Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Features of Host Cells: Cellular and Molecular Biology Review

As obligate intracellular parasites, viruses are completely dependent upon a host cell for their replication. They use energy generated by the host cell, and they exploit the host’s machinery to manufacture viral proteins. This chapter takes us for a tour inside a eukaryotic cell, highlighting the processes that viruses take advantage of during infection.

3.1 THE BASIC ORGANIZATION OF THE CELL

There are three domains of life—Bacteria, Archaea, and Eukarya. The organisms within these groups are divided depending on the presence or absence of a nucleus within the cell(s) of the organism. Prokaryotes are organisms without a nucleus to wall off their genetic material from the rest of the cell, while eukaryotes are organisms that contain a nucleus within their cells. All organisms within Bacteria and Archaea are prokaryotes, whereas Eukarya—as the name suggests—contains eukaryotes. Viruses exist that infect cells of all three domains. Most of the viruses that are discussed in this book infect humans and other animals, which are eukaryotes.

The defining characteristic of a eukaryotic cell is the nucleus, which is generally located in the center of a cell. Many structures, called organelles, are distributed in the liquid cytosol between the nucleus and the plasma membrane of the cell (Fig. 3.1). In the same way that each organ of our body performs a specialized function, the organelles within a cell each play a specific role in maintaining an operational cell.

Most organelles are composed of the same lipid membrane that creates the plasma membrane of the cell. This membrane is only two molecules thick and made of phospholipids. Phospholipids are a class of lipids; fats, oils, and waxes are other lipids. They are so named because the molecule has two parts: a polar head that contains a phosphate group, and a non-polar portion that is composed of two fatty acid lipid tails (Fig. 3.2A). As described in the Chapter 2, “Virus Structure and Classification,” Refresher on Chemical Bonds, water is a polar molecule and readily associates with other polar molecules. The phospholipid head of the phospholipid, being polar, is hydrophilic, while the nonpolar tails do not associate with water and are hydrophobic. Because of this amphipathic nature of a phospholipid, a group of phospholipids placed in an aqueous solution (such as the environment of a cell) will spontaneously assemble into a double layer of phospholipid molecules with the hydrophilic polar heads of the molecules facing the aqueous solution and the hydrophobic nonpolar tails associating with each other (Fig. 3.2B). This forms an effective barrier to prevent large molecules from escaping or...
Phospholipids and membrane structure. The plasma membrane of the cell is composed of many phospholipid molecules, which are amphipathic: the phosphate head is hydrophilic, while the fatty acid tails of the molecule are hydrophobic (A). As such, when they are placed in an aqueous solution, like that of the cell environment (B), they will spontaneously form a double layer, or bilayer, with the polar heads toward the aqueous solution and the fatty acid tails facing each other, forming a barrier between the extracellular and intracellular environment.

Electron micrographs of cell organelles. (A) The nucleus of a lung cell, surrounded by rER. Note the ribosomes on the rER that give it its characteristic rough appearance. (B) The Golgi complex of a lung cell, with its characteristic flattened sacs of membrane. (C) Two mitochondria within a lung cell. Images courtesy of Louisa Howard.

Gaining entry into the cell. In the same way that the plasma membrane acts as a barrier for the contents of the cell, most of the organelles within the cell also use a phospholipid bilayer to wall off their contents from the cytosol of the cell.

Word Origin: Endoplasmic Reticulum
From the Greek endo, meaning “within,” and Latin reticulum, meaning “little net”—the little network within the cytoplasm.

The rough endoplasmic reticulum (rER) is the first organelle encountered outside of the nucleus (Fig. 3.1). It is composed of connected sacs of membrane and is studded with ribosomes, giving it its characteristic “rough” appearance (Fig. 3.3A). Ribosomes make proteins after binding to messenger RNA (coming from the nucleus). They can be found attached to the rough ER or “free” (not attached) within the cytosol. Ribosomes attached to the rER will make protein that are subsequently transported into the lumen, or hollow inside, of the rER. Here, proteins are folded and modified; those that are modified with carbohydrates (including sugars) are known as glycoproteins, and proteins modified with lipids are termed lipoproteins. At the end of the rER, the proteins bud off in a pod of the rER membrane, known as a vesicle, that is transported to the Golgi complex (Fig. 3.1).

Word Origin: Golgi Complex
Also known as the Golgi apparatus or Golgi body, the Golgi complex was discovered in 1898 by Italian physician Camillo Golgi and named after him.

The Golgi complex is created from flattened sacs of membrane (Fig. 3.3B). The membrane vesicle inbound from the rER fuses with the Golgi to deliver the protein contents to the interior of the Golgi. The Golgi complex functions as a finishing and shipping company: the enzymes contained within it complete the protein modifications that began in the rER, and the proteins are then packaged into vesicles that travel to various locations within or outside the cell.
Some proteins are transported to the plasma membrane and released from the cell, while other proteins become permanently embedded into the plasma membrane.

At the Golgi complex, certain enzymes are packaged into specific vesicles called lysosomes. Lysosome enzymes are able to digest complex biological molecules that are delivered to the lysosome. These molecules can come from outside the cell in endosomes, which will be discussed in Section 3.2, or even from vesicles containing malfunctioning organelles. The 30+ enzymes found in the lysosome function best at a pH of ∼5, which is more acidic than the neutral pH of the cell (∼7.2), reducing the risk to the cell if the lysosome enzymes were to enter the cytosol.

