

# 11 NUCLEIC ACID STRUCTURE: A FOUNDATIONAL SUMMARY

## Executive Summary

Nucleic acids, specifically deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), are the fundamental molecules of life. They serve as the primary repositories and functional expressions of biological information, forming the blueprint that is transmitted from one generation to the next. DNA acts as the stable, long-term archive for genetic information, while RNA functions as the versatile messenger that translates this information into the functional machinery of the cell, such as proteins. Understanding the structure of these molecules is essential to comprehending how they perform these critical roles.

The following key principles provide a foundational overview of nucleic acid structure and stability:

- **DNA's Structure:** DNA is a right-handed double helix composed of two complementary strands that run in opposite directions (antiparallel). The molecule's hydrophilic sugar-phosphate backbones face outward, while the hydrophobic nitrogenous bases are stacked internally. This arrangement provides a stable framework for storing genetic information.
  - **Base Pairing Rule:** The two DNA strands are held together by a precise rule of complementarity. The purine Adenine (A) always pairs with the pyrimidine Thymine (T) via two hydrogen bonds, and the purine Guanine (G) always pairs with the pyrimidine Cytosine (C) via three hydrogen bonds.
  - **Genetic Information:** The specific sequence of these bases along a DNA strand constitutes the genetic code. The complementary nature of the double helix ensures that the information is stored reliably and provides a clear mechanism for its accurate replication, as each strand can serve as a template for synthesizing its partner.
  - **Structural Diversity:** DNA is not a static molecule. While the Watson-Crick B-form is the most common under physiological conditions, DNA can adopt other conformations, such as the wider A-form or the left-handed Z-form, depending on hydration levels and nucleotide sequence. Furthermore, specific sequences can fold into unusual structures like hairpins, three-stranded triplexes, or four-stranded tetraplexes, which may have roles in gene regulation.
  - **RNA's Role and Structure:** RNA primarily functions as a messenger (mRNA), carrying genetic instructions from DNA to the cell's protein-synthesis machinery. Although typically single-stranded, RNA can fold upon itself to create complex secondary structures, with the hairpin loop being the most common. This structural versatility allows RNA to perform a wide range of functions, including catalytic activity.
  - **Chemical Stability:** The stability of the DNA double helix is a critical chemical property. The helix can be unwound, or "denatured," by heat or extreme pH. The temperature at which this occurs (the melting temperature,  $T_m$ ) is largely determined by the molecule's Guanine and Cytosine (G+C) content. This is because  $G \equiv C$  base pairs exhibit stronger base-stacking interactions than  $A = T$  pairs, making G+C-rich DNA more stable.
-

---

# 1. The Blueprint of Life: Principles and Hierarchical Organization

Nucleic acids are the informational molecules that serve as the blueprint for life. They are the repositories of biological information, governing heredity and the functional expression of genetic traits. The transmission of this information from one generation to the next relies on the principle of molecular complementarity. This principle is unique to nucleic acids and allows them to form a consistent, stable structure over immense lengths, which is essential for reliably storing vast amounts of information.

To understand these complex molecules, their structure is best described using three hierarchical levels:

- **Primary Structure:** This is the most fundamental level, referring to the covalent structure of the molecule and, most importantly, the linear sequence of its nucleotides. It is this sequence that encodes the genetic information.
- **Secondary Structure:** This level describes any regular, stable structure adopted by some or all of the nucleotides. The iconic DNA double helix and the hairpins commonly found in RNA are the primary examples of secondary structure.
- **Tertiary Structure:** This refers to the complex, large-scale, three-dimensional folding of nucleic acids. Examples include the compact arrangement of DNA within chromosomes or the elaborate, functional shapes adopted by transfer RNA (tRNA) and ribosomal RNA (rRNA) molecules.

This organizational framework provides a basis for exploring the most famous of these structures: the DNA double helix.

## 2. Unraveling the Double Helix: The Discovery and Structure of DNA

The elucidation of DNA's structure in 1953 by James Watson and Francis Crick was a landmark achievement in biology, built upon several key discoveries. Experiments in the 1940s by Avery, MacLeod, and McCarty established DNA as the genetic material. In 1952, the Hershey-Chase experiment removed any remaining doubt that DNA, not protein, carried genetic information. Shortly thereafter, Erwin Chargaff's work revealed that in any given DNA molecule, the amount of Adenine equals Thymine ( $A=T$ ) and the amount of Guanine equals Cytosine ( $G=C$ ). The final piece of the puzzle came from Rosalind Franklin's X-ray diffraction patterns, which showed that DNA was a helix with distinct, repeating periodicities. Synthesizing this evidence, Watson and Crick proposed a model that accounted for all known properties of DNA.

