinotecan (Fig. 10.42). Following binding to the cell surface protein, the antibody–toxin conjugate was internalized inside the tumor cell. With this strategy, in six out of eight mice tumors disappeared entirely, while in the other two mice, the tumors were dramatically decreased in size. This exciting approach will have to be tried with larger numbers of animals before it can begin clinical trials. However, provided that it is possible to identify one or more proteins that are specifically induced by chemotherapeutic agents and are not found on the surfaces of nontumor cells, this procedure could become a general strategy that is used to treat a variety of different types of human cancer.

SUMMARY

A large number of proteins that have potential as therapeutic agents have been synthesized from cloned genes in bacteria. Because most of these proteins are from eukaryotic organisms, the strategy for the isolation of a target gene often involves isolating mRNA enriched in the messenger of interest, synthesizing a cDNA library, and subcloning the selected target cDNA into an appropriate expression vector. In some instances, novel and useful variants of these proteins can be constructed either by shuffling functional domains of related genes or by directed replacement of functional domains of the cloned gene. In addition, long-acting and stable variants of some therapeutic proteins have been synthesized.

In some instances, genetically engineered enzymes may be used as therapeutic agents. For example, both recombinant DNase I and alginate lyase have been used in an aerosol form to decrease the viscosity of the mucus found in the lungs of patients with cystic fibrosis. In addition, phenylalanine ammonia lyase may help patients with phenylketonuria as a replacement for phenylalanine hydroxylase, α_1-antitrypsin may be used to limit some infections, and glycosidases may be utilized to convert blood groups A, B, and AB to type O.

The development of recombinant DNA and monoclonal antibody technologies, combined with an understanding of the molecular structure and function of immunoglobulin molecules, has provided specific antibodies as therapeutic agents to treat various diseases. Antibody genes can be readily manipulated because the various functions of an antibody molecule are confined to discrete domains.

Drugs, prodrugs, or enzymes can be coupled to monoclonal antibodies or Fv fragments that are specific for proteins found only on the surfaces of certain cells, e.g., tumor cells. These antibody–drug or antibody–enzyme combinations act as therapeutic agents. However, if the therapy requires multiple treatments, the antibody component should be from a human source to prevent immunological cross-reactivity and sensitization of the patient. To achieve this, rodent monoclonal antibodies are "humanized" by substituting into human antibodies only the CDRs of the rodent monoclonal antibodies. In addition, it has become possible to produce and select human monoclonal antibodies in *E. coli* and in transgenic mice.

REFERENCES

Adams, G. P. and L. M. Weiner. 2005. Monoclonal antibody therapy of cancer. *Nat. Biotechnol.* **23:**1147–1157.

Alkawash, M. A., J. S. Soothill, and N. L. Schiller. 2006. Alginate lyase enhances antibiotic killing of mucoid *Pseudomonas aeruginosa* in biofilms. *APMIS* **114:**131–138.

Barbas, C. F., III, and D. R. Burton. 1996. Selection and evolution of high-affinity human anti-viral antibodies. *Trends Biotechnol.* **14:**230–234.

Bermúdez-Humarán, L. G., S. Nouaille, V. Zilberfarb, G. Corthier, A. Gruss, P. Langella, and T. Issad. 2007. Effect of intranasal administration of a leptin-secreting *Lactococcus lactis* recombinant on food intake, body weight, and immune response of mice. *Appl. Environ. Microbiol.* **73:**5300-5307.

Bird, R. E., and B. W. Walker. 1991. Single chain antibody variable regions. *Trends Biotechnol.* **9:**132–137.

Braat, H., P. Rottiers, D. W. Hommes, N. Huyghebaert, E. Remaut, J.-P. Remon, S. J. H. van Deventer, S. Neirynck, M. P. Peppelenbosch, and L. Steidler. 2006. A phase I trial with transgenic bacteria expressing interleukin-10 in Crohn's disease. *Clin. Gastroenterol. Hepatol.* **4:**754–759.

Brideau-Andersen, A. D., X. Huang, S.-C. C. Sun, T. T. Chen, D. Stark, I. J. Sas, L. Zadik, G. N. Dawes, D. R. Guptill, R. McCord, S. Govindarajan, A. Roy, S. Yang, J. Gao, Y. H. Chen, N. J. Ø. Skartved, A. K. Pedersen, D. Lin, C. P. Locher, I. Rebbarpragada, A. D. Jensen, S. H. Bass, T. L. S. Nissen, S. Viswanathan, G. R. Foster, J. A. Symons, and P. A. Patten. 2007. Directed evolution of gene-shuffled IFN-α molecules with activity profiles tailored for treatment of chronic viral diseases. *Proc. Natl. Acad. Sci. USA* **104:**8269–8274.

Brinkmann, U., L. H. Pai, D. J. FitzGerald, M. Willingham, and I. Pastan. 1991. B3(Fv)-PE38KDEL, a single-chain immunotoxin that causes complete regression of a human carcinoma in mice. *Proc. Natl. Acad. Sci. USA* **88:**8616–8620.

Burton, D. R. 1991. Human and

mouse monoclonal antibodies by repertoire cloning. *Trends Biotechnol.* **9:**169–175.

Chamow, S. M., and A. Ashkenazi. 1996. Immunoadhesins: principles and applications. *Trends Biotechnol.* **14:**52–60.

Chester, K. A., and R. E. Hawkins. 1995. Clinical issues in antibody design. *Trends Biotechnol.* **13:**294–300.

Collet, T. A., P. Roben, R. O'Kennedy, C. F. Barbas III, D. R. Burton, and R. A. Lerner. 1992. A binary plasmid system for shuffling combinatorial antibody libraries. *Proc. Natl. Acad. Sci. USA* **89:**10026–10030.

Cunningham, B. C., and J. A. Wells. 1991. Rational design of receptor-specific variants of human growth hormone. *Proc. Natl. Acad. Sci. USA* **88:**3407–3411.

Curnis, F., A. Sacchi, L. Borgna, F. Magni, A. Gasparri, and A. Corti. 2000. Enhancement of tumor necrosis factor α antitumor immunotherapeutic properties by targeted delivery to aminopeptidase N (CD13). *Nat. Biotechnol.* **18:**1185–1190.

de Vos, W. M. and J. Hugenholtz. 2004. Engineering metabolic highways in lactococci and other lactic acid bacteria. *Trends Biotechnol.* **22:**72–79.

Dübel, S. 2007. Recombinant therapeutic antibodies. *Appl. Microbiol. Biotechnol*. **74:**723–729.

Gram, H., L. A. Marconi, C. F. Barbas III, T. A. Collet, R. A. Lerner, and A. S. Kang. 1992. In vitro selection and affinity maturation of antibodies from a naive combinatorial immunoglobulin library. *Proc. Natl. Acad. Sci. USA* **89:**3576–3580.

Hanniffy, S., U. Wiedermann, A. Repa, A. Mercenier, C. Daniel, J. Fioamonti, H. Tlaskolova, H. Kozakova, H. Israelsen, S. Madsen, A. Vrang, P. Hols, J. Delcour, P. Bron, M. Kleerebezem, and J. Wells. 2004. Potential and opportunities for use of recombinant lactic acid bacteria in human health. *Adv. Appl. Microbiol.* **56:**1–64.

Harris, W. J. 1994. Humanizing monoclonal antibodies for in vivo use. *Anim. Cell Biotechnol.* **6:**259–279.

Hodgson, J. 1991. Making monoclonals in microbes. *Bio/Technology* **9:**421–425.

Holliger, P. and P. J. Hudson. 2005. Engineered antibody fragments and the rise of single domains. *Nat. Biotechnol.* **23:**1126–1136.

Huennekens, F. M. 1994. Tumor targeting: activation of prodrugs by enzyme-monoclonal antibody conjugates. *Trends Biotechnol.* **12:**234–239.

Huse, W. D., L. Sastry, S. A. Iverson, A. S. Kang, M. Alting-Mees, D. R. Burton, S. J. Benkovic, and R. A. Lerner. 1989. Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. *Science* **246:**1275–1281.

Jakobovits, A., R. G. Amado, X. Yang, L. Roskos, and G. Schwab. 2007. From XenoMouse technology to panitumumab, the first fully human antibody product from transgenic mice. *Nat. Biotechnol.* **25:**1134–1143.

Jean, F., L. Thomas, S. S. Molloy, G. Liu, M. A. Jarvis, J. A. Nelson, and G. Thomas. 2000. A protein-based therapeutic for human cytomegalovirus infection. *Proc. Natl. Acad. Sci. USA* **97:**2864–2869.

Kreitman, R. J. 1999. Immunotoxins in cancer therapy. *Curr. Opin. Immunol.* **11:**570–578.

Kufer, P., R. Lutterbrüse, and P. A. Baeuerle. 2004. A revival of bispecific antibodies. *Trends Biotechnol.* **22:**238–244.

Little, M., F. Breitling, S. Dübel, P. Fuchs, and M. Braunagel. 1995. Human antibody libraries in *Escherichia coli*. *J. Biotechnol.* **41:**187–195.

Liu, Q. P., G. Sulzenbacher, H. Yuan, E. P. Bennett, G. Pietz, K. Saunders, J. Spence, E. Nudelman, S. B. Levery, T. White, J. M. Neveu, W. S. Lane, Y. Bourne, M. L. Olsson, B. Henrissat, and H. Clausen. 2007. Bacterial glycosidases for the production of universal red blood cells. *Nat. Biotechnol.* **25:**454–464.

Liu, X., L. A. Lagenaur, D. A. Simpson, K. P. Essenmacher, C. L. Frazier-Parker, Y. Liu, S. S. Rao, D. H. Hamer, T. P. Parks, P. P. Lee, and Q. Xu. 2006. Engineered vaginal *Lactobacillus* strain for mucosal delivery of the human immunodeficiency virus inhibitor cyanovirin-N. *Antimicrob. Agents Chemother*. **50:**3250–3259.

LoBuglio, A. G., R. H. Wheeler, J. Trang, A. Haynes, K. Rogers, E. B. Harvey, L. Sun, J. Ghrayeb, and M. B. Khazaeli. 1989. Mouse/human chimeric monoclonal antibody in man: kinetics and immune response. *Proc. Natl. Acad. Sci. USA* **86:**4220–4224.

Lonberg, N. 2005. Human antibodies from transgenic animals. *Nat. Biotechnol.* **23:**1117–1125.

Lorenzen, N., P. M. Cupit, K. Einer-Jensen, E. Lorenzen, P. Ahrens, C. J. Secombes, and C. Cunningham. 2000. Immunoprophylaxis in fish by injection of mouse antibody genes. *Nat. Biotechnol.* **18:**1177–1180.

Marks, J. D., A. D. Griffiths, M. Malmqvist, T. P. Clackson, J. M. Bye, and G. Winter. 1992. By-passing immunization: building high affinity antibodies by chain shuffling. *Bio/Technology* **10:**779–783.

Mazor, Y., T. Van Blarcom, R. Mabry, B. L. Iverson and G. Georgiou. 2007. Isolation of engineered full-length antibodies from libraries expressed in *Escherichia coli*. *Nat. Biotechnol*. **25:**563–565.

Molina, A. 2008. A decade of rituximab: improving survival outcomes in non-Hodgkin's lymphoma. *Annu. Rev. Med*. **59:**237–250.

Mullinax, R. L., E. A. Gross, J. R. Amberg, B. N. Hay, H. H. Hogrefe, M. M. Kubitz, A. Greener, M. Alting-Mees, D. Ardourel, J. M. Short, J. A. Sorge, and B. Shopes. 1990. Identification of human antibody fragment clones specific for tetanus toxoid in a bacteriophage λ immunoexpression library. *Proc. Natl. Acad. Sci. USA* **87:**8095–8099.

Murata, K., T. Inose, T. Hisano, S. Abe, Y. Yonemoto, T. Yamashita, M. Takagi, K. Sakaguchi, A. Kimura, and T. Imanaka. 1993. Bacterial alginate lyase: enzymology, genetics and application. *J. Ferment. Bioeng*. **76:**427–437.

Nagata, S., H. Taira, A. Hall, L. Johnsrud, M. Streuli, J. Ecsödi, W. Boll, K. Cantell, and C. Weissmann. 1980. Synthesis in *E. coli* of a polypeptide with human leukocyte interferon activity. *Nature* **284:**316–320.

Osborn, B. L., L. Sekut, M. Corcoran, C. Poortman, B. Sturm, G. Chen, D. Mather, H. L. Lin, and T. J. Parry. 2002. Albutropin: a growth hormone-albumin fusion with improved phar-

macokinetics and pharmacodynamics in rats and monkeys. *Eur. J. Pharmacol.* **456:**149–158.

Plückthun, A. 1991. Antibody engineering: advances from the use of *E. coli* expression systems. *Bio/Technology* **9:**545–551.

Qiu, X.-Q., H. Wang, B. Cai, L.-L. Wang, and S.-T. Yue. 2007. Small antibody mimetics comprising two complementarity-determining regions and a framework region for tumor targeting. *Nat. Biotechnol.* **25:**921–929.

Queen, C., W. P. Schneider, H. E. Selick, P. W. Payne, N. F. Landolf, J. F. Duncan, N. M. Avdalovic, M. Levitt, R. P. Junghans, and T. A. Waldmann. 1989. A humanized antibody that binds to the interleukin 2 receptor. *Proc. Natl. Acad. Sci. USA* **86:**10029–10033.

Reichert, J. M. 2006. Trends in US approvals: new biopharmaceuticals and vaccines. *Trends Biotechnol.* **24:**293–298.

Reichert, J. M., C. J. Rosensweig, L. B. Faden, and M. C. Dewitz. 2005. Monoclonal antibody success in the clinic. *Nat. Biotechnol*. **23:**1073–1078.

Reiter, Y., U. Brinkmann, K. O. Webber, S.-H. Jung, and I. Pastan. 1994. Engineering interchain disulfide bonds into conserved framework regions of Fv fragments: improved biochemical characteristics of recombinant immunotoxins containing disulfide-stabilized Fv. *Protein Eng.* **7:**697–704.

Rubinfeld, B., A. Upadhyay, S. L. Clark, S. E. Fong, V. Smith, H. Koeppen, S. Ross, and P. Polakis. 2006. Identification and immunotherapeutic targeting of antigens induced by chemotherapy. *Nat. Biotechnol.* **24:**205–209.

Sampson, J. H., L. E. Crotty, S. Lee, G. E. Archer, D. M. Ashley, C. J. Wikstrand, L. P. Hale, C. Small, G. Dranoff, A. H. Friedman, H. S. Friedman, and D. D. Bigner. 2000. Unarmed, tumor-specific monoclonal antibody effectively treats brain tumors. *Proc. Natl. Acad. Sci. USA* **97:**7503–7508.

Sarkissian, C. H., Z. Shao, F. Blain, R. Peevers, H. Su, R. Heft, T. M. S. Chang, and C. R. Scriver. 1999. A different approach to treatment of phenylketonuria: phenylalanine degradation with recombinant phenylalanine ammonia lyase. *Proc. Natl. Acad. Sci. USA* **96:**2339–2344.

Söderlind, E., L. Strandberg, P. Jirholt, N. Kobayashi, B. Alexeiva, A.-M. Åberg, A. Nilsson, B. Jansson, M. Ohlin, C. Wingren, L. Danielsson, R. Carlsson, and C. A. K. Borrebaeck. 2000. Recombining germline-derived CDR sequences for creating diverse single-framework antibody libraries. *Nat. Biotechnol.* **18:**852–856.

Steidler, L., W. Hans, L. Schotte, S. Neirynck, F. Obermeier, W. Falk, W. Fiers, and E. Remaut. 2000. Treatment of murine colitis by *Lactococcus lactis* secreting interleukin-10. *Science* **289:**1352–1355.

Steidler, L., S. Neirynck, N. Huyghebaert, V. Snoeck, A. Vermeire, B. Goddeeris, E. Cox, J. P. Remon, and E. Remaut. 2003. Biological containment of genetically modified *Lactococcus lactis* for intestinal delivery of human interleukin 10. *Nat. Biotechnol.* **21:**785–789.

Subramanian, G. M., M. Fiscella, A. Lamousé-Smith, S. Zeuzem, and J. G. McHutchison. 2007. Albuferon α-2b: a generic fusion protein for the treatment of chronic hepatitis C. *Nat. Biotechnol.* **25:**1411–1419.

Taniguchi, T., Y. Fujii-Kuriyama, and M. Muramatsu. 1980. Molecular cloning of human interferon cDNA. *Proc. Natl. Acad. Sci. USA* **77:**4003–4006.

Ulmer, G. S., A. Herzka, K. J. Toy, D. L. Baker, A. H. Dodge, D. Sinicropi, S. Shak, and R. A. Lazarus. 1996. Engineering actin-resistant human DNase I for treatment of cystic fibrosis. *Proc. Natl. Acad. Sci. USA* **93:**8225–8229.

Vaughan, T. J., A. J. Williams, K. Pritchard, J. K. Osbourn, A. R. Pope, J. C. Earnshaw, J. McCafferty, R. A. Hodits, J. Wilton, and K. S. Johnson. 1996. Human antibodies with subnanomolar affinities isolated from a large non-immunized phage display library. *Nat. Biotechnol.* **14:**309–314.

Wilkenson, I. R., E. Ferrandis, P. J. Artymiuk, M. Teillot, C. Soulard, C. Touvay, S. L. Pradananga, S. Justice, Z. Wu, K. C. Leung, C. J. Strasburger, J. R. Sayers, and R. J. Ross. 2007. A ligand-receptor fusion of growth hormone forms a dimer and is a potent long-acting agonist. *Nat. Med.* **13:**1108–1113.

Wu, C., H. Ying, C. Grinnell, S. Bryant, R. Miller, A. Clabbers, S. Bose, D. McCarthy, R.-R. Zhu, L. Santora, R. Davis-Taber, Y. Kunes, E. Fung, A. Schwartz, P. Sakorafas, J. Gu, E. Tarcsa, A. Murtaza, and T. Ghayur. 2007. Simultaneous targeting of multiple disease mediators by a dual-variable-domain immunoglobin. *Nat. Biotechnol.* **25:**1290–1297.

Zahm, J.-M., C. Debordeaux, C. Maurer, D. Hubert, D. Dusser, N. Bonnet, R. A. Lazarus, and E. Puchelle. 2001. Improved activity of an actin-resistant DNase I variant on the cystic fibrosis airway secretions. *Am. J. Respir. Crit. Care Med.* **163:**1153–1157.

REVIEW QUESTIONS

1. Before the sequencing of the human genome, how would you have cloned and expressed a cDNA sequence encoding human IFN? You do not have a DNA hybridization probe for human IFN, although you have isolated a human cell line that can be induced to synthesize IFN approximately 100-fold over background levels. Explain your strategy.

2. What is the Fc portion of an antibody molecule? The Fab portion? The Fv portion? The CDR portion?

3. How are antibody light and heavy chains coordinately synthesized in *E. coli*?

4. How would you modify growth hormone to make it longer-acting?

5. Why would DNase I and alginate lyase be useful for treating cystic fibrosis?

6. How is the production of alginate lyase from a cloned gene detected in *E. coli* transformants?

7. What is a combinatorial cDNA library?

8. How is bacteriophage M13 used to select Fv fragments that bind to specific target antigens?

9. What are disulfide-stabilized and scFv molecules?

10. How are enzymes coupled to monoclonal antibodies or Fv fragments used as therapeutic agents?

11. How are mouse monoclonal antibodies "humanized"? Discuss the reasons for creating humanized monoclonal antibodies.

12. Describe a protocol for producing a therapeutic agent that targets and kills a specific cell type.

13. How would you engineer TNF-α to be a more specific and effective anticancer agent?

14. What would you do to make interleukin-10 more effective for treating inflammatory bowel disease?

15. How can the gene for DNase I be manipulated so that the enzyme becomes more effective for treating cystic fibrosis patients?

16. How would you develop a strategy to protect at-risk women from HIV infection?

17. How might low levels of phenylalanine be attained, other than with a phenylalanine-free diet, in patients with the human genetic disease phenylketonuria?

18. What types of genetic manipulations can be used to generate a very large bacterial library of highly specific single-chain human monoclonal antibodies?

19. How would you engineer a mouse so that it produces only human antibodies?

20. What is a bispecific diabody?

21. How would you design a short peptide so that it retains the antigen-binding specificity of an entire immunoglobulin molecule?

22. How would select antibodies against specific antigens in *E. coli?*

23. What are dual-variable-domain immunoglobulins?

24. How can you use a chemotherapy agent to facilitate the targeting of tumor cells with monoclonal antibodies?

11

Nucleic Acids as Therapeutic Agents

OFTEN, HUMAN DISORDERS, such as cancer, inflammatory conditions, and both viral and parasitic infections, result from the overproduction of a normal protein. Therapeutic systems using nucleotide sequences are being devised to treat these types of conditions. Theoretically, a small single-stranded nucleotide sequence (oligonucleotide) could hybridize to a specific gene or messenger RNA (mRNA) and diminish transcription or translation, respectively, thereby decreasing the amount of protein that is synthesized. An oligonucleotide that is designed to bind to a gene and block transcription is called an antigene oligonucleotide, and one that base pairs with a specific mRNA is called an antisense oligonucleotide. The binding of an oligonucleotide to a transcription factor that is responsible for the expression of a specific gene could lower both transcription and translation of the target gene. Double-stranded oligonucleotides that attach to DNA-binding proteins could prevent the activation of transcription of specific genes. Also, some synthetic RNA/DNA molecules called aptamers that bind to proteins that are not naturally nucleic acid-binding proteins and prevent them from functioning can be created. Ribozymes, which are natural RNA sequences that bind and cleave specific RNA molecules, could be engineered to target an mRNA and subsequently decrease the amount of a particular protein that is synthesized. In addition, interfering RNAs, small double-stranded RNA molecules that direct the sequence-specific degradation of mRNA, may be used instead of either antisense RNAs (or oligonucleotides) or ribozymes. The potential for nucleic acid therapeutic agents is considerable and is just now beginning to be realized.

Antisense RNA

To be an effective therapeutic agent, an antisense RNA must bind to a specified mRNA and prevent translation of the protein (Fig. 11.1A). The possibility of using an expression vector to produce an antisense RNA that suppresses a pathogenic condition has been examined. For example, episomally based expression vectors that carry the complementary DNA (cDNA)

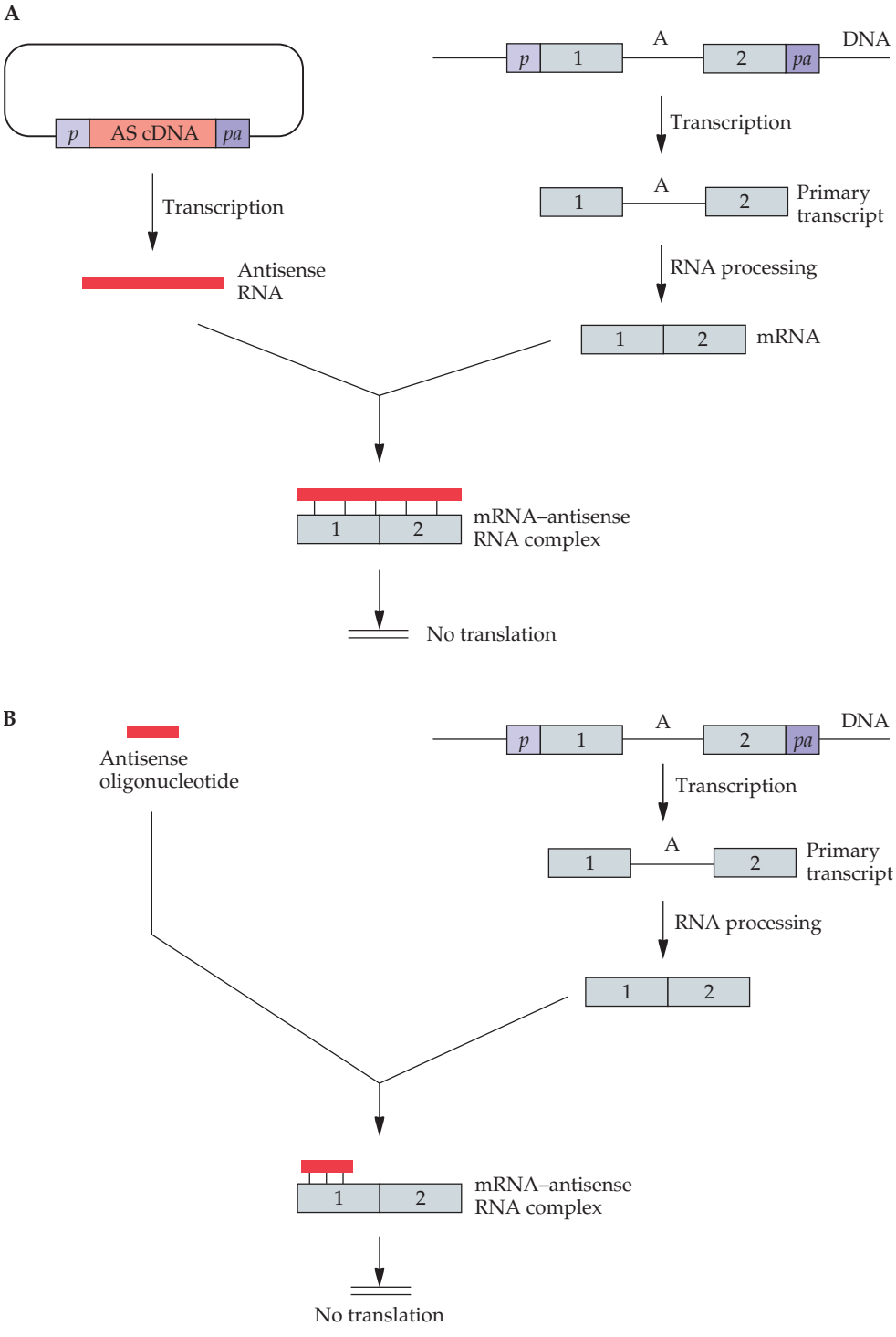


FIGURE 11.1 Inhibition of translation of specific mRNAs by antisense (AS) nucleic acid molecules. The promoter and polyadenylation regions are marked by *p* and *pa*, respectively; the intron is indicated by the letter A; and the exons are indicated by numbers (1 and 2). **(A)** A cDNA (AS gene) is cloned into an expression vector in reverse orientation, and the construct is transfected into a cell, where the AS RNA is synthesized. The AS RNA hybridizes to the target mRNA, and translation is blocked. **(B)** An AS oligonucleotide is introduced into a cell, and after it hybridizes with the target mRNA, translation is blocked.

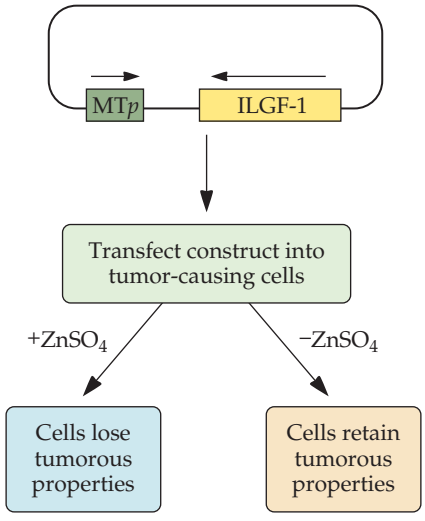


FIGURE 11.2 A cDNA for human insulin-like growth factor 1 (ILGF-1) cloned on a vector in the antisense orientation under the transcriptional control of a metallothionein promoter (MT*p*). Following transfection into tumor-causing cells, when low levels of $ZnSO_4$ are added, the cells have decreased tumorigenicity. The arrows above the gene and promoter indicate the normal direction of transcription. The origin of replication and other plasmid sequences have been omitted for clarity.

sequence for either insulin-like growth factor 1 or insulin-like growth factor 1 receptor were constructed with the cloned sequences oriented so that the transcripts were antisense rather than mRNA (sense) sequences. Insulin-like growth factor 1 is prevalent in malignant glioma, which is the most common form of human brain tumor. Excess production of insulin-like growth factor 1 receptor occurs in prostate carcinoma, which is a significant type of cancer in males. In both vectors, the reverse-oriented cDNAs are under the control of the metallothionein promoter, which is induced by low levels of $ZnSO_4$.

Cultured glioma cells were transfected with the vector that produces the antisense version of the insulin-like growth factor 1 mRNA. In the absence of $ZnSO_4$, the tumorous properties were retained; in contrast, when $ZnSO_4$ was added to the culture medium, these distinctive features were lost (Fig. 11.2). In another experiment, nontransfected glioma cells caused tumors after they were injected into rats, whereas glioma cells that had been transfected with antisense insulin-like growth factor 1 cDNA did not develop tumors.

When mice were injected with rat prostate carcinoma cells that were transfected with the insulin-like growth factor 1 receptor cDNA in the antisense orientation, they developed either small or no tumors, whereas large tumors were formed when mice were treated with either nontransfected or control-transfected rat prostate carcinoma cells. It was assumed that in both cases the antisense RNA hybridized with its complementary mRNA sequence and hindered translation of insulin-like growth factor 1 and insulin-like growth factor 1 receptor, thus preventing the proliferation of the cancer cells.

Antisense Oligonucleotides

The sequence-specific effectiveness of chemically synthesized antisense oligodeoxynucleotides (Fig. 11.1B) relies on hybridization to an accessible nucleotide sequence on the target mRNA, resistance to degradation by cellular nucleases, and ready delivery into cells. Oligonucleotides with about 15 to 24 nucleotides have sufficient specificity to hybridize to a unique mRNA. Potential mRNA target sites are determined by testing a set of antisense oligonucleotides with cells in culture that produce the target mRNA. Those antisense oligonucleotides that diminish the translation of the specified protein are selected. Proteomic analysis of cellular proteins that are labeled with fluorescent dyes during translation can be used to determine if the production of a particular protein is reduced in the presence of an antisense oligonucleotide. There are no general rules for predicting the best target sites in various RNA transcripts. Antisense oligonucleotides that are directed to the 5′ and 3′ ends of mRNAs, intron–exon boundaries, and regions that are naturally double stranded have all been effective.

Since oligodeoxynucleotides are susceptible to degradation by intracellular nucleases, it was important to find ways to synthesize molecules that are resistant to attack by nucleases without affecting the ability of the antisense oligonucleotide to hybridize to a target sequence. To this end, the backbone, pyrimidines, and sugar moiety have been modified (Fig. 11.3). Currently, the most extensively used antisense oligonucleotide has a sulfur group in place of the free oxygen of the phosphodiester bond (Fig. 11.3B). This modification is called a phosphorothioate linkage. Phosphorothioate antisense oligonucleotides are water soluble, polyanionic, and resistant to

endogenous nucleases. In addition, when a phosphorothioate antisense oligonucleotide hybridizes to its target site, the RNA–DNA duplex activates the endogenous enzyme ribonuclease H (RNase H), which cleaves the mRNA component of the hybrid molecule. Clinical trials with several phosphorothioate antisense oligonucleotides, which are considered to be "first-generation" therapeutic agents, have been initiated. Second-generation antisense oligonucleotides typically contain alkyl modifications at the 2′ position of the ribose (Fig. 11.3E) and are generally less toxic and more specific than phosphorothioate-modified molecules. Third-generation antisense oligonucleotides contain a variety of modifications within the ribose ring and/or the phosphate backbone, as well as being less toxic than either first- or second-generation antisense oligonucleotides. One phosphorothioate antisense oligonucleotide has been approved by the U.S. Food and Drug Administration (FDA) to treat cytomegalovirus infections of the retina in patients with acquired immune deficiency syndrome (AIDS). This particular antisense oligonucleotide, called fomivirsen and sold as Vitravene, is administered by injection of 330 µg in a volume of 50 µl directly into an affected eye after the application of a topical or local anesthetic. Fomivirsen treatment is typically once every 2 weeks for 4 weeks, followed by once

MILESTONE

Potent and Specific Genetic Interference by Double-Stranded RNA in *Caenorhabditis elegans*

A. Fire, S.-Q. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, and C. C. Mello

Nature **391:**806-811, 1998

The nematode worm *Caenorhabditis elegans* is studied as a model eukaryotic organism in part because the strains are inexpensive to breed and can be frozen. When the cells are thawed, they remain viable, allowing long-term storage. *C. elegans* has the advantage of being a multicellular eukaryotic organism that is simple enough to be studied in great detail. The developmental fate of every single somatic cell (959 in the adult hermaphrodite; 1,031 in the adult male) has been mapped out, and these patterns of cell lineage are largely invariant between individuals. In addition, *C. elegans* is one of the simplest organisms with a nervous system.

In the late 1990s, Andrew Fire, Craig Mello, and their colleagues were investigating how gene expression is regulated in *C. elegans*. When they injected worms with mRNA molecules encoding a *C. elegans* muscle protein, they did not observe any changes in the behavior of the worms. Injecting the antisense version of this mRNA also had no effect. However, when they injected sense and antisense RNAs together, they observed that the worms displayed peculiar twitching movements. Similar movements were seen in worms that completely lacked a functioning gene for the muscle protein.

Somehow the added double-stranded RNA molecule was silencing the expression of the gene carrying the same genetic information as that particular RNA. When double-stranded RNA molecules containing portions of the mRNA sequences for several other worm proteins were injected, the expression of these genes was also silenced.

From these experiments, Fire and Mello deduced that double-stranded RNA can silence genes, that this RNAi is specific for the gene whose sequence matches that of the injected RNA molecule, and that RNAi can spread between cells and even be inherited. In addition, since the injection of even tiny amounts of double-stranded RNA was sufficient to achieve an effect, Fire and Mello proposed that RNAi is a catalytic process.

Fire and Mello's discovery clarified many earlier confusing and apparently contradictory experimental observations and revealed a natural mechanism for controlling the flow of genetic information. Soon after this original report, other workers found interfering RNAs in a number of different systems from worms to mammals to plants. This work opened up a whole new field of research. Workers soon discovered that RNAi can regulate gene expression in hundreds of genes in our genome and that these small RNAs play an important role in animal and plant development and the control of cellular functions. RNAi also appeared to protect the genome against transposons and viruses, and it opened up exciting possibilities for use as a therapeutic agent. In 2006, Fire and Mello received the Nobel Prize in Physiology or Medicine for their pioneering work.

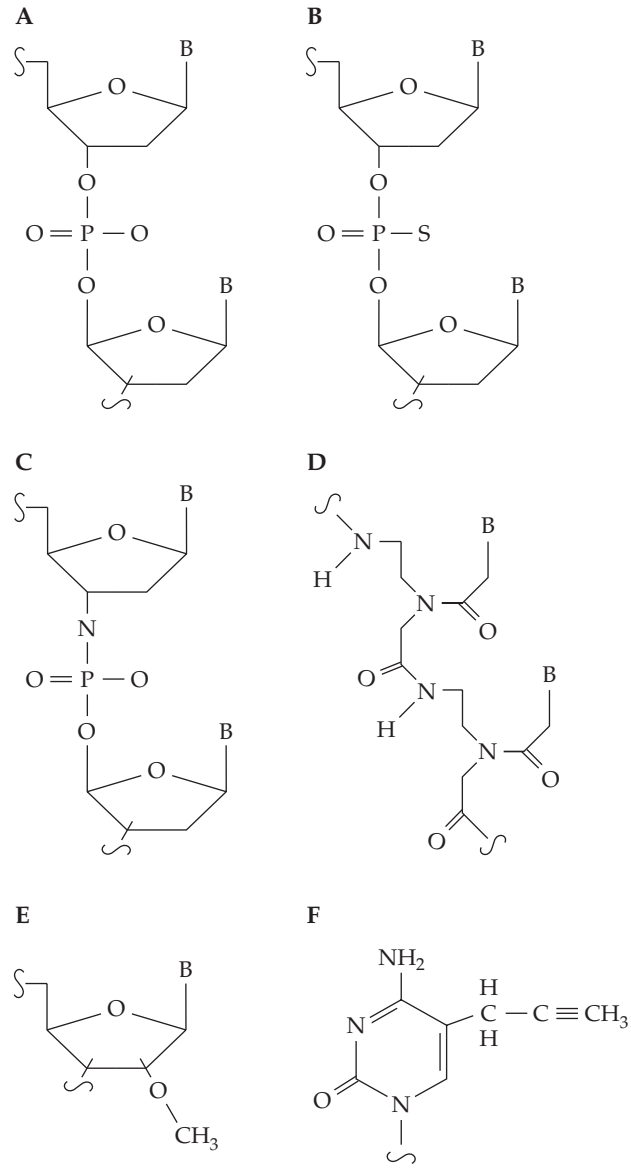


FIGURE 11.3 Modifications to antisense oligonucleotides. **(A)** Phosphodiester linkage; **(B)** phosphorothioate linkage; **(C)** phosphoramidite linkage; **(D)** polyamide linkage (peptide nucleic acid); **(E)** 2′-*O*-methyl ribose; **(F)** C-5 propynylcytosine.

every 4 weeks. Before treatment with fomivirsen is started, it is essential that the presence of cytomegalovirus be absolutely confirmed, since several other infective agents produce similar symptoms.

Antisense oligonucleotides with phosphoramidate and polyamide (peptide) linkages have been synhesized in the expectation that these molecules should be very resistant to nuclease degradation (Fig. 11.3C and D). Furthermore, as mentioned above, chemical groups have been added to the 2′ carbon of the sugar moiety and the 5 carbon (C-5) of pyrimidines to both enhance stability and facilitate the binding of the antisense oligonucleotide to its target site (Fig. 11.3E and F).

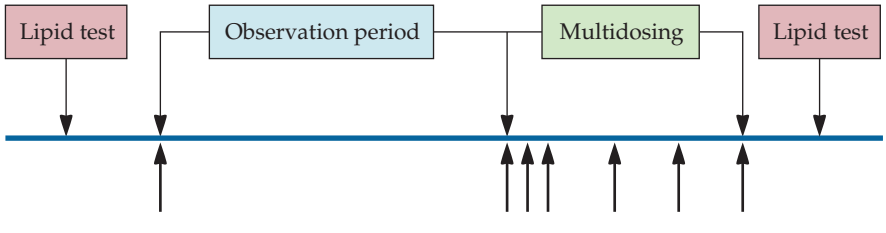
In one set of experiments, phosphoramidite antisense oligonucleotides were delivered by injecting muscle, followed by a short (less than a second)

electrical pulse. In this case, the antisense oligonucleotide formed highly stable duplexes with the target RNA and did not induce RNase H activity. Since only oligonucleotides targeted to the 5′ untranslated region, and not to the coding portion, of the mRNA were inhibitory, it was surmised that phosphoramidite antisense oligonucleotides prevented translation of the target mRNA.

Several preclinical trials have shown the usefulness of antisense oligonucleotides as therapeutic agents. For example, the narrowing (stenosis) of coronary and carotid arteries that leads to heart attacks and strokes, respectively, is often alleviated by angioplasty, which is a procedure that widens arteries by the insertion of an inflated balloon. However, arterial blockage recurs (restenosis) in about 40% of patients within 6 months because angioplasty induces a healing reaction, which stimulates the proliferation of smooth muscle cells and the secretion of an extracellular matrix in the inner layer of the artery at the site of the treatment. When phosphorothioate antisense oligonucleotides that targeted mRNAs for proteins that are essential for the mammalian cell cycle were applied to rat carotid arteries after angioplasty, restenosis was reduced by about 90%. In addition to postangioplasty restenosis, smooth muscle cell proliferation is implicated in atherosclerosis, hypertension, diabetes mellitus, and the failure of coronary bypass grafts. Presumably, these conditions might be controlled by similar antisense therapeutics.

In another study, a 20-nucleotide phosphorothioate antisense oligonucleotide complementary to the coding region of human apolipoprotein B was used to lower the level of low-density lipoprotein cholesterol in humans. High levels of low-density lipoprotein cholesterol have long been considered a significant risk factor for cardiovascular disease; high levels of apolipoprotein B are also likely associated with cardiovascular risk. Apolipoprotein B, which is produced in the liver, is an essential structural and receptor-binding component of all atherogenic (plaque-causing) lipoproteins. It plays a key role in low-density lipoprotein cholesterol transport and removal. Typically, statins, the most prescribed drug class in the world, are used to lower low-density lipoprotein cholesterol levels. While statins are effective for many individuals, some people continue to have high levels of both low-density lipoprotein cholesterol and apolipoprotein B. Therefore, a 20-nucleotide phosphorothioate antisense oligonucleotide complementary to the coding region of human apolipoprotein B was developed to be an adjunct to statin treatment. While this antisense oligonucleotide has been tested on only 36 individuals (Fig. 11.4), the results to date

FIGURE 11.4 Dosing regimen of an antisense oligonucleotide designed to lower low-density lipoprotein cholesterol. The entire procedure took approximately 10 weeks. The arrows below the horizontal line indicate the times when the antisense oligonucleotide was administered (subcutaneously). Doses of the antisense oligonucleotide ranged from 50 to 400 mg per injection.



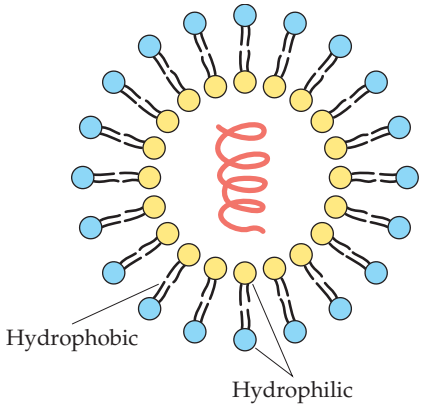


FIGURE 11.5 Schematic representation of a liposome carrying a nucleic acid.

are encouraging, and there is an expectation that this approach may lead to significant reductions in adverse cardiovascular events.

Antisense oligonucleotides have been tested to determine if they can control psoriasis, a disease of uncontrolled epidermal growth that causes red scaly, itchy patches to appear on the skin. Insulin-like growth factor I has been implicated in the pathogenesis of psoriasis because insulin-like growth factor I receptors are present in excess in psoriatic lesions. Thus, antisense oligonucleotide lowering of the mRNA for insulin-like growth factor I receptor might form the basis of a psoriasis therapy. In preliminary experiments, all of the different insulin-like growth factor receptor antisense oligonucleotides tested were 15 nucleotides in length. The potential antisense oligonucleotides were transfected into keratinocytes by using liposomes to facilitate antisense oligonucleotide cellular uptake (Fig. 11.5), and then the level of insulin-like growth factor I receptor mRNA was assessed. The three most active antisense oligonucleotides reduced the insulin-like growth factor I receptor protein concentration by 45 to 65%, while a random oligonucleotide had no effect on the amount of the protein. The selected antisense oligonucleotides were tested with athymic nude mice carrying human psoriatic lesional grafts. When the grafts were injected every 2 days for 20 days with antisense oligonucleotides complementary to the insulin-like growth factor I receptor mRNA, there was a significant reduction (58 to 69%) in both epidermis thickness and the cross-sectional area of the skin lesions. This result is very encouraging and suggests that skin diseases in which a normal protein is overproduced may be appropriate targets for antisense oligonucleotides that can be delivered topically.

Antisense oligonucleotides have also been used to inhibit the synthesis of the transcription factor forkhead box O1 (also called FOXO1) in mice. This protein increases the expression of phosphoenolpyruvate carboxykinase (PEPCK) and glucose-6-phosphatase (G6Pase), both of which are key enzymes in gluconeogenesis. Antisense oligonucleotides cause a reduction in the expression of these genes (approximately 50 to 60%) in both liver and fat tissues, but not in cardiac or skeletal muscle. Thus, the introduction of antisense oligonucleotides complementary to FOXO1 mRNA essentially mimics insulin action and therefore may bypass some of the defects in insulin signaling common among diabetics. To choose the most effective antisense oligonucleotide sequence, 80 different 20-nucleotide-long sequences complementary to various portions of mouse FOXO1 mRNA were tested, using mouse primary hepatocytes in culture, for the ability to inhibit FOXO1 mRNA expression. The positive effects that were reported represent the results of studies with the most effective of the original 80 oligonucleotides. It will now be of interest to determine whether the promising results observed in mice can be extended to humans.

Aberrant splicing of an mRNA occurs when a mutation in an intron is recognized by the RNA-processing system as an authentic splice site, and consequently, a portion of the intron is included as part of the processed mRNA (Fig. 11.6A). The presence of part of an intron disrupts the reading frame, and a truncated protein is produced. As a result, a disease condition may result from a diminished level of normal protein.

It was reasoned that an antisense oligonucleotide that targeted an aberrant splice site could likely prevent splicing at that site and increase the number of joining events between the correct intron–exon splice sequences. This notion was tested with a splice mutation in the second intron of the

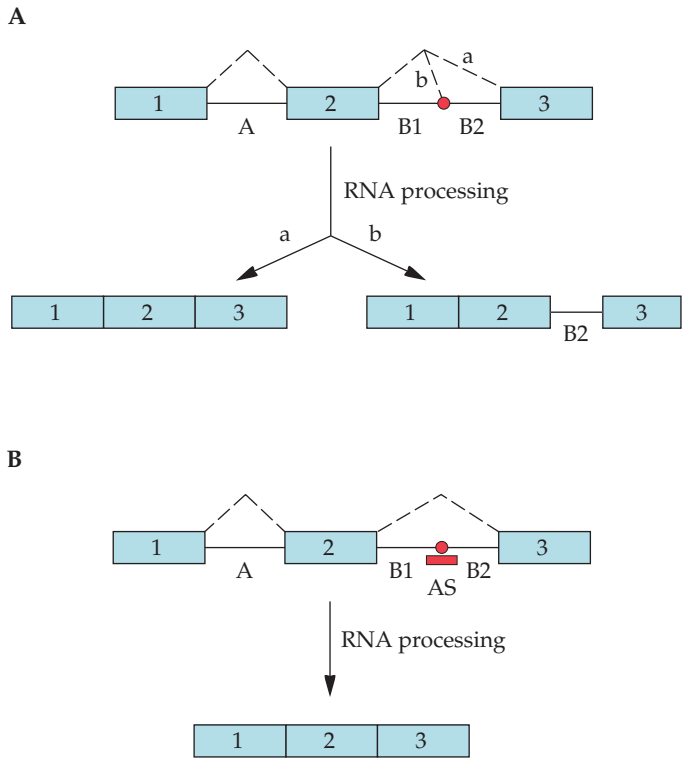


FIGURE 11.6 Correction of a mutant splice site with an antisense oligonucleotide. **(A)** The outcome of a mutation in a splice site is depicted. The numbers denote exons. The first intron is marked with the letter A, and the second intron contains a splice mutation (red dot) that divides the intron into two parts (B1 and B2). The dotted lines span the RNA segments that are removed during RNA processing. There are two possible splicing events: pathway a leads to a functional mRNA, and pathway b leads to an RNA that includes part of the second intron (B2). **(B)** An antisense oligonucleotide (AS) (shown as a red bar) that binds to the mutant splice site prevents RNA processing at the site, and consequently, only functional mRNA is produced.

β-globin gene (Fig. 11.6B). This mutation is responsible for one form of β-thalassemia, which is an inherited blood disorder that leads to loss of red blood cells (anemia). After cells that are homozygous for the intron 2 splice site mutation were transfected with a 2′-*O*-methyl phosphorothioate antisense oligonucleotide that targeted the mutant splice site, the number of normal β-globin chains was increased by about 50%, which theoretically would be beneficial to patients with this genetic defect. Further studies are required to determine if antisense rectification of splice site mutations is an effective therapeutic strategy for thalassemia and other conditions due to similar mutations.

It has recently been shown that it is possible to protect mice against retroviruses by injecting them, intravenously or intraperitoneally, with phosphorothioated antisense oligonucleotides that prevent the conversion of the viral RNA genome into double-stranded DNA. In this system, the added antisense oligonucleotide effectively blocks replication of the retrovirus. When an added antisense oligonucleotide binds to the junction of the polypurine tract (which is present in a broad range of different retroviruses) and the U3 element, a structure is formed which mimics the normal substrate for the virus-encoded enzyme RNase H (Fig. 11.7). This causes

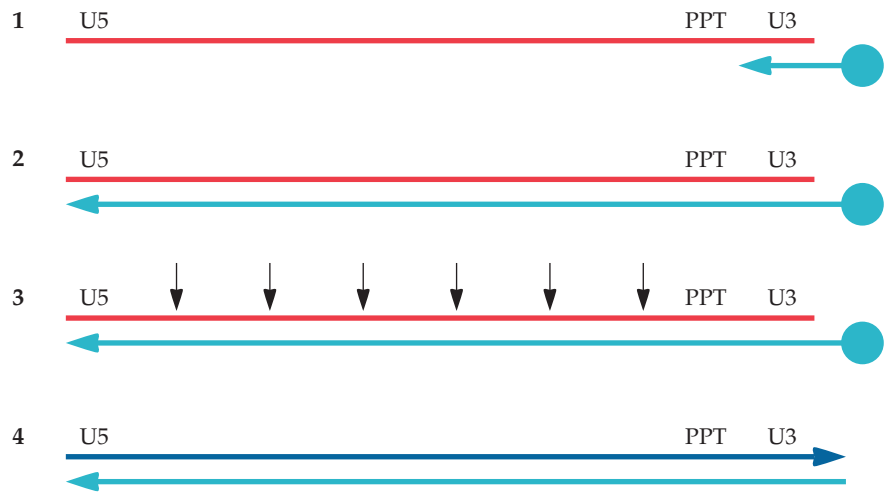


FIGURE 11.7 Schematic representation of the reverse transcription of retroviral RNA (red) to produce double-stranded DNA (the minus strand is light blue, and the plus strand is dark blue). In step 1, a transfer RNA (tRNA) (shown as a blue circle) primes synthesis of the minus strand. The RNA shows U3 and U5 elements and a polypurine tract (PPT). In step 2, a complete RNA–DNA duplex is formed. In step 3, RNase H digests the viral RNA in the RNA–DNA duplex into small pieces (the arrows indicate digestion sites). In step 4, a double-stranded DNA copy of the viral RNA is produced, which can exist as a provirus integrated into a cell's DNA. To block the formation of the minus strand and hence the synthesis of a double-stranded DNA version of the virus, an antisense oligonucleotide that hybridizes to the PPT region is added.

premature cleavage of the viral RNA, resulting in the virus being destroyed before reverse transcription occurs. This strategy has been shown to be effective in protecting mice from retroviruses. In principle, it should also work in humans and on a range of different retroviruses. However, a number of technical obstacles must be overcome before this approach is ready for testing in humans.

Ribozymes

Ribozymes are naturally occurring catalytic RNA molecules (RNA metalloenzymes) that are ~40 to 50 nucleotides in length and have separate catalytic and substrate-binding domains. Compared with protein therapeutics, an important advantage of ribozymes is that they are unlikely to evoke an immune response in a treated animal or human. The substrate-binding sequence combines by nucleotide complementarity and, possibly, non-hydrogen-bond interactions with its target sequence. The catalytic portion cleaves the target RNA at a specific site. By altering the substrate-binding domain, a ribozyme can be engineered to specifically cleave any mRNA sequence (Fig. 11.8). For therapeutic purposes, either hammerhead or hairpin ribozymes—named after the appearance of their secondary structure that results from intrastrand base pairing—may be used. However, some workers have suggested that hammerhead ribozymes are preferable because of their ability to more efficiently recognize, bind to, and cleave a range of different mRNAs.

In practice, an indirect strategy is often used for creating a therapeutic ribozyme, since the large-scale production of synthetic RNA molecules is

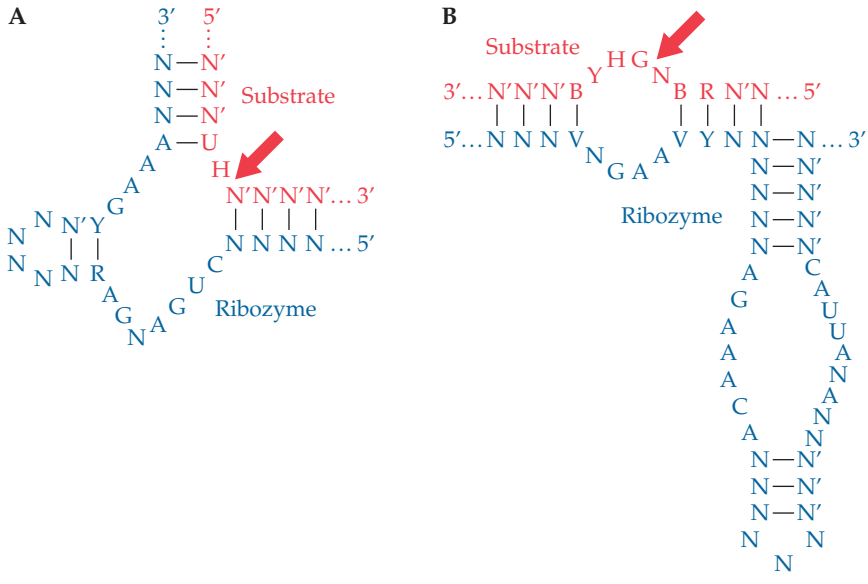


FIGURE 11.8 Two-dimensional representation of hammerhead **(A)** and hairpin **(B)** ribozyme–mRNA substrate complexes. The mRNA substrates and ribozymes are shown in red and blue, respectively. A, adenosine; C, cytosine; G, guanosine; U, uridine; Y, a pyrimidine nucleotide (C or U); R, a purine nucleotide (A or G); H, any nucleotide except G; B, any nucleotide except A; V, any nucleotide except U; N and N′, any complementary nucleotides. The arrows indicate the points of mRNA cleavage.

difficult and RNA molecules are susceptible to degradation after delivery to a target cell. One approach to overcome these drawbacks entails chemically synthesizing a double-stranded oligodeoxyribonucleotide with a ribozyme catalytic domain (~20 nucleotides) flanked by sequences that hybridize to the target mRNA after it is transcribed. The double-stranded form of the ribozyme oligodeoxyribonucleotide is cloned into a eukaryotic expression vector (usually a retrovirus). Cells are transfected with the construct, and the transcribed ribozyme cleaves the target mRNA, thereby suppressing the translation of the protein that is responsible for a disorder. Since most of the vectors that have been used cannot infect nondividing cells, target cells may be removed from a patient and then grown and transfected in culture before they are returned to the original tissue.

As an alternative to intracellular ribozyme production, ribozymes may be delivered directly to cells by injection or with liposomes, i.e., endogenous delivery. Directly delivered ribozymes may be chemically modified to protect them from rapid breakdown by nucleases. For example, the 2′ hydroxyl groups may be modified by alkylation or by substitution with either an amino group or a fluorine atom. These modifications increase the half-life of ribozymes in serum from minutes to days.

Under laboratory conditions, ribozymes can inhibit the expression of a variety of viral genes and significantly inhibit the proliferation of numerous organisms. For example, in cell culture, ribozymes inhibit the expression of (1) human cytomegalovirus transcriptional regulatory proteins, resulting in a 150-fold decrease in viral growth; (2) human herpes simplex virus type 1 transcriptional activator, resulting in a reduction of around 1,000-fold in viral growth; and (3) a reovirus mRNA encoding a protein required for viral proliferation. Moreover, a hammerhead ribozyme was designed to

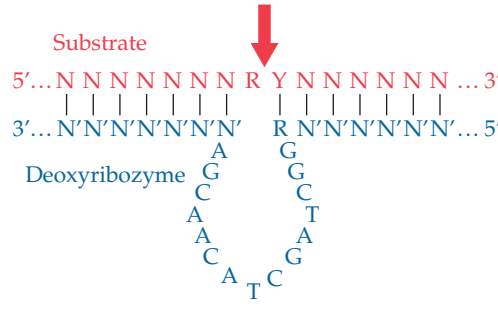
treat collagen-induced arthritis in mice. A plasmid-encoded ribozyme directed against the mRNA for tumor necrosis factor alpha, which is involved in rheumatoid arthritis, was intravenously injected into affected mice. Following this treatment, there was a significant reduction in the level of the tumor necrosis factor alpha mRNA, as well as a decrease in the collagen-induced arthritis. This type of ribozyme, used so far only with mice, has the potential to become a human therapeutic agent.

The development of resistance in humans to various chemical treatments is a persistent problem for the pharmaceutical industry. Generally, a single mutation that alters the target site is sufficient to void the action of a drug. However, with appropriate ribozyme-based therapeutics, ribozymes for a number of different sites could be used simultaneously, thereby cleaving an mRNA at different sites. The ability to cleave multiple sites on a single viral gene should make it less likely that any single viral mutation will confer resistance.

Deoxyribozymes

To date, no naturally occurring DNA equivalent of ribozymes, i.e., DNA enzymes (deoxyribozymes), has been discovered. However, oligodeoxynucleotides with catalytic activity have been synthesized. The best characterized and most studied of these deoxyribozymes is 10-23 RNase (Fig. 11.9). As a therapeutic agent, a catalytic oligodeoxynucleotide has some advantages over a ribozyme. DNA is approximately 1,000-fold more stable against hydrolytic destruction than protein and is nearly 100,000-fold more stable than RNA. In addition, deoxyribozymes are more efficient at binding and cutting mRNAs than are ribozymes. However, a deoxyribozyme cannot be produced continuously after the vector that encodes it is introduced into a particular tissue, because only ribozymes are produced from the DNA sequence. Therefore, deoxyribozymes must be delivered directly to affected cells. Proof-of-principle experiments have shown that deoxyribonucleotides are effective with cells in culture. For example, deoxyribozymes have been used to cleave mRNA transcribed from the growth-stimulating gene *myc*, which limits the growth of leukemia cells in culture. A deoxyribozyme has also been used to prevent mRNA from a gene called *Egr-1* from being expressed. Production of this mRNA is one reason for the failure of angioplasty—a procedure in which a balloon at the end of a catheter is used to

FIGURE 11.9 Two-dimensional representation of deoxyribozyme 10-23 RNase–mRNA substrate complex. The mRNA substrate and deoxyribozyme are shown in red and blue, respectively. A, adenosine; C, cytosine; G, guanosine; U, uridine; Y, a pyrimidine nucleotide (C or U); R, a purine nucleotide (A or G); N and N′, any complementary nucleotides. The arrow indicates the point of mRNA cleavage.



unclog arteries that contain atherosclerotic plaques. When the *Egr-1* mRNA is expressed, the recently unclogged artery is rapidly closed.

Chimeric RNA–DNA Molecules

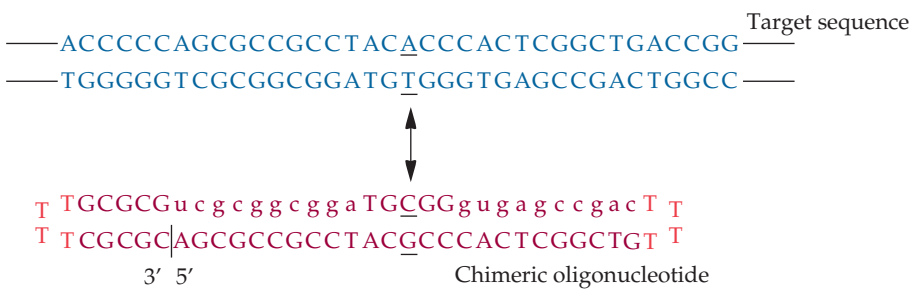
The ability to convert a mutant base pair of a gene to the wild-type (normal, or correct) version would reverse the consequences of many different genetic conditions. A strategy using a modified RNA–DNA oligonucleotide with 68 nucleotides (chimeric oligonucleotide) has been devised for this purpose. The composition of the chimeric oligonucleotide includes a single mixed oligonucleotide with ribonucleotides and deoxyribonucleotides in a duplex conformation with hairpin caps at the ends of the complementary segments and methylation of the oxygen of the 2′ carbon of the ribose sugars (Fig. 11.10). The rationale for this particular arrangement is based on various experimental observations. First, combined RNA–DNA molecules participate more readily than duplex DNA in homologous nucleic acid pairing reactions. Second, hairpin caps, which do not interfere with the pairing of homologous nucleic acid molecules, protect the molecule from exonucleases. Third, the 2′-O methylation of the ribose units shields the molecule from degradation by RNase H. In addition, the organization of the nucleotides of a chimeric oligonucleotide is important. Ten ribonucleotides flank a central core of five deoxyribonucleotides, and except for the correct base pair, this segment of the chimeric oligonucleotide has the same sequence as the target.

In cell culture, the feasibility of base pair correction with a chimeric oligonucleotide was examined with both a mutated cDNA sequence carried by a plasmid and a mutant site within a chromosomal sequence. In both instances, with high frequencies, the mutant sites were replaced by the correct base pair. However, more studies are required before chimeric oligonucleotides will become effective therapeutic agents.

Aptamers

Aptamers are nucleic acid sequences, RNA or DNA, that bind tightly to proteins, amino acids, drugs, or other molecules. They are typically 15 to 40

FIGURE 11.10 Correction of a single-base-pair mutation by a chimeric oligonucleotide. The double arrow points to the mutant site in the target sequence and the correct base pair in the chimeric oligonucleotide. The mutant and correct nucleotides are underlined. The uppercase letters represent deoxyribonucleotides, and the lowercase letters represent ribonucleotides. The nucleotides of the hairpin caps are shown in red. The vertical line and the 3′ and 5′ designations mark the 3′ and 5′ ends of the chimeric oligonucleotide. Adapted from Yoon et al., *Proc. Natl. Acad. Sci. USA* **93:**2071–2076, 1996.



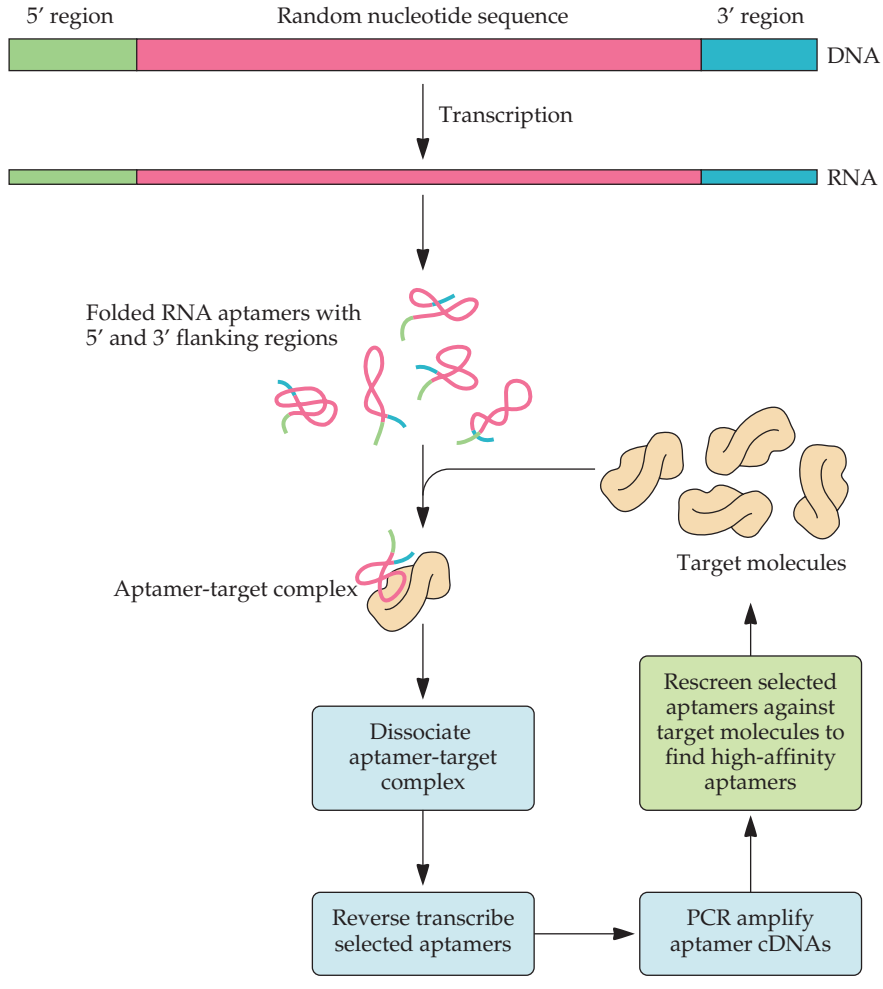


FIGURE 11.11 Overview of the SELEX procedure for selecting aptamers with high affinity to a target molecule (often a protein). The selected aptamers are typically cycled through this procedure 5 to 15 times.

nucleotides long, have highly organized secondary and tertiary structures, and bind with high affinity ($10^{-12} < K_d < 10^{-9}$, where K_d is the dissociation constant) to their target molecules. Aptamers are attractive as potential therapeutic agents because of their high specificity, relative ease of production, low or no immunogenicity, and long-term stability.

Aptamers that are directed against specific targets are typically selected by a procedure known as SELEX (*s*ystematic *e*volution of *l*igands by *ex*ponential enrichment), in which DNA or RNA ligands that bind to the target molecule are selectively enriched (Fig. 11.11). In this procedure, a random DNA sequence is cloned between two particular DNA sequences. The 3′ region contains an attachment site for reverse transcriptase primers, and the 5′ region contains an attachment site for a polymerase chain reaction (PCR) primer. The double-stranded DNA is converted to RNA using T7 RNA polymerase. The SELEX procedure combines several rounds of binding, partitioning, and amplification of selected nucleotide sequences from an initial pool of up to 10^{16} nucleotide sequence variants. The end

TABLE 11.1 Some proteins against which aptamers have been generated and the affinities of the aptamers for the proteins

Protein	K_d (nM)
Keratinocyte growth factor	0.0003
HIV type 1 reverse transcriptase	0.02
Transforming growth factor β1	0.03
P-Selectin	0.04
VEGF receptor	0.05
Platelet-derived growth factor	0.09
Immunoglobulin E	0.1
Extracellular signal-regulated kinase	0.2
CD4 antigen	0.5
HIV type 1 RNase H	0.5
Factor IXa	0.58
Angiogenin	0.7
Complement factor 5	1.0
Transforming growth factor β2	1.0
Secretory phospholipase A2	2.0
Thrombin	2.0
Angiopoietin 2	2.2
γ-Interferon	2.7
L-Selectin	3.0
Human neutrophil elastase	5.0
Tenascin C	5.0
Integrin	8.0
Hepatitis C virus NS3 protease	10.0
Factor VIIa	11.0
Yersinia pestis tyrosine phosphatase	18.0
Anti-insulin receptor antibody MA20	30.0
Trypanosoma cruzi cell adhesion receptor	172.0

result of this procedure is the selection of aptamers that bind to the target molecule with high affinity. Ultimately, the SELEX procedure yields one (or just a few) unique nucleic acid sequence(s) from the original mixture with high affinity for the target molecule. To make aptamers less sensitive to nuclease digestion, OH residues at the 2′ positions of purines may be replaced with 2′-*O*-methyl residues. In addition, aptamers may be capped at their 3′ end with a deoxythymidine residue. Table 11.1 lists some of the proteins against which aptamers have been generated, as well as the range of affinities of the aptamer for the target protein.

An aptamer known as pegaptanib received approval from the U.S. FDA in December of 2004 for use as a human therapeutic agent. Pegaptanib is a 30-nucleotide-long aptamer that targets vascular endothelial growth factor (VEGF) and binds to the protein with extremely high affinity (K_d = 0.05 nM). This secreted protein promotes the growth of new blood vessels by stimulating the endothelial cells that not only form the walls of the blood vessels, but also transport nutrients and oxygen to the tissues. When retinal pigment epithelial cells begin to senesce from lack of nutrition (ischemia), VEGF acts to stimulate the synthesis of new blood vessels (neovascularization).

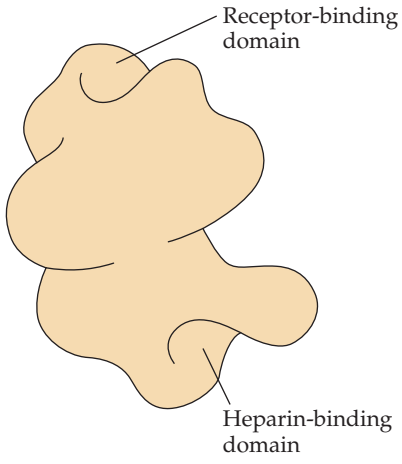


FIGURE 11.12 Schematic representation of protein VEGF, which contains both a receptor-binding domain and a heparin-binding domain.

However, this process is imperfect, and often the blood vessels do not form properly so that leakage results, causing scarring in the macular region of the retina with the eventual loss of central vision. These physiological changes contribute to age-related macular degeneration, a leading cause of blindness. Pegaptanib was selected to bind to one of the four isoforms of VEGF ($VEGF_{165}$) that is responsible for age-related macular degeneration. The drug is injected directly into the eye every 6 weeks, or about nine times a year. All four forms of VEGF have a receptor-binding domain (Fig. 11.12), while only $VEGF_{165}$ has a heparin-binding domain, which is the specific target for pegaptanib. Pegaptanib is a new type of therapeutic agent with an unusual specificity that can effectively suppress age-related macular degeneration—to date, approximately 95% of the patients receiving pegaptanib were 65 years of age or older.

The safety of aptamers used in clinical trials is a concern, especially when the optimal dose of a particular aptamer is not known. One way to overcome this problem is through the use of aptamer "antidotes." These molecules consist of short oligonucleotides whose sequences are complementary to the aptamers being tested. When antidotes are added to aptamers, they hybridize to the aptamers and inhibit their binding to the clinical target.

Interfering RNAs

Principles

The addition of double-stranded RNA to animal and plant cells reduces the expression of the gene from which the double-stranded RNA sequence is derived. This "gene silencing," which specifically reduces the concentration of a target mRNA by up to 90%, is reversible, since there is no change in the target cells' DNA. This phenomenon has been termed RNA interference (or "RNAi") and occurs naturally in virtually all eukaryotic organisms. RNAi appears to be the same phenomenon as gene silencing in animals or cosuppression in plants. Although all of its biological roles remain to be established, RNAi may protect both animals and plants from viruses and from the accumulation of transposons. A working model for RNAi has been formulated based on experimental analyses (Fig. 11.13). Following the introduction of a double-stranded RNA molecule into a cell, the double-stranded RNA is cleaved by the RNase III-like enzyme Dicer into single-stranded pieces of RNA, approximately 21 to 23 nucleotides in length, that have been called small interfering RNAs (siRNAs). The antisense strand of an siRNA is incorporated into an RNA-induced silencing complex (RISC) that binds to and then cleaves the mRNA. The specific binding of the siRNA to the mRNA that occurs is based on the complementarity of the two RNA sequences. The site of cleavage of the targeted mRNA is between nucleotides 10 and 11 relative to the 5′ end of the siRNA (antisense) guide strand. Consistent with this model, the transfection of mammalian cells in culture with duplexes of 21-nucleotide RNA can also mediate RNAi.

Despite the fact that many aspects of RNAi are still not completely understood, it could form the basis for new therapeutic agents. The use of short RNA duplexes may eventually provide an alternative approach to the use of antisense oligonucleotides or ribozymes.

The phenomenon of RNAi is expected to facilitate the development of a wide range of antiviral compounds and therapies utilizing specific siRNAs delivered to the appropriate target cell. Similarly, the expression of endogenous eukaryotic genes may be inhibited by plasmid-driven expression of short hairpin RNAs (shRNAs), which are similar in structure to the microRNAs that often normally regulate gene expression in eukaryotic cells. In fact, there are a number of reports of the use of either siRNA or shRNA to suppress virus replication in tissue culture (Table 11.2). Moreover, it has been successfully demonstrated that RNAi is effective in vivo (with mice), suggesting that, in principle, all viruses may be inactivated by RNAi.

Independently of how an siRNA or shRNA is introduced into a cell, it may have nonspecific effects. For example, introduction of these molecules may inadvertently activate innate cellular immune responses, such as the interferon response. In addition, siRNA or shRNA may also be complementary to nontarget mRNAs. However, several experimental approaches may be utilized to avoid these problems. (1) Off-target effects are most often observed when the siRNA or shRNA concentration is ≥100 nM. By lowering the concentration as much as possible (often to 20 nM or less), off-target

FIGURE 11.13 Overview of the process of RNA interference. Following introduction of double-stranded RNA (dsRNA) into a cell, the Dicer complex binds to the RNA and cleaves it into an siRNA containing approximately 21 bp. The antisense strand (red) becomes part of the RISC, directing the cleavage of the complementary mRNA.

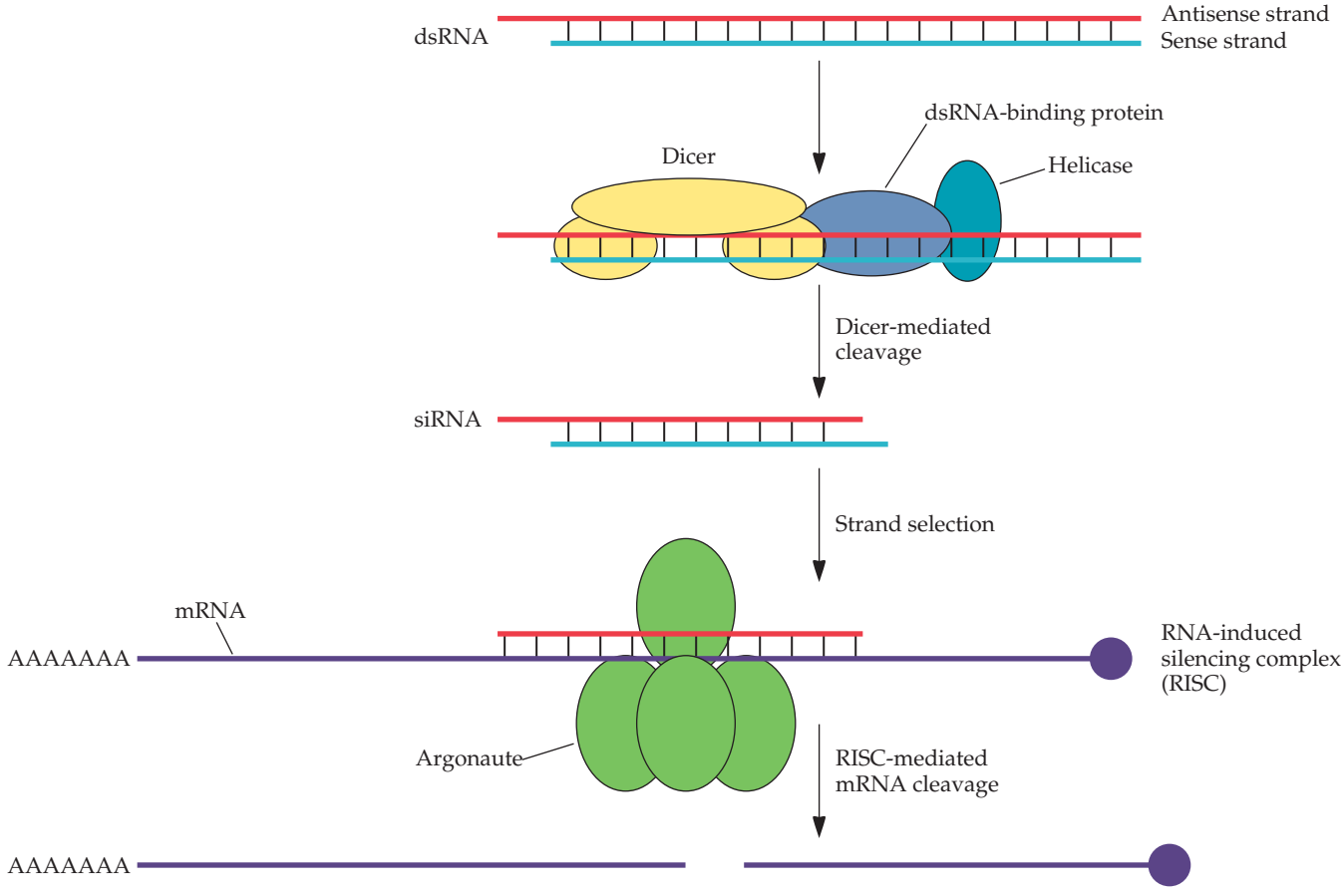


TABLE 11.2 Some examples of suppression of viral replication in tissue culture by RNAi

Severe acute respiratory syndrome-associated coronavirus
Hepatitis C virus
West Nile virus
Coxsackievirus B3
Foot-and-mouth disease virus
Hepatitis A virus
Human rhinovirus 6
Poliovirus
Respiratory syncytial virus
Human parainfluenza virus 3
Vesicular stomatitis virus
Influenza virus
Hepatitis delta virus
Rotavirus
HIV type 1
Hepatitis B virus
Herpes simplex virus type 1
Human cytomegalovirus
Epstein-Barr virus
Human herpesvirus 6B
Murine herpesvirus 68
Human papillomavirus type 18
JC virus

effects are often avoided. (2) Since the interferon response can be induced by double-stranded RNAs with as few as 11 bp that are perfectly complementary, siRNA or shRNA is designed to contain at least a 1-nucleotide bulge (where the bases on opposing strands are noncomplementary) near the center of the molecule (typically around 21 bp). (3) Since siRNA or shRNA can exert a toxic effect when it contains the sequence 5′-UGGC-3′, this sequence should be avoided. (4) Blunt-ended 27-bp RNA duplexes or 29-bp shRNAs with 2 nucleotides overhanging at the 3′ end are much more potent inducers of RNAi than 21-mer siRNAs. The greater level of effectiveness of the slightly longer RNAs may reflect the fact that they are first bound and cleaved by Dicer, which facilitates their entry into the RISC. Using these slightly longer RNAs at low concentrations should avoid side reactions associated with 21-mer siRNAs.

Unlike siRNAs, shRNAs are expressed in vivo as part of a genetic construct that includes a promoter sequence. This means that shRNAs need to be introduced by using strategies different than those used with siRNAs. Thus, shRNAs are typically delivered to their target cells by using viral vectors. Viral vectors that integrate into the chromosomal DNA are generally used when persistent long-term knockdown of gene expression is desired; the most popular choice is lentiviruses. However, with the use of all virus-based vectors, there are serious safety concerns that need to be addressed, and the type of promoter that is used also needs to be optimized.

Applications

Interfering RNAs have already found widespread use as tools in research that is directed toward understanding how gene expression is regulated in natural systems. One company that specializes in producing RNAi directed against human mRNAs advertises, "For each target (human) gene, we provide four plasmids each with a different short hairpin RNAi sequence (shRNA). Our experimentally verified design algorithm minimizes the risk of off target effects and ensures the maximum knock-down. At least one of the four shRNA plasmids will reduce the target mRNA levels in the transfected cells by >70%." With the ready availability of human shRNA libraries, new insights and understanding of many fundamental and disease processes should be rapidly forthcoming. It is hoped that this in turn will lead to a variety of new therapeutic agents and approaches.

In addition to the many reports of successful modification of the gene expression of cells in culture with RNAi, there are an increasing number of reports of the in vivo effectiveness of RNAi. By 2008, siRNAs against a wide range of proteins, viruses, and diseases had been successfully expressed in mice, including siRNAs directed against herpes simplex virus type 2, hepatitis B virus, hepatitis C virus, Huntington disease, metastatic Ewing sarcoma (a form of cancer), respiratory syncytial virus, hepatic cancer, transforming growth factor receptor 2, severe acute respiratory syndrome, heme oxygenase 1, keratinocyte-derived chemokine, tumor necrosis factor alpha, and human epidermal growth factor receptor 2. In addition, in 2008, three different RNAi therapeutics were being tested in clinical trials. All three of these therapeutics target VEGF or its receptor, a cause of age-related macular degeneration (see "Aptamers" above), and at that time, all had successfully completed either phase I or phase II clinical trials.

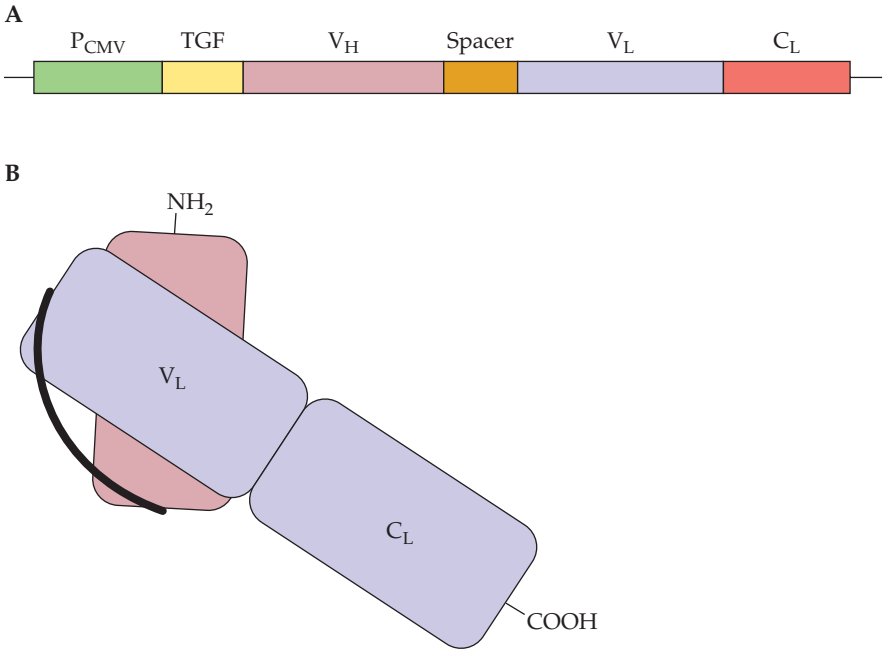


FIGURE 11.14 **(A)** DNA construct of a single-chain antibody against the hemorrhagic septicemia virus G protein. P_{CMV}, cytomegalovirus promoter; TGF, rainbow trout transforming growth factor beta signal peptide; V_H, variable domain from mouse hybridoma heavy chain; spacer, 42 nucleotides encoding a 14-amino-acid spacer connecting V_H and V_L; V_L, variable domain from mouse hybridoma light chain; C_L, constant part of the human light chain gene. **(B)** Schematic representation of the single-chain antibody against the hemorrhagic septicemia virus G protein. The curved black line represents the 14-amino-acid spacer.

Antibody Genes

A major problem with aquaculture is the loss of fish due to bacterial and viral infections. The chemical treatments that protect fish against these disease-causing pathogens are extremely costly and not particularly effective. One way to address this problem is to develop fish that synthesize protective antibodies against particular pathogens. The impetus for this approach came from the observation that rainbow trout could be protected against hemorrhagic septicemia virus by passive immunization through the injection of a monoclonal antibody against the G protein from the virus. Subsequently, a gene encoding a single-chain antibody directed against the hemorrhagic septicemia virus G protein was synthesized. The synthetic gene encoded variable regions from the mouse L and H chains of an anti-G protein monoclonal antibody with a human antibody constant domain fused to the 3′ end of the construct (Fig. 11.14). To mediate secretion in fish cells, the DNA was fused to the gene sequence encoding the signal sequence of rainbow trout transforming growth factor beta. The DNA construct was inserted into a eukaryotic expression vector under the transcriptional control of a constitutive cytomegalovirus promoter. The cloned gene–vector construct was injected into the circulatory system of rainbow trout. When the fish were challenged with the hemorrhagic septicemia virus 11 days after injection of the DNA construct, nearly all of them survived, whereas

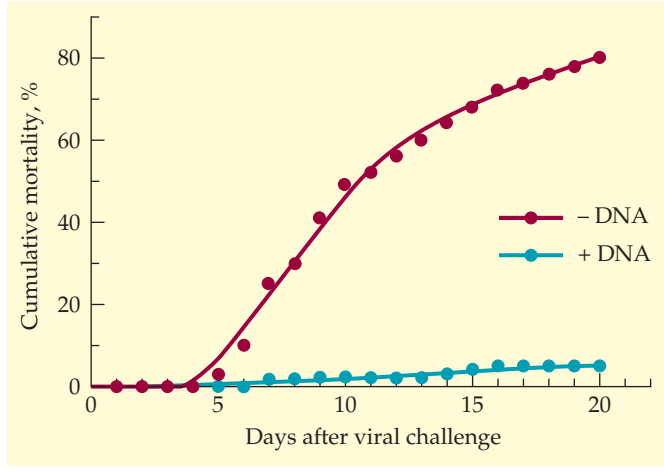


FIGURE 11.15 Survival of fish exposed to hemorrhagic septicemia virus 11 days after being injected with DNA encoding an antibody directed against the hemorrhagic septicemia virus G protein (+DNA) compared with control fish that did not receive this DNA construct (–DNA).

the control fish did not (Fig. 11.15). Thirty-nine days after the DNA injection, the plasma concentration of the single-chain antibody remained at a high level. Thus, instead of administering a purified antibody to provide an animal with passive immunity, it is possible to inject an animal with DNA encoding the antibody. Giving an animal the ability to synthesize a specific antibody may be more efficient than depending upon the animal's immune system to produce a similar antibody. However, since it may not be practical to inject large numbers of fish, the next step will be to create transgenic fish with antibody genes that confer passive immunity to various diseases.

Nucleic Acid Delivery

The ultimate effectiveness of any therapeutic agent depends upon the ability to deliver that agent to the tissues where it is required. Systemic introduction of a therapeutic agent often leads to the accumulation of very high levels in tissues where the agent is not required and sometimes results in serious side effects. To this end, viral vectors that deliver small nucleic acids to specific cellular targets have been developed. However, although virus-based gene delivery has been successful, a number of safety concerns have arisen in regard to the use of these vectors. Several approaches have recently been developed as an alternative to virus-based systems, as well as to the systemic introduction of target nucleic acids. There are several methods that have been used to deliver relatively small nucleic acids to animal cells. They include (1) intravenous injection; (2) local injection at the site of the pathology; (3) packaging into cationic liposomes (Fig. 11.5); (4) physical methods, like electroporation, sonoporation, or hydrodynamic pressure; and (5) a number of systems in which the nucleic acid is chemically conjugated to another molecule.

Human Gene Therapy

Most of the experience of managing genetic diseases has been gathered from inherited inborn errors of metabolism and, to a lesser extent, from

disorders with defective structural proteins. The strategies for treating genetic disorders include restrictive or supplemented diets, inhibition of enzyme reactions to prevent the accumulation of toxic molecules, removal of toxic molecules, replacement of defective or absent proteins, restoration of protein activity, selective protein removal, organ and bone marrow transplantation, and nucleic acid-based therapies (Table 11.3).

With the development of recombinant DNA technology, large numbers of enzymes and structural proteins are now available for therapeutic use. Infusions of β-glucosidase, β-galactosidase, α-L-iduronidase, and adenosine deaminase (ADA) in clinical trials have significantly reduced the adverse effects of Gaucher disease, Fabry disease, mucopolysaccharidosis I, and severe combined immunodeficiency disorder (SCID), respectively. This type of treatment, called enzyme replacement therapy, works well when either the enzyme or the structural protein (protein replacement therapy) is delivered to its biological site of action through the bloodstream.

Since the 1940s, when it was discovered that a gene from one strain of bacteria could be transferred to and expressed in another strain, researchers have contemplated the possibility that human genetic diseases might be cured in an analogous manner. Introduction of a normal gene into a cell with a defective gene ought to correct the disorder because the transferred

TABLE 11.3 Strategies for treating genetic disorders

Specially formulated diets
Restrictive diet to lower the intracellular level of a toxic molecule
Supplemented diet to replace a metabolic deficiency
Inhibition of enzyme reactions
Enzyme inhibitor to prevent the accumulation of a toxic molecule by blocking a step in a metabolic pathway that precedes the reaction with a defective enzyme
Removal of toxic molecules
Dialysis
Removal of excess cations (chelation)
Facilitation of excretion by binding a toxic molecule to a low-molecular-weight compound
Replacement of defective or missing product
Enzyme replacement therapy
Protein replacement therapy
Cofactor supplementation
Alteration of defective protein by small molecules
Restoration of partial protein function
Directed proteolytic degradation of defective protein
Transplantation
Replacement of a nonfunctional organ transplantation
Providing a required protein synthesized by blood cells (bone marrow transplantation)
Gene therapy
Rectification of a genetic defect with a functional gene
Nucleic acid therapy
Blocking translation of mRNA from a mutant gene with an oligonucleotide (antisense, ribozyme)
Correction of a gene mutation with an oligonucleotide

gene provides the required gene product. In theory, gene therapy should provide a persistent in vivo treatment, either in the tissue that is primarily affected by a gene mutation or in deficient cells that acquire a recombinant protein from distant cells that release the therapeutic protein into the circulatory system. Although gene therapy was intended as a cure for genetic disorders, gene products can also be used to treat cancers, infections, and various degenerative disorders.

In 1990, after exhaustive reviews by many different regulating panels in the United States, the first human gene therapy trial was initiated. Two young girls with ADA-deficient SCID received large doses of their own cells that had been engineered to carry a functional *ADA* gene. In both instances, the adverse symptoms were alleviated, indicating that this form of therapy is feasible. One of the patients has been free of SCID for more than 10 years, although she was regularly administered polyethylene glycol–ADA. The second ADA-deficient patient from the original trial and others from additional experiments have not had long-lasting cures. After the initial trials with gene therapy for ADA, a number of gene-based clinical protocols for various conditions were conducted. Unfortunately, these trials failed to establish the effectiveness of any of the treatments. Notwithstanding this lack of success, the trials provided a great deal of information about the methods of gene delivery, the duration of gene expression, and other technical features of gene therapy. Generally, despite the failure to correct a genetic disorder with an exogenous functional gene, this type of clinical research was considered safe. However, in September 1999, the attitude toward gene therapy dramatically changed. Jesse Gelsinger, a healthy 18-year-old with ornithine transcarbamylase deficiency, was given a large dose of a virus carrying the *OTC* gene as part of a clinical trial. Tragically, he died 4 days later of a massive immune response. In another trial, injection of a gene into the heart muscle of a patient with severe coronary artery disease was fatal. As a consequence, although these disastrous outcomes were not predictable, the rigorous requirements for human gene therapy experiments became even more stringent. Researchers, in addition, became disinclined to initiate new trials. However, in 2000, two infants with an X-linked form of SCID (SCID-X1) were successfully treated with the gene encoding the subunit (γc) that is part of various cytokine receptors. These patients were free of symptoms for 10 months and are being monitored to determine if the correction is permanent; of the 11 patients, 4 developed cancer. Also, hemophilic patients expressed an input gene encoding the blood coagulation factor IX for long periods and, importantly, enough of the protein was produced to reduce the extent of the condition from severe to mild.

Although in the broadest sense, the concept of human somatic cell gene therapy is straightforward, there are a number of critical biological considerations. For example, how will the cells that are to be targeted for correction be accessed? How will the therapeutic (remedial) gene be delivered? What proportion of the target cells must acquire the input gene to counteract the disease? Does transcription of the input gene need to be precisely regulated to be effective? Will overexpression of the input gene cause alternative physiological problems? Will the cells with the input gene be maintained indefinitely, or will repeated treatments be required?

Much of the research effort in human gene therapy has been directed to developing efficient and nonimmunogenic systems that deliver a therapeutic gene to a specified cell type. Both viral and nonviral strategies have

been examined in detail. Different types of viruses, including retroviruses, adenovirus, adeno-associated virus, herpes simplex virus, and vaccinia virus, have been engineered as gene transfer vectors. A virus-mediated gene delivery system uses the cell receptor recognition system of the virus for binding to a specific cell type, which is followed by the internalization of the vector DNA and its transport to the cell nucleus. Some viruses have mechanisms for integration of the vector DNA into a chromosome site, and with others, the input DNA is maintained as an extrachromosomal element. Since high efficiency of vector DNA transfer is important, the administered sample of viruses should contain mostly vector viruses with very few, and preferably no, infectious virus particles. To meet this goal, packaging cell lines were constructed for some virus vectors. These cultured cells carry genes that express viral proteins that are necessary for the formation of virus particles but are not capable of producing replication-competent (infectious) viruses. After transfection of a packaging cell line with a vector construct that is equivalent to the length of the wild-type virus genome and that carries the appropriate DNA sequence for packaging into a virus particle, the input DNA is replicated, assembled into viruses, and released into the cell medium. The noninfectious vector viruses are concentrated and prepared for use. Packaging cell lines have been devised for retrovirus vectors (Fig. 11.16) and other viral delivery systems. In some cases, disarmed viral and vector DNAs are cotransfected and only viruses with vector DNA are produced. In other instances, protocols allow the formation of both vector and infectious viruses with the separation of the two types of viruses before use. There are advantages and disadvantages to each of the major viral gene delivery systems. Some vector viruses transduce DNA at high efficiencies but the size of the insert is limited. Retrovirus vectors infect only dividing cells, which, without genetic modification, makes them ineligible for treating disorders of nondividing cells. Some viruses lack cell specificity. However, this particular shortcoming has been countered by designing viruses with cell-specific receptor-binding sites. Many vector viruses are immunogenic, which nullifies repeated treatment with the same viral strain. To overcome this problem, vector viruses with different antigenic determinants may be used for successive treatments. Currently, third- and fourth-generation varieties of vector viruses are being developed with distinctive features for specific illnesses.

Some successful gene therapy experiments using viral vectors have been reported. In phase I clinical trials, a small number of patients have been treated with a recombinant adeno-associated virus vector carrying a human cDNA encoding retinal-pigment epithelium-specific 65-kDa protein (RPE65). The protein encoded by the *RPE65* gene is an important part of the visual cycle, forming part of a pathway that regenerates the visual pigment after exposure to light. Individuals who lack this protein become deficient in 11-*cis* retinal, and their rod photoreceptors are unable to respond to light, eventually leading to blindness. Prior to undertaking these experiments, a strain of Baird dogs with a spontaneous defect in the *RPE65* gene was successfully treated, with a restoration of visual function in the treated dogs. In all instances, the recombinant viral vector carrying the restorative gene was introduced by surgery of the eye into the subretinal space. Although the phase I human experiments were aimed only at determining the safety of this treatment, an improvement in the vision of four of the six patients was observed. This encouraging result notwithstanding, a longer follow-up

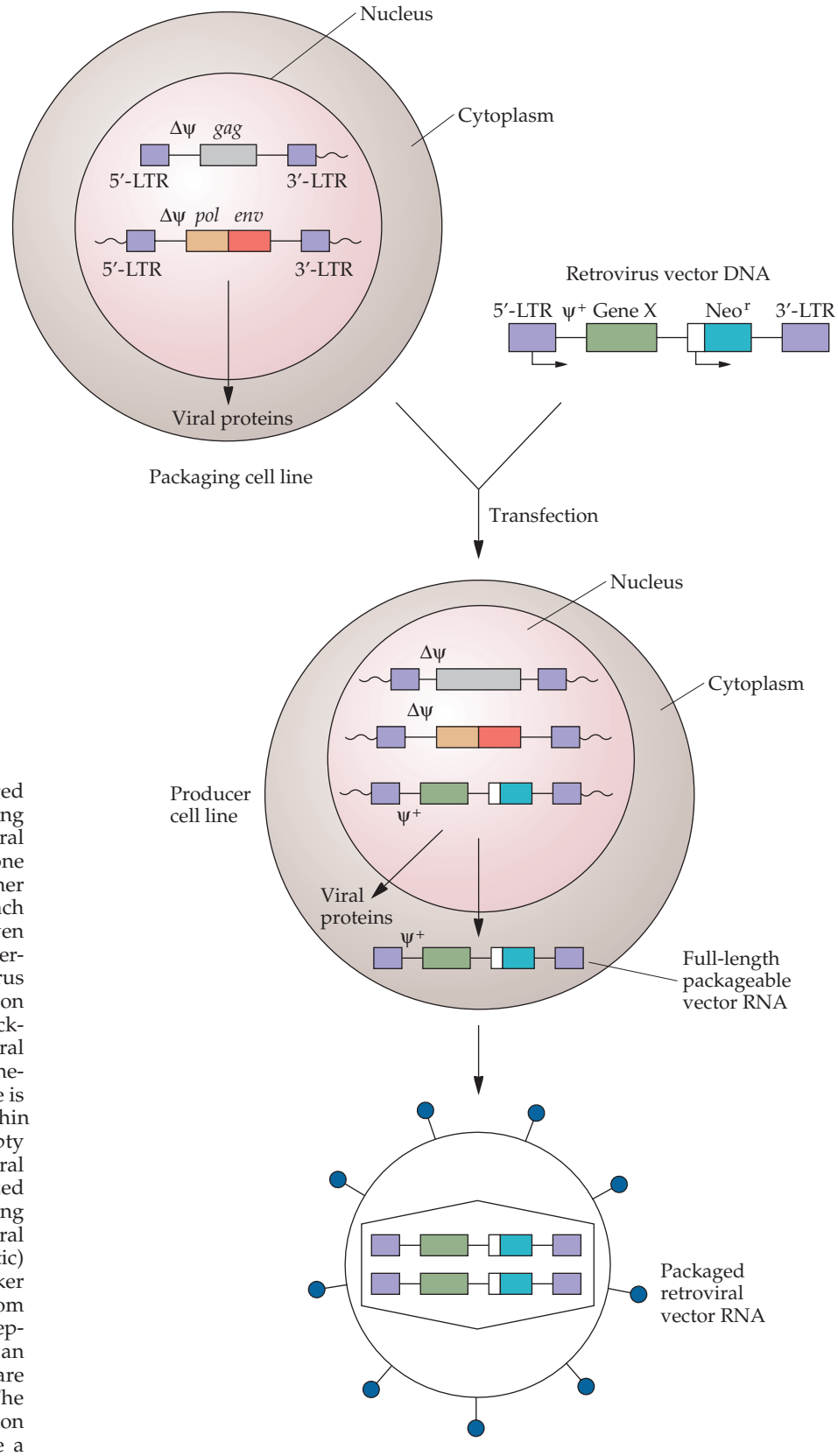


FIGURE 11.16 Production of packaged retrovirus vector RNA. The packaging cell line has two separate retroviral gene regions on its chromosomes; one contains the *gag* gene, and the other contains the *pol* and *env* genes. In each of these inserts, transcription is driven by sequences within the 5′ long terminal repeat (5′-LTR) region. Both virus DNA segments lack the encapsidation sequence (ψ^+) that is required for packaging a retroviral genome into a viral capsid. The packaging cell line synthesizes viral proteins, but because there is no encapsi dation sequence within either of the retroviral mRNAs, empty viral capsids are produced. The viral proteins continue to be synthesized after the transfection of a packaging cell line with a full-length retroviral vector carrying a remedial (therapeutic) gene (Gene X) and a selectable marker gene (Neo^r). The full-length RNAs from the retrovirus vector sequence are replicated, and because they have an encapsidation region (ψ^+), they are packaged into viral capsids. The released viral particles are replication defective because they do not have a *pol* gene.

period and clinical trials with many more subjects are required before this procedure can gain FDA approval. However, on the positive side, if the procedure is used to treat younger patients with less severe disease, greater improvements of visual function might be expected.

Notwithstanding the advantages of viral vectors, they are often immunogenic, costly to maintain, and difficult to produce on a large scale without high-level expertise. Consequently, various nonviral gene transfer systems have been devised. The least complicated nonviral gene delivery system is the introduction of pure (naked) DNA constructs directly into the cells of a target tissue. When plasmid DNA was injected into mouse skeletal muscle, some of the cells took up the DNA and a reporter gene was expressed for more than 50 days. However, this approach is limited to accessible tissues and requires large amounts of DNA. Pure DNA constructs that cover the surfaces of 1- to 3-μm-diameter gold particles have been propelled with a gene gun into skin cells (see chapter 18) and into subcutaneous tumor cells. Therapeutic genes delivered in this way were expressed in the targeted tissues. Surrounding a DNA construct with artificial lipid layers that form a lipid sphere with an aqueous core (liposome) facilitates the passage of the DNA through a cell membrane.

To avoid degradation of introduced DNA, DNA–molecular conjugates have been developed. With this approach, poly-L-lysine is chemically linked (conjugated) to a molecule that binds to a specific cell receptor. Next, DNA is added and combines with the poly-L-lysine to form a tightly compacted, twisted, solid ring. With the cell receptor-binding sites arrayed on the outside of the DNA–molecular conjugate, the complexes bound exclusively to the specified cells, but the frequency of transfection was low. To remedy this problem, in addition to a cell receptor-binding amino acid sequence, other short amino acid sequences that facilitate cell membrane fusion and internalization of the DNA–molecular conjugate and both protect the DNA from degradation and direct it to the cell nucleus have been combined into a single polypeptide. The addition of such a multifunctional protein to a DNA–molecular conjugate (Fig. 11.17) could enhance the efficiency of transfection. The current nonviral gene delivery systems have two major limitations: (1) the frequency of transfection is often too low to create a therapeutic effect and (2) the duration of therapeutic gene expression is too brief to provide an effective treatment.

A human artificial chromosome (HAC) would be an exciting therapeutic vector. The DNA-carrying capacity would be very large, which would allow the inclusion of several genes, each with a complete set of regulatory elements. This type of vector should have long-term stability and sustained expression of a therapeutic gene(s) within either a proliferating or a quiescent target cell. HACs (also called human engineered chromosomes) have been created in two ways (Fig. 11.18). First, HACs were assembled by ligation of individual chromosome components, including the chromosome ends (telomeres), centromere, and origins of replication. Telomere and centromere sequences were mixed with high-molecular-weight human DNA that had both origins of replication and a selectable gene marker in the presence of ligase. Cells were transfected with the DNA from the ligation mixture, and those with HACs were selected and maintained. A second method of forming a HAC entails paring down an existing human chromosome by deleting material from within each chromosome arm to form a "minichromosome." HACs have been formed that range in size from 0.7 to 400 megabases. However, before HACs are used for gene

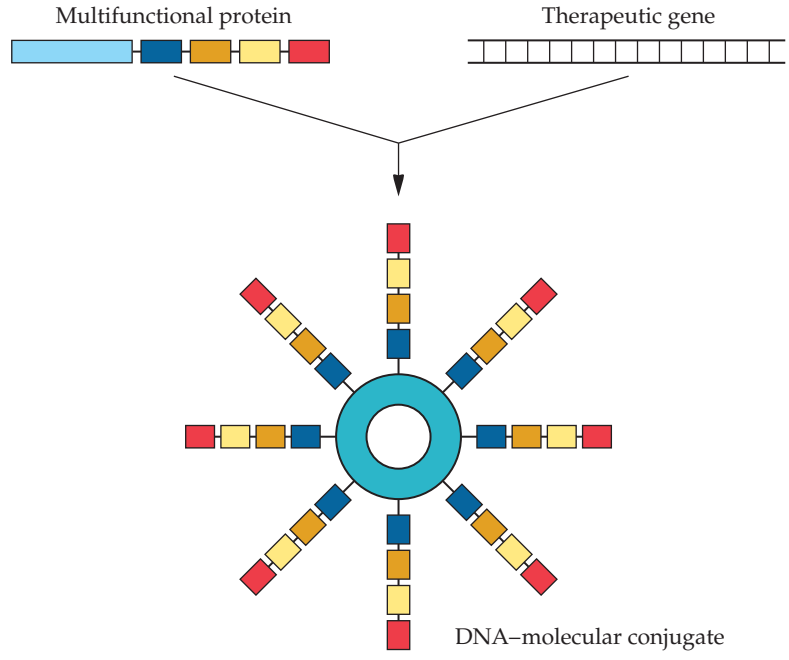


FIGURE 11.17 Schematic representation of a DNA–molecular conjugate delivery system. Short peptide sequences (motifs) that facilitate cell-specific binding (red), fusion with the cell membrane and internalization (yellow), protection of the DNA by preventing it from being routed to a lysosome (orange), and entry into the nucleus (dark blue) are attached to poly-L-lysine (light blue). The poly-L-lysine component combines with the DNA containing a therapeutic gene to form a condensed DNA–polylysine ring (medium blue) with the protein motifs facing outward.

therapy, certain issues must be addressed. For example, will HACs be efficiently introduced into the nuclei of target cells? Will effective levels of therapeutic gene expression be maintained for extended periods of time?

Before therapeutic genes are introduced in human beings, the efficacy of using a particular gene along with a specific delivery system is tested on small animals, typically mice. This is intended to ensure not only that the added gene relieves a particular ailment, but also that there are no unexpected side effects that occur as a consequence of the treatment. Recently, researchers reported the administration of a gene to dystrophic and normal mice that helped them to increase both muscle biomass and strength. The growth factor myostatin plays a critical role in regulating skeletal muscle mass. It negatively regulates both the number of myofibers formed in development and the postnatal growth of muscles. It was previously suggested that a number of neuromuscular disorders, including muscular dystrophies and age-related muscle disorders, might be "treated" using gene therapy approaches that prevent or lessen the inhibition of muscle growth by myostatin (Fig. 11.19). This could be achieved by the knockout of myostatin gene expression or by the overexpression of insulin-like growth factor 1, which can increase muscle size and strength. To this end, transgenic mice were created, from both dystrophic and normal mice, by a single postnatal intramuscular injection of adeno-associated virus that resulted in the overexpression of a gene encoding the myostatin inhibitor protein follistatin (Fig. 11.19). This single treatment enhanced muscle mass and strength in normal and dystrophic mice for

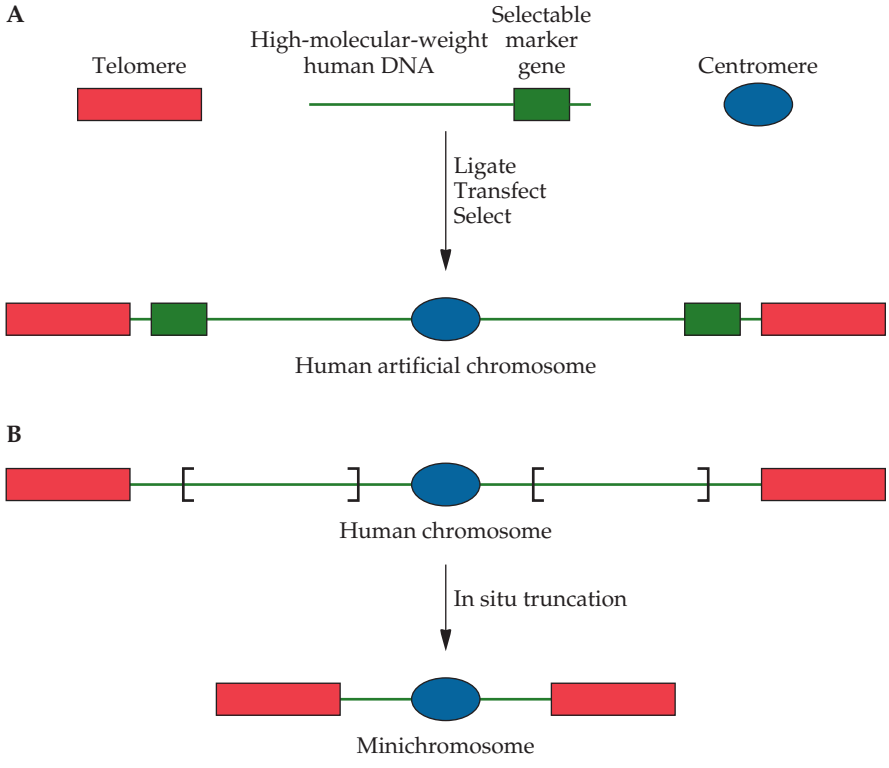


FIGURE 11.18 Construction of human artificial chromosomes (HACs). **(A)** Formation of a HAC in vitro by ligation. Chromosome elements (telomere [red] and centromere [blue]) and high-molecular-weight human DNA (green) that has origins of replication and a selectable marker gene (green rectangle) are placed in a ligation reaction mixture. After ligation, the DNA is introduced into cells in culture, and the selectable marker system is used to identify the cells with a functional HAC. The ligation products that do not form a complete chromosome are not maintained through successive cell divisions. **(B)** Formation of a HAC by truncation of a chromosome in situ. The square brackets denote DNA sections that are removed from the chromosome arms to produce a truncated chromosome (minichromosome). Chromosome DNA can be deleted by radiation of human cells and recovered as a minichromosome in a cell hybrid after fusion of rodent and irradiated human cells.

more than 2 years. This therapeutic strategy warrants serious consideration for clinical trials in the treatment of human muscle diseases. One potential concern, other than the safety of the viral vector, is that gene therapy approaches that are intended to treat muscle diseases by increasing muscle mass and strength might also be used for "gene doping" of healthy individuals to enhance athletic performance.

Targeting Systems

Lipids. While the effectiveness of siRNAs for specifically inhibiting gene expression in cultured cells has been demonstrated on numerous occasions, it is difficult to efficiently deliver these RNAs to tissues in vivo. One approach to overcome this difficulty has been to chemically couple an siRNA (at the terminal hydroxyl group of the sense strand RNA) to cholesterol (Fig. 11.20). The siRNA in question is complementary to an mRNA that encodes apolipoprotein B, a molecule involved in the metabolism of

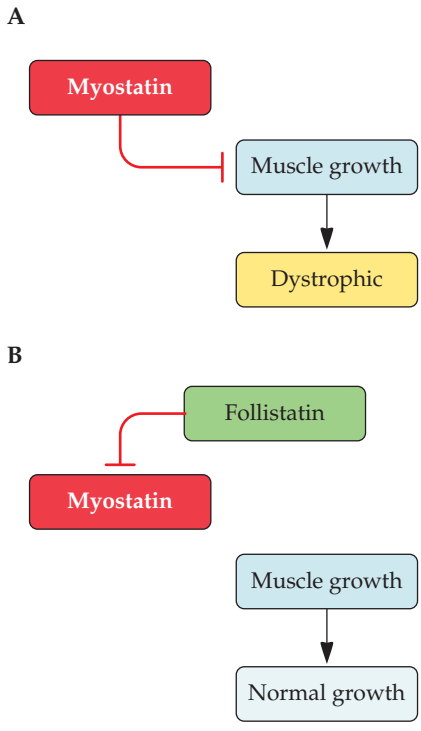


FIGURE 11.19 Schematic representation of the regulation of muscle growth and development by myostatin in nontransgenic mice **(A)** and in transgenic mice **(B)** that overexpress the protein follistatin.

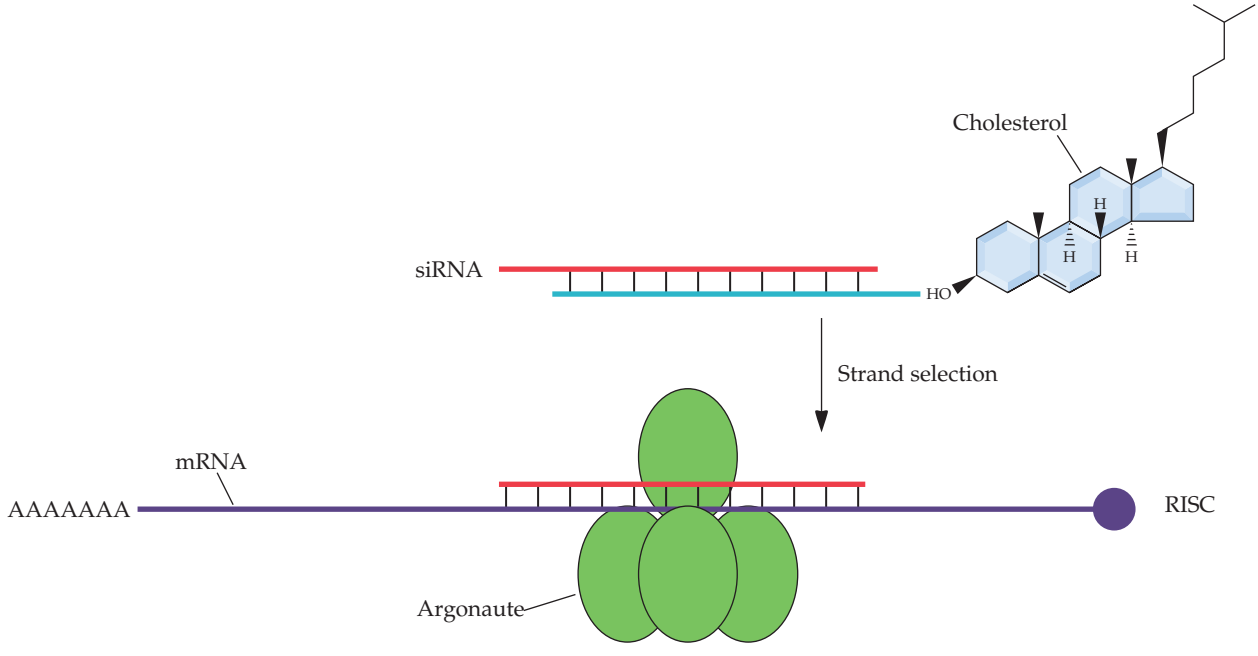
cholesterol. When the siRNA–cholesterol complex is intravenously injected into mice, it is taken up by the liver, jejunum (part of the small intestine), heart, kidney, lungs, and fat tissue cells. Once this complex is inside the tissue, the sense strand is destroyed and the antisense strand binds to the target mRNA.

With this approach, the level of apolipoprotein B was reduced by more than 50% in the liver and by 70% in the jejunum. This resulted in a significant decrease in the plasma apolipoprotein B level, as well as the total amount of cholesterol. This strategy is an important first step in the development of a method to therapeutically lower cholesterol levels in humans.

A number of other molecules, including some long-chain fatty acids and bile acids, may be used in place of cholesterol to mediate the uptake of siRNAs into cells. A critical factor in mediating the interaction between fatty acid-conjugated siRNAs and lipoprotein particles is the length of the fatty acid alkyl chain. Thus, docosanyl (C_{22}) and stearoyl (C_{18}) conjugates bind more tightly to high-density lipoprotein and subsequently silence gene expression more effectively in vivo than lauroyl (C_{12}) and myristoyl (C_{14}) conjugates. Studies are under way to improve the delivery of lipid-conjugated siRNAs to treat a wide range of diseases.

Bacteria. Bacteria that are normally found in association with various mammalian tissues and cells may be genetically engineered to produce therapeutic shRNAs. The engineered bacteria may then be used as vectors to deliver the therapeutic agent directly to the affected tissues. For example, a nonpathogenic strain of *Escherichia coli* was transformed with the plasmid vector TRIP containing the gene for the protein invasin, which permits

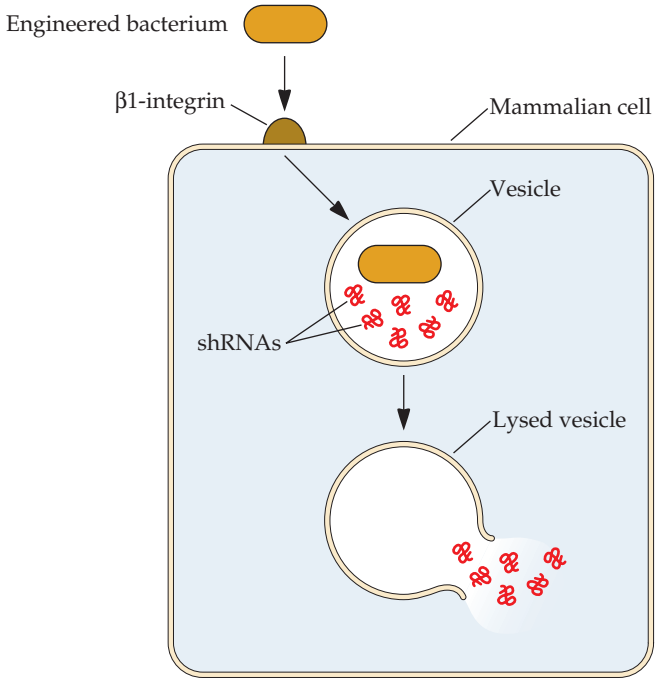
FIGURE 11.20 A conjugate of cholesterol and an siRNA in which the cholesterol is coupled through the 5′-OH of the sense strand of the siRNA. The cholesterol facilitates uptake of the siRNA into specific tissues. The antisense strand becomes part of the RISC and specifies where the mRNA is to be cleaved.



E. coli to enter β1-integrin-positive mammalian cells, and the gene *HlyA*, which encodes listeriolysin O, a protein that enables genetic material to escape from entry vesicles (Fig. 11.21). In addition, the TRIP vector carries an shRNA molecule under the control of a bacterial promoter directed against the mRNA produced by the cancer gene *CTNNB1*. As long as a bacterium is able to enter target mammalian cells and release shRNAs, the shRNA may be directed against any specific mRNAs. The *E. coli* cells act as a vector to transport the shRNAs to where they are required, e.g., cancer cells. This approach has been shown to work both for cancer cells in culture and with mice. With whole animals, the bacteria can be administered orally.

Collagen. The protein polymer collagen, isolated from calf dermis, can be digested, under acidic conditions, by the proteolytic enzyme pepsin to form subunits of approximately 300 kDa each. These rod-like proteins (approximately 300 nm in length by 1.5 nm in diameter) are positively charged and therefore readily interact with and bind to negatively charged siRNAs (Fig. 11.22). These "atelocollagen" particles protect siRNAs from nuclease digestion and also can be injected locally for tissue-targeting delivery of the siRNAs. For example, siRNA–atelocollagen complexes have been efficiently delivered to tumor cells in mice and, after injection, can exist in an intact form for at least 3 days. Furthermore, in mice, siRNA–atelocollagen complexes have been found to inhibit tumor growth in bone cells. This method of packaging siRNA promises to be both reliable and safe, depending on the tissue involved.

FIGURE 11.21 Use of a nonpathogenic strain of *E. coli* to deliver siRNAs to certain tissues. The bacterium was engineered to produce the protein invasin, which permits *E. coli* to enter β1-integrin-positive mammalian cells, as well as the gene *HlyA*, encoding listeriolysin O, which permits the shRNAs synthesized by the bacterium to be released inside the mammalian cell.



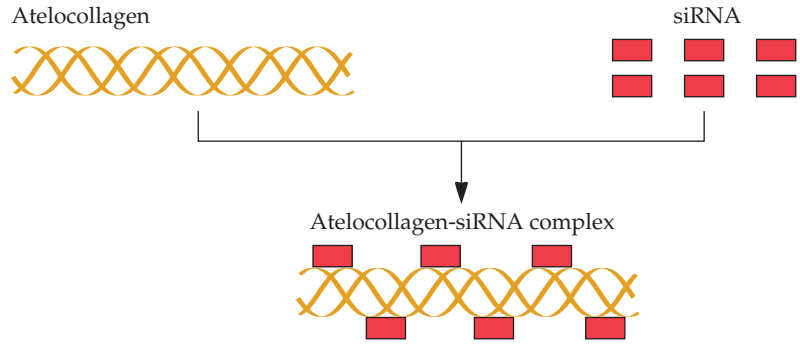
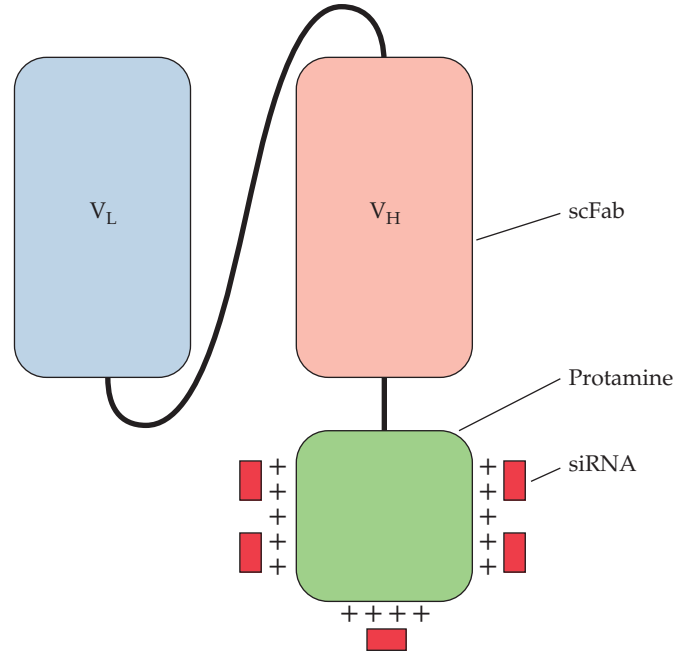


FIGURE 11.22 Negatively charged siRNAs bind to positively charged atelocollagen. The complex greatly facilitates the delivery of siRNAs to specific tissues.

Antibodies. It has become relatively straightforward to generate monoclonal antibodies against nearly any target protein; to humanize those antibodies, or their variable regions; and then to produce them in heterologous host cells. Moreover, at the DNA level, it is easy to fuse the antibody gene with the gene for another protein. With this in mind, the gene encoding single-chain Fab fragments that bind specifically to a protein called ErbB2, which is found on the surfaces of breast cancer cells, was fused to the gene for the positively charged nucleic acid-binding protein protamine. The fusion protein binds to the surfaces of cells expressing ErbB2 and at its C terminus carries the 51-amino-acid-long protamine, which readily binds to added siRNAs (Fig. 11.23). In one test of this system,

FIGURE 11.23 A single-chain Fab fragment directed against a mammalian cell surface protein is fused to the positively charged polypeptide protamine, which binds noncovalently to negatively charged siRNAs. The Fab fragment acts to deliver the siRNA to specific cells. Note that a conventional (two-chain) Fab fragment has also been used to deliver siRNAs.



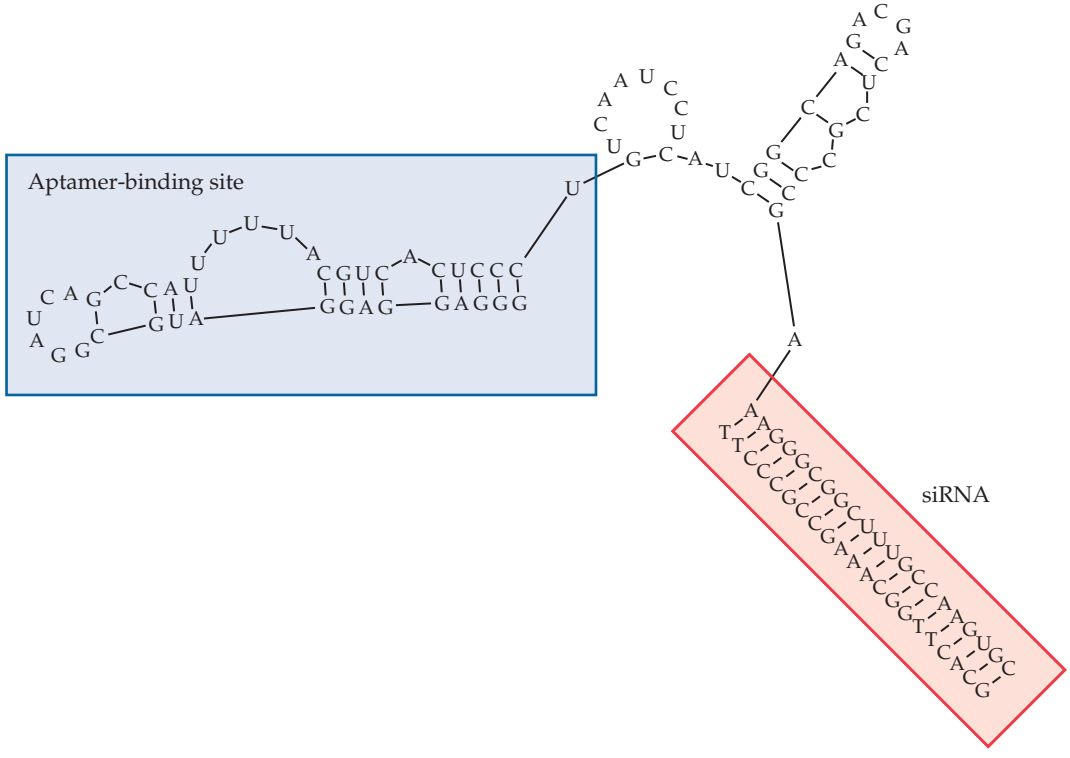


FIGURE 11.24 Schematic representation of the secondary structure of a chimeric RNA molecule consisting of an aptamer and an siRNA. The portion of the aptamer that binds to the target protein and the siRNA portion of the molecule are shaded.

an anti-human immunodeficiency virus (HIV) envelope Fab and an siRNA that is designed to cleave the HIV *gag* mRNA were employed. Using cells in culture, it was possible to reduce the amount of secreted Gag protein (the protein of the nucleocapsid shell around the RNA of a retrovirus) by >70%. This system also works in vivo when the construct is injected either intravenously or directly into tumors. The hope is that by using a combination of specific antibodies (or antibody fragments) that direct siRNAs only to certain cells, and siRNAs that selectively cleave specific target mRNAs, this system can be used to treat a wide range of diseases.

Aptamers. The binding specificity to a target antigen that is a central feature of the functioning of antibodies is also a property of aptamers. Thus, conjugating aptamers, which bind to specific cell surface proteins, to siRNAs that are designed to reduce the expression of certain mRNAs should provide another method of targeting siRNAs to specific tissues or cells. Also, since both aptamers and siRNAs are chemically synthesized RNA oligonucleotides, it should be simple and straightforward to synthesize chimeric RNA molecules that include both the binding specificity of an aptamer and an siRNA that targets a specific mRNA (Fig. 11.24). For example, an aptamer that binds selectively to a prostate-specific membrane antigen (found on prostate cancer cells) was first selected. Then, a 21-bp siRNA directed against mRNAs encoded by either of two genes that are necessary for prostate cells to survive was added to the aptamer sequence. Both activities (i.e., aptamer binding and RNAi) were maintained in the

chimeric molecule. The targeting aptamer did not impair the ability of the siRNA to silence the target gene, and the presence of the siRNA did not affect the ability of the aptamer to bind to its target. This simple but highly effective approach should be amenable to treating a wide range of human diseases provided that (1) silencing specific genes in a population produces therapeutic benefits and (2) there are surface receptors (usually proteins) that distinguish the target cell population and allow the siRNA to be internalized by the cell.

SUMMARY

A number of human disorders that result from the overproduction of a normal protein may be treated by using (1) nucleotide sequences that bind to a specific mRNA and prevent its translation, i.e., an antisense oligonucleotide; (2) RNA sequences that bind and cleave specific RNA molecules, i.e., ribozymes; (3) small RNA molecules, i.e., aptamers, that assume a highly organized secondary and tertiary structure and bind tightly to a wide range of molecules, including proteins, amino acids, and drugs; or (4) small double-stranded RNA molecules that direct the sequence-specific degradation of mRNA, i.e., interfering RNAs. These techniques may also be used to lessen or prevent diseases caused by pathogenic viruses and other disease-causing organisms.

The greatest impediment to the development of nucleic acid-based therapeutic agents is the difficulty in delivering these agents to their target tissue(s). Initially, workers used virus-based delivery systems with some success, although some safety concerns exist in regard to the use of these vectors. Other approaches for the delivery of nucleic acid-based therapeutic agents include intravenous injection; local injection at the site of the pathology; packaging the nucleic acid into cationic liposomes; physical methods, like electroporation, sonoporation, or hydrodynamic pressure; and conjugating the nucleic acid to another molecule, such as a lipid molecule, cholesterol, collagen, an antibody fragment, or an aptamer.

The development of effective treatments for genetic diseases has been elusive because, in many instances, the appropriate gene product cannot be provided to a patient. However, when a normal version of a gene has been identified and cloned, it may be possible that either it or a cDNA derivative can be used to correct the defect in affected individuals. Viral and nonviral systems have been developed for the delivery of therapeutic genes. Viral vectors take advantage of the ability of a virus to penetrate a specific cell, protect the DNA from degradation, and direct it to the cell nucleus. A number of viruses have been engineered for gene therapy applications. Packaging cell lines for some viral systems ensure that virtually no infectious viruses are present in a sample of vector viruses. Nonviral gene delivery systems include injection of pure DNA, bombardment of a target tissue with DNA-coated particles, and cellular uptake of DNA that is enclosed within a lipid envelope. In addition, HACs may find use as vectors for the long-term maintenance and expression of therapeutic genes in human cells. Generally, the major drawbacks of the current generation of gene therapy vector systems are immunogenicity, lack of cell specificity, inefficient gene transfer, and limited therapeutic gene expression.

REFERENCES

Agrawal, S. (ed.). 1996. *Antisense Therapeutics.* Humana Press Inc., Totowa, NJ.

Alton, E., and C. Kitson. 2000. Gene therapy for cystic fibrosis. *Exp. Opin. Investig. Drugs* **9:**1523–1535.

Bainbridge, J. W. B., A. J. Smith, S. S. Barker, S. Robbie, R. Henderson, K. Balaggan, A. Viswanathan, G. E. Holder, A. Stockman, N. Tyler, S. Petersen-Jones, S. S. Bhattacharya, A. J. Thrasher, F. W. Fitzke, B. J. Carter, G. S. Rubin, A. T. Moore, and R. R. Ali. 2008. Effect of gene therapy on visual function in Leber's congenital amaurosis. *N. Engl. J. Med.* **358:**2231–2239.

Bennett, J., and A. M. Maguire. 2000. Gene therapy for ocular disease. *Mol. Ther.* **1:**501–505.

Bramlage, B., E. Luzi, and F. Eckstein. 1998. Designing ribozymes for the inhibition of gene expression. *Trends Biotechnol.* **16:**434–438.

Burfeind, P., C. L. Chernicky, F. Rininsland, J. Ilan, and J. Ilan. 1996. Antisense RNA to the type I insulin-like growth factor receptor suppresses tumor growth and prevents invasion by rat prostate cancer cells in vivo. *Proc. Natl. Acad. Sci. USA* **93:**7263–7268.

Cairns, M. J., T. M. Hopkins, C. Witherington, L. Wang, and L.-Q. Sun. 1999. Target site selection for an RNA-cleaving catalytic DNA. *Nat. Biotechnol.* **17:**480–486.

Cavazzana-Calvo, M., S. Hacein-Bey, G. de Saint Basile, F. Gross, E. Yvon, P. Nusbaum, F. Selz, C. Hue, S. Certain, J. L. Casanova, P. Bousso, F. L. Deist, and A. Fischer. 2000. Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* **288:**669–672.

Choo, K. H. 2001. Engineering human chromosomes for gene therapy studies. *Trends Mol. Med.* **7:**235–237.

Chu, T. C., K. Y. Twu, A. D. Ellington, and M. Levy. 2006. Aptamer mediated siRNA delivery. *Nucleic Acids Res.* **34:**e73.

Chuah, M. K., D. Collen, and T. Vanden Driessche. 2001. Gene therapy for hemophilia. *J. Gene Med.* **3:**3–20.

Cole-Strauss, A., K. Yoon, Y. Xiang, B. C. Bryne, M. C. Rice, J. Gryn, W. K. Holloman, and E. B. Kmiec. 1996. Correction of the mutation responsible for sickle cell anemia by an RNA-DNA oligonucleotide. *Science* **273:**1386–1389.

Cullen, B. R. 2006. Enhancing and confirming the specificity of RNAi experiments. *Nat. Methods* **3:**677–681.

Davidson, B. L. 2006. All in the family. *Nat. Biotechnol*. **24:**951–952.

De Fougerolles, A., H.-P. Vornlocher, J. Maraganore, and J. Lieberman. 2007. Interfering with disease: a progress report on siRNA-based therapeutics. *Nat. Rev.* **6:**443–453.

Dillon, C. P., P. Sandy, A. Nencioni, S. Kissler, D. A. Rubinson, and L. Van Parijs. 2005. RNAi as an experimental and therapeutic tool to study and regulate physiological and disease processes. *Annu. Rev. Physiol.* **67:**147–173.

Elbashir, S. M., J. Harborth, W. Lendeckel, A. Yalcin, K. Weber, and T. Tuschi. 2001. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411:**494–498.

Elbashir, S. M., W. Lendeckel, and T. Tuschi. 2001. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* **15:**188–200.

Eng, C. M., M. Banikazemi, R. E. Gordon, M. Goldman, R. Phelps, L. Kim, A. Gass, J. Winston, S. Dikman, J. T. Fallon, S. Brodie, C. B. Stacy, D. Mehta, R. Parsons, K. Norton, M. O'Callaghan, and R. J. Desnick. 2001. A phase 1/2 clinical trial of enzyme replacement in Fabry disease: pharmacokinetic, substrate clearance, and safety studies. *Am. J. Hum. Genet.* **68:**711–722.

Faria, M., D. G. Spiller, C. Dubertret, J. S. Nelson, M. R. H. White, D. Scherman, C. Hélène, and C. Giovannangeli. 2001. Phosphoramidate oligonucleotides as potent antisense molecules in cells and in vivo. *Nat. Biotechnol.* **19:**40–44.

Fichou, Y., and C. Férec. 2006. The potential of oligonucleotides for therapeutic applications. *Trends Biotechnol.* **24:**563–570.

Field, A. K. 1998. Viral targets for antisense oligonucleotides: a mini review. *Antivir. Res.* **37:**67–81.

Haasnoot, E., M. Westerhout, and B. Berkhout. 2007. RNA interference against viruses: strike and counter-strike. *Nat. Biotechnol*. **25:**1435–1443.

Haidet, A. M., L. Rizo, C. Handy, P. Umapathi, A. Eagle, C. Shilling, D. Boue, P. T. Martin, Z. Sahenk, J. R. Mendell, and B. K. Kaspar. 2008. Long-term enhancement of skeletal muscle mass and strength by single gene administration of myostatin inhibitors. *Proc. Natl. Acad. Sci. USA* **105:**4318-4322.

Isaacs, F. J., D. J. Dwyer, and J. J. Collins. 2006. RNA synthetic biology. *Nat. Biotechnol.* **24:**545–554.

Johnson, W. E. 2007. Assisted suicide for retroviruses. *Nat. Biotechnol.* **25:**643–644.

Kay, M. A., C. S. Manno, M. V. Ragni, P. J. Larson, L. B. Couto, A. McClelland, B. Glader, A. J. Chew, S. J. Tai, R. W. Herzog, V. Arruda, F. Johnson, C. Scallan, E. Skarsgard, A. W. Flake, and K. A. High. 2000. Evidence for gene transfer and expression of factor IX in haemophilia B patients treated with an AAV vector. *Nat. Genet.* **24:**257–261.

Kim, D.-H., M. A. Behlke, S. D. Rose, M.-S. Chang, S. Choi, and J. J. Rossi. 2005. Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. *Nat. Biotechnol.* **23:**222-226.

Kono, K., Y. Torikoshi, M. Mitsutomi, T. Itoh, N. Emi, H. Yanagie, and T. Takagishi. 2001. Novel gene delivery systems: complexes of fusigenic polymer-modified liposomes and lipoplexes. *Gene Ther.* **8:**5–12.

Lee, J.-H., M. D. Canny, A. De Erkenez, D. Krilleke, Y.-S. Ng, D. T. Shima, A. Pardi, and F. Jucker. 2005. A therapeutic aptamer inhibits angiogenesis by specifically targeting the heparin binding domain of VEGF165. *Proc. Natl. Acad. Sci. USA* **102:**18902–18907.

