BIOMOLECULES
SEM-5, CC-12
PART-11, PPT-31

Part-11: Nucleic Acids-II

CONTENTS

1. *Hydrolysis of Nucleosides: Mechanism for acid catalyzed hydrolysis of nucleosides (both pyrimidine and purine types)*
2. *Primary Structure of DNAs and RNAs*
3. *Comparison of alkaline hydrolysis of DNA and RNA*
4. *Elementary idea of double helical structure of DNA (Watson-Crick model)*
5. *Complementary base pairing in DNA*
6. *Secondary Structure of RNAs*

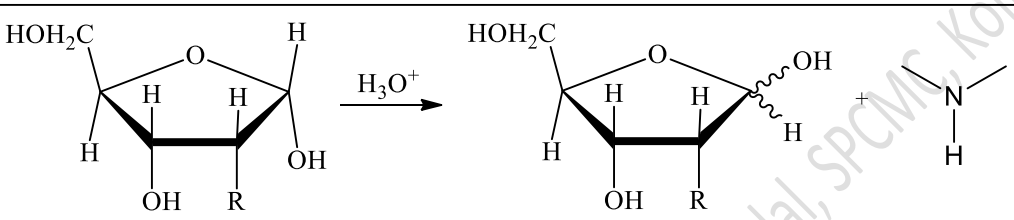
Dr. Kalyan Kumar Mandal
Associate Professor
St. Paul's C. M. College
Kolkata

BIOMOLECULES (PART-11, PPT-31)

Nucleic Acids-II

Hydrolysis of Nucleosides: Stability of *N*-Glycosidic Bond

The *N*-glycosidic bonds in most of the nucleosides obtained from RNA and DNA are usually highly stable in neutral and alkaline media (the intercellular pH is 7.0 to 7.4) but they are susceptible to hydrolysis in the presence of mineral and organic (HCO_2H, CH_3CO_2H, Cl_3CCO_2H, etc.) acids. The rate of hydrolysis depends on the hydrogen ion concentration. The reaction proceeds according to the general scheme as shown in Figure 1.



R = OH; D(-)-Ribose
R = H; 2-Deoxy-D(-)-Ribose

Figure 1:General scheme of hydrolysis of nucleosides

The ease of hydrolysis depends largely on the nature of the heterocyclic base (purine and pyrimidine) and the nature of the sugar unit. Purine nucleosides are much more susceptible to acid catalyzed hydrolysis than their pyrimidine counterparts. The rate of hydrolysis of different nucleosides corresponding to RNA and DNA follow the following order:

Deoxyguanosine > Deoxyadenosine > Guanosine > Adenosine >
Deoxycytidine ~ Deoxyuridine > Cytidine ~ Uridine

Purine ribonucleosides are hydrolyzed by heating with 0.1(N) HCl at 100 °C for 1 hour, whereas, pyrimidine nucleosides require boiling with 3(N) HCl at 125 °C for 4 hours. The rate of hydrolysis is also markedly affected by substituents in the heterocyclic ring. Glycosidic bonds in guanine derivatives are more sensitive to acids, as compared to those in adenine derivatives. Introduction of alkyl groups at positions 3 and 7 of the purine ring (alkylation of the nitrogens, N-3 and N-7) also causes a decrease in the glycosidic bond strength.

For example, 7-methyl-2´-deoxyguanosine (I; Figure 2) is hydrolyzed under mild conditions 10^4 times faster than 2´-deoxyguanosine (II). Again, the stability of the glycosidic bond is strongly dependent on the nature of the substituent at positions 2´ and 3´ in the sugar moiety of the nucleoside. Ribonucleosides are much more stable (10^2 to 10^3 times) toward hydrolysis than the corresponding deoxyribonucleosides.

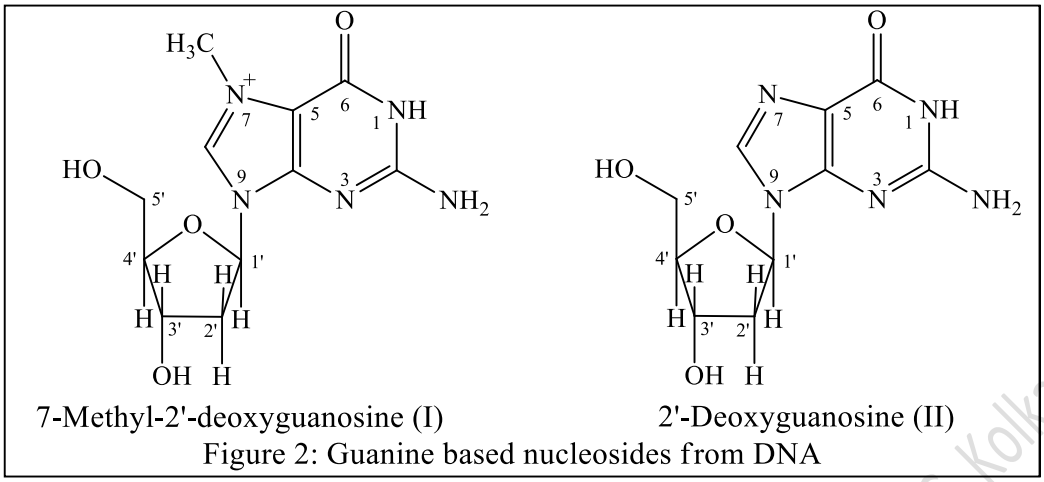


Figure 2: Guanine based nucleosides from DNA

Two different pathways have been outlined in Figures 3 & 4 regarding the mechanism of acid catalyzed hydrolysis of nucleosides. It is thought that pyrimidine nucleosides, thymine and uracil, tend to follow *pathway I*, via the intermediate formation of the Schiff bases as shown in Figure 3.

