**Smt. A.S.M. College for Women, Ballari.**
**DEPARTMENT OF ZOOLOGY**
**B.Sc. III Semester**
**Molecular Biology, Bioinstrumentation & Techniques in Biology**
Theory Notes

**Syllabus:**

**Unit 1:**

**Chapter 1: Process of transcription**

- Fine structure of gene ( cistron, recon & muton )
- RNA Polymerases- types & function.
- Transcription in prokaryotes and eukaryotes

**Chapter 2: Process of translation**

1. Genetic code and its salient features.
2. Translation in prokaryotes and eukaryotes.

**Unit 2:**

**Chapter 3: Regulation of gene expression- I**

- Regulation of gene expression in prokaryotes – Lac operon concept ( inducible ) and trp operon ( repressible ) in E. coli
- Regulation of gene expression in eukaryotes – role of chromatin ( euchromatin and heterochromatin ) in gene expression
- Post transcriptional modification : capping, splicing , polyadenyalation
- Concept of RNA editing (mRNA), gene silencing and RNAi.

**Chapter 4: Regulation of gene expression- II**

- Post translational modification: purpose, advantages and significance, glycosylation, methylation, phosphorylation & acetylation.

## Chapter 1: Process of transcription

### Introduction:

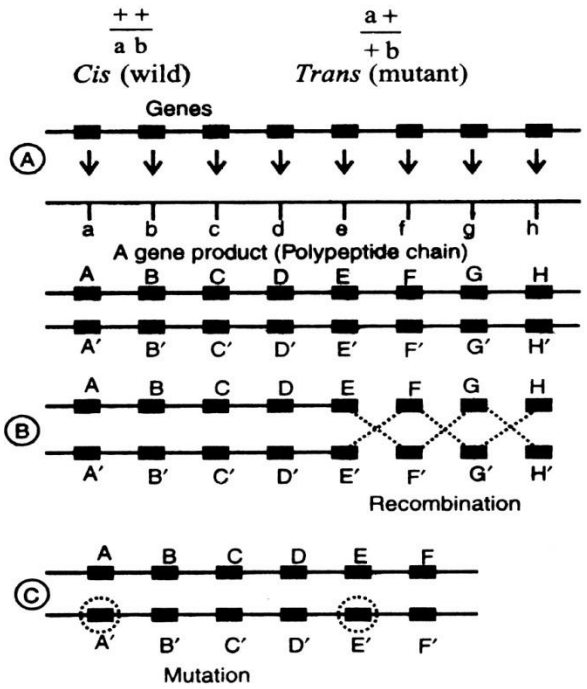
The genetic blueprint contained in the nucleotide sequence can determine the phenotype of an individual. The hereditary units, which are transmitted from one generation to the next generation, are called genes. A gene is a fundamental biological unit like atom which is the fundamental physical unit.

Mendel was the first scientist who proposed genes as particulate units and called them hereditary elements or factors. But the concept of gene has undergone a considerable change since Mendel's time.

### Fine Structure of Gene:

Benzer, in 1955, divided the gene into recon, muton and cistron which are the units of recombination, mutation and function within a gene. Several units of this type exist in a gene. In other words, each gene consists of several units of function, mutation and recombination. The fine structure of gene deals with mapping of individual gene locus.

On the basis of modern finding a gene can be defined as any part or segment of DNA that is assigned a particular function. This function may be the synthesis of protein, recombination or even mutation. Seymour Benzer (1957) has coined new terms to denote the relationship between DNA segments (genes) and its various functions.




- **Recon:** It is the smallest unit of DNA (gene) capable of undergoing crossing over and genetic recombination.
- **Muton:** It is the smallest unit of DNA (gene) which can undergo mutation.
- **Cistron:** It is the smallest unit of DNA (gene) which actually contains the genetic information. It contains information for synthesis of a polypeptide chain (protein or enzyme).

(Fig. A).

**A gene product (polypeptide chain)**: One-gene-one enzyme hypothesis of Beadle and Tatum was redefined by several workers in coming years. A single mRNA is transcribed by a single gene. Therefore, one-gene-one mRNA hypothesis was put forth. Exceptionally, a single mRNA is also transcribed by more than one gene and it is said to be polycistronic. Therefore, the concept has been given as one-gene-one protein hypothesis. The proteins are the polypeptide chain of amino acids translated by mRNA. Therefore, it has been correctly used as one-gene-one polypeptide hypothesis.

Moreover, genes are present within the genes and their cis-trans effect governs the function. Therefore, S. Benzer termed the functional gene as **cistron** (Fig. A). Crossing over within the

functional genes or cistron is possible. The cis and Trans arrangement of alleles may be written as above in fig A.

**Recombination:** In 1962, Benzer demonstrated that the crossing over or recombination occurs within a functional gene or cistron. In a cistron the recombinational units may be more than one. Thus, the smallest unit capable of undergoing recombination is known as **recon.**

**Mutation:** Benzer (1962) coined the term **muton** to denote the smallest unit of chromosome that undergoes mutational changes. Hence, muton may be defined as 'the smallest unit of DNA which may be changed in the nucleotide. Thus, changes at nucleotide level are possible (Fig. 2.6C). The smallest unit of muton is the nucleotide. Therefore, cistron is largest unit in size followed by recon and muton.

**RNA Polymerases- types & function**

**RNA Polymerases:** Ribonucleic Acid (RNA) Polymerase (RNAP) enzyme is a multi-subunit enzyme that applies its activity in the catalization of the transcription process of RNA synthesized from a DNA template.

