12083CH06

CHAPTER 6

MOLECULAR BASIS OF INHERITANCE

In the previous chapter, you have learnt the inheritance patterns and the genetic basis of such patterns. At the time of Mendel, the nature of those 'factors' regulating the pattern of inheritance was not clear. Over the next hundred years, the nature of the putative genetic material was investigated culminating in the realisation that DNA – deoxyribonucleic acid – is the genetic material, at least for the majority of organisms. In class XI you have learnt that nucleic acids are polymers of nucleotides.

Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are the two types of nucleic acids found in living systems. DNA acts as the genetic material in most of the organisms. RNA though it also acts as a genetic material in some viruses, mostly functions as a messenger. RNA has additional roles as well. It functions as adapter, structural, and in some cases as a catalytic molecule. In Class XI you have already learnt the structures of nucleotides and the way these monomer units are linked to form nucleic acid polymers. In this chapter we are going to discuss the structure of DNA, its replication, the process of making RNA from DNA (transcription), the genetic code that determines the sequences of amino acids in proteins, the process of protein synthesis (translation) and elementary basis of their regulation. The determination

of complete nucleotide sequence of human genome during last decade has set in a new era of genomics. In the last section, the essentials of human genome sequencing and its consequences will also be discussed.

Let us begin our discussion by first understanding the structure of the most interesting molecule in the living system, that is, the DNA. In subsequent sections, we will understand that why it is the most abundant genetic material, and what its relationship is with RNA.

6.1 The DNA

DNA is a long polymer of deoxyribonucleotides. The length of DNA is usually defined as number of nucleotides (or a pair of nucleotide referred to as base pairs) present in it. This also is the characteristic of an organism. For example, a bacteriophage known as ϕ×174 has 5386 nucleotides, Bacteriophage lambda has 48502 base pairs (bp), *Escherichia coli* has 4.6×10^6 bp, and haploid content of human DNA is 3.3×10^9 bp. Let us discuss the structure of such a long polymer.

6.1.1 Structure of Polynucleotide Chain

Let us recapitulate the chemical structure of a polynucleotide chain (DNA or RNA). A nucleotide has three components – a nitrogenous base, a pentose sugar (ribose in case of RNA, and deoxyribose for DNA), and a phosphate group. There are two types of nitrogenous bases – Purines (Adenine and Guanine), and Pyrimidines (Cytosine, Uracil and Thymine). Cytosine is common for both DNA and RNA and Thymine is present in DNA. Uracil is present in RNA at the place of Thymine. A nitrogenous base is linked to the OH of 1′C pentose sugar through a N-glycosidic linkage to form a nucleoside, such as adenosine or deoxyadenosine, guanosine or deoxyguanosine, cytidine or deoxycytidine and uridine or deoxythymidine. When a phosphate group is linked to OH of 5′C of a nucleoside through phosphoester linkage, a corresponding nucleotide (or deoxynucleotide depending upon the type of sugar present) is formed. Two nucleotides are linked through 3′-5′ phosphodiester linkage to form a dinucleotide. More nucleotides can be joined in such a manner to form a polynucleotide chain. A polymer thus formed has at one end a free

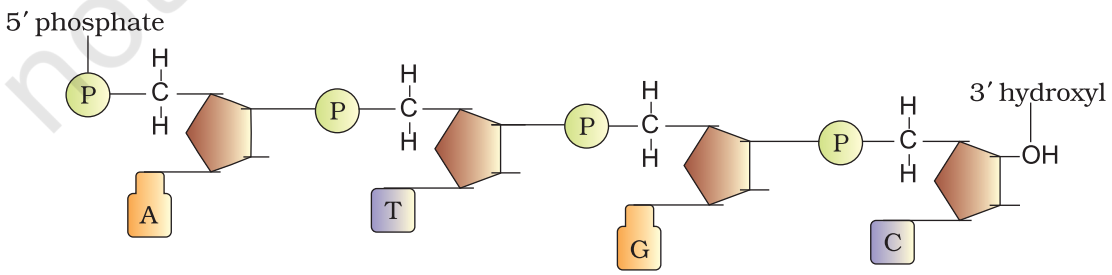


Figure 6.1 A Polynucleotide chain

phosphate moiety at 5′-end of sugar, which is referred to as 5'-end of polynucleotide chain. Similarly, at the other end of the polymer the sugar has a free OH of 3′C group which is referred to as 3′-end of the polynucleotide chain. The backbone of a polynucleotide chain is formed due to sugar and phosphates. The nitrogenous bases linked to sugar moiety project from the backbone (Figure 6.1).

