Nucleic acids are long chain polymers of nucleotides which are joined together by means of phosphodiester linkages. In phosphodiester bonds, one phosphoric acid molecule forms bonds with the 3′ carbon of one pentose molecule as well as with the 5′ carbon of a second pentose molecule.

Thus each sugar and phosphoric acid forms bonds with two phosphoric acid and pentose molecules; these linkages generate the sugar-phosphate backbone of the nucleic acids. The DNA and RNA bases are attached to the 1′ carbon of the pentose residues.

This assembly of phosphoric acid, pentose and organic base residues is known as polydeoxyribonucleotide in case of DNA and polyribonucleotide in case of RNA. Each nucleotide is composed of three distinct molecules: one molecule each of sugar, phosphoric acid and a nitrogenous base.

Nucleotides and Nucleic Acids:

Both DNA and RNA are known as nucleic acids. They have been given this name for the simple reason that they are made up of structures called nucleotides. Those nucleotides, themselves comprising a number of components, bond together to form the double-helix first discovered by the scientists James Watson and Francis Crick in 1956. This discovery won the two scientists the Nobel Prize. For now, when we discuss nucleic acids you should assume we are discussing DNA rather than RNA, unless otherwise specified.

Structure of Nucleic Acids:

Sugar:

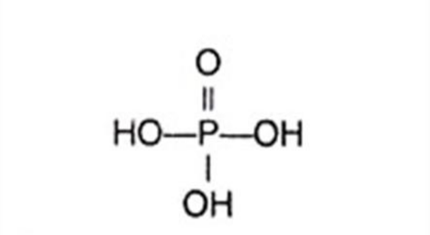
All nucleotides contain a 5-carbon sugar (pentose); the pentose ribose is found in RNA while deoxyribose is found in DNA. In deoxyribose molecules, one oxygen atom (O) is missing from 2′ position (Fig. below). The nucleic acids (NA) are named after the sugar present in them, for example,

Ribose + nucleic acid - Ribose-nucleic acid, commonly written as **“ribonucleic acid”** (RNA).

Deoxyribose + nucleic acid - de-oxy-ribose-nucleic acid, commonly written as **“deoxyribo-nucleic acid”** (DNA).

Phosphoric Acid:

Phosphoric acid (H_3PO_4) (Fig. below) is attached to each sugar at the 3′ and 5′ C positions to give rise to the sugar-phosphate backbone. Free nucleotides in the cell have 3 phosphate residues, generally attached to the 5′ C of the pentose. During the phosphodiester bond formation, two phosphate groups are removed from one of the two participating nucleotides.



$$HO-\overset{\overset{\displaystyle O}{\|}}{\underset{\underset{\displaystyle OH}{|}}{P}}-OH$$

Fig: Structural formula of phosphoric acid

Nitrogenous Bases:

The bases in nucleic acids are heterocyclic compounds containing nitrogen and carbon in their rings. The nitrogenous bases are of two types: pyrimidines and purines.

Pyrimidines:

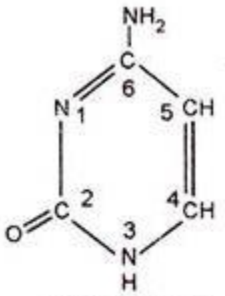
Pyrimidine ring is similar to the benzene ring, except it contains nitrogen in place of carbon at positions 1 and 3 (Fig. 3.3). They also contain a keto oxygen (=0) at the position 2. There are three common pyrimidines cytosine (C), thymine (T) and uracil (U).

Thymine contains two ketooxygens at positions 2 and 6 and a methyl group (-CH3) at position 5. Cytosine contains one keto oxygen at position 2 and an amino group ($-NH_2$) at position 6. These two pyrimidines are found in DNA, while another pyrimidine uracil occurs in RNA in the place of thymine. Uracil differs from thymine only in not having a methyl group at the position 5. Pyrimidines are associated with 1′ C of the sugar by the position 3.

