**Lecture 5.**

**Ecology of microorganisms. Microflora of the biosphere. Normal microbiota of the human organism. Action of external environmental factors (physical, chemical and biological) to microorganisms. Phages. Genetics of microorganisms, types of genetic variability. Biotechnology and genetic engineering**

**The purpose of the lecture:** To acquaint students with the ecology of microorganisms, the normal microflora of the human body, its role. Explain the characteristics of the impact of environmental factors (physical, chemical and biological) to microorganisms. To inform students about bacteriophages, their structure, characteristics and use in medical practice. To acquaint students with the genetic apparatus of bacteria, non-hereditary and hereditary variables, their mechanisms. To inform them about the genetics of viruses. To clarify the application of genetic engineering and genetic methods in diagnostics.

**Lecture plan:**

1. Ecology of microorganisms.

- Types of interactions between microorganisms. Symbiosis and its forms.

- Distribution of microorganisms in the environment. Sanitary indicator microorganisms and their characteristics. Soil, water and air microflora, sanitary indicators microorganisms. The presence of microbes in the metabolism in nature.

- Normal microflora of the human body, its role in physiological processes and pathology.

- Dysbacteriosis. Drugs used to restore normal flora. Eubiotics and probiotics.

- Gnotobiology

2. The effect of environmental factors on microorganisms.

- The effect of various physical factors (temperature, light, drought, pressure, radiation, ultrasound, etc.) on microorganisms.

- The effect of various chemicals on microorganisms, the main groups of disinfectants (surfactants, phenol, heavy metal salts, acids, alkalis, alcohol, etc.).

- The concept of disinfection and sterilization.

- The concept of asepsis and antiseptics.

- Sterilization methods: physical, chemical, mechanical

- The main groups of disinfectants

3. Bacteriophages.

- The nature, structure and characteristics of bacteriophages.

- Virulent and mild phages. Lysogeny, its mechanism. Phage conversion.

- Application of phages in practice.

4. Genetics of microorganisms.

- The concept of genetics of microorganisms, information about heredity and variability.

- Genetic apparatus of prokaryotes: replicon, operon, promoter concepts.

- Genotype and phenotype concepts. Bacterial chromosomes and plasmids. Types of plasmids (F, R, Hly, etc.).

- Forms of variability: non-hereditary (modification) and hereditary (genetic).

- Non-hereditary (modified) variability in bacteria.

- Hereditary variability in bacteria (mutation and recombination).

- Mutations, their types: spontaneous and inductive, gene and chromosome, direct and reversible mutations.

- Genetic recombination in bacteria. Transformation, transduction and conjugation. Their mechanism.

- Genetics of viruses.

- Genetic engineering, goals and objectives, practical application in microbiology. Application of genetic methods in diagnosis.

SURVIVAL OF MICROORGANISMS IN THE NATURAL ENVIRONMENT

The population of microorganisms in the biosphere remains roughly constant because the growth of microorganisms is balanced by the death of these organisms. The survival of an group within an environmental niche is ultimately nfluenced by successful competition for nutrients and by maintenance of a pool of all living cells, often composed of human cells and a consortium of different microorganisms (referred to as the microbiome or microbiota). Understanding competition for nutritional resources within a given microenvironment

is essential to understanding the growth, survival, and death of bacterial species (also known as physiology). Much of our understanding of microbial physiology has come from the study of isolated cultures grown under optimal conditions in laboratories (nutrient excess). However, most microorganisms compete in the natural environment under nutritional stress. Furthermore, a vacant environmental

microbial niche will soon be filled with a different microbiota composition. In the end, the complex interactions that ensure the survival of a specific microbiome is a balance between availability of nutrients and physiologic efficiency.

THE MEANING OF GROWTH

Growth is the orderly increase in the sum of all the components of an organism. The increaseb in size that results when a cell takes up water or deposits lipid or polysaccharide is not true growth. Cell multiplication is a consequence of binary fission that leads to an increase in the number of single bacteria making up a population, referred to as a culture.

**The Measurement of Microbial Concentrations**

Microbial concentrations can be measured in terms of cell concentration (the number of viable cells per unit volume of culture) or of biomass concentration (dry weight of cells per unit volume of culture). These two parameters are not always equivalent because the average dry weight of the cell varies at different stages of a culture. Nor are they of equal significance: For example, in studies of microbial genetics and the inactivation of microbes, cell concentration is the significant quantity; in studies on microbial biochemistry or nutrition, biomass concentration is the significant quantity.

**A. Viable Cell Count**

The **viable cell count** is typically considered the measure of cell concentration. For this, a 1-mL volume is removed from a bacterial suspension and serially diluted 10-fold followed by plating 0.1-mL aliquots on an agar medium. Each single invisible bacterium (or clump of

bacteria) will grow into a visible colony that can be counted. For statistical purposes, plates containing between 30 and 300 colonies give the most accurate data. The plate count × the dilution × 10 will give the number of colony forming units (CFU)/mL in the undiluted bacterial suspension. Using this method, dead bacteria within the suspension do not contribute to the final bacterial count.

**B. Turbidity**

For most purposes, the **turbidity** of a culture, measured by photoelectric means, can be related to the viable count using a **standard curve**. As an alternative a rough visual estimate is sometimes possible: For example, a barely turbid suspension of *Escherichia coli* contains about 107 cells per milliliter, and a fairly turbid suspension contains about 108 cells per milliliter. The correlation between turbidity and viable count can vary during the growth and death of a culture; cells may lose viability without producing a loss in turbidity of the culture.

**C. Biomass Density**

In principle, biomass can be measured directly by determining the dry weight of a microbial culture after it has been washed with distilled water. In practice, this procedure is cumbersome, and the investigator customarily prepares a standard curve that correlates dry weight with viable cell count. Alternatively, the concentration of biomass can be estimated indirectly by measuring an important cellular component such as protein or by determining the volume occupied by cells that have settled out of suspension.

 ENVIRONMENTAL CONTROL OF MICROBIAL GROWTH

The robust nature of uncontrolled microbial growth clearly presents a conflict with human life. To coexist with bacteria, higher species have to control bacterial growth. We, as humans, do this in a biologic context using an immune system and nutrient limitation. We also use physical methods to prevent exposure to microorganisms. Terms like **sterilization**, **disinfection**, **pasteurization**, and **aseptic** need to be precisely understood so as to articulate them in a proper sense. A list of these terms and their definitions are provided in Table 4-3.

As an example of the importance of understanding these terms, we speak of **sterilization** as the process of killing all the organisms, including spores, in a given preparation. Understanding this concept would be particularly important for surgical instruments because one would not want to introduce spores into the surgical site. By contrast, “disinfecting” these instruments may eliminate the bacteria but not the spores. Further, physically “cleaning” the instruments may not remove all of the bacteria and spores but simply decrease the bioburden on the instrument. The point is that an understanding of the terms used in Table 4-3 is critical to controlling the environmental impact of microorganisms in the context of human health.

STRATEGIES TO CONTROL BACTERIA AT THE ENVIRONMENTAL LEVEL

In medical microbiology one often considers the control of bacteria infecting humans with antibiotics as the gold standard in treating infections. While true, the real first line is to prevent exposure to infectious agents. For example, nearly 240,000 deaths annually occur worldwide as a result of neonatal tetanus. Yet this disease is very rare in developed countries. A major contributing factor is the inability to “sterilize” instruments (in addition to routine immunization with the tetanus vaccine) in many developing countries. If proper practices were used in underdeveloped regions, this disease could be substantially eliminated. Thus, one must understand methods of *sterilization*, *disinfection*, and *pasteurization*, among others. The techniques used to mitigate microbial infection should be understood at the mechanism of action level in order to apply them in the appropriate situation. Table 4-4 represents a nonexhaustive list of routinely used biocides. It is important to understand the terms **bacteriostatic** and **bactericidal** as defined in Table 4-4. The general mechanisms by which these biocides accomplish their antimicrobial activity are summarized in the following section.

GENERAL MECHANISMS OF BIOCIDE ACTION

**Disruption of the Cell Membrane or Wall**

The cell membrane acts as a selective barrier, allowing some solutes to pass through and excluding others. Many compounds are actively transported through the membrane, becoming concentrated within the cell. The membrane is also the site of enzymes involved in the biosynthesis of components of the cell envelope. Substances that concentrate at the cell surface may alter the physical and chemical properties of the membrane, preventing its normal functions and therefore killing or inhibiting the cell. The cell wall acts as a corseting structure (best characterized as a fishing net), protecting the cell against osmotic lysis. Thus, agents that destroy the wall (eg, lysozyme, which cleaves the sugar linkages of peptidoglycan) or prevent its normal synthesis (eg, penicillin, which interrupts peptidyl cross-linkages) may bring about lysis of the cell.