Maguire, A. M., F. Simonelli, E. A. Pierce, E. N. Pugh, F. Mingozzi, J. Bennicelli, S. Banfi, K. A. Marshall, F. Testa, E. M. Surace, S. Rossi, A. Lyubarsky, V. R. Arruda, B. Konkle, E. Stone, J. Sun, J. Jacobs, L. Dell'Osso, R. Hertle, J.-X. Ma, T. M. Redmond, X. Zhu, B. Hauck, O. Zelenaia, K. S. Shindler, M. G. Maguire, J. F. Wright, N. J. Volpe, J. W. McDonnell, A. Auricchio, K. A. High, and J. Bennett. 2008. Safety and efficacy of gene transfer for Leber's congenital amaurosis. *N. Engl. J. Med.* **358:**2240–2248.

Matzen, K., L. Elzaouk, A. A. Matskevich, A. Nitzsche, J. Heinrich, and K. Moelling. 2007. RNase H-mediated retrovirus destruction in vivo triggered by oligodeoxynucleotides. *Nat. Biotechnol.* **25:**669–674.

McNamara, J. O., II, E. R. Andrechek, Y. Wang, K. D. Viles, R. E. Rempel, E. Gilboa, B. A. Sullenger, and P. H. Giangrande. 2006. Cell type-specific delivery of siRNAs with aptamer-siRNA chimeras. *Nat. Biotechnol*. **24:**1005–1015.

Morris, M. C., L. Chaloin, F. Heitz, and G. Divita. 2000. Translocating peptides and proteins and their use for gene delivery. *Curr. Opin. Biotechnol.* **11:**461–466.

Nettelbeck, D. M., V. Jérôme, and R. Müller. 2000. Gene therapy: designer promoters for tumour targeting. *Trends Genet.* **16:**174–181.

Nimjee, S. M., C. P. Rusconi, and B. A. Sullenger. 2005. Aptamers: an emerging class of therapeutics. *Annu. Rev. Med.* **56:**555–583.

Oliveira, S., G. Storm, and R. M. Schiffelers. 2006. Targeted delivery of siRNA. *J. Biomed. Biotechnol.* **2006:**1–9.

Pendergast, P. S., H. N. Marsh, D. Grate, J. M. Healy, and M. Stanton. 2005. Nucleic acid aptamers for target validation and therapeutic applications. *J. Biomol. Technol*. **16:**224–234.

Putnam, D. A. 1996. Antisense strategies and therapeutic applications. *Am. J. Health Syst. Pharm.* **53:**151–160.

Rossi, J. J. 2004. A cholesterol connection in RNAi. *Nature* **432:**155–156.

Rossi, J. J. 2005. Receptor-targeted siRNAs. *Nat. Biotechnol*. **23:**682–683.

Rubanyi, G. M. 2001. The future of human gene therapy. *Mol. Aspects Med.* **22:**113–142.

Sampson, T. 2003. Aptamers and SELEX: the technology. *World Patent Info.* **25:**123–129.

Samuel, V. T., C. S. Choi, T. G. Phillips, A. J. Romanelli, J. G. Geisler, S. Bhanot, R. McKay, B. Monia, J. R. Shutter, R. A. Lindberg, G. I. Shulman, and M. M. Veniant. 2006. Targeting Foxo1 in mice using antisense oligonucleotide improves hepatic and peripheral insulin action. *Diabetes* **55:**2042–2050.

Scherer, L. J., and J. J. Rossi. 2003. Approaches for the sequence-specific knockdown of mRNA. *Nat. Biotechnol.* **21:**1457–1464.

Shahi, S., G. K. Shanmugasundaram, and A. C. Banerjea. 2001. Ribozymes that cleave reovirus genome segment Si also protect cells from pathogenesis caused by reovirus infection. *Proc. Natl. Acad. Sci. USA* **98:**4101–4106.

Siolas, D., C. Lerner, J. Burchard, W. Ge, P. S. Linsley, P. J. Paddison, G. J. Hannon, and M. A. Cleary. 2005. Synthetic shRNAs as potent RNAi triggers. *Nat. Biotechnol.* **23:**227–231.

Sioud, M., and D. R. Sorensen. 1998. A nuclease-resistant protein kinase Cα ribozyme blocks glioma cell growth. *Nat. Biotechnol.* **16:**556–561.

Snøve, O., and J. J. Rossi. 2006. Expressing short hairpin RNAs *in vivo*. *Nat. Methods* **3:**689–695.

Song, E., P. Zhu, S.-K. Lee, D. Chowdhury, S. Kussman, D. M. Dykxhoorn, Y. Feng, D. Palliser, D. B. Weiner, P. Shankar, W. A. Marasco, and J. Lieberman. 2005. Antibody mediated *in vivo* delivery of small interfering RNAs via cell-surface receptors. *Nat. Biotechnol.* **23:**709–717.

Soutschek, J., A. Akinc, B. Bramlage, K. Charisse, R. Constien, M. Donoghue, S. Elbashir, A. Geick, P. Hadwiger, J. Harborth, M. John, V. Kesavan, G. Lavine, R. K. Pandey, T. Racie, K. G. Rajeev, I. Röhl, I. Toudjarska, G. Wang, S. Wuschko, D. Bumcrot, V. Koteliansky, S. Limmer, M. Manoharan, and H.-P. Vornlocher. 2004. Therapeutic silencing of an endogenous gene by systematic administration of modified siRNAs. *Nature* **432:**173-178.

Takeshita, F., Y. Minakuchi, S. Nagahara, K. Honma, H. Sasaki, K. Hirai, T. Teratani, N. Namatame, Y. Yamamoto, K. Hanai, T. Kato, A. Sano, and T. Ochiya. 2005. Efficient delivery of small interfering RNA to bone-metastatic tumors by using atelocollagen *in vivo*. *Proc. Natl. Acad. Sci. USA* **102:**12177–12182.

Trang, P., A. Kilani, J. Kim, and F. Liu. 2000. A ribozyme derived from the catalytic subunit of RNase P from *Escherichia coli* is highly effective in inhibiting replication of herpes simplex virus 1. *J. Mol. Biol.* **301:**817–826.

Trang, P., M. Lee, E. Nepomuceno, J. Kim, H. Zhu, and F. Liu. 2000. Effective inhibition of human cytomegalovirus gene expression and replication by a ribozyme derived from the catalytic RNA subunit of RNase P from *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **97:**5812–5817.

Van Rij, R. P., and R. Andino. 2006. The silent treatment: RNAi as a defense against virus infection in mammals. *Trends Biotechnol.* **24:**186–193.

Wolfrum, C., S. Shi, N. K. Jayaprakash, M. Jayaraman, G. Wang, R. J. Pandey, K. G. Rajeev, T. Nakayama, K. Charrise, E. M. Ndungo, T. Zimmermann, V. Koteliansky, M. Manoharan, and M. Stoffel. 2007. Mechanisms and optimization of in vivo delivery of lipophilic siRNAs. *Nat. Biotechnol.* **25:**1149–1157.

Wraight, C. J., P. J. White, S. C. McKean, R. D. Fogarty, D. J. Venables, I. J. Liepe, S. R. Edmondson, and G. A. Werther. 2000. Reversal of epidermal hyperproliferation in psoriasis by insulin-like growth factor I receptor antisense oligonucleotides. *Nat. Biotechnol.* **18:**521–526.

Xiang, S., J. Fruehauf, and C. J. Li. 2006. Short hairpin RNA-expressing bacteria elicit RNA interference in mammals. *Nat. Biotechnol.* **24:**697–702.

Yoon, K., A. Cole-Strauss, and E. B. Kmiec. 1996. Targeted gene correction of episomal DNA in mammalian cells mediated by a chimeric RNA·DNA oligonucleotide. *Proc. Natl. Acad. Sci. USA* **93:**2071–2076.

Zamore, P. D., T. Tuschi, P. A. Sharp, and D. P. Bartel. 2001. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101:**25–33.

REVIEW QUESTIONS

1. How can antisense oligonucleotides be used to treat psoriasis?

2. What are ribozymes, and how can they be used as human therapeutic agents?

3. What are interfering RNAs, and how might they be used as human therapeutic agents?

4. What is an aptamer, and how is it used as a therapeutic agent?

5. How can antibody genes be used to confer passive immunity?

6. How can interfering RNAs be delivered to specific cells?

7. In developing a new nucleic acid-based therapeutic agent, how would you decide between antisense oligonucleotides, ribozymes, aptamers, and interfering RNA?

8. What are the key attributes of a therapeutic gene delivery system for humans?

9. How can the interferon response, which is usually induced by double-stranded RNA, be avoided when utilizing siRNAs as therapeutic agents?

10. What are some of the advantages and disadvantages of deoxyribozymes compared to ribozymes?

11. How can the progression of age-related macular degeneration be limited using RNA therapeutics?

12

Vaccines

VACCINATION PROTECTS A RECIPIENT from pathogenic agents by establishing an immunological resistance to infection. An injected or oral vaccine induces the host to generate antibodies against the disease-causing organism; therefore, during future exposures, the infectious agent is inactivated (neutralized, or killed), its proliferation is prevented, and the disease state is not established.

Over 200 years ago, in 1796, Edward Jenner experimentally tested the folklore-based notion that human infection with a mild cattle disease called cowpox would protect infected individuals against the human disease smallpox. Smallpox is an extremely virulent disease with a high death rate. If one survives, permanent disfigurement, mental derangement, and blindness often follow. Jenner inoculated James Phipps, an 8-year-old boy, with exudate from a cowpox pustule. In two separate trials after the initial vaccination, the boy was fully protected against human smallpox. This country doctor had discovered the principle of vaccination.

Communicable diseases such as tuberculosis, smallpox, cholera, typhus, bubonic plague, and poliomyelitis, have in the past been a scourge for humankind. With the advent of vaccination, antibiotics, and effective public health measures, these epidemic diseases have, for the most part, been brought under control (Table 12.1). Occasionally, however, protective measures become ineffective, and devastating new outbreaks occur. In 1991, a cholera epidemic struck Peru, producing, over the next 3 years, approximately 1 million infections and several thousand deaths. Also, for many current human and animal diseases, there are no vaccines. Today, more than 2 billion humans suffer from diseases that theoretically could be curtailed by vaccination. In addition, new diseases for which vaccines might be useful continue to emerge.

In recent years, in some developed countries, a small but vocal minority of individuals have refused to have their children vaccinated. These individuals argue that many of the previously common illnesses have been vanquished, and they fear the potential side effects of the vaccinations more than the disease itself. In addition, many of these people question modern medicine and instead prefer to rely upon so-called traditional or

TABLE 12.1 Annual cases in Canada from various diseases before and after the introduction of vaccines against the causative agents of the diseases

Disease	Annual no. of cases before vaccine was introduced	No. of cases in 2002
Polio	20,000	0
Diphtheria	9,000	0
Rubella	69,000	16
Mumps	52,000	197
Haemophilus influenzae type b infection	2,000	48
Whooping cough	25,000	2,557
Measles	300,000	7

natural therapies. In fact, the small number of individuals who are not vaccinated benefit from the fact that the vast majority of other people in society have been immunized, thereby making it difficult for many diseases to spread through a community. However, in communities where vaccination levels decrease below a certain level, there is a real danger of some traditional diseases making a comeback.

Modern vaccines typically consist of either a killed (inactivated) or a live, nonvirulent (attenuated) form of an infectious agent. Traditionally, the infectious agent is grown in culture, purified, and either inactivated or attenuated without, of course, losing the ability to evoke an immune response that is effective against the virulent form of the infectious organism. Notwithstanding the considerable success that has been achieved in creating effective vaccines against diseases such as German measles, diphtheria, whooping cough, tetanus, smallpox, and poliomyelitis, there are a number of limitations to the current mode of vaccine production.

- Not all infectious agents can be grown in culture, so no vaccines have been developed for a number of diseases.
- Production of animal and human viruses requires animal cell culture, which is expensive.
- Both the yield and rate of production of animal and human viruses in culture are often quite low, making vaccine production costly.
- Extensive safety precautions are necessary to ensure that laboratory and production personnel are not exposed to a pathogenic agent.
- Batches of vaccine may not be killed or may be insufficiently attenuated during the production process, thereby introducing virulent organisms into the vaccine and inadvertently spreading the disease.
- Attenuated strains may revert, a possibility that requires continual testing to ensure that the reacquisition of virulence has not occurred.
- Not all diseases (e.g., acquired immune deficiency syndrome [AIDS]) are preventable through the use of traditional vaccines.
- Most current vaccines have a limited shelf life and often require refrigeration to maintain potency. This requirement creates storage problems in countries with large, unelectrified rural areas.

Within the last 2 decades, recombinant DNA technology has provided a means of creating a new generation of vaccines that overcome the drawbacks

of traditional vaccines. The availability of gene cloning has enabled researchers to contemplate various novel strategies for vaccine development.

- Virulence genes could be deleted from an infectious agent that retains the ability to stimulate an immunological response. In this case, the genetically engineered agent could be used as a live vaccine without concern about reversion to virulence, because it is impossible for a whole gene to be reacquired spontaneously during growth in pure culture.
- Live nonpathogenic carrier systems that carry discrete antigenic determinants of an unrelated pathogenic agent could be created. In this form, the carrier system facilitates the induction of a strong immunological response directed against the pathogenic agent.
- For infectious agents that cannot be maintained in culture, the genes for the proteins that have critical antigenic determinants can be isolated, cloned, and expressed in an alternative host system, such as *Escherichia coli* or a mammalian cell line. These cloned gene proteins can be formulated into a vaccine.
- There are some infectious agents that do not damage host cells directly; instead, the disease condition results when the host immune system attacks its own (infected) cells. For these diseases, it may be possible to create a targeted cell-specific killing system. Although not a true vaccine, this type of system attacks only infected cells, thereby removing the source of the adverse immunological response. In these cases, the gene for a fusion protein is constructed. First, one part of this fusion protein binds to an infected cell. Then, the other part kills the infected cell.

Because of less stringent regulatory requirements, the first vaccines that were produced by recombinant DNA techniques were for animal diseases, such as foot-and-mouth disease, rabies, and scours, a diarrheal disease of pigs and cattle. In addition, many more animal vaccines are currently being developed. For human diseases, a large number of recombinant vaccines are currently in various stages of development, including clinical trials (Table 12.2).

Unfortunately, in comparison to the number of new therapeutic agents, very few recombinant-DNA-based vaccines have been developed. Why, according to the vaccine producers, does it take so long for new vaccines to come to the marketplace? First, while there were 25 major vaccine manufacturers worldwide in 1970, in 2005 there were only 5. Second, vaccines are currently viewed as "almost a commodity," with little financial incentive to develop new vaccines; in 2005, the worldwide market for preventive vaccines was approximately $8 billion. Third, the U.S. government is a major purchaser of vaccines, forcing discount prices and thereby decreasing the potential profit. Fourth, in 1980 in the United States, "good manufacturing practices" were introduced into vaccine production, causing manufacturing costs to increase dramatically. Fifth, the transition from conventional to newer processes for vaccine production is expensive and time-consuming (not including clinical trials), so that it is preferable to continue using a more established technology. On the positive side, the focus of the larger vaccine manufacturers on large-scale products has provided smaller biotechnology companies with a number of niche opportunities to develop and market new products. Finally, since most vaccines are intended to protect large

TABLE 12.2 Human disease agents for which recombinant vaccines are currently being developed

Pathogenic agent	Disease
Viruses	
Varicella-zoster viruses	Chicken pox
Cytomegalovirus	Infection in infants and immuno-compromised patients
Dengue virus	Hemorrhagic fever
Hepatitis A virus	High fever, liver damage
Hepatitis B virus	Long-term liver damage
Herpes simplex virus type 2	Genital ulcers
Influenza A and B viruses	Acute respiratory disease
Japanese encephalitis	Encephalitis
Parainfluenza virus	Inflammation of the upper respiratory tract
Rabies virus	Encephalitis
Respiratory syncytial virus	Upper and lower respiratory tract lesions
Rotavirus	Acute infantile gastroenteritis
Yellow fever virus	Lesions of heart, kidney, and liver
Human immunodeficiency virus	AIDS
Bacteria	
Vibrio cholerae	Cholera
E. coli enterotoxin strains	Diarrheal disease
Neisseria gonorrhoeae	Gonorrhea
Haemophilus influenzae	Meningitis, septicemic conditions
Mycobacterium leprae	Leprosy
Neisseria meningitidis	Meningitis
Bordetella pertussis	Whooping cough
Shigella strains	Dysentery
Streptococcus group A	Scarlet fever, rheumatic fever, throat infection
Streptococcus group B	Sepsis, urogenital tract infection
Streptococcus pneumoniae	Pneumonia, meningitis
Clostridium tetani	Tetanus
Mycobacterium tuberculosis	Tuberculosis
Salmonella enterica serovar Typhi	Typhoid fever
Parasites	
Onchocerca volvulus	River blindness
Leishmania spp.	Internal and external lesions
Plasmodium spp.	Malaria
Schistosoma mansoni	Schistosomiasis
Trypanosoma spp.	Sleeping sickness
Wuchereria bancrofti	Filariasis

populations, the very large amounts of money that companies often charge to treat a single individual with some of the newer therapeutic agents (see chapter 10) are unrealistic for pricing of a new vaccine. In fact, it is precisely

in many poorer countries, where most individuals cannot afford to pay very much for treatment or immunization, that vaccines are needed the most.

Subunit Vaccines

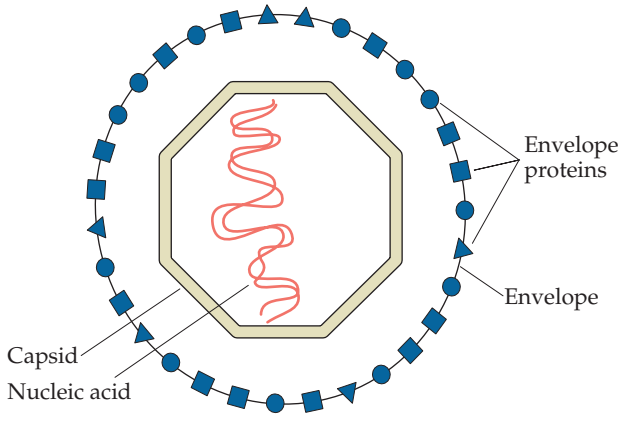
Vaccines generally consist of either killed or attenuated forms of the whole pathogenic agent. The antibodies elicited by these vaccines initiate an immune response to inactivate (neutralize) pathogenic organisms by binding to proteins on the outer surface of the agent. So, do vaccines need to contain the whole organism, or will specific portions of pathogenic organisms suffice? For disease-causing viruses, it has been shown that purified outer surface viral proteins, either capsid or envelope proteins (Fig. 12.1), are often sufficient for eliciting neutralizing antibodies in the host organism. Vaccines that use components of a pathogenic organism rather than the whole organism are called "subunit" vaccines; recombinant DNA technology is very well suited for developing new subunit vaccines.

There are advantages and disadvantages to the use of subunit vaccines. On the positive side, using a purified protein(s) as an immunogen ensures that the preparation is stable and safe, is precisely defined chemically, and is free of extraneous proteins and nucleic acids that can initiate undesirable side effects in the host organism. On the negative side, purification of a specific protein can be costly, and in certain instances, an isolated protein may not have the same conformation as it does in situ (within the viral capsid or envelope), with the result that its antigenicity is decreased. Obviously, the decision to produce a subunit vaccine depends on an assessment of several biological and economic factors.

Herpes Simplex Virus

Herpes simplex virus (HSV) has been implicated as a cancer-causing (oncogenic) agent, in addition to its more common roles in causing sexually transmitted disease, severe eye infections, and encephalitis, so prevention of HSV infection by vaccination with either killed or attenuated virus may

FIGURE 12.1 Schematic representation of an animal virus. Viruses generally consist of a relatively small nucleic acid genome (3 to 200 kb of either double- or single-stranded DNA or RNA) within a viral protein capsid that is sometimes, depending on the virus, surrounded by a protein-containing viral envelope (membrane).



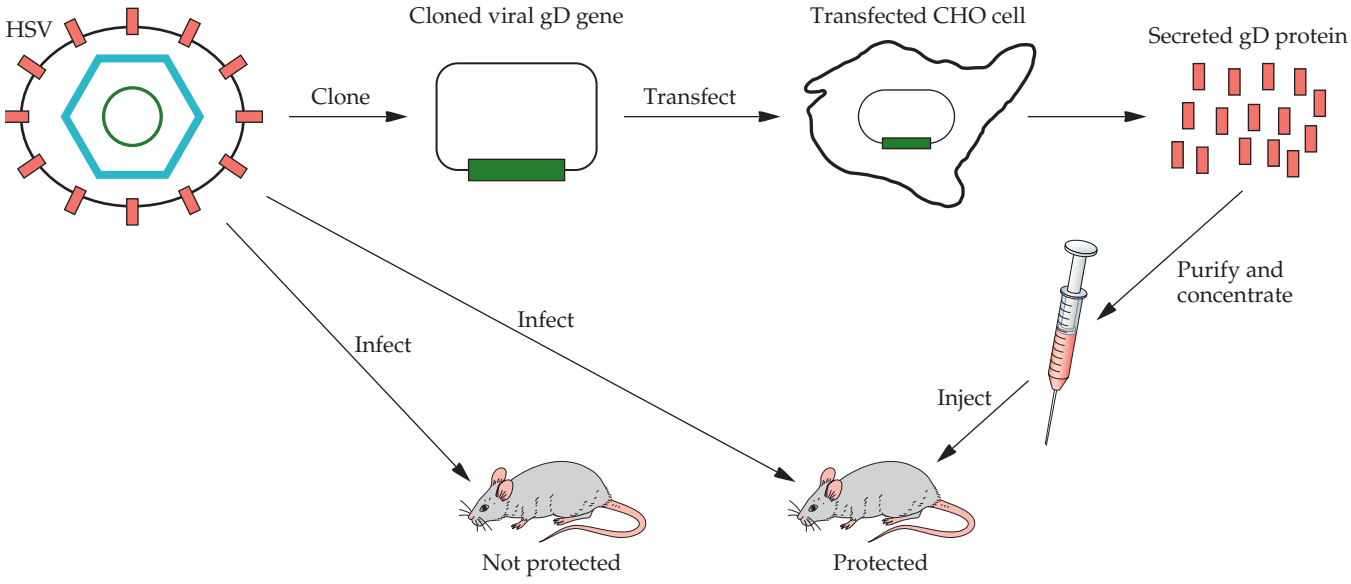


FIGURE 12.2 Schematic representation of the development of a subunit vaccine against HSV. The isolated HSV gD protein gene is used to transfect CHO cells. Then, the transfected cells are grown in culture and produce gD protein. Mice inoculated with the purified gD protein are protected against infection by HSV.

put the recipient at risk for cancer. Thus, protection against HSV would be best achieved by a subunit vaccine, which would not be oncogenic.

The primary requirement for creating any subunit vaccine is identification of the component(s) of the infectious agent that elicits antibodies that react against the intact form of the infectious agent. The HSV type 1 (HSV-1) envelope glycoprotein D (gD) is such a component, because after injection into mice, it elicits antibodies that neutralize intact HSV. The HSV-1 gD gene was isolated and then cloned into a mammalian expression vector and expressed in Chinese hamster ovary (CHO) cells (Fig. 12.2), which, unlike the *E. coli* system, properly glycosylate foreign eukaryotic proteins. The complete sequence of the gD gene encodes a protein that becomes bound to the mammalian host cell membrane (Fig. 12.3A). However, a membrane-bound protein is much more difficult to purify than a soluble one. Consequently, the gD gene was modified by removing the nucleotides encoding the C-terminal transmembrane-binding domain (Fig. 12.3B). The modified gene was then transformed into CHO cells, where the product was glycosylated and secreted into the external medium (Fig. 12.2). In laboratory trials, the modified form of gD was effective against both HSV-1 and HSV-2.

Foot-and-Mouth Disease

Foot-and-mouth disease virus (FMDV) has a devastating impact on cattle and swine and is extremely virulent, but for the most part, it has been possible to keep the negative effects of the virus to a minimum by using formalin-killed FMDV preparations as a vaccine. Approximately 1 billion doses of this killed-virus vaccine are used worldwide each year. The availability of the vaccine notwithstanding, in 2001, there was a major outbreak of foot-and-mouth disease in Europe in which tens of thousands of cattle were

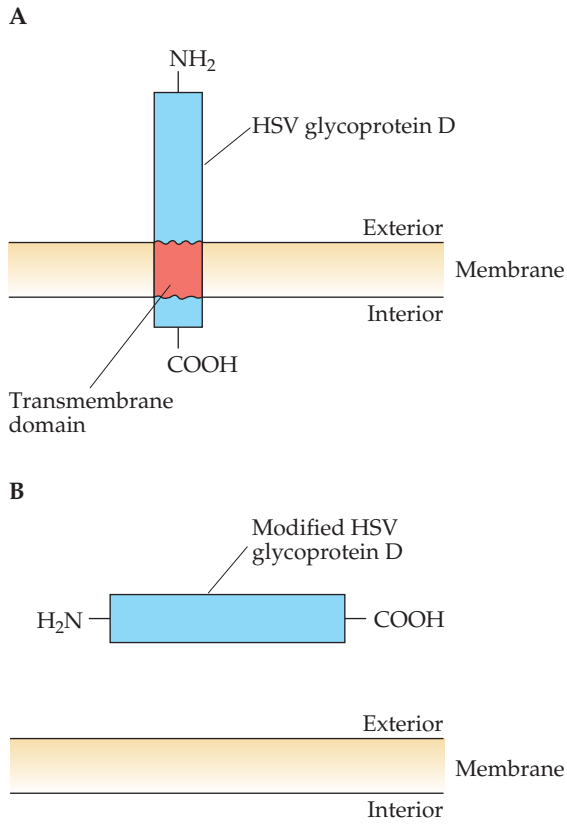


FIGURE 12.3 **(A)** Location in the envelope of HSV-1 gD with the transmembrane domain. **(B)** Extracellular location of a soluble gD without the transmembrane domain.

slaughtered and their carcasses were incinerated in an effort to prevent the virus from spreading.

Research on FMDV found that the major antigenic determinant that induces neutralizing antibodies is capsid viral protein 1 (VP1). Although purified VP1 is a much less potent antigen than intact viral particles, it can still elicit neutralizing antibodies by itself and therefore can protect animals from infection by FMDV. Thus, the gene for VP1 became a target for cloning.

The genome of FMDV is composed of single-stranded RNA (approximately 8,000 nucleotides long). Therefore, for recombinant DNA manipulations, it was necessary first to synthesize a double-stranded complementary DNA (cDNA) of the entire genome (Fig. 12.4). This cDNA was then digested with restriction enzymes, and the fragments were cloned in an *E. coli* expression vector. The product of the VP1 coding sequence was identified immunologically as part of a fusion protein under the control of the p^L promoter–*c*I repressor system. The fusion protein was 396 amino acids long and consisted of a portion of a stabilizing carrier protein, i.e., the bacteriophage MS2 replicase protein, as well as the entire coding sequence of the FMDV VP1 protein (Fig. 12.4). The fusion protein containing the VP1 protein fragment was able to generate neutralizing antibodies to FMDV.

A fusion protein, however, faces more government regulatory hurdles than intact VP1 would because of the potential immunogenic effects of the

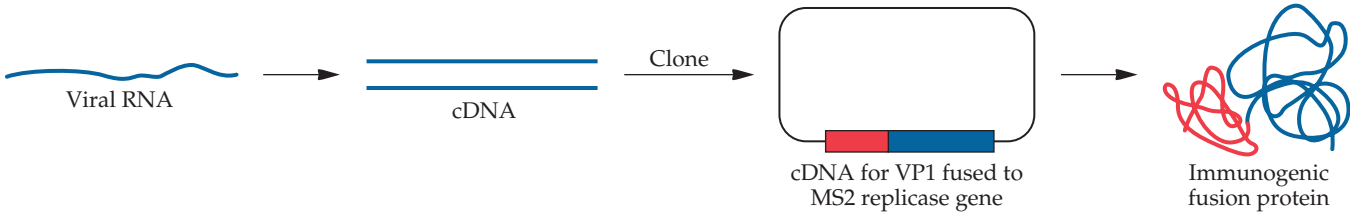


FIGURE 12.4 Schematic representation of the development of a subunit vaccine against foot-and-mouth disease. The entire viral RNA is made into cDNA, which is then digested with restriction enzymes. The DNA fragments are cloned into an expression vector in frame with the gene for the *E. coli* bacteriophage MS2 replicative protein. The plasmid constructs are used to transform *E. coli,* and then the stable fusion protein is isolated and used to inoculate animals.

non-VP1 component. Therefore, the VP1 sequence alone will have to be subcloned onto a different expression vector. Nevertheless, a subunit vaccine for foot-and-mouth disease could soon be ready for preclinical trials.

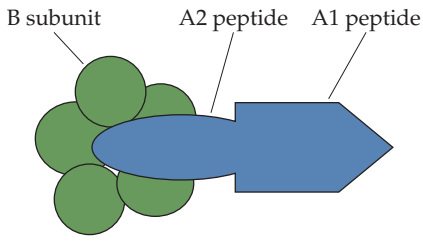
Cholera

The bacterium *Vibrio cholerae,* the causative agent of cholera, colonizes the small intestine and secretes large amounts of a hexameric enterotoxin, which is the actual pathogenic agent. This protein consists of one subunit, the A subunit, that has ADP ribosylation activity and stimulates adenylate cyclase, and five identical B subunits that bind specifically to an intestinal mucosal cell receptor (Fig. 12.5). The A subunit has two functional domains: the A_1 peptide, which contains the toxic activity, and the A_2 peptide, which joins the A subunit to the B subunits. Until a few years ago, a traditional cholera vaccine consisting of phenol-killed *V. cholerae* was in common use. This vaccine generated only moderate protection, typically lasting from about 3 to 6 months. More recently, a vaccine (Dukoral) consisting of heat-inactivated *V. cholerae* Inaba classic strain, heat-inactivated Ogawa classic strain, formalin-inactivated Inaba El Tor strain, formalin-inactivated Ogawa classic strain, and a recombinant cholera toxin B subunit, has come into use. The vaccine is taken orally (two doses 1 week apart), and it is claimed that an additional booster immunization is not required for about 2 years.

SARS

In 2003, there were more or less simultaneous outbreaks in several major cities, including Hong Kong, Singapore, and Toronto, of a new, unknown disease. The first case of this disease, severe acute respiratory syndrome, or SARS, was reported in Guangdong Province, southern People's Republic of China, in November 2002. Given the enormous frequency of air travel, the disease rapidly spread to 29 countries on five continents. With the assistance of the World Health Organization, authorities in affected regions immediately implemented strict infection control procedures, so that by mid-July 2003, the outbreak was effectively contained. However, this was not before a total of 8,096 SARS cases and 774 associated deaths were reported. Within a very short time, scientists had identified a novel coronavirus as the causative agent of the disease. The SARS virus contains a single-stranded plus-sense RNA genome of approximately 30 kb. In nature, the viral spike protein, which is inserted into the viral membrane, binds to

FIGURE 12.5 Schematic representation of hexameric cholera toxin. The A peptide is shown in blue, and the B peptide is shown in green.



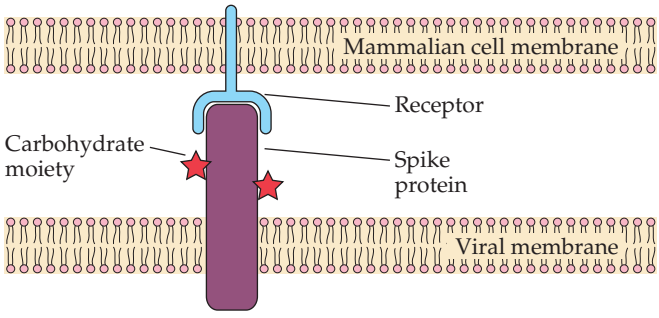


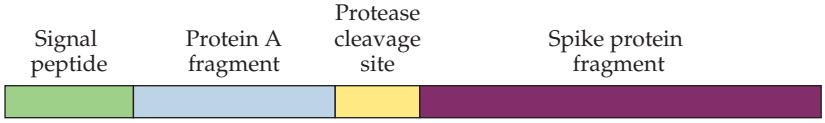
FIGURE 12.6 Schematic representation of the binding of the SARS virus glycosylated spike protein to a cellular outer surface protein receptor.

a receptor protein that is present on the surfaces of mammalian host cells (Fig. 12.6). Following the binding of the virus to the receptor, the viral and cell membranes can fuse, thereby facilitating the entry of the virus into the cell. The spike protein (or the external portion of the molecule) is an attractive candidate for the development of a subunit vaccine. In practice, it was found that the external portion of the spike protein (i.e., amino acids 318 to 510) could bind efficiently to the host cell receptor protein. Following the determination of the complete nucleotide sequence of the SARS virus in 2003, it was relatively straightforward to express a codon-optimized version of this 192-amino-acid peptide in CHO cells. In addition to encoding the 192-amino-acid spike peptide, the DNA construct introduced into the CHO cells also included a mammalian secretion signal, an N-terminal *(Staphylococcus aureus)* protein A purification tag, and a tobacco etch virus protease cleavage site (Fig. 12.7). The recombinant protein synthesized in CHO cells was secreted into the growth medium, purified by affinity chromatography on a column containing immobilized immunoglobulin G, and then digested with tobacco etch virus protease to remove the protein A purification tag. Using this construct, the spike protein fragment was readily synthesized and purified. To date, the fully glycosylated form of this subunit vaccine candidate has been shown to elicit a strong immune response in mice. It still remains to be seen whether it can protect immunized animals against infection with the SARS virus.

Staphylococcus aureus

The gram-positive bacterium *S. aureus* is a major cause of hospital-acquired infection. This bacterium produces a pore-forming toxin; is a leading cause of infections of the bloodstream, lower respiratory tract, and skin; and, because of the emergence of antibiotic-resistant strains, is a serious public health threat. To address the challenge of treating *S. aureus* infections,

FIGURE 12.7 The main features of a portion of the recombinant plasmid construct used to transfect CHO cells and to produce the fragment of the SARS spike protein that interacts with the host cell receptor.



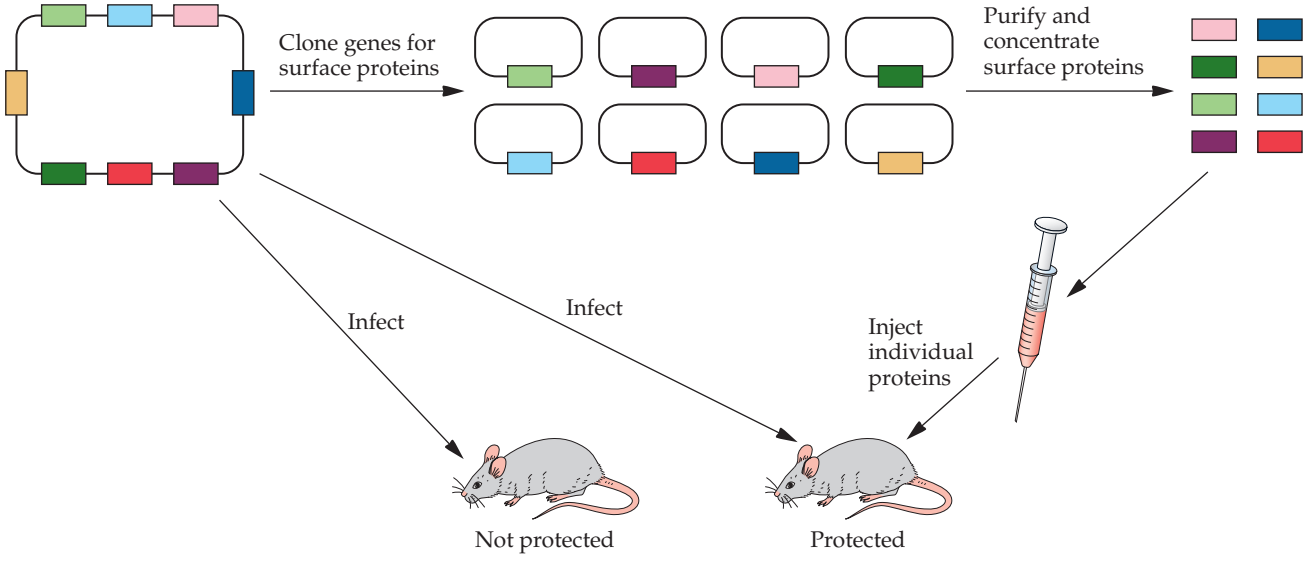


FIGURE 12.8 Development of a vaccine against *S. aureus*.

whole-cell attenuated or killed vaccines have been developed. However, these vaccines have not been particularly effective. Similarly, subunit vaccines composed of individual bacterial surface proteins generate immune responses that afford only partial protection when tested in experimental animals. However, a more effective subunit vaccine has recently been developed to protect individuals against *S. aureus* by combining several of the bacterium's antigens (Fig. 12.8). Starting with one disease-causing strain of *S. aureus*, 23 bacterial outer surface proteins were identified from genomic DNA sequence data. Then, the coding regions of these proteins, minus the signal sequences, were polymerase chain reaction (PCR) amplified and cloned into plasmid vectors that enabled the proteins to be expressed in *E. coli* with a poly-His tag at the N terminus of the protein (to facilitate the purification of the overexpressed protein). The proteins were expressed and purified, and mice were separately immunized with each of the 23 purified proteins. The immunized mice were subsequently challenged by injections of live disease-causing *S. aureus*. Many of the recombinant surface proteins generated an immune response that afforded partial protection against staphylococcal disease, with some proteins affording more protection than others. However, over the long term, immunization with individual surface proteins afforded only modest protection. A mixture of the four proteins that individually generated the most effective antibodies was used to immunize mice and was found to completely protect against the pathogen. The experimental design ensured that only common, and not strain-specific, *S. aureus* surface proteins were used to immunize mice. Thus, it is not surprising that the tetravalent subunit vaccine that was developed was effective against five different clinical isolates (strains) of *S. aureus*. This work represents an important first step in the development of an *S. aureus* vaccine.

Human Papillomavirus

Human papillomavirus is the causative agent of many common sexually transmitted diseases. While most of these infections are benign and often

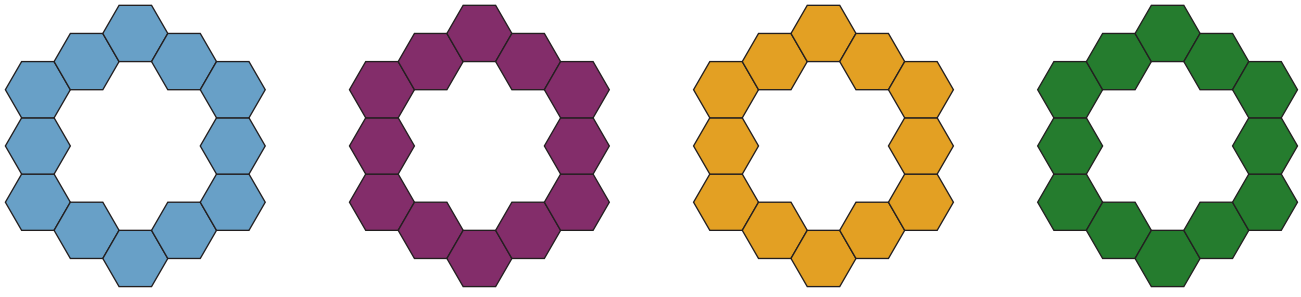


FIGURE 12.9 Schematic representation of the virus-like particles assembled from cloned and overproduced L1 proteins from the capsids of four different strains of human papillomavirus. These virus-like particles are the constituents of a commercial subunit vaccine against the virus.

asymptomatic, persistent infection with some strains of human papillomavirus is associated with the development of cervical and related cancers, as well as genital warts. Since human papillomavirus type 16 is associated with approximately 50% of cervical cancers, a vaccine that prevents human papillomavirus type 16 infection could significantly reduce the incidence of cervical cancer. Moreover, a vaccine that is directed against several different types of human papillomaviruses could effectively prevent nearly all human papillomavirus-induced cervical cancers. To put this into perspective, cervical cancer is the second most commonly diagnosed cancer among women worldwide, accounting for more than 250,000 deaths per year.

In June 2006, the U.S. Food and Drug Administration approved a vaccine that protects women against infection by human papillomavirus types 6, 11, 16, and 18, the types most frequently associated with cervical cancer and genital warts. By the end of 2006, this vaccine had been approved for use in more than 50 countries worldwide. The vaccine, called Gardasil, is quadrivalent, i.e., it contains virus-like particles assembled from the major capsid (L1) proteins of the above-mentioned four types of human papillomavirus (Fig. 12.9 and Box 12.1). It was previously shown that the L1 protein can self-assemble into virus-like particles that resemble papillomavirus virions, and these particles are highly immunogenic, inducing neutralizing antibodies directed against the whole live virus. The gene for the L1 protein from each of the four virus types was cloned and expressed in a recombinant *Saccharomyces cerevisiae* (yeast) strain. Following separate fermentations of the four yeast strains, the viral capsid proteins assembled into virus-like particles (i.e., the viral capsid without any other viral proteins or the viral nucleic acid). These virus-like particles were then purified and combined to form the quadrivalent vaccine.

Peptide Vaccines

The question arises as to whether a small discrete portion (domain) of a protein can act as an effective subunit vaccine and induce the production of neutralizing antibodies. Intuitively, one would expect that only the portions, or domains, of a protein that are accessible to antibody binding, that is, those on the exterior surface of the virus, would be immunologically important and that those located in inaccessible regions inside the virus particle could be ignored if they do not contribute to the conformation of

BOX 12.1

A Vaccine To Prevent Cervical Cancer

On 15 September 2007, the headline on the front page of *The Globe and Mail,* a Toronto, Canada, newspaper, read, "Should your daughter get the needle?" The article that followed related how the Canadian federal government, in conjunction with the Ontario provincial government, was funding a program that would offer free vaccinations against human papillomavirus to girls in grade 8 (typically 12- and 13-year-olds). The vaccine, which had received approval in the United States a year earlier, provides inoculated women with immunity against the viruses that are responsible for approximately 70% of all cervical cancers and 90% of genital warts. Grade 8 was chosen because, according to officials, it is before most girls become sexually active. The vaccine, sold under the brand name Gardasil, is given by needle in three doses over 6 months and is approved for females between the ages of 9 and 26. The rationale for giving the vaccine at such a young age is related to the fact that once a woman has been exposed to the four strains of the virus for which the vaccine provides protection, the vaccine will no longer be effective. The three doses of vaccine cost about $300 to $400, although those inoculated through this program received the vaccine free of charge. Boys can also get human papillomavirus infections, but testing is still under way to determine whether the vaccine works as well for them.

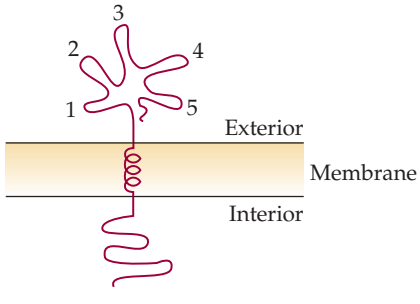
According to *The Globe and Mail,* "For many parents it's a no-brainer: anything that will protect their daughters from cancer … is worth the risks." However, at the same time, a small but vocal minority has expressed serious reservations about this program. On one hand, there are individuals who do not trust the medical establishment, the pharmaceutical companies, and/or the government. Others have expressed concerns that some girls will naively believe that this vaccine will protect them against any and all sexually transmitted diseases and use this as a rationale or excuse for becoming sexually active at an early age. Still others have questioned the potential side effects from the vaccine, despite the fact that extensive clinical trials have shown that they are quite rare. Notwithstanding the concerns of some individuals, the vaccine was initially offered through school inoculation programs to young females in the Canadian provinces of Newfoundland and Labrador, Prince Edward Island, Nova Scotia, and Ontario. However, by September 2008, all of the other provinces in Canada had decided to implement this program. The real benefits of the program (hopefully an enormous reduction in cervical cancer) may not be known for several decades; in the meantime, the debate will continue. Also, since the vaccine does not protect against all strains of human papillomavirus, it is essential that women continue to get an annual Pap test.

the immunogenic domain (Fig. 12.10). If this argument has validity, it is possible that short peptides that mimic epitopes (antigenic determinants) will be immunogenic and could be used as vaccines (peptide vaccines).

However, there are certain limitations to using short peptides as vaccines:

- To be effective, an epitope must consist of a short stretch of contiguous amino acids, which does not always occur naturally.
- The peptide must be able to assume the same conformation as the epitope in the intact viral particle.
- A single epitope may not be sufficiently immunogenic.

FIGURE 12.10 Generalized envelope-bound protein with external epitopes (1 to 5) that might elicit an immune response.



Foot-and-Mouth Disease

Potential epitopes of the soluble antigenic FMDV VP1 were identified from the X-ray crystallographic structure, and chemically synthesized domains of the protein were tested as candidate peptide vaccines. Peptides corresponding to amino acids 141 to 160, 151 to 160, and 200 to 213, which are located near the C-terminal end of VP1, and amino acids 9 to 24, 17 to 32, and 25 to 41, which are located near the N-terminal end of VP1, were each bound to a separate inert carrier protein (keyhole limpet hemocyanin) and injected into guinea pigs (Fig. 12.11). Very small peptides are usually rap-

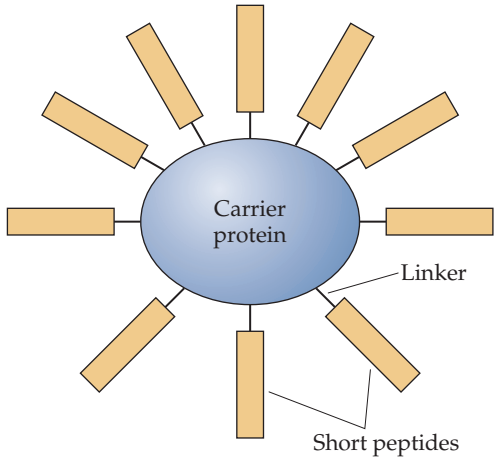


FIGURE 12.11 Structure of a peptide vaccine composed of identical short peptides bound to a carrier protein.

idly degraded unless they are bound to the surface of a larger carrier molecule. A single inoculation with peptide 141 to 160 elicited sufficient antibody to protect animals against subsequent challenges with FMDV. By contrast, inoculation with complete VP1 or peptide 9 to 24, 17 to 32, or 25 to 41 yielded lower levels of neutralizing antibodies.

In an additional experiment, a longer peptide consisting of amino acids 141 to 158 joined to amino acids 200 to 213 by two proline residues elicited high levels of neutralizing antibodies in guinea pigs, even when it was injected without any carrier protein. This "two-peptide" molecule was more effective than either of the single peptides alone and prevented FMDV proliferation in cattle, as well as in guinea pigs.

Although these results were promising, the amount (dose) of peptide material that had to be used to elicit an immunological response was approximately 1,000 times the amount of inactivated FMDV needed to elicit the same response. To overcome this problem, DNA encoding FMDV VP1 peptide 142 to 160 was linked to the gene encoding a highly immunogenic carrier molecule, hepatitis B virus core antigen (HBcAg). When the gene for this fusion protein was expressed in either *E. coli* or animal cells in culture, the protein molecules self-assembled into stable "27-nm particles," with the FMDV VP1 peptide located on the outer surface of the particle. These particles are highly immunogenic in laboratory animals. Therefore, HBcAg may be an effective carrier molecule for such short synthetic peptides. A comparison of the immunogenicities in guinea pigs of a variety of FMDV peptide vaccines, all of which contained the VP1 peptide 142 to 160 sequence, revealed that a fusion protein containing HBcAg and FMDV VP1 amino acids 142 to 160 was approximately 1/10 as immunogenic as inactivated FMDV particles, 35 times more immunogenic than a fusion protein containing *E. coli* β-galactosidase and FMDV VP1 amino acids 137 to 162, and 500 times more immunogenic than the free synthetic peptide composed of amino acids 142 to 160. Because synthetic peptides fused to HBcAg do not interfere with the assembly of the 27-nm hepatitis B virus-like particles, and because these particles are nearly as immunogenic as the intact virus from which the synthetic peptide was derived, this approach may become a general method for the delivery of peptide vaccines.

Malaria

The genus *Plasmodium* consists of approximately 125 known species of parasitic protozoa, 5 of which are known to infect humans and cause malaria. The *Plasmodium* life cycle is very complex. Sporozoites from the saliva of a biting female mosquito are transmitted to either the blood or the lymphatic system and then migrate to the liver and invade liver cells (hepatocytes) (Fig. 12.12). The parasite buds off the hepatocytes in merosomes containing hundreds or thousands of merozoites. These merosomes lodge in pulmonary capillaries and slowly disintegrate there, generally over 2 or 3 days, releasing merozoites. The merozoites invade the red blood cells, where the parasite divides several times to produce new merozoites, which then leave the red blood cells and travel within the bloodstream to invade new red blood cells. The parasite eventually forms gametocytes, which may be ingested by feeding mosquitoes. Fusion of the gametes that develop from gametocytes leads to the formation of new sporozoites in the mosquito that can infect new individuals, spreading the disease.

In the life cycle of the malaria parasite, it is the asexual blood-stage multiplication that is responsible for most of the acute symptoms of the disease. In areas where malaria is endemic, some individuals show considerable resistance to the disease despite the fact that when their blood is examined they are found to carry the parasite. This resistance to the worst symptoms of malaria was shown to be a result of an "antibody-dependent cellular-inhibition" mechanism that inhibits parasite development. In other words, some individuals who were infected with the malaria parasite made antibodies against a parasite protein that prevented the growth of the parasite. Following a detailed study, it was determined that the protective antibodies targeted merozoite surface protein 3. When this protein was examined in different strains of *Plasmodium,* it was observed that while the N-terminal part of the protein varied considerably from one strain to another, the C-terminal end of the protein was highly conserved among the various isolates of the parasite. It was therefore decided to chemically synthesize peptides that corresponded to small portions of the C terminus of merozoite surface protein 3. Human antibodies from individuals who were resistant to the parasite were affinity purified based upon their interaction with one or more of these peptides. The antibodies that bound to the peptides were then tested in an antibody-dependent cellular-inhibition assay. Antibodies directed against peptides B, C, and D (Fig. 12.13) had a major inhibitory effect on parasite growth. Based on the ability of peptides B, C, and D to bind to and select protective antibodies, a peptide representing amino acid residues 181 to 276 of merozoite surface protein 3 was chemically synthesized. This peptide is currently being tested in clinical trials as a novel malaria vaccine. While more research needs to be done, in the future, synthetic peptide vaccines could become highly specific, relatively inexpensive, safe, and effective alternatives to traditional vaccines.

Genetic Immunization: DNA Vaccines

Delivery

A novel strategy that elicits an antibody response without the introduction of an antigen has been developed. In this case, the gene encoding an antigenic protein is introduced into cells of a target animal, where the antigen is synthesized (Table 12.3). In the initial experiments, gold microprojectiles

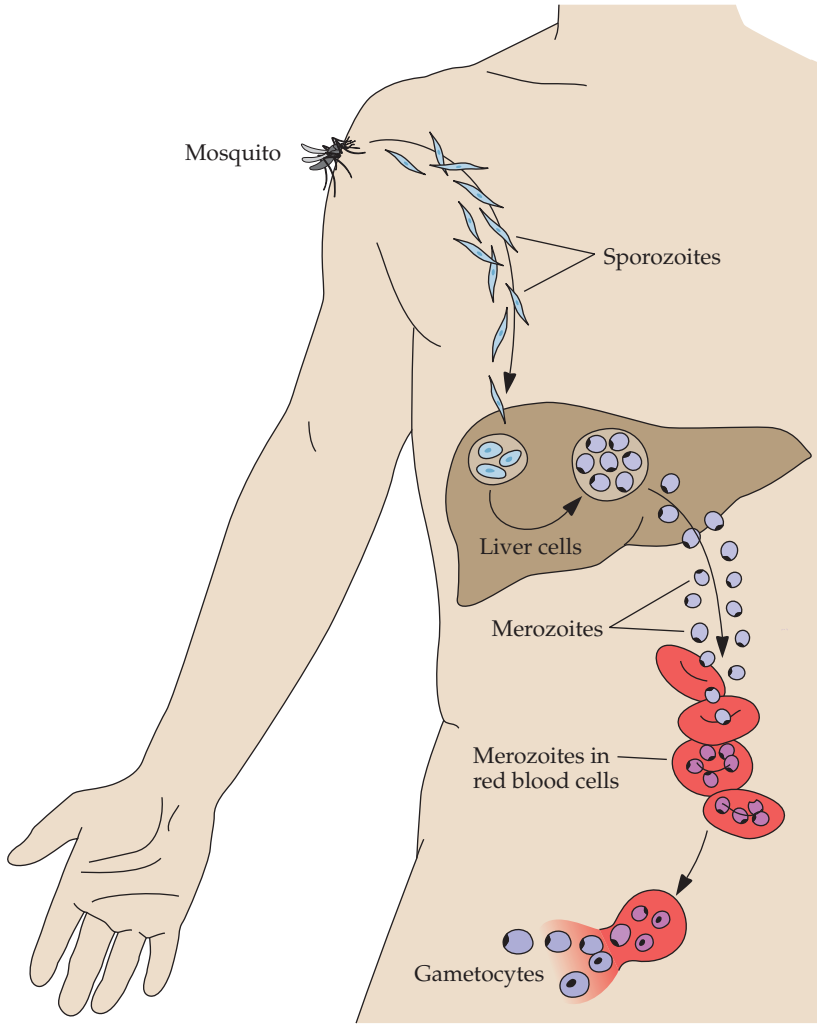


FIGURE 12.12 Infection of an individual with *Plasmodium falciparum* (a malaria-causing parasite) introduced by a mosquito.

were coated with *E. coli* plasmid DNA carrying an antigen gene under the transcriptional control of an animal virus promoter. A biolistic system was used to deliver the microprojectiles into cells in the ears of mice (see chapter 18 for a more detailed description of the biolistic system). Other workers introduced cloned cDNAs into mouse cells by injecting large amounts of the plasmid carrying the target DNA directly into the muscles of test animals. However, effective "genetic immunization" by direct injection into muscles (100 μg per mouse) requires 3 to 4 orders of magnitude more DNA than the biolistic delivery system (10 to 100 ng per mouse). One distinctive feature of genetic immunization is that the costly and time-consuming procedure of either purifying an antigen or creating a recombinant vaccine delivery vehicle is bypassed. Moreover, proteins produced by this procedure are more likely to be correctly posttranslationally modified than are proteins that are produced by different host organisms.

An advantage of genetic immunization, besides bypassing the need for purified protein antigens, is that it can trigger a response against only the

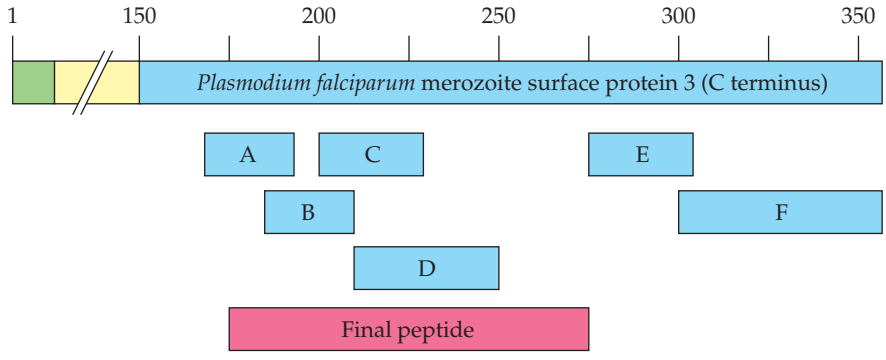


FIGURE 12.13 Schematic representation of *Plasmodium falciparum* merozoite surface protein 3 and peptides corresponding to portions of the C terminus. The peptides, labeled A to F, are drawn to scale, with the numbers above the whole protein indicating the amino acid number (counting from the N terminus). The "final peptide" is currently being tested in clinical trials for efficacy as a malaria vaccine.

protein encoded on the plasmid and not against the plasmid itself. In addition, when plasmid DNA is introduced into a mammalian system, only those genes (or cDNAs) under the control of eukaryotic regulatory signals will be transcribed and translated. Antibiotic resistance genes for maintaining the plasmid in *E. coli* will not be transcribed or translated, and the same vector can be used to deliver different proteins to an individual at the same time, or the administration of the same gene can be repeated a number of times.

The feasibility of genetic immunization has been examined in detail. In one series of experiments, mice were injected in the quadriceps of both legs with an *E. coli* plasmid carrying the cDNA for influenza A virus nucleoprotein under the transcriptional control of either a Rous sarcoma virus or a cytomegalovirus promoter. Although the expression of the nucleoprotein was too low to detect, nucleoprotein-specific antibodies were observed in the blood of the test mice 2 weeks after the initial injection. In comparison to control mice, the nucleoprotein-injected mice were significantly protected against the lethal effects of influenza virus infection (Fig. 12.14). Moreover, the nucleoprotein-injected mice were also protected against a different strain of influenza virus. This cross-protection is in sharp contrast to traditional influenza virus vaccines, which are directed against surface antigens of the virus, so that each vaccine is specific to a single strain of

TABLE 12.3 Advantages of genetic immunization over conventional vaccines

Cultivation of dangerous agents is not required.
Since genetic immunization does not utilize any viral or bacterial strains, there is no chance that an attenuated strain will revert to virulence.
Since no organisms are used, attenuated organisms that many cause disease in young or immunocompromised animals are not a problem.
Approach is independent of whether the microorganism is difficult to grow or attenuate.
Production is inexpensive because protein does not need to be produced or purified.
Storage is inexpensive because of the stability of DNA.
One plasmid could encode several antigens/vaccines, or several plasmids could be mixed together and administered at the same time.

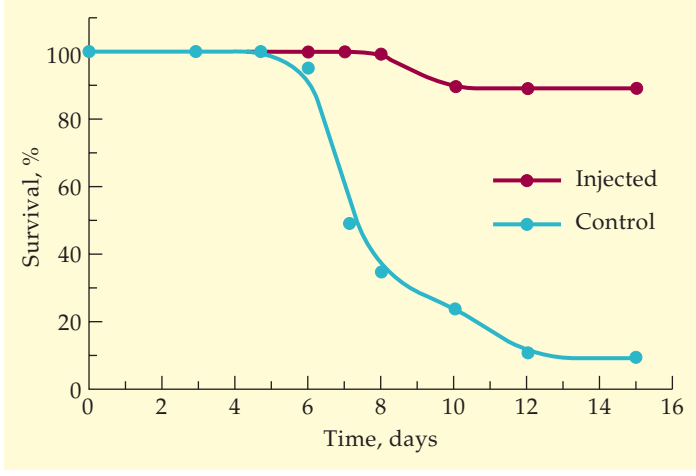


FIGURE 12.14 Survival of DNA-immunized mice. Injected mice were immunized with DNA that contained the influenza A virus nucleoprotein gene under the control of the Rous sarcoma virus promoter on an *E. coli* plasmid. The control mice were injected with plasmid DNA only. The *x* axis represents the number of days after the animals were challenged with the live influenza virus.

influenza virus. In addition, traditional vaccines work only as long as the antigens on the surface of the virus do not change. Unfortunately, the genes for the surface antigens mutate at a high rate, which creates significant differences among strains. Although core components of the virus, such as the nucleoprotein, are relatively invariant, they can activate the immune system by a mechanism that is different from that of surface antigens.

The fate of the introduced DNA is not known, and it could have the undesirable effect of integrating into the genome of the host cell, possibly disrupting an important gene. However, this risk is currently considered to be extremely low. It is more likely that the DNA will exist for a short period as a nonreplicating extrachromosomal element before it is degraded. To date, genetic immunization has been used primarily to induce immune responses in animals, and to a more limited extent in humans, against a number of pathogenic organisms, including influenza A virus, human immunodeficiency virus (HIV) type 1, bovine herpesvirus 1, rabies virus, *Plasmodium* species (which cause malaria), hepatitis B virus, hepatitis C virus, bovine rotavirus, bovine respiratory syncytial virus, pseudorabies virus, FMDV, Newcastle disease virus, *Clostridium tetani* (which causes tetanus), and *Mycobacterium tuberculosis* (which causes tuberculosis). Several human clinical trials using DNA vaccines are currently ongoing.

One of the problems with the use of DNA vaccines in large animals and humans compared to mice is that the transfection efficiency of introduced plasmid DNA is often insufficient to generate a protective immune response. One approach to deliver foreign DNA to animal cells utilizes biodegradable microscopic (0.3- to 1.5-μm) polymeric particles with a cationic surface that binds the plasmid DNA (Fig. 12.15). Plasmid DNA is bound to the surfaces of these "microparticles" and is slowly released over a period of 2 to 3 weeks after inoculation of an animal. Using microparticles, it was possible to achieve the same biological effect as with naked DNA with about 250-fold less DNA, demonstrating the potential of this approach. In addition, the level of antibodies induced by the expression of

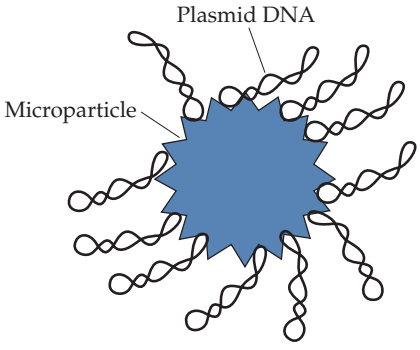


FIGURE 12.15 Schematic representation of the binding of plasmid DNA to the cationic surface of a polymeric microparticle.

plasmid-encoded genes bound to microparticles was significantly enhanced by (i) addition of the vaccine adjuvant aluminum phosphate and (ii) the use of nanoparticles 0.05 μm in diameter that were coated with poly-L-lysine. In contrast to naked DNA, DNA bound to microparticles induced potent cytotoxic T-lymphocyte responses at a low dose.

To date, most DNA vaccines have been delivered either by intramuscular or intradermal injection. Although these vaccines can induce a potent immune response, they do not induce mucosal immunity. Mucosal immunity can prevent pathogens from entering the body, while systemic immunity deals with pathogens only once they are inside the body. This is an important consideration because mucosal surfaces, including the respiratory, intestinal, and urogenital tracts, are the major sites of transmission of many infectious diseases. However, because of the protective barriers of the mucosal surfaces, traditional antigen-based vaccines are largely ineffective unless they are administered with specific agents that penetrate or bind to the mucosa, i.e., mucosal adjuvants.

Mucosal immunity induces a separate and distinct response from systemic immunity. The antibodies produced as part of the mucosal immune response restrict not only mucosal pathogens, but also microorganisms that initially colonize mucosal surfaces and then cause systemic disease. Many mucosal vaccines are live attenuated organisms that infect mucosal surfaces and are effective at inducing mucosal responses. Of these, oral polio vaccines and both attenuated *Salmonella enterica* serovar Typhi Ty21a and *Vibrio cholerae* vaccines are licensed for use in humans.

DNA vaccines that are designed for delivery to mucosal surfaces are similar in principle to those used for intramuscular or intradermal delivery. To increase plasmid uptake and decrease its subsequent degradation, various methods of formulating DNA have been tried. For example, cationic (positively charged) liposomes have been used to deliver DNA (which has a negatively charged phosphate backbone) to the respiratory tract, and DNA entrapment in biodegradable microparticles has been used for the oral delivery of foreign DNA. Moreover, to improve the potency of DNA vaccines for humans, a number of strategies have been devised, including using plasmids that, in addition to encoding a target gene, also express a cytokine(s), such as interleukin-2 (IL-2), IL-10, or IL-12 (which can act as an intercellular mediator in the generation of an immune response).

A range of systems, including liposomes, live vectors (bacteria and viruses), and a wide range of adjuvants that increase the immune response (bacterial toxins, carboxymethylcellulose, lipid derivatives, aluminum salts, and saponins), have been tried for delivery of DNA to different cell types. Of necessity, various optimization strategies are tested in mice before they are tried on larger animals and then on humans, with no guarantee that an approach that works well in mice will also be a successful strategy in humans. Nevertheless, given the many perceived advantages of genetic immunization over the use of conventional vaccines (Table 12.3), this has become a very active area of research. For example, electroporation has been used to increase the transfection of DNA encoding target antigens. With this approach, DNA is injected intramuscularly, and the skeletal muscle is immediately electrically stimulated with a pulse generator. Despite the fact that this procedure, which causes some patient discomfort, results in local tissue injury and inflammation, it is tolerated by patients without the need for any anesthesia, and there do not appear to be any long-term negative side effects to delivering DNA in this way. Most likely,

the electrical pulse increases the transfection efficiency of the added DNA, and it is becoming a method of choice for clinically administering DNA vaccines.

A modified strain of the invasive bacterium *Shigella flexneri* has been developed to facilitate the delivery of DNA into animal cells for genetic immunization (Fig. 12.16). *Shigella* can enter animal epithelial cells, escaping the phagocytic vacuole, and the bacterium can direct plasmid DNA to the nucleus of the host cell, where, if the introduced gene(s) contains a eukaryotic promoter, it is transcribed. *Shigella* is normally a pathogenic organism and would not be an acceptable DNA delivery system. Therefore, to use *Shigella*, it was first necessary to construct a nonpathogenic version of the wild-type organism by (1) engineering the bacterium to be toxin deficient and (2) making a deletion mutation in the *Shigella asd* gene, which encodes the enzyme aspartate β-semialdehyde dehydrogenase. This enzyme is normally involved in the synthesis of the bacterial cell wall constituent diaminopimelic acid; therefore, the mutant cannot grow unless diaminopimelic acid is added to the growth medium. *Shigella* strains with the *asd* mutation can invade animal epithelial cells and deliver their plasmid DNA; however, once present, the *Shigella* cells are unable to proliferate.

Determination of the safety of using *Shigella* as a vector for the delivery of DNA to animal cells must await the results of human trials; however, the results of experiments with guinea pigs are promising. The greatest potential advantage of this approach is that with the *Shigella* system, DNA for

MILESTONE

Protection against Foot-and-Mouth Disease by Immunization with a Chemically Synthesized Peptide Predicted from the Viral Nucleotide Sequence

J. L. Bittle, R. A. Houghten, H. Alexander, T. M. Shinnick, J. G. Sutcliffe, R. A. Lerner, D. J. Rowlands, and F. Brown
Nature **298:**30–33, 1982

Since Jenner developed the first vaccine over 200 years ago, most human vaccines against viral diseases have included partially killed or attenuated preparations of the disease-causing, or a similar nonpathogenic, version of the virus. While this approach has been undeniably effective and prevents the spread of a number of viral diseases, it is clearly limited. For example, not all viruses can be grown in culture, which precludes the development of vaccines against these viruses; production of traditional vaccines is expensive and potentially dangerous; and not all viral diseases are preventable through the use of these traditional vaccines. With the advent of molecular biotechnology, alternative strategies were examined for developing safer, less expensive, and more effective vaccines that would not have the limitations of using whole viruses, killed or attenuated, as vaccines. Since vaccines immunize individuals by priming their immune systems, it was thought that for some viruses short synthetic peptides might elicit the same antibody response as the antigenic determinants normally found on the external surface of the virus. These peptides were designed with the identical linear sequence of amino acids that made up the viral antigenic determinant in the first place. Of course, this approach could be expected to work only when the amino acids of an antigenic region (epitope) were contiguous.

Bittle et al. isolated and characterized the viral RNA for FMDV and then determined the sequence of VP1, the major antigenic protein of the virus. Based on other experiments, they reasoned that the major antigenic determinants on VP1 would probably be found at either the N- or C-terminal end of the protein. They then chemically synthesized a series of peptides based on the amino acid sequences of the N and C termini, chemically linked these peptides to carrier proteins, and then used them to inoculate rabbits and guinea pigs. With peptides from the C-terminal end of VP1 as antigens, the treated animals synthesized antibodies that protected them against disease from whole foot-and-mouth disease virus. This work established the principle that a protein domain(s) is sufficient to induce antibodies that can neutralize intact virus particles, and hence a new type of vaccine, which after the initial stages of development does not utilize or depend upon the disease-causing virus, is possible.

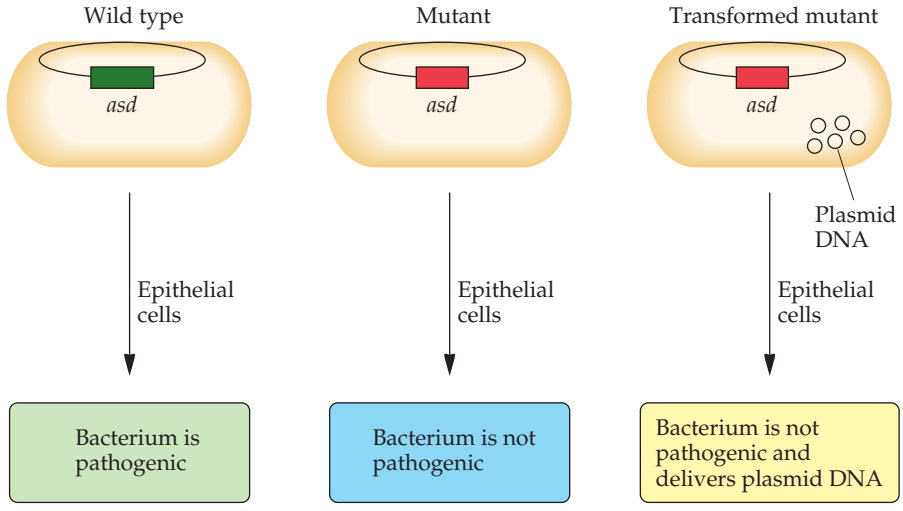
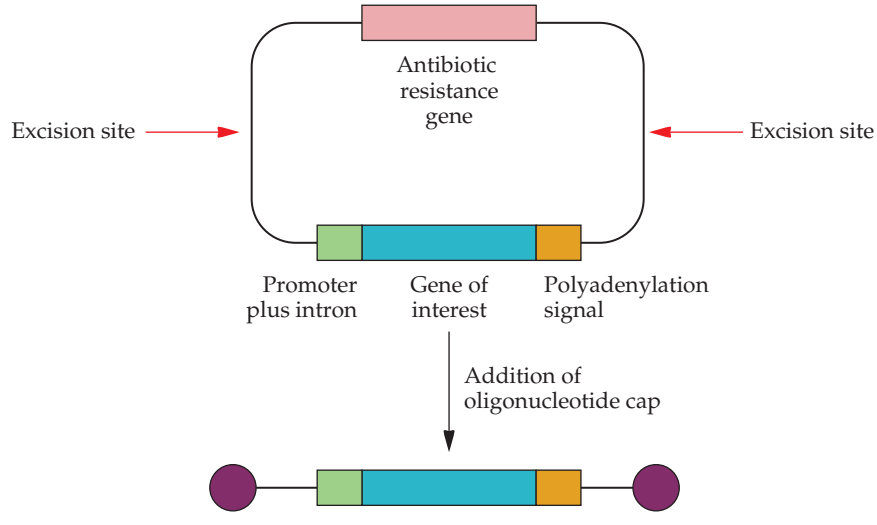


FIGURE 12.16 Use of nonpathogenic *S. flexneri* to deliver foreign DNA to mammalian epithelial cells. A strain of *Shigella* with a deletion mutation in the *asd* gene, which encodes the enzyme β-semialdehyde dehydrogenase, is unable to proliferate and can be used as a live vector.

vaccination may be delivered orally, greatly simplifying the delivery of a variety of vaccines.

Some limitations of plasmid-based vaccines are (1) the necessity for strong promoters that function in vivo to selectively transcribe the introduced DNA, (2) low levels of foreign-gene expression resulting from differences in codon usage between the introduced gene (often of viral, bacterial, or parasitic origin) and the animal being inoculated, and (3) the presence of antibiotic resistance genetic marker genes on the plasmid vector. To avoid the use of antibiotic resistance marker genes, researchers have developed a series of *m*inimalistic *i*mmunogenically *d*efined *g*ene *e*xpression (MIDGE) vectors (Fig. 12.17). Following insertion of the gene of interest into a MIDGE

FIGURE 12.17 Use of a MIDGE vector to produce a capped linear DNA sequence containing the gene of interest, a promoter, an intron (which facilitates expression of the gene of interest), and a polyadenylation signal.



vector, the antibiotic resistance gene is excised from the vector, and an oligonucleotide cap is added specifically to both ends of the linearized DNA that contains only a promoter/intron, the gene of interest, and a polyadenylation signal. The remaining portion of the plasmid is degraded by exonucleases. The capped ends are resistant to exonuclease digestion. The purified, capped linear MIDGE vector is then used directly for transfection. These vectors have been used successfully as a substitute for plasmid vectors.

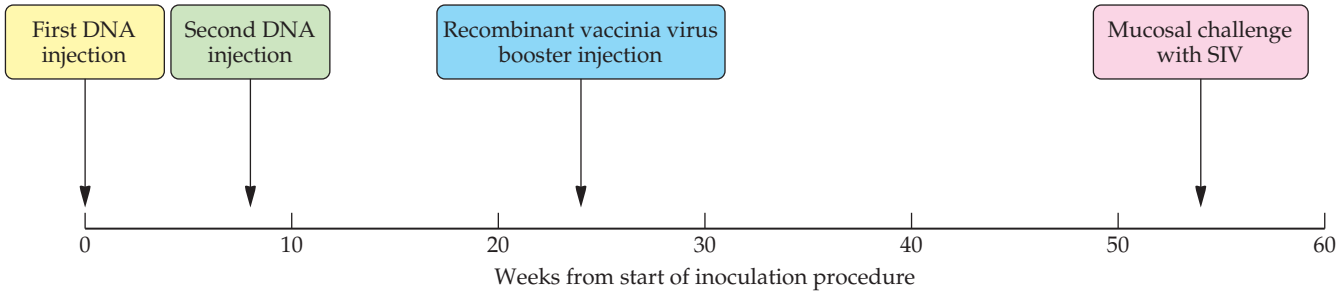
Vaccination of rhesus macaques (monkeys) with DNA encoding simian immunodeficiency virus proteins, followed by a booster with a modified vaccinia virus that encoded many of the same proteins, protected the monkeys against infection by simian immunodeficiency virus (Fig. 12.18). The DNA that was injected at 0 and 8 weeks expressed the simian immunodeficiency virus proteins Gag, Pol, Vif, Vpx, and Vpr, as well as the HIV type 1 proteins Env, Tat, and Rev. The recombinant vaccinia virus expressed the simian immunodeficiency virus proteins Gag and Pol and the HIV protein Env, all under the control of vaccinia virus promoters, and was administered at 24 weeks. The protection against simian immunodeficiency virus at 7 months after the booster with the recombinant modified vaccinia virus was much greater than protection with either treatment by itself. This procedure also confers immunity against a mucosal viral challenge. This feature is important because the site of entry of the virus into the simian host is effectively blocked. It is also noteworthy that the immunity was maintained for a long time.

Dental Caries

The gram-positive, facultatively anaerobic bacteria *Streptococcus mutans* and *Streptococcus sobrinus* are considered to be the primary causative agents of dental caries (tooth decay). These organisms colonize tooth surfaces and metabolize sucrose to produce lactic acid, which causes the tooth enamel to become vulnerable to decay. Sucrose is also used to produce a sticky, extracellular, dextran-based polysaccharide (glucan) that facilitates *Streptococcus* cells' adhering to one another and to tooth surfaces, forming plaque. It is the combination of plaque and acid that leads to tooth decay.

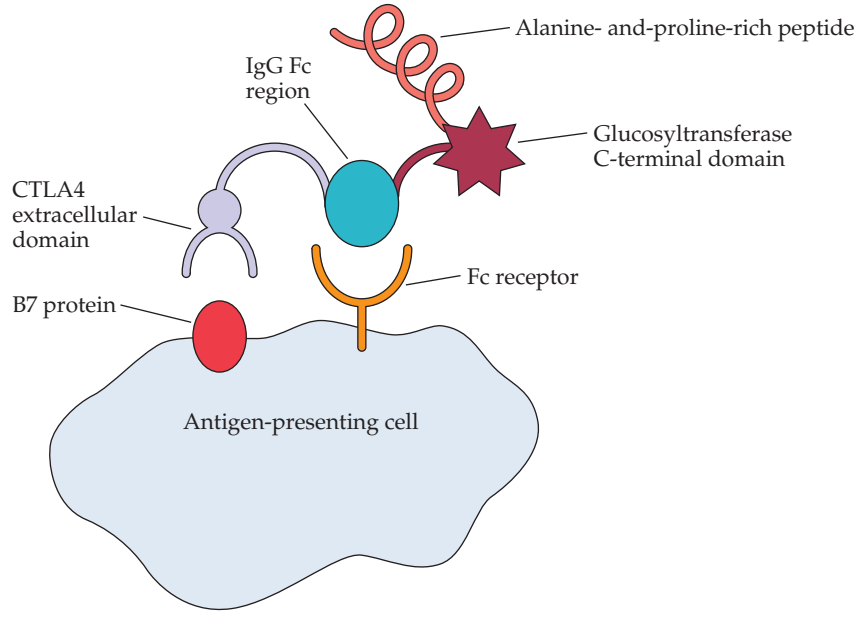
Two regions of one of the adhesion proteins found on the surfaces of *S. mutans* and *S. sobrinus* cells are important in the initial adherence of these bacteria to tooth surfaces: one sequence is rich in alanine residues, while the other is rich in proline residues. Another important component of the mechanism of tooth decay is the enzyme glucosyltransferase, which is responsible for the synthesis of glucan, an insoluble extracellular polymer

FIGURE 12.18 Vaccination regimen of rhesus monkeys with DNA containing simian immunodeficiency virus (SIV) genes and vaccinia virus carrying the same genes.



of glucose moieties. A DNA vaccine designed to prevent dental caries included the coding sequences for an alanine- and proline-rich peptide, as well as the C-terminal domain of a *Streptococcus* glucosyltransferase. This C-terminal domain is necessary for the binding of the glucosyltransferase to the bacterial cell surface. This DNA vaccine therefore encoded two separate peptides, both of which facilitate the binding of *Streptococcus* cells to the tooth surface. In an attempt to overcome the tendency of many DNA vaccines to induce only a weak immune response, the DNA vaccine construct contained two additional elements. First, the extracellular domain of cytotoxic T-lymphocyte antigen 4 (CTLA4), which binds to the B7 protein that is expressed on the surfaces of antigen-presenting cells, was included in the construct (Fig. 12.19). Second, the Fc region of an immunoglobulin G molecule, which can bind to the Fc receptor on the antigen-presenting cell, was included in this construct. The use of both of these peptides was designed to specifically target the multidomain fusion protein (Fig. 12.19) to immune system cells and thereby amplify the immune response and enhance the efficacy of the vaccine. In fact, rabbits immunized with this vaccine, either intranasally or intramuscularly, displayed a significantly enhanced, specific systemic and mucosal immune response compared to immunization with only the alanine- and proline-rich peptide fused to the C-terminal domain of glucosyltransferase. Moreover, it was subsequently shown that this DNA vaccine could provide significant protection against dental caries in rats that were challenged with *S. mutans* and *S. sobrinus*. Although the problem of delivering the DNA vaccine to humans still needs to be addressed so that its clinical efficacy can be tested, this is a very promising strategy that could be enormously beneficial to human populations.

FIGURE 12.19 Schematic representation of a multidomain protein encoded by a DNA vaccine. Two domains, the CTLA4 extracellular domain and the immunoglobulin G (IgG) Fc region, are designed to target the other two portions of the molecule, i.e., the glucosyltransferase C-terminal domain and the alanine-proline-rich peptide, to antigen-presenting cells.



Attenuated Vaccines

In some instances, genetic manipulation may be used to construct modified organisms (bacteria or viruses) that are used as live recombinant vaccines. These vaccines are either nonpathogenic organisms that have been engineered to carry and express antigenic determinants from a target pathogenic agent or engineered strains of pathogenic organisms in which the virulence genes have been modified or deleted. In these instances, as part of a bacterium or a virus, the important antigenic determinants are presented to the immune system with a conformation that is very similar to the form of the antigen in the disease-causing organism. Although successful in some cases, purified antigen alone often lacks the native conformation and elicits a weak immunological response.

Cholera

It is usually advantageous to develop a live vaccine, because they are generally much more effective than killed or subunit vaccines. The major requirement for a live vaccine is that no virulent forms be present in the inoculation material. With this objective in mind, a live cholera vaccine has been developed. Cholera, caused by the bacterium *V. cholerae,* is a fast-acting intestinal disease characterized by fever, dehydration, abdominal pain, and diarrhea. It is transmitted by drinking water contaminated with fecal matter. In developing countries, the threat of cholera is a real and significant health concern whenever water purification and sewage disposal systems are inadequate.

Since *V. cholerae* colonizes the surface of the intestinal mucosa, it was reasoned that an effective cholera vaccine should be administered orally and directed to this structure. With this in mind, a strain of *V. cholerae* was created with part of the coding sequence for the A_1 peptide deleted. This strain cannot produce active enterotoxin; therefore, it is nonpathogenic and is a good candidate for a live vaccine.

Specifically, in this experiment, a tetracycline resistance gene was incorporated into the A_1 peptide DNA sequence on the *V. cholerae* chromosome. This insertion inactivated the A_1 peptide activity and also made the strain resistant to tetracycline. Although the A_1 peptide sequence has been disrupted, the strain is not acceptable as a vaccine because the inserted tetracycline resistance gene can excise spontaneously, thereby restoring enterotoxin activity. Consequently, it was necessary to engineer a strain carrying a defective A_1 peptide sequence that could not revert (Fig. 12.20).

1. A plasmid containing the cloned DNA segment for the A_1 peptide was digested with the restriction enzymes ClaI and XbaI, each of which cut only within the A_1 peptide-coding sequence of the insert.
2. To recircularize the plasmid, an XbaI linker was added to the ClaI site and then cut with XbaI.
3. T4 DNA ligase was used to join the plasmid at the XbaI sites, thereby deleting a 550-base-pair segment from the middle of the A_1 peptide-coding region. This deletion removed 183 of the 194 amino acids of the A_1 peptide.
4. Then, by conjugation, the plasmid containing the deleted A_1 peptide-coding sequence was transferred into the *V. cholerae* strain carrying the tetracycline resistance gene within its A_1 peptide DNA sequence.

5. Recombination (a double crossover) between the remaining A_1 coding sequence on the plasmid and the tetracycline resistance gene-disrupted A_1 peptide gene on the chromosome replaced the chromosomal A_1 peptide-coding sequence with the homologous segment on the plasmid carrying the deletion.
6. After growth for a number of generations, the extrachromosomal plasmid, which is unstable in *V. cholerae,* was spontaneously lost.
7. Cells with an integrated defective A_1 peptide were selected on the basis of their tetracycline sensitivity. The desired cells no longer had the tetracycline resistance gene but carried the A_1 peptide sequence with the deletion.

A stable strain with an A_1 peptide sequence containing a deletion was selected in this way. This strain did not produce active enterotoxin but nevertheless retained all the other biochemical features of the pathogenic form of *V. cholerae;* that is, *V. cholerae* with an A_1 peptide containing a deletion is a good vaccine candidate because the bacterium that synthesizes only the A_2 and B peptides is as immunogenic as the native bacterium. When this strain was evaluated in clinical trials to test its effectiveness as a cholera vaccine, the results were equivocal. While the vaccine conferred nearly 90% protection against diarrheal disease in volunteers, it induced side effects in some of those who were tested. This strain may require modification at another chromosomal locus before it can be used as a vaccine.

Salmonella Species

Other attempts to engineer nonpathogenic strains of pathogenic bacteria that could be used as live vaccines have involved deletions in chromosomal regions that code for independent and essential functions. At least two deletions are preferred, because the probability that both sets of functions can be simultaneously reacquired is very small. It is assumed that a "doubly deleted" strain would have a limited ability to proliferate when it is used as a vaccine, thereby curtailing its pathogenicity while allowing it to stimulate an immunological response.

Strains of the genus *Salmonella* cause enteric fever, infant death, typhoid fever, and food poisoning. Therefore, an effective vaccine against these organisms is needed. Deletions in a number of different genes have been used to attenuate various *Salmonella* strains (Table 12.4). These mutations can be grouped into three basic categories: mutations in (1) biosynthetic genes, (2) regulatory genes, and (3) genes involved in virulence. In addition, strains with more than one deletion have been constructed. For example, one double-deletion strain has deletions in the *aro* genes, which encode enzymes involved in the biosynthesis of aromatic compounds, and in the *pur* genes, which encode enzymes involved in purine metabolism. These double-deletion strains, which can be grown on a complete and enriched medium that supplies the missing nutrients, generally establish only low-level infections, since their host cells contain only a very low level of the metabolites that they require for growth. Typically, their virulence is reduced by 100-fold or more. These attenuated *Salmonella* strains are effective oral vaccines for mice, sheep, cattle, chickens, and humans.

Deletion of the *dam* gene, which encodes DNA methylase, may be a highly effective approach to produce avirulent *Salmonella* strains. The *dam* gene is a master switch that regulates the expression of 20 to 40 different *Salmonella* regulatory proteins. Thus, when mice were immunized with

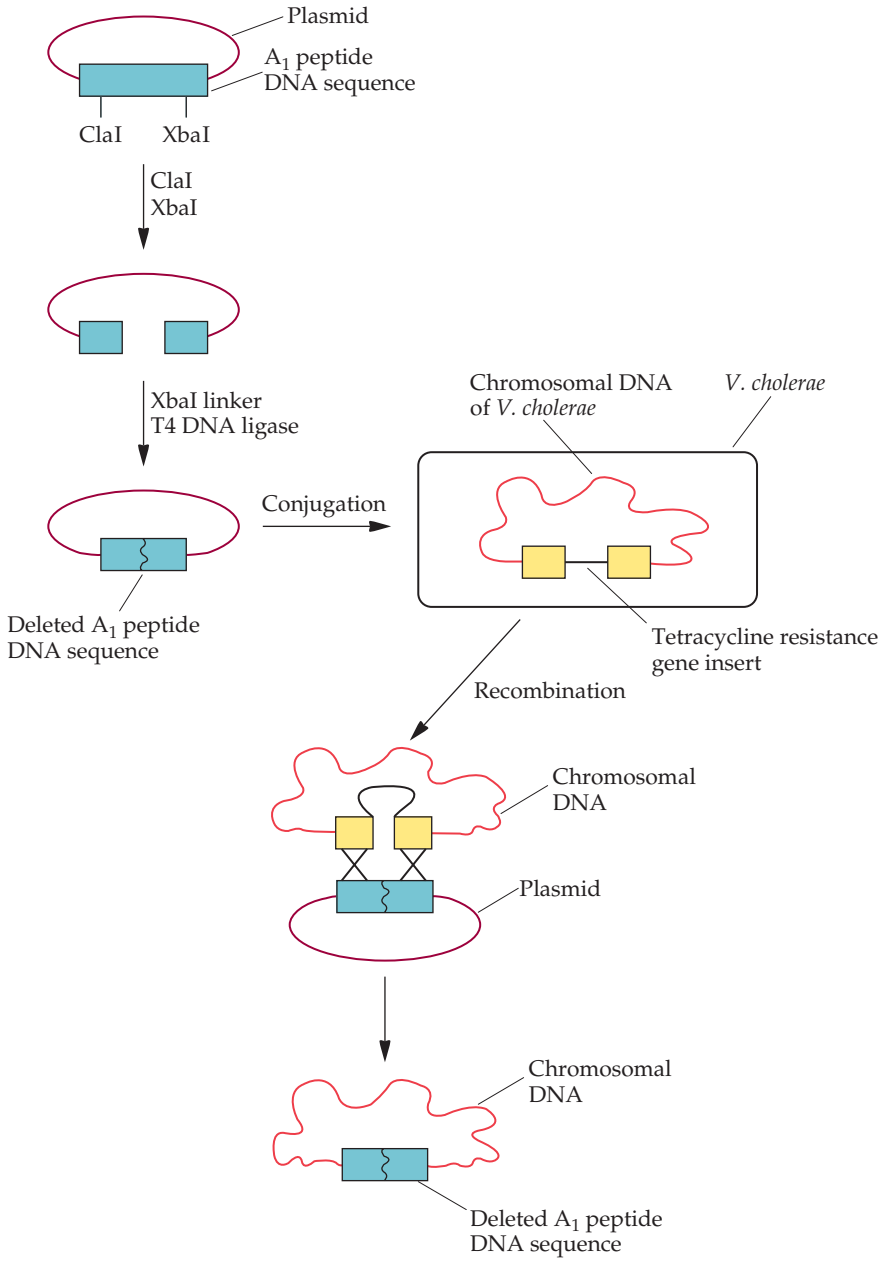


FIGURE 12.20 Strategy for deleting part of the cholera toxin A_1 peptide DNA sequence from a strain of *V. cholerae*. Note that the tetracycline resistance gene is introduced into the gene for the A1 peptide as part of a transposon. This construct no longer makes A1 peptide but cannot be used as a vaccine because it is possible for the transposon to be excised, with the result that A1 synthesis and pathogenicity are restored.

Dam-negative strains of *Salmonella,* they tolerated up to 10,000 times the normally lethal dose. Generally, pathogenic bacteria turn on many of their genes as briefly as possible to avoid detection and attack by the host's immune system. However, with Dam-negative strains, these genes are expressed for much longer periods, making it easier for the host immune system to detect and destroy the invading bacteria. Because many other

TABLE 12.4 Deleted genes and their functions in the development of attenuated strains of *Salmonella* spp.

Deleted gene	Gene function
galE	Synthesis of lipopolysaccharide; decrease toxicity from galactose
aroA, *aroC*, or *aroD*	Synthesis of chorismate, an aromatic amino acid precursor and a PABA precursor; PABA is involved in the synthesis of iron chelators
purA or *purE*	Synthesis of purines
asd	Peptidoglycan and lysine biosynthesis
phoP and *phoQ*	Regulation of acid phosphatases and genes necessary for survival in the microphage
cya	Encodes adenylate cyclase, which is involved in camp synthesis
crp	Enclosed camp receptor; regulates expression of proteins involved in transport and breakdown of carbohydrates and amino acids
cdt	Involved in tissue colonization by the bacterium
dam	Encodes DNA methylase; appears to be a master switch for 20–40 different virulence genes
htrA	Enclosed a stress-induced polypeptide; result in significantly reduced persistence in human tissues

cAMP, cyclic AMP; PABA, *p*-aminobenzoic acid.

gut-colonizing bacteria have *dam* genes, if this approach with *Salmonella* turns out to be as effective as is expected, it may be possible to utilize a similar protocol with a range of pathogenic bacteria.

Leishmania Species

Although the human immune system can respond to infections by protozoan parasites of the genus *Leishmania,* it has been difficult to develop an effective vaccine against these organisms. Attenuated strains of *Leishmania* are sometimes effective as vaccines; however, they often revert to virulence. Also, the attenuated parasite can persist for long periods in an infected but apparently asymptomatic individual. Such individuals can act as reservoirs for the parasite, which can be transferred to other people by an intermediate host. To overcome these problems, an attenuated strain of *Leishmania* that is unable to revert to virulence was created by targeted deletion of an essential metabolic gene, such as the one encoding dihydrofolate reductase–thymidylate synthase. In one of these attenuated strains, *Leishmania major* E10-5A3, the two dihydrofolate reductase–thymidylate synthase genes that are present in wild-type strains were replaced with the genes encoding resistance to the antibiotics G-418 and hygromycin. For growth in culture, it is necessary to add thymidine to the medium that is used to propagate the attenuated (but not the wild-type) strain. In addition, unlike the wild type, the attenuated strain is unable to replicate in macrophages in tissue culture unless thymidine is added to the growth medium (Fig. 12.21). Importantly, the attenuated strain survives for only a few days when inoculated into mice; in that time, it does not cause any disease. Moreover, this period is sufficient to induce substantial immunity against *Leishmania* in BALB/c mice after administration of the wild-type parasite (Fig. 12.22). Since the attenuated parasite did not establish a persistent infection or

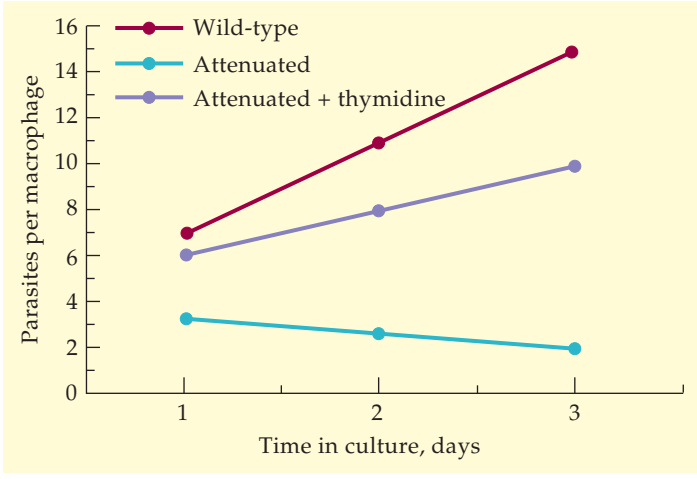


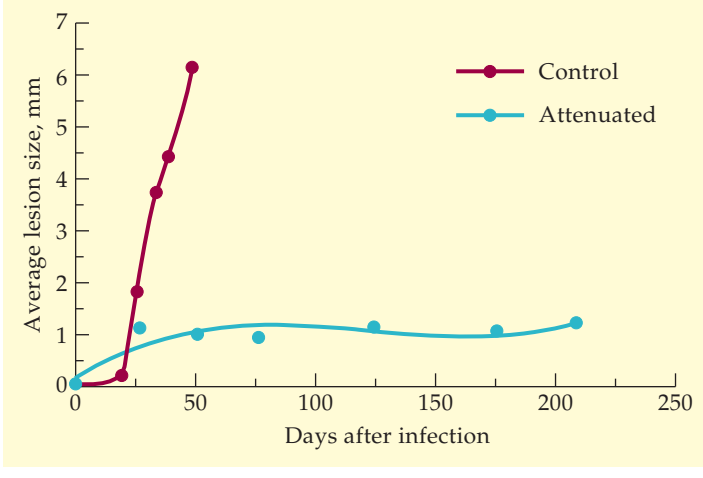
FIGURE 12.21 Proliferation of wild-type and attenuated *L. major* in mouse macrophages. At time zero, macrophages were infected with the same amount of stationary-phase *L. major*. The wild-type parasite and the attenuated parasite in the presence of thymidine were able to proliferate, while the attenuated strain did not proliferate in the absence of thymidine in the medium. Adapted from Titus et al., *Proc. Natl. Acad. Sci. USA* **92:**10267–10271, 1995.

cause disease, even in the most susceptible strains of mice tested, it is considered to be a strong candidate vaccine. Following additional experiments with animals, it should be possible to test whether this attenuated parasite is effective as a vaccine in humans.

Herpes Simplex Virus

As with other pathogenic organisms that have been developed as live vaccines, portions of the HSV genome have been deleted. Initially, it was

FIGURE 12.22 Immunity to virulent *L. major* induced in BALB/c mice inoculated with attenuated *L. major*. At time zero, mice that were previously inoculated with attenuated *L. major* were challenged with virulent *L. major,* and the sizes of the parasite-induced lesions were measured at various times. Control mice were not vaccinated with attenuated *L. major*. Adapted from Titus et al., *Proc. Natl. Acad. Sci. USA* **92:**10267–10271, 1995.



thought that a strong immune response could be obtained only if the virus was able to replicate. However, several vaccines based on nonreplicating viruses induce an immune response. Developing an avirulent HSV is important, because subunit vaccines so far have been unsuccessful in inducing immunity against the virus. To prepare a safe and efficacious live HSV vaccine, two deletions at different locations in the viral genome were generated independently and then combined to form a double-deletion virus. This strain is unable to proliferate in host cells, and the probability that both sets of functions can be simultaneously reacquired is very small. This replication-defective strain induces protective immunity that can reduce acute viral shedding and latent infection.

Vector Vaccines

Vaccines Directed against Viruses

Vaccinia virus, in the form of a live vaccine, has led to the eradication of smallpox globally. Vaccinia virus is a member of the poxvirus family. This completely sequenced virus has a double-stranded DNA genome that contains 187 kilobase pairs (kb) and encodes approximately 200 different proteins. Vaccinia virus DNA replicates within the cytoplasm of infected cells. Cytoplasmic, rather than nuclear, replication and transcription are possible because vaccinia virus DNA contains genes for DNA polymerase, RNA polymerase, and the enzymes to cap, methylate, and polyadenylate messenger RNA (mRNA). Thus, if a foreign gene is inserted into the vaccinia virus genome under the control of a vaccinia virus promoter, it will be expressed independently of host regulatory and enzymatic functions. The virus can infect humans and many other vertebrates, as well as invertebrates.

In addition to having a broad host range, vaccinia virus is well characterized at the molecular level, is stable for years after lyophilization (freeze-drying), and is usually a benign virus. For these reasons, it is a strong candidate as a vector vaccine. The function of a vector vaccine is to deliver and express cloned genes encoding antigens that elicit neutralizing antibodies against pathogenic agents. Unfortunately, the vaccinia virus genome is very large and lacks unique restriction sites. Therefore, it is not possible to insert additional DNA directly into the viral genome. Of necessity, the genes for specific antigens must be introduced into the viral genome by in vivo homologous recombination.

1. The DNA sequence coding for a specific antigen, such as HBcAg, is inserted into a plasmid vector immediately downstream of a cloned vaccinia virus promoter and in the middle of a nonessential vaccinia virus gene, such as the gene for the enzyme thymidine kinase (Fig. 12.23A).
2. This plasmid is used to transfect thymidine kinase-negative animal cells in culture, usually chicken embryo fibroblasts, that have previously been infected with wild-type vaccinia virus, which produces a functional thymidine kinase.
3. Recombination between DNA sequences that flank the promoter and the neutralizing antigen gene on the plasmid and the homologous sequences on the viral genome results in the incorporation of the cloned gene into the viral DNA (Fig. 12.23B). Although the recombination event is rare, the absence of thymidine kinase

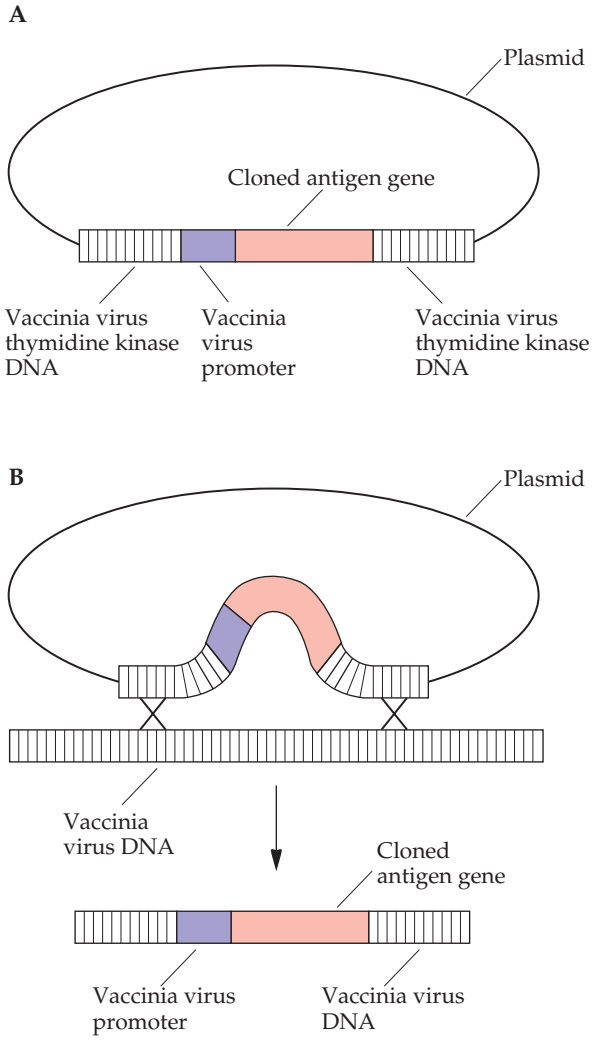


FIGURE 12.23 Method for the integration into vaccinia virus of a gene whose protein product, generally a viral antigen, elicits an immunological response. **(A)** Plasmid carrying a cloned expressible antigen gene. **(B)** A double-crossover event results in the integration of the antigen gene into vaccinia virus DNA.

activity in the host cells and the disruption of the thymidine kinase gene in the recombined virus render the host cells resistant to the otherwise toxic effects of bromodeoxyuridine. This selection scheme enriches for cell lines that carry a recombinant vaccinia virus.

4. The definitive selection of cells with a recombinant vaccinia virus is made by DNA hybridization with a probe for the antigen gene.

Since thymidine kinase-negative mutants of vaccinia virus arise spontaneously at a relatively high frequency of about 1 virus particle in 10^3 to 10^4, a selectable marker is often cotransferred with the target gene. This makes it much easier to distinguish a spontaneous thymidine kinase mutant from a mutant deliberately generated by homologous recombination. In other words, a virus with a spontaneous mutation would not carry the selectable marker, whereas a virus that underwent homologous recombination would.

The *neo* gene, which encodes the enzyme neomycin phosphotransferase II and confers resistance to the kanamycin analogue G-418, is often used as the selectable marker. This gene, unlike some other selectable markers, is quite stable once it is inserted into the vaccinia virus genome.

To avoid disrupting any vaccinia virus genes or the necessity of screening for selectable markers, a novel system has been devised in which every recombinant virus that can form a plaque will contain and express the target gene. Wild-type vaccinia virus contains a gene, *vp37,* that is responsible for the formation of plaques when the virus is grown on an animal cell monolayer (Fig. 12.24A). Deleting the *vp37* gene and replacing it with an *E. coli* marker gene (Fig. 12.24B) creates a vaccinia virus mutant that does not form plaques after 2 to 3 days of growth in cell culture. Target genes are introduced into the mutant vaccinia virus by homologous recombination with a transfer vector that carries the *vp37* gene, as well as the target gene (Fig. 12.24C). If homologous recombination between the nonplaque-forming mutant and the transfer vector occurs, the viruses that can form plaques have acquired the *vp37* gene. Also, the target gene is inserted into the vaccinia virus genome, and the selectable marker gene is lost. Since the *vp37* gene has been deleted in the mutant vaccinia virus, it is impossible for this mutation to revert to the wild type. Therefore, every virus that forms a plaque carries the desired construct. This procedure is simple and straightforward, is applicable to the cloning and expression of any target gene, does not require any extra marker genes, and does not disrupt any vaccinia virus genes.

A number of antigen genes have been successfully inserted into the vaccinia virus genome and subsequently expressed in animal cells in culture. These antigens include rabies virus G protein, hepatitis B surface antigen, Sindbis virus surface proteins, influenza virus NP and HA proteins, vesicular stomatitis virus N and G proteins, and HSV glycoproteins. Several recombinant vaccinia virus vehicles have been shown to be effective vaccines. For example, a recombinant vaccinia virus that expresses the HSV-1 gD (glycoprotein D) gene prevents herpes infections in mice. Another recombinant vaccinia virus that expresses the rabies virus surface antigen gene was able to elicit neutralizing antibodies in foxes, which are major carriers of rabies in Europe, and has been used in the field for some time, including in an area of approximately 10,000 km^2 in Belgium. The vaccinia–rabies virus glycoprotein recombinant virus vaccine that is presently on the market (Raboral) is a live viral vaccine containing 10^8 plaque-forming units (PFU), or live viral particles, per dose. It is constructed by insertion of the DNA copy coding for glycoprotein G of a rabies virus strain into the thymidine kinase gene of a strain of vaccinia virus. Once the vaccine is ingested by a fox, the vaccinia virus begins to replicate and express rabies glycoprotein G, which stimulates the development of immune responses to the rabies glycoprotein. This results in the production of neutralizing antibodies against the rabies virus in the immunized foxes. This immunity typically lasts about 12 months in cubs and 18 months in adult animals.

The use of vector vaccines constructed from vaccinia virus also offers the possibility of vaccinating individuals against several different diseases with one treatment. This may be achieved by using a recombinant vaccinia virus carrying cloned genes encoding a number of different antigens.

The timing of the production of a foreign protein whose gene is carried in a vaccinia virus depends on whether a vaccinia virus promoter functions

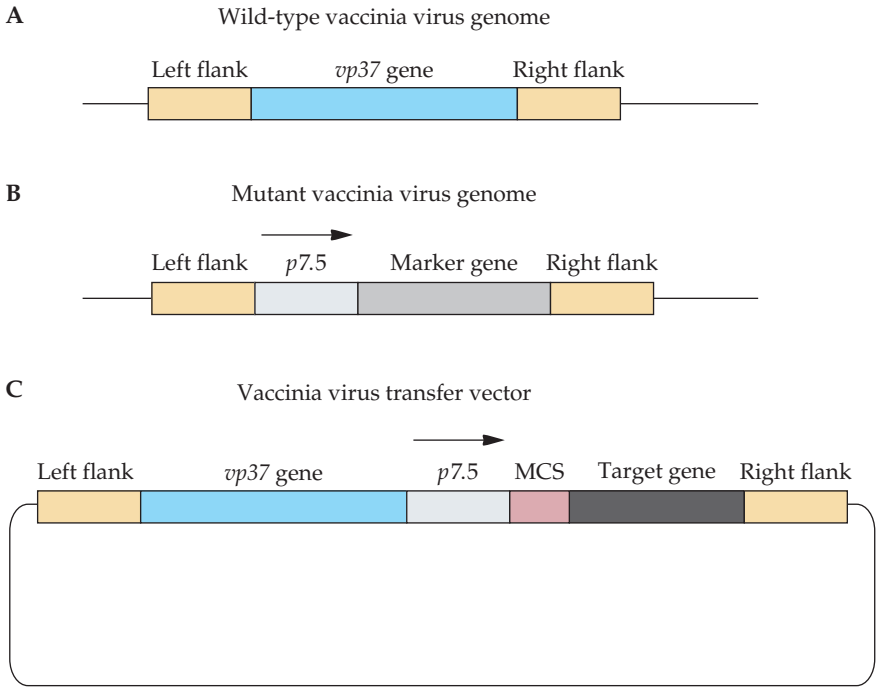


FIGURE 12.24 **(A)** Portion of a wild-type vaccinia virus genome that contains the *vp37* gene that is responsible for plaque formation in host cells. **(B)** Portion of a mutant vaccinia virus genome in which the *vp37* gene has been replaced by a marker gene. **(C)** Portion of a vaccinia virus transfer vector. "Left flank" and "right flank" refer to the DNA sequences that immediately precede and follow the *vp37* gene in the wild-type vaccinia virus genome. The native *vp37* promoter is part of the *vp37* gene sequence (not shown). MCS is a multiple cloning site with seven unique restriction enzyme sites. *p7.5* is a strong early/late vaccinia virus promoter. The target gene is inserted into the multiple cloning site. Subsequently, homologous recombination between the transfer vector (C) and the genomic DNA of the mutant virus (B) results in the replacement of the *E. coli* marker gene with the *vp37* gene, together with a target gene.

during the early or late phase of the infection cycle, and the strength of the promoter determines the amount of an antigen that is produced. For the most part, late promoters for an 11-kilodalton (kDa) protein (p11) and the cowpox virus A-type inclusion protein (pCAE) have been used to achieve high levels of foreign-gene expression. When genes encoding several different foreign proteins are inserted into one vaccinia virus, each is placed under the control of a different vaccinia virus promoter to avoid the possibility of homologous recombination between different portions of the virus genome that might cause the cloned genes to be lost.

A live recombinant viral vaccine has several advantages over killed virus or subunit vaccines. First, the virus can express the authentic antigen(s) in a manner that closely resembles a natural infection. Second, the virus can replicate within the host, thereby amplifying the amount of antigen that activates the release of antibodies from B cells (humoral response) and stimulates the production of T cells (cell-mediated immune response).

A disadvantage of using a live recombinant viral vaccine is that vaccination of an immunosuppressed host, such as an individual with AIDS, can lead to a serious viral infection. One way to avoid this problem may be to

insert the gene encoding human IL-2 into the viral vector. IL-2 enhances the response of the T cells of the immune system, enabling the recipient to limit the proliferation of the viral vector, and thereby decreases the possibility of an unwanted infection.

If the proliferation of vaccinia virus has deleterious effects in certain patients, it would be helpful to kill or inhibit it after vaccination. One approach is to create an interferon-sensitive vaccinia virus—wild-type vaccinia virus is relatively resistant to interferon—whose proliferation is curtailed. Such a virus vector would be susceptible to drug intervention if complications from vaccination with vaccinia virus vectors arose.

The basis of the resistance of vaccinia virus to interferon was not known until a vaccinia virus open reading frame (K3L) was found to encode a 10.5-kDa protein that has an amino acid sequence that is very similar to that of a portion of the 36.1-kDa host cell eukaryotic initiation factor 2a (eIF-2a). The N-terminal regions of both of these proteins contain 87 amino acids that are nearly identical. Moreover, this shared sequence contains a serine residue, amino acid 51, which in eIF-2a is normally phosphorylated by interferon-activated P1 kinase. When this serine residue in eIF-2a is phosphorylated in interferon-treated cells, protein synthesis, and therefore viral replication, is inhibited. Thus, vaccinia virus may avoid inhibition by interferon because the K3L protein acts as a competitive inhibitor of eIF-2a phosphorylation (Fig. 12.25). Therefore, deletion of all or a portion of the K3L gene from vaccinia virus should make the virus sensitive to interferon. A K3L-negative mutant of vaccinia virus was constructed by PCR mutagenesis of the K3L gene carried on a plasmid, followed by homologous recombination to replace the wild-type K3L sequence with the modified version. When the wild-type and mutant versions of vaccinia virus were tested for sensitivity to interferon, the mutant was 10 to 15 times more sensitive to interferon than was the wild-type version (Fig. 12.26). Reinsertion of the wild-type K3L sequence into the mutant virus restored the level of interferon sensitivity found in the wild type. This indicates that K3L is indeed involved in the interferon resistance phenotype of vaccinia virus. This work is an important step in the development of safer vaccinia virus vectors. Moreover, other interferon-resistant viruses may contain sequences comparable to K3L and therefore may be amenable to the construction of interferon-sensitive deletion mutants. Other, comparable approaches to the creation of attenuated strains of vaccinia virus have been developed. For example, a strain of vaccinia virus, constructed to have a mutation in the B8R gene, which encodes an interferon viroreceptor, is less pathogenic for mice than the parental viral strain.

Currently, several veterinary vaccinia virus-based vaccines have been licensed, and clinical studies to test their efficacies in preventing a number of human infectious diseases are under way. This technology is based in part on the development of an attenuated version of the vaccinia virus strain that had previously been used in the eradication of smallpox. To avoid any risk of the vector itself becoming a source of disease, some genetic information was removed from the virus genome so that the viral vector was highly attenuated. This attenuated virus has been used to express a number of viral antigens with the expectation that the recombinant virus would be an effective live vaccine. Protection has been achieved by cloning glycoproteins from porcine pseudorabies virus, hemagglutinin glycoproteins from equine influenza virus, a spike protein from the SARS virus, and a polyprotein of Japanese swine encephalitis virus and then vac-

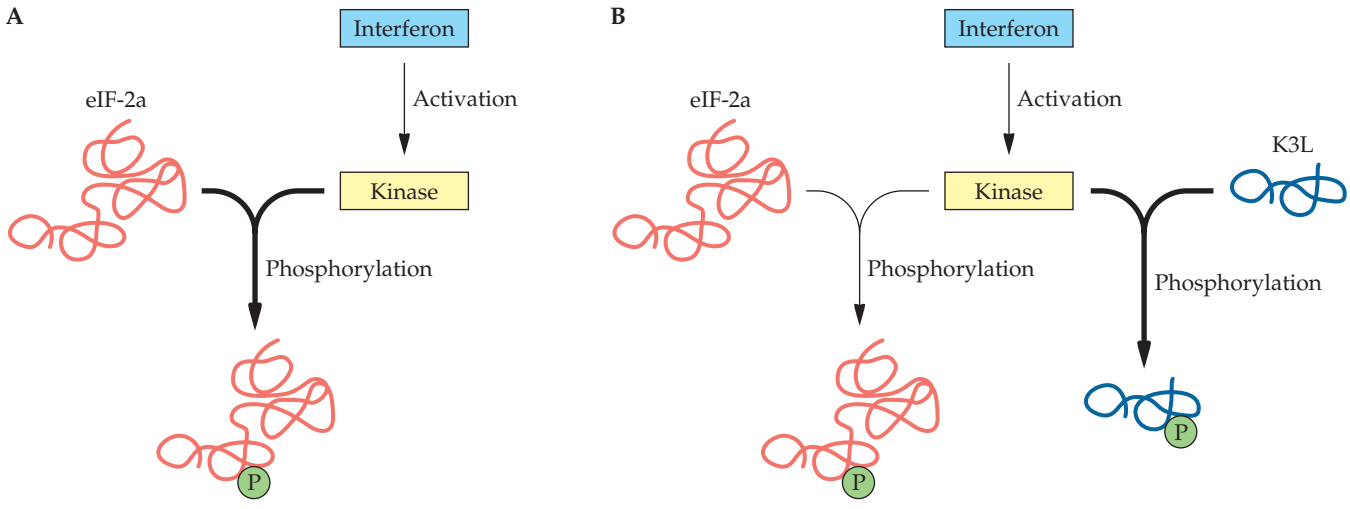


FIGURE 12.25 Competitive inhibition of the interferon-stimulated phosphorylation (inhibition) of eIF-2a by protein K3L, which is encoded by vaccinia virus and is nearly identical to a portion of eIF-2a. **(A)** In the presence of interferon, a kinase is activated that phosphorylates eIF-2a molecules and thereby prevents them from functioning. **(B)** When vaccinia virus protein K3L is also present, it is phosphorylated instead of the eIF-2a, so the eIF-2a remains active. The thickness of the arrows represents the relative flux through each pathway.

cinating humans to prevent transmission of these viruses. Based on the success of these attenuated vaccinia virus vaccines, it has been proposed that this virus be considered a general delivery system for a wide range of proteins.

For mass vaccination campaigns in developing countries, it would be advantageous to be able to deliver live vaccines in a simple, expeditious, and cost-effective manner. In addition, with mucosally transmitted pathogens, such as HIV, traditional vaccination routes may not induce mucosal immune responses sufficient to provide protective immunity. One possible alternative to traditional vaccination is aerosol immunization, which is potentially safer, easier, and less expensive to administer. To this end, researchers tested the abilities of two attenuated vaccinia virus-based vectors to be delivered effectively by aerosol immunization. In fact, it was found that aerosol delivery was both safe and effective, yielding long-lasting systemic and mucosal immune responses when delivered to rhesus macaques (monkeys). This approach still needs to be tested with humans; however, it could offer an effective means of inoculating large numbers of individuals in the future.

Although much of the work on the development of live viral vaccines has been done with vaccinia virus, other viruses, such as adenovirus, poliovirus, and varicella-zoster virus, are also being tested as potential vaccine vectors. Live attenuated poliovirus can be delivered orally, and such a mucosal vaccine, which is directed to receptors in the lungs or gastrointestinal tract, might also be useful against a range of diseases, including cholera, typhoid fever, influenza, pneumonia, mononucleosis, and rabies. However, the safety and efficacy of any apparently benign virus as a gene delivery and expression system must be firmly established before clinical trials are undertaken.

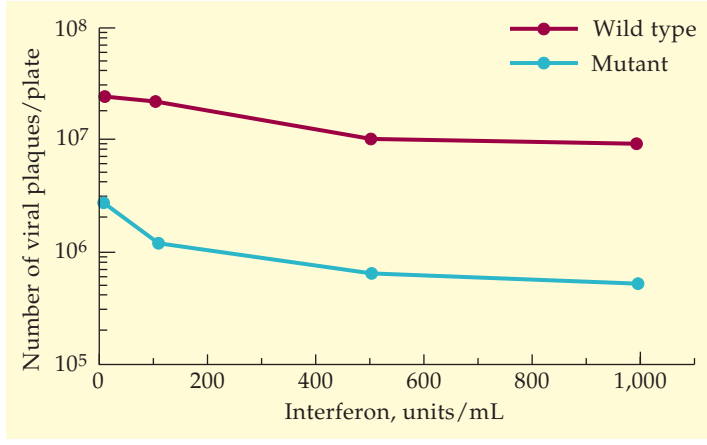


FIGURE 12.26 Sensitivity of wild-type ($K3L^+$) and mutant ($K3L^-$) vaccinia virus to interferon. Mouse L929 cells were pretreated with a mixture of mouse alpha and beta interferon for 24 hours before the virus was added. Plates without interferon had approximately 3×10^7 viral plaques per plate. Adapted from Paoletti and Tartaglia, U.S. patent 5,378,457, 1995.

Vaccines Directed against Bacteria

Since the discovery and subsequent widespread dissemination of antibiotics, only a modest amount of research has been directed toward the development of vaccines for bacterial diseases. However, there are good reasons for developing bacterial vaccines:

- Not all bacterial diseases are readily treated with antibiotics.
- The use of antibiotics over the last 40 years has resulted in the proliferation of bacterial strains that are resistant to several antibiotics.
- Reliable refrigeration facilities for the storage of antibiotics are not commonly available in many tropical countries.
- It is often difficult to ensure that individuals receiving antibiotic therapy undergo the full course of treatment.

Given the need to produce vaccines that will be effective against bacterial diseases, the question is, Which strategies are likely to be most effective? In instances where the disease-causing bacterium does not grow well in culture, the development of an attenuated strain is not feasible. For these bacteria, alternative approaches must be used. For example, *Rickettsia rickettsii*, a gram-negative obligately intracellular bacterium that causes Rocky Mountain spotted fever, does not grow in culture. In this case, a cloned 155-kDa protein that is a major surface antigen of *R. rickettsii* was used as a subunit vaccine and was found to protect immunized mice against infection by this disease-causing bacterium.

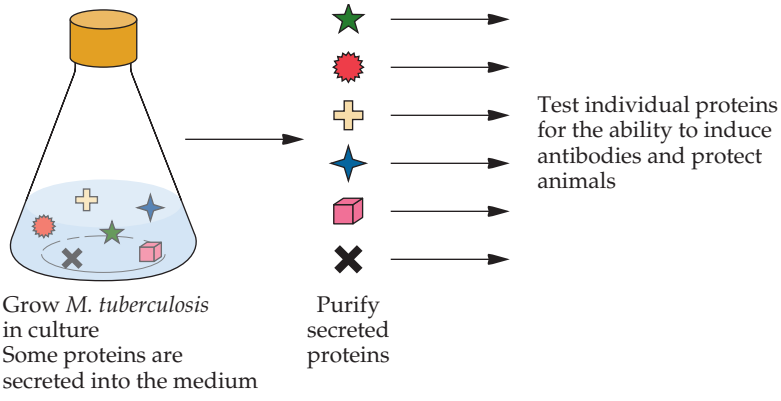
Tuberculosis. Tuberculosis, one of the most important infectious diseases worldwide, is caused by the bacterium *M. tuberculosis*. The bacterium can form lesions in any tissue or organ, which leads to cell death. The lungs are most commonly affected. Patients suffer fever and loss of body weight, and without treatment, tuberculosis is often fatal. It is estimated that approximately 2 billion people are currently infected with the organism and that approximately 2 million to 3 million deaths a year result from these infec-

tions. Over the past 50 years, antibiotics have been used to treat patients infected with *M. tuberculosis*. However, numerous multidrug-resistant strains of *M. tuberculosis* are now prevalent. In the United States, among HIV patients infected with an antibiotic-resistant strain of *M. tuberculosis*, there is a 50% mortality rate within 60 days. Consequently, a bacterial disease that was thought to be under control has become a serious public health problem in many parts of the world.

Currently, in some countries, bacillus Calmette-Guérin (BCG), an attenuated strain of *Mycobacterium bovis* that was developed between 1906 and 1919, is used as a vaccine against tuberculosis. However, the use of this vaccine has some drawbacks. First, live BCG cells can cause tuberculosis in immunocompromised individuals, such as AIDS patients. Second, individuals treated with BCG respond positively to a common tuberculosis diagnostic test, which makes it impossible to distinguish between individuals infected with *M. tuberculosis* and those inoculated with BCG cells. For these reasons, the BCG strain is not approved for use in a number of countries, including the United States.

In an attempt to determine whether a safer and more effective vaccine against tuberculosis might be developed, the extent of the immunoprotection elicited by purified *M. tuberculosis* extracellular proteins was examined. Following growth of the bacterium in liquid culture, 6 of the most abundant of the approximately 100 secreted proteins (Fig. 12.27) were purified. Each of these proteins was used separately and then in combination to immunize guinea pigs. The immunized animals were then challenged with an aerosol containing approximately 200 cells of live *M. tuberculosis*—a large dose for these animals. The animals were observed for 9 to 10 weeks before their lungs and spleens were examined for the presence of disease-causing organisms. In these experiments, some of the purified protein combinations provided a slightly lower level of protection against weight loss, death, and infection of lungs and spleen than did the live BCG vaccine. Prominent among the proteins that provided protection was the *M. tuberculosis* major secretory protein, a 30-kDa mycolytransferase also known as α-antigen, or antigen 85B. However, a DNA vaccine encoding

FIGURE 12.27 Schematic representation of the development of a multiprotein subunit vaccine for tuberculosis. The six most abundant secreted proteins from *M. tuberculosis* are purified from the growth medium and then tested for the ability to induce antibodies in guinea pigs. The immunized animals are subsequently challenged with *M. tuberculosis*.



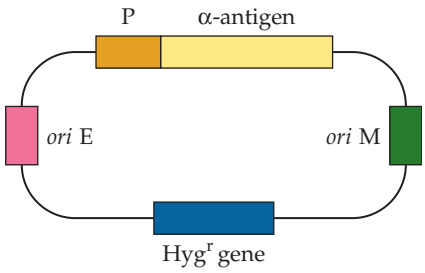


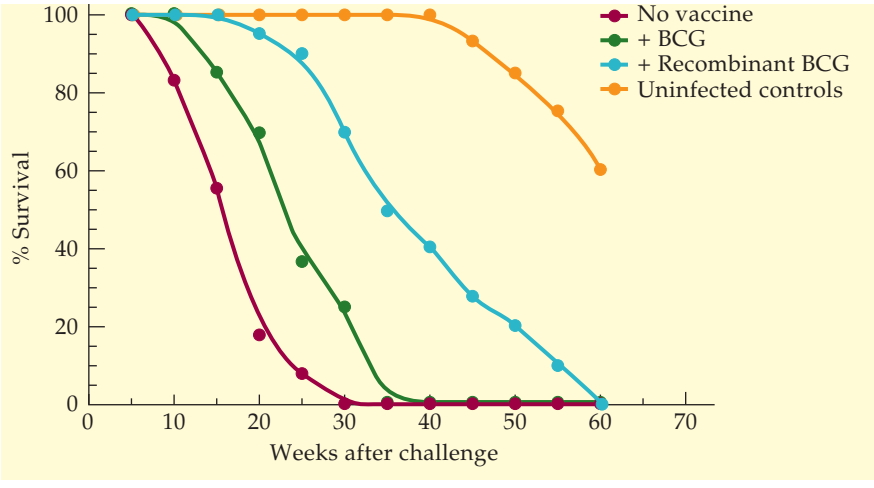
FIGURE 12.28 The plasmid construct used to transform BCG to make it a more effective vaccine. The plasmid is isolated from *E. coli* cells and then introduced into BCG by electroporation. *ori* E, *E. coli* origin of replication; *ori* M, *Mycobacterium* origin of replication; Hyg^r gene, hygromycin resistance gene (and its promoter); P and α-antigen, the promoter and the coding region of the 30-kDa secreted protein.

this protein was even less effective than the purified secreted protein. While this and possibly other *M. tuberculosis*-secreted proteins might eventually be part of a safe and efficacious vaccine for the prevention of tuberculosis in humans, it is necessary to develop a suitable delivery system for them. In theory, the optimal delivery system for an antigen that provides protection against tuberculosis should be (1) able to multiply in the mammalian host, (2) nonpathogenic, and (3) able to express and secrete the protective antigen. All of these requirements are satisfied by the available BCG strain. Therefore, an *E. coli*–mycobacterium shuttle vector that contained the gene for the 30-kDa protein (α-antigen) under the control of its own promoter was introduced into two different BCG strains (Fig. 12.28). Transformed cells produced 2.0- to 5.4-fold more 30-kDa protein than did nontransformed cells. In addition, despite the fact that the introduced genes were plasmid encoded and therefore potentially unstable, transformed cells continued to express a high level of 30-kDa protein after the vaccination of a test animal. In agreement with the hypothesis that the extracellular proteins of intracellular organisms are key immunoprotective molecules, guinea pigs immunized with transformed BCG strains had significantly fewer bacilli in their lungs and spleens. In addition, there were smaller and fewer lesions in their lungs, spleens, and livers, and the survival of the animals was significantly increased, compared with animals vaccinated with a nontransformed BCG strain (Fig. 12.29). This is the first report of a vaccine against tuberculosis that is more potent than the currently available commercial vaccine. This vaccine is currently in clinical trials; if it is successful, it could save tens of thousands of lives. Moreover, it is possible to prepare dried preparations of BCG in which individual bacteria form rod-like structures, 1 to 4 μm long and 0.2 to 0.4 μm in diameter, that may serve as the basis for a live bacterial vaccine that is delivered as an aerosol, thereby facilitating the inoculation of newborn infants.

Bacteria as Antigen Delivery Systems

Antigens that are located on the outer surface of a bacterial cell are more likely to be immunogenic than are those in the cytoplasm. Thus, one strategy

FIGURE 12.29 Survival of guinea pigs infected (challenged) with a pathogenic strain of *M. tuberculosis*. BCG is the traditional live bacterial vaccine strain.

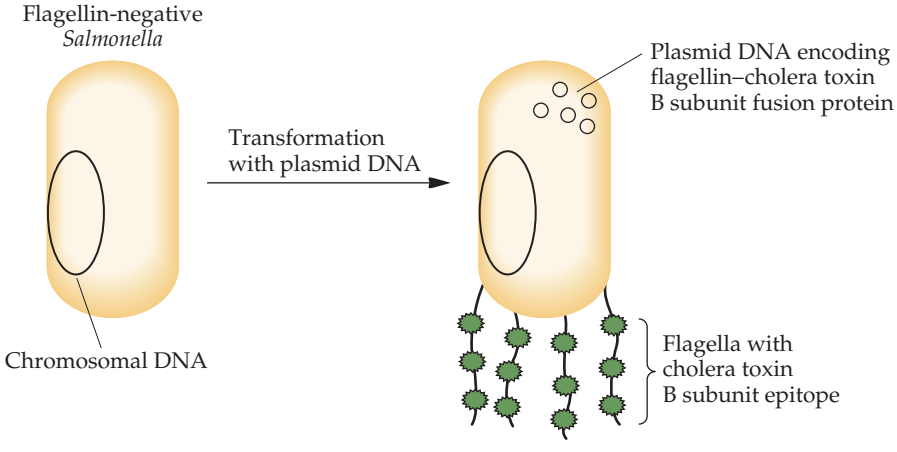


is to place a neutralizing antigen from a pathogenic bacterium on the surface of a live nonpathogenic bacterium. Flagella are made up of filaments of a single protein called flagellin. Under a microscope, they appear as conspicuous threadlike structures on the outer surfaces of some bacteria. If the flagella of a nonpathogenic organism could be made to carry a specific epitope from a pathogenic bacterium, protective immunogenicity might be easily achieved.

This strategy was used to engineer a cholera vaccine (Fig. 12.30). A synthetic oligonucleotide specifying an epitope of the cholera toxin B subunit was inserted into a portion of the *Salmonella* flagellin gene that varies considerably from one strain to another (hypervariable segment). The construct was then introduced into a flagellin-negative strain of *Salmonella*. The epitope, which consisted of amino acid residues 50 to 64 of the cholera toxin B subunit, was shown previously to elicit antibodies directed against intact cholera toxin. The chimeric flagellin functioned normally. Furthermore, the epitope was present at the flagellum surface. Immunization of mice by intraperitoneal injections of approximately 5×10^6 live or formalin-killed "flagellum-engineered" bacteria elicited high levels of antibodies directed against both the peptide, i.e., amino acids 50 to 64, and the intact cholera toxin molecule. Two or three different epitopes can be inserted into a single *Salmonella* flagellin gene, thereby creating a multivalent bacterial vaccine.

Attenuated *Salmonella* strains can be administered orally, which would enable them to deliver a range of bacterial, viral, and parasite antigens to the mucosal immune system. For this purpose, the choice of the promoter that drives the transcription of the foreign antigen is important. If too strong a promoter is used, the metabolic load might constrain bacterial proliferation. Moreover, unlike a closed system, such as a fermentation vessel, shifting the temperature or adding specific metabolites to induce foreign-gene expression is not possible when the bacterial vector is added to a host animal. On the other hand, promoters that respond to environmental signals may provide effective means of controlling the expression of the foreign antigen gene. For example, the *E. coli nirB* promoter, which is

FIGURE 12.30 Using *Salmonella* as an antigen delivery system and a flagellin–antigen fusion protein for presenting the antigen to the host immune system. A flagellin-negative strain of *Salmonella* was transformed with a plasmid containing a synthetic oligonucleotide specifying an epitope of the cholera toxin B subunit inserted into a hypervariable region of a *Salmonella* flagellin gene.



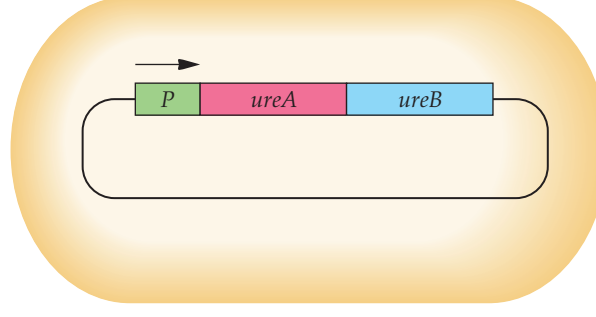
regulated by both nitrite and the oxygen tension of the environment, becomes active under anaerobic conditions. In one series of experiments, the *nirB* promoter was used to direct the expression of the nontoxic immunogenic fragment C of *C. tetani* toxin (tetanus toxin) in an attenuated strain of *Salmonella*. It is estimated that more than 1 million deaths per year in the developing world are the result of *C. tetani* infections. When the engineered *Salmonella* strain was grown aerobically in culture, tetanus toxin fragment C was not synthesized; however, following oral administration of the bacterium to test mice, fragment C was produced, and the animals generated antibodies against the peptide. Thus, the engineered *Salmonella* strain has potential as a live oral tetanus vaccine.

The spiral-shaped, gastrointestinal, and microaerophilic gram-negative bacterium *Helicobacter pylori* is widely distributed among human populations. It is believed to be the causative agent for a number of gastrointestinal diseases, including chronic gastritis, peptic ulcers, gastric lymphoma, and gastric cancer. Among infected individuals, which includes more than half of the world's population, about 10% are at risk of developing peptic ulcers. In recent years, the medical treatment for peptic ulcers has changed from antacids to antibiotics and proton pump inhibitors. The antibiotics eradicate the *H. pylori* infection, while the proton-pump inhibitors block the enzyme hydrogen–potassium ATPase, preventing the production of acid from the parietal cells at the gastric mucosa, which facilitates the healing of the mucosa.

Unfortunately, *H. pylori* is resistant to a number of commonly used antibiotics, including metronidazole, amoxicillin, erythromycin, and clarithromycin. Treatment of *H. pylori* requires multidrug regimens because the organism resides in a layer of mucus that acts as a barrier to antibiotic penetration. In addition, the necessary course of antibiotic treatment is too expensive for populations of less developed countries.

Colonization of the gastrointestinal tract by *H. pylori* is facilitated by the action of an *H. pylori*-encoded urease. This enzyme hydrolyzes urea to carbon dioxide and ammonia, thereby neutralizing stomach acid, making it possible for the bacterium to survive, bind, and function in the host. Urease is a cytosolic and surface-exposed nickel metalloenzyme and is one of the most abundantly expressed proteins in *H. pylori*. The enzyme comprises two subunits, A and B, that assemble into a complex [(αβ)3]4 supramolecular structure. Subunit B is more antigenic, making it a possible

FIGURE 12.31 Schematic representation of an attenuated strain of *S. enterica* serovar Typhi transformed with a plasmid encoding *H. pylori* urease subunits A and B under the transcriptional control of a *Salmonella* promoter *(P)*. The arrow indicates the direction of transcription.



vaccine candidate. To develop a vaccine that protects individuals against *H. pylori* infections, the genes encoding *H. pylori* urease subunits A and B were constitutively expressed under the control of a *Salmonella* promoter in a genetically deleted (attenuated) strain of *S. enterica* serovar Typhi (Fig. 12.31). Neither immunization with urease-expressing *S. enterica* serovar Typhi alone nor immunization with the purified urease enzyme plus an adjuvant conferred protection against challenge with a mouse-adapted strain of *H. pylori*. On the other hand, a vaccination protocol that combined both urease-expressing *S. enterica* serovar Typhi and urease plus an adjuvant was protective. While the success of this approach remains to be established in humans, these initial results are nevertheless encouraging and give hope that a human vaccine against *H. pylori* will be developed in the near future.

SUMMARY

Traditionally, vaccines have been either inactivated or attenuated infectious agents (bacteria or viruses) that are injected into an antibody-producing organism to produce immunity. There are a number of drawbacks to these vaccines. For example, not all pathogenic organisms can be grown to the large volumes needed to make a vaccine, there are safety concerns when large volumes of pathogenic organisms are being handled, attenuated strains may revert to the infectious state, inactivation may be incomplete, and shelf life is often dependent on refrigeration.

Recombinant DNA technology has been used in various ways to create reliable vaccines. Immunologically active, noninfectious agents are produced by deleting the genes that cause virulence; with this deletion, a live vaccine would never be able to revert to the infectious form. A gene(s) that encodes the major antigenic determinant(s) from a pathogenic organism can be cloned into the genome of a benign carrier organism (usually a virus or bacterium), which can be used as a vaccine without concern that any pathogenic organisms are present. The genes or segments of genes that encode the major antigenic determinants of pathogenic organisms can be cloned into expression vectors, and large amounts of the product can be harvested, purified, and used as a vaccine. With the last strategy, complete genes produce subunit vaccines, and cloned domains of the major antigenic determinants produce peptide vaccines. Peptide vaccines may also be produced by chemical peptide synthesis. As an alternative to using cloned antigenic proteins or peptides for inoculation, DNA constructs encoding the antigenic protein or peptide may be utilized. These DNA constructs may be delivered directly to animals or humans.

REFERENCES

Agarwal, K., and S. Agarwal. 2008. *Helicobacter pylori* vaccine: from past to future. *Mayo Clin. Proc.* **83:**169–175.

Amara, R. R., F. Villinger, J. D. Altman, S. L. Lydy, S. P. O'Neil, S. I. Staprans, D. C. Montefiore, Y. Xu, J. G. Herndon, L. S. Wyatt, M. A. Candido, N. L. Kozyr, P. L. Earl, J. M. Smith, H.-L. Ma, B. D. Grimm, M. L. Hulsey, J. Miller, H. M. McClure, J. M. McNicholl, B. Moss, and H. L. Robinson. 2001. Control of a mucosal challenge and prevention of AIDS by a multiprotein DNA/MVA vaccine. *Science* **292:**69–74.

