PART I

FOUNDATIONS OF BIOREGENERATIVE ENGINEERING

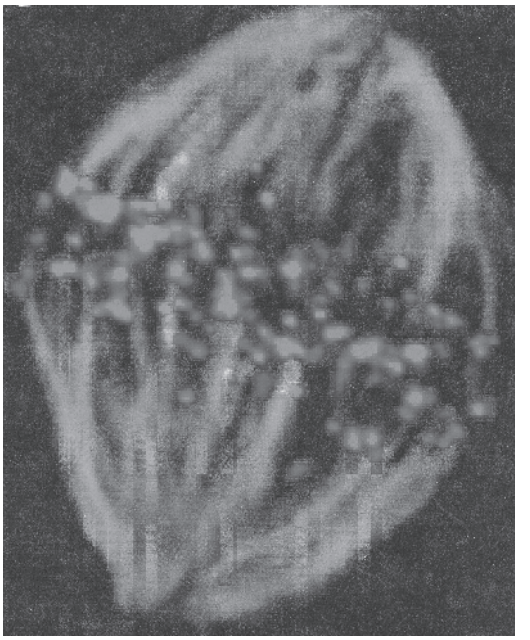
COPYRIGHTED MATERIAL

SECTION 1

MOLECULAR BASIS FOR BIOREGENERATIVE ENGINEERING

1

STRUCTURE AND FUNCTION OF MACROMOLECULES



Organization of chromosomes and microtubules in an epithelial cell in the metaphase. Green and red: immunochemically labeled tubulin and kinetochores, respectively. (Reprinted by permission from Kapoor TM et al: Chromosomes can congress to the metaphase plate before biorientation, *Science* 311:388–91, 2006. Copyright 2006, AAAS.) See color insert.

Bioregenerative Engineering: Principles and Applications, by Shu Q. Liu
Copyright © 2007 John Wiley & Sons, Inc.

A living organism is composed of several basic elements: water, electrolytes, nucleotides, amino acids, sugars, and lipids. Water is the most abundant substance in a living system, occupying about 70–85% of the total volume in most cells. All living organisms were originally developed in water, and all biochemical and enzymatic reactions in a cell take place in an aqueous environment. Thus, water is the most important element in a living organism.

A cell consists of a number of electrolytes, including sodium, potassium, calcium, magnesium, chloride, phosphate, sulfate, and bicarbonate. These electrolytes participate in fundamental processes, such as the establishment and maintenance of cell membrane and action potentials, regulation of biochemical and enzymatic reactions, control of muscular contraction and relaxation, formation of mechanical supporting and protection systems, and maintenance of the internal environment. The functions of these electrolytes will be discussed throughout the book where applicable.

Other elements, including nucleotides, sugars, amino acids, and lipids, participate in the formation of macromolecules, including deoxyribose and ribose nucleic acids (nucleotides and sugars), proteins (amino acids), phospholipids (lipids), and polysaccharides (sugars). These macromolecules are essential to the formation, survival, function, and regeneration of living organisms. This chapter focuses on the structure and function of these macromolecules.

DEOXYRIBONUCLEIC ACIDS (DNA)

DNA is the molecule for the transmission, processing, and storage of hereditary information. A DNA molecule is capable of replicating itself, a fundamental mechanism for the transmission of hereditary information from the mother to the daughter generation. A DNA molecule can be transcribed into various types of RNA, including messenger RNA (mRNA), ribosomal RNA (rRNA), and transfer RNA (tRNA). These RNA molecules participate in the translation of proteins. The translated proteins are transported to various compartments of a cell, serving as not only structural constituents, which provide the cell with shape and strength, but also enzymes and signaling molecules, which regulate cellular activities and functions.

Composition and Structure of DNA [1.1]

A DNA molecule is constituted by joining together a large number of nucleotides. A *nucleotide* is composed of three elements, including a base, a β-D-2-deoxyribose, and a phosphate group. There exist four types of base, which are nitrogen-containing ring compounds, including two pyrimidines—cytosine and thymine, denoted as C and T, respectively—and two purines—adenine and guanine, denoted as A and G, respectively (Fig. 1.1). A base is capable of forming a complex with a deoxyribose (Fig. 1.2), giving rise to a molecule known as *nucleoside*. Collectively, there are four types of nucleoside based on the four bases, including cytidine, thymidine, adenosine, and guanosine. With the addition of 1, 2, or 3 phosphate groups, a nucleoside is converted to a nucleotide, known as *nucleoside monophosphate*, *diphosphate*, or *triphosphate*, respectively (Fig. 1.3). The nomenclature for various individual nucleosides and nucleotides are listed in Table 1.1.

A complete DNA molecule is a double-stranded helical polymer of nucleotides. Each stand is composed of a pentose–phosphate backbone and bases positioned on the side of

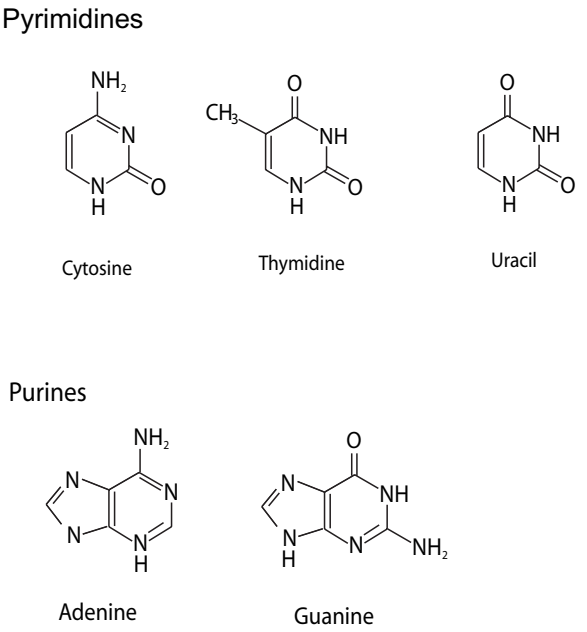


Figure 1.1. Chemical structure of pyrimidines (cytosine, thymine, and uracil) and purines (adenine and guanine) that constitute DNA and RNA.

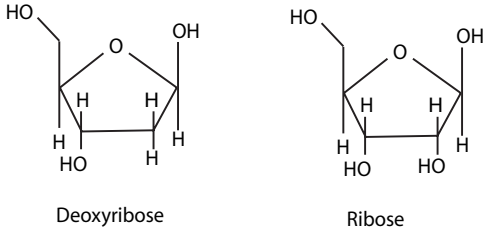
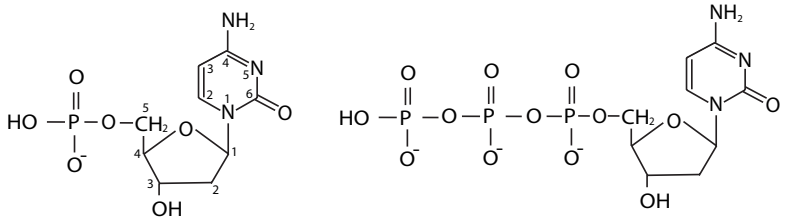


Figure 1.2. Chemical structure of deoxyribose and ribose.

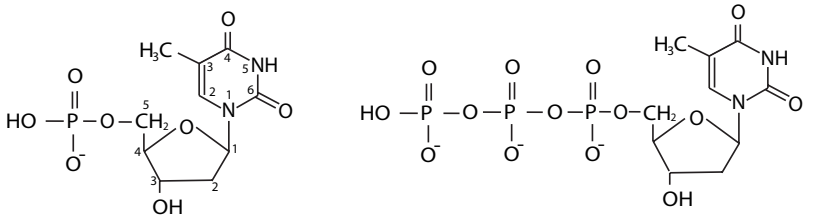
TABLE 1.1. Nomenclature of Nucleosides and Nucleotides for DNA

Types of Base	Cytosine (C)	Thymine (T)	Adenine (A)	Guanine (G)
Nucleosides	Deoxycytidine (dC)	Deoxythymidine (dT)	Deoxyadenosine (dA)	Deoxyguanosine (dG)
Nucleotides	Deoxycytidine 5′ monophosphate (dCMP)	Deoxythymidine 5′ monophosphate (dTMP)	Deoxyadenosine 5′ monophosphate (dAMP)	Deoxyguanosine 5′ monophosphate (dGMP)
	Deoxycytidine 5′ diphosphate (dCDP)	Deoxythymidine 5′ diphosphate (dTDP)	Deoxyadenosine 5′ diphosphate (dADP)	Deoxyguanosine 5′ diphosphate (dGDP)
	Deoxycytidine 5′ triphosphate (dCTP)	Deoxythymidine 5′ triphosphate (dTTP)	Deoxyadenosine 5′ triphosphate (dATP)	Deoxyguanosine 5′ triphosphate (dGTP)

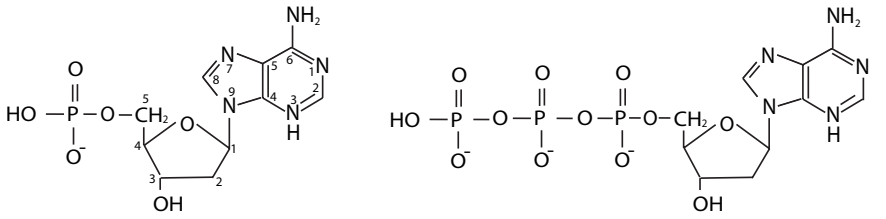
Deoxycytidine monophosphate and triphosphate



Deoxythymidine monophosphate and triphosphate



Deoxyadenosine monophosphate and triphosphate



Deoxyguanosine monophosphate and triphosphate

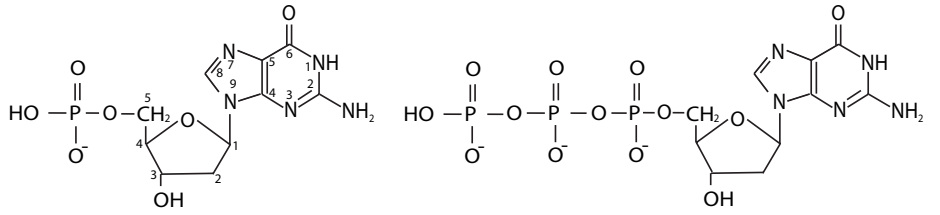


Figure 1.3. Chemical structure of the nucleotides that constitute DNA.

the backbone. The backbone is formed by joining the pentose molecules via covalent phosphodiester bonds. A phosphate group joins one pentose at the 5′ carbon and to the next pentose at the 3′ carbon (Fig. 1.4). The four bases are attached to the pentose–phosphate backbone and aligned on the same side of each DNA chain. The two strands of DNA are attached to each other via hydrogen bonds between the base pairs on the basis of the complementary principle, specifica, A with T and C with G. Note that a double-ringed purine base (A or G) is always paired with a single-ringed pyrimidine (T or C) (Fig. 1.5). A hydrogen bond is established between a positively charged H and a negatively

