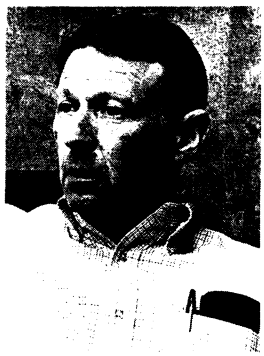
Лекция 10.
Генетическая инженерия

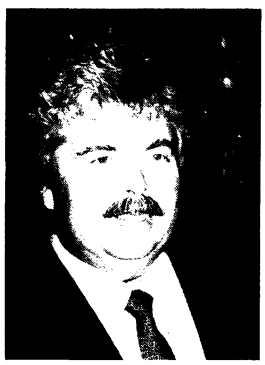
CHAPTER

Recombinant DNA Technology



Paul Berg

Stanley N. Cohen



Herbert Boyer

In our final chapter we describe a technology that in less than two decades has become fundamental to the advance of biochemistry. It helps to define present and future biochemical frontiers and illustrates many important principles of biochemistry. As the laws governing enzymatic catalysis, macromolecular structure, cellular metabolism, and information pathways continue to be elucidated, new research is directed at ever more complex biochemical processes. Cell division, the immune response, developmental processes in eukaryotes, vision, taste, oncogenesis, the cognitive processes in your brain as you read these words—all are orchestrated in an elaborate symphony of molecular and macromolecular interactions. As increasingly greater efforts are focused on understanding the biochemistry that underlies these processes, the real promise and implications of the biochemical journey begun in the nineteenth century become clear. Human beings not only can understand life, they can alter it.

The biochemical approach to understanding a complex biological process is to isolate and study the individual components in vitro with the goal of understanding the overall process in the whole organism. Perhaps the most fertile source for molecular insights into these processes lies in the cell's own information storehouse, its DNA. The sheer size of cellular chromosomes, however, presents us with an enormous barrier. How does one find and study a particular gene encoding a protein or RNA molecule with a molecular function we can only guess at, when that gene is only one of perhaps 100,000 genes scattered among the billions of base pairs that make up a mammalian genome? The answers began to appear in the mid-1970s.

Decades of advances in genetics, biochemistry, cell biology, and physical chemistry came together in the laboratories of Paul Berg, Herbert Boyer, and Stanley Cohen to yield techniques for locating, isolating, preparing, and studying small segments of DNA derived from much larger chromosomes. Taken together, these techniques are known as **DNA cloning.** DNA cloning has opened opportunities unimaginable just a few decades ago, including the identification and study of genes involved in almost every known biological process. These new methods are transforming basic research, agriculture, forensics, medicine, ecology, and many other fields, while at the same time presenting society with bewildering choices and serious ethical dilemmas.

Revolutionary as it is, this technology is grounded in the most fundamental biological and biochemical principles. The first two parts of this chapter outline these fundamentals, drawing on our understand-

ing of the chemistry and enzymes of nucleic acid metabolism described in the previous five chapters. We then turn to topics that help illustrate the range of applications and the potential of this technology.

DNA Cloning: The Basics

To clone means to make identical copies; it is a term that was once restricted to the procedure of isolating one cell from a larger population of cells, then allowing it to reproduce itself to generate many identical cells. In such a way, sufficient quantities of a single cell type were made available for study. By analogy, DNA cloning involves separating a specific gene or segment of DNA from its larger chromosome and attaching it to a small molecule of carrier DNA, then replicating this modified DNA thousands or even millions of times. The result is a selective amplification of that particular gene or DNA segment. Cloning a segment of DNA, either prokaryotic or eukaryotic, entails five general procedures:

1. A method for cutting DNA at precise locations. The discovery of sequence-specific endonucleases (restriction endonucleases) provided the necessary molecular scissors.
2. A method for joining two DNA fragments covalently. DNA ligase can do this.
3. Selection of a small molecule of DNA capable of self-replication. Segments of DNA to be cloned can be joined to plasmids or viral DNAs (cloning vectors). These composite DNA molecules containing covalently linked segments derived from two or more sources are called **recombinant DNAs.**
4. A method for moving recombinant DNA from the test tube into a host cell that can provide the enzymatic machinery for DNA replication.
5. Methods to select or identify those host cells that contain recombinant DNA.

The methods used to accomplish these and related tasks are collectively referred to as **recombinant DNA technology,** or more informally as **genetic engineering.** We now turn to these methods, with emphasis on their biochemical origins.

In this initial discussion we will focus on DNA cloning in the bacterium *E. coli,* which was the first organism used for recombinant DNA work and is still the most common host cell. *E. coli* has many advantages: its DNA metabolism (and many other biochemical processes) are well understood; many naturally occurring cloning vectors such as bacteriophages and plasmids associated with *E. coli* are well characterized; and effective techniques are available for moving DNA from one bacterial cell to another. DNA cloning in other organisms will be addressed later in the chapter.

Restriction Endonucleases and DNA Ligase Yield Recombinant DNA

Particularly important to recombinant DNA technology is a set of enzymes made available by decades of research on nucleic acid metabo-

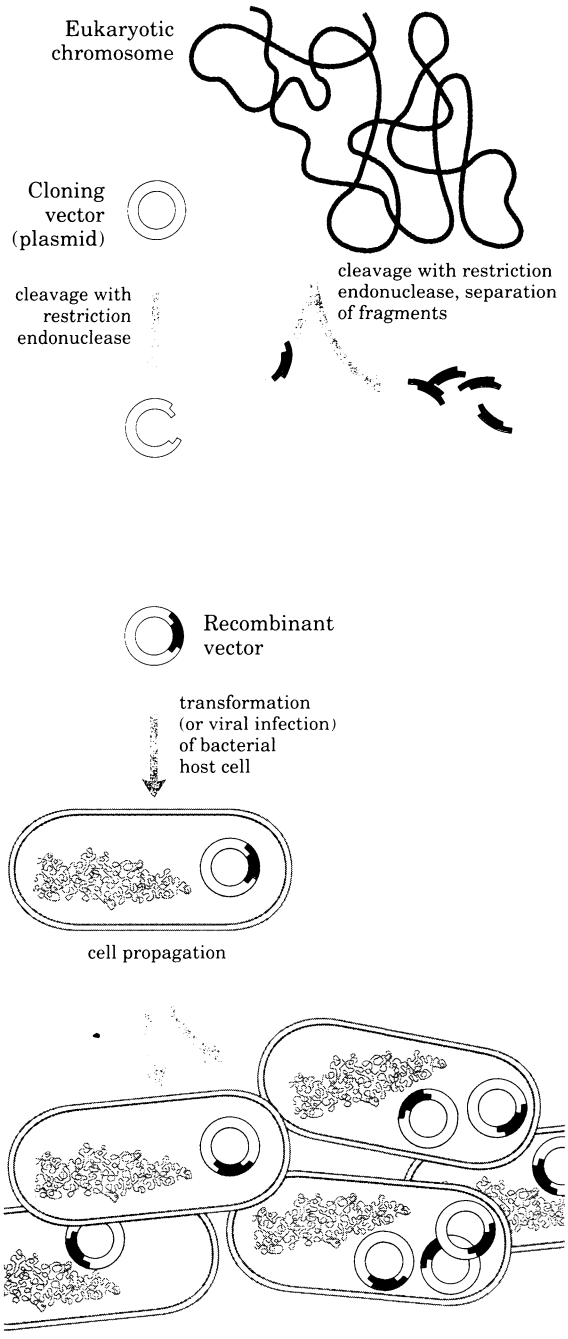


Figure 28–1 Schematic illustration of DNA cloning. A fragment of DNA of interest to the researcher is obtained by cleaving a eukaryotic chromosome with a restriction endonuclease. After isolating the fragment and ligating it to a cloning vector that has also been cleaved with a restriction endonuclease, the resulting recombinant DNA is introduced into a host cell where it can be propagated (cloned). Note that the size of the *E. coli* chromosome relative to that of a typical cloning vector such as a plasmid is much greater than depicted here.

lism (Table 28–1). Two enzymes in particular lie at the heart of the general approach to generating and propagating a recombinant DNA molecule as outlined in Figure 28–1. First, **restriction endonucleases** cleave DNA at specific sequences to generate a set of smaller fragments. Second, the DNA fragment to be cloned can be isolated and joined to a suitable cloning vector using **DNA ligase** to seal the DNA molecules together. The recombinant vector is then introduced into a host cell, which "clones" it as the cell undergoes many generations of cell divisions.

Table 28–1 Some of the enzymes used in recombinant DNA technology

Enzyme(s)	Function
Type II restriction endonucleases	Cleaving DNAs at specific base sequences
DNA ligase	Joining two DNA molecules or fragments
DNA polymerase I *(E. coli)*	Filling in gaps in duplexes by stepwise addition of nucleotides to 3′ ends
Reverse transcriptase	Making a DNA copy of an RNA molecule
Polynucleotide kinase	Adding a phosphate to the 5′-OH end of a polynucleotide to label it or permit ligation
Terminal transferase	Adding homopolymer tails to the 3′-OH ends of a linear duplex
Exonuclease III	Removing nucleotide residues from the 3′ ends of a DNA strand
Bacteriophage λ exonuclease	Removing nucleotides from the 5′ ends of a duplex to expose single-stranded 3′ ends
Alkaline phosphatase	Removing terminal phosphates from either the 5′ or 3′ end or both

Restriction endonucleases are found in a wide range of bacterial species. Werner Arber discovered that their biological function is to recognize and cleave foreign DNA (e.g., the DNA of an infecting virus); such DNA is said to be *restricted.* The cell's own DNA is not cleaved because the sequence recognized by the restriction endonuclease is methylated (and thereby protected) by a specific DNA methylase. The restriction endonuclease and the corresponding methylase in a bacterium are sometimes referred to as a **restriction-modification system.** There are three types of restriction endonucleases, designated I, II, and III. Types I and III are generally large, multisubunit complexes containing both the endonuclease and methylase activities. Type I restriction endonucleases cleave DNA at random sites that can be 1,000 base pairs or more from the recognition sequence. Type III enzymes cleave the DNA about 25 base pairs from the recognition sequence. Both types of enzyme move along the DNA in a reaction that requires the energy of ATP. The type II restriction enzymes, first isolated by Hamilton Smith, are simpler, require no ATP, and cleave the DNA within the recognition sequence itself. The extraordinary utility of the type II enzymes was first demonstrated by Daniel Nathans, and these are the enzymes used most widely for recombinant DNA work.

More than 800 restriction endonucleases have been discovered in different bacterial species. Over 100 different specific sequences are recognized by one or more of these enzymes. These sequences are almost always short (four to six base pairs, occasionally more) and palindromic (see Fig. 12–20). A sampling of sequences recognized by some type II restriction endonucleases is presented in Table 28–2. Note that the name of each enzyme consists of a three-letter abbreviation of the bacterial species from which it is derived (e.g., *Bam* for *Bacillus amyloliquefaciens*, *Eco* for *Escherichia coli*).

In a few cases, the interaction between a restriction endonuclease and its target sequence has been elucidated in exquisite molecular detail. The complex comprising the type II restriction endonuclease *Eco*RI and its target sequence is illustrated in Figure 28–2. DNA sequence recognition by *Eco*RI is mediated by 12 hydrogen bonds formed between the purines in the recognition site and six amino acid residues in the dimeric endonuclease (one Glu and two Arg residues in each subunit). Some restriction endonucleases cleave both strands of DNA so as to leave no unpaired bases on either end; these ends are often called **blunt ends** (Fig. 28–3a). Others make staggered cuts on the two DNA strands, leaving two to four nucleotides of one strand unpaired at each resulting end. These are referred to as **cohesive ends** or **sticky ends** (Fig. 28–3a) because they can base-pair with each other or with complementary sticky ends of other DNA fragments.

Table 28–2 Recognition sequences for some type II restriction endonucleases

*Bam*HI	(5′) G↓G A T C C (3′) C C T A G*G
*Cla*I	(5′) A↓T C G A T (3′) T A G C T*A
*Eco*RI	(5′) G↓A A T T C (3′) C T T A*A G
*Hae*III	(5′) G G↓C C (3′) C C G*G
*Hin*dIII	(5′) A↓A G C T T (3′) T T C G A*A
*Not*I	(5′) G C G G↓C C G C (3′) C G C C G G*C G
*Pst*I	(5′) C T G C A↓G (3′) G A C G T*C
*Pvu*II	(5′) C A G↓C T G (3′) G T C G A*C
*Sma*I	(5′) C C C↓G G G (3′) G G G C C*C
*Tth*111I	(5′) G A C N N N↓G T C (3′) C T G N N N*C A G

Arrows indicate the phosphodiester bonds cleaved by each restriction endonuclease. Asterisks indicate bases that are methylated by the corresponding methylase (where known). N denotes any base. The Roman numerals included in the enzyme names (e.g., *Bam*HI) distinguish different restriction endonucleases isolated from the same bacterial species rather than the type of restriction enzyme.

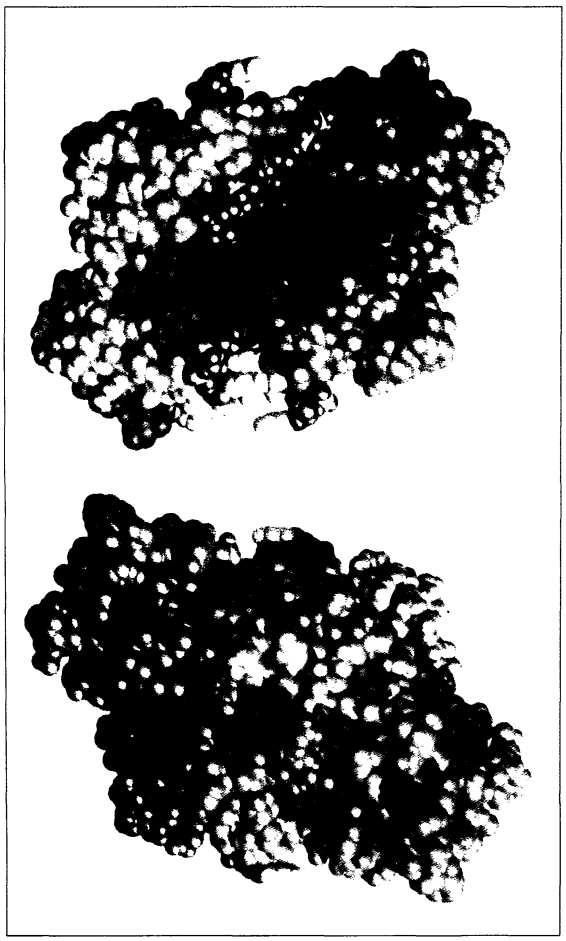
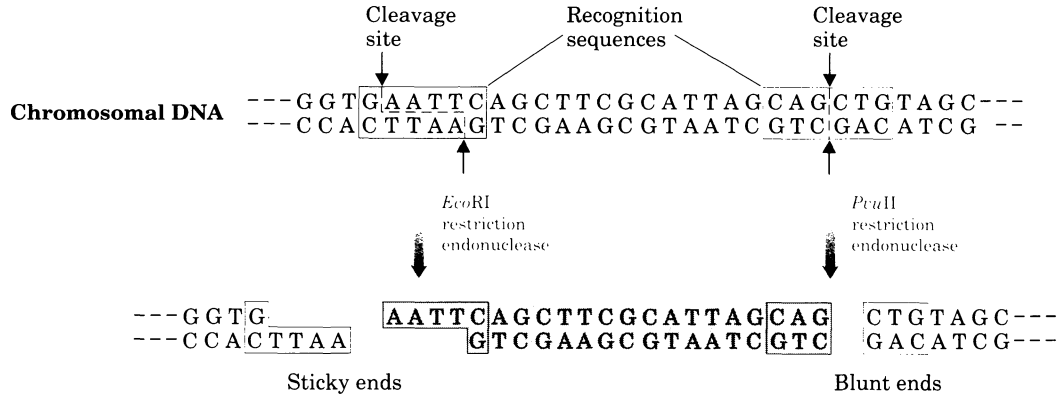
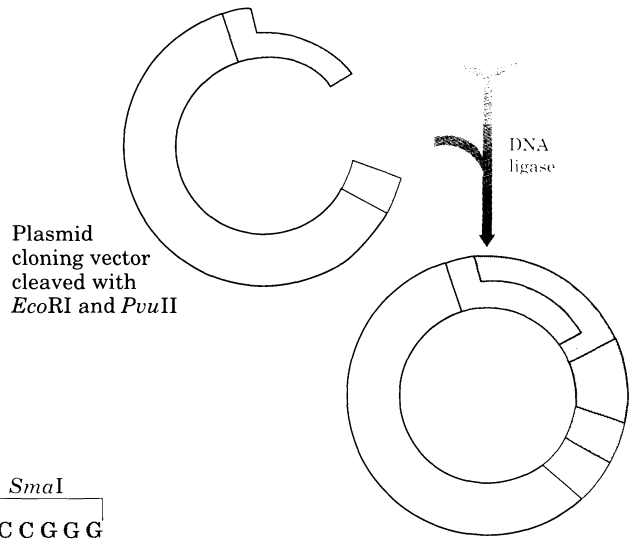
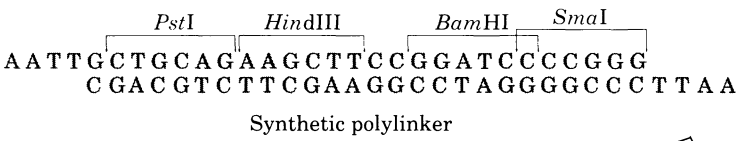


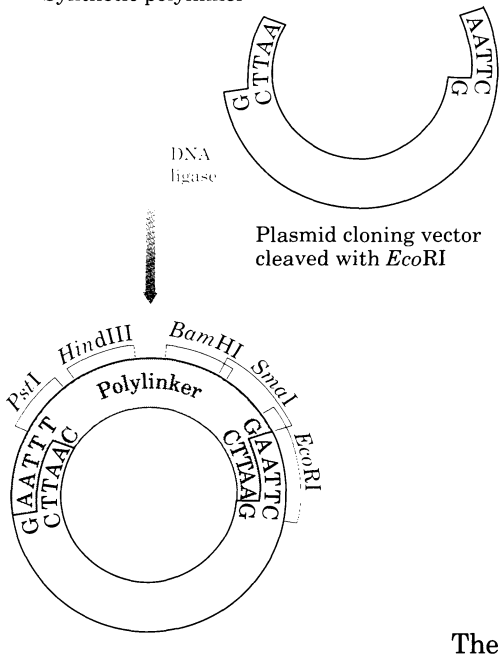
Figure 28–2 The interaction of *Eco*RI endonuclease with its target sequence. The dimeric enzyme (with its two subunits in gray and light blue) is shown bound to DNA. In the top view the DNA binding site is facing the viewer. In the bottom view the bound DNA is facing away from the viewer and is not visible. In the bound DNA, the bases that make up the recognition site for *Eco*RI are shown in red.







(a)



(b)

Figure 28–3 Cleavage of DNA molecules into reproducible fragments by restriction endonucleases. Restriction enzymes recognize and cleave only specific sequences, leaving either sticky ends (with protruding single strands) or blunt ends. **(a)** Cleavage of a DNA yields a characteristic set of fragments, and these fragments can be ligated to other DNAs such as the cleaved cloning vector (a plasmid) shown here. The ligation reaction is facilitated by the annealing of complementary sticky ends. DNA fragments with blunt ends are ligated at a lower efficiency than those with complementary sticky ends, and DNA fragments with different (noncomplementary) sticky ends generally are not ligated. **(b)** A synthetic DNA fragment containing the recognition sequences for several restriction endonucleases can be inserted into a plasmid after it has been cleaved by a restriction endonuclease; this creates a polylinker.

The average size of the DNA fragments produced by cleaving genomic DNA with a restriction endonuclease depends upon the frequency with which a particular restriction site occurs in a large DNA molecule; this in turn depends largely on the size of the recognition sequence. In a DNA molecule with a random sequence in which all four nucleotides are equally abundant, a 6 base pair sequence recognized by a restriction endonuclease such as *Bam*HI will occur on average once every 4^6, or 4,096, base pairs. Enzymes that recognize a 4 base pair

sequence will produce smaller DNA fragments; a recognition sequence of this size would be expected to occur on average every 256 base pairs. These sequences tend to occur less frequently than this because nucleotide sequences in DNA are not random and the four nucleotides are not equally abundant. The average size of the fragments produced by restriction endonuclease cleavage of a large DNA can be increased by simply not allowing the reaction to go to completion. Such an incomplete reaction is often called a partial digest.

Once a DNA molecule has been cleaved into fragments, a particular fragment that a researcher is interested in can be separated from the others by agarose gel electrophoresis (p. 347) or HPLC (p. 122). Because cleavage of a typical mammalian genome by a restriction endonuclease may yield several hundred thousand different fragments, isolation of a particular DNA fragment by electrophoresis or HPLC is often impractical. In these cases an intermediate step in the cloning of a specific gene or DNA segment of interest is the construction of a DNA library, described later in the chapter.

When the target DNA fragment is isolated, it is joined to a cloning vector using DNA ligase. The base-pairing of complementary sticky ends greatly facilitates the ligation reaction (Fig. 28–3). Because different restriction endonucleases usually generate different sticky ends, the efficiency of the ligation step is greatly affected by the endonucleases used to generate the DNA fragments. A fragment generated by *Eco*RI generally will not be linked to a fragment generated by *Bam*HI. Blunt ends can also be ligated, albeit less efficiently.

Before ligating two DNA fragments, it is often useful to add recognition sequences for a restriction endonuclease (other than that used to create the fragments) at the junction to permit cleavage of the ligated DNA at that location later on. This is often done by inserting a synthetic DNA fragment containing the required recognition sequence between the two DNA fragments. Such a synthetic DNA fragment is generally called a **linker.** A synthetic fragment containing recognition sequences for several restriction endonucleases is called a **polylinker** (Fig. 28–3b).

The importance of sticky ends in efficiently joining two DNA fragments in a desired manner was apparent in the earliest recombinant DNA experiments. Before restriction endonucleases were widely available, some workers found that sticky ends could be generated by the combined action of the bacteriophage λ exonuclease and terminal transferase (Table 28–1). The fragments to be joined were given complementary homopolymeric tails (Fig. 28–4, p. 990). This method was used by Peter Lobban and Dale Kaiser in 1971 in the first experiments to join naturally occurring DNA fragments. Similar methods were used soon after in the laboratory of Paul Berg to join DNA segments from simian virus 40 (SV40) to DNA derived from bacteriophage λ, thereby creating the first recombinant DNA molecule involving DNA segments from different species.

Cloning Vectors Amplify Inserted DNA Segments

Three types of cloning vectors—plasmids, bacteriophages, and cosmids—are commonly used in *E. coli.* **Plasmids** (see Fig. 23–3) are circular DNA molecules that replicate separately from the host chromosome. Naturally occurring bacterial plasmids range in size from 5,000 to 400,000 base pairs.

