SBT1102 – BIOCHEMISTRY

UNIT 1 CARBOHYDRATES

Introduction. Classification, Properties and Biological importance. Isomers, epimers, enantiomers,mutarotation, open chain and closed chain structures of glucose.

UNIT 2 AMINOACIDS AND PROTEINS

Aminoacids: classification- essential and non-essential amino acids, protein and non-protein amino acids, Zwitter ions. Proteins: Classification- based on i) shape and solubility and ii) increasing complexity of structure. Structure of proteins: primary, secondary, tertiary and quaternary, biological significance. Concept of isoelectric point and its significance.

UNIT 3 LIPIDS

Introduction, Classification, Properties and Biological importance. Fatty acid nomenclature and structure, Lipids in cell membrane Cholesterol and Steroids, Hormones - structure and function

UNIT 4 NUCLEIC ACIDS

Introduction- Nitrogeneous bases - Purines and Pyrimidines - Nucleosides and Nucleotides -- Structure of nucleic acids - DNA, RNA: m-RNA, t-RNA, r-RNA - Biological importance of nucleic acids. 16s rRNA and its significance.

UNIT 5 VITAMINS AND MINERALS

Vitamins: fat soluble and water soluble vitamins. Minerals: Micro and Macro minerals. Biological importance of vitamin and minerals, deficiency symptoms

Nucleic Acids – Introduction

The first isolation of what we now refer to as **DNA** was accomplished by Johann Friedrich Miescher 1870. He reported finding a weakly acidic substance of unknown function in the nuclei of human white blood cells, and named this material "nuclein". A few years later, Miescher separated nuclein into protein and nucleic acid components. In the 1920's nucleic acids were found to be major components of chromosomes, small gene-carrying bodies in the nuclei of complex cells.

Elemental analysis of nucleic acids showed the presence of phosphorus, in addition to the usual C, H, N & O. Unlike proteins, nucleic acids contained no sulfur. Complete hydrolysis of chromosomal nucleic acids gave inorganic phosphate, 2-deoxyribose (a previously unknown sugar) and four different heterocyclic bases (shown in the following diagram). To reflect the unusual sugar component, chromosomal nucleic acids are called deoxyribonucleic acids, abbreviated DNA. Analogous nucleic acids in which the sugar component is ribose are termed ribonucleic acids, abbreviated RNA. The acidic character of the nucleic acids was attributed to the phosphoric acid moiety.

Their functions include:

1. Serving as energy stores for future use in phosphate transfer reactions. These reactions are predominantly carried out by ATP.

2. Forming a portion of several important coenzymes such as NAD^+, $NADP^+$, FAD and coenzyme A.

3. Serving as mediators of numerous important cellular processes such as second messengers in signal transduction events. The predominant second messenger is cyclic-AMP (cAMP), a cyclic derivative of AMP formed from ATP.

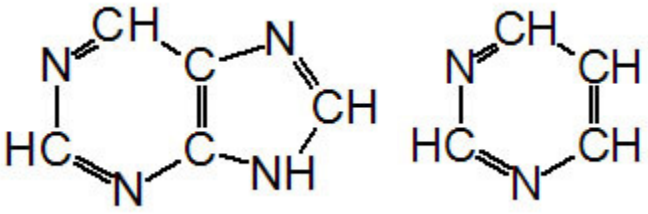
4. Serving as neurotransmitters and as signal receptor ligands. Adenosine can function as an inhibitory neurotransmitter, while ATP also affects synaptic neurotransmission throughout the central and peripheral nervous systems. ADP is an important activator of platelet functions resulting in control of blood coagulation.

5. Controlling numerous enzymatic reactions through allosteric effects on enzyme activity.

6. Serving as activated intermediates in numerous biosynthetic reactions. These activated intermediates include S-adenosylmethionine (S-AdoMet or SAM) involved in methyl transfer reactions as well as the many sugar coupled nucleotides involved in glycogen and glycoprotein synthesis.

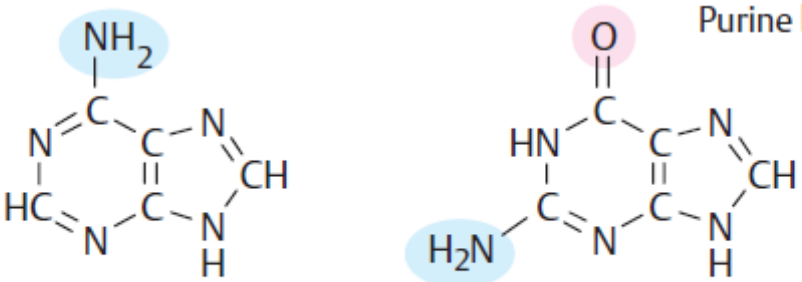
Nucleoside and Nucleotide Structure and Nomenclature

The nucleotides found in cells are derivatives of the heterocyclic highly basic, compounds, purine and pyrimidine.

