14 FROM DNA TO PROTEIN

Ricin and Your Ribosomes

In 2003, police acted on an intelligence tip and stormed a London apartment, where they found laboratory glassware and castor oil beans (Figure 14.1). They arrested a few young men and reminded the world that unconscionable people still view ricin as a bioweapon.

The castor oil plant (*Ricinus communis*) has ricin in all of its tissues, but its oil is valued as an ingredient in many plastics, paints, cosmetics, textiles, and adhesives. The oil —and ricin—is most concentrated in the seeds (beans), but the ricin is discarded when the oil is extracted.

A dose of ricin as small as a grain of salt can kill you; only plutonium and botulism toxin are more deadly. Researchers knew about ricin's lethal effects as long ago as 1888. During World War I, when deadly chlorine and mustard gases were wafting across battlefields, England and the United States investigated ricin's potential use as a weapon. Both countries shelved the research when the war ended.

Now fast-forward to 1969, at the height of the Cold War between Russia and the West. Georgi Markov, a Bulgarian writer, had defected to England. As he strolled down a busy London street, an assassin jammed the tip of a modified umbrella into one of Markov's legs. The tip held a tiny ball laced with ricin. Markov died in agony three days later.

Ricin is on stage once again. In 2004, traces were found in a United States Senate mailroom and State Department building, and in an envelope addressed to the White House. In 2005, the FBI arrested a man who had castor oil beans, substances consistent with ricin production, and an AK-47 stashed in a home in Florida.

How does ricin exert its deadly effects? *It inactivates ribosomes, the protein-building machinery of all cells.*

Ricin is a protein with two polypeptide chains. One chain helps ricin insert itself into cells. The other chain serves as an enzyme. Its catalytic action wrecks part of the ribosome where amino acids are assembled into proteins. It yanks adenine subunits out of an RNA molecule that is a crucial component of the ribosome's three-dimensional structure. Once that happens, the ribosome's shape unravels, protein synthesis stops, and cells spiral toward death. So does the individual. There is no antidote.

You can go about your business without ever knowing what a ribosome is or what it does. However, you also can recognize that protein synthesis is not a topic invented to torture biology students. It is something worth knowing about and appreciating for how it keeps us alive—and for appreciating anti-terrorism researchers who are working to keep us that way.

Figure 14.1 *Left*, castor oil plant seeds, source of the ribosome-busting ricin. *Right*, model for one of ricin's two polypeptide chains. This chain helps ricin penetrate living cells. The other one destroys the capacity for protein synthesis, and for life.

IMPACTS, ISSUES

Start with what you already know about DNA, the book of protein-building information in cells. The alphabet used to write the book seems simple enough—just A, T, G, and C, for the four nucleotide bases adenine, thymine, guanine, and cytosine. But how do you get from an alphabet to a "word"— a protein? The answer starts with the order, or sequence, of the four nucleotide bases in a DNA molecule.

As you know, when a cell replicates its DNA, the two nucleotide strands of the DNA double helix unwind from each other completely. At other times, however, enzymes selectively unwind the two strands in certain regions, which exposes the base sequences of genes. Most genes encode information about specific proteins.

It takes two big steps, **transcription** and **translation**, to get from the sequence of nucleotide bases in genes to the sequence of amino acids in a protein. In eukaryotic cells, the first step occurs inside the nucleus. A newly exposed DNA base sequence functions as a structural pattern, or a template, for making a strand of ribonucleic acid (RNA) from the cell's pool of free ribonucleotides.

The RNA moves into the cytoplasm, where it becomes translated. In this second step of protein synthesis, the RNA guides the assembly of amino acids into a new polypeptide chain. These are the chains that twist and fold into the three-dimensional shapes of proteins.

In short, RNA is transcribed on DNA templates, then RNA is translated into proteins:

DNA —*transcription*→ RNA —*translation*→ PROTEIN

How Would You Vote?

Ricin is difficult to disperse through the air and is unlikely to be used in a large-scale terrorist attack. However, ricin powder did turn up in a Senate office building. Scientists are working to develop a vaccine against ricin. If mass immunizations were to be offered, would you sign up to be vaccinated? See BiologyNow for details, then vote online.

Key Concepts

INTRODUCTION

Life depends on enzymes and other proteins. All proteins consist of polypeptide chains. The chains are sequences of amino acids that correspond to genes—sequences of nucleotide bases in DNA. The path leading from genes to proteins has two steps: transcription and translation.

TRANSCRIPTION

During transcription, the two strands of the DNA double helix are unwound in a gene region. Exposed bases of one strand become the template for assembling a single strand of RNA. Only one type of RNA transcript encodes the message that gets translated into protein. It is called messenger RNA. Section 14.1

CODE WORDS IN THE TRANSCRIPTS

The nucleotide sequence in DNA is read three bases at a time. Sixty-four base triplets correspond to specific amino acids and represent the genetic code.

The code words have been highly conserved through time. Only a few simple eukaryotes, prokaryotes, and prokaryote-derived organelles have slight variations on the genetic code. Section 14.2

TRANSLATION

During translation, amino acids are bonded together into a polypeptide chain in a sequence specified by base triplets in messenger RNA. Transfer RNA delivers amino acids one at a time to ribosomes. An RNA component of ribosomes catalyzes the chain-building reaction. Sections 14.3, 14.4

MUTATIONS IN THE CODE WORDS

Gene mutations introduce changes in protein structure, protein function, or both. The changes may lead to small or large variation in the shared traits that characterize individuals of a population. Section 14.5

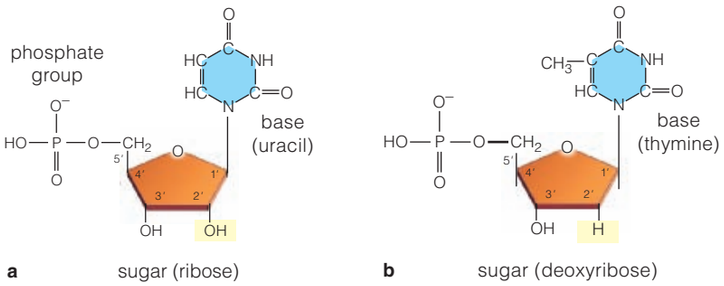
Links to Earlier Concepts

Once again you will meet up with the nucleic acids DNA and RNA (Section 3.7). Gene transcription has features in common with DNA replication, so you may wish to review Section 13.3 before you start. You will again consider how protein primary structure emerges (3.5), this time in the context of RNA interactions. The last section of this chapter will expand your knowledge of DNA repair mechanisms (13.3) and gene mutation (1.4, 3.6, 12.3, 12.8).