### 2.1. Key Features of the Watson-Crick Model

The Watson-Crick model describes what is now known as B-form DNA, the predominant structure found under physiological conditions. Its defining characteristics are:

---

- **Overall Shape:** A right-handed double helix, with two chains wound around a central axis.
- **Strand Orientation:** The two strands are **antiparallel**, meaning their 3'–5'-phosphodiester bonds run in opposite directions.
- **Backbone and Bases:** The hydrophilic sugar-phosphate backbones are located on the outside of the helix, exposed to water. The hydrophobic purine and pyrimidine bases are stacked on the inside, perpendicular to the long axis of the helix. The furanose ring of each deoxyribose is in the C-2' endo conformation.
- **Complementary Base Pairing:** The strands are held together by specific hydrogen bonds, forming the basis of information transfer.
  - Adenine (A) pairs with Thymine (T) using **two** hydrogen bonds.
  - Guanine (G) pairs with Cytosine (C) using **three** hydrogen bonds.
- **Structural Dimensions:**
  - The original model proposed 10 base pairs per turn, but subsequent measurements in aqueous solution refined this to **10.5 base pairs** per complete helical turn.
  - The distance per helical turn (pitch) is **36 Å** (3.6 nm).
  - The vertical distance between stacked base pairs is **3.4 Å**.
- **Major and Minor Grooves:** The offset pairing of the two strands creates two distinct grooves on the surface of the helix—a wider major groove and a narrower minor groove. These are critical for protein-DNA interactions, as the major groove is wide enough to allow proteins to access the base pairs and "read" the sequence via the unique chemical signatures presented by the edges of the bases.

## 2.2. The Forces of Stability

Several forces contribute to the stability of the DNA double helix. Ranked by their importance, these are:

- **Primary Stabilizing Force:** The primary stabilizing force is **base-stacking interactions**, a combination of hydrophobic and van der Waals forces between adjacent, stacked base pairs. It is a common misconception that hydrogen bonds are the main source of stability. Stacking between G≡C pairs is thermodynamically stronger than stacking between A=T pairs.
- **Charge Shielding:** The negatively charged phosphate groups in the backbone create electrostatic repulsion. **Metal cations** (like  $Mg^{2+}$ ) in solution shield these charges, stabilizing the duplex.
- **Hydrogen Bonds:** While the hydrogen bonds between base pairs are essential for the *specificity* of pairing (ensuring A pairs only with T, and G with C), they do **not** contribute significantly to the overall *stability* of the structure.

This standard B-form model is not the only way DNA can exist; the molecule's flexibility allows it to adopt several alternative conformations.

## 3. DNA's Dynamic Nature: Alternative Conformations

DNA is a remarkably flexible molecule. Rotations around the bonds in its sugar-phosphate backbone and different conformations of the deoxyribose sugar ring allow the double helix

to adopt structures other than the classic B-form. These variations are often dependent on the local nucleotide sequence and the surrounding chemical environment, particularly the degree of hydration.

The three best-characterized forms of DNA—A, B, and Z—have distinct properties.

Feature	A-Form	B-Form (Watson-Crick)	Z-Form
<b>Helical Sense</b>	Right-handed	Right-handed	Left-handed
<b>Diameter</b>	~26 Å	~20 Å	~18 Å
<b>Base pairs per helical turn</b>	11	10.5	12
<b>Base pair tilt</b>	20°	6°	7°

- **B-Form:** This is the standard Watson-Crick structure and is the most stable form for a random DNA sequence under normal physiological (hydrated) conditions.
- **A-Form:** This is a right-handed helix that is wider and more compact than the B-form. It is favored in solutions that are relatively devoid of water (dehydrated conditions).
- **Z-Form:** This is a left-handed helix that is more slender and elongated. Its backbone follows a zigzag path, giving it its name. This form is favored by specific sequences, particularly those with alternating purines and pyrimidines (e.g., repeating CGCG), and may play a role in regulating the expression of some genes or in genetic recombination.

This ability to form different structures based on sequence is a theme that extends to even more complex, non-helical DNA architectures.

## 4. Beyond the Helix: Complex DNA Architectures

Under certain conditions, specific DNA sequences can cause the molecule to abandon the standard double helix and fold into unusual three-dimensional structures. These architectures are not mere curiosities; they are believed to play important roles in cellular processes like gene regulation and recombination.

