Chemical organization of cells

The most frequent chemical elements of living cells are H, C, N, O, P, S that represent 99% of cell mass. This elements form the organic molecules: nucleic acids, proteins, carbohydrates, and lipids. Organic substances are represented by ***simple substances*** and ***polymers***. The polymers are macromolecules composed of building blocks named ***monomers***. Some of polymers – ***the homopolymers*** - are made from only one type of monomers (ex, cellulose, starch), other – ***the copolymers*** (***heteropolymers)*** – are composed from different types of monomers (ex, DNA, RNA, proteins). Proteins and nucleic acids are responsible for conveying genetic information. Nucleic acids carry the information, while proteins provide the means executing it. The sequence in which the individual building blocks are joined is the critical feature that determines the property of the resulting macromolecule.

A polymer has:

- a ***backbone*** consisting of a regularly repeating series of bonds;
- ***side-groups*** of characteristic diversity that stick out from the backbone.

The process of polymer's synthesis is called ***polymerization*** (or ***polycondensation***). The process of destruction of polymers is named ***hydrolysis***.

Carbohydrates

The general formula of carbohydrates is $C_x(H_2O)_y$, where *x* and *y* can be grater then 3. There are several types of carbohydrates: ***monosaccharides*** (glucose, ribose, deoxyribose), ***disaccharides*** (sucrose), ***oligosaccharides*** (different types of cellular receptors) and ***polysaccharides*** (cellulose, starch, glycogen). The monomers are linked by ***glycosidic bond***. The main functions of carbohydrates are structural, energetic and depositary.

Lipids

Lipids are organic substances insoluble in water, but soluble in nonpolar solvents (chloroform, ether). Lipids can be divided in ***structural lipids*** (lipids from membranes) and ***reserve lipids***.

Functions of lipids:

- ✓ energetic;
- ✓ structural;
- ✓ mechanical;
- ✓ hormonal;
- ✓ vitamins.

Proteins

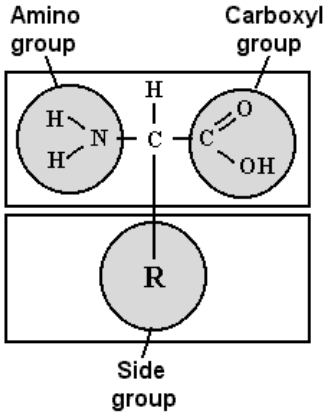

Fig. 1. The general structure of amino acids

Each protein consists of a unique sequence of amino acids. A free amino acid contains an ***amino group***, a ***carboxyl group*** (constant part) and a ***side group*** (variable part) (fig. 1). Depending on the structure of side group there are four types of amino acids: basic, acidic, neutral and hydrophobic. There are about 150 types of amino acids, but in the synthesis of proteins participate only 20.

Amino acids are joined into a chain by ***peptide bonds***, which are created by the condensation of carboxyl (COOH) group of one amino acid with the amino (NH_2) group of the next. A longer chain of amino acids joined in this manner is called a ***polypeptide***. The ***protein*** represents the functional unit, which may consist of one or more polypeptide chains.

A critical feature of protein is its ability to fold into a three dimensional ***conformation***. A major force underlying the acquisition of conformation in all proteins is the formation of ***noncovalent bonds***:

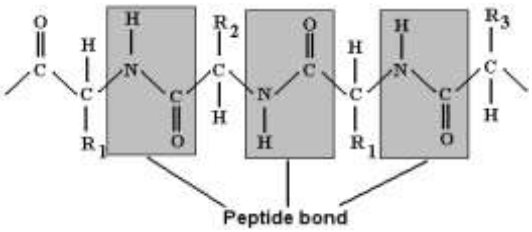

Fig. 2. Primary structure of proteins

- ***Hydrogen bonds*** - are weak electrostatic bonds that form between NH and C=O groups of the peptide backbone. Because many hydrogen bonds can be formed in a macromolecule, their averall contribution to the stability of the conformation can be substantial. But the weakness of the individual hydrogen bond allows them to be broken relatively easily under physical (high temperature) or physiological conditions (action of enzymes).
- ***Ionic bonds*** – occur between groups that have opposite charges (usually between acidic and basic amino acids).
- ***Hydrophobic bonds*** – occur between amino acids with apolar sidechains, which aggregate together to exclude water.
- ***van der Waals attractions*** – very weak interactions between atoms.

There is also a covalent bond - ***disulphide bond*** formed between the sulphur atoms of cysteine.

The conformation can be described in terms of several levels of structure.

