# Discovery of DNA Structure and Function: Watson and Crick

By: Leslie A. Pray, Ph.D. 
Citation: Pray, L. (2008) Discovery of DNA structure and function: Watson and Crick. Nature Education 1(1)
http://www.nature.com/scitable/topicpage/discovery-of-dna-structure-and-function-watson-397
[Novembre 2011]

The landmark ideas of Watson and Crick relied heavily on the work of other scientists. What did the duo actually discover?

Many people believe that American biologist James Watson and English physicist Francis Crick discovered DNA in the 1950s. In reality, this is not the case. Rather, DNA was first identified in the late 1860s by Swiss chemist Friedrich Miescher. Then, in the decades following Miescher's discovery, other scientists--notably, Phoebus Levene and Erwin Chargaff--carried out a series of research efforts that revealed additional details about the DNA molecule, including its primary chemical components and the ways in which they joined with one another. Without the scientific foundation provided by these pioneers, Watson and Crick may never have reached their groundbreaking conclusion of 1953: that the DNA molecule exists in the form of a three-dimensional double helix.

## The First Piece of the Puzzle: Miescher Discovers DNA

Although few people realize it, 1869 was a landmark year in genetic research, because it was the year in which Swiss physiological chemist Friedrich Miescher first identified what he called "nuclein" inside the nuclei of human white blood cells. (The term "nuclein" was later changed to "nucleic acid" and eventually to "deoxyribonucleic acid," or "DNA.") Miescher's plan was to isolate and characterize not the nuclein (which nobody at that time realized existed) but instead the protein components of leukocytes (white blood cells). Miescher thus made arrangements for a local surgical clinic to send him used, pus-coated patient bandages; once he received the bandages, he planned to wash them, filter out the leukocytes, and extract and identify the various proteins within the white blood cells. But when he came across a substance from the cell nuclei that had chemical properties unlike any protein, including a much higher phosphorous content and resistance to proteolysis (protein digestion), Miescher realized that he had discovered a new substance (Dahm, 2008). Sensing the importance of his findings, Miescher wrote, "It seems probable to me that a whole family of such slightly varying phosphorous-containing substances will appear, as a group of nucleins, equivalent to proteins" (Wolf, 2003).

More than 50 years passed before the significance of Miescher's discovery of nucleic acids was widely appreciated by the scientific community. For instance, in a 1971 essay on the history of nucleic acid research, Erwin Chargaff noted that in a 1961 historical account of nineteenth-century science, Charles Darwin was mentioned 31 times, Thomas Huxley 14 times, but Miescher not even once. This omission is all the more

remarkable given that, as Chargaff also noted, Miescher's discovery of nucleic acids was unique among the discoveries of the four major cellular components (i.e., proteins, lipids, polysaccharides, and nucleic acids) in that it could be "dated precisely... [to] one man, one place, one date."

## Laying the Groundwork: Levene Investigates the Structure of DNA

Meanwhile, even as Miescher's name fell into obscurity by the twentieth century, other scientists continued to investigate the chemical nature of the molecule formerly known as nuclein. One of these other scientists was Russian biochemist Phoebus Levene. A physician turned chemist, Levene was a prolific researcher, publishing more than 700 papers on the chemistry of biological molecules over the course of his career. Levene is credited with many firsts. For instance, he was the first to discover the order of the three major components of a single nucleotide (phosphate-sugar-base); the first to discover the carbohydrate component of RNA (ribose); the first to discover the carbohydrate component of DNA (deoxyribose); and the first to correctly identify the way RNA and DNA molecules are put together.