Hydrolysis of Nucleosides

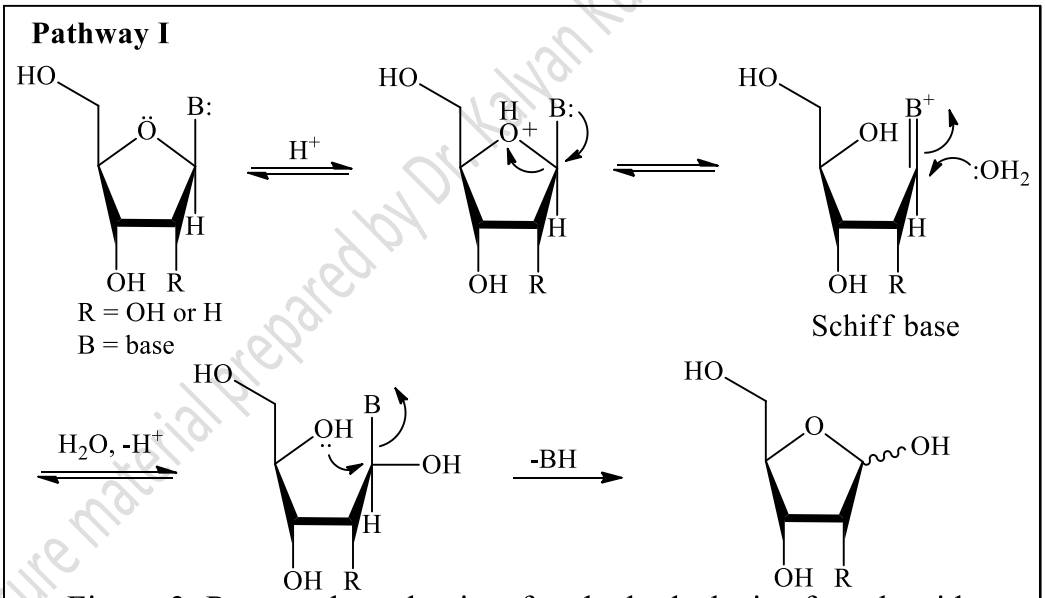


Figure 3: Proposed mechanism for the hydrolysis of nucleosides

On the contrary, purine nucleosides (adenosine and guanosine) and pyrimidine nucleoside, cytidine (cytosine nucleoside) follow *pathway-II* as shown in Figure 4 which involves an oxocarbenium ion intermediate.

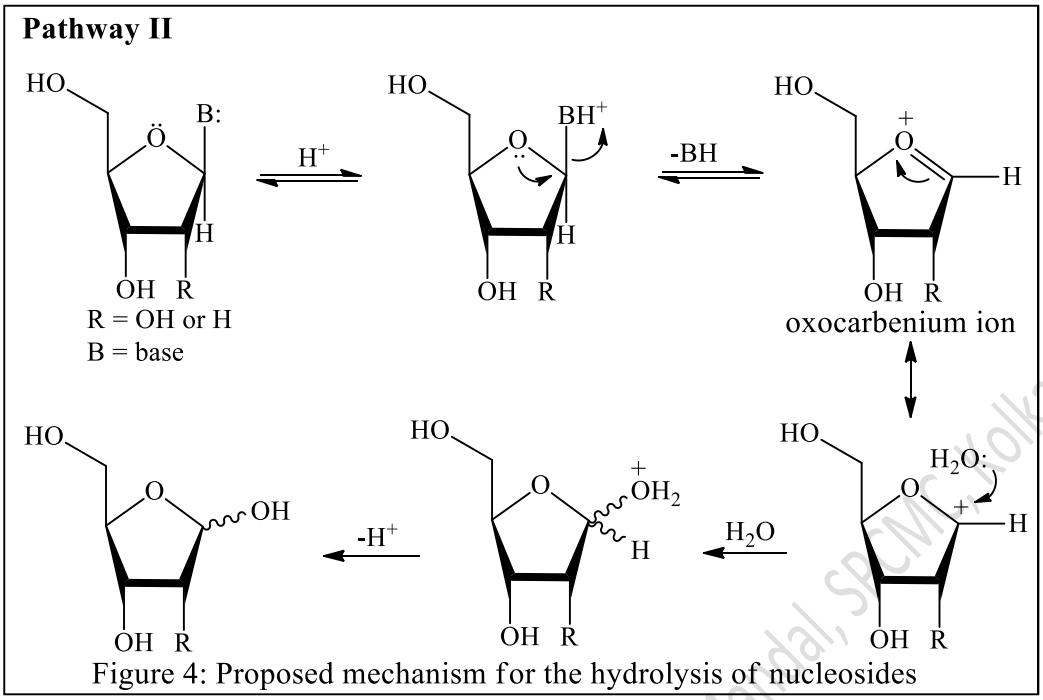


Figure 4: Proposed mechanism for the hydrolysis of nucleosides

It is generally observed that as the basicity of the heterocyclic base decreases, the tendency to follow *pathway I* increases. This is because the tendency to abstract proton by base decreases. Among pyrimidine bases, cytosine is unusually basic (pK_{aH+} = 4.6) and both thymine and uracil are very poor bases. Purine bases, adenine (pK_{aH+} = 4.3) and guanine (pK_{aH+} = 3.3) show basicity comparable to cytosine.

For purine bases the cleavage of the *N*-glycosidic bond is preceded by rapid protonation of the heterocyclic base, i.e., *pathway II* is followed. This scheme shows that in the rate-determining, unimolecular dissociation step, the base is present as the mono- or diprotonated state and it is then detached from the sugar moiety that leads to the formation of an oxocarbenium ion.

The presence of the electron withdrawing hydroxyl group at the 2´-position of ribose nucleus resists the formation of the oxocarbenium ion in the sugar moiety and slows down the rate of hydrolysis. Destabilization of the oxocarbenium ion in presence of an electron withdrawing group is the primary reason for that.This is why ribonucleosides resist hydrolysis more than the corresponding deoxyribonucleosides.

With adenosine, the first protonation takes place at N-1 and the second at N-7. With guanosine, the first protonation takes place at N-7 and the second at N-3. Protonation of the base moiety withdraws the electrons of the *N*-glycosidic bond toward N-9, leading to unimolecular fission of the bond. A resonance-stabilized oxocarbenium ion is obtained, onto which water rapidly attacks.

If *pathway I* is operating, then also the presence of the hydroxyl group at 2´-position of the sugar moiety discourages the Schiff base formation. The hydrolysis of the glycosidic bond in 2´-deoxyribocytidine is shown in Figure 5.

Protonation at N-7 of guanosine moiety makes the protonated nitrogen atom strongly electron withdrawing, and hence withdraws the electrons of the *N*-glycosidic bond toward N-9, leading to unimolecular fission of the bond. This electron withdrawal from the sugar moiety leads to unimolecular fission of the *N*-glycosidic bond. Consequently, guanosine or 2´-deoxyguanosine undergoes (mechanism shown in Figure 6) acid catalyzed hydrolysis at a faster rate.

