GENETICS

Information from Khan academy

A ROUGH SUMMARY

-but still quite long

DNA

Deoxyribonucleic acid

DNA is made of 4 nucleotides- Adenine, Thymine, Guanine and Cytosine. Adenine and Guanine are single-ring nucleotides, and Cytosine and Thymine are double-ringed nucleotides. Adenine and Thymine form hydrogen bonds, and Guanine and Cytosine form hydrogen bonds. Each nucleotide is made up of a phosphate and five carbon sugar backbone, and it also has a nitrogenous base, which is where genetic information is stored.

RNA

Ribonucleic acid

There are four types of RNA- messenger RNA (mRNA), transfer RNA(tRNA), ribosomal RNA(rRNA) and micro RNA(miRNA). RNA is still made up of four nucleotides- Adenine, Uracil, Guanine and Cytosine. Uracil is still a double-ringed nucleotide.

Nucleic acids, macromolecules made out of units called nucleotides, come in two naturally occurring varieties: **deoxyribonucleic acid** (**DNA**) and **ribonucleic acid** (**RNA**). DNA is the genetic material found in living organisms, all the way from single-celled bacteria to multicellular mammal. Some viruses use RNA, not DNA, as their genetic material, but aren't technically considered to be alive (since they cannot reproduce without help from a host).

In eukaryotes, such as plants and animals, DNA is found in the **nucleus**. In prokaryotes, such as bacteria, the DNA is not enclosed in a membranous envelope, although it's located in a specialized cell region called the **nucleoid**.

Many genes encode protein products, meaning that they specify the sequence of amino acids used to build a particular protein. Before this information can be used for protein synthesis, however, an RNA copy (transcript) of the gene must first be made. This type of RNA is called a **messenger RNA** (**mRNA**), as it serves as a messenger between DNA and the ribosomes, molecular machines that read mRNA sequences and use them to build proteins.

DNA and RNA are polymers, and are made up of monomers known as **nucleotides**. When these monomers combine, the resulting chain is called a **polynucleotide** (*poly-* = "many").

Each nucleotide is made up of three parts: a nitrogen-containing ring structure called a nitrogenous base, a five-carbon sugar, and at least one

phosphate group. The sugar molecule has a central position in the nucleotide, with the base attached to one of its carbons and the phosphate group (or groups) attached to another. Let's look at each part of a nucleotide in turn.

Bases include the pyrimidine bases (cytosine, thymine in DNA, and uracil in RNA, one ring) and the purine bases (adenine and guanine, two rings). The phosphate group is attached to the 5' carbon. The 2' carbon bears a hydroxyl group in ribose, but no hydroxyl (just hydrogen) in deoxyribose.

In addition to having slightly different sets of bases, DNA and RNA nucleotides also have slightly different sugars. The five-carbon sugar in DNA is called **deoxyribose**, while in RNA, the sugar is **ribose**. These two are very similar in structure, with just one difference: the second carbon of ribose bears a hydroxyl group, while the equivalent carbon of deoxyribose has a hydrogen instead.

Phosphate

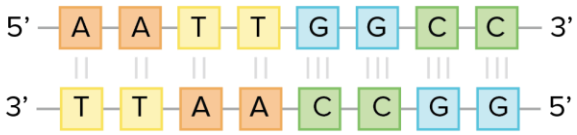
Nucleotides may have a single phosphate group, or a chain of up to three phosphate groups, attached to the 5' carbon of the sugar. Some chemistry sources use the term "nucleotide" only for the single-phosphate case, but in molecular biology, the broader definition is generally accepted

In a cell, a nucleotide about to be added to the end of a polynucleotide chain will bear a series of three phosphate groups. When the nucleotide joins the growing DNA or RNA chain, it loses two phosphate groups. So, in a chain of DNA or RNA, each nucleotide has just one phosphate group.

The two strands of the helix run in opposite directions, meaning that the 5′ end of one strand is paired up with the 3′ end of its matching strand. (This is referred to as **antiparallel** orientation and is important for the copying of DNA.)

So, can any two bases decide to get together and form a pair in the double helix? The answer is a definite no. Because of the sizes and functional groups of the bases, base pairing is highly specific: A can only pair with T, and G can only pair with C, as shown below. This means that the two strands of a DNA double helix have a very predictable relationship to each other.

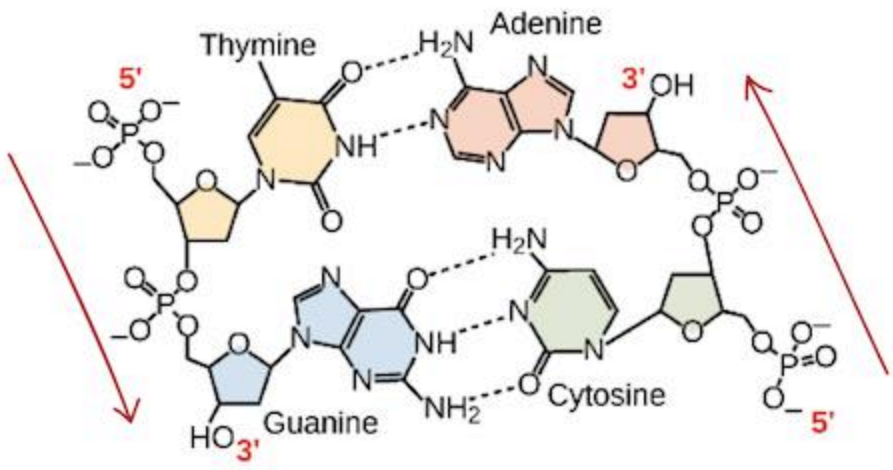
For instance, if you know that the sequence of one strand is 5'-AATTGGCC-3', the complementary strand must have the sequence 3'-TTAACCGG-5'. This allows each base to match up with its partner:



5'-AATTGGCC-3' 3'-TTAACCGG-5'

These two strands are complementary, with each base in one sticking to its partner on the other. The A-T pairs are connected by two hydrogen bonds, while the G-C pairs are connected by three hydrogen bonds.

When two DNA sequences match in this way, such that they can stick to each other in an antiparallel fashion and form a helix, they are said to be **complementary**.