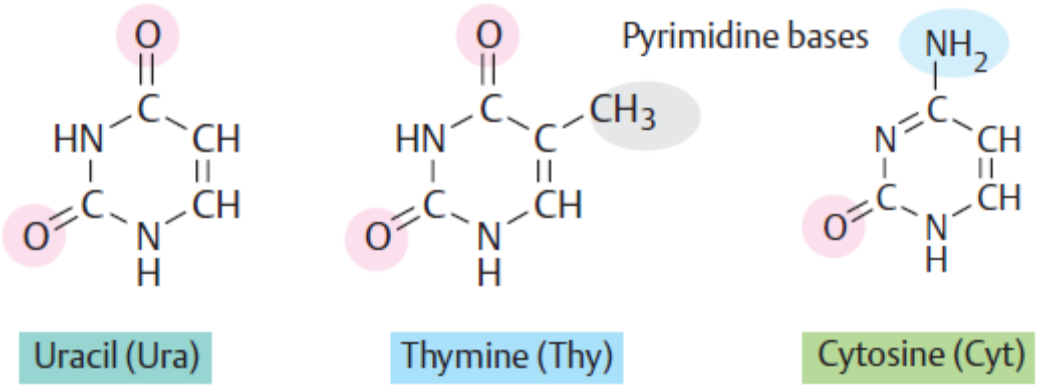


Purine **Pyrimidine**

Five of these bases are the main components of nucleic acids in all living creatures. The purine bases **adenine** and **guanine** and the pyrimidine base **cytosine** are present in both RNA *and* DNA. In contrast, **uracil** is only found in RNA. In DNA, **uracil** is replaced by thymine, the 5-methyl derivative of uracil.



Adenine (Ade) Guanine (Gua)



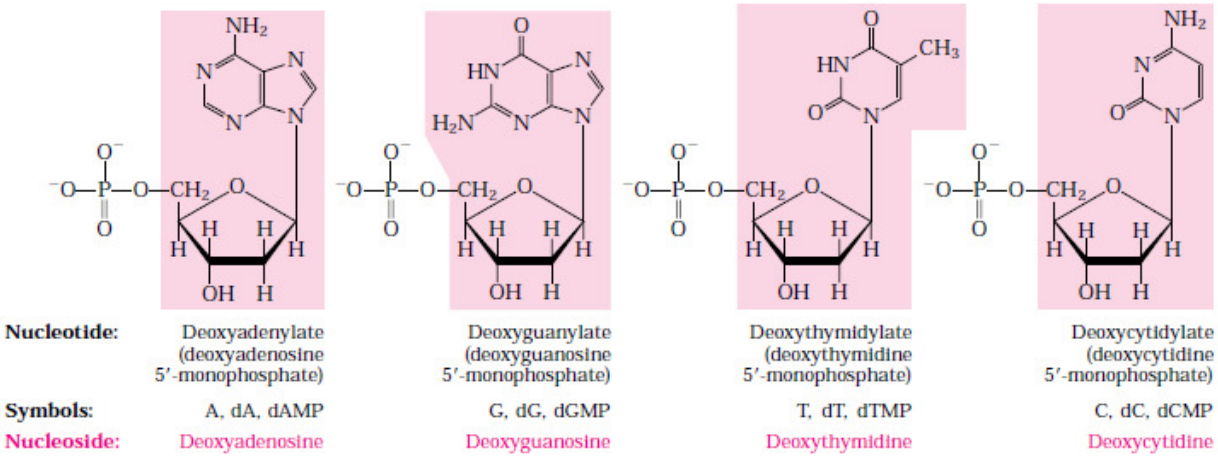
When a nucleic acid base is N-glycosidically linked to ribose or 2-deoxyribose, it yields a **nucleoside**. The nucleoside **adenosine** (abbreviation: A) is formed in this way from adenine and ribose, for example. The corresponding derivatives of the other bases are called *guanosine* (G), *uridine* (U), *thymidine* (T) and *cytidine* (C). When the sugar component is 2-deoxyribose, the product is a **deoxyribonucleoside.**

In the cell, the 5' OH group of the sugar component of the nucleoside is usually esterified with phosphoric acid. If the 5' phosphate residue is linked via an acid–anhydride bond to additional phosphate residues, it yields nucleoside diphosphates and triphosphates—e. g., ADP and ATP, which are important coenzymes in energy metabolism. All of these nucleoside phosphates are classified as **nucleotides**. In nucleosides and nucleotides, the pentose residues are present in the furanose form. The sugars and bases are linked by an *N*-**glycosidic bond** between the C-1 of the sugar and either the N-9 of the purine ring or N-1 of the pyrimidine ring. This bond always adopts the β-configuration.