- And therefore, RNA polymerase enzyme is responsible for the copying of DNA sequences into RNA sequences during transcription.
- The function of RNA polymerase is to control the process of transcription, through which copying of information stored in DNA into a new molecule of messenger RNA (mRNA.)
- During transcription, the RNA polymer is contemporary to the template DNA that is synthesized in the direction of 5′ to 3′.
- The enzyme RNA polymerase interacts with proteins to enable it to function in catalization of the synthesis of RNA.
- The collaborator proteins assist in enabling the specific binding of RNA polymerase, assist in the unwinding of the double chemical structure of DNA, moderate the enzymatic activities of RNA polymerase and to control the speed of transcription.
- The RNA polymerase enzyme has an interrupted mechanism whereby it continuously synthesizes RNA polymers of over four thousand bases per minute but they pause or stop occasionally to maintain fidelity.
- RNA polymerase is an enzyme that is responsible for copying a DNA sequence into an RNA sequence, during the process of transcription. As a complex molecule composed of protein subunits, RNA polymerase controls the process of transcription, during which the information stored in a molecule of DNA is copied into a new molecule of messenger RNA.
- RNA polymerases have been found in all species, but the number and composition of these proteins vary across taxa.
- For instance, bacteria contain a single type of RNA polymerase, while eukaryotes (multicellular organisms and yeasts) contain three distinct types.
- In spite of these differences, there are striking similarities among transcriptional mechanisms.
- For example, all species require a mechanism by which transcription can be regulated in order to achieve spatial and temporal changes in gene expression.

**Types of RNA polymerase**

Prokaryotic (Bacteria, viruses, archaea) organisms have a single type of RNA polymerase that synthesizes all the subtypes of RNA, while eukaryotes (multicellular organisms) have 5 different types of RNA polymerases which perform different functions in the synthesis of different RNA molecules.

**Prokaryotic RNA polymerase:**

- The prokaryotes have a single type of RNA polymerase (RNAP) which synthesizes all the classes of RNA, i.e. mRNA, tRNA, rRNA, sRNA.
- The RNA Polymerase molecule is made up of 2 domains and 5 subunits:
    1. Core and holoenzyme
    2. Subunits (β, β', α (αI and αII), ω,)
- The promoter is the sequence of DNA that is required for accurate and specific initiation of transcription, and also, it is the sequence of DNA to which RNA polymerase binds accurately to initiate transcription.
- The 'a' subunit is made up of two distinct domains. The N-terminal domain (a-NTD) and the C-terminal.
- The N-terminal is involved in dimerization forming a2 and further assembly of the RNA polymerase.
- The C-terminal domain functions such as binding to the Upstream Promoter (UP) DNA sequence at promoters for rRNA and tRNA genes and in communication with several transcriptional activators.
- Each of the subunit structure is as follows:
- Prokaryotic RNA Polymerase Subunits

| Subunit | Function |
|---|---|
| β | The β' + β form the catalytic center, responsible for RNA synthesis. |
| β' | The β' + β form the catalytic center, responsible for RNA synthesis. |
| α (αI and αII) | It is made up of the enzyme assembly, and it also binds the UP sequence in the promoter. |
| ω | It confers specificity for promoter; and binds to -10 and -35 sites in the promoter. |

**Eukaryotic RNA polymerase**

- There are 5 known types of RNA polymerases each responsible for the synthesis of specific subtypes of RNA. These include:
- RNA polymerase I that synthesizes a pre-rRNA 45S (35S in yeast), which matures and forms the major RNA sections of the ribosome.
- RNA polymerase II synthesizes precursors of mRNAs and most snRNA and microRNAs.
- RNA polymerase III synthesizes tRNAs, rRNA 5S, and other small RNAs found in the nucleus and cytosol.
- RNA polymerase IV and V found in plants are not well understood, however, they make siRNA. The plant chloroplast encodes the ssRNAPs and uses bacteria-like RNA Polymerase.
- Each of the nuclear RNA polymerases is a large protein molecule with about 8 to 14 subunits and the molecular weight is approximately 500,000 for each.
- They commonly have 3 subunits, a, b and b'. The largest subunits being b and b'.
- These subunits are used as catalytic promoters and for assembly of proteins.
- Each of these polymerases has a different function:

**RNA polymerase I**

- This enzyme is located in the nucleolus of the cell.
- It is a specialized nuclear substructure where the ribosomal RNA (rRNA) is synthesized by transcription and assembled into ribosomes.
- The rRNA is component elements of the ribosomes and is important in the process of translation.
- Therefore, RNA polymerase I synthesize almost all rRNAs except 5S rRNA.
- In yeast, the enzyme has a mass of 600kDa and 13 subunits.

**RNA polymerase II**

- This enzyme is located in the nucleus.
- Most organisms that possess RNA polymerase II have a 12-subunit RNAP II (with a mass of about 550 kDa)
- It is structurally made up of holoenzyme and mediators, with General Transcriptional factors (GTFs).
- They contain transcription factors and transcriptional regulators.
- It functions by synthesizing all proteins that code for the nuclear pre-mRNAs in eukaryotic cells (mRNAs in prokaryotic cells).
- It is responsible for transcribing most of the eukaryotic genes and especially found in human genes.

**RNA polymerase III**

- It is located in the nucleus.
- The RNA polymerase III has 14 or more distinct subunits with a mass of approximately 700 kDa.
- Its function is to transcribe transfer RNA (tRNA), ribosomal RNA (rRNA), and other small RNAs.
- Some of its target points are important for the normal functioning of the cell

**RNA polymerases IV and V**

- They are exclusively found in plants, and they perform combined action in the formation of small interfering RNA and heterochromatin in the cell nucleus.
- In Plants, the RNA polymerase is found in the chloroplast (plastids) and mitochondria, encoded by the mitochondrial DNA.
- These enzymes are much more related to bacterial RNA polymerase than to the nuclear RNA polymerase.
- Their function is to catalyze specific transcription of organelle genes.