During the early years of Levene's career, neither Levene nor any other scientist of the time knew how the individual nucleotide components of DNA were arranged in space; discovery of the sugar-phosphate backbone of the DNA molecule was still years away. The large number of molecular groups made available for binding by each nucleotide component meant that there were numerous alternate ways that the components could combine. Several scientists put forth suggestions for how this might occur, but it was Levene's "polynucleotide" model that proved to be the correct one. Based upon years of work using hydrolysis to break down and analyze yeast nucleic acids, Levene proposed that nucleic acids were composed of a series of nucleotides, and that each nucleotide was in turn composed of just one of four nitrogen-containing bases, a sugar molecule, and a phosphate group. Levene made his initial proposal in 1919, discrediting other suggestions that had been put forth about the structure of nucleic acids. In Levene's own words, "New facts and new evidence may cause its alteration, but there is no doubt as to the polynucleotide structure of the yeast nucleic acid" (1919).

Indeed, many new facts and much new evidence soon emerged and caused alterations to Levene's proposal. One key discovery during this period involved the way in which nucleotides are ordered. Levene proposed what he called a tetranucleotide structure, in which the nucleotides were always linked in the same order (i.e., G-C-T-A-G-C-T-A and so on). However, scientists eventually realized that Levene's proposed tetranucleotide structure was overly simplistic and that the order of nucleotides along a stretch of DNA (or RNA) is, in fact, highly variable. Despite this realization, Levene's proposed polynucleotide structure was accurate in many regards. For example, we now know that DNA is in fact composed of a series of nucleotides and that each nucleotide has three components: a phosphate group; either a ribose (in the case of RNA) or a deoxyribose (in the case of DNA) sugar; and a single nitrogen-containing base. We also know that there are two basic categories of nitrogenous bases: the purines (adenine [A] and guanine [G]), each with two fused rings, and the pyrimidines (cytosine [C], thymine [T], and uracil [U]), each with a single ring. Furthermore, it is now widely accepted that

RNA contains only A, G, C, and U (no T), whereas DNA contains only A, G, C, and T (no U) (Figure 1).