Hydrogen bonding between complementary bases holds DNA strands together in a double helix of antiparallel strands. Thymine forms two hydrogen bonds with adenine, and guanine forms three hydrogen bonds with cytosine.

Image modified from OpenStax Biology.

Properties of RNA

Ribonucleic acid (RNA), unlike DNA, is usually single-stranded. A nucleotide in an RNA chain will contain ribose (the five-carbon sugar), one of the four nitrogenous bases (A, U, G, or C), and a phosphate group.

Summary: Features of DNA and RNA

	DNA	RNA
Function	Repository of genetic information	Involved in protein synthesis and gene regulation; carrier of genetic information in some viruses

	DNA	RNA
Sugar	Deoxyribose	Ribose
Structure	Double helix	Usually single-stranded
Bases	C, T, A, G	C, U, A, G

Table modified from OpenStax Biology.

Prokaryotes are microscopic organisms belonging to the domains Bacteria and Archaea, which are two out of the three major domains of life. (Eukarya, the third, contains all **eukaryotes**, including animals, plants, and fungi.) Bacteria and archaea are single-celled, while most eukaryotes are multicellular.

Fossils show that prokaryotes were already here on Earth 3.53.53.5 billion years ago, and scientists think that prokaryotic ancestors gave rise to all of the life forms present on Earth today

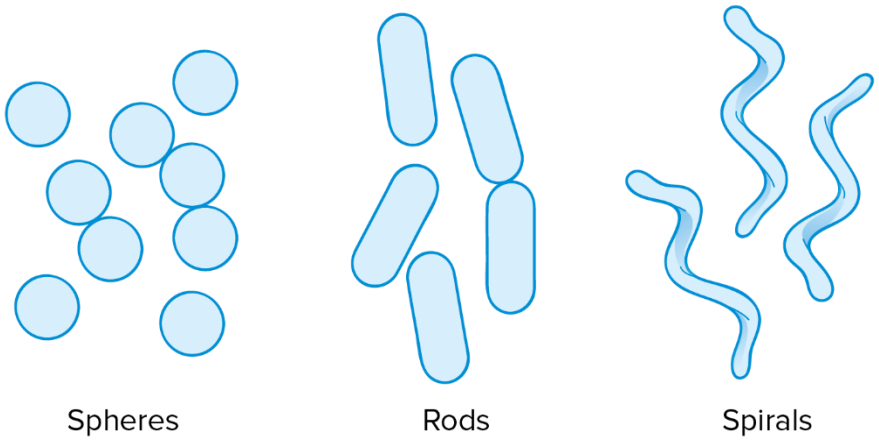
Prokaryotes vs. eukaryotes

Prokaryotes and eukaryotes are quite different. That may be obvious when comparing humans to bacteria. However, it's less obvious when we're comparing a bacterium to a yeast (which is tiny and unicellular, but eukaryotic). What actually separates these categories of organisms?

The most fundamental differences between prokaryotes and eukaryotes relate to how their cells are set up. Specifically:

- Eukaryotic cells have a **nucleus** while prokaryotic cells don't. This is the feature that formally separates the two groups.

- Eukaryotes usually have other membrane-bound organelles in addition to the nucleus, while prokaryotes don't.
- Cells in general are small, but prokaryotic cells are *really* small.
    - Many prokaryotic cells have sphere, rod, or spiral shapes



Prokaryotic cells are typically shaped as either spheres (called cocci), rods (called bacilli), or spirals.

Image modified from "Bacterial morphology diagram," by Mariana Ruiz Villareal (public domain).

REPLICATION OF DNA

The basic idea

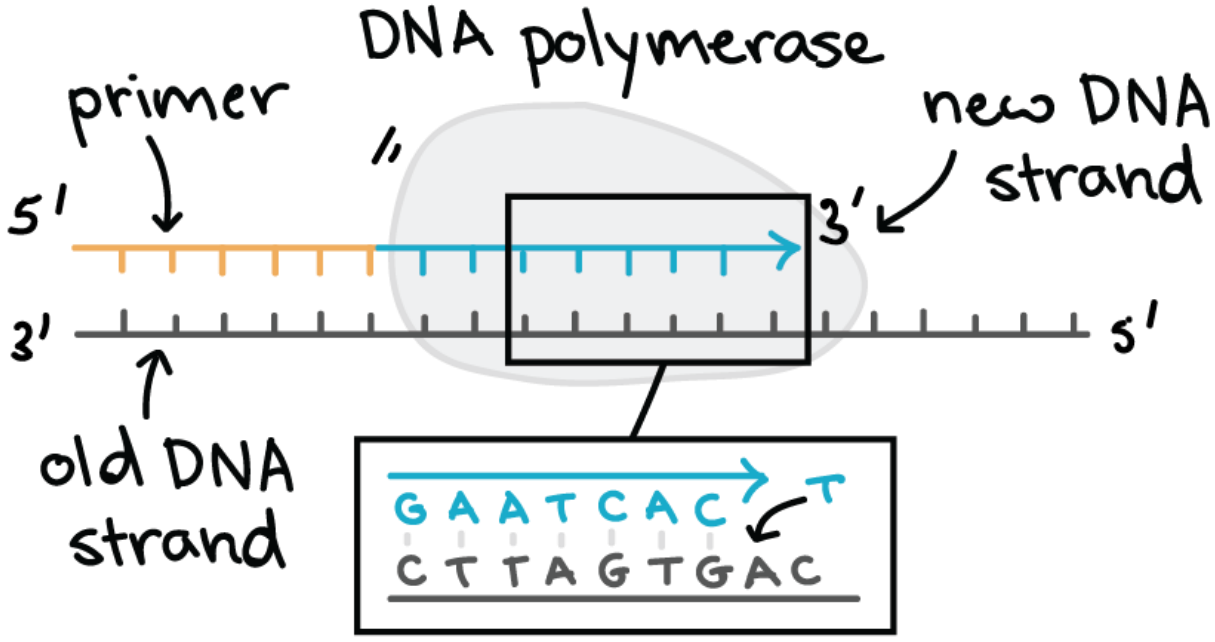
DNA replication is **semiconservative**, meaning that each strand in the DNA double helix acts as a template for the synthesis of a new, complementary strand.

This process takes us from one starting molecule to two "daughter" molecules, with each newly formed double helix containing one new and one old strand.