- ***The primary structure*** – the series of amino acids linked into polypeptide chain. This linear sequence is the essential information that is coded by the genetic material (fig. 2).
- The primary structure folds into ***secondary structure***, which describes the path that the polypeptide backbone of the protein follows in the space. **Hydrogen bonding** between groups on the same polypeptide chain causes the backbone to twist into a helix (***α-helix***) or sheet-like structure (***β-sheet***) (fig. 3).

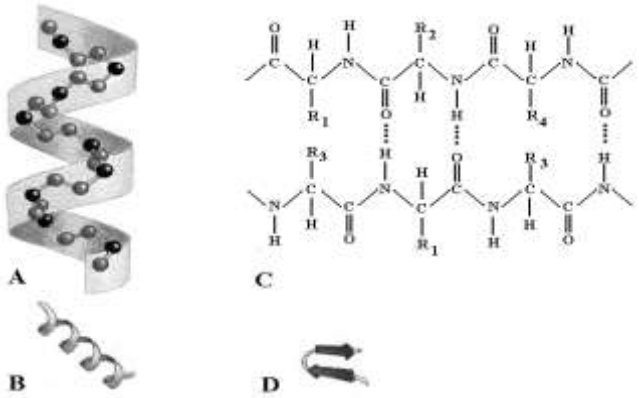

Fig. 3. A, B - α-helix; C, D - β-seets

Fig. 4. Tertiary structure of proteins

- ***The tertiary structure*** descries the organization in three dimensions of all the atoms in the polypeptide chain. This involves side chain interactions and packing of secondary structure motifs (a combination of α-helixes and β-sheets) (Fig. 4).
- The highest level of organization – ***quaternary structure*** – is recognized in ***multimeric proteins***, which consist of aggregates of more than one polypeptide chain. The individual polypeptide chains that make up a multimeric protein are called ***protein subunits***. Fig. 5 represents the structure of hemoglobin, which consists of two α chains ant two β chains.

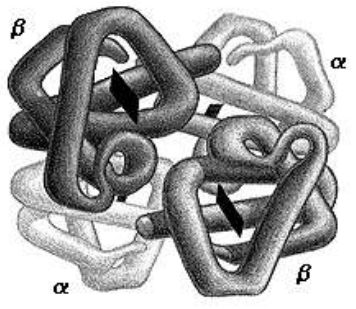

Fig. 5. Quaternary structure of hemoglobin

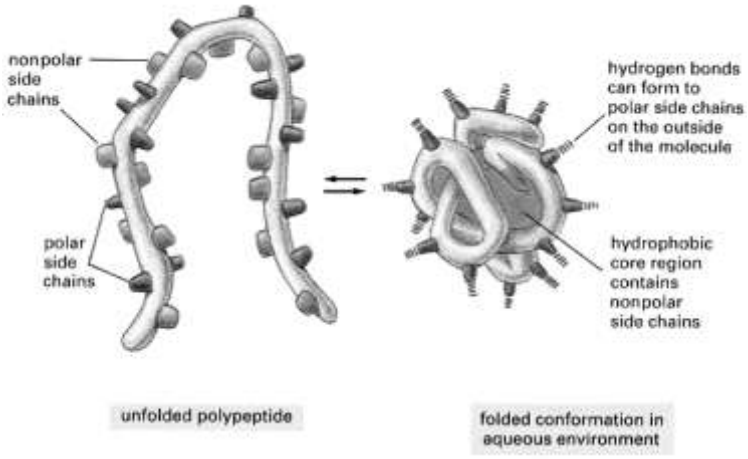

Fig. 6. Denaturation and renaturation of proteins

Some environmental factors (high or low temperature, pressure, light etc.) can convert the protein from the physiological conformation to some other (inactive) conformation by breaking of non-covalent bonds or disulphide bond – ***denaturation of protein***. If only non-covalent bonds are broken is possible the ***renaturation of protein*** by refolding of a previously unfolded protein into its native state (fig. 6).

The proteins are located everywhere in the cell and have many functions:

- ✓ *Catalytic* – serve as enzymes (DNA-polymerase, RNA-polymerase);
- ✓ *Hormonal* – regulation of metabolism (insulin);
- ✓ *Reception* – interaction with hormones or other signal molecules;
- ✓ *Transport* – (Na^+/K^+ pomp, hemoglobin);
- ✓ *Structural* – in membranes, ribosome etc.;
- ✓ *Immunologic* – antibodies (IgA, IgM, IgG).