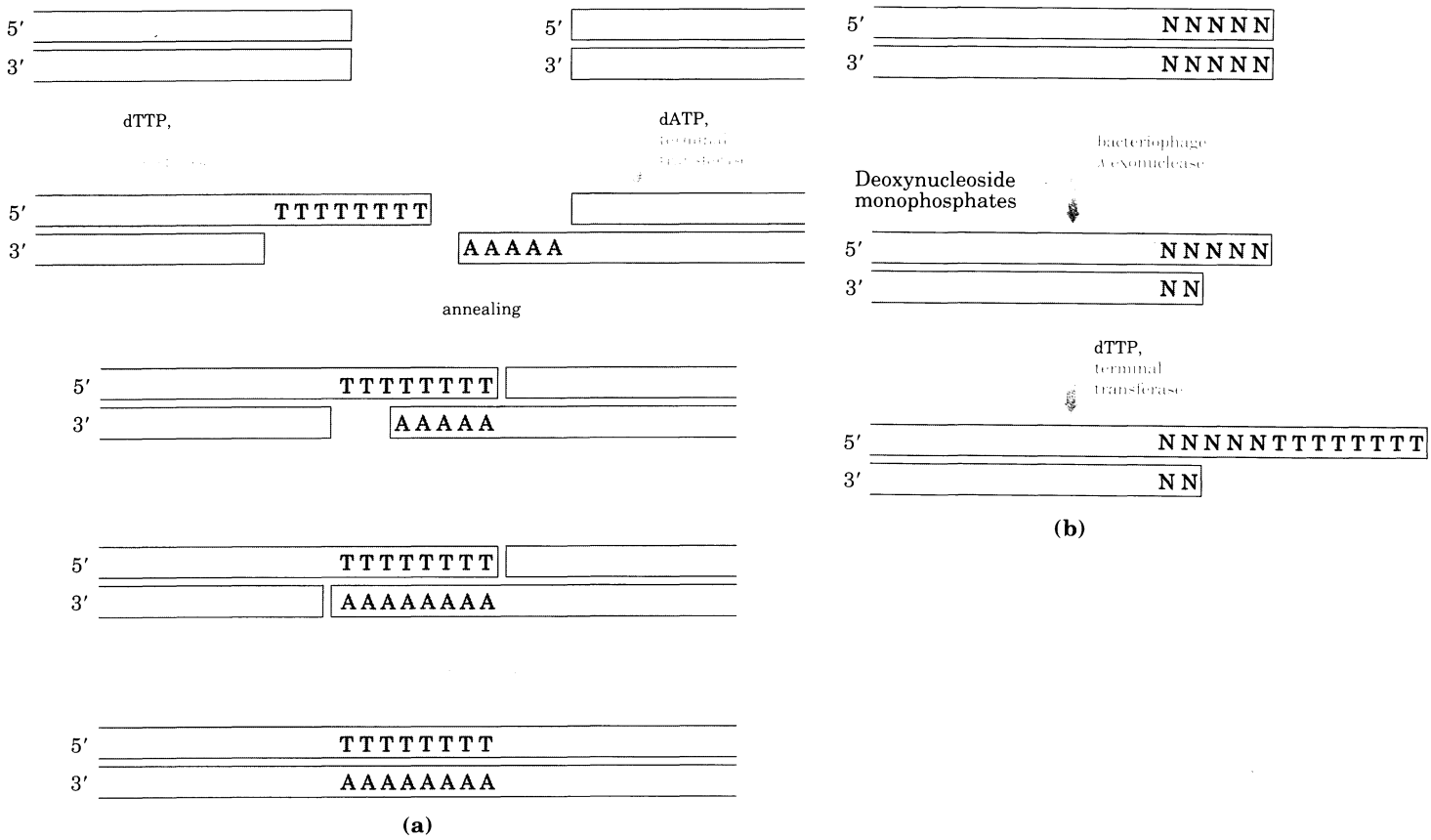


Figure 28–4 Sticky ends generated by terminal transferase can be used to join two DNA fragments. **(a)** Complementary homopolymeric tails are added to the ends of the two fragments to be joined, forming sticky ends. After annealing, the gaps are filled and the nicks sealed by the action of DNA polymerase I and DNA ligase (Chapter 24). **(b)** The optimal substrate for terminal transferase is the 3′ OH at the end of a single strand of DNA at least three nucleotides long. If the ends of the duplex DNA have a 5′ protruding single strand or are blunt ends, the λ exonuclease (which degrades DNA strands in the 5′→3′ direction) can be used to create a good substrate for terminal transferase. N denotes any base.

Plasmids can be introduced into bacterial cells by a process called **transformation.** To get the cells to take up the DNA, the cells and DNA are incubated together at 0 °C in a calcium chloride solution, then subjected to heat shock by rapidly shifting the cells to temperatures of 37 to 43 °C. For reasons not entirely understood, cells so treated become "competent" to take up the DNA. Because only a few cells take up the plasmid DNA, a method is needed for selecting those that do. The usual strategy is to build into the plasmid a gene that the host cell requires for growth under specific conditions. This makes a cell that contains the plasmid "selectable" if the cell is grown under those conditions. The gene, sometimes called a selectable marker, is often one that confers resistance to an antibiotic. Only those few cells that have been transformed by the recombinant plasmid will be antibiotic resistant and thus able to grow in the presence of the antibiotic.

Many different plasmid vectors suitable for cloning have been developed by modifying naturally occurring plasmids. Some of the important features of a cloning vector are illustrated by the *E. coli* plasmid pBR322 (Fig. 28–5): (1) the origin of replication is required to propagate the plasmid and helps maintain it at a level of 10 to 20 copies per cell; (2) two genes that confer resistance to different antibiotics allow the selection of cells that contain the plasmid or a recombinant version of it (Fig. 28–6); (3) several unique recognition sequences for different restriction endonucleases provide sites where the plasmid can be cut and foreign DNA inserted; and (4) an overall small size facilitates the plasmid's entry into cells. The efficiency of bacterial transformation decreases as plasmid size increases, and it is difficult to clone DNA segments longer than about 15,000 base pairs when plasmids are used as the vector.

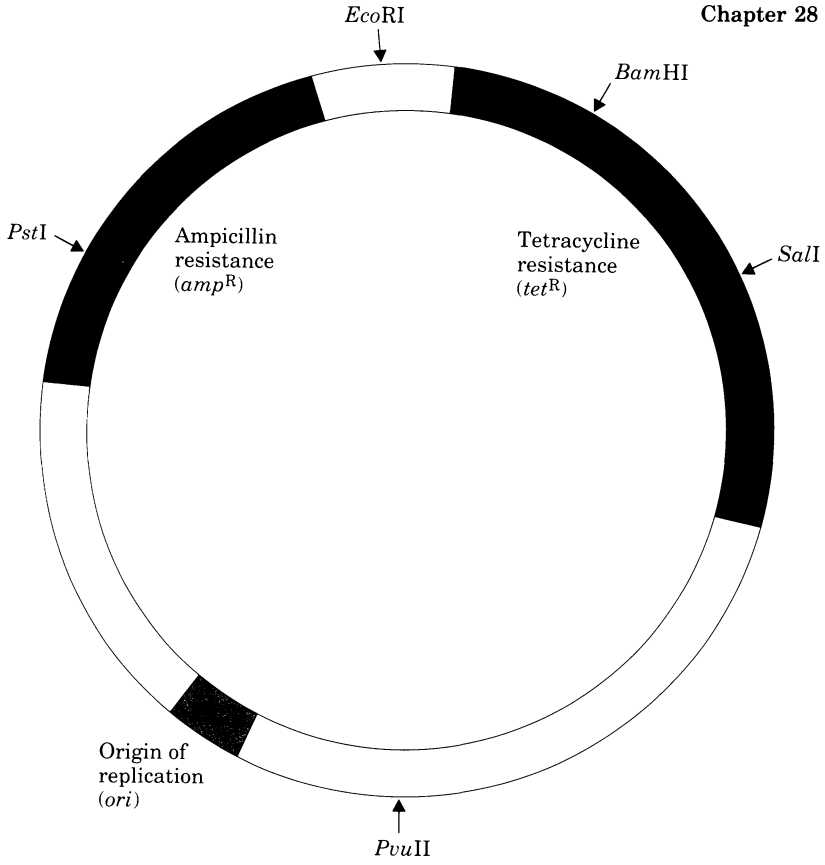


Figure 28–5 The constructed plasmid pBR322, showing the location of some important restriction sites, antibiotic-resistance genes, and the replication origin *(ori)*. This plasmid, constructed by Herbert Boyer and coworkers in 1977, was one of the early plasmids designed expressly for cloning in *E. coli*.

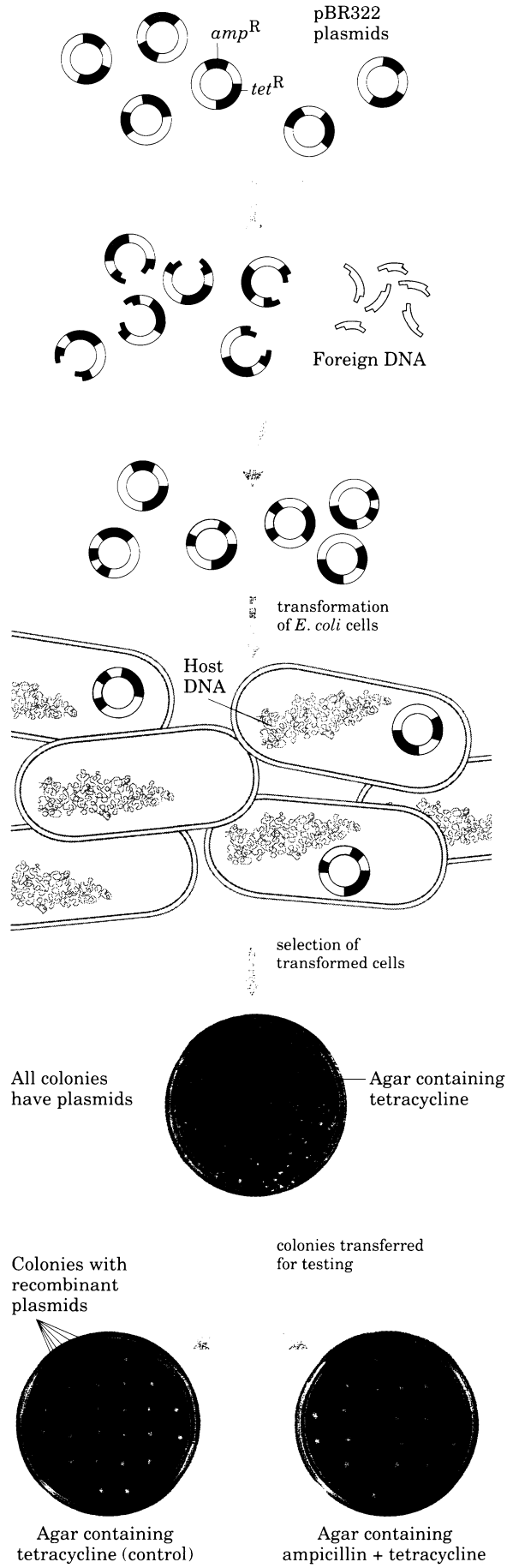


Figure 28–6 Cloning foreign DNA in *E. coli* with pBR322. If foreign DNA is inserted at the *Pst*I restriction site, the ampicillin-resistance element is disrupted and inactivated. After ligation of the DNA and transformation of *E. coli* cells, the cells are grown on agar plates containing tetracycline to select for those that have taken up a plasmid. By means of sterile toothpicks, individual colonies from these agar plates are transferred to the same position within a grid on two additional plates; one plate contains tetracycline (a control) and the other contains both tetracycline and ampicillin. Those cells that grow in the presence of tetracycline, but do not form colonies on the plate containing tetracycline plus ampicillin, contain recombinant plasmids (the ampicillin-resistance element is nonfunctional). Cells that contain pBR322 that was ligated without the insertion of a foreign DNA fragment retain ampicillin resistance and grow on both plates. Note that in this and other experiments that involve the use of two or more plate replicas an orienting mark is put on the back of each plate so that the colonies on different plates can be readily aligned.

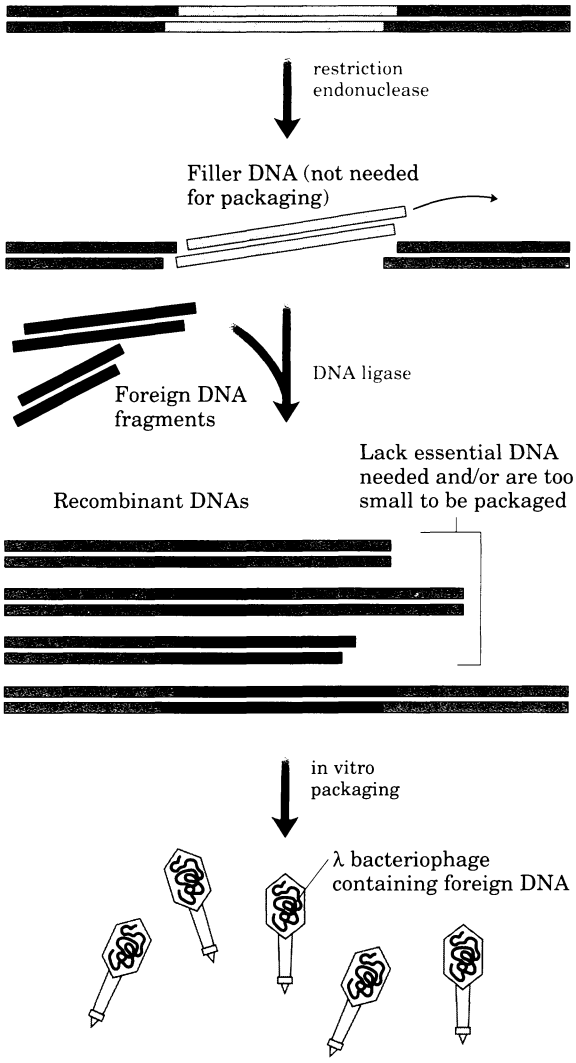


Figure 28–7 Bacteriophage λ cloning vectors. Recombinant DNA methods have been used to remove nonessential genes and certain restriction sites from the bacteriophage λ genome. The remaining genes (essential for bacteriophage production) are clustered in two large fragments at either end of the linear chromosome. Bacteriophage λ vectors generally have a piece of "filler" DNA in place of the eliminated genes to make the vector DNA large enough for packaging into phage particles. This filler can be replaced with foreign DNA in cloning experiments. Recombinants are packaged into viable phage particles only if they are of the appropriate size (i.e., they include an appropriately sized foreign DNA fragment) and include both of the essential λ DNA end fragments. The recombinant DNA molecules are packaged into phage particles in vitro.

Somewhat larger DNA segments can be cloned using **bacteriophage λ** as a vector. Bacteriophage λ has a very efficient mechanism for delivering its 48,502 base pairs of DNA into a bacterium. The general procedure for cloning DNA in bacteriophage λ (Fig. 28–7) is based on two key features of the λ genome: (1) about one-third of the genome is nonessential and can be replaced with foreign DNA, and (2) DNA will be packaged into infectious phage particles only if it is between 40,000 and 50,000 base pairs long. Bacteriophage λ vectors have been developed that can be readily cleaved into three pieces, two of which contain essential genes but which together are only about 30,000 base pairs long. Additional DNA must therefore be inserted between them to produce viable phage particles. Bacteriophage λ vectors permit the cloning of DNA fragments up to 23,000 base pairs long, and their design ensures that all viable phage particles will contain a foreign DNA fragment.

Once the bacteriophage λ fragments are ligated to foreign DNA fragments of suitable size, the resulting recombinant DNAs can be packaged into phage particles by adding them to crude bacterial cell extracts containing all the proteins needed to assemble a complete phage. This is called **in vitro packaging** (Fig. 28–7). The bacteriophage vector is now ready for insertion of the recombinant DNA into *E. coli* cells.

Cosmids are recombinant plasmids that combine useful features of both plasmids and bacteriophage λ. They are designed to permit the cloning of even larger DNA fragments (up to 45,000 base pairs). Cosmids (Fig. 28–8) are small (typically 5,000 to 7,000 base pairs), circular DNA molecules that contain (1) a plasmid origin of replication, (2) one or more selectable markers, (3) a number of unique restriction sites

Table 28–3 Types of cloning vectors used in *E. coli*

Type of vector	Method of introduction into *E. coli*	Method of propagation	Size of DNA fragment that can be cloned
Plasmids; modified by recombinant DNA techniques	Transformation; cells made competent to take up recombinant vector, then transformed cells selected using selectable marker	Plasmid replication	Up to 15,000 bp
Bacteriophage λ	Phage infection, following in vitro packaging of recombinant vector into phage particles	Phage replication	Up to 23,000 bp
Cosmids, constructed from plasmid and λ DNA genes	Either of above methods, depending on size of DNA fragment inserted; larger fragments require in vitro λ packaging	Plasmid-type replication	Up to 45,000 bp

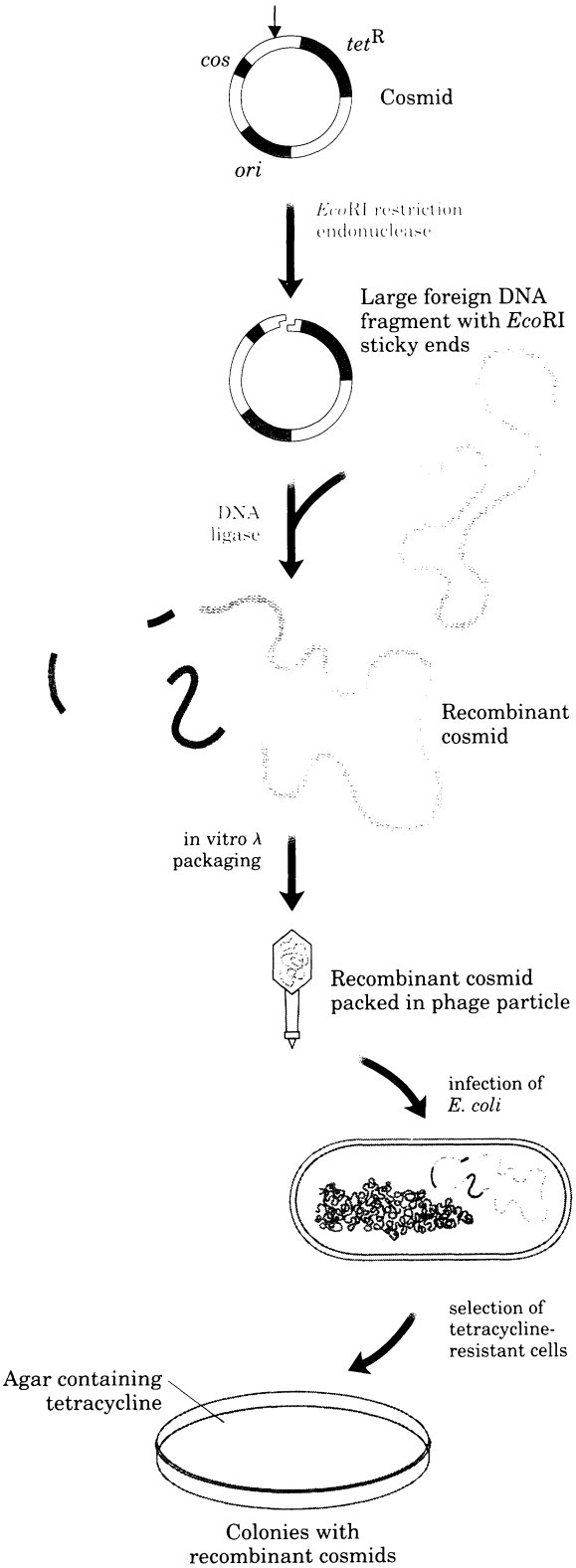


Figure 28–8 Cloning with cosmids. A cosmid contains a replication origin *(ori)* for propagation as a plasmid, and a *cos* site required for the packaging of DNA into λ phage particles. Unique restriction sites and antibiotic-resistance elements aid cloning and selection. Inserting a large foreign DNA fragment into the cosmid precludes bacterial transformation but, if the recombinant DNA molecule is a suitable size, it can be packaged into phage particles in vitro. The phage λ particles are then used to introduce the cosmid into bacterial cells, where the cosmid is propagated as a plasmid.

where foreign DNA can be inserted, and (4) a *cos* site (a DNA sequence in bacteriophage λ that is required for packaging).

Cosmids contain no other bacteriophage λ genes and can be propagated in *E. coli* like plasmids. When a large foreign DNA fragment is cloned into them, transformation of *E. coli* with these recombinant DNA constructs becomes difficult. If the cosmid contains a large enough insert of foreign DNA to be packaged into a phage particle, in vitro λ packaging systems permit the bacteriophage λ particle to be the vehicle for introducing the cosmid DNA efficiently into the bacterial cell. Once in the cell, the cosmid is again propagated as a plasmid, as it lacks the λ genes needed to make phage particles in the cell.

The three types of cloning vectors are summarized in Table 28–3.

Isolating a Gene from a Cellular Chromosome

Because a single gene is only a very small part of a chromosome, isolating a DNA fragment containing a particular gene often requires two procedures. First, a DNA library is constructed that contains many thousands of DNA fragments derived from a cellular chromosome. Second, the DNA fragment containing the gene of interest is identified by taking advantage of the one property that distinguishes it from the other DNA fragments—its sequence.

Cloning a Gene Often Requires a DNA Library

A **DNA library** is a collection of DNA fragments derived from the genome of a particular organism, with each fragment attached to a cloning vector. In short, genomic DNA is cleaved into thousands of fragments, and *all* of them are cloned. Thus, the total information content of an organism is represented by all the fragments in the library in much the same way as human knowledge is represented by all the volumes in a book library. A vector is prepared for cloning by cleaving purified vector DNA at an appropriate position with a restriction endonuclease. Genomic DNA to be cloned is reduced to appropriately sized fragments by partial digestion with a restriction endonuclease (usually the same one used to prepare the vector), followed by a procedure such as sucrose density gradient (isopycnic) centrifugation to remove fragments that are too large or too small to be cloned into the chosen vector. Fragments are then mixed with the cleaved vector DNA and ligated. The mixture is used to transform bacterial cells or is packaged

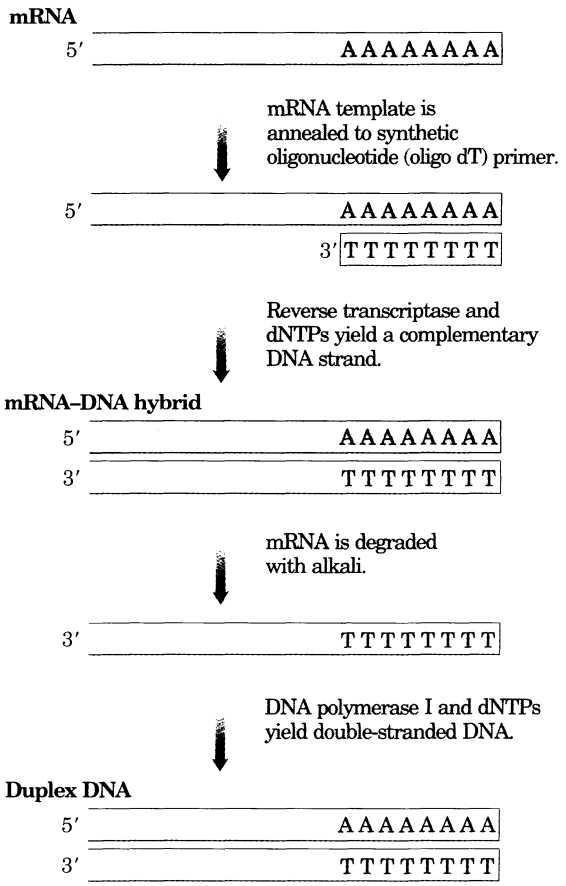


Figure 28–9 Constructing a cDNA library from mRNA. In practice, the mRNA from a cell will include transcripts from thousands of genes, and the cDNAs generated will be correspondingly heterogeneous. The duplex DNA produced by this method is inserted into an appropriate cloning vector.

into bacteriophage particles as described above. The final result is a large population of bacteria or bacteriophages each harboring a different recombinant DNA molecule. Ideally, nearly all of the DNA in the genome will be represented in the library, and libraries constructed in this way are referred to as **genomic libraries.** Each transformed bacterium grows into a colony or "clone" of cells, all of which have the same recombinant plasmid. In the case of bacteriophages, each type of recombinant phage creates a plaque—a clear region of lysed cells within a lawn of bacteria distributed evenly on an agar plate; all of the recombinant bacteriophages within a plaque are identical. The clone containing the particular gene a researcher is interested in must be identified within the thousands of clones in the library as described below. If the desired DNA is from a mammal with a genome of 3×10^9 base pairs of DNA and cosmids are used as cloning vectors, then the library must contain about 350,000 recombinant cosmids, each with a different insert of 35,000 to 45,000 base pairs, for there to be a 99% chance that any desired gene of unique sequence is represented in the library. The fragments in a genomic library derived from a higher eukaryote include not only genes but the noncoding DNA that makes up a large portion of many eukaryotic genomes.