**Protein Denaturation**

Proteins exist in a folded, three-dimensional state determined primarily by intramolecular noncovalent interactions such as ionic, hydrophobic, and hydrogen bonds or covalent disulfide linkages. This state is called the tertiary structure of the protein; it is readily disrupted by a number of physical (eg, heat) or chemical (eg, alcohol) agents, causing the protein to become nonfunctional. The disruption of the tertiary structure of a protein is called **protein denaturation**.

**Disruption of Free Sulfhydryl Groups**

Enzymes containing cysteine have side chains terminating in sulfhydryl groups. In addition to these, coenzymes such as coenzyme A and dihydrolipoate contain free sulfhydryl groups. Such enzymes and coenzymes cannot function unless the sulfhydryl groups remain free and reduced. Oxidizing agents thus interfere with metabolism by forming disulfide

linkages between neighboring sulfhydryl groups. Many metals such as mercuric ions likewise interfere by combining with sulfhydryls. There are sulfhydryl-containing enzymes in the cell, so oxidizing agents and heavy metals do widespread damage.

***Common Terms Related to Microbial Control ***

**Examples of Mechanisms That Can Reverse the Activity of Biocides**

******

***Germicidal Properties of Disinfectants and Antiseptic Agents***

******

Appreciating the growth and death of bacteria is fundamental to understanding the complex interaction that exists between pathogenic bacteria and their hosts. If unchecked by an intact immune system and nutrient limitation, logarithmic growth of bacteria would quickly outcompete the host for nutrients. The environmental control of microbial growth by biocides limits exposure to potentially pathogenic microorganisms. The concepts of sterilization, disinfection, pasteurization, and others, are central to bacterial control and ultimately to human health. In the end, understanding microbial growth and death is the first step toward effectively managing infectious diseases.

KEY CONCEPTS

1. Bacteria in humans exist as complex biosystems known as microbiota.

2. Quantification of bacterial cells can be accomplished using viable cell count, turbidity, and biomass.

3. Biomass and generation time are mathematically related.

4. Inoculating a single bacterial colony into a fixed volume of liquid medium is known as batch culture. In this system, bacterial growth exhibits four phases—lag, log, stationary, and death.

5. Some bacteria exist in a state that is defined as viable but not culturable.

6. Growth in continuous culture or as a biofilm more closely approximates bacterial growth within the human host.

7. Sterilization, disinfection, pasteurization, as well as other terms are critical to understanding and communicating the science of microbiology.

8. The general structures of biocides and mechanisms of action should be understood.

9. Depending on the mechanism of action, different biocides are bacteriostatic, bactericidal, and/or sporicidal.

10. Biocide activity is dependent on time and concentration. This activity can be reversed by agent removal, substrate competition, and agent inactivation.

***Human Microbiome in Health and Disease***

Up until the time of birth, the human fetus lives in a remarkably protected and for the most part sterile environment; however, this rapidly changes as the infant is exposed to bacteria, archaea, fungi, and viruses from the mother, other close contacts, and the environment. Over the next few years, communities of organisms (microbiota or normal flora) form on the surfaces of the skin, nares, oral cavity, intestines, and genitourinary tract. The focus of this chapter is to gain an understanding of the role these communities play in the metabolic and immunologic functions of healthy individuals, factors regulating the composition of these communities, and how disruption of these communities can result in disease states.

***Glossary of Terms***

******

**Role of the Microbiome in Disease**

If the normal microbiome characterizes health, then alterations in the microbiome can signify disease; this is a relationship we are only beginning to understand. In 1884 Robert Koch and Friedrich Loeffler defined the relationship between an organism and infection. The **Koch postulates** were based on the concept of one organism:one disease. Microbiome research has introduced a new concept of disease caused by a community of organisms rather than a single species of bacteria, and the influence extends beyond traditional “infectious” diseases to include immunologic and metabolic disorders such as inflammatory bowel disease, obesity, type 2 diabetes, and celiac disease. We are now at the forefront of a new era of redefining the concept of infectious diseases.

Disruption of the normal microflora (commonly referred to as **dysbiosis**) can lead to disease by the elimination of needed organisms or allowing the growth of inappropriate bacteria. For example, following exposure to antibiotics and suppression of the intestinal normal flora, *C. difficile* is able to proliferate and express enterotoxins, leading to inflammation of the colon **(antibiotic-associated colitis).** Another disease of the colon, **ulcerative colitis,** is associated with an increased level of bacteria producing mucin-degrading sulfatases, leading to degradation of the protective mucosal lining of the intestinal wall and stimulation of inflammatory immune responses. Individuals with an intestinal microbiota that is more efficient at breaking down complex carbohydrates internalize rather than void these nutrients; therefore they are susceptible to **obesity** and a predisposition to metabolic syndromes such as **type 2** **diabetes.** Not all patients genetically predisposed to **celiac** **disease,** which is an immune-mediated enteropathy precipitated by exposure to gluten proteins, are symptomatic. The intestinal microbiota of most individuals is composed of bacteria capable of digesting glutens, which may be sufficient to protect these genetically predisposed individuals. In the absence of these bacteria, disease may occur. Shifts in the skin microbiome are associated with progression to **chronic wound infections** and episodic exacerbations of **atopic dermatitis.** Alteration in the vaginal microbiome from relatively few predominant organisms to a heterogeneous mixed population is associated with the progression to **vaginitis.**

**Diagnostics and Therapeutics**

An understanding of the influence of dysbiosis on disease pathology can lead to both advanced diagnostic tests and paths for therapeutic intervention. Just as the presence of *Salmonella* or *Shigella* signifies disease, changes in the diversity and composition of the fecal microflora can also indicate susceptibility to or onset of disease. The most obvious example is *C. difficile* disease, which is a clinical disease preceded by a depletion of the normal flora because of antibiotic use. Interestingly, patients with chronic relapsing *C. difficile* infections are treated successfully by repopulating (some say **“repoopulating”**) the intestines with stool transplants from a healthy spouse or close relative, or with artificially created stool specimens consisting of a complex mixture of aerobic and anaerobic fecal organisms. More subtle alterations in the gut microbiome may predict development of diseases such as **necrotizing** **enterocolitis (NEC),** inflammatory bowel disease, and a predilection for obesity. NEC is a devastating intestinal disease that afflicts preterm infants. Prospectively collected stool samples from infants younger than 29 weeks’ gestational age who develop NEC demonstrate a distinct dysbiosis prior to the development of disease. Infants with early-onset disease have a dominance of Firmicutes (predominantly *Staphylococcus*), whereas infants with late-onset NEC have a dominance of Enterobacteriaceae. The effects of microbiome alterations have also been described for the pathogenesis of inflammatory bowel disease and colorectal cancer. Proliferation of bacteria such as *Akkermansia muciniphila* that produce mucin-degrading sulfatases is responsible for degradation of the intestinal wall lining. Additionally, an increase in members of the anaerobic family Prevotellaceae leads to upregulation of chemokine mediated inflammation. Enterotoxigenic *Bacteroides* *fragilis* can also induce T helper cell mediated inflammatory responses that are associated with colitis and are a precursor to colonic hyperplasia and colorectal tumors. Finally, *Methanobrevibacter smithii,* a minor member of the gut microbiome, enhances digestion of dietary glycans by *B. thetaiotaomicron* and other core intestinal bacteria, leading to accumulation of fat. Alterations of the microbiome leading to disease may not be characterized by the presence or absence of a specific microbe because more than one organism may provide the needed function. It is likely that future diagnostics will measure for the presence or absence of a specific gene product **(proteomics)** or metabolic function **(metabolomics)**.