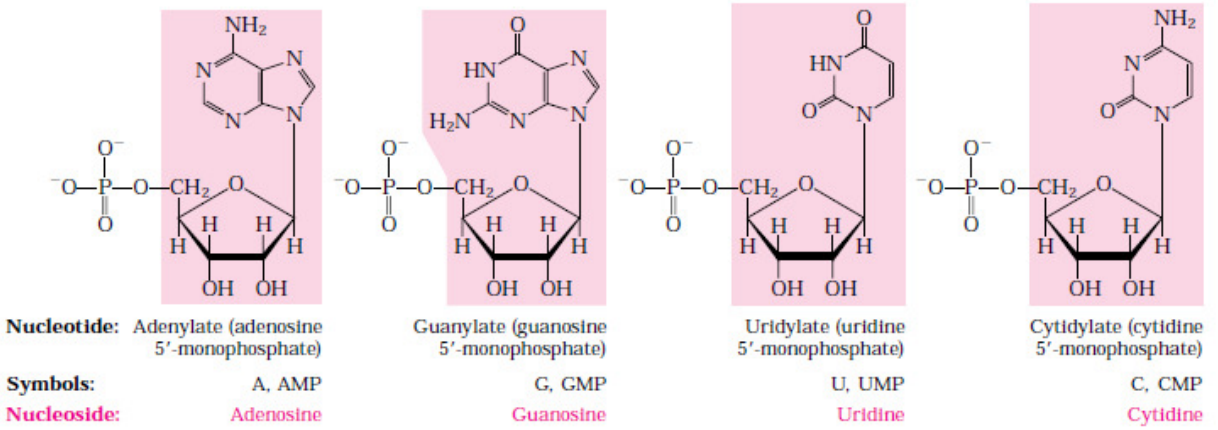
In the pentoses of nucleotidesand nucleosides the carbon numbers are given a prime (') designation to distinguish them from the numbered atoms of the nitrogenous bases. The base of a nucleotide is joined covalently (at N-1 of pyrimidines and N-9 of purines) in an *N*-*β*-glycosyl bond to the 1_ carbon of the pentose, and the phosphate is esterified

to the 5_ carbon. The *N*-*β*-glycosyl bond is formed by removal of the elements of water (a hydroxyl group from the pentose and hydrogen from the base), as in *O*-glycosidic bond formation.

Both DNA and RNA contain two major purine bases, **adenine** (A) and **guanine** (G), and two major pyrimidines. In both DNA and RNA one of the pyrimidines is **cytosine** (C), but the second major pyrimidine is not the same in both: it is **thymine** (T) in DNA and **uracil** (U) in RNA. Only rarely does thymine occur in RNA or uracil in DNA.