NUCLEIC ACIDS

There are two types of nucleic acids: ***DNA*** (***deoxyribonucleic acid***) and ***RNA*** (***ribonucleic acid***). A nucleic acid consist of a chemically linked sequence of monomers – ***nucleotides***. Each nucleotide contains a ***nitrogenous base***, a ***pentose sugar*** and a ***phosphate group***. Nitrogenous bases fall into two types: Purines (A, G) and Pyrimidines (C, T, U). The bases from DNA are A, G, C, **T** and the bases from RNA are A, G, C, **U**. In DNA pentose is 2-deoxyribose; whereas in RNA it is ribose (fig. 7).

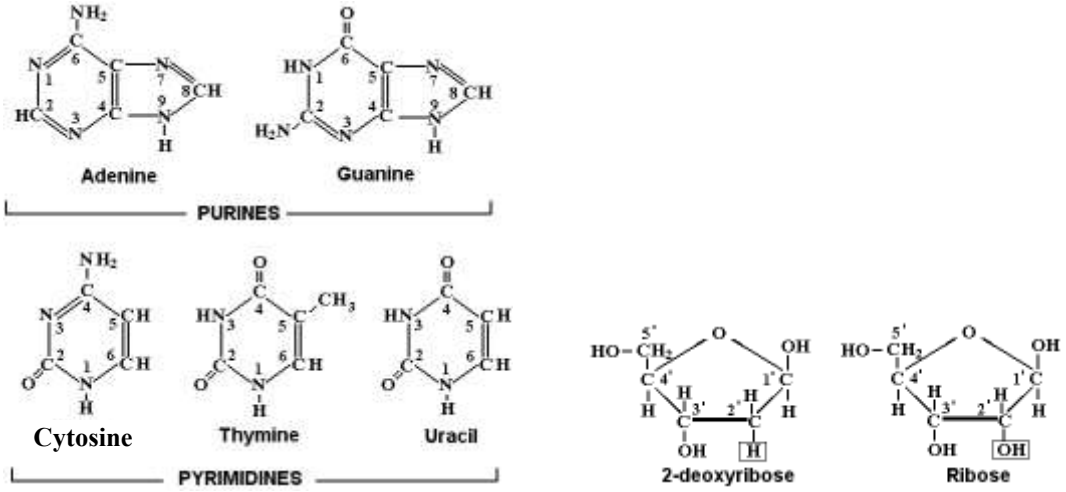

Fig. 7. Nitrogenous bases and pentose sugars – components of nucleotides

The base is linked to position 1' from the pentose ring by a glycosidic bond from N_1 of pyrimidines or N_9 of purines. A base linked to a sugar is called ***nucleoside***. When a phosphate group is added, the base-sugar-phosphate is called ***nucleotide*** (Tab. 1; Fig. 8)

Table 1. Bases, nucleosides and nucleotides

Base	Nucleoside	Nucleotide	Abbreviation (1P)		Abbreviation (2P)		Abbreviation (3P)	
			RNA	DNA	RNA	DNA	RNA	DNA
Adenine	Adenosine	Adenilic acid	AMP	dAMP	ADP	dADP	ATP	dATP
Guanine	Guanosine	Guanilic acid	GMP	dGMP	GDP	dGDP	GTP	dGTP
Cytosine	Cytidine	Cytidilic acid	CMP	dCMP	CDP	dCDP	CTP	dCTP
Thymine	Thymidine	Thymidilic acid		dTMP		dTDP		dTTP
Uracil	Uridine	Uridylic acid	UMP		UDP		UTP	