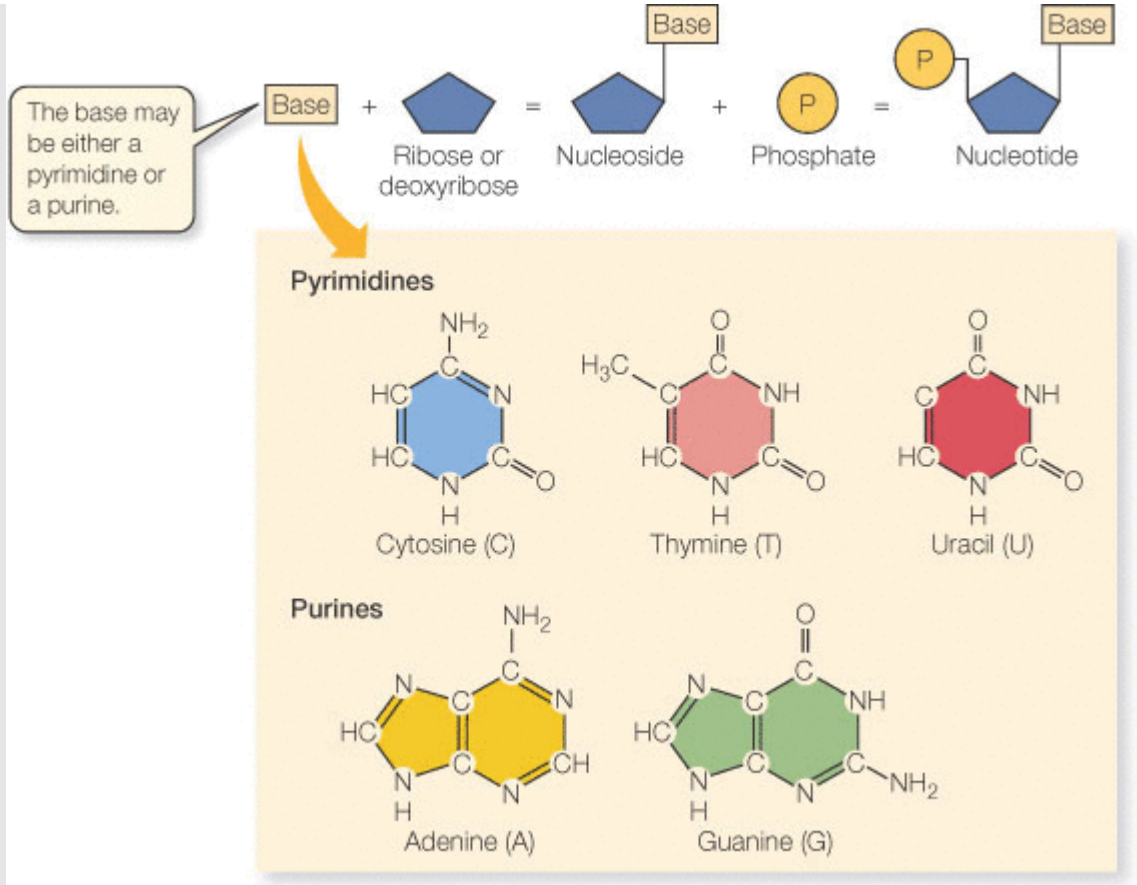


Figure 1: Nucleotides have three components.
A nucleotide consists of a phosphate group, a pentose sugar (ribose or deoxyribose), and a nitrogen-containing base, all linked together by covalent bonds. The nitrogenous bases have two different chemical forms: purines have two fused rings, and the smaller pyrimidines have a single ring.


# Strengthening the Foundation: Chargaff Formulates His "Rules"

Erwin Chargaff was one of a handful of scientists who expanded on Levene's work by uncovering additional details of the structure of DNA, thus further paving the way for Watson and Crick. Chargaff, an Austrian biochemist, had read the famous 1944 paper by Oswald Avery and his colleagues at Rockefeller University, which demonstrated that hereditary units, or genes, are composed of DNA. This paper had a profound impact on Chargaff, inspiring him to launch a research program that revolved around the chemistry of nucleic acids. Of Avery's work, Chargaff (1971) wrote the following:

*"This discovery, almost abruptly, appeared to foreshadow a chemistry of heredity and, moreover, made probable the nucleic acid character of the gene... Avery gave us the first text of a new language, or rather he showed us where to look for it. I resolved to search for this text."*

As his first step in this search, Chargaff set out to see whether there were any differences in DNA among different species. After developing a new paper chromatography method for separating and identifying small amounts of organic material, Chargaff reached two major conclusions (Chargaff, 1950). First, he noted that the nucleotide composition of DNA varies among species. In other words, the same nucleotides do not repeat in the same order, as proposed by Levene. Second, Chargaff concluded that almost all DNA--no matter what organism or tissue type it comes from--maintains certain properties, even as its composition varies. In particular, the amount of adenine (A) is usually similar to the amount of thymine (T), and the amount of guanine (G) usually approximates the amount of cytosine (C). In other words, the total amount of purines (A + G) and the total amount of pyrimidines (C + T) are usually nearly equal. (This second major conclusion is now known as "Chargaff's rule.") Chargaff's research was vital to the later work of Watson and Crick, but Chargaff himself could not imagine the explanation of these relationships--specifically, that A bound to T and C bound to G within the molecular structure of DNA (Figure 2).

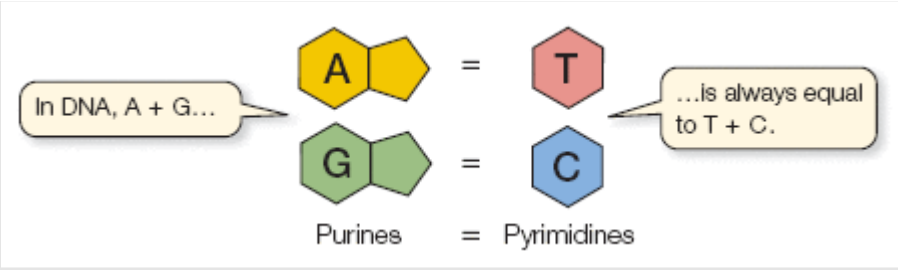


Figure 2: Chargaff's rule.
In DNA, the total abundance of purines is equal to the total abundance of pyrimidines.