A more specialized and exclusive DNA library can be constructed so as to include only those genes that are *expressed* in a given organism or even in certain cells or tissues. The critical difference between genes that are expressed and genes that are not is that the former are transcribed into RNA. The mRNA from an organism or certain cells derived from the organism is extracted, and **complementary DNAs (cDNAs)** are produced from the RNA in a multistep reaction catalyzed by reverse transcriptase (Fig. 28–9; Chapter 25). The resulting double-stranded DNA fragments are then inserted into a suitable vector and cloned, creating a population of clones called a **cDNA library.** Among the first eukaryotic genes characterized were those encoding globins, primarily because cloning of these genes was facilitated by making cDNA libraries from erythrocyte precursor cells, in which about half of the mRNA codes for globins.

Hybridization Identifies Specific Sequences in a DNA Library

After a DNA library has been generated, the challenge is to isolate one gene of interest from among the millions of other DNA segments represented in a DNA library. Because each gene has a unique nucleotide sequence, it often can be detected with the aid of a labeled (e.g., radioactive) DNA fragment that is complementary to it, called a **probe.** Typically, nitrocellulose paper is pressed onto an agar plate containing many individual bacterial colonies each containing a different recombinant DNA. Some cells from each colony adhere to the paper, forming a replica of the plate. The paper is treated with alkali to disrupt the cells and denature the DNA within, which remains localized to the region around the colony from which it came. The radioactive DNA probe is then added to the paper, annealing only to DNA containing the complementary gene. The labeled colony is visualized by autoradiography (Fig. 28–10). Detecting any nucleic acid with a labeled nucleic acid probe that is complementary to it is an application of nucleic acid hybridization (see Fig. 12–31), and many variations on this procedure have been developed to detect both DNA (Box 28–1, p. 996) and RNA.

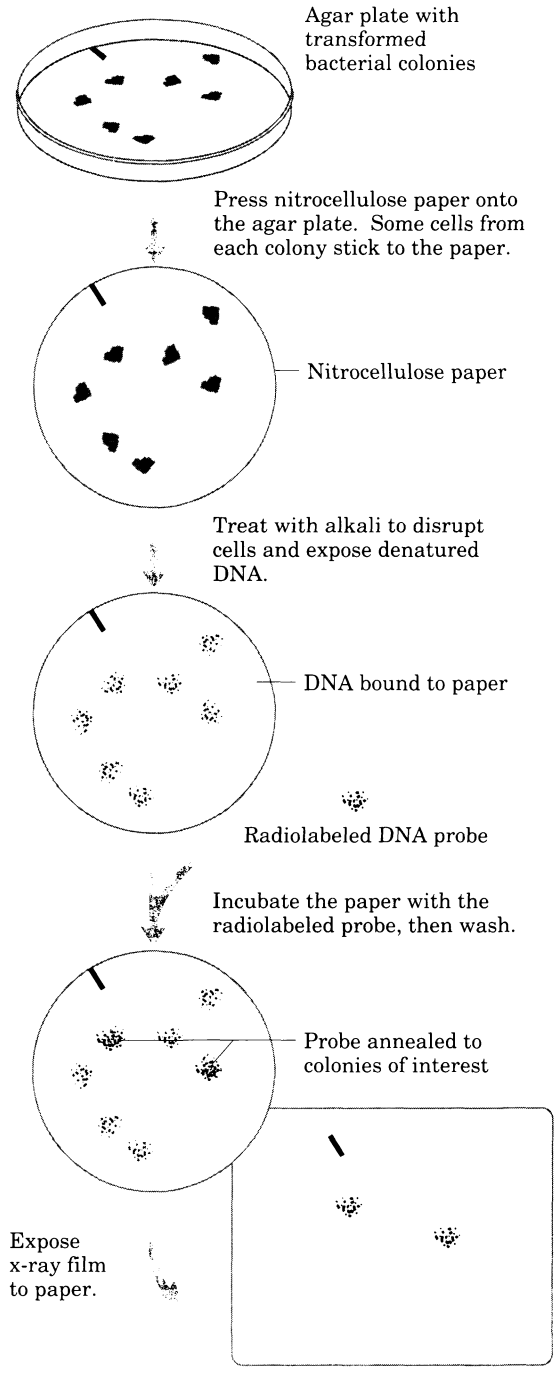


Figure 28–10 Identifying a clone with the desired DNA segment; a schematic diagram of a hybridization method. The radioactive DNA probe hybridizes only with homologous DNA. The annealed radioactive probe is revealed by autoradiography. When the labeled colonies have been identified, the corresponding colonies on the original agar plate can be used as a source of cloned DNA for further study.

Frequently, the limiting step in cloning a gene is finding or generating a complementary strand of nucleic acid to use as a probe. The origin of a probe depends on what is known about the gene under investigation. Sometimes a homologous gene cloned from another species can be used as a probe. Alternatively, if the protein product of a gene has been purified, probes can be designed and synthesized on the basis of its amino acid sequence and a knowledge of the genetic code (Fig. 28–11).

In this work, the synthesis and sequencing methods for proteins and DNA described in previous chapters come together and complement one another. Just as the sequence of an isolated protein can be used to isolate the corresponding gene, a cloned gene can be used to isolate its protein product. The sequence of an isolated gene can be used to design and chemically synthesize a short peptide that represents part of the protein product of the gene. Antibodies to this synthetic peptide can then be used to identify and purify the protein itself (Chapter 6) for further study.

Figure 28–11 Designing a probe to detect the gene for a protein of known amino acid sequence. Because the genetic code is degenerate, there is more than one possible DNA sequence that codes for any given amino acid sequence. As the correct DNA sequence cannot be known in advance, the probe is designed to be complementary to a region of the gene with minimal degeneracy. The oligonucleotides are synthesized with the sequence selectively randomized so that some contain either of the two possible nucleotides at each position of potential degeneracy (shaded in red). In the example shown, the synthesized oligonucleotide is actually a mixture of eight different sequences; one of the eight will complement the gene perfectly. All eight will match in at least 17 of 20 positions.

Known amino acid sequence $H_3\overset{+}{N}$---Gly — Leu — Pro — Trp — Glu — Asp — Met — Trp — Phe — Val — Arg---COO^-

Possible codons

(5′) GGA UUA CCA | UGG GAA GAC AUG UGG UUC GUA AGA (3′)
GGC UUG CCC | GAG GAU UUU GUC AGG
GGU CUA CCU | GUU CGA
GGG CUC CCG | GUG CGC
CUU CGU
CUG CGG

Region of minimal degeneracy

Synthetic probes UGG GA(A/G) GA(C/U) AUG UGG UU(C/U) GU

20 nucleotides long, 8 possible sequences

BOX 28–1 A New Weapon in Forensic Medicine

Traditionally, one of the most accurate methods for placing an individual at the scene of a crime has been a fingerprint. A new identification technique has become available based on methods developed for recombinant DNA technology. This is sometimes called DNA fingerprinting (also DNA typing or DNA profiling), and in some ways it is more powerful than any other identification method.

DNA fingerprinting is based on **sequence polymorphisms** that occur in the human genome (and the genome of every other organism). Sequence polymorphisms are slight sequence differences (usually single base-pair changes) that occur from individual to individual once every few hundred base pairs, on average. Each difference from the consensus human genome sequence is generally present in only a fraction of the human population, but every individual has some of them. Some of the sequence changes affect recognition sites for restriction enzymes, resulting in variation from individual to individual in the size of certain DNA fragments produced by digestion with a particular restriction enzyme. These size differences are referred to as **restriction fragment length polymorphisms,** or RFLPs. A probe for a sequence that is repeated several times in the human genome generally identifies a few of the thousands of DNA fragments generated when the human genome is digested with a restriction endonuclease.

The detection of RFLPs relies on a specialized hybridization procedure called **Southern blotting** (Fig. 1). DNA fragments from digestion of genomic DNA by restriction endonucleases are first separated according to size by electrophoresis in an agarose gel. The DNA fragments are denatured by soaking the gel in alkali, then transferred to nitrocellulose paper in such a way as to reproduce on the paper the distribution of fragments in the gel. The paper is then immersed in a solution containing a radioactively labeled DNA probe. Fragments to which the probe hybridizes are revealed by autoradiography, using procedures similar to those described in Figure 28–10.

The genomic DNA sequences used in these tests are generally regions containing repetitive DNA (short sequences repeated thousands of times in tandem; see p. 797), which are common in the genomes of higher eukaryotes. The number of repeated units in such DNA varies from individual to individual (except in the case of identical twins). If a suitable probe is chosen, the pattern of bands in such an experiment can be distinctive for each individual tested. If several probes are used, the test can be made so selective that it can positively identify a single individual in the human population. However, the Southern blot procedure requires relatively fresh DNA samples and larger amounts of DNA than are generally present at a crime scene. To increase sensitivity, RFLP analysis is being augmented by polymerase chain reaction (PCR) methods (see Fig. 28–12), which permit vanishingly small amounts of DNA to be amplified. The improved tests allow DNA fingerprints to be obtained from a single hair, a small semen sample from a rape victim, or from samples that might be months or even many years old.

Since their introduction in 1985, these methods have been developed to the point where they are proving decisive in court cases worldwide. In the example in Figure 1, the DNA from a semen sample obtained from a rape and murder victim was analyzed along with DNA samples from the victim and two suspects. Each of the DNA samples was cleaved into fragments and separated by gel electrophoresis. Radioactive DNA probes were used to identify a small subset of these fragments that contained sequences complementary to the probe. The sizes of the fragments identified varied from one individual to the next, as seen here in the different patterns for the three individuals (victim and two suspects) tested. One rape suspect's DNA exhibits a banding pattern identical to that of a semen sample taken from the victim. One probe was used here, but three or four different probes would be used to make a positive identification. Results have been used to help both convict and acquit suspects. DNA fingerprints can also be used to establish paternity with an extraordinary degree of certainty. The results of DNA fingerprinting have been successfully challenged in some court cases because of irregularities and a lack of controls in many early examples. The far-reaching impact of this technology on court cases will nevertheless continue to grow as standards are agreed upon and the methods become widely established in forensic laboratories.

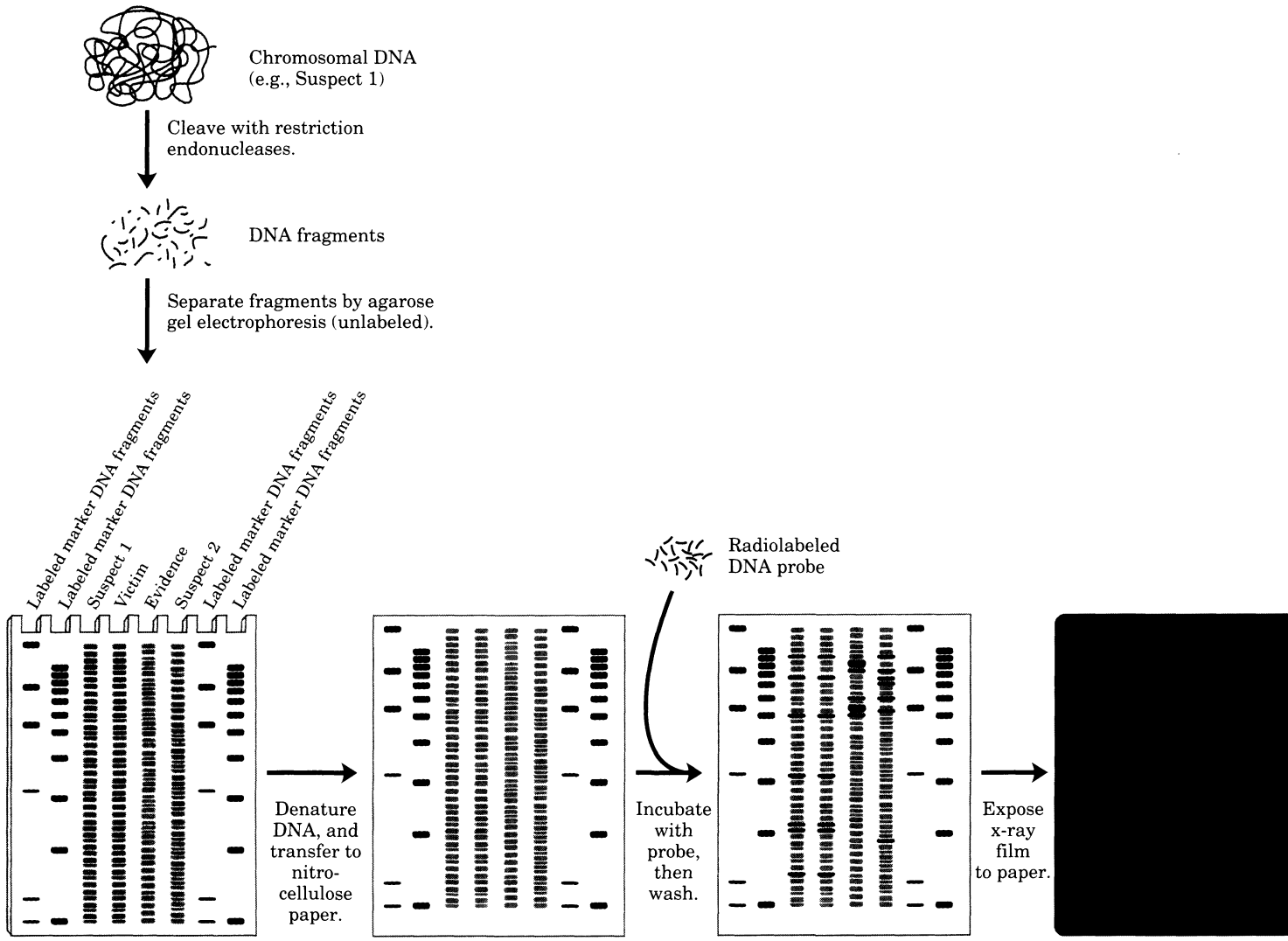


Figure 1 The Southern blot procedure, as applied to DNA fingerprinting.

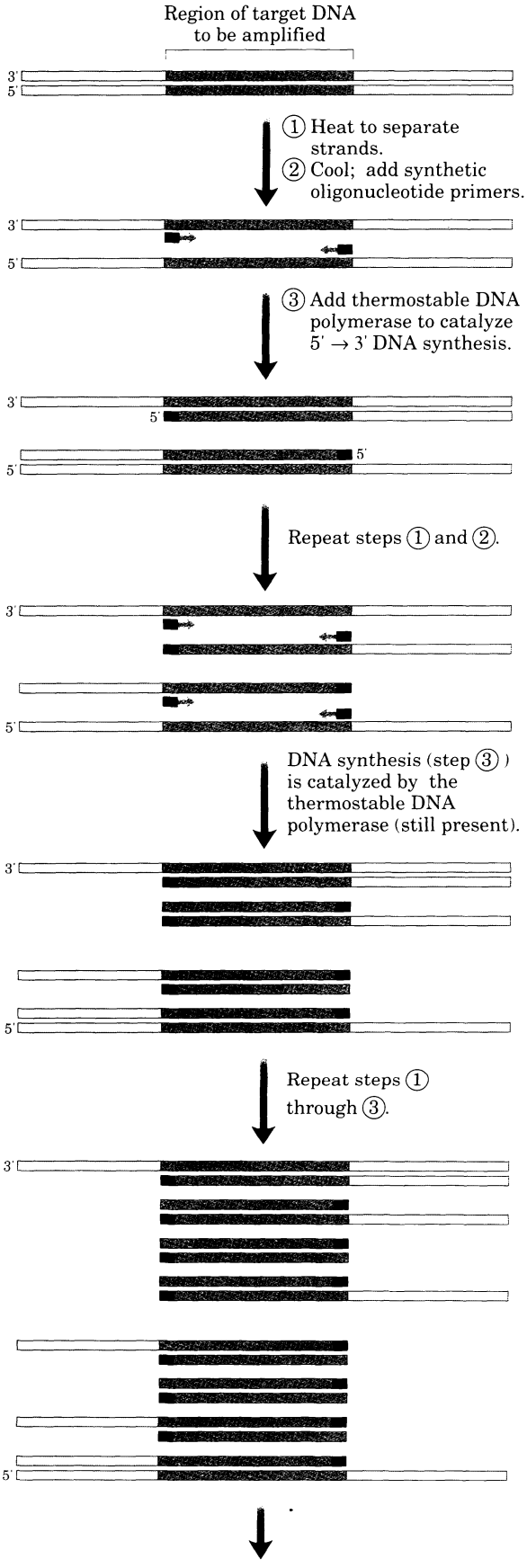


Figure 28–12 Amplifying a specific DNA segment with a polymerase chain reaction. DNA strands are separated by heating, then annealed to an excess of short synthetic DNA primers (blue) that flank the region to be amplified. After polymerization, the process is repeated for 25 or 30 cycles. The thermostable DNA polymerase *Taq*I (from *Thermus thermophilus,* bacteria that grow in hot springs) is not denatured by the heating steps.

Specific DNA Sequences Can Be Amplified

If one knows the sequence of at least part of a DNA segment to be cloned, cloning can be facilitated by amplifying the DNA segment in a process called a **polymerase chain reaction (PCR),** invented by Kary Mullis in 1984. Two oligonucleotides are synthesized, each complementary to a short sequence in one strand of the desired DNA segment and positioned just beyond the end of the sequence to be amplified; these synthetic oligonucleotides can be used as primers for replication of the DNA segment in vitro (Fig. 28–12). Isolated DNA containing the segment to be cloned is heated briefly to denature it, then cooled in the presence of a large excess of the synthetic oligonucleotide primers. A heat-stable DNA polymerase called *Taq*I and the four deoxynucleoside triphosphates are then added, and the primed DNA segment is selectively replicated. This process is repeated through 25 or 30 cycles, which can take only a few hours when automated, amplifying the DNA segment to the point where it can be readily isolated and cloned.

The PCR method is sensitive enough to detect as little as one DNA molecule in almost any type of sample. It has been used to clone DNA fragments from mummies and the remains of extinct animals such as the woolly mammoth, creating the new fields of molecular archaeology and molecular paleontology. In addition to its usefulness for cloning DNA, it is a potent new tool in forensic medicine (Box 28–1). It is also being used for detection of viral infections before they cause symptoms or elicit a detectable immune response and in prenatal diagnosis of a wide array of genetic diseases.

The Products of Recombinant DNA Technology

Normally, cloning a gene is only the first step in a much grander design. A cloned gene can be used to generate large amounts of its protein product. The amino acid sequence of the protein can be altered by introducing base-pair changes in the gene, a strategy that can be very powerful in addressing questions concerning protein folding, structure, and function. Increasingly sophisticated methods for moving DNA into and out of cells of all types are providing another avenue for studying gene function and regulation, and are allowing the introduction of new traits into plants and animals.

Our focus now turns to applications of DNA cloning, beginning with the proteins produced by cloned genes. We then describe cloning procedures used for a variety of eukaryotic cells, before finishing with an overview of the potential and implications of this technology.

mation of this type in which the introduced DNA is integrated into a cellular chromosome is called **integrative transformation;** it occurs at low frequency.

Transformation efficiencies can be increased by introducing cloned DNA on a self-replicating plasmid. A naturally occurring yeast plasmid called the 2 micron (2μ) plasmid has been engineered to create a variety of cloning vectors that incorporate a replication origin and other sequences needed for plasmid maintenance in yeast. Another type of plasmid that provides an increase in transformation frequencies contains a yeast chromosomal origin of replication (an autonomously replicating sequence, ARS; see p. 830). Such plasmids are somewhat unstable (and are lost from the yeast population) unless they also contain a centromere that permits them to function and segregate like yeast chromosomes during cell division.

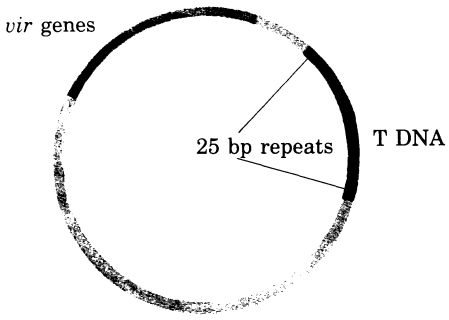
Recombinant plasmids are available that incorporate multiple genetic elements (replication origins, etc.), allowing them to be maintained in more than one species; for example, yeast or *E. coli*. Plasmids that can be propagated in cells of two or more different species are called **shuttle vectors.**

Cloning in Plants Is Aided by a Bacterial Parasite

The introduction of recombinant DNA into plants has enormous potential for agriculture, producing more nutritious and higher-yielding crops that are resistant to environmental stresses such as insect pests, disease, cold, and drought. Unlike animals, fertile plants of some species may be generated from a single transformed cell. Thus, a gene introduced into a plant cell may ultimately be transmitted to progeny through seed in successive generations. As with all systems, cloning in plant cells has its own peculiar problems. No naturally occurring plasmids have been found in plants to facilitate this process, and thus the most challenging task is getting DNA into plant cells.

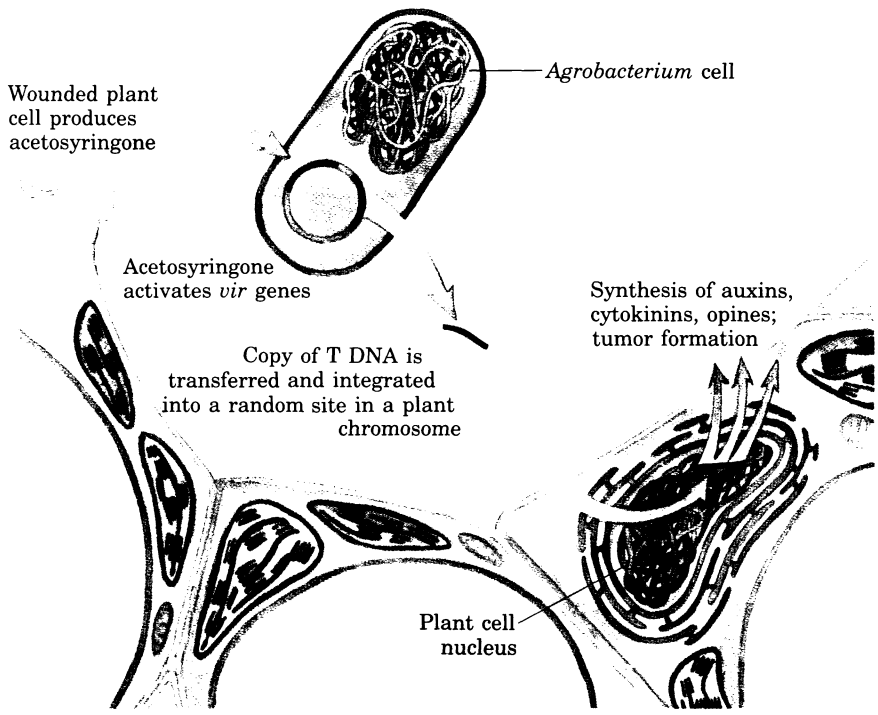
Fortunately, scientists have found an important and adaptable ally in the soil bacterium *Agrobacterium tumefaciens*. The bacterium invades plants at the site of a wound, transforming plant cells near the wound and inducing them to form a tumor called a crown gall. *Agrobacterium* contains a large (~200,000 base pair) plasmid called the **Ti plasmid** (Fig. 28–15). When the bacterium contacts a plant cell, a segment of this plasmid, called the T DNA (~23,000 base pairs), is transferred from the Ti plasmid to the plant cell nucleus and is integrated at a random position in one of the plant's chromosomes during transformation. This is a rare example of DNA transfer from a prokaryote to a eukaryote; it represents a natural genetic engineering process.