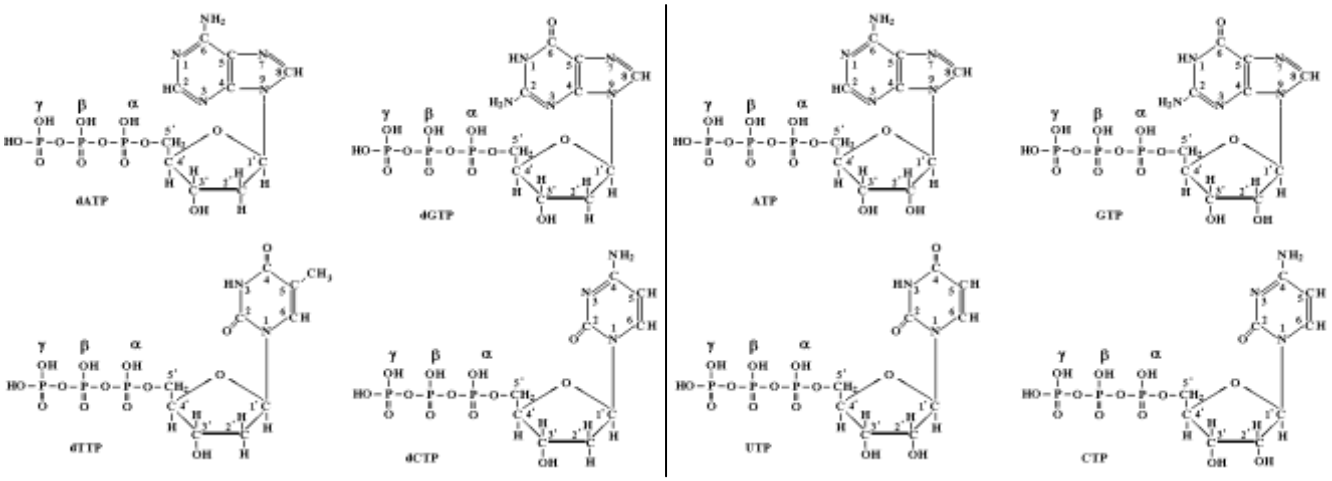

Fig. 8. Nucleoside triphosphates – monomers of DNA (left) and RNA (right)

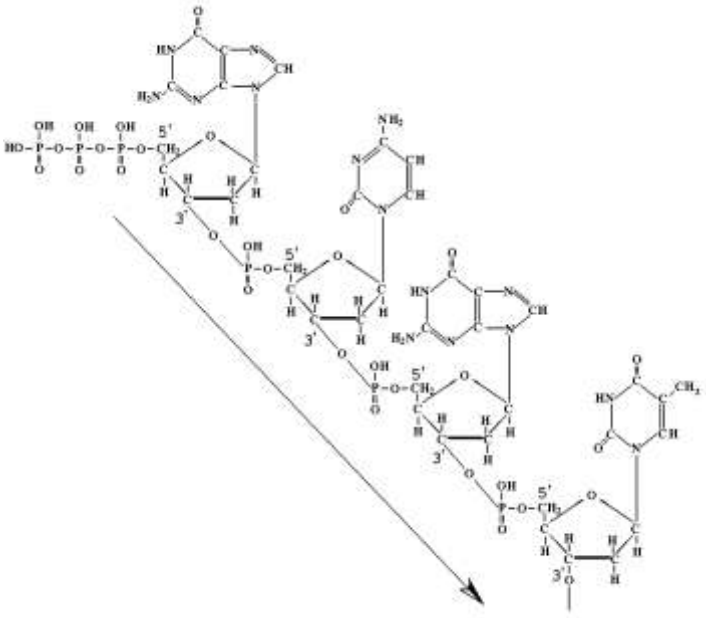

Fig. 9 Polynucleotidic chain

Nucleotides provide building blocks from which nucleic acids are constructed. The nucleotides are linked into a ***polynucleotide chain***. The 5' position of one pentose ring is connected to the 3' position of the next pentose ring via a phosphate group. The bond that link nucleotides is called ***phosphodiester bond*** (fig. 9)

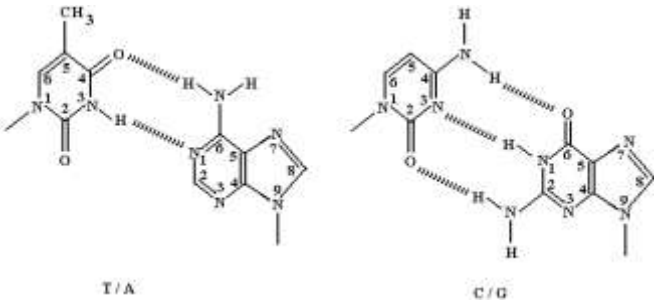

Fig. 10. Complementary base pairing

Number, type and sequence of nucleotides in polynucleotide chain represent ***the primary structure of nucleic acids***.

The molecule of DNA represents a double helix, which consists of two polynucletide chains associated by hydrogen bounding between the nitrogenous bases. There are 3 H-bonds between G and C and 2 H-bonds between A and T (fig. 10). These reactions are described as ***base pairing***, and the paired bases are said to be ***complementary***. The polynucleotide chains run in opposite directions – are ***antiparallel*** (fig. 11).