**Probiotics**

Probiotics are mixtures of bacteria or yeast that when ingested colonize and proliferate, even temporarily, the intestine. Consumers of probiotics believe they act by rebalancing the microbiome and its functions, such as enhancing digestion of food and modulating the individual’s innate and immune response. The most common reason people use over-the-counter probiotics is to promote and maintain regular bowel function and improve tolerance to lactose. Probiotics are commonly gram-positive bacteria (e.g., *Bifidobacterium, Lactobacillus*) and yeasts (e.g., *Saccharomyces*). Many of these microbes are found in ingestible capsules and as food supplements (e.g., yogurt, kefir). Probiotics have been used to treat *C. difficile*–associated diarrhea and inflammatory bowel disease, to provide protection from *Salmonella* and *Helicobacter pylori* disease, as therapy for pediatric atopic dermatitis and autoimmune diseases, and even for reduction in dental caries, although the value of probiotics for many of these conditions is unproven. Although probiotics are generally safe dietary supplements, many probiotics are ineffective. The species, mixture of species, and dose and viability of the probiotic organisms within a probiotic formulation influence its

potency, efficacy, and therapeutic potential. What is clear is that much like the use of complex artificial mixtures of organisms to treat recurrent *C. difficile* disease, carefully designed “smart probiotics” will likely be an important adjunct to medical therapy in the future.

 In the near future, with faster and cheaper DNA sequencing procedures, analysis of a person’s microbiome may become a routine diagnostic test for predicting and treating a wide range of diseases. However, a number of questions remain to be resolved: Can we predict disease in an individual by monitoring changes in the microbiome? Which changes are most important, taxonomic or genetic function? Can we prevent disease or treat disease by reestablishing a healthy microbiome? Can this be done by prescribing specific replacement microbes (e.g., fecal transplant) or with a universal mixture (probiotic)? Can the use of metabolic supplements **(prebiotics)** promote a healthy microbiota? Will use of antibiotics be replaced by use of “smart microbiome” therapies? Other questions include: What is the role of the host genome, environmental factors, and our hygienic practices in shaping the microbiome? What will be the informatic requirements for guiding diagnostics or therapeutics? Regardless of the answers to these and other questions, it is certain that we are witnessing the beginning of a new era of microbiology that can radically change our approach to prediction, diagnosis, and treatment of disease.

Microbial Genetics

The science of **genetics** defines and analyzes **heredity** of the vast array of structural and physiologic functions that form the properties of organisms. The basic unit of heredity is the **gene**, a segment of deoxyribonucleic acid (**DNA**) that encodes in its nucleotide sequence information for a specific physiologic property. The traditional approach to genetics has been to identify genes on the basis of their contribution to **phenotype**, or the collective structural and physiologic properties of an organism. A phenotypic property, be it eye color in humans or resistance to antibiotics in a bacterium, is generally observed at the level of the organism. The chemical basis for variation in phenotype is change in **genotype**, or alteration in the DNA sequence, within a gene or within the organization of genes.

**DNA** as the fundamental element of heredity was suggested in the 1930s from a seminal experiment performed by Frederick Griffith. In this experiment, killed virulent *Streptococcus pneumoniae* type III-S (possessing a capsule), when injected into mice along with living but nonvirulent type II-R pneumococci (lacking a capsule), resulted in a lethal infection from which viable type III-S pneumococci were recovered. The implication was that some chemical entity transformed the live, nonvirulent strain to the virulent phenotype. A decade later, Avery, MacLeod, and McCarty discovered that DNA was the transforming agent. This formed the foundation for molecular biology as we understand it today. Recombinant DNA technology was born in the 1960s and 1970s when investigations with bacteria revealed the presence of **restriction enzymes**, proteins that cleave DNA at specific sites, giving rise to DNA **restriction fragments**.

**Plasmids** were identified as small genetic elements carrying genes and capable of independent replication in bacteria and yeasts. The introduction of a DNA restriction fragment into a plasmid allows the DNA fragment to be amplified many times. Amplification of specific regions of DNA also can be achieved with bacterial enzymes using **polymerase** **chain reaction** (**PCR**) or other enzyme-based method of nucleic acid amplification. DNA amplified by these sources and digested with appropriate restriction enzymes can be inserted into plasmids. Genes can be placed under control of high-expression bacterial **promoters** that allow encoded proteins to be expressed at increased levels. Bacterial genetics

have fostered the development of **genetic engineering** not only in prokaryotes but also in eukaryotes. This technology is responsible for the tremendous advances in the field of medicine realized today.

NUCLEIC ACIDS AND THEIR ORGANIZATION IN EUKARYOTIC, PROKARYOTIC, AND VIRAL GENOMES

Genetic information in bacteria is stored as a sequence of **DNA bases.** Most DNA molecules are double stranded, with **complementary bases** (A-T; G-C) paired by hydrogen bonding in the center of the molecule. The orientation of the two DNA strands is **antiparallel**: One strand is chemically oriented in a 5′→3′ direction, and its complementary strand runs 3′→5′. The complementarity of the bases enables one strand (**template strand**) to provide the information for copying or expression of information in the other strand (**coding strand**). The base pairs are stacked within the center of the DNA double helix, and they determine its genetic information. Each turn of the helix has one major groove and one minor groove. Certain proteins have the capacity to bind DNA and regulate gene expression by interacting predominately with the major groove, where atoms comprising the bases are more exposed. Each of the four bases is bonded to phospho-2′-deoxyribose to form a **nucleotide**. The negatively charged phosphodiester backbone of DNA faces the solvent. The length of a DNA molecule is usually expressed in thousands of base pairs, or **kilobase** **pairs** (**kbp**). Whereas a small virus may contain a single DNA molecule of less than 0.5 kbp, the single DNA genome that encodes *Escherichia coli* is greater than 4000 kbp. In either case, each base pair is separated from the next by about 0.34 nm, or 3.4 × 10−7 mm, so that the total length of the *E coli* chromosome is roughly 1 mm. Because the overall dimensions of the bacterial cell are roughly 1000-fold smaller than this length, it is evident that a substantial amount of folding, or **supercoiling**, contributes to the physical structure of the molecule in vivo.

**Ribonucleic acid** (**RNA**) most frequently occurs in single- stranded form. The uracil base (U) replaces thymine base (T) in DNA, so the complementary bases that determine the structure of RNA are A-U and C-G. The overall structure of single-stranded RNA (ssRNA) molecules is determined by pairing between bases within the strand-forming loops, with the result that ssRNA molecules assume a compact structure capable of expressing genetic information contained in DNA. The most general function of RNA is communication of DNA gene sequences in the form of **messenger** **RNA** (**mRNA**) to **ribosomes**. These processes are referred to as **transcription** and **translation**. mRNA (referred to as +ssRNA) is transcribed as the RNA complement to the coding DNA strand. This mRNA is then translated by ribosomes. The ribosomes, which contain both **ribosomal RNA** (**rRNA**) and proteins, translate this message into the primary structure of proteins via aminoacyl-**transfer RNAs** (**tRNAs**). RNA molecules range in size from the small tRNAs, which contain fewer than 100 bases, to mRNAs, which may carry genetic messages extending to several thousand bases. Bacterial ribosomes contain three kinds of rRNA, with respective sizes of 120, 1540, and 2900 bases, and a number of proteins. Corresponding rRNA molecules in eukaryotic ribosomes are somewhat larger. The need for expression of an individual gene changes in response to physiologic demand, and requirements for flexible gene expression are reflected in the rapid metabolic turnover of most mRNAs. On the other hand, tRNAs and rRNAs—which are associated with the universally required function of protein synthesis—tend to be stable and together account for more than 95% of the total RNA in a bacterial cell. A few RNA molecules have been shown to function as enzymes (**ribozymes**). For example, the 23S RNA in the 50S ribosomal subunit catalyzes the formation of the peptide bond during protein synthesis.

**The Eukaryotic Genome**

The **genome** is the totality of genetic information in an organism.

Almost all of the eukaryotic genome is carried on two or more linear chromosomes separated from the cytoplasm within the membrane of the nucleus. **Diploid** eukaryotic cells contain two **homologues** (divergent evolutionary copies) of each chromosome. **Mutations**, or genetic changes, frequently cannot be detected in diploid cells because the contribution of one gene copy compensates for changes in the function of its homologue. Whereas a gene that does not achieve phenotypic expression in the presence of its homologue is **recessive**, a gene that overrides the effect of its homologue is **dominant**.

The effects of mutations can be most readily discerned in **haploid** cells, which carry only a single copy of most genes. Yeast cells (which are eukaryotic) are frequently investigated because they can be maintained and analyzed in the haploid state. Of the entirety of the human genome, only 2% is considered **coding DNA**, the rest is **noncoding DNA**.

Eukaryotic cells contain **mitochondria** and, in the case of plants, **chloroplasts**. Within each of these organelles is a circular molecule of DNA that contains a few genes whose function relates to that particular organelle. Most genes associated with organelle function, however, are carried on eukaryotic chromosomes. Many yeast contain an additional genetic element, an independently replicating 2-μm circle containing about 6.3 kbp of DNA. Such small circles of DNA, termed **plasmids** or **episomes**, are frequently associated with prokaryotes. The small size of plasmids renders them amenable to genetic manipulation and, after their alteration, may allow their introduction into cells. Therefore, plasmids are commonly used in genetic engineering.