# Putting the Evidence Together: Watson and Crick Propose the Double Helix

Chargaff's realization that A = T and C = G, combined with some crucially important X-ray crystallography work by English researchers Rosalind Franklin and Maurice Wilkins, contributed to Watson and Crick's derivation of the three-dimensional, double-helical model for the structure of DNA. Watson and Crick's discovery was also made possible by recent advances in model building, or the assembly of possible three-dimensional structures based upon known molecular distances and bond angles, a technique advanced by American biochemist Linus Pauling. In fact, Watson and Crick were worried that they would be "scooped" by Pauling, who proposed a different model for the three-dimensional structure of DNA just months before they did. In the end, however, Pauling's prediction was incorrect.

Using cardboard cutouts representing the individual chemical components of the four bases and other nucleotide subunits, Watson and Crick shifted molecules around on their desktops, as though putting together a puzzle. They were misled for a while by an erroneous understanding of how the different elements in thymine and guanine (specifically, the carbon, nitrogen, hydrogen, and oxygen rings) were configured. Only upon the suggestion of American scientist Jerry Donohue did Watson decide to make new cardboard cutouts of the two bases, to see if perhaps a different atomic configuration would make a difference. It did. Not only did the complementary bases

now fit together perfectly (i.e., A with T and C with G), with each pair held together by hydrogen bonds, but the structure also reflected Chargaff's rule (Figure 3).