Audran, R., M. Cachat, F. Lurati, S. Soe, O. Leray, G. Corradin, P. Druilhe, and F. Spertini. 2005. Phase I malaria vaccine trial with a long synthetic peptide derived from the merozoite surface protein 3 antigen. *Infect. Immun.* **73:**8017–8026.

Bittle, J. L., R. A. Houghten, H. Alexander, T. M. Shinnick, J. G. Sutcliffe, R. A. Lerner, D. J. Rowlands, and F. Brown. 1982. Protection against foot-and-mouth disease by immunization with a chemically synthesized peptide predicted from the viral nucleotide sequence. *Nature* **298:**30–33.

Blancou, J., M. P. Kieny, R. Lathe, J. P. Lecoq, P. P. Pastoret, J. P. Soulebot, and P. Desmettre. 1986. Oral vaccination of the fox against rabies using a live recombinant vaccinia vaccine. *Nature* **322:**373–375.

Blasco, R., and B. Moss. 1995. Selection of recombinant vaccinia viruses on the basis of plaque formation. *Gene* **149:**157–162.

Boothroyd, J. C., P. E. Highfield, G. A. M. Cross, D. J. Rowlands, P. A. Lowe, F. Brown, and T. J. R. Harris. 1981. Molecular cloning of foot and mouth disease virus genome and nucleotide sequences in the structural protein genes. *Nature* **290:**800–802.

Charles, I., and G. Dougan. 1990. Gene expression and the development of live enteric vaccines. *Trends Biotechnol.* **8:**117–121.

Chatfield, S. N., I. G. Charles, A. J. Makoff, M. D. Oxer, G. Dougan, D. Pickard, D. Slater, and N. F.

Fairweather. 1992. Use of the *nirB* promoter to direct the stable expression of heterologous antigens in *Salmonella* oral vaccine strains: development of a single-dose oral tetanus vaccine. *Bio/Technology* **10:**888–892.

Chow, M., R. Yabrov, J. Bittle, J. Hogle, and D. Baltimore. 1985. Synthetic peptides from four separate regions of the poliovirus type 1 capsid protein VP1 induce neutralizing antibodies. *Proc. Natl. Acad. Sci. USA* **82:**910–914.

Cichutek, K. 2000. DNA vaccines: development, standardization and regulation. *Intervirology* **43:**331–338.

Clarke, B. E., S. E. Newton, A. R. Carroll, M. J. Francis, G. Appleyard, A. D. Syred, P. E. Highfield, D. J. Rowlands, and F. Brown. 1987. Improved immunogenicity of a peptide epitope after fusion to hepatitis B core protein. *Nature* **330:**381–384.

Corbett, M., W. M. Bogers, J. L. Heeney, S. Gerber, C. Genin, A. Didierlaurent, H. Oostermeijer, R. Dubbes, G. Braskamp, S. Lerondel, C. E. Gomez, M. Esteban, R. Wagner, I. Kondova, P. Mooij, S. Ball-Jhagjhoorsingh, N. Beenhakker, G. Koopman, S. van der Burg, J.-P. Kraehenbuhl, and A. Le Pape. 2008. Aerosol immunization with NYVAC and MVA vectored vaccines is safe, simple, and immunogenic. *Proc. Natl. Acad. Sci. USA* **105:**2046-2051.

Cremer, K. J., M. Mackett, C. Wohlenberg, A. L. Notkins, and B. Moss. 1985. Vaccinia virus recombinant expressing herpes simplex virus type 1 glycoprotein D prevents latent herpes in mice. *Science* **228:**737–740.

Da Costa, X. J., C. A. Jones, and D. M. Knipe. 1999. Immunization against genital herpes with a vaccine virus that has defects in productive and latent infection. *Proc. Natl. Acad. Sci. USA* **96:**6994–6998.

DiMarchi, R., G. Brooke, C. Gale, V. Cracknell, T. Doel, and N. Mowat. 1986. Protection of cattle against foot-and-mouth disease by a synthetic peptide. *Science* **232:**639–641.

Ferguson, M. 1991. Progress towards rabies control. *Trends Biotechnol.* **9:**7–11.

Finkelstein, A., and R. F. Silva. 1989. Live recombinant vaccines for poultry. *Trends Biotechnol.* **7:**273–277.

Flexner, C., A. Hugin, and B. Moss. 1987. Prevention of vaccinia virus infection in immunodeficient mice by vector-directed IL-2 expression. *Nature* **330:**259–262.

Gómez-Duarte, O. G., D. Bumann, and T. F. Meyer. 1999. The attenuated *Salmonella* vaccine approach for the control of *Helicobacter pylori*-related diseases. *Vaccine* **17:**1667–1673.

Heithoff, D. M., R. L. Sinsheimer, D. A. Low, and M. J. Mahan. 1999. An essential role for DNA adenine methylation in bacterial virulence. *Science* **284:**967–970.

Horwitz, M. A., B.-W. E. Lee, B. J. Dillon, and G. Harth. 1995. Protective immunity against tuberculosis induced by vaccination with major extracellular proteins of *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. USA* **92:**1530–1534.

Horwitz, M. A., G. Harth, B. J. Dillon, and S. Maslesa-Galic. 2000. Recombinant bacillus Calmette-Guérin (BCG) vaccines expressing the *Mycobacterium tuberculosis* 30-kDa major secretory protein induce greater protective immunity against tuberculosis than conventional BCG vaccines in a highly susceptible animal model. *Proc. Natl. Acad. Sci. USA* **97:**13853–13858.

Horwitz, M. A., and G. Harth. 2003. A new vaccine against tuberculosis affords greater survival after challenge than current vaccine in the guinea pig model of pulmonary tuberculosis. *Infect. Immun.* **71:**1672-1679.

Jia, R., J. H. Guo, M. W. Fan, Z. Bian, Z. Chen, B. Fan, F. Yu, and Q. A. Xu. 2006. Immunogenicity of CTLA4 fusion anti-caries DNA vaccine in rabbits and monkeys. *Vaccine* **24:**5192-5200.

Kaper, J. B., H. Lockman, M. M. Baldini, and M. M. Levine. 1984. A recombinant live oral cholera vaccine. *Bio/Technology* **2:**345–349.

Kaper, J. B., H. Lockman, M. M. Baldini, and M. M. Levine. 1984. Recombinant nontoxinogenic *Vibrio cholerae* strains as attenuated cholera vaccine candidates. *Nature* **308:**655–658.

Kaper, J. B., J. G. Morris, Jr., and M. M. Levine. 1995. Cholera. *Clin. Microbiol. Rev.* **8:**48–86.

Kaslow, D. C., S. N. Isaacs, I. A. Quakyi, R. W. Gwadz, B. Moss, and D. B. Keister. 1991. Induction of *Plasmodium falciparum* transmission-blocking antibodies by recombinant vaccinia virus. *Science* **252:**1310–1313.

Kirnbauer, R., F. Booy, N. Cheng, D. R. Lowy, and J. T. Schiller. 1992. Papillomavirus L1 major capsid protein self-assembles into virus-like particles that are highly immunogenic. *Proc. Natl. Acad. Sci. USA* **89:**12180-12184.

Koutsky, L. A., K. A. Ault, C. M. Wheeler, D. R. Brown, E. Barr, F. B. Alvarez, L. M. Chiacchierini, and K. U. Jansen. 2002. A controlled trial of a human papillomavirus type 16 vaccine. *N. Engl. J. Med.* **347:**1645–1650.

Kupper, H., W. Keller, C. Kurz, S. Forss, H. Schaller, R. Franze, K. Strohmaier, O. Marquardt, V. G. Zaslavsky, and P. H. Hofschneider. 1981. Cloning of cDNA of major antigen of foot and mouth disease virus and expression in *E. coli*. *Nature* **289:**555–559.

Lasky, L. A., D. Dowbenko, C. C. Simonsen, and P. W. Berman. 1984. Protection of mice from lethal herpes simplex virus infection by vaccination with a secreted form of cloned glycoprotein D. *Bio/Technology* **2:**527–532.

Leitner, W. W., H. Ying, and N. P. Restifo. 2000. DNA and RNA-based vaccines: principles, progress and prospects. *Vaccine* **18:**765–777.

Lewis, P. J., and L. A. Babiuk. 1999. DNA vaccines: a review. *Adv. Virus Res.* **54:**129–188.

Londoño-Arcila, P., D. Freeman, H. Kleanthous, A. M. O'Dowd, S. Lewis, A. K. Turner, E. L. Rees, T. J. Tibbitts, J. Greenwood, T. P. Monath, and M. J. Darsley. 2002. Attenuated *Salmonella enterica* serovar Typhi expressing urease effectively immunizes mice against *Helicobacter pylori* challenge as part of a heterologousmucosal priming-parenteral boosting vaccination regimen. *Infect. Immun.* **70:**5096–5106.

Lowe, R. S., P. M. Keller, B. J. Keech, A. J. Davison, Y. Whang, A. J. Morgan, E. Kieff, and R. W. Ellis. 1987. Varicella-zoster virus as a live vector for the expression of foreign genes. *Proc. Natl. Acad. Sci. USA* **84:**3896–3900.

Manoj, S., L. A. Babiuk, and S. van Drunen Littel-van den Hurk. 2004. Approaches to enhance the efficacy of DNA vaccines. *Crit. Rev. Clin. Lab. Sci.* **41:**1–39.

McCluskie, M. J., and H. L. Davis. 1999. Mucosal immunization with DNA vaccines. *Microbes Infect*. **1:**685–698.

Mekalanos, J. J., and J. C. Sadoff. 1994. Cholera vaccines: fighting an ancient scourge. *Science* **265:**1387–1389.

Michel, M.-L., H. L. Davis, M. Schleef, M. Mancini, P. Tiollais, and R. G. Whalen. 1995. DNA-mediated immunization to the hepatitis B surface antigen in mice: aspects of the humoral response mimic hepatitis B viral infection in humans. *Proc. Natl. Acad. Sci. USA* **92:**5307–5311.

Miner, J. N., and D. E. Hruby. 1990. Vaccinia virus: a versatile tool for molecular biologists. *Trends Biotechnol.* **8:**20–25.

Moss, B. 1991. Vaccinia virus: a tool for research and vaccine development. *Science* **252:**1662–1667.

Nabel, G. J., and P. L. Felgner. 1993. Direct gene transfer for immunotherapy and immunization. *Trends Biotechnol.* **11:**211–215.

Newton, S. M. C., C. O. Jacob, and B. A. D. Stocker. 1989. Immunoresponse to cholera toxin epitope inserted in *Salmonella* flagellin. *Science* **244:**70–72.

Nussenzweig, R. S., and C. A. Long. 1994. Malaria vaccines: multiple targets. *Science* **265:**1381–1383.

Oehen, S., H. Hengartner, and R. M. Zinkernagel. 1991. Vaccination for disease. *Science* **251:**195–197.

Paoletti, E. 1996. Applications of pox virus vectors to vaccination: an update. *Proc. Natl. Acad. Sci. USA* **93:**11349–11353.

Paoletti, E., and J. Tartaglia. January 1995. Interferon sensitive recombinant poxvirus vaccine. U.S. patent 5,378,457.

Rice, J., C. H. Ottensmeier, and F. K. Stevenson. 2008. DNA vaccines: precision tools for activating effective immunity against cancer. *Nat. Rev. Cancer* **8:**108–120.

Sheridan, C. 2005. The business of making vaccines. *Nat. Biotechnol.* **23:**1359–1366.

Singh, M., M. Briones, G. Ott, and D. O'Hagan. 2000. Cationic microparticles: a potent delivery system for DNA vaccines. *Proc. Natl. Acad. Sci. USA* **97:**811–816.

Singh, S., S. Soe, J.-P. Mejia, C. Roussilhon, M. Theisen, G. Corradin, and P. Druilhe. 2004. Identification of a conserved region of *Plasmodium falciparum* MSP3 targeted by biologically active antibodies to improve vaccine design. *J. Infect. Dis.* **190:**1010–1018.

Sizemore, D. R., A. A. Branstrom, and J. C. Sadoff. 1995. Attenuated *Shigella* as a DNA delivery vehicle for DNA-mediated immunization. *Science* **270:**299–302.

Stover, C. K., V. F. de la Cruz, T. R. Fuerst, J. E. Burlein, L. A. Benson, L. T. Bennett, G. P. Bansal, J. F. Young, M. H. Lee, G. F. Hatfull, S. B. Snapper, R. G. Barletta, W. R. Jacobs, Jr., and B. R. Bloom. 1991. New use of BCG for recombinant vaccines. *Nature* **351:**456–460.

Stranger-Jones, Y. K., T. Bae, and O. Schneewind. 2006. Vaccine assembly from surface proteins of *Staphylococcus aureus*. *Proc. Natl. Acad. Sci. USA* **103:**16942–16947.

Tang, D.-C., M. DeVit, and S. A. Johnston. 1992. Genetic immunization is a simple method for eliciting an immune response. *Nature* **356:**152–154.

Tartaglia, J., and E. Paoletti. 1988. Recombinant vaccinia virus vaccines. *Trends Biotechnol.* **6:**43–46.

Titus, R. G., J. G. Gueiros-Filho, L. A. R. De Freitas, and S. M. Beverley. 1995. Development of a safe live *Leishmania* vaccine line by gene replacement. *Proc. Natl. Acad. Sci. USA* **92:**10267–10271.

Ulmer, J. B., J. J. Donnelly, S. E., Parker, G. H. Rhodes, P. L. Felgner, V. J. Dwarki, S. H. Gromkowski, R. R. Deck, C. M. DeWitt, A. Friedman, L. A. Hawe, K. R. Leander, D. Martinez, H. C. Perry, J. W. Shiver, D. L. Montgomery, and M. A. Liu. 1993. Heterologous protection against influenza by injection of DNA encoding a viral protein. *Science* **259:**1745–1749.

Weeratna, R. D., M. J. McCluskie, L. Comanita, T. Wu, and H. L. Davis. 2001. Optimization strategies for DNA vaccines. *Intervirology* **43:**218–226.

Xu, Q. A., F. Yu, M. W. Fan, Z. Bian, Z. Chen, B. Peng, R. Jia, and J. H. Guo. 2007. Protective efficacy of a targeted anti-caries DNA plasmid against cariogenic bacteria infections. *Vaccine* **25:**1191-1195.

Zakhartchouk, A. N., C. Sharon, M. Satkunarajah, T. Auperin, S. Viswanathan, G. Mutwiri, M. Petric, R. H.See, R. C. Brunham, B. B. Finlay, C. Cameron, D. J. Kelvin, A. Cochrane, J. M. Rini, and L. A. Babiuk. 2007. Immunogenicity of a receptor-binding domain of SARS coronavirus spike protein in mice: implications for a subunit vaccine. *Vaccine* **25:**136–143.

Zimet, G. D., M. L. Shew, and J. A. Kahn. 2008. Appropriate uses of cervical cancer vaccine. *Annu. Rev. Med.* **59:**223–236.

REVIEW QUESTIONS

1. Briefly describe a protocol for developing a vaccine against a toxin-producing bacterium.

2. What factors limit the usefulness of conventional vaccines?

3. As part of your work with an international animal health organization, you are given the task of developing a vaccine against a bovine virus that is the cause of tens of thousands of cattle deaths around the world annually. The viral genome consists of a 10-kb linear piece of single-stranded RNA with a poly(A) tail that encodes eight different proteins. The virus does not have a viral envelope, and the major antigenic determinant is the capsid protein viral protein 2. Outline an experimental strategy to develop a vaccine against this virus.

4. Discuss the development of peptide vaccines that are directed against viruses.

5. What is vaccinia virus, and how can it be used to produce unique live recombinant vaccines?

6. As an employee of the World Health Organization, you have to decide on the best strategy for eradicating rabies in wild animal populations. Assuming that you must choose between a peptide and a vaccinia virus-based vaccine, select one type of vaccine and justify your choice.

7. Discuss the advantages of a live recombinant viral vaccine over killed and subunit vaccines.

8. Discuss some of the different strategies that have been used to produce vaccines against cholera.

9. How would you develop a subunit vaccine against HSV?

10. How can bacteria be used as part of a DNA vaccine delivery system?

11. How can vaccinia virus be made more sensitive to interferon? Explain.

12. How would you develop a vaccine against *S. aureus?*

13. Suggest several methods that you could use to deliver DNA for genetic immunization to animal cells.

14. What are MIDGE vectors, and how can they be used to facilitate genetic immunization?

15. How would you develop a vaccine against human papillomavirus?

16. How would you develop an effective DNA vaccine against dental caries?

17. How would you improve the traditional vaccine against tuberculosis?

13

Synthesis of Commercial Products by Recombinant Microorganisms

To date, molecular biotechnology research has focused largely on the production of a range of different proteins, including enzymes that are used commercially. However, recombinant DNA techniques can also be used to enhance the production of low-molecular-weight compounds, such as vitamins, amino acids, dyes, precursors of biopolymers, and antibiotics. In these cases, the host microorganism is engineered to become a factory for the production of useful metabolites.

Restriction Endonucleases

Recombinant DNA technology would not be possible without a ready supply of different restriction endonucleases. Currently, more than 300 different restriction endonucleases are commercially available, with worldwide sales in 2007 in the range of $350 million. These enzymes occur naturally in many different microorganisms, including species that are aerobic, anaerobic, photosynthetic, diazotrophic, mesophilic, thermophilic, psychrophilic, and either slow or fast growing. For each of these organisms, a detailed fermentation protocol—specifying the temperature, pH, medium composition, and oxygen tension—has to be developed and optimized to achieve the maximum yield of the target restriction enzyme. To avoid having to maintain a large number of different microorganisms, stock a very wide range of microbial growth medium components, design several different types of fermenters, and spend an inordinate amount of time developing optimal growth conditions for a large number of different organisms (one major supplier of restriction enzymes lists 265 different enzymes in its catalog), investigators often clone restriction endonuclease genes into *Escherichia coli*. Exclusive use of *E. coli* allows bioengineers to standardize the production conditions for all restriction endonucleases. In addition, *E. coli* cells grow rapidly to high cell densities and can be engineered to significantly overexpress each target restriction enzyme.

Although the technology for isolating and expressing foreign genes in *E. coli* and some other host organisms is well established, it should be remembered that the host organism is a living entity that can be dramatically

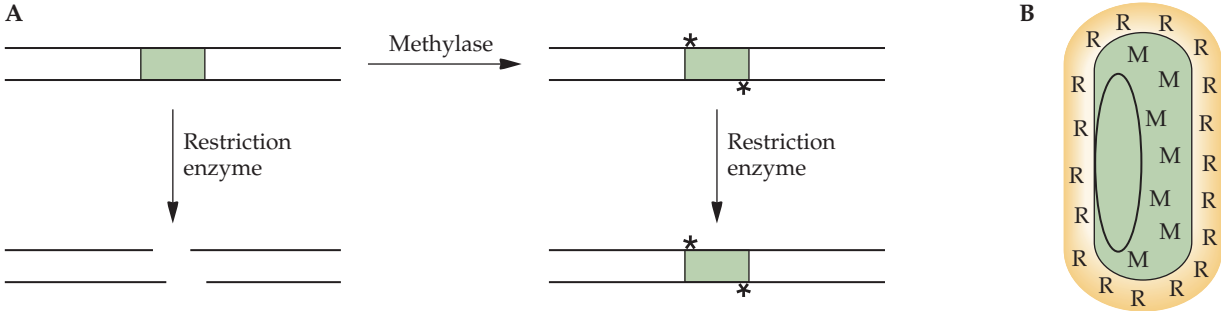
affected by the production or presence of a heterologous protein. For example, overexpression of a heterologous protein may drain the host organism of important metabolic resources and, as a result, adversely affect its growth. In addition, the presence of a heterologous protein may be lethal to the host. For example, restriction endonucleases digest DNA at sites that are present on all DNA molecules. As a result, an organism that expresses a cloned restriction endonuclease gene is likely to have its own DNA degraded unless a protection mechanism is present.

Microorganisms that make restriction endonucleases have evolved a self-protection system. Methylation of one or more of the bases of the DNA within the recognition sequence prevents the homologous restriction endonuclease from cutting the DNA at this site (Fig. 13.1A). Gram-negative microorganisms have an added mode of protection that entails localizing the restriction endonuclease within the periplasmic space and the methylation (modification) enzyme in the cell cytoplasm (Fig. 13.1B). This compartmentalization physically separates the restriction endonuclease from the DNA while ensuring that the modification enzyme has ready access to the chromosomal DNA. In addition to protecting the cell from the toxic effects of the restriction enzyme, this segregation provides a cellular defense against attack by any foreign DNA, such as DNA from a bacterial virus, that might enter the periplasm.

One way to circumvent the problem of host DNA degradation by heterologous restriction endonucleases is to clone and express the genes for both the restriction enzyme and its specific (cognate) modification enzyme in the host organism. Cloning both of these genes into the same organism is technically complex unless both the restriction endonuclease and methylation genes are close to each other on the chromosome. In addition, to prevent the digestion of the host DNA by the restriction endonuclease, it is imperative that, after transformation, the methylation activity be expressed prior to the production of the restriction endonuclease.

One of the first restriction enzyme genes to be cloned (into *E. coli*) encoded the enzyme PstI from *Providencia stuartii*, a gram-negative bacterium (Fig. 13.2). It is important to note that, for a particular genus and species, only some strains encode restriction enzymes. Therefore, in this strategy, in order to easily transform the host *E. coli* strain without degrading the input plasmid DNA, it was necessary to utilize an *E. coli* strain that was unable to synthesize the enzyme EcoRI.

FIGURE 13.1 **(A)** Protection of DNA from digestion by a restriction endonuclease by prior treatment with methylase, a methylating (modification) enzyme. The asterisks indicate the presence of a methylated base. **(B)** Cytoplasmic localization of modification enzyme (M) and periplasmic localization of restriction endonuclease (R) in gram-negative bacteria.



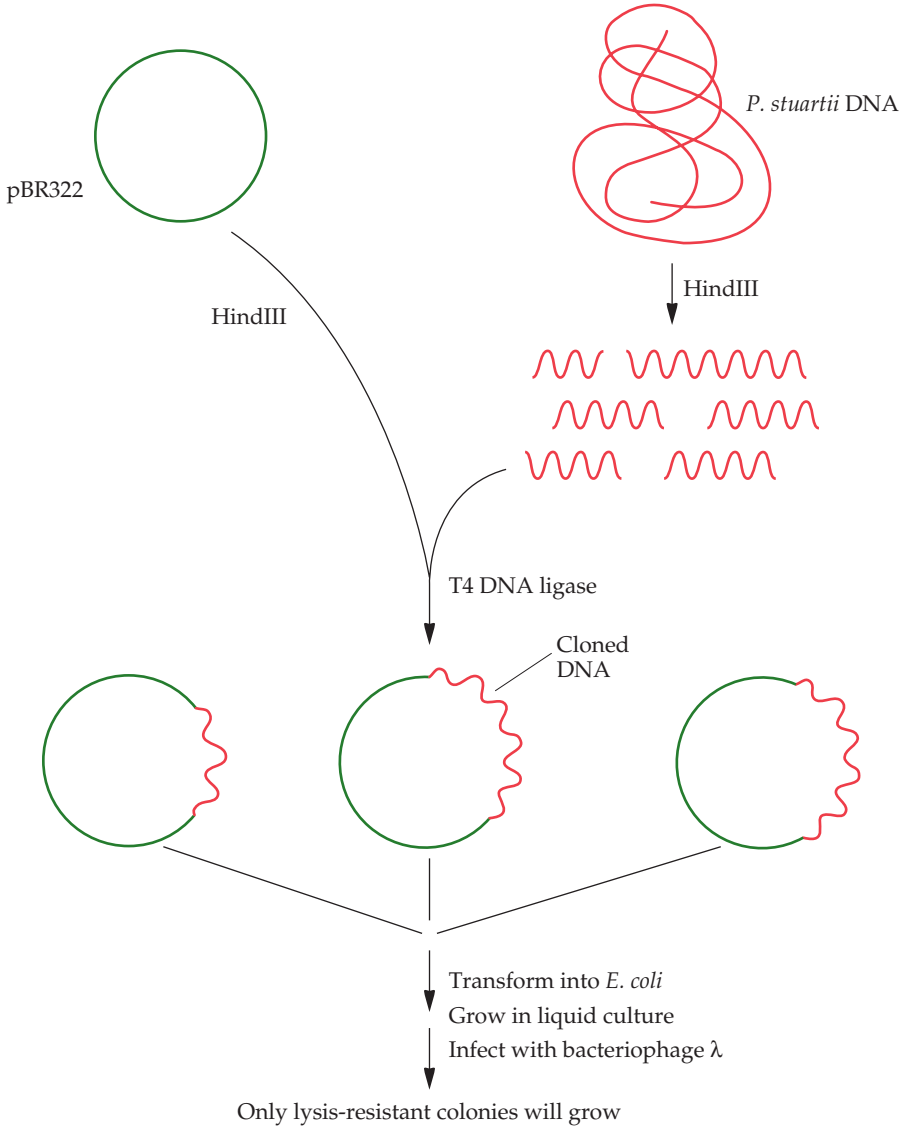


FIGURE 13.2 Method for cloning and selecting the gene for the restriction enzyme PstI. The *P. stuartii* chromosomal DNA is digested with HindIII and ligated into the HindIII site of plasmid pBR322. Transformants are grown in liquid medium before being infected with bacteriophage λ. The resistance of some transformants to lysis by λ is due to the presence and expression of a cloned PstI gene and its cognate methylation enzyme.

1. The chromosomal DNA from *P. stuartii* was digested with HindIII and ligated into the HindIII site on plasmid pBR322.
2. Following the introduction of the *P. stuartii* clone bank (library) into *E. coli* HB101 cells, transformants were grown in liquid medium before being infected with bacteriophage λ to test for the production of the restriction enzyme. When a restriction enzyme gene is expressed, the host cells become resistant to the lytic action of DNA bacteriophages such as λ because the restriction endonuclease extensively degrades the infecting bacteriophage DNA.

3. Transformants that were resistant to lysis by λ were grown, and samples were osmotically shocked to release the periplasmic proteins, which were assayed for PstI restriction enzyme activity.
4. Positive clones were assayed for PstI methylase activity.

One positive clone from this experiment contained, within a 4.0-kilobase-pair (kb) DNA fragment, an intact PstI restriction endonuclease and methylation operon, including the *P. stuartii* promoter. In this construct, the natural temporal order of synthesis—the methylation enzyme preceding the restriction endonuclease—was maintained. The level of the PstI restriction enzyme expressed in *E. coli* was approximately 10-fold higher than that in *P. stuartii*. As expected, PstI was localized in the periplasm, and the methylation enzyme was localized in the cytoplasm. Production of PstI using this *E. coli* clone is simpler and more efficient than production with *P. stuartii*.

Another strategy also has been used to isolate the genes for restriction and modification (methylation) enzyme systems. It was developed by a company that eventually became one of the world's leading suppliers of restriction enzymes and consists of the following steps.

1. A clone bank was made from the DNA of a donor organism that had a previously identified restriction endonuclease. The plasmid vector had at least one recognition site for the target restriction endonuclease.
2. The clone bank was introduced into *E. coli* by transformation. This step increased the amount of recombinant plasmid DNA and also allowed the expression of the modification enzyme.
3. Plasmid DNA was isolated from transformed cells that had been grown in liquid media under conditions that selected for the presence of the plasmid.
4. The plasmid DNA preparation was treated with the target restriction endonuclease.
5. *E. coli* cells were transformed with the restriction endonuclease-treated plasmid DNA preparation.

The rationale for this procedure is that the clones that carry and express the target modification enzyme will produce plasmid DNA that is resistant to digestion by the target restriction endonuclease because their DNA will be methylated at the recognition sites. For example, following transformation of *E. coli* by a pBR322–HindIII clone bank of *Desulfovibrio desulfuricans* DNA, plasmid DNA was isolated and digested with the restriction enzyme DdeI (Fig. 13.3). Plasmids that encode and express the DdeI modification enzyme are not digested by DdeI because the eight DdeI recognition sites of pBR322 are methylated. After the DdeI treatment, the remaining plasmid mixture is used to transform *E. coli*. Only intact circular plasmids yield transformants, and these carry the gene for a functional DdeI modification enzyme. All other plasmids are degraded by the restriction endonuclease. The resulting transformants must then be assayed for the DdeI restriction enzyme activity to determine which clones have the genes for both the modification enzyme and the restriction endonuclease. This strategy is effective for any restriction enzyme gene that is physically close to its modification enzyme gene—most restriction enzymes are encoded on the same operon as their cognate modification enzyme—and is cloned into a plasmid vector that has at least one recognition site for the target enzyme.

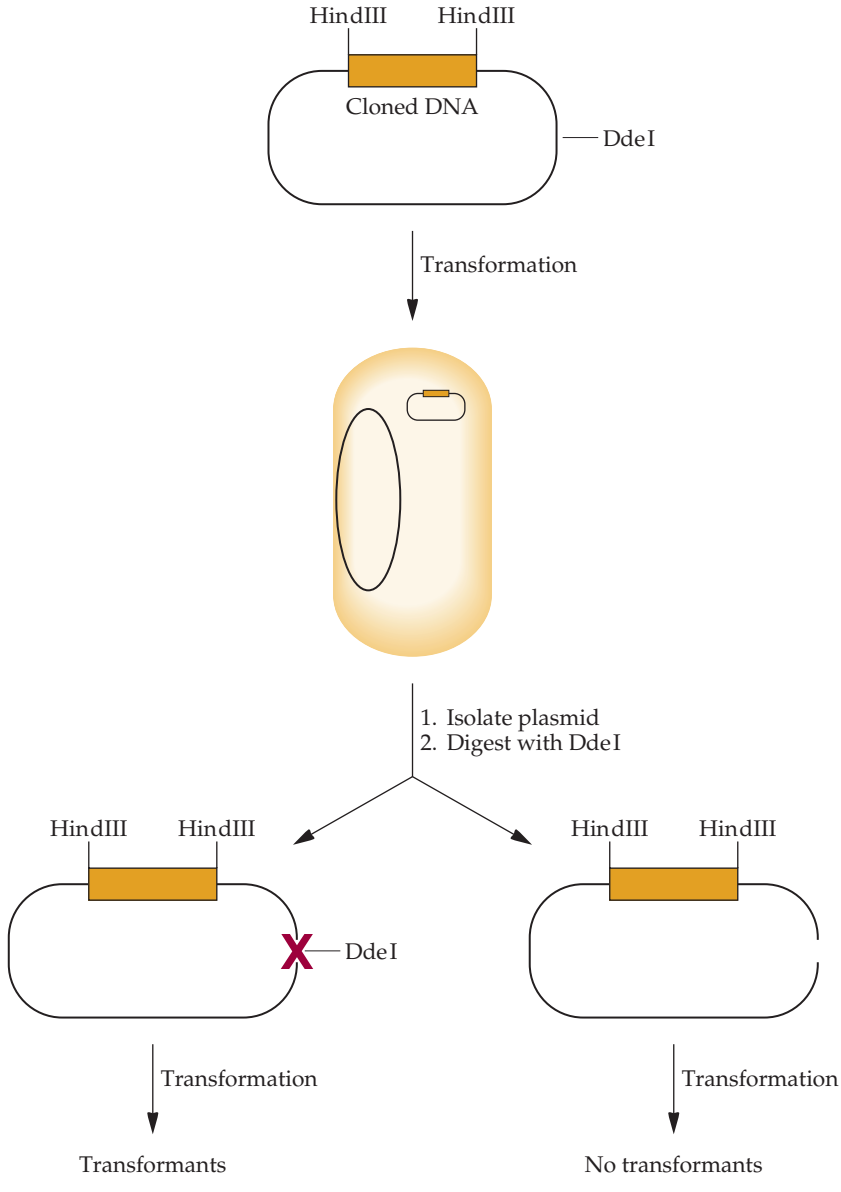


FIGURE 13.3 Scheme for cloning the DdeI modification enzyme. Plasmids with cloned *D. desulfuricans* DNA are used to transform *E. coli* cells. After the cells are grown in liquid culture, the plasmid DNA is isolated and digested with the restriction enzyme DdeI. Plasmids encoding the DdeI modification enzyme will be methylated at the DdeI restriction enzyme recognition site and will not be cut. The intact plasmid DNA is used to transform *E. coli* cells. The final transformants are assayed for the DdeI restriction enzyme and the corresponding methylase.

Lipase

Fatty stains are a persistent problem for the laundry industry. A combination of high temperature and high alkalinity can effectively emulsify and remove many fatty stains. However, these conditions often damage fabrics and also require large amounts of energy. On the other hand, the addition of lipases that are compatible with the wash conditions, such as the enzyme

produced by *Pseudomonas alcaligenes,* may provide an effective solution to this problem. Unfortunately, this enzyme is produced at such low levels that it is prohibitively expensive to use in the cleaning of laundry. Moreover, researchers have found it extremely difficult to overproduce the enzyme in a variety of heterologous hosts, including *Bacillus licheniformis, E. coli, Streptomyces lividans, Aspergillus niger,* and *Kluyveromyces lactis*. The difficulty in overexpressing the *P. alcaligenes* lipase may reflect the requirement for the simultaneous expression of another gene product that is involved in either the secretion or the stabilization of the bacterial lipase.

To isolate the lipase gene from *P. alcaligenes,* as well as any other gene whose expression was linked to lipase gene expression, the enzyme was first purified. The amino acid sequence of the N terminus was determined, and an oligonucleotide probe that corresponded to 11 amino acids from this region of the protein was synthesized. The oligonucleotide probe was used to screen a clone bank of *P. alcaligenes* DNA, with the result that a clone that contained all of the lipase gene and a portion of the additional gene was obtained. The DNA fragment encoding the lipase gene and a portion of the additional gene was used as a hybridization probe to screen another clone bank, with the result that the rest of the second gene was isolated. The two fragments were spliced together (Fig. 13.4), cloned into a broad-host-range expression vector, and used to transform *P. alcaligenes*. The lipase structural gene is called *lipA,* and the second (helper) gene is called *lipB*. When the vector was derived from a low-copy-number plasmid, the lipase activities of the transformants were four- to fivefold greater than that of the wild type, regardless of the presence or absence of the second gene *(lipB)*. However, with a high-copy-number plasmid, the lipase activities of the transformants were about 20-fold greater than that of the wild type in the absence of the *lipB* gene and approximately 35-fold greater than that of the wild type in the presence of the *lipB* gene. Since the lipase is secreted into the growth medium, very little purification should be necessary before using it in laundry detergent. Rather, the cells must be removed, and then the growth medium should be concentrated. However, when the production of the growth of this recombinant organism was scaled up from 10 liters to 10,000 liters, considerably less lipase activity was found than was expected. This was attributed to the production, and lack of removal, of large amounts of carbon dioxide in the growth medium. Modification of some of the operating parameters of this large fermenter resulted in a significant decrease in the level of dissolved carbon dioxide and decreased inhibition of lipase accumulation. This work went a long way toward providing lipase of sufficient quality and in sufficient quantity for use in laundry detergent.

Small Biological Molecules

With recombinant DNA technology, it is possible to modify metabolic pathways of organisms either by introducing new genes or by altering existing ones. The goal is to create an organism with a novel enzymatic activity that can convert an existing substrate into a commercial compound that with current technology can be produced only by a combination of chemical treatments and fermentation steps. Early metabolic-engineering experiments typically modified one or two genes in a biosynthetic pathway. However, with the knowledge of the complete DNA sequences of many bacterial genomes and the global information on the expression of bacterial

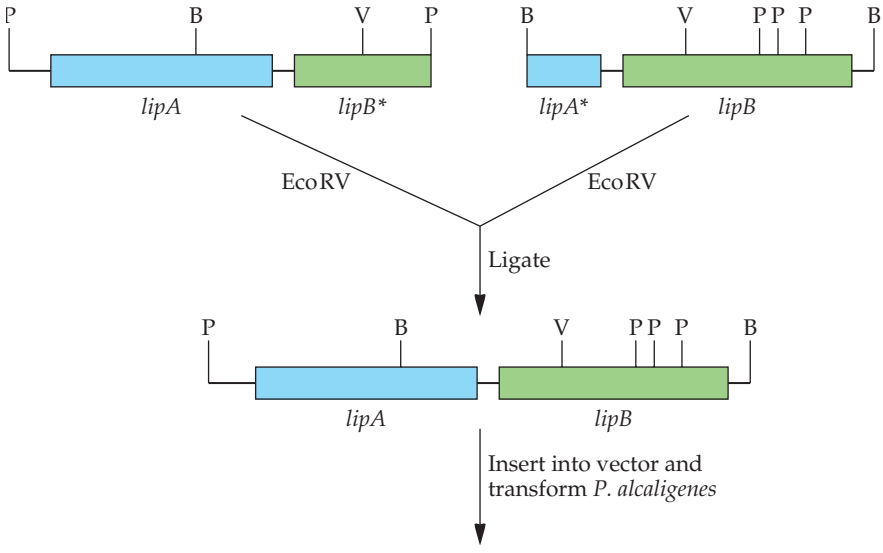


FIGURE 13.4 Splicing together portions of the lipase operon. P, PvuII; B, BclI; V, EcoRV. Partial genes are indicated by asterisks.

genes derived from studies that employ microarrays, proteomics, and metabolomics, it is possible to develop strategies to improve the yield of microbially produced molecules that involve the introduction or modification of an entire panel of related genes.

Synthesis of L-Ascorbic Acid

L-Ascorbic acid (vitamin C) is currently synthesized commercially by an expensive process starting with D-glucose that includes one microbial fermentation step and a number of chemical steps (Fig. 13.5). The last step in this process is the acid-catalyzed conversion of 2-keto-L-gulonic acid (2-KLG) to L-ascorbic acid. Biochemical studies of the metabolic pathways of a number of different microorganisms have shown that it may be possible to synthesize 2-KLG by a different pathway. For example, some bacteria *(Acetobacter, Gluconobacter,* and *Erwinia)* can convert glucose to 2,5-diketo-D-gluconic acid (2,5-DKG), and others *(Corynebacterium, Brevibacterium,* and *Arthrobacter)* have the enzyme 2,5-DKG reductase, which converts 2,5-DKG to 2-KLG.

The current procedure for synthesizing ascorbic acid could be improved by producing 2-KLG from glucose by cofermentation with suitable organisms. Unfortunately, cocultivation has problems of its own. For example, the two fermenting organisms might have different temperature and pH optima. The medium requirements and growth rates also might differ in such a way that the fermentation conditions are optimal for one organism and suboptimal for the other. This situation leads to the eventual "washout" (depletion or loss) of one of the organisms. Some of these incompatibilities may be overcome by utilizing a tandem fermentation process in which the two organisms are cultivated in succession (Fig. 13.6). Of course, this approach requires two separate fermentations rather than one, and if the organisms have different growth requirements, it is difficult to run the process on a continuous basis. Therefore, the best way to convert glucose into

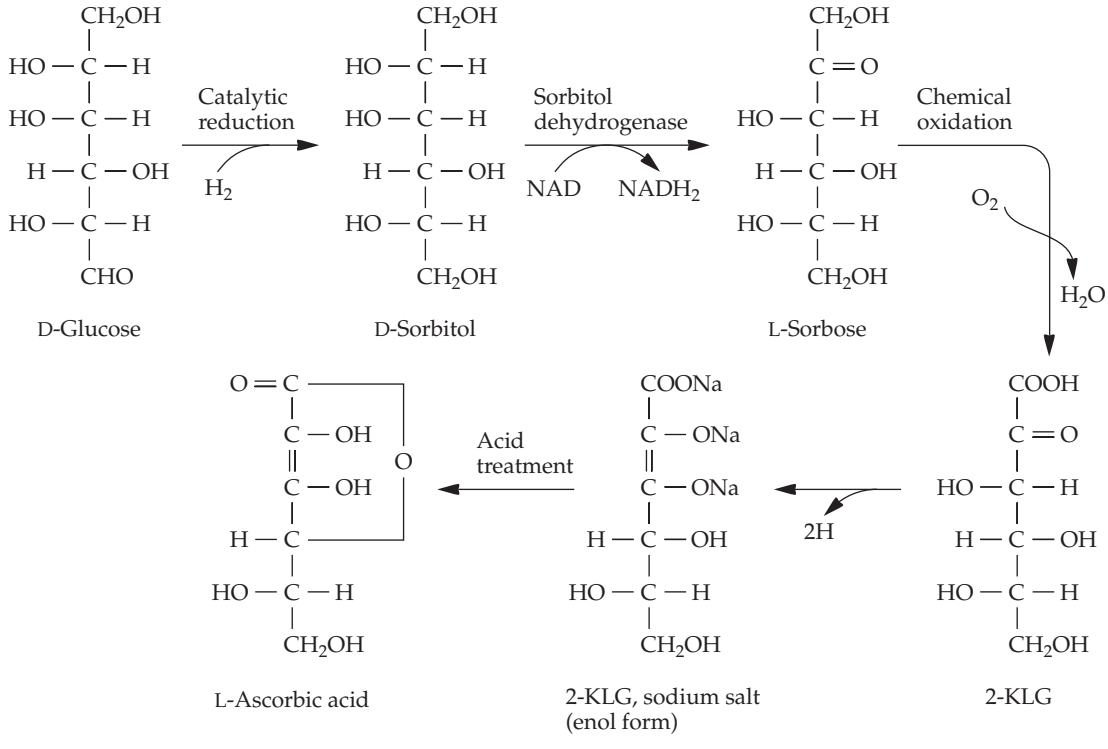


FIGURE 13.5 Commercial synthesis of L-ascorbic acid. Except for the microbial conversion of D-sorbitol to L-sorbose, all the steps are chemical reactions. The microbial conversion is carried out by *Acetobacter suboxydans,* which produces the enzyme sorbitol dehydrogenase.

2-KLG would be to engineer a single microorganism that carried all of the required enzymes. The conversion of D-glucose to 2,5-DKG by *Erwinia herbicola* includes several enzymatic steps, whereas the transformation of 2,5-DKG to 2-KLG by a *Corynebacterium* sp. requires only one. Consequently, the simplest strategy for constructing a single organism that is able to convert D-glucose to 2-KLG is to isolate the 2,5-DKG reductase gene from the *Corynebacterium* sp. and express it in *E. herbicola*.

The first step in cloning the 2,5-DKG reductase gene from the *Corynebacterium* sp. involved purifying the enzyme and determining the sequence of the first 40 amino acids from the N-terminal end of the molecule. On the basis of the known amino acid sequence, two 43-nucleotide DNA hybridization probes, each corresponding to a different portion of the protein molecule, were synthesized. Because 71% of the nucleotides in the *Corynebacterium* sp. are either G or C, the probes were designed to include, where possible, a G or C in the third position of all codons, thereby minimizing the extent of the mismatch between the probe and the target DNA. This approach was taken because at the time that this work was done, mixed probes were not readily available.

A *Corynebacterium* DNA clone bank was screened with these two probes. Any clones that hybridized with only one of the probes were discarded. It was assumed that any DNA that interacted with only one probe was probably not the target DNA. A clone that hybridized with both probes was isolated and then sequenced; it contained the 2,5-DKG reductase gene. The DNA sequences that were upstream of the ATG start signal were

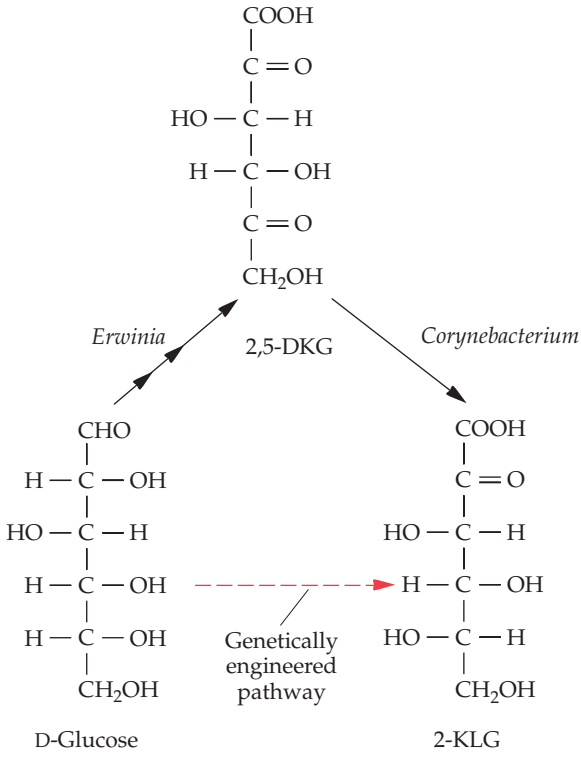


FIGURE 13.6 Biological production of 2-KLG. *Erwinia* has a set of three enzymes that synthesize 2,5-DKG from D-glucose. *Corynebacterium* has an enzyme that converts 2,5-DKG to 2-KLG. Thus, 2-KLG, which is the immediate precursor of L-ascorbic acid, can be produced from D-glucose either by cocultivating these two microorganisms or by genetically engineering *Erwinia* to express the enzyme from *Corynebacterium*, which converts 2,5-DKG to 2-KLG.

deleted and replaced with transcriptional and translational signals that function in *E. coli*, because the regulatory sequences from gram-positive microorganisms, such as *Corynebacterium* spp., are not efficiently utilized by *E. coli*. This construct expressed 2,5-DKG reductase activity in *E. coli* and subsequently was subcloned onto a broad-host-range vector, which was used to transform *E. herbicola*, which is able to use *E. coli* transcriptional and translational signals.

The transformed *Erwinia* cells were able to convert D-glucose directly to 2-KLG. The endogenous *Erwinia* enzymes, localized in the inner membrane of the bacterium, converted glucose to 2,5-DKG, and the cloned 2,5-DKG reductase, localized in the cytoplasm, catalyzed the conversion of 2,5-DKG to 2-KLG (Fig. 13.7). Thus, by genetic manipulation, the metabolic capabilities of two very dissimilar microorganisms were combined into one organism, which was able to produce the end product of the engineered metabolic pathway. This recombinant organism should be useful as a source of 2-KLG for the production of L-ascorbic acid, thereby replacing the first three steps of the currently used process (Fig. 13.5).

The commercial utility of the cloned 2,5-DKG reductase gene product might be improved by replacing certain amino acids of the enzyme to create mutants with increased catalytic activity and enhanced thermal stability. When the 2,5-DKG reductase gene was first isolated, the amino acid residues that contributed to the active site of this enzyme were not known.

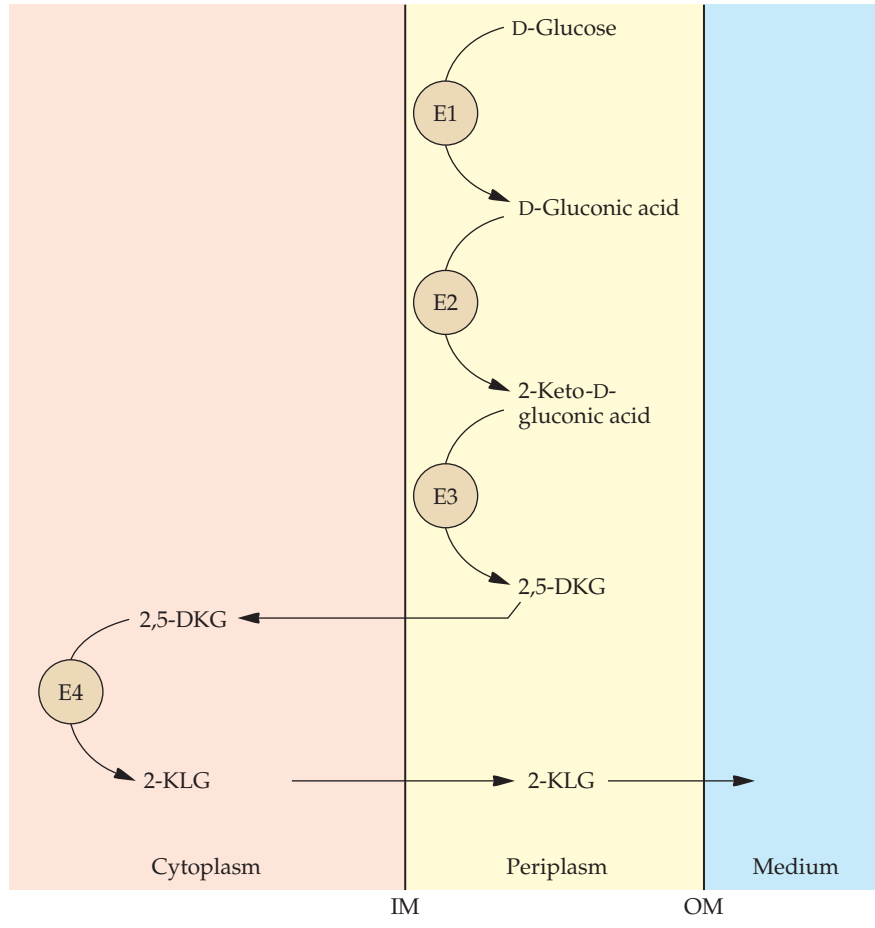


FIGURE 13.7 Conversion of D-glucose to 2-KLG by recombinant *E. herbicola*. The cellular locations of all of the component enzymes are indicated. The enzymes are denoted with the letter E and are numbered consecutively. Enzyme E4 is the introduced 2,5-DKG reductase. The major intermediates in the pathway are named. IM and OM denote the inner and outer membranes, respectively.

However, from the primary amino acid sequence, computer modeling predicted an enzyme structure with an eight-stranded α/β barrel (Fig. 13.8). This structure consisted of eight twisted parallel β-strands arranged close together, surrounded by eight α-helices that were joined to the β-strands through loops of various lengths. This folding pattern had previously been observed for 17 other enzymes whose crystal structures were known. By comparison with the structures of these other proteins, three of the loops that might be involved in substrate binding were identified (Fig. 13.8). Using oligonucleotide-directed mutagenesis, 12 different mutants, each with a single amino acid change in one of these loops, were constructed. Of the 12 mutants, 11 produced enzymes with a lower 2,5-DKG reductase specific activity than that of the native form of the enzyme. The 12th mutant, in which amino acid residue 192 was changed from glutamine to arginine, had approximately twice the activity of the native enzyme. Kinetic studies revealed that this increase in activity resulted from a 1.8-fold increase in the maximal rate of the enzyme-catalyzed reaction (V_{max}) and a 25% decrease in the Michaelis constant (K_m) of the enzyme-catalyzed reaction.

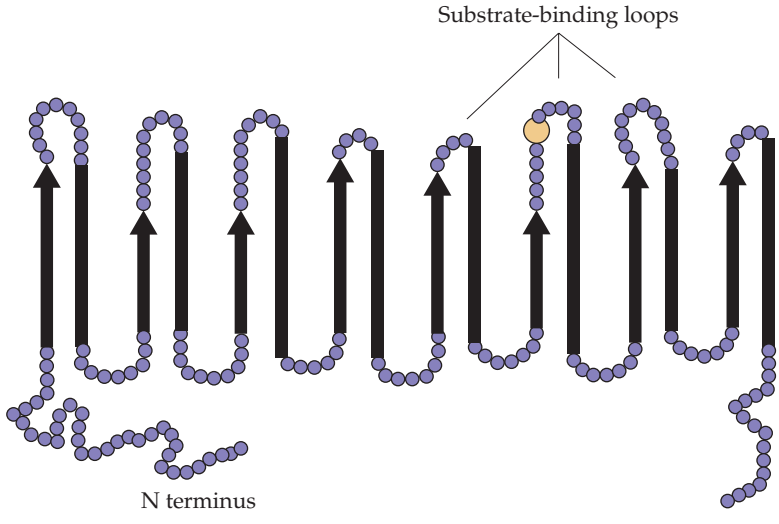


FIGURE 13.8 Predicted structure of 2,5-DKG reductase. The solid arrows indicate β-stranded regions, the solid bars are α-helical regions, and the circles are amino acid residues either at the N or C terminus or involved in loops connecting the β-strands to the α-helices. The three loops that may be involved in substrate binding are indicated. Amino acid residue 192 is shown in yellow.

The reaction catalyzed by 2,5-DKG reductase utilizes reduced nicotinamide adenine dinucleotide phosphate (NADPH) as a cofactor. However, the cellular concentration of reduced nicotinamide adenine dinucleotide (NADH) is usually about 10-fold greater than that of NADPH, while the financial cost of NADPH is about 10 times higher than that of NADH. To lower the cost of the bacterial production of ascorbic acid, it would be beneficial to engineer a version of 2,5-DKG reductase that used NADH instead of NADPH. The only structural difference between NADH and NADPH is the presence or absence of a phosphate group attached to the 2′ site of the adenine moiety. From the three-dimensional structure of 2,5-DKG reductase complexed with NADPH, it appears that 5 amino acid residues interact directly with the 2′ phosphate residue of NADPH. Using cassette mutagenesis (Fig. 13.9), a total of 40 different mutants of this enzyme were constructed; in each constructed mutant, 1 of the 5 amino acid residues that computer models suggested interacted with the 2′ phosphate residue was changed to a different amino acid. Following the expression, purification, and kinetic characterization of the 40 mutants in *E. coli,* it was observed that changing three of the five selected amino acids resulted in increases in 2,5-DKG reductase activity with NADH as the cofactor. In the best case, when the arginine residue at position 238 was changed to histidine, there was a sevenfold improvement over the wild type with NADH as a cofactor. Moreover, after two amino acid alterations were combined in one protein, an enzyme that showed even more activity with NADH included a change of the lysine residue at position 232 to glycine, as well as the change of arginine at position 238 to histidine. Also, when the best NADH-active mutant was combined with a double mutant which increased the binding of the substrate, even further improvements in activity with NADH were observed. The activity of the enzyme isolated from the final construct was 72 times higher than the activity of the wild-type enzyme. It now remains

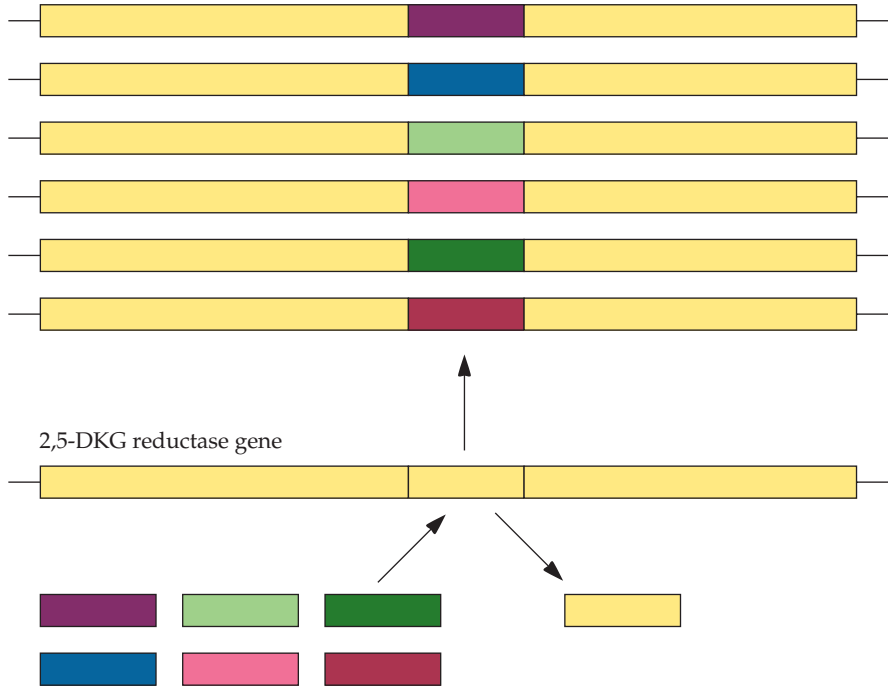


FIGURE 13.9 Mutagenesis of the *Corynebacterium* sp. 2,5-DKG reductase gene. The isolated gene, on a plasmid, was digested to remove a 57-bp DNA fragment. Then, chemically synthesized 57-bp DNA fragments that contained an alteration in the DNA sequence that coded for one of the amino acids to be changed was spliced into the gene in place of the original fragment. In addition, each 57-bp fragment (or mutagenesis cassette) contained a silent mutation that abolished a PstI site but did not alter the amino acid sequence. The silent mutation facilitated screening against the wild type.

to be seen whether this engineered enzyme can be used as the basis for the economically efficient biological synthesis of ascorbic acid.

Microbial Synthesis of Indigo

A number of bacteria, most notably *Pseudomonas* spp., have the ability to use a variety of organic compounds, such as naphthalene, toluene, xylene, and phenol, as their sole carbon source. In many instances, the genes encoding the enzymes for the degradation of these organic compounds are located on large, naturally occurring plasmids (typically 50 to 200 kb in size). Research on these bacteria often requires detailed genetic and biochemical studies, so that the genes encoding enzymes catalyzing important steps in the pathway can be targeted for modification. Occasionally, despite the original purpose of a study, an unexpected but useful discovery is made. For example, plasmid NAH7 has two separate and distinct operons that allow pseudomonads that contain this plasmid to grow on naphthalene as the sole carbon source. As a first step toward characterizing these genes, NAH7 plasmid DNA was digested with HindIII and the fragments were ligated with linear HindIII-digested plasmid pBR322. This clone bank was introduced into *E. coli* cells, and transformants were selected on the basis of their resistance to ampicillin and sensitivity to tetracycline. All

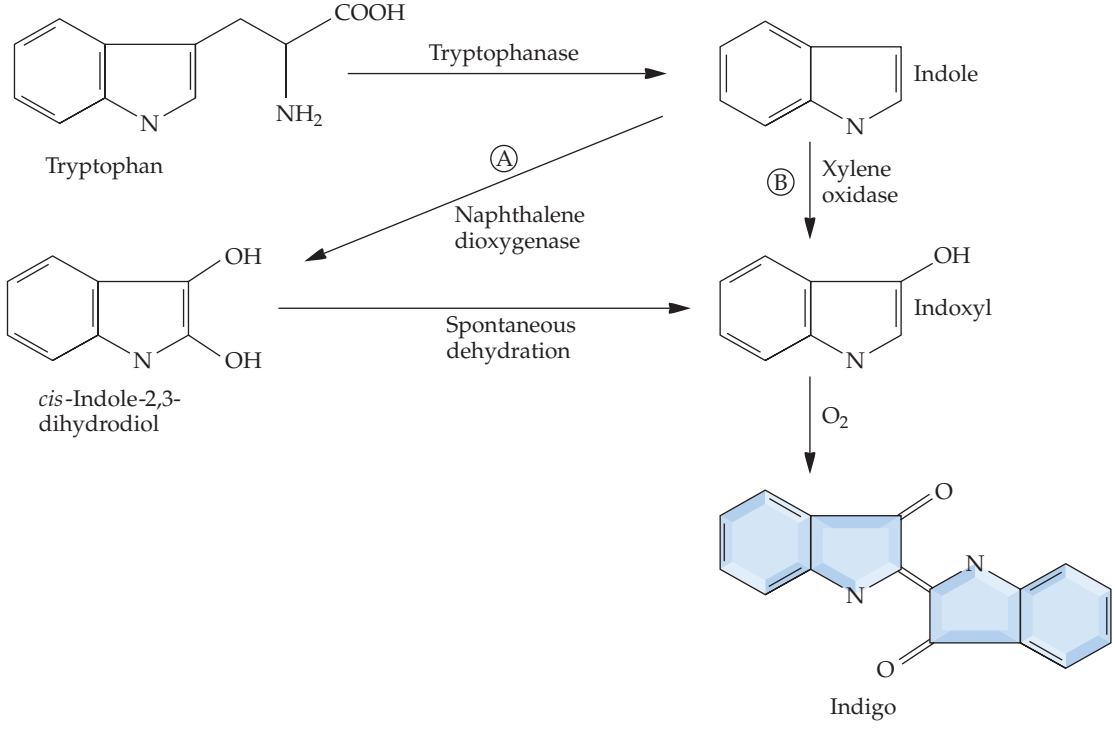
transformants were then tested for the production of nonvolatile metabolites that might result from the hydrolysis of radiolabeled naphthalene.

During the characterization of one of the transformants that had a 10.5-kb insert and could convert naphthalene to salicylic acid, it was observed that when the minimal growth medium contained tryptophan, it turned blue. A thorough analysis of the blue color revealed that the transformed *E. coli* cells were synthesizing the dye indigo. This synthesis was achieved in four steps (Fig. 13.10):

1. Conversion of tryptophan in the growth medium to indole by the enzyme tryptophanase, which is produced by the *E. coli* host cell
2. Oxidation of indole to *cis*-indole-2,3-dihydrodiol by naphthalene dioxygenase, which is encoded by the DNA that was cloned from the NAH7 plasmid
3. Spontaneous elimination of water
4. Air oxidation to form indigo

Thus, the combination of enzymes from two different pathways and two different organisms resulted in the synthesis of an unexpected compound, the dye indigo. In addition, introduction of the gene for the enzyme xylene oxidase, which is encoded in the TOL plasmid, can convert tryptophan to indoxyl, which then spontaneously oxidizes to indigo (Fig. 13.10).

FIGURE 13.10 Indigo biosynthesis from tryptophan in genetically engineered *E. coli*. Tryptophanase is an *E. coli* enzyme. In pathway A, the naphthalene dioxygenase is derived from the NAH plasmid; in pathway B, the xylene oxidase is from the TOL plasmid. *E. coli* transformants that synthesize indigo contain either pathway A or B, but not both pathways.



Indigo, a commercially important blue pigment that is used to dye both cotton and wool, was originally isolated from plants but is currently synthesized chemically. At present, approximately 13×10^6 kg of indigo, worth more than $250 million, is produced every year. Indigo, the coloring agent in blue jeans, is the largest-selling dye in the world. The ability to produce indigo from bacteria opens the possibility of developing an efficient and economical commercial microbial process for its production. This process would avoid the use of hazardous compounds, such as aniline, formaldehyde, and cyanide, which are needed in the chemical synthesis of indigo.

Despite the environmentally friendly nature of engineering bacteria to make indigo, at the present time, the chemical synthesis of the dye is less expensive, thwarting commercial schemes to produce indigo biologically. One approach to improving indigo production in *E. coli* involves engineering the host strain to overproduce tryptophan, the raw material for the process. However, despite improved efficiency when a tryptophan-overproducing *E. coli* strain is used, the overall process requires additional improvement before it is economically competitive with chemical synthesis.

Synthesis of Amino Acids

Amino acids are used extensively in the food industry as flavor enhancers, antioxidants, and nutritional supplements; in agriculture as feed additives; in medicine in infusion solutions for postoperative treatment; and in the chemical industry as starting materials for the manufacture of polymers and cosmetics (Table 13.1). It is estimated that more than 2.5 million tons of amino acids, worth more than $9 billion, were produced worldwide in 2008. L-Glutamic acid, which is used in the manufacture of the flavor enhancer monosodium glutamate, makes up around half of the total volume.

TABLE 13.1 Commercial applications of amino acids

Amino acid	Application(s)
Alanine	Flavor enhancer
Arginine	Therapy for liver diseases
Aspartic acid	Flavor enhancer; sweetener synthesis
Asparagine	Diuretic
Cysteine	Bread production; therapy for bronchitis; antioxidant
Glutamic acid	Flavor enhancer
Glutamine	Therapy for ulcers
Glycine	Sweetener synthesis
Histidine	Therapy for ulcers; antioxidant
Isoleucine	Intravenous solutions
Leucine	Intravenous solutions
Lysine	Feed additive; food additive
Methionine	Feed additive
Phenylalanine	Infusions; sweetener synthesis
Proline	Intravenous solutions
Serine	Cosmetics
Threonine	Feed additive
Tryptophan	Intravenous solutions; antioxidant
Tyrosine	Intravenous solutions; precursor for L-DOPA
Valine	Intravenous solutions

For the most part, amino acids are commercially produced either by extraction from protein hydrolysates or as fermentation products of either *Corynebacterium* or *Brevibacterium* spp., which are both nonsporulating gram-positive soil bacteria that secrete large amounts of amino acids into the growth medium. Traditionally, the productivity of these organisms has been improved by mutagenesis and subsequent screening for strains that overproduce certain amino acids. However, this way of developing new strains is slow and inefficient. By using detailed biochemical information about the enzymes that are involved in the biosynthesis of various commercially important amino acids, it is more expeditious to isolate and manipulate the specific genes encoding the key components of a particular pathway. However, this type of genetic engineering is not a simple matter. For example, the pathway(s) leading to the biosynthesis of certain amino acids contains a number of different enzymes, each of which may be either activated or inhibited by a number of metabolites present in the cell. This makes it difficult to know which enzyme(s) to manipulate in order to enhance the yield of the end product.

Because most broad-host-range plasmid vectors replicate only in gram-negative organisms, it is necessary to construct expression vectors that are specifically suited for *Corynebacterium* and *Brevibacterium* spp. Such cloning vehicles might take the form of *E. coli–Corynebacterium* shuttle vectors. The *E. coli* portion of the plasmid could encode resistance to the antibiotic tetracycline, chloramphenicol, or kanamycin. Because both *E. coli* and *Corynebacterium* spp. are susceptible to these antibiotics, they could be used as selectable markers in both organisms.

An efficient transformation protocol for *Corynebacterium glutamicum,* the species of *Corynebacterium* often used in these experiments, is still needed. Also, many *C. glutamicum* genes are not efficiently expressed in *E. coli*. Therefore, for selection schemes that depend on gene expression (e.g., complementation), the entire clone bank should be transformed into *C. glutamicum*. Unfortunately, the transformation frequency is very low when DNA is introduced into *C. glutamicum* by either direct transformation or electroporation. However, effective transformation of *C. glutamicum* is achieved when foreign DNA is introduced by conjugation or after the formation of protoplasts, i.e., after removal of the cell wall with lysozyme. The transformation of protoplasts is made possible by adding polyethylene glycol to facilitate the uptake of exogenous plasmid DNA.

Some progress has been made in increasing the amino acid output of *C. glutamicum*. For example, the synthesis of the essential amino acid tryptophan was enhanced by introducing into wild-type *C. glutamicum* cells a second copy of the gene encoding anthranilate synthetase, which is the rate-limiting enzyme in the normal tryptophan biosynthetic pathway (Fig. 13.11). The following protocol describes one way to isolate the anthranilate synthetase gene.

1. A library of *Brevibacterium flavum* chromosomal DNA was cloned into a *C. glutamicum–E. coli* shuttle vector and introduced into a mutant strain of *C. glutamicum* that produced no active anthranilic acid synthetase.
2. The mutant strain was unable to grow on minimal medium unless anthranilic acid was added; therefore, transformants were selected by their ability to grow in the absence of anthranilic acid.
3. The vector carrying the anthranilic acid synthetase gene was then transferred to a wild-type strain of *C. glutamicum*.

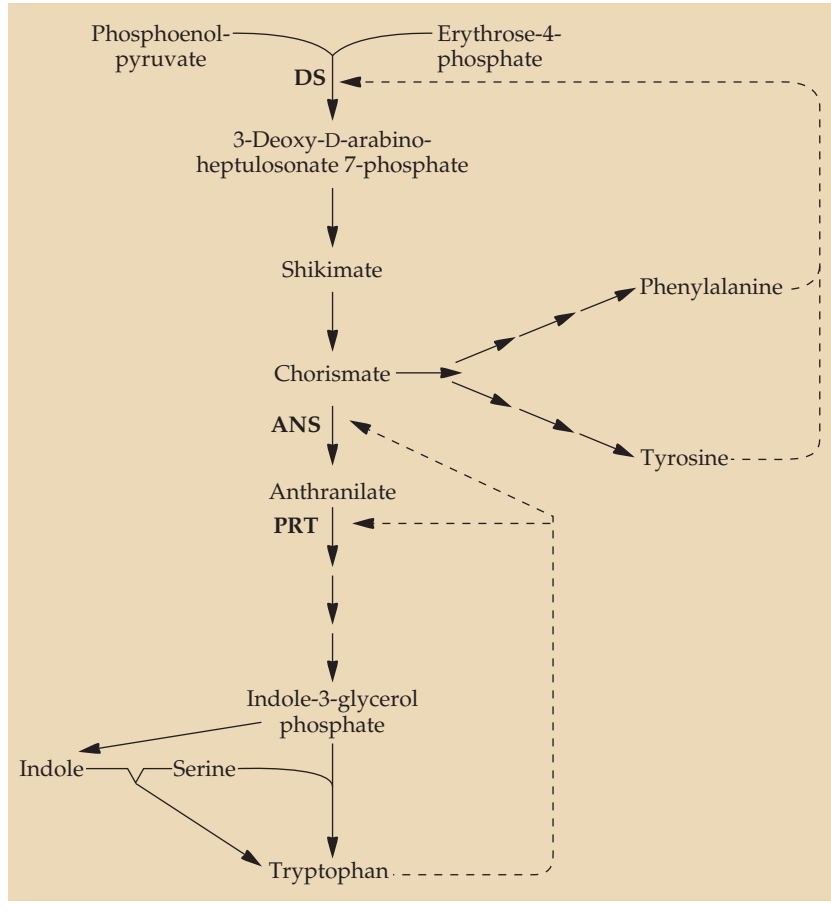


FIGURE 13.11 Simplified pathway and regulation of tryptophan biosynthesis in *C. glutamicum*. The abbreviations DS, ANS, and PRT represent the enzymes 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase, anthranilate synthase, and anthranilate phosphoribosyltransferase, respectively. The solid lines represent the synthetic pathways. The dashed lines denote feedback inhibition. Indole is produced in a side reaction and is converted into tryptophan by the action of tryptophan synthase β.

The amounts of tryptophan produced in the mutant and wild-type *C. glutamicum* strains—one without and one with the vector carrying the cloned anthranilic acid synthetase gene—were measured (Table 13.2). The cloned gene did indeed restore most of the capacity of the mutant to synthesize tryptophan. Moreover, the effect of adding this gene to wild-type *C. glutamicum* was much more dramatic, with the synthesis of tryptophan being increased by approximately 130%. This level of overproduction reflects more efficient utilization of available precursor material. Thus, by cloning an additional gene of an amino acid biosynthesis pathway into an organism, it was possible to generate much more of the end product. An even higher level of tryptophan production was achieved when modified genes for the three key enzymes, 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase, anthranilate synthase, and anthranilate phosphoribosyltransferase, were introduced into *C. glutamicum* cells (Fig. 13.11). The genes encoding these enzymes were mutagenized to render them insensitive to inhibition by the end product (feedback inhibition).

TABLE 13.2 Production of tryptophan under standard growth conditions by certain strains of *C. glutamicum*

Strain	Tryptophan concentration (mg/mL)
Mutant	0.00
Mutant with vector	0.34
Wild type	0.48
Wild type with vector	1.12

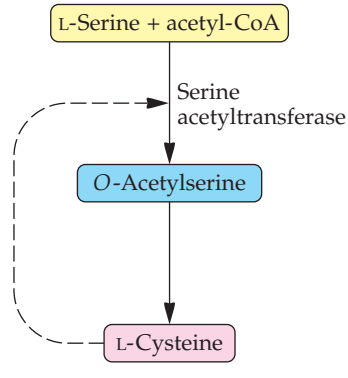
Adapted from Ozaki et al., U.S. patent 4,874,698, 1989.

An alternative to producing amino acids in *Corynebacterium* and *Brevibacterium* spp. is to produce them in *E. coli,* where both metabolic pathways and procedures for genetic manipulation are much better characterized. Its relative ease of manipulation makes *E. coli* an attractive host organism for this sort of metabolic engineering.

L-Cysteine, which is one of the most important amino acids in the pharmaceutical, food, and cosmetics industries, has traditionally been obtained by extracting it from acid hydrolysates of human hair and animal feathers. While a number of microorganisms are able to synthesize L-cysteine, high levels cannot be synthesized from glucose because L-cysteine feedback inhibits the enzyme serine (Ser) acetyltransferase, which catalyzes one of the steps in the biosynthesis of L-cysteine (Fig. 13.12). To try to remedy this situation, the methionine residue at position 256 in the *E. coli* serine acetyltransferase amino acid sequence was systematically changed to each of the other 19 amino acids. When *E. coli* was transformed with plasmids carrying *cysE* genes encoding altered serine acetyltransferase, several transformants with altered forms of serine acetyltransferase produced higher levels of L-cysteine than did the wild-type enzyme. Next, plasmids encoding the most effective serine acetyltransferase derivatives were used to transform a strain of *E. coli* that did not degrade L-cysteine. To improve on this modest success, complementary DNAs (cDNAs) encoding feedback inhibition-insensitive serine acetyltransferases from the plant *Arabidopsis thaliana* were expressed in a serine acetyltransferase-deficient and non-L-cysteine-utilizing *E. coli* strain. The transformants included several different strategies and produced a much higher level of L-cysteine than had previously been possible by merely manipulating the *E. coli* serine acetyltransferase gene. Whether the expression of this protein can be further improved and how this affects the synthesis of L-cysteine remain to be determined.

To rationally engineer a bacterium to modify its metabolism so that it overproduces a particular amino acid, it is essential to understand how many of the metabolic pathways of the bacterium are interrelated and regulated. Subsequent systematic reengineering of the metabolism of the organism can then be performed to achieve a much higher yield than would ever be possible by merely modifying one or two genes in the immediate biosynthetic pathway of a particular amino acid. At each stage in the development of the reengineered bacterium, it is possible to monitor the levels of a wide range of transcripts (using microarrays) and metabolites (metabolomics). Moreover, the mRNA and metabolite expression data allow the construction of a detailed computer model that can be used to predict the effects of other possible genetic manipulations on the production of the target amino acid. While this approach is relatively new, such

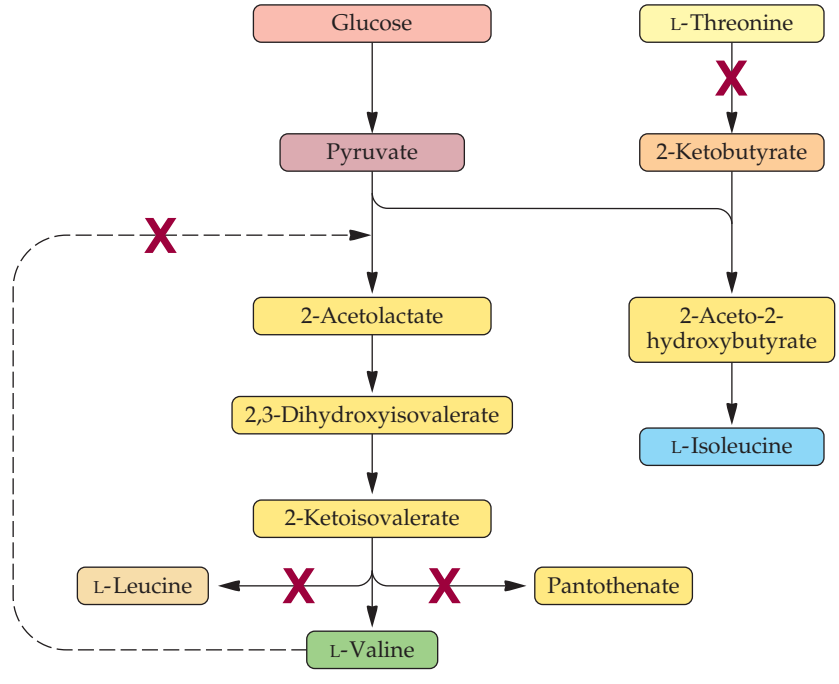
FIGURE 13.12 Biosynthesis of L-cysteine from L-serine and acetyl-CoA. The dashed arrow indicates feedback inhibition.



rational genetic engineering may be used to engineer bacterial strains for a variety of purposes, in addition to the production of specific amino acids.

To rationally engineer *E. coli* to overproduce L-valine, it was necessary to introduce a large number of mutations into the *E. coli* genome (Fig. 13.13). The flux from glucose to L-valine was first improved by abolishing feedback inhibition by L-valine of the enzyme that converts pyruvate to 2-acetolactate. Then, the subunits of this enzyme were overexpressed by replacing the endogenous promoter with a strong constitutive promoter. Next, the carbon flux toward L-valine was increased. The gene for the enzyme that converts L-threonine to 2-ketobutyrate was knocked out, as were the genes encoding enzymes that convert 2-ketoisovalerate to either pantothenate or L-leucine. In addition, some of the genes that encode enzymes that convert pyruvate to 2-acetolactate were amplified. Subsequently, the expression levels of the genes encoding enzymes that convert 2-acetolactate to L-valine via 2,3-dihydroxyisovalerate and 2-keto-isovalerate were all increased by modifying the regulatory regions of these genes. In addition, the *lrp* gene, encoding the leucine-responsive regulatory protein, was amplified, since some of genes in the L-valine pathway were under the positive transcriptional control of this protein. Finally, the expression of the two genes that encode proteins responsible for the export of L-valine from the bacterial cell was amplified, since it was thought that a low level of these proteins was limiting to L-valine production. Following the extensive metabolic engineering of *E. coli*, the final strain was able to

FIGURE 13.13 Simplified overview of the production of L-valine in rationally engineered *E. coli*. The red X's indicate pathways that were knocked out, the green arrows indicate pathways that were upregulated, and the dashed blue line indicates feedback inhibition. Some arrows represent a single enzymatic step, while other arrows represent several enzymatic steps. The interaction of glucose with other pathways, the export of L-valine from the cell, and several more complex regulatory steps are not shown.



produce 0.378 g of L-valine per g of glucose, which is higher than industrial strains of *C. glutamicum* that have been developed by repeated rounds of random mutagenesis and selection. Moreover, researchers believe that the rationally engineered strain can be further modified with resulting increases in L-valine productivity. Finally, it is important to note that the approach used here should in principle be useful for the overproduction of a wide range of metabolites from bacteria.

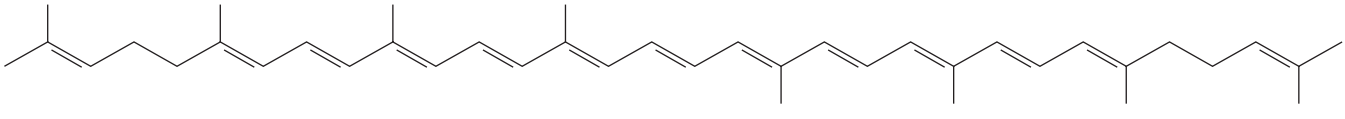
Microbial Synthesis of Lycopene

Lycopene ($C_{40}H_{56}$) is a bright red carotenoid pigment (Fig. 13.14) that is commonly found in tomatoes and other fruits, including watermelons, pink grapefruit, pink guavas, papayas, and rose hips. It is a powerful antioxidant that has been suggested to decrease low-density lipoprotein oxidation in humans and thereby to lower the risk of atherosclerosis and coronary heart disease. In addition, lycopene, and several carotenoids derived from lycopene, have been proposed as treatments for some types of cancer. It would be useful if lycopene (and carotenoids produced from lycopene) could be produced in microorganisms so that the large-scale-processing problems that exist when lycopene is isolated from tomatoes might be avoided. For the production of lycopene in *E. coli,* the 2-*C*-methyl-D-erythritol 4-phosphate pathway provides the precursors isopentenyl diphosphate and dimethylallyl diphosphate (Fig. 13.15). However, by introducing *Saccharomyces cerevisiae* genes encoding the mevalonate pathway under the control of *E. coli* transcriptional promoters, the levels of these precursor molecules are increased dramatically. By subsequently introducing *Pantoea agglomerans* (a gram-negative bacterium) genes encoding the biosynthesis of lycopene from the above-mentioned precursor metabolites, a relatively high level of lycopene can be produced (Fig. 13.15). With this engineered *E. coli* strain, which contains two additional biosynthetic pathways, it was possible to obtain approximately 60 mg of lycopene per liter of bacterial culture. While additional optimization of this system is still necessary before it can be the basis of a commercial system for lycopene production, this work is an important step in the development of such a system.

Increasing Succinic Acid Production

Succinic acid (succinate) is a dicarboxylic acid that is a component of the citric acid cycle; it is formed from fumarate and reacts to form succinyl-coenzyme A (CoA) (Fig. 13.16). At room temperature, pure succinic acid is a colorless and odorless solid that is moderately soluble in water. Succinic acid is used as a flavoring for foods and beverages and in the production of dyes, perfumes, lacquers, resins, and a variety of medicines. It is currently synthesized by the catalytic hydrogenation of malic acid or its anhydride; however, there is increasing interest in the production of succinic acid from renewable sources by microbial fermentation.

FIGURE 13.14 The chemical structure of lycopene.



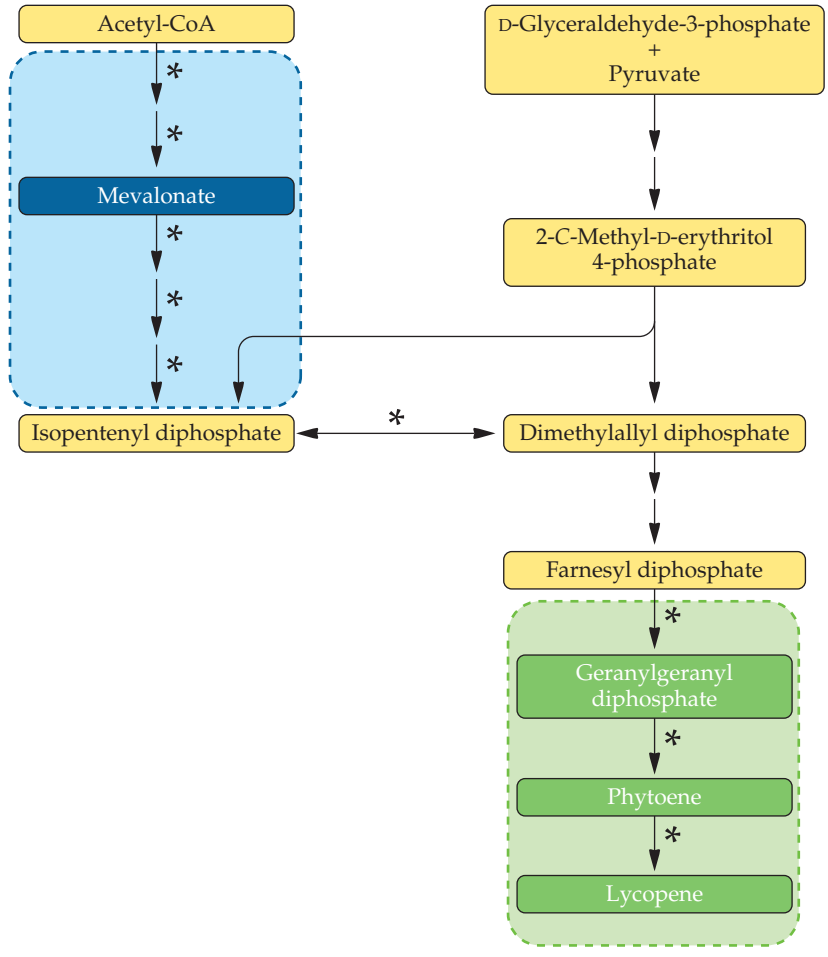


FIGURE 13.15 Overview of the production of lycopene in an engineered strain of *E. coli*. The yellow metabolites are normally produced in *E. coli*. The portion of the pathway highlighted in blue represents the enzymes encoded by *S. cerevisiae* genes, as well as some of the metabolites that they produce. The pathway highlighted in green represents enzymes encoded by *P. agglomerans* and the metabolites that they produce. Introduced enzymes are depicted, adjacent to the arrows, by asterisks.

Although a very large number of bacteria synthesize succinic acid, only a few of these organisms, including *Anaerobiospirillum succiniproducens*, *Actinobacillus succinogenes*, and *Mannheimia succiniproducens*, produce the metabolite at high levels. Unfortunately, at the same time that these anaerobic bacteria produce succinic acid, they also produce, and excrete, significant amounts of acetic, formic, and lactic acids. This not only reduces the yield of succinic acid, it also makes the purification process more difficult and costly. To increase the amount of succinic acid produced by *M. succiniproducens*, genes that were known to be involved in the synthesis of acetic, formic, and lactic acids from pyruvic acid were sequentially disrupted (Fig. 13.16), and each mutant was tested for the ability to synthesize succinic (and acetic, formic, and lactic) acid. Using this strategy, it was possible to engineer a strain, under the anaerobic conditions that are normal for the bacterium, to produce 13.5 g of succinic acid per liter of culture compared to 10.5 g per liter for the wild-type bacterium. At the same time, the production of formic and lactic acids was completely

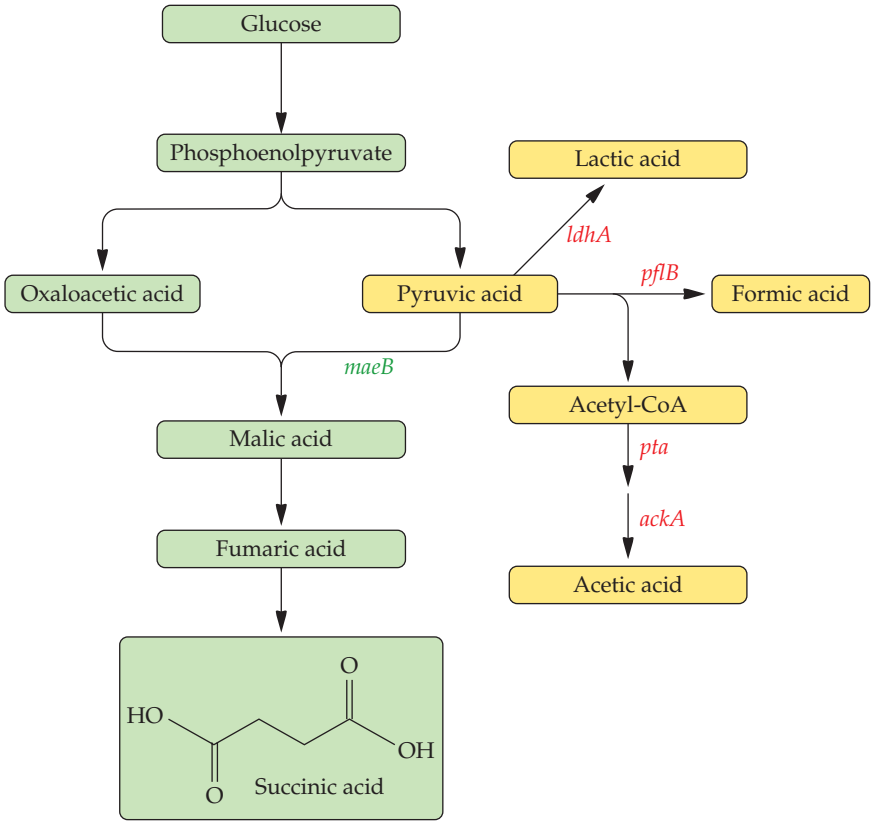


FIGURE 13.16 Overview of the metabolic pathways that lead to the formation of succinic, formic, lactic, and acetic acids under anaerobic conditions in *M. succiniproducens*. The metabolites indicated in green highlight the production of succinic acid from glucose. The metabolites shown in yellow depict the synthesis of unwanted side products. The four genes in red (i.e., *ldhA, pflB, pta,* and *ackA*) were mutated to decrease the flux through this portion of the pathway. The gene in green *(maeB)* may be overproduced in an effort to decrease pyruvic acid and increase succinic acid. *ldh,* lactate dehydrogenase; *pfl,* pyruvate formate lyase; *pta,* phosphotransacetylase; *ack,* acetate kinase; *mae,* malic enzyme.

abolished and the amount of acetic acid was significantly reduced (Fig. 13.17). In addition, when the cells that contained four separate mutations were grown in a fed-batch mode (see chapter 17), the yield of succinic acid increased to 52.4 g per liter while the amount of pyruvate was only 0.8 g per liter. Although this modified strain secreted pyruvic acid into the medium, it is technically simpler to remove pyruvic acid than acetic, formic, or lactic acid. Nevertheless, it is hoped that additional metabolic engineering (e.g., by overproducing MaeB) will both increase the amount of succinic acid and decrease the level of pyruvic acid that this strain produces so that the modified bacterium can be used as a biological "factory" for succinic acid synthesis.

Antibiotics

Since the discovery of penicillin in the late 1920s, more than 12,000 antibiotics with different specificities and a variety of modes of action have been isolated from various microorganisms. The universal use of antibiotics to treat bacterial diseases has resulted in an enormous improvement in human

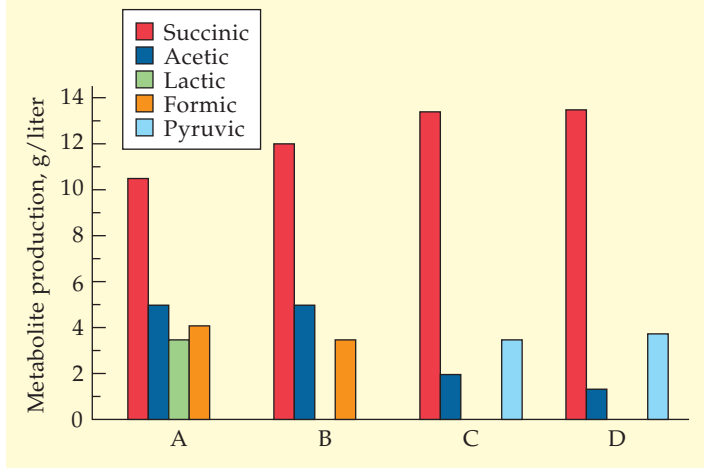


FIGURE 13.17 Formation of succinic, formic, lactic, acetic, and pyruvic acids by wild-type and mutant strains of *M. succiniproducens*. Bars: A, wild type; B, *ldhA* mutant; C, *ldhA* and *pflB* mutant; D, *ldhA, pflB, pta,* and *ackA* mutant.

health and has undoubtedly saved millions of lives. The majority of the most important antibiotics have been isolated from the gram-positive soil bacterium *Streptomyces,* although fungi and other gram-positive and gram-negative bacteria are also sources of antibiotics (Table 13.3). Worldwide, over 100,000 tons of antibiotics is produced per year, with annual gross sales of about $35 billion, including antibiotics used in animal feed and as animal growth promoters. The antibiotic market is driven by the sales of four leading drug classes: the cephalosporins (27%), macrolides (20%), quinolones (17%), and penicillins (17%). Together, these four drug classes account for more than 80% of global antibacterial sales.

An estimated 200 to 300 new antibiotics are discovered each year, primarily through labor-intensive research programs in which many thousands of different microorganisms are screened to find those that produce unique antibiotics. However, with the high costs of development and clinical testing, only the compounds that show significant therapeutic and economic promise are marketed. Therefore, only about 1 to 2% of newly discovered antibiotics are added annually to the disease-fighting arsenal. In fact, the pharmaceutical industry has been reluctant to invest in research and development in this area, and many companies have either abandoned or scaled down their efforts since 1999. In addition, to date, nearly all of the genetic improvements to industrially important antibiotic-producing strains have been achieved by the use of classical mutagenesis and selection. While the yields of antibiotics from many strains have been signifi-

TABLE 13.3 Some of the most common microbially synthesized antibiotics

Amikacin sulfate	Cefotaxime	Chlortetracycline	Kanamycin sulfate	Streptomycin sulfate
Amoxicillin	Cefoxitin	Clarithromycin	Lincomycin HCl	Teicoplanin
Ampicillin	Cefpodoxime proxetil	Clindamycin	Methicillin	Tetracycline HCl
Azithromycin	Ceftazidime	Erythromycin A	Oxytetracycline	Vancomycin HCl
Benzylpenicillin	Ceftriaxone	Flomoxef	Phenoxymethylpenicillin	
Cefaclor	Cefuroxime	Gentamicin sulfate	Rifampin	
Cefixime	Cephalexin	Imipenem	Spiramycin	

cantly improved—the original penicillin-producing fungus isolated by Alexander Fleming yielded 2 units per milliliter of culture, while the strains used today synthesize approximately 70,000 units per milliliter of culture—this yield improvement took many years and required the use of considerable manpower and financial resources. Recombinant DNA technology can have a positive impact on this endeavor in two ways. First, the technology can be used to develop new, structurally unique antibiotics with increased activities against selected targets and decreased side effects. Second, genetic manipulation can be used to relatively rapidly and inexpensively enhance yields and hence lower the cost of production of existing antibiotics.

For the genetic manipulation of *Streptomyces*, it is essential that it can be transformed and that the transformed cells can be readily selected. However, unlike *E. coli*, *Streptomyces* strains do not exist as individual cells but as extended aggregates called mycelial filaments. The cell wall must be removed to release individual cells (protoplasts) before DNA transformation (Fig. 13.18). Without this step, it would not be possible to distinguish transformed from nontransformed cells, because visible colonies on a solid medium would each have started from a cell aggregate rather than from an individual cell. Thus, colonies that grew in the presence of a selective antibiotic would contain a mixture of transformed and nontransformed cells. However, as a consequence of protoplast formation prior to transformation, all colonies that grow in the presence of a selective antibiotic contain only transformed cells. The uptake of plasmid DNA into *Streptomyces* protoplasts is enhanced by polyethylene glycol. Following transformation, the protoplasts are first plated onto a solid medium to enable the cell walls to regenerate and are then overlaid with a selective medium that often contains either neomycin or thiostrepton, both of which act as selection agents for transformed cells.

Cloning Antibiotic Biosynthesis Genes

The biosynthesis of an antibiotic may include 10 to 30 separate enzyme-catalyzed steps, so cloning all the genes for the synthesis of a particular antibiotic is not an easy task. One strategy for isolating the complete set of antibiotic biosynthesis genes consists of transforming one or more mutant strains that are unable to synthesize the antibiotic with DNA from a clone bank constructed from wild-type chromosomal DNA. Following the introduction of the clone bank DNA into mutant cells, transformants are screened for the ability to produce the antibiotic. Then, the plasmid DNA from the clone that supplies a functional gene and gene product, i.e., complements a mutant strain, is used as a DNA hybridization probe to screen another clone bank of wild-type chromosomal DNA (i.e., one in which the average-size fragment is around 10 kb) to isolate clones with regions that overlap the probe sequence. In this way, DNA segments that are adjacent to and usually bigger than the initial complementing DNA can be identified and cloned. A complete gene cluster can be reconstructed from the overlapping clones. If the antibiotic biosynthesis genes are clustered at a single site on the chromosomal DNA, the genes that are adjacent to the complementing gene are also likely to be involved in the biosynthesis of the target antibiotic. However, if the antibiotic biosynthesis genes are scattered in small clusters at different chromosomal locations, it is necessary to have at least one mutant per gene cluster to obtain a DNA clone that can be used to identify the rest of the genes in the cluster.

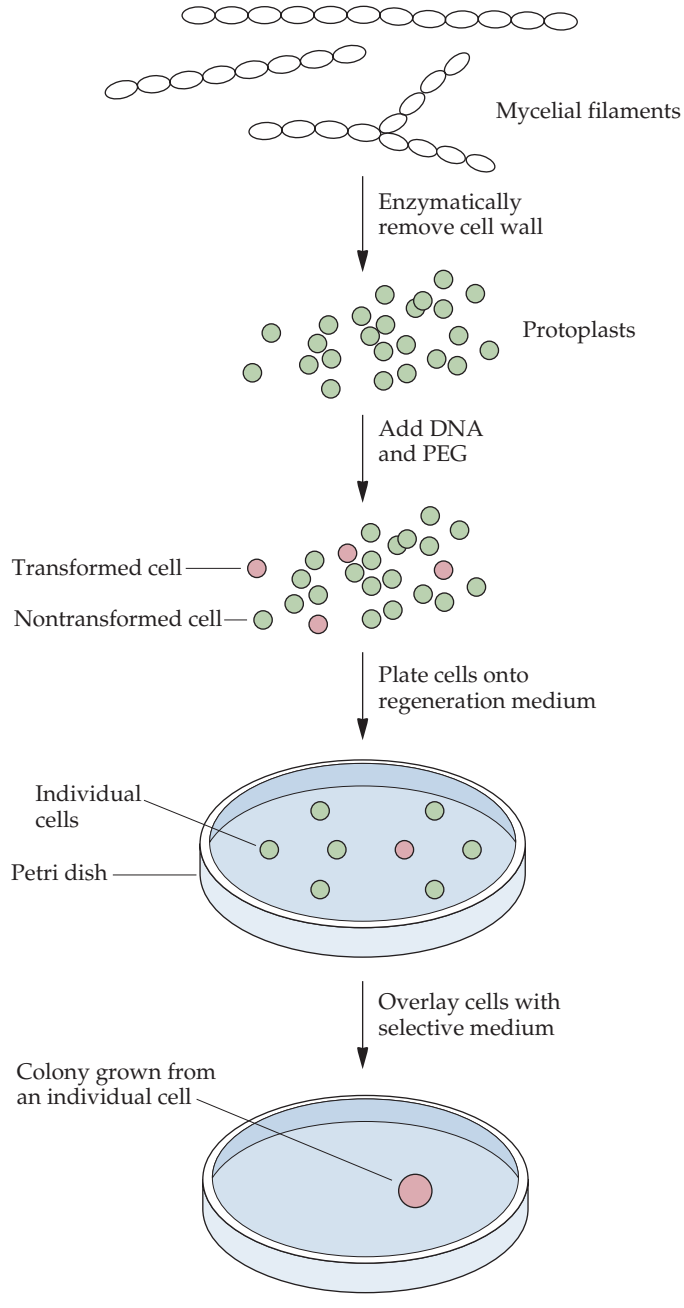


FIGURE 13.18 Schematic representation of DNA transformation and selection of transformants of *Streptomyces* strains. The pink circles represent transformed cells, and the green circles represent nontransformed cells. PEG, polyethylene glycol.

The complementation approach has been used to isolate some of the genes for the biosynthesis of the antibiotic undecylprodigiosin from *Streptomyces coelicolor* A3 (Fig. 13.19). In this case, the complementation assay is simple and entails scoring the color of the colonies. Colonies of wild-type organisms are red because of the presence of the antibiotic, and mutant colonies are cream colored. Complementation produces a red colony (Fig. 13.20).

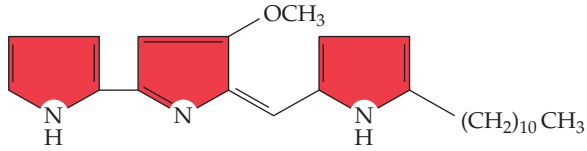
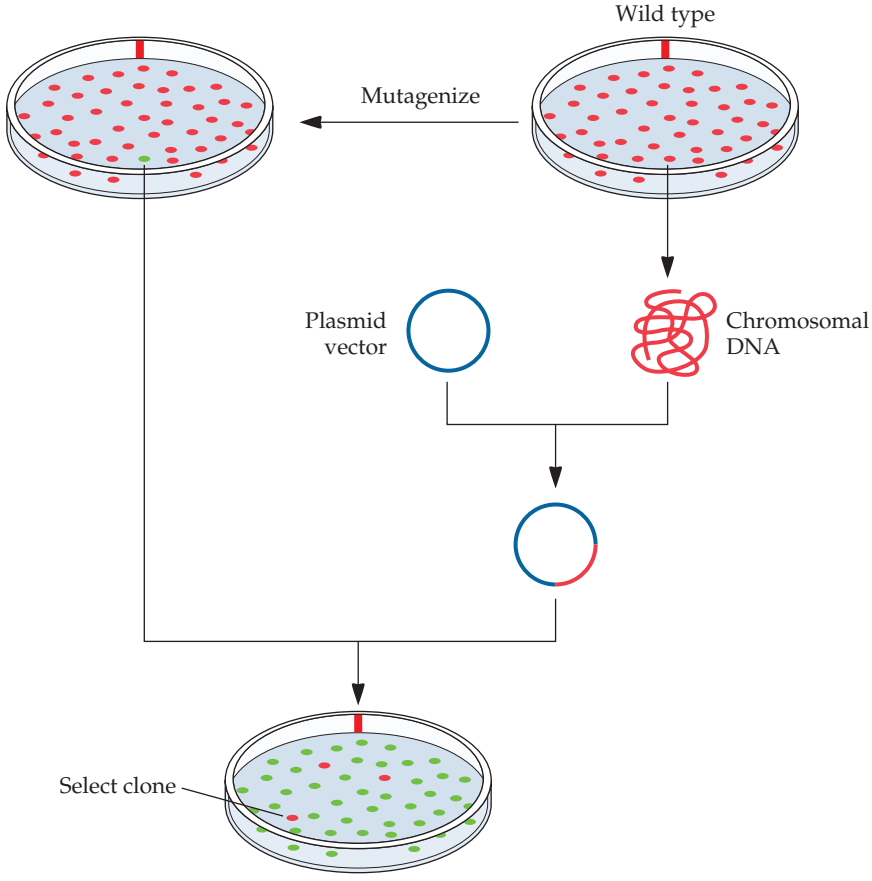


FIGURE 13.19 Chemical structure of the red antibiotic undecylprodigiosin.

In addition to cloning antibiotic biosynthesis genes by complementation, more direct strategies can be employed. One or more of the key enzymes in a biosynthetic pathway can be identified through either genetic or biochemical studies and then purified. The N-terminal amino acid sequence of the enzyme can then be determined, and with this information, oligodeoxyribonucleotide probes for the gene can be prepared. This approach has been used to isolate the gene for isopenicillin N synthetase from *Penicillium chrysogenum*. This enzyme catalyzes the oxidative

FIGURE 13.20 Method for cloning genes involved in the biosynthesis of the antibiotic undecylprodigiosin. Chromosomal DNA from wild-type antibiotic-producing cells is spliced into a *Streptomyces* cloning vector. The clone bank is used to transform a noncolored (i.e., non-antibiotic-producing) mutant of the wild type. Red transformants (in which the mutant has been complemented) are selected, and the plasmid DNA insert is characterized.



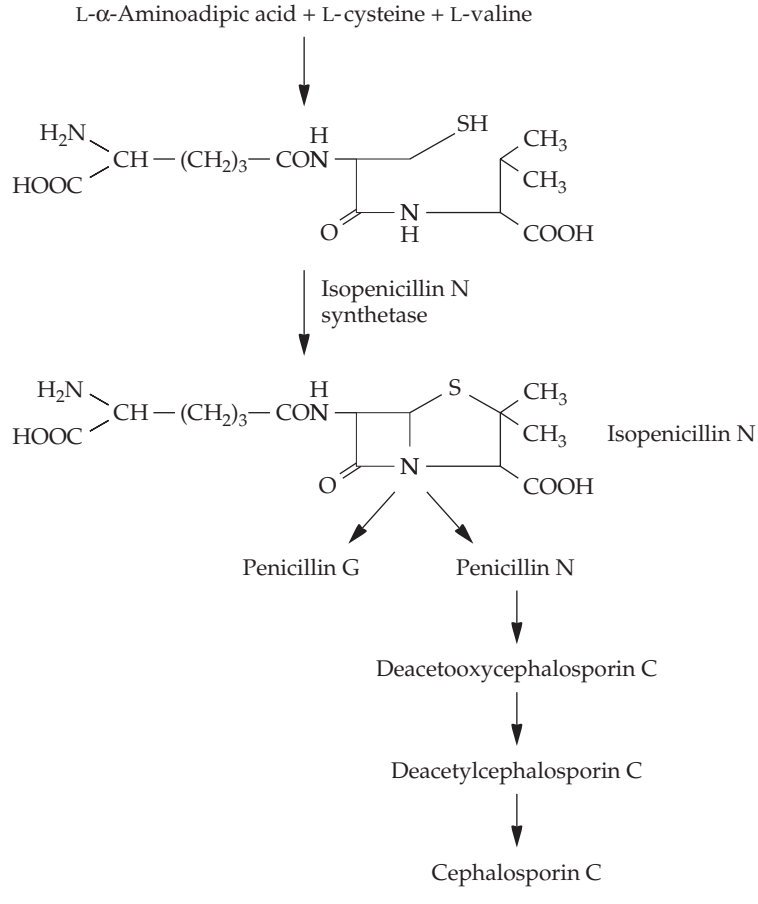
condensation of the compound δ-(L-α-aminodipyl)-L-cysteinyl-D-valine to isopenicillin N, a key intermediate in the biosynthesis of penicillins, cephalosporins, and cephamycins (Fig. 13.21).

Despite the difficulty, there are a number of examples of the cloning and transfer of large fragments of DNA encoding entire antibiotic biosynthetic pathways. In these cases, it is usually necessary to use a vector that can accept and maintain pieces of DNA as large as 100 kb. For this purpose, researchers have employed bacterial artificial chromosomes that have been engineered to replicate autonomously in *E. coli* and, when they are introduced into *Streptomyces,* to integrate into the chromosome.

Modulating Gene Expression in Streptomycetes

To improve the productivity of antibiotics produced by *Streptomyces* spp., it is desirable to utilize a regulatory expression system that can suppress the expression of the target gene(s) until the culture reaches a high cell density. It would also be beneficial if the system could be induced simply and inexpensively and if it could function in a range of different *Streptomyces* spp. One regulatable, high-expression *Streptomyces* system that was recently

FIGURE 13.21 Biosynthetic pathway for penicillins and cephalosporins in *P. chrysogenum*. Isopenicillin N synthetase catalyzes the synthesis of isopenicillin N from D-(L-α-aminoadipyl)-L-cysteinyl-D-valine. Isopenicillin N is a precursor in the synthesis of penicillin G, penicillin N, and cephalosporin C.



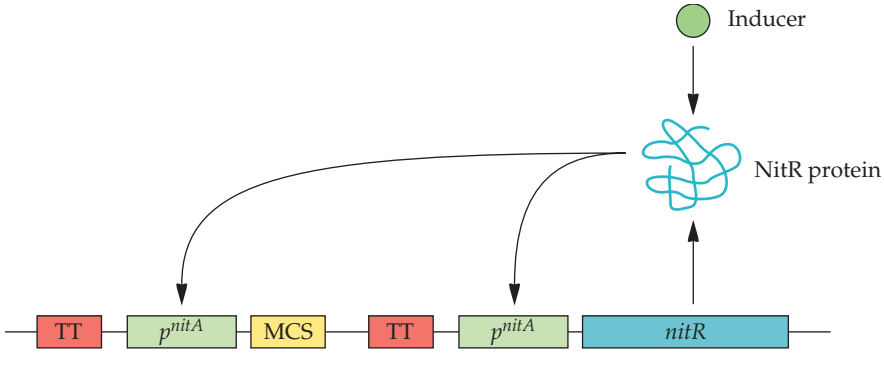
developed utilizes the nitrilase operon from the actinomycete *Rhodococcus rhodochrous*. In this bacterium, the expression of the enzyme nitrilase (encoded by the *nitA* gene under the control of the *nitA* promoter) is positively regulated by the protein NitR (encoded by *nitR,* which is present in the same operon as *nitA*). NitR forms a complex with the inducer, ε-caprolactam, before the complex binds to the *nitA* promoter, activating the synthesis of both proteins, NitA and NitR. However, to use this system as part of a high-expression vector, a transcription terminator was placed upstream of the *nitA* promoter (to prevent transcriptional read-through from other genes (Fig. 13.22); a multiple cloning site for inserting target DNA was placed downstream of the *nitA* promoter; downstream of the multiple cloning site (and the target gene), the synthetic operon ends with another transcription terminator; and finally, the vector contains the *nitR* gene under the control of the *nitA* promoter in a separate operon. When this expression vector is used, upon the addition of the inducer ε-caprolactam, the inducer–NitR protein complex activates the transcription of both the target gene(s) and *nitR*. The increased level of NitR results in a very high level of production of the target protein(s). While this system is still at a relatively early stage of development, with some target proteins, it has been possible to achieve expression levels as high as approximately 40% of all soluble protein. Moreover, the system functioned well when it was introduced into *S. coelicolor, Streptomyces avermitilis,* and *Streptomyces griseus,* bacterial strains that have all previously been used for antibiotic production. However, it still remains to be demonstrated that this expression system can be used as an effective method for increasing the yield of a commercially important antibiotic that is produced in a *Streptomyces* strain.

Synthesis of Novel Antibiotics

New antibiotics with unique properties and specificities may be produced by genetic manipulation of the genes involved in the biosynthesis of existing antibiotics. In one of the first experiments in which a novel antibiotic was produced, researchers began by examining the consequences of placing two slightly different antibiotic production pathways into one organism.

A *Streptomyces* plasmid (pIJ2303) carrying a 32.5-kb fragment of *S. coelicolor* chromosomal DNA contains all of the genes encoding the enzymes

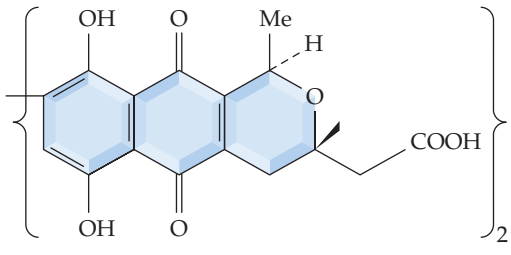
FIGURE 13.22 Overview of a regulatable expression system for use with *Streptomyces* spp. TT, transcription terminator; MCS, multiple cloning site; p^{nitA}, the *nitA* gene promoter. The *nitA* gene encodes nitrilase; *nitR* encodes a positive transcriptional regulatory protein; the inducer is ε-caprolactam.



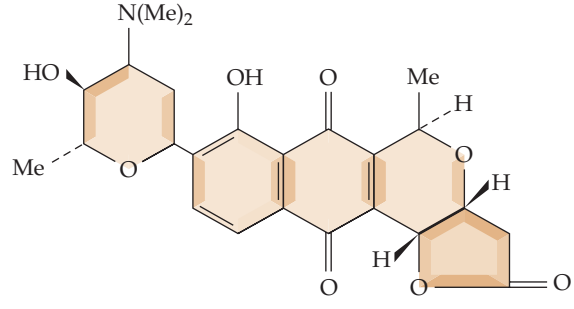
responsible for the biosynthesis of the antibiotic actinorhodine, starting from acetate. This antibiotic is a member of the family of antibiotics called isochromanequinones (Fig. 13.23). The intact plasmid and various subclones carrying portions of the 32.5-kb DNA fragment (e.g., pIJ2315) were introduced into either *Streptomyces* sp. strain AM-7161, which produces the related antibiotic medermycin, or *Streptomyces violaceoruber* B1140 or Tü22, both of which produce the related antibiotics granaticin and dihydrogranaticin.

Each of the antibiotics actinorhodine, medermycin, granaticin, and dihydrogranaticin (Fig. 13.23) functions as an acid–base indicator, conferring on a growing culture a characteristic color that depends on the pH of the medium (Table 13.4). The pH (and color), in turn, depends on the

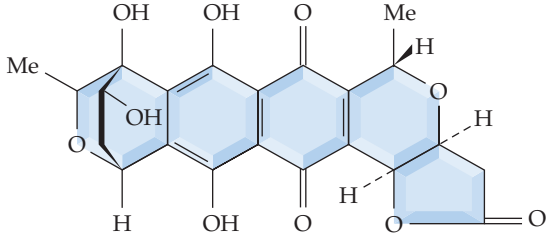
FIGURE 13.23 Structures of various isochromanequinone antibiotics produced by *Streptomyces* spp. Wild-type *S. coelicolor* and plasmid pIJ2303 encode actinorhodine, a *Streptomyces* sp. produces medermycin, and *S. violaceoruber* produces both granaticin and dihydrogranaticin. The hybrid antibiotics produced are mederrhodine A and dihydrogranatirhodine.



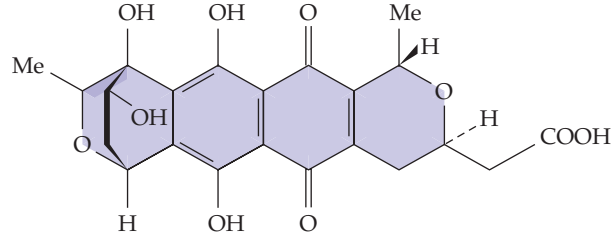
Actinorhodine



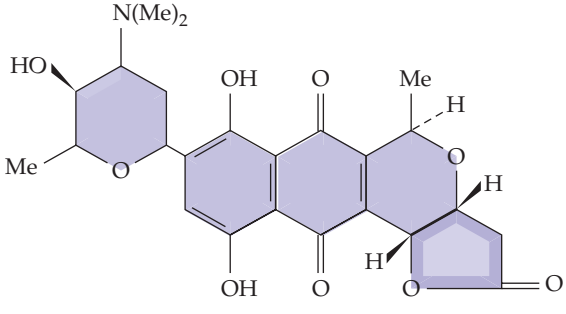
Medermycin



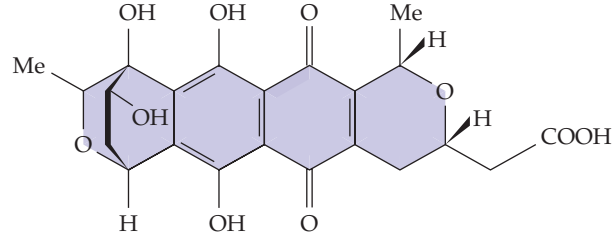
Granaticin



Dihydrogranaticin



Mederrhodine A



Dihydrogranatirhodine

TABLE 13.4 Antibiotics produced by various *Streptomyces* strains and those transformed with plasmids pIJ2303 and pIJ2315

Strain/plasmid	Color of culture		Antibiotic(s)
	Acidic	Alkaline	
S. coelicolor	Red	Blue	Actinorhodine
Streptomyces sp.	Yellow	Brown	Medermycin
Streptomyces sp./pIJ2303	Red	Blue	Medermycin, actinorhodine
Streptomyces sp./pIJ2315	Red	Purple	Mederrhodine A, medermycin
S. violaceoruber B1140	Red	Blue-purple	Granaticin, dihydrogranaticin
S. violaceoruber B1140/pIJ2303	Red	Blue-purple	Granaticin, dihydrogranaticin, actinorhodine
S. violaceoruber B1140 Tü22	Red	Blue-purple	Granaticin, dihydrogranaticin
S. violaceoruber B1140 Tü22/pIJ2303	Red	Blue-purple	Dihydrogranatirhodine, actinorhodine

Adapted from Hopwood et al., *Nature* **314:**642–644, 1985.

compound(s) being synthesized. Mutants of the *S. coelicolor* parental strain that are unable to produce actinorhodine are colorless. The appearance of new colors (in some cases) following the transformation of *Streptomyces* sp. strain AM-7161, or *S. violaceoruber* B1140, or *S. violaceoruber* Tü22 with a plasmid carrying either all or some of the genes encoding the enzymes that synthesize actinorhodine suggests that a novel antibiotic has been produced (Fig. 13.23 and Table 13.4). *Streptomyces* sp. strain AM-7161 and *S. violaceoruber* B1140 transformants containing pIJ2303 produce the antibiotics encoded by both the chromosomal and plasmid DNAs. However, when *S. violaceoruber* Tü22 is transformed with pIJ2303, a new antibiotic—dihydrogranatirhodine—is synthesized, along with actinorhodine. When *Streptomyces* sp. strain AM-7161 is transformed with pIJ2315, a second new antibiotic—mederrhodine A—is produced.

These new antibiotics represent minor structural variants of the preexisting antibiotics actinorhodine, medermycin, granaticin, and hydrogranaticin and probably arise when an intermediate compound from one biosynthetic pathway acts as a substrate for an enzyme from the other pathway. As the biochemistry of various antibiotic biosynthetic pathways has been better understood, it has become possible to design unique antibiotics by genetic manipulation of the genes encoding the relevant enzymes.

Engineering Polyketide Antibiotics

The term "polyketide" defines a class of antibiotics that are synthesized through the successive enzymatic condensation of small carboxylic acids, such as acetate, propionate, and butyrate. Polyketide drugs include the antibiotic erythromycin, the immunosuppressive drug FK506, and the cholesterol-lowering drug lovastatin. While various polyketides are produced by plants and fungi, the majority are produced by actinomycetes as secondary metabolites. To create new polyketide antibiotics, the functioning of the enzymes that synthesize these antibiotics must be understood before the genes encoding the enzymes can be manipulated.

Polyketide antibiotics are synthesized by a complex enzymatic mechanism analogous to that used for the synthesis of long-chain fatty acids. Each condensation cycle results in the formation, on a growing carbon chain, of a β-keto group. Polyketide synthesis consists of a number of steps that are each repeated several times, including ketoreduction, dehydration,

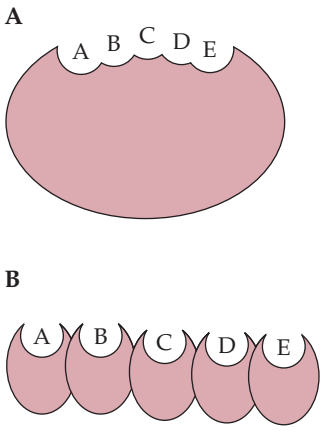
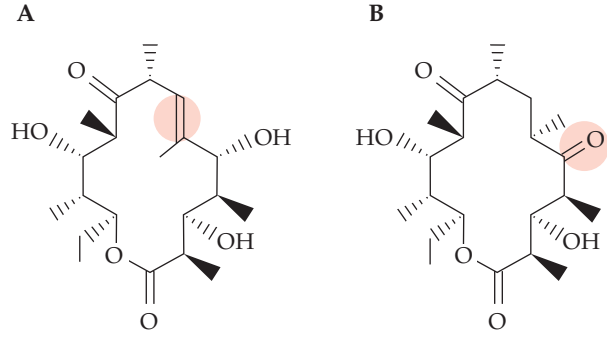


FIGURE 13.24 Schematic representation of polyketide synthase for aromatic polyketides. **(A)** The active site may be on a single polypeptide. **(B)** Alternatively, the enzyme can consist of assemblies of polypeptides with separate and distinct active sites. Both types of enzymes have different domains (regions A through E), and each domain has a separate enzymatic activity.

and enoylreduction of the β-group of the growing polyketide chain. There are two classes of polyketide synthases that are responsible for the synthesis of polyketide antibiotics (Fig. 13.24). The synthases that catalyze the biosynthesis of aromatic polyketides make up one class and generally consist of one polypeptide with an active site for each successive reaction (Fig. 13.24A). The second class includes synthases that are assemblies of several polypeptides that have separate and distinct active sites for every catalyzed step in polyketide biosynthesis (Fig. 13.24B). These enzymes have a number of different domains (regions A through E) (Fig. 13.24), and each domain has a separate enzymatic activity and active site catalyzing a particular step in the process. The complete synthesis of a polyketide antibiotic generally requires the participation of several of these multifunctional enzymes; together, they make up the subunits of the polyketide synthase.

If each of the enzymatic activities that is catalyzed by a domain on a multifunctional polyketide synthase subunit catalyzes only a single biochemical step in the pathway, the loss of any one activity should affect only a single step in the overall synthesis. Moreover, alteration of a catalytic domain whose function has been established should allow researchers to make predictable changes to the structure of the synthesized antibiotic. For example, a detailed knowledge of the genetics and biochemistry of the components involved in the synthesis of the antibiotic erythromycin allowed researchers to alter the biosynthetic genes in a predetermined manner and to produce predictably altered derivatives of erythromycin. Erythromycin is synthesized by *Saccharopolyspora erythraea*, and the entire 56-kb DNA segment that contains the *ery* gene cluster has been sequenced. The erythromycin polyketide synthase was altered in two different ways: either (1) a DNA region that encoded β-ketoreductase activity was deleted or (2) a DNA region encoding enoylreductase was mutated. With the β-ketoreductase deletion, the erythromycin intermediates that accumulated had a carbonyl moiety rather than a hydroxyl group at the C-5 carbon of the ring (Fig. 13.25). Similarly, with the enoylreductase mutation, a carbon–carbon double bond was introduced at positions C-6 and C-7 of the ring (Fig. 13.25). These experiments indicate that once the cluster of genes encoding the biosynthesis of a particular polyketide antibiotic has been isolated and characterized, it is possible, by altering specific DNA frag-

FIGURE 13.25 Altered erythromycin derivatives produced through genetic manipulation. **(A)** A mutation in an enoylreductase gene caused a carbon–carbon double bond to be introduced at positions C-6 and C-7 of the ring (highlighted). **(B)** A deletion in a β-ketoreductase gene caused the erythromycin to have a carbonyl moiety rather than a hydroxyl group at the C-5 carbon of the ring (highlighted). Adapted from Katz and Donadio, *Annu. Rev. Microbiol.* **47:**875–912, 1993.



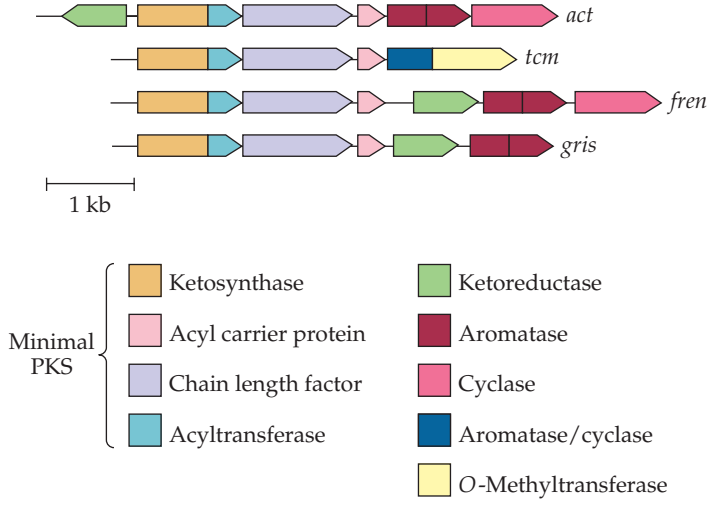
ments, to modify an antibiotic biosynthesis pathway and thereby to alter the structure of the antibiotic in a predictable manner. Moreover, it is possible to cut and splice DNA fragments and thereby shift polyketide synthase domains around to create novel polyketide antibiotics.

The gene clusters for aromatic polyketides all contain a set of three genes encoding a so-called minimal polyketide synthase. Each minimal polyketide synthase contains the activities for one cycle of polyketide chain elongation. The minimal module has a ketosynthase (with an acyltransferase domain), a chain length factor, and an acyl carrier protein. The minimal polyketide synthase is responsible for the synthesis of the aromatic polyketide backbone. Modifications to the basic structure are catalyzed by other enzymes acting in concert with the minimal polyketide synthase. The order of the modules in a polyketide synthase specifies the sequence of the distinct two-carbon units, and the number of modules determines the size of the polyketide chain. The genes encoding a complete set of these proteins are generally organized into a single cluster (Fig. 13.26). Each minimal polyketide synthase gene cluster encodes the synthesis of a particular antibiotic. By interchanging genes between clusters, new aromatic polyketide antibiotics have been created (Fig. 13.27). This experiment demonstrates the potential of using genetic manipulation to design and produce novel aromatic polyketide antibiotics. This approach promises to dramatically accelerate the process of discovery of new antibiotics.

Improving Antibiotic Production

In addition to being a means of developing new antibiotics, genetic engineering can be used to enhance the yields and rates of production of known antibiotics. The large-scale production of antibiotics by *Streptomyces* spp. is often limited by the amount of oxygen available to the cells. The low solubility of oxygen in aqueous media, combined with the highly dense nature

FIGURE 13.26 Gene clusters for the biosynthesis of the aromatic polyketide antibiotics actinorhodine *(act),* tetracenomycin *(tcm),* frenolicin *(fren),* and griseusin *(gris).* Each cluster contains genes encoding a minimal polyketide synthase (PKS), which is responsible for the synthesis of the polyketide backbone. The enzymes encoded by the other genes act to modify the growing polyketide chain. Each gene is shown pointed in the direction in which it is transcribed.



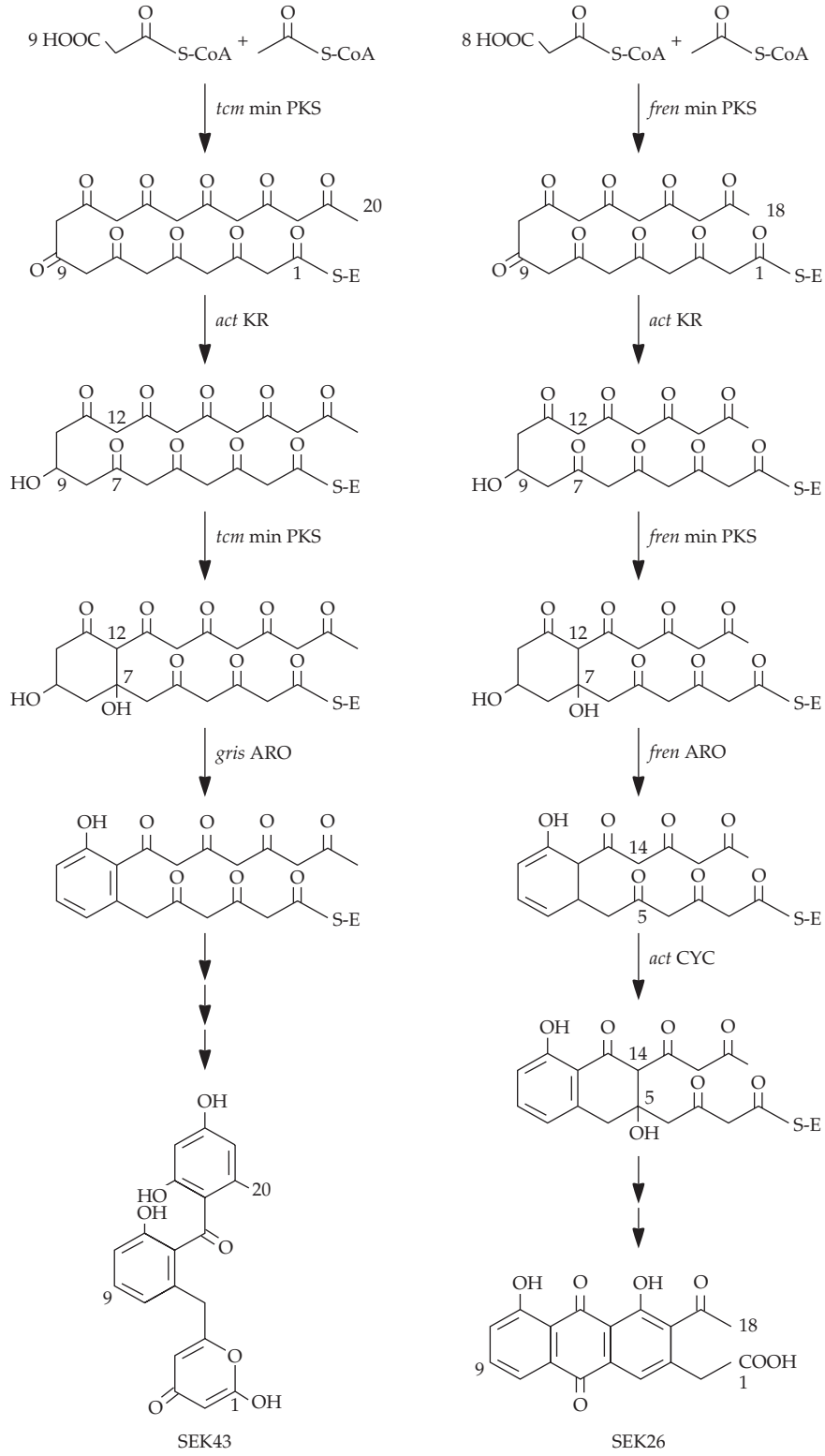


FIGURE 13.27 Theoretical biosynthetic pathway for the production of the rationally designed polyketides SEK43 and SEK26 by interchanging gene clusters. *act*, actinorhodine; *tcm*, tetracenomycin; *fren*, frenolicin; *gris*, griseusin; min PKS, minimal polyketide synthase; KR, β-ketoreductase; ARO, aromatase; CYC, cyclase.

of filamentous *Streptomyces* cultures, often results in an oxygen-depleted culture medium, a condition that causes poor cell growth and reduced antibiotic yield. To overcome this problem, it may be possible to improve the design of the bioreactors that are used to grow antibiotic-producing *Streptomyces* cultures and to develop, by genetic manipulation, *Streptomyces* strains that are better able to utilize the available oxygen. These two approaches are not mutually exclusive.

One strategy used by some aerobic microorganisms to cope with oxygen-poor environments is to synthesize a hemoglobin-like molecule that can sequester oxygen from the medium and then deliver it to the cells. For example, the aerobic bacterium *Vitreoscilla* produces a homodimeric heme protein that is functionally similar to eukaryotic hemoglobin. The gene for the *Vitreoscilla* hemoglobin was isolated and subcloned onto a *Streptomyces* plasmid vector. Following expression of the *Vitreoscilla* hemoglobin gene in *S. coelicolor,* the *Vitreoscilla* hemoglobin represented approximately 0.1% of the total cellular protein, even though the expression was controlled by the native *Vitreoscilla* hemoglobin gene promoter rather than by a *Streptomyces* promoter. When both transformed and nontransformed *S. coelicolor* cultures were grown in the presence of a low level of dissolved oxygen (i.e., approximately 5% saturation), the transformed cells with a functional *Vitreoscilla* hemoglobin produced 10 times more actinorhodine per gram (dry weight) of cells and had greater cell densities than did the nontransformed cells. The expression of the *Vitreoscilla* hemoglobin gene in oxygen-starved microbial cells may provide them with a general mechanism for obtaining sufficient oxygen to allow proliferation under otherwise limiting conditions.

The compound 7-aminocephalosporanic acid (7ACA) is synthesized chemically from the antibiotic cephalosporin C (Fig. 13.28) and is used as the starting material for the chemical synthesis of a number of cephem-type antibiotics (cephalosporins). Cephalosporins have few toxic effects on humans and protect against many different bacteria. Unfortunately, there is no known organism that can synthesize 7ACA. However, a novel 7ACA biosynthetic pathway has been constructed in the fungus *Acremonium chrysogenum,* which normally synthesizes only cephalosporin C. The genes involved in this novel engineered pathway consist of a cDNA that encodes D-amino acid oxidase and that comes from the fungus *Fusarium solani* and genomic DNA that encodes cephalosporin acylase and comes from the bacterium *Pseudomonas diminuta*. Both of these genes were subcloned separately onto an *A. chrysogenum* plasmid expression vector under the control of an *A. chrysogenum* promoter. In the first step of this new pathway, cephalosporin C is converted into the compound 7-β-(5-carboxy-5-oxopentanamido)cephalosporanic acid (keto-AD-7ACA) by D-amino acid oxidase (Fig. 13.28). Some of this product reacts with the hydrogen peroxide that is a by-product of the reaction to form 7-β-(4-carboxybutanamido) cephalosporanic acid (GL-7ACA) (Fig. 13.28). Cephalosporin C, keto-AD-7ACA, and GL-7ACA are each hydrolyzed by cephalosporin acylase to form 7ACA. However, in the absence of the D-amino acid oxidase step, only 5% of the cephalosporin C is converted to 7ACA; therefore, both enzymes are essential for high yields of 7ACA. Although the level of 7ACA that could be produced using this system was not sufficient to make this work the basis of a commercially viable process, it nevertheless demonstrates the feasibility of producing 7ACA biologically.

Medically important cephalosporins may be synthesized from either 7ACA or the related compound 7-aminodeacetoxycephalosporanic acid

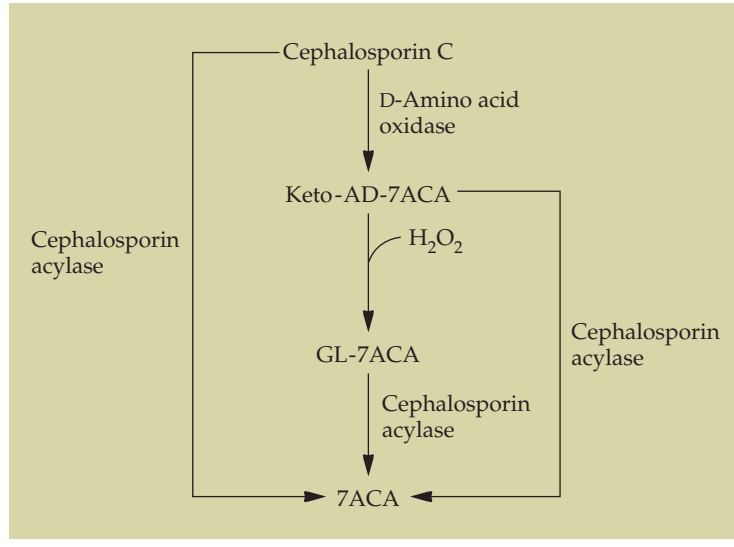


FIGURE 13.28 Genetically engineered biosynthetic pathway for the synthesis of 7ACA from cephalosporin C. The D-amino acid oxidase gene is from the fungus *F. solani,* and the cephalosporin acylase gene is from the bacterium *P. diminuta*.

(7ADCA). Disrupting the functioning of the *A. chrysogenum cefEF* gene results in the accumulation of large amounts of penicillin N (Fig. 13.29). Moreover, when this mutant strain is transformed with a *cefE* gene from *Streptomyces clavuligerus,* penicillin N is converted to deacetoxycephalosporin C (DAOC), which can then be converted to 7ADCA. At this stage of the development of this process, not all of the penicillin N is converted to DAOC; however, it may be possible to further increase the expression of the *cefE* gene. Thus, it is reasonable to expect that transgenic *A. chrysogenum* can eventually be engineered to produce large amounts of 7ADCA.

Many antibiotic-producing organisms are slow growing, require special growth conditions, or yield only small numbers of cells. To overcome these problems, *E. coli* was engineered to produce polyketide antibiotics at rates that are potentially useful for drug production. To do this, three genes (each 10 to 12 kb in length) encoding the components of the polyketide synthase from the bacterium *S. erythraea* were expressed in *E. coli*. Then, a *Bacillus subtilis* gene that produces an enzyme that attaches the cofactor phosophopantetheine to the polyketide synthase was cloned into the engineered *E. coli*. In addition, to supply the polyketide synthase with sufficient building blocks—propionyl-CoA and methylmalonyl-CoA—for polyketide synthesis, the *E. coli* gene encoding an enzyme that breaks down propionyl-CoA was inactivated and an *S. coelicolor* gene for propionyl-CoA carboxylase was introduced. Given the relative ease with which *E. coli* can be genetically manipulated and then grown in large-scale culture, this work may be a significant breakthrough for the development and production of new antibiotics.

Designer Antibiotics

In recent years, there has been an enormous proliferation in the prevalence of antibiotic-resistant bacterial infections. At present in the United States, more deaths are attributable to infections by methicillin-resistant strains of

Staphylococcus aureus than to human immunodeficiency virus/AIDS. An important component of *S. aureus* virulence is the carotenoid pigment staphyloxanthin that is synthesized by the bacterium. As a consequence of the large number of conjugated double bonds that this compound possesses, it can detoxify the reactive oxygen species that are produced by the host immune system in response to the bacterial infection. On the other hand, strains of *S. aureus* that do not contain staphyloxanthin are rapidly inactivated by the reactive oxygen species produced by host neutrophils. This observation suggests that the disruption of *S. aureus* staphyloxanthin synthesis might be a suitable target to prevent the proliferation and toxicity of the bacterium.

