MOLECULAR BIOLOGY

Nucleic Acid

DNA and RNA: bases, base pairing, structure, and the flow of genetic information, with a 10-question practice set with full solutions.

SUBJECT	LEVEL	INCLUDES
Biology	High school and early college	Practice set, answer key, revision sheet

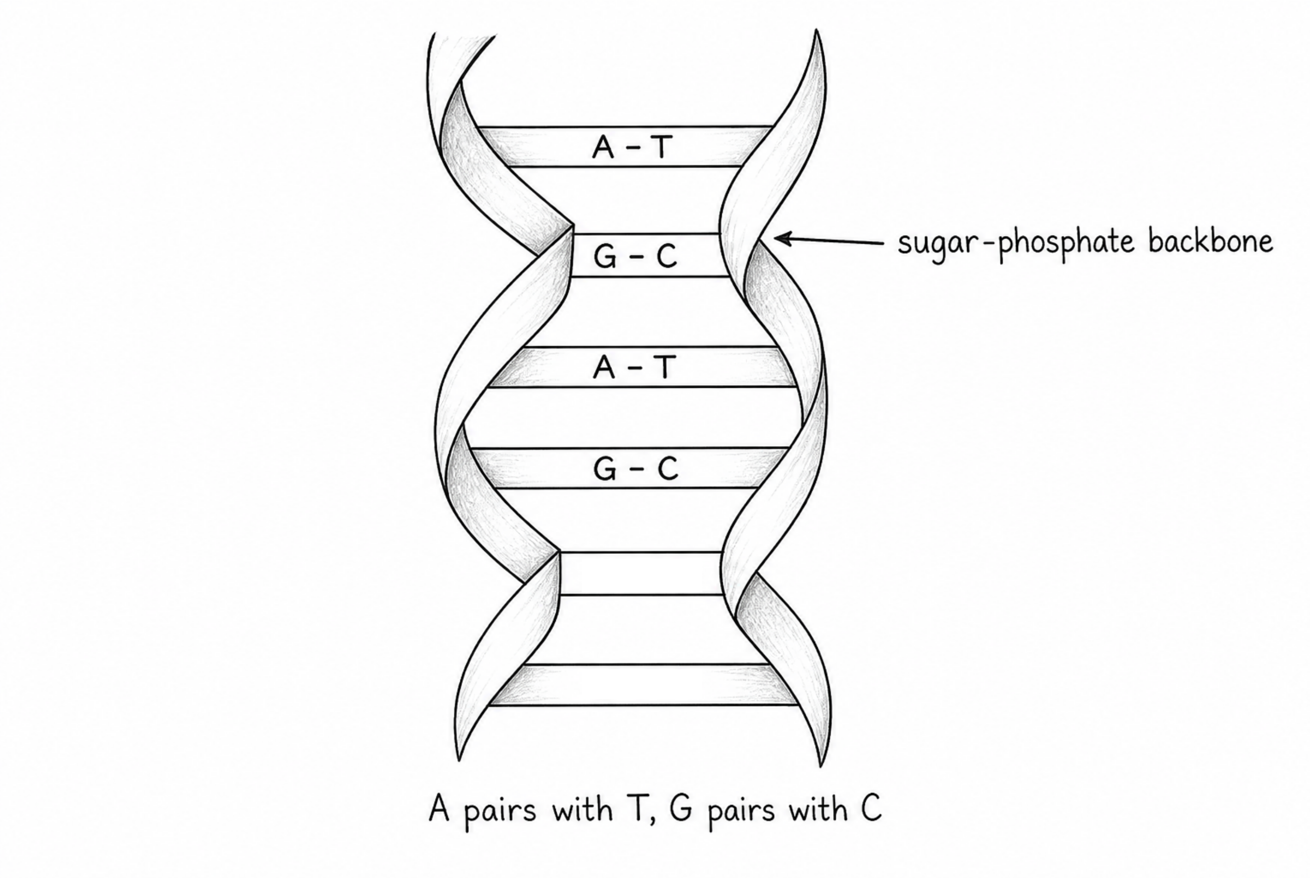
Read this note online, with updates: gauravtiwari.org/study-notes/nucleic-acid

INSIDE

Nucleic acids are polynucleotides, and along with polysaccharides and polypeptides, they form the true macromolecular fraction of every living cell. You'll find them working alongside proteins as the main constituents of chromosomes. Their central job is transmitting characters from one generation to the next, a phenomenon you know as **heredity**.

