

DNA (deoxyribonucleic acid) is double helical structure that was cracked by Watson and Crick based on the X-ray crystallography results. Each strand of a DNA helix is composed of repeating units of nucleotides. Nucleotide consists of 3 components: ribose or deoxyribose sugar, nitrogenous base (purines or pyrimidines) and phosphate.

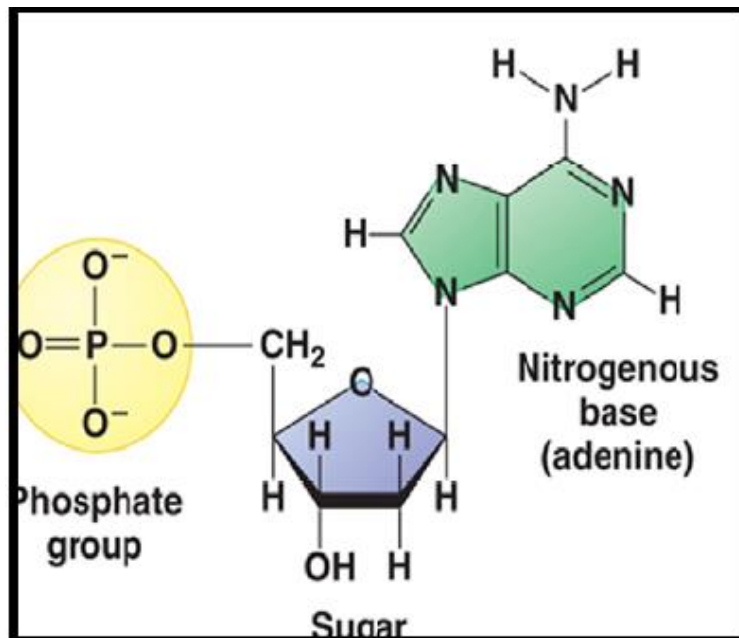


Fig.1. Structure of nucleotide

There are two types of purines- adenine and guanine. Pyrimidines are of three types- thymine, cytosine, and uracil. All nucleotides are common in DNA and RNA. But uracil is found in RNA and thymine is present only in DNA. DNA is negatively charged due to the presence of negatively charged phosphate groups. A nitrogenous base is linked to the pentose sugar through the N-glycosidic linkage. Two nucleotides are linked through 3'-5' phosphodiester linkage to form a dinucleotide. A polymer formed in such a manner has a free phosphate group at 5'-end of ribose sugar, which is referred to as 5'-end of polynucleotide chain. The other end of the polymer has a free 3'-OH group of the ribose sugar. This is referred to as 3'-end of the polynucleotide chain. The bonding between sugar and phosphates forms the backbone of a polynucleotide chain. The nitrogenous bases are linked to sugar moiety and project from the backbone.

The salient feature of the double helix structure of DNA are as follows-

- Two polynucleotides chains wrap around each other, where the backbone is constituted by sugar-phosphate, and bases project inside.
- The two DNA chains are antiparallel to each other. It means, if one chain has the polarity 5'-3', the other has 3'-5'.

- The two strands are coiled in right handed pattern.
- 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.

Packaging of DNA helix

Positively charged basic proteins that surround the DNA is known as histones. Histones are rich in basic amino acids such as lysine and arginine. Histones are organized to form a unit of eight molecules called as histone octamer. The DNA is negatively charged and is packaged by wrapping around the positively charged histone octamer. This forms a structure called nucleosome. A nucleosome typically contains 200 base pairs of DNA helix. Nucleosomes form the repeating unit of a structure called chromatin in nucleus. Chromatin are thread-like stained bodies seen in nucleus. The nucleosomes in chromatin appear as '**beads-on-string**' structure when viewed under electron microscope (EM). 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. At higher levels chromatin packaging requires additional proteins. These proteins are the **Non-Histone Chromosomal (NHC) proteins**. In a typical nucleus, loosely packed region of chromatin stains light and are referred to as **euchromatin**. The densely chromatin stains dark are called as **Heterochromatin**. Euchromatin is said to be transcriptionally active chromatin, whereas heterochromatin is inactive.