**Repetitive DNA**, which occurs in large quantities in eukaryotic cells, has been increasingly identified in prokaryotes. In eukaryotic genomes, repetitive DNA is infrequently associated with coding regions and is located primarily in extragenic regions. These short-sequence repeats (SSRs) or short tandemly repeated (STR) sequences occur in several to thousands of copies dispersed throughout the genome. The presence of prokaryotic SSRs and STRs is well documented, and some show extensive length-polymorphisms. This variability is thought to be caused by slipped-strand mispairing and is an important prerequisite for bacterial phase variation and adaptation. Many eukaryotic genes are interrupted by **introns**, intervening sequences of DNA that are missing in processed mRNA when it is translated. Introns have been observed in archaebacterial genes but with a few rare exceptions are not found in eubacteria.

**The Prokaryotic Genome**

Most prokaryotic genes are carried on the bacterial chromosome. And with few exceptions, bacterial genes are haploid. Genome sequence data from more than 340 microbial genomes demonstrate that most prokaryotic genomes (>90%) consist of a single circular DNA molecule containing from 580 kbp to more than 5220 kbp of DNA (Table 7-1). A few bacteria (eg, *Brucella melitensis*, *Burkholderia pseudomallei*, and *Vibrio cholerae*) have genomes consisting of two circular DNA molecules. Many bacteria contain additional genes on plasmids that range in size from several to 100 kbp. In contrast to eukaryotic genomes, 98% of bacterial genomes are coding sequences.

Covalently closed DNA circles (bacterial chromosomes and plasmids), which contain genetic information necessary for their own replication, are called **replicons** or **episomes**. Because prokaryotes do not contain a nucleus, a membrane does not separate bacterial genes from cytoplasm as in eukaryotes.

Some bacterial species are efficient at causing disease in higher organisms because they possess specific genes for pathogenic determinants. These genes are often clustered together in the DNA and are referred to as **pathogenicity** **islands**. These gene segments can be quite large (up to 200 kbp) and encode a collection of virulence genes. Pathogenicity islands (1) have a different G + C content from the rest of the genome; (2) are closely linked on the chromosome to tRNA genes; (3) are flanked by direct repeats; and (4) contain diverse genes important for pathogenesis, including antibiotic resistance, adhesins, invasins, and exotoxins, as well as genes that can be involved in genetic mobilization. Genes essential for bacterial growth (often referred to as “housekeeping genes”) can be carried on the chromosome or may be found on plasmids that carry genes associated with specialized functions. Many plasmids also encode genetic sequences that mediate their transfer from one organism to another (eg, those involved with sex pili) as well as others associated with genetic acquisition or rearrangement of DNA (eg, transposase). Therefore, genes with independent evolutionary origins may be assimilated by plasmids that are widely disseminated among bacterial populations. A consequence of such genetic events has been observed in the swift spread among bacterial populations of plasmid-borne resistance to antibiotics after their liberal use in hospitals.

**Transposons** are genetic elements that contain several genes, including those necessary for their migration from one genetic locus to another. In doing so, they create **insertion** **mutations**. The involvement of relatively short transposons (0.75–2.0 kbp long), known as **insertion elements**, produces the majority of insertion mutations. These insertion elements (also known as insertion sequence [IS] elements) carry only the genes for enzymes needed to promote their own transposition to another genetic locus but cannot replicate on their own. Almost all bacteria carry IS elements, with each species harboring its own characteristic IS elements. Related IS elements can sometimes be found in different bacteria, implying that at some point in evolution they have crossed species barriers. Plasmids also carry IS elements, which are important in the formation of high-frequency recombinant (**Hfr**) strains (see below). Complex transposons carry genes for specialized functions such as antibiotic resistance and are flanked by insertion sequences.

Transposons do not carry the genetic information required to encode their own replication, and therefore their propagation depends on their physical integration with a bacterial replicon. This association is fostered by enzymes that confer the ability of transposons to form copies of themselves; these enzymes may allow the transposons to integrate within the same replicon or an independent replicon. The specificity of sequence at the insertion site is generally low, so that transposons often seem to insert in a random pattern, but theytend to favor regions encoding tRNAs. Many plasmids are transferred among bacterial cells, and insertion of a transposon into such a plasmid is a vehicle that leads to the transposon’s dissemination throughout a bacterial population.

**The Viral Genome**

Viruses are capable of survival, but not growth, in the absence of a cell host. Replication of the viral genome depends on the metabolic energy and the macromolecular synthetic machinery of the host. Frequently, this form of genetic parasitism results in debilitation or death of the host cell. Therefore, successful propagation of the virus requires (1) a stable form that allows the virus to survive in the absence of its host, (2) a mechanism for invasion of a host cell, (3) genetic information required for replication of the viral components within the cell, and (4) additional information that may be required for packaging the viral components and liberating the resulting virus from the host cell. Distinctions are frequently made between viruses associated with eukaryotes and viruses associated with prokaryotes, the latter being termed **bacteriophage** or **phage**. When viral DNA is integrated into the eukaryotic genome, it is called a **provirus**; when a phage is integrated into a bacterial genome or episome, it is called a **prophage**. With more than 5000 isolates of known morphology, phages constitute the largest of all viral groups. Much of our understanding of viruses—indeed, many fundamental concepts of molecular biology—has emerged from investigation of bacteriophages.

Bacteriophages occur in more than 140 bacterial genera and in many different habitats. The nucleic acid molecule of bacteriophages is surrounded by a protein coat. Considerable variability is found in the nucleic acid of phages. Many phages contain double-stranded DNA (dsDNA); others contain double- stranded RNA (dsRNA), ssRNA, or single-stranded DNA (ssDNA). Unusual bases such as hydroxymethylcytosine are sometimes found in the phage nucleic acid. Bacteriophages exhibit a wide variety of morphologies. Many phages contain specialized syringe-like structures (tails) that bind to receptors on the cell surface and inject the phage nucleic acid into a host cell Phages can be distinguished on the basis of their mode of propagation. **Lytic phages** produce many copies of themselves as they kill their host cell. The most thoroughly studied lytic phages, the T-even (eg, T2, T4) phages of *E coli*, demonstrate the need for precisely timed expression of viral genes to coordinate events associated with phage formation. **Temperate** **phages** are able to enter a nonlytic **prophage** state in which replication of their nucleic acid is linked to replication of host cell DNA. Bacteria carrying prophages are termed **lysogenic** because a physiologic signal can trigger a lytic cycle resulting in death of the host cell and liberation of many copies of the phage. The best characterized temperate phage is the *E coli* phage λ (lambda). **Filamentous phages**, exemplified by the well-studied *E coli* phage M13, are exceptional in several respects. Their filaments contain ssDNA complexed with protein and are extruded from their bacterial hosts, which are debilitated but not killed by the phage infection. Engineering of DNA into phage M13 has provided single strands that are valuable sources for DNA analysis and manipulation.

REPLICATION

dsDNA is synthesized by **semiconservative replication**. As the parental duplex unwinds, each strand serves as a template (ie, the source of sequence information) for DNA replication. New strands are synthesized with their bases in an order complementary to that in the preexisting strands. When synthesis is complete, each daughter molecule contains one parental strand and one newly synthesized strand.

**Bacterial DNA**

The replication of bacterial DNA begins at one point and moves in both directions (ie, **bidirectional replication**). In the process, the two old strands of DNA are separated and used as templates to synthesize new strands (**semiconservative** **replication**). The structure where the two strands are separated and the new synthesis is occurring is referred to as the **replication fork**. Replication of the bacterial chromosom is tightly controlled, and the number of each chromosomes (when more than one is present) per growing cell falls between one and four. Some bacterial plasmids may have as many as 30 copies in one bacterial cell, and mutations causing relaxed control of plasmid replication can result in 10-fold higher copy numbers.

The replication of circular double-stranded bacterial DNA begins at the *ori* locus and involves interactions with several proteins. In *E coli*, chromosome replication terminates in a region called *ter*. The **origin** (*ori*) and **termination sites** (*ter*) for replication are located at opposite points on the circular DNA chromosome. The two daughter chromosomes are separated, or resolved, before cell division, so that each progeny cell gets one of the daughter DNAs. This is accomplished with the aid of **topoisomerases**, enzymes that alter the supercoiling of dsDNA. The topoisomerases act by transiently cutting one or both strands of the DNA to relax the coil and extend the DNA molecule. Because bacterial topoisomerases are essential and unique, they are targets of **antibiotics** (eg, quinolones). Similar processes used in the replication of bacterial chromosomes are used in the replication of plasmid DNA, except that, in some cases, replication is unidirectional.