**Functions of RNA Polymerase**

- Generally, the RNA molecule is a messenger molecule that is used to export information that is coded in DNA out of the cell nucleus, to synthesize proteins in the cell cytoplasm.
- RNA polymerase is used in the production of molecules that play a wide range of roles, of which one of its functions is to regulate the number and type of RNA transcript that is formed in response to the requirements of the cell.
- The RNA polymerase enzyme interacts with different molecular proteins, transcription factors, and signaling molecules on the carboxyl-terminal, which regulates its mechanisms, which play a major role in gene expression and gene specialization in multicellular (eukaryotic) organisms.
- The RNA enzyme also ensures irregularities and errors during the conversion of DNA to RNA (transcription). Such as ensuring that the right nucleotide is added to the newly synthesized RNA strand, inserting the right amino acid-base which is complementary to the template of the DNA strand.
- When the right nucleotides have been inserted, the RNA polymerase can then catalyze and elongate the RNA strand, at the same time, proofread the new strand and remove incorrect bases.
- RNA polymerase is also involved in the post-transcription modification of RNAs, converting them into functional molecules that facilitate the transportation of molecules from the nucleus to their site of action.
- Besides its role in the synthesis of proteins, RNA performs other functions such as
    - Protein coding
    - Regulation of gene expression
    - Act as enzymes
    - Formation of gametes by the non-coding RNA (ncRNA)
    - Production of regulatory molecules.

**Transcription in prokaryotes**

In both prokaryotes and eukaryotes, the second function of DNA (the first was replication) is to provide the information needed to construct the proteins necessary so that the cell can perform all of its functions. To do this, the DNA is “read” or transcribed into an mRNA molecule. The mRNA then provides the code to form a protein by a process called translation. Through the processes of transcription and translation, a

protein is built with a specific sequence of amino acids that was originally encoded in the DNA. This module discusses the details of transcription.

**The Central Dogma**: DNA Encodes RNA; RNA Encodes Protein

The flow of genetic information in cells from DNA to mRNA to protein is described by the central dogma (Figure 9.14), which states that genes specify the sequences of mRNAs, which in turn specify the sequences of proteins.

A flow chart shows DNA, with an arrow to RNA, which has an arrow to protein.
The central dogma states that DNA encodes RNA, which in turn encodes protein.

The copying of DNA to mRNA is relatively straightforward, with one nucleotide being added to the mRNA strand for every complementary nucleotide read in the DNA strand. The translation to protein is more complex because groups of three mRNA nucleotides correspond to one amino acid of the protein sequence. However, as we shall see in the next module, the translation to protein is still systematic, such that nucleotides 1 to 3 correspond to amino acid 1, nucleotides 4 to 6 correspond to amino acid 2, and so on.

**Transcription: from DNA to mRNA**

Both prokaryotes and eukaryotes perform fundamentally the same process of transcription, with the important difference of the membrane-bound nucleus in eukaryotes. With the genes bound in the nucleus, transcription occurs in the nucleus of the cell and the mRNA transcript must be transported to the cytoplasm. The prokaryotes, which include bacteria and Achaea, lack membrane-bound nuclei and other organelles, and transcription occurs in the cytoplasm of the cell. In both prokaryotes and eukaryotes, transcription occurs in three main stages: initiation, elongation, and termination.

**Initiation**

Transcription requires the DNA double helix to partially unwind in the region of mRNA synthesis. The region of unwinding is called a transcription bubble. The DNA sequence onto which the proteins and enzymes involved in transcription bind to initiate the process is called a promoter. In most cases, promoters exist upstream of the genes they regulate. The specific sequence of a promoter is very important because it determines whether the corresponding gene is transcribed all of the time, some of the time, or hardly at all (Figure 9.15).

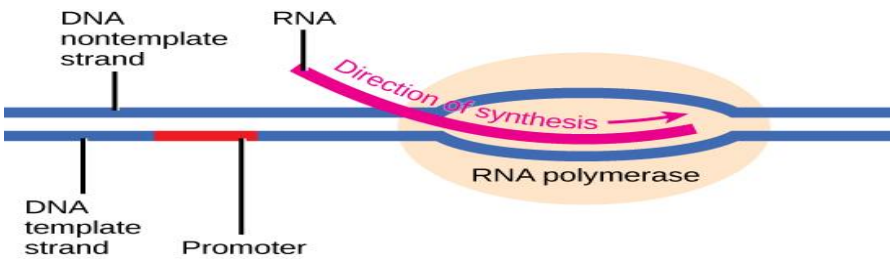



Illustration shows a template strand and non template strand of DNA, with a promoter section in red on the template strand. Downstream of the promoter is an RNA polymerase where RNA is being synthesized.

Figure 9.15 The initiation of transcription begins when DNA is unwound, forming a transcription bubble. Enzymes and other proteins involved in transcription bind at the promoter.

**Elongation**

Transcription always proceeds from one of the two DNA strands, which is called the template strand. The mRNA product is complementary to the template strand and is almost identical to the other DNA strand, called the non template strand, with the exception that RNA contains a uracil (U) in place of the thymine (T) found in DNA. During elongation, an enzyme called RNA polymerase proceeds along the DNA template adding nucleotides by base pairing with the DNA template in a manner similar to DNA replication, with the difference that an RNA strand is being synthesized that does not remain bound to the DNA template. As elongation proceeds, the DNA is continuously unwound ahead of the core enzyme and rewound behind it (Figure 9.16).

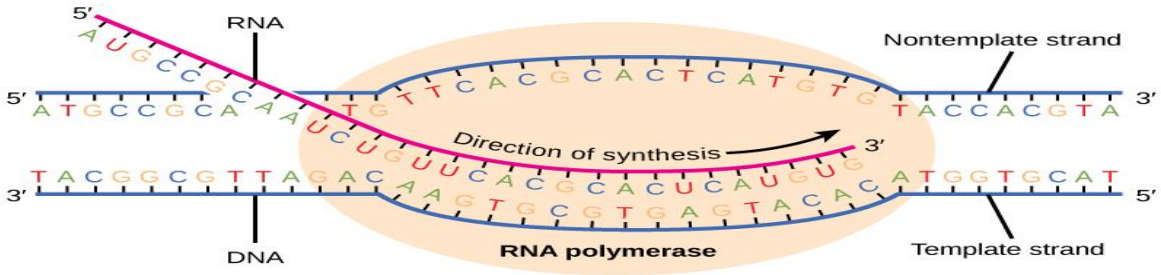



Illustration shows RNA synthesis by RNA polymerase. The RNA strand is synthesized in the 5' to 3' direction.

Figure 9.16 During elongation, RNA polymerase tracks along the DNA template, synthesizes mRNA in the 5′ to 3′ direction, and unwinds then rewinds the DNA as it is read.

