THE NUCLEUS

The nucleus contains the cell's **genetic information**—in other words, its DNA. The DNA in eukaryotes is found in several pieces. Each piece is associated with various proteins and is called a **chromosome**. So when we talk about chromosomes, we're really just talking about the pieces of DNA in the nucleus of a cell.

Notice that we specified **eukaryotes**. **Prokaryotes** also have DNA, even though they don't have a nucleus. Their DNA is found as one large circular chromosome floating around in the cytoplasm. The DNA is still a double helix; the double helix is just joined at both ends to make a circle. We'll talk more about prokaryotes and their chromosomes later.

Even though all chromosomes contain DNA, it is *not* true that all chromosomes are identical. The nucleotide base sequences in two different pieces of DNA—two different chromosomes—can be (and are) different.

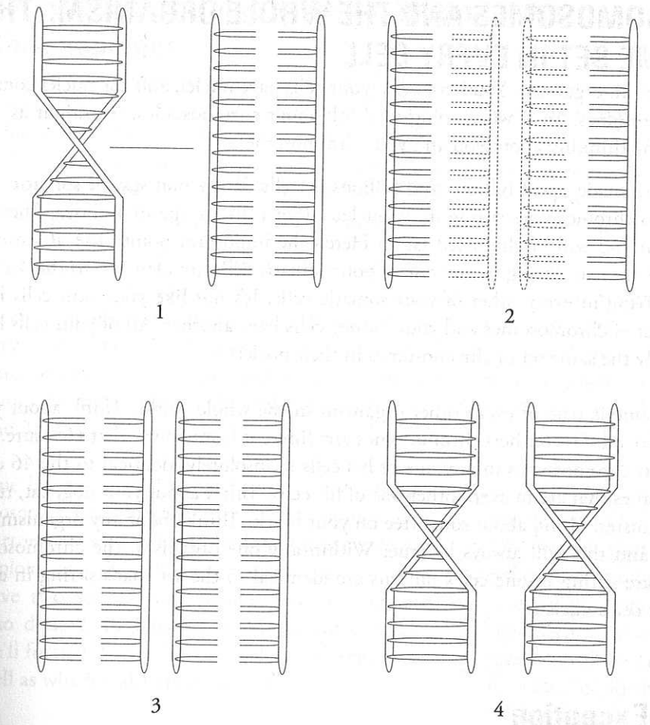
DNA Replicates Itself

One of the cool things about DNA is that it is able to make an exact copy of itself—in other words, it's able to **replicate**. This is necessary when cells want to divide. In order for the two new cells to be identical, the original cell first has to replicate its DNA, and then divide the replicated DNA evenly between the two new cells. We'll talk about that process a little later. Let's take a look now at how DNA replicates. (You might want to go back and do a quick review of DNA structure first.) There are four simple steps to replication.

> **All In the Family**
> An original strand of DNA is called the *parent DNA*. The two new DNA strands that form as a result of replication are called *daughter strands*. Replication begins at specific sites called *origins of replication*. The process of copying the parent DNA occurs in both directions from this origin site.

1. The double helix unwinds, and the two strands separate.
2. Next to each separated strand, an enzyme called **DNA polymerase** lines up nucleotides to form new second strands. The enzyme lines up nucleotides according to the base-pairing rules. Adenines are paired with thymines, and thymines are paired with adenines. Guanines are paired with cytosines, and cytosines are paired with guanines.
3. Hydrogen bonds form between the base pairs, forming the new "rungs" of the DNA ladder. Bonds form between the sugar-phosphate components of the newly aligned nucleotides so that each newly formed ladder has a new side too.
4. The new double-stranded molecules twist up into double helices.

Let's start with one DNA molecule. The strands separate, and for each strand, the cell makes a new complementary strand, and we end up with two new, but identical, DNA molecules. The original DNA molecule has replicated.



