Deoxyribonucleic Acid – Structure, Organization and Various Types of DNA

The chemical structure that carries genetic information is called DNA. In this article, you will learn about the structure, organization and various types of DNA. For exams, the phases of the cell cycle are of great importance. After reading this article, you can test your newly acquired knowledge with a few exam questions on this topic.

Design and Structure of DNA

Each cellular nucleus of the human body contains deoxyribonucleic acid (DNA), which carries the genetic information of humans (and most other organisms) on its 46 homologous chromosomes (proven by Oswald T. Avery in 1944). DNA is the “blueprint of the body”. The coding units on it are called genes. The filamentous macromolecule consists of a double helix in which two nucleotide strands are connected with each other through hydrogen bonding of their base pairs.
DNA consists of nucleotides. A **nucleotide** consists of a base (adenine, cytosine, guanine and thymine—the actual carriers of information), a sugar (with one less oxygen than the ribose from which it emerges, hence the name “deoxy”-ribose; DNA contains D-ribose or 2-deoxy-D-ribose) and a phosphate group. The length of a DNA segment is measured in the number of base pairs (bp) or nucleotides (nt). If the segment is larger, kilobases (1kb
= 1000 bp) or megabases (1Mb = 1000 kb) are used.

Nucleosides, Nucleotides: An Overview of Important Abbreviations

(Organic) bases: Adenine, guanine, thymine, cytosine and uracil. Other related bases: hypoxanthine and xanthine.

Tip: If you can recognize, differentiate and draw these bases, you can be sure to score some important points in your exams.

Nucleoside: organic base and a pentose (D-ribose or 2-deoxy-D-ribose): adenosine, guanosine, thymidine, cytidine and uridine, which are each abbreviated by their first letter (A, G, T, C, U). Other related bases are inosine and xanthosine.

Nucleotide: organic base, a pentose (D-ribose or 2-deoxy-D-ribose) and a phosphate. They are abbreviated with the nucleoside abbreviations (A, G, T, C, U) plus MP for a monophosphate, DP for a diphosphate, and TP for a triphosphate. Nucleotides containing 2-deoxy-D-ribose are additionally prefixed with a small d-. Examples: 5′-adenosine monophosphate = AMP; guanosine-3′-diphosphate = GDP; 5′-deoxyadenosine triphosphate = dATP.
- **Nucleoside monophosphate** (e.g., adenosine monophosphate = AMP): Nucleoside (organic base and pentose) and one phosphate group, which is esterified to the 5′ carbon, unless otherwise described.

- **Nucleoside diphosphate** (e.g., adenosine diphosphate = ADP): Nucleoside (organic base and pentose) and two phosphate groups, which are esterified to the 5′ carbon, unless otherwise described.

- **Nucleoside triphosphate** (e.g., adenosine triphosphate = ATP): Nucleoside (organic base and pentose) and three phosphate groups, which are esterified to the 5′ carbon, unless otherwise described.

In DNA, the **purine** bases adenine (A) and guanine (G) always pair with the **pyrimidine** bases thymine (T) and cytosine (C).

**Note:** The pyrimidine bases thymine and cytosine have a “y” in their name. (Find out more about nucleotide metabolism [here.](#)

In RNA, uracil instead of thymine is found predominantly (thymine is different from uracil only by a methyl group at position five). Adenine and thymine form two hydrogen bonds between each other; guanine and cytosine form three hydrogen bonds. All these bases are subject to **tautomerism**: for the base pairing, they must exist in the **keto** form.
Nucleotides create nucleotide chains, in which the mononucleotides form a 5′-3′ phosphodiester bond: the phosphate at position 5′ of one mononucleotide bonds with the deoxyribose at position 3′ of the other mononucleotide. Thus, a polynucleotide has a free 5′-phosphate end and a free 3′-hydroxyl end.

In DNA, two complementary nucleotide strands form a duplex. Depending on their orientation, they are called the 5′-3′ strand (coding strand or sense strand) and the 3′-5′ strand (non-coding strand, antisense strand or template strand). The sequence of nucleotide bases in the coding strand in 5′-3′ orientation corresponds to the complementary sequence of bases of the template strand in 3′-5′ orientation. The individual strands run antiparallel to each other.

Due to base pairing, the number of adenine present in one molecule of DNA is always equal to the number of thymine (A = T) and the number of guanine is always equal to that of cytosine (G = C) (Chargaff’s rule). This also implies that the sum of the purines (A + G) is always equal to the sum of the pyrimidines (T + C).
Since the application of this rule is sometimes tested in exams, here’s an example:

**Given:** C = 29 %. (Application (1): C = G; A = T)

**It follows:** G = 29 %. (Application (2): Purine A + G = Pyrimidine T + C; (A + G) + (T+C) = 100 %)

**Further:** A + 29 % = T + 29 %

**Therefore:** A and T each are 21 %; C and G are 29 %.

The DNA Double Helix

The two nucleotide strands (double strands) constituting DNA are not arranged in one plane but twisted into a **double helix.** James D. Watson and Francis H.C. Crick first described it correctly in 1953, and today, still, it is called the **Watson-Crick model.** In 1962, they received the Nobel Prize in Medicine.