(a) Deoxyribonucleotides



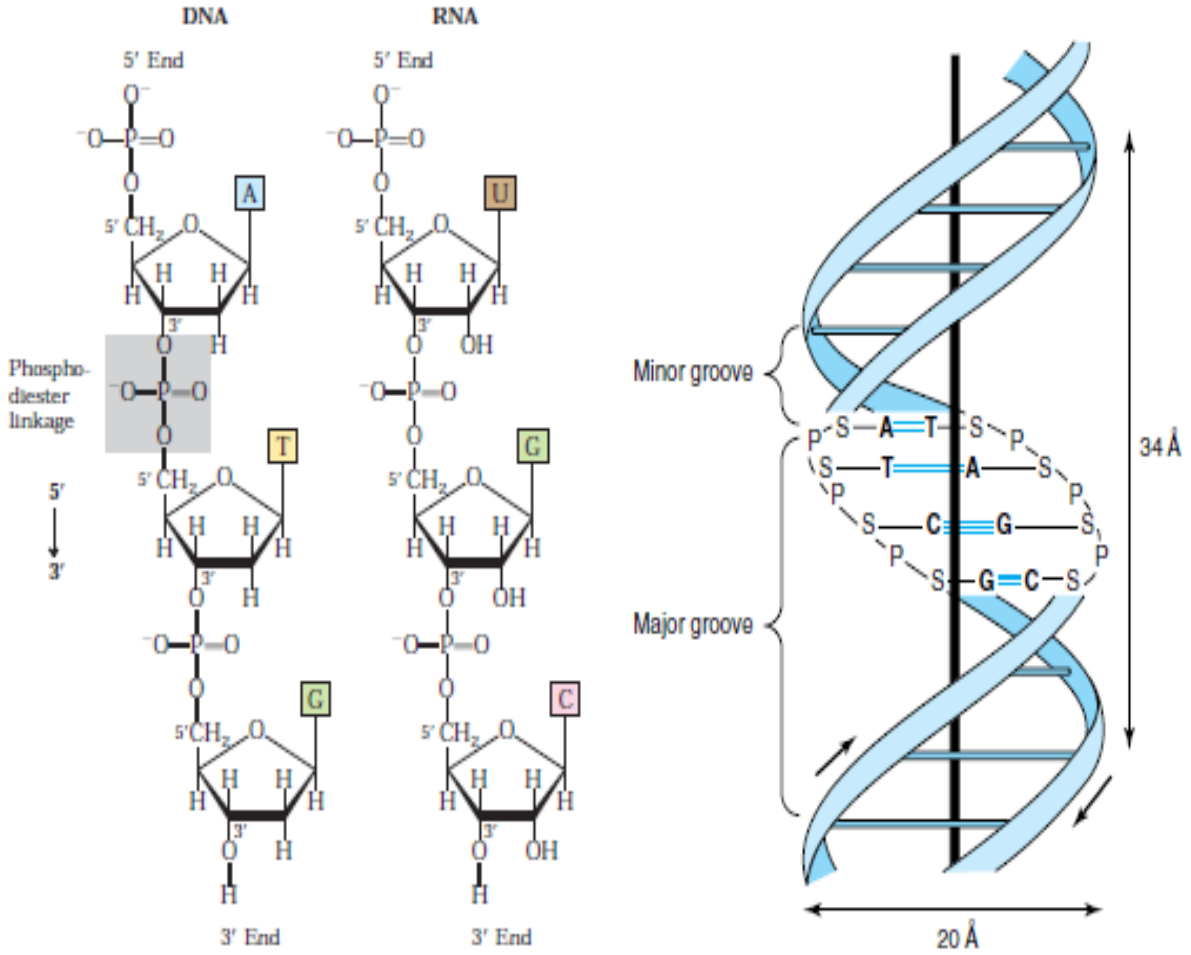
(b) Ribonucleotides

Nucleotide and Nucleic Acid Nomenclature

Base	*Nucleoside*	*Nucleotide*	*Nucleic acid*
Purines			
Adenine	Adenosine	Adenylate	RNA
	Deoxyadenosine	Deoxyadenylate	DNA
Guanine	Guanosine	Guanylate	RNA
	Deoxyguanosine	Deoxyguanylate	DNA
Pyrimidines			
Cytosine	Cytidine	Cytidylate	RNA
	Deoxycytidine	Deoxycytidylate	DNA
Thymine	Thymidine or deoxythymidine	Thymidylate or deoxythymidylate	DNA
Uracil	Uridine	Uridylate	RNA

The successive nucleotides of both DNA and RNA are covalently linked through phosphate-group “bridges,” in which the 5_-phosphate group of one nucleotide unit is joined to the 3_-hydroxyl group of the next nucleotide, creating a **phosphodiester linkage.** Thus the covalent backbones of nucleic acids consist of alternating phosphate and pentose residues, and the nitrogenous bases may be regarded as side groups joined to the backbone at regular intervals. The backbones of both DNA and RNA are hydrophilic.

By convention, the structure of a single strand of nucleic acid is always written with the 5’ end at the left and the 3’ end at the right—that is, in the 5’ n 3’ direction. Some simpler representations of this pentadeoxyribonucleotide are pA-C-G-T-AOH, pApCpGpTpA, and pACGTA. A short nucleic acid is referred to as an **oligonucleotide.** The definition of “short” is somewhat arbitrary, but polymers containing 50 or fewer nucleotides are generally called oligonucleotides. A longer nucleic acid is called a **polynucleotide.**

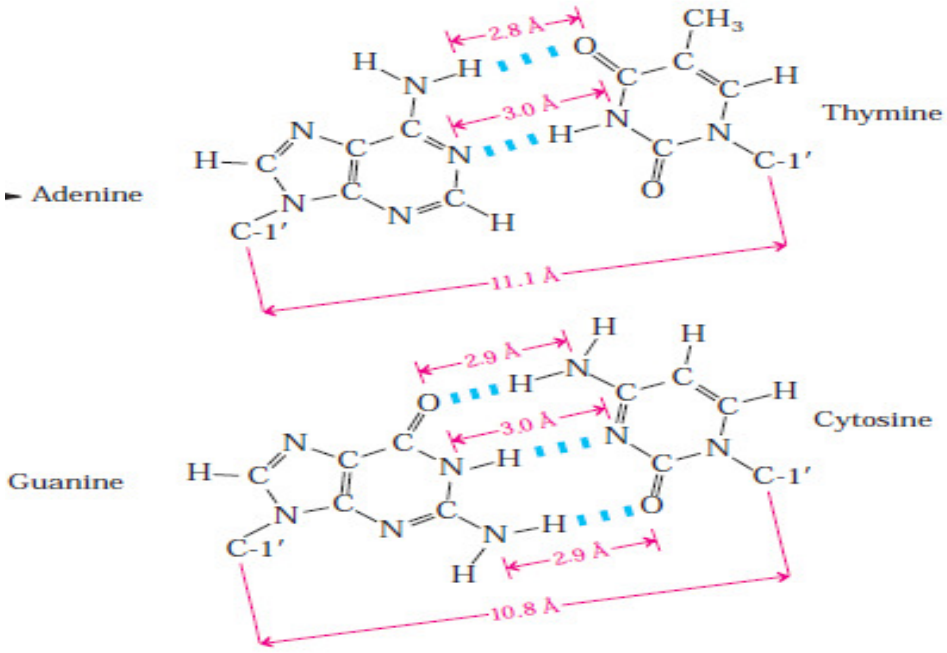


The existence of specific base-pairing interactions was discovered in the course of studies directed at determining the three-dimensional structure of DNA. Maurice Wilkins and Rosalind Franklin obtained x-ray diffraction photographs of fibers of DNA. The characteristics of these diffraction patterns indicated that DNA was formed of two chains that wound in a regular helical structure. From these and other data, James Watson and Francis Crick inferred a structural model for DNA that accounted for the diffraction pattern and was also the source of some remarkable insights into the functional properties of nucleic acids.