The T DNA encodes enzymes that convert plant metabolites to two classes of compounds important to the bacterium (Fig. 28–16). The first class consists of the plant growth hormones, auxins and cytokinins, which stimulate growth of the transformed plant cells to form the crown gall tumor. The second is a series of unusual amino acids called opines, a food source for the bacteria. The opines are produced at high concentrations in the tumor and secreted to the surroundings. They can only be metabolized by *Agrobacterium,* using enzymes encoded elsewhere on the Ti plasmid. In this manner the bacteria monopolize available nutrients by converting them to a form that does not benefit any other organism.



Ti plasmid

(a)

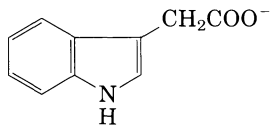


(b)

Figure 28–15 (a) The Ti (tumor-inducing) plasmid of *Agrobacterium tumefaciens*. **(b)** Acetosyringone, a phenolic compound produced by plants, increases in concentration in wounded plant cells and is released from the cells. *Agrobacterium* senses this compound, and the virulence *(vir)* genes on the Ti plasmid are expressed. The *vir* genes encode enzymes needed to introduce the T DNA into the genome of plant cells in the vicinity of the wound. A single-stranded copy of the T DNA is synthesized and transferred to the plant cell, where it is converted to duplex DNA and integrated into a chromosome. The T DNA encodes enzymes that synthesize growth hormones and opines (see Fig. 28–16) from common metabolites. Opines can be metabolized only by *Agrobacterium,* which uses them as a nutrient source. Expression of the T DNA genes by transformed plant cells thus leads to cell growth (tumor formation) and the diversion of plant cell nutrients to the invading bacteria.

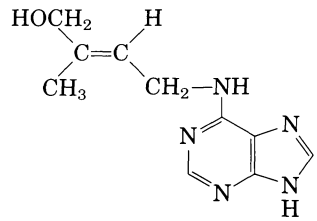
Figure 28–16 Metabolites produced in *Agrobacterium*-infected plant cells. Auxins and cytokinins are growth hormones. The most common auxin is indoleacetate, derived from tryptophan. Cytokinins are adenine derivatives. Opines generally are derived from amino acid precursors; at least 14 different opines have been identified in different species of *Agrobacterium*.

Auxins

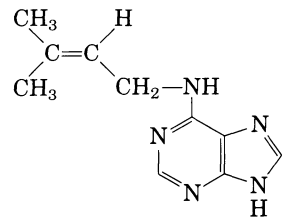


Indoleacetate

Cytokinins

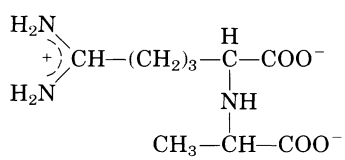


Zeatin

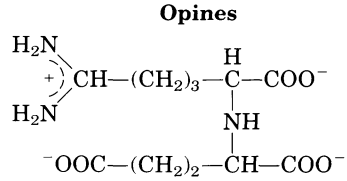


Isopentenyl adenine (i^6 Ade)

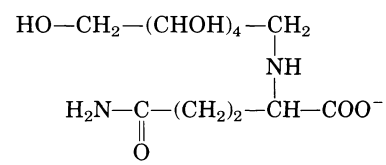
Opines



Octopine



Nopaline



Mannopine

The integration of the T DNA into a plant chromosome provides the vehicle necessary to introduce new genes into plants. The transfer of T DNA from *Agrobacterium* to the plant cell nucleus is mediated by two 25 base pair repeats that flank the T DNA and by the products of several genes, called virulence *(vir)* genes, located elsewhere on the Ti plasmid (Fig. 28–15).

The bacterial system that transfers T DNA into the plant genome can be harnessed to transfer recombinant DNA instead. A common cloning strategy employs an *Agrobacterium* that contains two different recombinant plasmids (Fig. 28–17). The first is a Ti plasmid from which the T DNA segment has been removed in the laboratory. The second is an *Agrobacterium–E. coli* shuttle vector that contains the T-DNA 25 base pair repeats flanking the gene that a researcher wants to introduce and a selectable marker (often a gene that can render plant cells resistant to an antibiotic such as kanamycin). The engineered *Agrobacterium* is used to infect a leaf (Fig. 28–17). Crown galls are not formed because the genes for auxin, cytokinin, and opine biosynthetic enzymes are not present on either plasmid. Instead, the *vir* gene products from the altered Ti plasmid direct the transformation of the plant cells by the gene flanked by the T-DNA 25 base pair repeats in the second plasmid. The transformed cells can be selected by growth on kanamycin and induced with growth hormones to form new plants.

The successful transfer of recombinant DNA into plants is vividly illustrated by an experiment in which the luciferase gene from fireflies was introduced into a tobacco plant (Fig. 28–18). Note that although tobacco is not high on the list of beneficial plants needing improvement, it is often used experimentally because it is particularly easy to transform with *Agrobacterium*.

The potential of this technology is not limited to the production of glow-in-the-dark plants. The same approach has been used to produce

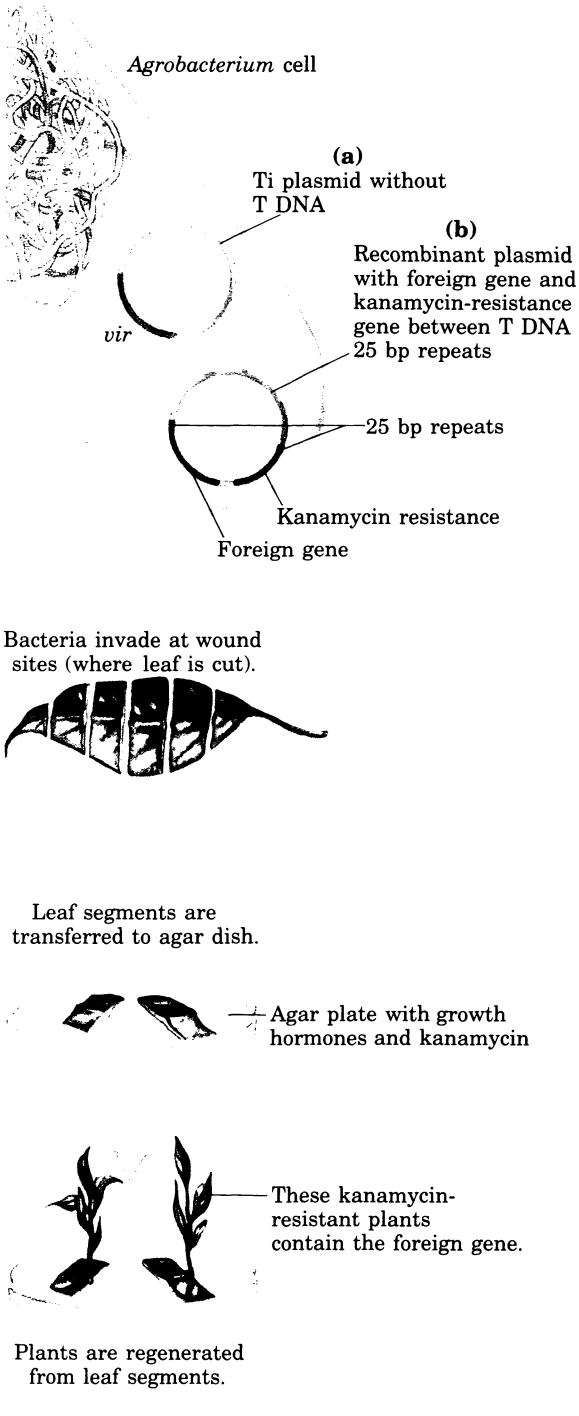


Figure 28–17 A two-plasmid strategy to create a recombinant plant. One plasmid **(a)** is the Ti plasmid, modified so that it lacks T DNA. The other plasmid **(b)** contains the gene of interest (e.g., the gene for the insect-killing protein described in Fig. 28–19) and an antibiotic-resistance (here, kanamycin-resistance) element flanked by T-DNA 25 base pair repeats. This plasmid also contains the replication origin needed for propagation in *Agrobacterium*. The bacteria invade at the site of a wound (the edge of the cut leaf). The *vir* genes on plasmid **(a)** mediate transfer into the plant genome of the segment of plasmid **(b)** flanked by the 25 base pair repeats. New plants are generated when the leaf (with transformed cells) is placed on an agar dish with controlled levels of plant growth hormones and kanamycin. Nontransformed plant cells are killed by the kanamycin. The gene of interest and the antibiotic-resistance element are normally transferred together, and thus plants that grow in the presence of this antibiotic generally contain the gene.

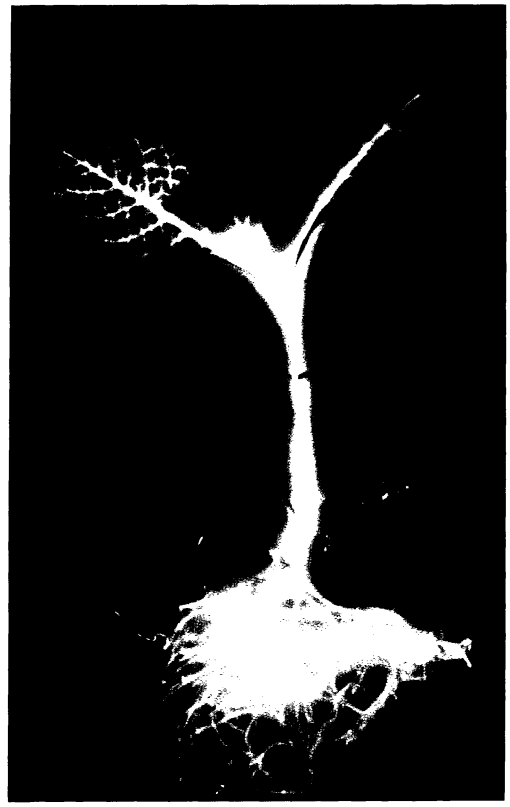


Figure 28–18 A tobacco plant in which the gene for firefly luciferase is expressed. Light was produced after the plant was watered with a solution containing luciferin, the substrate for this light-producing enzyme (see Box 13–3, Fig. 2). Don't expect glow-in-the-dark ornamental plants at your local nursery anytime soon; the light is actually quite weak and this photograph represents a 24 hour exposure. The real point—that this technology allows the introduction of new traits into plants—is nevertheless elegantly made.

plants that are resistant to herbicides, plant viruses, and insect pests (Fig. 28–19). Potential benefits include increases in yields and a reduction in the need for environmentally harmful agricultural chemicals.

Cloning in Animal Cells Points the Way to Human Gene Therapy

The transformation of animal cells with foreign genetic material offers an important mechanism for advancing knowledge about the structure and function of animal genomes, as well as for the generation of animals with new traits. This potential has spawned intensive research efforts that have produced increasingly sophisticated means for cloning in animals.

Most work of this kind requires a source of individual cells. Intact tissues are often difficult to keep alive and manipulate. Fortunately, many types of animal cells can be isolated and grown in a medium in the laboratory if their growth requirements are carefully met. Many cells grown in this kind of **tissue culture** maintain the differentiated properties they had in the whole tissue for weeks or even months.

There are several common methods for introducing DNA into an animal cell. Because no suitable plasmidlike vector is available, transformation requires the integration of the DNA into a host-cell chromosome.

In **spontaneous uptake,** a calcium phosphate–DNA precipitate is taken up by the cells, perhaps by endocytosis. Generally only about one in 10^2 to 10^4 cells is transformed in this procedure. Another direct method that is sometimes more efficient is to make the cells transiently permeable to DNA by exposing them to a brief high-voltage pulse in a technique called **electroporation.**

Figure 28–19 Tomato plants engineered to be resistant to some insect larvae. The plant on the right expresses a gene for a protein toxin, derived from the bacterium *Bacillus thuringiensis*. The protein, introduced by a protocol similar to that described in Fig. 28–17, is toxic to the larvae of some moth species that destroy tomato leaves. This protein, however, is harmless to humans and other organisms. The plant on the left is not genetically altered. Both plants were exposed to equal numbers of larvae. Insect resistance has also been genetically engineered in cotton and other plants.

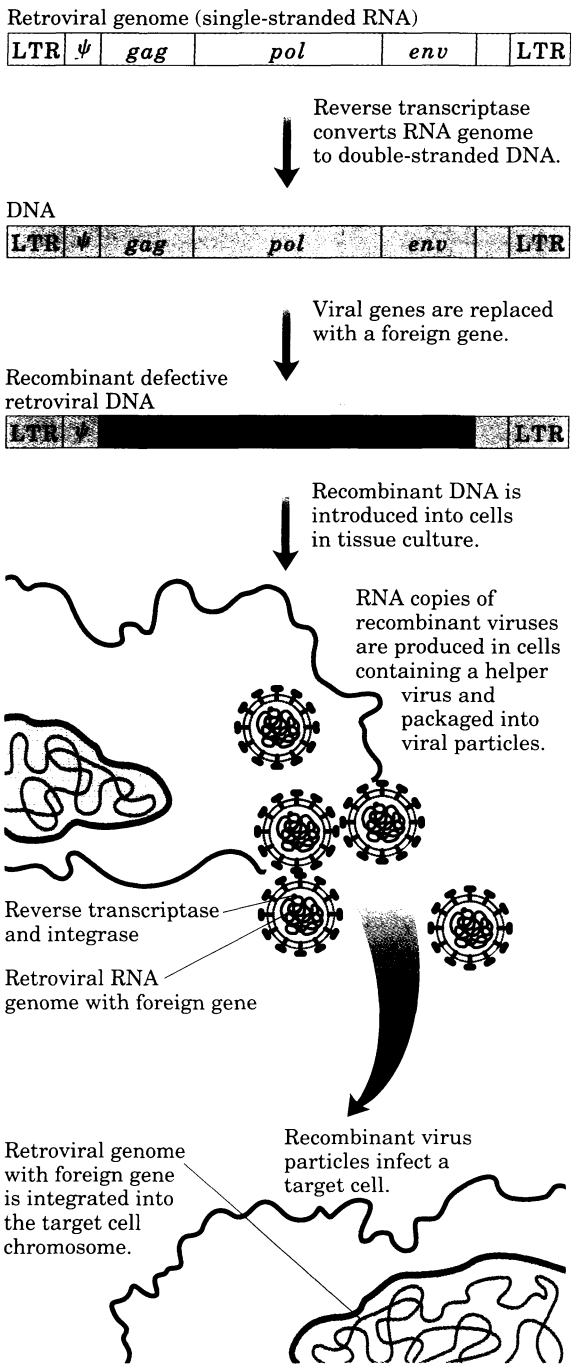


Figure 28–20 Cloning in mammalian cells with retroviral vectors.

Microinjection entails direct injection of DNA into the nucleus of a cell with the aid of a very fine needle. For skilled practitioners this method has a high success rate, but because cells must be injected one by one, the total number that can be treated is small.

A number of eukaryotic viruses sometimes integrate their DNA into a chromosome in the host cell. Some of these, in particular certain retroviruses (p. 882), have been modified to act as **viral vectors** to introduce foreign DNA into mammalian cells. The viruses have their own mechanisms for moving nucleic acid into cells, and transformation by this route can be very efficient. A simplified map of a typical retroviral genome is shown in Figure 28–20. When the virus enters a cell, its RNA genome is converted to DNA by reverse transcriptase and is then integrated into the host genome in a reaction mediated by the viral integrase. The long terminal repeat (LTR) sequences are required for integration of retroviral DNA in the host chromosome (see Fig. 25–30), and the ψ sequence is required to package the viral RNA in viral particles.

The *gag, pol,* and *env* genes of the retroviral genome can be replaced with foreign DNA. This recombinant DNA lacks the genes required for retroviral replication and assembly of viral particles. To assemble viruses with the recombinant genetic information, the DNA must be introduced into cells in tissue culture that are infected with a helper virus, which has the genes to produce virus particles but lacks the ψ sequence required for packaging. Within the cells, the recombinant DNA is transcribed, and the RNA is packaged. The resulting viral particles therefore contain only the recombinant viral RNA and can act as vectors to introduce this RNA into target cells. Viral reverse transcriptase and integrase enzymes (produced by the helper virus) are also packaged in the viral particle and are introduced into the target cells. Once the engineered viral genome is inside a cell, these enzymes create a DNA copy of the viral RNA genome and integrate it into a host chromosome. The integrated recombinant DNA effectively becomes a permanent part of the chromosome because the virus lacks the genes necessary to produce RNA copies of its genome and package them into new virus particles. In most cases the use of recombinant retroviruses is the best method for introducing DNA into large numbers of mammalian cells.

Transformation of animal cells by any of the above techniques is problematic for several reasons. The foreign DNA is generally inserted at chromosomal locations that vary randomly from cell to cell. When the foreign DNA contains a sequence homologous to a sequence on a host chromosome, the introduced DNA is sometimes targeted to that position and integrated by homologous recombination. The nonhomologous integrants still outnumber the targeted ones, however, by factors of 10^2 to 10^5. Some of these integration events are deleterious to the cell because they occur in and disrupt essential genes. Different integration sites can also greatly affect the expression of an integrated gene, because integrated genes are not transcribed equally well everywhere in the genome. Another targeting problem involves the class of cell to be transformed. If germ-line cells are altered, the alteration will be passed on to successive generations of the organism. If somatic cells alone are affected, the alteration will affect only the treated animal.

Despite these problems, this technology has been used extensively to study chromosome structure, as well as the function, regulation, and expression of genes in eukaryotic cells. The successful introduction of

recombinant DNA into an animal can again be illustrated by an experiment that altered an easily observable physical trait. The objective in this case was to alter germ-line cells in mice to create an inheritable change.

Microinjection of DNA into the nuclei of fertilized mouse eggs can produce efficient transformation (chromosomal integration). When the injected eggs are introduced into a female mouse and allowed to develop, the new gene is often expressed in some of the newborn mice. Those in which the germ line has been altered can be identified by testing their offspring. By careful breeding, a mouse line can be established in which all the mice are homozygous for the new gene or genes. Animals permanently altered in this way are called **transgenic.** This technology was used to introduce into mice the human growth hormone gene under the control of an inducible promoter. When fed a diet including the inducer, some of the mice that developed from injected embryos grew to an unusually large size (Fig. 28–21).

If mouse cells can be altered stably by recombinant DNA technology, so then can human cells. Introduction of DNA into human cells offers, for the first time, the potential for treating and even curing human genetic diseases that have been refractory to traditional therapies (Box 28–2, p. 1008). A major technological limitation in these efforts is our overall knowledge of the cellular metabolism that underlies many genetic diseases. As understanding improves, the ability to manipulate cellular metabolism by genetic engineering will improve. A contribution to this understanding may be made by the international project to sequence and map the entire human genome that will proceed through the 1990s. The technology needed to repair genetic defects brings with it the potential for altering human traits. Clearly, we are at a scientific crossroads that has far-reaching implications for the future of humankind.

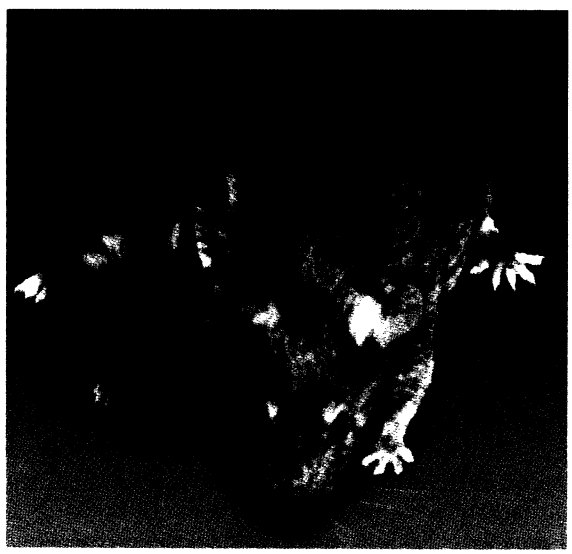


Figure 28–21 Cloning in mice. The gene for human growth hormone was introduced into the genome of the mouse on the right. Expression of the gene resulted in the greatly increased size of the mouse.

BOX 28–2 A Cure for Genetic Diseases

Human gene therapy is a reality in the 1990s. The experiments are going forward with an unprecedented level of oversight and regulation by governments and scientific review committees.

Because of the ethical issues inherent in this work, the objectives laid out by these review committees are narrowly defined; the experiments must meet strict ethical and practical criteria and are intended only to treat severe genetic disorders. First, the research is limited to somatic cells so that a treated individual cannot pass genetic alterations to offspring. Genetic engineering in human germ-line cells conjures up misguided past attempts to "improve" human beings, and evokes a wide range of objections on ethical grounds. Second, the risk to the patient must be outweighed by the potential therapeutic benefit. The inherent risk is exemplified by the possibility of random integration of DNA into a human chromosome leading to inactivation of a gene that regulates cell proliferation, effectively producing a cancer cell. For this reason the targets of the first gene therapy trials are among the most serious genetic diseases. Third, the target diseases must be limited to those that involve a known defect in a single gene, and the normal gene must be cloned and available. Fourth, the disease must involve cells that can be isolated from a patient, altered in tissue culture, and then reimplanted in the patient. This effectively limits the therapy to diseases involving cells of the skin or bone marrow, although some success has been achieved with other tissues such as liver. Fifth, the planned procedures must meet strict safety standards in animal trials before attempts are made with human beings.

The key experimental hurdle is the efficient introduction of DNA into a sufficient number of human cells in a form in which it can be expressed. Because very large numbers of cells must be transformed to have some hope of beneficial effects, research has focused on retroviral vectors (see Fig. 28–20). Expression of introduced genes has been highly variable in animal trials. In many cases, the introduced genes were expressed well in culture, then not at all when the cells were transferred to an animal. New strategies for gene expression are being developed.

Targets of human gene therapy include diseases that result from a functional lack of a single enzyme produced by a single gene (see Table 6–6). These include Lesch–Nyhan syndrome (p. 729), which occurs when hypoxanthine-guanine phosphoribosyltransferase is absent and results in mental retardation and severe behavioral problems. Two forms of severe immune deficiency, which result from a lack of adenosine deaminase (p. 729) or purine nucleoside phosphorylase, are also promising candidates. Work on correcting adenosine deaminase deficiency is already well advanced. Although these two diseases affect only a small number of people, they are very serious (people with severe immune deficiency soon die unless they are kept in a sterile environment), and in the case of adenosine deaminase deficiency the introduction of the missing gene activity into bone marrow cells does appear to have a beneficial effect.

Another effort is focused on new approaches to treating cancer. Immune-system cells known to be associated with tumors, called tumor-infiltrating lymphocytes, have been modified to produce a protein with demonstrated antitumor activity, called tumor necrosis factor (TNF). When reintroduced into a cancer patient the modified cells migrate to the tumor and the TNF they produce facilitates tumor shrinkage. Another approach is to remove and modify tumor cells themselves to produce TNF. When reintroduced into patients the modified cells stimulate the immune system to attack the cancer cells. In animal trials this approach has led to reduction or elimination of tumors and has left the animal immune to the cancer.

Additional genetic disorders that involve treatment of bone marrow cells include the genetic disorders of hemoglobin—sickle-cell anemia (p. 187) and thalassemia. These represent more formidable problems because hemoglobin is the product of more than one gene, and its expression must be limited to a small subfraction of bone marrow cells called the stem cells, which are the progenitors not only of erythrocytes but of granulocytes, macrophages, and platelets. Potential treatment of some more common genetic disorders must await development of methods to remove and replace cells from other tissues. For example, gene therapy for familial hypercholesterolemia, caused by a defect in cholesterol metabolism (p. 679) that can lead to heart attacks at an early age, requires the introduction and expression of a functional LDL receptor in hepatocytes. The prospect of curing such diseases holds great potential for alleviating human suffering.