14.1 How Is RNA Transcribed From DNA?

LINKS TO SECTIONS 3.7, 13.2, 13.3

*In **transcription,** the first step in protein synthesis, a sequence of nucleotide bases is exposed in an unwound region of a DNA strand. That sequence is the template upon which a single strand of RNA is assembled from adenine, cytosine, guanine, and uracil subunits.*



a **b**

c Example of base pairing between a DNA strand and a new RNA strand assembled on it during *transcription*:

DNA template

RNA transcript

d Example of base pairing between an old DNA strand and a new strand forming on it during *DNA replication*:

DNA template

new DNA strand

Figure 14.2 (**a**) Uracil, one of four ribonucleotides in RNA. The other three—adenine, guanine, and cytosine—differ only in their bases. Uracil compared with (**b**) thymine, a DNA nucleotide. (**c**) Base pairing of DNA with RNA during transcription, compared with (**d**) base pairing during DNA replication.

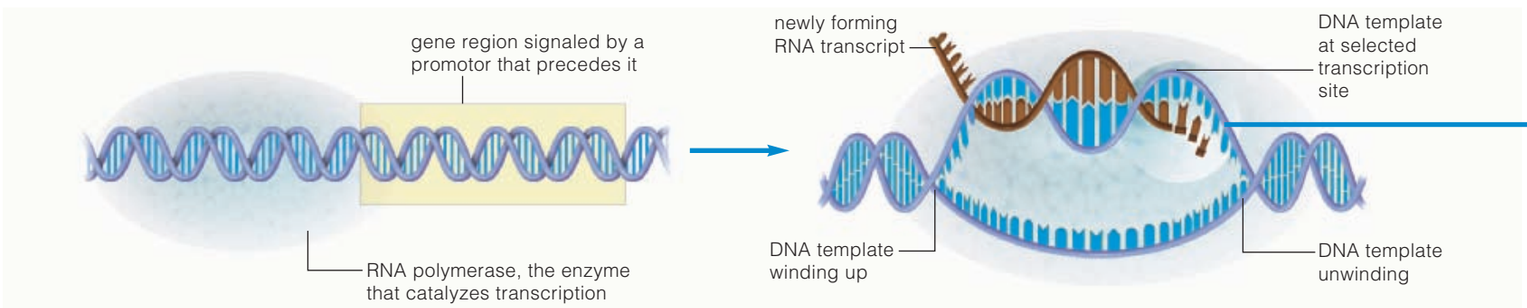
The chapter introduction may have left you with the impression that protein synthesis requires one class of RNA molecules. It actually requires three. When genes that specify proteins are transcribed, the outcome is **messenger RNA** (mRNA). *This is the only class of RNA that carries the protein-building codes.* **Ribosomal RNA** (rRNA) and **transfer RNA** (tRNA) are transcribed from different genes. The rRNA becomes a component of ribosomes, the structures in which polypeptide chains are assembled. The tRNA delivers amino acids one by one to ribosomes in the order specified by mRNA.

THE NATURE OF TRANSCRIPTION

An RNA molecule is almost but not quite like a single strand of DNA. It has four kinds of ribonucleotides, each with the five-carbon sugar ribose, one phosphate group, and one base. Three bases—adenine, cytosine, and guanine—are the same as those in DNA. In RNA, though, the fourth base is **uracil**, not thymine. Uracil, too, can pair with adenine, which means that a new RNA strand can base-pair with a DNA strand. Figure 14.2 is a simple way to think about this pairing.

Transcription *differs* from DNA replication in three respects. Only part of one DNA strand, not the whole molecule, is unwound and used as the template. The enzyme **RNA polymerase**, not DNA polymerase, adds ribonucleotides one at a time to the end of a growing strand of RNA. Also, transcription results in one free RNA strand, not a hydrogen-bonded double helix.

DNA contains many protein-coding regions. Each is transcribed separately, and each has its own START



a RNA polymerase initiates transcription at a promoter in DNA. After binding to a promoter, RNA polymerases recognize a base sequence in DNA as a template for making a strand of RNA from free ribonucleotides, which have the bases adenine, cytosine, guanine, and uracil.

b All through transcription, the DNA double helix becomes unwound in front of the RNA polymerase. Short lengths of the newly forming RNA strand briefly wind up with its DNA template strand. New stretches of RNA unwind from the template (and the two DNA strands wind up again).

Figure 14.3 ***Animated!*** Gene transcription. By this process, an RNA molecule is assembled on a DNA template. (**a**) Gene region of DNA. The base sequence along one of DNA's two strands (not both) is used as the template. (**b**–**d**) Transcribing that region results in a molecule of RNA.

and STOP signal. A **promoter** is a START signal, a base sequence in DNA to which RNA polymerases bind and prepare for transcription. After binding, an RNA polymerase recognizes a gene region and moves along it. It uses the gene's base sequence as a template for covalently bonding free ribonucleotides together in a complementary sequence, as in Figure 14.3. When it reaches a sequence that signals "the end" of the gene region, the new RNA is released as a free transcript.

FINISHING TOUCHES ON THE mRNA TRANSCRIPTS

In eukaryotic cells, mRNA transcripts are modified before leaving the nucleus. Just as a dressmaker may snip off some threads or put bows on a dress before it leaves the shop, so do cells tailor their "pre-mRNA." For instance, some enzymes attach a modified guanine "cap" to the start of a pre-mRNA transcript. Others attach about 100 to 300 adenine ribonucleotides as a tail to the other end. Hence its name, poly-A tail.

Later, the pre-mRNA's cap will bind to a ribosome. Enzymes will nibble off the tail from the tip on back. Thus each tail's length dictates how long a particular protein-building message will last in the cytoplasm.

A transcript's message gets processed even before it leaves the nucleus. Eukaryotic genes contain **exons**: protein-coding base sequences that are interrupted by noncoding sequences, or **introns**. Both are transcribed, but all introns are snipped out before the transcript reaches the cytoplasm (Figure 14.4). Either all exons are retained in a mature mRNA transcript or some are removed and the rest are spliced together in various combinations. By this **alternative splicing**, one gene can specify two or more proteins that differ slightly in form and function! Cells use different combinations of exons at different times. Alternative splicing was once considered to be a rare event. However, it may occur in half (or all) genes of the human genome. It helps explain how human cells can make hundreds of thousands of proteins from only 21,500 or so genes.

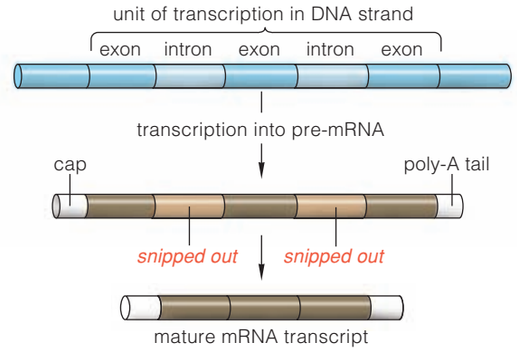
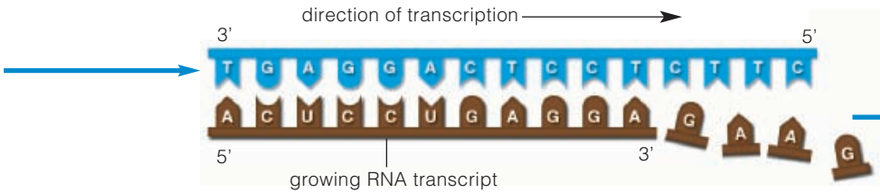


Figure 14.4 ***Animated!*** How pre-mRNA transcripts are processed into final form. Inside the nucleus, some or all introns are removed, and the transcript gets a cap and a tail.