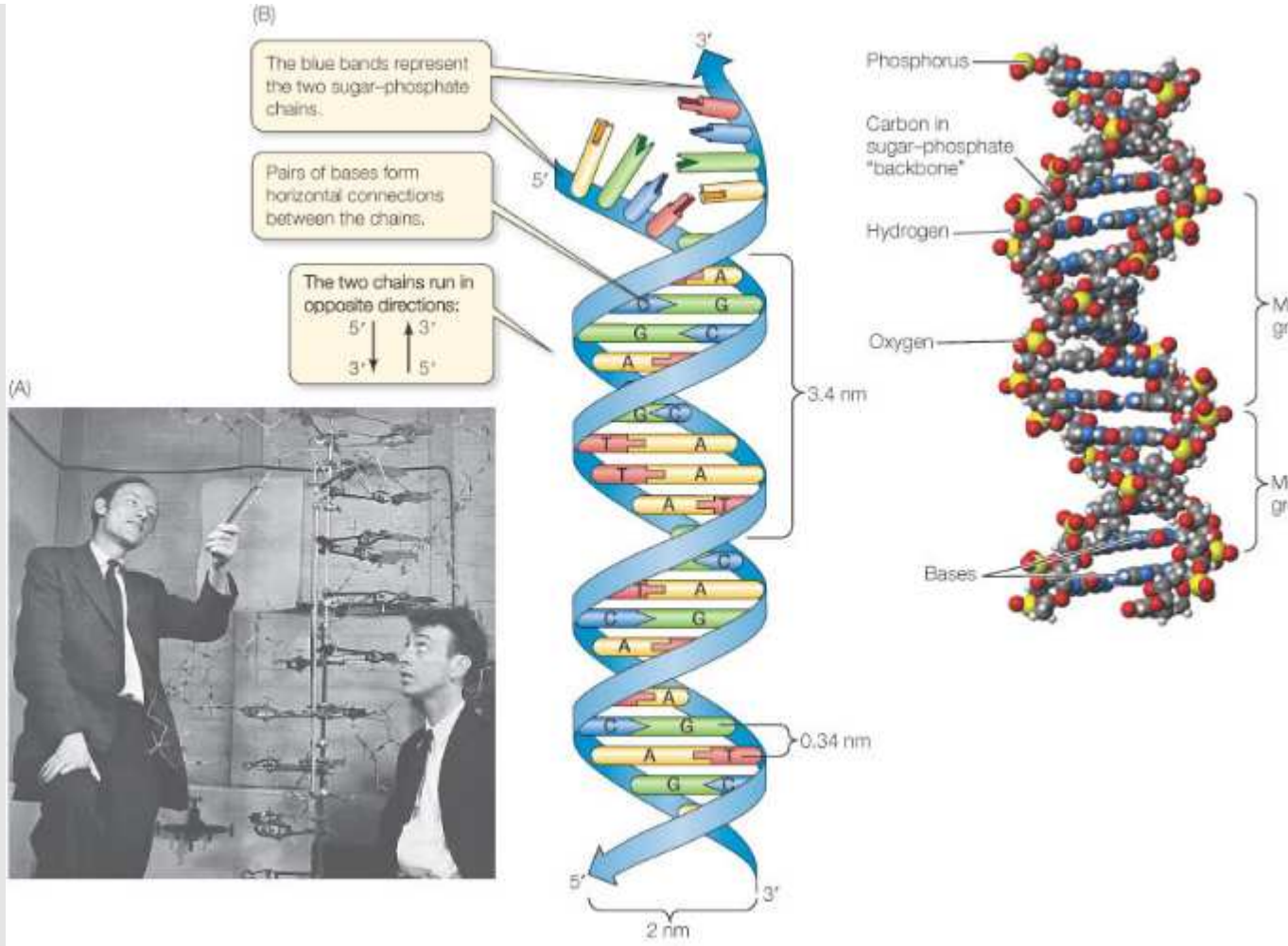


Figure 3: DNA is a double helix.
(A) Francis Crick (left) and James Watson (right) proposed that the DNA molecule has a double-helical structure. (B) Biochemists can now pinpoint the position of every atom in a DNA molecule. To see that the essential features of the original Watson-Crick model have been verified, follow with your eyes the double-helical chains of sugar-phosphate groups and note the horizontal rungs of the bases.


Although scientists have made some minor changes to the Watson and Crick model, or have elaborated upon it, since its inception in 1953, the model's four major features remain the same yet today. These features are as follows:

- DNA is a double-stranded helix, with the two strands connected by hydrogen bonds. A bases are always paired with Ts, and Cs are always paired with Gs, which is consistent with and accounts for Chargaff's rule.
- Most DNA double helices are right-handed; that is, if you were to hold your right hand out, with your thumb pointed up and your fingers curled around your thumb, your thumb would represent the axis of the helix and your fingers would represent the sugar-phosphate backbone. Only one type of DNA, called Z-DNA, is left-handed.

- The DNA double helix is anti-parallel, which means that the 5' end of one strand is paired with the 3' end of its complementary strand (and vice versa). As shown in Figure 4, nucleotides are linked to each other by their phosphate groups, which bind the 3' end of one sugar to the 5' end of the next sugar.
- Not only are the DNA base pairs connected via hydrogen bonding, but the outer edges of the nitrogen-containing bases are exposed and available for potential hydrogen bonding as well. These hydrogen bonds provide easy access to the DNA for other molecules, including the proteins that play vital roles in the replication and expression of DNA (Figure 4).