Structure of Nucleic Acids

The building blocks of nucleic acids are **nucleotides**. Each nucleotide has three chemically distinct components: a heterocyclic compound (the nitrogenous base), a monosaccharide (the sugar), and a phosphoric acid or phosphate residue.



The double helix: antiparallel sugar-phosphate backbones with A-T and G-C rungs holding the ladder together.

The nitrogenous bases you need to remember are **adenine, guanine, uracil, cytosine**, and **thymine**. Adenine and guanine are substituted purines, while the remaining three are substituted pyrimidines. The sugar in polynucleotides is either beta-D-ribose or beta-D-2-deoxyribose.

Here is the key distinction: a nucleic acid containing deoxyribose is called deoxyribonucleic acid (DNA), and one that contains ribose is called ribonucleic acid (RNA). Nucleic acids belong to the broader category of biomolecules that drive virtually every process in living organisms. With that foundation in place, let's explore each component in detail.

Purines, Pyrimidines, and Nucleosides

Before you can understand nucleic acids properly, you need to get comfortable with two classes of heterocyclic bases: the **purines** and the **pyrimidines**. Pyrimidines are six-membered heterocyclic rings, whereas purines contain both a five-membered and a six-membered ring fused together. Because nitrogen atoms sit within these rings, the compounds are classified as **heterocyclics** and behave as bases.

In nature, you'll encounter five important bases in nucleic acids: two purine bases (adenine and guanine) and three pyrimidine bases (cytosine, thymine, and uracil). These natural bases differ from each other in their ring substituents. Some modified versions, such as 5-fluorouracil, are used in cancer chemotherapy.

Every base has a lowermost nitrogen bonded to a hydrogen and two carbons. This -NH- group shares chemical similarities with an alcohol (-OH) group. Just as two sugar molecules can link when an -OH group of one monosaccharide reacts with another, a purine or pyrimidine can bond to a sugar molecule through that -NH- group.

When a purine or pyrimidine base links to a sugar molecule (usually beta-D-ribose or beta-D-2-deoxyribose), you get a **nucleoside**. The bond forms between C-1 of the sugar and either the purine nitrogen at position 9 or the pyrimidine nitrogen at position 1, releasing a water molecule in the process.

The naming of nucleosides reflects the parent base. For example, adenine combined with beta-D-ribose yields **adenosine**, while cytosine combined with beta-D-2-deoxyribose yields **deoxycytidine**.

Nucleotides

Nucleotides are phosphate esters of nucleosides. Think of them as nucleosides with at least one phosphate group attached. The ester linkage can form with the hydroxyl group at position 2, 3, or 5 of ribose, or at position 3 or 5 of deoxyribose. When two or more phosphate residues link together, you get a high-energy phosphate anhydride bond.