Acid Catalyzed Hydrolysis of Uridine

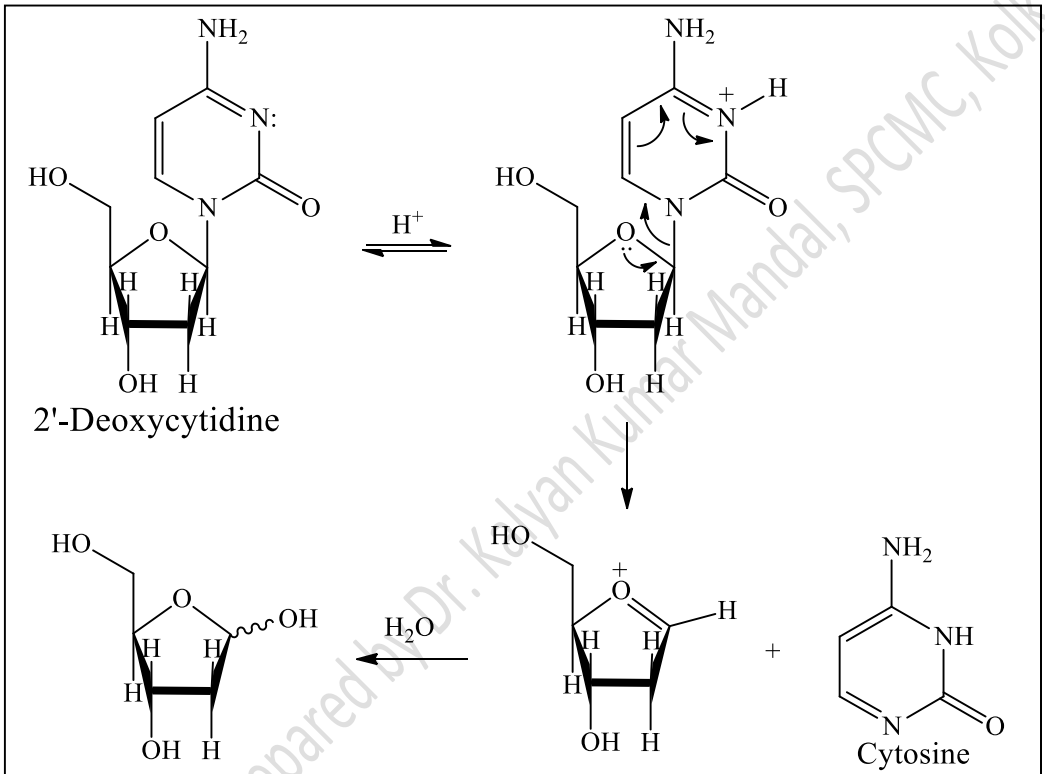


Figure 5: Mechanism for the acid catalyzed hydrolysis of 2'-deoxycytidine

Acid Catalyzed Hydrolysis of Guanosine

In the case of guanosine derivatives, protonation occurs at the nitrogen of the imidazole ring, N-7 is the site of maximum electron density to give monoprotonated species. For the deprotonated species, protonation occurs initially at N-7 and then at N-3.

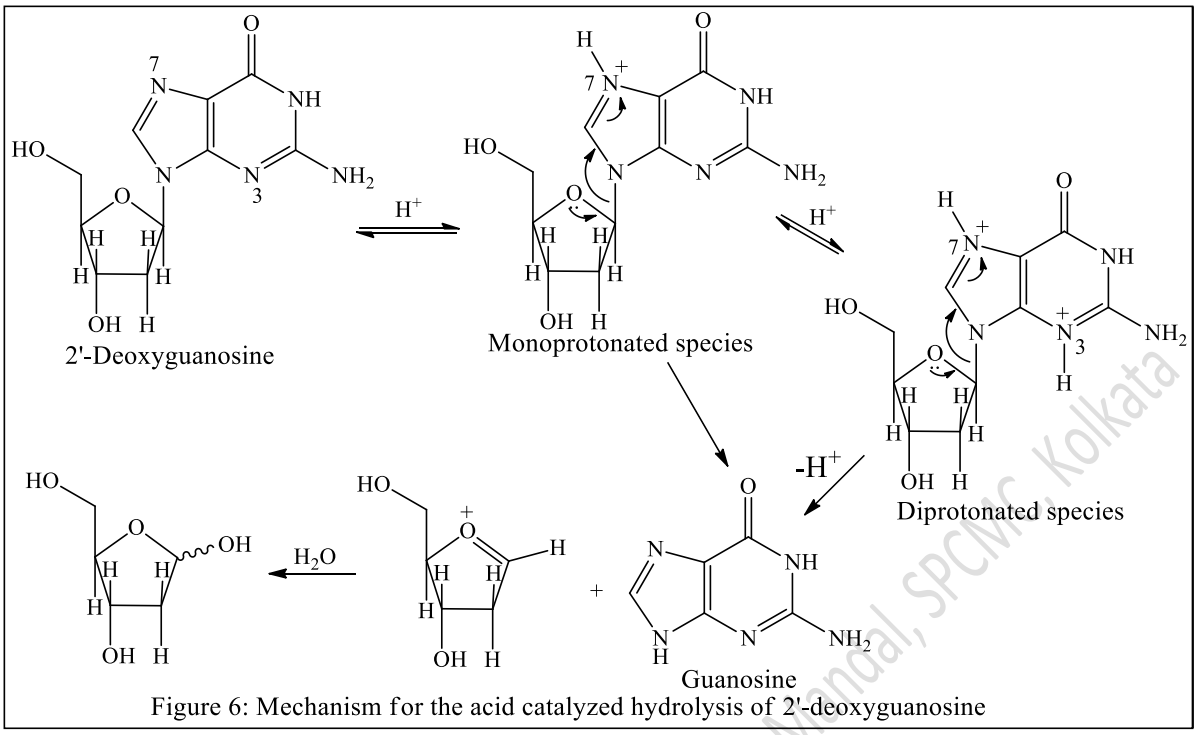


Figure 6: Mechanism for the acid catalyzed hydrolysis of 2'-deoxyguanosine

Hydrolysis of Nucleosides

Such a mechanism for the cleavage of the *N*-glycosidic bonds in purines is further supported by the extremely high labiality of nucleosides methylated at position 7 or 3. In this case, the positive charge at the respective nitrogen atom is fixed strongly enough. Therefore, such nucleosides are much less stable toward acid hydrolysis than the corresponding non-methylated analogues. Moreover, they can also be hydrolyzed in an alkaline medium.