**Termination**

Once a gene is transcribed, the prokaryotic polymerase needs to be instructed to dissociate from the DNA template and liberate the newly made mRNA. Depending on the gene being transcribed, there are two kinds of termination signals, but both involve repeated nucleotide sequences in the DNA template that result in RNA polymerase stalling, leaving the DNA template, and freeing the mRNA transcript.

On termination, the process of transcription is complete. In a prokaryotic cell, by the time termination occurs, the transcript would already have been used to partially synthesize numerous copies of the encoded protein because these processes can occur concurrently using multiple ribosomes (polyribosomes) (Figure 9.17). In contrast, the presence of a nucleus in eukaryotic cells precludes simultaneous transcription and translation.

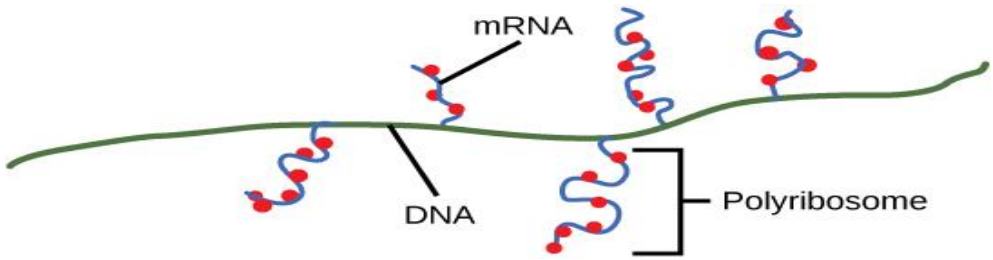


Illustration shows multiple mRNAs being transcribed off one gene. Ribosomes attach to the mRNA before transcription is done and begin making protein.

Figure 9.17 Multiple polymerases can transcribe a single bacterial gene while numerous ribosomes concurrently translate the mRNA transcripts into polypeptides. In this way, a specific protein can rapidly reach a high concentration in the bacterial cell.

**Transcription in Eukaryotes**

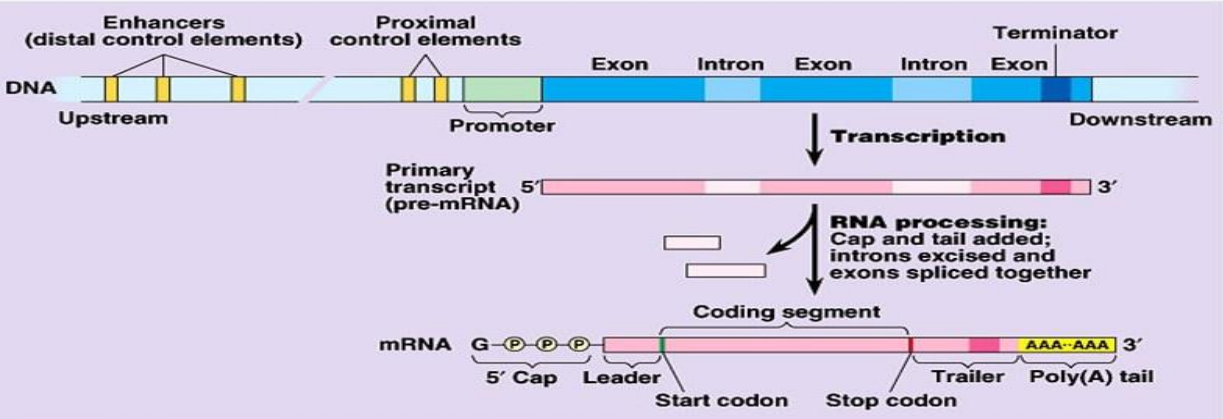
Transcription occurs in eukaryotes in a way that is similar to prokaryotes with reference to the basic steps involved. However, some major differences between them include:

Initiation is more complex.

Termination does not involve stem-loop structures.

Transcription is carried out by three enzymes (RNA polymerases I, II and III).

The regulation of transcription is more extensive than prokaryotes.




Enzymes involved in eukaryotic transcription.

Unlike prokaryotes where all RNA is synthesized by a single RNA polymerase, the nucleus of a eukaryotic cell has three RNA polymerases responsible for transcribing different types of RNA.

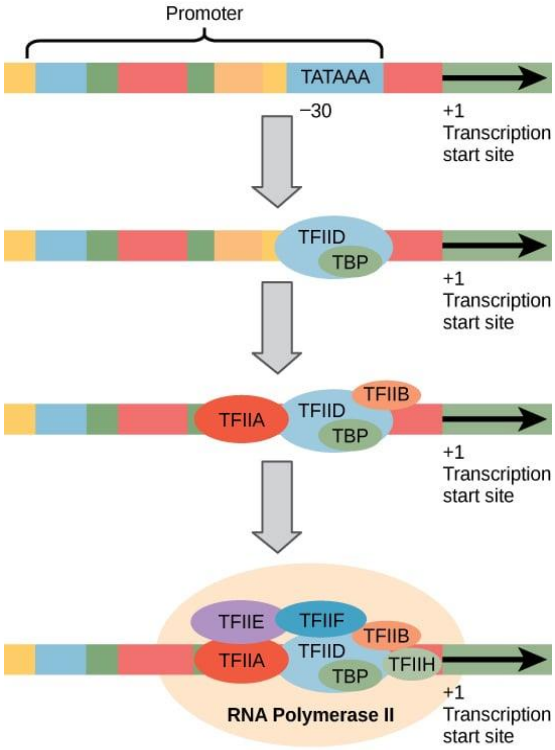
- **RNA polymerase I (RNA Pol I)** is located in the nucleolus and transcribes the 28S, 18S, and 5.8S rRNA genes.
- **RNA polymerase II (RNA Pol II)** is located in the nucleoplasm and transcribes protein-coding genes, to yield pre-mRNA, and also the genes encoding small nucleolar RNAs (snoRNAs) involved in rRNA processing and small nuclear RNAs (snRNAs) involved in mRNA processing, except for U6 snRNA.
- **RNA polymerase III (RNA Pol III)** is also located in the nucleoplasm. It transcribes the genes for tRNA, 5S rRNA, U6 snRNA, and the 7S RNA associated with the signal recognition particle (SRP) involved in the translocation of proteins across the endoplasmic reticulum membrane.
- Each of the three eukaryotic RNA polymerases contains 12 or more subunits and so these are large complex enzymes.
- The genes encoding some of the subunits of each eukaryotic enzyme show DNA sequence similarities to genes encoding subunits of the core enzyme of *E. coli* RNA polymerase.