The organelles described above facilitate the creation, modification, packaging, and transport of proteins. Viruses do not have their own organelles, so after gaining entry into the cell, viruses will take advantage of these organelles to manufacture the viral proteins necessary to create more infectious virions.

There are other important parts of the cell that are not directly involved in protein synthesis, and viruses will utilize these components, as well. The majority of cellular respiration, which generates cellular energy in the form of ATP, takes place within the mitochondria (singular: mitochondrion) of the cell (Fig. 3.3C). Viruses do not have their own mitochondria and so will use the ATP generated by the cell. Cells also have a cytoskeleton made of different-sized protein components: microtubules, intermediate filaments, and microfilaments, from largest to smallest diameter. In the same way that a human skeleton shapes the form of the body and provides support for its organs, these cytoskeletal components provide structure for the cell and its organelles (Fig. 3.4). They are also involved in the movement of vesicles within the cell, and the movement of the cells themselves. Some viruses use the cytoskeleton system for transport to different parts of the cell.

3.2 THE PLASMA MEMBRANE, EXOCYTOSIS, AND ENDOCYTOSIS

The plasma membrane is the primary zone of contact between the cell and the extracellular world. As such, this is the first place a virus interacts with a cell.

As mentioned above, the plasma membrane is made of a phospholipid bilayer. The current view of how the membrane is assembled is known as the fluid mosaic model, proposed by Singer and Nicolson. The “mosaic” part of the model refers to the presence of proteins suspended in the membrane bilayer. Many proteins, including glycoproteins, are embedded into the lipid bilayer (Fig. 3.5). Known as integral proteins, these proteins have a variety of functions, including being receptors for extracellular substances or facilitating the adhesion of one cell to another. Peripheral membrane proteins associate closely with the surface of the membrane but are not integrated within it. The “fluid” part of this model refers to the proteins and phospholipid molecules that are noncovalently associated with each other and are therefore not static within the membrane but move around freely. Cholesterol is a lipid that is found in the phospholipid bilayer that helps to maintain the fluidity and movement in the membrane. Cholesterol is also enriched in lipid rafts, portions of the membrane that contain integral proteins involved in transmitting signals to the interior of the cell.

The plasma membrane forms an effective barrier, but certain substances must be transported from one side of the membrane to the other. Certain integral proteins transport ions and small molecules into or out of the cell, but many molecules are too large for these channel or carrier proteins. To address this problem, eukaryotic cells export and import larger molecules by exocytosis and endocytosis. In the process of exocytosis, proteins packaged into secretory vesicles by the Golgi complex travel to the plasma membrane. The secretory vesicles fuse with the plasma membrane, releasing the vesicle contents to the cell exterior (Figs. 3.1 and 3.6A). The vesicle membrane, also composed of a phospholipid bilayer, becomes part of the plasma membrane.
In endocytosis, material from the cell exterior is enclosed in a cavity formed by the plasma membrane, which pinches off to form an endocytic vesicle. There are two broad categories of endocytosis: bulk-phase endocytosis and receptor-mediated endocytosis. In bulk-phase endocytosis, the cell forms a vesicle that engulfs whatever molecules are present in the extracellular fluid, and so the process is nonspecific. On the other hand, receptor-mediated endocytosis is initiated when specific ligands bind to receptors that are present on the cell surface. The cell has receptors on its surface for many biological factors, including growth factors and hormones. The cell imports the ligands by forming a vesicle that includes the membrane area with the receptors (Fig. 3.7A). These endocytic vesicles form in a specific area of the membrane called clathrin-coated pits. Clathrin is a protein that forms a honeycomb-shaped lattice on the intracellular membrane of the endocytic vesicle (Fig. 3.7B). A similar functioning protein, caveolin, forms membrane pits known as caveolae (singular: caveola).

Once inside the cell, the endocytic vesicle soon loses its clathrin or caveolin coating and fuses with a membrane vesicle known as an endosome. The early endosome becomes increasing acidic to form a late endosome, which then fuses with an enzyme-packed lysosome to degrade the contents of the endosome (Fig. 3.1).

**Phagocytosis** is a form of receptor-mediated endocytosis that is used by specialized cells to engulf entire cells. Amoebae use phagocytosis to ingest their prey via phagocytosis. In defense against pathogens, several immune system cells are able to phagocytose whole bacteria and dead cells.

To replicate, viruses must gain entry into a cell. Many viruses enter the cell through receptor-mediated or bulk-phase endocytosis and have mechanisms to escape from endosomes before they fuse with lysosomes. A few viruses are also able to gain entry into the cell via phagocytosis. These viral processes will be explained in detail in Chapter 4, “Virus Replication.”

### 3.3 THE CELL CYCLE

Viruses take advantage of the cell’s transcription and/or translation machinery in the process of virus replication. After gaining entry into a cell, a virus will need to replicate its nucleic acid genome and manufacture viral proteins in order to assemble new infectious virions. Different types of viruses use different aspects of the host cell; the basic cell processes will be discussed here, and the specifics of each virus type will be discussed in the following chapter.

The human genome is composed of over 3 billion nucleotides of DNA, arranged in a double-stranded format where
the phosphate and sugar portions of the nucleotides form the backbone of the strands and the nucleotide bases of one strand bind to the nucleotide bases of the other strand, forming a base pair (Fig. 3.8). Instead of having one long piece of nucleic acid, however, the DNA is broken up into pieces called chromosomes. Human cells are diploid, meaning that each cell has two copies of each chromosome, one passed along in the mother’s egg and the other from the father’s sperm (Fig. 3.9).