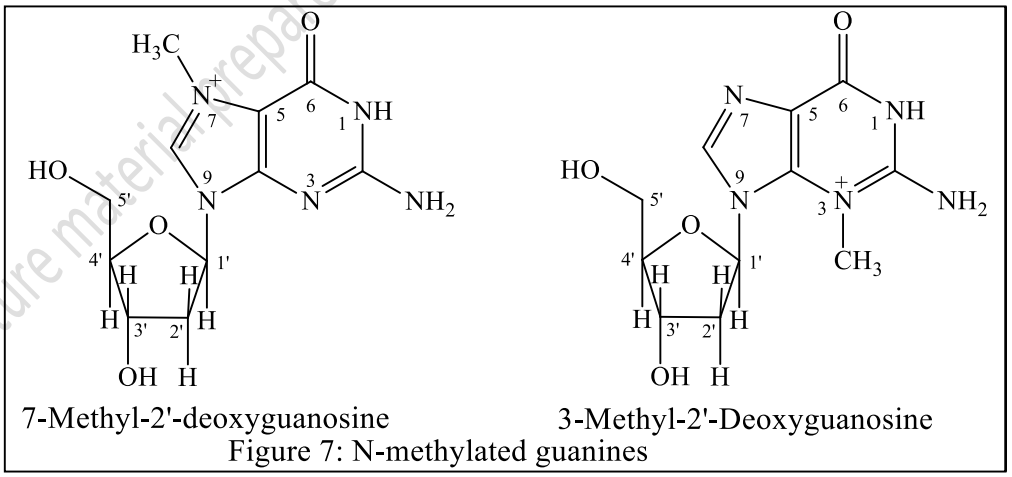


Figure 7: N-methylated guanines

Acid Catalyzed Hydrolysis of Adenosine

In the case of adenosine derivatives, protonation occurs at the nitrogen of the pyrimidine ring, N-1 is the site of maximum electron density to give monoprotonated species. For the deprotonated species, protonation occurs at initially at N-1 and then at N-7.

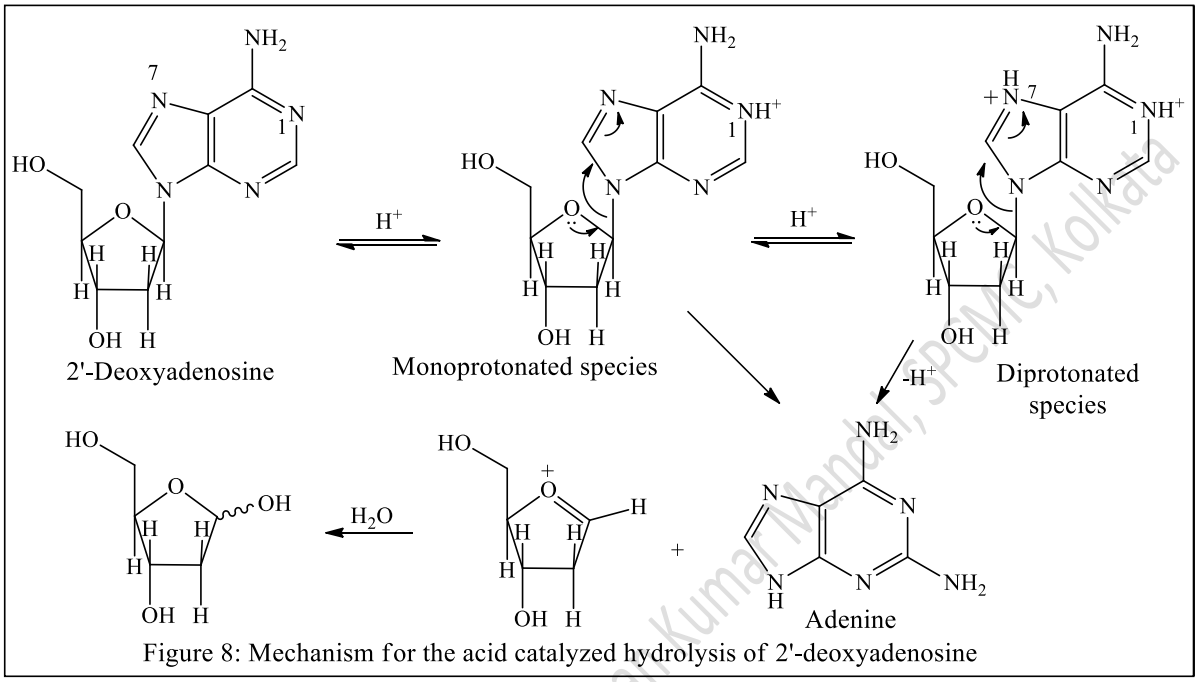


Figure 8: Mechanism for the acid catalyzed hydrolysis of 2'-deoxyadenosine

Hydrolysis of Nucleosides

It is to be noted that in the aminopurines, the position of protonation appears to be N-1 in adenine, whereas it is N-7 in guanine. This occurs because in adenosine amino groups (electron donating) at C-6 increases electron density at N-1 through electron donation. In case of guanosine, however, carbonyl group (at C-6) (electron withdrawing) decreases electron density at N-1. The effects are shown in Figure 9.

Presence of electron withdrawing carbonyl group in guanosine makes its purine moiety more electron withdrawing in both mono- and diprotonated species than that of the adenosine. The latter, however, contains the electron donating -NH_2 group which makes the purine moiety less electron withdrawing. Consequentlyfission of the *N*-glycosidic bond is comparatively slower in adenosine than that guanosine.

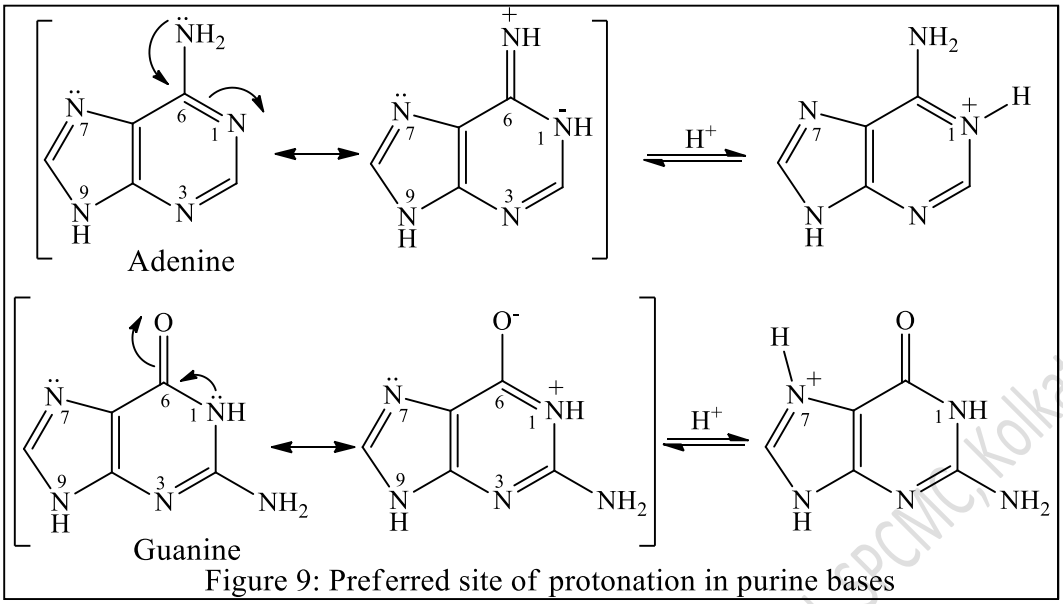


Figure 9: Preferred site of protonation in purine bases

This difference in site of protonation may contribute in a relatively higher hydrolysis rate observed for guanosine.

