Chapter 7: Genetics
Lesson 7.1: From DNA to Proteins

The spiral structure in the picture is a large organic molecule. Can you guess what it is? Here's a hint: molecules like this one determine who you are. They contain genetic information that controls your characteristics. They determine your eye color, facial features, and other physical attributes. What molecule is it?

You probably answered ''DNA.'' Today, it is commonly known that DNA is the genetic material in your cells. It was passed on to you from your parents and determines your characteristics. The discovery that DNA is the genetic material was an important milestone in molecular biology.

Lesson Objectives

- State the central dogma of molecular biology.
- Describe the structure of RNA, and identify the three main types of RNA.
- Give an overview of transcription.
- Describe the genetic code.
- Explain how translation occurs.

Vocabulary

- central dogma of molecular biology
- codon
- editing
- elongation
- exon
- genetic code
- initiation
- intron
- messenger RNA (mRNA)
- polyadentation
- promoter
- protein synthesis
- ribosomal RNA (rRNA)
- RNA
- RNA polymerase
- splicing
- termination
- transfer RNA (tRNA)
- transcription
- translation

Introduction

Your DNA, or deoxyribonucleic acid, contains the genes that determine who you are. How can this organic molecule control your characteristics? DNA contains instructions for all the proteins your body makes. Proteins, in turn, determine the structure and function of all your cells. What determines a protein's structure? It begins with the sequence of amino acids that make up the protein. Instructions for making proteins with the correct sequence of amino acids are encoded in DNA.

Central Dogma of Molecular Biology

DNA is found in chromosomes. In eukaryotic cells, chromosomes always remain in the nucleus, but proteins are made at ribosomes in the cell. How do the instructions in DNA get to the site of protein synthesis outside the nucleus? Another type of nucleic acid is responsible. This nucleic acid is RNA, or ribonucleic acid. RNA is a small molecule that can squeeze through pores in the nuclear membrane. It carries the information from DNA in the nucleus to a ribosome in the cell and then helps assemble the protein.

In short: **DNA → RNA → Protein**

Discovering this sequence of events was a major milestone in molecular biology. It is called the central dogma of molecular biology**,** the phase itself was coined by Francis Crick who stated "I called this idea the central dogma, for two reasons, I suspect. I had already used the obvious word hypothesis in the sequence hypothesis, and in addition I wanted to suggest that this new assumption was more central and more powerful." An overview of protein synthesis can be viewed at http://www.youtube.com/watch?v=-ygpqVr7_xs&feature=related (10:46).

RNA

DNA alone cannot ''tell" your cells how to make proteins. It needs the help of RNA, the other main player in the central dogma of molecular biology. Remember, DNA "lives" in the nucleus, but proteins are made on the ribosomes in the cytoplasm. How does the genetic information get from the nucleus to the cytoplasm? RNA is the answer.

RNA vs. DNA

RNA, like DNA, is a nucleic acid made up of nucleotides. However, RNA differs from DNA in several ways, see **Figure 7.1** below. In addition to being smaller than DNA, RNA also

- consists of one nucleotide chain instead of two,
- contains the nitrogen base uracil (U) instead of thymine,
- contains the sugar ribose instead of deoxyribose.