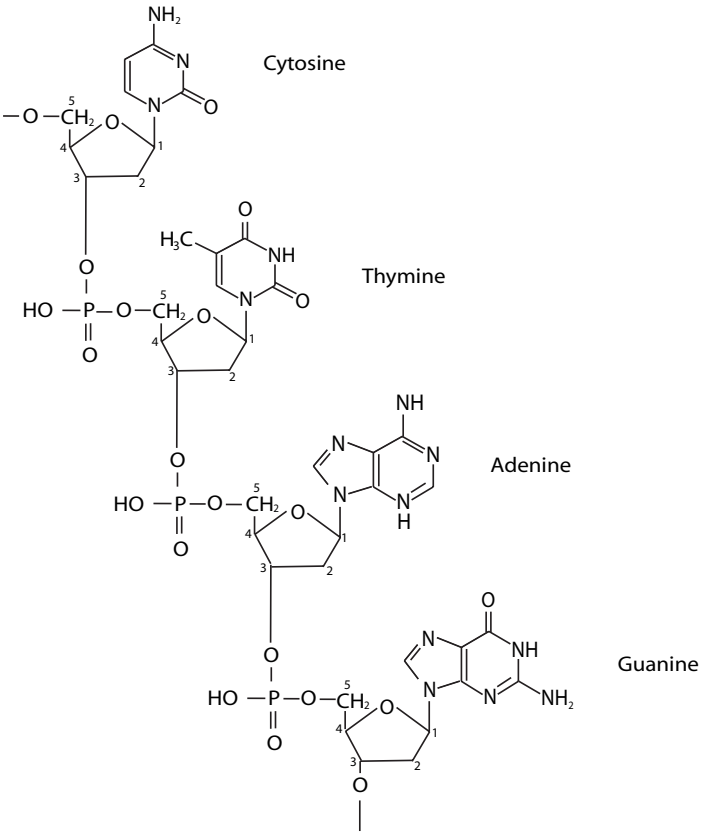


Figure 1.4. Formation of phosphodiester bonds between pentose molecules, constituting the backbone of DNA (based on bibliography 1.1).

charged *acceptor*, such as an O and N. A hydrogen bond is relatively weak, about 3% of the strength of a covalent bond. However, an array of hydrogen bonds, as found in the double-stranded helical DNA molecule, can be very strong if all hydrogen bonds are aligned on the same side of a DNA strand. For a double-stranded DNA molecule, the two sugar-phosphate backbones run in opposite directions, defined as an *antiparallel* arrangement. One strand is defined as the $5' \rightarrow 3'$ strand; the other, the $3' \rightarrow 5'$ strand.

The four bases, A, G, C, and T, are organized into a series of a large number of distinct sequences, each forming an independent functional unit known as the *gene*. In eukaryotic cells, genes are composed of coding sequences called exons and noncoding sequences called *introns*. The *exons* contain codons for specific proteins with each codon composed of three nucleotides that specify an amino acid. In contrast, the *introns* are sequences for the regulation of gene transcription. The intron sequences do not contain protein-coding sequences. The different sequences or structures of DNA are defined as *genotypes*, which determine the chemical, physical, and functional characteristics, or *phenotypes*, of various living organisms.

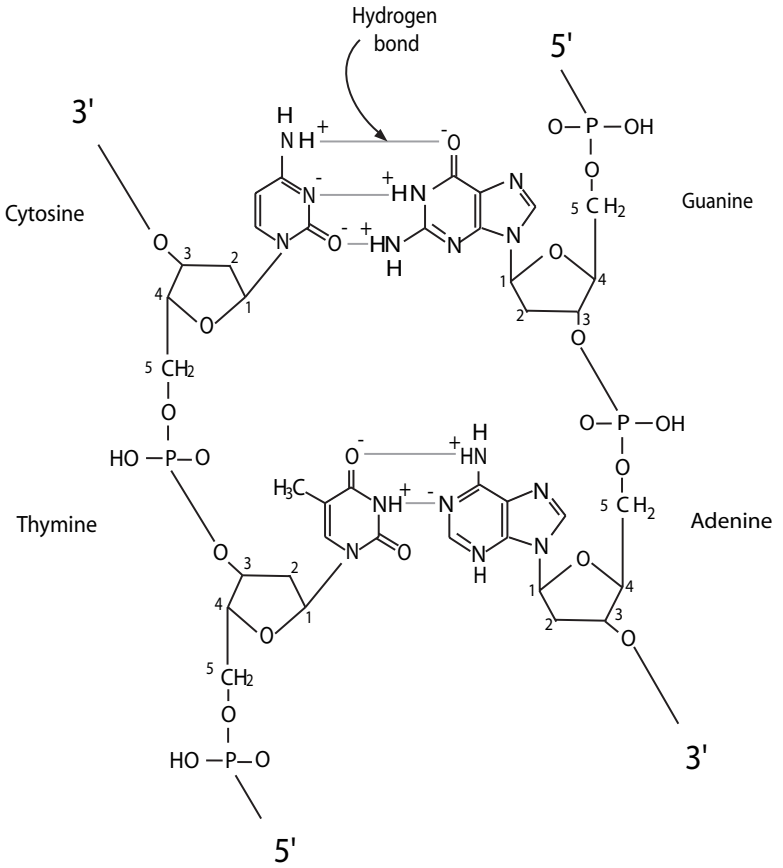


Figure 1.5. Formation of hydrogen bonds that link two single DNA strands into a double DNA helical structure (based on bibliography 1.1).

A gene carries genetic information in a particular form that can be stored, processed, copied, transcribed to generate messenger RNA (mRNA), and transmitted from the mother cells to the progeny. The double-stranded helical structure is ideal for such purposes. Since both nucleotide strands of a DNA molecule are complementary to each other, both strands carry identical genetic information. Such an organization ensures precise information transfer during DNA replication. When a DNA molecule is ready for replication, the two strands separate. Each strand can serve as a template for the synthesis of a new strand. A newly synthesized strand is identical to the complementary counterpart of the template strand. Thus, each daughter DNA is a pair of an original template and a newly synthesized strand. The daughter DNA molecules can in turn serve as templates for further DNA replication. In such a way, the genetic information can be stored, copied, and transmitted from generation to generation endlessly.

Organization of Chromosomes [1.2]

Each DNA molecule is packaged in a nuclear structure, known as *chromosome*. In humans, each cell contains 46 chromosomes, including a pair of sex chromosomes (XX

for female or XY for male), which are arranged in 23 pairs. One of each paired chromosomes is derived from each parent. Thus, each individual offspring inherits two copies of the same chromosome. The total chromosomes contain about 6.6×10^9 DNA base pairs. The spacing distance is about 3.4 Å per base pair. Each human cell contains 46 DNA molecules (one DNA molecule for each chromosome) with a total length of about 2 m. Given the small size of a cell nucleus (about 5–10 μm), the DNA molecules must be folded to fit in the nucleus. The folding and packaging of DNA are accomplished with the assistance of packaging proteins (primarily histones). DNA and the packaging proteins together form a structure called *chromatin*, which is organized to form a chromosome. Chromatin exists in two states: heterochromatin and euchromatin. *Heterochromatin* is a highly condensed form of chromatin and is not involved in RNA transcription, whereas *euchromatin* is less condensed and is ready for RNA transcription.

There are several levels of DNA packaging, including (1) formation of nucleosomes, which are DNA coils around protein cores; (2) folding of the nucleosome DNA into a fiber structure (~30 nm in diameter); (3) additional folding of the fiber DNA into thicker bundles (100–300 nm); and (4) formation of loop domains, each containing 15–100 kilo–base pairs (kb) (i.e., 15,000–100,000 base pairs). At the first level, a string-like DNA molecule is coiled around a series of core complexes of proteins known as *histones* to form nucleosomes. Each nucleosome contains about 170 base pairs (bp) of DNA. Uncoiled DNA fragment between two adjacent nucleosomes is referred to as *linker DNA*, which is about 30 bp in length. There are several types of histones (H), including H2A, H2B, H3, and H4, which constitute the core complex of nucleosomes. These histones contain positive charges, which neutralize the negative charges of the DNA phosphate groups, thus stabilizing the DNA–histone complex structures. Each human cell nucleus contains about 3×10^7 nucleosomes.

At higher levels, a string of DNA with coiled nucleosomes is folded and condensed into chromatin fibers about 30 nm in diameter. These fibers are further folded into thicker chromatin bundles. The folded chromatins form large loops, each containing thousands to millions of base pairs of DNA. The chromatin loops are organized into a chromosomal structure by nuclear matrix proteins, also known as *nonhistone chromosomal proteins*, which form chromosomal scaffolds. A well-characterized complex of nuclear matrix proteins is the condensin, which can be phosphorylated by the cyclin-dependent kinase-1/cyclin B complex and controls the final level of chromosomal condensation. With various levels of folding and confinement by nuclear matrix proteins, a DNA molecule is greatly reduced in length and well organized, allowing the fit of the molecule into a chromosome.

Functional Units of DNA [1.3]

All DNA molecules in the 23 pairs of chromosomes constitute the *genome*. Each DNA molecule within a chromosome is composed of several types of functional units, including the *genes*, a *centromere*, two *telomeres*, and numbers of *replication origins* (approximately 1 per 100,000 bp). A gene is a functional unit for the process and transmission of genetic information and for coding a specific protein. In total, there are more than 50,000 genes in the human genome. Each offspring individual inherits two copies of the same gene, one from the mother and the other from the father. Several terms, such as genetic locus

and alleles, are often used in genetic analysis. Genetic *locus* is defined as the chromosomal location of the two copies of each gene. *Alleles* are the forms of a gene at a genetic locus. Some genes exist in two or more alternate forms. Each gene is located at a specified locus of a chromosome. When the two copies of the gene at the same locus are identical, the individual who carries the gene is defined as a *homozygote*. When the two gene copies are different at the same locus, the individual is said a *heterozygote*. In humans, about 80% genetic loci contain homozygous genes and about 20% loci contain heterozygous genes.

Each gene encodes specific information for the transcription of an mRNA, which can be translated to a specific protein. The processes of mRNA transcription and protein translation are referred to as *gene expression*. At a given time, only a fraction of genes is expressed in the genome. The regulation of gene expression is a complicated process, involving a variety of signaling pathways and regulatory factors. In addition to the genes, there exist a large number of DNA sequences, which contain no information for protein coding in each DNA molecule. These noncoding sequences may participate in the regulation of gene stability and function, although the exact function remains poorly understood.

The *centromere* is a chromosomal structure that mediates the separation of a chromosome during mitosis and meiosis. In each DNA molecule or chromosome, the centromere is located at the point where a chromosome is attached to the microtubule-based spindle (Fig. 1.6). During mitosis, the centromeric regions of the two sister chromatids separate and are pulled by microtubules toward opposite poles. Each centromere region is composed of heterochromatin, which does not contain coding genes. A centromere contains

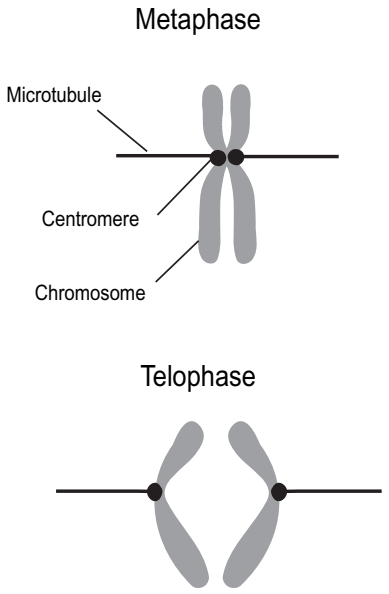


Figure 1.6. Location of centromere and separation of chromatids during mitosis (from metaphase to telophase). Based on bibliography 1.3.

a substructure called *kinetochore*, which binds microtubules and directs the movement of chromosome during mitosis. The DNA sequence of a centromere forms complexes with proteins, known as *centromere proteins*, including centromere protein (CENP)A, B, and C. These proteins regulate the function of the centromere. Centromere protein A [which has a molecular weight of 17 kilodaltons (17 kDa)] possibly mediates the formation of kinetochore and assists the attachment of centromere protein C to the kinetochore. Centromere protein B (65 kDa) may regulate the formation and organization of the centromeric heterochromatin. Centromere protein C (107 kDa) plays a critical role in the assembly of the kinetochore. Centromere proteins A and C are required for mitosis. The blockade of centromere protein C with neutralizing antibody, which is injected into to the cell nuclei, induces alterations in the kinetochore and cell arrest in mitosis.