Cells need to copy their DNA very quickly, and with very few errors (or risk problems such as cancer). To do so, they use a variety of enzymes and proteins, which work together to make sure DNA replication is performed smoothly and accurately.

DNA polymerase

One of the key molecules in DNA replication is the enzyme **DNA polymerase**. DNA polymerases are responsible for synthesizing DNA: they add nucleotides one by one to the growing DNA chain, incorporating only those that are complementary to the template.



Here are some key features of DNA polymerases:

- They always need a template
- They can only add nucleotides to the 3' end of a DNA strand
- They can't start making a DNA chain from scratch, but require a pre-existing chain or short stretch of nucleotides called a **primer**
- They **proofread**, or check their work, removing the vast majority of "wrong" nucleotides that are accidentally added to the chain

The addition of nucleotides requires energy. This energy comes from the nucleotides themselves, which have three phosphates attached to them (much like the energy-carrying molecule ATP). When the bond between

phosphates is broken, the energy released is used to form a bond between the incoming nucleotide and the growing chain.

[See the polymerization reaction]

In prokaryotes such as *E. coli*, there are two main DNA polymerases involved in DNA replication: DNA pol III (the major DNA-maker), and DNA pol I, which plays a crucial supporting role we'll examine later.

Starting DNA replication

How do DNA polymerases and other replication factors know where to begin? Replication always starts at specific locations on the DNA, which are called **origins of replication** and are recognized by their sequence.

E. coli, like most bacteria, has a single origin of replication on its chromosome. The origin is about 245245245 base pairs long and has mostly A/T base pairs (which are held together by fewer hydrogen bonds than G/C base pairs), making the DNA strands easier to separate.

Specialized proteins recognize the origin and open up the DNA. As the DNA opens, two Y-shaped structures called **replication forks** are formed, together making up what's called a **replication bubble**. The replication forks will move in opposite directions as replication proceeds.

Bacterial chromosome. The double-stranded DNA of the circular bacteria chromosome is opened at the origin of replication, forming a replication bubble. Each end of the bubble is a replication fork, a Y-shaped junction where double-stranded DNA is separated into two single strands. New DNA complementary to each single strand is synthesized at each

replication fork. The two forks move in opposite directions around the circumference of the bacterial chromosome, creating a larger and larger replication bubble that grows at both ends.

How does replication actually get going at the forks? **Helicase** is the first replication enzyme to load on at the origin of replication. Helicase's job is to move the replication forks forward by "unwinding" the DNA (breaking the hydrogen bonds between the nitrogenous base pairs).

Proteins called **single-strand binding proteins** coat the separated strands of DNA near the replication fork, keeping them from coming back together into a double helix.

Primers and primase

DNA polymerases can only add nucleotides to the 3' end of an existing DNA strand. (They use the free -OH group found at the 3' end as a "hook," adding a nucleotide to this group in the polymerization reaction.) How, then, does DNA polymerase add the first nucleotide at a new replication fork?

Alone, it can't! The problem is solved with the help of an enzyme called **primase**. Primase makes an RNA **primer**, or short stretch of nucleic acid complementary to the template, that provides a 3' end for DNA polymerase to work on. A typical primer is about five to ten nucleotides long. The primer *primes* DNA synthesis.

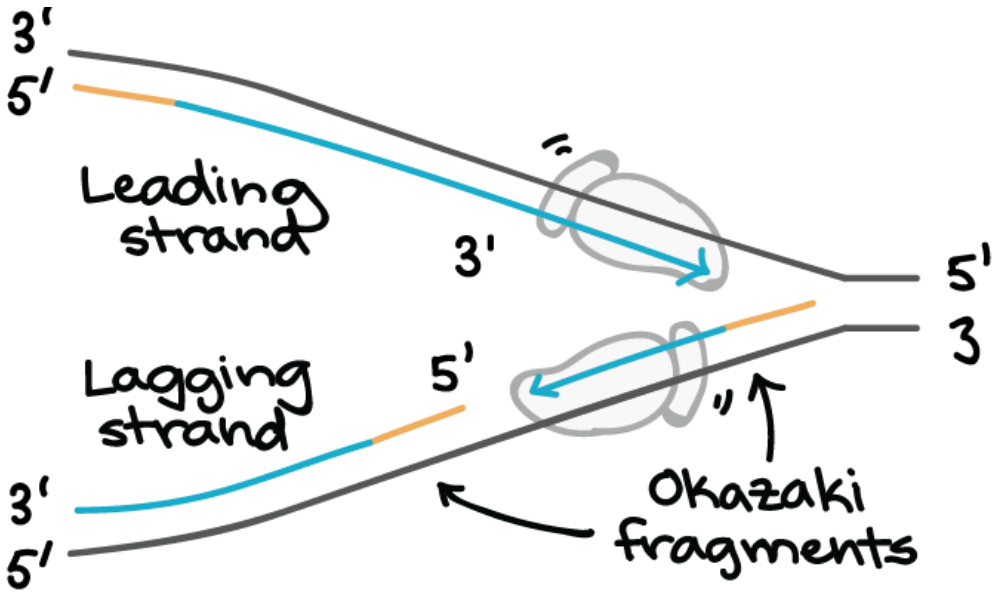
Once the RNA primer is in place, DNA polymerase "extends" it, adding nucleotides one by one to make a new DNA strand that's complementary to the template strand.

Leading and lagging strands

In *E. coli*, the DNA polymerase that handles most of the synthesis is DNA polymerase III. There are two molecules of DNA polymerase III at a replication fork, each of them hard at work on one of the two new DNA strands.

DNA polymerases can only make DNA in the 5' to 3' direction, and this poses a problem during replication. A DNA double helix is always anti-parallel; in other words, one strand runs in the 5' to 3' direction, while the other runs in the 3' to 5' direction. This makes it necessary for the two new strands, which are also antiparallel to their templates, to be made in slightly different ways.

One new strand, which runs 5' to 3' towards the replication fork, is the easy one. This strand is made continuously, because the DNA polymerase is moving in the same direction as the replication fork. This continuously synthesized strand is called the **leading strand**.



The other new strand, which runs 5' to 3' away from the fork, is trickier. This strand is made in fragments because, as the fork moves forward, the DNA polymerase (which is moving away from the fork) must come off and reattach on the newly exposed DNA. This tricky strand, which is made in fragments, is called the **lagging strand**.