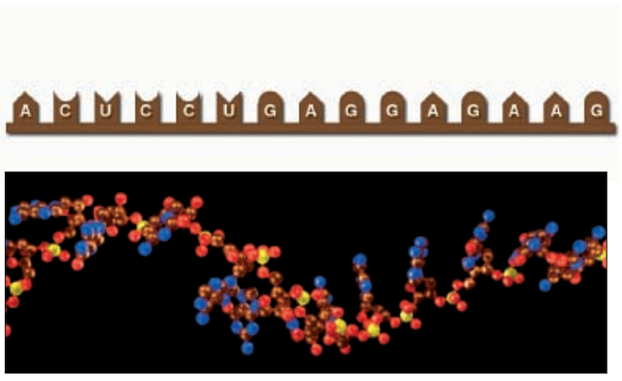
In gene transcription, a sequence of exposed bases on one of the two strands of a DNA molecule serves as a template for synthesizing a complementary strand of RNA.

RNA polymerases assemble the RNA from four kinds of ribonucleotides that differ in their bases: A, U, C, and G.

Before leaving the nucleus, each new mRNA transcript, or pre-mRNA, undergoes modification into final form.



c What happened in the gene region? RNA polymerase catalyzed the covalent bonding of ribonucleotides to one another to form an RNA strand. The base sequence in the new strand is complementary to the exposed bases on the DNA as a template. Many other proteins assist in transcription; compare Section 13.3.



d At the end of the gene region, the last stretch of the new transcript is unwound and released from the DNA template. Shown below it is a model for a transcribed strand of RNA.

14.2 The Genetic Code

The correspondence between genes and proteins is encoded in protein-building "words" in mRNA transcripts. Three nucleotide bases make up each three-letter word.

Figure 14.5*a* shows a bit of mRNA transcribed from a DNA template. To translate it, you have to know how many letters (bases) make each word (amino acid). That is what Marshall Nirenberg, Philip Leder, Severo Ochoa, and Gobind Korana figured out. After mRNA has docked at a ribosome, its bases are "read" *three at a time*. The base triplets in mRNA are **codons**. Figure 14.5*b* shows how their sequence corresponds to the amino acid sequence in a growing polypeptide chain.

There are sixty-four different codons even though there are only twenty amino acids in proteins (Figure 14.6). Why so many? Think it through. If the codon were only one nucleotide, mRNA could specify only four kinds of amino acids. Codons of two nucleotides could code for sixteen kinds of amino acids—still not enough. Mixes of three nucleotides could code for sixty-four kinds—more than enough.

Certain codons actually do specify more than one kind of amino acid. For instance, both GAA and GAG specify glutamate. Also, in most species, the first AUG in the transcript is a START signal for translating "three-bases-at-a-time." It also means methionine is the first amino acid in all new polypeptide chains. UAA, UAG, and UGA do not specify any amino acid. They are STOP signals that block further additions of amino acids to a new chain.

The set of sixty-four different codons is the **genetic code**, and it has been highly conserved through time. Prokaryotes, a few organelles derived from them, and some protists of ancient lineages have a few slightly variant codons. For instance, a few unique codons give mitochondria their own "mitochondrial code." We can predict that they are outcomes of gene mutations that did not alter the mix of proteins in adverse ways. The near-universal use of the genetic code indicates that there is little tolerance for variation.

The genetic code is a set of sixty-four different codons, which are nucleotide bases in mRNA that are "read" in sets of three. Different codons (base triplets) specify different amino acids.

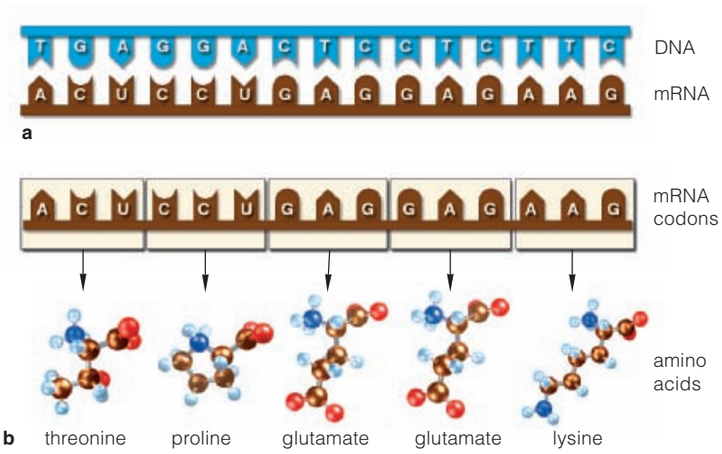


Figure 14.5 Example of the correspondence between genes and proteins. (**a**) An mRNA transcript of a gene region of DNA. Three nucleotide bases, equaling one codon, specify one amino acid. This series of codons (base triplets) specifies the sequence of amino acids shown in (**b**).

first base	second base: U	C	A	G	third base
U	phenylalanine	serine	tyrosine	cysteine	U
	phenylalanine	serine	tyrosine	cysteine	C
	leucine	serine	STOP	STOP	A
	leucine	serine	STOP	tryptophan	G
C	leucine	proline	histidine	arginine	U
	leucine	proline	histidine	arginine	C
	leucine	proline	glutamine	arginine	A
	leucine	proline	glutamine	arginine	G
A	isoleucine	threonine	asparagine	serine	U
	isoleucine	threonine	asparagine	serine	C
	isoleucine	threonine	lysine	arginine	A
	methionine (or START)	threonine	lysine	arginine	G
G	valine	alanine	aspartate	glycine	U
	valine	alanine	aspartate	glycine	C
	valine	alanine	glutamate	glycine	A
	valine	alanine	glutamate	glycine	G

Figure 14.6 ***Animated!*** *Right*, the near-universal genetic code. Each codon in mRNA is a set of three ribonucleotide bases. Sixty-one of these base triplets encode specific amino acids. Three are signals that stop translation.

The *left* vertical column (*brown*) lists choices for the first base of a codon. The *top* horizontal row (*light tan*) lists the second choices.The *right* vertical column (*dark tan*) lists the third. To give three examples, reading left to right, the triplet U|G|G corresponds to tryptophan. Both U|U|U and U|U|C correspond to phenylalanine.

14.3 The Other RNAs

Let's take stock. The codons in an mRNA transcript are the words in protein-building messages. Without translators, words that originated in DNA mean nothing; it takes the other two classes of RNA to synthesize proteins. Before getting into the mechanisms of translation, reflect on this overview of their structure and function.

Figure 14.7 shows the molecular structure for one of the tRNAs. All cells have pools of tRNAs and amino acids in their cytoplasm. Each tRNA has a molecular "hook," an attachment site for an amino acid. It has an **anticodon**, a ribonucleotide base triplet that can base-pair with a complementary codon in an mRNA transcript. When tRNAs bind to mRNA on a ribosome, the amino acid attached to each becomes positioned automatically in the order that the codons specify.

