

Structure of DNA

Introduction

In *Nucleic Acids* we learned that nucleotides are a combination of a base pair, a pentose sugar, and a phosphate. As nucleotides start linking together, the newest nucleotide with a phosphate group on its 5-carbon position bonds to the previous nucleotide at the 3-carbon position, forming a phosphodiester bond. Phosphodiester bonds are powerful and difficult to break. But we formed only a single strand of nucleic acids that way. Here, we're presenting eukaryotic DNA and exploring some features of **double-stranded DNA**.

While uncommon, double-stranded RNA follows the same rules described here, except uracil is substituted for thymine. However, going forward let's assume that DNA is double-stranded, while RNA is single-stranded.

Complementary and Antiparallel

Double-stranded is said to be complementary and antiparallel. It's the same genetic code—the exact identical sequence running past itself in opposite directions—and these two strands are connected to each other. We'll explore what this means exactly. Recall that the genetic code message is read 5' to 3'; a single strand of DNA goes in this direction. When there's double-stranded DNA, there will be a second copy of the message—completely identical—also read 5' to 3'. But the two strands are aligned so that they're going in opposing directions. If written out, one strand is going 5' left to 3' right, while the other is going 3' left to 5' right.

Since the second strand is **complementary** to the original, there'll be an easily deducible binding pattern. Complementary means, "A's with T's, G's with C's."

Adenosine and thymine always pair. There are **two hydrogen bonds** between them.

Guanine and cytosine also always pair. There are **three hydrogen bonds** between them.

Hydrogen bonds are weak and easy to break. That means separating DNA strands from each other requires little energy. This is good, because access to the information within is needed for replication and transcription. Conversely, the phosphodiester bonds are strong. This is also good, as we don't want our DNA falling apart. The pentose-phosphate backbone prevents all enzymes from the cell getting access to the nucleotides. It provides a protective shell, while the hydrogen bonds between nucleotides allows specialized enzymes to break those bonds to gain access to the code.

Reading top-down along the original strand yields the 5' to 3' message. If trying to read top-down along the second strand, it's quickly apparent there isn't identical code in that direction. But, if reading top down on the original strand, and bottom-up on the second (5' to 3'), the message IS the same.

The strands are **complementary** in that a T-A and G-C pairings are the only pairings allowed.

The strands are **antiparallel** because they both contain the same 5' to 3' message but move in opposite directions.

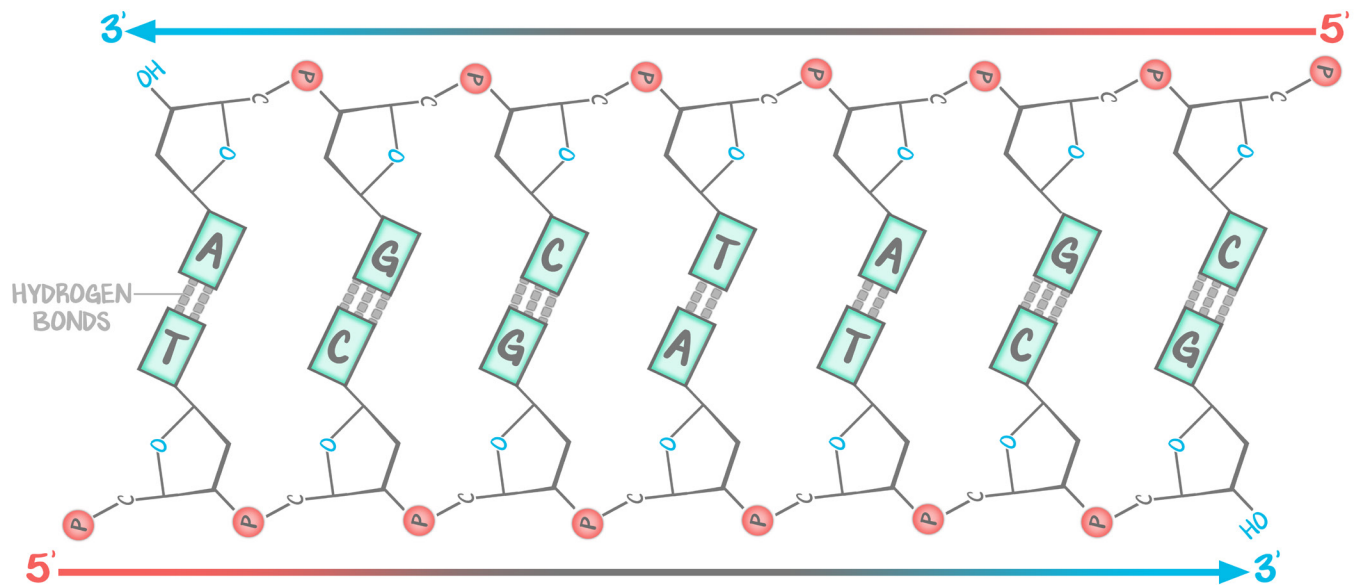


Figure 3.1: Complementary and Antiparallel

The entire genetic code is complementary and antiparallel, the exact same code duplicated twice. To illustrate this, we've chosen a sequence that is a palindrome. Choose any random stretch of DNA: it would not be identical in both directions, but if allowed to carry out to the ends of each strand, they would be.

Chargaff's Rules

The name "Chargaff" is irrelevant, but the concept is commonly tested. Since A always pairs with T, and G always pairs with C, it ensures an **identical quantity of purines and pyrimidines**. It also means that if the quantity of any ONE nucleotide in a DNA strand is known, the remaining three nucleotide quantities can be calculated.

If there are 15% A's, and A always pairs with T, it means there are also 15% T's. That's a total of 30% for the A-T pair, leaving 70% left over for the G-C pair. Since G always pairs with C, that 70% is broken down to 35% G's and 35% C's.

Helix

By the very nature of the hydrogen bonds, when there are two strands of DNA the **base-pairs align in the center** of the molecule. The only thing the cell sees is the **pentose-phosphate backbone**—a negatively charged and hydrophilic support structure exposed on either side of the double-stranded DNA. This protects the DNA and ensures that only the desired code is accessed. That's because specialized proteins are required to break the hydrogen bonds, partially reveal the code, allow access to it, then close it back up. Otherwise, unless it's being accessed for transcription or replication, the genetic code is safely stored away within the helix.

Nearly all DNA is **B-DNA**, which is the **right-handed** helix structure. Rarely, a left-handed Z DNA can be found, though its purpose is unknown. There are approximately **10 base pairs** per turn of the helix. This helical formation has **major grooves** (lots of space) and **minor grooves** (not as much space). The major-groove-minor-groove unit is repeated throughout the entire length of the DNA. The grooves serve as **protein-binding sites**, which can **modify gene translation**.