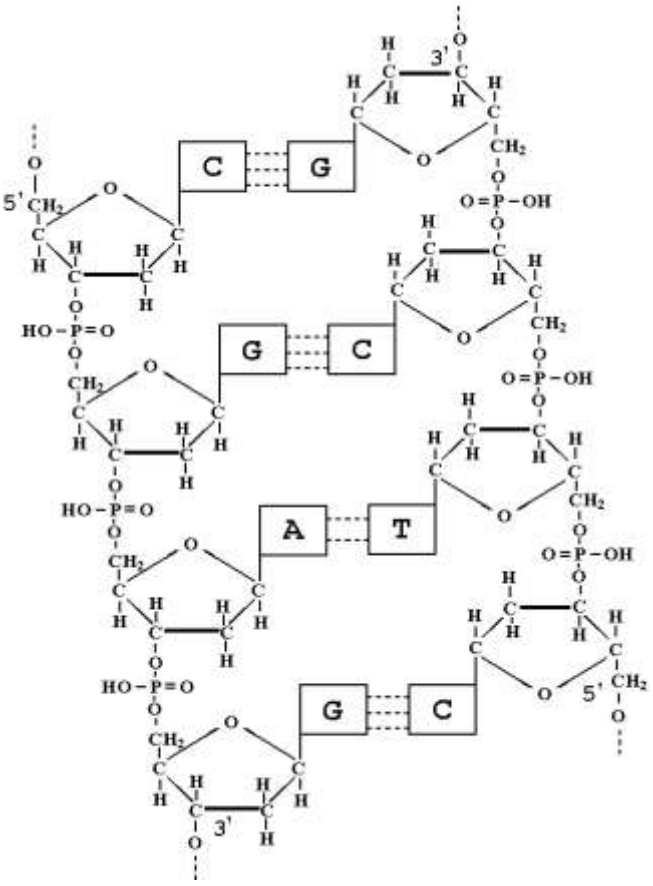

Fig. 11. The double helix of DNA – two antiparallel chains bounded by Hydrogen bonds between a purine and a pyrimidine

The double helix of antiparallel polynucleotide chains linked bay Hydrogen bonds between purines and pyrimidines represent the ***secondary structure of DNA***. This structure was established in 1953 by *James Watson* and *Francis Crick*. Each base pair is rotated ~36^0 around the axis of the helix relative to the next pair. Thus ~10 base pairs make a complete turn of 360^0. The twisting of the two strands around one another forms a double helix with ***narrow groove*** (~12 Ă across) and a ***wide groove*** (22 Ă across). The double helix is ***right-handed*** (fig. 12).

The complementarily bases in DNA assure:

- The stability of DNA;
- Replication;
- Transcription;
- Recombination;
- DNA repair.

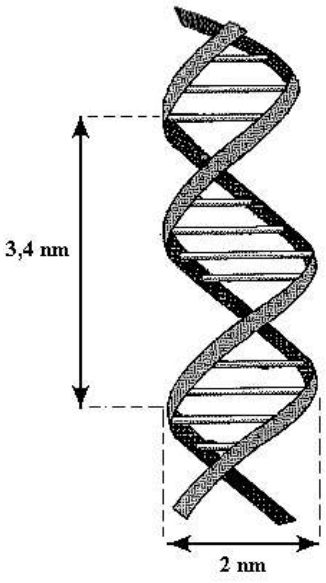

Fig. 12. Structure of B-DNA

Different conditions (usually heating in laboratory experiences or action of enzymes – DNA-helicases – in vivo) determine the conversion of the molecule from double-strand to the single-strand state. This process is called ***denaturation of DNA***. Slow cooling of DNA or action of enzymes may determine the ***renaturation of DNA*** – reassociation of denatured complementary single strand of a DNA double helix.

DNA can exist in more than one type of double-helical structure. Each family represents a characteristic type of double helix described by parameters *n* (number of nucleotides per turn) and *h* (the distance between adjacent repeating units) (Tab. 2). The B-form represents the general structure of DNA in the conditions of the living cell (fig. 12). The A-form is probably characteristic for hybrid duplex with one strand of DNA and one strand of RNA. This may occur during transcription.

The Z-form is only left-handed helix. Its name is taken from the zigzag path that the sugar-phosphate backbone follows along the helix. The Z-form occurs in polymers that have a sequence of alternating purines and pyrimidines.