There are sixty-four codons but not as many kinds of tRNAs. How do tRNAs match up with more than one type of codon? According to base-pairing rules, adenine pairs with uracil, and cytosine with guanine. However, in codon–anticodon interactions, these rules can loosen for the third base in a codon. This freedom in codon–anticodon pairing at a base is known as the "wobble effect." For example, AUU, AUC, and AUA specify isoleucine. All three codons can base-pair with one type of tRNA that hooks on to isoleucine.

Again, interactions between the tRNAs and mRNA take place at ribosomes. A ribosome has two subunits made of rRNA and structural proteins (Section 4.5 and Figure 14.8). In eukaryotic cells, they are built in the nucleus and moved to the cytoplasm. There, a large and small subunit converge as an intact, functional ribosome only when mRNA is to be translated.

LINK TO SECTION 4.5

Only mRNA carries DNA's protein-building instructions from the nucleus into the cytoplasm.

tRNAs deliver amino acids to ribosomes. Their anticodons base-pair with codons in the order specified by mRNA.

Polypeptide chains are built on ribosomes, each consisting of a large and small subunit made of rRNA and proteins.

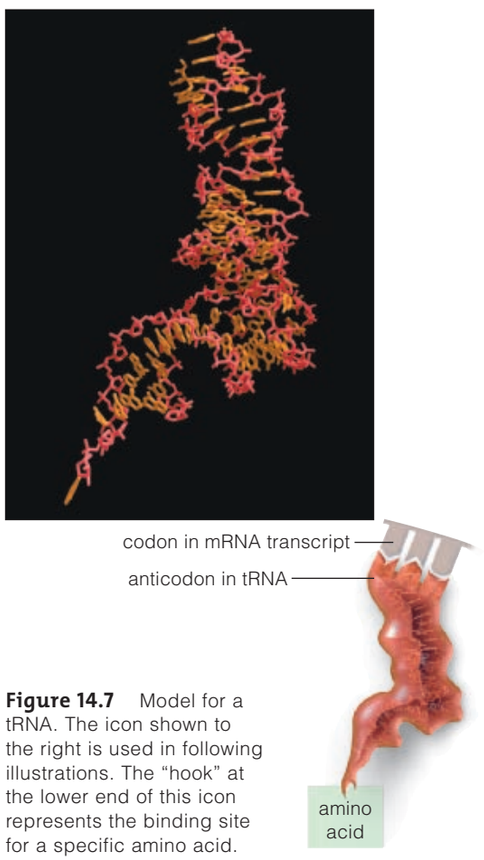


Figure 14.7 Model for a tRNA. The icon shown to the right is used in following illustrations. The "hook" at the lower end of this icon represents the binding site for a specific amino acid.

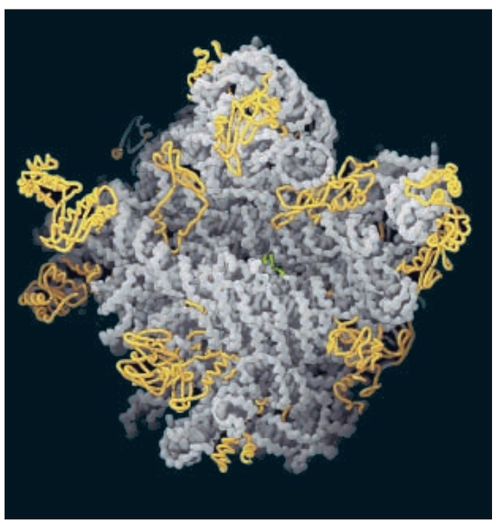


Figure 14.8 (**a**) Ribbon model for the large subunit of a bacterial ribosome. It has two rRNA molecules (*gray*) and thirty-one structural proteins (*gold*), which stabilize the structure. At one end of a tunnel through the large subunit, rRNA catalyzes polypeptide chain assembly. This is an ancient, highly conserved structure. Its role is so vital that the corresponding subunit of the eukaryotic ribosome, which is larger, may be similar in structure and function. (**b**) Model for the small and large subunits of a eukaryotic ribosome.