- However, four to seven other subunits of each eukaryotic RNA polymerase are unique in that they show no similarity either with bacterial RNA polymerase subunits or with the subunits of other eukaryotic RNA polymerases.

The basic mechanism of RNA synthesis by these eukaryotic RNA polymerases can be divided into the following phases:

**Initiation Phase**

During initiation, RNA polymerase recognizes a specific site on the DNA, upstream from the gene that will be transcribed, called a **promoter site** and then unwinds the DNA locally.




- Most promoter sites for RNA polymerase II include a highly conserved sequence located about 25–35 bp upstream (i.e. to the 5 side) of the start site which has the consensus TATA (A/T) A (A/T) and is called the TATA box.
- Since the start site is denoted as position +1, the TATA box position is said to be located at about position -25.
- The TATA box sequence resembles the -10 sequence in prokaryotes (TATAAT) except that it is located further upstream.
- Both elements have essentially the same function, namely recognition by the RNA polymerase in order to position the enzyme at the correct location to initiate transcription.
- The sequence around the TATA box is also important in that it influences the efficiency of initiation. Transcription is also regulated by upstream control elements that lie 5′ to the TATA box.
- Some eukaryotic protein-coding genes lack a TATA box and have an initiator element instead, centered around the transcriptional initiation site.
- In order to initiate transcription, RNA polymerase II requires the assistance of several other proteins or protein complexes, called general (or basal) transcription factors, which must assemble into a complex on the promoter in order for RNA polymerase to bind and start transcription.
- These all have the generic name of TFII (for Transcription Factor for RNA polymerase II).
- The first event in initiation is the binding of the transcription factor IID (TFIID) protein complex to the TATA box via one its subunits called TBP (TATA box binding protein).
- As soon as the TFIID complex has bound, TFIIA binds and stabilizes the TFIID-TATA box interaction. Next, TFIIB binds to TFIID.
- However, TFIIB can also bind to RNA polymerase II and so acts as a bridging protein. Thus,
- RNA polymerase II, which has already complexed with TFIIF, now binds.

- This is followed by the binding of TFIIE and H. This final protein complex contains at least 40 polypeptides and is called the **transcription initiation complex.**
- Those protein-coding genes that have an initiator element instead of a TATA box appear to need another protein(s) that binds to the initiator element.
- The other transcription factors then bind to form the transcription initiation complex in a similar manner to that described above for genes possessing a TATA box promoter.

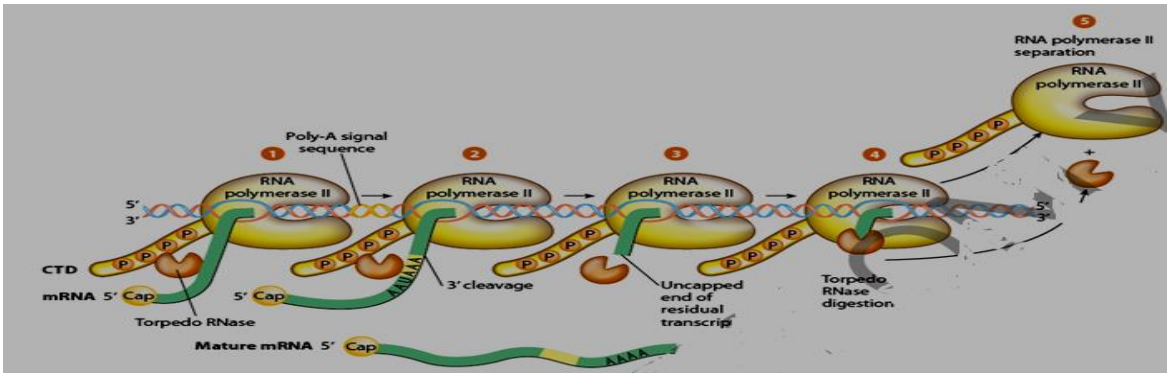
**Elongation Phase**

TFIIH has two functions:

1. It is a helicase, which means that it can use ATP to unwind the DNA helix, allowing transcription to begin.
2. In addition, it phosphorylates RNA polymerase II which causes this enzyme to change its conformation and dissociate from other proteins in the initiation complex.

- The key phosphorylation occurs on a long C-terminal tail called the C-terminal domain (CTD) of the RNA polymerase II molecule.
- Interestingly, only RNA polymerase II that has a non-phosphorylated CTD can initiate transcription but only an RNA polymerase II with a phosphorylated CTD can elongate RNA.
- RNA polymerase II now starts moving along the DNA template, synthesizing RNA, that is, the process enters the elongation phase.
- RNA synthesis occurs in the 5' → 3' direction with the RNA polymerase catalyzing a nucleophilic attack by the 3-OH of the growing RNA chain on the alpha-phosphorus atom on an incoming ribonucleoside 5-triphosphate.
- The RNA molecule made from a protein-coding gene by RNA polymerase II is called a primary transcript.

**Termination Phase**




- Elongation of the RNA chain continues until termination occurs.
- Unlike RNA polymerase in prokaryotes, RNA polymerase II does not terminate transcription at a specific site but rather transcription can stop at varying distances downstream of the gene.
- RNA genes transcribed by RNA Polymerase II lack any specific signals or sequences that direct RNA Polymerase II to terminate at specific locations.
- RNA Polymerase II can continue to transcribe RNA anywhere from a few bp to thousands of bp past the actual end of the gene.
- The transcript is cleaved at an internal site before RNA Polymerase II finishes transcribing. This releases the upstream portion of the transcript, which will serve as the

initial RNA prior to further processing (the pre-mRNA in the case of protein-encoding genes.)

