**Overview**

Deoxyribonucleic acid (DNA) stores information for the synthesis of specific proteins. DNA has deoxyribose as its sugar. DNA consists of a phosphate group, a sugar, and a nitrogenous base. The structure of DNA is a helical, double-stranded macromolecule with bases projecting into the interior of the molecule. These two strands are always complementary in sequence. One strand serves as a template for the formation of the other during DNA replication, a major source of inheritance. This unique feature of DNA provides a mechanism for the continuity of life. The structure of DNA was found by Rosalind Franklin when she used x-ray crystallography to study the genetic material. The x-ray photo she obtained revealed the physical structure of DNA as a helix.

DNA has a double helix structure. The outer edges are formed by alternating deoxyribose sugar molecules and phosphate groups, which make up the sugar-phosphate backbone. The two strands run in opposite directions, one going in a 3' to 5' direction and the other going in a 5' to 3' direction. The nitrogenous bases are positioned inside the helix structure like "rungs on a ladder," due to the hydrophobic effect, and stabilized by hydrogen bonding.

<table>
<thead>
<tr>
<th>Nitrogenous base</th>
<th>Nucleoside</th>
<th>Deoxynucleoside</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenine</td>
<td>Adenosine</td>
<td>Deoxyadenosine</td>
</tr>
<tr>
<td>Guanine</td>
<td>Guanosine</td>
<td>Deoxyguanosine</td>
</tr>
<tr>
<td>Thymine</td>
<td>5-Methyluridine</td>
<td>Deoxythymidine</td>
</tr>
<tr>
<td>Uracil</td>
<td>Uridine</td>
<td>Deoxyuridine</td>
</tr>
</tbody>
</table>

The structure of DNA has been elucidated by x-ray crystallography. The x-ray photo obtained by Rosalind Franklin revealed the physical structure of DNA as a helix.
The two strands run in opposite directions to form the double helix. The strands are held together by hydrogen bonds and hydrophobic interactions. The H-bonds are formed between the base pairs of the anti-parallel strands. The base in the first strand forms a H-bond only with a specific base in the second strand. Those two bases form a base-pair (H-bond interaction that keeps strands together and form double helical structure). The base-pairs in DNA are adenine-thymine (A-T) and cytosine-guanine (C-G). Such interactions provide us an understanding that nitrogen-containing bases are located inside of the DNA double helical structure, while sugars and phosphates are located outside of the double helical structure.

The component consisting of the base and the sugar is known as the nucleoside. DNA contains deoxyadenosine (deoxyribose sugar bonded to adenine), guanoside (deoxyribose sugar bonded to guanine), cytidine (deoxyribose sugar bonded to cytosine), and thymidine (deoxyribose sugar bonded to thymine). The linkage of the bonds between the base to the sugar is known as the beta-Glycosidic linkage. In purines, this occurs between the N-9 and C-1’ and in pyrimidines this occurs between the N-1 and C-1’. A nucleoside and a phosphate group make up a nucleotide. The bond between the deoxyribose sugar of the nucleoside and the phosphate group is a 3’-5’ phosphodiester linkage.

The bases, located inside the double helix, are stacked. Stacking bases interact with each other through the Van der Waals forces. Although the energy associated with a Van der Waals interaction is relatively small, in a helical structure, a large number of atoms are intertwined in such interactions and the net sum of the energy is quite substantial. The distance between two neighboring bases that are perpendicular to the main axis is 3.4 Å. The DNA structure is repetitive. There are ten bases per turn, that is the structure repeats after 34 Å, so every base has a 34° angle of rotation. The diameter of the double helix is approximately 20 Å.

An easy way to differentiate between Nucleosides and Deoxynucleosides is the atoms bonded to C-2 on the sugar unit. If the structure is a nucleoside, then C-2 bears two hydrogens. If it is a deoxynucleoside, then C-2 bears one hydrogen and one hydroxide group, in which the hydroxide group faces south.