In RNA, every nucleotide residue has an additional –OH group present at 2′-position in the ribose. Also, in RNA the uracil is found at the place of thymine (5-methyl uracil, another chemical name for thymine).

DNA as an acidic substance present in nucleus was first identified by Friedrich Meischer in 1869. He named it as 'Nuclein'. However, due to technical limitation in isolating such a long polymer intact, the elucidation of structure of DNA remained elusive for a very long period of time. It was only in 1953 that James Watson and Francis Crick, based on the X-ray diffraction data produced by Maurice Wilkins and Rosalind Franklin, proposed a very simple but famous **Double Helix** model for the structure of DNA. One of the hallmarks of their proposition was base pairing between the two strands of polynucleotide chains. However, this proposition was also based on the observation of Erwin Chargaff that for a double stranded DNA, the ratios between **Adenine** and **Thymine** and **Guanine** and **Cytosine** are constant and equals one.

The base pairing confers a very unique property to the polynucleotide chains. They are said to be complementary to each other, and therefore if the sequence of bases in one strand is known then the sequence in other strand can be predicted. Also, if each strand from a DNA (let us call it as a parental DNA) acts as a template for synthesis of a new strand, the two double stranded DNA (let us call them as daughter DNA) thus, produced would be identical to the parental DNA molecule. Because of this, the genetic implications of the structure of DNA became very clear.

The salient features of the Double-helix structure of DNA are as follows:

(i) It is made of two polynucleotide chains, where the backbone is constituted by sugar-phosphate, and the bases project inside.

(ii) The two chains have anti-parallel polarity. It means, if one chain has the polarity 5′→3′, the other has 3′→5′.

(iii) The bases in two strands are paired through hydrogen bond (H-bonds) forming base pairs (bp). Adenine forms two hydrogen bonds with Thymine from opposite strand and vice-versa. Similarly, Guanine is bonded with Cytosine with three H-bonds. As a result, always a purine comes opposite to a pyrimidine. This generates approximately uniform distance between the two strands of the helix (Figure 6.2).

(iv) The two chains are coiled in a right-handed fashion. The pitch of the helix is 3.4 nm (a nanometre is one billionth of a metre, that is 10^{-9} m) and there are roughly 10 bp in each