TRANSFER OF DNA

The haploid nature of the bacterial genome might be presumed to limit the genomic plasticity of a bacterium. However, the ubiquity of diverse bacteria in a complex microbiome provides a fertile gene pool that contributes to their remarkable genetic diversity through mechanisms of genetic exchange. Bacterial genetic exchange is typified by transfer of a relatively small fragment of a donor genome to a recipient cell followed by genetic recombination. Bacterial genetic recombination is quite unlike the fusion of gametes observed with eukaryotes; it demands that this donor DNA be replicated in the recombinant organism. Replication can be achieved either by integration of the donor DNA into the recipient’s chromosome or by establishment of donor DNA as an independent episome.

**Restriction and Other Constraints on Gene Transfer**

**Restriction enzymes** (restriction endonucleases) provide bacteria with a mechanism to distinguish between their own DNA and DNA from other biologic sources. These enzymes hydrolyze (cleave) DNA at restriction sites determined by specific DNA sequences ranging from 4 to 13 bases. Each bacterial strain that possesses a restriction system is able to disguise these recognition sites in its own DNA by modifying them through methylation of adenine or cytosine residues within the site. These restriction–modification systems fall into two broad classes: type I systems, in which the restriction and modification activities are combined in a single multisubunit protein, and type II systems, which consist of separate endonucleases and methylases. A direct biologic consequence of restriction can be cleavage of donor DNA before it has an opportunity to become established as part of a recombinant replicon, rendering the bacterium “immune” to incoming DNA.

Some plasmids exhibit a narrow host range and are able to replicate only in a closely related set of bacteria. Other plasmids, exemplified by some drug resistance plasmids, replicate in a broad range of bacterial genera. In some cases, two or more plasmids can stably coexist in a cell, but other pairs will interfere with the replication or partitioning. If two such plasmids are introduced into the same cell, one or the other will be lost at a higher than normal rate when the cell divides. This phenomenon is called **plasmid incompatibility**; two plasmids that cannot stably coexist belong to the same **incompatibility (Inc) group**, and two plasmids that can stably coexist belong to different Inc groups.

**Mechanisms of Recombination**

Donor DNA that does not carry information necessary for its own replication must recombine with recipient DNA to become established in a recipient strain. The recombination may be **homologous**, a consequence of close similarity in the sequences of donor and recipient DNA, or **nonhomologous**, the result of enzyme-catalyzed recombination between two dissimilar DNA sequences. Homologous recombination almost always involves exchange between genes that share common ancestry. The process requires a set of genes designated *rec*. Nonhomologous recombination depends on enzymes encoded by the integrated DNA and is most clearly exemplified by the insertion of DNA into a recipient to form a copy of a donor transposon.

The mechanism of recombination mediated by *rec* gene products is reciprocal: Introduction of a donor sequence into a recipient is mirrored by transfer of the homologous recipient sequence into the donor DNA. Increasing scientific attention is being paid to the role of **gene conversion**—the nonreciprocal transfer of DNA sequences from donor to recipient—in the acquisition of genetic diversity.

**Mechanisms of Gene Transfer**

The DNA composition of microorganisms is remarkably fluid. DNA can be transferred from one organism to another, and that DNA can be stably incorporated in the recipient, permanently changing its genetic composition. This process is called **horizontal gene transfer** (HGT) to differentiate it from the inheritance of parental genes, a process called **vertical** inheritance. Three broad mechanisms mediate efficient movement of DNA between cells—**conjugation**, **transduction**, and **transformation**. **Conjugation** requires donor cell-to recipient cell contact to transfer only one strand of DNA (Figure 7-6). The recipient completes the structure of dsDNA by synthesizing the strand that complements the strand acquired from the donor. In **transduction**, donor DNA is carried by a phage coat and is transferred into the recipient by the mechanism used for phage infection. **Transformation**, the direct uptake of “naked” donor DNA by the recipient cell, may be natural or forced. Forced transformation is induced in the laboratory, where, after treatment with high salt and temperature shock, many bacteria are rendered competent for the uptake of extracellular plasmids. The capacity to force bacteria to incorporate extracellular plasmids by transformation is fundamental to genetic engineering.

**A. Conjugation**

Plasmids are most frequently transferred by conjugation. Genetic functions required for transfer are encoded by the *tra* genes, which are carried by self-transmissible **plasmids**. Some self-transmissible plasmids can mobilize other plasmids or portions of the chromosome for transfer. In some cases, mobilization is achieved because the *tra* genes provide functions necessary for transfer of an otherwise nontransmissible plasmid (Figures 7-7 and 7-8). In other cases, the self-transmissible plasmid integrates with the DNA of another replicon

and, as an extension of itself, carries a strand of this DNA into a recipient cell. Genetic analysis of *E coli* was greatly advanced by elucidation of **fertility** factors carried on a plasmid designated F+. This plasmid confers certain donor characteristics upon cells; these characteristics include a sex pilus, an extracellular multimeric protein extrusion that attaches donor cells to recipient organisms lacking the fertility factor. A bridge between the cells allows a strand of the F+ plasmid, synthesized by the donor, to pass into the recipient, where the complementary strand of DNA is formed. The F+ fertility factor can integrate into numerous loci in the chromosome of donor cells. The integrated fertility factor creates high-frequency recombination (Hfr) donors from which chromosomal DNA is transferred (from the site of insertion) in a direction determined by the orientation of insertion. The rate of chromosomal transfer from Hfr cells is constant, and compilation of results from many conjugation experiments has allowed preparation of an *E coli* **genetic map** in which distances between loci are measured in number of minutes required for transfer in conjugation. A similar map has been constructed for the related coliform (*E coli*–like) bacterium *Salmonella typhimurium*, and comparison of the two maps shows related patterns of genomic organization. This type of mapping has now been replaced by high throughput genomic sequencing.

Integration of chromosomal DNA into a conjugal plasmid can produce a recombinant replicon—an **F** (fertility) prime, or **R** (resistance) prime, depending on the plasmid—in which the integrated chromosomal DNA can be replicated on the plasmid independently of the chromosome. This occurs when the integrated plasmid (eg, F) is bracketed by two copies of an IS element. Bacteria carrying gene copies, a full set on the chromosome and a partial set on a prime, are **partial** **diploids**, or **merodiploids**, and are useful for complementation studies. A wild-type gene frequently complements its mutant homologue, and selection for the wild-type phenotype can allow maintenance of merodiploids in the laboratory. Such strains can allow analysis of interactions between different **alleles**, genetic variants of the same gene. Merodiploids frequently are genetically unstable because recombination between the plasmid and the homologous chromosome can result in loss or exchange of mutant or wild-type alleles. This problem can frequently be circumvented by maintenance of merodiploids in a genetic background in which *recA*, a gene required for recombination between homologous segments of DNA, has been inactivated. Homologous genes from different organisms may have diverged to an extent that prevents homologous recombination between them but does not alter the capacity of one gene to complement the missing activity of another. For example, the genetic origin of an enzyme required for amino acid biosynthesis is unlikely to influence catalytic activity in the cytoplasm of a biologically distant host. A merodiploid carrying a gene for such an enzyme would also carry flanking genes derived from the donor organism. Therefore, conventional microbial genetics, based on selection of prime plasmids, can be used to isolate genes from fastidious organisms in *E coli* or *Pseudomonas aeruginosa.*

**B. Transduction**

Transduction is phage-mediated genetic recombination in bacteria. In simplest terms, a transducing particle might be regarded as bacterial nucleic acid in a phage coat. Even a lytic phage population may contain some particles in which the phage coat surrounds DNA derived from the bacterium rather than from the phage. Such populations have been used to transfer genes from one bacterium to another. Temperate phages are preferred vehicles for gene transfer because infection of recipient bacteria under conditions that favor lysogeny minimizes cell lysis and thus favors survival of recombinant strains. Indeed, a recipient bacterium carrying an appropriate prophage may form a repressor that renders the cell immune to lytic superinfection; such cells may still take up bacterial DNA from transducing particles. Transducing mixtures carrying donor DNA can be prepared under conditionsthat favor the lytic phage cycle. The size of DNA in transducing particles is usually no more than several percent of the bacterial chromosome, and therefore **cotransduction**—transfer of more than one gene at a time—is limited to linked bacterial genes. The speed and capacity by which phages recombine and replicate has made them central subjects for study of bacterial genetics and genetic engineering.