In the first committed step in the biosynthesis of *S. aureus* staphyloxanthin, the enzyme dehydrosqualene synthase condenses two molecules of farnesyl diphosphate to produce presqualene diphosphate (Fig. 13.30). In *S. aureus,* this compound is modified to yield dehydrosqualene, which is converted into 4,4′-diaponeurosporene and eventually into staphyloxanthin. Interestingly, the synthesis of cholesterol in humans also proceeds through presqualene diphosphate. In order to determine whether the *S. aureus* enzyme that catalyzed the first committed step could be a target for a designer antibiotic, the bacterial gene was cloned and overexpressed, the enzyme was purified to homogeneity, and its X-ray crystallographic structure was determined to 1.58-Å resolution. Based on the three-dimensional structure of the enzyme, as well as the presence of inhibitors of the human enzyme (from the cholesterol synthesis pathway) that performed the same function, a series of eight chemical inhibitors was designed, synthesized chemically, and tested. Three of the inhibitors, when tested at levels below 1 μM, dramatically blocked the conversion of farnesyl diphosphate to presqualene diphosphate (which is highly pigmented and therefore easy to visualize). In fact, one of the best inhibitors turned out to be a drug candidate that had already undergone preliminary testing in humans for its ability to lower cholesterol levels. Despite the fact that these results are preliminary, the compound that was selected using this strategy caused a 98% decrease in surviving *S. aureus* in infected mice. Thus, while much remains to be done, this work demonstrates that by choosing a nonconventional target it is possible to develop designer antibiotics that render pathogenic bacteria susceptible to human and animal immune systems.

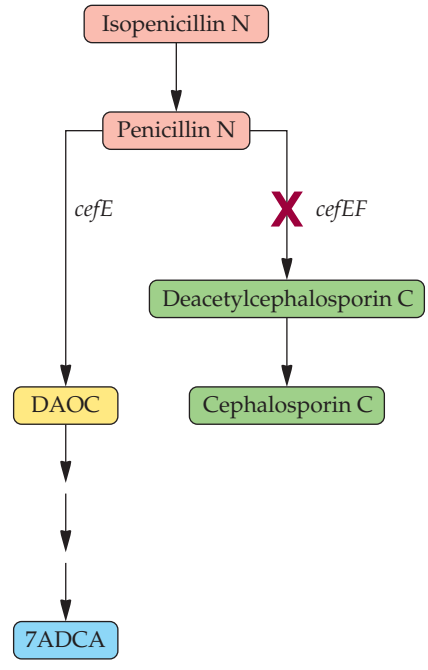


FIGURE 13.29 Synthesis of the compound 7ADCA in a genetically engineered strain of *A. chrysogenum*. The functioning of the endogenous *A. chrysogenum cefEF* gene is disrupted so that penicillin N accumulates. This strain is transformed with a *cefE* gene from *S. clavuligerus,* and the penicillin N is transformed into DAOC, which is subsequently converted to 7ADCA.

Biopolymers

Biopolymers are large, multiunit macromolecules synthesized by microorganisms, plants, and animals. Some of these polymers have physical and chemical properties that are useful to the food-processing, manufacturing, and pharmaceutical industries. The ability to genetically engineer organisms has stimulated researchers to design new biopolymers, replace synthetic polymers with biological equivalents, modify existing biopolymers to enhance their physical and structural characteristics, and find ways to increase the yields and decrease the costs of biopolymers produced by industrial processes.

Xanthan Gum

Xanthomonas campestris is a gram-negative obligatory aerobic soil bacterium that produces the commercially important biopolymer xanthan gum,

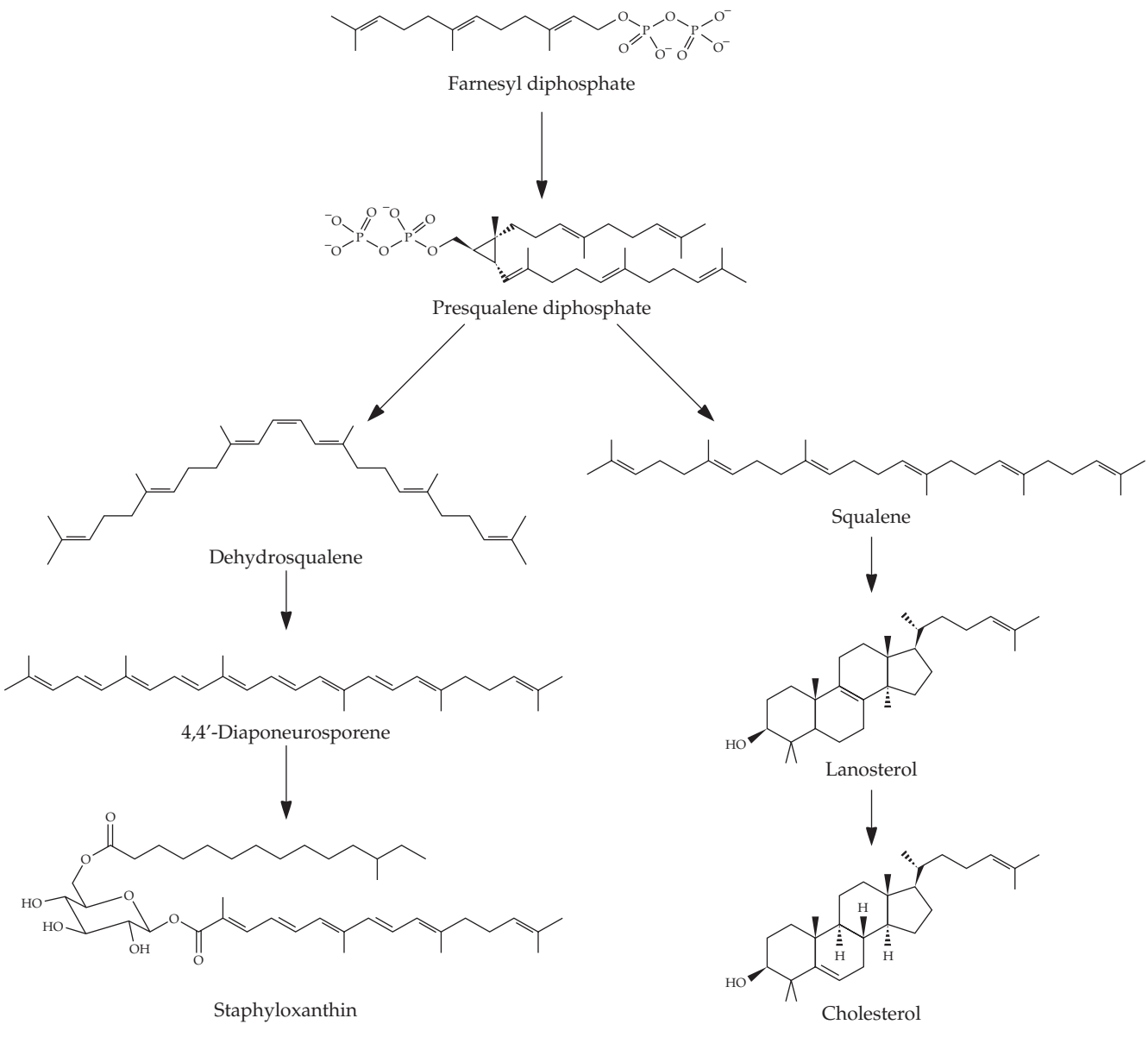


FIGURE 13.30 Simplified overview of the biosynthesis of staphyloxanthin in *S. aureus* and cholesterol in humans. The first step (the committed step) is the same in both pathways.

a high-molecular-weight exopolysaccharide, as a by-product of its metabolism. This polymer has a cellulosic backbone made up of a straight-chain polymer of glucose units. Each of its trisaccharide side chains includes one glucuronic acid and two mannose residues, which are attached to every second glucose residue of the backbone (Fig. 13.31). Xanthan gum has high viscosity, is stable in extreme physical and chemical environments, and exhibits physical and chemical properties similar to those of a plastic. In particular, its physical properties make it useful as a stabilizing, emulsifying, thickening, or suspending agent. For successful commercial production of xanthan gum, *X. campestris* should be grown on an inexpensive and

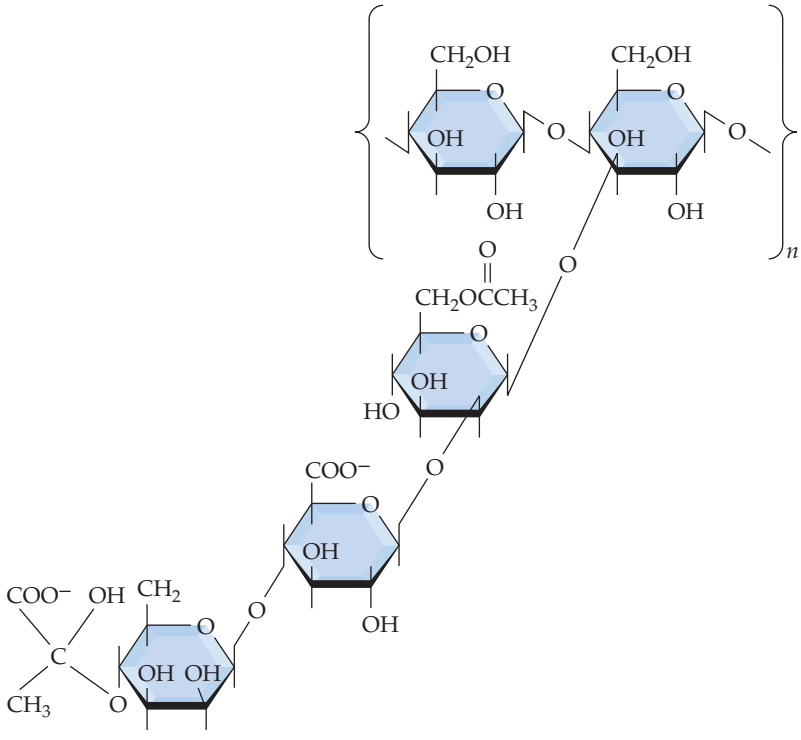


FIGURE 13.31 Structure of xanthan gum. The repeating unit, designated *n*, forms a chain of glucose units. The trisaccharide is attached to alternate glucose residues of the repeating chain.

plentiful carbon source. Wild-type *X. campestris* can efficiently utilize glucose, sucrose, and starch, but not lactose, as a carbon source.

Whey is a waste by-product of the cheese-making process that consists of water (94 to 95%), lactose (3.5 to 4%), and small amounts of protein, minerals, and low-molecular-weight organic compounds. Enormous quantities of whey are generated by the dairy industry, and its disposal is a major problem. In North America, whey has been used extensively as a "filler" in the preparation of prepared foods; however, with the increasing awareness that large numbers of individuals are lactose intolerant, it is imperative that alternative uses be found for this material. Moreover, disposing of whey by releasing it into rivers and lakes can deplete the amount of available oxygen, thereby killing many of the aquatic organisms. Transporting whey to landfill sites is exceptionally expensive, and potential groundwater contamination by the discarded whey is a major concern. Finally, the costs of removing the solid component of whey are prohibitive. Consequently, many schemes have been devised to use whey in creative ways.

Theoretically, whey could be used as a carbon source for growing industrially important microorganisms. With this in mind, *X. campestris* was genetically engineered to grow on whey. The *E. coli lacZY* genes, which encode the enzymes β-galactosidase and lactose permease, were cloned onto a broad-host-range plasmid under the transcriptional control of an *X. campestris* bacteriophage promoter (Fig. 13.32). This construct was introduced into *E. coli* and then transferred from *E. coli* to *X. campestris* by tripar-

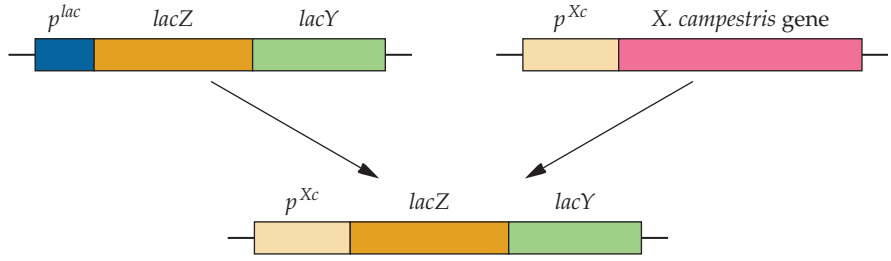


FIGURE 13.32 Engineering *E. coli lacZ* (encoding β-galactosidase) and *lacY* (encoding lactose permease) genes for constitutive expression in *X. campestris (Xc)*.

tite mating. Transformants maintained the plasmid, expressed the enzymes β-galactosidase and lactose permease at high levels, utilized lactose as the sole carbon source, and produced large amounts of xanthan gum with glucose, lactose, or whey as the carbon source (Table 13.5). By contrast, wild-type *X. campestris* produces large amounts of xanthan gum only when grown on glucose (Table 13.5). This system may well be able to convert a nuisance waste product into a substrate for the production of an economically valuable biopolymer.

Melanin

Melanins are a large, diverse family of light-absorbing biopolymers that are synthesized by animals, plants, bacteria, and fungi. It has been suggested that these pigments might be useful as topical sunscreens, sunlight-protective coatings for plastics, or additives for cosmetic products. Currently, melanins are obtained in small quantities either by extraction from natural sources or by chemical synthesis. However, recombinant DNA technology has made it possible to produce a range of melanins with different physical properties inexpensively and on a large scale.

Biochemically, melanins are irregular, somewhat random polymers that are composed of indoles, benzthiazoles, and amino acids. The first step in their synthesis, which is catalyzed by the copper-containing monooxygenase tyrosinase, is the oxidation of tyrosine to dihydroxyphenylalanine quinone. The final stages of the polymerization of melanin are nonenzymatic, and depending on the chemical nature of the nonquinone components that are incorporated into the polymeric structure (typically hydroxylated organic compounds), the end product can be black, brown, yellow, red, or violet.

The genes involved in melanin biosynthesis in the bacterium *Streptomyces antibioticus* have been isolated and analyzed. These genes were

TABLE 13.5 Production of xanthan gum by wild-type and transformed *X. campestris*

X. campestris	Amount of xanthan gum produced (μg/mL) with:		
	0.4% Glucose	0.4% Lactose	10% Whey
Wild type	3,530	245	224
Transformant	3,711	3,608	4,241

Adapted from Fu and Tseng, *Appl. Environ. Microbiol.* **56:**919–923, 1990.

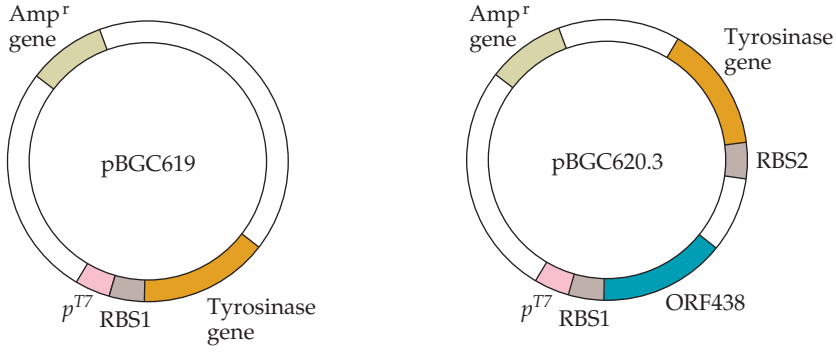
The amount of the product is expressed as micrograms per milliliter of culture grown on a minimal medium either 0.4% glucose or 0.4% lactose added or on diluted whey (10%), which contains approximately 0.44% lactose. The transformant carries the *E. coli lacZY* genes on a plasmid.

selected from a clone bank of *S. antibioticus* DNA on the basis of the ability to change color in the presence of specific compounds that were added to the medium. They consist of two open reading frames (ORFs), one encoding tyrosinase (molecular weight, 30,600) and one (ORF438) encoding a protein of unknown function with a molecular weight of approximately 14,800. To test whether both of these genes are required for melanin production, they were subcloned into an *E. coli* expression vector, where one construct contained only the tyrosinase gene and another carried both the tyrosinase and the ORF438 genes (Fig. 13.33). The vector with the tyrosinase gene directed the synthesis of a larger amount of tyrosinase than did the vector containing both the tyrosinase and the ORF438 genes. However, the amount of tyrosinase was irrelevant, because it turned out that melanin biosynthesis required the products of both genes. The protein encoded by ORF438 may act as a copper donor to apotyrosinase, the inactive precursor form of tyrosinase. Apotyrosinases are activated by acquiring copper ions. Under natural conditions, after dihydroxyphenylalanine quinone is produced by tyrosinase, a variety of low-molecular-weight compounds (nonquinones) can be incorporated into the final polymer. The chemical and physical nature, including the color, of the melanin that is formed after cloning of the key genes into *E. coli* may be manipulated to some extent, to form melanins with different properties, by the addition of different amounts of specific low-molecular-weight compounds to the medium.

Adhesive Protein

Researchers are trying to inexpensively produce an adhesive protein biopolymer, originally isolated from the blue mussel *Mytilus edulis*, in microbial cells. This biopolymer is an exceptionally strong, waterproof adhesive protein, called byssal adhesive, that enables the mussel to attach very tightly to a variety of surfaces. Following its secretion, the byssal adhesive becomes highly cross-linked (randomly), and consequently, the protein cannot be sequenced. Without this information, it is impossible to deduce nucleic acid sequences that might be used for the synthesis of DNA hybridization probes. However, it was possible to isolate an intracellular precursor form of the adhesive protein, called the 130-kilodalton (kDa) precursor protein,

FIGURE 13.33 *E. coli* expression plasmids carrying melanin biosynthesis genes. Plasmid pBGC619 contains the tyrosinase gene. Plasmid pBGC620.3 contains an ORF (ORF438) for melanin synthesis and the tyrosinase gene. Transcription of the cloned genes is under the control of the *E. coli* bacteriophage T7 promoter (p^{T7}). RBS1 and RBS2 denote two different ribosome-binding sites. The plasmids both carry genes that confer resistance to ampicillin (Amp^r).



MILESTONE

Production of 2-Keto-L-Gulonate, an Intermediate in L-Ascorbate Synthesis, by a Genetically Modified *Erwinia herbicola*

S. ANDERSON, C. B. MARKS, R. LAZARUS, J. MILLER, K. STAFFORD, J. SEYMOUR, D. LIGHT, W. RASTETTER, and D. ESTELL

Science **230:**144–149, 1985

Theoretically, prokaryotes that produce specific metabolites can be genetically manipulated in two different ways. First, the activity or amount of one or more of the enzymes in a pathway encoding the synthesis of a particular metabolite can be modified so that the amount of metabolite that the bacterium produces is increased. Second, foreign genes that produce enzymes that can use an endogenous metabolite as a substrate for the production of another metabolite not normally produced by the host bacterium can be introduced. While these sorts of manipulations are easy in theory, it is not necessarily easy to isolate and manipulate the required genes or to establish the appropriate conditions that enable complex biosynthetic pathways to function properly.

To create a bacterium that synthesized 2-keto-L-gulonic acid, which is the immediate precursor of commercially synthesized vitamin C, Anderson et al. isolated the *Corynebacterium* gene encoding the enzyme that converts 2,5-diketo-D-gluconic acid to 2-keto-L-gulonic acid and transferred it to an *Erwinia* sp., a bacterium that synthesizes 2,5-diketo-D-gluconic acid from D-glucose. The isolation of this gene was difficult because the enzyme had not been previously studied to any great extent. Therefore, before the gene could be isolated, the protein had to be purified and partially sequenced so that DNA hybridization probes based on the amino acid sequence of the protein could be designed. This work is an early example of what some workers have come to call metabolic engineering, which entails taking the genetic information for part of a metabolic pathway from one organism and transferring it into another organism to create a novel metabolic pathway.

that can be analyzed biochemically. It was found that the 130-kDa precursor protein is rich in serine, threonine, lysine, proline (Pro), and tyrosine; 60 to 70% of the amino acids contain a hydroxyl group. Most of the proline residues are hydroxylated to either 3- or 4-hydroxyproline (Hyp), and the majority of the tyrosines are hydroxylated to 3,4-dihydroxyphenylalanine (DOPA). Amino acid sequence analysis of the precursor protein further revealed that it is composed largely of repeating units that consist of a decapeptide with the sequence Ala-Lys-(Pro or Hyp)-Ser-(Tyr or DOPA)-Hyp-Hyp-Thr-DOPA-Lys; 7 of these 10 amino acids are hydroxylated.

The cDNA for the 130-kDa precursor adhesive protein was isolated from a cDNA library that was constructed with messenger RNA (mRNA) isolated from the gland that actively secretes the byssal adhesive. Both the adhesive protein and the cDNA have unusual features that might make cloning, expression, and production of a functional adhesive protein difficult in a heterologous host. First, the highly repetitive nature of the adhesive protein cDNA could make it unstable as a result of homologous recombination and subsequent loss of portions of the cloned sequence. Second, proline, lysine, and tyrosine represent about 70% of the amino acids of the protein; therefore, very high levels of synthesis may not be achievable because the corresponding intracellular aminoacyl-transfer RNA (tRNA) pools might be limiting.

When either complete or partial cDNAs for the adhesive protein were cloned onto yeast expression vectors and introduced into yeast cells, active novel forms of the adhesive protein, ranging from 20 to 100 kDa, were synthesized and represented a significant fraction (2 to 5%) of the total cell protein. Thus, there were no problems concerning either the stability of the cloned cDNA or the production of moderate amounts of the adhesive protein. Considerably higher expression levels were attained when a chemically synthesized adhesive protein gene sequence was expressed in *E. coli* (Fig. 13.34). In this case, repeating DNA units that encode the consensus

decapeptide repeat of the adhesive protein were used to construct a 600-base pair (bp) synthetic gene that encoded a protein with a molecular mass of approximately 25 kDa. The 30-bp repeat, the fundamental building block of the synthetic gene, consisted of codons optimized for *E. coli* expression. The synthetic gene was expressed at very high levels by using the T7 promoter. Notwithstanding the level of expression of the adhesive protein, most microorganisms are limited in their abilities to hydroxylate amino acids posttranslationally, so the final protein might be underhydroxylated. In fact, a number of tyrosine residues of the protein were not converted to DOPA, an outcome that limits the ability of the protein to form cross-links. This deficiency was overcome by creating an in vitro hydroxylation system that used a bacterial tyrosinase in the presence of ascorbic acid to hydroxylate the tyrosine residues to DOPA (Fig. 13.35). Ascorbic acid was included in the reaction mixture to prevent the premature oxidation of the DOPA residues to *o*-quinone. Oxidation must be controlled because it leads to cross-linking of the adhesive protein subunits. Like many other adhesives or glues, the protein adhesive must not be activated (cross-linked) before its actual use.

When the precursor form of the adhesive protein is oxidized, the cross-linked protein can bind to a variety of surfaces, including polystyrene, glass, hydrogel, and collagen. Moreover, the "strength" and specificity of the final adhesive can be manipulated by adding other proteins to the adhesive protein mixture before oxidation and cross-linking. By varying the kinds and amounts of the accessory proteins, adhesives with unique properties can be created. It is anticipated that biopolymeric adhesives will be used extensively in both medicine and dentistry.

More recently, researchers have isolated and expressed in *E. coli* the cDNA for the type 5 foot protein (adhesive protein) from the mussel *Mytilus galloprovincialis*. In this case, the protein that is produced in recombinant *E. coli* cells contains a tag of six histidine residues (to facilitate protein purification), as well as a slightly larger amount of DOPA than is found in the *M. edulis* protein. A higher number of DOPA residues is believed to result in a protein with greater adhesive properties. Unfortunately, the adhesive properties of this protein make it extremely difficult to purify (i.e., it sticks to everything), thereby limiting its commercial possibilities.

Rubber

Natural rubber, *cis*-1,4-polyisoprene, is an extensively used biopolymer that is obtained from a large number of different plants. The biosynthesis

FIGURE 13.34 Synthetic oligonucleotide used in the assembly of a gene for the bioadhesive protein produced by the mussel *M. edulis*. Two oligonucleotides are designed to base pair and form a DNA module with complementary extensions, which base pair to form a linear DNA molecule with a repeating sequence. The repeating units are joined by the enzyme T4 DNA ligase. The amino acid sequence encoded by the DNA repeat is shown. The decapeptide repeating unit contains three hydroxylated amino acids (Tyr, Ser, and Thr) plus three Pro residues which are subsequently hydroxylated.

5′ CCA ACC TAC AAA GCT AAG CCG TCT TAT CCG 3′
3′ TTT CGA TTC GGC AGA ATA GGC GGT TGG ATG 5′
Pro–Thr–Tyr–Lys–Ala–Lys–Pro–Ser–Tyr–Pro–Pro–Thr–Tyr

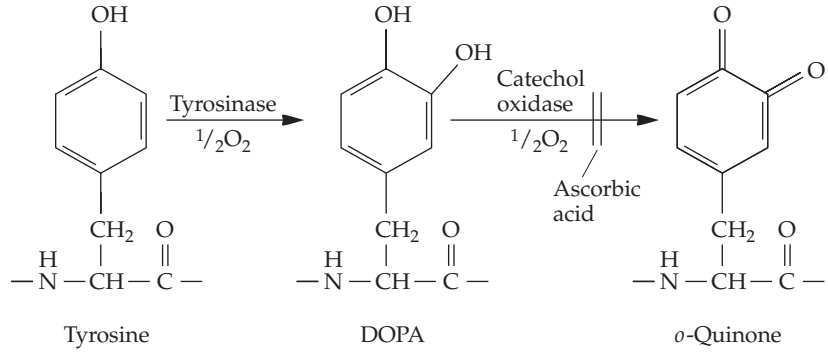


FIGURE 13.35 Pathway for in vitro posttranslational hydroxylation of some of the tyrosine residues in the *M. edulis* adhesive protein. Tyrosine is converted to DOPA by the action of the enzyme tyrosinase and then can be oxidized to *o*-quinone by either catechol oxidase or tyrosinase. The oxidation of DOPA to *o*-quinone can be prevented by the addition of ascorbic acid.

of rubber starts from simple sugars and requires approximately 17 enzyme-catalyzed steps, with the final step being the polymerization of isopentenyl pyrophosphate onto an allylic pyrophosphate. The last step is catalyzed by the enzyme rubber polymerase.

Studies have been undertaken to determine whether rubber can be synthesized by genetically engineered microorganisms. As an initial step in this direction, a cDNA library was constructed by using mRNA from the rubber-producing plant *Hevea brasiliensis*. This library was then screened with a short synthetic DNA hybridization probe whose sequence was based on the amino acid sequence of a portion of the rubber polymerase enzyme. Antibodies directed against the purified enzyme were used to prove unequivocally that the cloned cDNA expressed rubber polymerase. This cDNA clone can now be used, possibly in concert with other genes in the rubber synthesis pathway, in an attempt to produce natural rubber in a microbial system. Alternatively, it can be used as a source of rubber polymerase to develop an in vitro catalytic system. In either case, research that might lead to a new synthetic route for the production of rubber is under way.

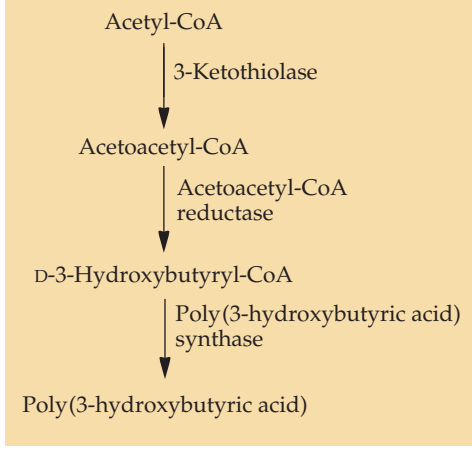
Polyhydroxyalkanoates

Polyhydroxyalkanoates are a class of biodegradable polymers that are produced by a number of different microorganisms, most notably *Alcaligenes eutrophus*, and used as an intracellular carbon and energy storage material. These compounds have thermoplastic or elastic properties, depending on the polymer composition, and are being considered for use in the synthesis of a range of biodegradable plastics. It has been estimated that by 2012 the U.S. market for biodegradable plastics will be around $1 billion per year.

Poly(3-hydroxybutyric acid) is the most thoroughly studied and characterized polyhydroxyalkanoate. Both the polymer and the *A. eutrophus* genes that encode its synthesis have been characterized. Poly(3-hydroxybutyric acid), its copolymer [poly(3-hydroxybutyrate-co-3-hydroxyvalerate)], and another polyhydroxyalkanoate [poly(3-hydroxyvaleric acid)] are produced commercially in the United Kingdom by the fermentation of *A. eutrophus*.

Although it is possible to produce poly(3-hydroxybutyric acid) as a by-product of the growth of *A. eutrophus,* the organism grows relatively slowly, requires a relatively low growth temperature (so that the fermentation vessel must be cooled), is difficult to lyse [making the purification of poly(3-hydroxybutyric acid) granules difficult], and utilizes only a limited number of carbon sources for growth (making production costs relatively high). On the other hand, when the genes for the biosynthesis of this polymer were transferred to *E. coli,* the resultant transformant grew rapidly to a high cell density and accumulated very large amounts (up to 95% of the dry cell weight) of poly(3-hydroxybutyric acid). Poly(3-hydroxybutyric acid) is synthesized from acetyl-CoA in three steps catalyzed by three enzymes (Fig. 13.36). The operon containing these genes was cloned into a plasmid as part of a 5.2-kb insert. Unfortunately, plasmids expressing the poly(3-hydroxybutyric acid) operon in *E. coli* were unstable. In the absence of selective pressure, such as the addition of antibiotics to the growth medium, about half of the *E. coli* cells lost the plasmid after approximately 50 generations. Plasmid loss of this magnitude, while not a major concern in small-scale batch cultures, becomes more of a problem with large-scale or continuous cultures (see chapter 17). This problem was overcome by inserting the *parB* genetic locus from another plasmid onto plasmids carrying the poly(3-hydroxybutyric acid) operon. This gene mediates plasmid stabilization by postsegregational killing of plasmid-free cells. The modified plasmids were quite stable even though the poly(3-hydroxybutyric acid) was produced constitutively, which places a metabolic load on the cells. An added benefit of producing poly(3-hydroxybutyric acid) in *E. coli* instead of *A. eutrophus* is that when the poly(3-hydroxybutyric acid) is recovered by extraction with an alkaline hypochlorite solution, the polymer is degraded to a much lesser extent than when it is produced in *A. eutrophus*. This is probably because most of the poly(3-hydroxybutyric acid) in *E. coli* is produced in a crystalline state, while in *A. eutrophus* it is amorphous. Nevertheless, the polymers extracted from the two organisms had identical polymer properties. In addition, *E. coli* transformants synthesizing poly(3-hydroxybutyric acid) produced very little acetate, which can be

FIGURE 13.36 Synthesis of poly(3-hydroxybutyric acid) from acetyl-CoA. The enzyme that catalyzes each of the reactions is shown to the right of the arrow.



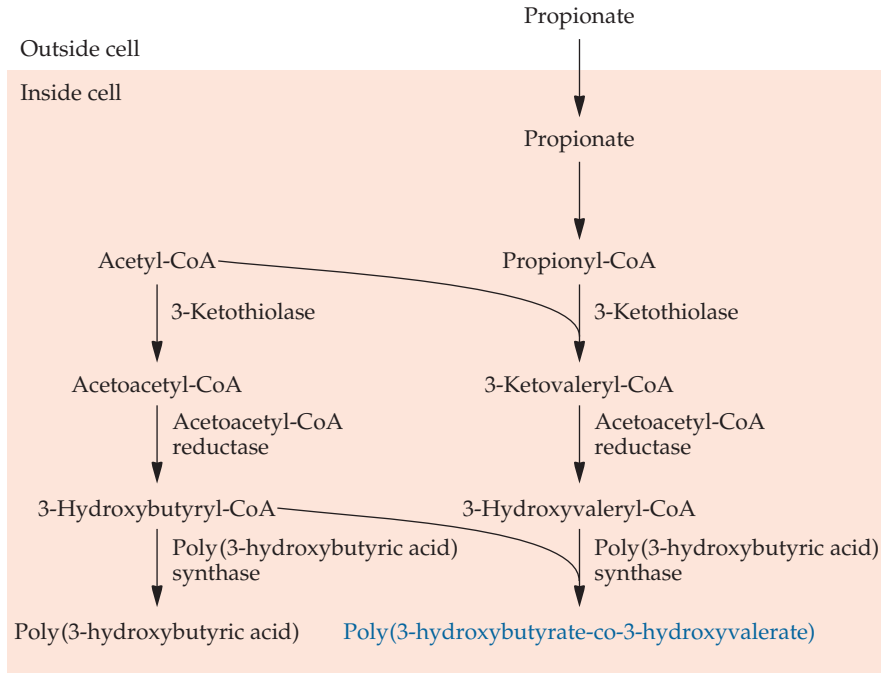


FIGURE 13.37 Model of the genetically engineered microbial synthesis of the copolymer poly(3-hydroxybutyrate-co-3-hydroxyvalerate).

deleterious for cell growth, presumably because all of the excess acetyl-CoA of the cell was converted to poly(3-hydroxybutyric acid) rather than acetate.

The copolymer poly(3-hydroxybutyrate-co-3-hydroxyvalerate) has properties that are similar to those of polypropylene. Consequently, there is considerable commercial interest in the biological production of poly-(3-hydroxybutyrate-co-3-hydroxyvalerate). However, *E. coli* strains that expressed the three polymer biosynthetic genes synthesized only poly-(3-hydroxybutyric acid) and not the copolymer. This limitation was overcome with *E. coli* cells that were mutated at both the *fadR* and *atoC* loci. The FadR protein is a negative regulator of fatty acid biosynthesis, and the *fadR* mutant activates the glyoxylate shunt, enhancing the capacity for energy metabolism and biosynthesis, which leads to a reduction of acetate excretion and improvement of the biomass yield. The *atoC* gene product is a positive regulator of fatty acid uptake, and the gene product from the *atoC* mutation turns on the synthesis of the proteins encoded by *atoA* and *atoD*, whose gene products facilitate the uptake of propionate from the growth medium into the cell. The propionate is converted to propionyl-CoA and then condensed with acetyl-CoA to form 3-ketovaleryl-CoA, which can be converted into 3-hydroxyvaleryl-CoA before its incorporation into the copolymer (Fig. 13.37). The amount of 3-hydroxyvalerate in the copolymer is dependent on the percentage of propionate used during the fermentation, but it never exceeds 40%.

In addition to engineering the composition of polyhydroxyalkanoate, to produce polymers with specific desired properties, it is also essential that the chain lengths of these polymers be regulated. Some polymerizing enzymes from bacteria such as *Ralstonia eutropha* yield primarily short-chain

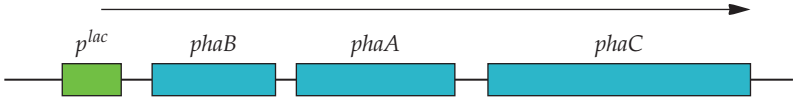
polymers with approximately 4 or 5 of the 3-hydroxyalkanoate monomeric units, while other enzymes from different bacteria, including *Pseudomonas oleovorans,* produce medium-chain polymers that include 6 to 14 monomeric units. In addition, mutants of the various steps of fatty acid oxidation can be generated and used to produce monomeric units with a modified composition compared with the wild-type strain. These modified monomers may then be incorporated into the polyhydroxyalkanoate. As a consequence of these manipulations, a wide range of polyhydroxyalkanoates with different physical and chemical properties have been synthesized.

It would be economically advantageous if bacteria could be engineered to efficiently produce polyhydroxyalkanoates using industrial waste products as a carbon source. To achieve this goal, polyhydroxybutyrate biosynthesis genes from a strain of an *Azotobacter* sp. were spliced onto a plasmid under the transcriptional control of the *lac* promoter (Fig. 13.38). The constructed plasmid was introduced into an *E. coli* strain that contained genes for the uptake and assimilation of lactose but that did not encode the *lac* repressor (see chapter 6). Thus, both lactose uptake and assimilation genes, as well as polyhydroxybutyrate biosynthesis genes, were expressed constitutively. The *E. coli* strain that carried the constructed plasmid was able to grow on either 25% lactose (a by-product of cheese making) or corn steep liquor (a by-product of corn [maize] processing) and to produce a significant level of polyhydroxybutyrate. By growing the transformed *E. coli* strain aerobically in a fed-batch culture (see chapter 17), after 24 hours, the cells accumulated polyhydroxybutyrate to 73% of their cell dry weight. Moreover, the physical properties of the polyhydroxybutyrate that was produced were similar to the properties of the polymer isolated from the *Azotobacter* sp. This engineered *E. coli* strain may be a suitable vehicle for producing a variety of polyhydroxyalkanoates from industrial waste products.

Hyaluronic Acid

Hyaluronic acid is a glycosaminoglycan, a polymer consisting of a repeating disaccharide unit of D-glucuronic acid and D-*N*-acetylglucosamine linked by β-1,4 and β-1,3 glycosidic bonds (Fig. 13.39), that in vivo can range in size from 5 to 20 kDa. This polymer is a component of the articular cartilage, where it is present as a coat around the cells; it is important in tissue hydrodynamics, movement, and cell proliferation; and it is used to treat osteoarthritis and to facilitate wound healing. Hyaluronic acid is also used as a component of some cosmetics and skin moisturizers. In 2005, the worldwide market for hyaluronic acid was a little over $1 billion, with most being supplied from rooster combs or the outer capsule of strains of group C *Streptococcus*. Both sources of hyaluronic acid can be problematic.

FIGURE 13.38 *Azotobacter* sp. genes encoding enzymes responsible for the biosynthesis of polyhydroxybutyrate under the transcriptional control of the *E. coli lac* promoter. The arrow indicates the direction of transcription. When a plasmid containing this construct is introduced into an *E. coli* strain that does not encode the *lac* repressor, the *pha* genes are expressed constitutively. *phaA,* 3-ketothiolase; *phaB,* acetoacetyl-CoA reductase; *phaC,* polyhydroxyalkanoate synthase.



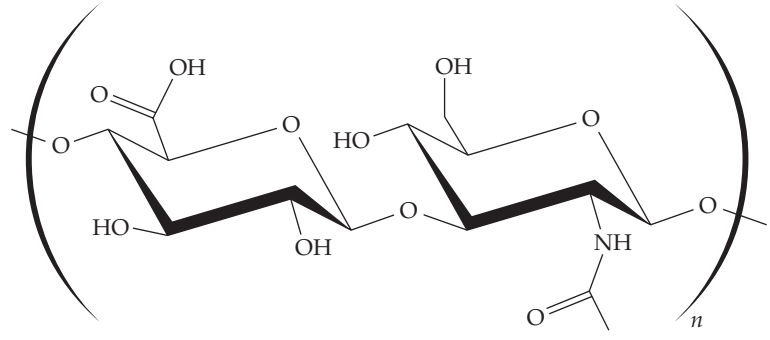
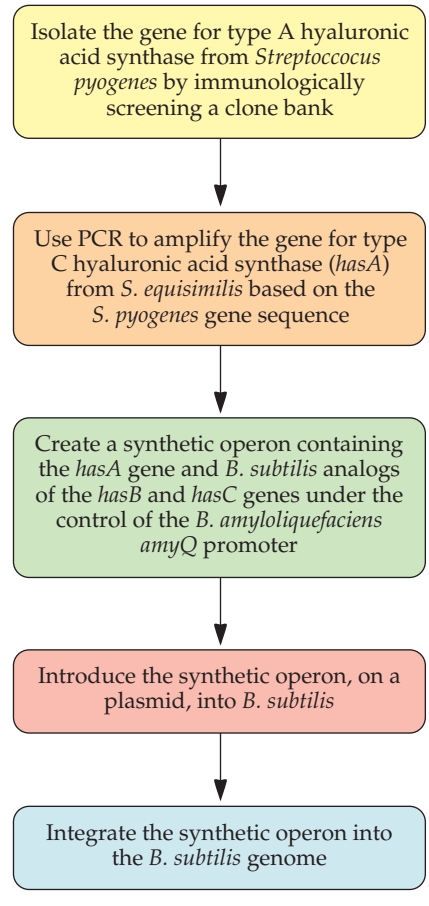


FIGURE 13.39 Structure of the repeating disaccharide unit of hyaluronic acid.

The rooster comb-based product can cause severe inflammation in individuals allergic to avian antigens, while the *Streptococcus*-based product is both difficult and expensive to produce. It would therefore be advantageous to have an alternative source of hyaluronic acid.

B. subtilis is a well-established industrial bacterium that can secrete large amounts of synthesized products while at the same time being very economical to grow on inexpensive medium on a large scale. In addition, *B. subtilis* does not produce any exo- or endotoxins or the enzyme

FIGURE 13.40 Flowchart of the engineering of *B. subtilis* to produce hyaluronic acid.



hyaluronidase (which degrades hyaluronic acid). The *Streptococcus equisimilis* gene encoding the last (and key) step in the synthesis of hyaluronic acid was isolated and then overexpressed in *B. subtilis* (Fig. 13.40), along with two *B. subtilis* genes that encode enzymes that provide the metabolites needed for the synthesis of hyaluronic acid. Following the large-scale growth in a bioreactor of this engineered *B. subtilis* strain, the amount of hyaluronic acid that was produced was comparable to the level produced by streptococcal strains (which grow more slowly), and the hyaluronic acid was secreted into the medium and not cell associated (as is the case with streptococcal strains), making it easier to isolate and purify. While this system may require some additional manipulation of the *B. subtilis* host strain to increase the yield of hyaluronic acid, this work is an important step toward the development of a commercial system for the bacterial production of hyaluronic acid.

SUMMARY

In addition to using bacteria as factories for the production of proteins, such as restriction enzymes, it is possible to modify the metabolic pathways of organisms, either by introducing new genes or by altering existing ones. In this way, various organisms can be genetically engineered for the production of a range of low-molecular-weight compounds, such as L-ascorbic acid, indigo, amino acids, antibiotics, lycopene, succinic acid, and the monomeric subunits of various biopolymers, such as xanthan gum, melanin, adhesive protein, rubber, polyhydroxyalkanoates, and hyaluronic acid. Here, the strategy is to insert the genes for one or more specific enzymes into the host organism by transformation with a vector-cloned gene construct. When they are expressed, the inserted genes encode a new pathway or augment a preexisting pathway for the synthesis of a specific compound. In addition, the biosynthesis of a desired compound may often be significantly increased by modulating the metabolic flux of the organism by turning on some pathways and blocking others. Several studies have shown that the creation of such unusual enzymatic pathways is technically feasible. Moreover, recombinant DNA technology has led to the development of new and more efficient synthetic routes for a variety of important compounds.

REFERENCES

Anderson, S., C. B. Marks, R. Lazarus, J. Miller, K. Stafford, J. Seymour, D. Light, W. Rastetter, and D. Estell. 1985. Production of 2-keto-L-gulonate, an intermediate in L-ascorbate synthesis, by a genetically modified *Erwinia herbicola*. *Science* **230:**144–149.

Bailey, J. E. 1991. Toward a science of metabolic engineering. *Science* **252:**1668–1675.

Banta, S., B. A. Swanson, S. Wu, A. Jarnagin, and S. Anderson. 2002. Alteration of the specificity of the cofactor-binding pocket of *Corynebacterium* 2,5-diketo-D-gluconic acid reductase A. *Protein Eng*. **15:**131–140.

Banta, S., B. A. Swanson, S. Wu, A. Jarnagin, and S. Anderson. 2002. Optimizing an artificial metabolic pathway: engineering the cofactor specificity of *Corynebacterium* 2,5-diketo-D-gluconic acid reductase for use in vitamin biosynthesis. *Biochemistry* **41:**6226–6236.

Berry, A. 1996. Improving production of aromatic compounds in *Escherichia coli* by metabolic engineering. *Trends Biotechnol.* **14:**250–256.

Breukink, E., I. Wiedemann, C. van Kraaij, O. P. Kuipers, H.-G. Sahl, and B. de Kruijff. 1999. Use of the cell wall precursor Lipid II by a pore-forming peptide antibiotic. *Science***286:**2361–2364.

Brooks, J. E., P. D. Nathan, D. Landry, L. A. Sznyter, P. Waite-Rees, C. L. Ives, L. S. Moran, B. E. Slatko, and J. E. Benner. 1991. Characterization of the cloned *Bam*HI restriction modification system: its nucleotide sequence, properties of the methylase, and expression in heterologous hosts. *Nucleic Acids Res.* **19:**841–850.

Cohen, G., D. Shiffman, M. Mevarech, and Y. Aharonowitz. 1990. Microbial isopenicillin N synthase genes: structure, function, diversity and evolution. *Trends Biotechnol.* **8:**105–111.

della-Cioppa, G., S. J. Garger, G. G. Sverlow, T. H. Turpen, and L. K. Grill. 1990. Melanin production in *Escherichia coli* from a cloned tyrosinase gene. *Bio/Technology* **8:**634–638.

Ensley, B. D., B. J. Ratzkin, T. D. Osslund, M. J. Simon, L. P. Wackett, and D. T. Gibson. 1983. Expression of naphthalene oxidation genes in *Escherichia coli* results in the biosynthesis of indigo. *Science* **222:**167–169.

Flores, N., J. Xiao, A. Berry, F. Bolivar, and F. Valle. 1996. Pathway engineering for the production of aromatic compounds in *Escherichia coli*. *Nat. Biotechnol.* **14:**620–623.

Floss, H. G. 1987. Hybrid antibiotics—the contribution of the new gene combinations. *Trends Biotechnol.* **5:**111–115.

Fu, J.-F., and Y.-H. Tseng. 1990. Construction of lactose-utilizing *Xanthomonas campestris* and production of xanthan gum from whey. *Appl. Environ. Microbiol.* **56:**919–923.

Gerritse, G., R. W. J. Hommes, and W. J. Quax. 1998. Development of a lipase fermentation process that uses a recombinant *Pseudomonas alcaligenes* strain. *Appl. Environ. Microbiol*. **64:**2644–2651.

Hahn, S. K., Y. K. Chang, and S. Y. Lee. 1995. Recovery and characterization of poly(3-hydroxybutyric acid) synthesized in *Alcaligenes eutrophus* and recombinant *Escherichia coli*. *Appl. Environ. Microbiol.* **61:**34–39.

Hazer, B., and A. Steinbüchel. 2007. Increased diversification of polyhydroxyalkanoates by modification reactions for industrial and medical applications. *Appl. Microbiol. Biotechnol.* **74:**1–12.

Herai, S., Y. Hashimoto, H. Higashibata, H. Maseda, H. Ikeda, S. Omura, and M. Kobayashi. 2004. Hyper-inducible expression system for streptomycetes. *Proc. Natl. Acad. Sci. USA* **101:**14031–14035.

Hillemann, D., A. Puhler, and W. Wolleben. 1991. Gene disruption and gene replacement in *Streptomyces* via single stranded DNA transformation of integration vectors. *Nucleic Acids Res.* **19:**727–731.

Hopwood, D. A., M. J. Bibb, C. J. Bruton, K. F. Chater, J. S. Feitelson, and J. A. Gil. 1983. Cloning *Streptomyces* genes for antibiotic production. *Trends Biotechnol.* **1:**42–48.

Hopwood, D. A., F. Malpartida, H. M. Kieser, H. Ikeda, J. Duncan, I. Fujii, B. A. M. Rudd, H. G. Floss, and S. Omura. 1985. Production of hybrid antibiotics by genetic engineering. *Nature* **314:**642–644.

Howard, K. A., C. Card, J. S. Benner, H. L. Callahan, R. Maunus, K. Silber, G. Wilson, and J. E. Brooks. 1986. Cloning the *Dde*I restriction-modification system using a two-step method. *Nucleic Acids Res.* **14:**7939–7951.

Hutchinson, C. R., H. Decker, K. Madduri, S. L. Otten, and L. Tang. 1993. Genetic control of polyketide biosynthesis in the genus *Streptomyces*. *Antonie Leeuwenhoek* **64:**165–176.

Hutchinson, C. R. 1994. Drug synthesis by genetically engineered microorganisms. *Bio/Technology* **12:**375–380.

Hutchinson, C. R., and I. Fujii. 1995. Polyketide synthase gene manipulation: a structure-function approach in engineering novel antibiotics. *Annu. Rev. Microbiol*. **49:**201–238.

Hwang, D. S., H. J. Yoo, J. H. Jun, W. K. Moon, and H. J. Cha. 2004. Expression of functional recombinant mussel adhesive protein Mgfp-5 in *Escherichia coli. Appl. Environ. Microbiol.* **70:**3352–3359.

Ikeda, M., K. Nakanishi, K. Kino, and R. Katsumata. 1994. Fermentative production of tryptophan by a stable recombinant strain of *Corynebacterium glutamicum* with a modified serine-biosynthetic pathway. *Biosci. Biotechnol. Biochem*. **58:**674–678.

Ishida, M., K. Miwa, S. Nakamori, and K. Sano. December 1989. Process for producing L-tryptophan. U.S. patent 4,885,245.

Isogai, T., M. Fukagawa, I. Aramori, M. Iwami, H. Kojo, T. Ono, Y. Ueda, M. Kohsaka, and H. Imanaka. 1991. Construction of a 7-aminocephalosporanic acid (7ACA) biosynthetic operon and direct production of 7ACA in *Acremonium chrysogenum*. *Bio/Technology* **9:**188–191.

Katz, L., and S. Donadio. 1993. Polyketide synthesis: prospects for hybrid antibiotics. *Annu. Rev. Microbiol*. **47:**875–912.

Kleinkauf, H., and H. von Dohren. 1990. Antibiotics—cloning of biosynthetic pathways. *FEBS Lett*. **268:**405–407.

Krämer, R. 1996. Genetic and physiological approaches for the production of amino acids. *J. Biotechnol.* **45:**1–21.

Lazarus, R. A., M. Hurle, S. Anderson, and D. B. Powers. December 1994. Enzymes for the production of 2-keto-L-gulonic acid. U.S. patent 5,376,544.

Lee, S. J., H. Song, and S. Y. Lee. 2006. Genome-based metabolic engineering of *Mannheimia succiniproducens* for succinic acid production. *Appl. Environ. Microbiol*. **72:**1939–1948.

Lee, S. Y., H. N. Chang, and Y. K. Chang. 1994. Production of poly(β-hydroxybutyric acid) by recombinant *Escherichia coli*. *Ann. N. Y. Acad. Sci.* **721:**43–53.

Lee, S. Y., K. S. Yim, H. N. Chang, and Y. K. Chang. 1994. Construction of plasmids, estimation of plasmid stability, and use of stable plasmids for the production of poly(β-hydroxybutyric acid) by recombinant *Escherichia coli*. *J. Biotechnol.* **32:**203–211.

Lee, S. Y., and H. N. Chang. 1995. Production of poly(3-hydroxybutyric acid) by recombinant *Escherichia coli* strains: genetic and fermentation studies. *Can. J. Microbiol.* **41:**207–219.

Liu, C.-I., G. Y. Liu, Y. Song, F. Yin, M. E. Hensler, W.-Y. Jeng, V. Nizet, A. H.-J. Wang, and E. Oldfield. 2008. A cholesterol biosynthesis inhibitor blocks *Staphylococcus aureus* virulence. *Science* **319:**1391–1394.

Magnolo, S. K., D. L. Leenutaphong, J. A. DeModena, J. E. Curtis, J. E. Bailey, J. L. Galazzo, and D. E. Hughes. 1991. Actinorhodin production by *Streptomyces coelicolor* and growth of *Streptomyces lividans* are improved by the expression of a bacterial hemoglobin. *Bio/Technology* **9:**473–476.

Martin, J. F. 1987. Cloning of genes involved in penicillin and cephalosporin biosynthesis. *Trends Biotechnol.* **5:**306–308.

McDaniel, R., S. Ebert-Khosla, D. A. Hopwood, and C. Khosla. 1995. Rational design of aromatic polyketide natural products by recombinant assembly of enzymatic subunits. *Nature* **375:**549–554.

Mermod, N., S. Harayama, and K. N. Timmis. 1986. New route to bacterial production of indigo. *Bio/Technology* **4:**321–324.

Nakamori, S., S.-I. Kobayashi, C. Kobayashi, and H. Takagi. 1998. Overproduction of L-cysteine and L-cystine by *Escherichia coli* strains with a genetically altered serine acetyltransferase. *Appl. Environ. Microbiol*. **64:**1607–1611.

Nikel, P. I., A. de Almeida, E. C. Melillo, M. A. Galvagno, and M. J. Pettinari. 2006. New recombinant *Escherichia coli* strain tailored for the

production of poly(3-hydroxybutyrate) from agroindustrial by-products. *Appl. Environ. Microbiol.* **72:**3949–3954.

Ozaki, A., R. Katsumata, and T. Oka. October 1989. Process for producing tryptophan. U.S. patent 4,874,698.

Park, J. H., K. H. Lee, T. Y. Kim, and S. Y. Lee. 2007. Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome analysis and *in silico* gene knockout simulation. *Proc. Natl. Acad. Sci. USA* **104:**7797–7802.

Pfeifer, B. A., S. J. Admiraal, H. Gramajo, D. E. Cane, and C. Khosla. 2001. Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science* **291:**1790–1792.

Piekarowicz, A., R. Yuan, and D. C. Stein. 1991. A new method for the rapid identification of genes encoding restriction and modification enzymes. *Nucleic Acids Res.* **19:**1831–1835.

Prieto, M. A., M. B. Kellerhals, G. B. Bozzato, D. Ragnovic, B. Witholt, and B. Kessler. 1999. Engineering of stable recombinant bacteria for production of chiral medium-chain-length poly-3-hydroxyalkanoates. *Appl. Environ. Microbiol*. **65:**3265–3271.

Ren, Q., N. Sierro, M. Kellerhals, B. Kessler, and B. Witholt. 2000. Properties of engineered poly-3-hydroxyalkanoates produced in recombinant *Escherichia coli* strains. *Appl. Environ. Microbiol*. **66:**1311–1320.

Rhie, H. G., and D. Dennis. 1995. Role of *fadR* and *atoC*(Con) mutations in poly(3-hydroxybutyrate-co-3-hydroxyvalerate) synthesis in recombinant *pha*$^+$ *Escherichia coli*. *Appl. Environ. Microbiol.* **61:**2487–2492.

Salerno, A. J., and I. Goldberg. 1993. Cloning, expression, and characterization of a synthetic analog to the bioadhesive precursor protein of the sea mussel *Mytilus edulis*. *Appl. Microbiol. Biotechnol.* **39:**221–226.

Schwarzer, A., and A. Puhler. 1991. Manipulation of *Corynebacterium glutamicum* by gene disruption and replacement. *Bio/Technology* **9:**84–87.

Sikora, L. A. January 1991. DNA fragment encoding a rubber polymerase and its use. U.S. patent 4,983,729.

Sosio, M., F. Giusino, C. Cappellano, E. Bossi, A. M. Puglia, and S. Donadio. 2000. Artificial chromosomes for antibiotic-producing actinomycetes. *Nat. Biotechnol*. **18:**343–345.

Stachelhaus, T., A. Schneider, and M. A. Marahiel. 1995. Rational design of peptide antibiotics by targeted replacement of bacterial and fungal domains. *Science* **269:**69–72.

Strausberg, R. L., and R. P. Link. 1990. Protein-based medical adhesives. *Trends Biotechnol.* **8:**53–57.

Takagi, H., N. Awano, S.-I. Kobayashi, M. Noji, K. Saito, and S. Nakamori. 1999. Overproduction of L-cysteine and L-cystine by expression of genes for feedback inhibition-sensitive serine acetyltransferase from *Arabidopsis thaliana* in *Escherichia coli*. *FEMS Microbiol. Lett*. **179:**453–459.

Terasawa, M., M. Fukushima, Y. Kurusu, and H. Yukawa. 1990. L-Tryptophan production by the application of high expressed tryptophanase in *Escherichia coli*. *Proc. Biochem. Int*. **25:**172–175.

Velasco, J., J. L. Adrio, M. A. Moreno, B. Díez, G. Soler, and J. L. Barredo. 2000. Environmentally safe production of 7-aminodeactoxycephalosporanic acid (7-ADCA) using recombinant strains of *Acremonium chrysogenum*. *Nat. Biotechnol*. **18:**857–861.

Walder, R. Y., J. L. Hartley, J. E. Donelson, and J. A. Walder. 1981. Cloning and expression of the *Pst*I restriction-modification system in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **78:**1503–1507.

Weber, J. M., J. O. Leung, S. J. Swanson, K. B. Idler, and J. B. McAlpine. 1991. An erythromycin derivative produced by targeted gene disruption in *Saccharopolyspora erythracea*. *Science* **252:**114–117.

Widner, B., R. Behr, S. Von Dollen, M. Tang, T. Heu, A. Sloma, D. Sternberg, P. L. DeAngelis, P. H. Weigel, and S. Brown. 2005. Hyaluronic acid production in *Bacillus subtilis*. *Appl. Environ. Microbiol.* **71:**3747–3752.

Yim, K. S., S. Y. Lee, and H. N. Chang. 1996. Synthesis of poly(3-hydroxybutyate-co-3-hydroxyvalerate) by recombinant *Escherichia coli*. *Biotechnol. Bioeng*. **49:**495–503.

Yoon, S.-H., J.-E. Kim, S.-H. Lee, H.-M. Park, M.-S. Choi, J.-Y. Kim, S.-H. Lee, Y.-C. Shin, J.-D. Keasling, and S.-W. Kim. 2007. Engineering the lycopene synthetic pathway in *E. coli* by comparison of the carotenoid genes of *Pantoea agglomerans* and *Pantoea ananatis*. *Appl. Microbiol. Biotechnol*. **74:**131–139.

REVIEW QUESTIONS

1. Describe a strategy for isolating the gene for the restriction endonuclease EcoRI.

2. Outline a strategy for cloning the gene for 2,5-diketo-D-gluconic acid reductase from *Corynebacterium* into *Erwinia*. Why is this useful?

3. Suggest a strategy for improving the commercial utility of a cloned 2,5-diketo-D-gluconic acid reductase gene.

4. How can indigo be produced in *E. coli?*

5. Outline a strategy for increasing the production of the amino acid tryptophan by *C. glutamicum*.

6. Suggest a strategy for isolating some of the genes that are involved in the biosynthesis of the antibiotic undecylprodigiosin, which is normally synthesized by *S. coelicolor*.

7. Why is it difficult to genetically transform various *Streptomyces* spp.? How can this difficulty be overcome?

8. Suggest a simple strategy for increasing the yield of an antibiotic by the genetic manipulation of the *Streptomyces* strain that produces the antibiotic.

9. Suggest an approach for producing modified versions of polyketide antibiotics, such as erythromycin.

10. How can an adhesive protein biopolymer that is normally produced by the blue mussel *M. edulis* be synthesized in *E. coli?*

11. Suggest a scheme for producing poly(3-hydroxybutyric acid) in *E. coli.*

12. What is whey? How can it be used to produce industrially important compounds?

13. How can *E. coli* be engineered to overproduce cysteine?

14. Suggest a scheme for the isolation of a lipase gene from the bacterium *P. alcaligenes*. How might this gene be used in a practical way?

15. How can the very large DNA fragments encoding antibiotic biosynthesis genes be introduced into host bacteria?

16. What strategies can be employed to produce large amounts of either 7ACA or 7ADCA in bacteria?

17. How can *E. coli* be engineered to produce lycopene?

18. How can the level of succinic acid produced by the bacterium *M. succiniproducens* be increased?

19. What is hyaluronic acid? How can it be produced in *B. subtilis?*

20. How can *E. coli* be genetically engineered to overproduce valine?

21. How can foreign proteins be expressed at high levels in *Streptomyces* spp.?

14

Bioremediation and Biomass Utilization

FOR CENTURIES, HUMANS BELIEVED that atmospheric, terrestrial, and aquatic systems were sufficient to absorb and break down wastes from population centers, industry, and farming. We now know that this is not true. Today, there are two fundamental problems. First, how do we dispose of the large quantities of wastes that are continually being produced? Second, how do we remove the toxic compounds that have been accumulating at dump sites, in the soil, and in water systems over the last few decades? Governments have tried to meet the challenge of environmental contamination by instituting antipollution regulations, but these rules often remain unenforced. Governments have also encouraged the three R's: reduce, reuse, and recycle.

Researchers are currently testing a number of technological strategies, including biotechnological schemes, to deal with large-scale wastes, such as lignocellulosics and toxic substances that persist in ecosystems.

The term "bioremediation" has been introduced to describe the process of using biological agents to remove toxic wastes from the environment. "Biomass" is the term used to describe the materials that are produced by the food and agricultural industries (e.g., starch and lignocellulosics) that were discarded as waste in the past. Biomass is now being considered as a source material for the production of a variety of economically important products.

Microbial Degradation of Xenobiotics

The problem of toxic waste disposal is enormous. Worldwide production in 1985 of just one chemical that is released into the environment—pentachlorophenol—was more than 50,000 tons. Incineration and chemical treatment have been used to break down many toxic chemicals, but these methods are costly and often create new environmental difficulties. With the discovery in the mid-1960s of a number of soil microorganisms that are capable of degrading xenobiotic ("unnatural," or synthetic; from the Greek *xenos,* meaning "foreign") chemicals, such as herbicides, pesticides, refrigerants, solvents, and other organic compounds, the notion that microbial

TABLE 14.1 Some *Pseudomonas* plasmids, their degradative pathways, and sizes

Name of plasmid	Compound(s) degraded	Plasmid size (kilobases)
SAL	Salicylate	60
SAL	Salicylate	68
SAL	Salicylate	72
SAL	Salicylate	83
TOL	Xylene and toluene	113
CAM	Camphor	225
XYL	Xylene	15
NAH	Naphthalene	69
OCT	Octane, D-camphor	~500
NAH7	Naphthalene, salicylate	83
PJP1	2,4-Dichlorophenoxyacetic acid	87
PJP2	2,4-Dichlorophenoxyacetic acid	54
PJB3	2,4-Dichlorophenoxyacetic acid	78
pP51	1,2-Di, 1,4-di-, and 1,2,4-trichlorobenzene	110
pAC31	3,5-Dichlorobenzoate	108
pAC25	3-Chlorobenzoate	102
pWW0	Xylene and toluene	117
pWW100	Biphenyl	200
pWWO	Xylene and toluene	176
pXYL-K	Xylene and toluene	135
pVI150	Phenol	>200
pNL1	Xylene, naphthalene, biphenyl	184
pAC27	3-Chlorobenzoate	110
pHMT112	Benzene	112
pTDN1	Aniline, *m*- and *p*-toluidine	79

Plasmids with the same name encode similar degradative pathways, even though they have different sizes and were described in different laboratories.

degradation might provide an economical and effective means of disposing of toxic chemical wastes gained credence.

Members of the genus *Pseudomonas* are the most predominant group of soil microorganisms that degrade xenobiotic compounds. Biochemical assays have shown that various *Pseudomonas* strains can break down and, as a consequence, detoxify hundreds of different organic compounds. In many cases, one strain can use any of several different related compounds as its sole carbon source.

The biodegradation of complex organic molecules generally requires the concerted efforts of several different enzymes. The genes that code for the enzymes of these biodegradative pathways are sometimes located in the chromosomal DNA, although they are more often found on large (approximately 50- to 200-kilobase) plasmids (Table 14.1). In some organisms, the genes that contribute to the degradative pathway are found in both chromosomal and plasmid DNA.

Degradative bacteria, in most cases, enzymatically convert xenobiotic, nonhalogenated aromatic compounds to either catechol (Fig. 14.1) or protocatechuate (Fig. 14.2). Then, through a series of oxidative cleavage reactions, catechol and protocatechuate are processed to yield either acetyl coenzyme A (acetyl-CoA) and succinate (Fig. 14.3) or pyruvate and acetaldehyde (Fig. 14.4), compounds that are readily metabolized by almost all

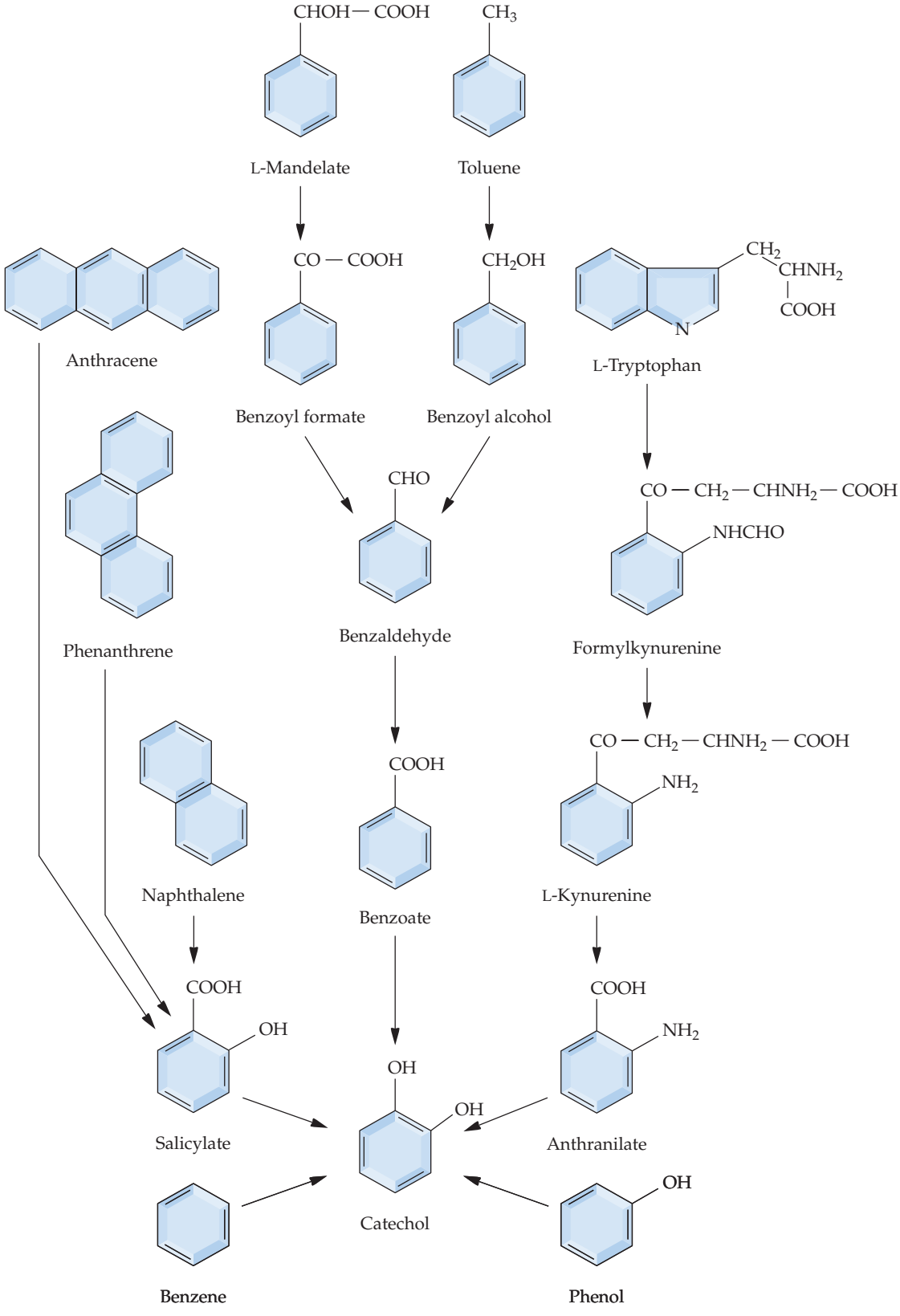


FIGURE 14.1 Pathways for the enzymatic conversion of aromatic compounds to catechol by degradative bacteria.

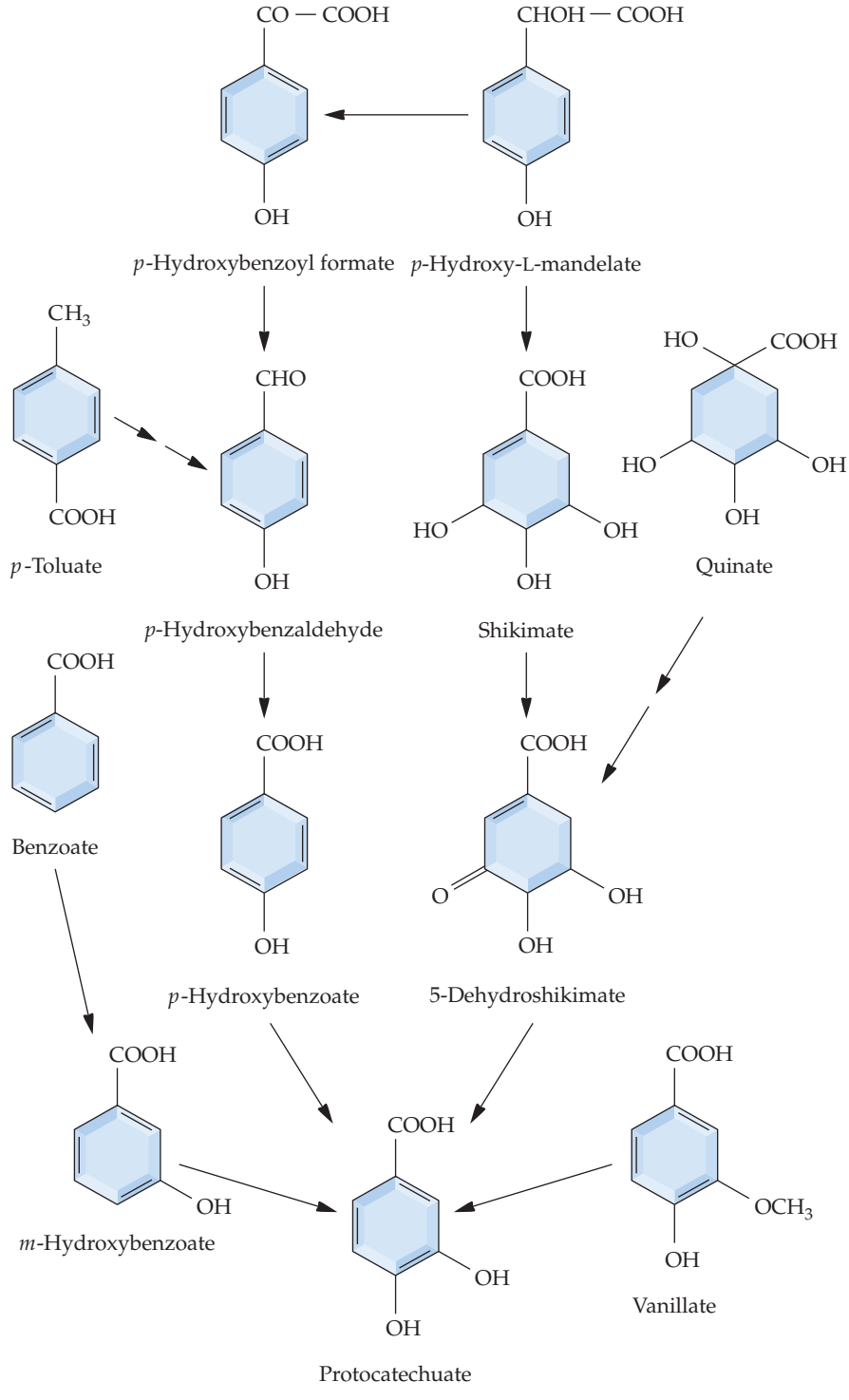


FIGURE 14.2 Pathways for the enzymatic conversion of aromatic compounds to protocatechuate by degradative bacteria.

organisms. Halogenated aromatic compounds, which are the main components of most pesticides and herbicides, are converted to catechol, protocatechuate, hydroquinones, or the corresponding halogenated derivatives

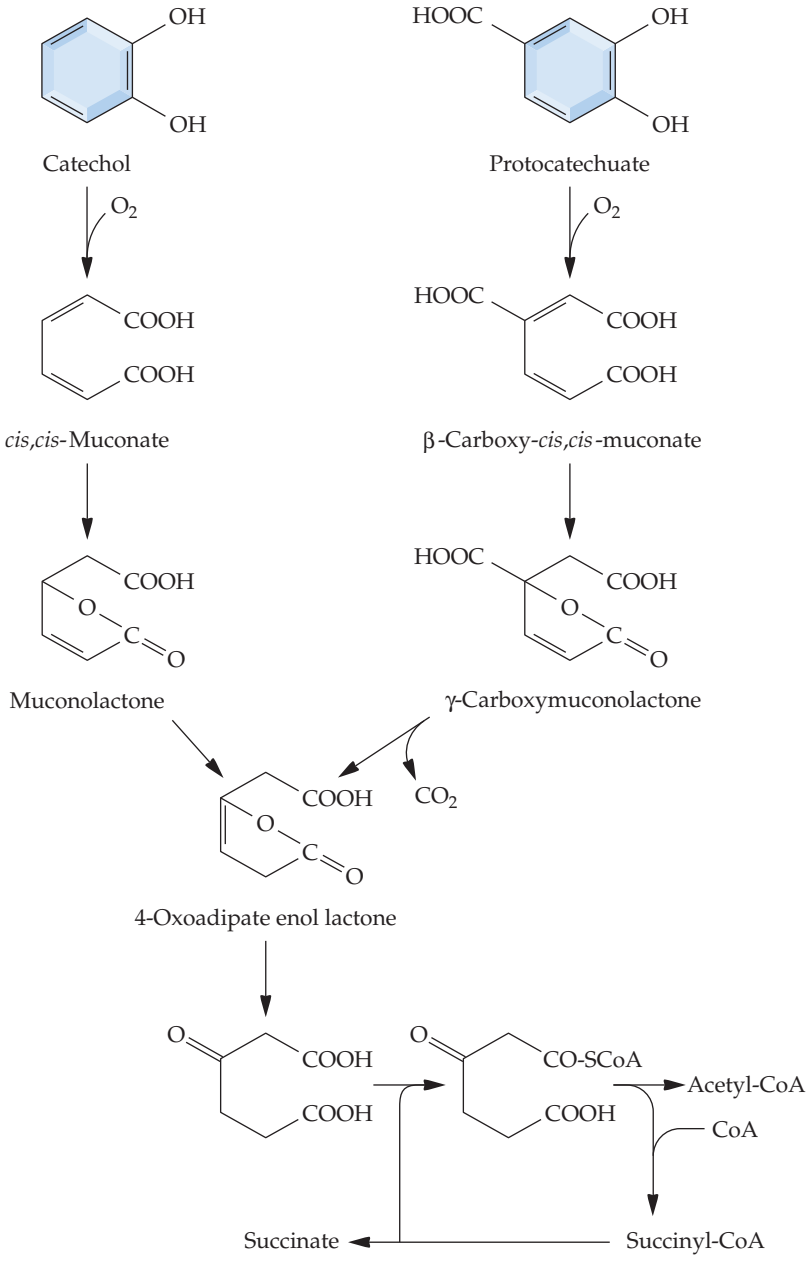


FIGURE 14.3 The *ortho*-cleavage pathway for enzymatic conversion of catechol and protocatechuate to acetyl-CoA and succinate.

by the same enzymes that degrade the nonhalogenated compounds. However, for the halogenated compounds, the rate of degradation is inversely related to the number of halogen atoms that are initially present on the target compound. Dehalogenation, the removal of a halogen substituent from an organic compound, is the critical requirement for detoxification and often occurs by a nonselective dioxygenase reaction that replaces the halogen on a benzene ring with a hydroxyl group. This step may occur either during or after the biodegradation of the original halogenated compound.

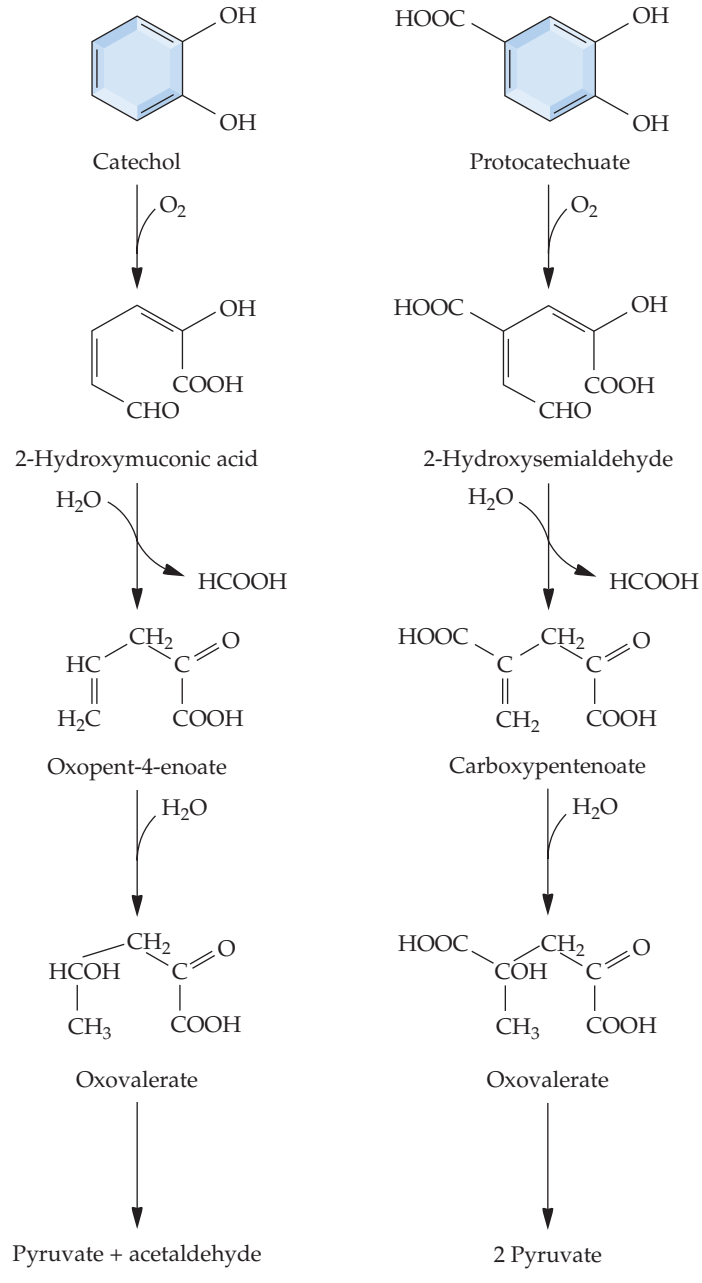


FIGURE 14.4 The *meta*-cleavage pathway for the enzymatic conversion of catechol and protocatechuate to pyruvate and acetaldehyde.

Genetic Engineering of Biodegradative Pathways

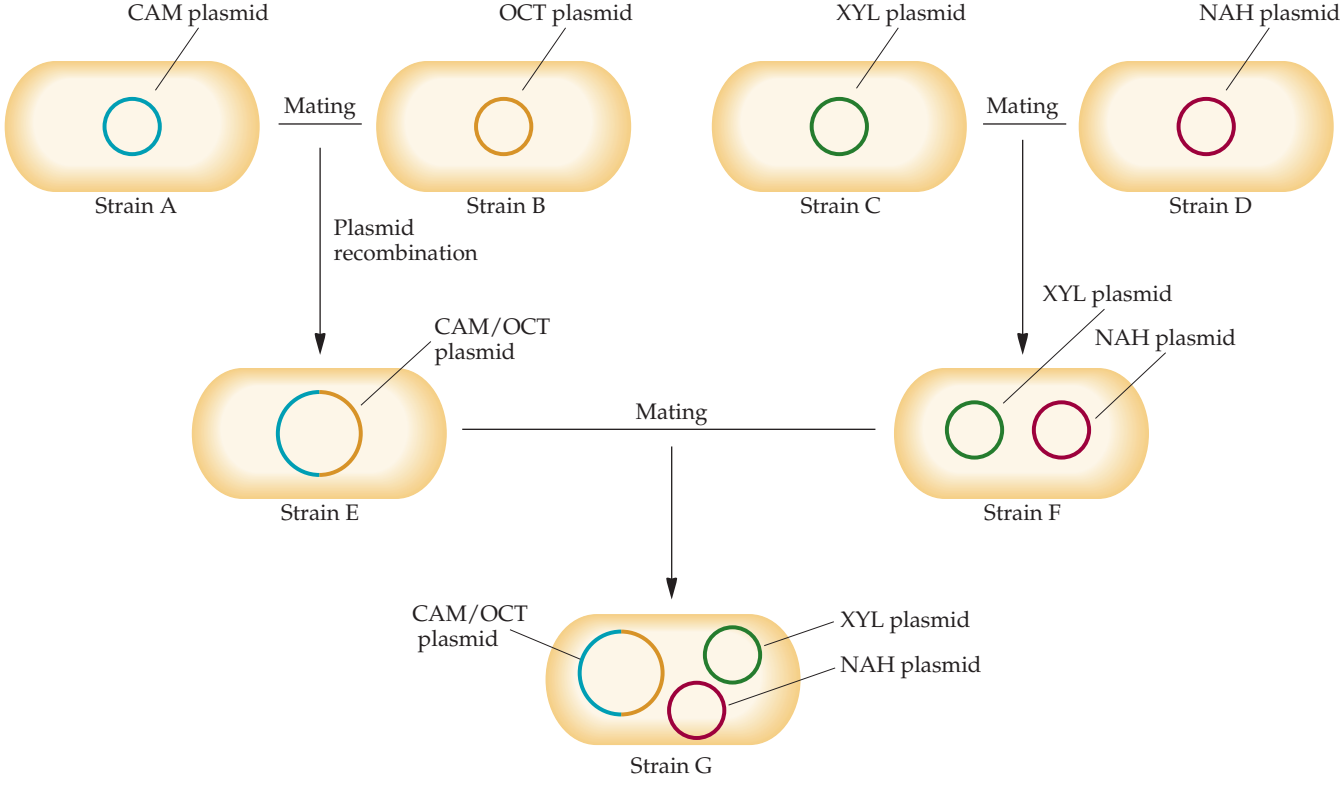
Despite the ability of many naturally occurring microorganisms to degrade a number of different xenobiotic chemicals, there are limitations to the biological treatment of these waste materials. For example, (1) no single microorganism can degrade all organic wastes; (2) high concentrations of some organic compounds can inhibit the activity or growth of degradative microorganisms; (3) most contaminated sites contain mixtures of chemicals, and an organism that can degrade one or more of the components of the mixture may be inhibited by other components; (4) many nonpolar

compounds adsorb onto particulate matter in soils or sediments and become less available to degradative microorganisms; and (5) microbial biodegradation of organic compounds is often quite slow. One way to address some of these problems is to transfer by conjugation into a recipient strain plasmids that carry genes for different degradative pathways (Fig. 14.5). If two resident plasmids contain homologous regions of DNA, recombination can occur and a single, larger "fusion" plasmid with combined functions can be created. Alternatively, if two plasmids do not contain homologous regions and, in addition, belong to different incompatibility groups, they can coexist within a single bacterium.

Manipulation by Transfer of Plasmids

Bacterial strains with expanded degradative capabilities were first created in the 1970s by Chakrabarty and coworkers. They used different plasmids to construct a bacterial strain that degraded a number of the hydrocarbon components found in petroleum (Fig. 14.5). This strain has been called a "superbug" because of its increased metabolic capabilities. The CAM

FIGURE 14.5 Schematic representation of the development of a bacterial strain that can degrade camphor, octane, xylene, and naphthalene. Strain A, which contains a CAM (camphor-degrading) plasmid, is mated with strain B, which carries an OCT (octane-degrading) plasmid. Following plasmid transfer and homologous recombination between the two plasmids, strain E carries a CAM and OCT biodegradative fusion plasmid. Strain C, which contains an XYL (xylene-degrading) plasmid, is mated with strain D, which contains an NAH (naphthalene-degrading) plasmid, to form strain F, which carries both of these plasmids. Finally, strains E and F are mated to yield strain G, which carries the CAM/OCT fusion plasmid, the XYL plasmid, and the NAH plasmid.



(camphor-degrading) plasmid was transferred by conjugation into a strain carrying the OCT (octane-degrading) plasmid. These two plasmids were incompatible and could not be maintained in the same cell as separate plasmids. However, when recombination occurred between the two plasmids, the resulting single plasmid was perpetuated and carried both camphor- and octane-degradative activities. The NAH (naphthalene-degrading) plasmid was transferred by conjugation into a strain carrying the XYL (xylene-degrading) plasmid. The NAH and XYL plasmids were compatible and could therefore coexist within the same host cell. Finally, the CAM/OCT fusion plasmid was transferred by conjugation into the strain carrying the NAH and XYL plasmids. The final result of these manipulations was the generation of a strain that grew better on crude oil than did any of the single-plasmid strains either alone or in combination.

Although this particular multiple-degradative strain has not been used to clean up oil spills, it has played a critical role in the development of the biotechnology industry. The inventor of this "superbug" was granted a U.S. patent describing its construction and use. This was the first patent ever granted for a genetically engineered microorganism and represented a watershed court decision, because it implied that biotechnology companies could protect their inventions in the same way as the chemical and pharmaceutical industries had in the past.

Most of the degradative bacteria that have been genetically manipulated by plasmid transfer are mesophiles, organisms that grow well only at temperatures between 20 and 40°C. However, rivers, lakes, and oceans that are polluted generally have temperatures that range from 0 to 20°C. To test whether bacteria with enhanced degradative abilities could be created for cold environments, a TOL (toluene-degrading) plasmid from a mesophilic *Pseudomonas putida* strain was transferred by conjugation into a facultative psychrophile, an organism with a low temperature optimum. The host psychrophile was able to degrade salicylate, but not toluene, and to use it as a sole carbon source at temperatures as low as 0°C. The transformed strain carried the introduced TOL plasmid and its own SAL (salicylate-degrading) plasmid and was able to use either salicylate or toluene as its sole carbon source at 0°C (Table 14.2). The wild-type (nontransformed)

TABLE 14.2 Generation times of wild-type (nontransformed) and transformed psychrophilic strains of *P. putida* on salicylate or toluate as the sole carbon source at various temperatures

Temperature (˚C)	Generation time (h) for:		
	Wild-type + salicylate	Transformant + salicylate	Transformant + toluate
37	No growth	No growth	No growth
30	2.2	2.5	2.0
25	2.1	3.2	1.3
20	2.6	3.8	1.9
15	3.2	4.2	2.9
10	6.3	5.6	3.3
5	13.9	12.9	12.2
0	28.6	18.1	24.4

Adapted from Kolenc et al., *Appl. Environ. Microbiol.* **54:**638–641, 1988.

The wild-type strain is unable to utilize toluate for growth at any temperature because it lacks the enzymes to metabolize the compound.

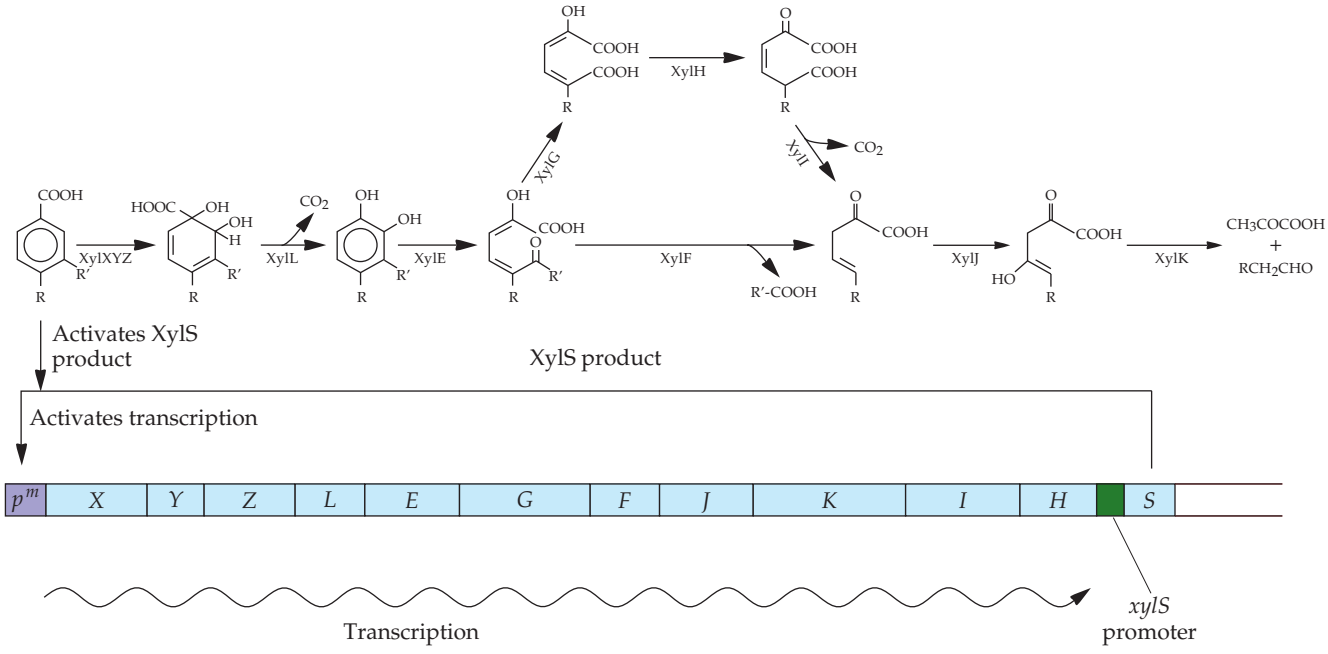


FIGURE 14.6 The *meta*-cleavage pathway and the *xyl* operon of the toluene- and xylene-degrading plasmid pWWO. Transcription of the *xyl* operon is controlled by the p^m promoter, which is regulated by the XylS gene product, which in turn must be activated by one of the initial pathway substrates. The genes from *xylX* to *xylH (X* to *H)* are under the control of the p^m promoter. The *xylS* gene, which is not part of this operon, is constitutively expressed. Some of the primary substrates are benzoate, where R and R′ = H; 3-methylbenzoate, where R = H and R′ = CH_3; 3-ethylbenzoate, where R = H and R′ = CH_2CH_3; and 4-methylbenzoate, where R = CH_3 and R′ = H. The *xylXYZ* genes encode toluene dioxygenase, *xylL* encodes dihydroxycyclohexadiene carboxylate dehydrogenase, *xylE* encodes catechol 2,3-dioxygenase, *xylF* encodes hydroxymuconic semialdehyde hydrolase, *xylG* encodes hydroxymuconic semialdehyde dehydrogenase, *xylH* encodes 4-oxalocrotonate tautomerase, *xylI* encodes 4-oxalocrotonate decarboxylase, *xylJ* encodes 2-oxopent-4-enoatehydratase, and *xylK* encodes 2-oxo-4-hydroxypentonate aldolase.

psychrophilic strain was unable to grow at any temperature when toluene (or toluate) was the only carbon source (not shown). This simple experiment indicates the feasibility of engineering psychrophilic degradative bacteria for use in the environment.

Manipulation by Gene Alteration

4-Ethylbenzoate. Bringing together different intact plasmid-based degradative pathways by conjugation is only one way to create bacteria with novel properties. It may also be possible to extend the degradative capability of a strain by altering the genes of an existing degradative pathway. The feasibility of this approach was examined for the toluene- and xylene-degrading pathway of plasmid pWWO. This plasmid encodes a "*meta*-cleavage" pathway involving 12 different genes and enables pseudomonads carrying the plasmid to utilize various alkylbenzoates as carbon sources (Fig. 14.6). The genes in the toluene–xylene pathway of pWWO are part of a single operon, called the *xyl* operon, under the control of the p^m promoter. Transcription from the p^m promoter, by RNA polymerase, is positively regulated by the *xylS* gene product, which is activated by most of the initial

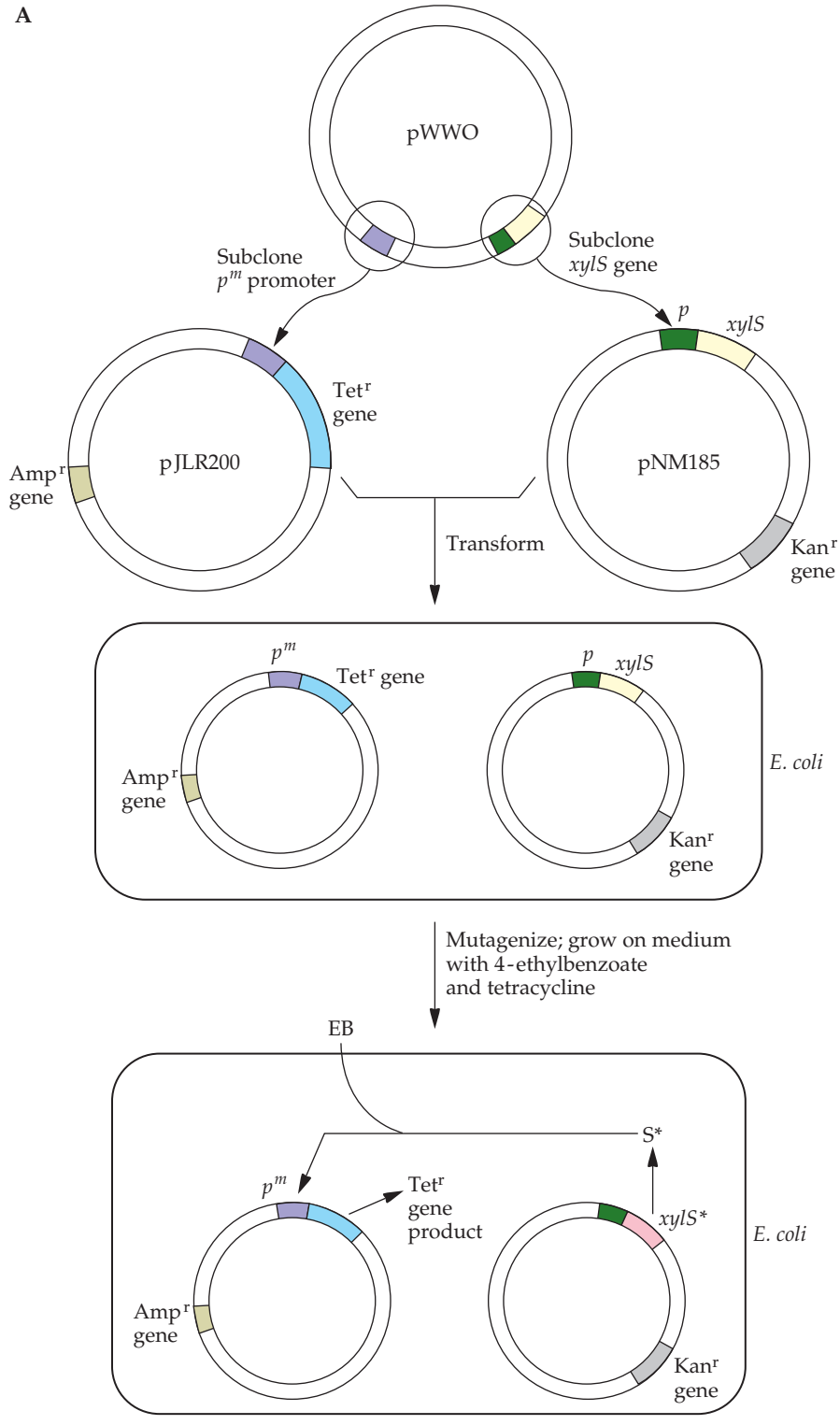


FIGURE 14.7 **(A)** Protocol used to create a XylS protein that is activated by 4-ethylbenzoate. The p^m promoter is cloned onto plasmid pBR322 and replaces the tetracycline resistance (Tet^r) gene promoter to form plasmid pJLR200; the *xylS* gene and its promoter are spliced onto a broad-host-range plasmid containing a kanamycin resistance (Kan^r) gene. *E. coli* is transformed with both of these plasmids. Transformants are selected by their resistance to both ampicillin (Amp^r) and kanamycin and then chemically mutagenized with ethyl methanesulfonate. Only cells with a mutation (S*) in the *xylS* gene that enables the XylS protein to be activated by 4-ethylbenzoate (EB) can grow on medium that contains both 4-ethylbenzoate and tetracycline, because only these cells are resistant to tetracycline.

substrates, such as benzoate and 3-methylbenzoate, of the pathway (Fig. 14.6). Detailed biochemical and genetic analyses showed that bacteria carrying pWWO could degrade 4-ethylbenzoate, albeit slowly, to 4-ethylcatechol, which accumulated in the medium, but no further. 4-Ethylcatechol

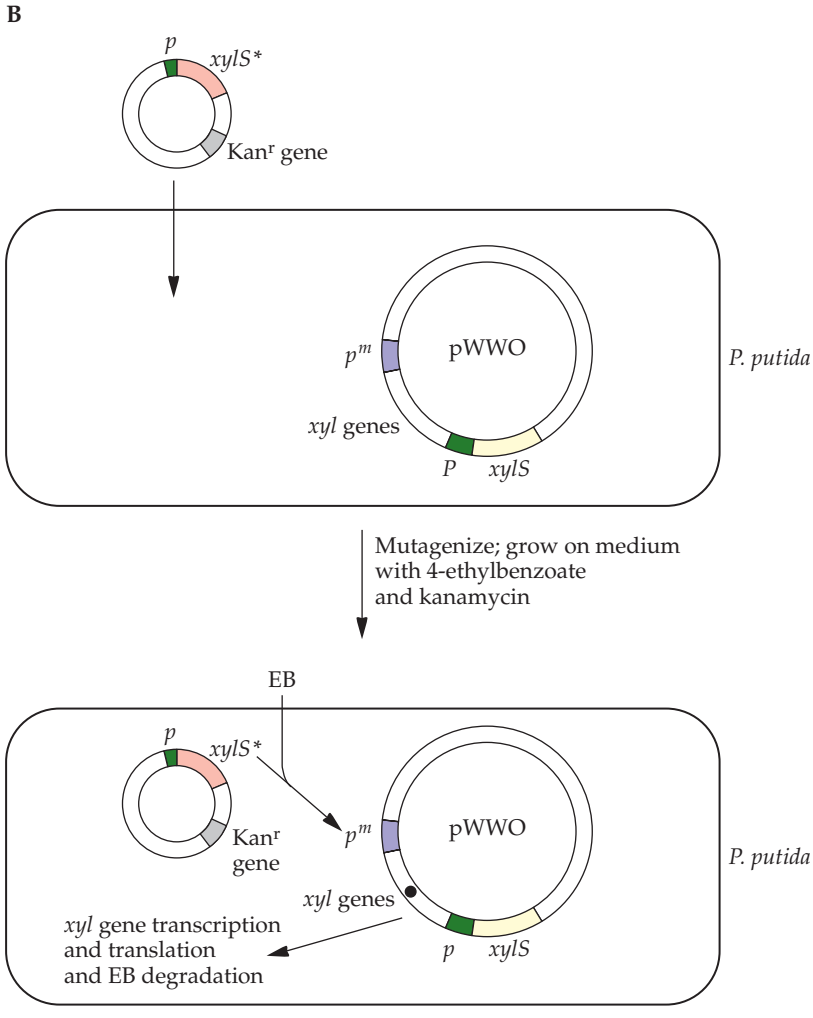


FIGURE 14.7 *(continued)* **(B)** Protocol used to create a modified catechol 2,3-dioxygenase that is not inhibited by 4-ethylcatechol. A *P. putida* strain carrying pWWO is transformed with a broad-host-range plasmid carrying the mutated *xylS** gene, whose product can activate the p^m promoter. Transformants are chemically mutagenized and then grown on minimal medium that contains 4-ethylbenzoate as the sole carbon source and kanamycin. Any cell that can grow on this medium carries a mutated catechol 2,3-dioxygenase gene. This mutation is indicated by a dot in the middle of the *xyl* gene cluster.

prevented its own degradation by inactivating one of the most important enzymes in the biodegradative pathway, catechol 2,3-dioxygenase, the product of the *xylE* gene. In addition, 4-ethylbenzoate, unlike most other alkylbenzoates, does not activate the XylS protein; consequently, transcription of the operon from the p^m promoter did not occur to any significant extent when 4-ethylbenzoate was the only substrate. Thus, there are two major problems with the naturally occurring *meta*-cleavage pathway system: (1) how to overcome the inactivation of an important enzyme in the degradative process by 4-ethylbenzoate and (2) how to induce transcription of the genes of this pathway with 4-ethylbenzoate as the inducer.

To find a mutant that could solve the second problem, a tetracycline resistance gene was placed under the control of the p^m promoter on one

plasmid, which also carried an ampicillin resistance gene. The *xylS* gene was cloned onto another plasmid carrying a kanamycin resistance gene. Transformants carrying both of these plasmids were selected on the basis of resistance to both ampicillin and kanamycin (Fig. 14.7A). The *Escherichia coli* cells carrying both of these plasmids were treated with the mutagen ethyl methanesulfonate and plated onto a medium containing both tetracycline and 4-ethylbenzoate. The only cells that could grow on this medium carried an altered XylS protein (S* in Fig. 14.7A) that could interact with 4-ethylbenzoate and cause the tetracycline resistance gene to be transcribed. Thus, the degradative pathway that includes this mutated *xylS* gene can be induced by 4-ethylbenzoate. To address the catechol 2,3-dioxygenase inactivation problem, the mutated *xylS* gene was subcloned onto a broad-host-range plasmid carrying a kanamycin resistance gene and introduced into *P. putida* cells carrying pWWO (Fig. 14.7B). The transformed cells were plated, at a high cell density, onto a minimal medium containing 4-ethylbenzoate as the sole carbon source, kanamycin to select for the presence of the plasmid, and ethyl methanesulfonate. Cells that were able to grow on this medium produced an altered form of the enzyme catechol 2,3-dioxygenase that was not inhibited by 4-ethylcatechol. Additional analysis confirmed that the catechol 2,3-dioxygenase gene on pWWO had been mutated and that mutant versions of both the *xylS* gene and the catechol 2,3-dioxygenase gene were required for the degradation of 4-ethylbenzoate.

An important aspect of this work is the fact that the two genes that were altered, i.e., those encoding XylS and catechol 2,3-dioxygenase, are the major determinants of the range of compounds that can be degraded by this pathway. The work with 4-ethylbenzoate demonstrates that by combining recombinant DNA technology, conventional mutagenesis, and the appropriate selection protocols, novel properties can be added to a degradative pathway.

Trichloroethylene. The compound trichloroethylene is widely used as a solvent and a degreasing agent, and as a result, it is one of the most common contaminants of soil and groundwater. Trichloroethylene persists in the environment for years, is a likely carcinogen, and is regulated in the United States under the Safe Water Drinking Act to a maximum contaminant level of 5 parts per billion. Unfortunately, anaerobic soil bacteria can reductively dehalogenate it to produce vinyl chloride, which is an even more toxic compound.

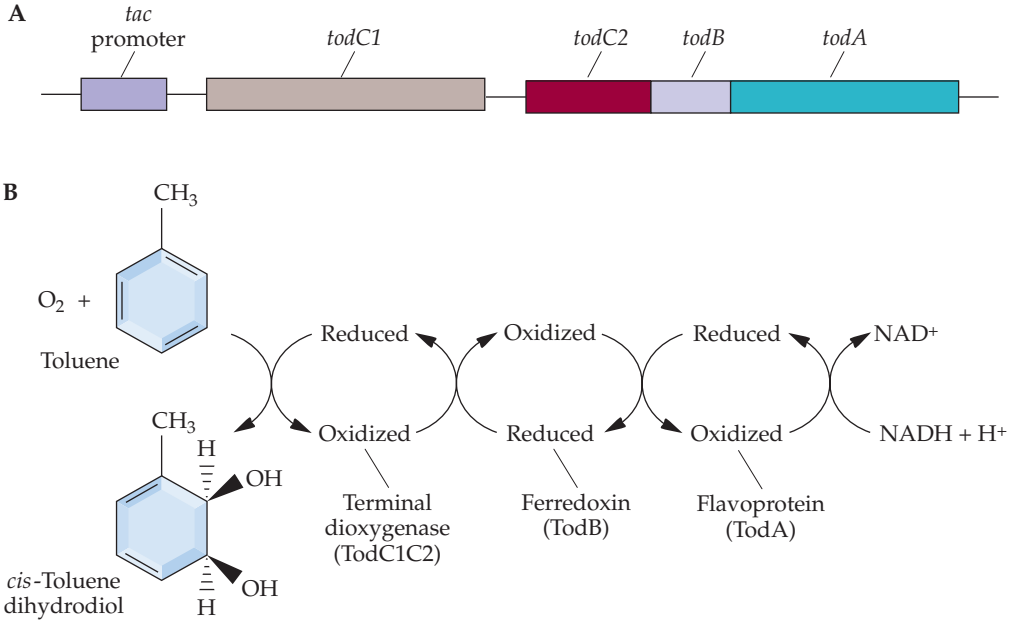
Studies showed that some of the strains of *P. putida* that could degrade aromatic compounds such as toluene could also degrade trichloroethylene. Genetic studies established that the complete *meta*-cleavage degradative pathway was not necessary to completely detoxify trichloroethylene. In fact, only the enzyme toluene dioxygenase, which normally catalyzes the oxidation of toluene to *cis*-toluene dihydrodiol, was required.

Four genes (Fig. 14.8A) are involved in the production of a functional toluene dioxygenase. These genes were isolated and expressed in *E. coli* under the control of the strong and inducible *tac* promoter. When *E. coli* cells carrying these genes were induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG), an inducer of the *tac* promoter, trichloroethylene was efficiently broken down to harmless compounds by the concerted enzymatic activities of the Tod proteins encoded by these four genes (Fig. 14.8B). Although the initial rates of trichloroethylene degradation were

lower with *E. coli* than with the original *P. putida* strain, the *E. coli* cells maintained these rates for longer periods than *P. putida*. It has been speculated that the basis for this difference may be that *E. coli* membranes are not as susceptible to damage from trichloroethylene as *P. putida* membranes are.

In a variation of this experiment, a hybrid *Pseudomonas* strain with elements of two separate degradative pathways was constructed. Bacterial strains that can degrade the compound biphenyl use the enzyme biphenyl dioxygenase. Biphenyl dioxygenase is a multicomponent enzyme encoded by four genes, *bphA1A2A3A4,* where *bphA1* encodes a large subunit of terminal dioxygenase (an iron-sulfur protein), *bphA2* encodes a small subunit of terminal dioxygenase, *bphA3* encodes ferredoxin, and *bphA4* encodes ferredoxin reductase (Fig. 14.9). BphA1 and BphA2 are associated as a heterotetramer and catalyze the introduction of two oxygen atoms into the biphenyl ring. Ferredoxin and ferredoxin reductase act as an electron transfer system from reduced nicotinamide adenine dinucleotide (NADH) to reduce the terminal dioxygenase. Biphenyl dioxygenase is quite similar in both structure and function to the enzyme toluene dioxygenase. Despite the similarities of their enzymes, biphenyl-utilizing pseudomonads cannot grow on toluene, and toluene-utilizing strains cannot grow on biphenyl. However, when the *bphA1* gene (coding for the large subunit of biphenyl dioxygenase) from *P. putida* KF715 was replaced by homologous recombination with the *todC1* gene (coding for the large subunit of the toluene dioxygenase) from *P. putida* F1, the resultant strain (Fig. 14.9) was able to degrade trichloroethylene (Table 14.3). In fact, the engineered strain grew well on a range of aromatic compounds and also was very efficient at

FIGURE 14.8 A cloned toluene dioxygenase operon under the control of the *tac* promoter in *E. coli*. **(A)** Toluene dioxygenase activity is due to the products of four genes (*todA, todB, todC1,* and *todC2*). *todA* encodes a flavoprotein that accepts electrons from NADH and transfers them to a ferredoxin encoded by *todB,* which reduces the terminal dioxygenase that is encoded by *todC1* and *todC2*. These genes are equivalent to the genes *xylXYZ* shown in Fig. 14.7. **(B)** Toluene is converted to *cis*-toluene dihydrodiol by the concerted enzymatic activities of the Tod proteins.



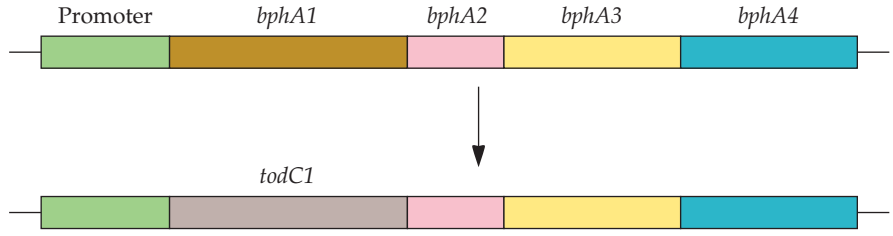


FIGURE 14.9 Creation of a hybrid *Pseudomonas* strain with elements of two separate degradative pathways by replacement, through homologous recombination, of the biphenyl dioxygenase *bphA1* gene with a toluene dioxygenase *todC1* gene.

degrading trichloroethylene, demonstrating that it is possible to rationally engineer bacterial strains that can degrade a number of different compounds. The creation of degradative bacteria with novel biological activities was achieved by slightly different means by the creation of chimeric versions of the *bphA1* gene (Fig. 14.10). In this case, one *bphA1* gene was from a strain of *Pseudomonas pseudoalcaligenes* with the ability to degrade only a narrow range of polychlorinated biphenyls (PCBs), while the other was from *Burkholderia cepacia,* which can degrade a very wide range of PCBs. Some of the hybrid genes encoded an enzyme with a wider degradative ability than either of the original enzymes. The native *bphA1* gene in *P. pseudoalcaligenes* was replaced with hybrid genes by homologous recombination. It now remains to be seen whether these engineered bacterial strains degrade a range of PCBs on a large scale.

Cell surface-expressed enzymes. Currently, detoxification of organophosphate pesticides in the environment is performed by chemical treatment, incineration, or burial in landfill sites. Each of these approaches has serious environmental drawbacks. It would therefore be advantageous if bacteria that are able to degrade these compounds could be utilized in place of the methods that are currently used. Several soil bacteria, including *Pseudomonas diminuta* MG and *Flavobacterium* spp., possess an enzyme, organophosphorus hydrolase, that catalyzes the hydrolysis of many of these pesticides, including methyl and ethyl parathion, paraoxon, chlorpyrifos (Dursban), coumaphos, cyanophos, and diazinon, to environmentally innocuous compounds. Unfortunately, these bacteria, as well as *E. coli* engineered to express organophosphorus hydrolase, degrade these pesticides relatively slowly because of the low rate of uptake into the bacterial cells. Thus, a novel approach was developed to solve this problem. *E. coli* cells were engineered to express organophosphorus hydrolase as part of a fusion pro-

TABLE 14.3 Growth of parental and engineered *Pseudomonas* strains on various aromatic compounds

Strain	Growth on:				
	Biphenyl	Diphenylmethane	Toluene	Benzene	Trichloroethylene
P. putida KF715	+++	+++	–	–	–
P. putida F1	–	–	+++	+++	+
P. putida KF715-D5	++	+	+++	+++	+++

Adapted from Suyama et al., *J. Bacteriol.* **178:**4039–4046, 1996.

In *P. putida* KF715-D5, the *bphA1* gene from *P. putida* KF715 is replaced with the *todC1* gene from *P. putida* F1. +++, good growth; ++ moderate growth; +, poor growth; –, very poor or no growth.

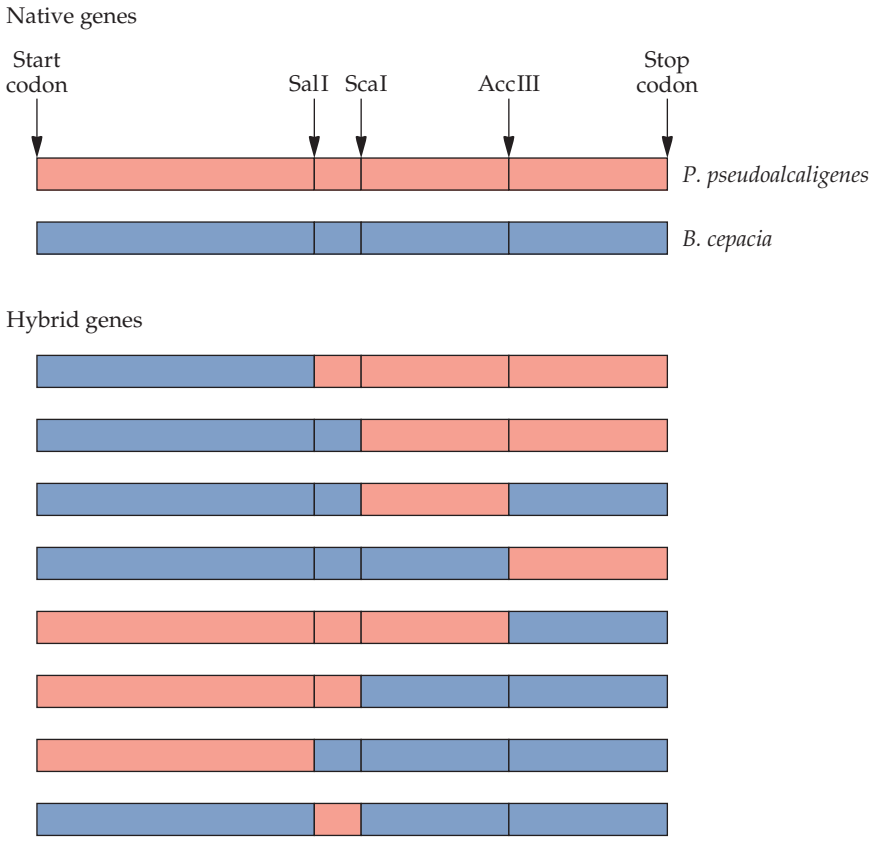


FIGURE 14.10 Chimeric (hybrid) *bphA1* genes constructed from *P. pseudoalcaligenes* and *B. cepacia bphA1* genes. The chimeras were constructed by digesting both of the isolated genes with the restriction enzymes shown and then recombining the resultant fragments as indicated. The various transformants carrying these constructs were subsequently assayed for enzymatic (degradative) activity.

tein that contained the *E. coli* lipoprotein signal peptide and the N-terminal portion of the lipoprotein and outer membrane protein A (Fig. 14.11). When this fusion protein was synthesized, it was localized on the outer surface of the bacterium (Fig. 14.12). This eliminated the problem of a low rate of pesticide uptake into the engineered bacterium. Cells with organophosphorus hydrolase on their outer surfaces had approximately seven times higher activity than cells that expressed a similar amount of the enzyme intracellularly. Moreover, the enzyme activity of cells with organophosphorus hydrolase on their surfaces was extremely stable. Nearly 100% of the activity was retained after 1 month at 37°C. This concept is quite promising; however, to date, it has been tested only on a laboratory scale.

Radioactive environments. The 26 countries worldwide that generate electricity from nuclear power plants, as well as those countries that have nuclear weapon programs, have generated thousands of radioactive waste sites. In the United States alone, there are more than 3,000 of these sites, and it has been estimated that, using currently available technology, the cleanup will require around $200 billion and take approximately 70 years. In addition to radioactivity, these sites also often have both organic and metal pollutants. While biodegradation of the organic pollutants is a logical first

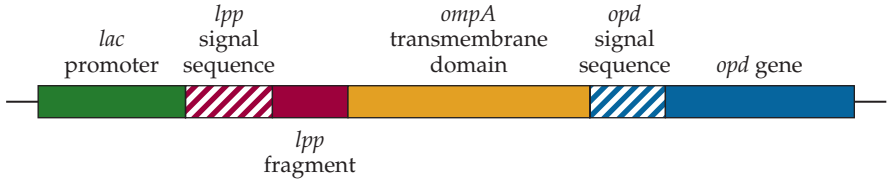
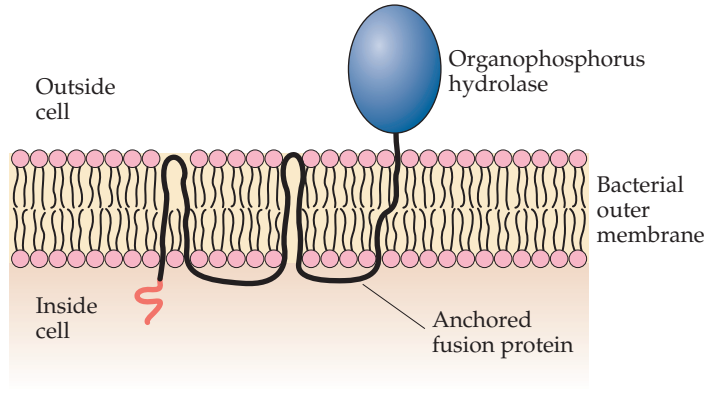


FIGURE 14.11 The DNA construct used to produce the fusion protein Lpp–OmpA–OPH. This construct includes the *E. coli lac* promoter, DNA encoding the *E. coli lpp* signal sequence and the first 9 amino acids of the mature *E. coli* lipoprotein, the portion of the *E. coli ompA* gene encoding the transmembrane domain, and the gene *(opd)* for the *Flavobacterium* sp. organophosphorus hydrolase (OPH) and its signal peptide.

step in their remediation, most microorganisms are highly sensitive to the damaging effects of the radiation (Fig. 14.13). Fortunately, the nonpathogenic soil bacterium *Deinococcus radiodurans* is naturally resistant to quite high levels of ionizing radiation. This resistance has been attributed to DNA repair processes that are exceptionally effective at repairing DNA damage. Moreover, any DNA that is introduced into *D. radiodurans,* either as part of a plasmid or inserted into the chromosome, is also protected against high levels of potentially damaging radiation. Since this bacterium can express foreign genes while growing in the presence of continuous radiation, it would appear to be an ideal candidate for the expression of bioremediating proteins in toxic environments that contain radioactive contaminants.

As a first step toward developing a system to remediate organic pollutants that are present in radioactive environments, the four genes that together code for toluene dioxygenase (Fig. 14.8) were placed on a plasmid under the control of a constitutive *D. radiodurans* promoter. The entire plasmid was then inserted into the chromosome of *D. radiodurans* by homologous recombination, a single crossover, between the chromosomal DNA and a chromosomal DNA fragment on the plasmid adjacent to the toluene dioxygenase genes. The integrated toluene dioxygenase was active

FIGURE 14.12 Schematic representation of the fusion protein Lpp–OmpA–OPH anchored in the *E. coli* outer membrane with organophosphorus hydrolase on the outside of the cell and therefore exposed to the external medium. Lpp (shown in red) includes the first 9 amino acids from the *E. coli* lipoprotein. OmpA (shown in black) includes the transmembrane domain (amino acids 46 to 159) of *E. coli* outer membrane protein A. OPH (shown in blue) includes *Flavobacterium* sp. organophosphorus hydrolase and its signal peptide.



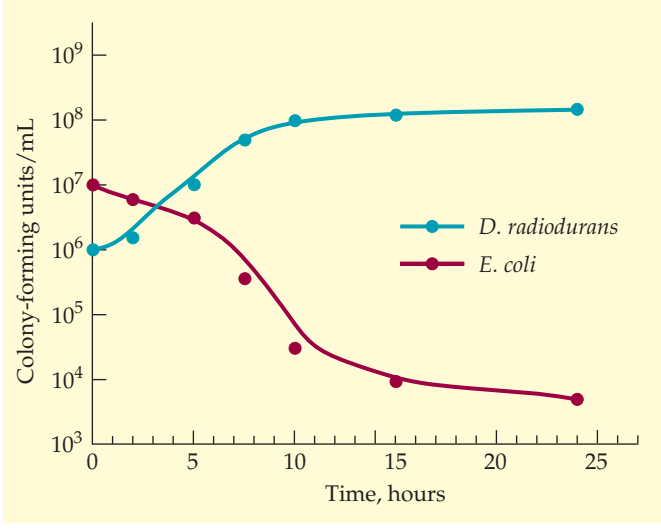


FIGURE 14.13 Effect of γ-irradiation on the growth of *E. coli* and *D. radiodurans*.

and conferred upon *D. radiodurans* the ability to degrade toluene, chlorobenzene, and 3,4-dichloro-1-butene irrespective of the presence or absence of high levels of ionizing radiation. The successful expression of toluene dioxygenase, an enzyme with several protein components and including metal and organic cofactors, suggests that many less complex biodegradative enzyme systems could also be expressed in *D. radiodurans*. Since *D. radiodurans* is tolerant of high levels of toluene, and a number of other organic compounds, once this bacterium has been genetically engineered to express the appropriate biodegradation pathway, it should be able to degrade a variety of organic pollutants in a radioactive environment. However, it remains to be seen how these engineered organisms behave under field conditions.

Nitroaromatics. For many years, a large number of different nitroaromatic compounds have been used industrially as dyes, plasticizers, explosives, solvents, and pesticides. Many of these compounds are recalcitrant to breakdown, persist in the environment, and are now considered to be toxic and sometimes carcinogenic pollutants. For example, the compound 4-nitrophenol (Fig. 14.14) is formed by the hydrolysis of the insecticide parathion and is considered to be a priority environmental pollutant, in part because it leads to numerous human health problems. Similarly, the compound 3-methyl-4-nitrophenol (Fig. 14.14) is a toxic breakdown product of the agricultural insecticide fenitrothion.

The bacterium *Burkholderia* sp. strain DNT facilitates the breakdown and detoxification of 2,4-dinitrotoluene. First, the enzyme 2,4-dinitrotoluene dioxygenase removes one nitro group to form 4-methyl-5-nitrocatechol (Fig. 14.14), which is then converted to 2-hydroxy-5-methylquinone by the enzyme 4-methyl-5-nitrocatechol monooxygenase. Unfortunately, the 4-methyl-5-nitrocatechol monooxygenase has a very narrow substrate range, which includes 4-nitrocatechol (Fig. 14.14), as well as 4-methyl-5-nitrocatechol. However, this enzyme cannot efficiently use either of the seemingly similar substrates 4-nitrophenol or 3-methyl-4-nitrophenol. In an effort to expand the substrate range of the enzyme, the *Burkholderia* sp.

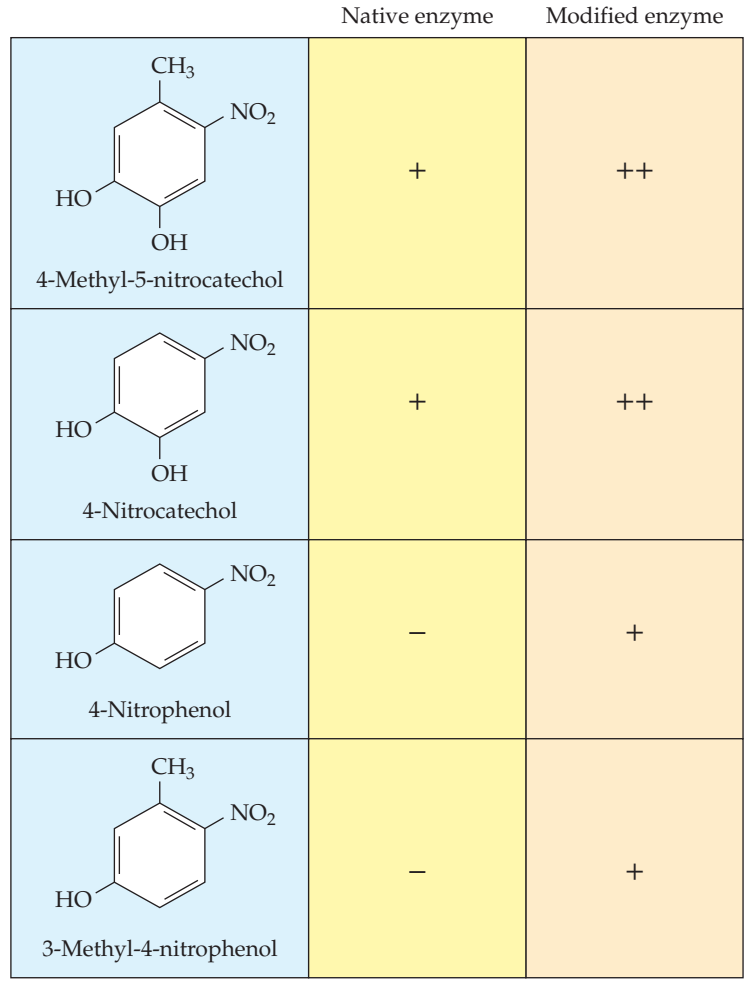


FIGURE 14.14 Breakdown of nitroaromatic compounds by a strain of *Burkholderia* sp. that produces either the native form or a modified form of the enzyme 4-methyl-5-nitrocatechol monooxygenase.

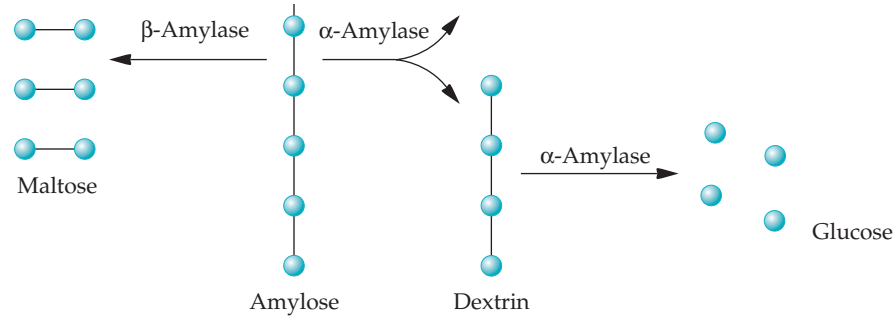
strain DNT 4-methyl-5-nitrocatechol monooxygenase gene was isolated and subjected to error-prone PCR, and the randomly mutated genes were cloned into a plasmid vector. The library of mutated 4-methyl-5-nitrocatechol monooxygenase genes was transferred into *E. coli* cells by electroporation, and the transformants were screened for activity on agar plates containing 4-nitrophenol. Cells that contained the wild-type gene turned the colonies light brown, while one transformant (of 3,000 tested) turned the colony dark brown, indicating that the 4-nitrophenol was being broken down. When the DNA sequence of the 4-methyl-5-nitrocatechol monooxygenase gene from the transformant that turned the medium dark brown was determined, two amino acids within the encoded protein had been altered. Amino acid 22 was changed from methionine to leucine, and amino acid 380 was changed from leucine to isoleucine. These two amino acid alterations resulted in the altered enzyme having 10 times greater activity toward 4-nitrophenol and 4 times greater activity toward 3-methyl-4-nitrophenol than the native form of the enzyme. In addition, the modified enzyme had about 50% more activity than the native enzyme toward

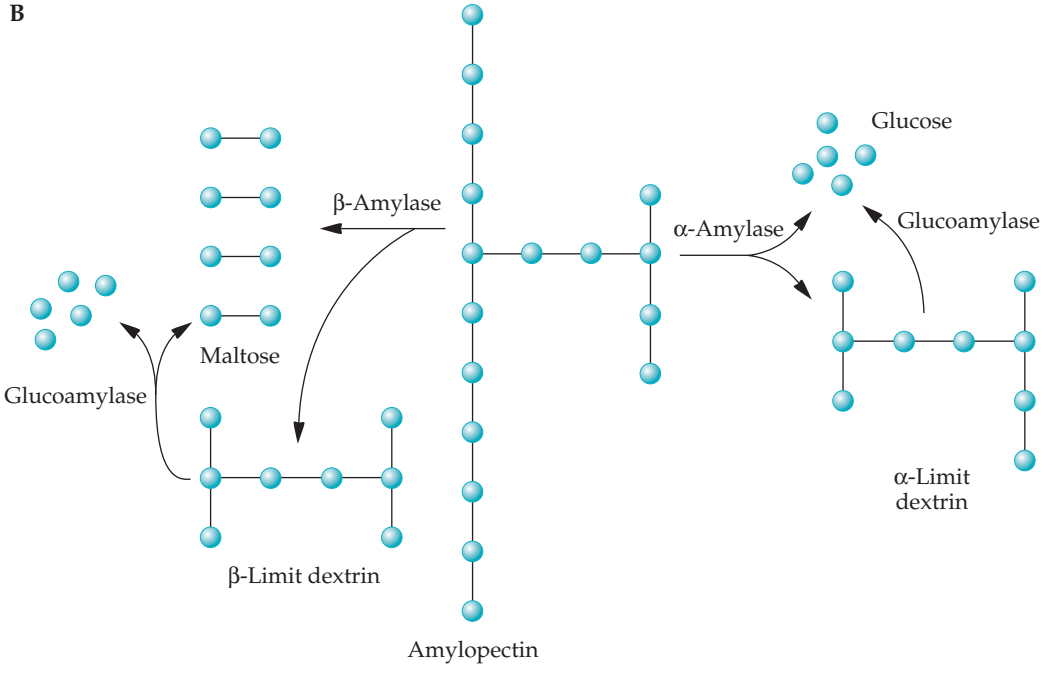
4-nitrocatechol and 4-methyl-5-nitrocatechol. The changes in activity that were found following the mutagenesis of the 4-methyl-5-nitrocatechol monooxygenase gene are an important first step in developing a bacterium that can effectively degrade 4-nitrophenol and 3-methyl-4-nitrophenol.

Utilization of Starch and Sugars

Starch, the major food reserve polysaccharide in plants, consists of a mixture of linear homopolymers (amylose) and branched homopolymers (amylopectin) of D-glucose. Amylose is made up of linear chains of 1×10^2 to 4×10^5 D-glucose residues linked by α-1,4 bonds (Fig. 14.15A). Amylopectin consists of short linear chains of approximately 17 to 23 glucose units that are linked by α-1,4 bonds and joined by 1,6 linkages and some 1,3 linkages to form a highly branched structure that contains 1×10^4 to 4×10^7 glucose residues (Fig. 14.15B). The degree of branching and the ratio of amylose to amylopectin vary with the source and age of the starch.

FIGURE 14.15 **(A)** Pathway for the enzymatic hydrolysis of amylose; **(B)** pathway for the enzymatic hydrolysis of amylopectin. The blue circles represent D-glucose residues.





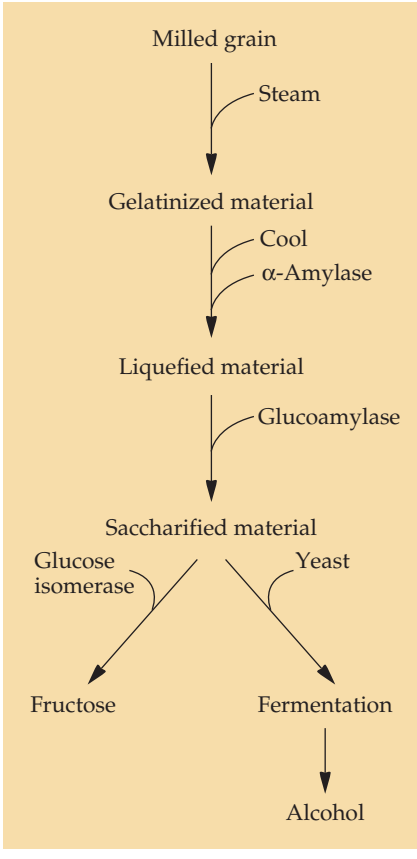


FIGURE 14.16 Industrial production of fructose and alcohol from starch.

Commercial Production of Fructose and Alcohol

The major use of starch is in the food and brewing industries, where it is usually hydrolyzed to low-molecular-weight components before being converted into other compounds, especially fructose and alcohol. The most important enzymes for breaking down and transforming starch are α-amylase, glucoamylase, and glucose isomerase. Together, these three enzymes account for approximately 30% of the cost of all enzymes currently used for industrial processes.

Fructose and alcohol are produced commercially from starch by multistep processes that include both enzymatic and nonenzymatic reactions, as follows (Fig. 14.16).

1. The procedure begins with the gelatinization of milled grain (often corn [maize], which is approximately 40% starch). This treatment—steam cooking under pressure—exposes the surface area of the starch, thereby making it more readily available for subsequent enzymatic hydrolysis. The product of this process has a gel-like consistency.
2. The gelatinized starch is cooled to 50 to 60°C, and α-amylase is added. In this liquefaction step, the gel-like starch is enzymatically digested by the hydrolysis of the available α-1,4-linkages to form low-molecular-weight polysaccharides. A high temperature is used because enzyme hydrolysis is more rapid at high temperatures and the enzyme does not efficiently penetrate the gelatinized starch at lower temperatures.
3. The final step for the release of glucose includes the addition of glucoamylase and results in the saccharification—complete hydrolysis—of the remaining polysaccharides, including both linear and cross-linked molecules.

The end product of these treatments is glucose, which may be either converted into alcohol as a result of fermentation by yeast cells or transformed into fructose by the enzyme glucose isomerase. The success of the latter enzymatic conversion has led to the replacement of sucrose by fructose, which is much cheaper than sucrose, as a sweetener for prepared foods and beverages in North America. The source of starch for commercial fructose production is generally corn, and the final product is called high-fructose corn syrup or high-fructose syrup, although it contains a mixture of fructose and glucose. High-fructose corn syrup is typically either 90% fructose (for baked goods), 55% fructose (in soft drinks), or 42% fructose (in sports drinks), with the remainder being glucose.

The enzyme α-amylase randomly hydrolyzes α-1,4-linkages in both amylose and amylopectin chains, yielding a mixture of glucose, maltose (two glucose molecules joined by an α-1,4 linkage), maltotriose (three glucose molecules joined by α-1,4 linkages), and a series of α-limit dextrins, which are the portions of the amylopectin chains that contain cross-links (Fig. 14.15). Although α-amylase can be isolated from a variety of microorganisms, it is commonly obtained from *Bacillus amyloliquefaciens* for industrial purposes.

For some applications, the enzyme β-amylase is used in addition to or in place of α-amylase to digest starch. By hydrolyzing alternate α-1,4 linkages from the ends of amylose and amylopectin chains, β-amylase cleavage yields primarily maltose residues and various β-limit dextrins.

MILESTONE

Microorganisms Having Multiple Compatible Degradative Energy-Generating Plasmids and Preparation Thereof

A. M. CHAKRABARTY
U.S. patent 4,259,444, 1981

Before the advent of recombinant DNA technology, one of the ways that DNA could be moved from one microorganism to another was by conjugation, in which entire plasmids were transferred between organisms. Chakrabarty transferred degradative plasmids—plasmids encoding all of the enzymes involved in the biodegradation of a particular compound—from one bacterium to another until he had constructed bacterial strains that contained several degradative plasmids. Each degradative pathway, encoded by the genes on a plasmid, degraded a different organic molecule. Starting with four separate bacteria, Chakrabarty and his coworkers constructed a single bacterium that contained the pathways to degrade camphor, octane, salicylate, and naphthalene. While this work was scientifically interesting and innovative in and of itself (especially since it was conducted in the early 1970s, before the development of most of the techniques of recombinant DNA that are now taken for granted), the key issue for the biotechnology industry was that this invention was awarded a U.S. patent in March 1981, nearly 9 years after the application was first filed. Following the landmark decision by the U.S. Supreme Court in Chakrabarty's favor, it was ruled that genetically engineered microorganisms were inventions that could be patented in the same way as any other invention. As much as any landmark scientific experiment, this court decision and the subsequent award of the patent has become a cornerstone of the development of the biotechnology industry.