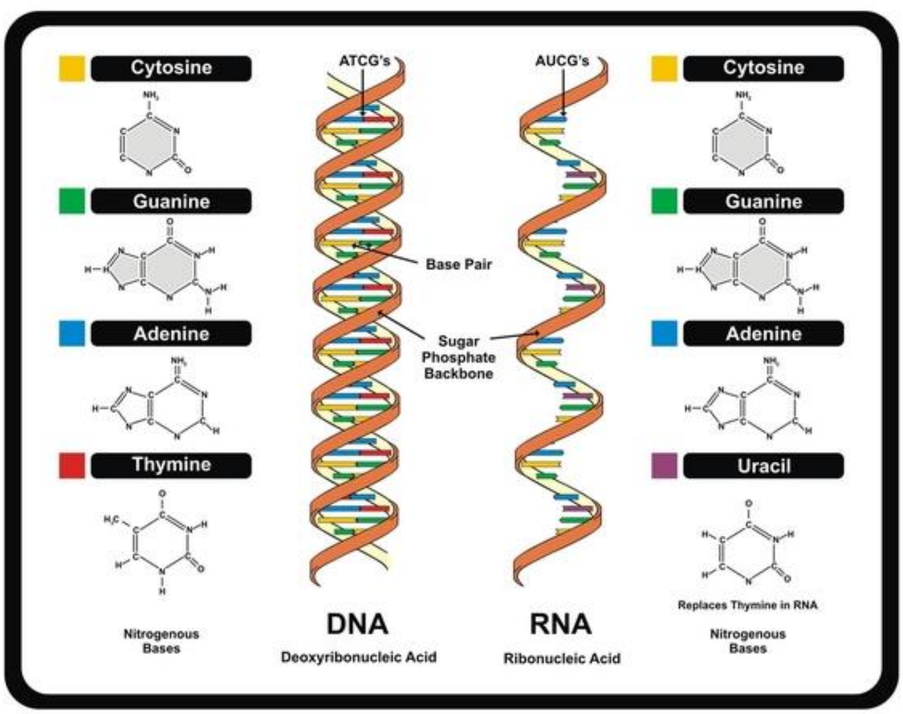

Figure 7.1: Structural comparison of the DNA molecule to the RNA molecule.

Types of RNA

There are three main types of RNA, all of which are involved in making proteins.

1. **Messenger RNA (mRNA)** copies the genetic instructions from DNA in the nucleus, and carries them to the ribosomes in the cytosol.

2. **Ribosomal RNA (rRNA)** helps form ribosomes, where proteins are assembled.

3. **Transfer RNA (tRNA)** brings amino acids to ribosomes, where they are joined together to form proteins.

All three types are shown in **Figure 7.2** below**.**

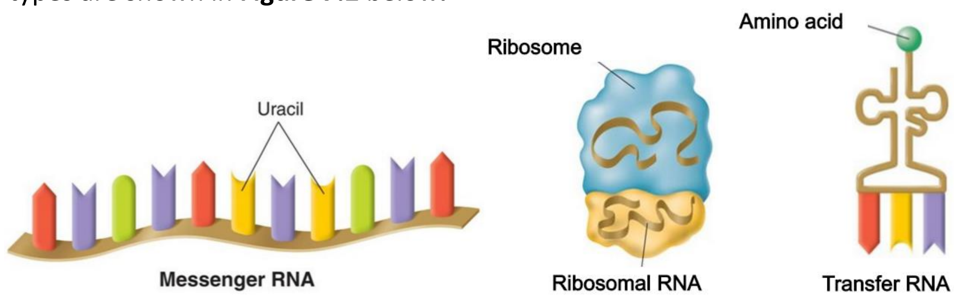

Figure 7.2 Shown are the three types of RNA and their roles: (1) mRNA copies the genetic message and carries it out of the nucleus, (2) rRNA is the main component of the ribosome and it site where proteins are synthesized, and (3) tRNA transfers the amino acids to the ribosome to build polypeptide chains, which will be folded into proteins. More on the roles of the RNAs will be discussed in the "Protein Synthesis" section of this textbook.

Protein Synthesis

The process in which cells make proteins is called protein synthesis. It actually consists of two processes: transcription and translation. Transcription takes place in the nucleus. It uses DNA as a template to make an RNA molecule, mRNA. The mRNA then leaves the nucleus and goes to a ribosome in the cell, where translation occurs. Translation reads the genetic code in mRNA and makes a protein.

The vocabulary of DNA, including the two processes involved in the central dogma, transcription and translation, is discussed at http://www.youtube.com/watch?v=s9HPNwXd9fk (18:23).

Transcription

To transcribe means "to paraphrase or summarize in writing." The information in DNA is transcribed - or summarized - into a smaller version - RNA - that can be used by the cell. This process is called transcription. Transcription is the first part of the central dogma of molecular biology: **DNA → RNA**. It is the transfer of genetic instructions in DNA to mRNA. During transcription, a strand of mRNA is made that is complementary to a strand of DNA. **Figure 7.3** shows how this occurs.