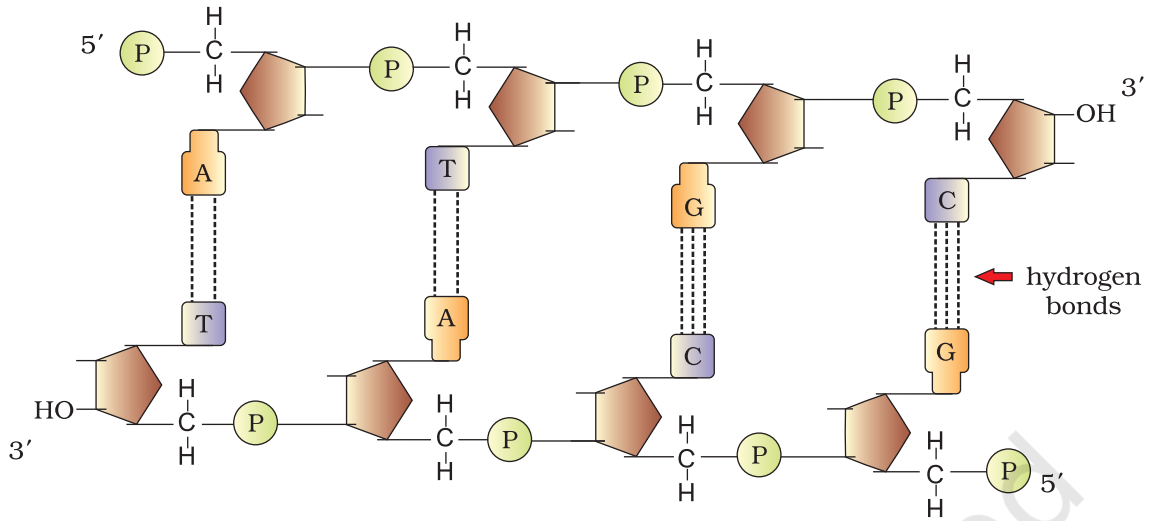


Figure 6.2 Double stranded polynucleotide chain

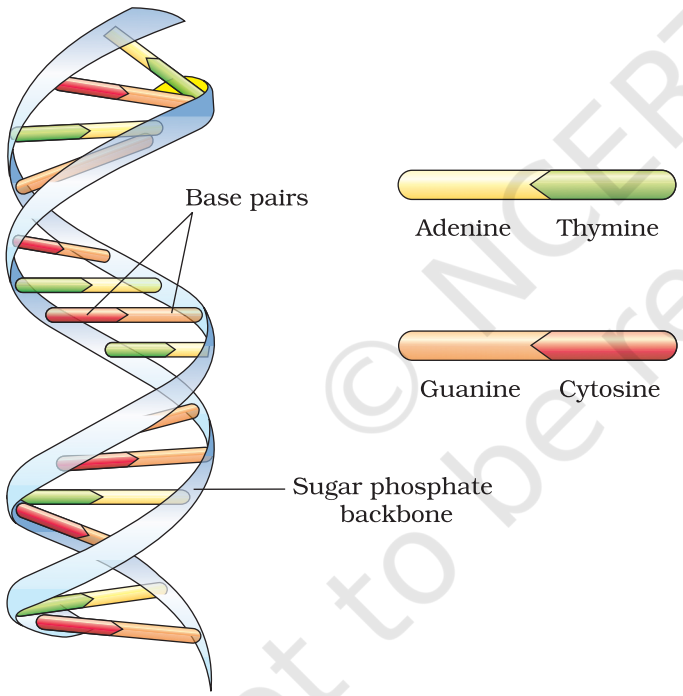


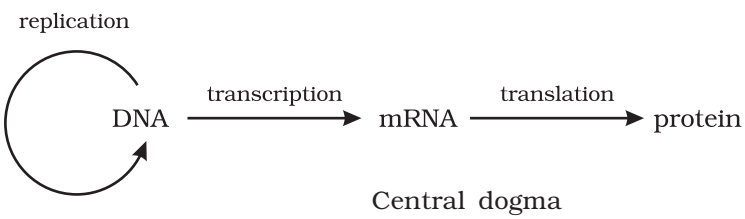
Figure 6.3 DNA double helix

turn. Consequently, the distance between a bp in a helix is approximately 0.34 nm.

(v) The plane of one base pair stacks over the other in double helix. This, in addition to H-bonds, confers stability of the helical structure (Figure 6.3).

Compare the structure of purines and pyrimidines. Can you find out why the distance between two polynucleotide chains in DNA remains almost constant?

The proposition of a double helix structure for DNA and its simplicity in explaining the genetic implication became revolutionary. Very soon, Francis Crick proposed the Central dogma in molecular biology, which states that the genetic information flows from DNA→RNA→Protein.



Central dogma

In some viruses the flow of information is in reverse direction, that is, from RNA to DNA. *Can you suggest a simple name to the process?*

6.1.2 Packaging of DNA Helix

Taken the distance between two consecutive base pairs as 0.34 nm (0.34×10^{-9} m), if the length of DNA double helix in a typical mammalian cell is calculated (simply by multiplying the total number of bp with distance between two consecutive bp, that is, 6.6×10^{9} bp × 0.34×10^{-9} m/bp), it comes out to be approximately 2.2 metres. A length that is far greater than the dimension of a typical nucleus (approximately 10^{-6} m). How is such a long polymer packaged in a cell?

If the length of E. coli DNA is 1.36 mm, can you calculate the number of base pairs in E.coli?