The enzyme glucoamylase hydrolyzes α-1,3, α-1,4, and α-1,6 linkages; however, because it is less efficient than α-amylase in cleaving α-1,4 linkages, it is usually used in conjunction with α-amylase. The major role of glucoamylase is digestion of the cross-links of amylopectin, which results in its complete breakdown to glucose. Glucoamylase and other enzymes are used to reduce the carbohydrate (limit dextrin) content of normal beers to produce the so-called light and dry varieties. Although glucoamylase digestion is usually performed prior to the onset of the fermentation, the two steps may be combined. A number of organisms produce glucoamylase, but for industrial purposes, it is usually obtained from the fungus *Aspergillus niger*.

Altering Alcohol Production

The enzymes that are used in the production of alcohol or fructose from milled grain are major components of the overall cost of the process. These enzymes are often used only once and then discarded. Thus, innovative approaches to the inexpensive large-scale production of the enzymes could lower the cost of alcohol or fructose production. There are several ways to achieve this end.

- Each of the enzymes could be overproduced in a fast-growing recombinant microorganism that utilizes an inexpensive substrate, thereby lowering the cost compared with production from native organisms.
- Variants of α-amylase, either naturally occurring or genetically manipulated, that function efficiently at 80 to 90°C could be used to allow the liquefaction step to be performed at this temperature. Heat-resistant α-amylase would speed the hydrolysis of gelatinized starch while decreasing the amount of energy that is required to cool the gelatinized starch to a temperature suitable for starch hydrolysis.

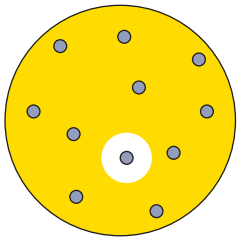


FIGURE 14.17 Detection of a *B. subtilis* clone expressing an α-amylase gene. The halo results from the degradation of the starch in the medium in the vicinity of the clone by secreted α-amylase.

- The α-amylase and glucoamylase genes could be altered so that each enzyme would have the same temperature and pH optimum, thereby enabling the liquefaction and saccharification steps to be performed under the same conditions.
- An enzyme that could efficiently degrade raw starch could be found or engineered, obviating the need for the gelatinization step and thereby saving a large amount of energy.
- A fermentation organism that can synthesize and secrete glucoamylase could be developed, eliminating the need to purify and add this enzyme during fermentation.

A considerable amount of research has been initiated to determine whether these possibilities are feasible.

Genes that code for α-amylase have been isolated from a number of organisms, including *B. amyloliquefaciens* and the high-temperature-tolerant bacterium *Bacillus stearothermophilus*. Briefly, chromosomal DNA was isolated, partially digested with the restriction enzyme Sau3AI, and then ligated to BamHI-digested pUB110 DNA. This plasmid has a unique BamHI site and carries a kanamycin resistance gene. The clone bank was transformed into *Bacillus subtilis,* which does not have α-amylase activity, and transformants were selected for resistance to kanamycin. All transformants were tested for the production and secretion of α-amylase as follows. After the transformants had formed colonies at 65°C on solid medium containing starch, the plates were exposed to iodine vapor. The colonies producing α-amylase were surrounded by a distinctive halo, or clear zone, indicating that the starch in the immediate vicinity of these cells had been hydrolyzed (Fig. 14.17). A positive starch–iodine test signifies that the transformed cell contains an α-amylase gene that is transcribed from its own promoter, because the vector does not carry a promoter. Also, a secretion signal is present, because the substrate is too large to enter the cell, and therefore the halo must be due to the activity of a secreted α-amylase. The availability of α-amylase genes from varied sources will enable researchers to carry out specific genetic modifications that suit the needs of specific industrial processes.

With the aim of bypassing the saccharification step during the production of alcohol from starch, researchers isolated a full-length glucoamylase complementary DNA (cDNA) from the fungus *Aspergillus awamori* and cloned it into a *Saccharomyces cerevisiae* plasmid under the control of the promoter and transcription terminator regulatory signals from the yeast enolase *(ENO1)* gene. A laboratory strain of *S. cerevisiae* that was transformed with the plasmid carrying the glucoamylase cDNA was able to express this activity and to ferment soluble starch to alcohol, thus demonstrating that the approach is feasible.

Unfortunately, this laboratory strain of *S. cerevisiae* has a number of properties that make it ill suited for use in a commercial process, including an inability to tolerate high levels of alcohol, inefficient expression of the glucoamylase cDNA, and loss of plasmids unless special conditions (selective pressure) are used for their maintenance. These problems, however, are not insurmountable. First, the level of glucoamylase expression was increased approximately fivefold by deleting a 175-base-pair (bp) negative regulatory region from the *ENO1* promoter on the plasmid. Second, the plasmid was modified by deleting its yeast origin of replication and adding a segment of DNA that is homologous to a yeast chromosomal site, thereby

converting it to an integrating vector. With this form of the plasmid, the complete glucoamylase construct was incorporated into a chromosomal site and stably maintained. Third, another *S. cerevisiae* strain (brewer's yeast) that tolerates high levels of alcohol was used as the host cell. The integrating vector was used to transform this yeast strain.

As a result of these modifications, investigators created two novel yeast strains that performed better than a naturally occurring amylolytic (starch-hydrolyzing) yeast, *Saccharomyces diastaticus*, which is closely related to *S. cerevisiae* and can hydrolyze and ferment soluble starch (Table 14.4). The performance of the brewer's strain, which contained the integrated glucoamylase gene, was superior to that of the laboratory strain with the same gene on a multicopy plasmid. This difference probably reflects plasmid instability with the concomitant loss of the introduced glucoamylase gene. Neither the laboratory strain nor the brewer's strain of *S. cerevisiae* was able to utilize soluble starch unless the strain was transformed with the cloned glucoamylase gene. In both the plasmid and integrated forms, the *A. awamori* glucoamylase cDNA was under the control of regulatory signals of *ENO1* from which the 175-bp negative regulatory region had been removed. The plasmid was maintained under selective pressure.

In an effort to produce large amounts of glucoamylase, several copies of the glucoamylase gene were integrated into the chromosomal DNA of the fungus *A. niger*. Surprisingly, there was no correlation between the number of copies of the glucoamylase gene in the chromosomal DNA and the amount of measured enzyme activity. On the other hand, the level of enzyme activity was strongly dependent upon the sites where the genes were inserted. Thus, merely increasing the gene copy number is not sufficient to produce a larger amount of active enzyme.

While most strains of the yeast *S. cerevisiae* encode only an intracellular glucoamylase, which is generally expressed only during sporulation, a few strains encode a secreted glucoamylase. Unfortunately, none of these enzymes contains a starch-binding domain. Thus, while *S. cerevisiae* glucoamylases can hydrolyze soluble starch (dextrins), they are unable to break down the larger molecules of insoluble starch. In an attempt to construct a strain of *S. cerevisiae* that could hydrolyze insoluble starch and therefore be of use in a variety of industrial processes, the coding region of a secreted *S. cerevisiae* glucoamylase gene was fused to a DNA fragment encoding a starch-binding domain from the fungus *A. niger* (Fig. 14.18). When this construct was expressed in *S. cerevisiae*, the presence of the starch-binding domain increased the ability of the enzyme to degrade

TABLE 14.4 Fermentation of soluble starch (25% [wt/vol]) by various yeast strains

Strain	Carbohydrate utilized (%)	Ethanol produced (g/liter)	Ethanol yield (g/g of substrate)
Laboratory	5	<0.1	0
Laboratory + gene on a plasmid	68	75.6	0.41
Brewer's	<1	3.1	0
Brewer's + integrated gene	93	118.2	0.48
S. diastaticus	43	44.2	0.38

Adapted from Cole et al., *Bio/Technology* **6**:417–421, 1988.

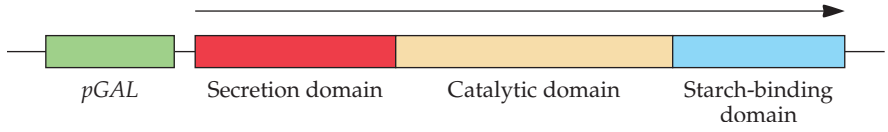


FIGURE 14.18 The genetic construct used to produce in *S. cerevisiae* a secreted glucoamylase containing a starch-binding domain (from *A. niger*). The construct is transcribed under the control of the yeast *pGAL* promoter in the direction indicated by the arrow.

insoluble starch by about sixfold. While this system is far from optimized, this simple genetic manipulation is an important initial step in the development of more efficient industrial systems for the production of alcohol from starch.

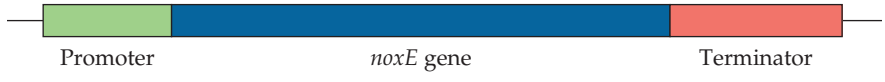
Low-ethanol wines. In recent years, many wine drinkers have expressed a strong preference for wines that contain only low levels of alcohol. In addition, consumer preferences have also moved to wines with a high flavor intensity that are prepared from fully matured grapes. However, the juice that is obtained from fully matured grapes generally contains a very high sugar concentration, which in turn produces wines with high levels of alcohol. A number of attempts have been made to engineer brewer's yeasts (*S. cerevisiae*) to reduce the ethanol content of the wine that is produced. Although a number of these approaches have successfully lowered the ethanol concentration, they typically cause the accumulation of undesirable side products. Nevertheless, one novel strategy seems to have avoided many of the problems of past efforts. By expressing in *S. cerevisiae* a gene (*noxE*) from the bacterium *Lactococcus lactis* that encodes an H2O-NADH oxidase (Fig. 14.19), it was possible to significantly alter some of the metabolic fluxes within the yeast cell. Transformed yeast cells that carried this gene within their chromosomal DNA had an intracellular NADH content that was about 75 to 80% lower than that of the native yeast strain and an oxidized nicotinamide adenine dinucleotide (NAD+) concentration that was 32 to 45% higher. At the same time, the reduced nicotinamide adenine dinucleotide phosphate (NADPH) and oxidized nicotinamide adenine dinucleotide phosphate (NADP+) ratios of the two yeast strains were identical. The transformed yeast showed a 15% decrease in the amount of ethanol produced and, unfortunately, increases of approximately threefold in the amounts of acetaldehyde and acetic acid that were produced, which

FIGURE 14.19 **(A)** Oxidation of NADH catalyzed by the H_2O-NADH oxidase encoded by the *noxE* gene from the bacterium *L. lactis;* **(B)** the bacterial *noxE* gene under the transcriptional control of the yeast glyceraldehyde 3-phosphate dehydrogenase gene promoter and the yeast phosphoglycerate kinase transcription terminator. The entire construct was integrated into the yeast chromosomal DNA.

A

$$4NADH + O_2 \xrightarrow{noxE} 4NAD^+ + 2H_2O$$

B



impair both growth and fermentation, as well as imparting unacceptable flavors to the wine. To overcome these side effects, researchers undertook a systematic study of the growth of this modified yeast strain under a wide range of conditions. It was observed that if oxygen was supplied to the growing cells only during the stationary phase, there was a 7% reduction in the ethanol yield compared to the native strain, without the inhibitory levels of acetaldehyde and acetone being formed. Researchers are now investigating whether the wine that is produced using the above-mentioned yeast strain is suitable for human consumption.

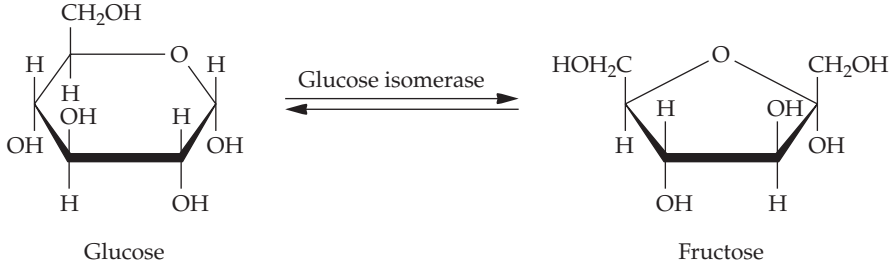
Improving Fructose Production

The enzyme glucose isomerase should really be called xylose/glucose isomerase, because it primarily catalyzes the conversion of the five-carbon sugar D-xylose to D-xylulose, with the conversion of D-glucose to D-fructose being a secondary or side reaction (Fig. 14.20). Kinetically, xylose/glucose isomerase has a lower k_{cat} (catalytic rate constant) and a higher K_m (binding constant) for glucose than for xylose, which means that xylose is bound more tightly to the enzyme than is glucose and that xylose is converted more rapidly to xylulose than glucose is converted to fructose.

Xylose/glucose isomerases are intracellular enzymes and, as such, do not yield the same quantities or purity of product as do the extracellular, or secreted, enzymes that are used in many industrial processes. Most industrial enzymes are used without any extensive purification. An extracellular enzyme preparation generally contains many fewer proteins than an intracellular extract. In addition, the preparation of an intracellular protein extract requires separation of the cells from the growth medium, mechanical disruption of the cells, and removal of cell debris following disruption. These factors lead to higher production costs for xylose/glucose isomerase than for many other industrial enzymes. One way to overcome this problem is to use a batch of xylose/glucose isomerase more than once. This recycling can be achieved by immobilizing the enzyme on a solid support, which both stabilizes the enzyme and facilitates its reuse.

The isomerization of glucose to fructose is a reversible reaction, and the final fructose content is dependent on the reaction temperature. The higher the temperature, the greater the fructose content in the final product. Most commercial processes use conversion temperatures of around 60°C. The enzyme is typically used in an immobilized state that is obtained by cross-linking it to itself with glutaraldehyde and then using it in a continuous process in a packed bed reactor. Under these conditions, a batch of enzyme can be used for approximately 150 to 200 days before it is discarded. Consequently, increasing the temperature optimum for the enzymatic

FIGURE 14.20 Conversion of glucose to fructose, catalyzed by glucose isomerase.



Glucose Fructose

activity and the thermostability of xylose/glucose isomerase is one way to make it more efficient.

The thermophilic bacterium *Thermus thermophilus* produces a xylose/glucose isomerase that not only is active at 95°C, but also is very stable at high temperatures. Therefore, this enzyme is a good candidate for use in industrial processes. Unfortunately, wild-type *T. thermophilus* does not produce large amounts of the enzyme. To circumvent this problem, the *T. thermophilus* xylose/glucose isomerase gene was isolated and expressed in *E. coli* and *Bacillus brevis* under the control of various promoters and ribosome-binding sites (Table 14.5). One of the constructs (the last one listed in Table 14.5) overproduced xylose/glucose isomerase more than 1,000-fold relative to the amount found in the original organism. Therefore, with this construct, high yields of a thermostable xylose/glucose isomerase can be produced for the industrial synthesis of fructose from glucose.

In addition, the substrate specificity of the enzyme can be enhanced. In one series of experiments, site-directed mutagenesis was used to change the nucleotides encoding either one or two amino acids of the xylose/glucose isomerase from the thermophilic organism *Clostridium thermosulfurogenes*. The targeted sites were selected for modification because of other evidence indicating that the corresponding amino acids were involved in substrate binding. Changing the tryptophan at amino acid residue 139 to phenylalanine or the valine at amino acid residue 186 to threonine produces a 1.7- to 2.6-fold increase in the catalytic efficiency (k_{cat}/K_m) of the enzyme toward glucose (Table 14.6). Moreover, these changes cause a two- to sevenfold reduction in the k_{cat}/K_m values of the enzyme toward xylose. When an enzyme has both of these amino acid changes, the k_{cat}/K_m value for glucose increases by 5.7-fold and the k_{cat}/K_m value for xylose decreases by 4.5-fold. The double amino acid modification changes an enzyme that was initially 17 times more reactive with xylose than with glucose to one that is now 1.5 times more reactive with glucose than it is with xylose. The shift in specificity that has been achieved, together with the thermostability of this xylose/glucose isomerase, should make it attractive for use in the industrial conversion of glucose to fructose.

Silage Fermentation

Crops such as grasses, corn, and alfalfa need to be preserved so that they can be used as animal feed many months after the crop is harvested. Traditionally, these crops are preserved by naturally occurring lactic acid bacteria that use

TABLE 14.5 Amounts of *T. thermophilus* xylose/glucose isomerase in different bacteria

Species	Plasmid copy no.	Promoter source	Source of ribosome-binding site	Enzyme activity (units/liter)
T. thermophilus	None	*T. thermophilus*	*T. thermophilus*	20
E. coli	200	*E. coli lac*	*T. thermophilus*	190
E. coli	20	*E. coli tac*	*T. thermophilus*	1,790
E. coli	20	*E. coli tac*	*E. coli*	3,260
E. coli	20	Phage T7 f10	*E. coli*	7,050
B. brevis	20	*B. brevis cwp*	*T. thermophilus*	1,400
B. brevis	20	*B. brevis cwp*	*T. thermophilus*	25,000

Adapted from Dekker et al., *Appl. Microbiol. Biotechnol.* **36:**727–732, 1992.

The first row represents data for the original enzyme-producing strain. All the other strains are transformants carrying the *T. thermophilus* xylose/glucose isomerase gene on a multicopy plasmid.

TABLE 14.6 Catalytic efficiency of wild-type and mutant xylose/glucose isomerases from *C. thermosulfurogenes*

Amino acid change(s)	Catalytic effiency (k_{cat}/K_m) ($min^{-1} mM^{-1}$) toward:	
	Glucose	Xylose
None (wild type)	5.8	97.2
Trp-139 → Phe	15	13.6
Val-186 → Thr	9.7	55.4
Trp-139 → Phe, Val-186 → Thr	32.9	21.6

Adapted from Meng et al., *Proc. Natl. Acad. Sci. USA* **88:**4015–4019, 1991.

the crop as a fermentation substrate to produce lactic and acetic acids. The resulting low pH restricts the growth and metabolic activity of other microorganisms and ensures that the crop is preserved. This preservation strategy is called the making of silage. Often, the numbers of lactic acid bacteria that are found on a fresh crop are quite small, so that a bacterial inoculum, typically *Lactobacillus plantarum,* must be added. Unfortunately, these bacterial inoculants are not especially effective when the amount of water-soluble carbohydrates in the fresh crop is insufficient to support both bacterial growth and lactic acid production.

To develop a bacterium that might be effective in silage fermentation, an α-amylase gene from a strain of *Lactobacillus amylovorus* that does not support silage fermentation was spliced into the *L. plantarum* gene for conjugated acid bile hydrolase *(cbh)* and integrated into the chromosomal DNA of a strain of *L. plantarum* (Fig. 14.21). The *cbh* gene is a dispensable gene when the bacterium is grown on silage; it encodes an enzyme that is active only when the bacterium is located in an animal's intestine. This work is an important first step in the development of *L. plantarum* strains that are more effective in the fermentation of silage from crops such as alfalfa, which contain a high level of starch.

Isopropanol Production

In order to decrease our reliance on nonrenewable petroleum products, it may be possible to engineer microorganisms to produce isopropanol from glucose (e.g., derived from starch). Isopropanol may be used directly as a fuel instead of methanol to esterify fats and oils to produce biodiesel, or it may be dehydrated to yield propylene, which is used to synthesize the polymer polypropylene. To engineer *E. coli* to produce isopropanol, the genes that were introduced into *E. coli* were based on the genes encoding the isopropanol biosynthesis pathway that exists in *Clostridium beijerinckii* (an organism that produces only moderate amounts of isopropanol but is difficult to grow and to manipulate genetically). The engineering of *E. coli* to produce isopropanol required the addition of four foreign genes (Fig. 14.22). The initial source of all four genes was *Clostridium acetobutylicum;* however, genes encoding the same activities from several other bacteria were also tested in an effort to obtain a transformed strain of *E. coli* that produced the greatest amount of isopropanol. The best combination of foreign genes encoding enzymes in the isopropanol biosynthesis pathway produced nearly three times as much isopropanol as the best reported strain of *C. beijerinckii,* indicating that this strain has significant potential for use in the industrial production of isopropanol.

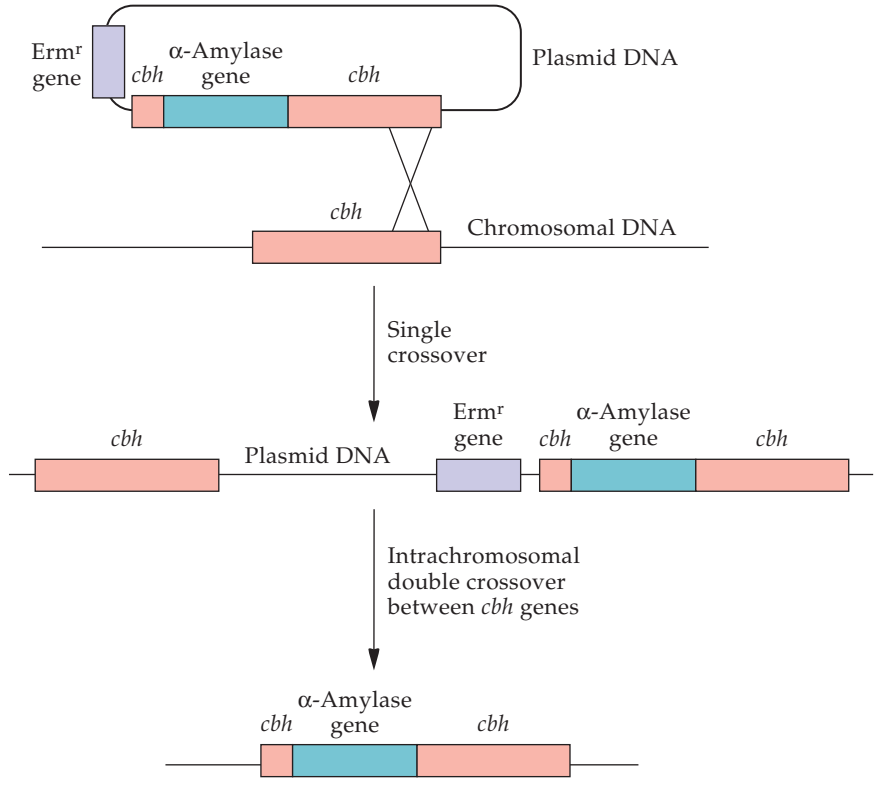


FIGURE 14.21 Chromosomal integration of an α-amylase gene into *L. plantarum*. The α-amylase gene is cloned into the *L. amylovorus cbh* gene on an *E. coli–L. plantarum* shuttle vector, which is then used to transform *L. plantarum*. Erythromycin-resistant and α-amylase-positive clones of *L. plantarum* result from a single crossover between the chromosomal and plasmid DNA at the *cbh* locus. After the growth of transformed *L. plantarum* for approximately 30 generations in the absence of selective pressure, intrachromosomal recombination resulted in the excision of the erythromycin resistance (Erm^r) gene, the chromosomal copy of the *cbh* gene, and the plasmid DNA. The final engineered *L. plantarum* carries only an α-amylase gene and no selectable marker genes.

Engineering Yeast Transcription

Conventional mutagenesis and selection have historically been used to improve the useful behavior of a range of microorganisms. Most prominently, from a biotechnological perspective, this approach has been utilized to develop microbial strains that overproduce specific antibiotics. The advantage of the approach is that it does not predetermine which gene(s) will be altered. However, conventional mutagenesis is a slow and tedious process requiring an inordinate amount of testing of mutated strains. Moreover, conventional mutagenesis and selection, or even sequential rounds of random mutagenesis of an isolated DNA fragment, introduce only a limited number of changes at each round of mutagenesis. Unfortunately, sometimes changing a fundamental property of a microorganism may require altering the expression of dozens or even hundreds of genes. While this cannot be achieved using the above-mentioned approaches, the reprogramming of a microorganism may be realized by mutagenizing one of the proteins that is responsible for regulating global transcription of the microorganism.

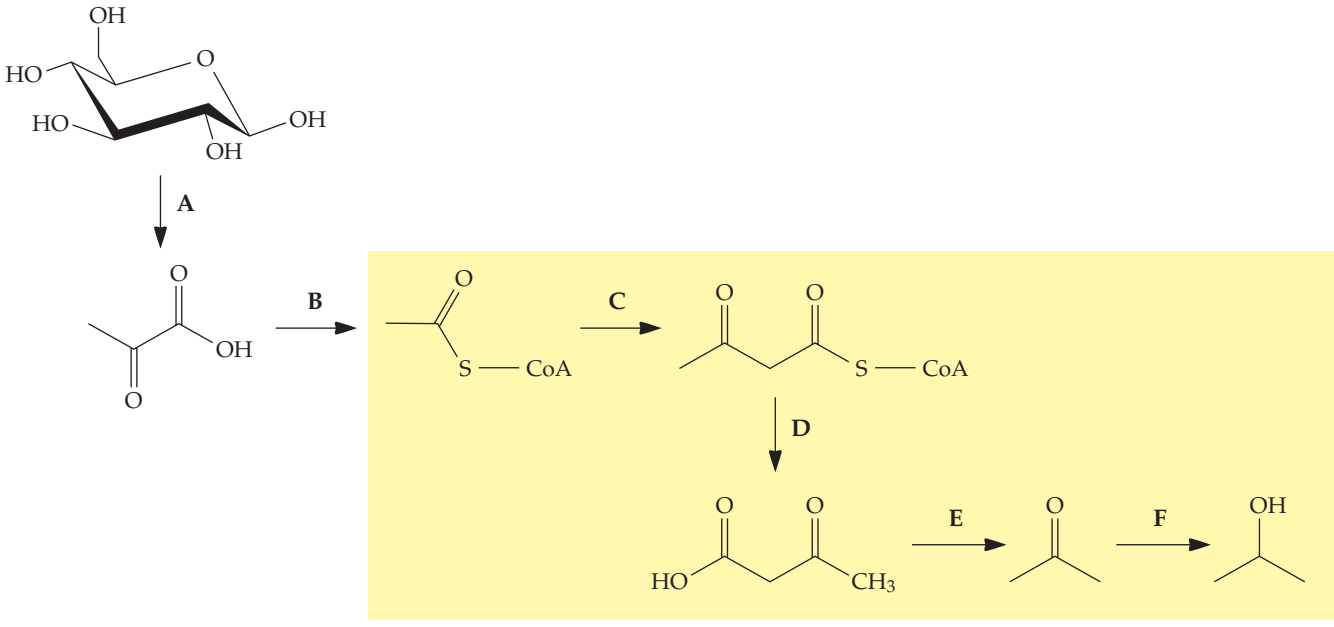


FIGURE 14.22 Engineering *E. coli* to produce isopropanol; a simplified overview of the biosynthetic pathway. The introduced portion of the pathway is highlighted in yellow. The enzymes involved in this pathway are as follows: several *E. coli* enzymatic steps are required for the conversion of glucose to pyruvate **(A);** pyruvate is converted to acetyl-CoA by an *E. coli* enzyme **(B);** acetyl-CoA acetyltransferase from *C. acetobutylicum* catalyzes the formation of acetoacetyl-CoA **(C);** acetoacetyl-CoA transferase from *E. coli* catalyzes the production of acetoacetate **(D);** acetoacetate decarboxylase from *C. acetobutylicum* catalyzes the synthesis of acetone **(E);** and secondary alcohol dehydrogenase from *C. beijerinckii* produces isopropanol from acetone **(F)**.

In order to engineer yeast to more efficiently produce alcohol, it is necessary that the yeast strain be able to tolerate high concentrations of both glucose and ethanol. To achieve this end, it is likely that the expression of a large number of yeast proteins must be altered. Moreover, it is by no means clear which proteins need to have their levels of expression either increased or decreased. Instead, one group of researchers undertook to reprogram a significant portion of yeast metabolism by generating a large number of randomly mutated yeast transcription factor genes. In a small number of cases, the modified transcription factor will alter yeast gene expression in a manner that increases the tolerance of yeast for high levels of both glucose and ethanol. In yeast, 15 different proteins bind to DNA and regulate the promoter specificity of RNA polymerase II. To dramatically modify yeast metabolism (Fig. 14.23), the gene encoding transcription factor *SPT15* was altered by error-prone PCR (see chapter 8). This amplification reaction produces a range of different mutagenized *SPT15* genes. All of the mutagenized *SPT15* genes were cloned onto a plasmid vector and then introduced into wild-type yeast *(S. cerevisiae)* cells. The transformed cells were grown on agar medium in the presence of 6% ethanol and 120 g/liter of glucose. Any transformant that grew better on this medium than the native yeast strain was then more fully characterized. One particular mutant displayed a higher level of cell viability and grew faster than the native yeast strain at ethanol concentrations of between 10 and 20%. When this strain was examined in detail, it was found that the *SPT15* gene had

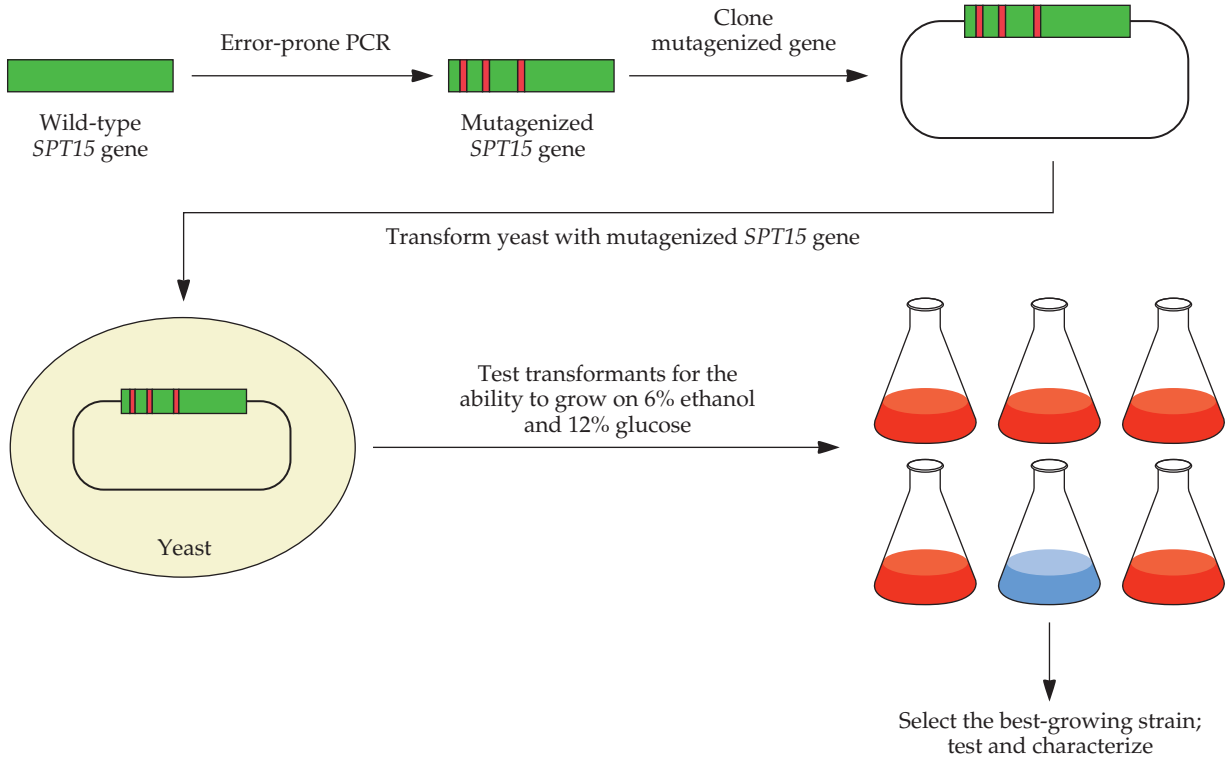


FIGURE 14.23 Engineering the yeast *S. cerevisiae* to tolerate high levels of glucose and ethanol. The yeast transcription factor *SPT15* gene was mutagenized by error-prone PCR, cloned into a plasmid vector, and used to transform a laboratory yeast strain. The transformants were selected for the ability to grow on medium containing both 6% ethanol and 12% glucose. The portions of the *SPT15* gene shown in red represent the introduced mutations.

three separate mutations, all of which were required for this activity. In addition, when the mutant was transcriptionally profiled (using microarray technology), several hundred genes were found to be differentially regulated compared to the native strain, with the majority of genes being upregulated. Interestingly, a detailed analysis of the genes whose expression was significantly altered did not reveal that a particular pathway or genetic network was primarily responsible for the observed reprogramming. Finally, the yeast strain that was used in these experiments was a standard laboratory strain, so that to incorporate this approach into a commercial process, it will be necessary to repeat the work with an industrial strain of yeast. If successful, this type of genetic manipulation could facilitate the development of strains of yeast that more efficiently convert glucose to ethanol.

Utilization of Cellulose

With an increasing world population and dramatic increases in the standard of living in parts of the developing world, meeting the growing worldwide demand for energy for heating, transportation, and industry has become a major challenge for all countries for the 21st century and beyond. Moreover, in addition to being sustainable, the future energy supply needs to be nonpolluting so that we may realize a reduction in

TABLE 14.7 Typical compositions of various lignocellulosic materials

Raw material	Amount (%) of:		
	Lignin	Cellulose	Hemicellulose
Pine wood	27.8	44.0	26.0
Birch wood	19.5	40.0	39.0
Sugarcane bagasse	18.9	33.4	30.0
Rice straw	12.5	32.1	24.0
Cotton	12.5	32.1	24.0

Adapted from Brown, *Philos. Trans. R. Soc. Lond. B* **300:**305–322, 1983.

greenhouse gas emissions. To this end, many countries around the world have begun to produce large amounts of alternative fuels in an effort to at least partially replace nonrenewable fossil fuels, such as oil and gas. In this regard, a major effort has been directed toward producing bioethanol. At the present time, Brazil produces large amounts of ethanol from the fermentation of sucrose derived from sugarcane, and the United States produces ethanol from corn starch. However, the reduction of greenhouse gases that results from the use of sugar- or starch-based ethanol is not as high as desired, and many socially conscious individuals have criticized the strategy of converting land from the production of food to the production of ethanol/energy. Moreover, sugar- or starch-based ethanol is unlikely to provide more than a small fraction of what we require. In this regard, in 2008, the Chinese government announced that it would not allow any further increase in starch-based ethanol production because of competing uses as food. If the world is going to produce ethanol on a large enough scale to significantly lower our use of fossil fuels, that ethanol will have to be produced from lignocellulosic waste products, such as corn stover, grasses, and wood chips. Thus, there is now, more than ever before, a tremendous amount of both political and scientific activity directed toward trying to produce ethanol from lignocellulosic materials.

Lignocellulosics

The polymers lignin, hemicellulose, and cellulose combine in various proportions to form a "lignocellulosic" structural support system for nearly all terrestrial plants (Table 14.7). This material constitutes a vast biomass that is often a waste product of agriculture, timber processing, and other human activity and needs to be disposed of in a safe and efficient manner or used as a resource. It has been estimated that annually ~10^{11} tons of these polymers is synthesized in the biosphere, with an energy content that is equivalent to around 640 billion tons of oil.

Lignocellulosic materials (cellulosics) have been grouped into three classes.

- Primary cellulosics include plants that are harvested specifically for cellulosic content, structural use, or feed value, e.g., cotton, timber, and hay.
- Agricultural waste cellulosics are the plant materials that remain after harvesting and processing, e.g., straw, corn stovers, rice hulls, sugarcane bagasse, animal manures, and timber residues.
- Municipal waste cellulosics encompass wastepaper and other discarded paper products.

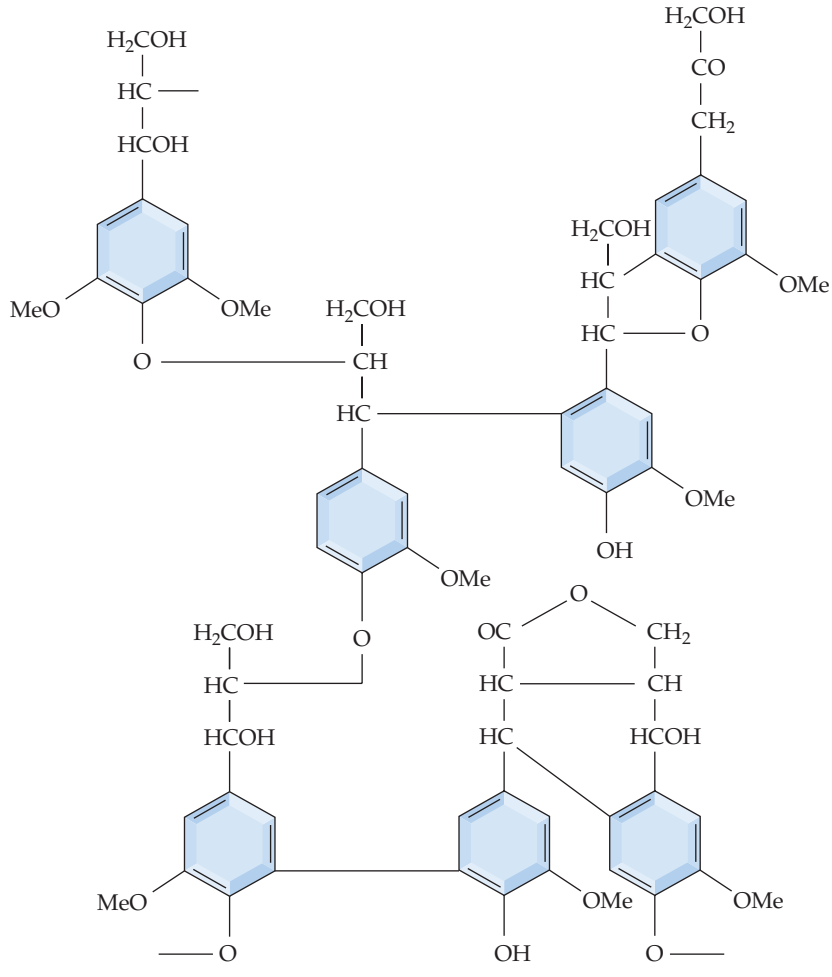
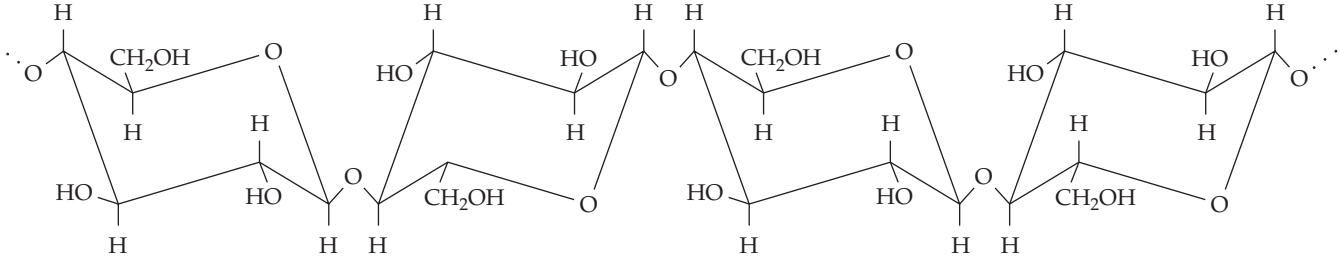


FIGURE 14.24 Schematic representation of lignin structure, showing some of the various possible linkages between the phenylpropane (a C_6 aromatic group attached to a C_3 alkyl chain) units. The phenylpropane units are linked in an unorganized, nonrepeating fashion.

Components of Lignocellulose

Lignin is a three-dimensional, globular, irregular, insoluble, high-molecular-weight (>10,000) polymer made up of phenylpropane subunits with no chains of regular repeating units or any bonds that are easily hydrolyzed either enzymatically or chemically (Fig. 14.24). The lignin polymer molecule

FIGURE 14.25 Structure of a portion of a cellulose chain. Glucose residues are joined head to tail by β-1,4 linkages.



has many different types of chemical linkages between aromatic phenylpropane units. The physical and chemical characteristics of lignin are generally attributed to the last step in lignin biosynthesis, the nonenzymatic free-radical-based joining of the phenylpropane units in a more or less random fashion. In plants, lignin is chemically bonded to hemicellulose and wraps around fibers composed of cellulose. Lignin is responsible for the rigidity of plants and for their resistance to mechanical stress and microbial attack.

Hemicelluloses are short-chain, heterogeneous polymers that contain both hexoses (six-carbon sugars, such as glucose, mannose, and galactose) and pentoses (five-carbon sugars, such as xylose and arabinose). The three major types of hemicelluloses are xylans, which have a backbone of poly-β-1,4-xylan, with side links to arabinose, glucuronic acid, and arabinoglucuronic acid; mannans, which are composed of glucomannans and galactomannans; and arabinogalactans. The origin of the lignocellulosic material usually defines the nature of the hemicelluloses. For example, xylan hemicellulose is particularly common in hardwoods, and glucomannans are characteristic of softwoods.

Cellulose, which is the simplest of the components found in lignocellulosic material, is the most abundant polymer in the biosphere. It is composed of long chains of D-glucose molecules linked in β-1,4 configuration (Fig. 14.25). Both cellulose and starch can be hydrolyzed to glucose, but their structures are very different. Starch is an energy storage molecule in which the glucose residues are linked in a manner that prevents a tightly ordered arrangement of the polymer chains. This open mesh-like structure is easily penetrated by water; as a result, starch is both water soluble and readily hydrolyzable by amylases and glucoamylases. By contrast, cellulose is a plant-supporting structural molecule. The glucose chains in cellulose are arranged in a manner that permits them to pack together in a crystal-like structure that is impervious to water. Consequently, the cellulose polymer is both insoluble and resistant to hydrolysis.

Nevertheless, cellulose is still a form of stored glucose, so it is the component of lignocellulosics that has the most potential for conversion into a variety of useful compounds, such as alcohol. However, before cellulose can be utilized, it must be released from its complex with lignin and hemicellulose. For most lignocellulosic materials, this separation requires treatment with either a strong acid or a strong base or the use of high temperature and pressure. Regardless of how the cellulose is separated from the lignocellulose complex, the energy that is necessary to achieve this adds significantly to the cost of the final product. However, since the annual production of lignocellulosic materials is huge, effective ways of enzymatically degrading cellulose and hemicellulose are being sought. Chemical and enzymatic methods for the selective degradation of lignin also are being investigated, with less success.

Isolation of Prokaryotic Cellulase Genes

A wide range of bacteria and fungi are naturally capable of degrading cellulose through the concerted action of several enzymes that collectively are referred to as cellulase. Aerobic microorganisms typically secrete large amounts of these cellulase enzymes into the medium outside of the cell. On the other hand, in anaerobic microorganisms, cellulase activity is often found as part of a multiprotein complex that is called a cellulosome (Fig. 14.26) that lies on the external surface of the cell (Fig. 14.27). Cellulases

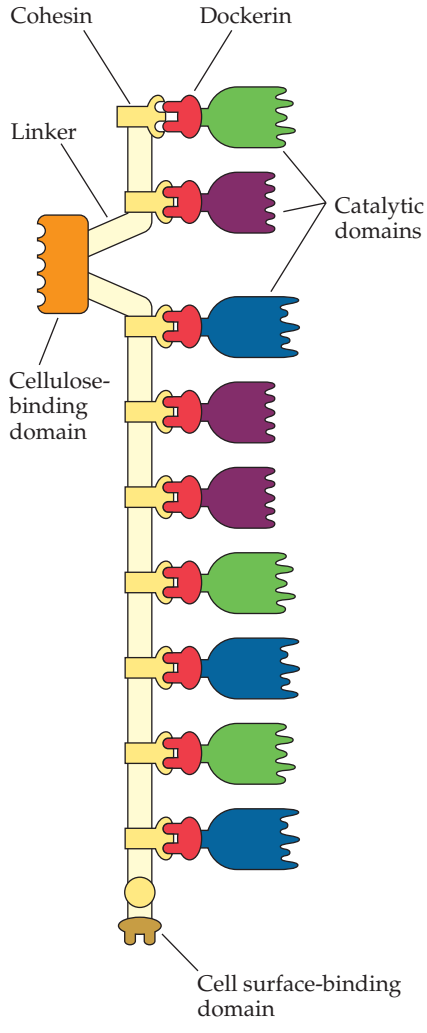


FIGURE 14.26 Schematic representation of a cellulosome protein complex. This complex contains both structural and enzymatically active protein components. The structural components include scaffoldin, a scaffolding protein (shown in yellow) that contains a strong cellulose-binding module (orange), and a number of cohesin molecules (nine are shown). The cohesin molecules act as binding sites (docking sites) for the dockerin proteins (red), which are attached to the catalytic domains (the active enzymatic components, which are shown in green, purple, and blue). Each cellulosome complex is anchored to the microbial cell surface through the cell surface-binding domain (brown).

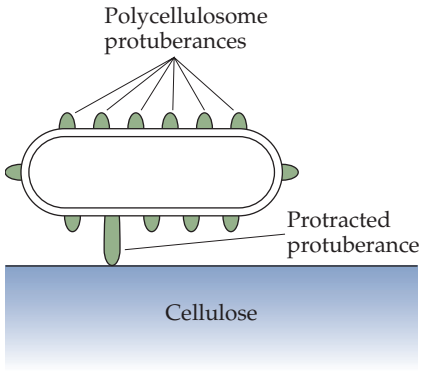


FIGURE 14.27 Interaction of a cellulolytic bacterium with cellulose. The bacterial cell surface has protuberance-like structures that contain multiple copies of the cellulosome (polycellulosome protuberances). On contact with the cellulose substrate, some of the polycellulosome protuberances protract dramatically and deposit their cellulosomes along the surface of the cellulose (detached from the bacterial surface).

consist of multiple copies of several enzymes with different enzymatic activities. including the following.

- Endoglucanase, which hydrolyzes β-1,4 linkages between adjacent glucose molecules within the amorphous (loosely packed) regions of the cellulose polymer, thereby breaking the chain in the middle
- Exoglucanase, which degrades the nicked cellulose chains from their nonreducing ends and produces glucose, cellobiose (two glucose units), and cellotriose (three glucose units)
- Cellobiohydrolase, which is often found in cellulolytic fungi and is a type of exoglucanase that removes units of 10 or more glucose residues from the nonreducing ends of the cellulose molecule
- β-Glucosidase, or cellobiase, which converts cellobiose and cellotriose to glucose

The breakdown of cellulose by microorganisms (either bacteria or fungi) that produce the various components of cellulase (Fig. 14.28) is slow and often incomplete. Therefore, genetic engineering strategies have been used in an attempt to create organisms with more effective cellulase activity. For this purpose, genes coding for the individual enzymatic functions of cellulase activity have been isolated from both prokaryotic and eukaryotic organisms.

Prokaryotic endoglucanase genes have been cloned by the following simple yet effective identification technique.

1. A clone bank of DNA from a cellulolytic prokaryote is constructed in *E. coli,* and the host cells are grown overnight on solid medium containing a selective antibiotic.
2. The colonies are then overlaid with agar containing carboxymethyl cellulose (CMC), a soluble derivative of cellulose, and the petri plates are incubated at 37°C for several hours. During this time, the CMC molecules that are present in the immediate vicinity of a colony that both synthesizes and secretes an endoglucanase are partially digested. Transformants that synthesize but do not secrete the cloned endoglucanase are not able to degrade the substrate, because it is too large to enter the cell.
3. The digested regions of the CMC are visualized by first flooding the petri plate with a solution of the dye Congo red, which is not toxic to the bacteria, followed by a wash with a solution of sodium chloride. Congo red selectively binds to high-molecular-weight cellulose chains and gives a red color; conversely, it binds weakly to low-molecular-weight polysaccharides and produces a yellow hue. The sodium chloride treatment stabilizes the binding of the dye. If a bacterial colony produces a secreted endoglucanase, it will be surrounded by a yellow halo; the background, where the CMC has not been degraded, will be red (Fig. 14.29).

This technique has been successfully used in the isolation of endoglucanase genes from *Streptomyces, Clostridium, Thermoanaerobacter, Thermomonospora, Erwinia, Pseudomonas, Cellvibrio, Ruminococcus, Cellulomonas, Fibrobacter,* and *Bacillus* species.

There is no convenient plate assay for detecting cells with a cloned exoglucanase gene, so immunological screening has been used to pick out the recombinant clones that express exoglucanase. Although this approach

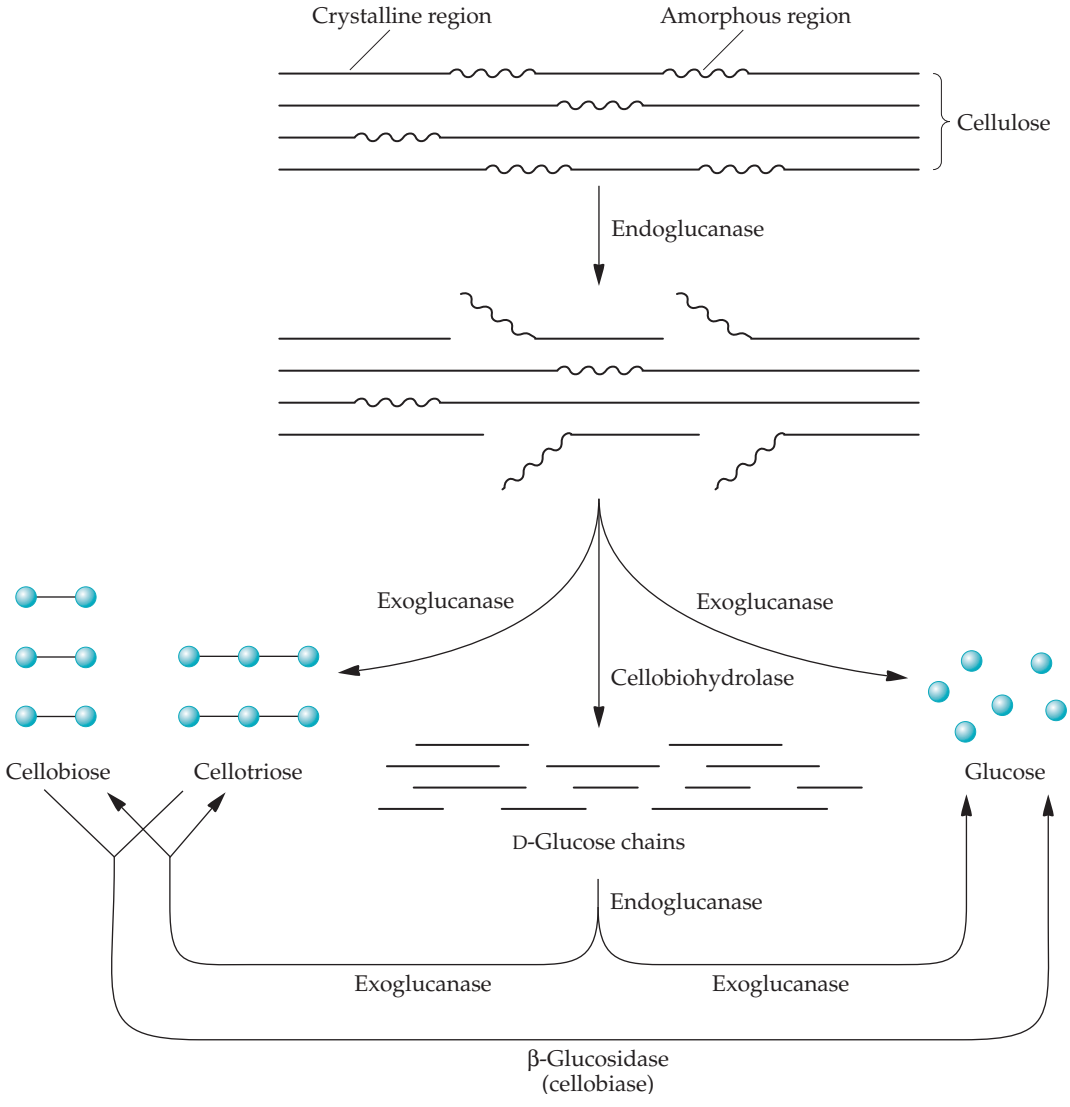


FIGURE 14.28 Enzymatic biodegradation of cellulose. Cellulose hydrolysis begins with the cleavage of β-1,4 linkages within the accessible amorphous regions of the cellulose chains by endo-glucanase(s). This reaction is followed by the removal of oligosaccharides from the partially cleaved cellulose chains by exoglucanase(s) and cellobiohydrolase(s). The degradation of cellulose is completed when the cellobiose and cellotriose are converted to glucose by β-glucosidase.

requires specific antibodies directed against the target protein, the protein does not have to be secreted to be detected. Recombinant cells can be lysed in situ, e.g., by exposure to chloroform vapor, before the cytoplasmic proteins are transferred to a nylon or nitrocellulose membrane for subsequent immunological testing. In these tests, replica plates are used to ensure that viable cells are available for further propagation and use.

Prokaryotic β-glucosidase genes have been isolated by transforming a clone bank from a β-glucosidase-producing microorganism into *E. coli* and then selecting for transformants that can grow on minimal media with cellobiose as the sole carbon source. Alternatively, clones that express β-glucosidase activity can be detected with a chromogenic substrate, such

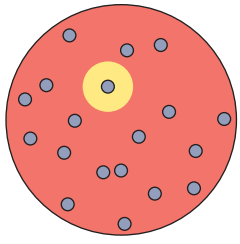


FIGURE 14.29 Detection of an *E. coli* clone expressing a bacterial endoglucanase gene. The yellow halo indicates the presence of a positive clone that degrades the soluble cellulose (CMC) in the medium in the vicinity of the clone via the secreted endoglucanase.

as 5-bromo-4-chloro-3-indolyl-β-D-glucopyranoside (BCIP), in the plating medium or with MacConkey–cellobiose agar. In both cases, β-glucosidase-positive colonies turn red.

Isolation of Eukaryotic Cellulase Genes

The strategy of DNA hybridization screening of either cDNA or genomic clone banks with a heterologous probe has not been particularly effective for isolating cellulase genes, because the sequences of cellulolytic enzymes from diverse sources are not very similar. To isolate the messenger RNAs (mRNAs) that encode cellulolytic enzymes from fungi or plants, a novel protocol had to be implemented. Unfortunately, these mRNAs usually constitute only a small fraction of the total mRNA population. Therefore, it is often necessary to enrich for the target mRNA or cDNA and to eliminate cDNA clones that do not carry the target sequences. To meet these ends, the technique of "differential hybridization" has been used for the isolation of a number of different induced eukaryotic cellulase genes, as follows (Fig. 14.30).

1. mRNA is isolated both from cells grown without cellulose (i.e., noninduced cells) and from cells grown in the presence of cellulose to enhance the synthesis of cellulase enzymes (i.e., induced cells).
2. Each mRNA population is fractionated on a sucrose gradient, and each fraction is translated in a cell-free system, either rabbit reticulocytes or wheat germ. A cellulose-induced protein(s) is identified, after separation of the cell-free translation products on a polyacrylamide gel, by the presence of unique bands that appear from the induced cells but not from the noninduced cells (Fig. 14.31). This step indicates which mRNA fractions contain messengers that are induced by the addition of cellulose.
3. The mRNA sucrose gradient fractions from cells that direct the synthesis of cellulose-induced proteins and the comparable fractions from the noninduced cells are used separately to program the synthesis of cDNA.
4. The cDNA sample from the induced-cell population is cloned into a plasmid vector, introduced into *E. coli*, replica plated, and then separately screened with labeled cDNA from both the induced and noninduced fractions as hybridization probes. Clones that hybridize only with the cDNA from the induced cells and not with cDNA from the noninduced cells potentially carry cellulose-induced genes and are characterized further.
5. To establish conclusively which of the positive cDNA clones encode cellulase enzymes, the DNA from these clones is introduced into an *E. coli* expression vector and, after *E. coli* is transformed with these constructs, the protein products are detected with antibodies to the enzymes of the cellulase complex.
6. The sequence of each of the positive cDNA clones is determined. In principle, this scheme can be used for the isolation of any induced eukaryotic gene(s).

Manipulation of Cellulase Genes

There are a variety of uses for cloned cellulase genes. In some cases, the cellulose-binding domain encoded by cellulase genes facilitates the purification of recombinant proteins. In other instances, the cellulolytic activity

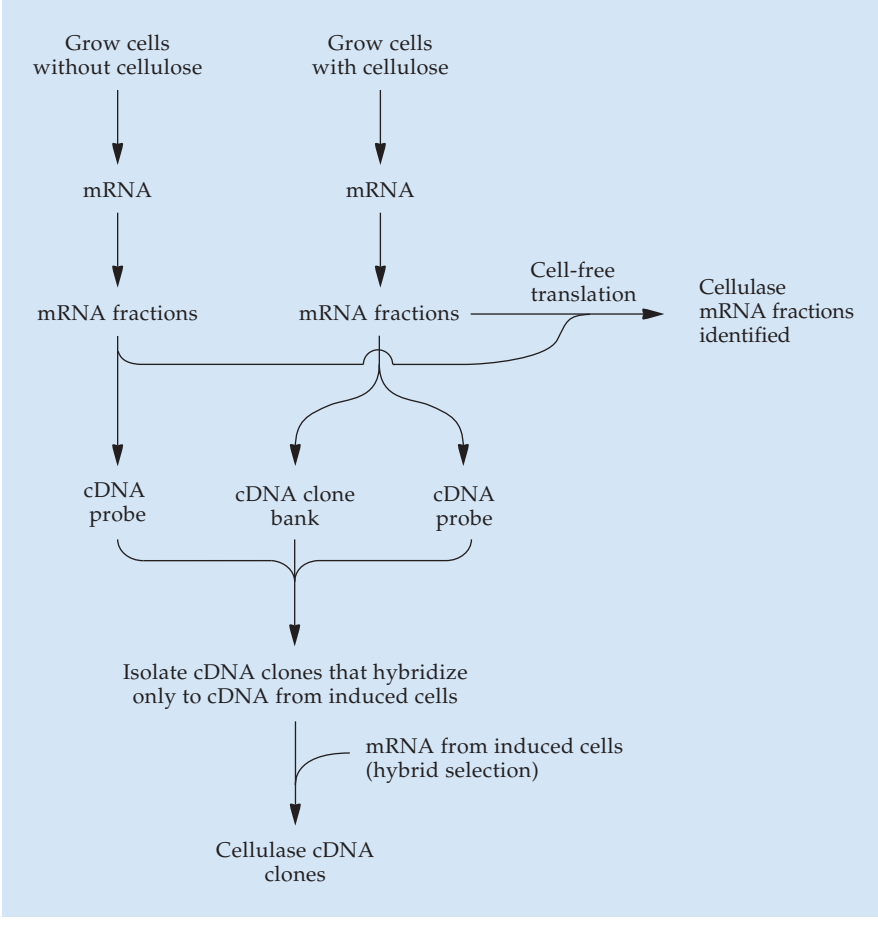


FIGURE 14.30 Differential hybridization for isolating eukaryotic cellulase cDNA clones.

is expressed in organisms that can convert waste cellulose into a commercial product, such as alcohol.

Many cellulase enzymes have three separate domains: a catalytic region; a hinge region that is often rich in proline, serine, or threonine residues; and a cellulose-binding region. The catalytic and binding domains act independently. This separation of functions can be exploited by cloning the DNA sequence that encodes the cellulose-binding domain as part of a fusion gene, where the other portion of the gene encodes a commercial protein. After expression of this fusion protein, it can be purified by passing a crude extract through a column packed with cellulose. Under these conditions, only the fusion protein will bind to the cellulose. Then the fusion protein, in homogeneous form, can be eluted from the column. The commercial protein can be retrieved by removal of the cellulose-binding domain by proteolytic cleavage. This system is similar in principle to immunoaffinity chromatography, except that it should be less expensive than using antibodies.

For convenience, most cellulase genes are initially cloned and expressed in *E. coli,* but other useful microorganisms might be developed by the introduction of cellulase genes. For example, *S. cerevisiae* and *Zymomonas mobilis,* which both efficiently convert simple sugars, such as glucose, into alcohol, have been used as hosts for the expression of cellulase genes. The idea

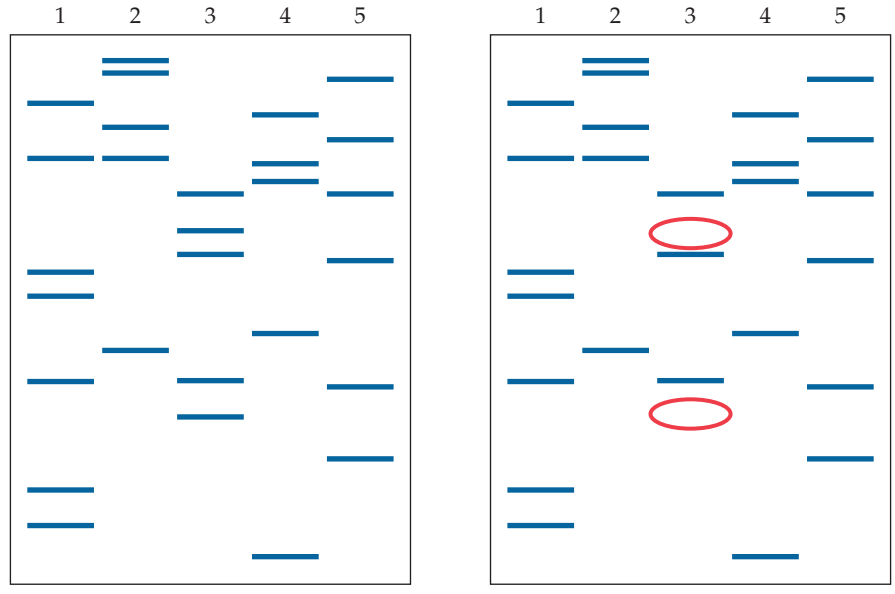


FIGURE 14.31 Schematic representation of stained polyacrylamide gels of translation products of mRNA fractions, i.e., proteins, following growth of cells either with or without cellulose. The numbers represent the fraction number following sucrose gradient centrifugation of mRNA. The red ovals on the gel without cellulose indicate where bands appear when there is growth in the presence of cellulose.

behind these studies was to test whether the presence of cellulase activity would enable these organisms to convert cellulose directly to alcohol.

In one series of experiments, endoglucanase and exoglucanase genes from the bacterium *Cellulomonas fimi* were each put under the control of an *S. cerevisiae* promoter and signal peptide sequence, subcloned onto the same plasmid vector, and introduced into *S. cerevisiae*. Some transformants secreted about 70% of each activity into the growth medium and were able to degrade the cellulose in filter paper and pretreated wood chips. The rate and extent of hydrolysis of both of these substrates were increased by adding β-glucosidase to the mixture, thereby decreasing the amount of cellobiose that accumulated and preventing end-product (feedback) inhibition of the endoglucanase and exoglucanase activities by cellobiose (Fig. 14.32). The role of β-glucosidase in the cellulolytic process has been examined in more detail. Cellobiose acts as a feedback inhibitor of cellulose hydrolysis, and glucose inhibits cellobiose cleavage. These two regulatory mechanisms may prevent complete enzymatic breakdown of cellulose. Instead of adding β-glucosidase to the medium, a β-glucosidase gene could be cloned into the host cell. To this end, a β-glucosidase gene was isolated from the cellulolytic fungus *Trichoderma reesei*, cloned onto a multiple-copy plasmid, and reintroduced into *T. reesei*. The transformant strain overproduced β-glucosidase activity 5.5-fold, and it degraded microcrystalline cellulose (Avicel), a cellulose derivative, 33% faster than the nontransformed strain. In addition, when a β-glucosidase gene from the yeast *Saccharomycopsis fibuligera* was expressed in *S. cerevisiae*, the transformed strain directed the enzyme to the periplasm. The β-glucosidase-producing *S. cerevisiae* strain was nearly as efficient at utilizing cellobiose to produce ethanol as the nontransformed *S. cerevisiae* strain was at producing alcohol from glucose. Thus, the presence of β-glucosidase genes enhances the enzy-

matic utilization of cellulose and suggests a simple strategy for genetically engineering more effective cellulolytic alcohol-producing microorganisms.

In addition to converting cellulosic wastes into useful materials, endoglucanase genes may have some novel uses. For example, when a wine-making yeast was transformed with an endoglucanase gene under the control of the constitutively expressed yeast actin gene promoter, the wine that was produced had an increased fruity aroma. This improvement was attributed to an increase in the amounts of at least 12 different volatile compounds, including ethyl propionate, 2-butanol, isoamyl acetate, isoamyl alcohol, and isobutyric acid. This type of genetic modification opens the possibility of engineering yeast strains that yield wines with particular desirable characteristics.

The participation of cellulase enzymes in an industrial process for the bioconversion of wastepaper to alcohol has also been examined. Wastepaper was partially digested by the addition of cellulase enzymes at 45°C; then, the released glucose was fermented by *S. cerevisiae* at 37°C. By extrapolation of small-scale results, yields of 400 liters of ethanol per ton of wastepaper were estimated. If all 100 million tons of wastepaper generated annually in North America were converted into ethanol and used as fuel, approximately 16% of the gasoline that is currently being used in North America could be saved.

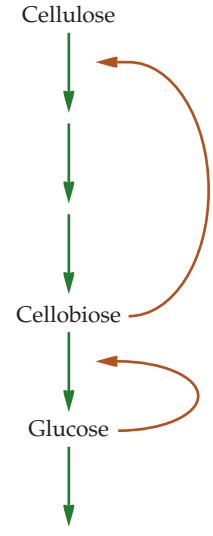


FIGURE 14.32 Metabolic control of cellulose hydrolysis. The green arrows indicate cellulose degradation. The red arrows indicate feedback inhibition by a particular metabolite of cellulose degradation. Relief of feedback inhibition may come from overproducing β-glucosidase, thereby decreasing the concentration of cellobiose, and removal of glucose by converting it into other metabolites.

Designer cellulosomes. Scientists hoping to exploit the activity of microbial cellulosomes to degrade cellulosic waste materials have attempted to manipulate some of the genes involved in the formation of this complex and create designer cellulosomes whose degradative activities are directed toward specific substrates. One of the key ingredients in the assembly and functioning of a cellulosome is the calcium-dependent high affinity (~10^9 M^{-1}) of the cohesin domain of the scaffoldin molecule for the dockerin domain (Fig. 14.26). Unfortunately, within a given species of microorganism, the dockerin domain binds to all of the cohesin domains with the same affinity. Thus, in nature, different catalytic domains attached to the dockerins are randomly incorporated into the cellulosome. However, it is possible, in the laboratory, to engineer cellulosomes that contain certain enzymatic activities designed to facilitate the degradation of specific substrates. In one series of experiments, designer cellulosomes were constructed that were more effective than the free enzymes (i.e., the enzymatic components not assembled into a cellulosome complex) and slightly less effective than native cellulosomes in degrading crystalline cellulose. However, these designer cellulosomes were not especially active at degrading straw (an agricultural waste product). When a xylanase gene was incorporated into the cellulosome complex, the ability of the designer cellulosome to degrade straw, which contains hemicellulose as well as cellulose, increased significantly (Fig. 14.33). Although the designer cellulosomes that have so far been constructed are not yet ready to be used in a commercial process, it is envisioned that these molecules could one day be the central component of a very large industry aimed at efficiently and economically converting cellulosic wastes into useable chemicals.

Zymomonas mobilis

Although industrial fermentations that produce alcohol are performed almost exclusively with *S. cerevisiae*, the bacterium *Z. mobilis* is a potentially useful organism for this purpose. *Zymomonas* is a gram-negative, rod-

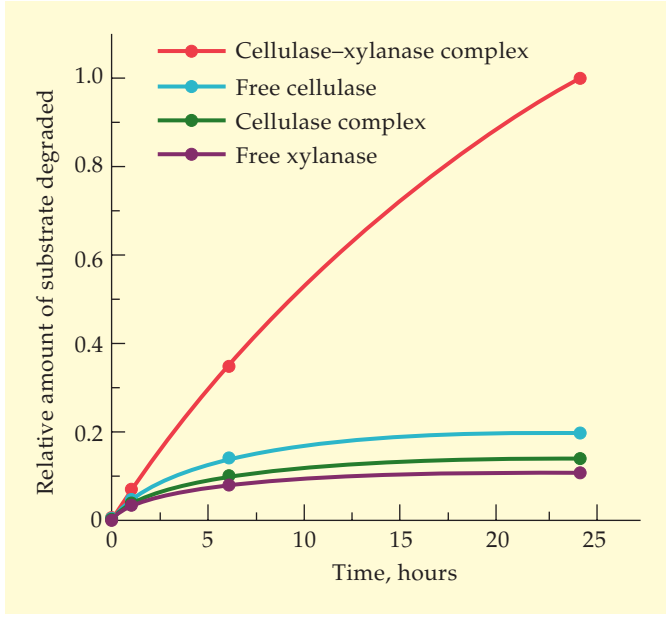


FIGURE 14.33 Digestion of straw by a cellulase-plus-xylanase designer cellulosome complex free cellulase enzyme, a cellulase-based designer cellulosome, and free xylanase enzyme. Although the data are presented as the relative amount of substrate degraded, in this experiment, the substrate was only partially degraded in all cases.

shaped organism that can ferment glucose, fructose, and sucrose and produce a relatively high yield of alcohol (Table 14.8). This high yield of ethanol is probably related to the fact that *Zymomonas* does not proliferate extensively (i.e., produce biomass) during fermentation. Since *Zymomonas* uses less substrate for biomass formation, more is available for ethanol production. In this regard, yeast produces 2 mol of ATP per mol of glucose, whereas *Zymomonas* uses a different pathway and produces only 1 mol of ATP per mol of glucose. Historically, *Zymomonas* has been used in tropical regions as a fermentative agent for the production of alcoholic beverages.

Zymomonas produces alcohol at a much higher rate than does *S. cerevisiae*, even though the organisms are similar in other features (Table 14.8).

TABLE 14.8 Comparison of *Z. mobilis* and *S. cerevisiae* as alcohol producers

Attribute	Value for:	
	Z. mobilis	Yeast
Conversion of sugar to ethanol (%)	96	96
Maximum ethanol concentration (%)	12	12
Ethanol productivity rate (g g^{-1} h^{-1})	5.67	0.67
Volumetric ethanol productivity rate (g $liter^{-1}$ h^{-1})	200	29
Sugar tolerance (%)	>40	>40
pH range for ethanol production	3.5–7.5	2–6.5
Optimum temperature (°C)	25–30	30–38

Adapted from Buchholz et al., *Trends Biotechnol* **5:**199–204, 1987.

The ethanol productivity rate was measured under batch fermentation conditions. The volumetric ethanol productivity rate was measured during continuous culture. Both strains yielded the same maximum ethanol concentration (12%) and had the same sugar tolerance (>40%).

However, there are biological and technical constraints that prevent *Zymomonas* from being used more widely for alcohol production. First, it can use only a limited number of carbon substrates for the production of alcohol. Second, broad-host-range cloning vectors and, as a consequence, foreign genes are difficult to maintain in this organism. Third, *Zymomonas* is naturally resistant to many of the more commonly used antibiotics, which precludes using the standard antibiotic resistance marker systems for cloning experiments.

Despite these difficulties, a number of foreign genes have been successfully introduced into and expressed in *Zymomonas*. Many of these experiments have focused on expanding the range of substrates that *Zymomonas* can utilize. For example, genes encoding enzymes that hydrolyze lactose, starch, cellulose, xylose, and cellobiose have all been introduced into *Zymomonas* (Table 14.9). Transformants were able to express all of these genes to some extent. However, in most of these cases, the transformed bacterium was unable to utilize the novel substrate as the sole carbon source.

In early studies directed toward developing strains of *Z. mobilis* that were capable of growth and ethanol production with xylose as a substrate, the bacterium was transformed with genes encoding the xylose utilization enzymes glucose/xylose isomerase and xylulokinase. However, these transformants were limited by their inability to further metabolize the pentoses (xylulose-5-phosphate, ribulose-5-phosphate, and ribose-5-phosphate) that are formed after xylose is assimilated (Fig. 14.34). To remedy this situation, *Z. mobilis* was transformed with a plasmid carrying two synthetic operons, one with two xylose assimilation genes and one with two pentose metabolism genes (Fig. 14.35). The pentose metabolism genes

TABLE 14.9 Some of the heterologous genes expressed in *Z. mobilis*

Enzyme encoded	Enzyme function
α-Amylase	Breakdown of starch to dextrins and glucose
Endo-1,4-β-D-glucanase	Breakdown of cellulose chains
β-D-Glucosidase	Breakdown of cellobiose to glucose
CMC	Breakdown of soluble cellulose
α-D-Galactosidase	Breakdown of raffinose, stachyose, and verbascose into glucose, galactose, sucrose, and fructose
Lac permease	Facilitates transport of lactose into bacterial cells
β-D-Galactosidase	Breakdown of lactose
Glucoamylase	Breakdown of starch and dextrins to glucose
Xylose isomerase	Conversion of xylose to xylulose
Xyulokinase	Conversion of L-xylulose to L-xylulose 5-phosphate
Xylose permease	Facilitates transport of xylose into bacterial cells
Transaldolase	Conversion of sedoheptulose 7-phosphate and D-glyceraldehyde 3-phosphate to yield D-erythrose 4-phosphate and D-fructose phosphate
Phosphomannose isomerase	Conversion of D-mannose 6-phosphate to D-fructose 6-phosphate
L-Arabinose isomerase	Conversion of L-arabinose to L-ribulose
L-Ribulokinase	Conversion of L-ribulose to L-ribulose 5-phosphate
L-Ribulose-phosphate-4-epimerase	Conversion of L-ribulose 5-phosphate to D-xylose 5-phosphate
Transketolase	Conversion of D-xylulose-5-phosphate (1) to sedoheptulose-7-phosphate and glyceraldehyde-3-phosphate and (2) to fructose-6-phosphate and glyceraldehyde-3-phosphate

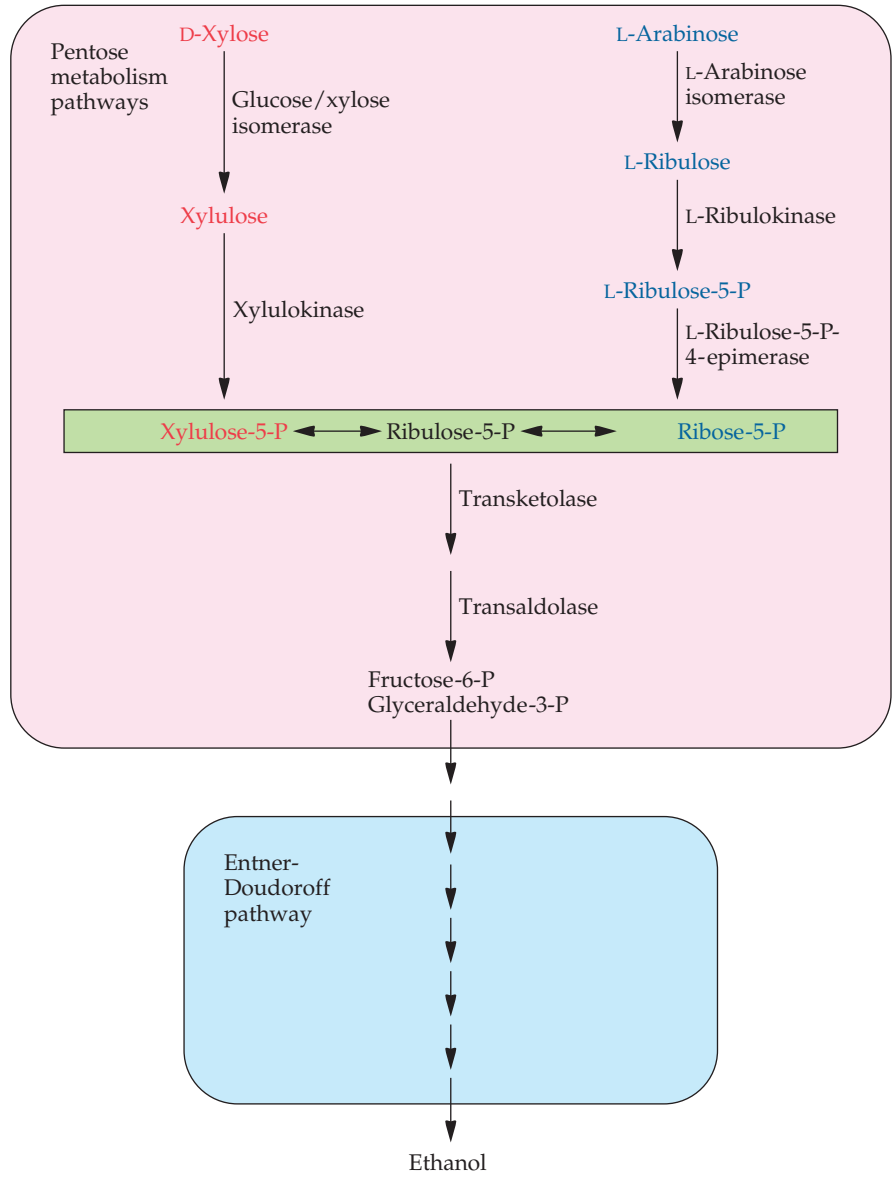


FIGURE 14.34 Schematic representation of the engineered assimilation of either xylose or arabinose by *Z. mobilis* and the engineered conversion of the resultant pentoses (shown in the green box) to ethanol.

encode the enzymes transketolase and transaldolase; both were placed under the control of the *Z. mobilis* enolase promoter. The xylose assimilation genes were placed under the transcriptional control of a strong constitutive promoter from the *Z. mobilis* gene for glyceraldehyde-3-phosphate dehydrogenase. Both constructs were cloned onto an *E. coli–Z. mobilis* shuttle vector, which was then used to transform *Z. mobilis*. As expected, the transformants assimilated xylose and converted the resulting pentoses that formed to fructose-6-phosphate and glyceraldehyde-3-phosphate, which, in turn, were readily converted to ethanol by the Entner-Doudoroff pathway of *Z. mobilis*. Moreover, the transformants grew efficiently on either glucose or xylose, as well as on glucose–xylose mixtures, and con-

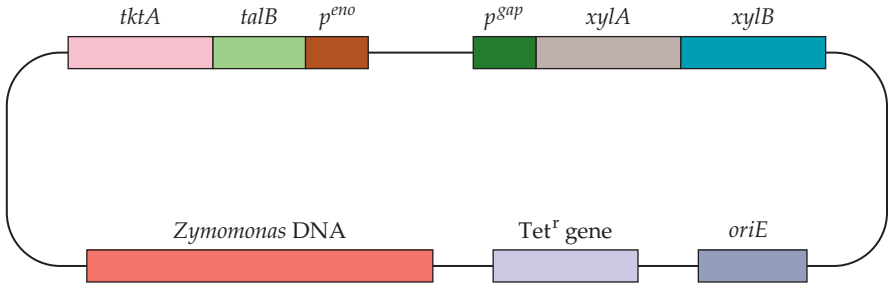


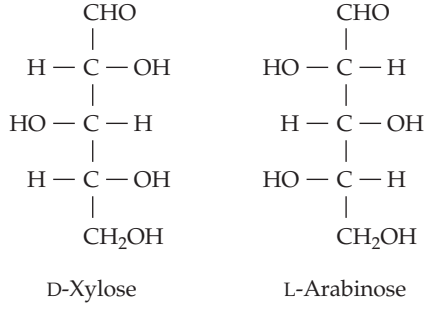
FIGURE 14.35 A *Zymomonas–E. coli* shuttle vector carrying one operon with genes encoding enzymes used for xylose assimilation *(xylA* and *xylB)* and another with genes encoding enzymes involved in pentose metabolism *(tktA* and *talB)*. p^{eno}, enolase promoter; p^{gap}, glyceraldehyde-3-phosphate dehydrogenase promoter; *xylA*, xylose isomerase gene; *xylB*, xylulokinase gene; *tktA*, transketolase gene; *talB*, transaldolase gene; Tet^r, tetracycline resistance gene; *oriE*, *E. coli* origin of replication. The *Zymomonas* DNA contains a *Zymomonas* origin of replication.

verted xylose to ethanol at high yield. This work demonstrates the feasibility of metabolically engineering *Z. mobilis* as an ethanol producer by using xylose, a waste material produced as a by-product of industrial processes, such as pulp and paper making, as a carbon source.

Xylose is the predominant pentose sugar in hardwoods, while arabinose (Fig. 14.36) is present in large amounts in various agricultural and other herbaceous plants. Some arabinose-containing plants, such as switchgrass, have been considered for use as dedicated energy crops, i.e., plants grown solely for use as sources of energy. Thus, it would be very useful if, in addition to a strain of *Z. mobilis* that can convert xylose to ethanol, an arabinose-fermenting *Z. mobilis* strain were also available. To develop such a strain, the arabinose assimilation genes L-ribulokinase, L-arabinose isomerase, and L-ribulose-5-phosphate-4-epimerase from *E. coli* were isolated and put under the transcriptional control of the constitutive *Z. mobilis* glyceraldehyde-3-phosphate dehydrogenase promoter (Fig. 14.37). Following expression of these genes, transformants produced the pentoses xylulose-5-phosphate, ribulose-5-phosphate, and ribose-5-phosphate. The plasmid used to transform *Z. mobilis* also contained two pentose metabolism genes encoding the enzymes transketolase and transaldolase under the transcriptional control of the constitutive *Z. mobilis* enolase promoter. The expression of these two genes catalyzed the conversion of the above-mentioned pentoses to fructose-6-phosphate and glyceraldehyde-3-phosphate.

The strategy that was used for this work was nearly identical to the strategy that was employed in the development of *Z. mobilis* strains able to utilize xylose as a carbon source. However, for *Z. mobilis* to utilize arabinose as a carbon source, arabinose rather than xylose assimilation genes were used. The metabolites that are produced starting with either xylose or arabinose are converted to ethanol by the Entner-Doudoroff pathway of *Z. mobilis* (Fig. 14.34). Moreover, the xylose- and arabinose-fermenting strains of *Z. mobilis* might be used together in a mixed bacterial culture for the conversion of the major sugars from certain agricultural residues into ethanol. To efficiently and economically convert lignocellulosic wastes to ethanol, it is necessary to convert all of the sugars, the pentoses as well as the hexoses, to ethanol. *Z. mobilis* lacks the pathways needed for the metabolism of mannose and galactose, which constitute a significant fraction of the

FIGURE 14.36 Chemical structures of D-xylose and L-arabinose.



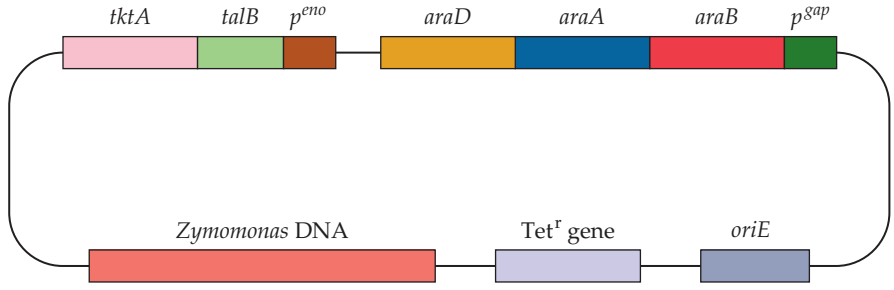


FIGURE 14.37 A *Zymomonas–E. coli* shuttle vector carrying one operon with genes encoding enzymes used for arabinose assimilation *(araB, araA,* and *araD)* and another with genes encoding enzymes involved in pentose metabolism *(tktA* and *talB). p^{eno}*, enolase promoter; *p^{gap}*, glyceraldehyde-3-phosphate dehydrogenase promoter; *araB*, L-ribulokinase gene; *araA*, L-arabinose isomerase gene; *araD*, L-ribulose-5-phosphate-4-epimerase gene; *tktA*, transketolase gene; *talB*, transaldolase gene; Tet^r, tetracycline resistance gene; *oriE*, *E. coli* origin of replication. The *Zymomonas* DNA contains a *Zymomonas* origin of replication.

hexoses in lignocellulosic material. Whether *Z. mobilis* can be engineered to be the organism of choice as part of a process of this sort is still an open question.

In addition to genetically engineering *Z. mobilis* to efficiently utilize xylose, arabinose, and glucose, researchers have utilized similar genetic approaches in an attempt to modify other microorganisms to utilize sugars derived from lignocellulosic materials to produce ethanol. Most of these efforts have been directed toward modifying *S. cerevisiae*, with some researchers attempting to engineer some strains of *P. putida*.

Instead of genetically engineering *Z. mobilis* to utilize xylose and arabinose and convert them into ethanol, some scientists have engineered *E. coli* to express some *Z. mobilis* genes so that it can produce ethanol. In this case, xylose and arabinose are converted to pyruvate by endogenous *E. coli* enzymes. Workers have mutated different *E. coli* genes to prevent the pyruvate from being converted into other, unwanted metabolites. Instead, the pyruvate is converted to ethanol by the enzymes pyruvate decarboxylase and alcohol dehydrogenase, with the genes for both of these enzymes coming from *Z. mobilis*. This sort of genetic modification has also been used to change the bacterium *Klebsiella oxytoca* into an ethanologenic organism. With both *E. coli* and *K. oxytoca*, the introduction of *Z. mobilis* genes yielded recombinant bacteria that were quite efficient in the laboratory at converting various sugars, both pentoses and hexoses, into alcohol. It now remains to be demonstrated whether any of these bacteria are effective on a large scale with an industrial substrate.

Since several naturally occurring yeast strains can utilize the range of sugars found in lignocellulosic materials, considerable effort has been directed to improving the performance of these strains. However, in contrast to the well-studied laboratory yeast strains, industrial strains are usually diploid or polyploid (and therefore not as easy to engineer), and their genetics are not especially well characterized or understood. Nevertheless, some industrial yeast strains that are able to tolerate the inhibitory compounds found in hydrolyzed lignocellulosic materials and ferment both hexoses and xyloses are being developed, and some researchers believe that within the next few years one or more of these organisms could

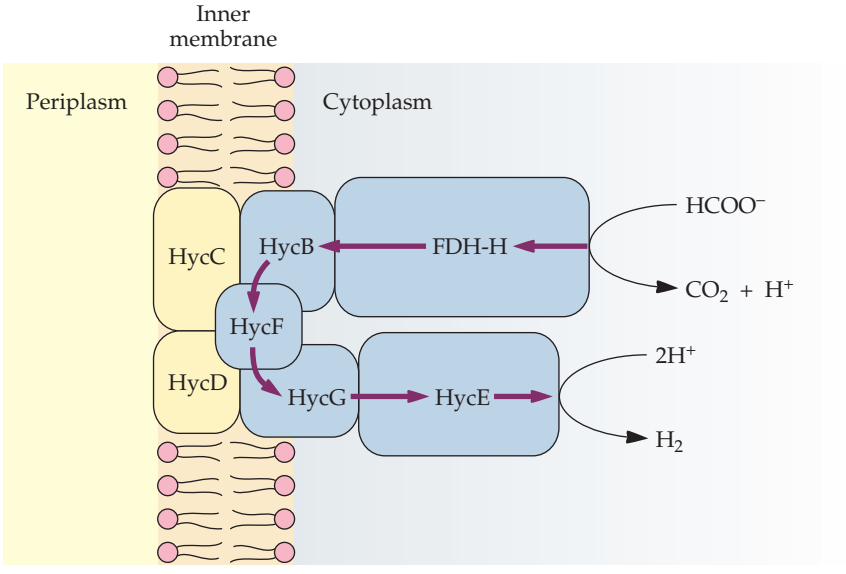


FIGURE 14.38 The major components of the *E. coli* formate hydrogen lyase. The purple arrows show the direction of the electron flux and the proteins through which the electrons flow (shown in blue). The Hyc proteins are the components of the formate hydrogen lyase system, which accept electrons from FDH-H (the large subunit of formate dehydrogenase). HycB is the small subunit of formate dehydrogenase and acts as the membrane docking target of FDH-H. HycC and HycD (shown in yellow) are integral membrane proteins and are thought to anchor the rest of the complex. HycE is the large subunit of hydrogenase 3, and HycG is the small subunit of hydrogenase 3. HycF is an electron transport protein.

become the cornerstone of an industrial process to convert lignocellulosic residues into ethanol.

Hydrogen Production

It has been known for some time that formic acid (formate) can be produced inexpensively, often as a by-product of the synthesis of other chemicals, such as acetic acid. Moreover, some types of bacteria are able to convert formate into hydrogen and carbon dioxide via the formate hydrogen lyase system. If the bacterial system that is responsible for hydrogen synthesis could be optimized, it might be possible to develop a practical system for the synthesis of hydrogen from biomass.

The *E. coli* formate hydrogen lyase system consists of a large number of different proteins, including those shown in Fig. 14.38, as well as others that specifically regulate the synthesis and maturation of these proteins. To overproduce the *E. coli* formate hydrogen lyase system and hence the amount of hydrogen, (1) the formate hydrogen lyase repressor gene, *hycA*, was inactivated and (2) the formate hydrogen lyase activator gene, *fhlA*, was overexpressed. These manipulations resulted in the large subunit of formate dehydrogenase (FDH-H) and the large subunit of hydrogenase 3 (HycE) being overexpressed 6.5- and 7.0-fold, respectively, compared to the wild type. These changes resulted in a nearly threefold increase in hydrogen productivity compared to the wild type. Additional enhancement of the amount of hydrogen produced was obtained by employing the engineered *E. coli* cells under anaerobic conditions in a bioreactor at a very high density

(96 g [dry weight] per liter). When the formate concentration was maintained below 25 mM, continuous hydrogen synthesis of 23.6 g of hydrogen per hour per liter was realized. This level of hydrogen production is sufficient for this system to be considered to have significant potential for commercial application.

SUMMARY

Bioremediation is the term that is applied to the use of microorganisms to clear the environment of contaminating substances. Many members of the bacterial genus *Pseudomonas* carry plasmids that encode enzymes capable of degrading aromatic and halogenated organic compounds. In most cases, a single plasmid carries the genes encoding enzymes for a specific degradative pathway. By combining plasmids from different pseudomonad strains within a single host, it is possible to create an organism with multiple degradation capabilities. In addition, by genetic manipulation, the range of substrates degraded by a particular enzymatic pathway can be extended.

Raw biological material is called biomass and is often used as a starting material in industrial processes. The use of milled grain for the production of alcohol or fructose requires a number of enzymatic steps. The enzymes that are used in these processes are often used only once and then discarded. To enhance enzymatic conversions and decrease costs, bacterial genes encoding enzymes that are thermostable, highly efficient catalytically, or tolerant of alcohol have been cloned, characterized, and tested.

To improve the commercial production of alcohol, some workers have genetically transformed the bacterium *Z. mobilis* with genes that allow it to utilize a broad range of compounds as carbon sources. Enzymes that degrade starch can also be used to facilitate the ability of microorganisms, such as *L. plantarum,* to ferment silage.

Often, as a consequence of processing biological material, large amounts of lignocellulose remain. This material generally has been treated as a waste product. However, there is now interest in using lignocellulose as a resource for carbon-containing compounds, especially glucose, that can be used in other processes. Retrieving glucose from lignocellulose is not an easy matter. Lignocellulose is a complex of lignin, hemicellulose, and cellulose; without harsh and expensive pretreatment, it is refractory to enzymatic degradation. Recent research has focused on characterizing the mechanism of breakdown of cellulose to glucose. The genes for endoglucanases, exoglucanases, and β-glucosidases from a variety of organisms have been cloned and characterized, but to date there has been little success in formulating a set of enzymes that efficiently degrades cellulose in vitro on a large scale.

REFERENCES

Alper, H., J. Moxley, E. Nevoigt, G. R. Fink, and G. Stephanopoulos. 2006. Engineering yeast transcription machinery for improved ethanol tolerance and production. *Science* **314:**1565–1568.

Barnett, C. C. 1991. Cloning and amplification of the gene encoding an extracellular β-glucosidase from *Trichoderma reesei*: evidence for improved rates of saccharification of cellulosic substrates. *Bio/Technology* **9:**562–567.

Bayer, E. A., L. J. W. Shimon, Y. Shoham, and R. Lamed. 1998. Cellulosomes—structure and ultrastructure. *J. Struct. Biol.* **124:**221–234.

Béguin, P. 1990. Molecular biology of cellulose degradation. *Annu. Rev. Microbiol.* **44:**219–248.

Brown, D. E. 1983. Lignocellulose hydrolysis. *Phil. Trans. R. Soc. Lond. B* **300:**305–322.

Buchholz, S. E., M. M. Dooley, and D. E. Eveleigh. 1987. *Zymomonas*—an alcoholic enigma. *Trends Biotechnol.* **5:**199–204.

Buchholz, S. E., and D. E. Eveleigh. 1990. Genetic modification of *Zymomonas mobilis*. *Biotechnol. Adv.* **8:**547–581.

Chakrabarty, A. M. March 1981. Microorganisms having multiple compatible degradative energy-generating plasmids and preparation thereof. U.S. patent 4,259,444.

Chen, W., and A. Mulchandani. 1998. The use of live biocatalysts for pesticide detoxification. *Trends Biotechnol.* **16:**71–76.

Cole, G. E., P. C. McCabe, D. Inlow, D. H. Gelfand, A. Ben-Bassat, and M. A. Innis. 1988. Stable expression of *Aspergillus awamori* glucoamylase in distiller's yeast. *Bio/Technology* **6:**417–421.

Cork, D. J., and J. P. Krueger. 1991. Microbial transformation of herbicides and pesticides. *Adv. Appl. Microbiol.* **36:**1–66.

Deanda, K., M. Zhang, C. Eddy, and S. Picataggio. 1996. Developing of an arabinose-fermenting *Zymomonas mobilis* strain by metabolic pathway engineering. *Appl. Environ. Microbiol.* **62:**4465–4470.

Dekker, K., A. Sugiura, H. Yamagata, K. Sakaguchi, and S. Udaka. 1992. Efficient production of thermostable *Thermus thermophilus* xylose isomerase in *Escherichia coli* and *Bacillus brevis*. *Appl. Microbiol. Biotechnol.* **36:**727–732.

Fierobe, H.-P., F. Mingardon, A. Mechaly, A. Bélaïch, M. T. Rinconi, S. Pagès, R. Lamed, C. Tardif, J.-P. Bélaïch, and E. A. Bayer. 2005. Action of designer cellulosomes on homogeneous versus complex substrates. Controlled incorporation of three distinct enzymes into a defined trifunc-

tional scaffoldin. *J. Biol. Chem.* **280:**16325–16334.

Fitzsimons, A., P. Hols, J. Jore, R. J. Leer, M. O'Connell, and J. Delcour. 1994. Development of an amylolytic *Lactobacillus plantarum* silage strain expressing the *Lactobacillus amylovorus* α-amylase gene. *Appl. Environ. Microbiol.* **60:**3529–3535.

Furukawa, K., A. Nishi, T. Watanabe, A. Suyama, and N. Kimura. 1998. Engineering microorganisms capable of efficient degradation of chlorinated environmental pollutants. *Rev. Toxicol.* **2:**179–187.

Ghosal, D., I.-S. You, D. K. Chatterjee, and A. M. Chakrabarty. 1985. Microbial degradation of halogenated compounds. *Science* **228:**135–142.

Hahn-Hägerdal, B., M. Galbe, M. F. Gorwa-Grauslund, G. Lidén, and G. Zacchi. 2006. Bio-ethanol—the fuel of tomorrow from the residues of today. *Trends Biotechnol.* **24:**549–556.

Hahn-Hägerdal, B., K. Kurhumaa, C. Fonseca, I. Spencer-Martins, and M. F. Gorwa-Grauslund. 2007. Towards industrial pentose-fermenting yeast strains. *Appl. Microbiol. Biotechnol.* **74:**937–953.

Hanai, T., S. Atsumi, and J. C. Liao. 2007. Engineering synthetic pathway for isopropanol production in *Escherichia coli*. *Appl. Environ. Microbiol*. **73:**7814–7818.

Heux, S., R. Cachon, and S. Dequin. 2006. Cofactor engineering in *Saccharomyces cerevisiae*: expression of a H_2O-forming NADH oxidase and impact on redox metabolism. *Metab. Eng*. **8:**303–314.

Heux, S., J.-M. Sablayrolles, R. Cachon, and S. Dequin. 2006. Engineering a *Saccharomyces cerevisiae* wine yeast that exhibits reduced ethanol production during fermentation under controlled microoxygenation conditions. *Appl. Environ. Microbiol.* **72:**5822–5828.

Innis, M. A., M. J. Holland, P. C. McCabe, G. E. Cole, V. P. Wittman, R. Tal, K. W. K. Watt, D. H. Gelfand, J. P. Holland, and J. H. Meade. 1985. Expression, glycosylation and secretion of an *Aspergillus* glucoamylase by *Saccharomyces cerevisiae*. *Science* **228:**21–26.

Kallio, P., A. Palva, and I. Palva. 1987. Enhancement of α-amylase production by integrating and amplifying the α-amylase gene of *Bacillus amyloliquefaciens* in the genome of *Bacillus subtilis*. *Appl. Microbiol. Biotechnol.* **27:**64–71.

Kennedy, J. F., V. M. Cabalda, and C. A. White. 1988. Enzymic starch utilization and genetic engineering. *Trends Biotechnol.* **6:**184–189.

Knowles, J., P. Lehtovaara, and T. Teeri. 1987. Cellulase families and their genes. *Trends Biotechnol.* **5:**255–261.

Kolenc, R. J., W. E. Inniss, B. R. Glick, C. W. Robinson, and C. I. Mayfield. 1988. Transfer and expression of mesophilic plasmid-mediated degradative capacity in a psychrotrophic bacterium. *Appl. Environ. Microbiol.* **54:**638–641.

Kumar, V., S. Ramakrishnan, T. T. Teeri, J. K. C. Knowles, and B. S. Hartley. 1992. *Saccharomyces cerevisiae* cells secreting an *Aspergillus niger* β-galactosidase grow on whey permeate. *Bio/Technology* **10:**82–85.

Lamed, R., J. Naimark, E. Morgenstern, and E. A. Bayer. 1987. Specialized cell surface structures in cellulolytic bacteria. *J. Bacteriol.* **169:**3792–3800.

Lange, C. C., L. P. Wackett, K. W. Minton, and M. J. Daly. 1998. Engineering a recombinant *Deinococcus radiodurans* for organopollutant degradation in radioactive mixed waste environments. *Nat. Biotechnol.* **16:**929–933.

Latorre-Garcia, L., A. C. Adam, P. Manzanares, and J. Polaina. 2005. Improving the amylolytic activity of *Saccharomyces cerevisiae* glucoamylase by the addition of a starch binding domain. *J. Biotechnol.* **118:**167–176.

Leonhartsberger, S., I. Korsa, and A. Böck. 2002. The molecular biology of formate metabolism in enterobacteria. *J. Mol. Microbiol. Biotechnol.* **4:**269–276.

Leungsakul, T., G. R. Johnson, and T. K. Wood. 2006. Protein engineering of the 4-methyl-5-nitrocatechol monooxygenase from *Burkholderia* sp. strain DNT for enhanced degradation of nitroaromatics. *Appl. Environ. Microbiol*. **72:**3933–3939.

Lynd, L. R., J. H. Cushman, R. J. Nichols, and C. E. Wyman. 1991. Fuel ethanol from cellulosic biomass. *Science* **251:**1318–1323.

Meng, M., C. Lee, M. Bagdasarian, and J. G. Zeikus. 1991. Switching substrate preference of thermophilic xylose isomerase from D-xylose to D-glucose by redesigning the substrate binding pocket. *Proc. Natl. Acad. Sci. USA* **88:**4015–4019.

Mingardon, F., A. Chanal, A. M. López-Contreras, C. Dray, E. A. Bayer, and H.-P. Fierobe. 2007. Incorporation of fungal cellulases in bacterial minicellulosomes yields viable, synergistically acting cellulolytic complexes. *Appl. Environ. Microbiol.* **73:**3822–3832.

Ohta, K., D. S. Beall, J. P. Mejia, K. T. Shanmugam, and L. O. Ingram. 1991. Genetic improvement of *Escherichia coli* for ethanol production: chromosomal integration of *Zymomonas mobilis* genes encoding pyruvate decarboxylase and alcohol dehydrogenase II. *Appl. Environ. Microbiol.* **57:**893–900.

Panesar, P. S., S. S. Marwaha, and J. F. Kennedy. 2006. *Zymomonas mobilis*: an alternative ethanol producer. *J. Chem. Technol. Biotechnol.* **81:**623–635.

Pérez-González, J. A., R. González, A. Querol, J. Sendra, and D. Ramón. 1993. Construction of a recombinant wine yeast strain expressing β-(1,4)-endoglucanase and its use in microvinification processes. *Appl. Environ. Microbiol*. **59:**2801–2806.

Quax, W. J., N. T. Mrabet, R. G. M. Luiten, P. W. Schuurhuizen, P. Stanssens, and I. Lasters. 1991. Enhancing the thermostability of glucose isomerase by protein engineering. *Bio/Technology* **9:**738–742.

Ramos, J. L., A. Wasserfallen, K. Rose, and K. N. Timmis. 1987. Redesigning metabolic routes: manipulation of TOL plasmid pathway for catabolism of alkylbenzoates. *Science* **235:**593–596.

Richins, R. D., I. Kaneva, A. Mulchandani, and W. Chen. 1997. Biodegradation of organophosphorus pesticides by surface-expressed organophosphorus hydrolase. *Nat. Biotechnol.* **15:**984–987.

Suyama, A., R. Iwakiri, N. Kimura, A. Nishi, K. Nakamura, and K. Furukawa. 1996. Engineering hybrid pseudomonads capable of utilizing a wide range of aromatic hydrocarbons and of efficient degradation of trichloroethylene. *J. Bacteriol.* **178:**4039–4046.

Tao, H., R. Gonzalez, A. Martinez, M. Rodriguez, L. O. Ingram, J. F. Preston, and K. T. Shanmugam. 2001. Engineering a homo-ethanol pathway in *Escherichia coli*: increased glycolytic flux and levels of expression of glycolytic genes during xylose fermentation. *J. Bacteriol.* **183:**2979–2988.

Timmis, K. N., R. J. Steffan, and R. Unterman. 1994. Designing microorganisms for the treatment of toxic wastes. *Annu. Rev. Microbiol.* **48:**525–557.

Van Rooyen, R., B. Hahn-Hägerdal, D. C. La Grange, and W. H. Van Zyl. 2005. Construction of cellulose-growing and fermenting *Saccharomyces cerevisiae* strains. *J. Biotechnol*. **120**:284–295.

Verdoes, J. C., A. D. van Diepeningen, P. J. Punt, A. J. M. Debets, A. H. Stouthamer, and C. A. M. J. J. van den Hondel. 1994. Evaluation of molecular and genetic approaches to generate glucoamylase overproducing strains of *Aspergillus niger*. *J. Biotechnol.* **36:**165–175.

Wayman, M., S. Chen, and K. Doan. 1992. Bioconversion of waste paper to ethanol. *Process Biochem.* **27:**239–245.

Winter, R. B., K.-M. Yen, and B. D. Ensley. 1989. Efficient degradation of trichloroethylene by a recombinant *Escherichia coli*. *Bio/Technology* **7:**282–285.

Witholt, B., M.-J. de Smet, J. Kingma, J. B. van Beilen, M. Kok, R. G. Lageveen, and G. Eggink. 1990. Bioconversions of aliphatic compounds by *Pseudomonas oleovorans* in multiphase bioreactors: background and economic potential. *Trends Biotechnol.* **8:**46–52.

Wong, W. K. R., C. Curry, R. S. Parekh, S. R. Parekh, M. Wayman, R. W. Davies, D. G. Kilburn, and N. Skipper. 1988. Wood hydrolysis by *Cellulomonas fimi* endoglucanase and exoglucanase coexpressed as secreted enzymes in *Saccharomyces cerevisiae*. *Bio/Technology* **6:**713–719.

Yoshido, A., T. Nishimura, H. Kawaguchi, M. Inui, and H. Yukawa. 2005. Enhanced hydrogen production from formic acid by formate hydrogen lyase-overproducing *Escherichia coli* strains. *Appl. Environ. Microbiol.* **71:**6762–6768.

Zhang, M., C. Eddy, K. Deanda, M. Finkelstein, and S. Picataggio. 1995. Metabolic engineering of a pentose metabolism pathway in ethanologenic *Zymomonas mobilis*. *Science* **267:**240–243.

Zylstra, G. J., L. P. Wackett, and D. T. Gibson. 1989. Trichloroethylene degradation by *Escherichia coli* containing the cloned *Pseudomonas putida* F1 toluene dioxygenase genes. *Appl. Environ. Microbiol.* **55:**3162–3166.

REVIEW QUESTIONS

1. How would you genetically engineer a bacterium to degrade trichloroethylene?

2. Outline a protocol that you would use to clone fungal cellulase genes.

3. Delineate the role of α-amylase and glucoamylase in the industrial production of alcohol. How might genetic manipulation of the genes encoding these enzymes be used to improve this process?

4. What is glucose isomerase? Why is it important? How and why would you modify the gene encoding this enzyme?

5. Elaborate some of the advantages and disadvantages of using *Z. mobilis* instead of *S. cerevisiae* for alcohol production. How would you improve the industrial performance of *Z. mobilis*?

6. How can *Z. mobilis* be engineered to produce ethanol from xylose and arabinose?

7. Starting with a *Pseudomonas* strain that can utilize phenol as its sole carbon source at 0°C, a *Pseudomonas* strain that can degrade anthracene to catechol at 35°C, and a *Pseudomonas* strain that can degrade *p*-toluene to protocatechuate at 35°C, suggest a strategy for developing a strain that can utilize phenol, anthracene, or *p*-toluene as its sole carbon source at 0°C.

8. Explain how a *Pseudomonas* strain that carries plasmid pWWO and does not normally degrade 4-ethylbenzoate can be genetically manipulated to hydrolyze this compound.

9. Suggest schemes for the isolation of prokaryotic endoglucanase and β-glucosidase genes.

10. What is a "superbug"?

11. How can *L. plantarum* be manipulated to improve its ability to ferment silage?

12. How can pesticide-degrading enzymes be expressed on the surface of a bacterium?

13. How would you degrade organic environmental pollutants in the presence of high levels of radioactivity?

14. How would you engineer yeast strains to more efficiently convert glucose into ethanol?

15. How would you expand the substrate range of a strain of *Burkholderia* sp. that normally degrades 2,4-dinitrotoluene?

16. How would you engineer glucoamylase to be more efficient in degrading starch?

17. How would you engineer *E. coli* to produce isopropanol?

18. What is a designer cellulosome? How is it produced?

19. How would you engineer *E. coli* to produce hydrogen gas from formic acid?

15

Plant Growth-Promoting Bacteria

UNDER NATURAL ENVIRONMENTAL CONDITIONS, successful plant growth and development and high crop yields depend on the genetic constitution of the crop species, suitable weather conditions, and soil components, including the availability of nutrients; the absence of growth-inhibitory substances, such as salt; the presence of certain beneficial microorganisms; and the absence of pathogenic ones (called phytopathogens, from *phyto*, meaning plant). Some beneficial indigenous soil bacteria and fungi act directly by providing a plant growth-enhancing product, and others act indirectly. The latter organisms inhibit the growth of pathogenic soil microorganisms, thereby preventing them from hindering plant growth.

The direct promotion of plant growth usually entails providing the plant with a compound that is synthesized by the bacterium, such as fixed nitrogen or a plant hormone. Also, these bacteria can facilitate the uptake by the plant of certain nutrients from the environment. The indirect promotion of plant growth occurs when plant growth-promoting bacteria lessen or prevent the deleterious effects of phytopathogenic organisms, either fungi or bacteria, i.e., they act as biocontrol agents. This activity is called antibiosis, and it either depletes a scarce resource required by the pathogen or produces a compound that impedes the growth of the phytopathogenic organism.

Direct stimulation of plant growth and development by plant growth-promoting bacteria can occur in several different ways. The bacteria can (1) fix atmospheric nitrogen to ammonia that is used by the plant; (2) synthesize siderophores that solubilize and sequester iron from the soil and provide it to plant cells; (3) synthesize phytohormones, such as auxin, cytokinin, or gibberellin, that enhance various stages of plant growth; (4) solubilize minerals, such as phosphorus, that are used by the plant; and (5) synthesize an enzyme that can modulate the level of the plant hormone ethylene. Any particular plant growth-promoting bacterium may utilize one or more of these mechanisms.

Much of the recent genetic research directed at creating microbial strains with augmented plant growth-promoting activity has focused on a few areas of study.

- Engineering of better biocontrol strains of bacteria to decrease the damage to plants from a variety of pathogens. This work is aimed at replacing some of the chemical pesticides that may become environmental pollutants.
- The use of bacteria to lower ethylene levels in plants. These studies are directed toward preventing high levels of ethylene from accumulating in plants and thereby decreasing the damage to the plant from a variety of environmental stresses, including drought, flooding, salt stress, and the presence of pathogens.
- The molecular basis of nitrogen fixation. This topic has been investigated thoroughly to determine whether it is possible to increase the level of microbial nitrogen fixation and consequently lessen the current dependency on chemical fertilizers for crop plants.
- Root nodule formation by symbiotic bacteria. This process has been studied with the aim of producing genetically engineered bacteria that can outcompete naturally occurring symbiotic bacteria.
- Microbial synthesis of iron-sequestering compounds (siderophores). These reactions are being characterized in the hope that it might be possible to produce beneficial strains that prevent the growth of phytopathogenic microorganisms.
- Manipulation of plant growth-promoting bacteria to facilitate phytoremediation (the use of plants to remediate contaminated environments).

Current research in this area mainly deals with plant growth-promoting bacteria rather than fungi. This is at least partly due to the fact that scientists have found it difficult or even impossible to grow many beneficial fungi in culture, so not only is it difficult to manipulate them in the laboratory, it is also extremely difficult to obtain large enough amounts of these organisms for inoculation of crops. In the past, bacterial fertilization had a dubious reputation. During the 1950s in the Soviet Union, more than 10 million hectares (about 39,000 square miles) of farmland were treated with diazotrophic (nitrogen-fixing) bacterial mixtures that consisted primarily of *Azotobacter chroococcum* and *Bacillus megaterium*. In these experiments, about 60% of the time, yields of various crops were increased by 10 to 20%. However, these field trials were poorly designed and not replicable, so many researchers were skeptical about the validity of the work and tended to discount the use of bacterial inoculants as fertilizing agents on a large scale. In recent years, considerable progress has been made toward understanding many of the mechanisms employed by plant growth-promoting bacteria. Thus, there is a much greater likelihood than in the past that results will be predictable and reproducible.

Growth Promotion by Free-Living Bacteria

Plant growth-promoting bacteria include a wide range of bacteria that are free-living or that form a symbiotic relationship with plants, such as *Rhizobium* and *Frankia*. While numerous free-living soil bacteria are considered to be plant growth-promoting bacteria (Table 15.1), not all bacterial

TABLE 15.1 Examples of successful agricultural plant growth stimulation by free-living plant growth-promoting bacteria

Bacterium	Plant(s)	Conditions
Azospirillum brasilense	Guinea grass, millet, sorghum, bean, wheat, barley, fountain grass, Sudan grass, corn, chickpea, fava bean, oat, rice	Field, greenhouse, hydroponic system
Azospirillum irakense	Winter wheat, corn	Field
Azospirillum lipoferum	Millet, sunflower, corn	Field, greenhouse
Azospirillum sp.	Wheat, corn, millet, mustard, rice, sorghum	Field, greenhouse
Azotobacter chroococcum	Barley	Growth chamber
Bacillus amyloliquefaciens	Tomato, pepper	Field
Bacillus cereus	Tomato, pepper	Field
Bacillus polymyxa	Wheat, sugar beet	Field
Bacillus pumilis	Tomato, pepper	Field
Bacillus subtilis	Tomato, pepper, peanut, onion	Field, growth chamber
Bacillus sp.	Sorghum, wheat	Field
Burkholderia vietnamiensis	Rice	Field
Enterobacter cloacae	Tomato, pepper, mung bean	Greenhouse
Pseudomonas cepacia	Winter wheat	Field, growth chamber
Pseudomonas chlororaphis	Spring wheat	Field, laboratory
Pseudomonas fluorescens	Winter wheat, potato, tomato, cucumber, blueberry	Field, greenhouse, growth chamber
Pseudomonas putida	Winter wheat, potato, canola, cucumber, lettuce, tomato, barley, oat	Field, greenhouse, growth chamber
Pseudomonas syringae	Bean	Greenhouse
Pseudomonas sp.	Canola, potato, rice, lettuce, cucumber, tomato, corn	Field, greenhouse, growth chamber, hydroponic system

strains of a particular genus and species have identical metabolic capabilities. Thus, for example, some *Pseudomonas putida* strains actively promote plant growth, while others have no measurable effect on plants.

The major applications of bacteria for improving plant growth include agriculture, horticulture, forestry, and environmental restoration (phytoremediation). In the past 20 years or so, based on a better understanding of the mechanisms employed by these bacteria and following a large number of successful laboratory and field studies, an increasing number of plant growth-promoting bacteria have been commercialized.

The mechanism most commonly invoked to explain the various effects of plant growth-promoting bacteria on plants is the production of phytohormones. Research in this area has focused on the role of a class of phytohormones called auxins. The most common and best-characterized auxin is indole-3-acetic acid (IAA), which stimulates in plants both rapid responses, such as increases in cell elongation, and long-term effects, such as increases in cell division and differentiation. Since both plants and plant growth-promoting bacteria can synthesize auxin, it is difficult for researchers to distinguish between plant responses that result from bacterial auxin synthesis and those that result from plant auxin synthesis. This uncertainty notwithstanding, there is considerable evidence to suggest that many plant growth-promoting bacteria facilitate plant growth by altering the hormonal balance within a plant.

In the early 1990s, it was discovered that many plant growth-promoting bacteria contain an enzyme that can modulate levels of the plant hormone ethylene. This enzyme, 1-aminocyclopropane-1-carboxylate

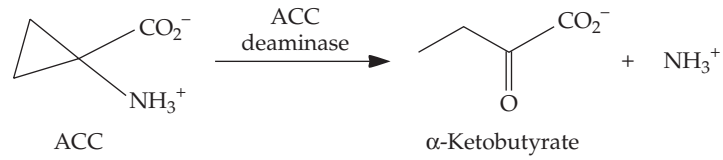
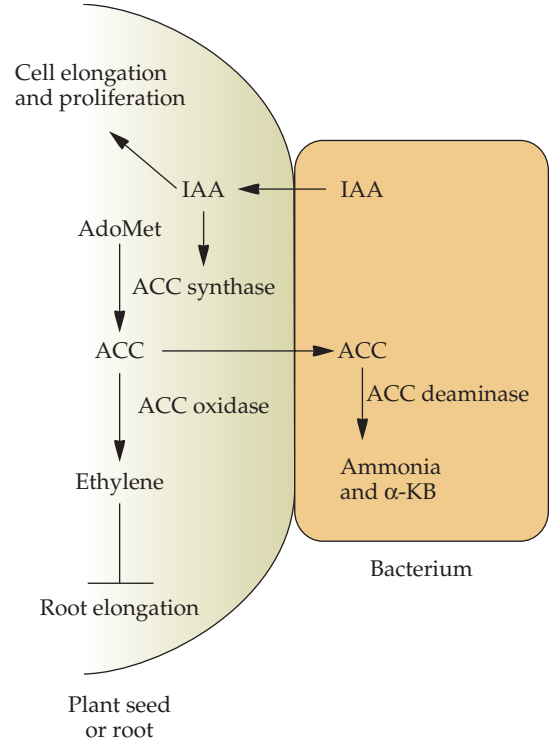


FIGURE 15.1 Cleavage of ACC to α-ketobutyrate and ammonia by ACC deaminase.

(ACC) deaminase, cleaves ACC, which is the immediate biosynthetic precursor of ethylene in plants (Fig. 15.1). As depicted in Fig. 15.2, the bacterium binds to seed coats or plant roots and then sequesters and cleaves ACC. As a result, the level of ethylene in the developing (or stressed) plant is lowered. In many plants, ethylene stimulates germination and breaks the

FIGURE 15.2 Schematic representation of the mechanisms by which an ACC deaminase-containing plant growth-promoting bacterium bound to either a seed or a plant root lowers the ethylene concentration and thereby prevents ethylene inhibition of root elongation. The arrows indicate chemical or physical steps in the mechanism, and the symbol ⊥ indicates inhibition of root elongation by ethylene. IAA is synthesized and secreted by a plant growth-promoting bacterium that is bound to the surface of either the seed or the root of a developing plant. After being taken up by the plant, together with the IAA from the plant, the bacterial IAA can stimulate either plant cell proliferation and elongation or the activity of the enzyme ACC synthase, which converts *S*-adenosylmethionine (AdoMet) to ACC. A significant portion of the ACC is exuded from plant roots or seeds, along with other small molecules normally present in seed or root exudates; taken up by the bacterium; and hydrolyzed by the enzyme ACC deaminase to ammonia and α-ketobutyrate (α-KB). This uptake and cleavage of ACC decreases the amount of ACC outside the plant. To maintain the equilibrium between internal and external ACC, the plant exudes more ACC. Consequently, the concentration of ACC, and therefore ethylene, in the plant is lowered. Adapted from Glick et al., *J. Theor. Biol.* **190:**63–68, 1998.



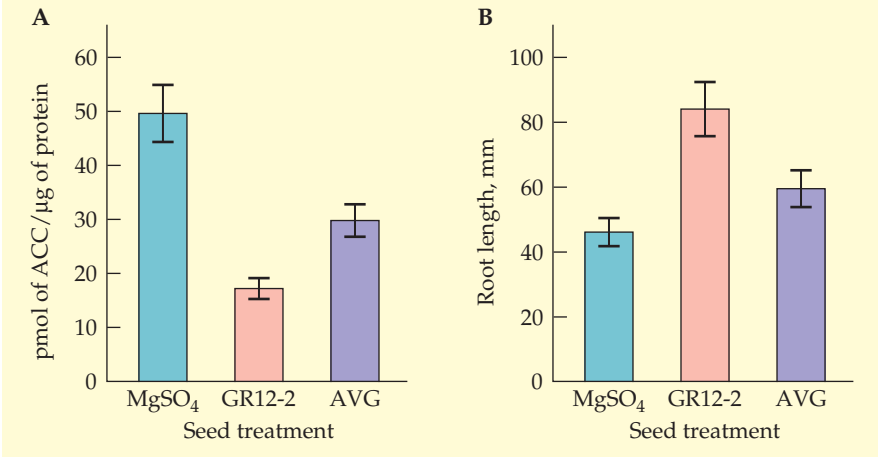
dormancy of the seeds; however, if the level of ethylene remains high after germination, a problem that is especially acute when plants are under stress, root elongation is inhibited. Thus, the ACC deaminase that is provided by a plant growth-promoting bacterium prevents the inhibition of root elongation (Fig. 15.3), and consequently, the plant produces longer roots during early development, resulting in a healthier and larger plant. In addition, many plant growth-promoting bacteria synthesize IAA. The IAA that is produced by the bacterium and taken up by the plant that is not used to promote plant cell elongation or proliferation stimulates the transcription of the enzyme ACC synthase in the plant. A greater amount of ACC synthase causes an increase in the level of ACC, eventually resulting in an increase in the ethylene concentration. When ACC deaminase activity is present, it prevents the buildup of ACC, even in the presence of high levels of IAA, so that the ethylene level does not become elevated to the point where plant growth is impaired.

In general, nitrogen fixation by free-living plant growth-promoting bacteria probably makes only a minor contribution to the growth of a plant. In fact, not all plant growth-promoting bacteria are diazotrophic, and many of those that are diazotrophic fix only limited amounts of nitrogen.

A number of plants use bacterial iron–siderophore complexes to obtain iron from the soil. Without this mechanism, plant growth in many soils would be severely limited, as iron is an essential plant nutrient. However, while bacterial siderophores undoubtedly contribute to the nutrition, and hence to the growth, of plants, in many instances this effect is small.

There is some controversy regarding the mechanism that plant growth-promoting bacteria use to facilitate the uptake of minerals such as phosphorus by a plant. On one hand, the increased mineral uptake in plants treated with plant growth-promoting bacteria may reflect a better-developed root system and an overall healthier plant. On the other hand, experiments with *Azospirillum* have shown that this organism enhances mineral uptake by secreting organic acids that can solubilize and bind some minerals.

FIGURE 15.3 Effect of treating canola seeds with ACC deaminase-containing plant growth-promoting bacteria on root ACC content **(A)** and root length **(B),** following growth of the plant for 4.5 days after the seeds were sown. The seeds were treated with $MgSO_4$ as a control, the ACC deaminase-containing bacterium *P. putida* GR12-2, or the chemical ethylene inhibitor 1-aminovinylglycine (AVG). The error bars indicate standard errors.



MILESTONE

A Model for the Lowering of Plant Ethylene Concentrations by Plant Growth-Promoting Bacteria

B. R. GLICK, D. M. PENROSE, and J. LI

J. Theor. Biol. **190:**63–68, 1998

In addition to the profound influence that ethylene has on normal plant growth and development, it is a stress hormone whose synthesis is increased when a plant is subjected to any one of a variety of environmental stresses. These stresses include mechanical trauma, pathogen infection, extremes of temperature, drought, flooding, salt, and the presence of environmental contaminants. Following periods of stress, the ethylene that is produced by the plant often exacerbates the effects of the stress. This can lead to plant senescence or death. Any chemical or biological treatment that lowers the amount of ethylene that is produced by a plant as a consequence of an environmental stress should therefore also decrease some of the resulting damage to the plant.

It has been known for many years that certain strains of bacteria can promote the growth of plants. At the time that this article was published, many of the mechanisms involved in the promotion of plant growth by bacteria had apparently been elucidated. However, there did not seem to be any one mechanism that could reliably and reproducibly promote the growth of a wide variety of plants under a range of different conditions. This was in spite of the fact that some strains of *Rhizobium*, some biocontrol strains, and some *Azospirillum* strains had been commercialized, albeit to a limited extent.

In this article, Glick and coworkers developed a conceptual framework to explain a number of empirical observations that their laboratory had reported beginning in 1994. In its simplest terms, the model that they elaborated suggested that some plant growth-promoting bacteria that were bound to plant tissues could act as a sink for some of the ACC that was produced by the plant in response to various types of stress. Since ACC is the immediate precursor of ethylene in all higher plants, the model predicted that lowering ACC levels before it could be converted to ethylene would limit some of the deleterious effects of a particular stressor on a plant. The model was tested in a growth chamber, then in a greenhouse, and eventually in field experiments. It was found that plant growth-promoting bacteria that contained active ACC deaminase could significantly decrease the inhibition of growth and damage to plants following exposure to either high salt levels, the presence of metals or organic contaminants, phytopathogens, flooding, or drought. Moreover, workers in many different laboratories around the world have found that this approach works well with a wide range of plants, including canola, tomato, lettuce, soybean, mung bean, Indian mustard, various grasses, wheat, pea, corn, and cotton. Thus, by either selecting or engineering plant growth-promoting bacteria to express ACC deaminase, the productivity of a range of crop plants can be improved dramatically.

As a better understanding of the mechanisms used by plant growth-promoting bacteria emerges, it will become possible to genetically engineer improved organisms that can stimulate the growth of a wide range of plants in a variety of environments.

Decreasing Plant Stress

In addition to its effect on seed germination and root elongation, ethylene mediates a wide range of plant responses and developmental steps. Ethylene is involved in tissue differentiation, formation of root and shoot primordia, lateral bud development, flowering initiation, anthocyanin synthesis, flower opening and senescence, fruit ripening and degreening, production of volatile organic compounds that are responsible for aroma formation in fruits, storage product hydrolysis, leaf and fruit abscission, and the response of plants to biotic and abiotic stress. In some processes, ethylene is stimulatory, while in others it is inhibitory.

The term "stress ethylene" describes the increase in ethylene biosynthesis associated with biological and environmental stresses and pathogen attack. The increased level of ethylene formed in response to trauma

inflicted by chemicals, temperature extremes, water stress, ultraviolet light, insect damage, disease, and mechanical wounding can be both the cause of some of the symptoms of stress (e.g., onset of wilting and increased senescence) and the inducer of responses that will enhance the survival of the plant under adverse conditions. Often, a small burst of ethylene is synthesized by plants within a few hours after an environmental stress. This low level of ethylene acts as a trigger to initiate the biosynthesis of a number of plant defense proteins. Subsequently, some 2 to 4 days after the onset of the stress, the plant produces a much larger burst of ethylene. It is this second peak of ethylene synthesis that is responsible for attenuating the deleterious effect(s) of the stress (Fig. 15.4).

While chemicals have been successfully used to control ethylene levels in plants, many of them are either expensive or potentially harmful to the environment. Consequently, ACC deaminase-containing plant growth-promoting bacteria have been tested to determine whether they could be used as an environmentally safe method for lowering plant ethylene levels.

Flooding is a common abiotic stress that affects many plants, often several times during the same growing season. Plant roots suffer a lack of oxygen as a consequence of flooding; this, in turn, causes deleterious effects, such as wilting (epinasty), inhibition of leaf chlorophyll synthesis (chlorosis), cell death (necrosis), and reduced fruit yield. Many plants respond to flooding by activating the transcription, in root cells, of some of the genes that code for isozymes of ACC synthase, the enzyme that converts the compound *S*-adenosylmethionine into ACC. This eventually results in an increase in the amount of ACC inside plant roots. However, since ACC oxidase cannot catalyze ethylene synthesis in the absence of oxygen, ACC is transported from the anaerobic environment of flooded roots into the aerobic shoots, where it is converted to ethylene (Fig. 15.5). The ethylene in the shoots causes plants to wilt, to lose biomass, and eventually (if the ethylene remains elevated for a prolonged time) to senesce and die. Treatment of tomato plants with ACC deaminase-containing plant growth-promoting bacteria significantly decreases the damage suffered by these plants due to stress ethylene brought on as a consequence of flooding (Fig. 15.6). These ACC deaminase-containing plant growth-promoting bacteria can act as a sink for ACC, lowering the level of ethylene that can be formed in the shoots and thereby protecting the tomato plants from a portion of the damage caused by flooding.

In addition to protecting plants from flooding damage, ACC deaminase-containing plant growth-promoting bacteria can also significantly decrease the damage to plants that is caused by drought, temperature extremes, high concentrations of salt, and a variety of environmental contaminants. For example, in greenhouse experiments, tomato plants treated with *Achromobacter piechaudii* ARV8, which contains ACC deaminase, were able to grow better in the presence of 86 mM salt (which is usually inhibitory to plant growth) than were tomato plants grown without the added bacterium in either the presence or the absence of salt (Fig. 15.7). This bacterium, which was isolated from a soil sample from the Arava region of the Negev desert in Israel, significantly lowered the level of stress ethylene produced by tomato plants in the presence of salt. More recently, several groups have shown that this approach can facilitate the growth of a range of crops in saline soils in the field, a problem that is endemic to about 25% of the world's arable land.

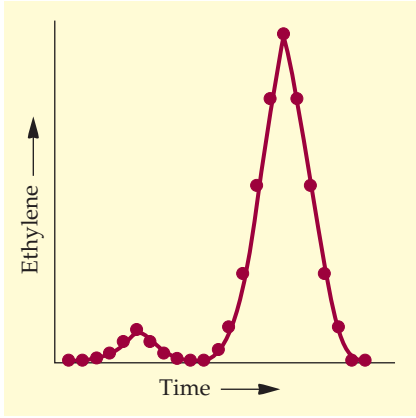
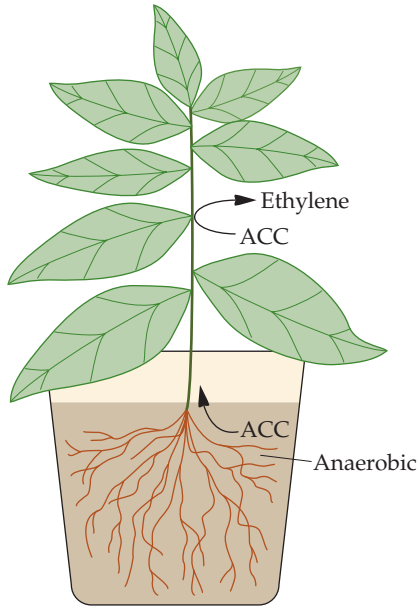


FIGURE 15.4 Time course of plant ethylene synthesis following environmental stress or fungal pathogen infection.

FIGURE 15.5 Schematic representation of a flooded potted plant, where the ACC that is produced in the roots as a consequence of the stress is unable to be converted to ethylene because of the absence of oxygen. The ACC is subsequently transported to the shoots, where oxygen is plentiful, and converted to ethylene, causing epinasty and loss of biomass.



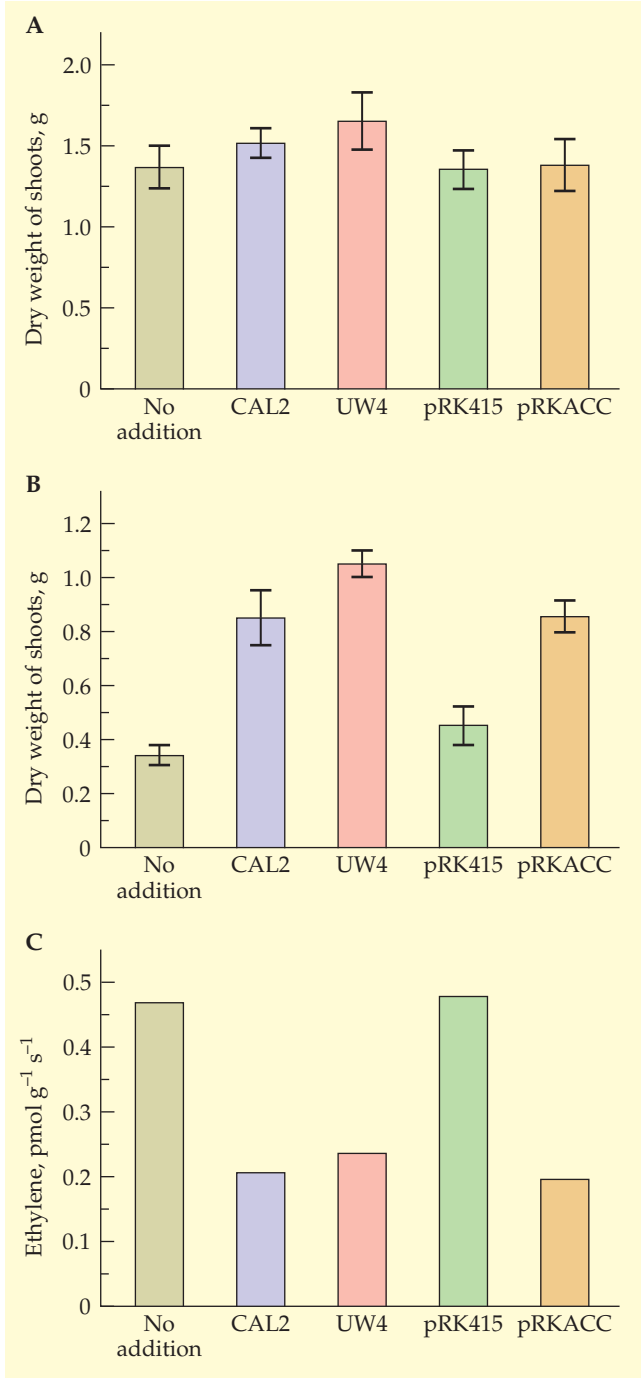


FIGURE 15.6 Effect of ACC deaminase-containing plant growth-promoting bacteria on flooded tomato plants. **(A and B)** Fifty-five-day-old tomato plants were either grown for an additional 9 days (A) or flooded for 9 days before the dry weight of the leaves and shoots was determined (B). **(C)** The amount of ethylene produced by 55-day-old tomato plant leaf stems (petioles) following 9 days of flooding was also measured. ACC deaminase activity is present in *Enterobacter cloacae* CAL2 and *P. putida* UW4, and in another *P. putida* strain transformed with the plasmid pRK415ACC, which carries a bacterial ACC deaminase gene. *P. putida* strain pRK415 has no ACC deaminase activity. The error bars indicate standard errors; the standard errors in panel C were negligible. Nonflooded plants produce approximately 0.07 pmol of ethylene g^{-1} s^{-1}, regardless of the presence or absence of bacteria. Adapted from Grichko and Glick, *Plant Physiol. Biochem.* **39:**11–17, 2001.

Increasing Phosphorus Availability

While a number of plant growth-promoting bacteria can synthesize and secrete organic acids that can dissolve inorganic phosphate in the environment, these organisms can rarely break down phytate, the complex compound (inositol hexaphosphate) that is the major chemical form of phosphorus within cereal grains and oilseeds. Several plants can produce phytases, enzymes that degrade phytate. However, the activity of these

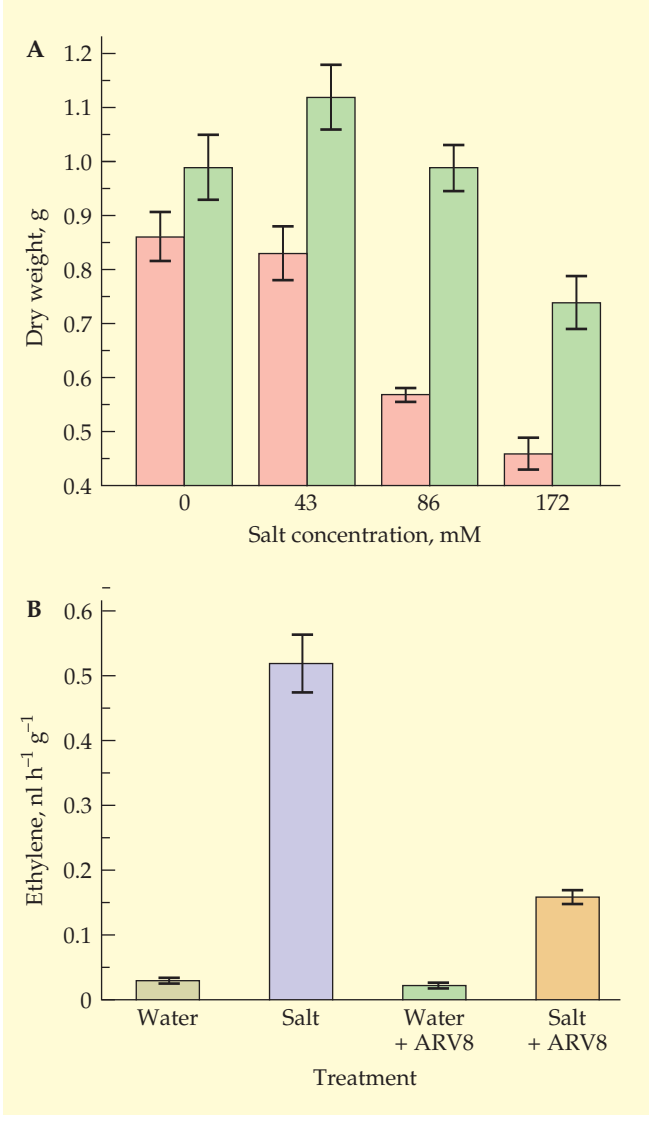


FIGURE 15.7 **(A)** The ACC deaminase-containing bacterium *A. piechaudii* ARV8 (green) increases the tolerance of tomato plants for salt compared to plants grown without the bacterium (red). **(B)** Treatment of tomato plants with salt causes an increase in the synthesis of stress ethylene. Ethylene production in the presence of salt is partially inhibited by *A. piechaudii* ARV8. The error bars indicate standard errors.

enzymes in plant roots is generally low, so these plants cannot efficiently utilize the phytate that is found in the soil. A gene encoding the enzyme phytase was isolated from the fungus *Aspergillus fumigatus*. This gene was inserted, using a transposon, into the chromosomal DNA of a strain of the bacterium *Bacillus mucilaginosus*, which can dissolve phosphorus from calcium phosphate. The transformed bacterial strain can express and secrete active phytase. In greenhouse experiments, this transformed bacterium was found to be superior to the wild-type strain in providing phosphorus to tobacco plants cultivated in its presence (Table 15.2). Importantly, the stimulation of plant growth that was observed in greenhouse experiments

TABLE 15.2 Growth of tobacco plants for 90 days in pots in the greenhouse with different treatments

Characteristic	Soil with no added bacteria	Soil plus wild-type bacterium	Soil plus transformed bacterium
Plant height (cm)	17.4	21.4	24.7
Plant dry weight (mg)	1,297	1,685	1,870
Leaf P content (μg/g)	710	732	800

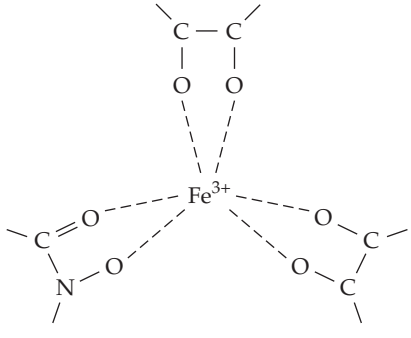
was also evident in the field, where the yield of tobacco plants increased by 19%. Notwithstanding these results, it may be some time before this work is commercialized because of political concerns about the use of genetically engineered bacteria in the environment.