The small fragments are called **Okazaki fragments**, named for the Japanese scientist who discovered them. The leading strand can be extended from one primer alone, whereas the lagging strand needs a new primer for each of the short Okazaki fragments.

The maintenance and cleanup crew

Some other proteins and enzymes, in addition the main ones above, are needed to keep DNA replication running smoothly. One is a protein called the **sliding clamp**, which holds DNA polymerase III molecules in place as they synthesize DNA. The sliding clamp is a ring-shaped protein and keeps the DNA polymerase of the lagging strand from floating off when it re-starts at a new Okazaki fragment.

Topoisomerase also plays an important maintenance role during DNA replication. This enzyme prevents the DNA double helix ahead of the replication fork from getting too tightly wound as the DNA is opened up. It acts by making temporary nicks in the helix to release the tension, then sealing the nicks to avoid permanent damage.

Finally, there is a little cleanup work to do if we want DNA that doesn't contain any RNA or gaps. The RNA primers are removed and replaced by DNA through the activity of **DNA polymerase I**, the other polymerase

involved in replication. The nicks that remain after the primers are replaced get sealed by the enzyme **DNA ligase**.

Summary of DNA replication in *E. coli*

Let's zoom out and see how the enzymes and proteins involved in replication work together to synthesize new DNA.

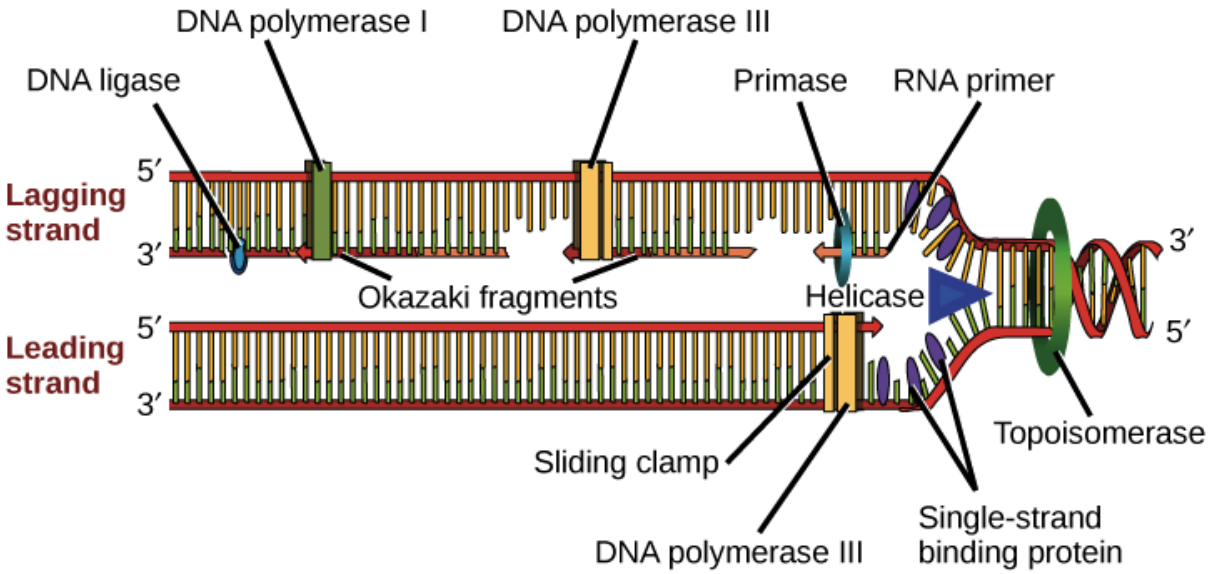


Illustration shows the replication fork. Helicase unwinds the helix, and single-strand binding proteins prevent the helix from re-forming. Topoisomerase prevents the DNA from getting too tightly coiled ahead of the replication fork. DNA primase forms an RNA primer, and DNA polymerase extends the DNA strand from the RNA primer. DNA synthesis occurs only in the 5' to 3' direction. On the leading strand, DNA synthesis occurs continuously. On the lagging strand, DNA synthesis restarts many times as the helix unwinds, resulting in many short fragments called "Okazaki fragments." DNA ligase joins the Okazaki fragments together into a single DNA molecule.

- **Helicase** opens up the DNA at the replication fork.

- **Single-strand binding proteins** coat the DNA around the replication fork to prevent rewinding of the DNA.
- **Topoisomerase** works at the region ahead of the replication fork to prevent supercoiling.
- **Primase** synthesizes RNA primers complementary to the DNA strand.
- **DNA polymerase III** extends the primers, adding on to the 3' end, to make the bulk of the new DNA.
- RNA primers are removed and replaced with DNA-by-DNA **polymerase I**.
- The gaps between DNA fragments are sealed by **DNA ligase**.

DNA replication in eukaryotes

The basics of DNA replication are similar between bacteria and eukaryotes such as humans, but there are also some differences:

- Eukaryotes usually have multiple linear chromosomes, each with multiple origins of replication. Humans can have up to 100,100,100,000000000 origins of replication.
- Most of the *E. coli* enzymes have counterparts in eukaryotic DNA replication, but a single enzyme in *E. coli* may be represented by multiple enzymes in eukaryotes. For instance, there are five human DNA polymerases with important roles in replication.
- Most eukaryotic chromosomes are linear. Because of the way the lagging strand is made, some DNA is lost from the ends of linear chromosomes (the telomeres) in each round of replication.