As this technology advances, however, so does the potential to alter other physical traits. For example, the introduction and expression of a single gene from the mouse Y chromosome (the *Sry* gene) into the genome of female (XX) mouse embryos causes them to develop into male mice. The need for continued societal involvement in debating the issues generated by this technology is obvious.

Recombinant DNA Technology Yields New Products and Choices

The products of recombinant DNA technology range from proteins to engineered organisms. Large amounts of commercially useful proteins can be produced by these techniques. Microorganisms can be designed for special tasks; plants or animals can be engineered with traits that are useful in agriculture. Some products of this technology have been approved for use and many more are in development. During the 1980s genetic engineering was transformed from a promising technology to a multibillion dollar industry. The first commercial product of recombinant DNA technology was human insulin, produced by Eli Lilly and Company and approved for human use by the U.S. Food and Drug Administration in 1982. Hundreds of companies have become involved in product development worldwide. Much of this growth has come in human pharmaceuticals, and some of the major classes of new products are listed in Table 28–4.

Table 28–4 Recombinant DNA products in medicine

Product category	Examples/Uses
Anticoagulants	Tissue plasminogen activator (TPA) activates plasmin, an enzyme involved in dissolving clots; effective in treating heart attack victims.
Blood factors	Factor VIII promotes clotting and is deficient in hemophiliacs. Use of factor VIII produced by recombinant DNA technology eliminates the risks associated with blood transfusions.
Colony stimulating factors	Immune system growth factors that stimulate leukocyte production; used to treat immune deficiencies and to fight infections.
Erythropoietin	Stimulates erythrocyte production; used to treat anemia in patients with kidney disease.
Growth factors	Stimulate differentiation and growth of various cell types; used to promote wound healing.
Human growth hormone	Used to treat dwarfism.
Human insulin	Used to treat diabetes.
Interferons	Interfere with viral reproduction; also used to treat some cancers.
Interleukins	Activate and stimulate different classes of leukocytes; possible uses in wound healing, HIV infection, cancer, immune deficiencies.
Monoclonal antibodies	Extraordinary binding specificity is used in diagnostic tests. Also used to transport drugs, toxins, or radioactive compounds to tumors as a cancer therapy; many other uses.
Superoxide dismutase	Prevents tissue damage from reactive oxygen species when tissues deprived of O_2 for short periods during surgery suddenly have blood flow restored.
Vaccines	Proteins derived from viral coats are as effective in "priming" an immune system as the killed virus more traditionally used for vaccines, but are safer. First developed was the vaccine for hepatitis B.

Erythropoietin is typical of the newer products. Erythropoietin is a protein hormone (M_r 51,000) that stimulates erythrocyte production. People with kidney disease often have a deficiency of this protein, a condition that leads to anemia. Erythropoietin produced by recombinant DNA technology can be used to treat these patients, reducing the need for repeated blood transfusions and their accompanying risks. Approved by the U.S. Food and Drug Administration in 1989, erythropoietin promises to be the most profitable pharmaceutical agent developed by recombinant DNA methods in the 1990s.

Other industrial applications of this technology are likely to continue developing. Enzymes produced by recombinant DNA technology are already used to produce detergents, sugars, and cheese. Engineered proteins are being used as food additives to supplement nutrition, flavor, and fragrance. Microorganisms are being engineered to extract oil and minerals from ground deposits, to digest oil spills, and to detoxify hazardous waste dumps and sewage. Engineered plants with improved resistance to drought, frost, pests, and disease are increasing crop yields and reducing the need for agricultural chemicals. The potential of this technology to benefit humankind and the world environment seems readily apparent yet sometimes hard to define, with the future rendered opaque by our still limited understanding of cellular metabolism and ecology.

Every major new technology comes with associated risks and a potential for unanticipated societal or environmental impact. As with the automobile and nuclear energy, economic, environmental, and ethical considerations will necessarily play an increasingly important role in determining how recombinant DNA technology is applied. One harbinger of this new relationship between biochemistry and society has been the debate in the United States and elsewhere over bovine growth hormone, which is used to increase milk production. In addition to some potential for added stress on the animals and concerns among consumers about the safety of the milk for human use, increasing milk production when a surplus already exists may have the effect of lowering prices and imposing economic hardship on dairy farmers.

Other issues raised by this technology promise to have a much broader impact. A particularly clear example can be seen in an array of new diagnostic procedures based on recombinant DNA technology. These are greatly increasing our ability to detect genetic diseases in an individual, often many years before the onset of symptoms or even before birth. The same technology that makes it possible to identify a criminal (Box 28–1) may be used to test individuals for a genetic predisposition to conditions such as Alzheimer's disease, hypercholesterolemia, asthma, and alcoholism. This information will permit better and earlier treatments, but the same information could be used to restrict individual access to health insurance (and thus health care), life insurance, and even certain jobs. The questions of who will have access to this information and how it will be used will grow in importance as more tests become widely available.

These are only some of the more straightforward examples. Release of genetically engineered organisms into the environment carries with it a level of risk that is sometimes difficult to evaluate. Human gene therapy (Box 28–2), with all of its promise, doubtless will present society with ethical dilemmas not yet anticipated. Issues of this kind must in the end foster a closer and more productive collaboration between science and the society it serves, as well as higher levels of scientific literacy in the general public, as we move toward the twenty-first century.

Summary

The study of gene structure and function has been greatly facilitated by recombinant DNA technology. The isolation of a gene from a large chromosome requires methods for cutting and joining DNA fragments, the availability of small DNA vectors that can replicate autonomously and into which the gene can be inserted, methods to introduce the vector with its foreign DNA into a cell in which it can be propagated to form clones, and methods to identify the cells containing the DNA of interest. Advances in this technology are revolutionizing many aspects of medicine, agriculture, and other industries.

The first organism used for DNA cloning was *E. coli*. Bacterial restriction endonucleases and DNA ligases provide the most important instruments for cutting DNA at specific sequences and joining DNA fragments. Bacterial cloning vectors include plasmids, bacteriophages, and cosmids. These permit the cloning of DNA fragments of different size ranges. In each case the vectors provide a replication origin for propagation in the bacterial host, and a selectable genetic element such as antibiotic resistance to facilitate the identification of cells harboring the recombinant vector. DNA is introduced into cells in viral vectors or by artificial methods that make the cell wall permeable.

The first step in cloning a gene is often the construction of a DNA library that includes fragments representing most of the genome of a given species. The library can be limited to expressed genes by cloning only the complementary DNA copies of isolated mRNAs to make a cDNA library. A specific segment of DNA can be amplified and cloned using the polymerase chain reaction. Clones containing a specific gene in a large library can be detected by hybridization with a radioactive probe containing the complementary nucleotide sequence.

Expression vectors provide the DNA sequences required for transcription, translation, and regulation of cloned genes. They allow the production of large amounts of cloned proteins for research and commercial purposes. Cloned genes also can be altered by site-directed mutagenesis, which is useful in studies of protein structure and function.

Yeast is sometimes used for cloning eukaryotic DNA and it has many of the same advantages as *E. coli*. Methods for cloning in plants and animals are producing a variety of organisms with altered traits. Plants that are resistant to disease, insects, herbicides, and drought are being produced with the aid of a natural gene transfer process promoted by the Ti plasmid of the parasitic soil bacterium *Agrobacterium tumefaciens*. Engineered DNA can be introduced into animal cells by microinjection or retroviral vectors. Such procedures have produced mice with new inheritable genetic traits. The technology extends to humans, and human gene therapy is now directed at treating human genetic diseases.

Further Reading

General

Hackett, P.B., Fuchs, J.A., & Messing, J.W. (1988) *An Introduction to Recombinant DNA Techniques,* 2nd edn, The Benjamin/Cummings Publishing Company, Menlo Park, CA.

Jackson, D.A., Symons, R.H., & Berg, P. (1972) Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **69,** 2904–2909.

The first recombinant DNA experiment linking DNA from two different organisms.

Lobban, P.E. & Kaiser, A.D. (1973) Enzymatic end-to-end joining of DNA molecules. *J. Mol. Biol.* **78,** 453–471.

Report of the first recombinant DNA experiment.

Sambrook, J., Fritsch, E.F., & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual,* 2nd edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

In addition to detailed protocols for a wide range of techniques, this three-volume set includes much useful background information on the biological, chemical, and physical principles underlying each technique.

Libraries and Gene Isolation

Arnheim, N. & Levenson, C.H. (1990) Polymerase chain reaction. *Chem. Eng. News* **68** (October 1), 36–47.

A broad overview of the technique and applications.

Arnheim, N. & Erlich, H. (1992) Polymerase chain reaction strategy. *Annu. Rev. Biochem.* **61,** 131–156.

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

This issue of Science *contains several other good articles on other aspects of biotechnology.*

Neufeld, P.J. & Colman, N. (1990) When science takes the witness stand. *Sci. Am.* **262** (May), 46–53.

Potential and problems of DNA fingerprinting.

Pääbo, S., Higuchi, R.G., & Wilson, A.C. (1989) Ancient DNA and the polymerase chain reaction. The emerging field of molecular archaeology. *J. Biol. Chem.* **264,** 9709–9712.

Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98,** 503–517.

The paper that introduced the Southern hybridization method.

Thornton, J.I. (1989) DNA profiling: new tool links evidence to suspects with high certainty. *Chem. Eng. News* **67** (November 20), 18–30.

Products of Recombinant DNA Technology

Bailey, J.E. (1991) Toward a science of metabolic engineering. *Science* **252,** 1668–1675.

An overview of efforts to reengineer entire metabolic pathways in microorganisms for commercial purposes.

Bains, W. (1989) Disease, DNA and diagnosis. *New Scientist* **122** (May 6), 48–51.

Botstein, D. & Shortle, D. (1985) Strategies and applications of *in vitro* mutagenesis. *Science* **229,** 1193–1201.

Buck, K. (1989) Brave new botany. *New Scientist* **122** (June 3), 50–55.

Prospects for genetic engineering in plants.

Culver, K.W., Osborne, W.R.A., Miller, A.D., Fleisher, T.A., Berger, M., Anderson, W.F., & Blaese, R.M. (1991) Correction of ADA deficiency in human T lymphocytes using retroviral-mediated gene transfer. *Transplant Proc.* **23,** 170–171.

Gene therapy of adenosine deaminase deficiency.

Hooykaas, P.J.J. & Schilperoort, R.A. (1985) The Ti-plasmid of *Agrobacterium tumefaciens:* a natural genetic engineer. *Trends Biochem. Sci.* **10,** 307–309.

Johnson, I.S. (1983) Human insulin from recombinant DNA technology. *Science* **219,** 632–637.

Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P., & Lovell-Badge, R. (1991) Male development of chromosomally female mice transgenic for *Sry*. *Nature* **351,** 117–121.

Recombinant DNA technology is used to demonstrate that a single gene directs the development of chromosomally female mice into males.

McLachlin, J.R., Cornetta, K., Eglitis, M.A., & Anderson, W.F. (1990) Retroviral-mediated gene transfer. *Prog. Nucleic Acid Res. Mol. Biol.* **38,** 91–135.

Murray, T.H. (1991) Ethical issues in human genome research. *FASEB J.* **5,** 55–60.

This issue also contains a number of other useful papers on the human genome project.

Palmiter, R.D., Brinster, R.L., Hammer, R.E., Trumbauer, M.E., Rosenfeld, M.G., Birnberg, N.C., & Evans, R.M. (1982) Dramatic growth of mice that develop from eggs microinjected with metallothionein–growth hormone fusion genes. *Nature* **300,** 611–615.

A description of how to make giant mice.

Sulston, J., Du, Z., Thomas, K., Wilson, R., Hillier, L., Staden, R., Halloran, N., Green, P., Thierry-Mieg, J., Qiu, L., Dear, S., Coulson, A., Craxton, M., Durbin, R., Berks, M., Metzstein, M., Hawkins, T., Ainscough, R., & Waterston, R. (1992) The *C. elegans* genome sequencing project: a beginning. *Nature* **356,** 37–41.

A description of strategies being developed to sequence large genomes.

The telltale gene. *Consumer Reports* **55** (July 1990), 483–488.

A good summary of new genetic screening techniques based on recombinant DNA, with a focus on ethical dilemmas posed by the technology.

Thompson, J. & Donkersloot, J.A. (1992) *N*-(Carboxyalkyl)amino acids: occurrence, synthesis, and functions. *Annu. Rev. Biochem.* **61,** 517–557.

A good summary of the structure and biological functions of opines.

Verma, I.M. (1990) Gene therapy. *Sci. Am.* **263** (November), 68–84.

Watson, J.D. (1990) The human genome project: past, present, and future. *Science* **248,** 44–48.

Weatherall, D.J. (1991) Gene therapy in perspective. *Nature* **349,** 275–276.

Westphal, H. (1989) Transgenic mammals and biotechnology. *FASEB J.* **3,** 117–120.

Zambryski, P., Tempe, J., & Schell, J. (1989) Transfer and function of T-DNA genes from *Agrobacterium* Ti and Ri plasmids in plants. *Cell* **56,** 193–201.

Problems

1. *Cloning*

(a) Draw the structure of the end of a linear DNA fragment that was produced by an *Eco*RI restriction digest (include those sequences remaining from the *Eco*RI recognition sequence).

(b) Draw the structure resulting from the reaction of this end sequence with DNA polymerase I and the four deoxynucleoside triphosphates.

(c) Draw the sequence produced at the junction if two ends with the structure derived in (b) are ligated.

(d) Draw the structure produced if the structure derived in (a) is treated with a nuclease that degrades only single-stranded DNA.

(e) Draw the sequence of the junction produced if an end with structure (b) is ligated to an end with structure (d).

(f) Draw the structure of the end of a linear DNA fragment that was produced by a *Pvu*II restriction digest (as in (a)).

(g) Draw the sequence of the junction produced if an end with structure (b) is ligated to an end with structure (f).

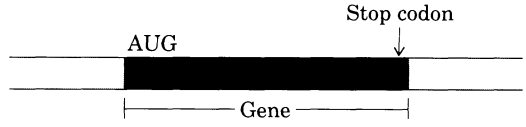
(h) Suppose you can synthesize a short duplex DNA fragment with any sequence you desire. With such a synthetic fragment and the procedures described in (a) through (g), design a protocol that will remove an *Eco*RI restriction site from a DNA molecule and incorporate a new *Bam*HI restriction site at approximately the same location. (Hint: See Fig. 28–3.)

(i) Design four different short synthetic DNA fragments that would permit ligation of structure (a) with a DNA fragment produced by a *Pst*I restriction digest. In one of these synthetic fragments, design the sequence so that the final junction contains the recognition sequences for both *Eco*RI and *Pst*I. In the second and third synthetic fragments, design the sequence so that the junction contains only the *Eco*RI or the *Pst*I recognition sequence, respectively. Design the sequence of the fourth fragment so that neither the *Eco*RI nor the *Pst*I sequence appears in the junction.

2. *Selecting for Recombinant Plasmids* When cloning a foreign DNA fragment into a plasmid it is often useful to insert the fragment at a site that interrupts a selectable marker (such as the tetracycline-resistance element of pBR322). The loss of function of the interrupted gene can be used to identify clones containing recombinant plasmids with foreign DNA. With a cosmid it is unnecessary to do this, yet one can easily distinguish cosmids that incorporate large foreign DNA fragments from those that do not. How are these recombinant cosmids identified?

3. *DNA Cloning* The plasmid cloning vector pBR322 (see Fig. 28–5) is cleaved with the restriction endonuclease *Eco*RI. An isolated DNA fragment from a eukaryotic genome (also produced by *Eco*RI cleavage) is added to the prepared vector and ligated. The mixture of ligated DNAs is then used to transform bacteria, and plasmid-containing bacteria are selected by growth in the presence of tetracycline. In addition to the desired recombinant plasmid, what other types of plasmids might be found among the transformed bacteria that are tetracycline resistant?

4. *Expressing a Cloned Gene* You have isolated a plant gene that encodes a protein in which you are interested. On the drawing below, indicate sequences or sites that you will need to get this gene transcribed, translated, and regulated in *E. coli*.



5. *Identifying the Gene for a Protein with a Known Amino Acid Sequence* Design a DNA probe that would allow you to identify the gene for a protein with the following amino-terminal amino acid sequence. The probe should be 18 to 20 nucleotides long, a size that provides adequate specificity if there is sufficient homology between the probe and the gene.

$H_3\overset{+}{N}$-Ala–Pro–Met–Thr–Trp–Tyr–Cys–Met–Asp–Trp–Ile–Ala–Gly–Gly–Pro–Trp–Phe–Arg–Lys–Asn–Thr–Lys– – –

6. *Cloning in Plants* The strategy outlined in Figure 28–17 employs *Agrobacterium* containing two separate plasmids. Suggest a reason why the sequences on the two plasmids are not combined on one plasmid.

7. *Cloning in Mammals* The retroviral vectors described in Figure 28–20 make it possible to integrate foreign DNA efficiently into a mammalian genome. Explain how these vectors, which lack genes for replication and viral packaging *(gag, pol, env)*, are assembled into infectious viral particles. Suggest why it is important that these vectors lack the replication and packaging genes.

Glossary

absolute configuration: The configuration of four different substituent groups around an asymmetric carbon atom, in relation to D- and L-glyceraldehyde.

absorption: Transport of the products of digestion from the intestinal tract into the blood.

acceptor control: The regulation of the rate of respiration by the availability of ADP as phosphate group acceptor.

accessory pigments: Visible light–absorbing pigments (carotenoids, xanthophyll, and phycobilins) in plants and photosynthetic bacteria that complement chlorophylls in trapping energy from sunlight.

acidosis: A metabolic condition in which the capacity of the body to buffer H^+ is diminished; usually accompanied by decreased blood pH.

actin: A protein making up the thin filaments of muscle; also an important component of the cytoskeleton of many eukaryotic cells.

activation energy ($\Delta G^{\ddagger}$): The amount of energy (in joules) required to convert all the molecules in 1 mole of a reacting substance from the ground state to the transition state.

activator: (1) A DNA-binding protein that positively regulates the expression of one or more genes; that is, transcription rates increase when an activator is bound to the DNA. (2) A positive modulator of an allosteric enzyme.

active site: The region of an enzyme surface that binds the substrate molecule and catalytically transforms it; also known as the catalytic site.

active transport: Energy-requiring transport of a solute across a membrane in the direction of increasing concentration.

activity: The true thermodynamic activity or potential of a substance, as distinct from its molar concentration.

activity coefficient: The factor by which the numerical value of the concentration of a solute must be multiplied to give its true thermodynamic activity.

acyl phosphate: Any molecule with the general chemical form $R-\overset{\overset{O}{\|}}{C}-OPO_3^{2-}$.

adenosine 3′,5′-cyclic monophosphate: See cyclic AMP.

adenosine diphosphate: See ADP.

adenosine triphosphate: See ATP.

adipocyte: An animal cell specialized for the storage of fats (triacylglycerols).

adipose tissue: Connective tissue specialized for the storage of large amounts of triacylglycerols.

ADP (adenosine diphosphate): A ribonucleoside 5′-diphosphate serving as phosphate group acceptor in the cell energy cycle.

aerobe: An organism that lives in air and uses oxygen as the terminal electron acceptor in respiration.

aerobic: Requiring or occurring in the presence of oxygen.

alcohol fermentation: The anaerobic conversion of glucose to ethanol via glycolysis. See also fermentation.

aldose: A simple sugar in which the carbonyl carbon atom is an aldehyde; that is, the carbonyl carbon is at one end of the carbon chain.

alkaloids: Nitrogen-containing organic compounds of plant origin; often basic, and having intense biological activity.

alkalosis: A metabolic condition in which the capacity of the body to buffer OH^- is diminished; usually accompanied by an increase in blood pH.

allosteric enzyme: A regulatory enzyme, with catalytic activity modulated by the noncovalent binding of a specific metabolite at a site other than the active site.

allosteric site: The specific site on the surface of an allosteric enzyme molecule to which the modulator or effector molecule is bound.

α helix: A helical conformation of a polypeptide chain, usually right-handed, with maximal intrachain hydrogen bonding; one of the most common secondary structures in proteins.

Ames test: A simple bacterial test for carcinogens, based on the assumption that carcinogens are mutagens.

amino acid activation: ATP-dependent enzymatic esterification of the carboxyl group of an amino acid to the 3′-hydroxyl group of its corresponding tRNA.

amino acids: α-Amino-substituted carboxylic acids, the building blocks of proteins.

amino-terminal residue: The only amino acid residue in a polypeptide chain with a free α-amino group; defines the amino terminus of the polypeptide.

aminoacyl-tRNA: An aminoacyl ester of a tRNA.

aminoacyl-tRNA synthetases: Enzymes that catalyze synthesis of an aminoacyl-tRNA at the expense of ATP energy.

aminotransferases: Enzymes that catalyze the transfer of amino groups from α-amino to α-keto acids; also called transaminases.

ammonotelic: Excreting excess nitrogen in the form of ammonia.

amphibolic pathway: A metabolic pathway used in both catabolism and anabolism.

amphipathic: Containing both polar and nonpolar domains.

amphoteric: Capable of donating and accepting protons, thus able to serve as an acid or a base.

anabolism: The phase of intermediary metabolism concerned with the energy-requiring biosynthesis of cell components from smaller precursors.

anaerobe: An organism that lives without oxygen. Obligate anaerobes die when exposed to oxygen.

anaerobic: Occurring in the absence of air or oxygen.

anaplerotic reaction: An enzyme catalyzed reaction that can replenish the supply of intermediates in the citric acid cycle.

angstrom (Å): A unit of length (10^{-8} cm) used to indicate molecular dimensions.

anhydride: The product of the condensation of two carboxyl or phosphate groups in which the elements of water are eliminated to form a compound with the general structure R—X(=O)—O—X(=O)—R, where X is either carbon or phosphorus.

anion-exchange resin: A polymeric resin with fixed cationic groups used in the chromatographic separation of anions.

anomers: Two stereoisomers of a given sugar that differ only in the configuration about the carbonyl (anomeric) carbon atom.

antibiotic: One of many different organic compounds that are formed and secreted by various species of microorganisms and plants, are toxic to other species, and presumably have a defensive function.

antibody: A defense protein synthesized by the immune system of vertebrates. See also immunoglobulin.

anticodon: A specific sequence of three nucleotides in a tRNA, complementary to a codon for an amino acid in an mRNA.

antigen: A molecule capable of eliciting the synthesis of a specific antibody in vertebrates.

antiparallel: Describing two linear polymers that are opposite in polarity or orientation.

antiport: Cotransport of two solutes across a membrane in opposite directions.

apoenzyme: The protein portion of an enzyme, exclusive of any organic or inorganic cofactors or prosthetic groups that might be required for catalytic activity.

apolipoprotein: The protein component of a lipoprotein.

asymmetric carbon atom: A carbon atom that is covalently bonded to four different groups and thus may exist in two different tetrahedral configurations.

ATP (adenosine triphosphate): A ribonucleoside 5′-triphosphate functioning as a phosphate group donor in the cell energy cycle; carries chemical energy between metabolic pathways by serving as a shared intermediate coupling endergonic and exergonic reactions.

ATP synthase: An enzyme complex that forms ATP from ADP and phosphate during oxidative phosphorylation in the inner mitochondrial membrane or the bacterial plasma membrane, and during photophosphorylation in chloroplasts.

ATPase: An enzyme that hydrolyzes ATP to yield ADP and phosphate; usually coupled to some process requiring energy.

attenuator: An RNA sequence involved in regulating the expression of certain genes; functions as a transcription terminator.

autotroph: An organism that can synthesize its own complex molecules from very simple carbon and nitrogen sources, such as carbon dioxide and ammonia.

auxin: A plant growth hormone.

auxotrophic mutant (auxotroph): A mutant organism defective in the synthesis of a given biomolecule, which must therefore be supplied for the organism's growth.

