27.2: DNA BASE PAIRS

Objectives

After completing this section, you should be able, given the necessary Kekulé structures, to show how hydrogen bonding can occur between thymine and adenine, and between guanine and cytosine; and to explain the significance of such interactions to the primary and secondary structures of DNA.

Study Notes

Watson and Crick received the Nobel Prize in 1962 for elucidating the structure of DNA and proposing the mechanism for gene reproduction. Their work rested heavily on X-ray crystallographic work done on RNA and DNA by Franklin and Wilkins. Wilkins shared the Nobel Prize with Watson and Crick, but Franklin had been dead four years at the time of the award (you cannot be awarded the Nobel Prize posthumously).

The history of Watson and Crick's proposed DNA model is controversial and a travesty of scientific ethics. Rosalind Franklin was deeply involved in the determination of the structure of DNA, and had collected numerous diffraction patterns. Watson attended a departmental colloquium at King's College given by Franklin, and came into possession of an internal progress report she had written. Both departmental colloquia and progress reports are merely methods of discussion between colleagues; works presented in these fora are not considered by scientists to be "published" works, and therefore are not in the public domain. Watson and Crick not only were aware of Franklin's work, but used her unpublished data, presented in confidence within her own college.

The final blow came about a year after the colloquium. Watson visited Wilkins at King's College, and Wilkins inexplicably handed over Franklin's diffraction photographs without her consent. Had Franklin's work not been secretly taken from her, she might quite possibly have solved the DNA structure before Watson and Crick, who at the time did not yet have their own photographs. This is truly one of the sadder episodes of questionable scientific ethics and discovery that I have ever encountered.

References

Kass-Simon, G., and P. Farnes. Women of Science: Righting the Record. Bloomington, IN: Indiana University Press, 1990.
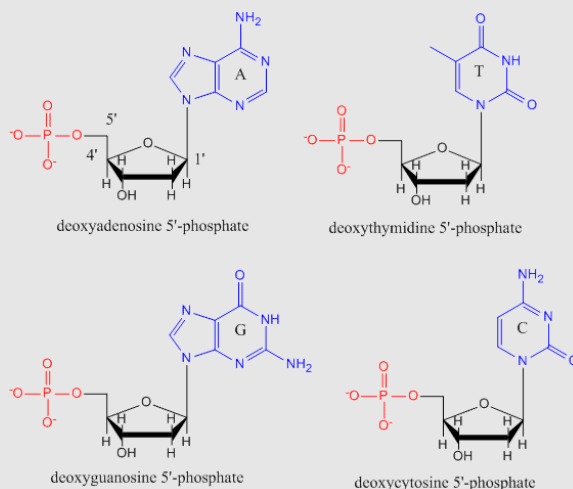
Maddox, B. Rosalind Franklin: The Dark Lady of DNA. New York: HarperCollins, 2002.

## INTERMOLECULAR FORCES IN NUCLEIC ACIDS

The nucleic acids RNA and DNA are involved in the storage and expression of genetic information in a cell. Both are polymers of monomeric nucleotides. DNA exists in the cell as double-stranded helices while RNA typically is a single-stranded molecule which can fold in 3D space to form complex secondary (double-stranded helices) and tertiary structures in a fashion similar to proteins. The complex 3D structures formed by RNA allow it to perform functions other than simple genetic information storage, such as catalysis. Hence most scientists believe that RNA preceded both DNA and proteins in evolution as it can both store genetic information and catalyze chemical reactions.

DNA

DNA is a polymer, consisting of monomers call deoxynucleotides. The monomer contains a simple sugar (deoxyribose, shown in black below), a phosphate group (in red), and a cyclic organic R group (in blue) that is analogous to the side chain of an amino acid.