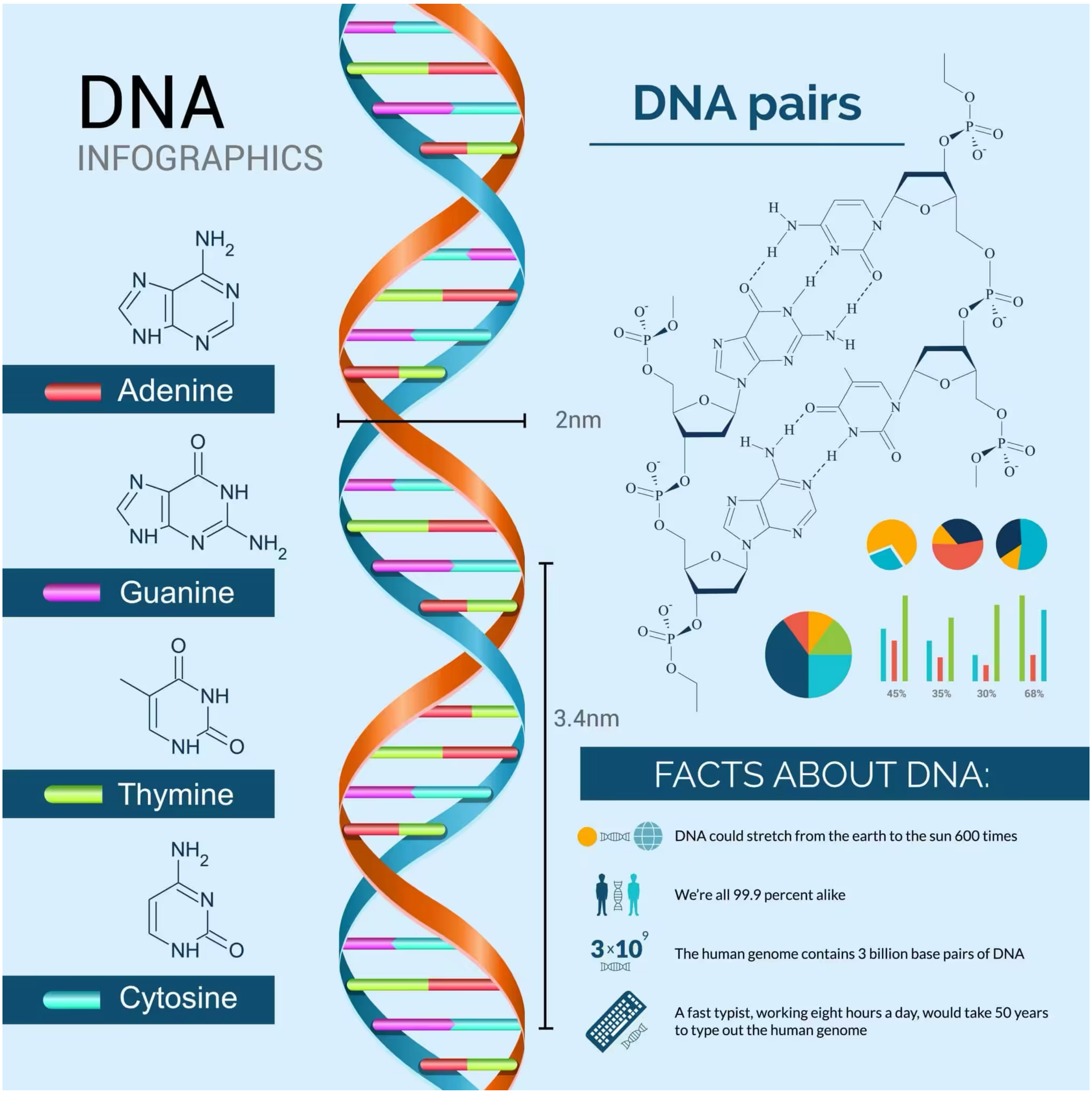
Why should you care about that bond? Because nucleotides drive energy transfer in your cells. **Adenosine diphosphate** (**ADP**) and **adenosine triphosphate** (**ATP**) are the most important examples. Part of the energy released from biological oxidation of the food you eat is stored in the phosphate anhydride bonds of ADP and ATP.

Your body releases that stored energy by reversibly hydrolysing these high-energy bonds, yielding about 35 kJ per mole of ATP. Every natural process, from muscle contraction and nerve impulse conduction to vision and your heartbeat, depends on the energy obtained from ATP. This is one reason why nucleic acid chemistry isn't just an abstract concept. It's the engine behind the energy currency that keeps you alive.

KEY CONCEPT

ATP is often called the ”energy currency” of the cell. Each ATP molecule stores energy in its phosphate anhydride bonds, releasing approximately 35 kJ/mol when hydrolysed. Your body recycles its entire supply of ATP roughly every 1-2 minutes during intense activity, which means you turn over your own body weight in ATP every single day.

Deoxyribonucleic Acid (DNA)



DNA Infographic: scientific research infographics set with charts and diagrams vector illustration

DNA is the molecule that started the dance of life on earth. It is a long polymer of deoxyribonucleotides, and its length is usually expressed as the number of nucleotides or base pairs it contains. The haploid content of human DNA is **3.3 x 10^9 base pairs**, which gives you a sense of how enormous this molecule truly is.