14.4 The Three Stages of Translation

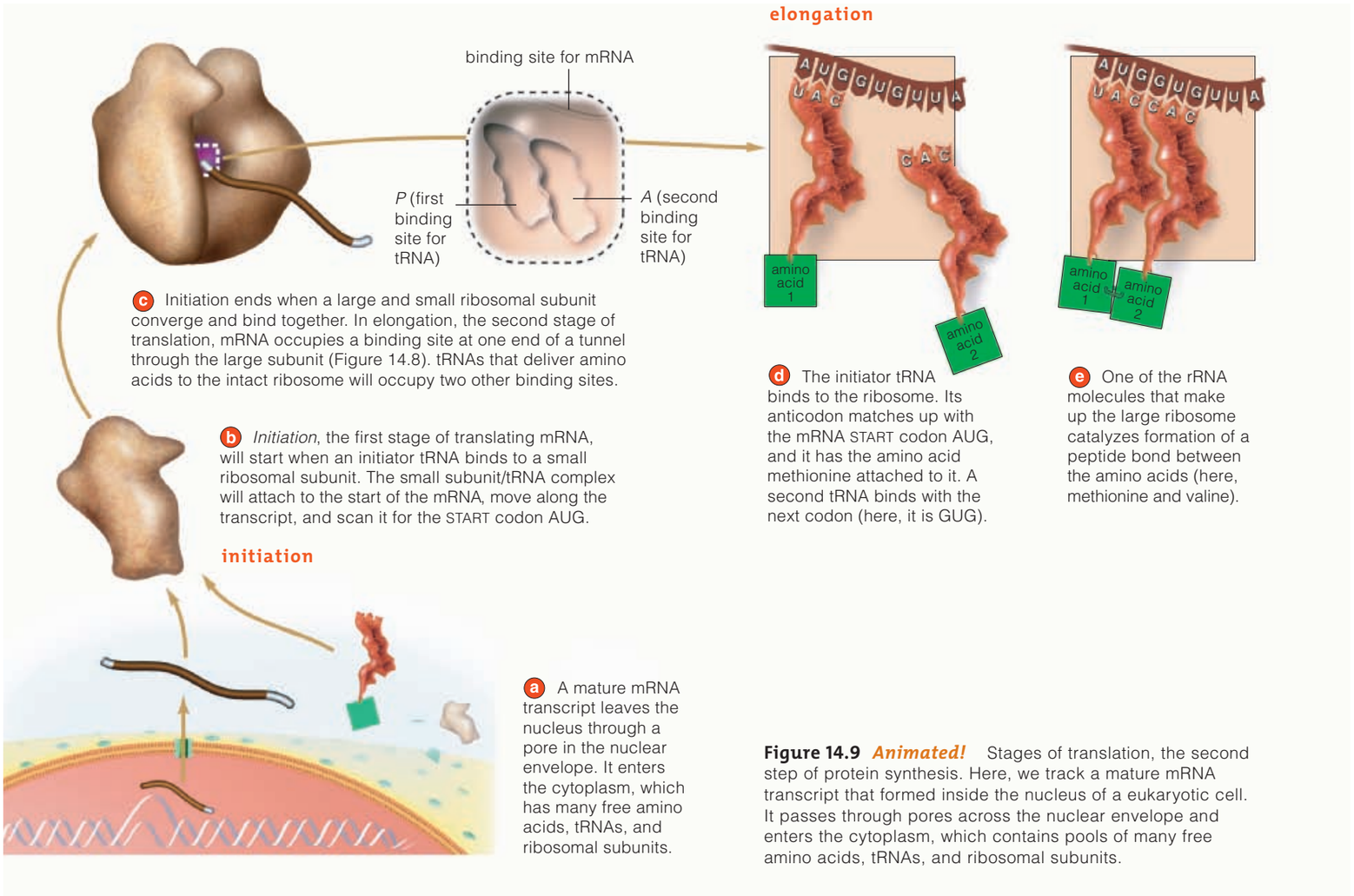
LINKS TO SECTIONS 3.5, 4.6

An mRNA transcript that encodes DNA's information about a protein enters an intact ribosome. There, its codons are translated into a polypeptide chain—a protein's primary structure (Section 3.5). Translation of the protein-building message proceeds through three continuous stages called initiation, elongation, and termination.

Only one kind of tRNA can start the *initiation* stage of translation. It alone has the anticodon UAC—which is complementary to the START codon of every mRNA transcript. The anticodon and codon meet up when this initiator tRNA binds to a small ribosomal subunit. Next, a large ribosomal subunit joins with the small subunit. Together, the initiator tRNA, the ribosome, and the mRNA transcript form an initiation complex (Figure 14.9*a–c*). The next stage can begin.

During the *elongation* stage, a polypeptide chain is synthesized while the mRNA passes between the two ribosomal subunits, a bit like a thread being moved through the eye of a needle. Many tRNA molecules deliver amino acids to the ribosome, and each binds to the mRNA in the order specified by their codons. One region of an rRNA molecule located at the center of the large ribosomal subunit is highly acidic, and it functions as an enzyme. It catalyzes the formation of peptide bonds between amino acids (Figure 14.9*d–f*).

Figure 14.9*g* shows how one peptide bond forms between the most recently attached amino acid and the next one brought to the ribosome. Here, you might wish to look once more at Section 3.5, which includes a step-by-step description of peptide bond formation during protein synthesis.



c Initiation ends when a large and small ribosomal subunit converge and bind together. In elongation, the second stage of translation, mRNA occupies a binding site at one end of a tunnel through the large subunit (Figure 14.8). tRNAs that deliver amino acids to the intact ribosome will occupy two other binding sites.

b *Initiation*, the first stage of translating mRNA, will start when an initiator tRNA binds to a small ribosomal subunit. The small subunit/tRNA complex will attach to the start of the mRNA, move along the transcript, and scan it for the START codon AUG.

d The initiator tRNA binds to the ribosome. Its anticodon matches up with the mRNA START codon AUG, and it has the amino acid methionine attached to it. A second tRNA binds with the next codon (here, it is GUG).

e One of the rRNA molecules that make up the large ribosome catalyzes formation of a peptide bond between the amino acids (here, methionine and valine).

a A mature mRNA transcript leaves the nucleus through a pore in the nuclear envelope. It enters the cytoplasm, which has many free amino acids, tRNAs, and ribosomal subunits.

Figure 14.9 ***Animated!*** Stages of translation, the second step of protein synthesis. Here, we track a mature mRNA transcript that formed inside the nucleus of a eukaryotic cell. It passes through pores across the nuclear envelope and enters the cytoplasm, which contains pools of many free amino acids, tRNAs, and ribosomal subunits.

During *termination*, the last stage of translation, the mRNA's STOP codon enters the ribosome. No tRNA has a corresponding anticodon. Proteins called release factors bind to the ribosome. Binding triggers enzyme activity that detaches the mRNA *and* the polypeptide chain from the ribosome (Figure 14.9*i–k*).

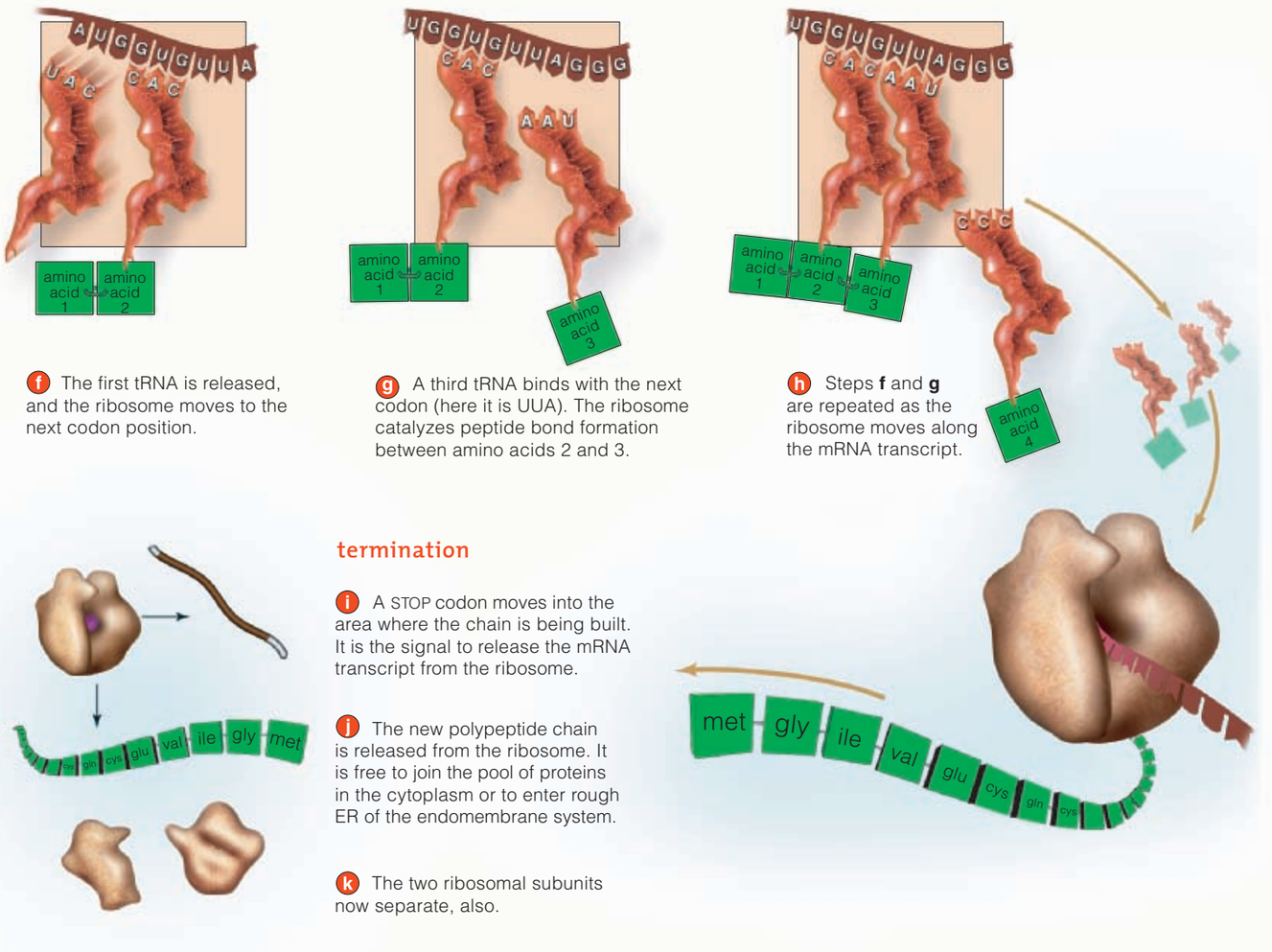
In cells that are quickly using or secreting proteins, you often see many clusters of ribosomes (polysomes) on an mRNA transcript, all translating it at the same time. This is what happens in unfertilized eggs, which usually stockpile mRNA transcripts in the cytoplasm in preparation for the cell divisions that lie ahead.