Deoxyribonucleic Acid: Primary Structure

Nucleotides bear the same relation to a nucleic acid that amino acids do to a protein; they are its monomeric units. The connecting links in proteins are amide groups and in nucleic acids they are phosphate ester linkages. Phosphate esters link the 3′-OH of one ribose (or deoxyribose) with the 5′-OH of another. Therefore, nucleic acids are polymers of nucleotides. Deoxyribo-nucleic acid (DNA) is a polymer of deoxyribonucleotides and is the storehouse of genetic information throughout all of nature (with the exception of certain viruses). The monomeric units of the DNA polymer are called residues.

This makes the nucleic acid a long unbranched chain with a "backbone" of sugar and phosphate units with heterocyclic bases protruding from the chain at regular intervals. The direction of the bases is indicated in the following way:

$$5' \leftarrow A - T - G - C - 3'$$

A four-residue segment (a tetranucleotide) of DNA is shown in Figure 10. This figure shows that the nucleotide residues in DNA are interconnected by phosphate groups that are esterified both to the 3′-OH group of one deoxyribose and the 5′-OH of another. The DNA polymer incorporates adenine, thymine, guanine, and cytosine as the nucleotide bases. Although only four residues are shown in Figure 10, a typical strand of DNA might be thousands or even millions of nucleotides long. *Each residue in a polynucleotide is differentiated by the identity of its base, and the sequence of bases encodes the genetic information in DNA.*

The DNA polymer is thus a backbone of alternating phosphates and 2′-deoxyribose groups to which are connected bases that differ from residue to residue. The ends of the DNA polymer

are labeled 3´or 5´, corresponding to the deoxyribose carbon to which the terminal hydroxy group is attached.

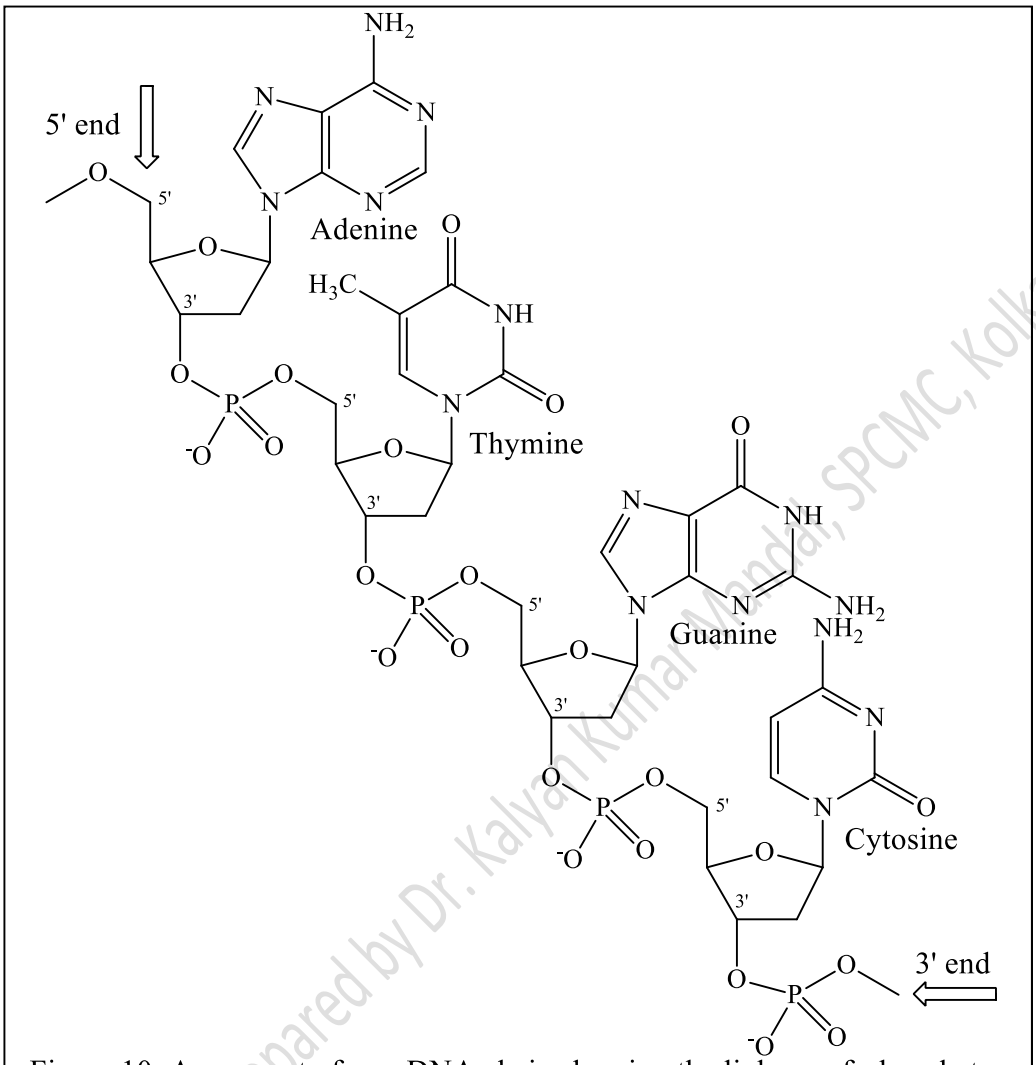


Figure 10: A segment of one DNA chain showing the linkage of phosphate ester groups with the 3'- and 5'-OH groups of deoxyribose units

Ribonucleic Acid: Primary Structure

It is the base sequence along the chain of DNA that contains the encoded genetic information. The sequence of bases can be determined using enzymatic methods and chromatography. Ribonucleic acid (RNA) polymers are conceptually much like DNA polymers, except that ribose, rather than 2´-deoxyribose, is the sugar. Three of the four bases in RNA are the same as in DNA. The fourth base, uracil, occurs in RNA instead of thymine, and some rare bases are found in certain types of RNA. Therefore, RNA has a similar structure (a tetranucleotide is shown in Figure 11) as that of DNA with two exceptions. A hydroxyl replaces a hydrogen atom at the 2´ position of each ribose unit and uracil replaces thymine.