Table 2. Parameters of some forms of DNA

Parameters of helix	A–DNA	B–DNA	Z–DNA
Sense	Right-handed	Right-handed	Left-handed
Base pair in turn	11	10,4	12
Distance between bases (Å)	2,9	3,4	3,7
Helical diameter (Å)	25,5	23,7	18,4

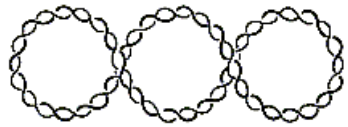

Fig. 13. Supercoiling of the circular DNA

The total length of DNA contained within the genome of an organism varies depending on its genetic complexity, and may vary from micrometers (viruses) or millimeters (bacteria), to several centimeters (in higher eukaryotes). Therefore, the DNA must be organized and packaged into higher order forms within the cell, and DNA molecules in vivo may acquire a compact shape by existing in circular forms, where the two ends of a linear DNA are covalently bound to each other. These circular DNA molecules can be twisted into supercoiled molecules to adopt an even more condensed configuration than the "relaxed" circular equivalent (fig.13.). In prokaryotic cells, this supercoiling of the DNA is critical for the packaging of the DNA into the cells. In prokaryotes the superhelical state of the DNA plays an important part in various biological processes including replication, transcription and recombination, by altering the accessibility of the DNA to proteins (fig. 14.).

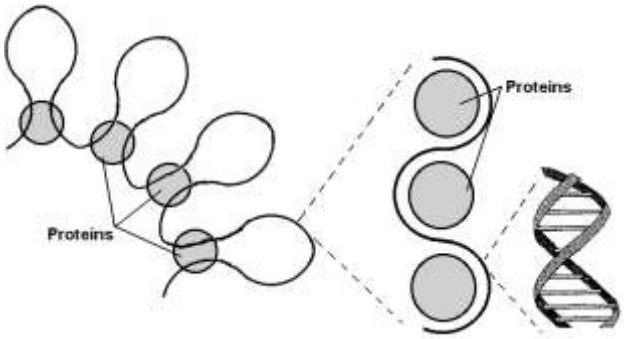

Fig. 14. Packing of DNA in prokaryotes

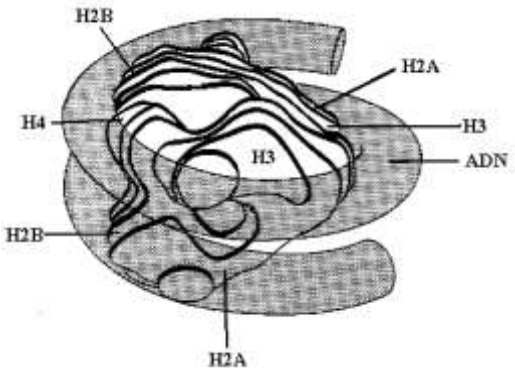

Fig. 15. Interaction of DNA with histones in eukaryotes

In eukaryotes the DNA is packaged in the nucleus (0.5mm in diameter) with the help of basic polypeptides called ***histones*** (fig. 15.). There are 5 classes oh histones: H1, H2A, H2B, H3, H4. Each DNA molecule is packaged in the form of a ***chromosome***. Chromosomes are composed of ***chromatin***, a densely staining material initially recognized in two different forms: highly condensed ***heterochromatin*** and more diffuse ***euchromatin***. Decondensed chromatin resembles "beads on a string". Each bead, called ***nucleosome***, contains about two supercoils of DNA wrapped around a core histone octamer (8 proteins). Interaction of DNA with proteins represents the ***tertiary structure of DNA***.

Proprieties of molecules of DNA

From the perspective of its function in replicating itself and being expressed as protein, the central propriety of the double helix is the ability to separate the two strands without disrupting covalent bonds. This makes it possible for the strands to separate and reform under physiological conditions. The specificity of the process is determined by complementary base pairing. Chemical structure and negative charging allow the interactions of DNA with other molecules. The main proprieties of DNA are the next.

Replication – synthesis of new molecules, which are identical with initial copy. That is possible due to the double strand molecule.

Repair – represents a range of cellular responses associated with restoration of primary structure of DNA, based on complementarity of bases.

Denaturation - the conversion of the molecule from double-strand to the single-strand state.

Renaturation – reassociation of denatured complementary single strand of a DNA double helix.

Coiling, supercoiling and *unwinding* - proprieties of double helix to changes its functional states.

Chargaff' rule – the amount of purines is equal with quantity of pyrimidines([G] = [C] and [A] = [T]).

Heterogeneity of DNA – in the same molecule.

Flexibility – changing of type of DNA (ex. From B-DNA – to A-DNA during transcription).

Functions of DNA

DNA is the primary genetic molecule of life and carries within its sequence the hereditary information that determines the structure of proteins in all eukaryotic and prokaryotic organisms. It contains the instructions by which cells grow, divide, and differentiate, and has provided a basis for the evolutionary process both within and between related species. DNA is the information-carrying material that comprises the genes, or units of inheritance, which are arranged in linear arrays along the chromosomes of the cell.