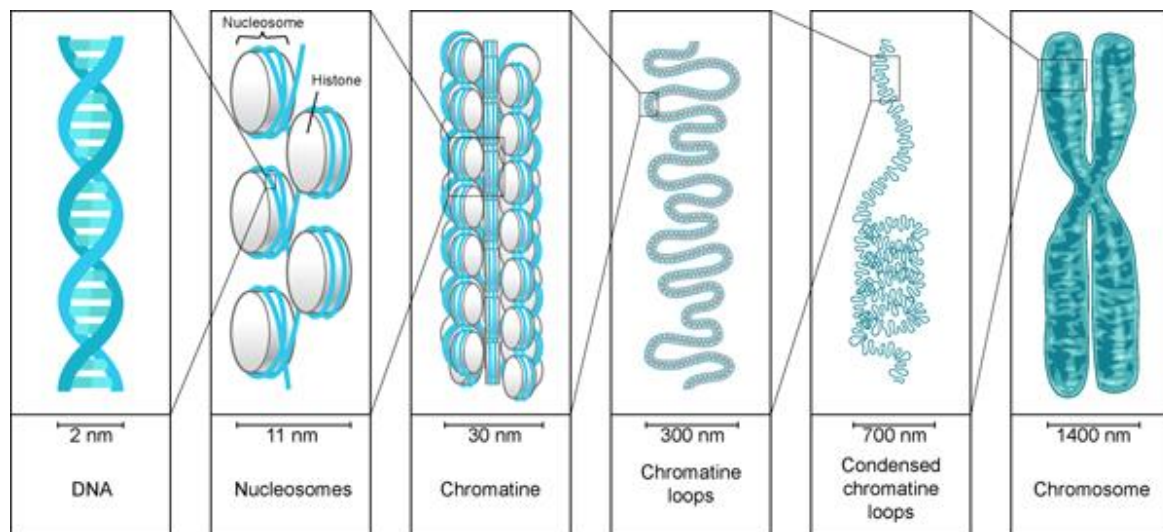


Fig.2. Packaging of DNA helix

DNA as a genetic material

Griffith performed an experiment known as transforming experiment. He used two strains of *Pneumococcus*. These two different strains were used to infect the mice. The two strains used were type III-S (smooth), that contains outer capsule made up of polysaccharide and type II-R (rough) strain do not contain capsule. The capsule protects the bacteria from the host immune system.

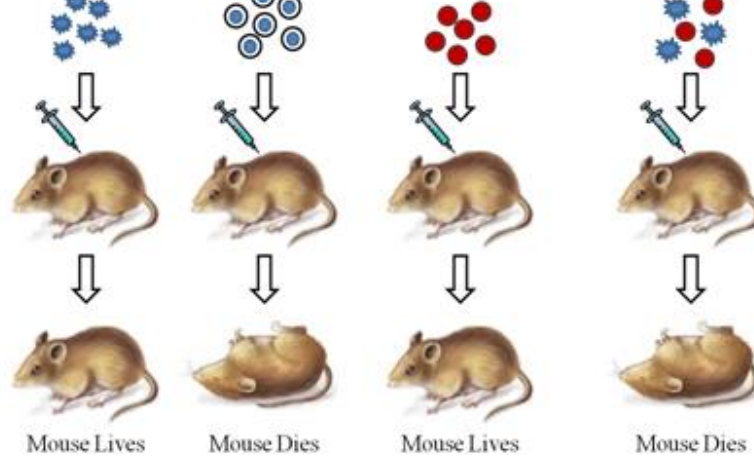


Fig.3. Griffith experiment

The Griffith experiment is explained below-

- Rough strain of *Pneumococcus* is injected in mouse. The mouse is alive.
- Smooth strain of *Pneumococcus* is injected in mouse. The mouse dies.
- When heat killed smooth strain of *Pneumococcus* is injected into mouse, the mouse is alive.
- In the last set of experiment, rough strain and heat killed smooth strain is injected into mouse. The mouse dies.

This proves that there is some transforming substance present in heat killed S strain that is converting or transforming the rough strain into virulent strain that is responsible for the death of the mouse. This transforming substance was later found to be DNA.

The genetic material is DNA

Alfred Hershey and Martha Chase (1952) performed an experiment to prove that DNA is the genetic material. They worked on bacteriophages, which are viruses that infect bacteria. When the bacteriophage attaches to the bacteria, its genetic material enters into the bacterial cell. The viral genetic material uses the bacterial cell to synthesize more viral particles. Hershey and Chase grew some viruses on a medium that contained radioactive phosphorus and other medium that contained radioactive sulfur. When radioactive phosphorus was present in the medium the viruses contained radioactive DNA but not radioactive protein. This is because DNA contains phosphorus, but protein does not. Similarly, when the growth medium contained radioactive sulfur the viruses contained radioactive protein but not radioactive DNA. This is because DNA does not contain sulfur.