The features of the Watson-Crick model of DNA deduced from the diffraction patterns are:

1. Two helical polynucleotide chains are coiled around a common axis. The chains run in opposite directions.
2. The sugar-phosphate backbones are on the outside and, therefore, the purine and pyrimidine bases lie on the inside of the helix.
3. The bases are nearly perpendicular to the helix axis, and adjacent bases are separated by 3.4 Å. The helical structure repeats every 34 Å, so there are 10 bases (= 34 Å per repeat/3.4 Å per base) per turn of helix. There is a rotation of 36 degrees per base (360 degrees per full turn/10 bases per turn).
4. The diameter of the helix is 20 Å.

Watson and Crick discovered that guanine can be paired with cytosine and adenine with thymine to form base pairs that have essentially the same shape. These base pairs are held together by specific hydrogen bonds. This base-pairing scheme was supported by earlier studies of the base composition of DNA from different species. In 1950, Erwin Chargaff reported that the ratios of adenine to thymine and of guanine to cytosine were nearly the same in all species studied.



The meaning of these equivalences was not evident until the Watson-Crick model was proposed, when it became clear that they represent an essential facet of DNA structure. The spacing of approximately 3.4 Å between nearly parallel base pairs is readily apparent in the DNA diffraction pattern. The stacking of bases one on top of another contributes to the stability of the double helix.

DNA: structure ‘

Deoxyribonucleic acids (DNAs) are polymeric molecules consisting of nucleotide building blocks. Instead of ribose, however, DNA contains 2’-deoxyribose, and the *uracil* base in RNA is replaced by *thymine.* The spatial structure of the two molecules also differs. The first evidence of the special structure of DNA was the observation that the amounts of adenine and thymine are almost equal in every type of DNA. The same applies to guanine and cytosine. The model of DNA structure formulated in 1953 explains these *constant base ratios:* intact DNA consists of *two* polydeoxynucleotide molecules (“strands”).

Each base in one strand is linked to a *complementary* base in the other strand by H-bonds. Adenine is complementary to thymine, and guanine is complementary to cytosine. One purine base and one pyrimidine base are thus involved in each **base pair**. The complementarity of A with T and of G with C can be understood by considering the H bonds that are possible between the different bases. Potential donors are amino groups (Ade, Cyt, Gua) and ring NH groups. Possible acceptors are carbonyl oxygen atoms (Thy, Cyt, Gua) and ring nitrogen atoms. *Two* linear and therefore highly stable bonds can thus be formed in A–T pairs, and *three* in G–C pairs. Base pairings of this type are only possible, however, when the *polarity* of the two strands differs—i. e., when they run in opposite directions.

In addition, the two strands have to be intertwined to form a **double helix**. Due to steric hindrance by the 2’-OH groups of the ribose residues, RNA is unable to form a double helix. The structure of RNA is therefore less regular than that of DNA. The conformation of DNA that predominates within the cell is known as **B-DNA**. Along the

whole length of the DNA molecule, there are two depressions—referred to as the “minor groove” and the “major groove”—that lie between the strands.

DNA: conformation ‘

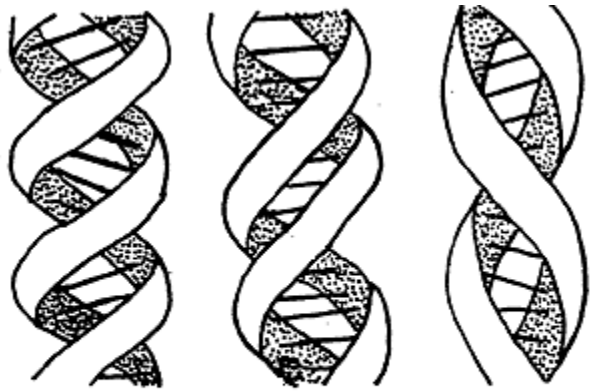
Investigations of synthetic DNA molecules have shown that DNA can adopt several different conformations. All of the DNA segments shown consist of 21 base pairs (bp) and have the same sequence. By far the most common form is **B-DNA**. This consists of two antiparallel polydeoxynucleotide strands intertwined with one another to form a **righthanded double helix**. The “backbone” of these strands is formed by deoxyribose and phosphate residues linked by phosphoric acid diester bonds.

In the B conformation, the aromatic rings of the nucleobases are stacked at a distance of 0.34 nm almost at right angles to the axis of the helix. Each base is rotated relative to the preceding one by an angle of 35°. A complete turn of the double helix (360°) therefore contains around 10 base pairs (abbreviation: bp), i. e., the *pitch* of the helix is 3.4 nm. Between the backbones of the two individual strands there are two grooves with different widths. The *major groove* is visible at the top and bottom, while the narrower *minor groove* is seen in the middle. DNA-binding proteins and transcription factors usually enter into interactions in the area of the major groove, with its more easily accessible bases.