- **Palindromes, Mirror Repeats, and Hairpins:** A **palindrome** is an inverted repeat with self-complementarity across opposite strands. This allows a single DNA strand to fold back on itself and base-pair to form a **hairpin** structure (a stem and a loop). If this occurs on both strands of a duplex, it forms a cross-shaped **cruciform** structure. In contrast, a **mirror repeat** has a symmetric sequence within a single strand and cannot form hairpins or cruciforms because it lacks self-complementarity.
- **Triplex DNA:** This structure involves three DNA strands. A third strand can bind within the major groove of a standard double helix through **Hoogsteen pairing**, a non-Watson-Crick form of hydrogen bonding. This is most common in regions with long stretches of only purines or only pyrimidines. The common C≡G•C<sup>+</sup> triplex is most stable at low pH because it requires a protonated cytosine to form the Hoogsteen pair.
- **Tetraplex DNA:** This four-stranded structure occurs readily in DNA sequences with a high proportion of **Guanine (G) residues**. Four guanine bases can hydrogen-bond to form a planar structure called a G-quartet. These quartets stack on top of each other to create a very stable "G tetraplex."

While DNA is the primary storage molecule, its close relative, RNA, is responsible for actively expressing this stored information.

## 5. RNA: The Versatile Messenger and Structural Player

While DNA serves as the permanent genetic archive within the cell, RNA is the active copy that carries out the instructions encoded in that archive. Its central role is to act as a messenger, ferrying information from the DNA in the nucleus to the cytoplasm, where proteins are synthesized.

### 5.1. Messenger RNA (mRNA) and the Flow of Information

- **Function:** Messenger RNA (mRNA) is the class of RNA that carries the genetic code from DNA to the ribosome, the site of protein synthesis.
- **Synthesis:** mRNA is synthesized using a DNA template in a process called **transcription**.
- **Types:** In eukaryotes, most mRNAs are **monocistronic**, meaning each mRNA molecule codes for only one polypeptide. In bacteria and archaea, mRNAs are often **polycistronic**, containing the code for two or more different polypeptides on a single molecule.

### 5.2. Unique Structural Features of RNA

RNA's structure is distinct from DNA's, which allows it to perform a much wider range of functions.

- **Strandedness:** RNA is typically a single-stranded molecule.
- **Default Conformation:** A single RNA strand tends to assume a **right-handed helical conformation** that is stabilized by base-stacking interactions, particularly between adjacent purine bases.
- **Secondary Structures:** Self-complementary sequences within a single RNA strand allow it to fold back on itself, forming complex secondary structures. Breaks in the helix caused by **mismatched or unmatched bases** result in structures like **bulges** and **internal loops**. The most common secondary structure is the **hairpin**, and specific short sequences, like "UUCG," are often found at the loop's end, forming particularly tight and stable structures that can act as starting points for the molecule's folding.
- **Double-Helical Form:** When an RNA strand pairs with a complementary RNA or DNA strand, the resulting duplex is always an **A-form** right-handed double helix, not the B-form helix typical of DNA.

The structural integrity of both DNA and RNA depends on non-covalent forces that can be disrupted, a key chemical property with biological significance.

## 6. The Chemistry of Stability: Denaturation and Renaturation

The non-covalent forces that hold the DNA double helix together, while strong in aggregate, can be disrupted by changes in the cellular environment, such as high heat or extremes of

---

pH. This process of unwinding is known as denaturation and is both reversible and fundamental to biological processes like DNA replication and transcription.

- **Denaturation (Melting):** This is the process where the double helix unwinds into two single strands. It is caused by the disruption of hydrogen bonds and base-stacking interactions. Importantly, **no covalent bonds are broken** during denaturation.
- **Renaturation (Annealing):** When physiological conditions are restored, separated strands can spontaneously rewind. This process occurs in two steps: first, a relatively slow **nucleation** step where the two strands find each other via random collision and form a short complementary duplex. This is followed by a much faster "**zippering**" step, where the remaining bases rapidly come into register.
- **Melting Temperature (T<sub>m</sub>):** The **T<sub>m</sub>** is a characteristic property of a DNA molecule, defined as the temperature at which half of the DNA has been denatured into single strands.
- **T<sub>m</sub> and Base Composition:** The T<sub>m</sub> of a DNA molecule is directly related to its G+C content. A higher percentage of G+C pairs results in a higher T<sub>m</sub>. This is primarily because G≡C base pairs have **stronger stacking interactions** than A=T pairs, contributing more to the overall stability of the helix.
- **Hyperchromic Effect:** The nitrogenous bases in single-stranded DNA absorb more ultraviolet (UV) light at a wavelength of 260 nm than when they are stacked in a double helix. This phenomenon, called the hyperchromic effect, provides a simple method for monitoring the process of DNA denaturation in the laboratory.

Ultimately, the elegant interplay between DNA's rigid chemical code and its dynamic structural potential is what makes it the universal and enduring blueprint of life.

---