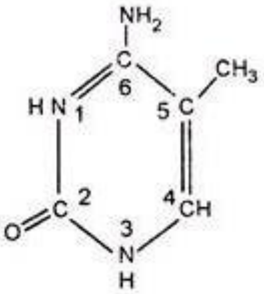
Purines:

Purines have two carbon-nitrogen rings. One of the rings is 6 membered (like pyrimidine), while the other is 5 membered; the two rings share their 4 and 5 C (Fig. 3.3). Both RNA and DNA contain the same two types of purines, viz., adenine (A) and guanine (G).

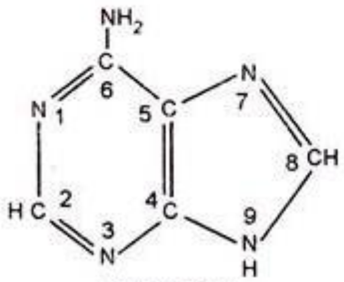
Adenine contains an amino group ($-NH_2$) at position 6, while in guanine this position is occupied by a keto oxygen (=0). In addition, guanine has an amino group at position 2. Both the purines contain nitrogen at positions 1, 3, 7 and 9. Purines associate with 1′ C of pentose sugar at their position 9 N.



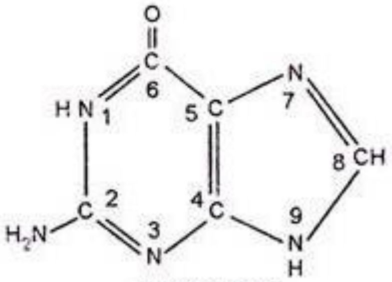
CYTOSINE (C)



5–METHYLCYTOSINE



ADENINE (A)
(6–Amino–purine)



GUANINE (G)
(2–Amino–6–hydroxypurine)

Nucleosides:

The combination of a base and a pentose is termed as nucleoside (Fig. below). The 1′ C of pentose attaches to the 3-position of a pyrimidine or at the 9-position of a purine (Fig. below). Nucleosides derived from ribose are called ribosides, while having de-oxy-ribose is known as de-oxy-riboside; the various nucleosides are as follows:

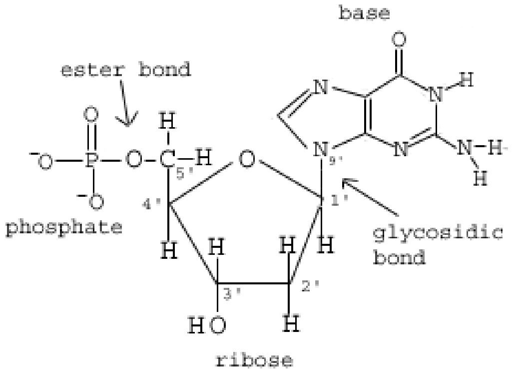
Base	Ribonucleoside	Deoxyribonucleoside
Adenine	Adenosine	Deoxyadenosine
Guanine	Guanosine	Deoxyguanosine
Cytosine	Cytidine	Deoxycitidine
Uracil	Uridine	—
Thymine	—	Thymidine (since only the deoxyriboside of thymine) occurs naturally)

Nucleotides:

A nucleotide consists of three things:

1. A nitrogenous base, which can be adenine, guanine, cytosine, or thymine (in the case of RNA, thymine is replaced by uracil).
2. A five-carbon sugar, called deoxyribose because it is lacking an oxygen group on one of its carbons.
3. One or more phosphate groups.

The nitrogen bases are pyrimidine in structure and form a bond between their 1' nitrogen and the 1' -OH group of the deoxyribose. This type of bond is called a glycosidic bond. The phosphate group forms a bond with the deoxyribose sugar through an ester bond between one of its negatively charged oxygen groups and the 5' -OH of the sugar ().



Polynucleotide Chain:

Nucleotides join together through phosphodiester bonds to yield polynucleotide chain. The phosphodiester bond formation occurs when the 3′ OH of a nucleotide reacts with the phosphoric acid residue attached to the 5′ C of another nucleotide giving rise to (5′ C – O – P – O – C3′) bound.

This liberates a pyrophosphate (P-P) since the nucleotides occur naturally as triphosphates. Several nucleotides become linked in this manner to form a nucleotide chain. Such a chain has a free OH at 5’C (-OH of the phosphate attached to the 5’C) at one end, and a free -OH at 3’C at its other end.