Biocontrol of Pathogens

Phytopathogens are an ongoing and serious agricultural problem that can reduce crop yields by 25 to 100%. This is an enormous loss of productivity. Currently, phytopathogen damage to crops is generally dealt with by the use of chemical agents, although other treatments have also been employed. For most bacterial diseases, plants may be symptomless for prolonged periods before changes in environmental conditions which favor the proliferation of the bacteria cause a rapid outbreak of disease. Under these conditions, severe damage can occur and destroy an entire crop. These field epidemics are difficult and costly to control.

Many of the chemicals that are used to control phytopathogens are hazardous to animals and humans, and they persist and accumulate in natural ecosystems. It is therefore desirable to replace these chemical agents with biological control agents that are more "friendly" to the environment. One approach for the control of phytopathogens is the development of transgenic plants that are resistant to one or more of them (see chapter 18). Alternatively, some plant growth-promoting bacteria can act as biocontrol agents to suppress or prevent phytopathogen damage, and a number of these biocontrol bacteria have been commercialized (Table 15.3). Plant growth-promoting bacteria can produce a variety of substances that limit damage to plants by phytopathogens. They include siderophores, antibiotics, other small molecules, and a variety of enzymes. This approach is still at an early stage of development but appears to have considerable potential. However, the ultimate utility of a strategy based on a particular mechanism can be assessed only under field conditions.

FIGURE 15.8 A six-coordinate iron–siderophore complex. Three bidentate functional groups on a siderophore molecule bind with ferric iron.



Siderophores

Iron is one of the most abundant minerals on Earth and is an essential requirement for living organisms. However, iron in the soil is unavailable for direct assimilation by microorganisms because ferric iron, or Fe(III), which is the predominant form in nature, is only sparingly soluble, i.e., its solubility is about 10^{-18} M at pH 7.4. This amount of soluble iron is much too small to support microbial growth. Consequently, to survive in this environment, soil microorganisms synthesize and secrete low-molecular-mass (~400- to 1,000-dalton) iron-binding molecules known as siderophores (Fig. 15.8). Siderophores bind Fe(III) with a very high affinity (dissociation constant $[K_d]$ = 10^{-20} to 10^{-50} M) and transport it back to cell

TABLE 15.3 Some commercially available biocontrol plant growth-promoting bacteria

Bacterium	Pathogen or disease	Crop(s)
Agrobacterium radiobacter	Crown gall disease caused by *A. tumefaciens*	Fruit trees, nut trees, and ornamental nursery stock
Azotobacter brasilense	Root rot and damping-off	Turf, forage crops, corn
Bacillus subtilis	*Rhizoctonia solani, Pythium* spp., *Fusarium* spp., *Alternaria* spp., and *Aspergillus* spp. that attack roots; also various seedling pathogens	Cotton, legumes, barley, tomato, rice
Bacillus amyloliquefaciens	*Fusarium* spp., *Rhizoctonia* spp.	Herbs, spices, vegetables, tree seedlings, ornamentals
Bacillus pumilis	Powdery mildew, downy mildew, *Fusarium* spp., *Phytophthora* spp., *Rhizoctonia* spp., *Sclerotinia* spp.	Fruits and vegetables, oak, maple, stored seeds
Burkholderia cepacia	*Fusarium* spp., *Pythium* spp., nematodes	Vegetables
Burkholderia cepacia	*Rhizoctonia* spp., *Fusarium* spp., *Pythium* spp., nematodes	Alfalfa, barley, beans, clover, corn, cotton, peas, grain sorghum, vegetables, wheat
Paenibacillus polymyxa	Damping-off, powdery mildew	Cucumber
Pseudomonas chlororaphis	*Fusarium*, leaf stripe, leaf spot, net blotch, spot blotch	Barley, oat
Pseudomonas fluorescens	Frost, *Erwinia amylovora, Pseudomonas tolassii*	Almond, apple, cherry, mushrooms, peach, pear, potato, strawberry, tomato
Pseudomonas syringae	*Botrytis cinerea*, *Penicillium* spp., *Mucor piriformis*, *Geotrichum candidum*	Citrus and pome fruit, potatoes
Streptomyces griseoviridis	*Fusarium* spp., *Alternaria brassicola*, *Phomopsis* spp., *Botrytis* spp., *Pythium* spp., and *Phytophthora* spp.	Field, ornamental, and vegetable crops
Streptomyces lydicus	Control of root rot and damping-off caused by *Fusarium, Rhizoctonia, Pythium, Phytophthora, Sclerotinia, Postia,* and *Verticillium;* also suppresses foliar diseases caused by *Botrytis*	Useful for protection of cuttings of a variety of plants; also used with turfgrass
Mixture of *B. subtilis, P. polymyxa, Bacillus circulans,* and *B. amyloliquefaciens*	A range of fungal damping-off diseases	Especially useful in hydroponic gardens

surface receptors, where it is taken into the cell. Once inside a cell, the iron is released and is then available to support microbial growth.

Plant growth-promoting bacteria can prevent the proliferation of fungal phytopathogens by producing siderophores that bind most of the Fe(III) in the area around the plant root (the rhizosphere). The resulting lack of iron prevents fungal pathogens from proliferating in the immediate vicinity. Fungal phytopathogens also synthesize siderophores, but these generally have a much lower affinity for iron than do the siderophores produced by plant growth-promoting bacteria. In effect, the plant growth-promoting bacteria outcompete fungal phytopathogens for available iron.

Unlike microbial phytopathogens, plants are not generally harmed by the localized depletion of iron in the soil caused by plant growth-promoting bacteria. Most plants can grow at much lower iron concentrations than microorganisms. In addition, some studies have shown that iron that has been sequestered by bacterial siderophores is taken up from the soil by the plant, to its benefit.

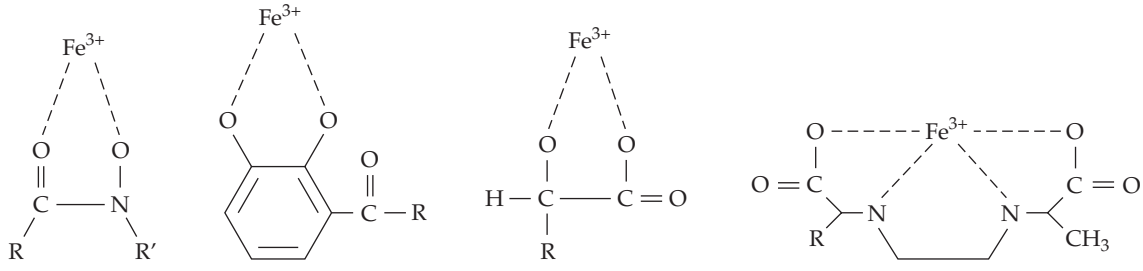


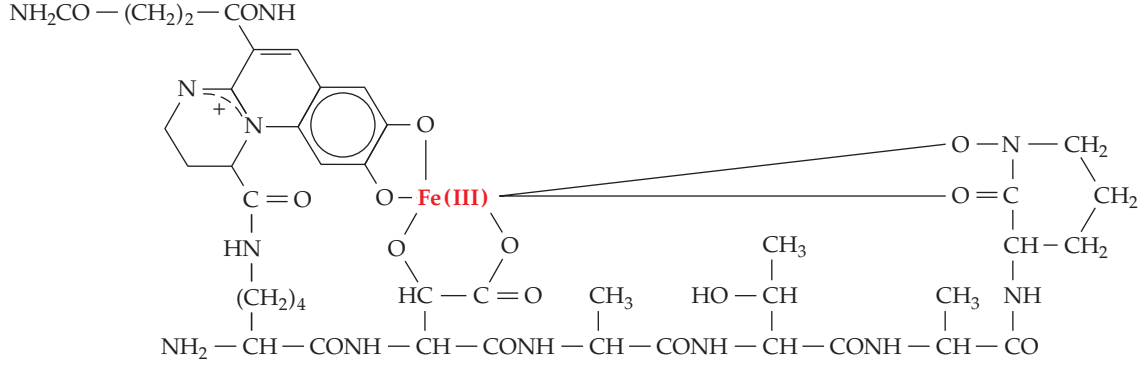
FIGURE 15.9 Iron-binding groups of microbial siderophores.

Because the sequestering of iron by a siderophore-producing bacterium can simultaneously prevent the proliferation of a number of different phytopathogenic microorganisms, siderophore genes are being examined to determine whether they can be used to create more effective biocontrol inoculants.

Siderophores generally have three functional, or iron-binding, groups connected by a flexible backbone, often a peptide. Each functional group usually presents two atoms of oxygen or, less commonly, two nitrogen atoms, that bind iron. In chemical terms, the functional groups are bidentate. Trivalent ferric iron can accommodate three of these groups to form a six-coordinate complex (Fig. 15.8). With some exceptions, the functional groups on microbial siderophores are either hydroxamates or catecholates (Fig. 15.9). Different combinations of functional groups may be present on a single siderophore. Other functional groups include carboxylate moieties, such as citrate, and ethylenediamine (Fig. 15.9). In general, hydroxamate-type siderophores are typical of fungi, and catecholates, which bind iron more tightly than hydroxamates, are common in bacterial siderophores. Plant siderophores, on the other hand, are linear hydroxy- and amino-substituted iminocarboxylic acids, such as mugineic acid and avenic acid.

One bacterial siderophore, called pseudobactin (Fig. 15.10), has been estimated to bind to Fe(III) with an affinity constant of approximately 10^{25} liters mol^{-1}. All fluorescent pseudomonads, so named because they produce a siderophore that fluoresces when excited by ultraviolet light, syn-

FIGURE 15.10 Structure of the siderophore called pseudobactin from *Pseudomonas* sp. strain B10. One molecule of Fe(III) is bound to the siderophore.



thesize structurally related siderophores that differ mainly in the number and configuration of the amino acids in the peptide chain that makes up the backbone.

The synthesis and regulation of pseudobactin in the plant growth-promoting bacterium *P. putida* WCS358 has been examined in detail. Mutagenesis was used to generate a set of 28 mutants that were defective for siderophore production. Two criteria were used for identifying the siderophore-deficient mutants: (1) lack of fluorescence under ultraviolet light and (2) inability to grow in the presence of bipyridyl, a molecule that sequesters most of the iron in the growth medium. When most of the iron is unavailable, only a cell that produces siderophores can grow.

A clone bank of *P. putida* WCS358 DNA was constructed in the broad-host-range cosmid vector pLAFR1 and was introduced by conjugation into each of the 28 siderophore mutants (Fig. 15.11). All of the resultant transformants were tested by complementation for restoration of fluorescence and/or the ability to grow in the presence of bipyridyl. Thirteen separate complementing cosmid clones, with an average insert size of 26 kilobase pairs (kb), were identified. After detailed analyses, these clones were found to represent at least five separate gene clusters.

One of these gene clusters has been studied further. It has a length of 33.5 kb and contains five transcriptional units with at least seven separate genes. Thus, like nitrogen fixation and nodulation, siderophore biosynthesis

FIGURE 15.11 Cloning genes involved in siderophore biosynthesis. The clone bank is constructed using the broad-host-range cosmid pLAFR1. The cells that have mutations in one of the genes involved in siderophore biosynthesis are unable to grow on medium containing bipyridyl, which sequesters all of the free iron in the medium. Cells with mutations in genes involved in siderophore biosynthesis are selected from the replica plate that does not contain bipyridyl. Transformants that can grow in the presence of bipyridyl are able to complement the mutation in one of the siderophore biosynthesis genes.

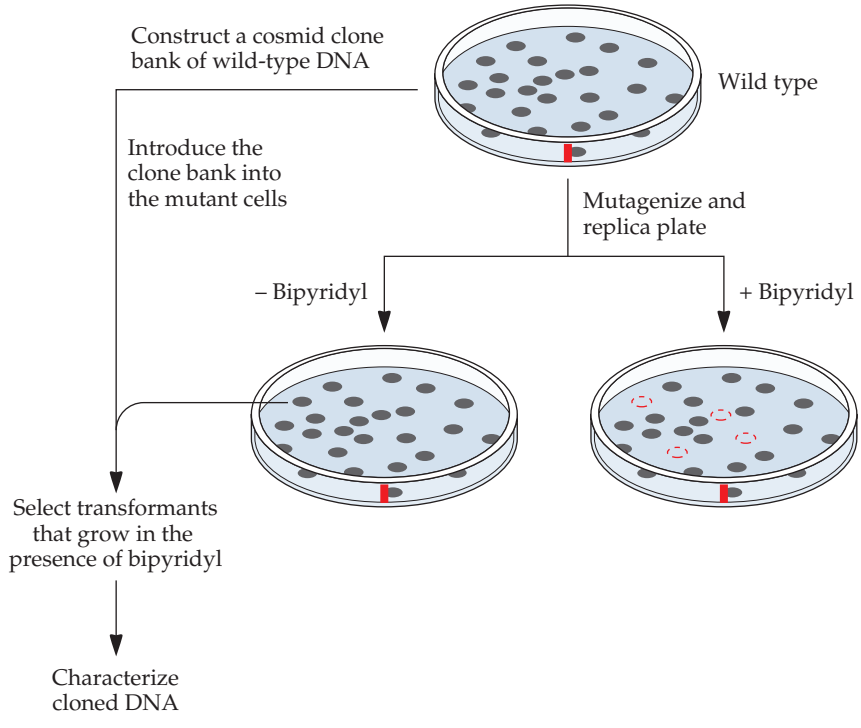


TABLE 15.4 Effect of additional copies of the *rpoD* gene that encodes ρ^{70} from *P. fluorescens* CHA0 on the ability of the bacterium to prevent damage to cucumber roots caused by the pathogenic fungus *P. ultimum*

Plant growth-promoting bacterium added	Average root fresh weight (mg)	
	Without *P. ultimum*	With *P. ultimum*
None	382	44
P. fluorescens CHA0	386	177
P. fluorescens CHA0 with vector	365	146
P. fluorescens CHA0 with vector and *rpoD* gene	371	335

Adapted from Schnider et al., *J. Bacteriol.* **177:**5387–5392, 1995.

In the absence of the plant growth-promoting bacterium *P. fluorescens* CHA0, the pathogen *P. ultimum* dramatically inhibits root growth. The presence of the *rpoD* gene on the plasmid vector enhanced the activity of the plant growth-promoting bacterium. Plants were grown for 2 weeks before their roots were measured.

is a complex process. Since each siderophore is encoded by a number of different genes, genetically engineering bacteria to produce modified siderophores is not a simple matter. However, there may be other ways to improve the effectiveness of plant growth-promoting bacteria as biocontrol agents. For example, it may be possible to extend the range of iron–siderophore complexes that one bacterial strain can utilize so that a genetically altered plant growth-promoting biocontrol bacterial strain could take up and use siderophores synthesized by other soil microorganisms, thereby giving it a competitive advantage. This was done by cloning the genes for iron–siderophore receptors from one plant growth-promoting control bacterium and introducing them into other strains.

Antibiotics

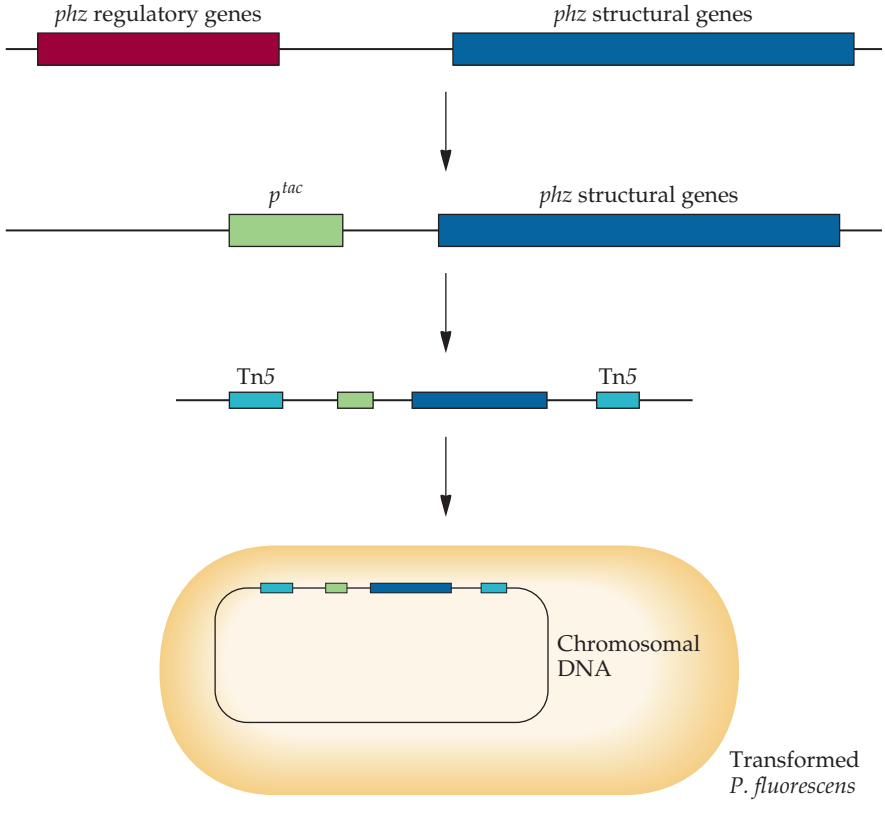
One of the most effective mechanisms by which a plant growth-promoting bacterium can prevent phytopathogen proliferation is the synthesis of antibiotics. For example, the antibiotics synthesized by biocontrol pseudomonads include agrocin 84, agrocin 434, 2,4-diacetylphloroglucinol, herbicolin, oomycin, phenazines, pyoluteorin, and pyrrolnitrin.

The biocontrol activity of a plant growth-promoting bacterium may be improved by providing it with genes that encode the biosynthesis of antibiotics that are normally produced by other bacteria. In this way, the range of phytopathogens that a single biocontrol bacterium can suppress can be extended. Moreover, by limiting the growth of other soil microorganisms, antibiotic-secreting plant growth-promoting bacteria should facilitate their own proliferation, since they will have fewer competitors for limited nutritional resources. In addition, genetic manipulation can be used to increase the amount of antibiotic that a bacterium synthesizes.

The production of a number of antifungal metabolites that are produced by pseudomonads appears to be controlled by a protein that acts as a global transcriptional regulator; therefore, it should be possible to enhance antibiotic production by modifying this global regulation. For example, antibiotic production was enhanced after *Pseudomonas fluorescens* CHA0 was transformed with a vector carrying the gene encoding the housekeeping RNA polymerase sigma-70 (σ^{70}). The modified strain was more effective at protecting cucumber plants against a root disease caused by the fungus *Pythium ultimum* (Table 15.4).

A single copy of the operon carrying all seven of the genes that encode the biosynthesis of the antibiotic phenazine-1-carboxylic acid (i.e., *phzABCDEFG*) was inserted into the chromosomal DNA of a plant growth-promoting bacterial strain of *P. fluorescens* (Fig. 15.12). The wild-type version of this bacterium, which does not synthesize phenazine-1-carboxylic acid, acts as a biocontrol agent against some fungal diseases. As indicated by a much larger zone of clearance of the fungal pathogen *P. ultimum* on solid medium, the engineered bacterium has a higher level of biocontrol activity than the wild type (Fig. 15.13). Also, the phenazine-1-carboxylic acid-producing bacterium prevented *P. ultimum*-caused damping-off disease in pea plants in soil. This work demonstrates the efficacy of this approach under greenhouse conditions; however, it remains to be demonstrated whether this altered bacterium is effective in the field.

FIGURE 15.12 Chromosomal insertion of the antibiotic phenazine-1-carboxylic acid operon *(phz)* into a biocontrol strain of *P. fluorescens*. The regulatory genes that normally control the expression of the seven biosynthetic genes were removed, and the entire operon was placed under the control of the *tac* promoter *(p^{tac})*. Since *P. fluorescens* does not utilize lactose as a carbon source, it does not encode the *lac* repressor, and in the absence of the *lac* repressor, any genes under the control of the *tac* promoter are expressed constitutively. The operon, under the control of the *tac* promoter, was inserted into a derivative of transposon Tn*5* adjacent to a kanamycin resistance gene (not shown) on a plasmid. Tn*5* facilitates integration of DNA into the chromosome of the host cell. Transconjugants in which the chromosomal insertion had not inactivated any important bacterial functions were tested for their effectiveness as biocontrol strains. The Tn*5* derivative is designed so that it does not easily pass from the biocontrol strain to other bacteria in the environment.



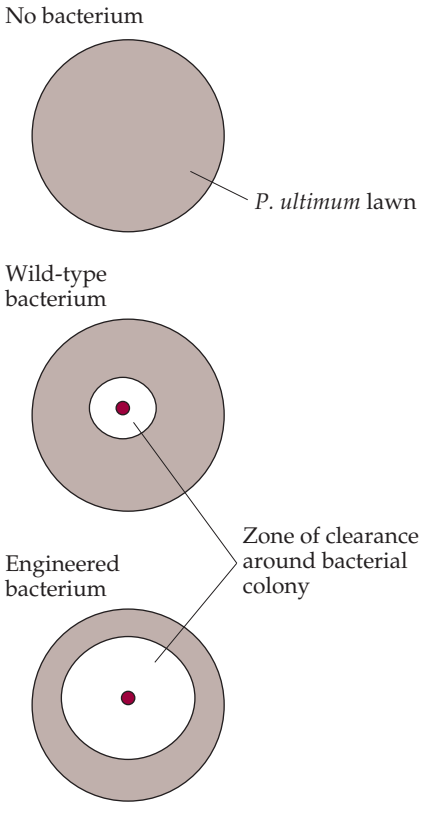


FIGURE 15.13 Effect of transforming a strain of *P. fluorescens* with genes expressing the antibiotic phenazine-1-carboxylic acid biosynthetic pathway on the ability of the bacterium to prevent the growth of the fungal pathogen *P. ultimum* on solid medium. The antifungal activity of the bacterium is proportional to the area of the zone of clearance around the center of the petri plate to which the bacterium is added.

At present, there is still only one commercially available genetically engineered biocontrol bacterial strain. A modified version of *Agrobacterium radiobacter* K84 has been marketed, first in Australia in 1989, and more recently all over the world, as a means of controlling crown gall disease, which is caused by the bacterium *Agrobacterium tumefaciens*. This disease affects almond trees and stone fruit trees, such as peach trees. The antibiotic agrocin 84, which is produced by *A. radiobacter,* is toxic to *A. tumefaciens*. However, agrocin 84-resistant strains of *A. tumefaciens* can develop if the plasmid carrying the genes for the biosynthesis of agrocin 84 is accidentally transferred from *A. radiobacter*. To avoid this possibility, the region of DNA responsible for plasmid transfer was removed from the agrocin 84 plasmid, pAgK84 (Fig. 15.14). As a result of this deletion, the *A. radiobacter* strain retains the capacity to act as a biocontrol agent, but it can no longer transfer the plasmid to pathogenic agrobacteria.

Enzymes

Some plant growth-promoting bacteria produce enzymes, such as chitinase, β-1,3-glucanase, protease, and lipase, that can degrade fungal cell walls and cause the fungal cells to lyse (Fig. 15.15). In one study, the incidence of plant disease caused by the phytopathogenic fungi *Rhizoctonia solani, Sclerotium rolfsii,* and *P. ultimum* was reduced by using a β-1,3-glucanase-producing strain of *Burkholderia cepacia*. In another study, the antifungal activities of three different strains of the plant growth-promoting bacterium *Enterobacter agglomerans* were attributed to a complex of four separate polypeptides that act together to degrade the chitin in fungal cell walls. When tested, these bacteria significantly decreased the damage to cotton plants following infection with *R. solani*. Moreover, Tn*5* mutants of *E. agglomerans* that were deficient in chitinase activity were unable to protect plants against damage caused by the fungal pathogen, indicating that the chitinase was the active element.

Many of the bacterial enzymes that can lyse fungal cells, including chitinases and β-glucanases, are encoded by a single gene. It should therefore be straightforward to isolate these genes and transfer them to plant growth-promoting bacteria to construct strains that produce fungus-degrading enzymes. In one series of experiments, a chitinase gene was isolated from the bacterium *Serratia marcescens* and then transferred into *Trichoderma harzianum* and *Rhizobium meliloti* cells. In both cases, the transformed microorganisms produced chitinase and displayed increased antifungal activity. When the *S. marcescens* chitinase gene was introduced into a strain of *P. fluorescens* that directly promotes plant growth, the transformant also stably expressed and secreted active chitinase and effectively controlled the phytopathogen *R. solani*.

Ice Nucleation and Antifreeze Proteins

One of the ways in which some pathogenic leaf bacteria, such as *Pseudomonas syringae,* damage plants is by synthesizing ice nucleation proteins. These proteins, which are produced at low temperatures, are present on the surface of the bacterium and act as sites that facilitate the formation of ice crystals at freezing temperatures. As the ice crystals grow, they can pierce the plant cells and cause irreparable damage. The bacteria benefit from this damage by gaining direct access to the nutrients from the lysed plant cells. In the absence of ice nucleation proteins on the leaf surface, a

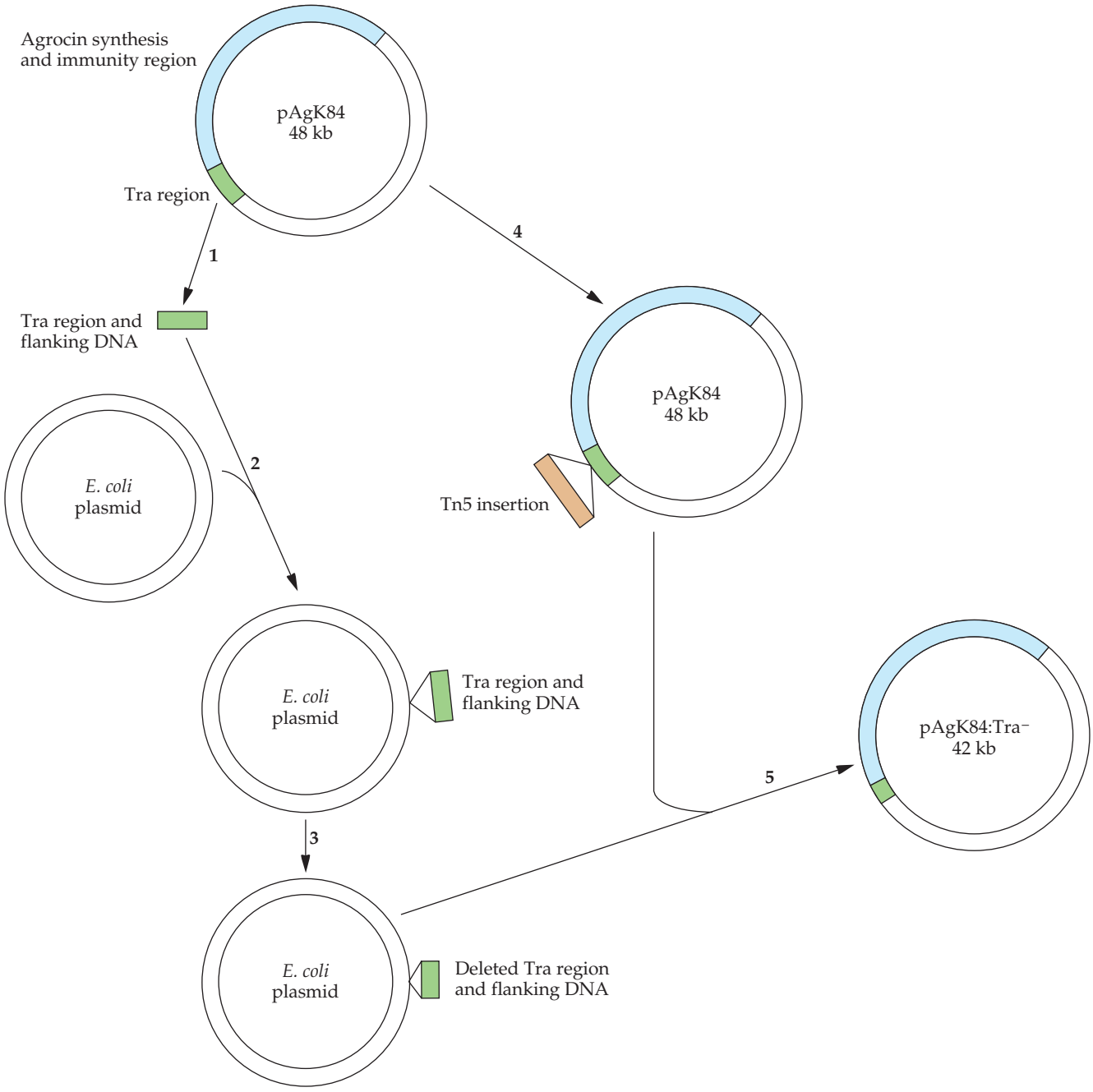


FIGURE 15.14 Construction of a transfer-deficient (Tra⁻) derivative of plasmid pAgK84 from *A. radiobacter,* which encodes both synthesis of and immunity to the antibiotic agrocin 84. Based on knowledge of the restriction enzyme map of pAgK84, a DNA fragment containing the transfer (Tra) region, together with some of the flanking DNA, is isolated **(1)** and spliced into an *E. coli* plasmid vector **(2).** By restriction enzyme digestion, approximately 80% of the Tra region and some of the flanking DNA (a total of about 6 kb) is deleted from the cloned DNA containing the Tra sequence **(3).** Homologous recombination of the *E. coli* plasmid containing the deleted Tra region with plasmid pAgK84, in which transposon Tn*5,* which carries a kanamycin resistance gene, has been inserted into the Tra region, is performed **(4).** This results in some derivatives of pAgK84 in which a portion of the Tra region of the plasmid has been deleted **(5)**. The resultant Tra⁻ mutant of pAgK84 can no longer be conjugationally transferred to other agrobacteria, although it is still able to synthesize and provide immunity to agrocin 84. None of the DNA fragments is shown to scale.

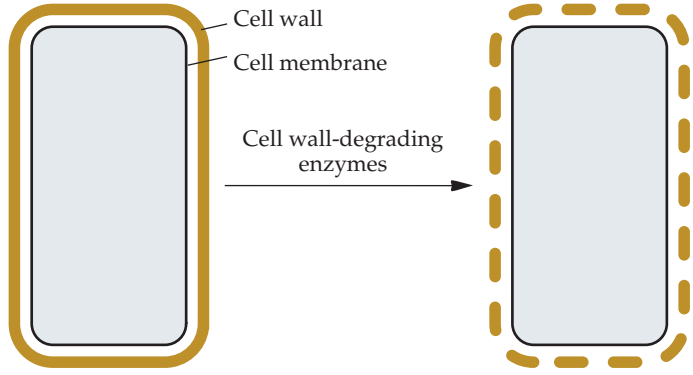
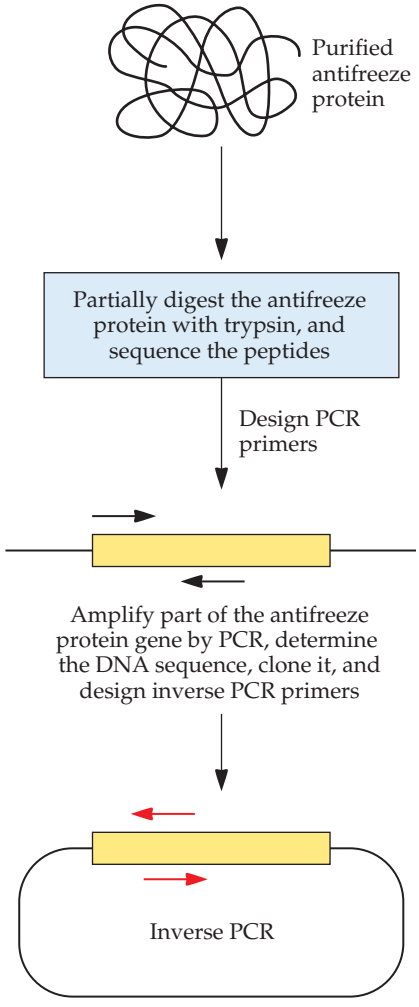


FIGURE 15.15 Schematic representation of a fungal cell wall being degraded by one or more enzymes produced by a biocontrol plant growth-promoting bacterium. Following the breakdown of the fungal cell wall, the cell is readily lysed.

FIGURE 15.16 Strategy for the isolation of a *Pseudomonas* antifreeze protein gene. PCR primers based on the partial amino acid sequence of the isolated protein were used to amplify a portion of the antifreeze protein gene. Inverse PCR primers were designed based on the DNA sequence of the amplified gene fragment.



brief overnight frost would not damage the plant because the water in a plant cell must usually be several degrees below the freezing point before ice crystals begin to form (i.e., it must be supercooled). One way to prevent freezing damage caused by *P. syringae* to susceptible crops, such as strawberries, is by spraying the plants, prior to the frost, with a mutant form of the ice-nucleating bacterium. Such a mutant, which may be constructed by either recombinant DNA manipulation or conventional mutagenesis and selection, lacks the ability to produce the ice nucleation protein, and therefore, ice crystals are not formed on the leaf surface. If a sufficient number of these "ice-minus" mutant bacteria are sprayed onto a susceptible plant, the mutant will displace the wild-type (ice-plus) bacteria, thereby preventing ice nucleation.

An important facet of the effectiveness of a biocontrol plant growth-promoting bacterium is its ability to persist and proliferate in the natural environment. In areas such as Canada, Scandinavia, Russia, and the northern United States, these organisms must survive long, cold winters and then grow at cool soil temperatures in the spring (~5 to 10°C). While microorganisms have a variety of adaptive strategies for thriving under adverse conditions, it may be possible to engineer organisms that are better able to deal with cold temperatures. Some soil bacteria that are also able to promote plant growth can both grow at 5°C and secrete antifreeze proteins into the surrounding medium when grown at low temperatures. A bacterial antifreeze protein regulates the formation of ice crystals outside the bacterium and thereby provides protection for the bacterium in the soil. In the presence of bacterial antifreeze protein, ice crystals still form, but their size is limited. In the absence of antifreeze proteins, ice crystals can grow to a large size and eventually puncture the bacterial cell wall and membrane, causing cell lysis. Ice crystals do not form inside the bacterium to any great extent. This is because at low temperatures bacteria decrease their volume by pumping some of their water from inside to outside the cell.

Recently, the gene for a bacterial antifreeze protein was isolated and characterized (Fig. 15.16). The strategy that was used to isolate this gene included purifying the protein, digesting it into small peptides with the proteolytic enzyme trypsin, determining the amino acid sequences of several of these peptides, and using those amino acid sequences to design polymerase chain reaction (PCR) primers for a portion of the antifreeze

MILESTONE

Cloning of *Rhizobium meliloti* Nodulation Genes by Direct Complementation of Nod⁻ Mutants

S. R. LONG, W. J. BUIKEMA, and F. M. AUSUBEL
Nature **298:**485–488, 1982

The isolation of a gene, when its basic features are unknown and there is no heterologous probe or antibody available, can be a daunting task. In these instances, it is often necessary to devise an innovative selection scheme. The selection scheme may be based on immunological detection of the target protein, determination of the activity of the target protein, DNA hybridization with an oligonucleotide probe whose sequence is deduced from the partial amino acid sequence of the purified target protein, or mutant complementation. Often, after a gene that encodes a particular function has been isolated, similar genes from other organisms can be isolated by using the first gene as a heterologous DNA hybridization probe. This approach depends on nucleotide sequence similarity between the probe and the target gene. This strategy generally works well for genes that are evolutionarily conserved, such as those that encode proteins involved in the process of nitrogen fixation, but it works poorly for many other genes, including those that encode proteins that degrade cellulose.

When Long and her coworkers set out to isolate and characterize the nodulation genes from *R. meliloti,* nodulation genes had never been isolated. Very little was known about how many genes were involved in this process or what products they might encode. As a starting point, these workers isolated and characterized several nodulation-defective mutants of *R. meliloti*. However, these studies had not given them any clue to the functions of these genes. Therefore, they set about selecting nodulation genes that could complement their nodulation-defective *R. meliloti* mutants. This work was further complicated by the fact that, at that time, clone banks were almost always constructed and maintained in *E. coli*. Therefore, to facilitate the assembly of the *R. meliloti* clone bank and its subsequent transfer to nodulation-defective mutant strains of *R. meliloti,* the researchers first constructed a broad-host-range cosmid vector in which there were large DNA fragments, with an average size of approximately 23 kb, that could be stably maintained in a number of different gram-negative bacteria. The large size of the DNA on the cosmid increased the likelihood that other genes involved in nodulation would be on the same DNA fragment as the complementing DNA sequence. After the cosmids had been transferred by conjugation to *R. meliloti,* the transformed bacteria were tested for their ability to nodulate alfalfa plants. In previous experiments, it had been found that even one nodulation-proficient bacterium in the presence of 10^4 defective ones could successfully nodulate alfalfa plants. Bacteria with cosmids carrying a complementing gene were isolated directly from the nodules.

In summary, these researchers were the first to isolate nodulation genes. They devised a clever and effective selection scheme. The broad-host-range cosmid vector that they developed has since been used many times by a large number of other researchers.

protein gene. Following the sequence determination of the PCR-amplified DNA, inverse PCR (Fig. 15.16) was used to obtain the remaining portion of the gene. It should be possible to transfer this gene to various strains of plant growth-promoting bacteria to create strains that can persist and proliferate at cold temperatures. Although there is currently no evidence to link antifreeze activity with the mechanism that bacteria use to function at low temperatures, it will be interesting to examine experimentally whether antifreeze protein activity is part of the adaptive strategy used by some bacteria for cold, as well as freezing, tolerance.

Ethylene

Fungal pathogens not only directly inhibit plant growth, they also cause the plant to synthesize stress ethylene, which causes some of the damage sustained by plants infected with fungal phytopathogens. For example, it is well known that exogenous ethylene often increases the severity of a fungal infection, whereas ethylene synthesis inhibitors significantly decrease the severity of a fungal infection.

TABLE 15.5 Reduction in the severity of damage (damping off) of cucumber caused by *P. ultimum*

Treatment	Seed germination rate (%)	Shoot fresh weight (g)	Root fresh weight (g)
None	100	3.48	2.85
P. ultimum	31	0.69	0.21
CHA0 + *P. ultimum*	79	2.11	0.95
CHA0/ACC + *P. ultimum*	87	2.27	1.27

Adapted from Wang et al., *Can. J. Microbiol.* **46:**898–907, 2000.

CHA0, the biocontrol bacterium *P. fluorescens* CHA0; CHA0/ACC, the biocontrol bacterium *P. fluorescens* CHA0 transformed with the ACC deaminase gene from *E. cloacae* UW4; UW4, *E. cloacae* UW4. The expression of a foreign ACC deaminase gene in *P. fluorescens* CHA0 results in an increase in the number of cucumber seeds that germinate in the presence of the fungal pathogen, as well as an increase in the fresh weight of both shoots and roots of the resulting plants.

As mentioned earlier, stress ethylene that is synthesized in response to fungal pathogen infection is produced in two peaks, a small one that occurs within a few hours after fungal infection and a much larger peak that occurs several days after fungal infection (Fig. 15.4). The first peak turns on the transcription of genes that encode proteins that protect plants against the pathogen, while the second peak is deleterious to the plant. Ideally, it would appear to be advantageous to allow the plant to synthesize the first, but not the second, peak of ethylene in response to a fungal pathogen. This is readily achieved using ACC deaminase-containing plant growth-promoting bacteria, since high-level expression of ACC deaminase does not occur until several hours after the appearance of increased amounts of ACC. Thus, these bacteria do not alter the first ethylene peak but significantly decrease the magnitude of the second peak.

A biocontrol bacterial strain, *P. fluorescens* CHA0, was transformed with the *P. putida* UW4 ACC deaminase gene, and the effect of this manipulation on the damage to cucumbers caused by *P. ultimum* was assessed. The ACC deaminase-containing biocontrol bacterial strain was more effective in lessening the damage than the wild-type biocontrol strain that did not possess the enzyme. Not only did plants inoculated with the ACC deaminase-transformed strains have greater root and shoot biomass than those treated with the wild-type biocontrol strain, but also the number of seeds that germinated in pathogen-containing soil was larger (Table 15.5). In addition, the ACC deaminase-transformed biocontrol strain reduced the extent of soft rot of potato slices, caused by the bacterial pathogen *Erwinia carotovora* subsp. *carotovora,* in sealed plastic bags by 50% compared with the wild-type biocontrol strain (Table 15.6). In effect, ACC deaminase acts synergistically with other mechanisms of biocontrol, such as the production of antibiotics or antifungal enzymes, to prevent phytopathogens from damaging plants.

Root Colonization

Depending upon the mechanism that a particular biocontrol bacterium uses to thwart the damage to plants caused by pathogenic microorganisms, it can be advantageous for the biocontrol strain to bind as tightly as possible to the plant root. One way to improve root colonization by biocontrol bacteria entails overexpressing the bacterial *sss* gene. This gene is normally thought to play a role in DNA rearrangements that regulate the transcrip-

tion of a gene(s) involved in the biosynthesis of cell surface components. When the *sss* gene was introduced on a plasmid into two strains of *P. fluorescens*, one that was normally a poor root colonizer and the other a good root colonizer, colonization of tomato roots was increased by approximately 28- and 12-fold, respectively. This work is an important first step in engineering more effective strains of biocontrol bacteria.

Nitrogen Fixation

Nitrogen gas (N_2), which makes up approximately 80% (by volume) of the air that we breathe, cannot be used directly by plants or animals to synthesize essential nitrogen-containing biomolecules, such as amino acids and nucleotides. Rather, it must first be converted (fixed) into ammonia. This conversion requires a high input of energy because the triple bond of N_2 (N≡N) is extremely stable. The energy for the biological fixation of nitrogen comes from the hydrolysis of large amounts of ATP. Similarly, the chemical (industrial) conversion of N_2 to ammonia uses a considerable amount of energy in the form of high temperature and pressure.

More than 100 million tons of fixed nitrogen is needed annually to sustain global food production. Synthetic (chemically produced) fertilizers account for about half of this nitrogen supply, and most of the remainder is derived from diazotrophic bacteria. No eukaryote is known to fix nitrogen. Chemical fertilizers have helped considerably in increasing crop yields, but their continual use has led to pollution problems as a result of runoff and to depletion of the nutrient reserves in the soil. Moreover, their cost has been rising steadily. These factors have provided an incentive for developing alternative sources of fixed nitrogen, including the development of diazotrophic microorganisms as "bacterial fertilizers."

A wide range of bacteria can fix nitrogen, and a number of them have potential as crop fertilizers. However, until a bacterial fertilizer has been shown conclusively to be as effective as a chemical formulation, there will be reluctance to change current practices, especially in those countries where the cost of chemical fertilizer is not significant relative to the value of the crop. For example, soybeans, which constitute the second largest crop in the United States in terms of both cash value and total acres planted, form a beneficial symbiotic relationship with the bacterium *Bradyrhizobium japonicum*. In this symbiosis, the bacteria provide the plant with fixed

TABLE 15.6 Effect of *P. fluorescens* CHA0 and CHA0/ACC on soft rot of potatoes by *E. carotovora* subsp. *carotovora*

Treatment	Weight of rotted potatoes (g/slice)
None	15.6
CHA0	14.5
CHA0/ACC	7.5

Adapted from Wang et al., *Can. J. Microbiol.* **46:**898–907, 2000.

For definitions of abbreviations, see Table 14.9. While the biocontrol strain *P. fluorescens* CHA0 does not significantly alter the extent of damage to potatoes due to the bacterial pathogen, expression of ACC deaminase lowers the level of stress ethylene and decreases the damage to the potatoes by approximately 50%. The total weight of each potato slice was approximately 20 g.

nitrogen and, in turn, receive photosynthetically fixed carbon from the plant. When plants are inoculated with specific strains of *B. japonicum*, the final yield of plant material can be increased by 25 to 50%, and the inoculated plants no longer require the addition of chemically fixed nitrogen. Although approximately 40% of the world's soybean crop is produced in just a few locales in the United States and agricultural practices tend to be similar throughout these locales, at present, only a small fraction of this crop is treated with *B. japonicum*. Most of these farmers continue to depend on the naturally occurring strains of *B. japonicum* in the soil and chemical fertilizers.

The most important microorganisms that are currently used agriculturally to improve the nitrogen content of plants include a range of rhizobial genera and species (Table 15.7). These bacteria are gram negative, flagellated, and rod shaped, and they form symbiotic relationships with legumes. Generally, each rhizobial species is specific for a limited number of plants and will not interact with plants other than its natural hosts (Table 15.7).

As part of their life cycle, rhizobial bacteria invade plant root cells and initiate a complex series of developmental changes that lead to the forma-

TABLE 15.7 Plant specificities of various rhizobial species

Bacterial species	Host plant(s)
Azorhizobium caulinodans	West African legume *(Sesbania rostrata)*
Bradyrhizobium elkanii	Soybean *(Glycine max)*, black-eyed pea *(Vigna unguiculata* subsp. *dekindtiana)*, mung bean *(Vigna radiata)*
Bradyrhizobium japonicum	Soybean *(G. max)*
Mesorhizobium amorphae	Desert false indigo *(Amorpha fruticosa)*
Mesorhizobium ciceri	Chickpea *(Cicer arietinum)*
Mesorhizobium chacoense	White carob tree *(Prosopis alba)*
Mesorhizobium huakuii	Chinese milk vetch *(Astragalus sinicus)*
Mesorhizobium loti	Lotus *(Lotus japonicus)*
Mesorhizobium mediterraneum	Chickpea (*C. arietinum*)
Mesorhizobium tianshanense	7 Legume species
Rhizobium sp. strain NGR234	>100 Tropical legume species
Rhizobium etli	Kidney bean *(Phaseolus vulgaris)*, mung bean *(V. radiata)*
R. etli bv. mimosae	Mimosa *(Mimosa affinis)*
Rhizobium galegae	Goat's rue *(Galega officinalis, Galega orientalis)*
Rhizobium gallicum	Common bean *(P. vulgaris)*
Rhizobium huautlense	Danglepod *(Sesbania herbacea)*
Rhizobium leguminosarum bv. phaseoli	Kidney bean, mung bean
R. leguminosarum bv. trifolii	Clover (*Trifolium* spp.)
R. leguminosarum bv. viciae	Pea *(Pisum sativum)*
Rhizobium sullae	Sweetvetch *(Hedysarum coronarium)*
Rhizobium tropici	Mimosoid trees (*Leucaena* spp.) and some tropical legume trees (*Macroptilium* spp.)
Sinorhizobium fredii	Soybean *(G. max)*
Sinorhizobium meliloti	Alfalfa *(Medicago sativa)*
Sinorhizobium morelense	White popinac *(Leucaena leucocephala)*

tion of a root nodule. Inside the root nodule, the bacteria proliferate and persist in a form that has no cell wall. The bacteria within the nodules fix atmospheric nitrogen by means of the enzyme nitrogenase. The structural and biochemical interactions between a symbiotic rhizobacterium and a host plant are quite intricate and mutually beneficial. Inside a nodule, nitrogenase is protected from the toxic effects of atmospheric oxygen in two ways. First, oxygen does not readily diffuse into a nodule. Second, the oxygen content within a nodule is regulated by the protein leghemoglobin. The heme moiety of this oxygen-binding protein is synthesized by the bacterium, and the globin portion of the molecule is encoded by a plant gene. The plant also provides the bacterium with photosynthetically fixed carbon, which the bacterium requires for growth. For its part, the plant benefits from this symbiotic relationship by receiving fixed nitrogen from the bacterium.

Nitrogenase

The renewed interest in diazotrophs as biological fertilizers overlapped the development of techniques for gene isolation and manipulation and provided the impetus for studying the biochemical and molecular biological aspects of nitrogen fixation. Initially, scientists believed that these studies would lead to the development of improved nitrogen-fixing organisms that would enhance crop yields. Some researchers even went so far as to suggest that bacterial genes for nitrogen fixation might be introduced directly into plants to enable them to fix their own nitrogen. Although this overly optimistic prediction has not materialized, a detailed understanding of the process of nitrogen fixation has emerged. And with this understanding, the possibility of improving the nitrogen-fixing activity of some diazotrophs by genetic manipulation is a little closer to becoming a reality.

Components of Nitrogenase

All known nitrogenases have two oxygen-sensitive components. Component I is a complex of two identical α-protein subunits (approximately 50,000 daltons each), two identical β-protein subunits (approximately 60,000 daltons each), 24 molecules of iron, 2 molecules of molybdenum, and an iron–molybdenum cofactor, often called FeMoCo (Fig. 15.17). Component II has two α-protein subunits (approximately 32,000 daltons each), which are not the same as the α-protein subunits of component I, and a number of associated iron molecules. The catalysis of nitrogen to ammonia requires the combination of components I and II, a complex of magnesium and ATP, and a source of reducing equivalents (reaction 1; the upward-pointing arrow indicates a gas and P_i is inorganic phosphate).

$$N_2 + 8H^+ + 8e^- + 16MgATP \rightarrow 2NH_3 + H_2\uparrow + 16MgADP + 16P_i \quad \textbf{(1)}$$

In addition to fixing nitrogen, the nitrogenase can reduce the gas acetylene to ethylene (reaction 2).

$$H—C\equiv C—H + 2H^+ \rightarrow H_2=C=C=H_2 \quad \textbf{(2)}$$

The measurement by gas chromatography of ethylene production as a function of time provides a convenient assay for nitrogenase activity. This assay can be performed with intact cells in solution (Fig. 15.18), bacteria associated with plant roots, crude cell extracts, or highly purified enzyme preparations.

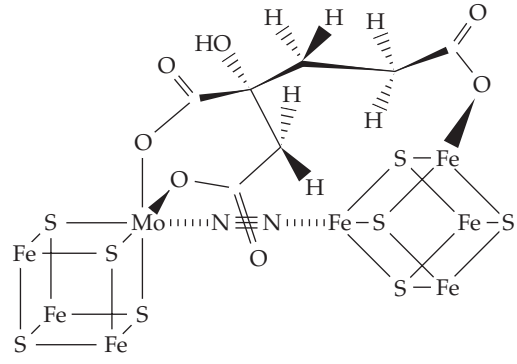


FIGURE 15.17 Structure of the iron–molybdenum cofactor bound to a molecule of dinitrogen (N_2).

Component I catalyzes the actual reduction of N_2, and component II donates electrons to component I. Both components are extremely sensitive to oxygen and can be rapidly and irreversibly inactivated when the oxygen concentration is too high. In addition to components I and II, the activity of a complete, functional nitrogenase depends on 15 to 20 additional accessory proteins. The roles of most of the accessory proteins have been delineated and include the transfer of electrons to component II and the biosynthesis of the iron–molybdenum cofactor that is a part of component I.

Genetic Engineering of the Nitrogenase Gene Cluster

Nitrogen fixation is a very complicated process requiring the concerted actions of a large number of different proteins. Therefore, it was not realistic to expect either that an intact single DNA fragment containing all the genetic information for nitrogen fixation could be readily cloned from a diazotrophic microorganism and transferred into a nondiazotrophic organism or that a recipient organism could maintain the physiological conditions needed for nitrogenase activity. Consequently, the most direct way to isolate the genes involved in nitrogen fixation (*nif* genes) was to identify and characterize those clones of a wild-type library that restore nitrogen fixation to various mutants of the original organism. This process is called genetic complementation.

The first *nif* genes identified by complementation were isolated from clone banks of the diazotroph *Klebsiella pneumoniae*. This well-studied organism is found in soil and water, as well as in the human intestine. The isolation protocol comprises the following steps (Fig. 15.19).

1. *K. pneumoniae* cells are treated with a dose of a mutagenic agent that allows approximately 0.1 to 1.0% of the cells to survive. Some of the mutagenized cells are able to grow on a minimal medium containing a source of fixed nitrogen, such as NH_4Cl, but do not grow in the absence of fixed nitrogen. These cells are likely to have a mutation in a *nif* gene and are designated Nif^-.
2. A clone bank that consists of chromosomal DNA from wild-type (Nif^+) *K. pneumoniae* cells is constructed in a broad-host-range plasmid expression vector and maintained in *Escherichia coli*.
3. The Nif^- *K. pneumoniae* cells are conjugated with the *E. coli* cells that carry the clone bank on a plasmid shuttle vector.

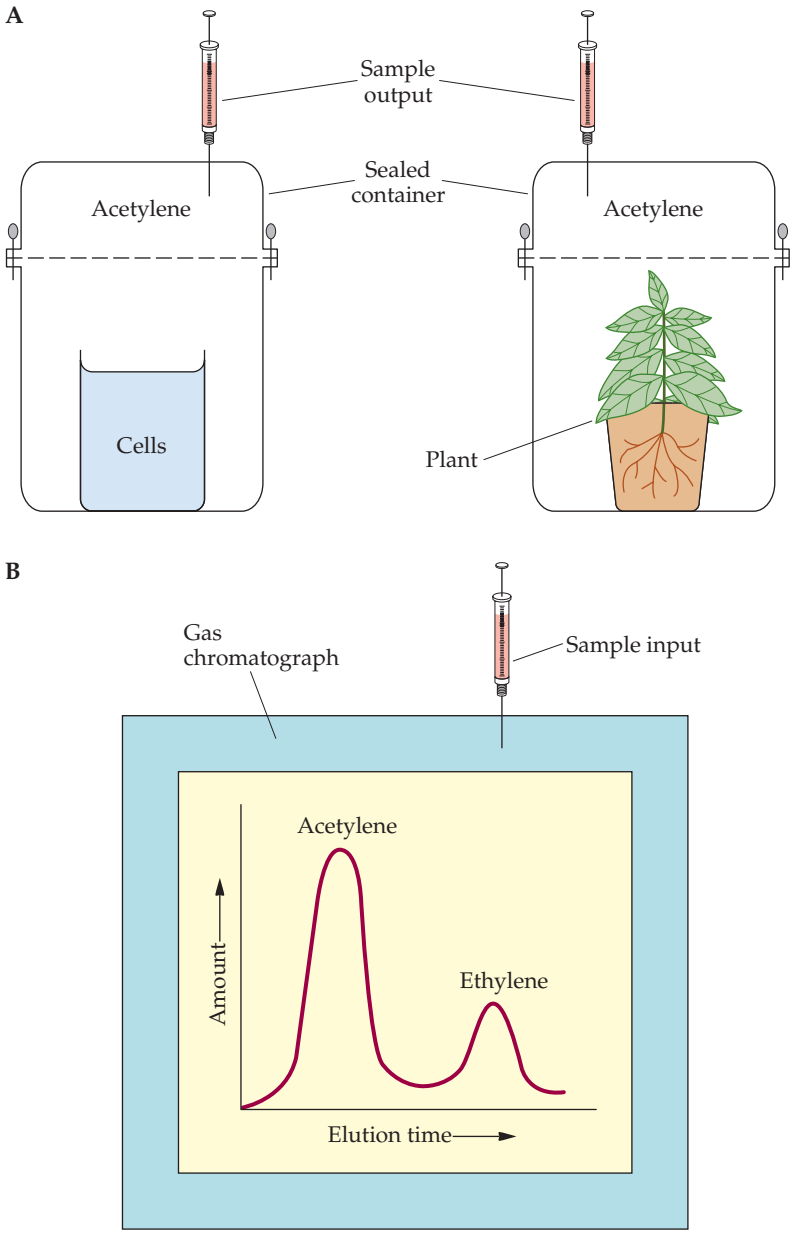


FIGURE 15.18 Assay for nitrogenase activity based on the conversion of acetylene to ethylene. **(A)** The source of the nitrogenase enzyme, for example, bacteria growing in culture, bacteria associated with plant roots, or a purified enzyme preparation (not shown), is placed in a sealed container under an atmosphere of acetylene. **(B)** Samples are periodically withdrawn from the sealed container, and the levels of acetylene and ethylene are measured by gas chromatography. The extent of nitrogenase activity is proportional to the amount of ethylene produced.

4. The transformed *K. pneumoniae* cells are selected for the acquisition of the Nif$^+$ phenotype by plating them onto a minimal medium that does not contain a source of fixed nitrogen. The only cells that grow under these conditions are Nif$^-$ *K. pneumoniae* cells containing a plasmid encoding and expressing the protein that is either missing or nonfunctional in the Nif$^-$ mutant.

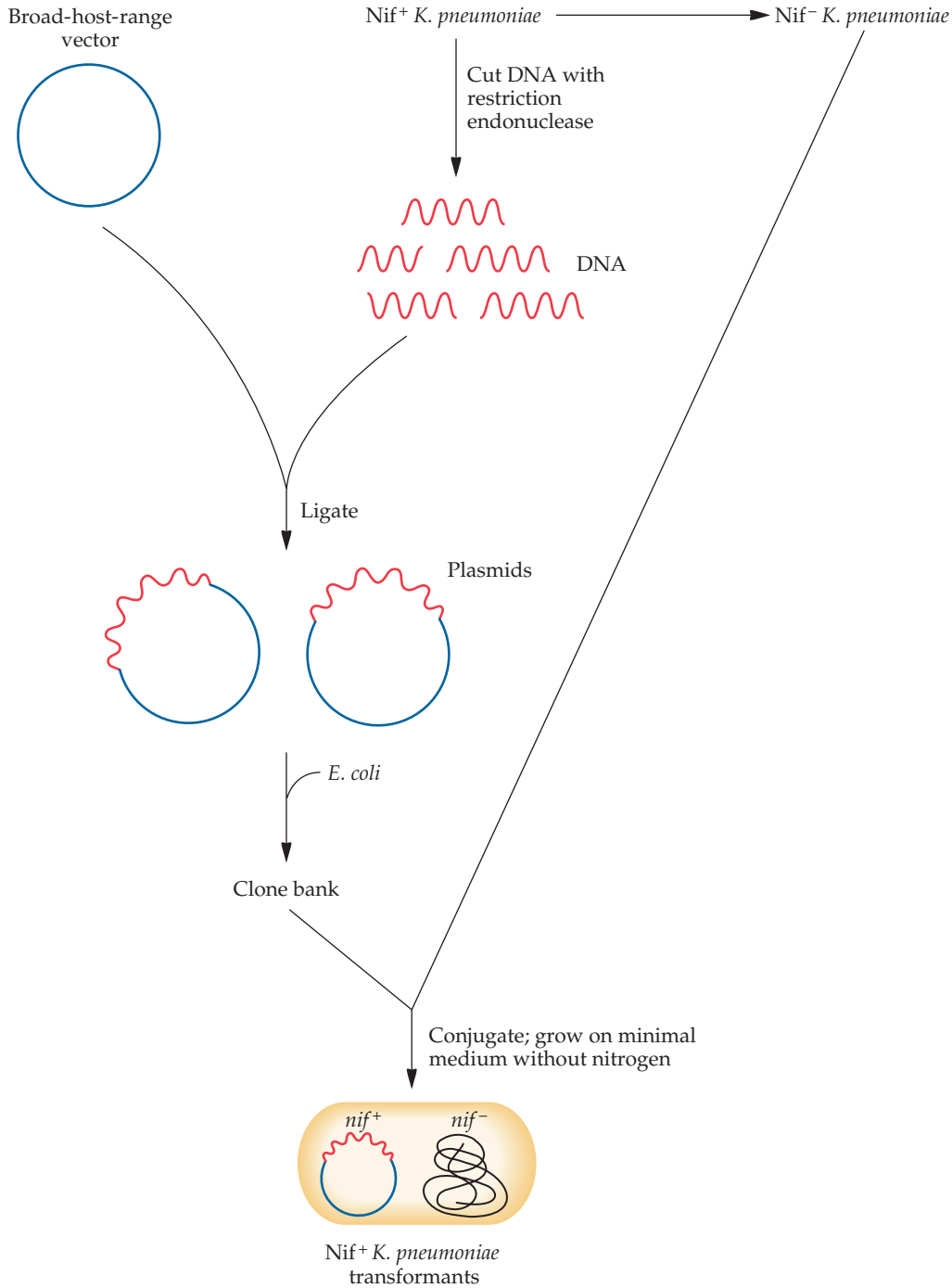


FIGURE 15.19 Procedure for isolating *nif* genes by genetic complementation. A clone bank that was constructed with Nif+ *K. pneumoniae* DNA is used to complement a Nif− *K. pneumoniae* strain. Transformants are selected for growth on minimal medium that does not contain fixed nitrogen.

The DNA fragment in the plasmid that complements the Nif− chromosomal mutation contains a *nif* gene that can be characterized more thoroughly and used to isolate other *nif* genes.

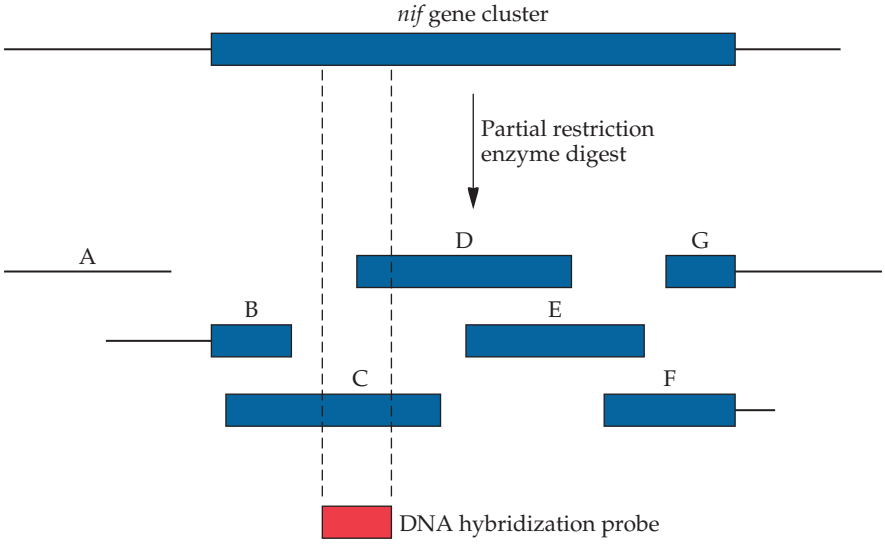


FIGURE 15.20 Partial restriction endonuclease digestion of the region of the chromosomal DNA encoding *nif* genes. The *nif* gene that was isolated by genetic complementation is used as a DNA hybridization probe (red). The probe and the DNA fragments from the partial digest are ordered as they would be in the chromosomal DNA. Colony hybridization of a clone bank consisting of the illustrated fragments, using the probe shown, would be expected to yield clones containing plasmids with fragments C and D. Using fragments C and D separately as probes to screen the same clone bank would be expected to yield clones containing plasmids with fragments B and E, respectively. Eventually, in this way, a large contiguous section of the host chromosome is isolated as a set of overlapping clones.

Two approaches have been used to isolate other genes that are involved in the nitrogen fixation process. First, the *K. pneumoniae* clone bank has been used to complement a series of independently derived Nif$^-$ mutants, increasing the likelihood that in each case a different *nif* gene will be isolated. Second, isolated *nif* genes have been used as DNA hybridization probes, which have then been used to screen a *K. pneumoniae* chromosomal DNA clone bank that carries large (7- to 10-kb) inserts (Fig. 15.20). The premise behind the latter scheme is based on the observation that in prokaryotic organisms many of the genes involved in one pathway are clustered on the chromosomal DNA and are often arranged in operons. Thus, DNA hybridization enables investigators to identify clones containing additional *nif* genes that are adjacent to the sequence initially isolated.

As a result of a considerable amount of research, the entire set of *nif* genes from *K. pneumoniae* has been isolated and characterized. These genes are arranged in a single cluster that occupies approximately 24 kb of the bacterial genome (Fig. 15.21). The cluster contains seven separate operons that together encode 20 distinct proteins (Table 15.8). All of the *nif* genes must be transcribed and translated in a concerted fashion, under the regulatory control of the *nifA* and *nifL* genes, to produce a functional nitrogenase. The NifA protein is a positive regulatory factor. It turns on the transcription of all of the *nif* operons except its own by binding to a specific DNA sequence (5′-TGT-N_{10}-ACA-3′) that is part of each promoter of each *nif* operon. There is a site on the DNA approximately 80 to 150 nucleotides upstream of each transcriptional start site where the NifA protein binds.

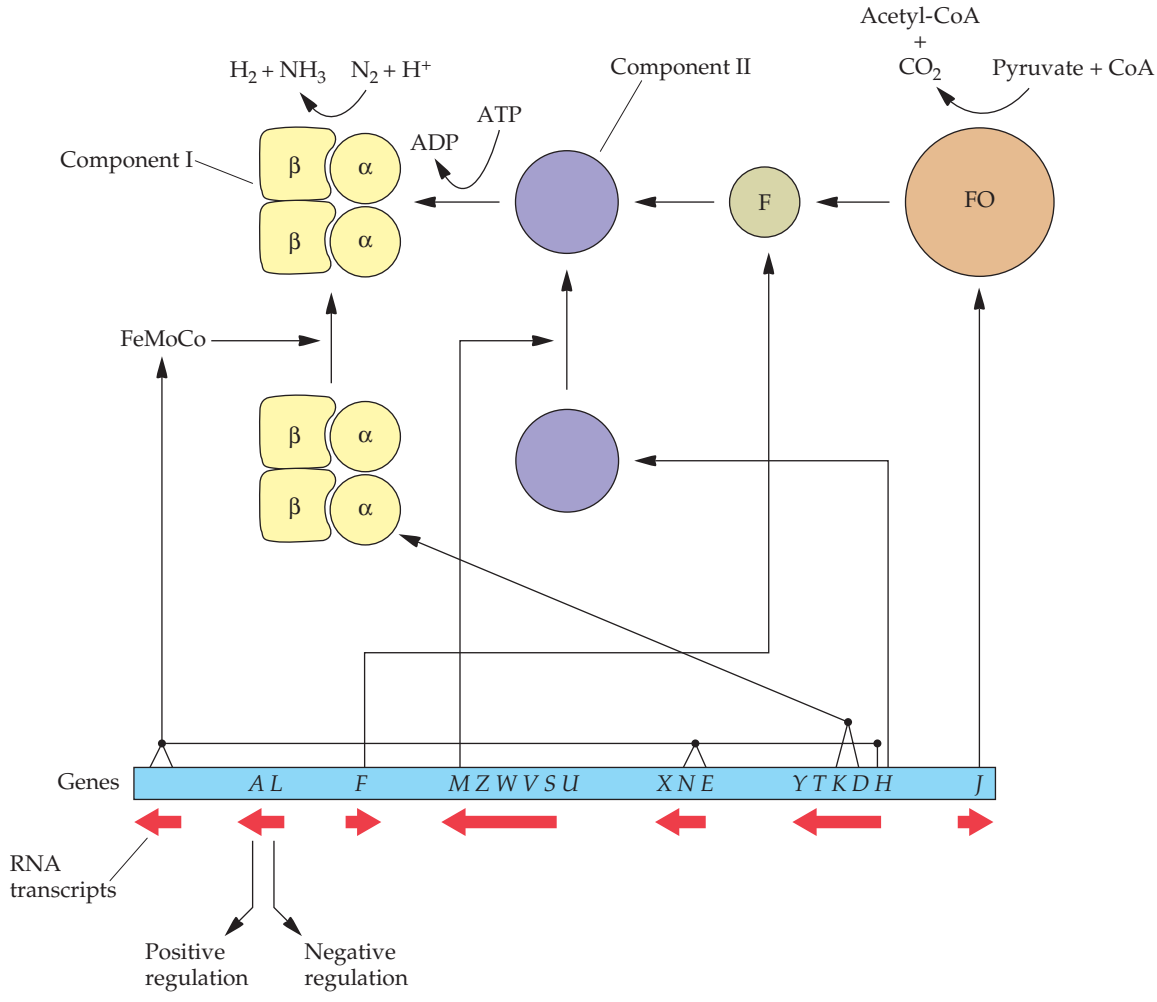


FIGURE 15.21 Arrangement of the genes in the *K. pneumoniae nif* gene cluster and some of the functions that they encode. The *nif* genes are represented by italic uppercase letters; each red arrow below sets of these letters denotes a specific *nif* operon and the direction of its transcription. The arrows pointing up and away from the gene designations show how some of the various gene products participate in nitrogen fixation. F, flavodoxin molecule; FO, pyruvate–flavodoxin oxidoreductase; FeMoCo, the iron–molybdenum cofactor; CoA, coenzyme A.

The DNA-bound NifA protein then interacts with a specific transcription initiation protein called sigma-54 (σ^{54}) before transcription from the *nif* promoter is initiated. The NifL protein is a negative regulatory factor. In the presence of either oxygen or high levels of fixed nitrogen, it acts as an antagonist of the NifA protein and, as a result, turns off the transcription of all other *nif* genes.

The bacterium *K. pneumoniae* does not make a major contribution to the overall global biological fixation of nitrogen. Therefore, to genetically engineer nitrogen fixation in soil bacteria that are more important in promoting plant growth, other *nif* genes have been cloned and characterized. To do this, the *nif* genes from *K. pneumoniae* have been used as DNA hybridization probes to isolate *nif* genes from clone banks of other diazotrophic microorganisms. Most diazotrophic organisms have a similar array of genes

encoding their nitrogen-fixing apparatus, and the DNA sequences of these genes do not vary much from one organism to another.

It may be possible to increase the amount of nitrogen fixed by diazotrophic organisms by manipulating the *nifA* and *nifL* genes. After researchers genetically engineered extra copies of the *nifA* gene into a *Sinorhizobium meliloti* strain, alfalfa plants that were inoculated with this transformant grew larger and produced more biomass than plants that were treated with the nontransformed strain. Similarly, it may be possible to engineer the *nifL* gene so that the NifL protein, the negative regulator, is less sensitive to the presence of fixed nitrogen. With this kind of deregulation, an organism would fix more nitrogen for its plant partner. However, not all nitrogen-fixing organisms have a NifL protein, so this sort of manipulation may be limited to only certain bacterial strains. In some organisms, the essential regions of NifL are part of NifA. Moreover, increasing the amount of nitrogen that an organism can fix also increases the amount of energy, usually in the form of fixed carbon, that is needed to power its metabolism. Consequently, an engineered microorganism that can fix a higher than normal level of nitrogen may lose its effectiveness as a plant growth-promoting agent because of a diminished growth rate.

Because of the complexity of nitrogen fixation by microorganisms, the simple addition of one or two *nif* genes will not confer on a nondiazotrophic recipient cell the ability to fix nitrogen. Moreover, genetic modification of plants with the entire 24-kb *nif* gene cluster would not be effective because the normal level of oxygen in the host cell would inactivate nitrogenase, and if this level were reduced, the host plant cell would probably die. In addition, the engineering of nitrogen fixation in plant cells requires resolving major, if not insurmountable, transcriptional, translational, and regulatory problems. For example, it is difficult to conceive how the regulation of nitrogen fixation could be achieved, since there are no plant promoters that respond to the NifA protein. Consequently, *nif* genes would not be turned on in such a transgenic plant. Each of the *nif* genes would also have to be under the control of separate promoters because plant cells cannot process multigene transcripts. The introduction of a functional nitrogen fixation capability into plants is therefore extremely unlikely.

TABLE 15.8 *K. pneumoniae* genes involved in nitrogen fixation and the functions of the proteins that they encode

nif gene	Function
D	Nitrogenase component I α subunit
K	Nitrogenase component I β subunit
H	Nitrogenase component II
F	Flavodoxin
J	Pyruvate:flavodoxin oxidoreductase
Q, B, N, E, V	FeMoCo synthesis
M	Processing of dinitrogenase reductase
A	Positive activator
L	Negative regulator
S	Maturation of component I
W, Z, T, Y, U, X	Other, less well-defined functions

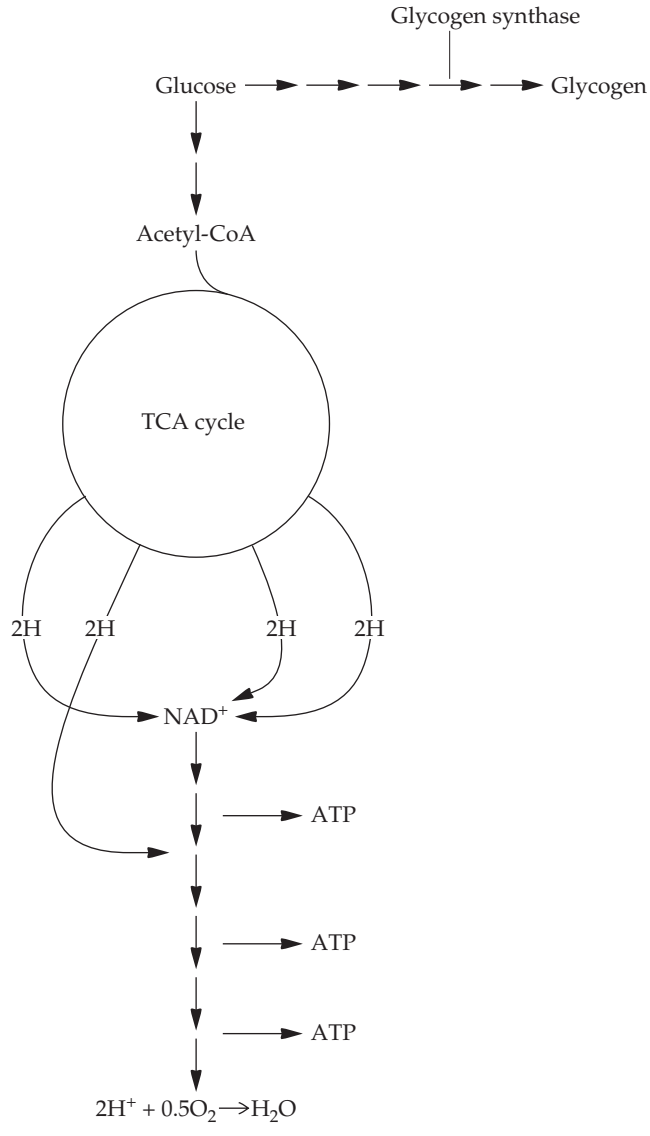


FIGURE 15.22 Schematic representation of the partitioning of rhizobial intracellular glucose between glycogen synthesis and respiration. A mutation in the gene for glycogen synthase prevents glycogen from being synthesized, so that all of the glucose enters the tricarboxylic acid (TCA) cycle. In this cycle, the acetyl group of acetyl coenzyme A (CoA) is enzymatically degraded to form carbon dioxide and hydrogen. The hydrogen (or the corresponding electrons) is fed into the electron transport chain, and a large portion of the energy released is conserved by the phosphorylation of ADP to ATP. The ATP is then available to "power" a large number of energy-requiring metabolic processes, including nitrogen fixation.

Engineering Improved Nitrogen Fixation

Engineering oxygen levels. The concentration of oxygen is a critical factor in determining the amount of nitrogen that is fixed by a rhizobial strain. On one hand, oxygen is inhibitory to nitrogenase and is a negative regulator of *nif* gene expression. On the other hand, oxygen is required for bacteroid respiration. This conundrum can be resolved by the introduction of leghe-

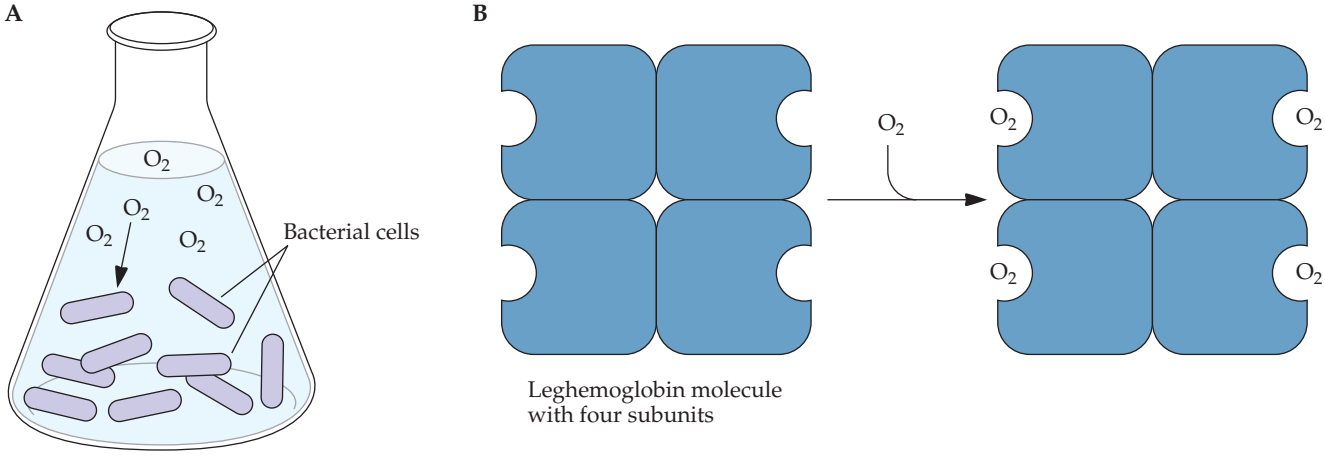
moglobin, which binds free oxygen tightly (Fig. 15.23) so that both the transcription of *nif* genes and the functioning of nitrogenase can proceed unimpaired. In fact, the addition of exogenous leghemoglobin to isolated bacteroids results in a dramatic increase in nitrogenase activity. Thus, it is possible to engineer more efficient strains of *Rhizobium* by overproducing leghemoglobin. Alternatively, since the globin portion of leghemoglobin is produced by the plant, it may be more efficient to transform rhizobial strains with genes encoding a bacterial equivalent of leghemoglobin.

Following the transformation of a strain of *Rhizobium etli* with a broad-host-range plasmid carrying the *Vitreoscilla* sp. (a gram-negative aerobic or microaerophilic bacterium) hemoglobin gene at low levels of dissolved oxygen (0.25 to 1.0%) in the growth medium, the rhizobial cells had a two- to threefold-higher respiratory rate than the nontransformed strain. These data suggest that free-living *R. etli* with a *Vitreoscilla* sp. hemoglobin gene may have a competitive advantage over nontransformed rhizobial strains in soil (which usually has a low level of oxygen). As has been observed in the laboratory for numerous other free-living bacteria, the hemoglobin-containing strain can grow to a greater extent because it is able to sequester oxygen and provide it to the reactions, where it is necessary for bacterial metabolism (nitrogenase activity is at its peak at this time).

In greenhouse experiments, when bean plants were inoculated with either nontransformed or hemoglobin-containing *R. etli*, the plants inoculated with the hemoglobin-containing strain had approximately 68% more nitrogenase activity. This difference in nitrogenase activity leads to a 25 to 30% increase in leaf nitrogen content at 40 to 50 days after infection and a 16% increase in the nitrogen content of the seeds that are produced. Thus, the expression of a bacterial hemoglobin gene may be advantageous to *Rhizobium* bacteria both when they are free-living and when they are in bacteroids as part of a symbiotic relationship with their host plant.

Modulating *nifH* and poly-β-hydroxybutyrate. In Mexico, most of the bean plants (the second most important crop in Mexico) are nodulated by *R. etli*.

FIGURE 15.23 **(A)** *R. etli* cells engineered to express a *Vitreoscilla* sp. hemoglobin gene bind low levels of dissolved oxygen, either from solution or from the soil. **(B)** The *Vitreoscilla* sp. hemoglobin protein binding to dissolved oxygen.



The most common strain of *R. etli* encodes three copies of the nitrogenase reductase (*nifH*) gene, each under the transcriptional control of a separate promoter. To increase the amount of nitrogenase, the strongest of the three *nifH* promoters (i.e., *PnifHc*) was coupled to the *nifHcDK* operon, which encodes the nitrogenase structural genes (where *nifHc* is one of the three *nifH* genes in this bacterium). The *nifHc* promoter is typically induced during nodule development. The *PnifHc–nifHcDK* construct was cloned into a broad-host-range plasmid and introduced into the wild-type strain of *R. etli*. The net result of this genetic manipulation was a significant increase in nitrogenase activity, plant dry weight, seed yield, and the nitrogen content of the seeds (Table 15.9). Moreover, this genetic manipulation worked as well or better when the *PnifHc–nifHcDK* construct was introduced into the large Sym plasmid from *R. etli* that contains all of the genetic information for nodulation and nitrogen fixation.

Biological nitrogen fixation requires a large amount of energy in the form of ATP. Thus, any mutation or genetic manipulation that increases the flux of carbon sources consumed by a bacterium through the citric acid cycle should be beneficial for nitrogen fixation (Fig. 15.22 and 15.24). This is because metabolism of glucose via the citric acid cycle results in the production of ATP. Consistent with this principle, it was observed that expression of the *PnifH$_c$–nifH$_c$DK* construct in a poly-β-hydroxybutyrate-negative strain of *R. etli* enhanced plant growth to an even greater extent than when this construct was expressed in a wild-type poly-β-hydroxybutyrate-positive strain. Finally, since no foreign genes were introduced into *R. etli*, the scientists who constructed these strains hope that the regulatory bodies in their country will view the manipulated strains as benign and approve them for widespread environmental use.

Hydrogenase

An undesirable side reaction of nitrogen fixation is the reduction of H^+ to H_2 (hydrogen gas) by nitrogenase. Energy in the form of ATP is wasted on the production of hydrogen, which is eventually lost to the atmosphere. Because of this side reaction, only 40 to 60% of the electron flux through the nitrogenase system is transferred to N_2, thereby significantly lowering the overall efficiency of the nitrogen-fixing process. Theoretically, if H_2 could be recycled to H^+, the extent of energy loss could be diminished and the nitrogen-fixing process would become more efficient. It is probably impossible to prevent this side reaction directly, because it is a consequence of the chemistry of the active site of the nitrogenase; hence,

TABLE 15.9 Symbiotic performance of genetically modified strains of *R. etli* in concert with common beans *(Phaseolus vulgaris)*

Bacterial strain	Nitrogenase activity (μmol of ethylene/h/g of nodule)	Plant dry wt (g/plant)	Seed yield (g/plant)	Seed N content (mg of N/g of seed)
Wild type	64.5	0.54	1.43	33.9
Wild type + extra *nifHDK*	72.7	0.66	1.56	41.4
Wild type + *nifH$_c$*	77.3	0.75	1.73	31.2
Wild type + *PnifH$_c$-nifH$_c$DK*	108.2	0.81	2.50	43.6

Adapted from Peralta et al., *Appl. Environ. Microbiol.* **70:**3272–3281, 2004.

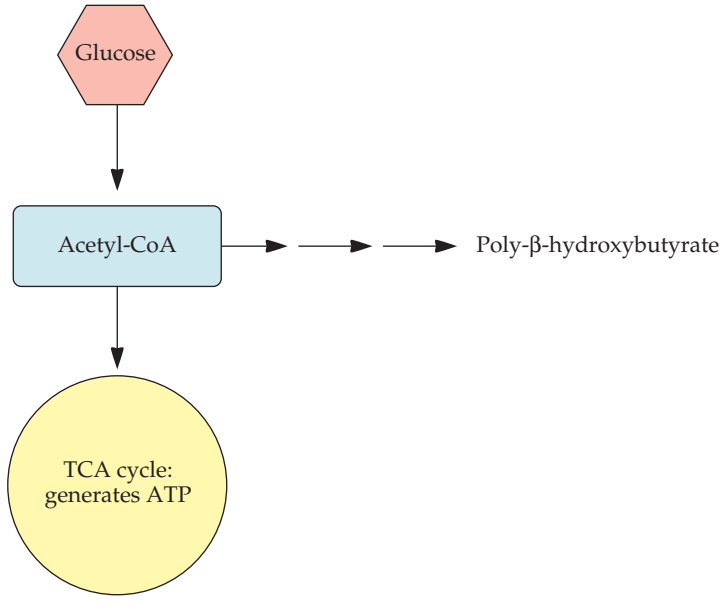


FIGURE 15.24 Conversion of glucose to acetyl coenzyme A (CoA), which can either enter the tricarboxylic acid (TCA) cycle, with the concomitant production of ATP, or be converted into the carbon storage polymer poly-β-hydroxybutyrate.

blocking the side reaction by altering nitrogenase would concomitantly inhibit nitrogenase activity.

Hydrogen Metabolism

In the mid-1970s, it was discovered that some strains of *B. japonicum* could use hydrogen as an energy source for growth under microaerophilic (low-oxygen-concentration) conditions. These strains have an enzyme called hydrogenase that is able to take up H_2 from the atmosphere and convert it into H^+ (Fig. 15.25). Experiments were undertaken to test whether the presence of hydrogenase in *B. japonicum* had an impact on the growth of soybean plants. Plants inoculated with strains that produce hydrogenase (Hup^+) had more biomass and nitrogen than plants that were treated with non-hydrogenase-producing (Hup^-) strains, despite higher levels of nitrogenase activity in the Hup^- strains (Table 15.10). From this and similar experiments, it was concluded that the presence of a hydrogen uptake system in a symbiotic diazotroph, such as *B. japonicum,* improves its ability to stimulate plant growth, presumably by binding and then recycling the hydrogen gas that is formed inside the nodule by the action of nitrogenase (Fig. 15.25). Within the nodule, the contribution of atmospheric hydrogen is negligible.

Although it is clearly beneficial to the plant to obtain its nitrogen from a symbiotic diazotroph that has a hydrogen uptake system, this trait is not common in naturally occurring rhizobial strains. In one study, it was found that the majority of naturally occurring *Rhizobium* and *Bradyrhizobium* strains examined were Hup^- (Table 15.11). In that study, the data were based on a small number of strains for each species except *B. japonicum*, for which over 1,400 strains were assayed. The conclusion that can be drawn

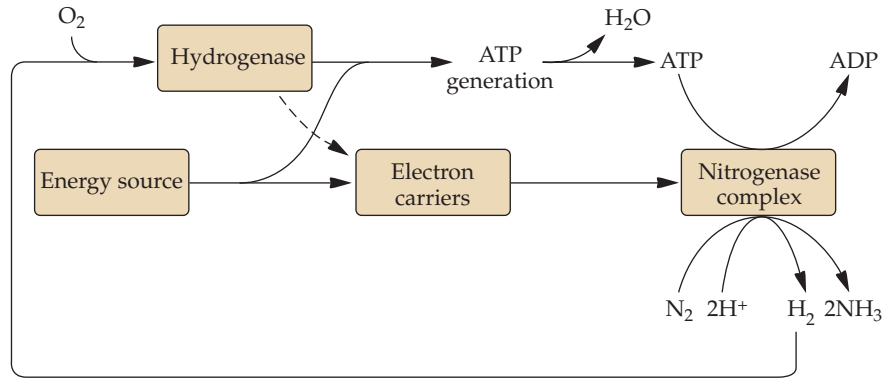


FIGURE 15.25 Recycling of the hydrogen gas that is produced as a by-product of nitrogen fixation. Hydrogen is generated by nitrogenase at the expense of ATP, but by using this pathway, the hydrogen can be recaptured by hydrogenase.

from this work is that commercial Hup^- rhizobial strains are prime candidates for transformation to a Hup^+ phenotype.

Genetic Engineering of Hydrogenase Genes

Although a considerable amount of effort has been directed over the past 30 years or so toward studying hydrogenases from both diazotrophic and nondiazotrophic microorganisms, an in-depth understanding of the structures and functions of these enzymes remains elusive. Many organisms have more than one hydrogenase, and many hydrogenases consist of more than a single polypeptide chain. Some hydrogenases are active only in the uptake of hydrogen from the atmosphere, whereas others, depending on the conditions, can also synthesize hydrogen. One result of this complexity is that the conversion of a Hup^- strain of *Rhizobium* into a Hup^+ strain may not be readily achieved by the introduction of just any hydrogenase gene. Rather, the introduced gene(s) must encode all of the enzyme's subunits and must be able to interact with the appropriate electron transport molecule within the host organism.

The most common strategy for isolating hydrogenase genes has been genetic complementation. The first hydrogenase gene to be isolated was

TABLE 15.10 Relative enzyme activities and growth-stimulating performance of a parental Hup^+ *B. japonicum* strain (SR) and three Hup^- mutants (SR1, SR2, and SR3)

B. japonicum strain	Relative nitrogenase activity	Relative hydrogenase activity	Relative plant dry weight	Relative nitrogen content
SR	1.00	1.00	1.00	1.00
SR1	1.27	0.01	0.81	0.93
SR2	1.13	0.01	0.74	0.91
SR3	1.23	0.01	0.65	0.85

Adapted from Albrecht et al., *Science* **203:**1255–1257, 1979.

Nitrogenase activity was assessed by monitoring the amount of acetylene that was reduced to ethylene as a function of time. Hydrogenase activity was measured by means of a hydrogen electrode. Plant dry weight included the weights of both the leaf material and root material. The nitrogen content was calculated as the fraction of the dry weight of the plant that was nitrogen. All values have been normalized relative to the parental strain.

TABLE 15.11 Percentages of native *Rhizobium*, *Sinorhizobium*, and *Bradyrhizobium* strains that have functional hydrogen uptake systems (Hup^+)

Bacterium	Hup^+ strains (%)
Rhizobium leguminosarum bv. leguminosarum	9.3
Rhizobium leguminosarum bv. trifolii	0
Rhizobium leguminosarum bv. phaseoli	0
Sinorhizobium meliloti	21
Bradyrhizobium japonicum	21
Bradyrhizobium sp.	91

Adapted from Evans et al., *Annu. Rev. Microbiol.* **41:**335–361, 1987.

the gene for an *E. coli* membrane-bound hydrogenase, and it was selected by complementation of an *E. coli* mutant that did not express this activity.

After the work with *E. coli,* hydrogenase *(hup)* genes from *B. japonicum* were isolated from a clone bank of wild-type DNA constructed in the broad-host-range cosmid vector pLAFR1 by complementation of *B. japonicum* Hup^- mutants. The presence of a hydrogenase that takes up hydrogen from the atmosphere in the complemented Hup^- mutant strains was indicated by the ability of the active hydrogenase to reduce the dye methylene blue in a hydrogen atmosphere. More detailed studies of the *B. japonicum hup* genes showed that they were organized into at least two, and possibly three, transcriptional units covering approximately 20 kb of the genome and including 18 separate genes. Subsequent work on the *hup* genes from *Rhizobium leguminosarum* has indicated that these genes are similar in both DNA sequence and gene organization to the *hup* genes from *B. japonicum*. Thus, the isolated *hup* genes from *B. japonicum* may be used as DNA hybridization probes to select homologous genes from a clone bank of *R. leguminosarum*.

Following the isolation of *R. leguminosarum hup* genes, and despite the complexity of this system, it has been possible to use cosmid vectors to transfer a complete set of uptake hydrogenase genes from a Hup^+ strain of *R. leguminosarum* to a Hup^- strain (Table 15.12). Plants treated with *R. leguminosarum* that had been transformed to Hup^+ grew larger and contained more nitrogen than the plants inoculated with the Hup^- parental strain (Table 15.12). Although hydrogenase genes have not received as much attention as *nif* genes, this simple gene transfer experiment is a convincing demonstration of the use of genetic manipulation to improve the ability of a diazotroph to stimulate plant growth.

More recently, one group of scientists modified the *hup* gene promoter in *R. leguminosarum* and in the process engineered a more efficient rhizobial

TABLE 15.12 Plant growth and nitrogen assimilation after the introduction of *hup* genes into a Hup^- strain of *R. leguminosarum*

Bacterial phenotype	Relative plant dry weight	Relative nitrogen amount	Relative leaf area	Relative nitrogen concentration
Hup^-	1.00	1.00	1.00	1.00
Hup^+	1.35	1.52	1.53	1.15

Adapted from Brewin and Johnston, U.S. patent 4,567,146, January 1986.
The data have been normalized relative to the Hup^- parental strain.

strain. In *R. leguminosarum,* 18 genes are associated with hydrogenase activity. There are 11 *hup* genes (Fig. 15.26) responsible for the structural components of the hydrogenase, the processing of the enzyme, and electron transport. There are also seven *hyp* (hydrogenase pleitropic) genes, which are involved in processing the nickel that is part of the active center of the enzyme. The *hup* promoter is dependent on the NifA protein (which is also required to activate the synthesis of *nif* genes), so that *hup* genes are expressed only within bacteroids. On the other hand, the *hyp* genes are transcriptionally regulated by an FnrN-dependent promoter, which is turned on by low levels of oxygen. Thus, the *hyp* genes are expressed both in bacteroids and microaerobically. By modifying the chromosomal DNA of *R. leguminosarum* and exchanging the *hup* promoter for an FnrN-dependent promoter (Fig. 15.26), a derivative of the original bacterium with an increased level of hydrogenase was created (Table 15.13). The engineered *R. leguminosarum* strain displayed a twofold increase in hydrogenase activity compared to the wild type, and no discernible amount of hydrogen gas was produced as a by-product of nitrogen fixation. This is expected to make this strain of *R. leguminosarum* much more effective at promoting plant growth and increasing plant nitrogen content. Moreover, regardless of whether nickel was added to the system, the amount of hydrogen evolved from nitrogen-fixing nodules was extremely low, indicating that virtually all of the hydrogen produced by nitrogenase was recycled. The reason that nickel was added to this system is that in many soils, the availability of nickel limits hydrogenase activity. In some soils, the level of nickel is so low that even if a naturally occurring strain contains hydrogenase genes, the hydrogenase activity may be so low as to be ineffectual. On

FIGURE 15.26 Replacement of the NifA-dependent *R. leguminosarum hup* promoter with an FnrN-dependent promoter. The *R. leguminosarum hyp* gene cluster is already under the transcriptional control of an FnrN-dependent promoter.

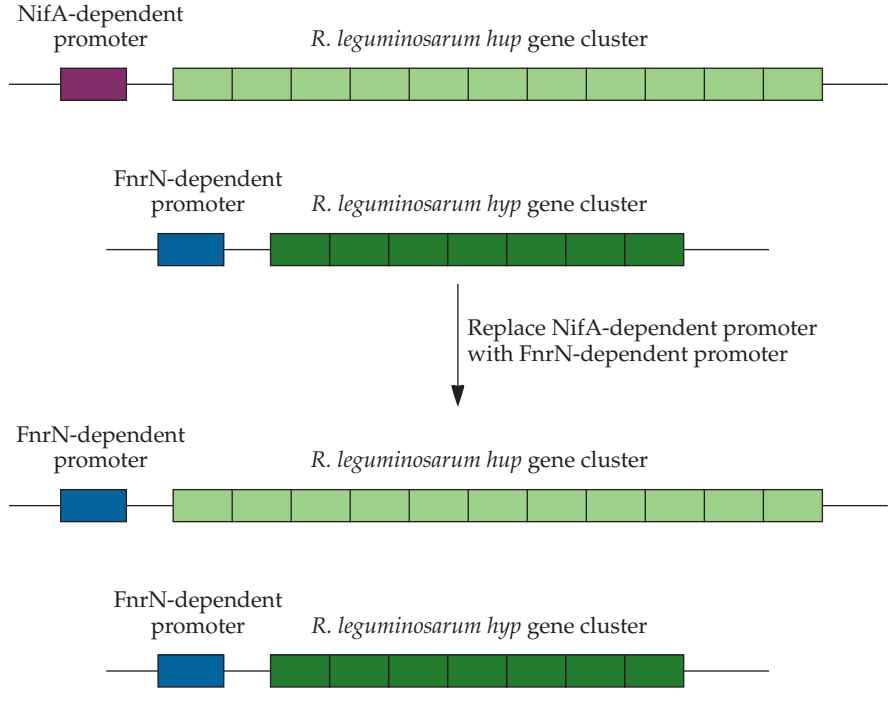


TABLE 15.13 Biological activities of wild-type and engineered (replacing the *hup* promoter with an FnrN-dependent promoter) strains of *R. leguminosarum*

Strain	Bacteroid hydrogenase activity (nmol of H_2/h/mg of protein)		Nodule H_2 evolution (mmol/H_2/g [fresh weight] of nodule)	
	-Ni	+Ni	-Ni	+Ni
Wild type	1,080	2,930	2.13	<0.25
Engineered strain	2,210	5,060	0.34	<0.25

Adapted from Brito et al., *Appl. Environ. Microbiol.* **68:**2461–2467, 2002, and Ureta et al., *Appl. Environ. Microbiol*. **71:**7603–7606, 2005.

the other hand, when the engineered *R. leguminosarum* was tested with various field soils, hydrogenase overproduction invariably overcame the limitation of low nickel levels, with the net result that the amount of fixed nitrogen, and hence plant productivity, was greater.

Nodulation

Competition among Nodulating Organisms

A major goal of agricultural biotechnology research is the development, by genetic manipulation, of *Rhizobium* strains that can increase plant productivity more effectively than naturally occurring strains. Many commercial inoculant strains that have been developed by mutation and selection to be superior nitrogen fixers are not very effective at establishing nodules on host plant roots when placed in competition with *Rhizobium* strains that are already present in the soil. Conversely, although many of the strains that are indigenous to the soil are highly successful in establishing nodules in competitive situations, they are not especially efficient at nitrogen fixation. Therefore, to make use of the commercial inoculant strains, either the nodulation capability of these strains must be enhanced or indigenous rhizobial strains must be inhibited.

Studies were undertaken to determine the genetic basis of this "competitiveness" with the aim of adding these particular genes to the strains that are used as inoculants. The nature of the competitive advantage of soil rhizobial species is not known, but it was reasoned that the indigenous bacteria might be more efficient at nodulation and, as a consequence, might prevent an inoculated strain from becoming established and forming its own nodules.

Genetic Engineering of Nodulation Genes

When scientists first attempted to isolate nodulation *(nod)* genes, the absence of any specific information about the biochemical or genetic basis of nodulation meant that a strategy had to be devised for the identification of the genes. Therefore, once again, genetic complementation was used. Nodulation-defective (Nod^-) mutants of *S. meliloti* were transformed with a clone bank of wild-type chromosomal DNA from *S. meliloti,* and those colonies that had acquired the ability to nodulate alfalfa roots were isolated (Fig. 15.27). More specifically, the steps of the procedure were as follows.

1. A clone bank of wild-type (Nod^+) *S. meliloti* was constructed by partial digestion of *R. meliloti* DNA with EcoRI and insertion into

the unique EcoRI site of the broad-host-range cosmid pLAFR1, which can carry up to 40 kb of cloned DNA.

2. The clone bank was packaged into bacteriophage λ, introduced into *E. coli,* and subsequently transferred to Nod⁻ mutants of *S. meliloti* by conjugation. The vector carries a tetracycline resistance gene that can be used as a selectable marker in both *E. coli* and *S. meliloti*.
3. After conjugation, suspensions of 200 to 300 transformed *S. meliloti* cells were tested for the ability to nodulate sterile alfalfa plants, with the expectation that only those transformants that carried and expressed a gene that complemented the nodulation defect in the *S. meliloti* host would produce nodules.
4. The bacteria that formed nodules on the test plants were recovered from within the nodules. These bacteria were then grown in culture and used to retrieve the vector carrying the complementing gene. The specific portion of the large insert DNA that carried the complementing gene was then subcloned onto another plasmid vector and analyzed further.
5. Once a single nodulation gene was identified, it was used as a DNA hybridization probe to identify adjacent regions of *S. meliloti* chromosomal DNA in a genomic library (Fig. 15.20).

The complete repertoire of nodulation genes from *S. meliloti* has been characterized. Detailed biochemical and genetic studies have revealed that nodulation and its regulation are complex processes that require the functioning of a large number of genes (Table 15.14). Some of the nodulation genes are highly conserved (common) among nodulating microorganisms, and others are species specific. The *nod* genes are grouped into three separate classes: common genes, host-specific genes, and the regulatory *nodD* gene. Thus, for example, the *nodABC* genes are common to all *Rhizobium* species and are structurally interchangeable. In most species, the *nodABC* genes are found on a single operon.

A number of events are now known to occur during nodulation. First, the *nodD* gene product, which is constitutively expressed, recognizes and binds to a flavonoid molecule, which is excreted by the roots of the potential host plants. Flavonoids are a class of plant phenolic molecules with a basic structure that consists of 15 carbons arranged as two aromatic rings connected by a 3-carbon bridge. They perform a number of different functions for the plant, such as pigmentation and defense against fungi or insects. The binding of flavonoids to the NodD gene product is one of the major determinants of rhizobial host specificity, because each rhizobial species recognizes only a limited number of flavonoid structures and each plant species produces its own specific set of flavonoid molecules (Table 15.15). In a limited number of instances, other small organic molecules, such as aldonic acids and betaines, that are exuded by plant roots or germinating seeds and are present in large amounts can interact with the NodD protein. Some strains, such as *R. leguminosarum* biovar (bv.) trifolii, have a very narrow host range, responding to only a few kinds of flavonoids, while others, such as *Rhizobium* sp. strain NGR234, have a very broad host range and respond to a much larger number of different flavonoids.

The binding of a flavonoid molecule activates the NodD gene product, presumably causing it to undergo a conformational change, and enables the flavonoid–NodD complex to attach to a nodulation promoter element

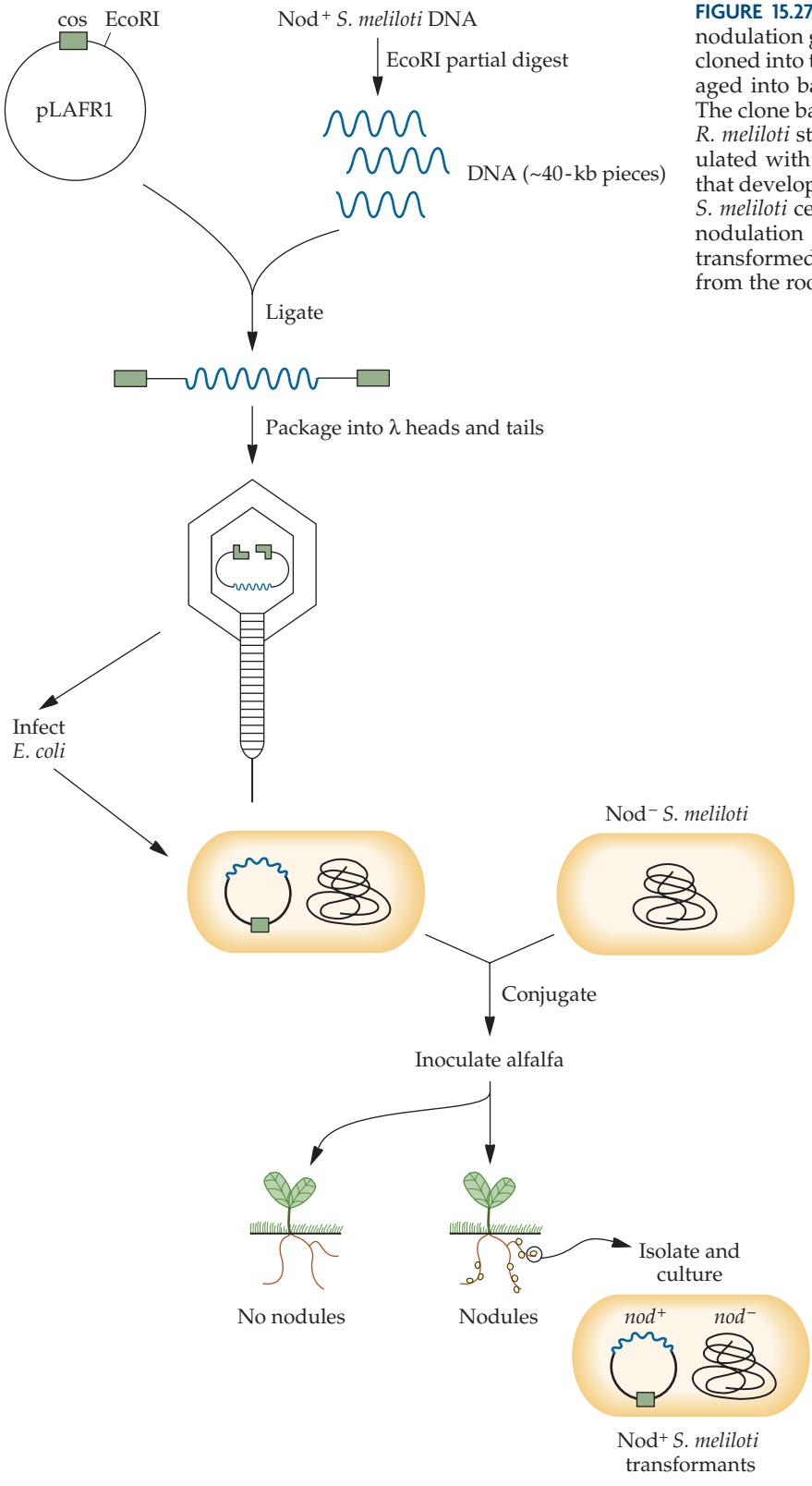


FIGURE 15.27 Procedure for the isolation of *S. meliloti* nodulation genes. The DNA from wild-type *S. meliloti* is cloned into the broad-host-range cosmid pLAFR1, packaged into bacteriophage λ, and introduced into *E. coli*. The clone bank is then transferred from *E. coli* to a Nod⁻ *R. meliloti* strain by conjugation. Alfalfa plants are inoculated with transformed Nod⁻ *S. meliloti*. Those plants that develop root nodules have been infected with Nod⁺ *S. meliloti* cells that presumably carry a complementing nodulation gene inserted into the cosmid vector. The transformed Nod⁺ *S. meliloti* cells can be isolated directly from the root nodules.

TABLE 15.14 Some *Rhizobium* nodulation gene products and their probable functions

Nod protein	Probable function
NodA	Common; cytoplasmic membrane; with NodB, stimulates cell division
NodB	Common; cytoplasmic membrane; with NodA, stimulates cell division
NodC	Common; outer membrane; chitin synthase
NodD	Common; positive transcriptional regulator; constitutive
NodE	Cytoplasmic membrane; β-ketoacyl synthase
NodF	Cytoplasmic; acyl carrier protein
NodG	Host-specific gene; dehydrogenase
NodH	Host-specific gene; sulfotransferase
NodIJ	Common; cytoplasmic membrane; capsular polysaccharide secretion
NodK	Affects onset of nodulation in some bradyrhizobia
NodL	Cytoplasmic membrane; acetyltransferase
NodM	D-Glucosamine synthetase
NodN	Unknown
NodO	Secreted, hemolysin
NodP	With NodQ; ATP sulfurylase
NodQ	With NodP; ATP sulfurylase
NodS	Methyltransferase
NodT	Outer membrane; secretion
NodU	Unknown
NodX	Cultivar specificity

Where biochemical or genetic evidence for the function of a particular protein is lacking, a possible function is assigned based on homology of the amino acid sequence to a protein of known sequence. Different rhizobial strains contain different subsets of these proteins. "Common" genes perform the same function in all species of rhizobia.

called a *nod* box. This promoter element is located upstream from all the nodulation genes except the *nodD* gene, and it activates the transcription of these genes.

The *nodABC* genes encode proteins that cause the plant root hair tips to swell and curl, an effect that is recognized as the initial step in the infection of the plant root by the bacterium. The bacteria synthesize an oligosaccha-

TABLE 15.15 Some legumes and the *nodD* gene inducers that they produce

Legume	Compound
Lupin *(Lupinus albus)*	Erythronic acid, tetronic acid
Alfalfa *(Medicago sativa)*	Stachydrine, trigonelline, luteolin, chrysoeriol, 4,4′-dihydroxy-2′-methoxychalcone, liquiritigenin, 7,4′-dihydroxyflavone
Clover *(Trifolium repens)*	7,4′-Dihydroxyflavone, geraldone, 4′-hydroxy-7-methoxyflavone
Common bean *(Phaseolus vulgaris)*	Delphinidin, kaempferol, malvidin, myricetin, petunidin, quercetin, eriodictyol, genistein, naringenin
Pea *(Pisum sativa)*	Apigenin, eriodictyol
Soybean *(Glycine max)*	Daidzein, genistein, coumesterol
Vetch *(Vicia sativa* subsp. *nigra)*	3,5,7,3′-Tetrahydroxy-4′-methoxyflavanone, 7,3′-dihydroxy-4′-methoxyflavanone, naringenin, 4,4′-dihydroxy-2′-methoxychalcone, liquiritigenin, 7,4′-dihydroxy-3′-methoxyflavanone, 5,7,4′-trihydroxy-3′-methoxyflavanone, 5,7,3′-trihydroxy-4′-methoxyflavanone naringenin

The inducers are released by either roots or germinating seeds.

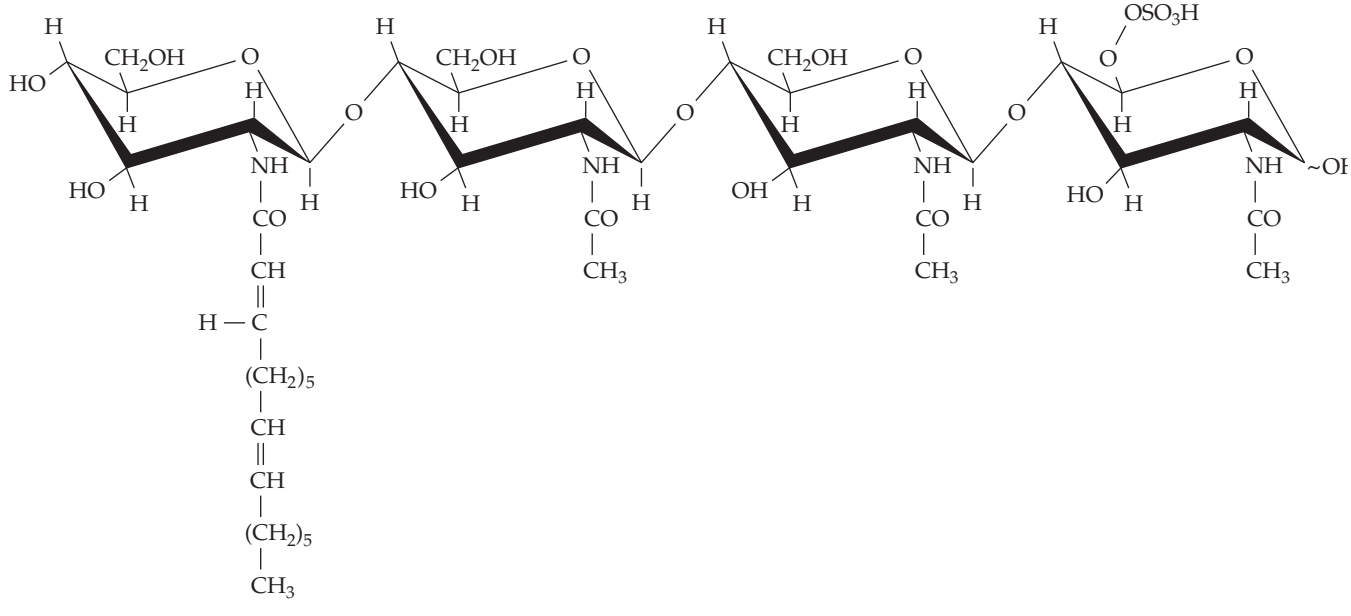
ride Nod factor that is modified by the NodH gene product and perhaps also by the NodQ and NodP gene products. This factor (Fig. 15.28) elicits in the plant a host-specific response that includes root hair curling and deformation and is essential for *Rhizobium* to induce nodules. After the initial change in the root morphology, the bacterium attaches to the root hair. Next, the bacterial cell penetrates the plant cell through an infection thread. Finally, depending on the bacterial strain, up to approximately 20 additional *nod* gene products are synthesized. These proteins, together with some plant-encoded proteins, contribute to the formation of the nodule.

DNA sequencing and computer analysis revealed that in a slow-growing variant of a *Bradyrhizobium* sp., the region of the DNA between the *nodD* and *nodABC* genes contained an open reading frame that the fast-growing form lacked. This open reading frame was designated *nodK*. When plants were inoculated with a *Bradyrhizobium* sp. that had a mutagenized *nodK* gene ($NodK^-$) and were compared with those treated with the wild-type strain ($NodK^+$), the onset of nodulation in the $NodK^-$ strain-treated plants was 5 days earlier, the nodulation number was doubled, and there was a 120% enhancement of plant yield.

To date, despite the fact that *nod* genes from several different rhizobial strains have been isolated and characterized, no simple genetic means has been devised for using *nod* genes to enable inoculated strains of *Rhizobium* to outcompete indigenous strains. Nevertheless, host specificity can be altered by the transfer of a *nodD* gene from a broad-specificity rhizobial strain to one with narrow specificity.

It has become clear that the process of nodulation is quite complicated. Thus, considerable additional effort will be required before it is possible to further enhance the competitiveness of rhizobial strains by genetic engineering.

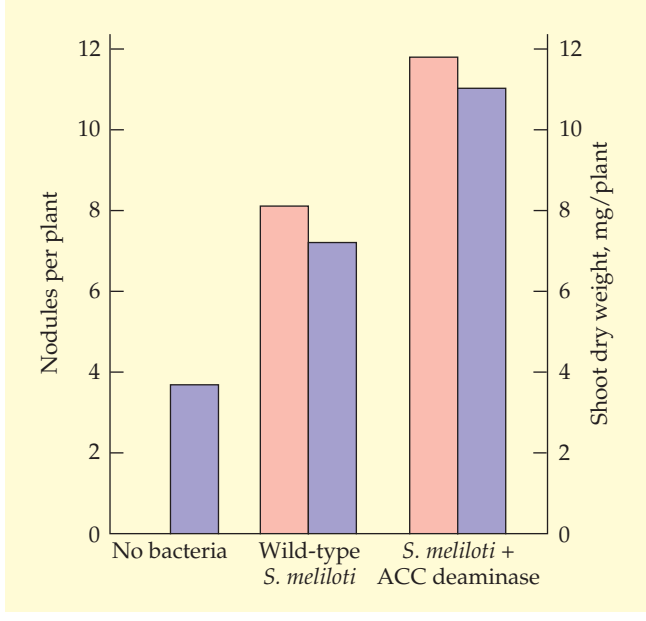
FIGURE 15.28 Proposed structure of a typical Nod factor, NodRm-1. This compound elicits a host-specific plant response that includes root curling and deformation.



Nodulation and Ethylene

Ethylene is often produced by plants following the initial stages of infection (eventually leading to nodule formation) by rhizobia. This small rise in the plant ethylene level is generally localized to a portion of the root and can inhibit, and therefore limit, subsequent rhizobial infection and nodulation. One way in which some strains of *Rhizobium* naturally increase the number of nodules that they can form on the roots of a host legume is to limit the rise in ethylene that occurs following the initial infection. Different *Rhizobium* species decrease ethylene levels either by synthesizing a small molecule called rhizobitoxin that chemically inhibits ACC synthase, one of the ethylene biosynthetic enzymes, or by producing ACC deaminase and removing some of the ACC before it can be converted to ethylene. The result of lowering the local level of ethylene is that both the number of nodules and the biomass of the plant are increased by 25 to 40%. Assays of isolated rhizobia indicate that in the field approximately 1 to 10% of rhizobial strains possess ACC deaminase. It should therefore be possible to increase the nodulation efficiency of *Rhizobium* strains that lack ACC deaminase by genetically engineering these strains with isolated *Rhizobium* ACC deaminase genes (and their regulatory regions). In fact, insertion of a single copy of an ACC deaminase gene from *R. leguminosarum* bv. viciae into the chromosomal DNA of a strain of *S. meliloti* that lacked this enzyme dramatically increased both the nodule numbers and biomass of host alfalfa plants (Fig. 15.29). While genetically engineered strains of *Rhizobium* may not be acceptable for use in the field in all jurisdictions at this time, as a result of this work, several commercial inoculant producers are already screening their more recently isolated *Rhizobium* strains for active ACC deaminase.

FIGURE 15.29 Increased ability of *S. meliloti* transformed with an ACC deaminase gene (and its regulatory region) to nodulate alfalfa (pink bars) and to promote plant growth (purple bars).



Phytoremediation

Scientists are currently trying to develop new and improved methods to deal with both inorganic and organic environmental contaminants. However, most of these procedures are either very expensive or not especially effective. One recently developed method of environmental cleanup is called phytoremediation. This procedure uses plants to remove or sequester hazardous substances from the environment or to destroy them. Phytoremediation of metals and other inorganic compounds may take one of several forms: phytoextraction, the absorption and concentration of metals from the soil into the roots and shoots of the plant; rhizofiltration, the use of plant roots to remove metals from effluents; or phytostabilization, the use of plants to reduce the spread of metals in the environment. Phytoremediation of organic compounds may occur by phytostabilization, i.e., reducing the spread of the organic material in the environment; phytostimulation, the stimulation of microbial biodegradation in the rhizosphere, the area around the roots of plants; or phytotransformation, the absorption and degradation of contaminants by the plant.

Following the testing of a large number of different plants, several plants that can naturally accumulate large amounts of metal have been identified and are being used to a limited extent for the phytoremediation of metals in the environment. These plants are called hyperaccumulators and are often found growing in soils with elevated metal concentrations. Unfortunately, plants that grow in the presence of very high concentrations of metals, even hyperaccumulating plants, are quite small. Depending upon the amount of metal at a particular site, it could take 15 to 20 years to completely remediate that site, even with hyperaccumulating plants. This is a time frame that is usually considered to be too long for practical application.

A number of different types of plants, such as many common grasses, as well as corn (maize), wheat, soybean, peas, and beans, are effective at stimulating the degradation of organic molecules in the rhizosphere. Typically, these plants all have extensive and fibrous root systems which form an extended rhizosphere. In addition to the biodegradation that takes place in the rhizosphere, several varieties of plants and trees can take up and degrade some organic contaminants. For example, plants with phytotransformation activity may contain nitroreductases, which are useful for degrading the explosive TNT (trinitrotoluene) and other nitroaromatics; dehalogenases for the degradation of chlorinated solvents and pesticides; and laccases that can degrade anilines, such as triaminotoluene.

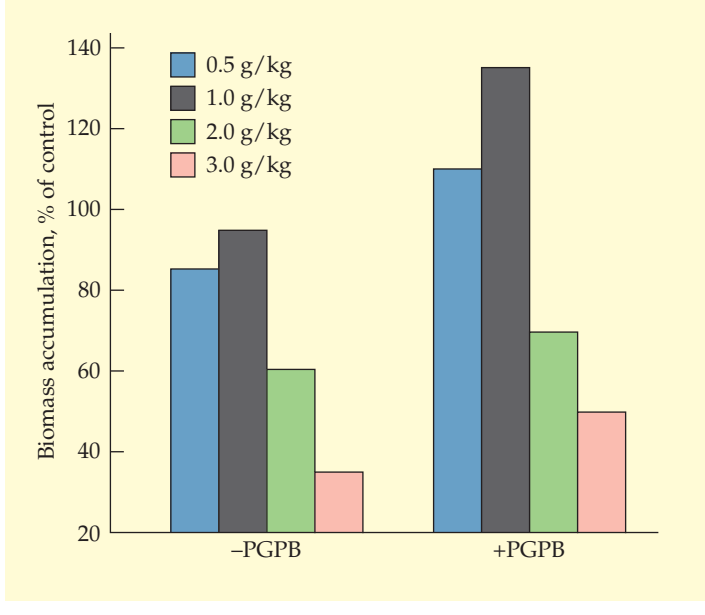
Engineering Strains That Facilitate Growth

Although using plants for remediation of persistent organic contaminants has advantages over other methods, many limitations exist for the large-scale application of this technology. For example, many plant species are sensitive to contaminants, so they grow slowly, and it is necessary to establish sufficient biomass for meaningful soil remediation. In addition, in most contaminated soils, the number of microorganisms is depressed, so that there are not enough bacteria either to facilitate contaminant degradation or to support plant growth. To remedy this situation, both degradative and plant growth-promoting bacteria may be added to the plant rhizosphere. Phytoremediation (i.e., degradation of organic compounds in the presence of plants) alone is not much faster than bioremediation (i.e., where biodeg-

radation of the organics is due to the activities of microorganisms independent of plants). On the other hand, cultivating plants together with plant growth-promoting bacteria allows the plants to germinate to a much greater extent and then to accumulate a larger amount of biomass in the presence of an environmental contaminant (Fig. 15.30). Typically, plant growth-promoting bacteria alleviate a portion of the stress imposed upon a plant by the presence of organic contaminants, and healthier plants are more efficient at breaking down organic contaminants.

In one study, plant growth-promoting bacteria that facilitate the phytoremediation of polycyclic aromatic hydrocarbons were developed and tested. Polycyclic aromatic hydrocarbons in the environment are of concern because of their toxic, mutagenic, and carcinogenic properties. The strain *Pseudomonas asplenii* AC, isolated from polycyclic aromatic hydrocarbon-contaminated soil, has plant growth-promoting activity, most likely due to its synthesis of indoleacetic acid. This strain was engineered to be more efficient at reducing stress in plants by transforming it with a bacterial gene for the enzyme ACC deaminase (and its regulatory region). The engineered strain was designated *P. asplenii* AC-1. The ability of the wild-type and transformed strains, as well as the transformed strain encapsulated in an alginate matrix (alginate is a biodegradable carbohydrate-based polymer), to promote the growth of canola plants grown in the greenhouse in soil containing polycyclic aromatic hydrocarbons was tested. In the presence of high levels of polycyclic aromatic hydrocarbons, the growth of canola plants was dramatically reduced. When strain AC was added to canola seeds, plant growth improved somewhat. Moreover, addition of strain AC-1, either in suspension or alginate encapsulated, dramatically improved plant growth (Fig. 15.31). These results suggest that plant growth in the

FIGURE 15.30 Growth of Kentucky bluegrass with (+) or without (–) plant growth-promoting bacteria (PGPB) in the presence of increasing amounts of polycyclic aromatic hydrocarbons. In every instance, the plants attain a significantly greater amount of biomass when plant growth-promoting bacteria are added to the seeds before they are planted.



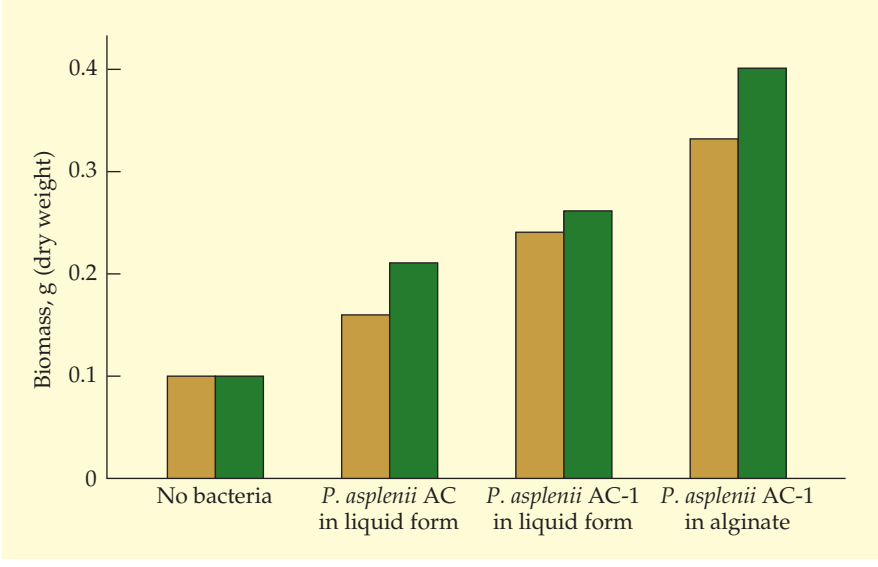


FIGURE 15.31 Canola root (brown) and shoot (green) biomass after 25 days of growth in soil containing 6 g of polycyclic aromatic hydrocarbons per kg. Adapted from Reed and Glick, *Can. J. Microbiol.* **51:**1061–1069, 2005.

presence of polycyclic aromatic hydrocarbons was facilitated by both bacterial indoleacetic acid and ACC deaminase. In addition, several factors may favor the alginate-encapsulated inoculant. As the alginate matrix dissolves, the encapsulated bacteria are released steadily over time, potentially allowing greater bacterial colonization of the plant roots, especially in the presence of polycyclic aromatic hydrocarbons, which can limit bacterial growth and persistence. Alginate encapsulation has also been reported to increase plasmid stability, which is important because the ACC deaminase gene was introduced into strain AC-1 on a broad-host-range plasmid. While these results are preliminary, they indicate that this approach may be useful in the cleanup of contaminated field sites.

Engineering Degradative Plasmids

In addition to engineering bacterial strains to facilitate plant growth in the presence of stressful contaminants, it is also possible to develop plant growth-promoting bacteria that can degrade some contaminants. In one study, scientists engineered a plant growth-promoting strain of *P. fluorescens* to be able to degrade 2,4-dinitrotoluene (Fig. 15.32). The compound 2,4-dinitrotoluene, which is an intermediate in the synthesis of both polyurethane and various explosives, is a problem pollutant. Its presence in the environment is widespread, and it is both toxic and carcinogenic. Several species of *Burkholderia* carry plasmids that encode enzymes that can break down 2,4-dinitrotoluene. However, some species of *Burkholderia* have been found to be either plant pathogens or opportunistic human pathogens, so that there has been reluctance to deliberately release any *Burkholderia* strains, even seemingly harmless ones, into the environment. Instead, using three minitransposons, all of the genes necessary for the complete degradation of 2,4-dinitrotoluene were introduced into the chromosomal DNA of a plant growth-promoting strain of *P. fluorescens*. In liquid culture,

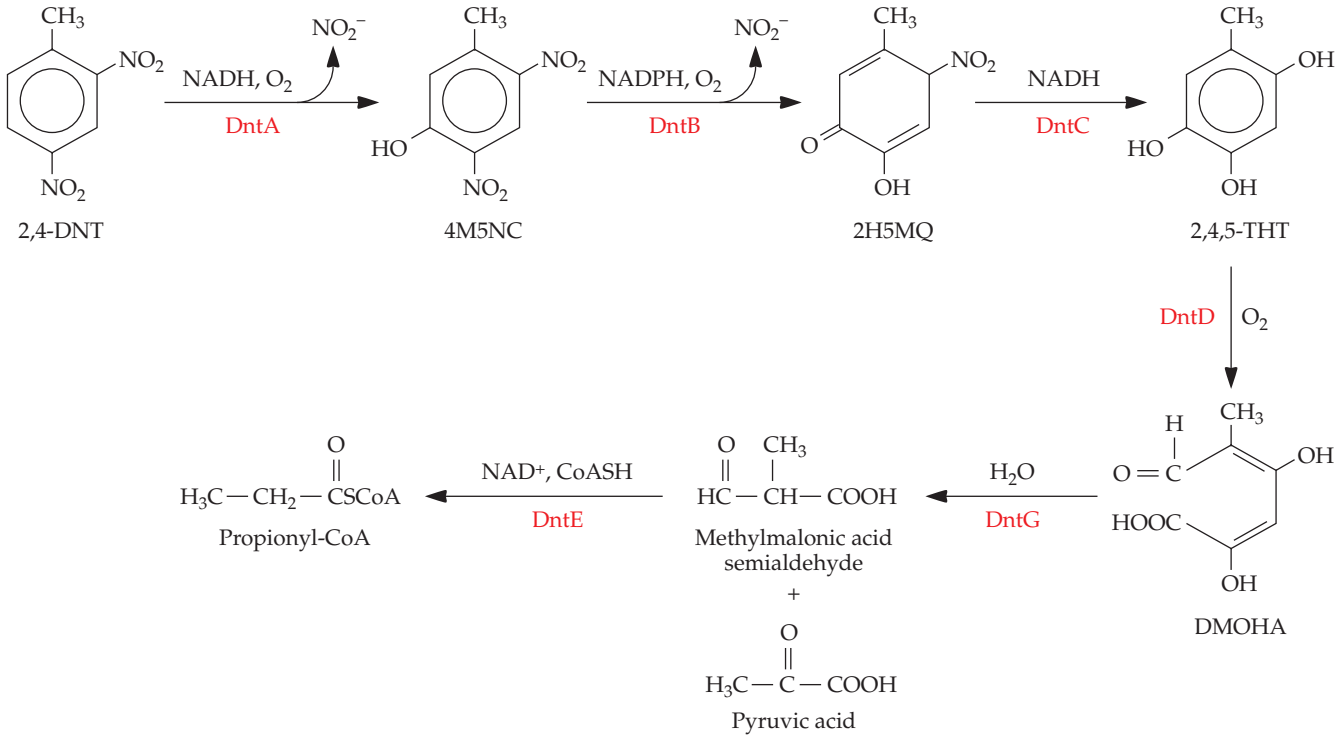


FIGURE 15.32 Pathway for 2,4-dintrotoluene (2,4-DNT) degradation. The enzymes involved in this pathway include DntA, 2,4-dinitrotoluene dioxygenase; DntB, 4-methyl-5-nitrocathecol (4M5NC) monooxygenase; DntC, 2-hydroxy-5-methylquinone (2H5MQ) reductase; DntD, 2,4,5-trihydroxytoluene (2,4,5-THT) oxygenase; DntG, 2,4-dihydroxy-5-methyl-6-oxo-2,4-hexadienoic acid (DMOHA) isomerase/4-hydroxy-2-keto-5-methyl-6-oxo-3-hexenoate hydrolase; and DntE, coenzyme A-dependent methylmalonate semialdehyde (CoASH) dehydrogenase.

the engineered *P. fluorescens* strain completely degraded toxic levels of 2,4-dinitrotoluene at both 28 and 10°C. Following inoculation of plant seeds, a much greater fraction of the seedlings survived in contaminated soil when the engineered rather than the wild-type *P. fluorescens* strain was used (Table 15.16).

Engineering Bacterial Endophytes

Some bacteria that normally bind to and proliferate on the roots of plants (rhizosphere bacteria) contain biodegradative plasmids that encode the enzymes for the complete breakdown of various organic contaminants (see chapter 13). However, the rhizosphere (the area around plant roots) is not always the environment that is most conducive to the degradation of these compounds. This is because when bacteria are attached to the root surface they are affected by soil pH, temperature, water content, and chemical composition, as well as the presence of amoebae, fungi, and other bacteria. Since many plants readily take up a wide range of organic compounds, it might be advantageous if the contaminant-degrading bacteria were localized within the plant roots rather than on the root surface (bacteria that can proliferate within plant tissues are called endophytes). To achieve this, one group of workers transferred, by conjugation, a plasmid containing biodegradative genes encoding enzymes that degrade toluene from a bacterium

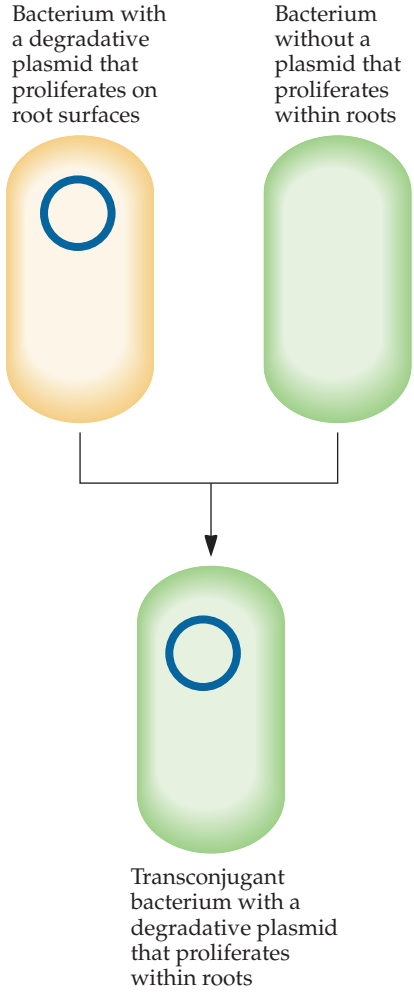
TABLE 15.16 Survival of plant seedlings 14 days after being planted in the presence of DNT

Presence of DNT	Added bacterial strain	Seedling survival (%)
–	None	65
+	None	4
+	Wild type	11
+	Engineered	42

The seeds were inoculated with the indicated bacterium before being planted. –, absent; +, present.

that binds only to root surfaces to an endophytic bacterium that can colonize the interior tissue of the plant root but does not normally degrade toluene (Fig. 15.33). The endophyte that cannot degrade toluene, the surface-colonizing bacterium that can degrade toluene, and the transconjugant endophyte that can degrade toluene were tested for the ability to

FIGURE 15.33 Conjugal transfer of a plasmid carrying toluene degradation genes from a root surface-colonizing bacterium (yellow) to a soil endophytic bacterium (green). In the presence of 3-week-old lupine plants, the resultant transconjugant is able to completely degrade toluene that is added to the soil.