An Important SAT Biology E/M Subject Test Word: Template

Think of a replicating DNA molecule. Its strands separate. Now think about one of the separated strands. That strand causes the formation of a new complementary strand with nucleotides that are ordered according to the base-pairing rules. Another way of saying all of that is as follows:

> Each DNA strand acts as a template for the formation of a new complementary strand.

When we say that one strand acts as a **template** for the formation of another, all we mean is that its nucleotide bases direct the construction of a complementary strand that forms alongside it.

46: The Magic Number
Every cell in the human body (excluding sex cells) has 46 chromosomes.

CHROMOSOMES AND THE WHOLE ORGANISM: THE SAME SET IN EVERY CELL

You are an organism. You have cells, your cells have nuclei, and the nuclei contain chromosomes. So if we're going to think about chromosomes, we might as well start by thinking about you and your chromosomes.

You are made up of billions and billions of cells. Every non-sex, or **somatic**, cell has 46 chromosomes sitting in its nucleus. (Sex cells are sperm and ova. They are special and will be discussed later.) Here's the important point: The 46 chromosomes that are sitting in any one of your somatic cells are identical to the 46 that are sitting in every other of your somatic cells. It's not like your skin cells have one set of chromosomes and your kidney cells have another. All of your cells have exactly the same set of chromosomes in their nuclei.

The same is true of every other organism in the whole world. Think about your best friend. His (or her) chromosomes are different from yours, that's for sure. But the 46 chromosomes in any one of his cells is absolutely identical to the 46 chromosomes that are in every other one of his cells. Think about your dog, cat, turtle, or hamster. Think about some tree on your block. Think about any organism you like, and this will always be true: Within any one organism, the chromosomes that are sitting in one cell's nucleus are identical to the set that's sitting in every other cell's nucleus.

An Exception

We just taught you that, within any individual, all cells have identical chromosomes. That's true. We also told you that two different individuals do not have the same sets of chromosomes. That's almost always true. The exception is identical twins. Each identical twin has exactly the same chromosomes as the other.

Chromosomes Come in Pairs: Homologous Chromosomes

We said that human cells have a total of 46 chromosomes. But these 46 chromosomes come in pairs: Each human somatic cell has 23 pairs of chromosomes. Many other species also have their chromosomes situated in pairs. Forty-six just happens to be the number of chromosomes in human cells; hence, human cells have 23 pairs of chromosomes. Another way of saying this is to say that human cells have two sets of chromosomes, and each set consists of 23 different chromosomes.

Think about the set of 23 chromosomes that came from Dad. We'll call them 1A, 2A, 3A . . . all the way to 23A. Now think about the set of 23 chromosomes that came from Mom. We'll call them 1B, 2B, 3B, etc. Chromosomes 1A and 1B are very similar (not exactly alike, but very similar). They form one of the 23 pairs of chromosomes in your cells, and we refer to them as being **homologous**. Chromosomes 1A and 1B form a **homologous pair**. So do chromosomes 2A and 2B, 3A and 3B, 4A and 4B, etc.

Twins
Identical twins have identical chromosomes. That is because identical twins are the result of a single early-stage embryo splitting in two. You can think of this as natural cloning. Each identical twin carries the same set of chromosomes from the original egg and sperm.

One last point: Cells that have two sets of chromosomes (all chromosomes have a homologous partner) are said to be **diploid**. So human cells are diploid, and the cells of any other organism that have two sets of homologous chromosomes are also diploid. In Chapter 8, Mendelian Genetics, we'll further discuss how sex cells are developed as well as which traits are passed on.

Mom's Eyes, Dad's Nose

One set of your chromosomes came from your father in a sperm cell. One set of those chromosomes came from your mother in an ovum. When the sperm and ovum joined in fertilization, the resulting cell had two sets of 23 chromosomes each, for a total of 46 chromosomes. From that single cell, every other cell in the body was derived, so every other cell in the body has 46 chromosomes.

What Do Chromosomes *Do*?

Okay, we know that the chromosomes found in the nucleus are DNA, and we know that DNA is the genetic information of a cell, but what does that mean, exactly? What is DNA's job?