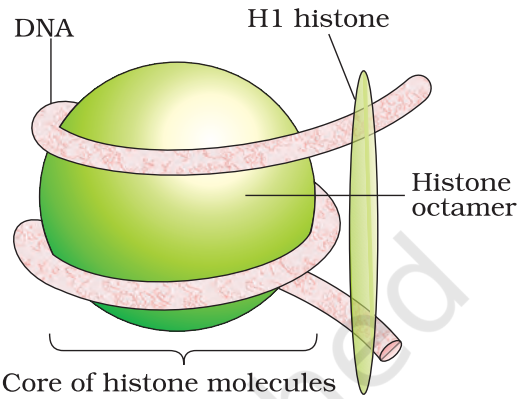


Figure 6.4a Nucleosome

In prokaryotes, such as, *E. coli*, though they do not have a defined nucleus, the DNA is not scattered throughout the cell. DNA (being negatively charged) is held with some proteins (that have positive charges) in a region termed as 'nucleoid'. The DNA in nucleoid is organised in large loops held by proteins.

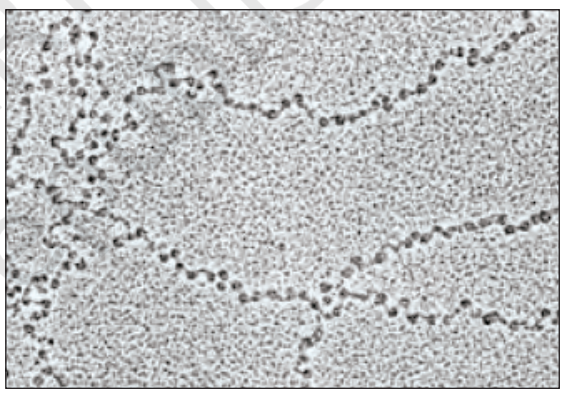


Figure 6.4b EM picture - 'Beads-on-String'

In eukaryotes, this organisation is much more complex. There is a set of positively charged, basic proteins called **histones**. A protein acquires charge depending upon the abundance of amino acids residues with charged side chains. Histones are rich in the basic amino acid residues lysine and arginine. Both the amino acid residues carry positive charges in their side chains. Histones are organised to form a unit of eight molecules called **histone octamer**. The negatively charged DNA is wrapped around the positively charged histone octamer to form a structure called **nucleosome** (Figure 6.4 a). A typical nucleosome contains 200 bp of DNA helix. Nucleosomes constitute the repeating unit of a structure in nucleus called **chromatin,** thread-like stained (coloured) bodies seen in nucleus. The nucleosomes in chromatin are seen as 'beads-on-string' structure when viewed under electron microscope (EM) (Figure 6.4 b).

Theoretically, how many such beads (nucleosomes) do you imagine are present in a mammalian cell?

The beads-on-string structure in chromatin is packaged to form chromatin fibers that are further coiled and condensed at metaphase stage of cell division to form chromosomes. The packaging of chromatin at higher level requires additional set of proteins that collectively are referred to as

Non-histone Chromosomal (NHC) proteins. In a typical nucleus, some region of chromatin are loosely packed (and stains light) and are referred to as **euchromatin**. The chromatin that is more densely packed and stains dark are called as **Heterochromatin**. Euchromatin is said to be transcriptionally active chromatin, whereas heterochromatin is inactive.

6.2 The Search for Genetic Material

Even though the discovery of nuclein by Meischer and the proposition for principles of inheritance by Mendel were almost at the same time, but that the DNA acts as a genetic material took long to be discovered and proven. By 1926, the quest to determine the mechanism for genetic inheritance had reached the molecular level. Previous discoveries by Gregor Mendel, Walter Sutton, Thomas Hunt Morgan and numerous other scientists had narrowed the search to the chromosomes located in the nucleus of most cells. But the question of what molecule was actually the genetic material, had not been answered.

Transforming Principle

In 1928, Frederick Griffith, in a series of experiments with *Streptococcus pneumoniae* (bacterium responsible for pneumonia), witnessed a miraculous transformation in the bacteria. During the course of his experiment, a living organism (bacteria) had changed in physical form.