The abbreviations **dsDNA** (double stranded DNA) and **ssDNA** (single stranded DNA) are
used to describe whether DNA is present in its double helix form or not.

The hydrophobic, positively charged bases with their hydrogen bonds are located inside; the neutral sugar rings and negatively charged phosphates form the negatively charged outside.

This characteristic is also of use in histochemistry for staining histological preparations: The blue-black dye hematoxylin is alkaline. It, therefore, tends to electrochemically attach to the acidic DNA of the nucleus, which makes the nuclei appear dark blue in the HE staining.

In vivo, the DNA double helixes are usually found in their B conformation (B-DNA). Rare forms are the left-handed Z form and the A form, which does not occur in vivo. The B form of the double helix is right-handed (referring to the sense in which it is wound around the axis). Its diameter is about 2 nm; the distance between two adjacent base pairs of a strand is 0.34 nm. One turn of the double helix covers ten base pairs, which results in a height of 3.4 nm. The stacking of the bases provides for further stabilization of the B form as the electrons of the overlapping bases exert so-called stacking forces. The helical structure forms a large groove and a small groove between its two strands.

Organization and Packaging of DNA: Histones, Nucleosomes and Chromatin

If the DNA molecule of a single human cell were laid out straight, without further packaging, it would be about six feet long. In order to fit into the nucleus and be protected against external influences and shearing forces, DNA must be in a condensed state. Various condensation methods have been observed, each involving certain proteins binding to the DNA molecules.

DNA: From Histone to Nucleosome and Solenoid to Chromatin

From small to large, the double helix undergoes the first compression through histone proteins, around which it winds in 1 and 2/3 turns. The alkaline histones are predominantly comprised of the alkaline, positively charged amino acids arginine and lysine, which allows for easy accumulation to the negatively charged backbone of the DNA via ionic interactions.

Like most proteins, the histones are synthesized in the cytosol. They are divided into five different classes: H1, H2A, H2B, H3 and H4. The DNA winds around a histone octamer, a structure formed by two molecules of each of the four histone proteins H2A, H2B, H3 and H4. The histone octamer and the 146 base pairs (bp) of the DNA wound around it form a nucleosome with a diameter of 11 nm.

The DNA found between the nucleosomes is called linker DNA and is up to 80 bp in length. The nucleosomes along with the linker DNA form a nucleosome strand. Histone molecules of the H1 class bind to the linker DNA. The nucleosome strand twists on itself and forms the 30 nm long solenoid strand. This fiber in turn forms chromatin loops, using various non-histone proteins, which form the macrostructure of the chromosomes in the nucleus.
For chromatin, we distinguish between **euchromatin** (Greek *eu* = well, good) and **heterochromatin**. The “good” **euchromatin** is less condensed and thus looser, which makes it more easily read by the enzymes for transcription; it is commonly found in metabolically active cells. **Heterochromatin** is maximally condensed, transcribed less often and can be found in cells with low metabolism.

### Histone Modification

Through modifications of the histones, the transcription of DNA can be altered. Alkaline histones, which mainly consist of the amino acids lysine and arginine, attach to the acidic DNA. Additional molecules can bind to their positively charged free groups:

- **Acetylation** refers to an acetyl group (chemical formula: COCH$_3$; a carbonyl and a methyl group) attaching to the lysine group of a histone octamer. Acetylation leads to an opening of the DNA and **increases** the rate of transcription.

- **Methylation** refers to the addition of methyl (CH$_3$) to a lysine or arginine group of a histone. Methylation may correlate with both an increased or a decreased transcription rate, depending on the methylated amino acid. DNA can also be methylated directly, which mostly occurs in regions with a high frequency of CpG sites, so-called **CpG islands**. As part of epigenetics, methylation of DNA can be inherited.

- **Phosphorylation** (attachment of a phosphate group) may occur on the free hydroxyl (OH) groups of some amino acids, i.e., on serine, threonine and tyrosine. Similar to methylation, this can have an effect of both increased and decreased transcription.

Other possible modifications are **ubiquitylation** and **ADP-ribosylation**.

### RNA: Types and Differences to DNA
Ribonucleic acid (RNA) functions as the genetic carrier of some viruses. In humans, RNA contributes in many different ways to the protein biosynthesis and is involved, among other things, in the transcription and translation but may also have a catalyzing or regulatory function:

- The **hnRNA** (heterogeneous nuclear RNA) can be found in the nucleus of eukaryotic cells as a precursor of the mature mRNA and is, therefore, also referred to as pre-mRNA.
- The **mRNA** (messenger RNA) is used as a complementary copy of a DNA strand during transcription and migrates from the nucleus to the cytosol to serve there as a template for translation.
- The **tRNA** (transfer RNA) acts as a mediating link during translation, bringing together the mRNA triplets and the amino acids for which they code.
- The **rRNA** (ribosomal RNA) along with some proteins forms the ribosome structure, whose subunits make translation possible.
- The **mtRNA** (mitochondrial RNA) consists of mitochondrial rRNA, tRNA and mRNA, which function in the same manner as the eukaryotic molecules of the
same name.
- The snRNA (small nuclear RNA), together with certain other proteins, can form a spliceosome, which can then initiate the splicing of certain genes.
- The snoRNA (small nucleolar RNA) act as a guide in the modification of rRNA and snRNA.
- The scRNA (small cytoplasmic RNA) belong to the signal recognition particles (SRP), which target proteins that are destined for the extracellular space and carry them to the endoplasmic reticulum for further transport.
- The siRNA (short interfering RNA) serves a regulatory function by binding to specific mRNA bases and inducing their degradation.
- The miRNA (micro RNA) serves a regulatory function by inhibiting the further processing of mRNA through specific base pairing.