Deoxyribonucleic Acid: Primary Structure

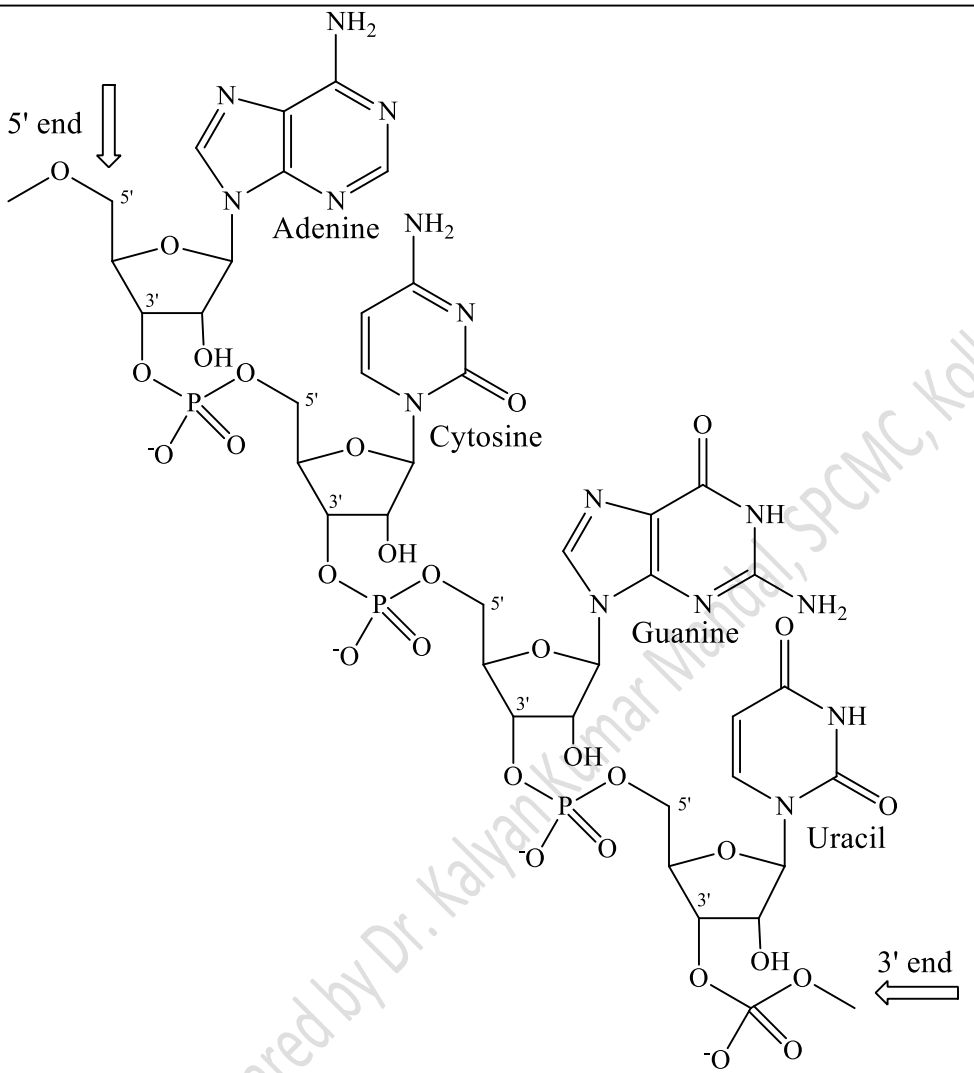


Figure 11: A segment of one RNA chain showing the linkage of phosphate ester groups with the 3'- and 5'-OH groups of ribose units (a tetranucleotide)

Alkaline Hydrolysis of RNA and DNA

When nucleotides are carefully hydrolyzed, ribose monophosphate may be isolated from the products. Therefore, the phosphoric acid is attached to the sugar residue in nucleotides. Examination of the nucleoside structures shows that the point of attachment may be 2´, 3´, 5´ in the ribose molecule, and 3´, 5´ in the deoxyribose molecule.

Alkaline hydrolysis of RNA gives a mixture of 2´- and 3´-phosphates. The mixture of 2´- and 3´-phosphates from RNAs has been shown to arise only in alkaline hydrolysates and has been explained on the basis of the formation of an intermediate 2´, 3´-cyclic phosphate. The proposed mechanism for the alkaline hydrolysis of RNA is shown in Figure 12.

Alkaline Hydrolysis of RNA

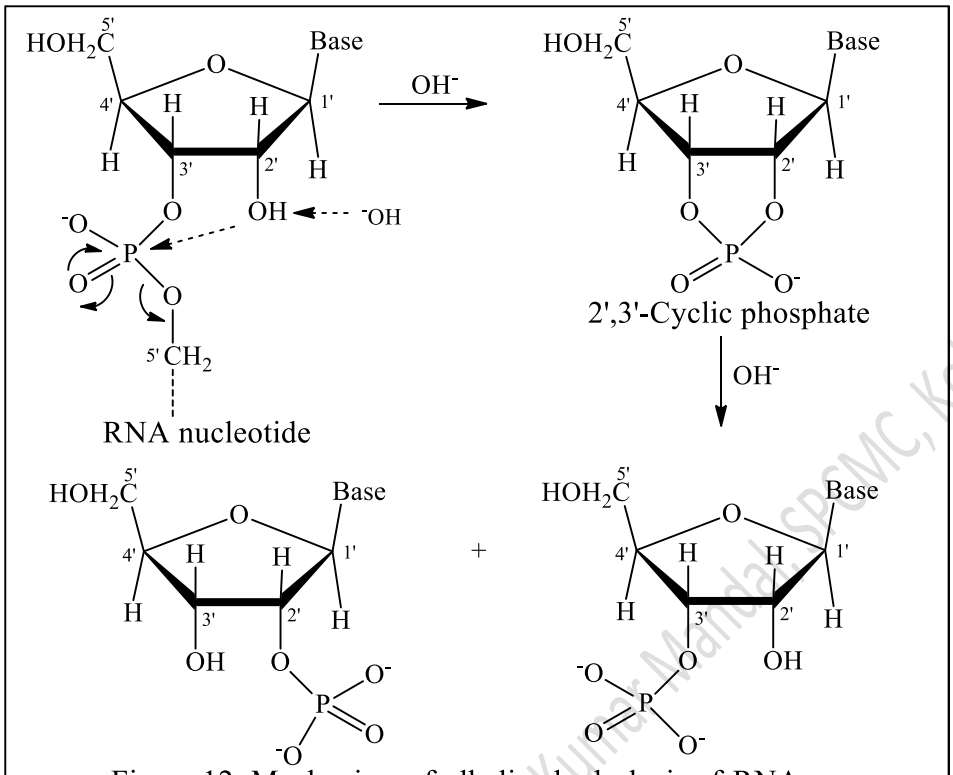


Figure 12: Mechanism of alkaline hydrolysis of RNA

Alkaline Hydrolysis of RNA and DNA

It can be said that only a ribonucleotide 5´-phosphate has an adjacent pair of *cis* hydroxyl groups (2´ and 3´). Hence, the 5´-phosphate can be readily distinguished from its isomers, 2´- and 3´-phosphates by means of periodic acid (HIO_4). General structure of both the nucleotide phosphates is shown in Figure 13.