When *Streptococcus pneumoniae* (pneumococcus) bacteria are grown on a culture plate, some produce smooth shiny colonies (S) while others produce rough colonies (R). This is because the S strain bacteria have a mucous (polysaccharide) coat, while R strain does not. Mice infected with the S strain (virulent) die from pneumonia infection but mice infected with the R strain do not develop pneumonia.

S strain ⟶ Inject into mice ⟶ Mice die

R strain ⟶ Inject into mice ⟶ Mice live

Griffith was able to kill bacteria by heating them. He observed that heat-killed S strain bacteria injected into mice did not kill them. When he

S strain (heat-killed) ⟶ Inject into mice ⟶ Mice live

S strain (heat-killed) + R strain (live) ⟶ Inject into mice ⟶ Mice die

injected a mixture of heat-killed S and live R bacteria, the mice died. Moreover, he recovered living S bacteria from the dead mice.

He concluded that the R strain bacteria had somehow been **transformed** by the heat-killed S strain bacteria. Some 'transforming principle', transferred from the heat-killed S strain, had enabled the R strain to synthesise a smooth polysaccharide coat and become virulent. This must be due to the transfer of the genetic material. However, the biochemical nature of genetic material was not defined from his experiments.

Biochemical Characterisation of Transforming Principle

Prior to the work of Oswald Avery, Colin MacLeod and Maclyn McCarty (1933-44), the genetic material was thought to be a protein. They worked to determine the biochemical nature of 'transforming principle' in Griffith's experiment.

They purified biochemicals (proteins, DNA, RNA, etc.) from the heat-killed S cells to see which ones could transform live R cells into S cells. They discovered that DNA alone from S bacteria caused R bacteria to become transformed.

They also discovered that protein-digesting enzymes (proteases) and RNA-digesting enzymes (RNases) did not affect transformation, so the transforming substance was not a protein or RNA. Digestion with DNase did inhibit transformation, suggesting that the DNA caused the transformation. They concluded that DNA is the hereditary material, but not all biologists were convinced.

Can you think of any difference between DNAs and DNase?

6.2.1 The Genetic Material is DNA

The unequivocal proof that DNA is the genetic material came from the experiments of Alfred Hershey and Martha Chase (1952). They worked with viruses that infect bacteria called bacteriophages.

The bacteriophage attaches to the bacteria and its genetic material then enters the bacterial cell. The bacterial cell treats the viral genetic material as if it was its own and subsequently manufactures more virus particles. Hershey and Chase worked to discover whether it was protein or DNA from the viruses that entered the bacteria.

They grew some viruses on a medium that contained radioactive phosphorus and some others on medium that contained radioactive sulfur. Viruses grown in the presence of radioactive phosphorus contained radioactive DNA but not radioactive protein because DNA contains phosphorus but protein does not. Similarly, viruses grown on radioactive sulfur contained radioactive protein but not radioactive DNA because DNA does not contain sulfur.

Radioactive phages were allowed to attach to *E. coli* bacteria. Then, as the infection proceeded, the viral coats were removed from the bacteria by agitating them in a blender. The virus particles were separated from the bacteria by spinning them in a centrifuge.

Bacteria which was infected with viruses that had radioactive DNA were radioactive, indicating that DNA was the material that passed from the virus to the bacteria. Bacteria that were infected with viruses that had radioactive proteins were not radioactive. This indicates that proteins did not enter the bacteria from the viruses. DNA is therefore the genetic material that is passed from virus to bacteria (Figure 6.5).