In certain conditions, DNA can adopt the **A conformation**. In this arrangement, the double helix is still right-handed, but the bases are no longer arranged at right angles to the axis of the helix, as in the B form. As can be seen, the A conformation is more compact than the other two conformations. The minor groove almost completely disappears, and the major groove is narrower than in the B form. A-DNA arises when B-DNA is dehydrated. It probably does not occur in the cell.

In the **Z-conformation**, which can occur within GC-rich regions of B-DNA, the organization of the nucleotides is completely different. In this case, the helix is *left-handed,* and the backbone adopts a characteristic *zig-zag* conformation (hence “Z-

DNA"). The Z double helix has a smaller pitch than B-DNA. DNA segments in the Z conformation probably have physiological significance, but details are not yet known.



Form	A	B	Z
helical twist	right	right	left
base pairs per turn	11	10	12
occurrence	RNA, DNA	DNA	DNA

Feature	B-DNA	A-DNA	Z-DNA
Type of helix	Right-handed	Right-handed	Left-handed
Helical diameter (nm)	2.37	2.55	1.84
Rise per base pair (nm)	0.34	0.29	0.37
Distance per complete turn (pitch) (nm)	3.4	3.2	4.5
Number of base pairs per complete turn	10	11	12
Topology of major groove	Wide, deep	Narrow, deep	Flat
Topology of minor groove	Narrow, shallow	Broad, shallow	Narrow, deep

RNA

RNA differs from DNA in both structural and functional respects. RNA has two major structural differences: each of the ribose rings contains a 2´-hydroxyl, and RNA uses uracil in place of thymine. RNA molecules are capable of base pairing, but generally will not form large regions of stable RNA-RNA double helix. RNA can act as a genetic material (although this role, at least for current organisms, seems to be restricted to viruses). Unlike DNA, RNA can form complex three-dimensional structures. As a result, RNA can also exhibit catalytic activity. The combination of the ability to store genetic information with the ability to catalyze reactions has resulted in a proposal for the origin

of life: the “RNA World”. The RNA world hypothesis proposes that RNA molecules once filled all of the roles of protein and nucleic acid macromolecules, and acted in both an information storage capacity and as the source of the enzymatic activity required for metabolic reactions. In general, RNA is less suited to acting as genetic material than DNA, and is less suited to forming efficient catalysts than proteins. Assuming that the RNA world once existed, nearly all of its functions have been taken over by other biological molecules. However, some vestiges of the RNA world may still exist. The vast majority of RNA functions are concerned with protein synthesis.

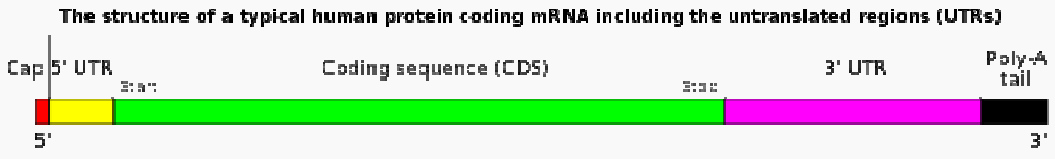
Characteristics

- RNA does not self replicate in order to multiply; instead it is encoded by DNA genes
- RNA is synthesized in order for the translation of DNA to be possible
- The DNA-RNA function is highly interdependable, i.e., if there is problem with DNA, there will be a problem with the RNA functions and vice versa (no RNA = no DNA translation can occur, thus DNA is useless without its RNA genes)

RNA genes of DNA encode for 3 major types of RNA:

- ribosomal RNA
- messenger RNA
- transfer RNA

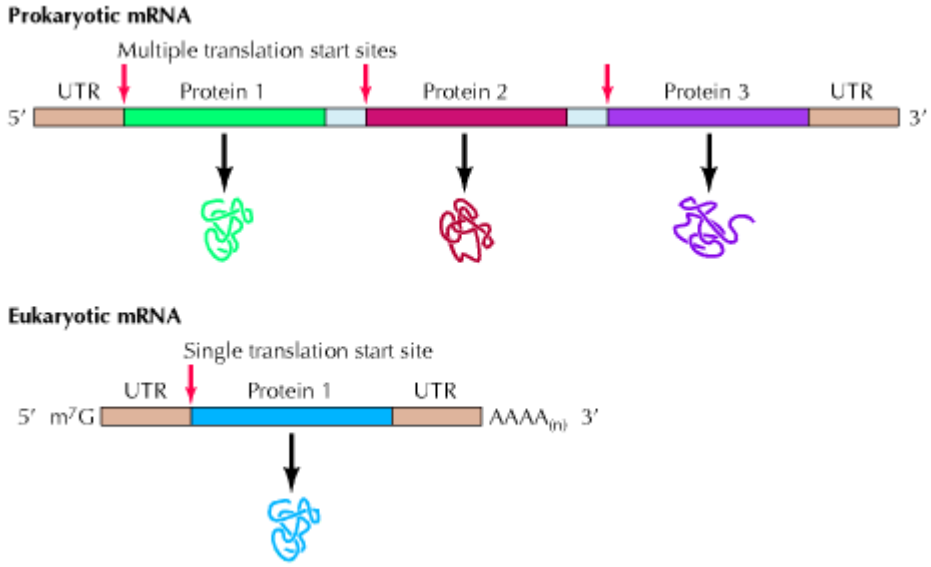
mRNA – messenger RNA



The structure of a mature eukaryotic mRNA. A fully processed mRNA includes a 5' cap, 5' UTR,coding region, 3' UTR, and poly(A) tail.

mRNA genes are the genes that encode only for proteins but this encoding has an RNA intermediate. The DNA is firstly transcribed into mRNA and subsequently translated into a protein product. So the mRNA genes are the genes that encode for mRNA in order to synthesize proteins. mRNA constitutes only the 5% of the total RNA.