You can watch an animation of the process at this link: http://bcs.whfreeman.com/webpub/Ektron/Hillis%20Principles%20of%20Life2e/Animated%20Tutorials/pol2e_at_1001_Transcription/pol2e_at_1001_Transcription.html.

.

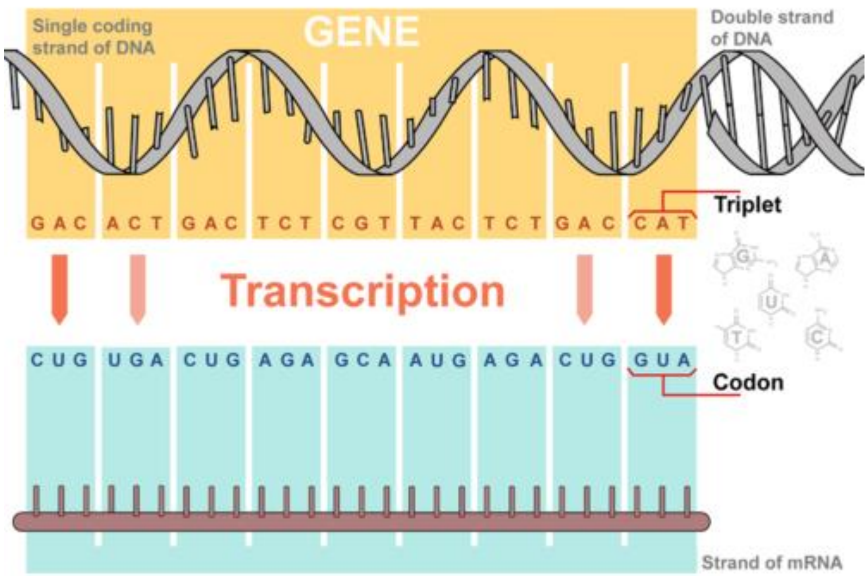

Figure 7.3 Overview of Transcription. Transcription uses the sequence of bases in a strand of DNA to make a complementary strand of mRNA. Triplets are groups of three successive nucleotide bases in DNA. Codons are complementary groups of bases in mRNA.

Steps of Transcription

Transcription takes place in three steps: initiation, elongation, and termination. The steps are illustrated in **Figure 7.4.**

1. Initiation is the beginning of transcription. It occurs when the enzyme RNA polymerase binds to a region of a gene called the promoter. This signals the DNA to unwind so the enzyme can copy the bases from a section of one of the DNA strands. The enzyme is ready to make a strand of mRNA with a complementary sequence of bases.
2. Elongation is the addition of nucleotides to the mRNA strand.
3. Termination is the ending of transcription, and occurs when RNA polymerase crosses a stop (termination) sequence in the gene. The mRNA strand is complete, and it detaches from DNA.

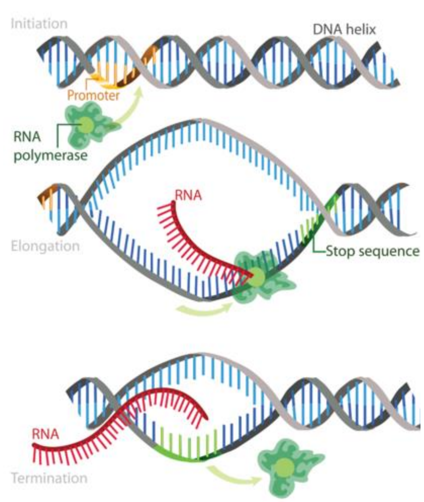

Figure 7.4 Steps of Transcription. Transcription occurs in the three steps - initiation, elongation, and termination - shown here.

Processing mRNA after Transcription

In eukaryotes, the new mRNA is not yet ready for translation. It must go through additional processing before it leaves the nucleus. This may include splicing, editing, and polyadenylation. These processes modify the mRNA in various ways. Such modifications allow a single gene to be used to make more than one protein.