deoxyadenosine 5'-phosphate

deoxythymidine 5'-phosphate

deoxyguanosine 5'-phosphate

deoxycytosine 5'-phosphate

Only four bases are used in DNA (in contrast to the 20 different side chains in proteins) which we will abbreviate, for simplicity, as A, G, C and T. They are bases since they contain amine groups that can accept protons. The polymer consists of a sugar - phosphate - sugar - phosphate backbone, with one base attached to each sugar molecule. As with proteins, the DNA backbone is polar but also charged. It is a polyanion. The bases, analogous to the side chains of amino acids, are predominately polar. Given the charged nature of the backbone, you might expect that DNA does not fold to a compact globular (spherical) shape, even if positively charged cations like Mg bind to and stabilize the charge on the polymer. Instead, DNA exists usually as a double-stranded (ds) structure with the sugar-phosphate backbones of the two different strands running in opposite directions (5'-3' and the other 3'-5'). The strands are held together by hydrogen bonds between bases on complementary strands. Hence like proteins, DNA has secondary structure but in this case, the hydrogen bonds are not within the backbone but between the "side chain" bases on opposing strands. It is actually a misnomer to call dsDNA a molecule, since it really consists of two different, complementary strands held together by hydrogen bonds. A structure of ds-DNA showing the opposite polarity of the strands is shown below.
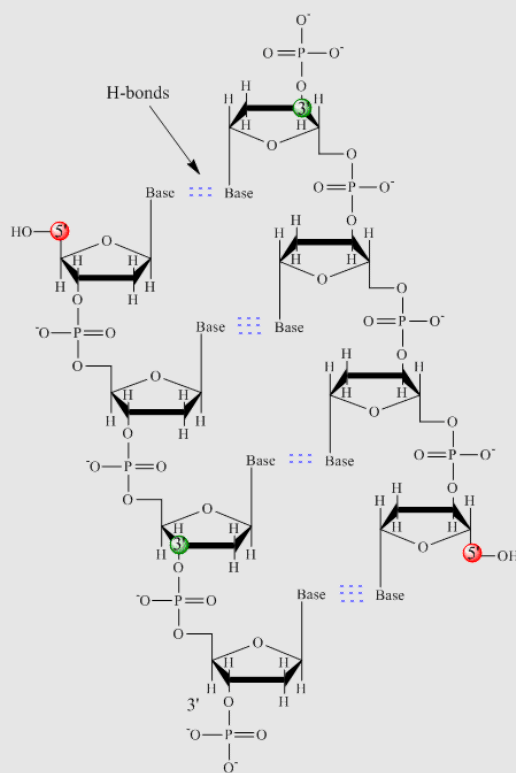
In double stranded DNA, the guanine (G) base on one strand can form three H-bonds with a cytosine (C) base on another strand (this is called a GC base pair). The thymine (T) base on one strand can form two H-bonds with an adenine (A) base on the other strand (this is called an AT base pair). Double-stranded DNA has a regular geometric structure with a fixed distance between the two backbones. This requires the bases pairs to consists of one base with a two-ring (bicyclic) structure (these bases are called purines) and one with a single ring structure (these bases are called pyrimidines). Hence a G and A or a T and C are not possible base pair partners.

DNA Base Pairing

A and T base pair
through Hydrogen bonds

G and C base pair
through Hydrogen bonds
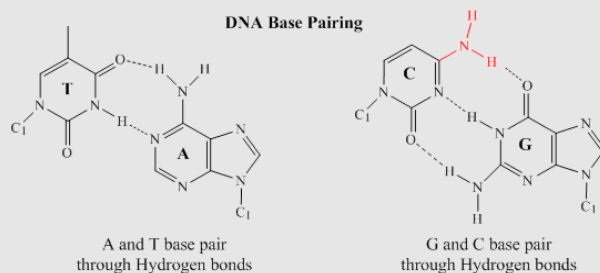
Double stranded DNA varies in length (number of sugar-phosphate units connected), base composition (how many of each set of bases) and sequence (the order of the bases in the backbone). The following links provide interactive Jmol models of dsDNA made by Angel Herráez, Univ. de Alcalá (Spain) and Eric Martz.

- Jmol model of ds-DNA with base pairs and H-bonds
- Jmol model of DNA strands and helical backbone

- Jmol model of DNA ends and parallelisms

Chromosomes consist of one dsDNA with many different bound proteins. The human genome has about 3 billion base pairs of DNA. Therefore, on average, each single chromosome of a pair has about 150 million base pairs and lots of proteins bound to it. dsDNA is a highly charged molecule, and can be viewed, to a first approximation, as a long rod-like molecule with a large negative charge. It is a polyanion. This very large molecule must somehow be packed into a small nucleus of a tiny cell. In complex (eukaryotic) cells, this packing problem is solved by coiling DNA around a core complex of four different pairs (eight proteins total) of histone proteins (H2A, H2B, H3, and H4) which have net positive charges. The histone core complex with dsDNA wound around approximately 2.5 times is called the nucleosome.

Jmol model of the nucleosome

DNA can adopt two other types of double-helical forms. The one discovered by Watson and Crick and found in most textbooks is called B-DNA. Depending on the actual DNA sequence and the hydration state of the DNA, it can be coaxed to form two other types of double-stranded helices, Z and A DNA. The A form is much more open then the B form.

The 3.2 billion base pairs of DNA in humans contains about 24,000 short stretches (genes) that encode different proteins. These genes are interspersed among DNA that helps determining if the gene is decoded into RNA molecules (see below) and ultimately into proteins. For a particular gene to be activated (or "turned on"), specific proteins must bind to the region of a particular gene. How can binding proteins find specific binding targets among the vast number of base pairs that to a first approximation have a repetitive sugar-phosphate-base repeat? The Jmol below shows how specificity can be achieved. When DNA winds into a double helix through base-pairs between AT and GC, hydrogen bond donors (amide Hs) and acceptors (Os) on the bases that are not used in intrastrand base pairing,are still available in the major and minor grove of the ds-DNA helix (see Jmol below). Unique base pair sequences will display unique patterns of H bond donors and acceptors in the major grove. These donors/acceptors can be recognized by specific DNA binding proteins which on binding can lead to gene activation.

- Jmol model of dsDNA showing unique H bond donors and acceptors in the major grove

CONTRIBUTORS AND ATTRIBUTIONS

- Dr. Dietmar Kennepohl FCIC (Professor of Chemistry, Athabasca University)
- Prof. Steven Farmer (Sonoma State University)
- Chris P Schaller, Ph.D., (College of Saint Benedict / Saint John's University)
- (Henry Jakubowski)

27.2: DNA Base Pairs is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.