In nature, **pathogenicity islands** are often transported by phages. For example, two phages transport pathogenicity islands responsible for converting a benign form of *V cholerae* into the pathogenic form responsible for epidemic cholera (see Chapter 17). These phages encode genes for cholera toxin (responsible for symptoms) and bundle-forming pili (responsible for attachment) that in combination substantially increase the virulence of *V cholerae*.

**C. Transformation**

As described above, forced transformation is typically thought of as a laboratory phenomenon. However, it is now clear that low-frequency **HGT** has been responsible for common mechanisms of antibiotic resistance among diverse species of bacteria. This is not surprising given the complex diversity and density of the intestinal flora or the biofilms that form on our teeth overnight. Couple this with the therapeutic administration of antibiotics that select for resistant organisms and a “perfect storm” exists for the spread of genetic material across species boundaries. In contrast to forced transformation (described above), natural competence is unusual among bacteria. Direct uptake of donor DNA by recipient bacteria depends on their **competence** for transformation. Naturally competent transformable bacteria, of medical importance, are found in several genera and include *Haemophilus influenzae*, *Neisseria gonorrhoeae*, *Neisseria meningitidis*, and *S. pneumoniae.* Natural transformation is an active process demanding specific proteins produced by the recipient cell. In addition, specific DNA sequences (**uptake sequences**) are required for uptake of the DNA. These uptake sequences are species specific, thus restricting genetic exchange to a single species. The DNA that is not incorporated can be degraded and used as a source of nutrients to support microbial growth. It is clear that genetic transformation is a major force in microbial evolution.

MUTATION AND GENE REARRANGEMENT

**Spontaneous Mutations**

Spontaneous mutations for a given gene in a wild-type background generally occur with a frequency of 10−6–10−8 in a population derived from a single bacterium (depending on the bacterial species and conditions used to identify the mutation). The mutations include **base substitutions**, **deletions**, **insertions**, and **rearrangements**. Base substitutions can arise as a consequence of mispairing between complementary bases during replication. In *E coli*, this occurs about once every 1010 times the DNA polymerase incorporates a nucleotide—a remarkably rare process. Occurrence of a mispaired base is minimized by enzymes associated with **mismatch repair**, a mechanism that essentially proofreads a newly synthesized strand to ensure that it perfectly complements its template. Mismatch repair enzymes distinguish the newly synthesized strand from the preexisting strand on the basis of methylation of adenine in GATC sequences of the preexisting strand. When DNA damage is too extensive, a special DNA repair system, the **SOS response**, rescues cells. The SOS response is a postreplication DNA repair system that allows DNA replication to bypass extensive DNA errors. Many base substitutions escape detection at the phenotypic level because they do not significantly disrupt the function of the gene product. For example, **missense mutations**, which result in substitution of one amino acid for another, may be without discernible phenotypic effect. On the other hand, **nonsense mutations** terminate synthesis of proteins and thus result in a protein truncated at the site of mutation. The gene products of nonsense mutations are inactive.

**Rearrangements** are the result of deletions that remove large portions of genes or even sets of genes. These large deletions involve recombination between directly repeated sequences (eg, IS elements) and almost never revert. Other mutations cause duplication, frequently in tandem, of comparable lengths of DNA. Such mutations usually are unstable and readily revert. Still other mutations can invert lengthy DNA sequences or transpose such sequences to new loci. Comparative gene maps of related bacterial strains have shown that such rearrangements can be fixed in natural populations. These observations point to the fact that linear separation of DNA loci on a bacterial chromosome does not completely disrupt possibilities for physical and chemical interaction among them.

**Mutagens**

The frequency of mutation is greatly enhanced by exposure of cells to mutagens. UV light is a **physical mutagen** that damages DNA by linking neighboring thymine bases to form dimers. Sequence errors can be introduced during enzymatic repair of this genetic damage. **Chemical mutagens** may act by altering either the chemical or the physical structure of DNA. Reactive chemicals alter the structure of bases in DNA. For example, nitrous acid (HNO2) substitutes hydroxyl groups for amino groups on DNA basis. The resulting DNA has altered template activity during subsequent rounds of replication. A **frameshift mutation** is a **genetic mutation** caused by **insertions** or **deletions** of a number of **nucleotides** in a DNA sequence that is not divisible by 3. This is caused by polymerase slippage and is favored by exposure to acridine dyes (eg, acridine orange), which can intercalate between bases. In general, the direct effect of chemical or physical mutagens is damage to DNA. The resulting mutations are introduced by the replication process and escape the repair enzymes described above. Mutations that change the activity of replication or repair enzymes can make a bacterium more susceptible to biologic mutagens and are referred to as *mutator strains*.

**Reversion and Suppression**

Regaining an activity lost as a consequence of mutation, termed **phenotypic reversion**, may or may not result from restoration of the original DNA sequence, as would be demanded by **genotypic reversion**. Frequently, a mutation at a second locus, called a **suppressor mutation**, restores the lost activity. In **intragenic suppression**, after a primary mutation has changed an enzyme’s structure so that its activity has been lost, a second mutation at a different site in the enzyme’s gene restores the structure required for activity. **Extragenic** **suppression** is caused by a second mutation lying outside the originally affected gene.

GENE EXPRESSION

The tremendous evolutionary separation of eukaryotic and prokaryotic genomes is illustrated by comparing their mechanisms of gene expression, which share only a small subset of properties. In both groups, genetic information is encoded in DNA, transcribed into mRNA, and translated on ribosomes through tRNA into the structure of proteins. The triplet nucleotide codons used in translation are generally shared, and many enzymes associated with macromolecular synthesis in the two biologic groups have similar properties. The mechanism by which the sequence of nucleotides in a gene determines the sequence of amino acids in a protein is largely similar in prokaryotes and eukaryotes and is as follows:

1. RNA polymerase forms a single polyribonucleotide strand, called mRNA, using DNA as a template; this process is called **transcription**. The mRNA has a nucleotide sequence complementary to a template strand in the DNA double helix if read in the 3′–5′ direction. Thus, an mRNA is oriented in a 5′–3′ direction.

2. Amino acids are enzymatically activated and transferred to specific adapter molecules of RNA, called tRNA. Each adapter molecule has a triplet of bases (**anticodon**) complementary to a triplet of bases on mRNA, and at one end its specific amino acid. The triplet of bases on mRNA is called the **codon** specific for that amino acid.

3. mRNA and tRNA come together on the surface of the ribosome. As each tRNA finds its complementary nucleotide triplet on mRNA, the amino acid that it carries is put into peptide linkage with the amino acid of the preceding tRNA molecule. The enzyme **peptidyltransferase** (which is actually the 23S rRNA, ie, a **ribozyme**) catalyzes the formation of the peptide bond. The ribosome moves along the mRNA with the nascent polypeptide growing sequentially until the entire mRNA molecule has been translated into a corresponding sequence of amino acids. This process, called **translation**, is diagrammed in Figure 7-10. In prokaryotes, genes associated with related functions are typically clustered in **operons**. Because there is no nucleus, transcription and translation is coupled, meaning that the nascent mRNA attaches to a ribosome and is translated at the same time it is transcribed. This coupled transcription and translation allows for the rapid response to changes in the environment. Likewise, the mRNA is rapidly turned over, having a half-life of the order of seconds to minutes.

In eukaryotes, clustering of related genes is unusual. **Enhancer sequences** are regions of eukaryotic DNA that increase transcription and may lie distantly upstream from the transcribed gene. Eukaryotic genes carry **introns**, DNA insertions that are not found in prokaryotic genes. Introns separate **exons**, the coding regions of eukaryotic genes. Transcribed introns are removed from eukaryotic transcripts during RNA processing, a series of enzymatic reactions that takes place in the nucleus. The mRNA of eukaryotes is polyadenylated at the 3′ end, protecting it from exonucleases so that it can traverse the nuclear membrane into the cytosol, where the ribosomes are located; in this case, translation is uncoupled from transcription. Because of this polyadenylation, eukaryotic mRNAs have half-lives of the order of hours to days.