- Splicing removes introns from mRNA (see **Figure 7.5**). Introns are regions that do not code for proteins. The remaining mRNA consists only of regions that do code for proteins, which are called exons. You can watch a video showing splicing in more detail at this link: http://vcell.ndsu.edu/animations/mrnasplicing/movie-flash.htm. Ribonucleoproteins are nucleoproteins that contain RNA. Small nuclear ribonuclearproteins are involved in pre-mRNA splicing.

- Editing changes some of the nucleotides in mRNA. For example, the human protein called APOB, which helps transport lipids in the blood, has two different forms because of editing. One form is smaller than the other because editing adds a premature stop signal in mRNA.

- Polyadenylation adds a "tail" to the mRNA. The tail consists of a string of A's (adenine bases). It signals the end of mRNA. It is also involved in exporting mRNA from the nucleus. In addition, the tail protects mRNA from enzymes that might break it down. See **Figure 7.6** below.

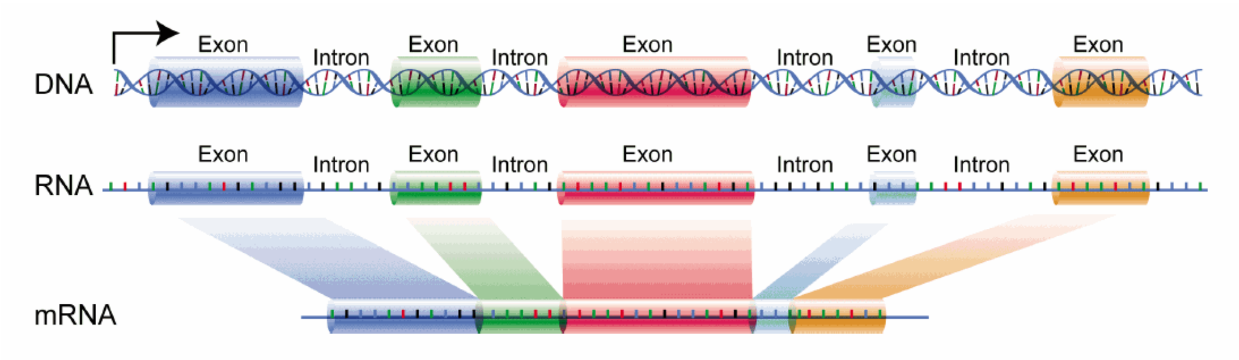

Figure 7.5 Splicing. Splicing removes introns from mRNA. UTR is an untranslated region of the mRNA.

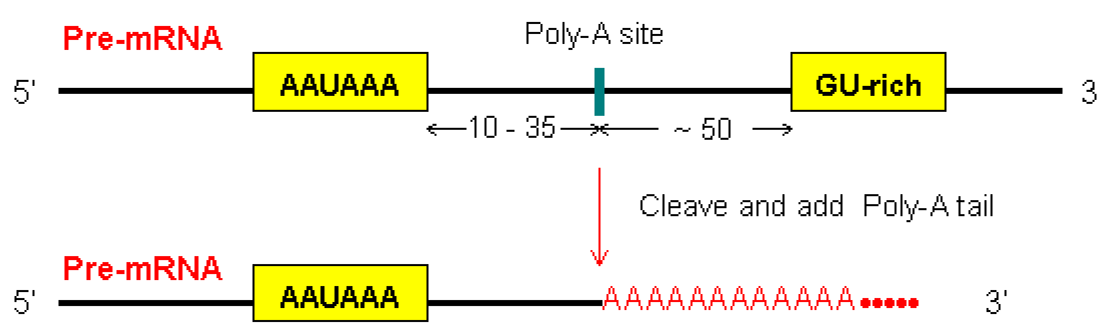

Figure 7.6 Polyadenylation. String of AAAAA bases added to end of mRNA molecule.

Translation

Translation is the second part of the central dogma of molecular biology: **RNA → Protein**. It is the process in which the genetic code in mRNA is read to make a protein. **Figure 7.7** shows how this happens. The sequence of nucleotides found on the newly transcribed mRNA molecule are instructions for the order in which amino acids should be joined to make polypeptide chains, which will eventually be folded into protein molecules with a variety of functions.