Terms and Naming

There are two types of nucleic acids, ribonucleic acids (RNA) and deoxyribonucleic acid (DNA). Recall that a nucleoside is a base + sugar. A Nucleotide is composed of a base + sugar + phosphate. The deoxy- prefix in Deoxyribonucleotides is the nomenclature used for DNA. The term ribonucleotides is employed when it is nomenclature for RNA, or in other words, C-2 on the sugar unit has an -OH group (versus deoxy which C-2 has 2 hydrogens). Symbols are used to simplify the names. For example, ATP (precursor of RNA). The "A" in the front signifies that the base is Adenine and the "T" in the middle signifies tri-phosphates. AMP on the other hand, also has an adenine, but the M signifies that the sugar is bound to a single phosphate group. Finally, in dAMP, the "d" signifies that it is a 2'-deoxyribo-, versus simply AMP means it is a ribonucleotide. In short, four nucleotide units of DNA are called deoxyadenylate, deoxyguanylate, deoxycytidylate, and thymidylate.
Early foundation for DNA structures

The primary structure of a nucleic acid is its covalent structure and nucleotide sequences. One of most important parts of determining the structure of DNA comes from the work of Erwin Chargaff and his colleagues in the late 1940s. They found that the four nucleotide bases of DNA of different organisms and that the amounts of certain bases are closely related. They concluded the following about the structure of DNA:

1. The base composition of DNA generally varies from one species to another.

2. DNA specimens isolated from different tissues of the same species have the same base composition.

3. The base composition of DNA in a given species does not change over time, nutritional states, or environment.

4. In all cellular DNA, regardless of the species, the number of adenine residues is equal to the number of thymine residue (A=T) and the number of guanine residues is equal to the number of cytosine residues (G=C).

Later in 1953, Rosalind Franklin and Maurice Wilkins used a powerful X-ray diffraction technique called X-ray crystallography to deduce the DNA structure. Photographs produced by the X-ray crystallography method are not actually pictures of molecules, however the spots and smudges produced by X-rays that were diffracted (deflected) as they passed through crystallized DNA. Crystallographers use mathematical equations to translate such patterns of spots into information about the three-dimensional shape of DNA. Franklin and Wilkins found that DNA molecules are helical with two periodicities along their long axis, a primary one of 3.4 A and a secondary one of 34 A.
Watson and Crick later based their model of DNA upon the data they were able to extract from Wilkins and Franklin's X-ray diffraction photo.

http://37days.typepad.com/37days/images/2008/03/02/franklin20dna20photo.jpg

They interpreted the pattern of spots on the X-ray photo to mean that DNA consisted of two chains and was helical in shape. Eventually, Watson and Crick formulated a DNA structure from the diffraction pattern of the x-ray photo and gave to incredible insight that is still accepted today. In this structure, they proposed that two helical DNA chains of opposite direction wound around the same axis to form a right handed double helix. The hydrophobic backbones form by phosphodiester bonds of alternating deoxyribose sugar and phosphate group that are faced outside of the helix, surrounded by aqueous environment. The furanose ring of each deoxyribose sugar is in the C-2' endo conformation. The purine and pyrimidine bases of both strands are stacked inside the double helix and stabilized by Van Der Waals interactions.

The double-helix has a diameter of 10 Å. Each adjacent base on one strand of the double-helix is 3.4 Å apart. Every 10 base-pairs constitutes a 360° turn in the helix, and the length of the helix is determined by 34 Å per 10 base-pairs.

**Orientation**

DNA molecules are asymmetrical, such property is essential in the processes of DNA replication and transcription. A double-stranded DNA molecule consists of two complementary but disjoint strands that are intertwined into a helix formation through a network of H bonds. Although both the right-handed and left-handed helices are among the allowed conformations, right-handed helices are energetically more favorable due to less steric hindrance between the side chains and the backbone. The direction of DNA is determined by the arrangement of the phosphate and deoxyribose sugar groups along the DNA backbone. One of the DNA ends terminates with the 3'-H group, whereas the other one terminates with the 5'-H group. All sequences of DNA are usually written from 5' to 3' termini. In a double-helix formation, the complementary DNA strands are oriented in opposite directions. DNA is a rather rigid molecule: at physiological conditions, DNA curves at the length scale of about 50 nm, which is 20 times the diameter of the double helix. More so, the alignment of the bases can indicate the global orientation of a DNA strand. For purine nucleotides (A and G) the most probable angle is approximately 88°, whereas for pyrimidine (C and T) that angle is approximately 105°.
Forces involved in DNA helices