Eukaryotic and prokaryotic ribosomes differ in many respects. Eukaryotic ribosomes are larger and have a sedimentation coefficient of 80S compared with the 70S sedimentation coefficient of prokaryotic ribosomes. The 40S and 60S eukaryotic ribosomal subunits are larger than the corresponding 30S and 50S ribosomal subunits of prokaryotes, and the eukaryotic ribosomes are relatively rich in protein. Significant differences are inherent in the sensitivity of the ribosomal activities to antibiotics (eg, tetracycline), many of which selectively inhibit prokaryotic, but not eukaryotic protein synthesis (see Chapter 9). It should be remembered, however, that **mitochondrial** ribosomes in eukaryotes resemble those from prokaryotes and may be susceptible to bacterial protein synthesis inhibitors.

**Regulation of Prokaryotic Gene Expression**

Specific proteins, the products of regulatory genes, govern expression of structural genes that encode enzymes. Transcription of DNA into mRNA begins at the **promoter**, the DNA sequence that binds RNA polymerase. The level of gene expression is determined by the ability of a promoter to bind the polymerase, and the intrinsic effectiveness of promoters differs widely. Further controls over gene expression are exerted by regulatory proteins that can bind to regions of DNA near promoters. Many prokaryotic structural genes that encode a related series of metabolic reactions are clustered on **operons**. This contiguous series of genes are expressed as a single mRNA transcript, and expression of the transcript may be governed by a single regulatory gene. For example, five genes associated with tryptophan biosynthesis are clustered in the *trp* operon of *E coli*. Gene expression is governed by attenuation, as described below, and is also controlled by repression: Binding of the amino acid tryptophan by a **repressor protein** gives it a conformation that allows it to attach to the *trp* **operator**, a short DNA sequence that helps to regulate gene expression. Binding of the repressor protein to the operator prevents transcription of the *trp* genes because the bacterium senses that there is sufficient tryptophan present and making more would not be in the best interests of the organism’s metabolic resources. **Repression** can be viewed as a course-control mechanism, an all-or-none approach to gene regulation. This form of control is independent of attenuation, a fine-tuning mechanism that also is used to govern *trp* gene expression. **Attenuation** is a regulatory mechanism of some biosynthetic pathways (eg, the tryptophan biosynthetic pathway) that controls the efficiency of transcription after transcription has been initiated but before mRNA synthesis of the operon’s genes takes place, especially when the end product of the pathway is in short supply. For example, under normal growth conditions, most *trp* mRNA transcripts terminate before they reach the structural genes of the *trp* operon. However, during conditions of severe tryptophan starvation, the premature termination of transcription is abolished, allowing expression of the operon at 10-fold higher levels than under normal conditions. The explanation for this phenomenon resides in the 162 bp regulatory sequence in front of the *trp* structural genes (Figure 7-11) referred to as the **leader sequence** or *trpL*.

The *trp* leader sequence can be transcribed into mRNA and subsequently translated into a 14 amino acid polypeptide with two adjacent tryptophan residues, a sequence that occurs at a very rare occurrence. At the end of *trpL* and upstream of the regulatory signals that control translation of the *trp* structural genes is a **Rho-independent terminator**. The DNA sequence of this region suggests that the encoded mRNA has a high probability of forming **stem loop secondary structures**. These have been named the **pause loop** (1:2), the **terminator loop**

(3:4), and the **antiterminator loop** (2:3). Attenuation of the *trp* operon uses the secondary structure of the mRNA to sense theamount of tryptophan in the cell (as trp-tRNA) according tothe model shown in Figure 7-11.Prevention of transcription by a repressor protein is called**negative control**. The opposite form of transcriptional regulation—initiation of transcription in response to binding of an**activator protein**—is termed **positive control**. Both forms ofcontrol are exerted over expression of the *lac* operon, genesassociated with fermentation of lactose in *E coli*. The operoncontains three structural genes. Transport of lactose into thecell is mediated by the product of the *lacY* gene. β-Galactosidase,the enzyme that hydrolyzes lactose to galactose and glucose, isencoded by the *lacZ* gene. The product of the third gene (*lacA*)is a transacetylase; the physiologic function of this enzyme forlactose utilization has not been clearly elucidated.

As a byproduct of its normal function, β-galactosidase produces allolactose, a structural isomer of lactose. Lactose itself does not influence transcriptional regulation; rather, this function is served by allolactose, which is the **inducer** of the *lac* operon because it is the metabolite that most directly elicits gene expression. In the absence of allolactose, the *lac* repressor, a product of the independently controlled *lacI* gene, exerts negative control over transcription of the *lac* operon by binding to the *lac* operator. In the presence of the inducer, the repressor is released from the operator, and transcription takes place.

Expression of the *lac* operon and many other operons associated with energy generation is enhanced by the binding of **cyclic AMP–binding protein** (**CAP**) to a specific DNA sequence near the promoter for the regulated operon. The protein exerts positive control by enhancing RNA polymerase activity. The metabolite that triggers the positive control by binding to CAP is 3′,5′-cyclic AMP (cAMP). This compound, formed in energy-deprived cells, acts through

CAP to enhance expression of catabolic enzymes that give rise to metabolic energy. Cyclic AMP is not alone in its ability to exert control over unlinked genes in *E coli*. A number of different genes respond to the nucleotide ppGpp (in which “pp” denotes phosphodiester

and “G” denotes guanine) as a signal of amino acid starvation, and unlinked genes are expressed as part of the SOS response to DNA damage.

GENETIC ENGINEERING

Engineering is the application of science to social needs. Over the past four decades, engineering based on bacterial genetics has transformed biology. Specified DNA fragments can be isolated and amplified, and their genes can be expressed at high levels. The nucleotide specificity required for cleavage by restriction enzymes allows fragments containing genes or parts of genes to be ligated (incorporated) into plasmids (vectors) that can in turn be used to transform bacterial cells. Bacterial colonies or **clones** carrying specified genes can be identified by **hybridization** of DNA or RNA with labele **probes** (similar to that shown in Figure 3-4). Alternatively, protein products encoded by the genes can be recognized either by enzyme activity or by immunologic techniques. Thus, genetic engineering techniques can be used to isolate virtually any gene, so that a biochemically recognizable property can be studied or exploited. Isolated genes can be used for a variety of purposes. **Site-directed mutagenesis** can identify and alter the DNA sequence of a gene. Nucleotide residues essential for gene

function can thus be determined and, if desired, altered. With hybridization techniques, DNA can be used as a probe that recognizes nucleic acids corresponding to the complementary sequence of its own DNA. For example, a latent virus in animal tissue can be detected with a DNA probe even in the absence of overt viral infection. The protein products of isolated viral genes offer great promise as vaccines because they can be prepared without genes that encode the replication of viral nucleic acid. For example, the capsid proteins of human papilloma virus have been cloned and expressed. These are referred to as noninfectious virus-like particles (VLPs) and form the basis for a vaccine against transforming strains of this virus. Moreover, proteins such as insulin that have useful functions can be prepared in large quantities from bacteria that express cloned genes.

**Preparation of DNA Fragments With Restriction Enzymes**

The genetic diversity of bacteria is reflected in their extensive range of **restriction enzymes**, which possess remarkable selectivity that allows them to recognize specific regions of DNA for cleavage. DNA sequences recognized by restriction enzymes are predominantly palindromes (inverted sequence repetitions). A typical sequence palindrome, recognized by the frequently used restriction enzyme *Eco*R1, is GAATTC; the inverted repetition, inherent in the complementarity of the G-C and A-T base pairs, results in the 5′ sequence TTC being reflected as AAG in the 3′ strand.

The length of DNA fragments produced by restriction enzymes varies tremendously because of the individuality of DNA sequences. The average length of the DNA fragment is determined in large part by the number of specific bases recognized by an enzyme. Most restriction enzymes recognize four, six, or eight base sequences; however, other restriction enzymes recognize 10, 11, 12, or 15 base sequences. Recognition of four bases yields fragments with an average length of 250 base pairs and therefore is generally useful for analysis or manipulation of gene fragments. Complete genes are frequently encompassed by restriction enzymes that recognize six bases and produce fragments with an average size of about 4 kbp. Restriction enzymes that recognize eight bases produce fragments with a typical size of 64 kbp and are useful for analysis of large genetic regions. Restriction enzymes that recognize more than 10 bases are useful for construction of a physical map and for molecular typing by pulsed-field gel electrophoresis.