Avogadro's number (N): The number of molecules in a gram molecular weight (a mole) of any compound (6.02×10^{23}).

back mutation: A mutation that causes a mutant gene to regain its wild-type base sequence.

bacteriophage (phage): A virus capable of replicating in a bacterial cell.

basal metabolic rate: The rate of oxygen consumption by an animal's body at complete rest, long after a meal.

base pair: Two nucleotides in nucleic acid chains that are paired by hydrogen bonding of their bases; for example, A with T or U, and G with C.

β conformation: An extended, zigzag arrangement of a polypeptide chain; a common secondary structure in proteins.

β oxidation: Oxidative degradation of fatty acids into acetyl-CoA by successive oxidations at the β-carbon atom.

bilayer: A double layer of oriented amphipathic lipid molecules, forming the basic structure of biological membranes. The hydrocarbon tails face inward to form a continuous nonpolar phase.

bile salts: Amphipathic steroid derivatives with detergent properties, participating in digestion and absorption of lipids.

binding energy: The energy derived from noncovalent interactions between enzyme and substrate or receptor and ligand.

biocytin: The conjugate amino acid residue arising from covalent attachment of biotin, through an amide linkage, to a Lys residue.

biomolecule: An organic compound normally present as an essential component of living organisms.

biopterin: An enzymatic cofactor derived from pterin and involved in certain oxidation–reduction reactions.

biosphere: All the living matter on or in the earth, the seas, and the atmosphere.

biotin: A vitamin; an enzymatic cofactor involved in carboxylation reactions.

bond energy: The energy required to break a bond.

branch migration: Movement of the branch point in branched DNA formed from two DNA molecules with identical sequences. See also Holliday intermediate.

buffer: A system capable of resisting changes in pH, consisting of a conjugate acid–base pair in which the ratio of proton acceptor to proton donor is near unity.

calorie: The amount of heat required to raise the temperature of 1.0 g of water from 14.5 to 15.5 °C. One calorie (cal) equals 4.18 joules (J).

Calvin cycle: The cyclic pathway used by plants to fix carbon dioxide and produce triose phosphates.

cAMP: See cyclic AMP.

CAP: See catabolite gene activator protein.

capsid: The protein coat of a virion or virus particle.

carbanion: A negatively charged carbon atom.

carbocation: A positively charged carbon atom; also called a carbonium ion.

carbon fixation reactions: In photosynthetic cells, the light-independent enzymatic reactions involved in the synthesis

of glucose from CO_2, ATP, and NADPH; also known as the dark reactions.

carboxyl-terminal residue: The only amino acid residue in a polypeptide chain with a free α-carboxyl group; defines the carboxyl terminus of the polypeptide.

carotenoids: Lipid-soluble photosynthetic pigments made up of isoprene units.

catabolism: The phase of intermediary metabolism concerned with the energy-yielding degradation of nutrient molecules.

catabolite gene activator protein (CAP): A specific regulatory protein that controls initiation of transcription of the genes producing the enzymes required for a bacterial cell to use some other nutrient when glucose is lacking.

catalytic site: See active site.

catecholamines: Hormones, such as epinephrine, that are amino derivatives of catechol.

cation-exchange resin: An insoluble polymer with fixed negative charges; used in the chromatographic separation of cationic substances.

cDNA: See complementary DNA.

central dogma: The organizing principle of molecular biology: genetic information flows from DNA to RNA to protein.

centromere: A specialized site within a chromosome, serving as the attachment point for the mitotic or meiotic spindle.

cerebroside: Sphingolipid containing one sugar residue as a head group.

channeling: The direct transfer of a reaction product (common intermediate) from the active site of one enzyme to the active site of a different enzyme catalyzing the next step in a sequential pathway.

chemiosmotic coupling: Coupling of ATP synthesis to electron transfer via an electrochemical H^+ gradient across a membrane.

chemotaxis: A cell's sensing of and movement toward, or away from, a specific chemical agent.

chemotroph: An organism that obtains energy by metabolizing organic compounds derived from other organisms.

chiral compound: A compound that contains an asymmetric center (chiral atom or chiral center) and thus can occur in two nonsuperimposable mirror-image forms (enantiomers).

chlorophylls: A family of green pigments functioning as receptors of light energy in photosynthesis; magnesium-porphyrin complexes.

chloroplasts: Chlorophyll-containing photosynthetic organelles in some eukaryotic cells.

chromatin: A filamentous complex of DNA, histones, and other proteins, constituting the eukaryotic chromosome.

chromatography: A process in which complex mixtures of molecules are separated by many repeated partitionings between a flowing (mobile) phase and a stationary phase.

chromosome: A single large DNA molecule and its associated proteins, containing many genes; stores and transmits genetic information.

chylomicron: A plasma lipoprotein consisting of a large droplet of triacylglycerols stabilized by a coat of protein and phospholipid; carries lipids from the intestine to the tissues.

cis and trans isomers: See geometric isomers.

cistron: A unit of DNA or RNA corresponding to one gene.

citric acid cycle: A cyclic system of enzymatic reactions for the oxidation of acetyl residues to carbon dioxide, in which formation of citrate is the first step; also known as the Krebs cycle or tricarboxylic acid cycle.

clones: The descendants of a single cell.

cloning: The production of large numbers of identical DNA molecules or cells from a single ancestral DNA molecule or cell.

closed system: A system that exchanges neither matter nor energy with the surroundings. See also system.

cobalamin: See coenzyme B_{12}.

codon: A sequence of three adjacent nucleotides in a nucleic acid that codes for a specific amino acid.

coenzyme: An organic cofactor required for the action of certain enzymes; often contains a vitamin as a component.

coenzyme A: A pantothenic acid–containing coenzyme serving as an acyl group carrier in certain enzymatic reactions.

coenzyme B_{12}: An enzymatic cofactor derived from the vitamin cobalamin, involved in certain types of carbon skeletal rearrangements.

cofactor: An inorganic ion or a coenzyme required for enzyme activity.

cognate: Describing two biomolecules that normally interact; for example, an enzyme and its normal substrate, or a receptor and its normal ligand.

cohesive ends: See sticky ends.

cointegrate: An intermediate in the migration of certain DNA transposons in which the donor DNA and target DNA are covalently attached.

colligative properties: Properties of solutions that depend on the number of solute particles per unit volume; for example, freezing-point depression.

common intermediate: A chemical compound common to two chemical reactions, as a product of one and a reactant in the other.

competitive inhibition: A type of enzyme inhibition reversed by increasing the substrate concentration; a competitive inhibitor generally competes with the normal substrate or ligand for a protein's binding site.

complementary: Having a molecular surface with chemical groups arranged to interact specifically with chemical groups on another molecule.

complementary DNA (cDNA): A DNA used in DNA cloning, usually made by reverse transcriptase; complementary to a given mRNA.

configuration: The spatial arrangement of an organic molecule that is conferred by the presence of either (1) double bonds, about which there is no freedom of rotation, or (2) chiral centers, around which substituent groups are arranged in a specific sequence. Configurational isomers cannot be interconverted without breaking one or more covalent bonds.

conformation: The spatial arrangement of substituent groups that are free to assume different positions in space, without breaking any bonds, because of the freedom of bond rotation.

conformation, β: See β conformation.

conjugate acid–base pair: A proton donor and its corresponding deprotonated species; for example, acetic acid (donor) and acetate (acceptor).

conjugate redox pair: An electron donor and its corresponding electron acceptor form; for example, Cu^+ (donor) and Cu^{2+} (acceptor), or NADH (donor) and NAD^+ (acceptor).

conjugated protein: A protein containing one or more prosthetic groups.

consensus sequence: A DNA or amino acid sequence consisting of the residues that occur most commonly at each position within a set of similar sequences.

conservative substitution: Replacement of an amino acid residue in a polypeptide by another residue with similar properties; for example, substitution of Glu by Asp.

constitutive enzymes: Enzymes required at all times by a cell and present at some constant level; for example, many enzymes of the central metabolic pathways. Sometimes called "housekeeping enzymes."

corticosteroids: Steroid hormones formed by the adrenal cortex.

cosmid: A cloning vector, used for cloning large DNA fragments; generally contains segments derived from bacteriophages and various plasmids.

cotransport: The simultaneous transport, by a single transporter, of two solutes across a membrane. See antiport, symport.

coupled reactions: Two chemical reactions that have a common intermediate and thus a means of energy transfer from one to the other.

covalent bond: A chemical bond that involves sharing of electron pairs.

cristae: Infoldings of the inner mitochondrial membrane.

cyclic AMP (cAMP): A second messenger within cells; its formation by adenylate cyclase is stimulated by certain hormones or other molecular signals.

cyclic electron flow: In chloroplasts, the light-induced flow of electrons originating from and returning to photosystem I.

cyclic photophosphorylation: ATP synthesis driven by cyclic electron flow through photosystem I.

cytochromes: Heme proteins serving as electron carriers in respiration, photosynthesis, and other oxidation–reduction reactions.

cytokinesis: The final separation of daughter cells following mitosis.

cytoplasm: The portion of a cell's contents outside the nucleus but within the plasma membrane; includes organelles such as mitochondria.

cytoskeleton: The filamentous network providing structure and organization to the cytoplasm; includes actin filaments, microtubules, and intermediate filaments.

cytosol: The continuous aqueous phase of the cytoplasm, with its dissolved solutes; excludes the organelles such as mitochondria.

dalton: The weight of a single hydrogen atom (1.66×10^{-24} g).

dark reactions: See carbon fixation reactions.

de novo pathway: Pathway for synthesis of a biomolecule, such as a nucleotide, from simple precursors; as distinct from a salvage pathway.

deamination: The enzymatic removal of amino groups from biomolecules such as amino acids or nucleotides.

degenerate code: A code in which a single element in one language is specified by more than one element in a second language.

dehydrogenases: Enzymes catalyzing the removal of pairs of hydrogen atoms from their substrates.

deletion mutation: A mutation resulting from the deletion of one or more nucleotides from a gene or chromosome.

denaturation: Partial or complete unfolding of the specific native conformation of a polypeptide chain, protein, or nucleic acid.

denatured protein: A protein that has lost its native conformation by exposure to a destabilizing agent such as heat or detergent.

deoxyribonucleic acid: See DNA.

deoxyribonucleotides: Nucleotides containing 2-deoxy-D-ribose as the pentose component.

desaturases: Enzymes that catalyze the introduction of double bonds into the hydrocarbon portion of fatty acids.

desolvation: In aqueous solution, the release of bound water surrounding a solute.

dextrorotatory isomer: A stereoisomer that rotates the plane of plane-polarized light clockwise.

diabetes mellitus: A metabolic disease resulting from insulin deficiency; characterized by a failure in glucose transport from the blood into cells at normal glucose concentrations.

dialysis: Removal of small molecules from a solution of a macromolecule, by allowing them to diffuse through a semipermeable membrane into water.

differential centrifugation: Separation of cell organelles or other particles of different size by their different rates of sedimentation in a centrifugal field.

differentiation: Specialization of cell structure and function during embryonic growth and development.

diffusion: The net movement of molecules in the direction of lower concentration.

digestion: Enzymatic hydrolysis of major nutrients in the gastrointestinal system to yield their simpler components.

diploid: Having two sets of genetic information; describing a cell with two chromosomes of each type.

dipole: A molecule having both positive and negative charges.

diprotic acid: An acid having two dissociable protons.

disaccharide: A carbohydrate consisting of two covalently joined monosaccharide units.

dissociation constant: (1) An equilibrium constant (K_d) for the dissociation of a complex of two or more biomolecules into its components; for example, dissociation of a substrate from an enzyme. (2) The dissociation constant (K_a) of an acid, describing its dissociation into its conjugate base and a proton.

disulfide bridge: A covalent cross link between two polypeptide chains formed by a cystine residue (two Cys residues).

DNA (deoxyribonucleic acid): A polynucleotide having a specific sequence of deoxyribonucleotide units covalently joined through 3′,5′-phosphodiester bonds; serves as the carrier of genetic information.

DNA chimera: A DNA containing genetic information derived from two different species.

DNA cloning: See cloning.

DNA library: A random collection of cloned DNA fragments that includes all or most of the genome of a given organism; also called a genomic library.

DNA ligase: An enzyme that creates a phosphodiester bond between the 3′ end of one DNA segment and the 5′ end of another.

DNA looping: The interaction of proteins bound at distant sites on a DNA molecule so that the intervening DNA forms a loop.

DNA polymerase: An enzyme that catalyzes template-dependent synthesis of DNA from its deoxyribonucleoside 5′-triphosphate precursors.

DNA replicase system: The entire complex of enzymes and specialized proteins required in biological DNA replication.

DNA supercoiling: The coiling of DNA upon itself, generally as a result of bending, underwinding, or overwinding of the DNA helix.

DNA transposition: See transposition.

domain: A distinct structural unit of a polypeptide; domains may have separate functions and may fold as independent, compact units.

double helix: The natural coiled conformation of two complementary, antiparallel DNA chains.

double-reciprocal plot: A plot of $1/V_0$ versus $1/[S]$, which allows a more accurate determination of V_{max} and K_m than a plot of V_0 versus [S]; also called the Lineweaver–Burk plot.

E_0': See standard reduction potential.

E. coli (Escherichia coli): A common bacterium found in the small intestine of vertebrates; the most well-studied organism.

electrochemical gradient: The sum of the gradients of concentration and of electric charge of an ion across a membrane; the driving force for oxidative phosphorylation and photophosphorylation.

electrochemical potential: The energy required to maintain a separation of charge and of concentration across a membrane.

electrogenic: Contributing to an electrical potential across a membrane.

electron acceptor: A substance that receives electrons in an oxidation–reduction reaction.

electron carrier: A protein, such as a flavoprotein or a cytochrome, that can reversibly gain and lose electrons; functions in the transfer of electrons from organic nutrients to oxygen or some other terminal acceptor.

electron donor: A substance that donates electrons in an oxidation–reduction reaction.

electron transfer: Movement of electrons from substrates to oxygen via the carriers of the respiratory (electron transfer) chain.

electrophile: An electron-deficient group with a strong tendency to accept electrons from an electron-rich group (nucleophile).

electrophoresis: Movement of charged solutes in response to an electrical field; often used to separate mixtures of ions, proteins, or nucleic acids.

elongation factors: Specific proteins required in the elongation of polypeptide chains by ribosomes.

eluate: The effluent from a chromatographic column.

enantiomers: Stereoisomers that are nonsuperimposable mirror images of each other.

end-product inhibition: See feedback inhibition.

endergonic reaction: A chemical reaction that consumes energy (that is, for which ΔG is positive).

endocrine glands: Groups of cells specialized to synthesize hormones and secrete them into the blood to regulate other types of cells.

endocytosis: The uptake of extracellular material by its inclusion within a vesicle (endosome) formed by an invagination of the plasma membrane.

endonuclease: An enzyme that hydrolyzes the interior phosphodiester bonds of a nucleic acid; that is, it acts at points other than the terminal bonds.

endoplasmic reticulum: An extensive system of double membranes in the cytoplasm of eukaryotic cells; it encloses secretory channels and is often studded with ribosomes (rough endoplasmic reticulum).

endothermic reaction: A chemical reaction that takes up heat (that is, for which ΔH is positive).

energy charge: The fractional degree to which the ATP/ADP/AMP system is filled with high-energy phosphate groups.

energy coupling: The transfer of energy from one process to another.

enhancers: DNA sequences that facilitate the expression of a given gene; may be located a few hundred, or even thousand, base pairs away from the gene.

enthalpy (H): The heat content of a system.

enthalpy change (ΔH): For a reaction, is approximately equal to the difference between the energy used to break bonds and the energy gained by the formation of new ones.

entropy (S): The extent of randomness or disorder in a system.

enzyme: A biomolecule, either protein or RNA, that catalyzes a specific chemical reaction. It does not affect the equilibrium of the catalyzed reaction; it enhances the rate of a reaction by providing a reaction path with a lower activation energy.

epimerases: Enzymes that catalyze the reversible interconversion of two epimers.

epimers: Two stereoisomers differing in configuration at one asymmetric center, in a compound having two or more asymmetric centers.

epithelial cell: Any cell that forms part of the outer covering of an organism or organ.

epitope: An antigenic determinant; the particular chemical group or groups within a macromolecule (antigen) to which a given antibody binds.

equilibrium: The state of a system in which no further net change is occurring; the free energy is at a minimum.

equilibrium constant (K_{eq}): A constant, characteristic for each chemical reaction; relates the specific concentrations of all reactants and products at equilibrium at a given temperature and pressure.

erythrocyte: A cell containing large amounts of hemoglobin and specialized for oxygen transport; a red blood cell.

Escherichia coli: See *E. coli*.

essential amino acids: Amino acids that cannot be synthesized by humans (and other vertebrates) and must be obtained from the diet.

essential fatty acids: The group of polyunsaturated fatty acids produced by plants, but not by humans; required in the human diet.

ethanol fermentation: See alcohol fermentation.

eukaryote: A unicellular or multicellular organism with cells having a membrane-bounded nucleus, multiple chromosomes, and internal organelles.

excited state: An energy-rich state of an atom or molecule; produced by the absorption of light energy.

exergonic reaction: A chemical reaction that proceeds with the release of free energy (that is, for which ΔG is negative).

exocytosis: The fusion of an intracellular vesicle with the plasma membrane, releasing the vesicle contents to the extracellular space.

exon: The segment of a eukaryotic gene that encodes a portion of the final product of the gene; a portion that remains after posttranscriptional processing and is transcribed into a protein or incorporated into the structure of an RNA. See intron.

exonuclease: An enzyme that hydrolyzes only those phosphodiester bonds that are in the terminal positions of a nucleic acid.

exothermic reaction: A chemical reaction that releases heat (that is, for which ΔH is negative).

expression vector: See vector.

facilitated diffusion: Diffusion of a polar substance across a biological membrane through a protein transporter; also called passive diffusion or passive transport.

facultative cells: Cells that can live in the presence or absence of oxygen.

FAD (flavin adenine dinucleotide): The coenzyme of some oxidation–reduction enzymes; it contains riboflavin.

fatty acid: A long-chain aliphatic carboxylic acid found in natural fats and oils; also a component of membrane phospholipids and glycolipids.

feedback inhibition: Inhibition of an allosteric enzyme at the beginning of a metabolic sequence by the end product of the sequence; also known as end-product inhibition.

fermentation: Energy-yielding anaerobic breakdown of a nutrient molecule, such as glucose, without net oxidation; yields lactate, ethanol, or some other simple product.

fibroblast: A cell of the connective tissue that secretes connective tissue proteins such as collagen.

fibrous proteins: Insoluble proteins that serve in a protective or structural role; contain polypeptide chains that generally share a common secondary structure.

fingerprinting: See peptide mapping.

first law of thermodynamics: The law stating that in all processes, the total energy of the universe remains constant.

Fischer projection formulas: See projection formulas.

5′ end: The end of a nucleic acid that lacks a nucleotide bound at the 5′ position of the terminal residue.

flagellum: A cell appendage used in propulsion. Bacterial flagella have a much simpler structure than eukaryotic flagella, which are similar to cilia.

flavin-linked dehydrogenases: Dehydrogenases requiring one of the riboflavin coenzymes, FMN or FAD.

flavin nucleotides: Nucleotide coenzymes (FMN and FAD) containing riboflavin.

flavoprotein: An enzyme containing a flavin nucleotide as a tightly bound prosthetic group.

fluid mosaic model: A model describing biological membranes as a fluid lipid bilayer with embedded proteins; the bilayer exhibits both structural and functional asymmetry.

fluorescence: Emission of light by excited molecules as they revert to the ground state.

FMN (flavin mononucleotide): Riboflavin phosphate, a coenzyme of certain oxidation–reduction enzymes.

footprinting: A technique for identifying the nucleic acid sequence bound by a DNA- or RNA-binding protein.

frame shift: A mutation caused by insertion or deletion of one or more paired nucleotides, changing the reading frame of codons during protein synthesis; the polypeptide product has a garbled amino acid sequence beginning at the mutated codon.

free energy (G): The component of the total energy of a system that can do work at constant temperature and pressure.

free energy of activation ($\Delta G^{\ddagger}$): See activation energy.

free-energy change (ΔG): The amount of free energy released (negative ΔG) or absorbed (positive ΔG) in a reaction at constant temperature and pressure.

free radical: See radical.

functional group: The specific atom or group of atoms that confers a particular chemical property on a biomolecule.

furanose: A simple sugar containing the five-membered furan ring.

fusion protein: (1) A family of proteins that facilitate membrane fusion. (2) The protein product of a gene created by the fusion of two distinct genes.

futile cycle: A set of enzyme-catalyzed cyclic reactions that results in release of thermal energy by the hydrolysis of ATP.

$\Delta G^{\circ\prime}$: See standard free-energy change.

gametes: Reproductive cells with a haploid gene content; sperm or egg cells.

gangliosides: Sphingolipids, containing complex oligosaccharides as head groups; especially common in nervous tissue.

gel filtration: A chromatographic procedure for the separation of a mixture of molecules on the basis of size; based on the capacity of porous polymers to exclude solutes above a certain size.

gene: A chromosomal segment that codes for a single functional polypeptide chain or RNA molecule.

gene expression: Transcription and, in the case of proteins, translation to yield the product of a gene; a gene is expressed when its biological product is present and active.

gene splicing: The enzymatic attachment of one gene, or part of a gene, to another.

general acid–base catalysis: Catalysis involving proton transfer(s) to or from a molecule other than water.

genetic code: The set of triplet code words in DNA (or mRNA) coding for the amino acids of proteins.

genetic information: The hereditary information contained in a sequence of nucleotide bases in chromosomal DNA or RNA.

genetic map: A diagram showing the relative sequence and position of specific genes along a chromosome.

genome: All the genetic information encoded in a cell or virus.

genotype: The genetic constitution of an organism, as distinct from its physical characteristics, or phenotype.

geometric isomers: Isomers related by rotation about a double bond; also called cis and trans isomers.

germ-line cell: A type of animal cell that is formed early in embryogenesis and may multiply by mitosis or may produce, by meiosis, cells that develop into gametes (egg or sperm cells).

globular proteins: Soluble proteins with a globular (somewhat rounded) shape.

glucogenic amino acids: Amino acids with carbon chains that can be metabolically converted into glucose or glycogen via gluconeogenesis.

gluconeogenesis: The biosynthesis of a carbohydrate from simpler, noncarbohydrate precursors such as oxaloacetate or pyruvate.

glycan: Another term for polysaccharide; a polymer of monosaccharide units joined by glycosidic bonds.

glycerophospholipid: An amphipathic lipid with a glycerol backbone; fatty acids are ester-linked to C-1 and C-2 of glycerol, and a polar alcohol is attached through a phosphodiester linkage to C-3.

glycolipid: A lipid containing a carbohydrate group.

glycolysis: The catabolic pathway by which a molecule of glucose is broken down into two molecules of pyruvate.

glycoprotein: A protein containing a carbohydrate group.

glycosaminoglycan: A heteropolysaccharide of two alternating units: one is either *N*-acetylglucosamine or *N*-acetylgalactosamine; the other is a uronic acid (usually glucuronic acid). Formerly called mucopolysaccharide.

glycosidic bonds: Bonds between a sugar and another molecule (typically an alcohol, purine, pyrimidine, or sugar) through an intervening oxygen or nitrogen atom; the bonds are classified as *O*-glycosidic or *N*-glycosidic, respectively.

glyoxylate cycle: A variant of the citric acid cycle, for the net conversion of acetate into succinate and, eventually, new

carbohydrate; present in bacteria and some plant cells.

glyoxysome: A specialized peroxisome containing the enzymes of the glyoxylate cycle; found in cells of germinating seeds.

Golgi complex: A complex membranous organelle of eukaryotic cells; functions in the posttranslational modification of proteins and their secretion from the cell or incorporation into the plasma membrane or organellar membranes.

gram molecular weight: The weight in grams of a compound that is numerically equal to its molecular weight; the weight of 1 mole.

grana: Stacks of thylakoids, flattened membranous sacs or discs, in chloroplasts.

ground state: The normal, stable form of an atom or molecule; as distinct from the excited state.

group transfer potential: A measure of the ability of a compound to donate an activated group (such as a phosphate or acyl group); generally expressed as the standard free energy of hydrolysis.

half-life: The time required for the disappearance or decay of one-half of a given component in a system.

haploid: Having a single set of genetic information; describing a cell with one chromosome of each type.