Friedrich Miescher first discovered DNA in 1869, naming it "nuclein" because he found it as an acidic substance in the nucleus. The real breakthrough came in 1953, when **James Watson and Francis Crick** used X-ray diffraction data from Maurice Wilkins and Rosalind Franklin to propose their now-famous **double helical model**. This work earned Watson, Crick, and Wilkins the Nobel Prize.

A cornerstone of their model was base pairing between the two polynucleotide strands. This was consistent with Erwin Chargaff's observation that in double-stranded DNA, the ratio of adenine to thymine and guanine to cytosine always equals one. You'll see this referred to as **Chargaff's rule**.

Here are the key features of the Watson-Crick model that you should know:

- DNA consists of two polynucleotide chains with a sugar-phosphate backbone on the outside and nitrogenous bases projecting inward.
- The two chains run in **anti-parallel** polarity: one strand runs 5' to 3', and the other runs 3' to 5'.
- Bases on opposite strands pair through hydrogen bonds. Adenine pairs with thymine via two H-bonds, and guanine pairs with cytosine via three H-bonds. A purine always sits opposite a pyrimidine, keeping the helix width uniform.
- The two chains coil in a **right-handed fashion**. The pitch of the helix is 3.4 nm, with roughly 10 base pairs per turn. The distance between consecutive base pairs is about 0.34 nm.
- Base pairs stack on top of each other within the helix. This stacking, combined with hydrogen bonding, gives the double helix its remarkable stability.

DNA Replication

One of the most critical functions of DNA is its ability to replicate itself before cell division. Watson and Crick's model immediately suggested how this could work: since each strand carries the complementary information of the other, the two strands can separate and each one can serve as a template for building a new partner strand.

This process is called **semiconservative replication**, and it was experimentally confirmed by Matthew Meselson and Franklin Stahl in 1958 using density-gradient centrifugation with ^{15}N-labeled *E. coli* DNA. Each daughter DNA molecule ends up with one original (parental) strand and one newly synthesized strand.

Replication doesn't happen randomly along the molecule. It begins at specific sites called **origins of replication**. In *E. coli*, there's a single origin (oriC), while human cells have thousands of origins spread across their 46 chromosomes. At each origin, the enzyme **helicase** unwinds the double helix, creating a replication fork that moves along the DNA.

The enzyme **DNA polymerase III** does the heavy lifting of adding new nucleotides to the growing strand. It reads the template strand in the 3' to 5' direction and synthesizes the new strand in the 5' to 3' direction. Because the two strands of DNA run anti-parallel, replication on one strand (the leading strand) proceeds continuously, while the other strand (the lagging strand) is synthesized in short fragments called **Okazaki fragments**. These fragments are later joined by **DNA ligase** to produce a continuous strand.

The accuracy of DNA replication is remarkable. DNA polymerase has a built-in proofreading ability, catching and correcting about 99% of mismatched bases during synthesis. Combined with post-replication mismatch repair systems, the final error rate drops to roughly 1 mistake per 10^9 to 10^{10} base pairs copied. That level of fidelity is essential for preserving genetic information across generations.

Packaging of DNA Within the Nucleus

If you stretched out the DNA from a single human cell, it would measure roughly 2.2 metres. The nucleus of a typical cell, on the other hand, is only about 10^{-6} m across. So how does such a long molecule fit inside such a tiny space?

Even in prokaryotes, which lack a well-defined nucleus, DNA is not scattered randomly through the cell. Because DNA carries a negative charge, it is organized into large loops held together by positively charged proteins in a region called the **nucleoid**.

Eukaryotes take packaging a step further with a set of positively charged, basic proteins called **histones**. These proteins are rich in lysine and arginine, two amino acids that carry positive charges in their side chains. Eight histone molecules come together to form a **histone octamer**. The role of histones is similar to how proteins serve structural functions throughout the cell.

The negatively charged DNA wraps around this positively charged histone octamer, forming a structure called a **nucleosome**. Each nucleosome holds about 200 base pairs of DNA. When you view nucleosomes under an electron microscope, they have a characteristic **"beads-on-string"** appearance.