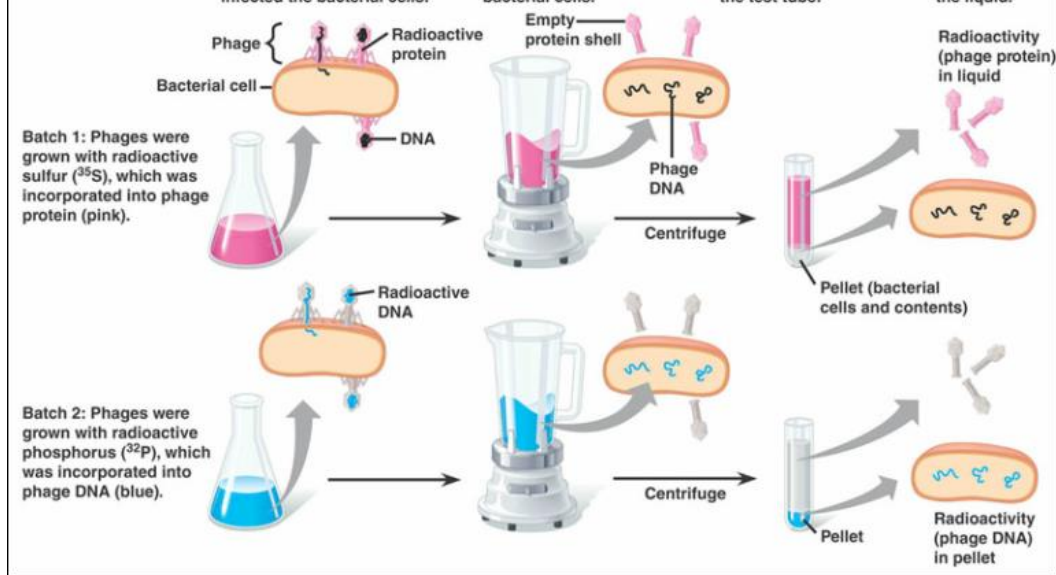


Fig.4. Hershey and chase experiment

Those bacteria were found to be radioactive only when they were infected with viruses that had radioactive DNA. This indicates that DNA was the material that passed from the virus to the bacteria. Bacteria infected with viruses containing 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. This experiment proves that DNA is the genetic material.

Central dogma of molecular biology

It is an explanation about how the flow of genetic information occurs in a biological system. This explains how DNA replicates and then gets converted into messenger RNA (mRNA) via transcription. Then this mRNA is translated to form proteins.

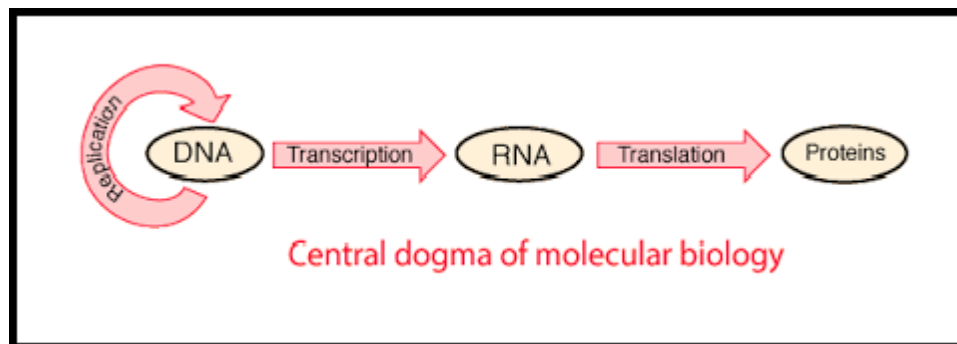


Fig.5. Central dogma of molecular biology

DNA replication

DNA replication is a process of producing two identical copies of DNA from a single DNA molecule. It is a process of biological inheritance. DNA is a double helix in which two strands are complementary to each other. These two strands of a helix separate at the time of replication to form two new DNA molecules. Out of the two strands of DNA formed, one is identical to one of the strand and another strand is complementary to the parent strand. This form of replication is known as **semi-conservative replication**. Before the cell enters the mitosis, the DNA is replicated in S phase of interphase. DNA polymerase is the most important enzyme involved in DNA replication. DNA replication is an energy dependent process. During the process of replication, the two DNA strands do not separate completely, the replication occurs within the small opening of the DNA helix known as **replication fork**. The DNA polymerase catalyzes the reaction in 5' to 3'. Consequently, on one strand (the template with polarity 3'-5'), the replication is continuous, while on the other (the template with polarity 5'-3'), it is discontinuous. The enzyme DNA ligase later joins the discontinuously synthesized fragments. The strand which is synthesized continuously is known as **leading strand** whereas the one synthesized discontinuously is known as **lagging strand**. Replication begins on the specific site on the DNA, known as **origin of replication**.