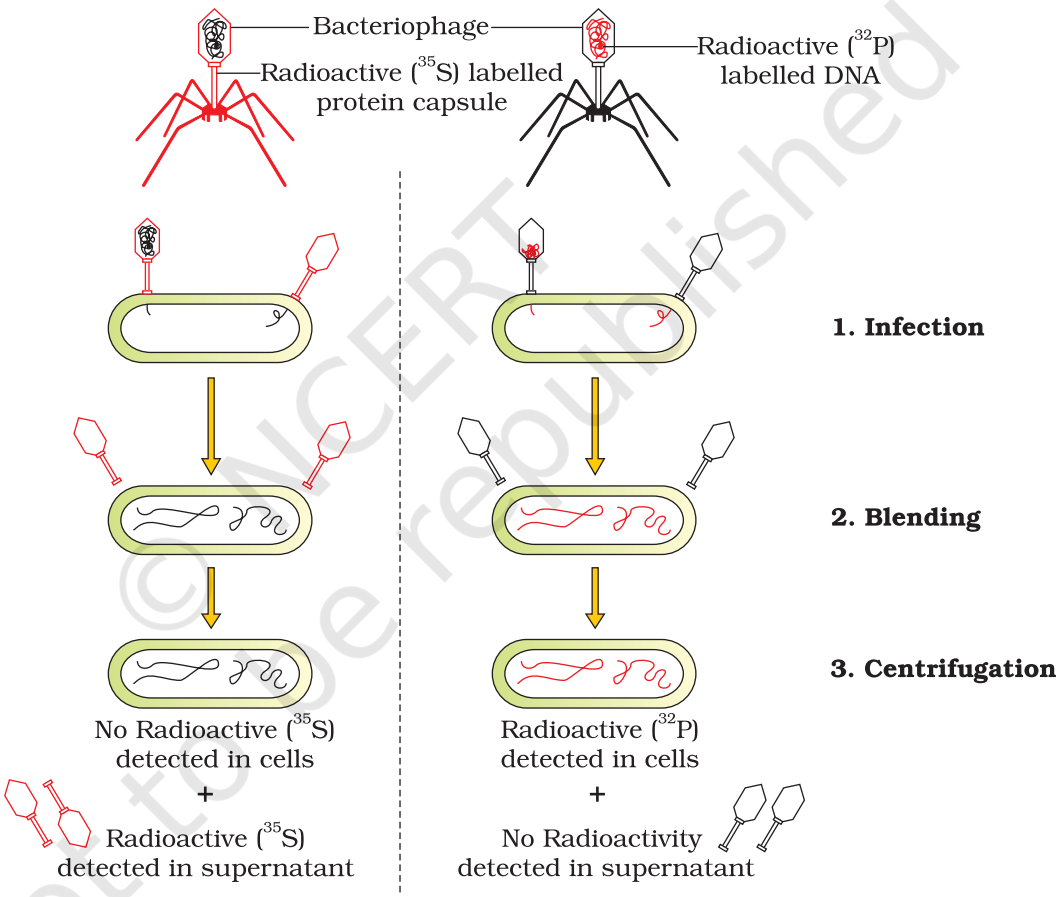


Figure 6.5 The Hershey-Chase experiment

6.2.2 Properties of Genetic Material (DNA versus RNA)

From the foregoing discussion, it is clear that the debate between proteins versus DNA as the genetic material was unequivocally resolved from Hershey-Chase experiment. It became an established fact that it is DNA that acts as genetic material. However, it subsequently became clear that

in some viruses, RNA is the genetic material (for example, Tobacco Mosaic viruses, QB bacteriophage, etc.). Answer to some of the questions such as, why DNA is the predominant genetic material, whereas RNA performs dynamic functions of messenger and adapter has to be found from the differences between chemical structures of the two nucleic acid molecules.

Can you recall the two chemical differences between DNA and RNA?

A molecule that can act as a genetic material must fulfill the following criteria:

(i) It should be able to generate its replica (Replication).

(ii) It should be stable chemically and structurally.

(iii) It should provide the scope for slow changes (mutation) that are required for evolution.

(iv) It should be able to express itself in the form of 'Mendelian Characters'.

If one examines each requirement one by one, because of rule of base pairing and complementarity, both the nucleic acids (DNA and RNA) have the ability to direct their duplications. The other molecules in the living system, such as proteins fail to fulfill first criteria itself.