FROM DNA TO PROTEIN

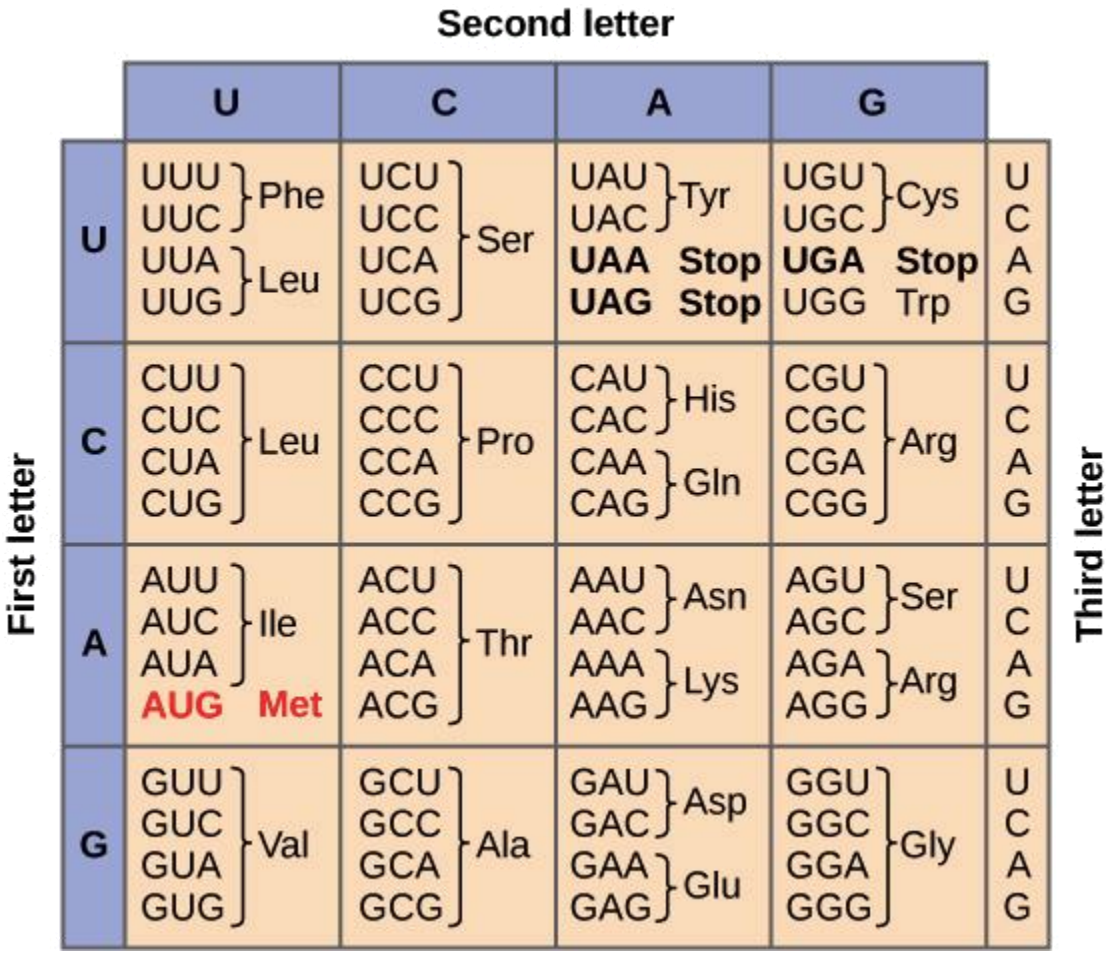
A gene is used to build a protein in a two-step process:

- Step 1: **transcription**! Here, the DNA sequence of a gene is "rewritten" in the form of RNA. In eukaryote, the RNA is processed (and often has a few bits snipped out of it) to make the final product, mRNA.
- Step 2: **translation**! In this stage, the mRNA is "decoded" to build a protein (or a chunk/subunit of a protein) that contains a specific series of amino acids.

The central dogma of molecular biology states that information flows from DNA (genes) to mRNA through the process of transcription, and then to proteins through the process of translation.

The genetic code

During translation, a cell “reads” the information in a messenger RNA (mRNA) and uses it to build a protein. Actually, to be a little more technical, an mRNA doesn’t always encode a whole protein. Instead, what we can confidently say is that it always encodes a **polypeptide**, or chain of amino acids.



This is a Genetic code table. Each three-letter sequence of mRNA nucleotides corresponds to a specific amino acid, or to a stop codon. UGA, UAA, and UAG are stop codons. AUG is the codon for methionine, and is also the start codon.

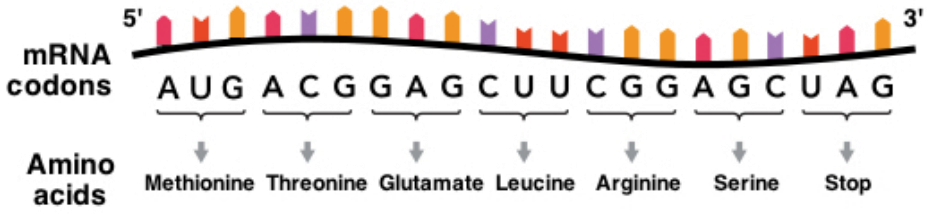
Image credit: "The genetic code," by OpenStax College, Biology (CC BY 3.0).

In an mRNA, the instructions for building a polypeptide are RNA nucleotides (As, Us, Cs, and Gs) read in groups of three. These groups of three are called **codons**.

There are 616161 codons for amino acids, and each of them is "read" to specify a certain amino acid out of the 202020 commonly found in

proteins. One codon, AUG, specifies the amino acid methionine and also acts as a **start codon** to signal the start of protein construction.

There are three more codons that do *not* specify amino acids. These **stop codons**, UAA, UAG, and UGA, tell the cell when a polypeptide is complete. All together, this collection of codon-amino acid relationships is called the **genetic code**, because it lets cells "decode" an mRNA into a chain of amino acids.



Each mRNA contains a series of codons (nucleotide triplets) that each specifies an amino acid. The correspondence between mRNA codons and amino acids is called the genetic code.

5' AUG - Methionine ACG - Threonine GAG - Glutamate CUU - Leucine CGG - Arginine AGC - Serine UAG - Stop 3'

Image modified from "RNA-codons-aminoacids," by Thomas Splettstoesser (CC BY-SA 4.0). The modified image is licensed under a CC BY-SA 4.0 license.