The components of the RNA can be deduced from its name: Ribonucleic acid, which contains D-ribose, unlike the deoxy-D-ribose of the DNA. RNA differs from DNA in other aspects as well: Its bases are adenine, cytosine, guanine and uracil. Uracil replaces thymine and pairs with adenine. The two bases differ only by one methyl group at position 5. Thymine is rarely found in RNA, one exception being (transfer) tRNA.

RNA is usually found as a single strand. Within the RNA strand, corresponding bases can pair through hydrogen bonds, and thus form, e.g., the cloverleaf structure of tRNA. Moreover, RNA can establish intermolecular base pairings, which gives it the form of a double helix with A conformation; this way, it is easier for lighter secondary structures similar to histone modifications (see above) to attach to it.
Other Forms of DNA

Mitochondrial DNA (mtDNA)

Mitochondria, the “powerhouses of the cell”, have their own DNA, the mitochondrial DNA (mtDNA; this fact is also seen as evidence of the endosymbiosis theory, which states that mitochondria were independent bacteria-like organisms before they were consumed by eukaryotic precursor cells to assume specific tasks for the cells).

The mtDNA is located inside of the mitochondrion (matrix), comprising 16 kb, in annular form or as a double strand. The mitochondrial genes on the mtDNA code for the compounds of the respiratory chain, for mitochondrial mRNA and tRNA.

Mitochondrial DNA is maternally inherited, i.e., from the mother. The paternal mitochondria are located at the neck of the sperm, which does not fully penetrate the egg. In addition, the egg has a degradation mechanism against paternal mitochondria.

Bacterial DNA

The DNA of prokaryotes (single-cell organisms without a nucleus, e.g., bacteria), is freely located in the cytoplasm in the form of a chromosome or plasmid. Both structures are circular and consist of a right-handed double helix, which is, due to reasons of energy and space, twisted additionally in left orientation around the helix axis.

This additional twist is carried out by topoisomerases that only occur in bacteria: DNA gyrase and topoisomerase IV. The entire gene expression is subject to fewer repair mechanisms than in humans because mutations are a desirable feature in the evolutionary strategy of the bacteria. Furthermore, through various mechanisms, bacteria can exchange DNA plasmids among themselves.

Viral DNA

The viral genome is either present as DNA or RNA, depending on the nucleic acid type. DNA viruses have a double-stranded, linear or circular DNA and often a large genome, which is relatively stable (for example: poxviruses, herpes viruses). Single-stranded DNA viruses are very rare (e.g., the parvovirus B19, which causes erythema infectiosum).

RNA viruses, on the other hand, have single-stranded (ss) RNA with limited genome size, which is subject to frequent mutation due to the lack of correction mechanisms. It is, however, very adaptable (e.g., flavivirus or HIV virus). Double-stranded RNA viruses are very rare (e.g., rotavirus).

Review Questions

The right answers can be found below the references.

1. In the case of the double-stranded DNA of a cell, the proportion of dGMP is about 20%. Approximately how high is the proportion of dAMP (according to the Chargaff rule)?
   
   A. 20 %
   B. 30 %
   C. 40 %
2. Which statements regarding the base pairing of nucleic acids are true?

A. Base pairing is only possible between deoxyribonucleotide strands.
B. The 2’OH groups of ribonucleotides prevent base pairing between two ribonucleic acid strands.
C. Thymine, cytosine and guanine have to be present in the lactam (keto) form for base pairing to occur.
D. During base pairing, hydrogen bonds (H-bonds) are formed between two opposite purine bases.
E. The formation of intermolecular hydrogen bonds within a nucleic acid is impossible for steric reasons.

3. Which statement about the histones is true?

A. The site of synthesis is the cell nucleus.
B. The histone octamer is composed of one H1, H2, H3 and H4 dimer.
C. Histones contain more alkaline than acidic amino acids in their amino acid chain.
D. Acetylation of the H4 histone protein D increases its inhibitory effect on transcription.
E. Histones control gene expression through DNA methylation.

4. Which statement on mitochondrial DNA (mtDNA) in humans is true? The mtDNA ...

A. ...is annular.
B. ...is associated with the histones H2A and H2B.
C. ...contains about the same number of exons and introns.
D. ...contains genes for cytosolic proteins.
E. ...is exclusively inherited paternally.

References


Correct answers: 1B, 2C, 3C, 4A

Legal Note: Unless otherwise stated, all rights reserved by Lecturio GmbH. For further legal regulations see our legal information page.