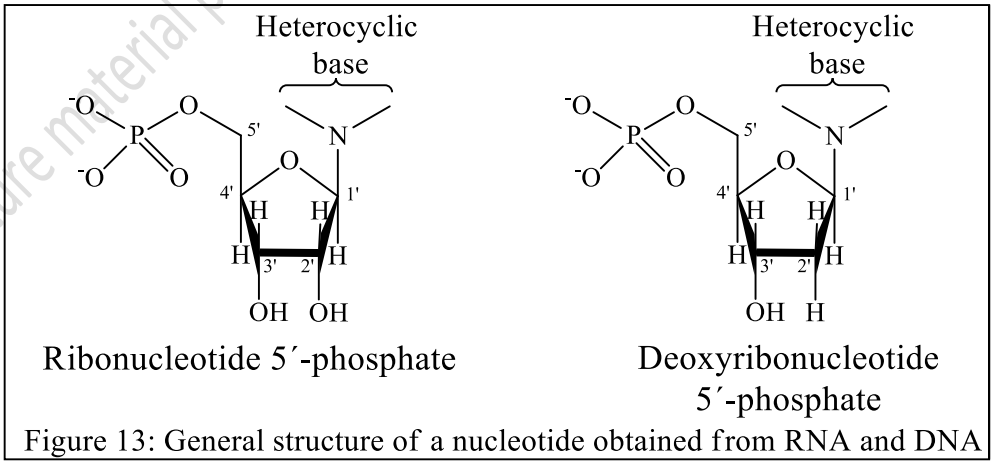


Figure 13: General structure of a nucleotide obtained from RNA and DNA

This method, however, fails to differentiate between deoxyribonucleoside 3´- and 5´-phosphates since neither of these contains a 2´-hydroxyl group. Alkaline hydrolysis of DNA is very slow. This is due to the absence of the 2´-hydroxyl group in deoxyribose, thereby

preventing the formation of the cyclic 2´,3´-phosphate which is readily formed from RNAs. This difference towards alkaline hydrolysis is used as a means of separating RNAs from DNAs.

Secondary Structure of DNA

It was the classic proposal of James Watson and Francis Crick (made in 1953 and verified shortly thereafter through the X-ray analysis by Maurice Wilkins) that gave a model for the secondary structure of DNA. This work earned Crick, Watson, and Wilkins the 1962 Nobel Prize in Physiology or Medicine. However, many believe that Rosalind Franklin, whose X-ray data was also key to solving the structure of DNA, should have shared the Nobel prize, but her death in 1958 precluded it.

The secondary structure of DNA is especially important because it enables to understand how genetic information is preserved, how it can be passed on during the process of cell division, and how it can be transcribed to provide a template for protein synthesis. Of prime importance to Watson and Crick's proposal was an earlier observation (made in the late 1940s) by Erwin Chargaff that certain regularities can be seen in the percentages of heterocyclic bases obtained from the DNA of a variety of species. Chargaff pointed out that for all species examined:

1. The total mole percentage of purines is approximately equal to that of the pyrimidines, that is, (%G + %A)/(%C + %T) ≈ 1.
2. The mole percentage of adenine is nearly equal to that of thymine (i.e., %A/%T ≈ 1), and the mole percentage of guanine is nearly equal to that of cytosine (i.e., %G/%C ≈ 1).
3. The ratio which varies from species to species is the ratio (%A + %T)/(%G + %C). He noted, moreover, that whereas this ratio is characteristic of the DNA of a given species, it is the same for DNA obtained from different tissues of the same animal and does not vary appreciably with the age or conditions of growth of individual organisms within the same species.

Watson and Crick also had X-ray data that gave them the bond lengths and angles of the purine and pyrimidine rings of model compounds. In addition, they had data from Franklin and Wilkins that indicated a repeat distance of 34 Å in DNA. Reasoning from these data, Watson and Crick proposed a double helix as a model for the secondary structure of DNA.The ratios of adenine to thymine, and guanosine to cytosine, in DNA are both 1.0. This observation has come to be known as *Chargaff's first parity rule*. The following regularities (with very few exceptions) in the composition of DNAs have been observed:

(a) A = T; (b) G = C

From that it follows that:

(c) A + G = T + C; (d) A + C = G + T

The Watson-Crick structure of DNA is shown in Figure 14. The structure has the following important features:

1. The structure of DNA contains *two* right-handed helical polynucleotide chains coiled around a common axis. The structure is therefore that of a *double helix*. The helix makes a complete turn every 10 nucleotide residues. The two polynucleotide chains run in opposite directions; that is, one chain runs in the 3′→5′ direction, and the other in the 5′→3′ direction.

2. The sugars and phosphates, which are rich in -OH groups and charges, are on the outside of the helix, where they can interact with solvent water or other hydrophilic compounds. The bases, which are hydrophobic, are buried in the interior of the double helix, away from water.

3. The chains are held together by hydrogen bonds between bases. *Each adenine (A) in one chain forms hydrogen bonds to a thymine (T) in the other, and each guanosine (G) in one chain forms hydrogen bonds to a cytosine (C) in the other*.

   The two chains are wound in such a manner that pyrimidine and purine bases points towards each other. Every purine in one chain is hydrogen-bonded to a pyrimidine in the other. For this reason, adenine (A) is said to be *complementary* to thymine (T), and guanine (G) is *complementary* to cytosine (C).

   The hydrogen-bonded A-T and G-C pairs are often referred to as Waston-Crick base pairs. Figure 15 provides a closer look at these Watson-Crick base pairs. It is to be noted that the A-T pair has about the same spatial dimensions as the G-C pair.

4. The planes of successive complementary base pairs are stacked, one on top of the other. The ring-planes of each pair of bases lie in the same plane and are perpendicular to the axis of the helix. The distance between each successive base-pair plane is 3.4 Å. Because the helix makes a complete turn every 10 residues, there is a distance of 10 x 3.4 = 34 Å along the helix per complete turn.

5. The double-helical structure of DNA results in two grooves that wrap around the double helix along its periphery. The larger groove is called the *major groove*, and the smaller is called the *minor groove.* These are shown in Figure 14. These grooves, particularly the major groove, are sites at which other macromolecules such as proteins are found to interact with DNA.

6. There is no intrinsic restriction on the sequence of bases in a polynucleotide. However, as a result of the base pairing, A-T and G-C, the sequence of one polynucleotide strand (the “Watson” strand) in the double helix is complementary to that in the other strand (the “Crick” strand). Thus, everywhere there is an A in one strand, there is a T in the other; everywhere there is a G in one strand, there is a C in the other.