This beads-on-string structure is further packaged into chromatin fibres, which coil and condense during the metaphase stage of cell division to form the chromosomes you see in textbook diagrams. This higher-level packaging requires an additional set of proteins collectively called **non-histone chromosomal (NHC) proteins**.

Different Forms of DNA

Scientists have identified more than a dozen structural forms of DNA, each named after a letter of the alphabet. Here are the ones you should know:

- **B-DNA** - The standard form with right-handed coiling and 10 base pairs per turn. This is the structure Watson and Crick described.
- **A-DNA** - Contains 11 base pairs per turn. The base pairs are tilted rather than perpendicular to the helical axis.

- **C-DNA** - Resembles B-DNA but has only 9 base pairs per turn.
- **D-DNA** - Similar to B-DNA but with 8 base pairs per turn.
- **Z-DNA** - The standout exception. Unlike all the forms above, Z-DNA demonstrates **left-handed coiling**.

DNA vs RNA: Key Differences

Students often confuse nucleic acid types, so a side-by-side comparison helps clarify the structural and functional distinctions between DNA and RNA. The table below covers the most important differences you'll encounter in exams and research contexts.

FEATURE	DNA	RNA
Full name	Deoxyribonucleic acid	Ribonucleic acid
Sugar	Beta-D-2-deoxyribose	Beta-D-ribose
Bases	A, T, G, C	A, U, G, C
Structure	Double-stranded (usually)	Single-stranded (usually)
Helix form	B-form double helix (standard)	Variable; forms secondary structures like hairpins and stem-loops
Location	Primarily nucleus (eukaryotes); nucleoid (prokaryotes)	Nucleus and cytoplasm
Function	Long-term storage of genetic information	Protein synthesis, gene regulation, catalysis
Stability	Highly stable; has repair mechanisms	Less stable; more chemically reactive
Unique base	Thymine (has a methyl group)	Uracil (lacks the methyl group)
Replication	Self-replicating	Synthesized from DNA template (transcription)
UV damage	Forms thymine dimers	Less prone to dimer formation

EXAM TIP

An easy way to remember the base pairing difference: DNA uses **T**hymine (think "DNA = **T**ough and stable"), while RNA uses **U**racil (think "RNA = **U**nstable and reactive"). The methyl group on thymine is one reason DNA resists chemical degradation better than RNA.

Ribonucleic Acid (RNA)

RNA is usually single-stranded, though you will encounter double-stranded RNA in certain viruses like the rice dwarf virus and reovirus. There is strong evidence that RNA was the first genetic material on earth. Essential life processes such as metabolism, translation, and splicing all evolved around RNA, and it served as both a genetic material and an enzymatic catalyst in primitive living systems.

However, because RNA acted as a catalyst, it was chemically reactive and inherently unstable. DNA eventually evolved from RNA through chemical modifications (thymine replacing uracil, for instance) and gained special repair mechanisms that made it far more stable. That stability is why DNA took over as the primary genetic material in most living organisms. Today, RNA still serves as the genetic material in certain groups of viruses.

There are three types of non-genetic RNA you should understand:

Messenger RNA (mRNA)

Discovered by Jacob and Monod in 1961, mRNA is synthesized in the nucleus and carries the genetic information needed for protein synthesis. Think of it as the blueprint that travels from your DNA to the ribosome, telling the cell exactly which protein to build.

Ribosomal RNA (rRNA)

rRNA is the largest type of RNA and makes up about 80% of the total cellular RNA. You'll find it in ribosomes, the molecular machines where protein synthesis actually takes place. It provides the structural framework that holds the ribosome together and plays a direct role in catalysing peptide bond formation.

Transfer RNA (tRNA)

Also known as **soluble RNA** or **adaptive RNA**, tRNA is the smallest type of RNA, accounting for roughly 10-15% of total cellular RNA. You'll find tRNA molecules in the cytoplasm, and there are as many types as there are amino acids used in protein synthesis (typically 20). Their job is to pick up specific amino acids from the cytoplasm and deliver them to the ribosome during translation.