Telomeres are DNA sequences found at the two ends of a DNA molecule. A telomere possesses several basic functions: (1) controlling the integrity of the DNA ends, (2) guiding the DNA replication machinery during DNA replication, and (3) providing signals that allow the DNA replication machinery to recognize the DNA ends without joining two DNA molecules mistakenly. In mammals, telomeres contain numerous repeats of the sequence 5′-TTAGGG-3′. Such an organization results in a unique structure at the two ends of each DNA molecule, rich in G in one strand while rich in C in the other strand. In mammalian cells, telomeres form complexes with nuclear proteins, such as telomere repeat factors (TRFs) 1 and 2, which play a critical role in regulating the stability and function of telomeres. Alterations in the binding of telomere repeat factor 2 to the telomeres cause an increase in the possibility of chromosome-to-chromosome fusion.

The origins of replication are sites along a DNA molecule where DNA replication is initiated. The DNA sequences of the replication origins have been characterized in lower levels of organisms, such as bacteria, yeast, and viruses. A replication origin sequence is about 300 bp in length. Such a structure allows the binding of initiator proteins and helicase, initiating the formation of a replication bubble and two replication forks. At the replication fork, the synthesis and annealing of a RNA primer activate a DNA polymerase, initiating DNA replication.

DNA Replication [1.4]

DNA replication is a process that synthesizes a copy of the entire genome based on the mother template during cell mitosis, ensuring the transmission of an identical genome from the mother to the daughter cell during mitosis. DNA replication is accomplished through several steps, including replication initiation, DNA extension, and sequence proofreading. Since *Escherichia coli* have been used for investigating the mechanisms of DNA replication, the process of DNA replication is described here on the basis of an *E. coli* model. The mechanisms of *E. coli* DNA replication are similar to those in eukaryotic cells.

Initiation. The replication of *E. coli* DNA is initiated at a specific site known as the *replication origin*, which is composed of several elements including a consensus 13-bp sequence and several binding sites for regulatory proteins including dnaA protein and helicase. On the binding of dnaA protein to the replication origin, a helicase binds to the replication origin and unwinds the double-stranded DNA, resulting in the formation of regionally separated single DNA strands with free bases. The two separation points flanking the replication origin are known as *replication forks.* With continuous separation of

the DNA double strands, the replication forks are dynamically moving away from the replication origin. The separation of the replication origin and the formation of the replication forks prepare the synthesis of DNA.

The initiation of DNA synthesis requires the presence of several components in *E. coli*: RNA primers (~30 bps in length) and DNA polymerases. A RNA primer is specific to the sequence of a replication origin and is synthesized by a RNA polymerase or primase. On the separation of a replication origin, a primase forms a complex with the template DNA as well as with several regulatory proteins, including dnaB, dnaT, priA, priB, and priC, leading to the synthesis of a specific RNA primer. The synthesis of a RNA primer induces the binding of a DNA polymerase. The synthesized RNA primer anneals to the replication origin, initiating DNA synthesis.

DNA Extension. DNA synthesis is a process by which the annealed RNA primer is elongated on the template DNA strand according to the base-pairing principle. Such a process requires the presence of several types of DNA polymerase and a DNA ligase. A bound, activated DNA polymerase is capable of selecting deoxynucleotides complementary to that of the template DNA and adding these deoxynucleotides to the RNA primer one at a time, resulting in the elongation of the daughter DNA molecule. The elongation of DNA occurs always in the $5' \rightarrow 3'$ direction. At each replication origin, there goes bidirectional DNA synthesis. At each replication fork, DNA synthesis is conducted simultaneously along the two separated DNA strands in opposite directions, because of different polarities of the two DNA strands (Fig. 1.7). As one DNA template directs DNA elongation toward the dynamically moving replication fork, a process defined as the *leading strand DNA elongation*, the other DNA template directs DNA synthesis in a direction away from the fork, which is defined as *lagging strand elongation*. The leading strand DNA elongation is continuous, whereas the lagging strand elongation is discontinuous; thus, DNA is synthesized segment by segment, due to the constraints of the fork moving direction and the DNA synthesis direction (Fig. 1.7).

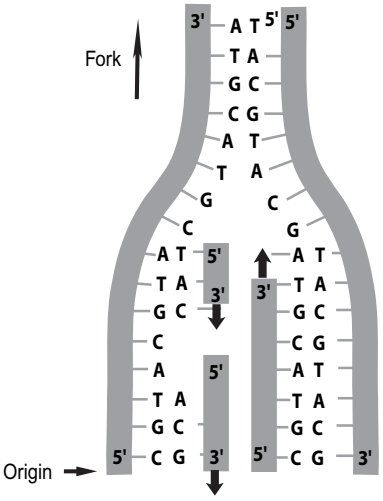


Figure 1.7. DNA replication along the two template DNA strands (based on bibliography 1.4).

Three types of DNA polymerase are found in *E. coli*: polymerases I, II, and III. DNA polymerase I possesses three functions: (1) catalyzing DNA extension in the 5′–3′ direction following a RNA primer on a DNA template, (2) serving as an exonuclease that eliminates mismatched deoxynucleotides and removes RNA primers on the lagging template following DNA extension, and (3) degrading double-stranded DNA in the $5' \rightarrow 3'$ direction. DNA polymerase II is involved in the repair of damaged DNA. DNA polymerase III possesses functions similar to those of DNA polymerase I, but with different target DNA strands. DNA polymerase III can act on both DNA strands, whereas DNA polymerase I works primarily on the lagging strand, completing DNA replication based on the template segments that have not been duplicated. DNA polymerase I also removes RNA primers following the completion of DNA extension. In addition, a DNA ligase is needed to join all newly synthesized DNA segments on the lagging template by catalyzing the formation of a phosphodiester bond between the 5′ phosphate of a nucleotide and the OH group of an adjacent nucleotide.

Proofreading. During DNA synthesis, incorrect nucleotides could be mistakenly inserted into the daughter DNA. These incorrect nucleotides are removed by an enzymatic process known as *DNA proofreading* or *DNA proofediting*. The enzymes that catalyze DNA extension, including DNA polymerases I and III, can serve as nucleases, which are responsible for the removal of incorrect nucleotides. These nucleases can recognize and excise mismatched bases by hydrolysis at the 5′ end of the mismatched nucleotide. Since these enzymes act in the $3' \rightarrow 5'$ direction, the excision of a nucleotide at the 5′ end will leave a free 3′ OH group in the preceding base, allowing the insertion of a correct nucleotide.

DNA Replication in Prokaryotic and Eukaryotic Cells. The processes described above are observed in prokaryotic *E. coli*. DNA replication is similar between prokaryotic and eukaryotic cells in many aspects, but there are differences. First, the time required for DNA replication differs between the two types of cell. A DNA replication–cell division cycle for *E. coli* is about 40 min, whereas that for eukaryotic cells is much longer. For instance, the cell division cycle is about 1.5 h in yeast and 24 h in mammalian cells. Second, eukaryotic cells contain usually multiple chromosomes. These cells must conduct DNA replication simultaneously in all chromosomes in a coordinated manner. An effective approach is to initiate DNA replication at multiple replication origins. Such a mechanism has been demonstrated by the existence of multiple DNA extension locations after a "pulse" exposure to radioactively labeled thymidine. In yeast, there exist about 400 replication origins in the 17 chromosomes. Third, the types of DNA polymerase are different between prokaryotic and eukaryotic cells. In eukaryotic cells, four types of DNA polymerases have been found: α, β, γ, and δ. The eukaryotic DNA polymerases α and δ are similar to prokaryotic polymerase II. DNA polymerase β may be responsible in DNA repair. The γ type may be involved in the replication of mitochondrial DNA.

RIBONUCLEIC ACID (RNA)

RNA Composition and Structure [1.5]

Ribonucleic acid (RNA) is a molecule that transmits genetic information from DNA to proteins. Similar to a DNA molecule, RNA is composed of linearly joined nucleotides,

TABLE 1.2. Nomenclature of Nucleosides and Nucleotides for RNA

Types of Base	Cytosine (C)	Uracil (U)	Adenine (A)	Guanine (G)
Nucleosides	Cytidine	Uridine	Adenosine	Guanosine
Nucleotides	Cytidine monophosphate (CMP)	Uridine monophosphate (UMP)	Adenosine monophosphate (AMP)	Guanosine monophosphate (GMP)
	Cytidine diphosphate (CDP)	Uridine diphosphate (UDP)	Adenosine diphosphate (ADP)	Guanosine diphosphate (GDP)
	Cytidine triphosphate (CTP)	Uridine triphosphate (UTP)	Adenosine triphosphate (ATP)	Guanosine triphosphate (GTP)

each including a nitrogenous base, a pentose, and a phosphate group. Unlike DNA, the pentose for RNA is β-D-ribose instead of β-D-2-deoxyribose (Fig. 1.2) and the four bases are cytosine, uracil, adenine, and guanine with uracil in place of thymine (Fig. 1.1). Furthermore, RNA is a single-stranded molecule and is relatively short compared to DNA molecules. The nomenclatures for various RNA nucleosides and nucleotides are listed in Table 1.2.

RNA is synthesized via DNA transcription, a process similar to DNA replication in certain aspects. To transcribe a RNA molecule, a DNA molecule opens locally into single-stranded forms. One of the two strands serves as a template for RNA synthesis according to the base-pairing principle. That is, for a deoxynucleotide A on the template, a ribonucleotide U (but not T) is added; and for a C on the template, a G is added. The RNA molecule is elongated by adding ribonucleotides one by one. These ribonucleotides are joined together via covalent bonds. The size and type of RNA transcribed from a region of DNA is controlled by proteins called *gene regulatory factors*. These proteins bind to specific sites of DNA and regulate the process of RNA transcription. For any given time, some genes are activated for RNA transcription, while others are not. The selection of gene activation is controlled by gene regulatory proteins, which are activated by upstream signaling pathways. The process of RNA synthesis stops when a DNA stop codon appears.

There exist three types of RNA molecules: messenger RNA (mRNA), transfer RNA (tRNA), and ribosome RNA (rRNA). An mRNA molecule is a copy of a specific gene that carries the information or codon for a protein. Thus, mRNA directs the translation or synthesis of a specific protein. An rRNA molecule serves as a machine for protein synthesis or translation with a specific mRNA as a template. A tRNA molecule is responsible for the transfer of amino acids to an rRNA that synthesizes proteins based on an mRNA transcript. The three types of RNA work coordinately in protein translation.

