

# Nucleic Acid Structure

- 1 [Nucleic Acid Structure](#)
- 2 [Nucleobases](#)
- 3 [Nucleobase Aromaticity](#)
- 4 [Base Pairing](#)
- 5 [DNA Double Helix structure](#)
- 6 [Translation](#)
- 7 [Non-Watson-Crick Base Pairing](#)
- 8 [Answers](#)

Nucleic acids are the biopolymers that are responsible for the storage and transmission of genetic information in all living organisms. The monomers, called nucleotides, that make up these macromolecules are composed of a five-membered carbohydrate ring (furanose), a nitrogen-containing base, and a phosphate group. The name of the nucleobase is determined by what groups are attached to it. When a hydrogen is attached, the resulting structure is called a nucleobase. When a ribose sugar is attached, the resulting structure is called a nucleoside. When a deoxyribose sugar is attached, the resulting structure is called a deoxyribonucleoside. When a phosphorylated sugar is attached, the structure is called a nucleotide. When nucleotide units are linked together, they form nucleic acids.

[Download the chemdraw file above](#)

## Nucleic Acid Structure

[Download the chemdraw file above](#)

B = nucleobase (see below)

This is a nucleic acid monomer of DNA which contains a 2'-deoxyribose sugar (note: RNA contains the sugar ribose).

Phosphodiester bonds can be formed by linking the 3' hydroxyl group of one nucleotide with the 5' phosphate group of another to make a nucleic acid polymer molecule.

[Download the chemdraw file above](#)

Notice that this nucleic acid has distinct ends. In this case, the left end is called the 5' terminus, and the right end is called the 3' terminus (defined by the carbons in the sugar).

## Nucleobases

Nucleotides can have one of five nucleobases bound to the 1' carbon. These nucleobases are aromatic nitrogen-containing bases. They can be double-ringed purines (guanine and adenine), or they can be single-ringed pyrimidines (cytosine, thymine, and uracil).

[Download the chemdraw file above](#)

Here are the purines guanine (G) and adenine (A).

[Download the chemdraw file above](#)

Here are the pyrimidines cytosine (C), thymine (T), and uracil (U). It is important to note that thymine is found only in DNA, while uracil is found only in RNA.

It is the sequence of nucleobases along a nucleic acid chain that encodes the genetic information.

## Nucleobase Aromaticity

DNA is the material that cells use to store genetic information. Because of this important function, the body has ensured that DNA molecules are stabilized through several mechanisms. One of these mechanisms is to ensure that the individual components of DNA are stable. One of these stabilized components of DNA is a nucleobase. Both the purine and pyrimidine nucleobases find additional stability through aromaticity. While this aromaticity is easily seen in adenine, it is not as readily apparent when looking the other nucleobase structures. It is helpful to look at these nucleobases' resonance structure to see its aromaticity.

Here are the purines.

[Download the chemdraw file above](#)

Here are the pyrimidines.

[Download the chemdraw file above](#)

[Download the chemdraw file above](#)

In both the purines and pyrimidines, Nitrogen feeds electron density into the ring's pi system sending excess density onto the electronegative Oxygen. This causes the Nitrogen to become electron poor and the Oxygen to become electron rich.

## Base Pairing

Nucleobases can pair with other nucleobases through hydrogen bonding to form a nucleic acid duplex with two complementary strands. Purines pair with pyrimidines. More specifically, guanine (G) pairs with cytosine (C); adenine (A) pairs with thymine (T) (or uracil in the case of RNA). Thus, G and C are said to be complementary, and A is complementary to T. These complementary base pairs are known as Watson-Crick base pairs. The surface at which the hydrogen bonding occurs is called the Watson-Crick surface.

G and C form three hydrogen bonding interactions, while A and T form two. The extra H-bond in the G-C base pairing makes the G-C interaction 50% stronger than the A-T interaction and, thus, harder to separate. The three H-bonds in G-C base pairing have an overall strength of 6 kcal/mol in comparison to the 4 kcal/mol strength of A-T pairing. Since G-C base pairing is stronger than A-T this means that G-C base pairs resist denaturation by heat more so than A-T base pairs. Thermophiles, bacteria that live in extremely high temperatures take advantage of this by incorporating more G-C pairs than A-T in their genomes. The base pairing interactions are shown below (only the donor lone pairs have been shown). Note that the interactions between the heteroatoms and hydrogen atoms during base pairing are described by the terms "donor" and "acceptor." The terms refer to H donation, NOT the donation of the electrons that constitute the bond. If the base is contributing a hydrogen, it is considered a "donor," and if the base is contributing a lone pair of electrons, it is considered an "acceptor." The relationships are noted in red in the figure.