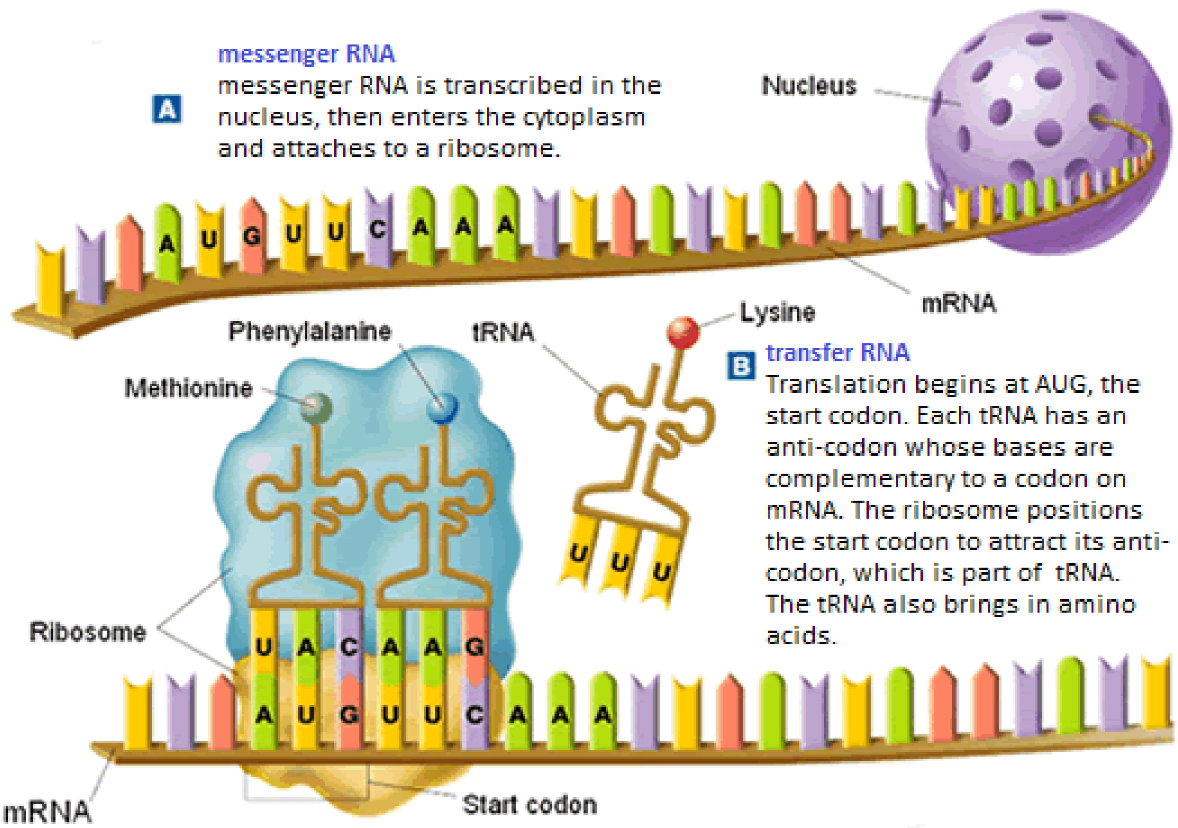

Figure 7.7: Translation. Translation of the codons in mRNA to a chain of amino acids occurs at a ribosome. Find the different types of RNA in the diagram. What are their roles in translation?

What exactly is read???

The Genetic Code

How is the information in a gene encoded? The answer is the genetic code. The genetic code consists of the sequence of nitrogen bases—A, C, G, T (or U)—in a polynucleotide chain. The four bases make up the ''letters'' of the genetic code. The letters are combined in groups of three to form code ''words,'' called codons. Each codon stands for (encodes) one amino acid, unless it codes for a start or stop signal. There are 20 common amino acids in proteins. There are 64 possible codons, more than enough to code for the 20 amino acids. The genetic code is shown in **Figure 7.8** on a codon chart. To see how scientists cracked the genetic code, go to this link: http://www.dnalc.org/view/16494-Animation-22-DNA-words-are-threeletters-long-.html. The Codon Wheel in the **Figure 7.9** is mostly used by students in Academic Biology to crack the genetic code.