RNA Transcription [1.5]

RNA transcription is a process that transfers genetic information from a gene to an mRNA, which is then translated to a protein. This sequence of processes is also known as *gene expression*. RNA transcription is similar to DNA replication except that (1) a different enzyme, the RNA polymerase, is used for RNA synthesis; (2) a uridine is used instead of thymidine; and (3) a single RNA strand is synthesized. Studies based on the *E. coli*

system have established that RNA transcription is accomplished through several steps, including initiation, elongation, and termination. Similar processes are found in eukaryotic RNA transcription with several exceptions (see below). Here, the E. coli system is used to describe the mechanisms of RNA transcription.

Initiation. A RNA molecule is transcribed from a specific gene and encodes information for the sequence of amino acids of a protein. RNA transcription is initiated at a DNA site next to a special DNA sequence known as a *promoter*, which is found in each gene and activated in response to the binding of transcription factors to specific DNA sequences known as enhancers. There are two types of promoter in the *E. coli* genome with consensus sequences TTGACAT and TATAAT, which are located at the −35 and −10 sites relative to the first base pair (+1) to be transcribed. On the binding of transcription factors, a RNA polymerase binds to a gene at the promoter region and unwinds the double-stranded DNA at the −10 site, forming an open promoter complex or "transcription bubble" (Fig. 1.8). Such a process initiates RNA transcription.

Elongation. The elongation of RNA is catalyzed by RNA polymerase, a complex enzyme that contains four subunits: α, β, β′, and σ subunits. The subunits can be separated from each another and bound to the other units. A complete complex of RNA polymerase subunits is required for the initiation of RNA transcription. RNA polymerase can recruit nucleotides and synthesize RNA based on a template gene according to the base-pairing principle. The enzyme can catalyze the addition of a nucleoside monophosphate to an existing RNA fragment while a diphosphate group is released from a recruited nucleoside triphosphate. The release of the diphosphate provides energy for the synthetic process. The "transcription bubble" continues to extend when RNA is synthesized (Fig. 1.8). RNA synthesis occurs in the $5' \rightarrow 3'$ direction.

Termination. RNA synthesis is terminated when the RNA polymerase recognizes a termination sequence on the DNA template strand. This sequence contains a GC-rich segment followed by a stretch of six or more adenosine nucleotides. The GC-rich segment of the synthesized RNA can fold regionally, forming a short, hairpin-shaped, double-stranded RNA segment. This segment is followed by U-rich segment. These structures serve as a termination signal. The RNA polymerase is then released and RNA synthesis terminated when the enzyme recognizes these structures.

RNA Transcription and Processing in Eukaryotes. While RNA transcription in eukaryotes is similar to that in prokaryotes in principle, there are several differences. Unlike prokaryotes, eukaryotes express at least three RNA polymerases, including polymerase I,



Figure 1.8. Schematic representation of RNA transcription (based on bibliography 1.5).

II, and III. These polymerases possess distinct functions. Polymerase I is for the synthesis of rRNA, polymerase II is for mRNA, whereas polymerase III is for tRNA.

In addition, transcribed RNAs in the nucleus are processed in eukaryotes before being delivered to the cytoplasm for protein translation. Immediately after the transcription of an mRNA molecule, an enzyme called *guanyltransferase* is activated, adding a 7-methylguanosine residue to the 5′ end of the mRNA transcript. A sequence at the 3′ end of the transcript, usually AAUAAA, activates an endonuclease, which cleaves the mRNA molecule (~20 bp) at the 3′ end. This step is followed by the activation of a poly(A) polymerase, inducing the addition of a poly(A) tail of 150–200 adenosine residues to the 3′ end of the mRNA. These processes generate a complete primary structure of an mRNA molecule.

Unlike prokaryotes, eukaryotes contain primary mRNAs that are composed of two types of sequence known as *exons* and *introns*. The exons contain protein coding regions and are used for protein translation, whereas the introns are sequences that interrupt the exons and do not contain protein coding information. An mRNA transcript must be spliced to remove the introns before protein translation occurs. A mature mRNA molecule is produced after the removal of the introns and is ready for protein translation. In eukaryotes, mRNA splicing occurs in a structure called *splicesome*, which is composed of splicing enzymes and associated factors. In addition to mRNAs, tRNAs, and rRNAs are also processed by splicing to remove introns.

PROTEINS

Protein Composition and Structure [1.6]

Proteins are molecules constituted with amino acids and are major components of living cells, constituting about 50% of the cell dry weight. There are two types of proteins in the cell: structural and regulatory proteins. Structural proteins participate in cell construction, giving a cell the shape, strength, and elasticity, whereas regulatory proteins control biological processes, such as cell-to-cell and cell-to-matrix communication, intracellular enzyme activation, signal transduction, control of gene expression, transport of necessary compounds for cell metabolism, cell division, cell differentiation, cell migration, and cell apoptosis. It is important to note that structural proteins are also involved in the regulation of cellular activities, which become clear in the following examples. Structural proteins include, for example, actin filaments, microtubules, and intermediate filaments, which form an integrated structure called *cytoskeleton*. While these proteins provide the cell with shape and strength, they play a critical role in the regulation of cell mitosis, migration, and adhesion. Regulatory proteins are found in the cell membrane, cytoplasm, and nucleus. Cell membrane proteins, such as growth factor receptors, adhesion molecules, and integrins, are responsible for cell-to-cell and cell-to-matrix interactions. Cytoplasmic proteins are mostly enzymes. Nuclear proteins are involved in the regulation of chromosomal organization and gene expression.

A protein consists of one or more peptides, which are constituted with 20 types of amino acid with distinct structure and chemical properties (Table 1.3). The combination of these amino acids gives rise to a variety of different peptides. The sequence of amino acids is specified by a corresponding gene. The length of a peptide varies widely, ranging from several amino acid residues, such as oxytocin, to about 25,000 residues, including titin. Most peptides are composed of 100–1000 amino acids.

TABLE 1.3. Amino Acids Found in Humans and Animals*

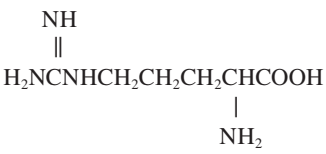
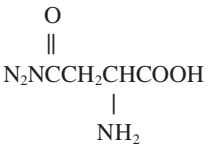
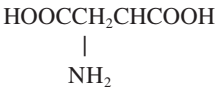
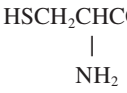
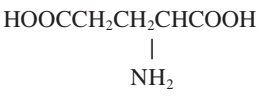
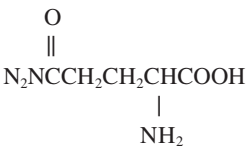
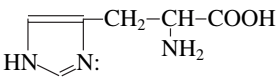
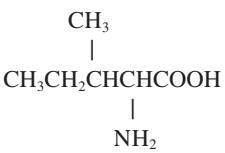
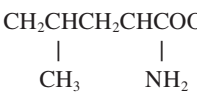
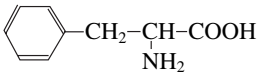
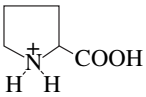
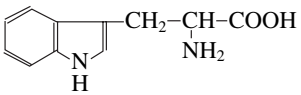
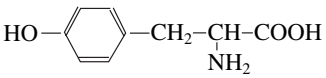
Amino Acids	Chemical Forms	Three-Letter Abbreviation	One-Letter Abbreviation
Alanine		Ala	A
Arginine		Arg	R
Asparagine		Asn	N
Aspartic acid		Asp	D
Cysteine		Cys	C
Glutamic acid		Glu	E
Glutamine		Gln	Q
Glycine		Gly	G
Histidine		His	H
Isoleucine		Ile	I
Leucine		Leu	L

TABLE 1.3. *Continued*

Amino Acids	Chemical Forms	Three-Letter Abbreviation	One-Letter Abbreviation
Lysine	$H_2NCH_2CH_2CH_2CH_2CHCOOH$ (NH$_2$ on CH)	Lys	K
Methionine	$CH_3SCH_2CH_2CHCOOH$ (NH$_2$ on CH)	Met	M
Phenylalanine		Phe	F
Proline		Pro	P
Serine	$HOCH_2CHCOOH$ (NH$_2$ on CH)	Ser	S
Threonine	$CH_3CH(OH)CHCOOH$ (NH$_2$ on CH)	Thr	T
Tryptophan		Trp	W
Tyrosine		Tyr	Y
Valine	$(CH_3)_2CHCHCOOH$ (NH$_2$ on CH)	Val	V

*Based on bibliography 1.6.

With the exception of one amino acid, proline, all amino acids are formed on the basis of a common structure, containing a central α-carbon atom bound with an amino group, a carboxyl group, a hydrogen atom, and a sidechain, the R group. Each amino acid has a distinct sidechain (Table 1.3). It is the sidechain that determines the unique chemical and physical properties of an amino acid. Proline has a cyclic sidechain linking the α carbon and the nitrogen, forming an imino group. Based on the structure and properties of the sidechains, amino acids can be classified into two major groups: charged and uncharged amino acids.

Charged amino acids include those with positive and negative charges. Amino acids with positive charges include arginine, histidine, and lysine, which are also known as *basic amino acids*. Those with negative charges include aspartic acid and glutamic acid,

also known as *acidic amino acids*. The charged amino acids are all hydrophilic in nature, capable of interacting with water, which determines the hydrophilic features of proteins.

The remaining 15 amino acids are all *uncharged*. Among these amino acids, five are polar amino acids, including serine, threonine, tyrosine, asparagines, and glutamine. The sidechains of these amino acids contain either polar hydrogen bond donors or acceptors, capable of interacting with water. These are considered hydrophilic amino acids. The remaining 10 uncharged amino acids possess nonpolar sidechains and interact poorly with water. These amino acids include glycine, alanine, valine, leucine, cysteine, methionine, praline, isoleucine, phenylalanine, and tryptophan. These are considered hydrophobic amino acids. Because of the versatility of the amino acids, a large number of proteins can be constructed.

Several amino acids can be modified by the addition of various chemical groups under the action of specific enzymes. For instance, a phosphate group can be added to serine, threonine, tyrosine, and histidine by a family of enzymes known as *protein kinases*, resulting in amino acid phosphorylation. Such a process plays a critical role in the regulation of a variety of cellular activities, such as cell division, adhesion, and migration. Other types of amino acid modification include the addition of the methyl group to lysine and the hydroxyl group to proline, as well as the formation of disulfate bonds between adjacent cysteine residues. These modifications are essential to the function and stability of proteins.

Protein Translation [1.7]

Protein translation is a process that synthesizes protein. The synthesis of proteins requires three basic elements: ribosomes, messenger RNA (mRNA), and transfer RNA (tRNA). Ribosomes are composed of rRNA components and regulatory proteins. There are three rRNA modules in *E. coli* ribosomes with distinct molecular sizes: 5*S*, 16*S*, and 23*S* (*S*: sedimentation velocity, a measure of molecular size). These molecules are transcribed from specific DNA sequences at specified locations of a DNA molecule. rRNA components and ribosomal proteins form two complex subunits. In prokaryotes, the two subunits are 50*S* and 30*S* in molecular size, whereas in eukaryotes they are 60*S* and 40*S*. These ribosomal structures serve as machineries for protein synthesis.