DNA Double Helix

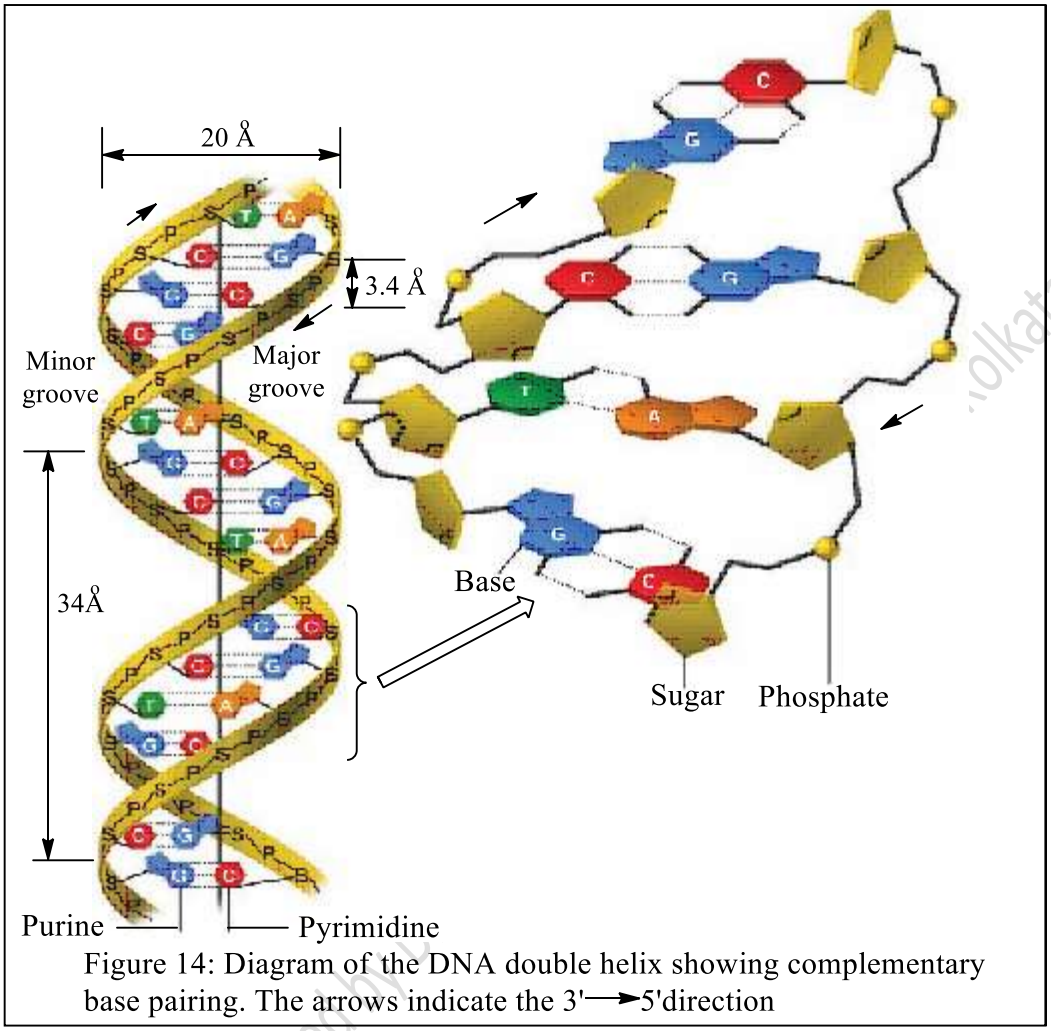
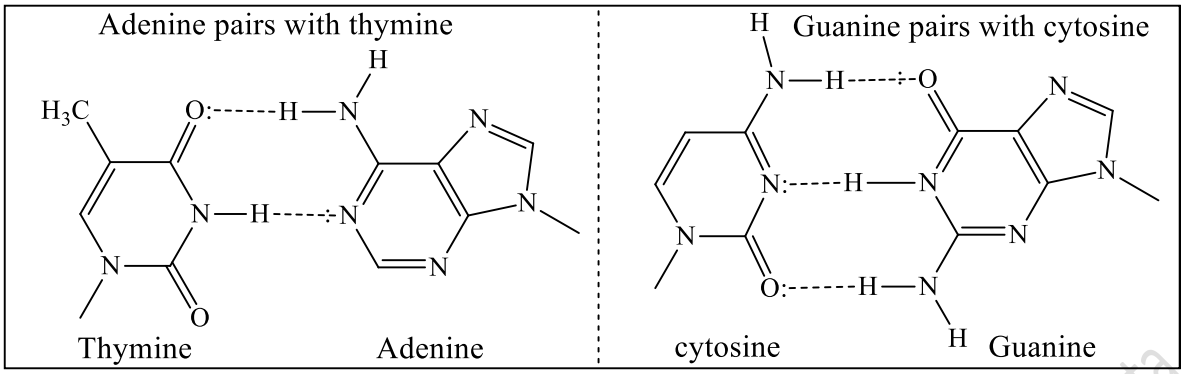


Figure 14: Diagram of the DNA double helix showing complementary base pairing. The arrows indicate the 3'⟶5'direction

Hydrogen-bonding complementarity in DNA accounts nicely for the Chargaff parity rule: if A always hydrogen bonds to T and G always hydrogen bonds to C, then the number of A residues must equal the number of T residues, and the number of G residues must equal the number of C residues. The A-T pair is held together by two hydrogen bonds and the G-C pair by three hydrogen bonds. Pairing of bases can occur only between a pyrimidine and a purine because of steric reason.

This structure also suggests a reasonable mechanism for the duplication of DNA during cell division: the two strands come apart, and a new strand is grown as a complement of each original strand. In other words, *the proper sequence of each new DNA strand during cellular reproduction is ensured by hydrogen-bonding complementarity* (Figure 15).



With DNAs, the sum of the keto-bases (G + T) is equal to the sum of the amino bases (A + C). The equivalence of A and T and of G and C are of paramount importance in connection with the secondary structure of DNAs. The molecular weights of DNAs have been determined by various physical methods. The values obtained range from about 10^6 to 10^9.

Secondary Structure of RNAs

It has been shown that the keto-bases (guanine and uracil) and the amino-bases (adenine and cytosine) are present in all RNAs in roughly equal amounts. Various methods have been used to determine the molecular weights of purified nucleic acids, e.g., end-group assay, ultracentrifugation, light scattering, etc. values obtained for RNAs range from about 2×10^4 to 2×10^6.

The secondary structure of RNAs has also been investigated. The results (X-ray analysis) appear to indicate that RNAs exist as single strands which contain helical segments stabilized by hydrogen bonding. There are, however, some example of RNAs which exist as double stands (double helical structure).