1st base	2nd base →	U	C	A	G
U		UUU (Phe/F) Phenylalanine	UCU (Ser/S) Serine	UAU (Tyr/Y) Tyrosine	UGU (Cys/C) Cysteine
		UUC (Phe/F) Phenylalanine	UCC (Ser/S) Serine	UAC (Tyr/Y) Tyrosine	UGC (Cys/C) Cysteine
		UUA (Leu/L) Leucine	UCA (Ser/S) Serine	UAA Ochre (*Stop*)	UGA Opal (*Stop*)
		UUG (Leu/L) Leucine	UCG (Ser/S) Serine	UAG Amber (*Stop*)	UGG (Trp/W) Tryptophan
C		CUU (Leu/L) Leucine	CCU (Pro/P) Proline	CAU (His/H) Histidine	CGU (Arg/R) Arginine
		CUC (Leu/L) Leucine	CCC (Pro/P) Proline	CAC (His/H) Histidine	CGC (Arg/R) Arginine
		CUA (Leu/L) Leucine	CCA (Pro/P) Proline	CAA (Gln/Q) Glutamine	CGA (Arg/R) Arginine
		CUG (Leu/L) Leucine	CCG (Pro/P) Proline	CAG (Gln/Q) Glutamine	CGG (Arg/R) Arginine
A		AUU (Ile/I) Isoleucine	ACU (Thr/T) Threonine	AAU (Asn/N) Asparagine	AGU (Ser/S) Serine
		AUC (Ile/I) Isoleucine	ACC (Thr/T) Threonine	AAC (Asn/N) Asparagine	AGC (Ser/S) Serine
		AUA (Ile/I) Isoleucine	ACA (Thr/T) Threonine	AAA (Lys/K) Lysine	AGA (Arg/R) Arginine
		AUG[A] (Met/M) Methionine	ACG (Thr/T) Threonine	AAG (Lys/K) Lysine	AGG (Arg/R) Arginine
G		GUU (Val/V) Valine	GCU (Ala/A) Alanine	GAU (Asp/D) Aspartic acid	GGU (Gly/G) Glycine
		GUC (Val/V) Valine	GCC (Ala/A) Alanine	GAC (Asp/D) Aspartic acid	GGC (Gly/G) Glycine
		GUA (Val/V) Valine	GCA (Ala/A) Alanine	GAA (Glu/E) Glutamic acid	GGA (Gly/G) Glycine
		GUG (Val/V) Valine	GCG (Ala/A) Alanine	GAG (Glu/E) Glutamic acid	GGG (Gly/G) Glycine

nonpolar | polar | basic | acidic | (stop codon)

Figure 7.8: The Genetic Code. To find the amino acid for a particular codon, find the cell in the table for the first and second bases of the codon. Then, within that cell, find the codon with the correct third base. For example CUG codes for leucine, AAG codes for lysine, and GGG codes for glycine.

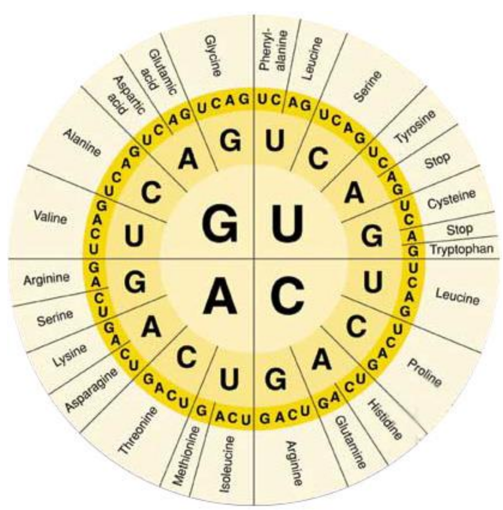

Figure 7.9: The Codon Wheel can also be used to decode the genetic code from an RNA sequence.