[Download the chemdraw file above](#)

Here is a better look at what is going on between a base pair, in this case G-C. Note that electron pairs must not be involved in resonant delocalization if they are to participate in H-bonding. Electron pairs that are in p orbitals and delocalized are indicated with arrows.

[Download the chemdraw file above](#)

Notice that cytosine is drawn here in its resonant form to provide stabilizing aromaticity, but that the hydrogen bonding is still the same.

One would like to tautomerize guanine in order to give it more stabilizing aromatic character as well (because the carbonyl gives the ring limited aromaticity); however, tautomerization of guanine produces a nucleobase which can no longer pair properly with cytosine (or any base for that matter). Tautomerization of guanine is actually one cause of cancer.

[Download the chemdraw file above](#)

## DNA Double Helix structure

The hydrogen bonding between complementary base pairs is what accounts for the double helical shape of DNA. As stated before, two complementary strands of DNA come together to form a nucleic acid duplex. These complementary strands are, importantly, antiparallel to each other. That is, one strand runs 5' to 3' while the other runs 3' to 5'. This helps DNA maximize stabilizing base pair interactions and brings about the double helical structure as the strands naturally twist around each other.

[Download the chemdraw file above](#)

Notice that these strands are antiparallel. Also, notice how the nucleobases are going into the plane of the page, like rungs on a ladder.

Now imagine looking down on a cross-section of the DNA backbone so that you are looking down on the face of the nucleobase rings, and the phosphate-sugar backbone is going into the plane of the page (like looking down on one rung of a ladder).

[Download the chemdraw file above](#)

You will notice in this image that there is a larger, outside curve and a smaller, inside curve. These curves form what are called major grooves and minor grooves, respectively, in the DNA helix. To understand how this works, imagine taking your ladder and bending it so that the rungs bulge out to one side (almost like a half cylinder). Now imagine twisting this bulging ladder in a counter-clockwise manner and you end up with your right handed DNA helix.

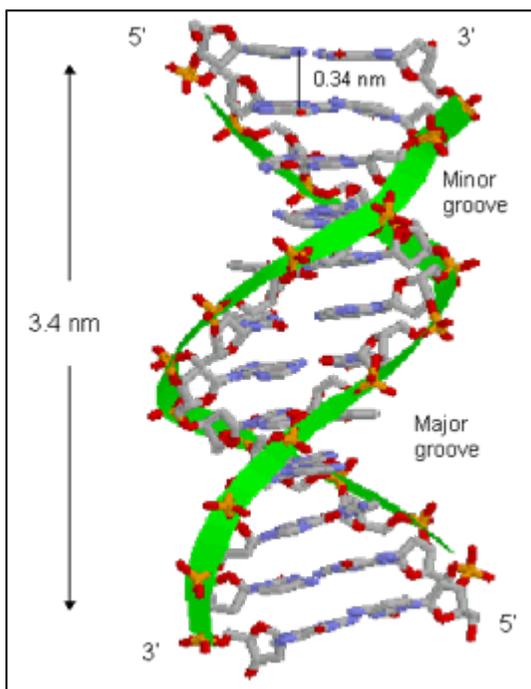


Image from <http://www.web-books.com/MoBio/Free/images/Ch3B3.gif>.

An important characteristic of the grooves is their involvement in allowing DNA to react with other molecules in the cell. You have already seen how most of electron density of the nucleobases is either involved in aromaticity or hydrogen bonding, leaving very few options when it comes to interaction with other molecules. A process of elimination leaves only the N7 and N3 sites on guanine and adenine as the only free nucleophiles that can interact with outside molecules. When designing anticancer drugs, they should target these sites because they are the only sites open for interaction. Notice that the diagram of the major and minor groove above, we see that N7 is found at the major groove and N3 is found at the minor groove. Proteins and other large macromolecules will bind to DNA in the major groove, while only smaller molecules, like drugs or harmful carcinogens, can bind to the minor groove.

## Translation

[Download the chemdraw file above](#)

An interesting fact in amino acid-coding codon sequencing is that there are 61 codons, but only around 40 anticodon containing tRNA molecules. The explanation can be understood by looking at the third codon and seeing that despite that codon, the same amino acid can be coded for between different triplet codons. The anticodon must not be too much specificity in the 5' position.