The DNA double helix is held together by two main forces: hydrogen bonds between complementary base pairs inside the helix and the Van der Waals base-stacking interaction.
Hydrogen bonds

Watson and Crick found that the hydrogen bonded base pairs, G with C, A with T, are those that best fit within the DNA structure. It is important to note that three hydrogen bonds can form between G and C, but only two bonds can be found in A and T pairs. This is why it is more difficult to separate DNA strands that contain more G-C pairs than A-T pairs. On the other hand, A-T pairs seem to destabilize the double helical structures. This conclusion was made possible by a known fact that in each species the G content is equal to that of C content and the T content is equal to that of A content.

Below is the link to the demo of the Hydrogen bondings between base pairs:

http://chemmac1.usc.edu/java/bases/basepairs.html

The three hydrogen bonds that constitute the linkage of Guanine(G) and Cytosine(C) consequently alters the thermal melting of DNA, which is dependent upon base compositions. With varying base composition the melting point of such molecule will either increase or decrease.

Denaturing and Annealing

Ultraviolet (UV) light can detect whether bases are stacked or unstacked. Stacked bases within the DNA structure facilitate shielding from light, therefore the absorbance of UV light of double helical DNA is much less than single stranded DNA. This characteristic is known as the hypochromic effect, in which less color is emitted from the double helix of DNA molecules.

The melting temperature ($T_m$) is the temperature in which DNA is half way between double stranded and of random sequence. The $T_m$ depends greatly on base composition. Since G-C base pairs are stronger due to more Hydrogen bonds, DNA with high G-C content will have a higher $T_m$ than that of DNA with greater A-T content.

When heat is applied to a double-stranded DNA, each individual strand will eventually separate (denature) because hydrogen bonds are disrupted between base pairs. Upon separation, the separated strands spontaneously reassociate to form the double helix again. This process is known as annealing.

In biological systems, both denaturing and annealing can occur. Helicases use chemical energy (from ATP) to disrupt the structure of double-stranded nucleic acid molecules. The study of the ability of DNA to reanneal within
the laboratory is important in discovering gene structure and expression.

**Complex Structures**

Complex structures can also be formed from single-stranded DNA. A stem-loop is formed when complementary sequences, within the same strand, pair to form a double helix. Hydrogen bonds between base pairs within the same strand occur. Often, these structures include mismatched bases, resulting in destabilization of the local structure. Such action can be important in higher-order folding, like in tertiary structures.

**Hypochromic Effect**

DNA absorbs very strongly at wavelengths close to UV light (~260 nm). A single stranded DNA will absorb more UV light than that of double-stranded DNA. DNA UV absorption decreases when it forms a double strand, this characteristic is an indication of DNA stability. With the increase in light energy, its structure and therefore its function will still remain intact since there is low disturbance to its structure.

The decreased absorbance observed with the DNA double helix with respect to the native and denatured forms is explained by the fact that the stacking of the nitrogenous bases that takes place with the double helix does not leave them as exposed to radiation and thus they are able to absorb less. The aromaticity of the nitrogenous bases (specifically in the purine and pyrimidine like ring structures) accounts for the absorption peak being at 260nm.

**Weak forces**

Various Weak Forces come together to stabilize the DNA structure.

- **Hydrogen bonds**, linkage between bases, although weak energy-wise, is able to stabilize the helix because of the large number present in DNA molecule.

- **Stacking interactions**, or also known as Van der Waals interactions between bases are weak, but the large amounts of these interactions help to stabilize the overall structure of the helix.
  - Double helix is stabilized by **hydrophobic effects** by burying the bases in the interior of the helix increases its stability; having the hydrophobic bases clustered in the interior of the helix keeps it away from the surrounding water, whereas the more polar surfaces, hence hydrophilic heads are exposed and interaction with the exterior water
  - Stacked base pairs also attract to one another through **Van der Waals forces** the energy associated with a single van der Waals interaction has small significant to the overall DNA structure however, the net effect summed over the numerous atom pairs, results in substantial stability.
  - Stacking also favors the conformations of rigid five-membered rings of the sugars of backbone.