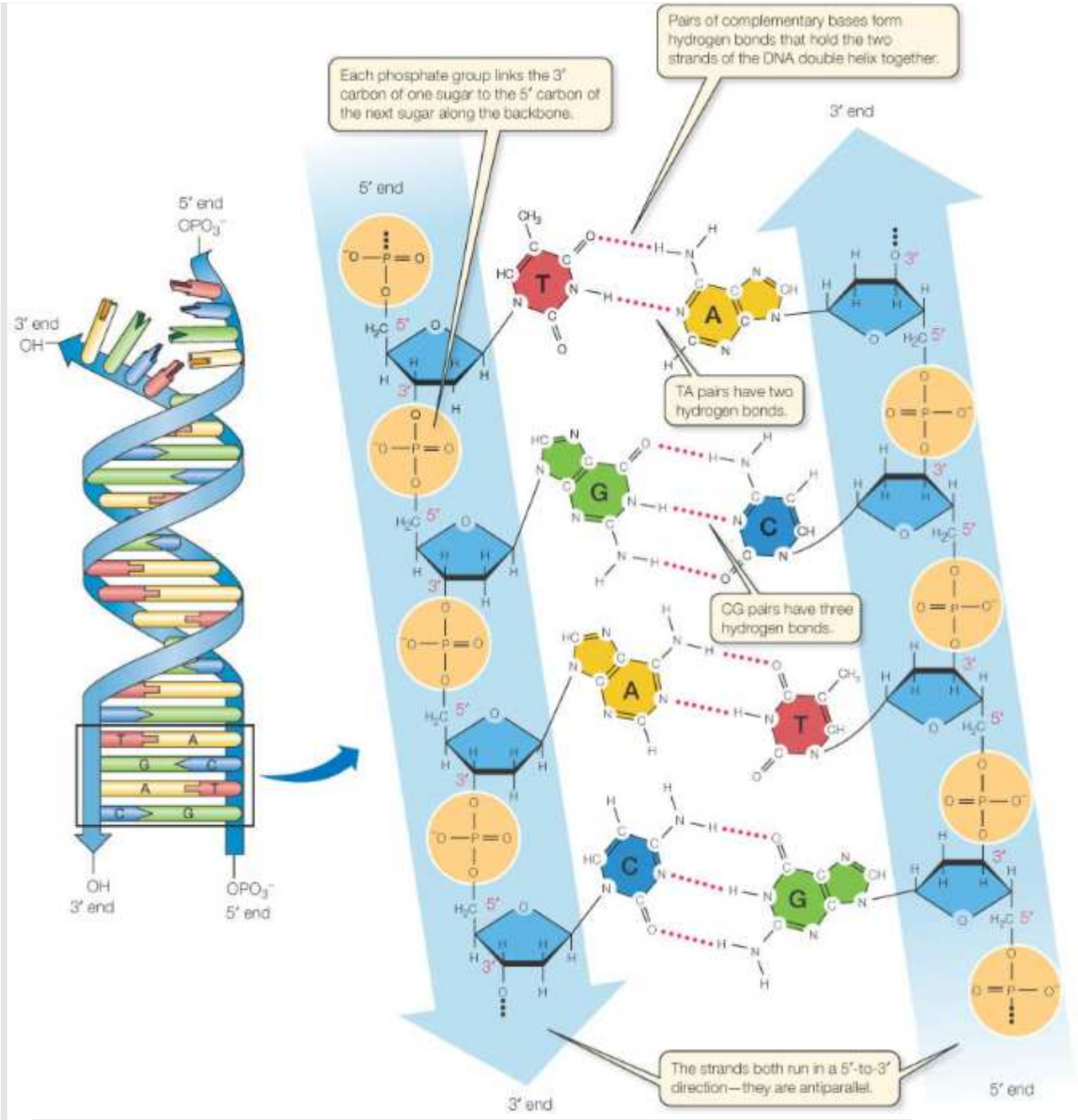


Figure 4: Base pairing in DNA is complementary.

The purines (A and G) pair with the pyrimidines (T and C, respectively) to form equal-sized base pairs resembling rungs on a ladder (the sugar-phosphate backbones). The ladder twists into a double-helical structure.