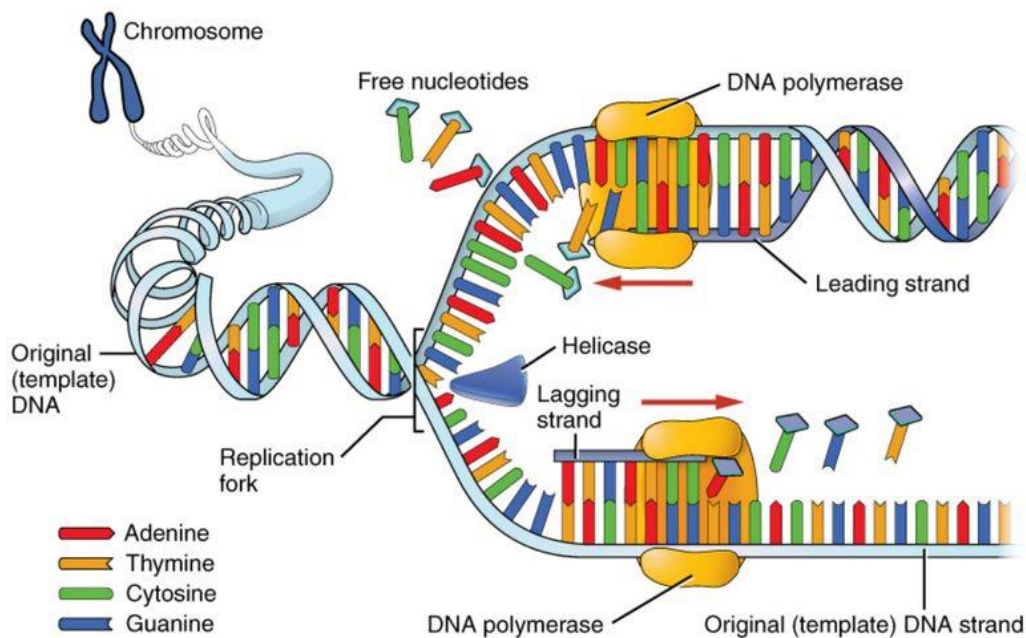


Fig.6. DNA replication

Transcription

It is a process of formation of RNA such as messenger RNA from DNA before gene expression or protein synthesis occurs. During transcription, one of the strands of DNA acts as a template for mRNA formation. The synthesis of mRNA occurs via RNA polymerase enzyme. Transcription usually occurs for a particular DNA segment which is required further for gene

structural gene and a terminator. DNA dependent RNA polymerase catalyzes the polymerization in 5'-3' direction. Promoter is the region where RNA polymerase binds. Terminator defines the end of transcription.