> DNA's job is to carry the instructions for making proteins.

In other words, it tells the cell in what order to connect amino acids to make proteins. But why just protein instructions? Why not carbohydrate instructions or lipid instructions? The answer is that enzymes are proteins. And if you can make enzymes, the enzymes can then make everything else. The enzymes will run all of the reactions needed to make everything else for a cell.

Human cells contain *a lot* of DNA. And not all of the DNA carries protein-building instructions. Scientists used to think that much of the DNA was not used at all. Today they are finding that this "non-coding" DNA may be involved in gene regulation. The portions of DNA that actually carry instructions (and regulatory sequences) for protein synthesis are called **genes**. So the chromosomes contain the genes that tell your body's cells how to make the enzymes (and other proteins) they need to function properly. But you don't go straight from DNA to protein. There's another step in there. Let's take a look below.

How Chromosomes Govern Protein Synthesis: Transcription and Translation

DNA contains genes, and genes tell your cells how to make protein. The step in between is the production of RNA. RNA is the "middleman" between DNA and protein.

DNA ⟶ RNA ⟶ PROTEIN

> This is known as the "central dogma of molecular biology." DNA to RNA to protein. In other words, DNA directs the synthesis of RNA, and RNA directs the synthesis of protein.

To refresh your memory, RNA is a nucleic acid that's similar to DNA.

- It has a sugar phosphate backbone, but the sugar is **ribose** (not deoxyribose, like in DNA).
- It is a polymer of nucleotides, but the bases are adenine, guanine, cytosine, and **uracil** (not thymine, like in DNA).
- It is **single-stranded** (not double-stranded, like DNA).

So DNA is a string of nucleotides, RNA is a string of nucleotides, and proteins are strings of amino acids. We can rewrite the central dogma to look like this.

DNA NUCLEOTIDES ⟶ RNA NUCLEOTIDES ⟶ AMINO ACIDS

Essentially, we're going from the language of nucleotides to the language of nucleotides, then from the language of nucleotides to the language of amino acids.

When making RNA from DNA, we're going from the language of nucleotides to the language of nucleotides. Whenever you copy something from one language to the same language, it is called a **transcription**. So DNA to RNA is a transcription.

If you missed class one day and borrowed a friend's notebook to copy over the notes into your notebook, that would be a transcription. Same information, same language, and a few slight differences (different handwriting, maybe a different pen color). Producing RNA from DNA is a similar situation. Same information (how to make protein), same language (nucleotides), slight differences (RNA is single-stranded, uses uracil, etc.).

DNA —transcription→ **RNA** ⟶ **Protein**

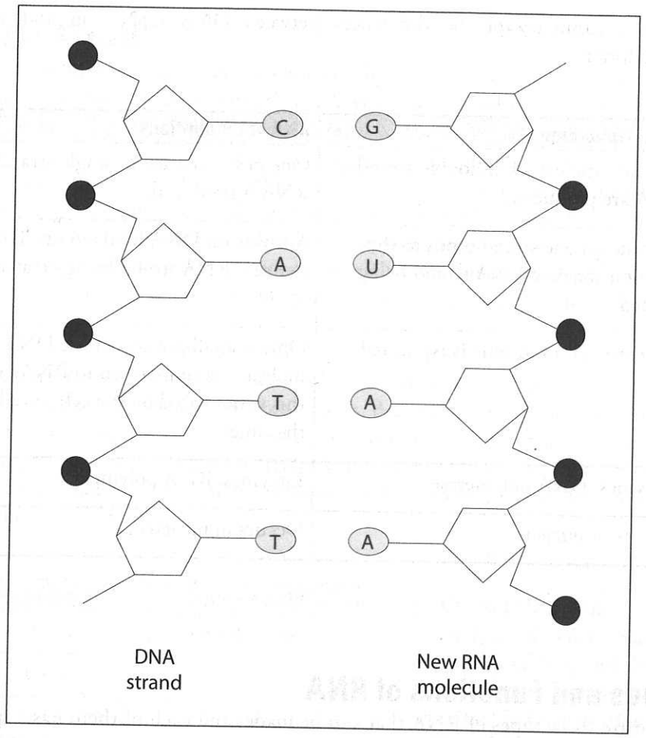
Now let's say you have a friend who lives in Russia who just HAS to have a copy of your notes. No problem. You can copy them over for your Russian friend, but what do you have to do first? Translate them. You need to translate them into a new language: Russian. Switching from one language to a new language is a translation. And when you switch from RNA (nucleotide language) to protein (amino acid language), that's a translation.