Thus a polynucleotide chain has a polarity of 5′ to 3′; it has a triphosphate group at its 5′-end and a free -OH group at its 3′-end. The growth of the chain occurs in 5′ to 3′ direction, that is, new nucleotides are added only to the free 3′ OH of polynucleotide. This constitutes the primary structure of DNA; it should be noted that it does not impose any restriction on the sequence of bases present in the chain.

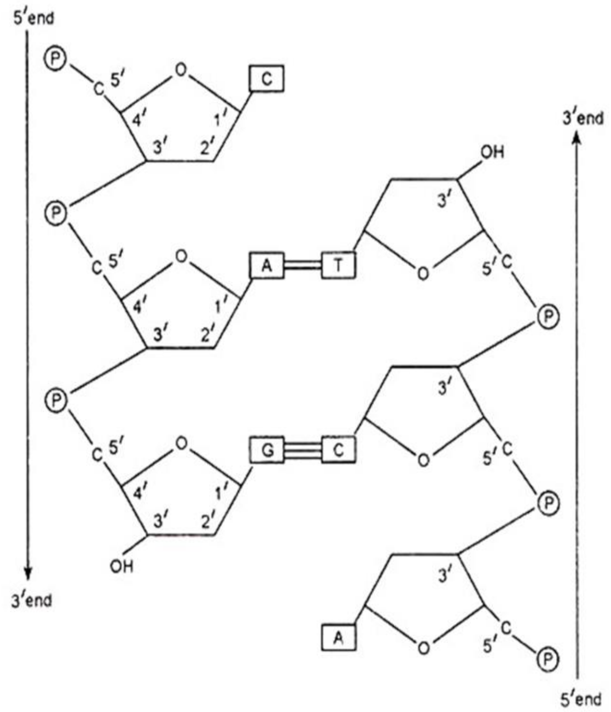
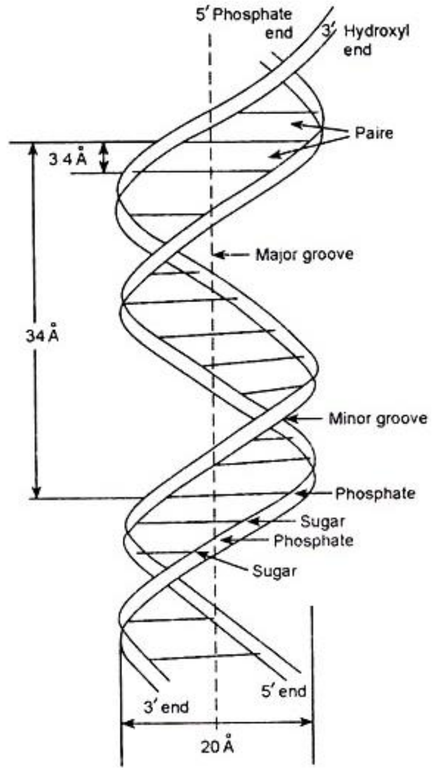


Fig: representation of the DNA double helix showing opposite polarities of the sugar-phosphate linkages in two strands

The Helical Structure of DNA:

The Helical Structure of DNA shows a single strand of DNA. However, as stated earlier, DNA exists as a double-helix, meaning two strands of DNA bind together.



As seen above, one strand is oriented in the 5' to 3' direction while the complementary strand runs in the 3' to 5' direction. Because the two strands are oppositely oriented, they are said to be anti-parallel to each other. The two strands bond through their nitrogen bases (marked A, C, G, or T for adenine, cytosine, and guanine). Note that adenine only bonds with thymine, and cytosine only bonds with guanine. The nitrogen bases are held together by hydrogen bonds: adenine and thymine form two hydrogen bonds; cytosine and guanine form three hydrogen bonds.

An important thing to remember about the structure of the DNA helix is that as a result of anti-parallel pairing, the nitrogen base groups face the inside of the helix while the sugar and phosphate groups face outward. The sugar and phosphate groups in the helix therefore make up the phosphate backbone of DNA. The backbone is highly negatively charged as a result of the phosphate groups.