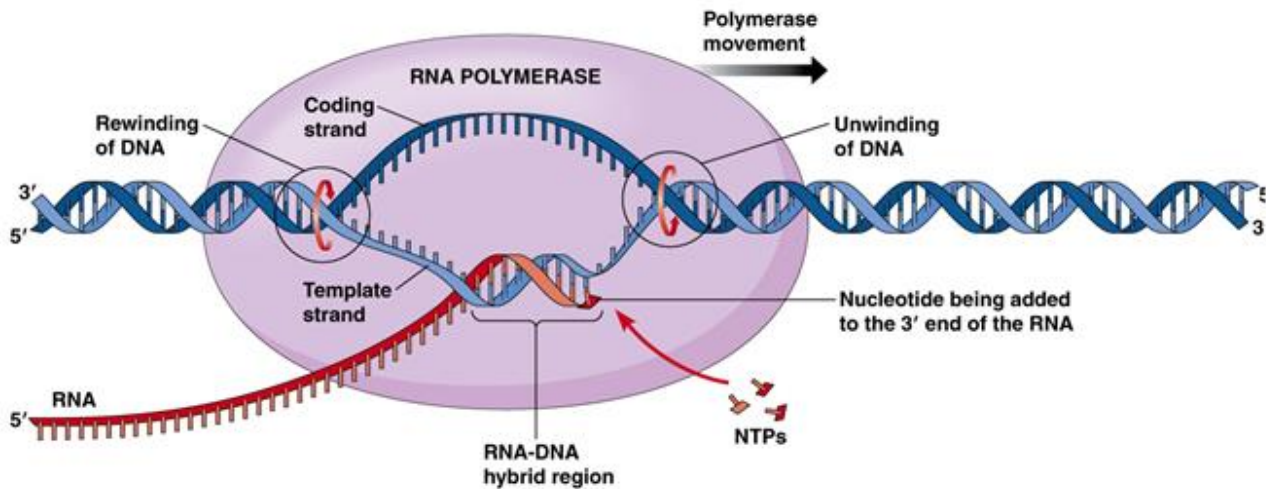


Fig.7. Process of transcription

Transcription consists of three steps- **initiation, elongation and termination.**

Initiation involves the binding of the RNA polymerase to promoter. Single type of DNA dependent RNA polymerase catalyzes the transcription of all types of RNA in bacteria.

Elongation is the process of addition of nucleotides to form the RNA.

Termination factor helps in termination of transcription. The RNA synthesized after transcription is known as primary transcript. The primary transcript undergoes modification such as splicing, capping, tailing etc. Primary transcript consist of introns and exons. The removal of introns is known as **splicing**. Addition of poly-A tail at the 3'-end of the RNA is known as **tailing**. In **capping** an unusual nucleotide (methyl guanosine triphosphate) is added to the 5'-end of RNA.

Some viruses have a property of reverse transcription. They are able to convert RNA template into DNA. The enzyme used is known as reverse transcriptase.

“For example: Human immunodeficiency virus that causes AIDS”.

Translation

This is the process of gene expression or protein synthesis that occurs in cytosol. Ribosomes are the cell organelles that are involved in protein synthesis. The messenger RNA formed by the process of transcription is decoded by ribosomes to form a polypeptide made up of amino acids. Messenger RNA is composed of polymer of nucleotides or **codon**. Each codon

Transfer RNA is involved in physically linking mRNA and the amino acid sequence of proteins.

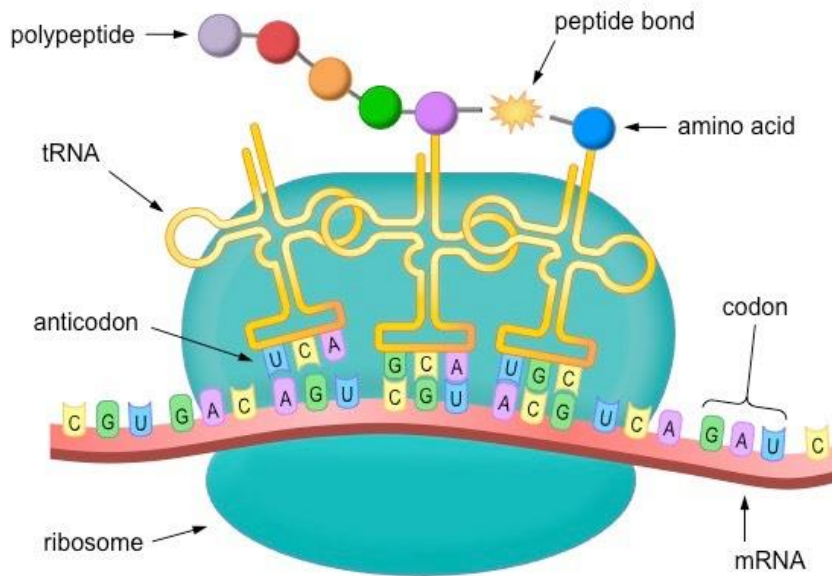


Fig.8. Translation set-up

It involves 4 main steps-

- Activation of amino acids- amino acids binds to specific tRNA molecule.
- Initiation of polypeptide synthesis- In capping an unusual nucleotide (methyl guanosine triphosphate) is added to the 5'-end of RNA
- Elongation of polypeptide synthesis- It involves the addition of amino acids to the growing polypeptide chains
- Termination of polypeptide synthesis- It involves the end of the translation of protein synthesis.

Genetic code

The set of rules by which information encoded in genetic material is translated into proteins in the living cells. Salient features of genetic code are as follows-

- The codon consists of three nucleotides. 61 codons code for 20 different amino acids.
- One codon codes for one amino acid.
- One amino acid can be coded by more than one codon.