Many newly formed polypeptide chains carry out their functions in the cytoplasm. Others have a special sequence of amino acids. The sequence is a shipping label that gets them into ribosome-studded, flattened sacs of rough ER (Section 4.6). In the organelles of the endomembrane system, the chains will take on final form before shipment to their ultimate destinations as structural or functional proteins.

Translation is initiated when a small ribosomal subunit and an initiator tRNA arrive at an mRNA transcript's START codon, and a large ribosomal subunit binds to them.

tRNAs deliver amino acids to a ribosome in the order dictated by the linear sequence of mRNA codons. A polypeptide chain lengthens as peptide bonds form between the amino acids.

Translation ends when a STOP codon triggers events that cause the polypeptide chain and the mRNA to detach from the ribosome.



f The first tRNA is released, and the ribosome moves to the next codon position.

g A third tRNA binds with the next codon (here it is UUA). The ribosome catalyzes peptide bond formation between amino acids 2 and 3.

h Steps **f** and **g** are repeated as the ribosome moves along the mRNA transcript.

i A STOP codon moves into the area where the chain is being built. It is the signal to release the mRNA transcript from the ribosome.

j The new polypeptide chain is released from the ribosome. It is free to join the pool of proteins in the cytoplasm or to enter rough ER of the endomembrane system.

k The two ribosomal subunits now separate, also.

14.5 Mutated Genes and Their Protein Products

LINKS TO SECTIONS 2.3, 3.6, 7.1

When a cell taps its genetic code, it is making proteins with precise structural and functional roles that keep it alive. If a gene changes, the mRNA transcribed from it may change and specify an altered protein. If the protein has a crucial role, the outcome will be a dead or abnormal cell.

Gene sequences can change. Sometimes one base gets substituted for another in the nucleotide sequence. At other times, an extra base is inserted or one is lost. Such small-scale changes in the nucleotide sequence of a DNA molecule are **gene mutations**, and they can alter the message that becomes encoded in mRNA. Cells have some leeway, because more than one codon can specify the same amino acid. For example, if UCU replaced UCC in an mRNA transcript, this might not be bad, because both codons specify serine. However, as the next examples show, many mutations result in proteins that function in an altered way or not at all.

COMMON GENE MUTATIONS

During DNA replication, recall, the wrong nucleotide may become paired with an exposed base on the DNA template and slip by proofreading and repair enzymes (Section 13.3). This type of mutation is a **base-pair substitution**. When the altered message is translated, it may call for the wrong amino acid or a premature STOP codon. Figure 14.10*b* shows how adenine *replaced* one thymine in the gene for beta hemoglobin, which can give rise to sickle-cell anemia (Section 3.6).

Figure 14.10*c* depicts another gene mutation, one in which a single base—thymine—was *deleted*. Again, DNA polymerases read base sequences in blocks of three. A deletion is one of the *frameshift* mutations; it shifts the "three-bases-at-a-time" reading frame. An altered mRNA is transcribed from the mutant gene, so an altered protein is the result.

Frameshift mutations fall in the broader categories of **insertions** and **deletions**. One or more base pairs become inserted into DNA or are deleted from it.

Other mutations arise from transposable elements, or **transposons**, that can jump around in the genome. Geneticist Barbara McClintock found that these DNA segments or copies of them move spontaneously to a new location in a chromosome or even to a different chromosome. When transposons land in a gene, they alter the timing or duration of its activity, or block it entirely. Their unpredictability can give rise to odd variations in traits. Figure 14.11 gives an example.

HOW DO MUTATIONS ARISE?

Many mutations happen spontaneously while DNA is being replicated. This is not surprising, given the swift pace of replication (about twenty bases per second in humans and a thousand bases per second in certain bacteria). DNA polymerases and DNA ligases can fix most mistakes (Section 13.3). But sometimes they go on assembling a new strand right over an error. The bypass can result in a mutated DNA molecule.

Not all mutations are spontaneous. A number arise after DNA is exposed to mutation-causing agents. To give an example, x-rays and other high-energy forms of **ionizing radiation** break chromosomes into pieces (Figure 14.12). Ionizing radiation damages DNA indirectly, also. When it penetrates living tissues, it leaves behind a long trail of destructive free radicals. Doctors and

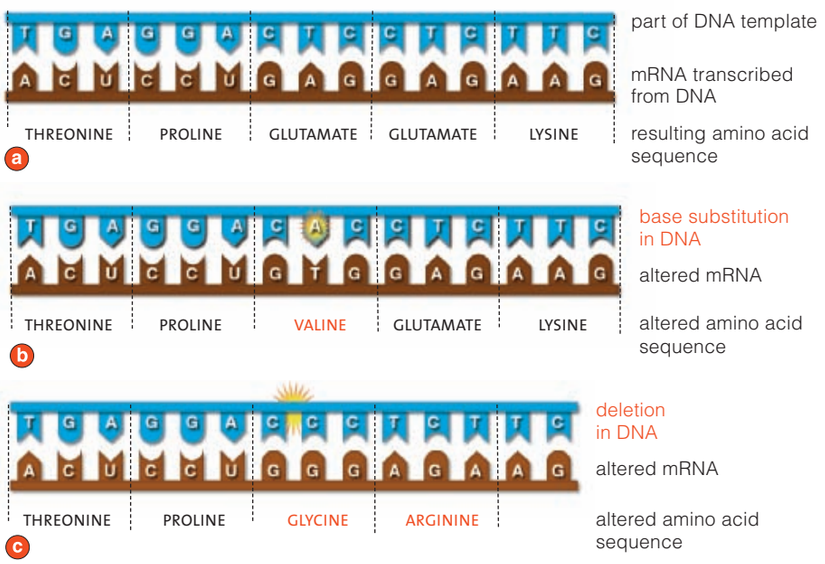


Figure 14.10 ***Animated!*** Example of gene mutation. (**a**) Part of a gene, the mRNA, and the specified amino acid sequence of the beta chain in hemoglobin. (**b**) A base-pair substitution in DNA replaces a thymine with an adenine. When the altered mRNA transcript is translated, valine replaces glutamate as the sixth amino acid of the new polypeptide chain. Sickle-cell anemia is the eventual outcome.