- This cleavage site is considered the "end" of the gene. The remainder of the transcript is digested by a 5′-exonuclease (called Xrn2 in humans) while it is still being transcribed by the RNA Polymerase II.
- When the 5′-exonulease "catches up" to RNA Polymerase II by digesting away all the overhanging RNA, it helps disengage the polymerase from its DNA template strand, finally terminating that round of transcription.

**RNA processing**

The primary eukaryotic mRNA transcript is much longer and localized into the nucleus, when it is also called heterogeneous nuclear RNA (hnRNA) or pre- mRNA.

It undergoes various processing steps to change into a mature RNA:

**Cleavage**

- Larger RNA precursors are cleaved to form smaller RNAs.
- Primary transcript is cleaved by ribonuclease-P (an RNA enzyme) to form 5-7 tRNA precursors.

**Capping and Tailing**

- Initially at the 5′ end a cap (consisting of 7-methyl guanosine or 7 mG) and a tail of poly A at the 3′ end are added.
- The cap is a chemically modified molecule of guanosine triphosphate (GTP).

**Splicing**

- The eukaryotic primary mRNAs are made up of two types of segments; non-coding introns and the coding exons.
- The introns are removed by a process called RNA splicing where ATP is used to cut the RNA, releasing the introns and joining two adjacent exons to produce mature mRNA.

**Chapter 2: Process of translation**

**Genetic code and its salient features.**

**Genetic code:** The genetic code can be defined as the set of certain rules using which the living cells translate the information encoded within genetic material (DNA or mRNA sequences). The ribosomes are responsible to accomplish the process of translation. They link the amino acids in an mRNA-specified (messenger RNA) order using tRNA (transfer RNA) molecules to carry amino acids and to read the mRNA three nucleotides at a time.

Genetic Code Table

The complete set of relationships among amino acids and codons is said to be a genetic code which is often summarized in a table.

| | U | | C | | A | | G | | |
|---|---|---|---|---|---|---|---|---|---|
| U | UUU | Phe | UCU | Ser | UAU | Tyr | UGU | Cys | U |
| | UUC | Phe | UCC | Ser | UAC | Tyr | UGC | Cys | C |
| | UUA | Leu | UCA | Ser | UAA | Stop | UGA | Stop | A |
| | UUG | Leu | UCG | Ser | UAG | Stop | UGG | Trp | G |
| C | CUU | Leu | CCU | Pro | CAU | His | CGU | Arg | U |
| | CUC | Leu | CCC | Pro | CAC | His | CGC | Arg | C |
| | CUA | Leu | CCA | Pro | CAA | Gln | CGA | Arg | A |
| | CUG | Leu | CCG | Pro | CAG | Gln | CGG | Arg | G |
| A | AUU | Ile | ACU | Thr | AAU | Asn | AGU | Ser | U |
| | AUC | Ile | ACC | Thr | AAC | Asn | AGC | Ser | C |
| | AUA | Ile | ACA | Thr | AAA | Lys | AGA | Arg | A |
| | AUG | Met | ACG | Thr | AAG | Lys | AGG | Arg | G |
| G | GUU | Val | GCU | Ala | GAU | Asp | GGU | Gly | U |
| | GUC | Val | GCC | Ala | GAC | Asp | GGC | Gly | C |
| | GUA | Val | GCA | Ala | GAA | Glu | GGA | Gly | A |
| | GUG | Val | GCG | Ala | GAG | Glu | GGG | Gly | G |

**Properties of genetic code are explained below.**

1. **Triplet code:** A codon or a code word is defined as a group of bases that specify an amino acid. There is strong evidence, which proves that a sequence of three nucleotides codes for an amino acid in the protein, i.e., the code is a ***triplet***.
   The four bases of nucleotide i.e., (A, G, C, and U) are used to produce three-base codons. The 64 codons involve sense codons (that specify amino acids). Hence, there are 64 codons for 20 amino acids since every codon for one amino acid means that there exist more than code for the same amino acid.
2. **Commaless code**: No room for punctuation in between which indicates that every codon is adjacent to the previous one without any nucleotides between them.
3. **Nonoverlapping code:** The code is read sequentially in a group of three and a nucleotide which becomes a part of triplet never becomes part of the next triplet.

   ***For example***
   5'-UCU-3' codes for Serine
   5'-AUG-3' codes for methionine

4. **Polarity:** Each triplet is read from 5' → 3' direction and the beginning base is 5' followed by the base in the middle then the last base which is 3'. This implies that the codons have a ***fixed polarity*** and if the codon is read in the reverse direction, the base sequence of the codon would reverse and would specify two different proteins.
5. **Degenerate code**: Every amino acid except tryptophan (UGG) and methionine (AUG) is coded by various codons, i.e., a few codons are synonyms and this aspect is known as the ***degeneracy of genetic code***. For instance, UGA codes for tryptophan in yeast mitochondria.
6. **Start and Stop Codons:** Generally, ***AUG codon*** is the initiating or start codon. The polypeptide chain starts either with eukaryotes (methionine) or prokaryotes (N-formylmethionine). On the other hand, ***UAG, UAA*** and ***UGA*** are called as termination codons or stop codons. These are not read by any tRNA molecules and they never code for any amino acids.
7. **Non-ambiguous and Universal:** The genetic code is non-ambiguous which means a specific codon will only code for a particular amino acid. Also, the same genetic code is seen valid for all the organisms i.e. they are universal.

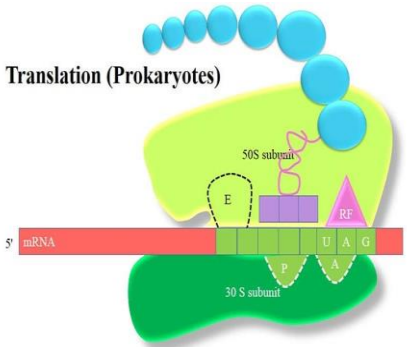
Exceptions to the Code

The genetic code is universal since similar codons are assigned to identical amino acids along with similar START and STOP signals in the majority of genes in microorganisms and plants. However, a few exceptions have been discovered and most of these include assigning one or two of the STOP codons to an amino acid.