2nd base	U	UUU Phenylalanine	UCU Serine	UAU Tyrosine	UGU Cysteine	U
		UUC Phenylalanine	UCC Serine	UAC Tyrosine	UGC Cysteine	C
		UUA Leucine	UCA Serine	UAA Stop	UGA Stop	A
		UUG Leucine	UCG Serine	UAG Stop	UGG Tryptophan	G
	C	CUU Leucine	CCU Proline	CAU Histidine	CGU Arginine	U
CUC Leucine		CCC Proline	CAC Histidine	CGC Arginine	C	
CUA Leucine		CCA Proline	CAA Glutamine	CGA Arginine	A	
CUG Leucine		CCG Proline	CAG Glutamine	CGG Arginine	G	
A	AUU Isoleucine	ACU Threonine	AAU Asparagine	AGU Serine	U	
	AUC Isoleucine	ACC Threonine	AAC Asparagine	AGC Serine	C	
	AUA Isoleucine	ACA Threonine	AAA Lysine	AGA Arginine	A	
	AUG Methionine (Start)	ACG Threonine	AAG Lysine	AGG Arginine	G	
G	GUU Valine	GCU Alanine	GAU Aspartic Acid	GGU Glycine	U	
	GUC Valine	GCC Alanine	GAC Aspartic Acid	GGC Glycine	C	
	GUA Valine	GCA Alanine	GAA Glutamic Acid	GGA Glycine	A	
	GUG Valine	GCG Alanine	GAG Glutamic Acid	GGG Glycine	G	
Nonpolar, aliphatic Polar, uncharged Aromatic Positively charged Negatively charged						

Fig.9. Genetic code

- The code is universal.

Regulation of gene expression

All the genes in the living cells is not active all the time. They become active when needed. Expression is controlled by genes are known as regulatory genes. Regulation in eukaryotes can occur at the following levels-

- Transcriptional level.
- Processing level.
- Transport of mRNA from nucleus to the cytoplasm.
- Translational level.

Lac operon or lactose operon

An operon consists of structural genes, operator genes, promoter genes, promoter genes, regulator genes, and repressor. Lac operon consist of lac Z, lac Y and lac A genes. Lac Z codes for galactosidase, lac Y codes for permease and lac A codes for transacetylase. When repressor molecules bind the operator, genes are not transcribed. When repressor does not bind the operator and instead inducer binds, transcription is switched on. In case of lac operon, lactose is an inducer. So, binding of lactose to operator, switch on the transcription.

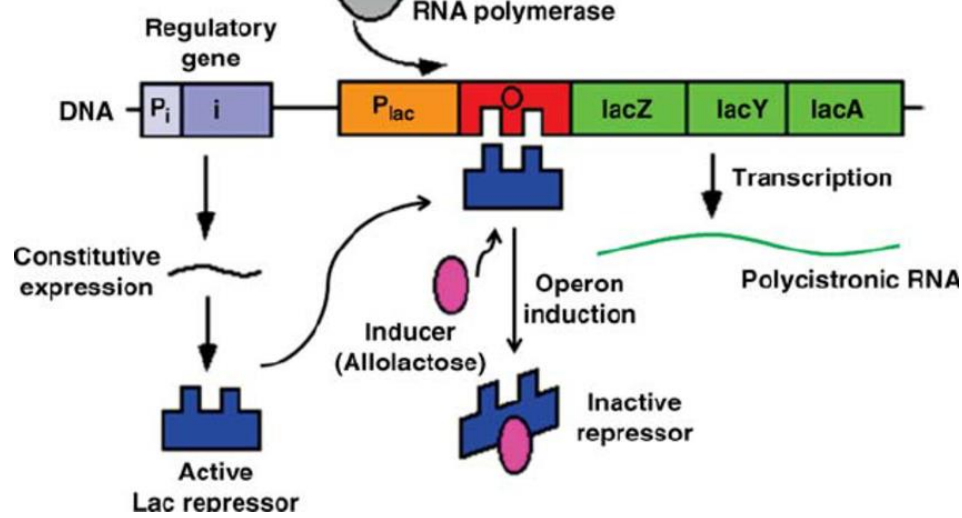


Fig.10. Lactose operon

Human genome project

The salient features of the human genome project are as follows-

- The human genome contains 3164.7 million nucleotide bases.
- The average gene consists of 3000 bases, but size varies.
- Humans are said to have about 30,000 genes.
- The functions are unknown for over 50 per cent of the discovered genes.
- Less than 2 per cent of the genome codes for proteins.
- Human genome consists of large portion of repeated sequences.
- Chromosome 1 has most genes (2968), and the Y has the fewest (231).