The DNA gives rise to nascent RNA in the nucleus. Addition of poly A tail to this nascent RNA makes this a pre-mRNA. The pre-mRNA has both introns and exons in it. Splicing removes the introns bringing the exons together to form the CDS. This is called the mature mRNA.



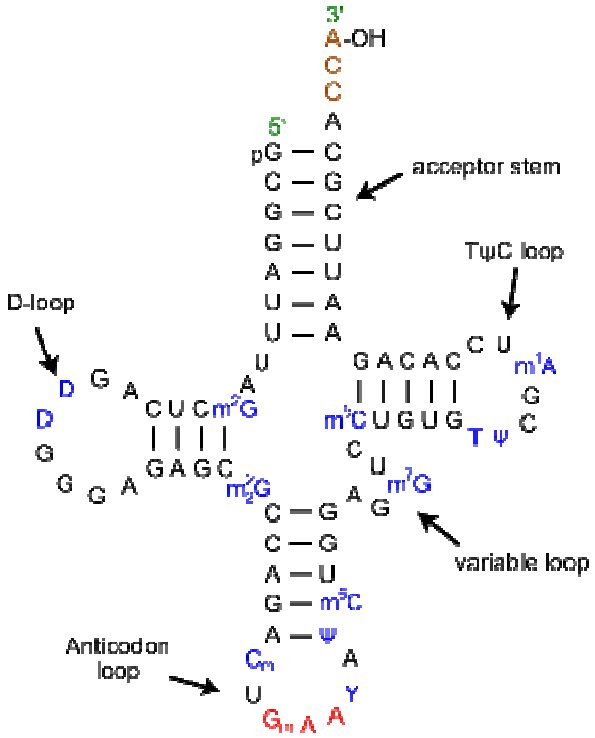
The prokaryotic mRNA is a polycistronic mRNA compared to the eukaryotic mRNA which is monocistronic mRNA. The prokaryotic mRNA has a leader sequence and a trailer sequence.

tRNA – transfer RNA

Transfer RNA is encoded by genes that also encode for the 5S size rRNA. RNA polymerase III is responsible for the transcription of these genes by binding on the promoter, situated about 100 base pairs downstream the Transcription Start Site -TSS, along with the Transcription Factors giving rise to the Transcription Initiation Complex. As soon as this complex is formed transcription process can begin and when the Transcription Complex faces an Adenine rich region transcription comes to an end as

this area is an indication for the gene end. tRNA constitutes 15% of the total RNA and is directly involved in the translation of the mRNA. More specificaly tRNA binds onto a specific amino acid and brings it along the translation site so that it is bound on the newly synthesized peptide.

- tRNA binds to its specific amino acid recognized by its side R chain in presence of the aminoacyl tRNA synthetase enzyme. The synthetase binds the 5'-CCA-OH-3' acceptor arm with the —COOH group of the amino acid.
- When the small ribosomal subunit faces an AUG codon on the mRNA it indicates the commencing of the peptide formation. As soon as the AUG codon is recognized then the first tRNA binds on the small ribosomal subunit and on the mRNA through its anticodon arm, giving rise to the Translation Initiation Complex designated as $tRNA_i^{met}$. Eventually the large ribosomal subunit binds on the complex indicating the initiation of the translation process. Translation always begins with the methionine amino acid on the newly synthesized peptide.



Analysis of the tRNA sequence suggests a cloverleaf secondary structure formed by regions of base pairing between the sections of the RNA strand, with this cloverleaf folding into the three-dimensional structure.

rRNA – ribosomal RNA

In bacteria (prokaryotes) there are three different ribosomal RNAs called 5S, 16S, and 23S. Eukaryotes have homologous RNAs called 5S, 28S, and 18S ribosomal RNAs. In addition, they have a 5.8S RNA that is homologous to one end of the prokaryotic 23S RNA. Both prokaryotic and eukaryotic ribosomes can be broken down into two subunits (the S in 16S represents Svedberg units), nt= length in nucleotides of the respective rRNAs.

Note that the S units of the subunits (or the rRNAs) cannot simply be added because they represent measures of sedimentation rate rather than of mass. The sedimentation rate of each subunit is affected by its shape, as well as by its mass. The nt units can be added as these represent the integer number of units in the linear rRNA polymers (for example, the total length of the human rRNA = 7216 nt).