**Physical Separation of Differently Sized DNA Fragments**

Much of the simplicity underlying genetic engineering techniques lies in the fact that **gel electrophoresis** permits DNA fragments to be separated on the basis of size (Figure 7-12): The smaller the fragment, the more rapid the rate of migration. The overall rate of migration and optimal range of size for separation are determined by the chemical nature of the

gel and by the degree of its cross-linking. Highly cross-linked gels optimize the separation of small DNA fragments. The dye **ethidium bromide** forms brightly fluorescent adducts as it binds to DNA, so that small amounts of separated DNA fragments can be visualized on gels (Figure 7-12A). Specific DNA fragments can be recognized by probes containing complementary sequences (Figure 7-12B and C).

**Pulsed-field gel electrophoresis** allows the separation of DNA fragments containing up to 100 kbp that are separated on high-resolution polyacrylamide gels. Characterizations of such large fragments have allowed construction of a physical map for the chromosomes from several bacterial species and have been invaluable in fingerprinting bacterial isolates associated with infectious disease outbreaks.

**Cloning of DNA Restriction Fragments**

Many restriction enzymes cleave asymmetrically and produce DNA fragments with **cohesive (sticky) ends** that may hybridize with one another. This DNA can be used as a donor with plasmid recipients to form genetically engineered recombinant plasmids. For example, cleavage of DNA with *Eco*R1 produces DNA containing the 5′ tail sequence AATT and the complementary 3′ tail sequence TTAA (Figure 7-13). Cleavage of a plasmid with the same restriction enzyme produces a linear fragment with cohesive ends

that are identical to one another. Enzymatic removal of the free phosphate groups from these ends ensures that they will not be ligated to form the original circular plasmid. Ligation in the presence of other DNA fragments containing free phosphate groups produces **recombinant plasmids**, which have DNA fragments as inserts in covalently closed circular DNA. Plasmids must be in a circular form to replicate in a bacterial host.

Recombinant plasmids may be introduced into a bacterial host, frequently *E coli*, by **forced transformation**. Alternatively, **electroporation** is a procedure that introduces DNA

into bacteria using an electrical gradient. Transformed cells may be selected on the basis of one or more drug resistance factors encoded by plasmid genes. The resulting bacterial population contains a **library** of recombinant plasmids carrying various cloned inserted restriction fragments derived from the donor DNA. Hybridization techniques may be used to identify bacterial colonies carrying specific DNA fragments (Figure 7-14), or, if the plasmid expresses the inserted gene, colonies can be screened for the gene product by an antibody specific for the protein produced.

CHARACTERIZATION OF CLONED DNA

**Restriction Mapping**

Manipulation of cloned DNA requires an understanding of its nucleic acid sequence. Preparation of a **restriction map** is the first step in gaining this understanding. A restriction map is constructed similar to a jigsaw puzzle from fragment sizes produced by **single digests**, which are prepared with individual restriction enzymes, and by **double digests**, which are formed with pairs of restriction enzymes. Restriction maps are also the initial step toward DNA sequencing because they identify fragments that will provide **subclones** (increasingly smaller fragments of DNA) that may be subjected to more rigorous analysis, which may involve DNA sequencing. In addition, restriction maps provide a highly specific information base that allows DNA fragments, identified on the basis of size, to be associated with specific gene function.

**Sequencing**

DNA sequencing displays gene structure and enables researchers to deduce the amino acid sequence of gene products. In turn, this information makes it possible to manipulate genes to understand or alter their function. In addition, DNA sequence analysis reveals regulatory regions that control gene expression and genetic “hot spots” particularly susceptible to evolutionary relationships that provide a framework for unambiguous classification of bacterial species. Such comparisons may facilitate identification of conserved regions that may prove particularly useful as specific hybridization probes to detect the organisms or viruses in clinical samples. The original method of DNA sequence determination used the **Maxam-Gilbert technique**, which relies on the relative chemical liability of different nucleotide bonds. The field now has largely moved to the **Sanger (dideoxy termination)** **method**, which interrupts elongation of DNA sequences by incorporating dideoxynucleotides into the sequences. Both techniques produce a nested set of oligonucleotides starting from a single origin and entail separation on a sequencing gel of DNA strands that differ by the increment of a single nucleotide. A polyacrylamide sequencing gel separates strands thatdiffer in length from one to several hundred nucleotides and reveals DNA sequences of varying lengths.

Four parallel lanes on the same gel reveal the relative length of strands undergoing dideoxy termination at adenine, cytidine, guanidine, and thymidine. Comparison of four lanes containing reaction mixes that differ solely in the method of chain terminating dideoxy nucleotide makes it possible to determine DNA sequence by the Sanger method (Figure 7-15). The relative simplicity of the Sanger method has led to its more general use, but the Maxam Gilbert technique is widely used because it can expose regions of DNA that are protected by specific binding proteins against chemical modification. DNA sequencing is greatly facilitated by genetic manipulation of *E coli* bacteriophage M13, which contains ssDNA. The replicative form of the phage DNA is a covalently closed

circle of dsDNA that has been engineered so that it contains a multiple cloning site that permits integration of specific DNA fragments that have been previously identified by restriction mapping. Bacteria infected with the replicative form secrete modified phages containing, within their protein coat, ssDNA **template** for elongation reactions. The origin for elongation is determined by a DNA **primer**, which can be synthesized by highly automated machines for **chemical oligonucleotide** **synthesis**. Such machines, which can produce DNA strands containing 75 or more oligonucleotides in a predetermined sequence, are essential for sequencing and for the modification of DNA by site-directed mutagenesis. Chemically synthesized oligonucleotides can serve as primers for polymerase chain reaction (**PCR**), a procedure that allows amplification and sequencing of DNA lying between the primers. Thus, in many instances, DNA need not be cloned in order to be sequenced or to be made available for engineering.

The study of biology has been revolutionized by the development of technology that allows sequencing and analysis of entire genomes, ranging from viruses to unicellular prokaryotic and eukaryotic microorganisms to humans. This has been facilitated by use of the procedure known as **shotgunning**. In this procedure, the DNA is broken into random smaller fragments to create a fragment library. These unordered fragments are sequenced by automated DNA sequencers and reassembled in the correct order using powerful computer software. A sufficient number of fragments are sequenced to ensure adequate coverage of the genome so that when they are assembled, most of the genome is represented without leaving too many gaps. (To achieve this, the entire genome is usually covered five- to eight-fold, leaving about 0.1% of the total DNA unsequenced.) After the random fragments have been assembled by areas of overlapping sequence, any remaining gaps can be identified and closed. Advanced data processing permits annotation of the sequence data in which putative coding regions, operons, and regulatory sequences are identified. Already, the genomes of a number of important microorganisms have been sequenced. The continued analysis of sequence data from important human pathogens, combined with studies on molecular pathogenesis, will facilitate our understanding of how these organisms cause disease and, ultimately, will lead to better prophylactic and therapeutic strategies.

SITE-DIRECTED MUTAGENESIS

Chemical synthesis of oligonucleotides enables researchers to perform controlled introduction of base substitution into a DNA sequence. The specified substitution may be used to explore the effect of a predesigned mutation on gene expression, to examine the contribution of a substituted amino acid to protein function, or—on the basis of prior information about residues essential for function—to inactivate a gene. Single-stranded oligonucleotides containing the specified mutation are synthesized chemically and hybridized to single-stranded phage DNA, which carries the wild-type sequence as an insert (Figure 7-16). The resulting partially dsDNA is enzymatically converted to the fully double- stranded replicative form. This DNA, which contains the wild-type sequence on one strand and the mutant sequence on the other, is used to infect a bacterial host by transformation. Replication results in segregation of wild-type and mutant DNA, and the double-stranded mutant gene can be challenge would ideally be to maintain biologically beneficial, engineered organisms in the environment rather than to eliminate them. However, this is not without social consequence. Among the examples of engineered organisms are *Pseudomonas* strains that produce a protein favoring formation of ice crystals. The value of these wild-type organisms is appreciated by ski slope owners, who have deliberately introduced the bacteria into the environment without arousing any public concern.

 An unfortunate side effect of the introduction of these organisms is that the ice crystals they promote can injure sensitive crops such as lettuce during seasons in which light frost is likely. Mutant bacteria that do not form ice crystals were designed by microbiologists who hoped that the mutant organisms might protect lettuce crops by temporarily occupying the niche normally inhabited by the ice-forming strains; however, attempts to use the mutant organisms in field studies were met with substantial protest, and studies were conducted only after lengthy and expensive legal delays. The legal precedents that have emerged from this and more recent related applications will establish guidelines for the progressive and beneficial use of genetic engineering techniques and facilitate determination of situations in which extreme caution is justified.