Biological Significance of Nucleic Acids

Beyond their roles in heredity and protein synthesis, nucleic acids participate in a wide range of biological processes that you should appreciate. Understanding these functions gives you a fuller picture of why nucleic acid research remains one of the most active fields in molecular biology.

Gene expression and regulation. The flow of genetic information follows what Francis Crick called the **central dogma**: DNA is transcribed into RNA, and RNA is translated into protein. But the process isn't a simple one-way street. Regulatory RNA molecules, including microRNA (miRNA) and small interfering RNA (siRNA), can silence specific genes by blocking or degrading their mRNA transcripts. This is a field called **RNA interference** (**RNAi**), and it earned Andrew Fire and Craig Mello the 2006 Nobel Prize in Physiology or Medicine.

Enzymatic activity. Certain RNA molecules, known as **ribozymes**, can catalyse chemical reactions much like enzymes made of protein. Thomas Cech and Sidney Altman discovered this in the early 1980s, winning the 1989 Nobel Prize in Chemistry. Ribozymes are central to the "RNA World" hypothesis, which proposes that life on earth began with self-replicating RNA molecules before DNA and proteins evolved.

Medical applications. Nucleic acid technology has transformed modern medicine. PCR (polymerase chain reaction, developed by Kary Mullis in 1983) amplifies tiny amounts of DNA for diagnostics. CRISPR-Cas9 gene editing, adapted from a bacterial immune system, allows scientists to precisely modify DNA sequences. And mRNA vaccines, like those developed for COVID-19 by Pfizer-BioNTech and Moderna, work by delivering synthetic mRNA that instructs your cells to produce a viral protein, triggering an immune response without using a live virus.

Practice Questions

Work each question before reading its solution. The set runs from direct recall and substitution to the applied questions that exams actually use to separate grades.

Q1. Name the 3 components of a nucleotide and the bond that links nucleotides into a strand.

SOLUTION

A pentose sugar (deoxyribose or ribose), a phosphate group, and a nitrogenous base. Phosphodiester bonds join the 3' carbon of one sugar to the 5' phosphate of the next, building a sugar-phosphate backbone with bases hanging inward like rungs.

Q2. List the 4 DNA bases, classify them, and state the pairing rules with bond counts.

SOLUTION

Purines (2 rings): adenine, guanine. Pyrimidines (1 ring): cytosine, thymine. A pairs with T through 2 hydrogen bonds; G pairs with C through 3. Purine always faces pyrimidine, keeping the helix a uniform width, and the extra G-C bond is why GC-rich DNA melts at higher temperature.

Q3. A DNA strand reads 5'-ATGCCGTA-3'. Write its complement, oriented correctly.

SOLUTION

Pair base by base: A with T, T with A, G with C, C with G, giving 3'-TACGGCAT-5' aligned under the original. Written from its own 5' end, the complementary strand reads 5'-TACGGCAT-3' (this particular sequence happens to read the same in both directions). The antiparallel bookkeeping, not the pairing itself, is what exams catch students on.

Q4. State 3 structural differences between DNA and RNA and the functional consequence of each.

SOLUTION

RNA uses ribose with a 2'-OH, which makes it chemically reactive and short-lived, fitting a disposable messenger; DNA's missing 2'-OH makes it stable for archival storage. RNA substitutes uracil for thymine, cheaper to make. RNA is typically single-stranded, freeing it to fold into functional 3D shapes, tRNAs and ribozymes, while DNA's duplex protects and pairs the archive.

Q5. Chargaff measured A = T and G = C in every DNA sample. Why was this the decisive clue for Watson and Crick?

SOLUTION

Equal amounts of A and T, and of G and C, across species whose overall composition varied wildly, screamed pairing: each A partnered with a T, each G with a C. The double helix with complementary strands explains the ratios automatically, and, as Watson and Crick noted in 1953, immediately suggests the copying mechanism: each strand templates the other.

Q6. Describe the B-DNA double helix in numbers: handedness, rise, pitch, and diameter.

SOLUTION

Right-handed, about 10 base pairs per turn, 0.34 nm rise per base pair, so a 3.4 nm pitch, and roughly 2 nm diameter, with major and minor grooves where proteins read the sequence. Human DNA totals about 3.2 billion base pairs, roughly 2 meters of it folded into each nucleus 6 μm across.