Apart from this, both the codons ***GUG*** and ***AUG*** may code for methionine as a starting codon, although GUG is meant for valine. This breaks the property of non-ambiguousness. Thus, it can be said that few codes often differs from the universal code or non-ambiguous code.

**Translation in prokaryotes**

- Translation involves translating the sequence of a messenger RNA (mRNA) molecule to a sequence of amino acids during protein synthesis.
- It is the process in which ribosomes in the cytoplasm or ER synthesize proteins after the process of transcription of DNA to RNA.




Ribosomes:

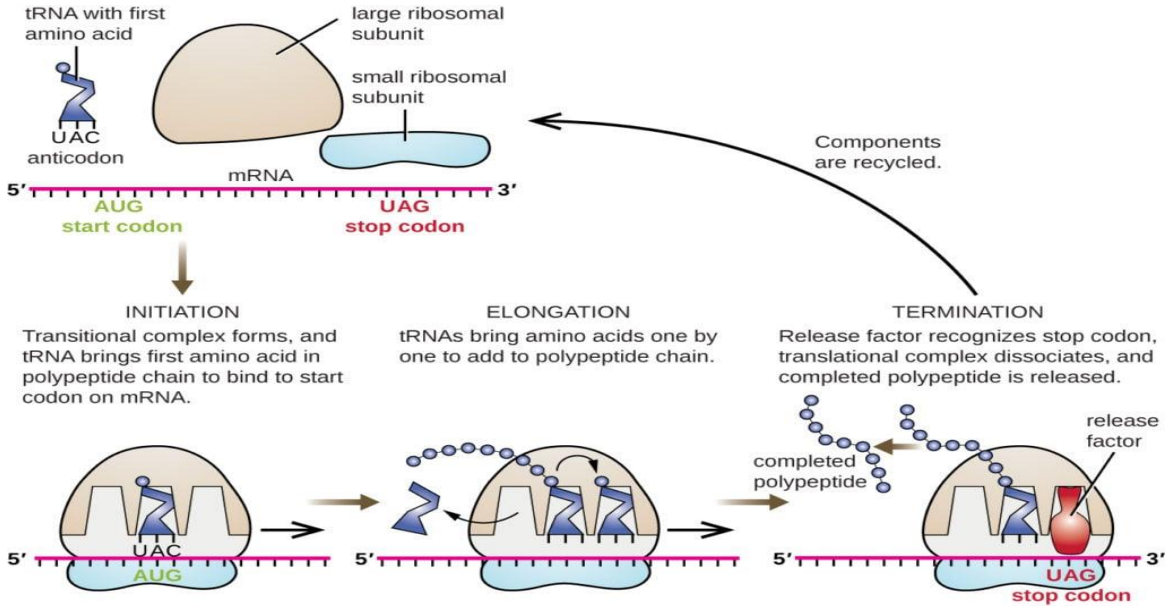
- Ribosomes exist normally as separate subunits that are composed of protein and rRNA.
- The subunits come together to form a ribosome when they bind to an mRNA, near its 5' end.
- On binding to an mRNA, the ribosome reads the nucleotide sequence from the 5' to 3' direction, synthesizing the corresponding protein from amino acids in an N-terminal (amino-terminal) to C-terminal (carboxyl terminal) direction.
- Ribosomes are located in the cytosol, either freely floating or associated with the endoplasmic reticulum.
- They serve to synthesize proteins.

Ribosomal site for protein translation

Each prokaryotic ribosome, shown schematically, has three binding sites for tRNAs.

1. **The aminoacyl-tRNA binding site** (or A site) is where, during elongation, the incoming aminoacyl-tRNA binds.
2. **The peptidyl-tRNA binding site** (or P site) is where the tRNA linked to the growing polypeptide chain is bound.
3. **The exit site** (or E site) is a binding site for tRNA following its role in translation and prior to its release from the ribosome.

All three sites (A, P and E) are formed by the rRNA molecules in the ribosome.




Protein synthesis (or translation) takes place in three stages:

1. Initiation
2. Elongation and
3. Termination.

- During initiation, the mRNA–ribosome complex is formed and the first codon (always AUG) binds the first aminoacyltRNA (called initiator tRNA).
- During the elongation phase, the other codons are read sequentially and the polypeptide grows by addition of amino acids to its C-terminal end.
- This process continues until a termination codon (Stop codon), which does not have a corresponding aminoacyl-tRNA with which to base pair, is reached.
- At this point, protein synthesis ceases (termination phase) and the finished polypeptide is released from the ribosome.

**Synthesis of aminoacyl-tRNAs**

Synthesis of aminoacyl-tRNAs is crucially important for two reasons:

1. Each amino acid must be covalently linked to a tRNA molecule in order to take part in protein synthesis, which depends upon the ‘adaptor’ function of tRNA to ensure that the correct amino acids are incorporated.

2. The covalent bond that is formed between the amino acid and the tRNA is a high energy bond that enables the amino acid to react with the end of the growing polypeptide chain to form a new peptide bond.

For this reason, the synthesis of aminoacyl-tRNA is also referred to as **amino acid activation.**

- Each tRNA molecule has a cloverleaf secondary structure with the anticodon accessible at the end of the anticodon stem loop.
- During synthesis of the aminoacyl-tRNA, the amino acid is covalently bound to the A residue of the CCA sequence at the 3' end.
- Each tRNA molecule carries only a single amino acid.
- The attachment of an amino acid to a tRNA is catalyzed by an enzyme called **aminoacyl-tRNA synthetase.**
- A separate aminoacyl-tRNA synthetase exists for every amino acid, making 20 synthetases in total.

The synthesis reaction occurs in two steps.