The Importance of the Hydrogen Bond:

Hydrogen bonding is essential to the three-dimensional structure of DNA. These bonds do *not,* however, contribute largely to the stability of the double helix. Hydrogen bonds are very weak interactions and the orientation of the bases must be just right for the interactions to take place. While the large number of hydrogen bonds present in a double helix of DNA leads to a cumulative effect of stability, it is the interactions gained through the stacking of the base pairs that leads to most of the helical stability.

Hydrogen bonding is most important for the specificity of the helix. Since the hydrogen bonds rely on strict patterns of hydrogen bond donors and acceptors, and because these structures must be in just the right spots, hydrogen bonding allows for only complementary strands to come together: A- T, and C-G. This complementary nature allows DNA to carry the information that it does.

Chargaff's Rule:

Chargaff's rule states that the molar ratio of A to T and of G to C is almost always approximately equal in a DNA molecule. Chargaff's Rule is true as a result of the strict hydrogen bond forming rules in base pairing. For every G in a double-strand of DNA, there must be an accompanying complementary C, similarly, for each A, there is a complementary paired T.

DNA is a Right-Handed Helix:

Each strand of DNA wraps around the other in a right-handed configuration. In other words, the helix spirals upwards to the right. One can test the handedness" of a helix using the right hand rule. If you extend your right hand with thumb pointing up and imagine you are grasping a DNA double helix, as you trace upwards around the helix with your fingers, your hand is moving up. In a left-handed helix, in order to have your hand move upwards with your thumb pointing up, you would need to use your left hand. DNA is always found in the right-handed configuration.

The Major and Minor Grooves:

As a result of the double helical nature of DNA, the molecule has two asymmetric grooves. One groove is smaller than the other. This asymmetry is a result of the geometrical configuration of the bonds between the phosphate, sugar, and base groups that forces the base groups to attach at 120 degree angles instead of 180 degrees. The larger groove is called the major groove while the smaller one is called the minor groove.

Since the major and minor grooves expose the edges of the bases, the grooves can be used to tell the base sequence of a specific DNA molecule. The possibility for such recognition is critical, since proteins must be able to recognize specific DNA sequences on which to bind in order for the proper functions of the body and cell to be carried out. As you might expect, the major groove is more information rich than the minor groove. This fact makes the minor groove less ideal for protein binding.

Characteristics of the DNA Double-Helix:

DNA will adopt two different forms of helices under different conditions--the B- and A-forms. These two forms differ in their helical twist, rise, pitch and number of base pairs per turn. The twist of a helix refers to the number of degrees of angular rotation needed to get from one base unit to another. In the B-form of helix, this is 36 degrees while in the A-form it is 33 degrees. Rise refers to the height change from one base pair to the next and is 3.4 angstroms in the B-form and 2.6 angstroms in the A-form. The pitch is the height change to get one full rotation (360 degrees) of the helix. This value is 34 angstroms in the B-form since there are ten base pairs per turn. In the A-form, this value is 28 angstroms since there are eleven base pairs per full turn.

Of the two forms, the B-form is far more common, existing under most physiological conditions. The A-form is only adopted by DNA under conditions of low humidity. RNA, however, generally adopts the A-form in situations where the major and minor grooves are closer to the same size and the base pairs are a bit tilted with respect to the helical axis.

Base Pairing in DNA:

The nitrogen bases form the double-strand of DNA through weak hydrogen bonds. The nitrogen bases, however, have specific shapes and hydrogen bond properties so that guanine and cytosine only bond with each other, while adenine and thymine also bond exclusively. This pairing of the nitrogen bases is called complementary. In order for hydrogen bonding to occur at all, a hydrogen bond donor must have a complementary hydrogen bond acceptor in the base across from it. Common hydrogen bond donors include primary and secondary amine groups or hydroxyl groups. Common acceptor groups are carbonyls and tertiary amines.