## Non-Watson-Crick Base Pairing

Sometimes, base pairs can interact in ways other than those described by Watson-Crick base pairing. One example is the Hoogsteen G-quadruplex structures found in the telomeres at the ends of chromosomes. G-quadruplexes are made up of four guanines that hydrogen bond with one another in a square-planar arrangement. In the middle of the G-quadruplex is a positively charged metal ion, which gives extra stability to the lone pairs of the oxygens not involved in hydrogen bonding. The large amount of hydrogen bonding in G-quadruplexes makes them extremely hard to degrade, which is why telomeres are composed of many of these structures. Often, cancerous cells are found to be defective in the formation of G-quadruplexes, compromising the integrity of the DNA.

[Download the chemdraw file above](#)

[Download the chemdraw file above](#)

There is also Hoogsteen base pairing found in DNA between nucleobases such as adenine and thymine and guanine and cytosine in the major groove. Basically, Hoogsteen pairing is normal Watson-Crick base pairing only flipped over, allowing for new geometry and new conformations such as the G-quadruplex described above. In the adenine and thymine Hoogsteen pairing, the nucleobases found on the two DNA strands are bound together by hydrogen bonds, using the N7 and exocyclic amine of the adenine and the carbonyl and the amide of the thymine. This frees the lone pairs on the adenine N1 and N3. In the guanine and cytosine Hoogsteen pairing, the nucleobases are bound together by hydrogen bonds using the N7 and carbonyl of the guanine and the amide and exocyclic amine of the cytosine. The lone pair on the guanine N3 is thus free.

[Download the chemdraw file above](#)

[Download the chemdraw file above](#)

Another example of base pairs interacting in atypical Watson-Crick manner is the Wobble Hypothesis of tRNA. The Wobble hypothesis was proposed by Crick in 1965. It clarifies the property of anticodons and the degenerative nature of the genetic code. The anticodon present in the tRNA can pair with more than one codon in the mRNA. This is true because the third base of the anticodon is non-specific to a particular base. This non-specific third base is called wobble base. This base can pair with a normal base and also with an atypical base, such as guanine with uracil. This solves the issue of the degenerative nature of the genetic sequence, namely, many amino acids are encoded by more than one codon.

For example, the anticodon CUG of leucine tRNA recognizes the codon GAC in the mRNA. Here the third base G of tRNA pairs with C of mRNA. G-C pairing is the normal base pairing. The CUG anticodon can also pair with another codon CUU which also codes for serine. Here the third base G of tRNA pairs with U of mRNA.

These depictions illustrate the difference between the normal G-C and A-U pairing and the 'wobble' pairing of G-U:

[Download the chemdraw file above](#)

## Drill Problems

**D22-1.** Double-stranded DNA can be heat denatured to produce single-stranded DNA as shown below. Based on your knowledge of nucleic acid interactions, circle the double-stranded DNA sequence that would require the highest temperature for denaturation. Give a brief explanation based on the structural differences in these sequences.

[Download the chemdraw file above](#)

**D22-3.** Single base mutations (point mutations) are recognized by DNA repair enzymes because they lead to basepair mismatches. However, tautomerism of nucleobases can sometimes mask this mispairing.

(a) Draw the structures that result from the amide tautomerizations of guanine and thymine.

(b) Based on your structures from part (a), what base pairings can such a tautomerization lead to for guanine? For thymine?

Fortunately, tautomerism accounts for less than 0.001% of the naturally occurring structures, so unrepaired mutations due to this phenomenon are kept infrequent.

**D22-2.** To obtain a better understanding of the relationship of DNA's structure to its function, researchers have designed some novel nucleoside bases, two of which are shown below.

[Download the chemdraw file above](#)

(a) Unlike C-G, iso-C-iso-G is expected to have only two hydrogen bonding interactions when these are base paired. Draw this base pairing interaction between iso-C and iso-G.

(b) Experiments have shown that iso-G prefers to base pair with thymidine. Show the tautomer of iso-G that would favor this interaction. In the same diagram, illustrate how thymidine would interact with this tautomeric iso-G.

**D22-4.** Uracil is a nucleobase used in place of thymine in RNA. The only difference between U and T is that U lacks the exocyclic methyl of T. Which remaining nucleobase (C, G, or A) do you suspect U base pairs best with? Draw this base pair.

## Answers

<a href="#">Drill22-1.mov</a>	
<a href="#">Drill22-2.mov</a>	
<a href="#">Drill22-3.mov</a>	