The first cell of a human being is the fertilized egg, or zygote, and all the cells that exist within an organism arise from the growth and division of previously existing cells. The cell cycle is the sequential stages through which a cell grows, replicates its DNA, and divides into two daughter cells. The cell cycle is divided into four phases (Fig. 3.10):

1. Gap 1, or G₁: Normal cellular growth occurs. Certain cells, such as neurons, will never continue the cell cycle and enter a stage known as Gap Zero (G₀). Cells that will divide continue to the next phase.
2. Synthesis, or S: The cell creates an additional copy of its chromosomes through DNA replication.
3. Gap 2, or G₂: The cell further enlarges and prepares for cell division.
4. Mitosis, or M: The two sets of chromosomes are separated as the one cell divides into two cells.

The cell cycle stage at which a virus infects a cell can be a crucial determinate of whether infection proceeds within the cell. Certain viruses require cells to be undergoing cell division because the viruses require the enzymes that are present during cell replication in order to replicate...
their own genomes. A number of viruses also interfere with the stages of the cell cycle to increase the efficiency of virus replication.

### 3.4 THE CENTRAL DOGMA OF MOLECULAR BIOLOGY: DNA REPLICATION

DNA replication, which occurs during S phase of the cell cycle, is the first tenet of the **Central Dogma of Molecular Biology**: DNA is replicated in the nucleus to create a copy of the DNA; DNA is transcribed into messenger RNA in the nucleus, and messenger RNA is translated by ribosomes in the cytosol to create a protein (Fig. 3.11).

DNA contains the hereditary information, and RNA is a temporary copy of a DNA gene. Ribosomes create a protein out of amino acids based upon the sequence of nucleotides within the RNA.

Consider the following analogy: you have a desktop computer at home in your bedroom with thousands of files on the hard drive. One of those files is a document that explains how to complete your final class project. The machine and supplies you need to complete your final project are located at school, however. Instead of taking your whole desktop computer with you, you copy the single file onto a USB drive and leave the house. Once you arrive at school, you read the instructions and use the machine and supplies to complete your final project. In this analogy, your hard drive is your DNA that contains thousands of genes, and the temporary copy that left the house (nucleus) is the mRNA. That temporary copy provided the instructions that were used to create your final class project (the protein) with the machine (ribosome) and its supplies (amino acids) found at school (in the cytoplasm).

---

**Word Origin: The Central Dogma: Replication, Transcription, and Translation**

The word *dogma* means a set of accepted principles. The Central Dogma of Molecular Biology is the main set of scientific principles that underlies the field of molecular biology, which deals with DNA, RNA, and proteins.

A *replicate* is an exact copy, and DNA replication is the process of copying DNA. To *transcribe* something is to rewrite it, and transcription is the process of creating a temporary RNA copy of an original DNA sequence. To *translate* something, on the other hand, is to change it from one language to another. Protein translation is the process of translating the language of RNA, made of nucleotides, into the language of proteins, made of amino acids.

The first part of the Central Dogma is DNA replication. The two strands of DNA are **antiparallel**, meaning that they are parallel to each other but going in opposite directions, much like the lanes of a two-way road. The directionality of the strand is determined by the position of the sugar deoxyribose in the nucleotide. The carbons within the nucleotide base are numbered, and the carbons within the sugar are also numbered but each number is followed by a prime symbol (similar to an apostrophe) to distinguish the sugar carbons from the carbon in
the base (Fig. 3.12A). In a growing strand of nucleic acid, the phosphate group of the nucleotide attaches to the sugar at the 5′ (pronounced “five prime”) carbon atom, and a new nucleotide is added to the 3′ (pronounced “three prime”) carbon of the sugar. This “forward” direction is referred to as 5′→3′ (“five prime to three prime”). All replication of DNA occurs in this forward direction.

Since the two strands of a DNA molecule are antiparallel, if one strand is going forward (5′→3′) left-to-right, then the other strand is going forward (5′→3′) from right-to-left. As such, the 5′ end of one strand is matched with the 3′ end of the other strand (Fig. 3.12B).

DNA replication occurs during the S phase of the cell cycle, when the chromosomes are replicated. During the process of DNA replication, cellular enzymes unwind the DNA molecule and separate the two DNA strands from each other. DNA replication is semiconservative: each current strand of DNA functions as a template for a new strand, and so a copied piece of DNA will be composed of one old and one new strand (Fig. 3.13A). After the two strands of the DNA separate, DNA polymerase is the enzyme that lays down the complementary nucleotides of the new strand of DNA, always in the 5′→3′ direction (Fig. 3.13B). Note that since the two strands of DNA are antiparallel, the old strand is read 3′→5′ while the new strand grows 5′→3′. DNA polymerase adds new nucleotides based upon the complementary base pair rules discussed in Chapter 1, “The World of Viruses,” and shown in Fig. 3.12A: adenine bonds with thymine, and cytosine bonds with guanine. Cellular DNA polymerases are DNA-dependent DNA polymerases because they synthesize DNA using a DNA template.

DNA polymerases have high fidelity, meaning that they do not often place an incorrect base in the growing strand of replicating DNA. They also have proofreading ability: in the same way you may type an incorrect letter on a keyboard and hit the “Backspace” key to replace it with the correct letter, DNA polymerase can reverse and replace an incorrectly placed nucleotide. DNA polymerase and repair enzymes can also cut out a section around an incorrect nucleotide and replace the section of DNA with the correct nucleotides. Taken together, DNA polymerase makes one mistake for every 1 million nucleotides copied, on average.