- **Charge-Charge Interactions** - refers to the electrostatic (ion-ion) repulsion of the negatively charged phosphate is potentially unstable, however the presence of Mg$^{2+}$ and cationic proteins with abundant Arginine and Lysine residues that stabilizes the double helix.

**Nitrogenous Bases**

Nitrogenous Bases are the foundational structure of DNA polymers, the structure of DNA polymers vary with the different attached nitrogenous bases.

Nitrogenous Bases can tautomerize between keto and enol forms. The aromaticity of the pyrimidine (Cytosine, Thymine, Uracil (RNA)) and purine (Adenine, Guanine) ring systems and their electron-rich nature of -OH and -NH$_2$ substituents able them to undergo keto-enol tautomeric shifts. The keto tautomer is called a lactam and the enol tautomer is called lactam. The lactam predominates at pH 7. Keto-enol tautomerization is the interconversion of a keto and enol involving the movement of a proton and the shifting of bonding electrons, hence the isomerism qualifies as tautomerism.
Keto-enol tautomerism is important in DNA structure because high phosphate-transfer potential of phosphoenolpyruvate results in the phosphorylated compound to be trapped in the less stable enol form, whereas dephosphorylation results in the keto form. Rare enol tautomers of bases guanine and thymine can lead to mutation because of the altered base-pairing properties.

**Base-stacking interactions**

The two strands of double-stranded DNA are held together by a number of weak interactions such as hydrogen bonds, stacking interactions, and hydrophobic effects. Of these, the stacking interactions between base pairs are the most significant. The strength of base stacking interactions depends on the bases. It is strongest for stacks of G-C base pairs and weakest for stacks of A-T base pairs. The hydrophobic effect stacks the bases on top of one another. The stacked base pairs attract one another through Van der Waals forces, typically from 2 to 4 kJ/mol. In addition, base stacking in DNA is favored by the conformations of the somewhat rigid five membered rings of the backbone phosphate-sugars. The base-stacking interactions, which are largely nonspecific with respect to the identity of the stacked base, make the major contribution to the stability of the double helix.

**Phosphodiester Bond**

Phosphodiester linkages form the covalent backbone of DNA. A phosphodiester bond is the linkage formed between the 3' carbon atom and the 5' carbon of the sugar deoxyribose in DNA.

The phosphate groups in a phosphodiester bond are negatively-charged. The pKa of phosphate groups are near 0, therefore they are negatively-charged at neutral pH (pH=7). This charge-charge repulsion forces the phosphates groups to take opposite positions of the DNA strands and is neutralized by proteins (histones), metal ions such as magnesium, and polyamines.

The tri-phosphate or di-phosphate forms of the nucleotide building are blocks, first have to be broken apart to release the energy require to drive an enzyme-catalyzed reaction for a phosphodiester bond to form and for the nucleotide to join. Once a single phosphate or two phosphates (pyrophosphates) break apart and participate in a catalytic reaction, the phosphodiester bond is formed.

An important role in repairing DNA sequences is due to the hydrolysis of phosphodiester bonds being catalyzed by phosphodiesterases, an enzyme that facilitates the repairs.
Secondary Structures of DNA

Base pairing of complementary nucleotides make up the secondary structure of DNA. A single-stranded DNA may participate in intramolecular base pairing between complementary base pairs and therefore make up secondary structure as well. Base pairing between Adenine (A)-Thymine (T) and Guanine (G)-Cytosine(C) are possible because these base pairs are similar in size. This means there is no "bulges" or "gaps the exist within the double helix.

Irregular placement of base pairs in a double helix will result in consequences that will render the macromolecule nonfunctional. Therefore if there is something wrong with the structure, signals will be alerted and DNA repair will work to fix damages.