How to read a Codon Chart and a Codon Wheel

Remember that RNA is the instructions for making proteins. Ribosomes use RNA to make amino acid chains that later get folded into proteins. The RNE tells ribosomes the sequence, or order, that the amino acids should be in to make certain proteins. Scientists have developed the Codon Chart and the Codon Wheel to read RNA fragments and figure out what amino acid sequence will be made. Every three bases in an RNA sequence are called a CODON. Each codon codes for only one amino acid.

How to read the Codon Chart:

The codon chart gives the same information as the codon wheel. The first letter of the codon is on the left side. Second letter is on the top, and the third letter is on the right side. Where all the points meet, you will find the amino acid that codon represents. Looking at the Codon Chart pictured in your book, try using the chart to read the RNA sequence: AUC-GUA-AGG.

How to read the Codon Wheel:

Read from the inside out!!! For example refer to the Codon Wheel in your book, if the RNA sequence is: AUGCCAGAU, the codons are AUG, CCA, and GAU. To read the codon AUG you begin at the center of the Codon Wheel at letter 'A', then move to the next ring of letters and find the letter 'U', next move to the next ring and find the letter 'G', and finally you reach the outermost ring which will provide the name of the amino acid that the RNA sequence coded for. In this case AUG codes for Methionine. Try and decode the last two codons: CCA and GAU on your own.

Reading the Genetic Code

As shown, the codon AUG codes for the amino acid methionine. This codon is also the start codon that begins translation. The start codon establishes the reading frame of mRNA. The reading frame is the way the letters are divided into codons. After the AUG start codon, the next three letters are read as the second codon. The next three letters after that are read as the third codon, and so on. This is illustrated in **Figure 7.10**. The mRNA molecule is read, codon by codon, until a stop codon is reached. UAG, UGA, and UAA are all stop codons. They do not code for any amino acids.

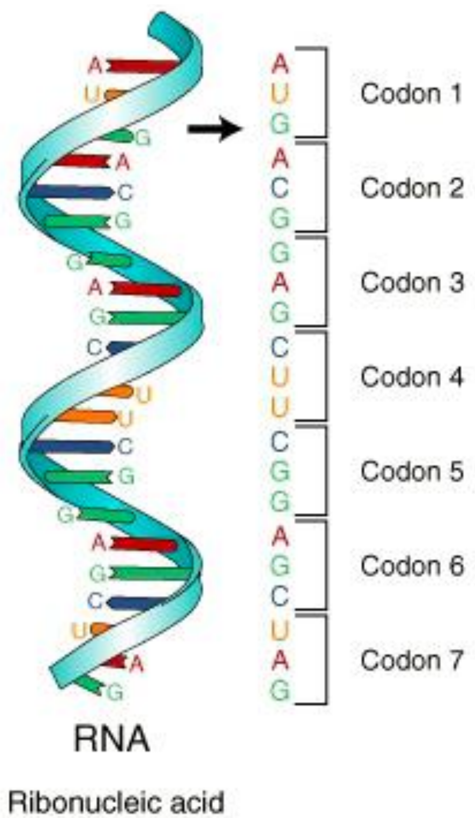

Figure 7.10: Reading the Genetic Code. The genetic code is read three bases at a time. Codons are the code words of the genetic code. Which amino acid does codon 2 in the drawing stand for?