Heterogeneity of DNA in eukaryotes

The total amount of DNA in the genome is characteristic for each living species known as its ***C-value***. There is enormous variation in the range of C-values (10^6-10^{11}). There is a discrepancy between genome size and genetic complexity, called ***C-value paradox***:

- There is an excess of DNA compared with the amount that could be expected to code for proteins.
- There are large variations in C-values between certain species whose apparent complexity does not vary much.

Eukaryotic DNA may be divided in some groups depending on their complexity:

- Nonrepetitive DNA – 45-56%;
- Moderately repetitive DNA – 8-30%;
- Highly repetitive DNA – 12-25%.

Nonrepetitive DNA consists of sequences existing in genome in one copy or a few copies and usually encodes information about proteins.

Moderately repetitive DNA consists of families of sequences that are not exactly the same, but are related. In human genome such sequences are represented by genes encoding for histones, rRNA, actins. Another sequences are ***Alu family*** – short fragments (~300 bp) distributed among the nonrepetitive sequences.

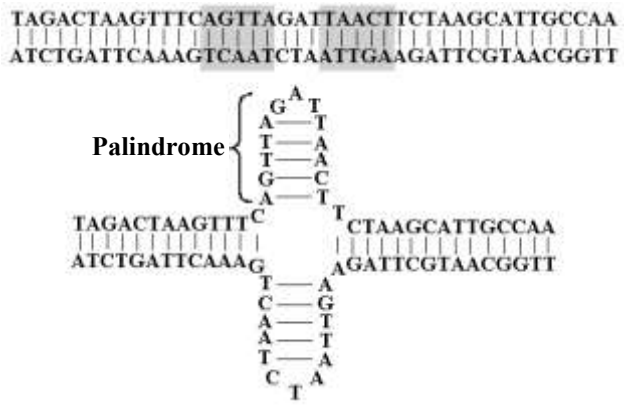

Fig. 16. The structure of palindrome and cruciform

Highly repetitive DNA represents very short sequences repeated many times in tandem in large clusters. They are also called ***simple sequences*** or ***satellite DNA***. The simple sequences are located mainly in the heterocromatine, especially around the centromer. The tandem clusters of satellite DNA are highly polymorphic, with wide variations between individuals. Such sequences, called ***minisatellites***, can be used to characterize individual genomes in the technique of "DNA fingerprinting".

Additional classes of highly repeated DNA are ***palindromes***, which consist of ***inverted repeats*** – a region of dyad symmetry (fig. 16). In a double-strand DNA, the complementary sequences on one strand have the opportunity to base pair only if the strand separates from its partner. As results a ***hairpin*** could be formed. The formation of two apposed hairpins creates a ***cruciform***. Palindromes are very important for interactions of DNA with proteins, they serve as indicators in the processes of initiation and termination of transcription, replication etc.

Ribonucleic acids

Molecules of RNA represent single-strand polymers and consist of nucleotides linked by phosphodiester bonds 3'-5' (some viral RNA may be double-stranded). The ***primary structure of RNA*** represents a sequence of nucleotides covalently linked through 3,5-phosphodiester bonds into an unbranched single-stranded chain. The four major bases in RNA are the purines adenine and guanine, and the pyrimidines cytosine and uracil (A, G, C, and U, respectively). The thymine nucleoside (T) in DNA is transcribed into U in RNA.

Regions of complementarity within a single-stranded RNA sequence can base-pair with one another to generate double-stranded regions of ***secondary structure*** (***hairpin loops***). U base-pairs with A, and G with C, via hydrogen bonding. Further folding can take place to form complex ***tertiary structures***. The cloverleaf structure of the tRNA molecule is possibly the best-defined RNA secondary structure (fig. 17).