A messenger RNA molecule is a sequence of linear nucleotides, carrying genetic codons that dictate the sequence of amino acids for a specific protein. The genetic codons are stored in a DNA molecule. DNA transcription transmits the genetic codons to mRNA molecules. Each amino acid is represented by a codon of three nucleotides. In other words, each 3-nucleotide codon determines a type of amino acid to be incorporated into a peptide at a specified site. Genetic codes for all amino acids found in humans are listed in Table 1.4. In addition to the codons for amino acids, each mRNA molecule contains stop codons, including UAG, UGA, and UAA, which signal the termination of peptide translation when recognized by a ribosome.

A tRNA molecule is responsible for the transport of an amino acid to a ribosome during protein synthesis. Transfer RNA is able to not only identify a specific amino acid but also recognize a corresponding codon on a mRNA molecule, ensuring the placement of the amino acid to an appropriate position. There are two functional domains in each tRNA molecule: an anticodon and an amino acid attachment site. The anticodon is a seven-nucleotide sequence that recognizes and binds to a mRNA site according to the complementary rule. The amino acid attachment site binds to an amino acid. Each amino acid

TABLE 1.4. Genetic Codes for Amino Acids*

Amino Acids	Genetic Codes					
Ala	GCU	GCC	GCA	GCG		
Arg	CGU	CGC	CGA	CGG	AGA	AGG
Asn	AAU	AAC				
Asp	GAU	GAC				
Cys	UGU	UGC				
Gln	CAA	CAG				
Glu	GAA	GAG				
Gly	GGU	GGC	GGA	GGG		
His	CAU	CAC				
Ile	AUU	AUC	AUA			
Leu	UUA	UUG	CUU	CUC	CUA	CUG
Lys	AAA	AAG				
Met	AUG					
Phe	UUU	UUC				
Pro	CCU	CCC	CCA	CCG		
Ser	UCU	UCC	UCA	UCG	AGU	AGC
Thr	ACU	ACC	ACA	ACG		
Trp	UGG					
Tyr	UAU	UAC				
Val	GUU	GUC	GUA	GUG		
Initiation codes	AUG	GUG				
Stop codes	UAA	UAG	UGA			

*Based on bibliography 1.7.

is carried by a specific tRNA molecule. Like rRNA, a tRNA molecule is coded by a specific gene at a specified location in a DNA molecule.

Protein translation is accomplished via three steps: initiation, elongation, and termination. These processes have been well understood in prokaryotes. Here, the prokaryote protein translation system is used as an example.

Initiation. The initiation of protein translation requires RNA molecules, including mRNA, rRNA, and tRNA, and rRNA-constituted ribosomes. In addition, three regulatory factors, termed *initiation factors* (IFs) 1, 2, and 3, are necessary. Several steps are involved for the initiation of protein translation. First, the translation of protein starts with the activation of IF3, which stimulates the binding of mRNA to the 30*S* subunit of ribosome. mRNA binding induces the attachment of the 50*S* subunit to the 30*S* subunit, forming a complete ribosome (note that the two ribosomal subunits are present in separate forms without protein synthesis when they are not activated). Second, IF2 is activated to bind GTP and fMet-tRNA, which is a translation initiator tRNA carrying *N*-formylmethionine (fMet). This combination stimulates the attachment of fMet-tRNA to a specific initiation codon (AUG or GUG) on the mRNA molecule localized to an rRNA site, known as the P site. An initiation codon is preceded by a sequence that can hybridize to rRNA. The GTP molecule provides energy for the ribosomal assembly process. When a phosphate group is removed from the GTP molecule, IF2 and IF3 are released from the ribosomal complex.

Elongation. *Elongation* is a process by which amino acids are added to the initiation fMet-tRNA one at a time. Such process is regulated by several protein elongation factors (EFs), including EF-Tu, EF-Ts, and EF-G. The elongation factor EF-Tu regulates the attachment of aminoacyl-tRNAs to a specific mRNA codon localized to an rRNA site adjacent to the P site, known as the *A site*. The binding of GTP to the aminoacyl-tRNA is required in this process for supplying energy. When GTP is hydrolyzed, elongation factor EF-Ts attaches to the ribosome, regulating the release of the EF-Tu-GDP complex, leaving tRNA at the A site. Note that at this state the first amino acid or a synthesized partial peptide is attached to the P site. An enzyme called *peptidyltransferase* catalyzes a process that transfer the amino acid or partial peptide from the P to the A site. At the same time, the elongation factor EF-G initiates a process that moves the mRNA molecule by three base pairs in the $5' \rightarrow 3'$ direction, which is associated with the release of the tRNA at the P site and the transfer of the peptide together with the associated tRNA from the A to the P site (Fig. 1.9).

Termination. *Termination* of protein translation is initiated when one of the three stop codons of the mRNA, specifically UAG, UAA, and UGA, appears at the A site. At least three regulatory proteins, known as *release factors* (RFs) 1, 2, and 3, regulate translation termination. These release factors can recognize and bind to a stop codon at the A site of the ribosome, inducing the release of synthesized peptide from the P site. The ribosome is then dissociated into two free subunits. Synthesized peptides undergo protein splicing and folding processes, eventually forming proteins with a three-dimensional structure.

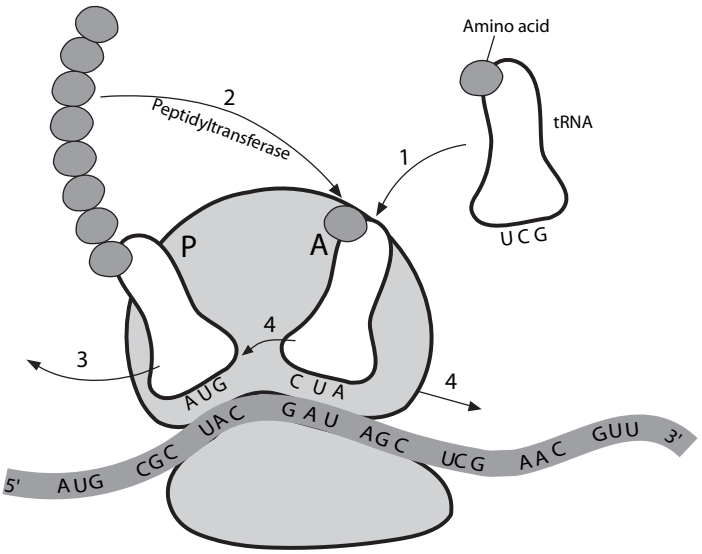


Figure 1.9. Schematic representation of protein translation. Several steps are involved in protein elongation: (1) recruitment of an aminoacyl-tRNA molecule to site A; (2) transfer of a partial peptide from site P to site A; (3) removal of the tRNA at site P; (4) movement of the peptide–tRNA complex from site A to site P; (5) movement of the rRNA complex toward the 3′ direction by three base pairs (based on bibliography 1.7).

Protein Folding and Architecture [1.8]

All proteins are folded into a three-dimensional structure after being translated. A protein is composed of one or more peptides, which are sequences of linearly jointed amino acids via peptide bonds. Three atoms from each amino acid, including the nitrogen from the amino group, the α-carbon, and the carbon from the carboxyl group, join together to form a central structure for each peptide called the *polypeptide backbone*. A peptide chain is synthesized in a ribosome based on the codons of an mRNA. The peptide end with a free amino group is referred to as the *N-terminus*, where peptide synthesis begins, and the end with a carboxyl group is the *C-terminus*. The counting of the number of amino acids starts at the *N*-terminus. Most peptide bonds can rotate freely, giving the flexibility of forming a variety of different conformations for proteins.

A peptide is usually folded into a three-dimensional (3D) structure, resulting in a final conformation that exhibits maximal stability and functionality. A denatured protein molecule under harsh conditions, such as extreme pH and a high concentration of urea, lose not only its conformation but also function. However, proteins are able to refold back to their original conformation and regain their functions under restored physiological conditions. Various chemical and physical features of the amino acids in a peptide chain determine the process of protein folding and the 3D protein conformation. The hydrophobic and hydrophilic features of amino acids play a critical role in controlling protein folding and conformation. The uncharged hydrophobic nonpolar sidechains of the amino acids intend to pack themselves in the core of a protein to minimize exposure to water molecules. The core of a protein contains the most conservative amino acids. In contrast, most charged and polar sidechains of amino acids are localized to the surface of a protein and capable of interacting with water molecules.

By X-ray diffraction and nuclear magnetic resonance spectroscopy, the structure of proteins can be determined to atomic accuracy. Protein structural analysis has demonstrated that proteins usually contain two types of secondary structure: α-helix and β-sheet. These structures are common in most proteins, although the overall conformation varies from protein to protein. An α-*helix* is a right-handed coiled structure with 3.6 residues per turn (Fig. 1.10). The helical structure is formed on the basis of hydrogen bonding between adjacent polar groups of the peptide backbone. A β-*sheet* is a structure containing several antiparallel segments of a single peptide, which is formed as a result

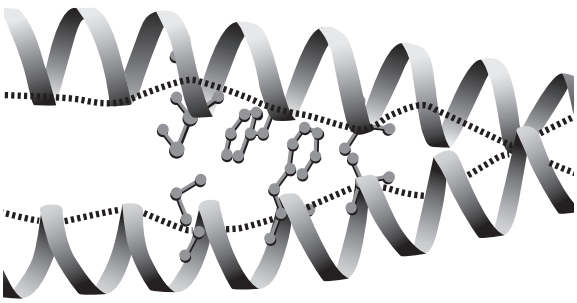


Figure 1.10. Crystallographic structure of a parallel α-helical coiled-coil dimer of the intermediate filament protein vimentin. (Reprinted from Burkhard P et al: *Trends Cell Biol* 11:82–8, 2001 by permission of Elsevier.)

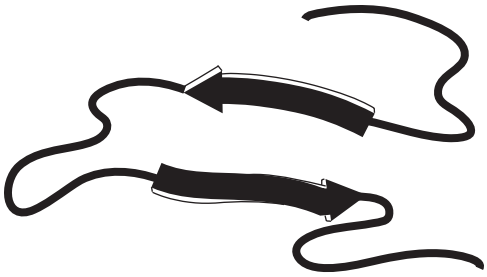


Figure 1.11. Schematic representation of the structure of a β-sheet-containing protein (based on bibliography 1.8).

of the back-and-forth turning of the peptide (Fig. 1.11). The antiparallel segments are linked together by lateral hydrogen bonds between the polar groups of the peptide backbone.

There exist peptide fragments that are not organized into regular structures like α-helices and β-sheets. These peptides exhibit disordered structures and can move more freely than a α-helix and β-sheet. Such structures are often found at the *N*- and *C*-termini. Collectively, α-helices, β-sheets, and irregular fragments are referred to as *secondary structures* (note that the amino acid sequences are known as the *primary structure*). These structures can be further organized into higher levels of 3D conformation. Examples of 3D protein structures include the coiled-coil structure, which is a complex of paired α-helices interacting laterally via hydrophobic bonds (Fig. 1.10), and the β-barrel structure, which is a cylindrical structure formed by a number of β-sheets. The formation of these superstructures enhances the stability of proteins. Furthermore, most proteins are composed of multiple peptides, which are integrated into a protein molecule. Proteins also form complexes such as dimmers and trimers. All these forms are essential for the functionality of proteins.