As a result of the double helical nature of DNA, the molecule has two asymmetric grooves. One groove is smaller than the other. This asymmetry is a result of the geometrical configuration of the bonds between the phosphate, sugar, and base groups that forces the base groups to attach at 120 degree angles instead of 180 degree. The larger groove is called the major groove, occurs when the backbones are far apart; while the smaller one is called the minor groove, occurs when they are close together.

Since the major and minor grooves expose the edges of the bases, the grooves can be used to tell the base sequence of a specific DNA molecule. The possibility for such recognition is critical, since proteins must be able to recognize specific DNA sequences on which to bind in order for the proper functions of the body and cell to be carried out. As you might expect, the major groove is more information rich than the minor groove, allowing the DNA proteins to interact with the bases. This fact makes the minor groove less ideal for protein binding.

![Visual Representation of Major and Minor Grooves in DNA Structure](image)

A form

These following features represented different characteristics of A-form DNA structure:

1. Most RNA and RNA-DNA duplex in this form
2. Shorter, wider helix than B.
3. Deep, narrow major groove not easily accessible to proteins
4. Wide, shallow minor groove accessible to proteins, but lower information content than major groove.
5. Favored conformation at low water concentrations
6. Base pairs tilted to helix axis and displaced from axis
7. Sugar pucker C3'-endo (in RNA 2'-OH inhibits C2'-endo conformation)
8. Right handed
9. Size is about 26 angstroms
10. Needs 11 base pairs per helical turn
11. Glycosyl bond conformation is Anti

B form

The double helical structure of normal DNA takes a right-handed form called the B-helix. It is about 20 angstroms with a C-2’ endo sugar pucker conformation. The helix makes one complete turn approximately every 10 base pairs (= 34 Å per repeat/3.4 Å per base). B-DNA has two principal grooves, a wide major groove and a narrow minor groove. Many proteins interact in the space of the major groove, where they make sequence-specific contacts with the bases. In addition, a few proteins are known to make contacts via the minor groove.

Z form

DNA sequences can flip from a B form to a Z form and vice versa. Z form of DNA is a more radical departure from the B structure; the most obvious distinction is the left-handed helical rotation.

The Z form is about 18 angstroms and there are 12 base pairs per helical turn, and the structure appears more slender and elongated. The DNA backbone takes on a zigzag appearance. Certain nucleotide sequences fold into left-handed Z helices much more readily than others. Prominent examples are sequences in which pyrimidines alternate with purines, especially alternating C and G or 5-methyl-C and G residues. To form the left-handed helix in Z-DNA, the purine residues flip to the syn conformation alternating with pyrimidines in the anti conformation. The major groove is barely apparent in Z-DNA, and the minor groove is narrow and deep. For pyrimidines, the sugar pucker conformation is C-2’ endo and for purines, it is a C-3’ endo.

Z-DNA formation occurs during transcription of genes, at transcription start sites near promoters of actively transcribed genes. During transcription, the movement of RNA polymerase induces negative supercoiling upstream and positive supercoiling downstream the site of transcription. The negative supercoiling upstream favors Z-DNA formation; a Z-DNA function would be to absorb negative supercoiling. At the end of transcription, topoisomerase relaxes DNA back to B conformation.
**Tertiary structure (3 dimensional)**

The tertiary structure of DNA molecule is made up of the two strands of DNA wind around each other. DNA double helix can be arranged in space, in a tertiary arrangement of strands.

- **Linking Number** (Lk) in a covalently closed circular DNA, where the two strands cannot be separated will result in a constant number of turns in a given molecule. Lk of DNA is an integral composed of two components:
  1) **Twist (Tw)**: number of helical turns of DNA strand
  2) **Writhe (Wr)**: number of supercoiled turns in DNA

Normally, DNA has Lk of about 25, meaning it is underwound. However, DNA can also be supercoiled with two “underwindings” which is made up of negative supercoils. This is much like the two “turns- worth” of a single stranded DNA and no supercoils. This kinds of interconversion of helical and superhelical turns in important in gene transcription and regulation.