Several other proteins and enzymes are involved in DNA replication. For instance, DNA polymerase cannot bind to a single-stranded portion of DNA, so when the two existing DNA strands are separated, an enzyme known as primase lays down a short complementary RNA fragment onto the DNA strand, creating a double-stranded portion to which DNA polymerase can bind. Other enzymes are also required for the process of replication: since DNA polymerase can only add to a nucleotide chain in the 5′→3′ direction, it can only create short fragments of DNA on one strand of the replicating DNA (known as the lagging strand), until the double-stranded DNA opens farther down the strand. The enzyme
ligase joins together these short fragments (known as Okazaki fragments) to create a contiguous DNA strand.

Several DNA viruses take advantage of cellular DNA polymerase and replication enzymes to replicate their genomes. Because DNA replication takes place within the nucleus, these viruses must gain entry into the nucleus to replicate their genomic DNA.

3.5 THE CENTRAL DOGMA OF MOLECULAR BIOLOGY: RNA TRANSCRIPTION AND PROCESSING

Sections of DNA called genes encode the information needed to create proteins. There are over 20,000 protein-encoding genes within the 46 chromosomes that constitute the human genome. There are three steps in the process of generating a protein from the information stored within DNA: transcription, RNA processing, and translation.

DNA replication occurs in the nucleus because DNA is located in the nucleus, and transcription, the process of creating a temporary RNA copy of the DNA, also occurs in the nucleus for the same reason (Fig. 3.14). A complex of transcription factor proteins binds the DNA immediately upstream of the gene start site at a location called a promoter. RNA polymerase II then associates with the transcription factors and the DNA (Fig. 3.15A). Transcription factors bind to specific sequences of DNA within the promoter region, ensuring that transcription of the DNA begins at the correct location. Other transcription factors can bind to enhancer regions that, as their
name suggests, can increase the rate of transcription. Unlike the promoter, enhancer regions can be thousands of nucleotides away, either upstream or downstream from the gene start site.

Cellular RNA polymerases are DNA-dependent RNA polymerases because they synthesize RNA based on a DNA template. Much in the same way that DNA polymerase uses a strand of DNA to create the complimentary strand, RNA polymerase uses the DNA template to create a strand of RNA, adding nucleotides in the 5′→3′ direction using the same complementary base pairing rules as DNA replication, except that the base uracil substitutes for thymine (Figs. 3.12B and 3.15B). Because only one of the two DNA strands, known as the template strand or antisense strand, acts as the template for RNA polymerase, the resulting RNA is single-stranded (Figs. 3.12B and 3.15B). RNA polymerase terminates transcription when it reaches a consensus sequence at the end of the gene. At this point, the RNA transcript is known as precursor messenger RNA (mRNA). It is termed “messenger RNA” because it is the message, encoded within the DNA, of how to create a specific protein.

RNA polymerases do not have as high fidelity as DNA polymerases and place an incorrect base on average once per 100,000 nucleotides transcribed, 10 times more often than DNA polymerase. These RNA polymerases are DNA-dependent RNA polymerases. Eukaryotic cells do not contain RNA-dependent RNA polymerases for the creation of mRNA, and so several types of RNA viruses encode their own RNA polymerases, with error rates of 1 in 100 to 1 in
100,000 nucleotides. A high mutation rate is the result of the low fidelity of several RNA viruses that encode their own RNA polymerase.

Following transcription, the precursor mRNA undergoes RNA processing, also known as posttranscriptional modification, to convert the precursor mRNA into mature mRNA. The first modification, which occurs while RNA polymerase is still transcribing the mRNA transcript, is the addition of a “cap” to the 5′-end of the transcript (Fig. 3.16). The 5′-cap consists of a methylated guanine nucleotide (known as 7-methylguanosine, m7G) that protects the 5′-end of the RNA transcript. Ribosomes will also bind to the 5′-cap to begin translation. The second modification is the addition of a 3′ poly(A) tail. The “tail” consists of 50–250 adenine nucleotides added to the 3′-end of the mRNA to protect the mRNA transcript. The final modification is the removal of introns by a process known as RNA splicing. Within most eukaryotic mRNA transcripts, there are sequences of mRNA that will not be translated into proteins. These sequences, known as introns, are removed during posttranscriptional modification, leaving behind the exons or coding sequences (Fig. 3.16). More than one protein can be created from a single mRNA through RNA splicing because different introns can be removed from an mRNA transcript, resulting in different RNA sequences and subsequently, different proteins. This process is known as alternative splicing. The mRNAs of some viruses, including HIV, also undergo alternative RNA splicing.

3.6 THE GENETIC CODE

Now processed, the mature mRNA transcript leaves the nucleus and is delivered to the ribosome, which is located in the cytosol. The ribosome acts as a protein factory, and the mature mRNA functions as the instructions for manufacturing. Proteins are made of amino acids, and most human proteins are 50–1000 amino acids in size. There are 20 different amino acids, and the sequence of mRNA determines the order in which the ribosome will assemble the amino acids into a protein.

The ribosome initially moves down the transcript one base at a time, reading the sequence in three-base words known as codons (Fig. 3.17). The ribosome starts translation, the assembly of a protein out of amino acids, when it encounters the start codon in the mRNA, which is the sequence AUG. The AUG codon is usually within the context of a slightly larger sequence, called the Kozak consensus sequence, which generally has the sequence GCCAACAUGG (the underlined adenine can also be a guanine). AUG codes for the amino acid methionine, and so all protein translation begins with methionine.