Changes in Protein Conformation [1.8]

Proteins undergo dynamic changes in their conformation. The atoms of a protein may move in a very fast speed in the order of 100 m/s within a nanometer range. Protein–protein interactions may induce conformational changes on the molecular scale. Such conformational changes play a critical role in the regulation of molecular activities. For instance, the binding of a growth factor to its receptor may cause a conformational change in the receptor, initiating autophosphorylation of the cytoplasmic receptor tyrosine kinase, which is associated with most growth factors. Conformational changes and autophosphorylation are critical processes for the activation of mitogenic signaling pathways. In the contractile apparatus of muscular cells, a conformational change in the myosin head, on the activation of a myosin molecule, is an essential process that initiates the sliding of actin filaments and the generation of forces. Conformational changes in protein kinases and phosphatases, such as the Src protein tyrosine kinase and the Src homology 2 domain-containing protein tyrosine phosphatase, control the state of molecular activation. It is commonly received that a protein conformational change is an essential process for the regulation of protein functions.

LIPIDS

There are various types of lipid molecules in mammalian cells, including phospholipids, glycolipids, steroids, and triglycerides. These lipid molecules play important roles in the construction of cellular structures and in the regulation of cellular functions. Lipids are the basic building blocks for cell and subcellular membranes, serve as hormones and intracellular signaling molecules, and contribute to the production of energy. The structure and function of common lipids are briefly described in the following sections.

Phospholipids [1.9]

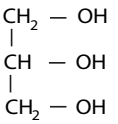
Phospholipids are the primary constituents of cell membrane. A phospholipid contains several basic elements, including an alcohol, two fatty acid chains, and a phosphate group, which can be bonded with another alcohol group. Based on the alcohol structure, phospholipids can be further divided into two subgroups: phosphoglycerides and sphingolipids. A *phosphoglyceride* contains a glycerol as an alcohol group, whereas a *sphingolipid* contains a sphingosine.

Phosphoglycerides. In a phosphoglyceride molecule, two of the three —OH groups of the alcohol (glycerol) molecule are bonded with fatty acids via ester bonds, and the remaining —OH group is esterified by a phosphate group. The phosphate group can be bonded with another chemical groups, which can be either inositol, serine, ethanolamine, or choline (Fig. 1.12). Each combination gives rise to a distinct phosphoglyceride, including phosphatidylinositol, phosphatidylserine, phosphatidylethanolamine, or phosphatidylcholine, respectively. It is important to note that some of these phosphoglycerides not only contribute to the construction of cell membrane, but also serve as signaling molecules. For instance, phosphatidylinositol (Fig. 1.13), when phosphorylated into phosphatidylinositol 3,4 biphosphates, play a critical role in the regulation of G-protein receptor-initiated signal transduction (see Chapter 5).

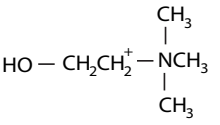
A phosphoglyceride molecule contains a polar alcohol head and nonpolar fatty acid tails, which render the molecule amphipathic in nature, that is, hydrophilic at the head and hydrophobic at the tail. Such a modular arrangement determines the form of phosphoglyceride aggregation while they are mixed in an aqueous solution. The hydrophilic head of phosphoglyceride interacts with the water molecules, while the hydrophobic tail intends to interact with the hydrophobic tails from different phosphoglyceride molecules, resulting in the spontaneous formation of a lipid bilayer with two water-contacting surfaces composed of hydrophilic heads and an internal layer composed of hydrophobic tails (Fig. 1.14). Indeed, all cell membranes are established on the basis of such a principle.

Sphingolipids. *Sphingolipids* are composed of a sphingosine, two fatty acid chains, and a phosphate group (Fig. 1.15). These molecules possess properties similar to those of phosphoglycerides and can be found in the membrane of many cell types. Sphingolipids can aggregate into microdomains in cell membranes, which may play a role in targeting specific proteins to the plasma membrane and in organizing membrane-associated signaling pathways. For instance, K^+ and other ion channels are localized to lipid microdomains on the cell surface. Sphingolipids can interact with ion channels and mediate the localization of the ion channels. Sphingolipids are also involved in the regulation of cellular

Glycerol



Choline



Phosphatidylcholine

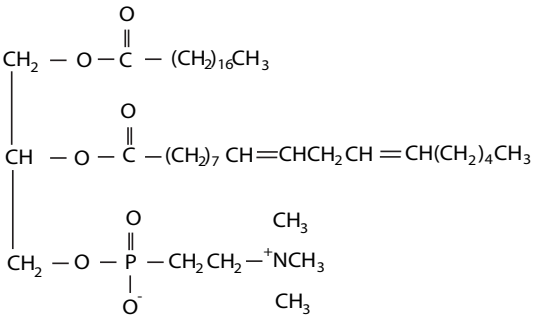


Figure 1.12. Chemical composition of glycerol, choline, and phosphatidylcholine (based on bibliography 1.9).

Phosphatidyl inositol (R = lipid molecule)

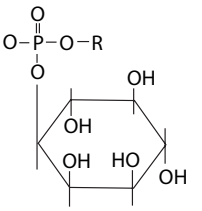


Figure 1.13. Chemical composition of phosphatidyl inositols (PI). Based on bibliography 1.9.

activities, such as cell apoptosis and proliferation. Some sphingolipid molecules such as ceramide and sphingosine promote cell apoptosis, whereas others such as sphingosine-1-phosphate induce cell proliferation. These observations demonstrate the involvement of sphingolipids in the regulation of signal transduction.

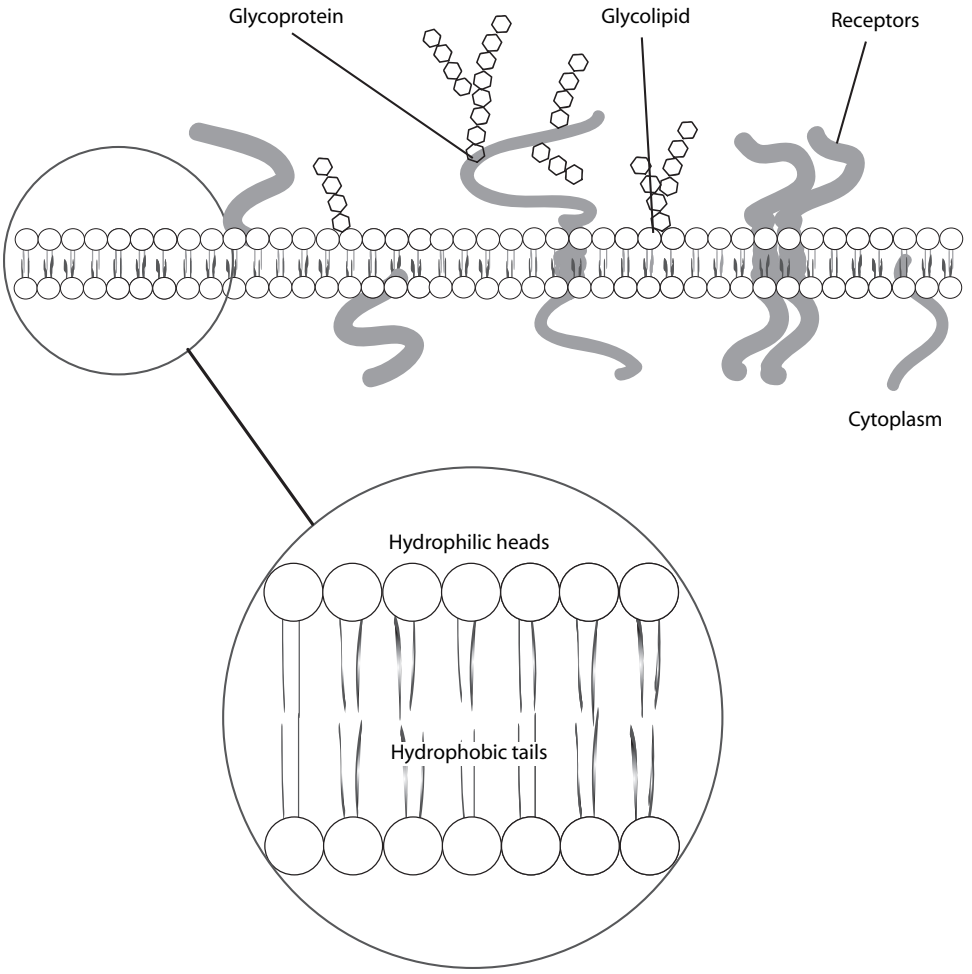


Figure 1.14. Constituents of a lipid bilayer.

Sphingolipids

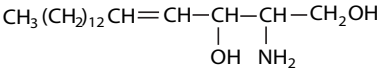


Figure 1.15. Chemical composition of the sphingolipid molecule sphingosine(based on bibliography 1.9).

Glycolipids [1.10]

Glycolipids are lipid molecules that contain carbohydrates or sugar residues, including glucose and galactose. Glycolipids are found in the membrane of all cell types and primarily distributed in the extracellular half of the lipid bilayer, constituting about 5% of the lipid molecules. These molecules often form aggregates via self-assembly. Given the

unique location of the glycolipids, it has been speculated that these molecules may play roles in regulating the interaction of cells with extracellular factors and may also serve to protect the cell from extreme chemical conditions.

Steroids [1.11]

Steroids are lipids that contain a structure with four fused rings, including three cyclohexane rings and a cyclopentane ring. Steroid compounds found in mammalian cells include cholesterol, vitamin D, progesterone, adrenocortical hormones, gonadal hormones, and bile salts. Cholesterol is the most abundant steroid in mammals and exists in a free form or a form esterified with fatty acids (Fig. 1.16). It participates in the construction of cell membrane structures, and is a basic molecule for the derivation of other steroids, such as vitamin D, adrenocortical hormones, gonadal hormones, and bile salts.

Because of the influence of blood cholesterol on atherogenesis, cholesterol has received much attention. A high level of blood cholesterol is referred to as *hypercholesterolemia*, a condition facilitating the development of atherogenesis. Cholesterol can be taken up from digested foods and also synthesized by the liver. The rate of cholesterol synthesis is dependent on the blood level of cholesterol, which controls cholesterol synthesis based on a negative feedback mechanism. Namely, a high level of blood cholesterol inhibits cholesterol synthesis. Dietary cholesterol can be absorbed into blood from the small intestine in a form known as *chylomicron*, which contains cholesterol, triglycerides, and apoproteins. Under the action of lipoprotein lipase, triglycerides are released from the chylomicrons, forming chylomicron remnants that consist of cholesteryl esters and apoproteins. The chylomicron remnants can be taken up by hepatocytes in the liver, where cholesterol is cleaved and released into the blood. Free cholesterol molecules can be taken up by cells for various purposes, including the construction of cell membrane, formation of bile, synthesis of hormones, and generation of endogenous lipoproteins.

In the liver, cholesterol can be used for generating endogenous low-density lipoproteins (LDL), which is circulated in the blood for 1–2 days and constitutes the major pool of plasma cholesterol (60–70% of total cholesterol). Circulating LDL is a major form that delivers cholesterol to needy cells. The release of cholesterol from a LDL molecule results in the formation of high-density lipoproteins (HDL), consisting of apoproteins and residual cholesterol. HDL can be reused in the liver to form LDL.