Overview of translation

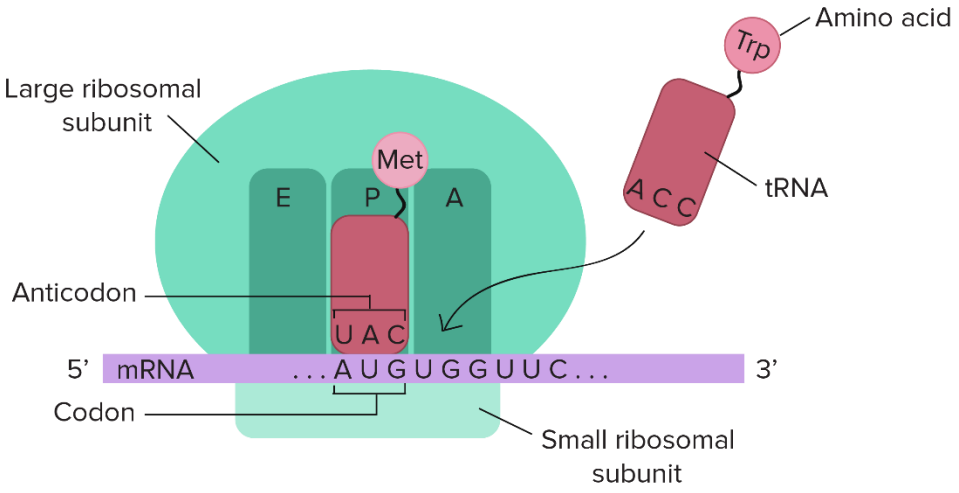
How is an mRNA "read" to make a polypeptide? Two types of molecules with key roles in translation are tRNAs and ribosomes.

Transfer RNAs (tRNAs)

Transfer RNAs, or **tRNAs**, are molecular "bridges" that connect mRNA codons to the amino acids they encode. One end of each tRNA has a

sequence of three nucleotides called an **anticodon**, which can bind to specific mRNA codons. The other end of the tRNA carries the amino acid specified by the codons.

There are many different types of tRNAs. Each type reads one or a few codons and brings the right amino acid matching those codons.



Ribosomes are composed of a small and large subunit and have three sites where tRNAs can bind to an mRNA (the A, P, and E sites). Each tRNA carries a specific amino acid and binds to an mRNA codon that is complementary to its anticodon.

Image modified from "Translation: Figure 3," by OpenStax College, Biology (CC BY 4.0).

Ribosomes

Ribosomes are the structures where polypeptides (proteins) are built. They are made up of protein and RNA (**ribosomal RNA**, or **rRNA**). Each ribosome has two subunits, a large one and a small one, which come together around an mRNA—kind of like the two halves of a hamburger bun coming together around the patty.

The ribosome provides a set of handy slots where tRNAs can find their matching codons on the mRNA template and deliver their amino acids. These slots are called the A, P, and E sites. Not only that, but the ribosome also acts as an enzyme, catalyzing the chemical reaction that links amino acids together to make a chain.

Steps of translation

Getting started: Initiation

In **initiation**, the ribosome assembles around the mRNA to be read and the first tRNA (carrying the amino acid methionine, which matches the start codon, AUG). This setup, called the initiation complex, is needed in order for translation to get started.

Extending the chain: Elongation

Elongation is the stage where the amino acid chain gets **long**er. In elongation, the mRNA is read one codon at a time, and the amino acid matching each codon is added to a growing protein chain.

Each time a new codon is exposed:

- A matching tRNA binds to the codon.
- The existing amino acid chain (polypeptide) is linked onto the amino acid of the tRNA via a chemical reaction.
- The mRNA is shifted one codon over in the ribosome, exposing a new codon for reading.

Elongation has three stages:

1) The anticodon of an incoming tRNA pairs with the mRNA codon exposed in the A site.

2) A peptide bond is formed between the new amino acid (in the A site) and the previously added amino acid (in the P site), transferring the polypeptide from the P site to the A site.

3) The ribosome moves one codon down on the mRNA. The tRNA in the A site (carrying the polypeptide) shifts to the P site. The tRNA in the P site shifts to the E site and exits the ribosome.

During elongation, tRNAs move through the A, P, and E sites of the ribosome, as shown above. This process repeats many times as new codons are read and new amino acids are added to the chain.

Finishing up: Termination

Termination is the stage in which the finished polypeptide chain is released. It begins when a stop codon (UAG, UAA, or UGA) enters the ribosome, triggering a series of events that separate the chain from its tRNA and allow it to drift out of the ribosome.

After termination, the polypeptide may still need to fold into the right 3D shape, undergo processing (such as the removal of amino acids), get shipped to the right place in the cell, or combine with other polypeptides before it can do its job as a functional protein.

GENE EXPRESSION & REGULATION

How is gene expression regulated?

There are various forms of **gene regulation**, that is, mechanisms for controlling which genes get expressed and at what levels. However, a lot of gene regulation occurs at the level of transcription.

Bacteria have specific regulatory molecules that control whether a particular gene will be transcribed into mRNA. Often, these molecules act by binding to DNA near the gene and helping or blocking the transcription enzyme, RNA polymerase.

In bacteria, genes are often found in operons

In bacteria, related genes are often found in a cluster on the chromosome, where they are transcribed from one **promoter** (RNA polymerase binding site) as a single unit. Such a cluster of genes under control of a single promoter is known as an **operon**. Operons are common in bacteria, but they are rare in eukaryotes such as humans.