The start codon sets the reading frame: instead of continuing to move down the mRNA transcript one base at a

---

**Study Break**

Pertaining to DNA and RNA architecture, explain what “five prime” and “three prime” mean and what these phrases have to do with DNA replication, transcription, and RNA processing.

---

**FIGURE 3.16** RNA processing. The precursor mRNA, produced by transcription, undergoes RNA processing while still within the nucleus. A 7-methylguanosine cap is added to the 5′ end of the transcript, while 50–250 adenine nucleotides are added to the 3′ end of the transcript. Introns are removed to produce a mature mRNA molecule that leaves the nucleus and is translated by ribosomes in the cytosol. The removal of different introns results in an mRNA transcript that is translated into a different protein.

**FIGURE 3.17** The flow of data, from DNA to protein. The antisense or template strand of DNA acts as a template to transcribe mRNA. The ribosome reads the mRNA in three nucleotide codons, beginning with the start codon, AUG, which codes for the amino acid methionine. The order of the bases within the codons determines which amino acid will be added to the growing protein by the ribosome.
time, the ribosome now reads the mRNA codons consecutively, three bases at a time (Fig. 3.18). The sequence of the triplet codon determines which amino acid is added next to the growing protein. When the ribosome reaches a stop codon, it falls off the mRNA, and the protein is complete. There are three variations of the stop codon: UGA, UAA, and UAG. The segment of mRNA before this starting point is not translated and is known as the 5′ untranslated region (5′ UTR) (Fig. 3.18B). Any mRNA past the stop codon will not be translated; this region is known as the 3′ UTR. The sequence from the start codon to the stop codon is known as an open reading frame because it is translatable.

**Genetic Code Analogy**

After binding to the mRNA, the ribosome begins translation at the start codon, AUG, and then moves down the mRNA transcript one codon (three nucleotides) at a time until it reaches a stop codon. Try finding the translated codons in the following sequence. The start codon—THE—will set the reading frame. The three stop codons are OKK, OOK, and OKO.

**FIGURE 3.19** The genetic code. The sequence of amino acids within a protein is determined by the nucleotide sequence of the mRNA. To use the table, find the first base of the codon in the leftmost column. Next, find the second base of the codon on the top row. The intersection of the column and row will be the target box in which the codon is located. Next, find the third base of the codon in the rightmost column to identify on which line of the target box the codon is located. Next to the codon sequence in the target box is the amino acid that corresponds to the codon. The list of amino acid abbreviations is located below the table. AUG, as the start codon, is in green and codes for methionine. The three stop codons are UAA, UAG, and UGA. Stop codons encode a release factor, rather than an amino acid, that causes translation to cease.

Based on the four nucleotides in RNA—adenine, guanine, cytosine, and uracil—there are 64 possible different 3-letter permutations (Fig. 3.19). There are only 20 amino acids that can be translated into proteins.
acids, however, and so some of the codons are redundant, meaning that two or more codons encode the same amino acid. There are three stop codons, which end translation and do not encode any amino acid.

Many scientists worked to decipher the genetic code. Robert W. Holley, Har Gobind Khorana, and Marshall W. Nirenberg shared a Nobel Prize in physiology or medicine in 1968 for their work in determining the “key” to deciphering the genetic code. The Table in Fig. 3.19 reveals which amino acids are encoded by each codon. The code is universal: all living things have the same 20 amino acids that are encoded by these codons, indicating that this system originated very early in the development of life and has been evolutionarily conserved over time. Being that viruses take advantage of the host translational machinery and ribosomes, viral mRNAs use these same codons to encode the same amino acids in their proteins as do living things.

### Study Break
Translate the two mRNA sequences found in Fig. 3.18 A and B.

### 3.7 THE CENTRAL DOGMA OF MOLECULAR BIOLOGY: PROTEIN TRANSLATION

Three major components are required for translation to occur: mRNA, the ribosome, and transfer RNAs (tRNAs). The ribosome is an organelle with two subunits—a small and large subunit (Fig. 3.20A)—that are made of another type of RNA, termed ribosomal RNA (rRNA), and over 50 proteins. Transfer RNAs are also made of RNA (Fig. 3.20B) and act as adaptors between the mRNA and the ribosome. It is the tRNA that brings amino acids to the ribosome so they may be joined together into a protein. Within a tRNA is an anticodon sequence that is complementary to the mRNA codon, and at the 3’ end of the tRNA is attached an amino acid. The mRNA codon sequence binds to the anticodon sequence within the tRNA, which has attached a specific amino acid (Fig. 3.20A). It is this adaptor molecule that actually determines which codon encodes which amino acid.

**Refresher: RNA**
- mRNA: messenger RNA. The temporary copy of DNA that will be translated by the ribosome.
- rRNA: ribosomal RNA. The ribosome is made of rRNA and proteins.
- tRNA: transfer RNA. Acting as an adaptor, it transports the amino acid to the ribosome and has an anticodon region that binds the mRNA codon.

There are three stages of translation (Fig. 3.21):

1. **Initiation**: the start of translation. The ribosome small unit, containing the tRNA holding methionine, binds at the 5’-cap of the processed mRNA molecule. It scans the mRNA until the start codon, AUG, is encountered.
In the same way that transcription factors were necessary for RNA polymerase II to bind to the DNA gene to be transcribed, an assortment of translation **initiation factors** assists in recruiting the ribosome and the first tRNA to the mRNA transcript. The large ribosomal subunit joins the small subunit.