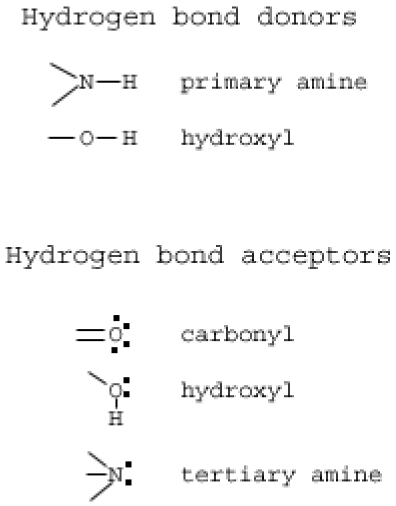


Figure: Common Hydrogen Bond Donors and Acceptors

There are three hydrogen bonds in a G:C base pair. One hydrogen bond forms between the 6' hydrogen bond accepting carbonyl of the guanine and the 4' hydrogen bond accepting primary amine of the cytosine. The second between the 1' secondary amine on guanine and the 3' tertiary amine on cytosine. And the third between the 2' primary amine on guanine and the 2' carbonyl on cytosine.

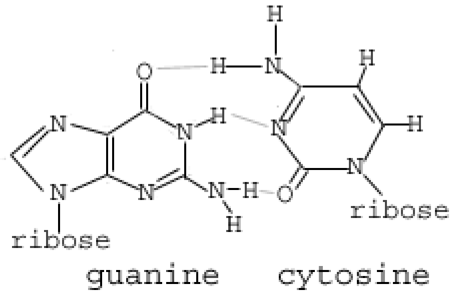


Figure: Guanine: Cytosine Base Pair

Between an A: T base pair, there are only two hydrogen bonds. One is found between the 6' primary amine of adenine and the 4' carbonyl of thymine. The other between the 1' tertiary amine of adenine and the 2' secondary amine of thymine ().

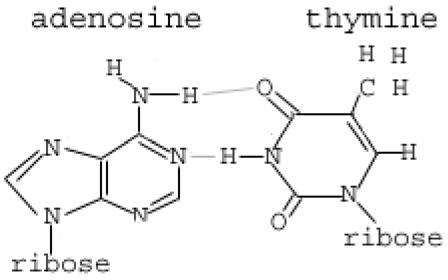


Figure: Adenine: Thymine Base Pair

STRUCTURE OF RNA:

RNA is a polymer of ribonucleotides held together by 3′, 5′-phosphodiester bridges. Although RNA has certain similarities with DNA structure, they have several specific differences

Pentose:

The sugar in RNA is ribose in contrast to deoxyribose in DNA.

Pyrimidine:

RNA contains the pyrimidine uracil in place of thymine (in DNA).

Single strand:

RNA is usually a single-stranded polynucleotide. However, this strand may fold at certain places to give a double-stranded structure, if complementary base pairs are in close proximity.

Chargaff's rule—not obeyed:

Due to the single-stranded nature, there is no specific relation between purine and pyrimidine contents. Thus the guanine content is not equal to cytosine (as is the case in DNA).

Susceptibility to alkali hydrolysis:

Alkali can hydrolyse RNA to 2′, 3′-cyclic diesters. This is possible due to the presence of a hydroxyl group at 2′ position. DNA cannot be subjected to alkali hydrolysis due to lack of this group.

Orcinol colour reaction:

RNAs can be histologically identified by orcinol colour reaction due to the presence of ribose.