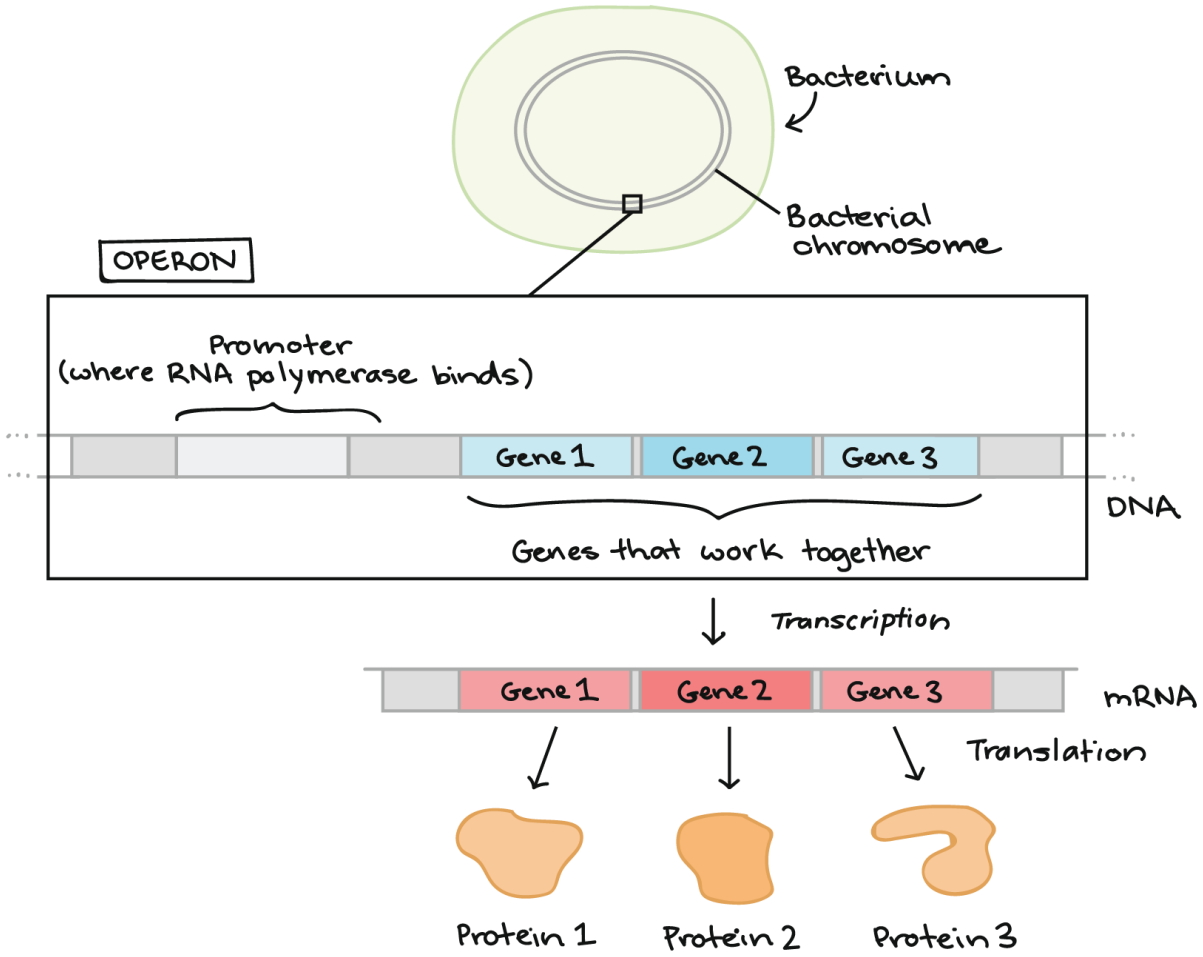


Diagram illustrating what an operon is. At the top of the diagram, we see a bacterial cell with a circular bacterial chromosome inside it. We zoom in on a small segment of the chromosome and see that it is an operon. The DNA of the operon contains three genes, Gene 1, Gene 2, and Gene 3, which are found in a row in the DNA. They are under control of a single promoter (site where RNA polymerase binds) and they are transcribed together to make a single mRNA that has contains sequences coding for all three genes. When the mRNA is translated, the three different coding sequences of the mRNA are read separately, making three different proteins (Protein 1, Protein 2, and Protein 3).

Note: The operon does not consist of just the three genes. Instead, it also includes the promoter and other regulatory sequences that regulate expression of the genes.

In general, an operon will contain genes that function in the same process. For instance, a well-studied operon called the *lac* operon contains genes that encode proteins involved in uptake and metabolism of a particular sugar, lactose. Operons allow the cell to efficiently express sets of genes whose products are needed at the same time.

Anatomy of an operon

Operons aren't just made up of the coding sequences of genes. Instead, they also contain **regulatory DNA sequences** that control transcription of the operon. Typically, these sequences are binding sites for **regulatory proteins**, which control how much the operon is transcribed. The promoter, or site where RNA polymerase binds, is one example of a regulatory DNA sequence.

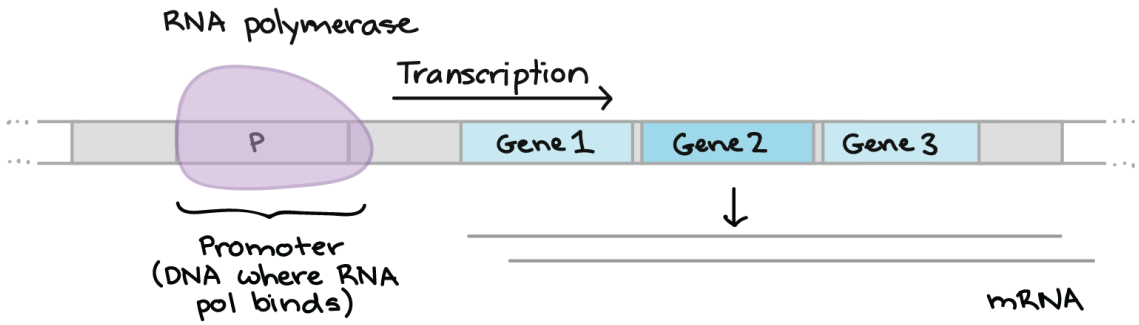


Diagram illustrating that the promoter is the site where RNA polymerase binds. The promoter is found in the DNA of the operon, upstream of

(before) the genes. When the RNA polymerase binds to the promoter, it transcribes the operon and makes some mRNAs.

Most operons have other regulatory DNA sequences in addition to the promoter. These sequences are binding sites for regulatory proteins that turn expression of the operon "up" or "down."

- Some regulatory proteins are **repressors** that bind to pieces of DNA called **operators**. When bound to its operator, a repressor reduces transcription (e.g., by blocking RNA polymerase from moving forward on the DNA).