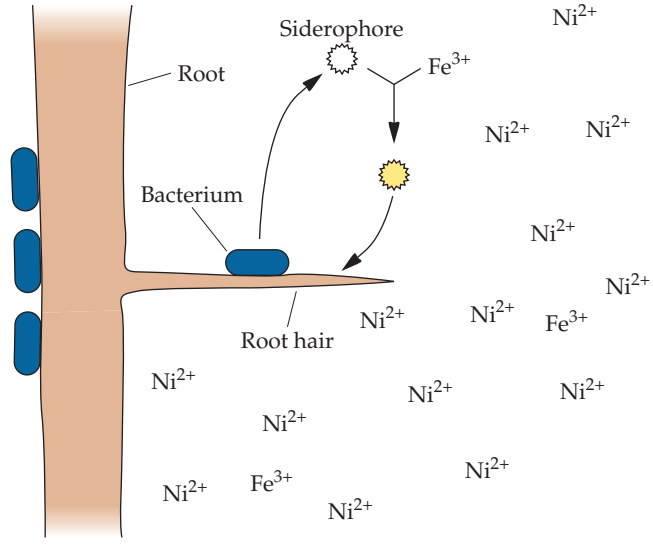


FIGURE 15.34 Schematic representation of a plant growth-promoting bacterium bound to a plant root. In the presence of large amounts of metal (such as nickel [Ni^{2+}]) in the environment, the plant has difficulty acquiring a sufficient amount of iron (Fe^{3+}) from the soil. However, the siderophore that is secreted by the bound bacterium has a very high affinity for iron, forming an iron–siderophore complex that can be taken up by the plant. Once inside the plant, the bacterial siderophore is cleaved, and the iron that is released is utilized in plant metabolic reactions.

degrade toluene in the presence of 3-week-old lupine plants. When either no bacterium or the endophyte that could not degrade toluene was present, the toluene remained intact and was toxic to the plants. The surface-colonizing bacterium that could degrade toluene removed some of the toluene from the soil and allowed the plant to grow to a limited extent. On the other hand, the transconjugant toluene-degrading endophytic strain completely degraded the toluene and protected its host against toluene toxicity. Given the fact that there is still widespread reluctance to deliberately release genetically engineered bacteria into the environment in many jurisdictions, transconjugants carrying naturally occurring biodegradative pathways may be advantageous in that they are not necessarily considered to be genetically modified bacteria.

Metals in the Environment

While plants grown in metal-contaminated soils can withstand some of the effects of high concentrations of metals within their tissues, two features of most plants result in a decrease in plant growth and viability. That is, in the presence of high levels of metals, most plants (1) synthesize stress ethylene and (2) become severely iron depleted. However, plant growth-promoting bacteria can relieve some of the effects of metals on plants. First, ACC deaminase-containing plant growth-promoting bacteria decrease the level of stress ethylene in a plant growing in soil that contains high levels of metal. Second, plants can take up and utilize complexes between bacterial siderophores and iron. In metal-contaminated soils, plants are generally unable to obtain enough iron because iron uptake is inhibited by the metal contaminant(s). Plant siderophores bind to iron with a much lower affinity

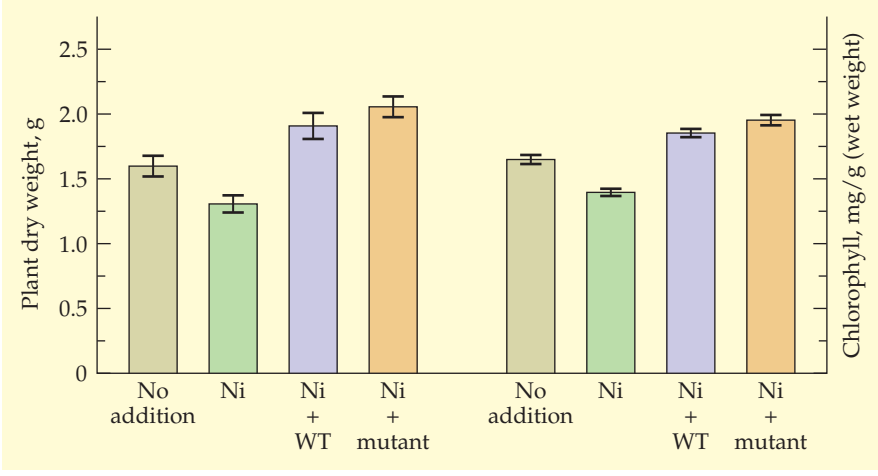


FIGURE 15.35 Effect of adding the wild-type plant growth-promoting bacterium *Kluyvera ascorbata* SUD165 (WT) or a siderophore-overproducing mutant of the bacterium, *K. ascorbata* SUD165/26 (mutant), on plant dry weight and chlorophyll concentration in the presence of nickel (Ni). The error bars indicate standard errors.

than bacterial siderophores, so plants are often unable to accumulate a sufficient amount of iron unless bacterial siderophores are present (Fig. 15.34).

Some metal-resistant bacterial strains promote plant growth in the presence of inhibitory levels of nickel, lead, or zinc and are therefore an effective adjunct to plants in phytoremediation studies. In one instance, a mutation that caused the overproduction of a bacterial siderophore was selected. When the wild-type bacterium and the siderophore-overproducing mutant were tested in the laboratory, the siderophore-overproducing mutant stimulated plant growth significantly more than the wild-type bacterium (Fig. 15.35). When the siderophore-overproducing mutant was tested in the field with the plant Indian mustard *(Brassica juncea)* in soil that had been contaminated with nickel over a period of many years, both the number of seeds that germinated in the nickel-contaminated soil and the size that the plants were able to attain were increased by 50 to 100%. Overall, there was a two- to fourfold increase in the amount of nickel removed from the soil by the addition of the mutant compared with the wild-type bacterium. On the other hand, the presence of the mutant bacterium had no measurable influence on the amount of nickel accumulated per milligram (dry weight) in either plant roots or shoots. Therefore, the bacterial plant growth-promoting effect in the presence of nickel is attributable to the increase in the amount of plant material and the number of plants.

Phytoremediation is still at an early stage of development and currently accounts for only a very small fraction of the total amount spent each year for the remediation of hazardous sites. However, the world remediation market, estimated in 2001 to be $25 billion to $30 billion, is expected to grow to nearly $100 billion by around 2010, and it is estimated that phytoremediation could account for up to 10% of this market.

At present, the largest number of sites being remediated contain organic contaminants, because organics are easier and less expensive than metals to remediate. However, the removal of metal contaminants from the environment is expected to receive more attention in the future.

SUMMARY

Many soil microorganisms have the ability to stimulate the growth of plants. With an eye to diminishing dependency on chemical fertilizers, the molecular mechanisms by which bacteria promote plant growth have been examined. Plant growth promotion may be achieved directly by the ability of the bacteria to fix nitrogen, sequester iron, facilitate phosphorus uptake, produce phytohormones that trigger responses in a growing plant, or enzymatically reduce potentially inhibitory levels of the plant hormone ethylene. Some soil bacteria can also stimulate plant growth indirectly by inhibiting the growth of phytopathogenic microorganisms.

Recently, considerable effort has been devoted to understanding and productively utilizing the ability of plant growth-promoting bacteria to facilitate plant growth by lowering plant ethylene levels. This occurs as a consequence of the action of the enzyme ACC deaminase, which breaks down ACC, the immediate precursor of ethylene in all higher plants.

Plants often respond to a variety of different environmental stresses by synthesizing ethylene, which can trigger a stress/senescence response in the plant. The increased level of ethylene synthesized in response to trauma inflicted by chemicals, temperature extremes, water stress, ultraviolet light, insect predation, disease, and mechanical wounding can be both the cause of some symptoms of stress and the inducer of responses that enhance the survival of the plant under adverse conditions. This seemingly paradoxical situation is explained by the presence of two bursts of ethylene synthesis following the stress. The first, small peak activates the synthesis of plant defense proteins, while the much larger peak of ethylene, which is synthesized later, can exacerbate the impact of the stressor.

The enzyme ACC deaminase, when present in plant growth-promoting bacteria, can act to modulate the level of ethylene in a plant and thereby decrease the deleterious effects of a variety of stressors. Bacteria that contain ACC deaminase activity can significantly decrease the inhibition of plant growth that is observed in the presence of high salt levels, phytopathogens, flooding, or drought. These bacteria can also be used as an adjunct in phytoremediation (environmental cleanup using plants) strategies that are designed to remove metals or organic contaminants from the environment.

Of the plant growth-promoting bacteria that have been studied in detail and are currently used in agricultural practice, much of the research has focused on rhizobial bacteria. These organisms form a complex, obligatory symbiotic relationship with specific plants.

The molecular basis of nitrogen fixation has been examined extensively. Nitrogenase, the nitrogen-fixing enzyme, has been characterized in detail. Molecular genetic studies have established that bacterial nitrogen fixation is a complex process that requires seven operons, with a total of 20 different proteins, that are coordinately regulated. This complexity has so far frustrated attempts to create plants that can fix nitrogen and has prevented transfer of the ability to fix nitrogen to other bacteria.

The amount of nitrogen that can be fixed by a rhizobial strain may be increased by genetic engineering of genes that indirectly affect nitrogen fixation. For example, nitrogen fixation may be increased either by inhibiting the synthesis of rhizobial glycogen, by modulating the level of oxygen within the bacterial cell, or by preventing rhizobial synthesis of the polymer poly-β-hydroxybutyrate, which normally acts as a carbon storage compound.

As part of the action of nitrogenase, hydrogen gas (H_2) is generated at the expense of ATP. Some *Rhizobium* strains possess the enzyme hydrogenase, which is able to recycle H_2 in vivo to H^+, an activity that increases the efficiency of nitrogen fixation. When strains are defective in hydrogenase activity, the ability to fix nitrogen and promote plant growth is diminished. With this in mind, hydrogenase genes have been cloned into strains of rhizobial bacteria that form symbiotic relationships with crop plants. Genetic engineering of hydrogenase genes can produce rhizobial strains with an enhanced ability to fix nitrogen.

Part of the interaction between symbiotic rhizobial strains and plants is the formation of nodules on the roots of plants that are the sites of bacterial nitrogen fixation. It has been reasoned that enhancing nodulation by genetic engineering will enable inoculated rhizobial strains to be more effective competitors for sites on the roots of target plants than indigenous strains. However, studies to date have shown that the genetic basis of nodulation is complex, involving a number of different genes, so that at present there is no simple way to manipulate this process genetically.

The indirect promotion of plant growth occurs when plant growth-promoting bacteria decrease or prevent the damage that is caused by either fungal or bacterial phytopathogens. The bacteria that act in this way are referred to as biocontrol bacterial strains. Some of the substances produced by biocontrol bacteria, such as siderophores, antibiotics, other small molecules, and various enzymes that can lyse fungal cell walls, help to limit the damage to plants by phytopathogens. The activity of biocontrol bacteria may be augmented by engineering these strains to be better root colonizers, more efficient at lowering plant ethylene levels, and more active producers of antibiotics or enzymes.

REFERENCES

Adams, M. W. W., L. E. Mortenson, and J. S. Chen. 1981. Hydrogenase. *Biochim. Biophys. Acta* **594:**105–176.

Albrecht, S. L., R. J. Maier, F. J. Hanus, S. A. Russell, D. W. Emerich, and H. J. Evans. 1979. Hydrogenase in *Rhizobium japonicum* increases nitrogen fixation by nodulated soybeans. *Science* **203:**1255–1257.

Arp, D. J. 1990. H_2 cycling in N_2 fixation: past, present, and future outlook, p. 67–76. *In* P. M. Gresshoff, L. E. Roth, G. Stacey, and W. E. Newton (ed.), *Nitrogen Fixation: Achievements and Objectives*. Chapman & Hall, New York, NY.

Barac, T., S. Taghavi, B. Borremans, A. Provoost, L. Oeyen, J. V. Colpaert, J. Vangronsveld, and D. van der Lelie. 2004. Engineered endophytic bacteria improve phytoremediation of water-soluble, volatile, organic pollutants. *Nat. Biotechnol.* **22:**583–588.

Bar-Ness, E., Y. Chen, Y. Hadar, H. Marschner, and V. Römheld. 1991. Siderophores of *Pseudomonas putida* as an iron source for dicot and monocot plants, p. 271–281. *In* Y. Chen and Y. Hadar (ed.), *Iron Nutrition and Interactions in Plants*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Brewin, N. J., and A. W. B. Johnston. January 1986. Synthetic plasmid and bacteria containing it. U.S. patent 4,567,146.

Brito, B., J. M. Palacios, J. Imperial, and T. Ruiz-Argüeso. 2002. Engineering the *Rhizobium leguminosarum* bv. viciae hydrogenase system for expression in free-living microaerobic cells and increased hydrogenase activity. *Appl. Environ. Microbiol.* **68:**2461-2467.

Burd, G. I., D. G. Dixon, and B. R. Glick. 1998. A plant growth promoting bacterium that decreases nickel toxicity in plant seedlings. *Appl. Environ. Microbiol*. **64:**3663–3668.

Burd, G. I., D. G. Dixon, and B. R. Glick. 2000. Plant growth-promoting bacteria that decrease heavy metal toxicity in plants. *Can. J. Microbiol*. **46:**237–245.

Cantrell, M. A., R. A. Haugland, and H. J. Evans. 1983. Construction of a *Rhizobium japonicum* gene bank and use in isolation of a hydrogen uptake gene. *Proc. Natl. Acad. Sci. USA* **80:**181–185.

Cheng, Z., E. Park, and B. R. Glick. 2007. 1-Aminocyclopropane-1-carboxylate (ACC) deaminase from *Pseudomonas putida* UW4 facilitates the growth of canola in the presence of salt. *Can. J. Microbiol.* **53:**912–918.

Chet, I., and J. Inbar. 1994. Biological control of fungal pathogens. *Appl. Biochem. Biotechnol*. **48:**37–43.

Crosa, J. H. 1989. Genetics and molecular biology of siderophore-mediated iron transport in bacteria. *Microbiol. Rev.* **53:**517–530.

Dekkers, L. C., I. H. M. Mulders, C. C. Phoelich, T. F. C. Chin-A-Woeng, A. H. M. Wijfjes, and B. J. J. Lugtenberg. 2000. The *sss* colonization gene of the tomato-*Fusarium oxysporum* f. sp. *radicis-lycopersici* biocontrol strain *Pseudomonas fluorescens* WCS365 can improve root colonization of other wild-type *Pseudomonas* spp. bacteria. *Mol. Plant-Microbe Interact*. **13:**1177–1183.

Evans, H. J., A. R. Harker, H. Papen, S. A. Russell, F. J. Hanus, and M. Zuber. 1987. Physiology, biochemistry, and genetics of the uptake hydrogenase in rhizobia. *Annu. Rev. Microbiol.* **41:**335–361.

Fischer, H.-M. 1994. Genetic regulation of nitrogen fixation in rhizobia. *Microbiol. Rev.* **58:**352–386.

Glick, B. R. 1995. The enhancement of plant growth by free-living bacteria. *Can. J. Microbiol*. **41:**109–117.

Glick, B. R. 2004. Teamwork in phytoremediation. *Nat. Biotechnol.* **22:**526–527.

Glick, B. R., C. L. Patten, G. Holguin, and D. M. Penrose. 1999. *Biochemical and Genetic Mechanisms Used by Plant Growth Promoting Bacteria*. Imperial College Press, London, United Kingdom.

Glick, B. R., C. B. Jacobson, M. M. K. Schwarze, and J. J. Pasternak. 1994. 1-Aminocyclopropane-1-carboxylic acid deaminase mutants of the plant growth-promoting rhizobacterium *Pseudomonas putida* GR12-2 do not stimulate canola root elongation. *Can. J. Microbiol*. **40:**911–915.

Glick, B. R., D. M. Penrose, and J. Li. 1998. A model for the lowering of plant ethylene concentrations by plant growth-promoting bacteria. *J. Theor. Biol*. **190:**63–68.

Glick, B. R., J. Zeisler, A. M. Banaszuk, J. D. Friesen, and W. G. Martin. 1981. The identification and partial characterization of a plasmid containing the gene for the membrane-associated hydrogenase from*E. coli*. *Gene* **15:**201–206.

Gresshoff, P. M., L. E. Roth, G. Stacey, and W. E. Newton (ed.). 1990. *Nitrogen Fixation: Achievements and Objectives*. Chapman & Hall, New York, NY.

Grichko, V. P., and B. R. Glick. 2001. Ethylene and flooding stress in plants. *Plant Physiol. Biochem*. **39:**1–9.

Grichko, V. P., and B. R. Glick. 2001. Amelioration of flooding stress by ACC deaminase-containing plant growth-promoting bacteria. *Plant Physiol. Biochem*. **39:**11–17.

Hennecke, H. 1990. Nitrogen fixation genes involved in the *Bradyrhizobium japonicum*-soybean symbiosis. *FEBS Lett*. **268:**422–426.

Higashi, S. 1993. (*Brady*)*Rhizobium*-plant communications involved in infection and nodulation. *J. Plant Res*. **106:**201–211.

Huang, X.-D., Y. El-Alawi, D. M. Penrose, B. R. Glick, and B. M. Greenberg. 2004. Responses of plants to creosote during phytoremediation and their significance for remediation processes. *Environ. Pollut.* **130:**453–463.

Huang, X.-D., Y. El-Alawi, J. Gurska, B. R. Glick, and B. M. Greenberg. 2005. A multi-process phytoremedia-

tion system for decontamination of persistent total petroleum hydrocarbons (TPHs) from soils. *Microchem. J.* **81:**139–147.

Jones, D. A., M. H. Ryder, B. G. Clare, S. K. Farrand, and A. Kerr. 1988. Construction of a Tra⁻ deletion mutant of pAgK84 to safeguard the biological control of crown gall. *Mol. Gen. Genet*. **212:**207–214.

Kloepper, J. W., R. Lifshitz, and M. N. Schroth. 1988. *Pseudomonas* inoculants to benefit plant production. *ISI Atlas Sci. Anim. Plant Sci*. **1:**60–64.

Layva, A., J. M. Palacios, T. Mozo, and T. Ruiz-Argüeso. 1987. Cloning and characterization of hydrogen uptake genes from *Rhizobium leguminosarum*. *J. Bacteriol*. **169:**4929–4934.

Lerouge, P., P. Roche, C. Faucher, F. Maillet, G. Truchet, J. C. Promé, and J. Dénarié. 1990. Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated oligosaccharide signal. *Nature* **344:**781–784.

Li, X., S. H. Yang, X. C. Yu, Z. X. Jin, W. D. Li, L. Li, and M. G. Li. 2005. Construction of transgenic *Bacillus mucilaginosus* strain with improved phytase secretion. *J. Appl. Microbiol.* **99:**878–884.

Li, X., Z. Wu, W. D. Li, R. Yan, L. Li, J. Li, Y. Li, and M. G. Li. 2007. Growth promoting effect of a transgenic *Bacillus mucilaginosus* on tobacco planting. *Appl. Microbiol. Biotechnol.* **74:**1120-1125.

Long, S. R., W. J. Buikema, and F. M. Ausubel. 1982. Cloning of *Rhizobium meliloti* nodulation genes by direct complementation of Nod⁻ mutants. *Nature* **298:**485–488.

Lynch, J. M. 1990. Beneficial interactions between microorganisms and roots. *Biotechnol. Adv*. **8:**335–346.

Ma, W., S. B. Sebestianova, J. Sebestian, G. I. Burd, F. C. Guinel, and B. R. Glick. 2003. Prevalence of 1-aminocyclopropane-1-carboxylate deaminase in *Rhizobium* spp. *Antonie van Leeuwenhoek* **83:**285–291.

Ma, W., T. C. Charles, and B. R. Glick. 2004. Expression of an exogenous 1-aminocyclopropane-1-carboxylate deaminase gene in *Sinorhizobium meliloti* increases its ability to nodulate alfalfa. *Appl. Environ. Microbiol*. **70:**5891–5897.

Maier, R. J., and E. W. Triplett. 1996. Toward more productive, efficient, and competitive nitrogen-fixing symbiotic bacteria. *Crit. Rev. Plant Sci*. **15:**191–234.

Marroquí, S., A. Zorreguieta, C. Santamaría, F. Temprano, M. Soberón, M. Megías, and J. A. Downie. 2001. Enhanced symbiotic performance by *Rhizobium tropici* glycogen synthase mutants. *J. Bacteriol*. **183:**854–864.

Marugg, J. D., M. van Spanje, W. P. M. Hoekstra, B. Schippers, and P. J. Weisbeek. 1985. Isolation and analysis of genes involved in siderophore biosynthesis in plant-growth-stimulating *Pseudomonas putida* WCS358. *J. Bacteriol*. **164:**563–570.

Marugg, J. D., H. B. Nielander, A. J. G. Horrevoets, I. van Megen, I. van Genderen, and P. J. Weisbeek. 1988. Genetic organization and transcriptional analysis of a major gene cluster involved in siderophore biosynthesis in *Pseudomonas putida* WCS358. *J. Bacteriol*. **170:**1812–1819.

Mayak, S., T. Tirosh, and B. R. Glick. 2004. Plant growth-promoting bacteria that confer resistance in tomato to salt stress. *Plant Physiol. Biochem*. **42:**565–572.

Monti, M. R., A. M. Smania, G. Fabro, M. E. Alvarez, and C. E. Argaraña. 2005. Engineering *Pseudomonas fluorescens* for biodegradation of 2,4-dinitrotoluene. *Appl. Environ. Microbiol.* **71:**8864–8872.

Morris, R. O. 1986. Genes specifying auxin and cytokinin biosynthesis in phytopathogens. *Annu. Rev. Plant Physiol.* **37:**509–538.

Muryoi, N., H. Kawahara, H. Obata, M. Griffith, and B. R. Glick. 2004. Cloning and expression of *afpA*, a gene encoding an antifreeze protein from the arctic plant growth-promoting rhizobacterium *Pseudomonas putida* GR12-2. *J. Bacteriol*. **186:**5661-5671.

Mylona, P., K. Pawlowski, and T. Bisseling. 1995. Symbiotic nitrogen fixation. *Plant Cell* **7:**869–885.

Nap, J.-P., and T. Bisseling. 1990. Developmental biology of a plant-prokaryote symbiosis: the legume root nodule. *Science* **250:**948–954.

Neilands, J. B., and S. A. Leong. 1986. Siderophores in relation to plant growth and disease. *Annu. Rev. Plant Physiol*. **37:**187–208.

O'Sullivan, D. J., and F. O'Gara. 1992. Traits of fluorescent *Pseudomonas* spp. involved in suppression of plant root pathogens. *Microbiol. Rev*. **56:**662–676.

Paau, A. S. 1991. Improvement of *Rhizobium* inoculants by mutation, genetic engineering, and formulation. *Biotechnol. Adv*. **9:**173–184.

Patten, C. L., and B. R. Glick. 1996. Bacterial biosynthesis of indole-3-acetic acid. *Can. J. Microbiol*. **42:**207–220.

Patten, C. L., and B. R. Glick. 2002. The role of bacterial indoleacetic acid in the development of the host plant root system. *Appl. Environ. Microbiol*. **68:**3795–3801.

Peralta, H., Y. Mora, E. Salazar, S. Encarnación, R. Palacios, and J. Mora. 2004. Engineering the *nifH* promoter region and abolishing poly-β-hydroxybutyrate accumulation in *Rhizobium etli* enhance nitrogen fixation in symbiosis with *Phaseolus vulgaris*. *Appl. Environ. Microbiol.* **70:**3272-3281.

Peters, J. W., K. Fisher, and D. R. Dean. 1995. Nitrogenase structure and function: a biochemical-genetic perspective. *Annu. Rev. Microbiol*. **49:**335–366.

Ramírez, M., B. Valderrama, R. Arrendondo-Peter, M. Soberón, J. Mora, and G. Hernández. 1999. *Rhizobium etli* genetically engineered for the heterologous expression of *Vitreoscilla* sp. hemoglobin: effects on free-living and symbiosis. *Mol. Plant-Microbe Interact*. **12:**1008–1015.

Reed, M. L. E., and B. R. Glick. 2004. Applications of free living plant growth-promoting rhizobacteria. *Antonie van Leeuwenhoek* **86:**1–25.

Reed, M. L. E., and B. R. Glick. 2005. Growth of canola (*Brassica napus*) in the presence of plant growth-promoting bacteria and either copper or polycyclic aromatic hydrocarbons. *Can. J. Microbiol.* **51:**1061–1069.

Rossen, L., E. O. Davis, and A. W. B. Johnston. 1987. Plant-induced expression of *Rhizobium* genes involved in host specificity and early stages of nodulation. *Trends Biol. Sci*. **12:**430–433.

Schnider, U., C. Keel, C. Blumer, J. Troxler, G. Défago, and D. Haas. 1995. Amplification of the housekeeping sigma factor in *Pseudomonas fluorescens* CHA0 enhances antibiotic production and improves biocontrol abilities. *J. Bacteriol*. **177:**5387–5392.

Spaink, H. 2000. Root nodulation and infection factors produced by rhizobial bacteria. *Annu. Rev. Microbiol*. **54:**257–288.

Spaink, H. P., C. A. Wijffelman, E. Pees, R. J. H. Okker, and B. J. J. Lugtenberg. 1987. *Rhizobium* nodulation gene *nodD* as a determinant of host specificity. *Nature* **328:**337–340.

Sprent, J. I. 1986. Benefits of *Rhizobium* to agriculture. *Trends Biotechnol*. **4:**124–129.

Stacey, G. 1995. *Bradyrhizobium japonicum* nodulation genetics. *FEMS Microbiol. Lett*. **127:**1–9.

Sun, X., M. Griffith, J. J. Pasternak, and B. R. Glick. 1995. Low temperature growth, freezing survival and production of antifreeze protein by the plant growth-promoting rhizobacterium *Pseudomonas putida* GR12-2. *Can. J. Microbiol*. **41:**776–784.

Timms-Wilson, T. M., R. J. Ellis, A. Renwick, D. J. Rhodes, D. V. Mavrodi, D. M. Weller, L. S. Thomashow, and M. J. Bailey. 2000. Chromosomal insertion of phenazine-1-carboxylic acid biosynthetic pathway enhances efficacy of damping-off disease control by *Pseudomonas fluorescens*. *Mol. Plant-Microbe Interact*. **13:**1293–1300.

Ureta, A.-C., J. Imperial, T. Ruiz-Argüeso, and J. M. Palacios. 2005. *Rhizobium leguminosarum* biovar viciae symbiotic hydrogenase activity and processing are limited by the level of nickel in agricultural soils. *Appl. Environ. Microbiol*. **71:**7603–7606.

van Rhijn, P., and J. Vanderleyden. 1995. The *Rhizobium*-plant symbiosis. *Microbiol. Rev*. **59:**124–142.

Wang, C., E. Knill, B. R. Glick, and G. Défago. 2000. Effect of transferring 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase genes into *Pseudomonas fluorescens* strain CHA0 and its *gacA* derivative CHA96 on their growth-promoting and disease-suppressive capacities. *Can. J. Microbiol*. **46:**898–907.

REVIEW QUESTIONS

1. Starting with a strain of *Bradyrhizobium japonicum* that can fix nitrogen and form a symbiotic relationship with soybean roots, and assuming that you do not have a DNA hybridization probe for *nod* genes, outline a scheme that you would use for isolating the cluster of nodulation genes from this organism.

2. How does glycogen synthesis affect the ability of a strain of *Rhizobium* to fix nitrogen?

3. How would you engineer a rhizobial strain to have a lower internal level of free oxygen? How might this affect the bacterium in the free-living state? As a bacteroid?

4. What is hydrogenase? How could it be used to improve the yield of alfalfa?

5. Suggest a strategy that you might employ to isolate all of the genes involved in nitrogen fixation from *Azotobacter vinelandii,* assuming that you do not have *nif* genes from other microorganisms to use as hybridization probes.

6. What might be the consequences of mutagenizing either *nifA* or *nifL* with respect to the amount of nitrogen that an organism can fix?

7. Discuss whether it is possible to genetically engineer plants to fix nitrogen.

8. What are siderophores? How could genetic manipulation of siderophore genes enable bacteria to enhance plant growth?

9. Suggest a scheme for isolating siderophore biosynthesis genes.

10. What are the advantages of microbial fertilizers over chemical fertilizers?

11. How do ACC deaminase-containing plant growth-promoting bacteria facilitate plant growth?

12. What is phytoremediation? How do plant growth-promoting bacteria affect phytoremediation?

13. How can *A. radiobacter* be engineered to be a more effective biocontrol agent?

14. Which enzymes secreted by plant growth-promoting bacteria contribute to their ability to act as biocontrol agents? How do these enzymes contribute to biocontrol?

15. How does poly-β-hydroxybutyrate affect the ability of a strain of *Rhizobium* to fix nitrogen?

16. How might endophytic bacteria be useful as part of a phytoremediation strategy?

17. Assuming that you do not have a DNA hybridization probe available, how would you isolate a bacterial antifreeze protein gene?

18. What strategies can be employed to increase the effectiveness of biocontrol bacterial strains?

19. What mechanisms do free-living plant growth-promoting bacteria use to facilitate plant growth?

16

Microbial Insecticides

Of all classes of organisms, insects have the largest number of described species (more than 750,000). Insects negatively affect humans in a variety of ways: they cause massive crop damage, and they act as vectors of both human and animal diseases. During the 1940s, a number of chemical insecticides were developed as a means of controlling the proliferation of noxious insect populations. One of these was the chlorinated hydrocarbon DDT (dichlorodiphenyltrichloroethane), which had originally been synthesized in the 1870s but was not recognized as an insecticide until the late 1930s. DDT proved to be exceptionally effective in killing and controlling many species of pests. Chlorinated hydrocarbons such as DDT function by attacking the nervous system and muscle tissue of insects. Later, other chlorinated hydrocarbons, such as dieldrin, aldrin, chlordane, lindane, and toxophene, were synthesized and applied on a massive scale against crop pests and insects that carry infectious agents.

Organophosphates, another class of chemical insecticides that includes malathion, parathion, and diazinon, were initially developed as chemical warfare agents. Now they are used to control insect populations by inhibiting the enzyme acetylcholinesterase, which hydrolyzes the nerve transmitter acetylcholine, thereby disrupting the functioning of motor neurons and neurons in the brain of the insect.

By the early 1960s, over 100 million acres of U.S. agricultural land was being treated annually with chemical insecticides. At about that time, researchers realized that chlorinated hydrocarbon insecticides, to a large extent, and organophosphate insecticides, to a lesser extent, had dramatic and immediate side effects and long-term and indirect effects on animals, ecosystems, and humans. Chlorinated hydrocarbons, exemplified by DDT, were found to persist in the environment for more than 20 years and to accumulate in increasing concentrations through food chains. This bioaccumulation in fatty tissues had a significant biological impact on many organisms. For example, in North America, many species of birds, including peregrine falcons, sparrow hawks, bald eagles, brown pelicans, and double-crested cormorants, underwent severe population declines.

During the 1950s, as the targeted insect pest populations became increasingly resistant to treatment with many chemical insecticides, higher concentrations of the insecticides were applied to control the pests. In addition, chemical insecticides were found to lack specificity; consequently, beneficial insects were killed along with those that were considered to be pests. In fact, in some instances, the natural enemies of the insect pest species were killed off more efficiently than the target organisms, with the bizarre result that pesticide treatment led to greater numbers of the insects.

Given all the drawbacks associated with the use of chemical insecticides, alternative means of controlling harmful insects have been sought. Using insecticides that are produced naturally by either microorganisms or plants was an obvious choice. On the positive side, these compounds are usually highly specific for a target insect species, biodegradable, and slow to select for resistance. But, on the negative side, their low potency and high cost of production limit their use for a variety of applications. Recombinant DNA technology provides an opportunity to overcome many of these negative attributes. In particular, the insecticidal activities of the bacterium *Bacillus thuringiensis* and insect baculovirus systems have been developed into safe, specific, and effective insecticides.

The worldwide market for pesticides is enormous: currently more than $30 billion per year and growing rapidly. Although biopesticides, mostly *B. thuringiensis*, make up only about 1% of this total, much of the expected growth in this field is likely to involve biopesticides.

Insecticidal Toxin of *B. thuringiensis*

Mode of Action and Use

A microbial insecticide can be a microbially produced toxic substance that kills an insect species or an organism that has the ability to fatally infect a specific target insect. The most studied, most effective, and most often utilized microbial insecticides are the toxins synthesized by *B. thuringiensis*. This bacterium comprises a large number of strains and subspecies, each of which produces a different toxin that can kill specific insects—there are more than 150 different subspecies of *B. thuringiensis* (Table 16.1). For

TABLE 16.1 Some properties of the insecticidal toxins from various strains of *B. thuringiensis*

***B. thuringiensis* strain or subspecies**	**Protoxin size (kDa)**	**Target insects**	**Serotype**
berliner	130–140	Lepidoptera	1
kurstaki KTO, HD-1	130–140	Lepidoptera	3
entomocidus 6.01	130–140	Lepidoptera	6
aizawai 7.29	130–140	Lepidoptera	7
aizawai IC 1	135	Lepidoptera, Diptera	7
kurstaki HD-1	71	Lepidoptera, Diptera	3
tenebrionis (san diego)	66–73	Coleoptera	8
morrisoni PG14	125–145	Diptera	8
israelensis	68	Diptera	14

Adapted from Lereclus et al., p. 37–69, in Entwistle et al. (ed.), *Bacillus thuringiensis, an Enviromental Biopesticide: Theory and Practice* (John Wiley & Sons, Chichester, United Kingdom, 1993).

example, *B. thuringiensis* subsp. *kurstaki* is toxic to lepidopteran larvae, including those of moths, butterflies, and skippers; cabbage worms; and spruce budworms. *B. thuringiensis* subsp. *israelensis* kills diptera, such as mosquitoes and blackflies. *B. thuringiensis* subsp. *tenebrionis* (also known as *B. thuringiensis* subsp. *san diego*) is effective against coleoptera (beetles), such as the potato beetle and the boll weevil. In addition, some subspecies of *B. thuringiensis* produce insecticidal toxins that are directed against hymenoptera (sawflies, wasps, bees, and ants), orthoptera (grasshoppers, crickets, and locusts), and mallophaga (lice).

The insecticidal activity (toxin) of *B. thuringiensis* subsp. *kurstaki* (first discovered in 1911) and other strains is contained within a very large structure called a parasporal crystal, which is synthesized during bacterial sporulation. Although no significant role on behalf of the bacterium has been attributed to the parasporal crystal structure, by synthesizing the crystal, the bacterium is "providing for its future" in that a dead insect provides sufficient nutrients to allow germination of the dormant spore. The parasporal crystal contains approximately 20 to 30% of the dry weight of a sporulated culture and usually consists mainly of protein (~95%) and a small amount of carbohydrate (~5%). About 150 different parasporal crystal proteins (Cry proteins) are known. The crystal is an aggregate of protein that can generally be dissociated by mild alkali treatment into subunits. The subunits can be further dissociated in vitro by treatment with β-mercaptoethanol, which reduces disulfide linkages (Fig. 16.1).

The insecticidal toxins from the *B. thuringiensis* strains were previously grouped into four major classes—CryI, CryII, CryIII, and CryIV—based on the insecticidal activity of the toxin. These proteins were further organized into subclasses (A, B, C, etc.) and subgroups (a, b, c, etc.). In the past few years, as increasing numbers of *B. thuringiensis* strains were isolated and their genes were characterized, it became clear that the original classification was unable to accommodate many of the newly discovered *B. thuringiensis* toxin genes. Therefore, a new system of *B. thuringiensis* gene classification was introduced.

In the current classification scheme (established in 1998), *B. thuringiensis* insecticidal (Cry) proteins are assigned designations based on their degree of evolutionary divergence, as estimated by certain mathematical algorithms. This scheme is readily visualized by constructing a phylogenetic tree based on the amino acid sequences of *B. thuringiensis* toxin proteins, i.e., Cry proteins (Fig. 16.2). Basically, the amino acid sequences of the proteins are compared, and if the proteins are identical, then they are 100% homologous. If only 50% of the amino acids are the same, then the proteins have 50% identity. The relationship among a set of protein sequences can be deduced and represented as a branched tree. The nodes (branch points) of the tree represent points of divergence. For the classification of the *B. thuringiensis* Cry proteins, a four-part naming system was devised. Demarcations, set at 95, 78, and 45% homology, show the boundaries that define the different nomenclature ranks. The name that is given to a particular toxin depends on the location of the node where the toxin protein sequence enters the tree relative to these set boundaries. A toxin that joins the tree to the left of the leftmost boundary is assigned a new primary rank (an Arabic numeral), one that joins the tree between the central and left boundaries is assigned a new secondary rank (an uppercase letter), one that joins the tree between the central and the right boundaries is assigned a new tertiary rank (a lowercase letter), and one that joins to the right of the

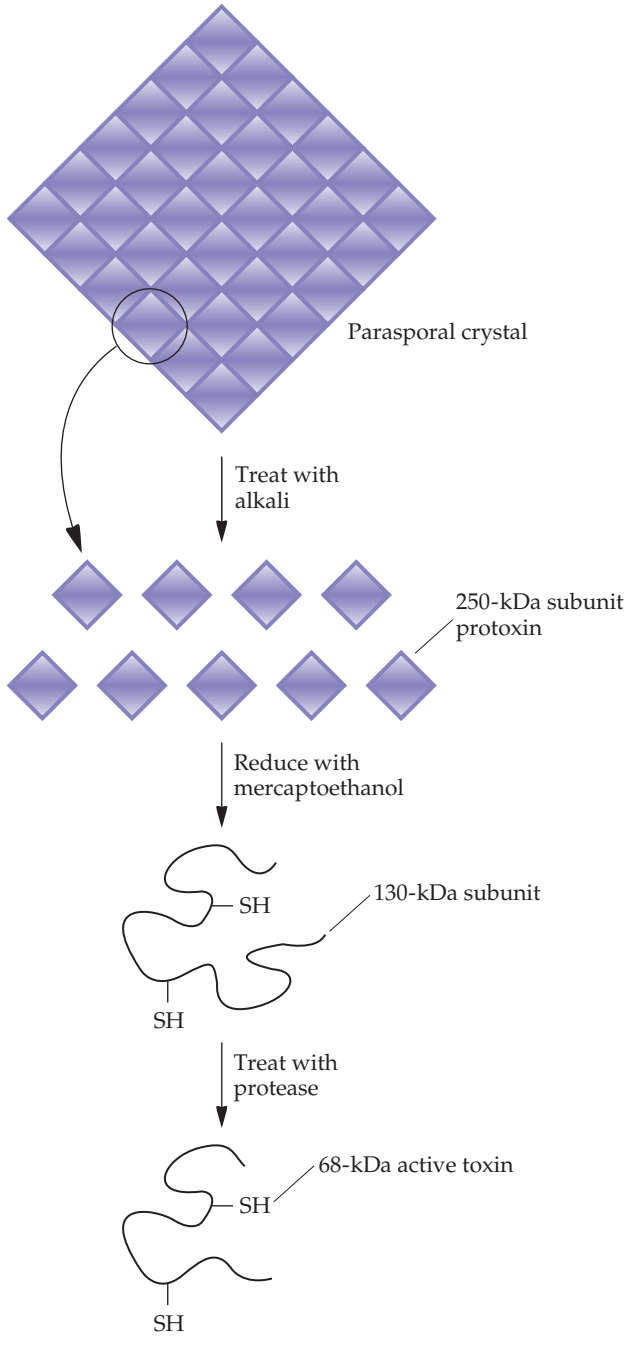


FIGURE 16.1 Schematic representation of a *B. thuringiensis* parasporal crystal composed of Cry1 protoxin protein. Each 250-kDa protein subunit of the parasporal crystal contains two 130-kDa polypeptides. (Molecular masses determined by polyacrylamide gel electrophoresis are approximations and do not always provide exact multiples.) Conversion of the 130-kDa protoxin into an active 68-kDa toxin requires the combination of a slightly alkaline pH (7.5 to 8) and the action of a specific protease(s), both of which are found in the insect gut. The activated toxin binds to protein receptors on the surface of the gut epithelial cell membrane.

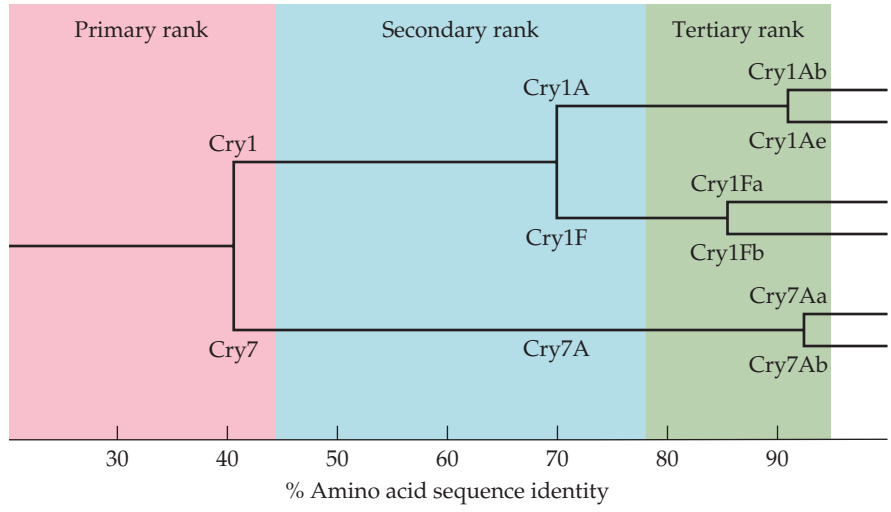


FIGURE 16.2 Schematic representation of a portion of the phylogenetic tree of *B. thuringiensis* insecticidal (Cry) proteins. The different background colors delineate the different levels of nomenclature ranks. Cry1 and Cry7 share less than 45% identity, Cry1A and Cry1F are between 45% and 78% identical, and Cry1Ab and Cry1Ae are between 78% and 95% identical. Adapted from Crickmore et al., *Microbiol. Mol. Biol. Rev*. **62:**807–813, 1998.

rightmost boundary is assigned a new quaternary rank (an Arabic numeral). For example, Cry proteins that are less than 45% homologous are given a number (e.g., Cry1 and Cry7) and are assigned to the primary rank. Cry proteins that are 45 to 78% identical to proteins of the primary rank are further designated with an uppercase letter (e.g., Cry1A and Cry1F). The complete Cry protein tree consists of the positions of all Cry proteins. This classification system is utilized throughout this chapter, even when referring to work that was published prior to the development of this system.

The parasporal crystal does not usually contain the active form of the insecticide. Rather, once the crystal has been solubilized, the protein that is released is generally a protoxin, a precursor of the active toxin. The protoxin of many of the Cry toxins that are directed against lepidoptera has a molecular mass of approximately 130 kilodaltons (kDa) (Fig. 16.1). When a parasporal crystal is ingested by a target insect, the protoxin is activated within its gut by the combination of alkaline pH (7.5 to 8.0) and specific digestive proteases, which convert the protoxin into an active toxin with a molecular mass of approximately 68 kDa (Fig. 16.1). In its active form, the toxic protein inserts itself into the membranes of the gut epithelial cells of the insect and creates an ion channel, which leads to an excessive loss of cellular ATP (Fig. 16.3). About 15 minutes after this ion channel forms, cellular metabolism ceases; the insect stops feeding within a few hours, becomes dehydrated, and eventually dies (in about 2 to 5 days). Because the conversion of the protoxin to the active toxin requires both alkaline pH and the presence of specific proteases, it is extremely unlikely that nontarget species, such as humans and farm animals, will be affected.

The mode of action of *B. thuringiensis* toxins imposes certain constraints on its application. To kill an insect pest, *B. thuringiensis* parasporal crystals must be ingested. Contact of the bacterium or the insecticidal toxin with the surface of the target organism has no effect on it. The requirement that the

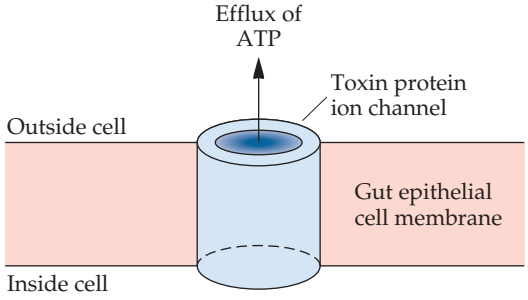


FIGURE 16.3 Insertion of the *B. thuringiensis* toxin into the membrane of an insect gut epithelial cell. The toxin forms an ion channel between the cell cytoplasm and the external environment.

insecticide be ingested, in part, limits the susceptibility of nontarget insects and other animals to the insecticide. *B. thuringiensis* is generally applied by spraying, so it is usually formulated with insect attractants to increase the probability that the target insect will ingest the toxin. However, insects that bore into plants or attack plant roots are less likely to ingest a *B. thuringiensis* toxin that has been sprayed on a host plant, so other strategies have been devised to control such pests. One approach is to create transgenic plants that carry and express a *B. thuringiensis* toxin gene so that they are protected from infestation throughout the growing season (see chapter 18).

It was recently discovered for gypsy moths (and suggested to possibly be the case for other insects, as well) that the *B. thuringiensis* toxin does not kill the larvae by itself as previously thought. Rather, bacteria that are part of the insect's gut microbial community are required for toxicity to the insect. Elimination of the insect's gut bacteria by oral administration of antibiotics abolished *B. thuringiensis* insecticidal activity, and reintroduction of an *Enterobacter* sp. that is normally part of the insect's gut microbial community restored this activity. The data indicate that the *B. thuringiensis* toxin enables the enteric bacteria to reach the insect hemocoel by permeabilizing the gut epithelium. In this way, the insect is killed much more rapidly than might otherwise be expected. The discovery that *B. thuringiensis* insecticidal activity depends on insect enteric bacteria should not have any significant effect on the efficacy or use of *B. thuringiensis*-based insecticides. However, this information may be important in the design and execution of some laboratory experiments intended to better understand the functioning of *B. thuringiensis* insecticidal strains and to facilitate the development of improved biological insecticides.

A limiting feature of the action of the *B. thuringiensis* toxin is that it can kill a susceptible insect only during a specific developmental stage. Therefore, the toxin must be applied when the pest population is at a particular stage in its life cycle (generally the larval stage). The other major impediment to more widespread application of *B. thuringiensis* subsp. *kurstaki* is that it costs from 1.5 to 3 times as much as chemical insecticides. The limitations and the cost notwithstanding, several subspecies of *B. thuringiensis* have been approved for use and have rapidly gained widespread acceptance (Table 16.2).

B. thuringiensis subsp. *kurstaki* was first discovered in 1901, although its commercial potential was largely ignored until 1951. Within recent decades, however, *B. thuringiensis* subsp. *kurstaki* has become the major means of

TABLE 16.2 Some subspecies of *B. thuringiensis* that have been approved for use in the field and some of their targets

B. thuringiensis subspecies	Targets
aizawai	Fruits and nuts, berries, peppers, tomato, root crops, tobacco, beans, corn, cotton, cabbage, eggplant, melons, cucumber, cauliflower, broccoli, ornamentals
kurstaki	Berries, fruits, nuts, melons, cucumber, squash, eggplant, tomato, broccoli, cabbage, kale, mustard, parsley, spinach, turnip, lettuce, stored grain, stored crop seed, ornamentals, cotton, celery, peanut, sugar beet, tobacco, avocado, onion, carrot, forestry products, grape, canola, sorghum, wheat, forage crops, corn, sunflower, root crops, cranberry
israelensis	Mosquito breeding habitat, including rice fields, ponds, pastures, ditches, salt marshes, tidal water, sewage lagoons, lakes; ornamental and nursery plants; mushrooms *(Agaricus bisporus)*
tenebrionis	Eggplant, tomato, potato, ornamentals

controlling the spruce budworm in Canada. In 1979, approximately 1% of the forest area in Canada that was treated with an insecticide to combat the spruce budworm (about 2 million hectares, or 8,000 square miles) was sprayed with *B. thuringiensis* subsp. *kurstaki*. The remainder of the treated forests were sprayed with chemical insecticides. By 1986, the use of *B. thuringiensis* subsp. *kurstaki* had increased dramatically. It was used to treat approximately 74% of the forests sprayed in that year for spruce budworm. In other countries, *B. thuringiensis* subsp. *kurstaki* has been used against tent caterpillars, gypsy moths, cabbage worms, cabbage loopers, and tobacco hornworms.

For the biological control (biocontrol) of insect pests, *B. thuringiensis* subsp. *kurstaki* is typically applied by spraying approximately 1.3×10^8 to 2.6×10^8 spores per square foot (1 square foot is equivalent to 0.093 m^2) of the target area. Administration of the spores is timed to coincide with the peak of the larval population of the target organism, because the parasporal crystals, being sensitive to sunlight, are short-lived in the environment. Under simulated conditions, sunlight degrades over 60% of the tryptophan residues of the parasporal crystal within a 24-hour period, thereby rendering the protein inactive. Depending on the amount of sunlight present, parasporal crystals may persist in the environment for as little as a day or as long as a month. The lack of persistence of the insecticidal protoxin in the natural environment means that natural selection of resistant insects is highly unlikely.

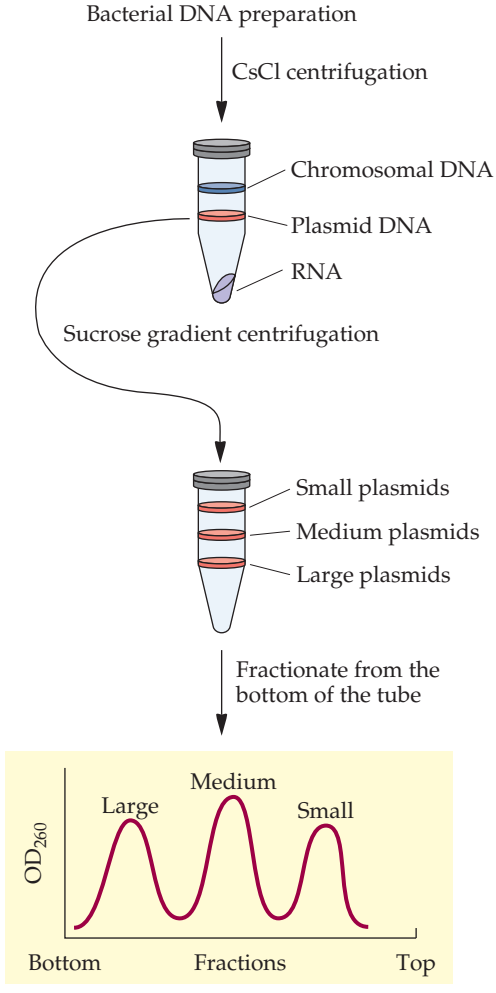
Toxin Gene Isolation

To develop *B. thuringiensis*-based insecticides that have greater potencies and broader host ranges, it is necessary to isolate and characterize the protoxin gene(s). For the initial isolation of insecticidal protoxin genes, the first step was to determine whether the toxin genes are located on a plasmid or on the chromosomal DNA. To test for plasmid-borne toxin genes, the source *B. thuringiensis* strain was conjugated with a strain that lacks insecticidal activity. If the latter strain acquired the ability to synthesize the insecticidal toxin, then the toxin gene(s) was most likely present on a plasmid, because the transfer of chromosomal DNA during conjugation is a rare event.

The procedure for isolating a protoxin-encoding DNA sequence is a familiar one. *B. thuringiensis* cells are grown in laboratory culture and lysed. The total cellular DNA is isolated and separated into plasmid and chromosomal DNA fractions by cesium chloride (CsCl) gradient centrifugation. When the protoxin gene is part of the genome, a clone bank is constructed from the chromosomal DNA. When the toxin gene(s) is plasmid encoded, the plasmid DNA can be further fractionated by sucrose gradient centrifugation, which separates different plasmids according to their sizes and enriches for the DNA that serves as the starting material for the isolation of a protoxin gene(s) (Fig. 16.4).

B. thuringiensis subsp. *kurstaki* contains an insecticidal protoxin gene on one of seven different plasmids that are approximately 2.0, 7.4, 7.8, 8.2, 14.4, 45, and 71 kilobase pairs (kb) in length. To determine which *B. thuringiensis* subsp. *kurstaki* plasmid carries the protoxin gene, following sucrose gradient centrifugation, the plasmid DNA sample is divided into three fractions that contain, respectively, the small (2.0-kb), medium-sized

FIGURE 16.4 Procedure for the isolation and partial enrichment of plasmid DNA fractions from a microorganism with a number of different plasmids, one of which encodes an insecticidal protoxin. OD_{260}, optical density at 260 nm.



(7.4-, 7.8-, 8.2-, and 14.4-kb), and large (45- and 71-kb) plasmids. The fraction with the small plasmid is discarded, because the plasmid is too small to encode a protein equivalent to the 130-kDa protoxin. A protein of this size requires at least 4.0 kb of coding DNA. The medium and large plasmid fractions are each partially digested with the restriction enzyme Sau3AI and then ligated into the BamHI site of plasmid pBR322. In the original experiments, these clone banks were transformed into *Escherichia coli*, and then the colonies were screened immunologically (see chapter 3) by the following procedure to detect clones that expressed a Cry protein and therefore carried a *cry* gene.

1. Colonies were transferred from agar plates to a nitrocellulose membrane.
2. The transferred colonies were lysed with organic solvents.
3. All available sites on the membrane to which primary and secondary antibodies could potentially bind (nonspecifically) were blocked by treating the membrane with bovine serum albumin (which bound to the nonspecific sites and prevented antibodies from binding to those sites).
4. The bovine serum albumin-treated membranes were treated with rabbit antiserum that contained antibodies against the insecticidal toxin. The antibodies bound only to the insecticidal toxin and not to any nonspecific sites on the membrane.
5. The membranes were washed to remove unbound antibodies and then treated with ^{125}I-labeled *Staphylococcus aureus* protein A, which bound only to the Fc portion of the bound antibodies and not to any nonspecific sites on the membrane.
6. Spots on the membrane corresponding to colonies that actively synthesized the insecticidal toxin were visualized by autoradiography.

The isolated protoxin gene was then used as a DNA hybridization probe to localize the *cry* gene to the 71-kb plasmid of *B. thuringiensis* subsp. *kurstaki*. Similar cloning and screening procedures have been used to isolate other *B. thuringiensis* toxin genes. However, given the current knowledge regarding sequence similarity among *B. thuringiensis* protoxin genes, the cloning and screening of these genes are more easily achieved by using polymerase chain reaction (PCR) and DNA hybridization techniques.

Engineering of *B. thuringiensis* Toxin Genes

Once the isolation and sequencing of a toxin gene were accomplished, the complete amino acid sequence was determined. Comparisons of amino acid sequences from other *B. thuringiensis* toxin proteins showed that a common toxic domain exists in these strains. Moreover, a subcloned segment of the complete protein-coding sequence produced a truncated protein that retained full insecticidal activity. Thus, an intact protoxin gene, a portion of one, or a chemically synthesized coding sequence can be used for further genetic manipulation.

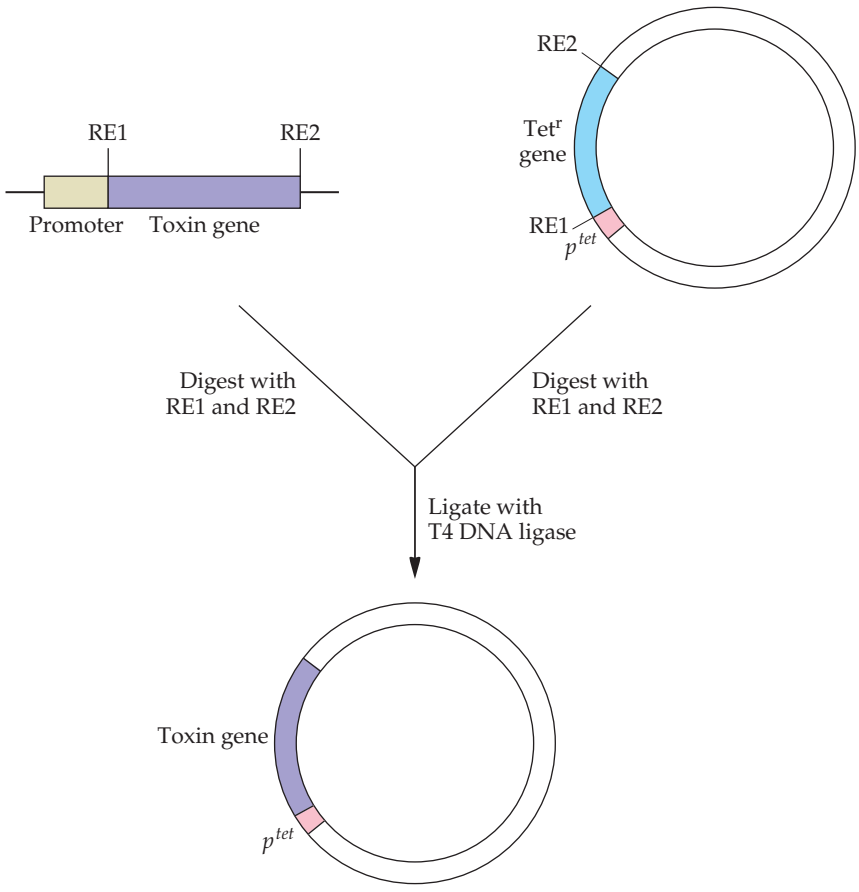
Synthesis during Vegetative Growth

Under normal conditions, most *B. thuringiensis* protoxin proteins are synthesized only during the sporulation phase of growth. In other words, only a portion of the growth cycle of the organism is devoted to parasporal

crystal production. It might therefore be advantageous, in terms of increased yield and decreased production time, to have the toxin gene transcribed and translated during vegetative growth. Furthermore, production of the insecticidal toxin during vegetative growth would permit the toxin to be synthesized by a continuous fermentation process, potentially significantly decreasing the cost of producing it. Continuous fermentations are carried out with smaller-scale—and therefore less expensive—bioreactors and downstream processing equipment than conventional batch fermentations. (See chapter 17 for additional details.)

During the sporulation of *B. thuringiensis*, a specific transcription initiation factor (sigma factor) interacts with the promoters of genes that are active only within this phase of the bacterial life cycle. This factor turns on the transcription of the messenger RNAs (mRNAs) that are unique to sporulation. In fact, when a *B. thuringiensis* toxin gene with its sporulation-specific promoter was cloned and expressed in *Bacillus subtilis*, *Bacillus megaterium*, or *B. thuringiensis*, gene transcription occurred only during

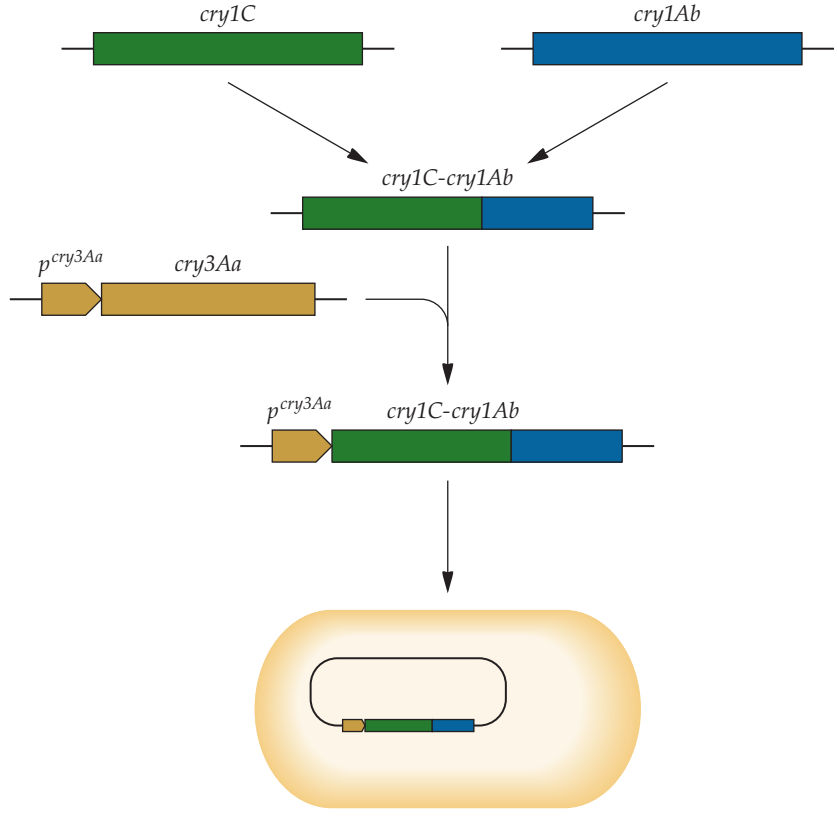
FIGURE 16.5 Procedure for subcloning the *B. thuringiensis* subsp. *kurstaki* insecticidal toxin gene so that it is expressed constitutively under the control of the promoter of the tetracycline resistance (Tet^r) gene (p^{tet}). The isolated *B. thuringiensis* toxin gene is removed from its promoter by digestion of the isolated DNA fragment with restriction enzymes RE1 and RE2. It is spliced by T4 DNA ligase into the plasmid vector downstream from p^{tet} in place of the tetracycline resistance gene, which has been removed by digestion with restriction enzymes RE1 and RE2.



sporulation. Thus, to express a *B. thuringiensis* insecticidal toxin during vegetative growth, it is necessary to place the toxin-producing gene(s) under the control of a promoter that is active during vegetative growth.

When a DNA fragment containing a toxin gene that lacked its native promoter was cloned into a plasmid under the control of a continuously active, constitutive promoter from a tetracycline resistance gene that had been originally isolated from a *Bacillus cereus* plasmid and reintroduced into *B. thuringiensis,* active toxin protein was produced continuously throughout the growth cycle, including both the vegetative and sporulation phases (Fig. 16.5). In addition, when the construct was used to transform a sporulation-defective mutant of *B. thuringiensis,* toxin synthesis occurred in the absence of sporulation. Under these conditions, toxin synthesis is more efficient than in wild-type cells, i.e., the final yield of protein is greater in the transformed cells, and less time and substrate are required to produce the toxin. A refinement of this system might entail integration of this vegetatively expressed toxin gene into the chromosomal DNA of the sporulation-defective *B. thuringiensis* host. This manipulation would ensure that the insecticidal toxin gene is not lost because of plasmid instability during a continuous fermentation process.

FIGURE 16.6 Construction of a strain of *B. thuringiensis* with greater potency and UV resistance. The C-terminal third of the *cry1Ab* gene was spliced together with the N-terminal two-thirds of the *cry1C* gene, all under the control of the *cry3Aa* promoter *(p^{cry3Aa}),* and then integrated into the chromosomal DNA of a sporulation-minus strain of *B. thuringiensis.* Adapted from Sanchis et al., *Appl. Environ. Microbiol.* **65:**4032–4039, 1999.



Unlike that of most other *B. thuringiensis* toxin protein *(cry)* genes, the expression of *cry3A* is normally controlled by a vegetative promoter, rather than by a sporulation-specific promoter. The *cry3A* gene encodes a toxin that is directed against coleopteran larvae. When a mutant strain of *B. thuringiensis* that was unable to form spores was transformed with a plasmid carrying a cloned *cry3A* gene, the insecticidal toxin was both overproduced and stabilized in comparison to when this protein was produced in the wild-type strain. This result suggests that other *cry* genes that are normally expressed only during sporulation could be placed under the control of the *cry3A* promoter and overproduced by expressing these constructs in a sporulation-defective *B. thuringiensis* mutant.

In one experiment, a chimeric *cry1C–cry1Ab* gene was constructed, placed under the transcriptional control of the vegetative *cry3A* promoter, and then integrated into the chromosomal DNA of a nonsporulating derivative of *B. thuringiensis* subsp. *kurstaki* (Fig. 16.6). The chimeric *cry1C–cry1Ab* gene consisted of approximately 2.2 kb of DNA from the *cry1C* gene and 1.3 kb of DNA from the *cry1Ab* gene. Although the mature toxin that is produced following proteolytic cleavage of the hybrid protoxin is identical to the toxin that is produced from the *cry1C* gene, this toxin was found to be considerably more active than Cry1C (Table 16.3). Thus, depending upon the insect tested, Cry1C–Cry1Ab was 3 to 34 times more active than Cry1C. This seemingly strange result probably occurs because of the increased stability to proteolytic digestion of the Cry1Ab portion of the hybrid protoxin protein, which is removed upon activation of the protoxin. The nonsporulating *B. thuringiensis* host strain had a disrupted *sigK* gene, which encodes the sigma factor σ^{28}, which is required for sporulation-specific transcription. Other workers have created nonsporulating *B. thuringiensis* strains by inserting modified protoxin genes into the late-stage sporulation gene $spoV_{Bt1}$. Since the chimeric protoxin Cry1C–Cry1Ab was encapsulated with the bacterial cells, the protein was considerably more resistant to the degradative effect of ultraviolet (UV) light, which rapidly inactivates the protoxin that is normally secreted outside of the bacterial cell during sporulation. In addition to increased potency and greater UV resistance, the environmental persistence of the nonsporulating mutant was significantly decreased compared with that of the sporulating wild-type strain. This may actually be an advantage, since it is less likely that the nonsporulating mutant will transfer any of its DNA to other organisms in the environment.

Broadening the Spectrum of Target Insects

Because many crops are attacked by more than one insect species, it would be advantageous, if feasible, to create microbial insecticides that are effective

TABLE 16.3 Activities of Cry1C and a chimeric Cry1C-Cry1Ab toxin against three insect species

Insect species	LC_{50} of Cry1C (ng)	LC_{50} of Cry1C-Cry1Ab (ng)
Spodoptera littoralis	378	103
Plutella xylostella	174	4.6
Ostrinia nubilalis	3,200	822

The LC_{50} values reflect the amount of insecticidal toxin (in nanograms) required to kill half of the insect population being tested under a defined set of conditions. The smaller the LC_{50}, the more potent the toxin.

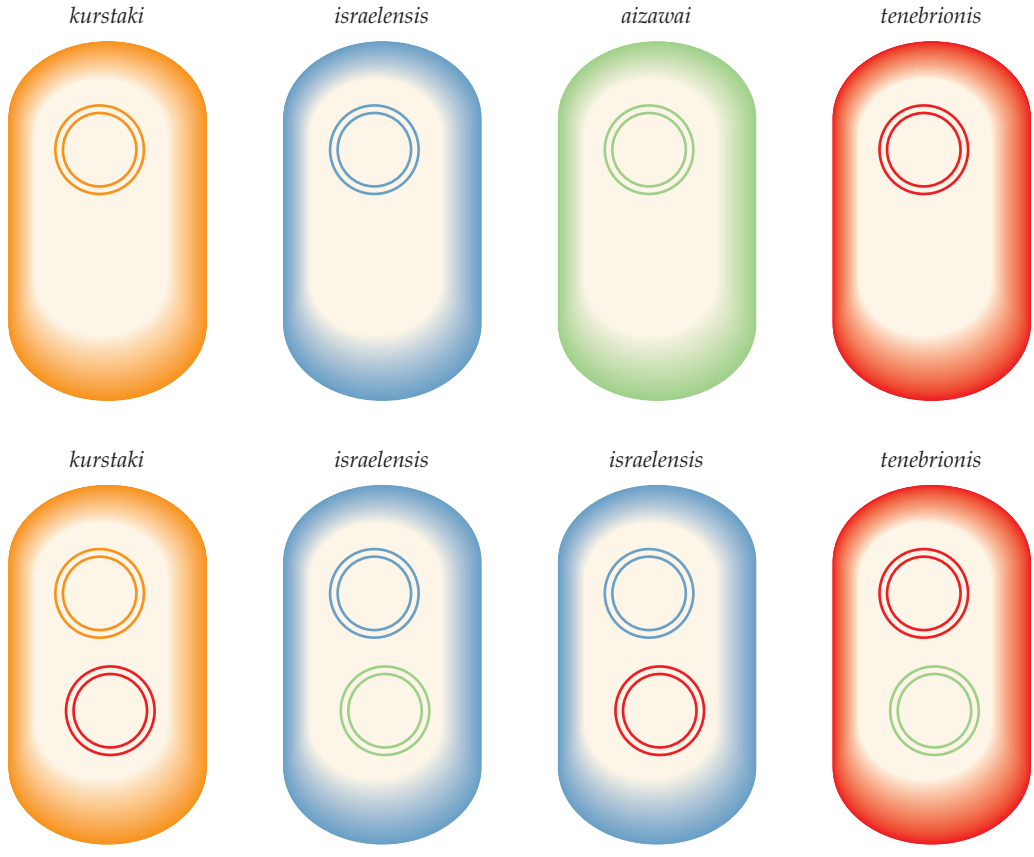


FIGURE 16.7 Naturally occurring and transformed subspecies of *B. thuringiensis*. The oval shape represents a bacterium, while the circle represents an insecticide-encoding plasmid. The plasmids are colored the same as the bacterium in which the toxin gene originated.

against a broad spectrum of target insects. Such a broad-specificity molecule could be obtained (1) by transferring the gene for a particular toxin, e.g., one against diptera, into a *B. thuringiensis* strain that normally synthesizes a different species-specific toxin, e.g., one against coleoptera; (2) by fusing portions of two different species-specific toxin genes to one another so that a unique dual-acting toxin (hybrid toxin) is produced; or (3) by modifying the portion of the insecticidal toxin that is responsible for binding to insect gut epithelial cell receptors.

Transferring *cry* genes. To test whether the spectrum of target insect pests could be widened, the insecticidal toxin genes from *B. thuringiensis* subsp. *aizawai* and *tenebrionis* were cloned into shuttle vectors that could be maintained in both *B. thuringiensis* and *E. coli*. These genetic constructs were then introduced by electroporation into *B. thuringiensis* subsp. *kurstaki*, *israelensis*, and *tenebrionis* (Fig. 16.7), and all the transformed strains were tested for toxicity to three different insect species.

In each case, the toxicity of the native host toxin protein(s) was maintained, and in most cases, the introduced toxin gene also expressed an active toxin with the same specificity as the toxin produced by the source bacterium (Table 16.4). In addition, and surprisingly, when the *B. thuringiensis*

subsp. *tenebrionis* toxin gene was introduced into *B. thuringiensis* subsp. *israelensis*, the resultant transformant was somewhat toxic to *Pieris brassicae*, the cabbage white butterfly, against which neither of the gene products alone has insecticidal activity.

In many instances, introduced plasmid vectors carrying isolated *cry* genes are unstable in *B. thuringiensis*. Often, in the absence of selective pressure, all or a portion of these plasmids are lost. The problem of plasmid instability with introduced genes was overcome by integrating cloned *cry* genes into the chromosomal DNA of the host cell. One group of researchers attempted to broaden the insect specificity of a strain of *B. thuringiensis* subsp. *kurstaki*, which normally carries five different insecticidal toxin genes, *cry1Aa*, *cry1Ab*, *cry1Ac*, *cry2Aa*, and *cry2Ab*. While the products of these *cry* genes are active against a variety of lepidopteran species, they are not effective against *Spodoptera* spp. Therefore, a *cry1Ca* gene, which is normally found only in *B. thuringiensis* subsp. *aizawai* and *entomocidus*, was introduced into the chromosomal DNA of the *B. thuringiensis* subsp. *kurstaki* host strain. The transformed *B. thuringiensis* subsp. *kurstaki* strain showed a sixfold increase in its ability to kill *Spodoptera exigua* (beet armyworm) larvae.

Modifying the loop regions of domain II. The toxic moiety of many Cry proteins is composed of three separate domains. Domain II is involved in the specific binding of the toxin to protein receptors that are found on the surfaces of insect midgut epithelial cells, although domain III may also play a role in receptor binding. Following binding, a portion of domain I, in the N-terminal region of the toxin, inserts into the membrane. It is believed that the interaction of portions of domain I from several toxin molecules interact to make up the pore. Domain III, which is located at the C-terminal end of the toxin molecule, is also thought to be involved in pore function.

Modification of *cry* genes to increase the binding of the Cry protein to receptors generally leads to an increase in insecticidal activity. In particular,

TABLE 16.4 Toxicities of naturally occurring and transformed subspecies of *B. thuringiensis* against the insects *Pieris brassicae* (cabbage white butterfly), *Aedes aegypti* (mosquito), and *Phaedon cochleariae* (beetle)

Source of toxin		Toxicity to:		
Host DNA	**Introduced DNA**	***Pieris***	***Aedes***	***Phaedon***
aizawai	None	++	+	–
israelensis	None	–	++	–
israelensis	*aizawai*	++	++	–
israelensis	*tenebrionis*	+	++	++
kurstaki	None	++	+	–
kurstaki	*tenebrionis*	++	+	++
tenebrionis	None	–	–	++
tenebrionis	*aizawai*	++	+	+

Adapted from Crickmore et al., *Biochem. J.* **270:**133–136, 1990.

In these experiments, the toxicity was graded as follows: ++, 0 to 5% of the leaf was consumed (for *Phaedon and Pieris*) or 100% mortality occurred within 1 hour (*Aedes*); +, 5 to 50% of the leaf was consumed (*Phaedon* and *Pieris*) or 50 to 100% mortality occurred within 24 hours (*Aedes*); –, 50% of the leaf was consumed (*Phaedon* and *Pieris*) or no mortality occurred within 24 hours (*Aedes*). The test plant was either cabbage leaf (for *Pieris*) or turnip leaf (for *Phaedon*).

FIGURE 16.8 Schematic representation of the three-dimensional structure of the Cry19Aa protein, highlighting loops 1 and 2, which were modified to alter the insecticide's specificity. Adapted from Abdullah and Dean, *Appl. Environ. Microbiol.* **70:**3769–3771, 2004, with permission.

modification of domain II is an effective means of increasing Cry toxicity to particular insects. In one series of experiments, researchers modified the insect specificity of Cry19Aa. This was done by directed mutagenesis of the *cry19Aa* gene, replacing a nucleotide sequence that encoded the amino acids Ser–Tyr–Trp–Thr in loop 1 of domain II with a sequence encoding Tyr–Gln–Asp–Leu–Arg and deleting a sequence in loop 2 encoding Tyr–Pro–Trp–Gly–Asp (Fig. 16.8). The decisions regarding which sequences to alter were based on computer models comparing the three-dimensional structure of Cry19Aa with the structure of Cry4Ba. These changes—alterations of both loop 1 and loop 2 were required—yielded a modified Cry19Aa protein whose insecticidal activity against the mosquito *Aedes aegypti* was increased more than 42,000-fold while its activity against other insects was essentially unchanged. This work suggests that it may be possible to rationally engineer various Cry toxins to have desired activities by manipulating specific amino acid sequences within the protein loops. However, even if the genetic manipulations are successful and designer-engineered Cry proteins are attainable, it remains to be seen whether the general public and the regulatory authorities in various countries will embrace this technology, which would include releasing genetically manipulated bacteria into the environment.

Improving Delivery of a Mosquitocidal Toxin

The *B. thuringiensis* subsp. *israelensis* insecticidal protein is highly toxic when ingested by mosquito larvae. Since 1982, it has been used successfully worldwide to control mosquitoes and blackflies. However, the parasporal crystal of this species sinks rapidly after it is sprayed on water, which effectively removes it from the feeding area of mosquito larvae and dramatically decreases its efficacy as a mosquitocide. To overcome this shortcoming, several approaches have been attempted. Currently, *B. thuringiensis* subsp.

israelensis insecticidal protein is available as granules or as slow-release rings or brickettes, which float on the surface of water. Another solution is to introduce the insecticidal toxin gene into organisms that are common food sources for mosquito larvae. Good candidate organisms for this purpose include *Synechocystis* and *Synechococcus* spp., which are photosynthetic cyanobacteria that proliferate near the water surface, where there is sufficient light for their growth and where mosquito larvae are normally found. Another organism with the potential to be a host for the expression of foreign insecticidal toxin genes is *Caulobacter crescentus,* an aquatic bacterium that generally is widely distributed throughout aquatic environments where mosquito larvae feed. The toxin gene from *B. thuringiensis* subsp. *israelensis* was introduced into and expressed in these organisms. In laboratory trials, the insecticidal toxin that was produced by either transformed cyanobacteria or *C. crescentus* was toxic to mosquito larvae. However, in field trials, transformed cyanobacteria or *C. crescentus* expressing *B. thuringiensis* insecticidal toxin genes had poor viability, and the cloned genes were expressed at a low level.

A possible alternative host for the expression of mosquitocidal *cry* genes is *Asticcacaulis excentricus,* a gram-negative aerobic bacterium that is found in aqueous environments near the surface of the water. In a series of experiments, *A. excentricus* was transformed with a broad-host-range plasmid vector that carried the genes for mosquitocidal toxin proteins produced by a strain of *Bacillus sphaericus* (a bacterium similar to *B. thuringiensis*) under the control of the *tac1* promoter, which is a variant of the *tac* promoter. This transformant produced insecticidal toxin proteins of 51 and 42 kDa and was almost as toxic to *Anopheles* and *Culex* mosquito larvae as the naturally occurring high-toxicity strains of *B. sphaericus*. However, unlike *B. sphaericus*, *A. excentricus* does not sink when it is sprayed onto

MILESTONE

Cloning and Expression of the *Bacillus thuringiensis* Crystal Protein Gene in *Escherichia coli*

H. E. Schnepf and H. R. Whiteley
Proc. Natl. Acad. Sci. USA **78:**2893–2897, 1981

Although it had been well known for a long time that the parasporal crystal that is produced by *B. thuringiensis* contained insecticidal activity, it took scientists many years before the conditions for solubilizing the crystal were discovered and the insecticidal protein toxin was isolated in a pure form. Moreover, although some *B. thuringiensis* mutants that were defective in the synthesis of the parasporal crystal were known, protocols for the genetic transformation of *B. thuringiensis* were not well developed. Therefore, when Schnepf and Whiteley decided to isolate *B. thuringiensis* insecticidal toxin protein genes, they were limited to screening *E. coli* transformants carrying *B. thuringiensis* DNA either immunologically, using antibodies directed against the whole crystal, or by the insecticidal activity of extracts of the transformants. Since evidence at the time suggested that the insecticidal toxin was probably plasmid encoded, clone banks were constructed from fractionated plasmid preparations with the idea of significantly enriching the clone banks for the presence of insecticidal toxin genes. Moreover, despite some concerns that antibodies against the whole (*B. thuringiensis*-produced and glycosylated) crystal protein might not interact with (*E. coli*-produced and nonglycosylated) crystal protein subunits in solution, this did not turn out to be a problem. Finally, extracts of the particulate fraction of *E. coli* transformants carrying the *B. thuringiensis* insecticidal toxin gene were found to be toxic to susceptible insects. This first cloning of a *B. thuringiensis* insecticidal toxin gene made it clear to workers in this field that these genes could be isolated in a straightforward manner and provided an impetus for increased activity both in the search for new strains of *B. thuringiensis* and in studies of the biochemistry of the insecticidal toxin.

ponds infested with mosquito larvae. Moreover, *A. excentricus* is inexpensive to produce, as it can be grown on much simpler media than either *B. sphaericus* or *B. thuringiensis*. It does not have a high level of protease activity, so the insecticidal toxin is not readily degraded. It is well adapted to environments such as those near the surface of standing water that are exposed to relatively high levels of UV light. Thus, *A. excentricus* cells should not be as sensitive to inactivation by UV light as those of either *B. sphaericus* or *B. thuringiensis*. However, to use a genetically engineered strain of *A. excentricus* to control mosquito populations in the environment, it will be necessary to integrate the insecticidal toxin genes into the chromosomal DNA without any antibiotic resistance genes.

Protecting Plant Roots

Insects that attack the roots of plants are not affected by *B. thuringiensis*-based insecticides that are sprayed onto leaves and shoots. However, it is possible to introduce the toxin gene from a *B. thuringiensis* strain into a bacterial species that colonizes the region adjacent to plant roots (the rhizosphere). The engineered bacteria could be introduced into the soil, where they would synthesize the insecticidal toxin and release it into the area immediately surrounding the plant roots, thereby conferring protection against root-attacking insects. In addition, as long as the engineered bacteria were able to persist in the soil, they would continue to synthesize the insecticidal toxin, thus obviating the need for repeated spraying of either biological or chemical insecticides. This approach has been tested on a small scale. The gene for the *B. thuringiensis* subsp. *kurstaki* insecticidal toxin was integrated into the chromosomal DNA of a strain of *P. fluorescens* that colonizes corn (maize) roots. The integration of the toxin gene was achieved as follows (Fig. 16.9).

1. A transposon Tn*5* element that had been cloned into a plasmid was genetically modified by altering portions of its left and right borders and deleting its transposase gene. Such an altered Tn*5* element cannot be excised from the plasmid, even by exogenous transposase, because the left and right borders are not are recognized by the transposase.
2. An isolated *B. thuringiensis* subsp. *kurstaki* insecticidal toxin gene was spliced into the middle of the altered Tn*5* element on the plasmid and placed under the control of a constitutive promoter.
3. A wild-type Tn*5* element was transposed into the chromosome of the root-colonizing strain of *P. fluorescens*.
4. The plasmid carrying the altered Tn*5* element with the inserted toxin gene was introduced into *P. fluorescens* carrying the integrated wild-type Tn*5* element.
5. Homologous recombination by means of a double crossover between the nontransposable Tn*5* element on the plasmid that carried the toxin gene and the chromosomally integrated wild-type Tn*5* led to the integration of the altered Tn*5* with the toxin gene into the chromosomal DNA, with the concomitant loss of the wild-type Tn*5* element.

In this form, the toxin gene is unlikely to be lost either during large-scale laboratory growth or after release of the engineered microorganism into the environment. Also, the probability of transfer of the toxin gene to

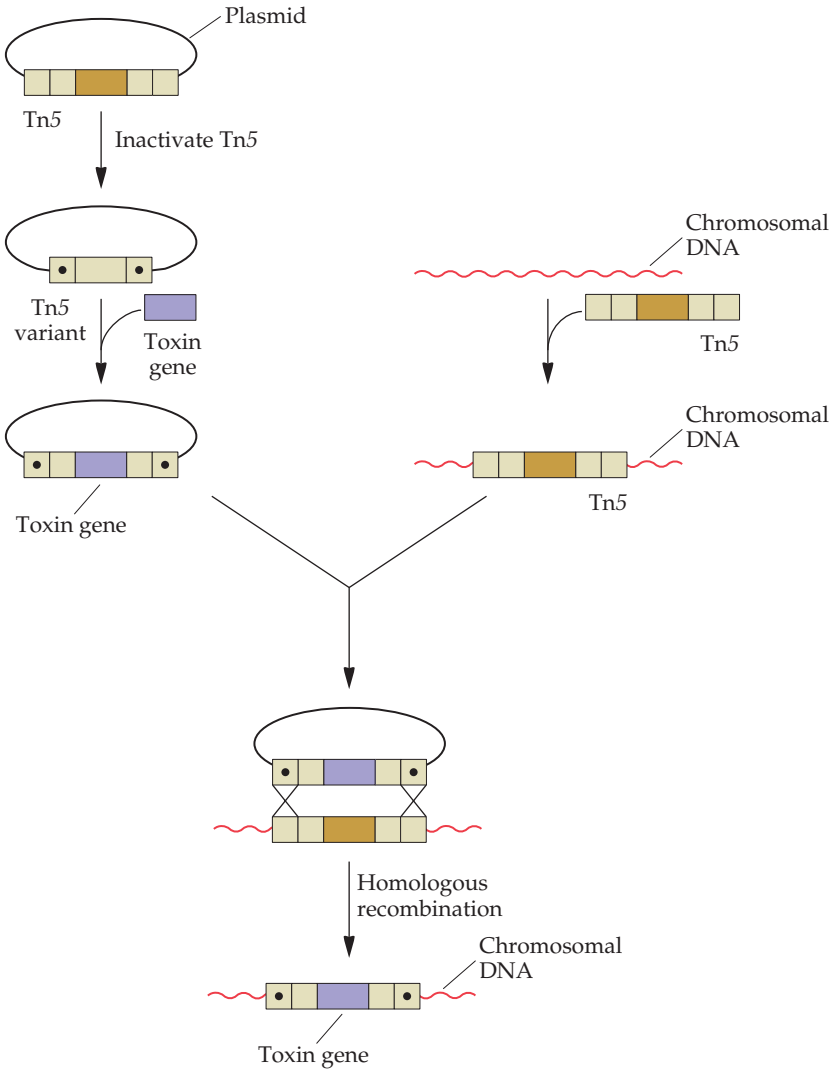


FIGURE 16.9 Procedure for the development of a genetically engineered *P. fluorescens* strain that carries a copy of the *B. thuringiensis* insecticidal toxin gene integrated into its chromosomal DNA. The *B. thuringiensis* insecticidal toxin gene is cloned into an excision-defective variant of Tn*5* on a plasmid. This construct is introduced into a *P. fluorescens* strain containing a wild-type Tn*5* sequence that has been integrated into its chromosomal DNA. By homologous recombination, the excision-defective Tn*5* element carrying the *B. thuringiensis* insecticidal toxin gene becomes integrated into the *P. fluorescens* chromosome.

other microorganisms in the environment is very low. Laboratory trials showed that the engineered *P. fluorescens* was toxic to tobacco hornworm larvae. However, the ability of this genetically manipulated microorganism to minimize root damage from insect predation remains to be tested in the greenhouse and in open-field trials.

In other laboratories, various *B. thuringiensis* insecticidal toxin genes have been introduced into the chromosomal DNA of a number of different microorganisms. For example, the *cry1Ac* genes were introduced into a strain of *P. fluorescens* and found to protect sugarcane plants against the sugarcane borer, *Eldana saccharina*. Also, when this gene was used to

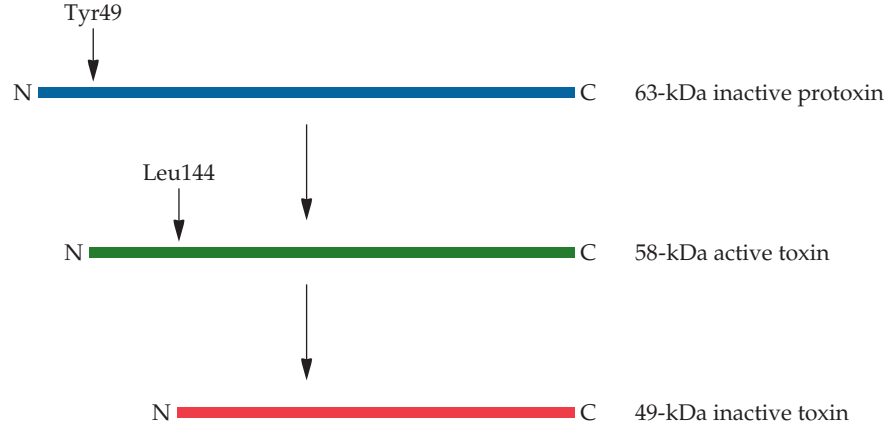
transform *Clavibacter xyli* subsp. *cynodontis,* a bacterium that normally inhabits the xylem of Bermuda grass, the bacterium protected corn plants from damage caused by the European corn borer, *Ostrinia nubilalis*.

Protoxin Processing

Occasionally, during the proteolytic processing of the protoxin to the active toxin, the insect midgut proteases continue to cleave the toxin protein and thereby render it inactive (Fig. 16.10). This degradation process occurred when the 63-kDa Cry2Aa1 protoxin was treated with the midgut juices of the gypsy moth *(Lymantria dispar)*. The midgut juices, which contained the protoxin-processing protease, first cleaved the protoxin on the C-terminal side of Tyr49, producing the active 58-kDa Cry2Aa1 toxin. However, continued incubation of the toxin with the midgut enzymes resulted in the cleavage of the toxin on the C-terminal side of Leu144. This second cleavage inactivated the toxin, producing an inactive 49-kDa protein, dramatically reducing its effectiveness. To ascertain that this result was not an artifact, researchers radiolabeled the Cry2Aa1 protoxin and showed that this excessive cleavage also occurred in vivo. To try to prevent the production of the inactive form of the toxin, the protoxin gene was altered in five different ways by site-directed mutagenesis. The amino acid residue in position 144 of the protoxin was changed from leucine to aspartic acid, alanine, glycine, histidine, or valine. All of the mutant proteins yielded a higher level of active toxin than the native form, and with the exception of the leucine-to-histidine change, the active mutant toxin proteins were no longer cleaved to an inactive form.

Since the C-terminal half of many *B. thuringiensis* insecticidal protoxins is not toxic to insects, it would be advantageous if that half of the protein could be eliminated. Then, the cellular resources that had previously gone into synthesizing the C-terminal half of the protoxin might be used to synthesize more of the active toxin, thereby increasing the amount of toxin that a bacterium might produce. Unfortunately, when such truncated *cry1* genes

FIGURE 16.10 Activation and subsequent cleavage of *B. thuringiensis* Cry2Aa1 protoxin by *L. dispar* midgut enzymes. Activation occurs by cleavage of the protoxin on the C-terminal side of Tyr49. Toxin inactivation occurs when the protein is cleaved on the C-terminal side of Leu144. When Leu144 is changed to one of several different amino acids, the toxin is both active and resistant to further proteolytic cleavage.



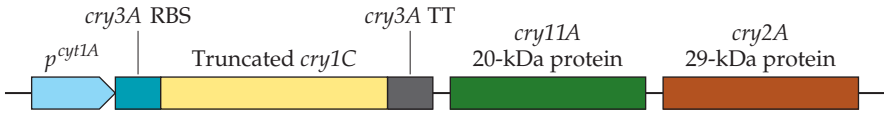


FIGURE 16.11 A genetic construct that produces a high level of stable truncated Cry1C protein. This construct includes the promoter from a *cyt1A* gene (p^{cyt1A}), a ribosome-binding site (RBS) from *cry3A* that stabilizes mRNA, a *cry3A* gene transcription terminator region (TT), a gene from the *cry11A* operon encoding a 20-kDa chaperone-like protein that stabilizes the structure of the Cry1C protein, and a gene from the *cry2A* operon encoding a 29-kDa protein that facilitates protoxin crystal formation. Each of the last two genes has its own promoter, ribosome-binding site, and transcription termination site (not shown).

were expressed in *B. thuringiensis*, the toxin yields were low and crystals did not form. To remedy this situation, several genetic elements that were known to enhance the synthesis and crystallization of "naturally truncated" Cry proteins were tested both separately and together in order to improve the stability and yield of truncated Cry1C (Fig. 16.11). The construct that significantly increased both the stability and yield of truncated Cry1C protein, with the truncated protein now forming crystals within sporulated cells, contained a number of different genetic elements, including the gene for a 20-kDa chaperone-like protein and a 29-kDa protein that facilitated protoxin crystal formation.

Preventing the Development of Resistance

When *B. thuringiensis* subsp. *kurstaki* is used as an insecticide in a controlled environment where there is no sunlight to rapidly break down the protoxin, e.g., when stored grain is treated to protect it against insect predation, resistant target insects develop within a few generations. This inherited resistance is typically due to an alteration in a midgut membrane protein that normally acts as a receptor for the *B. thuringiensis* subsp. *kurstaki* toxin. Resistant insects accumulate because the protoxin persists under these conditions and selects for resistant individuals. The lesson here is that the simplest way to avoid selecting for insects that are resistant to *B. thuringiensis* subsp. *kurstaki* in the absence of sunlight is to limit the use of this bacterium to field applications. However, extensive annual use, even under natural conditions, may result in a level of persistence high enough to allow selection to occur. Certainly, as larger quantities of *B. thuringiensis* are used over a wider geographical area, the probability that resistant strains of insects will be selected will increase. Various ways to avert this problem are being examined. These strategies, which may be utilized either with *B. thuringiensis* that is sprayed or with transgenic plants expressing the insecticidal toxin, include the following.

1. The use of two or more *B. thuringiensis* insecticidal toxins at the same time. Provided that the toxins bind to different receptors, it is extremely unlikely that an insect will develop resistance to both toxins at the same time. When this approach is used in transgenic plants, it is often called "gene pyramiding."
2. Application of a *B. thuringiensis* insecticidal toxin along with traditional chemical insecticides. The idea here is that almost no insect survives these two very different treatments, and resistance does not develop to either. Transgenic plants that produce a

B. thuringiensis insecticidal toxin are commonly treated with chemical insecticides. However, the number of chemical insecticide treatments is significantly reduced when the plants produce a *B. thuringiensis* insecticidal toxin. In Florida, nontransgenic corn plants often require as many as 10 sprayings of chemical insecticides per growing season. Plants that produce a *B. thuringiensis* insecticidal toxin are more likely to be sprayed with chemical insecticides only about three or four times a season.

3. Application of a *B. thuringiensis* insecticidal toxin at the same time as another biologically based insecticidal protein (typically isolated from plants; see chapter 18). Again, it is extremely unlikely that the target insects will survive both types of insecticides.
4. The use of two *B. thuringiensis* insecticidal toxins, one of which has had its toxin gene modified so that it binds to a different receptor than the other toxin.
5. The use of refugia (small tracts of land where the crop is not treated with the microbial insecticide). Approximately 20% of a crop is not sprayed with *B. thuringiensis* (or 20% is nontransgenic, with the remaining 80% of the plants being transgenic and producing a *B. thuringiensis* insecticidal toxin). The wild-type insects can proliferate in the absence of the *B. thuringiensis* insecticidal toxin, and only (a very small number of) mutant insects that are resistant to the high levels of *B. thuringiensis* insecticidal toxin survive in the presence of the toxin. Upon mating, the small number of resistant insects will all mate with sensitive insects, so that the next generation will contain mostly homozygous sensitive insects and a small number of heterozygous sensitive insects. This strategy assumes that resistance to the *B. thuringiensis* insecticidal toxin is inherited as a recessive trait. This approach has been used in the field for a number of years, with all of the available evidence indicating that little to no resistance to any *B. thuringiensis* insecticidal toxins has developed.

As noted above, fusion of the coding portions of the active regions of two different toxin genes is another way of generating a novel protein with extended toxicity. This idea has been examined in laboratory experiments. When a series of lepidopteran-specific hybrid toxins were constructed, some of them were more toxic than the products of either of the contributing genes by themselves, and in one case, a hybrid protein had acquired a totally new biological activity.

Generally, resistance to *B. thuringiensis* insecticidal toxins is the consequence of a mutation(s) that alters an insect midgut receptor protein(s) so that it no longer binds to the Cry protein. However, if a toxin gene were engineered so that the toxin bound to more than one midgut cell surface protein, then resistance might be less likely to arise, since it would require alterations to several proteins.

The insecticidal proteins Cry1Ca and Cry1Ea are both toxic to lepidoptera but have different species specificities. Cry1Ca is active against *S. exigua, Mamestra brassicae,* and *Manduca sexta,* while Cry1Ea is active only against *M. sexta*. In one experiment, hybrid Cry1Ca–Cry1Ea proteins were constructed and tested for their toxicities to different insect species, as well as for their abilities to bind to different receptors (Fig. 16.12). The hybrid toxin G27, which contained domain III from Cry1Ca, was toxic to *S. exigua*

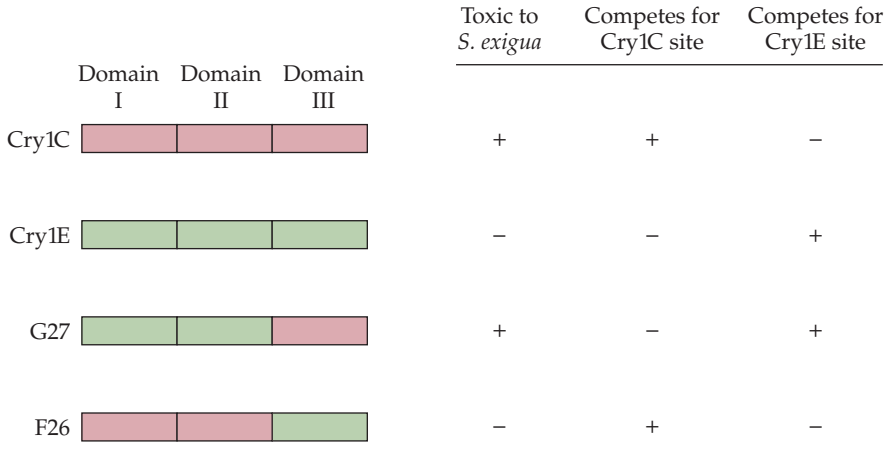


FIGURE 16.12 Toxicities and binding specificities of Cry1Ca, Cry1Ea, and hybrid toxins G27 and F26. A toxin–receptor-binding assay was used to determine binding specificities. Unlabeled Cry1Ca or Cry1Ea was added to a complex of the *S. exigua* midgut receptor protein and radiolabeled toxin, and the extent of binding of the radiolabeled toxin was determined. Adapted from Bosch et al., *Bio/Technology* **12:**915–918, 1994.

larvae even though it bound to the Cry1Ea receptor but not to the Cry1Ca receptor (Fig. 16.13). Conversely, the hybrid toxin F26 was not toxic to *S. exigua* larvae even though it bound to the Cry1Ca receptor. Since the Cry1Ca and G27 proteins bind to different insect midgut receptors (although both are toxic to *S. exigua*), either simultaneous or alternating treatments of *S. exigua* with these two *B. thuringiensis* insecticidal toxins might limit the development of strains that are resistant to the toxins. Resistance to both Cry1C and G27 would require mutations in two separate midgut proteins.

***B. thuringiensis* subsp. *israelensis* thwarts insect resistance.** In contrast to what has been observed with other strains of *B. thuringiensis*, no instances of field resistance of mosquitoes to *B. thuringiensis* subsp. *israelensis* have ever been reported, and only low levels of resistance have been observed in laboratory studies. This lack of insect resistance may reflect the fact that, in addition to synthesizing at least three different Cry proteins—Cry4A, Cry4B, and Cry11A—*B. thuringiensis* subsp. *israelensis* also produces Cyt1A, a highly hydrophobic endotoxin that is not at all homologous to any of the Cry proteins and appears to have a completely different mode of action. While Cry proteins bind to glycoproteins on the insect midgut epithelial membrane, the primary affinity of Cyt1A is the lipid component of the membrane, especially the unsaturated fatty acids. Cyt1A acts synergistically with the Cry proteins, and its presence may explain why mosquitoes do not develop resistance to the Cry proteins. In one series of experiments, using purified insecticidal proteins, it was demonstrated that with the addition of the Cyt1A protein, insects that had become resistant to Cry4A, Cry4B, and Cry11A (all of which are encoded by *B. thuringiensis* subsp. *israelensis*) were killed when they were treated with *B. thuringiensis* subsp. *israelensis*. Recent experiments suggest that, following the binding of Cyt1A to the midgut epithelial membrane, the protein can act as a receptor for some of the Cry proteins encoded by *B. thuringiensis* subsp.

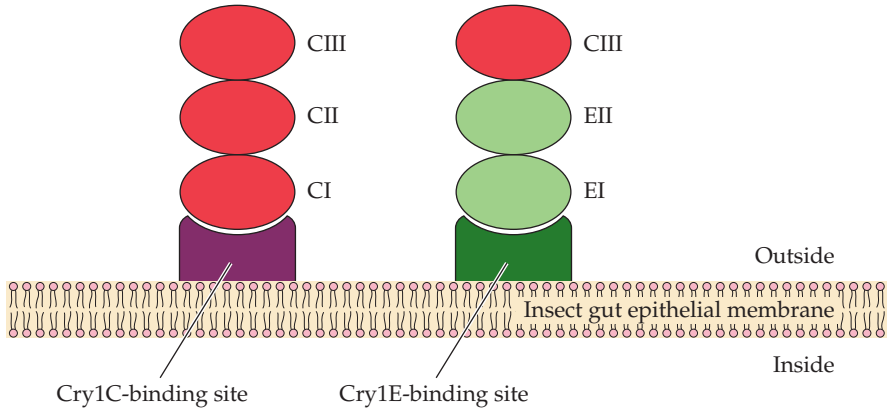


FIGURE 16.13 Schematic representation of Cry1C binding to the Cry1C receptor and hybrid G27 binding to the Cry1E receptor. For Cry1C, CI, CII, and CIII indicate that these domains all originate from Cry1C. For G27, EII and EIII indicate that these domains are from Cry1E.

israelensis. The reason that *B. thuringiensis* subsp. *israelensis* is a highly effective insect pathogenic bacterium may be because the strain not only carries several insecticidal proteins, but also contains a protein that acts as the receptor for these insecticidal proteins. It is therefore extremely unlikely that any target insect will be able to develop resistance to *B. thuringiensis* subsp. *israelensis*. To capitalize on the advantage that the Cyt1A protein provides to *B. thuringiensis* subsp. *israelensis*, genes for Cyt1A and Cry1Ac (which is typically found in *B. thuringiensis* subsp. *kurstaki* strains and targets lepidopteran larvae) were used to transform a strain of *B. thuringiensis* (Fig. 16.14). The combination of these two proteins in one strain was highly toxic to the diamondback moth (*Plutella xylostella*), a lepidopteran species. On the other hand, strains that expressed one or the other of these proteins, but not both, required extremely high levels of the proteins before any toxicity could be detected. These results suggest, in this case, that Cyt1A (which usually targets diptera) is uncharacteristically behaving as a receptor for Cry1Ac (a lepidopteran toxin). It will be exciting to ascertain whether this synergism can be extended to other Cry proteins combined with Cyt1A.

Other strategies that have been proposed as a means of avoiding the development of insects that are resistant to *B. thuringiensis* insecticidal toxins include alternating the strain of *B. thuringiensis* that is employed from one season to the next, alternating *B. thuringiensis* treatment with the use of chemical or other biological insecticides, or applying mixtures of different strains of *B. thuringiensis*.

Improved Biocontrol

Insects such the sugarcane borer *(E. saccharina)* that attack the internal regions of plants such as sugar cane are not affected by *B. thuringiensis*-based insecticides that are sprayed onto leaves and shoots. However, it is possible to introduce the toxin gene from a *B. thuringiensis* strain into a bacterium that colonizes either plant roots or interior surfaces. In these instances, the insecticidal toxin is delivered to the part of the plant that is normally attacked by the insect.

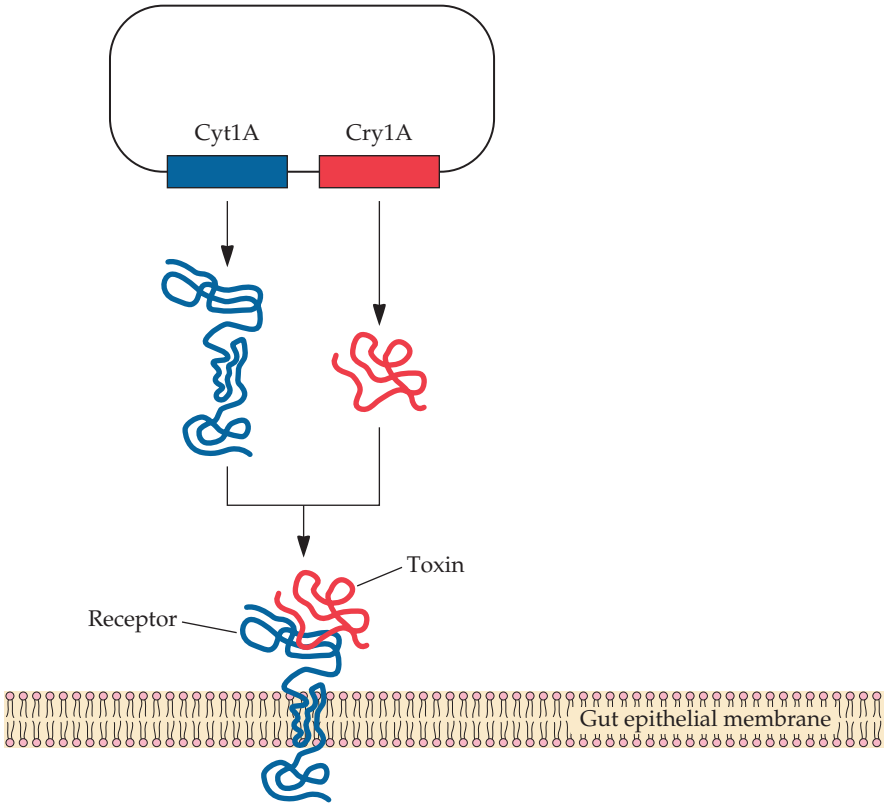


FIGURE 16.14 Schematic representation of an engineered strain of *B. thuringiensis* encoding both Cyt1A (which is from *B. thuringiensis* subsp. *israelensis* and becomes the membrane-bound receptor) and Cry1A (the lepidopteran-specific insecticidal toxin) proteins.

In one series of experiments, researchers genetically engineered two different *P. fluorescens* strains that, when fed to *E. saccharina* larvae, acted synergistically in limiting the proliferation of the insect. One *P. fluorescens* strain was engineered to express the *cry1Ac7* gene under the transcriptional control of the *tac* promoter, with the entire construct integrated into the host chromosomal DNA. The other *P. fluorescens* strain was engineered to

TABLE 16.5 Synergistic effects of *P. fluorescens* expressing Cry1Ac7 toxin and *P. fluorescens* expressing chitinase on sugarcane borer larvae

Concentration (mg/g of diet)		Insect mortality (%)	
Toxin-producing strain	Chitinase-producing strain	Day 2	Day 5
0	0	5.5	7.6
0.3	0	12.5	33.8
3.0	0	30.8	42.7
0	0.3	8.2	20
0	30.0	21.8	42.7
0.3	0.3	39.3	55.7
0.3	30.0	37.5	68.3

Adapted from Downing et al., *Appl. Environ. Microbiol.* **66:**2804–2810, 2000.
Day 2 and Day 5 indicate the number of days after the treatment was started.

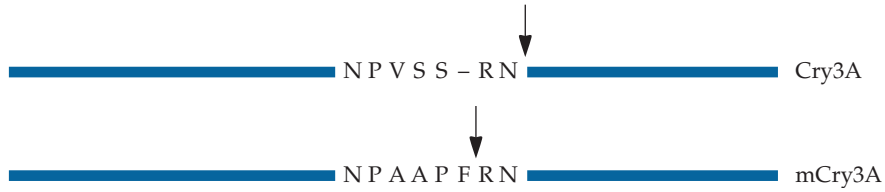


FIGURE 16.15 Location of α-chymotrypsin recognition and cut site within domain I of Cry3A and mCry3A. The letters represent the amino acid residues located between α-helices 3 and 4.

express a chitinase gene, originally isolated from the bacterium *Serratia marcescens*, also under the control of the *tac* promoter and integrated into the host chromosomal DNA. The chitinase is believed to cause perforations in the chitin-containing peritrophic membrane of the insect larvae, thereby lysing the membrane and killing the larvae, as well as increasing the accessibility of the midgut membranes to the *B. thuringiensis* insecticidal toxin. In laboratory tests, each of these *P. fluorescens* strains separately was toxic to *E. saccharina* larvae (Table 16.5). Moreover, when the two *P. fluorescens* strains were used together, there was significant synergism between the treatments so that low levels of both strains yielded a high level of insect mortality. Ideally a bacterial endophyte that can colonize the interior surfaces of the plant, rather than a surface-colonizing bacterial strain, would be preferred as a host strain, and of course, the efficacy of any construct must be demonstrated in the field, as well as in the laboratory.

Other workers have reported that the expression of a chitinase gene in a mosquitocidal strain of *B. sphaericus* (which is similar to *B. thuringiensis* subsp. *israelensis*) yielded a recombinant strain that was ~4,300 times more toxic than the wild type against a strain of the mosquito *Culex quinquefasciatus* that is considered to be resistant to the wild-type strain. The higher toxicity of the chitinase-expressing strain is thought to reflect the fact that chitinase digestion facilitates the interaction between the insecticidal toxin and its target cells.

While the Cry3A protein is an effective insecticide against the Colorado potato beetle *(Leptinotarsa decemlineata)*, it shows very little activity against the western corn rootworm *(Diabrotica undecempunctata howardi)*. Researchers speculated that the low level of activity against the western corn rootworm might reflect the fact that in this insect, proteolytic cleavage of the precursor form of the toxin is not sufficient for biological activity. That is, processing of Cry3A by the protease chymotrypsin might be necessary to increase the solubility and functional binding of the insecticide to the insect brush border membrane. It was found that the normal Cry3A chymotrypsin cleavage site (Fig. 16.15) was not efficiently cleaved in vitro. However, when a new enzyme recognition site for chymotrypsin was introduced into Cry3A (near the existing site), cleavage of the modified protein (mCry3A) by chymotrypsin increased substantially, and the protein solubility and insecticidal activity of the protein against the western corn rootworm, both in vitro and in vivo, also increased. Before a strain of *B. thuringiensis* that carries a gene for this modified Cry3A can be used in the environment, it will be necessary to elaborate its complete insect specificity and to ascertain (first in the laboratory) that this small modification in the structure of

TABLE 16.6 Some insect pests that are currently controlled with baculoviruses

Pest	Common name	Crop
Anticarsia gemmatalis	Velvetbean caterpillar	Soybean
Chrysomela scripta	Cottonwood leaf beetle	Trees
Cydia pomonella	Codling moth	Apple, walnut
Heliothis sp.	Cotton bollworm	Cotton, sorghum
Lymantria dispar	Gypsy moth	Deciduous trees
Mamestra brassicae	Cabbage moth	Vegetables
Neodiprion sertifer	European pine sawfly	Pine
Oryctes rhinoceros	Rhinoceros beetle	Coconut
Spodoptera exigua	Beet armyworm	Vegetables, flowers
Spodoptera littoralis	Egyptian cotton leaf worm	Cotton
Trichoplusia ni	Cabbage looper	Brassicas

Cry3A has not inadvertently generated any toxic activities against humans or other animals.

Activated Cry toxins bind to specific proteins (cadherins) on the surfaces of the microvilli of the insect midgut epithelial cells. Binding of toxin monomers to cadherins, which are transmembrane glycoproteins containing 12 cadherin repeating domains and one membrane-proximal extracellular domain (Fig. 16.16), facilitates the development of a multimeric form of the toxin monomers and formation of a pore in the membrane. Loss of cadherin or mutation of cadherin genes is generally associated with resistance to *B. thuringiensis*. A fragment of a cadherin protein containing the 12 repeating units and the membrane-proximal extracellular domain was mixed with Cry1A and fed to insect larvae. It was expected that the cadherin protein fragment would block the binding of the Cry1A protein to the midgut epithelial cells. Instead, its addition dramatically enhanced the Cry1A-induced insect mortality. The cadherin peptide fragment may first bind to microvilli and then attract Cry1A molecules, thereby increasing the probability of the toxin interacting with the bona fide receptor. It is thought that this approach, that is, the simultaneous application of Cry proteins and a peptide containing a portion of the receptor protein, will overcome or significantly delay the development of insect resistance by increasing Cry protein insect toxicity.

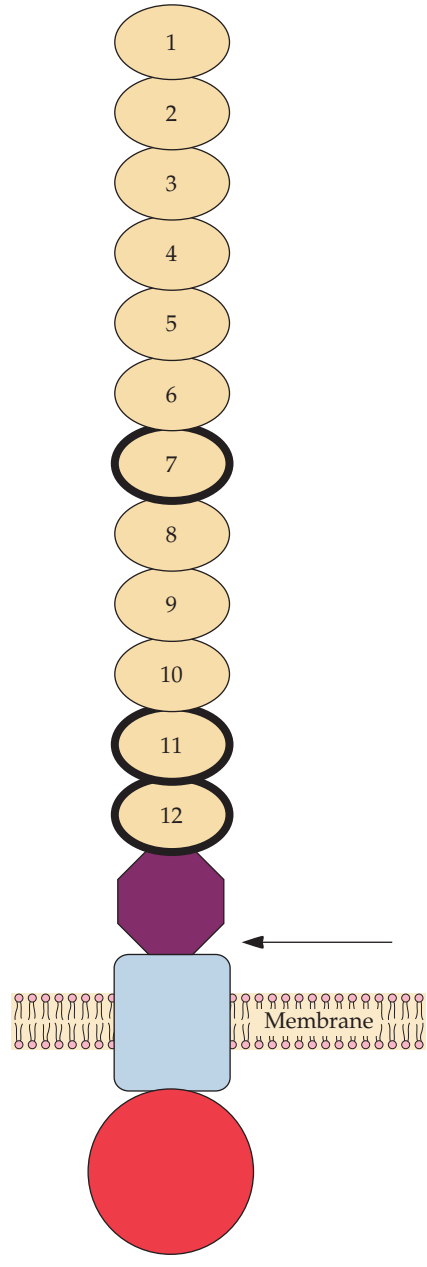


FIGURE 16.16 Schematic representation of a cadherin protein molecule embedded in a midgut epithelial membrane. Twelve cadherin domains are labeled. Domains 7, 11, and 12, which are highlighted, have been implicated as binding sites for Cry proteins. The purple region represents the membrane-proximal extracellular domain; the transmembrane domain is shown in light blue; and the cytoplasmic domain is shown in red. The arrow indicates the junction between the membrane-proximal extracellular domain and the transmembrane domain. The cadherin peptide analogue includes the 12 cadherin domains and the membrane-proximal extracellular domain.

Baculoviruses as Biocontrol Agents

Mode of Action

Baculoviruses are rod-shaped double-stranded DNA viruses that can infect and kill a large number of different invertebrate organisms. Subgroups of this viral family are pathogenic to several orders of insects, including the Lepidoptera, Hymenoptera, Diptera, Neuroptera, Trichoptera, Coleoptera, and Homoptera. In nature, some of these baculoviruses are important for the control of certain pest insects, and several have been registered for use as biological insecticides. Baculoviruses were used in North America against forest pests, such as the spruce sawfly *(Neodiprion sertifer)*, starting in the 1930s and ending with the advent of chemical pesticides in the 1960s. Baculoviruses continue to be used on a limited basis—approximately 0.1% of the money spent on pest control is directed toward baculoviruses—mostly by the forestry industry in an effort to control the gypsy moth, *L. dispar* (Table 16.6). It has been estimated that the costs associated with the development, production, and use of baculoviruses in developing countries

are much less than in industrialized countries. This is due to lower material and labor costs in developing countries, as well as to the smaller size of farms in some countries, lower levels of agricultural mechanization, and less expensive registration procedures. Nevertheless, even in the more developed countries of the world, the costs of the development and registration of a naturally occurring baculovirus are much less than the costs for a chemical insecticide. However, because baculoviruses have a high degree of host specificity, it is necessary to develop a large number of different baculoviruses to deal with different insect pests, whereas a single chemical pesticide might control all of them.

The vast majority of baculoviruses used as biological control agents are members of the genus *Nucleopolyhedrovirus,* and all subsequent discussion of baculoviruses refers to viruses from this genus. A baculovirus particle consists of a cylindrical nucleocapsid that surrounds the viral DNA. Often, in the nucleus of an infected cell, baculovirus particles are embedded in a crystalline protein matrix called an occlusion body. The occlusion body, or polyhedron, is largely composed of the protein polyhedrin. When an infected insect dies, millions of polyhedra are released. Upon ingestion by an insect, the polyhedra move to the midgut, where the alkaline environment facilitates the dissolution of the polyhedrin protein coat, releasing infectious nucleocapsids. The nucleocapsids are taken up by the insect midgut cells and then migrate through the cytoplasm to the nucleus, where the nucleocapsid is removed. After viral replication, which takes place in the nucleus, and nucleocapsid assembly, some nucleocapsids are released by budding through the plasma membranes of infected cells into the circulatory system of the insect. Consequently, the infection spreads to other cells throughout the insect. It usually takes about 10 rounds of viral replication, or about 5 to 9 days, for the insect to die. At that stage, about 25% of the dry weight of the insect consists of polyhedra.

A positive feature of using baculoviruses as biocontrol agents is that they generally have limited host ranges and do not affect nontarget organisms. However, this means that any particular baculovirus can be used to control only a limited number of insect pests. Since baculoviruses coevolved with their insect hosts over thousands of years, they are well adapted to avoid the insect's defense mechanisms, and resistance to these viruses develops only rarely, and much less frequently than resistance to *B. thuringiensis*.

Control of the European spruce sawfly *(Gilipinia hercyniae)* population in eastern Canada is the best example of insect control by a baculovirus. European sawfly populations were reduced to below economic threshold levels by 1943 and remain under control today. Ironically, the reason why some baculoviruses are not used commercially is related to the effectiveness of the virus. If a virus is effective at preventing proliferation of a particular insect species, the virus has to be applied only once every year or so, making it difficult for the industry to justify the high registration costs. Farmers and growers prefer to use a single insecticidal agent that can control many different insect pests rather than a number of different insecticides, so if baculoviruses are to be used more extensively, their limited host range needs to be expanded.

It has been known for some time that when insect cells are infected with two different strains of baculovirus at the same time, new variant viruses with slightly different specificities can form after the two starting viruses have replicated. These new viruses are the product of homologous

recombination between the two starting viruses. Detailed analysis of this phenomenon revealed that a region of DNA that is only 79 base pairs (bp) long and located within the p143 helicase gene was sufficient to permit homologous recombination between different baculoviruses. More importantly, this 79-bp DNA segment may be responsible for the host ranges of different baculoviruses. Therefore, alteration of some of the nucleotides within this 79-bp DNA segment may allow researchers to generate baculoviruses with modified (expanded) insect specificities.

Genetic Engineering for Improved Biocontrol

Baculoviruses are relatively slow in killing target insects. Depending on conditions, it can take from a few days to several weeks before the viral infection leads to the host's death. To remedy this ineffectiveness, several attempts have been made to enhance the virulence of baculoviruses by introducing foreign genes that either severely impair or kill the targeted insect species (Table 16.7). One approach has been to use a gene that disrupts the normal life cycle of the insect when it is expressed within the host insect cells.

During insect development, a reduction in the level of juvenile hormone in larvae initiates metamorphosis into pupae and leads to a cessation of larval feeding. The reduction in the juvenile hormone level is due to an increase in the amount of juvenile hormone esterase, an enzyme that converts the biologically active methyl ester form of juvenile hormone into an inactive acid form. Inhibition of juvenile hormone esterase activity leads to an in vivo accumulation of active juvenile hormone, so the larvae remain in the feeding stage longer, continue to grow, and eventually become giant larvae. Therefore, researchers reasoned that an experimentally induced increase in the supply of juvenile hormone esterase should lower the endogenous level of active juvenile hormone and cause a premature cessation of feeding. Basically, their premise was that shortening the duration of larval feeding would curtail the extent of crop damage.

To test this idea, the investigators first had to clone and express the gene for juvenile hormone esterase. This task was achieved by purifying the enzyme from the tobacco budworm *(Heliothis virescens)*, determining its amino acid sequence, synthesizing a DNA oligomer that corresponds to a portion of the esterase amino acid sequence, and then using this oligomer as a hybridization probe. The coding sequence for juvenile hormone esterase was isolated from an *H. virescens* complementary DNA (cDNA) library and inserted into the genome of a baculovirus under the control of baculovirus transcription signals. When the cabbage looper *(Trichoplusia ni)*

TABLE 16.7 Some genes that have been introduced into the baculovirus genome to increase insecticidal activity

Gene	Effect on host insect of introduced gene
Diuretic hormone	Reduced hemolymph volume
Juvenile hormone esterase	Feeding cessation
B. thuringiensis toxin	Feeding cessation
Scorpion toxin	Paralysis
Mite toxin	Paralysis
Wasp toxin	Premature melanization, low weight gain

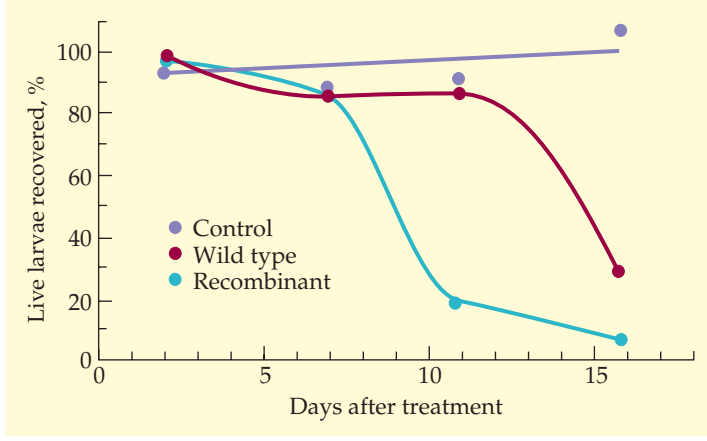


FIGURE 16.17 Survival of *T. ni* larvae after treatment of cabbage leaves with either wild-type or recombinant baculovirus expressing a scorpion neurotoxin gene. The control cabbage plants were treated only with insect larvae. The lower the percentage of live larvae recovered, the greater the killing of the larvae and the more effective the treatment. Plants were treated with baculovirus only once at the start of the experiment.

at the first larval instar stage was treated with this genetically modified baculovirus, the amount of juvenile hormone in the insect was reduced by the cloned juvenile hormone esterase, and larval feeding and growth were dramatically curtailed relative to feeding and growth by the control larvae that were treated with native baculovirus.

The usefulness of this approach for enhancing baculoviruses as general biocontrol agents has been questioned, because the reduction in larval feeding that is attributable to the effect of juvenile hormone esterase is confined to the first larval instar. Other stages of development are much less sensitive to this treatment. A baculovirus engineered to express juvenile hormone esterase would have to be applied when the majority of the target insect population was in its first larval instar stage, which, under natural conditions, is difficult to achieve.

Another approach for enhancing the effectiveness of baculoviruses as a pesticide is to incorporate into the viral genome an insect-specific toxin gene that, when expressed during the viral infection cycle, will yield a potent insect toxin. The gene that encodes the insect-specific neurotoxin produced by the North African fat-tailed scorpion (*Androctonus australis* Hector) was cloned into a baculovirus strain, and the genetically engineered virus was tested as a biological insecticide. This neurotoxin, which does not have any effect on mice, disrupts the flow of sodium ions in the neurons of targeted insects and eventually leads to paralysis and death. Laboratory-raised insects that were infected with a baculovirus carrying the scorpion neurotoxin gene caused 50% less damage to the leaves of test plants than did insects that had been treated with wild-type baculovirus.

When the cDNA for the toxin from the Israeli yellow scorpion *(Leiurus quinquestriatus hebraeus)* was cloned and introduced into the baculovirus *Autographa californica* multiple nuclear polyhedrosis virus, the time that it took to kill 50% of the insect larvae that were tested was reduced from 120 to 78 h. Moreover, 120 h after infection, the insect larvae treated with recombinant virus gained only one-third as much weight as larvae treated

with wild-type virus. Thus, this engineered baculovirus not only hastened the demise of the infected insect larvae, it also significantly decreased the ability of the insects to damage plants. Subsequent experiments have taken this system a step further. Researchers have studied the effects of different scorpion toxins, either separately or two at a time. They assessed whether either a combination of excitatory and depressant or alpha and depressant scorpion toxins would improve the efficacy of *A. californica* nuclear polyhedrosis virus, over a virus expressing only a single toxin, toward three different insect larvae. The best result was achieved by combined expression of the excitatory toxin and the depressant toxin. Under these conditions, the "effective time to paralysis" of *H. virescens* neonates was reduced to slightly less than 47 h. Additional improvements to this system have come from placing the scorpion toxin(s) under the transcriptional control of the *p-PCm* promoter, which contains the human cytomegalovirus minimal (*CMVm*) promoter ligated in *cis* with the polyhedrin upstream (*pu*) sequence. This results in a high level of expression of foreign genes at an early infection stage of the baculovirus.

Recently, a genetically engineered *A. californica* nuclear polyhedrosis virus that expresses the insect-specific neurotoxin from *Androctonus australis* was tested under field conditions. Interestingly, the modified baculovirus was even more effective in the field than in the laboratory studies that demonstrated a 25 to 50% reduction in the time it took to kill the insect pest *T. ni* (Fig. 16.17). In the field, the genetically engineered baculovirus killed the insect pests faster, decreased the damage to cabbage plants, and reduced the secondary cycle of infection (infections caused by the next generation of the virus) compared to the wild-type virus.

No matter how effective a particular genetically engineered baculovirus may be in small-scale experiments, a major hurdle to more widespread use is the difficulty and cost of propagating such viruses. Baculoviruses are obligate parasites; therefore, they must be grown either in living whole organisms or in insect cell culture. In more developed countries, the cost of baculovirus preparations, whether or not the virus has been genetically engineered, is not currently competitive with that of chemical insecticides. However, biological insecticides may become more appealing when the adverse environmental impact of chemical insecticides is factored into the cost-benefit analysis.

SUMMARY

Microbial insecticides are currently being developed as environmentally friendly biological substitutes for chemical pesticides. A number of subspecies of the bacterium *B. thuringiensis* produce a protoxin as part of a parasporal crystal that, after ingestion, kills specific insects. The transition from insecticidal protoxin to toxin occurs in the gut of the target insect and is mediated by the pH and digestive proteases in the gut. The death of the insect is the consequence of the formation of membrane channels in the gut cells, which allow ATP to escape and in turn lead to decreased cellular metabolism, cessation of feeding, dehydration, and eventually death. The *B. thuringiensis* toxins are highly specific for a limited number of insect species, nontoxic to nontarget species, and biodegradable. Consequently, they are unlikely to cause significant biological selection for resistant forms under normal conditions. These attributes make these biological insecticides effective agents for controlling insect damage to certain crops and preventing the proliferation of insects that act as vectors of human diseases.

The genes *(cry)* for various *B. thuringiensis* toxins have been cloned and characterized. By expressing a *B. thuringiensis cry* gene in a nonsporulating *Bacillus* strain, production of the insecticidal protein was achieved during vegetative growth, bypassing the need for parasporal crystal formation.

To expand the specificity of a *B. thuringiensis* toxin to other pest insects, toxin genes from different subspecies were cloned into plasmids and introduced into another *B. thuringiensis* strain, either on a broad-host-range plasmid or by integration into the chromosomal DNA of the host cell. In addition to expressing the toxicity of the original strain, the bacteria with

two different toxin genes sometimes showed an effect against a nontarget insect pest. In one study, it was found that modification of domain II of the Cry protein is an effective means of increasing its toxicity to particular insects. Similarly, a fusion protein consisting of two toxin domains from different *B. thuringiensis* toxin genes was constructed by genetic manipulation, and the fusion protein retained both toxic activities. In another study, the receptor-binding domain of one insecticidal toxin was combined with the toxin domain of another. It is thought that insect resistance is less likely to develop when such hybrid toxins are used. In addition, the simultaneous application of Cry proteins and a peptide containing a portion of the host Cry receptor protein increases Cry protein insect toxicity, another strategy that can be used to overcome or significantly delay the development of insect resistance to Cry proteins. A further strategy that can both improve biocontrol activity and serve to limit the development of *B. thuringiensis*-resistant insects is the use of *B. thuringiensis* toxins together with other insecticidal proteins, such as chitinase or the *B. thuringiensis* subsp. *israelensis* Cyt1A protein.

To ensure that *B. thuringiensis* spraying for the control of mosquitoes is effective, the *B. thuringiensis* toxin genes have been cloned into various microorganisms that live near the surfaces of ponds and are eaten by mosquito larvae. This strategy appears to be an effective means of delivering the *B. thuringiensis* toxin to the targeted insect. Also, rhizosphere bacteria that have been engineered with *B. thuringiensis* toxin genes lessen the damage caused by insects that attack the roots of plants.

Baculoviruses are pathogenic to many different species of insects, but each strain of baculovirus is specific to a small number of insect species. Although baculoviruses kill their host organisms, the process is usually considered to be too slow to be effective for controlling insects that attack crop plants. However, when certain genes are cloned into different strains of baculovirus, the virus can act as a delivery system for a gene that produces an insecticidal protein during the viral life cycle. Several tests of this strategy have been successful in laboratory trials. In addition, when a gene for a neurotoxin that kills insects was cloned into a baculovirus, the construct was effective in field trials.

REFERENCES

Abdullah, M. A. F., O. Alzate, M. Mohammad, R. J. McNall, M. J. Adang, and D. H. Dean. 2003. Introduction of *Culex* toxicity into *Bacillus thuringiensis* Cry4Ba by protein engineering. *Appl. Environ. Microbiol.* **69:**5343–5353.

Abdullah, M. A. F., and D. H. Dean. 2004. Enhancement of Cry19Aa mosquitocidal activity against *Aedes aegypti* by mutations in the putative loop regions of domain II. *Appl. Environ. Microbiol.* **70:**3769–3771.

Ananda Kumar, P., R. P. Sharma, and V. S. Malik. 1996. The insecticidal proteins of *Bacillus thuringiensis. Adv. Appl. Microbiol.* **42:**1–43.

Aronson, A. I., and Y. Shai. 2001. Why *Bacillus thuringiensis* insecticidal toxins are so effective: unique features of their mode of action. *FEMS Microbiol. Lett.* **195:**1–8.

Audtho, M., A. P. Valatis, O. Alzate, and D. H. Dean. 1999. Production of chymotrypsin-resistant *Bacillus thuringiensis* Cry2Aa1 δ-endotoxin by protein engineering. *Appl. Environ. Microbiol.* **65:**4601–4605.

Baum, J. A., and T. Malvar. 1995. Regulation of insecticidal crystal protein production in *Bacillus thuringiensis*. *Mol. Microbiol.* **18:**1–12.

Bosch, D., B. Schipper, H. van der Kleij, R. A. de Maagd, and W. J. Stiekema. 1994. Recombinant *Bacillus thuringiensis* crystal proteins with new properties: possibilities for resistance management. *Bio/Technology* **12:**915–918.

Bravo, A., S. S. Gill, and M. Soberón. 2007. Mode of action of *Bacillus thuringiensis* Cry and Cyt toxins and their potential for insect control. *Toxicon* **49:**423–435.

Broderick, N. A., K. F. Raffa, and J. Handelsman. 2006. Midgut bacteria required for *Bacillus thuringiensis* insecticidal activity. *Proc. Natl. Acad. Sci. USA* **103:**15196–15199.

Cai, Y., J. Yan, X. Hu, B. Han, and Z. Yuan. 2007. Improving the insecticidal activity against resistant *Culex quinquefasciatus* mosquitoes by expression of chitinase gene *chiAC* of *Bacillus sphaericus*. *Appl. Environ. Microbiol.* **73:**7744–7746.

Calogero, S., A. M. Albertini, C. Fogher, R. Marzari, and A. Galizzi. 1989. Expression of a cloned *Bacillus thuringiensis* delta-endotoxin gene in *Bacillus subtilis*. *Appl. Environ. Microbiol.* **55:**446–453.

Caramori, T., A. M. Albertini, and A. Galizzi. 1991. In vivo generation of hybrids between two *Bacillus thuringiensis* insect-toxin-encoding genes. *Gene* **98:**37–44.

Cerda, H., and M. G. Paoletti. 2004. Genetic engineering with *Bacillus thuringiensis* and conventional approaches for insect resistance in crops. *Crit. Rev. Plant Sci.* **23:**317–323.

Chejanovsky, N., N. Zilberberg, H. Rivkin, E. Zlotkin, and M. Guervitz. 1995. Functional expression of an alpha anti-insect scorpion neurotoxin in insect cells and lepidopterous larvae. *FEBS Lett.* **376:**181–184.

Chen, J., G. Hua, J. L. Jurat-Fuentes, M. A. Abdullah, and M. J. Adang. 2007. Synergism of *Bacillus thuringiensis* toxins by a fragment of a toxin-binding cadherin. *Proc. Natl. Acad. Sci. USA* **104:**13901–13906.

Cory, J. S., M. L. Hirst, T. Williams, R. S. Hails, D. Goulson, B. M. Green, T. M. Carty, R. D. Possee, P. J. Cayley, and D. H. L. Bishop. 1994. Field trial of a genetically improved baculovirus insecticide. *Nature* **370:**138–140.

Crickmore, N., C. Nicholls, D. J. Earp, T. C. Hodgman, and D. J. Ellar. 1990. The construction of *Bacillus thuringiensis* strains expressing novel entomocidal δ-endotoxin combinations. *Biochem. J.* **270:**133–136.

Crickmore, N., D. R. Zeigler, J. Fewitelson, E. Schnepf, J. Van Rie, D. Lereclus, J. Baum, and D. H. Dean. 1998. Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal crystal protein. *Microbiol. Mol. Biol. Rev.* **62:**807–813.

Croizier, G., L. Croizier, O. Argaud, and D. Poudevigne. 1994. Extension of *Autographa californica* nuclear polyhedrosis virus host range by interspecific replacement of a short DNA sequence in the p143 helicase gene. *Proc. Natl. Acad. Sci. USA* **91:**48–52.

De Maagd, R. A., A. Bravo, and N. Crickmore. 2001. How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect. *Trends Genet*. **17:**193–199.

Downing, K. J., G. Leslie, and J. A. Thomson. 2000. Biocontrol of the sugarcane borer *Eldana saccharina* by expression of the *Bacillus thuringiensis cry1Ac7* and *Serratia marcescens chiA* genes in sugarcane-associated bacteria. *Appl. Environ. Microbiol*. **66:**2804–2810.

Ferré, J., M. D. Real, J. Van Rie, S. Jansens, and M. Peferoen. 1991. Resistance to the*Bacillus thuringiensis* bioinsecticide in a field population of *Plutella xylostella* is due to a change in a midgut membrane receptor. *Proc. Natl. Acad. Sci. USA* **88:**5119–5123.

Ge, A. Z., N. I. Shivarova, and D. H. Dean. 1989. Location of the *Bombyx mori* specificity domain on a *Bacillus thuringiensis* δ-endotoxin protein. *Proc. Natl. Acad. Sci. USA* **86:**4037–4041.

Gelernter, W., and G. E. Schwab. 1993. Transgenic bacteria, viruses, algae and other microorganisms as *Bacillus thuringiensis* toxin delivery systems, p. 89–104. *In* P. F. Entwistle, J. S. Cory, M. J. Bailey, and S. Higgs (ed.), *Bacillus thuringiensis, an Environmental Biopesticide: Theory and Practice*. John Wiley & Sons, Chichester, United Kingdom.

Hammock, B. D., B. C. Bonning, R. D. Possee, T. N. Hanzlik, and S. Maeda. 1990. Expression and effects of the juvenile hormone esterase in a baculovirus vector. *Nature* **344:**458–461.

Held, G. A., L. A. Bulla, E. Ferrari, J. Hoch, A. I. Aronson, and S. A. Minnich. 1982. Cloning and localization of the lepidopteran protoxin gene of *Bacillus thuringiensis* subsp. *kurstaki. Proc. Natl. Acad. Sci. USA* **79:**6065–6069.

Herrera, G., S. J. Snyman, and J. A. Thomson. 1994. Construction of a bioinsecticidal strain of *Pseudomonas fluorescens* active against the sugarcane borer, *Eldana saccharina*. *Appl. Environ. Microbiol.* **60:**682–690.

Kalman, S., K. L. Kiehne, N. Cooper, M. S. Reynoso, and T. Yamamoto. 1995. Enhanced production of insecticidal proteins in *Bacillus thuringiensis* strains carrying an additional crystal protein gene in their chromosomes. *Appl. Environ. Microbiol*. **61:**3063–3068.

Lampel, J. S., G. L. Canter, M. B. Dimock, J. L. Kelly, J. J. Anderson, B. B. Uratani, J. S. Foulke, Jr., and J. T. Turner. 1994. Integrative cloning, expression, and stability of the *cryIA(c)* gene from *Bacillus thuringiensis* subsp. *kurstaki* in a recombinant strain of *Clavibacter xyli* subsp. *cynodontis*. *Appl. Environ. Microbiol*. **60:**501–508.

Lereclus, D., H. Agaisse, M. Gominet, and J. Chaufaux. 1995. Overproduction of encapsulated insecticidal crystal proteins in a *Bacillus thuringiensis spo0A* mutant. *Bio/Technology* **13:**67–71.

Lereclus, D., A. Delécuse, and M. M. Lecadet. 1993. Diversity of *Bacillus thuringiensis* toxins and genes, p. 37–69. *In* P. F. Entwistle, J. S. Cory, M. J. Bailey, and S. Higgs (ed.), *Bacillus thuringiensis, an Environmental Biopesticide: Theory and Practice.* John Wiley & Sons, Chichester, United Kingdom.