2. **Elongation**: the synthesis of the protein out of amino acids. The ribosome moves along the mRNA strand. For each codon, a tRNA with a complementary anticodon enters the ribosome, delivering the corresponding amino acid. The ribosome joins the growing amino acid strand to the new amino acid. It continues moving along the mRNA, one codon at a time, and a tRNA containing the corresponding amino acid enters the ribosome for each codon. The new amino acid is joined to the previous amino acids, elongating the amino acid chain.
3. **Termination**: the end of translation. When a stop codon in the mRNA is encountered by the ribosome, a release factor enters the ribosome and translation ceases. The now completed protein is released, and the ribosome falls off the mRNA.

As described above, many of these proteins undergo **posttranslational modification** in the rER and Golgi complex to add lipid or sugar residues to the molecule.

Eukaryotic mRNA is **monocistronic**, meaning that any mRNA transcript codes for only one protein. All viruses are dependent upon their host cells for the translation of their proteins, so viral mRNAs must conform to the biological constraints of the host cell machinery. Viruses have several tactics, however, to ensure the preferential transcription and translation of their viral mRNAs and proteins over those of the host.

---

### In-Depth Look: Peptides, Polypeptides, and Proteins

A peptide is a short chain of amino acids. A long peptide chain is known as a polypeptide. When a ribosome translates mRNA, it creates a polypeptide of amino acids.

So what is the difference between a polypeptide and a protein? A protein is a complete, functional entity. Some proteins are made of only one polypeptide chain, but other proteins are made of more than one polypeptide chain. Hemoglobin, the protein that transports oxygen in our red blood cells, is a protein composed of four polypeptide chains in total. While discussing protein translation in this chapter, we assume that the polypeptide created by the ribosome will be a functional protein, but keep in mind that many proteins are composed of more than one polypeptide chain.

---

3.8 **PROMOTION OF VIRAL TRANSCRIPTION AND TRANSLATION PROCESSES**

Viruses have evolved several mechanisms to ensure the successful transcription and translation of their gene products, necessary to create more infectious virions. Some viruses take advantage of the host splicing machinery to produce several mRNA transcripts from one precursor mRNA. HIV-1, for instance, produces most of its mRNAs from alternative splicing. Some viruses, like influenza, can snatch the 5′-caps from host mRNAs to gain the necessary cap for the viral mRNA, leaving the host mRNA untranslatable without a 5′-cap. Some viruses also create mRNAs that are translated into one long **polyprotein** that is then cleaved into several viral proteins after translation.

Other viruses have evolved tactics for protein translation that are not customarily used by the host. Ribosomes recognize and bind to the 5′-cap of mRNAs, but some viral mRNAs contain **internal ribosome entry sites** (IRES) that allow ribosomes to bind within the mRNA sequence, without a 5′-cap (Fig. 3.22A and B). IRES are used to initiate translation in internal sections of viral mRNA that are in a different reading frame, or when the virus has interfered with the normal translation process to inhibit host protein synthesis. **Ribosomal frameshifting** occurs when a ribosome pauses or meets a “slippery sequence” within a piece of mRNA. Instead of continuing to read the codons in frame, the ribosome encounters a problem and moves backward or forward one nucleotide. The result is that the ribosome begins translating a different reading frame than it was previously, producing a different mRNA and protein. Viruses like HIV utilize ribosomal frameshifting to encode several proteins within just one portion of DNA. A similar mechanism occurs with **termination suppression**, in which a stop codon is suppressed and the ribosome continues translating the mRNA, creating a polyprotein. **Ribosomal skipping** occurs while the ribosome is translating a viral mRNA. Viral proteins prevent the ribosome from joining a new amino acid to the growing protein chain, which releases the protein from the ribosome. Having not encountered a stop codon, however, the ribosome continues translating the remainder of the viral mRNA. The effect is that several viral proteins can be synthesized with only one piece of mRNA. **Leaky scanning** happens when a host ribosome encounters a start codon (AUG) within a piece of viral mRNA, but the Kozak consensus sequence (the nucleotides surrounding the start codon) is not in a favourable configuration. The ribosome may begin translation at this site, but the next ribosome that binds to the viral mRNA may continue past this weak start codon to begin translation at the next AUG encountered (Fig. 3.22C). The result is that one viral mRNA can encode two viral proteins. All of these transcription and translation processes promote the synthesis of viral proteins, often faster and with less energy than normal cellular mechanisms.

In addition to evolving mechanisms to ensure their mRNA is processed and recognized by host ribosomes, viruses have also evolved ways to interfere with the transcription and translation of host proteins:

1. **Some viruses can interfere with the host’s RNA polymerase II.** They can do so by interfering with transcription factors, by preventing the activation of the enzyme, or by causing the breakdown of the enzyme. Several herpesviruses have strategies to interfere with RNA polymerase II.

2. **Viruses can interfere with the processing of precursor RNA.** HIV-1 protein Vpr inhibits the splicing of host mRNA, and influenza inhibits the addition of a poly(A) tail to the host’s mRNA transcripts.

3. **Many viruses interfere with the export of the mRNA transcript from the nucleus.** They do so by interfering with or causing the breakdown of the proteins that export the processed mRNA from the nucleus.