(**c**) Deletion of the same thymine would be a frameshift mutation. The reading frame for the rest of the mRNA shifts, a different protein product forms, and it causes thalassemia—a different type of red blood cell disorder.

Figure 14.11 Barbara McClintock, who won a Nobel Prize for her research. She proved that transposons slip into and out of different locations in DNA. The curiously nonuniform coloration of kernels in strains of Indian corn (*Zea mays*) sent her on the road to discovery.

Several genes govern pigment formation and deposition in corn kernels, which are a type of seed. Mutations in one or more of these genes produce yellow, white, red, orange, blue, and purple kernels. However, as McClintock realized, *unstable* mutations can cause streaks or spots in *individual* kernels.

All of a corn plant's cells have the same pigment-encoding genes. But a transposon invaded a pigment-encoding gene before the plant started growing from a fertilized egg. While a kernel's tissues were forming, its cells could not make pigment, but the same transposon jumped out of the pigment-encoding gene in some of its cells. Descendants of *those* cells could make pigment. The spots and streaks in individual kernels are visual markers for those cell lineages.

Figure 14.12 Chromosomes from a human cell after exposure to gamma rays, a form of ionizing radiation. We can expect such broken pieces (*arrows*) to be lost during interphase, when DNA is being replicated. The extent of the chromosome damage in an exposed cell typically depends on how much radiation it absorbed.

dentists both use the lowest possible doses of x-rays to minimize the damage to a patient's DNA.

Nonionizing radiation excites electrons to a higher energy level. DNA absorbs one form, ultraviolet (UV) light. Two nucleotide bases in DNA—cytosine and thymine—are most vulnerable to excitation that can change base-pairing properties. UV light can induce adjacent thymine bases in a DNA strand to pair *with each other*, as a bulky dimer (page 217). At least seven gene products interact as a DNA repair mechanism to remove the dimer, which wrinkles the DNA. If DNA polymerase encounters a thymine dimer, it will make replication errors. Exposing unprotected skin to the sun invites thymine dimer formation in skin cells.

When thymine dimers are not repaired, they cause DNA polymerases to make even more errors during the next replication cycle. They are the original source of mutations that lead to certain cancers.

Natural and synthetic chemicals accelerate rates of gene mutations. For instance, **alkylating agents** can transfer charged methyl or ethyl groups to reactive sites in DNA. At these sites, DNA is more vulnerable to mistakes in base pairing and to mutation. Cancer-causing agents in cigarette smoke and many other substances exert their effects by alkylating DNA.

THE PROOF IS IN THE PROTEIN

When a mutation arises in a somatic cell of a sexually reproducing individual, its good or bad effects will not endure; it is not passed on to offspring. If it arises in a germ cell or a gamete, however, it may enter the evolutionary arena. It also may do so when it is passed on to offspring by asexual reproduction. Either way, *the protein product of such heritable mutations will have harmful, neutral, or beneficial effects on the individual's capacity to function in the prevailing environment.* The effects of uncountable mutations in millions of species have had spectacular evolutionary consequences—and that is a topic of later chapters.

A gene mutation is a permanent change in one or more bases in the nucleotide sequence of DNA. The most common types are base-pair substitutions, insertions, and deletions.

Exposure to harmful radiation and to chemicals in the environment can cause mutations in DNA.

A protein specified by a mutated gene may have harmful, neutral, or beneficial effects on the individual's capacity to function in the environment.

Summary

Introduction All enzymes and other proteins that are essential for life consist of polypeptide chains. Each chain, a linear sequence of amino acids, corresponds to nucleotide base sequences in DNA that form genes. The path from genes to proteins has two steps: transcription and translation (Figure 14.13).

Section 14.1 In eukaryotic cells, genes are transcribed in the nucleus and then translated cytoplasm. Both steps occur in the cytoplasm of prokaryotic cells, which have no nucleus. Enzymes unwind the two strands of a DNA double helix in a specific gene region. RNA polymerases covalently bond ribonucleotides one after another into a new RNA transcript, in an order complementary to the exposed bases on the DNA template. Adenine, guanine, cytosine, and uracil are the bases in ribonucleotides.

The mRNA transcript gets modified before it leaves the nucleus. Its 5′ end gets capped, and its 3′ end gets a poly-A tail, which paces how long the mRNA will stay intact in the cytoplasm. The introns between exons (the protein-coding portions of genes) are snipped out. The exons can be spliced together in different combinations.

Biology Now
Learn how genes are transcribed and transcripts are processed with the animation on BiologyNow.

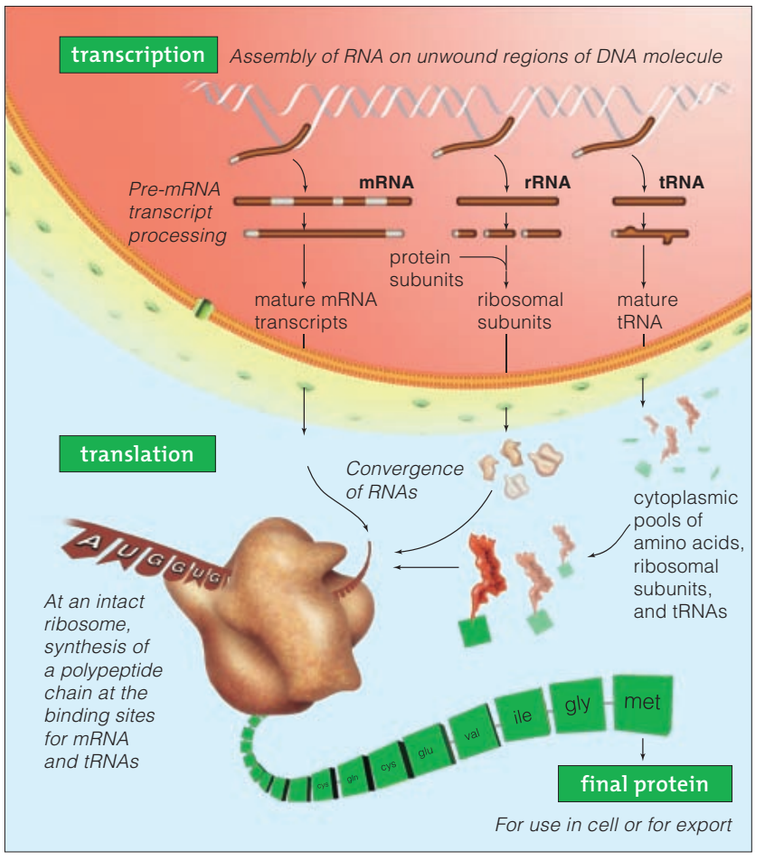


Figure 14.13 ***Animated!*** Summary of protein synthesis in eukaryotic cells. DNA is transcribed into RNA in the nucleus. RNA is translated in the cytoplasm. Prokaryotic cells do not have a nucleus; transcription and translation proceed in their cytoplasm.