The genetic material should be stable enough not to change with different stages of life cycle, age or with change in physiology of the organism. Stability as one of the properties of genetic material was very evident in Griffith's 'transforming principle' itself that heat, which killed the bacteria, at least did not destroy some of the properties of genetic material. This now can easily be explained in light of the DNA that the two strands being complementary if separated by heating come together, when appropriate conditions are provided. Further, 2′-OH group present at every nucleotide in RNA is a reactive group and makes RNA labile and easily degradable. RNA is also now known to be catalytic, hence reactive. Therefore, DNA chemically is less reactive and structurally more stable when compared to RNA. Therefore, among the two nucleic acids, the DNA is a better genetic material.

In fact, the presence of thymine at the place of uracil also confers additional stability to DNA. (Detailed discussion about this requires understanding of the process of repair in DNA, and you will study these processes in higher classes.)

Both DNA and RNA are able to mutate. In fact, RNA being unstable, mutate at a faster rate. Consequently, viruses having RNA genome and having shorter life span mutate and evolve faster.

RNA can directly code for the synthesis of proteins, hence can easily express the characters. DNA, however, is dependent on RNA for synthesis of proteins. The protein synthesising machinery has evolved around RNA. The above discussion indicate that both RNA and DNA can function as

genetic material, but DNA being more stable is preferred for storage of genetic information. For the transmission of genetic information, RNA is better.

6.3 RNA World

From foregoing discussion, an immediate question becomes evident – which is the first genetic material? It shall be discussed in detail in the chapter on chemical evolution, but briefly, we shall highlight some of the facts and points.

RNA was the first genetic material. There is now enough evidence to suggest that essential life processes (such as metabolism, translation, splicing, etc.), evolved around RNA. RNA used to act as a genetic material as well as a catalyst (there are some important biochemical reactions in living systems that are catalysed by RNA catalysts and not by protein enzymes). But, RNA being a catalyst was reactive and hence unstable. Therefore, DNA has evolved from RNA with chemical modifications that make it more stable. DNA being double stranded and having complementary strand further resists changes by evolving a process of repair.

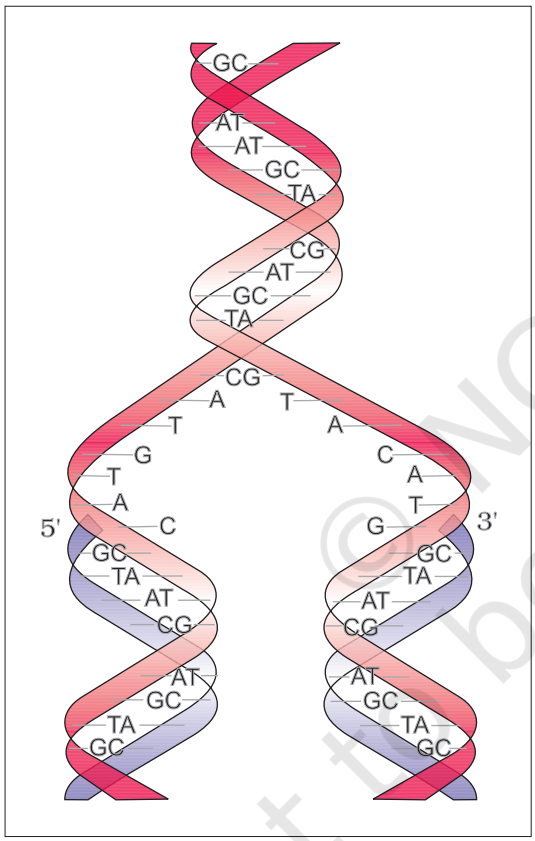


Figure 6.6 Watson-Crick model for semiconservative DNA replication

6.4 Replication

While proposing the double helical structure for DNA, Watson and Crick had immediately proposed a scheme for replication of DNA. To quote their original statement that is as follows:

"It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material" (Watson and Crick, 1953).

The scheme suggested that the two strands would separate and act as a template for the synthesis of new complementary strands. After the completion of replication, each DNA molecule would have one parental and one newly synthesised strand. This scheme was termed as **semiconservative** DNA replication (Figure 6.6).

6.4.1 The Experimental Proof

It is now proven that DNA replicates semiconservatively. It was shown first in *Escherichia coli* and subsequently in higher organisms, such as plants