Haworth perspective formulas: A method for representing cyclic chemical structures so as to define the configuration of each substituent group; the method commonly used for representing sugars.

helicase: An enzyme that catalyzes the separation of strands in a DNA molecule before replication.

helix, α: See α helix.

heme: The iron-porphyrin prosthetic group of heme proteins.

heme protein: A protein containing a heme as a prosthetic group.

hemoglobin: A heme protein in erythrocytes; functions in oxygen transport.

Henderson–Hasselbalch equation: An equation relating the pH, the pK_a, and the ratio of the concentrations of the proton-acceptor (A^-) and proton-donor (HA) species in a solution.

hepatocyte: The major cell type of liver tissue.

heteroduplex DNA: Duplex DNA containing complementary strands derived from two different DNA molecules with similar sequences, often as a product of genetic recombination.

heteropolysaccharide: A polysaccharide containing more than one type of sugar.

heterotroph: An organism that requires complex nutrient molecules, such as glucose, as a source of energy and carbon.

heterotropic enzyme: An allosteric enzyme requiring a modulator other than its substrate.

hexose: A simple sugar with a backbone containing six carbon atoms.

high-energy compound: A compound that on hydrolysis undergoes a large decrease in free energy under standard conditions.

high-performance liquid chromatography (HPLC): Chromatographic procedures, often conducted at relatively high pressures, using automated equipment that permits refined and highly reproducible profiles.

Hill reaction: The evolution of oxygen and the photoreduction of an artificial electron acceptor by a chloroplast preparation in the absence of carbon dioxide.

histones: The family of five basic proteins that associate tightly with DNA in the chromosomes of all eukaryotic cells.

Holliday intermediate: An intermediate in genetic recombination in which two double-stranded DNA molecules are joined by virtue of a reciprocal crossover involving one strand of each molecule.

holoenzyme: A catalytically active enzyme including all necessary subunits, prosthetic groups, and cofactors.

homeobox: A conserved DNA sequence of 180 base pairs encoding a protein domain found in many proteins that play a regulatory role in development.

homeodomain: The protein domain encoded by the homeobox.

homeostasis: The maintenance of a dynamic steady state by regulatory mechanisms that compensate for changes in external circumstances.

homeotic genes: Genes that regulate the development of the pattern of segments in the *Drosophila* body plan; similar genes are found in most vertebrates.

homologous genetic recombination: Recombination between two DNA molecules of similar sequence, occurring in all cells; occurs during meiosis and mitosis in eukaryotes.

homologous proteins: Proteins having sequences and functions similar in different species; for example, the hemoglobins.

homopolysaccharide: A polysaccharide made up of only one type of monosaccharide unit.

homotropic enzyme: An allosteric enzyme that uses its substrate as a modulator.

hormone: A chemical substance synthesized in small amounts by an endocrine tissue and carried in the blood to another tissue, where it acts as a messenger to regulate the function of the target tissue or organ.

hormone receptor: A protein in, or on the surface of, target cells that binds a specific hormone and initiates the cellular response.

hydrogen bond: A weak electrostatic attraction between one electronegative atom (such as oxygen or nitrogen) and a hydrogen atom covalently linked to a second electronegative atom.

hydrolases: Enzymes (proteases, lipases, phosphatases, nucleases, for example) that catalyze hydrolysis reactions.

hydrolysis: Cleavage of a bond, such as an anhydride or peptide bond, by the addition of the elements of water, yielding two or more products.

hydronium ion: The hydrated hydrogen ion (H_3O^+).

hydropathy index: A scale that expresses the relative hydrophobic and hydrophilic tendencies of a chemical group.

hydrophilic: Polar or charged; describing molecules or groups that associate with (dissolve easily in) water.

hydrophobic: Nonpolar; describing molecules or groups that are insoluble in water.

hydrophobic interactions: The association of nonpolar groups, or compounds, with each other in aqueous systems, driven by the tendency of the surrounding water molecules to seek their most stable (disordered) state.

hyperchromic effect: The large increase in light absorption at 260 nm occurring as a double-helical DNA is melted (unwound).

immune response: The capacity of a vertebrate to generate antibodies to an antigen, a macromolecule foreign to the organism.

immunoglobulin: An antibody protein generated against, and capable of binding specifically to, an antigen.

in vitro: "In glass"; that is, in the test tube.

in vivo: "In life"; that is, in the living cell or organism.

induced fit: A change in the conformation of an enzyme in response to substrate binding that renders the enzyme catalytically active; also used to denote changes in the conformation of any macromolecule in response to ligand binding such that the binding site of the macromolecule better conforms to the shape of the ligand.

inducer: A signal molecule that, when bound to a regulatory protein, produces an increase in the expression of a given gene.

induction: An increase in the expression of a gene in response to a change in the activity of a regulatory protein.

informational macromolecules: Biomolecules containing information in the form of specific sequences of different monomers; for example, many proteins, lipids, polysaccharides, and nucleic acids.

initiation codon: AUG (sometimes GUG in prokaryotes); codes for the first amino acid in a polypeptide sequence: *N*-formylmethionine in prokaryotes, and methionine in eukaryotes.

initiation complex: A complex of a ribosome with an mRNA and the initiating Met-tRNA^{Met} or fMet-tRNA^{fMet}, ready for the elongation steps.

inorganic pyrophosphatase: An enzyme that hydrolyzes a molecule of inorganic pyrophosphate to yield two molecules of (ortho) phosphate; also known as pyrophosphatase.

insertion mutation: A mutation caused by insertion of one or more extra bases, or a mutagen, between two successive bases in DNA.

insertion sequence: Specific base sequences at either end of a transposable segment of DNA.

integral membrane proteins: Proteins firmly bound to a membrane by hydrophobic interactions; as distinct from peripheral proteins.

intercalating mutagen: A mutagen that inserts itself between two successive bases in a nucleic acid, causing a frame-shift mutation.

intercalation: Insertion between two stacked aromatic or planar rings; for example, the insertion of a planar molecule between two successive bases in a nucleic acid.

interferons: A class of glycoproteins with antiviral activities.

intermediary metabolism: In cells, the enzyme-catalyzed reactions that extract chemical energy from nutrient molecules and utilize it to synthesize and assemble cell components.

intron (intervening sequence): A sequence of nucleotides in a gene that is transcribed but excised before the gene is translated.

ion channel: An integral membrane protein that provides for the regulated transport of a specific ion, or ions, across a membrane.

ion-exchange resin: A polymeric resin that contains fixed charged groups; used in chromatographic columns to separate ionic compounds.

ion product of water (K_W): The product of the concentrations of H^+ and OH^- in pure water: $K_W = [H^+][OH^-] = 1 \times 10^{-14}$ at 25 °C.

ionizing radiation: A type of radiation, such as x rays, that causes loss of electrons from some organic molecules, thus making them more reactive.

ionophore: A compound that binds one or more metal ions and is capable of diffusing across a membrane, carrying the bound ion.

iron–sulfur center: A prosthetic group of certain redox proteins involved in electron transfers; Fe^{2+} or Fe^{3+} is bound to inorganic sulfur and to Cys groups in the protein.

isoelectric focusing: An electrophoretic method for separating macromolecules on the basis of their isoelectric pH.

isoelectric pH (isoelectric point): The pH at which a solute has no net electric charge and thus does not move in an electric field.

isoenzymes: See isozymes.

isomerases: Enzymes that catalyze the transformation of compounds into their positional isomers.

isomers: Any two molecules with the same molecular formula but a different arrangement of molecular groups.

isoprene: The hydrocarbon 2-methyl-1,3-butadiene, a recurring structural unit of the terpenoid biomolecules.

isothermal: Occurring at constant temperature.

isotopes: Stable or radioactive forms of an element that differ in atomic weight but are otherwise chemically identical to the naturally abundant form of the element; used as tracers.

isozymes: Multiple forms of an enzyme that catalyze the same reaction but differ from each other in their amino acid sequence, substrate affinity, V_{max}, and/or regulatory properties; also called isoenzymes.

keratins: Insoluble protective or structural proteins consisting of parallel polypeptide chains in α-helical or β conformations.

ketogenic amino acids: Amino acids with carbon skeletons that can serve as precursors of the ketone bodies.

ketone bodies: Acetoacetate, D-β-hydroxybutyrate, and acetone; water-soluble fuels normally exported by the liver but overproduced during fasting or in untreated diabetes mellitus.

ketose: A simple monosaccharide in which the carbonyl group is a ketone.

ketosis: A condition in which the concentration of ketone bodies in the blood, tissues, and urine is abnormally high.

kinases: Enzymes that catalyze the phosphorylation of certain molecules by ATP.

kinetics: The study of reaction rates.

Krebs cycle: See citric acid cycle.

lagging strand: The DNA strand that, during replication, must be synthesized in the direction opposite to that in which the replication fork moves.

law of mass action: The law stating that the rate of any given chemical reaction is proportional to the product of the activities (or concentrations) of the reactants.

leader: A short sequence near the amino terminus of a protein or the 5′ end of an RNA that has a specialized targeting or regulatory function.

leading strand: The DNA strand that, during replication, is synthesized in the same direction in which the replication fork moves.

leaky mutant: A mutant gene that gives rise to a product with a detectable level of biological activity.

leaving group: The departing or displaced molecular group in a unimolecular elimination or a bimolecular substitution reaction.

lethal mutation: A mutation that inactivates a biological function essential to the life of the cell or organism.

leucine zipper: A protein structural

motif involved in protein–protein interactions in many eukaryotic regulatory proteins; consists of two interacting α helices in which Leu residues in every seventh position are a prominent feature of the interacting surfaces.

leukotrienes: A family of molecules derived from arachidonate; muscle contractants that constrict air passages in the lungs and are involved in asthma.

levorotatory isomer: A stereoisomer that rotates the plane of plane-polarized light counterclockwise.

ligand: A small molecule that binds specifically to a larger one; for example, a hormone is the ligand for its specific protein receptor.

light reactions: The reactions of photosynthesis that require light and cannot occur in the dark; also known as the light-dependent reactions.

Lineweaver–Burk equation: An algebraic transform of the Michaelis–Menten equation, allowing determination of V_{max} and K_m by extrapolation of [S] to infinity.

linking number: The number of times one closed circular DNA strand is wound about another; the number of topological links holding the circles together.

lipases: Enzymes that catalyze the hydrolysis of triacylglycerols.

lipid: A small water-insoluble biomolecule generally containing fatty acids, sterols, or isoprenoid compounds.

lipoate (lipoic acid): A vitamin for some microorganisms; an intermediate carrier of hydrogen atoms and acyl groups in α-keto acid dehydrogenases.

lipoprotein: A lipid–protein aggregate that serves to carry water-insoluble lipids in the blood. The protein component alone is an apolipoprotein.

low-energy phosphate compound: A phosphorylated compound with a relatively small standard free energy of hydrolysis.

lyases: Enzymes that catalyze the removal of a group from a molecule to form a double bond, or the addition of a group to a double bond.

lymphocytes: A subclass of leukocytes involved in the immune response. B lymphocytes synthesize and secrete antibodies; T lymphocytes either play a regulatory role in immunity or kill foreign and virus-infected cells.

lysis: Destruction of a cell's plasma membrane or of a bacterial cell wall, releasing the cellular contents and killing the cell.

lysogeny: One of two outcomes of the infection of a host cell by a temperate phage. It occurs when the phage genome becomes repressed and is replicated as part of the host DNA; infrequently it may be induced, and the phage particles so produced cause the host cell to lyse.

lysosome: A membrane-bounded organelle in the cytoplasm of eukaryotic cells; it contains many hydrolytic enzymes and serves as a degrading and recycling center for unneeded components.

macromolecule: A molecule having a molecular weight in the range of a few thousand to many millions.

mass-action ratio: For the reaction $aA + bB \rightleftharpoons cC + dD$, the ratio: $\frac{[C]^c\,[D]^d}{[A]^a\,[B]^b}$.

matrix: The aqueous contents of a cell or organelle (the mitochondrion, for example) with dissolved solutes.

meiosis: A type of cell division in which diploid cells give rise to haploid cells destined to become gametes.

membrane transport: Movement of a polar solute across a membrane via a specific membrane protein (a transporter).

messenger RNA (mRNA): A class of RNA molecules, each of which is complementary to one strand of DNA; carries the genetic message from the chromosome to the ribosomes.

metabolism: The entire set of enzyme-catalyzed transformations of organic molecules in living cells; the sum of anabolism and catabolism.

metabolite: A chemical intermediate in the enzyme-catalyzed reactions of metabolism.

metalloprotein: A protein having a metal ion as its prosthetic group.

metamerism: Division of the body into segments; in insects, for example.

micelle: An aggregate of amphipathic molecules in water, with the nonpolar portions in the interior and the polar portions at the exterior surface, exposed to water.

Michaelis–Menten constant (K_m): The substrate concentration at which an enzyme-catalyzed reaction proceeds at one-half its maximum velocity.

Michaelis–Menten equation: The equation describing the hyperbolic dependence of the initial reaction velocity, V_0, on substrate concentration, [S], in many enzyme-catalyzed reactions: $V_0 = \frac{V_{max}[S]}{K_m + [S]}$.

Michaelis–Menten kinetics: A kinetic pattern in which the initial rate of an enzyme-catalyzed reaction exhibits a hyperbolic dependence on substrate concentration.

microbodies: Cytoplasmic, membrane-bounded vesicles containing peroxide-forming and peroxide-destroying enzymes; include lysosomes, peroxisomes, and glyoxysomes.

microfilaments: Thin filaments composed of actin, found in the cytoplasm of eukaryotic cells; serve in structure and movement.

microsomes: Membranous vesicles formed by fragmentation of the endoplasmic reticulum of eukaryotic cells; recovered by differential centrifugation.

microtubules: Thin tubules assembled from two types of globular tubulin subunits; present in cilia, flagella, centrosomes, and other contractile or motile structures.

mitochondrion: Membrane-bounded organelle in the cytoplasm of eukaryotes; contains the enzyme systems required for the citric acid cycle, fatty acid oxidation, electron transfer, and oxidative phosphorylation.

mitosis: The multistep process in eukaryotic cells that results in the replication of chromosomes and cell division.

mixed-function oxidases (oxygenases): Enzymes, often flavoproteins, that use molecular oxygen (O_2) to simultaneously oxidize a substrate and a cosubstrate (commonly NADH or NADPH).

modulator: A metabolite that, when bound to the allosteric site of an enzyme, alters its kinetic characteristics.

molar solution: One mole of solute dissolved in water to give a total volume of 1,000 mL.

mole: One gram molecular weight of a compound. See Avogadro's number.

monoclonal antibodies: Antibodies produced by a cloned hybridoma cell, which therefore are identical and directed against the same epitope of the antigen.

monolayer: A single layer of oriented lipid molecules.

monoprotic acid: An acid having only one dissociable proton.

monosaccharide: A carbohydrate consisting of a single sugar unit.

mRNA: See messenger RNA.

mucopolysaccharide: An older name for a glycosaminoglycan.

multienzyme system: A group of related enzymes participating in a given metabolic pathway.

mutarotation: The change in specific rotation of a pyranose or furanose sugar or glycoside accompanying the equilibration of its α- and β-anomeric forms.

mutases: Enzymes that catalyze the transposition of functional groups.

mutation: An inheritable change in the nucleotide sequence of a chromosome.

myofibril: A unit of thick and thin filaments of muscle fibers.

myosin: A contractile protein; the major component of the thick filaments of muscle and other actin–myosin systems.

NAD, NADP (nicotinamide adenine dinucleotide, nicotinamide adenine dinucleotide phosphate): Nicotinamide-containing coenzymes functioning as carriers of hydrogen atoms and electrons in some oxidation–reduction reactions.

native conformation: The biologically active conformation of a macromolecule.

negative cooperativity: A phenomenon of some multisubunit enzymes or proteins in which binding of a ligand or substrate to one subunit impairs binding to another subunit.

negative feedback: Regulation of a biochemical pathway achieved when a reaction product inhibits an earlier step in the pathway.

neuron: A cell of nervous tissue specialized for transmission of a nerve impulse.

neurotransmitter: A low molecular weight compound (usually containing nitrogen) secreted from the terminal of a neuron and bound by a specific receptor in the next neuron; serves to transmit a nerve impulse.

nicotinamide adenine dinucleotide, nicotinamide adenine dinucleotide phosphate: See NAD, NADP.

ninhydrin reaction: A color reaction given by amino acids and peptides on heating with ninhydrin; widely used for their detection and estimation.

nitrogen cycle: The cycling of various forms of biologically available nitrogen through the plant, animal, and microbial worlds, and through the atmosphere and geosphere.

nitrogen fixation: Conversion of atmospheric nitrogen (N_2) into a reduced, biologically available form by nitrogen-fixing organisms.

nitrogenase complex: A system of enzymes capable of reducing atmospheric nitrogen to ammonia in the presence of ATP.

noncompetitive inhibition: A type of enzyme inhibition not reversed by increasing the substrate concentration.

noncyclic electron flow: The light-induced flow of electrons from water to $NADP^+$ in oxygen-evolving photosynthesis; it involves both photosystems I and II.

nonessential amino acids: Amino acids that can be made by humans and other vertebrates from simpler precursors, and are thus not required in the diet.

nonheme iron proteins: Proteins, usually acting in oxidation–reduction reactions, containing iron but no porphyrin groups.

nonpolar: Hydrophobic; describing molecules or groups that are poorly soluble in water.

nonsense codon: A codon that does not specify an amino acid, but signals the termination of a polypeptide chain.

nonsense mutation: A mutation that results in the premature termination of a polypeptide chain.

nonsense suppressor: A mutation, usually in the gene for a tRNA, that causes an amino acid to be inserted into a polypeptide in response to a termination codon.

nucleases: Enzymes that hydrolyze the internucleotide (phosphodiester) linkages of nucleic acids.

nucleic acids: Biologically occurring polynucleotides in which the nucleotide residues are linked in a specific sequence by phosphodiester bonds; DNA and RNA.

nucleoid: In bacteria, the nuclear zone that contains the chromosome but has no surrounding membrane.

nucleolus: A densely staining structure in the nucleus of eukaryotic cells; involved in rRNA synthesis and ribosome formation.

nucleophile: An electron-rich group with a strong tendency to donate electrons to an electron-deficient nucleus (electrophile); the entering reactant in a bimolecular substitution reaction.

nucleoplasm: The portion of a cell's contents enclosed by the nuclear membrane; also called the nuclear matrix.

nucleoside: A compound consisting of a purine or pyrimidine base covalently linked to a pentose.

nucleoside diphosphate kinase: An enzyme that catalyzes the transfer of the terminal phosphate of a nucleoside 5′-triphosphate to a nucleoside 5′-diphosphate.

nucleoside diphosphate sugar: A coenzymelike carrier of a sugar molecule, functioning in the enzymatic synthesis of polysaccharides and sugar derivatives.

nucleoside monophosphate kinase: An enzyme that catalyzes the transfer of the terminal phosphate of ATP to a nucleoside 5′-monophosphate.

nucleosome: Structural unit for packaging chromatin; consists of a DNA strand wound around a histone core.

nucleotide: A nucleoside phosphorylated at one of its pentose hydroxyl groups.

nucleus: In eukaryotes, a membrane-bounded organelle that contains chromosomes.

oligomer: A short polymer, usually of amino acids, sugars, or nucleotides; the definition of "short" is somewhat arbitrary, but usually less than 50 subunits.

oligomeric protein: A multisubunit protein having two or more identical polypeptide chains.

oligonucleotide: A short polymer of nucleotides (usually less than 50).

oligopeptide: A few amino acids joined by peptide bonds.

oligosaccharide: Several monosaccharide groups joined by glycosidic bonds.

oncogene: A cancer-causing gene; any of several mutant genes that cause cells to exhibit rapid, uncontrolled proliferation. See also proto-oncogene.

open reading frame: A group of contiguous nonoverlapping nucleotide codons in a DNA or RNA molecule that do not include a termination codon.

open system: A system that exchanges matter and energy with its surroundings. See also system.

operator: A region of DNA that interacts with a repressor protein to control the expression of a gene or group of genes.

operon: A unit of genetic expression consisting of one or more related genes and the operator and promoter sequences that regulate their transcription.

optical activity: The capacity of a substance to rotate the plane of plane-polarized light.

optimum pH: The characteristic pH at which an enzyme has maximal catalytic activity.

organelles: Membrane-bounded structures found in eukaryotic cells; contain

enzymes and other components required for specialized cell functions.

origin: The nucleotide sequence or site in DNA where DNA replication is initiated.

osmosis: Bulk flow of water through a semipermeable membrane into another aqueous compartment containing solute at a higher concentration.

osmotic pressure: Pressure generated by the osmotic flow of water through a semipermeable membrane into an aqueous compartment containing solute at a higher concentration.

oxidation: The loss of electrons from a compound.

oxidation, β: See β oxidation.

oxidation–reduction reaction: A reaction in which electrons are transferred from a donor to an acceptor molecule; also called a redox reaction.

oxidative phosphorylation: The enzymatic phosphorylation of ADP to ATP coupled to electron transfer from a substrate to molecular oxygen.

oxidizing agent (oxidant): The acceptor of electrons in an oxidation–reduction reaction.

oxygen debt: The extra oxygen (above the normal resting level) consumed in the recovery period after strenuous physical exertion.

oxygenases: Enzymes that catalyze reactions in which oxygen is introduced into an acceptor molecule.

palindrome: A segment of duplex DNA in which the base sequences of the two strands exhibit twofold rotational symmetry about an axis.

paradigm: In biochemistry, an experimental model or example.

partition coefficient: A constant that expresses the ratio in which a given solute will be partitioned or distributed between two given immiscible liquids at equilibrium.

pathogenic: Disease-causing.

pentose: A simple sugar with a backbone containing five carbon atoms.

pentose phosphate pathway: A pathway that serves to interconvert hexoses and pentoses and is a source of reducing equivalents and pentoses for biosynthetic processes; present in most organisms. Also called the phosphogluconate pathway.

peptidase: An enzyme that hydrolyzes a peptide bond.

peptide: Two or more amino acids covalently joined by peptide bonds.

peptide bond: A substituted amide linkage between the α-amino group of one amino acid and the α-carboxyl group of another, with the elimination of the elements of water.

peptide mapping: The characteristic two-dimensional pattern (on paper or gel) formed by the separation of a mixture of peptides resulting from partial hydrolysis of a protein; also known as peptide fingerprinting.

peptidoglycan: A major component of bacterial cell walls; generally consists of parallel heteropolysaccharides cross-linked by short peptides.

peripheral proteins: Proteins that are loosely or reversibly bound to a membrane by hydrogen bonds or electrostatic forces; generally water-soluble once released from the membrane.

permeases: See transporters.

peroxisome: Membrane-bounded organelle in the cytoplasm of eukaryotic cells; contains peroxide-forming and peroxide-destroying enzymes.

pH: The negative logarithm of the hydrogen ion concentration of an aqueous solution.

phage: See bacteriophage.

phenotype: The observable characteristics of an organism.

phosphodiester linkage: A chemical grouping that contains two alcohols esterified to one molecule of phosphoric acid, which thus serves as a bridge between them.

phosphogluconate pathway: An oxidative pathway beginning with glucose-6-phosphate and leading, via 6-phosphogluconate, to pentose phosphates and yielding NADPH. Also called the pentose phosphate pathway.

phospholipid: A lipid containing one or more phosphate groups.

phosphorolysis: Cleavage of a compound with phosphate as the attacking group; analogous to hydrolysis.

phosphorylation: Formation of a phosphate derivative of a biomolecule, usually by enzymatic transfer of a phosphate group from ATP.