Prokaryotes

In prokaryotes a small 30S ribosomal subunit contains the 16S ribosomal RNA. The large 50S ribosomal subunit contains two rRNA species (the 5S and 23S ribosomal RNAs). Bacterial 16S ribosomal RNA, 23S ribosomal RNA, and 5S rRNA genes are typically organized as a co-transcribed operon. There may be one or more copies of the operon dispersed in the genome (for example, *Escherichia coli* has seven).

Archaea contains either a single rDNA operon or multiple copies of the operon. The 3' end of the 16S ribosomal RNA (in a ribosome) binds to a sequence on the 5' end of mRNA called the Shine-Dalgarno sequence.

Eukaryotes

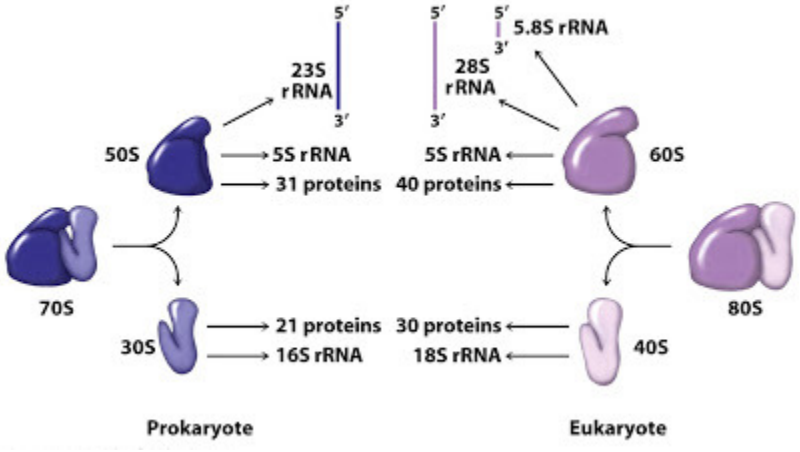
In contrast, eukaryotes generally have many copies of the rRNA genes organized in tandem repeats; in humans approximately 300–400 repeats are present in five

clusters (on chromosomes13, 14, 15, 21 and 22). Because of their special structure and transcription behaviour, rRNA gene clusters are commonly called "ribosomal DNA" (note that the term seems to imply that ribosomes contain DNA, which is not the case).

The 18S rRNA in most eukaryotes is in the small ribosomal subunit, and the large subunit contains three rRNA species (the 5S, 5.8S and 28S in mammals, 25S in plants, rRNAs).

Mammalian cells have 2 mitochondrial (12S and 16S) rRNA molecules and 4 types of cytoplasmic rRNA (the 28S, 5.8S, 18S, and 5S subunits). The 28S, 5.8S, and 18S rRNAs are encoded by a single transcription unit (45S) separated by 2 internally transcribed spacers. The 45S rDNA is organized into 5 clusters (each has 30-40 repeats) on chromosomes 13, 14, 15, 21, and 22. These are transcribed by RNA polymerase I. 5S occurs in tandem arrays (~200-300 true 5S genes and many dispersed pseudogenes), the largest one on the chromosome 1q41-42. 5S rRNA is transcribed by RNA polymerase III.

The tertiary structure of the small subunit ribosomal RNA (SSU rRNA) has been resolved by X-ray crystallography. The secondary structure of SSU rRNA contains 4 distinct domains — the 5', central, 3' major and 3' minor domains.



16S rRNA – Significance

The 16S rRNA gene is a section of prokaryotic DNA found in all bacteria and archaea. This gene codes for an rRNA, and this rRNA in turn makes up part of the ribosome. The first 'r' in rRNA stands for ribosomal. The ribosome is composed of two subunits, the large subunit (LSU) and the small subunit (SSU). These two subunits sandwich the mRNA as it feeds through the ribosome for translation. While there are also associated proteins helping to make up the functional units of the ribosome, in general, in bacteria, the SSU is coded for by the the 16S rRNA gene, and the LSU is coded for by the 23S rRNA & 5S rRNA genes.

The 16S rRNA gene is a commonly used tool for identifying bacteria for several reasons. First, traditional characterization depended upon phenotypic traits like gram positive or gram negative, bacillus or coccus, etc. Taxonomists today consider analysis of an organism's DNA more reliable than classification based solely on phenotypes. Secondly, researchers may, for a number of reasons, want to identify or classify only the bacteria within a given environmental or medical sample. While there is a homologous gene in eukaryotes, the 18S rRNA gene, it is distinct, thereby rendering the 16S rRNA gene a useful tool for extracting and identifying bacteria as separate from plant, animal, fungal, and protist DNA within the same sample. Thirdly, the 16S rRNA gene is relatively short at 1.5 kb, making it faster and cheaper to sequence than many other unique bacterial genes.

Ribosomes (and correspondingly the DNA that codes for them) have been mostly conserved over time, meaning that their structure has changed very little over time due to their important function, translating mRNA into proteins. But even within this gene there are parts that have been conserved more than others. This is due to the structure of the ribosome itself. With the way the ribosome folds, creating bonds with itself in some places (conserved regions) while other portions are looped and unbonded (hypervariable regions), the degree to which any portion of the gene is subject to mutations varies.