4. **Certain viral proteins can cause the degradation of host mRNA.** For instance, severe acute respiratory syndrome (SARS) is caused by a coronavirus named SARS-CoV,
which has a protein named nsp1 that induces the breakdown of host mRNAs. Interestingly, the SARS-CoV mRNA transcripts are protected from degradation.

5. Several viral proteins prevent translation of host mRNA. This can happen by interfering with host translation initiation factors or by removing the 5' caps from host mRNAs.

The host replication, transcription, and translation machinery is complex and involves a multitude of enzymes and molecules. Viruses must conform to the limitations of the host cell in order to replicate, but they have evolved many strategies to ensure the preferential transcription of their mRNA transcripts and efficient translation of viral proteins.
SUMMARY OF KEY CONCEPTS

Section 3.1 The Basic Organization of the Cell

- Each cell of a eukaryote has a central nucleus that separates its DNA from the rest of the cell. A prokaryote does not have a nucleus. All organisms within the domain Eukarya, including humans, have cells with nuclei, while those within Bacteria and Archaea are prokaryotes.

- Many organelles ensure the efficient functioning of the cell. Ribosomes make proteins and can be free within the cytosol or attached to the rER, which folds and modifies proteins into glycoproteins or lipoproteins after they have been made by the ribosome. These are shipped in vesicles to the Golgi complex, which packages them in vesicles to be delivered to the various parts of the cell or the extracellular space.

- Lysosomes are vesicles that are filled with enzymes that can digest complex biological molecules. These organelles break down nonfunctional organelles or material coming from outside the cell.

- Mitochondria are the powerhouses of the cell: they generate ATP. The cytoskeleton of the cell is made of protein components that lend support to the cell and its organelles. They are also involved in cell and organelle movement.

- Viruses use cell-generated ATP, take advantage of cellular organelles to manufacture viral proteins, and use the cytoskeleton for transport to different parts of the cell.

Section 3.2 The Plasma Membrane, Exocytosis, and Endocytosis

- The plasma membrane is composed of phospholipids, which have a polar head and nonpolar tails. As such, they are amphipathic: the phosphate head is hydrophilic and the fatty acid tails are hydrophobic.

- The fluid mosaic model describes the current thinking on the plasma membrane. This model states that the phospholipids of the membrane move freely within the membrane, as do the many integral proteins embedded in the membrane.

- The process of secreting large molecules, like proteins, from the cell is known as exocytosis. Endocytosis imports large molecules into the cell. Bulk-phase endocytosis forms a vesicle that encapsulates the extracellular fluid, while receptor-mediated endocytosis is initiated when ligands bind to receptors on the cell surface. The ligands are imported in vesicles that form in clathrin-coated pits.

- Endocytic vesicles soon fuse with endosomes that increase their acidity as they travel into the cell.

- In order to infect a cell, viruses must get through the plasma membrane. Some viruses that enter the cell through endocytosis must also escape from endosomes.

Section 3.3 The Cell Cycle

- Human DNA is diploid and divided into chromosomes.

- The cell cycle is the sequential stages through which a cell grows, replicates its DNA, and divides in two through the process of mitosis.

- The four phases of the cell cycle are Gap 1, Synthesis, Gap 2, and Mitosis.

- Some viruses require cells to be undergoing cell division because they require enzymes present during mitosis. Many viruses interfere with the stages of the cell cycle to increase the efficiency of viral replication.

Section 3.4 The Central Dogma of Molecular Biology: DNA Replication

- The Central Dogma of Molecular Biology states that DNA is replicated to create more DNA, DNA is transcribed into mRNA, and mRNA is translated by ribosomes to create proteins.

- DNA is composed of four different nucleotides (with bases adenine, cytosine, guanine, and thymine) bonded together. It is double stranded; each strand has directionality and the “forward” direction is termed 5′→3′. Because DNA is antiparallel, the 5′ end of one strand is matched with the 3′ end of the other strand.

- DNA replication is semiconservative, so each old strand acts as a template for the new DNA strand. DNA polymerase is the enzyme that reads the old strand in the 3′→5′ direction and creates the new strand out of nucleotides in the 5′→3′ direction. Nucleotides are added according to complementary base pair rules.

- DNA polymerases have high fidelity and make one error in every 1 million nucleotides added, on average.

- A few viral families take advantage of cellular DNA polymerases to replicate their DNA genomes.

Section 3.5 The Central Dogma of Molecular Biology: RNA Transcription and Processing

- The process of creating an mRNA copy of a portion of DNA is known as transcription. RNA polymerase II binds to transcription factors that assemble on the promoter of the gene, and the enzyme joins RNA nucleotides (adenine, guanine, cytosine, and uracil) in the 5′→3′ direction to create the precursor mRNA transcript.

- RNA polymerases have lower fidelity than DNA polymerases.

- Precursor mRNA is processed before leaving the nucleus. The transcript receives a 5′ 7-methylguanosine cap and a 3′ poly(A) tail, and introns are removed via RNA splicing.

Section 3.6 The Genetic Code

- The “genetic code” refers to which amino acids correspond to a sequence of processed mRNA.

- The eukaryotic ribosome translates an mRNA transcript in the 5′→3′ direction and begins at the start codon, AUG. This codon is found within a larger sequence
known as the Kozak consensus sequence. Translation will continue until a stop codon (UGA, UAA, or UAG) is reached.

- The region of mRNA that is translated is known as the open reading frame. The untranslated parts of the sequence are known as the 5′ UTR and 3′ UTR.
- There are 20 amino acids but 64 possible codons. Some codons are redundant and code for the same amino acid.