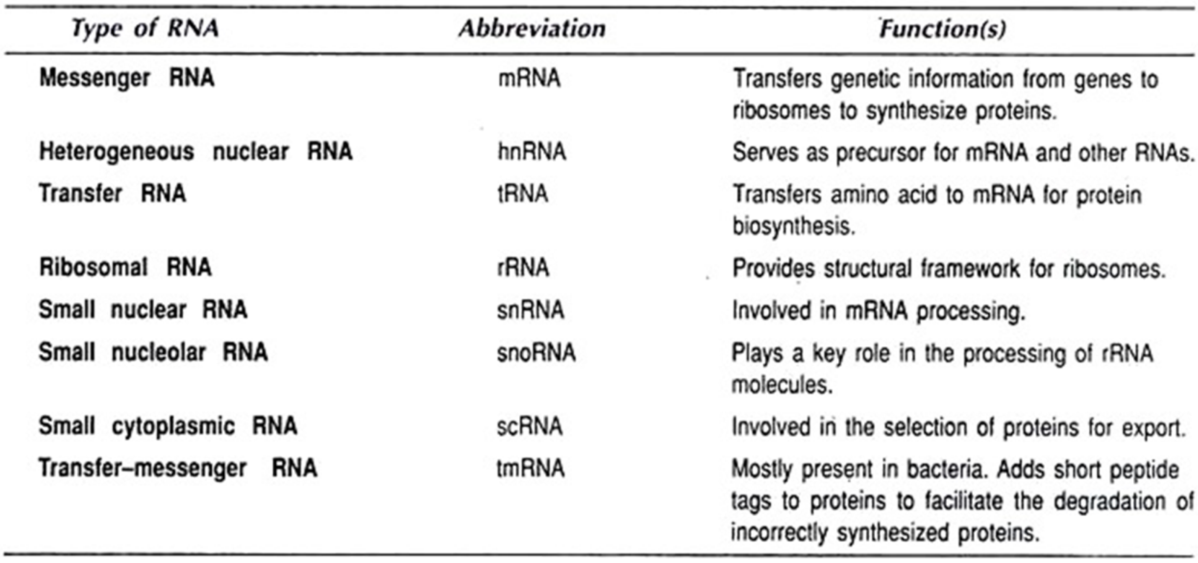
Types of RNA:

The three major types of RNAs with their respective cellular composition are given below

1. Messenger RNA (mRNA): 5-10%
2. Transfer RNA (tRNA): 10-20%
3. Ribosomal RNA (rRNA): 50-80%

Besides the three RNAs referred above, other RNAs are also present in the cells. These include heterogeneous nuclear RNA (hnRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA) and small cytoplasmic RNA (scRNA). The major functions of these RNAs are given in Table below.

Table: Cellular RNAs and their functions



Type of RNA	*Abbreviation*	*Function(s)*
Messenger RNA	mRNA	Transfers genetic information from genes to ribosomes to synthesize proteins.
Heterogeneous nuclear RNA	hnRNA	Serves as precursor for mRNA and other RNAs.
Transfer RNA	tRNA	Transfers amino acid to mRNA for protein biosynthesis.
Ribosomal RNA	rRNA	Provides structural framework for ribosomes.
Small nuclear RNA	snRNA	Involved in mRNA processing.
Small nucleolar RNA	snoRNA	Plays a key role in the processing of rRNA molecules.
Small cytoplasmic RNA	scRNA	Involved in the selection of proteins for export.
Transfer–messenger RNA	tmRNA	Mostly present in bacteria. Adds short peptide tags to proteins to facilitate the degradation of incorrectly synthesized proteins.

The RNAs are synthesized from DNA, and are primarily involved in the process of protein biosynthesis. The RNAs vary in their structure and function. A brief description on the major RNAs is given.

Messenger RNA (mRNA):

The mRNA is synthesized in the nucleus (in eukaryotes) as heterogeneous nuclear RNA (hnRNA). hnRNA, on processing, liberates the functional mRNA which enters the cytoplasm to participate in protein synthesis. mRNA has high molecular weight with a short half-life.

The eukaryotic mRNA is capped at the 5′-terminal end by 7-methylguanosine triphosphate. It is believed that this cap helps to prevent the hydrolysis of mRNA by 5′-exonucleases. Further, the cap may be also involved in the recognition of mRNA for protein synthesis.

The 3′-terminal end of mRNA contains a polymer of adenylate residues (20-250 nucleotides) which is known as poly (A) tail. This tail may provide stability to mRNA, besides preventing it from the attack of 3′-exonucleases. mRNA molecules often contain certain modified bases such as 6-methyladenylates in the internal structure.

Transfer RNA (tRNA):

Transfer RNA (soluble RNA) molecule contains 71-80 nucleotides (mostly 75) with a molecular weight of about 25,000. There are at least 20 species of tRNAs corresponding to 20 amino acids present in protein structure. The structure of tRNA (for alanine) was first elucidated by Holley. The structure of tRNA depicted in Fig. 2.14 resembles that of a clover leaf. tRNA contains mainly four arms, each arm with a base paired stem.