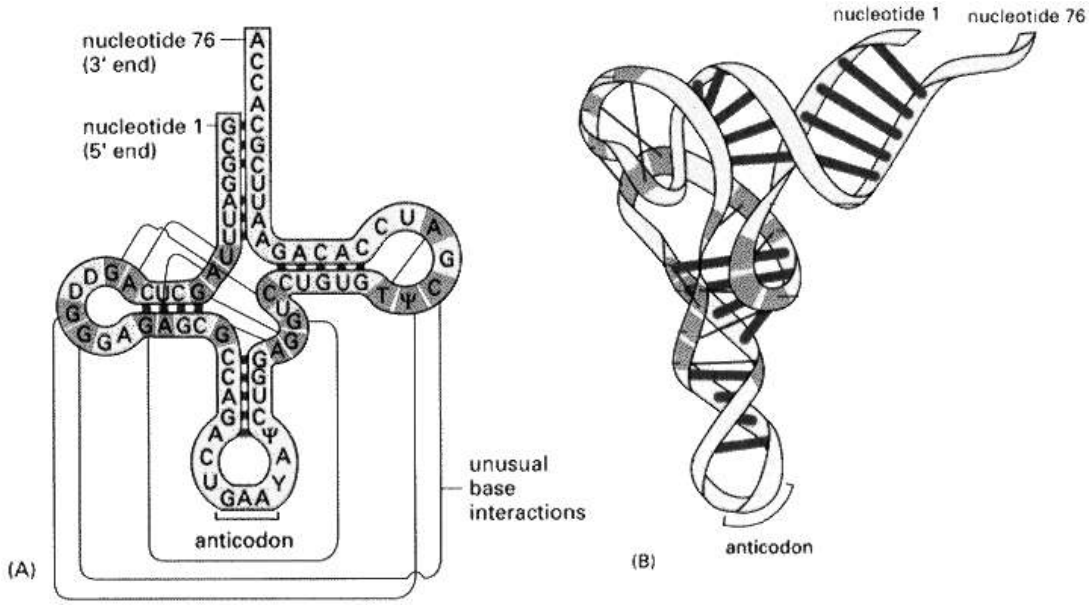

Fig. 17. A - Secondary structure of tRNA (cloverleaf); B – Tertiary structure of tRNA

Both prokaryotic and eukaryotic cells contain ***messenger RNA*** (mRNA), ***transfer RNA*** (tRNA) and ***ribosomal RNA*** (rRNA), all of which are involved in various ways in the accurate conversion of genetic information (in the form of DNA) into protein. This process in eukaryotic cells also involves two additional classes of RNA - ***heterogeneous nuclear*** (or pre-messenger) ***RNA*** (hNRNA or pre-mRNA) and ***small nuclear RNAs*** (snRNA). All cellular RNAs are synthesized by the process of ***transcription***, in which a complementary RNA copy of a DNA template is made.

Messenger RNA is a temporary complementary copy of the ***sense strand*** (anticoding strand) of protein-coding DNA. Messenger RNAs range in length from a few hundred to many thousands of nucleotides depending largely upon the size of the protein they encode. An mRNA contains a region that specifies the protein sequence (the protein-coding region), flanked on either side by untranslated regions (5' and 3' untranslated regions). The coding region represents a sequence of ***codons***, each of three nucleotides. The sequence of codons specifies the amino-acid sequence of the polypeptide chain, with each codon representing a particular amino acid. The mature mRNA molecule is transported to the cytoplasm where it is translated into protein on the ribosome.

Table 3. Types of rRNAs

Type of cell	Sedimentation coefficient	Length (bases)
Prokaryotes	5S	120
	16S	1540
	23S	2900
Eukaryotes	5S	120
	5,8S	160
	18S	1900
	28S	4700

Ribosomal RNA is a major structural component of the ribosome. It participates in the process of synthesis of proteins. There are some classes of rRNAs in prokaryotic and eukaryotic cells (Table 3). (Sedimentation coefficient characterizes the speed of sedimentation of particle during centrifugation. As the sedimentation coefficient is dependent on the mass of the particle it can be used to estimate molecular mass. Larger molecules have higher values of sedimentation coefficient. Sedimentation coefficients are generally expressed as Svedberg units [S]).

Transfer RNAs are small RNAs (70-80 nucleotides), which have a key role in translation (fig. 17). They act as "adaptor" molecules, each tRNA molecule having a 3-base ***anticodon*** complementary to one of the codons, and also having a site at which the appropriate amino acid is attached. During translation, the anticodon on the charged tRNA pairs with the codon on the mRNA within the ribosome, and the amino acid is then transferred from the tRNA to the growing polypeptide chain. Accuracy in the mRNA codont:tRNA anticodon interaction, and in the recognition of amino acids by tRNAs is central to the insertion of the appropriate amino acid into the polypeptide chain during translation. There are at least 20 types of tRNAs in the cell (according to the number of amino acids participating in translation). The maximal number of types of tRNAs is 61.

Small nuclear RNAs are present only in eukaryotes. They are parts of some nuclear enzymes and participate in the post-transcriptional modification of RNA.

Heterogeneous nuclear RNA represents the primary RNA transcripts, which are found in the nucleus. It is, as the name implies, a heterogeneous collection of RNA molecules which are on average some four to five times larger than mature mRNAs and in some cases more unstable.