**Section 3.7 The Central Dogma of Molecular Biology: Protein Translation**

- Within the ribosome, tRNAs act as adaptor molecules because their anticodon sequence recognizes the mRNA transcript. The corresponding amino acid is attached to the 3′ end of the tRNA molecule.
- There are three stages of translation: initiation, elongation, and termination.
- During initiation, translation initiation factors recruit the ribosome and tRNA to the mRNA transcript. The ribosome scans the mRNA until the start codon is encountered.
- During elongation, the ribosome moves along the mRNA. For each new codon, a tRNA carrying an amino acid enters the ribosome. The growing amino acid chain is bonded to the new amino acid.
- During termination, the ribosome encounters a stop codon and translation ends. The protein is complete.

**Section 3.8 Promotion of Viral Transcription and Translation Processes**

- Viruses have evolved many tactics to take advantage of the cellular transcription and translation machinery. Viruses use cellular transcription, RNA processing, and translation mechanisms to ensure the translation of their proteins. They also interfere with the transcription, RNA processing, and translation of host gene products to ensure the preferential translation of viral products.

**FLASH CARD VOCABULARY**

| Nucleus                  | Amphipathic         |
|--------------------------|----------------------|
| Prokaryote                | Rough endoplasmic reticulum |
| Eukaryote                | Ribosome             |
| Organelle                | Glycoprotein         |
| Phospholipid             | Lipoprotein          |
| Hydrophilic              | Vesicle              |
| Hydrophobic              | Golgi complex        |
| Lysosome                 | Semiconservative replication |
| Mitochondrion            | DNA polymerase       |

| Cytoskeleton             | Fidelity             |
|--------------------------|----------------------|
| Fluid mosaic model       | Primase              |
| Integral protein         | Ligase               |
| Peripheral protein       | Okazaki fragments    |
| Exocytosis               | Transcription factors|
| Endocytosis              | Promoter             |
| Clathrin-coated pit      | Enhancer             |
| Endosome                 | RNA polymerase II    |
| Base pair                | Transcript           |
| Chromosome               | RNA processing       |
| Diploid                  | 5′-methylguanosine cap |
| Zygote                   | 3′-Poly(A) tail      |
| Cell cycle               | RNA splicing         |
| Daughter cells           | Intron               |
| Mitosis                  | Exon                 |
| Central Dogma of Molecular Biology | Alternative splicing |
| DNA replication          | Codon                |
| Transcription            | Start codon          |
| Translation              | Stop codon           |
| Antiparallel             | Kozak consensus sequence |
| Reading frame            | Termination          |
| Open reading frame       | Translation initiation factors |
| 5′-Untranslated region   | Posttranslational modification |
| 3′-Untranslated region   | Monocistronic        |
| Redundant codons         | Polypeptide          |
| Genetic code             | Internal ribosome entry site |
| Transfer RNA             | Ribosomal frameshifting |
| Anticodon                | Ribosomal skipping   |
| Initiation               | Leaky scanning       |

**CHAPTER REVIEW QUESTIONS**

1. Describe the general process of expressing a gene (in a chromosome) into a protein. Where does each step take place within the cell?

2. A secreted protein has been synthesized by a ribosome. Describe the pathway it will take to leave the cell and what happens at each step.

3. What molecules do you know that are hydrophobic or hydrophilic?
4. Explain the fluid mosaic model of plasma membrane assembly.

5. Which cellular organelles or processes are utilized by viruses?

6. Describe what happens during each of the four stages of the cell cycle.

7. What is the Central Dogma of Molecular Biology?

8. Draw a double-stranded piece of DNA. Make sure to label the 5′- and 3′-ends. Now draw out the process of DNA replication, paying attention to the 5′- and 3′-ends and the direction that DNA Polymerase lays down the new strand.

9. Describe the three steps involved in RNA processing.

10. Use the genetic code in Fig. 3.19 to translate the following piece of mRNA: 5′- GCCGCCAUGGCAU AGCCGAUGACCCGGA -3′

11. Determine the 5′-UTR and 3′-UTR in the sequence above.

12. Describe what happens during the three stages of translation.

13. Explain at least three translational processes involving the ribosome that occur with viral translation but do not normally occur with cellular translation of a protein.

14. How do viruses ensure the preferential translation of their gene products over cellular gene products?

FURTHER READING

De Jesus, N., 2007. Epidemics to eradication: the modern history of poliomyelitis. Virol. J. 4, 70.

Fotin, A., Cheng, Y., Sliz, P., et al., 2004. Molecular model for a complete clathrin lattice from electron cryomicroscopy. Nature 432, 573–579.

Nirenberg, M., 2004. Historical review: deciphering the genetic code - a personal account. Trends Biochem. Sci. 29, 46–54.

Tarini, M., Cignoni, P., Montani, C., 2006. Ambient occlusion and edge cueing to enhance real time molecular visualization. IEEE Trans. Vis. Comput. Graph 12, 1237–1244.

Vereb, G., Szöllosi, J., Matkó, J., et al., 2003. Dynamic, yet structured: the cell membrane three decades after the Singer-Nicolson model. Proc. Natl. Acad. Sci. U.S.A. 100, 8053–8058.

Watson, J.D., Crick, F.H.C., 1953. Molecular structure of nucleic acids. A structure for deoxyribose nucleic acid. Nature 171, 737–738.