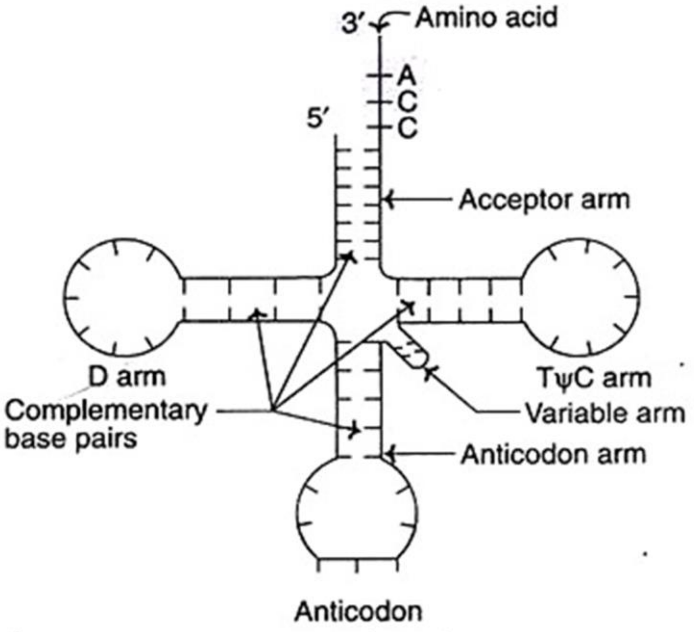


Fig: Structure of transfer RNA

1. **<u>The acceptor arm:</u>**

This arm is capped with a sequence CCA (5′ to 3′). The amino acid is attached to the acceptor arm.

2. **<u>The anticodon arm:</u>**

This arm, with the three specific nucleotide bases (anticodon), is responsible for the recognition of triplet codon of mRNA. The codon and anticodon are complementary to each other.

3. **<u>The D arm:</u>**

It is so named due to the presence of dihydrouridine.

4. **<u>The TΨC arm:</u>**

This arm contains a sequence of T, pseudouridine (represented by psi, Ψ) and C.

5. **<u>The variable arm:</u>**

This arm is the most variable in tRNA. Based on this variability, tRNAs are classified into 2 categories:

<u>Class I tRNAs:</u>

The most predominant (about 75%) form with 3-5 base pairs length.

<u>Class II tRNAs:</u>

They contain 13-20 base pair long arm.

Base pairs in tRNA:

The structure of tRNA is maintained due to the complementary base pairing in the arms.

The four arms with their respective base pairs are given below:

- ✓ The acceptor arm – 7 bp
- ✓ The TΨC arm – 5 bp
- ✓ The anticodon arm – 5 bp
- ✓ The D arm – 4 bp

Ribosomal RNA (rRNA):

The ribosomes are the factories of protein synthesis. The eukaryotic ribosomes are composed of two major nucleoprotein complexes-60S subunit and 40S subunit. The 60S subunit contains 28S rRNA, 5S rRNA and 5.8S rRNA while the 40S subunit contains 18S rRNA. The function of rRNAs in ribosomes is not clearly known. It is believed that they play a significant role in the binding of mRNA to ribosomes and protein synthesis.

Other RNAs:

The various other RNAs and their functions are summarised in Table above.

Catalytic RNAs—Ribozymes:

In certain instances, the RNA component of a ribonucleoprotein (RNA in association with protein) is catalytically active. Such RNAs are termed as ribozymes. At least five distinct species of RNA that act as catalysts have been identified. Three are involved in the self-processing reactions of RNAs while the other two are regarded as true catalysts (RNase P and rRNA). Ribonuclease P (RNase P) is a ribozyme containing protein and RNA component. It cleaves tRNA precursors to generate mature tRNA molecules.

RNA molecules are known to adapt tertiary structure just like proteins (i.e. enzymes). The specific conformation of RNA may be responsible for its function as biocatalyst. It is believed that ribozymes (RNAs) were functioning as catalysts before the occurrence of protein enzymes, during the course of evolution.