DNA —transcription→ **RNA** —translation→ **Protein**

So . . . we can say that RNA is *transcribed* from DNA, and protein is *translated* from RNA. Let's look at these processes in a little more detail.

WHERE RNA COMES FROM: TRANSCRIPTION

Remember DNA replication? The DNA molecule unwinds, and then each strand serves as a template for a new complementary strand of DNA. RNA **transcription** is very similar. The DNA molecule unwinds, and an enzyme called **RNA polymerase** creates a complementary strand of RNA, using *one* of the strands of DNA as a template. Guanine (on DNA) is paired with cytosine (on RNA). Cytosine (on DNA) is paired with guanine (on RNA). Thymine (on DNA) is paired with adenine (on RNA). And adenine (on DNA) is paired with uracil (on RNA).



The two strands (one DNA and one new RNA) do **NOT** then bond together to form a ladder. Remember: RNA is a **single-stranded** molecule. The new RNA molecule is released, and the DNA strand rejoins with its complementary partner.

One last thing to point out: In replication, the entire DNA molecule (the entire chromosome) is copied over. In transcription, only a portion of the DNA molecule is used (or transcribed) to make RNA. Remember that much of the DNA in human cells does *not* carry instructions for protein synthesis, and that the portions that *do* carry instructions for protein synthesis are called genes. So we need to transcribe only the genes. Furthermore, we need to transcribe only some of the genes—the ones that correspond to proteins needed by the cell at that time.

Here's a summary of the differences between DNA replication and RNA transcription.

DNA Replication	RNA Transcription
Two new molecules of double-stranded DNA are produced.	One new molecule of single-stranded RNA is produced.
Adenine on one strand binds to thymine on the new DNA strand being created.	Adenine on DNA binds to uracil on the new RNA strand being created.
The entire chromosome is replicated.	Only a small portion of the DNA molecule is transcribed to RNA, and this varies based on the cell's needs at the time.
Enzymes: DNA polymerase	Enzymes: RNA polymerase
Occurs in nucleus.	Occurs in nucleus.

Types and Functions of RNA

There are three types of RNA that can be made, and each of them has a special role in protein synthesis.

1. **mRNA**: *m* stands for "messenger." Messenger RNA is the RNA that actually carries the information for protein synthesis (in the form of a nucleotide sequence) from DNA in the nucleus to the ribosomes in the cytoplasm.
2. **rRNA**: *r* stands for "ribosomal." Ribosomal RNA interacts with the cell's ribosomes to make them functional. The ribosomes are made of rRNA and protein. The ribosomes are the site of protein synthesis.
3. **tRNA**: *t* stands for "transfer." Transfer RNA carries amino acids from the cytoplasm to the ribosomes during protein synthesis.

Replication and transcription occur in the nucleus. But translation (protein synthesis) occurs in the cytoplasm of the cell. The three types of RNA must leave the nucleus and enter the cytoplasm for protein synthesis to take place.

TRANSLATION

If you were going to translate your notes from English to Russian, you would need a translation dictionary, something that tells you how a word in English is written in Russian. The **translation** of protein is no different. When you move from the language of nucleotides to the language of amino acids, you still need a "dictionary"; something that tells you which nucleotides correspond to which amino acids. More specifically, you need something to tell you which *sequence of three* nucleotides corresponds to which *one* amino acid. A sequence of three nucleotides is called a **codon**, and the order of codons on mRNA specifies the order of amino acids in a protein. The dictionary for protein translation is called the **Genetic Code**.