**Quaternary structure and other unusual structure**

DNA is connected with histones and non-histone proteins to form the chromatin. The negative charge due to the phosphate group in DNA makes it relatively acidic. This negative charge binds to the basic histone groups.

**Histone Modification**

Recent studies provide that actively transcribed regions are characterized by specific modification pattern of histone. The experiments carried on by the dynamics of histone modification shows that there is a significant kinetic distinction between methylation, phosphorylation, and acetylation. This suggest that the roles of these modifications has different roles in gene expression patterns.

Histones are proteins which DNA wraps around and forms a chromatin. The basic unit of a chromatin is a nucleosome which are formed by histone octomer of 2 molecules of H2A, H2B, H3, and H4 along with 147 base pairs of DNA wrapped in a superhelix. The accessibility of DNA is regulated by higher-order chromatin structures that of which can be obtained by the packing of nucleosomes. It is believed that the N-Termini tail of the histone molecules contributes to the chromatin function in that it mediates inter-nucleosomal interactions and are involved in the recruitment of non-histone proteins to the chromatin. The N-termini tail directs interactions to the chromatin binders which is thought to be the driving force of modulate chromatin structure. However, there are other ways modifications can occur such as that observed by the unfolding or assembly of nucleosome and how it is involved in gene regulation. It is hoped that this can provided an explanation of epigenetic inheritance (Box 1) the there phenotypic differences in individual cannot be due to differences in DNA, such as that of monozygotic twins.

Epigenetic inheritance are changes in the gene activity that are not encoded by the DNA sequence. These changes include phosphorylation, methylation, ADP-ribosylation, SUMOylation, and ubiquitylation. These modifications can be considered active or repressive depending on their occurrence in active or silent genes. It is show that methylation can have different outcomes depending on the binders of the histone modifications. Nucleosome positioning are found to have an influence on the DNA sequence and may contribute to epigenetic inheritance.¹
Structural Variation in DNA

The Structural Variation in DNA is most due to:

1) **Varying deoxyribose conformations** (4 total conformations)

2) **Rotations about the contiguous bonds in the phosphodeoxyribose backbone** (between the $C_1$-$C_3$ and $C_5$-$C_6$)

3) **Free rotation about C1'-N-glycosyl bond** (resulting in syn or anti conformation)

Because of steric hindrance, purines bases in nucleotides are restricted to two stable conformations with respect to deoxyribose, called syn and anti. On the other hand, pyrimidines are generally restricted to the anti conformation because of steric interference between the sugar and the carbonyl oxygen at C-2 of the pyrimidine.

Comparison of A, B, and Z form of DNA

<table>
<thead>
<tr>
<th></th>
<th>A form</th>
<th>B form</th>
<th>Z form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helical sense</td>
<td>Right handed</td>
<td>Right handed</td>
<td>Left handed</td>
</tr>
<tr>
<td>Diameter</td>
<td>26 A</td>
<td>20 A</td>
<td>18 A</td>
</tr>
<tr>
<td>Base pairs per helical turn</td>
<td>11</td>
<td>10.5</td>
<td>12</td>
</tr>
<tr>
<td>Helix rise per base pair</td>
<td>2.6 A</td>
<td>3.4 A</td>
<td>3.7 A</td>
</tr>
<tr>
<td>Base tilt normal to the helix axis</td>
<td>$20^\circ$</td>
<td>$6^\circ$</td>
<td>$7^\circ$</td>
</tr>
<tr>
<td>Sugar pucker conformation</td>
<td>C-3’ endo</td>
<td>C-2’ endo</td>
<td>C-2’ endo for pyrimidines and C-3’ endo for purines</td>
</tr>
<tr>
<td>Glycosyl bond conformation</td>
<td>Anti</td>
<td>Anti</td>
<td>Anti for pyrimidine and syn for purines</td>
</tr>
</tbody>
</table>

References


Campbell and Reese's Biology, 7th Edition

Nelson and Cox's Lehninger Principles of Biochemistry, 5th Edition
Article Sources and Contributors


Image Sources, Licenses and Contributors


License

Creative Commons Attribution-Share Alike 3.0 Unported
http://creativecommons.org/licenses/by-sa/3.0/