**DNA structure and replication**

The structure of the DNA sugar-phosphate backbone gives the DNA strand *directionality*. This simply means that the two ends of the are different: One end is called the 5’ (5-prime) end and one end in called the 3’ (3-prime) end. This name refers to which carbon molecule on the nucleotide backbone the OH- group is attached.

None of this is very important to us, except to understand that the ends of a DNA strand are different.

**DNA replication**

**The Basic Replication Process.**

Semi-conservative replication produces two helices that contain one old and one new DNA strand.

**DNA replication** is semi-conservative. This means that each of the two strands in double-stranded DNA acts as a template to produce two new strands. Each “new” molecule contains one “old” strand that served as the template and one “new” strand that was constructed during the replication process.

Replication relies on complementary **base pairing**: adenine (A) always bonds with thymine (T) and cytosine (C) always bonds with guanine (G).



**More Detail: Enzymes**

The actual process is significantly more complex. We will now look at the process in slightly more detail:

DNA replication occurs through the help of several enzymes. One of those enzymes, called ***DNA helicase***, "unzips" the DNA molecule by breaking the hydrogen bonds that hold the two strands together. Each strand then serves as a template for a new *complementary strand* to be created. Complementary bases attach to one another (A-T and C-G).



The DNA template is used to create the complementary strand. The primary enzyme involved in this is ***DNA polymerase*** which joins nucleotides to synthesize the new complementary strand. DNA polymerase also proofreads each new DNA strand to make sure that there are no errors.

**Even More Detail: Leading and lagging strands**

DNA is made differently on the two strands at a replication fork.

One new strand, the *leading strand*, runs 5' to 3' towards the fork and is made continuously.

The other, the *lagging strand*, runs 5' to 3' away from the fork and is made in small pieces called *Okazaki fragments*.



Diagram of leading and lagging replication strands

**Summary:**

1. The first step in DNA replication is to ‘unzip’ the double helix structure of the DNA molecule. This is carried out by an enzyme called helicase which breaks the hydrogen bonds holding the complementary bases of DNA together

2. The separation of the two single strands of DNA creates a ‘Y’ shape called a replication ‘fork’. The two separated strands will act as templates for making the new strands of DNA.

3. One of the strands is oriented in the 3’ to 5’ direction (towards the replication fork), this is the leading strand. The other strand is oriented in the 5’ to 3’ direction (away from the replication fork), this is the lagging strand. As a result of their different orientations, the two strands are replicated differently:



*An illustration to show replication of the leading and lagging strands of DNA.
Image credit: Genome Research Limited*

**Leading Strand:**

4. A short piece of RNAcalled a primer (produced by an enzyme called primase) comes along and binds to the end of the leading strand. The primer acts as the starting point for DNA synthesis.

5. DNA polymerase binds to the leading strand and then ‘walks’ along it, adding new complementary nucleotide bases (A, C, G and T) to the strand of DNA in the 5’ to 3’ direction. This sort of replication is called continuous.

**Lagging strand:**

4. Numerous RNA primers are made by the primase enzyme and bind at various points along the lagging strand.

5. DNA polymerase adds chunks of DNA, called Okazaki fragments, to the lagging strand also in the 5’ to 3’ direction.

This type of replication is called discontinuous as the Okazaki fragments will need to be joined up later.

6. Once all of the bases are matched up (A with T, C with G), an enzyme called exonuclease strips away the primer(s). The gaps where the primer(s) were are then filled by yet more complementary nucleotides.

DNA Transcription

Transcription is the first step in gene expression. It involves copying a gene's DNA sequence to make an RNA molecule.

RNA (Ribonucleic Acid) is like a single DNA strand, with two key differences: The sugar in the RNA backbone is ribose rather than deoxyribose and more importantly (to us, at least) is that the thymine base is replaced by uracil. That is to say that ***RNA has A, C, G and U*** instead of A, C, G and T. The pairing rules are sill intact:

A with U

C with G

Transcription is performed by enzymes called ***RNA polymerases***, which link nucleotides to form an RNA strand (using a DNA strand as a template).

**The Basic Transcription Process:**

Transcribe means to copy a message into a written form. In biology the goal of transcription is to make an RNA copy of a gene's DNA sequence. The RNA copy, called messenger RNA, or mRNA, or transcript, carries the information needed to build a protein out of the nucleus, into the cytoplasm.



In transcription:

1. A region of DNA in unzipped and un wound. The region of opened-up DNA is called a **transcription bubble**.

2. One strand, the template strand, serves as a template for synthesis of a complementary mRNA transcript.

3. The mRNA molecule moves out of the nucleus.

4. The DNA molecule “rezips”.

**More Detail: Stages of transcription**

Transcription of a gene takes place in three stages: ***initiation, elongation, and termination.*** Here, we will briefly see how these steps happen in bacteria. You can learn more about the details of each stage (and about how eukaryotic transcription is different) in the [stages of transcription](https://www.khanacademy.org/science/biology/gene-expression-central-dogma/transcription-of-dna-into-rna/a/stages-of-transcription) article.

**1. Initiation:**  RNA polymerase binds to a sequence of DNA called the promoter, found near the beginning of a gene. Each gene (or group of co-transcribed genes, in bacteria) has its own promoter. Once bound, RNA polymerase separates the DNA strands, providing the single-stranded template needed for transcription.



**2. Elongation:** One strand of DNA, the template strand, acts as a template for RNA polymerase. As it "reads" this template one base at a time, the polymerase builds an mRNA molecule out of complementary nucleotides, making a chain that grows from 5' to 3'. The RNA transcript carries the same information as the non-template (coding) strand of DNA, but it contains the base uracil (U) instead of thymine (T).



**3. Termination:**  Sequences called **terminators** signal that the RNA transcript is complete. Once they are transcribed, they cause the transcript to be released from the RNA polymerase.

DNA Translation

During translation, the mRNA sequence is read in groups of three nucleotides. Each three-letter "word" or ***codon*** corresponds to an amino acid that's added to a polypeptide (protein or protein subunit).

The amino acids are carries by small segments of RNA, three nucleotides long, called ***transfer RNA*** or tRNA. The tRNA temporarily binds to the strand of mRNA while the amino acid it carries is attached to the growing peptide chain. Once the amino acid id attached, the tRNA detaches and goes back to pick up another amino acid.



This is the end goal of DNA and of genes. Of course, the real process is much more complex. The5re are many more enzymes involved and the sequences of DNA for any gene are on the order of 100 000 letters long. Furthermore, many of the letters are actually just gibberish, and the code must be spliced together to make sense. Consider the following simple example:

BQQWTHEDOGRSMAPQANANDAZAPTQMTETHEHAMFAUMNO

Decode the message above. Just like in a DNA sequence, all the words are three letters long. First, remove the "junk" letters. Second, put the remaining letters into groups of three, starting at the beginning.

Practice:

1. For the following DNA template strand segments, write the mRNA code that results from transcription:

A. CGGATTACAGTT

B. AGGTACCGCCA

C. ACCTATATTCCG

2. For the following DNA coding strand segments, write the mRNA code that results from transcription:

A. GTTCACGGA

B. CATGAGTTA

C. AAGCTCGCTG

3. What was the sequence of bases on the template DNA strand that results in the following mRNA sequence?

A. ACGGUGCU

B. GCCAUUAC

C. UUGGCAGC

4. What sequence of amino acids results from the following mRNA sequences?

A. CUUGAGGCUGCGUGAUG

B. AUGUUUUAUCAAUAAAUA

C. AGGCGAUCCAGCUGA