Because there are four possible nucleotide bases, and codons are groups of three bases, there are 64 ($4 \times 4 \times 4$) possible codons. Because there are only 20 different amino acids, some of the amino acids are coded for by more than one codon. The Genetic Code is nothing more, really, than a list of the 64 possible codons and the amino acids to which they correspond. The following chart shows a portion of the Genetic Code.

Codon	Amino Acid
AUG	methionine
CUU	leucine
GCA	alanine
UUG	leucine
CAG	glutamine
CGA	arginine

Cracking the Code

In 1961, Marshall Nirenberg, a chemist, began to explore codons and the amino acids to which they correspond. He made an RNA molecule that contained only nucleotides of uracil. This RNA, therefore, had the codon UUU repeated over and over. He then placed this into a test tube with all 20 amino acids, and found that a single kind of protein was translated: polyphenylalanine.

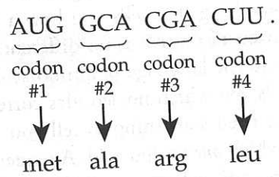
So, to figure out the order of amino acids in a protein, all you have to do is look at the sequence of codons on the mRNA. Suppose you had a piece of mRNA with the following sequence:

A U G G C A C G A C U U . . .

The codons are read in **nonoverlapping sequence**, like this:

AUG	GCA	CGA	CUU . . .
codon #1	codon #2	codon #3	codon #4

So, for this piece of mRNA, the amino acid sequence would be



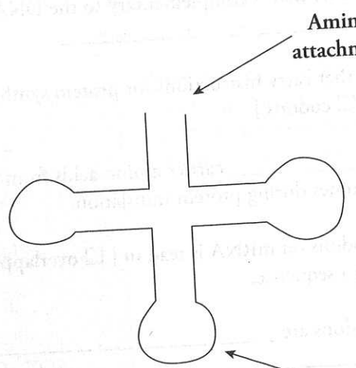
Now you can see how a nucleotide sequence on DNA can specify a nucleotide sequence on RNA, and how that same nucleotide sequence can specify an amino acid sequence in a protein.

Stop and Go

- The codon AUG (methionine) is known as the "start" codon because it's the first codon on all mRNA, and methionine is the first amino acid in all proteins.
- Three of the 64 possible codons do not specify an amino acid. They specify "stop." In other words, "stop translating, the protein is finished." The three stop codons are **UAA**, **UGA**, and **UAG**.

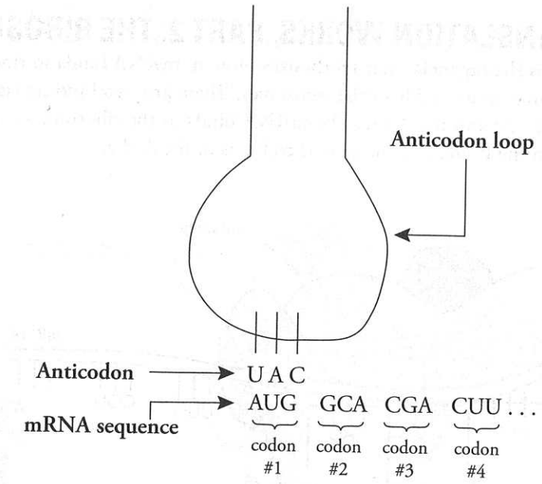
HOW TRANSLATION WORKS, PART 1: tRNA

tRNA is the molecule that carries amino acids from the cytoplasm to the ribosomes. It has a specific, three-dimensional shape (it looks somewhat like a pistol), and if it is flattened out, it has a "cloverleaf" shape, something like this (the nucleotide sequence has been eliminated for clarity):



The **anticodon** (contained within the **anticodon loop**) is a special region on the tRNA molecule that can base-pair with codons on mRNA. *The anticodon must be complementary to a codon to base-pair with it.*

For example, if a codon has the sequence AUG, the only tRNA molecules that can base-pair with that codon are the tRNAs that have UAC for an anticodon.