Steps of Translation

1) After mRNA leaves the nucleus, it moves to a ribosome (rRNA), which consists of rRNA and proteins.
2) mRNA attaches to the ribosome; the ribosome reads (translates) the sequence of codons in mRNA.
3) Translation begins with the AUG codon (start codon) on mRNA.
4) Molecules of tRNA bring amino acids to the ribosome in the correct sequence. tRNA matches its anti-codon and amino acid to the mRNA codon.
   - To understand the role of tRNA, you need to know more about its structure. Each tRNA molecule has an anticodon for the amino acid it carries. An anticodon is complementary to the codon for an amino acid. For example, the amino acid lysine has the codon AAG, so the anticodon is UUC. Therefore, lysine would be carried by a tRNA molecule with the anticodon UUC. Wherever the codon AAG appears in mRNA, a UUC anticodon of tRNA temporarily binds. While bound to mRNA, tRNA gives up its amino acid.
5) Then a second tRNA comes in and matches its anti-codon to an mRNA codon and the first tRNA releases its amino acid which will bond to the amino acid on the second tRNA with a peptide bond.
6) Third tRNA comes in and matches its anti-codon to an mRNA codon and the second tRNA releases its amino acid which will bond to the amino acid on the third tRNA with another peptide bond.
7) The process explained in steps 5 and 6 will continue until an amino acid for each mRNA codon after the start codon has been brought in by tRNA molecules, forming a polypeptide chain.
8) The polypeptide chain of amino acids keeps growing until a stop codon is reached. See **Figure 7.11** below. UGA is the stop codon.

After a polypeptide chain is synthesized, it may undergo additional processes. For example, it may assume a folded shape due to interactions among its amino acids. It may also bind with other polypeptides or with different types of molecules, such as lipids or carbohydrates. Many proteins travel to the Golgi apparatus to be modified for the specific job they will do.

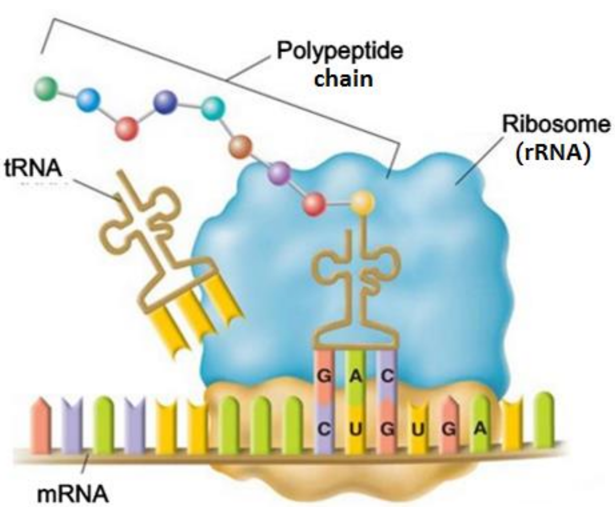

Figure 7.11: The polypeptide chain will keep growing during translation until a stop codon is reached.

Characteristics of the Genetic Code

The genetic code has a number of important characteristics.

- The genetic code is universal. All known living things have the same genetic code. This shows that all organisms share a common evolutionary history.
- The genetic code is unambiguous. Each codon codes for just one amino acid (or start or stop). What might happen if codons encoded more than one amino acid?
- The genetic code is redundant. Most amino acids are encoded by more than one codon

Lesson Summary

- The central dogma of molecular biology states that DNA contains instructions for making a protein, which are copied by RNA. RNA then uses the instructions to make a protein.
  In short: *DNA* → *RNA* → *Protein*.
- RNA differs from DNA in several ways. There three main types of RNA: messenger RNA (mRNA), ribosomal RNA (rRNA), and transfer RNA (tRNA). Each type plays a different in role in making proteins.
- Transcription is the *DNA* → *RNA* part of the central dogma of molecular biology. It occurs in the nucleus. During transcription, a copy of mRNA is made that is complementary to a strand of DNA. In eukaryotes, mRNA may be modified before it leaves the nucleus.
- The genetic code consists of the sequence of bases in DNA or RNA. Groups of three bases form codons, and each codon stands for one amino acid (or start or stop). The codons are read in sequence following the start codon until a stop codon is reached. The genetic code is universal, unambiguous, and redundant.
- Translation is the *RNA* → *protein* part of the central dogma. It occurs at a ribosome. During translation, a protein is synthesized using the codons in mRNA as a guide. All three types of RNA play a role in translation.

References/ Multimedia Resources

Textbook resource granted through licensure agreement with the CK-12 Foundation at
www.ck-12.org
CK-12 Foundation
3430 W. Bayshore Rd., Suite 101
Palo Alto, CA 94303
USA **http://www.ck12.org/saythanks**