Q7. State the central dogma with the 3 processes, their locations in a eukaryote, and 1 legal exception.

SOLUTION

DNA replicates (nucleus); DNA is transcribed into mRNA (nucleus); mRNA is translated into protein (ribosomes in the cytoplasm). Information flows DNA → RNA → protein. Retroviruses like HIV run reverse transcriptase, copying RNA back into DNA, the exception that biotechnology borrowed for making cDNA. Protein never back-translates to nucleic acid.

Q8. Why does semiconservative replication follow directly from the structure? What did Meselson and Stahl show?

SOLUTION

Because each strand fully specifies its partner, the duplex can unzip and each half template a new complement, yielding 2 daughters each with 1 old and 1 new strand. Meselson and Stahl (1958) grew bacteria on heavy nitrogen, switched to light, and watched density: after 1 generation all DNA was half-heavy, exactly the semiconservative prediction, ruling out conservative copying.

Q9. Compute: a gene's coding region is 1,200 base pairs. How many amino acids does it encode at most, and why fewer in practice?

SOLUTION

Three bases per codon: $1200/3 = 400$ codons, minus the stop codon leaves at most 399 amino acids. In practice eukaryotic genes carry introns that are spliced out, and the count refers only to exon sequence; the triplet code, though, is exact: 3 nucleotides, 1 amino acid, no overlap in standard genes.

Q10. ATP is ”energy currency” yet also a nucleotide. Explain the double life of nucleotides beyond information storage.

SOLUTION

ATP is adenosine, a ribose plus adenine, with 3 phosphates whose terminal bond powers most cellular work. NAD^+, FAD, and coenzyme A are nucleotide-built electron and group carriers; cyclic AMP is a hormone second messenger; GTP powers protein synthesis and signaling. The same building blocks run the cell’s ledger and its power grid, a strong hint that RNA-world chemistry came first.

Answer Key

QUICK ANSWERS

Q1 sugar + phosphate + base; phosphodiester links **Q2** A=T (2 bonds), G≡C (3 bonds) **Q3** 3’-TACGGCAT-5’ **Q4** ribose vs deoxyribose; U vs T; single vs double strand **Q5** the ratios implied 1-to-1 base pairing

Q6 right-handed, 10 bp/turn, 0.34 nm/bp, 2 nm wide **Q7** replication, transcription, translation; reverse transcription **Q8** each strand templates its complement; half-old daughters **Q9** ≤ 399: 3 bases per amino acid, minus stop **Q10** ATP, NAD^+, FAD, cAMP: nucleotides moonlight as energy and signals

Revision Sheet

The monomer

Nucleotide = sugar + phosphate + base; phosphodiester backbone, antiparallel strands.

DNA vs RNA

Deoxyribose/ribose, T/U, double/single strand: archive vs working copy.

Pairing

A=T (2 H-bonds), G≡C (3). Purine faces pyrimidine; GC-rich melts hotter.

Central dogma

DNA → RNA → protein; retroviruses add RNA → DNA.

B-DNA numbers

Right-handed, ~10 bp/turn, 0.34 nm/bp, 2 nm wide; 3.2 Gbp per human genome.

Beyond storage

ATP, GTP, NAD^+, FAD, cAMP: nucleotides as energy and signaling hardware.

About These Notes

This note is part of a free study-notes series on gauravtiwari.org covering mathematics, physics, chemistry, and biology: the concepts that keep deciding grades in high school, board, and entrance exams.

2008
Writing and teaching online since

2,000+
Articles published on gauravtiwari.org

125+
Free study notes in this series

M.Sc.
Mathematics, with physics and chemistry training

Gaurav Tiwari is a developer, educator, and founder of Gatilab. The study-notes series exists because most exam prep material is either a full textbook or a thin definition page, and revision needs something in between: one concept, complete, with the traps named and the practice attached.

GET THE FULL SERIES

Every note in the series is free to read online and free to download as a PDF like this one. Browse the complete collection at **gauravtiwari.org/free-study-notes-exam-prep** and pick the topics your next exam actually covers.

Gaurav Tiwari gauravtiwari.org