On the other end of the tRNA, an amino acid can attach. The amino acid that attaches corresponds with the codon that the tRNA anticodon can base-pair with. For example, in the situation above, the tRNA anticodon could base-pair with the codon AUG. The codon AUG codes for methionine. So the amino acid that would attach to that particular tRNA would be methionine. Here are some other examples:

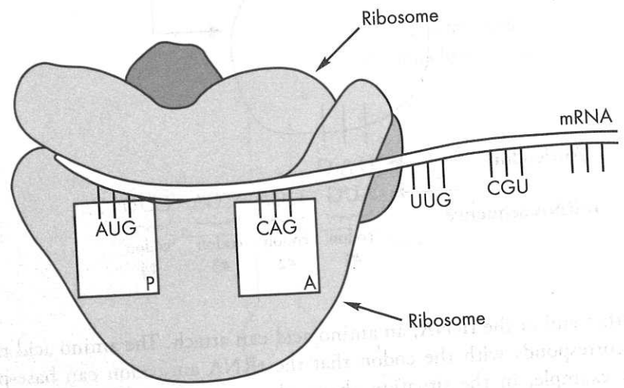
mRNA codon	tRNA anticodon	Attached amino acid
AUG	UAC	methionine
CUU	GAA	leucine
GCA	CGU	alanine
UUG	AAC	leucine
CAG	GUC	glutamine
CGA	GCU	arginine

tRNA: A Quick Review

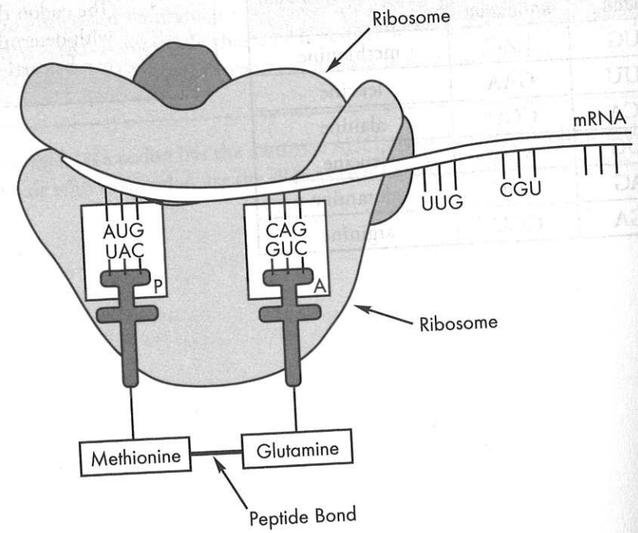
- tRNA molecules carry amino acids from the cytoplasm to the ribosomes.
- tRNA has an "anticodon" that's complementary to mRNA codons and that can base-pair with them.
- The codon that tRNA can base-pair with determines the amino acid that can be carried by the tRNA.

HOW TRANSLATION WORKS, PART 2: THE RIBOSOME

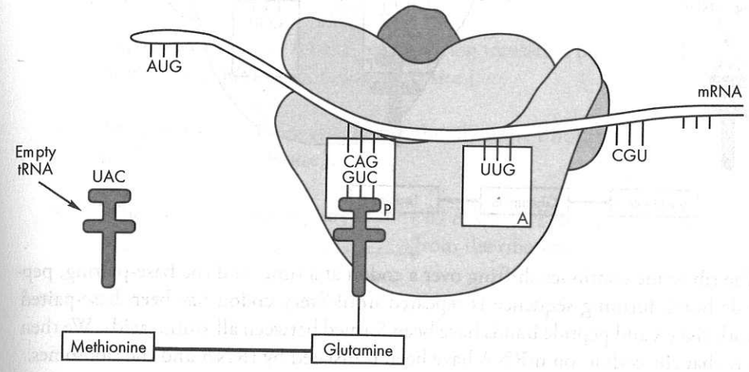
The **ribosome** is the organelle that synthesizes protein. mRNA binds to ribosomes, and tRNAs carry amino acids to the ribosomes. There are two binding sites on a ribosome: the P-site and the A-site. The mRNA binds to the ribosome so that the first codon is in the P-site, and the second codon is in the A-site.