Clinically, hypercholesterolemia can be divided into two groups: primary and secondary. The primary hypercholesterolemia is an inherited disease and is induced by genetic defects. In some patients, the cholesterol metabolic disorder is due to a single gene defect, which is called monogenic hypercholesterolemia. This type of disorder can be predicted on the basis of the Mendelian genetic mechanism; some members of a family inherit the

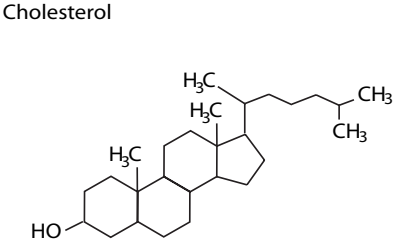


Figure 1.16. Chemical composition of cholesterol.

disease whereas others do not. In other patients, it is caused by a combination of multiple gene mutations and environmental stimulation. This type of cholesterol metabolic disorder is called polygenic hypercholesterolemia. In this case, the plasma cholesterol level of all family members may increase at some time during their lifespans, if the intake of saturated fats and cholesterol is high. Of all patients with hypercholesterolemia, the vast majority belong to the polygenic type. Secondary hypercholesterolemia is a complication of metabolic disorders, such as diabetes.

BIBLIOGRAPHY

1.1. Composition and Structure of DNA

Berg JM, Tymoczko JL, Stryer L: *Biochemistry*, 5th ed, Freeman, New York, 2002.

Jones ME: Pyrimidine nucleotide biosynthesis in animals: Genes, enzymes, and regulation of UMP biosynthesis, *Annu Rev Biochem* 49:253–79, 1980.

Kornberg A: Biologic synthesis of deoxyribonucleic acid, *Science* 131:1503–8, 1960.

Meselson M, Stahl FW: The replication of DNA in Escherichia coli, *Proc Natl Acad Sci USA* 44:671–82, 1958.

Watson JD, Crick FHC: Genetical implications of the structure of deoxyribonucleic acid, *Nature* 171:964–7, 1953.

Watson JD, Crick FHC: Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid, *Nature* 171:737–8, 1953.

Wilkins MHF, Stokes AR, Wilson HR: Molecular structure of deoxypentose nucleic acids, *Nature* 171:738–40, 1953.

1.2. Organization of Chromosomes

Grunstein M: Histones as regulators of genes, *Sci Am* 267(4):68–74B, 1992.

Heck MMS: Condensins, cohesins, and chromosome architecture: How to make and break a mitotic chromosome, *Cell* 91:5–8, 1997.

Kornberg RD: Chromatin structure: A repeating unit of histones and DNA, *Science* 184:868–71, 1974.

Koshland D, Strunnikov A: Mitotic chromosome condensation, *Annu Rev Cell Biol* 12:305–33, 1996.

Luger K, Mader AW, Richmond RK, Sargent DE, Richmond TJ: Crystal structure of the nucleosome core particle at 2.8 A resolution, *Nature* 389:251–60, 1997.

Paulson JR, Laemmli UK: The structure of histone-depleted metaphase chromosomes, *Cell* 12:817–28, 1977.

Richmond TJ, Finch JT, Rushton B, Rhodes D, Klug A: Structure of the nucleosome core particle at 7A resolution, *Nature* 311:532–7, 1984.

Saitoh Y, Laemmli UK: Metaphase chromosome structure: Bands arise from a differential folding path of the highly AT-rich scaffold, *Cell* 76:609–22, 1994.

Van Holde KE, Zlatanovai J: Chromatin higher order structure: Chasing a mirage, *J Biol Chem* 270:8373–6, 1995.

Huebert DJ, Bernstein BE: Genomic views of chromatin, *Curr Opin Genet Dev* 15(5):476–81, Oct 2005.

Zhao H, Dean A: Organizing the genome: Enhancers and insulators, *Biochem Cell Biol* 83(4):516–24, Aug 2005.

Luger K, Hansen JC: Nucleosome and chromatin fiber dynamics, *Curr Opin Struct Biol* 15(2):188–96, April 2005.

Gruenbaum Y, Margalit A, Goldman RD, Shumaker DK, Wilson KL: The nuclear lamina comes of age, *Nat Rev Mol Cell Biol* 6(1):21–31, Jan 2005.

Becskei A, Mattaj IW: Quantitative models of nuclear transport, *Curr Opin Cell Biol* 17(1):27–34, Feb 2005.

Gregory RI, Shiekhattar R: Chromatin modifiers and carcinogenesis, *Trends Cell Biol* 14(12):695–702, Dec 2004.

Wienberg J: The evolution of eutherian chromosomes, *Curr Opin Genet Dev* 14(6):657–66, Dec 2004.

1.3. Functional Units of DNA

Blackburn EH: Switching and signaling at the telomere, *Cell* 106:661–73, 2001.

Blattner ER, Plunkett G, Bloch CA, Perna NT, Burland V et al: The complete genome sequence of Escherichia coli K-12, *Science* 277:1453–62, 1997.

Carbon J: Yeast centromeres: Structure and function, *Cell* 37:351–3, 1984.

Clarke L: Centromeres of budding and fission yeasts, *Trends Genet* 6:150–4, 1990.

Greider CW: Telomeres do D-loop-Tloop, *Cell* 97:419–2, 1999.

Ingram VM: Gene mutations in human hemoglobin: The chemical difference between normal and sickle cell hemoglobin, *Nature* 180:326–8, 1957.

Fritsch EF, Lawn RM, Maniatis T: Molecular cloning and characterization of the human beta-like globin gene cluster, *Cell* 19:959–72, 1980.

Gilbert W, de Souza SJ, Long M: Origin of genes, *Proc Natl Acad Sci USA* 94:7698–703, 1997.

Henikoff S, Greene EA, Pietrokovski S, Bork P, Attwood TK, Hood L: Gene families: The taxonomy of protein paralogs and chimeras, *Science* 278:609–14, 1997.

International Human Genome Sequencing Consortium: Initial sequencing and analysis of the human genome, *Nature* 409:860–921, 2001.

Mouse Genome Sequencing Consortium: Initial sequence and comparative analysis of the mouse genome, *Nature* 420:520–62, 2002.

Pluta AE, MacKay AM, Ainsztein AM, Goldberg AG, Earnshaw WC: The centromere: Hub of chromosomal activities, *Science* 270:1591–4, 1995.

Rubin GM et al (and 54 others): Comparative genomics of the eukaryotes, *Science* 287:2204–215, 2000.

Schueler MG, Higgins AW, Rudd MK, Gustashaw K, Willard HE: Genomic and genetic definition of a functional human centromere, *Science* 294:109–15, 2001.

Sun X, Wahlstrom J, Karpen G: Molecular structure of a functional Drosophila centromere, *Cell* 91:1007–19, 1997.

Szostak JW, Blackburn EH: Cloning yeast telomeres on linear plasmid vectors, *Cell* 29:245–55, 1982.

Stoltzfus A, Spencer DE, Zuker M, Logsdon JM Jr, Doolittle WE: Testing the exon theory of genes: The evidence from protein structure, *Science* 265:202–7, 1994.

Venter JC et al (and 273 colleagues): The sequence of the human genome, *Science* 291:1304–51, 2001.

Wiens GR, Sorger PK: Centromeric chromatin and epigenetic effects in kinetochore assembly, *Cell* 93:313–6, 1998.

Zakian VA: Telomeres: Beginning to understand the end, *Science* 270:1601–7, 1995.

1.4. DNA Replication

Kornberg A, Baker TA: *DNA Replication*, 2nd ed, Freeman, New York, 1991.

Baltimore D: RNA-dependent DNA polymerase in virions of RNA tumour viruses, *Nature* 226:1209–11, 1970.

Temin HM, Mizutani S: RNA-dependent DNA polymerase in virions of Rous sarcoma virus, *Nature* 226:1211–3, 1970.

Baker TA, Bell SP: Polymerases and the replisome: Machines within machines, *Cell* 92:296–305, 1998.

Bell SP, Dutta A: DNA replication in eukaryotic cells, *Annu Rev Biochem* 71:333–74, 2002.

Benkovic SJ, Valentine AM, Salinas F: Replisome-mediated DNA replication, *Annu Rev Biochem* 70:181–208, 2001.

Frick DN, Richardson CC: DNA primases, *Annu Rev Biochem* 70:39–80, 2001.

Gilbert DM: Making sense of eukaryotic DNA replication origins, *Science* 294:96–100, 2001.

Hubscher U, Maga G, Spadari S: Eukaryotic DNA polymerases, *Annu Rev Biochem* 71:133–63, 2002.

Kunkel TA, Bebenek K: DNA replication fidelity, *Annu Rev Biochem* 69:497–529, 2000.

McEachern MJ, Krauskopf A, Blackburn EH: Telomeres and their control, *Annu Rev Genet* 34:331–58, 2000.

Ogawa T, Okazaki T: Discontinuous DNA replication, *Annu Rev Biochem* 49:421–57, 1980.

Stinthcomb D, Struhl K, Davis RW: Isolation and characterization of a yeast chromosomal replicator, *Nature* 282:39–43, 1979.

Waga S, Stillman B: Anatomy of a DNA replication fork revealed by reconstitution of SV 40 DNA replication in vitro, *Nature* 369:207–12, 1994.

Waga S, Stillman B: The DNA replication fork in eukaryotic cells, *Annu Rev Biochem* 67:721–51, 1998.

West SC: DNA helicases: New breeds of translocating motors and molecular pumps, *Cell* 86:177–80, 1996.

Batty DP, Wood RD: Damage recognition in nucleotide excision repair of DNA, *Gene* 241:193–204, 2000.

De Laat WL, Jaspers NGJ, Hoeijmakers JHJ: Molecular mechanism of nucleotide excision repair, *Genes Dev* 13:768–85, 1999.

Goodman MF: Error-prone repair DNA polymerases in prokaryotes and eukaryotes, *Annu Rev Biochem* 71:17–50, 2002.

Harfe BD, Jinks-Robertson S: DNA mismatch repair and genetic instability, *Annu Rev Genet* 34:359–99, 2000.

Hoeijmakers JHJ: Genome maintenance mechanisms for preventing cancer, *Nature* 411:366–74, 2001.

Khanna KK, Jackson SP: DNA double strand breaks: Signaling, repair and the cancer connection, *Nature Genet* 27:247–54, 2001.

Livneh Z: DNA damage control by novel DNA polymerases: translesion replication and mutagenesis, *J Biol Chem* 276:25639–42, 2001.

Modrich P: Strand-specific mismatch repair in mammalian cells, *J Biol Chem* 272:24727–30, 1997.

Sancar A: DNA excision repair, *Annu Rev Biochem* 65:43–81, 1996.

1.5. RNA Composition, Structure, and Transcription

Brenner S, Jacob F, Meselson M: An unstable intermediate carrying information from genes to ribosomes for protein synthesis, *Nature* 190:576–81, 1961.

Butler JE, Kadonaga JT: The RNA polymerase II core promoter: A key component in the regulation of gene expression, *Genes Dev* 16(20):2583–92, 2002.

Shilatifard A, Conaway RC, Conaway JW: The RNA polymerase II elongation complex, *Annu Rev Biochem* 72:693–715, 2003.

Bushnell DA, Westover KD, Davis RE, Kornberg RD: Structural basis of transcription: An RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms, *Science* 303:983–8, 2004.

Westover KD, Bushnell DA, Kornberg RD: Structural basis of transcription: Separation of RNA from DNA by RNA polymerase II, *Science* 303:1014–6, 2004.

Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, Gifford DK, Melton DA, Jaenisch R, Young RA: Core transcriptional regulatory circuitry in human embryonic stem cells, *Cell* 122:947–56, 2005.