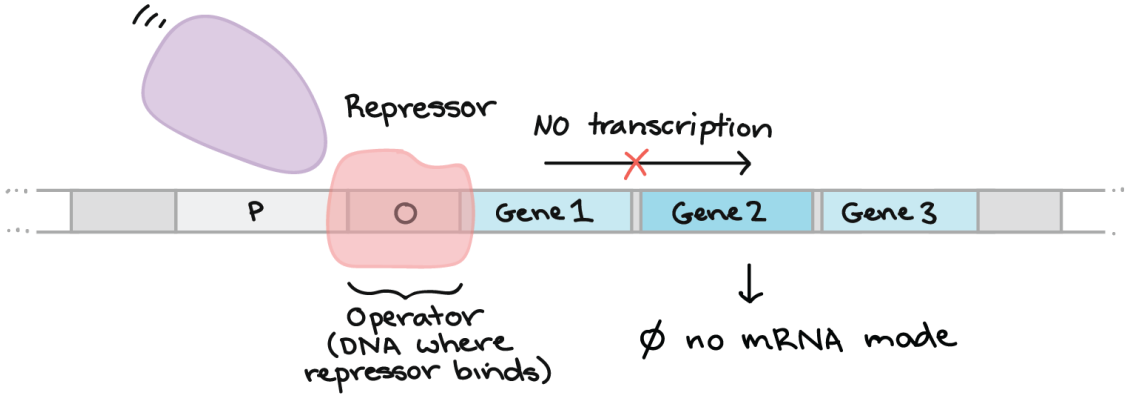


Diagram illustrating how a repressor works. A repressor protein binds to a site called on the operator. In this case (and many other cases), the operator is a region of DNA that overlaps with or lies just downstream of the RNA polymerase binding site (promoter). That is, it is in between the promoter and the genes of the operon. When the repressor binds to the operator, it prevents RNA polymerase from binding to the promoter and/or transcribing the operon. When the repressor is bound to the operator, no transcription occurs and no mRNA is made.

- Some regulatory proteins are **activators**. When an activator is bound to its DNA binding site, it increases transcription of the operon (e.g., by helping RNA polymerase bind to the promoter).

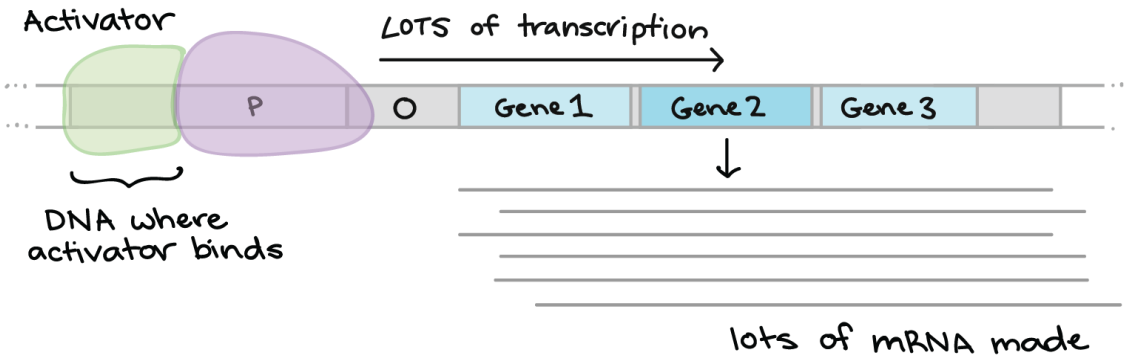


Diagram illustrating how an activator works. The activator protein binds to a specific sequence of DNA, in this case immediately upstream of (before) the promoter where RNA polymerase binds. When the activator binds, it helps the polymerase attach to the promoter (makes promoter binding more energetically favorable). This causes the RNA polymerase to bind firmly to the promoter and transcribe the genes of the operon much more frequently, leading to the production of many molecules of mRNA.

Where do the regulatory proteins come from? Like any other protein produced in an organism, they are encoded by genes in the bacterium's genome. The genes that encode regulatory proteins are sometimes called **regulatory genes**.

Many regulatory proteins can themselves be turned "on" or "off" by specific small molecules. The small molecule binds to the protein, changing its shape and altering its ability to bind DNA. For instance, an activator may only become active (able to bind DNA) when it's attached to a certain small molecule.

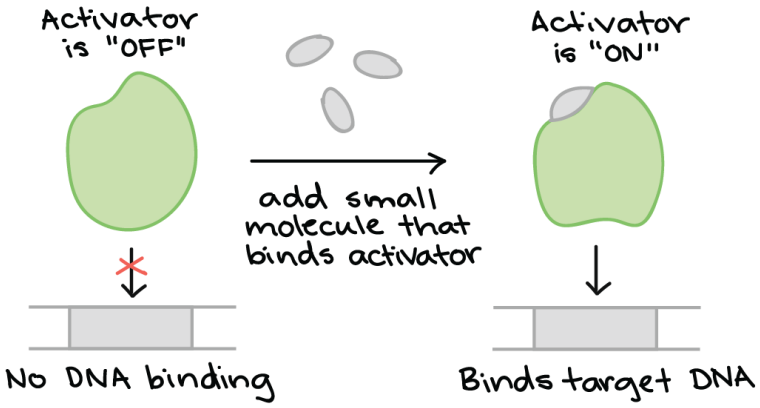


Diagram illustrating how a hypothetical activator's activity could be modulated by a small molecule. When the small molecule is absent, the activator is "off" - it takes on a shape that makes it unable to bind DNA. When the small molecule that activates the activator is added, it binds to the activator and changes its shape. This shape change makes the activator able to bind its target DNA sequence and activate transcription.

Operons may be inducible or repressible

Some operons are usually "off," but can be turned "on" by a small molecule. The molecule is called an **inducer**, and the operon is said to be **inducible**.

- For example, the *lac* operon is an inducible operon that encodes enzymes for metabolism of the sugar lactose. It turns on only when the sugar lactose is present (and other, preferred sugars are absent). The inducer in this case is allolactose, a modified form of lactose.

Other operons are usually "on," but can be turned "off" by a small molecule. The molecule is called a **corepressor**, and the operon is said to be **repressible**.

- For example, the *trp* operon is a repressible operon that encodes enzymes for synthesis of the amino acid tryptophan. This operon is expressed by default, but can be repressed when high levels of the amino acid tryptophan are present. The corepressor in this case is tryptophan.

These examples illustrate an important point: that gene regulation allows bacteria to respond to changes in their environment by altering gene expression (and thus, changing the set of proteins present in the cell).

Some genes and operons are expressed all the time

Many genes play specialized roles and are expressed only under certain conditions, as described above. However, there are also genes whose products are constantly needed by the cell to maintain essential functions. These **housekeeping genes** are constantly expressed under normal growth conditions ("constitutively active"). Housekeeping genes have promoters and other regulatory DNA sequences that ensure constant expression.