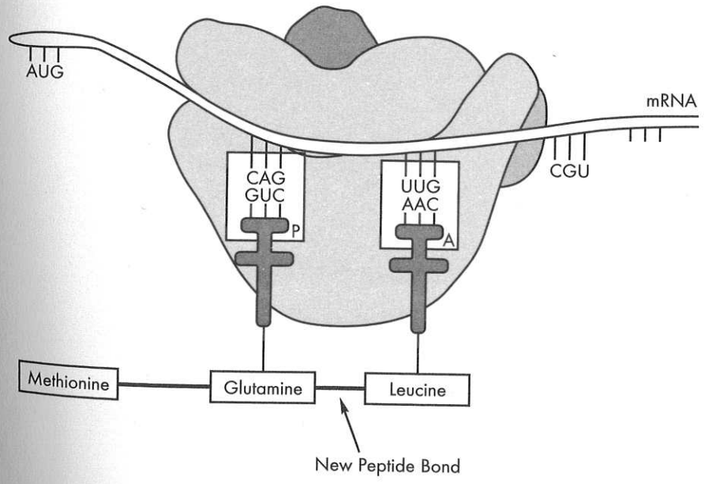
Once the mRNA is bound, tRNA, carrying the appropriate amino acids, comes and base-pairs with the codons on mRNA. Then the ribosome forms a **peptide bond** between the two amino acids.



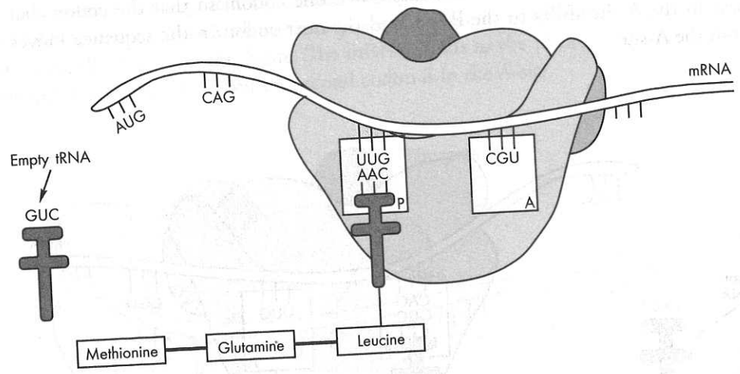
Once the peptide bond is formed, the first (and now empty) tRNA is released from the ribosome. It is free to return to the cytoplasm and bind to another amino acid. The ribosome, meanwhile, slides down one codon, so that the codon that was in the A-site shifts to the P-site, and the next codon in the sequence moves into the A-site.



Now the sequence is just repeated. Another tRNA (carrying the appropriate amino acid) moves in and base-pairs with the codon in the A-site. The ribosome forms a peptide bond between the two amino acids (the one attached to the tRNA in the P-site and the one attached to the tRNA in the A-site), and the tRNA in the P-site is now released and returns to the cytoplasm.



And the ribosome shifts over again.



The ribosome continues shifting over a codon at a time, and the base-pairing, peptide bond–forming sequence is repeated until every codon has been base-paired with tRNA and peptide bonds have been formed between all amino acids. We then say that the codons on mRNA have been translated by tRNA and the ribosomes.

When a stop codon (UAG, UGA, or UAA) appears in the A-site, the final tRNA is released from the ribosome, the completed protein is released, and translation is complete.

By the way, the "P" in P-site stands for **peptide**. This is where the growing peptide is attached to the ribosome. And the "A" in A-site stands for **amino acid**. This is where the next amino acid is added to the growing protein.