Myers LC, Kornberg RD: Mediator of transcriptional regulation, *Annu Rev Biochem* 69:729–49, 2000.

Naar AM, Lemon BD, Tjian R: Transcriptional coactivator complexes, *Annu Rev Biochem* 70:475–501, 2001.

Orphanides G, Reinberg D: A unified theory of gene expression, *Cell* 108:439–51, 2002.

Atchison ML: Enhancers: Mechanisms of action and cell specificity, *Annu Rev Cell Biol* 4:127–53, 1988.

Berger SL: Histone modifications in transcriptional regulation, *Curr Opin Genet Dev* 12:142–8, 2002.

Dvir A, Conaway JW, Conaway RC: Mechanism of transcription initiation and promoter escape by RNA polymerase II, *Curr Opin Genet Dev* 11:209–14, 2001.

Conaway JW, Shilatifard A, Dvir A, Conaway RC: Control of elongation by RNA polymerase II, *Trends Biochem Sci* 25:375–80, 2000.

Horn PJ, Peterson CL: Molecular biology. Chromatin higher order folding—wrapping up transcription, *Science* 297:1824–7, 2002.

McKenna NJ, O'Malley BW: Combinatorial control of gene expression by nuclear receptors and coregulators, *Cell* 108:465–74, 2002.

Naar AM, Lemon BD, Tjian R: Transcriptional coactivator complexes, *Annu Rev Biochem* 70:475–501, 2001.

Narlikar GJ, Fan HY, Kingston RE: Cooperation between complexes that regulate chromatin structure and transcription, *Cell* 108:475–87, 2002.

Pabo CO, Sauer RT: Transcription factors: Structural families and principles of DNA recognition, *Annu Rev Biochem* 61:1053–95, 1992.

Abelson J, Trotta CR, Li H: tRNA splicing, *J Biol Chem* 273:12685–8, 1998.

Blanc V, Davidson NO: C-to-U RNA editing: Mechanisms leading to genetic diversity, *J Biol Chem* 278:1395–8, 2003.

Maas S, Rich A, Nishikura K: A-to-I RNA editing: Recent news and residual mysteries, *J Biol Chem* 278:1391–4, 2003.

Madison-Antenucci S, Grams J, Hajduk SL: Editing machines: The complexities of trypanosome RNA editing, *Cell* 108:435–8, 2002.

Maniatis T, Tasic B: Alternative pre-mRNA splicing and proteome expansion in metazoans, *Nature* 418:236–43, 2002.

Moore MJ: Nuclear RNA turnover, *Cell* 108:431–4, 2002.

Proudfoot NJ, Furger A, Dye MJ: Integrating mRNA processing with transcription, *Cell* 108:501–12, 2002.

Ruskin B, Krainer AR, Maniatis T, Green MR: Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro, *Cell* 38:317–31, 1984.

Smith CW, Valcarcel J: Alternative pre-mRNA splicing: The logic of combinatorial control, *Trends Biochem Sci* 25:381–8, 2000.

1.6. Protein Composition and Structure

Branden C, Tooze J: *Introduction to Protein Structure*, 2nd ed, Garland, New York, 1999.

Chothia C, Finkelstein AV: The classification and origins of protein folding patterns, *Annu Rev Biochem* 59:1007–39, 1990.

Holm L, Sander C: Mapping the protein universe, *Science* 273:595–602, 1996.

Kendrew JC: The three-dimensional structure of a protein molecule, *Sci Am* 205:96–111, 1961.

Levitt M, Gerstein M, Huang E, Subbiah S, Tsai J: Protein folding: The endgame, *Annu Rev Biochem* 66:549–79, 1997.

Richardson JS: The anatomy and taxonomy of protein structure, *Adv Protein Chem* 34:167–339, 1981.

Sanger F: Sequences, sequences, and sequences, *Annu Rev Biochem* 57:1–28, 1988.

Umbarger HE: Amino acid biosynthesis and its regulation, *Annu Rev Biochem* 47:533–606, 1978.

1.7. Protein Translation

Ban N, Nissen P, Hansen J, Moore PB, Steitz TA: The complete atomic structure of the large ribosomal subunit at 2.4 A resolution, *Science* 289:905–20, 2000.

Dever TE: Gene-specific regulation by general translation factors, *Cell* 108:545–56, 2002.

Gray NK, Wickens M: Control of translation initiation in animals, *Annu Rev Cell Dev Biol* 14:399–458, 1998.

Moore PB, Steitz TA: The structural basis of large ribosomal subunit function, *Annu Rev Biochem* 72:813–50, 2003.

Moore PB, Steitz TA: The involvement of RNA in ribosome function, *Nature* 418:229–35, 2002.

Ogle JM, Ramakrishnan V: Structural insights into translational fidelity, *Annu Rev Biochem* 74:129–77, 2005.

Sachs AB: Cell cycle-dependent translation initiation: IRES elements prevail, *Cell* 101(3):243–5, 2000.

1.8. Protein Folding and Architecture

Burkhard P, Stetefeld J, Strelkov SV: Coiled coils: A highly versatile protein folding motif, *Trends Cell Biol* 11:82–8, 2001.

Crick FHC, Barnett L, Brenner S, Watts-Tobin RJ: General nature of the genetic code for proteins, *Nature* 192:1227–32, 1961.

Nirenberg M, Leder P: RNA codewords and protein synthesis, *Science* 145:1399–407, 1964.

Nirenberg MW, Matthaei JH: The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides, *Proc Natl Acad Sci USA* 47:1588–602, 1961.

Gahmberg CG, Tolvanen M: Why mammalian cell surface proteins are glycoproteins, *Trends Biochem Sci* 21:308–11, 1996.

Gething MJ, Sambrook J: Protein folding in the cell, *Nature* 355:33–45, 1992.

Schiene C, Fischer G: Enzymes that catalyse the restructuring of proteins, *Curr Opin Struct Biol* 10:40–5, 2000.

Sigler PB, Xu Z, Rye HS, Burston SG, Fenton WA, Horwich AL: Structure and function in GroEL-mediated protein folding, *Annu Rev Biochem* 67:581–608, 1998.

Udenfriend S, Kodukula K: How glycosylphosphatidylinositol-anchored membrane proteins are made, *Annu Rev Biochem* 64:563–91, 1995.

Zhang FL, Casey PJ: Protein prenylation: Molecular mechanisms and functional consequences, *Annu Rev Biochem* 65:241–69, 1996.

1.9. Phospholipids

Deguchi H, Yegneswaran S, Griffin JH: Sphingolipids as bioactive regulators of thrombin generation, *J Biol Chem* 279:12036–12042, 2004.

Martens JR, O'Connell K, Tamkun M: Targeting of ion channels to membrane microdomains: Localization of KV channels to lipid rafts, *Trends Pharmacol Sci* 25:16–21, 2004.

van Meer G, Burger KN: Sphingolipid trafficking—sorted out? *Trends Cell Biol* 2:332–7, 1992.

Wakil SJ, Stoops JK, Joshi VC: Fatty acid synthesis and its regulation, *Annu Rev Biochem* 52:537–79, 1983.

Lee A: Membrane structure, *Curr Biol* 11:R811–4, 2001.

McLaughlin S, Murray D: Plasma membrane phosphoinositide organization by protein electrostatics, *Nature* 438:605–11, 2005.

Behnia R, Munro S: Organelle identity and the signposts for membrane traffic, *Nature* 438:597–604, 2005.

Chalfant CE, Spiegel S: Sphingosine 1-phosphate and ceramide 1-phosphate: Expanding roles in cell signaling, *J Cell Sci* 118:4605–12, 2005.

Niggli V: Regulation of protein activities by phosphoinositide phosphates, *Annu Rev Cell Dev Biol* 21:57–79, 2005.

Rosen H, Goetzl EJ: Sphingosine 1-phosphate and its receptors: An autocrine and paracrine network, *Nature Rev Immunol* 5:560–70, 2005.

Branton D, Cohen CM, Tyler J: Interaction of cytoskeletal proteins on the human erythrocyte membrane, *Cell* 24:24–32, 1981.

Thompson TE, Tillack TW: Organization of glycosphingolipids in bilayers and plasma membranes of mammalian cells, *Annu Rev Biophys Biophys Chem* 14:361–86, 1985.

1.10. Glycolipids

Ramstedt B, Slotte JP: Membrane properties of sphingomyelins, *FEBS Lett* 531:33–7, 2002.

Cremesti AE, Goni FM, Kolesnick R: Role of sphingomyelinase and ceramide in modulating rafts: Do biophysical properties determine biologic outcome? *FEBS Lett* 531:47–53, 2002.

Lee A: Membrane structure, *Curr Biol* 11:R811–4, 2001.

Bush CA, Martin-Pastor M, Imberty A: Structure and conformation of complex carbohydrates of glycoproteins, glycolipids, and bacterial polysaccharides, *Annu Rev Biophys Biomol Struct* 28:269–93, 1999.

Stoffel W, Bosio A: Myelin glycolipids and their functions, *Curr Opin Neurobiol* 7:654–61, 1997.

Udenfriend S, Kodukula K: How glycosylphosphatidylinositol-anchored membrane proteins are made, *Annu Rev Biochem* 64:563–91, 1995.

Englund PT: The structure and biosynthesis of glycosyl phosphatidylinositol protein anchors, *Annu Rev Biochem* 62:121–38, 1993.

Paulson JC, Colley KJ: Glycosyltransferases. Structure, localization, and control of cell type-specific glycosylation, *J Biol Chem* 264:17615–8, 1989.

Low MG: Glycosyl-phosphatidylinositol: A versatile anchor for cell surface proteins, *FASEB J* 3:1600–8, 1989.

Thompson TE, Tillack TW: Organization of glycosphingolipids in bilayers and plasma membranes of mammalian cells, *Annu Rev Biophys Biophys Chem* 14:361–86, 1985.

Rauvala H, Finne J: Structural similarity of the terminal carbohydrate sequences of glycoproteins and glycolipids, *FEBS Lett* 97:1–8, 1979.

1.11. Steroids

Simons K, Toomre D: Lipid rafts and signal transduction, *Nature Rev Mol Cell Biol* 1:31–9, 2000.

Incardona JP, Eaton S: Cholesterol in signal transduction, *Curr Opin Cell Biol* 12:193–203, 2000.

Bruce C, Chouinard RA Jr, Tall AR: Plasma lipid transfer proteins, high-density lipoproteins, and reverse cholesterol transport, *Annu Rev Nutr* 18:297–330, 1998.

Lagrost L, Desrumaux C, Masson D, Deckert V, Gambert P: Structure and function of the plasma phospholipid transfer protein, *Curr Opin Lipidol* 9:203–9, 1998.

Simons K, Ikonen E: Functional rafts in cell membranes, *Nature* 387:569–72, 1997.

Yeagle PL: Lipid regulation of cell membrane structure and function, *FASEB J* 3:1833–42, 1989.

Kummerow FA: Modification of cell membrane composition by dietary lipids and its implications for atherosclerosis, *Ann NY Acad Sci* 414:29–43, 1983.

Griffiths AJF: *An Introduction to Genetic Analysis*, 6th ed, Freeman, New York, 1996.

Bretscher MS, Munro S: Cholesterol and the Golgi apparatus, *Science* 261:1280–1, 1993.