phosphorylation potential (ΔG_p): The actual free-energy change of ATP hydrolysis under the nonstandard conditions prevailing within a cell.

photochemical reaction center: The part of a photosynthetic complex where the energy of an absorbed photon causes charge separation, initiating electron transfer.

photon: The ultimate unit (a quantum) of light energy.

photophosphorylation: The enzymatic formation of ATP from ADP coupled to the light-dependent transfer of electrons in photosynthetic cells.

photoreduction: The light-induced reduction of an electron acceptor in photosynthetic cells.

photorespiration: Oxygen consumption occurring in illuminated temperate-zone plants, largely due to oxidation of phosphoglycolate.

photosynthesis: The use of light energy to produce carbohydrates from carbon dioxide and a reducing agent such as water.

photosynthetic phosphorylation: See photophosphorylation.

photosystem: In photosynthetic cells, a functional set of light-absorbing pigments and its reaction center.

phototroph: An organism that can use the energy of light to synthesize its own fuels from simple molecules such as carbon dioxide, oxygen, and water; as distinct from a chemotroph.

p*K*: The negative logarithm of an equilibrium constant.

plasma membrane: The exterior membrane surrounding the cytoplasm of a cell.

plasma proteins: The proteins present in blood plasma.

plasmalogen: A phospholipid with an alkenyl ether substituent on the C-1 of glycerol.

plasmid: An extrachromosomal, independently replicating, small circular DNA molecule; commonly employed in genetic engineering.

plastid: In plants, a self-replicating organelle; may differentiate into a chloroplast.

platelets: Small, enucleated cells that initiate blood clotting; they arise from cells called megakaryocytes in the bone marrow. Also known as thrombocytes.

pleated sheet: The side-by-side, hydrogen-bonded arrangement of polypeptide chains in the extended β conformation.

polar: Hydrophilic, or "water-loving"; describing molecules or groups that are soluble in water.

polarity: (1) In chemistry, the nonuniform distribution of electrons in a molecule; polar molecules are usually soluble in

water. (2) In molecular biology, the distinction between the 5′ and 3′ ends of nucleic acids.

polyclonal antibodies: A heterogeneous pool of antibodies produced in an animal by a number of different B lymphocytes in response to an antigen. Different antibodies in the pool recognize different parts of the antigen.

polylinker: A short, often synthetic, fragment of DNA containing recognition sequences for several restriction endonucleases.

polymerase chain reaction (PCR): A repetitive procedure that results in a geometric amplification of a specific DNA sequence.

polymorphic: Describing a protein for which amino acid sequence variants exist in a population of organisms, but the variations do not destroy the protein's function.

polynucleotide: A covalently linked sequence of nucleotides in which the 3′ hydroxyl of the pentose of one nucleotide residue is joined by a phosphodiester bond to the 5′ hydroxyl of the pentose of the next residue.

polypeptide: A long chain of amino acids linked by peptide bonds; the molecular weight is generally less than 10,000.

polyribosome: See polysome.

polysaccharide: A linear or branched polymer of monosaccharide units linked by glycosidic bonds.

polysome (polyribosome): A complex of an mRNA molecule and two or more ribosomes.

porphyrin: Complex nitrogenous compound containing four substituted pyrroles covalently joined into a ring; often complexed with a central metal atom.

positive cooperativity: A phenomenon of some multisubunit enzymes or proteins in which binding of a ligand or substrate to one subunit facilitates binding to another subunit.

posttranscriptional processing: The enzymatic processing of the primary RNA transcript, producing functional mRNA, tRNA, and/or rRNA molecules.

posttranslational modification: Enzymatic processing of a polypeptide chain after translation from its mRNA.

primary structure: A description of the covalent backbone of a polymer (macromolecule), including the sequence of monomeric subunits and any interchain and intrachain covalent bonds.

primary transcript: The immediate RNA product of transcription before any posttranscriptional processing reactions.

primase: An enzyme that catalyzes the formation of RNA oligonucleotides used as primers by DNA polymerases.

primer: A short oligomer (of sugars or nucleotides, for example) to which an enzyme adds additional monomeric subunits.

probe: A labeled fragment of nucleic acid containing a nucleotide sequence complementary to a gene or genomic sequence that one wishes to detect in a hybridization experiment.

processivity: For any enzyme that catalyzes the synthesis of a biological polymer, the property of adding multiple subunits to the polymer without dissociating from the substrate.

prochiral molecule: A symmetric molecule that can react asymmetrically with an enzyme having an asymmetric active site, generating a chiral product.

projection formulas: A method for representing molecules to show the configuration of groups around chiral centers; also known as Fischer projection formulas.

prokaryote: A bacterium; a unicellular organism with a single chromosome, no nuclear envelope, and no membrane-bounded organelles.

promoter: A DNA sequence at which RNA polymerase may bind, leading to initiation of transcription.

prophage: A bacteriophage in an inactive state in which the genome is either integrated into the chromosome of the host cell or (sometimes) replicated autonomously.

prostaglandins: A class of lipid-soluble, hormonelike regulatory molecules derived from arachidonate and other polyunsaturated fatty acids.

prosthetic group: A metal ion or an organic compound (other than an amino acid) that is covalently bound to a protein and is essential to its activity.

protein: A macromolecule composed of one or more polypeptide chains, each with a characteristic sequence of amino acids linked by peptide bonds.

protein kinases: Enzymes that phosphorylate certain amino acid residues in specific proteins.

protein targeting: The process by which newly synthesized proteins are sorted and transported to their proper locations in the cell.

proteoglycan: A hybrid macromolecule consisting of a heteropolysaccharide joined to a polypeptide; the polysaccharide is the major component.

proto-oncogene: A cellular gene, usually encoding a regulatory protein, that can be converted into an oncogene by mutation.

proton acceptor: An anionic compound capable of accepting a proton from a proton donor; that is, a base.

proton donor: The donor of a proton in an acid–base reaction; that is, an acid.

proton-motive force: The electrochemical potential inherent in a transmembrane gradient of H^+ concentration; used in oxidative phosphorylation and photophosphorylation to drive ATP synthesis.

protoplasm: A general term referring to the entire contents of a living cell.

purine: A nitrogenous heterocyclic base found in nucleotides and nucleic acids; containing fused pyrimidine and imidazole rings.

puromycin: An antibiotic that inhibits polypeptide synthesis by being incorporated into a growing polypeptide chain, causing its premature termination.

pyranose: A simple sugar containing the six-membered pyran ring.

pyridine nucleotide: A nucleotide coenzyme containing the pyridine derivative nicotinamide; NAD or NADP.

pyridoxal phosphate: A coenzyme containing the vitamin pyridoxine (vitamin B_6); functions in reactions involving amino group transfer.

pyrimidine: A nitrogenous heterocyclic base found in nucleotides and nucleic acids.

pyrimidine dimer: A covalently joined dimer of two adjacent pyrimidine residues in DNA, induced by absorption of UV light; most commonly derived from two adjacent thymines (a thymine dimer).

pyrophosphatase: See inorganic pyrophosphatase.

quantum: The ultimate unit of energy.

quaternary structure: The three-dimensional structure of a multisubunit protein; particularly the manner in which the subunits fit together.

R group: (1) Formally, an abbreviation denoting any alkyl group. (2) Occasionally, used in a more general sense to denote virtually any organic substituent (the R groups of amino acids, for example).

racemic mixture (racemate): An equimolar mixture of the D and L stereoisomers of an optically active compound.

radical: An atom or group of atoms possessing an unpaired electron; also called a free radical.

radioactive isotope: An isotopic form of an element with an unstable nucleus that stabilizes itself by emitting ionizing radiation.

radioimmunoassay: A sensitive and quantitative method for detecting trace amounts of a biomolecule, based on its capacity to displace a radioactive form of the molecule from combination with its specific antibody.

rate-limiting step: (1) Generally, the step in an enzymatic reaction with the greatest activation energy or the transition state of highest free energy. (2) The slowest step in a metabolic pathway.

reaction intermediate: Any chemical species in a reaction pathway that has a finite chemical lifetime.

reading frame: A contiguous and nonoverlapping set of three-nucleotide codons in DNA or RNA.

recombinant DNA: DNA formed by the joining of genes into new combinations.

redox pair: An electron donor and its corresponding oxidized form; for example, NADH and NAD^+.

redox reaction: See oxidation–reduction reaction.

reducing agent (reductant): The electron donor in an oxidation–reduction reaction.

reducing end: The end of a polysaccharide having a terminal sugar with a free anomeric carbon; the terminal residue can act as a reducing sugar.

reducing equivalent: A general or neutral term for an electron or an electron equivalent in the form of a hydrogen atom or a hydride ion.

reducing sugar: A sugar in which the carbonyl (anomeric) carbon is not involved in a glycosidic bond and can therefore undergo oxidation.

reduction: The gain of electrons by a compound or ion.

regulatory enzyme: An enzyme having a regulatory function through its capacity to undergo a change in catalytic activity by allosteric mechanisms or by covalent modification.

regulatory gene: A gene that gives rise to a product involved in the regulation of the expression of another gene; for example, a gene coding for a repressor protein.

regulatory sequence: A DNA sequence involved in regulating the expression of a gene; for example, a promoter or operator.

regulon: A group of genes or operons that are coordinately regulated even though some, or all, may be spatially distant within the chromosome or genome.

release factors: See termination factors.

releasing factors: Hypothalamic hormones that stimulate release of other hormones by the pituitary gland.

renaturation: Refolding of an unfolded (denatured) globular protein so as to restore native structure and protein function.

replication: Synthesis of a daughter duplex DNA molecule identical to the parental duplex DNA.

replisome: The multiprotein complex that promotes DNA synthesis at the replication fork.

repressible enzyme: In bacteria, an enzyme whose synthesis is inhibited when its reaction product is readily available to the cell.

repression: A decrease in the expression of a gene in response to a change in the activity of a regulatory protein.

repressor: The protein that binds to the regulatory sequence or operator for a gene, blocking its transcription.

residue: A single unit within a polymer; for example, an amino acid within a polypeptide chain. The term reflects the fact that sugars, nucleotides, and amino acids lose a few atoms (generally the elements of water) when incorporated in their respective polymers.

respiration: The catabolic process in which electrons are removed from nutrient molecules and passed through a chain of carriers to oxygen.

respiratory chain: The electron transfer chain; a sequence of electron-carrying proteins that transfer electrons from substrates to molecular oxygen in aerobic cells.

restriction endonucleases: Site-specific endodeoxyribonucleases causing cleavage of both strands of DNA at points within or near the specific site recognized by the enzyme; important tools in genetic engineering.

restriction fragment: A segment of double-stranded DNA produced by the action of a restriction endonuclease on a larger DNA.

restriction fragment length polymorphisms (RFLPs): Variations, among individuals in a population, in the length of certain restriction fragments within which certain genomic sequences occur. These variations result from rare sequence changes that create or destroy restriction sites in the genome.

retrovirus: An RNA virus containing a reverse transcriptase.

reverse transcriptase: An RNA-directed DNA polymerase in retroviruses; capable of making DNA complementary to an RNA.

ribonuclease: A nuclease that catalyzes the hydrolysis of certain internucleotide linkages of RNA.

ribonucleic acid: See RNA.

ribonucleotide: A nucleotide containing D-ribose as its pentose component.

ribosomal RNA (rRNA): A class of RNA molecules serving as components of ribosomes.

ribosome: A supramolecular complex of rRNAs and proteins, approximately 18 to 22 nm in diameter; the site of protein synthesis.

ribozymes: Ribonucleic acid molecules with catalytic activities; RNA enzymes.

RNA (ribonucleic acid): A polyribonucleotide of a specific sequence linked by successive 3′,5′-phosphodiester bonds.

RNA polymerase: An enzyme that catalyzes the formation of RNA from ribonucleoside 5′-triphosphates, using a strand of DNA or RNA as a template.

RNA splicing: Removal of introns and joining of exons in a primary transcript.

rRNA: See ribosomal RNA.

***S*-adenosylmethionine (adoMet):** An enzymatic cofactor involved in methyl group transfers.

salvage pathway: Synthesis of a biomolecule, such as a nucleotide, from intermediates in the degradative pathway for the biomolecule; a recycling pathway, as distinct from a de novo pathway.

saponification: Alkaline hydrolysis of triacylglycerols to yield fatty acids as soaps.

sarcomere: A functional and structural unit of the muscle contractile system.

satellite DNA: Highly repeated, nontranslated segments of DNA in eukaryotic chromosomes; most often associated with the centromeric region. Its function is not clear.

saturated fatty acid: A fatty acid containing a fully saturated alkyl chain.

second law of thermodynamics: The law stating that in any chemical or physical process, the entropy of the universe tends to increase.

second messenger: An effector molecule synthesized within a cell in response to an external signal (first messenger) such as a hormone.

secondary metabolism: Pathways that lead to specialized products not found in every living cell.

secondary structure: The residue-by-residue conformation of the backbone of a polymer.

sedimentation coefficient: A physical constant specifying the rate of sedimentation of a particle in a centrifugal field under specified conditions.

Shine–Dalgarno sequence: A sequence in an mRNA required for binding prokaryotic ribosomes.

shuttle vector: A recombinant DNA vector that can be replicated in two or more different host species. See also vector.

sickle-cell anemia: A human disease characterized by defective hemoglobin molecules; caused by a homozygous allele coding for the β chain of hemoglobin.

sickle-cell trait: A human condition recognized by the sickling of erythrocytes when exposed to low oxygen tension; occurs in individuals heterozygous for the allele responsible for sickle-cell anemia.

signal sequence: An amino-terminal sequence that signals the cellular fate or destination of a newly synthesized protein.

signal transduction: The process by which an extracellular signal (chemical, mechanical, or electrical) is amplified and converted to a cellular response.

silent mutation: A mutation in a gene that causes no detectable change in the biological characteristics of the gene product.

simple diffusion: The movement of solute molecules across a membrane to a region of lower concentration, unassisted by a protein transporter.

simple protein: A protein yielding only amino acids on hydrolysis.

site-directed mutagenesis: A set of methods used to create specific alterations in the sequence of a gene.

site-specific recombination: A type of genetic recombination that occurs only at specific sequences.

small nuclear RNA (snRNA): Any of several small RNA molecules in the nucleus; most have a role in the splicing reactions that remove introns from mRNA, tRNA, and rRNA molecules.

somatic cells: All body cells except the germ-line cells.

SOS response: In bacteria, a coordinated induction of a variety of genes as a response to high levels of DNA damage.

Southern blot: A DNA hybridization procedure in which one or more specific DNA fragments are detected in a larger population by means of hybridization to a complementary, labeled nucleic acid probe.

specific activity: The number of micromoles (μmol) of a substrate transformed by an enzyme preparation per minute per milligram of protein at 25 °C; a measure of enzyme purity.

specific heat: The amount of energy (in joules or calories) needed to raise the temperature of 1 g of a pure substance by 1 °C.

specific rotation: The rotation, in degrees, of the plane of plane-polarized light (D-line of sodium) by an optically active compound at 25 °C, with a specified concentration and light path.

specificity: The ability of an enzyme or receptor to discriminate among competing substrates or ligands.

sphingolipid: An amphipathic lipid with a sphingosine backbone to which are attached a long-chain fatty acid and a polar alcohol.

splicing: See gene splicing; RNA splicing.

standard free-energy change ($\Delta G°$): The free-energy change for a reaction occurring under a set of standard conditions: temperature, 298 K; pressure, 1 atm or 101.3 kPa; and all solutes at 1 M concentration. $\Delta G°'$ denotes the standard free-energy change at pH 7.0.

standard reduction potential (E_0'): The electromotive force exhibited at an electrode by 1 M concentrations of a reducing agent and its oxidized form at 25 °C and pH 7.0; a measure of the relative tendency of the reducing agent to lose electrons.

steady state: A nonequilibrium state of a system through which matter is flowing and in which all components remain at a constant concentration.

stem cells: The common, self-regenerating cells in bone marrow that give rise to differentiated blood cells such as erythrocytes and lymphocytes.

stereoisomers: Compounds that have the same composition and the same order of atomic connections, but different molecular arrangements.

sterols: A class of lipids containing the steroid nucleus.

sticky ends: Two DNA ends in the same DNA molecule, or in different molecules, with short overhanging single-stranded segments that are complementary to one another, facilitating ligation of the ends; also known as cohesive ends.

stop codons: See termination codons.

stroma: The space and aqueous solution enclosed within the inner membrane of a chloroplast, not including the contents within the thylakoid membranes.

structural gene: A gene coding for a protein or RNA molecule; as distinct from a regulatory gene.

substitution mutation: A mutation caused by the replacement of one base by another.

substrate: The specific compound acted upon by an enzyme.

substrate-level phosphorylation: Phosphorylation of ADP or some other nucleoside 5′-diphosphate coupled to the dehydrogenation of an organic substrate; independent of the electron transfer chain.

suicide inhibitor: A relatively inert molecule that is transformed by an enzyme, at its active site, into a reactive substance that irreversibly inactivates the enzyme.

suppressor mutation: A mutation that totally or partially restores a function lost by a primary mutation; located at a site different from the site of the primary mutation.

Svedberg (S): A unit of measure of the rate at which a particle sediments in a centrifugal field.

symbionts: Two or more organisms that are mutually interdependent; usually living in physical association.

symport: Cotransport of solutes across a membrane in the same direction.

synthases: Enzymes that catalyze condensation reactions in which no nucleoside triphosphate is required as an energy source.

synthetases: Enzymes that catalyze condensation reactions using ATP or another nucleoside triphosphate as an energy source.

system: An isolated collection of matter; all other matter in the universe apart from the system is called the surroundings.

telomere: Specialized nucleic acid structure found at the ends of linear eukaryotic chromosomes.

temperate phage: A phage whose DNA may be incorporated into the host-cell

genome without being expressed; as distinct from a virulent phage, which destroys the host cell.

template: A macromolecular mold or pattern for the synthesis of an informational macromolecule.

terminal transferase: An enzyme that catalyzes the addition of nucleotide residues of a single kind to the 3′ end of DNA chains.

termination codons: UAA, UAG, and UGA; in protein synthesis, signal the termination of a polypeptide chain. Also known as stop codons.

termination factors: Protein factors of the cytosol required in releasing a completed polypeptide chain from a ribosome; also known as release factors.

termination sequence: A DNA sequence that appears at the end of a transcriptional unit and signals the end of transcription.

terpenes: Organic hydrocarbons or hydrocarbon derivatives constructed from recurring isoprene units. They produce some of the scents and tastes of plant products; for example, the scents of geranium leaves and pine needles.

tertiary structure: The three-dimensional conformation of a polymer in its native folded state.

tetrahydrobiopterin: The reduced coenzyme form of biopterin.

tetrahydrofolate: The reduced, active coenzyme form of the vitamin folate.

thiamine pyrophosphate: The active coenzyme form of vitamin B_1; involved in aldehyde transfer reactions.

thioester: An ester of a carboxylic acid with a thiol or mercaptan.

3′ end: The end of a nucleic acid that lacks a nucleotide bound at the 3′ position of the terminal residue.

thrombocytes: See platelets.

thromboxanes: A class of molecules derived from arachidonate and involved in platelet aggregation during blood clotting.

thylakoid: Closed cisterna, or disc, formed by the pigment-bearing internal membranes of chloroplasts.

thymine dimer: See pyrimidine dimer.

tissue culture: Method by which cells derived from multicellular organisms are grown in liquid media.

titration curve: A plot of the pH versus the equivalents of base added during titration of an acid.

tocopherols: Forms of vitamin E.

topoisomerases: Enzymes that introduce positive or negative supercoils in closed, circular duplex DNA.

topoisomers: Different forms of a covalently closed, circular DNA molecule that differ only in their linking number.

toxins: Proteins produced by some organisms and toxic to certain other species.

trace element: A chemical element required by an organism in only trace amounts.

transaminases: See aminotransferases.

transamination: Enzymatic transfer of an amino group from an α-amino acid to an α-keto acid.

transcription: The enzymatic process whereby the genetic information contained in one strand of DNA is used to specify a complementary sequence of bases in an mRNA chain.

transcriptional control: The regulation of a protein's synthesis by regulation of the formation of its mRNA.

transduction: (1) Generally, the conversion of energy or information from one form to another. (2) The transfer of genetic information from one cell to another by means of a viral vector.

transfer RNA (tRNA): A class of RNA molecules (M_r 25,000 to 30,000), each of which combines covalently with a specific amino acid as the first step in protein synthesis.

transformation: Introduction of an exogenous DNA into a cell, causing the cell to acquire a new phenotype.

transgenic: Describing an organism that has genes from another organism incorporated within its genome as a result of recombinant DNA procedures.

transition state: An activated form of a molecule in which the molecule has undergone a partial chemical reaction; the highest point on the reaction coordinate.

translation: The process in which the genetic information present in an mRNA molecule specifies the sequence of amino acids during protein synthesis.

translational control: The regulation of a protein's synthesis by regulation of the rate of its translation on the ribosome.

translational repressor: A repressor that binds to an mRNA, blocking translation.

translocase: (1) An enzyme that catalyzes membrane transport. (2) An enzyme that causes a movement, such as the movement of a ribosome along an mRNA.

transpiration: Passage of water from the roots of a plant to the atmosphere via the vascular system and the stomata of the leaves.

transporters: Proteins that span a membrane and transport specific nutrients, metabolites, ions, or proteins across the membrane; sometimes called permeases.

transposition: The movement of a gene or set of genes from one site in the genome to another.

transposon (transposable element): A segment of DNA that can move from one position in the genome to another.

triacylglycerol: An ester of glycerol with three molecules of fatty acid; also called a triglyceride or neutral fat.

tricarboxylic acid cycle: See citric acid cycle.

triose: A simple sugar with a backbone containing three carbon atoms.

tRNA: See transfer RNA.

tropic hormone (tropin): A peptide hormone that stimulates a specific target gland to secrete its hormone; for example, thyrotropin produced by the pituitary stimulates secretion of thyroxine by the thyroid.

turnover number: The number of times an enzyme molecule transforms a substrate molecule per unit time, under conditions giving maximal activity at substrate concentrations that are saturating.

ultraviolet (UV) radiation: Electromagnetic radiation in the region of 200 to 400 nm.

uncoupling agent: A substance that uncouples phosphorylation of ADP from electron transfer; for example, 2,4-dinitrophenol.

uniport: A transport system that carries only one solute, as distinct from cotransport.

unsaturated fatty acid: A fatty acid containing one or more double bonds.

urea cycle: A metabolic pathway in vertebrates, for the synthesis of urea from amino groups and carbon dioxide; occurs in the liver.

ureotelic: Excreting excess nitrogen in the form of urea.

uricotelic: Excreting excess nitrogen in the form of urate (uric acid).

V_{max}: The maximum velocity of an enzymatic reaction when the binding site is saturated with substrate.

vector: A DNA molecule known to replicate autonomously in a host cell, to which a segment of DNA may be spliced to allow its replication; for example, a plasmid or a temperate-phage DNA.

viral vector: A viral DNA altered so that it can act as a vector for recombinant DNA.

virion: A virus particle.

virus: A self-replicating, infectious, nucleic acid–protein complex that requires an intact host cell for its replication; its genome is either DNA or RNA.

vitamin: An organic substance required in small quantities in the diet of some species; generally functions as a component of a coenzyme.

wild type: The normal (unmutated) phenotype.

wobble: The relatively loose base pairing between the base at the 3′ end of a codon and the complementary base at the 5′ end of the anticodon.

x-ray crystallography: The analysis of x-ray diffraction patterns of a crystalline compound, used to determine the molecule's three-dimensional structure.

zinc finger: A specialized protein motif involved in DNA recognition by some DNA-binding proteins; characterized by a single atom of zinc coordinated to four Lys residues or to two His and two Lys residues.

zwitterion: A dipolar ion, with spatially separated positive and negative charges.

zymogen: An inactive precursor of an enzyme; for example, pepsinogen, the precursor of pepsin.