Liu, J. T., M. J. Sui, D. D. Ji, I. H. Wu, C. C. Chou, and C. C. Chen. 1993. Protection from ultraviolet radiation by melanin of mosquitocidal activity of *Bacillus thuringiensis* var. *israelensis*. *J. Invertebr. Pathol.* **62:**131–136.

Liu, J. W., W. H. Yap, T. Thanabalu, and A. G. Porter. 1996. Efficient synthesis of mosquitocidal toxins in *Asticcacaulis excentricus* demonstrates potential of gram-negative bacteria in mosquito control. *Nat. Biotechnol*. **14:**343–347.

Maeda, S. 1995. Further development of recombinant baculovirus insecticides. *Curr. Opin. Biotechnol*. **6:**313–319.

McCutchen, B. F., P. V. Choudary, R. Crenshaw, D. Maddox, S. G. Kamita, N. Palekar, S. Volrath, E. Fowler, B. D. Hammock, and S. Maeda. 1991. Development of a recombinant baculovirus expressing an insect-selective neurotoxin: potential for pest control. *Bio/Technology* **9:**848–852.

Mettus, A. M., and A. Macaluso. 1990. Expression of *Bacillus thuringiensis* δ-endotoxin genes during vegetative growth. *Appl. Environ. Microbiol*. **56:**1128–1134.

Murphy, R. C., and E. S. Stevens. 1991. Cloning and expression of the *cryIVD* gene of *Bacillus thuringiensis* subsp. *israelensis* in the cyanobacterium *Agmenellum quadruplicatum* PR-6 and its resulting larvicidal activity. *Appl. Environ. Microbiol*. **58:**1650–1655.

Obukowicz, M. G., F. J. Perlak, K. Kusano-Kretzmer, E. J. Mayer, S. L. Bolten, and L. S. Watrud. 1986. Tn*5*-mediated integration of the delta-endotoxin gene from *Bacillus thuringiensis* into the chromosome of root-colonizing pseudomonads. *J. Bacteriol*. **168:**982–989.

Obukowicz, M. G., F. J. Perlak, K. Kusano-Kretzmer, E. J. Mayer, and L. S. Watrud. 1986. Integration of the delta-endotoxin gene of *Bacillus thuringiensis* into the chromosome of root-colonizing strains of pseudomonads using Tn*5*. *Gene* **45:**327–331.

Perez, C., L. F. Fernandez, J. Sun, J. L. Folch, S. S. Gill, M. Soberón, and A. Bravo. 2005. *Bacillus thuringiensis* subsp. *israelensis* Cyt1Aa synergizes Cry11Aa toxin by functioning as a membrane-bound receptor. *Proc. Natl. Acad. Sci. USA* **102:**18303–18308.

Priest, F. G. 1992. Biological control of mosquitoes and other biting flies by *Bacillus sphaericus* and *Bacillus thuringiensis*. *J. Appl. Bacteriol*. **72:**357–369.

Regev, A., H. Rivkin, B. Inceoglu, E. Gershburg, B. D. Hammock, M. Gurevitz, and N. Chejanovsky. 2003. Further enhancement of baculovirus insecticidal efficacy with scorpion toxins that interact cooperatively. *FEBS Lett.* **537:**106–110.

Sanchis, V., M. Gohar, J. Chaufaux, O. Arantes, A. Meier, H. Agaisse, J. Cayley, and D. Lereclus. 1999. Development and field performance of a broad-spectrum nonviable asporogenic recombinant strain of *Bacillus thuringiensis* with greater potency and UV resistance. *Appl. Environ. Microbiol*. **65:**4032–4039.

Sayyed, A. H., N. Crickmore, and D. J. Wright. 2001. Cyt1Aa from *Bacillus thuringiensis* subsp. *israelensis* is toxic to the diamondback moth, *Plutella*

xylostella, and synergizes the activity of Cry1Ac towards a resistant strain. *Appl. Environ. Microbiol*. **67:**5859–5861.

Schnepf, E., N. Crickmore, J. Van Rie, D. Lereclus, J. Baum, J. Feitelson, D. R. Zeiger, and D. H. Dean. 1998. *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiol. Mol. Biol. Rev*. **62:**775–806.

Schnepf, H. E., and H. R. Whiteley. 1981. Cloning and expression of the *Bacillus thuringiensis* crystal protein gene in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **78:**2893–2897.

Stewart, L. M. D., M. Hirst, M. L. Ferber, A. T. Merryweather, P. J. Cayley, and R. D. Possee. 1991. Construction of an improved baculovirus insecticide containing an insect-specific toxin gene. *Nature* **352:**85–88.

Thanabalu, T., J. Hindley, S. Brenner, C. Oei, and C. Berry. 1992. Expression of the mosquitocidal toxins of *Bacillus sphaericus* and *Bacillus thuringiensis* subsp. *israelensis* by recombinant *Caulobacter crescentus*, a vehicle for biological control of aquatic insect larvae. *Appl. Environ. Microbiol*. **58:**905–910.

Thiery, I., L. Nicholas, R. Rippka, and N. Tandeau de Marsac. 1991. Selection of cyanobacteria isolated from mosquito breeding sites as a potential food source for mosquito larvae. *Appl. Environ. Microbiol.* **57:**1354–1359.

Tomalski, M. D., and L. K. Miller. 1991. Insect paralysis by baculovirus-mediated expression of a mite neurotoxin gene. *Nature* **352:**82–84.

Tuan, S.-J., R. F. Hou, C.-F. Lee, and Y.-C. Chao. 2007. High level production of polyhedra in a scorpion toxin containing recombinant baculovirus for better control of insect pests. *Bot. Stud.* **48:**273–281.

Walters, F. S., C. M. Stacy, M. K. Lee, N. Palekar, and J. S. Chen. 2008. An engineered chymotrypsin/cathepsin G site in domain I renders *Bacillus thuringiensis* Cry3A active against western corn rootworm larvae. *Appl. Environ. Microbiol.* **74:**367–374.

Whalon, M. E. and B. A. Wingerd. 2003. Bt: mode of action and use. *Arch. Insect Biochem. Physiol.* **54:**200–211.

Wirth, M. C., G. P. Georghiou, and B. A. Federici. 1997. CytA enables CryIV endotoxins of *Bacillus thuringiensis* to overcome high levels of CryIV resistance in the mosquito, *Culex quinquefasciatus*. *Proc. Natl. Acad. Sci. USA* **94:**10536–10540.

Wood, H. A., and R. R. Granados. 1991. Genetically engineered baculoviruses as agents for pest control. *Annu. Rev. Microbiol.* **45:**69–87.

Xu, Y., M. Nagai, M. Bagdasarian, T. W. Smith, and E. D. Walker. 2001. Expression of the p20 gene from *Bacillus thuringiensis* H-14 increases Cry11A toxin production and enhances mosquito-larvicidal activity in recombinant gram-negative bacteria. *Appl. Environ. Microbiol*. **67:**3010–3015.

Yamamoto, T., S. Kalman, G. Powell, N. Cooper, and D. Cerf. 1998. Environmental release of genetically engineered *Bacillus thuringiensis*. *Rev. Toxicol*. **2:**157–166.

Yap, W. H., T. Thanabalu, and A. G. Porter. 1994. Expression of mosquitocidal toxin genes in a gas-vacuolated strain of *Ancylobacter aquaticus*. *Appl. Environ. Microbiol*. **60:**4199–4202.

REVIEW QUESTIONS

1. What are the advantages of biological insecticides over chemical insecticides?

2. Draw a simple phylogenetic tree that shows the relationship among Cry1Aa, Cry1Ab, Cry1Ba, Cry1Bb, Cry2Aa, and Cry2Ab.

3. Why is the *B. thuringiensis* toxin not toxic to humans?

4. Outline a strategy that you would use to isolate an insecticidal protoxin gene from *B. thuringiensis* subsp. *israelensis*. How would you use this gene in a practical way?

5. How would you determine whether a particular insecticidal protoxin gene is present on a plasmid or part of the chromosome of a *B. thuringiensis* strain?

6. How would you use genetic engineering to improve the usefulness of a particular *B. thuringiensis* protoxin?

7. If a *cry1C–cry1Ab* fusion gene encoding an insecticidal protoxin consists of approximately 2,200 bp of DNA from the *cry1C* gene and 1,300 bp of DNA from the *cry1Ab* gene, what is the advantage of synthesizing this fusion protoxin compared with the *cry1C* protoxin?

8. How can insect gut enzymes be limited to processing the *B. thuringiensis* insecticidal protoxin to the active toxin without degrading the toxin?

9. How would you engineer a Cry protein to lessen or avoid the development of insect resistance to this toxin?

10. Why is the bacterium *A. excentricus* an attractive host organism for the expression of *B. thuringiensis* insecticidal toxin genes?

11. How can the species range of an insecticidal *B. thuringiensis* strain be extended?

12. What is a truncated *B. thuringiensis* insecticidal protoxin?

13. Why is it unlikely that insects will ever develop resistance to *B. thuringiensis* subsp. *israelensis* strains?

14. What are cadherins, and how can they be used to enhance the toxicity of a particular Cry protein?

15. How would you improve the insecticidal properties of baculoviruses?

16. How might it be possible to expand the range of insects that are infected by a particular baculovirus?

17

Large-Scale Production of Proteins from Recombinant Microorganisms

THE PRODUCTION OF COMMERCIAL PRODUCTS that are synthesized by genetically engineered microorganisms requires the partnership of two kinds of experts. Molecular biologists are responsible for isolating, characterizing, modifying, and creating effectively expressed genes in microorganisms that can be used for industrial production, and biochemical engineers ensure that the genetically engineered form of a microorganism can be grown in large quantities under conditions that give optimal product yields. In the early days of molecular biotechnology, biologists naively thought that scale-up was simply a matter of multiplication; i.e., they believed that whatever conditions were found to be effective on a small scale would be equally effective on a large scale and that to achieve this it was merely necessary to use a larger reaction vessel with a correspondingly larger volume of medium.

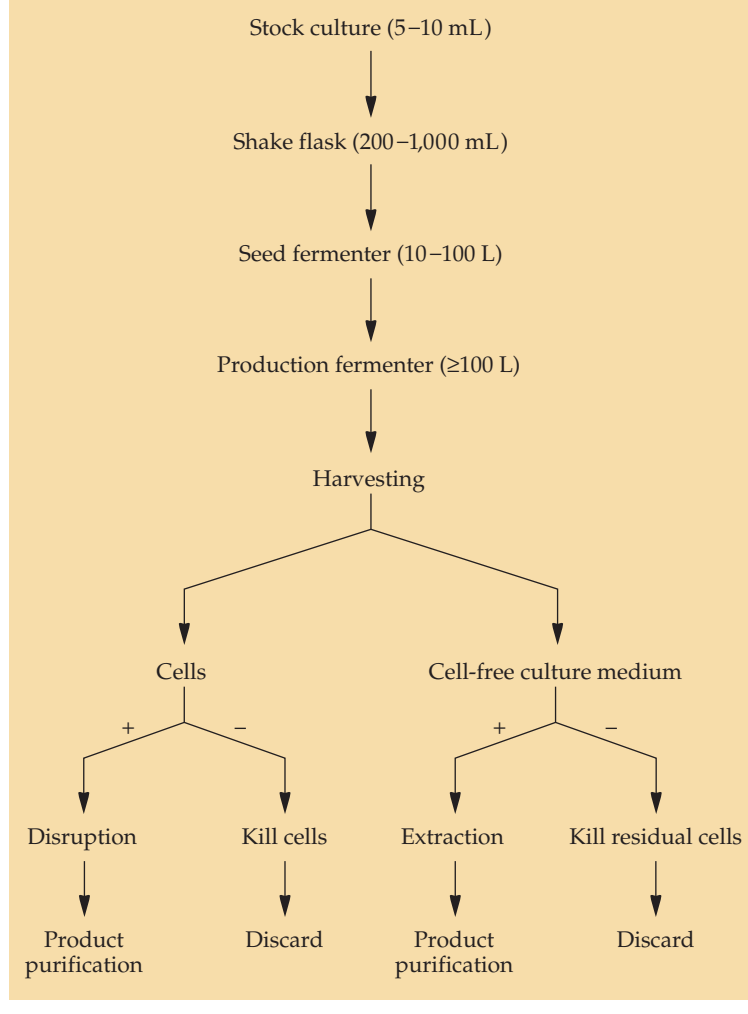
This simplistic view is far from reality. For example, good growth of aerobic microorganisms can usually be achieved in a standard 200-mL laboratory flask that is aerated with a mixer driven by a 300-watt motor. If the system were directly scaled up, a single 10,000-liter container would require a mixer with a 15-megawatt motor. Such a motor would be as large as a house, and the heat generated during stirring would boil the microorganisms. Although biochemical engineers may quibble about some aspects of this specific example, they all know that the industrial production of microorganisms is not merely a multiplication of bench scale conditions. For a start, increasing the size of the reaction vessel (bioreactor, or fermenter) is required for the large-scale growth of microorganisms, because it would be impractical to set up 50,000 individual culture flasks, each containing 200 mL, to obtain 10,000 liters of cell suspension.

A number of parameters must be precisely regulated to obtain maximum yields from either small (1- to 10-liter) or large (>1,000-liter) bioreactors. These parameters include the temperature, pH, rate and nature of mixing of the growing cells, and, with aerobic organisms, oxygen demand. Moreover, the optimal conditions generally change with each 10-fold increase in the volume of a bioreactor.

There are also other technical considerations. The design of the bioreactor is important. It should ensure adequate sterility and provide appropriate levels of containment of genetically engineered microorganisms. The reactor should also include probes that permit the accurate and continuous on-line monitoring of as many critical reaction parameters as possible so that adjustments can be made rapidly and easily throughout the course of the fermentation reaction (i.e., the growth of the microorganism). In addition, because sterilization may alter the composition of the medium (e.g., by destroying vitamins), it is important to ascertain that the medium composition is still optimal for maximal microbial growth following sterilization.

Generally, large-scale fermentation and product purification are stepwise processes (Fig. 17.1). A typical procedure begins with formulation and sterilization of the growth medium and sterilization of the fermentation equipment. The cells are grown first as a stock culture (5 to 10 mL), then in

FIGURE 17.1 Generalized scheme for a large-scale fermentation process. The commercial product is usually in either the cell or cell-free fraction, but not in both; consequently, one or the other of these fractions will be processed further (+) or discarded (–).



a shake flask (200 to 1,000 mL), and then in a seed fermenter (10 to 100 liters). Finally, the production fermenter (1,000 to 100,000 liters) is inoculated. After the fermentation step is completed, the cells are separated from the culture fluid by either centrifugation or filtration. If the product is intracellular, the cells are disrupted, the cell debris is removed, and the product is recovered from the debris-free fluid. If the product is extracellular, it is purified from the cell-free culture medium.

Principles of Microbial Growth

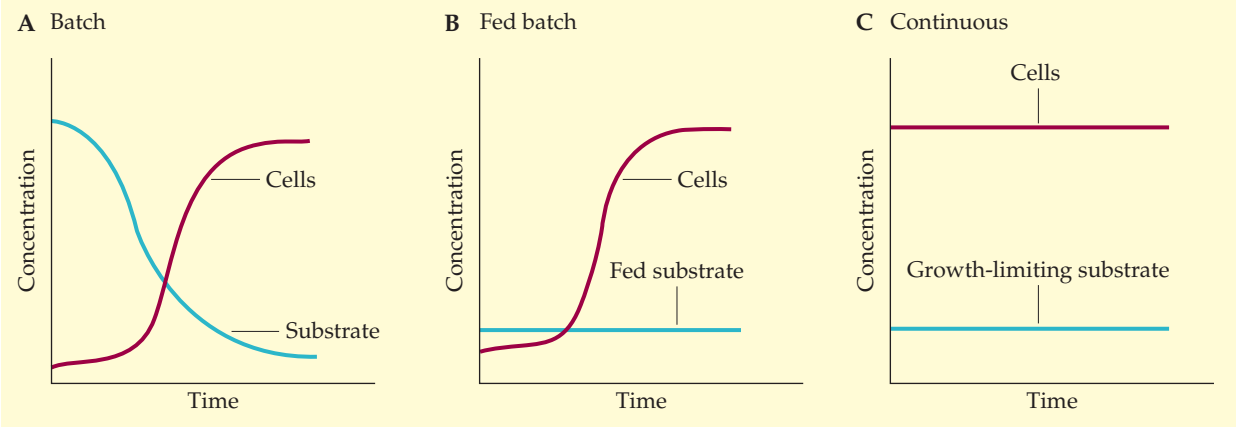
Microorganisms can be grown in batch, fed-batch, or continuous culture (Fig. 17.2). In batch fermentation, the sterile growth medium is inoculated with the appropriate microorganisms, and the fermentation proceeds without the addition of fresh growth medium. In fed-batch fermentation, nutrients are added incrementally at various times during the fermentation reaction; no growth medium is removed until the end of the process. In the continuous fermentation process, fresh growth medium is added continuously during fermentation, but there is also concomitant removal of an equal volume of spent medium containing suspended microorganisms. For each type of fermentation, oxygen (which is usually provided in the form of sterile air), an antifoaming agent, and, if required, acid or base are injected into the bioreactor as needed.

Batch Fermentation

During a batch fermentation, the composition of the culture medium, the concentration of microorganisms (biomass concentration), the internal chemical composition of the microorganisms, and the amount of either target protein or metabolite all change as a consequence of the state of cell growth, cellular metabolism, and availability of nutrients. Under these conditions, six typical phases of growth are usually observed: lag phase, acceleration phase, logarithmic (log) or exponential phase, deceleration phase, stationary phase, and death phase (Fig. 17.3).

Typically, there is no immediate increase in the numbers of cells after the inoculation into sterilized growth medium. This initial period is called

FIGURE 17.2 Schematic representation of the time course (progress curves) of cell concentration (mass) and substrate concentration in batch **(A)**, fed-batch **(B)**, and continuous **(C)** fermentations.



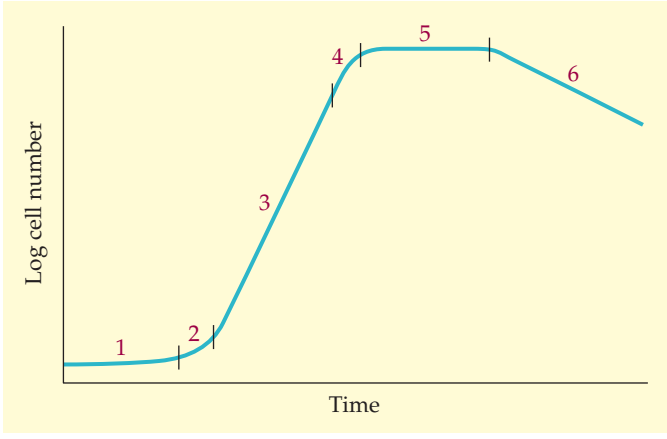


FIGURE 17.3 Pattern of microbial cell growth in a batch fermenter. The six phases of the growth cycle are (1) lag, (2) acceleration, (3) log (exponential), (4) deceleration, (5) stationary, and (6) death.

the lag phase. During the lag period, the microbial cells adapt to the new environmental conditions. The cells may have to adjust to a different pH or to a new level of available nutrients. As part of an adaptive response, previously unexpressed metabolic pathways may be induced. A lag phase generally occurs whenever the cells of the inoculum are derived from a culture that has stopped growing (i.e., has entered stationary phase) because of substrate limitation or product inhibition. These cells need time to reset their metabolic systems to adjust to the new medium. The length of the lag phase corresponds to how long the inoculated cells were in stationary phase and the extent to which the previous growth medium of the starting cells differed from the new, fresh culture medium. Conversely, when the inoculum is a cell culture from a growing cell population in log phase, a discernible lag phase may not occur and growth may begin immediately. Following the lag phase, the brief period when the rate of cell growth increases until log-phase growth is attained is called the acceleration phase.

During the log phase of growth, the cell mass undergoes several cell doublings and the specific growth rate of the culture remains constant. With excess substrate (nutrient supply) and no inhibition of growth by a compound that is present in the growth medium, the specific growth rate is independent of the substrate concentration. These changes and other related steps can be represented in mathematical form, making it possible for biochemical engineers to precisely model and then more easily scale up microbial cell growth. In this case, the rate of increase of the cell biomass with time, dX/dt, is the product of the specific growth rate, μ, and the biomass concentration, X:

$$dX/dt = \mu X$$

Similarly, the rate of increase of the cell number, dN/dt, is the product of the specific growth rate, μ, and the cell number, N:

$$dN/dt = \mu N$$

The specific growth rate, μ, is a function of the concentration of the limiting substrate (i.e., the carbon or nitrogen source), S; the maximum specific

growth rate, μ_{max}; and a substrate-specific constant, K_s. Both S and K_s are expressed in terms of concentration, e.g., in either grams or moles per liter:

$$\mu = \mu_{max} S/(K_s + S)$$

Sometimes, scientists refer to the doubling or generation time, t, of a culture rather than to its specific growth rate, μ, where $t = \ln 2/\mu$. The generation time of a culture is the length of time that it takes, under defined conditions, for the number of cells or the cell biomass to double. For the microorganisms that are commonly grown in culture, the value of μ_{max} varies from about 2.1 to 0.086 h^{-1} (reciprocal hours), which corresponds to doubling times of approximately 20 minutes to 8 hours.

When there is an excess of substrate (i.e., when $S >> K_s$), then $\mu = \mu_{max}$ and the maximal rate of log-phase growth of the culture occurs. In practice, the value of K_s is often so low that substrate levels equivalent to K_s are rarely encountered during log-phase growth. For example, for *Escherichia coli*, while the K_s for glucose is approximately 1 mg/liter, the initial glucose level is usually around 10,000 mg/liter. However, as the culture nears the end of the log phase, the concentration of the remaining substrate, S, is depleted and may even fall below the value of K_s. Under conditions where $S < K_s$, the microorganisms rapidly enter the deceleration phase. However, because of the large cell population at the end of the log phase, the substrate may be so rapidly assimilated that the deceleration phase is short-lived and not observable.

After either the depletion of a critical growth substance, such as the carbon source, from the medium or the accumulation of metabolic end products that inhibit growth, the increase in cell mass eventually ceases and the cells enter the stationary phase. During this phase, although the amount of biomass remains constant, cellular metabolism often changes dramatically; in some instances, compounds (secondary metabolites) that are of considerable commercial interest are synthesized. For example, antibiotics are usually produced during the stationary phase of the microbial growth cycle. The duration of the stationary phase depends on the particular organism and the conditions of growth.

In the death phase, the energy reserves of the cell are virtually exhausted, and metabolic activity ceases. For most commercial processes, the fermentation reaction is halted and the cells are harvested before the death phase begins.

Fed-Batch Fermentation

In fed-batch fermentations, substrate is added in increments at various times throughout the course of the reaction. These additions prolong both the log and stationary phases, thereby increasing the biomass and the amount of synthesis of stationary-phase metabolites, such as antibiotics. However, microorganisms in stationary phase often produce proteolytic enzymes (proteases), and these enzymes can degrade proteins synthesized by a genetically engineered microorganism. Therefore, when proteins are produced from a recombinant microorganism, it is important that the fermentation reaction not be allowed to reach this part of the growth cycle. Because it is often difficult to measure the substrate concentration directly during the fermentation reaction, other indicators that are correlated with the consumption of substrate, such as the production of organic acids,

MILESTONE

Expression of Intracellular Hemoglobin Improves Protein Synthesis in Oxygen-Limited *Escherichia coli*

C. KHOSLA, J. E. CURTIS, J. DEMODENA, U. RINAS, and J. E. BAILEY
Bio/Technology **8:**849–853, 1990

Because of the low solubility of oxygen in water, the growth of aerobic bacteria often becomes limited by the amount of dissolved oxygen in the fermentation broth. This problem is especially acute at high cell densities or when fermentations are performed on a large scale. To address this problem, chemical engineers have attempted to increase the rate of transfer of introduced oxygen to the liquid of the growth medium. Their approaches have included (1) sparging the growth medium with pure oxygen rather than with air; (2) introducing the air (or oxygen) under pressure; (3) adding chemicals, such as perfluorocarbons, to the fermentation broth to increase the solubility of oxygen; and (4) modifying the configuration of the fermentation vessel to optimize the aeration or agitation of the fermentation broth. While all of these solutions to the "oxygen problem" are somewhat effective, they are subject to a threshold effect beyond which it is impossible to introduce a sufficient amount of oxygen to improve the final yield of the fermentation.

As an alternative to these "hardware" solutions, Bailey and coworkers designed a biological system in which the host organism was modified so that it would be more efficient at using the low levels of oxygen that are normally present in the growth medium. They cloned a gene encoding a hemoglobin-like molecule from the gram-negative bacterium *Vitreoscilla* and introduced it into several different recombinant bacteria. The introduced bacterial hemoglobin bound oxygen from the environment and created a higher level of available oxygen within the cells, which resulted in an increase in growth and foreign-gene expression. This approach provided a clever biological solution to what at first glance seemed to be an almost insurmountable engineering problem.

changes in the pH, or the production of CO_2, can be used to estimate when additional substrate is needed. Generally, fed-batch fermentations require more monitoring and greater control than batch fermentations and are therefore used to a lesser extent. However, since they may be advantageous in the development of systems for the production of proteins from recombinant microorganisms, they are becoming increasingly popular.

The periodic addition of substrate to a growing microbial culture prolongs the log phase of growth and delays the onset of the stationary phase, which initiates cellular stress responses, the production of proteases, and other metabolic changes that affect the yield of a recombinant protein. Nevertheless, with continued cell growth, an increasing amount of the incoming substrate is needed for maintenance of the host cell metabolism. This means that fewer cellular resources are used for the synthesis of the recombinant protein or commercial metabolite(s). To ensure that the synthesis and stability of a recombinant protein are not impaired, increasing amounts of nutrients must be added to the growing culture. This may be done by carefully monitoring the fermentation reaction and adding substrates (carbon and nitrogen sources, together with trace elements) in increasing amounts as they are needed. Depending upon the particular microorganism, its genetic background, and the nature of the recombinant protein, a fed-batch fermentation strategy can increase the yield from 25% to more than 1,000% compared with batch fermentation.

Fed-batch processes are not limited to microbial cells but are also used with mammalian and insect cells in culture. This is important because (1) these cell culture systems are increasingly being used for the production of human therapeutic proteins, and (2) in the absence of fed-batch strategies, animal cells in batch culture are not very efficient in producing foreign proteins.

Continuous Fermentation

In a continuous fermentation, a steady-state condition, where $dX/dt = 0$, is attained when the total number of cells and the total volume in the bioreactor remain constant. In other words, under these conditions, the loss of cells due to outflow (product removal) is exactly balanced by the gain in new cells by growth (division). In more formal terms, for a continuous steady-state fermentation process, the dilution rate, *D,* is defined as the volumetric flow rate, *F,* divided by the constant liquid volume, *V,* in the bioreactor:

$$D = F/V$$

where D is equal to the specific growth rate, μ:

$$D = (dX/dt)(1/X) = \mu$$

To obtain hydrodynamically stable continuous cultures, the specific growth rate, μ, of the culture must be lower than the maximum attainable specific growth rate, μ_{max}. In practice, this condition is achieved by adjusting a pump that controls the volumetric flow rate, *F,* while keeping the volume of the culture within the bioreactor, *V,* constant.

The fundamental objective of industrial fermentations is to minimize costs and maximize yields. This goal can be achieved by developing the most efficient mode of fermentation for each particular process. Although the commercial use of continuous fermentation processes is typically limited to production of single-cell protein, antibiotics, and organic solvents, primarily because of the greater experience that scientists have with growing cells in batch mode, the cost of producing a specific amount of cell biomass by continuous culture is potentially much lower than producing the same amount by batch fermentation. The following factors account for the savings.

- Continuous fermentations use smaller bioreactors than batch fermentations to produce the same amount of product.
- After a large-scale batch fermentation is completed, large-scale equipment is needed for cell harvesting, cell breakage, and subsequent downstream processing (purification) of the protein or metabolite product that is produced by the microorganism. Continuously grown cells, however, are produced "a little bit at a time," so that the equipment required for cell harvesting, cell breakage, and downstream processing can be much smaller.
- Continuous fermentation, by definition, avoids the "down time" between batch runs, during which the bioreactor is prepared for reuse. A common hindrance to efficient industrial fermentation is the loss of productivity due to the down time of the bioreactor for repairs, cleaning, or sterilization. Continuous fermentations have less down time, because a single reaction can be maintained for a much longer period.
- The physiological state of the cells during continuous fermentation is more uniform, so that yields of product are more consistent. In batch fermentations, small differences in the timing of cell harvest, which coincides with the mid- to late log phase of growth, can lead to significant physiological differences.

Despite its merits, continuous fermentation has potential drawbacks that must be addressed before its use becomes more widespread.

- The duration of a continuous fermentation can be 500 to 1,000 hours, and therefore, some cells might lose recombinant plasmid constructs. Cells that lack plasmids usually have a smaller energy burden and divide faster than do plasmid-containing microorganisms, so that yields may decline with time because fewer cells are synthesizing the product protein. Integration of the cloned gene into the genome of the host organism avoids this problem.
- Maintenance of sterile conditions on an industrial scale for long periods is difficult. Furthermore, continuous processes need sterile backup equipment, a requirement that can greatly increase capital costs.
- The composition of culture medium that is used for industrial fermentations is not subject to the same level of quality assurance as that accorded laboratory medium components and therefore may vary from batch to batch. This variation can alter the physiology of the cells and decrease productivity.

Because batch fermentation has a proven history of reliability, there is reluctance to switch to another type of fermentation system, even though a continuous mode of operation is generally regarded as the most efficient fermentation strategy. Nevertheless, a number of researchers have recently developed, on the scale of a laboratory (up to 10 liters) or pilot plant (up to 1,000 liters), both continuous and fed-batch processes for the production of proteins from recombinant microorganisms. Therefore, it is probably only a matter of time before the use of continuous and fed-batch fermentations becomes more widespread in industry.

Maximizing the Efficiency of the Fermentation Process

Regardless of the type of fermentation process that is used to grow cells, it is necessary to monitor and control culture parameters, such as the dissolved oxygen concentration, pH, temperature, and degree of mixing. Changes in any one of these parameters can have a dramatic effect on the yield of cells and the stability of the protein product.

Optimal growth of *E. coli* cells and many other microorganisms that are used as hosts for cloned genes usually requires large amounts of dissolved oxygen. The maximal oxygen demand in a fermentation, Q_{max}, is dependent on the cell mass, X; the maximal specific growth rate, μ_{max}; and the growth yield based on oxygen consumed, Y_{O_2}, where

$$Q_{max} = X\mu_{max}/Y_{O_2}$$

Because oxygen is only sparingly soluble in water (0.0084 gram/liter at 25°C), it must be supplied continuously—generally in the form of sterilized air—to a growing bacterial culture. However, the introduction of air into a bioreactor produces bubbles, and if the bubbles are too large, the rate of transfer of oxygen to the cells is insufficient to support optimal growth. Thus, fermenter design should include provision for monitoring the dissolved-oxygen level of the culture, providing oxygen to the culture, and adequately mixing the culture to efficiently disperse the bubbles.

Most microorganisms grow optimally between pH 5.5 and 8.5. However, during growth in a bioreactor, cellular metabolites are released into the growth medium, a process that can change the pH of the medium. Therefore, the pH of the medium must be monitored and either acid or base must be added as needed to maintain a constant pH. Of course, the added acid or base must be well mixed into the fermentation broth so that the pH of the growth medium is the same throughout the entire reaction vessel.

Maintenance of the correct temperature is essential for the success of a fermentation reaction. Microorganisms grown at a temperature below the optimum grow slowly and have a reduced rate of cellular production (productivity). On the other hand, if the growth temperature is too high—but not lethal—there may be premature induction of the expression of the target protein, if it is under the control of a temperature-sensitive repressor, or induction of a heat shock (stress) response, which will produce cellular proteases that lower the yield of the protein product.

Adequate mixing of a microbial culture is essential for many aspects of a fermentation, including assurance of an adequate supply of nutrients to the cells and prevention of the accumulation of any toxic metabolic byproducts in local, poorly mixed regions of the bioreactor. Effective mixing is relatively easily attained with small-scale cultures, but it is one of the major problems when the scale of fermentation is increased.

Agitation of the fermentation broth also affects other factors, such as the rate of transfer of oxygen from the gas bubbles to the liquid medium and then from the medium to the cells, efficient heat transfer, accurate measurement of specific metabolites in the culture fluid, and efficient dispersion of added solutions, such as acids, bases, nutrients, or antifoaming agents. On these grounds, it might be concluded that the more mixing there is, the better the growth. However, excessive agitation of a fermentation broth can cause hydromechanical stress (shear), which damages larger microbial or mammalian cells, and a temperature increase, which may also decrease cell viability. Thus, a balance must be struck between the need to provide thorough mixing and the need to avoid damage to the cells.

There is an additional consideration for scaled-up fermentations that has nothing to do with the technical aspects of the process but depends instead on whether a genetically engineered microorganism is being used. In most countries, specific rules and regulations must be followed when genetically engineered microorganisms are grown on a large scale. Although most recombinant microorganisms are not hazardous, it is nevertheless important to ensure that they are not inadvertently released into the environment. Therefore, fail-safe systems are used to prevent accidental spills of live recombinant organisms and to contain them if they occur. Furthermore, all recombinant microorganisms must be treated by an approved procedure to render them nonviable before they are discharged from the production facility. The spent culture medium must also be treated to ensure that it does not contain viable organisms and that its disposal does not create an environmental problem.

High-Density Cell Cultures

A major objective of fermentation is to maximize the volumetric productivity, i.e., to obtain the largest amount of product in a given volume in as

short a time as possible. High cell densities are absolutely necessary for high productivity. Generally, when foreign proteins are produced by recombinant *E. coli,* the greater the final cell density, the greater the amount of product that is formed. In practice, cell concentrations of more than 50 (and in a few cases more than 150) grams (dry weight) of cells per liter of culture have been obtained with fed-batch cultures of recombinant *E. coli*. The dry weight of *E. coli* cells is approximately 20 to 25% of the wet weight.

One way to achieve a high density of *E. coli* cells is to optimize the growth medium. Some nutrients, including carbon and nitrogen sources, can inhibit cell growth if they are present at too high a concentration. Glucose is inhibitory above 50 grams per liter, ammonia is inhibitory above 3 grams per liter, iron is inhibitory above 1.15 grams per liter, magnesium is inhibitory above 8.7 grams per liter, phosphorus is inhibitory above 10 grams per liter, and zinc is inhibitory above 0.038 gram per liter. Therefore, merely increasing the amount of nutrients in the growth medium in batch culture does not necessarily yield a high cell density. In addition, since the nutrients in complex media, such as peptone or yeast extract, can vary from one batch of medium to another, fermentations that use complex media are not always reproducible.

Acetate, which can be inhibitory to cell growth, is produced by *E. coli* both when the cells are grown under oxygen-limiting conditions and in the presence of excess glucose. The acetate problem can be minimized by using glycerol instead of glucose as a carbon source, lowering the culture temperature, or using an *E. coli* strain that has been genetically engineered to shunt acetate into less toxic compounds (see below).

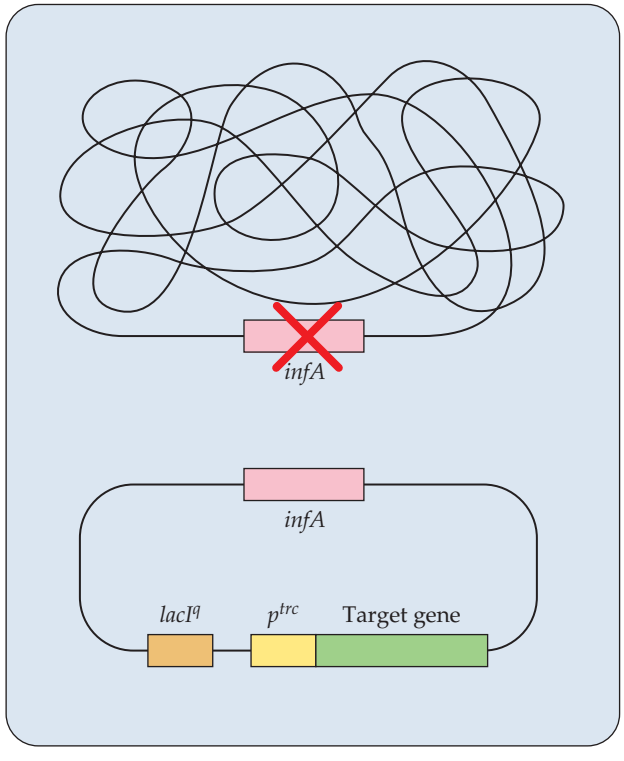
Oxygen may also become limited in high-density cell cultures. To overcome this problem, the rate of introduction of air (sparging), the agitation rate, or both can be increased. Also, pure oxygen rather than air, which is only approximately 20% oxygen, can be introduced into growing cell cultures. Cells can also be grown under pressure to increase the solubility of oxygen, which increases the rate of transfer of oxygen to the cells in the aqueous growth medium. Alternatively, expression in host cells of the *Vitreoscilla* hemoglobin gene in a number of different organisms has been shown to significantly increase the uptake of oxygen by growing cells and to thereby increase the amount of product formed (see chapter 6). For example, *Vitreoscilla* hemoglobin can enhance growth and heterologous protein production in *E. coli,* improve enzyme production in *Bacillus subtilis,* increase erythromycin production by *Saccharopolyspora erythraea,* improve the rate of degradation of benzoic acid by *Xanthomonas maltophilia,* and enhance the production of cephalosporin C by *Acremonium chrysogenum*.

High-density cell cultures are most readily attained in fed-batch cultures. The addition of nutrients following the depletion of some of the original nutrients may be constant, stepwise, or exponential. With constant-rate feeding, nutrients are added at the same rate throughout the fermentation. However, under these conditions, the specific growth rate continually declines. With stepwise feeding, increasing amounts of nutrients are added at higher cell concentrations. In this case, the specific growth rate decline is largely compensated for. With exponential feeding, nutrients are added at an exponential rate, with the result that a constant specific growth rate can be achieved. It is possible to automate the fed-batch addition of nutrients based on measuring the concentration of a growth-limiting substrate, such as glucose, in the culture medium during the fermentation process.

Increasing Plasmid Stability

The loss of plasmids during the large-scale growth of recombinant *E. coli* cells is a major industrial problem. Plasmid loss often limits the yield of plasmid-encoded recombinant proteins, especially when cells are grown in continuous culture. Plasmid instability in bacterial cultures is typically a consequence of the unequal distribution of plasmids to daughter cells during growth and cell division. Generally, once cells have lost a plasmid, they grow faster, with the result that cells lacking plasmids eventually dominate the culture. One approach to avoid this problem is to include an antibiotic resistance gene on the plasmid being used and then add that antibiotic to the culture medium. In addition to the obvious economic cost of the antibiotic, especially when dealing with large-scale cultures, disposal of spent growth medium is a potential environmental hazard in that both the antibiotic and antibiotic resistance genes may be released into the environment. One way to get around this problem is to delete an essential gene from the chromosomal DNA of the host bacterium and at the same time place this gene on the plasmid that is being stabilized. As a result, only plasmid-carrying cells can grow, making the bacterial strain totally dependent upon maintenance of the plasmid. In one example, the essential gene that was used encoded translation initiation factor 1 (Fig. 17.4). The target gene was placed under the transcriptional control of the strong and IPTG (isopropyl-β-D-thiogalactopyranoside)-inducible p^{trc} promoter.

FIGURE 17.4 Schematic representation of a bacterial cell in which the essential protein synthesis initiator factor 1 gene *(infA)* was deleted from the chromosome and included on a plasmid. The target gene is under the transcriptional control of the strong Trc promoter (p^{trc}) that is controlled by the *lac* repressor encoded by $lacI^q$, which overproduces the protein.



With this system, selection that utilizes antibiotics is no longer necessary, thereby decreasing both the cost and the environmental risk associated with large-scale fermentation.

Quiescent *E. coli* Cells

While it is possible to achieve high levels of foreign-gene expression in *E. coli* and other bacterial host cells, it is difficult to engineer recombinant bacteria to produce large amounts of a foreign protein and, at the same time, to grow to a high cell density. This is because a recombinant bacterial cell partitions its resources between production of the foreign protein and cell growth. It would be advantageous to be able to first grow cells to a high density and then to shift the allocation of available resources from growth to foreign-protein production. With this in mind, one group of workers engineered a quiescent cell expression system in which a plasmid-encoded protein is expressed in nongrowing but metabolically active cells. The quiescent state is established by the overexpression of Rcd, a regulatory protein, in an *hns* mutant *E. coli* host cell. The *hns* gene codes for a histone-like nucleoid-structuring protein. Cultures of the *hns* mutant of *E. coli* in which the *rcd* gene is induced gradually cease synthesizing host proteins but continue synthesizing plasmid-encoded foreign proteins for many hours after induction. In one study that utilized this system, the *rcd* gene was placed under the transcriptional control of the p^R promoter while the recombinant protein gene (encoding a single-chain antibody variable fragment [scFv]) was controlled by the p^L promoter (Fig. 17.5). The activities of both the p^R and p^L promoters are repressed by a temperature-sensitive *c*I repressor protein; in this system, the gene for this protein is encoded in the host chromosomal DNA. When cells are grown at 30°C, the *c*I repressor prevents transcription from both p^R and p^L. When the temperature is shifted to 42°C, the temperature-sensitive *c*I repressor protein is inactivated, and transcription can proceed from both p^R and p^L (see chapter 6). The temperature shift therefore causes the cells to become quiescent and at the same time to synthesize the recombinant scFv protein. In this particular case, the scFv protein contained a leader peptide that directed ~90% of the protein to be secreted into the growth medium. As shown in Table 17.1, in both batch and fed-batch modes, the quiescent cells produce less biomass and secrete considerably more of the scFv protein into the growth medium than do control *E. coli* host cells engineered to express scFv under the control of the p^L promoter. Understanding the commercial potential of this unique system, the scientists who developed this approach have applied for a patent to protect their intellectual property rights.

Protein Secretion

High-level cytoplasmic expression in *E. coli* of many different foreign proteins results in the formation of inclusion bodies consisting of insoluble improperly folded protein. Even when the foreign protein is soluble, purifying it from a cytoplasmic extract can be a major undertaking. In addition, sometimes proteins that are secreted into the growth medium are produced at a much higher level than when they are expressed in the cytoplasm. While these considerations make only a small difference in laboratory-scale experiments, they are of critical importance when foreign proteins are produced on a large scale.

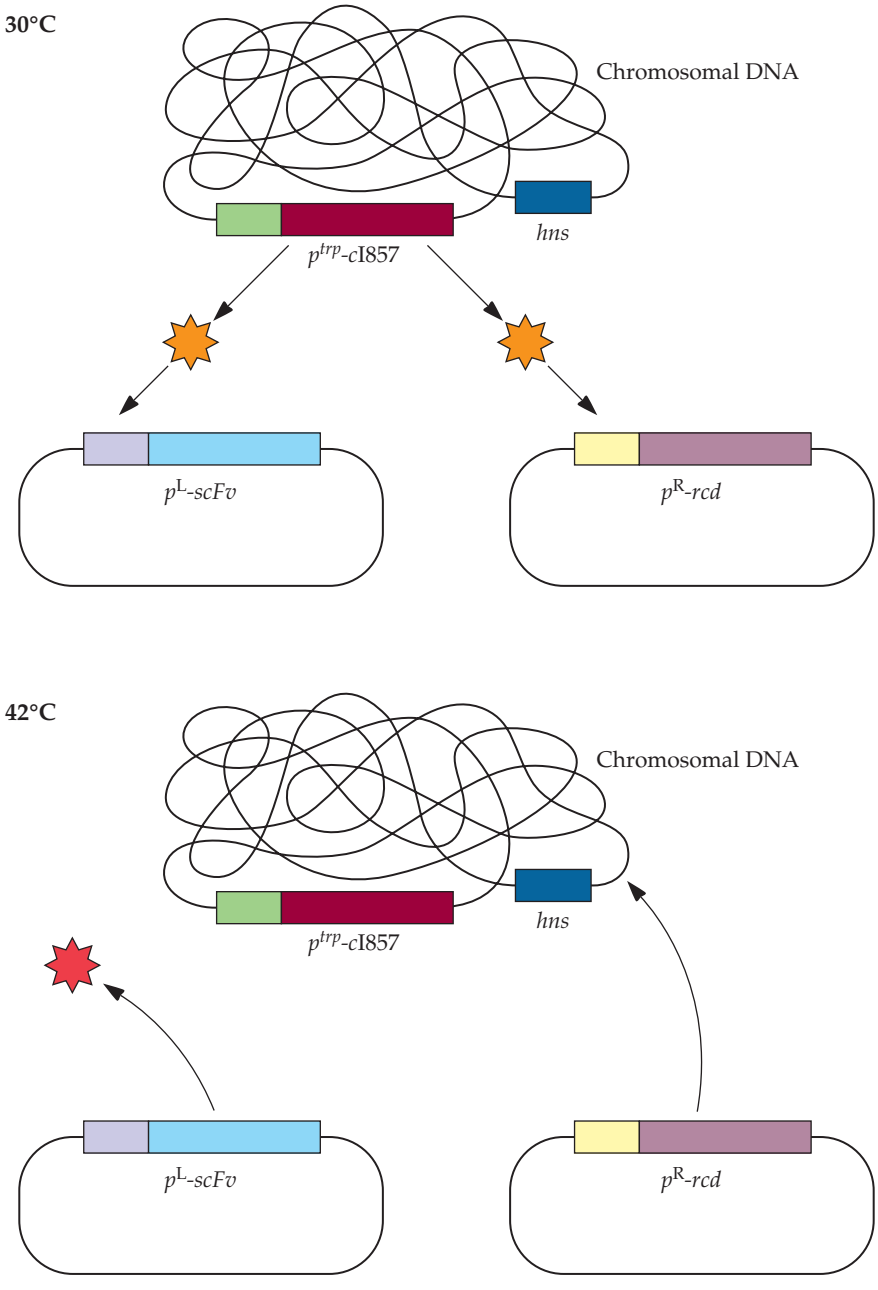


FIGURE 17.5 Synthesis of a single-chain antibody fragment (scFv) in "quiescent *E. coli* cells." At 30°C, the *c*I repressor, encoded by *c*I857 (inserted into the chromosomal DNA), binds to the operators of the p^R and p^L promoters and prevents transcription of the plasmid-encoded *scFv* and *rcd* genes. At 42°C, the temperature-sensitive *c*I repressor is inactivated so that transcription directed by the p^R and p^L promoters proceeds. Turning on the p^R promoter causes the Rcd protein to be synthesized, thereby causing the cells to become quiescent. At the same time, turning on the p^L promoter activates transcription of the gene encoding scFv. When the *rcd* gene is induced, a mutant of the *hns* gene causes the cessation of host cell protein synthesis.

TABLE 17.1 Cell growth and foreign protein (scFv) secreted into the growth medium for *E. coli* cells with induced quiescence compared to control (wild-type) cells

Cell type	Batch culture		Fed-batch culture	
	Cell growth (optical density at 600 nm)	scFv secreted (mg/liter)	Cell growth (optical density at 600 nm)	scFv secreted (mg/liter)
Quiescent	3.5	37	20	150
Control	20	13	80	35

One group of investigators observed that expression levels were quite low when they expressed several different foreign proteins, i.e., human granulocyte-macrophage colony-stimulating factor, α-interferon 2b (IFN-α2b), or scFv, under the transcriptional control of the strong p^m/*xylS* promoter/regulator system. The yields of human granulocyte-macrophage colony-stimulating factor and scFv, but not IFN-α2b, increased dramatically when the genes encoding these proteins were fused to a translocation signal sequence (Fig. 17.6). Interestingly, different translocation signal sequences were optimally effective with each of the proteins tested. To obtain a high level of expression of IFN-α2b, before assembling the construct, it was necessary to chemically synthesize the gene in order to eliminate the use of codons that are rarely used in *E. coli*. While the use of translocation signal sequences significantly stimulated the levels of expression of these three human proteins, depending on the protein and the translocation signal sequence, from 20 to 50% of the protein that was produced was found to be in an insoluble form. In order for this system to be used routinely for the large-scale production of human proteins in *E. coli*, a strategy that minimizes the extent of insoluble protein formation needs to be developed.

Reducing Acetate

It is often difficult to achieve high levels of foreign-gene expression and a high cell density at the same time because of the accumulation of harmful waste products, especially acetate, which inhibits both cell growth and protein production and also wastes carbon and energy resources. One strategy to reduce the inhibitory effects of acetate is to remove the acetate from the culture during the course of the fermentation. This may be achieved by several different methods, including continuous dialysis and the use of macroporous ion-exchange resins. However, these methods tend to remove nutrients that are necessary for cell growth along with the acetate.

Since acetate is often associated with the use of glucose as a carbon source, lower levels of acetate, and hence higher yields of protein, are generally obtained when fructose or mannose is used as a carbon source. Another strategy for reducing acetate accumulation in rich medium without impairing cell growth entails decreasing the glucose uptake rate of the cells by adding methyl α-glucoside, a glucose analogue, to the growing cells. The same effect has also been achieved by using an *E. coli* host cell that contained a mutation in *ptsG*, a gene encoding enzyme II in the glucose phosphotransferase system. In a comparison of batch cultures of wild-type and *ptsG* mutant *E. coli* cells in rich medium, with both carrying a plasmid expressing β-galactosidase activity, the wild-type cells attained a density of approximately 10 grams (dry weight) per liter, while the mutant cells attained more

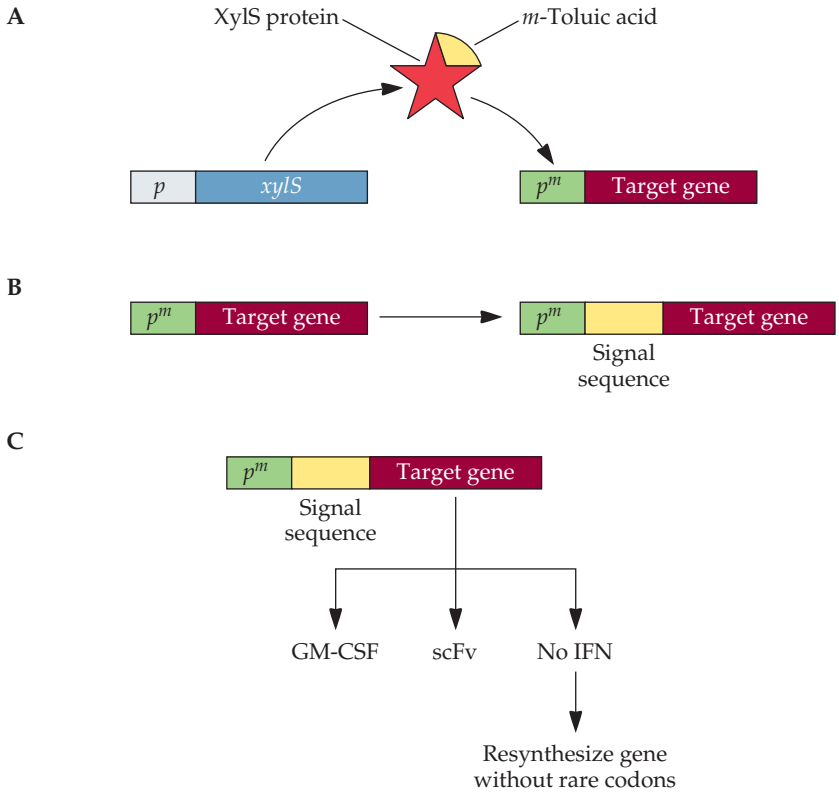


FIGURE 17.6 Engineering proteins for high-level expression following high-cell-density growth. **(A)** A target protein gene is placed under the transcriptional control of the strong and inducible (in this case by *m*-toluic acid) p^m promoter. **(B)** The target gene is fused to any one of several different secretion signal sequences. **(C)** Rare codons are removed from protein coding sequences that are not well expressed by chemically synthesizing the entire gene. GM-CSF, granulocyte-macrophage colony-stimulating factor; IFN, IFN-α2b.

than 15 grams (dry weight) per liter. At the same time, the mutant cells synthesized about 25% more β-galactosidase per gram (dry weight) of cells than the wild-type cells did. Overall, the *ptsG* mutant cells synthesized nearly twice as much β-galactosidase as did the wild-type cells.

Since it is often much easier and quicker to alter a particular host cell by genetic transformation than by mutagenesis and selection, alternative means of reducing acetate production in cells were developed. One of these methods includes introducing a gene (from *B. subtilis*) encoding the enzyme acetolactate synthase into *E. coli* host cells. This enzyme catalyzes the formation of acetolactate from pyruvate, thereby decreasing the flux through acetyl coenzyme A to acetate (Fig. 17.7). In practice, the acetolactate synthase genes are introduced into the cell on one plasmid, while the target gene (encoding the protein that is to be overexpressed in *E. coli*) is introduced on a second plasmid from a separate incompatibility group. The cells that were transformed with the acetolactate synthase genes produced 75% less acetate than the nontransformed cells and instead synthesized acetoin, which is approximately 50-fold less toxic to cells than acetate. The protein yield was also doubled.

An alternative strategy to converting acetate to acetoin is to redirect carbon flow to the tricarboxylic acid (TCA) cycle (citric acid cycle, or Krebs

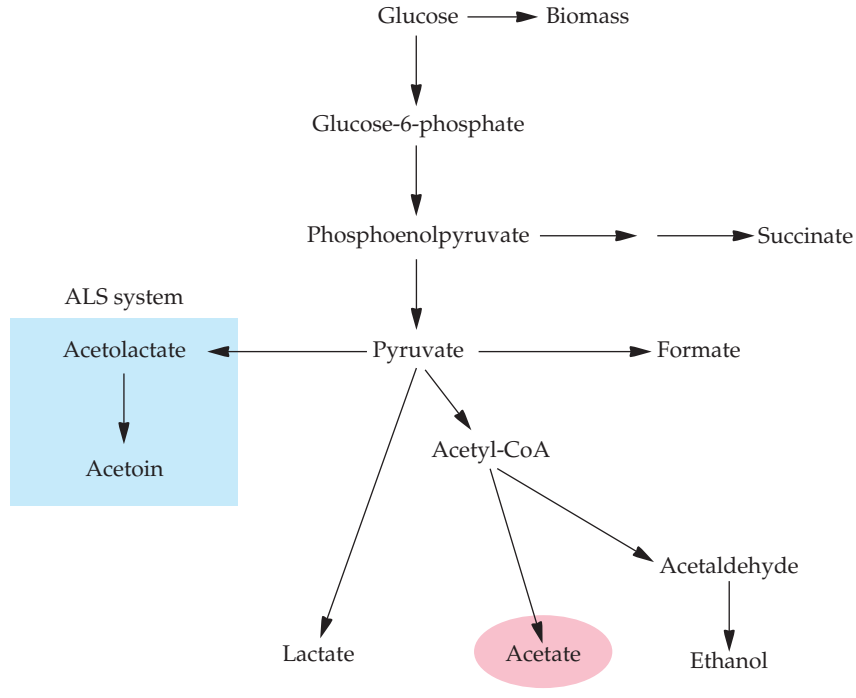


FIGURE 17.7 Schematic representation of the pathways for glucose metabolism in an *E. coli* strain that has been transformed with a plasmid carrying the genes for the protein subunits of acetolactate synthase (ALS). Note that the conversion of glucose to biomass is a multistep process. CoA, coenzyme A.

cycle). This is necessary because recombinant protein production decreases carbon flow in the TCA cycle as a consequence of the withdrawal of the intermediates that serve as protein precursor biochemicals. In one study, workers overexpressed the gene for the enzyme phosphoenolpyruvate carboxylase, which converts phosphoenolpyruvate to oxaloacetate, with the result that they obtained a 17% increase in the specific growth rate of the *E. coli* cells and a 44% decrease in acetate production. Unfortunately, overexpressing this enzyme also decreases the amount of glucose uptake by the bacterial cells and diminishes the growth rate. As an alternative approach to replenishing the TCA cycle, another group of researchers transformed *E. coli* host cells with the gene for the enzyme pyruvate carboxylase, which converts pyruvate directly to oxaloacetate (Fig. 17.8). Since *E. coli* does not normally contain pyruvate carboxylase, the gene was isolated from a strain of the gram-negative bacterium *Rhizobium etli.* With the introduction of pyruvate carboxylase, acetate levels were decreased, the cell yield was increased, and the amount of foreign protein synthesized was increased (Table 17.2). This result reflects the fact that the addition of pyruvate carboxylase allows *E. coli* cells to use the available carbon more efficiently, directing it away from acetate toward biomass and protein formation. Although it has not been tested extensively, it is thought that this strategy may be a generally effective method for increasing the level of expression of foreign proteins produced in *E. coli* host cells.

Similar to the strategy discussed above, the TCA cycle may also be replenished by converting aspartate to fumarate (Fig. 17.8). To do this, *E. coli* host cells were transformed with the gene for L-aspartate ammonia lyase

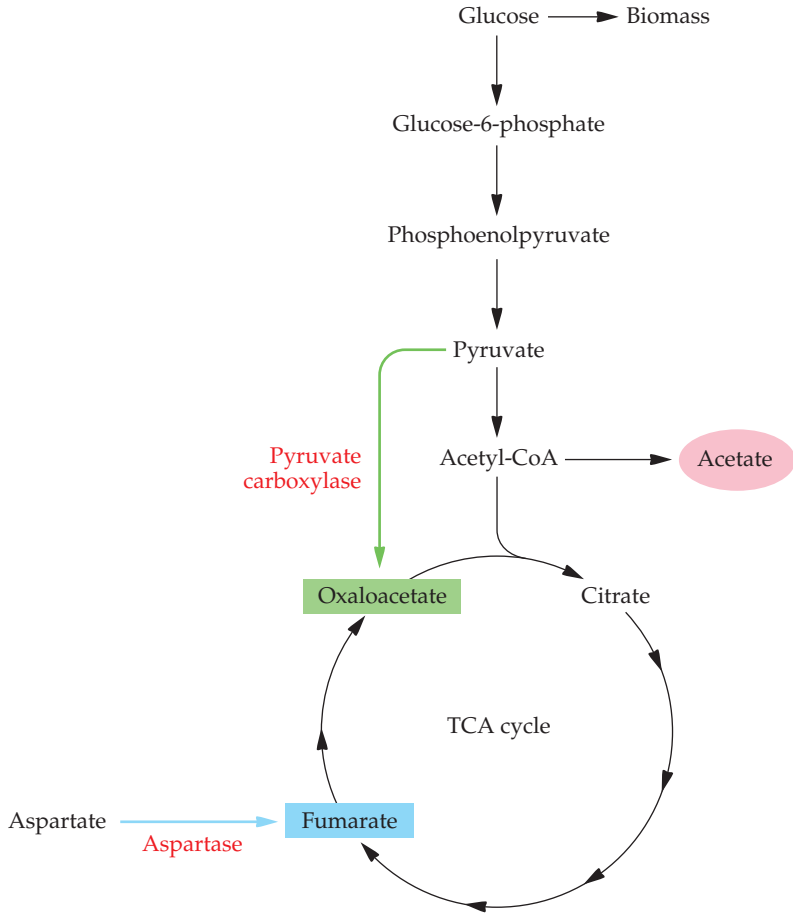


FIGURE 17.8 Replenishment of the TCA cycle in *E. coli* by the introduction of a gene from *R. etli* encoding pyruvate carboxylase. This avoids the conversion of pyruvate to acetate. The TCA cycle may also be replenished by the introduction of a gene encoding aspartase, converting aspartate in the medium to fumarate. Note that the conversion of glucose to biomass is a multistep process. CoA, coenzyme A.

(aspartase) under the control of the strong *tac* promoter on a stable low-copy-number plasmid. Aspartase activity is induced by the addition of IPTG at the mid- to late log phase of growth. The target recombinant protein is introduced on a separate plasmid. Using this system in minimal medium containing aspartate, the production of different recombinant proteins could be increased up to fivefold, with 30 to 40% more biomass production.

Bioreactors

A cursory examination of the biochemical engineering literature may give the impression that there are a limitless number of bioreactor designs. However, closer inspection reveals that virtually all of these designs fall into three fundamental classes:

- Stirred-tank reactors (STRs), which have internal mechanical agitation (Fig. 17.9A)
- Bubble columns, which rely on the introduction of air or another gas (sparging) for agitation (Fig. 17.9B)

- Airlift reactors, which have either an internal (Fig. 17.9C) or an external (Fig. 17.9D) loop; the mixing and circulation of the culture fluid in these reactors are the results of the motion of an introduced gas (usually air), which causes density differences within the different parts of the bioreactor.

The traditional, and by far the most commonly used, bioreactor is the STR. This type of bioreactor has several advantages over other bioreactor configurations.

- It has highly flexible operating conditions.
- It is readily available commercially.
- It provides efficient gas transfer to the growing microbial cells, or, in the words of fermentation engineers, the volumetric mass transfer coefficient, $k_L a$, of STRs is high.
- It has been used extensively by fermentation engineers and microbiologists for growing a variety of microorganisms.

In an STR, gas, usually air, is added to the culture medium under pressure through a device called a sparger, which can be either a ring with many small holes or a tube with a single orifice. Although sparging rings generate smaller bubbles and consequently create better initial gas distribution, sparging tubes are often preferred in many small-scale applications (<20 liters) because they are less likely to become plugged. Thorough dispersion of the gas within the bioreactor requires one or more impellers (agitators) in addition to the sparger. Mechanical agitation of the culture medium by the impellers breaks larger bubbles into smaller ones, disperses the bubbles throughout the medium, and enhances the residence time of the bubbles in the bioreactor. At high levels of agitation, the mean size of the bubbles in large bioreactors is essentially independent of the size of the holes in the sparger. The type of impeller, its rotational speed, and the physicochemical properties of the liquid phase are important factors that give rise to efficient gas dispersion. In large bioreactors, however, if the initial gas distribution from the sparger is not uniform across the tank, even vigorous agitation may not create a homogeneous gas environment.

Because of the corrosive or abrasive nature of many culture media and sterilization procedures, STRs are usually constructed from stainless steel or glass. The glass units are usually limited to laboratory-scale bioreactors that have a capacity of <50 liters.

One limitation on the size of a bioreactor is the ability of the system to efficiently remove heat that is generated as a consequence of either the metabolism of the growing cells or the energy input by agitation. Too much heat raises the temperature and alters the physiological state of the cells,

TABLE 17.2 Relative acetate levels, cell yields, and foreign-protein activities from *E. coli* cells transformed with an *R. etli* pyruvate carboxylase gene

E. coli strain	Acetate concentration	Cell yield	Foreign-protein activity
– Pyruvate carboxylase	1.0	1.0	1.0
+ Pyruvate carboxylase	0.43	1.41	1.68

Adapted from March et al., *Appl. Environ. Microbiol.* **68:**5620–5624, 2002.

The foreign protein was β-galactosidase, whose activity is relatively easy to quantify. The data have been normalized to the values for the *E. coli* strain without (–) pyruvate carboxylase.

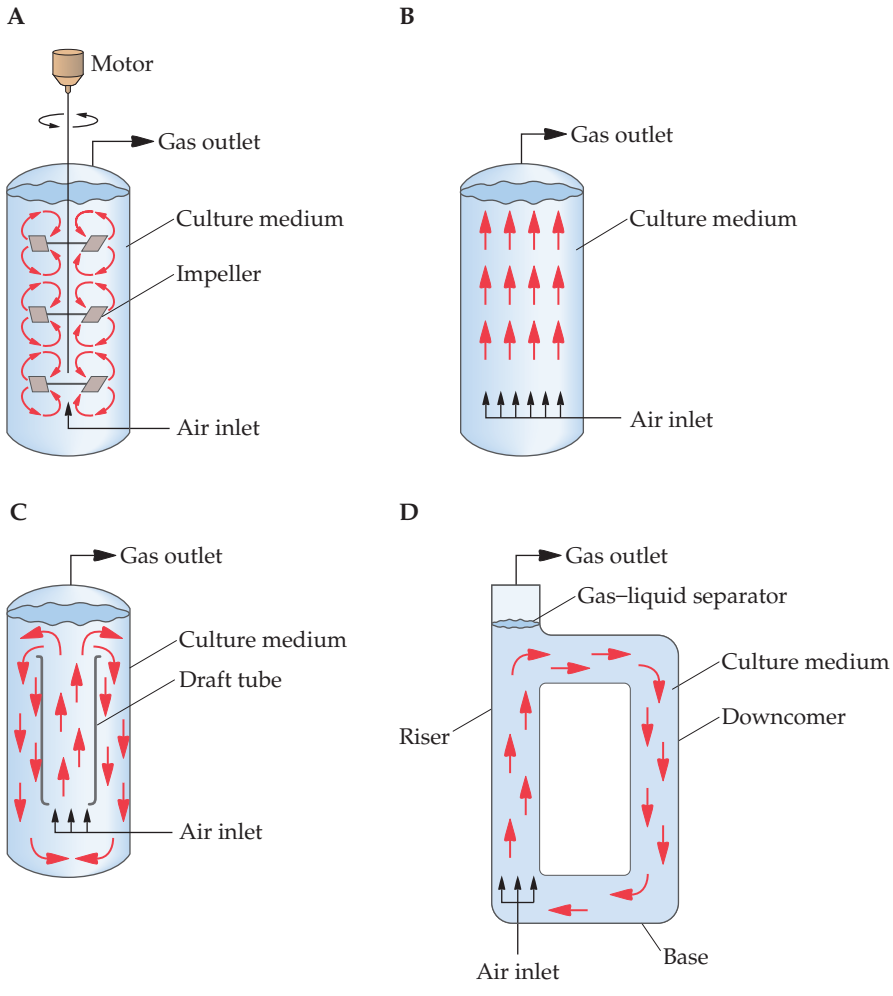


FIGURE 17.9 Simplified examples of bioreactor configurations. **(A)** STR; **(B)** bubble column; **(C)** internal-loop airlift reactor with a central draft tube; **(D)** external-loop airlift reactor. The arrows within the bioreactors indicate the direction of flow of the culture medium.

and decreases the product yield. Heat can be removed by using a cooling jacket around the reaction vessel or by internal coils. Although internal cooling coils are more effective than jackets in keeping the fermentation reaction close to the desired temperature, they can become fouled (coated) with microorganisms, which prevents cooling, and they sometimes interfere with the proper agitation of the fermentation broth.

Contamination of a fermentation with fungi or bacteria is usually disastrous. Therefore, bioreactors are designed so that they can be sterilized, usually with pressurized steam. There should be no internal dead spaces or surfaces that escape contact with the steam during sterilization. All seals, probes, and valves also must be readily steam sterilizable. When a bioreactor is designed, a trade-off is often made between a full set of ports for probes, which enables monitoring of the fermentation parameters, and fewer ports, which makes the maintenance of sterility easier.

The high level of agitation of the culture medium during a fermentation reaction often causes considerable foaming. Excessive foam can wet

the filter at the port through which the introduced air exits the bioreactor, thereby both decreasing the air flow and providing a potential pathway for the entry of contaminating cells. Either chemical antifoaming agents or mechanical foam breakers can be used to control foaming. However, the chemical agents can diminish the extent of microbial growth by preventing oxygen transfer or, in some cases, by inhibiting cellular enzymes. Furthermore, if an antifoaming compound is not removed before downstream processing, it can contaminate the final product. Foaming can also be controlled by providing sufficient "head space" within the bioreactor, i.e., space above the liquid in which the bubbles can dissipate. In practice, then, the "working volume," or actual volume of the culture, in an STR is typically only about 75% of the total volume of the bioreactor.

Many of the considerations that apply to STRs also apply to pneumatic reactors, such as bubble columns and airlift bioreactors. Thus, for example, sterility, constant pH, and constant temperature are key components of any fermentation, regardless of the precise configuration of the bioreactor.

The configurations of bubble columns and airlift bioreactors give them some distinct advantages over STRs. These pneumatic reactors are more energy efficient than STRs because agitation is provided by the injection of a stream of air—or another gas if anaerobic microorganisms are being grown—rather than by a mechanical stirrer. Also, with the elimination of the mixer shaft in these units, there is one less potential site of entry for contaminating organisms.

Pneumatic reactors generate a lower-shear environment than do STRs. Also, in airlift reactors, the shear stress is more evenly distributed throughout the vessel than in STRs. The reduction of shear forces is important for the following reasons.

- Genetically engineered microorganisms are often more susceptible to lysis when exposed to shear stress than are unmodified organisms, because the extra metabolic burden of synthesizing a foreign protein often causes genetically engineered microorganisms to form weakened cell walls.
- A frequent cellular response to the shear forces is decreased synthesis of all cellular proteins, including the recombinant protein.
- Shear stress can alter the physical and chemical properties of the cells so that the downstream processing steps become more difficult to perform. For example, the fermentation conditions can inadvertently increase the amount of surface polysaccharides that a microorganism produces and, as a consequence, can change the conditions for effective harvesting and lysis of these cells, making it more difficult to purify the target protein.

In bubble columns, the air is introduced under high pressure near the bottom, but the smaller bubbles coalesce into larger ones as they rise through the column, leading to uneven gas distribution. In addition, the use of high-pressure air tends to cause excessive foaming of the medium. These disadvantages restrict the flexibility or effective range of operating conditions, as well as the potential size of bubble columns.

Airlift bioreactors, however, can be readily adapted for either pilot plant or large-scale fermentation processes. In an airlift reactor, the gas is introduced into the bottom of a vertical channel (riser). Both the gas and liquid flow up the riser until they reach an open space at the top (gas–liquid separator), where the gas is at least partially disengaged from the liquid. The

degassed liquid, which is denser than the gassed liquid, descends in a separate vertical channel (downcomer) and moves along the base of the reactor until it reaches the bottom of the riser. In this way, the culture fluid and cells are continuously being circulated around the bioreactor.

There are two main types of airlift bioreactors: those that have a single container with internal baffles that create interior liquid circulation channels (internal-loop reactors) (Fig. 17.9C) and those that have an external loop so that the culture liquid circulates through separate, independent channels (external-loop reactors) (Fig. 17.9D). Internal-loop airlift reactors are simple in design, but once they are constructed, both the volume and the circulation rate are fixed for all fermentation processes. In contrast, the external loop of external-loop airlift reactors can be easily changed or modified, e.g., by altering its volume, to suit the requirements of different fermentations.

Airlift bioreactors are generally more efficient than bubble columns, especially for denser or more viscous suspensions of microorganisms. In airlift reactors, mixing is generally better, and bubble coalescence is not as big a problem as it is in bubble columns. In extremely large airlift fermenters, such as the 1,500,000-liter fermenter built by Imperial Chemical Industries, Ltd., in England for the production of single-cell protein, it takes a considerable amount of time for a cell to complete a full cycle through the reactor. To prevent the cells from becoming substrate depleted while they are traversing the bioreactor, there are multiple injection points for the introduction of substrate along the length of the unit.

Typical Large-Scale Fermentation Systems

When recombinant microorganisms are used to overproduce protein products, such as pharmaceuticals like insulin, cells are typically grown to mid- to late log phase of the growth cycle so that the target protein levels will be optimal. On the other hand, when recombinant microorganisms are used as "factories" to synthesize useful metabolites, such as antibiotics, host cells are commonly grown to the deceleration or stationary phase, where the synthesis of secondary metabolites is often optimal. Clearly, these sorts of differences must be considered when a large-scale fermentation process is being developed.

For maximal protein production, it is generally preferable to use cloned genes that are under the control of strong promoters that can be regulated. Initially, it was thought that constitutive expression of a cloned gene would be sufficient to obtain reasonable quantities of the product. However, experience has shown that continuous transcription and translation of a cloned gene drains energy from essential cell functions and slows cell growth. With an inducible system, the expression of a cloned gene can be confined to a specific period in the growth cycle of the microorganism. On these grounds, for optimal protein production, the process should consist of two separate stages. First, the cells are grown under optimum conditions to a relatively high cell density. Second, depending on the nature of the promoter that drives the cloned gene, transcription is induced either by shifting the temperature or by adding a chemical inducer, such as IPTG, to the medium.

A two-stage system is not easy to implement in a large bioreactor (>100 liters) because it is technically very difficult either to raise the temperature quickly, typically from 30 to 42°C, or to ensure that the chemical inducer is

rapidly and evenly mixed in a large vessel. Moreover, many chemical inducers, such as IPTG, are too expensive to be used on a large scale. However, as discussed below, this problem can be overcome by using two connected bioreactors (two-stage fermentation) so that the cells are grown in the first vessel and the induction is carried out in the second. Under these conditions, growth and induction are separately optimized, thereby increasing the overall amount of product formed per unit of time (productivity) of the fermentation.

Two-Stage Fermentation in Tandem Airlift Reactors

E. coli NM989, which carries the gene encoding the enzyme T4 DNA ligase under the transcriptional control of the p^L promoter and a temperature-sensitive *c*I repressor, was grown and induced in a two-stage airlift bioreactor. In this bacterial strain, the T4 DNA ligase gene was integrated in the chromosomal DNA, a location obviating any potential problems of plasmid instability that might occur during extended fermentation. The growth stage was carried out at 30°C in an external-loop airlift bioreactor that had a 10-liter working volume. The T4 DNA ligase gene was not expressed under these conditions. A second external-loop airlift bioreactor, with a working volume of approximately 5 liters at 42°C, was used for the induction stage (Fig. 17.10). The two bioreactors were linked by a transfer tube with a pump that controlled the continuous flow of cell suspension from the growth stage bioreactor (first) into the induction stage vessel (second). In addition, cells suspended in culture medium were removed from the induction stage bioreactor at a specific rate and prepared for downstream processing.

The maximal specific growth rate (μ_{max}) of the microbial culture was approximately 0.66 reciprocal hour in the first bioreactor and 0.54 reciprocal hour in the second. These values correspond to cell doubling times of 63 and 77 minutes, respectively. Fresh medium was continuously added at a rate of 2 liters per hour to the growth stage fermenter, and cell suspension (effluent) was simultaneously removed from the induction stage fermenter at the same rate. As a consequence of the liquid volumes of the two bioreactors, an average cell spent about 5 hours in the growth stage bioreactor and 2 hours in the induction stage bioreactor. The different residence times in the two phases of this fermentation process were necessary to optimize the number of cells produced and the yield and stability of the T4 DNA ligase. Generally, residence times can be altered as required by adjusting the relative working volumes in the two fermenters of a two-stage system and by adjusting the volumetric rate of input of nutrients into the first bioreactor.

The double-external-loop design of the airlift fermenter (Fig. 17.10) used in this work facilitated the adjustment of the working volumes of the two fermenters relative to one another. It also added versatility to the system, so that it was possible to obtain a variety of different growth conditions for different populations of recombinant cells. For T4 DNA ligase production, the best results were obtained when approximately 33 mL of cell suspension was transferred every minute from the growth stage bioreactor to the induction stage bioreactor. Because this amount of cell suspension was equivalent to only 0.67% of the volume of the induction stage bioreactor, the incoming cells underwent a virtually instantaneous temperature shift from 30 to 42°C. Nutrients in a concentrated form were added at a specific rate to the induction stage bioreactor throughout the

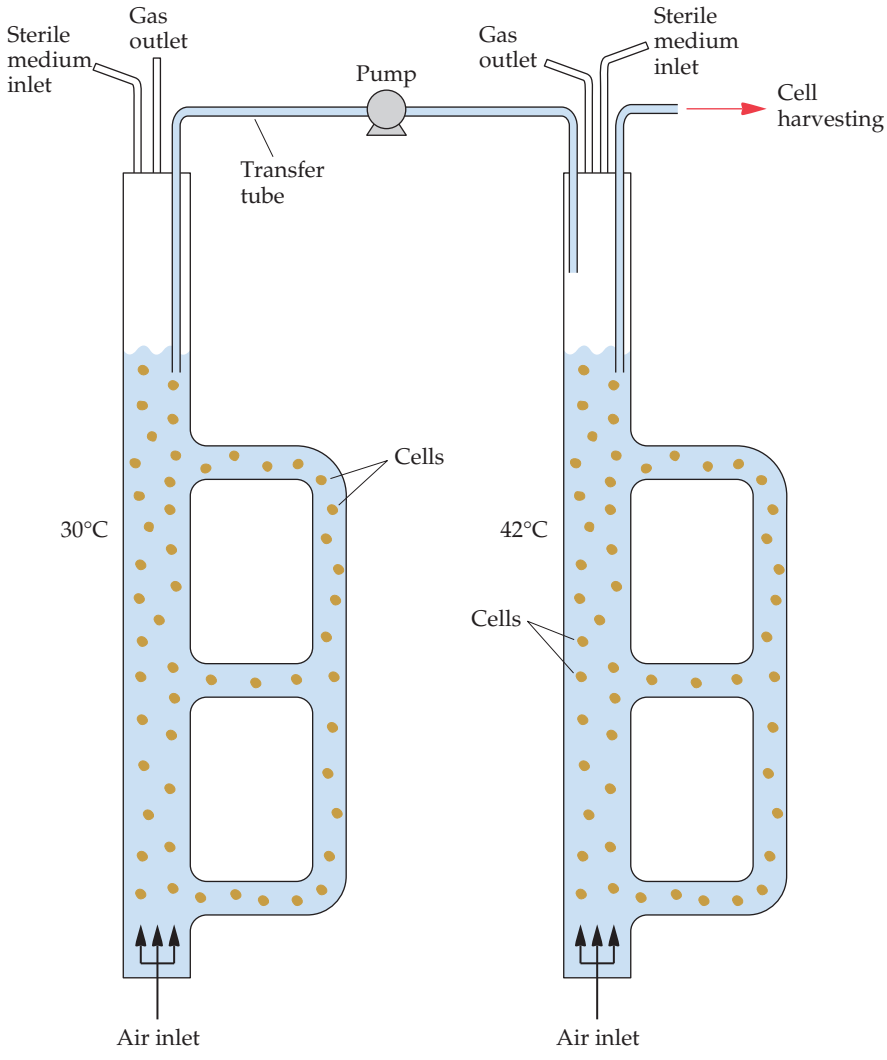


FIGURE 17.10 Two-stage airlift reactor used for the temperature-dependent induction of a protein product. Cells from the growth stage (left) fermenter, which is at 30°C, are pumped into the induction stage fermenter (right), which is at 42°C. Each bioreactor has a double external loop that is fitted with valves. By changing the valve settings, working volumes best suited for different fermentations can be created.

fermentation to keep the cells in this bioreactor in log phase. This action prevented the T4 DNA ligase from being degraded by the proteolytic enzymes that are normally synthesized during the deceleration and stationary phases.

With this continuous two-stage bioreactor, induced *E. coli* NM989 can be grown to a density of approximately 4 grams (dry weight) of cells per liter of culture. After induction, about 4% of the cell protein is T4 DNA ligase, an amount that corresponds to approximately 25,000 units of enzyme activity per gram (dry weight). This process can produce approximately 100,000 units of enzyme activity per liter of culture, or about 4,800,000 units per day. Assuming that it is possible to recover about 20% of the initial activity following purification of the enzyme and that the enzyme sells for about $0.25 per unit, then the final yield of purified

enzyme in 1 day will be worth about $240,000. Although these calculations do not consider all of the costs that go into the production of protein from a genetically engineered microorganism, it is clear that for highly valued products, small to moderate-size continuous fermentation systems can generate significant returns on the initial capital investment.

Two-Stage Fermentation in a Single Stirred-Tank Reactor

The tripartite fusion protein AGβ-Gal, which is used for immunological assays, was produced on a large scale in a single STR. The gene encoding AGβ-Gal was constructed by recombinant DNA techniques and encodes the five immunoglobulin G-binding regions of *Staphylococcus aureus* protein A, two immunoglobulin G-binding regions from *Streptococcus* sp. strain G148 protein G, and β-galactosidase from *E. coli*. The synthetic AGβ-Gal gene was placed under the control of the bacteriophage λ p^R promoter, which is regulated in the same manner as the p^L promoter; cloned into a plasmid that carries the gene for ampicillin resistance; and introduced by transformation into *E. coli*. The strain with the AGβ-Gal plasmid carries a second plasmid that has the genes for a temperature-sensitive *c*I repressor protein and a kanamycin resistance gene.

A 5-liter volume of these cells was grown in an STR at 30°C in the presence of both ampicillin and kanamycin—to provide selective pressure for the retention of both plasmids—and then used to initiate growth without antibiotics at 30°C in a 45-liter STR. The cell suspension in the 45-liter fermenter in turn served as an inoculum for a 600-liter STR, where the cells were grown at 30°C without antibiotics (Fig. 17.11). In general, to keep the cost of the process to a minimum, antibiotics are not added to large-scale microbial cultures. When the cell density in the 600-liter bioreactor reached the equivalent of about 4 grams (dry weight) per liter of growth medium, the temperature was shifted from 30 to 40°C to induce the expression of the AGβ-Gal protein. Under these conditions, it takes about 1 hour to reach 40°C. A temperature of 40°C rather than 42°C was chosen because the lower temperature was found to yield the same level of AGβ-Gal protein while allowing the cells to grow for a longer time. In other words, the low-temperature (40°C) conditions yielded a larger amount of protein product.

The specific activity of the AGβ-Gal protein increased for 2 hours after the initiation of the induction and then decreased. This decrease in activity was probably due to the synthesis of proteases by cells that had entered the deceleration and stationary phases of the growth cycle. In addition, 50% of the cells had lost their plasmids after growth for 4 hours at 40°C. These problems notwithstanding, after 4 hours at 40°C, the AGβ-Gal protein was approximately 20% of the dry weight of the total biomass. Considering the very high yield of the target protein that was produced by this strategy, it is probably not necessary to integrate the genes for the AGβ-Gal protein and the *c*I repressor into the chromosomal DNA of the *E. coli* host cell in an effort to increase the final yield.

Batch versus Fed-Batch Fermentation

In some instances, a simple fed-batch strategy can be used to produce both a high cell density and a high level of expression of a target protein (Fig. 17.12). For example, a plasmid carrying a gene encoding a hybrid protein that includes the insulin B peptide under the control of the *E. coli trp* promoter was introduced into a *trp*-minus mutant strain of *E. coli* that cannot

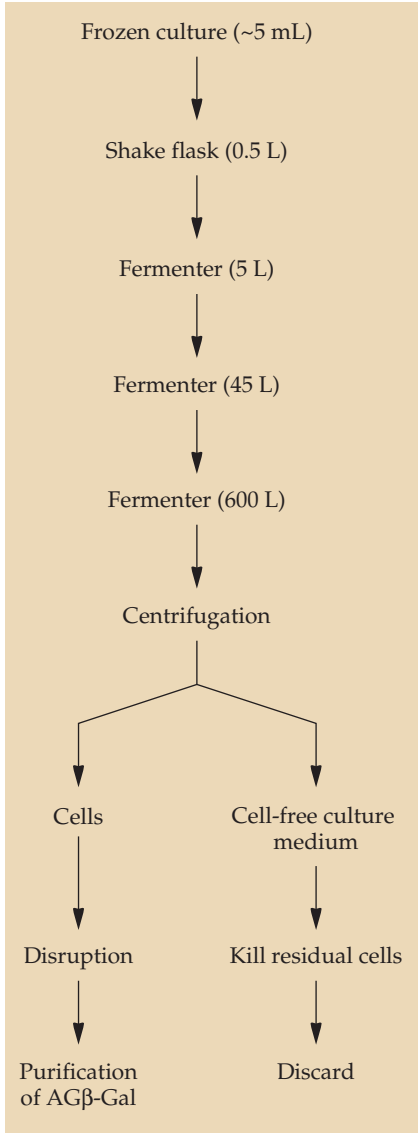


FIGURE 17.11 Scheme for the large-scale production of the protein AGβ-Gal. The volumes in parentheses indicate how much culture medium was used at each step of the process. The culture medium occupied about 60 to 75% of the total volume of each of the bioreactors (fermenters).

synthesize tryptophan, and the transformant was grown in media containing various amounts of tryptophan. At high levels of tryptophan, the synthesis of the target protein was repressed. However, after consumption of the tryptophan in the medium by the growing cells, synthesis of the target protein was induced. With this system, the addition of tryptophan to the medium resulted, in batch cultures, in increases in the amounts of both biomass and target protein produced. However, fed-batch fermentation was more effective than batch fermentation with or without added tryptophan (Table 17.3).

In another experiment, IFN-γ was produced in *E. coli*. Expression of the IFN-γ gene was controlled by the p^L promoter regulated by the tempera-

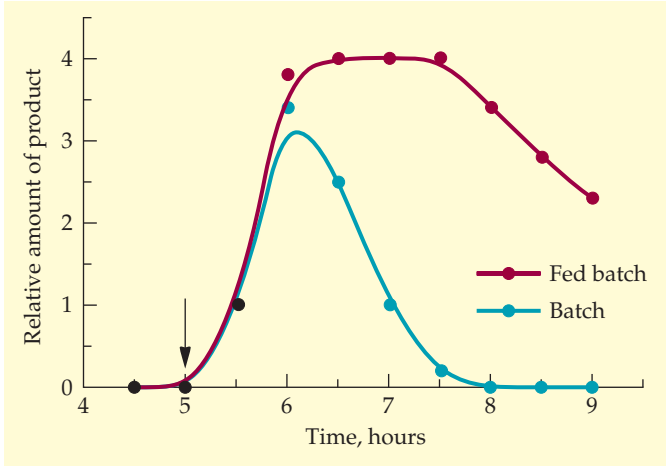


FIGURE 17.12 Schematic representation of the amount of foreign protein produced as a function of time by recombinant *E. coli* following induction (arrow) in the mid-log phase of growth. Prior to induction, no foreign protein is synthesized. After induction in batch mode, in the absence of additional nutrients, the cells soon enter stationary phase and synthesize proteases that degrade the foreign protein product. After induction in fed-batch mode, the added nutrients ensure that the cells remain in log phase for an extended time and do not produce any proteases until 1.5 to 2 hours later than the cells in batch mode; therefore, the foreign protein is more stable and is easier to recover. In addition, the provision of nutrients in fed-batch mode makes it less likely that plasmids carrying foreign genes will be lost than in batch mode. The time represents the number of hours from the start of the fermentation.

ture-sensitive *c*I repressor. Cells that contained this construct on a plasmid were grown in either batch or fed-batch mode (with both stepwise and constant-rate medium-feeding strategies). In the fed-batch mode, the addition of growth medium was carried out simultaneously with the temperature induction of the p^L promoter in the late exponential growth phase. Use of the fed-batch mode resulted in a significant increase in the length of the cell growth phase following induction. This fed-batch strategy enabled researchers studying this system to achieve a cell biomass that was 5-fold higher and a final IFN-γ concentration that was 23-fold higher than they were able to achieve in batch culture.

Another group studied the fermentation conditions that yielded the optimum expression of a monoclonal antibody (Fab) fragment directed

TABLE 17.3 Comparison of batch and fed-batch fermentations for the production of a fusion protein including the insulin B peptide

Product	Yield in fermentation system:		
	Batch	Batch + Trp	Fed batch
Biomass (g [DW]/liter)	6.7	12	20
Fusion protein ÷ total protein (%)	4.6	7.9	11
Total amount of fusion protein (g/liter)	0.17	0.53	1.21
Plasmid-bearing cells (%)	86	62	90

Adapted from Gosset et al., *Appl. Microbiol. Biotechnol.* **39:**541–546, 1993. DW, dry wieght. In the "Batch + Trp" fermentation, 0.1 g of tryptophan was added. In the fed-batch fermentation, 0.1 g of tryptophan was added every 2 hours, for a total of five times during the course of the 10-hour fermentation. A larger amount of tryptophan added to the batch fermentation did not increase the amount of either biomass or target protein produced.

against tetanus toxoid under the transcriptional control of the *E. coli lac* promoter. The plasmid construct included signal sequences that were inserted immediately upstream of the antibody (light and heavy) genes to target the antibody fragments to the *E. coli* periplasm. In this case, inexpensive lactose, rather than the considerably more expensive IPTG, could be used to induce the expression of the monoclonal antibody fragment gene, provided that the host strain of *E. coli* was Lac^+ and therefore able to take up lactose from the medium and convert it to glucose and galactose. In fact, allolactose, an isomer of D-lactose (and not D-lactose itself), is the actual inducer of the *E. coli lac* operon and is formed only if the host strain of *E. coli* contains a small amount of β-galactosidase (see chapter 6). An important facet of these experiments is that lactose not only acts as an inducer, but also is metabolized, providing an additional carbon source for the Fab-producing *E. coli* cells, thereby supporting both cell growth and product accumulation (Fig. 17.13). Again, in these experiments, fed-batch fermentation was clearly superior to batch fermentation, yielding both a larger amount of cell biomass and a greater concentration of Fab fragment.

Fed-batch fermentation strategies have also been successfully employed in the production of nonprotein products, such as poly(3-hydroxybutyrate) and poly(3-hydroxybutyrate-co-3-hydroxyvalerate), biopolymers with plastic-like properties (see chapter 13). In this case, the *E. coli* host cells carried a plasmid that contained polyhydroxyalkanoate biosynthesis genes from the bacterium *Alcaligenes latus*. By using an inexpensive growth medium, such as whey (a waste by-product of cheese making that consists mainly of lactose), to produce these polymers, it is hoped that a commercially viable product can be produced on a large scale.

Another group reported that instead of a single introduction of lactose in a fed-batch fermentation, to produce active protein fragments of human apolipoprotein(a), a strategy that employed continuous induction with lactose beneficially influenced the expression of the target protein. With a 75-liter fermenter, using lactose as the sole feed was not efficient for cell growth, presumably because the host strain of *E. coli* was unable to metabolize galactose. However, with a 1:50 ratio of lactose to glycerol, the target protein reached 16% of the total cellular protein. It remains to be seen whether continuous lactose induction will benefit the expression of proteins other than human apolipoprotein(a) and to what extent this process can be scaled up.

Harvesting Microbial Cells

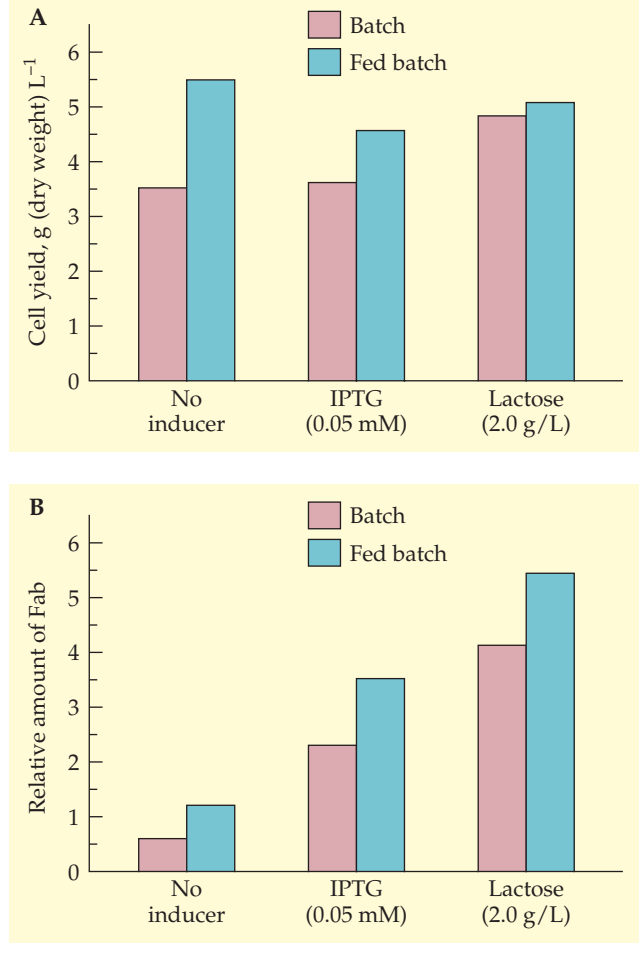
The first step in the process of purification of a product that is synthesized during microbial fermentation is the separation of the cells from the culture medium. Recombinant and native microbial cells can both be harvested with the same type of equipment. However, as a consequence of physiological changes, such as alterations in cell size and the production of extracellular polysaccharides, conditions that have been established for nontransformed cells may not be optimal for recombinant cells expressing a foreign protein.

For large volumes, either high-speed centrifugation, which is the current method of choice, or membrane microfiltration is used to separate cells from the growth medium. High-speed semicontinuous centrifuges have been specifically designed for the harvesting of microbial cells. The cell suspension is continuously fed into a running centrifuge, and the cells are

concentrated within it while the clarified medium is collected in an external container. When the centrifuge chamber is full of packed cells, the run is stopped and the cells are removed. The need to stop and then restart the procedure periodically, especially when large volumes are being processed, can be a major inconvenience. Furthermore, the cost of both the equipment and the power to run it, the release of the microorganisms into the air as aerosols during the harvesting procedure, and the difficulty in removing all the microbial cells from the spent medium also limit the use of this separation procedure.

Membrane filtration is an alternative method of separating the cells from the culture medium. Unfortunately, with traditional (dead-end) filtration, the microbial cells accumulate on the surface of the polymeric mem-

FIGURE 17.13 Induced recombinant protein production with batch and fed-batch *E. coli* cells grown to late logarithmic phase. The cells were either not induced, induced with 0.05 mM IPTG, or induced with 2.0 grams of lactose per liter. **(A)** In batch culture, the cells that were induced with lactose, which acts as both an inducer and a growth substrate, grew to a significantly greater extent than either the noninduced or the IPTG-induced cells. With fed-batch cells, the lesser extent of growth of induced versus noninduced cells probably reflects the resources that the induced cells direct to the synthesis of the Fab antibody fragment. **(B)** Under both batch and fed-batch conditions, more of the Fab antibody fragment was produced when the cells were induced with lactose than with IPTG. Fab antibody fragment synthesis in the absence of inducer represents incomplete repression of the *lac* promoter.



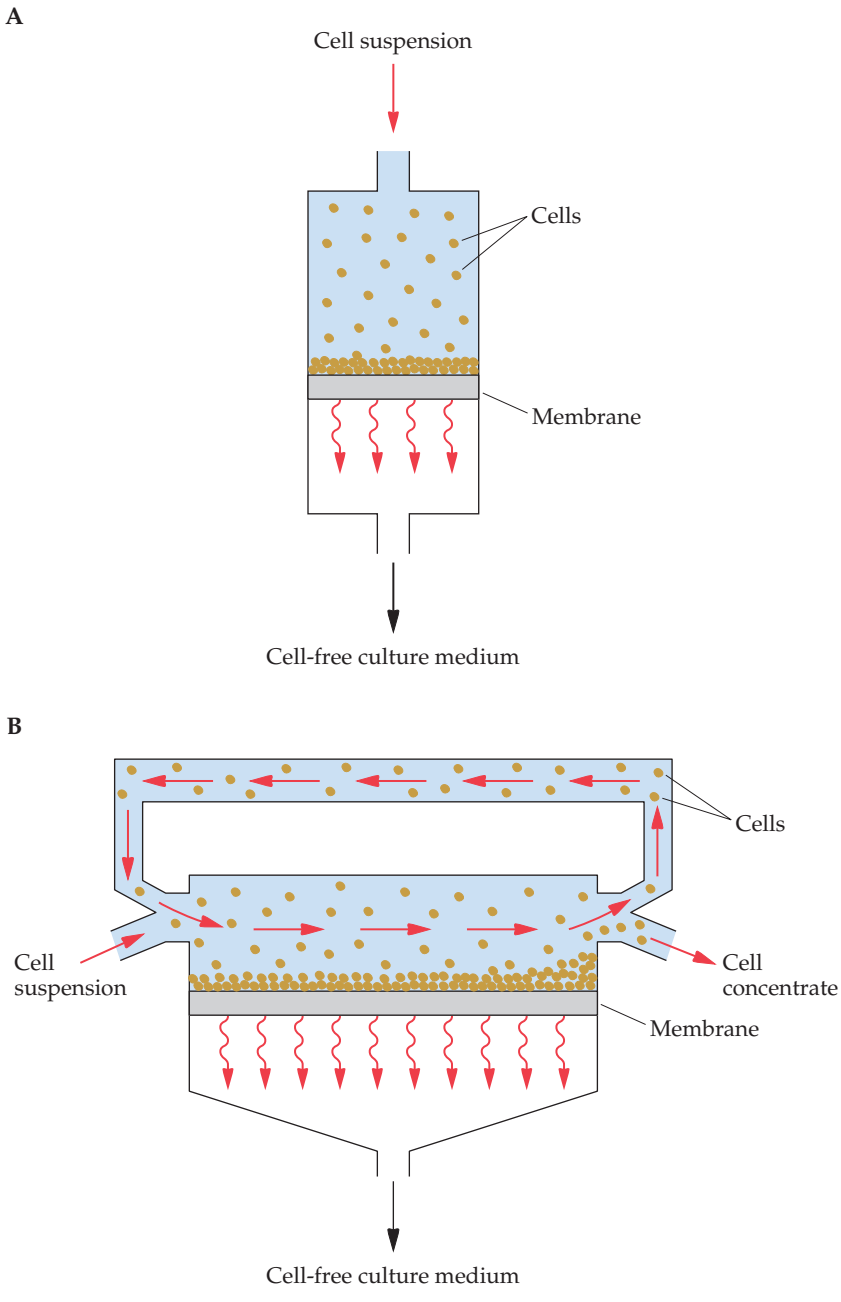


FIGURE 17.14 Membrane filtration systems for concentrating microbial cells. **(A)** Dead-end filtration; **(B)** cross-flow filtration. The arrows within each unit show the direction of the liquid flow.

brane filter. Consequently, the flow rate of spent medium through the membrane decreases rapidly (Fig. 17.14A). Increasing the pressure on the membrane enhances the flow for a short time; however, the cells still accumulate on the surface of the membrane and may even form a more compact and less permeable layer as a result of the pressure.

An alternative filtration technique entails passing the cell suspension at a high speed across the surface of the membrane (cross-flow filtration) (Fig. 17.14B). Under these conditions, only a very small fraction of the

circulating liquid actually goes through the membrane in any one pass. The remaining cell suspension acts to sweep the membrane clean of accumulated cells, so that the rate of liquid flow through the membrane does not decrease as rapidly as it does in dead-end filtration. In both dead-end and cross-flow filtration, the average pore size of the membranes is 0.2 to 0.45 μm. After many cycles in a cross-flow filtration system, almost all of the culture medium will have passed through the membrane. The use of cross-flow filtration is generally limited to laboratory-scale operations. Most industrial-scale operations still rely on centrifugation.

The next step in the purification process depends on both the nature and the location of the product. If the final product is a protein and it is in the culture medium, the medium is concentrated, often by ultrafiltration, and the target protein is purified by column chromatography or other standard procedures. If the product is a low-molecular-weight compound in the culture medium, it can be purified by the appropriate extraction procedures. Finally, if the product is in the cellular fraction, the cells must be disrupted (lysed) before the steps leading to product purification are initiated.

Disrupting Microbial Cells

A large number of chemical, biological, and physical methods have been developed for disrupting microbial cells. All of these procedures represent a compromise, because they must be vigorous enough to break the microbial cell walls yet gentle enough to ensure that the protein product is not denatured. There is no single set of conditions for cell wall lysis because the cell walls of diverse microbial species are composed of different polymers.

- In gram-positive bacteria, the cell wall is external to the cytoplasmic membrane and consists of a thick peptidoglycan layer of *N*-acetylglucosamine and *N*-acetylmuraminic acid residues cross-linked by oligopeptides.
- The cell wall of gram-negative bacteria has an outer membrane, a thin peptidoglycan layer, and a cytoplasmic membrane.
- The yeast cell wall is composed of a thick layer of partially phosphorylated mannans and β-glucan.

Cell wall composition and strength depend on culture conditions, the cellular growth rate, the phase of the growth cycle when the cells are harvested, how the concentrated cells are stored, and whether the isolated microorganism was expressing a cloned gene. All of these factors affect the cells' susceptibility to disruption.

The chemical methods that disrupt microbial cell walls include treatment with alkali, organic solvents, or detergents. If the protein product is stable at pH values from about 10.5 to 12.5, then bacterial cell lysis can easily be carried out on a large scale at low cost. For example, recombinant human growth hormone is efficiently released from *E. coli* by treatment with sodium hydroxide at pH 11. Few, if any, viable cells remain after alkali treatment, which obviates concerns about the inadvertent release of a genetically engineered microorganism from a production facility. Treatment with an organic solvent is a simple and inexpensive way to disrupt cells and has been used for the isolation of enzymes from yeasts. However, preliminary tests must be run to make sure that the proposed treatment does not denature the target protein. Detergents permeabilize bacterial cells by solubilizing cell membranes and membrane proteins. As a consequence of

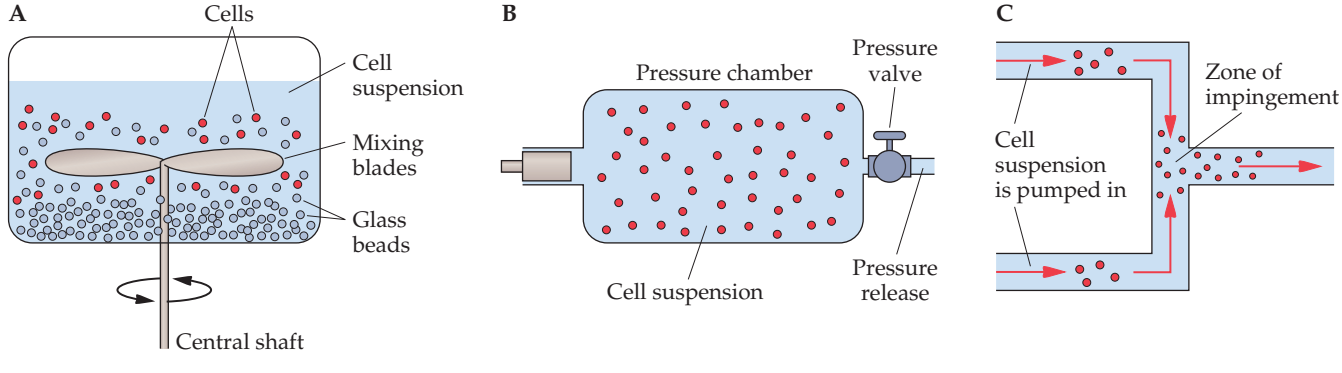
this activity, holes are formed, and proteins and other molecules are released from the cells. Unfortunately, detergents are expensive, frequently denature the protein product, and are often retained as contaminants throughout the purification process.

The major biological method for disrupting microbial cells is enzymatic lysis. For example, the cell walls of gram-positive bacteria are readily hydrolyzed by the enzyme lysozyme, which is isolated from egg whites; the cell walls of gram-negative bacteria are hydrolyzed by lysozyme and the metal-chelating agent ethylenediaminetetraacetic acid (EDTA); and the cell walls of yeasts are hydrolyzed by combinations of one or more of the following enzymes: β-1,3-glucanase, β-1,6-glucanase, mannanase, and chitinase. Enzymatic treatments are highly specific, and the conditions for lysis are mild. Currently, cost considerations limit the use of enzymes as cell lysis agents. However, the use of genetically engineered microorganisms for large-scale production of the enzymes that attack cell walls should make them less expensive.

Microbial cells can be physically disrupted either by nonmechanical methods, which include osmotic shock and repeated cycles of freezing and thawing, or by mechanical procedures, such as sonication, wet milling, high-pressure homogenization, and impingement. Generally, after treatment by a nonmechanical method, many of the cells remain intact. In contrast, mechanical disruption is highly efficient, which makes it the preferred choice. A sonicator that generates high-pressure sound waves that cause cell disruption by shear and cavitation (production of internal holes) is generally useful for small volumes.

Wet milling is quite commonly used for disrupting large quantities of cells (Fig. 17.15A). A concentrated cell suspension is pumped into the chamber of a high-speed agitator bead mill that is filled with an inert abrasive material, such as small glass beads (<1 mm in diameter) and is fitted with a central shaft that has a number of attached blades. When the device is turned on and the blades are put in motion, most of the cell disruption occurs as a consequence of the shear forces generated by the high-speed motion of the glass beads. Optimized cell disruption depends on both the number and configuration of the agitator disks, the agitator speed, the size of the glass beads, the number of glass beads, the cell concentration, the geometry of the grinding chamber, and the temperature. Bead mills have been successfully used to disrupt a wide range of different kinds of

FIGURE 17.15 Schematic representation of three methods of mechanical cell disruption of microbial cells. **(A)** Wet milling; **(B)** high-pressure homogenization; **(C)** impingement.



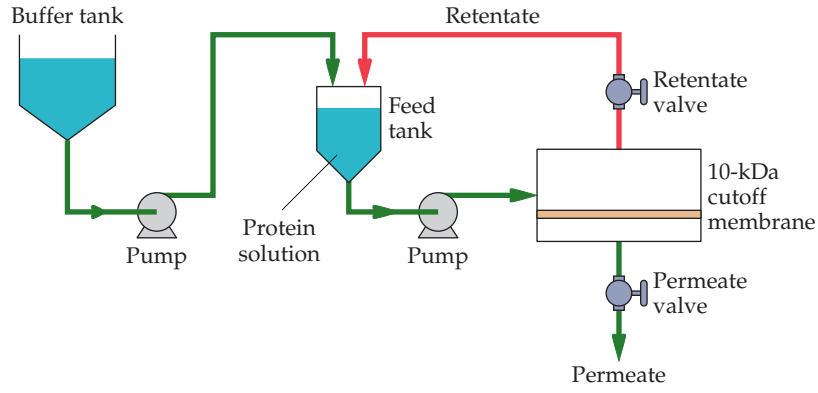


FIGURE 17.16 Schematic representation of large-scale ultrafiltration dialysis of a protein solution. The solution from the feed tank is pumped across the ultrafiltration membrane, with only a small fraction of the solution actually passing through the membrane and the rest being used to sweep the membrane clean of protein. The volume that passes through the membrane is matched by an equal volume of buffer added to the system. A membrane with a nominal molecular mass cutoff below the size of the target protein is used. For example, a 10-kDa cutoff membrane might be used to retain a 30-kDa protein while removing salt from the solution. Following dialysis, the target protein is found, in a dilute solution, in the feed tank. This solution may be concentrated by ultrafiltration using an identical setup except that no buffer tank is used. The arrows indicate the direction of liquid flow. The relative volume of permeate (liquid that passes through the membrane) compared with the retentate (retained liquid) is controlled by adjusting the permeate and retentate valves; this helps to keep the membrane relatively free of protein that might otherwise clog its pores.

microbial cell types and can readily break recombinant, as well as nonrecombinant, cells.

In the high-pressure homogenization process (Fig. 17.15B), concentrated cells are pumped into a valve assembly under high pressure, and the pressure is then rapidly decreased, causing the cells to lyse. This process can be customized for different microorganisms and protein products by changing the operating pressure, the design of the valve, the temperature of the cell suspension, or the number of times the cell mass is treated.

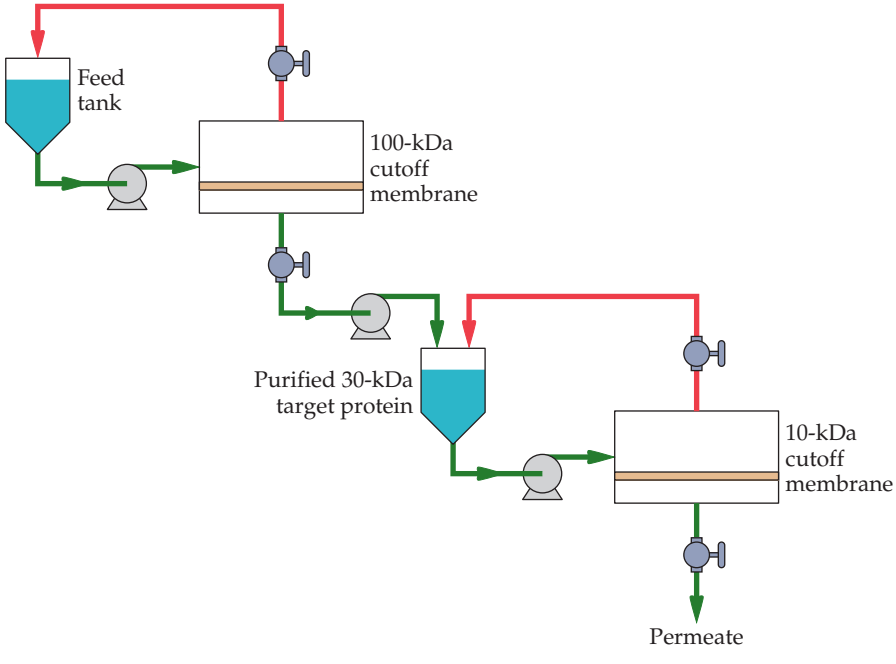
Impingement (Fig. 17.15C) is a cell disruption procedure in which a high-velocity stream of suspended cells under pressure hits either a stationary surface or a second fluid stream of suspended cells. The forces that are created at the point of contact disrupt the cells. With a device called a Microfluidizer, for example, two parallel streams of *E. coli* cells in suspension are directed toward one another. With this device, a high percentage of the cells are disrupted by a single passage through the unit. However, additional passages may be required for complete breakage of other cell types. Unlike high-pressure homogenizers and high-speed agitator bead mills, which generally require highly concentrated cell suspensions, this device can be used to disrupt cells in either dilute or concentrated preparations. The activities of cellular proteins are not significantly impaired by the technique. When the cell suspension is pretreated with low levels of lysozyme and then disrupted in the Microfluidizer at much lower than normal pressure and fluid velocity, the activities of labile proteins, which might be otherwise inactivated by the high pressure, are retained.

Downstream Processing

After cell disruption, cell debris is removed by either low-speed, high-capacity centrifugation or membrane microfiltration. The protein product is precipitated from the crude lysate, the clarified lysate, or the cell-free culture medium with organic solvents (alcohol or acetone) or ammonium sulfate. Under these conditions, the target protein is usually enriched approximately two- to fivefold. Unfortunately, the cost of the precipitant can add significantly to the cost of the process. Alternatively, the crude protein mixtures are concentrated and fractionated by cross-flow ultrafiltration through membranes that have a smaller average pore size than those used for either cell concentration or debris removal (Fig. 17.14B). This approach can be used with volumes ranging from 1 liter to several thousand liters and can be performed continuously, which means that large-volume systems are not required. Depending on the size and properties of the target protein, this method can yield 10- to 100-fold enrichment.

On a large scale, it is impractical to remove small molecules, such as salts or organic solvents, from protein solutions by conventional laboratory procedures. Consequently, the same apparatus that is used to concentrate proteins by ultrafiltration has been developed for the large-scale dialysis of proteins (Fig. 17.16). In addition, by using two different-size membranes,

FIGURE 17.17 Schematic representation of large-scale protein purification using ultrafiltration membranes. In the example shown, two membranes are used sequentially, a 100-kDa cutoff membrane and a 10-kDa cutoff membrane, in order to purify a 30-kDa protein. The solution, which contains a protein mixture, is pumped through the 100-kDa cutoff membrane, with the larger proteins being retained and the smaller proteins, including the 30-kDa target protein, passing through the membrane. This protein solution is next pumped through a 10-kDa cutoff membrane, with the 30-kDa target protein being retained. The arrows indicate the direction of the liquid flow. The symbols are identical to those found in Fig. 17.16.



large volumes of target proteins with different molecular masses may be fractionated by ultrafiltration (Fig. 17.17).

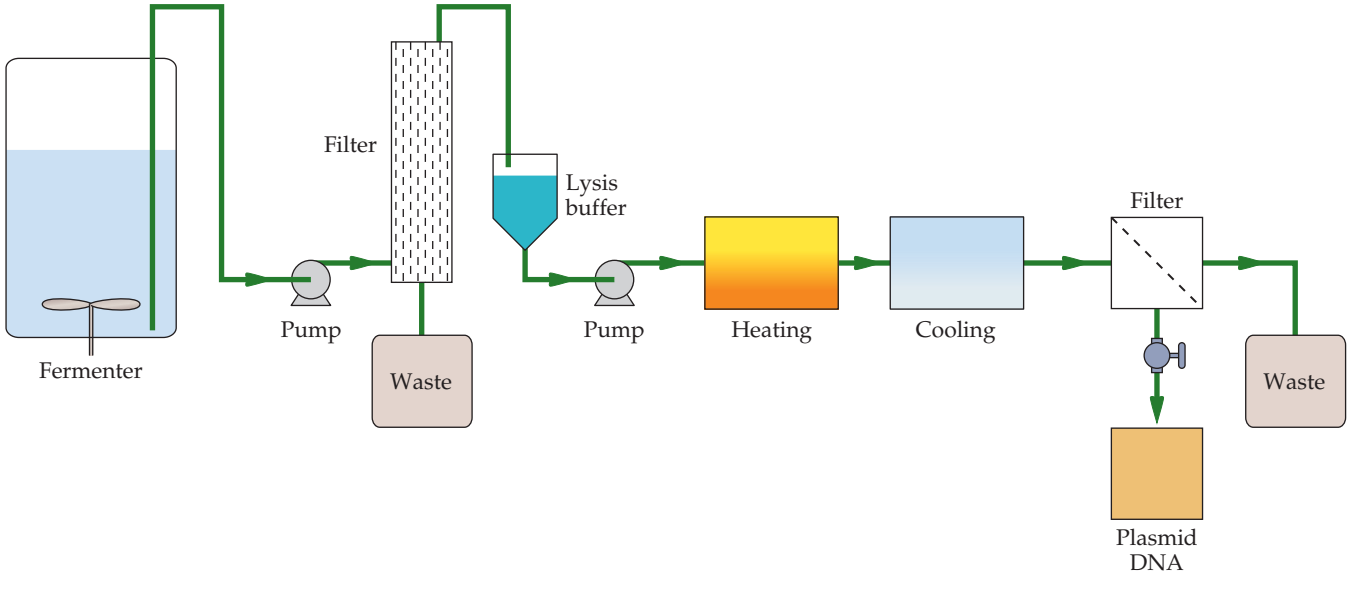
The required degree of purity of the final protein product depends on its end use. In some cases, such as enzymes for use with laundry detergent, crude preparations are satisfactory, but for other products, such as pharmaceutical proteins, additional purification procedures are required.

A number of proteins that are overproduced intracellularly are confined to insoluble particles (inclusion bodies) within the bacterium. After disruption, such inclusion bodies can be readily separated from the bulk of the remaining cell components. Initially, researchers found it difficult to solubilize inclusion bodies without irreversibly denaturing the protein, but protocols have now been devised to renature the proteins found in inclusion bodies. Of course, these additional steps increase the cost of the purification process.

Protein Solubilization

In some instances, overexpression of a target protein can result in the production of both soluble and insoluble forms, which complicates the purification process. For example, when human insulin-like growth factor I (IGF-I), a 7.6-kilodalton (kDa) peptide, was expressed in *E. coli* cells, approximately 90% of the recombinant protein was localized in the *E. coli* periplasm (soluble and insoluble) and about 10% was found in the external medium (soluble). To recover both forms of IGF-I from the periplasm, high concentrations of urea and dithiothreitol at alkaline pH were added to solubilize the insoluble forms of the peptide in situ. This treatment kills but does not lyse the cells. Consequently, the cytoplasmic proteins remain within the cells. The solubilization procedure produces a highly viscous solution, which precludes removing the cells and cell debris by centrifugation. Instead, an aqueous two-phase liquid extraction procedure that separates soluble and insoluble materials was developed for this purpose. Both

FIGURE 17.18 Schematic representation of a process for the continuous thermal lysis of *E. coli* cells for the large-scale preparation of purified plasmid DNA.



the in situ solubilization and aqueous two-phase liquid extraction procedure are highly efficient; 80 to 95% of the IGF-I is recovered from 10-, 100-, or 1,000-liter fermentations by these methods.

Large-Scale Production of Plasmid DNA

Gene therapy and genetic immunization protocols are being tested in a large number of clinical trials. Many of these trials require plasmids as vectors to deliver the remedial DNA to the patient. As these procedures become more routine and are extended to many more patients, it will be necessary to produce plasmid DNA, initially in the 5- to 20-kilobase size range, on a large scale and in a highly purified form, i.e., pharmaceutical grade.

A number of factors, including the choice of a host strain, such as *E. coli* K-12, that is safe and well characterized; growth conditions; defined medium; and the purification process that removes all of the undesired genomic DNA, RNA, proteins, lipids, and lipopolysaccharides, need to be considered for the large-scale production of plasmid DNA. Generally, to ensure that the plasmid DNA will be stable, a host strain should produce only low levels of nucleases. In addition, the presence of antibiotics in the culture medium should be avoided, since it is difficult to remove all traces of an antibiotic from the final preparation. For DNA isolation, the current method of choice for cell breakage is the alkaline-lysis procedure—treatment with 0.2 M NaOH and 1% sodium dodecyl sulfate—which breaks cells without disrupting plasmid DNA. The lysis solution must be added so that mixing is sufficient to lyse all of the cells and local pH extremes are avoided while the shear that is generated by the mixing does not damage the plasmid DNA. After cell lysis, the precipitates formed, which contain cell debris, denatured proteins, and nucleic acids, must be removed. This is most commonly done by centrifugation using fixed-angle rotors, a process which is difficult to scale up. After removal of the precipitated material, plasmid purification is best achieved by anion-exchange chromatography, followed by size exclusion (gel filtration) chromatography. In the industrial production of many recombinant proteins, high-pressure homogenizers are used to continuously disrupt cells. However, the use of such homogenizers typically results in the degradation of a large portion of the plasmid DNA as a consequence of the severe shear stress generated during cell breakage. To work out a process for the large-scale preparation of plasmid DNA, one group of researchers developed and optimized a unique continuous thermal-lysis protocol. In this procedure, cells were harvested after high-cell-density batch culture, filtered to remove most of the growth medium, and resuspended in lysis buffer containing lysozyme. After incubation at 37°C and gentle stirring for 20 minutes, the cells were heated to 70°C for 20 seconds and then filtered to remove the cell debris (Fig. 17.18). Using this procedure, researchers have reported obtaining 100 mg of high-quality plasmid DNA, free of contaminating chromosomal DNA, per liter of high-cell-density cell culture. The use of lysozyme is the most expensive component of this process. Nevertheless, as a consequence of the high plasmid yield, the avoidance of time-consuming and expensive centrifugation steps, the short processing time—17 liters of *E. coli* cells can be processed in ~45 minutes—and the ease with which the process can be scaled up, this approach is likely to provide a highly effective means of preparing large amounts of highly purified plasmid DNA.

SUMMARY

The large-scale production of genetically engineered organisms in industrial-size bioreactors (>1,000 liters) is not achieved merely by extrapolating directly from laboratory growth conditions (0.1 to 1.0 liter). The temperature, pH, rate and nature of mixing, oxygen demand for aerobic organisms, and nutrient levels must be taken into consideration when large bioreactors are being designed.

Microbial fermentations can be performed in several different ways. In batch fermentation, an inoculum of cells is added to fresh medium, and the fermentation is allowed to proceed without supplementation until the maximum amount of the target product is synthesized. Under these conditions, the cell culture passes through six phases of growth: lag, acceleration, log, deceleration, stationary, and death. Protein production is optimal during the log phase, whereas peak production of many low-molecular-weight products occurs during the stationary phase. It is important to monitor batch fermentations closely to ensure that the cells are harvested at the appropriate time. In fed-batch fermentation, growth medium is added at various intervals, usually to prolong the log phase of the fermentation process. Continuous fermentation entails adding fresh growth medium throughout the course of the fermentation and simultaneously removing cells and spent medium.

Each of these fermentation systems has disadvantages and advantages for large-scale production of recombinant products. Although continuous fermentation is still a relatively untried industrial process, the approach has some inherent benefits, making it likely that its use will become more widespread in the future.

One way to increase the amount of a recombinant protein is to grow transformants to as high a cell density as possible. The best way to obtain high-density cell cultures is to use a fed-batch fermentation strategy. Fed-batch fermentation has been directly compared with batch fermentation for the production of several different proteins, and in all cases examined, fed-batch fermentation has resulted in a higher yield of the target protein.

To obtain the largest amount of product in a given volume, it is helpful to attain high cell densities, to avoid the loss of the recombinant plasmid, to utilize *E. coli* cells that can become quiescent during foreign-gene expression, to secrete the foreign protein into the growth medium, and to avoid the formation of acetate in the medium.

There are three basic bioreactor configurations: STRs, bubble columns, and airlift reactors. Currently, STRs are used most frequently in industry, but interest in airlift bioreactors is increasing. In STRs, mixing is achieved by mechanical agitation. In airlift reactors, both aeration and mixing are performed by a gas, usually air, that is introduced through a sparger at the bottom of the vessel, with either internal baffles or external loops causing the fluid to circulate within the vessel. Bubble columns are similar to airlift reactors but lack the design features that cause the culture medium to circulate in the vessel. The problems of maintaining sterility, pH, temperature, and other fermentation parameters are overcome in different ways depending on the design of the bioreactor. Two-stage fermentation processes with either tandem airlift reactors or a single STR have been successfully used for the production of recombinant proteins.

When the product is present within the cells, the cells can be harvested by either centrifugation or cross-flow filtration and lysed chemically, enzymatically, or mechanically. The preferred forms of mechanical cell lysis include wet milling, high-pressure homogenization, and impingement. The product is then fractionated from the cell lysate. Ultrafiltration has been found to be an effective method for large-scale dialysis, concentration, and initial fractionation of proteins produced by recombinant organisms.

Scientists have begun to establish procedures for the large-scale isolation of plasmid DNA. These procedures must take into account the host cell and its growth and metabolism, plasmid size, cell lysis methods, and the complete removal of a number of potentially contaminating cell components.

REFERENCES

Aristidou, A. A., K. Y. San, and G. N. Bennett. 1995. Metabolic engineering of *Escherichia coli* to enhance recombinant protein production through acetate reduction. *Biotechnol. Prog.* **11:**475–478.

Bailey, J. E., and D. F. Olis. 1977. *Biochemical Engineering Fundamentals*. McGraw-Hill Book Co., New York, NY.

Charles, M. 1985. Fermentation scale-up: problems and possibilities. *Trends Biotechnol*. **3:**134–139.

Choi, J.-I., and S. Y. Lee. 1999. High-level production of poly(3-hydroxybutyrate-co-3-hydroxyvalerate) by fed-batch culture of recombinant *Escherichia coli*. *Appl. Environ. Microbiol*. **65:**4363–4368.

Datar, R. 1986. Economics of primary separation steps in relation to fermentation and genetic engineering. *Process Biochem*. **21:**19–26.

Donovan, R. S., C. W. Robinson, and B. R. Glick. 2000. Optimizing the expression of a monoclonal antibody fragment under the transcriptional control of the *Escherichia coli lac* promoter. *Can. J. Microbiol*. **46:**532–541.

Eiteman, M. A., and E. Altman. 2006. Overcoming acetate in *Escherichia coli* recombinant protein fermentations. *Trends Biotechnol*. **24:**530–536.

Engler, C. R. 1985. Disruption of microbial cells, p. 305–324. *In* C. L. Cooney, A. E. Humphrey, and M. Moo-Young (ed.), *Comprehensive Biotechnology*, vol. 2. Pergamon Press, Oxford, United Kingdom.

Giorgio, R. J., and J. J. Wu. 1986. Design of large scale containment facilities for recombinant DNA fermentations. *Trends Biotechnol*. **4:**60–65.

Gosset, G., R. de Anda, N. Cruz, A. Martínez, R. Quintero, and F. Bolivar. 1993. Recombinant protein production in cultures of an *Escherichia coli trp2* strain. *Appl. Microbiol. Biotechnol*. **39:**541–546.

Grund, G., C. W. Robinson, and B. R. Glick. 1991. Cross-flow ultrafiltration of proteins, p. 69–83. *In* M. D. White, S. Reuveny, and A. Shafferman (ed.), *Biologicals from Recombinant Microorganisms and Animal Cells: Production and Recovery*. Verlag Chemie, Weinheim, Germany.

Hägg, P., J. Wa de Pohl, F. Abdulkarim, and L. A. Isakson. 2004. A host/plasmid system that is not dependent on antibiotics and antibiotic resistance genes for stable plasmid maintenance in *Escherichia coli*. *J. Biotechnol.* **111:**17–30.

Hart, R. A., P. M. Lester, D. H. Reifsnyder, J. R. Ogez, and S. E. Builder. 1994. Large scale, in situ isolation of periplasmic IGF-I from *E. coli*. *Bio/Technology* **12:**113–117.

Khosla, C., J. E. Curtis, J. DeModena, U. Rinas, and J. E. Bailey. 1990. Expression of intracellular hemoglobin improves protein synthesis in oxygen-limited *Escherichia coli. Bio/Technology* **8:**849–853.

Kroner, K. H. 1986. Cross-flow filtration in the downstream processing of enzymes: current status. *Biotechnol. Forum* **3:**20–31.

Kroner, K. H., H. Nissinen, and H. Zeigler. 1987. Improved dynamic filtration of microbial suspensions. *Bio/Technology* **5:**921–926.

Lee, S. Y. 1996. High cell-density culture of *Escherichia coli*. *Trends Biotechnol*. **14:**98–105.

Levy, M. S., R. D. O'Kennedy, P. Ayazi-Shamlou, and P. Dunnill. 2000. Biochemical engineering approaches to the challenges of producing pure plasmid DNA. *Trends Biotechnol*. **18:**296–305.

Lim, H.-K., and K.-H. Jung. 1998. Improvement of heterologous protein productivity by controlling postinduction specific growth rate in recombinant *Escherichia coli* under control of the p^L promoter. *Biotechnol. Prog*. **14:**548–553.

Lim, H.-K., S.-G. Kim, K.-H. Jung and J.-H. Seo. 2004. Production of the kringle fragments of human apolipoprotein(a) by continuous lactose induction strategy. *J. Biotechnol.* **108:**271–278.

March, J. C., M. A. Eiteman, and E. Altman. 2002. Expression of an anaplerotic enzyme, pyruvate carboxylase, improves recombinant protein production in *Escherichia coli. Appl. Environ. Microbiol.* **68:**5620–5624.

McKillip, E. R., A. S. Giles, M. H. Levner, P. P. Hung, and R. N. Hjorth. 1991. Bioreactors for large-scale t-PA production. *Bio/Technology* **9:**805–812.

Mendoza-Vega, O., C. Hebert, and S. W. Brown. 1994. Production of recombinant hirudin by high cell density fed-batch cultivations of a *Saccharomyces cerevisiae* strain: physiological considerations during the bioprocess design. *J. Biotechnol*. **32:**249–259.

Merchuk, J. C. 1990. Why use airlift bioreactors? *Trends Biotechnol*. **8:**66–71.

Mukherjee, K. J., D. C. D. Rowe, N. A. Watkins, and D. K. Summers. 2004. Studies of single-chain antibody expression in quiescent *Escherichia coli. Appl. Environ. Microbiol*. **70:**3005–3012.

Park, T. H., J.-H. Seo, and H. C. Lim. 1991. Two-stage fermentation with bacteriophage λ as an expression vector in *Escherichia coli*. *Biotechnol. Bioeng*. **37:**297–302.

Paulson, D. J., R. L. Wilson, and D. D. Spatz. 1984. Cross-flow membrane technology and its applications. *Food Technol*. **December 1984:**77–87.

Prazeres, D. M. F., G. N. M. Ferreira, G. A. Monteiro, C. L. Cooney, and J. M. S. Cabral. 1999. Large-scale production of pharmaceutical-grade plasmid DNA for gene therapy: problems and bottlenecks. *Trends Biotechnol*. **17:**169–174.

Ramírez, D. M., and W. E. Bentley. 1995. Fed-batch feeding and induction policies that improve foreign protein synthesis and stability by avoiding stress response. *Biotechnol. Bioeng*. **47:**596–608.

Reuss, M. 1995. Stirred tank bioreactors, p. 207–255. *In* J. A. Asenjo and J. Merchuk (ed.), *Bioreactor System Design*. Marcel Dekker, Inc., New York, NY.

Riesenberg, D., and R. Guthke. 1999. High-cell-density cultivation of microorganisms. *Appl. Microbiol. Biotechnol*. **51:**422–430.

Robinson, D. K., C. P. Chan, C. Yu Ip, P. K. Tsai, J. Tung, T. C. Seamans, A. B. Lenny, D. K. Lee, J. Irwin, and M. Silberklang. 1994. Characterization of a recombinant antibody produced in the course of a high yield fed-batch process. *Biotechnol. Bioeng*. **44:**727–735.

Rowe, D. C. D., and D. K. Summers. 1999. The quiescent-cell expression system for protein synthesis in *Escherichia coli. Appl. Environ. Microbiol*. **65:**2710–2715.

Sauer, T., C. W. Robinson, and B. R. Glick. 1989. Disruption of native and recombinant *Escherichia coli* in a high-pressure homogenizer. *Biotechnol. Bioeng*. **33:**1330–1342.

Sayadi, S., M. Nasri, F. Berry, J. N. Barbotin, and D. Thomas. 1987. Effect of temperature on the stability of plasmid pTG201 and productivity of *xylE* gene product in recombinant *Escherichia coli*: development of a two-stage chemostat with free and immobilized cells. *J. Gen. Microbiol*. **133:**1901–1908.

Schleef, M. 1999. Issues for large-scale DNA manufacturing, p. 443–470. *In* H. J. Rehm and G. Reed (ed.), *Biotechnology: a Multi-Volume Comprehensive Treatise,* vol. 5a. *Recombinant Proteins, Monoclonal Antibodies, and Therapeutic Genes*. Wiley-VCH, New York, NY.

Schügerl, K., and A. Lübbert. 1995. Pneumatically agitated bioreactors, p. 257–303. *In* J. A. Asenjo and J. Merchuk (ed.), *Bioreactor System Design*. Marcel Dekker, Inc., New York, NY.

Schütte, H., and M.-R. Kula. 1990. Pilot- and process-scale techniques for cell disruption. *Biotechnol. Appl. Biochem*. **12:**599–620.

Seigel, R., and D. D. Y. Ryu. 1985. Kinetic study of instability of recombinant plasmid pPLc23trpAl in *E. coli* using two-stage continuous culture system. *Biotechnol. Bioeng*. **27:**28–33.

Siegel, M. H., H. Hallaille, and J. C. Merchuk. 1988. Air-lift reactors: design, operation, and applications. *Adv. Biotechnol. Processes* **7:**79–124.

Sletta, H., A. Tøndervik, S. Hakvåg, T. E. Vee Aune, A. Nedal, R. Aune, G. Evensen, S. Valla, T. E. Ellingsen, and T. Brautaset. 2007. The presence of N-terminal secretion signals leads to strong stimulation of the total expression levels of three tested medically important proteins during high-cell-density cultivations of *Escherichia coli*. *Appl. Environ. Microbiol.* **73:**906–912.

Strandberg, L., L. Andersson, and S.-O. Enfors. 1994. The use of fed batch cultivation for achieving high cell densities in the production of a recombinant protein in *Escherichia coli*. *FEMS Microbiol. Rev*. **14:**53–56.

Strandberg, L., K. Köhler, and S.-O. Enfors. 1991. Large-scale fermentation and purification of a recombinant protein from *Escherichia coli*. *Process Biochem*. **26:**225–234.

Strathman, H. 1985. Membranes and membrane processes in biotechnology. *Trends Biotechnol*. **3:**112–118.

Tanny, G. B., D. Mirelman, and T. Pistole. 1980. Improved filtration techniques for concentrating and harvesting bacteria. *Appl. Environ. Microbiol*. **40:**269–273.

Tutunjian, R. S. 1985. Scale-up considerations for membrane processes. *Bio/Technology* **3:**615–626.

Van Brunt, J. 1985. Scale-up: the next hurdle. *Bio/Technology* **3:**419–424.

Van Brunt, J. 1986. Fermentation economics. *Bio/Technology* **4:**395–401.

Wang, Z. W., Y. Chen, and Y.-P. Chao. 2006. Enhancement of recombinant protein production in *Escherichia coli* by coproduction of aspartase. *J. Biotechnol.* **124:**403–411.

White, M. D., B. R. Glick, and C. W. Robinson. 1995. Bacterial, yeast and fungal cultures: the effect of microorganism type and culture characteristics on bioreactor design and operation, p. 47–87. *In* J. A. Asenjo and J. Merchuk (ed.), *Bioreactor System Design*. Marcel Dekker, Inc., New York, NY.

Whitney, G. D., B. R. Glick, and C. W. Robinson. 1989. Induction of T4 DNA ligase in a recombinant strain of *Escherichia coli*. *Biotechnol. Bioeng*. **33:**991–998.

Yamane, T. 1995. Bioreactor operation modes, p. 479–509. *In* J. A. Asenjo and J. Merchuk (ed.), *Bioreactor System Design*. Marcel Dekker, Inc., New York, NY.

Zhu, K., H. Jin, Y. Ma, Z. Ren, C. Xiao, Z. He, F. Zhang, Q. Zhu, and B. Wang. 2005. A continuous thermal lysis procedure for the large-scale preparation of plasmid DNA. *J. Biotechnol.* **118:**257–264.

REVIEW QUESTIONS

1. What are the differences between batch, fed-batch, and continuous fermentations?

2. How has fed-batch fermentation been used to improve the production of the insulin B peptide, IFN-γ, and Fab fragment?

3. What parameters must be monitored and controlled in an optimized fermentation process?

4. What is the effect of a recombinant plasmid on the growth of microbial cells?

5. How does the mixing of a growing microbial culture affect the transfer of oxygen from the growth medium to the cells?

6. What are the advantages of a high cell density during a large-scale fermentation? What conditions lead to a high cell density during a large-scale fermentation?

7. What strategies can be employed to prevent acetate inhibition of the growth of recombinant *E. coli* strains?

8. What are the relative advantages and disadvantages of using an STR or an airlift fermenter?

9. Compare the growth and induction of a recombinant microbial culture using (1) two reactors in tandem and (2) a single reactor.

10. What is downstream processing?

11. What strategy would you use to purify a recombinant protein that is secreted into the growth medium?

12. What are the advantages and disadvantages of large-scale mechanical lysis of cells compared to chemical lysis?

13. How are cells mechanically disrupted using (1) wet milling, (2) high-pressure homogenization, and (3) impingement?

14. How are microbial cells concentrated after the fermentation stage of a biotechnological process? What are the advantages and disadvantages of these procedures?

15. How are large volumes of protein solutions partially purified using ultrafiltration?

16. How can the activity of an insoluble recombinant protein be recovered?

17. What factors should be considered for the large-scale isolation of plasmid DNA?

18. How can "quiescent *E. coli* cells" be engineered to produce large amounts of foreign protein?

19. What strategy would you use to ensure that a plasmid encoding a target protein is not lost during the large-scale growth of a recombinant bacterium?

20. What are some of the advantages of secreting a recombinant protein into the growth medium?