One of the ways that scientists have elaborated on Watson and Crick's model is through the identification of three different conformations of the DNA double helix. In other words, the precise geometries and dimensions of the double helix can vary. The most common conformation in most living cells (which is the one depicted in most diagrams of the double helix, and the one proposed by Watson and Crick) is known as B-DNA. There are also two other confirmations: A-DNA, a shorter and wider form that has been found in dehydrated samples of DNA and rarely under normal physiological circumstances; and Z-DNA, a left-handed confirmation. Z-DNA is a transient form of DNA, only occasionally existing in response to certain types of biological activity (Figure 5). Z-DNA was first discovered in 1979, but its existence was largely ignored until recently. Scientists have since discovered that certain proteins bind very strongly to Z-DNA, suggesting that Z-DNA plays an important biological role in protection against viral disease (Rich & Zhang, 2003).

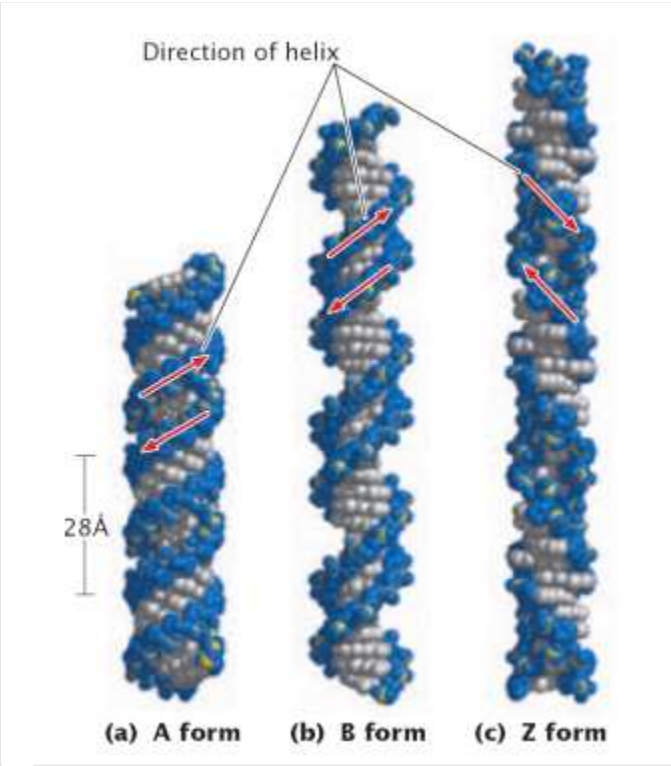


Figure 5: DNA can assume several different secondary structures.
These structures depend on the base sequence of the DNA and the conditions under which it is placed.


# Summary

Watson and Crick were not the discoverers of DNA, but rather the first scientists to formulate an accurate description of this molecule's complex, double-helical structure.

Moreover, Watson and Crick's work was directly dependent on the research of numerous scientists before them, including Friedrich Miescher, Phoebus Levene, and Erwin Chargaff. Thanks to researchers such as these, we now know a great deal about genetic structure, and we continue to make great strides in understanding the human genome and the importance of DNA to life and health.

# References and Recommended Reading

---


Chargaff, E. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia* **6**, 201–209 (1950)

---. Preface to a grammar of biology. *Science* **171**, 637–642 (1971)

Dahm, R. Discovering DNA: Friedrich Miescher and the early years of nucleic acid research. *Human Genetics* **122**, 565–581 (2008)

Levene, P. A. The structure of yeast nucleic acid. IV. Ammonia hydrolysis. *Journal of Biological Chemistry* **40**, 415–424 (1919)

Rich, A., &. Zhang, S. Z-DNA: The long road to biological function. *Nature Reviews Genetics* **4**, 566–572 (2003) (link to article)

Watson, J. D., & Crick, F. H. C. A structure for deoxyribose nucleic acid. *Nature* **171**, 737–738 (1953) (link to article)

Wolf, G. Friedrich Miescher: The man who discovered DNA. *Chemical Heritage* **21**, 10-11, 37–41 (2003)