1. The first step is the reaction of an amino acid and ATP to form an aminoacyl-adenylate (also known as aminoacyl-AMP).
2. In the second step, without leaving the enzyme, the aminoacyl group of aminoacyl-AMP is transferred to the 3' end of the tRNA molecule to form aminoacyl-tRNA

The overall reaction is:

**Amino acid + ATP + tRNA → aminoacyl-tRNA + AMP + $PP_i$**

**Initiation of protein synthesis:**

- The first codon translated in all mRNAs is the start codon or initiation codon, AUG which codes for methionine.
- Two different tRNAs are used for the two types of AUG codon; $tRNA_f^{Met}$ is used for the initiation codon and is called the initiator tRNA whereas $tRNA_m^{Met}$ is used for internal AUG codons.
- In prokaryotes the first amino acid of a new protein is N-formylmethionine (abbreviated fMet). Hence the aminoacyl-tRNA used in initiation is fMet-$tRNA_f^{Met}$.
- A short sequence rich in purines (5'-AGGAGGU-3'), called the **Shine–Dalgarno sequence**, lies 5' to the AUG initiation codon and is complementary to part of the 16S rRNA in the small ribosomal subunit.
- Therefore this is the binding site for the 30S ribosomal subunit which then migrates in a 3' direction along the mRNA until it encounters the AUG initiation codon.
- Initiation of protein synthesis requires proteins called initiation factors (IFs).
- In prokaryotes, three initiation factors (IF-1, IF-2 and IF-3) are essential.
- Because of the complexity of the process, the exact order of binding of IF-1, IF-2, IF-3, fMet-tRNAf is controversial.

**Steps Involved**

1. Initiation begins with the binding of IF-1 and IF-3 to the small (30S) ribosomal subunit. Their role is to stop the 30S subunit binding to the 50S subunit in the absence of mRNA and fMet-$tRNA_f^{Met}$ which would result in a nonfunctional ribosome.
2. The small subunit then binds to the mRNA via the Shine–Dalgarno sequence and moves 3' along the mRNA until it locates the AUG initiation codon.

3. The initiator tRNA charged with N-formylmethionine and in a complex with IF-2 and GTP (fMet-tRNAfMet/IF-2/GTP) now binds.
4. IF-3 is released.
5. The complex of mRNA, fMet-tRNAf Met, IF-1, IF-2 and the 30S ribosomal subunit is called the 30S initiation complex.
6. The large (50S) ribosomal subunit now binds, with the release of IF-1 and IF-2 and hydrolysis of GTP, to form a 70S initiation complex.

**Elongation of Protein Synthesis**

- At the start of the first round of elongation, the initiation codon (AUG) is positioned in the P site with fMet-tRNA$_f^{Met}$ bound to it via codon–anticodon base pairing.
- The next codon in the mRNA is positioned in the A site.
- Elongation of the polypeptide chain occurs in three steps called the elongation cycle, namely aminoacyl-tRNA binding, peptide bond formation and translocation:

**Aminoacyl-tRNA binding**

- The corresponding aminoacyl-tRNA for the second codon binds to the A site via codon–anticodon interaction.
- Binding of the aminoacyl-tRNA requires elongation factor EF-Tu and GTP which bind as an aminoacyl-tRNA/EF-Tu/GTP complex.
- Following binding, the GTP is hydrolyzed and the EF-Tu is released, now bound to GDP.
- Before the EF-Tu molecule can catalyze the binding of another charged tRNA to the ribosome, it must be regenerated by a process involving another elongation factor, EF-Ts.

This regeneration is called the EF-Tu–EF-Ts exchange cycle.

- First, EF-Ts binds to EF-Tu and displaces the GDP. Then GTP binds to the EF-Tu and displaces EF-Ts. The EF-Tu-GTP is now ready to take part in another round of elongation.

**Peptide bond formation**

- The second step, peptide bond formation, is catalyzed by peptidyl transferase.
- In this reaction the carboxyl end of the amino acid bound to the tRNA in the P site is uncoupled from the tRNA and becomes joined by a peptide bond to the amino group of the amino acid linked to the tRNA in the A site.

**Translocation**

- In the third step, a complex of elongation factor EF-G (also called translocase) and GTP (i.e. EF-G/GTP) binds to the ribosome.
- Three concerted movements now occur, collectively called translocation:

1. the deacylated tRNA moves from the P site to the E site
2. the dipeptidyl-tRNA in the A site moves to the P site, and
3. The ribosome moves along the mRNA (5’ to 3’) by three nucleotides to place the next codon in the A site.

- During the translocation events, GTP is hydrolyzed to GDP and inorganic phosphate, and EF-G is released ready to bind more GTP for another round of elongation.
- After translocation, the A site is empty and ready to receive the next aminoacyl tRNA.
- The A site and the E site cannot be occupied simultaneously. Thus the deacylated tRNA is released from the E site before the next aminoacyl-tRNA binds to the A site to start a new round of elongation.

- Elongation continues, adding one amino acid to the C-terminal end of the growing polypeptide for each codon that is read, with the peptidyl-tRNA moving back and forth from the P site to the A site as it grows.

**Termination of Protein Synthesis**

- Eventually, one of three termination codons (also called Stop codons) becomes positioned in the A site. These are UAG, UAA and UGA.
- Unlike other codons, prokaryotic cells do not contain aminoacyl-tRNAs complementary to
- Stop codons. Instead, one of two release factors (RF-1 and RF-2) binds instead.
- RF-1 recognizes UAA and UAG whereas RF-2 recognizes UAA and UGA. A third release factor, RF-3, is also needed to assist RF-1 or RF-2 interaction with the ribosome. Thus RF-1 + RF-3 or RF-2 + RF-3 bind depending on the exact termination codon in the A site.
- RF-1 (or RF-2) binds at or near the A site whereas RF-3/GTP binds elsewhere on the ribosome.
- The release factors cause the peptidyl transferase activity to transfer the polypeptide to a water molecule instead of to aminoacyl-tRNA, effectively cleaving the bond between the polypeptide and tRNA in the P site.

The free polypeptide now leaves the ribosome, followed by the mRNA and free tRNA, and the ribosome dissociates into 30S and 50S subunits ready to start translation again.

**********