Sections 14.2, 14.3 Only messenger RNA (mRNA) carries the protein-building information in DNA to ribosomes for translation. Its genetic message is written in codons, or sets of three nucleotides along an mRNA strand that specify an amino acid. There are sixty-four codons, a few of which act as START or STOP signals for translation. That set constitutes a highly conserved genetic code. A few variations in code words evolved among prokaryotes and prokaryote-derived organelles (e.g., mitochondria) and in a few ancient lineages of single-celled eukaryotes.

Translation requires three classes of RNAs. Transfer RNA (tRNA) molecules have anticodons that can bind briefly to complementary codons in mRNA. They also have a binding site for a free amino acid, which they deliver to ribosomes during protein synthesis. Different tRNAs reversibly bind different amino acids. Ribosomal RNA (rRNA) and proteins that stabilize it make up the two subunits that form ribosomes.

Biology Now
Explore the genetic code with the interaction on BiologyNow.

Section 14.4 During translation, peptide bonds form between amino acids in the order specified by codons in mRNA. Translation has three stages. In initiation, an initiator tRNA, two ribosomal subunits, and an mRNA converge as an initiation complex. In the elongation stage, tRNAs deliver amino acids to the intact ribosomes. Part of an rRNA molecule located in the ribosome's central region catalyzes peptide bond formation between amino acids. In the termination stage, a STOP codon and other factors trigger the release of mRNA and the new polypeptide chain. They also cause the ribosome's subunits to separate from each other.

Biology Now
Observe the translation of an mRNA transcript with the animation on BiologyNow.

Section 14.5 Gene mutations are heritable, small-scale changes in the base sequence of DNA. Major types are base-pair substitutions, insertions, and deletions. Many arise spontaneously as DNA is being replicated. Some arise after transposons jump to new locations in chromosomes; others arise after DNA is exposed to ionizing radiation or to chemicals in the environment. Mutations may cause changes in protein structure, protein function, or both.

Biology Now
Investigate the effects of mutation with the animation on BiologyNow.

Self-Quiz

Answers in Appendix II

1. DNA contains many different gene regions that are transcribed into different _____ .
   a. proteins
   b. mRNAs only
   c. mRNAs, tRNAs, rRNAs
   d. all of the above

2. An RNA molecule is typically _____ .
   a. a double helix
   b. single-stranded
   c. double-stranded
   d. triple-stranded

3. An mRNA molecule is synthesized by _____ .
   a. replication
   b. duplication
   c. transcription
   d. translation

4. Each codon specifies a(n) _____ .
   a. protein
   b. polypeptide
   c. amino acid
   d. mRNA

5. _____ different codons represent a near-universal genetic code.
   a. Twelve
   b. Twenty
   c. Thirty-four
   d. Sixty-four

6. Anticodons pair with _____ .
   a. mRNA codons
   b. DNA codons
   c. RNA anticodons
   d. amino acids

7. _____ can cause gene mutations.
   a. replication errors
   b. transposons
   c. ionizing radiation
   d. non-ionizing radiation
   e. b and c are correct
   f. all of the above

8. Match the terms with the most suitable description.
   ____ alkylating agent
   ____ chain elongation
   ____ exons
   ____ genetic code
   ____ anticodon
   ____ introns
   ____ codon
   a. protein-coding parts of a mature mRNA transcript
   b. base triplet for amino acid
   c. second stage of translation
   d. base triplet; pairs with codon
   e. one environmental agent that induces mutation in DNA
   f. set of 64 codons for mRNA
   g. noncoding part of pre-mRNA transcript, removed before translation

Additional questions are available on BiologyNow™

Critical Thinking

1. Using Figure 14.6, translate this nucleotide sequence in part of an mRNA transcript into an amino acid sequence:

5′—GGTTTCTTCAAGAGA—3′

2. Briefly review Section 13.3. Now suppose that DNA polymerase made a wrong base pairing while a crucial gene region of DNA was being replicated. DNA repair mechanisms did not kick in to fix the mistake. Here is the part of the DNA strand that contains the error:

...AATTCCGACTCCTATGG
...TTAAGGTTGAGGATACC

After the DNA molecule is replicated, two daughter cells form. One daughter cell is carrying the mutation and the other cell is normal. Develop a hypothesis to explain this observation.

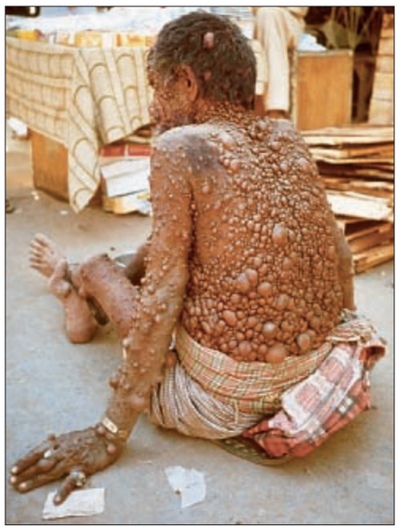


Figure 14.14 Soft skin tumors on an individual affected by the autosomal dominant disorder called neurofibromatosis.

3. *Neurofibromatosis* is a human autosomal dominant disorder caused by mutations in the *NF1* gene. It is characterized by the formation of soft, fibrous tumors in the peripheral nervous system and skin as well as abnormalities in muscles, bones, and internal organs (Figure 14.14).

Because the mutant allele is dominant, an affected child usually has an affected parent. Yet in 1991, scientists reported that a boy developed neurofibromatosis even though his parents did not. When they examined both copies of the boy's *NF1* gene, they found that the gene on the chromosome he inherited from his father contained a transposon. Neither father nor mother had a transposon in any of the copies of their *NF1* genes. Explain the cause of neurofibromatosis in the boy and how it arose.

4. Cigarette smoke is mostly carbon dioxide, nitrogen, and oxygen. The rest contains at least fifty-five different chemicals identified as carcinogenic, or cancer-causing, by the International Agency for Research on Cancer (IARC). When these carcinogens enter the bloodstream, enzymes convert them to a series of chemical intermediates that are easier to excrete. Some of the intermediates bind irreversibly to DNA. Propose one mechanism by which smoking cigarettes can cause cancer.

5. *Antisense drugs* may help us fight cancer and viral diseases, including SARS. These short mRNA strands are complementary to mRNAs that have been linked to these illnesses. Speculate on how these drugs work.

6. In some cases, the termination of transcription of prokaryotic DNA depends on the structure of the newly forming RNA transcript. The terminal end of an mRNA transcript often folds back tightly on itself and makes a hairpin-looped structure, like the one shown at right.

Why do you suppose that a "stem-loop" structure such as this stops transcription of prokaryotic DNA when the RNA polymerases reach it?

C
U—C
G—C
A—U
C—G
C—G
G—C
C—G
C—G
...CCCACAG—CAUUUUU...